



Universiteit
Leiden
The Netherlands

Developing metabolomics for a systems biology approach to understand Parkinson's disease

Willacey, C.C.W.

Citation

Willacey, C. C. W. (2021, September 8). *Developing metabolomics for a systems biology approach to understand Parkinson's disease*. Retrieved from <https://hdl.handle.net/1887/3209244>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3209244>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <https://hdl.handle.net/1887/3209244> holds various files of this Leiden University dissertation.

Author: Wallacey, C.C.W.

Title: Developing metabolomics for a systems biology approach to understand Parkinson's disease

Issue Date: 2021-09-08

**Developing metabolomics for
a systems biology approach to
understand Parkinson's
disease**

The publication of this thesis was financially supported by:

Leiden Academic Centre for Drug Research (LACDR)

Leiden University Libraries



Cover design: Alisa L. Willacey

Thesis layout: Cornelius C. W. Willacey

Printing: Ede Printservice

© Copyright, Cornelius C. W. Willacey, 2021

ISBN: 978-90-831713-8-8

All rights are reserved. No part of this book may be reproduced in any form or by any means without permission of the author.

Developing metabolomics for a systems biology approach to understand Parkinson's disease

Proefschrift

ter verkrijging van

de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op woensdag 8 september 2021

klokke 10:00 uur

door

Cornelius Carmichael William Willacey

Geboren te Nottingham, Verenigd Koninkrijk

in 1992

PROMOTOR

Prof. dr. T. Hankemeier

CO-PROMOTORS

Dr. A. C. Harms

Dr. R. M. T. Fleming

PROMOTIECOMMISSIE

Prof. dr. H. Irth (Chair)

Leiden University, the Netherlands

Prof. dr. J. A. Bouwstra (Secretary)

Leiden University, the Netherlands

Prof. Dr Christine Klein

University of Lübeck, Germany

Prof. dr. M. van der Stelt

Leiden University, the Netherlands

Prof. dr. C. Knibbe

Leiden University, the Netherlands

Prof. dr. B.M. Bakker

University of Groningen, the Netherlands

The research described in this thesis was performed at the division of Systems Biomedicine and Pharmacology of the Leiden Academic Centre for Drug Research (LACDR), Leiden University (Leiden, The Netherlands). The research was financially supported by the SysMedPD project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 668738.

Contents

Chapter 1	Introduction.....	1
	Metabolomics method development	
Chapter 2	LC-MS/MS analysis of the central energy and carbon metabolites in biological samples following derivatization by dimethylaminophenacyl bromide <i>Journal of Chromatography A (2019)</i>	29
Chapter 3	Metabolic profiling of material-limited cell samples by dimethylaminophenacyl bromide derivatization with UPLC-MS/MS analysis <i>Microchemical Journal (2020)</i>	73
	Biological application	
Chapter 4	A quantitative atlas of metabolites across regions of the rat brain <i>Manuscript submitted</i>	113
Chapter 5	Mechanistic model-driven exometabolomic characterisation of human dopaminergic neuronal metabolism <i>Manuscript in preparation</i>	167
	Conclusions	
Chapter 6	Concluding discussion and future prospects.....	263
Addendum	Summary.....	284
	Nederlandse samenvatting.....	288
	Curriculum vitae.....	293
	List of publications.....	295
	Acknowledgments.....	296

Chapter 1

Introduction

Systems biology in human health

Healthcare is constantly reaching new heights in the treatment of acute and chronic illnesses. These developments come in several forms; they can be economical¹, political², technological³ or scientific. With this ongoing progression, human longevity is lengthened with each generation that passes⁴. Mortality rates have decreased due to improved healthcare, understanding of diseases and therapeutic treatment options. For example, causal factors have been associated with diseases such as smoking with lung cancer^{5,6}, high fat diet with cardiovascular disease⁷ and alcohol with liver disease⁸. Moreover, landmark therapeutic treatments that vastly reduce mortality can be seen in the discovery of antibiotics, introduction of vaccines, and insulin treatment for diabetes. However, the reduced mortality rate comes at a cost with an increased morbidity rate. It has been estimated that the population of people in Europe aged 65 years and over will increase from 90.5 million in 2019 to 129.8 million in 2050 (shown in Figure 1.1)⁹. Unfortunately, the human body is limited due to time-dependent physiological changes, i.e. ageing. These changes occur during the maturation processes, both genetically and metabolically¹⁰. Many of the age-related illnesses are neurodegenerative diseases that result in the loss of function or homeostasis within the brain. Neurological disorders have been identified by the World Health Organization (WHO)¹¹ as a public health challenge and can manifest over a broad demographic in the form of conditions such as depression, schizophrenia, addiction and epilepsy. However, there are specific conditions that are associated with aging; these are neurodegenerative diseases including Parkinson's, Huntington's and Alzheimer's disease. As mentioned above, these diseases are becoming more prevalent as human longevity increases which in turn has created a surge in the demand for improved disease interpretation, comprehensive diagnostic procedures and available treatment options¹¹. Systems biology is a useful tool that has the ability to produce an encyclopedic evaluation of neurological disorders and diseases.

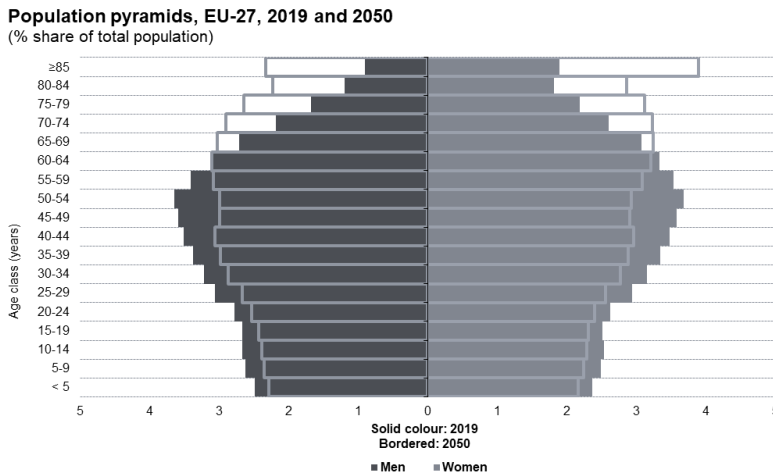


Figure 1.1. presents the aging demographic of the population in Europe across 27 nations between the year 2019 (shown in solid colours) and 2050 (shown in bordered colours). We see men in dark grey on the left and women in light grey on the right. This figure has been adapted from Eurostat⁹. [accessed on 24/11/2020]

Systems biology is a mathematical and computational research field that involves the construction of models to analyse and evaluate biological systems by integration of experimental and computational data^{12,13}. In recent years, there has been an incline in the application of systems biology due to the technological growth in the “omics” research fields¹⁴ and modelling approaches. Systems biology has the ability to detail biological systems with a breadth and depth that would otherwise be challenging, if not impossible, for the human mind¹⁵. It achieves this by not only investigating a single gene, protein or metabolite, but by assessing and evaluating the system holistically¹⁶. Models are usually constructed using data from genomics, transcriptomics, proteomics and, more recently, metabolomics. One modelling approach that is useful for improving the understanding of specific illnesses is genome-scale constraint-based modelling which uses genetic information to predict human metabolism (Step E, Figure 1.2). Constraint-based modelling is an approach that requires absolute quantitative metabolomics data to identify the physicochemical and biochemical bounds that exist within a biological system, identifying the steady-state metabolic fluxes¹⁷. After integration of omics data, the model can be used to predict metabolic exchange fluxes which provides insight into

the biological function¹⁸. Several models have been created using this principle, with the main modelling approach demonstrated by *Recon3D*¹⁹ which contains organ-specific data from several scientific disciplines. After the network is constructed, disease-specific models can be established. Examples of this have modelled human metabolism and gut microbiota in the virtual metabolic human (available online at VMH.life). After construction of these disease-specific models, they can be utilised for biomarker discovery, therapeutic treatment strategy identification and drug repurposing¹⁹. The general workflow of this thesis has been visually represented in Figure 1.2. We are going to focus on the metabolomics developments and application along with systems biology approaches to improve our understanding of the disease state and potentially identify new therapies of neurological disorders with a specific focus on Parkinson's disease (Step F, Figure 1.2).

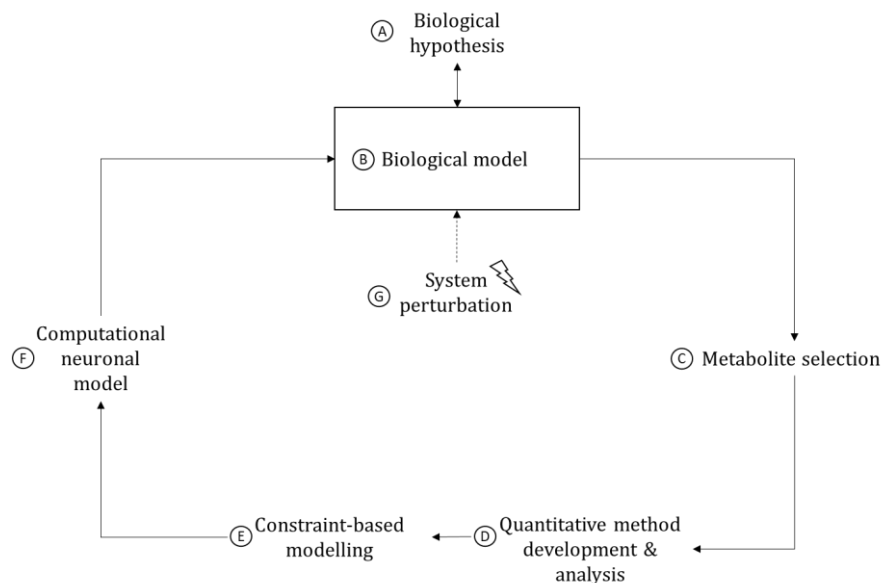


Figure 1.2. A schematic workflow of the systems biology approach with the use of metabolically constraint-based modelling. This figure highlights the workflow that is used throughout this thesis.

Parkinson's disease

Parkinson's disease (PD) is a progressive neurodegenerative condition which leads to a loss of fine motor movements, creating symptoms such as tremor, bradykinesia, postural instability and rigidity. In addition to these symptoms, PD patients also experience non-motor symptoms such as depression, memory loss and sleep disturbances²⁰. Parkinsonism and PD are often confused to be the same. However, Parkinsonism is the group term for neurological conditions which present movement disorders. A range of these conditions manifest with similar symptoms. Clinical determination of PD is only possible by post-mortem examination of the neural tissue, proving PD to be diagnostically challenging to physicians and specialists alike. Currently there are 6.9 million people in the world suffering from PD. This number is set to increase to 14.2 million by 2040, highlighting the importance for further research in the future²¹. There has been extensive research into PD for many years, though this has not translated into drug therapies on the market that are able to cure the condition. Symptomatic treatments are available such as levodopa, dopamine agonists, COMT inhibitors and anticholinergic agents, but these treatments do little to halt the progression of PD. Additionally, drugs such as dopamine agonists (i.e. ropinirole) can have undesirable adverse effects causing impulsive control disorders, such as gambling, binge eating and hypersexuality²².

Parkinson's disease is caused by the loss of dopaminergic neurons within the substantia nigra par compacta (SNpc); this area of dopaminergic neurons is allocated the term A9²³. The cause of PD is still unknown, but all associated causes lead to the depletion of the neurotransmitter dopamine which produces the distinctive motor symptoms that have been described above. Parkinson's disease has also been associated with mitochondrial dysfunction which is said to be present in approximately 10-20% of all PD patients²⁴. Currently, several genes have been linked to PD; PINK1²⁵, Parkin²⁶, LRRK2²⁷, SNCA²⁸, DJ-1²⁹, ATP13A2 and GBA^{30,31}. Here, we hypothesise (Step A, Figure 1.2) that the analysis of PD samples using metabolomics methods which target central carbon and energy metabolism will identify a disease phenotype. With this, we believe that we can identify biomarkers and create disease assays that can distinguish between the different genetic mutations that affect PD. This has opened new possibilities to understand the cause of PD and can identify new potential pathways for therapeutic targeting.

Metabolomics

1

Metabolomics is the study of the molecular phenotype of biological organisms; it is defined as the “the comprehensive study of the metabolome, the repertoire of biochemicals (or small molecules) present in cells, tissues, and body fluids” by Beger *et al.*³². The small molecules are usually those below 1.5 kDa including but not limited to amino acids, nucleic acids, organic acids, small peptides (dipeptides, tripeptides), sugars, fatty acids, hormones, minerals and vitamins. Our understanding of physiological function, disease states and therapeutic target sites has been derived from genomics, transcriptomics and proteomics. Metabolomics provides an alternative yet complementary technique with the other omics approaches. With techniques such as genomics, we have the ability to understand the potential of the biological system, however, metabolomics provides us with the ability to assess the functional status of the system³³.

The diverse physiochemical properties of metabolites within the human metabolome and differing matrices can provide challenges for analytical chemistry³⁴ from an identification and quantitation perspective. Within metabolomics, there are four main techniques used to quantify metabolites; these are Near-infrared spectroscopy (NIRS), ultraviolet–visible spectroscopy (UV-VIS), nuclear magnetic resonance (NMR) and mass spectrometry (MS)³⁵. The most commonly used techniques are MS and NMR. Both techniques are used to identify and quantify metabolites in biological samples, with NMR providing better quantitative results with high reproducibility. Additionally, NMR has the ability to elucidate the structure of metabolites, aiding identification of isomers, and analyse samples that are challenging to ionise in MS or require derivatization³⁶. However, another issue with the human metabolome is that some metabolites exist at very low concentrations – sub-nanoMolar. This is where MS emerges in superiority as it provides sensitivity when measuring low abundant metabolites in a quantitative manner. Moreover, MS is also better than NMR at identifying compounds in complex mixtures.

The metabolome reflects the combination of biological and environmental factors. The overall governance of the metabolome is directly influenced by the genome and, in turn, the transcriptome and proteome. The genome is influenced by factors such as genetic mutations, age, sex and ethnicity. Several diseases are associated with genetic risk factors across a broad range of illnesses, such as cancer, cardiovascular

disease and neurological disorders. All of these diseases manifest with an altered metabolome that are being studied extensively within the metabolomics field³⁷⁻³⁹. The second major influence on the metabolome is the exposome, i.e. the impact of environmental factors on a biological system^{40,41}. Common factors include: diet, gut microbiota, employment, drugs, exercise, geographical location, pollutants, cosmetics, smoking and alcohol consumption. The study of the exposome is complex and challenging due to the multifactorial effects on the metabolome. Researchers are faced with the underlying genetics plus these environmental factors experienced by the host over a lifetime. To truly understand the human metabolome, the interaction between both genetic and environmental factors must be considered. An example of these environmental effects includes the microbiota in the gut that has been associated with motor deficits and neuroinflammation in models relevant to PD⁴². Pollutants and toxins can also influence the human disease state as seen in PD that can be induced by exposure to the naturally occurring pesticide rotenone⁴³ or the illicit drug by-product MPTP⁴⁴. The impact of pesticide rotenone on the human metabolome⁴⁵ will also be discussed further in **chapter 2** where we use it as a chemical perturbation to mimic mitochondrial dysfunction (Step G, Figure 1.2). In this thesis, we quantitatively capture the broad metabolome, which requires the appropriate selection of metabolites that holistically capture the genome-exposome interaction.

Metabolite selection

The human metabolome is vast in size and new metabolites are being identified each year. As it currently stands, databases such as the Human Metabolome Database (HMDB 4.0) have identified 114,100 metabolites using a combination of measurements, expectations and predictions. Each metabolite has its own biological role, metabolic pathway(s), transport mechanism and physiological concentration⁴⁶. Metabolic pathways contain metabolites in an intricate and dynamic system that is constantly adapting to the physiological demands. Within the human body, metabolites range from core metabolites, such as amino acids, to TCA cycle metabolites that control energy production within the mitochondria, to metabolites that exist only to facilitate the intermediate stages of a pathway. Other metabolites are present in their metabolised form awaiting excretion via the liver and kidney,

such as sulphated dopamine. Despite these differences in role and function, the majority of the metabolites are essential to support and sustain life.

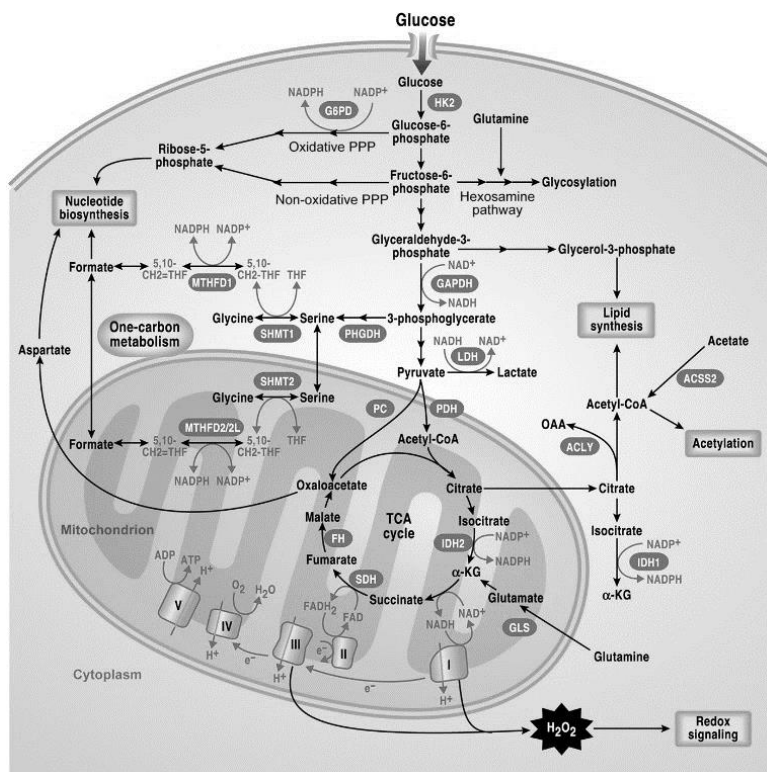


Figure 1.3. Metabolites associated with the energy processes within the cell. Glycolysis, the Krebs cycle and intermediate metabolites as well as amino acids involved in the energy process are detailed. Reprinted DeBerardinis and Chandel (2016)³⁷ © The Authors, some rights reserved; exclusive licensee AAAS. Distributed under a CC BY-NC 4.0 License.

Neurological disorders, as with all diseases, each have their own characteristic metabolic profile which can further vary in relation to an individual. These disease phenotypes respond to the chemical cues created by the metabolite levels in the corresponding pathways and functions. For this reason, the relevant metabolic targets must be specifically chosen for study (Step C, Figure 1.2). Burte *et al.* (2017) identified several metabolites associated with PD that are involved in the energy metabolism and Krebs cycle, which corresponds to PD being associated with

mitochondrial function²⁴; thus, metabolite selection must cover the energy metabolism. One of the main energy processes that occurs within the mitochondria is the electron-transport chain which involves a range of organic acids within the Krebs cycle, such as succinate and malate (Figure 1.3). Another important metabolic pathway is energy generation by glycolysis. This metabolite can enter the Krebs cycle in the form of oxaloacetate produced by pyruvate carboxylase from pyruvate or acetyl-CoA following conversion with pyruvate dehydrogenase. Pyruvate also has a role in anaerobic respiration where it can be converted into lactate for a simple energy supply. Other metabolites associated with the energy metabolism include amino acids, such as aspartic acid, glycine, serine, glutamine and glutamate, and other common metabolites, such as acylcarnitines³⁹ and *N*-acetylated amino acids^{39,47}. These metabolites are all investigated within this thesis, specifically in **chapter 2**, **chapter 3** and **chapter 5**.

Neurological disorders can be profiled using metabolites that are more specific to neurons. For example, neurotransmitters such as dopamine, GABA, serotonin, epinephrine, norepinephrine and glutamate have roles in the maintenance of homeostasis within neurons but also in the communication network across the neuronal cells. The brain is composed of several types of neuronal cells such as neurons and glial cells (oligodendrocytes, astrocytes, ependymal cells and microglia). The glial cells are distributed throughout the brain and are difficult to distinguish by metabolic profile. To date, the number of identified neurotransmitter molecules is over 100. The neurotransmitters are physicochemically diverse and exist in the form of amino acids (glutamate, D-serine and aspartic acid), monoamines, purines, neuropeptides (*N*-acetylaspartic acid) and others. Neurotransmitters such as GABA, glutamate and dopamine are key in the functionality of the substantia nigra and they exist in a complicated balance²³. Deciphering the function of these neurotransmitters can help improve the understanding of PD and other neurodegenerative diseases. In this thesis, key neurochemicals and neurotransmitters are investigated in **chapter 4** and **chapter 5**.

Biological samples in metabolomics

The metabolome can be measured from a range of biological fluids (matrices) that are extracted from human subjects; for example, blood serum and plasma⁴⁸, urine⁴⁵, faeces⁴⁹, sweat⁵⁰, tissue⁴⁸, semen⁵¹ and breast milk⁵². The most commonly used

biological matrix is blood plasma due to the safe and simple extraction procedures that exist. Furthermore, the majority of the metabolites that are excreted from cells and organs are transported through the blood, providing a broad overview of the metabolome. However, not all metabolites are excreted into the blood or pass the blood brain barrier; for example, tissue-specific metabolites and those excreted by the gut microbiota⁵³. Additionally, some metabolites exist in their modified forms, such as sulphated dopamine, or experience degradation, which provide analytical challenges.

In addition to human models, another common approach to study disease is animal modelling. One of the main strengths of animal models is the reduction of exposome influences which allows the researcher to focus on the genetic influence on the metabolome whilst providing a full organism with functional organs⁵⁴. Additionally, animal studies have fewer ethical considerations compared to studies that are designed in humans. Animal models unfortunately suffer from the fact that they do not fully represent the human physiological system and there are clear differences within the metabolome. Despite this, they provide a useful tool for scientific research.

In vitro cell culture work is similar to animal models by which they also minimise the exposome influence. In addition, it requires the least amount of ethical considerations. There are a range of cell lines available for the study of diseases; specific cell lines can be investigated according to the disease of interest, from oncology, endocrinology to neurology. Cell culture approaches allow the researcher to investigate single cell lines or use a co-culture in the attempt to make them more physiologically relevant. Cell cultures can also be 2-dimensional or 3-dimensional, with the latter said to provide a more physiologically realistic environment⁵⁵. However, cell culture has limitations which include existing in an artificial environment (cell culture media), variation in performance between scientists and simple cell lines not sufficiently representing complex organs. One approach to get closer to physiologically relevant cell lines is the use of induced pluripotent stem cell-derived cell lines which carry the genetics of an individual⁵⁶. A summary of the pros and cons of each study sample type is listed in Table 1.1.

In this thesis, we utilise cell culture and human urine in **chapter 2**, cell culture in **chapter 3**, animal models in **chapter 4**, and induced pluripotent stem cell-derived

neurons in **chapter 5** to demonstrate the need to match the most appropriate biological matrix with the hypothesis (Step B, figure 1.2).

Human	Animal	Cell line	iPSC
Examples	Examples	Examples	Examples
• Plasma • Serum • Urine • CSF • Faeces • Breast milk	• Primate • Canine • Rodent	• HEK292 (kidney) • SUIT-2 (pancreas) • PC-12 (adrenal) • HeLa (cervix)	• Dopaminergic neurons • Hepatocytes • Endothelial cells
Strengths	Strengths	Strengths	Strengths
• Multi-organ system • Organism of interest • Verbal evaluation allowed • Realistic exposome environment	• Multi-organ organism • Reflects human physiology closely • Suitable for early-phase research • Controlled exposome	• No ethical considerations • Transfection possible • Inexpensive • Ease of use • Availability	• Patient specific • More physiologically realistic • Co-culture possible
Weaknesses	Weaknesses	Weaknesses	Weaknesses
• More ethical considerations • Exposome influence • Information complexity	• More ethical considerations • Animal welfare • Human genetic difference	• Less physiological relevance • Artificial growth environment	• Artificial growth environment • Cell age is young • Time-consuming • More expertise required • Expensive

Table 1.1. Examples of common study sample type that can be used in metabolomics for the study of diseases. The strengths and weaknesses associated with each sample type are also detailed.

Metabolomics analysis

There are two main approaches used in metabolomics; untargeted and targeted. Untargeted is the global overview of the metabolism without a specific class or pathway of metabolites being identified – chemical unknowns⁵⁷. Using this approach provides a large amount of data that captures as much of the metabolome as possible. After this, the data can be compared to identify patterns in the human metabolism or disease state. However, as it currently stands, we are only able to identify < 2% of peak identified using the untargeted mass MS workflow⁵⁸. Additionally, the metabolites presented by untargeted metabolomics are not absolute quantitative values, thus, they cannot be integrated into constraint-based metabolic models in

systems biology. However, untargeted metabolism is highly desirable in exploratory research to identify new metabolites and pathways associated with illnesses. Within this thesis, we focus on the application of targeted metabolomics using absolute quantitative values.

Targeted metabolomics

Targeted metabolomics is where a specific metabolic pathway or class of metabolites are selected prior to analysis and the methods are optimised around the desired candidates to ensure accurate qualitative and quantitative results. An ideal instrumental setup for targeted quantitative analysis of abundant metabolites, such as amino acids, sugars and organic acids, from a biological matrix that is high in volume would be NMR³⁶. However, as mentioned previously, to truly understand the metabolome, you need to delve deeper using a range of biological matrices (some of which are low in volume/material-limited) with low concentrations of metabolites, thus LC-MS/MS using a triple quadrupole (QqQ) MS becomes the gold standard. One limitation of QqQ MS, is the lack of mass resolution, reducing the specificity and qualitative performance. There is where Quadrupole Time-of-Flight MS (qToF) MS provides a solution for this problem to provide high-resolution accurate mass data and improve the identification of metabolites at the expense of losing sensitivity compared to QqQ MS.

One of the main challenges with QqQ mass spectrometry is distinguishing metabolites with non-unique masses. One way this is addressed is the hyphenation with a separation science such as gas chromatography (GC), liquid chromatography (LC), supercritical fluid chromatography (SPF) or capillary electrophoresis (CE). With the utilisation of a separation science hyphenated to MS, it reduces the risk of mass interference from isomeric compounds, such as amino acids isoleucine and leucine. In this thesis, we focus on the use of LC-MS/MS. The most common system setup for metabolomics methods is LC hyphenated to MS.

A major issue with quantitative workflows using RPLC-MS/MS is matrix effect. This occurs during the electrospray ionisation (ESI), prior to MS detection. Matrix effect (Figure 1.4) is the result of several metabolites, salts or proteins eluting simultaneously, which can either lead to ion suppression (reduced charging of target analyte) or ion enhancement (increased charging of target analyte)⁵⁹. Furthermore,

negative ionisation is more susceptible to ion suppression when compared with positive ionisation. This unpredictable ionisation variable introduces challenges when it comes to accurate quantitation and reproducible analysis of the target analyte as the matrix changes (inter-sample). Moreover, the ion suppression can be so severe that the analyte of interest is suppressed below the detection level. One approach to characterise matrix effect is the use of isotopically labelled internal standard pairs for analytes of interest. For example, alanine-D₃ could be injected simultaneously as an analyte pair with alanine to characterise matrix effect. However, the cost of isotopically labelled metabolites is expensive and not all metabolites have an available isotopically labelled form. Therefore, analytical scientists tend to extend their internal standards to cover more than the analyte pair. Methods have been developed that use 10 internal standards to characterise matrix effect for 70 metabolites, however, this yields poor quantitative quality assurance due to the difference in elution time and physiochemical properties. The utilisation of the internal standard pair also does not reduce ion suppression, therefore, the issue of suppression below the limit of detection (LOD) still exists. Moreover, the issue of detection sensitivity is not addressed when it comes to the quantitation of material-limited samples.

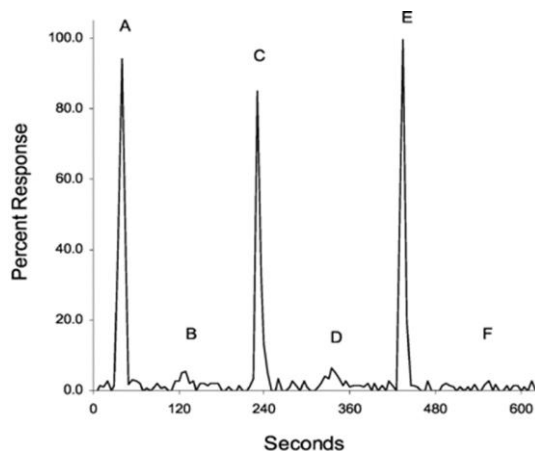


Figure 1.4. Signal response comparisons (m/z 195) for caffeine added to serum extracts prepared by solid-phase extraction, solvent extraction, and protein precipitation. (A, C, and E), 1 mg/L caffeine solution. (B), solid-phase extract with 1 mg/L caffeine added. (D), methylene chloride extract with 1 mg/L caffeine added. (F), serum protein precipitation extract with 1 mg/L caffeine added. Figure extracted from Annesley et al (2003)⁵⁹ and produced following the copyright permission from Oxford University Press Journals for personal thesis non-commercial use.

Material-limited samples

The use of modern techniques to assess the metabolome function in a more realistic biological environment compared to 2D cell culture has been increasing in recent years. Approaches such as 3D cell culture^{55,60}, human cell transplantation into animal models⁶¹ and microdialysate^{62,63} are becoming increasingly common. This has improved our physiological representation of the metabolome and enabled dynamic sampling in a high-throughput manner. However, this has led to a reduction in sample volume and reduced metabolite concentrations. Several approaches exist to increase sensitivity such as sheathless CE-MS⁶⁴ and nanoLC-ESI-MS⁶⁵ but these approaches are limited in their coverage or quantitative profile. To evaluate the metabolome of PD samples and other neurological disorders, sensitive quantitative analysis method are required that can still capture the broad metabolome⁶⁶. In

chapter 3, we use chemical derivatization as an approach to increase sensitivity of material-limited sample whilst maintaining the quantitative coverage.

Chemical derivatization

Above, we have discussed the main analytical methods using LC-MS for quantitative metabolomics and the need for improved quantitation and sensitivity. However, limitations exist that reduce the quantitation and detection. The main issues have been summarised here:

- Chromatography robustness
- Ionisation characteristics
- Metabolite physiochemical properties – suited to different separation sciences
- Metabolite stability
- Matrix effect
- Detection limit

One method that can be used to solve the issues listed above is chemical derivatization. This approach is our method of choice for accurate quantitation throughout this thesis.

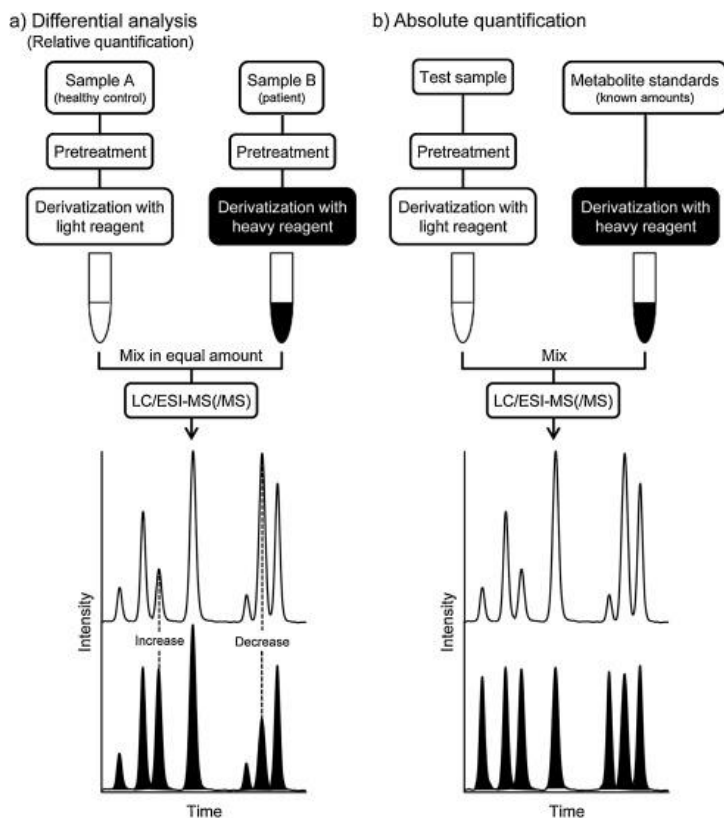


Fig 1.5. The schematic workflow used with ICD derivatization for qualitative differential analysis (a: left workflow) and absolute quantitative (b: right workflow). Figure extracted and modified from Higashi et al. (2016)⁶⁷ and produced following the copyright permission from Elsevier Journals for personal thesis non-commercial use.

Chemical derivatization is the process of adding a chemical group to the analyte to enhance the separation and detection within LC-MS. By changing the physiochemical properties of the analyte of interest, characteristics such as volatility, hydrophobicity, detectability, stability and polarity can be adjusted. GC-MS commonly uses chemical derivatization by alkylation and silylation, to improve the volatility and retention for metabolites such as amino acids and organic acids⁶⁸. Additionally, aliphatic metabolites have no UV or fluorescence properties – thus pose challenges with LC-UV⁶⁹. For this reason, derivatization can be used, introducing aromatic groups which have a UV absorbance. Three common reagents used for UV detection are benzoyl chloride⁶⁹, dansyl chloride⁷⁰, phenacyl bromide^{71,72}, and AccQ-tag^{73,74}. Additionally,

metabolites such as catecholamines (i.e. dopamine), have phenol groups which are highly vulnerable to degradation outside of the cell⁷⁵, presenting difficulties in the analytical community. Reagents such as benzoyl chloride have the ability to derivatize these functional groups thus stabilising the metabolite for analysis^{76,77}. Therefore, to stabilise, detect and quantify metabolites relevant to PD in a simple and cost-effect method, we were required to develop and apply this approach through this thesis (Step D, Figure 1.2).

Aim and scope of thesis

With our hypothesis that a mitochondrial phenotype can be identified in a subset of Parkinson's disease (PD) patients using metabolomics, we aim to develop and improve quantitative metabolomics methods using a targeted MS workflow and integrate the obtained data into constraint-based metabolic models for the study of PD. The association of PD with mitochondrial dysfunction and energy imbalance will create an identifiable metabolic phenotype. To achieve this, we must first target the appropriate metabolites associated with the central carbon and energy metabolism as well as the neurochemical communication and homeostasis. With this selection, we form the backbone of this thesis from which we can derive further understanding into the functional roles of the metabolism associated with PD. We pursue the need for the ability to detect and quantify metabolites with a method that has maximized coverage of all relevant targets and optimised quantitation. Thus, we focus on the development and application of targeted LC-MS/MS workflow, utilising chemical derivatization to achieve this.

In **chapter 2**, we aim to create an absolute quantitative method to study the energy and central carbon metabolism using a single separation and analysis technique. Chemical derivatization is a technique that can be employed to achieve this goal. This derivatization technique described simultaneously labels carboxylic acids, thiols and amines using the reagent dimethylaminophenacyl bromide (DmPABr) in a high-throughput, reliable single RPLC-MS/MS analysis with a 10-minute acquisition time using only positive ionization mode. Few published methods can target carboxylic acids and amines simultaneously – both of which form a large proportion of the human metabolome. In addition, quantitation is further enhanced by isotope-coded derivatization (ICD), which uses internal standards derivatized with an isotopically labelled reagent (DmPABr-D₆). Sixty-four central carbon and energy-related metabolites were detected and quantified from human urine and SUIT-2 cells, including amino acids, *N*-acetylated amino acids, metabolites from the Krebs cycle and pyruvate metabolism, acylcarnitines and medium-/long-chain fatty acids.

In **chapter 3**, the derivatization method described in **chapter 2** is applied to material-limited cell samples. Sensitivity is a common hindrance when faced with low sample concentrations. Previous studies have attempted to overcome this issue in the form of costly microscale separation such as CE and micro/nano-LC coupled to

mass spectrometers with low-diameter ionization emitter sources. By employing chemical derivatization, it is possible to improve chromatographic separation and enhance MS ionization. Favourable, sensitive and specific fragmentation is also achievable. Our novel method applies RPLC-MS/MS analysis to HepG2 cells, ranging from 250 cells to 1×10^5 cells, after fast and accessible derivatization DmPABr. The primary amine, secondary amine, thiol and carboxyl submetabolome are labelled, and we also utilize ICD as done previously. Thirty-seven metabolites were detected and quantified in a sub-10,000 HepG2 cells extract, with an additional 11 metabolites detected below LLOQ.

We discovered a lack of absolute quantitative metabolite reference values in relation to the mammalian brain whilst trying to study neurological disorders. Therefore, in **chapter 4**, we pursued the quantitation of neurochemicals across 25 regions of the rat brain. However, as highlighted previously, analytical methods have their pros and cons. Here, we utilised the benzoyl chloride derivatization technique as it has the ability to stabilise vulnerable catecholamines and capture neuroactive metabolites. To achieve this, we optimised LLE extraction and followed it with the derivatization LC-MS/MS technique. After the analysis, we obtained a comprehensive profile of 43 metabolites including important neurotransmitters such as dopamine, epinephrine, norepinephrine, GABA and serotonin. Additionally, we covered the urea cycle, and polyamine and tyrosine metabolism extensively. The brain regions investigated range from the frontal lobe to the brain stem, covering regions such as the orbitofrontal cortex, cerebral cortex, ventromedial prefrontal cortex and subcortical structure. After generation of the absolute quantitative reference values, we believe this data can be integrated into metabolic models, thus improving our understanding of the mammalian brain.

In **chapter 5**, utilising induced pluripotent stem cell (iPSC)-derived dopaminergic neurons, we conducted a multi-omics data investigation to understand the functionality and potentially identify vulnerabilities in Parkinson's disease. We then integrated the multi-omics data into a genome scale constraint-based reconstruction and analysis model that focused on the metabolism. With the focus being heavily dependent on absolute quantitative metabolomics, applying AccQ-tag derivatization with RPLC-MS. AccQ-Tag was used to capture the biogenic amine and neurochemical profile. Additionally, we used GC-MS to quantify sugars. With these values, the

Chapter 1

1

Recon3D reconstruction of the generic human metabolome was used to generate stoichiometrically and flux consistent constraint-based model of dopaminergic neuron metabolism. The metabolism was constrained using manual literature curation, transcriptomics, and the metabolomics input. With this, we constructed the *iNESC2DN* model that can be used for biomarker discovery, therapeutic treatment strategy identification and drug repurposing¹⁹.

Finally, we conclude this thesis with **chapter 6**, where we revisit the content of our work and address the future prospective of quantitative metabolomics in the application to human diseases and, specifically, neurological disorders. We also discuss the creation of disease-specific constraint-based models.

References

1. Brenner MH. Political economy and health. *Society and health*. 1995;1:211-246.
2. Navarro V, Muntaner C, Borrell C, et al. Politics and health outcomes. *The Lancet*. 2006;368(9540):1033-1037.
3. Barbash GI. New technology and health care costs--the case of robot-assisted surgery. *The New England journal of medicine*. 2010;363(8):701.
4. Mathers CD, Stevens GA, Boerma T, White RA, Tobias MI. Causes of international increases in older age life expectancy. *The Lancet*. 2015;385(9967):540-548.
5. Cornfield J, Haenszel W, Hammond EC, Lilienfeld AM, Shimkin MB, Wynder EL. Smoking and lung cancer: recent evidence and a discussion of some questions. *Journal of the National Cancer institute*. 1959;22(1):173-203.
6. Hecht SS. Cigarette smoking and lung cancer: chemical mechanisms and approaches to prevention. *The lancet oncology*. 2002;3(8):461-469.
7. Hooper L, Summerbell CD, Higgins JP, et al. Dietary fat intake and prevention of cardiovascular disease: systematic review. *Bmj*. 2001;322(7289):757-763.
8. Diehl AM. Liver disease in alcohol abusers: clinical perspective. *Alcohol*. 2002;27(1):7-11.
9. Eurostat. Ageing Europe - statistics on population developments. https://ec.europa.eu/eurostat/statistics-explained/index.php/Ageing_Europe_-_statistics_on_population_developments#Older_people_E2.80.94_population_overview. Published 2020. Accessed 24th November, 2020.
10. Hung C-W, Chen Y-C, Hsieh W-L, Chiou S-H, Kao C-L. Ageing and neurodegenerative diseases. *Ageing research reviews*. 2010;9:S36-S46.
11. Organization WH. *Neurological disorders: public health challenges*. World Health Organization; 2006.
12. Kitano H. Computational systems biology. *Nature*. 2002;420(6912):206-210.
13. Kitano H. Systems biology: a brief overview. *Science*. 2002;295(5560):1662-1664.

14. Vasilopoulou CG, Margarity M, Klapa MI. Metabolomic analysis in brain research: opportunities and challenges. *Frontiers in Physiology*. 2016;7:183.
15. Bui AA, Van Horn JD, Consortium NBKC. Envisioning the future of 'big data'biomedicine. *Journal of biomedical informatics*. 2017;69:115-117.
16. Ideker T, Galitski T, Hood L. A new approach to decoding life: systems biology. *Annual review of genomics and human genetics*. 2001;2(1):343-372.
17. Mao L, Nicolae A, Oliveira MA, He F, Hachi S, Fleming RM. A constraint-based modelling approach to metabolic dysfunction in Parkinson's disease. *Computational and structural biotechnology journal*. 2015;13:484-491.
18. Thiele I, Swainston N, Fleming RM, et al. A community-driven global reconstruction of human metabolism. *Nature biotechnology*. 2013;31(5):419-425.
19. Brunk E, Sahoo S, Zielinski DC, et al. Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nature biotechnology*. 2018;36(3):272.
20. Mhyre TR, Boyd JT, Hamill RW, Maguire-Zeiss KA. Parkinson's Disease. *Sub-cellular biochemistry*. 2012;65:389-455.
21. Dorsey ER, Bloem BR. The Parkinson pandemic—a call to action. *JAMA neurology*. 2018;75(1):9-10.
22. Klos KJ, Bower JH, Josephs KA, Matsumoto JY, Ahlskog JE. Pathological hypersexuality predominantly linked to adjuvant dopamine agonist therapy in Parkinson's disease and multiple system atrophy. *Parkinsonism & related disorders*. 2005;11(6):381-386.
23. Tritsch NX, Sabatini BL. Dopaminergic modulation of synaptic transmission in cortex and striatum. *Neuron*. 2012;76(1):33-50.
24. Luo Y, Hoffer A, Hoffer B, Qi X. Mitochondria: A Therapeutic Target for Parkinson's Disease? *International Journal of Molecular Sciences*. 2015;16(9):20704-20730.
25. Valente EM, Abou-Sleiman PM, Caputo V, et al. Hereditary early-onset Parkinson's disease caused by mutations in PINK1. *Science*. 2004;304(5674):1158-1160.

26. Kitada T, Asakawa S, Hattori N, et al. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature*. 1998;392(6676):605-608.
27. Zimprich A, Biskup S, Leitner P, et al. Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron*. 2004;44(4):601-607.
28. Gorostidi A, Bergareche A, Ruiz-Martínez J, et al. α -Synuclein Levels in Blood Plasma from LRRK2 Mutation Carriers. *PLoS One*. 2012;7(12):e52312.
29. Bonifati V, Oostra BA, Heutink P. Linking DJ-1 to neurodegeneration offers novel insights for understanding the pathogenesis of Parkinson's disease. *J Mol Med (Berl)*. 2004;82(3):163-174.
30. Horowitz M, Wilder S, Horowitz Z, Reiner O, Gelbart T, Beutler E. The human glucocerebrosidase gene and pseudogene: structure and evolution. *Genomics*. 1989;4(1):87-96.
31. Neudorfer O, Giladi N, Elstein D, et al. Occurrence of Parkinson's syndrome in type I Gaucher disease. *QJM*. 1996;89(9):691-694.
32. Beger RD, Dunn W, Schmidt MA, et al. Metabolomics enables precision medicine: "A White Paper, Community Perspective". *Metabolomics*. 2016;12(9):149.
33. Ivanisevic J, Siuzdak G. The role of metabolomics in brain metabolism research. *Journal of Neuroimmune Pharmacology*. 2015;10(3):391-395.
34. Johnson CH, Gonzalez FJ. Challenges and opportunities of metabolomics. *Journal of cellular physiology*. 2012;227(8):2975-2981.
35. Zhang A, Sun H, Wang P, Han Y, Wang X. Modern analytical techniques in metabolomics analysis. *Analyst*. 2012;137(2):293-300.
36. Markley JL, Brüschweiler R, Edison AS, et al. The future of NMR-based metabolomics. *Current Opinion in Biotechnology*. 2017;43:34-40.
37. DeBerardinis RJ, Chandel NS. Fundamentals of cancer metabolism. *Science Advances*. 2016;2(5).
38. Rasmiena AA, Ng TW, Meikle PJ. Metabolomics and ischaemic heart disease. *Clinical science*. 2013;124(5):289-306.

39. Thompson Legault J, Strittmatter L, Tardif J, et al. A Metabolic Signature of Mitochondrial Dysfunction Revealed through a Monogenic Form of Leigh Syndrome. *Cell Rep.* 2015;13(5):981-989.
40. Wild CP. The exposome: from concept to utility. *International journal of epidemiology.* 2012;41(1):24-32.
41. Wild CP. Complementing the genome with an “exposome”: the outstanding challenge of environmental exposure measurement in molecular epidemiology. In: AACR; 2005.
42. Sampson TR, Debelius JW, Thron T, et al. Gut Microbiota Regulate Motor Deficits and Neuroinflammation in a Model of Parkinson's Disease. *Cell.* 2016;167(6):1469-1480 e1412.
43. Garmier M, Carroll AJ, Delannoy E, et al. Complex I dysfunction redirects cellular and mitochondrial metabolism in Arabidopsis. *Plant Physiol.* 2008;148(3):1324-1341.
44. Langston JW, Ballard P, Tetrud JW, Irwin I. Chronic Parkinsonism in humans due to a product of meperidine-analog synthesis. *Science.* 1983;219(4587):979.
45. Willacey CCW, Naaktgeboren M, Lucumi Moreno E, et al. LC-MS/MS analysis of the central energy and carbon metabolites in biological samples following derivatization by dimethylaminophenacyl bromide. *Journal of Chromatography A.* 2019:460413.
46. Wishart DS, Feunang YD, Marcu A, et al. HMDB 4.0: the human metabolome database for 2018. *Nucleic acids research.* 2018;46(D1):D608-d617.
47. de Tommaso M, Ceci E, Pica C, et al. Serum levels of N-acetyl-aspartate in migraine and tension-type headache. *J Headache Pain.* 2012;13(5):389-394.
48. Chen J, Zhang J, Wu X, et al. Disordered Metabolic Profiling in Plasma and Tissues of Mice Infected with Artemisinin-Sensitive and -Resistant Plasmodium berghei K173 Determined by (1)H NMR Spectroscopy. *J Proteome Res.* 2019;18(5):1970-1993.
49. Karu N, Deng L, Slae M, et al. A review on human fecal metabolomics: Methods, applications and the human fecal metabolome database. *Anal Chim Acta.* 2018;1030:1-24.

50. Mena-Bravo A, De Castro ML. Sweat: a sample with limited present applications and promising future in metabolomics. *Journal of pharmaceutical and biomedical analysis*. 2014;90:139-147.
51. Longo V, Forleo A, Provenzano SP, et al. HS-SPME-GC-MS metabolomics approach for sperm quality evaluation by semen volatile organic compounds (VOCs) analysis. *Biomedical Physics & Engineering Express*. 2018;5(1):015006.
52. Marincola FC, Dessì A, Corbu S, Reali A, Fanos V. Clinical impact of human breast milk metabolomics. *Clinica Chimica Acta*. 2015;451:103-106.
53. Banks WA. The blood-brain barrier: connecting the gut and the brain. *Regulatory peptides*. 2008;149(1-3):11-14.
54. Harro J. Animal models of depression: pros and cons. *Cell and tissue research*. 2019;377(1):5-20.
55. Moreno EL, Hachi S, Hemmer K, et al. Differentiation of neuroepithelial stem cells into functional dopaminergic neurons in 3D microfluidic cell culture. *Lab Chip*. 2015;15(11):2419-2428.
56. Bellin M, Marchetto MC, Gage FH, Mummery CL. Induced pluripotent stem cells: the new patient? *Nature reviews Molecular cell biology*. 2012;13(11):713-726.
57. Roberts LD, Souza AL, Gerszten RE, Clish CB. Targeted metabolomics. *Current protocols in molecular biology*. 2012;98(1):30.32. 31-30.32. 24.
58. da Silva RR, Dorrestein PC, Quinn RA. Illuminating the dark matter in metabolomics. *Proceedings of the National Academy of Sciences*. 2015;112(41):12549-12550.
59. Annesley TM. Ion suppression in mass spectrometry. *Clinical chemistry*. 2003;49(7):1041-1044.
60. Trietsch SJ, Israels GD, Joore J, Hankemeier T, Vulto P. Microfluidic titer plate for stratified 3D cell culture. *Lab Chip*. 2013;13(18):3548-3554.
61. Anderson AJ, Haus DL, Hooshmand MJ, Perez H, Sontag CJ, Cummings BJ. Achieving stable human stem cell engraftment and survival in the CNS: is the future of regenerative medicine immunodeficient? *Regen Med*. 2011;6(3):367-406.
62. Gottas A, Ripel A, Boix F, Vindenes V, Morland J, Oiestad EL. Determination of dopamine concentrations in brain extracellular fluid using microdialysis

- with short sampling intervals, analyzed by ultra high performance liquid chromatography tandem mass spectrometry. *J Pharmacol Toxicol Methods*. 2015;74:75-79.
63. Wong JM, Malec PA, Mabrouk OS, Ro J, Dus M, Kennedy RT. Benzoyl chloride derivatization with liquid chromatography-mass spectrometry for targeted metabolomics of neurochemicals in biological samples. *J Chromatogr A*. 2016;1446:78-90.
64. Zhang W, Guled F, Hankemeier T, Ramautar R. Utility of sheathless capillary electrophoresis-mass spectrometry for metabolic profiling of limited sample amounts. *Journal of Chromatography B*. 2019;1105:10-14.
65. Luo X, Li L. Metabolomics of Small Numbers of Cells: Metabolomic Profiling of 100, 1000, and 10000 Human Breast Cancer Cells. *Analytical Chemistry*. 2017;89(21):11664-11671.
66. Willacey CC, Karu N, Harms AC, Hankemeier T. Metabolic profiling of material-limited cell samples by dimethylaminophenacyl bromide derivatization with UPLC-MS/MS analysis. *Microchemical Journal*. 2020:105445.
67. Higashi T, Ogawa S. Isotope-coded ESI-enhancing derivatization reagents for differential analysis, quantification and profiling of metabolites in biological samples by LC/MS: A review. *J Pharm Biomed Anal*. 2016;130:181-193.
68. Villas-Bôas SG, Smart KF, Sivakumaran S, Lane GA. Alkylation or silylation for analysis of amino and non-amino organic acids by GC-MS? *Metabolites*. 2011;1(1):3-20.
69. Sinjewel A, Swart E, Lingeman H, Wilhelm A. LC determination of propylene glycol in human plasma after pre-column derivatization with benzoyl chloride. *Chromatographia*. 2007;66(1-2):103-105.
70. Wu Y, Li L. Determination of total concentration of chemically labeled metabolites as a means of metabolome sample normalization and sample loading optimization in mass spectrometry-based metabolomics. *Analytical chemistry*. 2012;84(24):10723-10731.
71. Donáth-Nagy G, Vancea S, Imre S. Comparative study of captopril derivatization reaction by LC-UV, LC-MS and CE-UV methods. *Croatica Chemica Acta*. 2011;84(3):423-427.

72. Fujiwara T, Inoue R, Ohtawa T, Tsunoda M. Liquid-Chromatographic Methods for Carboxylic Acids in Biological Samples. *Molecules*. 2020;25(21):4883.
73. Jaudzems G, Guthrie J, Lahrichi S, Fuerer C. Total Amino Acids by UHPLC–UV in Infant Formulas and Adult Nutritionals, First Action 2018.06. *Journal of AOAC International*. 2019;102(5):1574-1588.
74. Zhou W, Zhanga XY, Duana GL. Liquid-Chromatography Quantitative Analysis of 20 Amino Acids after Derivatization with FMOC-Cl and Its Application to Different Origin Radix isatidis. *Journal of the Chinese Chemical Society*. 2011;58(4):509-515.
75. Meiser J, Weindl D, Hiller K. Complexity of dopamine metabolism. *Cell Communication and Signaling : CCS*. 2013;11:34-34.
76. Song P, Mabrouk OS, Hershey ND, Kennedy RT. In vivo neurochemical monitoring using benzoyl chloride derivatization and liquid chromatography-mass spectrometry. *Anal Chem*. 2012;84(1):412-419.
77. He J, Shi L, Liu S, Jia P, Wang J, Hu R. Ultrasound-mediated synthesis of N,N-bis(phenacyl)aniline under solvent-free conditions. *Monatshefte für Chemie - Chemical Monthly*. 2014;145(1):213-216.

Chapter 2

LC-MS/MS analysis of the central energy and carbon metabolites in biological samples following derivatization by dimethylaminophenacyl bromide

Cornelius C W Willacey, Martijn Naaktgeboren, Edinson Lucumi Moreno, Agnieszka B Wegrzyn, Daan Van der Es, Naama Karu, Ronan M T Fleming, Amy C Harms, and Thomas Hankemeier

Journal of Chromatography A (2019)
Volume 1608, 20 December 2019, 460413

Abstract

Recent advances in metabolomics have enabled larger proportions of the human metabolome to be analyzed quantitatively. However, this usually requires the use of several chromatographic methods coupled to mass spectrometry to cover the wide range of polarity, acidity/basicity and concentration of metabolites. Chemical derivatization allows in principle a wide coverage in a single method, as it affects both the separation and the detection of metabolites: it increases retention, stabilizes the analytes and improves the sensitivity of the analytes. The majority of quantitative derivatization techniques for LC-MS in metabolomics react with amines, phenols and thiols; however, there are unfortunately very few methods that can target carboxylic acids at the same time, which contribute to a large proportion of the human metabolome. Here, we describe a derivatization technique which simultaneously labels carboxylic acids, thiols and amines using the reagent dimethylaminophenacyl bromide (DmPABr). We further improve the quantitation by employing isotope-coded derivatization (ICD), which uses internal standards derivatized with an isotopically-labelled reagent (DmPABr-D₆). We demonstrate the ability to measure and quantify 64 central carbon and energy-related metabolites including amino acids, N-acetylated amino acids, metabolites from the TCA cycle and pyruvate metabolism, acylcarnitines and medium-/long-chain fatty acids. To demonstrate the applicability of the analytical approach, we analyzed urine and SUIT-2 cells utilizing a 15-minute single UPLC-MS/MS method in positive ionization mode. SUIT-2 cells exposed to rotenone showed definitive changes in 28 out of the 64 metabolites, including metabolites from all 7 classes mentioned. By realizing the full potential of DmPABr to derivatize and quantify amines and thiols in addition to carboxylic acids, we extended the coverage of the metabolome, producing a strong platform that can be further applied to a variety of biological studies.

Background

Metabolomics, the younger sibling of genomics and proteomics, is a fast-evolving field which has established itself as a promising approach for understanding biological variations within a range of matrices in humans, animals, microbes and plants¹⁻⁸. The quantitative profiling of metabolites in biological samples is challenging due to the vast number of metabolites, variation in physicochemical properties and the wide range of concentrations in samples. All of these factors result in large differences in the recovery, sensitivity and matrix interferences of these metabolites when analyzed by various methods. Nevertheless, recent advances in mass spectrometry have given scientists the ability to further understand the human metabolome and focus more closely on selected pathway analysis. When we study metabolic pathways, we experience the complexity as they can comprise of many chemical conversions and are intertwined, making a targeted assay with coverage of over 50 relevant metabolites highly beneficial for researchers in metabolism.

Mass spectrometry (MS) has the ability to identify and quantify the metabolome with current methods reaching sensitivities down to picomolar concentrations, even without any prior separation⁹. However, in the majority of cases, chromatography prior to MS is used to better address the challenges introduced by ion suppression, separation of isomers and in-source fragmentation. The three most common separation techniques, LC, GC and CE, have provided robust methodologies to better cover the human metabolome. Each of these techniques has been applied to numerous types of metabolites, and each technique has tailored advantages for specific types of metabolites. For example, UPLC-MS (RP & HILIC) provides coverage for a large proportion of the metabolome with the advantages of high-throughput, sensitivity, reliability and robustness¹⁰. Still, in LC-MS, metabolites can suffer from limited sensitivity, or poor separation of particularly polar metabolites. Quantification of metabolites with electrospray ionization (ESI)-MS can suffer from ion suppression¹¹. This interference can be corrected for by using coeluting isotopically-labelled internal standards, which are of limited availability and excessive costs.

Methods have been developed to combat these problems using advanced separation techniques and also chemical derivatization, which is the focus of this article. Chemical derivatization can be used to increase the separation resolution, sensitivity or to stabilize the metabolites, resulting in an increased metabolic coverage of MS-based metabolomics methods. For instance, benzoyl chloride is used to derivatize catecholamines and their metabolites to prevent oxidation and increase sensitivity in LC-MS¹². In a recent review, Higashi, Ogawa¹³ summarise the current techniques that are used for derivatization and conclude that isotope-coded derivatization (ICD) has the ability to enhance quantification in LC-MS(/MS). ICD is the process of labelling metabolites in a first sample with an unlabelled derivatization reagent and then using an isotopically labelled reagent to derivatize the same metabolite standards in a neat solution, i.e. pure solvent. This mixture, when added to the sample, can act as the corresponding internal standard (IS) for all analytes of interest. The benefit of this technique is the ability to introduce an isotopically labelled equivalent for all metabolites regardless of chemical structure complexity, which corrects for eventual ion suppression. Approaches such as ICD are important during derivatization workflows to compensate for possible matrix effects as the native matrix is altered due to derivatization. However, having an IS for each metabolite provides a tool to adjust for matrix interferences independent of the starting matrices. ICD can provide a cost-effective alternative when stable isotope IS are not available while still enabling improved trueness and precision. In this way, the derivatization reaction method is exploited in an additional manner next to modifying the separation and ionization of metabolites.

Several studies have utilized a range of reagents, some taking advantage of the ICD strategy to improve the quantitative performance^{12,14,15}. Typical examples are benzoyl chloride^{12,14} and dansyl chloride¹⁵ which both label amines, thiols, phenols and some alcohols. Another reagent, dimethylaminophenacyl bromide (DmPABr), has been applied previously to label carboxylic acid groups¹⁶. There were inconsistent reports about the reactivity of DmPABr. Guo, Li¹⁶ reported that DmPABr reacts only with carboxylic acids (i.e., not amines and thiols), and in a follow-up study Peng and Li acknowledged that it reacts also with nucleophiles at certain reaction conditions¹⁷. However, to conform with the aims of their method, liquid-liquid extraction (LLE) was applied to reduce the interference from amino acids and

derivatives, by excluding them altogether. The need for a reliable method that combines labelling of the amine, thiol and carboxylic acid functional groups has been highlighted by previous papers that have required two separate derivatization methods (DmPABr and dansyl chloride) to achieve the same coverage¹⁸.

In the current paper, we expand the utilization of the reagent DmPABr to simultaneously derivatize metabolites with carboxylic acid, amine and thiol functional groups. We did not apply LLE, and analyzed amino acids, N-acetylated amino acids, carnitines, and organic acids using LC-MS in positive ionization mode. We have examined and optimised the reaction conditions to reliably and repeatably derivatize a range of metabolites and analyze them in a single, highly sensitive quantitative method. The reaction mechanism is identical to that of the reagent phenacyl bromide with primary amines¹⁹, secondary amines^{20,21}, thiols^{22,23} and carboxylic acid-containing metabolites (derivatization example shown in Figure 2.1). First, we made adaptations to the method published by Guo, Li ¹⁶Peng, Li ¹⁷Stanislaus, Guo, Li ²⁴ to improve the metabolite coverage to include a wide range of central carbon and energy-related metabolites. Then, we developed a targeted quantitative UPLC-MS/MS method to allow for the sensitive analysis of these metabolites in a single 10-minute analysis. The final applied method was successfully validated for linearity, precision, limits of detection (LOD) and quantification (LOQ). By applying this method to human urine and *in vitro* experiments using human pancreatic cancer cells (SUIT-2), we could confirm the broad applicability of this methodology and biological relevance for the scientific community.

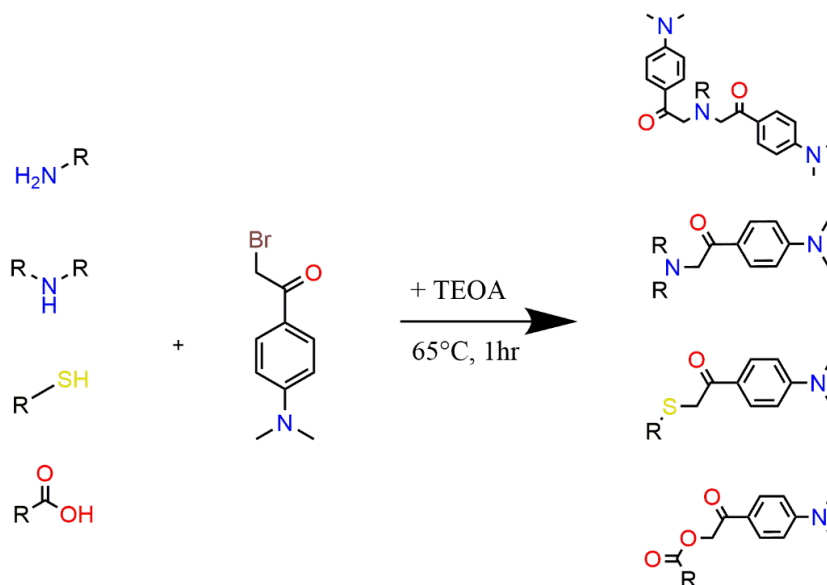


Fig 2.1. The derivatization reaction of DmPABr with the primary amine, secondary amine, thiol and carboxylic acid, respectively.

Materials and Methods

Chemicals

All of the chemicals were purchased from Sigma-Aldrich (St. Louis, USA) unless otherwise stated (abbreviations shown in Table 2.1). The LC-MS grade ACN was sourced from Actu-all Chemicals (Oss, The Netherlands) and de-ionized water was acquired using a Merck Milli-pore A10 purification system (Raleigh, USA). Stock solutions of 5 mg/mL Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val; 10mM NA-Ala, NA-Arg, NA-Asn, NA-Asp, NA-Cys, NA-Gln, NA-Glu, NA-Gly, NA-His, NA-Leu, NA-Lys, NA-Met, NA-Phe, NA-Pro, NA-Ser, NA-Thr, NA-Trp, NA-Tyr, NA-Val; 1 mg/mL AKG, CIT, FUM, ICIT, LAC, MAL, OXA, PYR, SUCC were made in 1:1 DMSO/DMF and stored at -80°C. Stock solutions of 2mg/mL CR; 1 mg/mL AA, AC, BTA, DEA, DDA, EA, EIA, FOR, OLA, OCA, PA, PPA, SA, UDA in 100% ACN (v/v) and stored at -80°C.

Absolute quantitative measurement of the energy metabolism

Metabolite	Abbreviation	Metabolite	Abbreviation
Alanine	Ala	N-acetylmethionine	NA-Met
Arginine	Arg	N-acetylphenylalanine	NA-Phe
Asparagine	Asn	N-acetylproline	NA-Pro
Aspartic acid	Asp	N-acetylserine	NA-Ser
Cysteine	Cys	N-acetylthreonine	NA-Thr
Glutamine	Gln	N-acetyltryptophan	NA-Trp
Glutamic acid	Glu	N-acetyltyrosine	NA-Tyr
Glycine	Gly	N-acetylvaline	NA-Val
Histidine	His	Alpha-Ketoglutaric acid	AKG
Isoleucine	Ile	Citric acids	CITS
Leucine	Leu	Fumaric acid	FUM
Lysine	Lys	Lactic acid	LAC
Methionine	Met	Malic acid	MAL
Phenylalanine	Phe	Oxaloacetic acid	OXA
Proline	Pro	Pyruvic acid	PYR
Serine	Ser	Succinic acid	SUCC
Threonine	Thr	Acetylcarnitine	AC
Tryptophan	Trp	Decanoylcarnitine	DC
Tyrosine	Tyr	Hexanoylcarnitine	HC
Valine	Val	Lauroylcarnitine	LC
N-acetylalanine	NA-Ala	Myristoylcarnitine	MC
N-acetylarginine	NA-Arg	Octanoylcarnitine	OC
N-acetylaspargine	NA-Asn	Palmitoylcarnitine	PC
N-acetylaspartic acid	NA-Asp	Propionylcarnitine	PPC
N-acetylcysteine	NA-Cys	Stearoylcarnitine	SC
N-acetylglutamine	NA-Gln	Arachidonic acid	AA
N-acetylglutamic acid	NA-Glu	Capric acid	DCA
N-acetylglycine	NA-Gly	Caprylic acid	OCA
N-acetylhistidine	NA-His	Dodecanoic acid	DDA
N-acetylisoleucine	NA-Ile	Oleic acid	OLA
N-acetylleucine	NA-Leu	Undecanoic acid	UDA
N-acetyllysine	NA-Lys	Creatinine	CR

Table 2.1. List of the abbreviations for the metabolites analyzed in this method

Derivatization reagent

The DmPABr reagent was purchased from BioConnect BV (Huissen, The Netherlands) and the internal standard DmPABr was synthesised following the published protocol by Guo, Li ¹⁶ using dimethyl sulphate-D₆ instead of dimethyl sulphate-¹³C₂. The structure of the reagent was confirmed using nuclear magnetic resonance (NMR). Also, with reference to the paper from Guo, Li ¹⁶, it is noted that the stability of the metabolites after reaction with DmPABr lasts for up to 6 months in a solution, and does not alter quantitative results²⁴. The DmPABr reagent was stored in ACN at -80°C to prevent the nucleophilic substitution reaction.

