CELLULAR AND MOLECULAR INVESTIGATIONS OF UNDIAGNOSED NEUROMETABOLIC DISORDERS

Submitted in application for award of Doctor of Philosophy (PhD)

Emma Reid

Centre for Translational Omics

Genetics and Genomic Medicine

Institute of Child Health

University College London

August 2016

DECLARATION

I, Emma Reid, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Where experimental or analytical work has been completed by others, this has been stated in the relevant section of this thesis. However, these instances are also listed below:

- Alignment of sequencing reads to the reference genome, variant calling (Section 2.4.11) and ExomeDepth analysis (Section 3.3.3) of the gene panel sequencing data was performed by Dr Chris Boustred (NE Thames Regional Genetics Service, GOSH, UK).
- Whole exome sequencing of patients X, 1, 4 and 5 was outsourced to BGI Genomics Hong Kong and the resulting raw data files were processed and aligned by Dr Chela James (GOSgene, ICH, UK) (Section 2.5).
- Whole exome sequencing of patient Y and processing of the resulting raw data was performed by Dr Olaf Bodamer (University of Miami, USA) (Section 6.4.1).
- Homozygosity mapping and whole exome sequencing of index PROSC family was performed by Dr Niklas Darin (The Queen Silvia Children's Hospital, Sweden) (Section 7.1.4).
- Sanger sequencing of additional PROSC and PNPO-deficient patients was performed by Dr Philippa Mills (ICH, UK) (Section 7.2).
- Sample processing for electron microscopy analysis was carried out by Elizabeth Latimer-Bowman (Histopathology Department, GOSH, UK) and imaging was performed by Glenn Anderson (Histopathology Department, GOSH, UK) (Section 2.14.2).
- Immunofluorescence staining of α-dystroglycan in a muscle sample from patient 23 was performed by the Dubowitz Neuromuscular Centre, ICH, UK (Section 3.3.6.1).
- Amino acid analyses of plasma, urine and CSF listed throughout Chapters 4, 5 and 6 were performed on a clinical basis by the Chemical Pathology Department, GOSH, UK.

London, August 2016

Emma Sarah Reid

ABSTRACT

Inborn errors of metabolism (IEM) affect 1 in 500 newborns causing significant disease-burden and mortality throughout childhood. However, despite extensive genetic and biochemical investigations the cause of disease remains unknown in up to 50% of patients with neurological symptoms; so-called neurometabolic disorders (NMD). The overarching aim of this thesis was to determine the cellular and molecular aetiologies for the clinical phenotypes seen in patients with undiagnosed NMD.

In order to improve the diagnosis of these disorders in clinical practice, a comprehensive targeted gene panel of 614 genes known to cause IEM was designed and a cohort of 44 patients was analysed. A definitive or probable genetic diagnosis was achieved in 53% of patients without a prior genetic diagnosis. Method optimisation and validation, comparison to other diagnostic strategies and the advantages and disadvantages of targeted sequencing are reviewed. Case reports, novel mutations/phenotypes and their contribution to the expansion of the literature are described.

Whole exome sequencing and functional characterisation was also undertaken for patients who had been extensively clinically investigated previously. Five patients identified with mutations in the mitochondrial glutamate transporter, *SLC25A22*, presenting with novel biochemical phenotypes are described and novel transporter functions are postulated. One patient diagnosed with a potassium channelopathy with biochemical abnormalities and anticonvulsant responses suggestive of an inborn error of vitmain B_6 metabolism is documented and the mechanisms underlying the generalised anticonvulsant effects of vitamin B_6 are postulated. Characterisation of a possible novel inborn error of lysine metabolism in a patient presenting with hyperlysinaemia and motor neuron disease is also discussed.

These studies also demonstrate the complexity of unravelling the relationship between genotype and phenotype and highlight the need for novel functional assays to assess the pathogenicity of sequence variants. Mass spectrometry-based assays were developed to enable characterisation of disorders affecting vitamin B_6 homeostasis, including pyridox(am)ine 5'-phosphate oxidase (PNPO), antiquitin and PROSC deficiency, the latter being a novel disorder. The differences between pyridoxine- and pyridoxal phosphate-responsive PNPO deficiency and fibroblast vitamer profiles in all patients were all investigated. Finally, multiple methodologies were employed with the aim of understanding the biological function of PROSC.

PUBLICATIONS

Some ideas, figures and methodologies have appeared previously in the following publications:

1. **Reid ES**, Williams H, Stabej PL, James C, Ocaka L, Bacchelli C, Footitt EJ, Boyd S, Cleary MA, Mills PB, Clayton PT (2015) Seizures Due to a KCNQ2 Mutation: Treatment with Vitamin B_6. *JIMD Reports* Oct 8. [epub ahead of print]

2. Oppici E, Fargue S, **Reid ES**, Mills PB, Clayton PT, Danpure CJ, Cellini B (2015) Pyridoxamine and pyridoxal are more effective than pyridoxine in rescuing folding-defective variants of human alanine:glyoxylate aminotransferase causing primary hyperoxaluria type I. *Human Molecular Genetics* 24(19):5500-11.

3. **Reid ES**, Papandreou A, Drury S, Boustred C, Yue WW, Wedatilake Y, Beesley C, Jacques TS, Anderson G, Abulhoul L, Broomfield A, Cleary M, Grunewald S, Varadkar SM, Lench N, Rahman S, Gissen P, Clayton PT, Mills PB (2016) Advantages and pitfalls of an extended gene panel for investigating complex neurometabolic phenotypes. *Brain* [in press].

ACKNOWLEDGEMENTS

Above all I would like to thank the patients, carers, relatives and other individuals who donated samples and their time, freely and without necessarily expecting benefit to themselves. I am very grateful to the organisations that funded this work: Great Ormond Street Hospital Children's Charity and UCL Impact.

I am indebted to my supervisors: Philippa Mills for her continual support both inside and outside the lab, not only helping me to grow as a scientist, but also in confidence as a person. Without her this work would not have been possible and I am very grateful to have had the opportunity to work in her lab. Paul Gissen I must thank for always playing the devil's advocate and pushing me to aim higher than I thought was possible. I also owe a great deal to Peter Clayton for sharing merely a slice of his brilliance with me whilst never failing to bamboozle me with hand-drawn biochemistry.

There are many other people that I have met during the course of my PhD who have been of enormous help to me. Firstly, I would like to thank Nick Lench for giving me the opportunity to learn in such a unique diagnostic laboratory, along with Suzie Drury and Chris Boustred for helping to make our panel a reality and always having time for my never-ending questions. I also thank the members of GOSgene for their generous collaboration and assistance with interpretation of the whole exome sequencing data. Thanks to Kevin Mills for introducing me to the wonders (and woes!) of mass spectrometry. I am also eternally grateful to the GOSH Metabolic Team for their help recruiting patients, enthusiasm in the project and for rekindling my passion for medicine. Many, many thanks to my colleagues and friends at the Institute of Child Health and particularly within the Biological Mass Spectrometry Centre, for their readiness to help and making each day in the lab a fun day.

Last but not least, I thank my family for their continual support and encouragement, and Malika for keeping me chained to that desk, no matter how much I protested!

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

α-AASA	α-aminoadipic semialdehyde
5-HIAA	5-hydroxyindoleacetic acid
AADC	Aromatic L-amino acid decarboxylase
aCGH	Array comparative genomic hybridisation
ACMG	American College of Medical Genetics and Genomics
ACN	Acetonitrile
ADP	Adenosine diphosphate
AED	Anti-epileptic drug
ALP	Tissue non-specific alkaline phosphatase
ALS	Amyotrophic lateral sclerosis
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
ATP	Adenosine triphosphate
BCA	Bicinchoninic acid
BFIS	Benign familial infantile seizures
BFNS	Benign familial neonatal seizures
BWA	Burrows-Wheeler Alignment
CCBL1	Cysteine conjugate-β lysase
CDG	Congenital disorder of glycosylation
cDNA	Complementary DNA
CK	Creatine kinase
CNV	Copy number variant
CoA	Coenzyme-A
CSF	Cerebrospinal fluid
CT	Computerised tomography
DAAO	D-amino acid oxidase
ddNTPs	Di-deoxynucleotides
DEPC	Diethylpyrocarbonate
dH_2O	Distilled water
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid

dNTPs	Deoxynucleotides
ECD	Enrichment Control DNA
EDTA	Ethylenediaminetetraacetic acid
EEG	Electroencephalography
ER	Endoplasmic reticulum
ERG	Electroretinography
ExAC	Exome Aggregation Consortium
FAD	Flavin adenine dinucleotide
$FADH_2$	Reduced flavin adenine dinucleotide
FBS	Fetal bovine serum
FMN	Flavin mononucleotide
FMOC-HCl	Fluorenylmethyloxycarbonyl chloride
GABA	γ-aminobutyric acid
GAD	Glutamate decarboxylase
GATK	Genome Analysis Toolkit
GC-MS	Gas chromatography coupled to mass spectrometry
GOSH	Great Ormond Street Hospital for Children NHS Foundation Trust
GPI	Glycosylphosphatidylinositol
GSA	Glutamate-γ-semialdehyde
GSH	Reduced glutathione
GSSG	Oxidised glutathione disulphide
HFBA	Heptafluorobutyric acid
HPI	Hyperprolinaemia type I
HPII	Hyperprolinaemia type II
HPLC	High-performance liquid chromatography
HPO	Human Phenotype Ontology
HVA	Homovanillic acid
IEM	Inborn error of metabolism
IVT	*In vitro* translation
KA	Kynurenic acid
KAC	α-keto-ϵ-aminocaproate
KCNQ2	Potassium voltage-gated channel subfamily Q member 2
KN	Kynurenine
LAMP2	Lysosome-associated membrane protein 2
LC-MS/MS	Liquid chromatography coupled to tandem mass spectrometry

m/z	Mass-to-charge ratio
MA	Malonic acid
$MgCl_2$	Magnesium chloride
miRNA	MicroRNA
MLPA	Multiplex ligation-dependent probe amplification
MMA	Methylmalonic acid
MRI	Magnetic resonance imaging
MRM	Multiple reaction monitoring
mRNA	Messenger RNA
MS/MS	Tandem mass spectrometry
mTORC1	Mammalian target of rapamycin complex 1
NAD^+	Nicotinamide adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide
$NADP^+$	Nicotinamide adenine dinucleotide phosphate
NADPH	Reduced nicotinamide adenine dinucleotide phosphate
NEE	Neonatal epileptic encephalopathy
NGS	Next-generation sequencing
NMD	Neurometabolic disorders
NMDA	N-methyl-D-aspartate receptor
o-AB	o-aminobenzaldehyde
OAT	Ornithine δ-aminotransferase
OMIM	Online Mendelian Inheritance in Man
P2C	L-Δ^1-piperideine-2-carboxylate
P5C	L-Δ^1-pyrroline-5-carboxylic acid
p62	Nucleoporin 62
P6C	L-Δ^1-piperideine-6-carboxylic acid
PA	4-pyridoxic acid
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDE	Pyridoxine dependent epilepsy
PK	Pyridoxal kinase
PL	Pyridoxal
PLP	Pyridoxal 5'-phosphate
PM	Pyridoxamine
PMP	Pyridoxamine 5'-phosphate

PN	Pyridoxine
PNH	Peripheral nerve hyperexcitability
PNP	Pyridoxine 5'-phosphate
PNPO	Pyridox(am)ine 5'-phosphate oxidase
PolyPhen-2	Polymorphism Phenotyping version 2
PPADS	Pyridoxal phosphate-6-azophenyl-2-4-disulfonic acid
PPNDS	Pyridoxal-5'-phosphate-6-(2'-naphthylazo-6'-nitro-4',8'-disulfonate)
PROSC	Proline synthetase co-transcribed [bacterial homolog]
qRT-PCR	Quantitative real-time polymerase chain reaction
RNA	Ribonucleic acid
ROS	Reactive oxygen species
SIFT	Sorting Intolerant from Tolerant
SLC25A22	Solute carrier family 25 member 22
SSIEM	Society for the Study of Inborn Errors of Metabolism
T_m	Melting temperature
TBE	Tris-Borate-EDTA
TBST	Tris-buffered saline and Tween 20
TCA	Tricarboxylic acid
TCA	Trichloroacetic acid
TFA	Trifluoroacetic acid
UDP	Uridine diphosphate
UDP-GalNAc	UDP-N-acetylglucosamine
UDP-GlcNAc	UDP-N-acetylgalactosamine
UPLC	Ultra-performance liquid chromatography
UPLC-MS/MS	Ultra-performance liquid chromatography coupled to tandem mass spectrometry
UTR	Untranslated region
VCF	Variant call file
VEP	Visual evoked potential
WES	Whole exome sequencing
WGS	Whole genome sequencing
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

1

INTRODUCTION

1.1 INBORN ERRORS OF METABOLISM (IEM)

1.1.1 *First description of IEM*

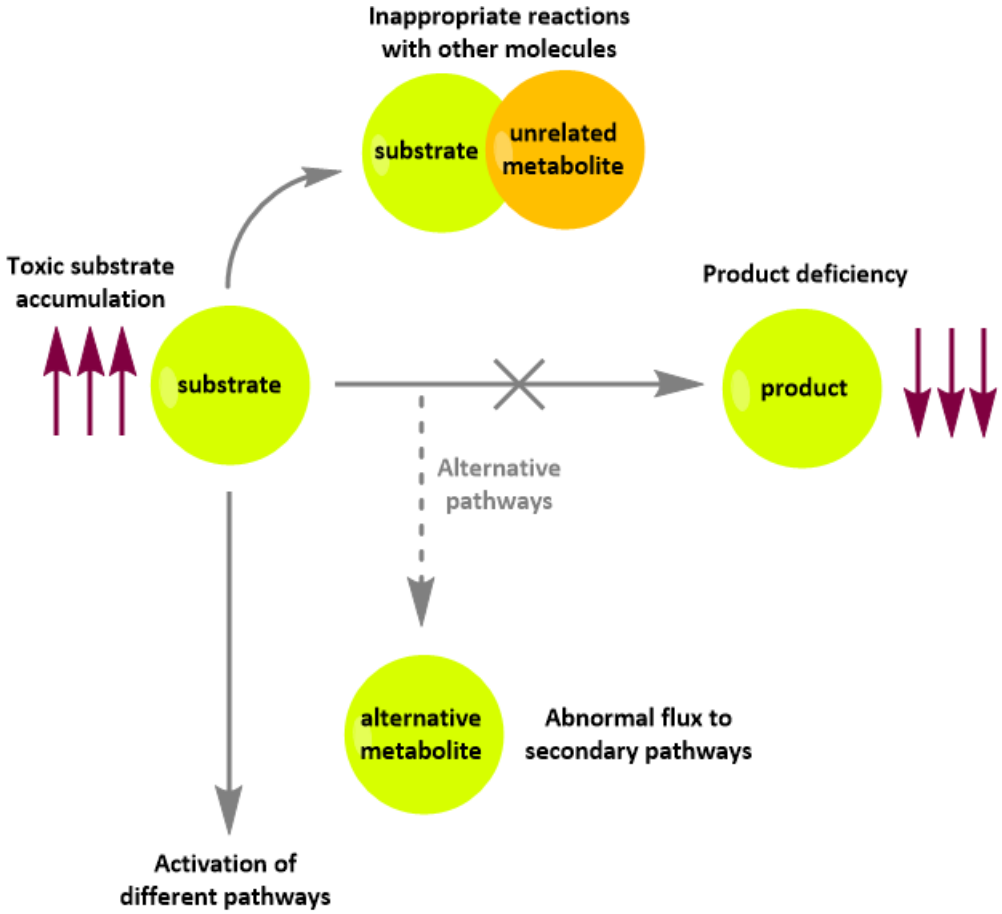
The concept of inborn errors of metabolism (IEM) was introduced at the beginning of the twentieth century by Sir Archibald Garrod (Garrod, 1908). Despite the existence of only very rudimentary analytical techniques and before the discovery of DNA, he proposed the existence of innumerable minor differences in biochemistry that make an individual unique. He also highlighted the fact that the anabolism and catabolism of all molecules including proteins, carbohydrates, fats and sugars, are catalysed by enzymes specific to each reaction, despite very limited evidence to support this hypothesis. Using the examples of albinism, alkaptonuria, cystinuria and pentosuria, Garrod assembled evidence to suggest that these disorders could be explained by a block in normal metabolic pathways due to a congenital deficiency of a specific enzyme. Indeed, he viewed metabolism as a continuous step-wise flux through intermediary molecules, each of which only has a transient existence. Unfortunately, Garrod's work was mostly ignored (particularly amongst his medical colleagues) for fifty years until La Du *et al.* (1958) demonstrated the absence of an enzyme, homogentisate 1,2-dioxygenase, in the liver of a patient with alkaptonuria. Integral to this progression within the field were technological advances including the advent of electrophoresis and chromatography, as well as the rediscovery and growing acceptance of Mendel's laws of inheritance. Indeed, the era of molecular medicine truly begun in 1949 when Pauling *et al.* demonstrated that genetic mutations could alter the structure of wild-type proteins. In his study of patients with sickle cell anaemia, he showed that the "sickling" of the red blood cells in affected patients was due to an altered conformation and charge of the haemoglobin molecules within. The next leap forward in the understanding of the pathogenesis of monogenic disorders came when it was shown that the altered properties of haemoglobin in patients with sickle cell anaemia is due to a single amino acid alteration within the polypeptide sequence (Ingram, 1956, 1957). However, this concept was soon extended beyond a select few disorders and in 1969 the assertion that genetic factors are involved in all diseases was made (McKusick, 1969), pre-dating the sequencing of the human genome by more than thirty years. Indeed, the vast genetic variation uncovered through the DNA sequencing revolution has

simply confirmed the remarkable insights made by Garrod almost 100 years previously; people are "chemically individual" (Garrod, 1908).

1.1.2 *Pathogenic mechanisms underlying IEM*

Since the recognition of alkaptonuria as the first IEM, more than 600 disorders have been described (Hamosh *et al.*, 2005). Indeed, what was once thought to be a few benign ("sport") medical conditions (Garrod, 1908), has been expanded to include a range of disorders that cumulatively affect up to 1 in 500 newborns and account for a significant proportion of morbidity and mortality both in childhood and the neonatal period (Chiaratti de Oliveira *et al.*, 2001). Collectively they are a diverse class of genetic disorders which are markedly heterogeneous, both clinically and genetically with disease involving almost any organ or system. The majority of IEM are caused by single-gene defects in pathways involved in the synthesis, metabolism, transport and/or storage of metabolites or larger molecules and result in insufficient, absent or abnormal conversion of substrates into products. These disorders typically follow an autosomal recessive mode of inheritance and can arise as a result of missense and nonsense mutations, insertions, deletions or genomic rearrangements. These mutations can occur in coding regions, intron-exon boundaries so as to affect correct gene splicing, regulatory regions, or intronic regions through the introduction of cryptic splice sites or the disruption of splicing enhancer sites. Due to the highly inerconnected nature of metabolism, mutations in different genes can affect the same pathway causing similar or even identical phenotypes. In most cases, the adverse biological effects of these mutations are propagated by five main mechanisms: toxic accumulation of substrates upstream of the pathway block, reduction of essential compounds downstream, activation of alternative pathways, diversion of metabolic flux to abnormal secondary pathways or secondary deficiencies of essential metabolites due to uncontrolled or inappropriate chemical reactions (Figure 1.1.1).

Figure 1.1.1: Pathogenesis of inborn errors of metabolism. Adapted from Lanpher *et al.* (2006).



1.1.3 *Classes of IEM*

In 1963, the Society for the Study of Inborn Errors of Metabolism (SSIEM) was founded to foster the exchange of ideas between professionals in different disciplines with an interest in inherited metabolic disease. The SSIEM later formulated a hierarchical classification which formed the basis of the designation of inborn errors of metabolism in the 11^{th} international classification of diseases (ICD11). Metabolic defects that affect the same or interconnected biochemical pathways, or indeed the appropriate metabolism of related molecules, often present with similar clinical features in patients. They are also typically diagnosed and monitored using the same routine or specialised assays and subject to comparable acute and long-term treatments. Therefore, this classification groups the $>$ 600 disorders into 15 broad classes according to their dysfunctional biochemical pathway, disease group and pathophysiological mechanisms. These classes are listed in Table 1.1.1, alongside examples of disorders within each class, the normal functions of these enzymes and the clinical and biochemical findings that may suggest one of these disorders.

Importantly, this summary also illustrates the vast intra- and inter-class clinical heterogeneity that complicates the diagnosis of these disorders.

Table 1.1.1: Classes of inborn errors of metabolism, examples of disorders and associated clinical and biochemical features.

Disorder class	Examples of disorders within this class	Normal function of enzymes within this class	Phenotypic and biochemical indicators that would suggest a disorder within this class
Amino acid and peptide metabolism	Urea cycle disorders, organic acidurias, phenylketonuria, maple syrup urine disease, alkaptonuria, hyperprolinaemia, hyperlysinaemia.	These enzymes catalyse the metabolism of amino acids and related molecules in the body. Their dysfunction results in abnormal levels of these and metabolically connected molecules. The urea cycle also functions to remove ammonia from the bloodstream and produce urea for excretion. Defects affecting this process or the provision of its substrates results in severe hyperammonaemia.	Hyperammonaemia (including acute crisis), avoidance of high-protein foods, vomiting, lethargy, neurodevelopmental abnormalities, seizures, abnormal organic acids in biofluids, abnormal amino acid concentrations in biofluids.
Carbohydrate metabolism	Glycogen storage disorders, galactosaemia, hereditary fructose intolerance, glucose transporter deficiency.	Carbohydrates are sugars ranging in size from monosaccharides (e.g. galactose and glucose) to polysaccharides (e.g. glycogen which is used for energy storage). The main function of carbohydrates is to be catabolised to produce cellular energy. Defects in the interconversion of carbohydrates and related molecules can cause abnormal storage of polysaccharides impairing normal function or accumulation of toxic sugar-derivatives.	Hypoglycaemia, hepatomegaly, hyperlipidaemia, growth retardation, exercise intolerance, renal failure, vomiting, neurodevelopmental abnormalities, seizures, abnormal storage material visible in tissue samples under electron microscopy.
Fatty acid and ketone body metabolism	Disorders of mitochondrial fatty acid oxidation, medium-chain acyl-CoA dehydrogenase (MCAD) deficiency, carnitine deficiency.	The oxidation of fatty acids by β-oxidation and the citric acid cycle is a major source of cellular ATP. Fatty acids are the major storage form of energy in humans which are normally broken down and utilised when glucose is unavailable. However, dysfunction of enzymes within these pathways means that this fat source cannot be metabolised to form energy, halting bodily processes.	Non-ketotic hypoglycaemia, cardiomyopathy, rhabdomyolysis, liver dysfunction, poor appetite, abnormal acylcarnitine profile.
Energy metabolism	Mitochondrial respiratory chain disorders (mutations in mitochondrial and nuclear DNA), mitochondrial membrane transport disorders, disorders of creatine metabolism.	Mitochondria are cellular organelles that generate ATP through oxidative phosphorylation carried out by the respiratory chain. Disease can result in defective energy production (isolated or multiple complex dysfunction), defects in the maintenance of mitochondrial DNA, mitochondrial protein synthesis or importation.	Multi-systemic disorder, proximal myopathy, cardiomyopathy, encephalopathy, seizures, neurodevelopmental abnormalities, characteristic MRI abnormalities (basal ganglia changes), elevated lactate in biofluids, reduced mitochondrial respiratory chain complex activity in muscle biopsy, abnormal copy number of mitochondrial DNA.

Metabolism of purines, pyrimidines and nucleotides	Aicardi-Goutieres syndrome, orotic aciduria.	Purines and pyrimidines are the building blocks of DNA and RNA, as well as playing a role in the regulation of cell metabolism (NADH), energy transport (ATP), cell signalling (cyclic AMP), formation of coenzymes (coenzyme A) and intermediates of phospholipid and carbohydrate metabolism.	Immunodeficiency, failure to thrive, seizures, neurodevelopmental abnormalities, renal abnormalities, anaemia, acute drug toxicity, abnormal organic acid concentrations in biofluids, abnormal uric acid concentrations in biofluids.
Metabolism of sterols	Disorders of bile acid biosynthesis, metabolism and transport, Smith-Lemli-Opitz syndrome.	Sterols are steroid alcohols. The most common is cholesterol which forms part of cellular membranes and is a precursor for bile acids (digestion, liver function, absorption of fat soluble vitamins), steroids (hormones) and hedgehog proteins (developmental signalling).	Cholestasis, fat-soluble vitamin malabsorption, liver disease, neurodevelopmental abnormalities, dysmorphic features/structural abnormalities, skin abnormalities, abnormal concentrations/isoforms of bile acids in biofluids, abnormal sterol analysis in biofluids.
Porphyrin and haem metabolism	Erythropoietic porphyria, X-linked sideroblastic anaemia.	Porphyrins are heterocyclic macrocycle organic compounds. Haem is most well known as the pigment in erythrocytes and cofactor of haemoglobin; it is synthesised from porphyrin precursors. It is the accumulation of these precursors caused by a deficiency in one of the enzymes in the metabolic pathway, that are toxic to the tissues.	Acute abdominal pain, vomiting, hypertension, tachycardia, motor neuropathy, photosensitivity, skin disease.
Lipid and lipoprotein metabolism	Familial hypercholesterolaemia, Sjogren-Larsson syndrome, Tangier disease.	Lipids function to store energy, act as signalling molecules and as structural components of cell membranes. Lipoproteins serve to emulsify lipid molecules, allowing them to move through aqueous environments (e.g. high- and low-density lipoproteins enable fats to be carried in the bloodstream).	Neurodevelopmental abnormalities, skin discolouration/abnormalities, cardiovascular disease, neuropathy, ocular abnormalities, abnormal cholesterol levels in biofluids, abnormal triglyceride levels in biofluids, alterations in the levels/ratios of lipoproteins in biofluids.

Congenital disorders of glycosylation and other disorders of protein modification	Disorders of N- and O-linked glycosylation, disorders of glycosphingolipid and glycosylphosphatidylinositol anchor glycosylation.	Glycosylation is the enzymatic attachment of a carbohydrate to another molecule, including proteins and lipids. This modification (occurring in 50% of proteins) enables correct protein folding and stability, cell-cell adhesion, antigen recognition, cell surface receptors and protein-protein interactions. Due to the variety of tissue proteins and lipids that are affected by defective glycosylation, dysfunction of these processes result in multi-systemic disorder.	Multi-systemic disorder, neurodevelopmental abnormalities, dysmorphic features/structural abnormalities, seizures, ocular abnormalities, failure to thrive, coagulopathy, hypoglycaemia, liver abnormalities, hyperphosphatasia, abnormal transferrin isoelectric focussing pattern, abnormal glycan analysis.
Lysosomal disorders	Mucopolysaccharidoses, oligosaccharidoses, sphingolipidoses, neuronal ceroid lipofuscinoses.	Lysosomes are cellular membrane-bound organelles that contain hydrolytic enzymes that degrade biomolecules (e.g. proteins, lipids, carbohydrates and nucleic acids). The dysfunction of these hydrolytic enzymes or proteins can impact on the transport of lysosome/lysosomal substrates, resulting in the toxic accumulation of waste products within cells.	Neurodevelopmental abnormalities, movement disorders, seizures, deafness, blindness, hepatosplenomegaly, cardiac abnormalities, abnormal lysosomal enzyme screen, abnormal storage material visible in tissue samples under electron microscopy.
Peroxisomal disorders	Zellweger spectrum disorder, defects of peroxisomal alpha-, beta- and omega-oxidation, peroxisomal biogenesis disorders.	Peroxisomes are cellular organelles, the main functions of which are catabolism of very long chain fatty acids (VLCFAs) and the biosynthesis of plasmologens. Dysfunction of the latter process causes profound abnormalities in the myelination of nerve cells. Impaired catabolism of VLCFAs results in a deficiency of substrates for the mitochondrial respiratory chain.	Neurodevelopmental abnormalities, dysmorphism, seizures, liver dysfunction, bone abnormalities, deafness, ocular abnormalities, abnormal very long chain fatty acid concentrations, abnormal plasmalogen concentrations in biofluids, abnormal bile acid concentrations in biofluids.
Neurotransmitter metabolism	Aromatic L-amino acid decarboxylase (AADC) deficiency, tyrosine hydroxylase deficiency, GABA transaminase deficiency.	Neurotransmitters transmit signals across a chemical synapse between neurons. They can be excitatory, inhibitory or a mixture of the two. Many are synthesised from amino acids through a small number of enzymatic conversions. Dysfunction of these enzymes causes imbalances of neurotransmitter concentrations and altered neuronal function/excitability.	Neurodevelopmental abnormalities, movement disorders, autonomic nervous system dysfunction, abnormal muscle tone and movements, seizures, oculogyric crises, abnormal neurotransmitter concentrations in biofluids.

Metabolism of vitamins and non-protein cofactors	Pyridox(am)ine 5'-phosphate oxidase (PNPO) deficiency, pyridoxine-dependent epilepsy, vitamin B_{12}-responsive methylmalonic aciduria, hereditary folate malabsorption.	Vitamins are essential cofactors for more than 200 enzymes in the human body. Consequently, a deficiency of the active cofactor caused by dysfunction of an enzyme critical for its metabolism, can have widespread effects. E.g. the active form of vitamin B_6 is a cofactor for $>$ 140 enzymes, many of which are involved in neurotransmitter metabolism.	Neurodevelopmental abnormalities, movement disorders, seizures, lethargy, vomiting, renal abnormalities, hepatomegaly, anaemia, abnormal neurotransmitter concentrations in biofluids, abnormal vitamin concentrations in biofluids.
Metabolism of trace elements and metals	Wilson disease, hypermanganesaemia, hereditary haemochromatosis.	Similarly to vitamins, trace elements are essential cofactors for enzyme activity and to maintain tertiary structures of proteins. Some have only a few functions (e.g. cobalt in the structure of vitamin B_{12} and iodine in the synthesis of thyroid hormone) but others are cofactors for many enzymatic reactions and thus their deficiency has widespread effects (e.g. iron, zinc, copper, manganese and selenium).	Neurodevelopmental abnormalities, liver dysfunction, movement disorders, abnormal blood cells, abnormal concentrations of trace elements in biofluids.
Metabolism of xenobiotics	Trimethylaminuria, dimethylglycinuria.	Xenobiotics are molecules not naturally produced by the body, not expected to be present within it or present in much higher concentrations than normal. The two metabolic disorders within this class result from dysfunction of enzymes which usually break down and then excrete molecules that are ingested in food. Instead, these accumulated chemicals are released in body fluids.	Unusual body odour.

1.1.4 *The importance of early diagnosis*

Timely diagnosis of IEM is crucial, especially for those disorders that are treatable or manageable. However, diagnosis can be complicated not only by the inordinate clinical variability and lack of genotype-phenotype correlations, but also by the fact that environmental factors can be central to determining each patient's phenotype (Dipple and McCabe, 2000; Scriver and Waters, 1999). These factors, alongside genetic variation/polymorphisms in other genes, explain the variable presentations within families sharing identical pathogenic mutations. Indeed, these interactions are thought to underlie the fact that a patient's genotype cannot typically predict the severity of disease, prognosis or treatment response. Once a diagnosis has been established, there are eight generic forms of treatment that can typically be undertaken:

1. **Restriction of the accumulating upstream metabolites**. If pathogenesis is being propagated by the accumulation of toxic metabolites upstream of the dysfunctional enzyme, dietary restriction of these compounds can be an effective treatment. One example is maple syrup urine disease, a disorder caused by a deficiency of the branched-chain α-oxoacid dehydrogenase complex which leads to a toxic accumulation of the branched chain amino acids and their keto-acids. Restriction of these branched chain amino acids (leucine, valine and isoleucine) from the diet followed by strict life-long adherence to a specially formulated diet can prevent the neurological damage associated with the disorder (Morton *et al.*, 2002). Indeed, similar preservation of neurological function and intellect can be observed in patients with phenylketonuria when phenylalanine is restricted from the diet; this finding lay the foundation for newborn screening for IEM (Guthrie and Susi, 1963).

2. **Removal of toxic substrates**. For some disorders, unlike those described above, more rapid removal of toxic compounds than can be achieved by dietary alteration is required to prevent severe neurological damage. One example can be seen in patients with urea cycle disorders. In normal individuals, the metabolism of protein results in the production of nitrogen which is then removed from the blood and converted to urea for urinary excretion. In patients affected by a deficiency of one of the six enzymes that together are responsible for the removal of ammonia from the blood, severe hyperammonaemia can occur which if untreated can result in respiratory distress, coma and death. Life-long dietary protein restriction is necessary to prevent excessive ammonia formation; however acute hyperammonaemic crises can still occur, especially within the neonatal period, after high-protein meals or periods of viral illness (Leonard and Morris, 2002). In severe cases including patients presenting with acute encephalopathy, haemodialysis or haemofiltration to remove

the ammonia may be required (Häberle *et al.*, 2012). Outside of acutely life-threatening scenarios, other compounds such as sodium benzoate can be used as adjunctive treatments to promote the excretion of nitrogenated molecules, thereby rapidly reducing ammonia accumulation (Leonard and Morris, 2002). Similar molecular scavenging treatments are used in a variety of other IEM including isovaleric aciduria and other organic acidaemias (Naglak *et al.*, 1988).

3. **Supplementation of the deficient downstream metabolites**. Where the pathogenic effects of a disorder are caused by low levels of essential metabolites downstream of the dysfunctional enzyme, supplementation of these products can ameliorate symptoms. Pyridox(am)ine 5'-phosphate oxidase deficiency is a primary defect in the synthesis of the form of vitamin B_6 that is a critical cofactor for more than 140 different enzymes in humans, many of which are involved in neurotransmitter metabolism. Hence, the addition of a supplement containing this vitamin to the diet of these patients causes the cessation of their seizure disorder and improves their neurological outcome (Mills *et al.*, 2014).

4. **Supplementation with competing metabolites**. Dietary supplementation is not limited to the replacement of deficient metabolites within the affected pathway; in some cases, treatment with metabolites that competitively inhibit enzyme activity or compete with the accumulating substrates for transport can be effective. One example is Wilson's disease, which is caused by defective excretion of carrier protein-bound copper into the bloodstream and secretion of excess concentrations of this metal into the bile. Therefore, copper accumulates in the liver causing chronic hepatitis, fibrosis and cirrhosis. Free copper also accumulates in the kidneys, eyes and brain causing neuropsychiatric symptoms (Bie *et al.*, 2007). Instead of being treated with chelation therapy, the majority of patients are now given zinc supplementation as it competes with copper for the same transporter at the gut mucosa, resulting in its excretion (Hoogenraad, 2006).

5. **Stimulation of residual enzyme activity**. As mentioned in Table 1.1.1, many enzymes depend on vitamins or trace metals for catalytic activity. In patients with mutations in these enzymes that do not completely abolish function, supplementation with these essential cofactors can increase the residual flux through the pathway and ameliorate symptoms. An example is methylmalonyl-coA mutase deficiency; the function of this enzyme is to catalyse the isomerisation of methylmalonyl-coA to succinyl-coA and requires vitamin B_{12} as a coenzyme. Affected patients can present with feeding problems, failure to thrive, hypotonia, developmental delay and acute metabolic decompensation. In a small proportion of patients, supplementation with vitamin B_{12} results in a marked decrease in the concentration of

methylmalonic acid in biofluids accompanied by a reduction in the severity of symptoms (de Baulny *et al.*, 2005). Supplementation with vitamin B_6 can be similarly effective in patients with cystathionine β-synthase deficiency (Clayton, 2006).

6. **Direct replacement of the dysfunctional enzyme**. In the case of many defects underlying IEM, the above treatment strategies are not possible because defects do not simply affect the interconversion of two metabolites. One such class of diseases are the lysosomal storage disorders, the majority of which are caused by dysfunctional enzymes which break down macromolecules within lysosomes. In recent years, enzyme-replacement therapy (typically in the form of intravenous or intrathecal administration of a recombinant form of the wild-type form of the deficient/defective enzyme) has been developed for the treatment of some IEM. Lysosomal enzymes have been particularly targeted due to the fact that some can be secreted and taken up by other cells, meaning that systemic administration can be efficacious. Clinical benefits of enzyme replacement therapy have been demonstrated in patients with early-stage Fabry disease, with favourable effects on heart and kidney function, decreased pain and increased quality of life (Lidove *et al.*, 2010). Similar prevention of progressive blood cell abnormalities, organomegaly and bone pathologies have been reported in patients with type I Gaucher disease (Weinreb *et al.*, 2002). However, one of the most significant limitations of these treatments is the fact that the recombinant enzymes cannot cross the blood-brain barrier or penetrate the bones. Therefore, the progression of bone and neurological disease is not affected in patients. However, work is ongoing to overcome these issues (Urayama, 2013).

7. **Replacement of the defective gene inside patient cells**. A complementary therapeutic technique which, similarly to enzyme replacement therapy, relies on the ability of cells to take up functional enzymes from another source is organ transplantation. The donor organ, or cells in the case of a haematopoietic stem cell transplant, contain the wild-type gene and therefore secrete the functional enzyme which can compensate for the systemic deficiency. Hurler syndrome is caused by a deficiency of α-L-iduronidase which functions to degrade heparan and dermatan sulphate within lysosomes. Patients are affected by progressive hepatic, cardiac, neurological and bone disease often leading to death before ten years of age (Pastores *et al.*, 2007). Treatment with a haematopoietic stem cell transplant from a heterozygous or unrelated donor is recommended in these patients as, unlike enzyme replacement therapy, functional enzymes can target the central nervous system (Aldenhoven *et al.*, 2015). Indeed, following a successful transplant a substantial clinical improvement is often seen, including the resolution of hepatosplenomegaly, improvements in cardiac,

pulmonary and hearing function, decreased coarseness of facial features and stabilisation of neuropsychological function (Souillet *et al.*, 2003). Similarly, liver transplantation in patients with urea cycle disorders can eradicate hyperammonaemia and result in improved neurodevelopmental outcomes (Kim *et al.*, 2013).

In contrast to delivering wild-type copies of the defective gene within the cells of a donor, direct insertion of DNA into patient cells can be achieved using gene therapy. This is typically accomplished by packaging the DNA in viral vectors, which takes advantage of the natural ability of viruses to introduce their DNA into host cells to enable their replication and the production of viral proteins. To date, gene therapy has been shown to be efficacious in multiple IEM including lysosomal storage disorders, urea cycle disorders and organic acidurias. One example can be seen in patients with Canavan disease, a hereditary leukodystrophy caused by inadequate catabolism of N-acetylaspartate in the brain. This results in the aberrant myelination of neurons and associated morphological changes in the brain that cause intellectual disability, loss of skills, seizures, macrocephaly and death in the third decade of life. Therapy using an adeno-associated viral vector delivered via intraparenchymal infusions resulted in normalisation of cerebral N-acetylaspartate concentrations as well as a stabilisation of brain atrophy and some reduction in seizure frequency (Leone *et al.*, 2012). Regardless of the disease and route of wild-type gene administration, children who are transplanted/treated at a younger age will invariably have a better clinical outcome as established organ damage is not typically reversible (Wynn, 2011; Ginocchio and Brunetti-Pierri, 2016).

8. **Symptom management.** Despite the multiple treatment options described above that are available for patients with certain IEM, many disorders remain for which treatment is essentially symptomatic or palliative. This can include anti-epileptic drug treatment for seizures, cochlear implants for sensorineural hearing loss, organ transplantation in situations of single-organ failure, enteral feeding to optimise nutritional intake, intravenous fluid administration at times of intercurrent illness and physiotherapy to maintain muscle tone and physical abilities (Parikh *et al.*, 2009).

A shared theme amongst the therapeutic options described above is the fact that early initiation of treatment invariably results in improved outcomes. Therefore, making diagnoses in patients in whom an IEM is suspected is critical to ensure that appropriate therapy can commence before permanent, particularly neurological, damage occurs.

As is evident in Table 1.1.1, neurological symptoms are the most common feature of IEM with global developmental delay often being the intial presenting feature. These IEM that include neurological dysfunction as a prominent manifestation are known as neurometabolic disorders (NMD). Clinical features may include epilepsy, movement disorders, neuro-cognitive deficits, behavioural abnormalities and psychiatric disorders (Karimzadeh, 2015). Indeed, there are more than 90 NMD in which this dysfunction may be treatable (Tarailo-Graovac *et al.*, 2016). Neurological signs can become apparent at any age between the newborn period and adulthood.

Whilst in some cases disorders can cause abnormal brain development *in utero* resulting in gross structural abnormalities (Klouwer *et al.*, 2015), in most cases symptoms are caused by progressive destruction of mental, motor and perceptual functions due to the accumulation of toxic molecules. This can be evident as a non-specific neurodevelopmental delay, be acutely precipitated by environmental stresses such as infection or vaccination, or have a latent period whereby a number of years of seemingly normal development are followed by the regression of skills once cellular damage has reached a certain threshold (van Karnebeek and Stockler, 2012; Filiano, 2006). Whilst the vast majority of reported NMD/IEM cases have presented during childhood ($\leq$ 16 years), it is becoming increasingly recognised that this likely reflects the more severe and easily diagnosed spectrum of these conditions and that phenotypically milder forms of these disorders do not manifest until later in adulthood, typically with psychiatric signs, dementia, mood or behavioural disorders (Sedel *et al.*, 2007). Indeed, a recent publication suggests that NMD are likely to account for a significant proportion of neurodegenerative diseases secondary to a genetic cause, with prevalence data ranging between 1 - 13% (Shevell *et al.*, 2003; Masri and Wahsh, 2014).

In addition to the presence of one or more of the neurological features described above, in order to invoke a clinical suspicion of an underlying NMD in a patient other indications suggesting a metabolic abnormality should be present. These can be in the form of clearly abnormal biochemical parameters including levels of amino acids, organic acids, acylcarnitines or lactate, an abnormal pattern of transferrin glycoforms or aberrant mitochondrial respiratory chain activity. However, episodes of acute metabolic decompensation, unusual dietary habits, multi-systemic involvement, dysmorphism, liver dysfunction (particularly hepatosplenomegaly) and failure to thrive are also suggestive features. Characteristic electroencephalography (EEG) or neuroimaging findings may also be present. Despite NMD not being a common cause of epilepsy, seizures are a frequent symptom of NMD. These seizures are often refractory to treatment with conventional anti-epileptic drugs but may have a good response to supplementation with vitamins,

cofactors or dietary manipulation (Papetti *et al.*, 2013). Although many different seizure types may occur including epileptic encephalopathy, status epilepticus, infantile spasms, generalised tonic-clonic and myoclonic seizures, an EEG showing burst-suppression or hypsarrhythmia may point towards an NMD (Youssef-Turki *et al.*, 2011). Similarly, the existence of white matter or basal ganglia abnormalities, a leukodystrophy, or atypical metabolite peaks upon examination using magnetic resonance spectroscopy are highly suggestive of an NMD.

The challenges when diagnosing these pathologies are largely attributable to the clinical and genetic heterogeneity (including often non-specific or atypical presentations early on in the disease course) and lack of clinical awareness of rare entities. Indeed, patients with suspected NMD are frequently referred to specialist centres and undergo extensive and often invasive diagnostic testing. Despite this, diagnostic delays or difficulties establishing a definitive diagnosis are commonly encountered, with many such patients attending secondary and tertiary neurology clinics remaining undiagnosed (Verity *et al.*, 2010).

1.3 Current methods for the diagnosis of IEM/NMD

Just as the advent of novel technologies have been critical in the discovery and understanding of the pathogenic mechanisms underlying IEM (Section 1.1.1), these same techniques have been utilised to revolutionise the diagnosis of these disorders in a clinical setting.

1.3.1 *Profiling of analytes in biofluids*

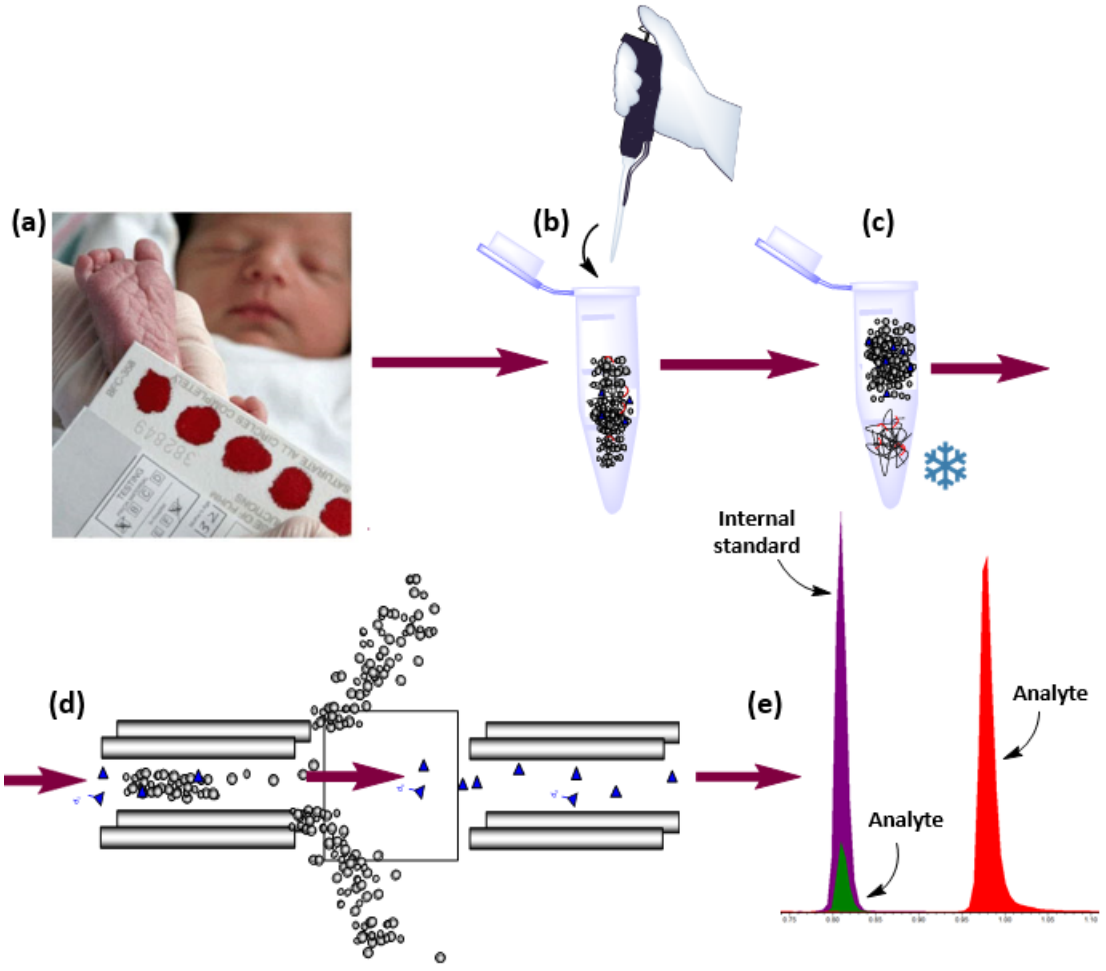
The primary method for the diagnosis, monitoring and investigation of IEM is through the analysis of metabolites. These assays are based on the fundamental principle underlying IEM, that dysfunction of an enzyme causes a block in a metabolic pathway that results in the accumulation of molecules upstream and the deficiency of molecules downstream of the block. A classic example is phenylketonuria, in which a deficiency of phenylalanine hydroxylase results in impaired conversion of phenylalanine to tyrosine. Hence, the analysis of amino acid concentrations in biological fluids from an affected patient will reveal high levels of phenylalanine and low levels of tyrosine (Vockley *et al.*, 2014). Indeed, many classes of metabolites may be analysed in this way, with particular patterns of abnormalities being diagnostic of a specific disorder or indicative of a wider sub-group of IEM. Relatively small groups of molecules sharing similar physicochemical properties and involved in related biological pathways are typically analysed concurrently; examples include organic acids, acylcarnitines, neurotransmitter amine metabolites, fatty acids and vitamins.

Many of these assays use mass spectrometry to quantify the concentrations of each analyte. Mass spectrometry is an analytical technique which allows the concurrent measurement of multiple analytes that are present at low concentrations in biological fluids including dried blood spots, whole blood, plasma, urine and cerebrospinal fluid (Figure 1.3.1). For some analytes such as acylcarnitines, tandem mass spectrometry (MS/MS) analysis is performed without prior chromatographic separation (Rinaldo *et al.*, 2008). When operated in scan mode, this technique facilitates the evaluation of the whole metabolite profile, as well as the detection of common drug artefacts and interfering compounds. However, greater analytical specificity can be achieved through the coupling of liquid chromatography separation to tandem mass spectrometry. This is typically performed using either high-performance (HPLC) or ultra-performance liquid chromatography (UPLC). The latter makes use of smaller stationary phase particles and can operate at higher pressures, resulting in superior analyte resolution and faster analysis when compared to HPLC. In contrast to MS/MS, HPLC-MS/MS and UPLC-MS/MS provide three levels of identification for each molecule, resulting in higher specificity and a superior ability to analyse complex mixtures. Firstly, HPLC/UPLC is used to separate compounds based on their physicochemical properties. The chromatographic system typically consists of a non-polar silica-based stationary phase and an eluting mobile phase; the composition of the mobile phase is then gradually altered to elute the analytes according to their polarity. Alteration of a number of parameters including the composition, gradient and flow rate of the mobile phase can change the ability to separate particular classes of molecules. Thus, the retention time of each molecule on the liquid chromatography column forms the first level of analyte identification.

Following chromatographic separation, the compounds sequentially enter the mass spectrometer and are ionised to form charged gas-phase ions. Electrospray ionisation is the most widely used technique as it is particularly suited to the analysis of polar molecules within a liquid matrix, as is commonly required for the analysis of biological fluids. Firstly, the liquid sample is pumped at high pressure through a fine capillary which is maintained at a high temperature and voltage to form a fine aerosol of highly charged droplets (Dass, 2007). Aided by a flow of heated nitrogen gas, the solvent within these droplets then evaporates to generate gas-phase ions (precursor ions). Depending on the physicochemical properties of the molecules to be measured, a positive or negative charge can be imparted to the ions which enables optimal sensitivity for the particular ion of interest. Following ionisation, the precursor ions enter the first quadrupole. This consists of four cylindrical metal rods aligned parallel to each other and through which a variable radiofrequency voltage can be applied (Pitt, 2009). This voltage can be tuned to allow the passage of ions with a specific mass-to-charge ratio (m/z) corresponding to that expected from the precursor ion of the analyte of interest. This selection constitutes the second level of analyte

identification. However, in a complex biological sample it is likely that multiple metabolites will produce an ion with a particular m/z. Therefore, the majority of mass spectrometry-based assays used within the clinical setting and this thesis use a triple quadrupole mass analyser. In this system, once selected precursor ions have passed through the first quadrupole they are bombarded with argon gas within a collision cell which fragments each species to form a product ion (Pitt, 2009).

Figure 1.3.1: The diagnosis of IEM using mass spectrometry coupled to liquid chromatography. **(a)** Metabolites are typically quantified in biofluids which can include dried blood spots, whole blood, plasma, urine and cerebrospinal fluid. **(b)** When preparing each sample, one or more internal standards are added to enable quantitation of each analyte. **(c)** Samples may also be subject to further processing steps including precipitation to remove contaminating proteins. **(d)** Each sample is then injected into a HPLC or UPLC system which separates the complex mixture based on the physicochemical properties of each molecule. Following this, they enter the mass spectrometer and are ionised to form precursor ions. Each precursor ion of interest is then selected and fragmented through collision with argon gas to form a number of smaller product ions. **(e)** These analytes are then identified based on their retention time and the mass-to-charge ratio (m/z) of their corresponding ions, and quantified by comparison of the signal to that of a known concentration of internal standard.



In an identical fashion to the first quadrupole, the second selects only product ions with the predicted m/z of the analyte of interest; this is the third level of identification. By combining the retention time and the m/z values of both the precursor and product ions metabolites can be identified and quantified.

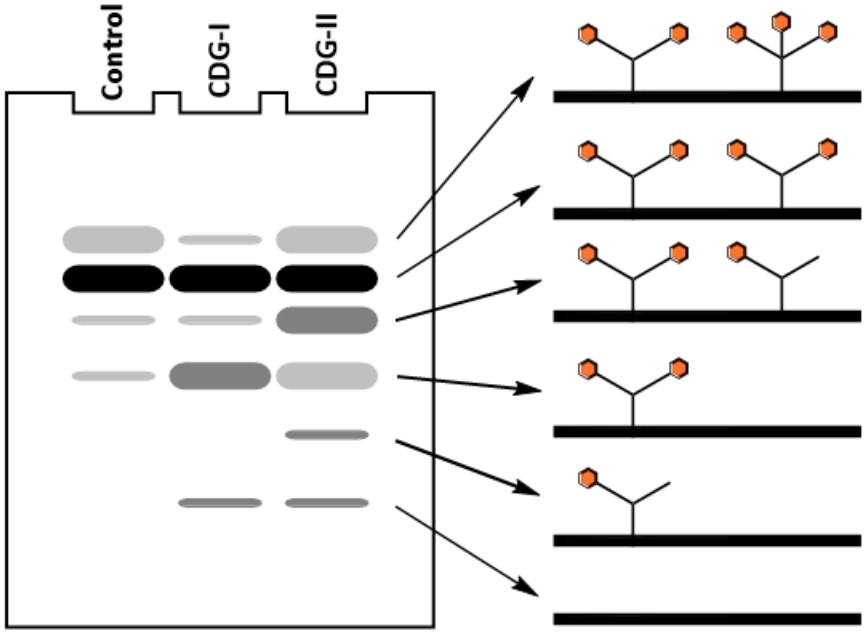
There are many advantages to these types of assays to measure the concentrations of various metabolites in biofluids, particularly within a clinical laboratory setting. These include high sample throughput, the ability to analyse multiple molecules of interest within a single assay and reduced cost compared to traditional radio-immunoassays. Therefore, the adoption and growth in the use of LC-MS/MS in these environments has been rapid (Grebe and Singh, 2011), with the majority of hospitals within the UK having at least one LC-MS/MS instrument and tertiary referral centres having dedicated departments housing multiple platforms. However, one limitation of these methodologies for the diagnosis of IEM is that findings may be non-specific and indicate a multitude of possible disorders. One example is an elevated concentration of alanine in plasma which can be indicative of long-term pyruvate accumulation, with concentrations above 450 μmol/L being used as a diagnostic factor for mitochondrial disorders (Wolf and Smeitink, 2002). However, the sensitivity of hyperalaninaemia for mitochondrial dysfunction is low as it may only be present in certain genetic defects or at times of physiological stress. In addition, mild abnormalities revealed by these classes of analyses can be overlooked or misinterpreted as being due to environmental factors. For example, non-specific aminoacdiuria can be a consequence of recent ingestion of a carbohydrate- or protein-rich meal or renal tubular dysfunction (Haas *et al.*, 2008).

1.3.2 *Specialised biochemical tests*

In contrast to the measurement of certain groups of metabolites in biofluids which is standard practice in the majority of hospitals for the diagnosis and monitoring of many IEM, there are a multitude of assays which are only available in specialist centres. This is usually because the required reagents are expensive, the assay is labour-intensive or the disorder for which the assay is used to diagnose is very rare. Examples of tests in this class include transferrin isoelectric focussing for congenital disorders of glycosylation (Figure 1.3.2), glycosaminoglycan analysis for mucopolysaccharidoses, measurement of globotriasylceramide for Fabry disease and α-aminoadipic semialdehyde quantitation for pyridoxine-dependent epilepsy.

Alternatively, the activity of certain enzymes can be assessed directly. These assays rely on the fundamental principle of IEM that dysfunction of an enzyme results in impaired conversion of the substrate to the corresponding product. It is the quantification of this activity that is measured

Figure 1.3.2: Transferrin isoelectric focussing patterns. The transferrin polypeptide has two N-linked polysaccharide chains which are branched with sialic acid residues (orange). Due to the fact that sialic acid has a negative charge, the differently glycosylated forms can be separated on a gel over a pH gradient. In normal individuals the most common oligosaccharide is that with four polysaccharide chains terminating in sialic acid residues. CDG type I (CDG-I) involves impaired synthesis of the lipid-linked oligosaccharide precursor or its transfer to the protein. CDG type II (CDG-II) is caused by dysfunctional processing of the protein-bound oligosaccharide chain.



in patient biofluids or tissue samples. One such group of diagnostic tests are the lysosomal enzyme screens which quantify the activity of between 7 and 11 enzymes in the leukocytes of patients presenting with suggestive features (e.g. neurodegeneration, hepatosplenomegaly or a cherry red spot on the macula of the eye). However, the majority of methods only examine the activity of a single enzyme at one time such as phosphomannomutase, arginase or galactokinase. These assays can be performed using substrates that are isotopically- or fluorescently-labelled and that can be detected using mass spectrometry and spectrofluorimetry, respectively (Blanchard *et al.*, 2008). Alternatively, colourmetric chemical reactions can be utilised and measured using a spectrophotometer (Bisswanger, 2014).

These specialised biochemical tests are beneficial as they allow the detailed examination of a specific disorder or small sub-group of disorders. In addition, enzyme assays provide a conclusive functional demonstration of the specific enzyme defect. Moreover, due to their availability only at a limited number of specialist centres, these tests are usually expensive and require a high degree of clinical suspicion to be requested. Whilst a positive result is diagnostic in an affected patient and the degree of residual enzyme activity may give some indication of prognosis (Koprivica

et al., 2000), identification of the pathogenic mutation(s) is desirable for pre-natal testing of future pregnancies within affected families.

1.3.3 *Single-gene Sanger sequencing*

The traditional approach for genetic diagnosis involves the amplification of all exons and intron-exon boundaries of a candidate gene using polymerase chain reaction (PCR) prior to Sanger sequencing. This method, also known as the "chain termination" or "dideoxy" sequencing, was first described by Sanger *et al.* (1977). Firstly, the DNA of interest is isolated and denatured to form single-stranded DNA. An oligonucleotide primer then binds to its complementary template sequence and free nucleotides are incorporated by DNA polymerase to facilitate *de novo* synthesis of a new DNA strand. However, in addition to normal deoxynucleotides (dNTPs), the reaction mix also contains di-deoxynucleotides (ddNTPs) which lack a 3'-hydroxyl group required for the formation of the phosphodiester bond between two adjacent nucleotides. These ddNTPs are also fluorescently labelled. The repetition of this process therefore results in the formation of multiple DNA fragments of differing lengths, each of which terminates in one of the four labelled ddNTPs.

In order to determine the sequence of each DNA strand, these fragments must then be aligned in order of size. In the vast majority of clinical laboratories, this analysis is performed by capillary electrophoresis. Each sample is loaded into a glass capillary and the migration of each fragment is initiated by an electric field. DNA has a backbone of negatively charged phosphate groups, thus the charge of a particular DNA fragment and the speed of its electrophoretic migration is determined by the number of nucleotides within it. The identity of the terminating amino acid is determined through laser excitation of each fragment when it reaches the end of the capillary, as the label of each ddNTP (A, T, G and C) emits light at a specific wavelength. The DNA sequence in the patient can then be aligned and compared to that of the wild-type sequence to detect any base alterations that may result in disease.

Sanger sequencing has been a mainstay of geneticists for over 30 years and remains the "gold-standard" DNA sequencing technique. With advances in automation, this methodology can be used to read continuous sequences of up to 1000 bp in length with a base-calling accuracy of up to 99.999% at a cost in the order of £0.35 per kilobase (Shendure and Ji, 2008). It can be used to identify the vast majority of mutations causing IEM including missense and nonsense mutations, insertions and deletions. However some pathogenic variants can be missed unless they are specifically searched for. These can include deep intronic splice variants as in leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation (van Berge *et al.*, 2014) or whole gene deletions and duplications as in Pelizaeus-Merzbacher

disease (Lee *et al.*, 2006). The latter can be identified using alternative methods for analysing copy number variants (CNVs) such as array comparative genomic hybridisation (aCGH) or multiplex ligation-dependent probe amplification (MLPA) and indeed, these investigations are often requested alongside standard sequencing. However, Sanger sequencing is accompanied by the obvious draw-back of requiring candidate genes for sequencing, which in turn often requires specialist clinical knowledge of the group of disorders in question and their genotype-phenotype correlations. Even if a specific disorder is suspected, clinical-grade single-gene sequencing is only available for a small subset of genes which often vary between centres. Indeed, advances in chemistry and bioinformatics, alongside the advent of "next-generation" sequencing methodologies have supplanted Sanger sequencing in many cases, especially for the analysis of large numbers of genes. These technologies will be described in Sections 1.3.4 and 1.3.5.

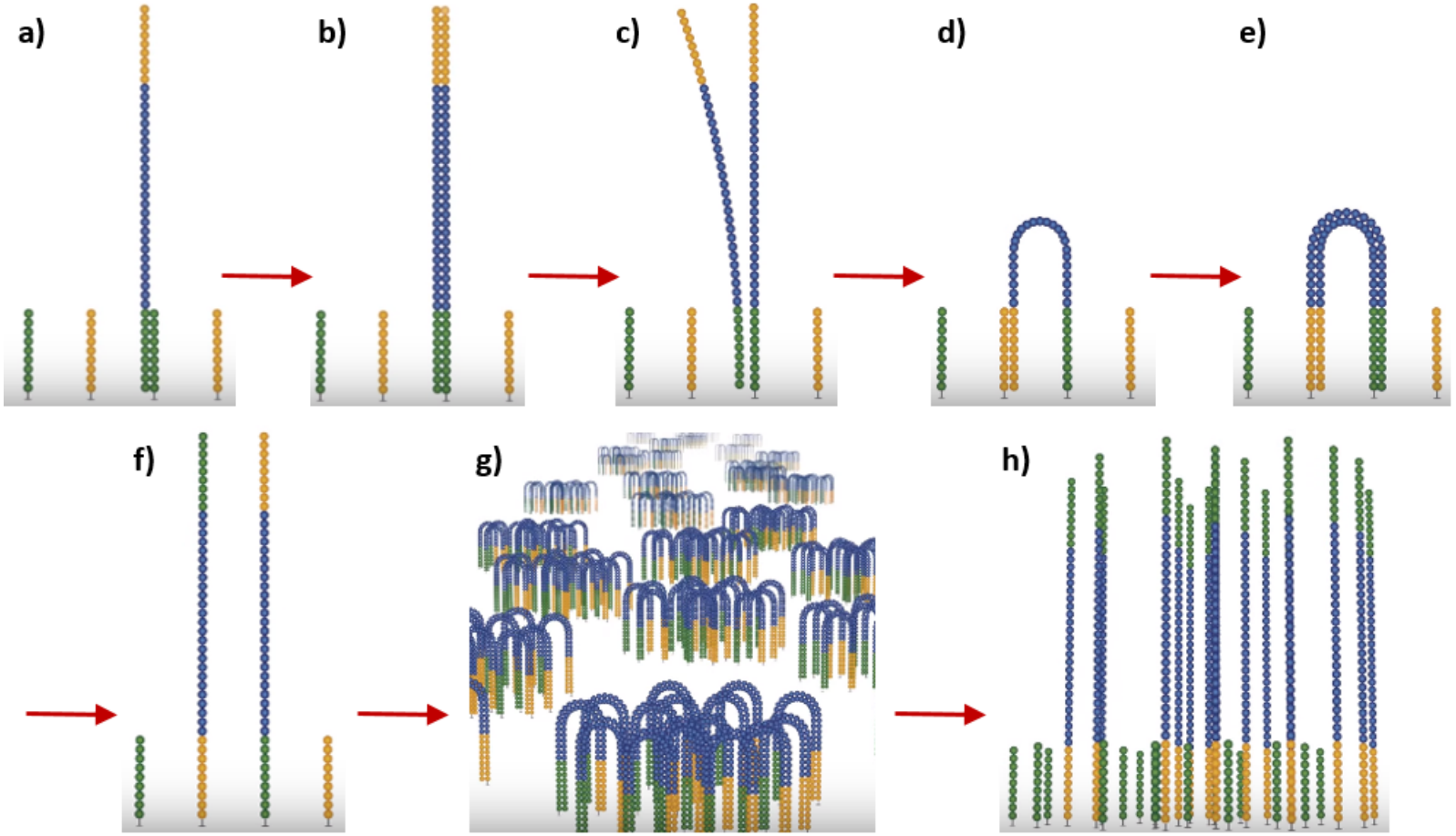
1.3.4 *Gene panels targeting multiple disease genes*

In 2001, Lander *et al.* published the sequence of the first human genome. This extraordinary collaborative effort took 15 years and cost almost 3 billion dollars to complete. In the fifteen years since this accomplishment, the demand for low-cost sequencing has driven the development of next-generation sequencing (NGS) technologies, facilitated by advances in microscopy, surface and nucleotide chemistry, computation and data storage (Shendure and Ji, 2008). This has occurred to such an extent that the latest platforms are capable of sequencing more than 45 human genomes in a single day for approximately $1000 each, with the data output for NGS applications more than doubling each year (Stein, 2010).

Although many NGS technologies exist, each making use of different biochemistry and detection methodologies, the Illumina sequencing platforms have emerged as the most widely-used choice in both clinical and research settings (Buermans and den Dunnen, 2014). Accordingly, the Illumina MiSeq and HiSeq 2500 platforms have been utilised throughout this thesis. The concept behind this technology is similar to Sanger sequencing in that DNA polymerase catalyses the sequential incorporation of fluorescently labelled dNTPs into a DNA template and each nucleotide is identified through laser excitation. The difference is that instead of sequencing only a single DNA fragment at one time, NGS performs this process concurrently across millions of small DNA fragments. Firstly, the sequencing library is prepared; this procedure depends on the capture methodology being used and is generally achieved using random fragmentation or enzymatic digestion. After isolating the desired genomic regions and fragmenting the DNA into strands of approximately 150 bp in length, adapters are annealed to the 3' and 5' end and the fragments are PCR amplified. These adapters contain (i) binding sites for the forward and

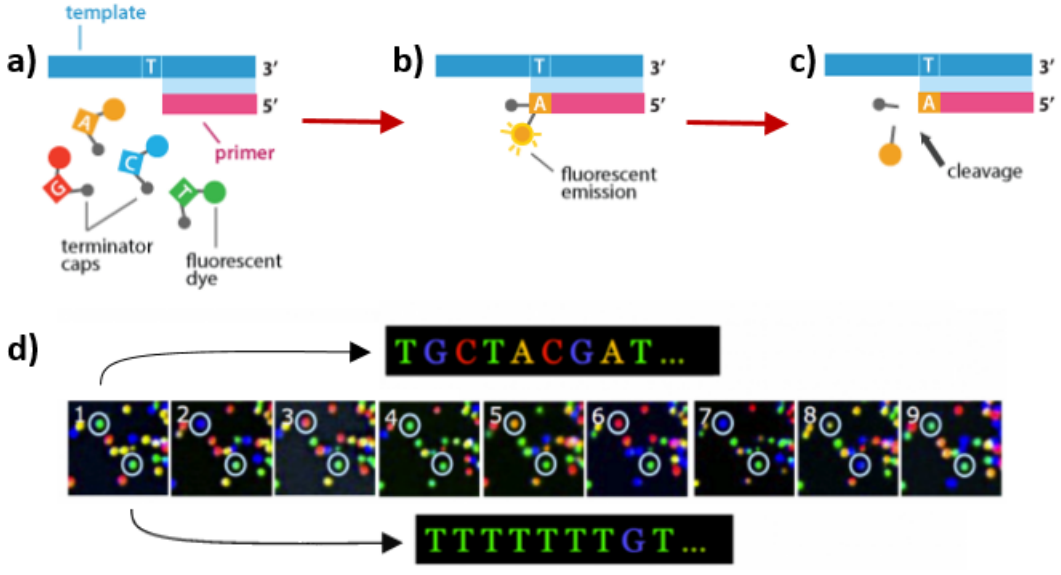
reverse sequencing primers, (ii) indexes to enable multiplexing of samples and the subsequent de-coding of data and (iii) regions complementary to oligos anchored to the flow cell (a glass slide containing small fluidic channels through which reagents can be pumped) to enable the immmobilisation of each DNA fragment on the solid sequencing platform. The library is then loaded into the sequencing flow cell and the DNA fragments are captured on the lawn of oligos which are complementary to either the 3' or 5' adapters. Each fragment is then amplified through bridge-amplification to generate distinct clonal clusters (Figure 1.3.3).

Figure 1.3.3: Cluster generation using bridge-amplification utilised by Illumina technologies. **(a)** The forward DNA strands bind to the complementary oligos on the flow cell surface. **(b)** DNA polymerase synthesises the complementary reverse strand. **(c)** The double-stranded molecule is denatured and the original forward template is washed away. **(d)** The strand folds over and the adapter region hybridises to the second type of oligo imbolised on the surface of the flow cell. **(e)** DNA polymerases again synthesise the complementary strand forming a double-stranded bridge. **(f)** The bridge is then denatured, resulting in two single-stranded copies of the DNA fragment that are tethered to the flow cell. **(g)** This process is then repeated multiple times and occurs simultaneously for millions of clusters, resulting in clonal amplification of all the DNA fragments. **(h)** The reverse strands are then cleaved and washed off leaving only the forward strands. Finally, the 3' ends are blocked to prevent unwanted priming. Adapted from www.illumina.com.



Sequencing then begins with the binding of the first sequencing primer to the immobilised DNA fragments and their extension to produce the first read. The four nucleotides (A, T, G and C) utilised in this method have been modified in two ways: each is reversibly attached to (i) a fluorescent label with a unique emission wavelength and (ii) a terminator group. With each cycle, the four nucleotides compete for addition to the growing DNA strand (Figure 1.3.4a). Only one is incorporated based on the sequence of the template. After the addition of each nucleotide, the clusters are excited by a laser and a characteristic fluorescent signal is emitted (Figure 1.3.4b). Both the fluorescent label and the terminator group are then cleaved and washed away (Figure 1.3.4c). The number of cycles determines the length of the read, whereas the emission wavelength along with the signal intensity determines each base call. For a given cluster, all identical strands are read simultaneously and millions of clusters are sequenced in a massively-parallel process (Figure 1.3.4d). The indexes of each DNA fragment are also sequenced in the same way which allows for the computational determination of which sequences belong to each patient when multiple samples have been multiplexed in the same run. Finally, each read (forward and reverse) is aligned to the reference genome for variant identification.

Figure 1.3.4: Sequencing-by-synthesis methodology on which Illumina technologies are based. **(a)** The reversibly terminated nucleotides and DNA polymerase are added. **(b)** Each complementary nucleotide is incorporated and excited sequentially to produce a fluorescent emission which is detected. **(c)** Both the label and terminator are cleaved and **(d)** the process is repeated to achieve massively-parallel sequencing.



This NGS technology has lead to a substantial increase in the identification of new disease-causing genes and the expansion of genotype-phenotype relationships. Indeed, increased reliability, robustness and the advent of novel capture techniques have facilitated its introduction into routine genetic diagnostics in the clinical setting. This has been particular useful for groups of disorders

characterised by great phenotypic and genetic heterogeneity such as intellectual disability (Vissers *et al.*, 2016) and epilepsy (Møller *et al.*, 2015) that require the analysis of many candidate genes. Generally, there are three ways to implement NGS for the diagnosis of disorders such as IEM/NMD: (i) targeted gene panel sequencing of a subset of genes, (ii) whole exome sequencing and (iii) whole genome sequencing. All share the ability to perform untargeted screening of multiple genes, which bypasses the need for specialist knowledge of individual candidate defects.

Targeted gene panel sequencing involves selecting a subset of protein-coding genes for analysis, typically those in which mutations are already known to cause a disorder within the group of interest. This approach has been increasing in prevalence over the past five years and is now broadly available in clinical genetics services of many healthcare systems worldwide including the National Health Service (www.ukgtn.nhs.uk). Whilst many panels exist targeting small subsets of genes often associated with specific phenotypic features such as inflammatory bowel disease, deafness or epilepsy (Kammermeier *et al.*, 2014; Tekin *et al.*, 2016; Lemke *et al.*, 2012; Trump *et al.*, 2016), currently panels for IEM are limited and focus on small subgroups of disease including congenital disorders of glycosylation and mitochondrial disorders (Jones *et al.*, 2013; DaRe *et al.*, 2013). Given the degree of phenotypic heterogeneity and overlap that is commonly seen within and between subgroups of IEM, an extended gene panel approach may be advantageous in this population.

Gene panel approaches have several advantages compared to whole exome or genome sequencing, including greater average sequencing depth (i.e. the number of reads in which any given nucleotide of a target sequence is represented) due to a reduction in the number of target regions. This parameter is critical for the identification of heterozygous variants and in the diagnosis of IEM, where many children born to non-consanguineous parents would be expected to harbour compound heterozygous mutations. Further advantages also include lowered cost per sample due to greater multiplexing capability and reduced time required for data processing. The clinical implications of any identified variants are also easier to interpret as only genes previously associated with a disease phenotype are sequenced. Indeed, this restriction reduces the identification of incidental findings and variants of uncertain significance (Table 1.3.1). However, these factors can also be disadvantageous as diagnostic success will only be achieved if the pathogenic gene is included in the panel design.

1.3.5 *Whole exome (WES) and genome (WGS) sequencing*

On a research basis, whole exome sequencing (WES) is the most widely used targeted sequencing method. The exome is defined as the part of the genome which when transcribed, remains within

the RNA after intron splicing and therefore constitutes the protein-coding regions. These regions represent less than 2% of the human genome but harbour the majority of known pathogenic variants. Thus, sequencing of these regions is a cost-effective alternative to whole genome sequencing (WGS) which is defined as the determination of the complete DNA sequence of an individual. Both techniques make use of the same NGS technologies as for gene panel sequencing, with Illumina platforms being the most popular due to their superior capacity for data output.

Many recent studies have successfully shown that WES or WGS can be used in a research environment to identify the genetic basis of rare disorders including IEM/NMD (Tarailo-Graovac *et al.*, 2016). Whilst WES, and to a much lesser extent WGS, are being adopted in many clinical centres internationally (Rabbani *et al.*, 2014), currently in the UK this approach is only offered by private companies or as part of a research study. In the case of screening offered by commercial companies, the prices of these investigations are often variable and turn-around times for research studies can range between three months and in excess of one year. However, three years ago Genomics England (funded by the UK Department of Health) launched the 100,000 Genomes Project which aims to perform WGS for NHS patients with rare diseases and those with cancer. This initiative is ongoing and to-date the genomes of 11,000 individuals have been sequenced (as of July 2016). Both WES and WGS have advantages compared to gene panel sequencing (Table 1.3.1), the most obvious of which is the ability to detect variants in novel disease-causing genes and therefore not relying on the prior identification of a likely disease class. However, this can also be a disadvantage as determining the significance of variants within genes that have not previously been associated with a disease in humans may not be possible, especially within a clinical diagnostic setting where further functional studies would be impractical.

Table 1.3.1: Comparison between targeted gene panel, whole exome (WES) and whole genome sequencing (WGS). -, no; ✓, yes. Adapted from Sun *et al.* (2015).

	Targeted gene panel	WES	WGS
Relative cost compared with WES	Depends on size of gene panel	1	~ 3
Coverage of target regions	Up to 100% when complemented with in-filling of inadequate regions using Sanger sequencing	97.5%	>97.5%
Analysis of novel disease genes	-	✓	✓
Intronic variants (> 30 bp from splice site)	-	-	✓
Incidental findings	-	✓	✓

One limitation of both gene panel sequencing and WES is that because they both target only exonic regions, any pathogenic mutations in intronic or regulatory regions will not be identified

(Table 1.3.1). In addition, due to NGS technologies being based on sequencing short (~150 bp) DNA fragments, if copy number variants (duplications, insertional transpositions, deletions or complex genomic rearrangements) are present that span more than one read or when the breakpoints are intronic they will not be detected. An advantage of WGS is that these types of potentially pathogenic structural changes can be detected as the coverage not only encompasses the entirety of each copy number variation, but also is more even across the genome (Gilissen *et al.*, 2014). The methods that exist to enable the detection of these variations are described in Section 3.3.3. However, whilst the advent of WES and WGS has revolutionised the field of molecular medicine, further guidelines regarding the reporting of medically actionable findings that are unrelated to the original indication for the test (e.g. *BRCA1* and *BRCA2* mutations) and advances in bioinformatic pipelines and data storage are required prior to the establishment of these technologies in routine clinical practice.

1.4 The importance of functional confirmation and characterisation of identified variants

As already discussed, the interpretation of the significance of the millions of variants being identified through WES or WGS presents a challenge in both a research and clinical setting (Stein, 2010). Indeed, the clinical significance of any sequence variant lies along a gradient, ranging from those in which the variant is almost certainly pathogenic to those that are almost certainly benign. Current guidelines suggest using a combination of *in silico* tools to determine the conservation of the mutated amino acid position across species and the frequency of each variant within healthy human populations; as well as incorporating functional consequence predictions, gene-specific information in the literature (e.g. annotation of critical protein domains and mutation hot-spots) and the phenotypic correlation between the patient being investigated and those previously reported with mutations in the same gene, to classify the potential pathogenicity of sequence variants (Richards *et al.*, 2015). Genetic testing of additional family members is also recommended where samples are available to confirm co-segregation of the variant with disease.

However, despite these recommendations, it has been reported that up to 27% of mutations cited in the medical literature as pathogenic have been misannotated and are in fact common polymorphisms (Bell *et al.*, 2011). These false assignments of pathogenicity can then have severely detrimental consequences on patient care as incorrect prognostic, therapeutic and reproductive advice may be given (MacArthur *et al.*, 2014). Indeed, the only way to conclusively support pathogenicity is through functional demonstration of the defect in patient-derived samples and this should be carried out whenever possible. However, each type of functional assay differs in

their ability to recapitulate the *in vivo* environment within each patient and therefore accurately determine any functional impairment. For example, determination of enzyme activity in a patient-derived skin biopsy or an animal disease model provides stronger evidence than the same experiment carried out using a protein generated through *in vitro* overexpression.

Unfortunately, functional assays are only established to assess the pathogenicity of variants in a minority of genes (Sections 1.3.1 and 1.3.2). Thus, novel biochemical methods are required not only to complement and support NGS findings, but also to further the understanding of the pathogenesis of novel disorders and the genotype-phenotype relationships that often complicate the treatment of patients with IEM. It is the development and application of such novel techniques, alongside the utilisation of next-generation sequencing technologies that will form the basis of my thesis.

1.5 AIMS OF THIS THESIS

The overarching aim of this thesis is to determine the cellular and molecular aetiologies for the clinical phenotypes observed in patients with undiagnosed neurometabolic disorders. This includes:

1. Developing and validating a targeted gene panel sequencing approach incorporating all genes known to cause inborn errors of metabolism to improve the diagnosis of these patients in clinical practice.

2. Using whole exome sequencing for the diagnosis of patients in which extensive genetic and biochemical testing has not identified a diagnosis.

3. Examining the heterogeneity of response to vitamin B_6 supplementation in patients with pyridox(am)ine 5'-phosphate oxidase deficiency in order to optimise treatment.

4. Utilising functional assays to characterise a novel inborn error of vitamin B_6 metabolism.

2

MATERIALS AND METHODS

2.1 MATERIALS

The following were purchased from Sigma Alrich (Poole, UK):

2-oxobutyrate, 2-oxoglutarate, acetic acid, acetonitrile, adenosine triphosphate (ATP), amidosulphobetanine-14 (ASB-14), ammonium acetate, ammonium bicarbonate, ampicillin, bacteriological agar, betaine, β-mercaptoethanol, borate buffer, calcium acetate, catalase, D-amino acid oxidase from porcine kidney, dimethyl sulphoxide (DMSO), dithioerythritol, D-lysine, flavin adenine dinucleotide, flavin mononucleotide (FMN), fluorenylmethyloxycarbonyl chloride (FMOC-Cl), formic acid, glycerol, heptafluorobutyric acid, hydrogen peroxide, iodoacetamide, kanamycin, kynurenic acid, kynurenic acid D5 [>98% atom %D], kynurenine, L-lysine, magnesium chloride ($MgCl_2$), methanol, Miller LB broth, mouse β-actin primary antibody, napthol AS-BI phosphate, ninhydrin, Nuclear Fast Red, o-aminobenzaldehyde, oxaloacetate, pararosaniline hydrochloride, perchloric acid, poly-D-lysine, potassium phosphate buffer, PROSC primary antibody, pyridoxal (PL), pyridoxal methyl D3 hydrochloride [>98% atom %D], pyridoxal 5'-phosphate (PLP), pyridoxamine (PM), pyridoxamine methyl D3 hydrochloride [>98% atom %D], pyridoxic acid (PA), pyridoxine (PN), pyridoxine 5'-phosphate (PNP), pyrophosphate buffer, pyruvate, Rainbow Marker, sheep serum, skim milk powder, sodium acetate, sodium hydroxide, sodium nitrate, sodium phosphate, thiourea, trichloroacetic acid, trifluoroacetic acid, Tris-HCl buffer, Triton X-100, urea.

The following were purchased from Thermo Fisher (Loughborough, UK):

1-Step Human Coupled IVT Kit (containing HeLa cell lysate, accessory proteins, reaction mix, nuclease-free water), CyroTube Vials, MicroAmp Fast Optical 96-well Reaction Plates, Alexa Fluor 610 secondary antibody, non-DEPC-treated nuclease-free water, Novex ECL Chemiluminescent Substrate Reagent Kit, NuPAGE Antioxidant, NuPAGE LDS Buffer, NuPAGE MOPS SDS Running Buffer, NuPAGE Novex 4-12% Bis-Tris Protein Gels, NuPAGE Sample Reducing Agent, Pierce BCA Protein Assay Kit, Pierce Bovine Serum Albumin Standard (Pre-diluted Set), protease inhibitor cocktail, pT7CFE1 vector, Qubit dsDNA Broad Range (BR) Assay Kit, Qubit dsDNA High Sensitivity (HS) Assay Kit, RIPA buffer, RNase H, RNase-free water (dH_2O), sequencing primers, SuperScript III First-Strand Synthesis System for RT-PCR, TaqMan Gene Expression Assay, TaqMan Gene Expression Master Mix, Tempus Blood RNA Tube, Tempus Spin RNA Isolation Kit.

The following were purchased from VWR (Lutterworth, UK):

Acetone, dimethylformamide, ethanol, hydrochloric acid, isopropanol, Oil Red O, paraformaldehyde, sterile glass coverslips, Sudan Black, Superfrost Plus Micro Slides.

The following were purchased from Invitrogen (Carlsbad, CA):

1 kilobase pair ladder, 100 base pair ladder, orange loading dye, SOC Medium, TOP10 chemically competent cells, TOPO TA Cloning Kit (containing TOPO vector, salt solution), Tris-Borate-EDTA (TBE) buffer, UltraPure agarose.

The following were purchased from New England Biolabs (Ipswich, MA):

ApaI, Buffer 4, calf intestinal phosphatase, EcoRI, EcoRI buffer, Exonuclease I, NdeI, Phusion High-Fidelity DNA polymerase kit (containing Phusion GC buffer, dNTPs, DMSO, Phusion DNA polymerase), SAP dilution buffer, shrimp alkaline phosphatase (SAP), T4 DNA ligase, T4 DNA ligase buffer.

The following were purchased from Agilent Technologies (Cedar Creek, TX):

Agilent High Sensitivity DNA Assay, Agilent SureSelect v4 (51Mb) kit, Custom HaloPlex Target Enrichment Kit, DpnI, Herculase II Fusion DNA Polymerase Kit (including Herculase II reaction buffer, dNTPs, primer 1 and 2, Herculase DNA polymerase), PfuUltra High-Fidelity DNA polymerase, QuikChange II XL Site-Directed Mutagenesis Kit (containing reaction buffer, dNTP mix, QuikSolution), TBST buffer, XL-1 Blue competent cells.

The following were purchased from Illumina (San Diego, CA):

HiSeq cartridge, Hybidization buffer (HT1), MiSeq cartridge, MiSeq v2 Reagent (300 cycles) Kit, PhiX virus library, Rapid SBS (3 x 50 cycle) Kit, Tris-HCl (pH 8.5) [0.1% Tween 20].

The following were purchased from Life Technologies (Paisley, UK):

Dulbecco's phosphate buffered saline (PBS), fetal bovine serum (FBS), HAMS F-10 media, penicillin, streptomycin, trypsin-EDTA.

The following were purchased from Agar Scientific (Stansted, UK):

Agar 100 Resin, alcoholic uranyl acetate, benzyldimethylamine, cacodylate buffer, copper grids, dodecenyl succinic anhydride, glutaraldehyde, methyl nadic andydride, osmium tetroxide, propylene oxide, Reynold's lead citrate.

The following were purchased from Qiagen (Hilden, Germany):

DNeasy Blood & Tissue Kit, DNeasy mini spin columns, QIAprep Spin Miniprep Kit, QIAquik Gel Extraction Kit, QIAshredder spin columns, Rnase A, RNeasy Mini Kit.

The following were purchased from Leica Biosystems (Newcastle, UK):

Bond Epitope Retrieval Solution 2, Bond primary antibody diluent, Bond Wash Solution, Eosin, Harris haematoxylin, Mayers haematoxylin, Mixed DAB Refine Solution, Post Primary Solution.

The following were purchased from Promega (Madison, WI):

5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) in 2% N, N-dimethyl formamide, Mass Spectrometry Grade Trypsin Gold, Tris-EDTA (TE) buffer.

The following were purchased from Waters (Manchester, UK):

[glu1]-fibrinopeptide B, enolase peptides standard, TruView LCMS Certified, Total Recovery Vials.

The following were purchased from Bioline (London, UK):

BIOTAQ DNA polymerase kit (including PCR reaction buffer, $MgCl_2$, Taq DNA polymerase), dNTP mixture.

The following were purchased from Applied Biosystems (Waltham, MA):

Big Dye sequencing buffer, Big Dye Terminator version 1.1.

The following were purchased from Bio-Rad (Hemel Hempstead, UK):

Dual-chambered cell couting slides, trypan blue.

C18 solid phase extraction columns were purchased from Biotage (Ystrad Mynach, UK), formaldehyde was purchased from Acquascience (Uckfield, UK), propylene glycol was purchased from Amresco (Solon, OH), Luxol Fast Blue Stain Kit was purchased from Atom Scientific (Hyde, UK), Vectashield aqueous mounting medium was purchased from Vector Laboratories (Peterborough, UK), AMPure XP beads were purchased from Beckman Coulter (High Wycombe, UK), donkey anti-rabbit IgG-HRP was purchased from Santa Cruz Biotechnology (Heidelberg, Germany), goat anti-mouse IgG-HRP was purchased from DAKO (Ely, UK), LAMP-2 antibody was purchased from Abcam (Cambridge, UK), p62 antibody was purchased from BD Biosciences (Oxford, UK). Pyridoxine D2 hydrochloride (5-hydroxymethyl-D2) [>98% atom %D] and α-aminoadipic acid D2 [>98% atom %D] were purchased from CDN Isotopes (Thaxted, UK), 4-pyridoxic acid D2 [>98% atom %D] was purchased from Buchem BV (Apeldoorn, The Netherlands), DL-proline D7 [>98% atom %D] was purchased from CK Isotopes (Leicester, UK). Pyridoxal 5'-phosphate D2 was kindly supplied as a gift by Professor Coburn, Department of Chemistry, Indiana University, Purdue University, Forte Wayne.

2.2 ETHICS STATEMENT

All patient-derived samples were obtained following the approval of the study (REC Ref. 13/LO/0168) by the ethics committee of Great Ormond Street Hospital for Children (GOSH), London, UK.

2.3 DNA EXTRACTION AND SANGER SEQUENCING

2.3.0.1 *Automated DNA extraction*

Genomic DNA was extracted from EDTA blood using an AutoGenFlex STAR automated system (AutoGen, Holliston, MA) according to the manufacturer's protocol at the NE Thames Regional Genetics Service Laboratories, GOSH, London.

2.3.0.2 *Manual DNA extraction*

Genomic DNA was extracted from EDTA blood using the DNeasy Blood & Tissue Kit as described by the manufacturer. 20 μL of proteinase K, 100 μL of blood and 100 μL of phosphate buffered saline (PBS) were mixed. 4 μL of RNase A (100 mg/ml) was added and incubated for 2 minutes at room temperature to remove contaminating RNA. 200 μL of Buffer AL was then added and samples were immediately vortexed and incubated at 56°C for 10 minutes. 200 μL of 100% ethanol was added and mixed to yield a homogeneous solution. The mixture was transferred to a DNeasy mini spin column, centrifuged at 6000 x g for 1 minute and the flow-through was discarded. 500 μL of Buffer AW1 was then added, centrifuged at 6000 x g for 1 minute and the flow-through was discarded. 500 μL of Buffer AW2 was added and centrifuged at 20,000 x g for 3 minutes to dry the DNeasy membrane as any residual ethanol can inhibit downstream reactions. The flow-through and collection tube was discarded. Columns were placed in clean microcentrifuge tubes and 200 μL of Buffer AE was applied directly to the column membrane. Columns were incubated at room temperature for 1 minute and centrifuged at 6000 x g for 1 minute to elute the DNA. DNA samples were stored long-term at -20°C.

2.3.1 *Primer design*

In order to design intronic primers for amplification of patient DNA by Polymerase Chain Reaction (PCR), genomic sequences of the gene of interest annotated with all known variants were downloaded from Ensembl (www.ensembl.org) and the region of interest (usually one or

more exons) and approximately 200 bp of flanking sequence was imported into Primer3 version 4.0.0 (www.primer3.ut.ee). Primers were picked using default settings. These included selecting primers between 18-23 bp in length, a G/C content between 30% - 70% and a predicted melting temperature (T_m) between 57 and 62°C. Primer-BLAST (www.ncbi.nlm.nih.gov/tools/primer-blast) was subsequently used to ensure that primers were specific for the gene of interest by comparing the chosen sequences to the human genome. Primers were redesigned if both forward and reverse primers were complementary not only to the region of interest, but also additional regions.

In the case of cDNA or plasmid sequencing, primers were designed in the same way as above with the exception that primers were designed to "walk" across the length of the sequence by generating overlapping PCR products to enable complete target coverage. All primer sequences and PCR conditions are detailed in the Appendices.

2.3.2 *Amplification of target genes from genomic DNA using Polymerase Chain Reaction (PCR)*

2.3.2.1 *PCR conditions*

Typical PCR reactions were carried out in a total volume of 25 μL using the conditions outlined in Table 2.3.1. When non-specific amplification was present, the volume of nuclease-free water could be reduced and PCR enhancing agents were added, alongside the optimisation of other reaction parameters (see Section 2.3.2.2). Negative controls were prepared for each reaction containing nuclease-free water instead of genomic DNA to check for any contamination. All samples were prepared and kept on ice until amplification.

Table 2.3.1: Typical PCR reaction mix.

Reagent	Volume (μL)
Nuclease-free water	15.55
10X PCR reaction buffer	2.5
$MgCl_2$ (25 mM)	0.75
dNTP mixture (10 mM)	2.5
Forward primer (10 μM)	1.25
Reverse primer (10 μM)	1.25
Taq DNA polymerase (5 U/μL)	0.2
Genomic DNA (100 ng)	1

Amplification was carried out using a Veriti 96-well Thermal Cycler (Thermo Fisher, Loughborough, UK). Typical cycling conditions are outlined in Table 2.3.2.

Table 2.3.2: Typical PCR cycling conditions. Amplification occurs in three stages: denaturation of the double-stranded DNA, annealing of primers to the complementary DNA at a variable temperature and extension of the DNA template copy by the 5' → 3' activity of *Taq* DNA polymerase. T_m; melting temperature (details of optimised T_m for each set of primers are detailed in the Appendices).

Step	Conditions
1	96°C for 5 mins
2	96°C for 30 secs
3	T_m °C for 30 secs
4	72°C for 30 secs
5	Repeat steps 2-4 34 times for a total of 35 cycles
6	72°C for 10 mins

2.3.2.2 *Visualisation of PCR products by agarose gel electrophoresis*

PCR products were analysed by agarose gel electrophoresis to determine both the specificity of amplification and the size of the products. Gels were prepared at different percentages depending on the requirements of downstream applications. Typically 1% (w/v) agarose gels were prepared using 1 g of UltraPure Agarose in 100 mL of 1X Tris-Borate-EDTA (TBE) buffer [containing 45 mM of Tris-HCL, 45 mM of boric acid and 10 mM of EDTA (pH 8.0)]. The gel was poured into a tray containing a comb and allowed to set at room temperature for 30 minutes. The gel was then placed into an electrophoresis tank containing 1X TBE buffer.

Differing amounts of PCR products were loaded into the agarose gel wells depending on the initial DNA concentration, PCR conditions and requirements of downstream applications. Typically, 5 μL of PCR product was mixed with 3 μL of orange loading dye and loaded into the agarose gel wells. 5 μL of 100 base pair or 1 kilobase pair ladder (1 μg/μL) was loaded into the first lane in order to estimate PCR product size. Electrophoresis was typically carried out at 95 V for between 30-60 minutes depending on the size of the PCR product. Bands were visualised using a ChemiDoc MP System (Bio-Rad, Hemel Hempstead, UK) coupled to Image Lab Software. If, after analysis, a distinct band was not produced then the reaction conditions were varied to optimise the PCR reaction. This included changing the annealing temperature, $MgCl_2$ concentration and the addition of 1.25 μL of 100% dimethyl sulphoxide (DMSO) and/or 2.5 μL of 5 M betaine.

2.3.3 *Sanger sequencing*

2.3.3.1 *Purification of PCR products using ExoSAP*

10 μL of PCR product was added to 0.75 μL of dH_2O, 0.5 μL of exonuclease I, 1 μL of shrimp alkaline phosphatase (SAP) and 0.25 μL of SAP dilution buffer. Reactions were mixed then incubated at 37°C for 15 minutes followed by 80°C for 15 minutes in a Veriti 96-well Thermal Cycler (Thermo Fisher). Samples were stored on ice until Sanger sequencing.

2.3.3.2 *Sanger sequencing*

3 μL of PCR product that had been cleaned according to the protocol above was added to with 0.5 μL of Big Dye Terminator version 1.1, 1.5 μL of 5X Sequencing Buffer, 4 μL dH_2O and 1 μL (5 pmol/μL) of forward or reverse primer. Each reaction was then cycled using the parameters outlined in Table 2.3.3.

Table 2.3.3: Sanger sequencing thermal cycling parameters.

Step	Conditions
1	95°C for 2 mins
2	95°C for 20 secs
3	50°C for 10 secs
4	60°C for 3 minutes
5	Repeat steps 2-4 34 times for a total of 35 cycles

2.3.3.3 *DNA precipitation*

2 μL of 3 M sodium acetate and 50 μL of 100% ethanol was added to each sequencing product and incubated for 20 minutes at room temperature after vortexing. Samples were then centrifuged at 12,000 x g for 40 minutes and the supernatant was discarded. An additional 50 μL of 70% ethanol was added to wash the DNA, samples were centrifuged at 12,000 x g for 10 minutes and the supernatant was discarded. Finally, the samples were centrifuged at 1000 x g upside down for 1 minute to remove residual ethanol. DNA was resuspended in 0.1X Tris-EDTA buffer and sequenced on an ABI DNA Sequencer (Life Technologies, Paisley, UK) at the NE Thames Regional Genetics Service Laboratories, GOSH, London. Sequence data was analysed using Sequencher 4.10.1 software (Gene Codes, Ann Arbor, MI) and compared to reference sequences downloaded from Ensembl (www.ensembl.org).

2.4 GENE PANEL SEQUENCING

2.4.1 *Qubit quantitation*

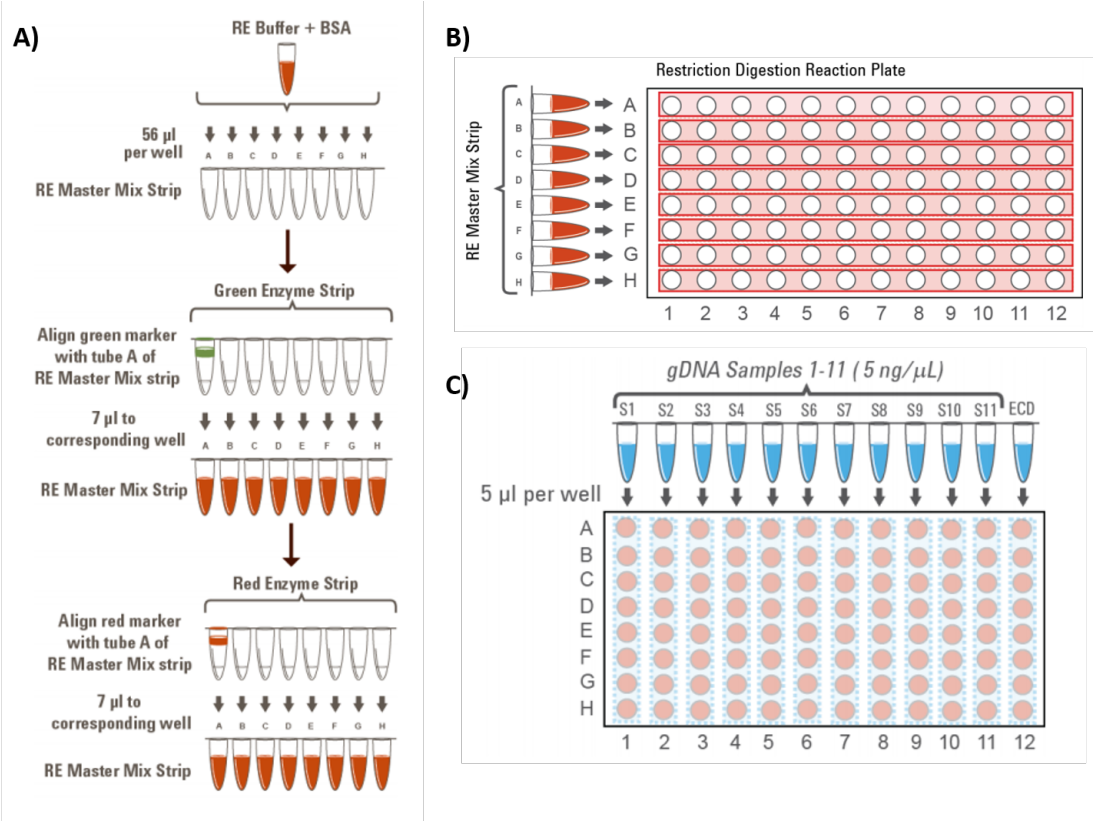
DNA was extracted as described in Section 2.3. DNA concentration was quantified prior to target capture using the Qubit dsDNA Broad Range (BR) Assay Kit according to manufacturer's instructions. Samples were quantified by mixing 2 μL of DNA and 198 μL of master mix [containing 1 μL of BR reagent and 199 μL of BR buffer per sample] and incubated for two minutes at room temperature. DNA concentration was then calculated using a Qubit 2.0 fluorometer (Thermo Fisher) and samples were diluted to a final concentration of 5 ng/μL with non-DEPC-treated nuclease-free water to a final volume of 45 μL.

2.4.2 *Target capture*

Library construction and capture hybridisation was performed using a Custom HaloPlex Target Enrichment Kit according to the manufacturer's protocol. DNA was processed in batches of 12 samples including one Enrichment Control DNA (ECD) sample to enable assessment of sample processing quality at two time points during the protocol. The 11 pre-diluted genomic DNA samples and 45 μL of ECD were digested in 8 different reactions by 16 restriction enzymes. Restriction enzyme master mixes were assembled in 0.2 mL 8-well strip tubes (Figure 2.4.1a). 476 μL of restriction enzyme buffer and 11.9 μL of bovine serum albumin solution were mixed and 56 μL of mixture was aliquoted into each well of an 8-well tube. 7 μL of the appropriate restriction enzymes from both the Green and Red Enzyme Strips were added to the master mixes (Figure 2.4.1a). The solutions were kept on ice throughout handling and mixed by gentle vortexing. 5 μL of restriction enzyme master mix was aliquoted row-wise into each well of a 96-well plate (Figure 2.4.1b). 5 μL of each genomic DNA sample was then aliquoted column-wise into each well of the 96-well plate (Figure 2.4.1c). The plate was sealed with adhesive plastic film, vortexed and briefly centrifuged. Each column then corresponds to one DNA sample digested in eight different restriction reactions.

The restriction digest plate was then placed in a thermal cycler and incubated at 37°C for 30 minutes followed by a low-temperature hold step at 8°C until restriction digestion validation could be carried out.

Figure 2.4.1: HaloPlex restriction enzyme digestion. A) Preparation of the restriction enzyme master mix strip for the 12-sample run. B) Aliquoting the restriction enzyme master mixes into the rows of a 96-well plate. Each row thus contains 5 μL per well of the same restriction enzyme combination. C) Aliquoting 5 μL of the 11 genomic DNA samples and the ECD sample column-wise into each well of the reaction plate.



2.4.3 *Restriction digestion validation*

Correct restriction enzyme digestion of the DNA was assessed using the Agilent High Sensitivity DNA Assay according to manufacturer's instructions. 4 μL of each of the eight Enrichment Control DNA (ECD) digestion reactions (Section 2.4.2) were transferred to 0.2 mL PCR tubes and incubated at 80°C for 5 minutes in a thermal cycler to inactivate the restriction enzymes. The eight digested ECD samples were processed and loaded onto a microfluidic chip (Agilent Technologies) and analysed within five minutes using an Agilent 2100 Bioanalyzer (Agilent Technologies) according to manufacturer's instructions. The ECD sample contains genomic DNA mixed with an 800 bp PCR product containing restriction sites for all the enzymes used in the digestion protocol. The presence of three prominent fragments (125 bp, 225 bp and 450 bp) over a smear of restriction DNA fragments between 100 and 2500 bp indicated correct sample

digestion. The undigested control showed genomic DNA bands larger that 2.5 kbp in size and a PCR product of 800 bp.

2.4.4 *Target enrichment and sample indexing*

Upon validation of correct restriction enzyme digestion, genomic DNA fragments were hybridised to the custom HaloPlex probe capture library. A hybridisation master mix was prepared by combining 650 μL of hybridisation solution and 260 μL of HaloPlex probes. 70 μL of master mix was then aliquoted into 12 0.2 mL PCR tubes. 10 μL of one specific indexing primer cassette was added to each tube containing master mix, using different indexes for each sample to be multiplexed. The identity of each Indexing Primer Cassette was noted for later sequence analysis. All 8 digestion reactions for each DNA sample were transferred to the appropriate hybridisation reaction tube and gently mixed to ensure inactivation of the enzymes. For the ECD sample, 32 μL of non-DEPC-treated nuclease-free water was also added in addition to the digested DNA samples to compensate for the volume removed for digest validation. The hybridisation reaction tubes were then incubated in a thermal cycler at 95°C for 10 minutes followed by 54°C for 16 hours overnight.

2.4.5 *Target DNA capture, ligation and elution*

Following target enrichment, the circularised target DNA-HaloPlex probe hybrids containing biotin, were captured using streptavidin beads. 520 μL of HaloPlex magnetic bead solution was transferred to a microcentrifuge tube and placed in a DynaMag-2 Magnet (Thermo Fisher) for 5 minutes. After the solution had cleared the supernatant was removed and 520 μL of capture solution was added to resuspend the beads. The hybridisation reactions were then removed from the thermal cycler and 40 μL of bead suspension was added to each reaction. The capture reactions were mixed thoroughly by pipetting up and down 15 times before incubating at room temperature for 15 minutes. Tubes were then transferred to an Agencourt SPRIPlate Super Magnet magnetic plate (Beckman Coulter, High Wycombe, UK) and the supernatant was allowed to clear for 30 seconds. The supernatant was removed, 100 μL of Wash Solution was added and the beads were resuspended by pipetting up and down 10 times. The tubes were then incubated at 46°C for 10 minutes in a thermal cycler and transferred again to the Agencourt SPRIPlate Super Magnet magnetic plate. The solution was allowed to clear and the supernatant was carefully discarded, taking care to remove any residual liquid.

Once the DNA had been captured, nicks in the circularised target DNA-HaloPlex probe hybrids were closed using DNA ligase. A DNA ligase master mix was prepared by adding 617.5 μL of ligation solution to 32.5 μL of DNA ligase. 50 μL of this master mix was then added to the beads in each DNA capture reaction tube. The beads were thoroughly resuspended by pipetting the mixture up and down 15 times. Subsequently the mixes were incubated at 55°C for 10 minutes in a thermal cycler. The tubes were then transferred to the magnetic plate, the solution was allowed to clear and the supernatant was carefully discarded. 100 μL of SSC Buffer was added to each tube and the beads were resuspended by pipetting up and down 10 times. The tubes were then returned to the Agencourt SPRIPlate Super Magnet magnetic plate, the solution was allowed to clear and the SSC Buffer was removed, making sure that no residual liquid remained. The beads were then resuspended in 25 μL of freshly prepared 50 mM sodium hydroxide and incubated at room temperature for 1 minute to elute the captured DNA.

2.4.6 *PCR amplification of target libraries*

Following DNA elution, the tubes were transferred to the Agencourt SPRIPlate Super Magnet magnetic plate and 20 μL of cleared supernatant from each tube was added to a PCR master mix tube (Table 2.4.1) kept on ice. The reactions were mixed by gentle vortexing and amplified using the cycling conditions outlined in Table 2.4.2.

Table 2.4.1: PCR reaction mix for amplification of HaloPlex target libraries. A Herculase II Fusion DNA Polymerase Kit (containing 100 mM dNTPs) was used for amplification.

Reagent	Volume (μL)
Nuclease-free water	16.1
5X Herculase II Reaction Buffer	10
dNTPs (25 mM each)	0.4
Primer 1 (25 μM)	1
Primer 2 (25 μM)	1
2 M acetic acid	0.5
Herculase II Fusion DNA Polymerase	1

Table 2.4.2: HaloPlex PCR cycling conditions.

Step	Conditions
1	98°C for 2 mins
2	98°C for 30 secs
3	60°C for 30 secs
4	72°C for 1 min
5	Repeat steps 2-4 16 times for a total of 17 cycles
6	72°C for 10 mins
7	8°C hold until purification step

2.4.7 *Purification of target libraries*

In this step, the amplified target DNA were purified using AMPure XP beads that have been allowed to come to room temperature for at least 30 minutes. 40 μL of each PCR reaction was transferred to a fresh 0.2 mL tube and the remaining volume was stored at -20°C for troubleshooting. The AMPure XP bead suspension was mixed well until the suspension appeared homogeneous. 100 μL of AMPure XP bead suspension and 40 μL of nuclease-free water was added to each 40 μL of PCR reaction and vortexed thoroughly. The samples were then incubated at room temperature for 5 minutes with shaking, then placed in the Agencourt SPRIPlate Super Magnet magnetic plate for 5 minutes to allow the solution to clear. The cleared solution was removed and discarded, taking care to avoid touching the beads. Keeping the tubes in the Agencourt SPRIPlate Super Magnet magnetic plate, 200 μL of 70% ethanol was added and any disturbed beads were allowed to settle for 30 seconds. The ethanol was then removed and an additional ethanol wash was performed. The tubes were air-dried with open lids at room temperature until the residual ethanol completely evaporated. The beads were resuspended in 40 μL of 10 mM Tris-HCl buffer (pH 8.0) and incubated at room temperature for 2 minutes to elute the DNA. The tubes were transferred to the Agencourt SPRIPlate Super Magnet magnetic plate for a final time and the solution was allowed to clear for 2 minutes. Finally, the cleared supernatant (approximately 40 μL) was transferred to a fresh tube and the beads were discarded.

2.4.8 *Validation of DNA enrichment*

Prior to sample pooling and sequencing, sample enrichment was validated using the Agilent High Sensitivity DNA Assay on the Agilent 2100 Bioanalyzer platform. Analysis was performed according to the manufacturer's instructions using 1 μL of each enriched sample library. Each amplicon of the libraries contains one target insert between 50-500 bp surrounded by 125 bp

of sequence motifs required for multiplexed sequencing using the Illumina platform. Therefore, microfluidic analysis showing amplicons ranging from 175-625 bp indicated correct library preparation. Smear analysis of amplicons in the range of 175-625 bp was used to determine the mean size of each library. DNA concentration was then quantified using the Qubit dsDNA High Sensitivity (HS) Assay Kit according to manufacturer's instructions.

2.4.9 *Library preparation for sequencing*

Each library to be sequenced was diluted in TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0) to a final concentration of 10 nM. The libraries were then pooled equally and the library mix was diluted to a concentration of 2 nM with TE buffer. The pooled libraries were re-quantified using the Qubit dsDNA HS Assay Kit (Section 2.4.1) and the concentration was adjusted to 2 nM. 10 μL of freshly prepared 0.2 N sodium hydroxide was then mixed with 10 μL of 2 nM pooled library DNA before being incubated at room temperature for 5 minutes to denature the DNA. The denatured DNA was subsequently diluted to 20 pM by the addition of 980 μL of pre-chilled Hybridization Buffer (HT1). Finally, 400 μL of 20 pM denatured DNA library was mixed with 600 μL of pre-chilled HT1, inverted several times to mix and the resulting 8 pM solution was kept on ice.

The internal sequencing control was prepared by mixing 2 μL of 10 nM PhiX virus library, 8 μL of 10 mM Tris-HCl, pH 8.5 containing 0.1% Tween 20 and 10 μL of freshly prepared 0.2 N sodium hydroxide. The solution was vortexed briefly and incubated at room temperature for 5 minutes to denature the DNA. The solution was then diluted to 8 pM as described above. Finally, 10 μL of 8 pM PhiX control was mixed with 990 μL of 8 pM denatured sample libraries and kept on ice until the samples were loaded onto the MiSeq or HiSeq cartridge.

2.4.10 *Next-generation sequencing*

Next-generation sequencing was performed on either the Illumina MiSeq or HiSeq 2500 platform (Illumina, San Diego, CA) according to manufacturer's instructions. In the case of the MiSeq platform, sequencing was performed using the reagents provided in the MiSeq v2 Reagent (300 cycles) Kit using 2 x 150 bp paired-end chemistry. In the case of the HiSeq 2500 platform, sequencing was performed using the reagents provided in the Rapid SBS (3 x 50 cycle) Kit in rapid run mode using 2 x 150 bp paired-end chemistry.

2.4.11 *Data analysis*

Sequencing reads were imported from FASTQ files and HaloPlex adaptors were trimmed using Cutadapt version 1.3 (Martin, 2011). Reads were aligned to the hg19 reference genome using the Burrows-Wheeler Alignment (BWA) (Li and Durbin, 2009) with default settings. Variants were then called using VarScan version 2.3 (Koboldt *et al.*, 2012) within our genomic regions of interest (coding exons plus 14bp of upstream and 6bp of downstream sequence) using a BED file generated using the UCSC Table Browser (Karolchik *et al.*, 2004). Variants were called according to a minimum of 30X coverage with five alternate reads, a Phred base quality of 20, and without strand bias. Copy number variants on the processed BAM files were called using an in-house read depth pipeline. Results were exported in VCF format for downstream analysis.

Variants were annotated based on Ensembl gene transcripts. Variants listed in the Ensembl (African, American or European population frequencies) (www.ensembl.org) or 1000 genomes database (www.browser.1000genomes.org/index.html) with a minor allele frequency of greater than 2% were filtered out. Prioritisation of the remaining variants was carried out based on the following criteria: selection of candidate genes based on patient's phenotype, predicted effect on gene function, conservation of amino acid position, and frequency of the variant in the publicly available databases. Selection of candidate genes based on patient's phenotype was carried out using Phenomizer (www.compbio.charite.de/phenomizer/), a resource based on the Human Phenotype Ontology, filtering by disease class or through manual literature searching. Details of other criteria used to prioritise non-synonymous variants can be found in Section 2.6.

2.5 WHOLE EXOME SEQUENCING (WES)

Whole exome sequencing (WES) was performed in generous collaboration with GOSgene at The Centre for Translational Genomics, UCL Institute of Child Health, London.

Whole exome sequencing was outsourced to BGI Genomics Hong Kong (www.genomics.cn). DNA was captured using the Agilent SureSelect v4 (51Mb) kit followed by sequencing using the Illumina HiSeq 2000 platform (Illumina, San Diego, CA). Three samples were mutliplexed per flowcell lane using 2 x 150bp paired-end chemistry to give an overall coverage of 100 times.

Processing of raw data files was carried out by Dr Chela James at GOSgene. Sequencing reads were processed using the Genome Analysis Toolkit (GATK) (McKenna *et al.*, 2010) recommended best practice of alignment to the reference human genome (human_g1k_v37) using the Burrows-Wheeler Aligner (Li and Durbin, 2009) with default settings. The alignment was then refined using Picard tools, GATK data processing tools, base quality score recalibration and indel

realignment. Single nucleotide polymorphism and insertion/deletion discovery and genotyping was performed using the GATK variant discovery pipeline including variant calling using Unified Genotyper followed by variant quality score recalibration (DePristo *et al.*, 2011). The resultant VCF files were filtered using Ingenuity Variant Analysis software (Ingenuity, Redwood City, CA) using differing parameters depending on the patient/family being investigated. Following this filtering the remaining variants were individually assessed to ensure they were not located within a segmental duplication, and finally BAM files were inspected using Integrative Genomics Viewer (IGV) software (Robinson *et al.*, 2011) to confirm the quality of the call. Non-synonymous variants were also prioritised and further investigation using the tools detailed in Section 2.6.

2.6 BIOLOGICAL INTERPRETATION USING BIOINFORMATICS TOOLS

When a potentially pathogenic variant was identified, sequence conservation was assessed by importing the protein sequence of the mutated gene in a range of species into Clustal Omega (www.ebi.ac.uk/Tools/msa/clustalo). Where possible, species were chosen that represented a large distance in evolutionary time such as *Escherichia coli* and/or *Saccharomyces cerevisiae*. Sequences were aligned using default settings.

In the case of amino acid substitutions, the effect on protein function was predicted using online tools such as SIFT (www.sift.bii.a-star.edu.sg/) and PolyPhen-2 (www.genetics.bwh.harvard.edu/pph2/). SIFT and PolyPhen-2 scores were determined for potentially pathogenic variants either through Ensembl (www.ensembl.org/) or direct input to the prediction tools.

In order to assess the frequency at which potentially pathogenic variants are present in the general population, multiple databases containing variant population statistics were searched including Ensembl (www.ensembl.org/), Exome Aggregation Consortium (www.exac.broadinstitute.org/) and 1000 Genomes (www.1000genomes.org/). If the minor allele of a variant was present at a frequency greater than 2%, especially in the ethnic sub-population of the affected patient, it was discarded.

2.7 CELL CULTURE

2.7.1 *Fibroblast cell culture conditions*

Fibroblasts were cultured in sterile 75 cm^2 flasks at 37°C in 5% CO_2 in 8 mL HAMS F-10 media supplemented with 10% fetal bovine serum (FBS), penicillin (100 U/ml) and streptomycin (100 μg/ml). Once cells were between 70 - 90% confluent the media was removed. Cells were

washed with 8 mL of Dulbecco's phosphate buffered saline (PBS) and detached using 3 mL 0.25% trypsin-EDTA for one minute. Trypsin was inactivated by addition of 13 mL of media and the resulting 16 mL of cell suspension was split equally between two 75 cm^2 flasks.

To harvest cells, the media was removed once cells were confluent and cells were trypsinised as above. Trypsin was inactivated by addition of 10 mL of media and then transferred to a 15 mL falcon tube. Cells were pelleted at 4500 x *g* for 5 minutes, the supernatant was removed and the cells were thoroughly washed and pelleted twice with 3 mL PBS. The remaining pellet was stored in 10 μL of PBS at -80°C to prevent degradation of proteins and metabolites before experimental use.

2.7.2 *Counting cells*

10 μL of cell suspension (Section 2.7.1) was mixed with 10 μL of 0.4% trypan blue in 0.81% sodium chloride and 0.06% potassium phosphate dibasic solution. 10 μL of mixture was loaded into dual-chambered slides and counted using a TC10 Automated Cell Counter. Cells were then diluted to the desired concentration in media, prior to harvesting or seeding (Section 2.7.1).

2.7.3 *Fibroblast cell storage*

Cells were harvested as above (Section 2.7.1) and the resulting pellet was resuspended in 900 μL of media. The cell suspension was added to 100 μL of 100% dimethyl sulphoxide in a CryoTube Vial and stored in a Mr. Frosty Freezing Container (Thermo Scientific) overnight. Vials were subsequently stored long-term at -196°C in liquid nitrogen.

2.7.4 *Protein assay*

Cellular protein concentration was determined using the Pierce BCA Protein Assay Kit according to manufacturer's instructions. BCA Working Reagent was prepared by mixing 50 parts of BCA Reagent A with 1 part of BCA Reagent B. All standards and samples were assembled in 96-well plates. 10 μL of each Pierce Bovine Serum Albumin Standard (Pre-Diluted Set) was added to the first 10 wells of the 96-well plate. 5 μL of each cell lysate was diluted with 5 μL of dH_2O in the following wells. 200 μL of BCA working reagent was then added to each standard and diluted cell lysate. The plate was mixed thoroughly on a plate shaker for 30 seconds before the plate was covered and incubated at 37°C for 30 minutes. The absorbance of each sample

was then measured at 555 nm using a Tecan Infinite 200 Microplate Reader (Tecan, Mannedorf, Switzerland).

The absorbance measurement of the blank standard was subtracted from all other standards and samples. The corrected absorbance of each standard was plotted against the concentration of each standard to yield a standard curve, from which the concentation of each lysate was determined.

2.8 cDNA ANALYSIS

2.8.1 *Isolation and purification of total RNA from fibroblasts*

Total RNA was purified from patient fibroblasts using the RNeasy Mini Kit. Fibroblasts were harvested as described in Section 2.7.1 and had been stored in DNase and RNase free falcon tubes at -80°C. Cells were resuspended in the residual 10 μL of PBS that the pellets were stored in and transferred to microcentrifuge tubes. Cells were then centrifuged for 5 minutes at 300 x g and the supernatant was removed. 600 μL of Buffer RLT was added to disrupt the cells and lysates were added to QIAshredder spin columns and centrifuged at 8000 x g for 2 minutes. 600 μL of 70% ethanol was added to the resulting homogenised lysate and mixed well. Samples were transferred to an RNeasy spin column, centrifuged for 15 seconds at 8000 x g and the flow-through was discarded. Columns were washed by addition of 700 μL of Buffer RW1, centrifuging for 15 seconds at 8000 x g and the flow-through was discarded. 500 μL of Buffer RPE was added, centrifuged for 2 minutes at 8000 x g and the flow-through was discarded. This latter centrifugation ensures that no ethanol is carried over during RNA elution as residual ethanol may interfere with downstream reactions. Spin columns were placed in clean microcentrifuge tubes and 50 μL of RNase-free water was added to the spin column membrane prior to centrifugation for 1 minute at 8000 x g to elute the RNA.

2.8.2 *Isolation and purification of total RNA from blood*

Total RNA was purified from whole blood using the Tempus Spin RNA Isolation Kit. 3 mL of blood was drawn directly into a Tempus Blood RNA Tube and immediately vigorously shaken for 10 seconds to ensure uniform contact with the Applied Biosystems Stabilizing Reagent. The mixture was then transferred to a 50 mL sterile falcon tube and 3 mL of PBS was added. The solution was then vortexed thoroughly for 30 seconds and centrifuged at 3,000 x g for 30 minutes at 4°C. After removing the supernatant, the pellet containing the RNA was resuspended in 400

μL of RNA purification resuspension solution before being stored on ice. Each sample was then purified using an RNA purification filter. The filters were primed initially by the addition of 100 μL of Wash Solution 1 followed by 400 μL of resuspended RNA. Samples were centrifuged at 16,000 x *g* for 30 seconds and the flowthrough was discarded. An additional 500 μL of Wash Solution 1 was then added, before centrifuging at 16,000 x *g* for 30 seconds and discarding the flowthrough. Samples were then washed with 500 μL of Wash Solution 2, centrifuged at 16,000 x *g* for 30 seconds and the flowthrough was discarded. This wash was then repeated. Samples were centrifuged again for 30 seconds at 16,000 x *g* to dry the membrane. To elute the RNA, 100 μL of Elution Solution was added and samples were incubated at 70°C for 2 minutes. Samples were centrifuged at 16,000 x *g* for 30 seconds and the resulting solution was re-applied to the filters before being centrifuged a second time for two minutes. Finally, 90 μL of RNA eluate was transferred to a clean microcentrifuge tube, taking care not to disturb any pelleted particulates from the column.

2.8.3 *cDNA synthesis*

cDNA synthesis was carried out using the SuperScript III First-Strand Synthesis System for RT-PCR. 8 μL of RNA, 1 μL of 50 μM oligo$[dT]_{20}$ and 1 μL of 10 mM dNTP mix were combined and incubated at 65°C for 5 minutes, then placed on ice for 1 minute. cDNA synthesis mix was made up containing 2 μL of 10X RT buffer, 4 μL of 25 mM $MgCl_2$, 2 μL of 0.1 M DTT, 1 μL of RNaseOUT (40 U/μL) and 1 μL of SuperScript II RT (200 U/μL). 10 μL of cDNA synthesis mix was added to each RNA/oligo$[DT]_{20}$ mixture and incubated for 50 minutes at 50°C. Reactions were then terminated at 85°C for 5 minutes and chilled on ice. 1 μL of RNase H was added and reactions were incubated for 20 minutes at 37°C. Resulting products containing RNA-free cDNA were stored at -20°C.

2.8.4 *PCR conditions*

The amplification of each whole cDNA sequence was carried out as described in (Table 2.8.1).

Table 2.8.1: Primers and conditions used to amplify and sequence whole cDNA. * 5% DMSO and 0.5 mM betaine were also added.

Gene	Sequence 5' → 3'	T_m (°C)	Polymerase	$MgCl_2$ (mM)	Product size (bp)
NDUFS1 F	CGAGGCCGCCATATTGAATA	58	Phusion	0	2537
NDUFS1 R	TCGTCACAAAGGTTATCAATGCT				
PROSC F	GGGGATGTGGAGAGCTGG	56	Taq	1.5*	846
PROSC R	CAGTATTCCCTGGCTCAGTG				

2.9 QUANTITATIVE POLYMERASE CHAIN REACTION (QRT-PCR)

Total RNA was purified from patient and control fibroblasts using the RNeasy Mini Kit as described in Section 2.8.1. The RNA concentration of each sample was determined using a NanoDrop 1000 (Thermo Scientific) and the RNA was stored at -80°C prior to cDNA synthesis. RNA samples were diluted to a final concentration of 500 ng RNA in 8 μL of RNase-free water before using the SuperScript III First-Strand Synthesis System to synthesise cDNA (Section 2.8.3).

Each qRT-PCR experiment was assembled in a MicroAmp Fast Optical 96-well Reaction Plate. As much as possible, reagents and plates were kept in the dark as the fluorescent-labelled probes are light-sensitive. Seven samples could be analysed in each assay; 4 μL of each patient's cDNA was aliquoted into each well of one row (12 wells). Three reaction master mixes were assembled consisting of 10 μL of 2X TaqMan Gene Expression Master Mix, 5 μL of RNase-free water and 1 μL of 20X TaqMan Gene Expression Assay per well. For each gene to be analysed a different TaqMan Gene Expression Assay was used: PROSC (Hs00200497_m1), β-actin (4333762T) and GAPDH (hs02758991_g1). For each cDNA sample, the expression of the three genes was analysed in four replicates. Thus, 16 μL of each master mix was added to four wells corresponding to each sample. Finally, 16 μL of each master mix was added to one well containing 4 μL of RNase-free water to act as a negative control. The plate was then sealed using adhesive film and centrifuged briefly to collect the contents of each well.

Each qRT-PCR reaction was cycled using the Applied Biosystems StepOne Real-Time PCR System using Standard Run Mode. Thermal cycling conditions were as follows: 50°C for 2 minutes, 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. The instrument was operated in Comparative C_T ($\Delta\Delta C_T$) mode using the same control sample in each assay and both β-actin and GAPDH as endogenous controls. After all assays were run, the threshold cycles (C_T) for each gene were averaged and all parameters were then re-analysed

using these standardised thresholds. Relative quantitation was carried out using the comparative C_T method ($2^{-\Delta\Delta CT}$ method) according to Schmittgen and Livak (2008).

2.10 WESTERN BLOTTING

Fibroblasts were cultured to confluence in 175 cm^2 flasks before being trypsinised and pelleted (Section 2.7.1). Cell pellets were washed three times with 3 mL of PBS and kept on ice until lysis. Cells were lysed through addition of 200 μL of lysis buffer [containing 40 μL of 5X RIPA buffer and 2 μL 100X protease inhibitor cocktail] and pipetted up and down to resuspend the pellet. Samples were incubated on ice for 10 minutes, centrifuged at 8600 x g for 5 minutes and the supernatant was collected. Protein concentration was determined using the Pierce BCA Protein Assay Kit (Section 2.7.4). Lysates were diluted in lysis buffer to yield 20 μg of protein in a final volume of 20 μL. To each diluted cell lysate, 5 μL of NuPAGE LDS Buffer and 2 μL of NuPAGE Sample Reducing Agent were added. Samples were mixed, heated at 90°C for 10 minutes and immediately returned to ice. Electrophoresis was carried out using a NuPAGE Novex 4-12% Bis-Tris Protein Gel. Gel wells were washed with 1X NuPAGE MOPS SDS Running Buffer to displace air bubbles and gel cassettes were assembled in an XCell SureLock Mini-Cell Electrophoresis System (Thermo Fisher). The Upper Buffer Chamber was filled with 200 mL of 1X NuPAGE MOPS SDS Running Buffer and 0.5 mL of NuPAGE Antioxidant. 10 μL of Rainbow Marker was loaded into the first well and samples were loaded into each of the following wells. The Lower Buffer Chamber was filled with 600 mL of 1X NuPAGE MOPS SDS Running Buffer and electrophoresis was carried out at a constant voltage of 180 V for 1 hour. The gel was removed from the cassette and the bottom foot and top fringes of the gel were removed using a clean scalpel. Gel transfer was carried out using an iBlot Dry Blotting System (Thermo Fisher) using the P3 setting for 13 minutes according to the manufacturer's instructions. The membrane was transferred to 50 mL of 5% skimmed milk powder in 1 X TBST [containing 50 mmol/L Tris-HCl, 300 mmol/L NaCl, 0.1% Tween 20 (pH 7.6)] and incubated for 1 hour at room temperature with gentle agitation to block the membrane. The milk was then discarded and replaced with 1.5 mL of PROSC primary antibody (1:100 in 5% TBST milk; HPA023646). The membrane was then incubated overnight at 4°C with gentle agitation. The following morning, the primary antibody solution was removed and the membrane was washed three times with 10 mL of 1X TBST. Each wash consisted of a 10 minute incubation with gentle agitation before replacing with clean 1X TBST. Once the washing of the membrane was complete, the TBST was replaced with 5 mL of donkey anti-rabbit IgG-HRP (sc-2317) diluted 1:5000 with 5% TBST milk. The membrane was then incubated for 1 hour at room temperature with gentle agitation. The

secondary antibody solution was then removed and the membrane was washed again three times with 10 mL of 1X TBST. The membrane was developed using the Novex ECL Chemiluminescent Substrate Reagent Kit. 500 μL of Reagent A and 500 μL of Reagent B were mixed and the membrane was subsequently washed in this solution for 5 minutes, taking care not to let the membrane dry out. The membrane was then imaged using a ChemiDoc MP System (Bio-Rad) coupled to ImageLab 4.1 software.

The membrane was again washed three times in 1X TBST, before being incubated with 10 mL of mouse beta-actin primary antibody (A1978) diluted 1:40,000 in 1X TBST for 1 hour at room temperature with gentle agitation. The primary antibody solution was then removed and the membrane was washed again three times in 1X TBST. The membrane was then incubated with 6 mL of goat anti-mouse IgG-HRP (P0447) diluted 1:3000 with 1X TBST for 1 hour at room temperature with gentle agitation. Finally, the membrane was washed three times in 1X TBST and developed and imaged following the same protocol as for the PROSC antibody. The membrane was then stored at 4°C in 1X TBST in case further imaging was required.

2.11 Measurement of B_6 vitamers and 4-pyridoxic acid using ultra-performance liquid chromatography tandem mass-spectrometry (UPLC-MS/MS)

The method described in this thesis for the quantitation of the B_6 vitamers in fibroblasts is based on those previously published from the department for the analysis of plasma (Footitt *et al.*, 2013).

2.11.1 *Fibroblast sample preparation*

Prior to mass spectrometry analysis fibroblast pellets were removed from storage at -80°C, resuspended in 50 μL of dH_2O and vortexed. Cells were then subjected to five freeze-thaw cycles in a methanol-dry ice mix and a 37°C water bath in order to lyse them. The lysates were then pelleted at 4500 x *g* for 5 minutes at 4°C and the resulting supernatant collected. Protein concentration was determined using the Pierce BCA Protein Assay Kit (Section 2.7.4). Proteins were precipitated by mixing 10 μL of cell lysate supernatant with 110 μL of master mix [containing 40 μL of dH_2O, 60 μL of 0.3N trichloroacetic acid (TCA) and 2 μL of each deuterated internal standard (concentrations detailed in Section 2.11.3)]. The sample was vortexed thoroughly for 30 seconds, left on ice in the dark for 60 minutes and centrifuged to pellet the precipitated protein.

The resulting supernatant, containing the B_6 vitamers, was transferred to a HPLC vial and placed in an autosampler, protected from light and kept at 4°C until sample injection.

2.11.2 *CSF sample preparation*

50 μL of CSF was added to 70 μL of master mix [containing 60 μL of 0.3N TCA and 2 μL of each deuterated internal standard (concentrations detailed in Section 2.11.3)]. Samples were vortexed thoroughly for 30 seconds, left on ice in the dark for 60 minutes and centrifuged to pellet the precipitated protein. The resulting supernatant, containing the B_6 vitamers, was transferred to a HPLC vial and placed in an autosampler, protected from light and kept at 4°C until sample injection.

2.11.3 *Quantification of B_6 vitamers and 4-pyridoxic acid*

Stock solutions of all the B_6 vitamers, PA and deuterated internal standards were made up using dH_2O and stored at -80°C to prevent degredation. During laboratory handling, all standards were kept on ice and protected from light. Quantification of the analytes in fibroblasts was achieved by spiking 2 μL of a known concentration [d_2-PLP; 100 nM, d_2-PA, d_3-PM, and d_2-PN; 10 nM, d_3-PL; 50 nM] of each internal standard into each sample. Calibration curves were constructed using different concentrations of the undeuterated vitamers (except PNP which is not commercially available) and the known concentrations of deuterated standards. The amounts of each endogenous analyte was then determined by ratioing the analyte signal area to that of its corresponding standard Table 2.11.1.

Table 2.11.1: B_6 vitamers and their corresponding internal standards used for quantitation.

Analyte	Deuterated standard	Deuterated standard concentration (nmol/L)	Analyte concentration expressed as:
PL	d_3-PL	50	nmol/L
PM	d_3-PM	10	nmol/L
PN	d_2-PN	10	nmol/L
PA	d_2-PA	10	nmol/L
PLP	d_2-PLP	100	nmol/L
PMP	d_2-PLP	100	nmol/L
PNP	d_2-PLP	100	"concentration units"

Stable isotopes of PMP and PNP were not commercially available at the time this assay was set-up, therefore these two vitamers were quantitated by ratioing to d_2-PLP. PNP was also not commercially available and thus calibration curves could not be generated. Although PNP and PLP should behave similarly due to their chemical similarity and difference of only one functional group, we cannot assume that their response within the mass spectrometer would be identical. Therefore, concentrations of PNP are expressed as "concentration units" instead of nmol/L. All results are corrected for cellular protein concentration using the Pierce BCA Protein Assay Kit according to manufacturer's instructions (Section 2.7.4).

2.11.4 *Identification of B_6 vitamers and 4-pyridoxic acid*

LC-MS/MS was performed using an Acquity Ultra Performance LC system linked to a triple quadrupole Xevo TQ-S instrument (Waters, Manchester, UK). Samples were separated on an Acquity UPLC HSS T3 column (1.8 μm x 2.1 mm x 50 mm) fitted with a HSS T3 VanGuard guard column (Waters) using the gradient described in Table 2.11.2. The flow rate was maintained at 0.4 mL/min.