Method validation and biological application

Method optimization and validation

The following performance parameters were assessed on all 64 metabolites in triplicate. Method optimization started with the selection of an appropriate alkaline solution at a range of concentrations, comparing triethylamine (TEA at 0, 50, 100, 150, 250, 300, 500 and 750 mM) and triethanolamine (TEOA, at 0, 200, 400, 650, 700, 750, 800 and 1000 mM). The reaction time was assessed for the selected time points 0, 5, 15, 30, 45, 60, 90, 180 and 240 minutes using TEOA (750 mM) incubated for one hour at 65°C. Due to the ability of water to react with the reagent acting as a nucleophile the reaction was assessed in the presence of water at 0%, 20%, 40%, 60%, 80% and 100%. The final optimized method used 750 mM solution of TEOA for derivatization at 65°C for one hour in a shaking incubator. The method was characterised by a matrix-free 8-point calibration line, and by determining the carry-over by a solvent injection blank after injecting the highest calibration level (calibration point 7: supplementary material Table S6). The calibration experiment was replicated (n=5). Matrix effect (ME) is defined as “the direct or indirect alteration or interference in response due to the presence of unintended analytes (for analysis) or other interfering substances in the sample”²⁵. The ME was calculated as the area of the internal standards in the neat solution against the area of the internal standard in the presence of the matrix. The method was also assessed for linearity of the calibration line (n=5) and LOD/LOQ. The LOD and LOQ were calculated using the following equations according to the ICH Q2R1 guidelines – σ being standard deviation of the signal in the blank injection:

Absolute quantitative measurement of the energy metabolism

$$\text{LOD} = (3.3 * \sigma) / \text{slope}$$

$$\text{LOQ} = (10 * \sigma) / \text{slope}$$

$$\text{ME} = (\text{Internal standard in neat solution} / \text{internal standard in matrix}) * 100$$

Urine validation samples

Urine from 10 healthy volunteers (aged 20-30) was collected and pooled and used for method optimization and validation. A volume of 10 μL of urine was transferred to an Eppendorf safe-lock vial (0.5 mL). The urine was dried in a Labconco SpeedVac (MO, United States). The dried content was reconstituted in 10 μL of DMSO/DMF to dissolve the remaining content. Then, 10 μL of triethanolamine (750 mM) was added to the vial, followed by 10 μL of DmPABr (82 mM). The sealed Eppendorf vial was placed into a shaking incubator for 60 minutes at 65°C to complete the derivatization. A total of 10 μL of formic acid (30 mg/mL) was added to the vial to quench the reaction with an additional 30 minutes in the shaking incubator. Then, 5 μL of DmPABr-D₆-labelled metabolites were then added (concentrations in supplementary material Table S6). Before vortexing, 45 μL of ACN was also added to the vial. The content was then transferred to an HPLC vial for analysis. The trueness and precision of the method was generated by using a pooled sample of urine collected from healthy urine donors. Samples were analyzed in repeated experiments on 3 separate days in replicates each day (n=5). Using this data, RSD calculations were performed to demonstrate the lack of variation in the derivatization conditions on separate days.

SUIT-2 oxidative stress analysis and validation

Human pancreatic cancer cells (SUIT-2) were cultured and placed into a 24-well plate, each containing 1×10^6 cells in 0.4 mL of culture media. The SUIT-2 cells were exposed to 1 nM, 10 nM and 100 nM of rotenone for 3, 8 and 24 hours (n=3). The cells were washed with PBS and transferred into an Eppendorf tube. The aliquots were centrifuged at 500 rpm to sediment the cells and the cell media was removed. To the cells, cooled methanol (80% v/v) was added to the cells before probe sonification, followed by centrifugation at 13,000 rpm to produce a protein precipitation; the supernatant was transferred to an Eppendorf safe-lock vial (1.5 mL) without disturbing the pellet. A volume equivalent to 2.5×10^5 cell supernatant was taken to total dryness in a speed vacuum concentrator. The following were added to the vial and vortexed between additions: 10 μL of DMSO/DMF (to first dissolve the dried

content), 10 μL of triethanolamine (750 mM) And 10 μL of DmPABr (82 mM). The sealed Eppendorf vial was placed into a shaking incubator for 60 minutes at 65°C to complete the derivatization. A volume of 10 μL of formic acid (30 mg/mL) was added to the vial to quench the reaction with an additional 30 minutes in the incubation. Finally, 5 μL of DmPABr-D₆-labelled metabolites were diluted in 45 μL of ACN and added to the vial. The content was then transferred to an HPLC vial for analysis.

LC-MS/MS analysis

Samples were analyzed by LC-MS using a Waters Acquity UPLC Class II (Milford, USA) coupled to an ABSciex QTrap 6500 series (Framingham, USA). The samples were run using scheduled multi-reaction monitoring (MRM) in positive mode with selected time windows. An injection of 1 μL was made per sample to minimise detector saturation and maintain desirable peak shape. The analytical column used was a Waters AccQ-tag C18 column (2.1mm x 100 mm, 1.8 μm , 180 Å), maintained at 60°C. Mobile phase solvent A was 0.1% v/v formic acid and 10 mM ammonium formate in water and mobile phase solvent B was 100% acetonitrile. Using the flow rate of 700 $\mu\text{L}/\text{min}$, the gradient profile is as follows: initial, 0.2% B; 1.5 min, 20% B; 4 min, 50% B; 6 min, 90% B; 10 min, 99.8% B; 13 min, 99.8% B; 13.1 min, 0.2% B and 15 min, 0.2% B. The last 6 minutes allow for column washing and equilibration prior to the next injection. The following parameters were used for the AB Sciex QTrap 6500 analysis (MRM transitions shown in Table S2 of the supplementary materials); electrospray ionization was used in positive mode at 4.5 kV. The gas temperature was 600°C. Automated peak integration was performed using AB Sciex MultiQuant Workstation Quantitative Analysis for QTrap; all peaks were visually inspected to ensure adequate integration.

Results & Discussion

Novel derivatization approach

DmPABr derivatization has been a successful method to support untargeted and targeted metabolomics platforms using ICD for carboxylic acid-containing metabolites. It was previously highlighted by Peng, Li¹⁷ that DmPABr also reacts with the amine group of asparagine and labels it twice (once on the acid and once on the amine group), however the double labelled metabolite was reported to be the minor

peak compared to the single derivatized form. As demonstrated in this paper, DmPABr can react with an amine group (once or twice) via a nucleophilic reaction in a quantitative manner, which is useful for LC-MS analysis. In comparison to another common reagent in LC-MS/MS, benzoyl chloride^{12,14}, the reagent DmPABr offers a more versatile and universal solution for derivatization. This reaction, however, is slower and forms a more stable bond, as DmPABr has the ability to react next to the amine group. It also reacts with carboxylic acids without forming an unstable anhydride as the bromine is attached to a methyl group rather to the acyl group. Therefore, we studied the ability of DmPABr to react with multiple functional groups such as the amine, carboxy and thiol groups and to use it for the metabolomics analysis of urine and cell samples.

Selection of metabolites and biological relevance

Utilizing the full capability of DmPABr allows us to extend from only derivatizing carboxylic acids to also targeting amines and thiols which broadens the applicability of the method significantly. To demonstrate this, we chose to measure central carbon and energy-related metabolites related to mitochondrial dysfunction as it requires the analysis of a broad range of chemically diverse metabolites. The key metabolites in aerobic respiration that are imperative for mammalian survival, such as α -ketoglutaric acid, citrates, succinic acid, fumaric acid, malic acid and oxaloacetic acid, are frequently used references to determine changes in mitochondrial function and cellular health in metabolomics studies in urine, plasma and *in vitro* models²⁶. N-acetylation has been known to increase during mitochondrial dysfunction due to the elevation of acetyl-CoA. Therefore, N-acetylated amino acids reflect mitochondrial dysfunction and energy metabolism, and in addition are highly relevant in urine as they remove excess amino acids from the body^{27,28}. Shifting of the energy balance from aerobic respiration to anaerobic respiration can also be noted by measurement of pyruvic acid and lactic acid which will both drastically increase²⁷. An additional target group, acylcarnitines, was also selected, as it reflects energy processes, particularly following beta-oxidation and when fatty acids are transported into the mitochondria^{27,29}. The last target group, fatty acids, is included to represent an alternative energy source by the mitochondria when sugars are inaccessible or depleted. Three common conditions that are often associated with mitochondrial

dysfunction are Parkinson's disease³⁰, Leigh's syndrome²⁷ and diabetes³¹; all of which hold extensive interest within the scientific community.

Optimization of reaction conditions

We investigated and optimised the method to analyze amines, thiols and carboxylic acids to inform on central carbon and energy metabolism. We have found three key factors that affect the derivatization of the functional groups mentioned above: alkalinity of reaction solution, reaction time, and the presence of water during the reaction. Figure 2.2A demonstrates that after 60 minutes the relative peak area did not increase significantly anymore, with high performance parameters (indicated in Table 2.2). This applied to all of the targeted functional groups: Ala (1° amine & carboxylic acids), NA-Asp (carboxylic acids), NA-Cys (thiol & carboxylic acid), PYR (α -keto acid) and AC (carboxylic acid). This derivatization time was considered acceptable in terms of metabolic coverage. We have utilized similar inert base catalysts as in previous articles published for DmPABr that target the carboxylic acid function group, which utilized either 750 mM triethanolamine (TEOA) ¹⁶ or 200 mM triethylamine (TEA) ²⁴. Variations in response using these bases are depicted in Figures 2.2B and 2.2C, leading to the selection of the appropriate conditions for derivatization of amine and thiol groups. Additional experiments indicated that 750 mM of TEOA was the optimum condition for consistent derivatization of metabolites in urine and cells (data not shown). TEOA (750 mM) also provided the most consistent derivatization indicated by identical values over the concentration range of 650 – 800 mM. TEOA was preferred because according to literature it causes less ion suppression in mass spectrometers, than TEA at the concentration tested ³².

Absolute quantitative measurement of the energy metabolism

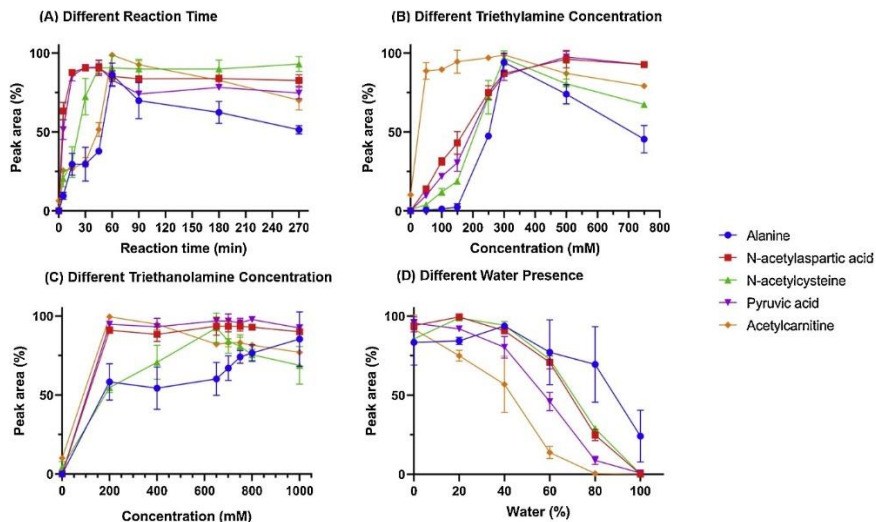


Figure 2.2. DmPABr reaction optimization shown for 5 metabolites (Ala – Blue; NA-Asp – Red; NA-Cys – Green; PYR – Purple; AC – Orange; $n=3$ per condition) representing the major classes selected in the method. (A) the effect of reaction time with 750 mM triethanolamine; (B) use of TEA as base; (C) use of TEOA as base; and (D) the reaction efficiency in the presence of water with 750mM triethanolamine. The data is presented as peak area normalized to the highest peak area.

Chapter 2

2

Analyte	LODs (nM)	LOQs (nM)	Fit (R ²)	RSD (%)	Carryover (%)	Analyte	LODs (nM)	LOQs (nM)	Fit (R ²)	RSD (%)	Carryover (%)
Ala	44.8	134	0.9938	7.7	0.01	NA-Met	0.11	0.33	0.9867	4.4	0.003
Arg	10.3	30.8	0.9944	6	0.004	NA-Phe	0.27	0.82	0.9965	4.8	0.003
Asn	4.01	12	0.9965	3.5	0.01	NA-Pro	15.6	46.8	0.9969	12.5	0.05
Asp	11.4	34	0.9915	9.9	0.02	NA-Ser	5.55	16.6	0.9919	9.6	0.01
Cys	49.5	148.6	0.9834	6	0.01	NA-Thr	10.4	31.3	0.9987	3.7	0.01
Gln	18.5	55.5	0.9913	12.3	0.02	NA-Trp	2	6	0.9965	1.1	0.003
Glu	43.5	130.5	0.9941	2.1	0.01	NA-Tyr	0.12	0.37	0.9901	11.5	0.01
Gly	932.4	2797	0.9952	1.8	0.32	NA-Val	0.52	1.56	0.9932	13	0.002
His	379.4	1138	0.9956	4.1	0.01	AKG	29.7	89.1	0.9909	6.2	0.08
Ile	9.28	27.9	0.9956	7.9	0.01	CITS	1001	3003	0.9988	1.7	0.12
Leu	11.3	33.9	0.9942	7.5	0.01	FUM	69.5	208.5	0.9949	3.9	0.03
Lys	142.8	428.3	0.9959	6.3	0.02	LAC	2192	6578	0.9927	4	0.01
Met	4.33	13	0.9903	0.6	0.01	MAL	43.8	131.5	0.9979	1.6	0.01
Phe	25.6	76.9	0.9997	10.2	0.02	OXA	14	42	0.9950	1.7	0.02
Pro	22.1	66.4	0.9870	3.6	0.003	PYR	55.8	167.3	0.9926	2.4	0.004
Ser	506	1518	0.9917	3.3	0.01	SUCC	23.1	69.4	0.9997	3.8	0.02
Thr	47.6	142.8	0.9950	2.9	0.01	AC	0.36	1.09	0.9951	12.8	0.0001
Trp	5.33	16	0.9978	9	0.02	DC	0.3	0.91	0.9964	8	0.01
Tyr	6.87	20.6	0.9966	7	0.02	HC	0.61	1.82	0.9903	10.2	0.001
Val	11.7	35.6	0.9939	10.5	0.01	LC	0.9	2.7	0.9988	11.3	0.004
NA-Ala	4.21	12.6	0.9941	4.3	0.002	MC	0.5	1.5	0.9996	8.7	0.01
NA-Arg	0.16	0.47	0.9989	9.1	0.001	OC	1.32	3.98	0.9998	5.5	0.01
NA-Asn	0.27	0.81	0.9882	11	0.02	PC	0.91	2.72	0.9970	7.1	0.02
NA-Asp	4.32	13	0.9942	3	0.002	PPC	2.85	8.56	0.9985	8.5	0.0001
NA-Cys	34.5	103.4	0.9975	4.6	0.06	SC	0.25	0.75	0.9993	8.4	0.03
NA-Gln	0.94	2.83	0.9979	6.7	0.001	AA	4.45	13.4	0.9995	1	0.04
NA-Glu	3.65	10.9	0.9979	1.5	0.02	OCA	137	411	0.9992	4.2	0.02
NA-Gly	3.32	9.97	0.9971	4.5	0.001	DCA	104	314	0.9860	2.6	0.01
NA-His	0.16	0.48	0.9917	3.7	0.001	DDA	15.3	46	0.9912	4.5	0.01
NA-Ile	27.7	83	0.9961	1.1	0.001	OLA	185	555	0.9960	4.4	0.02
NA-Leu	30.4	91.2	0.9946	1.5	0.001	UDA	12.9	38.7	0.9985	4.2	0.008
NA-Lys	72.6	217	0.9912	4.9	0.028	CR	548.2	1644	0.9943	1.6	0.001

Table 2.2. Information relating to the carry-over, limit of detection (LOD), limit of quantification (LOQ) and linearity of an 8-point calibration line in aqueous solution. RSD was calculated using calibration point 4 (concentration shown in supplementary Table S2).

Of the options for synthesized isotopically labelled DmPABr, D₆ was used in place of ¹³C₂ on the amine residue of DmPABr, as utilized for high-resolution MS by Guo, Li ¹⁶. With this, we have been able to introduce a mass difference of 6 Da, which is preferable for low-resolution MS compared to the previous addition of 2 Da, and less costly. The mass difference of 2 Da with the internal standard DmPABr-¹³C₂ provided cross-talk interference in the triple-quadrupole MS with the metabolites labelled once, such as long-chain fatty acids and N-acetylated amino acids. However, this was not as detrimental to the metabolites labelled more than once (such as amino acids) as a mass difference gain of 18 Da was observed for metabolites such as Ala (labelled thrice) when using DmPABr-D₆.

Other adaptations of the derivatization procedure were also evaluated, including the total elimination of the aqueous content prior to derivatization with DmPABr to improve reaction efficiency and decrease reaction variability. This was expected to be needed to create a quantitative method, unlike the previous published method that focused on identification. This variability and poor labelling efficiency may arise from the ability of water to act as a nucleophile under basic conditions (deprotonation). It has previously been noted that as little as 5% water content during the derivatization reaction has the ability to reduce the labelling efficiency by hydrolysing DmPABr [23]. It was also reported that the presence of water competes with the metabolites to react with DmPABr as a nucleophile and thus hydrolysing the reagent ³³. We investigated whether a low percentage of water would interfere with the labelling of the amine group. Figure 2.2D shows that the presence of 40% of water did not significantly change the derivatization efficiency. This may be explained by Stanislaus, Guo, Li ²⁴, where the increase in DmPABr concentration from 20 mg/ml¹⁶ to 40 mg/ml (excess reagent) resulted in 2-fold higher rate of the Sn2 reaction, ensuring completion of the reaction. The presence of water above 20% showed a decline in derivatization efficiency and in the presence of 100% water, the reaction is severely compromised. Therefore, we chose to conduct the reaction without the

presence of water to avoid the complications due to possible hydrolysis of the reagent.

Targeted LC-MS/MS method

The aims of the chromatographic method were to combine high-throughput analysis with sensitive measurement of a wide range of chemical classes. After derivatization with DmPABr, polar compounds which are hardly retained in RP could be retained and separated, hence eliminating the need for HILIC separation. Moreover, derivatization with DmPABr allows sensitive and universal analysis in positive ionization mode, instead of in two ionization modes. By using DmPABr, we introduce the tertiary amine group (Figure 2.3 that improves ionization hence enhances signals. Further improvement in intensity of measured metabolites can be gained by careful selection of MRMs. In Figure 2.4 A and B we illustrate that the metabolites that are labelled on the carboxylic acid show common and prominent fragments of 180.0 Da or 134.1 Da. These product fragments are ideal when measuring metabolites such as those involved in the TCA cycle as they lack nitrogen and are difficult to analyze in positive ionization mode without labelling. For metabolites which are labelled multiple times, such as amino acids, the fragments 180.0 Da and 134.1 Da are usually present (Fig. 2.4B) but are not selected as their signal is lower. Instead, a higher mass product ion giving a better signal is often seen. This results from derivatization twice on the amine group, and once on the acid moiety, yielding more sensitive fragments like 319.2 Da and 366.2 Da observed for arginine & alanine in Figure 2.4 C and 2.4 D, respectively. Having a common fragmentation pattern reduces the specificity of metabolite species but with adequate chromatography, this issue can be negated. Additionally, a qualifier transition can also be set which will provide a unique fragmentation pattern to identify the specific metabolite but this will provide a lower sensitivity due to more MS/MS events. For the complete method, the labelling pattern and the chosen MRM transitions are detailed in supplementary Table S2.

Absolute quantitative measurement of the energy metabolism

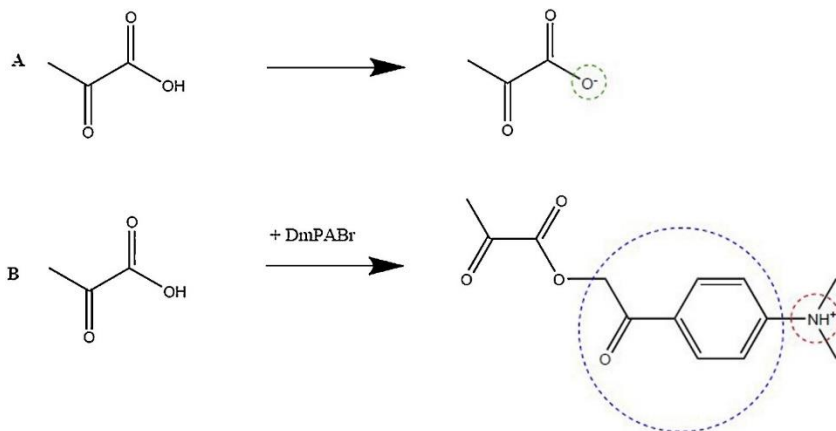


Figure 2.3. Demonstration of MS analysis of pyruvic acid: (A) in conventional negative ionization mode, holding a negative charge on the oxygen (green circle); (B) following DmPA labelling, producing a highly ionisable group (tertiary amine – red circle) and a higher retention group (blue circle).

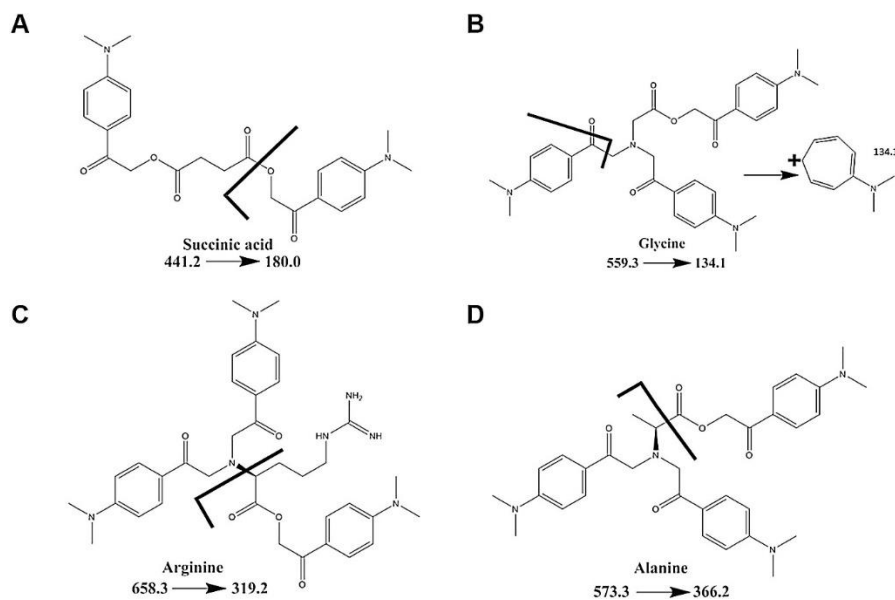


Figure 2.4. The common fragment formation of 180.0 Da & 134.1 Da: the product ion of the derivatization label and metabolite-specific product ions.

To demonstrate the applicability of the method on biological samples, SUIT-2 cell extracts were subjected to derivatization and analysis, resulting in wide representation of various chemical classes (Figure 2.5). The figure demonstrates that owing to strong retention of polar analytes (PYR, Gly) all 64 metabolites are detected in one run in positive ionization mode only within 8.4 min (latest elution, of arachidonic acid, AA). The derivatization leads to unique retention profile, such as the close elution between undecanoic acid (derivatized once) and leucine (derivatized three times), yet with baseline resolution between the isomers Ile and Leu. Another critical pair of isomers, CIT and ICIT present a common challenge in chromatography, and are not baseline resolved here (see Fig.2.5), therefore are reported as total citrates (CITS). In contrast, good separation was observed for N-acetylated amino acids, many of which elute early. The first peak to elute was creatinine (Cr) which is often used to normalize and report metabolites in urine³⁴. DmPABr can successfully derivatize creatinine, unlike the reagents utilized in commercially-available kits, that quantify non-derivatized creatinine (Biocrates AbsoluteIDQ® p180 Kit; Waters AccQ-Tag™).

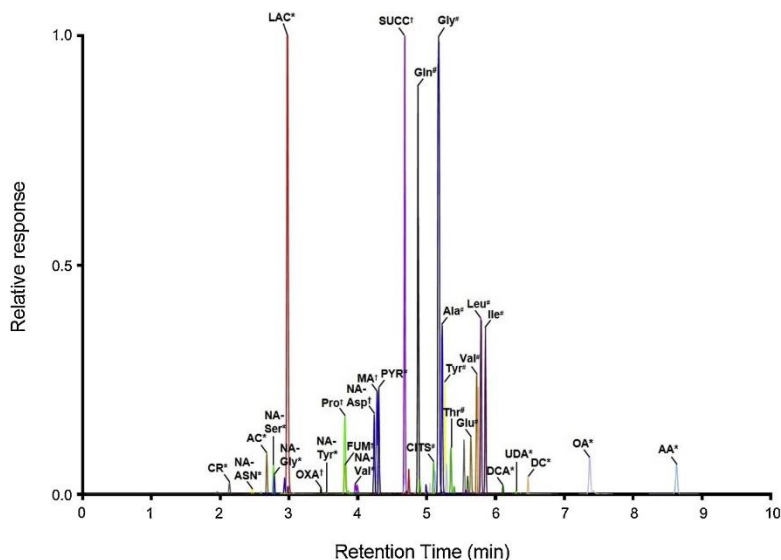


Figure 2.5. LC-MS/MS analysis of 64 metabolites after DmPABr derivatization in SUIT-2 cells. DmPA labelling pattern is also included (* = labelled once; † = labelled twice; # = labelled thrice)

Method performance in neat solutions

The methods performance incorporates the derivatization efficiency and the instrumental response. The method was validated for 64 metabolites that were deemed to be biologically relevant to assess the central energy and carbon metabolism. Using the ICD strategy, each metabolite had its corresponding DmPABr-D₆ internal standard to correct for ion suppression. This resulted in linear calibration lines for all metabolites in neat solutions (Table 2.2). All metabolites, including the amino acids which are derivatized by reacting 2-5 times with DmPABr showed a satisfactory linear calibration ($R^2 > 0.99$) except for Cys ($R^2 = 0.98$) and Pro ($R^2 = 0.98$). The RSD for all metabolites in neat solution was recorded below 12.8% also shown in Table 2.2.

N-acetylated amino acids also showed good analytical performance similar to their free amino acid counterparts. Table 2.2 also shows that the carry-over of the method was negligible (<0.05%). Looking at limits of detection (which are affected by the derivatization process itself), N-acetylated amino acids have a very low LOD, as recorded for NA-Asp (4 nM), NA-Cys (34 nM) and NA-Phe (0.3 nM), which are sufficient for their analysis in urine and cells. N-acetylated amino acids are used as the transport mechanism to excrete excess amino acids (particularly in the urine) that occur in relatively low concentrations when compared to free amino acids in urine^{28,35}. We have circumvented the issues of limited dynamic range of the detector in order to allow good quantitation of a wide concentration range of metabolites. As shown in Table 2.2, the LOD of Gly, His, Ser, CITS and LAC, were higher compared to the other metabolites in this method. This is due to intentional choice of less sensitive MRM channel, to reduce the signal and prevent detector saturation, counteracting the high physiological concentrations in urine or cells. Another intervention to prevent detector saturation took place, namely non-optimal ionization spray voltage throughout (4.5 kV vs. the optimal 5.5 kV). The application of this method to various matrices could benefit from tailoring the MS parameters as well as sample handling to improve the LOD.

Method performance in urine and cells

We applied the quantitative DmPABr method to urine and SUIT-2 cells (Tables 2.3-2.4). Table 2.3 shows the endogenous concentration of the metabolites measured in

urine from healthy males, after normalization to creatinine (measured in the same method). A total of 57 compounds were detected and quantified in urine. The compounds that were not detected include some carnitines and medium chain fatty acids as they do not occur or occur in low concentrations in healthy urine. All of the amino acids and N-acetylated amino acids that have been studied fall within the expected concentrations curated in HMDB³. Urine was assessed for intra-day and inter-day variability. Amino acids such as Ala, Ile, Trp had very low intra- and inter-day variability (all below 10%) and N-acetylated amino acids including NA-Asp, which is crucial for neurological studies, had an intra-day and inter-day variability of 2.2% and 5.4%, respectively. Creatinine had an intra-day and inter-day variability of 4.5% and 7.4%, respectively, which provides consistent normalization factor, if desired. Overall, the amino acids had a higher derivatization variability than other classes, which is probably due to the second step of derivatization on the 2° amine requiring more energy, compared to the single reaction with the carboxylic acid group.

Absolute quantitative measurement of the energy metabolism

Analyte	Urine Concentration ($\mu\text{mol}/\text{mmol}$ creatinine)	Intraday precision (%)	Interday precision (%)	Matrix effect (%)	Analyte	Urine Concentration ($\mu\text{mol}/\text{mmol}$ creatinine)	Intra- Day precision (%)	Inter- Day precision (%)	Matrix effect (%)
Ala	3.81	4.7	5.1	92.3	NA-Met	0.03	2.3	4	37.7
Arg	0.32	5.2	9.5	68.45	NA-Phe	0.05	3.1	4.2	51.7
Asn	1.11	4.9	7	74	NA-Pro	1.31	29.3	23.5	31
Asp	0.04	18.9	16.5	95	NA-Ser	0.25	2.4	4.1	66.5
Cys	11.5	12.9	21	91.2	NA-Thr	1.7	3.2	3.6	53.9
Gln	3.67	22.8	18.5	83.5	NA-Trp	0.2	3.6	3.7	45.5
Glu	0.12	5.5	6.4	81	NA-Tyr	0.047	5	9.8	36.8
Gly	136	1.4	4.2	67	NA-Val	2.07	3.7	13.2	16.6
His	76.3	5.2	5.7	42.7	AKG	0.5	10.6	12.8	71.3
Ile	0.23	6.7	7.7	80.3	CITS	7.47	13.1	11.5	73.9
Leu	0.52	6.4	7	85	FUM	0.13	6.7	7.7	144.4
Lys	1.22	23.1	28	42.2	LAC	0.89	3.3	8.9	43.5
Met	0.11	7.7	11.9	78.2	MAL	0.13	7.1	7.4	42.8
Phe	1.35	4.1	9.2	80.2	OXA	0.3	3.9	6.6	57.2
Pro	0.11	6.4	6.7	129.5	PYR	0.17	11.5	10.2	50.3
Ser	4.43	7.3	15.5	72.1	SUCC	5.31	2.7	3.4	39.4
Thr	1.21	6.3	16.1	70.2	AC	1.93	4.2	5	83.38
Trp	1.27	7.7	7.6	77.3	DC	0.001	9.5	7.8	93.1
Tyr	1.33	5.4	6.8	85	HC	ND	N/A	N/A	77.9
Val	0.62	5	7.3	78.9	LC	ND	N/A	N/A	83.6
NA-Ala	0.54	3.6	6.7	43.4	MC	ND	N/A	N/A	91.9
NA-Arg	0.67	1.3	1.5	67.5	OC	0.007	5.3	5	58.2
NA-Asn	2.01	7.7	7.4	54.8	PC	ND	N/A	N/A	96.3
NA-Asp	2.29	2.2	5.4	55.1	PPC	0.07	2.2	2.8	88.7
NA-Cys	1.45	4.1	10.4	27.4	SC	ND	N/A	N/A	95.4
NA-Gln	1.26	5.5	6.4	34.7	AA	0.02	9.4	13.5	103.6
NA-Glu	0.67	3.6	3	62.3	OCA	ND	N/A	N/A	66.7
NA-Gly	0.13	3.6	7.1	33.3	DCA	ND	N/A	N/A	90.7
NA-His	1.28	3	4.8	63.9	DDA	0.008	10.9	24.1	83.2
NA-Ile	0.14	11.5	10.3	46.7	OLA	0.006	11.2	16.3	95.8
NA-Leu	0.13	4.9	6.4	47.6	UDA	0.004	21.4	39.1	98.3
NA-Lys	0.3	24	22	60	CR*	N/A	4.5	7.4	38.1

* Creatinine was used for normalization

Chapter 2

Table 2.3. Method performance in urine of healthy men aged 18-30 to calculate concentration, intraday and interday precision calculated as %RSD. ND = Not Detected, N/A= Not Applicable

The application of the method to cells was conducted by measuring untreated SUIT-2 cells. The cells were assessed for intra-day variability and not inter-day due to practical considerations (Table 2.4). All 64 metabolites were detected from the intracellular environment in cell lysate. This provided an excellent readout on the energy state of the cells using metabolites involved in the TCA cycle (i.e., CITS, FUM & SUCC) and glycolysis (PYR & LAC).

Analyte	SUIT-2 Conc. (fmol/mg)	Intra-Day precision (%)	Matrix effect (%)	Analyte	SUIT-2 Conc. (fmol/mg)	Intra-Day precision (%)	Matrix effect (%)
Ala	408	4.9	121.1	NA-Met	0.022	1.8	77.6
Arg	243	10.3	75.8	NA-Phe	0.063	10	78.3
Asn	173	6.7	40.7	NA-Pro	0.111	9.6	78.6
Asp	56.8	5.7	75.1	NA-Ser	13.1	7.2	73.8
Cys	166	30.8	98.6	NA-Thr	5.66	5.7	74.2
Gln	2360	10.1	75.3	NA-Trp	0.034	10.8	52.7
Glu	474	16.8	83.3	NA-Tyr	0.126	6.5	87
Gly	2970	3.4	87.4	NA-Val	0.034	11.7	86.3
His	467	3.7	62.5	AKG	7.75	10	93.7
Ile	277	8.7	90.7	CITS	214	9.8	97
Leu	333	9.7	72.1	FUM	245	12.7	131.8
Lys	130	28.5	78.4	LAC	846	2.4	72.5
Met	127	18.7	87.9	MAL	114	3.8	75.1
Phe	309	9	70.2	OXA	18.2	13.6	75.4
Pro	206	10.3	114.5	PYR	191	1.4	78.1
Ser	270	24.9	105.1	SUCC	249	13.8	60.6
Thr	401	31.1	74.1	AC	12.1	13.8	86.4
Trp	59.7	10.7	97.8	DC	0.006	23.3	109.5
Tyr	284	7.7	66	HC	0.033	18.6	81.5
Val	307	10.8	76.4	LC	0.01	9.7	85.6
NA-Ala	3.05	3.7	82.1	MC	0.099	8	94.1
NA-Arg	0.699	8.6	74.2	OC	0.011	19.9	68
NA-Asn	159	9.8	64.8	PC	0.195	11.9	98.4

Absolute quantitative measurement of the energy metabolism

Analyte	SUIT-2 Conc. (fmol/mg)	Intra-Day precision (%)	Matrix effect (%)	Analyte	SUIT-2 Conc. (fmol/mg)	Intra-Day precision (%)	Matrix effect (%)
NA-Asp	56.6	2	76.7	PPC	0.379	7.4	91.2
NA-Cys	50.5	12	54.4	SC	0.067	9.4	96.5
NA-Gln	8.55	8.6	81.1	AA	0.457	19.5	98.2
NA-Glu	0.964	2.1	71	OCA	2.57	24.8	67
NA-Gly	3.75	5.6	53.3	DCA	1.16	27.3	93.7
NA-His	0.318	19.6	81.1	DDA	1.64	13	83
NA-Ile	0.126	8	56.1	OLA	6.84	14.7	88.1
NA-Leu	0.217	14	55	UDA	0.669	25.6	88.9
NA-Lys	0.045	31.5	84.2	CR	305	11.6	81.7

Table 2.4. Method performance in SUIT-2 cell to calculate concentration per mg and intra-day variability

Matrix effect was calculated for both urine and cell samples. The matrix effect was significant during the early eluting peaks such as the N-acetylated amino acids in urine. However, the matrix interferences were not as high during the analysis of cells. The presence of a matrix effect shows the importance of using the ICD technique to provide an internal standard for all metabolites.

Method application in cells to reflect energy metabolism

2 Exposing cells to rotenone is a well-established method to induce mitochondrial dysfunction by blocking complex I of the electron transport chain. This allows us to simulate disorders such as Parkinson's disease and Leigh's syndrome. Application of the method to SUIT-2 cells exposed to 100 nM rotenone showed the most changes in the central metabolism when compared to 1 nM and 10 nM (data shown in supplementary Figures S4 & S5 for 8 and 3 hour rotenone exposure). Out of the 64 metabolites, 28 showed significant changes including metabolites from all of the 7 classes (additional data shown in supplementary Table S3). The top ten most distinguishing metabolites were: Asn ($p = 0.0001$); AKG ($p = 0.0001$); Pro ($p = 0.0004$); CITS ($p = 0.0004$); PYR ($p = 0.0004$); OXA ($p = 0.001$); FUM ($p = 0.002$); AC ($p = 0.002$); SUCC ($p = 0.002$) and MC ($p = 0.003$) as shown in Figure 2.6. The two metabolites with the largest fold change, CITS and AKG, coincide with the shutdown of the TCA cycle by rotenone inhibition of complex I of the electron transport chain³⁶. The same reduction was seen in SUCC and FUM but to a lesser extent. Interestingly, we also identified changes in the N-acetylated amino acids such as NA-Glu ($p = 0.01$), NA-Ala ($p = 0.003$). N-acetylation of amino acids has been documented in mitochondrial dysfunction but has not been extensively studied due to difficulty with analysis^{27,37,38}. The results obtained with our novel method demonstrate its potential in studying the role of central carbon and energy metabolism such as mitochondrial dysfunction and Parkinson's disease³⁹.

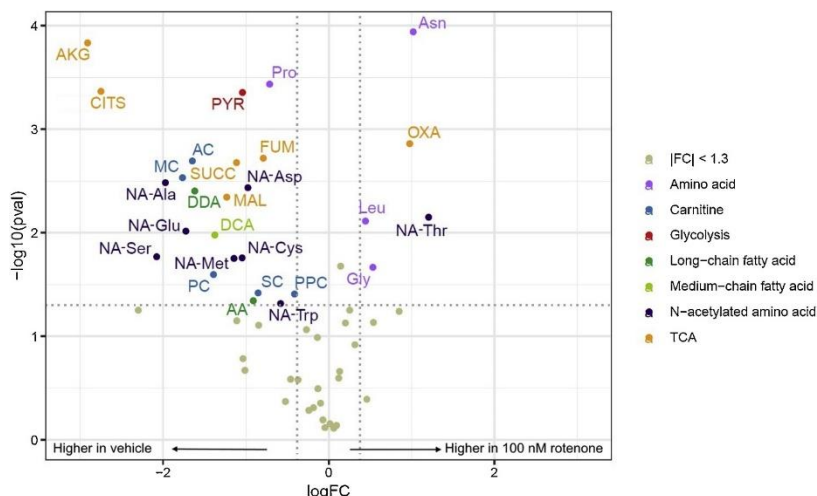


Figure 2.6. Volcano plot of SUIT-2 cells exposed to 100 nM rotenone for 24 hours vs control. All of the metabolites involved in the method show biological changes across all classes once treated with rotenone (amino acids – light purple; carnitines – blue; glycolysis – red; long-chain fatty acids – dark green; medium-chain fatty acid – light green; N-acetylated amino acids – dark purple; and TCA metabolites – orange).

Conclusion

The presented work expands the metabolite coverage of DmPABr by implementation of changes to the reaction conditions. Actually, the derivatization of several functional groups including carboxylic acids, primary amines, secondary amines and thiol groups was achieved in a consistent and robust way for the first time using DmPABr. This vastly improves the coverage of the method allowing for a higher proportion of the human metabolome to be targeted. We have demonstrated that using DmPABr derivatization to its full ability allows us to create a single RPLC-MS/MS analysis within 10 min acquisition time using only positive ionization mode. Since we used a targeted metabolomics method employing internal standards which were derivatized with stable isotope-labelled reagent, we can report each metabolite reliably with its absolute concentration. The great versatility of this approach was demonstrated by quantification of urine metabolites (normalized to DmPABr-derivatized creatinine). Applying the method to SUIT-2 cells exposed to

rotenone showed significant changes in almost 50% of the metabolites covered in this method, including common TCA and glycolysis metabolites and not-so-commonly studied N-acetylated amino acids. Understanding and documenting these biological and biochemical changes in the brain could prove invaluable for future research into neurodegenerative diseases, and requires investigation with a precise and robust quantitative analytical approach. A computational approach towards the prediction of derivatization of metabolites, and the prediction of retention for new metabolites, will further support the method application to cover a wider range of metabolites in complex matrices.

Acknowledgements

The author expresses thanks to: Jaco van Veldhoven for support during the synthesis of DmPABr-D₆; Alisa L. Willacey for advice and guidance during the finalization of the study and Alida Kindt for statistical support. This project was supported by the SysMedPD project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no, 668738.

References

1. Bayona LM, Videnova M, Choi YH. Increasing Metabolic Diversity in Marine Sponges Extracts by Controlling Extraction Parameters. *Marine drugs*. 2018;16(10).
2. Karu N, Deng L, Slae M, et al. A review on human fecal metabolomics: Methods, applications and the human fecal metabolome database. *Anal Chim Acta*. 2018;1030:1-24.
3. Wishart DS, Feunang YD, Marcu A, et al. HMDB 4.0: the human metabolome database for 2018. *Nucleic acids research*. 2018;46(D1):D608-d617.
4. Li T, Li E, Suo Y, et al. Energy metabolism and metabolomics response of Pacific white shrimp *Litopenaeus vannamei* to sulfide toxicity. *Aquatic toxicology (Amsterdam, Netherlands)*. 2017;183:28-37.
5. Karu N, Wilson R, Hamede R, et al. Discovery of Biomarkers for Tasmanian Devil Cancer (DFTD) by Metabolic Profiling of Serum. *J Proteome Res*. 2016;15(10):3827-3840.
6. Goldansaz SA, Guo AC, Sajed T, Steele MA, Plastow GS, Wishart DS. Livestock metabolomics and the livestock metabolome: A systematic review. *PLoS One*. 2017;12(5):e0177675.
7. Sumner LW, Lei Z, Nikolau BJ, Saito K. Modern plant metabolomics: advanced natural product gene discoveries, improved technologies, and future prospects. *Natural product reports*. 2015;32(2):212-229.
8. Ramirez-Gaona M, Marcu A, Pon A, et al. YMDB 2.0: a significantly expanded version of the yeast metabolome database. *Nucleic acids research*. 2017;45(D1):D440-d445.
9. Kantae V, Ogino S, Noga M, et al. Quantitative profiling of endocannabinoids and related N-acylethanolamines in human CSF using nano LC-MS/MS. *J Lipid Res*. 2017;58(3):615-624.
10. Zhou B, Xiao JF, Tuli L, Ransom HW. LC-MS-based metabolomics. *Molecular bioSystems*. 2012;8(2):470-481.
11. Lopes AS, Cruz EC, Sussulini A, Klassen A. Metabolomic Strategies Involving Mass Spectrometry Combined with Liquid and Gas Chromatography. *Advances in experimental medicine and biology*. 2017;965:77-98.

12. Song P, Mabrouk OS, Hershey ND, Kennedy RT. In vivo neurochemical monitoring using benzoyl chloride derivatization and liquid chromatography-mass spectrometry. *Anal Chem.* 2012;84(1):412-419.
13. Higashi T, Ogawa S. Isotope-coded ESI-enhancing derivatization reagents for differential analysis, quantification and profiling of metabolites in biological samples by LC/MS: A review. *J Pharm Biomed Anal.* 2016;130:181-193.
14. Wong JM, Malec PA, Mabrouk OS, Ro J, Dus M, Kennedy RT. Benzoyl chloride derivatization with liquid chromatography-mass spectrometry for targeted metabolomics of neurochemicals in biological samples. *J Chromatogr A.* 2016;1446:78-90.
15. Guo K, Li L. Differential ^{12}C -/ ^{13}C -isotope dansylation labeling and fast liquid chromatography/mass spectrometry for absolute and relative quantification of the metabolome. *Anal Chem.* 2009;81(10):3919-3932.
16. Guo K, Li L. High-Performance Isotope Labeling for Profiling Carboxylic Acid-Containing Metabolites in Biofluids by Mass Spectrometry. *Analytical Chemistry.* 2010;82(21):8789-8793.
17. Peng J, Li L. Liquid-liquid extraction combined with differential isotope dimethylaminophenacyl labeling for improved metabolomic profiling of organic acids. *Anal Chim Acta.* 2013;803:97-105.
18. Sun Y, Chen D, Liu J, et al. Metabolic profiling associated with autophagy of human placenta-derived mesenchymal stem cells by chemical isotope labeling LC-MS. *Experimental cell research.* 2018;372(1):52-60.
19. Egorov M, Delpech B, Aubert G, et al. A concise formation of N-substituted 3,4-diarylpyrroles – synthesis and cytotoxic activity. *Organic & Biomolecular Chemistry.* 2014;12(9):1518-1524.
20. Hu Q, Zhang Z, Liu Y, Imamoto T, Zhang W. ZnCl_2 -Promoted Asymmetric Hydrogenation of β -Secondary-Amino Ketones Catalyzed by a P-Chiral Rh-Bisphosphine Complex. *Angewandte Chemie International Edition.* 2015;54(7):2260-2264.
21. He J, Shi L, Liu S, Jia P, Wang J, Hu R. Ultrasound-mediated synthesis of N,N-bis(phenacyl)aniline under solvent-free conditions. *Monatshefte für Chemie - Chemical Monthly.* 2014;145(1):213-216.

22. Yousaf TI, Lewis ES. Group transfers. 1. Enolate structures contributing to the transition state for nucleophilic substitution on α -substituted carbonyl compounds. *Journal of the American Chemical Society*. 1987;109(20):6137-6142.
23. Loghmani-Khouzani H, Poorheravi MR, Sadeghi MMM, Caggiano L, Jackson RFW. α -Fluorination of β -ketosulfones by Selectfluor™ F-TEDA-BF₄. *Tetrahedron*. 2008;64(30):7419-7425.
24. Stanislaus A, Guo K, Li L. Development of an isotope labeling ultra-high performance liquid chromatography mass spectrometric method for quantification of acylglycines in human urine. *Analytica Chimica Acta*. 2012;750:161-172.
25. Shah VP, Midha KK, Findlay JW, et al. Bioanalytical method validation--a revisit with a decade of progress. *Pharmaceutical research*. 2000;17(12):1551-1557.
26. Hallan S, Afkarian M, Zelnick LR, et al. Metabolomics and Gene Expression Analysis Reveal Down-regulation of the Citric Acid (TCA) Cycle in Non-diabetic CKD Patients. *EBioMedicine*. 2017;26:68-77.
27. Thompson Legault J, Strittmatter L, Tardif J, et al. A Metabolic Signature of Mitochondrial Dysfunction Revealed through a Monogenic Form of Leigh Syndrome. *Cell Rep*. 2015;13(5):981-989.
28. Mardinoglu A, Shoaie S, Bergentall M, et al. The gut microbiota modulates host amino acid and glutathione metabolism in mice. *Mol Syst Biol*. 2015;11(10):834.
29. Schoeman JC, Hou J, Harms AC, et al. Metabolic characterization of the natural progression of chronic hepatitis B. *Genome Med*. 2016;8(1):64.
30. Celardo I, Martins LM, Gandhi S. Unravelling mitochondrial pathways to Parkinson's disease. *British Journal of Pharmacology*. 2014;171(8):1943-1957.
31. Adams SH. Emerging perspectives on essential amino acid metabolism in obesity and the insulin-resistant state. *Adv Nutr*. 2011;2(6):445-456.
32. Wu Z, Gao W, Phelps MA, Wu D, Miller DD, Dalton JT. Favorable effects of weak acids on negative-ion electrospray ionization mass spectrometry. *Analytical chemistry*. 2004;76(3):839-847.

33. Nallu M, Subramanian M, Vembu N, Akber Hussain A. Effect of binary aqueous-organic solvents on the reaction of phenacyl bromide with nitrobenzoic acid(s) in the presence of triethylamine. *International Journal of Chemical Kinetics*. 2004;36(7):401-409.
34. Knobel RB, Smith JM. Laboratory blood tests useful in monitoring renal function in neonates. *Neonatal network : NN*. 2014;33(1):35-40.
35. Kawase T, Nagasawa M, Ikeda H, Yasuo S, Koga Y, Furuse M. Gut microbiota of mice putatively modifies amino acid metabolism in the host brain. *Br J Nutr*. 2017;117(6):775-783.
36. Garmier M, Carroll AJ, Delannoy E, et al. Complex I dysfunction redirects cellular and mitochondrial metabolism in Arabidopsis. *Plant Physiol*. 2008;148(3):1324-1341.
37. de Tommaso M, Ceci E, Pica C, et al. Serum levels of N-acetyl-aspartate in migraine and tension-type headache. *J Headache Pain*. 2012;13(5):389-394.
38. Gerlo E, Van Coster R, Lissens W, Winckelmans G, De Meirleir L, Wevers R. Gas chromatographic-mass spectrometric analysis of N-acetylated amino acids: the first case of aminoacylase I deficiency. *Anal Chim Acta*. 2006;571(2):191-199.
39. Lewitt PA, Li J, Lu M, Beach TG, Adler CH, Guo L. 3-hydroxykynurenine and other Parkinson's disease biomarkers discovered by metabolomic analysis. *Mov Disord*. 2013;28(12):1653-1660.

Supplementary information

Table S1: List of abbreviations

Metabolite	Abbreviation	Metabolite	Abbreviation
Alanine	Ala	N-acetylmethionine	NA-Met
Arginine	Arg	N-acetylphenylalanine	NA-Phe
Asparagine	Asn	N-acetylproline	NA-Pro
Aspartic acid	Asp	N-acetylserine	NA-Ser
Cysteine	Cys	N-acetylthreonine	NA-Thr
Glutamine	Gln	N-acetyltryptophan	NA-Trp
Glutamic acid	Glu	N-acetyltyrosine	NA-Tyr
Glycine	Gly	N-acetylvaline	NA-Val
Histidine	His	Alpha-Ketoglutaric acid	AKG
Isoleucine	Ile	Citric acids	CITS
Leucine	Leu	Fumaric acid	FUM
Lysine	Lys	Lactic acid	LAC
Methionine	Met	Malic acid	MAL
Phenylalanine	Phe	Oxaloacetic acid	OXA
Proline	Pro	Pyruvic acid	PYR
Serine	Ser	Succinic acid	SUCC
Threonine	Thr	Acetylcarnitine	AC
Tryptophan	Trp	Decanoylcarnitine	DC
Tyrosine	Tyr	Hexanoylcarnitine	HC
Valine	Val	Lauroylcarnitine	LC
N-acetylalanine	NA-Ala	Myristoylcarnitine	MC
N-acetylarginine	NA-Arg	Octanoylcarnitine	OC
N-acetylaspargine	NA-Asn	Palmitoylcarnitine	PC
N-acetylaspartic acid	NA-Asp	Propionylcarnitine	PPC
N-acetylcysteine	NA-Cys	Stearoylcarnitine	SC
N-acetylglutamine	NA-Gln	Arachidic acid	AA
N-acetylglutamic acid	NA-Glu	Capric acid	DCA
N-acetylglycine	NA-Gly	Caprylic acid	OCA
N-acetylhistidine	NA-His	Dodecanoic acid	DDA
N-acetylisoleucine	NA-Ile	Oleic acid	OLA
N-acetylleucine	NA-Leu	Undecanoic acid	UDA
N-acetyllysine	NA-Lys	Creatinine	CR

Chapter 2

Table S2. Target analyte list with MRM parameters for the AB Sciex Qtrap 6500 and labelling pattern including the number of DmPA-tags per class.

Metabolites	Retention time (mins)	Precursor (m/z)	Product (m/z)	Collision Energy (eV)	Labelling pattern			
					1° Amine	2° Amine	Thiol	Acid
CR	2.12	275.1	134.1	40	1	0	0	0
NA-Arg	2.35	378.2	180	30	0	0	0	1
NA-Asn	2.46	336.2	134.1	25	0	0	0	1
NA-Gln	2.54	350.2	134.1	25	0	0	0	1
NA-Ser	2.62	309.1	180	20	0	0	0	1
AC	2.64	365.1	134.1	35	0	0	0	1
NA-Gly	2.74	279.1	180	15	0	0	0	1
NA-Thr	2.82	323.2	180	20	0	0	0	1
PPC	2.94	379.2	134.1	35	0	0	0	1
NA-Ala	2.99	293.1	180	15	0	0	0	1
LAC	2.99	252.0	180	5	0	0	0	1
NA-Lys	3.3	511.2	134.1	25	2	0	0	1
NA-Pro	3.35	319.2	70	30	0	0	0	1
NA-Tyr	3.41	385.2	180	15	0	0	0	1
NA-His	3.42	520.2	271.2	45	0	0	0	1
OXA	3.46	455.0	134	35	0	0	0	2
NA-Met	3.65	353.3	180	15	0	0	0	1
NA-Val	3.68	321.2	180	15	0	0	0	1
Pro	3.8	438.1	289.1	25	0	1	0	1
FUM	3.8	439.1	134	30	0	0	0	2
HC	3.95	421.4	140.1	35	0	0	0	1
NA-Leu	4.04	335.2	180	15	0	0	0	1
NA-Trp	4.07	408.3	180	15	0	0	0	1

Absolute quantitative measurement of the energy metabolism

Metabolites	Retention time (mins)	Precursor (m/z)	Product (m/z)	Collision Energy (eV)	Labelling pattern			
					1° Amine	2° Amine	Thiol	Acid
NA-Ile	4.08	335.2	180	15	0	0	0	1
NA-Phe	4.12	369.2	180	15	0	0	0	1
NA-Asp	4.23	498.1	180	30	0	0	0	2
PYR	4.28	250.1	180	15	0	0	0	1
MAL	4.31	457.1	134.1	35	0	0	0	2
NA-Cys	4.34	486.2	134.1	35	0	0	1	1
NA-Glu	4.35	512.2	134.1	35	0	0	0	2
His	4.38	639.2	244.2	15	2	0	0	1
OC	4.6	449.1	134.1	35	0	0	0	1
SUCC	4.68	441.1	134	30	0	0	0	2
Arg	4.71	658.2	319.2	35	2	0	0	1
Gln	4.91	630.2	340.2	30	2	0	0	1
Asn	4.94	616.2	339.2	30	2	0	0	1
AKG	5.1	469.2	180	30	0	0	0	2
CITS	5.1	676.1	180	10	0	0	0	3
DC	5.11	477.5	140.1	35	0	0	0	1
Ser	5.15	589.2	408.2	10	2	0	0	1
Gly	5.2	559.2	134.1	10	2	0	0	1
Ala	5.23	573.2	366.2	20	2	0	0	1
Tyr	5.28	665.2	458.2	30	2	0	0	1
Thr	5.31	603.2	422.2	25	2	0	0	1
Cys	5.4	766.2	470.2	25	2	0	1	1
LC	5.54	505.5	140.1	35	0	0	0	1
Lys	5.56	476.6	134.1	30	4	0	0	1
Trp	5.56	688.2	340.2	30	2	0	0	1
Asp	5.6	778.2	392.2	30	2	0	0	2
Met	5.62	633.2	426.2	25	2	0	0	1

Metabolites	Retention time (mins)	Precursor (m/z)	Product (m/z)	Collision Energy (eV)	Labelling pattern			
					1° Amine	2° Amine	Thiol	Acid
OCA	5.65	306.2	180	20	0	0	0	1
Glu	5.68	792.2	585.2	35	2	0	0	2
Val	5.75	601.2	394.2	30	2	0	0	1
Phe	5.77	649.2	442.2	30	2	0	0	1
Leu	5.83	615.2	408.2	30	2	0	0	1
Ile	5.87	615.2	408.2	30	2	0	0	1
MC	5.92	533.6	140.1	35	0	0	0	1
DCA	6.08	334.3	180	20	0	0	0	1
PC	6.26	561.6	140.1	35	0	0	0	1
UDA	6.27	348.3	180	20	0	0	0	1
DDA	6.44	362.3	180	20	0	0	0	1
SC	6.55	589.6	140.1	35	0	0	0	1
OLA	7.34	444.5	180	20	0	0	0	1
AA	8.58	474.5	180	20	0	0	0	1
CR-IS	2.11	281.1	140.1	40	1	0	0	0
NA-Arg-IS	2.35	384.2	186	30	0	0	0	1
NA-Asn-IS	2.4	342.2	140.1	25	0	0	0	1
NA-Gln-IS	2.48	356.2	140.1	25	0	0	0	1
NA-Ser-IS	2.56	315.1	186	20	0	0	0	1
AC-IS	2.64	371.1	140.1	35	0	0	0	1
NA-Gly-IS	2.7	285.1	186	15	0	0	0	1
NA-Thr-IS	2.79	329.2	186	20	0	0	0	1
NA-Ala-IS	2.94	299.1	186	15	0	0	0	1
PPC-IS	2.94	385.2	140.1	35	0	0	0	1
LAC-IS	2.95	258.1	186	15	0	0	0	1
NA-Pro-IS	3.3	325.2	76	30	0	0	0	1
NA-Lys-IS	3.3	523.2	140.1	25	2	0	0	1

Absolute quantitative measurement of the energy metabolism

Metabolites	Retention time (mins)	Precursor (m/z)	Product (m/z)	Collision Energy (eV)	Labelling pattern			
					1° Amine	2° Amine	Thiol	Acid
NA-Tyr-IS	3.36	391.2	186	15	0	0	0	1
NA-His-IS	3.4	365.2	140.1	45	0	0	0	1
OXA-IS	3.43	467.1	140	35	0	0	0	2
NA-Met-IS	3.6	359.3	186	15	0	0	0	1
NA-Val-IS	3.61	327.2	186	15	0	0	0	1
Pro-IS	3.78	450.1	295.1	25	0	1	0	1
FUM-IS	3.78	451.1	140	30	0	0	0	2
HC-IS	3.95	427.4	140.1	35	0	0	0	1
NA-Leu-IS	3.96	341.2	186	15	0	0	0	1
NA-Ile-IS	4	341.2	186	15	0	0	0	1
NA-Trp-IS	4.03	414.3	186	15	0	0	0	1
NA-Phe-IS	4.08	375.2	186	15	0	0	0	1
NA-Asp-IS	4.2	510.1	186	30	0	0	0	2
PYR-IS	4.24	256.1	186	15	0	0	0	1
MAL-IS	4.24	469.1	140	35	0	0	0	2
NA-Glu-IS	4.29	524.2	140.1	35	0	0	0	2
His-IS	4.34	657.2	250.2	40	2	0	0	1
OC-IS	4.6	455.0	140.1	35	0	0	0	1
NA-Cys-IS	4.65	498.2	140.1	35	0	0	1	1
SUCC-IS	4.65	453.1	140	30	0	0	0	2
Arg-IS	4.7	676.2	331.2	35	2	0	0	1
Asn-IS	4.86	634.2	351.2	30	2	0	0	1
Gln-IS	4.87	648.2	352.2	30	2	0	0	1
AKG-IS	5.07	481.2	186	30	0	0	0	2
CITS-IS	5.07	694.1	186	40	0	0	0	3
Ser-IS	5.08	607.2	420.2	25	2	0	0	1
DC-IS	5.11	483.5	140.1	35	0	0	0	1

Metabolites	Retention time (mins)	Precursor (m/z)	Product (m/z)	Collision Energy (eV)	Labelling pattern			
					1° Amine	2° Amine	Thiol	Acid
Gly-IS	5.13	577.2	140.1	30	2	0	0	1
Ala-IS	5.18	591.2	378.2	25	2	0	0	1
Tyr-IS	5.23	683.2	470.2	30	2	0	0	1
Thr-IS	5.25	621.2	434.2	25	2	0	0	1
Cys-IS	5.4	395.6	154.2	25	2	0	1	1
Lys-IS	5.5	491.6	140.1	30	4	0	0	1
Trp-IS	5.51	706.2	352.2	30	2	0	0	1
Asp-IS	5.54	802.2	410.2	30	2	0	0	2
LC-IS	5.54	511.5	140.1	35	0	0	0	1
Met-IS	5.56	651.2	438.2	25	2	0	0	1
Glu-IS	5.6	810.2	597.2	35	2	0	0	2
OCA-IS	5.65	312.21	186	20	0	0	0	1
Val-IS	5.69	619.2	406.2	30	2	0	0	1
Phe-IS	5.72	667.2	454.2	30	2	0	0	1
Leu-IS	5.76	633.2	420.2	30	2	0	0	1
Ile-IS	5.82	633.2	420.2	30	2	0	0	1
MC-IS	5.92	539.6	140.1	35	0	0	0	1
DCA-IS	6.08	340.3	186	20	0	0	0	1
PC-IS	6.26	567.6	140.1	35	0	0	0	1
UDA-IS	6.27	354.3	186	20	0	0	0	1
DDA-IS	6.44	368.3	186	20	0	0	0	1
SC-IS	6.55	596.7	140.1	35	0	0	0	1
OA-IS	7.34	450.5	186	20	0	0	0	1
AA-IS	8.58	480.5	186	20	0	0	0	1

Absolute quantitative measurement of the energy metabolism

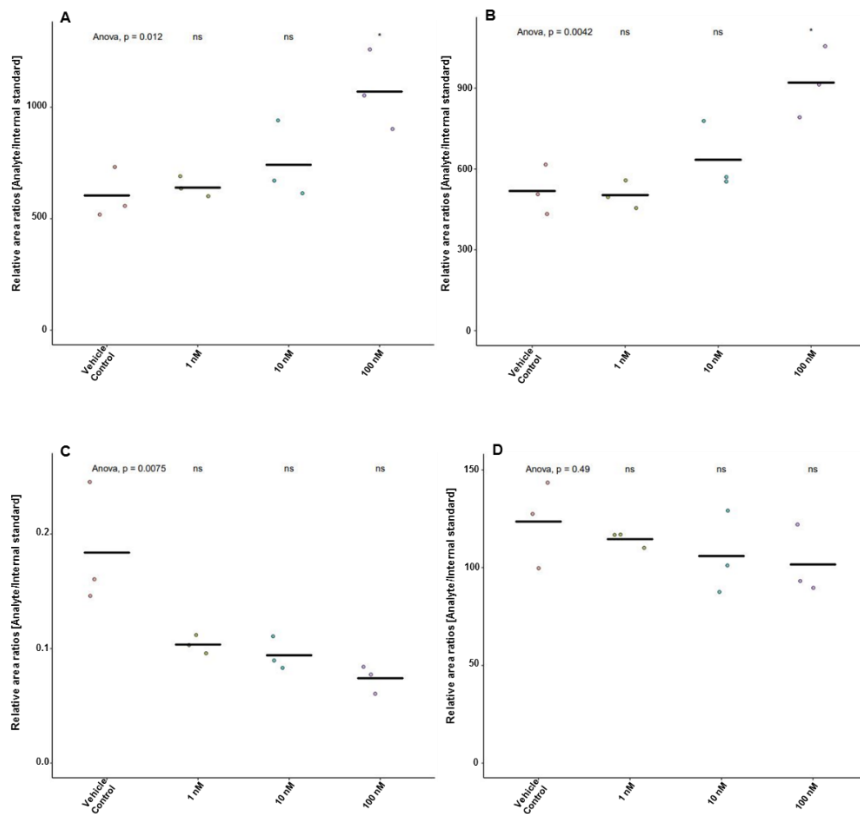
Table S3. Concentration changes between the SUIT-2 cell vehicle control and cells exposed to 100 nM rotenone. The values were used to create the volcano plot shown in figure 6.