Table 2.11.2: Mobile phase gradient profile for separation of B_6 vitamers in fibroblasts. HFBA; heptafluorobutyric acid.

Time (minutes)	A (3.7% acetic acid + 0.01% HFBA)	B (100% methanol)	Curve
0.00	97.5	2.5	0
0.40	97.5	2.5	6
3.75	50.0	50.0	6
3.76	0.1	99.9	11
4.26	97.5	2.5	1
5.00	97.5	2.5	1

Transitions for each analyte were determined by direct infusion into the mass spectrometer and subsequently detected using multiple reaction monitoring (MRM) in positive ion mode. All B_6 vitamers and their excretion product 4-pyridoxic acid (PA) could be uniquely identified based on their retention time and the m/z ratios of their corresponding parent and daughter ions Table 2.11.3. From these parent and daughter ions for each analyte we were able to then calculate their respective fragmentation patterns. These suggested losses of H_2O for PL; NH_3 and H_2O for PM; two molecules of H_2O for PN and PA; HPO_3 and H_2O for PLP; and H_2O, HPO_3 and NH_3 for PMP. As PNP was not available for direct infusion into the mass spectrometer to allow experimental determination of any transitions, an additional transition was added to the MRM

using a predicted fragmentation pattern based on that of PLP and PN (i.e. a loss of phosphate and H_2O). A 15 μL volume of spiked sample mixture was injected on to the mass spectrometer every 5 minutes and the data was acquired using MassLynx software (Waters, Manchester, UK). Once the method parameters were fully investigated and validated, data processing was subsequently completed using QuanLynx to reduce variability in results due to human error.

Table 2.11.3: Molecular weights, mass transitions, cone voltages, collision energies and retention times of the different B_6 vitamers and internal standards. All analytes were detected using a Xevo TQ-S mass spectrometer in positive ion mode using MRM.

Analyte	Molecular weight (g/mol)	Parent ion (m/z)	Daughter ion (m/z)	Cone voltage (V)	Collision energy (V)	Retention time (mins)
Pyridoxal	167.16	168.10	150.05	21	12	0.93
d_3-pyridoxal	170.16	171.10	153.05	21	12	0.93
Pyridoxamine	168.19	169.12	134.04	22	20	0.95
d_3-pyridoxamine	171.19	172.12	137.04	22	20	0.95
Pyridoxine	169.18	170.09	134.04	27	19	1.01
d_2-pyridoxine	171.18	172.09	136.04	27	19	1.01
Pyridoxic acid	183.16	184.06	147.99	18	18	0.85
d_2-pyridoxic acid	185.16	186.06	149.99	18	18	0.85
Pyridoxal phosphate	247.14	248.00	150.01	27	16	0.81
d_2-pyridoxal phosphate	249.14	250.00	152.01	27	16	0.81
Pyridoxamine phosphate	248.17	249.04	134.05	27	22	0.69
Pyridoxine phosphate	249.16	250.04	134.04	27	16	1.42

2.12 MEASUREMENT OF PLP-CYSTEINE CONJUGATES USING UPLC-MS/MS

PLP can react with cysteine to form a thiazolidine complex (Ponticelli *et al.*, 1983; Terzuoli *et al.*, 1998) with a predicted molecular weight of 350 g/mol.

2.12.1 *Sample preparation*

Prior to mass spectrometry analysis fibroblast pellets were removed from -80°C storage, re-suspended in 60 μL of dH_2O and vortexed. Mixtures were freeze-thawed, pelleted and the

supernatant was collected as above (Section 2.11.1). Protein concentration was determined using the Pierce BCA Protein Assay Kit (Section 2.7.4). Proteins were precipitated by mixing 45 μL of cell lysate supernatant with 90 μL of ice-cold 100% methanol. The samples were incubated on ice in the dark for ten minutes and centrifuged to pellet the precipitated protein. The resulting supernatant was transferred to a HPLC vial and placed in an autosampler, protected from light and kept at 4°C until sample injection.

2.12.2 *Identification of PLP-cysteine conjugate*

LC-MS/MS was performed using a Waters Alliance 2795 LC system linked to a triple quadrupole Micro Quattro instrument (MicroMass, Waters, UK). Samples were separated on a HS F5 column (3 μm x 10 cm x 2.1 mm) fitted with a HS F5 guard column (Supelco, Sigma Aldrich) using the gradient described in Table 2.12.1. The flow rate was maintained at 0.2 mL/min.

Table 2.12.1: Mobile phase gradient profile for the detection of PLP-cysteine in fibroblasts.

Time (minutes)	A (100% methanol)	B (3.7% acetic acid)	Curve
0.00	2.5	97.5	1
2.00	2.5	97.5	6
10.00	50.0	50.0	6
12.00	2.5	97.5	11
14.00	2.5	97.5	11

The mass spectrometer was operated in positive ion MRM mode and PLP-cysteine was identified using a parent ion of 350.98 m/z, a daughter ion of 219.03 m/z, a cone voltage of 27 V and a collision voltage of 19 V.

2.13 QUANTIFICATION OF PNPO ENZYME ACTIVITY USING UPLC-MS/MS

2.13.1 *Sample preparation*

The cell pellets were removed from -80°C storage, resuspended in 50 μL of dH_2O and vortexed. Mixtures were then subjected to three freeze-thaw cycles in a methanol-dry ice mix and a 37°C water bath to gently lyse cells. The lysates were then pelleted at 4500 x *g* for 5 minutes at 4°C. Protein concentration was determined using the Pierce BCA Protein Assay Kit (Section 2.7.4) and the supernatants were diluted with dH_2O to an equal starting concentration. 10 μL

of diluted supernatant was then transferred into five labelled eppendorfs corresponding to the five time-points used to measure the enzyme activity [0, 0.5, 1, 2 and 3 hours]. 110 μL of buffer [containing 60 μL of 40 mM potassium phosphate buffer (pH 7.6), 12 μL of 3mM ATP, 12 μL of 3 mM $MgCl_2$, 12 μL of 15 μM FMN, 4 μL of 3 μM d_2-PN and 10 μL dH_2O] was added and the reaction tubes were covered in foil to protect the vitamers from light.

The 0 hour reaction was stopped by adding 120 μL of 0.3N TCA [containing 4 μL of 3 μM d_3-PLP and 4 μL of 3 μM d_3-PL] and vortexing thoroughly for 30 seconds. The remaining tubes were incubated with shaking at 37°C for the appropriate times, and subsequently stopped with 120 μL of 0.3N TCA containing the deuterated vitamers. After stopping each time-point reaction, each tube was left on ice in the dark for 45 minutes to allow any Schiff bonds to be broken and protein to precipitate, before centrifuging at 17,900 x g for 10 minutes to pellet the proteins present in the reaction mixture. The resulting supernatant, which contains the B_6 vitamers, was transferred to a HPLC vial and placed in an autosampler, protected from light and kept at 4°C until sample injection.

LC-MS/MS quantitation of the relevant vitamers was carried out using the method described previously in Section 2.11.4, with the addition of one transition. This corresponds to the predicted fragmentation of d_2-pyridoxine phosphate (i.e. a loss of HPO_3 and H_2O) to allow the measurement of the conversion of d_2-PN to d_2-PNP and d_2-PLP by pyridoxal kinase (PK) and pyridox(am)ine phosphate oxidase (PNPO), respectively (Table 2.13.1).

Calibration curves were constructed using different concentrations of the analytes d_2-PLP and d_2-PN and known concentrations of other deuterated standards (Table 2.13.2). d_2-PNP is not commercially available, thus this analyte was quantitated using the standard curve for d_2-PLP and results were expressed as "concentration units".

Table 2.13.1: Molecular weights, mass transitions, cone voltages, collision energies and retention times of the analytes quantified in the PNPO/PK enzyme assay. All analytes were detected using a Xevo TQ-S mass spectrometer in positive ion mode using MRM.

Analyte	Molecular weight (g/mol)	Parent ion (m/z)	Daughter ion (m/z)	Cone voltage (V)	Collision energy (V)	Retention time (mins)
Pyridoxal	167.16	168.1	150.05	21	12	0.93
Pyridoxamine	168.19	169.12	134.04	22	20	0.95
Pyridoxine	169.18	170.09	134.04	27	19	1.01
d_2-pyridoxine	171.18	172.09	136.04	27	19	1.01
Pyridoxic acid	183.16	184.06	147.99	18	18	0.85
Pyridoxal phosphate	247.14	248.00	150.01	27	16	0.81
d_2-pyridoxal phosphate	249.14	250.00	152.01	27	16	0.81
Pyridoxamine phosphate	248.17	249.04	134.05	27	22	0.69
Pyridoxine phosphate	249.16	250.04	134.04	27	16	1.42
d_2-pyridoxine phosphate	251.16	252.04	136.04	27	16	1.42

Table 2.13.2: Internal standards used to quantitate analytes and construct calibration curves for the PK/PNPO enzyme assay.

Analyte	Deuterated standard	Deuterated standard concentration (nmol/L)	Analyte concentration expressed as:
d_2-PN	d_3-PL	100	nmol/L
d_2-PNP	d_3-PLP	100	"concentration units"
d_2PLP	d_2-PLP	100	nmol/L

2.14 TINCTORAL STAINING, IMMUNOFLUORESCENCE AND ELECTRON MICROSCOPY OF PATIENT FIBROBLASTS

2.14.1 *Tinctoral staining and immunohistochemistry*

2.14.1.1 *Cell immobilisation*

Once confluent, fibroblasts were detached from 75 cm^2 flasks (as described in Section 2.7.1) and resuspended in 2 mL of Ham's F10 medium. Cells were immobilised as a monolayer onto Superfrost Plus Micro Slides by aliquoting 300 μL of suspension into six sample chambers of a Thermo Scientific Cytospin 4 and spinning at 450 rpm for 10 minutes. Slides were allowed to air-dry for 5 minutes prior to fixation and staining.

2.14.1.2 *Haematoxylin and Eosin to examine cell morphology*

Slides were fixed in 95% ethanol for 5 minutes and rinsed in running water. Slides were stained with Harris haematoxylin for 30 seconds under gentle agitation. Haematoxylin staining was differentiated briefly in a solution of 1% hydrochloric acid in 70% ethanol and nuclear staining intensity was checked macroscopically. Slides were run under warm tap water to convert nuclear colouration from red/purple to blue and counterstained in 1% eosin for 10 seconds.

2.14.1.3 *Oil Red O to stain neutral lipids*

Oil Red O working solution was made by dissolving 1 g Oil Red O in 5 mL acetone and mixing with 100 mL 70% ethanol. The solution was allowed to settle for 24 hours before filtration and use. Slides were stained in Oil Red O working solution for one hour then rinsed in 70% ethanol followed by water. Nuclei were stained in Mayers haematoxylin for one minute and washed in running water until nuclei appeared blue.

2.14.1.4 *Sudan Black to stain phospholipids*

Sudan Black working solution was made by dissolving 1 g Sudan Black in 100 mL propylene glycol in a boiling water bath, before cooling and filtering. Slides were fixed in 4% formal calcium [100 mL of 40% w/w formaldehyde, 20 g of calcium acetate and 900 mL of dH_2O] and stained in Sudan Black working solution for two hours. Sudan Black was flooded off with 85% propylene glycol and staining was differentiated by washing two further times with 85% propylene glycol. Slides were subsequently washed well in 50% propylene glycol followed by running water. Slides were counterstained with Nuclear Fast Red and washed again in water.

2.14.1.5 *Luxol Fast Blue to stain myelin/sphingomyelin*

Slides were fixed in 95% ethanol for 5 minutes. Staining was carried out using the Luxol Fast Blue Stain Kit with some modifications. 0.1% Luxol Fast Blue in Acidified Methanol was filtered and placed in a thick glass staining trough at 60°C. Slides were incubated in the pre-heated solution for 2 hours at 60°C, then washed in 70% Denatured Ethanol and rinsed in tap water. Subsequently, slides were differentiated in 0.05% Lithium Carbonate Solution, briefly rinsed in 95% alcohol and washed in tap water. Slides were stained in Cresyl Violet Solution for 10-12 minutes, washed in tap water and differentiated in Cresyl Violet Differentiator for 4 seconds with gentle agitation and washed again in water.

2.14.1.6 *Lysosome-associated membrane protein 2 (LAMP2)*

Immunohistochemical staining was performed using a Leica Bond-Max autostainer (Leica Microsystems, Milton Keynes, UK). Firstly, slide details were entered onto the Bond-Max computer system, each slide was labelled with unique barcoded labels and reagents were uploaded onto the machine. Slides were placed in a Bond-Max slide tray, each slide was covered with a Bond-Max slide cover and inserted into the autostainer. LAMP-2 [H4B4] primary antibody was diluted 1:100 in Bond primary antibody diluent. Antibody detection was carried out with the Bond Polymer Refine Detection using Protocol F according to the manufacturer's protocol. 150 μL of each reagent was dispensed for all steps of the protocol. Slides were washed in alcohol and washed in Bond Wash Solution at room temperature for 5 minutes. Heat-induced epitope retrieval using Bond Epitope Retrieval Solution 2 (pH 7.0) was carried out for 20 minutes. Slides were washed in Bond Wash Solution at 35°C for 3 minutes after antigen retrieval. Slides were then blocked with 3-4% hydrogen peroxide at room temperature for 10 minutes and then incubated with the LAMP-2 primary antibody at room temperature for 15 minutes. Slides were washed in Bond Wash Solution and incubated with Post Primary Solution at room temperature for 8 minutes. Slides were washed in Bond Wash Solution and incubated in Mixed DAB Refine Solution at room temperature for 10 minutes. Slides were washed in dH_2O prior to counterstaining with Mayer's haematoxylin for 5 minutes at room temperature.

2.14.1.7 *Dehydration, clearing and mounting*

After staining, slides were dehydrated and cleared by sequential washing through a series of alcohol (70%, 90%, 100% and 100%) followed by two washes in xylene. Finally, slides were coverslipped using a Leica Multistainer and Coverslipper (Leica Microsystems).

2.14.1.8 *Visualisation and image capture*

Cells were visualised using a Nikon Optiphot microscope and images were captured using a Leica DMD108 Digital Microimaging Device.

2.14.2 *Electron microscopy*

Sample processing was carried out by Elizabeth Latimer-Bowman and electron microscopy imaging was carried out by Glenn Anderson (Histopathology Department, GOSH, London).

One confluent 75 cm^2 flask of fibroblasts per patient was pelleted as described in Section 2.7.1. Pellets were then fixed in 5 mL of 2.5% glutaraldehyde buffered with 100 mM cacodylate buffer (pH 7.2) for 24 hours at room temperature. Samples were placed on a sample rotator for 5 minutes, the buffered glutaraldehyde was removed and replaced with 5 mL of fresh 100 mM cacodylate buffer (pH 7.2). Samples were placed on a sample rotator for 5 minutes, the buffer solution was removed and replaced with 5 mL of 1% osmium tetroxide. Samples were incubated on a sample rotator for 1 hour at room temperature to allow for the secondary fixation of lipids with the samples. The osmium solution was then removed, discarded and washed twice with 5 mL of fresh 100 mM cacodylate buffer (pH 7.2) for 5 minutes. Once the buffer solution was removed, samples were dehydrated through graded alcohol washes (70%, 90% and two washes of 100%), incubating the samples for 10 minutes on a sample rotator at each step. The final 100% alcohol wash was removed, replaced with reagent grade propylene oxide and incubated for 10 minutes on a sample rotator. This was subsequently replaced with fresh propylene oxide for a second 10 minute wash. The propylene oxide was removed and replaced with 1 mL of fresh propylene oxide and 1 mL of freshly-made resin [12 g of Agar 100 Resin, 8 g of dodecenyl succinic anhydride, 5 g of methyl nadic anhydride and 0.38 g of benzyldimethylamine]. Samples were then incubated for 1 hour at room temperature on a sample rotator to allow resin infiltration. The solution was removed and 1 mL of fresh resin was added. Any residual propylene oxide was allowed to evaporate before the samples were again incubated for 1 hour on a sample rotator. Finally, the solution was replaced with fresh resin and incubated overnight at room temperature on a sample rotator. Samples were then transferred to embedding moulds and the resin was polymerised overnight at 60°C.

Ultrathin 90 nm sections were cut with a diamond knife on a Leica Ultracut UCT Ultramicrotome, placed on copper grids and stained with alcoholic uranyl acetate for 5 minutes. Sections were then rinsed in 50% ethanol followed by dH_2O and allowed to dry for 5 minutes. Finally,

sections were stained with Reynold's lead citrate for 15 minutes before being rinsed in dH_2O. Examination was carried out using a JEOL 1400 transmission electron microscope.

2.14.3 *Nucleoporin p62 (p62) immunofluorescence*

Immunofluorescence was carried out using the method of Yasin *et al.* (2013). Sterile glass coverslips were covered with 1 N hydrochloric acid and shaken for 1 hour at room temperature. Coverslips were removed from the acid, laid on Whatman filter paper in a glass dish and autoclaved for 20 minutes at 120°C. Treated coverslips were transferred into wells of a 24-well plate using sterile forceps. 500 μL of 5 μg/mL poly-D-lysine was added to each well and the plate was incubated at 37°C for 1 hour. Coverslips were washed in PBS three times for 5 minutes and the 24-well plate was left to dry at room temperature for 1 hour.

Fibroblasts were trypsinised, pelleted and resuspended in media as previously (Section 2.7.1). Fibroblasts were seeded on the coverslips at a density of 5 x 10^4 cells per coverslip and incubated at 37°C for 6 hours to allow the cells to adhere. 500 μL of media was added to each well and incubated overnight at 37°C to culture the cells to confluence. The following morning, the media was aspirated and 400 μL of 4% paraformaldehyde was added for 15 minutes at room temperature to fix the cells. 300 μL of paraformaldehyde was removed, 400 μL of PBS was added and the plate was stored at 4°C prior to staining.

PBS was removed from the wells and 400 μL of blocking solution [containing 0.1% Triton X-100, 10% Sheep serum in PBS] was added. The plate was incubated at room temperature for 1 hour with shaking and the blocking solution was then removed. Autophagosomes were stained by incubating cells with 200 μL of p62 primary antibody (1:50) at 4°C overnight. After removing the primary antibody, cells were washed twice with 500 μL of PBS for 10 minutes. PBS was removed and cells were then incubated in 200 μL of Alexa Fluor 610 secondary antibody (1:240) in the dark for 1 hour at room temperature. The secondary antibody was removed and cells were again washed with 500 μL of PBS for 10 minutes. Slides were washed twice in PBS and mounted in Vectashield aqueous mounting medium containing 4',6-diamidino-2-phenylindole/DAPI. Fibroblasts were visualised and images captured using a Leica DMLB fluorescent microscope.

CellProfiler 2.1.1 (Carpenter *et al.*, 2006) was used to analyse the immunofluorescence images obtained. CellProfiler uses analysis pipelines to complete counting tasks which contain configurable modules that can be optimised to suit a specific task. Representative images with different cell densities and staining intensities were selected to determine the most accurate settings for the analysis pipeline. Cells were counted by quantitation of nuclear DAPI staining using minimum and maximum diameter settings of 70 and 300 pixels, respectively. The threshold correction

factor was set to 1 using the Otsu Global thresholding method in two-class mode. Staining of p62 punctae was also counted using a diameter range of 1 to 30 pixels and a threshold correction factor of 2.2. The Otsu Global thresholding method in three-class mode was used with the middle intensity assigned to the background to avoid counting background as cells. All other pipeline settings were kept as default.

2.15 Measurement of glutamate-γ-semialdehyde using UPLC-MS/MS

Glutamate-γ-semialdehyde is structurally similar to α-aminoadipic semialdehyde which is routinely measured in our laboratory by HPLC-MS/MS of FMOC derivatives of each molecule (Mills *et al.*, 2006).

2.15.1 *Fibroblast sample preparation*

Prior to mass spectrometry analysis fibroblast pellets were lysed in 50 μL of dH_2O as previously (Section 2.11.1). 10 μL of cell lysate was then mixed in a HPLC vial with 10 μL of 0.1 mM d_3-aminoadipic acid, 40 μL of dH_2O, 125 μL of 0.1 M borate buffer (pH 10.4) and 125 μL of 6 mM fluorenylmethyloxycarbonyl chloride (FMOC-Cl) in acetone. Care was taken to add the FMOC-Cl to the vials last and to immediately cap and vortex the samples as the FMOC-Cl is water sensitive. Vials were placed in an autosampler, protected from light and kept at 4°C until sample injection.

2.15.2 *Identification and semi-quantification of glutamate-γ-semialdehyde*

LC-MS/MS was performed using the method used in our laboratory to measure α-aminoadipic semialdehyde for the diagnosis of pyridoxine-dependent epilepsy (Mills *et al.*, 2006). The Micro Quattro instrument was operated in positive ion MRM mode and glutamate-γ-semialdehyde (GSA) was identified using a parent ion of 352.30 m/z, a daughter ion of 130.10 m/z, a cone voltage of 10 V and a collision voltage of 8 V. A 10 μL volume of cell lysate mixture was injected on to the mass spectrometer every 5 minutes and the data was acquired using MassLynx software (Waters). Data processing was completed using QuanLynx (Waters). Analysis of the amount of GSA in fibroblasts was achieved by spiking 10 μL of 0.1 mM d_3-aminoadipic acid into each sample (Section 2.15.1). The amount of GSA was determined by ratioing the signal area to that of the ^{15}N-aminoadipic acid. GSA is not commercially available, therefore calibration curves

could not be generated. Thus, concentrations of GSA are expressed as "concentration units". All results are corrected for cellular protein concentration using the Pierce BCA Protein Assay Kit according to manufacturer's instructions.

2.16 MOLECULAR CLONING OF CYSTEINE CONJUGATE-β LYASE (CCBL1)

2.16.1 *cDNA clone information*

A full length cDNA clone of *CCBL1* in the vector pME18SFL3 (Accession No. AK314427) was purchased from the Biological Resource Center, National Institute of Technology and Evaluation, Japan. Upon arrival, the clone was resuspended in 200 μL 0.1X Tris-EDTA buffer prior to transformation into XL-1 Blue cells (Section 2.16.3.1).

2.16.2 *Molecular biology media*

2.16.2.1 *Solid LB medium*

Solid LB medium was made by mixing 8 g of Miller LB Broth and 6 g of bacteriological agar in 400 mL of dH_2O before autoclaving for 20 minutes at 120°C. The medium was allowed to cool to 55°C and ampicillin or kanamycin antibiotic (stock concentration 50 mg/mL) was then added to a final concentration of 50 μg/mL. 40 mL aliquots were then poured into sterile petri dishes and left to set for 30 minutes at room temperature. If blue-white screening was required then 32 μL of 50 mg/mL 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) in 2% N, N-dimethyl formamide was evenly spread on each plate and incubated at 37°C for 30 minutes prior to bacterial plating.

2.16.2.2 *Liquid LB broth*

Liquid LB broth was made by mixing 3 g of Miller LB Broth in 150 mL of dH_2O before autoclaving for 20 minutes at 120°C. The broth was allowed to cool to room temperature and ampicillin or kanamycin antibiotic (stock concentration 50 mg/mL) was then added to a final concentration of 50 μg/mL.

2.16.3 *Bacterial transformation of E. coli cells*

2.16.3.1 *XL-1 Blue Cells*

XL-1 Blue competent cells and β-mercaptoethanol were thawed on ice. 40 μL of cells was added to 0.68 μL of β-mercaptoethanol in a pre-chilled 1.5 mL eppendorf, then left on ice for 10 minutes, swirling every two minutes. 1 μL of resuspended plasmid was added to the cells and gently mixed prior to incubation on ice for 30 minutes. Cells were then heat shocked at 42°C for 45 seconds in a water bath, then placed on ice for two minutes. 500 μL of SOC Medium that had been pre-warmed to 37°C was added and the mixture incubated at 37°C for one hour with shaking. Cells were streaked on LB agar plates containing 50 μg/mL of ampicillin at different densities (10, 50, 100 and 200 μL). Plates were allowed to dry for five minutes and incubated at 37°C overnight, upside down to prevent condensation from falling onto the agar.

2.16.3.2 *TOP10 Chemically Competent Cells*

One vial containing 50 μL of TOP10 chemically competent cells per transformation were thawed on ice. 2 μL of TOPO cloning reaction was then added to the cells, gently mixed and incubated on ice for 30 minutes. Cells were then heat shocked at 42°C for 30 seconds, then placed on ice. 500 μL of SOC Medium that had been pre-warmed to 37°C was added and the mixture incubated at 37°C for one hour with shaking (200 rpm). LB agar plates containing 50 μg/mL of kanamycin and coated with X-Gal for blue-white screening, were pre-warmed at 37°C for 30 minutes. Cells were then streaked at different densities (10, 30, 80 and 150 μL). Plates were allowed to dry for five minutes and incubated at 37°C overnight, upside down to prevent condensation from falling onto the agar.

2.16.4 *Liquid culture of E. coli transformants and preparation of glycerol stocks*

Single isolated *E. coli* colonies were picked using sterile inoculation loops following overnight growth on agar plates and incubated in 5 mL of LB liquid media (Section 2.16.2.2) containing 50 μg/mL of ampicillin or kanamycin at 37°C overnight with shaking. Where blue-white screening was used, only white colonies were picked. Glycerol stocks were prepared by mixing 850 μL of each liquid culture with 150 μL 100% sterile glycerol. Stocks were stored at -80°C for future use.

2.16.5 *Preparation of plasmid DNA*

The remaining 4.15 mL of liquid culture was centrifuged at 16,000 x *g* for one minute to pellet the *E. coli* cells. The supernatant was removed and the pellet resuspended in the 20 μL that remained. Plasmid DNA was then extracted using the QIAprep Spin Miniprep Kit according to manufacturer's instructions. 250 μL of Buffer P1 containing 100 μg/ml RNase A was used to resuspend the pellet until no cell clumps were visible. Cells were then lysed by addition of 250 μL of Buffer P2 followed by gentle inversion of the mixture 4-6 times. 350 μL of Buffer N3 was then added and each tube was immediately inverted 4-6 times to avoid localised precipitation. Samples were then centrifuged at 16,000 x *g* for 10 minutes to form a compact white pellet. Supernatants were applied to QIAprep spin columns by pipetting and the columns were centrifuged at 16,000 x *g* for 60 seconds. The flow-through was discarded and the columns were washed by the addition of 500 μL of Buffer PB to remove trace nuclease activity. Samples were centrifuged at 16,000 x *g* for 60 seconds and the flow-through was again discarded. 750 μL of Buffer PE was added, samples were centrifuged at 16,000 x *g* for 60 seconds and the flow-through was discarded. Samples were then centrifuged at 16,000 x *g* for an additional 60 seconds to remove residual buffer that contains ethanol and thus may inhibit subsequent enzymatic reactions. Finally, the QIAprep columns were placed in clean 1.5 mL microcentrifuge tubes and the DNA was eluted by adding 30 μL of Buffer EB to the centre of the membrane of each QIAprep spin column. This was allowed to stand for 60 seconds prior to centrifuging at 16,000 x *g* for 60 seconds. Plasmidic DNA concentration was subsequently determined using a NanoDrop 1000 (Thermo Scientific) and samples were stored at -20°C.

2.16.6 *Sequencing plasmid DNA*

Plasmid DNA was diluted to a final concentration of 100 ng/μL using dH_2O. Sequencing was then carried out using Sanger sequencing as previously described (Section 2.3.3.2) using 3 μL of plasmid DNA per sequencing reaction. Primers used to confirm the sequence of *CCBL1* clones are detailed in Table 2.16.1.

Table 2.16.1: Details of "walking" primers used to verify the sequence of the CCBL1 clones. These primers are designed to "walk" across the clone sequence by overlapping the PCR products to enable sequencing of the full clone sequence to ensure that the sequence of the clone is as expected. F, forward; R, reverse.

Name	Sequence 5' → 3'
1 F	GGATCGACTACAACCCCTGG
2 F	TGACGAAGATCCTGGCAAGT
3 F	GGGTCGTCCTGTGTTTGT
4 F	GTGTTCTCCAGGGAAGAGCT
5 F	CCAGAACTCCGTCTTCCACT
6 F	GAAGATGCCTGACTTGCCTG
1 R	CAAGTTCACGACGTCATGCT
2 R	GGTCTATCTCCTGACCCAGC
3 R	ATTCTGGATGGGACCCGG
4 R	GCAATGCTGATGTGCTGGT
5 R	CAAAGTAGCTGCTGGGTTGG
6 R	TCTTGATCATCCACTTGACGAA
7 R	AGTTCCACCTTCCACTTCCG

2.16.7 *Engineering NdeI sites*

Primers were designed to immediately flank the start and stop codon of the *CCBL1* cDNA with *NdeI* restriction enzyme sites and five additional random bases to allow correct restriction enzyme cleavage. Primers are detailed in Table 2.16.2.

Engineering of *NdeI* sites was carried out using Phusion High-Fidelity DNA polymerase. In this polymerase, a DNA binding domain is fused to a *Pyrococcus*-like proofreading polymerase. This activity is necessary to avoid the introduction of unwanted mutations whilst amplifying large targets. The reaction mix detailed in Table 2.16.3 was assembled on ice, taking care to add the Phusion High Fidelity DNA polymerase last. Each reaction was then cycled using the parameters outlined in Table 2.16.4 using a Veriti 96-well Thermal Cycler (Thermo Fisher, Loughborough, UK).

Table 2.16.2: NdeI site engineered primers to allow insertion of cDNA into pT7CFE1-CHis expression vector. *NdeI* sites are shown underlined and the random five bases required for proper restriction enzyme cleavage are shown in bold.

Name	Sequence 5' → 3'
Forward	**GCATG**CATATGATGGCCAAACAGCTGCAGGCC
Reverse	**GCATG**CATATGCTAGAGTTCCACCTTCCACTTCC

Table 2.16.3: NdeI site engineering reaction mix.

Reagent	Volume (μL)
Plasmid DNA (20 ng/ul)	1
5X Phusion GC Buffer	10
Forward primer (10 μM)	2.5
Reverse primer (10 μM)	2.5
dNTPs (10 mM)	1
DMSO	1.5
dH_2O	31
Phusion DNA Polymerase	0.5

Table 2.16.4: NdeI site engineering reaction conditions.

Step	Conditions
1	98°C for 2 mins
2	98°C for 10 secs
3	61°C for 30 secs
4	72°C for 60 secs
5	Repeat steps 2-4 24 times for a total of 25 cycles
6	72°C for 10 mins

2.16.8 *TOPO TA cloning*

2.16.8.1 *Addition of 3' overhangs*

Post PCR amplification, 0.2 μL of Taq polymerase was added to each 50 μL reaction. Samples were mixed well, centrifuged and incubated at 72°C for 10 minutes in a Veriti 96 Well Thermal Cycler. Tubes were immediately transfered to ice before proceeding with the TOPO cloning reaction.

2.16.8.2 *TOPO cloning reaction*

4 μL of PCR product, 1 μL of salt solution and 1 μL of TOPO vector were gently mixed and incubated at room temperature for 30 minutes before placing the reaction on ice. TOP10 competent cells were subsequently transformed with the resulting ligation reaction as previously described (Section 2.16.3.2). Orientation of plasmid inserts was carried out by restriction enzyme digestion with EcoRI and NdeI. Details of reagents and digestion conditions are given in Table 2.16.5

Table 2.16.5: Conditions for restriction enzyme digestion of plasmids.

	EcoRI	**NdeI** (CCBL1 in TOPO 2.1)	**NdeI** (pT7CFE1 vector)	**ApaI**
Plasmid DNA (μL)	3	8	2	3
Buffer (μL)	EcoRI buffer (1.5)	Buffer 4 (1.5)	Buffer 4 (1.5)	Buffer 4 (1.5)
Restriction enzyme (μL)	EcoRI (1.5)	NdeI (1.5)	NdeI (1.5)	ApaI (1.5)
dH_2O (μL)	9	4	10	7.5
BSA (μL)	-	-	-	1.5
Incubation temperature (°C)	37	37	37	25
Incubation time (hours)	3	2	2	2

2.16.9 *Site-directed mutagenesis*

Primers were designed to introduce the c.814C>T mutation present in our patient into the wild-type *CCBL1* clone. Mutagenic primers were designed using the following criteria:

1. Both mutagenic primers must contain the desired mutation and anneal to the same sequence on opposite strands.

2. For optimal efficiency, primers should have a minimum of 40% G/C content and terminate in one or more G/C bases.

3. The mutation must lie in the centre of the primers with 10-15 bases of correct flanking sequence.

4. Primers should be between 25 and 45 bases in length with a melting temperature (T_m) of ≥78°. Where:

$$T_m = 81.5 + 0.41(\%GC) - (675/N) - \%mismatch$$

The reaction mix detailed in Table 2.16.7 was assembled on ice, taking care to add the PfuUltra High Fidelity DNA polymerase last. Each reaction was then cycled using the parameters outlined in Table 2.16.8. The total size of the TOPO 2.1 vector with the addition of the *CCBL1* insert is

Table 2.16.6: Site-directed mutagenesis primers to allow insertion of c.814C>T mutation into CCBL1 clone. Mutated bases are shown underlined.

Name	Sequence 5' → 3'
Forward	ATCATGAAGCACCTGCGGACC<u>A</u>TGCACCAGAACTCCGTCTT
Reverse	AAGACGGAGTTCTGGTGCA<u>T</u>GGTCCGCAGGTGCTTCATGAT

~5.3 kb. PfuUltra High Fidelity DNA Polymerase has an extension rate of 1 kb/min, thus the extension time is six minutes.

Table 2.16.7: Site-directed mutagenesis reaction mix using QuikChange II XL Site-Directed Mutagenesis Kit. Volumes of plasmid DNA were adjusted accordingly for a final concentration of 10 ng/reaction, thus volume is denoted as X μL.

Reagent	Volume (μL)
10X Reaction buffer	5
Forward primer (50 ng/μL)	2.5
Reverse primer (50 ng/μL)	2.5
dNTP mix	1
QuikSolution	3
PfuUltra High Fidelity DNA polymerase (2.5 U/μL)	1
Plasmid DNA (10 ng)	X
dH_2O	To a final volume of 50 μL

Table 2.16.8: Mutant strand synthesis reaction conditions.

Step	Conditions
1	95°C for 1 min
2	95°C for 50 secs
3	60°C for 50 secs
4	68°C for 6 mins
5	Repeat steps 2-4 17 times for a total of 18 cycles
6	68°C for 7 mins

Following temperature cycling, reactions were placed on ice for 2 minutes to rapidly cool to <37°C. 1 μL of DpnI was then added to each product, mixed, briefly centrifuged to ensure the contents were at the bottom of the tube and incubated at 37°C for one hour. The DpnI endonuclease (target sequence: 5'-Gm^{6}ATC-3') is specific for methylated and hemimethylated DNA and is used to digest the parental DNA template and to select for mutation containing synthesized DNA. Reactions were stored at -20°C overnight.

2.16.10 *Dephosphorylation to prevent plasmid re-ligation*

2 μL of restriction enzyme buffer, 1 μL of calf intestinal phosphatase, 17 μL of dH_2O and 15 μL of digested DNA (Table 2.16.5) were mixed and incubated at 37°C for 3 hours. Products were stored at -20°C overnight.

2.16.11 *Gel extraction*

DNA was run on a 1% agarose gel (Section 2.3.2) at 95V for 1.5-2 hours to ensure good separation of different sized DNA fragments. DNA fragments were visualised using an ultraviolet transilluminator and excised from the agarose gel with a clean, sharp scalpel. Gel slices were weighed in clean microcentrifuge tubes and DNA was extracted using the QIAquik Gel Extraction Kit. Three volumes of Buffer QG was added to each volume of gel (e.g. 300 μL of Buffer QG to each 100 mg of gel). Samples were incubated at 50°C for 10 minutes in a water bath and were vortexed every 2-3 minutes to make sure the gel slice had completely dissolved. One gel volume of isopropanol was added to the samples, the mixture was applied to QIAquick columns and centrifuged for 1 minute at 17,900 x *g* to bind DNA. The flow-through was discarded, 500 μL of Buffer QG was added, centrifuged for 1 minute at 17,900 x *g* and the flow-through was discarded. To wash the columns, 750 μL of Buffer PE was added and the columns were centrifuged for 1 minute at 17,900 x *g*. The flow-through was discarded and the columns were centrifuged for an additional 1 minute at 17,900 x *g* to remove residual ethanol which may inhibit downstream reactions. The columns were placed into clean microcentrifuge tubes and DNA was eluted by addition of 30 μL of Buffer EB and centrifuging for 1 minute at 17,900 x *g*. A 5 μL aliquot of the purified DNA was then analysed on an agarose gel (Section 2.3.2).

2.16.12 *Ligation*

For optimal ligation using T4 DNA ligase a molar ratio of 1:3 vector to insert and between 50-200 ng of vector DNA is required. In order to calculate the optimal amount of insert DNA the following formula is used:

$$\text{ng of insert} = \frac{\text{ng of insert} \times \text{size of insert in kb}}{\text{size of vector in kb}} \times \frac{\text{molar ratio of vector}}{\text{insert}}$$

The appropriate volumes of pT7CFE1 vector and insert DNA were mixed, incubated at 45°C for 5 minutes and then incubated on ice for 5 minutes. 1 μL of T4 DNA ligase, 2 μL of ligase buffer and 1.6 μL of dH_2O were added, the reactions were mixed and incubated at room temperature for 2 hours. Enzymes were then heat inactivated at 65°C for 10 minutes and reactions were stored on ice until transformation of TOP10 chemically competent cells (Section 2.16.3.2).

2.16.13 *Protein expression using 1-Step Human Coupled IVT Kit*

DNA was concentrated and RNase A (present from QIAprep Spin MiniPrep Kit) was removed by adding 1.5 μL of 3 M sodium acetate (pH 5.2) and 30 μL of 100% ethanol to 15 μL of plasmidic DNA (CCBL1 in pT7CFE1 vector). Samples were mixed and incubated at -20°C for 15 minutes, prior to centrifugation at 14,000 x *g* for 15 minutes. The supernatant was removed and 30 μL of 70% ethanol was added. Samples were centrifuged at 14,000 x *g* for 5 minutes, the supernatant was removed and samples were air-dried for 5 minutes at room temperature. DNA was resuspended in nuclease-free dH_2O to a specific desired final concentration.

Protein expression was carried out using the 1-Step Human Coupled IVT Kit according to the manufacturer's instructions. All reagents were thawed on ice and reactions were assembled in nuclease-free microcentrifuge tubes. 12.5 μL of HeLa cell lysate, 2.5 μL of accessory proteins, 5 μL of reaction mix, 2 μL of plasmidic DNA (CCBL1 in pT7CFE1 vector) and 3 μL of nuclease-free dH_2O were mixed together gently at room temperature. Reactions were incubated at 30°C for between 90 minutes and 6 hours and maintained on ice for subsequent downstream applications.

2.17 Q-TOF ANALYSIS OF IN VITRO PROTEIN EXPRESSION

Proteins expressed using the 1-Step Human Coupled IVT Kit (Section 2.16.13) were lyophilised overnight using a freeze-drier pror to resuspending in 20 μL of 100 mM Tris-HCl (pH 7.8) containing 2% (w/v) amidosulphobetanine-14 (ASB-14), 6 M urea and 2 M thiourea. 1.5 μL of 200 mM dithioerythritol in 100 mM Tris-HCl (pH 7.8) was added to each sample, vortexed and incubated with shaking at room temperature for 60 minutes. 3 μL of 200 mM iodoacetamide in 100 mM Tris-HCl (pH 7.8) was added to each sample, vortexed and incubated with shaking at room temperature for 30 minutes whilst protected from light. 155 μL of dH_2O and 1 μg of Mass Spectrometry Grade Trypsin Gold was added and samples were incubated at 37°C overnight to digest the proteins.

Residual salts, detergents, urea, thiourea, iodoacetamide and dithioerythritol were removed from samples prior to mass spectrometry analysis using C18 solid phase extraction columns. C18 columns were first primed with 1 mL of 50% acetonitrile (ACN) [0.1% TFA], followed by 1 mL of 0.1% TFA. Samples which had been diluted 1:1 with 0.2% trifluoroacetic acid (TFA) to aid binding of samples to the solid silica phase were subsequently applied to the primed columns. The eluents were collected in siliconised microcentrifuge tubes before being re-applied to the columns and collected again. Residual salts present in the samples were removed by washing the columns with 1 mL 3% ACN [0.1% TFA]. Peptides were eluted in 500 μL 50% ACN [0.1%

TFA] and collected in clean siliconised microcentrifuge tubes. Eluted peptides were lyophilised overnight.

Peptide pellets were resuspended in 300 μL 3% ACN containing 0.1% TFA and 50 pmol/μL enolase peptides standard (MassPREP) prior to centrifugation at 16,000 x g for 10 minutes. The supernatant was then transferred to a TruView LCMS Certified, Total Recovery Vial. 1 μL of sample was injected onto a nanoACQUITY ultra high performance liquid chromatography system coupled to a SYNAPT G2-Si mass spectrometer (Waters). Peptides were separated using low pH reverse phase chromatography on an ACQUITY UPLC Peptide BEH C18 nanoACQUITY Column (130 Å, 1.7 μm, 75 μm x 150 mm). The column was maintained at 35°C and the flow rate was 400 nL/minute. The mobile phases used were as follows: A, 0.1% formic acid with 5% DMSO and B, 0.1% formic acid in 100% acetonitrile with 5% DMSO. Peptides were diluted 1:10 during trapping with mobile phase A and concentrated and desalted onto a nanoACQUITY UPLC Symmetry C18 Trap Column (100 Å, 5 μm, 180 μm x 20 mm). Chromatographic separation was achieved using the following gradient: 3% mobile phase B, increasing linearly to 40% over 40 minutes and increasing linearly to 85% over the following 2 minutes. The mobile phase composition was held for a further 2 minutes, before returning to the initial conditions for 15 minutes of re-equilibration.

Mass spectrometry analysis was performed over 60 minutes on a SYNAPT G2-Si mass spectrometer in UDMS^E positive ion electrospray ionisation mode and operated in V-mode. One second alternating high and low energy scans were performed at a capillary voltage of 3.0 kV, sampling cone voltage of 40 V and a source temperature of 70°C over a mass range of 50-2000 Da in resolution analyser mode. Ion mobility separation was performed to separate similar precursor ions using a wave velocity of 650 m/s and a wave height of 40 V. Low energy scans were performed using a collision energy of 0 V and the high energy scans were performed using a gradient of collision energies, optimised depending on the ion mobility bin. The collision energy was 13.6 V from 0-20 ion mobility bins increasing linearly to 49.1 V at 120 mobility bins, followed by another linear gradient to 54.1 V at 200 mobility bins. Every 60 seconds [glu1]-fibrinopeptide B was delivered via an auxiliary pump at a flow rate of 300 nL/min as a lock mass to permit real-time recalibration by correcting m/z shifts arising from instrumental drift.

Raw data was imported into ProteinLynx GlobalServer version 3.0.1 (Waters, Manchester, UK) in order to identify peptide masses corersponding to the fragmentation ion data. Mass corrections were applied based on the [glu1]-fibrinopeptide B mass delivered via an auxillary pump. Spectra were compared to the UniProt reviewed human proteome (www.uniprot.org/downloads) using the following searching parameters: two fragment ions matched per peptide, four fragment ions per protein, two peptides per protein and one missed enzymatic cleavage. Protein concentrations

were determined by comparison to a known concentration (50 pmol/μL) of spiked enolase peptide standard.

2.18 Quantification of CCBL1 enzyme activity with respect to kynurenine using UPLC-MS/MS

Experiments were undertaken to determine whether the overexpressed CCBL1 enzyme had kynurenine aminotransferase activity. All reagents were made up in 200 mM sodium phosphate buffer (pH 7.5). 2.5 μL of IVT mix was transferred into five labelled eppendorfs corresponding to the five time-points used to measure the enzyme activity [0, 15, 30, 60, 120 minutes]. 47.5 μL of master mix [containing 5 μL of 50 μM kynurenic acid, 2 μL of 50 μM oxaloacetate, 2 μL of 1 μM PLP, 38.5 μL of 200 mM sodium phosphate (pH 7.5)] was then added. The 0 minute reaction was stopped by adding 50 μL of 0.3 TCA and vortexing thoroughly for 30 seconds. The remaining tubes were incubated with shaking at 37°C for the appropriate times, and subsequently stopped with 0.3N TCA. Tubes were then left on ice in the dark for 45 minutes to allow any protein to precipitate, before centrifuging at 10,000 rpm for 10 minutes to pellet the proteins present in the reaction mixture. The resulting supernatant was transferred to a HPLC vial and placed in an autosampler, protected from light and kept at 4°C until sample injection.

Stock solutions of kynurenine (KN), kynurenic acid (KA) and d_5-kynurenic acid were made up using dH_2O and stored at -80°C to prevent degradation. Quantification was achieved by spiking 2 μL of 3 μM d_5-KA to each reaction. Calibration curves were constructed using different concentrations of KN and KA, whilst keeping the concentration of d_5-KA identical to that spiked into each reaction. Concentrations of KN and KA were then determined by ratioing the analyte signal area to that of d_5-KA. Analyte concentrations were expressed as nmol/L.

LC-MS/MS was performed using an Acquity Ultra Performance LC sytem linked to a triple quandrupole Xevo TQ-S instrument (Waters, Manchester, UK). An Acquity UPLC HSS T3 column (1.8 μm x 2.1 mm x 50 mm) fitted with a HSS T3 VanGuard guard column (Waters) was used with a mixture of mobile phase A (100% methanol) and B (4 mM ammonium acetate pH 2.1 with formic acid) at a flow rate of 0.5 mL/minute. Details of the mobile phase gradients are shown in Table 2.18.1.

Analytes were detected using the mass spectrometer in positive ion mode using multiple reaction monitoring (MRM). Kynurenine and kynurenic acid could be identified based on their retention time and the m/z ratios of their corresponding parent and daughter ions (Table 2.18.2).

Table 2.18.1: Mobile phase gradient profile for separation of kynurenine and kynurenic acid in fibroblasts.

Time (minutes)	A (100% methanol)	B (4 mM ammonium acetate pH 2.1 with formic acid)	Curve
0.00	0.2	99.8	0
1.00	0.2	99.8	6
2.00	85.0	15.0	6
5.00	85.0	15.0	2
6.00	0.2	99.8	6
8.00	0.2	99.8	11

Table 2.18.2: Mass transitions, cone voltages, collision energies and retention times of analytes quantified using the CCBL1 enzyme assay. All analytes were detected using a Xevo TQ-S mass spectrometer in positive ion mode using MRM.

Analyte	Molecular weight (g/mol)	Parent ion (m/z)	Daughter ion (m/z)	Cone voltage (V)	Collision energy (V)	Retention time (mins)
Kynurenine	208.22	209.14	192.08	20	8	1.84
Kynurenic acid	189.17	190.20	144.00	30	12	2.11
d_5-kynurenic acid	194.17	195.20	148.96	30	12	2.11

A 15 μL volume of each CCBL1 enzyme assay time-point mixture was injected on to the mass spectrometer every 8 minutes and the data was acquired using MassLynx software (Waters). Data processing was completed using QuanLynx software (Waters).

2.19 QUANTIFICATION OF ENDOGENOUS KYNURENINE AND KYNURENIC ACID USING UPLC-MS/MS

5 μL neat urine was added to 53 μL dH_2O and 2 μL 3 μM d_5-KA. 60 μL 0.3N TCA was then added to each sample before being centrifuged to pellet any precipitated protein. The resulting supernatant was transferred to a HPLC vial and placed in an autosampler, protected from light and kept at 4°C until sample injection. Quantitation of the endogenous KN and KA was performed using the UPLC-MS/MS method detailed in Section 2.18. Results were corrected for urinary creatinine concentration using the method described by (Mills *et al.*, 2010).

2.20 COLOURMETRIC ASSAYS FOR DETECTION OF TERTIARY AMINES

2.20.1 *Urinary ninhydrin assay*

200 μL of urine was added to 300 μL of 3N perchloric acid and 200 μL of 2% (w/v) ninhydrin. Samples were heated for 5 minutes at 100°C and centrifuged at 16,000 x *g* for 10 minutes. Absorbance was measured at 492 nm and 555 nm. Results were subsequently corrected by urinary creatinine concentration using the method described by (Mills *et al.*, 2010).

2.20.2 *Urinary o-aminobenzaldehyde (o-AB) assay*

250 μL of urine was added to 250 μL of 40 mM o-aminobenzaldehyde (o-AB) in 200 mM potassium phosphate (pH 8). Samples were incubated at 37°C overnight and the absorbance was measured at 450 nm using a spectrophotometer. Results were subsequently corrected for urinary creatinine concentration.

2.20.3 *Detection of in vitro translation reaction products using o-aminobenzaldehyde (o-AB)*

Experiments were designed to determine whether CCBL1 produced by an IVT reaction could convert lysine to an equilibrium mixture of α-keto-ϵ-aminocaproate and Δ^1-piperideine-2-carboxylate. The latter of which could be expected to produce a coloured reaction product with o-AB.

Mixes containing 5 mM L-lysine, 70 μM PLP, 100 mM α-keto acid, 100 mM potassium phosphate buffer (pH 8.0) and 10 μL CCBL1 IVT reaction mix in a final volume of 250 μL were prepared. Four different reactions were performed, each using a different α-keto acid:

2-oxobutyrate, 2-oxoglutarate, pyruvate or oxaloacetate. An additional sample with no α-keto acid substrate was also prepared. Each sample was incubated at 30°C for 4 hours and terminated by the addition of 11.4 μL 3N TCA and 113.6 μL 100% ethanol. 250 μL of the de-proteinised supernatant was then mixed with 250 μL of o-AB mix [containing 40 mM o-AB in 200 mM potassium phosphate buffer (pH 8.0)]. The samples were incubated at 37°C overnight and the absorbance was subsequently analysed at 450 nm.

2.21 Synthesis of α-keto-ϵ-aminocaproate and Δ^1-piperideine-2-carboxylate and detection of these compounds and lysine

Two methodologies were used to synthesise an equilibrium mixture of α-keto-ϵ-aminocaproate and Δ^1-piperideine-2-carboxylate from D-lysine using D-amino acid oxidase from porcine kidney. In the first (based on Konno (1998)), 300 μL 0.133 M pyrophosphate buffer (pH 8.3), 700 IU catalase, 300 μL 0.1 M D-lysine, 200 μL 0.1 mM flavin adenine dinucleotide, 100 μL 70% methanol and 10 units of D-amino acid oxidase (in 100 μL of dH_2O) was incubated at 37°C for 2 hours. The reaction was terminated by the addition of 1 mL 0.3 N TCA. In the second method (based on Visser *et al.* (2012)), mixes containing 25 mM ammonium bicarbonate buffer (pH 8.3), 1 mM D-lysine, 0.25 mg D-amino acid oxidase and 900 units of catalase in a total volume of 100 μL were prepared. 0.1 mM flavin adenine dinucleotide was added to half of the reactions. Each reaction was terminated by the addition of 100 μL 0.3 N TCA. Subsequently, each reaction mix was desalted using C18 solid phase extraction columns prior to direct infusion into the Xevo TQ-S mass spectrometer. Columns were primed with 1 mL 50% acetonitrile (ACN) [0.1% TFA], followed by 1 mL 0.1% TFA. Samples were added to the primed columns, and the eluents were reapplied to the columns. Salts were removed by washing the columns with 0.1% TFA. Other molecules were eluted in 1 mL 50% ACN [0.1% TFA] followed by 100% methanol. The transitions identified are shown in Table 2.21.1. LC-MS/MS was performed using the same mass spectrometer and column as for the quantification of the B_6 vitamers (Section 2.11) as many of the methods published for the analysis of amino acids (such as lysine) use columns with similar physicochemical properties. The gradient used is shown in Table 2.21.2 and the flow rate was maintained at 0.6 mL/min. Analytes were detected in positive ion mode using multiple reaction monitoring.

Table 2.21.1: Optimal transitions for the detection of lysine and Δ^1-piperideine-2-carboxylate (P2C). *, predicted compound.

Analyte	Parent ion (m/z)	Daughter ion (m/z)	Cone voltage (V)	Collision voltage (V)	Retention time (mins)
Lysine	147.10	56.08	28	24	0.58
Lysine	147.10	84.09	28	14	0.58
P2C*	128.10	55.04	40	20	0.54
P2C*	128.10	82.08	40	14	0.54

Table 2.21.2: Mobile phase gradient profile for the separation of lysine metabolites. HFBA, heptafluorobutyric acid.

Time (minutes)	A (3.7% acetic acid + 0.01% HFBA)	B (100% acetonitrile)	Curve
0.00	99.8	0.2	0
0.40	99.8	0.2	6
3.75	50.0	50.0	6
3.76	0.1	99.9	11
4.25	0.1	99.9	11
4.26	97.5	2.5	1
5.20	97.5	2.5	1

To assess the activity of CCBL1 towards lysine, 5 μL of IVT mix containing overexpressed CCBL1 protein (Section 2.16.13) was transferred into five labelled eppendorfs corresponding to five time-points [0, 30, 60, 120 and 180 minutes]. 45 μL of master mix containing 1 μM L-lysine, 1 mM oxaloacetate, 40 μM PLP and 1000 units of catalase was then added. Each sample was incubated at 37°C for the appropriate time before terminating with 50 μL of 0.3 N TCA. Samples were then precipitated and injected as previously (Section 2.18), before analysis using the parameters described above (Tables 2.21.1 and 2.21.2).

3

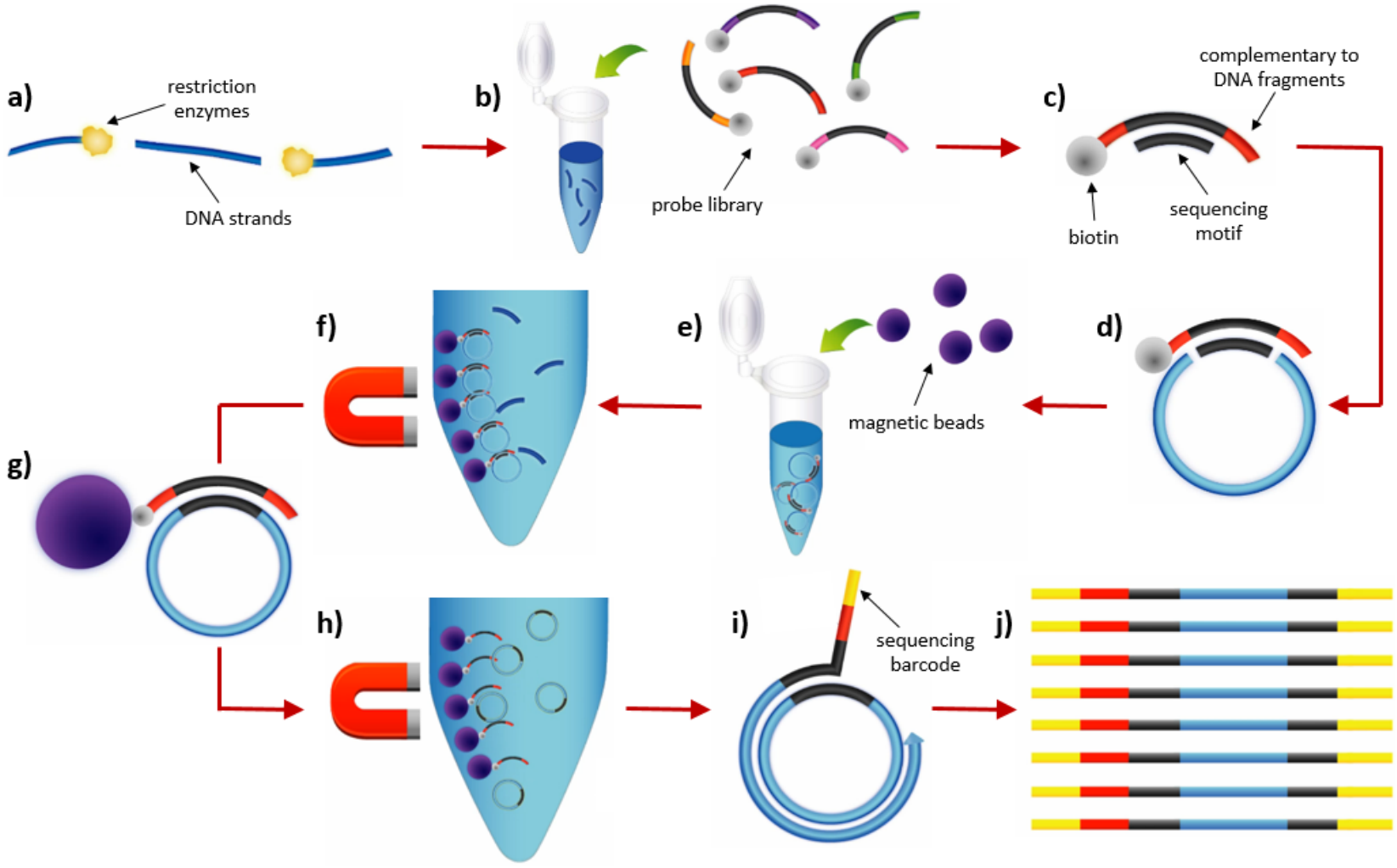
VALIDATION AND USE OF GENE PANEL SEQUENCING TECHNOLOGY FOR THE DIAGNOSIS OF INBORN ERRORS OF METABOLISM

3.1 INTRODUCTION

As described in Chapter 1, there are many challenges to the accurate diagnosis of IEM, mainly attributable to the clinical and genetic heterogeneity, the often-atypical presentation early in the disease course, the lack of clinical awareness of rare entities and the unavailability of certain diagnostic tests in non-specialist centres. Hence, patients with suspected IEM, especially those with accompanying neurological signs, are often referred to specialist centres early in life. However, despite this, there are very often diagnostic delays or difficulties establishing the correct IEM diagnosis. Indeed, despite extensive genetic and biochemical investigations the underlying aetiology remains undetermined in many cases, and in up to 50% of patients at some tertiary referral centres. Next-generation sequencing technologies are ideal for the genetic diagnosis of patients such as these, in which an unbiased approach is beneficial. Both WES and WGS have been proven to be effective for the diagnosis of IEM in a research setting (Tarailo-Graovac *et al.*, 2016; Dinwiddie *et al.*, 2013; Stranneheim *et al.*, 2014). However, the identification of causal variant(s) is a challenge requiring expert analysis and collaboration within a multidisciplinary team. Indeed, much of the time taken for WES and WGS is primarily due to data analysis, which is often prohibitive in clinical laboratories. Ideally, for those disorders which are amenable to treatment, diagnosis should be more rapid than currently achievable using WES or WGS so as to minimise the impact of any irreversible damage that may be caused. Examples include directing early vitamin supplementation with pyridoxal 5'-phosphate in patients with pyridox(am)ine 5'-phosphate oxidase deficiency (Mills *et al.*, 2014) or with biotin and thiamine in patients with biotin-thiamine-responsive basal ganglia disease (Alfadhel *et al.*, 2013). Misdiagnosis or diagnostic delay of these disorders can lead to a variety of adverse outcomes including neuropsychological dysfunction, mental retardation and death.

The objective of this chapter was to design and investigate the utility of a comprehensive gene panel targeting all IEM-causing genes and use it to investigate patients presenting with complex neurometabolic phenotypes. An Agilent HaloPlex Target Enrichment System was used which makes use of restriction enzyme digestion and custom-designed probes to capture the genes of interest prior to sequencing using an Illumina platform (Figure 3.1.1).

Figure 3.1.1: Targeted gene capture using HaloPlex technology. **(a)** Each DNA sample is fragmented using restriction enzymes. **(b)** A library of custom-designed probes which target the genes of interest are then added. **(c)** Each probe is an oligonucleotide designed to hybridise to both ends of a targeted DNA restriction fragment. The ends of the probe are complementary to the desired fragment sequence, the middle part contains a sequencing-specific motif and each is biotinylated to facilitate capture. **(d)** The probes hybridise to the target fragments, forming a circular shape. **(e)** Magnetic streptavidin beads are added, **(f)** the circular DNA fragments are captured and the non-circular fragments are removed. **(g)** DNA ligase closes the DNA circles and **(h)** the circular DNA fragments are then eluted from the beads. **(i)** PCR occurs using universal primers containing a barcode which allows for sample multiplexing. **(j)** Multiple PCR cycles generates barcoded linear fragments ready for sequencing.



The panel's effectiveness in establishing diagnoses, clinical implications and contribution to expansion of known phenotypes are discussed. The value of *in silico* tools commonly used by research and clinical diagnostic laboratories to predict the pathogenicity of sequence variants are also considered.

3.2 METHODS

3.2.1 *Patient recruitment*

A total of forty four patients were recruited from metabolic medicine clinics at Great Ormond Street Hospital for Children. For panel validation purposes, fourteen patients with a known genetic diagnosis were recruited to verify the panel's ability to identify pathogenic variants (Table 3.3.4). Additionally, thirty patients presenting with neurometabolic disease and lacking a specific genetic diagnosis were selected for targeted sequencing (Table 3.3.6 and Table 3.3.10). Written informed consent was obtained in all cases.

3.2.2 *Design of amplicons to enable maximal capture of IEM genes using HaloPlex enrichment*

The design of the IEM gene panel was based on genes listed on the SSIEM Classification of Inborn Errors of Metabolism (2011) (www.ssiem.org) with additions from recent publications, existing smaller disease-specific panels [Mitome200-Nuclear (Baylor College of Medicine) and Mitochondrial Disorders Panel (ARUP Laboratories)] and suggestions from the GOSH Metabolic Team. The final design contained 614 genes known to cause IEM and covered 16 broad sub-classes of this group of disorders (Table 3.2.1). A complete list of the genes included can be found in (Appendix 9.1). The gene accession numbers and coordinates were downloaded from the UCSC Genome Browser (www.genome.ucsc.edu) and the SureDesign system (Agilent Technologies Inc. USA) was used to design the probe library for the genes of interest. The design was based on a read length of 150bp on an Illumina platform and targeted all coding exons of the target genes as well as 25bp of flanking intronic sequence. This created a 1.43Mbp design with a predicted target coverage of 99.55%. 6471bp were predicted to not be covered. The failure to completely cover a region of interest can be attributed to repeated regions in the flanking sequences, lack of restriction fragments of appropriate size, or fragments too large relative to the read length leading to partial sequencing of fragments. When using the SureDesign software, amplicons are selected from only one DNA strand for each target and amplicons containing known polymorphisms are excluded. Coverage was optimised, in collaboration with Agilent, by including amplicons from

both strands of each target that was not initially fully covered and increasing the number of probes to capture all known haplotypes in order to increase target coverage. By implementing these changes, predicted coverage of 39% of the 6471bp that would not have been covered by the original design was possible, without reducing the stringency of the design.

Table 3.2.1: The sixteen classes of IEM represented within the gene panel and the distribution of the 614 genes amongst them. Note that due to the highly interconnected and heterogeneous nature of metabolism, mutations in one gene can give rise to multiple different disorders, leading to an apparently higher total number of genes than stated in the text.

Disorder Class	Number of genes
Amino acid and peptide metabolism	98
Carbohydrate metabolism	43
Fatty acid and ketone body metabolism	18
Energy metabolism	154
Metabolism of purines, pyrimidines and nucleotides	33
Metabolism of sterols	22
Porphyrin and haem metabolism	9
Lipid and lipoprotein metabolism	32
Congenital disorders of glycosylation and other disorders of protein modification	76
Lysosomal disorders	54
Peroxisomal disorders	24
Neurotransmitter metabolism	8
Metabolism of vitamins and non-protein cofactors	40
Metabolism of trace elements and metals	24
Metabolism of xenobiotics	2
Other disorders	11

3.2.3 *Target capture, library sequencing and variant calling*

Genomic DNA was extracted from EDTA blood using the AutoGenFlex STAR automated system according to the manufacturer's instructions (Section 2.3). The materials and methods used for HaloPlex target enrichment, next-generation sequencing using the Illumina platform and variant calling are described in Section 2.4. Sequence variants with putatively deleterious effects were confirmed by Sanger sequencing as described in Section 2.3.3.2. Primer sequences are detailed in Appendix 9.1.1.

3.3.1 *Analysis pipeline for the indentification of potentially pathogenic variants*

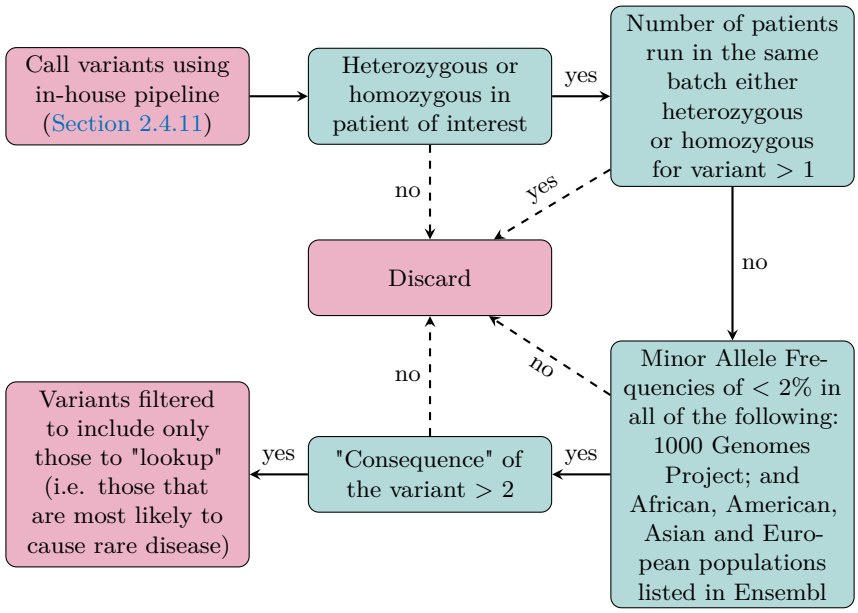
Raw FASTQ files were aligned to the hg19 reference genome and variants were called using the method described in Section 2.4.11. Variants were initially filtered using the standard pipeline used by the North East Thames Regional Genetics Service Laboratories (summarised in Figure 3.3.1). Firstly, the variant must be present in a homozygous or heterozygous state in the patient of interest and not identified in any other patient analysed in the same sequencing run. An exception to this can be made if siblings or other family members are analysed together. Due to the rarity of individual IEM (even the most common, phenylketonuria and medium-chain-acyl-CoA dehydrogenase deficiency, have an incidence of approximately 1 in 15,000), it is highly unlikely that multiple unrelated patients with the same genetic defect and even less-so the same mutation would be analysed in the same batch. Nevertheless, this analysis algorithm is not fixed and multiple different iterations of analysis with varied parameters can be performed, thus a good potentially pathogenic variant (e.g. strong phenotypic correlation with other patients with mutations in that gene, low minor allele frequency etc.) may not be discarded based on a single criterion. Similar flexibility can also be employed in the case of a candidate variant that is identified in a heterozygous state in another unrelated patient.

Secondly, only variants with a minor allele frequency of less than 2% are selected as only variants that are rare in the general population are likely to cause rare metabolic diseases. Finally, the variants are ranked by their predicted consequence score (Table 3.3.1). This is an arbitrary score that predicts the likelihood of a type of mutation causing disease; for example a synonymous variant in the 3'UTR region of a protein is less likely to cause disease than an exonic stop gain mutation. These filtering parameters are termed the "Lookup" filter and determine which of the $\sim$ 1500 variants identified in each patient are worthy of further scrutiny.

Table 3.3.1: Consequence scores of gene panel variants. The set of consequence terms is defined by the Sequence Ontology (http://www.sequenceontology.org/). Each variant is mapped to the reference genome and a rule-based approach is used to predict the effect that each variant may have on the gene transcript. Note that each variation may have a different effect in different transcripts, hence more than one term is often applied to each variant. These terms have been ranked from 1 - 5 based on their predicted effect on function, with 1 being the most likely to be benign and 5 being the most likely to be damaging.

Consequence terms (i.e. the variant type)	Consequence ranking
Stop lost	5
Stop gained	5
Frameshift variant	4
Splice donor variant	4
Splice acceptor variant	4
Splice region variant	3
In-frame insertion	3
In-frame deletion	3
Missense variant	3
Initiator codon variant	3
Stop retained variant	3
Transcription factor binding site variant	2
Synonymous variant	2
3' untranslated region variant	1
5' untranslated region variant	1
Upstream gene variant	1
Downstream gene variant	1
Regulatory region variant	1
Non-coding exon variant	1
Nonsense-mediated decay transcript variant	1
Non-coding transcript variant	1
Intron variant	1

Figure 3.3.1: "Lookup" filtering pipeline applied to variants. Minor allele frequencies were collated from 1000 Genomes Project (www.1000genomes.org/) and Ensembl (www.ensembl.org/) using data from African, American, Asian and European populations. Types of variants are ranked on their predicted consequence (i.e. a 3' UTR variant has a ranking of 1, whereas a stop gain variant has a ranking of 5).



After applying this filter, approximately 40 candidate variants remained for each patient. In a routine diagnostic setting it is important to aim to minimise the time required for manual data interrogation to reduce the costs and turn-around time for the assay. An analysis pipeline was developed to help further filter these variants (shown in Figure 3.3.2). This algorithm was based on a number of parameters assessed using the tools detailed in Section 2.6, including:

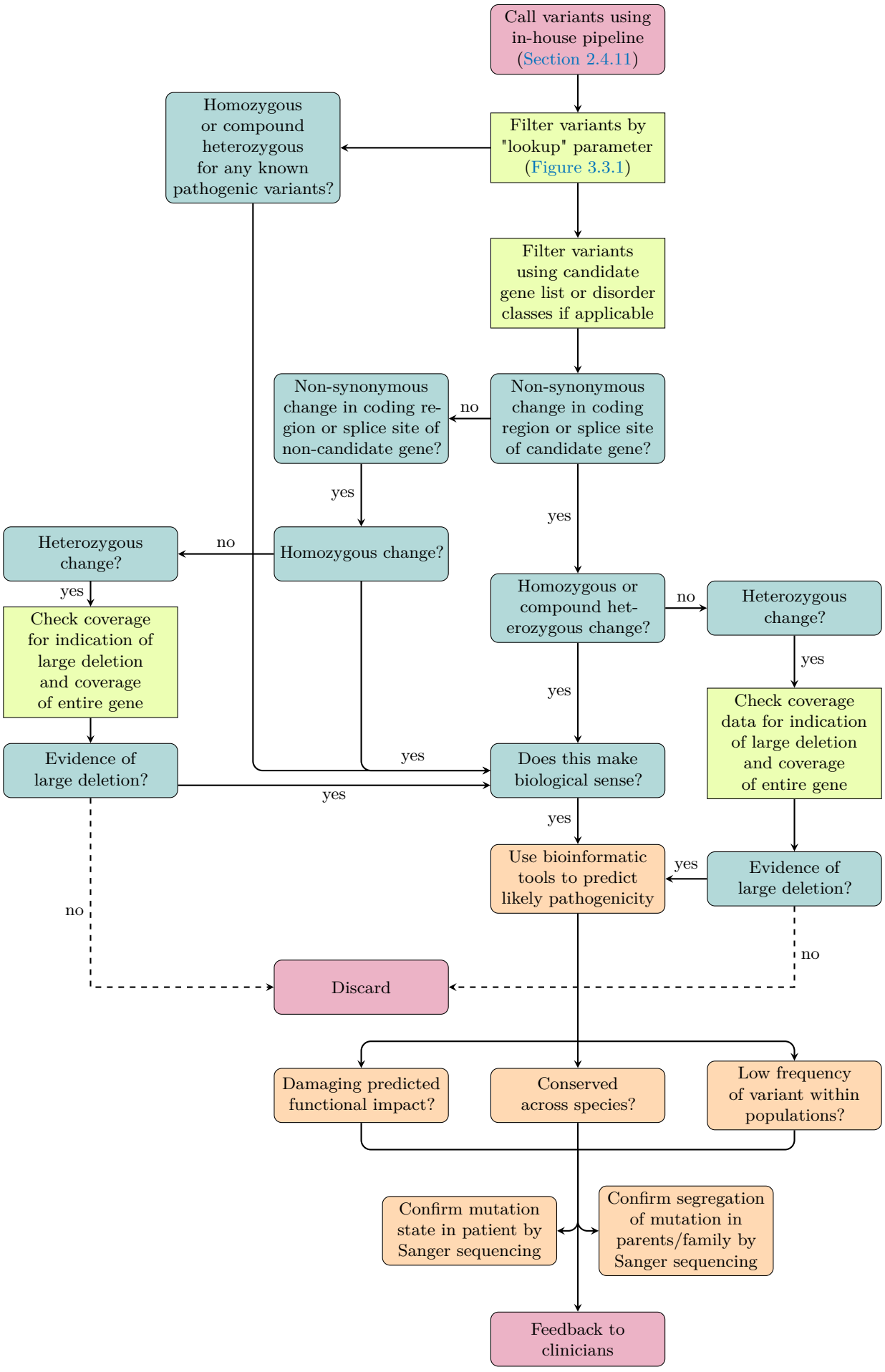
- Whether the variant had been reported as pathogenic previously;
- The frequency of the variant within populations;
- Predicted functional impact of the variant with nonsense and frameshift mutations, single or multi-exon deletions, and variants affecting the canonical splice sites (± 1 or 2 bases from econ boundaries) or initiation codons predicted to cause the greatest loss of function;
- Conservation of the mutated residue amongst species;
- Correlation of the patient's phenotype with the clinical and biochemical features of other cases reported in the literature with pathogenic mutations in the same gene;
- Segregation of the variant within the patient's family (where samples were available);

Highly conserved DNA sequences are thought to have functional value as particular bases or amino acids have been maintained by evolution despite speciation. The further back in the

phylogenetic tree a residue occurs, the more highly conserved it is considered to be. Deleterious sequence variants are also thought to be rare in the population as their presence reduces an organism's fitness. Variants such as this are removed from the population as the probability of survival and subsequent reproduction are reduced. SIFT and PolyPhen-2 were the two online tools used to predict the functional impact of variants. Both tools consider sequence homology and physico-chemical amino acid properties to make predictions. PolyPhen-2 additionally uses annotations of functionally important domains, 3D structures from the Protein Data Bank and a number of other tools. SIFT provides qualitative predictions (either "tolerated" or "deleterious") with normalised probabilities that amino acid changes are tolerated between 0-1. Substitutions with a score < 0.05 are classified as "deleterious". PolyPhen-2 also provides qualitative predictions (one of "probably damaging", "possibly damaging" or "benign") and a probability score between 0-1. Substitutions with a score nearer 1 are more confidently predicted to be deleterious.

Initially, the data was searched for any homozygous or compound heterozygous mutations that are known to be pathogenic and would be consistent with the phenotype of the patient (Figure 3.3.2). If none were identified, clinical phenotypic information was considered, either by using a candidate gene list or filtering by disease class. Candidate gene lists were generated using Phenomizer, a database based on Human Phenotype Ontology (HPO) terms, in which the clinical and/or biochemical features of a particular patient were input to produce a list of all the disorders and genes associated with them that could fit the phenotypic profile (Köhler *et al.*, 2009). The HPO provides a standardised vocabulary of phenotypic features seen in human disease. It is collated using information from Online Mendelian Inheritance in Man (OMIM) (www.omim.org) and the medical literature and contains approximately 10,000 terms (Kohler *et al.*, 2014). Alternatively, if applicable, variants were searched for within a certain disease class. For example, if the patient had lactic acidosis or abnormal muscle respiratory chain activity, variants within genes known to cause mitochondrial disorders were interrogated first.

Figure 3.3.2: Analysis pipeline for IEM gene panel.



Where phenotypic filtering was possible, the resulting reduced number of variants were examined further in order to determine whether any homozygous or compound heterozygous non-synonymous variants were present that may explain the patient's phenotype. If plausible variants were identified then SIFT and PolyPhen-2 scores were used to predict their functional impact, multiple sequence alignments were generated using Clustal Omega to assess the cross-species conservation of the mutated residue in the case of single nucleotide variants, and databases of population statistics (Ensembl, Exome Aggregation Consortium and 1000 Genomes) were interrogated to determine the minor allele frequency of the variant. If, after consultation with the clinician looking after the patient, the variants were deemed likely to be pathogenic then Sanger sequencing was used to confirm the mutation state in the patient and segregation in the parents or other family members if DNA was available. When phenotypic filtering did not reveal any candidate variants, all other variants meeting the minimal criteria for potentially causing rare autosomal recessive disease (see "Lookup" filtering (Figure 3.3.1)) were also examined.

Once candidate variants had been identified, they were secondarily ranked according to American College of Medical Genetics and Genomics (ACMG) guidelines (Richards *et al.*, 2015). These recommendations were compiled to enable the objective classification of sequence variants into five classes using criteria informed by expert opinion and empirical data: pathogenic, likely pathogenic, uncertain significance, likely benign or benign. These classes are determined based on the strength of the evidence supporting pathogenicity from different sources including *in silico* predictions, population databases and the scientific and medical literature.

3.3.2 *Optimisation of target coverage and sequencing metrics*

The gene panel was designed to target 7522 exons and 25 base pairs of flanking intronic sequence of the 614 genes of interest. This created a 1.43Mbp design with a predicted coverage of 99.5%. The sequencing parameters then needed to be optimised in order to maximise the utility of the panel for clinical use. Read depth (i.e. the average number of times a base has been sequenced) is critical for the ability to distinguish heterozygous and homozygous variations. In order to confidently call a heterozygous variant, the DNA must be sequenced enough times so that there is a reasonable probability of observing both alleles (Bentley *et al.*, 2008). Traditionally, a read depth of 30X was considered to be the threshold below which coverage is inadequate for accurate variant calling. Practically however, covering all exons and intron-exon boundaries above a 30X threshold is often not achievable, especially with a panel this large. This is due to various factors including variation in DNA capture, partial sequencing of fragments, repetitive regions causing poor PCR amplification and variation in performance of the chosen sequencing platform

(Berglund *et al.*, 2013). As the use of NGS has become more widespread, guidelines have changed and laboratories tend to use varying minimum thresholds, sometimes as low as 10X (Rehm *et al.*, 2013). Two of the most frequently used sequencing platforms in both research and diagnostic laboratories, the Illumina MiSeq and HiSeq platforms, were used to sequence the capture regions of our patients' DNA. These platforms share the same functional chemistry; the major difference is that the HiSeq can generate at least one order of magnitude more sequencing reads than the MiSeq, thereby allowing greater multiplexing whilst maintaining target coverage. In some cases, coverage of inadequate regions may be boosted simply by sequencing fewer samples per batch thus allocating each sample more sequencing capacity leading to a higher average read depth. However, a balance must be struck between the potential increase in data quality and the consequent increase in cost per sample.

In order to ascertain the number of samples that could be run concurrently, initial studies were performed using the Illumina MiSeq platform and involved varying the number of samples pooled per sequencing lane in order to maximise target coverage. Graphical and numerical representations of the sequencing metrics obtained are shown in Figure 3.3.3 and Table 3.3.2, respectively. When seven patient samples were pooled, a mean depth of 174.0X sequencing coverage was observed and 94.0% of the targeted regions were successfully sequenced with a depth of at least 20X. On average 1099 (range: 1033 - 1157) variants were identified in each patient with an average of 39 (range: 22 - 53) variants remaining after preliminary "Lookup" filtering. When four patient samples were pooled, a mean depth of 462.4X sequencing coverage was observed and 97.1% of the targeted regions were successfully sequenced with a depth of at least 20X. On average 1060 (range: 973 - 1301) variants were identified in each patient with an average of 32 (range: 18 - 61) variants remaining after filtering. In contrast, when 24 patient samples were pooled and sequenced using Illumina HiSeq platform, a mean depth of 924.5X sequencing coverage was observed and 97.8% of the targeted regions were successfully sequenced with a depth of at least 20X. On average 1562 (range: 1456 - 1658) variants were identified in each patient with an average of 28 (range: 14 - 38) variants remaining after filtering. This revealed that when sequencing seven patients in one MiSeq run, fewer bases were covered at read depths higher than 30X when compared to running 24 samples on the HiSeq or four samples on the MiSeq (Figure 3.3.3).

Figure 3.3.3: Comparison of read depth parameters between differing batches of samples analysed using either the MiSeq or HiSeq platform.

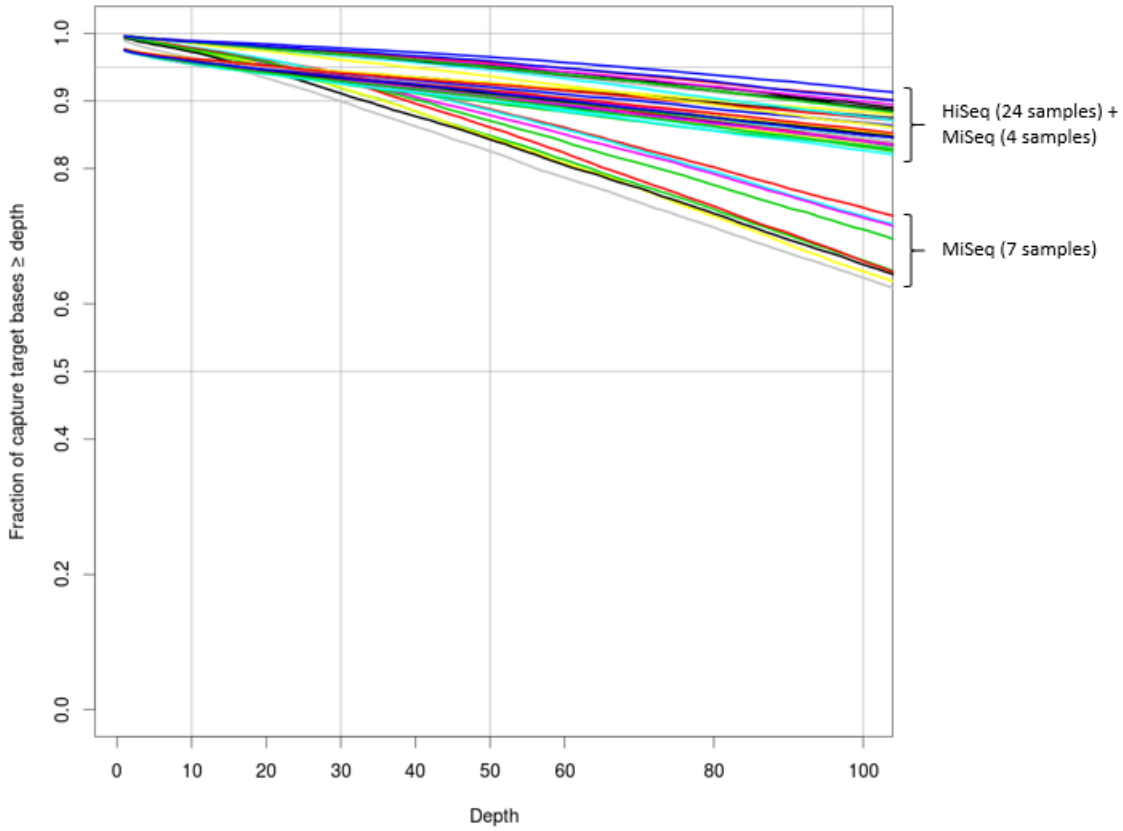


Table 3.3.2: Percentage and range of alignment for all samples sequenced on the Illumina HiSeq/MiSeq platforms in the regions of interest (ROI). The mean percentage and range of exons covered at greater than 10X, 20X, 30X, 50X and 100X are also shown. Finally, the mean and range of base coverage across the samples is shown.

Sequencing platform (number of samples)	% ROI coverage >10X		% ROI coverage >20X		% ROI coverage >30X		% ROI coverage >50X		% ROI coverage >100X		Mean base coverage per sample	
	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
MiSeq (7)	96.5	94.9 - 97.3	94.0	90.9 - 95.4	91.0	86.3 - 93.2	84.2	76.5 - 88.0	66.0	52.7 - 72.4	174.0	128 - 200.4
MiSeq (4)	98.0	97.9 - 98.1	97.1	96.7 - 97.4	96.2	95.4 - 96.7	94.2	92.9 - 95.1	88.2	86.0 - 89.8	462.4	406 - 488.1
HiSeq (24)	98.3	98.0 - 98.5	97.8	97.4 - 98.1	97.2	96.8 - 97.7	96.2	95.6 - 97.0	93.4	92.3 - 94.9	924.5	969.6 - 1006.0

In summary, sequencing 24 patients in one HiSeq run resulted in double the average coverage than sequencing four patients in one MiSeq run. However despite this, the proportion of target regions meeting either a 20X or 30X threshold was not significantly different (Figure 3.3.3). A cost analysis between these two options revealed that one large HiSeq run was more than £100 cheaper per patient than analysing many small cohorts of patients on the MiSeq platform (Table 3.3.3) and indeed within the diagnostic setting use of the HiSeq is preferential due to less machine-time being required. Finally, the analysis of increased number of patients in the same batch increases the power of the "Lookup" filter as the chance that sequencing artefacts and common polymorphisms will be identified in more than one patient and thus discarded is increased, leading to fewer candidate genes requiring further scrutiny (Section 3.3.2). Therefore, sequencing 24 patients per HiSeq run was determined to be superior for gene panel analysis.

Table 3.3.3: Comparison of the costs required to prepare and sequence four patient samples on the MiSeq platform and 24 patient samples on the HiSeq platform.

Reagent	Cost of MiSeq run multiplexing four patients (cost per patient)	Cost of HiSeq run multiplexing 24 patients (cost per patient)
MiSeq Reagent Kit v2 (300 cycles)	£165	-
HiSeq Rapid SBS Kit v2 (50 cycles x 3)	-	£41
HaloPlex 48-patient Custom Kit	£270	£270
Other required reagents	£10	£10
Total per patient	£445	£321

3.3.3 *Panel validation using patients with a known diagnosis and methods for the identification of insertions and deletions*

In order to validate the sensitivity and specificity of the gene panel methodology, fourteen patients with a known genetic diagnosis were recruited to this study and analysed in a blinded manner. Using the panel and pipeline outlined in Section 3.3.1, we were able to identify 71% (15/21) of the pathogenic sequence variants previously identified in these patients including seven heterozygous missense, four homozygous missense, one homozygous nonsense, two heterozygous splice site mutations and one heterozygous combined insertion/deletion event. This still left 29% of mutations which were not detected; these included four deletions ranging in size from 2 bp to approximately 6 kb, one of which was in a homozygous state, one heterozygous combined insertion/deletion event and one homozygous missense change (Table 3.3.4).

Table 3.3.4: Patients analysed using the gene panel who had a prior known genetic diagnosis. -, not applicable due to SIFT and PolyPhen-2 only predicting the functional consequence of missense mutations; *, reported at a low minor allele frequency in Ensembl; [1], Identified using gene panel analysis, in-house CNV analysis or not identified.

Patient	Gender	Phenotype	Disorder	Gene	Nucleotide change	Amino acid change	Identified[1]	SIFT	PolyPhen-2	Reference
1	Male	Hyperammonaemia	Arginosuccinic aciduria	*ASL*	c.925G>A	p.Gly309Arg	Panel	Deleterious	Benign	Linnebank *et al.* (2002)
					c.919-2A>G	Splicing errors	Panel	-	-	Novel
2	Female	Lysosomal storage disorder	Hurler syndrome	*IDUA*	c.1469T>C; c.1469T>C	p.Leu490Pro; p.Leu490Pro	Panel	Tolerated	Benign	Bach *et al.* (1993)
3	Female	Epilepsy	Pyridoxine-dependent epilepsy	*ALDH7A1*	c.950C>T	p.Ser317Leu	Panel	Deleterious	Possibly damaging	Mills *et al.* (2010)
					∼6kb deletion encompassing exons 14-17	∼6kb deletion encompassing exons 14-17	CNV	-	-	Mills *et al.* (2010)
4	Male	Epilepsy	Pyridox(am)ine phosphate oxidase deficiency	*PNPO*	c.98A>T	p.Asp33Val	Panel	Deleterious	Possibly damaging	Schmitt *et al.* (2010)
					c.264-21_264-1delinsC	Splicing errors	CNV	-	-	Mills *et al.* (2014)
5	Male	Cholestasis	Bile acid synthesis defect	*CYP7B1*	c.1249C>T; c.1249C>T	p.Arg417Cys; p.Arg417Cys	Panel	Deleterious	Probably damaging	Goizet *et al.* (2009)
6	Female	Dystonia	Hypermanganesemia with dystonia, polycythemia and cirrhosis	*SLC30A10*	c.292_402del; c.292_402del	p.Val98_Phe134del; p.Val98_Phe134del	No	-	-	Tuschl *et al.* (2012)
7	Unknown	Mitochondrial disease	Leigh syndrome due to COX deficiency	*SURF1*	c.312_320del10insAT;	-	Panel	-	-	Wedatilake *et al.* (2013)
					c.751+5G>A	Splicing errors	Panel	-	-	

8	Unknown	Mitochondrial disease	Combined oxidative phosphorylation deficiency 8	*AARS2*	c.2033G>A	p.Arg678Gln	Panel	Tolerated	Benign	Novel*
					c.1195A>C	p.Asn399His	Panel	Deleterious	Benign	Novel*
9	Male	Mitochondrial disease	Mitochondrial DNA depletion syndrome 8A/B	*RRM2B*	c.165G>A	p.Met55Ile	Panel	Deleterious	Possibly damaging	Novel
					deletion of exons 4-6	deletion of exons 4-6	CNV	-	-	Novel
10	Female	Mitochondrial disease	Mitochondrial complex I deficiency	*NDUFS2*	c.875T>C	p.Met292Thr	Panel	Deleterious	Possibly damaging	Tuppen *et al.* (2010b)
					c.840_842del	p.Glu280-281del	CNV	-	-	Novel
11	Unknown	Mitochondrial disease	Leigh syndrome	*BCS1L*	c.385G>A; c.385G>A	p.Gly129Arg; p.Gly129Arg	Panel	Deleterious	Possibly damaging	Tuppen *et al.* (2010a)
12	Female	Mitochondrial disease	Thiamine metabolism dysfunction syndrome 2	*SLC19A3*	c.517A>G; c.517A>G	p.Asn173Asp; p.Asn173Asp	Panel	Tolerated	Probably damaging	Fassone *et al.* (2013)
13	Male	Folate metabolism defect	Folate transporter defect	*SLC46A1*	c.198C>A; c.198C>A	p.Cys66*; p.Cys66*	Panel	-	-	Novel
14	Female	Riboflavin transporter defect	Brown-Vialetto-Van Laere syndrome	*SLC52A2*	c.916G>A; c.916G>A	p.Gly306Arg; p.Gly306Arg	No	Deleterious	Probably damaging	Johnson *et al.* (2012)

A limitation of gene panel methodology and next-generation sequencing (NGS) is its inability to easily detect insertions or deletions (indels) larger than approximately ten base pairs or if the indel spans more than one short read. Multiplex ligation-dependent probe amplification (MLPA) or array comparative genome hybridization (aCGH) are often requested alongside NGS for these purposes. However, there are a number of ways in which these structural variations can be detected from NGS data, including: *de novo* assembly, read splitting and inconsistencies of insert sizes in pair-end mapping (Ye *et al.*, 2009). Although *de novo* assembly is thought to be the most accurate method, it is extremely problematic to apply to the human genome due to the abundance and size of repetitive regions. Other algorithmic methods exist to identify the break points of large deletions from short paired-end reads, such as Pindel (Ye *et al.*, 2009). However, these types of methods have not been implemented or validated in the clinical setting and would only work if the breakpoints were covered by the gene panel design.

Deletions can also be detected by the analysis of read depth as this parameter is proportional to the underlying genomic copy number. Read depth analysis is particularly effective for whole exome or gene panel data because it does not rely on sequencing into or near the breakpoints. These approaches compare the number of reads mapping to a certain genomic region with the expected number under a certain statistical model which is specific to each algorithm. Similarly to aCGH methodology, any deviations from these predictions indicate a possible copy number variation (CNV). Due to the extensive variability in capture efficiency across exons, greater CNV calling accuracy is achieved when the test sample is compared to a group of reference samples sequenced using the same methodology and in the same batch. Multiple tools available online, typically in the form of R packages, use different statistical models to calculate read depth ratios. Both CNV-seq (Xie and Tammi, 2009) and ExomeCNV (Sathirapongsasuti *et al.*, 2011) assume a normal distribution of read count ratios. cn.MOPS assumes that read counts are distributed according to a mixture of Poisson distributions (Klambauer *et al.*, 2012) and ExomeDepth uses a beta-binomial model (Plagnol *et al.*, 2012). There are three main advantages in estimating copy number variation using these packages compared to aCGH. Firstly, since the depth of coverage shares a linear relationship with genomic copy number, the ability to recognise CNVs is more accurate when the read depth at a particular position is high. Secondly, certain methods allow the precise determination of the breakpoints because they do not rely on predefined probes. Thirdly, because aCGH and MLPA use predefined probes, they cannot detect alleles within patients that are incompatible with probe binding. Nevertheless, aCGH and MLPA are still considered the "gold standard" methodologies for the detection of CNVs. This is due to the sample preparation protocols being less complex and labour-intensive, a consensus

analytical pipeline being established across clinical laboratories and a reduction in the number of false-positives when compared to NGS.

The ability of both ExomeDepth and cn.MOPS to detect the five copy number variants not identified by the standard pipeline were tested. All raw data was aligned to the hg19 reference genome and exon locations within our genes of interest were defined. ExomeDepth analysis was carried out by Dr Chris Boustred, GOSH using the R package (www.cran.r-project.org/web/packages/ExomeDepth/index.html) essentially as described by Plagnol *et al.* (2012) and documentation available at the Comprehensive R Archive Network. None of the five variants were identified. Analysis using cn.MOPS (www.bioconductor.org/packages/release/bioc/html/cn.mops.html) was similarly performed according to Klambauer *et al.* (2012) and documentation available online. Patient 3, who has pyridoxine-dependent epilepsy, harbours a heterozygous deletion encompassing exons 14 - 17 of *ALDH7A1* that had been confirmed by aCGH prior to this study commencing. cn.MOPS analysis called a heterozygous deletion on chromosome 5 (chr5:125882015-125885985). This region encompasses the majority of exon 15 and 17, and the entirety of exon 16 of *ALDH7A1*. Although a deletion was called in the correct region, the number and location of the deleted exons did not correspond to those identified by aCGH. The remaining four variants were not identified.

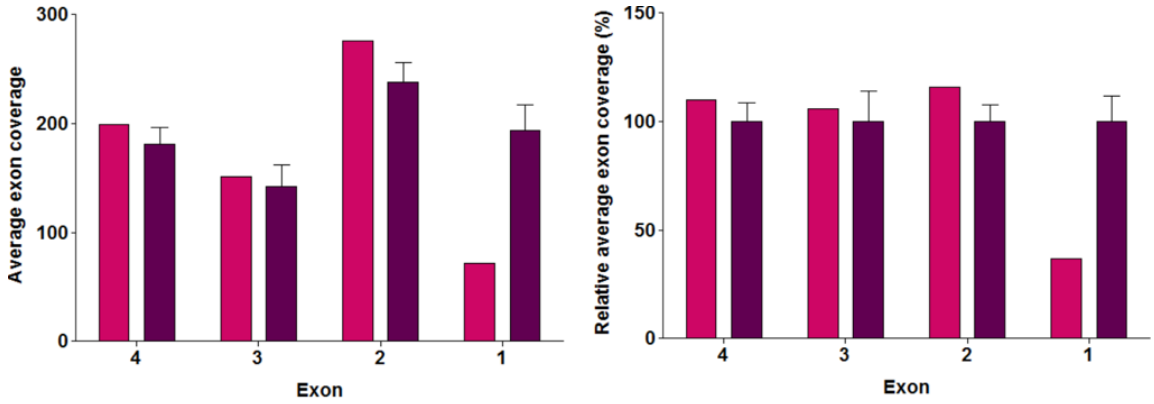
One of the main limitations of tools such as those described above is the necessity for a reference set of at least six samples to have been processed with the test sample using the same capture methodology. Their power to detect CNVs is also largely affected by sample to sample variability caused during target capture or sequencing (Plagnol *et al.*, 2012). In addition, these tools are best suited to the detection of CNVs whose size is greater than one exon as they are only called if there is approximately a 50% reduction in expected reads i.e. a heterozygous exonic deletion, or a 50% increase in expected reads i.e. an exonic duplication.

Read depth can also be assessed manually by using the Integrative Genomics Viewer (IGV) (Robinson *et al.*, 2011) to interrogate the BAM file in the genomic region of a strong candidate gene. In the case of a homozygous deletion one would expect to see a gap in coverage or half the number of amplicons in a region in the case of a heterozygous deletion. Although this method has a lower specificity, the sensitivity is high as small increases or decreases in coverage can be observed and quantified. However, this is generally not feasible when HaloPlex capture has been used because amplicons are generated by restriction enzyme digestion as opposed to random mechanical shearing and thus HaloPlex results in many reads covering identical or overlapping genomic regions. As part of this study a method was therefore developed for identifying CNVs based on comparison of the average coverage of each exon of a candidate gene across multiple samples. Similarly to the algorithmic methods described above, the method also requires a cohort

of reference samples with the use of higher numbers of samples affording greater CNV calling accuracy. However, unlike the methods already described, these reference samples can be taken from any batch sequenced using the same Illumina platform using the same capture methodology. As the number of reads (and therefore the coverage) is inversely proportional to the number of samples sequenced per lane, correcting for the sequencing capacity allocated to each sample allows for direct comparisons between them. Firstly, candidate genes were identified using the pipeline described in Section 3.3.1. The average read depth of each exon of the candidate gene in the patient sample and multiple (>10) unrelated patients was then extracted. Read depth parameters were corrected for the total sequencing capacity allocated to each sample (i.e. if four samples were analysed in one MiSeq run and each sample used 25%, 20%, 27% and 28% of the MiSeq capacity, then the average read depth would be divided by 0.25, 0.2, 0.27 and 0.28, respectively). The data for all the "control" samples was then averaged and compared to the data from the patient of interest. The data was also expressed as a percentage for ease of viewing and analysis. A 100% reduction indicated a homozygous deletion; whilst a 50% reduction indicated a heterozygous deletion of a whole of an exon. A deletion of a large region (i.e. < one exon) was more difficult to interpret but corresponded to a reduction in coverage.

Using this method we were able to detect the presence of all heterozygous deletions and insertion/deletion events in the patients with a prior genetic diagnosis which could then be accurately determined by Sanger sequencing of the affected exon. The addition of this CNV analysis method to the analysis pipeline increased the proportion of pathogenic sequence variants identified to 90% (19/21). The main disadvantage of this method is its inability to detect homozygous or compound heterozygous CNVs, as read depth analysis would only be carried out in a targeted manner once a potentially pathogenic variant has been identified in a good candidate gene. Indeed, one of the mutations that was not detected was the homozygous deletion of 111 bp within the *SLC30A10* gene in patient 6. However, when the coverage of this gene was subsequently examined, the deletions was clearly evident (Figure 3.3.4). The other was a homozygous missense variant in the *SLC52A2* gene; this was not identified because the gene was not included in the 614-gene design. Indeed, mutations in *SLC52A2* were only identified as causing a neurometabolic disorder in late-2012 (Haack *et al.*, 2012). Encouragingly, no other candidate genes were identified in this patient, illustrating the stringency of our methodology for identifying potentially pathogenic variants. Many of the mutations identified in these patients have not been reported in the literature previously (Table 3.3.4).

Figure 3.3.4: Deletion analysis of *SLC30A10* in patient 6 showing a reduction in coverage of exon 1.



3.3.4 *Identification of potentially pathogenic variants in patients without a prior genetic diagnosis*

Following the validation of the gene panel sequencing and analysis methodology, 30 patients with neurometabolic disease and lacking a specific genetic diagnosis were examined. The clinical, biochemical and genotypic data of these patients is summarised in Tables 3.3.5 3.3.6, 3.3.7 and 3.3.10. Ages ranged from 1 to 20 years (mean 7.2 years, median 6 years). Of the 30 patients, only 9 had abnormal biochemical parameters pointing towards a specific underlying IEM. Despite this, all patients had undergone extensive previous testing including multiple standard and specialised biochemical blood tests, invasive investigations such as skin biopsies, muscle biopsies and lumbar punctures, as well as targeted candidate gene testing. Gene panel sequencing identified a definitive or likely genetic diagnosis that could at least partially explain the clinical phenotype in 53.3% of patients (16/30).

3.3.5 *Analysis of patients with strong biochemical indicators*

Of the 9 patients (patients 15 - 22 and 31, Tables 3.3.5 and 3.3.6) with previous biochemical testing indicating a specific diagnosis, identification of pathogenic variants was possible for 88.9% (8/9). Parental DNA to check segregation or whether mutations were in *cis* or *trans* within the patient was not available for these cases. Pathogenic variants were not identified in patient 31 in whom biochemical testing was suggestive of a diagnosis of hyperprolinaemia type II. Candidate gene sequencing subsequently revealed that this patient was homozygous for a complex insertion/deletion event in the *ALDH4A1* gene (c.411_424delinsCGGCCC; p.Pro138Glyfs*13).

In the remaining cases, one heterozygous insertion, six heterozygous and four homozygous missense mutations were identified. The genetic diagnoses were consistent with the biochemical findings in all cases. In the majority of cases, two pathogenic variants were identified in each candidate gene. The three patients in which this was not the case are described below.

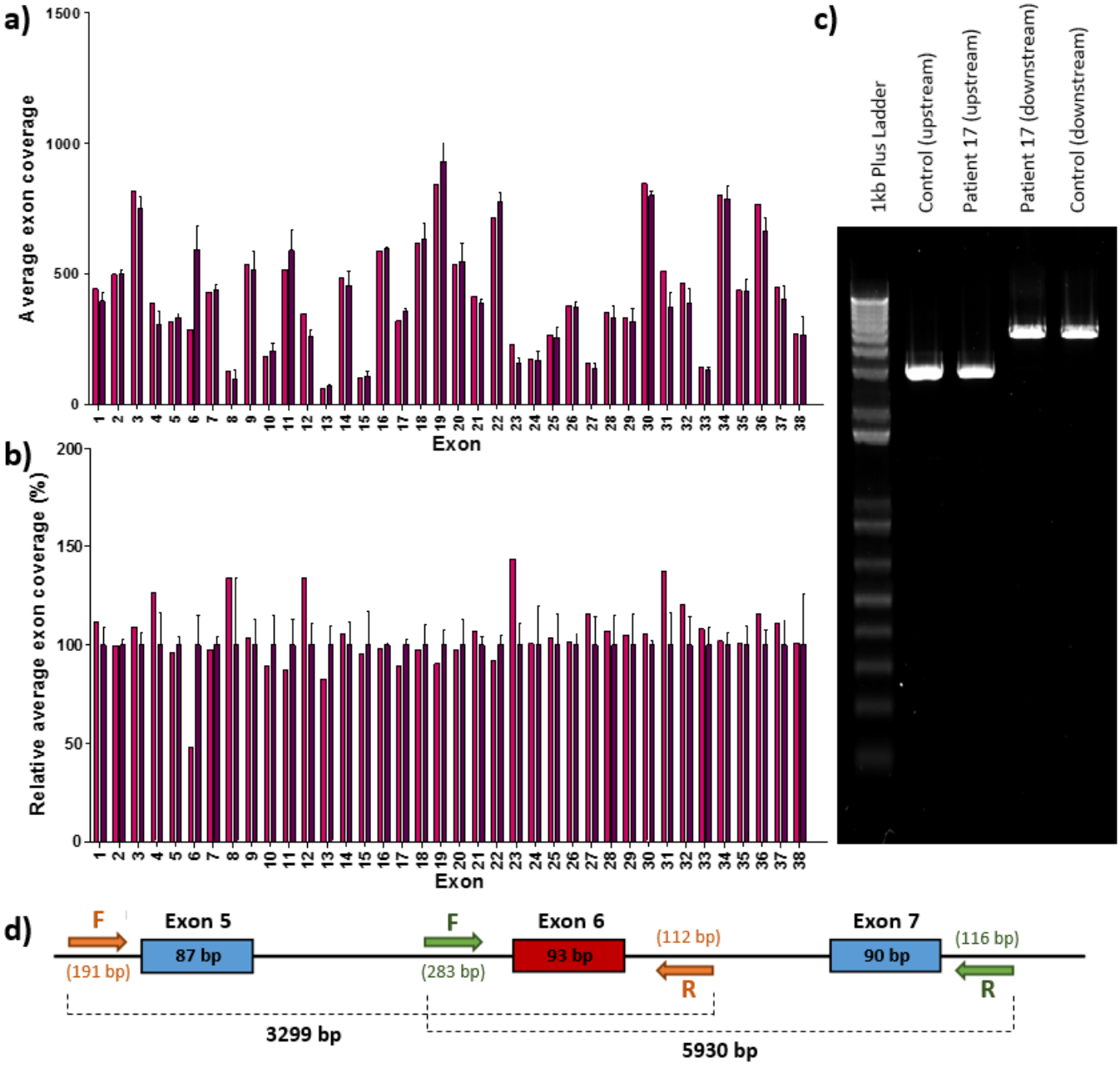
Three heterozygous variants in the *HLCS* gene were identified in patient 15; the first is a novel insertion of one base causing a frameshift and premature stop codon (p.Val512Cysfs*65) and the remaining two are novel missense variants. Both p.Pro709Leu and p.Val641Met are predicted to be deleterious and probably damaging by SIFT and PolyPhen-2, respectively. However, p.Val641Met is reported at a minor allele frequency of 0.48%, with four individuals homozygous for this variant reported in the ExAC database. Indeed, when classifying the variants according to American College of Medical Genetics and Genomics (AMCG) guidelines, this variant is predicted to be of uncertain significance (Table 3.3.7).

A novel, likely pathogenic heterozygous variant (p.Val151Met) in the *UMPS* gene was identified in patient 16. This patient was brought to the attention of the metabolic team as his older sibling was born prematurely and died due to hyperammonaemia. There were no clinical concerns regarding the patient, however urinary organic acid analysis showed moderately elevated orotic acid levels between 9 - 48 μmol/mmol creatinine (ref: 0 - 5). Despite 95% of the *UMPS* gene being covered at a read depth of $>$ 30X, a second variant was not identified. Heterozygous carriers of mutations in *UMPS* have been reported previously to have urinary orotic acid levels between 2 and 7 times higher than that of controls (Suchi *et al.*, 1997). In contrast, levels in affected individuals are in the order of 200 times higher than the reference range (Grohmann *et al.*, 2015). Thus, the patient's phenotypic features would be consistent with carrier status for the *UMPS* mutation.

Similarly, only one variant was identified in patient 17. A biochemical diagnosis of carbamoyl phosphate synthetase I deficiency had been made on the basis of a severely reduced carbamoyl phosphate synthetase activity in a liver biopsy. In concordance with this, a known pathogenic heterozygous missense change (p.His337Arg) was identified in exon 10 of *CPS1* (Aoshima *et al.*, 2001). A second pathogenic mutation was not identified within *CPS1*; however, only 98.5% of the gene was covered with a read depth $>$ 30X and exon 21 had regions that were only covered at a read depth of 3X. Therefore standard PCR of this exon was performed to examine the regions that were insufficiently covered but no mutations were detected. Given that a second variant was not identified, deletion screening was carried out which showed reduced coverage of exon 6 in the patient, indicating a possible deletion (Figure 3.3.5a+b). Standard PCR and Sanger sequencing of the region using intronic primers revealed a sequence identical to the reference, indicating that if a deletion was present the breakpoints were not exonic. Long-range PCR was undertaken using

Phusion DNA polymerase and primers located upstream of exon 5 and downstream of exon 7. Whilst this revealed the same sized DNA fragments in both the patient and control, the results were inconclusive as a small deletion could not be ruled out Figure 3.3.5c. Unfortunately, RNA was not available for this patient so it was not possible to sequence the cDNA to evaluate the entire gene for single nucleotide variants or gross structural variants.

Figure 3.3.5: Deletion analysis of *CPS1* in patient 17 showing a reduction in coverage of exon 6. **(a)** average exon coverage corrected for sequencing capacity and **(b)** expressed as a percentage. Pink, patient; purple, controls (n=10). **(c)** Long-range PCR of both exons and introns adjacent to exon 6 as described in Section 3.3.14. **(d)** Schematic illustrating the position of the long-range PCR primers used for amplification. The distance of each primer from the beginning/end of each exon is shown in brackets.



3.3.6 *Identification of likely pathogenic variants in patients without an indicative biochemical profile*

Of the remaining 22 patients without a biochemical marker pointing towards a specific diagnosis, a molecular genetic diagnosis that could at least partially explain the clinical-biochemical phenotype was found for 36.4% of these (8/22) (patients 23 - 30, Table 3.3.6). In these 8 patients 11 variants were identified; three of which are known to be pathogenic. The variants included three heterozygous and four homozygous missense mutations, one nonsense mutation, one splice site mutation, one duplication leading to a frameshift and a premature stop codon and one deletion/insertion event leading to a frameshift and a premature stop codon. All patients presented in childhood and had neurological and biochemical abnormalities suggestive of a neurometabolic disorder. Two variants were identified in each candidate gene. All mutations were confirmed by Sanger sequencing in the proband and segregation within family members was carried out where possible. Classification of the variants according to ACMG criteria is given in Table 3.3.7. Detailed clinical case reports and contributions to the expansion of genotypes and phenotypes associated with each disorder are described below.

Table 3.3.5: Clinical details of patients without a prior genetic diagnosis who had pathogenic variants identified through gene panel sequencing. Details of identified mutations can be found in Table 3.3.6. 5-HIAA, 5-hydroxyindoleacetic acid; 5-MTHF, 5-methyltetrahydrofolate; BH_4, tetrahydrobiopterin; BMI, body mass index; CK, creatine kinase; DHA, docosahexaenoic acid; DHCA, 3α,7α-dihydroxycholestanoic acid; GSD, glycogen storage disease; HVA, homovanillic acid; IgA, immunoglobulin A; IgG, immunoglobulin G; IgM, immunoglobulin M; THCA, trihydroxycholestanoic acid; VSD, ventricular septal defect.

Patient	Age of patient	Gender	Primary clinical phenotype	Other phenotypic features	Relevant specialist investigations	Diagnosis	Gene
15	4	Male	Organic acidaemia, no further information available.	Short stature, asthma, development unremarkable.	Elevated 3-methylcrotonylglycine (112 μmol/mmol creatinine), methylcitrate (66 μmol/mmol creatinine) and 3-hydroxyisovalerate in urine. Normal biotinidase activity	Holocarboxylase synthetase deficiency	*HLCS*
16	1	Male	Sibling born prematurely and passed away due to hyperammonaemia. No clinical concerns.	N/A	Moderately elevated urinary orotic acid between 9-48 μmol/mmol creatinine (ref: 0-5)	Orotic aciduria	*UMPS*
17	15	Male	At 3.5 years: lethargy, vomiting, alkalosis and hyperammonaemia. Learning and behavioural difficulties	Day 2 of life: lethargy and irritability, presumed sepsis but negative cultures. Ammonia not measured.	Raised glutamine in plasma. Low carbomyl-phosphate synthase activity in liver biopsy (0.15 mmol/hr/mg protein)	Carbomoyl-phosphate synthetase I deficiency	*CPS1*
18	17	Female	At 3 years: short stature, high BMI, distended abdomen.	Hepatomegaly, no documented learning difficulties.	Low GSD debranching enzyme activity of 0.07 μmol/min/g protein (ref: 0.3-3.0)	Glycogen storage disease type III	*AGL*
19	9	Female	Galactosaemia picked up through newborn screening and treated early	Normal development apart from mild difficulties in mathematics	Low gal-1-P-uridyltransferase activity of 1.8 μmol/hr/g Hb (ref: 18-40), elevated galactose-1-phosphate of 2.4 μmol/g Hb (ref: <0.1)	Galactosaemia	*GALT*

20	5	Male	Global delay, one of monozygotic twins.	Hypotonia, brachycephaly, long face. Brain MRI: delayed myelination, lack of white matter bulk.	Elevated plasma lysine ranging between 439-449 μmol/L (ref: 100-300), elevated CSF lysine of 67 μmol/L (ref: 10-32)	Hyperlysinaemia	*AASS*
21	5	Male	Global delay, well-controlled epilepsy, one of monozygotic twins.	Hypotonia, brachycephaly, long face.	Elevated plasma lysine ranging between 440-780 μmol/L (ref: 100-300), elevated CSF lysine of 92 μmol/L (ref: 10-32)	Hyperlysinaemia	*AASS*
22	20	Female	Sensorineural hearing loss, ataxia, neurological regression, similarly affected sister.	Scoliosis, constipation. Brain MRI: leukodystrophy.	Bile acid analysis and skin fibroblast studies suggestive of a peroxisomal biogenesis defect. Elevated C26:C22 ratio of 0.038 (ref: 0 - 0.026), elevated phytanate of 20.21 (ref: 0 - 15), elevated pristanate of 30.11 (ref: 0 - 2), low DHA of 58 (ref: 75 - 180). Presence of THCA, DHCA and C29 dicarboxylic acid on bile acid analysis. Absence of peroxisomes by immunofluorescence microscopy.	Peroxisome biogenesis disorder	*PEX6*
23	11	Female	Developmental delay, ataxia, horizontal nystagmus.	Microcephaly, retinal dystrophy.	CSF: low 5-MTHF and high HVA and BH_4. Blood: Elevated prolactin, alanine, intermittently high CK and plasma lactate. Muscle: Normal respiratory chain enzymes	Muscular dystrophy-dystroglycanopathy	*POMGNT1*
24	6	Male	Neonatal jitteriness, developmental delay, autism.	Joint hypermobility	Persistent methylmalonic and malonic aciduria	Combined malonic and methylmalonic aciduria	*ACSF3*
25	9	Male	Congenital ataxia, diplegia, drop attacks with no obvious EEG correlate.	Brain MRI: Abnormal signal in caudate and lentiform nuclei bilaterally.	Plasma: mildly raised alanine and normal lactate. CSF: low 5-HIAA	Spinocerebellar ataxia 28 / Autosomal recessive spastic ataxia 5	*AFG3L2*

26	4	Male	Developmental delay, subsequent regression with progressive dyskinetic movement disorder and dysphagia.	Sensorineural deafness. Brain MRI: High T2 signal in the basal ganglia and cerebellar atrophy.	Raised 3-methylglutaconic acid with normal 3-methylglutaric acid levels	3-methylglutaconic aciduria with deafness, encephalopathy and Leigh-like syndrome	*SERAC1*
27	2	Male	Global severe developmental delay, tonic seizures.	Multi-organ malformations including VSD and Hirschprung's disease, dysmorphism. Brain MRI: Dandy-Walker malformation and reduced white matter bulk.	Recurrent hypoglycaemia, hypogammaglobulinaemia, hyperphosphatasia	Hyperphosphatasia with mental retardation syndrome 3	*PGAP2*
28	6	Male	Developmental delay, microcephaly, lower limb hyper-reflexia.	Brain MRI: lack of white matter bulk.	Abnormal transferrin isoelectric focusing (type I pattern)	Late infantile neuronal ceroid lipofuscinosis	*TPP1*
						Hereditary fructose intolerance	*ALDOB*
29	6	Male	Global developmental delay, sensorineural hearing loss.	Neonatal acute liver failure which subsequently resolved. Recurrent hypoglycaemia and recurrent infections.	Abnormal isoelectric focusing (type I pattern), normal phosphomannomutase and phosphomannisomerase. Low IgA/IgM, normal IgG and lymphocyte subsets	Galactose epimerase deficiency	*GALE*
30	2	Male	Microcephaly, developmental delay.	Dysplastic kidneys.	Neonatal lactic acidosis, high plasma triglycerides, elevated urine thymidine and uracil, low plasma urate and detectable thymine	Dihydropyrimidinase deficiency	*DPYS*

Table 3.3.6: Pathogenic or likely pathogenic variants identified through gene panel sequencing in patients without a prior genetic diagnosis. -, not applicable due to SIFT and PolyPhen-2 only predicting the functional consequence of missense mutations; *, siblings; [1], second mutation not identified; [2], parental DNA was unavailable but Sanger sequencing identified the same homozygous mutation in a similarly affected sister; [3], parental DNA was unavailable but Sanger sequencing identified the same homozygous pathogenic mutation in a brother who also had an abnormal type I transferrin isoelectric focussing pattern.

Patient	Gene	Nucleotide change	Amino acid change	Segregation confirmed	SIFT	PolyPhen-2	Minor allele frequency (ExAC)	Reference	Phenotype fully explained
15	HLCS	c.2126C>T	p.Pro709Leu	No	Damaging	Probably damaging	0%	Novel	Yes
		c.1921G>A	p.Val641Met		Damaging	Possibly damaging	0.48%	Novel	
		c.1533dupT	p.Val512Cysfs*65		-	-	0.00082%	Novel	
16[1]	*UMPS*	c.451G>A	p.Val151Met	No	Damaging	Probably damaging	0.00082%	Novel	Yes
17[1]	*CPS1*	c.1010A>G	p.His337Arg	No	Damaging	Probably damaging	0%	Aoshima *et al.* (2001)	No
18	*AGL*	c.2590C>T	p.Arg864*	No	-	-	0.0083%	Shen *et al.* (1996)	Yes
		c.2590C>T	p.Arg864*						
19	*GALT*	c.563A>G	p.Gln188Arg	No	Damaging	Probably damaging	0.13%	Reichardt *et al.* (1991)	Yes
		c.584T>C	p.Leu195Pro		Damaging	Benign		Reichardt *et al.* (1992)	
20*	*AASS*	c.965G>A	p.Arg322His	No	Tolerated	Probably damaging	0%	Novel	No
		c.965G>A	p.Arg322His						
21*	*AASS*	c.965G>A	p.Arg322His	No	Tolerated	Probably damaging	0%	Novel	No
		c.965G>A	p.Arg322His						
22	*PEX6*	c.2734G>A	p.Ala912Thr	No[2]	Damaging	Probably damaging	0.0025%	Novel	Yes
		c.2743G>A	p.Ala912Thr						

23	*POMGNT1*	c.373C>G	p.Arg125Gly	No	Tolerated	Benign	0.0017%	Novel	Yes
		c.1539+1G>A	Splicing errors		-	-	0.092%	Yoshida *et al.* (2001)	
24	*ACSF3*	c.1453A>C	p.Ser485Arg	Yes	Damaging	Probably damaging	0%	Novel	Yes
		c.1453A>C	p.Ser485Arg						
25	*AFG3L2*	c.1067T>G	p.Leu356Arg	No	Damaging	Probably damaging	0%	Novel	Yes
		c.1067T>G	p.Leu356Arg						
26	*SERAC1*	c.1850delinsCA	p.Ile617Thrfs*6	Yes	-	-	0%	Novel	Yes
		c.1850delinsCA	p.Ile617Thrfs*6						
27	*PGAP2*	c.560C>T	p.Ala187Val	Yes	Damaging	Probably damaging	0.0025%	Novel	Yes
		c.560C>T	p.Ala187Val						
28	*TPP1*	c.887G>A	p.Gly296Asp	No	Damaging	Probably damaging	0.0032%	Novel	Yes
		c.887G>A	p.Gly296Asp						
	ALDOB	c.178C>T	p.Arg60*	No[3]	-	-	0.0091%	Santer *et al.* (2005)	Yes
		c.178C>T	p.Arg60*						
29	*GALE*	c.280G>A	p.Val94Met	No	Damaging	Probably damaging	0.00083%	Wohlers *et al.* (1999)	Yes
		c.284G>A	p.Gly95Asp		Damaging	Probably damaging	0%	Novel	
30	*DPYS*	c.144_151dupGCTGCGGG	p.Val51Glyfs*50	No	-	-	0%	Novel	No
		c.144_151dupGCTGCGGG	p.Val51Glyfs*50						

Table 3.3.7: Classification of the variants identified in the 30 patients without a prior genetic diagnosis according to the recommendations of the American College of Medical Genetics and Genomics (AMCG) guidelines. Definitions of the criteria used to classify the strength of the evidence to support pathogenicity are shown in Table 3.3.8.

Gene	Variant	Very strong evidence	Strong evidence	Moderate evidence	Supporting evidence	Final classification
HLCS	c.2126C>T; p.Pro709Leu			PM2	PP2, PP3, PP4	Uncertain significance
HLCS	c.1921G>A; p.Val641Met			PM1, PM2	PP2, PP3, PP4	Likely pathogenic
HLCS	c.1533dupT; p.Val512Cysfs*65	PVS1		PM1, PM2	PP2, PP4	Pathogenic
UMPS	c.451G>A; p.Val151Met			PM1, PM2	PP2, PP3, PP4	Likely pathogenic
CPS1	c.1010A>G; p.His337Arg		PS1, PS3	PM1, PM2	PP2, PP3, PP4	Pathogenic
AGL	c.2590C>T; p.Arg864*		PS1, PS3	PM2	PP2, PP4	Pathogenic
GALT	c.563A>G; p.Gln188Arg		PS1, PS3	PM1, PM2, PM3	PP2, PP3, PP4	Pathogenic
GALT	c.584T>C; p.Leu195Pro		PS1, PS3	PM1, PM2, PM3	PP2, PP3, PP4	Pathogenic
AASS	c.965G>A; p.Arg322His			PM2	PP1, PP2, PP3, PP4	Likely pathogenic
POMGNT1	c.373C>G; p.Arg125Gly			PM2, PM3	PP2, PP3, PP4	Likely pathogenic
POMGNT1	c.1539+1G>A; Splicing errors	PVS1	PS3	PM2	PP2, PP4	Pathogenic
DPYS	c.144_151dupGCTGCGGG; p.Val51Glyfs*50	PVS1		PM2	PP2, PP4	Pathogenic
ACSF3	c.1453A>C; p.Ser485Arg			PM1, PM2	PP1, PP2, PP3, PP4	Likely pathogenic
PEX6	c.2734G>A; p.Ala912Thr			PM1, PM2	PP1, PP2, PP3, PP4	Likely pathogenic

AFG3L2	c.1067T>G; p.Leu356Arg			PM1, PM2	PP1, PP2, PP3, PP4	Likely pathogenic
SERAC1	c.1850_1851delinsCA; p.Ile617Thrfs*6	PVS1		PM2	PP1, PP2, PP4	Pathogenic
PGAP2	c.560C>T; p.Ala187Val			PM2	PP1, PP2, PP3, PP4	Likely pathogenic
TPP1	c.887G>A; p.Gly296Asp			PM1, PM2	PP2, PP3, PP4	Likely pathogenic
ALDOB	c.178C>T; p.Arg60*	PVS1	PS3	PM2	PP1, PP2, PP3, PP4	Pathogenic
GALE	c.280G>A; p.Val94Met		PS1, PS3	PM1, PM2	PP2, PP3, PP4	Pathogenic
GALE	c.284G>A; p.Gly95Asp			PM1, PM2, PM3	PP2, PP3, PP4	Likely pathogenic
EARS2	c.760A>G; p.Thr254Ala			PM1, PM2	PP2, PP3, PP4	Likely pathogenic
IDH2	c.673G>A; p.Asp225Asn			PM2	PP3, PP4	Uncertain significance

Table 3.3.8: Criteria for classifying potentially pathogenic variants according to ACMG Guidelines. These criteria were applied to the variants identified in our patient cohort (Table 3.3.7).

Disorder Class	Number of genes
PVS1	Null variant (nonsense, frameshift, canonical ± 1 or 2 splice sites, initiation codon, single or multiexon deletion) in a gene where loss of function is a known mechanism of disease
PS1	Same amino acid change as a previously established pathogenic variant regardless of nucleotide change
PS3	Well-established *in vitro* or *in vivo* functional studies supportive of a damaging effect on the gene or gene product
PM1	Located in a mutational hot spot and/or critical and well-established functional domain without benign variation
PM2	Absent from controls (or at extremely low frequency if recessive) in Exome Sequencing Project, 1000 Genomes Project or Exome Aggregation Consortium
PM3	For recessive disorders, detected *in trans* with a pathogenic variant
PP1	Co-segregation with disease in multiple affected family members in a gene definitively known to cause disease
PP2	Missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease
PP3	Multiple lines of computational evidence support a deleterious effect on the gene or gene product (conservation, evolutionary, splicing impact, etc.)
PP4	Patient's phenotype or family history is highly specific for a disease with a single genetic aetiology

3.3.6.1 *Patient 23*

Patient 23 is an 11 year old female. She was born by normal vaginal delivery at 36 weeks gestation following spontaneous rupture of membranes with a birth weight of 1900g. She was nursed in special care for two weeks and discharged home at three weeks of age on formula feeds. At eight weeks of age she was admitted to hospital in Germany with pneumonia. She was noted to be hypotonic, displayed abnormal movements and required tube feeding. Investigations at this time showed elevated plasma and CSF lactate (7.0 and 2.7 nmol/L) and intermittently raised creatine kinase (CK). A muscle biopsy was performed but only histology was carried out at this time. She was given a working diagnosis of a "likely mitochondrial disorder" and started on various vitamin supplements including coenzyme Q10, riboflavin and carnitine. These were of no benefit and were gradually discontinued. The family moved to London and she was referred to the Metabolic Medicine Department at GOSH at four years of age. Her main presenting problems were poor appetite, developmental delay, frequent falls and limited exercise tolerance. Although she could only speak two words, she appeared to understand language and could indicate her needs by pointing. She had an ataxic gait and myopic astigmatism, left

esotropia with manifest horizontal nystagmus but normal ophthalmological electrodiagnostic testing. Further investigation undertaken at this age confirmed raised plasma and CSF lactate, whilst CK and white cell ubiquinone were normal. She was still considered to have a possible mitochondrial disorder and had annual monitoring of cardiac and renal function which remained normal. Her vision has improved slightly with time although nystagmus and squint remain present. A brain MRI at seven years of age was normal. At this age, a muscle biopsy was repeated and showed normal respiratory chain enzymology and normal histology. Chromosomal analysis was also normal. Cultured fibroblasts demonstrated normal fatty acid oxidation and pyruvate dehydrogenase, although just under the normal range, was considered normal. CSF neurotransmitter analysis identified low 5-methyltetrahydrofolate so calcium folinate therapy was commenced. Her condition has remained fairly static. Over time she continues to have good general health, her appetite has improved and she no longer requires tube feeding. Her growth is normal; she can walk and run although she is ataxic. Her appearance is somewhat coarse with hirsutism.

DNA analysis revealed compound heterozygous variants in the *POMGNT1* gene. Parental DNA was not available to determine the segregation of each variant. The first was a splice site variant (c.1539+1G>T) affecting the invariant GT donor site. This mutation has been reported previously to generate two aberrant mRNA transcripts; a read-through of intron 17 with a premature stop codon at position 484 and a skipping of exon 17 resulting in an in-frame deletion of 42 amino acids (Yoshida *et al.*, 2001). This mutation is found in 37/38 alleles from 14 families of Finnish origin and represents a common founder mutation in this population (Diesen *et al.*, 2004). The second variant identified was a missense change (p.Arg125Gly) which is predicted to be tolerated and benign by SIFT and PolyPhen-2, respectively. It is reported with a minor allele frequency of 0.0017%. Two other variants affecting this amino acid residue have been reported in the ExAC database, p.Arg125Gln and p.Arg125Trp. These have a minor allele frequency of 0.0025% and 0.0017%, respectively. Therefore in total, six individuals have been identified with heterozygous variants that alter the Arg125 amino acid in *POMGNT1* giving a combined minor allele frequency of 0.0058%. Whilst the presence of *POMGNT1* has not been reported in lower organisms, a multiple sequence alignment of higher organisms showed that Arg125 is conserved from humans to zebrafish.

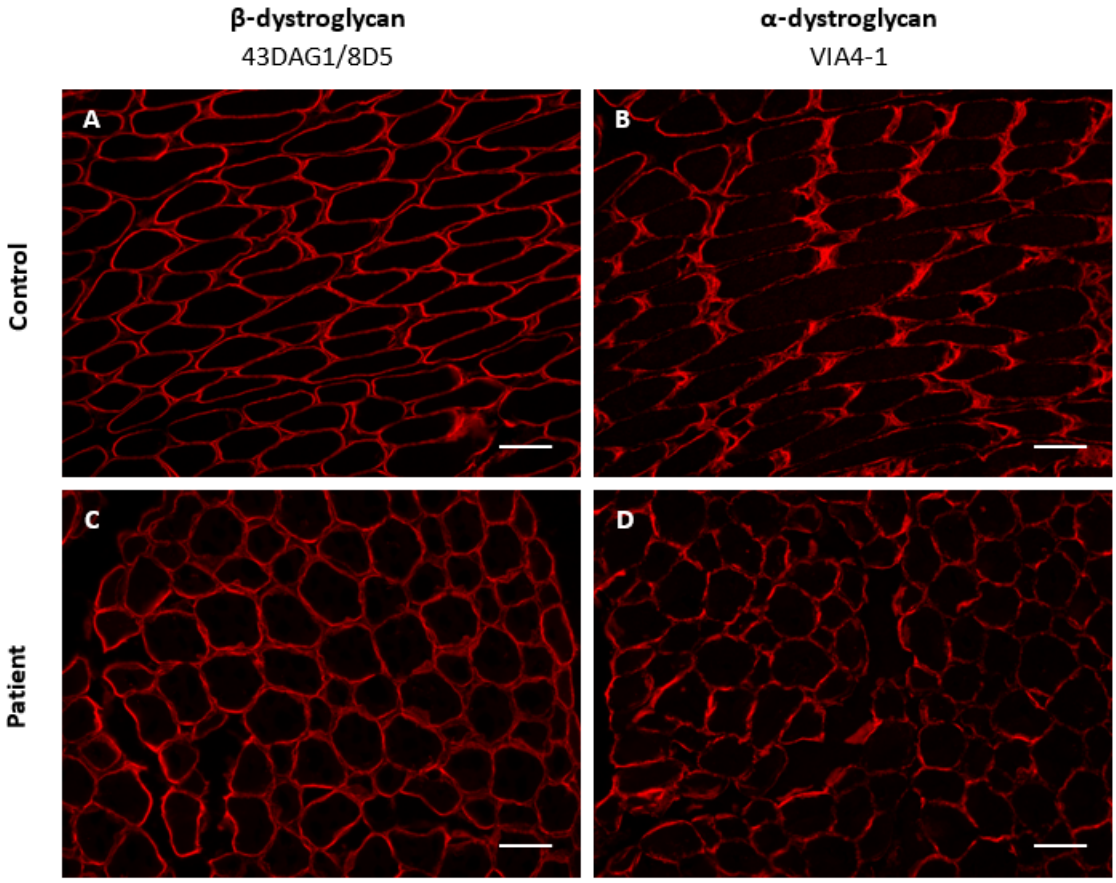
The dystroglycan complex, comprised of α- and β-dystroglycan, links the extracellular matrix to intracellular actin cables and provides structural integrity in muscle tissues. For correct α-dystroglycan function the protein must be O-glycosylated; POMGNT1 participates in O-mannosyl glycan synthesis by transferring N-acetylglucosamine to the O-linked mannose (Martin, 2007). Patients with mutations in this gene typically present with neonatal hypotonia, mental retardation,

moderate to severe muscle weakness, seizures, myopia, glaucoma and retinal hypoplasia (Yiş *et al.*, 2014). Brain MRI in these patients typically display abnormalities including cobblestone lissencephaly, ventriculomegaly, white matter changes, cerebellar cysts and brainstem hypoplasia. However, similarly to our patient, normal brain imaging has been reported on multiple occasions (Yiş *et al.*, 2014).

Given the glycosylation abnormalities present in these patients, immunofluorescence staining of the glycosylated epitope of α-dystroglycan in muscle is typically absent or reduced, alongside myopathic changes including the presence of small round muscle fibres of variable size. Therefore, upon the identification of the two *POMGNT1* variants in our patient, the muscle biopsy taken at seven years of age was re-evaluated by Dr Tom Jacques, Histopathology Department, GOSH. As identified at the time of intial evaluation, there was increased variation in fibre size ranging from 7 - 37 μm (normal: 31 - 39 μm) due to scattered small angular and rounded fibres, as well as prominent lipid accumulation for the age of the patient. All other histological and immunohistochemical investigations undertaken at the time were normal. In view of the potential diagnosis, immunohistochemistry and immunofluorescence staining of α-dystroglycan using the VIA4 antibody was completed (Dubowitz Neuromuscular Centre, ICH). Allowing for the artefactual vacuolation in the biopsy, no diagnostic loss of α-dystroglycan was reported (Figure 3.3.6). However, this does not exclude an abnormality of O-glycosylation, with patients having been reported in the literature with only very mild changes detected using this method (Clement *et al.*, 2008; Jimenez-Mallebrera *et al.*, 2009). Indeed, evidence suggests that the degree of hypoglycosylation does not consistently correlate with clinical severity (Jimenez-Mallebrera *et al.*, 2009).