Metabolites	Vehicle mean conc.	100 nM Rotenone mean	LogFC	pval
Asn	6.61	7.63	1.02	0.0001
AKG	3.44	0.535	-2.91	0.0001
Pro	12.5	11.8	-0.713	0.0004
CITS	-0.209	-2.96	-2.75	0.0004
PYR	8.44	7.40	-1.04	0.0004
OXA	4.17	5.15	0.979	0.001
FUM	12.8	12.0	-0.789	0.002
AC	1.64	-0.00632	-1.65	0.002
SUCC	6.63	5.52	-1.11	0.002
MC	-5.87	-7.64	-1.76	0.003
NA.Ala	4.41	2.44	-1.97	0.003
NA.Asp	6.74	5.77	-0.98	0.004
DDA	11.3	9.70	-1.62	0.004
MAL	8.29	7.06	-1.23	0.005
NA-Thr	1.24	2.45	1.21	0.007
Leu	10.2	10.7	0.444	0.008
NA-Glu	6.03	4.31	-1.72	0.01
DCA	9.46	8.09	-1.37	0.01
NA-Ser	7.01	4.93	-2.08	0.02
NA-Cys	9.13	8.08	-1.04	0.02
NA-Met	-1.75	-2.90	-1.14	0.02
Tyr	8.44	8.59	0.145	0.02
Gly	-2.37	-1.84	0.535	0.02
PC	-4.16	-5.55	-1.39	0.03
SC	-1.60	-2.45	-0.852	0.04
PPC	0.822	0.408	-0.412	0.04
AA	-0.645	-1.56	-0.91	0.045

Metabolites	Vehicle	100 nM	LogFC	pval
	mean	Rotenone		
	conc.	mean		
NA-Trp	1.39	0.811	-0.580	0.048

Absolute quantitative measurement of the energy metabolism

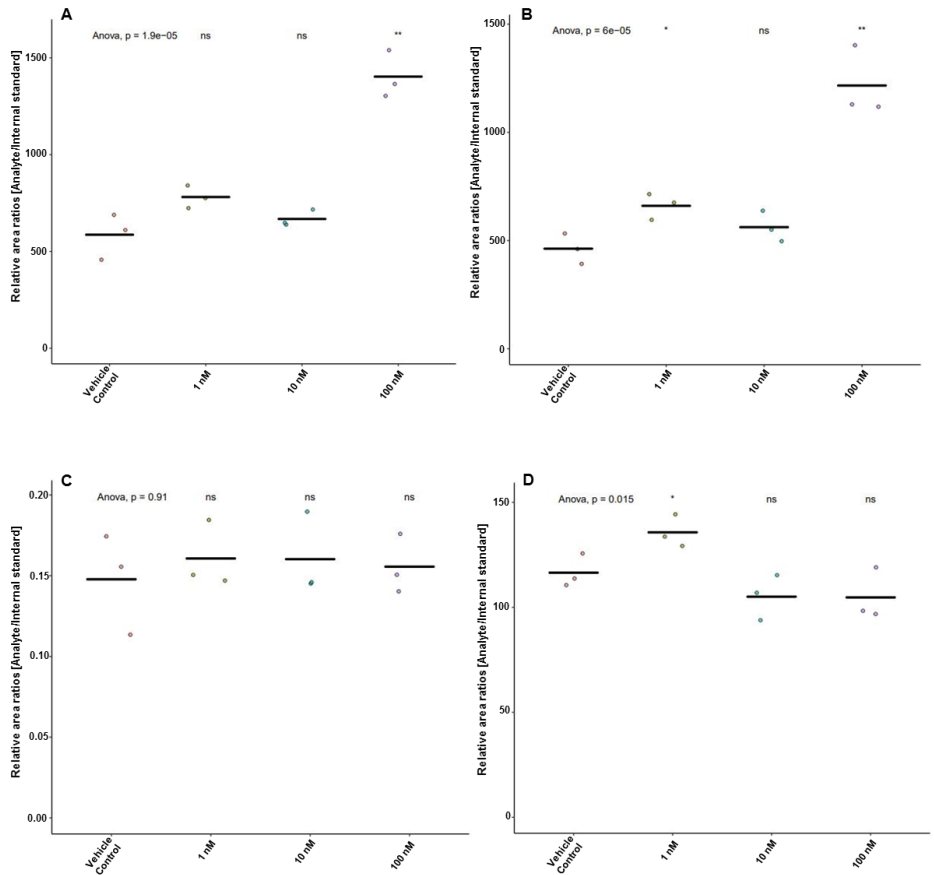
Figure S4. SUIT-2 cells exposed to rotenone for 8 hours. The plots show four metabolites (A: PYR, B: MAL, C: PC, and D: SUCC), at four conditions: vehicle control, 1 nM rotenone, 10 nM rotenone and 100 nM rotenone. ns = not significant.



2

Chapter 2

Figure S5. SUI-2 cells exposed to rotenone for 3 hours. The plots show four metabolites (A: PYR, B: MAL, C: PC, and D: SUCC), at four conditions: vehicle control, 1 nM rotenone, 10 nM rotenone and 100 nM rotenone. ns = not significant.



Absolute quantitative measurement of the energy metabolism

*Table S6. Metabolite standard concentrations (μM), utilized for calibration plots. * concentration of the internal standard.*

Metabolites	Cal 1	Cal 2	Cal 3*	Cal 4	Cal5	Cal 6	Cal 7
Ala	9.29	18.57	37.14	74.29	148.57	297.14	594.29
Arg	1.16	2.32	4.64	9.29	18.57	37.14	74.29
Asn	4.18	8.36	16.71	33.43	66.86	133.71	267.43
Asp	2.32	4.64	9.29	18.57	37.14	74.29	148.57
Cys	13.93	27.86	55.71	111.43	222.86	445.71	891.43
Gln	13.93	27.86	55.71	111.43	222.86	445.71	891.43
Glu	0.93	1.86	3.71	7.43	14.86	29.71	59.43
Gly	69.64	139.29	278.57	557.14	1114.29	2228.57	4457.14
Ile	0.93	1.86	3.71	7.43	14.86	29.71	59.43
Leu	1.39	2.79	5.57	11.14	22.29	44.57	89.14
Lys	6.96	13.93	27.86	55.71	111.43	222.86	445.71
Met	0.46	0.93	1.86	3.71	7.43	14.86	29.71
Phe	2.79	5.57	11.14	22.29	44.57	89.14	178.29
Pro	0.23	0.46	0.93	1.86	3.71	7.43	14.86
Ser	13.93	27.86	55.71	111.43	222.86	445.71	891.43
Thr	5.57	11.14	22.29	44.57	89.14	178.29	356.57
Trp	2.79	5.57	11.14	22.29	44.57	89.14	178.29
Tyr	3.71	7.43	14.86	29.71	59.43	118.86	237.71
Val	1.86	3.71	7.43	14.86	29.71	59.43	118.86
His	27.86	55.71	111.43	222.86	445.71	891.43	1782.86
NA-Ala	0.31	0.62	1.23	2.46	4.93	9.86	19.71
NA-Gly	0.15	0.29	0.58	1.16	2.32	4.65	9.3
NA-Ile	0.07	0.15	0.3	0.6	1.2	2.39	4.79
NA-Leu	0.07	0.15	0.29	0.58	1.16	2.33	4.66
NA-Pro	0.04	0.07	0.14	0.28	0.57	1.13	2.26
NA-Val	0.14	0.27	0.55	1.1	2.19	4.38	8.76
NA-Phe	0.03	0.06	0.13	0.25	0.5	1.01	2.01
NA-Trp	0.05	0.11	0.22	0.44	0.87	1.74	3.49
NA-Tyr	0.03	0.06	0.11	0.22	0.45	0.9	1.79
NA-Asp	0.74	1.49	2.97	5.94	11.89	23.78	47.56
NA-Glu	0.25	0.51	1.01	2.03	4.06	8.11	16.23
NA-Arg	0.27	0.53	1.07	2.14	4.28	8.55	17.11

Chapter 2

2

Metabolites	Cal 1	Cal 2	Cal 3*	Cal 4	Cal5	Cal 6	Cal 7
NA-His	0.07	0.14	0.27	0.55	1.1	2.19	4.39
NA-Lys	0.14	0.27	0.54	1.09	2.18	4.36	8.72
NA-Ser	0.12	0.23	0.47	0.94	1.87	3.74	7.49
NA-Thr	0.31	0.62	1.23	2.46	4.93	9.86	19.71
NA-Cys	0.4	0.81	1.61	3.22	6.45	12.9	25.79
NA-Met	1.44	2.89	5.78	11.55	23.11	46.22	92.43
NA-Asn	1.03	2.05	4.1	8.21	16.42	32.83	65.66
NA-Gln	0.51	1.01	2.03	4.05	8.1	16.2	32.4
PYR	0.7	1.39	2.79	5.57	11.14	22.29	44.57
AKG	2.32	4.64	9.29	18.57	37.14	74.29	148.57
MAL	0.46	0.93	1.86	3.71	7.43	14.86	29.71
LAC	4.64	9.29	18.57	37.14	74.29	148.57	297.14
CIT/ICIT	58.04	116.07	232.14	464.29	928.57	1857.14	3714.29
SUCC	1.86	3.71	7.43	14.86	29.71	59.43	118.86
FUM	0.23	0.46	0.93	1.86	3.71	7.43	14.86
OXA	0.46	0.93	1.86	3.71	7.43	14.86	29.71
AC	0.93	1.86	3.71	7.43	14.86	29.71	59.43
OC	0.02	0.05	0.09	0.19	0.37	0.74	1.49
PPC	0.22	0.45	0.89	1.79	3.57	7.14	14.29
OCA	0.46	0.93	1.86	3.71	7.43	14.86	29.71
DCA	0.46	0.93	1.86	3.71	7.43	14.86	29.71
DDA	0.01	0.03	0.06	0.11	0.22	0.45	0.89
UDA	0.05	0.09	0.19	0.37	0.74	1.49	2.97
DC	0.22	0.45	0.89	1.79	3.57	7.14	14.29
HC	0.22	0.45	0.89	1.79	3.57	7.14	14.29
LC	0.22	0.45	0.89	1.79	3.57	7.14	14.29
MC	0.22	0.45	0.89	1.79	3.57	7.14	14.29
PC	0.22	0.45	0.89	1.79	3.57	7.14	14.29
SC	0.22	0.45	0.89	1.79	3.57	7.14	14.29
OLA	0.23	0.46	0.93	1.86	3.71	7.43	14.86
AA	0.05	0.09	0.19	0.37	0.74	1.49	2.97
CR	29.02	58.04	116.07	232.14	464.29	928.57	1857.14

Chapter 3

Metabolic profiling of material-limited cell samples by dimethylaminophenacyl bromide derivatization with UPLC-MS/MS analysis

Cornelius C W Willacey, Naama Karu, Amy C Harms and Thomas Hankemeier

Microchemical Journal (2020)
Volume 159, December 2020, 105445

Abstract

The ability to dissect the intracellular metabolome is vital in the study of diverse biological systems and models. However, limited cell availability is a challenge in metabolic profiling due to the low concentrations affecting the sensitivity. This is further exacerbated by modern technologies such as 3D microfluidic cell culture devices that provide a physiologically realistic environment, compared to traditional techniques such as cell culture in 2D well-plates. Attempts to address sensitivity issues have been made via advances in microscale separation such as CE and micro/nano-LC coupled to mass spectrometers with low-diameter ionization emitter sources. An alternative approach is sample derivatization, which improves the chromatographic separation, enhances the MS ionization, and promotes favourable fragmentation in terms of sensitivity and specificity. Although chemical derivatization is widely used for various applications, few derivatization methods allow sensitive analysis below 1×10^4 cells. Here, we conduct RPLC-MS/MS analysis of HepG2 cells ranging from 250 cells to 1×10^5 cells, after fast and accessible derivatization by dimethylaminophenacyl bromide (DmPABr), which labels the primary amine, secondary amine, thiol and carboxyl submetabolome, and also utilizes the isotope-coded derivatization (ICD). The analysis of 1×10^4 HepG2 cells accomplished quantification of 37 metabolites within 7-minute elution, and included amino acids, N-acetylated amino acids, acylcarnitines, fatty acids and TCA cycle metabolites. The metabolic coverage includes commonly studied metabolites involved in the central carbon and energy-related metabolism, showing applicability in various applications and fields. The limit of detection of the method was below 20 nM for most amino acids, and sub 5 nM for the majority of N-acetylated amino acids and acylcarnitines. Good linearity was recorded for derivatized standards in a wide biological range representing expected metabolite levels in 2-10,000 cells. Intraday variability in 5×10^3 HepG2 cells was below 20% RSD for concentrations measured of all but two metabolites. The method sensitivity at the highest dilution of cell extract, 250 HepG2 cells, enabled the quantification of twelve metabolites and the detection of three additional metabolites below LLOQ. Where possible, performance parameters were compared to published methodologies that measure cell extract samples. The presented work shows a proof of concept for harnessing a derivatization method for sensitive analysis of material-limited biological samples. It

Miniaturization using chemical derivatization

offers an attractive tool with further potential for enhanced performance when coupled to low-material suitable technologies such as CE-MS and micro/nano LC-MS.

Background

The study of the metabolome provides an important insight into biochemical processes within an organism in a range of environments. The field of metabolomics has been fast-evolving, and delivered quantitative and qualitative analysis of metabolites in various matrices from humans ¹, animals ², plants ³ and microbes ⁴, among others. Metabolomics analysis offers diagnostic support ⁵, and improves our understanding of disease mechanisms ⁶, therapeutic response ⁷ and off-target drug action ⁸. Improvements in technology and knowledge create opportunities for new approaches to study intricate and dynamic biological systems, and many of these new approaches are the analysis of volume-limited samples and low concentration samples. Volume-limited and low concentration samples in metabolomics include microdialysate ⁹, CSF ¹⁰, microfluidic cell culture ¹¹, region specific tissue sampling ¹², blood and interstitial fluid collected by microneedle-arrays, and similar low-volume devices ¹³. Metabolomics analysis of cells poses challenges due to the low availability of cell content, multiple analysis methods required in order to measure metabolites from different classes, and limited number of methods that offer accurate quantitation.

Over the past decade, 3D microfluidic cell cultures grew more popular as it provide a more realistic biological environment compared to conventional 2D culture techniques ^{14,15} and also offer high-throughput and dynamic sampling ^{14,16}. The majority of the devices used in microfluidic cell cultures are below 1×10^4 cell count, but not down to single-cell, as this represents a different field of study. Cell cultures are widely used for the research of various health conditions as they offer advantages in sample availability for multiple sets of experiments, fewer ethical considerations and more controlled conditions compared to limited clinical samples from patients. Unfortunately, the study of the intracellular metabolic profile is limited due to the aforementioned reasons, which are mainly low sensitivity and difficulty in the accurate quantitation of a wide range of relevant metabolites.

The metabolomics community tends to apply two analytical approaches in mass spectrometry to address volume/material-limited sample sensitivity issues. The most common approach is selection of advantageous technology and instruments to achieve the required application, and the less common approach is chemical derivatization to modify the analytes and improve the analysis performance ¹⁷. The

former approach harnesses the advancements in technology by optimizing the separation technique, ionization interfaces or selecting the appropriate mass spectrometer design. Classic methods for the analysis of volume-limited samples use CE-MS ¹⁸, UPLC-MS ¹⁹, microLC-MS ²⁰ and nanoLC-MS ²¹. These techniques are often coupled to advanced ESI sources such as sheathless interfaces in CE ^{22,23} and micro-/nano-ESI emitters in LC-MS applications ^{21,24,25}. Despite miniaturized LC methods being available, a limited number of studies have used them to measure metabolites, and they are more common within the field of pharmacology and environmental sciences ²⁶. The latter approach, chemical derivatization, promotes sensitivity and accuracy in several ways: increased selectivity and resolution between interfering peaks (ion suppressors; isobaric and isomeric compounds); improved peak-shape; enhanced ionization efficiency, and more favorable ionization behaviour. Most derivatization reagents often increase the hydrophobicity of metabolites when the labelling group is relatively large (e.g. benzene rings) resulting in higher retention of metabolites on a reverse-phase column, requiring higher organic content in order to elute. The higher organic solvent content is more suitable for efficient ionization (i.e. improved desolvation), allowing more ions to enter the MS, thus promoting higher sensitivity ¹⁷. Chemical derivatization has been instrumental in GC-MS for several decades to improve volatility, separation and sensitivity ²⁷, and there has been a recent resurgence in modern analytical applications using non-GC methods. Chemical derivatization strategies such as benzoyl chloride ²⁸, dansyl chloride ²⁹, dimethylaminophenacyl bromide (DmPABr) ³⁰ and N-dimethyl-amino naphthalene-1-sulfonyl chloride (Dns-Cl) ³¹ are commonly referenced and applied to label specific functional groups. Recently, Lkhagya *et al.* compared the sensitivity gain that can be achieved in LC-MS/MS by different derivatization reagents, Dansyl, OPA, Fmoc, Dabsyl and Marfey's, when applied to metabolically characterize a medicinal Chinese herb ³². They showed that each reagent has its own strength in producing a sensitivity gain, and the main limitation was metabolome coverage. The derivatization strategies mentioned above also employ the isotope-coded derivatization (ICD) approach ¹⁷ in which the metabolites of interest are labelled by both a derivatization reagent and an isotopically-labelled reagent, generating an internal standard for each metabolite, with full coverage and in a cost-effective manner. While most publications of methods that target volume/material-limited samples discuss the sensitivity enhancements achieved via introduction of the chemical label, only a few

publications offer methodical evaluation of the sensitivity gain over conventional approaches ^{28,29,33}. Despite its advantages, derivatization techniques suffer from some limitations. They involve time-consuming processes, require additional processing steps (risk of errors), and depend on labelling efficiency (reproducibility of recovery), which also limits the coverage according to the reagent reactivity with the functional groups. Fortunately, chemical reagents have been developed to cover the majority of functional groups found within the human metabolome. The reagents benzoyl chloride, dansyl chloride and Dns-Cl label metabolites containing amine, phenol and thiol functional groups.

In a recent publication, we demonstrated a method that expands the functional group coverage of DmPABr to label primary amines (twice), secondary amines (once), thiols (once) and carboxylic acids (once) (derivatization reaction shown in Fig 3.1), further enhancing the quantitative coverage of the human metabolome ³⁴. However, we intentionally reduced the ionization and collision energy efficiency to allow quantitation, within the dynamic range of the detector, of high abundance metabolites in urine and cells in high numbers. In the presented work, we show a proof-of-concept for the analysis of cells in the microfluidic range (below 1×10^4 cells) following derivatization with the reagent DmPABr. Additionally, we evaluate the performance of the targeted quantitative method applying the DmPABr reagent against commonly utilised methods. We demonstrate absolute quantification of the central carbon and energy metabolism in a low cell-count sample of human HepG2 cells, which are commonly used to demonstrate analytical methods due to their robustness and ease of use. Two million HepG2 cells were lysed and further diluted to solutions containing 1×10^5 to 250 cells, representing the microfluidic cell culture range. Cell dilution is a common approach, however, it does not address additional limitations in microfluidic cell culture devices, mainly discrepancies in metabolite concentrations per cell number extract, as stated by Gunda *et al.* ³⁵. With regards to metabolic coverage, we selected the metabolites due to their wide range of physicochemical properties, ability to be derivatized by DmPABr, applicability to human diseases, and coverage by other previously-published volume-limited sample analyses, in order to provide a fair comparison. The metabolites covered within this method include amino acids, N-acetylated amino acids, acylcarnitines and organic acids. We showcase the capability of the DmPABr derivatization method to provide a

Miniaturization using chemical derivatization

sensitive quantitative analysis of low numbers of HepG2 cells without the need for miniaturised separation and ionization techniques.

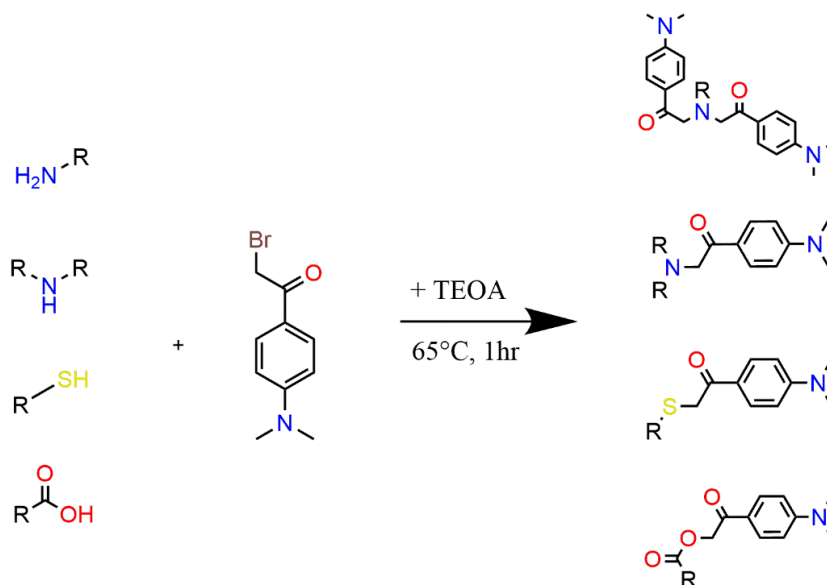


Fig 3.1. The derivatization reaction of DmPABr with the primary amine, secondary amine, thiol and carboxylic acid, respectively.

Materials and Methods

Chemicals

All chemicals were purchased from Sigma-Aldrich (St. Louis, USA) unless stated otherwise. Stock solutions of 5 mg/mL L-alanine (Ala), L-arginine (Arg), L-asparagine (Asn), L-aspartic acid (Asp), L-cysteine (Cys), L-glutamine (Gln), L-glutamic acid (Glu), glycine (Gly), L-histidine (His), L-isoleucine (Ile), L-leucine (Leu), L-lysine (Lys), L-methionine (Met), L-phenylalanine (Phe), L-proline (Pro), L-serine (Ser), L-threonine (Thr), L-tryptophan (Trp), L-tyrosine (Tyr), L-valine (Val) and creatinine (CR) were solubilized in DMSO/DMF (1:1 v/v) and were stored at -80 °C. Additionally, 1 mg/mL N-acetylalanine (NA-Ala), N-acetylarginine (NA-Arg), N-acetylaspargic acid (NA-Asp), N-acetylglutamine (NA-Gln), N-acetylglycine (NA-Gly), N-acetylmethionine (NA-Met), N-acetylthreonine (NA-Thr), N-acetyltryptophan (NA-Trp), N-acetyltyrosine (NA-Tyr), N-acetylvaline (NA-Val), α -ketoglutaric acid (AKG), citric/isocitric acid (CITS), fumaric acid (FUM), lactic acid (LAC), malic acid (MAL), oxaloacetic acid (OXA), pyruvic acid (PYR), succinic acid (SUCC), acetylcarnitine (C2:0-carnitine), decanoylcarnitine (C10:0-carnitine), hexanoylcarnitine (C6:0-carnitine), lauroylcarnitine (C12:0-carnitine), myristoylcarnitine (C14:0-carnitine), octanoylcarnitine (C8:0-carnitine), palmitoylcarnitine (C16:0-carnitine), propionylcarnitine (C3:0-carnitine) and stearoylcarnitine (C18:0-carnitine) were solubilized in DMSO/DMF (1:1 v/v) and stored at -80 °C. Undecanoic acid (C11:0), dodecanoic acid (C12:0), octanoic acid (C8:0) and decanoic acid (C10:0) were solubilized at 1 mg/mL in ACN. The LC-MS grade ACN, DMSO and DMF were sourced from Actua-all Chemicals (Oss, The Netherlands). Dimethylaminophenacyl bromide (DmPABr) was procured from BioConnect BV (Huissen, The Netherlands) and DmPABr- $^{13}\text{C}_2$ was purchased from Nova Medical Testing (Alberta, Canada). In addition, the list of chemical identifiers (ChEBI IDs) can be found in supplementary Table S1.

HepG2 sample collection and preparation

The HepG2 cells were seeded and cultured at 37 °C under 5 % CO₂, harvested after 5 days and rinsed with PBS at 37 °C. The HepG2 cells were then separated into Eppendorf vials containing 2×10^6 cells per vial, and stored at -80 °C until sample preparation. Sample preparation consisted of reconstitution immediately in 1 mL of

water/methanol (1:4 v/v), followed by 5 minutes of sonication and vortexing to lyse the cells. The cells were then centrifuged at 13,000 rpm for 10 minutes at 4 °C to allow protein precipitation using an Eppendorf 5427R Centrifuge (Hamburg, Germany). The supernatant was transferred to a new Eppendorf vial and further diluted using water/methanol (1:1 v/v) to the equivalent cell contents of 1×10^5 , 5×10^4 , 2.5×10^4 , 1×10^4 , 5×10^3 , 2.5×10^3 and 1×10^3 , 500 and 250.

Derivatization of HepG2 cells

Triplicates of HepG2 cell supernatant containing the equivalent cell volume ranging from 250 to 1×10^5 were dried in a Labconco SpeedVac (MO, United States). Each dried sample was reconstituted immediately in 10 μ L of DMSO/DMF (1:1 v/v) to dissolve both polar and apolar metabolites, followed by the addition of 10 μ L triethanolamine (750 mM) and 20 μ L DmPABr (40 mg/mL). The content was then kept at 65 °C for 1 hour, followed by the addition of 10 μ L formic acid (30 mg/mL), and further 30 minutes at 65 °C to quench the reaction. After this, 5 μ L of DmPA- $^{13}\text{C}_2$ labelled metabolite internal standard and 45 μ L acetonitrile were added, bringing the total volume up to 100 μ L. The stability of DmPABr derivatized samples were demonstrated previously³⁰.

Chromatography conditions

The LC method conditions were detailed previously³⁴ with further adaptations. The method modifications focused on the retention times of the internal standards as the DmPA-D₆ was changed to DmPA- $^{13}\text{C}_2$ resulting in co-elution with each metabolite. The target metabolites were separated using a Waters Acquity UPLC Class II (Milford, USA) on an AccQ-tag C18 column [2.1 x 100 mm, 1.4 μ m (Milford, USA)] kept at 60 °C, using gradient elution at a flow rate of 0.7 mL/min. Mobile phase A consisted of water containing 10 mM ammonium formate and 0.1% formic acid, and mobile phase B was 100% acetonitrile. The gradient profile was as follows; starting at 0.2 % B; linear increase to 20 % B at 1.5 min, 50 % B at 4.0 min, 90 % B at 6.0 min, 99.8 % B at 10.0 min and maintained until 13.0 min, then back to start conditions at 13.1 min, equilibrating until 15.0 min. The flow of the first 1.2 minutes was diverted to waste to prevent the DMSO/DMF peak from entering the mass spectrometer. The autosampler was maintained at 10 °C, and the injection volume was 1 μ L.

Mass spectrometry and data generation

Chapter 3

An AB Sciex QTrap 6500 mass spectrometer (Framingham, USA) was operated in positive ionization mode to accommodate the tertiary amine introduced by the derivatization reagent. The MS parameters were set as follows: curtain gas - 30.0 psi; collision gas - medium; ionization voltage - 5500 V; temperature - 600°C; ion source gas 1 at 60.0 psi; ion source gas 2 at 50.0 psi.

MRM optimization was achieved per analyte by independently derivatizing each analyte and then conducting direct infusion in compound optimization analysis mode. The MRM channels were optimized for entrance potential, declustering potential and exit potential. For each analyte, a unique fragmentation pattern was favoured, and the most abundant product ion was selected to provide the optimal sensitivity. The full details of the DmPABr derivatized metabolites, MRM parameters, and MS conditions can be found in supplementary Table S2.

The data was integrated using the AB Sciex MultiQuant Workstation Quantitative Analysis for QTrap. Automatic integration was used where possible and with a manual visual inspection conducted to ensure reliable integration. The data was assessed using peak area ratios ($P_{\text{analyte}}/P_{\text{Internal standard}}$). For statistical analysis, Microsoft Excel and GraphPad Prism 8 (La Jolla, CA) was used. We assessed the method using four independently made matrix-free calibration lines. We conducted a calibration concentration starting at the same concentration listed in supplementary Table S3, and diluted by 2-fold until the LOD was reached for both methods. Additionally, we used the equivalent of 5×10^3 cells ($n = 3$) to assess the intraday variability of the method. All concentrations reported represent the intracellular concentration of the extracted HepG2 cells. The LOD and LOQ were calculated using the following equations using the ICH Q2 guidelines (σ = standard deviation of the lowest calibration point):

$$\text{LOD} = (3.3 * \sigma) / \text{slope}$$

$$\text{LOQ} = (10 * \sigma) / \text{slope}$$

Results & Discussion

Separation profile advantages of derivatization

Derivatization with DmPABr prior to RPLC offers advantages in the separation profile of the targeted metabolites, that otherwise may co-elute, or elute early

alongside some high-abundance compounds which may act as ion suppressors. Other analytical techniques, such as CE-MS, may exhibit similar separation issues during sensitive analysis of cells ²². In addition, compromised peak-shape is an issue that often arises during the separation of amino acids ³⁶ and organic acids on a HILIC column ³⁷. HILIC methods have been established that measure amino acids, organic acids ³⁸ and acylcarnitines ³⁹ with good peak-shape, yet they usually require longer acquisition time, and they do not offer universal coverage within one injection because both positive and negative ionisation mode are required.

Figure 3.2 presents the chromatograms of the different MRM channels for amino acids measured quantitatively in 5×10^3 HepG2 cell extracts. Following derivatization, the chromatogram shows ideal peak shape of amino acids that usually suffer from early elution and poor peak-shape on RPLC. Moreover, isomeric metabolites such as leucine and isoleucine can be baseline resolved (see supplementary figure S2). The peak width at half height measured for alanine, N-acetylaspartic acid, leucine and isoleucine was 1.071, 1.001, 0.943 and 0.909 seconds, respectively. This demonstrates that derivatization with DmPABr followed by RPLC can compete with CE in terms of peak width. However, sharp peaks also require suitable mass spectrometers to record a sufficient number of data points across a peak, using small scan times. Processing large batches of samples using small time windows can be challenging due to retention shifts. Fortunately, the retention time repeatability of this method was high for all metabolites, for example, the retention time relative standard deviation for alanine, N-acetylaspartic acid and myristoylcarnitine was 0.014 %, 0.016 %, 0.034 %, respectively, in three measurements of 1×10^4 HepG2 cells extracts along 22.5 hours. In comparison, separation techniques such as CE may experience migration time RSD between 2% and 3% ²². Using HILIC separation, the analysis of apolar metabolites and organic acids have posed challenges due to non-Gaussian peak shape. Fig 3.3 demonstrates how derivatization provides greater retention and improved peaks shapes for such problematic organic acids, including lactic acid (monocarboxylic acid), oxaloacetic acid (ketoacid) and succinic acid (dicarboxylic acid). In addition, aromatic amino acids and acylcarnitines also suffer from poor peak shape on HILIC, yet after derivatization they behave more favourably on RPLC. Another peak shape parameter assessed here was the asymmetry, which generally showed very good results. For example, in neat calibration solution, the asymmetry factor of phenylalanine, tryptophan, lactic acid, succinic acid,

Chapter 3

palmitoylcarnitine and steroylcarnitine was 0.91, 1.19, 0.98, 1.15, 1.06 and 1.10, respectively (additional asymmetry factors for neat standards are shown in supplementary Table S5).

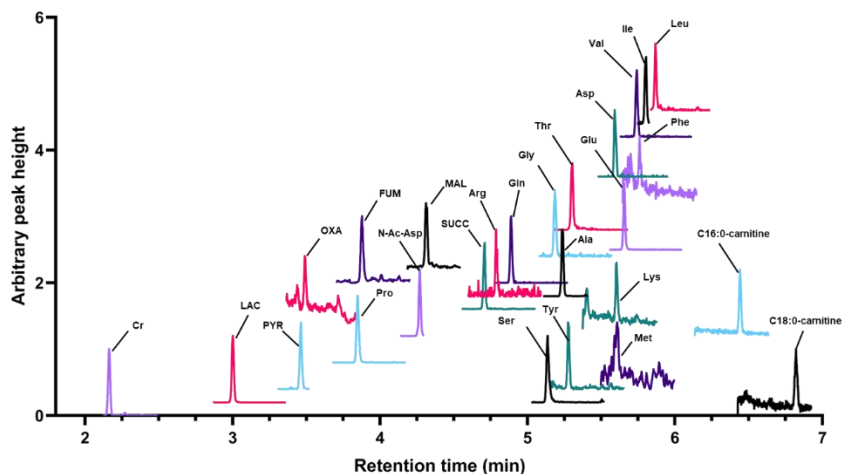


Fig 3.2. LC-MS/MS analysis of 5×10^3 HepG2 cells shown in multi-reaction monitoring (MRM) in positive ionisation mode after derivatization with DmPABr. Only the metabolites above the LOQ are shown in this chromatogram. The peak intensity of each signal was scaled to a uniform height and does not represent actual peak height.

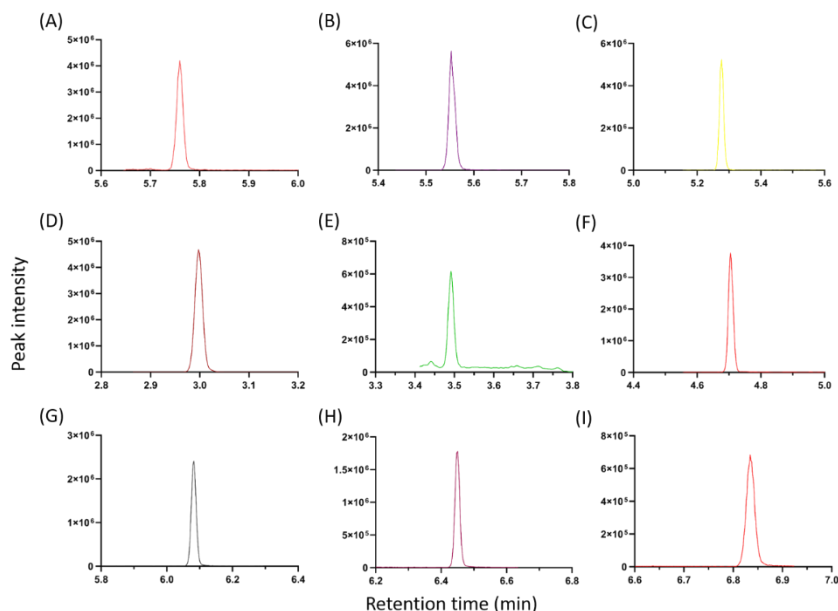


Fig 3.3. The extracted ion chromatogram of the derivatized metabolites measured in matrix-free solution showing the midpoint calibration concentration (supplementary Table S3). The metabolites shown are phenylalanine (A), tryptophan (B), tyrosine (C), lactic acid (D), oxaloacetic acid (E), succinic acid (F), palmitoylcarnitine (G), myristoylcarnitine (H) and steroylcarnitine (I). The specific MRM transitions are given in supplementary Table S2.

Method performance in matrix-free standard solutions

The general performance of the DmPABr method was already evaluated in previous work ³⁴. Here, we demonstrate the method suitability for metabolomics analysis of HepG2 cells, which requires tailored optimisation and modified calibration ranges. The performance parameters summarized in Table 3.1 are the linear range, coefficient of determination and repeatability expressed as the relative standard deviation of quadruplets of the middle calibration point. The linearity of the majority of metabolites in the low concentration range were above R^2 of 0.990, except for asparagine, glutamic acid, phenylalanine, tryptophan, N-Ac-methionine, oxaloacetic acid and dodecanoic acid, but they were deemed to be within an acceptable range for

Chapter 3

consideration. Interestingly, the metabolites with an $R^2 < 0.990$ belong to a range of chemical classes with a range of physiochemical properties, further supporting our earlier observation that the variability is not due to derivatization efficiency of specific functional groups³⁴. Furthermore, the lower concentration ranges (sub 250 nM) are prone to higher variability which may explain the compromise in linearity. Overall, the method exhibited good repeatability ($n = 4$) at the middle calibration point which represents an estimate of 5×10^3 HepG2 cells. The RSD was below 20 % for all metabolites measured in neat solutions by the method, providing consistent quantitative results. For example, metabolites such as alanine, lactic acid and lauroylcarnitine had an RSD of 2.6 %, 0.8 % and 1.2 %, respectively, demonstrating the low variability in different functional groups including primary amine, carboxylic acid and quaternary amine.

Metabolite	Linear range (nM)	R^2	RSD (%)	Metabolite	Linear range (nM)	R^2	RSD (%)
Alanine	70-1060	0.998	2.6	N-acetylthreonine	2-1250	0.999	14.7
Arginine	10-600	0.994	16.8	N-acetyltryptophan	5-450	0.997	13.3
Asparagine	60-530	0.988	3.7	N-acetyltyrosine	0.5-250	0.998	4.2
Aspartic acid	80-700	0.993	16.3	N-acetylvaline	5-1100	0.997	11
Cysteine	400-7000	0.995	14.8	α -Ketoglutaric acid	20-1400	0.994	8.3
Glutamine	50-1700	0.991	11.5	Citrates	450-30000	0.993	7.7
Glutamic acid	60-2350	0.984	5.1	Fumaric acid	60-1900	0.987	5.8
Glycine	500-60000	0.999	7	Lactic acid	500-20000	0.999	0.8
Histidine	900-14000	0.991	12.1	Malic acid	5-950	0.998	9
Isoleucine	10-250	0.995	9.6	Oxaloacetic acid	30-1900	0.989	5.7
Leucine	10-350	0.995	3.8	Pyruvic acid	20-1400	0.995	1.3
Lysine	500-7000	0.991	3.3	Succinic acid	50-1900	0.997	2.7
Methionine	30-950	0.998	6	Acetylcarnitine	30-1900	0.997	15.7
Phenylalanine	40-700	0.984	4.7	Decanoylcarnitine	1-1800	1	4.2
Proline	60-1900	0.999	3.7	Hexanoylcarnitine	10-1800	0.999	8.2
Serine	100-3500	0.99	10	Lauroylcarnitine	1-1800	0.999	1.2
Threonine	20-700	0.99	12.2	Myristoylcarnitine	5-1800	0.999	4.6
Tryptophan	40-700	0.985	10.1	Octanoylcarnitine	5-190	0.999	8.6

Metabolite	Linear range (nM)	R ²	RSD (%)	Metabolite	Linear range (nM)	R ²	RSD (%)
Tyrosine	60-950	0.994	3	Palmitoylcarnitine	5-1800	0.999	5.9
Valine	30-500	0.991	3.6	Propionylcarnitine	5-1800	0.999	1.9
N-acetylalanine	5-1250	0.999	2.4	Stearoylcarnitine	5-1800	1	3
N-acetylarginine	0.5-1100	1	11.5	Decanoic acid	5-1900	0.997	1.7
N-acetylaspartic acid	1-750	0.992	12.1	Octanoic acid	30-3700	0.997	5.1
N-acetylglutamine	5-1150	0.997	3.5	Dodecanoic acid	120-900	0.98	8.6
N-acetylglycine	1-600	0.998	15.1	Undecanoic Acid	5-750	0.994	6.5
N-acetylmethionine	2-1450	0.988	15.4	Creatinine	50-7000	1	7

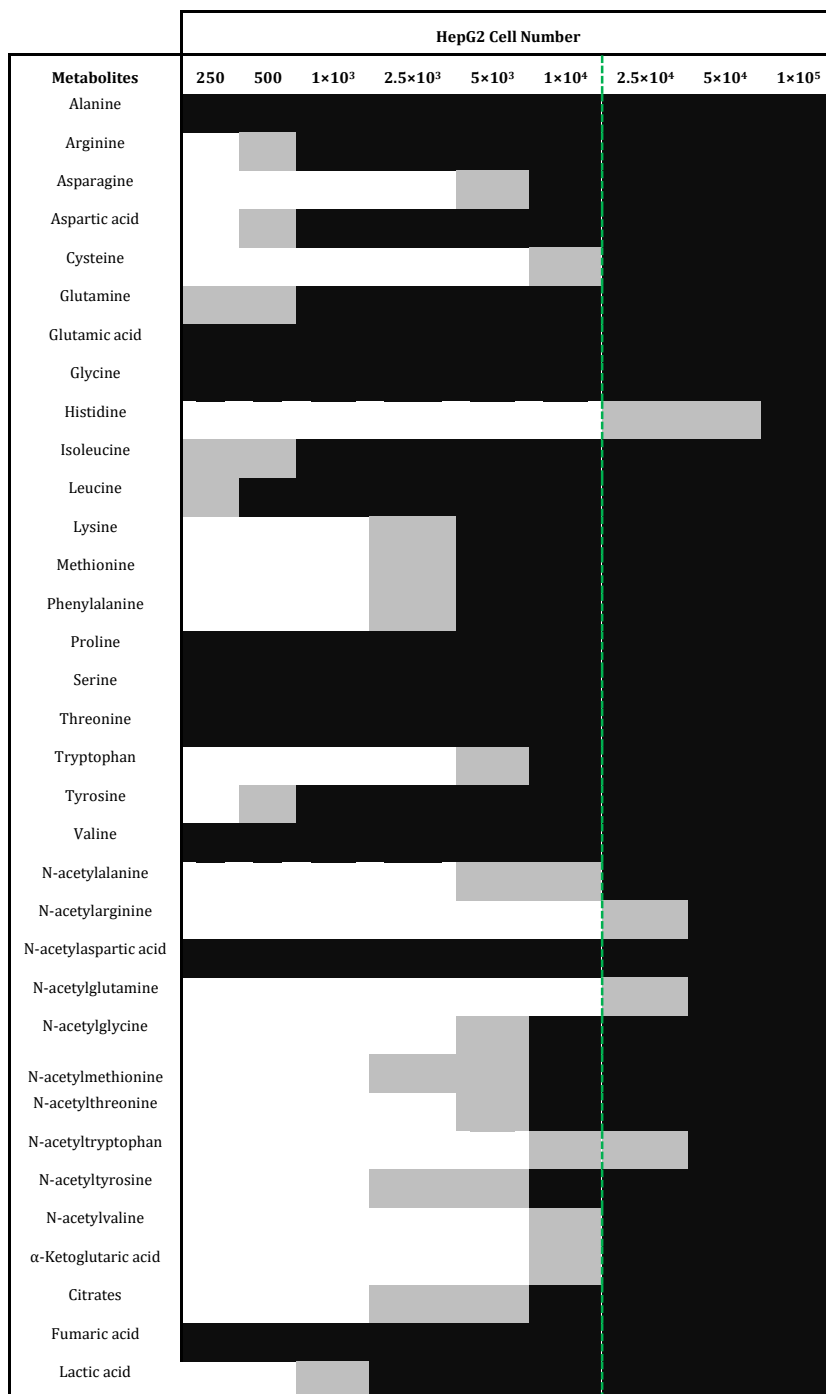
Table 3.1. Summary of the method performance showing the linear range, linearity and RSD of the method in neat solution. The RSD was assessed at the midpoint concentration of the low concentration calibration line.

Method performance across varying dilutions of HepG2 cells

To address the needs of microfluidics cells analysis, where good performance is required below 1×10^4 cells, the quantitative metabolic coverage was measured in cellular extracts equivalent to the cellular content of 250 to 1×10^5 HepG2 cells. Table 3.2 presents the metabolites that could be quantified and detected (below LLOQ) across a range of cell extract dilutions, ranging from 1×10^5 cells extract down to dilution containing 250 cells (equivalent to less than a cell loaded on the column). All of the amino acids, except histidine, were detected below 1×10^4 cells. Histidine is the only metabolite within this method that is double charged, making the metabolite more vulnerable to in-source fragmentation, thus reducing the sensitivity in limited MRM setup. Additionally, 13 amino acids were quantified in 1×10^3 HepG2 cells, and alanine, glutamic acid, glycine, proline, serine, threonine and valine were quantified in 250 cells. Unlike amino acids, the majority of N-acetylated amino acids exist in relatively low concentrations within the cells. Nevertheless, 8 out of 14 metabolites included in the method were successfully detected in the 1×10^4 cells extract, and the mitochondria active N-acetylaspartic acid could be detected in 250 HepG2 cells. N-acetylated amino acids can be found in high concentrations in the extracellular environment, which is an interesting direction to further investigate the applicability

of the current method in low cell numbers ⁴⁰. The 9 acylcarnitines targeted in this method were quantified in 5×10^4 cells, and four in 1×10^4 cells. Additionally, all acyl carnitines (except acetylcarnitine) could be detected in 1×10^4 cells. The method also covers organic acids and, as mentioned previously, the main strength of DmPA-labelling of organic acids is achieved by the addition of a tertiary amine, resulting in higher sensitivity despite susceptibility of the unlabelled metabolite to ion suppression in the ESI source ⁴¹. TCA cycle intermediates were detected with very good sensitivity. α -ketoglutarate, citrates, malic acid, oxaloacetic acid and pyruvic acid were detected in 1×10^4 cells, and fumaric acid, malic acid and pyruvic acid were further quantified in 250 cells. The quantified concentrations of these metabolites agrees with previously published data showing that within the TCA cycle, α -ketoglutarate and oxaloacetic acid are present at lower concentrations, hence are more challenging to quantify ³⁴. The quantitation of energy and central carbon-related metabolites can improve our understanding of the health and functionality of cells, and applying this to 3D microfluidic cells provides an accurate and true recording of the physiological environment.

Miniaturization using chemical derivatization



3

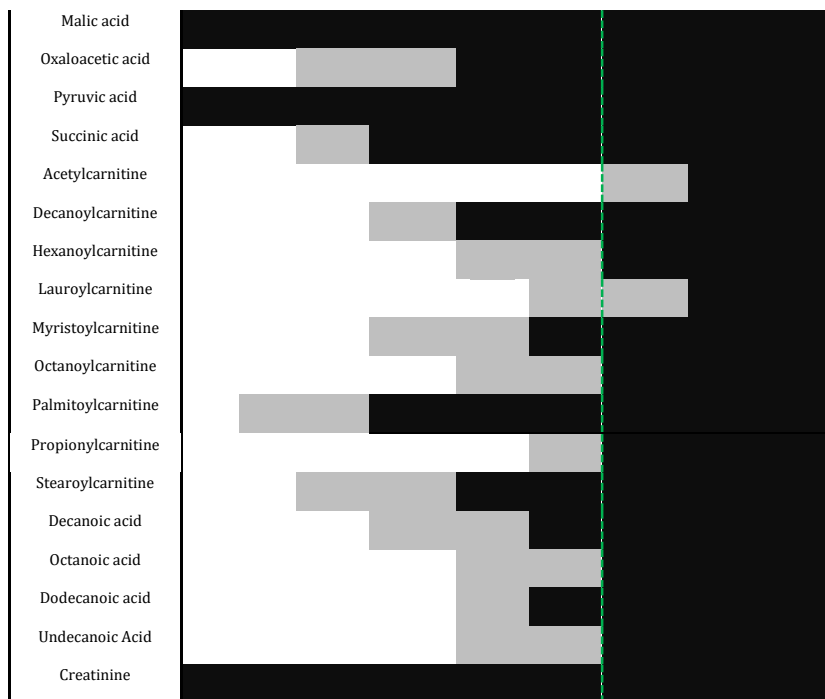


Table 3.2. Detection and quantitation of metabolites by the DmPABr derivatization method, applied to a range of HepG2 cell numbers ($250-1 \times 10^5$). The shaded cells represents the detection in that dilution of cells: Black, $>LLOQ$; Grey, $<LLOQ$ and $>LOD$; white, $<LOD$. The dotted green line shows the different cells number zones of microfluidic cell culture number (left of line) and macroscopic cell culture (right of line).

Table 3.2 demonstrates that the method can quantitatively analyze a range of metabolites with varying functional groups and physicochemical properties in the range of HepG2 cell counts. The linearity of calculated concentration along the range of cell dilutions is depicted for selected metabolites in Fig 3.4, and further detailed in supplementary Table S4. These plots visualise the applicability of DmPABr derivatization to microfluidic cell culture ranges which are $\text{sub-}1 \times 10^4$ cells. Generally, good linearity is observed throughout the range of cell dilutions, apart from specific cases where linearity was limited for the lower range of cell count. This behaviour is not unexpected due to solvation and ionisation efficiency and still aligns well the aim of the work. This effect is also observed in supplementary figure S1 that

shows the total ion chromatograms of the MRM channels recorded when the method was applied on 3 different cell dilutions.

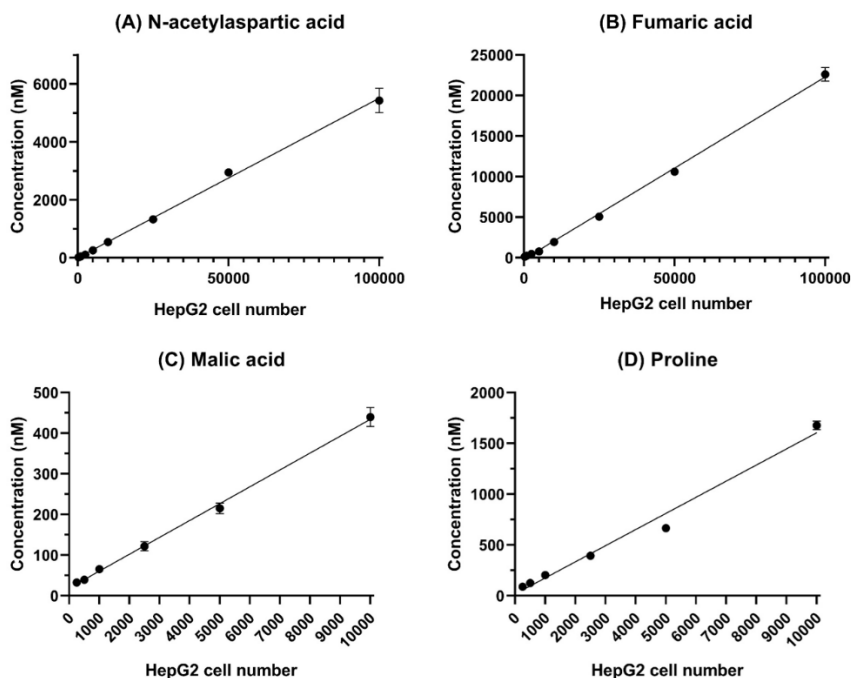


Fig 3.4. Quantification of selected metabolites in a range of cell counts. A) *N*-acetylaspartic acid and B) fumaric acid measured in 250 to 1×10^5 HepG2 cells ($n = 3$); C) malic acid and D) proline measured in 250 to 1×10^4 HepG2 cells ($n = 3$).

Quantitative results in 5000 HepG2 cell extract

The DmPABr LC-MS/MS method presented here was adapted from our previously published method ³⁴, by optimizing the MRM parameters and increasing ionisation voltages, to increase the sensitivity across the metabolites range. Table 3.3 summarizes the absolute quantitation of central carbon and energy-related metabolites in 5×10^3 HepG2 cell extracts (equivalent of 5 cells on column) by employing the ICD approach using DmPA- $^{13}\text{C}_2$ -labelled metabolites as a corresponding internal standard. Amino acids such as arginine and phenylalanine can be quantified despite low abundance (41.0 and 88.5 nM, respectively). The

mitochondrial abundant metabolite N-acetyl aspartate, which is associated with several diseases including Parkinson's disease, Canavan disease and Leigh's syndrome, was quantified at 261.6 nM. However, other N-acetylated amino acids, such as N-acetylglycine, N-acetylmethonine and N-acetylthreonine, were detected below the LLOQ. TCA cycle intermediates and pyruvic acid were also captured by the method at these cell number ranges, which could further support the study of energy metabolism within low cell numbers in a physiological-presenting environment using 3D microfluidic cell culture. After optimisation of the ionization voltage and collision energy from the previously published method, the LOD of metabolites such as serine improved from 506 nM to 23.4 nM, glycine from 932.4 to 25.7 nM, and N-acetylthreonine from 10.4 to 2.2 nM. Similarly, the LOD of α -ketoglutarate decreased from 29.7 down to 15.6 nM³⁴. The majority of late eluters showed the most sensitivity gain compared to early eluters, probably due to improved desolvation conditions owing to higher organic solvents, as discussed previously. The asymmetry factor during the measurement of 5×10^3 HepG2 cell extract for alanine, N-acetylaspatic acid, glutamine, leucine, isoleucine, succinic acid and malic acid was 0.96, 0.95, 0.96, 1.00, 1.13, 0.88 and 0.94, respectively, demonstrating a close-to-Gaussian profile, without a specific tendency for tailing or fronting. The variability of the 5×10^3 HepG2 cells measurements observed for almost all metabolites was well below RSD of 20%. Higher RSD values were recorded for decanoylcarnitine and hexanoylcarnitine (34.5 % and 66 %, respectively), probably due to the increased background noise (as discussed previously). Nonetheless, we chose to include and present this data to identify required improvements that may further increase the sensitivity and repeatability. These metabolites could warrant the use of MS³ which provides the ability to reduce background noise and increase sensitivity⁴².

Miniaturization using chemical derivatization

Metabolite	Conc. (nM)	LOD (nM)	Asymmetry factor	RSD (%)	Metabolite	Conc. (nM)	LOD (nM)	Asymmetry factor	RSD (%)
Alanine	666	14.9	0.96	2.1	N-acetylthreonine	<LOQ	2.2	0.86	19.1
Arginine	41	2.9	0.74	10.4	N-acetyltryptophan	ND	8.3	N/A	N/A
Asparagine	<LOQ	43.4	1.59	3.4	N-acetyltyrosine	<LOQ	0.2	1.91	14.8
Aspartic acid	595.4	57.4	1.33	11.9	N-acetylvaline	ND	0.7	N/A	N/A
Cysteine	ND	366.4	N/A	N/A	α -Ketoglutaric acid	ND	15.6	N/A	N/A
Glutamine	285.8	32.8	0.96	6.3	Citrates	<LOQ	181.4	4.02	3
Glutamic acid	2138.7	39.6	1.39	9.1	Fumaric acid	774.3	9.3	2.32	5
Glycine	1404.1	25.7	1.48	0.8	Lactic acid	662.5	70.5	0.90	4.9
Histidine	ND	803.7	N/A	N/A	Malic acid	215.2	5.5	0.94	6
Isoleucine	56.9	4.2	1.13	5.2	Oxaloacetic acid	82.5	10.3	1.14	2.6
Leucine	95	4.8	1.00	7.5	Pyruvic acid	309.3	11.5	0.92	3.9
Lysine	553	17.9	1.41	0.4	Succinic acid	253.2	21.7	0.88	2.5
Methionine	85.6	3.8	0.82	16.2	Acetylcarnitine	ND	21.1	ND	N/A
Phenylalanine	88.5	3.7	2.47	7.5	Decanoylcarnitine	4.1	0.7	3.08	34.5
Proline	666.6	7.4	1.32	3.6	Hexanoylcarnitine	<LOQ	4.2	1.80	66
Serine	704	23.4	2.06	2.6	Lauroylcarnitine	ND	1	ND	N/A
Threonine	348.8	11.7	1.72	5.1	Myristoylcarnitine	<LOQ	2.4	1.00	11.1
Tryptophan	<LOQ	26.7	0.63	1.2	Octanoylcarnitine	<LOQ	2.8	0.60	12.7
Tyrosine	114.3	9.7	0.86	5.6	Palmitoylcarnitine	4.9	0.2	0.79	14.9
Valine	108	12.4	0.93	6.1	Propionylcarnitine	ND	3.6	N/A	N/A
N-acetylalanine	<LOQ	6.8	1.01	5.2	Stearoylcarnitine	4.4	3.6	1.53	15.3
N-acetylarginine	ND	1.4	N/A	N/A	Decanoic acid	<LOQ	1.4	0.94	2.5
N-acetylaspatic acid	261.6	1.1	0.95	9	Octanoic acid	<LOQ	25	0.95	10.8
N-acetylglutamine	ND	3.9	N/A	N/A	Dodecanoic acid	<LOQ	80.7	1.14	10.6
N-acetyl glycine	<LOQ	2.1	0.64	14.5	Undecanoic Acid	<LOQ	3.8	1.11	6
N-acetylmethionine	<LOQ	0.9	1.83	12.3	Creatinine	2167.7	36.8	0.83	9.8

Table 3.3. Method performance of derivatized metabolites in the analysis of 5×10^3 HepG2 cells. Repeatability is expressed as %RSD of concentration for sample measured in triplicate intraday from different samples; <LOQ, the metabolite was detected but falls below the low limit of quantitation; ND, the metabolite was not detected in an extract from 5×10^3 cells; N/A, not applicable.

Sensitivity compared to commonly used methods

Several chromatographic techniques have been applied in the pursuit of sensitive metabolite analysis of volume-limited samples. The use of HILIC-MS is a common approach for measuring amino acids and organic acids from cell lysate. Liu *et al.*⁴³ quantified 107 metabolites in Huh-7 cells with the use of 10 internal standards in a 25-minute HILIC-MS/MS method. The method achieved amino acids LODs of 30 nM for phenylalanine (vs. 3.7 nM by DmPABr), 1000 nM for tryptophan (vs. 26.7 nM with DmPABr), 3000 nM for glycine (vs. 25.7 nM for DmPABr). Additionally, organic acids had a LODs of 330 nM for alpha-ketoglutarate (vs. 15.6 nM for DmPABr), 200 nM for succinic acid (vs. 21.7 nM for DmPABr), and 250 nM for malic acid (vs. 5.5 nM for DmPABr). This shows a significant increase in sensitivity compared to HILIC-MS/MS methods and a reduced analysis time. In a recent work by Zhang *et al.*²², sheathless CE-MS enabled the detection of amine-containing metabolites down to 500 HepG2 cell extracts. This method achieved LODs of 4.5 nM for alanine (vs. 14.9 nM by DmPaBr), 1.0 nM for glutamic acid (vs. 39.6 nM by DmPaBr), 5.7 nM for glutamine (vs. 32.8 nM by DmPaBr), 7.9 nM for tryptophan (vs. 26.7 nM by DmPaBr), and 2.9 nM for valine (vs. 12.4 nM by DmPaBr). This demonstrates that sheathless CE-MS is more sensitive to amino acids than DmPABr, however it requires an advanced separation technology that is less robust than RPLC, has less universal coverage of the metabolome, and the lack of internal standard coverage reduces quantitative performance. Additionally, it should be noted that different calculations were used to obtain the LOD, and the sheathless CE approach used signal-to-noise extrapolation. The sheathless CE-MS approach also struggles with the separation and sensitive detection of organic acids due to the lack of positively ionisable groups. This is another advantage that DmPA-labelling achieves by introducing a tertiary amine onto organic acids, thus enabling sensitive detection in positive ionisation mode (for example, malic acid and pyruvate at 5.5 nM and 11.5 nM LOD, respectively). GC-MS is another approach utilized to measure amino acids and organic acids from cell lysate, yet it can be compromised by lower sensitivity. The method applied by Danielsson *et al.*⁴⁴ provides varied metabolic coverage, but with minimal use of internal standards (seven). The few reported LOD values were 540 nM, 10 nM and 30 nM for serine, phenylalanine and succinic acid, respectively, compared to 23.4 nM, 3.7 nM and 21.7 nM detected using DmPABr labelling (which minimises internal standard cost by applying ICD). Luo, Li⁴⁵ used dansyl-labelling derivatization prior

to nanoLC-MS, and detected 1620 ± 148 metabolite peak pairs from the amine/phenome submetabolome. This method also uses the chemical isotope labelling approach, creating internal standards for each metabolite for qualitative investigation, unlike the use in our work that allows quantitative analysis.

Conclusion

The presented work demonstrates an approach for sensitive metabolomics analysis of a low-cell number sample. Chemical derivatization by DmPABr, followed by a LC-MS/MS targeted analysis, allowed absolute quantification of 37 metabolites in a diluted extract of 1×10^4 HepG2 cells (equivalent of 10 cells on column), 27 metabolites in a diluted extract of 5×10^3 HepG2 cells (equivalent of 5 cells on column), 18 metabolites in a diluted extract of 1×10^3 HepG2 cells (equivalent of 1 cell on column) and 12 metabolites in a diluted extract of 250 HepG2 cells (an equivalent of 0.25 cells on column). The method was evaluated using chemically diverse metabolites of high biological importance that were already implicated in several health conditions. Owing to the ability of the DmPABr reagent to label a broad selection of metabolites, the method can be further expanded to a wider selection of metabolites, matrices and applications, and further optimized for greater sensitivity. This aligns with the growing need for sensitive quantification of material-limited samples, and can be successfully achieved by combining with micro/nano-LC or CE coupled to nanoESI-MS/MS.

Acknowledgements

The author expresses thanks to Dr. Wei Zhang at Leiden University for culturing and providing the HepG2 cells. This project was supported by the SysMedPD project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no, 668738.

References

1. Thompson Legault J, Strittmatter L, Tardif J, et al. A Metabolic Signature of Mitochondrial Dysfunction Revealed through a Monogenic Form of Leigh Syndrome. *Cell Rep.* 2015;13(5):981-989.
2. Karu N, Wilson R, Hamede R, et al. Discovery of Biomarkers for Tasmanian Devil Cancer (DFTD) by Metabolic Profiling of Serum. *J Proteome Res.* 2016;15(10):3827-3840.
3. Huberty M, Martis B, Van Kampen J, et al. Soil Inoculation Alters Leaf Metabolic Profiles in Genetically Identical Plants. *Journal of Chemical Ecology.* 2020:1-11.
4. Jozefczuk S, Klie S, Catchpole G, et al. Metabolomic and transcriptomic stress response of Escherichia coli. *Molecular systems biology.* 2010;6(1):364.
5. Madsen R, Lundstedt T, Trygg J. Chemometrics in metabolomics—a review in human disease diagnosis. *Analytica chimica acta.* 2010;659(1-2):23-33.
6. Shah SH, Kraus WE, Newgard CB. Metabolomic profiling for the identification of novel biomarkers and mechanisms related to common cardiovascular diseases: form and function. *Circulation.* 2012;126(9):1110-1120.
7. Tatar Z, Migne C, Petera M, et al. Variations in the metabolome in response to disease activity of rheumatoid arthritis. *BMC musculoskeletal disorders.* 2016;17(1):353.
8. Van Esbroeck AC, Janssen AP, Cognetta AB, et al. Activity-based protein profiling reveals off-target proteins of the FAAH inhibitor BIA 10-2474. *Science.* 2017;356(6342):1084-1087.
9. Hadrévi J, Ghafouri B, Sjörs A, et al. Comparative metabolomics of muscle interstitium fluid in human trapezius myalgia: an in vivo microdialysis study. *European journal of applied physiology.* 2013;113(12):2977-2989.
10. Trushina E, Dutta T, Persson X-MT, Mielke MM, Petersen RC. Identification of altered metabolic pathways in plasma and CSF in mild cognitive impairment and Alzheimer's disease using metabolomics. *PLoS One.* 2013;8(5):e63644.

11. Moreno EL, Hachi S, Hemmer K, et al. Differentiation of neuroepithelial stem cells into functional dopaminergic neurons in 3D microfluidic cell culture. *Lab Chip*. 2015;15(11):2419-2428.
12. Choi WT, Tosun M, Jeong H-H, et al. Metabolomics of mammalian brain reveals regional differences. *BMC Systems Biology*. 2018;12(8):127.
13. Taylor RM, Miller PR, Ebrahimi P, Polsky R, Baca JT. Minimally-invasive, microneedle-array extraction of interstitial fluid for comprehensive biomedical applications: transcriptomics, proteomics, metabolomics, exosome research, and biomarker identification. *Laboratory animals*. 2018;52(5):526-530.
14. Wevers NR, Van Vught R, Wilschut KJ, et al. High-throughput compound evaluation on 3D networks of neurons and glia in a microfluidic platform. *Sci Rep*. 2016;6(1):1-10.
15. Beurivage C, Naumovska E, Chang YX, et al. Development of a gut-on-a-chip model for high throughput disease modeling and drug discovery. *International Journal of Molecular Sciences*. 2019;20(22):5661.
16. Kane KI, Moreno EL, Hachi S, et al. Automated microfluidic cell culture of stem cell derived dopaminergic neurons. *Sci Rep*. 2019;9(1):1-12.
17. Higashi T, Ogawa S. Isotope-coded ESI-enhancing derivatization reagents for differential analysis, quantification and profiling of metabolites in biological samples by LC/MS: A review. *J Pharm Biomed Anal*. 2016;130:181-193.
18. Ramautar R, Somsen GW, de Jong GJ. CE-MS for metabolomics: Developments and applications in the period 2016–2018. *Electrophoresis*. 2019;40(1):165-179.
19. Nassar AF, Wu T, Nassar SF, Wisnewski AV. UPLC–MS for metabolomics: a giant step forward in support of pharmaceutical research. *Drug discovery today*. 2017;22(2):463-470.
20. Yi X, Leung EKY, Bridgman R, Koo S, Yeo K-TJ. High-sensitivity micro LC-MS/MS assay for serum estradiol without derivatization. *Journal of Applied Laboratory Medicine*. 2016;1(1):14-24.
21. Kantae V, Ogino S, Noga M, et al. Quantitative profiling of endocannabinoids and related N-acylethanolamines in human CSF using nano LC-MS/MS. *J Lipid Res*. 2017;58(3):615-624.

22. Zhang W, Guled F, Hankemeier T, Ramautar R. Utility of sheathless capillary electrophoresis-mass spectrometry for metabolic profiling of limited sample amounts. *Journal of Chromatography B*. 2019;1105:10-14.
23. Hirayama A, Tomita M, Soga T. Sheathless capillary electrophoresis-mass spectrometry with a high-sensitivity porous sprayer for cationic metabolome analysis. *Analyst*. 2012;137(21):5026-5033.
24. Chetwynd AJ, David A. A review of nanoscale LC-ESI for metabolomics and its potential to enhance the metabolome coverage. *Talanta*. 2018;182:380-390.
25. Chetwynd AJ, David A, Hill EM, Abdul-Sada A. Evaluation of analytical performance and reliability of direct nanoLC-nanoESI-high resolution mass spectrometry for profiling the (xeno) metabolome. *Journal of Mass Spectrometry*. 2014;49(10):1063-1069.
26. Maciel EVS, de Toffoli AL, Sobieski E, Nazário CED, Lanças FM. Miniaturized liquid chromatography focusing on analytical columns and mass spectrometry: A review. *Analytica Chimica Acta*. 2020;1103:11-31.
27. Villas-Bôas SG, Smart KF, Sivakumaran S, Lane GA. Alkylation or silylation for analysis of amino and non-amino organic acids by GC-MS? *Metabolites*. 2011;1(1):3-20.
28. Wong JM, Malec PA, Mabrouk OS, Ro J, Dus M, Kennedy RT. Benzoyl chloride derivatization with liquid chromatography-mass spectrometry for targeted metabolomics of neurochemicals in biological samples. *J Chromatogr A*. 2016;1446:78-90.
29. Guo K, Li L. Differential ¹²C-/¹³C-isotope dansylation labeling and fast liquid chromatography/mass spectrometry for absolute and relative quantification of the metabolome. *Anal Chem*. 2009;81(10):3919-3932.
30. Guo K, Li L. High-Performance Isotope Labeling for Profiling Carboxylic Acid-Containing Metabolites in Biofluids by Mass Spectrometry. *Analytical Chemistry*. 2010;82(21):8789-8793.
31. Guo H, Jiao Y, Wang X, Lu T, Zhang Z, Xu F. Twins labeling-liquid chromatography/mass spectrometry based metabolomics for absolute quantification of tryptophan and its key metabolites. *J Chromatogr A*. 2017;1504:83-90.