Our patient's phenotype is very consistent with a mild α-dystroglycanopathy caused by mutations in *POMGNT1*. Up to 30% of patients are able to walk independently and similarly to our patient, speak some words (Yiş *et al.*, 2014). Indeed there has been a case report with an even milder presentation; she has normal intellect and attends university as well as having normal glycosylated α-dystroglycan staining. Whilst no clear genotype-phenotype correlations have been found in the α-dystroglcanopathies a phenotype with mental retardation with a normal MRI is not normally associated with mutations in *POMGNT1* but instead with mutations in *POMT1*, *POMT2* and *ISPD* which highlights the difficulty in detecting some of the α-dystroglycanopathies.

Figure 3.3.6: Skeletal muscle biopsy analysis in patient 23 and a healthy control. (A and C) Immunofluorescence staining of β-dystroglycan using the 43DAG1/8D5 antibody. **(B and D)** Immunofluorescence staining of α-dystroglycan using the VIA4-1 antibody against the glycosylated epitope. Scale bars: 50 μm.



3.3.6.2 *Patient 24*

Patient 24 is a six year old male and the third child born to consanguineous parents. The mother had previously suffered one still birth and there was a family history of two cousins living abroad who had seizures and developmental delay. The patient presented with delayed motor milestones in the first year of life, accompanied by autistic features and significant speech and language delay. He was seen in the Metabolic Medicine Department at GOSH at two years of age as routine investigations for developmental delay showed methylmalonic acid (MMA) in urine. He was not dysmorphic, had normal growth, normal brain MRI and had never had a metabolic decompensation. Further detailed metabolic investigations showed multiple abnormalities including: elevated urinary MMA concentration on two occasions over the course of four years ranging between 36 - 89 μmol/mmol creatinine (ref: 0 - 30), elevated plasma MMA of 2.85 μmol/L (ref: 0 - 0.29), low plasma homocysteine of 4 μmol/L (ref: 5 - 15), mild generalised aminoaciduria and elevated urinary N-acetyl-β-D-glucosaminidase of 76 unit/μmol.

A number of biochemical parameters were also found to lie within the normal range, including: methylcitrate/creatinine ratio, vitamin B_{12}, blood spot carnitine profile and CSF amino acids. Now at the age of six, he is autistic but is making slow developmental progress. Next-generation sequencing using the gene panel revealed a novel homozygous missense mutation (p.Ser485Arg) in *ACSF3*. Sanger sequencing confirmed that the variant segregated within the family with both parents being heterozygous for p.Ser485Arg. This variant is predicted to be deleterious and probably damaging by SIFT and PolyPhen-2, respectively. The variant lies with the second motif region of ACSF3. This region is conserved amongst the acyl-coA synthetase family of enzymes which function to activate fatty acids for intermediary metabolism and is predicted to be involved in conformational change and catalytic function (Hiltunen *et al.*, 2010). The mutated Ser485 residue is conserved in all other known ACSF3 family members from humans through to zebrafish and *Xenopus* (Figure 3.3.7).

Figure 3.3.7: Multiple sequence alignment of the ACSF3 protein across species. (*) positions with have a single fully conserved residue, (:) conservation between amino acids with strongly similar properties, (.) conservation between amino acids with weakly similar properties. The alignment was generated using Clustal Omega.

```
Human          DGQYWIRGRTSVDIIKTGGYKVSALEVEWHLLAHPSITDVAVIGVPDM
Gorilla        DGQYWIRGRTSVDIIKTGGYKVSALEVEWHLLAHPSITDVAVIGVPDM
Cow            DGCYWIRGRTSVDIIKSGGYKVSALEVERLLLAHPSITDVAVIGVPDM
Horse          DGRYWIRGRTSVDIIKSGGYKISALEVERLLLAHPSITDVAVIGVPDM
Cat            DGRYWIRGRTSVDIIKSGGYKISALEVERLLLAHPSITDVAVIGVPDM
Dog            DGMYWIRGRTSVDIIKSGGYKISALEVERLLLTHPSITDVAVIGVPDM
Mouse          DARYWIRGRTSVDIIKTGGYKVSALEIERHLLAHPSITDVAVIGVPDM
Cod            DGVYWIMGRTSVDIIKSGGYKISALDVERQLLAHPDITDVAVIGVQDQ
Zebrafish      DGVYWIMGRTSVDIIKSGGYKISALDVERHLLAHPDITDVAVIGAPDA
Xenopus        DGTYWILGRTSVDIIKSGGYKVSALEVERHLLGHPSITDVAVIGAPDV
               *. *** *********:****:***::*  ** **.********. *
```

ACSF3 encodes a mitochondrial methylmalonyl-coA and malonyl-coA synthetase and is postulated to catalyse the first step of intramitochondrial fatty acid synthesis (Sloan *et al.*, 2011). Mutations in this gene causing combined malonic and methylmalonic aciduria were first described in two back-to-back publications in 2011 (Sloan *et al.*; Alfares *et al.*). Since then two other reports have been published bringing the total number of cases to 17 (de Sain-van der Velden *et al.*, 2016; Pupavac *et al.*, 2016). Whilst a significant proportion of patients presented after the fourth decade of life with memory problems, seizures and T2 hyperintensities on brain MRI (Sloan *et al.*, 2011), the majority of individuals present in childhood with symptoms suggestive of a metabolic disorder. These include ketoacidosis, hypoglycaemia, metabolic acidosis, failure to thrive, elevated transaminases, seizures, encephalopathy, microcephaly, developmental delay, dystonia and hypotonia. Some patients are clinically asymptomatic, only being identified though

urine screening programs (Alfares *et al.*, 2011). The majority of patients have methylmalonic acid (MMA) concentrations at least ten times the upper limit of the reference range in both plasma and urine, with concentrations ranging between 5 - 50 μmol/L and 30 - 600 μmol/mmol creatinine, respectively (Sloan *et al.*, 2011). In the case of patient 24, plasma MMA concentrations were comparable to the previously described patients with the mildest elevation, being ten-fold elevated above the reference range. In contrast, his urinary MMA concentrations were much more mildly elevated than any cases in the literature being only increased 2.5-fold relative to the reference range. In some patients, MMA is the only accumulated metabolite; however, others also have detectable malonic acid (MA) in their urine (Pupavac *et al.*, 2016). In patient 24, MA was not consistently seen in urine and when it was present it was not quantified (as this is not routinely done in the laboratory where the sample was analysed). These apparently milder biochemical abnormalities are mirrored by his milder clinical features, presenting only with developmental delay and autism. Given its conservation, location in a functionally important part of the protein, predicted damaging impact and the consistent biochemical findings it is likely that this variant is pathogenic. Thus, this case expands the genotypic/phenotypic spectrum of a recently described disease.

3.3.6.3 *Patient 25*

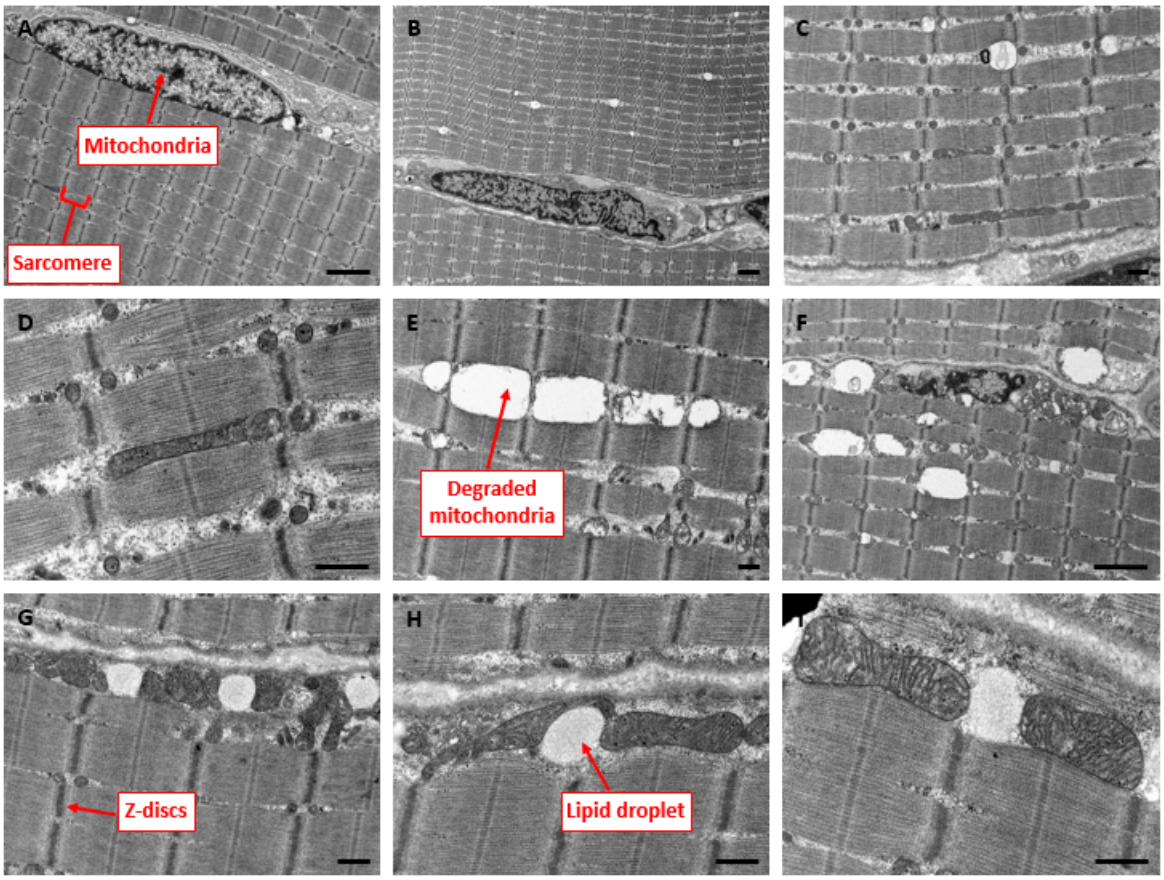
Patient 25 is the second child born to first cousin Pakistani parents. The perinatal history was unremarkable, he was bottle-fed and gained weight satisfactorily. He had the usual immunisations and his parents were not concerned about his progress until the second year of life. He did not start walking until he was 18 months old and when he did, he was very unsteady and had frequent falls. He was slow to start speaking and said his first words at 2 1/2 years of age. At 4 1/2 years of age he had short phrases in the parent's language of Punjabi/Urdu but very little English. He started school at this time but had difficulty in keeping up with the other children and suffered frequent drop attacks consisting of episodes of acute onset marked hypotonia associated with noisy breathing and cyanosis. At 4 years of age, he was not able to use a spoon but could scribble with a pencil despite not holding it with a pincer grasp. Examination at 9 years of age demonstrated generalised dystonia with variable upper limb tone and clearly increased lower limb tone, sustained ankle clonus and brisk reflexes throughout. His voluntary tongue movements were uncoordinated and he could not blow up his cheeks. He had difficulties in following with his eyes but there were no abnormal eye movements. Major clinical features were noted to be ataxia with diplegia, developmental delay and drop attacks. Numerous investigations had been carried out with the aim of establishing a diagnosis for this patient. Brain MRI demonstrated abnormal signal in the caudate and lentiform nuclei as well as some abnormal signal in the midbrain. This

was suggestive of a metabolic disorder and perhaps more specifically, a mitochondrial abnormality. Multiple EEGs have been normal. Extensive metabolic investigations including CSF lactate, neurotransmitter analysis, white cell enzymes and muscle mitochondrial respiratory chain enzymes were normal. Pyruvate dehydrogenase was found to be slightly low in skin fibroblasts but much higher than the affected range.

Sequencing identified a homozygous missense variant in *AFG3L2* (p.Leu356Arg), which encodes the catalytic subunit of the mitochondrial-AAA (m-AAA) protease. The variant is predicted deleterious and probably damaging by SIFT and PolyPhen-2, respectively. The Leu356 residue is conserved from humans to yeast and is part of an 18-amino acid region that is 100% conserved across this evolutionary time. m-AAA proteases can be formed from homo-oligomeric complexes of AFG3L2 or hetero-oligomeric complexes of AFG3L2 and paraplegin (*SPG7*). Mutations in *AFG3L2* can therefore result in much more severe phenotypes than those in *SPG7* as all m-AAA protease isoenzymes are affected (Elleuch *et al.*, 2006). These complexes are located within the inner mitochondrial membrane are responsible for degrading misfolded proteins and regulating ribosome assembly by promoting the maturation of proteins such as MrpL32, a critical component of the large subunit of mitochondrial ribosomes (Gerdes *et al.*, 2012). These functions are thought to be propagated by the recognition of the folded state of substrate proteins (i.e. in the case of MrpL32, degradation begins at the unstructured N-terminus but cannot proceed beyond the tightly-folded, cysteine-containing C-terminal region and the protein dissociates resulting in correct N-terminal processing) (Bonn *et al.*, 2011).

Mutations in *AFG3L2* can cause two distinct disorders, autosomal dominant spinocerebellar ataxia 28 (characterised by adult-onset dysarthria, ptosis and cerebellar ataxia) and autosomal recessive spastic ataxia 5, of which there has only been one reported case (Pierson *et al.*, 2011). This case involved two affected siblings born to consanguineous parents presenting with spasticity before the second year of life, tonic-clonic and myoclonic seizures, dystonia, dysarthria, dysphagia and motor degeneration. One sibling died at 13 years of age due to complications as a result of pneumonia. The other patient eventually lost the ability to ambulate and developed spastic paraparesis, oculomotor apraxia and stimulus-induced myoclonus. He underwent extensive testing with a view to achieving a diagnosis. He had normal cognition; brain MRI showed cerebellar atrophy and nerve conduction studies revealed an axonal sensorimotor neuropathy. A muscle biopsy examined by light microscopy was normal, however electron microscopy showed misplaced mitochondria associated with large lipid droplets. Similar electron microscopy findings of large lipid droplets, degraded mitochondria and some mitochondria with a disorganised cristae structure were also identified in patient 25 (Figure 3.3.8). However, the significance of these findings is uncertain.

Figure 3.3.8: Electron microscopy of skeletal muscle (quadriceps) from patient 25. **(A)** Control electron microscopy of skeletal muscle showing a normal muscle architecture with normal mitochondrial localisation and sarcomere structure. **(B-I)** Electron microscopy of skeletal muscle from patient 25. Mitochondria were normally localised near the Z-discs and most had normal structure. Muscle architecture and sarcomere sturcture was normal. Some lipid droplets were noted but this was not deemed to be outside normal limits. **(E-F)** Evidence of degraded mitochondria was noted, usually near the basal lamina, although this was likely artefactual. **(G-I)** A small proportion of mitochondria had a disorganised cristae structure, usually associated with close proximity to lipid droplets, of uncertain significance. Scale bars: **C, D, E, G, H, I**: 500 nm; **A, B, F**: 2 μm.



The p.Leu365Arg identified in our patient is in close proximity to p.Lys354Ala which has been found to severely impact upon AFG3L2 activity (Pierson *et al.*, 2011). These functional studies were performed in yeast lacking the ability to process MrpL32, abolishing the synthesis of mitochondrial-encoded respiratory chain subunits and impairing aerobic respiration. The p.Lys354Ala mutation abolished respiratory growth to a greater extent than the p.Tyr616Cys mutation identified in the two affected siblings (Pierson *et al.*, 2011). The p.Lys354Ala mutation is situated within the Walker A motif which binds ATP in the ATPase domain of AFG3L2. The mutated Leu365 in our patient immediately follows this Walker A motif but still resides within the ATP-binding ATPase domain. Indeed, the substitution of a hydrophobic leucine residue to

a larger electronically charged arginine residue may be predicted to cause a disruption of the tertiary structure in this region and potentially impacting on ATP binding.

Collectively, this evidence suggests that this variant is likely to be pathogenic. Patient 25 had a milder clinical course than the patients described by Pierson *et al.* (2011). This is likely explained by the fact that the p.Leu356Arg mutation is not predicted to affect the Walker A motif; this sequence motif has a highly conserved 3D structure with the lysine residue being invariant and critical for phosphate binding. This lysine compensates for the negative charges of the bound nucleotide, thereby stabilising the hydrolysis transition state and facilitating ATPase activity (Rosa and Nelson, 2011). On the other hand, the p.Leu356Arg mutation would likely disrupt ATP binding but not directly impact the catalytic mechanism, resulting in a lesser degree of functional impairment. In summary, we have identified a novel, likely pathogenic mutation in *AFG3L2* which, if proven by functional studies, would be only the second reported case of this disorder. In addition, the phenotypic spectrum would be widened to include significant intellectual disability as well as independent ambulation. Finally, after having extensive investigations and multiple EEGs in order to determine the cause of his drop attacks, these results also suggest that they may be due to stimulus-induced myoclonus.

3.3.6.4 *Patient 26*

Patient 26 is the first child born to first cousin consanguineous Afghan parents. He was able to sit by 10 months of age and crawl by 12 months. At 18 months of age, he was able to stand for a few seconds and take a few steps. He was babbling with double syllable babble although he never developed any words. Following a febrile illness at two years of age, he suffered acute developmental regression and developed a progressive dyskinetic movement disorder with dystonia and choreoathetoid movements. Brain MRI and CT demonstrated progressive abnormal signal in the basal ganglia and cerebellum. Sensorineural deafness was diagnosed at the age of two years and three months. The patient was also affected by dysphagia with failure to thrive requiring gastrostomy feeding and renal tubular dysfunction. Urinary organic acid analysis revealed an elevated 3-methylglutaconate with a normal level of 3-methylglutarate. The concentration of these analytes were not determined as it is not standard practice in the laboratory where these samples were analysed (Chemical Pathology Department, GOSH, UK). Orotate was also mildly raised at 8 μmol/mmol creatinine (ref: 0 - 5). A repeat analysis showed a consistent mildly raised 3-methylglutaconate with a normal 3-methylglutarate, although orotate was normal. Plasma amino acid analysis demonstrated very strongly raised methionine and elevated proline. Alanine was also elevated alongside mildly raised phenylalanine and tyrosine, possibly secondary to liver dysfunction. Lactate was also occasionally mildly elevated. A muscle biopsy showed

normal mitochondrial respiratory chain enzyme activities as well as normal mitochondrial DNA sequencing. Unfortunately, he died at four years of age due to multi-organ failure associated with infection. NGS analysis using the gene panel identified one novel homozygous mutation in *SERAC1* (c.1850delinsCA) which is predicted to cause a frameshift and the subsequent introduction of a premature stop codon (p.Ile617Thrfs*6).

Only 20 mutations in 28 patients with *SERAC1* mutations have been reported to date, all with a fairly homogenous clinical phenotype including 3-methylglutaconic aciduria, sensorineural deafness, encephalopathy and progressive Leigh-like (bilateral basal ganglia involvement on brain MRI) syndrome. During the first year of life, affected children typically suffer feeding problems and subsequent failure to thrive requiring gastrostomy feeding. Hypotonia and transient liver involvement are also seen in some cases. By two years of age, children reach developmental stasis or regress, ultimately being dependent on care-givers for all aspects of daily living. This can be largely attributed to progressive deafness, lack of speech development, dystonia and spasticity (Wortmann *et al.*, 1993). Lactate and alanine can also be elevated in plasma, leading to suspicion of a mitochondrial disorder. Our patient's phenotype is therefore consistent with a diagnosis of this disorder.

The exact function of SERAC1 still remains largely unknown and thus the pathogenic mechanism of mutations affecting its function are elusive. *SERAC1* encodes a phospholipase consisting of 654 amino acids with a single transmembrane region at the N-terminus anchoring the protein at the interface between the mitochondria and the endoplasmic reticulum and a conserved lipase domain suggesting a role in lipid metabolism (Wortmann *et al.*, 2012). Upon identification of the underlying gene defect responsible for this disorder, functional studies were undertaken using patient tissues (Wortmann *et al.*, 2012). Phospholipid analysis showed elevated levels of phosphatidylglycerol-34:1 and lower concentrations of phosphatidylglycerol-36:1 species in patient fibroblasts. Filipin staining also demonstrated an accumulation of unesterified cholesterol (Rodríguez-García *et al.*, 2016). Finally, morphological examination of patient muscle biopsies using electron microscopy revealed aggregates of degrading mitochondria. Collectively this data indicates that SERAC1 plays a role in phosphatidylglycerol remodelling, cholesterol trafficking and normal mitochondrial function.

The mutation identified in our patient results in truncation of SERAC1 by 32 amino acids. Many truncating mutations have been described including two similarly affecting the carboxy-terminal region resulting in the protein lacking the last 45 and 13 amino acids (Wortmann *et al.*, 2012). Given the pathogenic effect of smaller truncations than that seen in our patient it is plausible to conclude that the p.Ile627Thrfs*6 mutation is pathogenic. This finding was subsequently used to offer pre-natal testing for this family.

3.3.6.5 *Patient 27*

This patient was the first child born to consanguineous parents from Kuwait, presenting with a multisystemic and dysmorphic syndrome. Dysmorphic features such as microcephaly, retro- and micrognathia, a cleft of soft and hard palate were noted early in life. Organ malformations affected the brain (Dandy-Walker malformation), heart (doubly committed subarterial ventricular septal defect), lung (bilateral pulmonary hypoplasia) and gut (Hirschsprung's disease). He suffered early-onset asymmetrical tonic seizures responding to levetiracetam, but non-responsive to pyridoxine or pyridoxal 5'-phosphate. Continuous oxygen supplementation was required due to central apnoeas. Severe peripheral and bulbar neuropathy, with motor more pronounced than sensory, was also noted but assumed to have been secondarily acquired. He also had a very disrupted sleep pattern. Initially some very slow developmental progress was seen but, particularly following a prolonged stay in intensive care, he regressed. There was no speech development. Hyperphosphatasia, hypoketotic hypoglycaemia and hypogammaglobulinaemia were noted biochemically. Transferrin isoelectric focussing was normal. The child unfortunately died following an acute enterocolitis. A novel homozygous missense mutation was identified in the fourth transmembrane domain of *PGAP2* (p.Ala187Val). This variant is reported to have a minor allele frequency of 3/12052 alleles (0.0024%) and is predicted to be deleterious and probably damaging by SIFT and PolyPhen-2, respectively. Sequence alignment of *PGAP2* across species shows that Ala187 is conserved in mammals and birds but not zebrafish. Sanger sequencing confirmed that both parents were heterozygous for this mutation.

PGAP2 has a role in the synthesis of the glycosylphosphatidylinositol (GPI) anchor, a glycolipid which functions to tether approximately 150 proteins to lipid bilayers. These proteins have diverse functional properties and include membrane-associated enzymes, adhesion molecules, immunologically important proteins, antigens and receptors (Ferguson *et al.*, 2009). One example is alkaline phosphatase, a cell-surface hydrolase whose increased concentration in serum is indicative of a GPI-anchor defect. GPI-anchor biosynthesis occurs in three steps. Firstly, the GPI precursor is assembled in the endoplasmic reticulum (ER) membrane. Newly synthesised proteins are then imported into the lumen of the ER, the GPI-addition signal peptide is cleaved and the GPI moiety is added to the carboxy-terminus of the protein by a transamidation reaction. Finally, lipid remodelling and modifications of carbohydrate side chains are carried out in the ER and Golgi apparatus. PGAP2 functions in the third step of this pathway to reacylate the GPI with a saturated fatty acid (Kinoshita, 2014). This is the last enzymatic modification prior to correct expression of the GPI-anchored protein on the outer leaflet of the plasma membrane. Without this correct modification, the lyso-GPI-anchored proteins are transported to the cell

surface and secreted where they are susceptible to cleavage by phospholipase D (Tashima *et al.*, 2006).

To date, there have only been three reports detailing patients with mutations in *PGAP2* in nine individuals from four families (Krawitz *et al.*, 2013; Hansen *et al.*, 2013; Jezela-Stanek *et al.*, 2016). Many features described in these cases are also present in our patient including: microcephaly, cleft palate, Dandy-Walker malformation, heart septal defect, Hirschsprung's disease, epilepsy, hypotonia, profound developmental delay, disrupted sleep pattern, respiratory insufficiency and hyperphosphatasia. Elevated serum alkaline phosphatase activity was evident for all patients ranging between 1700 - 4455 U/L. The levels in our patient were consistent with these reports (1192 U/L – 9120 U/L (ref: 145 - 320)).

Defects of GPI-anchor biosynthesis are difficult to diagnose, partly due to the relatively recent description and a lack of clinical awareness of these disorders. Firstly, all patients have a normal transferrin isoelectric focussing pattern which often directs focus away from a possible congenital disorder of glycosylation. Secondly, elevated alkaline phosphatase is more commonly considered an indicator of liver or bone disorders and thus overlooked, especially if the patient does not fit these phenotypes and other liver function tests and calcium and phosphate levels are normal. Indeed, in the case of our patient, his markedly elevated levels were thought to be a laboratory mistake. In some inborn errors of GPI-anchor synthesis, lower expression of GPI-anchored markers such as CD55 and CD59 are observed in patient cells such as lymphoblastoid cell lines which can be discerned using flow cytometry (Sutherland *et al.*, 2007). This is not the case in patients with *PGAP2* mutations. To assess pathogenicity, all previous reports have performed *in vitro* analyses in Chinese hamster ovary (CHO) cell lines through transfection with mutant constructs. Under these conditions, introduction of the wild-type PGAP2 protein was able to restore surface expression of CD55 and CD59 to a greater extent than any of the PGAP2 proteins containing the patient mutations (Krawitz *et al.*, 2013; Hansen *et al.*, 2013; Jezela-Stanek *et al.*, 2016). Although effective, these functional studies for examining the effects of mutations in *PGAP2* are extremely time- and labour-intensive and would not be possible in the clinical setting; however, given the phenotypic correlation with the previously described cases, we are confident that the p.Ala187Val mutation detected is pathogenic. Additional features in this patient which have not been described previously include retro- and micrognathia, pulmonary hypoplasia, central apnoeas, hypoketotic hypoglycaemia and hypogammaglobulinaemia. However, additional cases are required to determine whether these features represent an expansion of the phenotypic spectrum of PGAP2-deficicency.

3.3.6.6 *Patient 28*

Patient 28 is a 7 year old male born to consanguineous first cousin parents, presenting mainly with developmental delay. He sat at 9 months and achieved independent walking at 23 months of age. He was affected by speech and language delay despite normal vision and hearing. Examination revealed microcephaly (< 2nd centile), increased tone and hyper-reflexia in the lower limbs with equivocal plantar responses but no ankle clonus. Brain MRI demonstrated a thin corpus callosum but otherwise unremarkable intracranial appearances. Neurometabolic testing showed a type-I transferrin isoelectric focussing pattern, suggesting an underlying congenital disorder of glycosylation. Gene panel sequencing revealed a novel homozygous variant in the *TPP1* gene (p.Gly296Asp), alongside a homozygous known pathogenic mutation in *ALDOB* (p.Arg60*) (Santer *et al.*, 2005).

TPP1 encodes tripeptidyl peptidase I, a protein which degrades peptides within lysosomes. Mutations were first described to cause late-infantile neuronal ceroid lipofuscinosis, typically impairing motor and mental development early in childhood causing movement disorders, intellectual decline, seizures and visual impairment. However, more recently a spectrum of disease severity corresponding to different mutations and degrees of residual enzyme activity have been described (Dy *et al.*, 2015). Subsequent to the identification of the potentially pathogenic homozygous *TPP1* variant, the corresponding enzyme activity was measured as 28 nmol/hr/mg protein (ref: 42 - 339) in patient 28. Homozygote patients typically have values in the range of 0.4 - 26 (Chemical Pathology, GOSH). This patient's clinical severity and progression is at the milder end of the spectrum compared to other patients with mutations in *TPP1* as he has never suffered seizures or visual loss. p.Gly296Asp is predicted to be deleterious and probably damaging by SIFT and PolyPhen-2, respectively. Whilst the presence of TPP1 has not been reported in lower organisms, the mutated Gly296 residue is conserved from humans to Xenopus (Figure 3.3.9).

Figure 3.3.9: Multiple sequence alignment of the TPP1 protein from various species. (*) positions with have a single fully conserved residue, (:) conservation between amino acids with strongly similar properties, (.) conservation between amino acids with weakly similar properties. The alignment was generated using Clustal Omega.

```
Human          RAGIEASLDVQYLMSAGANISTWVYSSPGRHEG-QEPFLQWLMLLS-NESALP
Gorilla        RAGIEASLDVQYLMSAGANISTWVYSSPGRHEG-QEPFLQWLMLLS-NESALP
Cat            RAGIEASLDVEYLMSAGANISTWVYSSPGRHES-QEPFLQWLLVLS-NESALP
Dog            RAGIEASLDVEYLMSAGANISTWVYSSPGRHES-QEPFLQWLLLLS-NESALP
Mouse          RAGIEASLDVEYLMSAGANISTWVYSSPGRHEA-QEPFLQWLLLLS-NESSLP
Zebrafish      KAGIEASLDVEYIMSSGANISTWVFTNPGRHES-QEPFLQWMLLLS-NMSAVP
Coelacanth     RAGLEASLDVEYIMSAGANISTWVFSNPGRYES-QEPFLDWLLLLS-NMSFIP
Lamprey        RAGLEASLDIEYIMSLGANISTWFWSTPGRRET-QEPFLQWLLALS-NTSSVP
Xenopus        RAGLEASLDVEYIMSTGANISTWVFSNPGKWHEMMEQDADWGGGILWDRGRAP
               :**:*****::*:** *******.:::.**: .   *   :*   :  : .  *
```

Mutations in *ALDOB* cause hereditary fructose intolerance, a disorder characterised by recurrent vomiting, abdominal pain and hypoglycaemia when fructose or sucrose is added to the diet at the time of weaning. Patients that survive infancy develop a natural avoidance of sweets and fruit, however long-term fructose exposure can result in liver failure, renal tubulopathy and growth retardation (Ali *et al.*, 1998). Commonly, a hypoglycosylated pattern of transferrin isoforms is identified and, as in our patient, misinterpreted as an indicator of a type-I CDG (Adamowicz *et al.*, 2007). Upon the identification of this mutation, further enquiries were made to the responsible consultant who identified that the patient's brother, although developmentally normal, also had an abnormal transferrin pattern. Both children were described to be "very much savoury boys". Sanger sequencing confirmed that both siblings were homozygous for the p.Arg60Ter mutation. Both children have normal liver function tests and hepatic ultrasound findings, suggesting that dietary self-selection has prevented serious liver dysfunction. Nevertheless, this finding has important clinical implications as they should not receive any sugar-containing medicines, especially not a fructose or sorbitol infusion which may be fatal.

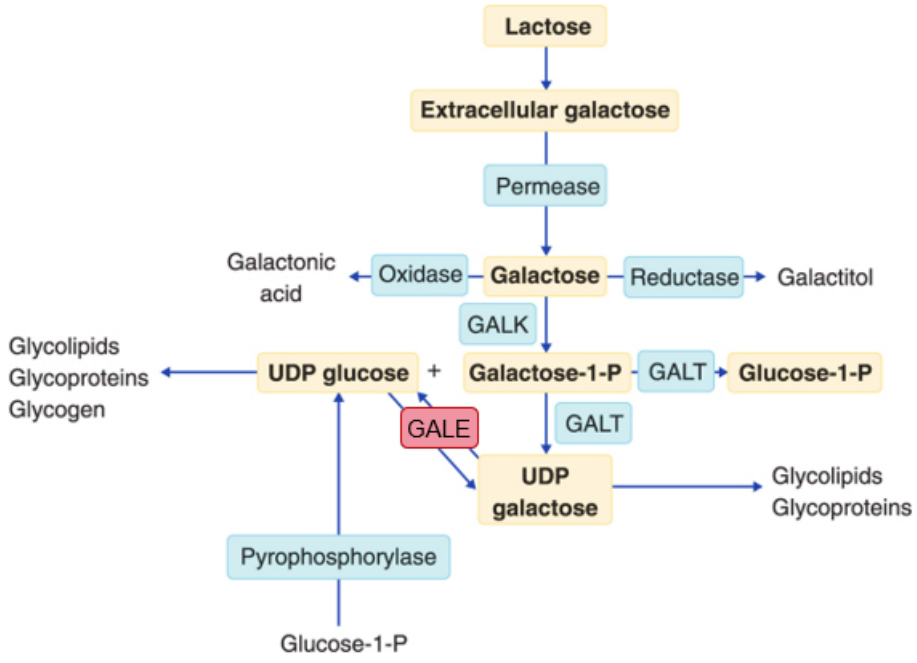
3.3.6.7 *Patient 29*

Patient 29 is a 6 year old male who presented with neonatal jaundice, lethargy, deranged liver function and hypoglycaemic episodes which subsequently subsided. Transferrin isoelectric focussing revealed a Type-I pattern, suggestive of a congenital disorder of glycosylation. At 8 months of age he was found to have severe bilateral sensorineural hearing loss and eventually required the insertion of cochlear implants. Post-infancy he remained clinically stable for approximately three years, but in view of his previous medical history, routine immunisations were delayed and remain incomplete. At four years of age he became unwell with pyrexia and hyptonia within two hours of pneumococcal immunisation, followed by recurrent upper and lower respiratory tract infections for several months. Immunological investigations revealed low levels of plasma IgA and IgM in the context of normal IgG, pneumococcus-specific antibodies and lymphocyte subsets, possibly reflective of transient hypogammaglobulinaemia of infancy.

Compound heterozygous mutations were found affecting adjacent amino acids in exon 3 of the *GALE* gene. The first (p.Val94Met) is a known pathogenic mutation (Wohlers *et al.*, 1999) and the second is a novel missense mutation (p.Gly95Asp). Both variants are predicted to be deleterious and probably damaging by SIFT and PolyPhen-2, respectively. *GALE* encodes UDP-galactose-4'-epimerase, an enzyme which catalyses two analogous reactions: the interconversion between UDP-galactose and UDP-glucose and between UDP-N-acetylgalactosamine (UDP-GlcNAc) and UDP-N-acetylglucosamine (UDP-GalNAc). Each pair of molecules are epimers, differing only in the position of one -OH group at one stereogenic centre within the sugar moiety. Mutations in this

gene cause galactosaemia type III which can present with very variable symptoms. In its mildest form, the enzyme deficiency is restricted to the circulating blood cells and considered clinically benign. On the other hand, the severe generalised form presents similarly to classical galactosaemia with jaundice, vomiting, hypotonia, sensorineural deafness, failure to thrive and hepatomegaly. To date, 22 pathogenic mutations have been described (Timson, 2006). Unlike galactosaemia caused by mutations in *GALT* and *GALK*, treatment with a galactose-free diet is not recommended as patients cannot use the endogenous pathway to synthesise UDP-galactose (Figure 3.3.10). Instead, a galatose-restricted diet and supplementation with N-acetylgalactosamine can be used to slow disease progression (Kingsley *et al.*, 1986).

Figure 3.3.10: Metabolic pathway showing the interconversions of galactose in the human body. Mutations in *GALT* and *GALK* cause galactosaemia type I and type II, respectively. Mutations in *GALE* (shown in red), as in our patient, cause galactosaemia type III.



p.Val94Met was the first mutation identified in a patient with the severe, generalised form of galactosaemia type III. At this time, activity of the mutant enzyme was quantified as 5% with respect to UDP-galactose and 24% with respect to UDP-GalNAc (Wohlers *et al.*, 1999). Experiments have since demonstrated that this mutation does not significantly impact on substrate binding but dramatically decreases the V_{max} of the enzyme with respect to both substrates (Wohlers and Fridovich-Keil, 2000). The 3D structure of GALE containing the p.Val94Met mutation has been experimentally solved (Thoden *et al.*, 2001) and has revealed that the Val94 residue prevents the substrates rotating out of the active site before undergoing modification of the sugar moiety, hence preventing binding that does not result in the desired product. The

mutation of this residue alters the shape of the substrate pocket and allows free rotation of the sugar moieties, thereby affecting catalysis. The effect on enzyme activity is not as dramatic when UDP-GalNAc is used as a substrate compared when UDP-galactose is used. This can be explained by the fact that the bulkier sugar moieties of UDP-GalNAc can adopt fewer non-productive binding conformations, thereby increasing the chances of effective catalysis. The effects of p.Gly95Asp are likely to be similar to that of p.Val94Met as this amino acid also resides within the binding pocket of the active site.

This case highlights how clinicians can be misguided by abnormal biochemical findings. A type-I transferrin isoelectric focussing pattern is suggestive of a congenital disorder of glycosylation, with over 40 defects known to cause this specific abnormality (Scott *et al.*, 2014). However, an abnormal transferrin pattern is seen in other disorders including untreated galactosaemia type III (Walter *et al.*, 1999). Given the patients's clinical features, galatosaemia was considered and functional studies were carried out at another specialist centre. Galactose-1-phosphate uridylyltransferase activity (GALT) was measured as normal on two occassions. Galactose-1-phosphate was not measured, but would have been expected to be elevated, suggesting a defect in the metabolism of galactose (Walter *et al.*, 1999). Following this, results stated that galactosaemia had been "excluded", when in fact only one of the three possible enzyme activities had been determined to be normal. Thus, in addition to considering all explanations for a particular biochemical abnormality, it is crucial to critically interrogate the results of any specialist investigations. A congenital disorder of glycosylation was also similarly assumed to be the diagnosis in patient 28 due an abnormal type-I transferrin pattern (Section 3.3.6.6).

3.3.7 *Patients with a diagnosis not fully explaining the phenotype*

In three patients within our cohort, mutations were identified that were likely to explain only a minority of their phenotypic features. Thus, it can be assumed that the remaining features are attributable to other, as yet unidentified genetic defects.

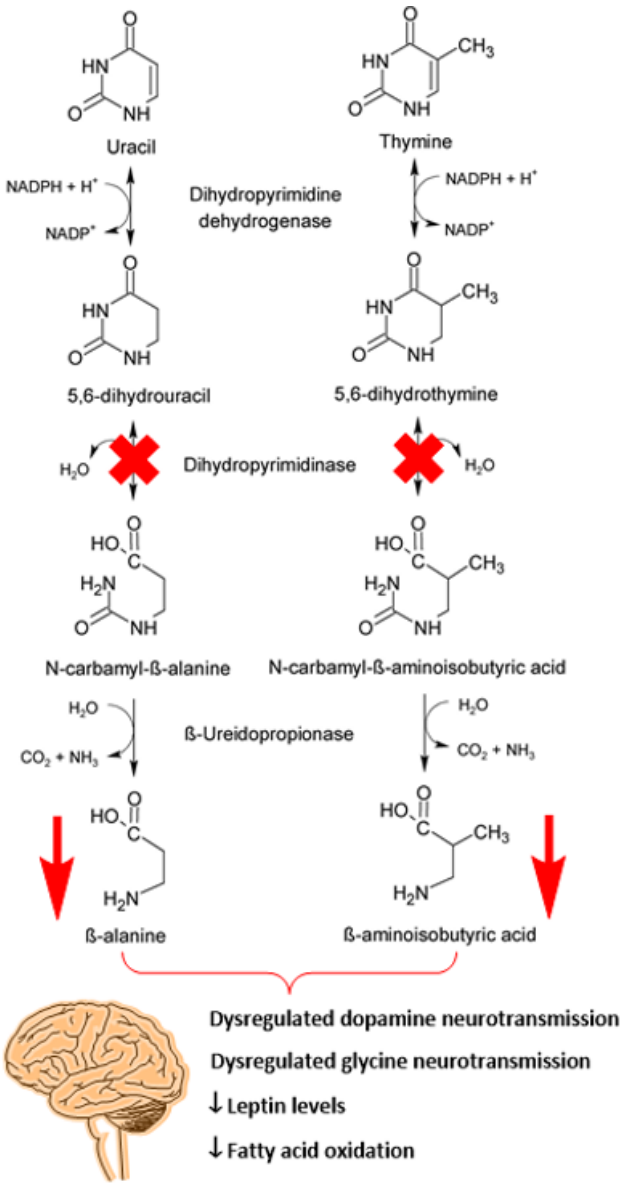
Patients 20 and 21 are monozygotic twins and were found to be homozygous for a novel missense variant in the *AASS* gene, encoding α-aminoadipic semialdehyde synthase. Mutations in this gene cause hyperlysinaemia. Although some patients have been reported to present with hypotonia, seizures or mild developmental delay, approximately 50% are asymptomatic (van Gelderen and Teijema, 1973). Given this, some controversy exists as to whether hyperlysinaemia should be considered a disease or a benign metabolic variant. There are two factors which make the evidence for pathogenicity of *AASS* mutations relatively weak. Firstly, the vast majority of the described cases, as well as our patients, are born to consanguineous parents which increases

the probability of harbouring damaging mutations in more than one gene. Secondly, the range of neurological phenotypes and severity of disease reported is very broad despite all cases having no functional AASS protein. Both twins had global developmental delay, hypotonia, brachycephaly and elevated lysine in plasma and CSF. One also had epilepsy that was well-controlled with AEDs. These features would be consistent with some reported cases of hyperlysinaemia (Houten *et al.*, 2013) , although the lysine concentrations seen in our patients are much lower than those described in this report. The missense variant in our patients (p.Arg322His) is conserved from humans through to Xenopus, zebrafish and coelacanths indicating its functional importance and predicted to be tolerated and probably damaging by SIFT and PolyPhen-2, respectively. It is likely that this variant is responsible for the elevated lysine concentrations in these children. However, it does not explain their global developmental delay, epilepsy, hypotonia or dysmorphism and therefore whole exome or genome sequencing would be an appropriate second tier test for these patients.

Patient 24 is a two year old male and the fourth child born of first cousin consanguineous parents. They are originally from Bangladesh and have two fit and healthy girls aged 3 and 10 years old. There had been a previous neonatal death in the family of a baby born at 28 weeks who sadly died at 8 months of age; he had renal dysplasia, bronchiolitis and raised triglycerides. A diagnosis of neutral lipid storage disease was considered but never verified. Patient 24 was born at 32 weeks after an emergency caesarean section. He was antenatally diagnosed with dysplastic kidneys and anhydramnios on the 28-week scan. He was intubated and ventilated at birth. In the very early neonatal period his renal function initially deteriorated but then gradually improved in the following days. He had lactic acidosis with a maximum concentration of 8.4 nmol/L, which subsequently resolved. He has been followed-up at the Metabolic Medicine Department at GOSH but no diagnosis has been made. His most recent assessment showed small kidneys with a normal renal function, hepatomegaly, elevated triglycerides and eczema. He is developmentally delayed but is making progress. A brain MRI showed a reduction in the commissural calibre. He has microcephaly ($<$ 0.4th centile), his weight is on the 25th centile and height is on the 0.9th centile. He also has elevated thymidine and uracil suggestive of dihydropyrimidine dehydrogenase (DPD) deficiency. The pyrimidine degradation pathway functions to convert uracil and thymidine to β-alanine and β-aminoisobutyric acid, respectively. This is catalysed by three enzymes: dihydropyrimidine dehydrogenase (DPD), dihydropyrimidinase (DPYS) and β-ureidopropionase, respectively. Mutations have been reported in genes encoding all three of these enzymes causing varying degrees of neurological symptoms in affected patients. A novel homozygous duplication resulting in a frameshift and premature stop codon (p.Val51Glyfs*50) in the *DPYS* gene was identified in this patient. Similarly to hyperlysinaemia, approximately 50% of patients with DPYS

deficiency are asymptomatic and are only diagnosed by biochemical testing. However, in others neurological abnormalities including intellectual disability, seizures, hypotonia, microcephaly and autistic features are seen. Some patients also present with gastrointestinal problems including gastroesophageal reflux, cyclic vomiting, villous atrophy and failure to thrive (van Kuilenburg *et al.*, 2010). Pathogenesis is proposed to be propagated by the decrease of β-alanine and β-aminoisobutyric acid which regulate dopamine and glycine neurotransmission, leptin levels and fatty acid oxidation (Begriche *et al.*, 2008; Ericson *et al.*, 2010) (Figure 3.3.11).

Figure 3.3.11: Schematic illustrating the pathogenic mechanism underlying DPYS deficiency.



Although certain features of patient 24 may at least in part be due to his DPYS deficiency (developmental delay, microcephaly and pyrimidine abnormalities), his remaining phenotypic features are likely to be attributable to other, as yet unidentified genetic defects. Other diagnoses

that were considered include neutral lipid storage disease, however due to the lack of availability of clinical Sanger sequencing of these genes this has yet to be done. Mutations in adipose triglyceride lipase (*PNPLA2*) or its activator, 1-acylglycerol-3-phosphate O-acyltransferase (*ABHD5*) cause neutral lipid storage disease with myopathy and Chanarin-Dorfman syndrome, respectively. Only *ABHD5* was included in our gene panel design. This had an overall coverage of 94.9% above 30X, however, two exons contained small regions that were not covered by any reads. Nevertheless, the coverage metrics of these regions were extremely similar across the other patient samples, suggesting that this lack of coverage is likely sequence specific (e.g. GC-rich regions) rather than due to the presence of a mutation in this patient. Indeed, no variants or putative CNVs were identified in *ABHD5* in patient 24. Although *PNPLA2* was not included at the time of patient analysis, this gene could be added in further iterations of the panel.

3.3.8 *Patients with potential diagnoses called into question by experimental evidence*

A final point of consideration emphasised by this study is the importance of stringent evaluation of potential pathogenicity and confirmatory sequencing. Gene panel sequencing identified candidate variants in three patients that were subsequently called into question or refuted based on additional genetic or functional evidence.

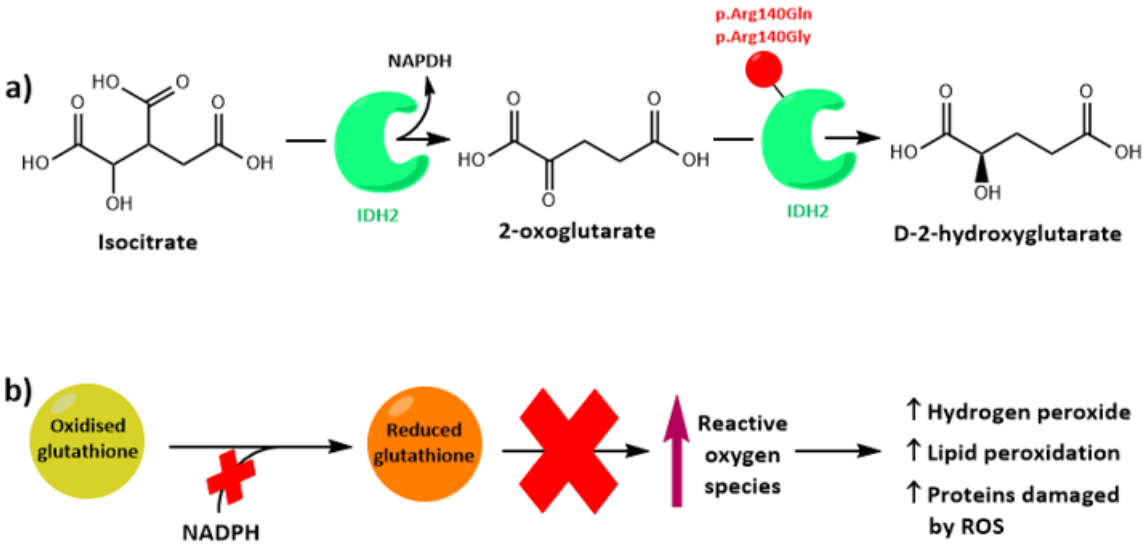
3.3.8.1 *Patient 38*

Patient 38 is a 6 year old female who initially presented with neonatal seizures that were described to have responded to treatment with vitamin B_6. She is severely developmentally delayed and suffered regression from three years of age. She has microcephaly, sensorineural deafness and a movement disorder with chorea and non-epileptic myoclonic jerks. Biochemically, she has been shown to excrete urinary methylmalonic acid (MMA). Sequencing identified a homozygous variant in *IDH2* (c.673G>A; p.Asp225Asn) which is predicted tolerated and possibly damaging by SIFT and PolyPhen-2, respectively. Both parents were confirmed to be heterozygous for the variant by Sanger sequencing. The variant is reported at a minor allele frequency of 0.02% and is conserved from humans to yeast. ACMG guidelines suggest that this variant should be classified as having uncertain significance (Table 3.3.7).

IDH2 encodes the mitochondrial NADP-dependent isocitrate dehydrogenase. Heterozygous mutations affecting specific residues in this gene (Arg140) allow it to convert 2-oxoglutarate (the product of the normal reaction) to D-2-hydroxyglutarate (Kranendijk *et al.*, 2010). Hence these mutations give rise to type II D-2-hydroxyglutaric aciduria. This would not be expected with homozygous mutations causing a loss of function of isocitrate dehydrogenase activity and indeed,

was not seen in patient 38. One of the main functions of the *IDH2* gene product is to provide NADPH in the mitochondrion which is important for the regeneration of reduced glutathione from the oxidised form, thus providing protection from reactive oxygen and nitrogen species. The IDH2 knockout mouse has increased hydrogen peroxide, increased malondialdehyde production (a product of lipid peroxidation) and carbonylated proteins (derived from proteins attacked by reactive oxygen species) (Kim *et al.*, 2014). The urinary MMA excretion identified in this patient may also have been explained by the fact that cobalamin can be inactivated by peroxynitrite as well as other reactive oxygen or nitrogen species (Mukherjee and Brasch, 2011).

Figure 3.3.12: Hypothesised pathogenic mechanism of loss of function mutations in *IDH2*. **(a)** The normal function of IDH2 is to catalyse the conversion of isocitrate to 2-oxoglutarate, resulting in the formation of NADPH. In the case of D-2-hydroxyglutaric aciduria, mutation of the Arg140 residue changes the specificity of this enzyme, resulting in the conversion of 2-oxoglutarate to form D-2-hydroxyglutarate. **(b)** In the case of mutations causing a loss of IDH2 function, there would be predicted to be a deficiency of NADPH within the mitochondria. Subsequently, reduced glutathione cannot be regenerated, resulting in an increase of reactive oxygen species and ROS-related damage.



Following this finding, urinary conjugated metabolites of vitamin E were measured as a marker of oxidative stress using the method developed by Sharma *et al.* (2013). No difference was seen between patient 38 and a cohort of controls, providing no evidence for a disturbance in redox state in this patient. The significance, if any, of this variant remains unknown and further genetic and biochemical investigations will be required.

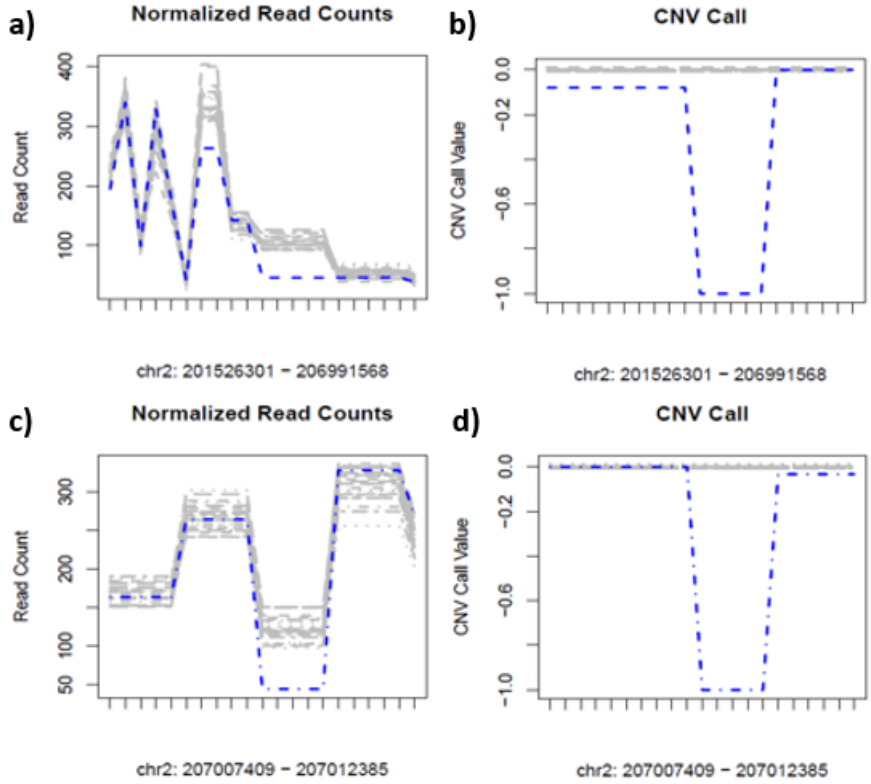
3.3.8.2 *Patient 39*

This patient is the third child of four children born to consanguineous parents. He was diagnosed with generalised tonic-clonic seizures at 8 months of age, however when looking retrospectively

seizures had occurred earlier in life. At this age he also had marked motor delay and hypotonia. He subsequently developed feeding difficulties, gastroesophageal reflux, recurrent aspiration pneumonia and had very limited developmental progress. At latest review aged 11 years, he is wheelchair-bound, unable to sit, can chew on his fingers but has no speech or understanding. He has increased tone throughout all limbs with brisk reflexes, as well as soft dysmorphic features with deep-set eyes and a slightly beaked nose. He continues to have seizures although these are relatively well-controlled with antiepileptic drugs. His older brother was similarly affected and died without a diagnosis at the age of 11 years. A number of biochemical investigations have yielded normal results including: plasma and CSF lactate, plasma amino acids, urinary organic acids and acylcarnitines. Brain MRI has been carried out on three occasions. At one year of age there was a generalised lack of white matter bulk. A repeat scan at three years of age showed an increase in the prominence of the periventricular spaces, early signs of an evolving leukodystrophy in the peritrigonal white matter with stable morphology of the corpus callosum and brain stem; appearances similar to those of his affected older sibling at a similar age. Finally, an MRI with contrast at eight years of age demonstrated subtle new changes at the level of the corticospinal tract, an increase in the number of perivascular spaces but the periventricular deep white matter abnormalities were unchanged. New diffuse brain leptomeningeal involvement was also noted. The radiological appearances as a whole were largely those of a failure of formation of normal white matter and not those of a progressive leukoencephalopathy.

No potentially pathogenic variants were identified in this patient using the standard analysis pipeline. Thus, deletion screening was attempted and two putative deletions were identified in *NDUFS1* using cn.MOPS (Figure 3.3.13). Mutations in this gene are a common cause of mitochondrial respiratory chain complex I deficiency and typically present with a severe and rapidly progressive leukoencephalopathy, in which patients share similar MRI features to those seen in our patient. Other features can include severe lactic acidosis, seizures, recurrent vomiting, hypotonia, macrocytic anaemia and ocular abnormalities Unfortunately, a muscle biopsy to determine the activity of the mitochondrial respiratory chain complexes had not been performed.

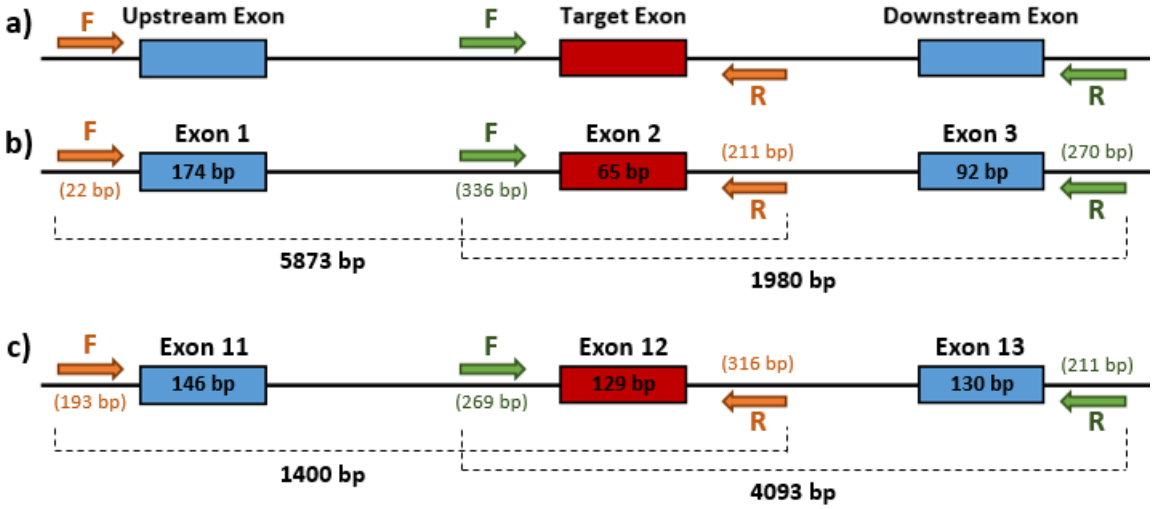
Figure 3.3.13: CNV analysis of patient 39 using cn.MOPS. Two putative heterozygous deletions in *NDUFS1* were identified in **(a-b)** exon 2 and **(c-d)** exon 12. The x-axis of each graph represents the genomic position. The y-axis represents **(a and c)** the read depth and **(b and d)** the copy number at the position of the call (i.e. 0 is wild-type, -1 is a heterozygous deletion, -2 is a homozygous deletion).



Following this finding, intronic primers were designed to amplify exon 2 and 12 of *NDUFS1*. Sanger sequencing revealed a sequence identical to the reference, indicating that if a deletion was present the breakpoints were not exonic. Long-range PCR was attempted using primers located upstream and downstream of the adjacent exons in order to sequence the adjacent intronic regions and identify any breakpoints that were present (Figure 3.3.14). Amplification of these regions was problematic as, despite the use of multiple DNA polymerases and PCR-enhancers, a single DNA fragment of the expected size could not be obtained.

Therefore, in order to investigate the presence of these deletions further, RNA was extracted from whole blood. cDNA was generated and the entire *NDUFS1* gene was amplified using Phusion DNA polymerase and the primers detailed in Appendix 2.8.1. The size of the wild-type *NDUFS1* cDNA is predicted to be 2537 bp, which agrees with the product seen in Figure 3.3.15. Exon 2 is 65 bp and exon 12 is 130 bp in length. If, as predicted by the cn.MOPS analysis, two heterozygous deletions were responsible for the disease in this family (i.e. one deletion was inherited from each parent) two DNA fragments of 2472 bp and 2407 bp would be evident (Table

Figure 3.3.14: Long range PCR methodology. (a) Schematic illustrating the locations of the two sets of primers designed to amplify the complete intronic regions adjacent to the deleted exon in order to identify any deletion breakpoints. Primer pairs are shown in orange and green, respectively. F, forward primer; R, reverse primer. Primer positions for the amplification of genomic regions surrounding **(b)** exon 2 and **(c)** exon 12. The distance of each primer from the beginning/end of each exon is shown in brackets.



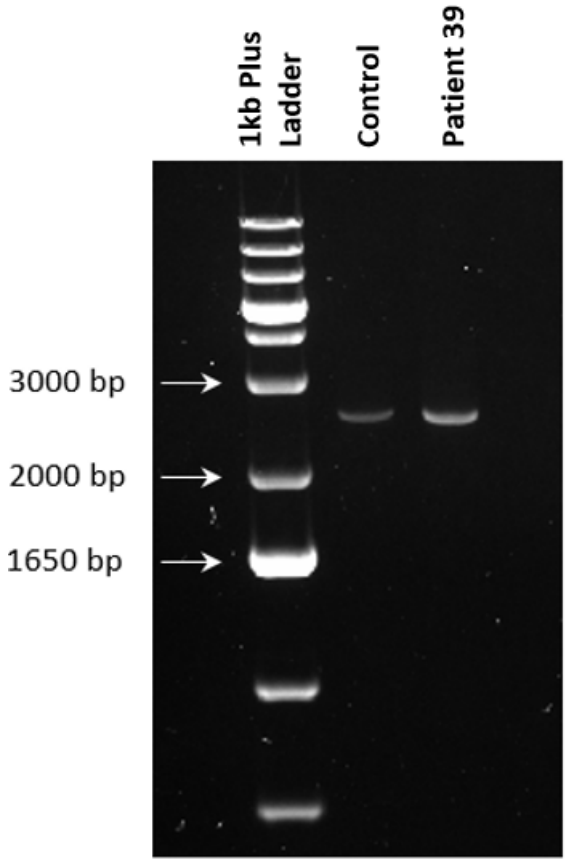
3.3.9). In patient 39 there was only one DNA fragment of the same size as the control indicating that no exonic deletions were present (Figure 3.3.15).

Table 3.3.9: Predicted sizes of wild-type and mutated NDUFS1 cDNA sequences.

	Wild-type *NDUFS1* (bp)	Exon 2 (bp)	Exon 13 (bp)	Predicted size of cDNA fragment (bp)
Wild-type	2537	-	-	2537
Heterozygous deletion of exon 2	2537	65	-	2472
Heterozygous deletion of exon 12	2537	-	130	2407
Compound heterozygous deletion of exon 2 and 12	2537 2537	65 -	- 130	2472 2407

Technological or genomic variations in the depth of coverage across chromosomal regions can lead to high false-discovery rates when detecting CNVs (Klambauer *et al.*, 2012). cn.MOPS claims to reduce this false discovery rate by modelling the depth of coverage independently at each genomic position and has been shown to be efficient for large-scale exome and genome sequencing data. However, these algorithms may not be sufficient to normalise for the extreme variability in read depth produced by the restriction enzyme-generated amplicons used by HaloPlex capture.

Figure 3.3.15: Agarose gel electrophoresis illustrating the amplification of the full-length *NDUFS1* cDNA from control and patient blood. Additional gel electrophoresis was performed in which the cDNA products were run further to determine whether a small difference in fragment size could be determined. No difference was observed, therefore the clearest image is shown.



3.3.8.3 *Patient 43*

Patient 43 presented to his local hospital at five weeks of age with failure to thrive, poor feeding, hypoglycaemia and neurological dysfunction. He was noted to have marked hypotonia, brisk tendon reflexes and an initially compensated metabolic acidosis. Investigations revealed a normal echocardiogram and brain MRI showed agenesis of the corpus callosum and colpocephaly with normal myelination. He was subsequently transferred to intensive care where a full septic screen was negative. Serum ammonia was 89 μmol/L (ref: $<$ 55 μmol/L) and he was commenced on thiamine, carnitine, biotin and pyridoxine. One week later he developed a decompensated metabolic acidosis with worsening hypotonia, poor cry, diminished spontaneous movement and spontaneous eye movement with a conjugate gaze. Sodium bicarbonate treatment was initiated with subsequent worsening of hypernatraemia. He was then noted to have prolonged clotting time indicating significant liver dysfunction. Over the following days after transfer to GOSH he continued to deteriorate with significant acidosis (pH of 6.9, lactate of 18 μmol/L) requir-

ing tris-hydroxymethyl aminomethane and renal replacement therapy. He required significant inotropic support and multiple blood product transfusions due to a coagulopathy. Following a cardiorespiratory arrest and in view of a mitochondrial myopathy being the most likely diagnosis, care was withdrawn and the patient died. A muscle biopsy taken at this time showed cytochrome oxidase (COX)-negative fibres and severely reduced mitochondrial complex I and IV activities.

Sequencing identified a homozygous variant in *EARS2* (c.760A>G; p.Thr254Ala). It was predicted to be deleterious and probably damaging by SIFT and PolyPhen-2, respectively. The residue is also conserved between humans and yeast, suggesting its functional importance, and is not reported in any population database. *EARS2* encodes glutamyl-tRNA synthetase 2 which catalyses the ligation of glutamate to tRNA molecules.

There have only been six reports detailing 18 patients with mutations in this gene causing early-onset leukoencephalopathy with thalamus and brainstem involvement and high lactate levels in body fluids (LTBL). It has been shown that these patients can be can be affected by a spectrum of disease severity ranging from a severe phenotype resulting in death in the neonatal period (Danhauser *et al.*, 2016) and a relatively mild phenotype characterised by elevated lactate and psychomotor developmental delay, who show some spontaneous clinical improvement including regression of MRI lesions and falling lactate concentrations (Steenweg *et al.*, 2012). There were various similarities when comparing our patient to those reported in the literature, especially with regard to his biochemical parameters. All patients had grossly elevated lactate concentrations, COX-negative fibres, ragged-red and ragged-blue fibres. Patients have been described to have variable mitochondrial complex deficiencies including complex I, III and IV deficiencies (Steenweg *et al.*, 2012). Patient 43 had severely reduced complex I and IV enzyme activity. MRI features are highly specific for the disorder including dysgenesis/agenesis of the corpus callosum and symmetrical abnormalities of the cerebral and cerebellar white matter, thalami, midbrain and brainstem. These features are highly consistent with those observed in our case. Patients have also been described to have varying degrees of liver dysfunction including elevated transaminases and α-fetoprotein, hepatomegaly, steatosis, fibrosis and cholestasis (Talim *et al.*, 2013). A post-mortem liver biopsy from our patient showed macro- and micro-vesicular steatosis and marked cholestasis. His grossly prolonged clotting time also indicated a significant liver abnormality. Clinical features in those patients with a poor outcome include hypotonia, dystonia, spastic tetraparesis, bradykinesia, ptosis, opthalmoplegia, severe visual impairment, reliance on tube-feeding and absence of speech.

Despite having a phenotype consistent with cases of *EARS2* mutations and the variant being classified as likely pathogenic using ACMG criteria (Table 3.3.7), this same variant was subsequently independently identified in a second unrelated family by another research group

in our department. The affected members of this family presented with a similar phenotype; however, extensive screening within the family identified the variant in a homozygous state in an unaffected sibling, calling the pathogenicity of this variant into question. Functional studies are currently ongoing.

3.3.9 *Patients who remained undiagnosed despite gene panel sequencing*

Despite our panel's demonstrated efficacy, there were also limitations when using this approach. Of the 22 patients without a biochemical marker pointing towards a specific diagnosis, 63.6% (14/22) remained undiagnosed after gene panel sequencing (patients 31 - 44). The clinical, biochemical and MRI findings in these patients are described in Table 3.3.10. In 35.7% (5/14) of patients a diagnosis was reached after our study had concluded. As detailed in Section 3.3.5, a diagnosis of hyperprolinaemia type II was confirmed in patient 31 by Sanger sequencing. Patient 35 was suspected to have an inborn error of vitamin B_{12} metabolism as biochemical investigations revealed persistently low plasma vitamin B_{12} concentrations. Gene panel sequencing identified one heterozygous variant in *CUBN* (c.8741C>T; p.Ala2914Val), predicted deleterious and probably damaging. Read depth analysis did not show any evidence of a CNV in this gene. Subsequent Sanger sequencing of genes known to cause vitamin B_{12} confirmed the presence of the variant we had identified but also revealed a heterozygous small deletion (c.328_332del) in *CUBN* and a heterozygous known pathogenic variant (c.290T>C; p.Met97Thr) in the *GIF* gene. Re-analysis of the data revealed that in the case of the deletion, the affected genomic region was covered by nine HaloPlex reads and only one contained the 4 bp deletion; thus, the frequency of the alternative allele did not meet the 20% cut-off to be reported as a possible variant. In addition, the first iteration of the variant calling pipeline developed by the NE Thames Regional Genetics Service at GOSH that was used to analyse our data, only called variants that were covered at a read depth greater than 30X. Similarly, the missense mutation in the *GIF* gene was covered by 12 reads and whilst present in $>$ 20% (25%) of amplicons, it was not called due to inadequate coverage (i.e. $<$ 30X). In the following iterations of the variant calling pipeline, this coverage threshold was dramatically reduced in order to detect variants such as these. In the case of patient 35, it remains unclear as to which of the identified variants are pathogenic.

In two patients, genetic findings associated with neurological abnormalities but not inborn errors of metabolism have been identified. Microarray analysis revealed a *de novo* deletion at 7q36.2 in patient 40. This includes *DPP6* which is known to be associated with microcephaly and mental retardation (Liao *et al.*, 2013). Although this may explain some of the patient's phenotypic features, it was thought that he may have a second genetic defect so he has been

recruited into the Deciphering Developmental Disorders whole exome sequencing project. Patient 41 was found to have a mutation in the *KANSL1* gene. Mutations in this gene have been shown to result in Koolen-de Vries syndrome (Zollino *et al.*, 2012). This protein is involved in histone acetylation and the syndrome is characterised by developmental delay, intellectual disability and distinctive facial features. Indeed, patient 41 shared a number of these features including upslanting palpebral fissures, a large broad nose, abnormal hair pigmentation and coarsening facial features. Interestingly, although they were both suspected to have a neurometabolic disease, neither had strong biochemical indicators so these findings are somewhat unsurprising.

Finally, patient 44 was found to have mutations in the *BCKDK* gene, causing branched-chain keto-acid dehydrogenase kinase deficiency. Affected patients were first described in 2012 by Novarino *et al.*, presenting with autism, epilepsy, intellectual disability and reduced concentrations of branched-chain amino acids. This diagnosis was not made by our panel because the gene was not included in our 614-gene panel as at the time of design a disorder had not been associated with this gene. In any future versions or formats of the gene panel, genes such as these could be added to the design.

Whilst in the remaining nine undiagnosed patients the sequencing metrics suggested that the capture efficiency and depth of coverage was good, mutations may have been missed due to inefficient capture of GC-rich regions or low coverage due to complexity. It is also plausible that the disease-causing genes were not included in the design, the causative mutations were intronic or homozygous/compound heterozygous CNVs were responsible for disease. Therefore these patients are being considered for further genetic testing including whole exome and whole genome sequencing.

Table 3.3.10: Patients who remained undiagnosed after gene panel sequencing. Some of these patients have subsequently had diagnoses established via array CGH or candidate gene testing.

Patient	Age	Gender	Primary neurological phenotype	Other phenotypic features	Relevant specialist investigations	Eventual diagnosis
31	6	Female	Developmental delay, absence seizures.	Bilateral sensorineural deafness.	Grossly elevated plasma proline, elevated n-pyrrole-2-carboxyglycine confirming hyperprolinaemia type II.	Hyperprolinaemia type II (homozygous complex insertion/deletion event in *ALDH4A1*, c.411_424delinsCGGCCC; p.Pro138GlyfsTer13)
32	7	Female	Learning difficulties, delayed motor milestones, reduced exercise tolerance, responsive to intramuscular vitamin B12 injections.	Joint hypermobility.	Methylmalonic aciduria, high plasma homocysteine, normal muscle biopsy.	Not yet reached
33	8	Male	Global developmental delay, neurological regression, dysphagia, epilepsy.	Alopecia, gastro-oesophageal reflux disease, neutropenia, treated with granulocyte-colony stimulating factor, platelet dysfunction. Brain MRI: Leigh-like changes in the basal ganglia and brainstem.	Intermittently elevated plasma lactate but normal CSF lactate, low plasma manganese.	Not yet reached
34	5	Female	Episodes of severe ketotic hypoglycaemia with seizures.	N/A	Normal acylcarnitines and plasma amino acids. Slightly low fructose-1,6-bisphosphatase activity.	Not yet reached
35	3	Female	Developmental delay and regression, dysphagia.	Brain MRI: delayed myelination	Low vitamin B12.	Multiple mutations, only one of which was picked up by gene panel (*CUBN*, p.Ala2194Val)
36	5	Female	One of similarly affected siblings, parental consanguinity. Developmental delay, reduced exercise tolerance, joint hypermobility.	Dysmorphic features, pancreatic insufficiency and fat malabsorption.	Several raised plasma amino acids. Muscle histology suggestive of a mitochondrial disorder but normal respiratory chain enzymes.	Not yet reached

37	8	Female	One of similarly affected siblings, parental consanguinity. Developmental delay, reduced exercise tolerance, joint hypermobility.	Pancreatic insufficiency and fat malabsorption.	Several raised plasma amino acids	Not yet reached
38	6	Female	Global delay, microcephaly, movement disorder with chorea and non-epileptic myoclonic jerks.	Previous faltering growth. Renal tubular acidosis on $NaHCO_3$ supplements. Brain MRI: delayed myelination.	N/A	Not yet reached
39	12	Male	Global delay, seizures, dysphagia. Sibling with similar features.	Dysmorphism. Brain MRI: leukodystrophy.	Persistent low arginine but normal lactate, carnitine profile and urinary organic acids.	Not yet reached. Unconfirmed *NDUFS1* deletions
40	5	Male	Global delay, retinal dystrophy, dystonic extensor spasms, epilepsy.	Gastro-oesophageal reflux disease, hip dislocation, scoliosis. Brain MRI: progressive cerebral and cerebellar atrophy.	EEG features consistent with electrical status epilepticus during sleep (ESES).	Microarray: deletion at 7q36.2. *De novo* change which includes the *DPP6* gene which is known to be associated with neurological disorders
41	15	Female	Developmental delay, paroxysmal episodes of gasping, opisthotonus and discomfort related to food ingestion.	Distinct facial features, abnormal maculae on optical coherence tomography and slightly swollen optic discs, mild scoliosis and hypermobility. Brain MRI: non-progressive ventricular dilation.	Abnormal VEP/ERG.	Diagnosis of Koolen-de Vries syndrome made by geneticists. *KANSL1*: c.1635-3T>C
42	6	Male	Global delay, four-limb motor disorder with variable increased tone.	Xp21 in-frame deletion within dystrophin gene. Sister with similar phenotype but without the dystrophin deletion. Brain MRI: arachnoid cyst.	High creatine kinase. Very long chain fatty acids (VLCFA): moderately raised C26 and C26:C22 ratio.	Not yet reached
43	1	Male	Marked hypotonia and hyper-reflexia at birth. Faltering growth, poor feeding, hypoglycaemia and lethargy at five months.	Rapid evolution to multi-organ failure and passed away shortly afterwards. Brain MRI: agenesis of the corpus callosum and colpocephaly.	Persistent lactic acidosis. Ammonia: 88. Normal mitochondrial genome analysis, no POLG common mutations identified.	*EARS2* variants picked up the panel. Functional work underway to establish its significance
44	7	Male	Global delay, acquired microcephaly.	Brain MRI: lack of white matter bulk and delayed myelination.	Persistently low levels of branched-chain amino acids in plasma in CSF.	Mutations in *BCKDK* gene diagnosed by candidate gene testing. Gene not on panel.

3.4 ADVANTAGES AND PITFALLS OF AN EXTENDED GENE PANEL FOR INVESTIGATING COMPLEX NEUROMETABOLIC PHENOTYPES

The data presented in this chapter shows that many suspected cases of inborn errors of metabolism will have mutations in known genes that are amenable to targeted screening and that when used in combination with clinical and biochemical evidence it can be an invaluable front-line approach within the clinical setting. Candidate gene sequencing is useful in cases that present with disease-specific features such as microcephaly, intellectual disability, syndactyly of the second and third toes and elevated 7-dehydrocholesterol in cases of Smith-Lemli-Opitz syndrome (Bianconi *et al.*, 2015) or elevated blood concentrations of phenylalanine in phenylketonuria (Vockley *et al.*, 2014). However, for the majority of IEM, a gene-by-gene Sanger sequencing approach is not economical or efficient due to the genetic heterogeneity involved. For example, an abnormal transferrin isoelectric focussing pattern suggests the presence of a congenital disorder of glycosylation (CDG); however, there are more than 100 genetic disorders known to cause CDG with more than 40 defects of N-glycosylation causing an abnormal transferrrin pattern, rendering sequential candidate gene sequencing extremely time-consuming and often futile (Scott *et al.*, 2014). Given this extreme genetic and phenotypic heterogeneity seen in IEM, a wider screening approach is necessary to facilitate timely genetic diagnosis. The advent of next-generation sequencing surmounts these issues and is enabling us to achieve the identification of causative genes in patients at an accelerating rate.

The gene panel methodology described in this chapter was shown to be a powerful tool that enhances the diagnostic ability in the clinical setting. The panel includes 614 genes which was comprehensive of all the genetic defects known to cause neurometabolic disorders in June 2013. Therefore, this methodology shares similarities with whole exome sequencing approaches for diagnosing these conditions in a clinical setting but with the added advantage of higher coverage of target regions (Sun *et al.*, 2015). This approach was used to investigate a cohort of 44 patients with a wide array of, often non-specific, neurometabolic symptomatology. We successfully identified 44 causal or likely pathogenic variants in 31 of the patients investigated, of which 24 were novel. Many of the 30 patients without a prior genetic diagnosis had biochemical indicators that may suggest a certain class of disorders (e.g. intermittently raised plasma lactate which would be suggestive of a mitochondrial disorder). However, only 30% had laboratory results that indicated a more specific diagnosis, further indicating that candidate gene sequencing would be unlikely to result in diagnoses in this patient cohort. In this sub-group of patients with strong biochemical abnormalities, pathogenic or likely pathogenic variants were identified in 88.9% of cases. This suggests that an extended panel approach with subsequent focus on

the candidate gene(s) could be an initial cost-effective approach for patients with suggestive biochemical findings. This is particularly pertinent where biochemical results point towards a disorder for which routine genetic testing is not available or when abnormalities suggest a group of disorders which can have multiple genetic aetiologies (e.g. mutations in twelve different *PEX* genes can cause a peroxisome biogenesis disorder as seen in patient 22 (Steinberg *et al.*, 1993)). In the remaining 21 patients with variable clinical presentations and no obvious indicators suggesting specific genes, likely pathogenic variants that could, at least partially, explain the observed phenotypes were identified in eight cases (38.1%). Despite the fact that some mutations failed to account for the entirety of the clinical picture, the diagnosis rate of our panel was higher than most panels targeting smaller sub-sets of human disease with cohorts of patients exhibiting less phenotypic heterogeneity, which typically range between 10 - 30% (Trump *et al.*, 2016; Kammermeier *et al.*, 2014; Jones *et al.*, 2013; Sommen *et al.*, 2016).

The data presented not only expands the genotypic and phenotypic spectrum of many well-described and recently documented neurometabolic disorders, but also re-empahsises the complexity of diagnosing patients with IEM and the advantages of an untargeted approach. Recent studies have shown that mutations at distinct disease loci in a patient can lead to complicated phenotypes, with 4.6% of participants having blended phenotypes resulting from two single-gene defects (Yang *et al.*, 2014). The existence of pathogenic variants in two genes in one patient is not surprising given that it has been shown that individuals harbour approximately 13,500 single nucleotide variants which alter the amino acid sequence within their exome (Marian, 2014). Blended phenotypes were also seen in our patient cohort with 9.1% (4/44) having two or more disorders, with this higher proportion likely occurring due to the high degree of consanguinity in our cohort. These issues further contribute to the complexity of accurately diagnosing IEM and highlight the power of NGS as a clinical tool for establishing molecular diagnoses. Conversely, they emphasise the need for diagnosticians to perform elaborate and accurate clinical phenotyping and not over-rely on sequencing results, especially when the identified gene defects do not fully account for the observed clinical picture. Futhermore, we have highlighted how common it is for clinicians investigating neurometabolic disorders to be misguided by investigation results with consequential diagnostic delays. Finally, although applied on a paediatric cohort, our approach would arguably be even more useful in adult populations, where neurometabolic phenotypes can be even more atypical, presentations more variable and biochemical phenotypes even more subtle (Sedel *et al.*, 2007).

3.4.1 *Difficulties in the interpretation of the pathogenicity of novel variants*

This study also highlighted the difficulties encountered in the interpretation of novel variants. Of the pathogenic or likely pathogenic variants identified, 24 were novel and included one splice site mutation, six insertions and/or deletions and 17 missense variants. Whilst splice site, nonsense and frameshift variants are typically expected to have large effects on protein function, the interpretation of novel missense variants can often be challenging. *In silico* prediction tools can give indications of variant pathogenicity but other factors should also be considered (Richards *et al.*, 2015). In this chapter SIFT and PolyPhen-2 were used, both of which are based on multiple sequencing alignments but with the latter also incorporating structure-based prediction algorithms. As each tool uses a slightly different methodology, the interpretation of variants can differ; hence, combinations of *in silico* tools should be used. However, the best combination is often gene-dependent (Leong *et al.*, 2015). Mutation Taster, Combined Annotation Dependent Depletion and Likely Ratio Test are additional tools that are becoming more popular for these type of functional consequence predictions (Schwarz *et al.*, 2010; Kircher *et al.*, 2014; Chun and Fay, 2009). In this study, such discordance between prediction tools is evident not only for novel variants such as p.Asn322His in *AASS* and p.Ser485Arg in *ACSF3*, but also for variants of experimentally established pathogenicity such as p.Gly309Arg in *ASL* (Linnebank *et al.*, 2002) and p.Leu195Pro in *GALT* (Reichardt *et al.*, 1992). Furthermore, both SIFT and PolyPhen-2 classified a known pathogenic mutation in *IDUA* (p.Leu490Pro) as tolerated and benign, respectively (Bach *et al.*, 1993). Previous studies have also indicated the inability of these online prediction tools to correctly predict the pathogenicity of all variants analysed, hence increased awareness and caution in interpretation of these results is warranted (Walters-Sen *et al.*, 2015). Nevertheless, *in silico* tools, alongside population and evolutionary conservation data, remain valuable in filtering large numbers of variants identified using NGS platforms but further evidence to support or refute pathogenicity should be sought where appropriate. Indeed, in this emerging era of genomic medicine, protein structural analysis to examine pathogenicity is increasing utility and feasibility on a large-scale (Yue *et al.*, 2014). This involves mapping the mutated proteins onto the known structures of human wild-type structures or those of homologues from other species. The potential effects of each variant on bonding interaction, packing and secondary structures due to the amino acid substitution can then be inspected using *in silico* prediction software. Should the 3-dimensional structures of more proteins with the human exome be experimentally solved and become publicly available, this approach would be a valuable aid towards novel variant analysis. Finally, it is crucial that the experimental evidence accompanying reports of rare pathogenic variants are revisited and treated with scrutiny. One study has reported that 27%

of mutations cited in the medical literature as pathogenic are common polymorphisms or have been misannotated (Bell *et al.*, 2011). Indeed, false assignments of pathogenicity may result in incorrect prognostic, therapeutic and reproductive advice being given and having severely detrimental consequences on patient care (MacArthur *et al.*, 2014).

3.4.2 *Summary: The future of IEM diagnosis*

With decreasing NGS costs and the advent of the 100,000 Genomes Project, it is plausible that WES and WGS will supersede the use of gene panels in the clinical diagnostic setting in the near future. However, many challenges need to be addressed prior to this implementation, including difficulties in interpreting the overwhelming amounts of data generated (WGS will identify approximately 4 million variants per patient) and uncertainties regarding clinically reportable and actionable findings (Dewey *et al.*, 2014). Moreover, WES and WGS have proven invaluable in the identification of novel genes in many disorders including inborn errors of metabolism (Howard *et al.*, 2014; Novarino *et al.*, 2012; van Karnebeek *et al.*, 2014), but such findings are often not currently actionable within the diagnostic setting. Elucidating the significance of these variants is not possible without functional characterisation in appropriate settings and models, which is often expensive and beyond the capacity of most clinical diagnostic laboratories. Until such challenges can be surpassed, the use of a targeted gene panel approach provides a rapid and cost-effective method of evaluating patients with neurometabolic disorders and enables more timely diagnosis and prompt treatment initiation in these conditions. Indeed, our panel has been incorporated into the GOSHome Gene Panel (www.labs.gosh.nhs.uk/media/759058/goshome_v7.pdf) which targets approximately 5000 known disease-causing genes. This panel is based on an Agilent SureSelect design, meaning newly-identified genes can be easily added to future iterations and large-scale deletion screening using tools such as ExomeDepth can be more effectively implemented as sequencing fragments are generated using random mechanical shearing.

4

SEIZURES DUE TO A KCNQ2 MUTATION - TREATMENT WITH VITAMIN B_6

4.1 WHOLE EXOME SEQUENCING FOR THE DIAGNOSIS OF IEM/NMD

As demonstrated in Chapter 3, many patients with a suspected IEM remain undiagnosed even following extensive genetic testing. In cases such as these, and in patients presenting with atypical phenotypes suspected to represent a novel disorder, whole exome sequencing (WES) can be used to perform genetic screening of the protein-coding regions of the genome. Advantages of this approach when compared to standard Sanger or gene panel sequencing is the ability to detect variants in novel disease-causing genes and therefore not relying on the prior identification of candidate genes or groups of disorders.

As part of this thesis, ten families comprising fifteen affected children underwent WES. All had been comprehensively phenotyped and extensively genetically and biochemically investigated previously by the Metabolic Medicine Team at GOSH. A definitive or potential diagnosis that could be further investigated using functional studies was identified in four families. These findings are described and their implications in relation to the medical literature are discussed in Chapters 4, 5 and 6. The clinical details of the six families in whom a genetic diagnosis was not made are given in Table 4.1.1. These patients will not be discussed further due to the constraints of this thesis format. However, it is possible that potentially pathogenic variants were not identified in these cases because they lie within intronic or regulatory regions, disease is due to large CNVs or the mutated regions were not adequately captured by the SureSelect kit or covered by sequencing reads.

Table 4.1.1: Patients who underwent whole exome sequencing and remained undiagnosed.

Family	Number of affected patients	Clinical and biochemical phenotype	Hypothesised diagnoses
1	2	Failure to thrive, diarrhoea, vomiting, ketotic hypoglycaemia, hypermobile joints, ear infections and croup, eosinophilic colitis/duodenitis, predominance of unconjugated bile acids in urine.	Abnormality of bile acid metabolism, abnormal liver X receptor (LXR) or farnesoid X receptor (FXR) signalling.
2	2	Early-onset movement disorder, generalised dystonia, frequent epileptiform activity with multifocal features on EEG, undetectable CSF PLP whilst off-treatment, muscle biopsy in-keeping with myopathic features.	Inborn error of vitamin B_6 metabolism or homeostasis.
3	1	Episodes of encephalopathy, ataxia, dysarthria, regression from 22 months of age, dry and scaly skin rash, skin and neurological problems improved by nicotinamide treatment, mild aminoaciduria.	Inborn error of nicotinamide metabolism, abnormality of NAD transport to NAD-dependent enzymes, defect of kynurenine metabolism, mitochondrial disorder.
4	1	Failure to thrive, hypotonia, interstitial lung disease, chronic cough and rhinitis, abnormal shape of lower ribs, short stature with growth hormone resistance, developmental delay, protein-losing enteropathy, calcification in basal ganglia and subcortically, gastro-oesophageal reflux with funcoplication, abnormal fat distribution, bruises easily, lethargy, mild dysmorphism, low albumin, elevated alanine transaminase.	Congenital disorder of glycosylation, Aicardi-Goutieres syndrome, abnormality of growth hormone/IGF1 axis, surfactant deficiency, cutis laxa

5	2	Seizures, developmental delay, abnormal MRI, neuropsychiatric abnormalities, spasticity, motor impairment, neurodegeneration with brain and eye abnormalities.	Neurometabolic disorder.
6	1	Dilated cardiomyopathy with biventricular dysfunction and pulmonary hypertension, perimembranous ventricular septal defect, pulmonary valve dysplasia, mild dysmorphic features, generalised hypotonia, extreme joint laxity, umbilical hernia, cutis laxa, rocker bottom feet with prominent heels, segmentation anomaly of cervical spine, acute metabolic decompensation following chest infection complicated by low cardiac output, severe hepatic and moderate renal dysfunction, mildly decreased mitochondrial complex IV activity, vitamin D deficiency.	Noonan's syndrome, connective tissue disorder, neuromuscular disorder, mitochondriopathy, congenital disorder of glycosylation.

4.2 INTRODUCTION

In this chapter, the case of a girl (patient X) whose neonatal-onset seizure disorder appeared to respond to a multivitamin supplement containing pyridoxine is presented. At this time she also had a high plasma (670 nmol/L; ref: 15 – 73 nmol/L) but low CSF (12 nmol/L; ref: 14 – 92 nmol/L) concentration of PLP, resulting in a high plasma:CSF ratio. Abnormalities such as these are often indicative of a primary inborn error of vitamin B_6 metabolism or a secondary defect affecting PLP availability.

4.2.1 *Vitamin B_6-dependent and -responsive disorders*

Since the mid-1900s it has been known that dietary vitamin B_6 deficiency causes seizures in infants. This was particularly noted during an epidemic of seizures in American infants between 1951 and 1953 who were fed a milk formula deficient in vitamin B_6 (Coursin, 1954). These seizures remit upon administration of physiological doses of vitamin B_6 (0.2 - 0.5 mg for infants and 0.5 - 1.5 mg for children) and is now extremely uncommon, especially in developed countries (Ohtahara *et al.*, 2011).

In contrast, genetic disorders disrupting the normal metabolism of dietary vitamin B_6 cause severe seizures remitting only upon administration of pharmacological doses of vitamin B_6 (at least 10 times higher that physiological requirements). These disorders and their pathogenesis are described in detail in Chapter 7. Pyridoxine dependent epilepsy (PDE) was first described by Hunt *et al.* (1954). Patients typically present with antiepileptic drug-resistant seizures in the neonatal period that respond dramatically to pyridoxine (PN) and remain seizure free on this treatment. This metabolic defect has been shown to be due to a deficiency of α-aminoadipic semialdehyde (α-AASA) dehydrogenase (antiquitin), an enzyme on the lysine catabolic pathway (Mills *et al.*, 2010). The accumulating upstream metabolite, L-Δ^1-piperideine-6-carboxylate (P6C) forms an adduct with pyridoxal 5'-phosphate (PLP), the active form of vitamin B_6, rendering it inactive as a cofactor. P6C is in equilibrium with α-AASA and it is measurement of these compounds, primarily in urine but also in plasma and CSF, that now forms the biochemical basis for diagnosis, alongside molecular genetic analysis of *ALDH7A1*. In the majority of patients with PDE, treatment with intravenous PN (50 or 100 mg single dose) followed by a maintenance oral dosing regimen of 5 – 30 mg/kg/day (maximum 200 mg/day) results in seizure resolution. Patients with pyridox(am)ine 5'-phosphate oxidase (PNPO) deficiency, are also now being recognised as responding to PN; despite early patient cohorts only showing a clinical response to PLP (Mills *et al.*, 2014).

Only children with hypophosphatasia due to mutations in tissue non-specific alkaline phosphatase (*ALPL*) have been documented to have a high plasma:CSF ratio as seen in patient X (Baumgartner-Sigl *et al.*, 2007; de Roo *et al.*, 2014). These ratios have been documented to be as high as 36.1 prior to commencing PN treatment and 24.0 in a patient on 135 mg/day PN. In contrast, our patient had a higher ratio of 55.8. Alkaline phosphatase is required for calcium uptake, bone mineralisation and cellular uptake of PLP (hence why patients with hypophosphatasia have abnormal PLP levels); patients typically present with severe bone disease which can be fatal during infancy (Baumgartner-Sigl *et al.*, 2007). Patients with severe disease may also have AED-resistant seizures that can be ameliorated with PN. However, there are milder forms of the disease which can present in childhood or adulthood with rickets, premature loss of deciduous teeth and repeated bone fractures.

A low CSF PLP concentration alone is a more non-specific finding that can be seen in a number of disorders. Low CSF and plasma PLP concentrations are indicative of both PNPO and antiquitin deficiencies (Mills *et al.*, 2005; Ormazabal *et al.*, 2008; Footitt *et al.*, 2013; Stockler *et al.*, 2011). Two other classes of disorder have been described to cause vitamin B_6-responsive seizures, namely hyperprolinaemia type II and disorders of glycosylphosphatidylinositol (GPI) anchor biosynthesis. Given the pathophysiology of these conditions, we might expect patients to have low CSF PLP, however there are no reports of patients having this measured. Hyperprolinaemia type II, like PDE, results in an accumulation of a metabolite (L-Δ^1-pyrroline-5-carboxylate (P5C)) which forms an adduct with PLP, rendering it inactive and causing seizures. Patients have grossly elevated proline concentrations in plasma and elevated P5C in plasma, urine and CSF (Flynn *et al.*, 1989). Disorders of GPI anchor biosynthesis have only recently been identified over the past five years to cause a subclass of congenital disorders of glycosylation. To date, mutations have been identified in eleven genes: *PIGV*, *PIGO*, *PIGL*, *PIGW*, *PIGT*, *PIGA*, *PIGM*, *PIGN*, *PIGY PGAP2* and *PGAP3*. Alkaline phosphatase is normally GPI anchored to the cell surface and must be correctly located to facilitate PLP to enter the brain. Thus, these disorders lead to hyperphosphatasia (elevated serum alkaline phosphatase) and epileptic seizures in almost all cases. Indeed, treatment with high-dose PN monotherapy resulted in complete cessation of seizures in two cases (Thompson *et al.*, 2006; Kuki *et al.*, 2013). Despite affecting a common metabolic pathway, GPI anchor defects are characterised by vast additional phenotypic diversity including characteristic dysmorphic facial features, cleft palate, short terminal phalanges, occular abnormalities, sensorineural deafness, heart defects, renal abnormalities, Hirschsprung disease and cerebellar atrophy (Almeida *et al.*, 2006; Krawitz *et al.*, 2010, 2012, 2013; Ohba *et al.*, 2014; Howard *et al.*, 2014; Chiyonobu *et al.*, 2014; Fujiwara *et al.*, 2015; Lam *et al.*, 2015; Tarailo-Graovac *et al.*, 2015; Ilkovski *et al.*, 2015).

The patient described in this chapter was referred to the metabolic clinic at GOSH because of her biochemical abnormalities indicative of an inborn error of vitamin B_6 metabolism. Urinary α-AASA analysis and sequencing of the *PNPO* gene excluded both antiquitin and PNPO deficiency and her clinical phenotype was not typical of any of the disorders described above.

4.2.2 *Vitamin B_6 to treat idiopathic epilepsy*

Since the recognition that vitamin B_6 plays a critical role in normal brain function, many trials of supplementation with PN or PLP in children with epilepsy of various aetiologies have been carried out, with variable outcomes. Many studies have demonstrated that an average of 13% of patients with West syndrome (a severe epilepsy syndrome with patients typically affected by infantile spasms, hypsarrythmia and mental retardation) are responsive to either PN or PLP (Table 4.2.1) (French *et al.*, 1965; Hansson and Hagberg, 1968). Indeed, in Japan high-dose vitamin B_6 treatment is the initial therapy given to children with this diagnosis (Tsuji *et al.*, 2007), despite the disorder having a multitude of possible aetiologies. The addition of high-dose PN treatment to current AED regimens also improved response rates and decreased time taken for seizure cessation in children with recurrent convulsions due to infection (Jiao *et al.*, 1997). Vitamin B_6 treatment has also been shown to be effective in between 5 - 9.5% of children with idiopathic epilepsy, with PLP being more effective than PN (Wang *et al.*, 2005; Mishra *et al.*, 2010).

Table 4.2.1: Efficacy of vitamin B_6 treatment in West syndrome (modified from Ohtahara *et al.* (2011)). None of these reports stated that a maximum threshold of vitamin B_6 that could be administered to each patient per day was used. PN, pyridoxine; PLP, pyridoxal 5'-phosphate.

Reference	PN/PLP	Dose per day	Efficacy
Fukuyama *et al.* (1980)	PLP	10-30 mg/kg	3/64 (4.7%)
Izuora and Iloeje (1989)	PN	150-1200 mg	0/9 (0.0%)
Ito *et al.* (1991)	PN	10-50 mg/kg	1/20 (5.0%)
Pietz *et al.* (1993)	PN	100-300 mg/kg	5/17 (29.4%)
Yoshida (1993)	PLP	10-70 mg/kg	9/59 (15.3%)
Suzuki *et al.* (1996)	PLP	20-50 mg/kg	2/25 (8.0%)
Takuma and Seki (1996)	PLP	20-50 mg/kg	3/28 (10.7)
Ohtsuka *et al.* (2000)	PLP	30-400 mg	30/216 (13.9%)
Toribe (2001)	PLP	20-50 mg/kg	6/50 (12.0%)
Wang *et al.* (2005)	PN	10-20 mg/kg	2/13 (15.4%)
	PLP	30-50 mg/kg	4/13 (30.1%)
Total			65/501 (13.0%)

A recent review of 216 West syndrome patients treated with PLP between 1969 and 1998 at Okayama University Hospital revealed many interesting findings including higher response rates amongst patients receiving higher-dose supplementation and with an unknown seizure aetiology (Ohtahara *et al.*, 2011). Whilst the genetic and biochemical basis for the response of many of these patients has not been investigated, the proportion responding is so high it is highly unlikely that a significant proportion have antiquitin or PNPO deficiency. One explanation may be genetic heterogeneity in genes involved in vitamin B_6 metabolism. This is supported by the variability in B_6 requirements amongst patients and efficacy in children with various seizure semiologies and aetiologies. Another may be that a B_6-response is age-dependent and related to changing requirements for neurotransmitter synthesis by B_6-dependent enzymes. Supporting this, successful withdrawal of PN or PLP has only been reported to occur after two years of age (Ohtahara *et al.*, 2011). Whole exome sequencing is one method that can be used to investigate the genetic aetiologies of idiopathic cases such as these. This approach was used to investigate the underlying aetiology of disease in patient X.

4.3 CASE REPORT

Patient X, a daughter of unrelated parents, was born by spontaneous labour at 38^{+6} weeks after an uneventful pregnancy. Good foetal movements were reported. The baby had hiccoughs during the last trimester, although similar movements were also reported during the mother's first pregnancy with an unaffected child.

She was born in good condition and discharged on day three of life. Prior to this she suffered two episodes of facial reddening, stiffening and then becoming pale; these were associated with feeding and thus assumed to be reflux. One day post discharge, she had episodes of choking and cyanosis, associated with stiffening after which she became floppy. Further seizures on day 4 were documented at her local hospital; these were accompanied by "cycling" movements of her arms with oxygen saturations dropping to 68%, lasting less than one minute. A full septic screen including a lumbar puncture was negative and she was given a loading dose of phenobarbitone. Seizures continued following this requiring control with phenobarbitone, phenytoin and lorazepam. At nine days of age she had a normal cranial ultrasound and CT scan but an electroencephalogram (EEG) demonstrated some asymmetry with larger amplitude responses on the right and abnormal paroxysmal components. A brain MRI on day 28 demonstrated normal brain structures with appropriate maturation, but some increased signal intensity in the subthalamic nuclei around the lateral geniculate nuclei bilaterally.

A series of biochemical investigations between seven days and one month of age demonstrated mild, but likely insignificant, abnormalities of plasma amino acids (Table 4.3.1). Urine analysis showed widespread mild elevation of multiple amino acids (Table 4.3.2) and organic acid analysis revealed mildly elevated 2-oxoglutarate and pyruvate (not quantified) interpreted as a possible renal tubule leak or immaturity. Other analytes found to be high were gamma-glutamyl transferase 256 U/L (ref: 12 – 43 U/L) and alkaline phosphatase 342 U/L (ref: 129 – 291 U/L), a common finding in children on anticonvulsants.

Table 4.3.1: Plasma amino acid abnormalities identified in patient X with a suspected inborn error of metabolism. The samples were analysed by high-performance liquid chromatography (HPLC) as part of the patient's clinical care. Only amino acids which were reported to be abnormal on at least one occasion are shown. -, reported as normal. All results are expressed in μmol/L.

Amino acid	Age at time of sampling			Reference range
	7 days	30 days	4 months 5 days	
Glutamine	876	882	-	400 - 800
Lysine	87	-	-	100 - 300
Phenylalanine	33	-	-	35 - 100
Arginine	-	172	-	40 - 120
Ornithine	-	146	238	25 - 120
Taurine	-	168	216	40 - 140
Glutamate	-	173	264	25 - 130
Glycine	-	-	341	100 - 300
Serine	-	-	431	90 - 290
Alanine	-	-	499	150 - 450

By six weeks of age seizures continued to occur sporadically, beginning with both eyes staring towards the corner of the room, mouth pouting, clonic movements of both limbs and respiratory grunting sounds. A repeat EEG at this time demonstrated abnormal delta activities and intermittently occurring angular or sharp waves, mainly anteriorly whilst at rest. Conversely, when she cried or had been alerted, the recording was of lower voltage without sharp waves but the content was abnormal. At two months of age she was commenced on 0.3 mL of DaliVit multivitamin oral drops per day; a dose which contained 0.25 mg of PN. Seizures were reported to have ceased six days after this, at a time when she was also receiving 6.7 mg/kg/day of phenytoin and 11 mg/kg/day of carbamazepine. Further biochemical testing at three months of age, revealed a plasma PLP level of 670 nmol/L (ref: 15 – 73 nmol/L) but a CSF PLP level of 12 nmol/L (ref: 14 – 92 nmol/L). This high plasma:CSF PLP gradient suggested an abnormality of vitamin B_6 metabolism, thus she was commenced on 5 mg/kg/day of PN at four months of

Table 4.3.2: Urine amino acid abnormalities identified in patient X with a suspected inborn error of metabolism. Samples were analysed by ion exchange chromatography with post column ninhydrin derivatisation using the Biochrom 30+ amino acid analyser as part of the patient's clinical care. Only amino acids which were reported to be abnormal on at least one occasion are shown. -, reported as normal. All results are expressed in μmol/mmol creatinine.

Amino acid	30 days	Reference ranges
Glycine	1254	300 - 950
Serine	192	25 - 95
Threonine	61	10 - 45
Alanine	167	30 - 130
Glutamine	247	40 - 120
Lysine	115	5 - 20
Cystine	41	5 - 35
Taurine	316	30 - 55
Histidine	263	50 - 155
Aspartate	61	10 - 45

age. Phenytoin was then weaned with no increase in seizure frequency. Urinary α-AASA was not elevated and no mutations were detected in *ALDH7A1* or *PNPO*.

Throughout the following two years she continued to have seizures, mainly during intercurrent illness, requiring 15 mg/kg/day of carbamazepine for control, in addition to PN supplementation. An MRI and EEG were unremarkable, therefore weaning of carbamazepine was carried out over a period of two months. Three days after weaning she had two generalised seizures lasting between 3 – 4 minutes, consisting of tongue biting, stiffening, going pale, grunting, frothing at the mouth and becoming floppy afterwards. The carbamazepine was recommenced at her original dose but she became very ataxic (a side effect of this medication) therefore the dose was halved (7.3 mg/kg/day). Neurotransmitter analysis at four years of age revealed a slightly low level of methyltetrahydrofolate of 41 nmol/L (ref: 52 – 178 nmol/L), thus she was started on calcium folinate (7.5 mg/day).

Since commencing PN treatment at four months of age her dose had been increased to 15.6 mg/kg/day in addition to 7.3 mg/kg/day of carbamazepine, in line with weight gain. At seven years old, she was developmentally delayed with minimal expressive language and attends a school for children with special needs but remains healthy except for seizures in the context of intercurrent illness. A recent EEG has shown a change to a left temporal lobe focus.

Whole exome sequencing was performed as described in Section 2.5 and filtered to look for variants inherited in both an autosomal dominant and recessive manner (Section 4.5). Confirmation of the pathogenic *KCNQ2* mutation was carried out using the PCR conditions outlined in Tables 4.4.1 and 4.4.2. Visualisation of amplification products, PCR clean-up and Sanger sequencing were performed as described in Section 2.3.2.

Table 4.4.1: Composition of the reaction mix used to amplify the mutation-containing region of *KCNQ2*. Amplification was carried out using a GoTaq DNA Polymerase Kit (Promega, Maddison, WI).

Reagent	Volume (μL)
Nuclease-free water	16.35
5X PCR reaction buffer	6
$MgCl_2$ (25 mM)	2.4
dNTP mixture (10 mM)	0.6
Forward primer (5 μM)	1
Reverse primer (5 μM)	1
GoTaq DNA polymerase (5 U/μL)	0.15
Genomic DNA (75 ng)	1

Table 4.4.2: Touchdown PCR cycling conditions used to amplify the mutation-containing region of *KCNQ2*.

Step	Conditions
1	95°C for 3 mins
2	95°C for 30 secs
3	64°C for 30 secs (-1°C per cycle)
4	72°C for 1 min
5	Repeat steps 2-4 9 times for a total of 10 cycles
6	95°C for 30 secs
7	54°C for 30 secs
8	72°C for 1 min
9	Repeat steps 6-8 24 times for a total of 25 cycles
10	71°C for 7 mins

4.5.1 *Autosomal recessive filtering*

Whole exome sequencing data was initially filtered and analysed assuming that the patient's disorder had been inherited in an autosomal recessive manner. Variants classified as "pathogenic" or "likely pathogenic" according to computed American College of Medical Genetics and Genomics (ACMG) Guidelines (detailed in Tables 3.3.7 and 3.3.8) were included. Only variants associated with a loss of gene function, including missense, splice site, insertions, deletions or variants affecting start or stop codons were considered. Five missense variants were identified in five genes (Table 4.5.1). Of these, only *CHD2*, encoding chromodomain helicase DNA binding 2, is known to cause a seizure phenotype when mutated. *De novo* mutations in this gene cause childhood-onset epileptic encephalopathy with a typical onset of seizures in the second year of life. Patients typically develop normally in the first year of life before presenting with developmental delay, followed by the onset of myoclonic, absence and tonic-clonic seizures (Thomas *et al.*, 2015). The majority of patients (up to 70%) have seizures which are refractory to treatment with anti-epileptic drugs (Thomas *et al.*, 2015). *CHD2* encephalopathy is also characterised by extreme clinical photosensitivity and self-induced photic seizures.

The clinical phenotype of our patient would be an atypical presentation for a child affected by *CHD2* encephalopathy. Tonic-clonic seizures commenced on day four of life and she has never demonstrated any photosensitivity or photoparoxysmal response. Her seizures additionally showed a good response to anti-epileptic drugs, particularly carbamazepine. Furthermore, *CHD2* encephalopathy is typically caused by *de novo* mutations in *CHD2*, whereas the variant detected (p.Ser1407Thr) was in a homozgous state in our patient and heterozygous in both parents. The same variant was also detected in one of four unrelated samples run in the same sequencing batch. This missense variant has not been previously associated with disease and has been reported in publicly available databases at a minor allele frequency of approximately 1% (1000 Genomes Project, 0.88%; NHLBI Exome Sequencing Project, 1.22%, ExAC, 1.53%). However, this frequency rises to 3.24% in South Asian populations. Our family is of Asian origin and hence this variant was considered unlikely to be pathogenic in this case.

The other variants identified were present in genes that have been associated with disorders, but were not consistent with the clinical phenotype of patient X (Table 4.5.1). The *LCT* gene encodes lactase which catalyses the metabolism of lactose to form glucose and galactose. Mutations in this gene cause lactose intolerance resulting in abdominal pain, diarrhoea and nausea upon the consumption of lactose-containing dairy products. Although its physiological role

remains unclear, *PNMA2* has been tentatively linked to paraneoplastic neurological syndrome (i.e. indirect remote effects of cancer on the nervous system) due to the detection of this protein using antibodies in the serum of affected patients (Hoffmann *et al.*, 2008). Finally, *TRIP11* encodes Golgi microtubule-associated protein 210 in which mutations cause achondrogenesis type 1A. This disorder is characterised by extremely short limbs, narrow chest, short ribs, and bones that are poorly ossified and fracture-prone. Indeed, affected infants usually die before or soon after birth (Smits *et al.*, 2010).

Table 4.5.1: Whole exome sequencing data was filtered to show autosomal recessive variants. Chr, chromosome; n/a, not applicable due to both SIFT and PolyPhen-2 only predicting functional impact for simple variants that change the protein sequence; -, not stated in Ingenuity Variant Analysis software.

Chr	Position	Reference Allele	Patient Allele	Gene	Protein Variant	Het/Hom	SIFT	PolyPhen	dbSNP ID	1000 Genomes Frequency	NHLBI ESP Frequency	ExAC Frequency
2	136552328	T	G	*LCT*	p.Glu1665Ala	Het	Tolerated	Benign	-	-	-	-
2	136562522	C	T	*LCT*	p.Asp1427Asn	Het	Tolerated	Possibly damaging	776196678	-	-	-
8	26365425	G	A	*PNMA2*	p.Arg283Cys	Hom	Tolerated	Possibly damaging	78548714	0.92	0.02	0.49
14	92473994	T	G	*TRIP11*	p.Glu506Ala	Hom	Damaging	Probably damaging	2273186	1.54	0.93	0.66
15	93545488	T	A	*CHD2*	p.Ser1407Thr	Hom	Tolerated	Benign	61756301	0.88	1.22	1.53

4.5.2 *Variants in genes known to cause inborn errors of vitamin B_6 metabolism*

One of the most striking features seen in our patient, which directed further metabolic investigations, was her abnormally high plasma:CSF PLP ratio as well as her apparent improvement in seizure control on PN treatment. Antiquitin and PNPO deficiency were both ruled out biochemically and/or genetically. The high plasma PLP level in patient X whilst on only 0.07 mg/kg/day PN was noteworthy; being much higher than levels reported in healthy individuals and similar to levels in adults taking 0.63 mg/kg/day PN (Midttun *et al.*, 2005), and in fact more comparable to children 8.0 mg/kg/day for treatment of PDE or 30 mg/kg/day for PNPO deficiency (Footitt *et al.*, 2013) (Table 4.5.2). Moreover, the plasma:CSF PLP ratio of 55.8 was very striking; being much higher than the upper limit of 4.4 in paediatric patients with neurological disease (Footitt *et al.*, 2011). The only other genetically defined disorder in which a high plasma:CSF PLP ratio has been documented is hypophosphatasia due to mutations in alkaline phosphatase (*ALPL*).

Table 4.5.2: Comparison of plasma PLP concentrations in patient X and in patients described in theliterature on differing doses of PN supplementation for various conditions. [1], Midttun *et al.* (2005); [2], patients 5 and 6 from Footitt *et al.* (2013); [3], patients 1 and 2 from Footitt *et al.* (2013); PDE, pyridoxine dependent epilepsy; PNPO, pyridox(am)ine 5'-phosphate oxidase.

Patient	PN supplementation	Plasma PLP (nmol/L)	Reference range (nmol/L)
Patient X	0.25 mg/day (0.07 mg/kg/day)	670.0	15.0 - 73
Cardiovascular patients[1]	40 mg/day (0.63 mg/kg/day)	234.0 - 585.0	17.0 - 102.3
PDE patients[2]	200 mg/day (8 mg/kg/day)	587.9 - 603.3	46.0 - 321.0
PNPO-deficient patients[3]	30 mg/kg/day	580.0 - 632.6	46.0 - 321.0

Given these metabolic abnormalities, whole exome sequencing data was also scrutinised for potentially pathogenic variants in genes known to cause inborn errors of vitamin B_6 metabolism, a high plasma:CSF PLP ratio or hyperphosphatasia; namely, *PNPO*, *ALDH7A1*, *ALPL* and genes involved in glycosylphosphatidylinositol (GPI) anchor synthesis. Twenty variants were identified in the candidate genes (Table 4.5.3), including eleven synonymous variants and nine single amino acid substitutions.

Table 4.5.3: Autosomal recessive variants present in vitamin B_6-related genes in patient X. Splice site prediction was performed using tools available at (www.fruitfly.org/seq_tools/splice.html) (Reese *et al.*, 1997). Chr, chromosome; n/a, not applicable.

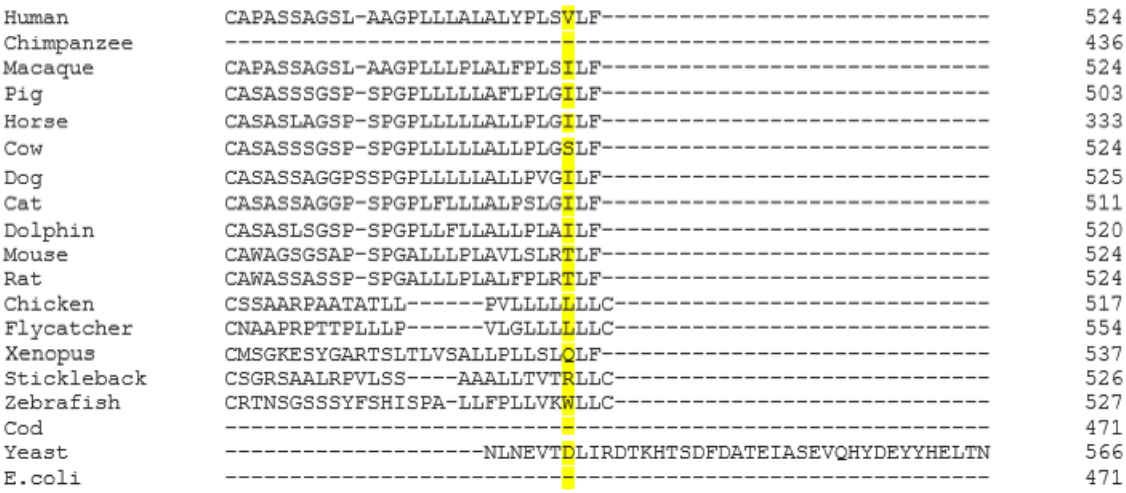
Chr	Position	Reference Allele	Sample Allele	Gene	Protein Variant	Splicing effect	Het/Hom	SIFT	PolyPhen	dbSNP ID	1000 Genomes Frequency	NHLBI ESP Frequency	ExAC Frequency
1	21889635	T	C	*ALPL*	p.Ser110Ser	None	Hom	n/a	n/a	1780316	92.95	92.17	94.50
1	21904131	T	C	*ALPL*	p.Val522Ala	n/a	Het	Tolerated	Benign	34605986	7.19	7.61	11.07
1	77634948	G	A	*PIGK*	p.Tyr124Tyr	None	Hom	n/a	n/a	1779199	98.66	98.32	99.55
1	77685042	T	C	*PIGK*	p.Thr16Ala	n/a	Het	Damaging	Benign	12723684	19.17	17.97	12.28
2	46839477	T	C	*PIGF*	p.Ala109Ala	None	Het	n/a	n/a	1824050	18.03	21.76	32.75
3	196674307	C	T	*PIGZ*	p.Met487Ile	n/a	Het	Damaging	Benign	17855662	3.04	5.29	6.66
3	196674749	C	T	*PIGZ*	p.Arg340Gln	n/a	Hom	Tolerated	Benign	4916589	63.22	60.77	59.11
3	196674879	A	G	*PIGZ*	p.Leu297Leu	None	Hom	n/a	n/a	12636891	59.25	49.22	53.81
3	196674916	A	C	*PIGZ*	p.Ala284Ala	None	Het	n/a	n/a	1147240	57.73	45.76	45.96
3	196674972	C	T	*PIGZ*	p.Ala266Thr	n/a	Hom	Tolerated	Benign	574365	80.51	57.58	76.99
3	196674973	T	C	*PIGZ*	p.Ala265Ala	None	Hom	n/a	n/a	573708	81.75	61.94	77.16
4	509850	T	C	*PIGG*	p.Ser197Ser	None	Het	n/a	n/a	11726338	2.06	2.72	2.21
4	515489	G	A	*PIGG*	p.Ala296Ala	None	Het	n/a	n/a	13115344	20.33	18.87	13.76
4	517376	T	C	*PIGG*	p.Leu448Leu	None	Het	n/a	n/a	76662266	22.42	19.64	15.18

4	517461	T	C	*PIGG*	p.Cys477Arg	n/a	Het	Tolerated	Benign	7666425	22.42	19.65	15.18
4	517622	C	T	*PIGG*	p.Ala530Ala	None	Het	n/a	n/a	13150531	22.42	19.09	21.86
4	520853	G	A	*PIGG*	p.Val566Ile	n/a	Het	Tolerated	Benign	13114026	17.83	15.67	13.81
4	527677	T	C	*PIGG*	p.Ile748Thr	n/a	Het	Tolerated	Benign	34623004	20.79	19.26	15.02
4	533001	T	C	*PIGG*	p.Phe799Ser	n/a	Het	Tolerated	Benign	1127410	20.87	19.31	15.04
17	37830900	A	G	*PGAP3*	p.Val104Val	None	Hom	n/a	n/a	2941504	57.79	65.76	69.83

All of the variants were listed in dbSNP, reported to have minor allele frequencies of at least 2%. *In silico* splice site prediction revealed that none of the synonymous variants would be predicted to affecting correct mRNA splicing. All of the missense variants were predicted to be benign by PolyPhen-2 and 7/9 were also predicted to be tolerated by SIFT. The remaining two variants were p.Thr16Ala in *PIGK* and p.Met487Ile in *PIGZ*. The first variant is reported in publicly available databases at a minor allele frequency of approximately 15% (1000 Genomes Project, 19.17%; NHLBI Exome Sequencing Project, 17.97%, ExAC, 12.28%). The second variant occurs with a minor allele frequency of approximately 5% (1000 Genomes Project, 3.04%; NHLBI Exome Sequencing Project, 5.29%, ExAC, 6.66%). In addition, two other missense variants affecting amino acid 487 have been described (p.Met487Thr and p.Met487Leu) indicating that this residue is not likely to have high importance for protein function.

The only variant of the 20 detected for which there is evidence of a physiological effect is p.Val522Ala in *ALPL*. This variant is annotated as a gain of function as a genome-wide association study investigating factors associated with recurrent kidney stones identified this variant both as being correlated with the formation of kidney stones and with increased serum levels of alkaline phosphatase protein and increased enzyme activity (Oddsson *et al.*, 2015; Orimo *et al.*, 2001). This variant is very common with minor allele frequencies of over 10% (1000 Genomes Project, 7.19%; NHLBI Exome Sequencing Project, 7.61%, ExAC, 11.07%). However, this frequency rises to 17.89% in South Asian populations. Thus, reference ranges for serum alkaline phosphatase concentrations are likely to already incorporate this variation. In addition, this variant was detected in a homozygous state in one of the patient's parents and in a heterozygous state in one of four unrelated individuals run in the same sequencing batch. Nevertheless, a gain of function of *ALPL* would not be consistent with the biochemistry seen in our patient. ALP is a glycophosphatidylinositol (GPI) anchored protein which functions as a PLP phosphatase, allowing unphosphorylated species to enter the brain (Whyte *et al.*, 1985; Giocondi *et al.*, 2008). Thus, an increase in circulating alkaline phosphatase may be expected to result in a decrease in plasma PLP and less available to enter the brain. Furthermore, this residue is extremely poorly conserved with other hydrophobic as well as polar amino acids found at this position (Figure 4.5.1). The observed substitution of a valine for an alanine residue is a very minimal change with the removal of a -$CHCH_3$ group, resulting in an amino acid with very similar physicochemical properties. Collectively the evidence, i.e. the high frequency in the general population, effect on circulating ALP levels (Oddsson *et al.*, 2015; Orimo *et al.*, 2001), poor conservation across species and minimal change in amino acid properties, suggests that this variant would be unlikely to result in the elevated plasma and reduced CSF concentrations of PLP reported in patient X.

Figure 4.5.1: Multiple sequence alignment of the *ALPL* gene across species. The lack of sequence conservation of the site of the p.Val522Ala variant is highlighted in yellow. The alignment was generated using Clustal Omega.



4.5.3 *Autosomal dominant filtering*

Given that no plausible variants were identified that fitted an autosomal recessive pattern of inheritance (i.e. homozygous or compound heterozygous variants) and that there was no family history of seizures, the data was re-analysed to look for *de novo* variants. Data was scrutinised using the same classification and inclusion criteria as for the autosomal recessive analysis. Ten variants were identified in nine genes, including seven missense variants, one insertion and two deletions (Table 4.5.4). Many of the variants were identified in genes that encode proteins of unknown function or that have not yet been associated with disease. The first is *BAIAP2L2*, whose protein product is involved in receptor-mediated endocytosis (Veltman *et al.*, 2011). *KIF12* is a member of the kinesin superfamily of microtubule-associated molecular motors that play an important role in intracellular transport and cell division (Chen *et al.*, 2007). *FOXRED2* is thought to encode a flavoprotein which may function in endoplasmic reticulum-associated degradation of non-native proteins. *RNF145* encodes ring finger protein 145, a protein of unknown function (Shim *et al.*, 2011). Finally, *WDR26* enncodes a member of the WD repeat protein family. It is thought that this protein plays a role in cell signal transduction and leukocyte migration (Zhu *et al.*, 2004; Sun *et al.*, 2011).

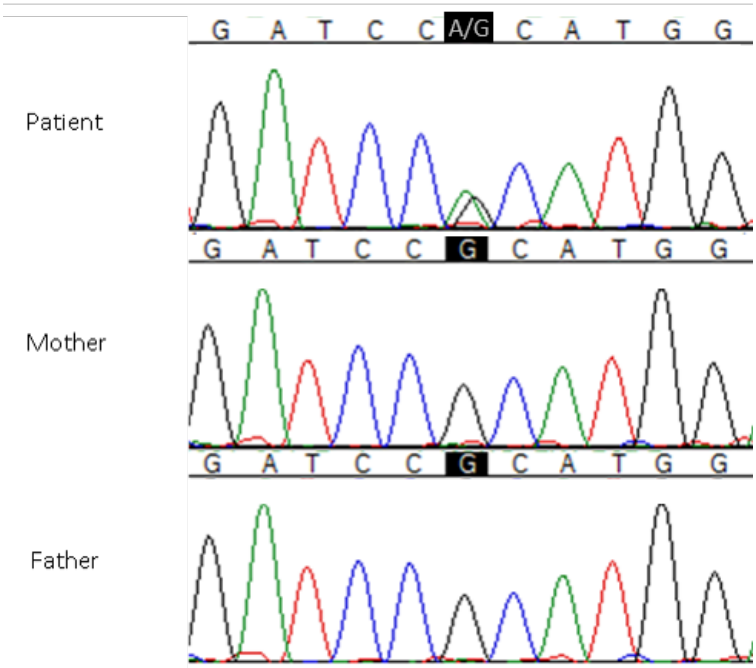
Table 4.5.4: Whole exome sequencing data was filtered to show autosomal dominant and *de novo* variants. Chr, chromosome; n/a, not applicable due to both SIFT and PolyPhen-2 only predicting functional impact for simple variants that change the protein sequence; -, not stated in Ingenuity Variant Analysis software.

Chr	Position	Reference Allele	Sample Allele	Gene	Protein Variant	SIFT	PolyPhen	dbSNP ID	1000 Genomes Frequency	NHLBI ESP Frequency	ExAC Frequency
1	224621770	CCA	-	*WDR26*	p.Gly25del	n/a	n/a	-	-	-	-
4	3076604	CAGCAGCAG	-	*HTT*	p.Gln38_Gln40del	n/a	n/a	757035717	-	-	-
5	158630629	-	T	*RNF145*	p.Asn13fs*44	n/a	n/a	762761147	-	-	-
9	116854149	G	C	*KIF12*	p.Pro512Ala	Deleterious	Probably damaging	530640960	0.26	-	0.32
10	6262778	C	G	*PFKFB3*	p.Leu241Val	Tolerated	Benign	758281350	-	-	-
19	55879672	C	T	*IL11*	p.Arg112His	Tolerated	Probably damaging	4252548	1.2	1.67	4.99
20	62076073	C	T	*KCNQ2*	p.Arg210His	Damaging	Probably damaging	-	-	-	-
22	36894150	C	T	*FOXRED2*	p.Ala424Thr	Tolerated	Benign	149345662	-	0.02	0.01
22	38383171	G	A	*BAIAP2L2*	p.Pro407Ser	Tolerated	Benign	142001534	2.16	2.44	3.75
22	38483174	A	T	*BAIAP2L2*	p.Ser406Thr	Tolerated	Benign	78489217	1.68	-	1.81

The remaining four genes in which variants have been identified have been associated with human disease, three of which (*PFKFB3*, *IL11* and *HTT*) were not consistent with the clinical phenotype of patient X. *PFKFB3* encodes a bifunctional enzyme with both 6-phosphofructo-2-kinase and fructose-2,6-biphosphatase activity. This protein is required for cell cycle progression and prevention of apoptosis. It functions as a regulator of cyclin-dependent kinase 1, linking glucose metabolism to cell proliferation and survival in tumor cells. Hence increased PFKFB3 activity has been associated with multiple types of cancer (Chen *et al.*, 2016). Although mutations in *IL11* have not been associated with disease in humans, this interleukin has been reported to have diverse roles including effects in osteoclastogenesis, neurogenesis, adipogenesis, promotion of stem cell development, haematopoiesis, immunological activity and anti-inflammatory pathways. As such, altered levels of this protein and/or its potential as a treatment candidate in multiple diseases such as haemophilia A, Crohn's disease, rheumatoid arthritis and multiple types of cancer have been reported (Negahdaripour *et al.*, 2016). Finally, the *HTT* gene encodes huntingtin, a protein in which mutations cause Huntington's disease. This disorder is characterised by progressive neurodegeneration with chorea, dystonia, incooridination, cognitive decline and behavioural difficulties with onset typically in mid- to late-adulthood. Disease is caused by an expansion of a trinucleotide repeat $(\mathrm{CAG})_n$ encoding a string of glutamine residues within the protein. In normal individuals the number of glutamine residues ranges between 9 - 36, whereas affected patients have more than 37. Although juvenile forms of Huntington's disease exist, they are associated with greater degrees of expansion (typically $>$ 60 repeats) (Quarrell *et al.*, 2013). Therefore, this is highly unlikely to be the cause of disease in patient X given that the identified mutation results in the deletion of three glutamine residues.

The only gene in which mutations are known to cause a seizure phenotype was *KCNQ2*, encoding a voltage-gated potassium channel that is expressed in the brain. *De novo* mutations in *KCNQ2* are known to cause a range of epileptic disorders including neonatal epileptic encephalopathy and benign familial neonatal seizures. The variant identified in our patient (p.Arg210His) is a known pathogenic mutation that has been reported previously in four patients to cause a severe epileptic encephalopathy (Weckhuysen *et al.*, 2013; Numis *et al.*, 2014). Sanger sequencing of the affected region of *KCNQ2* confirmed the presence of the mutation in the patient and its absence in both parents, thus confirming the expected *de novo* inheritance pattern (Figure 4.5.2).

Figure 4.5.2: *KCNQ2* sequence analysis of affected family using Sanger sequencing. Patient X shows a heterozygous missense change from G to A at position 629 of the cDNA (c.629G>A) causing a change of amino acid 210 from arginine to histidine (p.Arg210His).



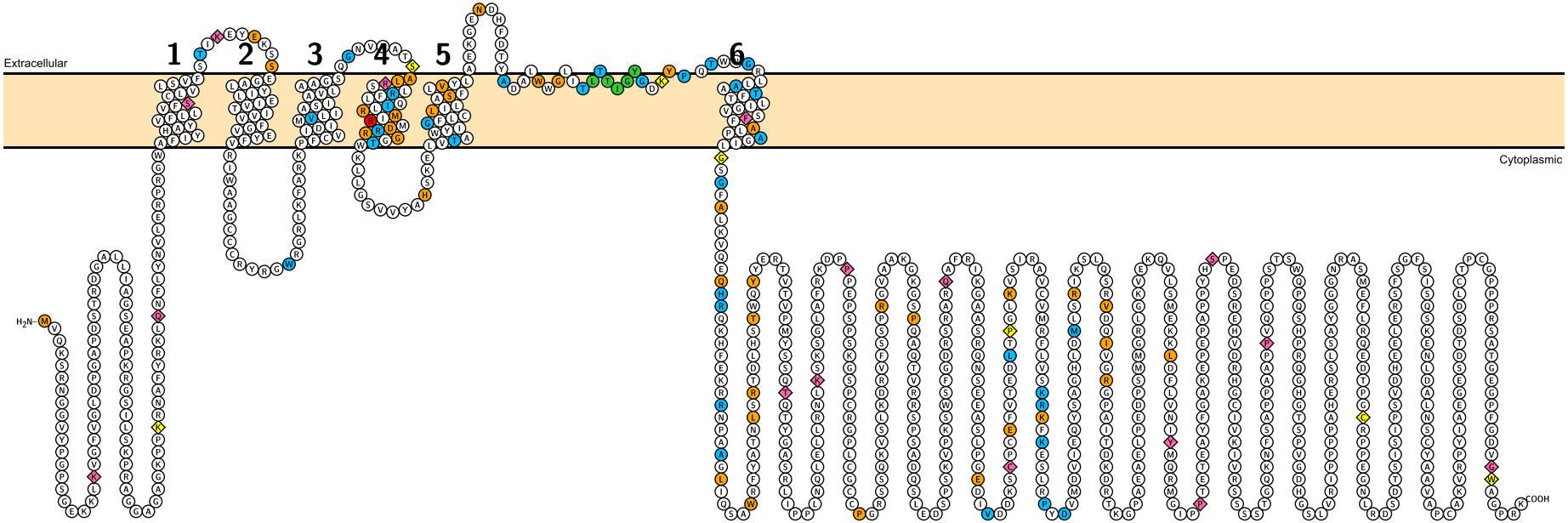
4.5.4 *Function of KCNQ2*

KCNQ2 encodes a voltage-gated potassium channel expressed in the brain (Biervert *et al.*, 1998) and plays a critical role in determining response to synaptic inputs and sub-threshold electroexcitability of neurons. Typical features of voltage-gated potassium channels include six transmembrane domains (S1-S6) and a pore loop between S5 and S6 (Hartmann *et al.*, 1991) (Figure 4.5.3). This pore loop domain contains a conserved sequence of eight amino acids (T/SxxTxGxG) which confers selectivity for potassium ions (Miller, 2000). There is also a long intracellular C terminus which contains highly conserved domains important for correct channel tetramerisation, stability, trafficking and insertion into the neuronal membrane (Haitin and Attali, 2008). The Arg210 residue which is mutated in patient X lies within the transmembrane domain (S4) and the implications of this substitution are described in Section 4.5.7. Voltage-gated ion channels contain positively charged residues that can move within the cell membrane, thus sensing the electric field within it. In turn, these movements then produce a current forcing a conformational change of the channel and opening or closing a gate that controls the flow of ions through the pore (Cha *et al.*, 1999). This voltage-sensing region lies within the S4 transmembrane domain and consists of several well-conserved, evenly spaced arginine and lysine residues with mostly non-polar amino acids between them. Site-directed mutagenesis studies have shown that

a reduction of the net positive charge within the voltage-sensing region causes a reduction in gating charge and in the steepness of the potential dependence of activation (Stühmer *et al.*, 1989).

All known potassium channels are tetramers of identical or homologous α-subunits. KCNQ2 and its homologue KCNQ3, form a single functional heteromeric potassium channel that facilitate currents 11-fold higher than either homomeric channel (i.e. channels formed of either four KCNQ2 or KCNQ3 monomers) (Singh *et al.*, 2003) and assemble in a 1:1 stoichiometric ratio (Hadley *et al.*, 2003). It is these heteromeric and homomeric channels that control a slow potassium current, termed the M-current. The M-current was first discovered in bullfrog sympathetic ganglion cells and regulates the ability of a neuron to fire an action potential (Brown and Adams, 1980). When a neuron is initially polarised the likelihood that KCNQ2/KCNQ3 channels will open increases, leading to the generation of an outward potassium current. This potassium efflux counteracts the sodium influx generated by an action potential and thus, a full action potential is prevented. Inhibition of the M-current by neurotransmitters or drugs leads to neuronal hyperexcitability and subsequently, seizures (Marrion, 1997).

Figure 4.5.3: Schematic representation of KCNQ2. The six transmembrane domains (S1-S6), pore loop domain containing the potassium selectivity sequence and the long intracellular C-terminus are illustrated. Pore loop domain containing the potassium selectivity sequence (green). Mutations were curated using the Human Genome Mutation Database and manual literature searching. Positions at which single nucleotide (i.e. missense or nonsense) mutations are known to cause either benign familial neonatal seizures (orange) or neonatal epileptic encephalopathy (blue), insertions (yellow) and deletions (pink) are highlighted. The mutated arginine at position 210 identified in our patient is also depicted (red). Figure created using Protter (www.wlab.ethz.ch/protter/) (Omasits *et al.*, 2014).



4.5.5 *Phenotypic spectrum of KCNQ2 mutations*

Dominant mutations in *KCNQ2* result in a range of (mostly epileptic) disorders including peripheral nerve hyperexcitability (PNH), neonatal epileptic encephalopathy (NEE) and 60-70% of benign familial neonatal seizures (BFNS) (Weckhuysen *et al.*, 2012).