32. Lkhagva A, Shen C-C, Leung Y-S, Tai H-C. Comparative study of five different amine-derivatization methods for metabolite analyses by liquid chromatography-tandem mass spectrometry. *Journal of Chromatography A*. 2020;1610:460536.
33. Song P, Mabrouk OS, Hershey ND, Kennedy RT. In vivo neurochemical monitoring using benzoyl chloride derivatization and liquid chromatography-mass spectrometry. *Anal Chem*. 2012;84(1):412-419.
34. Willacey CCW, Naaktgeboren M, Lucumi Moreno E, et al. LC-MS/MS analysis of the central energy and carbon metabolites in biological samples following derivatization by dimethylaminophenacyl bromide. *Journal of Chromatography A*. 2019:460413.
35. Gunda V, Yu F, Singh PK. Validation of Metabolic Alterations in Microscale Cell Culture Lysates Using Hydrophilic Interaction Liquid Chromatography (HILIC)-Tandem Mass Spectrometry-Based Metabolomics. *PLoS One*. 2016;11(4):e0154416.
36. Prinsen H, Schiebergen-Bronkhorst BGM, Roeleveld MW, et al. Rapid quantification of underivatized amino acids in plasma by hydrophilic interaction liquid chromatography (HILIC) coupled with tandem mass-spectrometry. *Journal of inherited metabolic disease*. 2016;39(5):651-660.
37. Heaton JC, McCalley DV. Some factors that can lead to poor peak shape in hydrophilic interaction chromatography, and possibilities for their remediation. *Journal of Chromatography A*. 2016;1427:37-44.
38. Ibáñez AB, Bauer S. Analytical method for the determination of organic acids in dilute acid pretreated biomass hydrolysate by liquid chromatography-time-of-flight mass spectrometry. *Biotechnol Biofuels*. 2014;7(1):145-145.
39. Peng M, Liu L, Jiang M, et al. Measurement of free carnitine and acylcarnitines in plasma by HILIC-ESI-MS/MS without derivatization. *Journal of Chromatography B*. 2013;932:12-18.
40. Mardinoglu A, Shoaie S, Bergentall M, et al. The gut microbiota modulates host amino acid and glutathione metabolism in mice. *Mol Syst Biol*. 2015;11(10):834.
41. Annesley TM. Ion suppression in mass spectrometry. *Clinical chemistry*. 2003;49(7):1041-1044.

Chapter 3

42. Quinete N, Bertram J, Reska M, Lang J, Kraus T. Highly selective and automated online SPE LC–MS3 method for determination of cortisol and cortisone in human hair as biomarker for stress related diseases. *Talanta*. 2015;134:310-316.
43. Liu Q, Cai J, Nichols RG, et al. A Quantitative HILIC–MS/MS Assay of the Metabolic Response of Huh-7 Cells Exposed to 2, 3, 7, 8-Tetrachlorodibenzo-p-Dioxin. *Metabolites*. 2019;9(6):118.
44. Danielsson AP, Moritz T, Mulder H, Spéjel P. Development and optimization of a metabolomic method for analysis of adherent cell cultures. *Analytical biochemistry*. 2010;404(1):30-39.
45. Luo X, Li L. Metabolomics of Small Numbers of Cells: Metabolomic Profiling of 100, 1000, and 10000 Human Breast Cancer Cells. *Analytical Chemistry*. 2017;89(21):11664-11671.

Supplementary information*Table S1. List of the ChEBI identifiers for the metabolites investigated in this methodology*

Metabolite	ChEBI ID	Metabolite	ChEBI ID
Alanine	16977	N-acetylthreonine	45826
Arginine	16467	N-acetyltryptophan	70976
Asparagine	17196	N-acetyltyrosine	21563
Aspartic acid	17053	N-acetylvaline	21565
Cysteine	17561	α -Ketoglutaric acid	30915
Glutamine	18050	Citrates	30769/30887
Glutamic acid	16015	Fumaric acid	18012
Glycine	15428	Lactic acid	28358
Histidine	15971	Malic acid	6650
Isoleucine	17191	Oxaloacetic acid	30744
Leucine	15603	Pyruvic acid	32816
Lysine	18019	Succinic acid	15741
Methionine	16643	Acetylcarnitine	57589
Phenylalanine	17295	Decanoylcarnitine	68830
Proline	17203	Hexanoylcarnitine	70749
Serine	17115	Lauroylcarnitine	77086
Threonine	16857	Myristoylcarnitine	84634
Tryptophan	16828	Octanoylcarnitine	73039
Tyrosine	17895	Palmitoylcarnitine	73067
Valine	16414	Propionylcarnitine	28867
N-acetylalanine	40992	Stearoylcarnitine	73074
N-acetylarginine	40521	Decanoic acid	30813
N-acetylaspartic acid	21547	Octanoic acid	28837
N-acetylglutamine	21553	Dodecanoic acid	30805
N-acetylglycine	40410	Undecanoic Acid	32368
N-acetylmethionine	21557	Creatinine	16737

Chapter 3

Table S2. Retention time and MRM parameters for the measurement of the metabolites covered within this method. The ICD generated internal standards are noted by the addition of -IS to the metabolite name.

Metabolite	Precursor	Product	RT	CE
Alanine	573.2	366.2	5.23	20
Arginine	658.2	319.2	4.71	35
Asparagine	616.2	339.2	4.94	30
Aspartic acid	778.2	392.2	5.6	30
Cysteine	766.2	134.1	5.6	25
Glutamine	630.2	340.2	4.91	20
Glutamic acid	792.2	585.2	5.68	35
Glycine	559.2	134.1	5.2	30
Histidine	400.83	134.1	4.98	40
Isoleucine	615.2	408.2	5.83	30
Leucine	615.2	408.2	5.87	30
Lysine	476.6	134.1	5.5	30
Methionine	633.2	426.2	5.62	25
Phenylalanine	649.2	442.2	5.77	30
Proline	438.1	289.1	3.8	25
Serine	589.2	408.2	5.15	25
Threonine	603.2	422.2	5.31	25
Tryptophan	688.2	340.2	5.56	30
Tyrosine	665.2	458.2	5.28	30
Valine	601.2	394.2	5.75	30
N-acetylalanine	293.13	180	2.99	15
N-acetyl glycine	279.11	180	2.74	15
N-acetylvaline	321.19	180	3.68	15
N-acetyltryptophan	408.27	180	4.07	15
N-acetyltyrosine	385.23	180	3.41	15
N-acetyl asparagine	498.14	180	4.26	30
N-acetylarginine	378.24	180	2.35	30
N-acetylthreonine	323.16	180	2.82	20
N-acetylmethionine	353.25	180	3.65	15

Metabolite	Precursor	Product	RT	CE
N-acetylglutamine	350.18	84	2.51	25
Pyruvic acid	250.05	134	3.55	30
α -Ketoglutaric acid	469.2	134	4.79	30
Malic acid	457.07	134	4.31	35
Lactic acid	252.07	180	2.99	30
Citric/isocitric acid	676.101	180	5.1	20
Succinic acid	441.07	134	4.68	20
Fumaric acid	439.06	134	3.8	30
Oxaloacetic acid	455.06	134	3.46	35
C2:0-carnitine	365.11	134	2.64	20
C8:0-carnitine	449.04	134	4.6	35
C3:0-carnitine	379.21	134	2.94	35
C16:0-carnitine	561.6	134	6.26	35
C18:0-carnitine	589.7	134	6.55	35
C14:0-carnitine	533.6	134	5.92	35
C6:0-carnitine	421.4	134	3.95	35
C10:0-carnitine	477.5	134	5.11	35
C12:0-carnitine	505.5	134	5.54	35
Octanoic acid	306.21	180	5.65	20
Decanoic acid	334.26	180	6.08	20
Dodecanoic acid	362.32	180	6.44	20
Undecanoic acid	348.29	180	6.27	20
Creatinine	275.12	134.1	2.12	40
IS-Alanine	579.2	370.2	5.23	25
IS-Arginine	664.2	321.2	4.71	35
IS-Asparagine	622.2	343.2	4.94	30
IS-Aspartic acid	786.2	396.2	5.6	30
IS-Cysteine	774.2	136.1	5.6	25
IS-Glutamine	636.2	344.2	4.91	30
IS-Glutamic acid	800.2	591.2	5.68	35
IS-Glycine	565.2	136.1	5.2	30
IS-Isoleucine	621.2	412.2	5.83	30
IS-Leucine	621.2	412.2	5.87	30

Metabolite	Precursor	Product	RT	CE
IS-Lysine	481.6	136.1	5.5	30
IS-Methionine	639.2	430.2	5.62	25
IS-Phenylalanine	655.2	446.2	5.77	30
IS-Proline	442.1	291.1	3.8	25
IS-Serine	595.5	412.2	5.15	25
IS-Threonine	609.2	426.2	5.31	25
IS-Tryptophan	694.2	344.2	5.56	30
IS-Tyrosine	671.2	462.2	5.28	30
IS-Valine	607.2	398.2	5.75	30
IS-Histidine	645.2	136.1	4.98	40
IS-N-acetylalanine	295.1	182	2.99	15
IS-N-acetyl glycine	281.1	182	2.74	15
IS-N-acetylvaline	323.2	182	3.68	15
IS-N-acetyltryptophan	410.3	182	4.07	15
IS-N-acetyltyrosine	387.2	182	3.41	15
IS-N-acetylaspartic acid	502.1	182	4.26	30
IS-N-acetylarginine	380.2	182	2.35	30
IS-N-acetylthreonine	325.2	182	2.82	20
IS-N-acetylmethionine	355.3	182	3.65	15
IS-N-acetylglutamine	352.2	84	2.54	25
IS-Pyruvic acid	252.1	136	3.55	30
IS-Alpha-Ketoglutaric acid	473.2	136	4.79	30
IS-Malic acid	461.1	136	4.31	35
IS-Lactic acid	254.1	182	2.99	30
IS-Citric/Isocitric acid	682.1	182	5.1	40
IS-Succinic acid	445.1	136	4.68	30
IS-Fumaric acid	443.1	136	3.8	30
IS-Oxaloacetic acid	459.1	136	3.46	35
IS-C2:0-carnitine	367.1	136	2.64	35
IS-C8:0-carnitine	451	136	4.6	35
IS-C3:0-carnitine	381.2	136	2.94	35
IS-Octanoic acid	308.2	182	5.56	20
IS-Decanoic acid	336.2	182	6.08	20

Miniaturization using chemical derivatization

Metabolite	Precursor	Product	RT	CE
IS-Dodecanoic acid	364.3	182	6.44	20
IS-Undecanoic acid	350.3	182	6.27	20
IS-Creatinine	277.1	136	2.12	40
IS-C16:0-carnitine	563.6	136	6.26	35
IS-C18:0-carnitine	592.7	136	6.55	35
IS-C14:0-carnitine	535.6	136	5.92	35
IS-C6:0-carnitine	423.4	136	3.95	35
IS-C10:0-carnitine	479.5	136	5.11	35
IS-C12:0-carnitine	507.5	136	5.54	35

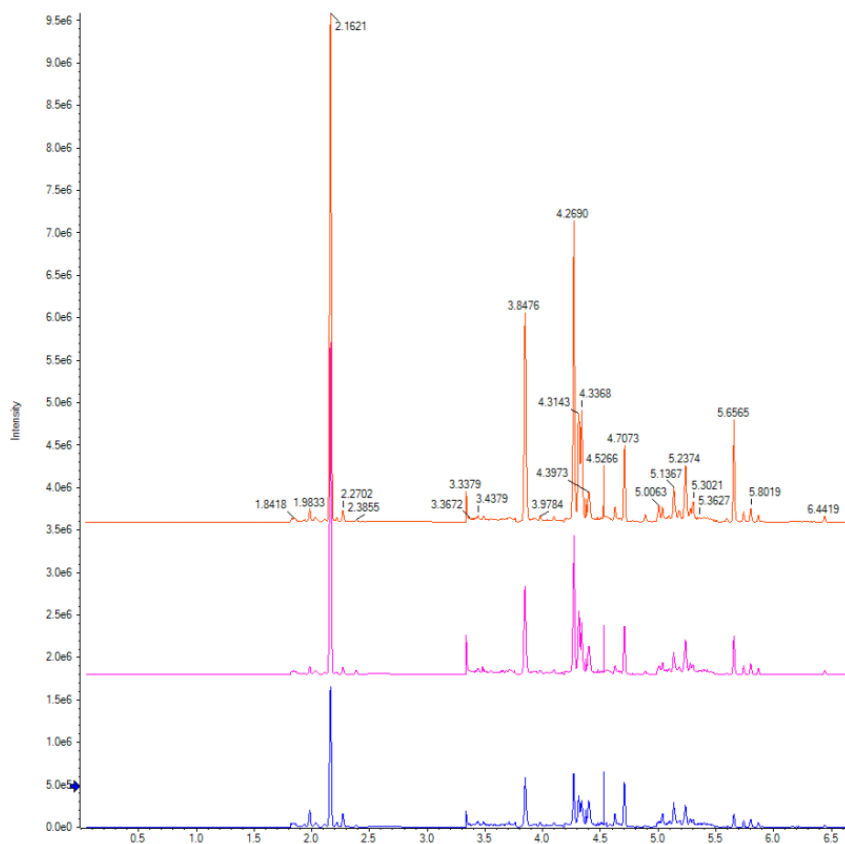
Chapter 3

Table S3. Calibration stock concentration of the metabolites measured within this method

Metabolite	Stock concentration (μM)	Metabolite	Stock concentration (μM)
Alanine	594.29	N-acetylthreonine	19.71
Arginine	74.29	N-acetyltryptophan	3.49
Asparagine	267.43	N-acetyltyrosine	1.79
Aspartic acid	148.57	N-acetylvaline	8.76
Cysteine	891.43	α-Ketoglutaric acid	148.57
Glutamine	891.43	Citrates	3714.29
Glutamic acid	59.43	Fumaric acid	14.86
Glycine	4457.14	Lactic acid	297.14
Histidine	1782.86	Malic acid	29.71
Isoleucine	59.43	Oxaloacetic acid	29.71
Leucine	89.14	Pyruvic acid	32.4
Lysine	445.71	Succinic acid	118.86
Methionine	29.71	Acetylcarnitine	59.43
Phenylalanine	178.29	Decanoylcarnitine	29.71
Proline	14.86	Hexanoylcarnitine	14.29
Serine	891.43	Lauroylcarnitine	14.29
Threonine	356.57	Myristoylcarnitine	14.29
Tryptophan	178.29	Octanoylcarnitine	1.49
Tyrosine	237.71	Palmitoylcarnitine	14.29
Valine	118.86	Propionylcarnitine	14.29
N-acetylalanine	19.71	Stearoylcarnitine	14.29
N-acetylarginine	17.11	Decanoic acid	29.71
N-acetylaspatic acid	47.56	Octanoic acid	29.71
N-acetylglutamine	32.4	Dodecanoic acid	0.89
N-acetylglycine	9.3	Undecanoic Acid	2.97
N-acetylmethionine	92.43	Creatinine	1857.14

Miniaturization using chemical derivatization

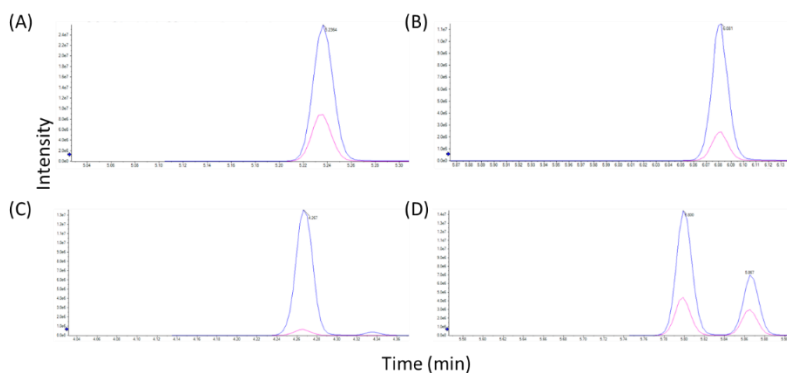
Fig S1. LC-MS/MS Total Ion Count chromatograms of a range of dilutions of HepG2 cells, derivatized by DmPABr. Top trace, 1×10^4 cells; middle trace, 5×10^3 cells; bottom trace, 2.5×10^3 cells.



3

Chapter 3

Fig S2. Extracted ion chromatogram from a neat standard solution showing the co-elution of the analytes (blue) with the internal standard (pink) using the isotope-coded derivatization approach. A) alanine; B) myristoylcarnitine; C) N-acetylated aspartic acid; D) isoleucine and leucine (left to right).



Miniaturization using chemical derivatization

Table S4. Summary of the metabolites cell coverage range and linearity of the cell concentrations across a range of dilution

Metabolite	Cell number range	R ²	Metabolite	Cell number range	R ²
Alanine	250-1e5	0.985	N-acetylthreonine	1e4-1e5	0.999
Arginine	1e3-1e5	0.991	N-acetyltryptophan	5e4-1e5	N/A
Asparagine	1e4-1e5	0.99	N-acetyltyrosine	1e4-1e5	0.99
Aspartic acid	1e3-1e5	0.998	N-acetylvaline	2.5e4-1e5	0.981
Cysteine	2.5e4-1e5	0.984	α -Ketoglutaric acid	2.5e4-1e5	0.982
Glutamine	1e3-1e5	0.978	Citrates	1e4-1e5	0.995
Glutamic acid	250-1e5	0.99	Fumaric acid	250-1e5	0.999
Glycine	250-1e5	0.985	Lactic acid	2.5e3-1e5	0.995
Histidine	1e5	N/A	Malic acid	250-1e5	1
Isoleucine	1e3-1e5	0.978	Oxaloacetic acid	5e3-1e5	0.989
Leucine	500-1e5	0.975	Pyruvic acid	250-1e5	0.985
Lysine	5e3-1e5	0.976	Succinic acid	2.5e3-1e5	0.995
Methionine	5e3-1e5	0.979	Acetylcarnitine	5e4-1e5	N/A
Phenylalanine	5e3-1e5	0.976	Decanoylcarnitine	5e3-1e5	0.997
Proline	250-1e5	1	Hexanoylcarnitine	2.5e4-1e5	0.992
Serine	250-1e5	0.986	Lauroylcarnitine	5e4-1e5	N/A
Threonine	250-1e5	0.982	Myristoylcarnitine	1e4-1e5	0.998
Tryptophan	1e4-1e5	0.979	Octanoylcarnitine	2.5e4-1e5	0.981
Tyrosine	1e3-1e5	0.98	Palmitoylcarnitine	2.5e3-1e5	0.997
Valine	250-1e5	0.978	Propionylcarnitine	2.5e4-1e5	0.986
N-acetylalanine	2.5e4-1e5	0.997	Stearoylcarnitine	2.5e3-1e5	0.997
N-acetylgarginine	5e4-1e5	N/A	Decanoic acid	1e4-1e5	0.987
N-acetylaspatic acid	250-1e5	0.998	Octanoic acid	2.5e4-1e5	0.986
N-acetylglutamine	5e4-1e5	1	Dodecanoic acid	1e4-1e5	0.992
N-acetylglucose	1e4-1e5	0.973	Undecanoic Acid	2.5e4-1e5	0.999
N-acetylmethionine	1e4-1e5	0.997	Creatinine	250-1e5	0.997

Chapter 3

Table S5. Summary of the metabolite asymmetry factors from the measurement of neat calibration standard at the midpoint concentration.

Metabolite	Asymmetry factor	Metabolite	Asymmetry factor
Alanine	1.11	N-acetylthreonine	1.17
Arginine	1.05	N-acetyltryptophan	0.86
Asparagine	0.97	N-acetyltyrosine	1.17
Aspartic acid	1.22	N-acetylvaline	1.08
Cysteine	0.90	α -Ketoglutaric acid	0.96
Glutamine	1.15	Citrates	1.92
Glutamic acid	1.31	Fumaric acid	1.59
Glycine	1.14	Lactic acid	0.98
Histidine	1.19	Malic acid	1.14
Isoleucine	1.04	Oxaloacetic acid	1.28
Leucine	0.97	Pyruvic acid	0.97
Lysine	1.01	Succinic acid	1.15
Methionine	1.01	Acetylcarnitine	1.11
Phenylalanine	0.91	Decanoylcarnitine	1.25
Proline	1.14	Hexanoylcarnitine	0.98
Serine	1.89	Lauroylcarnitine	0.88
Threonine	0.93	Myristoylcarnitine	0.99
Tryptophan	1.19	Octanoylcarnitine	1.27
Tyrosine	1.31	Palmitoylcarnitine	1.06
Valine	1.01	Propionylcarnitine	1.09
N-acetylalanine	1.05	Stearoylcarnitine	1.10
N-acetylarginine	1.04	Decanoic acid	1.18
N-acetylaspatic acid	1.15	Octanoic acid	1.03
N-acetylglutamine	1.12	Dodecanoic acid	0.97
N-acetylglycine	1.26	Undecanoic Acid	0.86
N-acetylmethionine	1.08	Creatinine	0.92

Chapter 4

A quantitative atlas of metabolites across regions of the rat brain

Cornelius C W Willacey, Maria E Secci, Marcus W Meinhardt, Tom Schilperoord,
Isabelle Kohler, Amy C Harms, Naama Karu[#], Wolfgang H Sommer[#] and Thomas
Hankemeier[#]

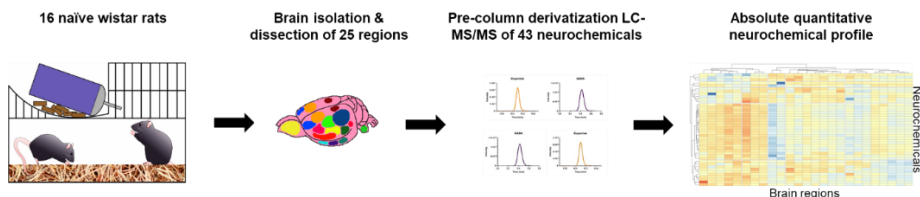
Submitted

[#] Authors equally contributed to the manuscript

Abstract

The mammalian brain is an extremely complex organ, comprising of a multitude of regions with varying functionalities as well as genome, transcriptome, proteome and metabolome profiles. Despite the expanding research of the brain and the advancing analytical technology, quantitative metabolic mapping of the brain is rather limited. Here, we present a brain quantitative metabolic atlas of the healthy adult rat (n=16), a preferred animal model for human brain research. The atlas provides absolute quantitative values for 43 metabolites that are of biochemical importance in the brain or are associated with CNS health conditions. They represent several chemical classes, and include amino acids and neuroactive derivatives such as major neurotransmitters, polyamines, and antioxidants. We used a pre-column benzoyl chloride derivatization followed by UPLC-MS/MS analysis to accurately and sensitively measure the regions. While the current metabolic coverage is a demonstrator for the applicability of the analytical method, it can be further modified to include more metabolites of interest, taking advantage of the simple stabilisation of metabolites that are vulnerable to degradation. The quantified metabolites across 25 brain regions reflected the specialised function of several regions, showing high agreement between distinct neural composition and neurotransmitter abundance within the cells. As a proof of concept for biochemical interpretation, significant patterns were also highlighted along the metabolic pathways of tyrosine, tryptophan, urea and polyamine production. The metabolic atlas reference dataset can be further utilised as quantitative reference levels to compare to new studies and identify perturbations in relevant pathways or diseases. The dataset can also be integrated into genome-scale metabolic models, to further define neuronal networks and the connectome.

Synopsis



To provide metabolic reference values of the brain that can contribute to our understanding of the cellular homeostasis network and neural communication, neurochemicals were analysed in 25 brain regions from 16 healthy adult Wistar rats. To achieve this, we performed liquid-liquid extraction followed by pre-column derivatization with benzoyl chloride with UPLC-MS/MS.

- Absolute quantitation on 43 neurochemicals was performed.
- Key metabolic pathways were covered such as the tyrosine, polyamine, urea cycle and tryptophan metabolism.
- A distinct metabolic profile was identified between regions of the brain.
- The neurochemicals quantified are associated with a range of neurological disorders.
- Reference values have been provided for clinical use and metabolic model integration using a systems approach.

Background

Metabolomics has the potential to decipher the contributing factors towards diseases of the CNS ¹⁻³ via the study of homeostasis, cell signalling, oxidative stress and communication ⁴. The mammalian brain is the most complex organ, and it operates by an intricate, interconnected and synergistic network via chemical neurotransmission with localised regions designated for specific and specialised roles. In turn, some regions have a more distinct genome, transcriptome, proteome and metabolome and higher expression of specific neurons. With recent advancements in metabolomics-based technologies, it is possible to further explore the molecular phenotype of each brain region. Leading databases such as KEGG ⁵, BioCyc ⁶, and VMH ⁷, relate metabolic pathways to the genome, transcriptome and proteome, to understand cellular function, mechanism of diseases and therapeutic target sites ⁸. Although, there has been continuous discussion about limited knowledge on the metabolic composition of the mammalian brain ^{9,10} followed by a noticeable absence of quantitative metabolomics data ^{11,4}. The majority of quantitative reference data is curated for blood, cerebrospinal fluid, and urine ^{12,13}. The lack of organ-specific reference data hinders the next step in research, particularly when attempting to combine data to constraint-based models. For obvious reasons, healthy human brain metabolomics work is not possible. Nevertheless, even in murine models, several of the approaches to capture the brain metabolome have focused on specific diseases or measured limited regions with minimal sensitivity for important neurotransmitters ^{9,14,15}. Quantitative measurement of amino acids has provided useful information but this alone does not detail the communication profile, i.e., neurotransmitters. Furthermore, this approach was also limited to the analysis of the prefrontal cortex, striatum, hippocampus and cerebellum ¹⁶. A more practical approach to explore the brain utilises measurements of the cerebrospinal fluid ¹⁷⁻¹⁹. This has been important in the detection and understanding of key metabolites associated with human diseases, but is only indicative of the brain metabolome as a whole. One of the studies that addresses this limitation comes from Choi et al. (2018)²⁰ who mapped the mouse brain using quantitative analysis of amino acids, lipids, peptides, nucleotides, etc. to investigate the four brain regions; frontal cortex, hippocampus, cerebellum, and olfactory bulb. Moreover, global metabolomics has been utilised to map eight mouse-derived brain

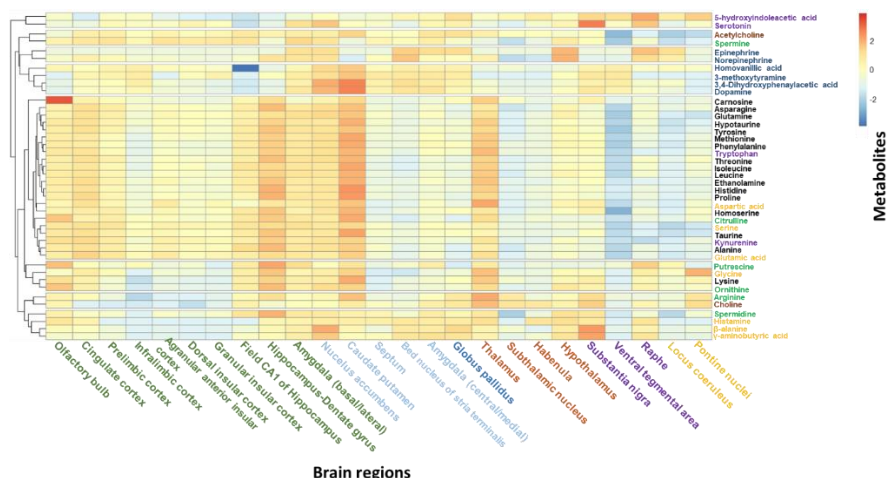
regions ⁴ without reporting absolute quantitative concentrations, hampering the use of this information in mechanistic models.

In this study, we developed and optimized a method for the analysis of neurometabolites in rat brain tissue based on sample derivatization with benzoyl chloride ²¹. Due to the relatively high instability of neurotransmitters, specifically the catecholamines in the extracellular environment, ²², samples were derivatized to stabilise the metabolites while increasing the overall analytical sensitivity and chromatographic performance.

The developed method has led to the most comprehensive, quantitative metabolic neurochemical profile of the rat brain to date in terms of brain regions measured (n = 25), number of animals (n = 16) and metabolites absolutely quantified (n = 43). The concentration profile of 43 neurochemicals was conducted across the cerebral cortex, striatum, diencephalon, midbrain and pons. Furthermore, we focused on neurochemicals which are of interest in the research of CNS activity and diseases. We also documented the essential regions associated with the key brain pathways for control of emotion, behaviour, memory and movement, such as mesolimbic, limbic and nigrostriatal. Our atlas lays the groundwork for quantitative data integration into metabolic models which can vastly improve understanding of the functionality and pathophysiology of the mammalian brain.

Results

Metabolic profile of healthy rat brain



4 Figure 4.1: Neurochemical metabolome of healthy rodent brains. The column represents each individual brain region and the row represents the metabolites quantified. Due to the large number of brain regions, the median concentration values between the 16 samples are shown. The brain regions have been arranged by their bregma coordinates from the frontal lobe to the brain stem. The data has been clustered based on their metabolite concentration similarity using hierarchical clustering with complete linkage. The brain regions are coloured by their associated hierarchy: green, cortex; light blue, striatum; blue, pallidum; orange, diencephalon; purple, midbrain; yellow, pons. The figure is scaled using z-score across the metabolites: red represents a high concentration of the metabolite and blue represents a low concentration of metabolites. The metabolites are coloured by their associated metabolic pathways: purple, tryptophan metabolism; brown, cholinergic metabolism; blue, tyrosine metabolism; yellow, neuroactive amines; green, urea cycle and polyamine metabolism; black, biogenic amines.

Figure 4.1 provides an overview of median metabolite concentrations across the brain regions (n=25), via a heatmap of all quality-approved compounds (n=43) in the rat samples (n=16). The metabolite concentration values are reported for each rat in Tables S8-12. Hierarchical clustering with complete linkage created seven main clusters of closely-associated metabolites: serotonin and its turnover product 5-

HIAA; acetylcholine and spermine; norepinephrine and its turnover product epinephrine; dopamine metabolism into 3-MT, DOPAC and HVA; amino acids and derivatives (possibly driven by concentration similarity); ornithine and its derivative putrescine, together with glycine and lysine; arginine and choline; histamine, spermidine, β -alanine and GABA (all neuroactive). The brain regions in the presented heatmap are arranged by parent regions based on physical proximity. When hierarchical clustering of the brain regions was performed according to the metabolic profile (Fig. S2), some of the formed clusters corresponded with the parent regions (e.g. insular regions as part of the cortex). The clustering of brain regions is also demonstrated in a unsupervised analysis using principal component analysis (Fig. S1).

Neurotransmitter abundance across the brain

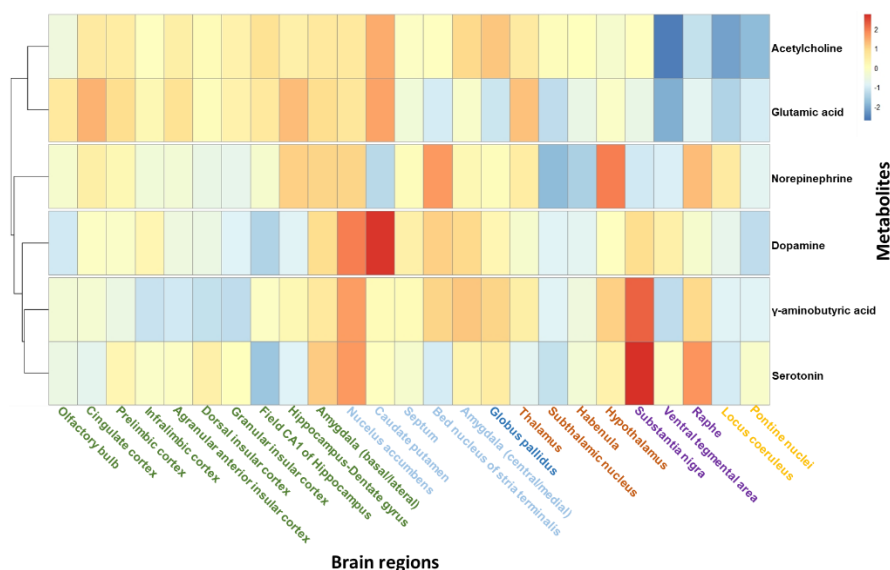


Figure 4.2: Heat map showing the relative concentrations of acetylcholine, glutamate, norepinephrine, serotonin, dopamine and GABA across all 25 brains regions. The heat map has been arranged by the distance from the bregma coordinates and clustered by metabolite concentration similarity. The figure is scaled using z-score to identify the movement from the mean across the metabolites: red represents a high concentration of the metabolites and blue represents a low concentration of metabolites. The brain

regions are coloured by their associated hierarchy: green, cortex; light blue, striatum; blue, pallidum; orange, diencephalon; purple, midbrain; yellow, pons.

Figure 4.2 presents the heatmap obtained for the neurotransmitters acetylcholine, glutamate, norepinephrine, serotonin, dopamine and GABA. The concentrations of each metabolite per region are detailed in Tables S8-12. The brain regions were ordered using the distance from the bregma coordinates, starting with the positive distances on the left, to negative distances on the right. A distinct pattern was identified in the concentration of neurotransmitters across the rat brain. For example, regions within the brainstem have a low concentration of acetylcholine and glutamate, while a higher concentration of serotonin were measured in the raphe. Serotonin was more abundant in the substantia nigra (mean value 157.3 ± 77.6 ng/g wet brain tissue) compared to dopamine (mean value 145.2 ± 43.6 ng/g wet brain tissue). The serotonin/dopamine ratio had been documented and discussed before by Cragg, Hawkey, Greenfield ²³ using fast-scan cyclic voltammetry with carbon-fibre microelectrodes. The substantia nigra receives serotonin innervations from the dorsal and medial raphe nuclei, providing input to dopaminergic dendrites. Interestingly, the frontal cortex has higher concentrations of acetylcholine and glutamate, while the olfactory bulb has a lower concentration in acetylcholine. The concentrations of dopamine and GABA are high across the midbrain which is consistent with the known distribution of dopaminergic neurons and GABAergic neuron projections within this area. Overall, the areas with the highest concentration of dopamine are the nucleus accumbens, caudate putamen and substantia nigra. The caudate putamen and substantia nigra are involved in the nigrostriatal pathway, which has a role in the regulation of movement where dopaminergic neurons are heavily involved. The nucleus accumbens, which is involved in the mesolimbic pathway, is another dopaminergic neuron projection dense region.

Comparison of brain regions along metabolic pathways

To demonstrate the applicability of the metabolic atlas in biochemical research, metabolite differences between selected brain regions were investigated along important metabolic pathways that are captured in the method. These brain regions are interlinked in their functionality, such as the regulation of movement (caudate putamen and substantia nigra) or are involved in similar health conditions, such as anxiety and post-traumatic stress disorder (amygdala, hypothalamus and infralimbic

cortex). The possibilities of utilising metabolic concentrations in biochemical interpretation are further expanded in the discussion.

Chapter 4

Tyrosine metabolism

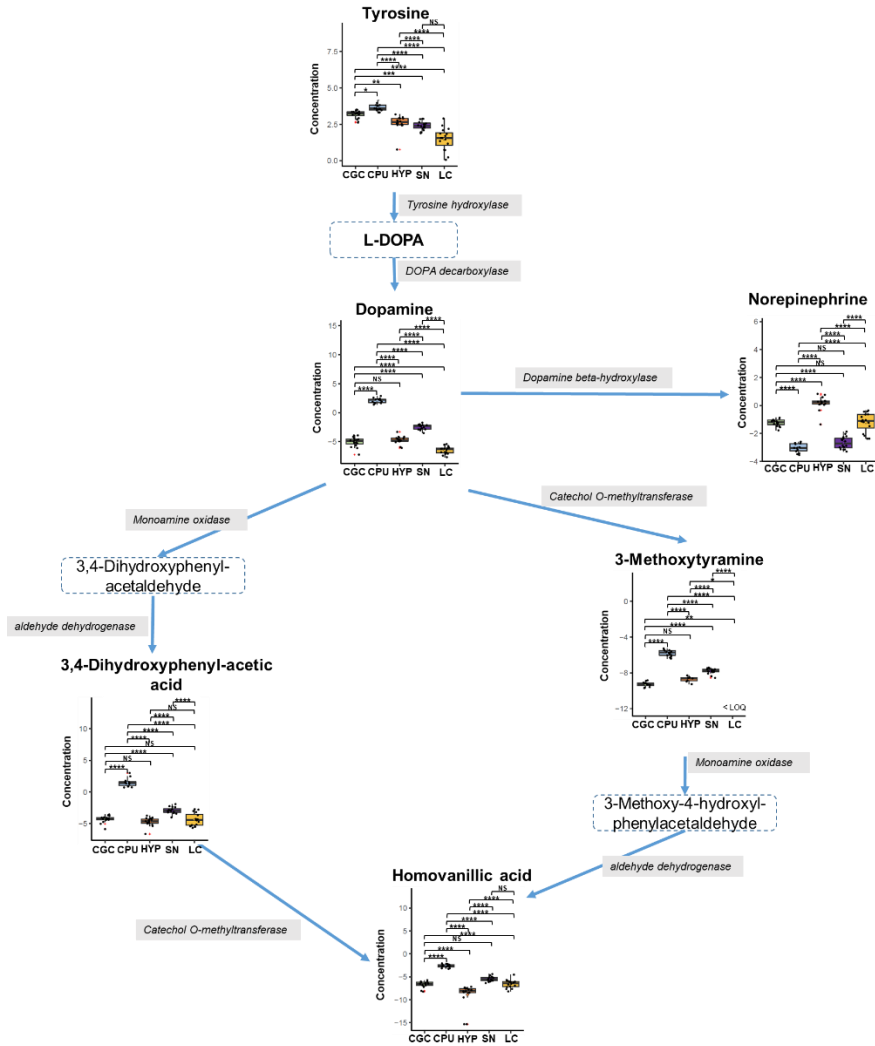
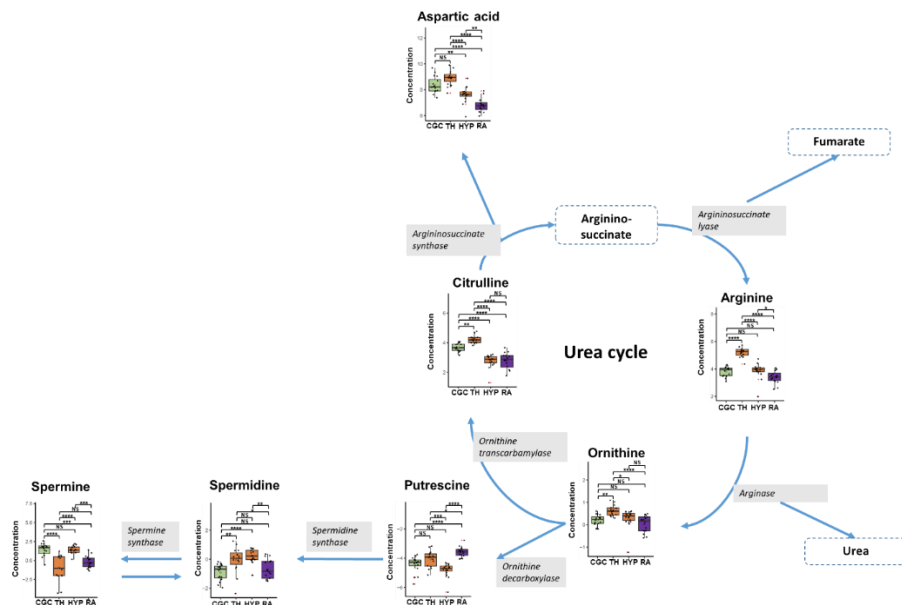


Figure 4.3: Metabolite concentrations along the tyrosine pathway across five brain regions: cingulate cortex (CGC), caudate putamen (CPU), hypothalamus (HYP), substantia nigra (SN) and locus coeruleus (LC). Metabolite concentrations are expressed as log2 transformed values (ng/g wet brain tissue). The five brain regions were compared using one-way ANOVA followed by Tukey's HSD post-hoc test. Significance levels were ns = $q > 0.05$, * = $q \leq 0.05$, ** = $q \leq 0.01$, *** = $q \leq 0.001$ and **** = $q \leq 0.0001$.

Seven metabolites in the tyrosine pathway were selected for this method, including the precursor tyrosine, its catecholamine neurotransmitter products, and a few of their degradation products as shown in Figure 4.3. Dramatic concentration differences were recorded for all metabolites between many of the selected brain regions. Tyrosine concentration was significantly different between most of the compared brain regions, decreasing from the caudate putamen through the hypothalamus, substantia nigra and, finally, the locus coeruleus. The concentration of dopamine was significantly higher in the caudate putamen and substantia nigra compared with the hypothalamus and locus coeruleus. The dopamine metabolites 3-MT, DOPAC and their metabolite HVA repeated the same pattern of differences between the brain regions. Moreover, the ratios describing the conversion into each product were reversed to the precursors in terms of relation between brain regions, apart from HVA/3MT which resembled the HVA. The above pattern recorded for 3-MT, DOPAC and HVA was the opposite to that of dopamine metabolites norepinephrine and epinephrine, showing higher metabolite levels in the hypothalamus and locus coeruleus, compared to the caudate putamen and substantia nigra. This suggests lower expression and activity of the enzyme dopamine β -hydroxylase, which metabolises dopamine to norepinephrine within the two regions. The turnover from dopamine to norepinephrine as depicted in a boxplot demonstrated not only a better within-region similarity, but also highlights the region with the highest turnover (LC).

*Figure 4.4. Metabolite concentrations along the urea cycle and polyamine metabolism across four brain regions: (CGC), thalamus (TH), hypothalamus (HYP) and raphe (RA). Metabolite concentrations are expressed as log2 transformed values (ng/g wet brain tissue). The four brain regions were compared using one-way ANOVA followed by Tukey's HSD post-hoc test. Significance levels were ns = $q > 0.05$, * = $q \leq 0.05$, ** = $q \leq 0.01$, *** = $q \leq 0.001$ and **** = $q \leq 0.0001$.*

124



spermine. This difference can be explained by the reversible enzymatic conversion between spermine and spermidine.

Chapter 4

Metabolic ratios

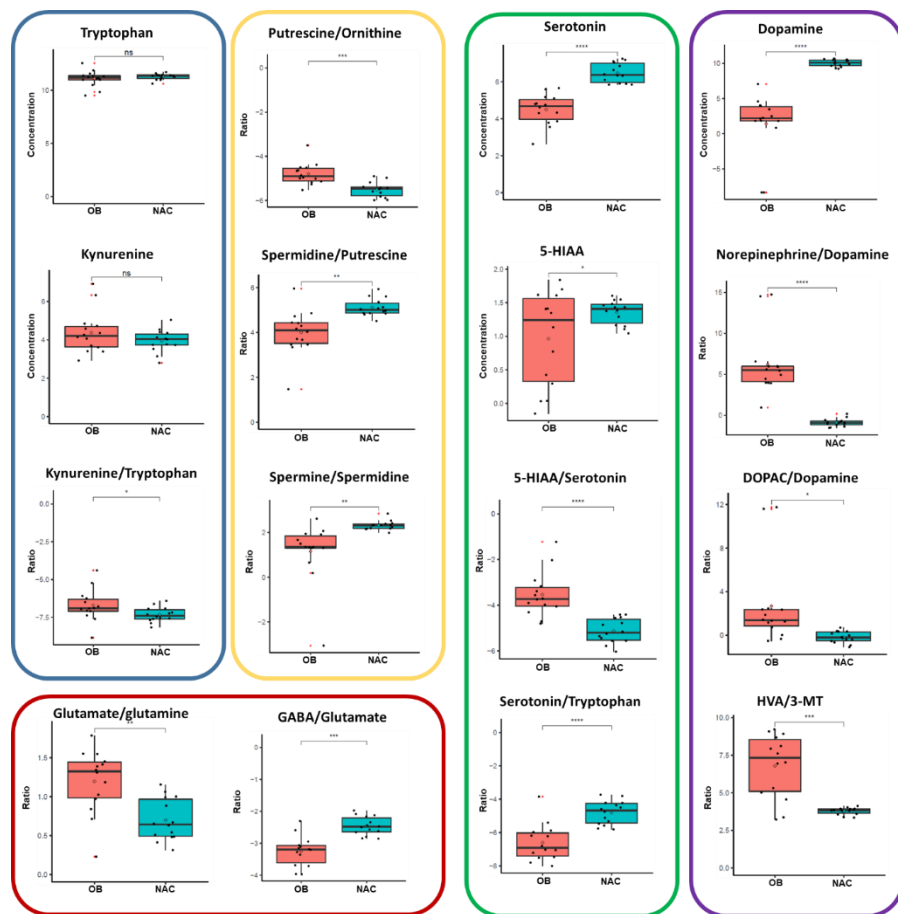


Figure 4.5. Metabolic turnover ratios of metabolites captured in the olfactory bulb (OB) and nucleus accumbens (NAC). Metabolite concentrations are expressed as log₂ transformed values (ng/g wet brain tissue). The data has been statistically compared using a paired t-test with Benjamini-Hochberg FDR correction. Significance levels were ns = $q > 0.05$, * = $q \leq 0.05$, ** = $q \leq 0.01$, *** = $q \leq 0.001$ and **** = $q \leq 0.0001$. The boxes indicate the metabolic pathways as follows: blue; tryptophan metabolism, yellow; polyamine metabolism, green; serotonin metabolism, purple; dopamine metabolism and red; GABA metabolism.

An advantage of quantifying metabolites along the same pathway is the ability to estimate an enzymatic conversion between a precursor and a product even without

isotopically-labelled flux analysis. Of course, it should be taken into account that what might appear as conversion of a free amino acid into another might merely represent post-translational modification of an amino acid residue of a protein, followed by proteolysis. Therefore, we concentrated only on well-established direct conversions, such as the turnover of serotonin to 5HIAA, dopamine to norepinephrine and putrescine to spermidine. Based on the local expression of the converting enzymes, and the neuronal composition of different brain regions, we hypothesise that unique turnover ratios will be identified. These conversion rates must be calculated per animal and not from the mean of a group. Paired *t*-test with Benjamini-Hochberg FDR correction was calculated on the olfactory bulb and nucleus accumbens, as shown in Figure 4.5. Tryptophan and kynurenine are not significantly different, however when we evaluate the turnover ratio kynurenine/tryptophan, we see that there is a significant difference ($q \leq 0.05$). This is a simple demonstration of how a ratio can portray significant differences that are not observed by the study of a single metabolite. Metabolic turnovers can also be used to reduce variation as seen with 5-HIAA in the olfactory bulb. When we look at the turnover ratio of 5-HIAA/serotonin, we have a reduced variation and greater significance ($q \leq 0.0001$). The exit from the urea cycle to the polyamine pathway is shown with the ratio ornithine/putrescine, which is significantly greater in the olfactory bulb than in the nucleus accumbens ($q \leq 0.001$). The polyamine turnover can then be explored using the ratios spermidine/putrescine and spermine/spermidine; both turnovers are lower in the olfactory bulb. Using the ratio GABA/glutamate, we can correlate this to the presence of glutamatergic or GABAergic neurons. We can also gather information regarding GABA synthesis by identifying the glutamate/glutamine ratio. Furthermore, this approach can be applied to the norepinephrine/dopamine ratio, which can be correlated to noradrenergic and dopaminergic neurons. The concentration of serotonin is lower in the nucleus accumbens compared to the olfactory bulb ($q \leq 0.0001$). However, the turnover of serotonin to 5-HIAA by the enzyme monoamine oxidase is higher in the olfactory bulb ($q \leq 0.0001$). In addition, this enzyme activity of monoamine oxidase can also be observed in the turnover of dopamine to DOPAC ($q \leq 0.05$) and 3-MT to homovanillic acid ($q \leq 0.001$).

Discussion

The presented study generated a quantitative neurochemical atlas of adult rat brain using an absolute quantitative methodology. Here, we objectively explore various brain regions ranging from the frontal cortex to the brainstem, of which each has specific roles, to allow exploration of regulatory functions and further associations with neurological diseases and psychiatric disorders.

As discussed earlier, the quantitation of neurochemicals is specifically challenging due to their broad range of physicochemical properties, chemical instability of neurotransmitters, and low concentrations requiring higher sensitivity. With the use of advanced techniques, as with the demonstrated pre-column derivatization followed by UHPLC-MS/MS analysis, it is possible to achieve accurate and reliable quantitation of the metabolic composition of each brain region. There were previous attempts to measure the healthy mammalian brain, although with limited metabolic coverage or brain regional coverage, and without addressing the stability issues of neurotransmitters. For example, Chen et al. (2016)²⁴ investigated the metabolic profile of the whole brains from six 4-week-old C56BL6 mice, which is of limited value for understanding the diversity and complexity across the brain. Nevertheless, these results can be somewhat compared to our metabolic atlas by using the mean values across brain regions. Using all of the comparable metabolites, we saw a similar order of magnitude concentrations are observed for amino acids, which are essential components in all cells (Asn 7429 ± 68 vs. 5718 ± 3120 ; Gln $165,427 \pm 10647$ vs. $217,651 \pm 108,235$; Tyr 3966 ± 278 vs. 5714 ± 2969 , all in ng metabolite per g wet tissue), and dopamine (137.52 ± 12.42 vs. 214.62 ± 681.87 ng/g wet tissue). After taking the mean values of the metabolite concentrations as quantified from each of the 25 brain regions, our standard deviations were larger due to the changes in metabolite abundance across the brain as seen in Figure 4.1. Kaplan et al. (2013)²⁵ studied specific metabolites across the healthy rat brain, profiling on broader parts, such as the entire prefrontal cortex, without separating it into distinct anatomical regions as we report in the metabolic atlas. Therefore, only direct comparison of parent regions can be conducted and this yields dissimilar values, yet the metabolic abundance profile shows likeness in a few cases (for example, DOPAC is reported at 84.1 ng/g wet tissue at the prefrontal cortex, and mean of the regions in the atlas is 52.2 ng/g wet tissue). The comparison of literature values to the metabolic atlas

reference concentrations is challenging due to the factors listed above. In addition, it is clear that research of such a complex and diverse organ as the brain, requires that metabolites are quantified on a more localised and region-specific basis. Brain regions have their own distinct biological function, such as the cingulate cortex role in endocrine function, emotional learning and motivation ²⁶, and the granular insular cortex role in visceral sensory function ²⁷. Moreover, different brain regions are associated with different diseases, such as the cingulate cortex association with schizophrenia and depression ²⁸, and the infralimbic cortex association with anxiety and post-traumatic stress disorder ²⁹.

The metabolic atlas provides distinct neurochemical profiles of brain regions, with some similarity between regions that share neighbouring anatomical location and neuronal pathways (for example, in regions of the cortex; agranular, dorsal and granular insular cortexes). It is clear that metabolic pathways, such as the tyrosine metabolism, urea cycle and polyamine metabolism, are differentially expressed. A clear evidence is the high concentrations of dopamine in brain regions that are known to have a high proportion of dopaminergic innervation. These regions include the caudate putamen and substantia nigra, which are part of the nigrostriatal pathway involved in the regulation of movement and associated with diseases such as Parkinson's ^{23,30,31}. The metabolic atlas can also provide biochemical insights via metabolic correlations that are associated with human health and disease. Two amino acids, β -alanine and arginine, positively correlate in most regions with neurotransmitters, GABA and choline, respectively. The relationship between β -alanine and GABA is well documented and has been shown in the cerebellum, specifically in Purkinje cells ³², and across the majority of the brain regions ²⁰. The metabolic atlas also showed that β -alanine correlated with dopamine in the nucleus accumbens, consistent with the report by Ericson et al (2010) ³³. A distinct dipeptide is carnosine, which measured high levels in the olfactory bulb, also aligns with Margolis et al (1974)³⁴ and Sharma et al (2015)³⁵ who reported that carnosine synthase I is expressed at a high level in the olfactory bulb. Interestingly, metabolic pathways such as the polyamine metabolism can also be used as biomarkers for brain trauma and stroke ^{36,37}. Although polyamine metabolism has a crucial role in cellular homeostasis and ROS scavenging, the full mechanism is still not totally understood

³⁸.

Chapter 4

The metabolic atlas can also be utilised to crudely outline the type of neurons that are present in the different regions of the brain, for example, the presence of acetylcholine and epinephrine prove the existence of cholinergic and adrenergic neurons. Furthermore, this can be correlated with genomics ³⁹, transcriptomics ⁴⁰ and proteomics ³⁵ data that have been mapped in murine atlases. Various types of cells exist across brain regions, including glial cells and neurons. The presented metabolic atlas includes representative neurotransmitters that are produced or affected by the different neurons: acetylcholine (cholinergic neurons), dopamine (dopaminergic), norepinephrine (noradrenergic), GABA (GABAergic), glutamate (glutamatergic) and serotonin (serotonergic neurons). Brain regions and neuron cell types are associated with different CNS diseases and psychological conditions, including Alzheimer's disease, Attention Deficit Hyperactivity Disorder, addiction, depression, Huntington's disease and Parkinson's disease ⁴¹⁻⁴⁴. These illnesses can be related to specific regions and neuron cell types or distributed broadly across the brain. Therefore, the breadth of coverage of the metabolic atlas enables integration into systems biology-based models, such as genome-scale constraint-based models, to further understand the connectivity and function of the brain and, in turn, diseases ^{45,46}.

4

Within the two pathways highlighted in our work, the tyrosine metabolism and urea cycle combined with polyamine metabolism demonstrate the ability to categorise regions based on presumed neuron abundance. For example, higher dopamine concentration in comparison to norepinephrine and epinephrine concentrations, suggests a region that is rich in dopaminergic neurons that express the enzyme tyrosine hydroxylase but lack the enzyme dopamine β -hydroxylase, which converts dopamine into norepinephrine ²². However, adrenergic neurons contain both enzymes, producing dopamine as well as norepinephrine and epinephrine. There are higher levels of dopaminergic neurons and dopamine in the midbrain, controlling functionalities such as movement and emotional regulation (motivation, impulsiveness and pleasure). Lower concentration of dopamine were recorded in the hypothalamus (Fig 4.3), which governs hormonal regulation. In the regions of the Pons, such as locus coeruleus, we see a low concentration of dopamine and a high concentration of norepinephrine. This coincides with the locus coeruleus being one of the main point of origin for noradrenergic neurons ⁴⁷. Within the supporting information, we also see that regions mainly associated with emotional regulation,

such as the raphe, which is governed by serotonergic neurons, show the same low abundance of dopamine.⁴⁸ Here, we also see the second highest concentration of serotonin across all brain regions after the substantia nigra (shown in Tables S8-12). The above examples demonstrate the application of neurotransmitter profiling to characterise the brain connectome. Similarly, mapping of metabolites along the urea cycle and polyamine metabolism can assist in understanding the biochemistry of the brain, as perturbations in these metabolites were associated with neurological disorders: stroke³⁶, Huntington's disease^{49,50}, and Alzheimer's disease⁵¹.

Interestingly, by gauging the metabolic ratios between certain metabolites within this method, we can start to form correlations to pathway activation and enzyme function. Our findings show that a single metabolite does not always identify significant differences between regions in subtle pathways, like the tryptophan metabolism where the metabolite has more than one function; however, we can identify differences when we explore the metabolic ratio. The metabolic ratios can evaluate precursor metabolites involved in the synthesis of neurotransmitters such as GABA. GABA can be synthesised from two pathways, the polyamine degradation pathway^{52,53} and the GABA shunt⁵⁴. GABA is formed from both spermidine and putrescine via the enzymes diamine oxidase and polyamine oxidase, respectively, through the intermediate 4-aminobutyraldehyde⁵². At the same time, it is produced by glutamate via the enzyme glutamate decarboxylase⁵⁴. This highlights the importance of understanding the urea cycle and polyamine metabolism when exploring the functions and presence of neurotransmitters. This is also seen with the formation of glutamate from glutamine or the TCA cycle. To expand understanding of the GABA shunt, the method by Wallacey et al (2019)⁵⁵ can be used in the future as this quantitatively captures the TCA cycle. In the nucleus accumbens, we see a higher concentration of GABA and, in turn, we see a higher turnover of glutamate to GABA. Metabolic ratios can be predicted using metabolites that share the same enzyme as shown with the metabolites that share the enzyme monoamine oxidase.

The evaluation of neuronal health and communication can be accelerated by metabolic models that integrate omics data. Moreover, the study of the intracellular metabolic content and the extracellular environment would provide further information relating to the connectivity of brain regions. The metabolic atlas of the brain regions includes the concentrations of important metabolites, reflecting only

the intracellular metabolic content without distinguishing between synaptically transmitted neurotransmitters. A clear limitation of this approach is the absence of indication whether the presence of intracellular neurotransmitters correlates with extracellular synaptic release. Nevertheless, the high abundance of neurotransmitters in the brain regions suggests high likelihood that it is involved in neurotransmission. There are metabolomics approaches that can circumvent this issue of neurotransmission evaluation by measuring microdialysate ^{56,57} or CSF ⁵⁸. These approaches have their own limitations, in the form of low concentrations that requires high instrumental sensitivity, and the chemical instability of the targets under the analytical conditions with adequate derivatization of the vulnerable functional groups.

Quantitatively mapping the metabolic concentrations across the mammalian brain provides the scientific community with important reference data to support the interpretation of homeostasis and neuronal communication. Furthermore, the reference data can help in the identification of perturbations in the metabolic profile of murine models as a response to various CNS conditions. Of course, this needs to be further adapted and translated into human models, as the physiology differs. Analytically, the neurochemical reference values can assist in benchmarking method performance, and encouraging integration into genome-scale metabolic models. Utilising metabolomics in a quantitative fashion, which has been described as “the final piece in the omics puzzle”, enables exploration of the brain architecture and function throughout these dynamic networks ^{4,8,59}.

Materials and Methods

Chemicals

All chemicals were purchased from Sigma-Aldrich (St. Louis, USA) unless stated otherwise. Table S4 shows the ChEBI IDs of all targeted metabolites. The LC-MS grade ACN was sourced from Actu-all Chemicals (Oss, The Netherlands) and de-ionised water was produced using a Merck Milli-pore A10 purification system (Raleigh, USA).

Standard solutions

Stock solutions of 3-methoxytyramine, acetylcholine, aspartic acid, β -alanine, choline, dopamine, epinephrine, γ -aminobutyric acid, histamine, homoserine,

homovanillic acid, kynurenine, 3,4-dihydroxyphenylacetic acid, ornithine, putrescine, serotonin (5-HT), spermidine, spermine, tryptophan and tyrosine were prepared using 1 mg in 1 mL of de-ionised water, vortexed and stored at -80 °C. Similarly, stock solutions of 2 mg/mL were made for 5-HIAA, arginine, asparagine, ethanolamine, leucine and norepinephrine; 5 mg/mL for carnosine, citrulline, glutamic acid, isoleucine, methionine, phenylalanine, proline, serine, taurine and threonine; 10 mg/mL for alanine, cysteine, glutamine, glycine, histidine, lysine and valine.

Sample collection of rat brain regions

Male Wistar rats (Charles River Laboratories, Germany) were treated as approved under the ethical guidelines for the care and use of laboratory animals and were approved by the local animal care committee (Regierungspräsidium Karlsruhe, Germany). Sixteen male rats with initial weight of 300-350 g were given a 12-hour light/dark cycle (9:00 – 21:00) and *ad libitum* access to food and water. At the selected time points (17 and 19.5 weeks), the rats were euthanised, brains were quickly removed, snap-frozen in isopentane at -40 °C, and stored at -80 °C. Coronal brain sections (100 µm) of all brain matter were generated using a cryostat (Leica Biosystems CM3050S) and were immediately dissected using micropunch tools with different dimensions depending on the area of interest (0.75-1.25 mm diameter, Stoelting). Table 1 summarizes in detail, which micropunch was used for the respective brain regions as well as the means tissue weight of all micropunched regions. Bilateral samples were obtained under a magnifying lens using anatomical landmarks from Paxinos, Watson ⁶⁰.

Brain region	Punch	Mean brain weight [mg]
OB - olfactorius bulbus	black	15.6
CgC - cingulate cortex	red	25.6
PrLC - prelimbic cortex	red	14.5
ILC - infralimbic cortex	red	11.7
GI - granular insular cortex	yellow/red	14.0

Chapter 4

Brain region	Punch	Mean brain weight [mg]
DI - dorsal insular cortex	yellow/red	13.7
AI - agranular insular cortex	yellow/red	14.5
NAC - nucleus accumbens	red	19.9
CPU - caudate putamen	black	50.6
S - septum	yellow	13.7
BNST - bed nucleus of stria terminalis	red	9.7
GP - globus pallidus	red	13.9
HYP - hypothalamus	red	15.4
CMA - amygdala (central/medial)	red	11.4
BLA - amygdala (basal/lateral)	red	20.9
HB - habenula	yellow	11.7
HC - hippocampus	red/black	45.4
CA1 - field CA1 of hippocampus	red	15.6
Th - thalamus	red/black	30.0
STh - subthalamic nucleus	yellow	10.6
SN - substantia nigra	yellow	11.5
VTA - ventral tegmental area	yellow	7.3
RA - raphe	yellow	8.7
LC - locus coeruleus	yellow	9.3
PONS - pontine nuclei	red	9.1

Table 1: Detailed description of micropunches used for every region and the mean tissue weight.

Sample preparation of rat brain regions

Liquid-liquid extraction

The protocol for liquid-liquid extraction was based on the original Bligh & Dyer protocol ^{61,62}. Freeze-dried brain samples were kept at -80°C until sample preparation. The whole sample preparation was performed on ice. To each brain sample, 400 μL MeOH and 125 μL H_2O were added, together with the stable-isotope-labelled metabolites 2-aminobutyric acid- d_6 , 3-(4-hydroxy-3-methoxy)-ethyl- d_4 -amine, alanine- $^{13}\text{C}^{15}\text{N}$, arginine- $^{13}\text{C}^{15}\text{N}$, asparagine- $^{13}\text{C}^{15}\text{N}$, aspartic acid- $^{13}\text{C}^{15}\text{N}$, beta-Alanine- D_4 , epinephrine- D_3 , glutamic acid- $^{13}\text{C}^{15}\text{N}$, glutamine- $^{13}\text{C}^{15}\text{N}$, glycine- $^{13}\text{C}^{15}\text{N}$, histamine- D_4 , isoleucine- $^{13}\text{C}^{15}\text{N}$, leucine- $^{13}\text{C}^{15}\text{N}$, lysine- $^{13}\text{C}^{15}\text{N}$, methionine- $^{13}\text{C}^{15}\text{N}$, ornithine- D_6 , serine- $^{13}\text{C}^{15}\text{N}$, threonine- $^{13}\text{C}^{15}\text{N}$, tryptophan- $^{13}\text{C}^{15}\text{N}$, tyrosine- $^{13}\text{C}^{15}\text{N}$ and valine- $^{13}\text{C}^{15}\text{N}$ as internal standards. Samples were homogenised using a bullet blender at speed 9 for 3 min, with 0.5 mm stainless steel beads (Next Advance, USA). Following homogenization, samples were centrifuged for 5 min at $800 \times g$ at 4°C (Eppendorf 5415r centrifuge, country) and 450 μL of the homogenate was transferred into new tubes (1.5 mL). To these new tubes, 460 μL chloroform, 250 μL H_2O and 60 μL MeOH were added. The samples were vortexed for 30 s and left on ice for 10 min, followed by centrifugation for 10 min at $2,000 \times g$ at 4°C . The top aqueous layer (450 μL) was removed from the samples and transferred into new tubes (0.5 mL Eppendorf vials). Derivatization took place immediately after to reduce metabolite degradation.

Derivatization

The derivatization protocol was adapted from Wong et al (2012)²¹. Rat brain supernatant was transferred into a fresh vial (200 μL) and dried using a Labconco SpeedVac concentrator (MO, United States). The dried samples were reconstituted in 10 μL H_2O whilst maintained on ice. To start the derivatization reaction, 10 μL of 100 mM sodium carbonate (pH 9.4) was added, followed by 10 μL of 2% benzoyl chloride in ACN (v/v), and vortexed immediately for 10 seconds triggering the spontaneous $\text{S}_{\text{N}}1$ reaction at room temperature. The reaction was quenched by the addition of 20 μL H_2O with 1 % sulphuric acid after 5 minutes. To the quenched mixture, 50 μL H_2O was added to reduce the organic content. The samples were vortexed again and transferred to a glass injection prior to LC-MS analysis.

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) Analysis

Analysis of the brain samples was conducted using an Agilent infinity Class II 1290 UHPLC system (Agilent Technologies, Germany) coupled to an SCIEX QTrap 6500 mass spectrometer (SCIEX, Massachusetts, USA). Five microliters of sample were injected on a T3 HSS C18 column (1.0 x 100 mm, 1.8 μ m; Waters Technologies, UK) kept at 30 °C. The autosampler was maintained at 10 °C. Mobile phase A consisted of water containing 10 mM ammonium formate and 0.1% formic acid (v/v), and mobile phase B was 100 % ACN. The flow rate was 100 μ L/min and the gradient profile was as follows: initial, 0% B; 0.01 min, 15% B; 0.5 min, 17% B; 14 min, 55% B; 14.5 min, 70% B; 18 min, 100% B; 20 min, 100% B and 22 min, 0% B. The mass spectrometer was operated in positive ionization mode at the following conditions: temperature 600 °C, curtain gas pressure 30.0 psi, collision gas set on medium, ionization voltage 5500 V, ion source gas 1 pressure 60.0 psi, and ion source gas 2 pressure 50.0 psi. Data was acquired using selected reaction monitoring (SRM), as detailed in Table S5. The samples were analysed against an eight-point calibration line with a two-fold dilution between each point. Five system suitability tests were injected before each batch and compared before the data analysis was continued. Nine quality control samples made of the pooled supernatant were injected per batch every 8-10 samples.

Statistical analysis

SCIEX MultiQuant was used for integration of the chromatographic peaks, confirmed by manual inspection. All concentrations were calculated across the calibration curves. The calibration line linear ranges are documented in Table S2. With the integrated data, statistical analysis was performed using R (<http://cran.r-project.org/>). The calculation of specific concentrations was calculated in R after assessing all analytical data for accuracy and linearity. Due to the large number of samples in this study (25 brain regions per 16 rats), they were analysed by LC-MS/MS in 10 continuous batches. To ensure data quality within and across batches, replicates of a pooled brain samples were measured in each batch for further quality analysis. In addition, each batch contained an independent metabolite calibration line to improve quantitation of the neurochemicals. Using an in-house quality control tool, we assessed the variation in metabolite area ratios between batches. Metabolites that fell below the lower limit of detection (LLOQ) were replaced with

values using minimum imputing during statistical analysis. The minimum imputing value was calculated using the method LLOQ/3. Sample outliers were not removed from the presented data to provide a realistic documentation of the biological and method variation. Before further analysis, samples were normalised to the corresponding brain sample wet weight. For each metabolite, a one-way ANOVA followed by Tukey's HSD post-hoc test was conducted to compare between the brain regions per animal. In addition, comparison of metabolite concentrations between brain regions were conducted using Student's *t*-test. False-discovery rate (FDR) correction using Benjamini-Hochberg. PCA was performed after log₂ transformation and scaling.

Data quality assessment

The performance parameters were repeatability, linear range and linearity and are detailed in Table S2. Metabolites with a quality control sample interday repeatability relative standard deviation (RSD) of the area ratio below 20% were retained, except for epinephrine (RSD = 27.5 %) and homoserine (RSD = 23.7 %). We compared rats that were 17 weeks old (n = 6) and 19.5 weeks old (n = 10). A Student's *t*-test was conducted followed by correction for multiple comparisons (Benjamini-Hochberg FDR). We could therefore confirm that the neurochemical concentration is not significantly different between the two ages represented in our study (FDR $q > 0.05$). This was true for all regions except the hippocampus, the field CA1 of the hippocampus, dorsal insular cortex and thalamus (Table S3). The hippocampus regions both had 15 metabolites identified as significantly different (FDR $q < 0.05$) with the majority consisting of amino acids; the dorsal insular cortex had choline and citrulline identified as significant, and the thalamus had 5-HIAA identified as significant. Overall, the results show that we have established an accurate methodology that can be used to quantitatively map the adult rat brain as a model for mammalian brain. After the assessment, 43 neurochemicals were identified in a reliable quantifiable concentration (metabolite ChEBI identifiers shown in Table S4).

Acknowledgements

The author expresses thanks to Elisabeth Röbel, CIMH, for dissecting the brains and Nanda Koopman at Leiden University for assisting in the analysis of the brain samples. This project was supported by the SysMedPD and Sybil-AA projects, with

Chapter 4

both receiving funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 668738 and 668863, respectively.

References

1. Burte F, Houghton D, Lowes H, et al. metabolic profiling of Parkinson's disease and mild cognitive impairment. *Mov Disord.* 2017.
2. Havelund JF, Heegaard NHH, Faergeman NJK, Gramsbergen JB. Biomarker Research in Parkinson's Disease Using Metabolite Profiling. *Metabolites.* 2017;7(3).
3. Graham SF, Holscher C, Green BD. Metabolic signatures of human Alzheimer's disease (AD): 1 H NMR analysis of the polar metabolome of post-mortem brain tissue. *Metabolomics.* 2014;10(4):744-753.
4. Ivanisevic J, Epstein AA, Kurczy ME, et al. Brain region mapping using global metabolomics. *Chemistry & biology.* 2014;21(11):1575-1584.
5. Kanehisa M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* 2019;28(11):1947-1951.
6. Romero P, Wagg J, Green ML, Kaiser D, Krummenacker M, Karp PD. Computational prediction of human metabolic pathways from the complete human genome. *Genome Biology.* 2004;6(1):R2.
7. Noronha A, Modamio J, Jarosz Y, et al. The Virtual Metabolic Human database: integrating human and gut microbiome metabolism with nutrition and disease. *Nucleic acids research.* 2018;47(D1):D614-D624.
8. Vasilopoulou CG, Margarity M, Klapa MI. Metabolomic analysis in brain research: opportunities and challenges. *Frontiers in Physiology.* 2016;7:183.
9. Dumas M-E, Davidovic L. Metabolic phenotyping and systems biology approaches to understanding neurological disorders. *F1000prime reports.* 2013;5.
10. Rae C. Metabolomics in neuroscience and disease. In: *Future Medicine;* 2014.
11. Gonzalez-Riano C, Garcia A, Barbas C. Metabolomics studies in brain tissue: A review. *Journal of Pharmaceutical and Biomedical Analysis.* 2016;130:141-168.
12. Wishart DS, Feunang YD, Marcu A, et al. HMDB 4.0: the human metabolome database for 2018. *Nucleic acids research.* 2018;46(D1):D608-d617.
13. Psychogios N, Hau DD, Peng J, et al. The Human Serum Metabolome. *PLoS One.* 2011;6(2):e16957.

14. Wu L, Niu Z, Hu X, et al. Regional cerebral metabolic levels and turnover in awake rats after acute or chronic spinal cord injury. *The FASEB Journal*. 2020;34(8):10547-10559.
15. Salek RM, Xia J, Innes A, et al. A metabolomic study of the CRND8 transgenic mouse model of Alzheimer's disease. *Neurochemistry international*. 2010;56(8):937-947.
16. Pinto MCX, de Paiva MJN, Oliveira-Lima OC, et al. Neurochemical study of amino acids in rodent brain structures using an improved gas chromatography–mass spectrometry method. *Journal of Chemical Neuroanatomy*. 2014;55:24-37.
17. Nicholson JK, Holmes E, Kinross JM, Darzi AW, Takats Z, Lindon JC. Metabolic phenotyping in clinical and surgical environments. *Nature*. 2012;491(7424):384-392.
18. Noga MJ, Dane A, Shi S, et al. Metabolomics of cerebrospinal fluid reveals changes in the central nervous system metabolism in a rat model of multiple sclerosis. *Metabolomics*. 2012;8(2):253-263.
19. Trushina E, Dutta T, Persson X-MT, Mielke MM, Petersen RC. Identification of altered metabolic pathways in plasma and CSF in mild cognitive impairment and Alzheimer's disease using metabolomics. *PLoS One*. 2013;8(5):e63644.
20. Choi WT, Tosun M, Jeong H-H, et al. Metabolomics of mammalian brain reveals regional differences. *BMC Systems Biology*. 2018;12(8):127.
21. Wong JM, Malec PA, Mabrouk OS, Ro J, Dus M, Kennedy RT. Benzoyl chloride derivatization with liquid chromatography-mass spectrometry for targeted metabolomics of neurochemicals in biological samples. *J Chromatogr A*. 2016;1446:78-90.
22. Meiser J, Weindl D, Hiller K. Complexity of dopamine metabolism. *Cell Communication and Signaling : CCS*. 2013;11:34-34.
23. Cragg S, Hawkey C, Greenfield S. Comparison of serotonin and dopamine release in substantia nigra and ventral tegmental area: region and species differences. *Journal of neurochemistry*. 1997;69(6):2378-2386.
24. Chen J, Hou W, Han B, et al. Target-based metabolomics for the quantitative measurement of 37 pathway metabolites in rat brain and serum using hydrophilic interaction ultra-high-performance liquid

- chromatography-tandem mass spectrometry. *Anal Bioanal Chem.* 2016;408(10):2527-2542.
25. Kaplan KA, Chiu VM, Lukus PA, et al. Neuronal metabolomics by ion mobility mass spectrometry: cocaine effects on glucose and selected biogenic amine metabolites in the frontal cortex, striatum, and thalamus of the rat. *Anal Bioanal Chem.* 2013;405(6):1959-1968.
 26. Devinsky O, Morrell MJ, Vogt BA. Contributions of anterior cingulate cortex to behaviour. *Brain.* 1995;118(1):279-306.
 27. Cechetto DF, Saper CB. Evidence for a viscerotopic sensory representation in the cortex and thalamus in the rat. *Journal of Comparative Neurology.* 1987;262(1):27-45.
 28. Cotter D, Mackay D, Landau S, Kerwin R, Everall I. Reduced glial cell density and neuronal size in the anterior cingulate cortex in major depressive disorder. *Archives of general psychiatry.* 2001;58(6):545-553.
 29. Giustino TF, Maren S. The Role of the Medial Prefrontal Cortex in the Conditioning and Extinction of Fear. *Frontiers in Behavioral Neuroscience.* 2015;9(298).
 30. Mhyre TR, Boyd JT, Hamill RW, Maguire-Zeiss KA. Parkinson's Disease. *Sub-cellular biochemistry.* 2012;65:389-455.
 31. Biggio G, Casu M, Corda M, Di Bello C, Gessa G. Stimulation of dopamine synthesis in caudate nucleus by intrastriatal enkephalins and antagonism by naloxone. *Science.* 1978;200(4341):552-554.
 32. Yeh H, Moises H, Waterhouse B, Woodward D. Modulatory interactions between norepinephrine and taurine, beta-alanine, gamma-aminobutyric acid and muscimol, applied iontophoretically to cerebellar Purkinje cells. *Neuropharmacology.* 1981;20(6):549-560.
 33. Ericson M, Clarke RB, Chau P, Adermark L, Söderpalm B. β -alanine elevates dopamine levels in the rat nucleus accumbens: antagonism by strychnine. *Amino Acids.* 2010;38(4):1051-1055.
 34. Margolis FL. Carnosine in the primary olfactory pathway. *Science.* 1974;184(4139):909-911.
 35. Sharma K, Schmitt S, Bergner CG, et al. Cell type- and brain region-resolved mouse brain proteome. *Nature Neuroscience.* 2015;18(12):1819-1831.

36. Testai FD, Gorelick PB. Inherited metabolic disorders and stroke part 2: homocystinuria, organic acidurias, and urea cycle disorders. *Archives of neurology*. 2010;67(2):148-153.
37. Zahedi K, Huttinger F, Morrison R, Murray-Stewart T, Casero RA, Strauss KI. Polyamine catabolism is enhanced after traumatic brain injury. *Journal of neurotrauma*. 2010;27(3):515-525.
38. Ha HC, Sirisoma NS, Kuppusamy P, Zweier JL, Woster PM, Casero RA. The natural polyamine spermine functions directly as a free radical scavenger. *Proceedings of the National Academy of Sciences*. 1998;95(19):11140-11145.
39. Lein ES, Hawrylycz MJ, Ao N, et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature*. 2007;445(7124):168-176.
40. Ortiz C, Navarro JF, Jurek A, Märtin A, Lundeberg J, Meletis K. Molecular atlas of the adult mouse brain. *Science Advances*. 2020;6(26):eabb3446.
41. Reinikainen KJ, Paljärvi L, Huuskonen M, Soininen H, Laakso M, Riekkinen PJ. A post-mortem study of noradrenergic, serotonergic and GABAergic neurons in Alzheimer's disease. *Journal of the neurological sciences*. 1988;84(1):101-116.
42. Koob GF, Sanna PP, Bloom FE. Neuroscience of addiction. *Neuron*. 1998;21(3):467-476.
43. Kalivas PW. Neurocircuitry of addiction. *Neuropsychopharmacology: the fifth generation of progress*. 2002;95:1357-1366.
44. Godwin-Austen RB, Smith NJ. Comparison of the effects of bromocriptine and levodopa in Parkinson's disease. *J Neurol Neurosurg Psychiatry*. 1977;40(5):479-482.
45. Brunk E, Sahoo S, Zielinski DC, et al. Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nature biotechnology*. 2018;36(3):272.
46. Thiele I, Swainston N, Fleming RM, et al. A community-driven global reconstruction of human metabolism. *Nature biotechnology*. 2013;31(5):419-425.
47. Sara SJ. The locus coeruleus and noradrenergic modulation of cognition. *Nature reviews neuroscience*. 2009;10(3):211-223.

48. Huang KW, Ochandarena NE, Philson AC, et al. Molecular and anatomical organization of the dorsal raphe nucleus. *Elife*. 2019;8:e46464.
49. Bichell TJV, Wegrzynowicz M, Tipps KG, et al. Reduced bioavailable manganese causes striatal urea cycle pathology in Huntington's disease mouse model. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*. 2017;1863(6):1596-1604.
50. Chiang M-C, Chen H-M, Lee Y-H, et al. Dysregulation of C/EBP α by mutant Huntingtin causes the urea cycle deficiency in Huntington's disease. *Human molecular genetics*. 2007;16(5):483-498.
51. Hansmannel F, Sillaire A, Kamboh MI, et al. Is the urea cycle involved in Alzheimer's disease? *Journal of Alzheimer's Disease*. 2010;21(3):1013-1021.
52. Seiler N. CHAPTER 8 - POLYAMINE METABOLISM AND FUNCTION IN BRAIN. In: Osborne NN, ed. *Selected Topics from Neurochemistry*. Amsterdam: Pergamon; 1985:147-165.
53. Seiler N. Catabolism of polyamines. *Amino Acids*. 2004;26(3):217-233.
54. Yogeewari P, Sriram D, Vaigundaragavendran J. The GABA shunt: an attractive and potential therapeutic target in the treatment of epileptic disorders. *Current drug metabolism*. 2005;6(2):127-139.
55. Willacey CCW, Naaktgeboren M, Lucumi Moreno E, et al. LC-MS/MS analysis of the central energy and carbon metabolites in biological samples following derivatization by dimethylaminophenacyl bromide. *Journal of Chromatography A*. 2019:460413.
56. Gottas A, Ripel A, Boix F, Vindenes V, Morland J, Oiestad EL. Determination of dopamine concentrations in brain extracellular fluid using microdialysis with short sampling intervals, analyzed by ultra high performance liquid chromatography tandem mass spectrometry. *J Pharmacol Toxicol Methods*. 2015;74:75-79.
57. Cannazza G, Carrozzo MM, Cazzato AS, et al. Simultaneous measurement of adenosine, dopamine, acetylcholine and 5-hydroxytryptamine in cerebral mice microdialysis samples by LC-ESI-MS/MS. *J Pharm Biomed Anal*. 2012;71:183-186.
58. Han X, Min M, Wang J, et al. Quantitative profiling of neurotransmitter abnormalities in brain, cerebrospinal fluid, and serum of experimental

- diabetic encephalopathy male rat. *Journal of Neuroscience Research*. 2018;96(1):138-150.
59. Ivanisevic J, Siuzdak G. The role of metabolomics in brain metabolism research. *Journal of Neuroimmune Pharmacology*. 2015;10(3):391-395.
60. Paxinos G, Watson C. *The rat brain in stereotaxic coordinates: hard cover edition*. Elsevier; 2006.
61. Bligh EG, Dyer WJ. A rapid method of total lipid extraction and purification. *Canadian journal of biochemistry and physiology*. 1959;37(8):911-917.
62. Sündermann A, Eggers LF, Schwudke D. Liquid Extraction: Bligh and Dyer. In: Wenk MR, ed. *Encyclopedia of Lipidomics*. Dordrecht: Springer Netherlands; 2016:1-4.

Supplementary information

Table S1. The relative standard deviation of metabolites from the pooled brain samples for quality control.

Metabolite	QC _{RSD} (%)	Metabolite	QC _{RSD} (%)
3-Methoxytyramine	10.4	Homoserine	23.7
3,4-Dihydroxyphenylacetic acid	13.0	Homovanillic acid	13.0
5-Hydroxyindoleacetic acid	9.3	Hypotaurine	5.2
Acetylcholine	7.6	Isoleucine	2.3
Alanine	3.1	Kynurenine	10.3
Arginine	6.1	Leucine	2.1
Asparagine	5.9	Lysine	2.8
Aspartic acid	6.5	Methionine	2.0
β -Alanine	5.4	Norepinephrine	5.4
Carnosine	7.4	Ornithine	8.2
Choline	8.3	Phenylalanine	3.1
Citrulline	8.5	Proline	10.9
Cysteine	13.5	Putrescine	4.2
Dopamine	9.3	Serine	5.3
Epinephrine	27.5	Serotonin	17.2
Ethanolamine	12.5	Spermidine	9.1
Gamma-Aminobutyric acid	2.8	Spermine	10.3
Glutamic acid	5.1	Taurine	7.8
Glutamine	4.8	Threonine	5.4
Glutathione	8.4	Tryptophan	2.8
Glycine	5.4	Tyramine	25.9
Histamine	14.2	Tyrosine	1.8
Histidine	7.6	Valine	2.0

Chapter 4

Table S2. Analytical results showing the linear equation, linear range and R² value of the metabolites.

Metabolites	Linear equation	Linear range (μM)	R ²
3-Methoxytyramine	$y = -0.017 + 0.59 \times X$	0 - 9.82	0.997
3,4-Dihydroxyphenylacetic acid	$y = -5e-04 + 0.059 \times x$	0 - 10.14	0.99
5-Hydroxyindoleacetic acid	$y = 0.02 + 0.017 \times X$	0 - 210	0.993
Acetylcholine	$y = 0.01 + 0.09 \times x$	0 - 11	0.992
Alanine	$y = 0.0049 + 0.01 \times x$	0 - 1122	1
Arginine	$y = 0.0089 + 0.018 \times x$	0 - 475	1
Asparagine	$y = 0.0015 + 0.012 \times x$	0 - 151	0.999
Aspartic acid	$y = 0.0073 + 0.016 \times x$	0 - 75	0.998
Beta-Alanine	$y = 0.00007 + 0.01 \times x$	0 - 224	1
Carnosine	$y = -0.022 + 0.017 \times x$	0 - 221	0.997
Choline	$y = 0.0043 + 0.024 \times x$	0 - 8.7	0.995
Citrulline	$y = 0.0071 + 0.0041 \times x$	0 - 713	0.998
Dopamine	$y = 0.0012 + 0.069 \times x$	0 - 23.4	0.999
Epinephrine	$y = -0.0019 + 0.029 \times x$	0 - 6.00	0.997
Ethanolamine	$y = -2.1 + 1 \times x$	0 - 410	0.997
Gamma-Aminobutyric acid	$y = 0.01 + 0.017 \times x$	0 - 48.5	0.998
Glutamic acid	$y = 0.00082 + 0.019 \times x$	0 - 340	0.999
Glutamine	$y = -0.064 + 0.026 \times x$	0 - 1370	1
Glycine	$y = 0.017 + 0.011 \times x$	0 - 2663	1
Histamine	$y = 0.011 + 0.028 \times x$	0 - 27.2	0.994
Histidine	$y = -0.41 + 0.039 \times x$	0 - 955	0.995
Homoserine	$y = 0.005 + 0.0075 \times x$	0 - 84.0	0.997
Homovanillic acid	$y = 0.00019 + 0.11 \times x$	0 - 22.0	0.998
Hypotaurine	$y = 0.001 + 0.009 \times x$	0 - 620	1
Isoleucine	$y = 0.015 + 0.046 \times x$	0 - 190	0.999
Kynurenine	$y = 0.002 + 0.38 \times X$	0 - 48.0	0.997
Leucine	$y = 0.0055 + 0.0095 \times x$	0 - 76.2	1
Lysine	$y = 0.0037 + 0.0093 \times x$	0 - 548	0.999
Methionine	$y = -0.0072 + 0.06 \times x$	0 - 168	1
Norepinephrine	$y = -0.00057 + 0.029 \times x$	0 - 23.72	0.996
Ornithine	$y = 0.037 + 0.053 \times x$	0 - 29.7	0.999
Phenylalanine	$y = -0.0026 + 0.18 \times x$	0 - 303	0.999
Proline	$y = 0.015 + 0.03 \times x$	0 - 217	1

Absolute quantitative neurochemical brain atlas

Metabolites	Linear equation	Linear range (μM)	R ²
Putrescine	$y = -0.0044 + 0.31 \times x$	0 - 12.4	1
Serine	$y = 0.17 + 0.024 \times x$	0 - 475	0.999
Serotonin	$y = 0.0097 + 0.18 \times x$	0 - 9.5	0.994
Spermidine	$y = -0.42 + 0.99 \times x$	0 - 27.5	0.994
Spermine	$y = -0.051 + 0.33 \times x$	0 - 9.88	0.969
Taurine	$y = -0.012 + 0.0092 \times x$	0 - 1598	0.998
Threonine	$y = 0.04 + 0.0053 \times x$	0 - 840	0.999
Tryptophan	$y = -0.016 + 0.064 \times x$	0 - 490	1
Tyrosine	$y = 0.011 + 0.16 \times x$	0 - 552	1
Valine	$y = -0.012 + 0.032 \times x$	0 - 854	1

Table S3. FDR corrected student t-test values comparing the metabolite concentrations in the brain regions between 17 and 19.5 week old rats.

3-Methoxytyramine	0.98	0.64	0.99	0.78	0.41	0.9	0.99	0.65	0.96	0.42	0.95	0.03	0.93	1	0.61	0.99	0.62	0.6	0.98	0.88	0.67	0.84	0.96	0.87	0.56
3,4-Dihydroxyphenylacetic acid	0.98	0.54	0.99	0.23	0.28	0.8	0.99	0.66	0.43	0.69	0.95	0.78	0.76	1	0.61	0.99	0.99	0.6	0.99	0.11	0.26	0.97	0.79	0.68	0.56
5-Hydroxyindoleacetic acid	0.98	0.39	0.99	0.28	0.57	0.9	0.99	0.67	0.97	0.8	0.99	0.75	0.76	1	0.61	0.99	0.62	0.73	0.98	0.81	0.26	0.84	0.54	0.03	0.56
Acetylcholine	0.59	0.89	0.99	0.76	0.31	0.8	0.99	0.12	0.95	0.4	0.99	0.78	0.99	1	0.61	0.79	0.83	0.78	0.98	0.86	0.26	0.95	0.5	0.78	0.67
Alanine	0.98	0.49	0.99	0.37	0.06	0.95	0.99	0.12	0.22	0.47	0.95	0.19	0.66	1	0.61	0.79	0.53	0.79	0.98	0.55	0.28	0.84	0.5	0.49	0.56
Arginine	0.59	0.39	0.99	0.19	0.1	0.86	0.99	0.41	0.2	0.35	0.95	0.03	0.66	1	0.61	0.79	0.53	0.6	0.98	0.57	0.26	0.84	0.5	0.92	0.56
Asparagine	0.59	0.39	0.99	0.17	0.03	0.9	0.99	0.22	0.45	0.35	0.95	0.03	0.66	1	0.61	0.79	0.53	0.6	0.98	0.58	0.26	0.84	0.5	0.66	0.56
Aspartic acid	0.59	0.79	0.99	0.19	0.25	0.95	0.99	0.4	0.43	0.35	0.95	0.03	0.76	1	0.61	0.79	0.53	0.6	0.99	0.58	0.26	0.84	0.5	0.56	0.56
aminopropanoic acid	0.98	0.45	0.99	0.17	0.02	0.95	0.99	0.18	0.45	0.81	0.95	0.05	0.76	1	0.61	0.99	0.57	0.6	0.98	0.75	0.28	0.84	0.5	0.53	0.56
Carnosine	0.98	0.41	0.99	0.25	0.03	0.86	0.99	0.12	0.96	0.35	0.95	0.11	0.76	1	0.61	0.82	0.99	0.6	0.98	0.56	0.26	0.84	0.5	0.58	0.56
Choline	0.98	0.89	0.99	0.64	0.93	0.95	0.99	0.02	0.16	0.6	0.95	0.16	0.66	1	0.61	0.79	0.54	0.79	0.98	0.55	0.29	0.99	0.5	0.72	0.56
Citrulline	0.59	0.39	0.99	0.37	0.15	0.91	0.99	0.05	0.39	0.35	0.95	0.07	0.76	1	0.61	0.79	0.53	0.6	0.98	0.55	0.26	0.84	0.64	0.9	0.62
Dopamine	0.65	0.65	0.99	0.88	0.27	0.8	0.99	0.4	0.97	0.35	0.99	0.97	0.99	1	0.61	0.99	0.9	0.91	0.98	0.84	0.26	0.97	0.5	0.55	0.56
Epinephrine	NA	0.96	0.99	0.88	0.67	NA	NA	0.5	NA	0.6	0.95	NA	0.76	1	0.61	0.79	0.87	0.67	0.98	0.55	0.33	0.84	0.56	0.48	0.56
Ethanolamine	0.59	0.78	0.99	0.17	0.03	0.95	0.99	0.38	0.2	0.35	0.95	0.07	0.66	1	0.61	0.79	0.53	0.6	0.99	0.55	0.26	0.95	0.5	0.67	0.56
Gamma-Aminobutyric acid	0.98	0.79	0.99	0.25	0.11	0.86	0.99	0.19	0.2	0.75	0.95	0.15	0.77	1	0.61	0.79	0.99	0.79	0.98	0.55	0.26	0.95	0.5	0.67	0.56
Glutamic acid	0.66	0.84	0.99	0.28	0.27	0.95	0.99	0.65	0.53	0.28	0.95	0.1	0.97	1	0.61	0.79	0.99	0.67	0.98	0.55	0.26	0.94	0.5	0.37	0.56
Glutamine	0.98	0.49	0.99	0.23	0.25	0.9	0.99	0.12	0.43	0.35	0.95	0.04	0.76	1	0.61	0.79	0.56	0.6	0.98	0.55	0.26	0.95	0.5	0.32	0.56
Glycine	0.98	0.54	0.99	0.23	0.03	0.95	0.99	0.16	0.16	0.35	0.95	0.04	0.72	1	0.61	0.79	0.53	0.6	0.98	0.57	0.26	0.84	0.5	0.71	0.56
Histamine	0.72	0.49	0.99	0.69	0.41	0.37	0.99	0.77	0.97	0.35	0.95	0.78	0.97	1	0.61	0.79	0.62	NA	0.99	0.55	0.26	0.97	0.5	0.38	0.56
Histidine	0.66	0.39	0.99	0.17	0.02	0.9	0.99	0.4	0.2	0.52	0.95	0.03	0.66	1	0.61	0.79	0.53	0.6	0.99	0.55	0.26	0.84	0.5	0.93	0.56
Homoserine	0.66	0.39	0.99	0.63	0.41	0.86	0.99	0.5	0.2	0.35	0.95	0.7	0.76	1	0.77	0.79	0.53	0.62	0.63	0.55	0.29	0.84	0.5	0.44	0.56
Homovanillic acid	0.66	0.39	0.99	0.76	0.5	0.9	0.99	0.96	0.97	0.35	0.95	0.83	0.66	1	0.61	0.79	0.9	0.6	0.99	0.55	0.52	0.84	0.69	0.09	0.66
Hypotaureine	0.72	0.39	0.99	0.6	0.03	0.86	0.99	0.75	0.53	0.35	0.95	0.97	0.66	1	0.61	0.79	0.53	0.6	0.98	0.64	0.26	0.84	0.5	0.94	0.56
Isoleucine	0.98	0.39	0.99	0.25	0.05	0.9	0.99	0.18	0.2	0.75	0.95	0.04	0.66	1	0.61	0.79	0.53	0.6	0.99	0.55	0.36	0.95	0.56	0.68	0.57
Kynurenine	0.59	0.68	0.99	0.49	0.93	0.86	0.99	0.38	0.97	0.92	0.95	0.06	0.76	1	0.9	0.79	0.54	0.6	0.98	0.55	0.8	0.94	0.5	0.37	0.56
Leucine	0.98	0.54	0.99	0.25	0.06	0.9	0.99	0.12	0.2	0.81	0.95	0.06	0.66	1	0.61	0.79	0.53	0.6	0.98	0.55	0.29	0.99	0.55	0.8	0.56
Lysine	0.59	0.39	0.99	0.19	0.05	0.49	0.99	0.74	0.96	0.35	0.95	0.03	0.66	1	0.61	0.79	0.53	0.6	0.98	0.56	0.26	0.84	0.5	0.23	0.56
Methionine	0.59	0.39	0.99	0.17	0.03	0.86	0.99	0.5	0.43	0.6	0.95	0.03	0.66	1	0.61	0.82	0.53	0.6	0.98	0.55	0.26	0.84	0.5	0.89	0.56
N-Acetylputrescine	0.98	0.49	0.99	0.6	0.93	0.16	0.99	0.38	0.43	0.81	0.95	0.19	0.76	1	0.61	0.79	0.53	0.78	0.98	0.55	0.28	0.95	0.71	0.58	0.56
Norepinephrine	0.98	0.4	0.99	0.17	0.07	0.99	0.99	0.43	0.96	0.69	0.95	0.16	0.76	1	0.61	0.99	0.53	0.6	0.98	0.75	0.26	0.94	0.5	0.61	0.56
Ornithine	0.98	0.39	0.99	0.69	0.17	0.9	0.99	0.12	0.2	0.58	0.95	0.06	0.76	1	0.94	0.79	0.53	0.8	0.98	0.81	0.45	0.84	0.55	0.97	0.56
Phenylalanine	0.65	0.49	1	0.17	0.03	0.9	0.99	0.22	0.21	0.6	0.95	0.03	0.66	1	0.61	0.79	0.53	0.6	0.99	0.55	0.26	0.84	0.5	0.92	0.56
Proline	0.66	0.39	0.99	0.23	0.02	0.86	0.99	0.22	0.26	0.7	0.95	0.04	0.66	1	0.61	0.79	0.53	0.6	0.98	0.55	0.27	0.84	0.54	0.59	0.56
Putrescine	0.98	0.39	0.99	0.31	0.11	0.86	0.99	0.79	0.96	0.81	0.95	0.48	0.76	1	0.61	0.79	0.83	0.79	0.99	0.55	0.66	0.95	0.69	0.8	0.95
Serine	0.52	0.39	0.99	0.28	0.03	0.86	0.99	0.18	0.55	0.35	0.95	0.89	0.66	1	0.61	0.79	0.53	0.6	0.99	0.55	0.28	0.84	0.5	0.5	0.56
Serotonin	0.9	0.98	0.99	0.49	0.51	0.95	0.99	0.5	0.55	0.35	0.95	0.78	0.76	1	0.65	0.79	0.83	0.78	0.98	0.89	0.29	0.84	0.55	0.69	0.56
Spermidine	0.74	0.64	0.99	0.17	0.18	0.9	0.99	0.75	0.45	0.35	0.95	0.75	0.76	1	0.61	0.79	0.83	0.6	0.98	0.78	0.29	0.84	0.5	1	0.56
Spermine	0.98	0.96	0.99	0.19	0.17	0.9	0.99	0.75	0.43	0.55	0.95	0.21	0.66	1	0.61	0.99	0.83	0.6	0.99	0.55	0.29	0.84	0.5	0.71	0.56
Threonine	0.98	0.45	0.99	0.25	0.19	0.9	0.99	0.38	0.55	0.35	0.95	0.1	0.72	1	0.61	0.82	0.53	0.6	0.98	0.75	0.28	0.84	0.5	0.98	0.56
Tryptophan	0.59	0.39	0.99	0.17	0.03	0.86	0.99	0.43	0.38	0.45	0.95	0.03	0.66	1	0.61	0.99	0.53	0.6	0.98	0.27	0.26	0.84	0.5	0.69	0.57
Tyrosine	0.59	0.39	0.99	0.17	0.03	0.86	0.99	0.5	0.74	0.35	0.95	0.03	0.66	1	0.61	0.99	0.57	0.6	0.98	0.55	0.26	0.84	0.5	0.18	0.56
Valine	0.98	0.39	0.99	0.29	0.08	0.86	0.99	0.38	0.43	0.79	0.95	0.07	0.66	1	0.61	1	0.53	0.6	0.98	0.55	0.36	0.94	0.56	0.97	0.57
Aganthal insular cortex																									
Amgydala (centrolateral)																									
Amgydala (basolateral)																									
Bed nucleus of stria terminalis																									
Field CA1 of hippocampus																									
Cingulate cortex																									
Caudate putamen																									
Dorsal insular cortex																									
Granular insular cortex																									
Globus pallidus																									
Habenula																									
Hippocampus																									
Hypothalamus																									
Infralimbic cortex																									
Isthmus of cingulate cortex																									
Nucleus accumbens																									
Nucleus caudatus																									
Olfactory bulb																									
Parabrachial nucleus																									
Perirhinal cortex																									
</																									

Chapter 4

Table S4. ChEBI identifiers for metabolites used in this method, including the metabolite class and associated pathways.

Metabolite	ChEBI ID	Metabolite Class	Pathway
3-Methoxytyramine	1582	Dopamine metabolite	Tyrosine metabolism
3,4-Dihydroxyphenylacetic acid	41941	Dopamine metabolite	Tyrosine metabolism
5-Hydroxyindoleacetic acid	27823	Serotonin metabolite	Tryptophan metabolism
Acetylcholine	15355	Neurotransmitter	Cholinergic
Alanine	16977	Amino acid	
Arginine	29016	Amino acid	Urea cycle and polyamine metabolism
Asparagine	17196	Amino acid	
Aspartic acid	22660	Amino acid and neurotransmitter	
Beta-Alanine	16958	Amino acid and neurotransmitter	
Carnosine	15727	Antioxidant	
Choline	15354	Acetylcholine precursor	Cholinergic
Citrulline	18211	Biogenic amine	Urea cycle and polyamine metabolism
Cysteine	17561	Amino acid and neurotransmitter	
Dopamine	18243	Neurotransmitter	Tyrosine metabolism
Epinephrine	33568	Neurotransmitter	Tyrosine metabolism
Ethanolamine	16000	Biogenic amine	
Gamma-Aminobutyric acid	16865	Neurotransmitter	GABAergic
Glutamic acid	18237	Amino acid and neurotransmitter	
Glutamine	18050	Amino acid and neurotransmitter precursor	
Glycine	15428	Amino acid and neurotransmitter	
Histamine	18295	Neurotransmitter	
Histidine	15971	Amino acid and histamine precursor	
Homoserine	15699	Amino acid	
Homovanillic acid	545959	Dopamine metabolite	Tyrosine metabolism
Hypotaurine	16668	Antioxidant	
Isoleucine	24898	Amino acid	
Kynurenine	16946	Biogenic amine	Tryptophan metabolism
Leucine	25017	Amino acid	
Lysine	18019	Amino acid	
Methionine	16811	Amino acid	
Norepinephrine	18357	Neurotransmitter	Tyrosine metabolism
Ornithine	15729	Biogenic amine	Urea cycle and polyamine metabolism
Phenylalanine	28044	Amino acid	
Proline	26271	Amino acid	
Putrescine	17148	Polyamine	Urea cycle and polyamine metabolism
Serine	17822	Amino acid (D-isomer neurotransmitter)	
Serotonin	28790	Neurotransmitter	Tryptophan metabolism
Spermidine	16610	Polyamine	Urea cycle and polyamine metabolism
Spermine	15098	Polyamine	Urea cycle and polyamine metabolism
Taurine	15891	Neurotransmitter	
Threonine	16857	Amino acid	
Tryptophan	27897	Amino acid	Tryptophan metabolism
Tyrosine	18186	Amino acid	Tyrosine metabolism
Valine	16414	Amino acid	

Absolute quantitative neurochemical brain atlas

Table S5. Analytical parameters showing the product and precursor ions, collision energy, retention time, declustering potential and entrance/exit voltage for the metabolites and internal standards.

Metabolites	Collision						
	Precursor ion (m/z)	Product ion (m/z)	Retention time (min)	energy (eV)	Entrance potential	Exit potential	Declustering potential
2-Aminobutyric acid-D ₆	214	105	5.05	20	10	6	30
3-(4-hydroxy-3-methoxy)-ethyl-D ₄ -amine	276	105	7.48	20	10	6	30
3,4-Dihydroxyphenylacetic acid	394	105	14.87	20	10	6	30
3-Methoxytyramine	376	105	15.33	20	10	6	30
5-Hydroxyindoleacetic acid	313	146	11.2	20	10	6	30
Acetylcholine	146	87	1.08	15	10	6	30
Alanine	194	105	4.07	20	10	6	21
Alanine- ¹³ C ¹⁵ N	198	105	4.07	20	10	6	21
Arginine	279	105	3.1	35	10	12	41
Arginine- ¹³ C ¹⁵ N	289	105	3.1	35	10	12	41
Asparagine	237	105	3.09	37	10	6	26
Asparagine- ¹³ C ¹⁵ N	243	105	3.09	37	10	6	26
Aspartic acid	238	105	3.37	10	10	6	30
Aspartic acid- ¹³ C ¹⁵ N	243	105	3.37	10	10	6	30
Beta-Alanine	194	105	3.9	20	10	6	30
Beta-Alanine-D ₄	198	105	3.9	20	10	6	30
Carnosine	331	110	2.96	20	10	6	30
Choline	104	60	0.87	20	10	6	30
Citrulline	280	105	3.36	20	10	6	30
Dopamine	466	105	16.8	27	10	6	51
Epinephrine	496	105	16.15	27	10	6	56
Epinephrine-D ₃	499	105	16.15	20	10	6	30
Ethanolamine	166	105	3.41	20	10	6	30

Chapter 4

Metabolites	Precursor ion (m/z)	Product ion (m/z)	Retention time (min)	Collision	Entrance potential	Exit potential	Declustering potential
				energy (eV)			
Gamma-Aminobutyric acid	208	105	4.32	10	10	6	30
Glutamic acid	252	105	3.64	20	10	6	26
Glutamic acid- ¹³ C ¹⁵ N	258	105	3.64	20	10	6	30
Glutamine	251	105	3.2	20	10	6	30
Glutamine- ¹³ C ¹⁵ N	260	105	3.2	20	10	6	30
Glycine	180	105	3.47	17	10	12	11
Glycine- ¹³ C ¹⁵ N	183	105	3.47	17	10	12	11
Histamine	216	105	3.14	20	10	6	30
Histamine-D ₄	220	105	3.14	20	10	6	30
Histidine	260	105	2.86	31	10	6	30
Homoserine	224	105	3.33	20	10	6	30
Homovanillic acid	304	105	11.76	20	10	6	30
Hypotaurine	214	105	2.86	20	10	6	30
Isoleucine	236	105	7.99	25	10	14	16
Isoleucine- ¹³ C ¹⁵ N	243	105	7.99	25	10	14	16
Kynurenine	417	122	11.93	20	10	6	36
Leucine	236	105	8.34	25	10	14	16
Leucine- ¹³ C ¹⁵ N	243	105	8.34	25	10	14	16
Lysine	355	188	7.55	20	10	6	30
Lysine- ¹³ C ¹⁵ N	363	194	7.55	20	10	6	30
Methionine	254	105	6.35	15	10	6	30
Methionine- ¹³ C ¹⁵ N	260	105	6.35	15	10	6	30
Norepinephrine	482	105	15.7	27	10	6	56
Ornithine	341	174	6.93	15	10	6	30
Ornithine-D ₆	347	180	6.93	15	10	6	30
Phenylalanine	270	120	8.51	10	10	6	30

Absolute quantitative neurochemical brain atlas

Metabolites	Collision						
	Precursor ion (m/z)	Product ion (m/z)	Retention time (min)	energy (eV)	Entrance potential	Exit potential	Declustering potential
Proline	220	105	4.72	20	10	6	30
Putrescine	297	105	8.03	30	10	6	30
Serine	210	105	3.17	20	10	6	30
Serine- ¹³ C ¹⁵ N	214	105	3.17	20	10	6	30
Serotonin	385	264	14.87	20	10	6	100
Spermidine	458	162	10.65	20	10	6	71
Spermine	619.6	497	12.41	25	10	18	31
Taurine	230	105	3.03	10	10	6	30
Threonine	224	105	3.6	10	10	6	36
Threonine- ¹³ C ¹⁵ N	229	105	3.6	10	10	6	36
Tryptophan	309	159	8.72	10	10	6	30
Tryptophan- ¹³ C ¹⁵ N	322	171	8.72	10	10	6	30
Tyrosine	390	105	12.89	30	10	6	51
Tyrosine- ¹³ C ¹⁵ N	400	105	12.89	30	10	6	51
Valine	222	105	6.33	30	10	6	30
Valine- ¹³ C ¹⁵ N	228	105	6.33	30	10	6	30

Table S6. Calculation table showing the difference tyrosine metabolism concentrations between the regions Cingulate cortex (CGC), caudate putamen (CPU), hypothalamus (HYP), substantia nigra (SN) and locus coeruleus (LC) with one-way ANOVA using the posthoc test of Tukey, and paired t-test with FDR adjustment with Benjamini-Hochberg.

Variable	Comparison	Tukey HSD	T-test	FDR adj.
Tyrosine	CPU-CGC	0.0383802	0.000012	2.82E-05
Tyrosine	HYP-CGC	0.0065248	0.003537	0.004879
Tyrosine	SN-CGC	0.0002474	1.78E-07	6.7E-07
Tyrosine	LC-CGC	0	7.97E-09	4.56E-08
Tyrosine	HYP-CPU	0.0000002	1.22E-06	3.63E-06
Tyrosine	SN-CPU	0	1.03E-08	5.58E-08
Tyrosine	LC-CPU	0	5.21E-10	5.08E-09
Tyrosine	SN-HYP	0.867557	0.257104	0.286335
Tyrosine	LC-HYP	0.0000001	0.000132	0.000244
Tyrosine	LC-SN	0.000006	6.03E-05	0.000121
Dopamine	CPU-CGC	0	4.47E-12	1.34E-10
Dopamine	HYP-CGC	0.5993419	0.382896	0.413759
Dopamine	SN-CGC	0	4.17E-07	1.38E-06
Dopamine	LC-CGC	0.0000002	0.000972	0.00148
Dopamine	HYP-CPU	0	6.15E-14	5.27E-12
Dopamine	SN-CPU	0	3.05E-12	1.09E-10
Dopamine	LC-CPU	0	4.9E-17	2.1E-14
Dopamine	SN-HYP	0	2.87E-07	1.02E-06
Dopamine	LC-HYP	0	1.04E-05	2.52E-05
Dopamine	LC-SN	0	1.66E-12	6.48E-11
Norepinephrine	CPU-CGC	0	1.52E-10	2.04E-09
Norepinephrine	HYP-CGC	0	1.05E-06	3.17E-06
Norepinephrine	SN-CGC	0	6.68E-09	3.98E-08
Norepinephrine	LC-CGC	0.9999564	0.624624	0.65357
Norepinephrine	HYP-CPU	0	1.56E-11	3.04E-10
Norepinephrine	SN-CPU	0.2308768	0.056935	0.067473
Norepinephrine	LC-CPU	0	1.46E-07	5.66E-07
Norepinephrine	SN-HYP	0	9.47E-10	7.39E-09
Norepinephrine	LC-HYP	0	0.0001	0.000192
Norepinephrine	LC-SN	0	1.72E-05	3.87E-05
DOPAC	CPU-CGC	0	4.04E-12	1.33E-10

Absolute quantitative neurochemical brain atlas

Variable	Comparison	Tukey HSD	T-test	FDR adj.
DOPAC	HYP-CGC	0.4760077	0.07036	0.082023
DOPAC	SN-CGC	0.0000039	2.59E-05	5.63E-05
DOPAC	LC-CGC	0.9992935	0.934282	0.943076
DOPAC	HYP-CPU	0	1.09E-14	1.17E-12
DOPAC	SN-CPU	0	2.79E-11	4.99E-10
DOPAC	LC-CPU	0	7.3E-10	6.27E-09
DOPAC	SN-HYP	0	6.79E-06	1.69E-05
DOPAC	LC-HYP	0.6260875	0.215665	0.2422
DOPAC	LC-SN	0.0000015	0.00052	0.000829
3-MT	CPU-CGC	0	2.67E-08	1.23E-07
3-MT	HYP-CGC	0.8492226	0.444833	0.475893
3-MT	SN-CGC	0.0000192	4.57E-05	9.48E-05
3-MT	LC-CGC	0.0019659	0.001997	0.002894
3-MT	HYP-CPU	0	1.16E-07	4.73E-07
3-MT	SN-CPU	0.0001365	3.81E-09	2.59E-08
3-MT	LC-CPU	0	7.31E-09	4.24E-08
3-MT	SN-HYP	0.0000003	0.000121	0.000228
3-MT	LC-HYP	0.0401984	0.131591	0.150541
3-MT	LC-SN	0	2.28E-07	8.38E-07
HVA	CPU-CGC	0	7E-11	1.11E-09
HVA	HYP-CGC	0.0000619	0.006398	0.00855
HVA	SN-CGC	0.0811306	0.000408	0.000659
HVA	LC-CGC	0.9986749	0.70737	0.732807
HVA	HYP-CPU	0	1.04E-06	3.16E-06
HVA	SN-CPU	0.0000009	6.84E-10	5.98E-09
HVA	LC-CPU	0	2.16E-09	1.6E-08
HVA	SN-HYP	0	0.00022	0.000383
HVA	LC-HYP	0.0000214	0.003306	0.004592
HVA	LC-SN	0.1498828	0.000205	0.000359

Chapter 4

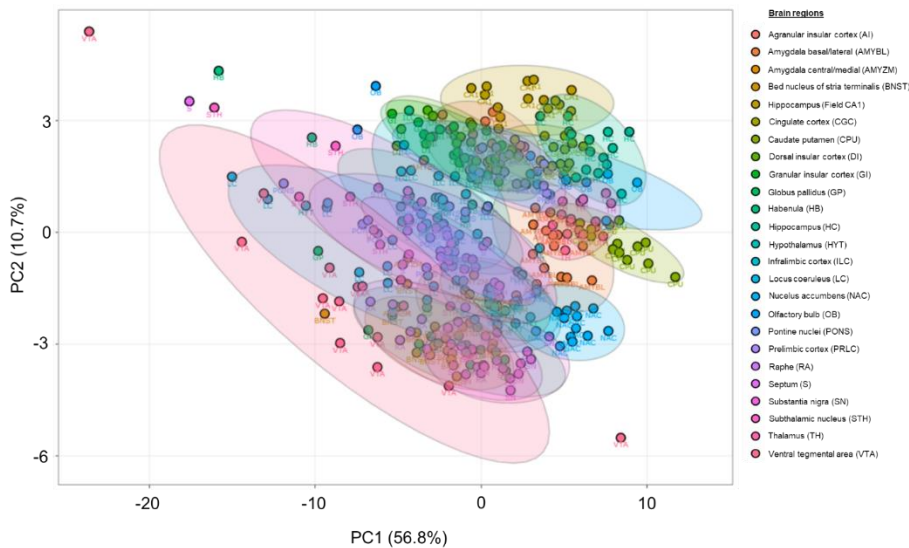
Table S7. Calculation table showing the difference in the urea cycle metabolism concentrations between the regions cingulate cortex (CGC), thalamus (TH), hypothalamus (HYP) and raphe (RA) with one-way ANOVA using the posthoc test of Tukey, and paired t-test with FDR adjustment with Benjamini-Hochberg.

Variable	Comparison	Tukey HSD	Paired T-test	FDR adj.
Aspartic acid	HYP-CGC	0.004077	0.00582947	0.009115
Aspartic acid	TH-CGC	0.0653043	0.03385037	0.044787
Aspartic acid	RA-CGC	0	6.9645E-06	2.07E-05
Aspartic acid	TH-HYP	0.0000007	0.00016906	0.000363
Aspartic acid	RA-HYP	0.0015311	0.00079501	0.001497
Aspartic acid	RA-TH	0	8.2037E-09	7.06E-08
Arginine	HYP-CGC	0.9972692	0.96499371	0.972533
Arginine	TH-CGC	0	1.42E-07	6.78E-07
Arginine	RA-CGC	0.0606083	0.00420978	0.006831
Arginine	TH-HYP	0	9.917E-06	2.87E-05
Arginine	RA-HYP	0.0374952	0.06026586	0.077356
Arginine	RA-TH	0	8.819E-09	7.34E-08
Citrulline	HYP-CGC	0.0000002	2.6051E-05	6.65E-05
Citrulline	TH-CGC	0.0030859	6.7557E-06	2.03E-05
Citrulline	RA-CGC	0.0000003	1.286E-05	3.65E-05
Citrulline	TH-HYP	0	2.5047E-07	1.08E-06
Citrulline	RA-HYP	0.9968472	0.70833689	0.758195
Citrulline	RA-TH	0	2.6866E-09	2.89E-08
Ornithine	HYP-CGC	0.9118277	0.66940796	0.722624
Ornithine	TH-CGC	0.004658	5.5294E-05	0.00013
Ornithine	RA-CGC	0.4497391	0.14030167	0.173195
Ornithine	TH-HYP	0.0288558	0.03103136	0.041482
Ornithine	RA-HYP	0.14963	0.15637509	0.190306
Ornithine	RA-TH	0.0000345	0.00017963	0.00038
Putrescine	HYP-CGC	0.1429626	0.02577924	0.035191
Putrescine	TH-CGC	0.2026575	0.17064441	0.20573
Putrescine	RA-CGC	0.000061	0.00020668	0.00043
Putrescine	TH-HYP	0.0006571	0.00208103	0.003603
Putrescine	RA-HYP	0	2.6379E-06	8.51E-06
Putrescine	RA-TH	0.038648	0.00430864	0.006948
Spermidine	HYP-CGC	0.0000593	1.2728E-06	4.56E-06

Absolute quantitative neurochemical brain atlas

Variable	Comparison	Tukey HSD	Paired T-test	FDR adj.
Spermidine	TH-CGC	0.00434	0.00366133	0.006038
Spermidine	RA-CGC	0.8038884	0.42072138	0.465863
Spermidine	TH-HYP	0.6212349	0.2898807	0.338413
Spermidine	RA-HYP	0.0012834	0.00176518	0.003119
Spermidine	RA-TH	0.0495354	0.01420687	0.020251
Spermine	HYP-CGC	0.999999	0.7111175	0.758195
Spermine	TH-CGC	0.0000004	0.00052039	0.001009
Spermine	RA-CGC	0.0004895	0.00044222	0.000871
Spermine	TH-HYP	0.0000004	0.00019483	0.000409
Spermine	RA-HYP	0.0004671	4.9895E-05	0.00012
Spermine	RA-TH	0.1879476	0.13648489	0.169294

Figure S1. Principal component analysis (PCA) of the 25 brain regions.



Chapter 4

Figure S2. Neurochemical metabolome of healthy rodent brains (n = 16). The column represents each individual brain region and the row represents the metabolites quantified. The data has been clustered by brain region based on the metabolic profile similarity and the metabolites have been clustered based on their concentration similarity.

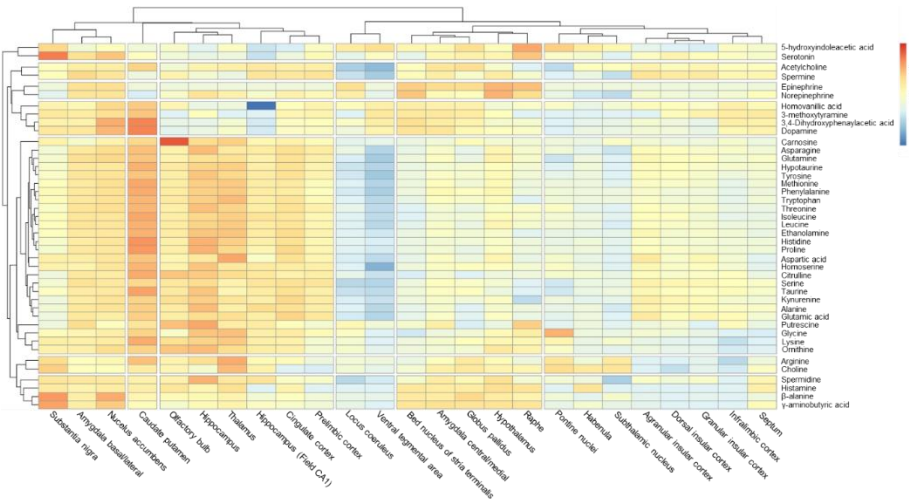


Table S8. Concentrations of metabolites from five brain regions: agranular insular cortex, amygdala (basal/lateral), amygdala (central/medial), bed nucleus of stria terminalis and field CA1 of hippocampus (ng/g wet tissue).

Metabolites	Agranular insular cortex	Amygdala (basal/lateral)	Amygdala (central/medial)	Bed nucleus of stria terminalis	Field CA1 of hippocampus
Acetylcholine	158.972 ± 24.9415	174.155 ± 39.054	185.96 ± 37.1674	135.215 ± 25.9035	183.979 ± 22.1108
Alanine	20420.7 ± 3000.06	27745.6 ± 5534.36	12358.5 ± 3026.47	8293.67 ± 1533.19	37355.7 ± 7680.86
Arginine	7912.27 ± 1965.3	13204.6 ± 3881.48	12753.8 ± 3601.68	9053.85 ± 2619.74	14307.8 ± 3212.39
Asparagine	5533.52 ± 1107.83	8400.29 ± 1974.44	4628.76 ± 1152.11	3359.36 ± 880.641	7788.9 ± 1882.06
Aspartic acid	156605 ± 65421.6	165076 ± 43719.2	105205 ± 33867.7	67302.4 ± 19963.6	114678 ± 39944.2
Beta Alanine	8.69351 ± 1.50709	23.9169 ± 5.77705	27.1273 ± 6.44169	28.1133 ± 9.93445	12.4067 ± 2.73043
Carnosine	885.961 ± 227.071	1045.15 ± 311.863	794.976 ± 161.697	636.325 ± 145.751	1589.97 ± 431.861
Choline	11719.6 ± 2473.44	17513.8 ± 4596.38	19086 ± 4513.93	16082.7 ± 4855.82	21496 ± 3550.78
Citrulline	9878.6 ± 2220.45	13390.7 ± 3931.37	7568.32 ± 2092.7	4476.52 ± 1087.49	13850.1 ± 2654.01
Dopamine	23.1735 ± 23.7747	164.397 ± 79.0223 0.982547 ±	190.337 ± 85.4983	216.385 ± 106.235	3.15742 ± 2.73132 0.0531321 ±
Epinephrine	0.00034076 ± 0	1.08419	1.25766 ± 1.32894	2.64968 ± 2.2492	0.186512
Ethanolamine	6257.54 ± 3168.77	10966 ± 4353.85	3608.02 ± 1282.28	1987.09 ± 664.009	10945.4 ± 5491.55
Gamma-Aminobutyric acid	41746 ± 5418.05	77595.3 ± 16720.9	91444.4 ± 14538.5	84275.1 ± 17789.4	57793.2 ± 10301.5
Glutamic acid	736144 ± 392360	727387 ± 191005	330401 ± 104720	202130 ± 57152.9	653966 ± 246743
Glutamine	230920 ± 38275.5	342059 ± 61039.7	188746 ± 40287.2	143678 ± 28532.6	254783 ± 54305.9
Glycine	21776.4 ± 4319.35	32165.5 ± 6902.9	18955.9 ± 3446.52	15036.5 ± 3576.7	34773.8 ± 8840.4
Histamine	1.21692 ± 4.86465	2.22502 ± 2.18496	7.53272 ± 5.47346	17.9942 ± 11.7084	1.86687 ± 7.46444
Histidine	5212.84 ± 1210.45	8547.77 ± 2416.86	4279.37 ± 924.335	3440.39 ± 744.522	7512.43 ± 2535.5
Homoserine	427.916 ± 72.2616	476.405 ± 134.768	350.479 ± 105.847	297.981 ± 103.374	453.063 ± 117.321
Homovanillic acid	11.9776 ± 3.00515 3.15727 ±	15.0576 ± 6.38347	23.6929 ± 9.78977	35.5936 ± 19.8978	0.708928 ± 1.13577
Hypotaurine	0.864003	5.09547 ± 1.26089	2.13711 ± 0.585498	1.76199 ± 0.639967	3.78816 ± 1.32983
Isoleucine	1285.51 ± 207.054	1926.48 ± 450.082	1086.83 ± 279.511	823.099 ± 227.839	1846.84 ± 476.466
Kynurenine	15.321 ± 14.8495	14.0112 ± 5.18582	7.37899 ± 3.96571	4.15748 ± 1.99057	12.6655 ± 4.14046
Leucine	3652.08 ± 627.218	5607.22 ± 1207.15	3064.59 ± 679.56	2300.3 ± 597.798	4960.07 ± 1215.25
Lysine	9397.6 ± 2221	16010.7 ± 4576.26	11105.8 ± 2837.07	9157.32 ± 2803.19	14934.2 ± 3803.93
Methionine	2814.17 ± 617.602	4282.02 ± 1096.27	2757.01 ± 647.53	2093.83 ± 599.463	3832.48 ± 1106.7
Norepinephrine	206.77 ± 49.0701	493.129 ± 125.124	286.102 ± 66.389	875.249 ± 333.371	226.541 ± 73.7065
Ornithine	535.912 ± 222.989	784.573 ± 163.797	587.466 ± 234.835	494.206 ± 108.433	773.364 ± 146.277
Phenylalanine	3420.93 ± 714.569	5410.67 ± 1267.99	3697 ± 822.432	2777.68 ± 676.394	4984.75 ± 1315.41
Proline	4719.67 ± 981.488	9730.82 ± 2933.9	3354.54 ± 754.124	2498.81 ± 597.644	6355.83 ± 2527.52
Putrescine	19.1361 ± 7.89752	35.6063 ± 9.68012	27.5947 ± 8.14714	19.8089 ± 2.98606	31.1389 ± 10.4186

Serine	38670.5 ± 6291.33	54249.4 ± 11084.6	27767.1 ± 6489.08	18289.5 ± 3805.68	49072.9 ± 8563.71
Serotonin	38.2679 ± 24.9346	65.4906 ± 26.1903	40.4537 ± 16.8427	29.98 ± 20.9473	16.6772 ± 8.70513
Spermidine	450.204 ± 158.561	783.427 ± 314.482	660.419 ± 262.441	434.302 ± 193.06	851.082 ± 350.243
Spermine	3949.01 ± 1434.65	5033.21 ± 2438.55	2324.76 ± 1059.74	1252.03 ± 657.357	3325.73 ± 1331.4
Taurine	184744 ± 30803	254478 ± 64011.3	126910 ± 22588.4	87975.3 ± 19075.8	274284 ± 64513
Threonine	153861 ± 29368.8	231371 ± 49848.5	126249 ± 32667.4	95971.1 ± 21033.4	192283 ± 38836.6
Tryptophan	1515.45 ± 303.823	2344.07 ± 507.363	1543.86 ± 321.592	1170.63 ± 281.265	2118.63 ± 509.549
Tyrosine	5502.74 ± 1292.39	8277.86 ± 1888.76	4824.1 ± 1015.07	3689.18 ± 981.707	7358.29 ± 2027.42
Valine	2261.26 ± 397.93	3414.61 ± 807.676	1833.4 ± 471.546	1474.91 ± 357.103	3195.73 ± 796.129
3-Methoxytyramine	0.595814 ±	2.83065 ±			0.159074 ±
3,4-Dihydroxyphenylacetic acid	0.975758	0.834558	3.5757 ± 1.27649	3.99317 ± 0.865987	0.412127
5-Hydroxyindoleacetic acid	35.8977 ± 21.1692	106.146 ± 52.7125	152.162 ± 58.0275	303.444 ± 112.566	7.16996 ± 3.58714
	1.85406 ±	2.13679 ±			
	0.421251	0.660286	3.0261 ± 0.595062	3.34928 ± 1.04111	1.2748 ± 0.1977

Table S9. Concentrations of metabolites from five brain regions: cingulate cortex, caudate putamen, dorsal insular cortex, granular insular cortex and globus pallidus (ng/g wet tissue).

Metabolites	Cingulate cortex	Caudate putamen	Dorsal insular cortex	Granular insular cortex	Globus pallidus
Acetylcholine	176.885 ± 34.2676	220.316 ± 38.8522	145.284 ± 37.7802	158.56 ± 28.4178	195.558 ± 50.0478
Alanine	31963.3 ± 5328.43	45741.6 ± 6919.51	17454.2 ± 3616.67	16642.5 ± 3769.31	7198.27 ± 2007.49
Arginine	12557.3 ± 2881.26	28727.4 ± 9177.13	6856.98 ± 1590.58	8028.89 ± 2132.5	18456.9 ± 8350.53
Asparagine	8976.77 ± 1152.61	11931.6 ± 2229.38	4972.42 ± 986.273	4542.75 ± 848.257	3140.12 ± 999.461
Aspartic acid	242580 ± 121980	362283 ± 244582	111314 ± 46117.4	146252 ± 68956	85695.4 ± 37196.6
Beta Alanine	16.5517 ± 2.01394	20.6594 ± 3.85917	8.08548 ± 1.80796	8.80726 ± 1.9214	999.461 ± 38.5794
Carnosine	1857.78 ± 483.785	3561.17 ± 831.296	705.068 ± 135.926	730.884 ± 144.615	11.6989 ± 1267.5
Choline	14205.4 ± 5407.54	35901.3 ± 33167.4	10264.2 ± 2579.23	11799.9 ± 3829.62	441.902 ± 21887.7
Citrulline	11473.9 ± 2180.14	24606.9 ± 4650.7	8208.9 ± 1956.48	8826.9 ± 2219.82	6311.49 ± 4797.56
Dopamine	26.6111 ± 13.2396	3412.47 ± 1005.27	13.2417 ± 8.53102	8.17184 ± 5.16257	1453.82 ± 56.7327
Epinephrine	0.00034076 ± 0	0.00034076 ± 0	0.0349825 ± 0.138567	0.00034076 ± 0	25.6333 ± 1.25413
Ethanolamine	19710.7 ± 12927.7	40233.8 ± 27932.5	3673.43 ± 1618.68	4298.78 ± 2709.41	1.07752 ± 3017.49
Gamma-Aminobutyric acid	54588.6 ± 8482.86	65634.3 ± 9007.05	37446.9 ± 6848.43	38884.5 ± 7335.29	1391.75 ± 92641
Glutamic acid	1.15645e+06 ± 545507	1.17599e+06 ± 404705	447415 ± 109268	513937 ± 132330	24230.7 ± 190966
Glutamine	338799 ± 52711.3	451707 ± 93930.4	201162 ± 42228.5	190067 ± 32161.6	96556.5 ± 132075

Absolute quantitative neurochemical brain atlas

Glycine	37058.5 ± 5700.22	52127.2 ± 11002	19521.4 ± 4012.77	17724 ± 4607.49	20421.2 ± 6463.02
Histamine	2.35283 ± 2.80255	2.81139 ± 0.980411	0.31583 ± 0.63934	0.310383 ± 0.794856	7.5352 ± 4.73455
Histidine	9963.9 ± 2310.46	23084.4 ± 4813.47	4506.48 ± 898.08	4301.96 ± 843.417	1271.82
Homoserine	454.695 ± 119.865	672.562 ± 174.082	344.634 ± 104.41	406.27 ± 81.5484	351.217 ± 156.268
Homovanillic.acid	9.82979 ± 3.79143	152.635 ± 39.3104	9.9473 ± 5.23554	6.83877 ± 3.23147	23.5218 ± 7.17038
Hypotaurine	5.48044 ± 1.26616	9.88593 ± 2.04759	0.666876	2.52763 ± 0.753303	1.91219 ± 0.675484
Isoleucine	1990.74 ± 344.357	3102.21 ± 568.966	1197.64 ± 313.193	1144.62 ± 325.459	990.836 ± 297.288
Kynurenine	23.87 ± 7.49545	25.3663 ± 9.29199	13.2583 ± 14.3705	10.5465 ± 6.12963	4.78938 ± 2.22949
Leucine	5628.27 ± 1076.45	8652.51 ± 1613.21	3158.88 ± 688.107	3120.35 ± 826.955	2674.87 ± 876.169
Lysine	14931.3 ± 3266.13	29567.5 ± 7676.46	8494.41 ± 2576.99	8279.85 ± 1941.02	13182.8 ± 4725.48
Methionine	4215.39 ± 902.402	7700.54 ± 1630.38	2442.53 ± 578.179	2496.47 ± 583.816	2644.49 ± 928.561
Norepinephrine	364.82 ± 59.6798	105.972 ± 23.0184	179.037 ± 48.7902	172.746 ± 37.139	263.374 ± 119.896
Ornithine	769.393 ± 130.63	1129.71 ± 230.296	513.932 ± 190.411	463.864 ± 127.389	499.881 ± 156.22
Phenylalanine	5085.28 ± 914.74	8516.99 ± 1678.68	2902.57 ± 615.635	2968.91 ± 710.085	3681.19 ± 1246.7
Proline	13589.1 ± 3635.66	28943.8 ± 5486.12	4008.8 ± 847.663	3737.96 ± 818.992	2581.04 ± 901.106
Putrescine	21.7945 ± 6.9398	31.9378 ± 7.61175	19.2621 ± 5.95057	15.5955 ± 4.46451	14.6196 ± 4.01968
Serine	53322.5 ± 8620.15	63921 ± 9616.6	33593.6 ± 6576.45	31969.7 ± 7382	23421.5 ± 7445.87
Serotonin	28.0253 ± 10.6664	34.3478 ± 7.99479	42.4078 ± 17.0353	36.9333 ± 11.5015	53.665 ± 27.2761
Spermidine	411.873 ± 147.486	627.91 ± 189.277	389.529 ± 142.548	441.358 ± 166.585	606.396 ± 350.322
Spermine	3064.93 ± 1440.93	577.772 ± 321.972	2821.81 ± 1154.71	2975.21 ± 1204.07	869.163 ± 490.549
Taurine	357446 ± 69998.6	879223 ± 166161	154259 ± 30812.6	142092 ± 24233.6	131377 ± 39774.3
Threonine	234543 ± 37766.3	340997 ± 56921.9	139802 ± 28815.6	133724 ± 30964.5	112650 ± 35887.5
Tryptophan	2380.93 ± 392.692	3775.72 ± 811.171	1325.92 ± 254.334	1358.2 ± 260.405	1532.75 ± 487.92
Tyrosine	8352.57 ± 1489.99	11941.8 ± 2612.9	4697.01 ± 1018.79	4610.54 ± 975.796	4552.55 ± 1419.19
Valine	3390.46 ± 506.811	4405.12 ± 716.511	2125.95 ± 587.514	1899.81 ± 543.598	1557.64 ± 393.828
3-Methoxytyramine	1.03721 ± 0.625142	15.4688 ± 3.70875	0.969123 ± 0.956178	1.24313 ± 0.970403	2.2528 ± 1.02326
3,4-Dihydroxyphenylacetic acid	44.6101 ± 14.1924	2584.21 ± 1447.61	20.8747 ± 7.77151	16.3223 ± 7.18972	95.7149 ± 35.999
5-Hydroxyindoleacetic acid	1.37456 ± 0.287738	2.01957 ± 0.48101	1.54948 ± 0.309255	1.50654 ± 0.280439	4.31572 ± 1.2141

Chapter 4

Table S10. Concentrations of metabolites from five brain regions: habenula, hippocampus, hypothalamus, infralimbic cortex and locus coeruleus (ng/g wet tissue).