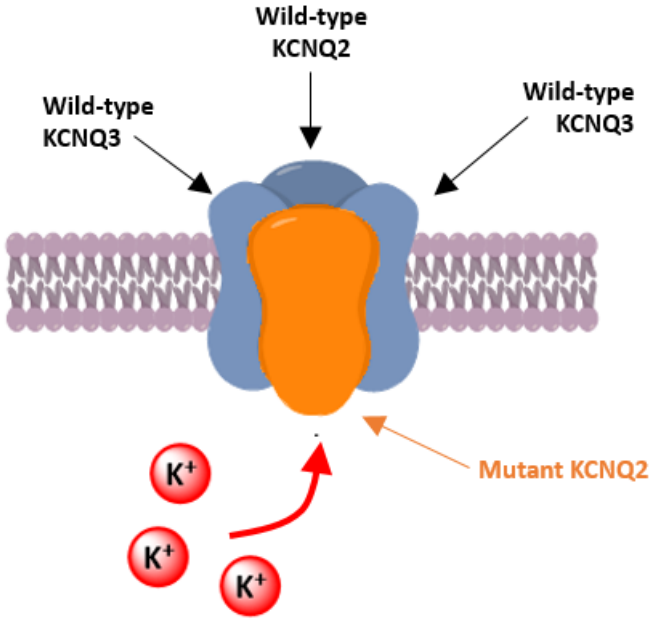
BFNS are characterised by recurrent seizures commencing within the first few days following birth, which are typically tonic-clonic in nature although other seizure types may be observed. Interictal EEG activity is usually normal or demonstrates only mild abnormalities and cognitive development is typically normal. The seizures then resolve spontaneously between one and twelve months of age. However, approximately 15% of patients will be affected by epilepsy later in life (Massingale and Buttross, 1993). A second phenotypically overlapping syndrome, benign familial infantile seizures (BFIS), is differentiated from BFNS by the age of seizure onset occurring after three months of age, although seizure course and developmental outcome is comparable (Franzoni *et al.*, 2005). NEE is a more severe seizure disorder characterised by intractable seizures and a burst-suppression pattern on EEG, evolving to multifocal epileptiform activity with transient T1 and T2 hyperintensities of the basal ganglia. Although seizures may remit by five years of age, the majority of patients suffer neurological impairment and severe psychomotor developmental delay presenting with axial hypotonia and/or spastic quadriplegia. Additionally, a proportion of patients with NEE do not survive their first years of life (Dalen Meurs-van der Schoor *et al.*, 2014; Weckhuysen *et al.*, 2013; Kato *et al.*, 2013). In BFNS, *KCNQ2* mutations are inherited in a classical autosomal dominant pattern from an affected parent or occur *de novo*. Whilst mutations causing NEE are primarily *de novo*, as is the case in patient X (Bellini *et al.*, 1993). However, recently it has been shown that a *KCNQ2* mutation can be inherited from an apparently healthy parent and cause the more severe NEE phenotype in the affected child. In cases such as these, investigations have demonstrated that one parent had mosaicism, with between 5 - 30% of cells carrying the *KCNQ2* mutation (Milh *et al.*, 2015).

PNH is characterised by spontaneous and continuous muscle overactivity in the form of myokymia (continuous undulating movements of distal skeletal muscle), fasciculations, cramps or other symptoms. The underlying aetiology can be heterogeneous, although the most commmon is autoimmune-mediated PNH, where the production of antibodies directed against voltage-gated potassium channels occurs (Hart *et al.*, 2002). Two mutations affecting the same residue in *KCNQ2* (p.Arg207Gln and p.Arg207Trp) have been described to cause myokymia by altering the excitability of peripheral motorneurons with p.Arg207Trp also associated with neonatal seizures (Dedek *et al.*, 2001; Blumkin *et al.*, 2012).

4.5.6 *Location and pathogenic mechanisms of KCNQ2 mutations*

As illustrated in Figure 4.5.3, mutations causing infantile seizures have been identified in all regions of KCNQ2 and there is no clear relationship between the mutation location and whether a child will be affected by BFNS or NEE. In addition to those shown, more than 15 splice site mutations and four large deletions encompassing either entire *KCNQ2* exons or whole genes including *KCNQ2* have also been described (Bellini *et al.*, 1993). Many studies have sought to examine the electrophysiological effects of particular *KCNQ2* mutations, the functional role of each channel region and whether any genotype-phenotype correlations exist. It has been suggested that the clinical severity of disease may be related to the extent of potassium channel impairment. The majority of mutations in *KCNQ2*, particularly those associated with BFNS, cause varying degrees of loss of function of the mutant allele resulting in haploinsufficiency. Indeed, a 20 - 30% reduction in KCNQ2/3 current is sufficient to increase neuronal excitability under physiological conditions (Schroeder *et al.*, 1998; Maljevic *et al.*, 2008) and result in epileptic seizures during the neonatal period. However, a minority of mutations cause a dominant-negative effect on protein function thereby reducing the current through the channel by more than the expected 50%. These effects are caused when the assembly of mutant and wild-type subunits results in the formation of non-functional tetramers.

Figure 4.5.4: Conformation of tetrameric KCNQ2/KCNQ3 channels in patients with a *de novo* mutation in *KCNQ2*.



One study introduced mutations causing either NEE or BFNS into *Xenpus* oocytes and compared their effects using voltage clamping (Orhan *et al.*, 2014). All mutant alleles showed a

loss of function when expressed in a 1:1:2 ratio with wild-type KCNQ2 and KCNQ3 channels, as would be expected in an affected individual (Figure 4.5.4). A dominant-negative effect was found in 5/7 (71%) of the mutations causing NEE that were analysed, compared with only 4 of more than 50 (8%) known BFNS mutations (Orhan *et al.*, 2014), suggesting that this mechanism may underlie the pathogenesis of the severe epileptic phenotype. More specifically, mutations within the voltage-sensing S4 domain such as that in patient X typically cause large shifts in voltage-dependent activation and a slowing of both activation and deactivation kinetics, suggesting that affected channels fail to open in response to depolarisation.

4.5.7 *Clinical comparison of patient X with patients previously reported with a p.Arg210His mutation in KCNQ2*

The evidence that the mutation found in patient X causes epileptic encephalopathy is strong. Four patients have been reported previously to have the same *de novo* p.Arg210His mutation (Weckhuysen *et al.*, 2013; Numis *et al.*, 2014). All had NEE presenting with tonic asymmetrical seizures within the first two days of life accompanied by apnoea and cyanosis. Seizures occurred multiple times per day becoming almost continuous at times; this was so severe in one patient that intubation and transfer to intensive care was required on day five of life. Other seizure types were also observed including: hemi-clonic, asymmetric tonic, tonic-vibratory, seizures with bradycardia and status epilepticus. EEGs at seizure onset revealed multifocal epileptic activity which evolved into a burst-supression pattern in all but one patient, whereas most recent follow-up EEGs (where available) showed slow background activity in all cases. MRI was only available for one case and showed diffuse hypomyelination with marked thinning of the corpus callosum at two months of age. Similarly, patient X presented with seizures characterised by cyanosis and apnoea. Brain MRI revealed hyperintensities of the subthalamic nucleus and geniculate nuclei, similar to other cases, albeit with different mutations, with hyperintensities of the basal ganglia (Weckhuysen *et al.*, 2012). The tonic stiffening and choking seen in the first months of life, associated with asynchronous discontinuity of EEG activities have parallels with the severe electro-clinical phenotype described in NEE. However, in patient X the EEG abnormalities remained mild, with resolution of any discontinuity within two weeks and normal findings during initial treatment with vitamin B_6. All EEG changes were non-specific and no distinctive electroclinical pattern was discernible. As expected for patients with NEE, all children described in the literature were developmentally delayed with severe axial or generalised hypotonia, poor motor skills and some were unable to stand (Weckhuysen *et al.*, 2013; Numis *et al.*, 2014). One patient also had intermittent dystonic posturing and swallowing difficulties

requiring nasogastric tube feeding (Weckhuysen *et al.*, 2013). None of these four patients were trialled on vitamin B_6. In contrast, delay was moderate but largely intellectual in nature with minimal expressive language in patient X. However, she was able to walk independently and attend a school for children with special needs. It is impossible to know whether this somewhat milder developmental delay is the result of the prolonged vitamin B_6 supplementation in patient X.

All children were trialled on multiple AEDs with only minimal or temporary response (Weckhuysen *et al.*, 2013; Numis *et al.*, 2014). Similarly to patient X, two became seizure-free on administration of carbamazepine at two months and four months of age. In another, a combination of valproate, topiramate, clobazam, dexamethasone and the ketogenic diet resulted in seizure freedom at four months of age (Weckhuysen *et al.*, 2013). The fourth patient showed no improvement after ten days of carbamazepine treatment and died shortly afterwards due to respiratory failure in the context of infection (Numis *et al.*, 2014). Carbamazepine acts to stabilise the inactive state of voltage-gated sodium channels and blocks the movement of sodium ions through the channel during an action potential. The resulting decrease in neuron excitability prevents the development of seizure activity until the AED dissociates. Although seemingly unrelated, voltage-gated sodium channels co-localise with KCNQ potassium channels at critical locations in neuronal membranes including the axonal initial segment and node of Ranvier in the hippocampus, cerebral and cerebellar cortex, the ventral horn and the sciatic nerve (Pan *et al.*, 2006). Thus it has been proposed, and recently confirmed (Nguyen *et al.*, 2012), that modulation of sodium channels can alter the activation kinetics and voltage dependence of KCNQ channels.

Further evidence of the pathogenicity of this mutation is the recent identification of a second mutation affecting the same residue (p.Arg210Cys) in a seven year old female described to have KCNQ2-related epileptic encephalopathy (Mercimek-Mahmutoglu *et al.*, 2015). From the limited phenotypic information available, her presentation appears atypical with seizure onset at three years of age. Seizures were of generalised tonic-clonic and absence type and she was affected by global developmental delay, hyperactivity and aggressive behaviour. EEG showed generalised spike-and-waves, MRI showed thin white matter and magnetic resonance spectroscopy was normal. Although this patient, along with the others described with a mutation affecting this residue, are affected by the more severe epileptic encephalopathy phenotype, the age of seizure onset in this case was much later. A possible explanation for this is the different physical properties of each amino acid involved.

Indeed, two different mutations affecting another arginine residue in close proximity to Arg210, also within the voltage-sensing S4 domain, can result in either a mild or severe phenotype. Substitution with a polar uncharged residue (p.Arg213Gln) showed a much more severe functional

defect when compared to that with a bulky aromatic residue (p.Arg213Trp) which is associated with BFNS (Miceli *et al.*, 2013). In the case of variants affecting the Arg210 residue, it appears that substitution with a positively charged aromatic residue (histidine) may be more damaging than the smaller thiol group (cysteine), despite both amino acids being unfavoured in transmembrane α-helical structures (Pace and Scholtz, 1998). We can thus hypothesise that maintaining the structural integrity and charge interactions of the S4 domain is critical for correct channel function and that the Arg210 residue plays a crucial role in voltage-sensing.

4.5.8 *Other patients with KCNQ2 mutations showing a response to vitamin B_6 treatment*

To our knowledge no metabolic abnormalities have been documented in patients with *KCNQ2* mutations. One reason for this may be that the majority of these children present to neurology clinics, where mild abnormalities, for example the amino acid variations seen in our patient, would be reported as normal because results are not dramatically different from the reference ranges. Indeed these results may be irrelevant and were, in addition to mildly elevated 2-oxoglutarate and pyruvate levels, proposed to be due to an age-associated renal tubule immaturity. Serum hepatic enzymes were also elevated, although this is a common finding in children taking AEDs with one study finding 50% of patients taking carbamazepine having elevated alkaline phosphatase levels (Hussein *et al.*, 2013). However, the abnormal vitamin B_6 metabolite levels and apparent response to pyridoxine treatment in patient X are novel findings and more difficult to explain. Furthermore, evidence suggests that this apparent clinical improvement in response to vitamin B_6 treatment in the neonatal period and early infancy has not led to a dependency on supplementation for seizure control later in life. Indeed, since genetic diagnosis, weaning of pyridoxine has commenced and her dose has been halved with no increase in seizures.

In order to determine whether vitamin B_6 had been reported as being beneficial for other patients with mutations in *KCNQ2*, a comprehensive literature review of reports indexed in PubMed describing patients with *KCNQ2* mutations was performed using the terms "KCNQ2" and "epilepsy". Ten reports were identified detailing 23 patients with *KCNQ2* mutations that have been trialled on vitamin B_6, either transiently or on a long-term basis (Appendix 9.3). Three of the 23 patients were reported to have had a clinical response to varying degrees. The first, a patient with a 1.5 Mb terminal deletion of the long arm of chromosome 20 which encompasses *KCNQ2*, carried a clinical diagnosis of pyridoxine-dependent epilepsy for seven years (Mefford *et al.*, 2012). This patient suffered from tonic seizures from two weeks of age accompanied by hypsarrhythmia on EEG whilst on treatment with phenobarbital. Upon intravenous administration of 100 mg of pyridoxine, a 95% reduction of epileptiform activity was observed within one minute, leading to

a diagnosis of PDE. At seven years of age he remained on 200 mg pyridoxine monotherapy and free of clinical seizures but affected by severe developmental delay. The second was a child with benign familial neonatal seizures who presented at four days of age with clonic and tonic seizures. At one year of age, the patient was developmentally normal and although suffered sporadic minor breakthrough seizures, was well-maintained on levetiracetam. However, the patient was also reported to have been treated acutely with pyridoxal 5'-phosphate in the initial period (Allen *et al.*, 2014). Finally, a patient with neonatal epileptic encephalopathy commencing at two days of age with generalised tonic-clonic seizures and multifocal epileptic activity on EEG was reported, in which a combination of topiramate, vigabatrin and pyridoxine controlled seizures (Weckhuysen *et al.*, 2012).

There is very little phenotypic information regarding the three patients that showed a clinical response. There was no similarity in mutations with one chromosomal deletion encompassing the entirety of *KCNQ2*, one insertion in the transmembrane S2 domain (p.Val143_Arg144insGlnTyrPhe Val) and one missense mutation in the transmembrane S4 domain (p.Ile205Val). Developmental outcome was also variable. One patient had BFNS which ceased at one year of age and neurodevelopmental examination remained normal (Allen *et al.*, 2014). One had NEE and was with moderate mental retardation, with only the ability to follow two commands at eight years of age (Weckhuysen *et al.*, 2012). The final patient whose seizures were controlled on PN monotherapy was severely developmentally delayed with minimal expressive language, incontinence and blindness (Mefford *et al.*, 2012). Additionally, 6/23 of the patients were treated with vitamin B_6 during the first month of life; however, responses to each AED were not stated. Whilst in the majority of patients identified no clinical improvement was noted, this may be related to the length of trial they received and other AEDs that were being taken concurrently. Current guidelines recommend that on an acute basis in individuals experiencing clinical seizures, 100 mg of PN should be administered intravenously whilst monitoring the EEG, oxygen saturation and vital signs for several hours. If no clinical response is demonstrated, this dose should then be repeated up to a maximum of 500 mg. Alternatively, a trial can be performed by administering 30 mg/kg/day of PN orally for at least one week (Gospe, 1993). There is also increasing evidence that vitamin B_6, given either as PN or PLP, can result in improved seizure control in idiopathic epilepsy (Ohtahara *et al.*, 2011). However, the mechanisms underlying this response in clinically and genetic diverse patients are currently unknown and will be explored in Sections 4.5.9, 4.5.10 and 4.5.11.

4.5.9 *PLP is required for the synthesis of inhibitory neurotransmitters*

PLP, either administered directly or formed within the liver through the metabolism of PN by pyridoxal kinase and pyridox(am)ine 5'-phosphate oxidase, may act as a general anticonvulsant by promoting the synthesis of inhibitory neurotransmitters. It is well known that uncontrolled or abnormal alterations in the ratios of neurotransmitters in the brain can affect the propensity for seizure activity. Excitatory and inhibitory neurotransmission are largely controlled by glutamate and γ-aminobutyric acid (GABA), respectively. Indeed, GABA release at inhibitory synapses plays an integral role in suppressing both the origin and spread of seizure activity. As such, drugs which enhance the activity of GABA receptors, inhibit GABA degradation or the uptake of GABA from the extracellular space have antiepileptic properties (Treiman, 2001).

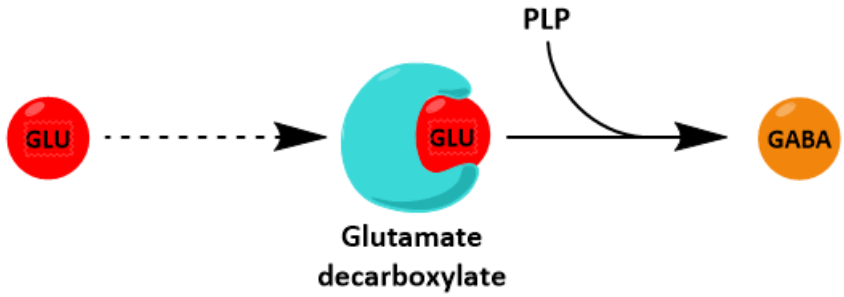
There are three types of GABA receptors, $GABA_A$, $GABA_B$ and $GABA_C$ (Briggs *et al.*, 2011). Both $GABA_A$ and $GABA_C$ are ionotropic receptors activated by the binding of GABA and allow the flow of chloride and bicarbonate to cross the cell membrane. In contrast, $GABA_B$ receptors are metabotropic and linked to potassium channels by G-proteins; the binding of GABA results in an efflux of potassium ions across the cell membrane. Although acting through different mechanisms, all function to hyperpolarise neurons, preventing opening of voltage-gated sodium channels and generation of action potentials. Thus, voltage-dependent calcium channels do not open and neurotransmitters are not released.

There is considerable evidence that seizures can alter GABAergic signalling and result in the depletion of GABA. CSF GABA concentrations are significantly lower in patients with intractable seizures and to a greater extent in those with generalised tonic-clonic and complex partial seizures compared to those with simple partial seizures (Wood *et al.*, 1979), indicating a correlation with disease severity. Reduced levels of GABA within the brain have also been associated with worse seizure control (Petroff *et al.*, 1996). Loss of GABAergic neurons is often seen in focal epilepsies and has been recapitulated in many inducible mouse models (Fritschy *et al.*, 1999; Ni *et al.*, 2005). Experiments in rats have also shown that seizures reduce the amplitude of GABA inhibitory postsynaptic currents (Isaeva *et al.*, 2009), abolish normal receptor subunit expression (Laurén *et al.*, 2005) and alter the expression of cation-chloride cotransporters altering the effect of $GABA_A$ receptor activation from hyperpolarising to depolarising (Li *et al.*, 2008).

The presence of a pathogenic *KCNQ2* mutation results in the impairment of the M-current, leading to a state of neuronal hyperexcitability and the propagation of seizures. This would therefore be predicted, as in other epileptic disorders, to result in a depletion of GABA and/or other electrophysiological changes that result in ineffective inhibitory GABAergic signalling. GABA is synthesised from glutamate by the enzyme glutamate decarboxylase (GAD) for which

PLP is a cofactor (Figure 4.5.5). Inhibition of this enzyme by PLP-γ-glutamyl hydrazone has been shown to result in fatal convulsions that could be rescued by administration of PLP (Tapia and Pasantes, 1971). Homozygous knock-outs of GAD in mice have increased susceptibility to chemically induced seizures (Asada *et al.*, 1996) or develop spontaneous seizures which are easily precipitated by mild stress (Kash *et al.*, 1997), depending on the genetic background of the mice. *In vivo* studies of GAD from human brain tissue demonstrate that both isoforms (GAD_{65kDA} and GAD_{67kDa}) show increased enzymatic activity with increasing concentrations of PLP. Taken together, it is plausible that supplementation of patients with either PN or PLP may favour the conversion of glutamate to GABA in the brain, leading to anticonvulsant effects.

Figure 4.5.5: Glutamate decarboxylase converts glutmate to form GABA using PLP as a cofactor.



4.5.10 *PLP as an ion channel antagonist*

In addition to PLP being an essential cofactor for enzymes involved in neurotransmitter metabolism, recent work has also shown that PLP may act directly on ion channels. Indeed, synthetic vitamin B_6 derivatives and PLP itself have been found to be effective antagonists of P2X receptors. These receptors belong to the P2 receptor family which were identified in 1978 and found to be activated by the binding of adenosine 5'-triphosphate (ATP) (Burnstock, 2007). Within this family there are two sub-classes; the ionotropic P2X and metabotropic P2Y receptors, containing seven and eight subtypes, respectively. P2Y receptors are G protein-coupled, activated by the binding of ATP or uridine 5'-triphosphate (UTP) and mediate cell proliferation, differentiation and death during development and regeneration (Burnstock, 2007). P2X receptors are cation-permeable ion channels which allow the rapid and non-selective passage of sodium, potassium and calcium ions resulting in neuron depolarisation in response to the binding of ATP. Amongst the P2X class of receptors, there is one receptor ($P2X_7R$) which has been implicated in the pathogenesis of epilepsy (Engel *et al.*, 2012; Henshall *et al.*, 2013). The activation of $P2X_7R$ occurs only when ATP concentrations are in the mM range as, unlike other members of the

P2X family, it has a relatively low affinity for the ligand (Engel *et al.*, 2012). However, these extremely elevated extracellular concentrations can occur when ATP is released from damaged or overstimulated cells during seizure activity (Wieraszko and Seyfried, 1989). Prolonged $P2X_7R$ activation causes increased neuronal excitability, inflammation, oxidative stress and apoptosis through the action of diverse pathways. In addition, the activation of this particular receptor has been shown to promote the assembly of a pore within plasma membranes permeable to hydrophilic molecules up to 800 Da in size. Although reversible, this pore formation has been hypothesised to be involved in mediating the cytotoxic effects of $P2X_7R$ activation (Sperlágh *et al.*, 2006). Studies have demonstrated that $P2X_7R$ knockout mice have reduced seizure severity compared to wild-type animals subjected to chemical-induced seizures (Solle *et al.*, 2001). Furthermore, in mice affected by refractory seizures, a combination of treatment with both lorazepam and a $P2X_7R$ antagonist controlled seizures, when neither could solely achieve the same result (Engel *et al.*, 2012).

The two B_6 derivatives based on the synthetic modification of PLP found to be effective antagonists of these receptors are pyridoxal phosphate-6-azophenyl-2-4-disulfonic acid (PPADS) and pyridoxal-5'-phosphate-6-(2'-naphthylazo-6'-nitro-4',8'-disulfonate) (PPNDS). The former is a non-selective antagonist which blocks P2X and P2Y receptors with a similar potency (North and Surprenant, 2000). The latter is selective for $P2X_1$ receptors and estimated to be seven-fold more potent than PPADS (Lambrecht *et al.*, 2000). PLP has been shown to be as effective at inhibiting $P2X_7R$ activity as PPADS, requiring a concentration of 10 μmol/L to induce 50% inhibition. Two other studies showed that PLP could block P2X receptor activation in the rat vagus nerve, vas deferens (Trezise *et al.*, 1994) and cardiomyocytes (Wang *et al.*, 1999). Although PLP levels are tightly controlled in all cells (including those of the brain), individuals on pyridoxine supplementation have gross elevations of multiple B_6 vitamers in their CSF. CSF PLP levels are also typically at the upper end of reference ranges in these patients, regardless of whether the indication for supplementation is PDE, PNPO deficiency or a disease of unknown aetiology (Jaeger *et al.*, 2016; van der Ham *et al.*, 2012). Indeed, a neonate with an unknown B_6-responsive disorder was reported to have CSF concentrations of 3776 nmol/L pyridoxal (ref: 14.8 - 42.5), 29.6 nmol/L pyridoxamine (ref: 0.1 - 0.5) and 18,881 nmol/L PN (ref: <0.03) whilst on treatment with 30 mg/kg/day PN (van der Ham *et al.*, 2012). These vitamers can then be converted to PLP by the action of pyridoxal kinase and pyridox(am)ine 5'-phosphate oxidase. Thus, supplementation with PN results in a sustained increase in PLP levels within the brain, may result in the curtailing of $P2X_7R$ activation caused by seizure-induced ATP production and subsequently the clinical improvement seen in patient X. Finally, it is possible that the action of PLP is not only limited to P2X receptors and to date, no study has investigated whether PLP

may modulate the KCNQ family of potassium channels. Further work is warranted to investigate the potentially widespread action of vitamin B_6 on ion channels.

4.5.11 *Oxidative stress in the propagation of epilepsy and PLP as an antioxidant*

A final potential mechanism which may underlie the generalised anticonvulsant effect of vitamin B_6 is the role of these compounds as antioxidants. Although the contribution of oxidative stress to disease pathogenesis is established in many neurological conditions such as Parkinson's disease, Alzheimer's disease and amyotrophic lateral sclerosis (Hwang, 2013; Markesbery, 1997; Barber *et al.*, 2006), the role of oxidative stress resulting from excessive free-radical production in epilepsy initiation and propagation has only recently been recognised.

Oxidative stress is caused by an imbalance between the production of damaging reactive oxygen species (ROS) and the antioxidant systems which act to defend against them. ROS are chemically reactive oxygen-containing molecules which are generated during normal cellular metabolism (e.g. superoxide radicals ($\cdot O_2^-$), hydrogen peroxide (H_2O_2), hydroxyl radicals ($\cdot OH$) and singlet oxygen ($^1O_2$)) and cause cellular damage through the chemical modification of various biomolecules such as proteins, lipid and nucleic acids which ultimately result in cell death. Protein oxidation leads to functional changes of a number of enzymes, often rendering them less catalytically active, more sensitive to heat inactivation and with an altered susceptibility to proteolytic degradation (Stadtman, 2001). Lipids, particularly unsaturated lipids which contain multiple double bonds and reactive hydrogen atoms, are easily susceptible to peroxidation. Peroxidation of phospholipids within lipid bilayers has been shown to induce conformational and electrostatic changes which both increases the area per lipid and reduces the thickness of the bilayer. Together this increases membrane permeability and reduces membrane integrity (Wong-Ekkabut *et al.*, 2007). Finally, DNA can be oxidised by ROS; this occurs most readily at guanine residues forming 8-hydroxy-2-deoxyguanosine. Consequences of this oxidation include a propensity for mutagenesis if the oxidised residues are not correctly repaired by base excision repair, and the possibility of epigenetic alterations resulting in the inappropriate repression or activation of genes (Mikhed *et al.*, 2015).

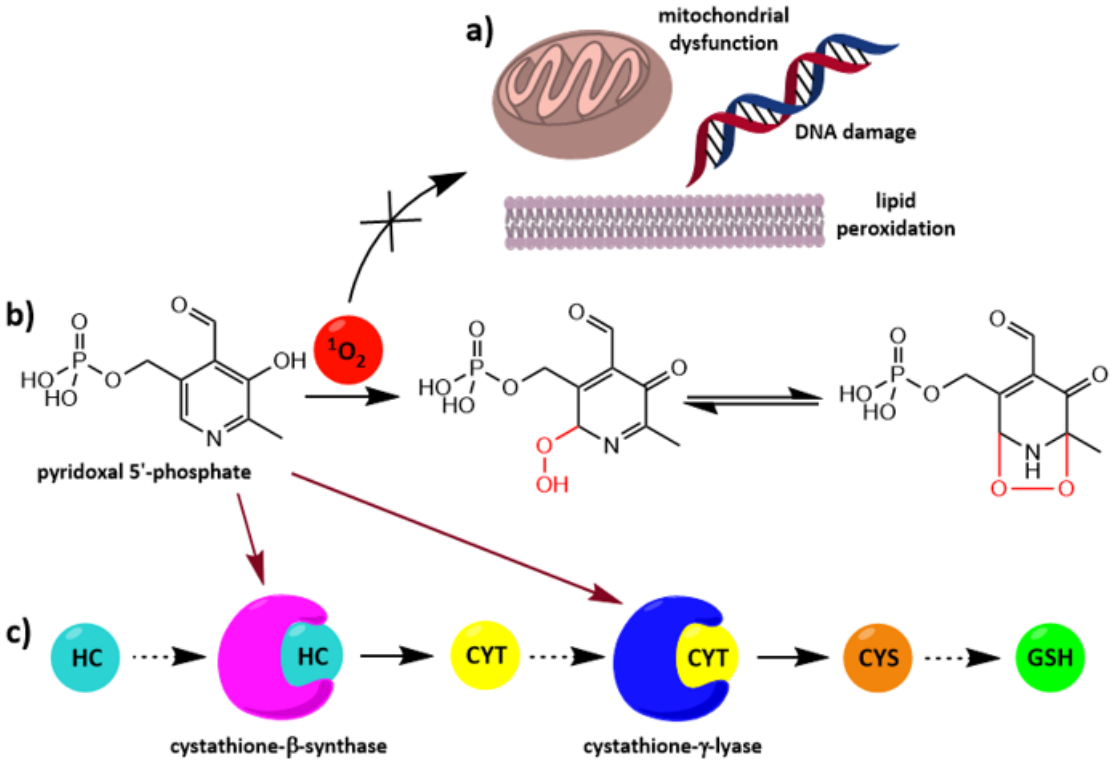
The antioxidant defence systems including enzymes (e.g. superoxide dismutase, catalase, glutathione peroxidase, glutathione reductase and peroxiredoxins) and non-enzymatic molecules (e.g. vitamin C, vitamin E and reduced glutathione) (Matés, 2000) have been well documented. However, more recently the antioxidant properties of vitamin B_6 are becoming increasingly recognised (Figure 4.5.6). Multiple vitamers including pyridoxine, pyridoxamine and PLP are able to scavenge superoxide radicals induced by hydrogen peroxide treatment, reduce lipid

peroxidation and prevent ROS-induced changes in mitochondrial transmembrane potential in U937 monocytes (Kann and Kovács, 2007). Analogous experiments were also performed in endothelial cells, yielding similar results (Mahfouz *et al.*, 2009). B_6 vitamers have also been shown to prevent the death of yeast cells from oxidative damage, with PLP, pyridoxamine 5'-phosphate and pyridoxamine showing a higher antioxidant activity than vitamin C (Chumnantana *et al.*, 2005), detoxifying superoxide and/or singlet oxygen via reactivity of the pyridine ring and phenolic hydroxyl group (Ohta and Foote, 2002) (Figure 4.5.6). However, the quenching of ROS comes at a price, with collisions between B_6 vitamers and singlet oxygen resulting in vitamer degradation (Natera *et al.*, 2012). A similar finding has been shown for folates, which can also act as antioxidants with depletion upon collision with ROS being found to occur (Rezk *et al.*, 2003). Indeed, ROS-induced degradation of PLP and folate may explain the low concentrations of PLP and 5-methyltetrahydrofolate seen in the CSF of patient X. If excessive ROS production due to unregulated neuronal firing results in the CSF PLP deficiency seen in our patient, it is intuitive that PN supplementation should correct this abnormality.

In addition to direct quenching of ROS, PLP functions as a cofactor for more than 140 enzyme-catalysed reactions in the human body. One example is the synthesis of cysteine from homocysteine which requires the action of two PLP-dependent enzymes, cystathione-β-synthase and cystathione-γ-lysase (Meier *et al.*, 2001) (Figure 4.5.6). Cysteine is a unique proteinogenic amino acid which is susceptible to oxidation within proteins to form disulphide bonds which are important for maintaining protein secondary and tertiary structure. In addition, cysteine is essential for the synthesis of reduced glutathione (GSH) which is able to neutralise reactive oxygen species through the donation of a reducing equivalent, resulting in the formation of oxidised glutathione disulphide (GSSG) (Meister, 1994). GSH is then regenerated from GSSG by glutathione reductase. Additionally, GSH is essential for the detoxification of peroxidised lipids and xenobiotics and the reduction of hydrogen peroxide to water by glutathione S-transferase and glutathione peroxidase, respectively (Meister, 1994). Thus, PLP is critical for the optimal function of the glutathione antioxidant defence system. Indeed, the mechanisms described above have been corroborated in animal studies where vitamin B_6 status altered oxidative stress markers. One study in mice fed a diet supplemented with homocysteine thiolactone to induce oxidative stress, found that this effect was exacerbated in mice given a B_6-deficient diet (Hsu *et al.*, 2015).

Although the antioxidant effects of vitamin B_6 in epilepsy have yet to be evaluated, the efficacy of other antioxidants as antiepileptic agents have been repeatedly demonstrated. Research has shown that melatonin, in addition to its role in regulating circadian biology, can also act as a free-radical scavenger and it has been shown that levels are reduced in children with epilepsy (Ardura *et al.*, 2010). Treatment with this indoleamine prevents seizure activity, promotes glutathione

Figure 4.5.6: Function of vitamin B_6 as an antioxidant. **(a)** Oxidative stress can result in mitochondrial dysfunction, DNA damage and lipid peroxidation, which can be ameliorated by the quenching of ROS by vitamin B_6. **(b)** Pyridoxal 5'-phosphate (and the other B_6 vitamers) can react with singlet oxygen (red) which leads to addition at C5 and proton transfer. The equilibrium partner is then formed by solvolysis or direct solvent addition leading to hydrogen migration and elimination of water. **(c)** Pyridoxal 5'-phosphate can also act as a cofactor for cystathione-β-synthase which catalyses the conversion of homocysteine (HC) to form cystathione (CYT), as well as cystathione-γ-lysae which catalyses the conversion of CYT to form cysteine (CYS). The production of cysteine is essential for the synthesis of reduced glutathione (GSH).

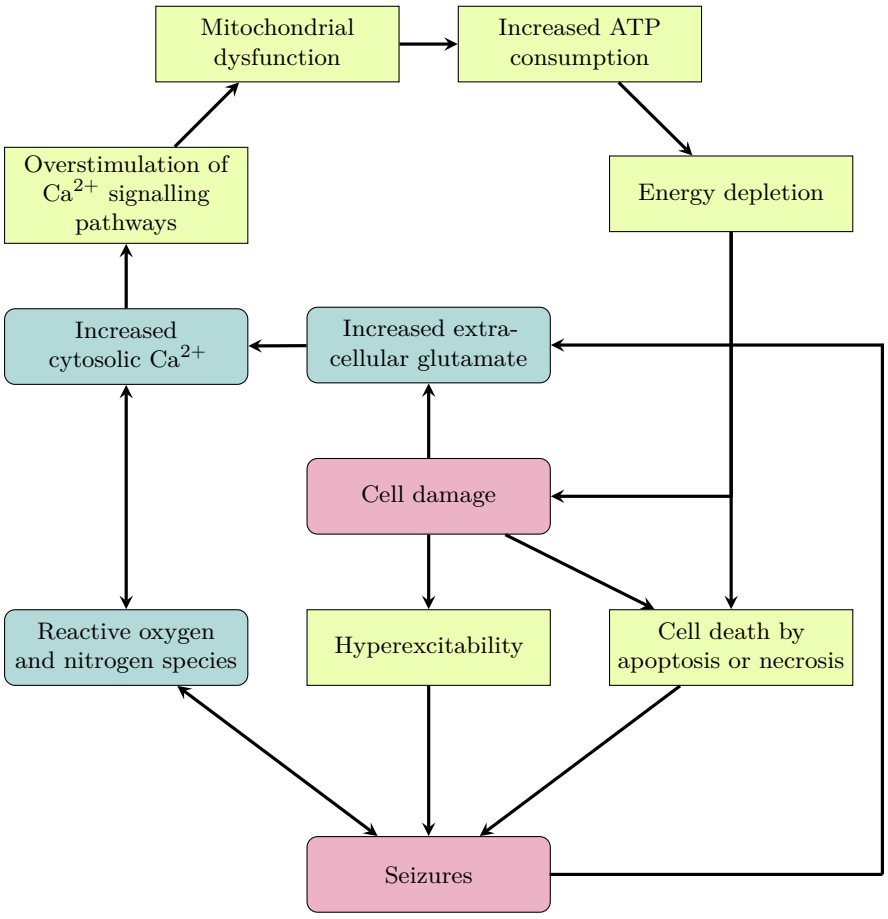


function and suppresses ROS production, lipid peroxidation and nucleotide oxidation (Mohanan and Yamamoto, 2002). Dietary intake of other endogenous antioxidants such as lipoic acid also produced similar results with the additional benefit of an increase in antioxidant enzyme activities (Rochette *et al.*, 2013). Ascorbate and α-tocopherol are the active forms of vitamin C and E, respectively. Both are dietary antioxidants whose mechanism of action is not fully understood and antiepileptic potential are somewhat controversial. Although ascorbate supplementation ameliorates convulsions and has neuroprotective effects, studies have indicated that when iron concentrations are high (e.g. an abundance of haemoglobin following a haemorrhage) ascorbate promotes oxidation (Sadrzadeh and Eaton, 1988). Effects of α-tocopherol include preventing damage to the blood-brain-barrier through the inhibition of lipid peroxidation (Levy *et al.*, 1992).

Fundamental to the utilisation of antioxidants as adjunctive treatments is the understanding of the paradigm of how ROS contributes to seizure pathogenesis. The brain is highly susceptible

to oxidative stress because it contains high levels of polyunsaturated fatty acids, high iron concentrations which catalyse the formation of hydroxyl radicals and a low catalase activity which normally functions to eliminate hydrogen peroxide (Mariani *et al.*, 2005). Additionally, mitochondrial oxidative phosphorylation is the principle source of cellular superoxide; thus high ATP requirements to maintain ionic gradients across neuronal membranes and for neurotransmission are counterbalanced by concurrent high ROS levels (Kann and Kovács, 2007). It has been hypothesised that a cascade of events following seizures produces a self-perpetuating vicious cycle of oxidative stress (Figure 4.5.7).

Figure 4.5.7: Proposed mechanism of reactive oxygen and nitrogen species generation and cellular damage following epileptic seizures.



Excessive neuronal firing as occurs during seizure activity, results in the release of glutamate (the major excitatory neurotransmitter in the brain) into the intracellular space. This glutamate then results in the activation of NMDA (N-methyl-D-aspartate) and AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid) receptors and an influx of calcium ions into neurons (DeLorenzo *et al.*, 1998). One of the many roles of mitochondria is to serve as a calcium ion buffer to maintain cellular homeostasis. The increase in calcium concentration within the mitochondria following seizures results in the activation of enzymes within the tricarboxylic acid cycle (McCormack and

Denton, 1993), increased reduced substrates for oxidative phosphorylation (NADH and $FADH_2$), increased respiratory chain activity and increased ROS generation. High levels of cellular calcium can also activate other enzymes such as nitric oxide synthase which result in the production of reactive nitrogen species (closely related to ROS and causing similar oxidative stress) (Alderton *et al.*, 2001). These species can inhibit complex IV activity resulting in the inappropriate transfer of electrons to oxygen, as well as promote further synaptic glutamate release (Brookes *et al.*, 2004). Calcium overload also triggers the activation of mitochondrial permeability transition pores causing the membrane to become permeable to molecules smaller than 1.5 kDa and the loss of molecules such as cytochrome c. This can then result in cell death by apoptosis or necrosis (Zoratti and Szabò, 1995).

Many studies in rat and mouse models of epilepsy have shown neuronal loss, mitochondrial dysfunction, decreased reduced glutathione, increased lipid peroxidation, protein oxidation (Shin *et al.*, 2008) and DNA damage (Tang *et al.*, 1998) following seizures. Increases in the binding of oxidative stress-activated transcription factors such as NF-κB have also been observed in seizure models (Rong and Baudry, 1996), as have increases in the activities of multiple antioxidant enzymes such as catalase, superoxide dismutase and glutathione peroxidase after chemically-induced seizures (Tejada *et al.*, 2007), suggesting an attempt to counteract ROS overproduction. Evidence of mitochondrial dysfunction and oxidative stress have also been identified in patients with various types of seizures including temporal lobe epilepsy, idiopathic generalised epilepsy, complex partial seizures (Rowley and Patel, 2013; Menon *et al.*, 2012; Rumià *et al.*, 2013).

Finally, as well as seizures themselves causing oxidative stress, there is evidence that treatment with AEDs including valproate, phenytoin, phenobarbital and carbamazepine can exacerbate this phenomenon (Schulpis *et al.*, 2006; Varoglu *et al.*, 2010). Patient X has been on carbamazepine treatment from two months of age; dosage was increased in line with weight gain to a maximum of 15 mg/kg/day with good seizure control. Multiple studies have investigated the effects of carbamazepine and concluded that treatment increases plasma peroxide, lipid peroxidation in plasma and DNA damage markers in leukocytes, whilst it reduces superoxide dismutase, glutathione peroxidase and catalase activity in erythrocytes (Hamed *et al.*, 2004; Aycicek and Iscan, 2007; Varoglu *et al.*, 2010; Niketić *et al.*, 1995). Therefore, it is possible that the treatment of patient X and the general epileptic population with conventional AEDs may promote oxidative stress which may be ameliorated by treatment with vitamin B_6 or other antioxidants. In addition, treatment with PN or PLP improve the clinical outcome of patients by curtailing the cycle of mitochondrial and neuronal dysfunction caused by seizure-induced oxidative stress.

4.6 SUMMARY

In this chapter, the case of a girl whose severe neonatal seizures showed an apparent response to pyridoxine treatment and had an abnormally high plasma:CSF PLP ratio was presented. This case confirms the utility of whole exome sequencing for the diagnosis of childhood epilepsy, revealing that this patient's neonatal seizures and developmental disorder were caused by a *de novo* mutation in *KCNQ2*. However, with the benefit of hindsight, this potassium channelopathy may have been diagnosed using a targeted epilepsy gene panel such as that described by Lemke *et al.* (2012). Targeted gene panel sequencing has advantages over whole exome sequencing, including increased depth of coverage, more robust coverage of regions of interest and reduced incidental findings (de Koning *et al.*, 2015). In this case, whole exome sequencing was employed due to the patient's atypical clinical presentation and biochemical findings; it was hypothesised that she had a novel inborn error of vitamin B_6 metabolism. Indeed, if this approach had not been undertaken, the possibility that patient X harboured mutations or polymorphisms in genes affecting vitamin B_6 metabolism could not have been refuted. However, diagnostic panels remain a timely and cost-effective alternative for the diagnostic work-up of children with neonatal or infantile epilepsies.

Despite the potential benefits of vitamin B_6 as an anticonvulsant, there are well-documented adverse effects that can occur in patients taking high doses. Treatment with PLP can commonly cause nausea, vomiting, diarrhoea, loss of appetite and mild elevations in serum alanine aminotransferase (Ohtahara *et al.*, 2011; Toribe, 2001). Two reports have also documented rare complications including rhabdomyolysis and an increase in seizure frequency (Wang *et al.*, 2005). High-dose PLP can also cause more severe liver dysfunction. A child with PNPO deficiency treated with 100 mg/kg/day PLP due to repeated encephalopathy developed hepatic cirrhosis, fibrosis and portal hypertension; liver function tests improved substantially when his dose was halved but did not entirely normalise (Sudarsanam *et al.*, 2014). Liver toxicity was also noted in a child with homocysteinuria who developed hepatitis following an increase of his PLP dose to 1000 mg/day (Yoshida *et al.*, 1985). In contrast, long-term high-dose PN supplementation can cause sensory and motor neuropathy (Schaumburg *et al.*, 1983). Treatment can also be pro-convulsant in some circumstances; a newborn at risk of PDE due to having an affected sibling was treated prophylactically with PN but developed status epilepticus. Antiquitin deficiency was subsequently excluded and seizures stopped with pyridoxine cessation, therefore seizures were due to pyridoxine toxicity (Hartmann *et al.*, 2011). Ideally patients with intractable epilepsy, including those with channelopathies, should be trialled on vitamin B_6. In those showing a response, liver function

and nerve conduction should be tested periodically. A trial of discontinuation should also be carried out to confirm that there is a real necessity for PN/PLP supplementation.

Finally, we hypothesise that the anticonvulsant effect of B_6 vitamers may be more universal than previously thought and represents a promising adjunctive treatment for patients, not only with channelopathies, but also the wider epileptic population. Whole exome sequencing of patient X, combined with a review of the literature, identified a small group of patients with mutations in *KCNQ2* that were described to have varying degrees of clinical response to vitamin B_6 treatment. It is possible that this response may be due to the mechanisms described in this chapter, including replenishing the pool of PLP needed for the synthesis of inhibitory neurotransmitter, direct antagonist action on ion channels and antioxidant action of excess reactive oxygen species generated by increased and abnormal neuronal firing. Unfortunately, information that may facilitate understanding of the differences between "responders" and "non-responders" such as PN/PLP dosage, length of B_6 trial, age of patient during B_6 trial and other AEDs being taken concurrently are not available. Indeed, with the exception of cases of West syndrome, reports often assume any observed anticonvulsant effect of vitamin B_6 to be coincidental or simply not discussed further. This may be due to the lack of a dramatic response or the low response rate amongst patients. Additional well-controlled trials are needed to determine optimal treatment regimens and identify which patients may benefit from vitamin B_6 treatment.

5

MUTATIONS IN SLC25A22: EXPANDING THE SPECTRUM OF MUTATIONS AND THE CLINICAL AND BIOCHEMICAL PHENOTYPE

5.1 INTRODUCTION

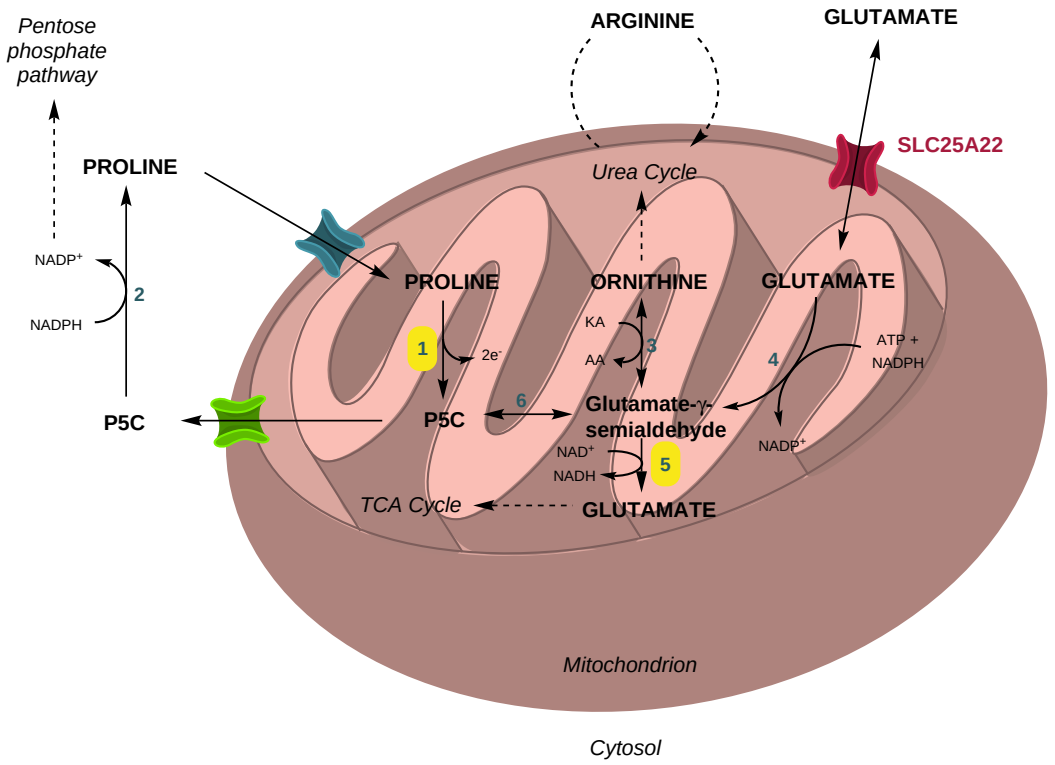
In this chapter, whole exome sequencing was used to investigate the underlying aetiology of disease in two unrelated families comprising a total of three affected patients. The proband in both families presented with features of early infantile epileptic encephalopathy with severe refractory seizures and hypotonia. The first proband was referred to the metabolic clinic because of persistent hyperprolinaemia. In contrast, the two patients from the second family were referred due to fibroblast vacuolation upon ultrastructural examination, of unknown cause and significance. It was also noteworthy that the younger sibling in this family presented with a much milder phenotype, having no history of seizures at the time of presentation.

5.1.1 *Causes of hyperprolinaemia*

Hyperprolinaemia is defined as an excess accumulation of proline in the body. Mutations in *PRODH* and *ALDH4A1* cause hyperprolinaemia type I (HPI) and II (HPII), respectively. *PRODH* encodes proline dehydrogenase, which catalyses the first step in the catabolism of proline within the mitochondria forming Δ^1-pyrroline-5-carboxylate (P5C) (Figure 5.1.1). This can then undergo non-enzymatic hydrolysis of the pyrroline ring to form glutamate-γ-semialdehyde (GSA). *ALDH4A1* encodes P5C dehydrogenase which catalyses the oxidation of GSA to form glutamate (Figure 5.1.1). This reaction is required to connect the urea cycle with the tricarboxylic acid cycle within the mitochondria (Geraghty *et al.*, 1998). A deficiency of either enzyme results in an accumulation of cellular proline which can be detected in blood, urine and CSF. The two disorders can be differentiated biochemically; HPI presents solely with elevated proline levels, whilst HPII is also accompanied by urinary P5C excretion which is undetectable in control samples. The degree of proline accumulation also differs with HPI patients having concentrations between 3- and 10-fold higher than controls, whereas those in HPII patients are between 10- and 15-fold elevated.

The proband in the first family described in this chapter had persistently elevated proline concentrations in plasma, urine and CSF. In contrast, the proband in the second family was

Figure 5.1.1: Primary genetic defects causing hyperprolinaemia. Schematic illustrating the synthesis, degradation and interconversions of proline within the mitochondrion. Mutations in proline dehydrogenase (1; *PRODH*) cause hyperprolinaemia type I and those in P5C dehydrogenase (5; *ALDH4A1*) cause hyperprolinaemia type II. 1, proline dehydrogenase; 2, P5C reductase; 3, ornithine-δ-aminotransferase; 4, P5C synthase; 5, P5C dehydrogenase; 6, non-enzymatic reaction.



not initially reported to have hyperprolinaemia as an indication for whole exome sequencing. However, scrutiny of first-line biochemical screening results following genetic diagnosis also revealed elevated proline levels. Despite this, proline concentrations were not as high as those typical of classical hyperprolinaemia and the early onset and severity of seizures accompanied by ocular abnormalities were atypical for either HPI and HPII; therefore these diagnoses were deemed unlikely.

Hyperprolinaemia can also be secondary to other genetic and environmentally-acquired conditions. An example of the former can be seen in patients with mitochondrial disorders. Despite encompassing more than 250 distinct gene defects, lactic acidosis is present as a biochemical abnormality in between 30 - 50% of mitochondrial disorders (Koenig, 2008). This occurs when the mitochondria's capacity to metabolise pyruvate through the Krebs cycle and oxidative phosphorylation to produce ATP is impaired. To compensate for this reduced ATP synthesis, glycolysis and therefore pyruvate production is increased. The addition of hydrogen ions (resulting from impaired oxidative phosphorylation) to pyruvate then generates lactic acid. Lactic acid can non-competitively inhibit proline dehydrogenase activity resulting in hyperprolinaemia (Kowaloff

et al., 1977). This inhibition, which occurs readily at the concentrations observed in patients with mitochondrial disorders (5 - 10 mmol/L), also causes a secondary decrease in the formation of ornithine, citrulline and arginine due to reduced flux through the proline catabolic pathway with the mitochondria (Dillon *et al.*, 1999). An alternative hypothesis for the elevated proline concentrations seen in patients with mitochondrial disorders is that an abnormal mitochondrial redox state results in an altered reaction equilibrium of the proline dehydrogenase enzyme, thereby affecting intracellular proline concentrations (Wood, 1987).

Proline concentrations can also be high in patients with other conditions in which elevated venous lactic acid can be present including hypoxia, renal failure, drug toxicity, sepsis and neurodegenerative diseases (Kang *et al.*, 2001; Koenig, 2008). A mitochondrial disorder was thought to be the most plausible differential diagnosis in both families, however, mitochondrial respiratory chain analysis was uninformative in the probands from both families. Thus, whole exome sequencing was performed.

5.2 METHODS

Whole exome sequencing was performed as described in Section 2.5 and filtered to look for variants inherited in an autosomal recessive manner. Confirmation of the pathogenic *SLC25A22* mutations were carried out using PCR conditions outlined in Section 2.3.2 with the addition of 5% DMSO. Standard PCR cycling conditions were used with an annealing temperature of 66°C for the region surrounding the p.Thr56Pro mutation and 68°C for that surrounding the p.Ala296Thr mutation. Fibroblasts were cultured from patient skin punch biopsies in the Chemical Pathology Department, GOSH. These were then transferred to our laboratory and cultured as described in Section 2.7 before carrying out electron microscopy, tinctorial and immunohistochemical staining (Section 2.14). Analysis of intracellular glutamate-γ-semialdehyde using UPLC-MS/MS was carried out using the method described in Section 2.15.

5.3 CASE REPORTS

5.3.1 *Patient 1*

A 2 year old boy (patient 1) presented to our clinic for evaluation of epilepsy, severe developmental delay, and hyperprolinaemia. His parents are first cousins of Afghan origin. He was born at term following an uneventful pregnancy with a birth weight of 3.3 kg. At birth he was noted to be non-specifically neurologically abnormal with central and peripheral hypotonia and delayed

feeding. However, at six weeks of age he developed myoclonic and tonic-clonic seizures, dystonic movements and a failure to fix and follow. At two months of age, his main symptoms were severe global developmental delay, visually unresponsive behaviour with impaired cortical responses to light and a myoclonic and generalised tonic-clonic seizure disorder. Brain MRI performed at 2 years of age showed frontotemporal hypoplasia, globally delayed myelination, delayed temporal pole myelination, prominent cerebellar folia and a small splenium. Extensive genetic and biochemical investigations had been carried out (Table 5.3.1). Amino acid profiles were consistently abnormal in plasma, urine and CSF (detailed in Results). Plasma proline was between 322 – 1195 μmol/L (ref: 85 – 290 μmol/L). At the age of 2 years and 8 months, a series of plasma amino acid determinations were done – after a 6 hour fast and at 1h, 2h and 4h postprandially. A muscle biopsy showed increased lipid and glycogen but neither ragged red nor cytochrome oxidase-negative fibres, and normal activities of respiratory chain complexes. Ammonia had also been slightly raised on one occasion (50; ref: <40 μmol/L) and white cell ubiquinone was low (24; ref: 37 – 133 pmol/mg). At the age of 7 years, patient 1's seizures are well controlled on a combination of levetiracetam (40 mg/kg/day), sodium valproate (20 mg/kg/day) and pyridoxine (15 mg bd). At his most recent clinic assessment at the age of 7 years and 11 months, he was in a wheelchair and globally severely developmentally delayed, although he continues to make slow progress. He appeared to have a full range of eye movements but could not fix or follow. He had no dystonic movements, whilst reflexes were elicitable in the upper limbs they were absent in the lower limbs.

5.3.2 *Patients 4 and 5*

Patient 4, a 7-year old girl, presented to our clinic for evaluation of epilepsy and severe developmental delay. Her parents are first cousins of Syrian origin. The patient is the second child of her parents, with a healthy older brother and an affected younger sister. Within the first month of life she developed a febrile illness followed by seizures after the fever had settled. The seizures continued so sodium valproate was commenced, with some clinical improvement. Clonazepam was subsequently added. Her current treatment is sodium valproate 260 mg b.d. and clonazepam 500 μg once at night. She continues to have brief seizures lasting less than a minute at least once a week. These are usually tonic seizures and sometimes occur in clusters several times per day. The patient started crawling and sitting independently at the age of 2 years and 6 months. At 7 years she could babble but had no identifiable words. Developmental progress is very slow although she has never suffered developmental regression. On physical examination she appeared globally hypotonic although the lower limb reflexes were brisk and plantar responses

were equivocal. An EEG at 6 years of age showed symmetrical diffuse irregular slow wave activity. Brain MRI also performed at 6 years demonstrated some symmetrical signal abnormalities of the insular cortex bilaterally and adjacent capsular white matter. Delayed temporal pole myelination, prominent cerebellar folia and a small splenium were also noted. Appearances suggested a possible underlying neurometabolic or neurogenetic disorder, with no specific features. Plasma proline was consistently raised above 350 μmol/L (ref: 85 – 290 μmol/L).

Patient 5, the younger sister of patient 4, presented to our clinic alongside her sister for further investigation. Although some features of her condition resemble those of her sister, she is much more mildly affected. At the time of examination at 3 years of age, her main presenting symptom was developmental delay. She has never had any seizures. She started walking at 3 years and six months later was able to take ten steps. At this age she had almost ten words, including "mum" and "dada". Although having progressed much further than her sister, she is delayed compared to her peers at school and is making very slow developmental progress. She has also never suffered periods of developmental regression. Physical examination demonstrated hypotonia with normal reflexes. Her ankle and other joints appeared lax. She walked with a broad-based gait and very flat feet. She had hypermetropia, astigmatism and a right convergent squint. She is microcephalic and a brain MRI at 2 years of age demonstrated delayed temporal pole myelination, prominent cerebellar folia and a small splenium. Plasma proline was normal and no other amino acid abnormalities were identified. Absence seizures were detected at seven years of age.

Both sisters have had extensive biochemical investigations (Table 5.3.1). Ammonia has been mildly raised in the plasma of patient 4 (56; ref: < 40 μmol/L). Abnormalities of amino acid concentrations were also noted (detailed in Results). Patient 4 had a muscle biopsy, which showed increased lipid, increased variation in fibre size due to scattered small fibres and mild myopathic features. Activities of the respiratory chain complexes were normal.

Table 5.3.1: Results of metabolic and genetic investigations carried out in patients 1, 4 and 5 at the time that whole exome sequencing was performed. nd; not measured or not stated in clinical notes; 5-HIAA, 5-hydroxyindoleacetic acid; 5-MTHF, 5-methyltetrahydrofolate; BH2, dihydrobiopterin; BH4, tetrahydrobiopterin; HVA, homovanillic acid; TSH, thyroid-stimulating hormone. Specific amino acid abnormalities are detailed in Figures 5.4.4 - 5.4.7.

Test	Patient 1	Patient 4	Patient 5
White cell lysosomal enzymes	Normal	Normal	Normal
Urine glycosaminoglycans (GAGs)	nd	Normal	Normal
Vacuolated lymphocytes	Normal	Normal	Normal
Full blood count	Normal	Normal	nd
DNA repair studies	nd	Normal	nd

Plasma PLP	nd	Normal	nd
Very long chain fatty acids (VLCFA)	Normal	Normal	Normal
Plasma glucose	Normal	Normal	nd
Plasma lactate	Normal	Normal	Normal
Parathyroid hormone	nd	Normal	nd
Biotinidase	Normal	Normal	Normal
Vitamin A	nd	Normal	nd
Vitamin E	nd	Normal	nd
Carnitines	Normal	Normal	nd
Creatine kinase (CK)	Normal	Normal	nd
White cell ubiquinone	Low: 24 (ref: 37-133 pmol/mg)	nd	nd
Purines	nd	Normal	nd
Transferrin isoelectric focussing	Normal	Normal	nd
Mitochondrial DNA mutations	nd	Normal	nd
Urine organic acids	Normal	nd	Normal
CSF neurotransmitters	Low BH4: 21 (ref: 27-105 nmol/L). Normal HVA, 5-HIAA, Neopterin, BH2 and 5-MTHF.	nd	All normal
CSF PLP	Low: 37 (ref: 44-89 nmol/L)	nd	Normal
CSF glucose/lactate	Normal	nd	Normal
CSF amino acids	Abnormal	nd	Abnormal
Plasma amino acids	Abnormal	Mild abnormalities	Normal
Urine amino acids	Mild abnormalities	nd	Mild abnormalities
Liver function tests	Normal	Normal	nd
Array CGH	nd	Normal	Normal
Fragile X analysis	nd	Normal	nd
Thyroid function test	Normal	Raised TSH: 8.9 (ref: 0-6 mU/L)	Normal
Calcium	Normal	nd	Normal
Phosphate	Slightly raised: 2.0 (ref: 1.2-1.8)	nd	Normal
Electrolytes	nd	nd	Normal
Fibroblast electron microscopy	nd	Cytoplasmic vacuoles (mostly empty but some with electron dense lamellar bodies)	Cytoplasmic vacuoles (mostly empty)
Muscle biopsy	Increased lipid and glycogen. Normal respiratory chain analysis and histology enzymes.	Increased variation in fibre size due to scattered small fibres. Increased lipid. Mild myopathic features. Normal respiratory chain enzymes.	nd

5.4 RESULTS AND DISCUSSION

5.4.1 *Whole exome sequencing*

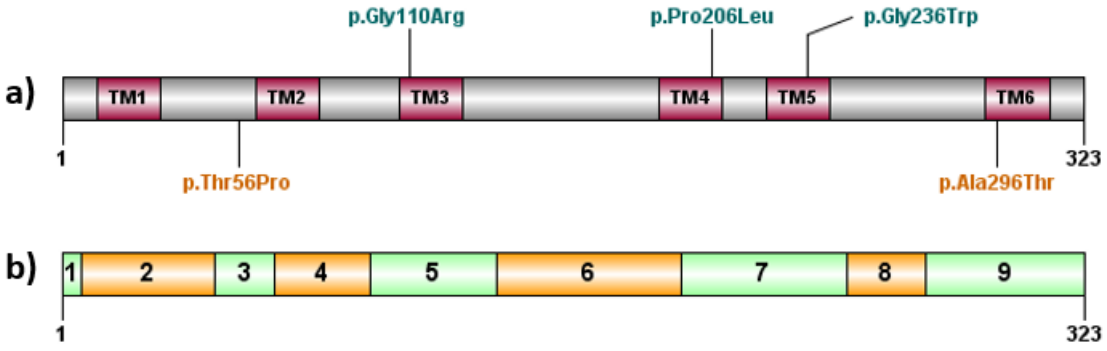
Whole exome sequencing was employed to identify the aetiology of disease in two unrelated families that had been extensively biochemically investigated and for whom it had not been possible to identify a diagnosis. Data was filtered and analysed assuming that each patients' disorder had been inherited in an autosomal recessive manner. Each family was sequenced and analysed separately. Common variants with a minor allele frequency of $> 3\%$ in the 1000 Genomes Project or NHLBI ESP databases were excluded, as were variants with a read depth of less than 5. Variants classified as "pathogenic" or "likely pathogenic" according to computed American College of Medical Genetics and Genomics (ACMG) Guidelines (detailed in Tables 3.3.7 and 3.3.8) were included. Only variants associated with a loss of gene function including missense, splice site, insertions, deletions or variants affecting start or stop codons were considered. In patient 1 this generated a list of 186 variants in 108 genes. This included, 126 missense variants, 34 deletions, 20 insertions and 6 variants affecting splice sites. In patients 4 and 5, 17 homozygous variants were identified, including 16 missense changes and one affecting a splice donor site. Variants were then further filtered to those associated with the biological term "epilepsy". In the case of patients 4 and 5, any potentially pathogenic variants had to be present in the same state in both siblings to be considered (Table 5.4.1). In patient 1, only one variant in *SLC25A22* was identified. A variant in *SLC25A22* was also identified in patients 4 and 5. However, three other variants were also identified in the following genes: (i) *LOX* which encodes lysyl oxidase with secondarily reduced levels of this enzyme being implicated in occipital horn syndrome and Menkes disease, (ii) *DRD4* which encodes for dopamine receptor D4 and has been identified as a risk factor for developing attention deficit-hyperactivity disorder, (iii) *SMPD1* which encodes acid sphingomyelinase and mutations in this gene cause Niemann-Pick disease type A/B. Although this latter disorder shares some clinical similarities with those seen in patients 4 and 5 (e.g. developmental delay and hyptonia), overall the phenotypic correlation was poor and after further detailed consideration (discussed in Section 5.4.6.1) this variant was not deemed to be relevant to disease pathogenesis.

Table 5.4.1: Whole exome sequencing data after having been filtered to show autosomal recessive variants associated with the biological term "epilepsy". -, not stated in Ingenuity Variant Analysis software.

Chr	Position	Reference Allele	Sample Allele	Gene	Protein Variant	Het/Hom	SIFT	PolyPhen	dbSNP ID	1000 Genomes Frequency	NHLBI ESP Frequency	ExAC Frequency
Patient 1												
11	794494	T	G	SLC25A22	p.Thr56Pro	Hom	Damaging	Probably damaging	-	-	-	-
Patients 4 and 5												
1	152573420	C	A	LCE3C	p.Asn71Lys	Hom	-	Benign	138652729	0.44	0.61	0.21
5	121413581	C	G	LOX	p.Glu34Gln	Hom	Tolerated	Benign	777431502	-	-	-
5	137688497	G	A	KDM3B	p.Ala5Thr	Hom	-	Possibly damaging	200772506	0.88	-	7.30
11	640157	C	A	DRD4	p.Pro303Gln	Hom	Tolerated	-	752533625	-	-	0.58
11	792001	C	T	SLC25A22	p.Ala296Thr	Hom	Damaging	Probably damaging	773825846	-	-	0.001
11	6412635	G	A	SMPD1	p.Val113Met	Hom	Damaging	-	142215226	0.06	0.06	0.08

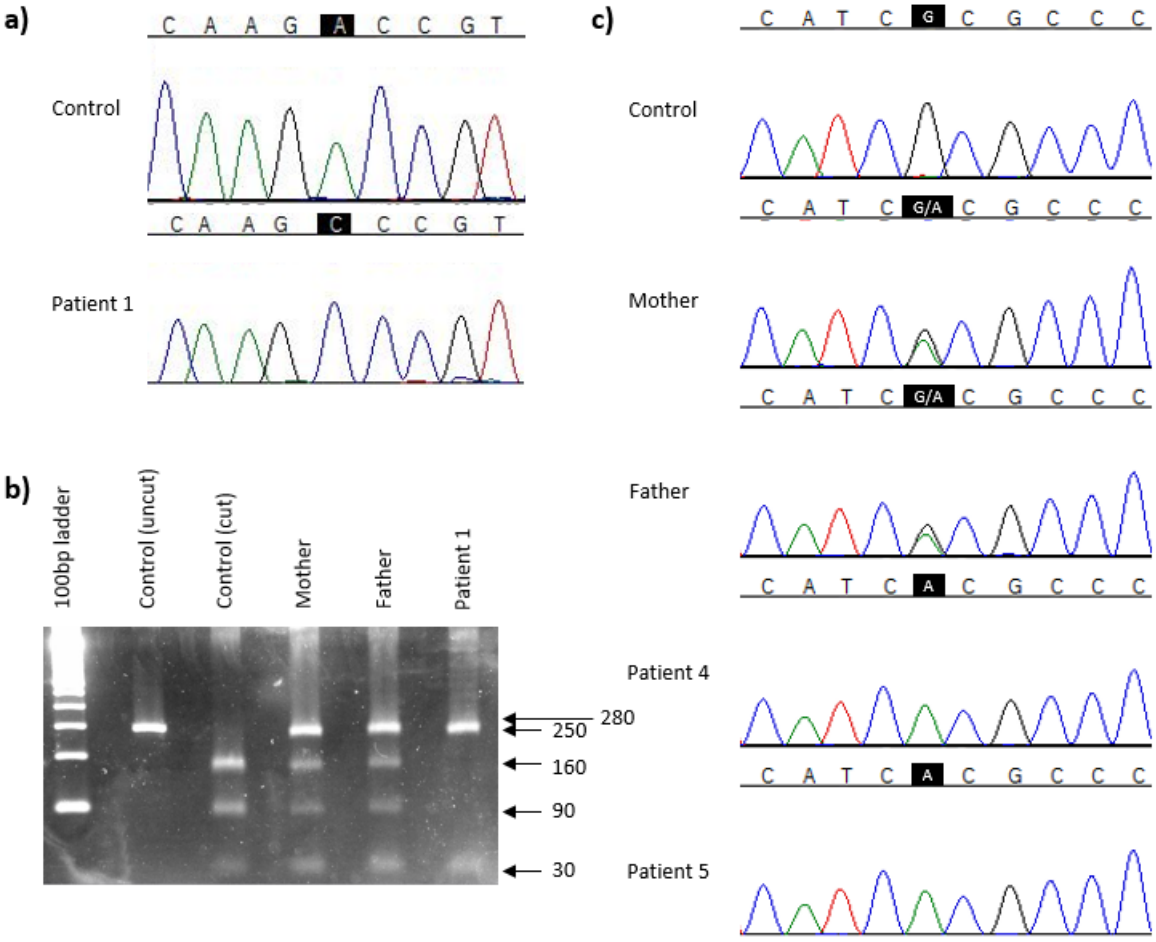
Homozygous sequence changes in *SLC25A22*, encoding solute carrier family 25 member 22, were identified in all three patients and deemed to be the best candidate for pathogenesis in both families due to the phenotypic similarities between our patients and those described previously in the literature (reviewed in detail in Table 5.4.3) and the predicted functional impact of each variant. Mutations in this gene are known to cause neonatal epileptic encephalopathy and migrating partial seizures of infancy. Only three pathogenic mutations in *SLC25A22* have been described previously in nine patients (Molinari *et al.*, 2005, 2009; Poduri *et al.*, 2013; Cohen *et al.*, 2014), all of which are single amino acid substitutions that alter key residues involved in the transport mechanism of the protein. Two novel mutations in *SLC25A22* were identified in patient 1 (c.166A>C; p.Thr56Pro) and patients 4 and 5 (c.886G>A; p.Ala296Thr) (Figure 5.4.1).

Figure 5.4.1: Schematic representation of SLC25A22. (a) Six transmembrane domains of SLC25A22. Mutations reported in the literature causing a seizure disorder are shown in blue and the mutated residues in our families are shown in orange. **(b)** Regions corresponding to each of the nine exons of *SLC25A22* (aligned to transmembrane domains above).



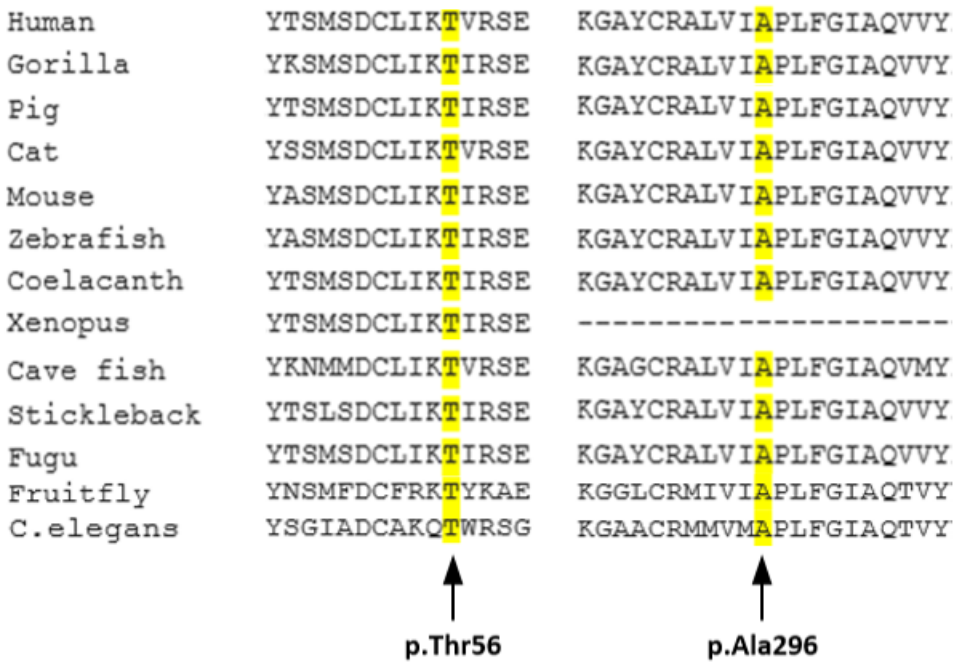
Both variants were predicted damaging by SIFT and probably damaging by PolyPhen-2. Neither mutation has been reported in the 1000 Genomes Project or NHLBI Exome Sequencing Project databases. p.Thr56Pro has also not been reported in the ExAC database but p.Ala296Thr has been identified in a heterozygous state in one European individual, resulting in a minor allele frequency of 0.0012%. Segregation of the p.Thr56Pro mutation was confirmed in the parents of patient 1 by restriction enzyme digestion with HpyCH4III (Figure 5.4.2b). This restriction enzyme recognises ACNGT sequences, cutting between the N and G nucleotides. Digestion of PCR products containing the wild-type sequence would be expected to result in three DNA fragments of 30, 90 and 160 bp. In contrast, the presence of the p.Thr56Pro mutation would be expected to abolish the second HypCH4III restriction site, therefore resulting in two fragments of 30 and 250 bp. Segregation of p.Ala296Thr in the parents of patients 4 and 5 was confirmed by Sanger sequencing (Figure 5.4.2c).

Figure 5.4.2: *SLC25A22* mutation analysis in affected families. **(a)** Sanger sequencing illustrating the p.Thr56Pro mutation in patient 1. **(b)** HpyCH4III restriction enzyme digestion of patient 1 and parental DNA. Both parents have a heterozygous pattern and patient 1 is homozygous for the mutation. **(c)** Sanger sequencing illustrating the homozygous p.Ala296Thr mutation in patients 4 and 5. Both parents are heterozygous for the mutation.



Both mutations also affect highly conserved regions of the protein with each residue being conserved from humans to *Drosophila* and *C. elegans*. There is no known ortholog of SLC25A22 in bacterial species. No potentially pathogenic changes in genes known to cause hyperprolinaemia, namely *PRODH* and *ALDH4A1*, were identified in the the whole exome sequencing data for any of the patients.

Figure 5.4.3: Conservation of mutation sites in SLC25A22. Multiple sequence alignment of the *SLC25A22* gene across species generated using Clustal Omega illustrating the conservation of the sites of the p.Thr56Pro and p.Ala296Thr mutations, respectively.



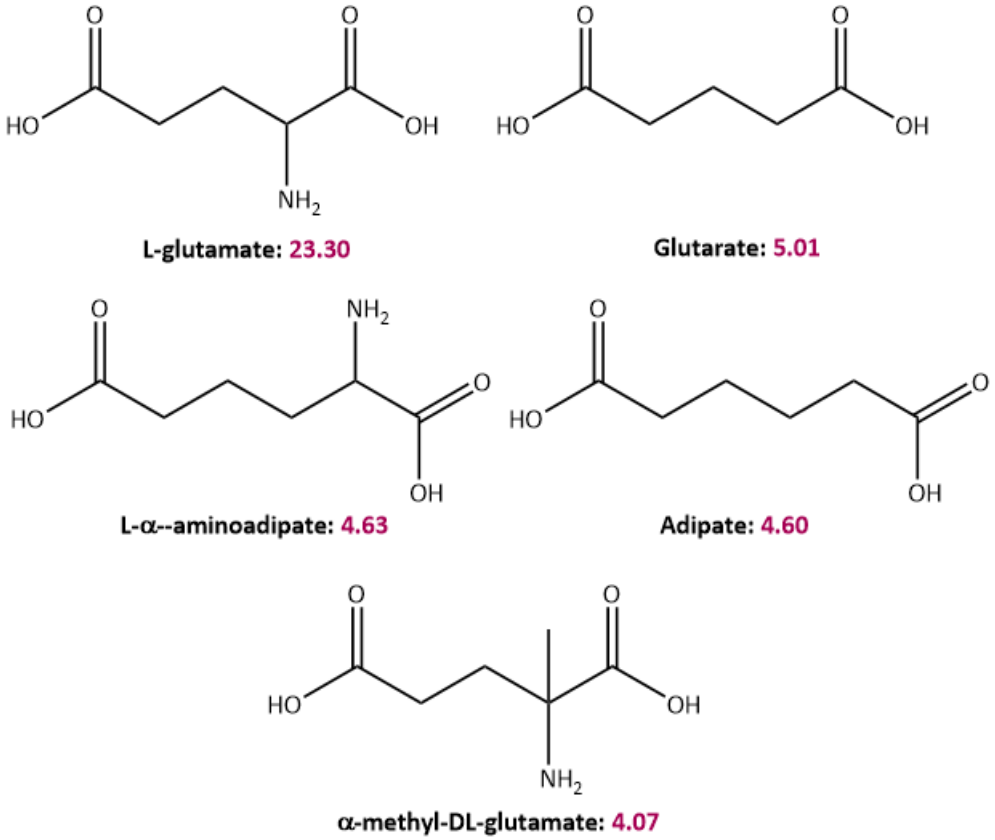
5.4.2 *Function of SLC25A22*

In humans, the *SLC25* gene family encodes 53 mitochondrial carriers that function to import and/or export a variety of metabolites from the mitochondria (Palmieri, 2014). In 2002, Fiermonte *et al.* identified two proteins that could catalyse the transport of glutamate across the mitochondrial membrane, SLC25A22 and SLC25A18. These transporters share 63% sequence identity to each other and 33% to the C-terminal domain of the aspartate/glutamate carrier. SLC25A22 is more highly expressed than SLC25A18 in all human tissues with the exception of the brain, where their expression is equal. The substrate specificities and transport abilities of each protein were assayed by reconstituting each in proteoliposomes. Both were able to catalyse glutamate uniport (i.e. transport in one direction) and exchange (i.e. concurrent influx and efflux). When comparing the kinetic characteristics of the two transporters, SLC25A22 had a higher K_m and V_{max} and thus a lower affinity for glutamate. It has therefore been hypothesised that SLC25A22 only becomes active to cope with higher cytosolic concentrations of glutamate and regulate the transport rate of this amino acid into the mitochondria to satisfy tissue-specific needs, for example after a protein-rich meal.

Further studies using liposomes confirmed that SLC25A22 was not only able to facilitate the uptake of [U-^{14}]-glutamate but, albeit at a reduced rate, also other molecules with a structural

similarity to glutmate (Figure 5.4.4). These molecules are structurally similar to glutamate with the following modifications: the removal of the amine group (-NH_2), addition of a methylene group (-CH_2-) to the carbon backbone, the addition of a methylene group and removal of the amine group, or the addition of a methyl group at the α-position (-CH_3).

Figure 5.4.4: Substrates transported by SLC25A22 and the exchange rates with respect to each substrate. Exchange rates were quantified as μmol/min/g protein (purple text).



Very few members of the SLC25 family have had their 3-dimensional structure solved due to difficulties in crystallisation. However, sequence alignment analysis and site-directed mutagenesis studies have provided further information regarding structural-functional links and transport mechanisms. All members of the SLC25 protein family share a common membrane topology which consists of three homologous domains. Each domain contains two hydrophobic transmembrane α-helices separated by a highly conserved hydrophillic region protruding into the mitochondrial matrix, each containing a short helix (Figure 5.4.5a). The six transmembrane α-helices form a funnel-shaped cavity in the inner mitochondrial membrane through which the transported substrates can pass. Each of the three hydrophillic regions contains the following sequence motif: PX[D/E]XX[K/R]X[K/R] (20-30 residues) [D/E]GXXXX[W/Y/F][K/R]G (Palmieri, 2014). The first segment of this motif is located at the C-terminus of the odd-numbered helices and the third segment is located at the N-terminus of the even-numbered helices. The proline residue beginning

each motif acts to sharply kink each odd-numbered helix, bringing the sequence motifs into close proximity (Figure 5.4.5). The charged residues are then able to form a network of salt bridges that closes the transporter cavity on the side of the mitochondrial matrix. It has also been proposed that all SLC25 transporters share a similarly located substrate binding site formed by residues within the three even-numbered helices. These residues are centrally located in each helix (one and a half helix turns above the salt bridge network) and protrude into transporter cavity (Palmieri and Pierri, 2010). These residues also determine substrate specificity, for example the presence of R[D/E] amino acids within the fourth helix is specific for amino acid carriers. A second proposed salt bridge network formed by the motif [F/Y][D/E]XX[R/K] is hypothesised to close the transporter cavity on the cytosolic side, facilitating bidirectional transport (Robinson *et al.*, 2008).

In order to predict the possible effect of the mutations identified in our patients, their position within the protein was examined. Residue Thr56 is located within the first matrix α-helix. In patient 1, this residue is mutated to proline, which is likely to affect this secondary structure as proline is a known helix-breaker. In turn, the Thr56Pro mutation may cause structural instability of the first domain which consists of the first and second transmembrane helices and the first matrix α-helix. Ala296 is located within the sixth transmembrane helix and faces the hydrophobic core of the inner mitochondrial membrane. In the siblings patients 4 and 5, this residue is mutated to a threonine residue. The introduction of a polar residue into a hydrophobic area might lead to a destabilisation of the helical structure. Sequence alignment of all 53 human mitochondrial carriers revealed that six other SLC25 proteins (SLC25A27, SLC25A31, SLC25A32, SLC25A38, SLC35A44 and SLC25A46) also have a threonine at the equivalent position (Figure 5.4.6a). However, an alignment of these six proteins revealed that this amino acid is not part of a conserved region within these transporters. In contrast, when comparing the two glutamate transporters (SLC25A22 and SLC25A18), Ala296 and the surrounding the residues are highly conserved indicating that this region is likely to play an important role in glutamate transport.

Figure 5.4.5: Schematic representation of SLC25 transporter family. **(a)** Schematic representation illustrating the membrane topology shared by members of the SLC25 transporter family. Each contains six transmembrane helices (H1 - H6) and the structure can be divided into three repeated domains (Repeat 1 - 3). Each domain contains two transmembrane helices connected by a matrix loop which contains a short helix (h12, h34 and h56, respectively). The three domains are joined by short loops that protrude into the intermembrane space. Adapted from Palmieri (2014). **(b)** Membrane topology of SLC25A22. Mutations reported in the literature causing a seizure disorder are shown in blue and the mutated residues in our families are shown in orange.

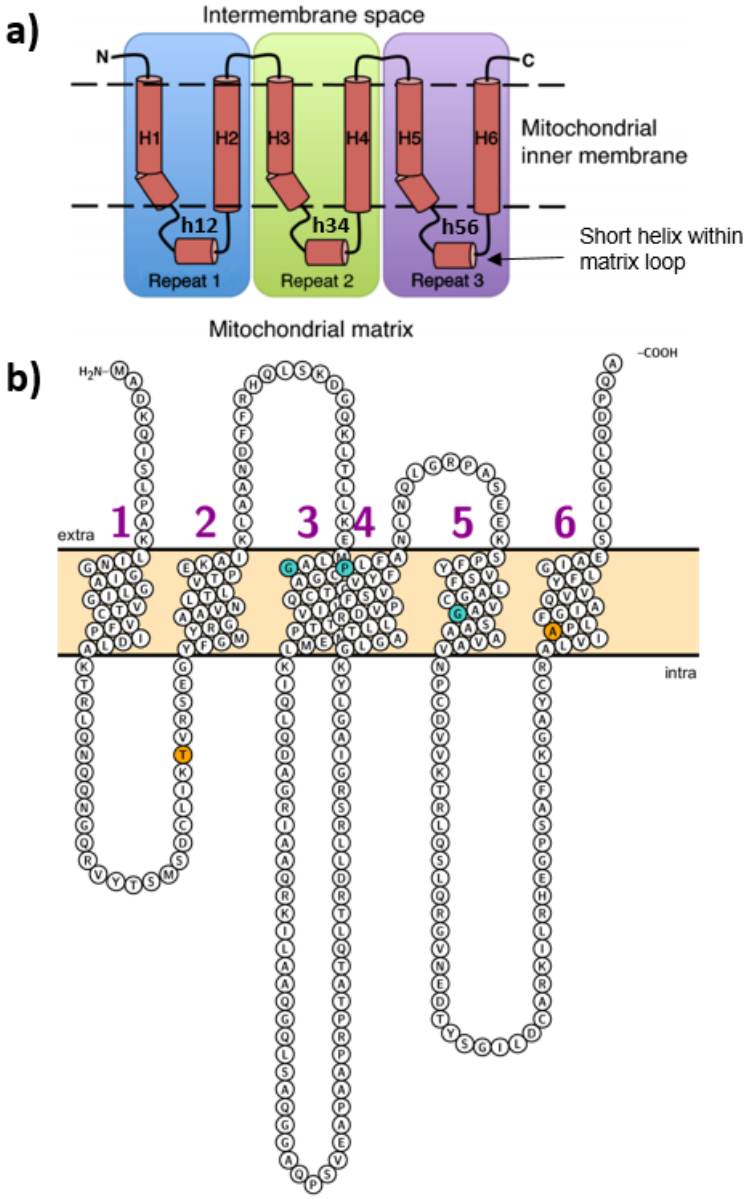


Figure 5.4.6: Conservation of p.Ala296 position across SLC25 family. **(a)** Multiple sequence alignment of the 53 human mitochondrial carriers (SLC25 family) showing the amino acid residues corresponding to the p.Ala296 position in SLC25A22 (yellow). Proteins with a threonine at this position are highlighted in blue and SLC25A22 is in green. **(b)** Multiple sequence alignment of the six SLC25 proteins with a threonine at this position (yellow) compared to SLC25A22 (green).

5.4.3 *Phenotypic correlation with other patients with SLC25A22 deficiency*

Only three pathogenic mutations in *SLC25A22* have been described previously in nine patients (Molinari *et al.*, 2005, 2009; Poduri *et al.*, 2013; Cohen *et al.*, 2014) (summary in Table 5.4.2). Prior to the description of the patients in this chapter, all previously described cases had early-onset intractable seizures with a burst-suppression pattern on EEG, hypotonia, occular abnormalities (especially abnormal VEP), generalised brain atrophy and abnormalities of the corpus callosum evolving into a severe encephalopathy with spasticity and a poor developmental outcome. All patients underwent extensive biochemical testing to investigate the possibility of a metabolic disorder causing disease pathogenesis. However, no abnormalities were reported in any of the nine patients.

Table 5.4.2: Phenotypic comparison of our five patients compared to those described in the literature with mutations in *SLC25A22*. ERG, electroretinography; VEP, visual evoked potential; nd, not done; *, patient may have been too young to detect abnormalities.

Patient	Consanguineous (country of origin)	Age at presentation	Age at seizure onset	Main presenting features	Electroencephalogram (EEG)	Eye abnormalities	Brain imaging	Mutation
Patient 1	Yes (Afghanistan)	Birth	6 weeks	Refractory seizures + hypotonia	Abnormal	No response to light, abnormal VEP	Small splenium, prominent cerebellar folia, frontotemporal hypoplasia, delayed myelination.	c.166A>C; p.Thr56Pro
Patient 2	Yes (Afghanistan)	Birth	10 weeks	Refractory seizures + hypotonia	Multifocal	Occasional visual fixation but not following	nd	c.166A>C; p.Thr56Pro
Patient 3	Yes (Afghanistan)	Birth	10 weeks	Refractory seizures + hyptonia	Multifocal	Occasional visual fixation but no following	nd	c.166A>C; p.Thr56Pro

Patient 4	Yes (Syria)	First month	First month	Refractory seizures + hypotonia	6 years: Symmetrical diffuse irregular slow wave activity	Normal	Small splenium, prominent cerebellar folia, delayed temporal pole myelination. Symmetrical signal abnormalities of the insular cortex bilaterally and adjacent capsular white matter.	c.886G>A; p.Ala296Thr
Patient 5	Yes (Syria)	3 years	7 years	Developmental delay + hypotonia	nd	Hypermetropia, an astigmatism and right convergent squint	Small splenium, prominent cerebellar folia, delayed temporal pole myelination.	c.886G>A; p.Ala296Thr
Molinari *et al.* (2005)	Yes (Israel)	First 48 hours (n=4)	First 48 hours (n=4)	Refractory seizures + hypotonia	Burst-suppression (n=2)	Abnormal ERG (n=2), VEP (n=2)	Brain atrophy (n=2), subarachnoid enlargement (n=1), cerebellar hypoplasia (n=1), hypoplastic corpus callosum with abnormal splenium (n=1).	c.617C>T; p.Pro206Leu
Molinari *et al.* (2009)	Yes (Algeria)	Birth	5 days	Refractory seizures + hypotonia	8 days: Burst suppression. 6 months: Hypsarrhythmia	Abnormal ERG, VEP	Cerebellar hypoplasia, abnormal corpus callosum abnormal gyration of both temporo-parieital poles with lack of differentiation of white and grey matter.	c.706G>T; p.Gly236Trp
Poduri *et al.* (2013)	Yes (Saudi Arabia)	1 week (n=1), 2 weeks (n=1)	1 week (n=1), 2 weeks (n=1)	Refractory seizures + hypotonia (n=2)	1 month: Delta brush pattern (n=1), positive spikes and high-voltage focal spikes (n=2)	Normal (n=1), abnormal VEP (n=1)	Normal (n=1), delayed myelination and thin corpus callosum (n=1).	c.328G>C; p.Gly110Arg
Cohen *et al.* (2014)	Yes (Israel)	First months (n=1), 2 weeks (n=1)	First months (n=1), 2 weeks (n=1)	Refractory seizures + hypotonia (n=2)	Burst-suppression (n=2)	No response to light or movement (n=2), retinal pigmentation (n=1)	Thin corpus callosum and brain atrophy (n=2), increased Virchow-Robin spaces and increased subarachnoid spaces (n=1).	c.617C>T; p.Pro206Leu

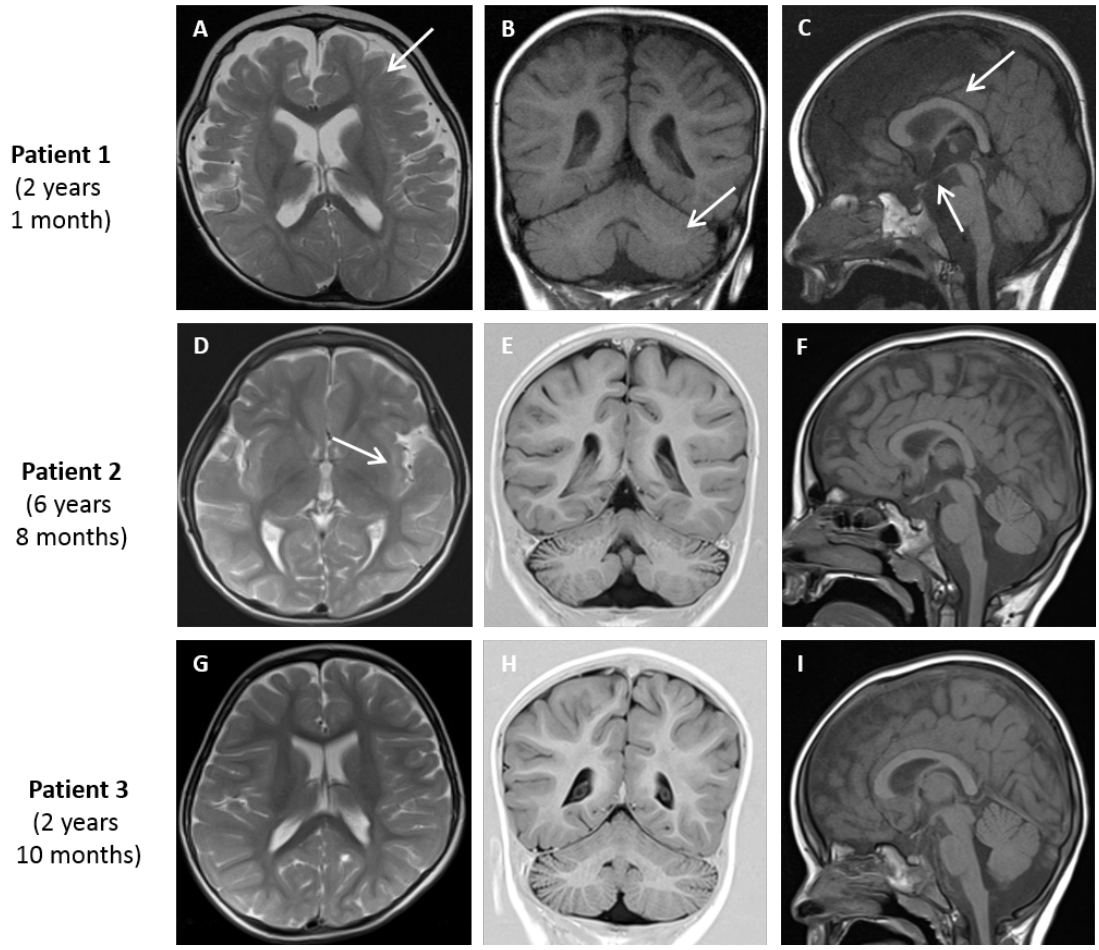
The severe seizures affecting these patients have been proposed by Molinari *et al.* (2009) to be due to an accumulation of glutamate, a powerful excitatory neurotransmitter, in brain cells. This accumulation is predicted to occur to a greater extent within astrocytes than neurons as the expression of the mitochondrial aspartate/glutamate carriers (*SLC25A12* and *SLC25A13*) is restricted to neurons (Berkich *et al.*, 2007). Thus, the lack of glutamate uptake in astrocytes would not be compensated for by the aspartate/glutamate carriers, leading to cytosolic glutamate accumulation followed by inappropriate release into the synaptic cleft. As described previously in the context of oxidative stress (Section 4.5.11 and Figure 4.5.7), glutamate release causes activation of NMDA and AMPA receptors and excessive neuronal calcium influx. The result is a vicious cycle of damage including the generation of reactive oxygen and nitrogen species, mitochondrial dysfunction, cell death and neuronal hyperexcitability resulting in the propagation of seizures.

Despite harbouring novel mutations, patients 1 and 4 share similar clinical features to those described previously (Table 5.4.2). These include refractory seizures and psychomotor developmental delay. Patient 1 also displays an impaired cortical response to light and abnormal VEP seen in 63% (5/8) patients. Patient 5 however, presented atypically with features not present in her sister including hypermetropia, an astigmatism and a right convergent squint. Although visual pathway dysfunction is common in patients with *SLC25A22* mutations, these mild eye abnormalities have not been reported previously; indeed, they may be unrelated to her genetic diagnosis. Most notably though, she is much more mildly affected than her sister and did not present with seizures until much later on in life. This also contrasts to the other patients described thus far, the vast majority of which presented with seizures within the first two weeks of life. Indeed, all of the new patients described in this chapter begin seizing later than those reported in the literature (one month - seven years). Our second family was also atypical with regard to their developmental outcome. All patients reported in the literature developed a severe encephalopathy with spasticity rendering them wheelchair-bound and unable to achieve developmental milestones. In contrast, patient 4 could sit and crawl independently at 2 years and 6 months of age and babble at 7 years of age. Consistent with her other clinical features, patient 5 was not affected as severely. Although making very slow progress, at 3 years of age she was able to take ten steps independently and had a limited vocabulary of approximately ten words. As described above (Section 5.4.2) this milder neurological impairment may be due to greater residual activity of the p.Ala296Thr mutation relative to the other variants. This hypothesis will however require corroboration with functional studies.

Previously reported MRI changes in patients with *SLC25A22* mutations include cerebellar hypoplasia, dysmorphism of the corpus callosum with hypoplasia and abnormal splenium, sub-

arachnoid enlargement and generalised brain atrophy (Table 5.4.2). The findings in our patients were very similar. Scans performed after two years of age all showed a small splenium, delayed temporal pole myelination and prominent cerebellar folia indicating cerebellar hypoplasia. In addition, we noted that in all of our patients who had imaging investigations, the optic chiasma was small, consistent with post-retinal dysfunction (Figure 5.4.7).

Figure 5.4.7: Brain magnetic resonance imaging of SLC25A22-deficient patients. (A, D and G) T2-weighted images. **(B, C, F and H)** T1-weighted images. **(E and H)** T1-weighted images with inversion recovery sequence. Patient 1 (2 y 1m), Patient 4 (6y 8m) and Patient 5 (2y 10m). All patients show frontotemporal hypoplasia/atrophy **(A, D and G)** and prominence of cerebellar folia, consistent with cerebellar hypoplasia/atrophy **(B, E and H)**. All patients also have a small splenium, with the splenium of the corpus callosum smaller than the genu and a small optic chiasma (smaller than the mammillary body) **(C, F and I)**. Abnormal appearance of the insular cortex is also noted in Patient 4 **(D)**.

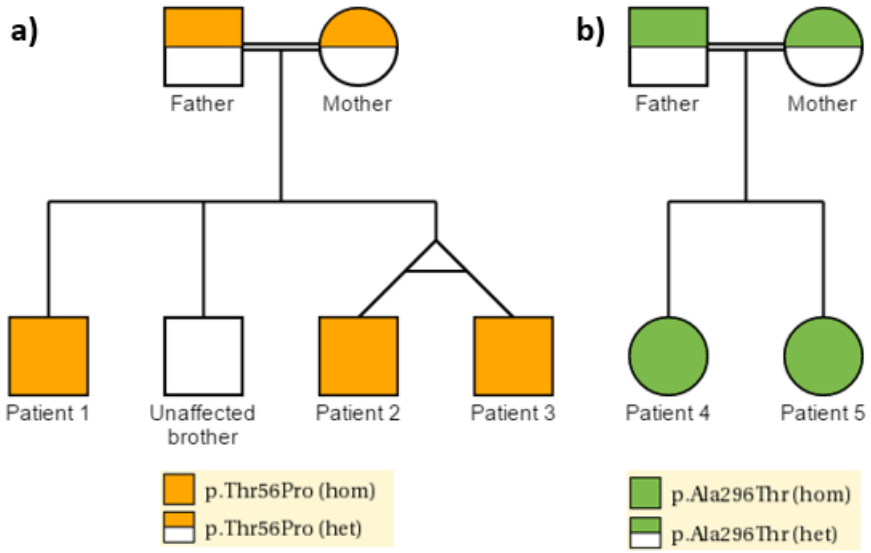


5.4.3.1 *Patients 2 and 3*

After patient 1 was diagnosed, his mother had a subsequent pregnancy. All of the evidence, including the whole exome sequencing data and strong phenotypic correlation between patient 1 and the nine patients with mutations in *SLC25A22* previously described, was critically con-

sidered before prenatal diagnosis was offered. This Sanger sequencing (performed by the North East Thames Regional Genetics Service, GOSH, UK) indicated that the foetus had the same homozygous p.Thr56Pro mutation in *SLC25A22* as patient 1 and the parents elected to terminate the pregnancy. His mother then had a second pregnancy with twins and prenatal diagnosis indicated that both foetuses (patients 2 and 3) had homozygous *SLC25A22* mutations. However, on this occasion the parents decided not to terminate the pregnancy. Pedigrees illustrating the relationship between the affected patients in each family are shown in Figure 5.4.8.

Figure 5.4.8: Pedigrees of both families found to have mutations in *SLC25A22*. The mutation status of the unaffected brother in family 1 is not known.



Since it could be hypothesised that abnormal mitochondrial glutamate levels related to the SLC25A22 mutations could lead to impaired mitochondrial redox balance with increased ROS production and on the basis that their affected brother had low levels of ubiquinone in his white cells, the twins were commenced on supplemental ubiquinone (10 mg/kg/day) as a mitochondrially targeted antioxidant from birth. Initially they appeared less severely affected than their brother. Whereas at 6 weeks patient 1 was profoundly hypotonic, visually inattentive and had both seizures and dystonic movements, at 6 weeks the twins had milder, and mainly axial, hypotonia. They were both smiling and fixing and had no abnormal movements. Both appeared slightly dysmorphic, with a smooth philtrum and low set ears. Sadly, at the age of 16 weeks, both twins developed seizures with a flexor spasm-like semiology and multifocal EEG. In addition to ubiquinone, they are currently both on levetiracetam (40 mg/kg/day). At last review aged almost 12 months old, both twins continue to have generalized tonic-clonic seizures once or twice a week and have developmental stasis. Patient 2 occasionally fixes and smiles and patient 3 occasionally fixes but does not smile. They are hypotonic with no head control and have normal tendon reflexes. Most recent amino acid quantitation at 11 months of age

identified an elevated proline concentration in patient 2 of 418 μmol/L (ref: 85 – 290 μmol/L), whereas in patient 3 levels remained within the control range (274 μmol/L). Other mild amino acid abnormalities were also detected in both twins. Given that patients 2 and 3 were prenatally diagnosed, extensive biochemical investigations similar to those carried out for patients 1, 4 and 5 were not performed. The amalgamated prevalence of each symptomatology in all patients described to date are summarised in 5.4.3.

Table 5.4.3: Clinical phenotype and demographics of patients with *SLC25A22* mutations.

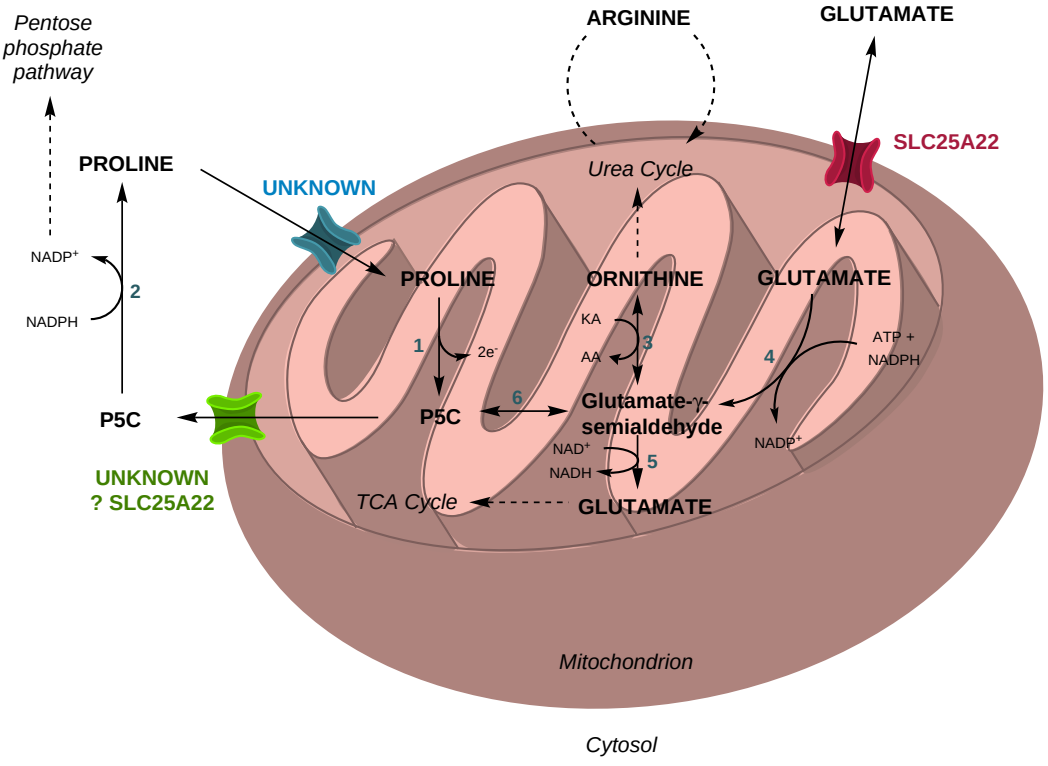
Clinical features and demographics	Incidence
Gender	8/14 male; 6/14 female
Country of origin	Afghanistan 3/14; Syria 2/14; Israel 6/14; Algeria 1/14, Saudi Arabia 2/14
Parental consanguinity	14/14 (100%)
Age at seizure onset	First week; 6/14 (43%) 1 week - 1 month; 3/14 (21%) 1 - 3 months; 4/14 (29%) 7 years; 1/14 (7%)
Seizure type	Neonatal/early infantile epileptic encephalopathy; 11/14 (79%) Migrating partial seizures of infancy; 2/14 (14%) Absence seizures in childhood 1/14 (7%)
Response to AEDs	None; 5/11 (45%) Partial; 5/11 (45%) Yes; 1/11 (9%)
Distinct EEG features	Burst-suppression; 5/11 (45%) Multifocal; 4/11 (36%) Delta brush pattern; 1/11 (9%) Irregular slow wave activity; 1/11 (9%)
Visual evoked potential (VEP)	Abnormal; 5/8 (63%) Normal; 3/8 (37%)
Electroretinogram (ERG)	Abnormal; 2/4 (50%) Normal 2/4 (50%)
Other ocular findings described	No fixing/following; 2/11 (18%) No response to light; 5/11 (45%) Retinal pigmentation; 1/11 (9%) Hypermetropia, astigmatism and convergent squint; 1/11 (9%) Normal; 2/11 (18%)
Microcephaly	7/7 (100%)
MRI	Hypoplastic corpus callosum/splenium; 8/11 (72%) Cerebellar hypoplasia/prominent folia; 5/11 (45%) Delayed myelination; 4/11 (36%) Brain atrophy; 4/11 (36%) Subarachnoid enlargement; 2/11 (18%) White matter abnormalities; 2/11 (18%) Frontotemporal hypoplasia; 1/11 (9%) Normal; 1/11 (9%)

Biochemical findings	Persistently elevated plasma proline; 2/5 (40%) Intermittently elevated plasma proline; 1/5 (20%) Intermittently elevated plasma ornithine or arginine; 4/5 (80%) Vacuolated fibroblasts (mostly empty); 3/3 (100%)

5.4.4 *Catabolism, cycling and synthesis of proline*

Prior to the identification of patients 1 - 5, biochemical abnormalities have not been associated previously with this disorder. This is not the case in the patients we have described. One patient was referred because of hyperprolinaemia (patient 1) and in another two it was detected on first line investigations (patients 2 and 4). However, it was uncertain whether these elevated proline levels were an incidental unrelated finding or could be explained by mutations in *SLC25A22*. The metabolic pathways involved in catabolism, cycling and synthesis of proline are shown in Figure 5.4.9.

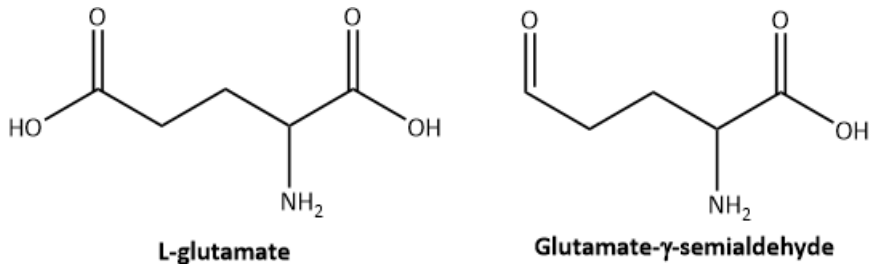
Figure 5.4.9: Proline biosynthetic and catabolic pathways. Schematic illustrating the synthesis, degradation and interconversions of proline, pyrroline-5-carboxylate (P5C), glutamate-γ-semialdehyde (GSA) and glutamate within the mitochondrion. 1, proline dehydrogenase; 2, P5C reductase; 3, ornithine-δ-aminotransferase; 4, P5C synthase; 5, P5C dehydrogenase; 6, non-enzymatic reaction.



Proline enters the mitochondrion from the cytosol and intramitochondrial FAD-dependent proline dehydrogenase converts proline to L-Δ^1-pyrroline-5-carboxylate (P5C). P5C can then either be transported out of the mitochondrion or be further metabolised. The transporter

responsible for the efflux of P5C has not been identified (Miller *et al.*, 2009). However, P5C in the mitochondrion can undergo non-enzymatic hydrolysis which opens the ring structure to generate glutamate-γ-semialdehyde (GSA), its equilibrium partner. Given the structural similarity of GSA and glutamate (Figure 5.4.10), it is possible that GSA, as well as glutamate, may be transported out of the mitochondrion by SLC25A22. Glutamate-γ-semialdehyde is structurally identical to glutamate with the exception of the γ-carboxylic acid (-COOH) which is exchanged with an aldehyde group (-COH). This substrate was not included in those tested by Fiermonte *et al.* (2002). Subsequently in the cytosol GSA could then spontaneously re-cyclise to form P5C.

Figure 5.4.10: Structural similarity of L-glutamate and glutamate-γ-semialdehyde.



Once in the cytosol P5C can be converted back to proline by P5C reductase, thus completing the proline/P5C shuttle and the transfer of reducing equivalents from cytosolic NADPH to the mitochondrial respiratory chain. Alternatively GSA, which is in equilibrium with P5C, can be further metabolised to either glutamate or ornithine and arginine. In the pathway to glutamate, P5C dehydrogenase uses nicotinamide adenine dinucleotide (NAD^+) to remove two electrons from GSA during the conversion (Arentson *et al.*, 2012). The pathway from P5C/GSA to ornithine proceeds via ornithine δ-aminotransferase (OAT), with the subsequent production of arginine occurring within the urea cycle.

Proline can also be synthesised from glutamate. This involves three enzymatic steps; the first two of which are catalysed by P5C synthase, a bifunctional enzyme that exhibits glutamate kinase and γ-glutamyl phosphate reductase activities (Pérez-Arellano *et al.*, 2010). Glutamate kinase uses ATP to generate γ-glutamyl phosphate, which is subsequently reduced by γ-glutamyl phosphate reductase using NADPH to produce GSA. GSA then non-enzymatically cyclizes to form P5C and subsequently is reduced to proline in the cytosol by P5C reductase. All the intramitochondrial metabolic interconversions can have an effect on plasma amino acid concentrations with defects in these enzymes causing characteristic abnormalities. Indeed, plasma proline is elevated in proline dehydrogenase and P5C dehydrogenase deficiencies, whilst proline, ornithine, arginine (and citrulline) are all low in P5C synthase deficiency (Baumgartner *et al.*, 2005). In contrast, OAT deficiency causes variable age-dependent abnormalities with plasma ornithine, arginine and citrulline being deficient in infancy (often accompanied by hyperammonaemia due to urea

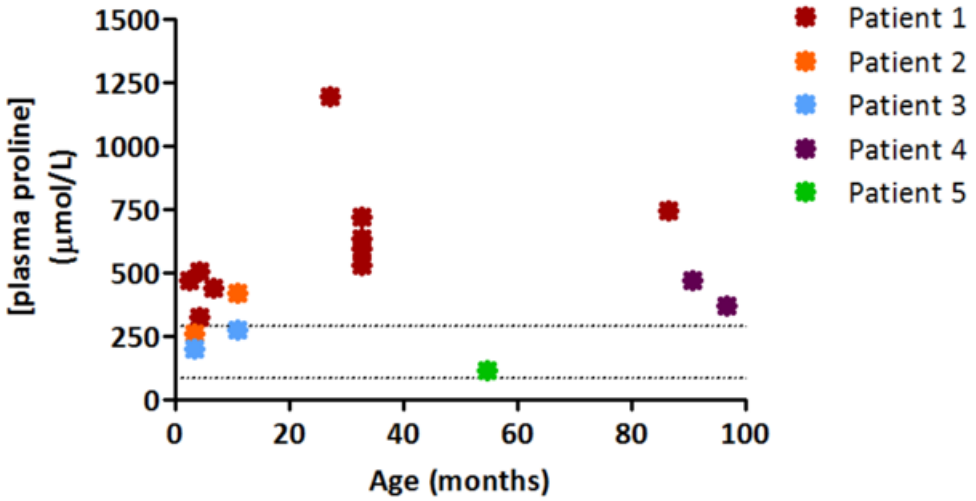
cycle impairment) but elevated after a year of age. This can be explained by the fact that OAT not only catalyses the transamination of P5C to form ornithine, but also the reverse reaction from ornithine to P5C which predominates after the first year of life. Therefore, mutations in *SLC25A22* may cause elevated proline concentrations in biofluids by reducing flux through the proline/P5C shuttle if GSA cannot exit the mitochondria.

5.4.5 *Amino acid abnormalities in patients with SLC25A22 deficiency*

In addition to the plasma proline levels which were consistently elevated in two patients and elevated in the most recent sample from one of the pre-natally diagnosed twins (Figure 5.4.11), other plasma amino acid abnormalities were also detected in multiple samples from the other patients. Amongst the most consistently elevated were glutamate, ornithine and arginine (Table 5.4.4 and Table 5.4.5). A defect in the export of GSA/P5C or glutamate from the mitochondrion could be expected to result in reduced catabolism of proline and increased production of glutamate and/or ornithine and arginine. The fact that 60% (3/5) of patients had elevated plasma proline concentrations and 80% (4/5) had elevated glutamate and/or ornithine and arginine on multiple occasions supports the hypothesis that the export of GSA/P5C and/or glutamate from the mitochondrion may be disturbed in SLC25A22 deficiency. This hypothesis requires that SLC25A22 catalyses the efflux as well as influx of GSA/P5C and/or glutamate into the mitochondria. As described in Section 5.4.2, this bidirectional transport of glutamate is supported by studies in reconstituted liposomes (Fiermonte *et al.*, 2002). An obvious concern regarding these experiments is the possibility that transporters may insert into liposomal membranes in both orientations leading to false reporting of bidirectional transport activity. However, kinetic data has shown that using this methodology, the transporters unidirectionally inserted into the liposomal membrane, suggesting that the observed bidirectional transport activity of SLC25A22 *in vivo* in not artefactual (Palmieri *et al.*, 1995).

The reasons for the lack of hyperprolinaemia in patients 3 and 5 are unclear, although the age of sampling for amino acid analysis may be important for detecting this mild abnormality. As described in the context of OAT deficiency, during the neonatal and young infant periods the flux of amino acid conversion within the mitochondria is proline → ornithine, but in older children the direction of flux is reversed (Figure 5.4.9). In order to investigate this further the plasma concentrations of proline, ornithine, glutamate and glutamine from patient 1 were plotted to examine the changes in these analytes over time, as this was the only patient for which amino acid concentrations had been quantified on more than two occassions over an extended period of time. As expected, there was a trend for proline concentrations to increase with age, whilst

Figure 5.4.11: Plasma proline concentrations in patients 1 - 5 with mutations in *SLC25A22*. Multiple data points at 32.8 months in patient 1 represent a pre- and post-prandial series of samples. Dotted lines represent reference ranges. All analysis was performed in the Chemical Pathology Department, GOSH, UK.



glutamate concentrations decreased (Figure 5.4.12). This may explain the normal plasma proline in patient 5 (4 years 6 months) as she was younger than her sister (7 years 6 months) at the time of sampling. However, patient 1 was much younger than this when proline was found to be elevated, suggesting that this effect may be mutation-specific. Indeed, there are likely to be additional environmental factors influencing plasma proline concentrations in these patients as, despite sharing the same genotype and treatment regimen, only one of the affected twins (patients 2) had an elevated proline concentration. These factors may include seizure control, antiepileptic drug treatment, diet and polymorphisms in mitochondrial genes or those that are involved in the synthesis, catabolism or transport of amino acids.

Moreover, if SLC25A22 can indeed catalyse the efflux of GSA/P5C, the accumulation of P5C inside the mitochondria of patients with SLC25A22-deficiency may contribute to seizure propagation. Neurotransmitter analysis undertaken at the time of CSF amino acid quantitation revealed a low concentration of pyridoxal 5'-phosphate (PLP) in patient 1. As described in Section 4.2.1, hyperprolinaemia type II is caused by a deficiency of P5C dehydrogenase (Figure 5.4.9) resulting in an accumulation of P5C. In the brain, excess P5C inactivates PLP by Knoevenagel condensation. If, as we have proposed, levels of P5C are high in the mitochondria then PLP could be inactivated in this site. Thus, as in hyperprolinaemia type II, this secondary PLP deficiency may be contributing to the propagation of seizures (Farrant *et al.*, 2001). Interestingly, patient 5 had not presented with seizures at the time of sampling and had a normal CSF PLP concentration, which again may be explained by the milder predicted impact of the p.Ala296Thr mutation. CSF analysis was not performed in patients 2 - 4.

Figure 5.4.12: Plasma amino acid analysis in patient 1. **(a)** proline, **(b)** ornithine, **(c)** glutamine and **(d)** glutamate. Multiple data points at 32.8 months represent pre- and post-prandial samples. Dotted lines represent reference ranges. All analysis was performed in the Chemical Pathology Department, GOSH, UK.

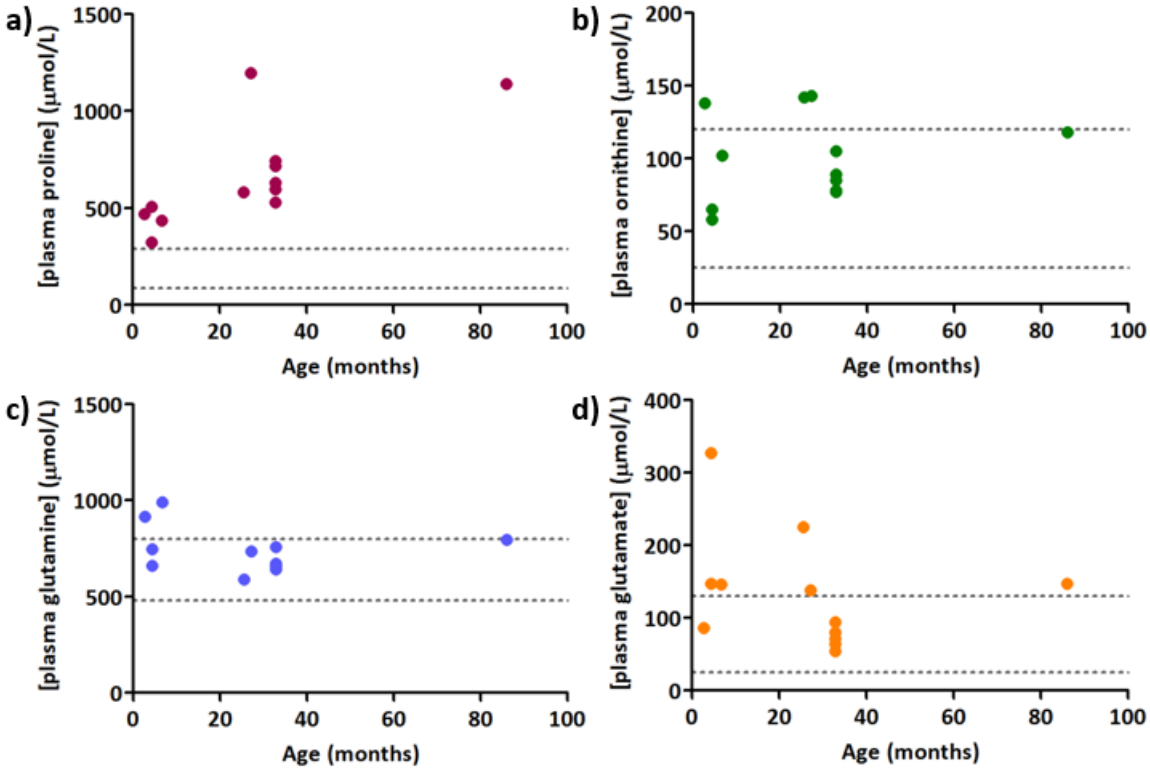


Table 5.4.4: Profile of plasma amino acids for patients 1 - 5. Samples were analysed by high-performance liquid chromatography (HPLC) as part of patients clinical care. Amino acid concentrations elevated above the reference range are shown in orange and those below the reference range are shown in blue. * Amino acids were not quantified at this time.

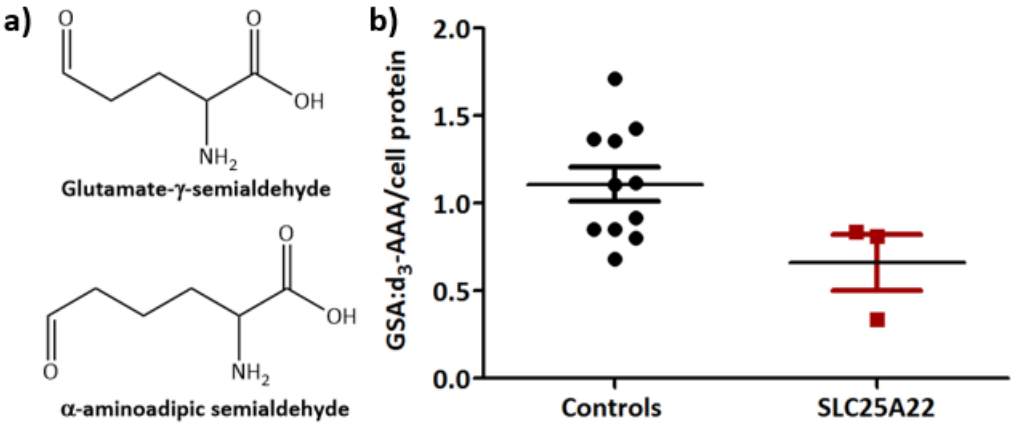
	Patient 1 (µmol/L)											Reference range (µmol/L)
	2 months 19 days	4 months 10 days	4 months 11 days	6 months 23 days	2 years 3 months 4 days	2 years 8 months 25 days	2 years 8 months 25 days	2 years 8 months 25 days	2 years 8 months 25 days	2 years 8 months 26 days	7 years 2 months 18 days	
Glycine	221	226	372	239	434	179	182	205	178	200	363	100 – 330
Serine	197	203	260	182	629	189	201	237	212	232	278	90 – 290
Threonine	135	100	168	143	420	161	187	242	199	236	287	70 – 220
Proline	468	322	506	437	1195	530	595	717	632	743	1139	85 – 290
Leucine	119	91	79	132	265	151	159	255	179	223	170	65 – 220
Isoleucine	71	59	54	71	157	90	89	138	95	115	106	26 – 100
Valine	181	144	162	214	480	294	297	393	319	365	314	90 – 300
Alanine	329	306	535	329	913	236	278	491	289	407	758	150 – 450
Glutamine	914	659	745	991	737	642	656	757	671	759	796	480 – 800
Arginine	101	53	72	112	228	91	97	142	139	141	126	40 – 120
Ornithine	138	58	65	102	143	78	85	89	77	105	118	25 – 120
Lysine	233	125	122	211	470	214	242	378	281	351	359	100 – 300
Methionine	22	22	31	20	48	19	25	41	26	34	34	10 – 60
Taurine	109	44	40	94	71	48	136	54	35	49	91	40 – 140
Phenylalanine	53	37	42	40	100	45	47	74	54	63	57	35 – 100
Tyrosine	73	74	55	64	146	54	65	132	84	113	100	30 – 120
Tryptophan	34	28	26	28	71	36	31	55	34	42	31	30 – 80
Histidine	77	66	72	80	168	72	74	94	74	87	96	30 – 150
Asparagine	56	47	63	51	124	55	59	92	66	79	71	-
Aspartate	9	4	9	4	32	2	3	3	2	3	5	-
Glutamate	86	147	327	146	138	71	94	54	64	80	147	25 – 130

Table 5.4.5: (Continued) Profile of plasma amino acids for patients 1 - 5. Samples were analysed by high-performance liquid chromatography (HPLC) as part of patients clinical care. Amino acid concentrations elevated above the reference range are shown in orange and those below the reference range are shown in blue. * Amino acids were not quantified at this time.

	Patient 2 (µmol/L)		Patient 3 (µmol/L)		Patient 4 (µmol/L)		Patient 5 (µmol/L)	Reference range (µmol/L)
	3 months 16 days	11 months	3 months 16 days	11 months	7 years 6 months 22 days	8 years 0 months 20 days	4 years 6 months 25 days	
Glycine	265	164	216	170	385	427	222	100 – 330
Serine	249	168	187	135	275	210	166	90 – 290
Threonine	250	116	185	94	185	150	134	70 – 220
Proline	261	418	197	274	467	368	115	85 – 290
Leucine	144	192	108	98	205	82	103	65 – 220
Isoleucine	101	115	73	62	130	60	52	26 – 100
Valine	211	319	167	191	339	176	188	90 – 300
Alanine	582	328	321	285	716	820	230	150 – 450
Glutamine	848	614	764	493	851	711	757	480 – 800
Arginine	188	166	159	135	129	115	88	40 – 120
Ornithine	132	60	118	43	142	107	53	25 – 120
Lysine	357	280	280	182	300	124	155	100 – 300
Methionine	40	32	30	21	40	18	18	10 – 60
Taurine	98	53	69	58	111	101	64	40 – 140
Phenylalanine	61	95	52	68	60	37	42	35 – 100
Tyrosine	113	106	86	63	125	57	46	30 – 120
Tryptophan	83	59	60	43	29	21	45	30 – 80
Histidine	86	87	71	70	104	83	80	30 – 150
Asparagine	113	81	81	59	78	*	53	-
Aspartate	9	8	7	5	5	*	5	-
Glutamate	150	139	117	53	100	149	95	25 – 130

In order to investigate whether SLC25A22 can transport GSA or P5C, the concentration of GSA in fibroblasts from SLC25A22-deficient patients were analysed by mass spectrometry. This method was based on the method of Mills *et al.* (2006) to measure α-AASA in patients with pyridoxine-dependent epilepsy (Section 2.15). α-AASA is structurally and physicochemically similar to GSA, differing only in the addition of a -CH_2 group (Figure 5.4.13a). Although not statistically significant, the levels of GSA in the fibroblasts of the three SLC25A22-deficient patients were found to be lower than in control cells (Figure 5.4.13b). This would be consistent with the hypothesis that GSA/P5C cannot exit the mitochondrion and therefore is converted into other metabolites including glutamate and ornithine.

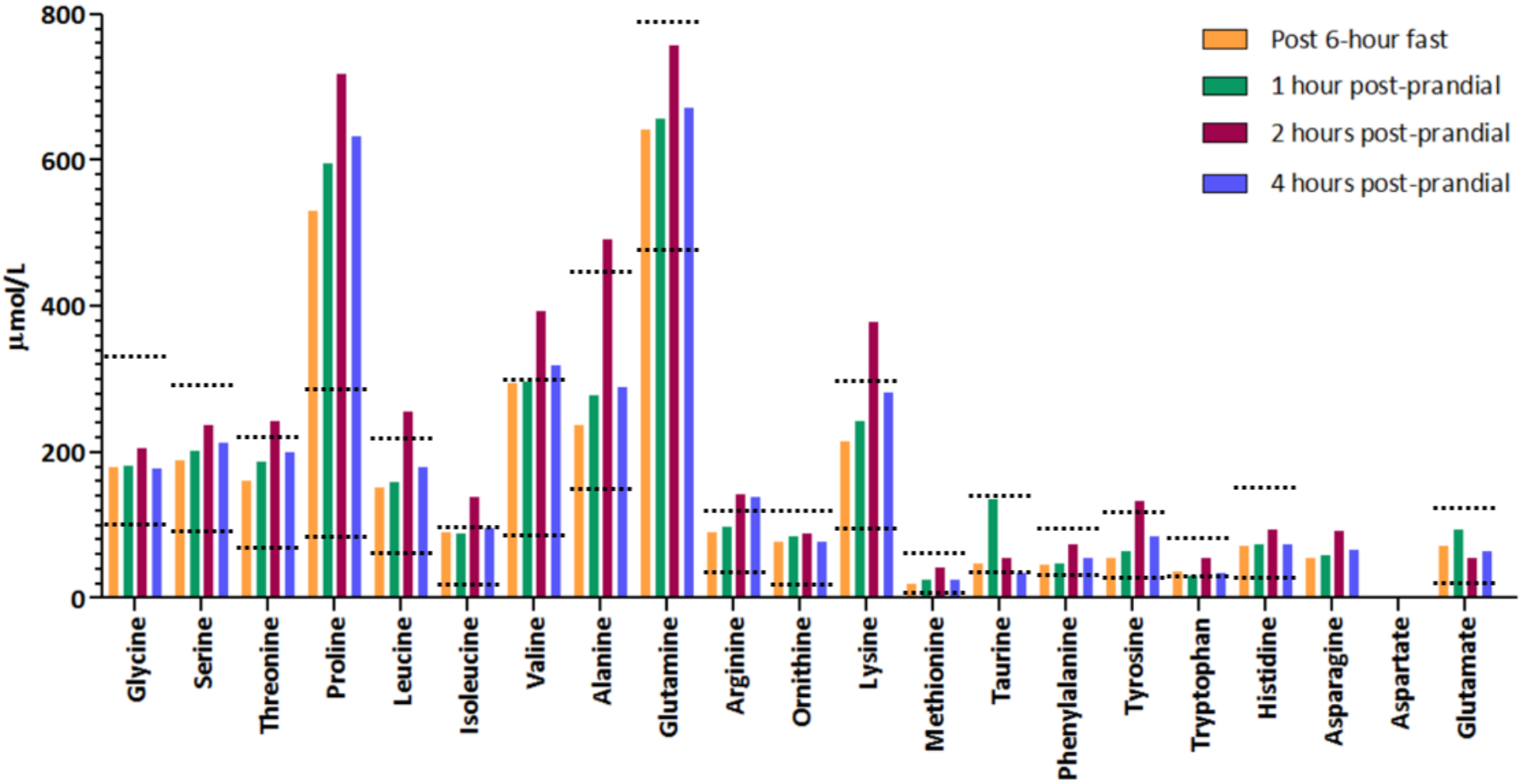
Figure 5.4.13: Measurement of glutamate-γ-semialdehyde (GSA) in control and SLC25A22-deficient fibroblasts. Calibration curves could not be constructed so values were normalised against an internal standard, d_3-aminoadipic acid (d_3-AAA).



5.4.5.1 *Secondary amino acid abnormalities due to glutamate deficiency*

With a view to identifying factors which may have been contributing to the consistent hyperprolinaemia that was presumed to be related to disease pathogenesis in patient 1, a series of post-prandial plasma amino acids were quantified at 2 years 8 months of age. Levels were measured following a 6 hour fast and then 1, 2 and 4 hours post-prandial. Only proline was elevated in the fasting sample, but in the post-prandial samples the concentrations of several amino acids were mildly elevated including leucine, isoleucine, alanine, valine, lysine, tyrosine arginine and threonine (Figure 5.4.14). Random plasma samples from patients 2 - 4 also showed variable elevations of the same amino acids in addition to glutamate, arginine and ornithine. A random plasma sample from patient 5 showed no abnormality. These samples may also have been taken post-prandially.

Figure 5.4.14: Plasma amino acid concentrations in patient 1 at 2 years and 8 months of age in response to feeding. Samples were taken after a six hour fast and 1, 2 and 4 hours after eating. Dotted lines represent the reference ranges used in the clinical laboratory where these analyses were performed (Chemical Pathology Laboratory at GOSH). There were no reference ranges for asparagine and aspartate.



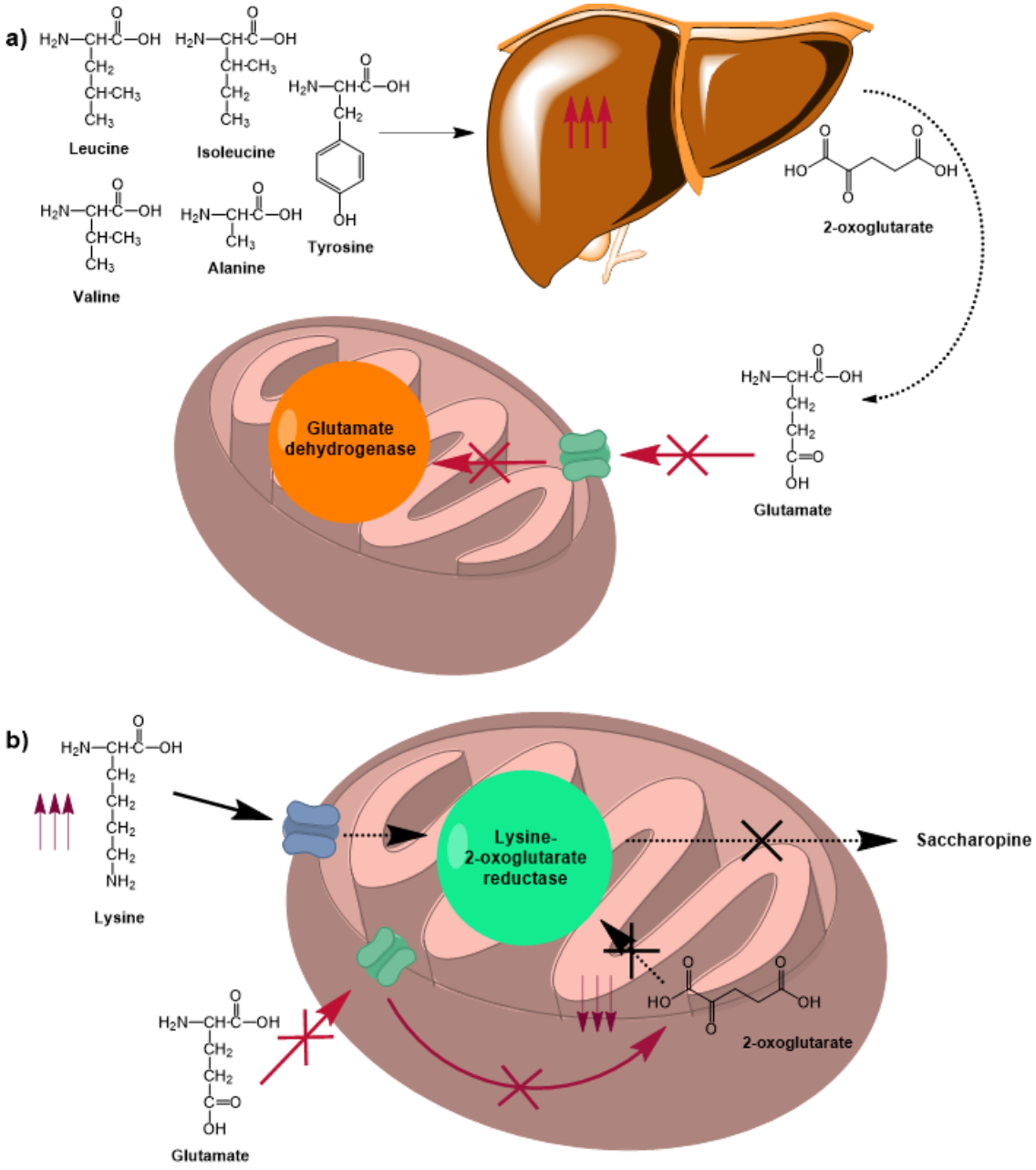
The increase in plasma amino acid concentrations following a protein- or carbohydrate-rich meal is a well-documented effect (Schmid *et al.*, 1992; Frame, 1958; Yokogoshi and Wurtman, 1986). These increases are not uniform, with the absolute increases ranging between 20 - 400 μmol/L. Maximal levels are typically observed between 90 and 120 minutes after the ingestion of a meal. The majority of the post-prandial increases observed in patient 1 were within the ranges reported by others, with the exception of proline, alanine and tyrosine which were outside both the reference ranges reported at GOSH and others published in the literature.

The potential mechanism underlying this widespread post-prandial elevation of amino acids in patient 1 and how it may occur as a consequence of the SLC25A22 transporter defect were then considered. In the fed state, amino acids are the preferred fuel of the liver. The influx from the portal vein leads to high concentrations in the cytosol of the hepatocytes. Leucine, isoleucine, valine, alanine and tyrosine can all undergo reversible transamination with 2-oxoglutarate to produce glutamate (Figure 5.4.15a). These high levels of glutamate need to be transported into the mitochondria for deamination by glutamate dehydrogenase to re-form 2-oxoglutarate, a role that is likely fulfilled by SLC25A22. The transport of glutamate to the mitochondria is therefore critical for replenishing the 2-oxoglutarate levels required for the adequate metabolism of the high concentrations of amino acids entering the liver after a meal. Therefore, when SLC25A22 is dysfunctional, cytosolic deamination will be impaired and hence the plasma concentrations of amino acids such as leucine, isoleucine, valine, alanine and tyrosine will be higher than normal in this post-prandial period, as observed in patient 1. In the case of lysine, the first step in the major catabolic pathway occurs in the mitochondrion but requires 2-oxoglutarate (Blemings *et al.*, 1998) (the catabolic pathways of lysine metabolism are described in detail in Section 6.1). Reduced glutamate influx may lead to reduced 2-oxoglutarate and hence reduced hepatic catabolism of lysine (Figure 5.4.15b).

5.4.5.2 *Urinary and cerebrospinal fluid (CSF) amino acid abnormalities*

Urinary amino acid analysis was only carried out in patients 1 (on two occasions) and patient 5. The first sample from patient 1 at two months of age demonstrated generalised aminoaciduria of uncertain significance. However, it is noteworthy that proline was elevated 6.5-fold higher than the upper limit of the reference range compared to the 3.4-fold average of the other amino acids that were increased in concentration. Nevertheless, the remaining two samples showed consistently elevated proline and glutamate (Table 5.4.6), similar to those observations made in plasma samples. Taurine, tyrosine and aspartate concentrations were also observed to be mildly elevated in urine samples which was not seen in plasma analysis. Urinary amino acid concentrations can be non-specifically affected by renal tubular dysfunction (which can be inherited or secondarily

Figure 5.4.15: Potential mechanisms underlying the secondary amino acid abnormalities identified in patients with SLC25A22 deficiency. **(a)** Impaired cytosolic transamination of amino acids by 2-oxoglutarate. **(b)** Lysine catabolism may be impaired by a lack of 2-oxoglutarate.



acquired), diurnal variation, sample storage, infection and mild vitamin/trace metal deficiencies. Indeed, the significance of these mild amino acid abnormalities in patients 1 and 5 is uncertain.

Table 5.4.6: Profile of urinary amino acids for patients 1 and 5. Samples were analysed by ion exchange chromatography with post column ninhydrin derivatisation using the Biochrom 30+ amino acid analyser as part of patients clinical care. Amino acid concentrations elevated above the reference range are shown in orange and those below the reference range are shown in blue. * Amino acids were not quantified at this time.

	Patient 1 (µmol/mmol creatinine)	Reference range (µmol/mmol creatinine)	Patient 1 (µmol/mmol creatinine)	Reference range (µmol/mmol creatinine)	Patient 5 (µmol/mmol creatinine)	Reference range (µmol/mmol creatinine)
	2 months 21 days		2 years 1 month 20 days		4 years 6 months 22 days	
Glycine	1190	300 – 950	374	250 – 626	277	100 – 400
Serine	316	25 – 95	19	20 – 100	67	20 – 50
Threonine	147	10 – 45	19	10 – 45	39	10 – 25
Proline	719	5 – 110	10	0 – 3	5	0 – 3
Leucine	17	5 – 20	2	3 – 10	9	3 – 10
Isoleucine	9	2 – 40	<1	2 – 10	3	1 – 10
Valine	13	2 – 10	<1	2 – 8	12	2 – 8
Alanine	382	30 – 130	68	30 – 80	60	20 – 80
Glutamine	363	40 – 120	76	30 – 120	146	30 – 120
Arginine	4	2 – 15	<1	2 – 20	3	2 – 10
Ornithine	14	5 – 15	5	2 – 10	3	2 – 10
Lysine	172	5 – 20	27	5 – 30	76	5 – 30
Cystine	24	5 – 35	12	5 – 25	13	5 – 20
Methionine	*	*	11	3 – 15	9	3 – 10
Taurine	607	30 – 55	147	30 – 105	429	40 – 200
Phenylalanine	17	5 – 20	17	5 – 15	15	4 – 15
Tyrosine	51	5 – 30	25	5 – 15	25	5 – 20
Tryptophan	16	1 – 15	<1	1 – 5	10	1 – 6
Histidine	314	50 – 155	203	100 – 300	139	100 – 300
Aspartate	105	10 – 45	50	10 - 35	39	10 – 30
Glutamate	14	2 – 15	24	0 - 3	10	0 - 3

Both patients also underwent CSF amino acid analysis which revealed low levels of glutamate in both patients and elevated proline in patient 1 (Table 5.4.7). This CSF glutamate deficiency was somewhat paradoxical given the elevated concentrations of this amino acid in urine and in approximately half of plasma samples. Two excitatory amino acid transporters, EEAT1 and EEAT3, localised to ependymal and choroid plexus epithelial cells have recently been shown to transport L-glutamate from the CSF to the brain in rats (Akanuma *et al.*, 2015). It has been suggested that the seizures seen in patients with SLC25A22 deficiency are propagated by the accumulation of glutamate in the cytosol of astrocytes and its subsequent inappropriate release into the synaptic cleft (Molinari *et al.*, 2009). It has also been reported that the processes of astrocytes assembled next to the CSF take up glutamate via EAAT1 and EAAT2. This uptake is increased in a dose-dependent manner by the presence of glutamate, mediated by increased cell-surface expression of EAAT1 (Duan *et al.*, 1999). It is therefore plausible that inappropriate glutamate release from astrocytes triggers an increase in EAAT1 cell-surface expression, not only on astrocytes, but also ependymal and choroid plexus epithelial cells resulting in increased glutamate uptake from the CSF and the low levels we observe in our patients.

A second possibility is based on the hypothesis that the direction of glutamate flux through SLC25A22 may be dependent on the intra- and extra-mitochondrial conditions. As also described in relation to mutations in *KCNQ2* (Section 4.5.11), glutamate is the major excitatory neurotransmitter in the mammalian brain. Unlike astrocytes, neurons are unable to carry out *de novo* synthesis of glutamate from glucose. Instead, astrocytes release glutamine into intercellular compartments within the brain which can then be taken up by neurons and recycled back to glutamate by glutaminase (Bak *et al.*, 2006). This enzyme is solely localised to the mitochondria meaning that the synthesised glutamate must be exported to the cytosol to be packaged into synaptic vesicles (Daikhin and Yudkoff, 2000). Thus, dysfunction of SLC25A22 may lead to glutamate being trapped inside the neuronal mitochondria and to the low levels seen in patients 1 and 5.

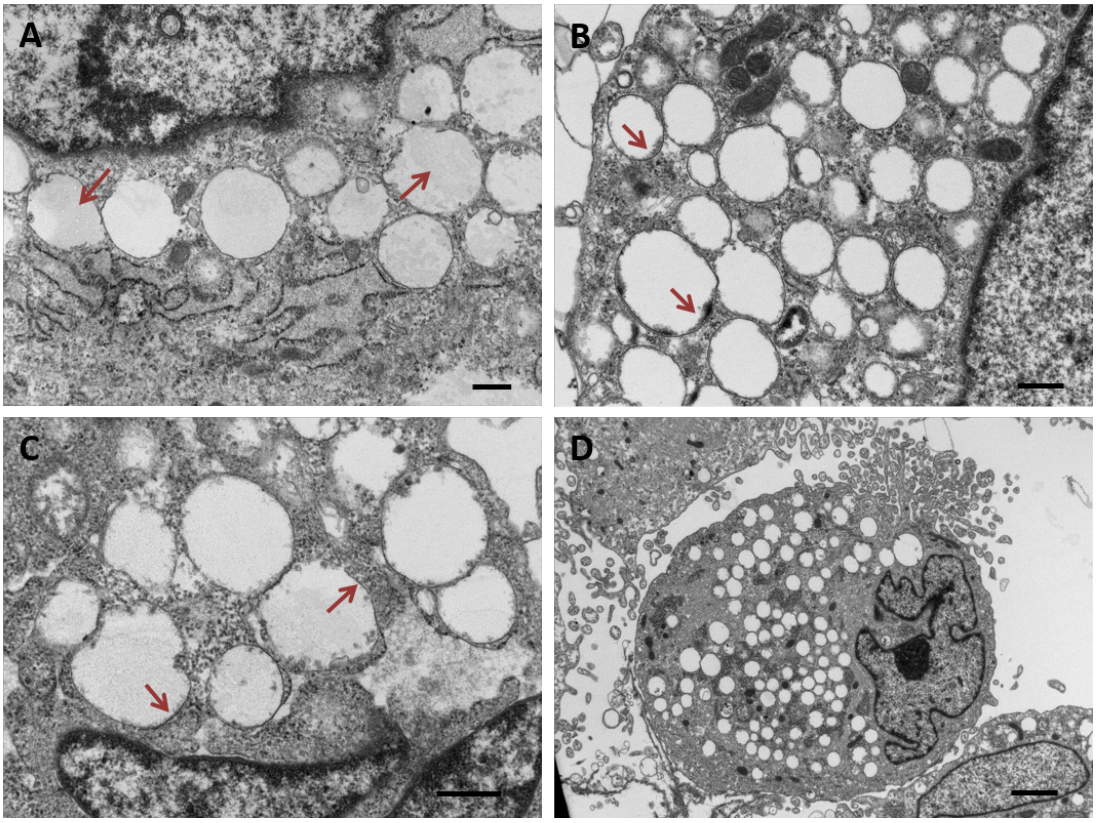
Table 5.4.7: Profile of CSF amino acids for patients 1 and 5. Samples were analysed by high-performance liquid chromatography (HPLC) as part of patients clinical care. Amino acid concentrations elevated above the reference range are shown in orange and those below the reference range are shown in blue.

	Patient 1 (µmol/L)	Patient 5 (µmol/L)	Reference range (µmol/L)
	2 months 19 days	4 years 6 months 25 days	
Glutamate	1	1	5 – 17
Threonine	56	36	10 – 45
Serine	72	31	10 – 81
Glycine	4	4	4 – 14
Proline	12	<1	0 – 1
Leucine	22	14	4 – 18
Isoleucine	11	6	3 – 8
Valine	26	16	5 – 25
Alanine	31	25	16 – 36
Glutamine	621	458	420 – 600
Arginine	22	19	15 – 40
Ornithine	7	3	3 – 13
Lysine	27	18	10 – 32
Methionine	3	2	2 – 6
Taurine	9	5	3 – 10
Histidine	14	14	3 – 18
Tryptophan	2	3	1 – 3
Phenylalanine	9	7	5 – 15
Tyrosine	21	10	5 - 15

5.4.6 *Investigation of fibroblast vacuolation identified in patients with SLC25A22 mutations*

In addition to amino acid analyses which, although revealing subtle abnormalities remained non-diagnostic, three patients had additional extensive biochemical investigations as part of the diagnostic work-up for a potential metabolic disorder (Table 5.3.1). This included the examination of fibroblasts obtained from a skin biopsy by electron microscopy to investigate any potential storage or mitochondrial disorder. Observation of the cellular ultrastructural morphology revealed extensive vacuolation with the presence of empty, single membrane-bound vacuoles being evident in all patients (Figure 5.4.16). This is a characteristic feature of defects of lysosomal enzymes and can indicate the presence of a storage material, with each type having a distinctive ultrastructural appearance (Alroy and Ucci, 2006). The vacuoles were of a regular size and present in the vast majority of cells examined. At higher magnification, some of the vacuoles in the SLC25A22-deficient patients clearly contained electron-lucent or fine fibrillar material or had remnants of this material adjacent to the membrane (Figure 5.4.16). These appearances were consistent with the lysosomal accumulation of lipids.

Figure 5.4.16: Ultrastructural features of the vacuoles identified in patients with *SLC25A22* mutations. (A and B) Vacuoles contain (or contain remnants of) fine fibrillar and electron-lucent material. **(C)** Vacuoles are bound by a single membrane. **(D)** Vacuoles are regularly sized. Ultrastructure of control fibroblasts can be seen in Figure 5.4.17. Scale bars: **A-C**: 500 nm; **D**: 2 μm.

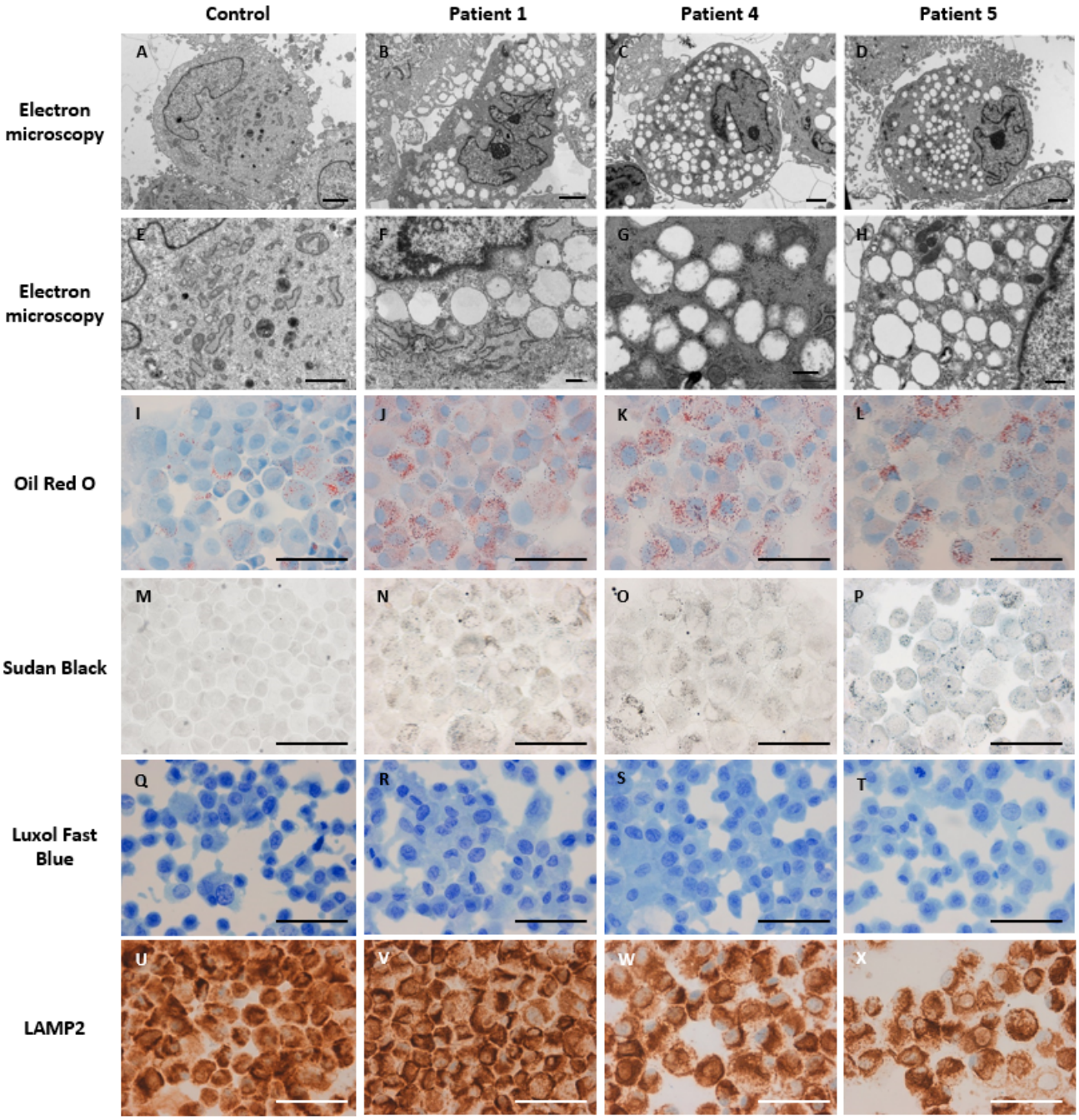


5.4.6.1 *Histological staining and interrogation of whole exome data to investigate the contents of the vacuoles in patient fibroblasts*

In order to investigate whether any soluble storage material had been washed away during sample processing, fibroblasts were re-examined using both tinctorial stains and immunohistochemical methods (Figure 5.4.17). The histological identification of lipids is challenging because droplets can dissolve upon fixation or staining with alcohol-based chemicals. Hence, when carrying out each stain, care was taken not to fix cells with alcohol and to minimise the time slides were washed in alcoholic solutions. Fibroblasts from all patients showed a punctate staining pattern using Oil Red O (Section 2.14.1.3) and Sudan Black (Section 2.14.1.4) indicating the presence of both neutral lipids and phospholipids, respectively. Luxol Fast Blue staining was negative for the presence of sphingomyelin (Section 2.14.1.5). LAMP2 immunostaining was comparable to controls providing no evidence to support any lysosomal dysfunction (Section 2.14.1.6). These results were consistent with the vacuoles seen by electron microscopy containing soluble lipid that had been washed away during sample processing. Two possibilities were then considered:

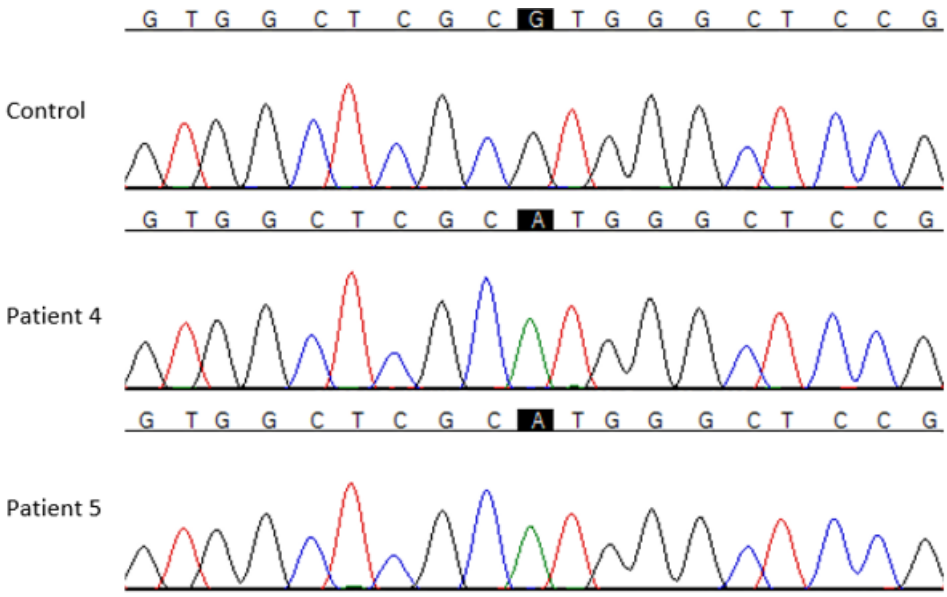
both families have a second metabolic disorder leading to the production of lipid vacuoles or that the accumulation was secondary to SLC25A22 deficiency.

Figure 5.4.17: Structural and ultrastructural examinations of SLC25A22-deficient fibroblasts. Fibroblast cell culture from control **(A, E, I, M, Q, U)**, patient 1 **(B, F, J, N, R, V)**, patient 4 **(C, G, K, O, S, W)** and patient 5 **(D, H, L, P, T, X)**. Ultrastructural examination by electron microscopy revealed widespread, almost exclusively empty vacuoles in all patients **(A-H)**. Patients showed excess accumulation of lipid (**I-L**, Oil Red O; **M-P**, Sudan Black). Luxol Fast Blue staining was negative for the presence of sphingomyelin (**Q-T**). LAMP2 immunostaining was comparable to controls (**U-X**). Scale bars: **A-E**: 2 μm; **F-H**: 500 nm; **I-X**: 100 μm.



In light of the possibility that a second metabolic disorder was present in the patients described in this chapter, whole exome sequencing data was scrutinised for variants in genes known to cause lysosomal storage disorders and a homozygous variant was identified in *SMPD1* (c.340G>A; p.Val114Met) in patients 4 and 5.

Figure 5.4.18: Electropherograms showing SMPD1 mutation analysis in patient 4 and 5. Sanger sequencing confirmed the presence of a homozygous p.Val114Met mutation in *SMPD1* in patients 4 and 5.



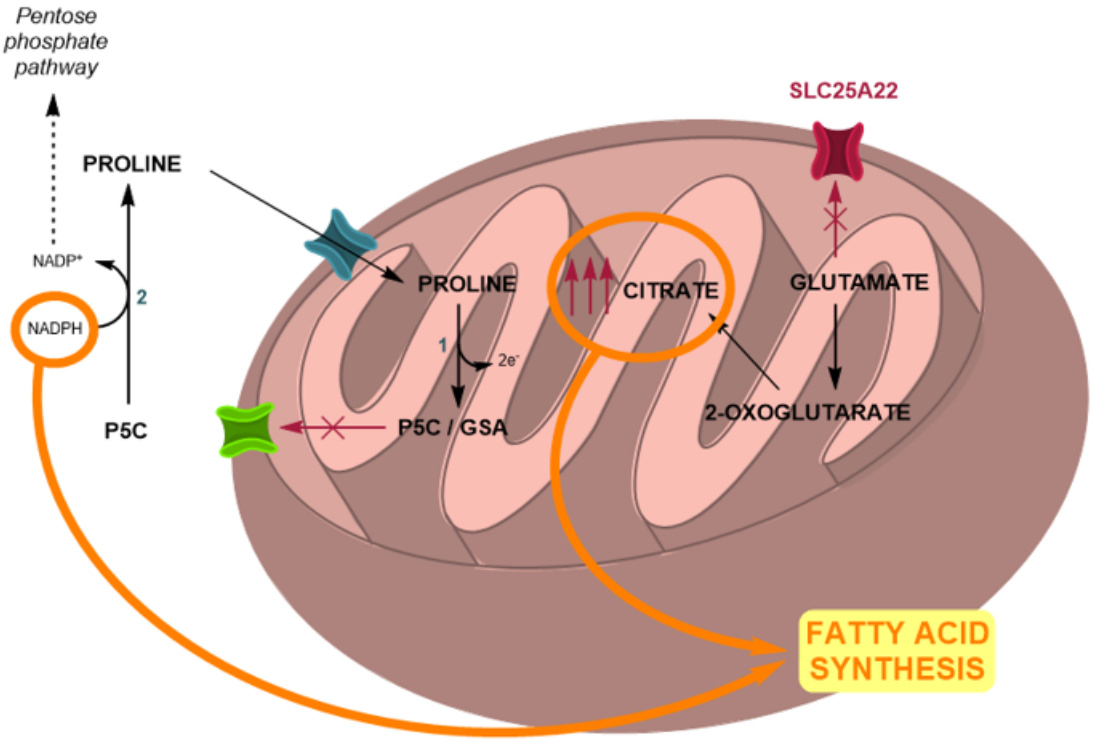
The variant is reported in publicly available databases at a minor allele frequency of less than 1% (1000 Genomes Project, 0.06%; NHLBI Exome Sequencing Project, 0.06%, ExAC, 0.076%) and is predicted tolerated and benign by SIFT and PolyPhen-2, respectively. Segregation of the variant and confirmation of its presence in a homozygous state in both patients was confirmed by Sanger sequencing (Figure 5.4.18). Mutations in this gene cause Niemann-Pick disease type A/B which presents either in early infancy or mid-childhood with a combination of the following features: hepatosplenomegaly, psychomotor regression, interstitial lung disease, delayed bone mineralisation, thrombocytopenia and a cherry-red spot in the macula of the eye (Irun *et al.*, 2013). Our patients do not share any of these features. Leukocyte acid sphingomyelinase (encoded by *SMPD1*) activity was subsequently measured as 0.43 nmol/hr/mg protein (ref: 0.86 – 2.8) by the Chemical Pathology Department, GOSH. Although this activity is lower than that reported in heterozygotes (ref: 0.7 – 1.3), it is higher than that reported in affected homozygotes (ref: 0.08 – 0.18). Luxol fast blue (LFB) staining has been used previously to detect sphingomyelin, the storage material that accumulates in Niemann-Pick type A/B (Aronson and Volk, 2013). Given that the vacuolation was present to a comparable degree in all patients and there was no evidence of punctate LFB staining indicating an accumulation of sphingomyelin (Figure 5.4.17), it was concluded that this *SMPD1* variant was not pathogenic.

Accumulation of neutral lipids can occur in defects of adipose triglyceride lipase (*PNPLA2*) or its activator, 1-acylglycerol-3-phosphate O-acyltransferase (*ABHD5*) causing neutral lipid storage disease with myopathy and Chanarin-Dorfman syndrome, respectively. Lipids can also accumulate intracellularly in patients with Wolman disease, deficiencies of carnitine, carnitine palmitolytransferase and mitochondrial fatty acid oxidation enzymes. Whole exome sequencing data was scrutinised for potentially pathogenic variants in any of the genes responsible for these diseases. None were identified, therefore we concluded that our patients do not harbour a second genetic defect resulting in excess lipid production. Following the exclusion of a secondary metabolic defect in our patients, two possibilities regarding how lipid vacuole accumulation could be secondary to SLC25A22 deficiency were considered.

5.4.6.2 *Vacuolation due to impaired transport of reducing equivalents*

One hypothesis would be that an excess of nicotinamide-adenine dinucleotide phosphate (NADPH) and citrate within the cytosol may drive increased lipid synthesis and the second was that SLC25A22-deficiency may drive an increase in autophagy (Figure 5.4.19).

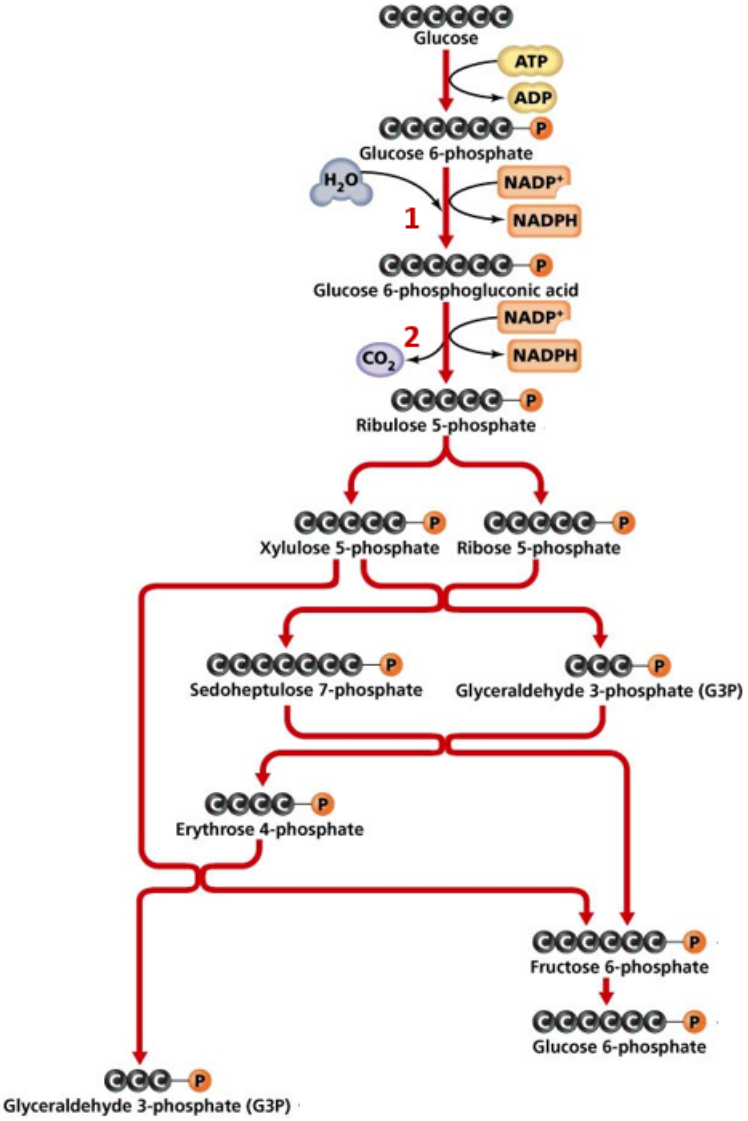
Figure 5.4.19: Proposed mechanism of lipid synthesis based on the impaired transport of reducing equivalents and accumulation of citrate in patients with SLC25A22 deficiency. 1, proline dehydrogenase; 2, P5C reductase.



In both proliferating and quiescent fibroblasts the pentose phosphate pathway is the major source of cytosolic NADPH through the oxidation of glucose (Lemons *et al.*, 2010). These

reducing equivalents are then transferred to the mitochondria via the proline/P5C shuttle. Firstly, hydride ions (H^-) from NADPH are transferred to proline and subsequently to the respiratory chain upon the conversion back to P5C by proline dehydrogenase (Hagedorn and Phang, 1983). The P5C then exits the mitochondria and is converted back to proline by the cytosolic P5C reductase. The proline/P5C shuttle can also regulate cytosolic and mitochondrial redox states because the production of $NADP^+$ by P5C reductase increases the activity of glucose-6-phosphate dehydrogenase which catalyses the rate-limiting step of the pentose phosphate pathway. Both this enzyme and 6-phosphoglutaconate dehydrogenase, which catalyses the following enzymatic conversion, produce one molecule of NADPH through the utilisation of $NADP^+$ (Figure 5.4.20). Therefore, this upregulation of glucose-6-phosphate dehydrogenase by $NADP^+$ from the proline/P5C shuttle, results in increased production of cytosolic NADPH.

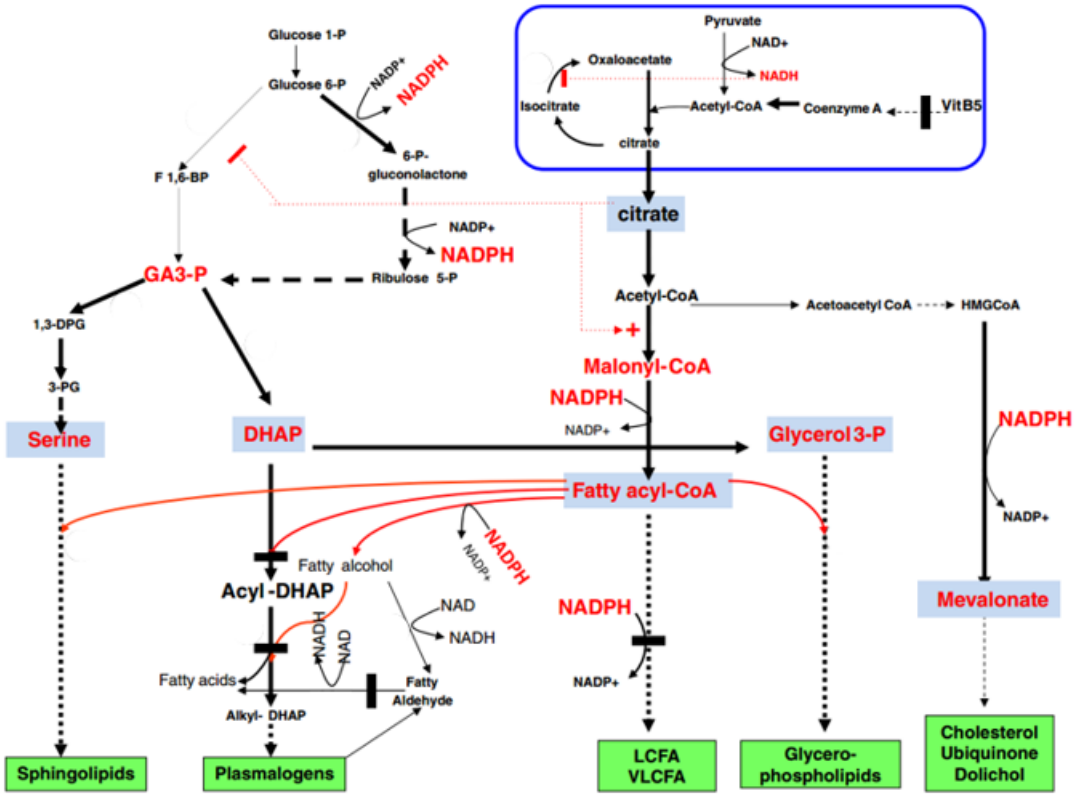
Figure 5.4.20: Schematic of the pentose phosphate pathway. 1, glucose-6-phosphate dehydrogenase; 2, 6-phosphoglutaconate dehydrogenase.



Given that one molecule of succinate entering the tricarboxylic acid (TCA) cycle generates 100 times more ATP than the proline/P5C shuttle, it is unlikely that this mechanism represents a major energy source. However, this may be untrue in certain conditions, including situations where the TCA cycle is not operating maximally. This may be the case in SLC25A22 deficiency due to the abnormal transport of glutamate. Within the mitochondria, glutamate dehydrogenase catalyses the conversion of glutamate to 2-oxoglutarate which then enters the TCA cycle. Thus, as in glutamate dehydrogenase deficiency (Nissen *et al.*, 2015), a lack of 2-oxoglutarate could cause TCA cycle dysfunction, possibly increasing the functional importance of the proline/P5C shuttle. If, as we have proposed, SLC25A22 facilitates the export of P5C (or its equilibrium partner GSA) from the mitochondria, mutations would result in shuttle impairment and an increase of cytosolic NADPH/$NADP^+$.

A second hypothesis is that citrate may accumulate in patients with SLC25A22 deficiency. When quiescent cells are supplemented with stable isotope-labelled glutamine, approximately 15% of the cellular citrate became labelled (Lemons *et al.*, 2010). The only known mechanism to generate the specific configuration of labelled citrate identified is through the conversion of glutamine → glutamate → 2-oxoglutarate → citrate; this corresponds to a reversed direction of TCA cycle flux. If glutamate is trapped inside the mitochondrion a similar increase in citrate synthesis could occur. Under normal conditions, glucose is converted to pyruvate by glycolysis in the cytosol. Pyruvate is then converted to acetyl-coA followed by citrate within the mitochondria. This citrate is then exported to the cytosol where it is cleaved by ATP citrate lyase back to acetyl-coA. Carboxylation then occurs by acetyl-coA carboxylase to form malonyl-coA. This is the first step in the synthesis of a range of lipids including sphingolipids, plasmalogens, fatty acids and glycerophospholipids (Figure 5.4.21).

Figure 5.4.21: Schematic illustrating the requirements of citrate and $NADP^+$/NADPH for the synthesis of fatty acids and complex lipids. Figure adapted from Lamari *et al.* (2013)).



Similarly, NADPH is a critical cofactor for many enzymes involved in lipid synthesis (Figure 5.4.21). Thus, it would be intuitive that any increase in the cytosolic concentrations of NADPH or citrate could result in an increase in lipid synthesis. Indeed, one of the postulated mechanisms for the lipid accumulation often seen in muscle biopsies of patients with mitochondrial respiratory chain disorders is the build-up of reducing equivalents and export of TCA cycle intermediates such as citrate. It is therefore possible that, providing that SLC25A22 can catalyse glutamate exchange *in vivo*, impairment of the transporter could lead to conditions favouring increased lipid synthesis in fibroblasts.

5.4.6.3 *Vacuolation due to increased autophagy*

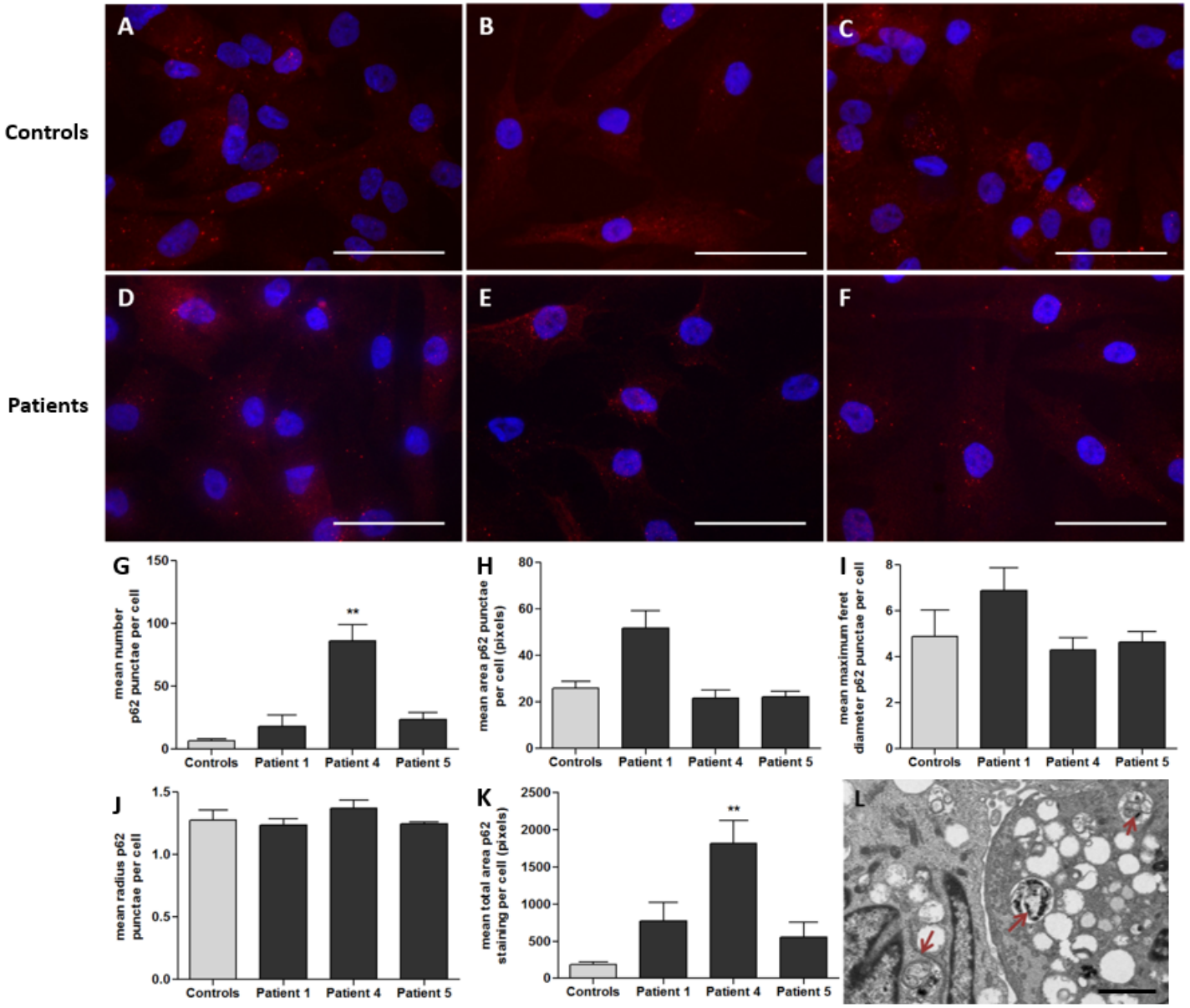
A second hypothesis for the formation of the vacuoles evident in patient fibroblasts is that abnormal mitochondrial metabolism triggers an autophagy response. Autophagy is an intracellular pathway which is required for the elimination of damaged organelles and intracellular pathogens, as well as the recycling of cytoplasmic macromolecules (Levine and Kroemer, 2008). Three distinct types of autophagy have been identified to-date: chaperone-mediated autophagy, microautophagy and macroautophagy. These share common degradative mechanisms within the lysosome but differ in

the way that substrates are delivered to the organelle. The predominant type, macroautophagy, is a synonym for autophagy and involves the formation of a double membrane structure around the material targeted for destruction. This autophagosome then travels through the cytoplasm and fuses with a lysosome, which uses acidic hydrolytic enzymes to degrade the contents of each autophagosome (Klionsky, 2007). Amino acids are known to influence autophagy through the regulation of the mammalian target of rapamycin complex 1 (mTORC1), a protein complex which functions as a sensor of nutrient, energy and redox state within cells and controls protein translation accordingly (Hay and Sonenberg, 2004). Amino acids suppress autophagy, whereas their deficiency stimulates the pathway (Mortimore and Schworer, 1977). This increase in autophagy under conditions of stress promotes the recycling of nutrients to enable the production of glucose, signalling molecules and neurotransmitters, as well as the maintenance of immune and endocrine function (Wu, 2013). In many cell types, the majority of the stimulatory/inhibitory effects can be attributed to concentrations of leucine and, to a lesser extent, the other branched chain amino acids (Lynch, 2001). However, *in vitro* studies have shown that pheochromocytoma cells supplemented with low concentrations of glutamate undergo an autophagic pro-survival response which can be accompanied by cytoplasmic vacuolation, reminiscent of that observed in the fibroblasts of our patients (Stamoula *et al.*, 2015). Thus, it is possible that the dysregulation of glutamate levels both within the cytosol and mitochondria of SLC25A22-deficient cells may result in increased autophagy.

Accumulation of nucleoporin 62 (p62) can be used as a marker for the induction of the autophagy pathway. Therefore to investigate whether mitochondrial dysfunction caused by *SLC25A22* mutations may drive an increase in autophagy/mitophagy and account for the fibroblast vacuolation, the amount and subcellular localisaation of p62 was investigated in patient fibroblasts by immunofluorescence (Section 2.14.3). Proteins can be targeted for autophagic degradation through ubiquitination, a process which is dependent on the binding of p62 to targeted substrates (Kim *et al.*, 2008). This complex then associates with proteins within the autophagosome membrane via an LC3-interacting region prior to transporting the cargo to the lysosome (Lamark *et al.*, 2009). Four representative immunofluorescence images per case were analysed using CellProfiler 2.1.1. Image analysis showed that the mean feret diameter and radius of p62 punctae remained constant between controls and patients with *SLC25A22* mutations. However, the mean number and area of autophagosomes appeared to be increased in patient cells indicating a possible increase in autophagy or mitophagy due to mitochondrial dysfunction (Figure 5.4.22). This increase in number of p62-positive autophagosomes was significant ($p = 0.002$) in patient 4. Similarly, when the results were taken together, the mean total area of p62 staining per cell was increased in all patients compared to controls and significantly in patient 4

($p = 0.002$). In order to investigate this further, electron microscopy images were scrutinised. In all patients, the majority of membrane-bound vacuolar structures were empty. However, some vacuoles in patient 4 contained cellular debris or electron-dense lamellar bodies (Figure 5.4.22L). This would be consistent with the p62 staining and indicates a greater degree of autophagy occurring in these cells (Ylä-Anttila *et al.*, 2009). Further work is needed to determine whether this trend is a consistent observation and how it is affected by stressing mitochondria or blocking autophagic degradation using bafilomycin or chloroquine.

Figure 5.4.22: Immunofluorescence of the autophagy marker nucleoporin 62 (p62). **(A)** Control 1, **(B)** Control 2, **(C)** Control 3, **(D)** Patient 1, **(E and L)** Patient 4, **(F)** Patient 5. Blue, nuclear staining with DAPI; red, p62 punctae. Analysis of mean **(G)** number, **(H)** area, **(I)** maximum ferret diameter, **(J)** radius and **(K)** and total area of p62 staining per cell (** p = 0.002). **(L)** Electron microscopy of patient 4. Scale bars: **A-F**: 50 nm; **L**: 2 μm.



5.4.7 *Investigation of the efficacy of ubiquinone treatment for patients with SLC25A22 deficiency*

An additional noteworthy biochemical finding for patient 1 was that he had a low white cell ubiquinone level of 24 pmol/mg (ref: 37 - 133). This test was not performed in patients 4 or 5. Our hypothesis for the elevated proline levels observed in patients with SLC25A22-deficiency is that disruption of the proline/P5C shuttle results in reduced flux through proline dehydrogenase. Reducing equivalents from proline are transferred via a flavin adenine dinucleotide (FAD) cofactor to an acceptor. Although detailed kinetic studies of human proline dehydrogenase have not been carried out to-date, characterisation of proline utilization A (PutA) in Gram-negative bacteria and proline dehydrogenase from *Thermus thermophilus* have provided mechanistic insights. Both enzymes share high sequence identity with human proline dehydrogenase, particularly within the active site region. The conversion of proline to P5C by proline dehydrogenase occurs via a ping-pong mechanism, meaning that the first product is formed in a reductive half-reaction and released prior to the second substrate binding and the completion of the oxidative half-reaction. In the former, proline binds to the active site and is oxidised to P5C with the concurrent reduction of the FAD cofactor to $FADH_2$. In the oxidative half-reaction these electrons are then transferred from $FADH_2$ to ubiquinone situated within the membrane, which binds to to the enzyme at an alternative site (Moxley *et al.*, 2011). Although remaining unproven, it is likely that proline dehydrogenase in humans also uses a similar mechanism to transfer electrons from $FADH_2$ to ubiquinone in the respiratory chain. The pair of electrons can also result in the reduction of oxygen to form superoxide (Tanner, 2008). Indeed, ubiquinone could be being consumed due to impaired redox balance, reduced proline flux somehow leading to increased losses of ubiquinone, or a combination of the two.

Ubiquinone plays a central role in the respiratory chain within mitochondria by carrying electrons from complex I and II to complex III (Crane *et al.*, 1957). Inborn errors affecting ubiquinone biosynthesis typically present with brain involvement including seizures, cognitive impairment and cerebellar signs (Quinzii *et al.*, 2007). Unfortunately, ubiquinone supplementation in patients 2 and 3 did not ameliorate the onset of a seizure disorder despite being commenced at birth. This indicates that respiratory chain dysfunction secondary to ubiquinone deficiency within the mitochondria is not a major pathogenic mechanism in SLC25A22 deficiency; this is also supported by normal activity of the respiratory chain enzymes in muscle in patients 1 and 4.

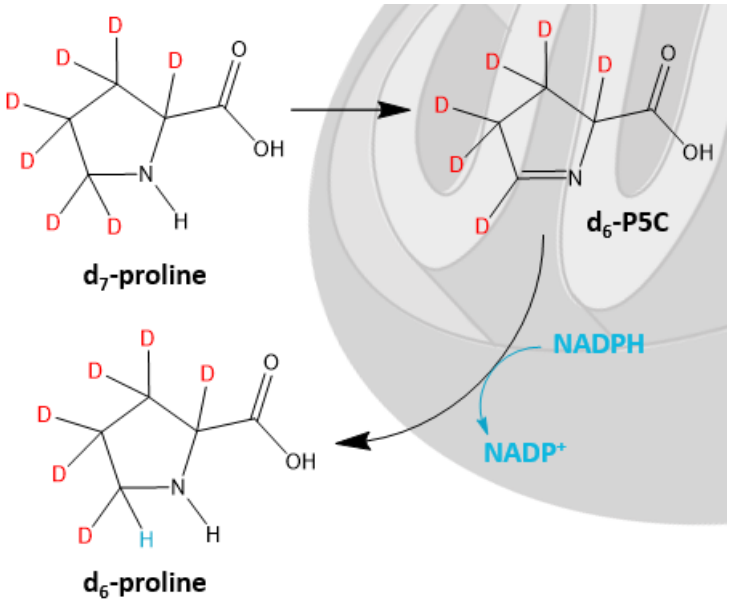
5.5 SUMMARY

In this chapter, the cases of five children from two unrelated families found to have a severe developmental disorder accompanied by early-onset seizures in the majority of cases are described. Whole exome sequencing was utilised to identify mutations in *SLC25A22* in three patients including the proband in each family. Three siblings were subsequently prenatally diagnosed by Sanger sequencing.

This cohort again demonstrates the power of using whole exome sequencing for the identification of patients presenting with non-classical phenotypes and for providing potential new insights into protein function. We have shown that mutations in *SLC25A22* can not only cause a neonatal-onset seizure disorder, but also a developmental disorder presenting without seizures until seven years of age. In this particular patient, the absence of seizures but presence of developmental delay and fibroblast vacuolation would likely have resulted in screening using a metabolic next-generation sequencing gene panel as opposed to an epilepsy panel. Thus, it is likely that this disorder remains underdiagnosed as only patients with a seizure disorder will currently be screened for mutations in *SLC25A22*. In addition, we have demonstrated the presence of biochemical abnormalities including hyperprolinaemia, low CSF glutamate and lipid accumulation in our cohort that may provide useful diagnostic indicators in future patients. These abnormalities may have not been reported previously due to the specific investigations not being carried out or amino acid concentrations not being dramatically elevated above reference ranges. Increased cytosolic lipid synthesis could indicate an accumulation of cytosolic NADPH and this, alongside hyperprolinaemia suggests impairment of the proline/P5C shuttle. Thus we hypothesise that SLC25A22 functions *in vivo* as a mitochondrial P5C/GSA transporter in addition to its previously described function as a glutamate transporter. Future studies on additional cohorts of patients presenting with these features will be required to determine the full phenotypic range associated with this gene. Finally, further functional studies to determine whether or not SLC25A22 can catalyse the transport or exchange of P5C or GSA will need to be undertaken. This could be carried out through the reconstitution of SLC25A22 into liposomes as descibed by Molinari *et al.* (2005) and/or by measuring of flux through the proline/P5C shuttle in patient cells using stable isotopes. The latter could be achieved by culturing fibroblasts in media containing d_7-proline; this isotope would then enter the mitochondria and undergo enzymatic oxidation by proline dehydrogenase. The creation of the double bond in P5C would result in the removal of one of the deuterium atoms and the formation of d_6-P5C. If SLC25A22 can transport GSA, the d_6-P5C would be expected to non-enzymatically hydrolyse to form d_6-GSA before undergoing efflux to the cytosol and re-cyclisation to d_6-P5C. P5C reductase would then complete

the proline/P5C shuttle. However, in this case the replacing hydrogen atoms would be derived from NADPH rather than a stable isotope (deuterium). Thus, the recycled proline would be in the d_6 form (Figure 5.5.1). The ratio between d_7- and d_6-proline could then be used to assess any differences in function of the proline/P5C shuttle in patient cells.

Figure 5.5.1: Schematic showing the conversion of d_7-proline to d_6 proline that would occur if SLC25A22 transports P5C/GSA.



6

CYSTEINE CONJUGATE β-LYASE: A PROTEIN OF MANY FUNCTIONS?

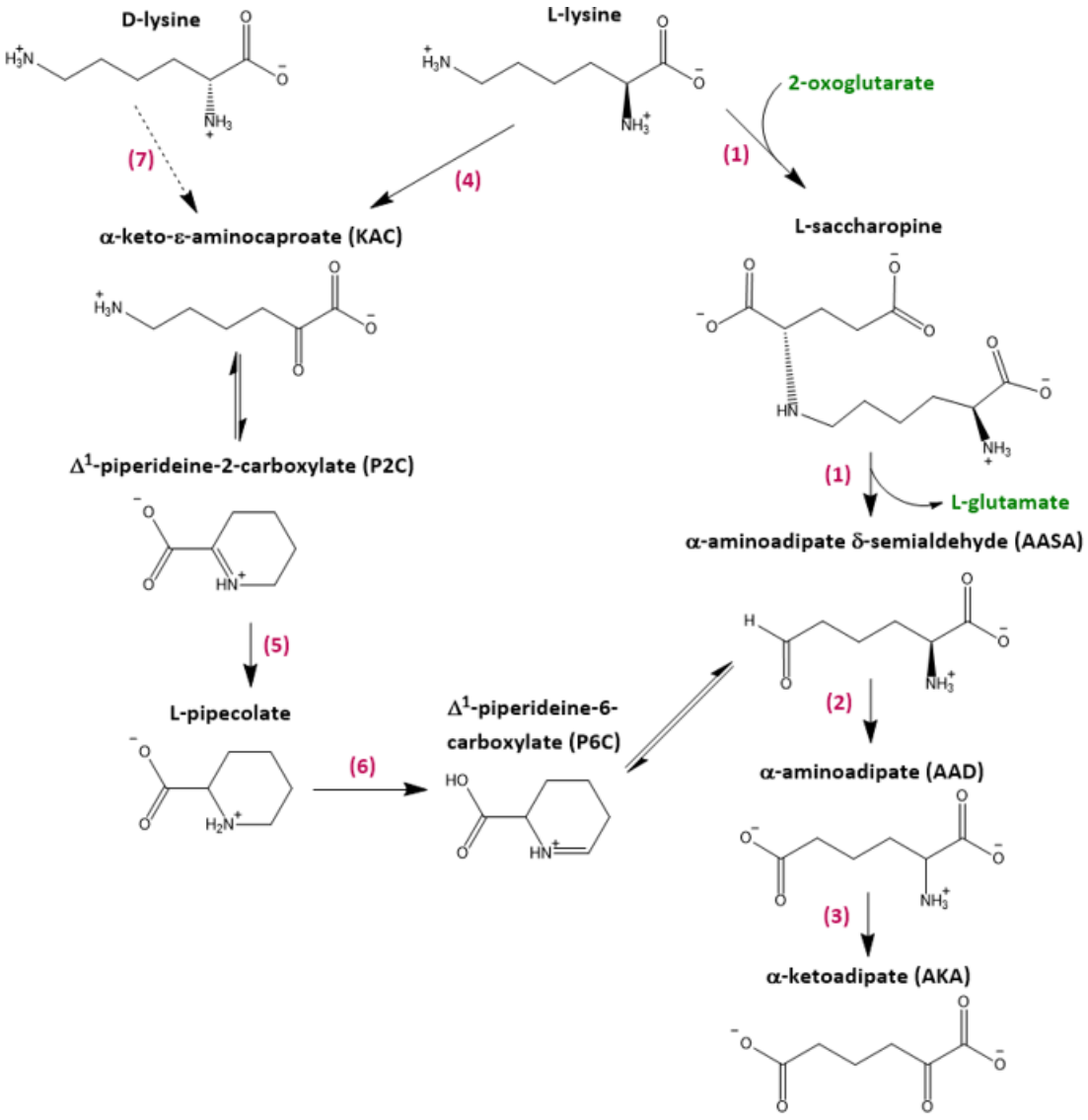
6.1 INTRODUCTION

In this chapter, whole exome sequencing was used to investigate the underlying aetiology of disease in a patient Y presenting with a movement disorder and developmental delay, alongside persistently elevated lysine concentrations in blood, urine and CSF.

Lysine is one of the nine essential proteinogenic amino acids in humans; its side chain consists of a charged aliphatic group ($(CH_2)_4NH_2$). It is an essential precursor for many molecules which are important for the optimal function of the mammalian central nervous system including the excitatory neurotransmitter glutamate (Papes *et al.*, 2001), carnitine which is required for the β-oxidation of fatty acids (Tanphaichitr *et al.*, 1971) and intermediates such as crotonyl-CoA which can provide carbon-units to the tricarboxylic acid (TCA) cycle for ATP production (Sauer *et al.*, 2011).

The degradation of lysine occurs via two distinct pathways (Figure 6.1.1). The first, termed the *saccharopine pathway*, predominates in extracerebral tissues and the foetal brain. The second, termed the *pipecolate pathway*, is the major catabolic pathway in the adult brain. Indeed, enzymatic assays have demonstrated a reduction in the capacity of the saccharopine pathway and an increase in that of the pipecolate pathway throughout gestation and neonatal life within the brain (Rao *et al.*, 1992). This suggests that pipecolic acid or the other intermediate metabolites of this pathway may play a role in neuronal development. Although the two pathways are typically spatially isolated, they share common intermediates and converge to form a common degradative mechanism (Figure 6.1.1). The catabolism of lysine is also unique because transamination (i.e. the transfer of the α-amino group to an α-keto acid acceptor which occurs early-on in the catabolism of most amino acids) is irreversible. This is due to the production of a compound that spontaneously and rapidly cyclises, thereby not favouring the reverse transamination reaction. This occurs regardless of whether the initial catabolic steps involve the conversion of the α-amino group to a keto group (pipecolate pathway) or the conversion of the ϵ-amino group to an aldehyde (saccharopine pathway).

Figure 6.1.1: Catabolic pathways of of lysine metabolism. The saccaropine pathway (right) is the major pathway in tissues outside of the brain. The pipecolate pathway (left) is the major pathway within the mammalian brain. **1**, L-lysine-α-ketoglutarate reductase/saccaropine dehydrogenase (LKR/SDH); **2**, α-aminoadipate semialdehyde dehydrogenase (AASDH); **3**, α-aminoadipate aminotransferase/kynurenine aminotransferase II (AADT/KATII), **4**; currently unknown (converts α-amino group of L-lysine to a keto group); **5**, ketimine reductase μ crystallin (CRYM/KR); **6**, L-pipecolate oxidase (POX); **7**, D-amino acid oxidase.



Given the persistent elevation of lysine in the biofluids of the patient described in this chapter (patient Y), extensive genetic testing of genes known to be involved in the metabolism of lysine had been performed prior to the beginning of this PhD project. Sanger sequencing had failed to identify mutations in **(1)** α-aminoadipic semialdehyde synthase (*AASS*), **(activator of 2)** L-aminoadipate-semialdehyde dehydrogenase-phosphopantetheinyl transferase (*AASDHPPT*),

(2) L-aminoadipate-semialdehyde dehydrogenase (*AASDH*), **(6)** pipecolic acid oxidase (*PPOX*) and **(7)** D-amino acid oxidase (*DAO*).

Despite well-established causes of hyperlysinaemia (mutations in *AASS*) being excluded, mild elevations of lysine are often seen in disorders which result in decreased availability of α-ketoglutarate (required for the formation of saccharopine) such as urea cycle disorders, propionic acidaemia and methylmalonic aciduria (Kamoun *et al.*, 2002). Furthermore, the movement disorder exhibited by patient Y shared similarities with that seen in patients with amyotrophic lateral sclerosis (ALS). Levels of lysine have been shown to be significantly elevated in the plasma of patients with ALS, a neurodegenerative disease which affects approximately 5000 patients in the UK and results in the progressive loss of motor functions due to muscle paralysis and ultimately, death. This accumulation was particularly pronounced in patients with early-onset symptoms, although the reason for this biochemical abnormality is unclear (Cecchi *et al.*, 2014). Due to the failure of targeted single-gene sequencing to find a diagnosis for patient Y, whole exome sequencing was performed.

6.2 METHODS

Sanger sequencing of the *CCBL1* gene was performed using the primer conditions detailed in Appendix 9.5.1 using the standard PCR conditions outlined in Section 2.3.2. The details of the individual protocols used to sub-clone the *CCBL1* cDNA sequence into the pT7CFE1-CHis vector, introduce the c.814G>A mutation and carry out *in vitro* translation of the proteins are detailed in Section 2.16. Analysis of the proteins contained within each *in vitro* translation assay was carried out using label-free mass spectrometry (Section 2.17).

Quantitation of the enzyme activity of CCBL1 with respect to kynurenine was determined using the method described in Section 2.18. Endogenous KN and KA were also quantified in urine using this UPLC-MS/MS method (Section 2.19). The methods for the detection of tertiary amines in urine using both o-aminobenzaldehyde and ninhydrin are described in Section 2.20. Finally, the methods used to synthesise α-keto-ϵ-aminocaproate/Δ^1-piperideine-2-carboxylate from D-lysine using D-amino acid oxidase and to assess the ability of the overexpressed CCBL1 enzymes to catalyse the conversion of L-lysine to KAC/P2C are described in Section 2.21.

6.3 CASE REPORT

An 18 year old female (patient Y) presented to the metabolic clinic at GOSH. Her parents are first cousins of Pakistani origin and they have five other children. One of these children (the

older brother of the proband) also had difficulties with walking and speech. At 20 years of age he could only use 2 - 3 word sentences. However, examination did not reveal spastic paraparesis nor did he have any amino acid abnormalities. In addition, his disorder did not appear to be progressive in nature. There was also a family history of movement disorders with three cousins of the proband being affected by an inability to walk and speech problems, although they resided in India and more detailed clinical information was not available. Patient Y had movement problems beginning in infancy with walking compromised by marked scissoring of her legs at 2 years of age. This abnormal gait persisted until she fell and broke her ankle at 5 years of age, following which she could only walk with support. Her upper limb movements were abnormal with consequential functional impairment; she could feed herself with a spoon and hold a glass with two hands but could not wash her face or do up buttons. She had difficulty with speech and particularly word pronunciation throughout her life and tended to point to indicate her needs. She also had problems with poor sleep. Clinical examination at 18 years of age revealed profound muscle weakness, hypertonia, hyperreflexia, spastic paraparesis and speech difficulties. Electrophysiology indicated impaired upper and lower motor neuron dysfunction. She had dry, wrinkled skin without the typical flaking of ichthyosis. Brain MRI showed widespread cerebral and cerebellar atrophy with some high signal in the perotrigonal white matter; this was non-specific and compatible with a neurometabolic or neurodegenerative disorder. Biochemical investigations revealed significantly elevated levels of lysine in plasma, urine and CSF and she was subsequently diagnosed with hyperlysinaemia (Tables 6.3.1, 6.3.2 and 6.3.3). This finding was both striking and persistent, with all samples taken showing the same abnormality. Pyridoxine supplementation (20 mg bd) was started as elevated plasma and CSF threonine concentrations raised the possibility of a PLP deficiency; the conversion of lysine to α-keto-ϵ-aminocaproate/ Δ^1-piperideine-2-carboxylate is also believed to be PLP-dependent. This increased her sleep quality and reduced vomiting frequency but no improvement in motor function was observed. Haemoglobin was also noted to be low, associated with microcytic hyperchromic cells and low ferritin which was treated with iron supplementation. Unfortunately at 23 years of age she presented with severe cachexia as a result of progressive motor neuron disease-related weakness compounded by a refusal to eat which resulted in her death shortly afterwards.

Metabolic investigations which yielded normal results included fatty aldehyde NAD^+ oxidoreductase (excluding Sjörgren-Larsson syndrome), plasmalogen biosynthesis, peroxisomal fatty acid β-oxidation, localisation of peroxisomal enzymes, plasma very long chain fatty acids, urine glycosaminoglycans, urine organic acids, leukocyte ubiquinone, free and acylcarnitines, leukocyte lysosomal enzymes and urine α-AASA (excluding pyridoxine-dependent epilepsy due to antiquitin deficiency).

Table 6.3.1: Profile of plasma amino acids for patient Y. Samples were analysed by high-performance liquid chromatography (HPLC) as part of the patient's clinical care. Amino acid concentrations elevated above the reference range are shown in orange and those below the reference range are shown in blue. * Amino acids were not quantified at this time.

	Patient Y (µmol/L)					Reference range (µmol/L)
	17 years 1 month 16 days	17 years 8 months 30 days	17 years 11 months 8 days	18 years 2 months 14 days	18 years 11 months 7 days	
Glycine	261	250	244	297	268	100 – 330
Serine	138	131	98	135	134	90 – 290
Threonine	247	257	215	257	283	70 – 220
Proline	240	335	446	323	218	85 – 290
Leucine	94	97	65	80	81	65 – 220
Isoleucine	56	49	35	43	47	26 – 100
Valine	177	178	142	158	158	90 – 300
Alanine	411	320	344	415	260	150 – 450
Glutamine	839	824	607	840	847	480 – 800
Arginine	66	31	63	73	61	40 – 120
Ornithine	63	74	47	30	35	25 – 120
Lysine	624	504	433	524	606	100 – 300
Methionine	27	29	26	30	32	10 – 60
Taurine	93	72	60	51	73	40 – 140
Phenylalanine	54	52	47	65	53	35 – 100
Tyrosine	53	54	43	58	54	30 – 120
Tryptophan	53	44	21	27	42	30 – 80
Histidine	89	82	84	94	83	30 – 150
Asparagine	91	59	*	*	64	-
Aspartate	3	6	*	*	2	-
Glutamate	58	63	75	25	17	25 – 130

Table 6.3.2: Profile of urinary amino acids in patient Y. Samples were analysed by ion exchange chromatography with post column ninhydrin derivatisation using the Biochrom 30+ amino acid analyser as part of patients clinical care. Amino acid concentrations elevated above the reference range are shown in orange.

	Patient Y (µmol/mmol creatinine)		Reference range (µmol/mmol creatinine)
	17 years 1 month 16 days	17 years 11 months 8 days	
Glycine	920	2066	100 – 400
Serine	174	338	20 – 50
Threonine	147	318	10 – 25
Proline	2	68	0 – 3
Leucine	9	13	3 – 10
Isoleucine	2	4	1 – 10
Valine	7	13	2 – 8
Alanine	67	131	20 – 80
Glutamine	388	435	30 – 120
Arginine	9	28	2 – 10
Ornithine	3	26	2 – 10
Lysine	841	1746	5 – 30
Cystine	45	107	5 – 20
Methionine	3	27	3 – 10
Taurine	54	59	40 – 200
Phenylalanine	14	33	4 – 15
Tyrosine	19	46	5 – 20
Tryptophan	9	16	1 – 6
Histidine	282	479	100 – 300
Aspartate	26	30	10 – 30
Glutamate	8	6	0 - 3

Table 6.3.3: Profile of amino acids in the cerebrospinal fluid of patient Y. The sample was analysed by high-performance liquid chromatography (HPLC) as part of patients clinical care. Amino acid concentrations elevated above the reference range are shown in orange and those below the reference range are shown in blue.

	Patient Y (µmol/L)	Reference range (µmol/L)
	17 years 1 month 16 days	
Glutamate	4	5 – 17
Threonine	61	10 – 45
Serine	22	10 – 81
Glycine	8	4 – 14
Proline	< 1	0 – 1
Leucine	8	4 – 18
Isoleucine	3	3 – 8
Valine	13	5 – 25
Alanine	26	16 – 36
Glutamine	564	420 – 600
Arginine	16	15 – 40
Ornithine	1	3 – 13
Lysine	176	10 – 32
Methionine	2	2 – 6
Taurine	4	3 – 10
Histidine	15	3 – 18
Tryptophan	3	1 – 3
Phenylalanine	11	5 – 15
Tyrosine	9	5 - 15

6.4 RESULTS AND DISCUSSION

6.4.1 *Whole exome and confirmatory Sanger sequencing*

Whole exome sequencing was performed by Dr Olaf Bodamer, University of Miami, USA. The data was analysed to look for variants that fitted an autosomal recessive pattern of inheritance in the affected family. Twenty eight missense variants were identified, many of which were in genes that encode proteins of unknown function or that have not yet been associated with disease. *MROH9*, *EGFLAM*, *AKNA*, *ADAMTS12*, *MYO1H*, *BLZF1* and *AMOTL2* are protein-coding genes whose physiological role is unknown. *OR6A2* and *OR1L6* encode olfactory receptors, in which mutations have not been associated with disease. The protein encoded by *PSMD9* forms part of the 26S proteasome complex which functions to cleave peptides in an ATP/ubiquitin-dependent pathway (Watanabe *et al.*, 1998). *NDUFA8* encodes a component of mitochondrial respiratory chain complex I (Szklarczyk *et al.*, 2011). *KDM2B* encodes a histone demethylase and *ITGA1* encodes the alpha 1 subunit of integrin receptors, forming a cell-surface receptor for collagen and lamin (Andricovich *et al.*, 2016; Lee *et al.*, 2007). The protein encoded by *ASTN2* is expressed in the brain and is thought to function in neuronal migration based on mouse studies

(Wilson *et al.*, 2010). *DAB2* encodes a clathrin adaptor protein and is key component in the trafficking of the cystic fibrosis transmembrane conductance regulator (Fu *et al.*, 2012). *BARX1* encodes a transcription factor with many diverse functions including stomach smooth muscle development and intestinal rotation (Jayewickreme and Shivdasani, 2015). *TRPV2* encodes an ion channel which acts as a mechanosensor, thermosensor and lipid sensor. These diverse function enable roles in axon growth, thermosensation and gastrointestinal transit (Shibasaki, 2016).

Table 6.4.1: Whole exome sequencing data was filtered to show autosomal recessive variants in patient Y. All variants were homozygous in patient Y. Chr, chromosome; -, not present in database.

Chr	Position	Reference Allele	Sample Allele	Gene	Protein Variant	SIFT	PolyPhen	dbSNP ID	NHLBI ESP Frequency	ExAC Frequency
1	169347686	G	A	*BLZF1*	p.Arg196Gln	Tolerated	Benign	1064274	2.10	0.85
1	170967395	G	A	*MROH9*	p.Val526Ile	Deleterious	Benign	151291051	-	0.02
3	134078284	G	A	*AMOTL2*	p.Arg649Cys	Deleterious	Probably damaging	140029062	0.10	0.12
5	33881252	G	A	*ADAMTS12*	p.Thr154Met	Benign	Probably damaging	117518215	3.30	2.19
5	38370532	G	A	*EGFLAM*	p.Arg227His	Tolerated	Benign	199795131	1.10	0.62
5	39382748	T	C	*DAB2*	p.Lys417Arg	Tolerated	Benign	-	-	-
5	52097489	A	C	*ITGA1*	p.Thr325Pro	Deleterious	Benign	-	-	-
6	7583703	G	A	*DSP*	p.Asp1471Asn	Tolerated	Possibly damaging	41302885	-	-
7	158380295	C	A	*PTPRN2*	p.Ala23Ser	Tolerated	Benign	-	-	0.60
8	102631911	G	A	*GRHL2*	p.Val415Ile	Benign	Probably damaging	3779617	2.00	1.29
8	29053708	G	T	*KIF13B*	p.Pro53Gln	Deleterious	Probably damaging	200573525	0.20	0.43
9	96715374	C	A	*BARX1*	p.Gly107Trp	Tolerated	Benign	-	-	0.29
9	117124785	G	A	*AKNA*	p.Ala606Val	Tolerated	Benign	41278657	0.70	2.18
9	117849138	C	T	*TNC*	p.Arg291His	Tolerated	Possibly damaging	141281085	0.12	0.0008

9	119448979	C	T	*ASTN2*	p.Glu36Lys	Tolerated	Benign	10983304	1.20	-
9	124910416	C	G	*NDUFA8*	p.Arg119Pro	Tolerated	Possibly damaging	-	-	0.0008
9	125512943	C	T	*OR1L6*	p.Arg273Trp	Deleterious	Benign	117703463	0.90	1.43
9	126132826	G	T	*CRB2*	p.Trp498Cys	Tolerated	Possibly damaging	144803819	0.04	0.08
9	131598099	C	T	*CCBL1*	p.Val272Met	Deleterious	Probably damaging	-	-	0.003
11	6816792	T	C	*OR6A2*	p.Ile50Val	Tolerated	Benign	61741824	1.70	3.02
12	109839021	C	T	*MYO1H*	p.Arg216Cys	Deleterious	Probably damaging	146988917	-	0.02
12	121416864	C	T	*HNF1A*	p.Ala98Val	Tolerated	Benign	1800574	2.00	3.67
12	121880564	C	T	*KDM2B*	p.Glu825Lys	Tolerated	Benign	201943632	0.01	-
12	122353798	A	G	*PSMD9*	p.Lys93Glu	Deleterious	Probably damaging	-	-	-
16	53682949	C	T	*RPGRIP1L*	p.Arg744Gln	Deleterious	Benign	2302677	6.20	4.63
17	16335357	A	G	*TRPV2*	p.Met578Val	Deleterious	Possibly damaging	-	-	0.29
17	16843729	G	A	*TNFRSF13B*	p.Ala181Val	Tolerated	Benign	72553883	0.20	0.004
19	1469469	C	-	*APC2*	p.Ala2057Thr	Tolerated	Benign	189440287	4.80	2.47

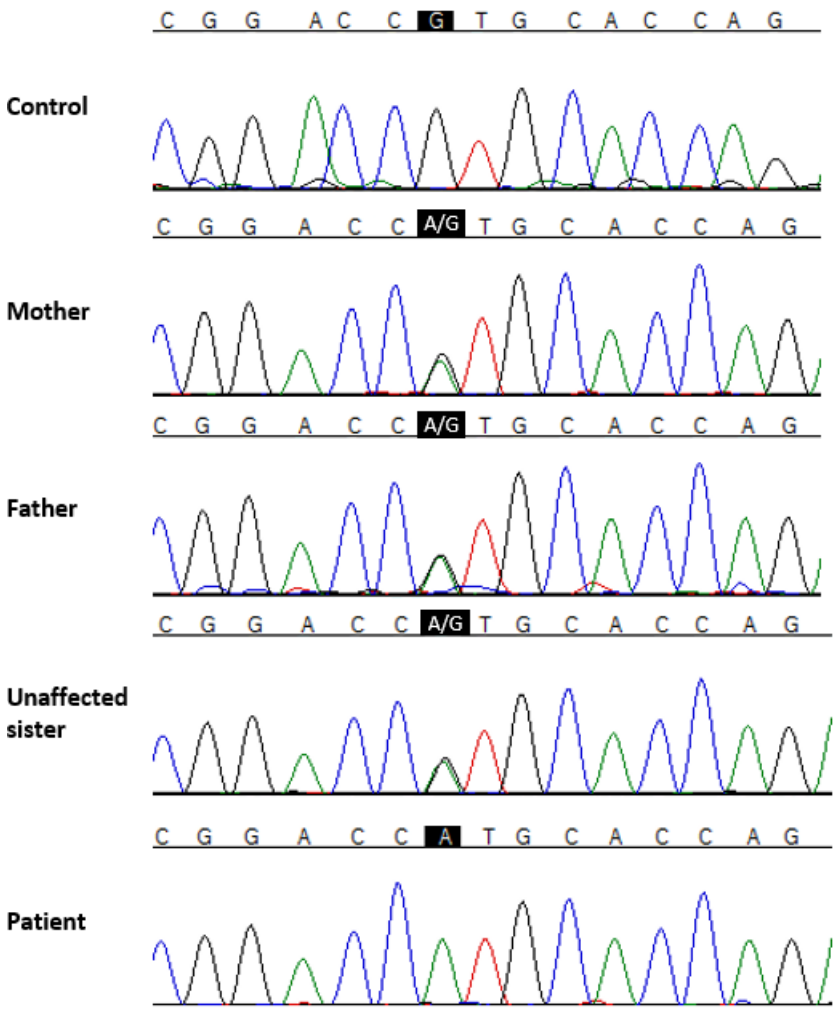
The remaining variants were identified in genes that have been associated with human disease, but the respective disorders were not consistent with the phenotypic features seen in patient Y. Mutations in *GRHL2* cause sensorineural deafness autosomal dominant type 28 and ectodermal dysplasia/short stature syndrome (Vona *et al.*, 2013; Petrof *et al.*, 2014). Mutations in *TNC* also cause sensorineural deafness autosomal dominant type 56 (Zhao *et al.*, 2013). A genome-wide associated study identified an association at a locus corresponding to the *KIF13B* gene with corticobasal degeneration. This disorder shares similarities with Parkinson's disease with typical onset between 50 - 70 years of age, movement and cognitive decline (Kouri *et al.*, 2015). Mutations in the *DSP* gene cause a spectrum of disorders presenting with cardiac and/or cutaneous disease ranging from lethal epidermolysis bullosa to cardiomyopathy (Boyden *et al.*, 2016). Mutations in *HNF1A* cause maturity-onset diabetes of the young type 3 (Bellanné-Chantelot *et al.*, 2008). Similarly, antibodies generated against *PTPRN2* have been identified in patients with type I diabetes (Hoppu *et al.*, 2006). Mutations in *CRB2* cause two different diseases presenting solely with renal abnormalities: focal segmental glomerulosclerosis 9 and ventriculomegaly with cystic kidney disease (Lamont *et al.*, 2016). Mutations in *RPGRIP1L* can also cause multiple different disorders including COACH syndrome (cerebellar hypo/aplasia, mental retardation, ataxia, ocular coloboma and hepatic fibrosis), Joubert syndrome 7 (molar tooth sign on brain MRI, oculomotor apraxia, ptosis, nystagmus, cerebellar ataxia and nephronophthisis) and Meckel syndrome 5 (anencephaly, occipital encephalocele, postaxial polydactyly, cleft lip and palate, microphthalmia, severe cystic kidney disease and bowing of the long bones) (Delous *et al.*, 2007). Mutations in *APC2* causes Sotos syndrome, a disorder characterised by a language disorder, learning difficulties, hyperactivity and macrocephaly (Almuriekhi *et al.*, 2015). Finally, mutations in *TNFRSF13B* cause common variable immunodeficiency 2 and immunoglobulin A deficiency 2. These disorders are characterised by recurrent infections, splenomegaly and an increased risk of malignancies.

Of the variants identified, only one offered a potential explanation for the hyperlysinaemia phenotype. This was a missense variant (c.814G>A; p.Val272Met) in the *CCBL1* gene encoding cysteine conjugate β-lyase. p.Val272Met is predicted to be deleterious and probably damaging by SIFT and PolyPhen-2, respectively. Futhermore, the mutated Val272 residue is conserved from humans to fruit flies (Figure 6.4.1). Segregation of this variant was confirmed by Sanger sequencing (Figure 6.4.2).

Figure 6.4.1: Multiple sequence alignment of the CCBL1 protein across species. (*) positions with have a single fully conserved residue, (:) conservation between amino acids with strongly similar properties, (.) conservation between amino acids with weakly similar properties. The alignment was generated using Clustal Omega.

```
Human          KVGWVLGPDHIMKHLRTVHQNSVFHCPTQSQAAVAESFEREQLLFRQPSS
Chimpanzee     KVGWVLGPDHIMKHLRTVHQNSVFHCPTQSQAAVAESFEREQLLFRQPSS
Pig            KVGWVLGPDRLLRHLRTVHQNSVFHCPTQAQAAVAQSFEREQLHFGQPSS
Dog            KVGWVLGPDSLVKHLRTVHQNSIYHCPTQGQAAVAQSFQHEQLHFGQPSS
Chicken        KVGWTVGPNRLLQHLRTVHQNSVYHCATAAQDAVAKGFQRELEHYGKPDS
Mouse          KVGWVMGPDNIMKHLRTVHQNSIFHCPTQAQAAVAQCFEREQQHFGQPSS
Xenopus        KVGWALGPDHLLKHLRTVHQNSVYHCATAAQEAVAQGLQRELEQLGKPDS
Zebrafish      KVGWAIGSGHIMKHLKTVHQNSVYHCATAAQEAISVGFQREYDVFGTEDS
Fruit fly      KIGWAYGPAELIRNLQMVHQNSVYTCPTPLQEGVARSFEVELARLGQPES
               *:**. *   :::.*: *****:: * *  * .::  :: *       .*
```

Figure 6.4.2: Segregation of the p.Val272Met mutation in the *CCBL1* gene in the family of patient Y.



Following the identification of this variant, sequencing of the *CCBL1* gene was carried out for an unrelated patient presenting with similar phenotypic features including hyperlysinaemia, seizures and spastic parparesis. His detailed case history is described in (Yiannikas and Cordato,

1996). No potentially pathogenic variants were identified in any of the 14 exons of the *CCBL1* gene.

6.4.2 *Rationale underlying CCBL1 as a potential candidate*

CCBL1, alongside the three other enzymes known to exhibit kynurenine aminotransferase activity and which may also have lysine aminotransferase activity, are candidates for the enzyme which catalyses the first step in the pipecolate pathway of lysine metabolism in the brain. It has been known for nearly 40 years that the injection of L-lysine into the brain of rats results in the production of labelled L-pipecolate (Chang, 1978). L-lysine is first oxidised to form α-keto-ϵ-aminocaproate which cyclises to form Δ^1-piperideine-2-carboxylate (P2C); the enzyme responsible for this catalysis is unknown (Figure 6.1.1). *In vivo*, P2C is rapidly reduced to form L-pipecolate by ketimine reductases.

The removal of the α-amino group of L-lysine to form α-keto-ϵ-aminocaproate is likely to be catalysed by either an L-amino acid oxidase or an aminotransferase. Many L-amino acid oxidases have been examined for their ability to deaminate L-lysine. However, *in vitro* experiments have shown L-lysine to be either a poor or completely ineffective substrate of these enzymes (Murthy and Janardanasarma, 1999; Mason *et al.*, 2004; Wiemann *et al.*, 2005; Urban *et al.*, 1988). Despite aminotransferases tranditionally not being thought to act upon L-lysine, when isotope-labelled glutamate is incubated with brain slices from mice the isotope was detected within the α-amino group of L-lysine, suggesting this is not the case (Papes *et al.*, 2001). The likely explanation for this finding is that L-lysine and α-keto-ϵ-aminocaproate can undergo reversible transamination using glutamate as a carrier of the amine group and being made possible by the high concentrations of glutamate within the brain. One family of enzymes which are known to exhibit broad substrate and α-keto acid specificities are the kynurenine aminotransferases. Given their catalysis of many reactions, they are known by multiple synonyms (Table 6.4.2).

The kynurenine aminotransferases are already of interest to neurologists given their ability to catalyse the transamination of kynurenine to form kynurenic acid. Kynurenic acid is not only the sole known endogenous antagonist of the N-methyl-D-aspartate subclass of glutamate receptors, but also is an antagonist of the α7-nicotinic acetylcholine receptor (Han *et al.*, 2009a). Indeed, abnormal concentrations of kynurenic acid have been documented in many neurodegenerative disorders including Alzheimer's disease, Huntington's disease, schizophrenia and dementia (Beal *et al.*, 1990; Widner *et al.*, 2000; Erhardt *et al.*, 2007; Guillemin *et al.*, 2005). In addition, the β-lyase activity of CCBL1 can mediate the toxicity of sulphur-containing fragments released from halogenated alkene-derived cysteine S-conjugates. Indeed, the cysteine S-conjugate of dopamine

Table 6.4.2: Synonyms of members of the kynurenine aminotransferase enzyme family. The name by which they are referred to in this chapter is highlighted in bold.

Gene	Enzyme names
CCBL1	**Cysteine conjugate β-lyase isozyme 1 (CCBL1)** Glutamine transaminase K (GTK) Kynurenine aminotransferase I (KAT I)
CCBL2	**Cysteine conjugate β-lyase isozyme 2 (CCBL2)** Glutamine transaminase L (GTL) Kynurenine aminotransferase III (KAT III)
AADAT	**α-aminoadipate aminotransferase (AADT)** Kynurenine aminotransferase II (KAT II)
GOT2	**Mitochondrial aspartate aminotransferase (AspAT)** Glutamic-oxaloacetic transaminase 2 Kynurenine aminotransferase IV (KAT IV)

has been demonstrated to be a substrate of CCBL1, which may play an important role in the pathogenesis of Parkinson's disease (Cooper *et al.*, 2008).

In vitro experiments have been carried out to investigate the ability of recombinant human kynurenine aminotransferases to catalyse the transamination of L-lysine (Han *et al.*, 2009a; Hallen *et al.*, 2013). In alkaline conditions (pH 9.2) L-lysine was found to be 5% as effective as L-glutamine (one of the substrates exhibiting the highest activity) as a substrate of CCBL2. The activity of CCBL1 under the same conditions was not measured. However, under more physiological conditions (pH 7.4) the activity of CCBL1 and CCBL2 with respect to L-lysine is less than 1%. The higher activity at more alkaline pH's may be explained by a mechanism similar to that observed in the pathogenesis of pyridoxine-dependent epilepsy (Section 4.2.1) and hyperprolinaemia type II. In these disorders, P6C or P5C complex with pyridoxal 5'-phosphate (PLP) through a Knoevenagel condensation thereby inactivating PLP-dependent enzymes. An analogous inactivation would be expected to occur upon the production of P2C, resulting in product inhibition of the PLP-dependent aminotransferase. Imines such as P2C, P5C and P6C exist in equilibrium with their enamine tautomers (Lu and Lewin, 1998). At higher pH's the equilibrium shifts to favour the enamine structure which is less reactive with PLP as it is less nucleophilic. In summary, it is possible that α-keto-ϵ-aminocaproate in the brain is produced by the action of a kynurenine aminotransferase. Although both CCBL1 and CCBL2 are present in the mammalian brain (Han *et al.*, 2009a), it remains unclear whether these enzymes can catalyse the transamination of L-lysine at physiologically relevant pH values. However, the identification of a homozygous variant in *CCBL1* in patient Y, whose phenotype included persistent hyperlysinaemia as a prominent feature, indicates that CCBL1 may function to catalyse this enzymatic reaction.

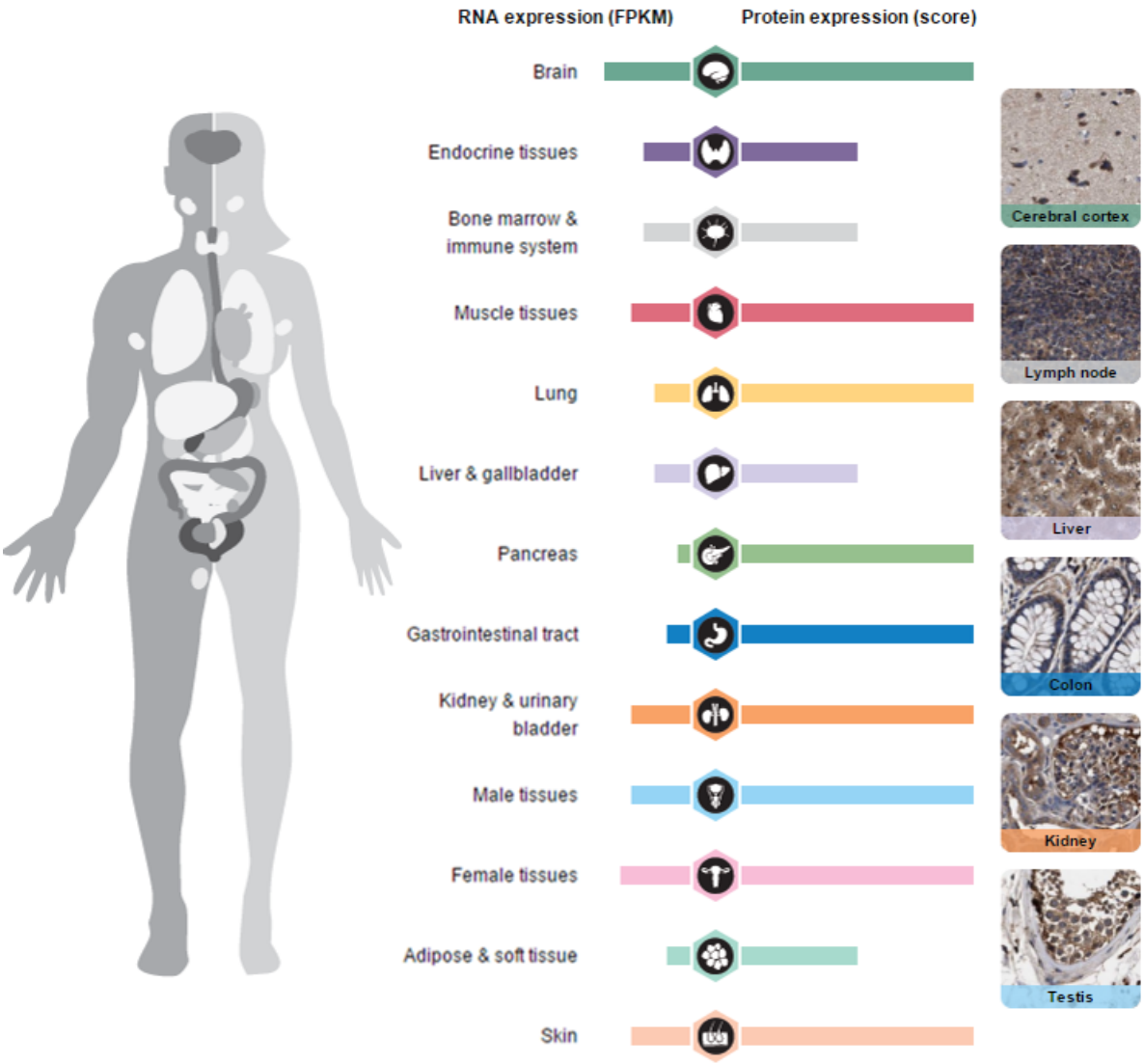
6.4.3 *Molecular cloning of CCBL1*

In order to determine whether or not the identified variants in *CCBL1* were pathogenic in patient Y, functional studies to examine the effect on enzymatic activity were required. Patient fibroblasts were available; however, these cells metabolise L-lysine via the saccharopine pathway (Struys and Jakobs, 2010), therefore examining the ability of endogenous CCBL1 to catalyse the conversion of L-lysine to α-keto-ϵ-aminocaproate/P2C in patient cells was not possible. An additional factor complicating the characterisation of CCBL1 *in vivo* is the fact that there are at least four distinct enzymes which can catalyse transamination of kynurenine (Table 6.4.2). Although CCBL1 is ubiquitously expressed (Figure 6.4.3), the other kynurenine aminotransferases are also highly expressed in the skin making the contributions of each enzyme activity difficult to discriminate.

Therefore in order to determine whether the p.Val272Met variant has any functional impact on the CCBL1 protein, the cDNA sequence of *CCBL1* was cloned in order to overexpress the protein. Firstly, the cDNA clone corresponding to the NM_004059 transcript in the pME18SFL3 vector (Section 2.16.1) was transformed into XL-1 Blue cells (Section 2.16.3.1) and sequenced to confirm the sequence was that of the wild-type gene (Section 2.16.6). NdeI restriction enzyme sites were then engineered at the 5' and 3' ends of the transcript using PCR without disrupting the open reading frame (Section 2.16.7). The resultant amplification products were sub-cloned into the TOPO 2.1 vector (Section 2.16.8), transformed into One Shot TOP10 Chemically Competent *E. coli* (Section 2.16.3.2) and correct insertion was validated by restriction enzyme digestion with EcoRI. Digestion with EcoRI would be expected to result in the excision of the *CCBL1* cDNA fragment (1281 bp) from the TOPO 2.1 vector (3956 bp). Indeed all, with the exception of clone 17, revealed that the correct cDNA fragment had been inserted (Figure 6.4.4a).

Subsequently, site-directed mutagenesis was used to introduce the c.814G>A variant into the wild-type *CCBL1* cDNA sequence cloned into the TOPO 2.1 vector and the resultant site-direct mutagenesis product was transformed into TOP10 cells (Section 2.16.9). Sanger sequencing confirmed the presence of the desired mutation (Figure 6.4.4b). Prior to cloning into the multiple cloning site of the pT7CFE1-CHis vector, which is suggested for use for optimal protein expression with the HeLa cell lysate-based 1-step human *in vitro* protein translation (IVT) kit, wild-type and mutant *CCBL1* cDNA in the TOPO 2.1 vector were digested with NdeI to facilitate cloning (Section 2.16.8). The empty pT7CFE1-CHis vector was also digested using NdeI and dephosphorylated to prevent re-ligation (Section 2.16.10). The products were run on an agarose gel, the linear pT7CFE1-CHis vector and the gel fragments containing the CCBL1 cDNA inserts were excised and the DNA was extracted from the gel (Section 2.16.11 and Figure

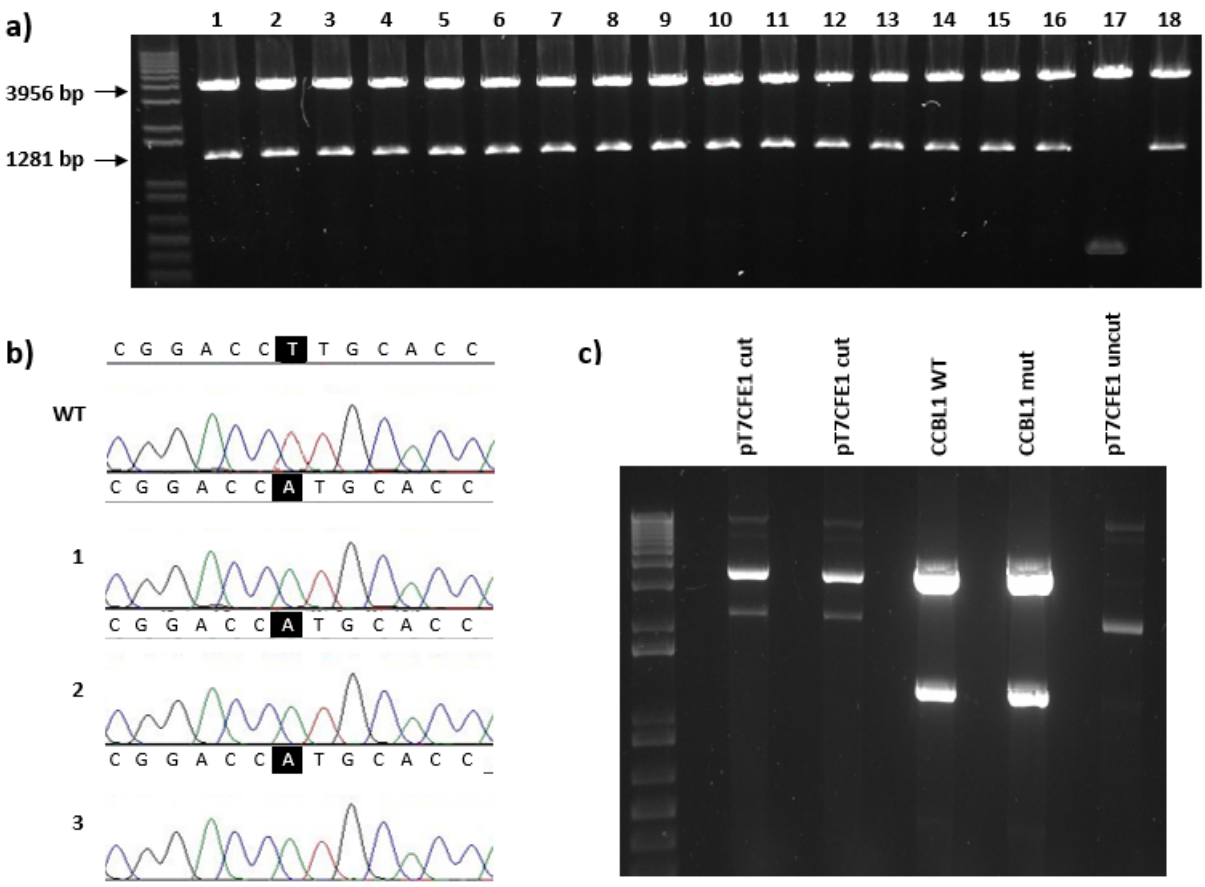
Figure 6.4.3: Profile of protein and RNA expression of CCBL1 in human tissues. Data was extracted from the Human Protein Atlas (www.proteinatlas.org). The size of each coloured bar indicates the degree of expression in the respective tissue. FPKM; Fragments Per Kilobase of transcript per Million mapped reads.



6.4.4c). The purified DNA was analysed on an agarose gel and the concentration was quantified prior to ligation into pT7CFE1-CHis and transformation into XL-1 Blue cells (Section 2.16.12).

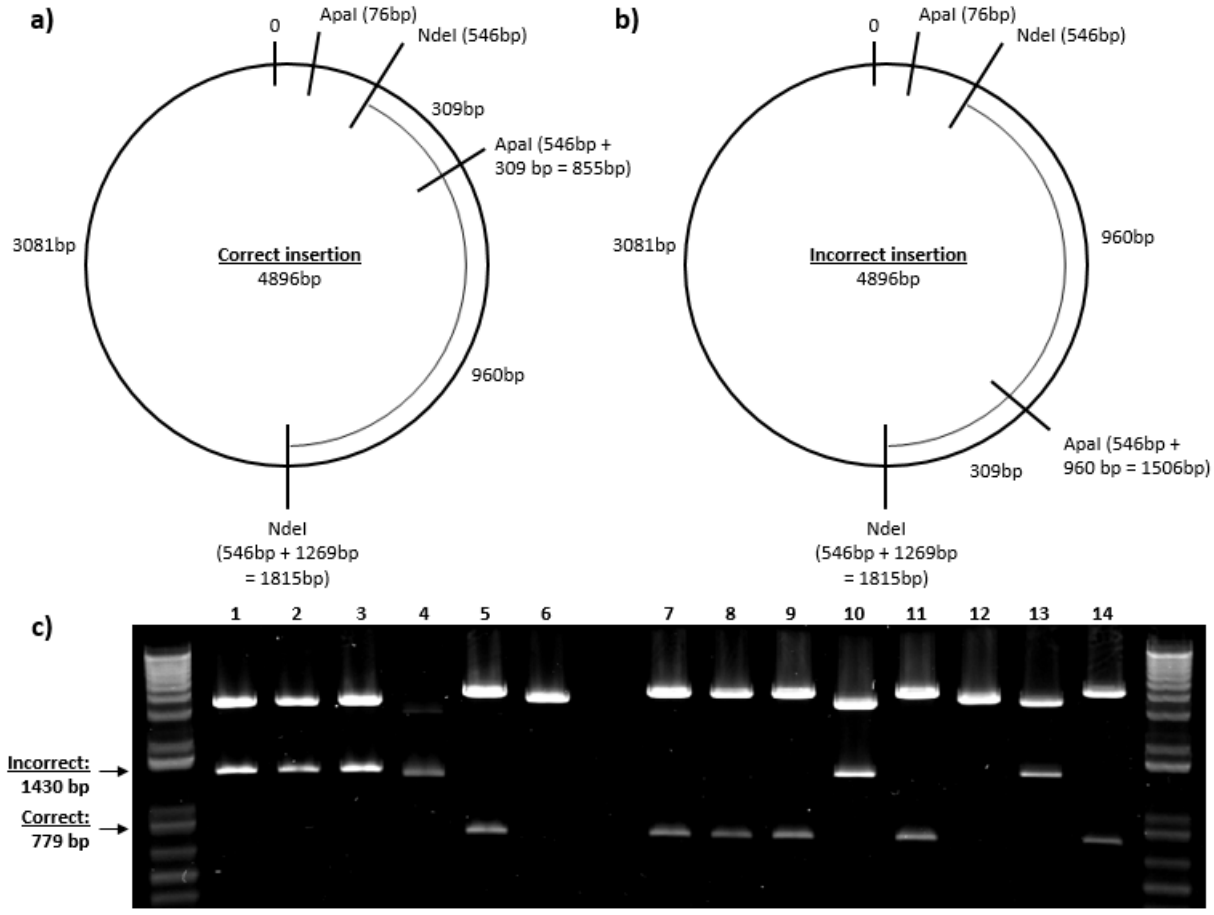
The direction in which the CCBL1 cDNA had been inserted into the vector was then analysed by restriction enzyme digestion with ApaI (Section 2.16.8). The *CCBL1* cDNA insert sequencing contains one ApaI restriction enzyme site 309 bp from the cDNA start codon. The pT7CFE1 vector also contains one ApaI restriction enzyme site 76 bp from the T7 promoter region. Therefore, if the cDNA sequence was cloned in the forward direction, a 779 bp fragment (309 bp [ApaI site in *CCBL1* insert from cDNA start codon] + 546 bp [NdeI site where cDNA sequence has been inserted from the T7 promoter] - 76 bp [Apa I site in pT7CFE1 vector from

Figure 6.4.4: Molecular cloning of the CCBL1 cDNA. (a) Digestion of TOPO 2.1 clones with EcoRI to check for the presence of the CCBL1 cDNA insert. **(b)** Sanger sequencing confirming the introduction of the c.814G>A mutation by site-directed mutagenesis in all three mutated clones. WT, wild-type. **(c)** Analysis of the pT7CFE1-CHis vector and the TOPO 2.1 plasmids containing the CCBL1 inserts digested using NdeI (left). The linear pT7CFE1-CHis and both the wild-type and mutated CCBL1 cDNA fragments were excised from the agarose gel (right).



T7 promoter]) would be identified upon digestion with ApaI. This analysis revealed one plasmid containing the wild-type and five plasmids containing the mutant cDNA sequence cloned in the forward orientation (Figure 6.4.5). Each insert of clones 5 (wild-type) and 7, 8, 9, 11 and 14 (mutant) were determined to be in-frame with the initiator methionine residue by Sanger sequencing and these were used for subsequent studies.

Figure 6.4.5: Validating the orientation of the CCBL1 cDNA insert within the pT7CFE1-CHis vector. Schematics illustrating the restriction enzyme sites and the predicted sizes of the DNA fragments that would result for the **(a)** correct and **(b)** incorrect orientation. **(c)** Agarose gel electrophoresis showing the results of restriction enzyme digestion of each clone with ApaI. Clone 5 (wild-type) and 7, 8, 9, 11 and 14 (mutant) revealed a fragment of 779 bp indicating the correct orientation of the insert.

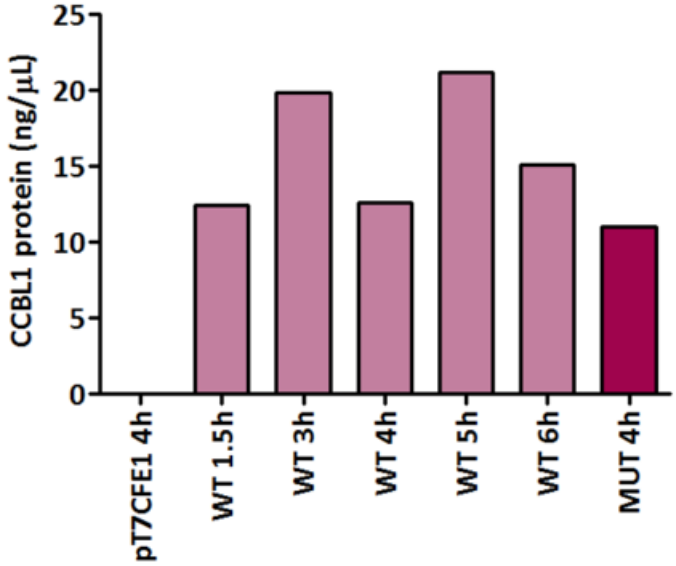


6.4.4 *In vitro protein translation and analysis of the products*

Prior to the expression of protein for enzyme activity quantitation, the relationship between the amount of CCBL1 protein produced and incubation time of the IVT reaction was investigated using label-free mass spectrometry (Section 2.17). This technique is an untargeted approach which can be used to identify and quantify proteins within a biological sample. This analysis showed that a five hour incubation resulted in maximal translation of the CCBL1 protein, producing a total of 21.2 ng/μl of CCBL1 protein (Figure 6.4.6). Thereafter, all experiments were performed using this parameter. Reactions incubated with only the empty pT7CFE1-CHis vector did not identify any CCBL1 protein, indicating it is not a component of the accessory proteins or HeLa cell lysate that is used during the IVT reaction. Moreover comparison of the reaction containing

the vector with the mutant sequence with that of the vector containing wild-type protein, did not show any reduction in the concentration of CCBL1 produced, suggesting that p.Val272Met does not result in mRNA decay.

Figure 6.4.6: Effect of incubation time on the amount of CCBL1 produced by the *in vitro* translation system. WT, pT7CFE1-CHis vector containing the wild-type CCBL1 sequence; MUT, pT7CFE1-CHis vector containing the mutant (p.Val272Met) CCBL1 sequence; pT7CFE1, empty pT7CFE1-CHis vector not containing any inserted sequence.



Concurrently to the quantitation of CCBL1, 898 proteins were identified within the IVT reaction mix (Appendix 9.5). Analysis of this list of proteins using the WEB-based Gene Set Analysis Toolkit (WebGestalt), revealed Gene Ontology terms that were over-represented when compared to the reference genome (i.e. if <1% of the proteins encoded by the human genome play a role in RNA binding, but 10% of the proteins identified in our sample were associated with this process, then the corresponding Gene Ontology term would be considered over-represented) (Zhang *et al.*, 2005) (Table 6.4.3). Unsurprisingly, these included genes involved in the initiation, elongation and termination of protein translation, mRNA generation and the binding of proteins, RNA, small molecules, nucleosides, nucleoside phosphates and nucleotides. Importantly, none of the other kynurenine aminotransferases nor enzymes catalysing steps of the saccharopine or pipecolate pathways of L-lysine metabolism were present within the IVT reaction mix.

Table 6.4.3: Pathway analysis of *in vitro* reaction mix. Solutions were analysed by label-free mass spectrometry and showthe number of identified proteins corresponding to each Gene Ontology term.

Gene Ontology terms	Number of proteins identified
Biological process and molecular function	
Establishment of protein localisation to endoplasmic reticulum	46
Cellular macromolecule (catabolic process)	109
mRNA (metabolic process)	116
mRNA (catabolic process)	59
Nuclear-transcribed mRNA (catabolic process)	57
Nonsense-mediated decay	52
Translational initiation	66
Translational elongation	52
Translational termination	45
Small molecule binding	205
Protein binding	389
Organic cyclic compound binding	315
Heterocyclic compound binding	314
Nucleoside binding	156
Nucleoside phosphate binding	196
RNA binding	114
Nucleotide binding	195
Purine nucleoside triphosphate binding	153
Structural constituent of ribosome	45
Cellular component	
Macromolecular complex	284
Intracellular	542
Cytoplasmic	495
Non-membrane-bounded organelle	272

6.4.5 *Quantitation of kynurenine aminotransferase activity*

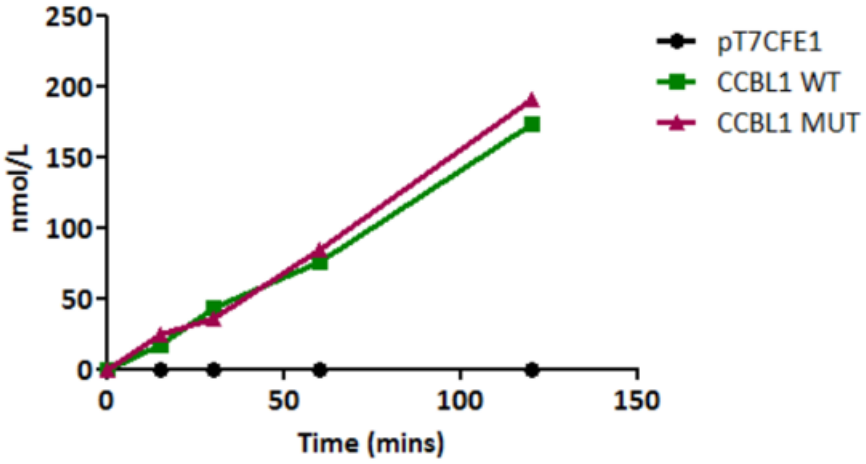
Prior to examining whether or not CCBL1 can catalyse the first enzymatic step in the pipecolate pathway of lysine metabolism, the kynurenine aminotransferase activity of both the wild-type and mutant CCBL1 was determined in order to determine whether synthesis using the IVT kit produced active enzymes (Section 2.18). The UPLC-MS/MS method used was adapted from that of Han *et al.* (2004) which uses high-performance liquid chromatography coupled to an ultraviolet detector (HPLC-UV). A summary of the differences between the two methods can be found in Table 6.4.4.

Table 6.4.4: Comparison of the method published by Han *et al.* (2004) and our UPLC-MS/MS method used for the investigation of the kynurenine aminotransferase activity of CCBL1.

	Han *et al.* (2004)	UPLC-MS/MS	Reason for change
Reaction volume	50 μL	50 μL	-
Kynurenine concentration	5 mM	5 μM	Greater sensitivity of UPLC-MS/MS
α-keto acid substrate	α-ketobutyrate	Oxaloacetate	
α-keto acid concentration	2 mM	2 μM	Greater sensitivity of UPLC-MS/MS
PLP concentration	40 μM	40 nM	Greater sensitivity of UPLC-MS/MS
Potassium phosphate buffer (pH 7.5) concentration	200 mM	200 mM	-
Enzyme	2 μg recombinant human CCBL1	2.5 μL IVT reaction mix containing approximately 50 ng CCBL1 protein	Purpose of assay.
Reaction incubation	10 mins at 45°C	2 hours at 37°C	Ability to measure activity over extended time-course to detect subtle differences between enzymes. More physiological conditions.
Reaction termination	0.8 M formic acid	0.3 N TCA	Established for many assays within the laboratory including the B_6 vitamer analysis.
Analysis method	HPLC-UV	UPLC-MS/MS	Greater sensitivity of UPLC-MS/MS

CCBL1 proteins containing both the wild-type and mutant sequences were expressed and the ability of each enzyme to catalyse the conversion of kynurenine (KN) to kynurenic acid (KA) was assayed. Both wild-type and mutant CCBL1 enzymes were active and had similar activities, with the rate of KA production being 190.2 and 212.5 pmol/hr/μg protein for the wild-type and mutant protein, respectively (Figure 6.4.7). No conversion of KN to KA was evident when the reaction mix was incubated with only the pT7CFE1 vector containing no *CCBL1* cDNA insert, indicating that the translated CCBL1 protein was responsible for catalysis.

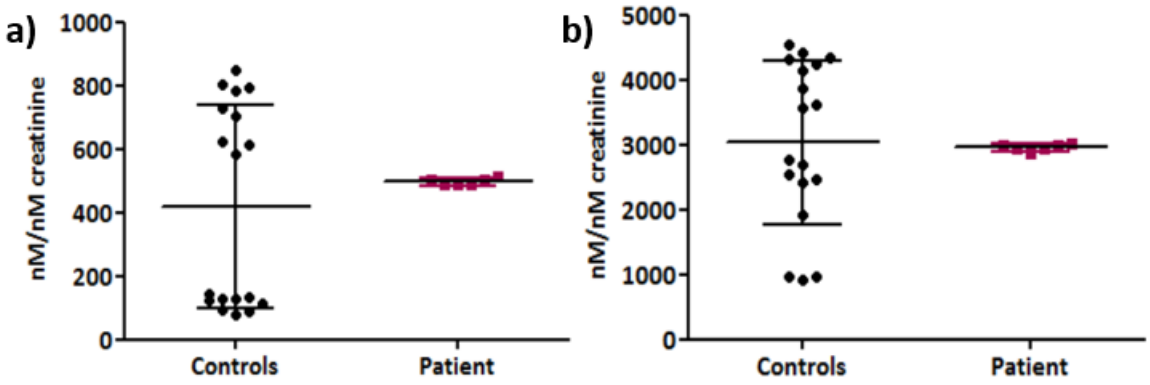
Figure 6.4.7: Assay of kynurenine aminotransferase activity of CCBL1 expressed using the *in vitro* translation system.



6.4.6 *Quantitation of kynurenine and kynurenic acid in patient urine*

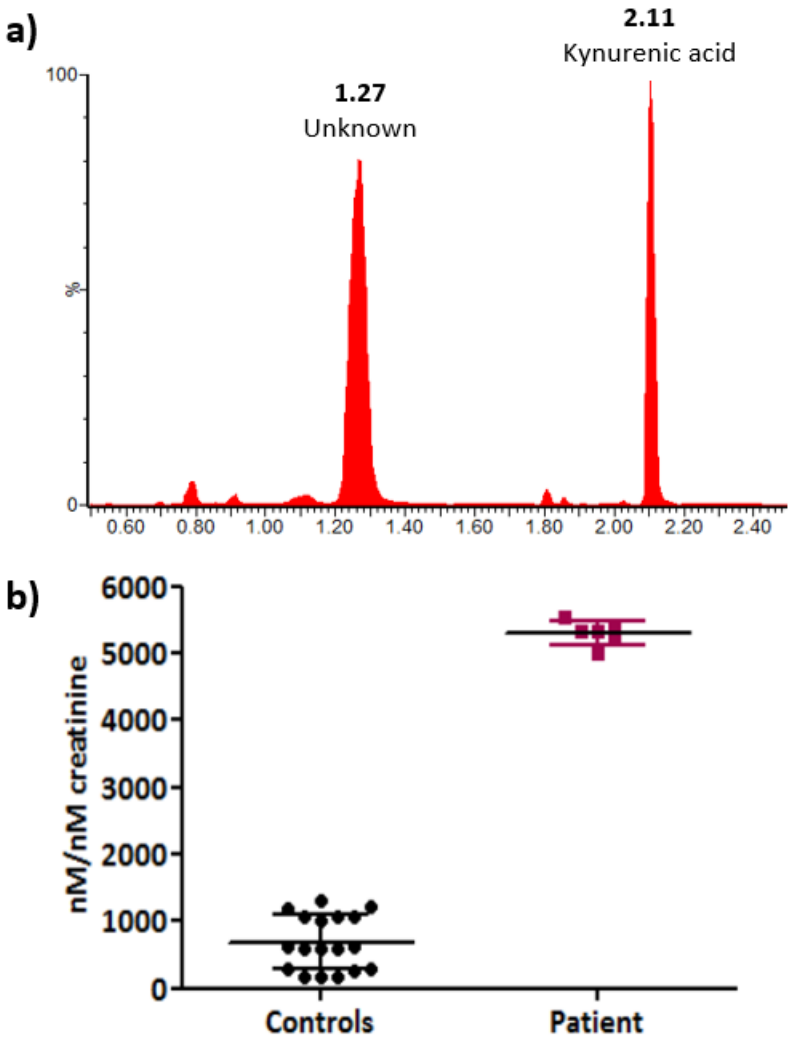
Endogenous KN and KA was also quantified in urine using this UPLC-MS/MS method (Section 2.19). The concentrations of KN and KA in the urine of patient Y were within the range of the six paediatric control samples analysed (Figure 6.4.8). Indeed, the urinary ranges and degree of variability of both KN and KA concentrations were comparable to those that have been reported using alternative methodologies (Zhao *et al.*, 2011; Marcos *et al.*, 2016).

Figure 6.4.8: Quantitation of endogenous kynurenine and kynurenic acid in urine. **(a)** Kynurenine and **(b)** kynurenic acid concentrations in the urine of patient Y and controls (n = 6). Two aliquots of patient urine was prepared and each sample was analysed in triplicate (three injections into the UPLC-MS/MS). Error bars represent ± standard deviation.



However, an additional peak in the KA chromatogram that had a different retention time to KA was identified in urine which was found to be present in larger amounts in the patients' sample (Figure 6.4.9).

Figure 6.4.9: Quantitation of additional peak in kynurenic acid channel in urine. (a) Chromatogram illustrating the additional peak observed in the kynurenic acid channel (190.20 > 144.00) in patient urine. **(b)** Quantitation of the additional peak in control (n = 6) and patient urine. Two aliquots of patient urine was prepared and each sample was analysed in triplicate (three injections into the UPLC-MS/MS). Error bars represent ± standard deviation.



Molecules can exist as structural isomers (which share the same molecular formula but differ in their arrangement) and sterioisomers (which have the same molecular formula and sequence of bonded atoms but differ in the three-dimensional orientation of the atoms in space). However, given that the second peak eluted 0.84 minutes prior to that of kynurenic acid, it is more likely that this peak represents an isobaric species (an unrelated compound with the same m/z ratio) rather than an isomer of KA (Figure 6.4.10a).

The Human Metabolome Database (HMDB) is a freely available resource containing experimental MS/MS data for over 800 compounds (Wishart *et al.*, 2007). Searching this extensive database for molecules that share the 190.20 > 144.00 transition, corresponding to the loss of a carboxylic acid group, identified homocitrulline, N-acetylglutamic acid and 3-indolepropionic

acid as potential candidates. The structure and chemical properties of each molecule can give an indication of their retention time. In the case of reverse phase HPLC, generally the retention time of a compound increases with decreasing water solubility and increasing numbers of carbon atoms (Hanai and of Chemistry , Great Britain). In addition, neutral and charged species typically show the least retention, followed by acidic and basic molecules (Law, 1990). These parameters for each candidate are shown in Table 6.4.5 (Wishart *et al.*, 2007). Given that KA and 3-indolepropionic acid share a similar bicyclic structure, water solubility, carbon content and pKa value, it is likely that their retention times would be similar. Conversely, homocitrulline and N-acetylglutamic acid have approximately 10- and 20-fold higher solubility in water than KA, respectively. This factor alone is likely to result in these compounds being less retained on the column and being plausible candidates for the identity of the isobaric species.

Table 6.4.5: Chemical characteristics of the possible compounds sharing the 190.20 > 144.00 transition. All parameters were extracted from the Human Metabolome Database.

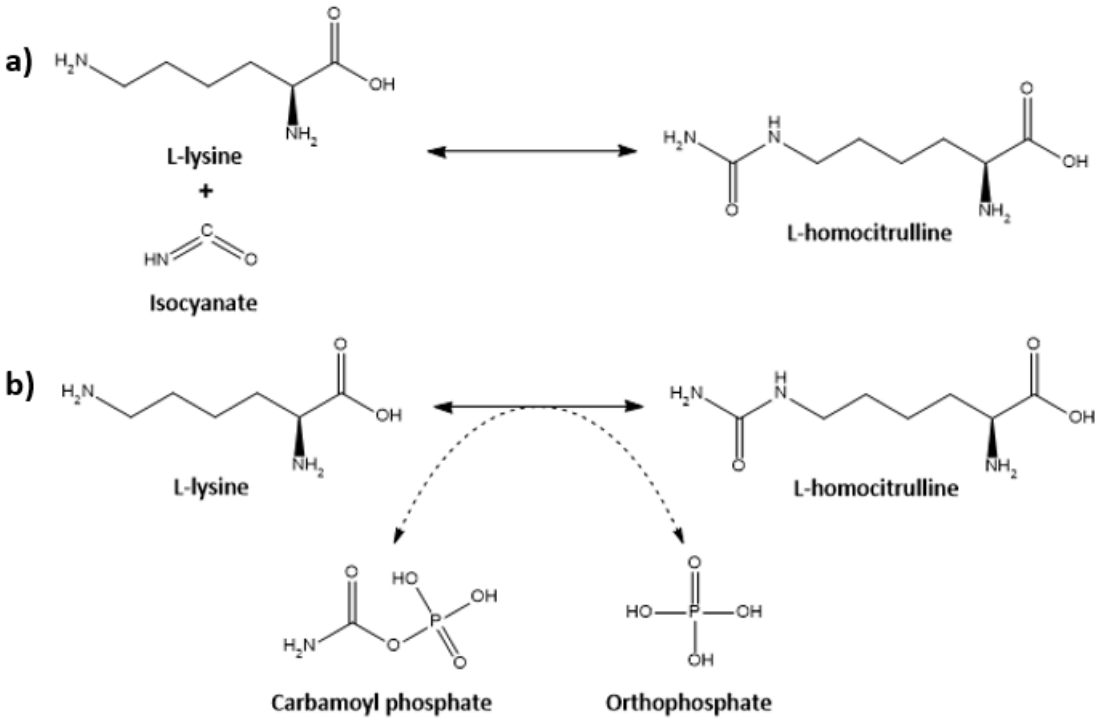
Compound	Water solubility (mg/mL)	Number of carbon atoms	Physiological charge	pKa (strongest acid)
Kynurenic acid	0.95	10	-1	3.17
3-indolepropionic acid	0.73	11	-1	4.8
N-acetylglutamic acid	18.6	7	-2	3.43
Homocitrulline	12.1	7	0	2.35

N-acetylgluatamic acid is synthesised from glutamate and acetyl-CoA by N-acetylglutamate synthase. In humans it is an essential allosteric cofactor for carbamoyl phosphate synthetase I, the first enzyme of the urea cycle (Caldovic and Tuchman, 2003). Indeed, N-acetylglutamic is typically undetectable in plasma and urine of healthy subjects regardless of age (Tavazzi *et al.*, 2005). In contrast, homocitrulline is readily quantified in urine with control ranges being between 0 - 10 μmol/mmol creatinine (Al-Dirbashi *et al.*, 2006; Korman *et al.*, 2004) and is elevated in patients with urea cycle disorders. Homocitrulline is an amino acid which is formed through the carbamoylation of lysine. This can occur spontaneously through a reaction with cyanate, which exists in equilibrium with urea in the human body (Figure 6.4.10b). Alternatively, it can be synthesised enzymatically by lysine carbamoyltransferase catalysing the reaction of lysine with carbamoyl phosphate (Figure 6.4.10c). Therefore, the synthesis of homocitrulline is enhanced when either lysine or carbamoyl phosphate is accumulated. Examples of this phenomenon are observed in hyperlysinaemia type I (Heiden *et al.*, 1978) and type II (Carson *et al.*, 1968), lysinuric protein intolerance (Habib *et al.*, 2013) and hyperornithinaemia, hyperammonaemia and homocitrullinuria (HHH) syndrome (Martinelli *et al.*, 2015). Patient Y did not have any symptoms of a urea cycle disorder including hyperammonaemia, episodes of metabolic decompensation or

protein intolerance. Therefore, it is possible that the hyperlysinaemia present in our patient may result in increased production and urinary excretion of homocitrulline which is observed as an isobaric species in the KA MS/MS channel.

Homocitrulline has recently been made available as a standard from Santa Cruz Biotechnology. In the future, this compound should be purchased and injected into the UPLC-MS/MS using the same chromatography conditions as described in Section 2.18 to determine its retention time. It is also worth noting that the urine sample from patient Y had been stored at -20°C for five years following the patient's death. Therefore, it is possible that the unknown compound may be present as a consequence of sample degradation and this should be considered in any further experiments.

Figure 6.4.10: Synthesis of homocitrulline. L-homocitrulline is generated through the carbamylation of L-lysine. This can occur **(b)** through the reaction with cyanate (which is equilibrium with urea under physiological conditions) or **(c)** reaction with carbamoyl phosphate.

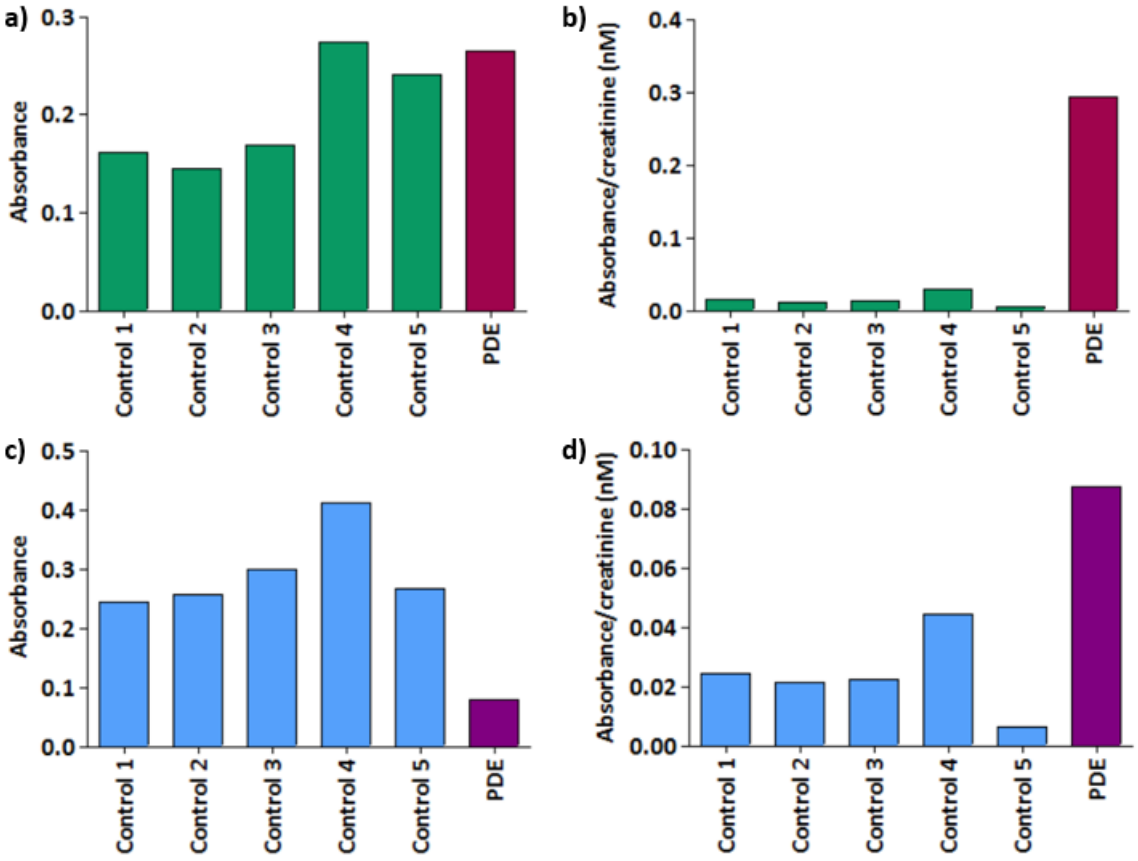


6.4.7 *Attempts to measure activity of CCBL1 towards lysine*

In order to assess the activity of the CCBL1 enzyme towards lysine, spectrophotometric methodologies were evaluated. Both o-aminobenzaldehyde (o-AB) (Namwat *et al.*, 2002) and ninhydrin (Kim *et al.*, 1994) are known to react with tertiary amines such as Δ^1-piperideine-2-carboxylate (P2C) to produce a coloured product. A positive control, i.e. a urine sample from a patient

with pyridoxine-dependent epilepsy where α-AASA and its equilibrium partner Δ^1-piperideine-6-carboxylate (P6C) accumulates, was used as P6C shares a similar structure to P2C. Both are tertiary amines and simply differ in the position of the double bond within the piperideine ring. Although methods were not available to determine the concentration of P6C in this urine sample, the concentration of α-AASA was quantified as part of the routine diagnostic service as 286.8 μmol/mmol creatinine (ref: < 4) (Yi Yang, ICH, UK). Five age-matched paediatric controls were also analysed. Both methods demonstrated an increased absorbance in the AASA-positive urine upon correction for urinary creatinine concentration, although derivatisation with o-AB gave greater discrimination between samples (Figure 6.4.11).

Figure 6.4.11: Colourmetric reactions of AASA-positive urine to detect tertiary amines. Absorbance of urine when reacted with o-aminobenzaldehyde **(a)** and corrected for creatinine concentration **(b)**. Absorbance of urine when reacted with ninhydrin **(c)** and corrected for creatinine concentration **(d)**. Aliquots of the same urine samples were used for both derivitisation reactions. PDE, pyridoxine-dependent epilepsy.



Accordingly, the same derivatisation method using o-AB was then used to assess the conversion of lysine to P2C by wild-type and mutant CCBL1. Four different α-keto acids, 2-oxobutyrate, 2-oxoglutarate, pyruvate and oxaloacetate were tested as the amine group acceptor. However, no significant production of a molecule causing a colourmetric reaction was observed using any of the assay conditions.

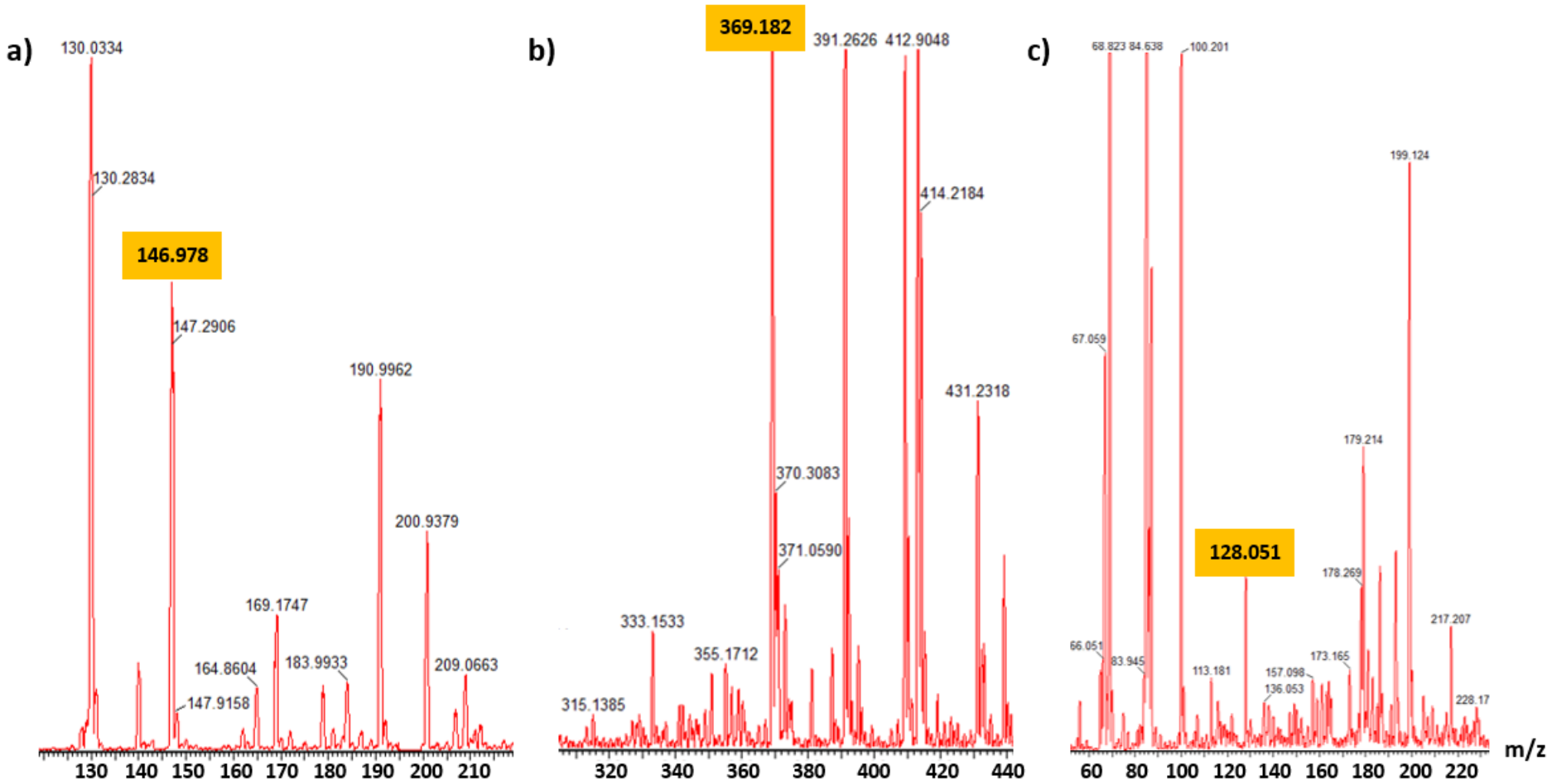
It was possible that no enzymatic activity was identified because the colourmetric assay did not have adequate sensitivity to detect the low concentrations of P2C being produced. Therefore, mass spectrometry-based assays were investigated. Neither α-keto-ϵ-aminocaproate (KAC), the product of the proposed reaction, nor its equilibrium partner Δ^1-piperideine-2-carboxylate (P2C) are commercially available for infusion and therefore determination of mass spectrometry parameters required for analysis, thus attempts were made to synthesise them enzymatically. Two publications report methods for the reaction of D-alanine and D-proline with D-amino acid oxidase (Konno, 1998; Visser *et al.*, 2012). These were adapted accordingly (Table 6.4.6). Each solution was desalted prior to direct infusion into the mass spectrometer (Section 2.21).