Metabolites	Habenula	Hippocampus	Hypothalamus	Infralimbic cortex	Locus coeruleus
Acetylcholine	138.168 ± 43.6301	165.02 ± 19.5724	124.348 ± 32.1786	131.279 ± 23.4067	76.8679 ± 34.4889
Alanine	9861.23 ± 3303.42	52040.5 ± 7740.68	12150.8 ± 2852.52	14341.1 ± 3074.6	7004.92 ± 3155.62
Arginine	9534.58 ± 2879.58	23889 ± 9208.28	13313.4 ± 4219.52	5315.42 ± 1131.98	10711.1 ± 5664.62
Asparagine	2872.13 ± 858.762	12577.7 ± 2458.18	5971.57 ± 1385.47	4610.16 ± 940.931	3085.58 ± 1297.01
Aspartic.acid	68398.6 ± 26941.6	197864 ± 76019.1	142104 ± 60497.4	88115.4 ± 25959.9	56901.5 ± 23118
Beta.Alanine	8.29832 ± 2.9573	21.4994 ± 3.37471	27.049 ± 6.90186	7.71419 ± 1.91731	10.0583 ± 4.09067
Carnosine	652.886 ± 198.304	3242.25 ± 708.84	901.386 ± 223.326	895.362 ± 211.056	568.995 ± 261.535
Choline	20588.8 ± 5609.09	23319.5 ± 21114.3	18593.6 ± 5101.23	11200.4 ± 2844.8	14378.6 ± 5234.86
Citrulline	7545.54 ± 3061.68	18059.4 ± 3371.51	6197.47 ± 1520.1	8894.43 ± 2256.46	9791.95 ± 15641.4
Dopamine	8.3133 ± 5.26903 0.170341 ±	11.0573 ± 11.0047	31.8595 ± 15.0168	66.4429 ± 54.6684 0.286713 ±	9.44439 ± 4.20101
Epinephrine	0.427837	0.00034076 ± 0	13.4651 ± 5.89928	0.789978	1.19952 ± 1.1403
Ethanolamine	1794.87 ± 704.67	24201.7 ± 10566	3519.15 ± 1380.41	2820.77 ± 1042.83	1281.98 ± 506.393
Gamma.Aminobutyric.acid	48111.1 ± 13335.7	66716.8 ± 8432.94	86435.7 ± 18729.2	40337.8 ± 8448.07	46304 ± 17958.8
Glutamic.acid	248042 ± 81758.8	913402 ± 219150	367987 ± 119743	449663 ± 139960	150508 ± 63147
Glutamine	136255 ± 42974.3	347705 ± 57770.3	240249 ± 61866.3	170557 ± 30971.1	100171 ± 38702.5
Glycine	18834.8 ± 5371.71	56089.1 ± 8057.97	28688.7 ± 7246.27	16321.9 ± 3877.35	27209.6 ± 11030.7 0.165903 ±
Histamine	31.9683 ± 31.3848	2.99243 ± 6.54254	49.7542 ± 19.7137	1.72495 ± 6.67775	0.482703
Histidine	3452.33 ± 955.246	17478.6 ± 4203.01	5234.3 ± 1333.36	3973.9 ± 814.586	3478.14 ± 3008.33
Homoserine	293.176 ± 106.317	675.657 ± 135.457	364.284 ± 95.8522	357.934 ± 102.212	278.106 ± 117.518
Homovanillic.acid	1.58565 ± 1.40712	1.88106 ± 1.39842	3.35833 ± 1.83234 2.87818 ±	25.2676 ± 9.34704	11.8782 ± 8.5344
Hypotaurine	2.10924 ± 1.14796	6.64859 ± 2.57326	0.834111	2.19391 ± 0.58954	1.6441 ± 1.75978
Isoleucine	973.13 ± 293.239	2445.6 ± 497.43	1415.99 ± 404.631	921.589 ± 239.288	940.536 ± 618.208
Kynurenine	3.77544 ± 2.5795	26.656 ± 13.8735	6.96688 ± 2.86564	7.00106 ± 2.68412	4.57299 ± 8.94593
Leucine	2581.96 ± 734.521	7312.12 ± 1380.91	4355.54 ± 1178.41	2569.87 ± 532.767	2374.01 ± 1275.1
Lysine	8745.92 ± 2804.35	23501.1 ± 6415.74	13573.9 ± 3968.45	6696.12 ± 1521.44	10181.5 ± 4238.08
Methionine	1890.51 ± 556.412	5895.08 ± 1447.09	3210.28 ± 931.119	1988.54 ± 444.251	1744.77 ± 885.39
Norepinephrine	103.54 ± 55.0668	566.341 ± 89.6003	999.84 ± 274.484	216.104 ± 51.1323	402.347 ± 168.232
Ornithine	461.006 ± 132.254	1141.85 ± 211.925	829.383 ± 178.856	441.306 ± 164.044	641.028 ± 472.706
Phenylalanine	2584.44 ± 737.314	7172.83 ± 1542.6	4249.39 ± 1215.51	2519.93 ± 481.125	2313.05 ± 1217.06
Proline	2635.07 ± 741.731	21375.3 ± 5036.44	4424.24 ± 1215.15	3291.85 ± 631.772	2129.14 ± 1192.86
Putrescine	19.7033 ± 7.63821	47.9962 ± 12.4828	16.3951 ± 4.22596	22.5052 ± 6.21627	26.3023 ± 11.6753
Serine	16937.4 ± 5651.8	53195.2 ± 18206.4	20693.6 ± 5828.76	28437.2 ± 6130.65	10340.4 ± 6168.73
Serotonin	33.3762 ± 21.9242	22.8848 ± 11.0523	46.0678 ± 25.7497	31.6357 ± 12.7856	41.7417 ± 39.7936

Absolute quantitative neurochemical brain atlas

Spermidine	347.248 ± 160.279	1223.41 ± 325.251	893.652 ± 262.148	288.032 ± 114.994	216.448 ± 168.103
Spermine	675.306 ± 496.16	1103.67 ± 887.547	2864.24 ± 966.583	1943.31 ± 942.294	220.797 ± 183.648
Taurine	73285.3 ± 23928.4	567768 ± 113883	79730.1 ± 19118.2	146086 ± 28453.5	41047.1 ± 20256.7
Threonine	95874.2 ± 29435.3	308542 ± 53943.2	139969 ± 34226.8	115903 ± 23252.1	96878.6 ± 37932.1
Tryptophan	1148.85 ± 323.024	3249.14 ± 722.551	1755.03 ± 483.408	1142.61 ± 207.142	1071.82 ± 633.689
Tyrosine	3471.26 ± 1022.47	11146 ± 2308.85	5699 ± 1467.65	3975.1 ± 793.388	2887.27 ± 1498.56
Valine	1627.38 ± 503.157	4379.46 ± 661.883	2350.29 ± 612.083	1675.14 ± 429.542	1550.13 ± 861.186
	0.376592 ±	0.516739 ±			0.416013 ±
3-Methoxytyramine	0.878296	0.411636	1.04471 ± 1.05857	1.82324 ± 1.57208	0.983928
3,4-Dihydroxyphenylacetic acid	15.0785 ± 6.45539	12.6215 ± 4.61137	34.4959 ± 13.2209	141.318 ± 114.406	51.2535 ± 37.9151
		1.65129 ±	2.91746 ±	2.79776 ±	
5-Hydroxyindoleacetic acid	3.63961 ± 1.13172	0.320298	0.907593	0.928344	3.77981 ± 1.52331

Chapter 4

Table S11. Concentrations of metabolites from five brain regions: nucleus accumbens, olfactory bulb, pontine nuclei, prelimbic cortex and raphe (ng/g wet tissue).

Metabolites	Nucleus accumbens	Olfactory bulb	Pontine nuclei	Prelimbic cortex	Raphe
Acetylcholine	177.901 ± 38.4363	114.866 ± 33.275	88.4805 ± 27.41 9020.19 ±	173.629 ± 34.514	94.5811 ± 22.2539 9117.19 ±
Alanine	30622.9 ± 4853.34	26305.3 ± 11955.1	3447.32	24726.9 ± 4128.74	2247.88 9628.19 ±
Arginine	15653.4 ± 4481.52	17693.9 ± 9438.56	18867 ± 6296.07 3767.94 ±	9680.18 ± 2953.79	2934.15 4128.48 ±
Asparagine	9309.18 ± 1430.42	6685.86 ± 2837.42	1462.67 95934.6 ±	7623.51 ± 1217.04	1084.43 78975.6 ±
Aspartic acid	197207 ± 62116.5	157733 ± 86537.9	36816.2 12.1537 ±	154405 ± 67457.4	31940.6 22.1656 ±
Beta.Alanine	63.7059 ± 11.4387	23.4305 ± 16.288	4.94786 844.747 ±	11.6887 ± 1.80135	5.42069 822.183 ±
Carnosine	1823.51 ± 330.184	31456.9 ± 34994.7	325.997 23828.7 ±	1360.04 ± 256.101	214.301 18952.6 ±
Choline	19272.3 ± 5167.82	20974.4 ± 7868.5	8682.65 8122.23 ±	12343.4 ± 3314.43	3337.13 6469.82 ±
Citrulline	12849.8 ± 2409.92	20322 ± 13201.4	3729.87 3.58438 ±	12022 ± 2271.01	2269.41 20.7751 ±
Dopamine	1085.77 ± 365.483	16.268 ± 34.3918	2.35233 0.527166 ±	32.1129 ± 25.6579	7.63125 6.18125 ±
Epinephrine	1.07023 ± 1.19154	0.32484 ± 1.02108	1.03566 2093.92 ±	0.000791855	3.71557 1764.76 ±
Ethanolamine	12144.4 ± 5689.87	11490.7 ± 9701.95	796.816 44691.8 ±	8028.69 ± 4448.91	777.757 82140.8 ±
Gamma.Aminobutyric acid	110943 ± 16336.1	57504.8 ± 25083.9	15833.9 205776 ±	50065.1 ± 8815.82	18093.8 237309 ±
Glutamic acid	621347 ± 157150	552392 ± 247945	82744.7 95328.8 ±	730098 ± 205197	82670.2 164756 ±
Glutamine	372831 ± 51431.3	236507 ± 93195.3	28198.8 64799.1 ±	264763 ± 39769.9	45140.3 36296.5 ±
Glycine	35146.2 ± 6542.55	32443 ± 31899.2	23354.6	30047.3 ± 5871.53	9979.16 24.1797 ±
Histamine	5.80121 ± 2.69468	25.5235 ± 22.1847	0.000761343 ± 0 3066.27 ±	0.360308 ± 0.813683	15.9175 3834.56 ±
Histidine	9452.69 ± 2150.28	8869.12 ± 5357.03	1016.15 398.272 ±	6491.57 ± 1350.95	991.668 360.217 ±
Homoserine	489.315 ± 126.649	414.561 ± 122.392	187.874	474.065 ± 141.103	129.124 5.8841 ±
Homovanillic acid	97.8893 ± 12.0437	19.1944 ± 13.3579	8.9546 ± 6.81338 1.36863 ±	18.3685 ± 3.80686	2.11213 2.0168 ±
Hypotaurine	5.09311 ± 1.11932	5.03698 ± 2.84029	0.515183 1085.43 ±	3.65632 ± 0.624271	0.677665 1104.41 ±
Isoleucine	1981.27 ± 492.631	2528.16 ± 2749.92	406.931	1583.9 ± 373.844	282.61 2.46727 ±
Kynurenine	16.3416 ± 6.37375	29.1928 ± 31.9169	5.1026 ± 3.98981	13.7396 ± 3.98437	2.2517 3126.11 ±
Leucine	5604.7 ± 1363.97	6857.12 ± 7071.86	2719.51 ± 885.8 15275.4 ±	4440.07 ± 1022.83	720.187 13517.7 ±
Lysine	16234.2 ± 3479.83	19241.1 ± 8299.96	7345.07 2151.43 ±	10882.4 ± 2097.39	4453.84 2179.06 ±
Methionine	4933.08 ± 1256.05	4897.32 ± 3925.76	848.416 158.492 ±	3257.57 ± 690.664	635.242 653.029 ±
Norepinephrine	584.528 ± 167.016	210.904 ± 76.3982	62.8139 617.457 ±	305.637 ± 48.4211	182.841 691.672 ±
Ornithine	1020.61 ± 219.81	1141.66 ± 428.869	220.025 2766.18 ±	644.721 ± 296.96	171.991 2741.37 ±
Phenylalanine	6083.03 ± 1284.09	6093.27 ± 5070.92	958.964	4132.51 ± 875.479	695.159

Absolute quantitative neurochemical brain atlas

Proline	9782.41 ± 2568.14	11099.2 ± 8472.17	2522.08 ±	6815.52 ± 1494.58	2795.98 ±
			877.419		723.041
Putrescine	22.7081 ± 4.4809	39.666 ± 10.9008	14.9952 ±	21.6166 ± 6.54335	37.9712 ±
			5.45182		9.71125
Serine	44008.5 ± 7036.2	45363.1 ± 16938.1	11223.7 ±	47151.2 ± 6757.2	12933.7 ±
			3458.47		2748.92
Serotonin	92.0468 ± 35.7517	25.9371 ± 13.0941	35.9664 ±	37.8491 ± 12.3351	101.716 ±
			21.5807		64.708
Spermidine	807.423 ± 287.629	823.76 ± 675.48	403.971 ±	511.903 ± 198.303	493.291 ±
			239.648		233.968
Spermine	4005.21 ± 1430.34	2078.06 ± 1631.89	340.641 ±	4667.23 ± 2009.88	1034.64 ±
			227.895		619.939
Taurine	296200 ± 53220.6	284599 ± 136828	44648.4 ±	227958 ± 36335	47910.5 ±
			14484.5		11950.5
Threonine	242927 ± 32228.5	192072 ± 81097.5	147749 ±	194875 ± 39530.5	130205 ±
			60371.1		32486.1
Tryptophan	2557.7 ± 464.891	2474.95 ± 1288.68	1234.53 ±	1851.14 ± 339.05	1009.28 ±
			401.369		255.505
Tyrosine	8413.53 ± 1493.66	7325.54 ± 3159.13	3318.82 ±	6567 ± 1145.8	4036.74 ±
			1296.67		1028.79
Valine	3112.89 ± 570.189	3852.08 ± 3790.78	1806.05 ±	2739.9 ± 436	1962.56 ±
			744.705		527.715
3-Methoxytyramine	7.12739 ± 1.62212	0.988628	0.172435 ±	1.06161 ± 0.99864	1.18997 ±
			0.46557		1.77071
3,4-Dihydroxyphenylacetic acid	976.406 ± 270.726	25.4538 ± 32.9048	12.4972 ±	60.2958 ± 27.7897	35.0891 ±
			10.9419		14.8275
5-Hydroxyindoleacetic acid	0.302986	0.935789	6.37339 ±	2.49312 ± 0.556062	7.4911 ±
			2.91409		2.64128

Chapter 4

Table S12. Concentrations of metabolites from five brain regions: septum, substantia nigra, subthalamic nucleus, thalamus and ventral tegmental area (ng/g wet tissue).

Metabolites	Septum	Substantia nigra	Subthalamic nucleus	Thalamus	Ventral tegmental area
	135.554 ±	141.364 ±	147.818 ±	173.287 ±	87.7003 ±
Acetylcholine	42.3912	24.8703	43.9152	21.4608	94.8893
	12982.1 ±	11216.5 ±	6147.42 ±	24493.6 ±	
Alanine	3414.32	2175.87	2529.23	3423.05	5453.7 ± 5099.76
	8955.74 ±	22585.8 ±	17511.1 ±	33030.1 ±	10961.2 ±
Arginine	3095.52	7192.26	7453.56	7671.21	8721.87
	3768.94 ±	4064.67 ±	2289.36 ±	7862.78 ±	2467.68 ±
Asparagine	1120.7	776.878	916.053	898.556	2485.57
	64303.6 ±	173167 ±	67117.5 ±	342326 ±	57689.4 ±
Aspartic.acid	27286.1	62940.6	36924.1	130637	48912.7
	11.2629 ±	78.9253 ±	13.8556 ±	25.197 ±	11.5286 ±
Beta.Alanine	3.35513	15.2423	6.93014	7.89352	9.66957
	662.783 ±	1806.23 ±	1150.36 ±	3676.51 ±	782.564 ±
Carnosine	213.016	609.539	458.115	958.204	753.096
	15954.3 ±	26730.9 ±	21805.9 ±	33531.4 ±	
Choline	4480.36	4838.69	8203.34	9603.52	16716 ± 20216.4
	5770.69 ±	6333.88 ±	5743.94 ±	16554.4 ±	5046.36 ±
Citrulline	1642.02	1466.26	1994.69	3159.8	4650.74
	77.1255 ±	145.173 ±	7.15894 ±	22.2423 ±	88.1404 ±
Dopamine	35.6035	43.5657	5.73189	13.7155	74.0207
	0.412285 ±	0.403209 ±	0.281557 ±	0.436804 ±	0.728507 ±
Epinephrine	0.690775	0.764629	0.579828	0.665228	1.0525
	2443.5 ±	4634.79 ±	2164.32 ±	22680.5 ±	
Ethanolamine	1113.3	1884.06	1428.86	9248.76	1402.61 ± 1171.5
Gamma.Aminobutyric acid	63074.4 ±	134966 ±	42328.5 ±	74327.1 ±	51478.4 ±
	16469.4	20662.8	14926.8	9637.69	52295.4
	289570 ±	297975 ±	191460 ±	954283 ±	
Glutamic.acid	119812	107146	102379	284638	143840 ± 135923
	170356 ±	165700 ±	108103 ±	346080 ±	87735.4 ±
Glutamine	50139	28467.8	44853.2	37137.8	82090.7
	14981.5 ±	29182.5 ±	18785.9 ±	59092.2 ±	19185.5 ±
Glycine	4485.46	6364.23	6774.96	6508.26	16486.5
	3.62362 ±	9.3332 ±	3.78297 ±	50.4631 ±	
Histamine	4.05061	6.24084	6.11408	25.7226	10.05 ± 31.4831
	3650.09 ±	4884.31 ±	2900.89 ±	12861.5 ±	2960.27 ±
Histidine	1006.2	1139.92	1014.16	2567.07	2513.82
	310.669 ±	462.035 ±	283.443 ±	509.911 ±	244.941 ±
Homoserine	128.546	128.848	164.266	50.1137	207.314
	14.0249 ±	22.0263 ±	4.33919 ±	2.37363 ±	34.0534 ±
Homovanillic.acid	6.49057	8.61756	3.44579	1.24548	28.8892
	2.00881 ±	2.35249 ±	1.23895 ±	7.18556 ±	1.02985 ±
Hypotaurine	0.56206	0.620103	0.615422	1.51779	1.65152
	914.11 ±	1396.1 ±	808.772 ±	2591.02 ±	758.278 ±
Isoleucine	254.46	322.424	328.185	423.957	728.869
	5.17378 ±	15.4285 ±	3.46619 ±	15.0323 ±	11.6354 ±
Kynurenine	1.85068	41.1177	2.21863	7.25279	37.4386
	2501.34 ±	3689.04 ±	1937.21 ±	7361.63 ±	1898.58 ±
Leucine	702.445	740.998	762.465	1343.01	1747.54

Absolute quantitative neurochemical brain atlas

	7407.73 ±	18291.7 ±	11803.4 ±	25517.7 ±	8727.84 ±
Lysine	2234.78	4983.93	5161.64	5067.27	7420.91
		2889.94 ±	1696.12 ±	5907.28 ±	1490.79 ±
Methionine	2248 ± 673.652	749.116	801.627	1133.28	1448.39
	263.298 ±	139.279 ±	86.5197 ±	367.937 ±	190.817 ±
Norepinephrine	86.4125	44.258	47.7059	74.4545	195.356
	419.498 ±	744.614 ±	489.863 ±	1033.77 ±	532.557 ±
Ornithine	114.621	137.952	169.532	177.618	462.125
	2754.72 ±	4086.75 ±	2414.73 ±	7463.5 ±	1995.05 ±
Phenylalanine	830.157	938.891	1020.69	1280.36	1938.86
	2859.86 ±	3632.8 ±	1956.51 ±	13831.5 ±	1904.91 ±
Proline	824.756	862.246	828.304	3273.01	1848.16
	17.7438 ±	17.8071 ±	18.0273 ±	28.1669 ±	
Putrescine	5.64527	4.77939	10.107	10.7512	21.515 ± 11.8501
	24382.3 ±	21320.4 ±	16763.7 ±	38174.2 ±	9671.41 ±
Serine	7063.17	4478.35	5500.11	5858.35	8264.49
	29.2718 ±	157.249 ±	18.4931 ±	30.0821 ±	55.1209 ±
Serotonin	12.0165	77.5567	9.82114	16.9682	72.5741
	502.115 ±	776.144 ±	353.284 ±	810.628 ±	263.499 ±
Spermidine	235.007	296.104	437.31	419.101	157.306
	1654.02 ±	1112.98 ±	207.207 ±	875.303 ±	168.232 ±
Spermine	845.076	425.408	172.76	732.434	218.189
	114930 ±	137431 ±	80749.4 ±	215039 ±	43003.2 ±
Taurine	33681.4	27511.7	29851.1	46450.4	42608.9
	109252 ±	139060 ±	91888.6 ±	271433 ±	90372.6 ±
Threonine	32490.4	30663.2	32024.8	29823.2	92506.9
	1251.22 ±	1756.16 ±	1155.51 ±	3502.49 ±	
Tryptophan	383.286	382.713	471.815	469.987	916.966 ± 803.25
	3898.35 ±	4942.7 ±	2895.25 ±	9704.05 ±	
Tyrosine	1163.78	1040.97	1148.24	1238.91	2384.49 ± 2348
	1599.33 ±	2084.91 ±	1208.35 ±	3439.46 ±	1265.33 ±
Valine	434.511	441.15	405.322	385.001	1127.31
	2.90899 ±	3.77405 ±	0.515119 ±	0.535187 ±	4.01339 ±
3-Methoxytyramine	0.752505	0.740801	0.995858	0.542754	4.50581
3,4-					
Dihydroxyphenylacetic acid	102.334 ±	117.232 ±	15.8703 ±	30.1396 ±	
	35.5321	43.2602	8.03395	8.70261	173.8 ± 156.472
5-Hydroxyindoleacetic acid	2.14027 ±	4.77231 ±	2.96337 ±	2.55508 ±	5.45705 ±
	0.555773	0.705951	0.870568	0.383124	4.59809

Chapter 5

Mechanistic model-driven exometabolomic characterisation of human dopaminergic neuronal metabolism

German Preciat*, Jennifer Modamio*, Cornelius C.W. Willacey*, Agnieszka Wegrzyn, Edinson L. Moreno, Fatima L. Monteiro, Diana El Assal, Miguel A.P. Oliveira, Can Gulersonmez, Laurent Heirendt, Zhi Zhang, Ben Cousins, Hulda S. Haraldsdóttir, Susanne Zach, Siham Hachi, Amy Harms, Santosh Vempala, Bastian Hengerer, Jens C. Schwamborn, Enrico Glaab, Ines Thiele, Thomas Hankemeier[†] and Ronan M.T. Fleming[†]

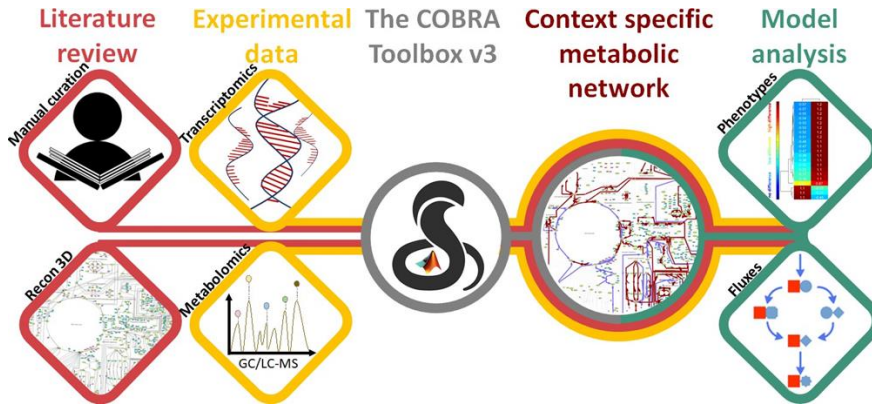
In preparation

^{#†} Authors equally contributed to the manuscript

Abstract

Patient-derived cellular models are a powerful approach to study human disease, especially neurodegenerative diseases, such as Parkinson's disease, where affected primary neurons, e.g., substantia nigra dopaminergic neurons, are almost inaccessible. Induced pluripotent stem cell-derived models of midbrain-specific dopaminergic neurons are increasingly used to investigate Parkinson's disease. Starting with the comprehensive generic reconstruction of human metabolism, Recon3D, we generated the first constraint-based, genome-scale, *in silico* model of human dopaminergic neuronal metabolism (iNESC2DN). Transcriptomic data, obtained by RNA sequencing, and quantitative exometabolomic data, obtained by targeted mass spectrometry-based metabolomics were generated for *in vitro* neuroepithelial stem cell-derived cultures and supplemented by extensive manual curation of the literature on dopaminergic neurons. The predictions of the iNESC2DN model are consistent with neurobiochemical prior information and in concordance with measured fluxes of uptake and secretion of many extracellular metabolites by dopaminergic neurons *in vitro*. We leverage it to rank order the most important metabolite concentrations to quantify to maximally reduce the uncertainty associated with current predictions of normal dopaminergic neuronal metabolism *in vitro*, as well as optimally design experiments to measure metabolic perturbations associated with Parkinson's Disease. Finally, the iNESC2DN model provides a foundation for future targeted metabolomic and tracer-based metabolomic analyses of dopaminergic neurons. This illustrates the synergy between constraint-based computational modelling of metabolism and biology-driven quantitative bioanalytical chemistry.

Mechanistic model of dopaminergic neuron metabolism



Background

Patient specific induced pluripotent stem cells (iPSCs)-derived, neuroepithelial stem cells (NESC)^{39, 46, 6, 17}, differentiated into neurons³³, offer an accessible approach to study neurodegenerative disorders *in vitro*. These neurons exhibit certain features, such as extensive arborisation and spontaneous electrophysiological activity¹⁹, that mimic nigrostriatal dopaminergic neurons, the cell type most vulnerable to degeneration in Parkinson's Disease (PD)²⁵. It has been hypothesised that this selective vulnerability is due to an imbalance between the high energy demand of, for example, maintaining tonic electrophysiological activity, and low energy supply as a result of, for example, mitochondrial dysfunction^{29,7}. Therefore, characterisation of the normal metabolic status of a dopaminergic neuron is of major interest but has not yet been reported.

*CO*nstraint-Based Reconstruction and Analysis (COBRA)²⁸ provides a mathematical and mechanistic computational modelling framework for experimental design, integrative analysis of prior biochemical knowledge with experimental data as well as the generation of novel hypotheses. In particular, quantitative bioanalytical chemistry^{27, 35, 32} has been effectively combined with constraint-based modelling of metabolism³ to enable context-specific biochemical interpretation of metabolomic data, e.g., to discover differences in glycolytic versus oxidative metabolism in different lymphoblastic leukaemia cell lines⁴, and to characterise metabolic changes influencing pluripotency and cell fate in stem cells⁹.

In this study, Recon3D⁸, the most comprehensive generic human metabolic reconstruction to date, was rendered context-specific by a combination of manual curation and omics data integration, to generate a constraint-based model of metabolism in human neuroepithelial stem cell-derived dopaminergic neurons, denoted iNESC2DN. Manual literature curation and transcriptomic data were used to establish the activity, or inactivity, of a core set of metabolic genes and reactions. In parallel, liquid chromatography-mass spectrometry (LC-MS) and gas chromatography-mass spectrometry (GC-MS) were used to quantify biogenic amines and organic acids in fresh and spent culture media from NESC-derived dopaminergic neurons in macroscopic cell culture. Different subsets of the obtained exometabolomic data, were used to refine the iNESC2DN model and test its predictions. The predicted metabolite uptake and secretion fluxes of the iNESC2DN

model were broadly consistent with bioanalytical quantification of metabolite consumption and secretion fluxes. A novel approach was developed to predict the most informative extracellular metabolites to target for future bioanalytical quantification as well as predict the effect of condition-specific metabolic perturbations as a mean to design future targeted metabolomic and tracer-based metabolomic experiments. Taken together, the iNESC2DN model provides a foundation for a systems approach to investigate metabolic dysfunction in patient-derived cellular models of PD, and the approach taken can serve as a template for the study of other neurodegenerative diseases.

Materials and methods

The following summary is complemented by essential methodological details as Supporting Information.

***In vitro* experiments.** The iPSC derived NESC were differentiated towards midbrain-specific dopaminergic neurons using an established protocol³³. Calcium imaging and automated image analysis using an established pipeline¹⁹ was used to assess electrophysiological activity at day 23 of differentiation (Figure S4). Additionally, at day 23 of differentiation, transcriptomic and exometabolomic data were generated from separate *in vitro* cultures using the same differentiation protocol³³ (Figure S4). Transcriptomic data was obtained by RNA-sequencing. Targeted exometabolomic data was generated from fresh and spent culture media, for 74 biogenic amines and amino acids, using an established LC-MS method²⁴, and for 24 organic acids by adapting an established GC-MS platform¹ (supplementary information section S1).

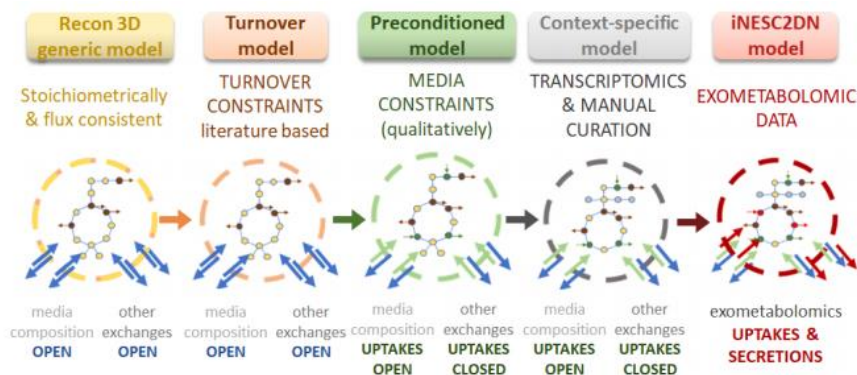


Figure 5.1: Overview of the model generation pipeline. From the Recon3D metabolic model⁸, a turnover model was generated by the integration of constraints representing the minimum cellular turnover of key metabolites. A preconditioned model was then generated by applying qualitative media constraints. Transcriptomic and manually curated data revealed the active and inactive reactions and genes in the cell culture and in dopaminergic neurons, which were integrated to generate a context-specific model using a model extraction algorithm⁴³. The final iNESC2DN model, used for design of future experiments, includes all exometabolomically derived constraints on uptake and secretion reactions.

Reconstruction. Following an established protocol⁴⁰, the generic human metabolic reconstruction, Recon2⁴¹, was refined with additional manual curation of metabolic literature specific to dopaminergic neurons, and included in an update to the generic human metabolic reconstruction, Recon3D⁸. Further manual curation was performed to define active and inactive reactions and genes, transport reactions, degradation pathways and quantitative constraints necessary to represent the requirement for molecular turnover in a non-growing, non-dividing dopaminergic neuron. When specific information on dopaminergic neurons was not present in the literature, information from other neuronal types, cerebral tissue, or rodent data was used (supplementary information section S2).

Model generation. A stoichiometrically consistent, flux consistent, constraint-based metabolic model, specific to *in vitro* NESC-derived dopaminergic neurons, was generated using the results of manual curation combined with transcriptomic and

exometabolomic data. Active and inactive genes, obtained from manual curation, or transcriptomic data, or both, were constrained in Recon3D, with manual curation given priority if a discrepancy arose⁴⁰. This integration was completed using the COBRA Toolbox¹⁶, a software tool for modelling genome-scale biochemical networks and integrative analysis of omics data in a network context. In particular, FASTCORE⁴³, was used as the model extraction algorithm. Models were refined by comparing biochemical literature with the results of Flux Balance Analysis²⁶. The workflow for model generation is illustrated in Figure 5.1 and described in more detail in supplementary information section S3 and section 4.

Model testing. Two test models were generated, termed ModelUpt and ModelSec, which included a subset of quantitative metabolomic data as constraints on uptake reaction fluxes or secretion reaction fluxes, respectively, while the excluded metabolomic data was used for comparison with model predictions. Flux Variability Analysis (FVA)²⁰ and uniform sampling¹⁵, were used to test the ability of these test models to predict the fluxes of extracellular secretion or uptake reactions, or both (supplementary information section S5). Uniform sampling provides a quantitative prediction of the probability of each quantitative flux value, between the same minimum and maximum flux predicted by flux variability analysis, assuming that each feasible steady-state flux vector is equi-probable.

Experimental design. Three distinct pipelines were developed that use the iNESC2DN model for experimental design. An *uncertainty reduction* pipeline rank orders exchanged metabolites according to those whose quantitative exometabolomic measurement would maximally shrink the feasible steady-state solution space. A *phenotypic perturbation* pipeline rank orders exchange reactions according to those whose rates are predicted to be most likely to change in response to a perturbation to an internal reaction rate. Finally, a *tracer-based metabolomic* pipeline was used to identify the non-elementary conserved moieties¹⁴ exchanged across the boundary of the iNESC2DN model that could be isotopically labelled to quantify the activity of metabolic pathways of specific importance to dopaminergic neurons (supplementary information section S6).

Results

Experimental characterisation. Differentiated neurons were identified by TUB β III immunoreactivity and those also positive for tyrosine hydroxylase indicated the presence of neurons capable of converting tyrosine to L-DOPA, the penultimate step in dopamine synthesis (Figure 5.6a). Analysis of calcium imaging data revealed spontaneously active neurons (Figure 5.6b, c, d). In the transcriptomic data, fragments were detected from 18,530 genes, but only 12,698 of these were sufficiently abundant to be considered expressed. That is, above a threshold of one Fragment per Kilobase of exon per Million reads³⁴. Of the expressed genes, 1,202 could be mapped to metabolic genes in Recon3D and were considered active, unless manual curation of the literature revealed otherwise. The selected metabolomic platforms target a total of 98 metabolites present in Recon3D. In the spent medium, only 50 metabolites were quantified above the lower limit of detection. However, the iNESC2DN model contains 49 metabolites with constraints on their corresponding exchange reaction fluxes as there was one two measured metabolite (Glutaric acid) that could not be integrated with the model as there are no stoichiometrically and flux consistent reactions that correspond to them in Recon3D (supplementary information section S7).

Dopaminergic neuronal reconstruction and model generation. Literature curation revealed evidence for the activity, or inactivity, of 252 metabolic genes (Table S-1) and 445 metabolic reactions (Table S-2) in dopaminergic neurons. Turnover constraints were added to represent the maintenance of a dopaminergic neuron (supplementary information section S4.3, Table S-2). Subsequently, differences in metabolite concentrations over time, were either converted into constraints on exchange reaction fluxes to generate a context-specific model, or kept independent from the model generation pipeline and used to test *in silico* model predictions (supplementary information section S7). Exometabolomic concentration changes for two metabolites (L-proline and serine), could not be directly integrated with the draft context-specific model as it became infeasible, therefore, as described in supplementary information section S4.6, relaxation of exometabolomic constraints on reactions corresponding to these two metabolites was required. The iNESC2DN model, i.e., the context-specific model using all exometabolomic constraints compatible with a feasible model, consists of 1,791 biochemical

reactions, between 828 unique metabolites, representing the activity of 1,853 metabolic genes from 90 biological pathways. In addition, the model contained 246 exchange reactions, 20 for uptake of metabolites from the media, 161 to secrete metabolites into spent media, 63 reversible exchange reactions (e.g., for transport of water), which were open, and 5 ionic external reactions, e.g., for sodium, calcium and potassium (Table S-3), which were closed as the model currently ignores ion transport associated with electrophysiological activity.

Model testing. A model generated using quantitative exometabolomic data on the uptake of metabolites (ModelUpt) could reasonably well quantitatively predict the flux of most secretion reactions, determined from exometabolomic data on metabolite secretion. Likewise, a model constrained with exometabolomic data on secretion reactions (ModelSec) could reasonably well predict the flux of most uptake reactions, determined from exometabolomic data. In both cases, the peak of the sample distribution for each exchange reaction, obtained from uniform sampling, was substantially better at quantitatively predicting the independent exometabolomic data, when compared with the ranges of exchange fluxes determined by flux variability analysis. Figure 5.2 illustrates representative comparisons for either uptakes or secretions, while Figures S12 and S13 illustrate comparisons for all reactions. In Figure 5.2, the measured secretion reaction fluxes were within the range predicted by flux variability analysis of ModelUpt for 26 metabolites, as in (a) and (b), and outside the range for 3 metabolites, as in (c). The measured uptake reaction fluxes were within the range predicted by flux variability analysis of ModelSec for 14 metabolites, as in (d) and (e), and outside the range for 5 metabolites, as in (f).

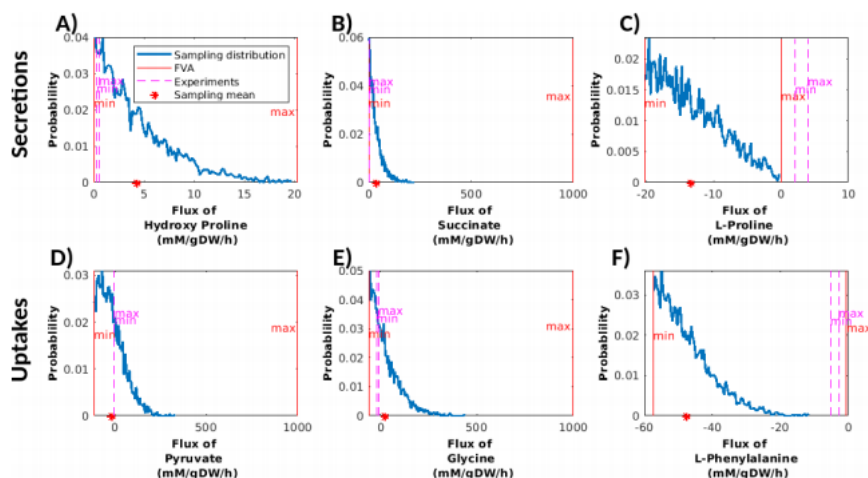


Figure 5.2: Comparison of predicted and measured metabolite exchange reaction rates. An uptake constrained model (ModelUpt) was tested for its ability to predict measured rates of 30 secreted metabolites, with three representatives illustrated in (a-c). A secretion constrained model (ModelSec) to test its ability to predict measured rates of 19 metabolites taken up from the fresh medium, with three representatives illustrated in (d-f). A measured range for each exchange reaction rate (pink) was obtained from quantitative exometabolomic measurements and includes one standard deviation of measurement uncertainty. Predicted probability of exchange reaction flux obtained by uniform sampling (dark blue). Predicted exchange reaction flux, derived from the mean of the sampling distribution (red star). Predicted maximum and minimum fluxes obtained by flux variability analysis (FVA).

Model characterisation. The iNESC2DN model has the potential to secrete 161 metabolites (Table S-2), including hydrophilic metabolites such as sugars, amino acids, carboxylic acids, keto acids, and nucleobases/nucleosides/nucleotides, while the lipophilic metabolites include free fatty acids, oxylipins, sterol lipids, sphingolipids, prenol lipids and fat soluble vitamins. The properties of these metabolites are an analytical chemistry consideration when selecting or developing targeted platforms for future exometabolomic experiments (Table S-3). Out of 161 metabolites predicted to be secreted, 17 were expected based on their assignment as active reactions during manual curation (Table S-2). A minimal set of reactions required to satisfy the constraints on the iNESC2DN model, e.g., turnover constraints,

Mechanistic model of dopaminergic neuron metabolism

is predicted to consist of 363 reactions (Table S-2). These reactions are involved in major metabolic pathways and pathways specific to neurons and dopaminergic neurons (Figure S14). Of the minimal reactions, about half (151/363) were manually curated to be active in dopaminergic neurons, with 32 involved the metabolism of dopamine. Twenty minimal reactions correspond to exchange reactions including 7 metabolites that can be taken up or secreted. The other 13 minimal uptake reactions predict the set of minimal medium metabolites for an *in vitro* dopaminergic neuron. These metabolites are glucose, the major source of energy, inorganic phosphate, ammonia, reduced glutathione, hydrogen carbonate and 15 amino acids, 9 of which are essential. Two amino acids, glutamine and arginine, are predicted to be conditionally essential with respect to dopaminergic neurons, as their uptake is essential for the feasibility of the model, but they can be synthesised by other tissues in the human body.

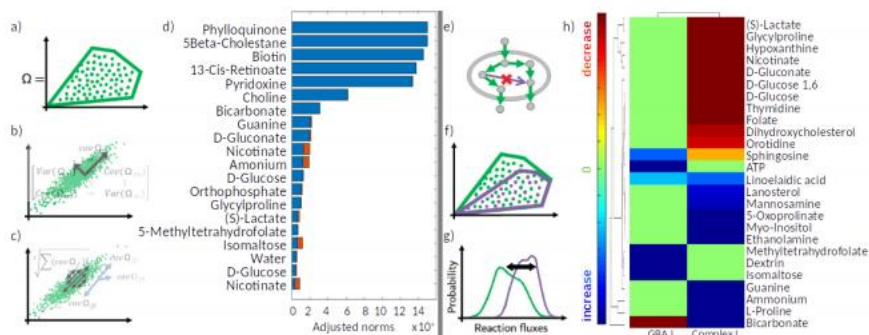


Figure 5.3: Experimental design. Uncertainty reduction. a) The steady-state flux space, $\Omega := \{v \in \mathbb{R}^n \mid Sv = 0, l \leq v \leq u\}$, of the iNESC2DN model was sampled. b) The covariance matrix of the sampled flux vectors $v \in \Omega$ was computed. c) The Euclidean norm for each row of the covariance matrix was calculated. d) The most informative exchange metabolites to measure were rank ordered by decreasing size of the Euclidean norm (blue), after taking into account the reduction in uncertainty (red) associated with measurement of higher ranked metabolites. The variance reduction due to cumulative measurement of higher ranked metabolite exchanges (orange) is taken into account in the ranking. Phenotypic perturbation. e) In the iNESC2DN model, certain internal reaction rates were perturbed, by changing reaction bounds, to represent, e.g., a gene deletion or a decrease in the maximum rate of a reaction. f) The steady-state flux space of the original and perturbed models are sampled. g) A two-sample Kolmogorov-Smirnov test was used to test for significant differences between the control and perturbed flux probability distributions. h) Significantly perturbed reactions were hierarchically clustered according to the magnitude of the increase (blue) or decrease (red), in the mean of the flux probability distribution for each exchanged metabolite.

Exometabolomic experimental design. Using the uncertainty reduction pipeline, we rank ordered 20 unmeasured exchange metabolites by the degree to which their measurement would shrink the feasible set of steady-state flux vectors for the iNESC2DN model (Figure 5.3a and Table S-3). The three top informative extracellular metabolites identified were phylloquinone (Vitamin K1, phyQ), 5-betacholestane-3-alpha (link), which is a bile acid synthesis pathway intermediate, and biotin (btn), which is a small vitamin molecule that acts as a cofactor in oxidative metabolism. The phenotypic perturbation pipeline predicted a set of exchange reactions that consistently vary as a result of knock-out of either the GBA1 gene, encoding

lysosomal and cytoplasmic glucocerebrosidase, or complete inhibition of mitochondrial complex 1 (Figure 5.3b).

Tracer-based metabolomic experimental design. A subset of the iNESC2DN model was atomically resolved using the COBRA Toolbox v3.0¹⁶. Specifically, a submodel was generated from the majority (1,091/1,533) internal reactions where a balanced atom mapping could be algorithmically predicted using the Reaction Decoder Tool³¹, including manual correction of R-group specification in appropriate substrate-reactant pairs, not previously done for the atom mappings reported in Recon3D⁸. No balanced atom mappings could be computed for 442 reactions in the iNESC2DN model, as at least one molecular structure was not available for each reaction or the corresponding reaction was unbalanced. In the submodel, a total of 215 conserved moieties, their corresponding chemical structures and moiety subnetworks were identified. Using this subModel we predicted the non-trivial conserved moieties associated with all of the metabolites that could be taken up from the fresh medium, which therefore could be used in future tracer-based metabolomic experiments. For example, a conserved moiety, with molecular formula N_4C_4 , is predicted to be taken up from the medium within hypoxanthine and is present in 90 different metabolites in the subModel including 5-Methylthioadenosine, AMP, ATP and hexanoyl coenzyme A, each of which have the potential to be secreted by the iNESC2DN model.

Discussion

Advances in constraint-based reconstruction and analysis. Completion of this study required several advances in constraint-based reconstruction and analysis. For example, this modelling approach is most commonly applied to biochemical systems where one predicts a feasible steady state flux vector that also satisfies a biologically motivated cellular objective, e.g., maximisation of biomass production flux for an exponentially growing culture of bacteria²⁶. However, neither substantia nigra dopaminergic neurons nor differentiated dopaminergic neurons divide, and it is not known what the cellular objective is for such neurons. Therefore, we added new constraints that enforce certain internal reactions, or combinations thereof, to operate above a certain flux, e.g., constraints on the turnover rate for metabolites and constraints representing the energetic requirements for biomass maintenance and electrophysiological signalling. As no cellular objective is assumed, uniform

sampling⁴² was applied, e.g., to reliably predict the sets of reaction fluxes that vary most in response to a PD relevant perturbation. This required the development and application of a novel algorithm, guaranteed to uniform sampling of the steady state solution space of high-dimensional metabolic networks¹⁵, such as those derived from Recon3D.

Another example of a novel advance in constraint-based reconstruction and analysis is our pipeline (Figure 5.1) to generate a constraint-based metabolic model of a non-dividing cell that starts with the most comprehensive generic metabolic network to date, Recon3D⁸, and integrates biochemical, transcriptomic, exometabolomic and manually curated data. It allows the generation of a variety of *in silico* models of neuronal metabolism, in a more comprehensive manner than previously described methods²³ and models¹⁸. The pipeline is sufficiently flexible that it can be used to generate context-specific, genome-scale metabolic models using data from dopaminergic neurons with different genetic backgrounds and different conditions, e.g., mitochondrial monogenic PD patient-derived cultures (e.g., PINK1) and isogenic control cultures exposed to mitochondrial stressors.

Biochemical interpretation of well predicted metabolic characteristics.

Variants of the iNESC2DN model performed well at quantitatively predicting metabolite secretion fluxes, given quantitative bounds on metabolite uptake fluxes, and at quantitatively predicting metabolite uptake fluxes, given quantitative bounds on metabolite secretion fluxes (Figure 5.2). Of the analysis methods tested, uniform sampling of steady state fluxes yielded the best predictions of quantitative secretion fluxes, especially for proline, putrescine and asparagine. The iNESC2DN model predicts the potential to uptake or secrete many metabolites that are not constrained by our quantitative exometabolomic data. Of the unmeasured metabolites predicted to be secreted by the iNESC2DN model, at least 30 are specifically associated with neuronal disorders (Table S-3), e.g., increased pyroglutamic acid is an indicator of glutathione deficiency and is associated with brain toxicity due to formation of amino acid adducts and dopamine quinones¹³.

All vulnerable neuronal populations in Parkinson's disease seem to either use monoamine neurotransmitters, such as dopamine (dopaminergic neurons within substantia nigra pars compacta), norepinephrine (noradrenergic neurons in locus coeruleus), and serotonin (serotonergic neurons in raphe nucleus), or produce

cytosolic monoamines, such as the cholinergic neurons of the dorsal motor nucleus of the vagus^{36, 25}. Also, high levels of cytosolic monoamines are hypothesised to underlie selective degeneration, since vulnerable neuronal populations generally include a catecholamine-derived neuromelanin pigment^{36, 44}. Consistent with phenylalanine being the precursor of monoamine neurotransmitters, the iNESC2DN model includes a high representation of reactions from the phenylalanine, tyrosine and tetrahydrobiopterin metabolism.

Most of the cellular phenotypic traits that are shared between vulnerable neuronal populations in Parkinson's disease, can be associated with a metabolic burden³⁶. Such neurons require a high supply of energy in order to meet the demand to tonically propagate action potentials over a large axonal arbour and for the synthesis, release and reuptake of neurotransmitters^{47, 36}. This intrinsic need to produce and consume a large amount of energy is thought to makes these neurons especially vulnerable to any impairment of energy metabolism^{47, 45}, therefore mitochondrial deficits could drive pathogenesis in Parkinson's disease^{38, 37, 36}. Consistent with this, in the iNESC2DN model we observe an increased representation of reactions related with oxidative phosphorylation, mitochondrial transport and the metabolism of cofactors, such as NAD metabolism.

The predicted minimal medium, which is the minimum number of metabolites required to be taken up by the model, consists of typical energetic substrates, essential amino acids and certain nonessential amino acids. In particular, L-glutamine is a non-essential amino acid that can be converted into nucleotides that then serve as a source of energy. Recently, a novel link has been described between glycolysis and mitochondrial dysfunction, which is mediated by reductive carboxylation of L-glutamine¹². A decrease in utilisation of reduced nicotinamide adenine dinucleotide (NADH) by the mitochondrial respiratory chain results in cytosolic reductive carboxylation of glutamine and thereby cytosol-confined NADH recycling. It is not known if this mechanism is of particular interest for PD. The minimal set of active reactions also predicted the activity of many reactions in dopamine metabolism, reflecting the importance of these reactions within the metabolic network. The set of minimal medium metabolites for an *in vitro* dopaminergic neuron provides a basis for the rational design of defined fresh medium specific for neuronal cell cultures⁵.

Biochemical interpretation of poorly predicted metabolic characteristics.

When disparate biochemical information from different experiments is integrated into a constraint-based model, they must be made consistent because inconsistent constraints will lead to an infeasible model, that is, one which does not admit any steady state flux. Therefore, it is important to achieve a balance between integration of further prior information, with the aim to improve quantitative predictions, and the risk of infeasibility due to inconsistency. As an example, the glutathione transferase (VMH link) reaction in dopamine metabolism was manually curated to be active since it is present in dopaminergic neurons¹¹, but it was excluded during model generation as it was not part of any flux consistent pathway in the generic Recon3D model. This indicates a metabolic pathway that requires future manual curation in the next iteration of the generic human metabolic reconstruction. Furthermore, qualitative metabolic predictions were made concerning secreted metabolites that may be important for dopaminergic neurons, but they were not tested because the targeted metabolomic platforms were initially chosen before the model existed.

Exometabolomic concentration changes for L-proline and serine, could not be directly integrated with the preconditioned model without making it inconsistent with the existence of a steady state flux. Recon3D allows reversible transport of the conditionally essential amino acid, L-proline. Prior to addition of exometabolomic data, the context-specific model includes extracellular transport reactions for L-proline, e.g., via proton symport PROT2r, but does not require secretion of L-proline, only either uptake or secretion is required. Therefore, when the exometabolomic data, which observes secretion of L-proline, is attempted to be integrated with the draft context-specific model, it may not be, and in this case is not feasible to obtain a steady state flux that secretes L-proline. Therefore, relaxation of the exometabolomic constraint, to permit L-proline uptake by the model, is required to render the model feasible. The situation is the same for serine, a non-essential amino acid. Essentially the common issue here is that data on the presence of gene products only provides information that the corresponding reaction may be active, but not the direction that the corresponding reaction is active in.

Relationship between *in vivo*, *in vitro* and *in silico*. Manual curation of the literature focused on quantification of neuronal molecular composition, turnover

fluxes, active genes, active reactions and inactive reactions specific to neurons, and substantia nigra dopaminergic neurons in particular. In parallel, we integrated transcriptomic and metabolomic data from human neuroepithelial stem cell derived neurons in macroscopic culture. As such, the iNESC2DN model is an *in silico* model that particularly emphasises the properties of human substantia nigra dopaminergic neurons and the properties of human neuroepithelial stem cell derived neurons³³. The macroscopic culture is a state of the art *in vitro* model of a human substantia nigra dopaminergic neuron *in vivo*. However, a single DN emanating from the substantia nigra is characterised by a massive axonal arbour²¹, much larger than other neuronal types, and projects to ~200k terminals in the striatum²⁹. In contrast, the *in vitro* neurons do have extensive neuronal projections, but not to the same extent as *in vivo*. Like this morphological divergence, there may be a molecular divergence between the *in vivo* neuron, on which manual curation was based, and on the *in vitro* neuronal culture used for generation of transcriptomic and metabolomic data, which is not pure culture of DN and may have a different extracellular metabolome. It will be interesting to compare this version of the iNESC2DN model with future versions generated using protocols already in development for generation of higher purity dopaminergic neuronal cultures.

Exometabolomic experimental design. Algorithmic experimental design was used to propose designs that optimise the information obtained in future exometabolomic and tracer-based metabolomic experiments. Algorithmic design of exometabolomic experiments enables optimal selection and development of targeted mass spectrometry platforms for future analyses. This is important as one targeted analytical platform cannot quantify the concentration of all of the metabolites within the iNESC2DN model (supplementary information section S15). Our uncertainty reduction pipeline rank orders unmeasured exchanged metabolites by the degree to which their measurement would shrink the feasible set of steady-state flux vectors. The top ranked metabolites include biotin, which is known to be enriched in select areas of the central nervous system, including the substantia nigra²². Phylloquinone interacts with the N-terminus of alpha-synuclein, inhibits fibril formation *in vitro* and is being investigated with a view toward development of new therapies targeting alpha-synuclein aggregation¹⁰.

Design of tracer-based metabolomic experiments. The pipeline for tracer-based metabolomic experiment design is hampered by the absence of molecular structures for some reactants³⁰, e.g., those with R groups in the structure, as they precluded the atomic resolution of all reactions in the iNESC2DN model. However, it was possible to atom map the majority of internal reactions, which permitted the identification of the majority of conserved moieties¹⁴ in the iNESC2DN model. Identification of conserved moieties has strong potential for use in design of tracer-based metabolomic experiments⁴⁶. By isotopically labelling any single atom in a conserved moiety, one can use the iNESC2DN model to predict the reachable set of metabolites that could contain that isotopic label, or any other isotopically labelled atom in the same conserved moiety. For a single conserved moiety, this approach for the design of an isotopic labelling strategy has been explored with the related concept of an elementary metabolic unit². This will facilitate future study of metabolic pathways particularly significant for identified by our exometabolomic approach in more detail.

Conclusions

We have developed the first, mechanistic, genome-scale, metabolic model of a pluripotent stem cell derived dopaminergic neuronal culture, denoted iNESC2DN. It combines extensive manual curation of biochemical literature with genome-scale quantification of transcripts and extracellular metabolite concentration changes. The model also atomically resolves metabolic transformations at genome-scale. Variants of the model, tested against subsets of independent exometabolomic data, could quantitatively predict metabolite uptake and secretion fluxes for many fresh and spent medium metabolites. With a view towards future metabolomic experiments to refine the model in an iterative systems biology cycle, we demonstrate its utility for experimental design of targeted metabolomic and tracer-based metabolomic experiments. As such, the iNESC2DN model establishes a solid foundation for comparative analyses of neuroepithelial stem cell derived dopaminergic neurons from PD patients and controls via mechanistic model-driven metabolomic and tracer-based metabolomic approaches, and we expect, that this strategy will be very useful also for other neurodegenerative diseases.

Acknowledgements

The authors would like to thank Sylvain Arreckx and Thomas Pfau for helping with computational issues, Maïke Aurich for helping with MetaboTools, Swagatika Sahoo explanation of the reconstruction process. ELM, MO, and DE were supported Aides à la Formation-Recherche Training allowances granted to by the Fonds National de la Recherche Luxembourg. ZZ and EG were supported by the Fonds Nationale de la Recherche, Luxembourg, as part of the BMBF-funded e:Med project MitoPD (INTER/BMBF/13/04) and the CORE INTER project MiRisk-PD (FNR11676395). LH has been supported by the FNR OPEN Grant (FNR/O16/11402054). The bioinformatic analyses presented in this paper were carried out in part using the HPC facilities of the University of Luxembourg (see <http://hpc.uni.lu>). The provision of NESC from the StemBANCC project is gratefully acknowledged. JM, ELM, GP, JS, CW, TH and RF received funding from the European Union's Horizon 2020 research and innovation programme, for the project SysMedPD, under grant agreement no. 668738.

References

1. Rodrigo D. A.M. Alves et al. "Global Profiling of theMuscleMetabolome:Method Optimization, Validation and Application to Determine Exercise-Induced Metabolic Effects". In: *Metabolomics* 11.2 (Apr. 1, 2015), pp. 271–285. doi: 10.1007/s11306-014-0701-7.
2. Maciek R. Antoniewicz, Joanne K. Kelleher, and Gregory Stephanopoulos. "Elementary Metabolite Units (EMU): A Novel Framework for Modeling Isotopic Distributions". In: *Metabolicengineering* 9.1 (Jan. 2007), pp. 68–86. doi: 10.1016/j.ymben.2006.09.001.
3. Maik K. Aurich, Ronan M. T. Fleming, and Ines Thiele. "MetaboTools: A Comprehensive Toolbox for Analysis of Genome-Scale Metabolic Models". In: *Frontiers in Physiology* 7 (2016). doi: 10.3389/fphys.2016.00327.
4. Maik K. Aurich et al. "Prediction of Intracellular Metabolic States from Extracellular Metabolomic Data". In: *Metabolomics* 11.3 (Aug. 14, 2014), pp. 603–619. doi: 10.1007/s11306-014-0721-3.
5. Cedric Bardy et al. "Neuronal Medium That Supports Basic Synaptic Functions and Activity of Human Neurons in Vitro". In: *Proceedings of the National Academy of Sciences of the United States of America* 112.20 (May 19, 2015), E2725–E2734. doi: 10.1073/pnas.1504393112. pmid:25870293.
6. Milena Bellin et al. "Induced Pluripotent Stem Cells: The New Patient?" In: *Nature Reviews Molecular Cell Biology* 13.11 (Nov. 2012), pp. 713–726. doi: 10.1038/nrm3448.
7. J Paul Bolam and Eleftheria K Pissadaki. "Living on the Edge with Too Many Mouths to Feed: Why Dopamine Neurons Die". In: *Movement Disorders* 27.12 (Sept. 2012), pp. 1478–1483. doi: 10.1002/mds.25135. pmid: 23008164.
8. Elizabeth Brunk et al. "Recon3D Enables a Three-Dimensional View of Gene Variation in Human Metabolism". In: *Nature Biotechnology* 36 (Feb. 19, 2018), p. 272. url: <http://dx.doi.org/10.1038/nbt.4072>.
9. Sriram Chandrasekaran et al. "Comprehensive Mapping of Pluripotent Stem Cell Metabolism Using Dynamic Genome-Scale Network Modeling". In: *Cell Reports* 21.10 (Dec. 5, 2017),pp. 2965–2977. doi: 10.1016/j.celrep.2017.07.048. pmid: 29212039.

10. Fernanda Luna da Silva et al. "Vitamins K Interact with N-Terminus - Synuclein and Modulate the Protein Fibrillization in Vitro. Exploring the Interaction between Quinones and alpha-Synuclein". In: *Neurochemistry International* 62.1 (Jan. 1, 2013), pp. 103–112. doi: 10.1016/j.neuint.2012.10.001.
11. Alexies Dagnino-Subiabre et al. "Glutathione TransferaseM2-2 Catalyzes Conjugation of Dopamine and Dopa o-Quinones". In: *Biochemical and Biophysical Research Communications* 274.1(July 21, 2000), pp. 32–36. doi: 10.1006/bbrc.2000.3087.
12. Edoardo Gaude et al. "NADH Shuttling Couples Cytosolic Reductive Carboxylation of Glutamine with Glycolysis in Cells with Mitochondrial Dysfunction". In: *Molecular Cell* 69.4 (Feb. 15,2018), 581–593.e7. doi: 10.1016/j.molcel.2018.01.034. pmid: 29452638,29452638.
13. Greenblatt, James M. and Brogan, Kelly. *Integrative Therapies for Depression: Redefining Models for Assessment, Treatment and Prevention*. Dec. 1, 2015. url: <https://w/www.crcpress.com / Integrative - Therapies - for - Depression - Redefining - Models - for - Assessment /Greenblatt-Brogan/p/book/9781498702294> (visited on 02/01/2018).
14. Hulda S. Haraldsdóttir and Ronan M. T. Fleming. "Identification of Conserved Moieties in Metabolic Networks by Graph Theoretical Analysis of Atom Transition Networks". In: *PLOS Computational Biology* 12.11 (Nov. 21, 2016), e1004999. doi: 10.1371/journal.pcbi.1004999.
15. Hulda S. Haraldsdóttir et al. "CHRR: Coordinate Hit-and-Run with Rounding for Uniform Sampling of Constraint-Based Models". In: *Bioinformatics* 33.11 (Jan. 6, 2017), pp. 1741–1743. doi: 10.1093/bioinformatics/btx052.
16. Laurent Heirendt et al. "Creation and Analysis of Biochemical Constraint-Based Models: The COBRA Toolbox v3.0". In: *Nature Protocols (accepted)* (2018). url: <https://arxiv.org/abs/1710.04038> (visited on 07/28/2017).
17. Ji-feng Kang, Bei-sha Tang, and Ji-feng Guo. "The Progress of Induced Pluripotent Stem Cells as Models of Parkinson's Disease". In: *Stem Cells International* 2016 (2016). doi: 10.1155/2016/4126214. pmid: 26880962.
18. Nathan E Lewis et al. "Large-Scale in Silico Modeling of Metabolic Interactions between Cell Types in the Human Brain". In: *Nature Biotechnology* 28.12 (Dec. 2010), pp. 1279–1285. doi:10.1038/nbt.1711.

19. E. Lucumi Moreno et al. "Differentiation of Neuroepithelial Stem Cells into Functional Dopaminergic Neurons in 3D Microfluidic Cell Culture". In: *Lab on a Chip* 15.11 (May 20, 2015), pp. 2419–2428. doi: 10.1039/C5LC00180C.
20. R. Mahadevan and C. H. Schilling. "The Effects of Alternate Optimal Solutions in Constraint-Based Genome-Scale Metabolic Models". In: *Metabolic Engineering* 5.4 (Oct. 2003), pp. 264–276. doi: 10.1016/j.ymben.2003.09.002.
21. Wakoto Matsuda et al. "Single Nigrostriatal Dopaminergic Neurons Form Widely Spread and Highly Dense Axonal Arborizations in the Neostriatum". In: *The Journal of Neuroscience* 29.2 (Jan. 14, 2009), pp. 444–453. doi: 10.1523/JNEUROSCI.4029-08.2009. pmid: 19144844.
22. Bruce E. McKay, Michael L. Molineux, and Ray W. Turner. "Biotin Is Endogenously Expressed in Select Regions of the Rat Central Nervous System". In: *The Journal of Comparative Neurology* 473.1 (May 17, 2004), pp. 86–96. doi: 10.1002/cne.20109.
23. Monica L. Mo, Bernhard Ø Palsson, and Markus J. Herrgård. "Connecting Extracellular Metabolomic Measurements to Intracellular Flux States in Yeast". In: *BMC Systems Biology* 3.1 (Mar. 25, 2009), p. 37. doi: 10.1186/1752-0509-3-37. pmid: 19321003.
24. Marek J. Noga et al. "Metabolomics of Cerebrospinal Fluid Reveals Changes in the Central Nervous System Metabolism in a Rat Model of Multiple Sclerosis". In: *Metabolomics* 8.2 (Apr. 2011), pp. 253–263. doi: 10.1007/s11306-011-0306-3.
25. Miguel A. P. Oliveira et al. "Embryonic Development of Selectively Vulnerable Neurons in Parkinson's Disease". In: *npj Parkinson's Disease* 3.1 (June 26, 2017), p. 21. doi: 10.1038/s41531-017-0022-4.
26. Jeffrey D. Orth, Ines Thiele, and Bernhard Ø Palsson. "What Is Flux Balance Analysis?" In: *Nature Biotechnology* 28.3 (Mar. 2010), pp. 245–248. doi: 10.1038/nbt.1614.
27. G. Paglia et al. "Monitoring Metabolites Consumption and Secretion in Cultured Cells Using Ultra-Performance Liquid Chromatography Quadrupole-Time of Flight Mass Spectrometry (UPLC-Q-ToF-MS)." In: *Analytical and Bioanalytical Chemistry* 402.3 (2012), pp. 1183–98. doi:10.1007/s00216-011-5556-4. pmid: 22159369.