Table 6.4.6: Differences between published and novel methods for the reaction of D-amino acids with DAAO. Methods published by Konno (1998) and Visser *et al.* (2012) for the reaction of D-amino acids with D-amino acid oxidase and the adaptations made to these assays for the oxidation of D-lysine prior to direct infusion into the mass spectrometer. -, identical to original method; *, FAD was added to half of the reactions during the optimisation process; DAAO, D-amino acid oxidase.

	Konno (1998)	Method for MS/MS infusion	Visser *et al.* (2012)	Method for MS/MS infusion
Buffer	40 mM pyrophosphate (pH 8.3)	-	25 mM ammonium bicarbonate (pH 8.3)	-
Catalase	700 IU	-	900 IU	-
D-amino acid	D-alanine	D-lysine	D-proline	D-lysine
D-amino acid concentration	30 mM	-	1 mM	-
Flavin adenine dinucleotide	20 mM	-	None	20 mM*
70% methanol	100 μL (final 7% (v/v))	-	None	-
DAAO	0.1 mL homogenised tissue supernatant containing DAAO	10 units in 100 μL dH_2O	0.25 mg	-
Reaction incubation	1 hour at 37°C	2 hours at 37°C	1 hour at 37°C	2 hours at 37°C
Total reaction volume	1 mL	1 mL	100 μL	100 μL
Reaction termination	1 mL 10% TCA	1 mL 0.3 N TCA	700 μL ice cold acetonitrile	100 μL 0.3 N TCA

The mixtures were analysed in both positive and negative mode using a scan window between 50 - 800 m/z. L-lysine was commercially available and this compound was directly infused to generate transitions (Figure 6.4.12a). Using the method adapted from Konno (1998), no peak with an m/z corresponding to either KAC or P2C was seen. Derivatisation of the assay mixture with fluorenylmethyloxycarbonyl chloride (FMOC-Cl) was then performed to increase the sensitivity of analytes containing a primary or secondary amine group. Nevertheless, only an ion corresponding to the derivatised lysine was observed (m/z 369.18) (Figure 6.4.12b). Using the method adapted from Visser *et al.* (2012), an ion with a m/z of 128.10 which was a possible match for P2C was identified (Figure 6.4.12c).

Figure 6.4.12: D-amino acid oxidase reaction products. MS scan spectra illustrating the peaks resulting from the reaction of D-amino acid oxidase with D-lysine. Highlighted in orange are the m/z ratios corresponding to **(a)** D/L-lysine, **(b)** FMOC-lysine and **(c)** P2C (potentially).



This was subsequently tuned and the four transitions obtained are as detailed in Table 6.4.7. The chromatography and peak shape generated by each transition were evaluated using a range of mobile phases (with varying concentrations) including: dH_2O, formic acid, ammonium acetate, acetic acid, acetonitrile (with and without formic acid), methanol (with and without formic acid). A combination of 3.7% acetic acid and 100% acetonitrile was found to give a good and reproducible peak shape (Figure 6.4.13). The two transitions for each compound generating the highest sensitivity under these UPLC-MS/MS conditions were selected (Figure 6.4.13). The final assay parameters can be found in (Section 2.21).

Table 6.4.7: Transitions identified to detect lysine and Δ^1-piperideine-2-carboxylate (P2C). *, predicted compound.

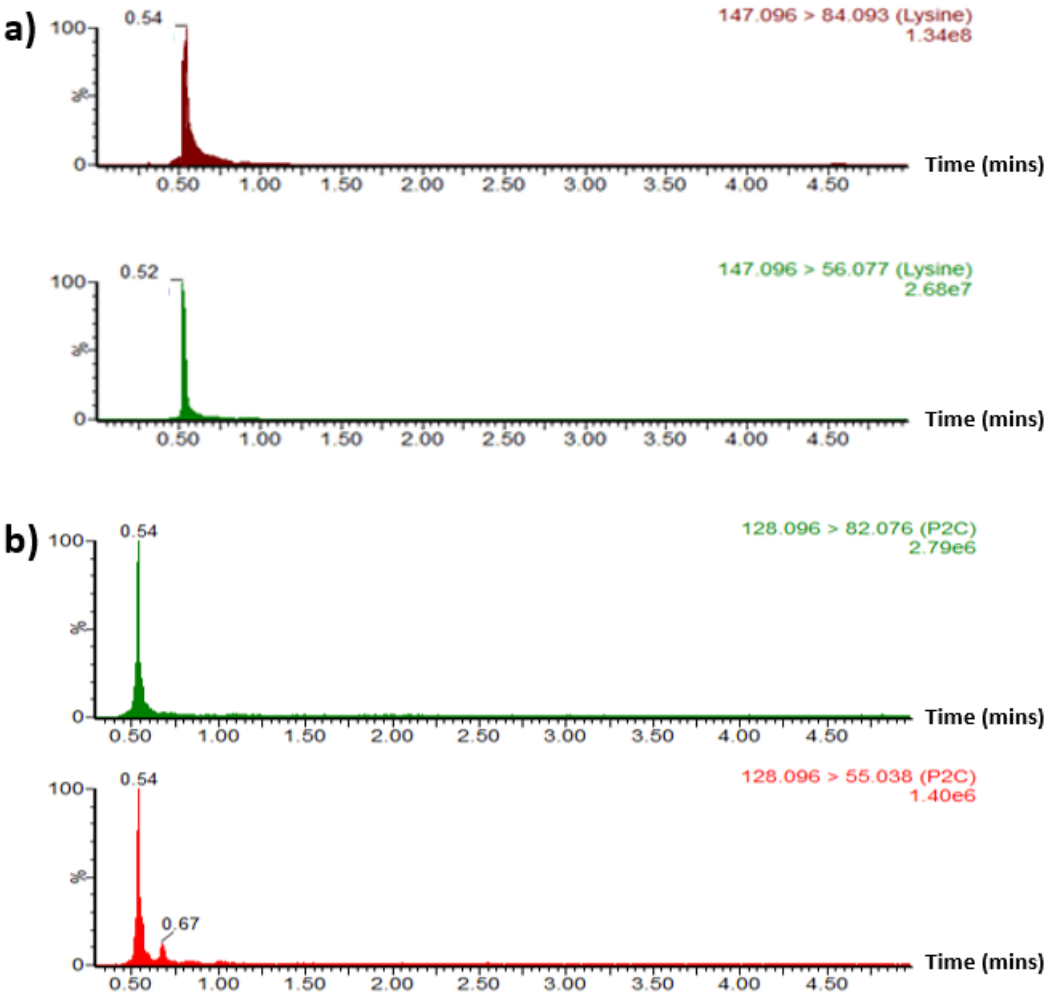
Analyte	Parent ion (m/z)	Daughter ion (m/z)	Cone voltage (V)	Collision voltage (V)	Retention time (mins)
Lysine	147.10	41.91	28	22	0.58
Lysine	147.10	56.08	28	24	0.58
Lysine	147.10	67.06	28	20	0.58
Lysine	147.10	74.02	28	20	0.58
Lysine	147.10	84.09	28	14	0.58
P2C*	128.10	55.04	40	20	0.54
P2C*	128.10	82.08	40	14	0.54
P2C*	128.10	86.96	40	6	0.54
P2C*	128.10	100.08	40	14	0.54

Investigation of the possible conversion of L-lysine to KAC/P2C was carried out under varying conditions. Firstly, an enzyme assay identical to that performed for the assessment of CCBL1 activity towards kynurenine was carried out using lysine as a substrate (Section 2.18). Secondly, an enzyme assay based on the method described in Hallen *et al.* (2013) which suggested the use of an alternative buffer was performed (Section 2.21). No reduction in lysine concentration nor increase in that of the possible P2C transition was observed within a three hour time-course when using either method.

There are two possibilities as to why we found no evidence to suggest that CCBL1 catalyses the first step of the pipecolate pathway of L-lysine metabolism: either CCBL1 does not catalyse this reaction or the conditions used in our experiments didn't facilitate this activity. Future experiments to further investigate the potential role of CCBL1 in lysine metabolism should consider the following issues:

- The concentration of CCBL1 protein synthesised using the *in vitro* translation (IVT) system (approximately 20 ng/μL) may not have been sufficient to produce enough P2C for detection using the spectrophotometric or mass spectrometry-based assays. A high-yield IVT kit

Figure 6.4.13: Chromatograms illustrating the transitions identified through the enzymatic conversion of D-lysine by D-amino acid oxidase. The two transitions for each analyte with superior peak shape and signal intensity are shown. textbf(a) Lysine and **(b)** theoretical Δ^1-piperideine-2-carboxylate (P2C).



which can produce up to 100 times more overexpressed protein than the system used in this chapter is commercially available. Therefore, this kit may be used to produce higher concentrations of CCBL1 for further experiments.

- Although D-amino acid oxidase (DAAO) has been shown to catalyse the oxidative deamination of D-lysine to form KAC/P2C, this reaction proceeds extremely slowly as a result of product inhibition (Yagi *et al.*, 1969). Oxidised DAAO forms a charge-transfer complex with the enamine form of P2C; this involves an association of the two molecules in which a fraction of electronic charge is transferred between them (Nishina *et al.*, 1991). The reactions described in this chapter were performed at pH 8.3 as these conditions have been repeatedly published for the reaction of other D-amino acids with DAAO. However, lowering the pH at which the reaction is carried out may reduce the product inhibition of DAAO as the imine form of P2C would be expected to predominate (Nishina *et al.*, 1991).

Therefore, solutions with wider ranges of buffering capacity should be assessed in order to maximise the chances of detecting KAC or P2C.

- Previous studies examining the ability of members of the kynurenine aminotransferase enzyme family to catalyse the transamination of L-lysine have demonstrated an increased catalytic ability in strongly alkaline conditions (Han *et al.*, 2009a; Hallen *et al.*, 2013). Thus, enzyme activity may be more readily observed at higher pH's. However, these conditions would not be representative of those present physiologically *in vivo*.

- Namwat *et al.* (2002) performed functional characterisation of the *VisA* gene in *Streptomyces virginiae*, encoding an enzyme which catalyses the conversion of L-lysine to KAC/P2C. Enzyme activity was optimal at 35°C and between pH 7.5 - 8.5. Multiple different α-keto acid amino acceptors were investigated and revealed that at a concentration of 100 nM 2-oxobutyrate resulted in the highest activity, whereas at 10 nM 2-oxohexanoate and 2-oxovalerate were optimal. Although for the majority of α-keto acids tested, using higher concentrations resulted in enhanced enzyme activity, this was not always the case. This may be explained by the fact that some α-keto acids can have inhibitory effects at high concentrations (Bruntner and Bormann, 1998). If CCBL1 similarly catalyses this reaction in humans, activity would likely be highly dependent on the α-keto acid substrate. During initial experiments, 100 nM of 2-oxobutyrate, 2-oxoglutarate, pyruvate and oxaloacetate were tested as α-keto substrates for CCBL1. Thus, production of KAC/P2C may not have been observed because either the substrates were not efficient amine acceptors or the concentrations used were inhibiting CCBL1 activity. Further experiments using lower (10 nM) concentrations of a wide range of substrates including 2-oxohexanoate and 2-oxovalerate should be performed.

- There is experimental evidence for the existence of three CCBL1 mRNA transcripts encoding proteins of differing lengths (Figure 6.4.14). The first to be identified was CCBL1_001, which encodes a polypeptide of 422 amino acids (Perry *et al.*, 1995). The second, CCBL1_002, lacks the third coding exon resulting in a smaller 372 amino acid isoform (Gerhard *et al.*, 2004). Finally, CCBL1_201 uses an alternate upstream translation start codon, resulting in an additional 94 N-terminal amino acids compared to CCBL1_001. Although the latter has been deposited in both the National Center for Biotechnology Information (NCBI) and Ensembl genome databases the sequence was identified as part of a large-scale cDNA sequencing project for which neither the full study methodology nor results are publicly available. Therefore further confirmation of its existence *in vivo* is required (Han *et al.*, 2010). *In silico* prediction suggests that the additional N-terminal region may contain a

putative mitochondrial targeting sequence (Claros and Vincens, 1996). The CCBL1_001 transcript was used in the studies described in this chapter due to its established expression and activity in human tissues. Although the Val272 residue which is mutated in patient Y does not lie within this N-terminal region of CCBL1_201, it is possible that this residue may interact with others in this region to confer a different biological role to CCBL1_001 *in vivo*. Indeed, the additional N-terminal amino acids may result an altered tertiary protein structure which in turn may confer different catalytic capabilities, including activity towards L-lysine. Further studies should be performed to investigate the potentially different physiological roles of this protein transcript.

Figure 6.4.14: Sequence alignment of the three human CCBL1 transcripts. Green, additional N-terminal 94-amino acid sequence in CCBL1_201; orange, sequence of 50 amino acids which is absent in the CCBL1_002 transcript.

```
CCBL1_201      MFRTAAAISVHLVGPLQGRKAGAPLTWFLHQTQVAEKNVGSPAPRPTSHLGLANPQERKV    60
CCBL1_001      ------------------------------------------------------------    0
CCBL1_002      ------------------------------------------------------------    0

CCBL1_201      PQGPQSPCEVPAAEGILRTGGAKLPIQGHPAKLAMAKQLQARRLDGIDYNPWVEFVKLAS    120
CCBL1_001      ----------------------------------MAKQLQARRLDGIDYNPWVEFVKLAS    26
CCBL1_002      ----------------------------------MAKQLQARRLDGIDYNPWVEFVKLAS    26
                                                  *************************

CCBL1_201      EHDVVNLGQGFPDFPPPDFAVEAFQHAVSGDFMLNQYTKTFGYPPLTKILASFFGELLGQ    180
CCBL1_001      EHDVVNLGQGFPDFPPPDFAVEAFQHAVSGDFMLNQYTKTFGYPPLTKILASFFGELLGQ    86
CCBL1_002      EHDVVNLGQGFPDFPPPDFAVEAFQHAVSGDFMLNQYTKTF------------------    67
               ***************************************

CCBL1_201      EIDPLRNVLVTVGGYGALFTAFQALVDEGDEVIIIEPFFDCYEPMTMMAGGRPVFVSLKP    240
CCBL1_001      EIDPLRNVLVTVGGYGALFTAFQALVDEGDEVIIIEPFFDCYEPMTMMAGGRPVFVSLKP    146
CCBL1_002      -------------------------------VIIIEPFFDCYEPMTMMAGGRPVFVSLKP    96
                                              ****************************

CCBL1_201      GPIQNGELGSSSNWQLDPMELAGKFTSRTKALVLNTPNNPLGKVFSREELELVASLCQQH    300
CCBL1_001      GPIQNGELGSSSNWQLDPMELAGKFTSRTKALVLNTPNNPLGKVFSREELELVASLCQQH    206
CCBL1_002      GPIQNGELGSSSNWQLDPMELAGKFTSRTKALVLNTPNNPLGKVFSREELELVASLCQQH    156
               ************************************************************

CCBL1_201      DVVCITDEVYQWMVYDGHQHISIASLPGMWERTLTIGSAGKTFSATGWKVGWVLGPDHIM    360
CCBL1_001      DVVCITDEVYQWMVYDGHQHISIASLPGMWERTLTIGSAGKTFSATGWKVGWVLGPDHIM    266
CCBL1_002      DVVCITDEVYQWMVYDGHQHISIASLPGMWERTLTIGSAGKTFSATGWKVGWVLGPDHIM    216
               ************************************************************

CCBL1_201      KHLRTVHQNSVFHCPTQSQAAVAESFEREQLLFRQPSSYFVQFPQAMQRCRDHMIRSLQS    420
CCBL1_001      KHLRTVHQNSVFHCPTQSQAAVAESFEREQLLFRQPSSYFVQFPQAMQRCRDHMIRSLQS    326
CCBL1_002      KHLRTVHQNSVFHCPTQSQAAVAESFEREQLLFRQPSSYFVQFPQAMQRCRDHMIRSLQS    276
               ************************************************************

CCBL1_201      VGLKPIIPQGSYFLITDISDFKRKMPDLPGAVDEPYDRRFVKWMIKNKGLVAIPVSIFYS    480
CCBL1_001      VGLKPIIPQGSYFLITDISDFKRKMPDLPGAVDEPYDRRFVKWMIKNKGLVAIPVSIFYS    386
CCBL1_002      VGLKPIIPQGSYFLITDISDFKRKMPDLPGAVDEPYDRRFVKWMIKNKGLVAIPVSIFYS    336
               ************************************************************

CCBL1_201      VPHQKHFDHYIRFCFVKDEATLQAMDEKLRKWKVEL     516
CCBL1_001      VPHQKHFDHYIRFCFVKDEATLQAMDEKLRKWKVEL     422
CCBL1_002      VPHQKHFDHYIRFCFVKDEATLQAMDEKLRKWKVEL     372
               ********************************
```

6.5 SUMMARY

In this chapter we have described a patient affected by severe spastic paraparesis, speech difficulties and persistent hyperlysinaemia. Disease was progressive with severe cachexia resulting in her death at 23 years of age. After extensive genetic and biochemical investigations, combined homozygosity mapping and whole exome sequencing identified a homozygous missense variant in the *CCBL1* gene. This was deemed to be a plausible candidate for pathogenesis in this family given the known role of this enzyme in the homeostasis of kynurenic acid in the brain and postulated function in lysine metabolism.

We have established a LC-MS/MS method for the quantitation of kynurenine and kynurenic acid, which may also be used for the assessment of kynurenine aminotransferase activity in patients with other neurological disorders. Although it has been determined that the p.Val272Met variant identified in our patient does not adversely affect kynurenine aminotransferase activity, the possibility of CCBL1 having a role in L-lysine metabolism has not been ruled out. Indeed, further experimentation as described above may provide more insights into the validity of this hypothesis. Furthermore, the structure of recombinant CCBL1 has already been solved by X-ray crystallography (Han *et al.*, 2009b). Expert analysis of these models using *in silico* tools to predict the localisation and conformation of any solvent pockets or cavities that could bind L-lysine may also provide supplementary supportive or contradictory evidence.

Finally, the possibility remains that the variants in *CCBL1* are not pathogenic in this family. Disease may be caused by copy number variants or mutations in intronic or regulatory regions that were not detectable by whole exome sequencing. If more detailed and rigorous scrutiny of the genetic data obtained to-date did not reveal any alternative candidate mutations, a whole genome sequencing approach may be beneficial in this case. Another possibility is that this family is affected by two recessive disorders. Recent studies have suggested that 4.6% of patients have a blended phenotype (Yang *et al.*, 2014); however in cohorts comprising a high degree of consanguinity such as that analysed using our gene panel, this proportion was much higher at 9.1%. The patient described in this chapter had a brother who, although also had difficulties with walking and speech, did not have spastic paraparesis or any amino acid abnormalities. Three cousins were also described to have an inability to walk and speech problems, although they resided in India and more detailed clinical information was not available. Therefore, it is possible that two recessive disorders are present in this family and the hyperlysinaemia identified in the proband may be caused by a secondary genetic defect. Indeed, this metabolic abnormality may not be contributing to pathogenicity nor be a result of the defect underlying the motor and speech difficulties in this family. In order to examine this possibility further, the segregation

of the *CCBL1* variant and any other candidate mutations should be carried out in as many extended family members as possible including the affected cousins.

In summary, this case again demonstrates the difficulties encountered when interpreting next-generation sequencing data without functional characterisation of candidate variants. These issues are even more pronounced when confirmatory assays are not already established or a novel gene or protein function is being postulated. In the future we will need to find better, more high-throughput methodologies that will enable us to confirm the effects of variants of uncertain significance identified by whole exome or whole genome sequencing.

7

FUNCTIONAL CHARACTERISATION OF INBORN ERRORS OF VITAMIN B_6 METABOLISM USING NOVEL ASSAYS

7.1 INTRODUCTION

With the increasing prevalence of next-generation sequencing for the diagnosis of patients where traditional genetic and biochemical testing has been uninformative, there is an ever-growing need for functional assays to determine if the identified variants are pathogenic. This chapter describes the adaptation of an ultra-performance liquid chromatography tandem mass spectrometry (UPLC-MS/MS) method to measure B_6 vitamer profiles in human fibroblasts and quantify PNPO enzyme activity. Its application to exploring the mechanisms and abnormalities underlying disease in both patients with and without a diagnosis are described. The pathogenic aetiology of a novel disorder, PROSC deficiency, was also examined using conventional molecular techniques and the newly established mass spectrometry-based methods.

7.1.1 *Vitamin B_6 metabolism and the importance of pyridoxal 5'-phosphate*

Epilepsy is the most common neurological disorder of childhood, with around one child in every two hundred suffering with the condition (Shinnar and Pellock, 2002). For most children that have epilepsy there is not a clear aetiology and approximately a third of all affected children have seizures that cannot be controlled with anti-epileptic drugs (AEDs) (Sisodiya *et al.*, 2002). These children have an increased risk of death and may need to undergo high-risk invasive surgery. Inborn errors of metabolism are rarely found to be the cause of epilepsy. However, seizures are a common finding in children with neurometabolic disorders (NMD) and can be due to one of many pathogenic mechanisms including: energy deficiency, toxicity, impaired neuronal function, disturbance of neurotransmitter systems, vitamin/cofactor disturbances or associated brain malformations (Wolf *et al.*, 2005). A well known group of NMD are the seizure disorders in which patients respond to treatment with vitamin B_6.

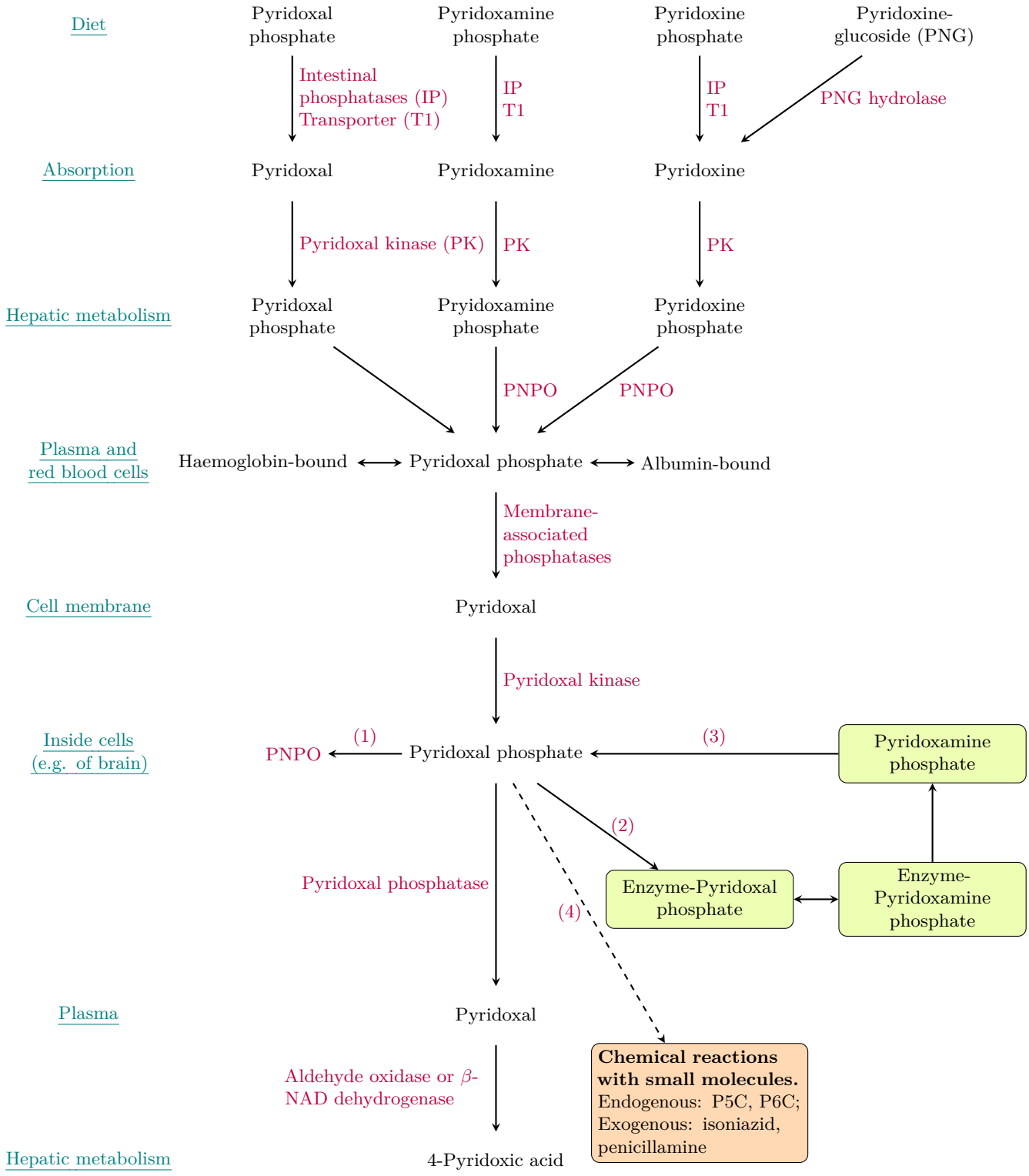
7.1.1.1 *Pathway of vitamin B_6 metabolism*

Vitamin B_6 is present in the body as 6 vitamers: pyridoxine (PN), pyridoxamine (PM) and pyridoxal (PL) and their 5' phosphate esters. However, only pyridoxal 5'-phosphate (PLP) is the

active vitamer and has cofactor activity. Vitamin B_6 is present in a variety of foods including meats, pulses, cereals, vegetables and fruit (Reynolds, 1988; Roth-Maier *et al.*, 2002). Some is also derived from the intestinal flora, where it is then taken up by colonocytes via a specific carrier-mediated mechanism (Said *et al.*, 2008). The majority of animal-derived B_6 consists of PLP and pyridoxamine 5'-phosphate (PMP), whilst that from human milk primarily exists as PLP or PL. In contrast plants contain mostly PN, pyridoxine 5'-phosphate (PNP) and the glycosylated form of pyridoxine, pyridoxine-5'-β-D-glucoside (PNG). Enterocytes, and cell membranes in general, are not permeable to the phosphorylated B_6 vitamers, therefore these are hydrolysed to their free bases by intestinal phosphatases prior to absorption in the upper small intestine. PNG is hydrolysed to PN by pyridoxine-5'-β-D-glucoside hydrolase, and the brush border membrane lactase phlorizin hydrolase (Mackey *et al.*, 2004). This uptake occurs via a carrier-mediated system, although the molecular identity of this transporter (T1) remains unknown (Said, 2004). PL, PM and PN are then trapped within the enterocytes through phosphorylation by pyridoxal kinase (PK), prior to being dephosphorylated by plasma membrane phosphatases and released into the portal vein to be transported to the liver. Once localised in the hepatocytes, the vitamers are again phosphorylated by PK; PMP and PNP are then converted to PLP by cytosolic pyridox(am)ine 5'-phosphate oxidase (PNPO). This PLP is then released into the blood where it is transported to the relevant tissues; the PLP is bound to erythrocyte haemoglobin or the lysine-190 amino acid residue of plasma albumin in order to protect the vitamer from hydrolysis (Bohney *et al.*, 1992). However, only the free bases can cross the blood-brain barrier and therefore PLP is first cleaved to PL by the ectoenzyme tissue non-specific membrane-associated alkaline phosphatase and transported into the CSF by an active transport mechanism. Uptake of the free vitamins from CSF into brain cells then occurs via a similar mechanism. Once in the brain cell, PL is phosphorylated by pyridoxal kinase, thereby trapping the active cofactor (Spector, 1978). Approximately 50% of the dietary intake of vitamin B_6 is excreted in the urine as 4-pyridoxic acid, a compound which is synthesised when unbound pyridoxal is oxidised by aldehyde oxidase or β-NAD dehydrogenase (Schwartz and Kjeldgaard, 1951; Stanulović *et al.*, 1976).

In addition to the *de novo* synthesis of PLP from dietary sources, PLP can also be recycled from the degraded apoenzymes within which it was functioning as a cofactor (González *et al.*, 2007). This "salvage pathway" functions to maintain PLP homeostasis by catalysing the interconversions of the six vitamers using the enzymes described above; pyridoxal kinase converts the free vitamers to their phosphorylated forms and PNP and PMP can be oxidised by PNPO to form PLP. The later reaction is particularly important as PMP is the major vitamer released from the degradation of apoenzymes which catalyse transamination reactions (Figure 7.1.1).

Figure 7.1.1: Enzymes and transporters involved in the metabolism of dietary vitamin B_6 and the synthesis and homeostasis of pyridoxal 5'-phosphate. (1) PNPO is subject to feedback inhibition from PLP. (2) PLP forms Schiff base interactions with the ϵ-amino group of lysine residues within proteins which is essential for cofactor activity. (3) The "salvage pathway" recycles the PLP cofactor from degraded enzymes by transamination reactions. (4) PLP can react with small molecules rendering it unavailable as a cofactor. These molecules include L-Δ1-pyrroline-5-carboxylic acid (P5C) and Δ1-piperideine-6-carboxylic acid (P6C), causing hyperprolinaemia type II and antiquitin deficency, respectively and drugs such as isoniazid and penicillamine.



7.1.1.2 *Inborn errors of metabolism resulting in a deficiency of PLP*

As described above, all dietary forms of B_6 can be converted to PLP which is a cofactor for more than 140 enzymes in humans, many of which are involved in the synthesis and degradation of neurotransmitters. PLP has excellent electron sink properties that make it a versatile organic catalyst; meaning that the molecule contains a group that can pull electrons from a reactive centre, stabilising an electron-deficient intermediate or transition state. In the case of PLP, the aldehyde group forms the electron sink which reacts with endogenous and exogenous nucleophiles (Christen and Mehta, 2001). All PLP-dependent enzymes, with the exception of glycogen phosphorylase, act upon amino acids or amines. In the brain, PLP-dependent enzymes are involved in the metabolism of many amino acid and amine neurotransmitters, including dopamine, serotonin, GABA, glutamate, d-serine, glycine and taurine. These enzymes are also important in the synthesis of neuroprotective compounds such as kynurenic acid (Table 7.1.1).

Table 7.1.1: PLP-dependent enzymes important for normal neurological function. Adapted from Surtees *et al.* (2006)).

Enzyme	Function
Aromatic amino acid decarboxylase	Dopamine and serotonin synthesis
Branched chain amino acid, 2-oxoglutarate aminotransferase	Glutamate synthesis
GABA transaminase	GABA catabolism
Glutamate decarboxylase	GABA synthesis
Glycine cleavage system	Glycine catabolism
Kynureninase	Quinolinic acid synthesis
Kynurenine aminotransferase	Kynurenic acid synthesis
L-serine racemase	D-serine synthesis

Given the central importance of PLP in amino acid and neurotransmitter metabolism it is not surprising that inborn errors of metabolism which affect availability of PLP present with a neurological phenotype frequently involving seizures; these include pyridoxine dependent epilepsy (PDE) due to mutations in *ALDH7A1*, hyperprolinaemia type II, hypophosphatasia and pyridox(am)ine 5'-phosphate oxidase (PNPO) deficiency. Both PDE and hyperprolinaemia type II are secondary PLP disorders, resulting in the accumulation of L-Δ^1-piperideine-6-carboxylic acid (P6C) and L-Δ^1-pyrroline-5-carboxylic acid (P5C) respectively, which react with PLP thereby sequestering it and rendering it inactive. Hypophosphatasia however, directly affects the vitamin B_6 pathway, resulting in a deficiency of the tissue non-specific alkaline phosphatase which is required for the dephosphorylation of circulating PLP to PL therefore allowing it to cross the

blood-brain barrier. The main B_6 responsive seizure disorders, PDE and PNPO deficiency, are described in more detail in Section 7.1.2 and Section 7.1.3.

7.1.2 *Antiquitin (ALDH7A1) deficiency*

Pyridoxine-dependent epilepsy (PDE) due to mutations in *ALDH7A1* and dysfunction of the antiquitin protein has already been briefly described in Section 4.2.1. A deficiency of α-aminoadipic semialdehyde (α-AASA) dehydrogenase, otherwise known as antiquitin, results in the accumulation of L-Δ^1-piperideine-6-carboxylate (P6C). This metabolite then forms an adduct with PLP rendering it inactive as a cofactor (Mills *et al.*, 2006). More than 200 patients with genetically confirmed PDE have been reported in the literature, although many more patients are being diagnosed in clinical practice. More than 80 pathogenic mutations have been reported in *ALDH7A1*, with one missense mutation (p.Glu399Gln) accounting for approximately 30% published alleles (Mills *et al.*, 2010). Classically, patients present with seizures beginning within the first month of life which are unresponsive to conventional AEDs but respond, at least partially, to treatment with PN. Global developmental delay or intellectual disability is present in up to 75% of patients (van Karnebeek *et al.*, 2016) and the severity of this impairment does not correlate with age at seizure onset or diagnosis, seizure type or biochemical abnormalities at presentation (Bok *et al.*, 2012). Biochemically, elevations of α-AASA, P6C and pipecolic acid can be identified. Patients may also be misdiagnosed with hypoxic ischaemic encephalopathy in the newborn period as infants are commonly born in poor condition with evidence of foetal distress. Diagnosis can be complicated by the multiple metabolic abnormalities that can be present during the acute phase including hypoglycaemia, lactic acidosis, electrolyte disturbances, coaguolpathy and abnormal plasma and CSF amino acid concentrations. Similarly, the presence of structural brain abnormalities such as ventriculomegaly, corpus callosum and cerebellar hypoplasia, white matter abnormalities and delayed myelination may also increase the list of differential diagnoses (van Karnebeek *et al.*, 2016).

In the vast majority of patients, treatment with a single intra-venous dose of 100 mg PN, followed by an oral regimen of 5 - 10 mg/kg/day results in seizure resolution without the need for additional AED treatment. However, breakthrough seizures during periods of intercurrent illness are common (Mills *et al.*, 2010; Stockler *et al.*, 2011). Given the high prevalence of adverse developmental outcomes, PN has been administered to pregnant mothers and newborn infants before the onset of seizures. Although neonatal seizures were prevented, neurodevelopmental delay persisted, albeit to a lesser degree than siblings who were antenatally untreated (Rankin *et al.*, 2007; Bok *et al.*, 2010). This suggests that there may be a secondary mechanism besides

the deficiency of PLP within the brain, causing neurological sequelae. Indeed, reducing the potentially neurotoxic accumulation of P6C/α-AASA by decreasing flux through the lysine catabolic pathway by dietary lysine restriction and arginine supplementation has been shown to improve behavioural and language skills in some patients (van Karnebeek *et al.*, 2012; Coughlin *et al.*, 2015).

7.1.3 *Pyridox(am)ine 5'-phosphate oxidase (PNPO) deficiency*

PNPO deficiency is a primary disorder which directly affects PLP synthesis. It is caused by autosomal recessive loss-of-function mutations in *PNPO*, whose function is described in Section 7.1.1.1. The molecular basis of PNPO deficiency and initial cohort of patients were first defined in 2005 by Mills *et al.*. Classically these children were born prematurely and presented with seizures shortly after birth, including myoclonic jerks and severe tonic-clonic seizures, associated with an EEG showing a burst suppression pattern. Within the early cohorts of patients it was thought that the disorder was associated with typical biochemical findings that could be attributed to the impaired activity of PLP-dependent enzymes including aromatic L-amino acid decarboxylase, threonine dehydratase, glycine cleavage enzyme and ornithine δ-aminotransferase (Mills *et al.*, 2005). However, this has since been shown not to hold true and the biochemical phenotype is in fact much more broad, with many patients having unremarkable metabolic investigations (Hoffmann *et al.*, 2007). In the few patients who have had CSF PLP measured prior to supplementation, levels were reduced (Ormazabal *et al.*, 2008) although normal concentrations have also been reported (Levtova *et al.*, 2015).

One of the most noticeable initial differences between children with PNPO and those with antiquitin deficiency was that PNPO deficient neonates responded to treatment with PLP with a complete cessation of seizures within hours, but did not respond to trials of PN. This is intuitive because PLP can be transported from the gut through all the necessary conversions without any participation of PNPO whereas PN given either intra-venously or orally, requires the activity of PNPO to produce PLP in the brain (Figure 7.1.1). However, three recent studies have shown that some patients with PNPO deficiency do in fact respond to PN treatment, with a few experiencing a clinical deterioration upon switching to PLP (Mills *et al.*, 2014; Plecko *et al.*, 2014; Jaeger *et al.*, 2016). Collectively, they identified twenty patients harbouring nine mutations in differing homozygous or compound heterozygous combinations. The residual PNPO activity of each of these mutations was examined using either a cell-free *in vitro* translation asssay or transfection into Chinese hamster ovary cells but no clear relationship between higher activity

and PN-responsiveness, as would be expected if this response to PN was simply due to residual PNPO activity allowing conversion through the pathway, was found.

7.1.4 *PROSC deficiency*

Prior to the beginning of this PhD project, combined homozygosity mapping and whole exome sequencing identified a homozygous nonsense mutation in *PROSC* (proline synthetase co-transcribed [bacterial homolog]) in a consanguineous pedigree of Syrian origin containing three children with vitamin B_6-responsive seizures (unpublished data; Dr Niklas Darin, The Queen Silvia Children's Hospital, Sweden). Following this, sequencing of the *PROSC* gene was carried out in a cohort of children whose epilepsy had shown some response to treatment with PN or PLP and antiquitin and PNPO deficiency had been excluded genetically and/or biochemically. Four additional patients were identified harbouring an additional nonsense, three missense and two splice site mutations (unpublished data; Dr Philippa Mills, ICH, UK). All patients presented with seizures on the first day of life, with the exception of one patient who presented at one month of age. All also had acquired microcephaly and differing degrees of developmental delay. Plasma PLP concentrations were elevated in all patients that were tested; although all were on PN supplementation at the time, these levels were at least four-fold higher than PDE patients taking a similar PN dose.

PROSC encodes a ubiquitously expressed cytosolic protein of unknown function first identified in 1999 by Ikegawa *et al.*. Orthologs of *PROSC* are highly conserved and present in bacteria, plants and yeast, indicating an important role in cellular function. All orthologs share a PLP-binding site which is postulated to bind PLP through a Schiff base linkage to a conserved lysine residue (Ito *et al.*, 2013). Indeed, the structure of *PROSC* orthologs are extremely similar to the N-terminal domain of bacterial alanine racemase and eukaryotic ornithine decarboxylase. However, unlike alanine racemase, in *PROSC* orthologs the PLP is solvent exposed rather than interacting with a tyrosine residue within the polypeptide which is critical for catalytic racemisation activity (Watanabe *et al.*, 2002; Ito *et al.*, 2013). Supporting this, studies of the bacterial homologue, *YggS*, showed that the protein has no racemisation, transamination, deamination or decarboxylation activity against proteinogenic amino acids (Ito *et al.*, 2013). A more recent study has suggested that *Yggs* may be involved in PLP homeostasis (Prunetti *et al.*, 2016). The *YggS*-deficient *E. coli* examined by Prunetti *et al.* (2016) accumulated pyridoxine 5'-phosphate (PNP) and showed concentration-dependent toxicity towards pyridoxine (PN), therefore it has been hypothsised that PROSC plays a role in vitamin B_6 homeostasis. The exact function of PROSC and possible pathogenic mechanisms underlying its deficiency were examined in this study.

7.1.5 *Maintenance of PLP concentrations and prevention of damaging reactions*

The chemical properties of PLP make it highly reactive. In addition to facilitating its cofactor activity, the reactive aldehyde group at the 4'-position of PLP can react with many endogenous and exogenous nucleophiles within cells. This includes condensation reactions with hydrazines and sulphydryl compounds such as isoniazid and penicillamine, respectively. Other small molecules such as P6C and P5C which accumulate in patients with antiquitin deficiency and hyperprolinaemia type II, respectively, can inactivate PLP through a Knoevenagel condensation with the C_3 carbon. Given that PLP can form aldimines with primary and secondary amines, reactions with proteins can also occur; if unregulated it is thought that this may cause impairment of enzymatic activity (Vermeersch *et al.*, 2004). Similarly, PLP can form thiazolidine adducts with molecules such as cysteine resulting in the accumulation of damaged metabolites (Liu *et al.*, 2013). Given this high degree of chemical reactivity, free pools of PLP must be tightly controlled to both ensure that supplies are adequate for apoenzyme function and damaging interactions are minimised.

As detailed in Section 7.1.1.1, the major synthesis and catabolic pathways of vitamin B_6 metabolism are well characterised, however the mechanisms by which PLP concentrations within tissues are tightly controlled are still not fully understood. Many studies have shown that although 4-pyridoxic acid excretion and plasma vitamer concentrations increase when subjects are fed a diet containing a high B_6 content, PLP concentrations within the tissues including muscle liver and brain remain unchanged (Coburn *et al.*, 1991; Schaeffer *et al.*, 1995). This tight regulation is likely to be due to multiple concurrent processes. When exogenous supplies of vitamin B_6 are sufficient, PLP concentrations are balanced by the degree of PLP binding to cellular proteins, the transport of the precursor vitamers and phosphatase activity. Both PNPO and PK are also subject to negative feedback inhibition by PLP (i.e. high PLP concentrations result in lower activity) (di Salvo *et al.*, 2015). Pyridoxal kinase also plays a role by compartmentalising the phosphorylated vitamers as they cannot cross the plasma membrane. It has been suggested that channelling, thereby avoiding the release of PLP into the cytoplasm and the occurrence of damaging side-reactions, is the mechanism responsible for effective transfer of PLP to apoenzymes (di Salvo *et al.*, 2011). In addition to the active site, PNPO contains a secondary allosteric site which binds PLP tightly and whose exact function is unknown (Musayev *et al.*, 2003). Crystal structures of PNPO show the presence of a putative tunnel between the active site and the secondary PLP site, suggesting that the vitamer may be transferred across the protein without contacting the solvent (Safo *et al.*, 2005). The mechanism by which PLP is delivered to their

target enzymes is still unknown. However, some mechanistic insight into how this occurs has been gained from the study of patients with PROSC deficiency (Section 7.1.4).

7.1.6 *Aims of this chapter and advantages of using mass spectrometry-based assays for the evaluation of patients with inborn errors of vitamin B_6 metabolism*

The aim of this chapter was to adapt existing mass spectrometry-based methods to enable the examination of B_6 vitamer profiles in cells and develop a novel method for the quantitation of PNPO enzyme activity to examine mutation pathogenicity in patient cells. High-performance liquid chromatography (HPLC) (Bisp *et al.*, 2002; Rybak and Pfeiffer, 2004) and liquid chromatography-tandem mass spectrometry (LC-MS/MS) (Footitt *et al.*, 2013) techniques have already been developed to quantify all of the B_6 vitamers and 4-pyridoxic acid. The main advantages of these methods compared to traditional methods for evaluating vitamin B_6 status include the ability to measure all of the vitameric forms of vitamin B_6 and greater analytical sensitivity, which may reveal subtle abnormalities indicative of specific known or novel disorders. Using these assays, the differences between PNPO-deficient patients responsive to PN and PLP were examined and the B_6 vitameric profiles in patients with PROSC deficiency were determined. The pathogenic aetiology underlying PROSC deficiency was also explored further using molecular biology, cell culture and additional mass-spectrometry-based techniques.

7.2 METHODS

The methods described in this chapter were used to characterise fibroblast cell lines derived from eight control patients, three children with confirmed or suspected PNPO deficiency, one child with ALDH7A1 deficiency, three children with PROSC deficiency and one parent who was heterozygous for a mutation in *PROSC*. Control fibroblasts were derived from patients having a skin biopsy for the diagnostic work-up of a metabolic disease but with no evidence of a disorder affecting vitamin B_6 metabolism. Of the three patients with confirmed or suspected PNPO deficiency, two were genetically confirmed to have PNPO deficiency; one was responsive to PLP (PNPO 1) and the other to PN (PNPO 3). In the final patient (PNPO 2) who was responsive to PLP, despite carrying a clinical diagnosis of PNPO deficiency, a definitive diagnosis had yet to be made because despite extensive cDNA sequencing only one heterozygous frameshift mutation had been identified. Details of mutations in each patient are shown in Table 7.2.1. The mutation analysis of these patients was carried out by Dr Philippa Mills prior to this study commencing.

Table 7.2.1: Mutations identified in the vitamin B_6-responsive for which fibroblasts were available for analysis in this study. (?), second mutation not identified.

Patient	Gene	Vitamer response	Nucleotide change	Amino acid change
PNPO 1	*PNPO*	PLP	c.364-1G>A; c.148G>A + c.364-1G>A; c.148G>A	Splice error; p.Glu50Lys + Splice error; p.Glu50Lys
PNPO 2	*PNPO*	PLP	c.641dupA; (?)	p.Val215Glyfs*14; (?)
PNPO 3	*PNPO*	PN	c.674G>A; c.674G>A	p.Arg225His; p.Arg225His
PDE 1	*ALDH7A1*	PN	c.1195G>C; c.1195G>C	p.Glu399Gln; p.Glu399Gln
PROSC 1	*PROSC*	PLP	c.524T>C; c.524T>C	p.Leu175Pro; p.Leu175Pro
PROSC 2	*PROSC*	PLP	c.207+1G>A; c.320-2A>G	Splice error; Splice error
PROSC 3	*PROSC*	PN	c.233C>G; c.233C>G	p.Ser78*; p.Ser78*
PROSC het	*PROSC*	n/a	c.233C>G	p.Ser78*

The ultra-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) method described in Section 2.11 was used to examine the B_6 vitamer profiles in patient fibroblasts and CSF. The enzyme activity of pyridoxal kinase (PK) and pyridox(am)ine 5'-phosphate oxidase (PNPO) was assessed as described in Section 2.13.

In order to examine the subcellular location of the accumulated PLP in PROSC-deficient cells, fibroblasts were harvested and lysed (Section 2.11) and the protein concentration was determined (Section 2.7.4). 20 μL of supernatant was added to an Amicon Ultra-0.5 mL 3 kDa Centrifugal Filter (Millipore) and centrifuged at 14,000 x *g* for 30 minutes. The B_6 vitamers in the resulting fractions (>3 kDa and <3 kDa) and in the unfiltered supernatant were analysed as described previously (Section 2.11). Results were corrected for the unfiltered protein concentration.

The PLP-cysteine thiazolidine compound was synthesised by mixing 300 μL of 4.4 mM PLP, 300 μL of 61 mM L-cysteine and 2.4 mL of dH_2O. The solution was then incubated in the dark at 37°C for 3 hours. The method described in Section 2.12 was utilised to analyse the PLP-cysteine content of the fibroblast samples.

In order to examine the effect of the *PROSC* mutations on protein translation, cDNA analysis was undertaken. Total RNA was isolated from control and patient (PROSC 1 PROSC 2 and PROSC 3) fibroblasts using the method described in Section 2.8.1 and diluted to a final concentration of approximately 200 ng/μL. Messenger RNA (mRNA) was then reverse-transcribed to produce approximately 1 μg/μL of complementary DNA (cDNA) (Section 2.8.3). cDNA corresponding to the *PROSC* and *HPRT1* genes was amplified using the primers listed in Appendix 2.8.1. Each fragment was excised and cloned into the TOPO 2.1 vector (Sections 2.16.11 and 2.16.8) prior to Sanger sequencing (Section 2.3.3.2). Western blotting and real-time polymerase chain reaction (qRT-PCR) were carried out using the methods outlined in Section 2.10 and Section 2.9, respectively. In order to assess the growth of patient cells, $2.0x10^4$ cells/mL were

seeded in 25 cm^2 flasks and cultured for seven days in media depleted or repleted of pyridoxine. Cells were counted using the method described in Section 2.7.2.

7.3 RESULTS AND DISCUSSION

7.3.1 *Method development of UPLC-MS/MS method for B_6 vitamer quantitation*

Patients with suspected PNPO deficiency are often on PN or PLP supplementation, therefore it can be difficult to determine the pathogenicity of novel sequence variants within *PNPO* by measuring plasma vitamer profiles. We therefore aimed to directly quantify PNPO enzyme activity from patient fibroblasts and subsequently use this method to determine the differences in vitamer profiles between disorders affecting vitamin B_6 metabolism. The UPLC-MS/MS method described in this chapter for the analysis of B_6 vitamers was adapted from that already published for the quantitation of these analytes in plasma by Footitt *et al.* (2013). In addition to enabling the examination of fibroblasts, *in vitro* conditions can be modified to gain greater insights into disease pathogenesis. The method described by Footitt *et al.* (2013) has a limit of detection of 1 nM for all vitamers, which was not deemed sensitive enough for these purposes. The method was transferred to an Acquity Ultra Performance LC system linked to a triple quadrupole Xevo TQ-S instrument. Cone and collision energies were varied to provide optimal sensitivity and deuterated pyridoxamine (d_3-PM) was added as an additional internal standard to increase the accuracy of endogenous PM quantitation. The run time and injection volumes were reduced from 25 minutes and 25 μL to 5 minutes and 15 μL, respectively. Details of the parameters that were optimised are given below.

7.3.1.1 *Mobile phase*

The successful retention and separation of the B_6 vitamers is dependent on the mobile phase composition. In order to accurately determine the concentrations of any analyte using mass spectrometry, the compound must produce a single, reproducible peak without cross-talk between channels. An Acquity UPLC HSS T3 column (1.8 μm x 2.1 mm x 50 mm) fitted with a HSS T3 VanGuard guard column (Waters) was used for analysis as it was functionally analogous to the HS F5 column used by Footitt *et al.* (2013). This column consists of a trifunctional C_{18} alkyl phase bonded at a density that promotes the retention and separation of small, water-soluble polar compounds. The phase is also tolerant of the low pH's and highly aqueous mobile phases needed to retain the B_6 vitamers on the column. All LC-MS/MS methods analysing the B_6 vitamers use 3.7% acetic acid with variable concentrations of the ion-pairing reagent, heptafluorobutyric

acid (HFBA) as the aqueous phase, with either acetonitrile or methanol as the solvent phase (Midttun *et al.*, 2005; Footitt *et al.*, 2013; Albersen *et al.*, 2015). Methanol was determined to produce a better peak shape and sensitivity for the majority of vitamers when compared to acetonitrile (Figure 7.3.1a). The concentration of HFBA required within the aqueous phase was also investigated over a range of 0 - 0.1% (Figure 7.3.1c). An increase in HFBA concentration was found to improve the chromatography of the phosphorylated vitamers whilst having a deleterious effect on the unphosphorylated species. A composition of 3.7% acetic acid containing 0.01% HFBA was determined to be the best compromise achieving acceptable sensitivity for all analytes.

A flow rate of 0.8 mL/min was examined as is the standard methodology in our laboratory (Manwaring *et al.*, 2013); however, the vitamers were not well-retained on the column with all eluting within 0.5 minutes. Reducing the flow rate to 0.4 mL/min increased the retention times of the vitamers to between 0.69 - 1.42 minutes without compromising the chromatography (transitions and mobile phase gradient detailed in Table 2.11.3 and Table 2.11.2, respectively).

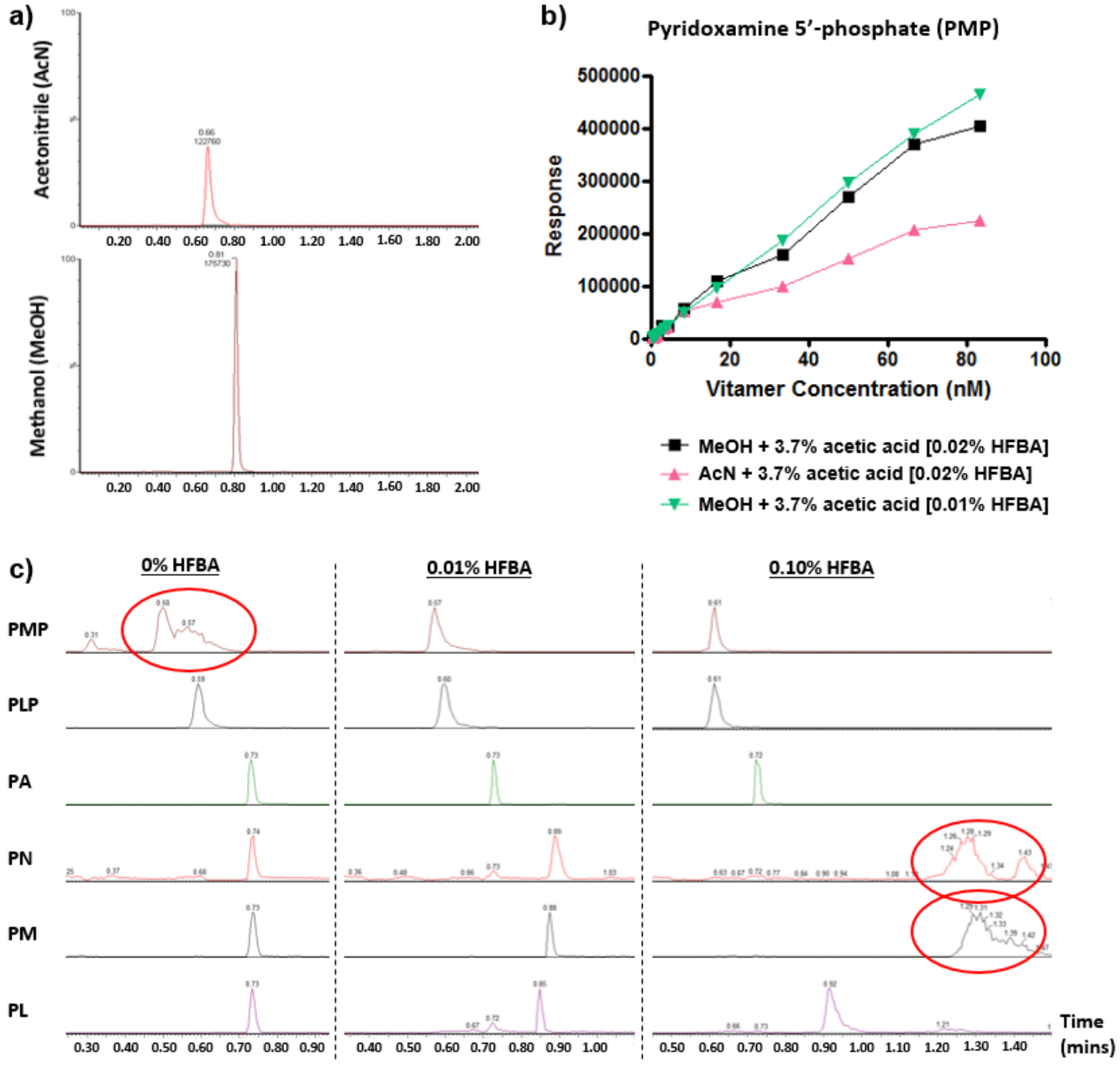
7.3.1.2 *Transitions*

The B_6 vitamers and 4-pyridoxic acid were detected using multiple reaction monitoring (MRM) mode. The transitions (i.e. the precursor/product ions) and the cone/collision voltages used to generate them were optimised by infusion of each analyte within the mobile phase composition at the time of elution of the analyte from the column. Several possible transitions were identified for each vitamer and one was then selected for the final method based on those which gave superior response (final conditions in Table 2.11.3). The mass spectrometer was operated in positive ion mode and a scan segment of 3.6 minutes from 0.4 to 4.0 minutes. The total sample run time was 5 minutes, including time for column re-equilibration. All analytes could be differentiated based on both their transitions (m/z ratios) and retention times, without cross-talk between ion pairs originating from different vitamers.

7.3.1.3 *Linearity*

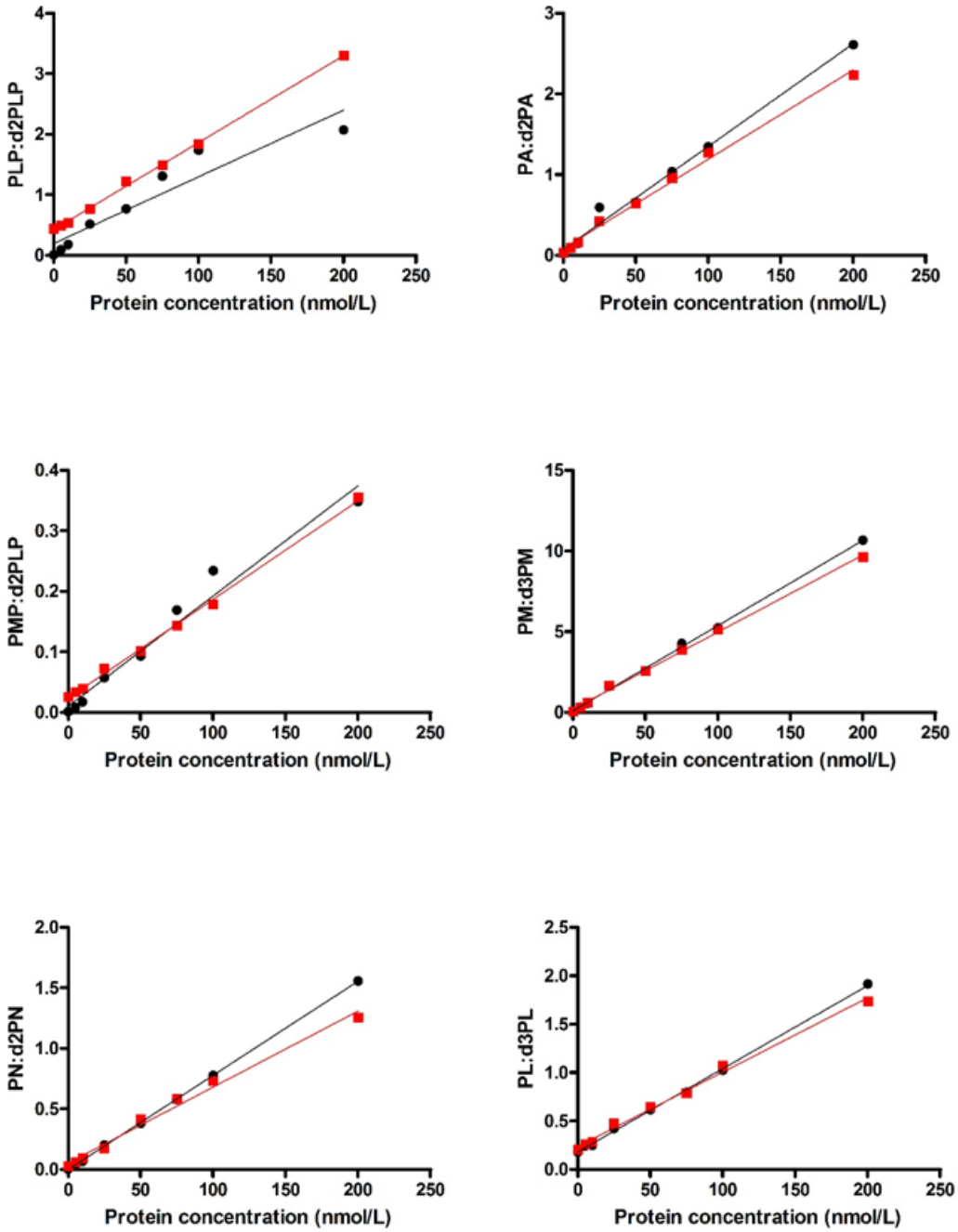
The linear range of each vitamer using three different mobile phase compositions (i) 3.7% acetic acid [0.02% HFBA] + acetonitrile, (ii) 3.7% acetic acid [0.01% HFBA] + methanol and (iii) 3.7% acetic acid [0.01% HFBA] + methanol was investigated. The latter provided the greatest linearity and sensitivity not only for PMP (Figure 7.3.1b), but also for all other vitamers. Calibration curves were constructed for each B_6 analyte in both water and a pooled mixture of fibroblast cell lysates. A range from 0 - 200 nmol/L was chosen because it is comparable to previous vitamer measurements in cell lines such as Chinese hamster ovary cells (Fargue *et al.*, 2013). Values of r^2 equal to 0.99 were achieved for each vitamer using linear regression. All vitamers had a linear

Figure 7.3.1: Optimisation of the parameters required for the quantitation of the B_6 vitamers in fibroblasts. **(a)** The effect of acetonitrile and methanol as mobile phase B on the chromatography and sensitivity of vitamer detection. **(b)** Linearity of the response for pyridoxamine 5'-phosphate (PMP) using an injection volume of 15 μL and different mobile phase compositions. **(c)** Different concentrations of heptafluorobutyric acid (HFBA) were added to mobile phase A [3.7% acetic acid] in order to improve chromatographic peak shape.



response up to 200 nmol/L, with the exception of PLP which remained linear up to 100 nmol/L. The limits of detection for this assay were determined to be 10 pmol/L for PA, PM, PN and PL, 100 pmol/L for PLP and 1 nmol/L for PMP. The final method for the quantitation of the endogenous B_6 vitamers in fibroblasts is described in Section 2.11.

Figure 7.3.2: Calibration curves of B6 analytes. Curves were generated in water (black) and cell lysate (red).



7.3.2 *Examining vitamin B_6 metabolism and homoeostasis in patients with B_6-responsive seizure disorders*

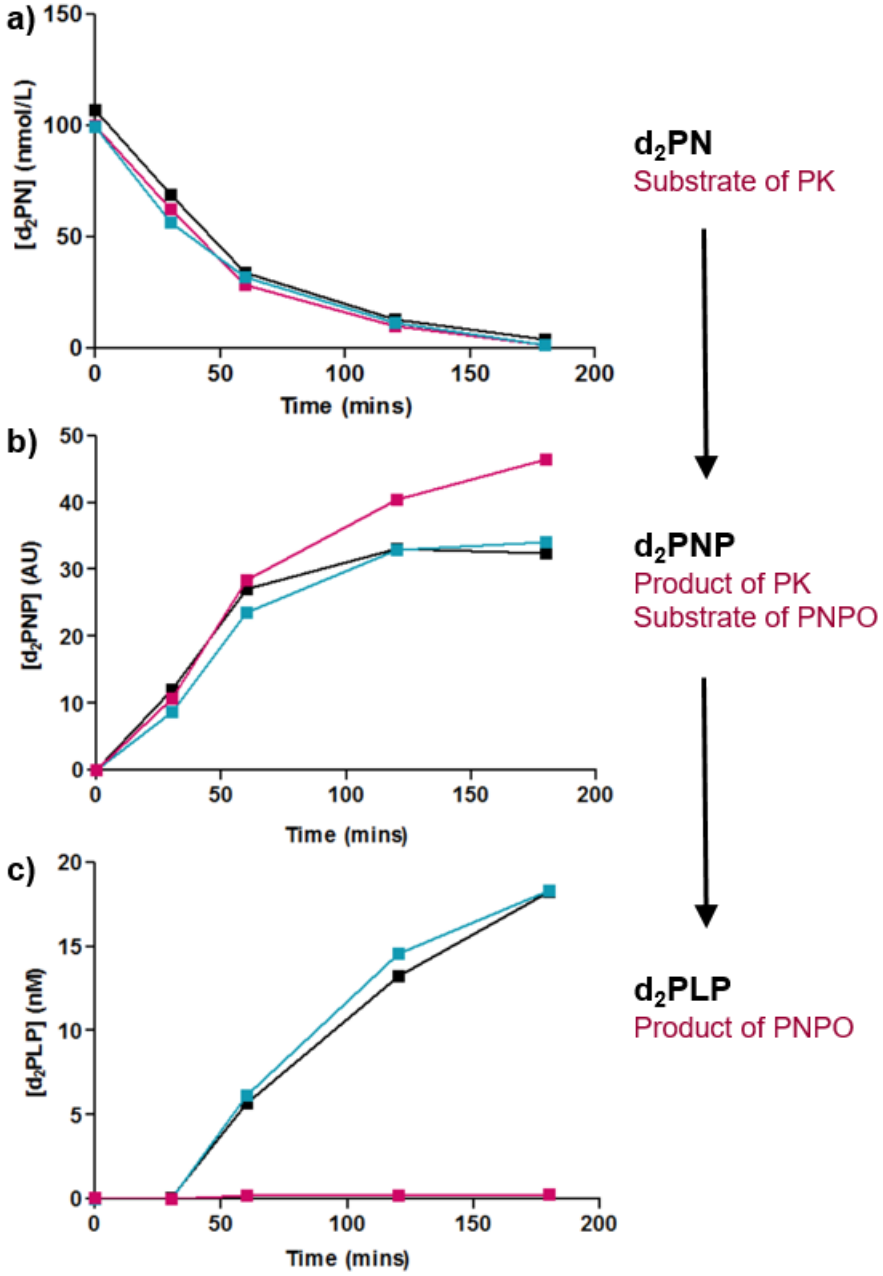
7.3.2.1 *Enzymatic assay of PNPO activity in control fibroblasts*

A diagnosis of PNPO can be suspected on the basis of a severe early-onset seizure disorder which shows a clinical response (in the form of a reduction or cessation of seizures) to either PN or PLP. However, the clinical response can be difficult to interpret as patients are typically receiving treatment with multiple AEDs concurrently and may be affected by other sequelae (e.g. complications of prematurity). A low concentration of PLP in the CSF prior to commencing treatment is also indicative but often not measured. The identification of two sequence variants in the *PNPO* gene is also strong evidence but their pathogenicity must be determined if the variants are novel and not known to be disease-causing. To date, no method has been described which facilitates the quantitation of PNPO activity in patient cells to confirm the diagnosis of PNPO deficiency and the pathogenicity of any sequence variants identified in *PNPO*. Therefore, the UPLC-MS/MS method described in Section 2.13 was established to facilitate the analysis of PNPO activity directly rather than indirectly (i.e. scrutinising subtle differences in B_6 vitamer profiles). The assay described in this chapter is a coupled assay of both pyridoxal kinase (PK) and PNPO enzyme activity (Figure 7.3.3), therefore the experimental conditions needed to facilitate the activity of both enzymes *in vivo*. The pH optimum of PNPO is between 9.0 - 10.0 (Wada and Snell, 1961), whereas that for PK is 6.0 (Neary and Diven, 1970). Flavin mononucleotide is also an essential cofactor for PNPO activity (Musayev *et al.*, 2003). Similarly, human PK requires monovalent (K^+) and divalent cations (Mg^{2+}) for its activation and ATP to donate the phosphate group required for phosphorylation (Musayev *et al.*, 2007).

Initial studies had been performed previously (unpublished data; Dr Philippa Mills, ICH, UK) which demonstrated that potassium phosphate buffered at pH 7.6 provided adequate activity for both enzymes. This buffer also contributed sufficient K^+ ions for optimal PK activity. In addition, the presence of phosphate ions inhibits the action of phosphatases present in the fibroblast lysate that otherwise function to dephosphorylate the deuterated B_6 vitamers. The chosen concentrations of FMN, $MgCl_2$ and ATP were based on those previously described (Wada and Snell, 1961; Neary and Diven, 1970; Musayev *et al.*, 2003, 2007).

In control fibroblast lysates, once exogenous d_2-PN is added it is converted to d_2-PNP by PK with almost complete disappearance of the substrate within a two hour period. This d_2-PNP accumulates linearly over the first 60 minutes before reaching a plateau and is then converted to d_2-PLP by PNPO. The increase in d_2-PLP concentration is linear between 30 - 60

Figure 7.3.3: Conversion of deuterated vitamers measured during the coupled PNPO enzyme assay. Graphical representation of the conversion of **(a)** d_2-pyridoxine (d_2PN), to **(b)** d_2-pyridoxine 5'-phosphate (d_2-PNP) and then to **(c)** d_2-pyridoxal 5'-phosphate (d_2-PLP). Black, control; blue, PROSC-deficient, red; PNPO-deficient.



minutes following an initial lag phase (Figure 7.3.3). PNPO activity was therefore quantified by measuring the increase in d_2-PLP concentration over this time period and was expressed in terms of pmol/hr/mg protein. Control PNPO activity ranged from 50 - 300 pmol/hr/mg protein (Figure 7.3.6). The findings in patients with PNPO, antiquitin and PROSC deficiency are discussed in Sections 7.3.3, 7.3.5 and 7.3.7, respectively.

7.3.2.2 *B_6 vitamer profiles in control fibroblasts*

After having established a direct method for the analysis of PNPO enzyme activity, this mass spectrometry-based assay was adapted to quantitate the endogenous B_6 vitamers in the fibroblasts of patients with PNPO, antiquitin and PROSC deficiency and determine the correlation with published findings in plasma. Initial experiments were carried out with cells that had been cultured in standard media containing 971 nM pyridoxine hydrochloride (Figure 7.3.4). The major vitameric species in control cells were PL (4 - 20 nmol/g cell protein), PLP (8 - 15 nmol/g cell protein) and PMP (5 - 35 nmol/g cell protein) whereas the concentrations of PM, PN and PNP were less than 1 nmol/g cell protein. Experiments were then performed to investigate the effect of variable concentrations of PN in the culture media on the cellular B_6 vitamer composition (Figure 7.3.5). In control fibroblasts, the levels of all vitamers remained unchanged, with the exception of PMP which showed an approximate 5-fold increase, upon culturing in media deficient in PN. This likely reflects an increase in the recycling of PMP from apoenzymes when no exogenous source of vitamin B_6 is present. The B_6 vitamer profiles of fibroblasts derived from patients with PNPO, antiquitin and PROSC deficiency cultured in conditions of PN depletion and repletion are discussed in Sections 7.3.3, 7.3.5 and 7.3.7, respectively.

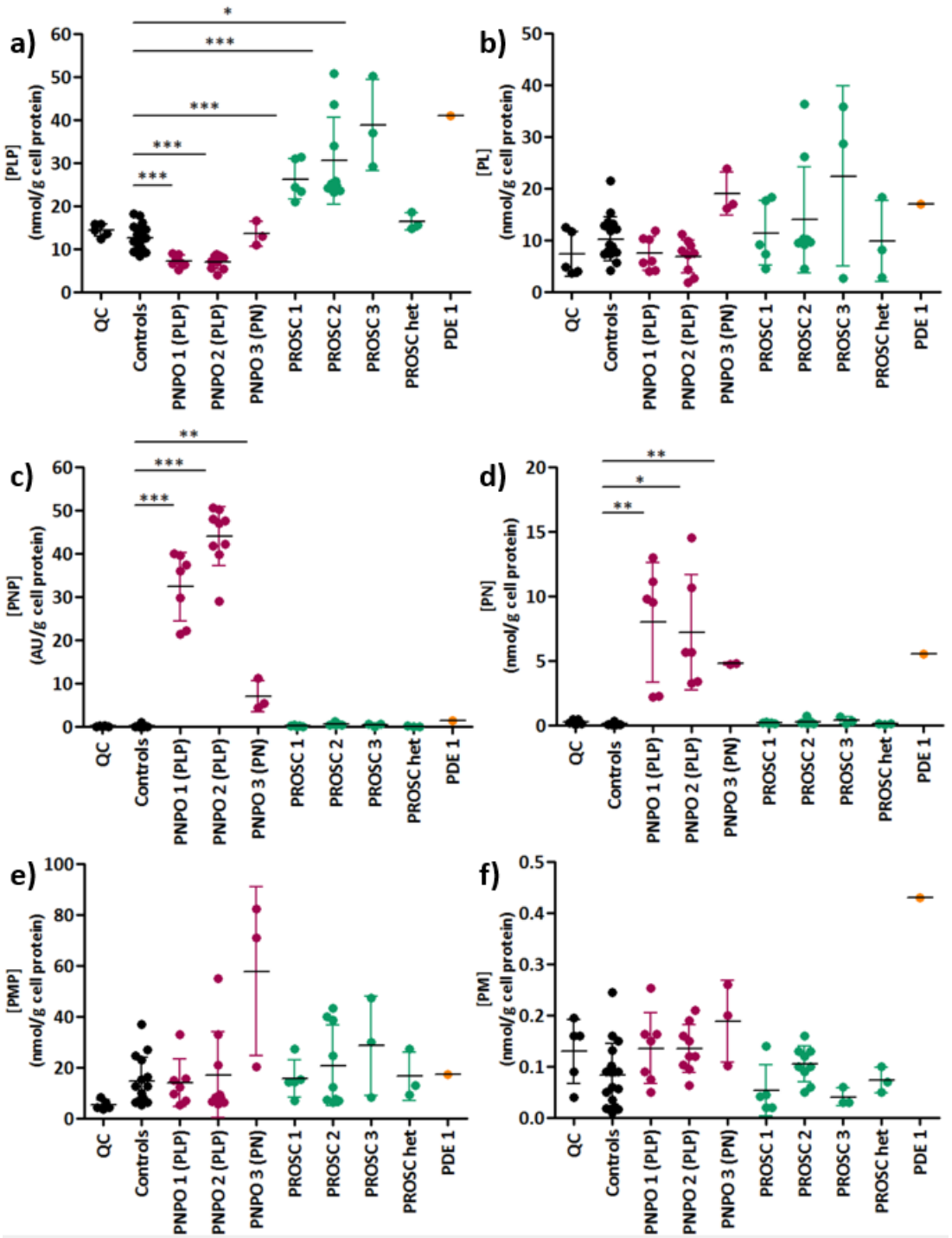
7.3.3 *Differences between PN- and PLP-responsive patients with PNPO deficiency*

To date, more than twenty disease-causing mutations in PNPO have been identified with nine having been associated with PN-responsiveness in twenty patients (Mills *et al.*, 2014; Plecko *et al.*, 2014; Jaeger *et al.*, 2016). Some, such as the p.Arg225His mutation present in the patient we have examined (PNPO 3), are associated exclusively with a response to PN with a switch to PLP resulting in status epilepticus in some patients (Plecko *et al.*, 2014). This mutation is commonly present as at least one of the pathogenic alleles in PN-responsive patients and affects a highly conserved region within the active site of PNPO (Musayev *et al.*, 2009). Indeed, the mutated residue is conserved across multiple species between humans and *E. coli*. Whilst residual activity associated with this mutation has been measured as being between 8% and undetectable, both

used artificial systems that may have underestimated activity due to the potential stabilisation of the mutant protein *in vivo* (Mills *et al.*, 2014; Plecko *et al.*, 2014).

The observed PN-response suggests that the p.Arg225His mutation does allow for some residual enzyme activity. Indeed, examinations of patient fibroblasts provided some evidence to support this hypothesis (Figure 7.3.4 and Figure 7.3.5). When grown in conditions of B_6 repletion, the PNPO-deficient fibroblasts did not show any significant differences in the concentrations of PL, PM or PMP compared to control cells. The latter was at first sight somewhat unexpected given that PM and PMP have been reported in plasma and CSF samples from patients with PNPO deficiency (Footitt *et al.*, 2013; Jaeger *et al.*, 2016). However, it can be explained by the fact that PN is the sole vitamin B_6 source within the cell culture medium. Indeed, all PNPO-deficient cell lines showed a significant increase in the concentrations of PN and PNP. The PN concentrations were elevated to the same extent in both the PLP- and PN-reponsive cells (2.5 - 15 nmol/g cell protein). However, the accumulation of PNP occurred to a greater extent in the PLP-responsive fibroblasts (PNPO 1 and 2; 20 - 50 AU/g cell protein) compared to those derived from the PN-responsive patient (PNPO 3; 4 - 11 AU/g cell protein). In these conditions of B_6 repletion, only the two PLP-responsive cell lines (PNPO 1 and 2) demonstrated a deficiency of PLP. This greater degree of PNP accumulation and the presence of PLP deficiency despite supplementation with concentrations of PN in excess of those required for growth indicates a greater dysfunction of the PNPO enzyme in the patients with PLP-responsive seizures (PNPO 1 and 2). These findings suggest that the p.Arg225His mutation allows for some flux through the pathway, in agreement with the findings published by (Mills *et al.*, 2014).

Figure 7.3.4: Comparison of the B_6 vitamer profiles of patient fibroblasts with that of controls grown in standard media. Media contains 971 nM pyridoxine hydrochloride. Each point represents the mean vitamer concentration of one flask of cells analysed in triplicate within one quantitation experiment. A quality control (QC) cell line was analysed in every experiment in order to confirm correct instrument functioning and sample preparation. The treatment response of each PNPO-deficient patient is shown in brackets. Error bars represent ± standard deviation. **(a)** PLP, pyridoxal 5'-phosphate; **(b)** PL, pyridoxal; **(c)** PNP, pyridoxine 5'-phosphate; **(d)** PN, pyridoxine; **(e)** PMP, pyridoxamine 5'-phosphate; **(f)** PM, pyridoxamine. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.



When fibroblasts were cultured in media depleted of vitamin B_6, the cells from PLP-responsive patients still had elevated levels of PN and PNP although concentrations were reduced between 6 - to 10-fold and 4- to 6-fold, respectively. However the accumulation of PN and PNP seen in PNPO 3 (PN-reponsive) under conditions of B_6 repletion was ameliorated when the fibroblasts were depleted of vitamin B_6. Under conditions of extreme vitamin B_6 deficiency, PNPO not only functions to generate sufficient PLP through the metabolism of PNP and PMP, but also to recycle the cofactor from PMP released by degraded apoenzymes (Musayev *et al.*, 2003). Fibroblasts from all patients showed an increase in PMP concentrations compared to the same cell lines grown in media supplemented with PN, indicative of increased recycling of PMP from apoenzymes when no exogenous B_6 source is present. However, in the PLP-responsive cell lines, the accumulation of PMP was more marked compared to the PNPO 3 (PN-reponsive) cell line indicating an impairment of this function. PM levels also appeared to be elevated in some patients (PNPO 2) but reduced in others (PNPO 3); this is likely due to analytical variability when measuring these extremely low analyte concentrations. Finally, compared to the control fibroblasts (6 - 16 nmol/g cell protein) the PLP concentrations in the PNPO-deficient cell lines were below (PNPO 2; 3 nmol/g cell protein) or at the lower boundary of the control range (PNPO 1 and 3; 6 nmol/g cell protein).

Direct quantitation of PNPO enzyme activity also provided evidence that the p.Arg225His mutation results in a higher residual enzyme activity than the mutations harboured by the two PLP-responsive PNPO patients (Figure 7.3.6). Fibroblasts from all three PNPO-deficient patients showed an identical reduction in d_2-PN concentration but a marginally increased accumulation of d_2-PNP relative to controls due to the action of PK and the lack of active PNPO to convert it into d_2-PLP, although this was not statistically significant. Control PNPO activity ranged from 50 - 300 pmol/hr/mg protein. The residual enzyme activity in the PN-responsive cells (11 pmol/hr/mg protein) was 6.7% of the mean activity detectable in controls and 20% that of the control cell line with the lowest activity, whereas it was undetectable in the cells from the PLP-responsive PNPO-deficient patients (Figure 7.3.6).

Although a degree of residual enzymatic function would account for the therapeutic response to PN, the reasons underlying the intolerance to PLP therapy remain unknown. PNPO is known to be strongly inhibited by free cellular PLP due to the toxic effects of its cellular accumulation (Choi *et al.*, 1987). Given that the p.Arg225His mutation is located within a site that can bind PLP, the mutant protein may suffer complete inhibition by the administration of PLP. It is possible that supplementation of these patients with much lower doses of PLP than conventionally required to treat patients with PNPO deficiency may be sufficient to correct the pathogenic

Figure 7.3.5: Comparison of the B_6 vitamer profiles of patient fibroblasts with that of controls grown in standard media and media depleted of pyridoxine. Dark green, standard media; light green, media depleted of pyridoxine. The treatment response of each PNPO-deficient patient is shown in brackets. The results of one experiment are shown where each patient was analysed in triplicate. Error bars represent ± standard deviation. **(a)** PLP, pyridoxal 5'-phosphate; **(b)** PL, pyridoxal; **(c)** PNP, pyridoxine 5'-phosphate; **(d)** PN, pyridoxine; **(e)** PMP, pyridoxamine 5'-phosphate; **(f)** PM, pyridoxamine.

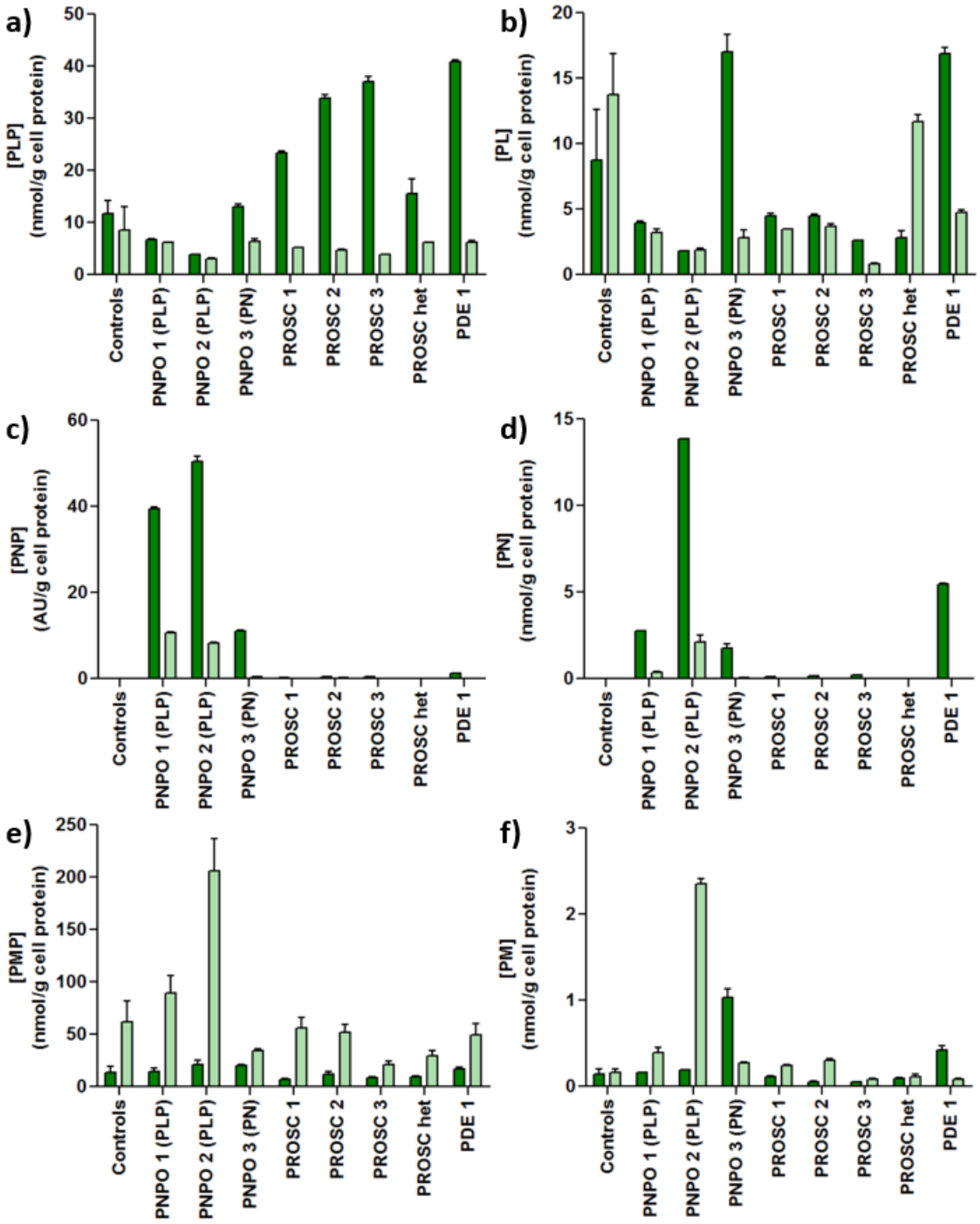
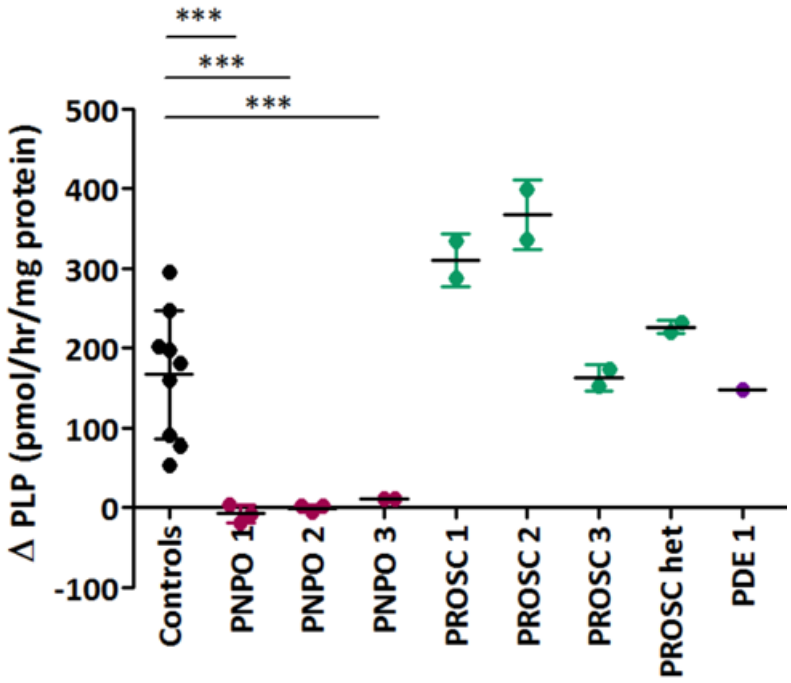


Figure 7.3.6: PNPO enzyme activity quantitation in patient fibroblasts. All cells from PNPO-deficient patients showed a significant ($p < 0.001$) reduction in enzyme activity.



deficiency of the active vitamer without encountering enzymatic inhibition. Further *in vivo* studies are required to investigate these possibilities.

Additional questions that remain include why some patients with identical mutations respond to PN and others do not and why there appears to be little correlation between *in vitro* residual activity and response to either PN or PLP. For example p.Arg225Cys (9% activity), p.Arg225His (8% activity) and p.Asp33Val (44% activity) are predominantly PN-responsive genotypes, whereas patients harbouring the p.Arg229Trp mutation (15% activity) have been predominantly treated with PLP (Mills *et al.*, 2014). The latter may be due to variability amongst assay methodologies as already noted (Mills *et al.*, 2014; Plecko *et al.*, 2014). In addition, *in vitro* measurements may not be representative of the interactions occurring *in vivo*. Environmental factors affecting the levels of PLP within the brain as well as the clinical condition of the affected neonate are also likely to influence the therapeutic response. These include polymorphisms within the *ALPL* gene encoding tissue non-specific alkaline phosphatase, dietary intake of vitamin B_6 and riboflavin, prematurity, breast feeding, age at the time of each treatment trial and the degree of any previous insults (e.g. poor condition requiring intubation at birth or prolonged status epilepticus) (Mills *et al.*, 2014). As well as influencing treatment response, these factors likely play a role in the overall clinical outcome of each patient making further research into these areas critical.

7.3.4 *Confirmation of PNPO deficiency in the background of only partially informative genetic analyses*

One of the patients examined in this study (PNPO 2) was clinically diagnosed as having PNPO deficiency due to her severe epilepsy beginning at five hours of life which was completely unresponsive to conventional antiepileptic drugs and pyridoxine, but a therapeutic breakthrough was achieved when PLP was administered at 20 days of age. A case report describing her clinical history has been published (Raimondi *et al.*, 2015). Whilst standard Sanger sequencing of the *PNPO* gene identified a single heterozygous frameshift mutation resulting in a premature stop codon (p.Val215Glyfs*14), it was not possible to identify a second mutation despite isolating mRNA from patient fibroblasts to analyse cDNA. A tentative diagnosis of PNPO deficiency could therefore only be given (Dr Philippa Mills, ICH).

Mass spectrometry-based analysis revealed that this patient not only had negligible PNPO enzyme activity (Figure 7.3.6), but also showed the characteristic deficiency of PLP, accumulation of both PN and PNP under conditions of B_6 depletion and repletion and accumulation of PMP under conditions of stress (Figure 7.3.4 and Figure 7.3.5). Indeed, in PNPO 2 these B_6 abnormalities were evident to a much greater extent than in PNPO 1 (the other PLP-responsive patient), potentially suggesting an even more severe functional impairment of PNPO which would be inconsistent with heterozygote mutation status. Thus, these investigations confirm the diagnosis of PNPO deficiency in this patient. Given that cDNA analysis was unremarkable and the wild-type allele could be visualised when the frameshift mutation was identified by Sanger sequencing, the possibility of an exonic or whole gene deletion could be excluded. Instead, the second pathogenic mutation is likely to reside in an intronic or regulatory genomic region.

7.3.5 *Abnormalities in a patient with antiquitin deficiency*

The B_6 vitamer profile of one patient with genetically confirmed antiquitin deficiency was examined (Figure 7.3.4 and Figure 7.3.5). Concentrations of PL, PNP and PMP were found to be within the control range. A mildly elevated concentration of PM was identified in cells supplemented with PN, however, similarly to the cell lines from other patients, this finding was not recapitulated in fibroblasts grown in depleted media. This variability observed across these experiments is likely due to analytical variability when measuring these extremely low analyte concentrations. In contrast to fibroblasts from patients with PNPO deficiency who had PLP levels comparable to or below that of controls, these cells showed elevated PLP levels under conditions of PN supplementation. These levels were 3-fold higher than those seen in control cells

and comparable to those identified in PROSC-deficient patients (discussed further in Section 7.3.7). The concentration of PN was also elevated 55-fold relative to control cells, similar to the accumulation observed in PNPO-deficient cells (22- to 145-fold elevated above controls). However, when the antiquitin-deficient cells were cultured in media depleted of B_6, no vitamer abnormalities were detected.

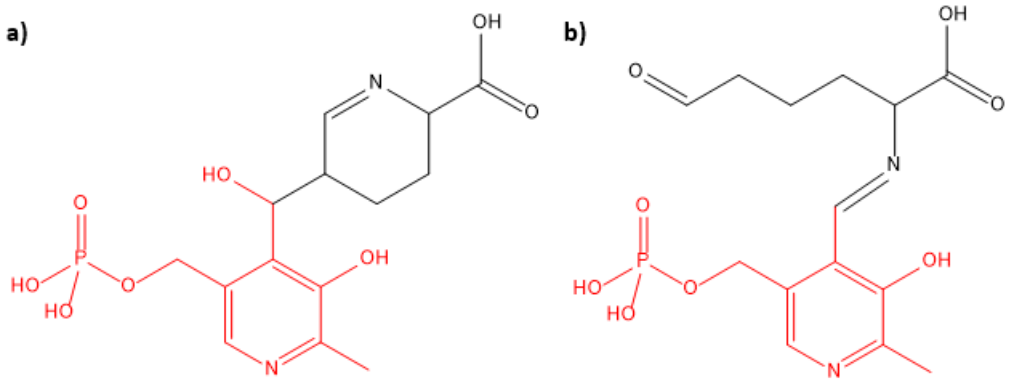
Studies examining the concentrations of the B_6 vitamers in the plasma of patients with PDE whilst on PN supplementation have revealed that levels of PLP, PL and pyridoxic acid (PA) are strongly elevated in the vast majority of cases (Footitt *et al.*, 2013; Mathis *et al.*, 2016). Concentrations of PN have also been documented as markedly elevated in a sub-group (71%) of patients (Mathis *et al.*, 2016). However, the vitamer profiles in these patients did not differ from those of other patients on PN-treatment, including those affected by epileptic encephalopathy or hypophosphatasia. Therefore, these findings have been suggested to reflect the intake of supra-physiological doses of vitamin B_6. Similarly, the variability in PN levels has been postulated to be due to the interval between sampling and the intake of the last PN dose.

PA is not detectable in fibroblasts as the enzymes which function to convert PL to PA for urinary excretion (aldehyde oxidase or β-NAD dehydrogenase) are not expressed in these cells. PL is also not elevated in the fibroblasts, likely due to differences in the identity, distribution and activity of phosphatases that acts to dephosphorylate PLP to from PL between plasma and fibroblasts. In contrast, the elevated PLP and PN reported in plasma was recapitulated in the fibroblasts from PDE 1. As suggested by Mathis *et al.* (2016), the elevated PN may simply be due to a shorter interval between the last media change (containing excess PN) and harvesting the fibroblasts. Alternatively, the PDE cells may increase the uptake of PN from the media to correct for the physiological deficiency of PLP and to maintain adequate function of B_6-dependent enzymes.

As discussed in Section 7.1.2, endogenous PLP undergoes a Knoevenagel condensation reaction with the Δ^1-piperideine-6-carboxylic acid (P6C) that accumulates in patients with antiquitin deficiency (Figure 7.3.7a) (Mills *et al.*, 2006). This chemical reaction consists of the nucleophilic addition of a carbanion molecule to a carbonyl group followed by a dehydration reaction in which a molecule of water is lost. It could also be hypothesised that PLP may react with the equilibrium partner α-aminoadipic semialdehyde (α-AASA) through a Schiff base interaction (Figure 7.3.7b). It is known that PLP forms a Schiff base with lysine residues within proteins (Toney, 2005). Therefore, given the structural similarities between lysine and α-AASA, it is possible that the latter could partake in a similar interaction with PLP. Indeed, as part of the study identifying the pathogenic aetiology of PDE (Mills *et al.*, 2006), urine samples from patients were analysed for the presence of this complex. The complex was not identified (unpublished data; Dr Philippa

Mills, ICH, UK), however the conditions used to prepare the urine samples were different to those used in this chapter. Schiff bases are known to be hydrolysed under acidic conditions; indeed, these conditions are used for the measurement of total PLP in biofluids (Albersen *et al.*, 2015; Footitt *et al.*, 2013). However, bonds formed by Knoevenagel condensations are less labile and unlikely to be broken by acid hydrolysis. Thus, the increased PLP may be derived from the accumulated PLP-AASA and/or PLP-P6C complexes. An alternative hypothesis is that the physiological deficiency of PLP in PDE cells resulted in a compensatory upregulation of PNPO enzyme activity resulting in apparently greater concentrations of PLP upon TCA precipitation of the samples. However, this hypothesis was not supported by the direct quantitation of PNPO enzyme activity in PDE cells, which was found to be within the control range (Figure 7.3.6). Finally, it is important to note that these cells were only analysed on one occasion and thus statistical analysis could not be performed. Indeed, these results only represent the findings in one patient with antiquitin deficiency and caution should be taken when drawing conclusions. Further work analysing a larger cohort of patients would be beneficial to investigate the hypotheses presented above.

Figure 7.3.7: Undesired reactions of PLP in patients with antiquitin deficiency. (a) PLP (red) can react with P6C (black) by a Knoevenagel condensation reaction. **(b)** One hypothesis is that PLP (red) may also react with α-AASA (black) forming a Schiff base.



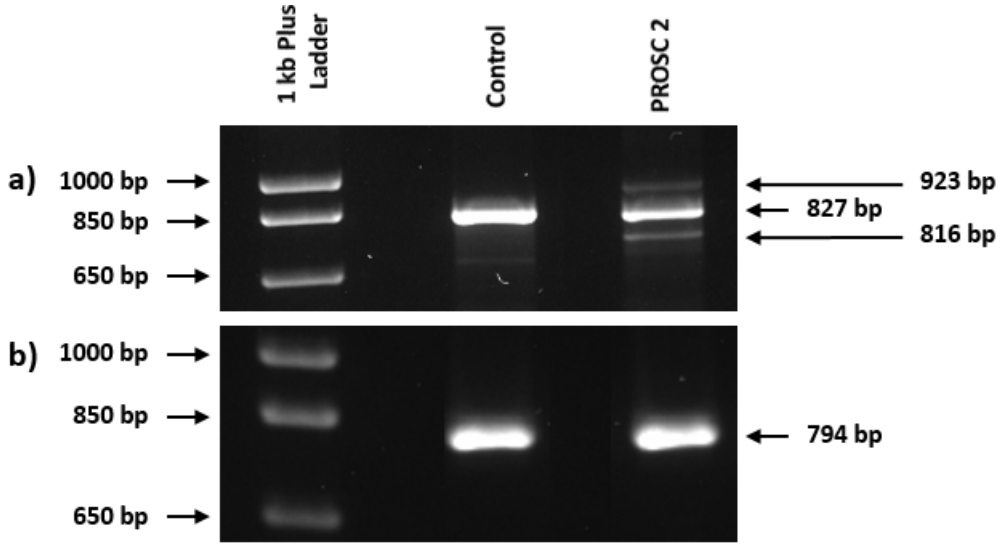
7.3.6 *Investigation of the effects of mutations in PROSC on protein transcription and translation*

Three patients with genetically confirmed PROSC-deficiency, a novel disorder of vitamin B_6 metabolism, were examined; one patient had a homozygous missense mutation (PROSC 1; p.Leu175Pro), one was compound heterozygous for two splice site mutations (PROSC 2; c.207+1G>A and c.320-2A>G) and one had a homozygous nonsense mutation (PROSC 3; p.Ser78*) (Table 7.2.1). The effect of each mutation on transcription and translation of the

PROSC gene product were characterised as were the effects of the two splice site mutations in PROSC 2 on the cDNA product.

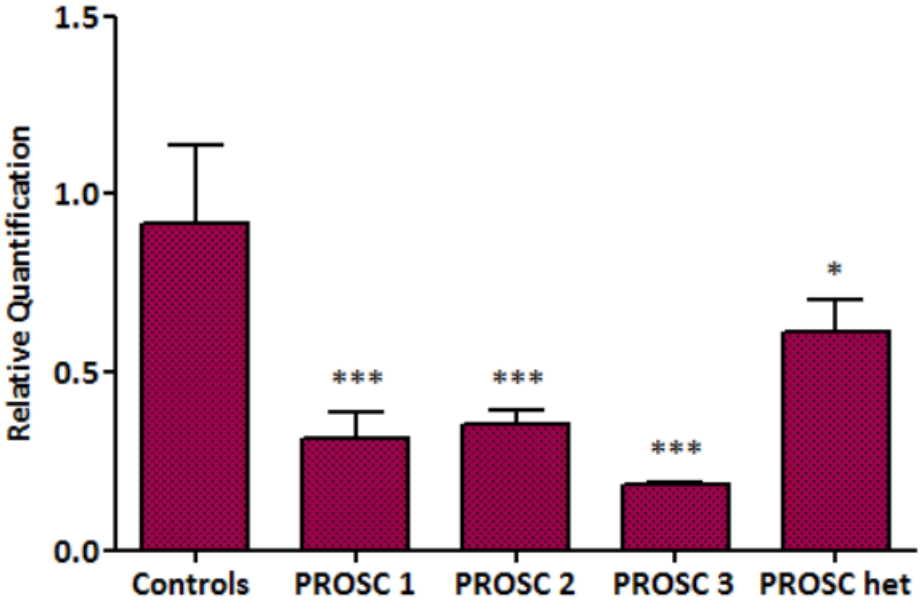
cDNA was generated from RNA extracted from fibroblasts and the *PROSC* gene was subsequently amplified and sequenced (Section 2.8). A housekeeping gene, hypoxanthine phosphoribosyltransferase 1 (*HPRT1*), was also amplified to check the integrity of the RNA and corresponding cDNA. A single fragment corresponding to the size expected for the region of *HPRT1* amplified was identified in all samples, indicating that the reverse-transcription of total RNA to generate cDNA had been successful. cDNA amplification of *PROSC* from control cells resulted in a 846 bp fragment which corresponds to the predicted size of the wildtype *PROSC* gene in control fibroblasts. In contrast, three DNA fragments of different sizes were identified in PROSC 2 (Figure 7.3.8). In order to investigate these further, each fragment was excised and cloned into the TOPO 2.1 vector prior to Sanger sequencing (Sections 2.16.11 and 2.16.8). Real-time polymerase chain reaction (qRT-PCR) (Section 2.9) and Western blotting (Section 2.10) were also carried out to investigate the effects of mutations on gene expression and protein translation (Figures 7.3.9 and 7.3.10). Taken together, these investigations revealed some intuitive and some unexpected findings.

Figure 7.3.8: cDNA generated from RNA extracted from fibroblasts. (a) *PROSC* and **(b)** *HPRT1* cDNA amplified from total RNA extracted from fibroblasts.



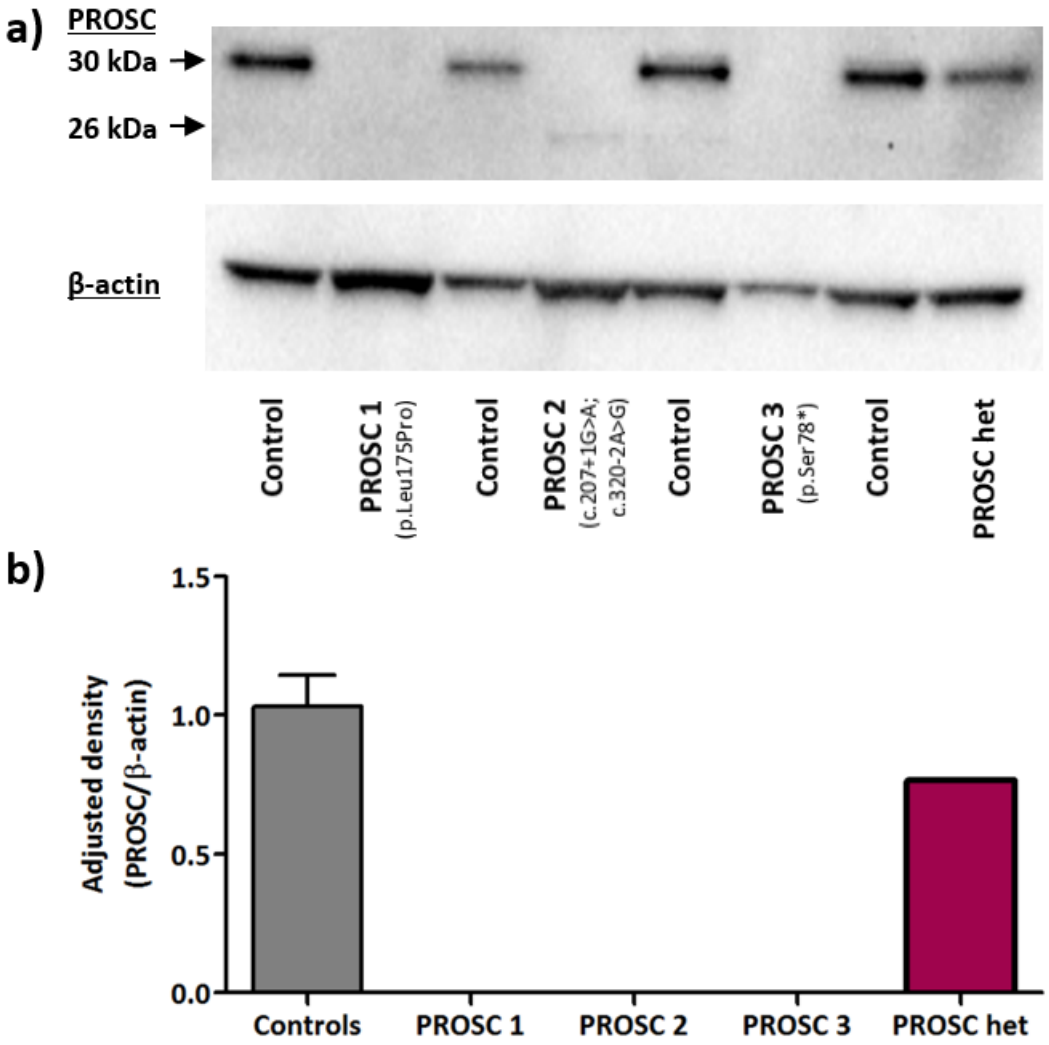
Firstly, the presence of the nonsense mutation (p.Ser78*) in PROSC 3 would be expected to result in nonsense-mediated decay and indeed a significant, approximately 5-fold, reduction in PROSC mRNA expression was demonstrated (Figures 7.3.9). PROSC protein was also undetectable (Figure 7.3.10). As expected as a carrier of this mutation, the mother of this patient showed intermediate levels of mRNA and protein expression.

Figure 7.3.9: qRT-PCR of PROSC mRNA in fibroblasts. Fibroblasts were grown in standard media containing 971 nM pyridoxine hydrochloride. Relative quantitation was carried out using the comparative C_T ($2^{-\Delta\Delta CT}$) method with β-actin and GAPDH as reference genes and analysed in quadruplicate. Statistical analysis was performed using Welch's unequal variances t-test. Data is presented as means ± s.d. All patient cells showed a significant ($p<0.0001$) reduction in PROSC mRNA expression, with PROSC het cells having an intermediate reduction ($p=0.0110$). Controls, n=8; patients, n=3.



Analysis of cDNA extracted from PROSC2 revealed that c.207+1G>A and c.320-2A>G affect DNA splicing and result in decreased mRNA expression (Figures 7.3.9). The c.207+1G>A mutation was found to cause not only the inclusion of the second intron which would be predicted to result in nonsense-mediated decay due to the introduction of a premature stop codon (p.Val70Ilefs*6; 923bp cDNA fragment), but also the skipping of the second exon and an in-frame deletion of 36 amino acids (p.Asp34_Tyr69del; 816bp cDNA fragment) which corresponds to a protein approximately 26 kDa in size (Figure 7.3.10). Western blotting was carried out using the HPA023646 antibody (Sigma) which binds to a stretch of 75 amino acids encompassing exon five and approximately half of exon six of the *PROSC* gene (Section 2.10). A protein product corresponding to the p.Val70Ilefs*6 mutation was not detected as the protein is truncated upstream of the antibody binding site (Figure 7.3.10). Indeed, if this mutation results in nonsense-mediated decay as predicted, this protein product would not be detectable using Western blotting regardless of the antibody epitope. Splice site mutations can lead to exon skipping, intron retention and insertions or deletions of variable size due to the use of cryptic splice sites. In the classic model of spliceosome assembly, only a single intron was thought to be removed at a time in a progressive 5' to 3' direction along the mRNA (Lang and Spritz, 1983). If this was the case, it would be expected that mutations of the 5' splice donor site would result

Figure 7.3.10: Western blot of the PROSC protein in patient fibroblast lysates. **(a)** Western blot of PROSC protein alongside a β-actin loading control. **(b)** Densitometry quantification of the PROSC protein was carried out using ImageJ software. A detailed description of the method used can be found at www.lukemiller.org/index.php/2010/11/analyzing-gels-and-western-blots-with-image-j/.



exclusively in retention of the following intron. However, this is now known not to be the case; indeed many mutations of the canonical +1G residue have been reported to causing skipping of the preceding exon in other disorders including sialidosis, congenital afibrinogenemia and phenylketonuria (Penzel *et al.*, 2001; Attanasio *et al.*, 2003; Marvit *et al.*, 1987). This impact on transcription is determined by many factors including the size of the mutated intron, the order of intron removal within each specific gene and surrounding genomic elements (Takahara *et al.*, 2002). Although most splice site mutations only produce a single outcome, there are many reports of a single mutation generating multiple products. This can be attributed to "major" and "minor" splicing pathways in which the order that exons are removed differs (Schwarze *et al.*, 1999).

Sequencing of the transcript corresponding to the c.320-2A>G mutation also revealed somewhat unexpected results. Instead of causing the skipping of exon five, only the first ten exonic amino acids are deleted due to the use of a cryptic splice site (p.Ala107_Thr116del; 827bp cDNA framgent). A protein product corresponding to the p.Ala107_Thr116 mutation was not detected due to the deletion's location within the antibody binding site (Figure 7.3.10). Western blotting using an alternative antibody binding to a stretch of 16 amino acids within exon two was also performed (SAB1105316, Sigma), however a specific band could not be obtained (data not shown). These cDNA and protein findings illustrate the importance of experimental confirmation of the effects of mutations predicted to affect correct splicing and caution against the over-reliance on *in silico* prediction tools (Baralle and Baralle, 2005).

The missense mutation (p.Leu175Pro) identified in a homozygous state in PROSC 1 affects a residue conserved across mammalian species, but also in Yggs (bacterial homolog) and YBL036C (yeast ortholog). It also lies within the predicted PLP-binding barrel domain similar to that at the N-terminus of bacterial alanine racemase and eukaryotic ornithine decarboxylase (Ito *et al.*, 2013). Unexpectedly, the p.Leu175Pro mutation significantly decreases PROSC mRNA expression (Figures 7.3.9) and results in no detectable PROSC protein in patient fibroblasts (Figure 7.3.10). Although the immunogen sequence of the antibody encompasses the mutated residue, this would not necessarily be expected to affect antibody binding. The binding affinity of an antibody for its corresponding epitope can be variable (Saper, 2009). Indeed, polyclonal antibodies such as the one utilised in this study, are more tolerant of minor antigen changes such as polymorphisms or missense mutations than monoclonal antibodies (Lipman *et al.*, 2005). Furthermore, the protein findings were mirrored at the mRNA level, suggesting that this decreased expression of PROSC is not an artefact of the experimental methodology. Although unusual, missense mutations have been reported to reduce mRNA expression (Nguyen *et al.*, 2011). Since point mutations are unlikely to induce a dramatic reduction in mRNA stability and result in reduced levels, another possible explanation is that the mutation increases the affinity for the binding of a microRNA (miRNA) or creates an miRNA binding site which causes RNA degredation. MicroRNAs are small non-coding regulatory RNA molecules which function to regulate the expression of complementary mRNAs. miRNAs typically silence mRNA molecules by either (i) cleavage of mRNA strands, (ii) destabilising the mRNA by shortening the poly(A) tail or (iii) reducing translation efficiency by ribosomes (Fabian *et al.*, 2010). The p.Leu175Pro mutation lies within a predicted binding site of hsa-miR-202-5p miRNA (Paraskevopoulou *et al.*, 2013). Very little is known about this miRNA except that it is highly expressed in Sertoli cells and plays a role in testis development (Wainwright *et al.*, 2013). The significance of this is uncertain but it is

possible that the missense mutation present in PROSC 1 affects mRNA stability through the modification of miRNA binding.

An additional factor to consider at the protein level is the known propensity of the introduction of proline residues to disrupt correct secondary and tertiary structure formation, particularly within cytoplasmic proteins (Li *et al.*, 1996). JPred4, a tool for secondary structure prediction (Drozdetskiy *et al.*, 2015), suggests that the p.Leu175 residue is located immediately prior to a β-sheet motif, at the end of a sequence of five amino acids connecting it to a preceding α-helix. Proline disfavours the β-sheet structure due to its incompatible dihedral angle and lack of one potential hydrogen bond donor (Li *et al.*, 1996). Perhaps more importantly, because of its cyclic structure, proline induces a turn or bend in the structure; this often functions to bring α-helices and β-sheets into close proximity to allow for the formation of more complex tertiary and quaternary conformations. Within this sequence of five amino acids, a proline residue is already present two amino acids upstream of our mutated leucine. Therefore, it may be expected that the two proline residues in close proximity result in a marked instability of the encoded protein due to an aberrant tertiary structural conformation.

7.3.7 *PROSC deficiency affects PLP homeostasis*

The analyses described above demonstrated that all mutations present in the three PROSC-deficient patients cause decreased expression of each mutant mRNA and protein. The most consistent phenotypic feature in these patients was an early-onset seizure disorder which responded to treatment with either PN or PLP. To exclude PNPO deficiency, the enzyme activity of PNPO in fibroblasts from the PROSC-deficient patients was quantified (Figure 7.3.6). PROSC 1 and PROSC 2 had activity of 310 and 367 pmol/hr/mg protein, respectively. These values were marginally above the reference range of 50 - 300 pmol/hr/mg protein, although the increase was not statistically significant and likely reflects the relatively small number of control fibroblast lines available for analysis. The PNPO activity in PROSC 3 and PROSC het cells was within the control range.

In addition, as part of their diagnostic work-up, three patients had their B_6 vitamers analysed in plasma whilst receiving supplementation. Whilst these patients had high levels of PLP, unlike patients with PNPO deficiency (Footitt *et al.*, 2013), they do not accumulate PN, PM, PNP and PMP in plasma (unpublished data; Dr Philippa Mills, ICH, UK). Although not as high as those seen in patients with hypophosphatasia due to mutations in tissue non-specific alkaline phosphatase (Whyte *et al.*, 1988), PLP levels were 4 - 15 times higher than those reported for

PNPO and PDE patients receiving comparable supraphysiological doses of vitamin B_6 (Table 7.3.1).

Table 7.3.1: Plasma PLP levels in patients with vitamin B6-responsive disorders being treated with variable doses of either pyridoxine (PN) or pyridoxal 5'-phosphate (PLP). * Data taken from Footitt *et al.* (2013).

Patient	Vitamer dosage	Plasma PLP levels (nmol/L)
Control range	None	46 - 321
Treatment with pyridoxine (PN)		
Pyridoxine-dependent epilepsy (PDE)*	100 mg twice a day	11 - 604
Undiagnosed seizure disorder*	100 mg twice a day	877
PROSC deficiency	200 mg twice a day	2600
Treatment with pyridoxal phosphate (PLP)		
PNPO deficiency*	30 mg/kg/day	580 - 633
PLP-responsive seizures*	30 mg/kg/day	710
Partially PLP-responsive seizures*	30 mg/kg/day	478
PROSC deficiency	45 mg/kg/day	2166
PROSC deficiency	20 mg/kg/day	2750

B_6 vitamer analysis was then carried out in patient fibroblasts to determine whether the findings in biofluids (i.e. plasma) would be recapitulated intracellularly. Similarly to the profiles in plasma, elevated levels of PLP were identified in all PROSC-deficient fibroblasts when cultured in media containing excess concentrations of PN (Figure 7.3.4). However, the accumulation was not as striking as that observed in plasma, only being between 2- and 3-fold higher than control fibroblasts cultured under identical conditions. The PLP concentrations in the PROSC het cells which harbour the p.Ser78* mutation in a heterozygous state (15 - 18 nmol/g cell protein) were approximately half of those in the PROSC 3 cells which are homozygous for the same mutation (30 - 50 nmol/g cell protein). These intermediate concentrations support the hypothesis that dysfunction of the PROSC protein results in dysregulation of intracellular PLP concentrations. The concentration of PL was also marginally higher in patient cells and, whilst not statistically significant, it likely reflects endogenous phosphatase activity on elevated PLP levels. All other vitamer concentrations remained within control ranges, mirroring the findings in plasma. However, when cells were cultured under conditions of B_6 depletion, levels of PLP were maintained between 4 - 6 nmol/g cell protein similar to those seen in control cells (Figure 7.3.5). This, along with the normalisation of the PLP deficiency in PNPO-deficient cells (Section 7.3.3)

suggests that mechanisms exist to tightly regulate PLP in conditions where vitamin B_6 is scarce. Indeed, no other vitamer abnormalities were identified under these conditions.

In addition to the plasma samples that had been analysed previously (unpublished data; Dr Philippa Mills, ICH, UK) and the cultured fibroblasts from the three patients described above, two cerebrospinal fluid (CSF) samples from PROSC 2 (c.207+1G>A and c.320-2A>G) whilst off-treatment were available for analysis. In order to check the validity of the UPLC-MS/MS method (Section 2.11) for the quantitation of the B_6 vitamers CSF, seven samples from the Neurometabolic Unit, National Hospital for Neurology and Neurosurgery where PLP had been measured were analysed (Table 7.3.2). The results for PLP were found to agree with those from the previous analyses. The concentrations of the other vitamers were compared to the reference ranges reported by Albersen *et al.* (2015), which is the only study to report reference ranges for all vitamers (with the exception of PNP) in children.

PROSC 2 was receiving 40 mg/kg/day of PLP for seizure control but underwent a withdrawal trial at 17 months of age. The dosage was tapered over a two week period and two CSF samples were taken 48 hours after this cessation. PL concentrations were slightly below the reference range reported by (Albersen *et al.*, 2015). In contrast, PLP levels were markedly decreased in both samples at 5.5 and 6.0 nmol/L (ref: 11 - 34) (Table 7.3.2). Combined with the patients' dramatic seizure resolution in response to vitamin B_6 treatment, this suggests a cerebral PLP deficiency and supports the hypothesis that PROSC is involved in vitamin B_6 homeostasis.

Table 7.3.2: CSF B_6 vitamer profiles of a PROSC patient (not on B_6 supplementation) and of patients who have been investigated for possible PLP deficiency or neurotransmitter abnormalities. All concentrations are expressed in nmol/L with the exception of PNP which is stated in "concentration units". * Albersen *et al.* (2015), ** Patient was receiving 40 mg/kg/day which was then tapered and eventually stopped over a two week period, CSF was taken 48 hours after this cessation. [1], the patients on PN supplementation all had PL concentrations that were greater than the 200 nmol/L upper limit of quantitation. n/a, not applicable; nd, not detected; nm, not measured; PLP, pyridoxal 5'-phosphate; PL, pyridoxal; PA, 4-pyridoxic acid; PN, pyridoxine; PNP, pyridoxine 5'-phosphate; PMP, pyridoxamine 5'-phosphate; PM, pyridoxamine; 5-HIAA; 5-hydroxyindoleacetic acid; 5-MTHF, 5-methyltetrahydrofolate.

Patient	Age	Supplementation	Neurotransmitter phenotype	Clinical phenotype	PLP	PL	PA	PN	PNP	PMP	PM
Reference range*	1y - 18y	None	n/a	n/a	11 - 34	16 - 56	<0.09 - 3	nd	nm	nd	0.3 - 0.9
PROSC 2	17m	48h after PLP cessation**	nd	Neonatal-onset seizures	**6.0, 5.5**	15.1, 14.9	nd	0.05, 0.08	0.92, 0.82	0.88, 1.05	0.13, 0.18
C8	13m	None	Marginally raised 5-HIAA, PLP within ref range	Near respiratory arrest, dysmorphic features, global developmental delay	20.9	28.2	nd	0.04	1.52	1.05	0.20
C3	15m	None	Neurotransmitters and PLP within ref ranges	Developmental impairment	23.3	23.5	nd	0.05	0.84	0.88	0.21
C6	17m	None	Elevated 5-HIAA and PLP within ref range	Acute-on-chronic neurodegeneration	25.6	41.1	nd	0.08	0.56	0.53	0.27
A8	17m	PN	Low 5-MTHF, high PLP, interfering peak on PLP chromatogram	Seizures, cardiac arrest	**81.2**	**> 200**[1]	nd	**5.87**	1.00	1.58	0.66

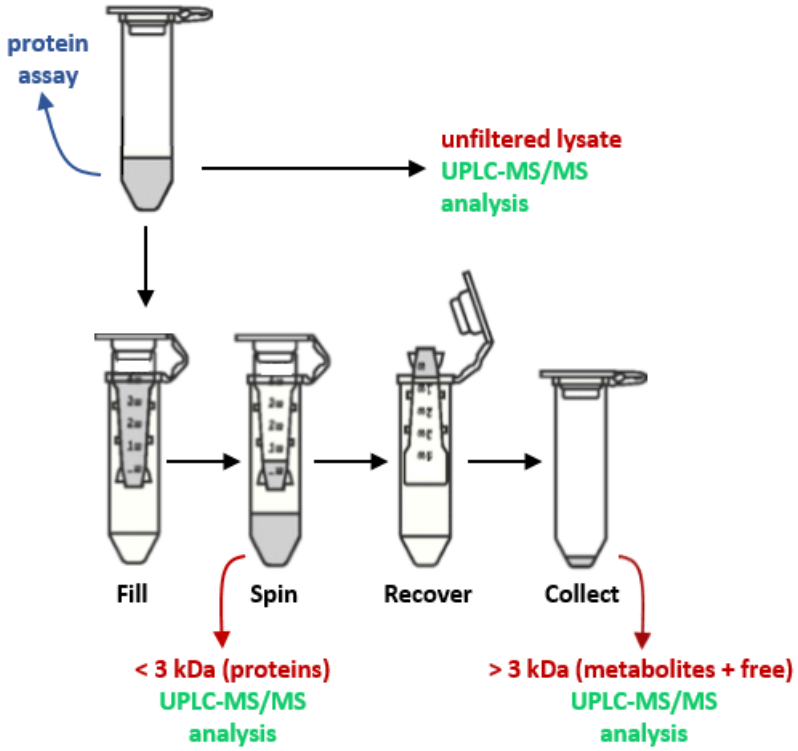
A2	16m	PN	Elevated PLP	Pyridoxine-dependent epilepsy	**82.9**	$\gg$ **200**[1]	nd	**41.53**	0.98	1.41	2.22
A7	14m	None	Undetectable PLP, marginally raised 5-HIAA	Seizures, developmental delay	**1.06**	**91.8**	nd	0.22	0.66	1.05	0.44
A1	18m	PN	Elevated PLP	B_6-dependent epilepsy	**75.7**	$\gg$ **200**[1]	nd	$\gg$ **200**	0.98	1.05	4.92

7.3.8 *Characterisation of potential pathogenic mechanisms underlying PROSC deficiency*

The elevated PLP concentrations observed in the plasma of patients on vitamin B_6 supplementation and in fibroblasts grown in conditions of PN excess, combined with the CSF PLP deficiency observed in PROSC 2 prior to treatment initiation suggested that pathogenesis in PROSC deficiency could be due to abnormal transport of PLP into the brain or dysfunction of PLP homeostasis. PROSC is a soluble cytoplasmic protein and has a binding site which binds PLP through a Schiff base linkage but does not affect the quaternary structure of the protein upon binding (Prunetti *et al.*, 2016). Its structural properties would therefore not be consistent with a transmembrane protein which functions to facilitate the transport of PLP into the brain. In addition, the lack of a conformational change upon the binding of PLP does not support a catalytic function of PROSC; rather, it suggests that the protein may simply bind PLP and act as a carrier protein. Indeed, it is possible that the normal function of PROSC is to carry newly-synthesised PLP from pyridoxal kinase and PNPO to the B_6-dependent enzymes that require it as a cofactor, thereby protecting PLP from unwanted reactions in the cytosol. Dysfunction of PROSC may then be expected to result in uncontrolled reactions of PLP with proteins, amino acids and metabolites such as cysteine (Terzuoli *et al.*, 1998) and carnosine (Vistoli *et al.*, 2013), thus rendering it inactive as a cofactor. If this was indeed the case, it would be expected that PLP would not be preferentially bound to proteins or metabolites. In order to investigate whether this was the case, cell lysates were fractionated using a 3 kDa cut-off filter. Each fraction containing proteins ($>$ 3 kDa) and metabolites ($<$ 3 kDa) was then prepared and analysed. The method developed in this chapter uses trichloroacetic acid to precipitate the proteins within the matrix being analysed. However, this also functions to break any Schiff bonds binding PLP to proteins or other molecules. Therefore, although it was possible to discriminate between protein- and metabolite-bound PLP, the determination of the proportion of free PLP within cells was not possible and was included in the $<$ 3kDa fraction (Figure 7.3.11).

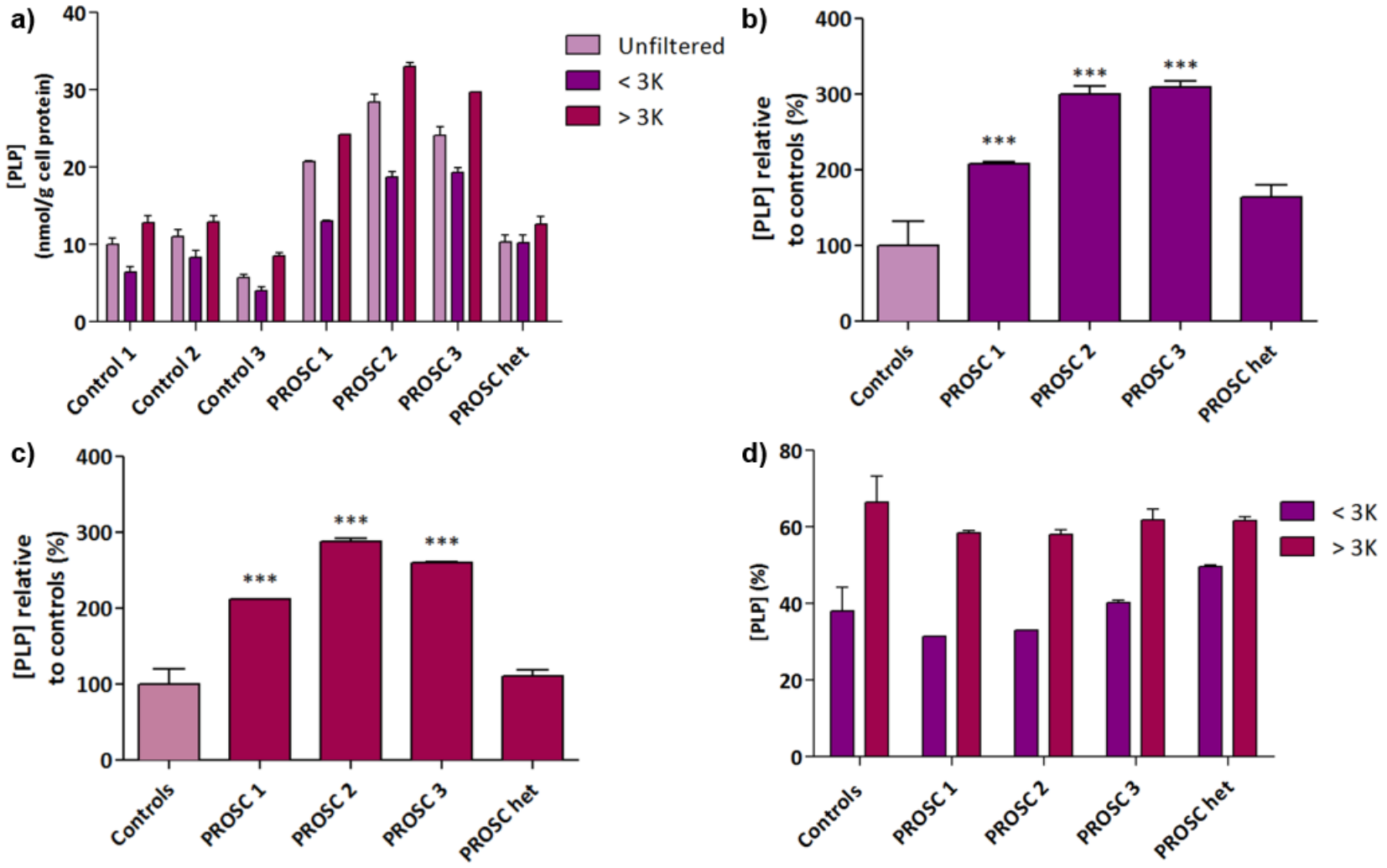
As observed in the whole cell lysate vitamer analysis (Section 7.3.7), all PROSC-deficient cell lines showed an accumulation of PLP when compared to controls (Figure 7.3.12a). The concentration of PLP in the $<$ 3 kDa and $>$ 3 kDa fractions should equate to the total concentration in the unfiltered sample. However, as illustrated in Figure 7.3.12a this was not the case. This was because each fraction could not be corrected for the protein concentration in the respective fraction as no protein was detectable in the $<$ 3 kDa solution. Therefore, correction of the PLP concentration in these samples for a negligible amount of cell protein would result in an extremely elevated artefactual quantification. To negate this problem, the PLP concentrations in each fraction were corrected for the total protein concentration in the unfiltered sample (Figure 7.3.11);

Figure 7.3.11: Preparation of cell lysate fractions from control and PROSC-deficient fibroblasts. Fibroblasts were harvested, lysed and the protein concentration of the unfiltered lysate was determined using a protein assay. The unfiltered, < 3 kDa and > 3 kDa fractions were analysed using UPLC-MS/MS before being corrected for the protein concentration of the unfiltered sample.



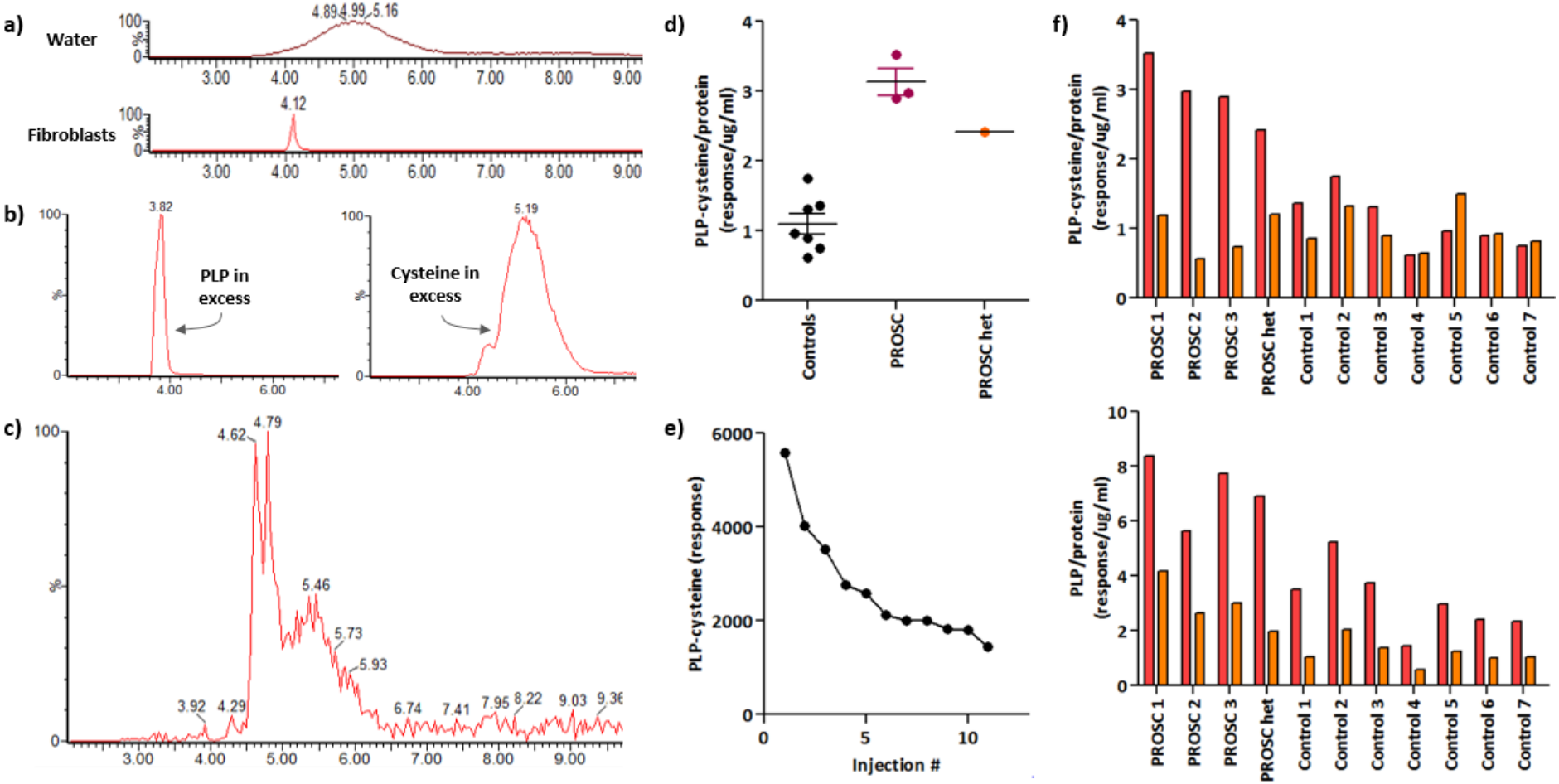
this allowed each patient to be directly compared. The concentrations of PLP within both the < 3 kDa and > 3 kDa fractions were significantly increased in all PROSC-deficient cell lines ($p < 0.001$) when compared to that of control fibroblasts (Figure 7.3.12a-c). When considering the uncorrected PLP concentrations, approximately 60% of the vitamer was detected in the > 3kDa fraction suggesting that it was bound to proteins and other large molecules (Figure 7.3.12d). This distribution in the PROSC-deficient fibroblasts was similar to that seen in control cells.

Figure 7.3.12: Distribution of PLP in fibroblast cell lysate fractions from PROSC patients and controls. **(a)** Concentrations of PLP in each fraction corrected for the protein concentration of the unfiltered sample. **(b)** and **(c)** Concentration of PLP in each patient compared to controls in the < 3kDa and >3kDa fraction, respectively. **(d)** Distribution of PLP between > 3kDa and < 3kDa fractions.All patients had significantly increased PLP concentrations ($p < 0.001$) relative to controls (n=3).