28. Bernhard Ø Palsson. *Systems Biology: Constraint-Based Reconstruction and Analysis*. Cambridge, England: Cambridge University Press, Jan. 26, 2015. 550 pp.
29. Eleftheria K. Pissadaki and J. Paul Bolam. "The Energy Cost of Action Potential Propagation in Dopamine Neurons: Clues to Susceptibility in Parkinson's Disease". In: *Frontiers in Computational Neuroscience* 7 (Mar. 18, 2013), p. 13. doi: 10.3389/fncom.2013.00013. pmid:23515615.
30. German A. Preciat Gonzalez et al. "Comparative Evaluation of Atom Mapping Algorithms for Balanced Metabolic Reactions: Application to Recon3D". In: *Journal of Cheminformatics* 9 (2017), p. 39. doi: 10.1186/s13321-017-0223-1.
31. Syed Asad Rahman et al. "Reaction Decoder Tool (RDT): Extracting Features from Chemical Reactions". In: *Bioinformatics* 32.13 (Jan. 7, 2016), pp. 2065–2066. doi: 10.1093/bioinformatics/btw096. pmid: 27153692.
32. Rawi Ramautar et al. "Human Metabolomics: Strategies to Understand Biology". In: *Current Opinion in Chemical Biology. In Vivo Chemistry • Analytical Techniques* 17.5 (Oct. 2013), pp. 841–846. doi: 10.1016/j.cbpa.2013.06.015.
33. Peter Reinhardt et al. "Derivation and Expansion Using Only Small Molecules of Human Neural Progenitors for Neurodegenerative Disease Modeling". In: *PLoS ONE* 8.3 (Mar. 22, 2013). Ed. By Marcel Daadi, e59252. doi: 10.1371/journal.pone.0059252.
34. Kirti Sharma et al. "Cell Type- and Brain Region-Resolved Mouse Brain Proteome". In: *Nature Neuroscience* advance online publication (Nov. 2, 2015). doi: 10.1038/nn.4160.
35. P. A. Steininger et al. "Change of the Metabolomic Profile during Short-Term Mononuclear Cell Storage". In: *Vox Sanguinis* (Jan. 1, 2017), n/a–n/a. doi: 10.1111/vox.12482.
36. David Sulzer and D. James Surmeier. "Neuronal Vulnerability, Pathogenesis and Parkinson's Disease". In: *Movement disorders : official journal of the Movement Disorder Society* 28.1 (Jan. 2013), pp. 41–50. doi: 10.1002/mds.25095. pmid: 22791686.

37. D. James Surmeier, José A. Obeso, and Glenda M. Halliday. "Selective Neuronal Vulnerability in Parkinson Disease". In: *Nature Reviews Neuroscience* 18.2 (Jan. 20, 2017), pp. 101–113. doi:10.1038/nrn.2016.178.
38. D. James Surmeier and David Sulzer. "The Pathology Roadmap in Parkinson Disease". In: *Prion* 7.1 (Jan. 1, 2013), pp. 85–91. doi: 10.4161/pri.23582.
39. Kazutoshi Takahashi et al. "Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors". In: *Cell* 131.5 (Nov. 30, 2007), pp. 861–872. doi: 10.1016/j.cell.2007. 11.019.
40. Ines Thiele and Bernhard Ø Palsson. "A Protocol for Generating a High-Quality Genome-Scale Metabolic Reconstruction". In: *Nature Protocols* 5.1 (Jan. 2010), pp. 93–121. doi: 10.1038/nprot.2009.203.
41. Ines Thiele et al. "A Community-Driven Global Reconstruction of Human Metabolism". In: *Nature Biotechnology* 31.5 (May 2013), pp. 419–425. doi: 10.1038/nbt.2488.
42. Ines Thiele et al. "Candidate Metabolic Network States in Human Mitochondria. Impact of Diabetes, Ischemia, and Diet." In: *Journal of Biological Chemistry* 280.12 (Mar. 25, 2005), pp. 11683–11695. doi: 10.1074/jbc.M409072200.
43. Nikos Vlassis, Maria Pires Pacheco, and Thomas Sauter. "Fast Reconstruction of Compact Context-Specific Metabolic Network Models". In: *PLoS Comput Biol* 10.1 (Jan. 16, 2014), e1003424. doi: 10.1371/journal.pcbi.1003424.
44. Eberhard Weihe et al. "Three Types of Tyrosine Hydroxylase-Positive CNS Neurons Distinguished by Dopa Decarboxylase and VMAT2 Co-Expression". In: *Cellular and molecular neurobiology* 26.0 (2006), pp. 659–678. doi: 10.1007/s10571-006-9053-9. pmid: 16741673.
45. Peter Wellstead and Mathieu Cloutier. "An Energy Systems Approach to Parkinson's Disease". In: *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 3.1 (Jan. 1, 2011), pp. 1–6. doi: 10.1002/wsbm.107.
46. Wolfgang Wiechert. "13C Metabolic Flux Analysis". In: *Metabolic Engineering* 3.3 (July 2001), pp. 195–206. doi: 10.1006/mben.2001.0187.
47. Weihai Ying. "NAD⁺ and NADH in Brain Functions, Brain Diseases and Brain Aging". In: *Frontiers in Bioscience: A Journal and Virtual Library* 12 (2007), pp. 1863–1888. pmid: 17127427.

Supporting Information

Part I

Methods

S1 Experiments

S1.1 Cell culture

An overview of the experimental approach is given in Figure S4.



Figure S4: Experimental protocol overview.

Human neuroepithelial stem cells (hNESC) were differentiated into midbrain dopaminergic neurons. The cell number in each culture well was counted on day 1, 13, 19 and estimated for day 23. Spent media samples for metabolomic analyses were collected at days 10, 13, 19 and 23. Samples were analysed with both GC-MS and LC-MS. At day 23, live cells were subjected to calcium imaging followed by immunostaining assays, and collection of parallel samples for transcriptomic analysis. The media composition at the various stages of cell culture were as follows; Maintenance stage (red): maintenance medium containing ascorbic acid, purmorphamine (PMA) and the aminopyrimidine CHIR-99021(CHIR). Differentiation stage (green): differentiation medium containing ascorbic acid, Brain-derived neurotrophic factor (BDNF), glial cell-derived neurotrophic factor (GDNF), Transforming Growth Factor Beta 3 ($TGF\beta 3$), dbcAMP and PMA. Maturation stage (blue): differentiation media without PMA.

S1.1.1 Human neuroepithelial stem cell-derived dopaminergic neuronal differentiation.

A human neuroepithelial stem cell line from a healthy human donor (Identifier: 3.0.0.10.0 Acronym: hNEscs K7/ NPBScs/NEs, wild-type) was maintained and differentiated into DN_s, using an established protocol⁵⁴, summarised below.

N2B27 Medium preparation. The culture medium, denoted *N2B27 medium*, was used as the basis to prepare both maintenance and differentiation media and was

Mechanistic model of dopaminergic neuron metabolism

obtained by mixing equal amounts of Neurobasal medium (Invitrogen/Life Technologies) and DMEM/F12 medium (Invitrogen/Life Technologies) supplemented with 1% penicillin and streptomycin (Life Technologies), 2 mM L-glutamine (Life Technologies), 0.5 X B27 supplement without Vitamin A (Life Technologies) and 0.5 X N2 supplement (Life Technologies). The final concentration of the media composition is fully detailed in Table S-4.

Plate coating. Nunc cell-culture treated 6-well plates (ThermoFisher scientific, Roskilde, Denmark) were coated with 1% Matrigel (Discovery Labware, Inc., Two Oak Park, Bedford, MA, USA, Catalogue number 354277, lot number 3318549) in 600 μ L of knockout DMEM (1X) medium.

Cell seeding and maintenance. At the time of cell seeding, the knockout DMEM (1X) medium from the coating step, was removed from each well and the K7 hNESC line was seeded in three replicate wells. The medium to maintain the hNESC in culture, denoted *maintenance medium* (red in Figure S4: Overview of the experimental protocol), is based on N2B27 medium with 0.5 μ M PMA (Enzo life sciences), 3 μ M CHIR (Axon Medchem) and 150 μ M ascorbic acid (Sigma Aldrich). The cell seeding was done by preparing 5×10^6 million cells/mL in 50% matrigel in maintenance medium and adding 200 μ L of this preparation to obtain approximately 0.2 mm or 200 μ m thick layer of cells in three dimensions within Matrigel, with 4×10^5 cells per well. After the Matrigel and cell mixture was added to the well, the plate was incubated for 2 min at 37 °C to gelate the matrigel layer, the plate was then taken out of the incubator and 2.8 mL of maintenance medium was added and the plate was incubated at 37 °C and 5% CO₂ for 48 h.

Neuronal differentiation and maturation. The *differentiation medium with PMA* preparation to induce the differentiation of hNESC towards midbrain dopaminergic neurons consisted of N2B27 medium with 200 μ M ascorbic acid, 0.01 ng/ μ L BDNF (Peprotech), 0.01 ng/ μ L GDNF (Peprotech), 0.001 ng/ μ L TGF β 3 (Peprotech), 2.5 μ M dbcAMP (Sigma Aldrich) and 1 μ M PMA. This medium preparation was completely replaced every 2 days during the next 6 days of culture in the differentiation process. For the maturation of differentiated neurons, PMA is required to be absent from the differentiation medium. This *differentiation medium*

without PMA was used from day 9 onwards and 50% media replacement every 2 days for 3 weeks.

S.1.1.2 Microscopy and calcium imaging

To monitor cellular morphology during differentiation, bright field images were acquired every 48h for 23 days of differentiation using a Zeiss Axiovert 40 CFL microscope equipped with a cooled charge-coupled device based camera (Zeiss AxioCam MRm, Zeiss). At day 23 in culture, calcium imaging was done with a Fluo-4 AM green-fluorescent calcium indicator dye. After removing the differentiation medium, 1 mL of 5 μ M cell permeant Fluo-4 AM (Invitrogen/Life Technologies, F14201) in neurobasal medium, was added to selected wells of a 6-well plate at room temperature. Full frame fluorescence images, of size 2560 $\times$ 2160 pixels, were acquired using an epifluorescence microscope (Leica DMI6000 B, Germany) equipped with a cooled sCMOS camera (Neo 5.5, Andor technology, UK) and both were controlled with Micro-manager (version 1.4)¹³. Images were sampled at a rate of approximately 10 Hz for about 2 min, stored as image stacks and analysed off-line using MATLAB (release 2013b; Mathworks). To automatically detect the neurons, we used the ADINA toolbox¹² (<https://bitbucket.org/jakirkham/adina-toolbox-v0.1/src>), which is a set of MATLAB functions specifically developed for the analysis of calcium imaging data. This includes a segmentation step where regions of interest corresponding to individual neurons are selected. For each segmented neuron, we measured fluorescence traces as relative changes in fluorescence intensity over time.

S1.1.3 Immunofluorescence staining assay

Immunostaining for a dopaminergic marker, tyrosine hydroxylase (TH) and a pan neuronal marker, Class III β -tubulin (TUB β III) were used to identify differentiated dopaminergic neurons. Immunostaining for tyrosine hydroxylase (TH) positive differentiated neurons was performed on wells of a 6-well plate after day 25 of differentiation. Differentiated cells were fixed with 4 % PFA in 1 $\times$ phosphate-buffered saline (PBS) (15 min), followed by permeabilisation with 0.05% Triton-X 100 in 1 $\times$ PBS (3 min on ice), and blocking with 10% fetal calf serum (FCS) in 1 $\times$ PBS (1 h). After washing with 1 $\times$ PBS, the primary antibodies mouse anti-TUB β III (1:1000, Covance, Germany), rabbit anti TH (1:1000, Santa Cruz biotechnology, Germany) and chicken anti-GFAP (1:1000, Merck Millipore, Germany), were

incubated for 90 min at 25 °C. After washing with 1× PBS, the secondary antibodies Alexa Fluor 488 Goat Anti-Rabbit (1:1000, Invitrogen), Alexa Fluor 568 Goat Anti-Mouse (1:1000, Invitrogen), Alexa Fluor 647 Goat Anti-chicken (1:1000, Invitrogen) and Hoechst 33342 to stain DNA (1:10000, Invitrogen), were incubated overnight at 4 °C. After washing with 1× PBS, confocal images of areas of selected wells were acquired, using a confocal microscope (Zeiss LSM 710).

S1.2 Transcriptomic analyses

S1.2.1 Cell culture

A human neuroepithelial stem cell line from a healthy donor was maintained and differentiated into DN_s, using an established protocol⁵⁴, described in supplementary information section S1.1.1, with the following adaptations. The hNE_{SC}s were cultivated in mTESR1 medium (StemCell technologies, #05850) on 6-well dishes coated with Matrigel (Corning, #354263). The media composition, to the extent that it has been defined by the manufacturer, is detailed in Table S-4. At 23 days of the protocol (Figure S4), the percentage of TH positive cell was estimated between 15-20%. Since protein content per cell can vary from 2.46×10^{-5} to $4.71 \times 10^{-5} \mu\text{g}/\text{cell}$, protein content was measured using a Bradford protein assay.

S1.2.2 RNA preparation

RNA extraction The Ambion Magmax™-96 total RNA isolation kit (Life Sciences) was used for RNA extraction. Magnetic beads were used to isolate nucleic acids. Afterwards, the samples were washed and purified with DNAase. The RNA obtained was eluted in 50 μM elution buffer. Fragment Analyzer (Aligent Technologies Inc.) was used to measure RNA quality and concentration.

RNA-sequencing protocol RNA-sequencing data was generate from a hNE_{SC}-derived dopaminergic neuronal cell culture at day 23 in culture. The sequencing library preparation was done using 200 ng of total RNA input with the TrueSeq RNA Sample Prep Kit v3-Set B (RS-122-2002, Illumina Inc, San Diego, CA) producing a 275 bp fragment including adapters in average size. In the final step before sequencing, twelve individual libraries were normalised and pooled together using the adapter indices supplied by the manufacturer. Pooled libraries have then been clustered on

the cBot Instrument (Illumina Inc, San Diego, CA) using the TruSeq SR Cluster Kit v3-cBot-HS (GD-401-3001, Illumina Inc, San Diego, CA) sequencing was then performed as 78 bp, single reads and 7 bases index read on an Illumina HiSeq3000 instrument using the TruSeq SBS Kit HS- v3 (50-cycle) (FC-401-3002, Illumina Inc, San Diego, CA).

S1.3 Analysis of RNA sequencing data

The raw RNA-seq data were analysed with a custom-made RNA-seq analysis pipeline, which included publicly available software (SAMtools, version 0.1.18; FASTX-Toolkit, version 0.0.14)³⁶ and custom-made python scripts. The RNA-seq analysis pipeline consists of six main steps: (i) quality control for the raw RNA-seq reads; (ii) preprocessing of the raw RNA-seq reads to remove adapters and low-quality sequences; (iii) alignment of the reads to the human reference genome; (iv) assembly of the alignments into transcripts and (v) quantification of the expression levels of each gene. Briefly, the raw RNA-seq reads (length 52 nucleotides, single-end) of each sample were checked by FastQC (version 0.11.2) to determine the read quality. Adapter sequences and low quality sequences were removed by cutadapt (version 1.10)³⁹ using default settings. Reads with length less than 25 nucleotides were excluded from further analysis. Next, the alignment of RNA-seq reads against the human reference genome (NCBI build37.2, downloaded from iGenome of Illumina, https://support.illumina.com/sequencing/sequencing_software/igenome.html) was performed using TopHat2 (version 2.0.13)³⁰. Alignment results were processed by Cufflinks (version 2.2.1)⁶⁷ for assembly of transcripts with default parameter settings. The quantification of gene expression was estimated by normalised FPKM (Fragments per kilobase of transcript per Million mapped reads) and counts at gene level by cuffnorm (version 2.2.1)⁶⁷. In order to obtain one expression value per gene, we used the transcript with the largest average expression as representative for the corresponding gene, since measurements for low-abundance transcripts are less reliable. In case of replicated genes, the maximum value expression from replicates was averaged.

S1.4 Exometabolomic data

Table S-3 contains a list of target metabolites analysed with both LC-MS and GC-MS platforms: 75 biogenic amines and amino acids, and 24 organic acids. Aspartic acid is targeted in both platforms, therefore a total of 98 metabolites were targeted.

S1.4.1 LC-MS profiling of biogenic amines and amino acids

The analysis of 75 biogenic amines (Table S-3) was performed with an established LC-MS method⁴⁶. Briefly, 15 μL of culture medium was extracted by adding 400 μL of ice-cold methanol, 55 μL of ice-cold milliQ water, 10 μL of tris(2-carboxyethyl)phosphine (TCEP; $1\mu\text{g}/\mu\text{L}$) and 10 μL of a mixture of stable isotope labelled internal standards (Table S-3). The samples were vortexed for 10-20 seconds and centrifuged at $16000\times g$ and 4 $^{\circ}\text{C}$ for 10 min. For the the calibration samples, 80 μL of each calibrant sample was mixed with 10 μL of TCEP ($1\mu\text{g}/\mu\text{L}$) and 10 μL of internal standard mix and extracted with 400 μL of ice-cold methanol as for medium samples. After centrifugation, all supernatants were transferred into 1.5 mL tubes and the liquid extracts were evaporated in a vacuum concentrator (Labconco, Kansas City, MO, USA) to dryness. The dried extracts were first dissolved in 80 μL borate buffer (pH 9) and mixed with 20 μL of pure acetonitrile containing 3 $\mu\text{g}/\mu\text{L}$ AccQ-Tag derivatisation reagent (Waters, Etten-Leur, Netherlands) to start the chemical derivatisation of the primary and/or secondary amine groups. The derivatisation reaction was performed at 55 $^{\circ}\text{C}$ for 30 min in a temperature-controlled orbital shaker (VWR Incubating Microplate Shaker, Germany). After completion of derivatisation, the samples were centrifuged at $16000\times g$ and 4 $^{\circ}\text{C}$ for 2 min and 80 μL of the supernatant was transferred into LC vials for sample injection. 1 μL of the liquid extract was injected onto the analytical column for the analysis.

All measurements were performed with a Waters Acquity ultra-high pressure liquid chromatography (UPLC) (Milford, MA, USA) hyphenated with Agilent 6460 triple-quadrupole mass spectrometer (Palo Alto, CA, USA). Chromatographic separation was achieved on a Water Acquity HSS T3 C18 UPLC column ($2.1\times 100\text{ mm}$, 1.7 μm) and the metabolites were identified based on their retention time and via multiple reaction monitoring (MRM) transitions from their protonated precursor ions of the AccQ-Tag derivates into common product ion of 171 m/z with corresponding linear ranges and LOD values (Table S-3). The peak detection and integration were performed with Agilent MassHunter Quantitative Software v7.0 (Palo Alto, CA, USA).

For the concentration determination, the calibration lines were drawn on a concentration range over three orders of magnitude (0.1 μM - 100 μM). The calibrant and internal standards were spiked into blank solvent (methanol/water, 80%/20%; v/v) in which the area ratio of each target analyte to its corresponding internal standard was used to define the ordinate values of the calibration curve. In total, sixteen calibration points were selected where each six calibration point covers one order of magnitude (e.g. 100 nM - 1 μM : 100 nM, 200 nM, 400 nM, 600 nM, 800 nM, 1 μM). After linear regression, the linear response range for each metabolite was determined by Pearson's correlation coefficient ($R^2 \geq 0.95$) except dopamine and levodopa. The linear equation for each calibration line was used to convert area ratios obtained in samples into absolute quantities (in μM) by using a Macro formula in Microsoft Office 2010.

S1.4.2 GC-MS profiling of polar metabolites

Twenty-four polar metabolites (Table S-3) were analysed in culture media using a modified version of an in-house built GC-MS platform³. Because of the high abundance of D-glucose and L-lactic acid in culture media, samples were diluted 1:299 (v/v) in milliQ water. Fifty microliters of both diluted and non-diluted culture medium was extracted with 425 μL of an extraction solvent (methanol/water, 94%/6%; volume/volume) containing stable isotope labelled internal standards (Table S-3). After vortexing the samples for 10 min on a multivortex, the samples were centrifuged at 16000 $\times g$ and 4 $^{\circ}\text{C}$ for 10 min. Four hundred microliters of the supernatant was transferred into a 1.5 mL tube and the solvent was evaporated in a vacuum concentrator (Labconco, Kansas City, MO, USA). Dry samples were resuspended for the oximation reaction in 35 μL of pyridine containing methoxyamine hydrochloride (15 $\mu\text{g}/\mu\text{L}$) and kept at 30 $^{\circ}\text{C}$ for 90 min. After the oximation of the aldehyde groups on reducing sugars and organic acids, samples were further derivatised with silylation reaction for 60 min in an orbital shaker (VWR, Germany). This reaction was carried out by adding 40 μL of MSTFA (N-methyl-N-trimethylsilylacetamide) into the samples. Subsequently, samples were centrifuged at 16000 $\times g$ and room temperature for 5 min and 70 μL of the supernatant was transferred into silanized glass inserts. The GC-MS measurements were performed on an Agilent 7890A GC System coupled to a single 759 quadruple

5975C Mass Selective Detector. One microliter of sample was injected with splitless injection. The analytes were separated on an Agilent HP-5MS Ultra Inert capillary GC column (30 m, 250 μm ID, 0.25 μm film thickness). Metabolite identification was carried out by using the retention time of the chemical standards and mass spectral similarity of the fragmentation pattern with NIST MS Search Software (v2.0). The metabolite quantification was performed based on the specific fragment ion for each polar metabolite (Table S-3). Both peak extraction and integration were performed by using the vendor's software (Agilent MassHunter Quantitative software v5.0). The concentrations were determined by spiking eleven (for diluted samples) or six (for non-diluted samples) different concentration values of the chemical standards on a 50 μL of diluted (300x) or non-diluted mixture of study samples (e.g. quality control sample). The area ratio of each target analyte to its corresponding internal standard was used to measure the ordinate values for the calibration lines. After linear regression, the linear response range for each metabolite was determined by Pearson's correlation coefficient ($R^2 \geq 0.95$) except alpha-ketoglutaric acid. The concentration values are then calculated using Macro formula in Microsoft Office 2010.

S2 Reconstruction

S2.1 Active and inactive genes and reactions

A *context-specific metabolic model* contains only the set of reactions active in a particular context. Therefore, we assembled a core set of genes and metabolic reactions known to be active or inactive in dopaminergic neurons *in vivo* or in hNESC-derived dopaminergic neurons in culture. A core set of active genes (Table S-1), as well as active and inactive reactions (Table S-2) was obtained either from manual curation of the literature or from transcriptomic data. Manual curation, described below, was focused on the physiological and biochemical literature on dopamine metabolism, dopaminergic neuronal transporters, central carbon metabolism, mitochondria-associated reactions and genes. In addition, manual curation of the literature was used to determine the need for addition or deletion of external reactions that are required for modelling non-equilibrium steady-state fluxes in dopaminergic neuronal metabolism. The list of genes, established by manual curation to be metabolically active, was combined with the aforementioned

transcriptomic data and used to generate the context-specific model through gene-protein-reaction associations⁶⁴.

S2.2 Dopamine metabolism

A key characteristic of dopaminergic neurons is their ability to synthesise, degrade and release dopamine. Therefore, manual reconstruction of dopamine metabolism was emphasised. In Recon2⁶⁵, there were already 75 tyrosine related reactions distributed in 6 subsystems. This content was extended with information from a comprehensive literature review of dopamine metabolism⁴¹ and additional manual curation of the literature (Table S-2), according to an established protocol⁶⁴.

S2.3 Dopaminergic neuronal transporters

The metabolic identity of a cell is strongly influenced by its ability to transport particular metabolites across its extracellular membrane, and in the metabolic model, this is represented by constraints on the corresponding exchange reactions, which define the boundary conditions of the model. To start the reconstruction of dopaminergic neuronal transporters, we began with the 1550 extracellular transport reactions in Recon 2.04⁴⁷, which correspond to 255 genes as identified by gene-protein reaction associations. Almost half (120/255) of the genes associated with extracellular transport reactions were manually curated. Manual curation of the experimental literature primarily involved the identification of transporters present in human substantia nigra pars compacta tissue or cell cultures of dopaminergic neurons through *in situ* hybridisation, RT-PCR, immunohistochemistry or immunoblotting. When human data was not found, data from rat or mouse was included instead. Additionally, when data specific for dopaminergic neurons or substantia nigra pars compacta was not found, evidence for transporters being present in neurons in general, astrocytes or blood brain barrier was used instead. After a review of metabolic genes active in the brain, only the genes specific for dopaminergic neurons or substantia nigra pars compacta were included in the list of active genes.

S3 Constraint generation

S3.1 Biomass maintenance and turnover constraints

Stoichiometric specification of biomass composition¹⁵, as well as cellular synthesis and turnover requirements is an essential component for the specification of the objective function in constraint-based modelling. However, fully differentiated dopaminergic neurons do not replicate and therefore, it is sufficient if lipid, nucleic acid, and amino acid synthesis meet the demand for their turnover. Therefore, we adapted an established methodology⁶⁴ to define the minimal biomass maintenance and turnover requirements for dopaminergic neurons. This required manual curation of the neurochemical literature to extract biomass precursor turnover rates, fractional biomass composition, and identification of key degradation reactions for dopaminergic neurons. Where human data was not obtained, rodent data was used. Where human substantia nigra pars compacta dopaminergic neuronal data was not obtained, other neuronal data was used.

Biomass composition. TheRecon3D⁷ biomass maintenance reaction was decomposed into its constituent biomass precursors. The fractional composition of a human substantia nigra pars compacta dopaminergic neuron was obtained by following several steps. First, the lipid and water fractional composition of a human substantia nigra pars compacta dopaminergic neuron was assumed to be the same as the one from a 55 year human cerebral cortex grey matter (39.6% dry weight of lipids and 60.4% dry weight of non-lipid residues and 82.3% wet weight water content)⁴⁹. Furthermore, we used the protein wet weight (w/w) fractional composition for human substantia nigra (99 mg/g w/w)⁵ to calculate the protein dry weight (DW) fractional composition. RNA and DNA dry weight fractional compositions for human substantia nigra (grey matter) were readily available in the literature (3,29 $\mu\text{g}/\text{mg}$ DW of RNA and 1,81 $\mu\text{g}/\text{mg}$ DW of DNA)³⁵. Based on the relative concentrations of the different neuronal lipids, amino acids, and nucleic acids, the overall dry weight (DW) fractional human neural tissue composition was estimated to be 39.60% lipid, 55.93% protein, 0.18% DNA, 0.33% RNA, and 3.96% others⁴⁸.

The fractional composition (%) of each biomass precursor was converted into a reaction rate ($\mu\text{mol}/\text{gDW}/\text{h}$ that is micromole per gram dry weight per hour). These values were then converted into fluxes ($\mu\text{mol}/\text{gDW}/\text{h}$) taking in consideration an

Chapter 5

experimental time of 48h and a value for the dry weight of an hNESC-derived neuron ($DW(DN)$, gDW/cell). The latter was calculated using

$$DW(DN) := \text{Protein content} / \text{protein percentage}$$

where the protein content ($\mu\text{g}/\text{cell}$) was obtained from the results of the aforementioned Bradford assay (0.0002459 to 0.00047053 $\mu\text{g}/\text{cell}$) and the protein percentage was based on the calculated fractional protein composition of a human substantia nigra pars compacta dopaminergic neuron. This is a coarse-grained approximation of neuronal lipid, amino acid, and nucleic acid maintenance requirements converted into $\mu\text{mol}/\text{gDW}/\text{h}$. The protein content ranged from 0.0002459 to 0.00047053 $\mu\text{g}/\text{cell}$. The protein percentage obtained from the literature is 55.93% protein⁴⁸. Therefore, the dry weight of a single dopaminergic neuronal cell was estimated to be 6.4×10^{-10} gDW/cell, with a range of $4.4 - 8.41 \times 10^{-10}$.

Key degradation reactions. Using the neurochemical literature, the degradation pathway, or pathways, for each biomass precursor were identified and the first reaction in each degradation pathways was identified in Recon3D. For example, as reviewed in³¹, phosphatidylserine is exclusively localised in the cytoplasmic leaflet of neuronal and astrocytic membranes, forming protein docking sites for signalling pathways. The phosphatidylserine decarboxylase enzyme is able to decarboxylate the serine moiety of phosphatidylserine to form phosphatidylethanolamine. Although one of the fatty acyl groups of phosphatidylserine can also be hydrolysed to convert phosphatidylserine into lysophosphatidylserine, this is quantitatively a minor pathway.

Turnover rates

The turnover rate of a metabolite reflects the rate at which that metabolite is replaced in a tissue, given by the reciprocal of the turnover time⁴⁰. Turnover rates are commonly expressed as half-lives and represent the time, expressed in hours, required for half of the precursor to be replaced³³. Metabolite half-lives $t_{1/2}$ were collected from the literature³³ and converted into turnover rates λ with

$$\lambda := \ln(2) / t_{1/2} \quad (1)$$

and then interpreted as constraints on degradation reaction rates. Turnover rates were converted to the same unit ($\mu\text{mol/gDW/h}$) as reaction rate, and applied as constraints, as described in supplementary information section S4.3.

S3.2 Maximum metabolite uptake constraints

Only the constituents of the defined fresh medium, plus some reversible extracellular transport reactions including water, carbon dioxide and oxygen, were permitted to be taken up by the model. That is, lower bounds on the corresponding exchange reactions were set by assuming that the maximum uptake rate is equal to the metabolite concentration in the fresh medium, divided by the duration of the interval being modelled (Table S-3). This is always an overestimate of the actual metabolite uptake rate, because it effectively assumes that the concentration of each uptaken metabolite is zero at the end of the time interval.

S3.3 Exometabolomically derived exchange reaction rate ranges

For each pair of quantitative measurements for the same metabolite at two time points, an exometabolomically derived exchange reaction rate, v_{exp} , was estimated for the corresponding metabolite by assuming a constant rate of change of metabolite concentration with respect to time and setting this rate of change of metabolite concentration to be equal to the experimentally measured flux when scaled appropriately with respect to the calculated dry cell mass in culture. The total dry weight for the *in vitro* cell culture (at day 19 and 23) was estimated from the product of the dry weight of a single neuron in culture times the number of cells in the cellular culture.

Cell number. The cell number in each culture well was measured on days 1, 13, 19, but not day 21 or 23 in culture. Therefore, the evolution of cell number with respect to time was estimated using a cubic spline fit to the measured cell numbers (Figure S5). Exometabolomic data was collected at day 9, 13, 19 and 23. However, only exometabolomic data from day 19 and 23 were used to quantitatively constrain the models. This is consistent with the established differentiation protocol used, where a 30-45% increase in cell number is observed during the first five days and therefore a steady state assumption was not considered valid during the early period in cell culture.

Steady state assumption. The assumption of a metabolic steady state for the differentiated dopaminergic neuronal cell culture in the latter period in cell culture is based on two observations. Firstly, in contrast to earlier stages, the cell number does not alter significantly in the last five days in culture (<3-4% increase). Secondly, it is known that the rate of neuronal differentiation reaches toward a plateau toward the end of the period in culture⁵⁴.

Exchange reaction rate. In the model, the unit of flux is $\mu\text{mol/gDW/hr}$, while the unit of metabolite concentration change is μmol . In order to transform an extracellular metabolite concentration change into a lower bound on the corresponding exchange reaction flux, we assumed that

$$\text{flux} = \text{metabolite concentration change} / \text{cell dry weight} \times \text{cell number} \times \text{time}$$
 based on an established approach, implemented in Metabotools⁴, a software suite for integration of metabolomic data with constraint-based models, integrated into the COBRA Toolbox²⁶.

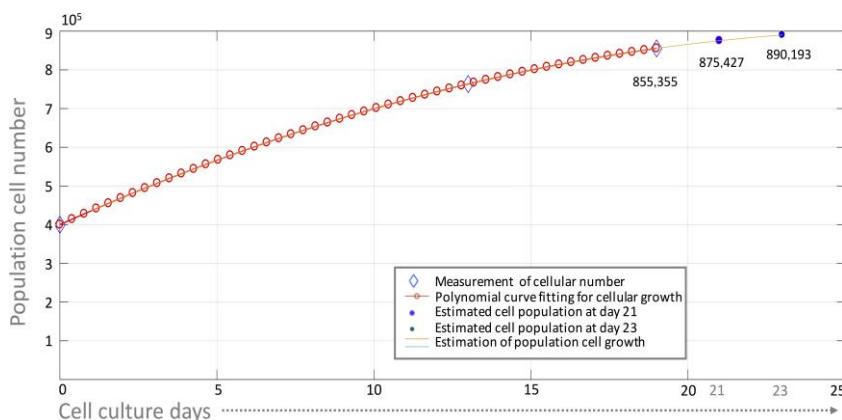


Figure S5: Measured and estimated cell numbers during neuronal differentiation Cell culture numbers were measured at seeding (day 0) and day 13 and 19 of differentiation. The cell culture was seeded with a density of 400k cells per well. The cell number at day 21 and 23 was estimated by interpolation in order to enable normalisation of metabolic uptake and secretion rates.

Exchange reaction rate ranges. The measurements of metabolite concentration changes and the aforementioned cell culture parameters are associated with

measurement uncertainty. Therefore, this measurement uncertainty was propagated to uncertainty in the estimation of experimental exchange reaction rates. The exometabolically derived exchange reaction ranges was therefore set to be between $v_{exp} - \sigma$, and $v_{exp} + \sigma$ where σ denotes the standard deviation in estimated experimental exchange reaction rates. For details on the exometabolically derived exchange flux ranges for each measured metabolite, see Table S2.

S4 Model generation

A context-specific, flux-consistent, constraint-based metabolic model representative of the hNESC derived dopaminergic neuronal *in vitro* cell culture was generated using an overall approach based on established and novel Constraint-Based Reconstruction and Analysis methods²⁶.

S4.1 An overview of constraint-based modelling

All constraint-based modelling predictions are derived from optimisation problems, typically formulated in the form:

$$\begin{aligned} \min_{v \in \mathbb{R}^n} \quad & \psi(v) \\ \text{s.t.} \quad & Sv = 0 \\ & l \leq v \leq u, \end{aligned} \tag{2}$$

where $S \in \mathbb{R}^{m \times n}$ is a stoichiometric matrix of m metabolites and n reactions representing a biochemical 908 network, $v \in \mathbb{R}^n$ is the vector representing the flux through all of the reactions in a network and $\psi: \mathbb{R}^n \rightarrow \mathbb{R}$ is an objective function, which is typically convex. In a constraint-based metabolic model of reaction fluxes, the set of feasible steady-state flux vectors forms a polyhedral convex solution space, defined by the equality and inequality constraints in Equation (2), enabling optimisation of a variety of convex objective functions over this set.

The matrix S can be split horizontally into two matrices corresponding to internal, $N \in \mathbb{Z}^{m \times k}$, and external, $B \in \mathbb{R}^{m \times (n-k)}$, reactions, with corresponding internal and external rate vectors, $z \in \mathbb{R}^k$ and $w \in \mathbb{R}^{n-k}$. While all internal reactions are characterised by being mass and charge balanced, external reactions are, on the other hand, imbalanced reactions. External reactions are classified in sink, demand or exchange reactions. A demand reaction allows the accumulation of a compound. A sink reaction allows the production of a metabolite. Finally, A exchange reaction allows the

exchange of a metabolite across the extracellular boundary of a system, providing a mechanism to transfer metabolites between the environment and the extra-cellular fluid. Such reactions are distinct from transport reactions, which transfer metabolites between compartments within the model, including the extracellular compartment. Exchange reactions are added to a model to allow certain metabolites to be exchanged across the boundary of the system at variable rates.

The linear equality, $Sv = 0$ in Equation (2), represents mass balance for all the metabolites. This means, for each metabolite the rate of metabolite consumption is equal to the rate of metabolite production. In Equation (2), $Sv = 0$ implies that $Nz = -Bw$ where internal production plus external input equal internal consumption plus external output. For certain intracellular metabolites, those not exchanged across the boundary of the system, we assume they are at a steady-state, so we have $N_i z = 0$, where N_i denotes the i^{th} row of the internal stoichiometric matrix. Additional linear inequalities keep reaction rates between lower and upper bounds, l and u , respectively.

S4.2 The generic human metabolic model: Recon3.0 model

Given a generic reconstruction of human metabolism, not specific to any organ, tissue or cell type, a generic model of human metabolism was generated, using an established procedure²⁶.

Recon3D⁷ is the latest and most comprehensive, manually-curated genome-scale reconstruction of human metabolism. Additionally, Recon3D provides information about gene-protein-reaction associations which associate each metabolic gene with the corresponding enzyme or enzyme complex and reaction in a Boolean manner. The largest stoichiometrically and flux consistent²⁶ part of Recon3D, termed Recon3.0model, was used as a generic model for generation of dopaminergic neuronal metabolic models. This generic model is divided into 9 cellular compartments and currently encompasses 2,248 open reading frames and 10,600 metabolic reactions involving 5,835 unique metabolites.

In each metabolic reaction, v_i is constrained between a lower and an upper bound, $lb \leq v_i \leq ub$. The default reaction lower and upper bounds are commonly set based on

model characteristics and constraints value. Lower and upper bounds were set to include fluxes from metabolite concentration in the media, e.g., glucose flux rate based on media composition ($-5,430.74 \mu\text{mol/gDW/hr}$). Reactions can be reversible or irreversible. A reaction is said to be reversible in the case where it has a negative *lb* and a positive *ub*. When the *lb* is set to zero and the *ub* is a positive number the reaction proceeds in the forward direction. Similarly, when the *ub* is zero and the *lb* is a negative number the reaction occurs in the backward direction. The same works for exchange reactions: if a metabolite is taken up, the corresponding exchange reaction has zero as *ub* and a negative number as *lb*, whereas if it is secreted, the *ub* is a positive number and the *lb* is set to zero.

S4.3 Generation of the turnover model

Given the generic model, a turnover model was generated by applying constraints on the turnover rates of certain key cellular constituents of a dopaminergic neuron.

As described in supplementary information section S3.1, the minimum turnover requirement of a dopaminergic neuron was obtained from the literature and used to constrain the model as follows. When a biomass precursor was associated with a single degradation reaction, this reaction was set to irreversible in the direction of degradation, and 0.75 times the degradation rate *d* was set as the lower bound on that degradation reaction. A 25% relaxation of the lower bounds from the estimated degradation rate was used as standard to account for uncertainty in the data⁶⁶. For the example of phosphatidylserine in supplementary information section S3.1, a lower bound was set on the phosphatidylserine decarboxylase reaction. When a biomass precursor could be metabolised through a reversible reaction, one direction of which corresponded to catabolism, this reaction was split into a pair of irreversible reactions and the turnover constraint applied to the catabolic direction.

When a biomass precursor could be degraded by more than one reaction, the sum total rates of degradation by all degradation reactions, was set to be greater than 0.75 times the degradation rate *d*, via an inequality of the form

$$v_1 + v_2 + \dots + v_n \geq 0.75 \times d, \quad (3)$$

with due consideration of reaction directionality. Support for inequalities, such as Equation 3, within constraint-based modelling problems, has been fully implemented

within the COBRA Toolbox²⁶. In total, this approach resulted in 21 turnover constraints on single degradation reactions, and a further 8 turnover constraints, each on a set of degradation reactions, when the metabolite could be degraded by more than one pathway.

S4.4 Generation of the preconditioned model

Given the turnover model, a preconditioned model was generated by applying maximum metabolite uptake constraints and biochemically motivated constraints on certain otherwise reversible exchange reactions.

As described in supplementary information section S3.2, for each metabolite present in the fresh medium, the maximum metabolite uptake rate ($\mu\text{mol/gDW/hr}$) was calculated (Table S-3). This was then used to set the lower bound on the corresponding exchange reaction. The lower bounds on all other metabolite uptake reactions were set to zero, to reflect the assumption that no other metabolites, except those in the defined medium, were accessible to the *in vitro* culture. For each consumed metabolite, the upper bound on the corresponding exchange reaction is, by definition, set to zero.

Furthermore, some bounds on otherwise reversible exchange reactions were manually set to satisfy specific characteristics of a (neuronal) cell culture, e.g., the production of oxygen and glucose were disallowed by setting the upper bound on the corresponding exchange reaction to zero, so that only uptake became possible in the model. Also, the ionic transport reactions for sodium, calcium, potassium and iron, were closed ($l_j = u_j = 0$). Of the components present in the fresh medium, 50 were used to qualitatively constraint the uptakes of the model. Furthermore, a set of reactions associated to dopaminergic neuronal metabolism were qualitatively constrained such as the production of neuromelanin, ATP, dopamine, GABA etc.

S4.5 Generation of the context-specific models

Given the preconditioned model, a context-specific model was generated by application of manually curated and omics derived constraints on the activity or inactivity of certain genes and reactions.

A core set of Recon3D genes and reactions were extracted by extensive manual curation of biochemical studies as being active or inactive in human dopaminergic neurons from the substantia nigra (supplementary information section S8). In addition to the specific data extracted from the literature, RNA-sequencing data from hNESC-derived dopaminergic neuronal *in vitro* cell culture, which is sensitive at genome-scale, was mapped into Recon3D to identify the genes that should be active in a dopaminergic neuronal reconstruction (supplementary information section S1.2).

A metabolic network formed from the set of core reactions alone is not necessarily flux consistent, that is, some reactions may not admit a non-zero steady-state flux. Therefore, we used the FASTCORE algorithm⁷¹, implemented in the COBRA Toolbox⁵⁷ to generate a compact, flux-consistent model. This model returns a minimal number of extra reactions, beyond the core set, that are required to ensure the flux-consistency of the model. Therefore, the output is a context-specific, flux-consistent model.

S4.6 Generation of exometabolomically constrained models

Given the context-specific model, a set of exometabolomically constrained models were generated by selective application of constraints derived from quantitative measurements of fresh and spend cell culture medium. Exometabolomically derived exchange reaction ranges, for metabolites measured to be taken up from the fresh medium were used to generate an *uptake constrained model* (ModelUpt) to test its ability to predict the measured secreted metabolites. Exometabolomically derived exchange reaction ranges, for metabolites measured to be secreted into the medium were used to generate a *secretion constrained model* (ModelSec) to test its ability to predict uptaken metabolites. Furthermore, for each one of the measured metabolites, a *leave-one-out model* was generated, from the iNESC2DN, to test the ability to predict exchange of one metabolite measured, but left out of the set used to constrain the

model. The iNESC2DN model was generated from the context-specific model by including all constraints derived from the metabolomic data on uptaken and secreted metabolites. It is this iNESC2DN model that was subsequently used then to design future metabolomic experiments.

A given vector of exometabolomically derived exchange reaction rates $v_{exp} \in \mathbb{R}^n$, obtained as described in supplementary information section S3.3, may not be consistent with the feasible set of steady state fluxes defined in 2. Specifically, inconsistent with the set defined by the steady state constraint ($Sv = 0$) as well as the lower and upper bounds on each reaction $l \leq v \leq u$. Should this occur, we fitted the model to the experimental data, relaxing the constraints on the bounds of 2, that admits a steady state flux $v \in \mathbb{R}^n$, using the following quadratic optimisation problem

$$\begin{aligned} \min_{v, p, q \in \mathbb{R}^n} \quad & (v_{exp} - v)^T \text{diag}(w_{exp})^{-1} (v_{exp} - v) + p^T \text{diag}(w_l) p + q^T \text{diag}(w_u) q \\ \text{s.t.} \quad & Sv = 0, \\ & l - p \leq v \leq u + q, \\ & 0 \leq p \\ & 0 \leq q \end{aligned} \tag{4}$$

where $p \in \mathbb{R}^{n_{\geq 0}}$ and $q \in \mathbb{R}^{n_{\geq 0}}$ are non-negative variables that permit relaxation of the lower and upper bound constraints, respectively. This formulation also allows for different weights to be input as parameters to Problem 4 to penalise deviation from experimentally measured mean fluxes, with $w_{exp} \in \mathbb{R}^{n_{\geq 0}}$, penalise relaxation of lower bounds, with $w_l \in \mathbb{R}^{n_{\geq 0}}$ and penalise relaxation of upper bounds, with $w_u \in \mathbb{R}^{n_{\geq 0}}$. For example, if the experimentally measured flux is actually the mean flux, then one could set w_{exp} to be the inverse of the variance on the experimental flux measurements, thereby increasing the penalty on deviation from an experimentally measured mean flux where the variance is lower. Certain lower or upper bounds might not be realistic to be relaxed, e.g., an essential amino acid can always be taken up but never secreted, therefore the upper bound on the corresponding exchange reaction must be zero.

S5 Model testing and characterisation

A selection of constraint-based modelling techniques were used to test the ability of *in silico* models to predict independent exometabolomic data. Let $\Omega := \{v \in \mathbb{R}^n \mid Sv = 0, l \leq v \leq u\}$ denote the set of steady-state flux vectors consistent with the constraints, and $\dim(\Omega)$ denote the number of linearly independent dimensions of this set. In Flux Balance Analysis (FBA) one obtains an optimal flux vector by choosing a coefficient

vector $c \in \mathbb{R}^n$, representing a biologically motivated linear objective $\psi(v) := c^T v$ (e.g. ATP production, dopamine secretion, etc.) and enforcing $v \in \Omega$. Since the correct coefficient vector is not known for neurons, we used alternative approaches to explore Ω . Flux Variability Analysis (FVA)³⁸, was used to find the *flux ranges* for each reaction rate by choosing a coefficient vector $c \in \mathbb{R}^n$ with one non-zero entry, then minimising and maximising $\psi(v) := c^T v$, for each reaction in turn. FVA was implemented in a computationally efficient manner using the fastFVA algorithm²¹, within the COBRA Toolbox²⁶.

In addition, uniform sampling⁷² was used to generate an unbiased characterisation of the set of steady-state flux vectors Ω . Uniform sampling provides a quantitative prediction of the probability of each quantitative flux value, between the same minimum and maximum flux predicted by flux variability analysis, assuming that each feasible steady-state flux vector is equiprobable. Unlike FBA and FVA, uniform sampling does not use an objective function when predicting steady-state fluxes. Uniform sampling was implemented using the Coordinate Hit-and-Run with Rounding (CHRR) algorithm²⁵, within the COBRA Toolbox²⁶, using the parameters $nSkip = \dim(\Omega)^2$ and $nSamples = 8 \times \dim(\Omega)$, which represent the number of samples skipped between stored samples, and the total number of stored samples obtained, respectively. FVA and uniform sampling were used to test whether various *in silico* models could predict the outcome of independent exometabolomic analyses (see Figure 5.2, and for all measured exchanged metabolites more comprehensively, see Figure S12 and S13).

Sparse Flux Balance Analysis approximately minimises the function $\psi(v) := \|v\|_0$ subject to $v \in \Omega$, and was used to predict the minimum number of reactions that are required to be active to satisfy the known metabolic demands on a dopaminergic neuron, as represented by the aforementioned dopaminergic neuron specific constraints on the steady-state flux space Ω . The SparseFBA algorithm is implemented in the COBRA Toolbox²⁶.

S6 Experiment design

S6.1 Exometabolomic experiment design

Using uniform sampling, we calculated the flux ranges of exchange reactions to predict a list of secreted metabolites to consider for targeted metabolomic analyses in future exometabolomic experiments. Furthermore, we prioritised this list of secreted metabolites by rank ordering it using two novel pipelines for optimal design of future exometabolomic experiments.

In the uncertainty reduction pipeline, illustrated in Figure 5.3a-c, our goal is to select the k rows of the covariance matrix that will explain as much uncertainty in the model as possible, as suggested by the uniform samples. More precisely, we seek to select a subset of k rows such that the k -dimensional volume spanned by these vectors is maximised. Computing this subset exactly is a difficult computational problem, and we instead use a heuristic, iterative method that greedily selects the row that has the maximum Euclidean distance to the subspace spanned by the rows selected already. For the first row, this reduces to selecting the row of the covariance matrix with largest Euclidean norm. The largest Euclidean norm of the covariance matrix of exchange reaction fluxes, was used to rank order the metabolic exchanges contributing the most uncertainty to iNESC2DN predictions (Figure 5.3d).

In the phenotypic perturbation pipeline, illustrated in Figure 5.3e-h, we predicted the exometabolomic changes most likely to occur following perturbation to internal reaction rates, as illustrated in (Figure 5.3h). Starting with the iNESC2DN model, the effects of two different perturbations were predicted: (i) deletion of the glucocerebrosidase (GBA1) gene, the gene most commonly associated with PD, and (ii) complete inhibition of mitochondrial complex I. For each reaction, a two-sample Kolmogorov-Smirnov test was used to check whether the sampled fluxes for perturbed and control models came from different distributions, with a 5% significance level. This test uses the maximum absolute difference between the cumulative distribution functions of the distributions of the two data vectors. For the j^{th} reaction, the test statistic is

$$D_j^* := \max_{v_j} (|F\hat{c}(v_j) - F\hat{p}(v_j)|)$$

where $F_{\hat{c}}(v_j)$ is the proportion of control model sample flux values less than or equal to v_j and $F_{\hat{p}}(v_j)$ is the proportion of perturbed model sample flux values less than or equal to v_j .

S6.2 Conserved moieties and atom mapping

Where available, a molecular structure was obtained for each metabolite from the Virtual Metabolic Human database (www.vmh.life)⁴⁷ and atom to atom mappings for each of the internal reaction of the iNESC2DN model were obtained using an atom mapping algorithm, the Reaction Decoder Tool⁵³, which performed optimally in a benchmarking exercise⁵². The atom mappings were then used to identify the metabolite structural moieties that are conserved despite all of the metabolic transformations in the iNESC2DN model. Moiety identification used an established algorithm²⁴, implemented within the COBRA Toolbox²⁶.

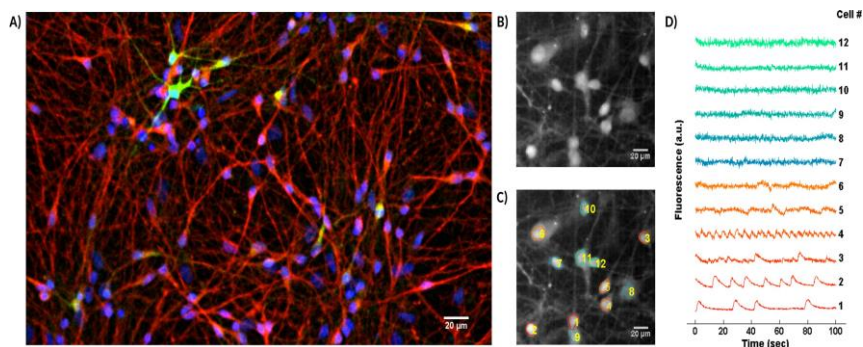


Figure S6: Immunostaining and calcium imaging.

Immunostaining of differentiated neurons and calcium imaging of spontaneously firing human neuroepithelial stem cell differentiated into dopaminergic neurons. (a) Immunostaining of a representative well at day 23, showing neurons positive for nuclei with Hoechst (blue), TUBIII (red) and TH (green); scale bar 20 μm. (b) Mean frame of a field of view of representative neurons. (c) Automatic segmentation of neurons. (d) Fluorescence traces showing the spontaneous activity of individual segmented neurons.

Part II

Results

S7 Experimental results

S7.1 Cell culture

The differentiation of hNECs into dopaminergic neurons was successfully accomplished. Differentiated neurons were identified by TUB β III immunoreactivity. Neurons positive for tyrosine hydroxylase (TH) confirmed the presence of neurons capable of converting tyrosine to L-DOPA, the penultimate step in dopamine synthesis (Figure S6a). Analysis of calcium imaging data revealed spontaneously active neurons (Figure S6b, c, d).

S7.2 Transcriptomic data

Transcriptomic data contains a range of gene expression values. Some of the low expression values are certainly attributable to experimental noise or aborted transcripts, but for borderline expression values, it is a challenge to divide the corresponding genes into expressed or not expressed. Each gene with less than zero FPKM, on base-two logarithmic scale, was considered not expressed⁵⁹. Each gene with FPKM higher than a threshold of zero, on base-two logarithmic scale, was considered expressed. Out of the 18,530 unique genes with expression levels reported in the transcriptomic data, 12,698 were considered to be expressed, based on the aforementioned threshold. However, only 1,202 were mapped into Recon3D (metabolic genes) and therefore included in the model. To test the viability of the selected transcriptomic data expressed in the *in vitro* culture and selected in Recon3D, a receiver operating characteristic (ROC) curve²³ was generated to qualitatively compare the expressed and not-expressed assignments from hNEC-derived dopaminergic neurons transcriptomic data against the active and inactive assignments for manually curated dopaminergic neuronal genes (supplementary information section S8 below), which we assume to be a true representation of dopaminergic neuronal gene expression (Figure S7). If a gene was considered to be active by manual curation and was also found to be expressed in transcriptomic data, it was considered a true positive (TP). The proportion of true positives that were correctly classified as positive, is given by the *true positive rate* (TPR)

$$TPR = TP / TP + FN'$$

Mechanistic model of dopaminergic neuron metabolism

where TN, FP and FN denote true negatives, false positives and false negatives, respectively. Likewise, the *false positive rate* (FPR) is

$$FPR = FP / FP + TN.$$

The true and false positive rates can vary depending on the threshold applied to distinguish between an expressed and a not-expressed gene. In the reconstruction, genes expressed above the threshold were assigned to be metabolically active and genes expressed below the threshold were not included as in the model, unless the corresponding reactions had to be included to generate a flux consistent model.

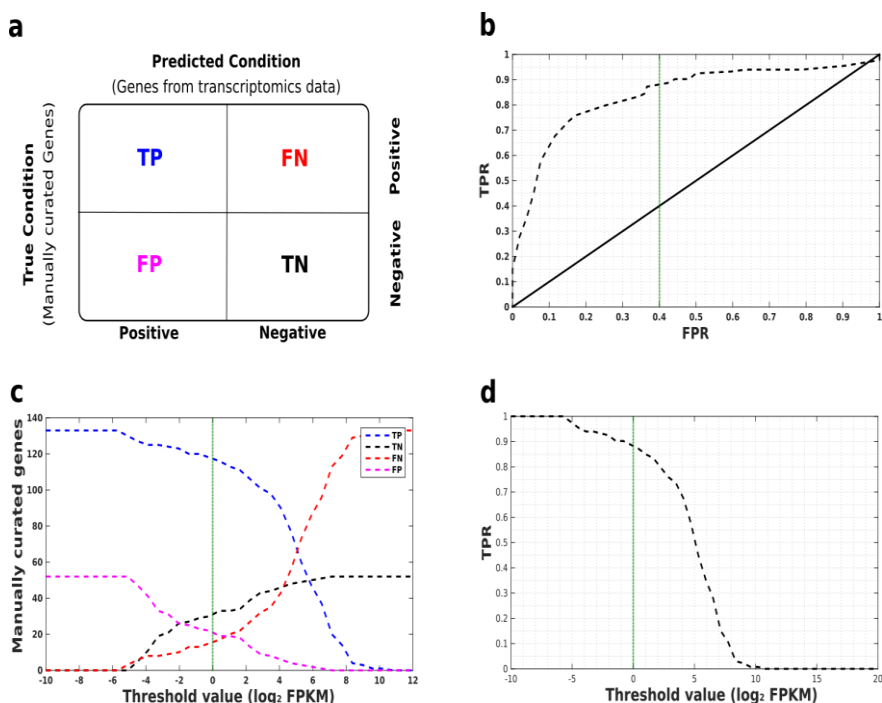


Figure S7: Manually curated genes compared with transcriptomic data.

(a) Confusion matrix illustrating the performance of the transcriptomic classification into active and inactive genes. TP - True Positive, TN - True Negative, FN - False Negative, FP - False Positive. (b) Receiver operating characteristic (ROC) curve. TPR - True Positive Rate, FPR - False Positive Rate. (c) Number of manually curated genes per threshold for each condition. (d) A true positive rate of 0.9 corresponded to a threshold

Chapter 5

value of zero Fragments Per Kilobase of transcript per Million mapped reads (FPKM), on base-two logarithmic scale (green vertical line).

S7.3 Exometabolomic analysis

Medium characterisation. The manufacturers specification identifies a total of 57 different metabolites in the fresh culture medium, of which 24 were amines, 12 vitamins, 16 inorganic salts, 1 lipid, 2 nucleotides and 2 organic acids (Table S-4). Out of these 57 metabolites, 50 were present in the stoichiometrically and flux consistent generic model (Recon3Dmodel) and were omitted from further consideration. Of the omitted 7 metabolites, four inorganic salts (nitrate, vanadium, manganese and copper) were not present in the reconstruction (Recon3D), and a further three (magnesium, cyanocobalamin and selenite) did not correspond to any stoichiometrically and flux consistent reaction.

By running fresh medium samples in the organic acid (GC-MS) and amine (LC-MS) platforms, we were able to measure the absolute concentrations of both organic acids (glucose and pyruvic acid) and 22 of the 24 amines, known to be in the medium. This enabled us to test the concordance between the specifications of the medium manufacturer and the actual concentrations (Figure S8). Reduced glutathione and L-cystine are two amines that cannot be detected by the LC-MS platform. The remaining medium components are mainly inorganic salts and they were not measured.

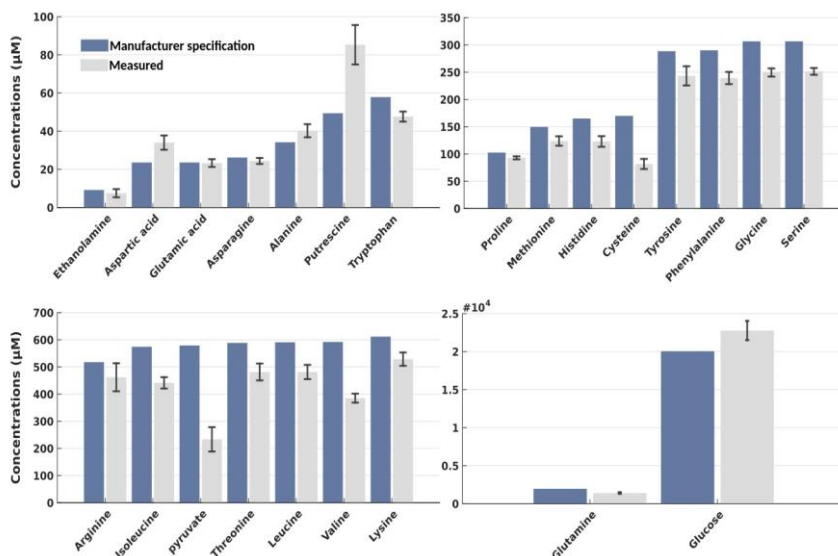


Figure S8: Validation of specified fresh medium concentrations.

Metabolite concentrations specified by the medium manufacturer (blue) compared to the absolute concentrations measured by mass spectrometry (grey). Some quantified metabolite concentrations, e.g., for cysteine, pyruvic acid, valine, aspartic acid and putrescine, significantly deviate from the manufacturers specifications.

S8 Reconstruction and manual curation

Active genes. Based on literature curation, a total of 252 metabolic genes were established to be active in dopaminergic neuronal metabolism (Table S-1). Out of these 252 genes, 20 are related to transport reactions, 6 to dopamine metabolism, 124 associated to mitochondria and 102 to central carbon metabolism. Significant effort was made to manually curate transport reactions as their presence or absence help to establish the idiosyncratic boundary conditions for any particular cell type. Out of the genes for transport reactions in Recon 2.04, biochemical literature on approximately half (118/255) were individually manually curated. From these 118 transporters, 20 were found to be present in the substantia nigra or substantia nigra pars compacta in human, mouse, or rat.

Active reactions. Based on literature curation, a total of 420 unique reactions were found to be active in dopaminergic neurons (Table S-2). However, 10 of these were excluded from the model generation process as they were either stoichiometrically or flux inconsistent. Included in these 420 reactions there are 69 from dopamine metabolism, 31 from an intersection the medium metabolites with the transport reactions associated with the 24 transport genes, 8 demand reactions for biomass precursors, 53 exchange reactions corresponding to medium metabolites, 5 exchange reactions from the metabolites being newly synthesised according to the exometabolomic data, and 160 reactions related to mitochondria and central carbon metabolism.

Inactive genes and reactions. Based on literature curation, a total of 148 genes were deemed to be inactive in neurons (Table S-1). These inactive genes were used to help determine the cutoff between expressed and non-expressed genes in the analysis of transcriptomic data (supplementary information section S1.2). Based on manual curation, a total of 211 metabolic reactions were considered to be inactive in the brain (Table S-2) and therefore were excluded from the model.

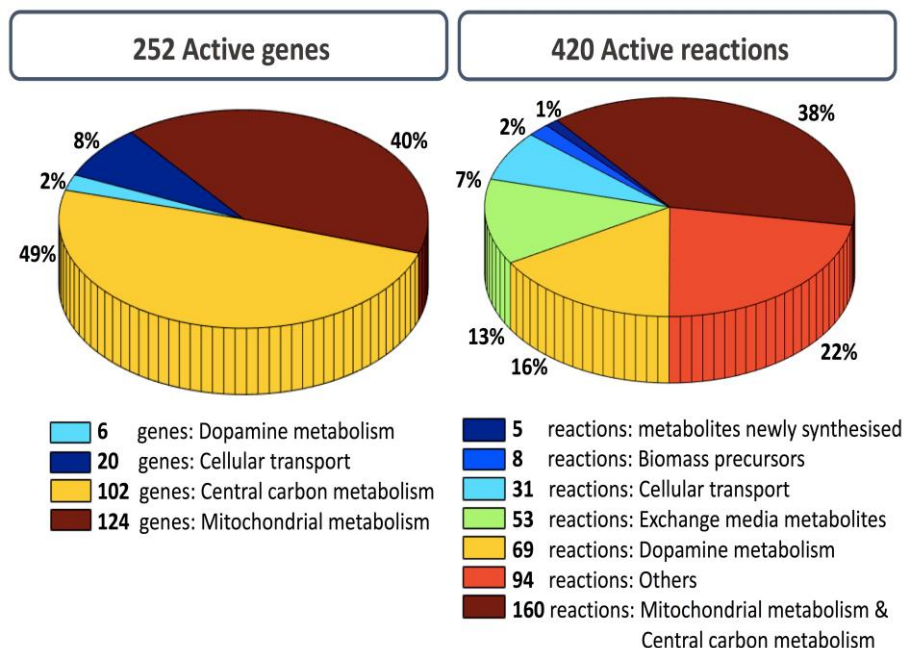


Figure S9: Classification of active reactions and genes by manual curation.

This manual curation result is partly a reflection of the availability of biochemical information on certain pathways, e.g., central metabolism, and partly a reflection of the pathways that were targeted for manual curation, e.g., dopamine metabolism.

Dopamine metabolism. Following manual curation of the literature^{41, 58, 43, 44, 76, 14, 10, 69, 27, 45}, 9 metabolites and 49 reactions were added to dopamine metabolism during the generation of Recon3D from Recon 2.04. These are 11 transport, 11 exchange, 19 metabolic and 8 demand reactions. In Recon3D, dopamine metabolism now includes 122 reactions in total (Figure S10). Out of these 122 reactions, we were able to collect evidence for the occurrence of 77 reactions in dopaminergic neurons that were also included in our reconstruction as active reactions: 42/49 newly added reactions and 35/73 dopamine-related reactions already present in Recon 2.04 (21 reactions from 'Tyrosine metabolism' subsystem, 1 reaction from 'Miscellaneous' subsystem, 4 reactions from 'Transport, extracellular' subsystem, 7 reactions from 'Exchange/demand reaction' subsystem, 1 reactions from 'Tetrahydrobiopterin metabolism' subsystem, 1 reaction from 'Phenylalanine metabolism' subsystem). For

many reactions (45/122) no clear information was found, therefore they were not included as active or inactive. Further information can be found in Table S-2.

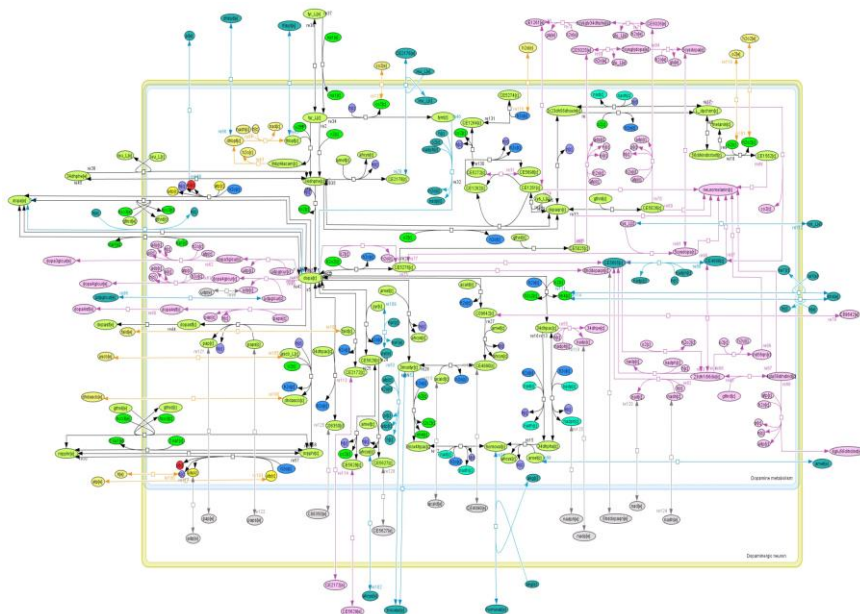


Figure S10: Reconstruction of dopamine metabolism.