The fact that PLP concentrations were found to be elevated in both low and high molecular weight fractions, suggests that the supra-physiological cellular concentrations of PLP react non-specifically with a range of intracellular molecules. One of the molecules that PLP is known to spontaneously react with is cysteine, forming a PLP-cysteine thiazolidine molecule (Buell and Hansen, 1960; Liu *et al.*, 2013). In order to investigate whether this was occurring in patients with PROSC deficiency, mass spectrometry-based methods were established to detect this molecule in fibroblasts. The formation of this thiazolidine complex has been reported to occur rapidly upon the incubation of a solution containing PLP and cysteine, being completed within 30 minutes at room temperature and 10 minutes at 45°C (Głowacki *et al.*, 2016). However, various factors affect the efficiency of the reaction including molar ratio of PLP:cysteine, pH and temperature. A solution of the PLP-cysteine thiazolidine complex was prepared (Section 7.2) and analysed before being spiked into a matrix of fibroblast cell lysate. The retention time and peak shape of the thiazolidine compound was very different in an aqueous solution, where a broad peak was eluted between 4 - 6 minutes, compared to that spiked into a matrix of cell lysate where a single sharp peak was observed with a retention time of 4.12 minutes (Figure 7.3.13a). Indeed, similar differences in chromatography were also observed dependent on whether PLP or cysteine was in excess within the reaction mix, with conditions of PLP excess resulting in superior chromatography (Figure 7.3.13b).

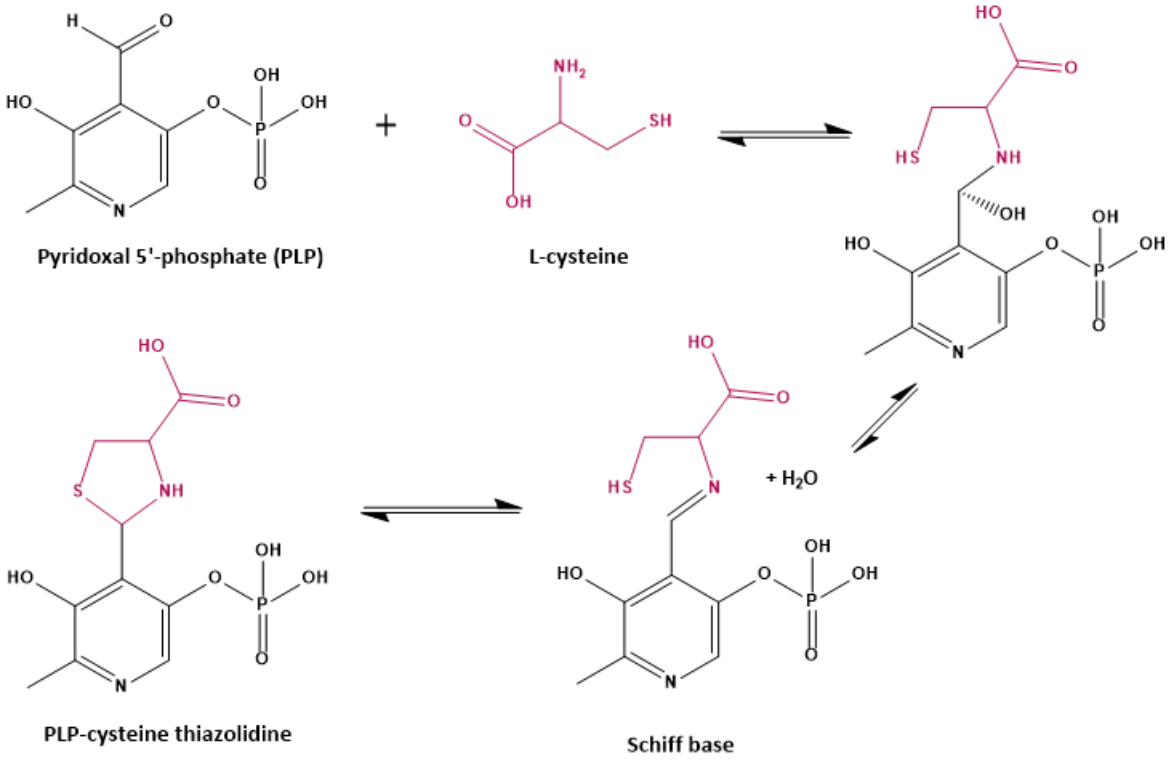
Figure 7.3.13: Experiments investigating the presence of PLP-cysteine conjugates within patient fibroblasts. **(a)** Chromatography of the PLP-cysteine conjugate in water and spiked into fibroblast cell lysate. **(b)** Differing chromatographic peak shape depending on the ratio of PLP:cysteine in solution. **(c)** Endogenous PLP-cysteine in fibroblasts show co-eluting peaks when the cells are precipitated with acid, as for the B_6 vitamer method (Section 2.11). **(d)** Concentration of PLP-cysteine in fibroblasts precipitated using methanol with an injection volume of 2 μL. **(e)** Decreasing PLP-cysteine response over time. **(f)** Samples were run twice sequentially. Graphs illustrate the decrease in response from the first injection (red) to the second (orange).



Despite this variability, fibroblasts from control and PROSC-deficient patients were analysed to determine whether any endogenous PLP-cysteine could be detected. The cell lysates were prepared using the same methodology as for the quantitation of the B_6 vitamers, with the exception that 50 μL of cell lysate supernatant was used to generate a more concentrated solution for analysis. Several peaks that shared the same parent/daughter ions and had similar retention times were detected; presumably isomeric forms of the PLP-cysteine conjugate. It has been shown that the thiazolidine exists as two diastereoisomers differing in the orientation of the carboxyl group, which are in equilibrium via the Schiff base (Ponticelli *et al.*, 1983; Terzuoli *et al.*, 1998) (Figure 7.3.14). The interconversion of the two diastereoisomers at a neutral pH of 6.4 occurs so rapidly that distinct nuclear magnetic resonsance signals from the two diastereoisomers cannot be discerned, however at much lower pH's ($\sim$ pH 3.0) similar to the solutions used for sample preparation and the mobile phase for the mass spectrometry analysis, this interconversion is significantly slower. If the reaction also shows first order kinetics, the rate of interconversion via the Schiff base will occur faster at higher concentrations. Therefore, the peaks could correspond to a mixture of the two diastereoisomers and the Schiff base (Figure 7.3.13c), given the low physiological concentrations of the thiazolidine complex in fibroblasts and low pH generated by the protein precipitation using 0.3N trichloroacetic acid and 3.7% acetic acid used as a component of the mobile phase.

Given the effects of assay conditions such as pH on the stability of PLP-cysteine, fibroblast samples were also prepared by using methanol precipitation (method described in Section 2.12). Analysis using this method resulted in a single peak and demonstrated an apparent increase in the concentration of the thiazolidine in PROSC-deficient cells (Figure 7.3.13d). However, upon greater scrutiny is was apparent that the response corresponding to the PLP-cysteine thiazolidine was decreasing over time (Figure 7.3.13e). This instability has also been documented by others (Buell and Hansen, 1960). Given that the PROSC-deficient cells were analysed first, it is unknown whether or not the increased concentrations identified in these fibroblasts were an artefact of the instability of the analyte. All samples were analysed twice with approximately five hours between each analysis (Figure 7.3.13f). Despite being kept in the dark at 4°C between analyses, PLP concentrations were comparably reduced in all samples. In contrast, the PLP-cysteine concentrations were reduced in the PROSC samples but not in control samples. The significance of these findings, if any, are uncertain. Therefore the development of novel, more robust analytical methods is warranted to further investigate the hypothesis that pathogenesis in PROSC deficiency is caused by uncontrolled reactions of PLP with proteins and small molecules such as cysteine and carnosine (Vistoli *et al.*, 2013). Indeed, very recently a novel method using hydrophilic interaction liquid chromatography followed by ultra-violet detection has been described using

Figure 7.3.14: Formation of the PLP-cysteine thiazoldine complex. Firstly, PLP and cysteine react via a condensation reaction with the loss of water to form an addition compound which is in equilibrium with the Schiff base. This imine molecule then cyclises to form the thiazolidine complex.

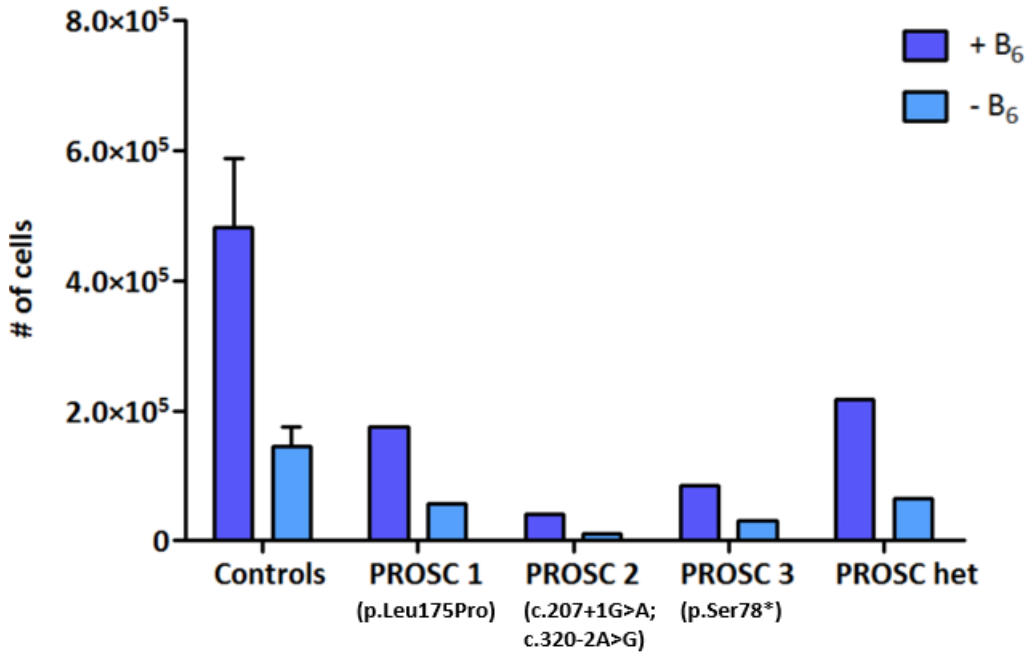


PLP as a derivatising agent for the analysis of cysteine and homocysteine, which may be adapted and utilised to analyse endogenous concentrations of the PLP-cysteine thiazolidine complex (Głowacki *et al.*, 2016).

Finally, in order to further investigate the apparent inability of patients with PROSC deficiency to appropriately regulate PLP homeostasis, the growth of control and PROSC-deficient fibroblasts in standard media (repleted) and media depleted of PN was examined. Fibroblasts were seeded at the same density and allowed to grow for seven days before counting the cells (Section 7.2). In standard media which contains 971 nM pyridoxine hydrochloride, the PROSC-deficient cells showed a 2.5- to 10-fold reduction in growth compared to control cells. Furthermore, PROSC het (p.Ser78*, heterozygous) fibroblasts demonstrated an intermediate reduction in growth compared to PROSC 3 cells (p.Ser78*, homozygous) despite the growth rate of the PROSC het cells being 2-fold reduced when compared to control levels.

However, when control fibroblasts were cultured in media depleted of a source of vitamin B_6, growth was reduced by approximately 70%. This is intuitive as the sub-optimal functioning of cellular B_6-dependent enzymes would be expected to result in a deficiency of metabolites essential for growth and cell division. This trend was also reflected in the PROSC-deficient and PROSC

Figure 7.3.15: Fibroblast growth in media repleted and depleted of vitamin B_6. Controls, n=3; patients, n=1.



het fibroblasts with growth being reduced by between 63% and 74% in conditions of B_6 depletion compared to repletion. Indeed, the trends observed in repleted media were also recapitulated in depleted media, with PROSC 2 cells showing the most severe growth impairment followed by PROSC 3 and PROSC 1. This growth deficiency in all patients with *PROSC* mutations, independent of mutation type, suggests that under conditions which usually promote cell viability, these cells experience PN toxicity which may be caused by an accumulation of toxic PLP-reacted molecules and inefficient functioning of PLP-dependent apoenzymes.

7.3.9 *Current hypotheses regarding PROSC deficiency*

The most consistent finding amongst patients with PROSC deficiency was an apparent inability to regulate PLP levels. In one patient, CSF PLP deficiency was identified prior to the initiation of vitamin B_6 supplementation. Combined with the patients' universal dramatic seizure resolution in response to supplementation with PN or PLP, this suggested a cerebral PLP deficiency. However, when these patients were treated with doses of PN or PLP typically used for the treatment of PNPO deficiency or pyridoxine-dependent epilepsy, abnormally high concentrations of PLP where detected in plasma. Indeed, this abnormality was also identified in patient-derived fibroblasts grown in conditions of PN excess. Further investigation revealed that these supra-physiological cellular concentrations of PLP react non-specifically with a range of intracellular molecules including proteins and metabolites. Finally, PROSC-deficient fibroblasts exhibited reduced

growth rates under conditions that usually promote cell viability, suggesting a susceptibility to PN-toxicity. Taken together, these results suggest that PROSC plays a role in cellular homeostasis of PLP in humans.

In eukaryotic cells, the concentration of free PLP is tightly controlled at approximately 1 μM; however, this is not high enough to meet the demand of the multiple B_6-dependent apoenzymes (di Salvo *et al.*, 2011). Thus, the question of how cells supply sufficient PLP with high enough specificity to avoid non-specific attack on nucleophilic molecules remains. As touched on in Section 7.1.5, several mechanisms have been proposed related to how cells ensure that PLP is guided to apoenzymes whilst preventing damaging accumulation of the free vitamer. These include:

1. Significant substrate inhibition of pyridoxal kinase and PNPO by PLP prevents its accumulation (Zhao and Winkler, 1995; Safo *et al.*, 2006).

2. Free cellular PLP is dephosphorylated by phosphatases to form PL, which is subsequently converted to 4-pyridoxic acid by aldehyde oxidase and NAD-dependent dehydrogenases for urinary excretion.

3. Circulating PLP is protected from hydrolysis by phosphatases through protein binding. In plasma, in excess of 95% of circulating PLP is bound to the Lys190 residue of human serum albumin (Bohney *et al.*, 1992). PLP also binds to the N-terminal amino acid of the β-chain of deoxygenated haemoglobin in erythrocytes which decreases its affinity for oxygen (Benesch *et al.*, 1982).

4. It has been proposed that amino acids may carry PLP to apoenzymes and function to provide a pool of bioavailable PLP. Each PLP-amino acid aldimine would react with a corresponding apoenzyme whose substrate has a similar structure to allow for PLP transfer within the active site. This hypothesis is largely based on the observation that, similarly to our molecular weight cut-off experiments, a significant proportion of cellular PLP in *E. coli* cells is "free" (i.e. not protein-bound) (Fu *et al.*, 2001).

5. Both pyridoxal kinase and PNPO have also been proposed to have a channelling capacity, protecting PLP from the cytosolic environment and interacting with each apoenzyme. Structural and *in vitro* solution studies have revealed that one PLP molecule binds at a non-catalytic site within a surface cleft of each PNPO monomer (Safo *et al.*, 2005). This site lies approximately 11 angstroms from the active site with a putative tunnel connecting the two, allowing the spatial transfer of PLP without contacting the solvent. Indeed, crystallographic studies have illustrated that a series of conformational changes occur following PNP binding

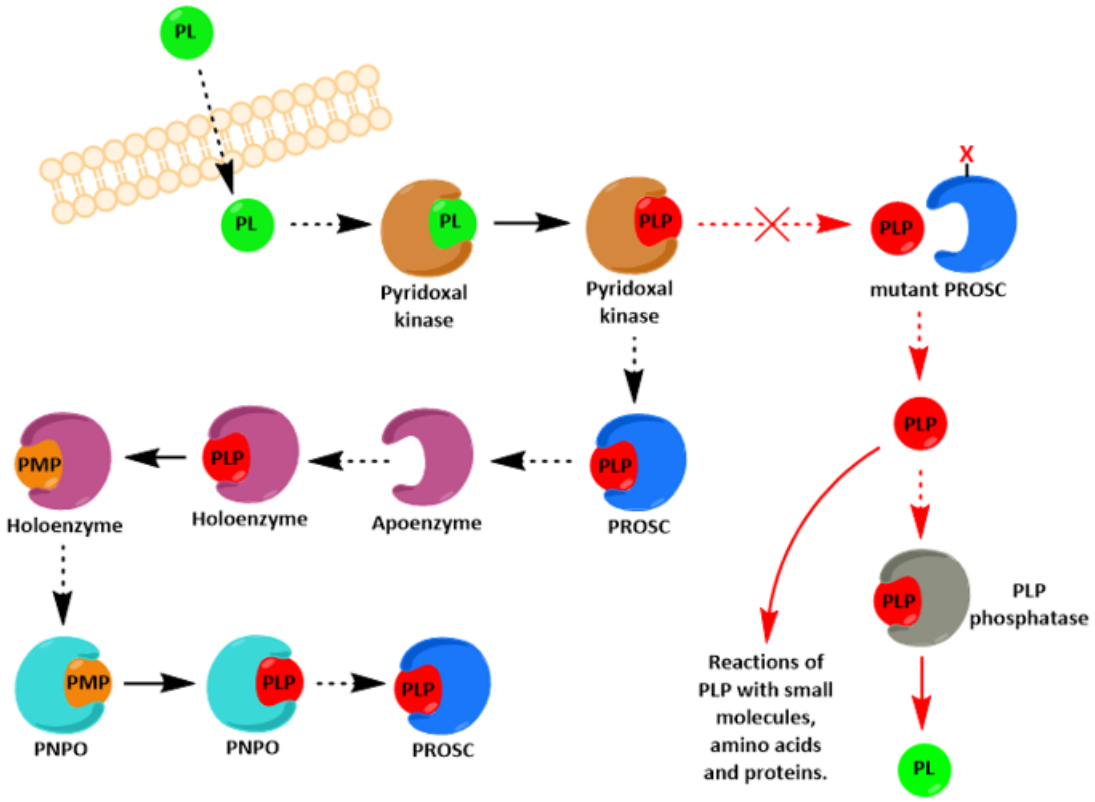
at the active site that enlarge the putative channel to such an extent that PLP could pass through (Safo *et al.*, 2005). Studies have shown that more efficient conversion of certain apoenzymes to holoenzymes occurs when PLP bound to the non-catalytic site of PNPO was introduced to the cell extract when compared to free PLP (Yang and Schirch, 2000). It has also been demonstrated that pyridoxal kinase can form complexes with multiple B_6-dependent enzymes including aspartate aminotransferase, alanine aminotransferase and glutamate decarboxylase accompanied by the transfer of PLP (Kim *et al.*, 1988; Cheung *et al.*, 2003). However, this hypothesis remains somewhat controversial as it requires that these two structurally dissimilar proteins are able to recognise and bind to over 100 different enzymes, encompassing many different topological folds.

It is highly unlikely that these mechanisms (if they indeed occur *in vivo*) are mutually exclusive. Rather, the predominance of each is likely to be specific to each apoenzyme and determined by its cellular location. Whilst the exact mechanisms underlying this function remain unknown, we propose that PROSC is a PLP-carrier protein which protects PLP from endogenous phosphatase activity as well as preventing unwanted side-reactions and ensuring a sufficient reservoir for PLP-dependent apoenzymes.

One possible mechanism that we propose is that, once pyridoxal has entered the cell and been phosphorylated by pyridoxal kinase, the PLP is transferred to the PLP binding site within the β-barrel motif of PROSC. Thereby ensuring that the concentration of free cellular PLP is maintained at low enough levels to prevent inappropriate reactions with other reactive molecules (including amino acids, peptides and proteins) and also its dephosphorylation to pyridoxal by endogenous phosphatases. When required, the PLP is then transferred from PROSC to B_6-dependent apoenzymes where it forms a Schiff base with specific lysine residues within their active sites. Following catalysis PMP may be formed which can be recycled by PNPO back to PLP and transferred again to PROSC (Figure 7.3.16).

As the results described in this chapter show, mutations in *PROSC* affect the homeostasis of vitamin B_6. This is likely due to the inability of the mutant protein to protect PLP (Figure 7.3.16). This not only results in the cofactor being dephosphorylated to pyridoxal, but due to the highly reactive nature of PLP, the vitamer reacts with small molecules such as cysteine and carnosine as well as the N-terminal amine of peptides and proteins. Whilst the total amount of PLP (both bound and unbound) in patient cells and biofluids is high when supplemented with supra-physiological concentrations of vitamin B_6, under normal conditions the amount of bioavailable PLP is insufficient in these patients. This results in decreased activity of enzymes requiring PLP for effective catalysis and an increase in the concentration of their substrates. In

Figure 7.3.16: Schematic illustrating the hypothesised function of PROSC and the implications of its deficiency. Steps involving the interconversion/chemical alteration of the B_6 vitamer are shown as solid arrows. Those that simply involve vitamer transfer between enzymes are shown as dashed arrows. The implications of a deficiency or dysfunction of PROSC are depicted in the right-hand pathway (red arrows). PL, pyridoxal; PLP, pyridoxal 5'-phosphate; PMP, pyridoxamine 5'-phosphate.



patients carrying these mutations this presents as an epilepsy that responds to treatment with vitamin B_6.

7.3.10 *Evaluation of the UPLC-MS/MS methods developed in this chapter*

7.3.10.1 *Effect of cell culture conditions on the quantitation of endogenous B_6 vitamer concentrations*

Whilst the method described in this chapter consistently identified trends in B_6 vitamer profiles between disease groups and individual patients, a degree of variability was noted in the absolute vitamer concentrations (Figure 7.3.4). It is likely that a significant proportion of this variability arose as a result of differences in passage number as each cell line was received at different times throughout this project. Endeavours were made to control for this parameter, although the passage number between patients varied by up to five. Reports have shown that increased time in serial sub-culture results in increased cellular senescence, apoptosis, decreased fidelity of DNA polymerase activity and alterations in metabolism including mitochondrial function (Mammone *et al.*, 2006; Linn *et al.*, 1976; Shmookler Reis and Goldstein, 1983). To date, there have been no reports describing the effects of passage number on vitamin B_6 metabolism. However, this metabolic pathway is facilitated by the action of multiple enzymes including pyridoxal kinase, PNPO and intracellular phosphatases (Figure 7.1.1) and it is possible that the activity of these enzymes may be affected by the time each cell line has spent in sub-culture, thereby changing the vitamer profile. In addition to variations in the age of each fibroblast line, cellular metabolism can be affected by culture conditions. For example, the turnover rate of proteins within cultured cells has been estimated to be 1% per hour (Eagle *et al.*, 1959). The degradation and turnover of B_6-dependent holoenzymes would be expected to result in the release of PMP which subsequently is recycled using the salvage pathway to form PLP. In addition, small variations in factors such as the confluence of each culture, time between media changes and media volume will affect the amount of PN taken up by each cell. Although this would not be expected to affect the intracellular vitamer composition, the absolute concentrations of each vitamer could be altered.

7.3.10.2 *Advantages of the direct quantitation of PNPO activity in patient fibroblasts*

Current published methods for investigating the pathogenicity of variants in *PNPO* have necessitated the overexpression of each mutant protein in an artificial system (Mills *et al.*, 2014; Plecko *et al.*, 2014). This requires various molecular techniques including site-directed mutagenesis, cloning and transfection which are both costly and time-consuming. In addition, these methods may not accurately recapitulate endogenous post-translational modifications, dimerisation and folding; they also do not take into account the genetic background of each patient. Whilst, the method we have described requires the patient to undergo a skin biopsy, once the fibroblast cell line has been established the pathogenicity of sequence variants in *PNPO* can be rapidly

determined. Furthermore, as a coupled assay, it is possible to not only measure PNPO activity but also that of pyridoxal kinase, for which mutations have yet to be described. In addition, supplementation of different compounds including B_6 vitamers (e.g PN, PLP and PM) and cofactors (e.g. flavin mononucleotide) can be trialled *in vitro* to provide optimal enzyme function and potentially improve treatments for patients.

7.4 SUMMARY

In summary, methods have been developed to enable the functional assessment of patients with known or suspected abnormalities of vitamin B_6 metabolism with a view to providing personalised medicine and a greater understanding of pathogenic mechanisms. These methods have enabled the confirmation of a diagnosis of PNPO deficiency in the background of only partially informative genetic analyses and provided evidence to suggest that PROSC is a protein essential for intracellular PLP homoeostasis in humans. Supplementary B_6 vitamer data should be gathered from additional PNPO-deficient patients to further explore the reasons underlying differences in treatment response. Finally, additional research should be undertaken to better understand the function of PROSC and its deficiency. Generation of a knock-out animal model followed by in-depth disease characterisation would be a valuable endeavour.

8

CONCLUSIONS AND FUTURE WORK

This thesis has utilised next-generation sequencing technology and novel mass spectrometry-based techniques to determine the cellular and molecular aetiologies for the clinical phenotypes observed in patients with undiagnosed neurometabolic disorders (NMD). Inborn errors of metabolism (IEM), of which NMD constitute a large sub-group, affect up to 1 in 500 newborns and account for a significant proportion of morbidity and mortality both in childhood and the neonatal period (Chiaratti de Oliveira *et al.*, 2001). Indeed, defects in more than 600 genes are known to cause these disorders. However, extreme genetic and phenotypic heterogeneity amongst IEM often results in diagnostic delays, with potential consequential adverse neurological outcomes. Ultimately it is hoped that research within these groups of patients will impact upon children and their families by improving diagnosis and treatment options, thereby having a positive impact on quality of life.

The work presented in Chapter 3 demonstrates the utility of an extended gene panel sequencing approach to improve the diagnosis of patients with undiagnosed suspected NMD in clinical practice. Indeed, this study highlighted findings that may have important implications not only for the diagnosis of patients with complex NMD, but also the interpretation of routine biochemical tests in wider patient populations. These include the propensity for patients to present with atypical or wide-ranging clinical features, blended phenotypes resulting from more than one single-gene defect or harbour mutations in genes causing disorders that are extremely rare or recently described. However, the difficulties in predicting the functional consequences of novel sequence variants were also emphasised. As noted by others (Walters-Sen *et al.*, 2015), *in silico* prediction tools were unable to consistently correctly assign pathogenicity. Future work is warranted to consider the advantage and practicality of techniques such as computational protein structural analysis to improve diagnostic specificity in situations where there is no scope for definitive *in vitro* assays.

This thesis also explores whole exome sequencing for the examination of patients in which extensive genetic and biochemical testing has not identified a diagnosis, and molecular biology techniques to explore their underlying aetiology. Chapter 5 presents the cases of five children from two unrelated families affected by a severe developmental disorder accompanied by, in the majority of cases early-onset seizures, who were found to have mutations in *SLC25A22*. Despite being known to function as a mitochondrial glutamate transporter, no biochemical abnormalities had been identified in any of the nine patients described in the literature. In contrast, multiple abnormalities including hyperprolinaemia, low CSF glutamate and cellular lipid accumulation

were observed in these cases. Taken together, these features indicated that SLC25A22 may also function *in vivo* as a mitochondrial glutamate γ-semialdehyde transporter. In order to substantiate this hypothesis, SLC25A22 should be reconstituted into liposomes as described by Molinari *et al.* (2005) to determine whether this mitochondrial protein can indeed catalyse the transport or exchange of this intermediate of proline metabolism.

Unlike the previous section in which novel phenotypes led to the proposition of alternative protein functions, Chapter 6 aimed to determine whether sequence variants in *CCBL1*, a gene that has not been previously associated with disease, were pathogenic in a patient with a movement disorder and hyperlysinaemia. These variants were deemed to be good candidates due to the known role of this protein in the homeostasis of kynurenic acid in the brain and postulated function in lysine metabolism. Mass spectrometry-based studies revealed that the variants had no effect on enzyme activity towards kynurenine but, despite trying several different approaches in earnest, the possibility of CCBL1 playing a role in lysine metabolism has not been excluded. Indeed, further studies to interrogate the crystal structure of this protein for possible binding sites of lysine would be beneficial. In addition, evaluation of the effect of these sequence variants on the other biological functions of CCBL1, including the metabolism of the cysteine conjugates of certain halogenated alkenes and alkanes may reveal additional insights (Han *et al.*, 2009b).

Much of this thesis focusses on the metabolism of vitamin B_6 and the implications of genetic defects affecting the homeostasis of these vitameric species. These disorders are a well known group of NMD typically presenting with seizures which are intractable to treatment with conventional antiepileptic drugs but respond to high-dose vitamin B_6 therapy. Novel mass spectrometry-based assays identified subtle differences in vitamer concentrations between patients with pyridox(am)ine 5'-phosphate oxidase (PNPO) deficiency depending on their response to treatment. This not only paves the way for personalised medicine in these groups of patients to optimise the efficacy and minimise the systemic toxicity of treatment, but also furthers the understanding of the mechanisms underlying these differing therapeutic outcomes. Ongoing research to address the latter would be best performed using larger patient cohorts and each assay should be adapted to enable the analysis of dried blood spots for ease of patient sampling.

Nevertheless, as presented and discussed in Chapters 4 and 7, many patients present to clinicians with seizures that are responsive to vitamin B_6 but do not have PNPO or antiquitin deficiency. In the former, the diagnosis of an apparently-responsive patient with a potassium channelopathy required the detailed consideration of mechanisms underlying the anticonvulsant effect of B_6 vitamers in the wider epileptic population. The reasons behind this clinical response are likely multi-factorial including the requirement of pyridoxal 5'-phosphate (PLP) as a cofactor in the synthesis of GABA (the major inhibitory neurotransmitter in the brain), the ability

of B_6 vitamers to quench the reactive oxygen species released due to unregulated neuronal firing, and the inhibited activation of $P2X_7$ receptors which in turn reduces excitotoxicity and neuroinflammation (Henshall *et al.*, 2013). These hypotheses suggest that further research into these mechanisms using patch clamping in cell-based systems or by assessing the neuropathological effects of vitamin B_6 treatment in an established epileptic mouse model (Löscher, 2011), may lay the foundation for novel treatment protocols for intractable seizures. In contrast, Chapter 7 includes the investigation of the pathogenic mechanism underlying a novel inborn error of vitamin B_6 metabolism, PROSC deficiency. Patients with this disorder were found to have an inability to regulate their cellular PLP levels, suggesting that the PROSC protein functions to chaperone PLP to the enzymes which require it as an essential cofactor. The severe seizures affecting these patients are hypothesised to be propagated by the uncontrolled reactions of PLP with proteins and metabolites, not only resulting in a physiological PLP deficiency but also of small molecules (e.g. cysteine) and the impaired function of attacked enzymes. The detection of stable complexes of PLP with other metabolites in order to provide evidence to support this theory proved to be technically challenging. However, the adaptation of novel methods that use hydrophillic interaction liquid chromatography and ultra-violet detection for the quantitation of physiochemically similar molecules would be a valuable endeavour (Głowacki *et al.*, 2016).

The work contained within this thesis has provided a genetic diagnosis for more than twenty patients who, despite extensive genetic and often invasive biochemical testing over many years, remained undiagnosed. It has also expanded the genotypic and phenotypic spectrum of many neurometabolic disorders and resulted in the implementation of comprehensive clinical genetic analysis for inborn errors of metabolism, which it is hoped will benefit many children in the future. However, nine patients remained undiagnosed following gene panel sequencing and whole exome sequencing failed to identify a diagnosis in six families. In these cases, a whole genome sequencing approach should be employed as disease may be due to mutations in genes that have not been associated with disease previously or are not amenable to targeted capture methodologies. Indeed, the underlying aetiology of disease may be due to a mutation that, due to a lack of supporting bioinformatic and functional evidence, are currently classified as variants of uncertain significance. As the number of individuals who have had their exome or genome sequenced rapidly increases, databases containing population frequency data will become more populated, which will aid in the interpretation of patient data. However, there is a need for the development of high-throughput functional assays to assess the potential pathogenicity of variants that have not been associated with disease previously. Indeed, as healthcare shifts towards a paradigm of "personalised medicine" (i.e. using each individual's genetic profile to guide decisions made in regard to the prevention, diagnosis and treatment of disease), the complexity of unravelling the

relationship between genotype and phenotype of both rare and common disorders will remain at the forefront.

9

APPENDICES

9.1 CONFIRMATION OF GENE PANEL FINDINGS BY SANGER SEQUENCING

Table 9.1.1: Primers and conditions used to amplify and sequence gene panel findings. Primers to amplify *TPP1* are not listed because confirmatory Sanger sequencing was performed on a clinical basis at North East Thames Regional Genetic Service, GOSH, UK.

Gene	Sequence 5' → 3'	T_m (°C)	$MgCl_2$ (mM)	Product size (bp)
POMGNT1 (splice F)	GTCCATGTCTGCCAGCTCT	64	1.5	250
POMGNT1 (splice R)	CCCAAGGTTACATGGCTAGC			
POMGNT1 (missense F)	TGTTTCAAGCAGCTGGTGTT	58	1.5	232
POMGNT1 (missense R)	ACTTCTGGTGAGTTGGTGTCA			
ACSF3 (F)	CTGGCATAGCTGTTTCTCCG	62	1.5	389
ACSF3 (R)	GACTCATCTGCAGTCGTCTAA			
PEX6 (F)	TGCCAACTCTGTTTCTTCCTG	60	1	293
PEX6 (R)	CCTCAAACTCCTGGGCTCAA			
AFG3L2 (F)	TGTTCTACCATAGCTCAGATGTT	60	1.5	370
AFG3L2 (R)	AGGGCCATCTCTAGCAAGTG			
SERAC1 (F)	CCCATTCGGCCTCTTTCAGT	62	1.5	383
SERAC1 (R)	TACAGCGCTTGAAGGGAGAA			
PGAP2 (F)	AATTACCTGCCCTCGGTGAG	60	1.5	534
PGAP2 (R)	TTTTCTTCTGGGCTGCCTTG			
DPYS (F)	TCCGGATTTGCAGCCTGA	64	1	583
DPYS (R)	GACCCCAGCGAAGAGAATCT			
GALE (F)	GCATTGCCAAGGACTAAAACC	64	1.5	201
GALE (R)	CTAGTGTCTGTGCCCTGTCC			
ALDOB (F)	GGTCTTCTCCCTGGAACAC	62	1.5	500
ALDOB (R)	GATGGAAAAGGGTGAGAAGAGA			

Disease group / disease	***Gene name***
1. Disorders of amino acid and peptide metabolism	
1.1. Urea cycle disorders and inherited hyperammonaemias	
1.1.1. Carbamoylphosphate synthetase I deficiency	*CPS1*
1.1.2. N-Acetylglutamate synthetase deficiency	*NAGS*
1.1.3. Ornithine transcarbamylase deficiency	*OTC*
1.1.4. Citrullinaemia type1	*ASS1*
1.1.5. Argininosuccinic aciduria	*ASL*
1.1.6. Argininaemia	*ARG1*
1.1.7. HHH syndrome	*SLC25A15*
1.1.8. Citrullinemia Type 2	*SLC25A13*
1.1.9. Hyperinsulinemic hypoglycemia and hyperammonemia	*GLUD1*
1.1.10. Hyperammonemia	*CA5A*
1.2. Organic acidurias	
1.2.1. Glutaric aciduria	
1.2.1.1. Glutaric aciduria type I	*GCDH*
1.2.1.2. Glutaric aciduria type III	*C7orf10*
1.2.2. Propionic aciduria	*PCCA* *PCCB*
1.2.3. Methylmalonic aciduria	
1.2.3.1. Methylmalonyl-CoA mutase deficiency	*MUT*
1.2.3.2. Methylmalonyl-CoA epimerase deficiency	*MCEE*
1.2.4. Isovaleric aciduria	*IVD*
1.2.5. Methylcrotonylglycinuria	*MCCC1* *MCCC2*
1.2.6. Methylglutaconic aciduria	
1.2.6.1. Methylglutaconic aciduria type I	*AUH*
1.2.6.2. Methylglutaconic aciduria type II, Barth syndrome	*TAZ*
1.2.6.3. Methylglutaconic aciduria type III, Costeff syndrome	*OPA3*
1.2.6.4. Methylglutaconic aciduria type IV	
1.2.6.5. Methylglutaconic aciduria type V	*DNAJC19*
1.2.6.6. Methylglutaconic aciduria with deafness, encephalopathy and Leigh-like syndrome (MEGDEL)	*SERAC1*
1.2.7. 3-Hydroxy-3-methyl-glutaric aciduria	*HMGCL*
1.2.8. 2-Methylbutyric aciduria	*ACADSB*
1.2.9. 2-Methyl-3-hydroxybutyric aciduria, HSD10 disease	*HSD17B10*
1.2.10. 3-Oxothiolase deficiency	*ACAT1*
1.2.11. Isobutyric aciduria	*ACAD8*
1.2.12. Methacrylic aciduria	*HIBCH*
1.2.13. 3-Hydroxyisobutyric aciduria	*ALDH6A1*
1.2.14. Methylmalonate semialdehyde dehydrogenase deficiency	*ALDH6A1*
1.2.15. L-2-hydroxyglutaric aciduria	*L2HGDH*
1.2.16. D-2-hydroxyglutaric aciduria	
1.2.16.1. D-2-hydroxyglutarate dehydrogenase deficiency	*D2HGDH*
1.2.16.2. Mitochondrial isocitrate dehydrogenase deficiency	*IDH2*

1.2.17.	Aminoacylase deficiency	
1.2.17.1.	Aminoacylase 1 deficiency	*ACY1*
1.2.17.2.	Aminoacylase 2 deficiency	*ASPA*
1.2.18.	Methylmalonate semialdehyde dehydrogenase deficiency	*ALDH6A1*
1.2.19.	Combined methylmalonic and malonic aciduria	*ACSF3*
1.2.20.	Malonyl-CoA decarboxylase deficiency	*MLYCD*

1.3. Disorders of the metabolism of branched-chain amino acids not classified as organic acidurias

1.3.1.	Branched-chain amino acid transferase	*BCAT1*
		BCAT2
1.3.2.	Maple syrup urine disease	
1.3.2.1.	BCKD E1 alpha subunit of deficiency	*BCKDHA*
1.3.2.2.	BCKD E1 beta subunit of deficiency	*BCKDHB*
1.3.2.3.	Dihydrolipoamide branched chain transacylase deficiency	*DBT*

1.4. Disorders of phenylalanine or tyrosine metabolism

1.4.1.	Phenylalanine hydroxylase deficiency	*PAH*
1.4.2.	Tyrosinaemia type II	*TAT*
1.4.3.	4-hydroxyphenylpyruvate dioxygenase deficiency	*HPD*
S	Tyrosinaemia type III	
S	Hawkinsinuria	
1.4.4.	Alkaptonuria	*HGD*
1.4.5.	Tyrosinaemia type I	*FAH*

1.5. Disorders of the metabolism of sulphur amino acids

1.5.1.	Methionine adenosyltransferase I/III deficiency	*MAT1A*
1.5.2.	Glycine N-methyltransferase deficiency	*GNMT*
1.5.3.	S-adenosylhomocysteine hydrolase deficiency	*AHCY*
1.5.4.	Cystathionine beta-synthase deficiency	*CBS*
1.5.5.	Cystathionase deficiency	*CTH*
1.5.6.	Isolated sulfite oxidase deficiency	*SUOX*
1.5.7.	Methionine synthase deficiency-cblG	*MTR*
1.5.8.	Methionine synthase reductase deficiency-cblE	*MTRR*

1.6. Disorders of histidine, tryptophan or lysine metabolism

1.6.1.	Histidinaemia	*HAL*
1.6.2.	Urocanase deficiency	*UROC1*
1.6.3.	Glutamate formiminotransferase deficiency	*FTCD*
1.6.4.	Tryptophanaemia	*TDO2*
1.6.5.	Hyperlysinaemia	*AASS*
		PTPRZ1
1.6.6.	2-Aminoadipic aciduria	*DHTKD1*
1.6.7.	2-Oxoadipic aciduria	*DHTKD1*
1.6.8.	Hydroxykynureninuria	*KYNU*
1.6.9.	Hydroxylysinuria	*AGPHD1*
		AGXT2L2

1.7. Disorders of serine, glycine or glycerate metabolism

1.7.1.	Phosphoglycerate dehydrogenase deficiency	*PHGDH*
1.7.2.	Phosphoserine phosphatase deficiency	*PSPH*

1.7.3.	Phosphoserine aminotransferase deficiency	*PSAT1*
1.7.4.	Nonketotic hyperglycinaemia	
1.7.4.1.	P protein deficiency	*GLDC*
1.7.4.2.	T protein deficiency	*AMT*
1.7.4.3.	H protein deficiency	*GCSH*
1.7.5.	Sarcosinaemia	*SARDH*
1.7.6.	D-glyceric aciduria	*GLYCTK*

1.8. Disorders of ornithine or proline metabolism

1.8.1.	Ornithine aminotransferase deficiency	*OAT*
1.8.2.	Hyperprolinaemia type I	*PRODH*
1.8.3.	Hyperprolinaemia type II	*ALDH4A1*
1.8.4.	Hypoprolinaemia, Cutis laxa, autosomal recessive, type IIIa	*ALDH18A1*
1.8.5.	Cutis laxa, autosomal recessive, type IIb/IIIb	*PYCR1*

1.9. Disorders of amino acid transport

1.9.1.	Lysinuric protein intolerance	*SLC7A7*
1.9.2.	Cystinuria	*SLC3A1* *SLC7A9*
1.9.3.	Cystinuria-hypotonia syndrome (contiguous gene defect)	
1.9.4.	Hartnup disease	*SLC6A19*
1.9.5.	Iminoglycinuria	*SLC36A2*
1.9.6.	Lowe syndrome	*OCRL*
1.9.7.	Hypotonia-cystinuria syndrome	*SLC3A1*
1.9.8.	Hypotonia-cystinuria syndrome	*PREPL*

1.10. Other disorders of amino acid metabolism

1.10.1.	Glutamine deficiency, congenital	*GLUL*

1.11. Disorders of the gamma-glutamyl cycle

1.11.1.	Glutathionuria	*GGT1*
1.11.2.	Cysteinylglycinase deficiency	*DPEP1*
1.11.3.	Oxoprolinuria	*OPLAH*
1.11.4.	Gamma-glutamylcysteine synthetase deficiency	*GCLC*
1.11.5.	Glutathione synthetase deficiency	*GSS*

1.12. Other disorders of peptide metabolism

1.12.1.	Prolidase deficiency	*PEPD*
1.12.2.	Carnosinaemia	*CNDP1*
1.12.3.	Homocarnosinosis	

1.13. Other disorders of amino acid and protein metabolism

2. Disorders of carbohydrate metabolism

2.1. Disorders of galactose metabolism

2.1.1.	Classical galactosaemia	*GALT*
2.1.2.	Galactokinase deficiency	*GALK1*
2.1.3.	Uridine diphosphate galactose-4-epimerase deficiency	*GALE*

2.2. Disorders of fructose metabolism

2.2.1.	Essential fructosuria	*KHK*
2.2.2.	Hereditary fructose intolerance	*ALDOB*

2.3. Disorders of pentose metabolism

2.3.1.	Essential pentosuria	*DCXR*
2.3.2.	Ribose-5-phosphate isomerase deficiency	*RPIA*
2.3.3.	Transaldolase deficiency	*TALDO1*

2.4. Disorders of glycerol metabolism

2.4.1.	Glycerol kinase deficiency	*GK*
2.4.2.	Complex glycerol kinase deficiency due to contiguous gene deletion	

2.5. Disorders of glyoxylate metabolism

2.5.1.	Primary hyperoxaluria type I	*AGXT*
2.5.2.	Primary hyperoxaluria type II	*GRHPR*

2.6. Disorders of glucose transport

2.6.1.	Glucose transporter 1 deficiency (blood-brain barrier)	*SLC2A1*
2.6.2.	Glucose transporter 2 deficiency	*SLC2A2*
	S Fanconi-Bickel syndrome	
2.6.3.	Glucose/galactose malabsorption	*SLC5A1*

2.7. Disorders of gluconeogenesis

2.7.1.	Fructose-1,6-bisphosphatase deficiency	*FBP1*
2.7.2.	Pyruvate carboxylase deficiency	*PC*
2.7.3.	Phosphoenolpyruvate carboxykinase deficiency	*PCK1*

2.8. Glycogen storage disorders

2.8.1.	Glycogen storage disease type 1a, von Gierke	*G6PC*
2.8.2.	Glycogen storage disease type 1b, von Gierke	*SLC37A4*
2.8.3.	Glycogen storage disease type II, Pompe	*GAA*
2.8.4.	Glycogen storage disease type III, Cori	*AGL*
2.8.5.	Glycogen storage disease type IV, Andersen	*GBE1*
2.8.6.	Glycogen storage disease type V, McArdle	*PYGM*
2.8.7.	Glycogen storage disease type VI, Hers	*PYGL*
2.8.8.	Glycogen storage disease type VII, Tarui	*PFKM*
2.8.9.	Glycogen storage disease type IX	
2.8.9.1.	Hepatic phosphorylase kinase deficiency	*PHKA2*
2.8.9.2.	Hepatic and muscle phosphorylase kinase deficiency	*PHKB*
2.8.9.3.	Hepatic phosphorylase kinase deficiency with cirrhosis	*PHKG2*
2.8.9.4.	Muscle phosphorylase kinase deficiency	*PHKA1*
2.8.9.5.	Cardiac muscle phosphorylase kinase deficiency	*PRKAG2*
2.8.10.	Glycogen storage disease type X	*PGAM2*
2.8.11.	Glycogen storage disease type XI	*SLC2A2*
2.8.12.	Glycogen storage disease type XIV	*PGM1*
2.8.13.	Glycogen storage disease type XV	*GYG1*
2.8.14.	Glycogen storage disease type 0a, liver	*GYS2*
2.8.15.	Glycogen storage disease type 0b, muscle	*GYS1*
2.8.16.	Other glycogen storage disease	
2.8.16.1.	Muscle LDH deficiency	*LDHA*
2.8.16.2.	Aldolase A deficiency	*ALDOA*
2.8.16.3.	Beta-enolase deficiency	*ENO3*
2.8.16.4.	Phosphoglycerate kinase deficiency	*PGK1*
2.8.17.	Unspecified glycogen storage disease	

2.9. Other carbohydrate disorders

2.9.1.	Lactose intolerance	*LCT*
2.9.2.	Disaccharide intolerance 1	*SI*
2.9.3.	Trehalase deficiency	*TREH*

3. Disorders of fatty acid and ketone body metabolism

3.1. Disorders of lipolysis

3.1.1.	**Neutral lipid storage disease**	*ABHD5*

3.2. Disorders of carnitine transport and the carnitine cycle

3.2.1.	Carnitine transporter deficiency	*SLC22A5*
3.2.2.	Carnitine palmitoyltransferase I (CPTI) deficiency	*CPT1A*
3.2.3.	Carnitine acylcarnitine translocase deficiency	*SLC25A20*
3.2.4.	Carnitine palmitoyltransferase II (CPTII) deficiency	*CPT2*

3.3. Disorders of mitochondrial fatty acid oxidation

3.3.1.	Very long - chain acyl CoA dehydrogenase deficiency	*ACADVL*
3.3.2.	Mitochondrial trifunctional protein deficiency	*HADHA* *HADHB*
3.3.3.	Medium - chain acyl CoA dehydrogenase deficiency	*ACADM*
3.3.4.	Short - chain acyl CoA dehydrogenase deficiency	*ACADS*
3.3.5.	3-alpha-hydroxyacyl- CoA dehydrogenase deficiency	*HADH*
3.3.6.	Multiple acyl-CoA dehydrogenase deficiency	
3.3.6.1.	Electron transfer flavoprotein deficiency, alpha chain	*ETFA*
3.3.6.2.	Electron transfer flavoprotein deficiency, beta chain	*ETFB*
3.3.6.3.	ETF-ubiquinone oxidoreductase deficiency	*ETFDH*

3.4. Disorders of ketone body metabolism

3.4.1.	3-Hydroxy-3-Methylglutaryl-CoA synthase deficieny	*HMGCS2*
3.4.2.	Succinyl-CoA:3-Oxoacid-CoA transferase (SCOT) deficiency	*OXCT1*
3.4.3.	Cytosolic acetoacetyl-CoA thiolase deficiency	*ACAT1*

3.5. Other disorders of fatty acid and ketone body metabolism

3.5.1.	Malonyl CoA decarboxylase deficiency	*MLYCD*

4. Disorders of energy metabolism

4.1. Disorders of pyruvate metabolism

4.1.1.	Pyruvate dehydrogenase complex deficiency	
4.1.1.1.	Pyruvate dehydrogenase E1α subunit deficiency	*PDHA1*
4.1.1.2.	Pyruvate dehydrogenase E1β subunit deficiency	*PDHB*
4.1.1.3.	Dihydrolipoyl transacetylase deficiency	*DLAT*
4.1.1.4.	Dihydrolipoyl dehydrogenase deficiency	*DLD*
4.1.1.5.	Pyruvate dehydrogenase E3 binding protein deficiency	*PDHX*
4.1.1.6.	Pyruvate dehydrogenase kinase deficiency	*PDK1* *PDK2* *PDK3* *PDK4*
4.1.1.7.	Pyruvate dehydrogenase phosphatase deficiency	*PDP1* *PDP2* *PDPR*

4.1.1.8.	Pyruvate dehydrogenase deficiency, unspecified	

4.2. Disorders of the citric acid cycle

4.2.1.	2-Oxoglutarate dehydrogenase deficiency	*OGDH* *DLST*
4.2.2.	Fumarase deficiency	*FH*

4.3. Mitochondrial respiratory chain disorders (caused by nuclear mutations only)

4.3.1.	OXPHOS structural subunits	
4.3.1.1.	Complex I	*NDUFS1* *NDUFS2* *NDUFS3* *NDUFS4* *NDUFS6* *NDUFS7* *NDUFS8* *NDUFV1* *NDUFV2* *NDUFA1* *NDUFA2* *NDUFA9* *NDUFA10* *NDUFA11* *NDUFA12* *NDUFB3* *NDUFB9*
4.3.1.2.	Complex II	*SDHA* *SDHB* *SDHC* *SDHD*
4.3.1.3.	Complex III	*UQCRB* *UQCRQ*
4.3.1.4.	Complex IV	*COX4I2* *COX6B1* *COX7B*
4.3.1.4.	Complex V	*ATP5E* *ATP5A1*
4.3.2.	OXPHOS assembly factors	
4.3.2.1.	Complex I	*NDUFAF1* *NDUFAF2* *NDUFAF3* *NDUFAF4* *NDUFAF6* *NDUFAF5* *NUBPL* *FOXRED1* *ACAD9*

4.3.2.2.	Complex II	*SDHAF1*
		SDHAF2
4.3.2.3.	Complex III	*BCS1L*
		HCCS
		TTC19
4.3.2.4.	Complex IV	*SURF1*
		SCO2
		SCO1
		COX10
		COX15
		LRPPRC
		FASTKD2
		ETHE1
		TACO1
		COA5
		COX14
		COX20
4.3.2.5.	Complex V	*ATPAF2*
		TMEM70
4.3.3.	Required for mtDNA maintenance	*POLG*
		POLG2
		C10orf2
		SLC25A4
		TYMP
		DGUOK
		TK2
		SUCLA2
		SUCLG1
		MPV17
		RRM2B
4.3.4.	Required for mitochondrial gene expression	*PUS1*
		MTO1
		MRPS16
		MRPS22
		MRPL3
		GFM1
		TSFM
		TUFM
		AARS2
		DARS2
		EARS2
		FARS2
		HARS2
		IARS2
		LARS2
		MARS2

		RARS2
		SARS2
		YARS2
		TRMU
		MTFMT
		MTPAP
		C12orf65
		RMND1
4.3.5.	Defective Fe-S/lipoic acid biosynthesis	*ISCU*
		FXN
		NFU1
		BOLA3
		LIAS
		ABCB7
		GLRX5
4.3.6.	Disorders of CoQ10 biosynthesis	*PDSS1*
		PDSS2
		COQ2
		ADCK3
		COQ9
		COQ6
		COQ4
4.3.7.	Secondary CoQ10 deficiency	*APTX*
		SETX
		ETFDH
4.3.8.	Disorders of mitochondrial solute import	*SLC25A3*
		SLC25A12
		SLC25A22
		SLC25A38
4.3.9.	Disorders of mitochondrial protein import	*TIMM8A*
		DNAJC19
		GFER
		PNPT1
4.3.10.	Disorders of mitochondrial membrane lipids	*TAZ*
		AGK
		SERAC1
4.3.11.	Disorders of mitochondrial dynamics, fusion and fission	*MFN2*
		OPA1
		DNM1L
		MFF
4.3.12.	Miscellaneous disorders/unknown function	*AIFM1*
		TMEM126A
		SPG7
		HSPD1
		AFG3L2
4.3.12.1.	Hyperoxaluria Type III	*HOGA1*

4.3.12.1.	Charcot-Marie-Tooth disease, recessive intermediate, B (Lysyl-tRNA synthetase mutations)	*KARS*
4.3.12.1.	Spinocerebellar ataxia-7	*ATXN7*
4.3.12.1.	Succinyl CoA:3-oxoacid CoA transferase deficiency	*OXCT1*
4.3.12.1.	Parkinson disease 6, early onset	*PINK1*
4.3.12.1.	Hypotonia-cystinuria syndrome	*PPM1B*
4.3.12.1.	Wolfram syndrome 1	*WFS1*
4.3.12.1.	Wolfram syndrome 2	*CISD2*
4.3.13.	Disorders of creatinine metabolism	
4.3.13.1.	Creatine transporter deficiency	*SLC6A8*
4.3.13.2.	Guanidinoacetate methyltransferase deficiency	*GAMT*
4.3.13.3.	Arginine:glycine amidinotransferase deficiency	*GATM*

5. Disorders in the metabolism of purines, pyrimidines and nucleotides

5.1. Disorders of purine metabolism

5.1.1.	Primary idiopathic gout	*ABCG2*
5.1.2.	Familial juvenile hyperuricaemic nephropathy	*UMOD*
5.1.3.	Adenylosuccinate lyase deficiency	*ADSL*
5.1.4.	AICAR transformylase deficiency	*ATIC*
5.1.5.	Adenosine deaminase deficiency	*ADA*
5.1.6.	Deoxyguanosine kinase deficiency	*DGUOK*
5.1.7.	Myoadenylate deaminase deficiency	*AMPD1*
5.1.8.	Lesch-Nyhan syndrome	*HPRT1*
5.1.9.	Adenine phosphoribosyl transferase deficiency	*APRT*
5.1.10.	Phosphoribosyl pyrophosphate synthetase 1 defects	*PRPS1*
5.1.11.	Inosine triphosphatase deficiency	*ITPA*
5.1.12.	Adenosine deaminase superactivity	
5.1.13.	Purine nucleoside phosphorylase deficiency	*PNP*
5.1.14.	Mitochondrial Ribonucelotide Reductase subunit 2 deficiency	*RRM2B*
5.1.15.	Xanthinuria type I	*XDH*
5.1.16.	Xanthinuria type II	*XDH* *AOX1*
5.1.17.	Thiopurine S-methyltransferase deficiency	*TPMT*

5.2. Disorders of pyrimidine metabolism

5.2.1.	Orotic aciduria	*UMPS*
5.2.2.	Pyrimidine - 5 - nucleotidase deficiency	*NT5C*
5.2.3.	Dihydroorotate dehydrogenase deficiency	*DHODH*
5.2.4.	Uridine-5'-monophosphate hydrolase superactivity	*NT5C3*
5.2.5.	Thymidine phosphorylase deficiency	*TYMP*
5.2.6.	Thymidine kinase 2 deficiency	*TK2*
5.2.7.	Dihydropyrimidine dehydrogenase deficiency	*DPYD*
5.2.8.	Dihydropyrimidinase deficiency	*DPYS*
5.2.9.	Beta-ureidopropionase deficiency	*UPB1*
5.2.10.	Hyper-beta-alaninaemia	
5.2.11.	Beta-aminoisobutyrate-pyruvate transaminase deficiency	

5.3. Disorders of nucleotide metabolism

5.3.1.	Aicardi-Goutières Syndrome (AGS)	
5.3.1.1.	AGS1	*TREX1*
5.3.1.2.	AGS2	*RNASEH2B*
5.3.1.3.	AGS3	*RNASEH2C*
5.3.1.4.	AGS4	*RNASEH2A*
5.3.1.5.	AGS5	*SAMHD1*
5.3.1.6.	AGS6	*ADAR*
5.3.2.	RNASET2-deficient cystic leukoencephalopathy	*RNASET2*

6. Disorders of the metabolism of sterols

6.1. Disorders of sterol biosynthesis

6.1.1.	Mevalonate kinase deficiency	*MVK*
6.1.2.	Smith - Lemli - Opitz syndrome	*DHCR7*
6.1.3.	X-linked dominant chondrodysplasia punctata 2	*EBP*
6.1.4.	Congenital hemidysplasia with ichtyosiform erythroderma and limb defects	*NSDHL*
6.1.5.	Desmosterolosis	*DHCR24*
6.1.6.	Lathosterolosis	*SC5DL*
6.1.7.	Greenberg skeletal dysplasia	*LBR*
6.1.8.	Antley-Bixler syndrome	
6.1.8.1.	Antley-Bixler syndrome with disordered steroidogenesis	*POR*
6.1.8.2.	Antley-Bixler syndrome type without disordered steroidogenesis	*FGFR2*
6.1.9.	Sterol-C4-methyl oxidase deficiency	*MSMO1*

6.2. Disorders of bile acid biosynthesis

6.2.1.	3- β-hydroxysterol Δ5-oxidoreductase/isomerase deficiency	*HSD3B7*
6.2.2.	Δ4-3-oxysterol 5β-reductase deficiency	*AKR1D1*
6.2.3.	Oxysterol 7-alpha-hydroxylase deficiency	*CYP7B1*
6.2.4.	Cholesterol 7-alpha-hydroxylase deficiency	*CYP7A1*
6.2.5.	Cerebrotendinous xanthomatosis	*CYP27A1*
6.2.6.	Bile acid amidation defect	*BAAT*
6.2.7.	Bile acid CoA ligase deficiency	*SLC27A5*

6.3. Disorders of bile acid metabolism and transport

6.3.1.	Bilirubin UDP-glucuronosyltransferase 1 deficiency	*UGT1A1*
6.3.2.	Byler disease	*ATP8B1*
6.3.3.	Progressive familial intrahepatic cholestasis type 2	*ABCB11*
6.3.4.	Progressive familial intrahepatic cholestasis type 3	*ABCB4*

6.4. Other disorders in the metabolism of sterols

6.4.1.	X-linked ichthyosis	*STS*

7. Disorders of porphyrin and haem metabolism

7.1.1.	Acute neuropathic porphyrias	
7.1.1.1.	Acute intermittent porphyria	*HMBS*
7.1.1.2.	Variegate porphyria	*PPOX*
7.1.1.3.	Hereditary coproporphyria	*CPOX*
7.1.1.4.	Acute hepatic porphyria	*ALAD*

7.1.2.	Porphyrias with erosive photodermatosis	
7.1.2.1.	Porphyria cutanea tarda	*UROD*
7.1.2.2.	Congenital erythropoietic porphyria	*UROS*
7.1.3.	Porphyrias with acute painful photosensitivity	
7.1.3.1.	Erythropoietic protoporphyria	*FECH*
7.1.3.2.	X-linked dominant protoporphyria	*ALAS2*
7.1.3.3.	X-linked sideroblastic anaemia (XLSA)	*ALAS2*

8. Disorders of lipid and lipoprotein metabolism

8.1. Inherited hypercholesterolaemias

8.1.1.	Disorder of low density lipoprotein receptor	*LDLR*
8.1.2.	Sitosterolaemia	*ABCG5* *ABCG8*
8.2.1.	Autosomal dominant hypercholesterolemia-3	*PCSK9*
8.2.1.	Autosomal recessive hypercholesterolemia	*LDLRAP1*

8.2. Inherited hypertriglyceridaemias

8.2.1.	Familial chylomicronaemia	
8.2.1.1.	Familial lipoprotein lipase deficiency	*LPL*
8.2.1.2.	Familial apolipoprotein C - II deficiency	*APOC2*
8.2.2.	Familial hypertriglyceridaemia	*APOA5* *LIPI*

8.3. Inherited mixed hyperlipidaemias

8.3.1.	Familial dysbetalipoproteinaemia	*APOE*
8.3.2.	Familial combined hyperlipoproteinaemia	*USF1*
8.3.3.	Hepatic lipase deficiency	*LIPC*

8.4. Disorders of high density lipoprotein metabolism

8.4.1.	Apolipoprotein A-I deficiency	*APOA1*
8.4.2.	Tangier disease	*ABCA1*
8.4.3.	Lecithin cholesterol acyltransferase deficiency	*LCAT*
8.4.3.1.	Fish-eye disease	
8.4.3.2.	Norum disease	
8.4.4.	Familial hyperalphalipoproteinaemia	*CETP*

8.5. Inherited hypolipidaemias

8.5.1.	Familial abetalipoproteinaemia	*MTTP*
8.5.2.	Familial hypobetalipoproteinaemia	*APOB*
8.5.3.	Anderson disease	*SAR1B*
8.5.3.	Scavenger receptor class B type I deficiency	*SCARB1*

8.6. Other disorders of lipid and lipoprotein metabolism

8.6.1.1.	Sjøgren - Larsson syndrome	*ALDH3A2*
8.6.1.2.	Pancreatic triacylglycerol lipase deficiency	*PNLIP*
8.6.1.3.	Pancreatic colipase deficiency	*CLPS*

8.7. Unspecified disorders of lipid and lipoprotein metabolism

8.8. Disorders of complex lipid synthesis

8.8.1.	Serine palmitoyl transferase deficiency	*SPTLC1*
8.8.2.	Serine palmitoyl transferase deficiency	*SPTLC2*
8.8.3.	Fatty acid 2-hydroxylase deficiency	*FA2H*

8.8.4.	Phosphatidate phosphatase deficiency	*LPIN1*
8.8.5.	Phospholipase A2 deficiency	*PLA2G6*
8.8.6.	PHARC syndrome	*ABHD12*
8.8.7.	Choline kinase deficiency	*CHKB*
8.8.8.	GM3 synthase deficiency	*ST3GAL5*
8.8.9.	Acylglycerol kinase deficiency (Senger syndrome)	*AGK*

9. Congenital disorders of glycosylation and other disorders of protein modification

S CDG

9.1. Disorders of protein N-glycosylation

9.1.1.	Phosphomannomutase 2 deficiency	*PMM2*
9.1.2.	Phosphomannose isomerase deficiency	*MPI*
9.1.3.	Glucosyltransferase 1 deficiency	*ALG6*
9.1.4.	Mannosyltransferase 6 deficiency	*ALG3*
9.1.5.	Mannosyltransferase 8 deficiency	*ALG12*
9.1.6.	Glucosyltransferase 2 deficiency	*ALG8*
9.1.7.	Mannosyltransferase 2 deficiency	*ALG2*
9.1.8.	UDP-GlcNAc:Dol-P-GlcNac-P transferase deficiency	*DPAGT1*
9.1.9.	Mannosyltransferase 1 deficiency	*ALG1*
9.1.10.	Mannosyltransferase 7-9 deficiency	*ALG9*
9.1.11.	Flippase of Man5GlcNAc2-PP-Dol deficiency	*RFT1*
9.1.12.	N-acetylglucosaminyltransferase deficiency	*MGAT2*
9.1.13.	Glucosidase 1 deficiency	*GLS*
9.1.14.	TUSC3-CDG	*TUSC3*
9.1.15.	SRD5A3-CDG	*SRD5A3*
9.1.16.	Mannosyltransferase 1 deficiency	*ALG1*
9.1.17.	Congenital myasthenic sydrome	*ALG14*
9.1.17.	Congenital myasthenic sydrome	*GFPT1*
9.1.17.	ALG13-CDG	*ALG13*
9.1.17.	ALG11-CDG	*ALG11*
9.1.17.	ALG3-CDG	*ALG3*
9.1.17.	ALG9-CDG	*ALG9*
9.1.17.	IAP-CDG	*MAGT1*
9.1.17.	MOGS-CDG	*MOGS*
9.1.17.	MAN1B1-CDG	*MAN1B1*
9.1.17.	ST3GAL3-CDG	*ST3GAL3*

9.2. Disorders of protein O-glycosylation

9.2.1.	O-xylosylglycan synthesis deficiencies	
9.2.1.1.	Multiple exostoses type I	*EXT1*
9.2.1.2.	Multiple exostoses type II	*EXT2*
9.2.1.3.	Beta-1,4-galactosyltransferase 7 deficiency	*B4GALT7*
9.2.2.	O-N-acetylgalactosaminylglycan synthesis deficiencies	
9.2.2.1.	Polypeptide N-acetylgalactosaminyl transferase deficiency	*GALNT3*
9.2.2.2.	GALNT12-CDG	*GALNT12*

9.2.2.3.	COSMC-CDG	*C1GALT1C1*
9.2.3.	O-xylosyl/N-acetylgalactosaminylglycan synthesis deficiencies	*SLC35D1*
9.2.4.	O-mannosylglycan synthesis deficiencies	
9.2.4.1.	Protein-O-mannosyltransferase 1 deficiency	*POMT1*
9.2.4.2.	Protein-O-mannosyltransferase 2 deficiency	*POMT2*
9.2.4.3.	Protein-O-mannose beta-1,2-N-acetylglucosaminyltransferase deficiency	*POMGNT1*
9.2.4.4.	Fukutin deficiency	*FKTN*
9.2.4.5.	Fukutin-related protein deficiency	*FKRP*
9.2.4.6.	N-acetylglucosaminyltransferase-like protein deficiency	*LARGE*
9.2.4.7.	O-fucose-specific beta-1,3-N-acetylglucosaminyltransferase deficiency	*LFNG*
9.2.4.8.	O-fucose-specific beta-1,3-N-glucosyltransferase deficiency	*B3GALTL*
9.2.4.8.	LFNG-CDG	*LFNG*
9.2.4.8.	B4GALT7-CDG	*B4GALT7*
9.2.4.8.	B3GAT3-CDG	*B3GAT3*
9.2.4.8.	CHSY1-CDG	*CHSY1*
9.2.4.8.	CHST3-CDG	*CHST3*
9.2.4.8.	CHST14-CDG	*CHST14*
9.2.4.8.	CHST6-CDG	*CHST6*

9.3. Disorders of glycosphingolipid and glycosylphosphatidylinositol anchor glycosylation

9.3.1.1.	Lactosylceramide alpha-2,3-sialyltransferase deficiency	*ST3GAL5*
9.3.1.2.	Phosphatidylinositolglycan, class M deficiency	*PIGM*
9.3.1.3.	Hyperphosphatasia	*PIGV*
9.3.1.4.	Hyperphosphatasia	*PIGO*
9.3.1.4.	PIGA-CDG	*PIGA*
9.3.1.4.	PIGL-CDG	*PIGL*
9.3.1.4.	PIGN-CDG	*PIGN*
9.3.1.4.	PGAP2-CDG	*PGAP2*

9.4. Disorders of multiple glycosylation and other glycosylation pathways

9.4.1.	GDP-Man:Dol-P mannosyltransferase deficiency	*DPM1*
9.4.2.	Lec35 deficiency	*MPDU1*
9.4.3.	Beta-1,4-galactosyltransferase 1 deficiency	*B4GALT1*
9.4.4.	UDP-GlcNAc epimerase/kinase deficiency	*GNE*
9.4.5.	CMP-sialic acid transporter deficiency	*SLC35A1*
9.4.6.	GDP-fucose transporter deficiency	*SLC35C1*
9.4.7.	Dolichol pathway deficiencies	
9.4.7.1.	Dolichol kinase deficiency	*DOLK*
9.4.8.	Conserved oligomeric Golgi (COG) complex deficiency	
9.4.8.1.	Component of COG complex 7 deficiency	*COG7*
9.4.8.2.	Component of COG complex 1 deficiency	*COG1*
9.4.8.3.	Component of COG complex 8 deficiency	*COG8*
9.4.8.3.	Component of COG complex 4 deficiency	*COG4*

9.4.8.3.	Component of COG complex 5 deficiency	*COG5*
9.4.8.3.	Component of COG complex 6 deficiency	*COG6*
9.4.9.	V-ATPase deficiencies	
9.4.9.1.	V0 subunit A2 of vesicular H(+)-ATPase deficiency	*ATP6V0A2*
9.4.9.2.	COPII component SEC23B	*SEC23B*

9.5. Disorders of protein ubiquitinylation

9.6. Other disorders of protein modification

9.6. Other CDGs

9.6.1.	SLC35A2-CDG	*SLC35A2*
9.6.1.	G6PC3-CDG	*G6PC3*
9.6.1.	CDG2K	*TMEM165*
9.6.1.	Retinitis pigmentosa	*DHDDS*
9.6.1.	DMP3-CDG	*DPM3*

10. Lysosomal disorders

10.1. Mucopolysaccharidoses

10.1.1.	MPS I, Hurler, Scheie disease	*IDUA*
10.1.2.	MPS II, Hunter disease	*IDS*
10.1.3.	MPS III, Sanfilippo disease	
10.1.3.1.	MPS IIIA, Sanfilippo A disease	*SGSH*
10.1.3.2.	MPS IIIB, Sanfilippo B disease	*NAGLU*
10.1.3.3.	MPS IIIC, Sanfilippo C disease	*HGSNAT*
10.1.3.4.	MPS IIID, Sanfilippo D disease	*GNS*
10.1.4.	MPS IV, Morquio disease	
10.1.4.1.	MPS IVA, Morquio A disease	*GALNS*
10.1.4.2.	MPS IVB, Morquio B disease	*GLB1*
10.1.5.	MPS VI, Maroteaux - Lamy disease	*ARSB*
10.1.6.	MPS VII, Sly disease	*GUSB*
10.1.7.	MPS IX, Natowicz	*HYAL1*

10.2. Oligosaccharidoses

10.2.1.	Alpha - D – mannosidosis	*MAN2B1*
10.2.2.	Beta - D – mannosidosis	*MANBA*
10.2.3.	Sialidosis	*NEU1*
10.2.4.	Aspartylglucosaminuria	*AGA*
10.2.5.	Fucosidosis	*FUCA1*
10.2.6.	Schindler disease	*NAGA*

10.3. Sphingolipidoses

10.3.1.	GM1-gangliosidosis	*GLB1*
10.3.2.	GM2-gangliosidosis	
10.3.2.1.	GM2-gangliosidosis 0-variant, Sandhoff disease	*HEXB*
10.3.2.2.	GM2-gangliosidosis B-variant, Tay-Sachs disease	*HEXA*
10.3.2.3.	GM2-gangliosidosis AB-variant	*GM2A*
10.3.3.	Gaucher disease	*GBA*
10.3.4.	Krabbe disease	*GALC*
10.3.5.	Metachromatic leukodystrophy	*ARSA*
10.3.6.	Prosaposin deficiency	*PSAP*

10.3.6.1. Saposin A deficiency
10.3.6.2. Saposin B deficiency
10.3.6.3. Saposin C deficiency
10.3.6.4. Saposin D deficiency
10.3.7. Fabry disease *GLA*
10.3.8. Farber disease *ASAH1*
10.3.9. Niemann-Pick disease type A or B *SMPD1*
10.3.10. Niemann-Pick disease type C
10.3.10.1. Niemann-Pick disease type C1 *NPC1*
10.3.10.2. Niemann-Pick disease type C2 *NPC2*

10.4. Ceroid lipfuscinoses, neuronal (CLN)

10.4.1. CLN1, Santavuori-Haltia disease *PPT1*
10.4.2. CLN2, Jansky-Bielschowsky disease *TPP1*
10.4.3. CLN3, Batten Spielmeyer-Vogt disease *CLN3*
10.4.4. CLN4A, Kufs disease recessive type *CLN6*
10.4.5. CLN4B Kufs disease dominant type *DNAJC5*
10.4.6. CLN5 Finnish variant *CLN5*
10.4.7. CLN6 *CLN6*
10.4.8. CLN7 *MFSD8*
10.4.9. CLN8, Northern epilepsy type *CLN8*
10.4.10. CLN9
10.4.11. CLN10 *CTSD*

10.5. Lysosomal export disorders

10.5.1. Cystinosis *CTNS*
10.5.2. Salla disease/infantile sialic acid storage disease *SLC17A5*

10.6. Other lysosomal disorders

10.6.1. Mucolipidosis II, I-cell disease *GNPTAB*
10.6.2. Mucolipidosis III, Pseudo-Hurler polydystrophy *GNPTG*
10.6.3. Mucolipidosis IV *MCOLN1*
10.6.4. Multiple sulphatase deficiency *SUMF1*
10.6.5. Wolman/cholesterol ester storage disease *LIPA*
10.6.6. Pompe disease, GSD type II *GAA*
10.6.7. Sialuria *GNE*
10.6.8. Danon disease *LAMP2*
10.6.9. Cathepsin-related disorders
10.6.9.1. Galactosialidosis *CTSA*
10.6.9.2. Papillon-Lefèvre syndrome *CTSC*
10.6.9.3. Pycnodysostosis *CTSK*
10.6.10. Hermansky-Pudlak Syndrome *HPS1*

11. Peroxisomal disorders

11.1. Disorders of peroxisome biogenesis

PEX1
PEX2
PEX3
PEX5
PEX6

	PEX10
	PEX12
	PEX13
	PEX14
	PEX16
	PEX19
	PEX26
11.2. Rhizomelic chondrodysplasia punctata	
11.2.1. Rhizomelic chondrodysplasia punctata type 1	*PEX7*
11.2.2. Rhizomelic chondrodysplasia punctata type 2	*GNPAT*
11.2.3. Rhizomelic chondrodysplasia punctata type 3	*AGPS*
11.3. Disorders of peroxisomal alpha-, beta and omega-oxidation	
11.3.1. X-linked adrenoleukodystrophy	*ABCD1*
11.3.2. Peroxisomal acyl-CoA oxidase 1 deficiency	*ACOX1*
11.3.3. Peroxisomal D-bifunctional protein deficiency	*HSD17B4*
11.3.4. Sterol carrier protein deficiency	*SCP2*
11.3.5. Alpha-methylacyl-CoA racemase deficiency	*AMACR*
11.3.6. Refsum disease	*PHYH*
11.4. Other peroxisomal disorders	
11.4.1. Primary hyperoxaluria type I	*AGXT*
11.4.2. Acatalasaemia	*CAT*
11.4.3. Mulibrey nanism	*TRIM37*
12. Disorders of neurotransmitter metabolism	
12.1. Disorders in the metabolism of biogenic amines	
12.1.1. Tyrosine hydroxylase deficiency	*TH*
12.1.2. Aromatic L-amino acid decarboxylase deficiency	*DDC*
12.1.3. Dopamine beta-hydroxylase deficiency	*DBH*
12.1.4. Monoamine oxidase	*MAOA*
12.2. Disorders in the metabolism of gamma-aminobutyrate	
12.2.1. Succinic semialdehyde dehydrogenase deficiency	*ALDH5A1*
12.2.2. GABA transaminase deficiency	*ABAT*
12.3. Other disorders of neurotransmitter metabolism	
12.3.1. Dopamine transporter deficiency syndrome	*SLC6A3*
12.3.1. Brain Dopamine–Serotonin Vesicular Transport Disease	*SLC18A2*
13. Disorders in the metabolism of vitamins and (non-protein) cofactors	
13.1. Disorders of folate metabolism and transport	
13.1.1. Hereditary folate malabsorption	*SLC46A1*
13.1.2. Cerebral folate deficiency due to FOLR1 deficiency	*FOLR1*
13.1.3. Dihydrofolate reductase deficiency	*DHFR*
13.1.4. Methylenetetrahydrofolate reductase deficiency	*MTHFR*
13.2. Disorders of cobalamin absorption, transport and metabolism	
13.2.1. Intrinsic factor deficiency	*GIF*
13.2.2. Enterocyte intrinsic factor receptor deficiency	

13.10.1. Riboflavin transporter deficiency	*SLC25A1*
13.10.1. Riboflavin transporter deficiency	*SLC25A2*
13.10.1. Riboflavin transporter deficiency	*SLC25A3*

14. Disorders in the metabolism of trace elements and metals

14.1. Disorder of copper metabolism

14.1.1. Menkes syndrome	*ATP7A*
14.1.1.1. Occipital horn syndrome	
14.1.2. Wilson disease	*ATP7B*

14.2. Disorder of iron metabolism

14.2.1. Hereditary haemochromatosis	
14.2.1.1. Hereditary haemochromatosis Type 1	*HFE*
14.2.1.2. Hereditary haemochromatosis Type 2	*HFE2*
	HAMP
14.2.1.3. Hereditary haemochromatosis Type 3	*TFR2*
14.2.1.4. Hereditary haemochromatosis Type 4	*SLC40A1*
14.2.2. Acoeruloplasminaemia	*CP*
14.2.3. Neurodegeneration with brain iron accumulation (NBIA)	*PANK2*
	PLA2G6
	C19orf12
	FA2H
	WDR45
	ATP13A2

14.3. Disorder of zinc metabolism

14.3.1. Acrodermatitis enteropathica	*SLC39A4*
14.3.2. Hyperzincemia and hypercalprotectinemia	

14.4. Disorder of phosphate, calcium and vitamin D metabolism

14.5. Disorder of magnesium metabolism

14.5.1. Hypermagnesaemia	
14.5.1.1. Hypermanganesemia with dystonia, polycythemia, and cirrhosis	*SLC30A10*
14.5.2. Primary hypomagnesaemia	
14.5.2.1. Hypomagnesaemia type 1, intestinal	*TRPM6*
14.5.2.2. Hypomagnesaemia type 2, renal	*FXYD2*
14.5.2.3. Hypomagnesaemia type 3, renal	*CLDN16*
14.5.2.4. Hypomagnesaemia type 4, renal	*EGF*
14.5.2.5. Hypomagnesaemia type 5, renal with ocular involvement	*CLDN19*
14.5.2.6. Hypomagnesaemia type 6, renal	*CNNM2*
14.5.2.7. Gitelman syndrome	*SLC12A3*
14.5.3. Secondary hypomagnesaemia	
14.5.4. Hypomagnesaemic tetany	
14.5.5. Hypomagnesaemia with cerebellar atrophy, hypotonia, strabismus, developmental delay, short stature, mild skeletal dysplasia, and connective tissue abnormalities	*SLC39A8*

Figure 9.3.1: Literature review of reports indexed in PubMed detailing patients with autosomal dominant *KCNQ2* mutations that had been trialled either transiently or on a long-term basis with pyridoxine or pyridoxal 5'-phosphate. ACTH, adrenocorticotropic hormone; B6, vitamin B_6 (not stated whether PN or PLP); DZP, diazepam; CBZ, carbamazepine; CLB, clobazam; CZP, clonazepam; KD, ketogenic diet; FA, folinic acid; LEV, levetiracetam; LZP, lorazepam; MDZ, midazolam; NZP, nitrazepam; PB, phenobarbital; PHT, phenytoin; PLP, pyridoxal 5'-phosphate; PN, pyridoxine; TPM, topiramate; VGB, vigabatrin; VPA, valproate; ZNS, zonisamide.

	Allen et al. *Epilepsia* 2014; 55(9):e99–e105	Mefford et al. *Am J Med Genet Part A* 2012; 158A:3190–3195	Weckhuysen et al. *Ann Neurol* 2012; 71:15–25
Phenotype	Benign familial neonatal seizures	? Pyridoxine-dependent epilepsy	Neonatal epileptic encephalopathy
Mutation	c.419_430dup (p.Val143_Arg144insGlnTyrPheVal)	1.5 Mb terminal deletion of the long arm of chromosome 20	c.613A>G (p.Ile205Val)
Inheritance/family history	Maternal (affected mother)	No family history of epilepsy or developmental disability.	*De novo* mutation. Paternal aunt with seizures from day 7 to 4 years.
Sex	F	M	M
Perinatal and early development	Normal	Precipitous after a 36 week pregnancy complicated by frequent Braxton-Hicks contractions.	Normal
Seizure onset	Day 4	2 weeks	Day 2. Last 2 months of pregnancy rhythmical jerking similar to seizures.
Seizure type and frequency	Clonic mainly, but also tonic. Clusters multiple/day or days seizure-free. Minor cyanosis (hypoxia) in initial period. Subsequently, mainly tonic with upper limb involvement. Lasting 1–4 min, multiple/day, then intermittent clusters (couple/day), weeks and months seizure–free.	Reddening and tonic stiffening of arms, lasting approx 1 min. Initially sporadic but by 8 weeks of age were occurring 4-6/day.	Generalized tonic with clonic components, lip smacking, back arching, and apnoea. Multiple seizures daily.

AEDs and response	**PLP (used acutely)**. LEV (some response, required dose increases). Other drugs used but ineffective: PB, MDZ, LZP.	**PB 15mg/day and PN 100mg/day (after EEG showing seizure reduction after PN – no recurrence of clinical seizures). PN increased to 200mg/day. At 11 months, PB discontinued and PN reduced to 150mg/day. PN discontinued after diagnosis made.**	VGB (initially reduced seizures and normalized EEG with 7 weeks seizure freedom). **Combination TPM, VGB, and PN controlled seizures** (episode of status epilepticus at 3 months; seizure-free from 9 months until 8 years). Partially effective other drugs: MDZ. Other drugs used but ineffective: PB, FA, betamethasone, VPA.
Seizure outcome	Sporadic breakthrough minor seizures	Free of clinical seizures.	Seizure free from 9 months to 8 years with 2 recent tonic seizures
Current AEDs	LEV	None	TPM
EEG findings	At onset: Bilateral independent high amplitude sharp waves, degree of bisynchrony and periodicity (1 Hz at times) normal background. 3.5 months: Normal.	Hypsarrhythmia with frequent electrodecremental episodes. The following day the EEG was repeated and 100 mg of pyridoxine was administered intravenously which was followed within 1 min by a 95% reduction in the epileptiform activity. 14 weeks: Slow with multifocal epileptic activity. Video EEG telemetry (11 months): Independent and synchronous bifrontal spike and wave discharges and that the nocturnal episodes were confusional arousals from slow wave sleep. 5 years: Multifocal epileptiform activity with discontinuous background.	7 days: Multifocal epileptic activity most frequently seen in left temporal and right frontal regions. One seizure with nystagmus and intermittent bilateral clonic jerks. Ictal changes showed diffuse attenuation with multifocal spikes. 9 months: Normal.
Additional features, examination and developmental follow-up	Normal	Pendular nystagmus, delay in acquisition of milestones. 2 years: Hypotonic, legally blind, developmental level between 9-12 months. Noted to have nightly events of	Moderate mental retardation. Regression with status epilepticus. Not rolling at 6 months. Walked at 16 months; 30 single words at 4 years. At 8 years: follows 2 commands, reads small

		inconsolable crying and screaming developing between 2- and 3 hr after falling asleep and lasting from one to 90 min. 5 years: Minimal expressive language. 7 years: Nocturnal arousals resolved. Minimal expressive language, lack of bladder and bowel control, dysconjugate gaze, rotary nystagmus, axial hypotonia, increased tendon reflexes.	words. No use of toys. Poor fine motor skills.
Neuroimaging	MRI (3 weeks): Normal.	MRI: Slight delay in myelinisation. MRI (11 months): Normal	CT scan (2 days): Subdural haemorrhage. Generalized hypodense cerebral parenchyma suggestive of hypoxia (but not confirmed on MRI at 11 days). Head circumference just above 50th percentile (4 years). MRI (11 days): ↑T2 globus pallidus and thalamus. Normal white matter volume; ↑T2 ↓T1 signal in periventricular white matter; ↑T2 signal paralleling the posterior limbs of the internal capsules. MRI (3 years 6 months): Bilateral hippocampal enlargement (L>R) with ↑T2 left hippocampus
	Borgatti et al. *Neurology* 2004; 63:57–65	**Dedek et al. *Epilepsy Research* 2003; 54:21-27**	**Martin et al. *Human Molecular Genetics* 2014; 23(12): 3200–3211**
Phenotype	Benign familial neonatal seizures, epileptic encephalopathy, and profound mental retardation	Benign familial neonatal seizures	Ohtahara Syndrome
Mutation	c.1620G>A (p.Lys526Asn)	c.954C>G (p.Ser247Trp)	c.827C>T (p.Thr276Ile)
Inheritance/family history	Affected mother and two younger sisters	Affected mother	*De novo* mutation
Sex	F	M	M

Perinatal and early development	Born at 40 weeks by caesarean section due to podalic presentation	Born by caesarean section due to prolonged delivery period and symptoms of foetal distress	Born at 41 weeks by emergency caesarean section due to failure to progress
Seizure onset	Day 3	Day 3	Day 1
Seizure type at onset	Clonic seizures lasting about 40 seconds. Subsequently right sided clonic and tonic-clonic seizures with oro-alimentary automatism.	Left or right head deviation, and upper and lower limb involvement.	Cyanotic episodes, then more obvious seizures (up to 200 per day). Tonic spasms (more than 20 per day) lasting 2-3 minutes.
AEDs and response	Not achieved complete seizure control. Partially effective other drugs: ACTH. Other drugs used but ineffective: PB, VGB, benzodiazepines, PHT, VPA, CZP, PN, immunoglobulin.	ACTH (7 weeks – immediate improvement of EEG background activity and a progressive reduction of abnormalities. Seizure frequency decreased and seizures stopped at age 13 weeks). Other drugs used but ineffective: PB, PHT, PN, VGB.	TPM, DZP, and NZP (fits initially continued, but at 5 months were less severe, they ceased by 17 months). Other drugs used but ineffective: CZP, VGB, FA, PLP
Seizure outcome	Many polymorphic seizures a day (mainly tonic, atypical absence, and audiogenic startles).	Seizure free	Seizure free
Current AEDs	Not stated.	Not stated.	TPM and NZP.
EEG at onset	Ictal: Multifocal epileptiform abnormalities asynchronous over both cerebral hemispheres. Last EEG: Multifocal and bilateral asynchronous epileptiform abnormalities more evident over the left fronto-temporal areas.	Day 8: Theta–delta activity, synchronous, incomplete sleep spindles and sharp waves. Ictal: Asynchronous background activity with suppression-burst patterns and multiple paroxysmal abnormalities with random asynchronous attenuation patterns. 2.5 years: Background activity is normal, and epileptiform abnormalities are only intermittently evident over the bilateral central regions.	Day 1: Very abnormal with a discontinuous pattern, 2-3 seconds of EEG attenuation interrupted by burst of irregular slow/fast and sharp waves, the bursts occurring, both synchronously and asynchronously. There were a few periods of more continuous EEG, focal sharp waves were seen, mainly over the left mid central to parietal cortex. Day 14: The predominantly sleep EEG was more abnormal, characterized by synchronous and asynchronous bursts of irregular slow waves, with spike transients interrupted by 1-2 seconds of EEG attenuation (almost burst-

			suppression), with multifocal spikes standing out during the periods of EEG attenuation. The quasi burst suppression pattern persisted until age 3 months, after which the EEG became more continuous in wake and sleep. 20 months: More stable, dominated by drug-induced fast activity, with no epileptiform discharges. 4 years: Though continuous, again showed multifocal sharp/spike wave discharges, maximal over the right mid temporal cortex, spreading in sleep to the right fronto-central areas.
Additional features, examination and developmental follow-up	Severe spastic tetraparesis and profound mental retardation without any language capability.	Psychomotor development was delayed; he now has head control and shows social smiling, but is not able to sit without support. Muscle hypotonia and dystonic features are present.	Severe developmental delay. By age 4, he could sit with support, hold objects, indicate 'yes or no' by facial gestures, and showed some hand regard and environmental awareness. His head circumference has moved from the 75th to the 25th percentile over time. He is small with height/length falling from the 0.4th percentile, weight remaining at the 0.4th percentile.
Neuroimaging	MRI: Thin corpus callosum with moderate white matter reduction and slightly enlarged lateral ventricles	CT (day 7): Normal. CT (day 10): Mild, diffuse white matter hypodensity which was mostly obvious in the supratentorial compartment as well as signs of diffuse oedema. MRI (18 days): Normal CT (30 days): Normal MR spectroscopy: Normal	MRI: Generalized mild reduction in white matter bulk with a thin corpus callosum, but was otherwise normal.

	Numis et al. *Neurology* 2014; 82:368-370	Saitsu et al. *Ann Neurol* 2012; 72(2):298-300		
		Case # 1469	Case # 1654	Case # 1754
Phenotype	Neonatal epileptic encephalopathy	Ohtahara Syndrome	Ohtahara Syndrome	Ohtahara Syndrome
Mutation	c.1734 G>C (p.Met578Ile)	c.1010C>G (p.Ala337Gly)	c.341C>T (p.Thr114Ile)	c.794C>T (p.Ala265Val)
Inheritance/family history	Family histories were all negative for epilepsy or mental retardation	*De novo* mutation	*De novo* mutation	*De novo* mutation
Sex	Not stated.	M	F	M
Perinatal and early development	Born at 34 weeks. Lack of visual fixation, decreased spontaneous movements, and axial hypotonia.	Not reported	Not reported	Not reported
Seizure onset	Day 4	Day 7	Day 0	Day 1
Seizure type at onset	Tonic head, conjugate eye, and mouth deviation, associated with unilateral tonic abduction of the limbs. Apnoea and desaturation requiring oxygen administration.	Tonic seizures, vomiting. Complex partial seizures since age 5.	Tremor of the upper extremities then generalised convulsions with cyanosis. Complex partial seizures since age 5.	Apnoeic spell, then tonic spasms with right opsoclonus like movement.
AEDs and response	CBZ (effective – seizure free within 2 weeks). Other drugs used but ineffective: PB, LEV, TPM, VGB, CLB, CZP, KD, PN, PLP, FA.	High dose PB (seizure-free and burst-suppression disappeared. Other drugs used but ineffective: B6, ZNS.	ZNS (seizure free). Other drugs used but ineffective: B6, CZP, PHT.	Intractable. Other drugs used but ineffective: B6, ZNS, VPA, CZP, CBZ.
Seizure outcome	Seizure free	Complex partial seizures	Complex partial seizures	Intractable seizures.
Current AEDs	Not stated.	Not mentioned.	Not mentioned.	Not mentioned.
EEG at onset	Interictal: Lack of organization and physiologic features with almost-continuous multifocal epileptiform abnormalities intermixed with random asynchronous attenuations.	Initial: Burst-suppression	Initial: Burst-suppression	Initial: Burst-suppression

	Ictal: Low-voltage fast activity followed by recruiting spikes or theta rhythms arising mainly from the central regions of either hemisphere, followed by focal spike-wave complexes and prolonged focal or diffuse postictal attenuation.			
Additional features, examination and developmental follow-up	Severe psychomotor delay, quadriplegia, axial hypotonia with appendicular hypertonia, and a tendency to opisthotonic posturing.	Severe mental retardation, no pyramidal signs, no meaningful words, able to crawl, stand with support.	Profound mental retardation, spastic quadriplegia, developmental quotient 10, bed-ridden, smiling.	Developmental delay, no eye pursuit. Myoclonus at the bilateral upper extremities.
Neuroimaging	MRI (20 and 33 days of life (with spectroscopy)): Progressive diffuse hypomyelination with marked thinning of the corpus callosum; T1 signal prolongation in the lentiform nuclei that normalized on day 33 of life.	Not mentioned.	Not mentioned.	Not mentioned.
	Kato et al. *Epilepsia* 2013; 54(7):1282–1287			
	Patient 205	**Patient 14**	**Patient 110**	**Patient 272**
Phenotype	Ohtahara Syndrome	Early onset epileptic encephalopathy	Ohtahara Syndrome	Ohtahara Syndrome
Mutation	c.650C>A (p.Thr217Asn)	c.794C>T (p.Ala265Val)	c.794C>T (p.Ala265Val)	c.854C>A (p.Pro285His)
Inheritance/family history	*De novo* mutation	*De novo* mutation	*De novo* mutation	Maternally transmitted. Mother: idiopathic epilepsy since neonate, medicated with VPA
Sex	F	M	M	F
Perinatal and early development	Not reported.	Not reported	Not reported	Not reported
Seizure onset	Day 0	Day 2	Day 2	Day 0
Seizure type at onset	At onset: Pale face for tens of seconds. Day 1: Eye deviation to left	At onset: Facial flushing and eye fixation. Day 3: Tonic seizures (daily).	At onset: No cry, poor suck, stiffening and arching with eye rolling.	At onset: Poor feeding. Day 2: Irritability with hypoxia.

	followed by tonic seizures (0.5-1/h).		Day 5: Left-sided seizures.	Day 8: Tonic seizures (1-4/day).
AEDs and response	High dose PB (seizure-free). Other drugs used but ineffective: PLP, ZNS.	CBZ (seizure-free). Partially effective other drugs: DZP, MDL, high-dose PB. Other drugs used but ineffective: PLP, VPA.	Intractable. Other drugs used but ineffective: PB, PLP, CLB, MDZ, VGB.	VPA (seizure-free). Other drugs used but ineffective: PLP, PB.
Seizure outcome	Seizure-free.	Seizure-free from 16 months.	Intractable.	Seizure-free from 3 months.
Current AEDs	Not mentioned.	Not mentioned.	Not mentioned.	Not mentioned.
EEG at onset	Day 1: Burst-suppression, asymmetric. Transition to other EEG findings: Diffuse high-amplitude spike-and-slow wave, multifocal spikes.	Day 5: Multifocal sharp waves. Transition to other EEG findings: Focal spikes moving loci at each recording.	Day 5: Burst-suppression, brief suppression. Transition to other EEG findings: Rolandic spike waves, normal background activity.	Day 12: Burst-suppression, asymmetric.
Additional features, examination and developmental follow-up	No head control, no words, social smile, profound mental retardation, spastic quadriplegia, myoclonus, opisthotonic posture.	Sitting alone at 6 y, moderate mental retardation, spastic diplegia with dyskinesia, athetotic/dystonic movement.	Hypotonia, jerking of left arm, stiffening and arching. Development unknown. Died at 3m.	DQ 35, rolling-over at 18m, no words, moderate mental retardation, hypotonic quadriplegia.
Neuroimaging	On globus pallidus, ↑T1 signal at 2 d, ↑T1 & T2 at 1 m, ↑T2 at 6 m, normal at 2 y	Mildly delayed myelination, ↑T2 signal on GP at 7 m & 2 y	Normal at 0 m.	↑T1 & T2 signal on GP at 12 d, ↑T2 at 3 m.
	Kato et al. *Epilepsia* 2013; 54(7):1282–1287			**Milh et al. *Orphanet Journal of Rare Diseases* 2013; 8:80**
	Patient 168	**Patient 178**	**Patient 17**	
Phenotype	Ohtahara Syndrome	Ohtahara Syndrome	Ohtahara Syndrome	**Six patients described that have been treated with vitamin B_6 during the first month of life. Responses to each AED**
Mutation	c.881C>T (p.Ala294Val)	c.997C>T (p.Arg333Trp)	c.1689C>G (p.Asp563Glu)	
Inheritance/family history	*De novo* mutation	*De novo* mutation.	*De novo* mutation. Mother: symptomatic localisation-related epilepsy	

			controlled with CBZ since school age.	**not stated.**
Domain in protein	Transmembrane domain (S4)	C-terminal region	C-terminal region	
Sex	M	M	F	
Perinatal and early development	Not reported.			
Seizure onset	Week 1	Day 1	Day 1	
Seizure type at onset	At onset: Convulsion-like movements. 1 month: Asymmetric tonic seizures (10/day).	At onset: Tonic seizures (status epilepticus). Transition to partial seizures (eyes rolling up).	At onset: Poor feeding with cyanosis. Day 3: Tonic seizures followed by facial clonic seizures (>10/day). Transition to tonic seizures or generalised tonic-clonic convulsions (1/week).	
AEDs and response	TPM (seizure-free). Partially effective other drugs: ZNS, VPA. Other drugs used but ineffective: PLP, CZP, PB.	ZNS (almost seizure-free). Partially effective other drugs: VPA, lidocaine. Other drugs used but ineffective: DZP, PB, PHT, PLP.	CBZ and CZP (seizure-free). Partially effective other drugs: PHT, PB. Other drugs used but ineffective: PLP, VPA, NZP.	
Seizure outcome	Seizure-free from 6 months.	Seizure-free. Only one seizure in 10 years.	No seizures since 10 y, but relapsed at 24 y after a year of drug withdrawal.	
Current AEDs	Not mentioned.	Not mentioned.	Not mentioned.	
EEG at onset	< 1 months: Burst-suppression. 3 months: Evolution to hypsarrythmia.	Day 42: Burst-suppression, brief suppression. Transition to other EEG findings: Multifocal spikes in F4 and T3-5.	Day 4: Burst-suppression, asymmetric. Transition to other EEG findings: Frontal continuous fast wave and multifocal spikes. No paroxysms since 10 y.	
Additional features, examination and developmental follow-up	No rolling over, no words, profound mental retardation, hypotonic quadriplegia, mild	Speaking two-word-sentences, walking alone, severe mental retardation, left upper	Walking alone at 22 m, a few words at 4–5 y, moderate mental retardation with	

	motor deterioration.	extremity slightly rigid, tooth-grinding.	autistic features.
Neuroimaging	↑T2 signal on globus pallidus at 3 m & 9 m, moderate frontal atrophy with delayed myelination and thin corpus callosum at 9 m.	No images obtained.	Normal CT at 0 m (no images obtained).

Table 9.4.1: Primers and conditions used to amplify and sequence SLC25A22. The first set of primers used for the nested PCR amplification of exons 4 and 6 are indicated by a *.

Exon	Sequence 5' → 3'	T_m (°C)	$MgCl_2$ (mM)	Product size (bp)
1 - 2 F	CTCAAGGCCTCCTCCACC	64	1.5	420
	AGCTAAGCGCGAGAAGGC			
1 - 2 R	CCGTTTCCCTTCTGTCAAAC			
	GCTTTACCGCTCAACCGTT			
3 F	TGAGGACTTGGCCTCTTCTATC	66	1.5	280
3 R	ACGTCCACGCTCACACAC			
4* F	CAGGAGGCAAGTCCTGG	50	1.5	279
	AGCTAAGCGCGAGAAGGC			
4* R	AAGGGGCTGAAGACAGGC			
	GCTTTACCGCTCAACCGTT			
4 F	CTCCCCACTCAGGCCAC	62	2.0	208
4 R	GGGTGGACCCATCCTTTATC			
5 F	CCCCACAGGCCTGCATCT	64	1.5	374
	AGCTAAGCGCGAGAAGGC			
5 R	AGGCTGCTGTCTCCTCTTC			
	GCTTTACCGCTCAACCGTT			
6* F	GAAGAGGTGAGGGCGAGG	50	1.5	393
	AGCTAAGCGCGAGAAGGC			
6* R	CACCACAGAGAAGGGGACAT			
	GCTTTACCGCTCAACCGTT			
6 F	CCTGCCCAGTTCGCACAG	62	1.0	280
6 R	TGGGGTGGGCAGGCC			
7 F	ACGCTGCTCAGGTAGGAGG	66	1.5	495
7 R	CTGTGGAGGAAGGACGAAAG			
8 F	TCAACCCCTGTGATGGTCAG	62	1.5	247
	AGCTAAGCGCGAGAAGGC			
8 R	TGGAGGAAGGACGAAAGGG			
	GCTTTACCGCTCAACCGTT			
9 F	CTCTGGGATCCTGGACTGTG	66	1.5	345
9 R	GAGGGGTCTTCCCTTGCTC			

Table 9.5.1: Primers and conditions used to amplify and sequence CCBL1. All exons were amplified using a T_m of 64°C, 1.5 mM $MgCl_2$ and 5% DMSO. All primers are tailed at the 5' end (forward primers: AGCTAAGCGCGAGAAGGC and reverse primers: GCTTTACCGCTCAACCGTT).

Gene	Sequence 5' → 3'	T_m (°C)	$MgCl_2$ (mM)	Product size (bp)
1 F	TAGGAACGGGAGAGTGTGTG	64	1.5	322
1 R	GTCACCTCTCGAAACCCACT			
2 F	GCCCCGGAAGTGACGTCA	64	1.5	292
2 R	GTTAGCCGACCCTCCCAG			
4 F	CACTTGGTGCTATTCCTGGC	64	1.5	408
4 R	GCTGCAGTTAAAGAGGGCTC			
5 F	CTAGTGAGAGTGCAGTGGGA	64	1.5	309
5 R	CCAAGATCGTGCCAGCCA			
6 F	TGCTGGAGTCAAGAGGAACC	64	1.5	395
6 R	TGGAATCTGGCAGTCTGTGT			
7 F	TCTAAACAGGCCGAAGTCCA	64	1.5	293
7 R	GCTGTGTGGGCTTGAAGTC			
8 F	GGAGGAGGTGACACTTAAACC	64	1.5	389
8 R	AGTGAGCCGTGATCATGTCA			
9 F	AGGAGTTTGAGGCCAGTGC	64	1.5	249
9 R	GCCACCATACCCAGCTTTTG			
10 F	AGGTGCTAGAATTGCCTGGA	64	1.5	480
10 R	TCCTTTTCCAACTCCCAGGG			
11 F	AAGGTGGGGATAGATGCTGG	64	1.5	348
11 R	AGTCTTGGTCCTCAGTGCTC			
12 F	TGACCTCCAGTGTTCCGC	64	1.5	399
12 R	ATCTGCAGGGCTCCATAACC			
13 F	CTATGCGTGTCCCCTGGG	64	1.5	487
13 R	GGGCAGATGGACACACAGAT			
14 F	CAGAGCCAGGTGAAGAGGG	64	1.5	679
14 R	CCTGGGCAAGAGTGAGACT			
15 F	GAACCTCTCTGTCCCCTCC	64	1.5	700
15 R	GCCAGGGAAGGTGAGGTTAT			
16 F	TCTAGGTTGGGGAAGATGCT	64	1.5	678
16 R	CAGGAGGTGGAGGTTGCAAT			

9.6 LIST OF PROTEINS IDENTIFIED THROUGH QTOF ANALYSIS OF THE IVT KIT

Accession number	Protein name
A0A024QZX5	Serpin B6
A0A024R4E5	High density lipoprotein binding protein (Vigilin), isoform CRA_a
A0A024R571	EH domain-containing protein 1
A0A075B752	Annexin
A0A087WSW2	Zinc finger protein 43
A0A087WSW9	Thioredoxin reductase 1, cytoplasmic
A0A087WSY9	Thioredoxin reductase 1, cytoplasmic
A0A087WTT1	Polyadenylate-binding protein
A0A087WUD7	Sec1 family domain-containing protein 2
A0A087WUL0	Bifunctional ATP-dependent dihydroxyacetone kinase/FAD-AMP lyase (cyclizing)
A0A087WUL2	Proteasome subunit beta type-3 (Fragment)
A0A087WUS0	40S ribosomal protein S24
A0A087WV23	SH3 domain-binding glutamic acid-rich-like protein 3
A0A087WV47	Ig gamma-1 chain C region
A0A087WV48	Uncharacterized protein
A0A087WV55	Complement C1q tumor necrosis factor-related protein 8
A0A087WV77	Ubiquitin-40S ribosomal protein S27a
A0A087WVN4	Farnesyl pyrophosphate synthase (Fragment)
A0A087WVQ6	Clathrin heavy chain
A0A087WW66	26S proteasome non-ATPase regulatory subunit 1
A0A087WWU8	Tropomyosin alpha-3 chain
A0A087WXM6	60S ribosomal protein L17 (Fragment)
A0A087WY61	Nuclear mitotic apparatus protein 1
A0A087WYR3	Tumor protein D54
A0A087WYT3	Prostaglandin E synthase 3
A0A087WZ27	Zinc finger protein 90
A0A087WZ65	Ankyrin-3 (Fragment)
A0A087WZV1	Heterogeneous nuclear ribonucleoprotein A/B
A0A087X0N0	Trafficking kinesin-binding protein 1
A0A087X0X3	Heterogeneous nuclear ribonucleoprotein M
A0A087X1I3	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial
A0A087X1S2	Nuclease-sensitive element-binding protein 1
A0A087X1X7	Elongation factor 1-delta
A0A087X1Z3	Proteasome activator complex subunit 2
A0A087X2D0	Serine/arginine-rich-splicing factor 3
A0A087X2G1	ATP-dependent RNA helicase DDX1
A0A087X2I1	26S protease regulatory subunit 10B
A0A0A0MQW3	Serpin B13
A0A0A0MR45	Calpastatin
A0A0A0MR47	Neurotrophin receptor-interacting factor homolog
A0A0A0MRI6	Regulator of G-protein-signaling 7
A0A0A0MRK8	Fucose mutarotase
A0A0A0MRM2	Nebulin-related-anchoring protein
A0A0A0MSI0	Peroxiredoxin-1 (Fragment)
A0A0A0MSN4	Angiotensin-converting enzyme
A0A0A0MSN9	Negative elongation factor E (Fragment)
A0A0A0MSQ0	Plastin-3
A0A0A0MSR7	Protein TBATA
A0A0A0MSS8	Aldo-keto reductase family 1 member C3
A0A0A0MSW4	Phosphatidylinositol transfer protein beta isoform
A0A0A0MTN3	Glutathione S-transferase Mu 3
A0A0A0MTS2	Glucose-6-phosphate isomerase (Fragment)
A0A0A6YYG9	Protein ARPC4-TTLL3
A0A0B4J1Z1	Serine/arginine-rich-splicing factor 7
A0A0B4J2A4	3-ketoacyl-CoA thiolase, mitochondrial
A0A0B4J2B4	40S ribosomal protein S15
A0A0B4J2C3	Translationally-controlled tumor protein
A0A0B4J2F2	Protein LOC102724428
A5A3E0	POTE ankyrin domain family member F
A6NCQ0	ADP-sugar pyrophosphatase
A6NI03	Putative tripartite motif-containing protein 64B
A6NLN1	Polypyrimidine tract binding protein 1, isoform CRA_b
A6NM15	Putative COBW domain-containing protein 7
A6ZIE3	MUC1 isoform M9
A8MU27	Small ubiquitin-related modifier 3
A8MU58	Aminoacyl tRNA synthase complex-interacting multifunctional protein 2
A8MUS3	60S ribosomal protein L23a

A8MXH2	Nucleosome assembly protein 1-like 4 (Fragment)
A8MXP9	Matrin-3
A8MZ26	EF-hand calcium-binding domain-containing protein 9
B0QY89	Eukaryotic translation initiation factor 3 subunit L
B0QYC2	Interleukin-2 receptor subunit beta (Fragment)
B0QZ18	Copine-1
B1AK85	F-actin-capping protein subunit beta
B1AKR6	Dynein light chain roadblock-type 1
B1ALD0	AP complex subunit beta
B1ANR0	Polyadenylate-binding protein
B3KNJ4	SUMO-1 activating enzyme subunit 1, isoform CRA_a
B3KQ25	Proteasome activator complex subunit 3
B4DQU5	Ras-related protein Rab-11A
B4DUR8	T-complex protein 1 subunit gamma
B4DXD0	Translin-associated factor X-interacting protein 1
B4DXW1	Actin-related protein 3
B5MCI0	Smoothelin
B5MCX3	Septin-2
B5MDF5	GTP-binding nuclear protein Ran
B5ME19	Eukaryotic translation initiation factor 3 subunit C-like protein
B7Z4W5	Cysteine conjugate-beta lyase cytoplasmic (Glutamine transaminase K, kyneurenine aminotransferase), isoform CRA_b
B7Z574	Calpastatin
B7Z645	Heterogeneous nuclear ribonucleoprotein Q
B7Z6Z4	Myosin light polypeptide 6
B7Z7P8	Eukaryotic peptide chain release factor subunit 1
B8ZZ45	Glycosyltransferase-like domain-containing protein 1
B8ZZA1	Prothymosin alpha
B8ZZL8	10 kDa heat shock protein, mitochondrial
B9A041	Malate dehydrogenase, cytoplasmic
B9ZVP7	60S ribosomal protein L23
C9IYI6	Tyrosine-protein phosphatase non-receptor type 7 (Fragment)
C9J0D1	Histone H2A
C9J0K6	Sorcin
C9J0Q5	Nuclear receptor corepressor 2
C9J1T2	Transforming protein RhoA (Fragment)
C9J1V9	HCG2043275
C9J1Z8	ADP-ribosylation factor 5 (Fragment)
C9J4W5	Eukaryotic translation initiation factor 5A (Fragment)
C9J6D1	Nucleosome assembly protein 1-like 4 (Fragment)
C9J6G3	Geranylgeranyl pyrophosphate synthase (Fragment)
C9J9K3	40S ribosomal protein SA (Fragment)
C9J9W2	LIM and SH3 domain protein 1 (Fragment)
C9JA08	60S ribosomal export protein NMD3
C9JIF9	Acylamino-acid-releasing enzyme
C9JJ34	Ran-specific GTPase-activating protein (Fragment)
C9JNW5	60S ribosomal protein L24
C9JY79	Non-erythrocytic beta-spectrin 4
C9JZI7	Nucleosome assembly protein 1-like 4 (Fragment)
D3DSM8	Formimidoyltransferase-cyclodeaminase
D3DVN5	Transmembrane protein 175
D3YTB1	60S ribosomal protein L32 (Fragment)
D6R967	Inorganic pyrophosphatase 2, mitochondrial (Fragment)
D6RA82	Annexin
D6RAF8	Heterogeneous nuclear ribonucleoprotein D0 (Fragment)
D6RAN4	60S ribosomal protein L9 (Fragment)
D6RDG3	Transcription factor BTF3 (Fragment)
D6RFW5	UDP-glucuronosyltransferase 2A1
D6RG13	40S ribosomal protein S3a (Fragment)
D6RJC3	Type II inositol 3,4-bisphosphate 4-phosphatase (Fragment)
E5RFP0	NudC domain-containing protein 2
E5RHG6	Tubulin-specific chaperone A
E5RI99	60S ribosomal protein L30 (Fragment)
E5RIT6	60S ribosomal protein L26-like 1 (Fragment)
E5RJH5	Beta-enolase
E5RJR5	S-phase kinase-associated protein 1
E7ENH0	Pregnancy-specific beta-1-glycoprotein 8
E7EPB3	60S ribosomal protein L14
E7EPB6	Cystic fibrosis transmembrane conductance regulator (Fragment)
E7EQG2	Eukaryotic initiation factor 4A-II
E7EQL5	Cytoplasmic dynein 1 intermediate chain 2 (Fragment)
E7EQV9	Ribosomal protein L15 (Fragment)
E7ERU0	Dystonin
E7ERW8	Protein diaphanous homolog 1
E7EUU4	Eukaryotic translation initiation factor 4 gamma 1
E7EVA0	Microtubule-associated protein
E9PAV3	Nascent polypeptide-associated complex subunit alpha, muscle-specific form
E9PB90	Hexokinase

E9PCP0	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-3
E9PCY7	Heterogeneous nuclear ribonucleoprotein H
E9PDE8	Heat shock 70 kDa protein 4L
E9PEZ3	Protein diaphanous homolog 1
E9PHA6	DNA mismatch repair protein Msh2
E9PIY6	ADP-ribosylation factor GTPase-activating protein 2 (Fragment)
E9PIZ5	Thioredoxin reductase 1, cytoplasmic (Fragment)
E9PJD9	60S ribosomal protein L27a
E9PJK5	FTS and Hook-interacting protein
E9PK25	Cofilin-1
E9PKD5	26S protease regulatory subunit 6A (Fragment)
E9PKG1	Protein arginine N-methyltransferase 1
E9PKW4	Sulfotransferase
E9PKZ0	60S ribosomal protein L8 (Fragment)
E9PLG2	26S protease regulatory subunit 6A (Fragment)
E9PLH9	GDP-L-fucose synthase (Fragment)
E9PLI6	Probable RNA-binding protein EIF1AD (Fragment)
E9PLK3	Puromycin-sensitive aminopeptidase
E9PM95	Apolipoprotein L2 (Fragment)
E9PMI6	Methylosome subunit pICln
E9PNW8	Fatty acyl-CoA reductase 1 (Fragment)
E9PPP3	Phosphatidylinositol 4-phosphate 3-kinase C2 domain-containing subunit alpha (Fragment)
E9PPV6	Serpin H1
E9PR30	40S ribosomal protein S30
E9PR53	Protein FAM212B
E9PRQ5	Puromycin-sensitive aminopeptidase (Fragment)
E9PRQ7	UBX domain-containing protein 1
E9PRY8	Elongation factor 1-delta
F2Z2V0	Copine-1 (Fragment)
F2Z2Y4	Pyridoxal kinase
F2Z388	60S ribosomal protein L35
F2Z393	Transaldolase
F5GY55	DNA damage-binding protein 1
F5GYZ6	LIM domain only protein 3 (Fragment)
F5H039	Gephyrin
F5H126	Synaptotagmin-7
F5H4R6	Nucleosome assembly protein 1-like 1
F5H5G6	Fanconi anemia-associated protein of 100 kDa (Fragment)
F5H6X6	Neutral alpha-glucosidase AB
F5H7S3	Tropomyosin alpha-1 chain
F5H7Y1	T-complex protein 1 subunit alpha (Fragment)
F6U211	40S ribosomal protein S10
F6X3S4	Uncharacterized protein
F8VPD4	CAD protein
F8VRK7	60S acidic ribosomal protein P0
F8VTQ5	Heterogeneous nuclear ribonucleoprotein A1 (Fragment)
F8VXH9	Poly(rC)-binding protein 2 (Fragment)
F8VXI9	ARF GTPase-activating protein GIT2
F8VY02	Endoplasmic reticulum resident protein 29
F8VZ49	Heterogeneous nuclear ribonucleoprotein A1 (Fragment)
F8VZX2	Poly(rC)-binding protein 2
F8W1A4	Adenylate kinase 2, mitochondrial
F8W1R7	Myosin light polypeptide 6
F8W6G6	Leukocyte immunoglobulin-like receptor subfamily B member 3
F8W775	Neuronal cell adhesion molecule
F8W9U3	Alpha/beta hydrolase domain-containing protein 14B
F8W9U4	Microtubule-associated protein
F8WBD4	Zinc finger protein with KRAB and SCAN domains 5
G3V1C5	Interleukin 18 binding protein, isoform CRA_a
G3V203	60S ribosomal protein L18
G3V295	Proteasome subunit alpha type
G3V3G9	DDB1- and CUL4-associated factor 8
G3V483	G2/M phase-specific E3 ubiquitin-protein ligase
G3V5F4	Mitochondrial basic amino acids transporter (Fragment)
G5E9C7	Dual-specificity mitogen-activated protein kinase kinase 2
G5E9G0	60S ribosomal protein L3
G8JLD5	Dynamin-1-like protein
H0Y2S9	Myosin phosphatase Rho-interacting protein (Fragment)
H0Y325	Nesprin-1 (Fragment)
H0Y586	Proteasome subunit alpha type-7 (Fragment)
H0Y5B4	60S ribosomal protein L36a
H0Y5H9	Serpin B4 (Fragment)
H0Y8C6	Importin-5 (Fragment)
H0YAK9	Nephronectin (Fragment)
H0YCQ8	Eukaryotic translation initiation factor 3 subunit M (Fragment)
H0YEM1	Poly(U)-binding-splicing factor PUF60 (Fragment)
H0YFX9	Histone H2A (Fragment)

H0YGI8	Stress-induced-phosphoprotein 1 (Fragment)
H0YI98	Dynactin subunit 2 (Fragment)
H0YIV4	Nucleosome assembly protein 1-like 1 (Fragment)
H0YJC0	26S protease regulatory subunit 10B (Fragment)
H0YKC5	Deoxyuridine 5'-triphosphate nucleotidohydrolase, mitochondrial (Fragment)
H0YKN4	Annexin A2
H0YL52	Tropomyosin alpha-1 chain (Fragment)
H0YL69	Proteasome subunit alpha type (Fragment)
H0YL80	Tropomyosin alpha-1 chain (Fragment)
H0YLA4	Sorbitol dehydrogenase
H0YN26	Acidic leucine-rich nuclear phosphoprotein 32 family member A
H0YNG5	Alpha-mannosidase
H3BPS8	Fructose-bisphosphate aldolase (Fragment)
H3BQI1	Dynein light chain roadblock-type 2
H3BQZ9	Adenine phosphoribosyltransferase
H3BR35	Eukaryotic peptide chain release factor GTP-binding subunit ERF3A (Fragment)
H7BXP1	NF-kappa-B inhibitor-interacting Ras-like protein 2
H7C084	Bifunctional purine biosynthesis protein PURH (Fragment)
H7C0A3	Protein ARPC4-TTLL3 (Fragment)
H7C144	Alpha-actinin-4 (Fragment)
H7C307	Ubiquitin carboxyl-terminal hydrolase 40 (Fragment)
I3L0K2	Thioredoxin domain-containing protein 17
I3L121	Clusterin-associated protein 1 (Fragment)
I3L1Y9	FLYWCH family member 2
I3L2F9	Uncharacterized protein
I3L397	Eukaryotic translation initiation factor 5A (Fragment)
I6L894	Ankyrin-2
I6L9D5	FAM114A2 protein
J3KN39	NACHT, LRR and PYD domains-containing protein 2
J3KNQ3	26S proteasome non-ATPase regulatory subunit 13
J3KPE3	Guanine nucleotide-binding protein subunit beta-2-like 1
J3KQ18	D-dopachrome decarboxylase
J3KQ32	Obg-like ATPase 1
J3KQG3	EPH receptor A10, isoform CRA_b
J3KR24	Isoleucine--tRNA ligase, cytoplasmic
J3KRH3	Hexosaminidase D (Fragment)
J3KRP9	Protein TANC2 (Fragment)
J3KSD5	Lethal(2) giant larvae protein homolog 2 (Fragment)
J3KSH8	Hematological and neurological-expressed 1 protein (Fragment)
J3KTE4	Ribosomal protein L19
J3KTF8	Rho GDP-dissociation inhibitor 1 (Fragment)
J3QLI9	Small nuclear ribonucleoprotein Sm D1
K7EJ78	40S ribosomal protein S15
K7EJR3	26S proteasome non-ATPase regulatory subunit 8 (Fragment)
K7ELJ7	Calpain small subunit 1
K7ELL7	Glucosidase 2 subunit beta
K7EMW4	Nicalin
K7EP46	Thimet oligopeptidase
K7EPC4	Histone acetyltransferase KAT2A (Fragment)
K7EQA9	Hsp90 co-chaperone Cdc37 (Fragment)
K7EQJ8	GTPase Era, mitochondrial (Fragment)
K7EQX3	AP-1 complex subunit mu-1 (Fragment)
K7ESK7	SUMO-activating enzyme subunit 2 (Fragment)
K7ESP6	Queuine tRNA-ribosyltransferase (Fragment)
M0QXK4	40S ribosomal protein S19 (Fragment)
M0QYI7	Protein fuzzy homolog
M0QZN2	40S ribosomal protein S5
M0R080	DnaJ homolog subfamily B member 1 (Fragment)
M0R0E2	Zinc finger protein 432 (Fragment)
M0R0F0	40S ribosomal protein S5 (Fragment)
M0R0P7	60S ribosomal protein L18a
M0R0P8	Unconventional myosin-IXb
M0R117	60S ribosomal protein L18a
M0R246	Zinc finger protein 134
O00148	ATP-dependent RNA helicase DDX39A
O00170	AH receptor-interacting protein
O00204	Sulfotransferase family cytosolic 2B member 1
O00231	26S proteasome non-ATPase regulatory subunit 11
O00232	26S proteasome non-ATPase regulatory subunit 12
O00299	Chloride intracellular channel protein 1
O00303	Eukaryotic translation initiation factor 3 subunit F
O00429	Dynamin-1-like protein
O00487	26S proteasome non-ATPase regulatory subunit 14
O00571	ATP-dependent RNA helicase DDX3X
O14818	Proteasome subunit alpha type-7
O14929	Histone acetyltransferase type B catalytic subunit
O14980	Exportin-1

O15031	Plexin-B2
O15050	TPR and ankyrin repeat-containing protein 1
O15061	Synemin
O15067	Phosphoribosylformylglycinamidine synthase
O15144	Actin-related protein 2/3 complex subunit 2
O15260	Surfeit locus protein 4
O15371	Eukaryotic translation initiation factor 3 subunit D
O43143	Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15
O43175	D-3-phosphoglycerate dehydrogenase
O43182	Rho GTPase-activating protein 6
O43237	Cytoplasmic dynein 1 light intermediate chain 2
O43242	26S proteasome non-ATPase regulatory subunit 3
O43399	Tumor protein D54
O43432	Eukaryotic translation initiation factor 4 gamma 3
O43707	Alpha-actinin-4
O43776	Asparagine--tRNA ligase, cytoplasmic
O60256	Phosphoribosyl pyrophosphate synthase-associated protein 2
O60506	Heterogeneous nuclear ribonucleoprotein Q
O60664	Perilipin-3
O60701	UDP-glucose 6-dehydrogenase
O60814	Histone H2B type 1-K
O60869	Endothelial differentiation-related factor 1
O60884	DnaJ homolog subfamily A member 2
O60942	mRNA-capping enzyme
O75145	Liprin-alpha-3
O75306	NADH dehydrogenase [ubiquinone] iron-sulfur protein 2, mitochondrial
O75340	Programmed cell death protein 6
O75369	Filamin-B
O75531	Barrier-to-autointegration factor
O75874	Isocitrate dehydrogenase [NADP] cytoplasmic
O76003	Glutaredoxin-3
O94915	Protein furry homolog-like
O95336	6-phosphogluconolactonase
O95373	Importin-7
O95433	Activator of 90 kDa heat shock protein ATPase homolog 1
O95833	Chloride intracellular channel protein 3
O96019	Actin-like protein 6A
P00338	L-lactate dehydrogenase A chain
P00374	Dihydrofolate reductase
P00441	Superoxide dismutase [Cu-Zn]
P00488	Coagulation factor XIII A chain
P00491	Purine nucleoside phosphorylase
P00492	Hypoxanthine-guanine phosphoribosyltransferase
P00558	Phosphoglycerate kinase 1
P00568	Adenylate kinase isoenzyme 1
P00966	Argininosuccinate synthase
P01011	Alpha-1-antichymotrypsin
P02545	Prelamin-A/C
P02751	Fibronectin
P02786	Transferrin receptor protein 1
P04075	Fructose-bisphosphate aldolase A
P04080	Cystatin-B
P04083	Annexin A1
P04350	Tubulin beta-4A chain
P04406	Glyceraldehyde-3-phosphate dehydrogenase
P04792	Heat shock protein beta-1
P04818	Thymidylate synthase
P05198	Eukaryotic translation initiation factor 2 subunit 1
P05386	60S acidic ribosomal protein P1
P05387	60S acidic ribosomal protein P2
P05388	60S acidic ribosomal protein P0
P05455	Lupus La protein
P06132	Uroporphyrinogen decarboxylase
P06576	ATP synthase subunit beta, mitochondrial
P06703	Protein S100-A6
P06732	Creatine kinase M-type
P06733	Alpha-enolase
P06748	Nucleophosmin
P07195	L-lactate dehydrogenase B chain
P07237	Protein disulfide-isomerase
P07355	Annexin A2
P07437	Tubulin beta chain
P07737	Profilin-1
P07741	Adenine phosphoribosyltransferase
P07814	Bifunctional glutamate/proline--tRNA ligase
P07900	Heat shock protein HSP 90-alpha
P07951	Tropomyosin beta chain

P07954	Fumarate hydratase, mitochondrial
P08107	Heat shock 70 kDa protein 1A/1B
P08238	Heat shock protein HSP 90-beta
P08243	Asparagine synthetase [glutamine-hydrolyzing]
P08708	40S ribosomal protein S17
P08758	Annexin A5
P09104	Gamma-enolase
P09960	Leukotriene A-4 hydrolase
P09972	Fructose-bisphosphate aldolase C
P0CG38	POTE ankyrin domain family member I
P10599	Thioredoxin
P10632	Cytochrome P450 2C8
P10809	60 kDa heat shock protein, mitochondrial
P11021	78 kDa glucose-regulated protein
P11142	Heat shock cognate 71 kDa protein
P11413	Glucose-6-phosphate 1-dehydrogenase
P11586	C-1-tetrahydrofolate synthase, cytoplasmic
P11766	Alcohol dehydrogenase class-3
P11908	Ribose-phosphate pyrophosphokinase 2
P11940	Polyadenylate-binding protein 1
P12004	Proliferating cell nuclear antigen
P12268	Inosine-5'-monophosphate dehydrogenase 2
P12814	Alpha-actinin-1
P12956	X-ray repair cross-complementing protein 6
P13010	X-ray repair cross-complementing protein 5
P13489	Ribonuclease inhibitor
P13639	Elongation factor 2
P13667	Protein disulfide-isomerase A4
P13804	Electron transfer flavoprotein subunit alpha, mitochondrial
P13929	Beta-enolase
P14174	Macrophage migration inhibitory factor
P14618	Pyruvate kinase PKM
P14625	Endoplasmin
P14868	Aspartate--tRNA ligase, cytoplasmic
P15085	Carboxypeptidase A1
P15121	Aldose reductase
P15311	Ezrin
P15531	Nucleoside diphosphate kinase A
P15880	40S ribosomal protein S2
P15927	Replication protein A 32 kDa subunit
P16152	Carbonyl reductase [NADPH] 1
P17066	Heat shock 70 kDa protein 6
P17174	Aspartate aminotransferase, cytoplasmic
P17655	Calpain-2 catalytic subunit
P17812	CTP synthase 1
P17844	Probable ATP-dependent RNA helicase DDX5
P17858	ATP-dependent 6-phosphofructokinase, liver type
P17980	26S protease regulatory subunit 6A
P17987	T-complex protein 1 subunit alpha
P18124	60S ribosomal protein L7
P18206	Vinculin
P18669	Phosphoglycerate mutase 1
P19338	Nucleolin
P20042	Eukaryotic translation initiation factor 2 subunit 2
P20073	Annexin A7
P20290	Transcription factor BTF3
P20618	Proteasome subunit beta type-1
P20848	Putative alpha-1-antitrypsin-related protein
P20929	Nebulin
P21333	Filamin-A
P21980	Protein-glutamine gamma-glutamyltransferase 2
P22102	Trifunctional purine biosynthetic protein adenosine-3
P22234	Multifunctional protein ADE2
P22314	Ubiquitin-like modifier-activating enzyme 1
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1
P23025	DNA repair protein complementing XP-A cells
P23258	Tubulin gamma-1 chain
P23284	Peptidyl-prolyl cis-trans isomerase B
P23378	Glycine dehydrogenase (decarboxylating), mitochondrial
P23381	Tryptophan--tRNA ligase, cytoplasmic
P23396	40S ribosomal protein S3
P23526	Adenosylhomocysteinase
P23921	Ribonucleoside-diphosphate reductase large subunit
P24534	Elongation factor 1-beta
P24666	Low molecular weight phosphotyrosine protein phosphatase
P25705	ATP synthase subunit alpha, mitochondrial
P25786	Proteasome subunit alpha type-1

P25787	Proteasome subunit alpha type-2
P25788	Proteasome subunit alpha type-3
P25815	Protein S100-P
P26038	Moesin
P26373	60S ribosomal protein L13
P26447	Protein S100-A4
P26639	Threonine--tRNA ligase, cytoplasmic
P26640	Valine--tRNA ligase
P26641	Elongation factor 1-gamma
P27348	14-3-3 protein theta
P27635	60S ribosomal protein L10
P27695	DNA-(apurinic or apyrimidinic site) lyase
P27797	Calreticulin
P28065	Proteasome subunit beta type-9
P28066	Proteasome subunit alpha type-5
P28070	Proteasome subunit beta type-4
P28074	Proteasome subunit beta type-5
P29350	Tyrosine-protein phosphatase non-receptor type 6
P29401	Transketolase
P29508	Serpin B3
P30041	Peroxiredoxin-6
P30043	Flavin reductase (NADPH)
P30044	Peroxiredoxin-5, mitochondrial
P30046	D-dopachrome decarboxylase
P30048	Thioredoxin-dependent peroxide reductase, mitochondrial
P30050	60S ribosomal protein L12
P30084	Enoyl-CoA hydratase, mitochondrial
P30085	UMP-CMP kinase
P30086	Phosphatidylethanolamine-binding protein 1
P30101	Protein disulfide-isomerase A3
P30153	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform
P30740	Leukocyte elastase inhibitor
P31150	Rab GDP dissociation inhibitor alpha
P31327	Carbamoyl-phosphate synthase [ammonia], mitochondrial
P31689	DnaJ homolog subfamily A member 1
P31939	Bifunctional purine biosynthesis protein PURH
P31946	14-3-3 protein beta/alpha
P31947	14-3-3 protein sigma
P31948	Stress-induced-phosphoprotein 1
P31949	Protein S100-A11
P32119	Peroxiredoxin-2
P32320	Cytidine deaminase
P32942	Intercellular adhesion molecule 3
P32969	60S ribosomal protein L9
P33991	DNA replication licensing factor MCM4
P33993	DNA replication licensing factor MCM7
P34897	Serine hydroxymethyltransferase, mitochondrial
P34931	Heat shock 70 kDa protein 1-like
P34932	Heat shock 70 kDa protein 4
P35221	Catenin alpha-1
P35241	Radixin
P35268	60S ribosomal protein L22
P35270	Sepiapterin reductase
P35520	Cystathionine beta-synthase
P35579	Myosin-9
P35580	Myosin-10
P35913	Rod cGMP-specific 3',5'-cyclic phosphodiesterase subunit beta
P35998	26S protease regulatory subunit 7
P36578	60S ribosomal protein L4
P36871	Phosphoglucomutase-1
P36952	Serpin B5
P37108	Signal recognition particle 14 kDa protein
P37802	Transgelin-2
P38117	Electron transfer flavoprotein subunit beta
P38606	V-type proton ATPase catalytic subunit A
P38646	Stress-70 protein, mitochondrial
P38919	Eukaryotic initiation factor 4A-III
P39019	40S ribosomal protein S19
P39023	60S ribosomal protein L3
P39748	Flap endonuclease 1
P40121	Macrophage-capping protein
P40227	T-complex protein 1 subunit zeta
P40426	Pre-B-cell leukemia transcription factor 3
P40763	Signal transducer and activator of transcription 3
P40925	Malate dehydrogenase, cytoplasmic
P40926	Malate dehydrogenase, mitochondrial
P41091	Eukaryotic translation initiation factor 2 subunit 3

P41250	Glycine--tRNA ligase
P41252	Isoleucine--tRNA ligase, cytoplasmic
P42338	Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit beta isoform
P42695	Condensin-2 complex subunit D3
P43121	Cell surface glycoprotein MUC18
P43490	Nicotinamide phosphoribosyltransferase
P43626	Killer cell immunoglobulin-like receptor 2DL1
P43686	26S protease regulatory subunit 6B
P45974	Ubiquitin carboxyl-terminal hydrolase 5
P46777	60S ribosomal protein L5
P46778	60S ribosomal protein L21
P46779	60S ribosomal protein L28
P46781	40S ribosomal protein S9
P46940	Ras GTPase-activating-like protein IQGAP1
P47712	Cytosolic phospholipase A2
P47897	Glutamine--tRNA ligase
P47914	60S ribosomal protein L29
P48357	Leptin receptor
P48643	T-complex protein 1 subunit epsilon
P48741	Putative heat shock 70 kDa protein 7
P49189	4-trimethylaminobutyraldehyde dehydrogenase
P49321	Nuclear autoantigenic sperm protein
P49327	Fatty acid synthase
P49368	T-complex protein 1 subunit gamma
P49411	Elongation factor Tu, mitochondrial
P49588	Alanine--tRNA ligase, cytoplasmic
P49721	Proteasome subunit beta type-2
P49756	RNA-binding protein 25
P49773	Histidine triad nucleotide-binding protein 1
P49915	GMP synthase [glutamine-hydrolyzing]
P50395	Rab GDP dissociation inhibitor beta
P50454	Serpin H1
P50747	Biotin--protein ligase
P50990	T-complex protein 1 subunit theta
P50991	T-complex protein 1 subunit delta
P51665	26S proteasome non-ATPase regulatory subunit 7
P51858	Hepatoma-derived growth factor
P52209	6-phosphogluconate dehydrogenase, decarboxylating
P52597	Heterogeneous nuclear ribonucleoprotein F
P52788	Spermine synthase
P52888	Thimet oligopeptidase
P52907	F-actin-capping protein subunit alpha-1
P53396	ATP-citrate synthase
P53675	Clathrin heavy chain 2
P53814	Smoothelin
P53999	Activated RNA polymerase II transcriptional coactivator p15
P54136	Arginine--tRNA ligase, cytoplasmic
P54577	Tyrosine--tRNA ligase, cytoplasmic
P54652	Heat shock-related 70 kDa protein 2
P54725	UV excision repair protein RAD23 homolog A
P54727	UV excision repair protein RAD23 homolog B
P55010	Eukaryotic translation initiation factor 5
P55060	Exportin-2
P55072	Transitional endoplasmic reticulum ATPase
P55795	Heterogeneous nuclear ribonucleoprotein H2
P55884	Eukaryotic translation initiation factor 3 subunit B
P56192	Methionine--tRNA ligase, cytoplasmic
P56211	cAMP-regulated phosphoprotein 19
P56537	Eukaryotic translation initiation factor 6
P57077	MAP3K7 C-terminal-like protein
P57737	Coronin-7
P60174	Triosephosphate isomerase
P60228	Eukaryotic translation initiation factor 3 subunit E
P60709	Actin, cytoplasmic 1
P60842	Eukaryotic initiation factor 4A-I
P60866	40S ribosomal protein S20
P60891	Ribose-phosphate pyrophosphokinase 1
P60900	Proteasome subunit alpha type-6
P60903	Protein S100-A10
P60953	Cell division control protein 42 homolog
P61006	Ras-related protein Rab-8A
P61088	Ubiquitin-conjugating enzyme E2 N
P61160	Actin-related protein 2
P61163	Alpha-centractin
P61204	ADP-ribosylation factor 3
P61221	ATP-binding cassette sub-family E member 1
P61247	40S ribosomal protein S3a

P61353	60S ribosomal protein L27
P61604	10 kDa heat shock protein, mitochondrial
P61758	Prefoldin subunit 3
P61970	Nuclear transport factor 2
P61978	Heterogeneous nuclear ribonucleoprotein K
P61981	14-3-3 protein gamma
P62081	40S ribosomal protein S7
P62136	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit
P62191	26S protease regulatory subunit 4
P62195	26S protease regulatory subunit 8
P62241	40S ribosomal protein S8
P62244	40S ribosomal protein S15a
P62249	40S ribosomal protein S16
P62258	14-3-3 protein epsilon
P62263	40S ribosomal protein S14
P62266	40S ribosomal protein S23
P62269	40S ribosomal protein S18
P62277	40S ribosomal protein S13
P62280	40S ribosomal protein S11
P62304	Small nuclear ribonucleoprotein E
P62306	Small nuclear ribonucleoprotein F
P62318	Small nuclear ribonucleoprotein Sm D3
P62328	Thymosin beta-4
P62424	60S ribosomal protein L7a
P62701	40S ribosomal protein S4, X isoform
P62714	Serine/threonine-protein phosphatase 2A catalytic subunit beta isoform
P62753	40S ribosomal protein S6
P62829	60S ribosomal protein L23
P62851	40S ribosomal protein S25
P62854	40S ribosomal protein S26
P62857	40S ribosomal protein S28
P62899	60S ribosomal protein L31
P62906	60S ribosomal protein L10a
P62913	60S ribosomal protein L11
P62937	Peptidyl-prolyl cis-trans isomerase A
P62942	Peptidyl-prolyl cis-trans isomerase FKBP1A
P62979	Ubiquitin-40S ribosomal protein S27a
P63104	14-3-3 protein zeta/delta
P63151	Serine/threonine-protein phosphatase 2A 55 kDa regulatory subunit B alpha isoform
P63220	40S ribosomal protein S21
P63244	Guanine nucleotide-binding protein subunit beta-2-like 1
P63267	Actin, gamma-enteric smooth muscle
P67775	Serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform
P67809	Nuclease-sensitive element-binding protein 1
P67936	Tropomyosin alpha-4 chain
P68036	Ubiquitin-conjugating enzyme E2 L3
P68104	Elongation factor 1-alpha 1
P68363	Tubulin alpha-1B chain
P68366	Tubulin alpha-4A chain
P68371	Tubulin beta-4B chain
P78371	T-complex protein 1 subunit beta
P78417	Glutathione S-transferase omega-1
P78527	DNA-dependent protein kinase catalytic subunit
P80723	Brain acid soluble protein 1
P82970	High mobility group nucleosome-binding domain-containing protein 5
Q00526	Cyclin-dependent kinase 3
Q00688	Peptidyl-prolyl cis-trans isomerase FKBP3
Q00796	Sorbitol dehydrogenase
Q00839	Heterogeneous nuclear ribonucleoprotein U
Q00889	Pregnancy-specific beta-1-glycoprotein 6
Q01105	Protein SET
Q01469	Fatty acid-binding protein, epidermal
Q01518	Adenylyl cyclase-associated protein 1
Q01813	ATP-dependent 6-phosphofructokinase, platelet type
Q02750	Dual specificity mitogen-activated protein kinase kinase 1
Q02790	Peptidyl-prolyl cis-trans isomerase FKBP4
Q02878	60S ribosomal protein L6
Q04446	1,4-alpha-glucan-branching enzyme
Q04760	Lactoylglutathione lyase
Q04828	Aldo-keto reductase family 1 member C1
Q04917	14-3-3 protein eta
Q04941	Proteolipid protein 2
Q05586	Glutamate receptor ionotropic, NMDA 1
Q05639	Elongation factor 1-alpha 2
Q05BV3	Echinoderm microtubule-associated protein-like 5
Q06323	Proteasome activator complex subunit 1
Q06830	Peroxiredoxin-1

Q08J23	tRNA (cytosine(34)-C(5))-methyltransferase
Q12774	Rho guanine nucleotide exchange factor 5
Q12931	Heat shock protein 75 kDa, mitochondrial
Q12955	Ankyrin-3
Q12965	Unconventional myosin-Ie
Q13057	Bifunctional coenzyme A synthase
Q13136	Liprin-alpha-1
Q13162	Peroxiredoxin-4
Q13200	26S proteasome non-ATPase regulatory subunit 2
Q13263	Transcription intermediary factor 1-beta
Q13283	Ras GTPase-activating protein-binding protein 1
Q13315	Serine-protein kinase ATM
Q13347	Eukaryotic translation initiation factor 3 subunit I
Q13409	Cytoplasmic dynein 1 intermediate chain 2
Q13509	Tubulin beta-3 chain
Q13541	Eukaryotic translation initiation factor 4E-binding protein 1
Q13885	Tubulin beta-2A chain
Q14005	Pro-interleukin-16
Q14011	Cold-inducible RNA-binding protein
Q14019	Coactosin-like protein
Q14116	Interleukin-18
Q14152	Eukaryotic translation initiation factor 3 subunit A
Q14195	Dihydropyrimidinase-related protein 3
Q14204	Cytoplasmic dynein 1 heavy chain 1
Q14240	Eukaryotic initiation factor 4A-II
Q14258	E3 ubiquitin/ISG15 ligase TRIM25
Q14315	Filamin-C
Q14376	UDP-glucose 4-epimerase
Q14444	Caprin-1
Q14568	Putative heat shock protein HSP 90-alpha A2
Q14669	E3 ubiquitin-protein ligase TRIP12
Q14674	Separin
Q14687	Genetic suppressor element 1
Q14764	Major vault protein
Q14774	H2.0-like homeobox protein
Q14974	Importin subunit beta-1
Q15021	Condensin complex subunit 1
Q15029	116 kDa U5 small nuclear ribonucleoprotein component
Q15046	Lysine--tRNA ligase
Q15056	Eukaryotic translation initiation factor 4H
Q15084	Protein disulfide-isomerase A6
Q15181	Inorganic pyrophosphatase
Q15365	Poly(rC)-binding protein 1
Q15366	Poly(rC)-binding protein 2
Q15691	Microtubule-associated protein RP/EB family member 1
Q15843	NEDD8
Q16543	Hsp90 co-chaperone Cdc37
Q16555	Dihydropyrimidinase-related protein 2
Q16576	Histone-binding protein RBBP7
Q16658	Fascin
Q16719	Kynureninase
Q2VPJ9	Leucine-rich repeat-containing protein 75B
Q32Q12	Nucleoside diphosphate kinase
Q3V6T2	Girdin
Q3ZCM7	Tubulin beta-8 chain
Q4VXU2	Polyadenylate-binding protein 1-like
Q53EL6	Programmed cell death protein 4
Q53SF7	Cordon-bleu protein-like 1
Q562R1	Beta-actin-like protein 2
Q58FF7	Putative heat shock protein HSP 90-beta-3
Q58FF8	Putative heat shock protein HSP 90-beta 2
Q58FG0	Putative heat shock protein HSP 90-alpha A5
Q5H907	Melanoma antigen family D, 2, isoform CRA_d
Q5JQF8	Polyadenylate-binding protein 1-like 2
Q5JXB2	Putative ubiquitin-conjugating enzyme E2 N-like
Q5SR53	Putative uncharacterized protein PIK3CD-AS1
Q5T4S7	E3 ubiquitin-protein ligase UBR4
Q5T6H7	Xaa-Pro aminopeptidase 1
Q5TBR0	Sialic acid synthase (Fragment)
Q5TFE4	5'-nucleotidase domain-containing protein 1
Q5TIA1	Meiosis inhibitor protein 1
Q5TID7	Coiled-coil domain-containing protein 181
Q68CZ2	Tensin-3
Q6AWC2	Protein WWC2
Q6DN14	Multiple C2 and transmembrane domain-containing protein 1
Q6FI81	Anamorsin
Q6GQQ9	OTU domain-containing protein 7B

Q6IN84	rRNA methyltransferase 1, mitochondrial
Q6P2Q9	Pre-mRNA-processing-splicing factor 8
Q6P575	Putative inactive beta-glucuronidase protein GUSBP11
Q6PGP7	Tetratricopeptide repeat protein 37
Q6PKG0	La-related protein 1
Q6S8J3	POTE ankyrin domain family member E
Q6TGC4	Protein-arginine deiminase type-6
Q6ZMR3	L-lactate dehydrogenase A-like 6A
Q6ZP82	Coiled-coil domain-containing protein 141
Q6ZSI9	Calpain-12
Q71U36	Tubulin alpha-1A chain
Q7KZF4	Staphylococcal nuclease domain-containing protein 1
Q7L2H7	Eukaryotic translation initiation factor 3 subunit M
Q7Z406	Myosin-14
Q7Z692	Carcinoembryonic antigen-related cell adhesion molecule 19
Q7Z6J0	E3 ubiquitin-protein ligase SH3RF1
Q7Z6Z7	E3 ubiquitin-protein ligase HUWE1
Q7Z794	Keratin, type II cytoskeletal 1b
Q86VP6	Cullin-associated NEDD8-dissociated protein 1
Q86VS8	Protein Hook homolog 3
Q86W92	Liprin-beta-1
Q86WI1	Fibrocystin-L
Q86YR6	POTE ankyrin domain family member D
Q8IZF4	Probable G-protein coupled receptor 114
Q8IZS6	Tctex1 domain-containing protein 3
Q8IZT9	Protein FAM9C
Q8N163	Cell cycle and apoptosis regulator protein 2
Q8N1D0	Beckwith-Wiedemann syndrome chromosomal region 1 candidate gene B protein
Q8N1L9	Basic leucine zipper transcriptional factor ATF-like 2
Q8N3C0	Activating signal cointegrator 1 complex subunit 3
Q8N5J2	Protein FAM63A
Q8N7S5	Phosphoinositide phospholipase C
Q8N9F8	Zinc finger protein 454
Q8NC51	Plasminogen activator inhibitor 1 RNA-binding protein
Q8NGP4	Olfactory receptor 5M3
Q8TF72	Protein Shroom3
Q8WUM4	Programmed cell death 6-interacting protein
Q8WVJ2	NudC domain-containing protein 2
Q8WW27	Putative C->U-editing enzyme APOBEC-4
Q8WXA9	Splicing regulatory glutamine/lysine-rich protein 1
Q8WXG6	MAP kinase-activating death domain protein
Q8WXG9	G-protein coupled receptor 98
Q92466	DNA damage-binding protein 2
Q92526	T-complex protein 1 subunit zeta-2
Q92538	Golgi-specific brefeldin A-resistance guanine nucleotide exchange factor 1
Q92688	Acidic leucine-rich nuclear phosphoprotein 32 family member B
Q92841	Probable ATP-dependent RNA helicase DDX17
Q92878	DNA repair protein RAD50
Q92905	COP9 signalosome complex subunit 5
Q92928	Putative Ras-related protein Rab-1C
Q92973	Transportin-1
Q96A99	Pentraxin-4
Q96BD8	Spindle and kinetochore-associated protein 1
Q96DG6	Carboxymethylenebutenolidase homolog
Q96F86	Enhancer of mRNA-decapping protein 3
Q96FQ6	Protein S100-A16
Q96HE9	Proline-rich protein 11
Q96LR5	Ubiquitin-conjugating enzyme E2 E2
Q96NW4	Ankyrin repeat domain-containing protein 27
Q96P64	Arf-GAP with GTPase, ANK repeat and PH domain-containing protein 4
Q96PY6	Serine/threonine-protein kinase Nek1
Q96R06	Sperm-associated antigen 5
Q96S07	Proline-rich protein 25
Q96TA1	Niban-like protein 1
Q99426	Tubulin-folding cofactor B
Q99436	Proteasome subunit beta type-7
Q99497	Protein deglycase DJ-1
Q99523	Sortilin
Q99714	3-hydroxyacyl-CoA dehydrogenase type-2
Q99829	Copine-1
Q99832	T-complex protein 1 subunit eta
Q9BPU6	Dihydropyrimidinase-related protein 5
Q9BPZ3	Polyadenylate-binding protein-interacting protein 2
Q9BQE3	Tubulin alpha-1C chain
Q9BR76	Coronin-1B
Q9BRP8	Partner of Y14 and mago
Q9BTT0	Acidic leucine-rich nuclear phosphoprotein 32 family member E

Q9BUF5	Tubulin beta-6 chain
Q9BVA1	Tubulin beta-2B chain
Q9BVG4	Protein PBDC1
Q9BWD1	Acetyl-CoA acetyltransferase, cytosolic
Q9BWT3	Poly(A) polymerase gamma
Q9BYX7	Putative beta-actin-like protein 3
Q9BZ29	Dedicator of cytokinesis protein 9
Q9H082	Ras-related protein Rab-33B
Q9H0J4	Glutamine-rich protein 2
Q9H1A4	Anaphase-promoting complex subunit 1
Q9H2P9	Diphthine synthase
Q9H361	Polyadenylate-binding protein 3
Q9H5J0	Zinc finger and BTB domain-containing protein 3
Q9H6Y2	WD repeat-containing protein 55
Q9HB71	Calcyclin-binding protein
Q9HBG6	Intraflagellar transport protein 122 homolog
Q9HC52	Chromobox protein homolog 8
Q9HCK5	Protein argonaute-4
Q9NPA0	ER membrane protein complex subunit 7
Q9NQI0	Probable ATP-dependent RNA helicase DDX4
Q9NQW7	Xaa-Pro aminopeptidase 1
Q9NR45	Sialic acid synthase
Q9NR48	Histone-lysine N-methyltransferase ASH1L
Q9NSC5	Homer protein homolog 3
Q9NTI7	Protein FAM212B
Q9NVE5	Ubiquitin carboxyl-terminal hydrolase 40
Q9NVI1	Fanconi anemia group I protein
Q9NXD2	Myotubularin-related protein 10
Q9NYL9	Tropomodulin-3
Q9NZA1	Chloride intracellular channel protein 5
Q9NZL4	Hsp70-binding protein 1
Q9P258	Protein RCC2
Q9P2H3	Intraflagellar transport protein 80 homolog
Q9P2J5	Leucine--tRNA ligase, cytoplasmic
Q9UBT2	SUMO-activating enzyme subunit 2
Q9UHV9	Prefoldin subunit 2
Q9UI15	Transgelin-3
Q9UKX2	Myosin-2
Q9UKY7	Protein CDV3 homolog
Q9ULV0	Unconventional myosin-Vb
Q9UMS4	Pre-mRNA-processing factor 19
Q9UMX0	Ubiquilin-1
Q9UQ80	Proliferation-associated protein 2G4
Q9Y230	RuvB-like 2
Q9Y265	RuvB-like 1
Q9Y266	Nuclear migration protein nudC
Q9Y281	Cofilin-2
Q9Y2T7	Y-box-binding protein 2
Q9Y2Z0	Suppressor of G2 allele of SKP1 homolog
Q9Y3I0	tRNA-splicing ligase RtcB homolog
Q9Y3L3	SH3 domain-binding protein 1
Q9Y3U8	60S ribosomal protein L36
Q9Y490	Talin-1
Q9Y4Y9	U6 snRNA-associated Sm-like protein LSm5
Q9Y536	Peptidyl-prolyl cis-trans isomerase A-like 4A/B/C
Q9Y570	Protein phosphatase methylesterase 1
Q9Y617	Phosphoserine aminotransferase
Q9Y6N3	Calcium-activated chloride channel regulator family member 3
Q9Y6U3	Adseverin
S4R435	Protein RPS10-NUDT3 (Fragment)

Table 9.7.1: Primer sequences used to sequence across the whole *PROSC* cDNA.

Primer	Sequence
1F	GCTAGTGGCGGTCAGCAAAACC
2F	GGTTAAAGGTTATGGTCCAGAT
3F	GGACAAGTGCGCAGCAGACGTGAA
1R	GGGCTGGATGGCTGGGAGATCC
2R	AGCATGAAGAGATTGGGGACAGCC
3R	CATGCCCATGCTCAGCTCAACC
4R	CCATTGCCTGAAGCAAGCTTCC

BIBLIOGRAPHY

M. Adamowicz, R. Płoski, D. Rokicki, *et al.* (2007). Transferrin hypoglycosylation in hereditary fructose intolerance: using the clues and avoiding the pitfalls. *Journal of Inherited Metabolic Disease*, 30(3):407.

S. Akanuma, T. Sakurai, M. Tachikawa, *et al.* (2015). Transporter-mediated L-glutamate elimination from cerebrospinal fluid: possible involvement of excitatory amino acid transporters expressed in ependymal cells and choroid plexus epithelial cells. *Fluids and Barriers of the CNS*, 12:11.

O. Y. Al-Dirbashi, Z. N. Al-Hassnan, M. S. Rashed (2006). Determination of homocitrulline in urine of patients with HHH syndrome by liquid chromatography tandem mass spectrometry. *Analytical and Bioanalytical Chemistry*, 386(7-8):2013–2017.

M. Albersen, M. Bosma, J. J. M. Jans, *et al.* (2015). Vitamin B6 in plasma and cerebrospinal fluid of children. *PloS One*, 10(3):e0120972.

M. Aldenhoven, R. F. Wynn, P. J. Orchard, *et al.* (2015). Long-term outcome of Hurler syndrome patients after hematopoietic cell transplantation: an international multicenter study. *Blood*, 125(13):2164–2172.

W. K. Alderton, C. E. Cooper, R. G. Knowles (2001). Nitric oxide synthases: structure, function and inhibition. *The Biochemical Journal*, 357(3):593–615.

M. Alfadhel, M. Almuntashri, R. H. Jadah, *et al.* (2013). Biotin-responsive basal ganglia disease should be renamed biotin-thiamine-responsive basal ganglia disease: a retrospective review of the clinical, radiological and molecular findings of 18 new cases. *Orphanet Journal of Rare Diseases*, 8:83.

A. Alfares, L. D. Nunez, K. Al-Thihli, *et al.* (2011). Combined malonic and methylmalonic aciduria: exome sequencing reveals mutations in the ACSF3 gene in patients with a non-classic phenotype. *Journal of Medical Genetics*, 48(9):602–605.

M. Ali, P. Rellos, T. M. Cox (1998). Hereditary fructose intolerance. *Journal of Medical Genetics*, 35(5):353–365.

N. M. Allen, M. Mannion, J. Conroy, *et al.* (2014). The variable phenotypes of KCNQ-related epilepsy. *Epilepsia*, 55(9):e99–105.

A. M. Almeida, Y. Murakami, D. M. Layton, *et al.* (2006). Hypomorphic promoter mutation in PIGM causes inherited glycosylphosphatidylinositol deficiency. *Nature Medicine*, 12(7):846–851.

M. Almuriekhi, T. Shintani, S. Fahiminiya, *et al.* (2015). Loss-of-function mutation in APC2 causes sotos syndrome features. *Cell Reports*, 10(9):1585–1598.

J. Alroy, A. A. Ucci (2006). Skin biopsy: a useful tool in the diagnosis of lysosomal storage diseases. *Ultrastructural Pathology*, 30(6):489–503.

J. Andricovich, Y. Kai, W. Peng, *et al.* (2016). Histone demethylase KDM2B regulates lineage commitment in normal and malignant hematopoiesis. *The Journal of Clinical Investigation*, 126(3):905–920.

T. Aoshima, M. Kajita, Y. Sekido, *et al.* (2001). Novel mutations (H337R and 238-362del) in the CPS1 gene cause carbamoyl phosphate synthetase I deficiency. *Human Heredity*, 52(2):99–101.

J. Ardura, J. Andres, J. R. Garmendia, *et al.* (2010). Melatonin in epilepsy and febrile seizures. *Journal of Child Neurology*, 25(7):888–891.

B. W. Arentson, N. Sanyal, D. F. Becker (2012). Substrate channeling in proline metabolism. *Frontiers in Bioscience*, 17:375–388.

S. M. Aronson, B. W. Volk (2013). *Inborn Disorders of Sphingolipid Metabolism: Proceedings of the Third International Symposium on the Cerebral Sphingolipidoses.* Elsevier.

H. Asada, Y. Kawamura, K. Maruyama, *et al.* (1996). Mice lacking the 65 kDa isoform of glutamic acid decarboxylase (GAD65) maintain normal levels of GAD67 and GABA in their brains but are susceptible to seizures. *Biochemical and Biophysical Research Communications*, 229(3):891–895.

C. Attanasio, A. David, M. Neerman-Arbez (2003). Outcome of donor splice site mutations accounting for congenital afibrinogenemia reflects order of intron removal in the fibrinogen alpha gene (FGA). *Blood*, 101(5):1851–1856.

A. Aycicek, A. Iscan (2007). The effects of carbamazepine, valproic acid and phenobarbital on the oxidative and antioxidative balance in epileptic children. *European Neurology*, 57(2):65–69.

G. Bach, S. M. Moskowitz, P. T. Tieu, *et al.* (1993). Molecular analysis of Hurler syndrome in Druze and Muslim Arab patients in Israel: multiple allelic mutations of the IDUA gene in a small geographic area. *American Journal of Human Genetics*, 53(2):330–338.

L. K. Bak, A. Schousboe, H. S. Waagepetersen (2006). The glutamate/GABA-glutamine cycle: aspects of transport, neurotransmitter homeostasis and ammonia transfer. *Journal of Neurochemistry*, 98(3):641–653.

D. Baralle, M. Baralle (2005). Splicing in action: assessing disease causing sequence changes. *Journal of Medical Genetics*, 42(10):737–748.

S. C. Barber, R. J. Mead, P. J. Shaw (2006). Oxidative stress in ALS: a mechanism of neurodegeneration and a therapeutic target. *Biochimica Et Biophysica Acta*, 1762(11-12):1051–1067.

M. R. Baumgartner, D. Rabier, M. Nassogne, *et al.* (2005). Delta1-pyrroline-5-carboxylate synthase deficiency: neurodegeneration, cataracts and connective tissue manifestations combined with hyperammonaemia and reduced ornithine, citrulline, arginine and proline. *European Journal of Pediatrics*, 164(1):31–36.

S. Baumgartner-Sigl, E. Haberlandt, S. Mumm, *et al.* (2007). Pyridoxine-responsive seizures as the first symptom of infantile hypophosphatasia caused by two novel missense mutations (c.677T>C, p.M226T; c.1112C>T, p.T371I) of the tissue-nonspecific alkaline phosphatase gene. *Bone*, 40(6):1655–1661.

M. F. Beal, W. R. Matson, K. J. Swartz, *et al.* (1990). Kynurenine pathway measurements in Huntington's disease striatum: evidence for reduced formation of kynurenic acid. *Journal of Neurochemistry*, 55(4):1327–1339.

K. Begriche, J. Massart, A. Abbey-Toby, *et al.* (2008). Beta-aminoisobutyric acid prevents diet-induced obesity in mice with partial leptin deficiency. *Obesity*, 16(9):2053–2067.

C. J. Bell, D. L. Dinwiddie, N. A. Miller, *et al.* (2011). Carrier testing for severe childhood recessive diseases by next-generation sequencing. *Science Translational Medicine*, 3(65):65ra4.

C. Bellanné-Chantelot, C. Carette, J.-P. Riveline, *et al.* (2008). The type and the position of HNF1A mutation modulate age at diagnosis of diabetes in patients with maturity-onset diabetes of the young (MODY)-3. *Diabetes*, 57(2):503–508.

G. Bellini, F. Miceli, M. V. Soldovieri, *et al.* (1993). KCNQ2-Related Disorders. In R. A. Pagon, M. P. Adam, H. H. Ardinger, *et al.*, editors, *GeneReviews(®)*. University of Washington, Seattle, Seattle (WA).

R. Benesch, R. E. Benesch, S. Kwong, *et al.* (1982). Labeling of hemoglobin with pyridoxal phosphate. *The Journal of Biological Chemistry*, 257(3):1320–1324.

D. R. Bentley, S. Balasubramanian, H. P. Swerdlow, *et al.* (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, 456(7218):53–59.

E. C. Berglund, C. M. Lindqvist, S. Hayat, *et al.* (2013). Accurate detection of subclonal single nucleotide variants in whole genome amplified and pooled cancer samples using HaloPlex target enrichment. *BMC Genomics*, 14(1):856.

D. A. Berkich, M. S. Ola, J. Cole, *et al.* (2007). Mitochondrial transport proteins of the brain. *Journal of Neuroscience Research*, 85(15):3367–3377.

S. E. Bianconi, J. L. Cross, C. A. Wassif, *et al.* (2015). Pathogenesis, epidemiology, diagnosis and clinical aspects of Smith-Lemli-Opitz syndrome. *Expert Opinion on Orphan Drugs*, 3(3):267–280.

P. Bie, P. Muller, C. Wijmenga, *et al.* (2007). Molecular pathogenesis of Wilson and Menkes disease: correlation of mutations with molecular defects and disease phenotypes. *Journal of Medical Genetics*, 44(11):673–688.

C. Biervert, B. C. Schroeder, C. Kubisch, *et al.* (1998). A potassium channel mutation in neonatal human epilepsy. *Science*, 279(5349):403–406.

M. R. Bisp, M. V. Bor, E. Heinsvig, *et al.* (2002). Determination of vitamin B6 vitamers and pyridoxic acid in plasma: development and evaluation of a high-performance liquid chromatographic assay. *Analytical Biochemistry*, 305(1):82–89.

H. Bisswanger (2014). Enzyme assays. *Perspectives in Science*, 1(1–6):41–55.

S. Blanchard, M. Sadilek, C. R. Scott, *et al.* (2008). Tandem mass spectrometry for the direct assay of lysosomal enzymes in dried blood spots: application to screening newborns for mucopolysaccharidosis I. *Clinical Chemistry*, 54(12):2067–2070.

K. P. Blemings, T. D. Crenshaw, N. J. Benevenga (1998). Mitochondrial lysine uptake limits hepatic lysine oxidation in rats fed diets containing 5, 20 or 60% casein. *The Journal of Nutrition*, 128(12):2427–2434.

L. Blumkin, A. Suls, T. Deconinck, *et al.* (2012). Neonatal seizures associated with a severe neonatal myoclonus like dyskinesia due to a familial KCNQ2 gene mutation. *European Journal of Paediatric Neurology*, 16(4):356–360.

J. P. Bohney, M. L. Fonda, R. C. Feldhoff (1992). Identification of Lys190 as the primary binding site for pyridoxal 5'-phosphate in human serum albumin. *FEBS letters*, 298(2-3):266–268.

L. A. Bok, J. V. Been, E. A. Struys, *et al.* (2010). Antenatal treatment in two Dutch families with pyridoxine-dependent seizures. *European Journal of Pediatrics*, 169(3):297–303.

L. A. Bok, F. J. Halbertsma, S. Houterman, *et al.* (2012). Long-term outcome in pyridoxine-dependent epilepsy. *Developmental Medicine and Child Neurology*, 54(9):849–854.

F. Bonn, T. Tatsuta, C. Petrungaro, *et al.* (2011). Presequence-dependent folding ensures MrpL32 processing by the m-AAA protease in mitochondria. *The EMBO journal*, 30(13):2545–2556.

L. M. Boyden, C. Y. Kam, A. Hernández-Martín, *et al.* (2016). Dominant de novo DSP mutations cause erythrokeratodermia-cardiomyopathy syndrome. *Human Molecular Genetics*, 25(2):348–357.

S. W. Briggs, A. S. Galanopoulou, S. W. Briggs, *et al.* (2011). Altered GABA signaling in early life epilepsies, altered GABA signaling in early life epilepsies. *Neural Plasticity*, 2011:e527605.

P. S. Brookes, Y. Yoon, J. L. Robotham, *et al.* (2004). Calcium, ATP, and ROS: a mitochondrial love-hate triangle. *American Journal of Physiology Cell Physiology*, 287(4):C817–833.

D. A. Brown, P. R. Adams (1980). Muscarinic suppression of a novel voltage-sensitive K+ current in a vertebrate neurone. *Nature*, 283(5748):673–676.

C. Bruntner, C. Bormann (1998). The Streptomyces tendae Tü901 L-lysine 2-aminotransferase catalyzes the initial reaction in nikkomycin D biosynthesis. *European Journal of Biochemistry*, 254(2):347–355.

M. V. Buell, R. E. Hansen (1960). Reaction of pyridoxal-5-phosphate with aminothiols. *Journal of the American Chemical Society*, 82(23):6042–6049.

H. P. J. Buermans, J. T. den Dunnen (2014). Next generation sequencing technology: Advances and applications. *Biochimica et Biophysica Acta*, 1842(10):1932–1941.

G. Burnstock (2007). Physiology and pathophysiology of purinergic neurotransmission. *Physiological Reviews*, 87(2):659–797.

L. Caldovic, M. Tuchman (2003). N-acetylglutamate and its changing role through evolution. *The Biochemical Journal*, 372(2):279–290.

A. E. Carpenter, T. R. Jones, M. R. Lamprecht, *et al.* (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biology*, 7(10):R100.

N. A. Carson, B. G. Scally, D. W. Neill, *et al.* (1968). Saccharopinuria: a new inborn error of lysine metabolism. *Nature*, 218(5142):679.

M. Cecchi, P. Messina, L. Airoldi, *et al.* (2014). Plasma amino acids patterns and age of onset of amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis & Frontotemporal Degeneration*, 15(5-6):371–375.

A. Cha, G. E. Snyder, P. R. Selvin, *et al.* (1999). Atomic scale movement of the voltage-sensing region in a potassium channel measured via spectroscopy. *Nature*, 402(6763):809–813.

Y. E. Chang (1978). Lysine metabolism in the rat brain: the pipecolic acid-forming pathway. *Journal of Neurochemistry*, 30(2):347–354.

L. Chen, J. Zhao, Q. Tang, *et al.* (2016). PFKFB3 control of cancer growth by responding to circadian clock outputs. *Scientific Reports*, 6:24324.

Q. Chen, G. S. Lakshmikanth, J. A. Spudich, *et al.* (2007). The localization of inner centromeric protein (INCENP) at the cleavage furrow is dependent on Kif12 and involves interactions of the N terminus of INCENP with the actin cytoskeleton. *Molecular Biology of the Cell*, 18(9):3366–3374.

P. Y. Cheung, C. C. Fong, K. T. Ng, *et al.* (2003). Interaction between pyridoxal kinase and pyridoxal-5-phosphate-dependent enzymes. *Journal of Biochemistry*, 134(5):731–738.

A. Chiaratti de Oliveira, A. M. dos Santos, A. M. Martins, *et al.* (2001). Screening for inborn errors of metabolism among newborns with metabolic disturbance and/or neurological manifestations without determined cause. *São Paulo Medical Journal*, 119(5):160–164.

T. Chiyonobu, N. Inoue, M. Morimoto, *et al.* (2014). Glycosylphosphatidylinositol (GPI) anchor deficiency caused by mutations in PIGW is associated with West syndrome and hyperphosphatasia with mental retardation syndrome. *Journal of Medical Genetics*, 51(3):203–207.

S. Y. Choi, J. E. Churchich, E. Zaiden, *et al.* (1987). Brain pyridoxine-5-phosphate oxidase. Modulation of its catalytic activity by reaction with pyridoxal 5-phosphate and analogs. *The Journal of Biological Chemistry*, 262(25):12013–12017.

P. Christen, P. K. Mehta (2001). From cofactor to enzymes. The molecular evolution of pyridoxal-5'-phosphate-dependent enzymes. *Chemical Record*, 1(6):436–447.

R. Chumnantana, N. Yokochi, T. Yagi (2005). Vitamin B6 compounds prevent the death of yeast cells due to menadione, a reactive oxygen generator. *Biochimica Et Biophysica Acta*, 1722(1):84–91.

S. Chun, J. C. Fay (2009). Identification of deleterious mutations within three human genomes. *Genome Research*, 19(9):1553–1561.

M. G. Claros, P. Vincens (1996). Computational method to predict mitochondrially imported proteins and their targeting sequences. *European Journal of Biochemistry*, 241(3):779–786.

P. T. Clayton (2006). B6-responsive disorders: a model of vitamin dependency. *Journal of Inherited Metabolic Disease*, 29(2-3):317–326.

E. M. Clement, C. Godfrey, J. Tan, *et al.* (2008). Mild POMGnT1 mutations underlie a novel limb-girdle muscular dystrophy variant. *Archives of Neurology*, 65(1):137–141.

S. P. Coburn, P. J. Ziegler, D. L. Costill, *et al.* (1991). Response of vitamin B-6 content of muscle to changes in vitamin B-6 intake in men. *The American Journal of Clinical Nutrition*, 53(6):1436–1442.

R. Cohen, L. Basel-Vanagaite, H. Goldberg-Stern, *et al.* (2014). Two siblings with early infantile myoclonic encephalopathy due to mutation in the gene encoding mitochondrial glutamate/H+ symporter SLC25A22. *European Journal of Paediatric Neurology*, 18(6):801–805.

A. J. L. Cooper, J. T. Pinto, B. F. Krasnikov, *et al.* (2008). Substrate specificity of human glutamine transaminase K as an aminotransferase and as a cysteine S-conjugate beta-lyase. *Archives of Biochemistry and Biophysics*, 474(1):72–81.

C. R. Coughlin, C. D. M. van Karnebeek, W. Al-Hertani, *et al.* (2015). Triple therapy with pyridoxine, arginine supplementation and dietary lysine restriction in pyridoxine-dependent epilepsy: Neurodevelopmental outcome. *Molecular Genetics and Metabolism*, 116(1-2):35–43.

D. B. Coursin (1954). Convulsive seizures in infants with pyridoxine-deficient diet. *Journal of the American Medical Association*, 154(5):406–408.

F. L. Crane, Y. Hatefi, R. L. Lester, *et al.* (1957). Isolation of a quinone from beef heart mitochondria. *Biochimica Et Biophysica Acta*, 25(1):220–221.

Y. Daikhin, M. Yudkoff (2000). Compartmentation of brain glutamate metabolism in neurons and glia. *The Journal of Nutrition*, 130(4S Suppl):1026S–31S.

C. Dalen Meurs-van der Schoor, M. van Weissenbruch, M. van Kempen, *et al.* (2014). Severe neonatal epileptic encephalopathy and KCNQ2 mutation: neuropathological substrate? *Frontiers in Pediatrics*, 2:136.

K. Danhauser, T. B. Haack, B. Alhaddad, *et al.* (2016). EARS2 mutations cause fatal neonatal lactic acidosis, recurrent hypoglycemia and agenesis of corpus callosum. *Metabolic Brain Disease*, 31(3):717–721.

J. T. DaRe, V. Vasta, J. Penn, *et al.* (2013). Targeted exome sequencing for mitochondrial disorders reveals high genetic heterogeneity. *BMC Medical Genetics*, 14:118.

C. Dass (2007). Modes of Ionization. In *Fundamentals of Contemporary Mass Spectrometry*, pages 15–65. John Wiley & Sons, Inc.

H. O. de Baulny, J. F. Benoist, O. Rigal, *et al.* (2005). Methylmalonic and propionic acidaemias: management and outcome. *Journal of Inherited Metabolic Disease*, 28(3):415–423.

T. J. de Koning, J. D. H. Jongbloed, B. Sikkema-Raddatz, *et al.* (2015). Targeted next-generation sequencing panels for monogenetic disorders in clinical diagnostics: the opportunities and challenges. *Expert Review of Molecular Diagnostics*, 15(1):61–70.

M. G. A. de Roo, N. G. G. M. Abeling, C. B. Majoie, *et al.* (2014). Infantile hypophosphatasia without bone deformities presenting with severe pyridoxine-resistant seizures. *Molecular Genetics and Metabolism*, 111(3):404–407.

M. G. M. de Sain-van der Velden, M. van der Ham, J. J. Jans, *et al.* (2016). A new approach for fast metabolic diagnostics in CMAMMA. *JIMD Reports*, [Epub ahead of print].

K. Dedek, B. Kunath, C. Kananura, *et al.* (2001). Myokymia and neonatal epilepsy caused by a mutation in the voltage sensor of the KCNQ2 K+ channel. *Proceedings of the National Academy of Sciences of the United States of America*, 98(21):12272–12277.

R. J. DeLorenzo, S. Pal, S. Sombati (1998). Prolonged activation of the N-methyl-d-aspartate receptor–Ca2+ transduction pathway causes spontaneous recurrent epileptiform discharges in hippocampal neurons in culture. *Proceedings of the National Academy of Sciences of the United States of America*, 95(24):14482–14487.

M. Delous, L. Baala, R. Salomon, *et al.* (2007). The ciliary gene RPGRIP1L is mutated in cerebello-oculo-renal syndrome (Joubert syndrome type B) and Meckel syndrome. *Nature Genetics*, 39(7):875–881.

M. A. DePristo, E. Banks, R. Poplin, *et al.* (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, 43(5):491–498.

F. E. Dewey, M. E. Grove, C. Pan, *et al.* (2014). Clinical interpretation and implications of whole-genome sequencing. *JAMA*, 311(10):1035–1045.

M. L. di Salvo, R. Contestabile, M. K. Safo (2011). Vitamin B(6) salvage enzymes: mechanism, structure and regulation. *Biochimica Et Biophysica Acta*, 1814(11):1597–1608.

M. L. di Salvo, I. Nogués, A. Parroni, *et al.* (2015). On the mechanism of Escherichia coli pyridoxal kinase inhibition by pyridoxal and pyridoxal 5'-phosphate. *Biochimica Et Biophysica Acta*, 1854(9):1160–1166.

C. Diesen, A. Saarinen, H. Pihko, *et al.* (2004). POMGnT1 mutation and phenotypic spectrum in muscle-eye-brain disease. *Journal of Medical Genetics*, 41(10):e115.

E. L. Dillon, D. A. Knabe, G. Wu (1999). Lactate inhibits citrulline and arginine synthesis from proline in pig enterocytes. *American Journal of Physiology - Gastrointestinal and Liver Physiology*, 276(5):G1079–G1086.

D. L. Dinwiddie, L. D. Smith, N. A. Miller, *et al.* (2013). Diagnosis of mitochondrial disorders by concomitant next-generation sequencing of the exome and mitochondrial genome. *Genomics*, 102(3):148–156.

K. M. Dipple, E. R. McCabe (2000). Phenotypes of patients with "simple" Mendelian disorders are complex traits: thresholds, modifiers, and systems dynamics. *American Journal of Human Genetics*, 66(6):1729–1735.

A. Drozdetskiy, C. Cole, J. Procter, *et al.* (2015). JPred4: a protein secondary structure prediction server. *Nucleic Acids Research*, 43:W389–W394.

S. Duan, C. M. Anderson, B. A. Stein, *et al.* (1999). Glutamate induces rapid upregulation of astrocyte glutamate transport and cell-surface expression of GLAST. *The Journal of Neuroscience*, 19(23):10193–10200.

M. E. Dy, K. B. Sims, J. Friedman (2015). TPP1 deficiency: Rare cause of isolated childhood-onset progressive ataxia. *Neurology*, 85(14):1259–1261.

H. Eagle, K. A. Piez, R. Fleischman, *et al.* (1959). Protein turnover in mammaliar cell cultures. *The Journal of Biological Chemistry*, 234(3):592–597.

N. Elleuch, C. Depienne, A. Benomar, *et al.* (2006). Mutation analysis of the paraplegin gene (SPG7) in patients with hereditary spastic paraplegia. *Neurology*, 66(5):654–659.

T. Engel, A. Jimenez-Pacheco, M. T. Miras-Portugal, *et al.* (2012). P2X7 receptor in epilepsy; role in pathophysiology and potential targeting for seizure control. *International Journal of Physiology, Pathophysiology and Pharmacology*, 4(4):174–187.

S. Erhardt, L. Schwieler, L. Nilsson, *et al.* (2007). The kynurenic acid hypothesis of schizophrenia. *Physiology & Behavior*, 92(1–2):203–209.

M. Ericson, R. B. C. Clarke, P. Chau, *et al.* (2010). Beta-alanine elevates dopamine levels in the rat nucleus accumbens: antagonism by strychnine. *Amino Acids*, 38(4):1051–1055.

M. R. Fabian, N. Sonenberg, W. Filipowicz (2010). Regulation of mRNA translation and stability by microRNAs. *Annual Review of Biochemistry*, 79:351–379.

S. Fargue, G. Rumsby, C. J. Danpure (2013). Multiple mechanisms of action of pyridoxine in primary hyperoxaluria type 1. *Biochimica et Biophysica Acta*, 1832(10):1776–1783.

R. D. Farrant, V. Walker, G. A. Mills, *et al.* (2001). Pyridoxal phosphate de-activation by pyrroline-5-carboxylic acid. Increased risk of vitamin B6 deficiency and seizures in hyperprolinemia type II. *The Journal of Biological Chemistry*, 276(18):15107–15116.

E. Fassone, Y. Wedatilake, C. J. DeVile, *et al.* (2013). Treatable Leigh-like encephalopathy presenting in adolescence. *BMJ Case Reports*, 2013:bcr2013200838.

M. A. Ferguson, T. Kinoshita, G. W. Hart (2009). Glycosylphosphatidylinositol Anchors. In A. Varki, R. D. Cummings, J. D. Esko, *et al.*, editors, *Essentials of Glycobiology*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), 2nd edition.

G. Fiermonte, L. Palmieri, S. Todisco, *et al.* (2002). Identification of the mitochondrial glutamate transporter. Bacterial expression, reconstitution, functional characterization, and tissue distribution of two human isoforms. *The Journal of Biological Chemistry*, 277(22):19289–19294.

J. J. Filiano (2006). Neurometabolic diseases in the newborn. *Clinics in Perinatology*, 33(2):411–479.

M. P. Flynn, M. C. Martin, P. T. Moore, *et al.* (1989). Type II hyperprolinaemia in a pedigree of Irish travellers (nomads). *Archives of Disease in Childhood*, 64(12):1699–1707.

E. J. Footitt, P. T. Clayton, K. Mills, *et al.* (2013). Measurement of plasma B6 vitamer profiles in children with inborn errors of vitamin B6 metabolism using an LC-MS/MS method. *Journal of Inherited Metabolic Disease*, 36(1):139–145.

E. J. Footitt, S. J. Heales, P. B. Mills, *et al.* (2011). Pyridoxal 5'-phosphate in cerebrospinal fluid; factors affecting concentration. *Journal of inherited metabolic disease*, 34(2):529–538.

E. G. Frame (1958). The levels of individual free amino acids in the plasma of normal man at various intervals after a high-protein meal. *Journal of Clinical Investigation*, 37(12):1710–1723.

E. Franzoni, R. Bracceschi, M. C. Colonnelli, *et al.* (2005). Clinical features of benign infantile convulsions: familial and sporadic cases. *Neurology*, 65(7):1098–1100.

J. H. French, B. B. Grueter, R. Druckman, *et al.* (1965). Pyridoxine and infantile myoclonic seizures. *Neurology*, 15:101–113.

J. M. Fritschy, T. Kiener, V. Bouilleret, *et al.* (1999). GABAergic neurons and GABA(A)-receptors in temporal lobe epilepsy. *Neurochemistry International*, 34(5):435–445.

L. Fu, A. Rab, L. P. Tang, *et al.* (2012). Dab2 is a key regulator of endocytosis and post-endocytic trafficking of the cystic fibrosis transmembrane conductance regulator. *The Biochemical Journal*, 441(2):633–643.

T. F. Fu, M. di Salvo, V. Schirch (2001). Distribution of B6 vitamers in Escherichia coli as determined by enzymatic assay. *Analytical Biochemistry*, 298(2):314–321.

I. Fujiwara, Y. Murakami, T. Niihori, *et al.* (2015). Mutations in PIGL in a patient with Mabry syndrome. *American Journal of Medical Genetics Part A*, 167A(4):777–785.

Y. Fukuyama, K. Tada, M. Kawamura, *et al.* (1980). Recent topics in the field of vitamin B6 therapy (round table discussion). *Shoni Naika*, 12:2062–2072.

A. Garrod (1908). The Croonian Lectures on inborn errors of metabolism, lecture II: alkaptonuria. *The Lancet*, 172(4427):1–7.

M. T. Geraghty, D. Vaughn, A. J. Nicholson, *et al.* (1998). Mutations in the Delta1-Pyrroline 5-Carboxylate Dehydrogenase Gene Cause Type II Hyperprolinemia. *Human Molecular Genetics*, 7(9):1411–1415.

F. Gerdes, T. Tatsuta, T. Langer (2012). Mitochondrial AAA proteases–towards a molecular understanding of membrane-bound proteolytic machines. *Biochimica Et Biophysica Acta*, 1823(1):49–55.

D. S. Gerhard, L. Wagner, E. A. Feingold, *et al.* (2004). The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). *Genome Research*, 14(10B):2121–2127.

C. Gilissen, J. Y. Hehir-Kwa, D. T. Thung, *et al.* (2014). Genome sequencing identifies major causes of severe intellectual disability. *Nature*, 511(7509):344–347.

V. M. Ginocchio, N. Brunetti-Pierri (2016). Progress toward improved therapies for inborn errors of metabolism. *Human Molecular Genetics*, 25(R1):R27–35.

M. C. Giocondi, B. Seantier, P. Dosset, *et al.* (2008). Characterizing the interactions between GPI-anchored alkaline phosphatases and membrane domains by AFM. *Pflügers Archiv: European Journal of Physiology*, 456(1):179–188.

C. Goizet, A. Boukhris, A. Durr, *et al.* (2009). CYP7B1 mutations in pure and complex forms of hereditary spastic paraplegia type 5. *Brain*, 132(Pt 6):1589–1600.

E. González, D. Danehower, M. E. Daub (2007). Vitamer levels, stress response, enzyme activity, and gene regulation of Arabidopsis lines mutant in the pyridoxine/pyridoxamine 5'-phosphate oxidase (PDX3) and the pyridoxal kinase (SOS4) genes involved in the vitamin B6 salvage pathway. *Plant Physiology*, 145(3):985–996.

S. M. Gospe (1993). Pyridoxine-Dependent Epilepsy. In R. A. Pagon, M. P. Adam, H. H. Ardinger, *et al.*, editors, *GeneReviews(®)*. University of Washington, Seattle, Seattle (WA).

R. Głowacki, J. Stachniuk, K. Borowczyk, *et al.* (2016). Quantification of homocysteine and cysteine by derivatization with pyridoxal 5'-phosphate and hydrophilic interaction liquid chromatography. *Analytical and Bioanalytical Chemistry*, 408(7):1935–1941.

S. K. Grebe, R. J. Singh (2011). LC-MS/MS in the clinical laboratory – where to from here? *The Clinical Biochemist Reviews*, 32(1):5–31.

K. Grohmann, H. Lauffer, P. Lauenstein, *et al.* (2015). Hereditary orotic aciduria with epilepsy and without megaloblastic anemia. *Neuropediatrics*, 46(2):123–125.

G. J. Guillemin, S. J. Kerr, B. J. Brew (2005). Involvement of quinolinic acid in aids dementia complex. *Neurotoxicity Research*, 7(1-2):103–123.

R. Guthrie, A. Susi (1963). A simple phenylalaine method for detecting phenylketonuria in large populations of newborn infants. *Pediatrics*, 32:338–343.

T. B. Haack, C. Makowski, Y. Yao, *et al.* (2012). Impaired riboflavin transport due to missense mutations in SLC52A2 causes Brown-Vialetto-Van Laere syndrome. *Journal of Inherited Metabolic Disease*, 35(6):943–948.

R. H. Haas, S. Parikh, M. J. Falk, *et al.* (2008). The in-depth evaluation of suspected mitochondrial disease. *Molecular Genetics and Metabolism*, 94(1):16–37.

A. Habib, Z. Md Yunus, N. A. Azize, *et al.* (2013). Hyperexcretion of homocitrulline in a Malaysian patient with lysinuric protein intolerance. *European Journal of Pediatrics*, 172(9):1277–1281.

J. K. Hadley, G. M. Passmore, L. Tatulian, *et al.* (2003). Stoichiometry of expressed KCNQ2/KCNQ3 potassium channels and subunit composition of native ganglionic M channels deduced from block by tetraethylammonium. *The Journal of Neuroscience*, 23(12):5012–5019.

C. H. Hagedorn, J. M. Phang (1983). Transfer of reducing equivalents into mitochondria by the interconversions of proline and delta 1-pyrroline-5-carboxylate. *Archives of Biochemistry and Biophysics*, 225(1):95–101.

Y. Haitin, B. Attali (2008). The C-terminus of Kv7 channels: a multifunctional module. *The Journal of Physiology*, 586(7):1803–1810.

A. Hallen, J. F. Jamie, A. J. L. Cooper (2013). Lysine metabolism in mammalian brain: an update on the importance of recent discoveries. *Amino Acids*, 45(6):1249–1272.

S. A. Hamed, M. M. Abdellah, N. El-Melegy (2004). Blood levels of trace elements, electrolytes, and oxidative stress/antioxidant systems in epileptic patients. *Journal of Pharmacological Sciences*, 96(4):465–473.

A. Hamosh, A. F. Scott, J. S. Amberger, *et al.* (2005). Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Research*, 33(Database Issue):D514–D517.

Q. Han, T. Cai, D. A. Tagle, *et al.* (2010). Structure, expression, and function of kynurenine aminotransferases in human and rodent brains. *Cellular and Molecular Life Sciences*, 67(3):353–368.

Q. Han, J. Li, J. Li (2004). pH dependence, substrate specificity and inhibition of human kynurenine aminotransferase I. *European Journal of Biochemistry*, 271(23-24):4804–4814.

Q. Han, H. Robinson, T. Cai, *et al.* (2009a). Biochemical and structural properties of mouse kynurenine aminotransferase III. *Molecular and Cellular Biology*, 29(3):784–793.

Q. Han, H. Robinson, T. Cai, *et al.* (2009b). Structural insight into the inhibition of human kynurenine aminotransferase I/glutamine transaminase K. *Journal of Medicinal Chemistry*, 52(9):2786–2793.

T. Hanai, R. S. of Chemistry (Great Britain) (1999). *HPLC: A Practical Guide*. Royal Society of Chemistry.

L. Hansen, H. Tawamie, Y. Murakami, *et al.* (2013). Hypomorphic mutations in PGAP2, encoding a GPI-anchor-remodeling protein, cause autosomal-recessive intellectual disability. *American Journal of Human Genetics*, 92(4):575–583.

O. Hansson, B. Hagberg (1968). Effect of pyridoxine treatment in children with epilepsy. *Acta Societatis Medicorum Upsaliensis*, 73(1):35–43.

I. K. Hart, P. Maddison, J. Newsom-Davis, *et al.* (2002). Phenotypic variants of autoimmune peripheral nerve hyperexcitability. *Brain*, 125(Pt 8):1887–1895.

H. Hartmann, M. Fingerhut, C. Jakobs, *et al.* (2011). Status epilepticus in a neonate treated with pyridoxine because of a familial recurrence risk for antiquitin deficiency: pyridoxine toxicity? *Developmental Medicine and Child Neurology*, 53(12):1150–1153.

H. A. Hartmann, G. E. Kirsch, J. A. Drewe, *et al.* (1991). Exchange of conduction pathways between two related K+ channels. *Science*, 251(4996):942–944.

N. Hay, N. Sonenberg (2004). Upstream and downstream of mTOR. *Genes & Development*, 18(16):1926–1945.

J. Häberle, N. Boddaert, A. Burlina, *et al.* (2012). Suggested guidelines for the diagnosis and management of urea cycle disorders. *Orphanet Journal of Rare Diseases*, 7:32.

C. v. d. Heiden, M. Brink, P. K. d. Bree, *et al.* (1978). Familial hyperlysinaemia due to L-lysine alpha-ketoglutarate reductase deficiency: Results of attempted treatment. *Journal of Inherited Metabolic Disease*, 1(3):89–94.

D. C. Henshall, M. Diaz-Hernandez, M. T. Miras-Portugal, *et al.* (2013). P2X receptors as targets for the treatment of status epilepticus. *Frontiers in Cellular Neuroscience*, 7:237.

J. K. Hiltunen, Z. Chen, A. M. Haapalainen, *et al.* (2010). Mitochondrial fatty acid synthesis–an adopted set of enzymes making a pathway of major importance for the cellular metabolism. *Progress in Lipid Research*, 49(1):27–45.

G. F. Hoffmann, B. Schmitt, M. Windfuhr, *et al.* (2007). Pyridoxal 5'-phosphate may be curative in early-onset epileptic encephalopathy. *Journal of Inherited Metabolic Disease*, 30(1):96–99.

L. A. Hoffmann, S. Jarius, H. L. Pellkofer, *et al.* (2008). Anti-Ma and anti-Ta associated paraneoplastic neurological syndromes: 22 newly diagnosed patients and review of previous cases. *Journal of Neurology, Neurosurgery, and Psychiatry*, 79(7):767–773.

T. U. Hoogenraad (2006). Paradigm shift in treatment of Wilson's disease: Zinc therapy now treatment of choice. *Brain and Development*, 28(3):141–146.

S. Hoppu, T. Härkönen, M. S. Ronkainen, *et al.* (2006). IA-2 antibody isotypes and epitope specificity during the prediabetic process in children with HLA-conferred susceptibility to type I diabetes. *Clinical and Experimental Immunology*, 144(1):59–66.

S. M. Houten, H. Te Brinke, S. Denis, *et al.* (2013). Genetic basis of hyperlysinemia. *Orphanet Journal of Rare Diseases*, 8:57.

M. F. Howard, Y. Murakami, A. T. Pagnamenta, *et al.* (2014). Mutations in PGAP3 impair GPI-anchor maturation, causing a subtype of hyperphosphatasia with mental retardation. *American Journal of Human Genetics*, 94(2):278–287.

C. C. Hsu, C. H. Cheng, C. L. Hsu, *et al.* (2015). Role of vitamin B6 status on antioxidant defenses, glutathione, and related enzyme activities in mice with homocysteine-induced oxidative stress. *Food & Nutrition Research*, 59:25702.

A. D. Hunt, J. Stokes, W. W. McCrory, *et al.* (1954). Pyridoxine dependency: report of a case of intractable convulsions in an infant controlled by pyridoxine. *Pediatrics*, 13(2):140–145.

R. R. S. Hussein, R. H. Soliman, A. M. Abdelhaleem Ali, *et al.* (2013). Effect of antiepileptic drugs on liver enzymes. *Beni-Suef University Journal of Basic and Applied Sciences*, 2(1):14–19.

O. Hwang (2013). Role of oxidative stress in Parkinson's disease. *Experimental Neurobiology*, 22(1):11–17.

S. Ikegawa, M. Isomura, Y. Koshizuka, *et al.* (1999). Cloning and characterization of human and mouse PROSC (proline synthetase co-transcribed) genes. *Journal of Human Genetics*, 44(5):337–342.

B. Ilkovski, A. T. Pagnamenta, G. L. O'Grady, *et al.* (2015). Mutations in PIGY: expanding the phenotype of inherited glycosylphosphatidylinositol deficiencies. *Human Molecular Genetics*, 24(21):6146–6159.

V. M. Ingram (1956). A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. *Nature*, 178(4537):792–794.

V. M. Ingram (1957). Gene mutations in human haemoglobin: the chemical difference between normal and sickle cell haemoglobin. *Nature*, 180(4581):326–328.

P. Irun, M. Mallén, C. Dominguez, *et al.* (2013). Identification of seven novel SMPD1 mutations causing Niemann-Pick disease types A and B. *Clinical Genetics*, 84(4):356–361.

E. Isaeva, D. Isaev, R. Khazipov, *et al.* (2009). Long-term suppression of GABAergic activity by neonatal seizures in rat somatosensory cortex. *Epilepsy Research*, 87(2-3):286–289.

M. Ito, T. Okuno, H. Hattori, *et al.* (1991). Vitamin B6 and valproic acid in treatment of infantile spasms. *Pediatric Neurology*, 7(2):91–96.

T. Ito, J. Iimori, S. Takayama, *et al.* (2013). Conserved pyridoxal protein that regulates Ile and Val metabolism. *Journal of Bacteriology*, 195(24):5439–5449.

G. I. Izuora, S. O. Iloeje (1989). Pyridoxine therapy on Nigerian children with infantile spasms. *East African Medical Journal*, 66(8):525–530.

B. Jaeger, N. G. Abeling, G. S. Salomons, *et al.* (2016). Pyridoxine responsive epilepsy caused by a novel homozygous PNPO mutation. *Molecular Genetics and Metabolism Reports*, 6:60–63.

C. D. Jayewickreme, R. A. Shivdasani (2015). Control of stomach smooth muscle development and intestinal rotation by transcription factor BARX1. *Developmental Biology*, 405(1):21–32.

A. Jezela-Stanek, E. Ciara, D. Piekutowska-Abramczuk, *et al.* (2016). Congenital disorder of glycosylphosphatidylinositol (GPI)-anchor biosynthesis—The phenotype of two patients with novel mutations in the PIGN and PGAP2 genes. *European Journal of Paediatric Neurology.*

F. Y. Jiao, D. Y. Gao, Y. Takuma, *et al.* (1997). Randomized, controlled trial of high-dose intravenous pyridoxine in the treatment of recurrent seizures in children. *Pediatric Neurology*, 17(1):54–57.

C. Jimenez-Mallebrera, S. Torelli, L. Feng, *et al.* (2009). A comparative study of alpha-dystroglycan glycosylation in dystroglycanopathies suggests that the hypoglycosylation of alpha-dystroglycan does not consistently correlate with clinical severity. *Brain Pathology*, 19(4):596–611.

J. O. Johnson, J. R. Gibbs, A. Megarbane, *et al.* (2012). Exome sequencing reveals riboflavin transporter mutations as a cause of motor neuron disease. *Brain*, 135(9):2875–2882.

M. A. Jones, D. Rhodenizer, C. da Silva, *et al.* (2013). Molecular diagnostic testing for congenital disorders of glycosylation (CDG): detection rate for single gene testing and next generation sequencing panel testing. *Molecular Genetics and Metabolism*, 110(1-2):78–85.

J. Kammermeier, S. Drury, C. T. James, *et al.* (2014). Targeted gene panel sequencing in children with very early onset inflammatory bowel disease–evaluation and prospective analysis. *Journal of Medical Genetics*, 51(11):748–755.

P. Kamoun, V. Richard, D. Rabier, *et al.* (2002). Plasma lysine concentration and availability of 2-ketoglutarate in liver mitochondria. *Journal of Inherited Metabolic Disease*, 25(1):1–6.

P. B. Kang, J. V. Hunter, E. M. Kaye (2001). Lactic acid elevation in extramitochondrial childhood neurodegenerative diseases. *Journal of Child Neurology*, 16(9):657–660.

O. Kann, R. Kovács (2007). Mitochondria and neuronal activity. *American Journal of Physiology Cell Physiology*, 292(2):C641–657.

P. Karimzadeh (2015). Approach to neurometabolic diseases from a pediatric neurological point of view. *Iranian Journal of Child Neurology*, 9(1):1–16.

D. Karolchik, A. S. Hinrichs, T. S. Furey, *et al.* (2004). The UCSC Table Browser data retrieval tool. *Nucleic Acids Research*, 32(Database issue):D493–496.

S. F. Kash, R. S. Johnson, L. H. Tecott, *et al.* (1997). Epilepsy in mice deficient in the 65-kDa isoform of glutamic acid decarboxylase. *Proceedings of the National Academy of Sciences of the United States of America*, 94(25):14060–14065.

M. Kato, T. Yamagata, M. Kubota, *et al.* (2013). Clinical spectrum of early onset epileptic encephalopathies caused by KCNQ2 mutation. *Epilepsia*, 54(7):1282–1287.

S. Köhler, M. H. Schulz, P. Krawitz, *et al.* (2009). Clinical diagnostics in human genetics with semantic similarity searches in ontologies. *The American Journal of Human Genetics*, 85(4):457–464.

H. R. Kim, H. W. Rho, J. W. Park, *et al.* (1994). Assay of ornithine aminotransferase with ninhydrin. *Analytical Biochemistry*, 223(2):205–207.

I. K. Kim, A. K. Niemi, C. Krueger, *et al.* (2013). Liver transplantation for urea cycle disorders in pediatric patients: a single-center experience. *Pediatric Transplantation*, 17(2):158–167.

P. K. Kim, D. W. Hailey, R. T. Mullen, *et al.* (2008). Ubiquitin signals autophagic degradation of cytosolic proteins and peroxisomes. *Proceedings of the National Academy of Sciences of the United States of America*, 105(52):20567–20574.

S. Kim, S. Y. Kim, H. J. Ku, *et al.* (2014). Suppression of tumorigenesis in mitochondrial NADP(+)-dependent isocitrate dehydrogenase knock-out mice. *Biochimica Et Biophysica Acta*, 1842(2):135–143.

Y. T. Kim, F. Kwok, J. E. Churchich (1988). Interactions of pyridoxal kinase and aspartate aminotransferase emission anisotropy and compartmentation studies. *The Journal of Biological Chemistry*, 263(27):13712–13717.

D. M. Kingsley, K. F. Kozarsky, L. Hobbie, *et al.* (1986). Reversible defects in O-linked glycosylation and LDL receptor expression in a UDP-Gal/UDP-GalNAc 4-epimerase deficient mutant. *Cell*, 44(5):749–759.

T. Kinoshita (2014). Biosynthesis and deficiencies of glycosylphosphatidylinositol. *Proceedings of the Japan Academy Series B, Physical and Biological Sciences*, 90(4):130–143.

M. Kircher, D. M. Witten, P. Jain, *et al.* (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics*, 46(3):310–315.

G. Klambauer, K. Schwarzbauer, A. Mayr, *et al.* (2012). cn.MOPS: mixture of Poissons for discovering copy number variations in next-generation sequencing data with a low false discovery rate. *Nucleic Acids Research*, page gks003.

D. J. Klionsky (2007). Autophagy: from phenomenology to molecular understanding in less than a decade. *Nature Reviews Molecular Cell Biology*, 8(11):931–937.

F. C. C. Klouwer, K. Berendse, S. Ferdinandusse, *et al.* (2015). Zellweger spectrum disorders: clinical overview and management approach. *Orphanet Journal of Rare Diseases*, 10:151.

D. C. Koboldt, Q. Zhang, D. E. Larson, *et al.* (2012). VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Research*, 22(3):568–576.

M. K. Koenig (2008). Presentation and diagnosis of mitochondrial disorders in children. *Pediatric Neurology*, 38(5):305–313.

S. Kohler, S. C. Doelken, C. J. Mungall, *et al.* (2014). The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. *Nucleic Acids Research*, 42(D1):D966–D974.

R. Konno (1998). Methods for the detection of D-amino-acid oxidase. *Biological Procedures Online*, 1:27–31.

V. Koprivica, D. L. Stone, J. K. Park, *et al.* (2000). Analysis and classification of 304 mutant alleles in patients with type 1 and type 3 Gaucher disease. *American Journal of Human Genetics*, 66(6):1777–1786.

S. H. Korman, N. Kanazawa, B. Abu-Libdeh, *et al.* (2004). Hyperornithinemia, hyperammonemia, and homocitrullinuria syndrome with evidence of mitochondrial dysfunction due to a novel SLC25A15 (ORNT1) gene mutation in a Palestinian family. *Journal of the Neurological Sciences*, 218(1-2):53–58.

N. Kouri, O. A. Ross, B. Dombroski, *et al.* (2015). Genome-wide association study of corticobasal degeneration identifies risk variants shared with progressive supranuclear palsy. *Nature Communications*, 6:7247.

E. M. Kowaloff, J. M. Phang, A. S. Granger, *et al.* (1977). Regulation of proline oxidase activity by lactate. *Proceedings of the National Academy of Sciences of the United States of America*, 74(12):5368–5371.

M. Kranendijk, E. A. Struys, E. van Schaftingen, *et al.* (2010). IDH2 mutations in patients with D-2-hydroxyglutaric aciduria. *Science*, 330(6002):336.

P. M. Krawitz, Y. Murakami, J. Hecht, *et al.* (2012). Mutations in PIGO, a member of the GPI-anchor-synthesis pathway, cause hyperphosphatasia with mental retardation. *American Journal of Human Genetics*, 91(1):146–151.

P. M. Krawitz, Y. Murakami, A. Rieß, *et al.* (2013). PGAP2 mutations, affecting the GPI-anchor-synthesis pathway, cause hyperphosphatasia with mental retardation syndrome. *American Journal of Human Genetics*, 92(4):584–589.

P. M. Krawitz, M. R. Schweiger, C. Rödelsperger, *et al.* (2010). Identity-by-descent filtering of exome sequence data identifies PIGV mutations in hyperphosphatasia mental retardation syndrome. *Nature Genetics*, 42(10):827–829.

I. Kuki, Y. Takahashi, S. Okazaki, *et al.* (2013). Vitamin B6-responsive epilepsy due to inherited GPI deficiency. *Neurology*, 81(16):1467–1469.

B. N. La Du, V. G. Zannoni, L. Laster, *et al.* (1958). The nature of the defect in tyrosine metabolism in alcaptonuria. *The Journal of Biological Chemistry*, 230(1):251–260.

C. Lam, G. A. Golas, M. Davids, *et al.* (2015). Expanding the clinical and molecular characteristics of PIGT-CDG, a disorder of glycosylphosphatidylinositol anchors. *Molecular Genetics and Metabolism*, 115(2-3):128–140.

F. Lamari, F. Mochel, F. Sedel, *et al.* (2013). Disorders of phospholipids, sphingolipids and fatty acids biosynthesis: toward a new category of inherited metabolic diseases. *Journal of Inherited Metabolic Disease*, 36(3):411–425.

T. Lamark, V. Kirkin, I. Dikic, *et al.* (2009). NBR1 and p62 as cargo receptors for selective autophagy of ubiquitinated targets. *Cell Cycle*, 8(13):1986–1990.

G. Lambrecht, J. Rettinger, H. G. Bäumert, *et al.* (2000). The novel pyridoxal-5'-phosphate derivative PPNDS potently antagonizes activation of P2X(1) receptors. *European Journal of Pharmacology*, 387(3):R19–21.

R. E. Lamont, W.-H. Tan, A. M. Innes, *et al.* (2016). Expansion of phenotype and genotypic data in CRB2-related syndrome. *European Journal of Human Genetics*, [Epub ahead of print].

E. S. Lander, L. M. Linton, B. Birren, *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature*, 409(6822):860–921.

K. M. Lang, R. A. Spritz (1983). RNA splice site selection: evidence for a 5' leads to 3' scanning model. *Science*, 220(4604):1351–1355.

B. Lanpher, N. Brunetti-Pierri, B. Lee (2006). Inborn errors of metabolism: the flux from Mendelian to complex diseases. *Nature Reviews Genetics*, 7(6):449–460.

H. B. Laurén, F. R. Lopez-Picon, E. R. Korpi, *et al.* (2005). Kainic acid-induced status epilepticus alters GABA receptor subunit mRNA and protein expression in the developing rat hippocampus. *Journal of Neurochemistry*, 94(5):1384–1394.

B. Law (1990). Use of silica with reversed-phase type eluents for the analysis of basic drugs and metabolites. *Trends in Analytical Chemistry*, 9(1):31–36.

H. J. Lee, S. Y. Kim, J. M. Koh, *et al.* (2007). Polymorphisms and haplotypes of integrinalpha1 (ITGA1) are associated with bone mineral density and fracture risk in postmenopausal Koreans. *Bone*, 41(6):979–986.

J. A. Lee, K. Inoue, S. W. Cheung, *et al.* (2006). Role of genomic architecture in PLP1 duplication causing Pelizaeus-Merzbacher disease. *Human Molecular Genetics*, 15(14):2250–2265.

J. R. Lemke, E. Riesch, T. Scheurenbrand, *et al.* (2012). Targeted next generation sequencing as a diagnostic tool in epileptic disorders. *Epilepsia*, 53(8):1387–1398.

J. M. S. Lemons, X. J. Feng, B. D. Bennett, *et al.* (2010). Quiescent fibroblasts exhibit high metabolic activity. *PLoS Biol*, 8(10):e1000514.

J. V. Leonard, A. A. M. Morris (2002). Urea cycle disorders. *Seminars in Neonatology*, 7(1):27–35.

P. Leone, D. Shera, S. W. McPhee, *et al.* (2012). Long-term follow-up after gene therapy for Canavan disease. *Science Translational Medicine*, 4(165):165ra163.

I. U. S. Leong, A. Stuckey, D. Lai, *et al.* (2015). Assessment of the predictive accuracy of five in silico prediction tools, alone or in combination, and two metaservers to classify long QT syndrome gene mutations. *BMC Medical Genetics*, 16:34.

B. Levine, G. Kroemer (2008). Autophagy in the pathogenesis of disease. *Cell*, 132(1):27–42.

A. Levtova, S. Camuzeaux, A. M. Laberge, *et al.* (2015). Normal cerebrospinal fluid pyridoxal 5'-phosphate level in a PNPO-deficient patient with neonatal-onset epileptic encephalopathy. *JIMD Reports*, 22:67–75.

S. L. Levy, W. M. Burnham, A. Bishai, *et al.* (1992). The anticonvulsant effects of vitamin E: a further evaluation. *The Canadian Journal of Neurological Sciences*, 19(2):201–203.

H. Li, R. Durbin (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14):1754–1760.

S. C. Li, N. K. Goto, K. A. Williams, *et al.* (1996). Alpha-helical, but not beta-sheet, propensity of proline is determined by peptide environment. *Proceedings of the National Academy of Sciences of the United States of America*, 93(13):6676–6681.

X. Li, J. Zhou, Z. Chen, *et al.* (2008). Long-term expressional changes of Na+ -K+ -Cl- co-transporter 1 (NKCC1) and K+ -Cl- co-transporter 2 (KCC2) in CA1 region of hippocampus following lithium-pilocarpine induced status epilepticus (PISE). *Brain Research*, 1221:141–146.

C. Liao, F. Fu, R. Li, *et al.* (2013). Loss-of-function variation in the DPP6 gene is associated with autosomal dominant microcephaly and mental retardation. *European Journal of Medical Genetics*, 56(9):484–489.

O. Lidove, M. L. West, G. Pintos-Morell, *et al.* (2010). Effects of enzyme replacement therapy in Fabry disease–a comprehensive review of the medical literature. *Genetics in Medicine*, 12(11):668–679.

S. Linn, M. Kairis, R. Holliday (1976). Decreased fidelity of DNA polymerase activity isolated from aging human fibroblasts. *Proceedings of the National Academy of Sciences*, 73(8):2818–2822.

M. Linnebank, E. Tschiedel, J. Häberle, *et al.* (2002). Argininosuccinate lyase (ASL) deficiency: mutation analysis in 27 patients and a completed structure of the human ASL gene. *Human Genetics*, 111(4-5):350–359.

N. S. Lipman, L. R. Jackson, L. J. Trudel, *et al.* (2005). Monoclonal versus polyclonal antibodies: distinguishing characteristics, applications, and information resources. *ILAR Journal*, 46(3):258–268.

P. Liu, M. P. Torrens-Spence, H. Ding, *et al.* (2013). Mechanism of cysteine-dependent inactivation of aspartate/glutamate/cysteine sulfinic acid a-decarboxylases. *Amino Acids*, 44(2):391–404.

W. Löscher (2011). Critical review of current animal models of seizures and epilepsy used in the discovery and development of new antiepileptic drugs. *Seizure*, 20(5):359–368.

S. P. Lu, A. H. Lewin (1998). Enamine/imine tautomerism in alpha,beta-unsaturated-alpha-amino acids. *Tetrahedron*, 54(50):15097–15104.

C. J. Lynch (2001). Role of leucine in the regulation of mTOR by amino acids: revelations from structure-activity studies. *The Journal of Nutrition*, 131(3):861S–865S.

D. G. MacArthur, T. A. Manolio, D. P. Dimmock, *et al.* (2014). Guidelines for investigating causality of sequence variants in human disease. *Nature*, 508(7497):469–476.

A. D. Mackey, R. J. McMahon, J. H. Townsend, *et al.* (2004). Uptake, hydrolysis, and metabolism of pyridoxine-5'-beta-D-glucoside in Caco-2 cells. *The Journal of Nutrition*, 134(4):842–846.

M. M. Mahfouz, S. Q. Zhou, F. A. Kummerow (2009). Vitamin B6 compounds are capable of reducing the superoxide radical and lipid peroxide levels induced by H2O2 in vascular endothelial cells in culture. *International Journal for Vitamin and Nutrition Research*, 79(4):218–229.

S. Maljevic, T. V. Wuttke, H. Lerche (2008). Nervous system KV7 disorders: breakdown of a subthreshold brake. *The Journal of Physiology*, 586(7):1791–1801.

T. Mammone, D. Gan, R. Foyouzi-Youssefi (2006). Apoptotic cell death increases with senescence in normal human dermal fibroblast cultures. *Cell Biology International*, 30(11):903–909.

V. Manwaring, W. E. Heywood, R. Clayton, *et al.* (2013). The identification of new biomarkers for identifying and monitoring kidney disease and their translation into a rapid mass spectrometry-based test: evidence of presymptomatic kidney disease in pediatric Fabry and type-I diabetic patients. *Journal of Proteome Research*, 12(5):2013–2021.

J. Marcos, N. Renau, O. Valverde, *et al.* (2016). Targeting tryptophan and tyrosine metabolism by liquid chromatography tandem mass spectrometry. *Journal of Chromatography A*, 1434:91–101.

A. Marian (2014). Sequencing your genome: what does it mean? *Methodist DeBakey Cardiovascular Journal*, 10(1):3–6.

E. Mariani, M. C. Polidori, A. Cherubini, *et al.* (2005). Oxidative stress in brain aging, neurodegenerative and vascular diseases: an overview. *Journal of Chromatography B, Analytical Technologies in the Biomedical and Life Sciences*, 827(1):65–75.

W. R. Markesbery (1997). Oxidative stress hypothesis in Alzheimer's disease. *Free Radical Biology & Medicine*, 23(1):134–147.

N. V. Marrion (1997). Control of M-current. *Annual Review of Physiology*, 59:483–504.

M. Martin (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal*, 17(1):pp. 10–12.

P. T. Martin (2007). Congenital muscular dystrophies involving the O-mannose pathway. *Current Molecular Medicine*, 7(4):417–425.

D. Martinelli, D. Diodato, E. Ponzi, *et al.* (2015). The hyperornithinemia-hyperammonemia-homocitrullinuria syndrome. *Orphanet Journal of Rare Diseases*, 10:29.

J. Marvit, A. G. DiLella, K. Brayton, *et al.* (1987). GT to AT transition at a splice donor site causes skipping of the preceding exon in phenylketonuria. *Nucleic Acids Research*, 15(14):5613–5628.

J. M. Mason, M. D. Naidu, M. Barcia, *et al.* (2004). IL-4-induced gene-1 is a leukocyte L-amino acid oxidase with an unusual acidic pH preference and lysosomal localization. *Journal of Immunology*, 173(7):4561–4567.

A. Masri, S. A. Wahsh (2014). Manifestations and treatment of epilepsy in children with neurometabolic disorders: A series from Jordan. *Seizure*, 23(1):10–15.

T. W. Massingale, S. Buttross (1993). Survey of treatment practices for neonatal seizures. *Journal of Perinatology*, 13(2):107–110.

D. Mathis, L. Abela, M. Albersen, *et al.* (2016). The value of plasma vitamin B6 profiles in early onset epileptic encephalopathies. *Journal of Inherited Metabolic Disease*, [Epub ahead of print].

J. M. Matés (2000). Effects of antioxidant enzymes in the molecular control of reactive oxygen species toxicology. *Toxicology*, 153(1-3):83–104.

J. G. McCormack, R. M. Denton (1993). Mitochondrial Ca2+ transport and the role of intramitochondrial Ca2+ in the regulation of energy metabolism. *Developmental Neuroscience*, 15(3-5):165–173.

A. McKenna, M. Hanna, E. Banks, *et al.* (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9):1297–1303.

V. A. McKusick (1969). On lumpers and splitters, or the nosology of genetic disease. *Perspectives in Biology and Medicine*, 12(2):298–312.

H. C. Mefford, J. Cook, S. M. Gospe (2012). Epilepsy due to 20q13.33 subtelomere deletion masquerading as pyridoxine-dependent epilepsy. *American Journal of Medical Genetics Part A*, 158A(12):3190–3195.

M. Meier, M. Janosik, V. Kery, *et al.* (2001). Structure of human cystathionine beta-synthase: a unique pyridoxal 5'-phosphate-dependent heme protein. *The EMBO Journal*, 20(15):3910–3916.

A. Meister (1994). Glutathione-ascorbic acid antioxidant system in animals. *The Journal of Biological Chemistry*, 269(13):9397–9400.

B. Menon, K. Ramalingam, R. V. Kumar (2012). Oxidative stress in patients with epilepsy is independent of antiepileptic drugs. *Seizure*, 21(10):780–784.

S. Mercimek-Mahmutoglu, J. Patel, D. Cordeiro, *et al.* (2015). Diagnostic yield of genetic testing in epileptic encephalopathy in childhood. *Epilepsia*, 56(5):707–716.

F. Miceli, M. V. Soldovieri, P. Ambrosino, *et al.* (2013). Genotype-phenotype correlations in neonatal epilepsies caused by mutations in the voltage sensor of K(v)7.2 potassium channel subunits. *Proceedings of the National Academy of Sciences of the United States of America*, 110(11):4386–4391.

O. Midttun, S. Hustad, E. Solheim, *et al.* (2005). Multianalyte quantification of vitamin B6 and B2 species in the nanomolar range in human plasma by liquid chromatography-tandem mass spectrometry. *Clinical Chemistry*, 51(7):1206–1216.

Y. Mikhed, A. Görlach, U. G. Knaus, *et al.* (2015). Redox regulation of genome stability by effects on gene expression, epigenetic pathways and DNA damage/repair. *Redox Biology*, 5:275–289.

M. Milh, C. Lacoste, P. Cacciagli, *et al.* (2015). Variable clinical expression in patients with mosaicism for KCNQ2 mutations. *American Journal of Medical Genetics Part A*, 167A(10):2314–2318.

C. Miller (2000). An overview of the potassium channel family. *Genome Biology*, 1(4):reviews0004.1–reviews0004.5.

G. Miller, A. Honig, H. Stein, *et al.* (2009). Unraveling delta1-pyrroline-5-carboxylate-proline cycle in plants by uncoupled expression of proline oxidation enzymes. *The Journal of Biological Chemistry*, 284(39):26482–26492.

P. B. Mills, S. S. M. Camuzeaux, E. J. Footitt, *et al.* (2014). Epilepsy due to PNPO mutations: genotype, environment and treatment affect presentation and outcome. *Brain*, 137(5):1350–1360.

P. B. Mills, E. J. Footitt, K. A. Mills, *et al.* (2010). Genotypic and phenotypic spectrum of pyridoxine-dependent epilepsy (ALDH7A1 deficiency). *Brain*, 133(7):2148–2159.

P. B. Mills, E. Struys, C. Jakobs, *et al.* (2006). Mutations in antiquitin in individuals with pyridoxine-dependent seizures. *Nature Medicine*, 12(3):307–309.

P. B. Mills, R. A. H. Surtees, M. P. Champion, *et al.* (2005). Neonatal epileptic encephalopathy caused by mutations in the PNPO gene encoding pyridox(am)ine 5'-phosphate oxidase. *Human Molecular Genetics*, 14(8):1077–1086.

D. Mishra, V. Kalra, R. Seth, *et al.* (2010). Efficacy of pyridoxine in early-onset idiopathic intractable seizures in children. *Indian Journal of Pediatrics*, 77(11):1252–1256.

R. S. Møller, H. A. Dahl, I. Helbig (2015). The contribution of next generation sequencing to epilepsy genetics. *Expert Review of Molecular Diagnostics*, 15(12):1531–1538.

P. V. Mohanan, H. Yamamoto (2002). Preventive effect of melatonin against brain mitochondria DNA damage, lipid peroxidation and seizures induced by kainic acid. *Toxicology Letters*, 129(1-2):99–105.

F. Molinari, A. Kaminska, G. Fiermonte, *et al.* (2009). Mutations in the mitochondrial glutamate carrier SLC25A22 in neonatal epileptic encephalopathy with suppression bursts. *Clinical Genetics*, 76(2):188–194.

F. Molinari, A. Raas-Rothschild, M. Rio, *et al.* (2005). Impaired mitochondrial glutamate transport in autosomal recessive neonatal myoclonic epilepsy. *American Journal of Human Genetics*, 76(2):334–339.

G. E. Mortimore, C. M. Schworer (1977). Induction of autophagy by amino-acid deprivation in perfused rat liver. *Nature*, 270(5633):174–176.

D. H. Morton, K. A. Strauss, D. L. Robinson, *et al.* (2002). Diagnosis and treatment of maple syrup disease: a study of 36 patients. *Pediatrics*, 109(6):999–1008.

M. A. Moxley, J. J. Tanner, D. F. Becker (2011). Steady-state kinetic mechanism of the proline:ubiquinone oxidoreductase activity of proline utilization A (PutA) from Escherichia coli. *Archives of Biochemistry and Biophysics*, 516(2):113–120.

R. Mukherjee, N. E. Brasch (2011). Kinetic studies on the reaction between cob(I)alamin and peroxynitrite: rapid oxidation of cob(I)alamin to cob(II)alamin by peroxynitrous acid. *Chemistry*, 17(42):11723–11727.

S. N. Murthy, M. K. Janardanasarma (1999). Identification of L-amino acid/L-lysine alpha-amino oxidase in mouse brain. *Molecular and Cellular Biochemistry*, 197(1-2):13–23.

F. N. Musayev, M. L. Di Salvo, T. P. Ko, *et al.* (2003). Structure and properties of recombinant human pyridoxine 5'-phosphate oxidase. *Protein Science*, 12(7):1455–1463.

F. N. Musayev, M. L. di Salvo, T.-P. Ko, *et al.* (2007). Crystal Structure of human pyridoxal kinase: Structural basis of M+ and M2+ activation. *Protein Science*, 16(10):2184–2194.

F. N. Musayev, M. L. Di Salvo, M. A. Saavedra, *et al.* (2009). Molecular basis of reduced pyridoxine 5'-phosphate oxidase catalytic activity in neonatal epileptic encephalopathy disorder. *The Journal of Biological Chemistry*, 284(45):30949–30956.

M. Naglak, R. Salvo, K. Madsen, *et al.* (1988). The treatment of isovaleric acidemia with glycine supplement. *Pediatric Research*, 24(1):9–13.

W. Namwat, H. Kinoshita, T. Nihira (2002). Identification by heterologous expression and gene disruption of VisA as L-lysine 2-aminotransferase essential for virginiamycin S biosynthesis in Streptomyces virginiae. *Journal of Bacteriology*, 184(17):4811–4818.

J. Natera, W. Massad, N. A. García (2012). The role of vitamin B6 as an antioxidant in the presence of vitamin B2-photogenerated reactive oxygen species. A kinetic and mechanistic study. *Photochemical & Photobiological Sciences*, 11(6):938–945.

J. T. Neary, W. F. Diven (1970). Purification, properties, and a possible mechanism for pyridoxal kinase from bovine brain. *The Journal of Biological Chemistry*, 245(21):5585–5593.

M. Negahdaripour, N. Nezafat, Y. Ghasemi (2016). A panoramic review and in silico analysis of IL-11 structure and function. *Cytokine & Growth Factor Reviews*, [Epub ahead of print].

H. M. Nguyen, H. Miyazaki, N. Hoshi, *et al.* (2012). Modulation of voltage-gated K+ channels by the sodium channel B1 subunit. *Proceedings of the National Academy of Sciences of the United States of America*, 109(45):18577–18582.

N. Y. N. Nguyen, M. J. Maxwell, L. M. Ooms, *et al.* (2011). An ENU-induced mouse mutant of SHIP1 reveals a critical role of the stem cell isoform for suppression of macrophage activation. *Blood*, 117(20):5362–5371.

H. Ni, Y. W. Jiang, T. Bo, *et al.* (2005). c-Fos, N-methyl-d-aspartate receptor 2C, GABA-A-a1 immonoreactivity, seizure latency and neuronal injury following single or recurrent neonatal seizures in hippocampus of Wistar rat. *Neuroscience Letters*, 380(1–2):149–154.

V. Niketić, S. Ristić, Z. S. Saicić, *et al.* (1995). Activities of antioxidant enzymes and formation of the glutathione adduct of hemoglobin (Hb ASSG) in epileptic patients with long-term antiepileptic therapy. *Farmaco*, 50(11):811–813.

Y. Nishina, K. Sato, K. Shiga (1991). Isomerization of delta 1-piperideine-2-carboxylate to delta 2-piperideine-2-carboxylate on complexation with flavoprotein D-amino acid oxidase. *Journal of Biochemistry*, 109(5):705–710.

J. D. Nissen, K. Pajęcka, M. H. Stridh, *et al.* (2015). Dysfunctional TCA-Cycle Metabolism in Glutamate Dehydrogenase Deficient Astrocytes. *Glia*, 63(12):2313–2326.

R. A. North, A. Surprenant (2000). Pharmacology of cloned P2X receptors. *Annual Review of Pharmacology and Toxicology*, 40:563–580.

G. Novarino, P. El-Fishawy, H. Kayserili, *et al.* (2012). Mutations in BCKD-kinase lead to a potentially treatable form of autism with epilepsy. *Science*, 338(6105):394–397.

A. L. Numis, M. Angriman, J. E. Sullivan, *et al.* (2014). KCNQ2 encephalopathy: delineation of the electroclinical phenotype and treatment response. *Neurology*, 82(4):368–370.

A. Oddsson, P. Sulem, H. Helgason, *et al.* (2015). Common and rare variants associated with kidney stones and biochemical traits. *Nature Communications*, 6:7975.

C. Ohba, N. Okamoto, Y. Murakami, *et al.* (2014). PIGN mutations cause congenital anomalies, developmental delay, hypotonia, epilepsy, and progressive cerebellar atrophy. *Neurogenetics*, 15(2):85–92.

B. K. Ohta, C. S. Foote (2002). Characterization of endoperoxide and hydroperoxide intermediates in the reaction of pyridoxine with singlet oxygen. *Journal of the American Chemical Society*, 124(41):12064–12065.

S. Ohtahara, Y. Yamatogi, Y. Ohtsuka (2011). Vitamin B(6) treatment of intractable seizures. *Brain & Development*, 33(9):783–789.

Y. Ohtsuka, T. Ogino, T. Asano, *et al.* (2000). Long-term follow-up of vitamin B(6)-responsive West syndrome. *Pediatric Neurology*, 23(3):202–206.

U. Omasits, C. H. Ahrens, S. Müller, *et al.* (2014). Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics*, 30(6):884–886.

G. Orhan, M. Bock, D. Schepers, *et al.* (2014). Dominant-negative effects of KCNQ2 mutations are associated with epileptic encephalopathy. *Annals of Neurology*, 75(3):382–394.

H. Orimo, H. J. Girschick, M. Goseki-Sone, *et al.* (2001). Mutational analysis and functional correlation with phenotype in German patients with childhood-type hypophosphatasia. *Journal of Bone and Mineral Research*, 16(12):2313–2319.

A. Ormazabal, M. Oppenheim, M. Serrano, *et al.* (2008). Pyridoxal 5'-phosphate values in cerebrospinal fluid: reference values and diagnosis of PNPO deficiency in paediatric patients. *Molecular Genetics and Metabolism*, 94(2):173–177.

C. N. Pace, J. M. Scholtz (1998). A helix propensity scale based on experimental studies of peptides and proteins. *Biophysical Journal*, 75(1):422–427.

F. Palmieri (2014). Mitochondrial transporters of the SLC25 family and associated diseases: a review. *Journal of Inherited Metabolic Disease*, 37(4):565–575.

F. Palmieri, C. Indiveri, F. Bisaccia, *et al.* (1995). Mitochondrial metabolite carrier proteins: purification, reconstitution, and transport studies. *Methods in Enzymology*, 260:349–369.

F. Palmieri, C. L. Pierri (2010). Structure and function of mitochondrial carriers - role of the transmembrane helix P and G residues in the gating and transport mechanism. *FEBS Letters*, 584(9):1931–1939.

Z. Pan, T. Kao, Z. Horvath, *et al.* (2006). A common ankyrin-G-based mechanism retains KCNQ and NaV channels at electrically active domains of the axon. *The Journal of Neuroscience*, 26(10):2599–2613.

F. Papes, M. J. Surpili, F. Langone, *et al.* (2001). The essential amino acid lysine acts as precursor of glutamate in the mammalian central nervous system. *FEBS Letters*, 488(1–2):34–38.

L. Papetti, P. Parisi, V. Leuzzi, *et al.* (2013). Metabolic epilepsy: an update. *Brain & Development*, 35(9):827–841.

M. D. Paraskevopoulou, G. Georgakilas, N. Kostoulas, *et al.* (2013). DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. *Nucleic Acids Research*, 41(Web Server issue):W169–173.

S. Parikh, R. Saneto, M. J. Falk, *et al.* (2009). A modern approach to the treatment of mitochondrial disease. *Current Treatment Options in Neurology*, 11(6):414–430.

G. M. Pastores, P. Arn, M. Beck, *et al.* (2007). The MPS I registry: design, methodology, and early findings of a global disease registry for monitoring patients with Mucopolysaccharidosis Type I. *Molecular Genetics and Metabolism*, 91(1):37–47.

L. Pauling, H. A. Itano, S. J. Singer, *et al.* (1949). Sickle cell anemia a molecular disease. *Science*, 110(2865):543–548.

R. Penzel, J. Uhl, J. Kopitz, *et al.* (2001). Splice donor site mutation in the lysosomal neuraminidase gene causing exon skipping and complete loss of enzyme activity in a sialidosis patient. *FEBS Letters*, 501(2-3):135–138.

S. Perry, H. Harries, C. Scholfield, *et al.* (1995). Molecular cloning and expression of a cDNA for human kidney cysteine conjugate beta-lyase. *FEBS letters*, 360(3):277–280.

G. Petrof, A. Nanda, J. Howden, *et al.* (2014). Mutations in GRHL2 result in an autosomal-recessive ectodermal Dysplasia syndrome. *American Journal of Human Genetics*, 95(3):308–314.

O. A. Petroff, D. L. Rothman, K. L. Behar, *et al.* (1996). Low brain GABA level is associated with poor seizure control. *Annals of Neurology*, 40(6):908–911.

T. M. Pierson, D. Adams, F. Bonn, *et al.* (2011). Whole-exome sequencing identifies homozygous AFG3L2 mutations in a spastic ataxia-neuropathy syndrome linked to mitochondrial m-AAA proteases. *PLoS Genetics*, 7(10).

J. Pietz, C. Benninger, H. Schäfer, *et al.* (1993). Treatment of infantile spasms with high-dosage vitamin B6. *Epilepsia*, 34(4):757–763.

J. J. Pitt (2009). Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *The Clinical Biochemist Reviews*, 30(1):19–34.

V. Plagnol, J. Curtis, M. Epstein, *et al.* (2012). A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics*, 28(21):2747–2754.

B. Plecko, K. Paul, P. Mills, *et al.* (2014). Pyridoxine responsiveness in novel mutations of the PNPO gene. *Neurology*, 82(16):1425–1433.

A. Poduri, E. L. Heinzen, V. Chitsazzadeh, *et al.* (2013). SLC25A22 is a novel gene for migrating partial seizures in infancy. *Annals of Neurology*, 74(6):873–882.

F. Ponticelli, R. Pagani, M. C. Missale, *et al.* (1983). Synthesis and NMR properties of 2-(4-pyridyl) thiazolidine-4-carboxylic acids. *The Italian Journal of Biochemistry*, 32(2):92–101.

I. Pérez-Arellano, F. Carmona-Alvarez, A. I. Martínez, *et al.* (2010). Pyrroline-5-carboxylate synthase and proline biosynthesis: from osmotolerance to rare metabolic disease. *Protein Science*, 19(3):372–382.

L. Prunetti, B. El Yacoubi, C. R. Schiavon, *et al.* (2016). Evidence That COG0325 Proteins are involved in PLP Homeostasis. *Microbiology*, [Epub ahead of print].

M. Pupavac, X. Tian, J. Chu, *et al.* (2016). Added value of next generation gene panel analysis for patients with elevated methylmalonic acid and no clinical diagnosis following functional studies of vitamin B12 metabolism. *Molecular Genetics and Metabolism*, 117(3):363–368.

O. W. J. Quarrell, M. A. Nance, P. Nopoulos, *et al.* (2013). Managing juvenile Huntington's disease. *Neurodegenerative Disease Management*, 3(3):267–276.

C. M. Quinzii, S. DiMauro, M. Hirano (2007). Human coenzyme Q10 deficiency. *Neurochemical Research*, 32(4-5):723–727.

B. Rabbani, M. Tekin, N. Mahdieh (2014). The promise of whole-exome sequencing in medical genetics. *Journal of Human Genetics*, 59(1):5–15.

F. Raimondi, P. Mills, P. T. Clayton, *et al.* (2015). A preterm neonate with seizures unresponsive to conventional treatment. *BMJ Case Reports*, 2015:bcr2015209743.

P. M. Rankin, S. Harrison, W. K. Chong, *et al.* (2007). Pyridoxine-dependent seizures: a family phenotype that leads to severe cognitive deficits, regardless of treatment regime. *Developmental Medicine and Child Neurology*, 49(4):300–305.

V. V. Rao, X. Pan, Y. F. Chang (1992). Developmental changes of L-lysine-ketoglutarate reductase in rat brain and liver. *Comparative Biochemistry and Physiology B, Comparative Biochemistry*, 103(1):221–224.

M. G. Reese, F. H. Eeckman, D. Kulp, *et al.* (1997). Improved splice site detection in Genie. *Journal of Computational Biology*, 4(3):311–323.

H. L. Rehm, S. J. Bale, P. Bayrak-Toydemir, *et al.* (2013). ACMG clinical laboratory standards for next-generation sequencing. *Genetics in Medicine*, 15(9):733–747.

J. K. Reichardt, J. W. Belmont, H. L. Levy, *et al.* (1992). Characterization of two missense mutations in human galactose-1-phosphate uridyltransferase: different molecular mechanisms for galactosemia. *Genomics*, 12(3):596–600.

J. K. Reichardt, S. Packman, S. L. Woo (1991). Molecular characterization of two galactosemia mutations: correlation of mutations with highly conserved domains in galactose-1-phosphate uridyl transferase. *American Journal of Human Genetics*, 49(4):860–867.

R. D. Reynolds (1988). Bioavailability of vitamin B-6 from plant foods. *The American Journal of Clinical Nutrition*, 48(Suppl 3):863–867.

B. M. Rezk, G. R. M. M. Haenen, W. J. F. van der Vijgh, *et al.* (2003). Tetrahydrofolate and 5-methyltetrahydrofolate are folates with high antioxidant activity. Identification of the antioxidant pharmacophore. *FEBS Letters*, 555(3):601–605.

S. Richards, N. Aziz, S. Bale, *et al.* (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*, 17(5):405–424.

P. Rinaldo, T. M. Cowan, D. Matern (2008). Acylcarnitine profile analysis. *Genetics in Medicine*, 10(2):151–156.

A. J. Robinson, C. Overy, E. R. S. Kunji (2008). The mechanism of transport by mitochondrial carriers based on analysis of symmetry. *Proceedings of the National Academy of Sciences of the United States of America*, 105(46):17766–17771.

J. T. Robinson, H. Thorvaldsdóttir, W. Winckler, *et al.* (2011). Integrative genomics viewer. *Nature Biotechnology*, 29(1):24–26.

L. Rochette, S. Ghibu, C. Richard, *et al.* (2013). Direct and indirect antioxidant properties of alpha-lipoic acid and therapeutic potential. *Molecular Nutrition & Food Research*, 57(1):114–125.

M. E. Rodríguez-García, E. Martín-Hernández, A. M. de Aragón, *et al.* (2016). First missense mutation outside of SERAC1 lipase domain affecting intracellular cholesterol trafficking. *Neurogenetics*, 17(1):51–56.

Y. Rong, M. Baudry (1996). Seizure activity results in a rapid induction of nuclear factor-kappa B in adult but not juvenile rat limbic structures. *Journal of Neurochemistry*, 67(2):662–668.

M. B. D. l. Rosa, S. W. Nelson (2011). An interaction between the Walker A and D-loop motifs Is critical to ATP hydrolysis and cooperativity in bacteriophage T4 Rad50. *Journal of Biological Chemistry*, 286(29):26258–26266.

D. A. Roth-Maier, S. I. Kettler, M. Kirchgessner (2002). Availability of vitamin B6 from different food sources. *International Journal of Food Sciences and Nutrition*, 53(2):171–179.

S. Rowley, M. Patel (2013). Mitochondrial involvement and oxidative stress in temporal lobe epilepsy. *Free Radical Biology & Medicine*, 62:121–131.

J. Rumià, F. Marmol, J. Sanchez, *et al.* (2013). Oxidative stress markers in the neocortex of drug-resistant epilepsy patients submitted to epilepsy surgery. *Epilepsy Research*, 107(1-2):75–81.

M. E. Rybak, C. M. Pfeiffer (2004). Clinical analysis of vitamin B(6): determination of pyridoxal 5'-phosphate and 4-pyridoxic acid in human serum by reversed-phase high-performance liquid chromatography with chlorite postcolumn derivatization. *Analytical Biochemistry*, 333(2):336–344.

S. M. Sadrzadeh, J. W. Eaton (1988). Hemoglobin-mediated oxidant damage to the central nervous system requires endogenous ascorbate. *The Journal of Clinical Investigation*, 82(5):1510–1515.

M. K. Safo, F. N. Musayev, M. L. di Salvo, *et al.* (2006). Crystal structure of pyridoxal kinase from the Escherichia coli pdxK gene: implications for the classification of pyridoxal kinases. *Journal of Bacteriology*, 188(12):4542–4552.

M. K. Safo, F. N. Musayev, V. Schirch (2005). Structure of Escherichia coli pyridoxine 5'-phosphate oxidase in a tetragonal crystal form: insights into the mechanistic pathway of the enzyme. *Acta Crystallographica Section D, Biological Crystallography*, 61(5):599–604.

H. M. Said (2004). Recent advances in carrier-mediated intestinal absorption of water-soluble vitamins. *Annual Review of Physiology*, 66:419–446.

Z. M. Said, V. S. Subramanian, N. D. Vaziri, *et al.* (2008). Pyridoxine uptake by colonocytes: a specific and regulated carrier-mediated process. *American Journal of Physiology Cell Physiology*, 294(5):C1192–1197.

F. Sanger, S. Nicklen, A. R. Coulson (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 74(12):5463–5467.

R. Santer, J. Rischewski, M. von Weihe, *et al.* (2005). The spectrum of aldolase B (ALDOB) mutations and the prevalence of hereditary fructose intolerance in Central Europe. *Human Mutation*, 25(6):594.

C. B. Saper (2009). A guide to the perplexed on the specificity of antibodies. *The Journal of Histochemistry and Cytochemistry*, 57(1):1–5.

J. F. Sathirapongsasuti, H. Lee, B. A. J. Horst, *et al.* (2011). Exome sequencing-based copy-number variation and loss of heterozygosity detection: ExomeCNV. *Bioinformatics*, 27(19):2648–2654.

S. W. Sauer, S. Opp, G. F. Hoffmann, *et al.* (2011). Therapeutic modulation of cerebral l-lysine metabolism in a mouse model for glutaric aciduria type I. *Brain*, 134(1):157–170.

M. C. Schaeffer, D. Gretz, J. D. Mahuren, *et al.* (1995). Tissue B-6 vitamer concentrations in rats fed excess vitamin B-6. *The Journal of Nutrition*, 125(9):2370–2378.

H. Schaumburg, J. Kaplan, A. Windebank, *et al.* (1983). Sensory neuropathy from pyridoxine abuse. A new megavitamin syndrome. *The New England Journal of Medicine*, 309(8):445–448.

R. Schmid, E. Schulte-Frohlinde, V. Schusdziarra, *et al.* (1992). Contribution of postprandial amino acid levels to stimulation of insulin, glucagon, and pancreatic polypeptide in humans. *Pancreas*, 7(6):698–704.

B. Schmitt, M. Baumgartner, P. B. Mills, *et al.* (2010). Seizures and paroxysmal events: symptoms pointing to the diagnosis of pyridoxine-dependent epilepsy and pyridoxine phosphate oxidase deficiency. *Developmental Medicine and Child Neurology*, 52(7):e133–142.

T. D. Schmittgen, K. J. Livak (2008). Analyzing real-time PCR data by the comparative CT method. *Nature Protocols*, 3(6):1101–1108.

B. C. Schroeder, C. Kubisch, V. Stein, *et al.* (1998). Moderate loss of function of cyclic-AMP-modulated KCNQ2/KCNQ3 K+ channels causes epilepsy. *Nature*, 396(6712):687–690.

K. H. Schulpis, C. Lazaropoulou, S. Regoutas, *et al.* (2006). Valproic acid monotherapy induces DNA oxidative damage. *Toxicology*, 217(2-3):228–232.

R. Schwartz, N. O. Kjeldgaard (1951). The enzymic oxidation of pyridoxal by liver aldehyde oxidase. *Biochemical Journal*, 48(3):333–337.

J. M. Schwarz, C. Rödelsperger, M. Schuelke, *et al.* (2010). MutationTaster evaluates disease-causing potential of sequence alterations. *Nature Methods*, 7(8):575–576.

U. Schwarze, B. J. Starman, P. H. Byers (1999). Redefinition of exon 7 in the COL1A1 gene of type I collagen by an intron 8 splice-donor–site mutation in a form of osteogenesis imperfecta: influence of intron splice order on outcome of splice-site mutation. *The American Journal of Human Genetics*, 65(2):336–344.

K. Scott, T. Gadomski, T. Kozicz, *et al.* (2014). Congenital disorders of glycosylation: new defects and still counting. *Journal of Inherited Metabolic Disease*, 37(4):609–617.

C. R. Scriver, P. J. Waters (1999). Monogenic traits are not simple: lessons from phenylketonuria. *Trends in Genetics*, 15(7):267–272.

F. Sedel, N. Baumann, J. C. Turpin, *et al.* (2007). Psychiatric manifestations revealing inborn errors of metabolism in adolescents and adults. *Journal of Inherited Metabolic Disease*, 30(5):631–641.

G. Sharma, D. P. Muller, S. M. O'Riordan, *et al.* (2013). Urinary conjugated a-tocopheronolactone–a biomarker of oxidative stress in children with type 1 diabetes. *Free Radical Biology & Medicine*, 55:54–62.

J. Shen, Y. Bao, H. M. Liu, *et al.* (1996). Mutations in exon 3 of the glycogen debranching enzyme gene are associated with glycogen storage disease type III that is differentially expressed in liver and muscle. *The Journal of Clinical Investigation*, 98(2):352–357.

J. Shendure, H. Ji (2008). Next-generation DNA sequencing. *Nature Biotechnology*, 26(10):1135–1145.

M. Shevell, S. Ashwal, D. Donley, *et al.* (2003). Practice parameter: evaluation of the child with global developmental delay: report of the Quality Standards Subcommittee of the American Academy of Neurology and The Practice Committee of the Child Neurology Society. *Neurology*, 60(3):367–380.

K. Shibasaki (2016). Physiological significance of TRPV2 as a mechanosensor, thermosensor and lipid sensor. *The Journal of Physiological Sciences*, 66(5):359–65.

S. Shim, W. Lee, H. Chung, *et al.* (2011). Amyloid beta-induced FOXRED2 mediates neuronal cell death via inhibition of proteasome activity. *Cellular and Molecular Life Sciences*, 68(12):2115–2127.

E. J. Shin, K. H. Ko, W. K. Kim, *et al.* (2008). Role of glutathione peroxidase in the ontogeny of hippocampal oxidative stress and kainate seizure sensitivity in the genetically epilepsy-prone rats. *Neurochemistry International*, 52(6):1134–1147.

S. Shinnar, J. M. Pellock (2002). Update on the epidemiology and prognosis of pediatric epilepsy. *Journal of Child Neurology*, 17(Suppl 1):S4–17.

R. J. Shmookler Reis, S. Goldstein (1983). Mitochondrial DNA in mortal and immortal human cells. Genome number, integrity, and methylation. *The Journal of Biological Chemistry*, 258(15):9078–9085.

N. A. Singh, P. Westenskow, C. Charlier, *et al.* (2003). KCNQ2 and KCNQ3 potassium channel genes in benign familial neonatal convulsions: expansion of the functional and mutation spectrum. *Brain*, 126(12):2726–2737.

S. M. Sisodiya, W. R. Lin, B. N. Harding, *et al.* (2002). Drug resistance in epilepsy: expression of drug resistance proteins in common causes of refractory epilepsy. *Brain*, 125(1):22–31.

J. L. Sloan, J. J. Johnston, I. Manoli, *et al.* (2011). Exome sequencing identifies ACSF3 as the cause of Combined Malonic and Methylmalonic Aciduria. *Nature Genetics*, 43(9):883–886.

P. Smits, A. D. Bolton, V. Funari, *et al.* (2010). Lethal skeletal dysplasia in mice and humans lacking the golgin GMAP-210. *The New England Journal of Medicine*, 362(3):206–216.

M. Solle, J. Labasi, D. G. Perregaux, *et al.* (2001). Altered cytokine production in mice lacking P2X(7) receptors. *The Journal of Biological Chemistry*, 276(1):125–132.

M. Sommen, I. Schrauwen, G. Vandeweyer, *et al.* (2016). DNA diagnostics of hereditary hearing loss: a targeted resequencing approach combined with a mutation classification system. *Human Mutation*, 37(8):812–819.

G. Souillet, N. Guffon, I. Maire, *et al.* (2003). Outcome of 27 patients with Hurler's syndrome transplanted from either related or unrelated haematopoietic stem cell sources. *Bone Marrow Transplantation*, 31(12):1105–1117.

R. Spector (1978). Vitamin B6 Transport in the Central Nervous System: In Vitro Studies. *Journal of Neurochemistry*, 30(4):889–897.

B. Sperlágh, E. S. Vizi, K. Wirkner, *et al.* (2006). P2X7 receptors in the nervous system. *Progress in Neurobiology*, 78(6):327–346.

E. R. Stadtman (2001). Protein oxidation in aging and age-related diseases. *Annals of the New York Academy of Sciences*, 928:22–38.

E. Stamoula, T. Vavilis, E. Aggelidou, *et al.* (2015). Low dose administration of glutamate triggers a non-apoptotic, autophagic response in PC12 cells. *Cellular Physiology and Biochemistry*, 37(5):1750–1758.

M. Stanulović, V. Jeremić, V. Leskovac, *et al.* (1976). New pathway of conversion of pyridoxal to 4-pyridoxic acid. *Enzyme*, 21(4):357–369.

M. E. Steenweg, D. Ghezzi, T. Haack, *et al.* (2012). Leukoencephalopathy with thalamus and brainstem involvement and high lactate 'LTBL' caused by EARS2 mutations. *Brain*, 135(5):1387–1394.

L. D. Stein (2010). The case for cloud computing in genome informatics. *Genome Biology*, 11(5):207.

S. J. Steinberg, G. V. Raymond, N. E. Braverman, *et al.* (1993). Peroxisome Biogenesis Disorders, Zellweger Syndrome Spectrum. In R. A. Pagon, M. P. Adam, H. H. Ardinger, *et al.*, editors, *GeneReviews(®)*. University of Washington, Seattle, Seattle (WA).

W. Stühmer, F. Conti, H. Suzuki, *et al.* (1989). Structural parts involved in activation and inactivation of the sodium channel. *Nature*, 339(6226):597–603.

S. Stockler, B. Plecko, S. M. Gospe, *et al.* (2011). Pyridoxine dependent epilepsy and antiquitin deficiency: clinical and molecular characteristics and recommendations for diagnosis, treatment and follow-up. *Molecular Genetics and Metabolism*, 104(1-2):48–60.

H. Stranneheim, M. Engvall, K. Naess, *et al.* (2014). Rapid pulsed whole genome sequencing for comprehensive acute diagnostics of inborn errors of metabolism. *BMC Genomics*, 15:1090.

E. A. Struys, C. Jakobs (2010). Metabolism of lysine in alpha-aminoadipic semialdehyde dehydrogenase-deficient fibroblasts: evidence for an alternative pathway of pipecolic acid formation. *FEBS Letters*, 584(1):181–186.

M. Suchi, H. Mizuno, Y. Kawai, *et al.* (1997). Molecular cloning of the human UMP synthase gene and characterization of point mutations in two hereditary orotic aciduria families. *American Journal of Human Genetics*, 60(3):525–539.

A. Sudarsanam, H. Singh, B. Wilcken, *et al.* (2014). Cirrhosis associated with pyridoxal 5'-phosphate treatment of pyridoxamine 5'-phosphate oxidase deficiency. *JIMD Reports*, 17:67–70.

Y. Sun, C. A. Ruivenkamp, M. J. Hoffer, *et al.* (2015). Next-Generation Diagnostics: Gene Panel, Exome, or Whole Genome? *Human Mutation*, 36(6):648–655.

Z. Sun, X. Tang, F. Lin, *et al.* (2011). The WD40 repeat protein WDR26 binds G-beta-gamma and promotes G-beta-gamma-dependent signal transduction and leukocyte migration. *The Journal of Biological Chemistry*, 286(51):43902–43912.

R. Surtees, P. Mills, P. Clayton (2006). Inborn errors affecting vitamin B $_6$ metabolism. *Future Neurology*, 1(5):615–620.

D. R. Sutherland, N. Kuek, J. Davidson, *et al.* (2007). Diagnosing PNH with FLAER and multiparameter flow cytometry. *Cytometry Part B, Clinical Cytometry*, 72(3):167–177.

Y. Suzuki, T. Kita, T. Mano, *et al.* (1996). Outcome of initial treatment with high-dose vitamin B6, valproate sodium or clonazepam in West syndrome. *No To Hattatsu*, 28(5):398–402.

R. Szklarczyk, B. F. J. Wanschers, S. B. Nabuurs, *et al.* (2011). NDUFB7 and NDUFA8 are located at the intermembrane surface of complex I. *FEBS Letters*, 585(5):737–743.

K. Takahara, U. Schwarze, Y. Imamura, *et al.* (2002). Order of intron removal influences multiple splice outcomes, including a two-exon skip, in a COL5A1 acceptor-site mutation that results in abnormal pro-a1(V) N-propeptides and Ehlers-Danlos Syndrome Type I. *American Journal of Human Genetics*, 71(3):451–465.

Y. Takuma, T. Seki (1996). Combination therapy of infantile spasms with high-dose pyridoxal phosphate and low-dose corticotropin. *Journal of Child Neurology*, 11(1):35–40.

B. Talim, A. Pyle, H. Griffin, *et al.* (2013). Multisystem fatal infantile disease caused by a novel homozygous EARS2 mutation. *Brain*, 136(2):e228–e228.

L. Tang, R. J. Reiter, Z. R. Li, *et al.* (1998). Melatonin reduces the increase in 8-hydroxy-deoxyguanosine levels in the brain and liver of kainic acid-treated rats. *Molecular and Cellular Biochemistry*, 178(1-2):299–303.

J. J. Tanner (2008). Structural biology of proline catabolism. *Amino Acids*, 35(4):719–730.

V. Tanphaichitr, D. W. Horne, H. P. Broquist (1971). Lysine, a precursor of carnitine in the rat. *Journal of Biological Chemistry*, 246(20):6364–6366.

R. Tapia, H. Pasantes (1971). Relationships between pyridoxal phosphate availability, activity of vitamin B6-dependent enzymes and convulsions. *Brain Research*, 29(1):111–122.

M. Tarailo-Graovac, C. Shyr, C. J. Ross, *et al.* (2016). Exome sequencing and the management of neurometabolic disorders. *New England Journal of Medicine*, 374(23):2246–55.

M. Tarailo-Graovac, G. Sinclair, S. Stockler-Ipsiroglu, *et al.* (2015). The genotypic and phenotypic spectrum of PIGA deficiency. *Orphanet Journal of Rare Diseases*, 10:23.

Y. Tashima, R. Taguchi, C. Murata, *et al.* (2006). PGAP2 is essential for correct processing and stable expression of GPI-anchored proteins. *Molecular Biology of the Cell*, 17(3):1410–1420.

B. Tavazzi, G. Lazzarino, P. Leone, *et al.* (2005). Simultaneous high performance liquid chromatographic separation of purines, pyrimidines, N-acetylated amino acids, and dicarboxylic acids for the chemical diagnosis of inborn errors of metabolism. *Clinical Biochemistry*, 38(11):997–1008.

S. Tejada, A. Sureda, C. Roca, *et al.* (2007). Antioxidant response and oxidative damage in brain cortex after high dose of pilocarpine. *Brain Research Bulletin*, 71(4):372–375.

D. Tekin, D. Yan, G. Bademci, *et al.* (2016). A next-generation sequencing gene panel (MiamiOtoGenes) for comprehensive analysis of deafness genes. *Hearing Research*, 333:179–184.

L. Terzuoli, R. Leoncini, R. Pagani, *et al.* (1998). Some chemical properties and biological role of thiazolidine compounds. *Life Sciences*, 63(14):1251–1267.

J. B. Thoden, T. M. Wohlers, J. L. Fridovich-Keil, *et al.* (2001). Molecular basis for severe epimerase deficiency galactosemia. X-ray structure of the human V94m-substituted UDP-galactose 4-epimerase. *The Journal of Biological Chemistry*, 276(23):20617–20623.

R. H. Thomas, L. M. Zhang, G. L. Carvill, *et al.* (2015). CHD2 myoclonic encephalopathy is frequently associated with self-induced seizures. *Neurology*, 84(9):951–958.

M. D. Thompson, A. Killoran, M. E. Percy, *et al.* (2006). Hyperphosphatasia with neurologic deficit: a pyridoxine-responsive seizure disorder? *Pediatric Neurology*, 34(4):303–307.

D. J. Timson (2006). The structural and molecular biology of type III galactosemia. *IUBMB Life*, 58(2):83–89.

M. D. Toney (2005). Reaction specificity in pyridoxal phosphate enzymes. *Archives of Biochemistry and Biophysics*, 433(1):279–287.

Y. Toribe (2001). High-dose vitamin B(6) treatment in West syndrome. *Brain & Development*, 23(7):654–657.

D. M. Treiman (2001). GABAergic mechanisms in epilepsy. *Epilepsia*, 42(3):8–12.

D. J. Trezise, N. J. Bell, B. S. Khakh, *et al.* (1994). P2 purinoceptor antagonist properties of pyridoxal-5-phosphate. *European Journal of Pharmacology*, 259(3):295–300.

N. Trump, A. McTague, H. Brittain, *et al.* (2016). Improving diagnosis and broadening the phenotypes in early-onset seizure and severe developmental delay disorders through gene panel analysis. *Journal of Medical Genetics*, 53(5):310–307.

T. Tsuji, A. Okumura, H. Ozawa, *et al.* (2007). Current treatment of West syndrome in Japan. *Journal of Child Neurology*, 22(5):560–564.

H. A. L. Tuppen, J. Fehmi, B. Czermin, *et al.* (2010a). Long-term survival of neonatal mitochondrial complex III deficiency associated with a novel BCS1L gene mutation. *Molecular Genetics and Metabolism*, 100(4):345–348.

H. A. L. Tuppen, V. E. Hogan, L. He, *et al.* (2010b). The p.M292T NDUFS2 mutation causes complex I-deficient Leigh syndrome in multiple families. *Brain*, 133(10):2952–2963.

K. Tuschl, P. T. Clayton, S. M. Gospe, *et al.* (2012). Syndrome of hepatic cirrhosis, dystonia, polycythemia, and hypermanganesemia caused by mutations in SLC30A10, a manganese transporter in man. *American Journal of Human Genetics*, 90(3):457–466.

A. Urayama (2013). Toward the successful delivery of lysosomal enzymes across the blood–brain barrier. *Clinical and Experimental Neuroimmunology*, 4(2):228–238.

P. Urban, I. Chirat, F. Lederer (1988). Rat kidney L-2-hydroxyacid oxidase. Structural and mechanistic comparison with flavocytochrome b2 from baker's yeast. *Biochemistry*, 27(19):7365–7371.

L. van Berge, E. M. Hamilton, T. Linnankivi, *et al.* (2014). Leukoencephalopathy with brainstem and spinal cord involvement and lactate elevation: clinical and genetic characterization and target for therapy. *Brain*, 137(4):1019–1029.

M. van der Ham, M. Albersen, T. J. de Koning, *et al.* (2012). Quantification of vitamin B6 vitamers in human cerebrospinal fluid by ultra performance liquid chromatography-tandem mass spectrometry. *Analytica Chimica Acta*, 712:108–114.

H. H. van Gelderen, H. L. Teijema (1973). Hyperlysinaemia. Harmless inborn error of metabolism? *Archives of Disease in Childhood*, 48(11):892–895.

C. D. van Karnebeek, W. S. Sly, C. J. Ross, *et al.* (2014). Mitochondrial carbonic anhydrase VA deficiency resulting from CA5A alterations presents with hyperammonemia in early childhood. *American Journal of Human Genetics*, 94(3):453–461.

C. D. M. van Karnebeek, H. Hartmann, S. Jaggumantri, *et al.* (2012). Lysine restricted diet for pyridoxine-dependent epilepsy: first evidence and future trials. *Molecular Genetics and Metabolism*, 107(3):335–344.

C. D. M. van Karnebeek, S. Stockler (2012). Treatable inborn errors of metabolism causing intellectual disability: a systematic literature review. *Molecular Genetics and Metabolism*, 105(3):368–381.

C. D. M. van Karnebeek, S. A. Tiebout, J. Niermeijer, *et al.* (2016). Pyridoxine-dependent epilepsy: an expanding clinical spectrum. *Pediatric Neurology*, 59:6–12.

A. B. P. van Kuilenburg, D. Dobritzsch, J. Meijer, *et al.* (2010). Dihydropyrimidinase deficiency: Phenotype, genotype and structural consequences in 17 patients. *Biochimica et Biophysica Acta*, 1802(7–8):639–648.

A. O. Varoglu, A. Yildirim, R. Aygul, *et al.* (2010). Effects of valproate, carbamazepine, and levetiracetam on the antioxidant and oxidant systems in epileptic patients and their clinical importance. *Clinical Neuropharmacology*, 33(3):155–157.

D. M. Veltman, G. Auciello, H. J. Spence, *et al.* (2011). Functional analysis of Dictyostelium IBARa reveals a conserved role of the I-BAR domain in endocytosis. *The Biochemical Journal*, 436(1):45–52.

C. Verity, A. M. Winstone, L. Stellitano, *et al.* (2010). The epidemiology of progressive intellectual and neurological deterioration in childhood. *Archives of Disease in Childhood*, 95(5):361–364.

J. J. Vermeersch, S. Christmann-Franck, L. V. Karabashyan, *et al.* (2004). Pyridoxal 5'-phosphate inactivates DNA topoisomerase IB by modifying the lysine general acid. *Nucleic Acids Research*, 32(18):5649–5657.

W. F. Visser, N. M. Verhoeven-Duif, T. J. de Koning (2012). Identification of a human trans-3-hydroxy-L-proline dehydratase, the first characterized member of a novel family of proline racemase-like enzymes. *The Journal of Biological Chemistry*, 287(26):21654–21662.

L. E. L. M. Vissers, C. Gilissen, J. A. Veltman (2016). Genetic studies in intellectual disability and related disorders. *Nature Reviews Genetics*, 17(1):9–18.

G. Vistoli, D. De Maddis, V. Straniero, *et al.* (2013). Exploring the space of histidine containing dipeptides in search of novel efficient RCS sequestering agents. *European Journal of Medicinal Chemistry*, 66:153–160.

J. Vockley, H. C. Andersson, K. M. Antshel, *et al.* (2014). Phenylalanine hydroxylase deficiency: diagnosis and management guideline. *Genetics in Medicine*, 16(2):188–200.

B. Vona, I. Nanda, C. Neuner, *et al.* (2013). Confirmation of GRHL2 as the gene for the DFNA28 locus. *American Journal of Medical Genetics Part A*, 161A(8):2060–2065.

H. Wada, E. E. Snell (1961). The enzymatic oxidation of pyridoxine and pyridoxamine phosphates. *The Journal of Biological Chemistry*, 236:2089–2095.

E. N. Wainwright, J. S. Jorgensen, Y. Kim, *et al.* (2013). SOX9 regulates microRNA miR-202-5p/3p expression during mouse testis differentiation. *Biology of Reproduction*, 89(2):34.

J. H. Walter, R. E. P. Roberts, G. T. N. Besley, *et al.* (1999). Generalised uridine diphosphate galactose-4-epimerase deficiency. *Archives of Disease in Childhood*, 80(4):374–376.

L. C. Walters-Sen, S. Hashimoto, D. L. Thrush, *et al.* (2015). Variability in pathogenicity prediction programs: impact on clinical diagnostics. *Molecular Genetics & Genomic Medicine*, 3(2):99–110.

H. Wang, M. Kuo, M. Chou, *et al.* (2005). Pyridoxal phosphate is better than pyridoxine for controlling idiopathic intractable epilepsy. *Archives of Disease in Childhood*, 90(5):512–515.

X. Wang, K. Dakshinamurti, S. Musat, *et al.* (1999). Pyridoxal 5'-phosphate is an ATP-receptor antagonist in freshly isolated rat cardiomyocytes. *Journal of Molecular and Cellular Cardiology*, 31(5):1063–1072.

A. Watanabe, T. Yoshimura, B. Mikami, *et al.* (2002). Reaction mechanism of alanine racemase from Bacillus stearothermophilus: x-ray crystallographic studies of the enzyme bound with N-(5'-phosphopyridoxyl)alanine. *The Journal of Biological Chemistry*, 277(21):19166–19172.

T. K. Watanabe, A. Saito, M. Suzuki, *et al.* (1998). cDNA cloning and characterization of a human proteasomal modulator subunit, p27 (PSMD9). *Genomics*, 50(2):241–250.

S. Weckhuysen, V. Ivanovic, R. Hendrickx, *et al.* (2013). Extending the KCNQ2 encephalopathy spectrum: clinical and neuroimaging findings in 17 patients. *Neurology*, 81(19):1697–1703.

S. Weckhuysen, S. Mandelstam, A. Suls, *et al.* (2012). KCNQ2 encephalopathy: emerging phenotype of a neonatal epileptic encephalopathy. *Annals of Neurology*, 71(1):15–25.

Y. Wedatilake, R. M. Brown, R. McFarland, *et al.* (2013). SURF1 deficiency: a multi-centre natural history study. *Orphanet Journal of Rare Diseases*, 8:96.

N. J. Weinreb, J. Charrow, H. C. Andersson, *et al.* (2002). Effectiveness of enzyme replacement therapy in 1028 patients with type 1 Gaucher disease after 2 to 5 years of treatment: a report from the Gaucher Registry. *The American Journal of Medicine*, 113(2):112–119.

M. P. Whyte, J. D. Mahuren, K. N. Fedde, *et al.* (1988). Perinatal hypophosphatasia: tissue levels of vitamin B6 are unremarkable despite markedly increased circulating concentrations of pyridoxal-5'-phosphate. Evidence for an ectoenzyme role for tissue-nonspecific alkaline phosphatase. *The Journal of Clinical Investigation*, 81(4):1234–1239.

M. P. Whyte, J. D. Mahuren, L. A. Vrabel, *et al.* (1985). Markedly increased circulating pyridoxal-5'-phosphate levels in hypophosphatasia. Alkaline phosphatase acts in vitamin B6 metabolism. *The Journal of Clinical Investigation*, 76(2):752–756.

B. Widner, F. Leblhuber, J. Walli, *et al.* (2000). Tryptophan degradation and immune activation in Alzheimer's disease. *Journal of Neural Transmission*, 107(3):343–353.

S. Wiemann, A. Kolb-Kokocinski, A. Poustka (2005). Alternative pre-mRNA processing regulates cell-type specific expression of the IL4l1 and NUP62 genes. *BMC Biology*, 3:16.

A. Wieraszko, T. N. Seyfried (1989). Increased amount of extracellular ATP in stimulated hippocampal slices of seizure prone mice. *Neuroscience Letters*, 106(3):287–293.

P. M. Wilson, R. H. Fryer, Y. Fang, *et al.* (2010). Astn2, a novel member of the astrotactin gene family, regulates the trafficking of ASTN1 during glial-guided neuronal migration. *The Journal of Neuroscience*, 30(25):8529–8540.

D. S. Wishart, D. Tzur, C. Knox, *et al.* (2007). HMDB: the Human Metabolome Database. *Nucleic Acids Research*, 35(Database issue):D521–526.

T. M. Wohlers, N. C. Christacos, M. T. Harreman, *et al.* (1999). Identification and characterization of a mutation, in the human UDP-galactose-4-epimerase gene, associated with generalized epimerase-deficiency galactosemia. *American Journal of Human Genetics*, 64(2):462–470.

T. M. Wohlers, J. L. Fridovich-Keil (2000). Studies of the V94M-substituted human UDPgalactose-4-epimerase enzyme associated with generalized epimerase-deficiency galactosaemia. *Journal of Inherited Metabolic Disease*, 23(7):713–729.

N. I. Wolf, T. Bast, R. Surtees (2005). Epilepsy in inborn errors of metabolism. *Epileptic Disorders*, 7(2):67–81.

N. I. Wolf, J. A. M. Smeitink (2002). Mitochondrial disorders: a proposal for consensus diagnostic criteria in infants and children. *Neurology*, 59(9):1402–1405.

J. Wong-Ekkabut, Z. Xu, W. Triampo, *et al.* (2007). Effect of lipid peroxidation on the properties of lipid bilayers: a molecular dynamics study. *Biophysical Journal*, 93(12):4225–4236.

J. H. Wood, T. A. Hare, B. S. Glaeser, *et al.* (1979). Low cerebrospinal fluid gamma-aminobutyric acid content in seizure patients. *Neurology*, 29(9):1203–1208.

J. M. Wood (1987). Membrane association of proline dehydrogenase in Escherichia coli is redox dependent. *Proceedings of the National Academy of Sciences of the United States of America*, 84(2):373–377.

S. B. Wortmann, A. P. de Brouwer, R. A. Wevers, *et al.* (1993). MEGDEL Syndrome. In R. A. Pagon, M. P. Adam, H. H. Ardinger, *et al.*, editors, *GeneReviews(®)*. University of Washington, Seattle, Seattle (WA).

S. B. Wortmann, F. M. Vaz, T. Gardeitchik, *et al.* (2012). Mutations in the phospholipid remodeling gene SERAC1 impair mitochondrial function and intracellular cholesterol trafficking and cause dystonia and deafness. *Nature Genetics*, 44(7):797–802.

G. Wu (2013). Functional amino acids in nutrition and health. *Amino Acids*, 45(3):407–411.

R. Wynn (2011). Stem cell transplantation in inherited metabolic disorders. *ASH Education Program Book*, 2011(1):285–291.

C. Xie, M. T. Tammi (2009). CNV-seq, a new method to detect copy number variation using high-throughput sequencing. *BMC Bioinformatics*, 10:80.

K. Yagi, K. Okamura, M. Naoi, *et al.* (1969). Reaction of D-amino-acid oxidase with D-lysine. *Journal of Biochemistry*, 66(5):581–589.

E. S. Yang, V. Schirch (2000). Tight binding of pyridoxal 5'-phosphate to recombinant Escherichia coli pyridoxine 5'-phosphate oxidase. *Archives of Biochemistry and Biophysics*, 377(1):109–114.

Y. Yang, D. M. Muzny, F. Xia, *et al.* (2014). Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA*, 312(18):1870–1879.

S. A. Yasin, A. M. Ali, M. Tata, *et al.* (2013). mTOR-dependent abnormalities in autophagy characterize human malformations of cortical development: evidence from focal cortical dysplasia and tuberous sclerosis. *Acta Neuropathologica*, 126(2):207–218.

K. Ye, M. H. Schulz, Q. Long, *et al.* (2009). Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics*, 25(21):2865–2871.

U. Yiş, G. Uyanik, D. M. Rosendahl, *et al.* (2014). Clinical, radiological, and genetic survey of patients with muscle-eye-brain disease caused by mutations in POMGNT1. *Pediatric Neurology*, 50(5):491–497.

C. Yiannikas, D. Cordato (1996). Familial hyperlysinemia in a patient presenting with progressive spastic paraparesis. *Neurology*, 47(3):846.

P. Ylä-Anttila, H. Vihinen, E. Jokitalo, *et al.* (2009). Monitoring autophagy by electron microscopy in Mammalian cells. *Methods in Enzymology*, 452:143–164.

H. Yokogoshi, R. J. Wurtman (1986). Meal composition and plasma amino acid ratios: effect of various proteins or carbohydrates, and of various protein concentrations. *Metabolism: Clinical and Experimental*, 35(9):837–842.

A. Yoshida, K. Kobayashi, H. Manya, *et al.* (2001). Muscular dystrophy and neuronal migration disorder caused by mutations in a glycosyltransferase, POMGnT1. *Developmental Cell*, 1(5):717–724.

I. Yoshida, Y. Sakaguchi, M. Nakano, *et al.* (1985). Pyridoxal phosphate-induced liver injury in a patient with homocystinuria. *Journal of Inherited Metabolic Disease*, 8(2):91.

R. Yoshida (1993). High dose pyridoxal phosphate therapy for West syndrome: Clinical study of 59 patients with special attention to side effects. *Journal of Tokyo Women's Medical College*, 63(10):1156–1184.

I. B. Youssef-Turki, I. Kraoua, S. Smirani, *et al.* (2011). Epilepsy aspects and EEG patterns in neuro-metabolic diseases. *Journal of Behavioral and Brain Science*, 1(2):69–74.

W. W. Yue, D. S. Froese, P. E. Brennan (2014). The role of protein structural analysis in the next generation sequencing era. *Topics in Current Chemistry*, 336:67–98.

B. Zhang, S. Kirov, J. Snoddy (2005). WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Research*, 33(Web Server issue):W741–W748.

G. Zhao, M. E. Winkler (1995). Kinetic limitation and cellular amount of pyridoxine (pyridoxamine) 5'-phosphate oxidase of Escherichia coli K-12. *Journal of Bacteriology*, 177(4):883–891.

J. Zhao, H. Chen, P. Ni, *et al.* (2011). Simultaneous determination of urinary tryptophan, tryptophan-related metabolites and creatinine by high performance liquid chromatography with ultraviolet and fluorimetric detection. *Journal of Chromatography B, Analytical Technologies in the Biomedical and Life Sciences*, 879(26):2720–2725.

Y. Zhao, F. Zhao, L. Zong, *et al.* (2013). Exome sequencing and linkage analysis identified tenascin-C (TNC) as a novel causative gene in nonsyndromic hearing loss. *PloS One*, 8(7):e69549.

Y. Zhu, Y. Wang, C. Xia, *et al.* (2004). WDR26: a novel G-beta-like protein, suppresses MAPK signaling pathway. *Journal of Cellular Biochemistry*, 93(3):579–587.

M. Zollino, D. Orteschi, M. Murdolo, *et al.* (2012). Mutations in KANSL1 cause the 17q21.31 microdeletion syndrome phenotype. *Nature Genetics*, 44(6):636–638.

M. Zoratti, I. Szabò (1995). The mitochondrial permeability transition. *Biochimica Et Biophysica Acta*, 1241(2):139–176.