Dopamine metabolism in Recon2.04 (green, blue) was refined and updated with newly added reactions (pink).

S9 Constraints

S9.1 Exometabolomic data

We performed a targeted metabolomic analysis and quantified metabolic differences with respect to time. Two platforms were selected, capable of detecting a total of 98 metabolites (74 metabolites from amines and amino acids platform and 24 from organic acid platform). Of the 50 metabolites, measured above the lower limit of detection, all were present in Recon3D. However one metabolite, glutaric acid (VMH), was not present in the stoichiometrically and flux consistent subset of Recon3D, designated Recon3.0model, and therefore, they were not further considered for computational modelling. From the remaining 49 metabolites, 24 were not previously present in the fresh medium 11.

S10 Generated models

Mechanistic model of dopaminergic neuron metabolism

The main characteristics of the different models generated along this process is given in Table S-5. A total of 6 main models were generated. Model generic correspond to Recon3.0model. The first four features define the composition of the metabolic network based on transcriptomic data and manual curation of reactions and genes. The last seven features define the range of fluxes through the network by differential application of constraints. A brief summary of each model is given below:

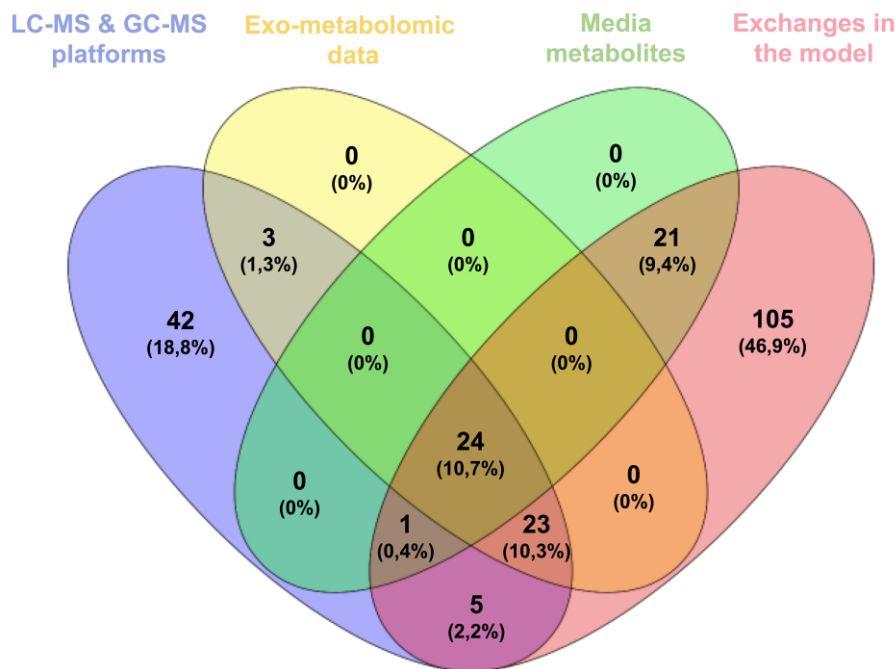


Figure S11: Venn diagram summarising metabolites.

A total of 98 metabolites were targeted by the selected LC-MS and GC-MS platforms (blue). Of these 98, only 50 were detected in the spent medium obtained to generate exometabolomic data (yellow). Of these 50 metabolites, 49 could be used to constrain the model as only one metabolite was not present in the stoichiometrically and flux consistent subset of Recon3D. Of these 49 metabolites, 25 were present in the fresh medium and 24 were synthesised by the cells and secreted into the spent medium. The final iNESC2DN model contains exchanges reactions for 179 metabolites (red), therefore there still remains 126 metabolites to target with exometabolomic platforms developed and applied in future.

3. *Preconditioned model*: Turnover model with added qualitative constraints for 46 fresh media metabolites. Exchange reactions corresponding to fresh media metabolites were open with a *lb* set less than zero, to the negative of the maximum concentration of each metabolite in the fresh media. Uptake reactions corresponding to metabolites not in the fresh media were closed (*lb* set to zero) (see Table S-3).
4. *Draft context-specific model*: Preconditioned model with added manual curation and integration of context-specific transcriptomic data, to generate a model consisting of 1776 metabolic reactions (see Tables S-2).
5. *iNESC2DN model*: Context-specific model with exometabolomic constraints added for 30 metabolites taken up and 19 metabolites secreted. This omits exometabolomic constraints on L-proline and L-serine, as the model would otherwise be infeasible.
6. *ModelUpt*: Context-specific model with added exometabolomic constraints for 30 metabolites measured to be taken up from the medium (see Table S-2).
7. *ModelSec*: Context-specific model with added exometabolomic constraints for 19 metabolites measured to be secreted into the medium (see Table S-2).

S11 Model testing

S11.1 Prediction of metabolite secretion rates

Figure 5.2a-c provides a comparison of measured and predicted metabolite secretion rates, for 3 representative metabolites, using ModelUpt, a context-specific model quantitatively constrained with exometabolomically derived constraints on secretion reaction rates. Figure S12 provides this comparison for all 29 metabolites measured to be secreted into the medium by the cell culture.

S11.2 Prediction of metabolite uptake rates

Figure 5.2d-f provides a comparison of measured and predicted metabolite uptake rates, for 3 representative metabolites, using ModelUpt, a context-specific model quantitatively constrained with exometabolomically derived constraints on uptake reaction rates. Figure S12 provides this comparison for all 19 metabolites measured to be uptaken from the fresh medium by the cell culture.

S12 Model characterisation

Figure S14 compares the number of reactions from each metabolic subsystem present in the Recon3.0 model, the iNESC2DN model and the minimal set of reactions required to be active to satisfy all of the iNESC2DN constraints. Figure 5.3b and 5.3h illustrate the reactions within the iNESC2DN model that are predicted to be perturbed as a result of GBA1 knockout and Complex I inhibition, respectively.

Part III

Discussion

S13 Characteristics of the iNESC2DN model

Tetrahydrobiopterin metabolism is an example of a metabolic subsystem critical for dopaminergic neuronal metabolism that is also represented in the iNESC2DN model. Tetrahydrobiopterin is absolutely required for the synthesis of monoamine neurotransmitters²⁸, including dopamine. The tetrahydrobiopterin biosynthetic pathway is thought to involve up to eight different proteins that support six alternate *de novo* and two alternate salvage pathways. The expression of these genes is highly correlated with each other and is particularly enriched within monoaminergic

Chapter 5

neurons²⁸. The first and limiting step in the tetrahydrobiopterin biosynthetic pathway is catalyzed by GTP cyclohydrolase I (GCH1). The expression of this gene is generally low and particularly heterogeneous across different populations of monoamine-containing neurons in humans and rodents, although GCH1 expression is a characteristic of nigrostriatal dopaminergic neurons²⁸. The *in vitro* culture expresses GCH1 ($\log_2(\text{RPKM}) > 2$) and therefore the corresponding Recon3D reaction gtp cyclohydrolase I (GTPCI) is also included in the iNESC2DN model. GTP cyclohydrolase I produces dihydroneopterin triphosphate (ahdt) and formic acid (for). In a subsequent reaction, 6-pyruvoyltetrahydropterin synthase (PTHPS) converts dihydroneopterin triphosphate (ahdt) into 6-pyruvoyltetrahydropterin (6pthp), which is a direct precursor of tetrahydrobiopterin (thbpt).

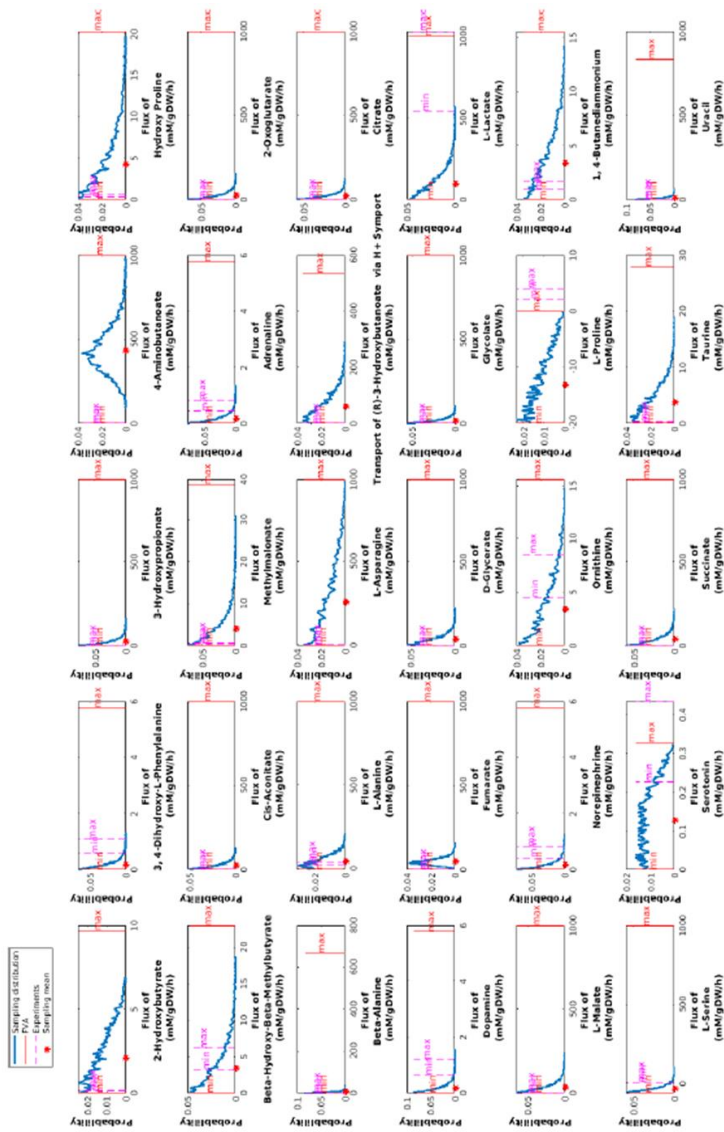


Figure S12: Comparison of predicted and measured metabolite secretion rates. An uptake constrained model (ModellUpt) was tested for its ability to predict measured rates of 29 secreted metabolites, with three representatives illustrated in (a-c). A measured range for each exchange reaction rate (pink) was obtained from quantitative exometabolic measurements and includes one standard deviation of measurement uncertainty. Predicted probability of exchange reaction flux obtained by uniform sampling (dark blue). Predicted exchange reaction flux, derived from the mean of the sampling distribution (red dot). Predicted maximum and minimum fluxes obtained by flux variability analysis (orange). The measured secretion reaction fluxes were within the range predicted by flux variability analysis of ModellUpt for 26 metabolites, and outside the range for 3 metabolites.

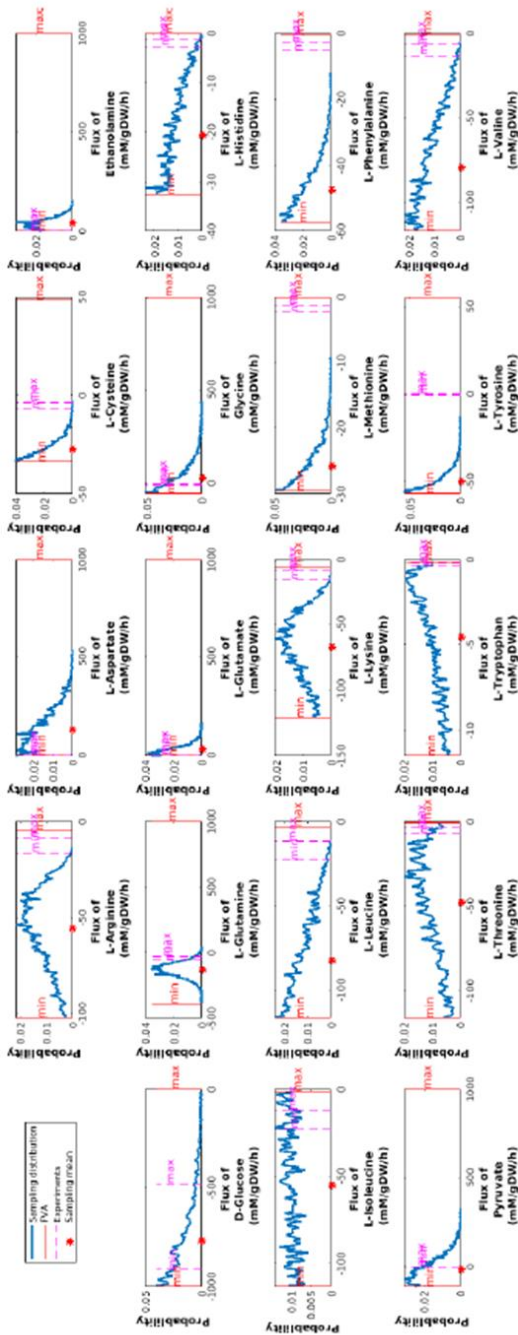


Figure S13: Comparison of predicted and measured metabolite uptake rates. A secretion constrained model (ModelSec) to test its ability to predict measured rates of 19 metabolites measured to be taken up from the fresh medium. A measured range for each exchange reaction rate (pink) was obtained from quantitative exometabolomic measurements and includes one standard deviation of measurement uncertainty. Predicted probability of exchange reaction flux obtained by uniform sampling (dark blue). Predicted exchange reaction flux, derived from the mean of the sampling distribution (red dot). Predicted maximum and minimum fluxes obtained by flux variability analysis (orange). The measured uptake reaction fluxes were within the range predicted by flux variability analysis of ModelSec for 14 metabolites and outside the range for 5 metabolites.

Mechanistic model of dopaminergic neuron metabolism

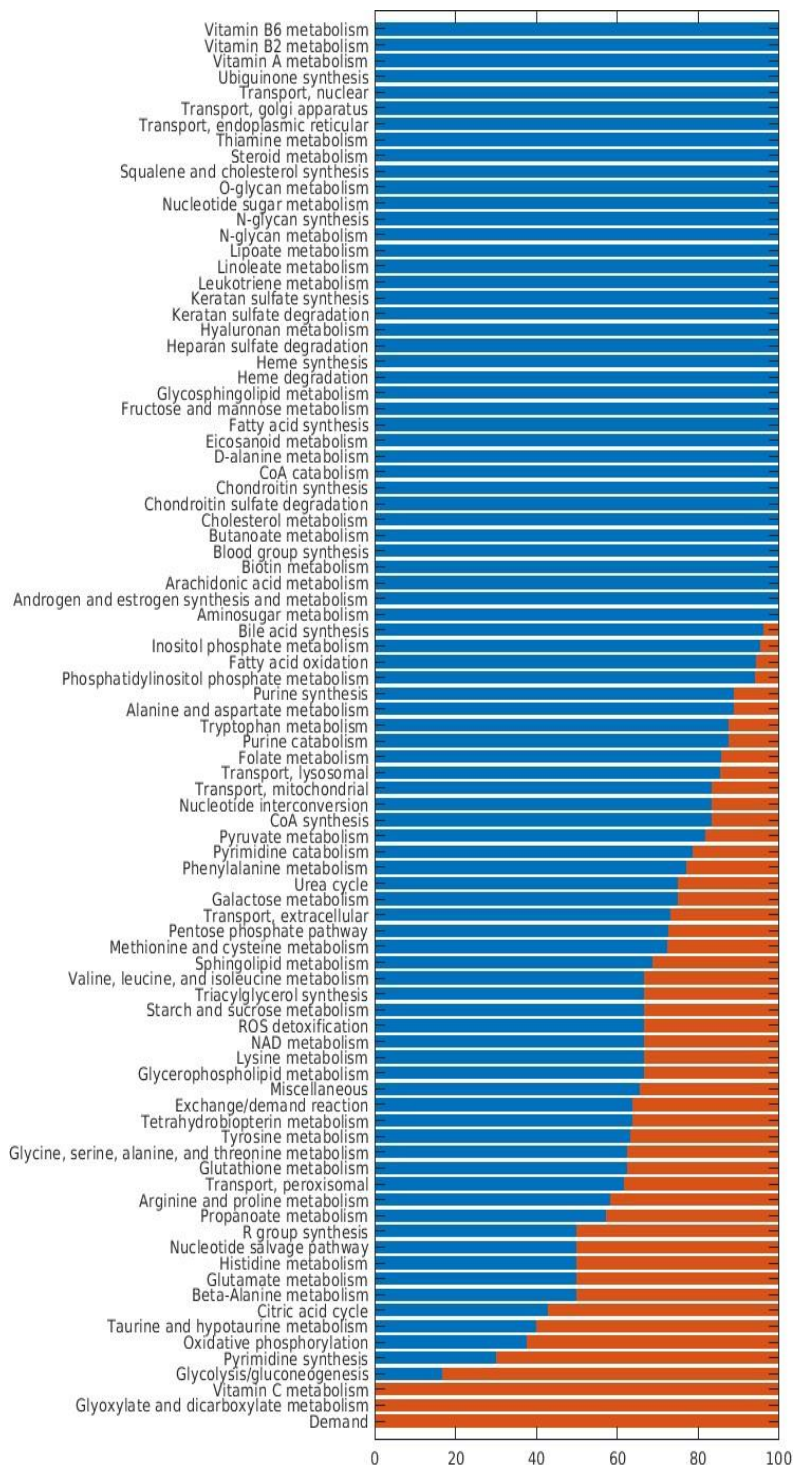


Figure S14: Minimal flux metabolic subsystems of the iNESC2DN model.

Comparison of the fraction of reactions in (red) and out (blue) of the minimal flux vector in each metabolic subsystem of the iNESC2DN model.

S14 Metabolic assays

The quality of the experimental measurements was assessed by comparing the measured and supplier reported concentration values (Fig. S8). Based on this comparison, for most of the measured metabolites, the measurements were obtained within a similar concentration value (20%) reported in fresh culture media. However, some measurements (e.g. cysteine, pyruvic acid, valine, aspartic acid and putrescine) demonstrated larger differences between measured and supplier's concentrations. There can be several explanations for the discrepancy between specified and measured quantities. Some compounds may undergo easily reactions with other elements/compounds in their environment due to their reductive and/or oxidative nature. For instance, the thiol group of the cysteine is reactive both with oxidants and reductants as well as it has high affinity for metals⁵¹. Therefore, it can be very difficult to determine the quantity of free cysteine residues after protein hydrolysis unless the thiol group of the molecule is not stabilised by extensive chemical derivatisation procedures¹. This could easily explain the observation of a lower value for the measured concentration in comparison to the reported value 8. Under oxidative and low pH conditions of the incubator, free cysteine may have been converted into different compounds which were not targeted during our analysis. The discrepancy can be due to the incomplete knowledge on the composition and concentration of medium supplements. B21 medium supplement contains a confidential amount of putrescine which may have contributed to the large increase in the measured concentrations.

It has been previously shown that the concentration of pyruvic acid can be reduced in blood by its reaction with bisulphite-binding substances. This can be also true for the pyruvic acid in fresh culture medium where there are several substances with the bisulphite-binding characteristics. In general, it is challenging to determine an overall reference point for the absolute concentrations as many factors may contribute to the in accuracy of the measurements. One of these factors is the difference in measurement uncertainties between different analytical methods. Previously, the measurement accuracy of ELISA against a validated tandem LC-MS

method has been evaluated and several differences were observed. The analytical assays that were used to report the concentration values in the medium samples by the suppliers may exhibit a difference in uncertainty in comparison to our analytical platforms. At this point, there is still a lack of certain standardisation for absolute concentration measurements among different analytical laboratories which have to be compiled under Good Laboratory Practice.

The absolute quantification of fluxes is advantageous for the training and validation of a metabolic model. However, the high chemical diversity and broad concentration range of the endogenous metabolites complicates the simultaneous quantification of many metabolites with one analytical methodology⁵⁶. Due to its high selectivity and large dynamic range, mass spectrometry in combination with conventional separation techniques such liquid chromatography (LC), gas chromatography (GC) or capillary electrophoresis (CE) are the most popular methodologies in quantitative metabolomics.

S15 Analytical chemistry considerations

It is not possible for one analytical platform to quantify the concentration of all of the metabolites predicted to be secreted by the iNESC2DN model, because of the differences in their solubility, mass, endogenous concentration levels and volatility²². For selective and sensitive detection of diverse chemical classes of metabolites, several mass spectrometry (MS) based analytical platforms have been developed and reviewed extensively⁶. For the coverage of a broad group of hydrophilic metabolites such as amines and amino acids, carboxylic acids, sugars, nucleobases/nucleosides/nucleotides several hydrophilic interaction chromatography coupled mass spectrometry (HILIC-MS) platforms have been published and applied in cell culture and body fluids^{50, 19, 75, 77, 70}. For the analysis of specific hydrophilic classes, it has been also a possibility to combine several chemical derivatisation techniques based on the functional group of the metabolites and achieve more retention and sensitivity on traditional reverse-phase liquid chromatography coupled mass spectrometry (RPLC-MS)^{60, 73, 20}.

For the separation of sugars and volatile metabolites gas chromatography coupled mass spectrometry (GC-MS) is a complementary method with its separation

Chapter 5

efficiency for the stereoisomers of pentose and hexose sugar^{32, 29, 68}. Nevertheless, GC-MS analysis requires chemical derivatisation for sugars which is suited for reducing the amount of several conformational isomers as well as increasing their volatility. For lipophilic metabolites, as similar to hydrophilic metabolites, there are analytical methods with broad coverage of different lipid classes^{55, 74, 8, 11, 63, 9}. However, for certain lipid classes, such as oxylipins and isoprostanes^{61, 16, 2}, sphingolipid⁴², bile acids^{17, 34}, cholesterol and cholesterol esters³⁷ and prenol lipids¹⁸ more targeted RPLC-MS methods have been developed and validated in the literature.

To be able to generate the absolute concentration values of the metabolites, an efficient calibration strategy should be included with the authentic chemical and stable isotope labelled internal standards for each target metabolite or the target chemical class⁶². For this purpose the predictions of the *in silico* model will be an important asset to define the key target metabolites in the metabolic network for the calibration samples and thereby helping the right choice of the analytical platforms for the design of metabolomic experiments.

Supplementary tables.

Table S1. Literature curation showing evidence of activity, or inactivity of 252 genes.

EntrezGeneID (Active)	EntrezGeneID (Inactive)
38	586
39	1717
43	3156
226	6566
240	9194
241	10449
292	5409
384	128
412	222
498	1109
622	1544
840	1551
1152	1553
1160	1555
1583	1557
1588	1559
1593	1562
1607	1572
1630	1576
1737	1577
1738	1582
1743	1585
1757	1586
2023	1621
2026	2180
2052	2571
2074	2572
2108	2632
2109	2670
2110	2819
2222	2880
2246	3028
2261	3158
2271	3283

EntrezGeneID (Active)	EntrezGeneID (Inactive)
2539	3284
2597	3292
2629	3293
2645	3294
2646	3295
2720	3764
2739	5091
2747	5161
2820	5409
2821	5834
2932	6241
3098	6296
3099	6518
3101	6519
3170	6530
3479	6542
3763	6580
3939	7299
3945	7350
4047	7923
4190	8529
4191	8630
4929	9154
4967	10005
5053	10165
5106	10202
5160	10249
5162	10965
5209	10998
5211	11001
5214	11283
5223	23597
5226	26002
5315	26027
5337	26227
5527	28965
6342	51144
6515	51170
6821	51171

Mechanistic model of dopaminergic neuron metabolism

EntrezGeneID (Active)	EntrezGeneID (Inactive)
6853	51478
6888	51703
7084	54988
8050	55856
8802	55902
8803	57016
8878	57030
10059	64078
10135	64816
10797	64902
10858	66002
11332	79611
22934	80221
23761	81616
25796	83884
51552	84532
54550	93034
55669	93517
57084	114570
57194	116280
57704	117140
64080	121210
64802	122970
80210	123880
115827	126410
126328	134530
130752	197320
548596	246210
100133941	257200
18	267020
217	284490
219	340810
223	341390
224	345280
501	348160
539	376500
847	641370
1327	277
1329	278

EntrezGeneID (Active)	EntrezGeneID (Inactive)
1337	7364
1339	7365
1340	9426
1345	10720
1346	54575
1347	54577
1349	54579
1350	54657
1351	64711
1384	91227
1431	94033
1468	114770
1537	116080
2744	116180
2746	124980
2805	125210
2806	133690
2876	153200
2879	155180
2936	158840
3417	203610
3418	253180
3419	266740
3420	338600
3421	339220
4199	348930
4200	360200
4512	389020
4513	441280
4514	646490
4519	653600
4535	728440
4536	100140000
4537	100530000
4538	100530000
4539	102720000
4540	
4541	
4694	

Mechanistic model of dopaminergic neuron metabolism

EntrezGeneID (Active)	EntrezGeneID (Inactive)
4695	
4696	
4697	
4698	
4700	
4701	
4702	
4704	
4705	
4706	
4707	
4708	
4709	
4710	
4711	
4712	
4713	
4714	
4715	
4716	
4717	
4718	
4719	
4720	
4722	
4723	
4724	
4725	
4726	
4728	
4729	
4731	
4942	
6389	
6390	
6391	
6392	
6648	
7351	

EntrezGeneID (Active)	EntrezGeneID (Inactive)
7381	
7384	
7385	
7386	
7388	
7991	
8402	
8604	
8659	
9016	
9167	
9377	
9481	
10400	
10476	
10873	
10935	
10975	
23530	
27089	
27165	
29796	
51079	
54539	
55967	
56267	
79751	
80025	
83733	
84701	
125965	
126328	
170712	
341947	
349565	
374291	
1644	
4128	
4129	

Mechanistic model of dopaminergic neuron metabolism

EntrezGeneID (Active)	EntrezGeneID (Inactive)
6818	
7054	
9588	
19	
366	
3767	
4891	
5947	
6505	
6529	
6535	
6538	
6546	
6581	
6582	
6833	
10060	
81539	
84889	
206358	
6511	
6531	
6571	

Chapter 5

Table S2. Metabolic reactions (445) in dopaminergic neurons.

Reaction abbreviation	Reaction description	Reaction formula
EX_phe_L[e]	exchange reaction for L-phenylalanine	phe_L[e]<=>
EX_tyr_L[e]	L-Tyrosine exchange	tyr_L[e]<=>
EX_dopa[e]	Dopamine exchange	dopa[e]<=>
EX_34dhphe[e]	3,4-Dihydroxy-L-phenylalanine exchange	34dhphe[e]<=> 34dhphe[c]
EX_glu_L[e]	Predicted informative metabolites	glu_L[e]<=>
EX_dopasf[e]	Dopamine 3-O-sulfate exchange	dopasf[e]<=>
EX_dopa4sf[e]	Dopamine 4-O-sulfate exchange	dopa4sf[e]<=>
EX_dopa4glcur[e]	Dopamine 4-O-glucuronide exchange	dopa4glcur[e]<=>
EX_dopa3glcur[e]	Dopamine 3-O-glucuronide exchange	dopa3glcur[e]<=>
r0399	L-Phenylalanine,tetrahydrobiopterin:oxygen oxidoreductase [4-hydroxylating]	o2[c]+ thbpt[c]+ phe_L[c]-> tyr_L[c]+ thbpt4acam[c]
TYR3M02	tyrosine 3-monooxygenase	o2[c]+ thbpt[c]+ tyr_L[c]-> 34dhphe[c]+ thbpt4acam[c]
DHPR	6,7-dihydropteridine reductase	h[c]+ nadh[c]+ dhppt[c]-> nad[c]+ thbpt[c]
3HLYTCL	3-Hydroxy-L-tyrosine carboxy-lyase	h[c]+ 34dhphe[c]-> co2[c]+ dopa[c]
DOPAt4_2_r	Dopamine reversible transport in via sodium symport [1:2]	2.0 na1[e]+ dopa[e]<=> 2.0 na1[c]+ dopa[c]
DOPASfT	Dopamine 3-O-sulfate transport [diffusion]	dopasf[c]-> dopasf[e]
DOPA4SfT	Dopamine 4-O-sulfate transport [diffusion]	dopa4sf[c]-> dopa4sf[e]
DOPA4GLCURt	Dopamine 4-O-glucuronide transport	dopa4glcur[c]+h2o[c]+ atp[c]-> dopa4glcur[e]+ adp[c]+ pi[c]+ h[c]
DOPA3GLCURt	Dopamine 3-O-glucuronide transport	dopa3glcur[c]+ h2o[c]+ atp[c]-> dopa3glcur[e]+ adp[c]+ pi[c]+ h[c]
TYRCBOX	L-Tyrosine carboxy-lyase	h[c]+ tyr_L[c]-> co2[c]+ tym[c]
DOPAc	formation of dopamine	o2[c]+ h[c]+ nadph[c]+ tym[c]-> h2o[c]+ nadp[c]+ dopa[c]
DOPAVESSEC	Dopamine secretion via secretory vesicle [ATP driven]	h2o[c]+ atp[c]+ dopa[c]-> h[c]+ adp[c]+ pi[c]+ dopa[e]
42A12BOOX	4-[2-Aminoethyl]-1,2-benzenediol:oxygen oxidoreductase[deaminating][flavin-containing]	h2o[c]+ o2[c]+ dopa[c]-> h2o2[c]+ nh4[c]+ 34dhpac[c]
34DHPEAR	3,4-Dihydroxyphenylethanol:NADP+ reductase	34dhpac[c]+ nadph[c]+ h[c]-> 34dhphe[c]+ nadp[c]
34DHPLACOX	3,4-Dihydroxyphenylacetaldehyde:NAD+ oxidoreductase	h2o[c]+ nad[c]+ 34dhpac[c]-> 2.0 h[c]+ nadh[c]+ 34dhpha[c]
34DHPLACOX_N	3,4-Dihydroxyphenylacetaldehyde:NADP+ oxidoreductase	h2o[c]+ nadp[c]+ 34dhpac[c]<=> 2.0 h[c]+ nadph[c]+ 34dhpha[c]
ADP_		
EX_34dhpha[e]	exchange for 34dhpha	34dhpha[e]<=>
DOPASULT	Dopamine Sulfotransferase	paps[c]+ dopa[c]-> h[c]+ pap[c]+ dopasf[c]
DOPASULT4	Dopamine 4-O-Sulfotransferase	paps[c]+ dopa[c]-> h[c]+ pap[c]+ dopa4sf[c]
UDPG4DOPA	Dopamine 4-O-Glucuronidation	udpglcur[c]+ dopa[c]-> udp[c]+ dopa4glcur[c]
UDPG3DOPA	Dopamine 3-O-Glucuronidation	udpglcur[c]+ dopa[c]-> udp[c]+ dopa3glcur[c]
RE3201C	RE3201	o2[c]+ 56dihindlcrbxlt[c]-> h2o2[c]+ CE1562[c]
DCT	Dopachrome tautomerase	L_dpchrm[c]<=> 56dihindlcrbxlt[c]
DOPACHRMISO	L-dopachrome isomerase 1	o2[c]+ 2.0 2c23dh56dhoxin[c]-> 2.0 h2o[c]+ 2.0 L_dpchrm[c]
RE3198C	RE3198	h[c]+ nadh[c]+ L_dpchrm[c]<=> nad[c]+ 2c23dh56dhoxin[c]
DOPACHRMDC	L-dopachrome decarboxylation	L_dpchrm[c]-> co2[c]+ CE4888[c]
DOPAQNISO1	Dopaquinone isomerase 1	dopaqn[c]<=> h[c]+ 2c23dh56dhoxin[c]

Mechanistic model of dopaminergic neuron metabolism

Reaction abbreviation	Reaction description	Reaction formula
RE1917C	RE1917	dopaqn[c] + CE1261[c] <=> 34dhphe[c] + CE1262[c]
TYRDHINDOX	Tyrosinase: 5,6-dihydroxyindole oxygen oxidoreductase	2.0 CE4888[c] + o2[c] -> 2.0 h2o[c] + 2.0 ind56qn[c]
DACT	2,3-dihydro-1H-indole-5,6-dione tautomerization	23dh1i56dio[c] <=> CE4888[c]
NADPQNOXR	DT-diaphorase: NADP quinone oxireductase	nadph[c] + h[c] + 23dh1i56dio[c] -> nadp[c] + CE5665[c]
NADQNOXR	DT-diaphorase: NAD quinone oxireductase	nadh[c] + h[c] + 23dh1i56dio[c] -> nad[c] + CE5665[c]
LACROX	leukoaminochrome autooxidation	CE5665[c] + o2[c] -> 23dh1i56dio[c] + h2o2[c]
RE2526C	RE2526	h[c] + nadph[c] + CE4888[c] <=> nadp[c] + CE5665[c]
RE2296C	RE2296	gthrd[c] + CE5276[c] <=> CE5025[c]
CE5025t	transport of 5-S-glutathionyl-dopamine	CE5025[c] + atp[c] + h2o[c] -> CE5025 [e] + adp[c] + pi[c] + h[c]
EX_CE5025[e]	5-S-glutathionyl-dopamine exchange	CE5025[e] <=>
DOPAQCYS	Dopamine-o-quinone cysteine addition	cys_L[c] + CE5276[c] <=> 5cysdopa[c]
5CYSDDOPAt	transport of 5-S-cysteinyldopamine	5cysdopa[c] <=> 5cysdopa [e]
EX_5cysdopa[e]	5-S-cysteinyldopamine exchange	5cysdopa[e] <=>
RE1916C	RE1916	gthrd[c] + dopaqn[c] <=> CE5026[c]
CE5026t	transport of 5-S-glutathionyl-L-DOPA	CE5026[c] + h2o[c] + atp[c] <=> CE5026[e] + adp[c] + pi[c] + h[c]
EX_CE5026[e]	5-S-glutathionyl-L-DOPA exchange	CE5026[e] ->
CE1261t	transport of 5-S-cysteinyldopa	CE1261[c] <=> CE1261 [e]
EX_CE1261[e]	CE1261 exchange	CE1261[e] <=>
4GLU56DIHDINDt	transport of 4-s-glutathionyl-5,6-dihydroxyindoline	4glu56dihdind[c] + h2o[c] + atp[c] <=> 4glu56dihdind[e] + adp[c] + pi[c] + h[c]
EX_4glu56dihdind[e]	4-s-glutathionyl-5,6-dihydroxyindoline exchange	4glu56dihdind[e] <=>
DOPAtu	Dopamine uniport	dopa[e] <=> dopa[c]
RE1918C	RE1918	dopa[c] + acald[c] <=> h2o[c] + C09642[c]
C09642te	salsolinol transport uniport	C09642[c] <=> C09642[e]
EX_C09642[e]	C09642[e] exchange	C09642[e] <=>
CE5629t	transport of 6,7-dihydroxy-1,2,3,4-tetrahydroisoquinoline	CE5629[c] <=> CE5629[e]
EX_CE5629[e]	exchange for 6,7-dihydroxy-1,2,3,4-tetrahydroisoquinoline	CE5629[e] <=>
RE1921C	RE1921	CE5626[c] <=> 2.0 h[c] + co2[c] + CE5629[c]
RE1919C	RE1919	pyr[c] + dopa[c] <=> h2o[c] + CE5626[c]
RE3095C	RE3095	h2o2[c] + dopa[c] <=> 2.0 h2o[c] + CE5276[c]
RE2130C	RE2130	dopa[c] + fald[c] <=> h2o[c] + CE2172[c]
CE2172t	transport of 6,7-dihydroxy-1,2,3,4-tetrahydroisoquinoline	CE2172[c] <=> CE2172[e]
EX_CE2172[e]	exchange for 6,7-dihydroxy-1,2,3,4-tetrahydroisoquinoline	CE2172[e] <=>
DOPACCL	dopamine o-quinone cyclization	CE5276[c] -> CE5665[c]
DM_CE1261[c]	neuromelanin production from 5-S-cysteinyldopa	
DM_4glu56dihdind[c]	neuromelanin production from 4-s-glutathionyl-5,6-dihydroxyindoline	
DM_5cysdopa[c]	neuromelanin production from 5-S-cysteinyldopamine	
DM_CE1562[c]	neuromelanin production from 5,6-indolequinone-2-carboxylate	
DM_CE4888[c]	neuromelanin production from CE4888	

Chapter 5

Reaction abbreviation	Reaction description	Reaction formula
DM_CE5025[c]	neuromelanin production from 5-S-glutathionyl-dopamine	
DM_CE5026[c]	neuromelanin production from 5-S-glutathionyl-L-DOPA	
DM_ind56qn[c]		
ALAt4		na1[e]+ ala_L[e] -> na1[c]+ ala_L[c]
ARGtD		arg_L[e] <=> arg_L[c] h2o[c]+ atp[c]+ 2 na1[e]+ ascb_L[e] -> adp[c]+ h[c]+ pi[c]+ 2 na1[c]+ ascb_L[c]
ASCSVCTtc		
ASCBt4		na1[e]+ ascb_L[e] <=> na1[c]+ ascb_L[c]
ASNt4		na1[e]+ asn_L[e] -> na1[c]+ asn_L[c]
ASPDt6		h[e]+ 3 na1[e]+ asp_D[e]+ k[c] -> h[c]+ 3 na1[c]+ asp_D[c]+ k[e]
ASPt6		h[e]+ 3 na1[e]+ k[c]+ asp_L[e] -> h[c]+ 3 na1[c]+ asp_L[c]+ k[e]
CAt7r		3 na1[e]+ ca2[c] <=> 3 na1[c]+ ca2[e]
CHOLtu		chol[e] <=> chol[c]
CREATt4_2_r		2 na1[e]+ creat[e] <=> 2 na1[c]+ creat[c]
GLNt4		na1[e]+ gln_L[e] -> na1[c]+ gln_L[c]
GLUt6		h[e]+ 3 na1[e]+ k[c]+ glu_L[e] -> h[c]+ 3 na1[c]+ glu_L[c]+ k[e]
GLYt2r		h[e]+ gly[e] <=> h[c]+ gly[c]
GLYt4		na1[e]+ gly[e] -> na1[c]+ gly[c]
HMR_9613		2 na1[e]+ 4abut[e]+ cl[e] -> 2 na1[c]+ 4abut[c]+ cl[c]
HMR_9614		na1[e]+ dopa[e] -> na1[c]+ dopa[c]
HOMt4		na1[e]+ hom_L[e] -> na1[c]+ hom_L[c]
KCC2t		nh4[e]+ cl[e] <=> nh4[c]+ cl[c]
KCCt		k[e]+ cl[e] <=> k[c]+ cl[c]
LEUt4		na1[e]+ leu_L[e] -> na1[c]+ leu_L[c]
LYStiDF		lys_L[e] -> lys_L[c]
METt4		na1[e]+ met_L[e] -> na1[c]+ met_L[c]
PHEt4		na1[e]+ phe_L[e] -> na1[c]+ phe_L[c]
PROt2r		h[e]+ pro_L[e] <=> h[c]+ pro_L[c]
PROt4		na1[e]+ pro_L[e] -> na1[c]+ pro_L[c]
SELMETHte		na1[e]+ selmeth[e] -> na1[c]+ selmeth[c]
SERt4		na1[e]+ ser_L[e] -> na1[c]+ ser_L[c]
THRt4		na1[e]+ thr_L[e] -> na1[c]+ thr_L[c]
r1492		k[c] -> k[e]
r1518		h2o[c]+ atp[c]+ lnc[c] -> adp[c]+ h[c]+ pi[c]+ lnc[e]
r2471		h[e]+ ser_L[e] -> h[c]+ ser_L[c]
DM_pe_hs[c]		
EX_btn[e]		
EX_ascb_L[e]		
EX_chol[e]		
EX_fol[e]		

Mechanistic model of dopaminergic neuron metabolism

Reaction abbreviation	Reaction description	Reaction formula
EX_glc_D[e]		
EX_gly[e]		
EX_gthrd[e]		
EX_glu_L[e]		
EX_etha[e]		
EX_hxan[e]		
EX_tyr_L[e]		
EX_phe_L[e]		
EX_ala_L[e]		
EX_pro_L[e]		
EX_thr_L[e]		
EX_asn_L[e]		
EX_ile_L[e]		
EX_his_L[e]		
EX_lys_L[e]		
EX_ser_L[e]		
EX_asp_L[e]		
EX_Lcystin[e]		
EX_pnto_R[e]		
EX_inost[e]		
EX_thm[e]		
EX_pydxn[e]		
EX_pyr[e]		
EX_ribflv[e]		
EX_thymd[e]		
EX_ca2[e]		
EX_cl[e]		
EX_arg_L[e]		
EX_M02482[e]		
EX_cys_L[e]		
EX_k[e]		
EX_na1[e]		
EX_hco3[e]		
EX_gln_L[e]		
EX_inlc[e]		
EX_leu_L[e]		
EX_met_L[e]		
EX_val_L[e]		
EX_trp_L[e]		

Chapter 5

Reaction abbreviation	Reaction description	Reaction formula
EX_HC02172[e]		
EX_ncam[e]		
EX_ptrc[e]		
EX_pi[e]		
EX_so4[e]		
EX_lipoate[e]		
EX_M02887[e]		
EX_fe3[e]		
EX_nh4[e]		
EX_gal[e]		
EX_orn[e]		
EX_lac_L[e]		
EX_cit[e]		
EX_CE2028[e]		
EX_succ[e]		
EX_o2[e]	exchange reaction for oxygen	$o2[e] \rightleftharpoons$
H2Ot	H2O transport via diffusion	$h2o[e] \rightleftharpoons h2o[c]$
O2t	o2 transport [diffusion]	$o2[e] \rightleftharpoons o2[c]$
CO2t	CO2 transporter via diffusion	$co2[e] \rightleftharpoons co2[c]$
EX_co2[e]	CO2 exchange	$co2[e] \rightleftharpoons$
EX_h2o[e]	H2O exchange	$h2o[e] \rightleftharpoons$
GBAl	glucocerebrosidase	$h2o[l] + gluside_hs[l] \rightarrow crm_hs[l] + glc_D[l]$
GTPCI	GTP cyclohydrolase I	$h2o[c] + gtp[c] \rightarrow h[c] + for[c] + ahdt[c]$
PTHPS	6-pyruvoyltetrahydropterin synthase	$ahdt[c] \rightarrow ppp[c] + 6p[thp[c]$
SPR	sepiapterin reductase	$2.0\ h[c] + 2.0\ nadph[c] + 6p[thp[c] \rightarrow 2.0\ nadp[c] + thbpt[c]$
34HPLtm		$h[c] + 34hpl[c] \rightleftharpoons h[m] + 34hpl[m]$
4ABUTtm		$4abut[c] \rightleftharpoons 4abut[m]$
DAGK_hs		
AACOAT		
ABTArm		$akg[m] + 4abut[m] \rightleftharpoons glu_L[m] + succsa[m]$
ABUTt2r	$h[e] + 4abut[e] \rightleftharpoons h[c] + 4abut[c]$	
ABUTt4_2_r	$2\ na1[e] + 4abut[e] \rightleftharpoons 2\ na1[c] + 4abut[c]$	
ACGLUtm		$acglu[c] \rightleftharpoons acglu[m]$
ACHEe		
ADK1m		$atp[m] + amp[m] \rightleftharpoons 2\ adp[m]$
AGTtm		$ala_L[m] + gb[c] \rightarrow pyr[m] + gly[m]$
AKGDm		$akg[m] + coa[m] + nad[m] \rightarrow co2[m] + nadh[m] + succoa[m]$
AKGMALtm		$akg[m] + mal_L[c] \rightleftharpoons akg[c] + mal_L[m]$
ALAt2r		$h[e] + ala_L[e] \rightleftharpoons h[c] + ala_L[c]$

Mechanistic model of dopaminergic neuron metabolism

Reaction abbreviation	Reaction description	Reaction formula
ALOX5		
AMET12m		$\text{amet}[c] + \text{ahcys}[m] \rightleftharpoons \text{ahcys}[c] + \text{amet}[m]$
APOC_LYS_BTNP m		$\text{h2o}[m] + \text{apoC_Lys_btn}[m] \rightarrow \text{apoC}[m] + \text{biocyt}[m]$
ARGNm		$\text{h2o}[m] + \text{arg_L}[m] \rightarrow \text{urea}[m] + \text{orn}[m]$
ASPLUUm		$\text{h}[c] + \text{glu_L}[c] + \text{asp_L}[m] \rightarrow \text{h}[m] + \text{glu_L}[m] + \text{asp_L}[c]$
ASPNATm		$\text{asp_L}[m] + \text{accoa}[m] \rightarrow \text{h}[m] + \text{coa}[m] + \text{Nacasp}[m]$
ASPTAm		$\text{akg}[m] + \text{asp_L}[m] \rightleftharpoons \text{glu_L}[m] + \text{oaa}[m]$
ATPS4mi		$4 \text{ h}[c] + \text{adp}[m] + \text{pi}[m] \rightarrow 3 \text{ h}[m] + \text{h2o}[m] + \text{atp}[m]$
BDHm		$\text{nad}[m] + \text{bbh}[m] \rightleftharpoons \text{h}[m] + \text{nadh}[m] + \text{acac}[m]$
C02712tm		$\text{C02712}[c] \rightleftharpoons \text{C02712}[m]$
C09642te		$\text{C09642}[c] \rightleftharpoons \text{C09642}[e]$
C160CPT1		
C160CPT2		$\text{coa}[m] + \text{pmtcrn}[m] \rightarrow \text{crn}[m] + \text{pmtcoa}[m]$
CATm		$2 \text{ h2o2}[m] \rightarrow \text{o2}[m] + 2 \text{ h2o}[m]$
CHSTEROLt		$\text{h2o}[c] + \text{atp}[c] + \text{chsterol}[c] \rightarrow \text{adp}[c] + \text{h}[c] + \text{pi}[c] + \text{chsterol}[e]$
CK		$\text{atp}[m] + \text{creat}[m] \rightleftharpoons \text{adp}[m] + \text{pcreat}[m]$
CKc		
CSm		$\text{h2o}[m] + \text{accoa}[m] + \text{oaa}[m] \rightarrow \text{h}[m] + \text{coa}[m] + \text{cit}[m]$
CSNAT2m		$\text{coa}[m] + \text{pcrn}[m] \rightleftharpoons \text{ppcoa}[m] + \text{crn}[m]$
CSNATm		$\text{coa}[m] + \text{acrn}[m] \rightleftharpoons \text{accoa}[m] + \text{crn}[m]$
CYOOm2i		$\text{o2}[m] + 4 \text{ ficytC}[m] \rightarrow 4 \text{ h}[m] + 2 \text{ h2o}[m] + 4 \text{ ficytC}[m]$
CYOOm3i		$7.92 \text{ h}[m] + \text{o2}[m] + 4 \text{ ficytC}[m] \rightarrow 1.96 \text{ h2o}[m] + 4 \text{ h}[c] + 4 \text{ ficytC}[m] + 0.02 \text{ o2s}[m]$
CYOR_u10mi		$2 \text{ h}[m] + 2 \text{ ficytC}[m] + \text{q10h2}[m] \rightarrow 4 \text{ h}[c] + \text{q10}[m] + 2 \text{ ficytC}[m]$
DURIK1m		$\text{atp}[m] + \text{duri}[m] \rightarrow \text{h}[m] + \text{adp}[m] + \text{dump}[m]$
ENO		
ETF		$\text{fadh2}[m] + \text{etfox}[m] \rightarrow \text{fad}[m] + \text{etfrd}[m]$
EX_4abut[e]		
EX_arachd[e]		
EX_atp[e]		
EX_bbh[e]		
EX_chsterol[e]		
EX_gln_L[e]		
EX_h[e]		
EX_h2o2[e]		
EX_hco3[e]		
EX_k[e]		
EX_lneldc[e]		
EX_lnlncg[e]		
EX_met_L[e]		

Reaction abbreviation	Reaction description	Reaction formula
EX_na1[e]		
EX_nh4[e]		
EX_pi[e]		
EX_strdnc[e]		
EX_thr_L[e]		
FADH2ETC		fadh2[m]+ q10[m] -> fad[m]+ q10h2[m] 7 h2o[m]+ 7 coa[m]+ 7 nad[m]+ 7 fad[m]+ pmtcoa[m] -> 7 h[m]+ 7 nadh[m]+ 8 accoa[m]+ 7 fadh2[m]
FAOXC160		
FBA		
FE2DMT1		h[e]+ fe2[e] -> h[c]+ fe2[c]
FUMm		h2o[m]+ fum[m] <=> mal_L[m]
G3PD2m		fad[m]+ glyc3p[c] -> fadh2[m]+ dhap[c]
G6PDH2c		
G6PDH2r		
GAPD		
GBA		
GBA2e		
GLCt1r		
GLNtm		gln_L[c] -> gln_L[m]
GLUNm		h2o[m]+ gln_L[m] -> glu_L[m]+ nh4[m]
GLUt2m		h[c]+ glu_L[c] <=> h[m]+ glu_L[m]
GLUVESSEC		
GLYKm		atp[m]+ glyc[m] -> h[m]+ adp[m]+ glyc3p[m]
GND		
GNDc		
GTHO		
GTHOm		h[m]+ nadph[m]+ gthox[m] -> nadp[m]+ 2 gthrd[m]
GTHP		
GTHPm		h2o2[m]+ 2 gthrd[m] -> 2 h2o[m]+ gthox[m]
H2CO3Dm		h2o[m]+ co2[m] -> h[m]+ hco3[m]
H2O2tm		h2o2[c] -> h2o2[m]
H2OGLYAQPt	h2o[e]+ glyc[e] <=> h2o[c]+ glyc[c]	
HEX1		
HISTAtu	hista[e] <=> hista[c]	
Htmi		h[c] -> h[m]
ICDHxm		nad[m]+ icit[m] -> akgl[m]+ co2[m]+ nadh[m]
ICDHyl		
ICDHyrn		nadp[m]+ icit[m] <=> nadph[m]+ akgl[m]+ co2[m]
L_LACt2r		
L_LACtcm		lac_L[c] -> lac_L[m]

Mechanistic model of dopaminergic neuron metabolism

Reaction abbreviation	Reaction description	Reaction formula
LDH_L		
LDH_Lm		$\text{nad}[\text{m}] + \text{lac_L}[\text{m}] \rightleftharpoons \text{h}[\text{m}] + \text{nadh}[\text{m}] + \text{pyr}[\text{m}]$
LGTHL		
LNSTLSr		
MDH		
MDHm		$\text{nad}[\text{m}] + \text{mal_L}[\text{m}] \rightleftharpoons \text{h}[\text{m}] + \text{nadh}[\text{m}] + \text{oaa}[\text{m}]$
ME2m		$\text{nadp}[\text{m}] + \text{mal_L}[\text{m}] \rightarrow \text{nadph}[\text{m}] + \text{co2}[\text{m}] + \text{pyr}[\text{m}]$
MTHFD2m		$\text{nad}[\text{m}] + \text{mlthf}[\text{m}] \rightleftharpoons \text{nadh}[\text{m}] + \text{methf}[\text{m}]$
NADH2_u10mi		$5 \text{ h}[\text{m}] + \text{nadh}[\text{m}] + \text{q10}[\text{m}] \rightarrow 4 \text{ h}[\text{c}] + \text{nad}[\text{m}] + \text{q10h2}[\text{m}]$
NMNATn		
OCOAT1m		$\text{acac}[\text{m}] + \text{succoa}[\text{m}] \rightarrow \text{aacoa}[\text{m}] + \text{succ}[\text{m}]$
ORNtDF		$\text{orn}[\text{e}] \rightarrow \text{orn}[\text{c}]$
P45011A1m		$\text{h}[\text{m}] + \text{nadph}[\text{m}] + 2 \text{ o2}[\text{m}] + \text{chsterol}[\text{m}] \rightarrow 2 \text{ h2o}[\text{m}] + \text{nadp}[\text{m}] + 4\text{mptnl}[\text{m}] + \text{prgnlone}[\text{m}]$
P45027A11m		$\text{h}[\text{m}] + \text{nadph}[\text{m}] + \text{o2}[\text{m}] + \text{xoltriol}[\text{m}] \rightarrow \text{h2o}[\text{m}] + \text{nadp}[\text{m}] + \text{xoltetrol}[\text{m}]$
P45027A12m		$\text{nadp}[\text{m}] + \text{xoltetrol}[\text{m}] \rightarrow \text{h}[\text{m}] + \text{nadph}[\text{m}] + \text{thcholst}[\text{m}]$
P45027A13m		$\text{nadph}[\text{m}] + \text{o2}[\text{m}] + \text{thcholst}[\text{m}] \rightarrow \text{h2o}[\text{m}] + \text{nadp}[\text{m}] + \text{thcholstoic}[\text{m}]$
P45027A14m		$\text{h}[\text{m}] + \text{nadph}[\text{m}] + \text{o2}[\text{m}] + \text{xol7ah2}[\text{m}] \rightarrow \text{h2o}[\text{m}] + \text{nadp}[\text{m}] + \text{xol7ah3}[\text{m}]$
P45027A15m		$\text{h}[\text{m}] + \text{nadph}[\text{m}] + \text{o2}[\text{m}] + \text{xol7ah3}[\text{m}] \rightarrow 2 \text{ h2o}[\text{m}] + \text{nadp}[\text{m}] + \text{xol7ah2al}[\text{m}]$
P45027A16m		$\text{nadph}[\text{m}] + \text{o2}[\text{m}] + \text{xol7ah2al}[\text{m}] \rightarrow \text{h2o}[\text{m}] + \text{nadp}[\text{m}] + \text{dhcholestanate}[\text{m}]$
P45027A1m		$\text{h}[\text{m}] + \text{nadph}[\text{m}] + \text{o2}[\text{m}] + \text{chsterol}[\text{m}] \rightarrow \text{h2o}[\text{m}] + \text{nadp}[\text{m}] + \text{xol27oh}[\text{m}]$
PDHm		$\text{coa}[\text{m}] + \text{nad}[\text{m}] + \text{pyr}[\text{m}] \rightarrow \text{co2}[\text{m}] + \text{nadh}[\text{m}] + \text{accoa}[\text{m}]$
PEFLIPm		$\text{h2o}[\text{c}] + \text{atp}[\text{c}] + \text{pe_hs}[\text{c}] \rightarrow \text{adp}[\text{c}] + \text{h}[\text{c}] + \text{pi}[\text{c}] + \text{pe_hs}[\text{m}]$
PETOHHm_hs		$3 \text{ amet}[\text{m}] + \text{pe_hs}[\text{m}] \rightarrow 3 \text{ h}[\text{m}] + 3 \text{ ahcys}[\text{m}] + \text{pchol_hs}[\text{m}]$
PFK		
PGI		
PGK		
PGL		
PGLc		
PGM		
PPM		
PSDm_hs		$\text{h}[\text{m}] + \text{ps_hs}[\text{m}] \rightarrow \text{co2}[\text{m}] + \text{pe_hs}[\text{m}]$
PYK		
r0022		$\text{nad}[\text{m}] + 2 \text{ gthrd}[\text{m}] \rightleftharpoons \text{h}[\text{m}] + \text{nadh}[\text{m}] + \text{gthox}[\text{m}]$
r0081		$\text{akg}[\text{m}] + \text{ala_L}[\text{m}] \rightleftharpoons \text{glu_L}[\text{m}] + \text{pyr}[\text{m}]$
r0083		$\text{h}[\text{m}] + \text{HC01434}[\text{m}] \rightarrow \text{akg}[\text{m}] + \text{co2}[\text{m}]$
r0321		$\text{coa}[\text{m}] + \text{acac}[\text{m}] + \text{atp}[\text{m}] \rightarrow \text{aacoa}[\text{m}] + \text{amp}[\text{m}] + \text{ppi}[\text{m}]$
r0399		

Chapter 5

Reaction abbreviation	Reaction description	Reaction formula
r0407		
r0408		
r0409		
r0423		$\text{nadp[m]} + \text{icit[m]} \rightarrow \text{h[m]} + \text{nadph[m]} + \text{HC01434[m]}$
r0425		$\text{nad[m]} + \text{icit[m]} \rightleftharpoons \text{h[m]} + \text{nadh[m]} + \text{HC01434[m]}$
r2535m		$\text{hom_L[m]} \rightleftharpoons \text{hom_L[c]}$
r2539		$\text{L2aadp[m]} + \text{L2aadp6sa[c]} \rightleftharpoons \text{L2aadp[c]} + \text{L2aadp6sa[m]}$
RBK		
RBK_D		
RE0124C		
RE1530M		$\text{dgtp[m]} + \text{duri[m]} \rightleftharpoons \text{h[m]} + \text{dgdp[m]} + \text{dump[m]}$
RE1804M		$\text{nad[m]} + \text{xol7ah3[m]} \rightleftharpoons \text{h[m]} + \text{nadh[m]} + \text{xol7ah2al[m]}$ $\text{h2o[m]} + \text{nadp[m]} + \text{xol7ah2al[m]} \rightleftharpoons 2 \text{h[m]} + \text{nadph[m]} + \text{dhcholestanate[m]}$
RE1807M		
RE2625M		$\text{nad[m]} + \text{xoltetrol[m]} \rightleftharpoons \text{h[m]} + \text{nadh[m]} + \text{CE4872[m]}$ $\text{h2o[m]} + \text{nadp[m]} + \text{CE4872[m]} \rightleftharpoons 2 \text{h[m]} + \text{nadph[m]} + \text{thcholstoic[m]}$
RE2626M		
RE3251M		$\text{h[m]} + \text{nadph[m]} + \text{o2[m]} + \text{xoltetrol[m]} \rightarrow \text{h2o[m]} + \text{nadp[m]} + \text{CE4874[m]}$
RPE		
RPEc		
RPI		
SARDHm		$\text{fad[m]} + \text{sarcs[m]} + \text{thf[m]} \rightarrow \text{fadh2[m]} + \text{gly[m]} + \text{mlthf[m]}$
SPODMm		$2 \text{h[m]} + 2 \text{o2s[m]} \rightarrow \text{o2[m]} + \text{h2o2[m]}$
SRTNtu	$\text{srt[n[e]} \rightleftharpoons \text{srt[n[c]}$	
STS1		
SUCD1m		$\text{fad[m]} + \text{succ[m]} \rightleftharpoons \text{fadh2[m]} + \text{fum[m]}$
SUCOASm		$\text{coa[m]} + \text{atp[m]} + \text{succ[m]} \rightleftharpoons \text{adp[m]} + \text{pi[m]} + \text{succoa[m]}$
TALA		
TAUPAT1c	$\text{h[e]} + \text{taur[e]} \rightarrow \text{h[c]} + \text{taur[c]}$	
TKT1		
TKT2		
TMDK1m		$\text{atp[m]} + \text{thymd[m]} \rightarrow \text{h[m]} + \text{adp[m]} + \text{dtmp[m]}$
DM_atp_c_		
DM_4glu56dihdi		
nd[c]	4-S-Glutathionyl-5,6-Dihydroxyindoline	
DM_CE5025[c]	5-S-Glutathionyl-Dopamine	
DM_CE5026[c]	5-S-Glutathionyl-L-Dopa	
DM_CE1562[c]	5,6-Indolequinone-2-Carboxylate	
DM_ind56qn[c]	Indole-5,6-Quinone	
DM_ascb_L[c]	L-Ascorbate	
sink_asn_L[c]	L-Asparagine	

Mechanistic model of dopaminergic neuron metabolism

Reaction abbreviation	Reaction description	Reaction formula
DM_no2[c]	Nitrite	
sink_thmtp[c]	Thiamine-Triphosphate	
DM_6hddopaqn[c]	6-Hydroxydopamine-Quinone	
sink_chol[c]	Choline	
DM_CE4888[c]	Dopaminochrome	
DM_gm1_hs[n]	Ganglioside Gm1	
sink_pre_prot[r]	Glycophosphatidylinositol (Gpi)-Anchored Protein Precursor	
sink_fe3[c]	Iron (Fe3+)	
sink_citr[c]	L-Citrulline	
sink_lnlc[c]	Linoleate	
sink_lnlcoa[c]	Linoleic Coenzyme A	
DM_C02712[c]	N-Acetylmethionine	
sink_nadp[c]	Nicotinamide Adenine Dinucleotide Phosphate	
sink_nad[c]	Nicotinamide Adenine Dinucleotide	
sink_odecoa[c]	Octadecenoyl Coenzyme A (N-C18:1 Coenzyme A)	
sink_pmtcoa[c]	Palmitoyl Coenzyme A (N-C16:0 Coenzyme A)	
DM_pchol_hs[c]	Phosphatidylcholine	
DM_pcreat[c]	Phosphocreatine	
sink_phyQ[c]	Phylloquinone	
DM_K_c_	Potassium	
DM_Ser_Gly_Ala_X_Gly_ly_	Protein-Linked Serine Residue (Glycosaminoglycan Attachment Site)	
sink_Ser_Gly_Ala_X_Gly[r]	Protein-Linked Serine Residue (Glycosaminoglycan Attachment Site)	
DM_na1[c]	Sodium	
DM_sprm_c_	Spermine	
sink_stcoa[c]	Stearoyl Coenzyme A (N-C18:0 Coenzyme A)	
DM_taur[c]	Taurine	
DM_thm[m]	Thiamin	
sink_thmpp[c]	Thiamine Diphosphate	
sink_Tyr_ggn[c]	Tyr-194 Of Apo-Glycogenin Protein (Primer For Glycogen Synthesis)	
DM_ca2[c]	Calcium	
DM_chsterol[c]	cholesterol	
DM_pnto_R	(R)-Pantothenate	
DM_CE1261[c]	5-S-Cysteinylidopa	
sink_CE1273[c]	5Beta-Cholestane-3Alpha,7Alpha,12Alpha,24S,25-Pentol	
sink_crvnc[c]	Cervonic Acid, C22:6 N-3	
sink_dec dicoa[e]	Decadienoyl Coenzyme A	
sink_c101coa[c]	Decenoyl Coenzyme A	
DM_pail35p_hs[n]	1-Phosphatidyl-1D-Myo-Inositol 3,5-Bisphosphate	

Chapter 5

Reaction abbreviation	Reaction description	Reaction formula
DM_5cysdopa[c]	5-S-Cysteinyldopamine	
sink_retfa[c]	Fatty Acid Retinol	
HMR_1735		$o2[c] + h[c] + nadph[c] + chsterol[c] \rightarrow h2o[c] + nadp[c] + xol24oh[c]$
CL5_hs		$cdpdag_hs[c] + pglyc_hs[c] \rightarrow h[c] + cmp[c] + clpn_hs[c]$
PIK4		$atp[c] + pail_hs[c] \rightarrow h[c] + adp[c] + pail4p_hs[c]$
PSDm_hs		$h[m] + ps_hs[m] \rightarrow co2[m] + pe_hs[m]$
HMR_0653		$amet[c] + pe_hs[c] \rightarrow 2\ h[c] + ahcys[c] + M02686[c]$
PCHOLP_hs		$h2o[c] + pchol_hs[c] \rightarrow h[c] + pa_hs[c] + chol[c]$
PLA2_2		$h2o[c] + pchol_hs[c] \rightarrow h[c] + Rtotal2[c] + lpchol_hs[c]$
SMS		$pchol_hs[c] + crm_hs[c] \rightarrow dag_hs[c] + sphmyln_hs[c]$
HMR_0795		$h2o[c] + sphmyln_hs[c] \rightarrow h[c] + crm_hs[c] + cholp[c]$
NTD4		$h2o[c] + cmp[c] \rightarrow pi[c] + cytd[c]$
AMPDA		$h2o[c] + h[c] + amp[c] \rightarrow nh4[c] + imp[c]$
NTD7		$h2o[c] + amp[c] \rightarrow pi[c] + adn[c]$
NTD9		$h2o[c] + gmp[c] \rightarrow pi[c] + gsn[c]$
NTD2		$h2o[c] + ump[c] \rightarrow pi[c] + uri[c]$
ASPTA		$akg[c] + asp_L[c] \rightleftharpoons glu_L[c] + oaa[c]$
FPGS2		$atp[c] + glu_L[c] + 5thf[c] \rightarrow h[c] + adp[c] + pi[c] + 6thf[c]$
FPGS3		$atp[c] + glu_L[c] + 6thf[c] \rightarrow h[c] + adp[c] + pi[c] + 7thf[c]$
FPGS4		$atp[c] + 4\ glu_L[c] + dhf[c] \rightarrow 3\ h2o[c] + h[c] + adp[c] + pi[c] + 5dhf[c]$
FPGS5		$atp[c] + glu_L[c] + 5dhf[c] \rightarrow h[c] + adp[c] + pi[c] + 6dhf[c]$
FPGS6		$atp[c] + glu_L[c] + 6dhf[c] \rightarrow h[c] + adp[c] + pi[c] + 7dhf[c]$
FPGS8		$10fthf5glu[c] + atp[c] + glu_L[c] \rightarrow 10fthf6glu[c] + h[c] + adp[c] + pi[c]$
FPGS9		$10fthf6glu[c] + atp[c] + glu_L[c] \rightarrow 10fthf7glu[c] + h[c] + adp[c] + pi[c]$
r1382		$6\ atp[c] + 6\ glu_L[c] + thf[c] \rightarrow 6\ h[c] + 6\ adp[c] + 6\ pi[c] + 7thf[c]$
FPGS7		$10fthf[c] + atp[c] + 4\ glu_L[c] \rightarrow 10fthf5glu[c] + 3\ h2o[c] + h[c] + adp[c] + pi[c]$
FPGS		$atp[c] + 4\ glu_L[c] + thf[c] \rightarrow 3\ h2o[c] + h[c] + adp[c] + pi[c] + 5thf[c]$
HMR_9726		$glu_L[c] + 5fthf[c] \rightarrow thf[c] + forglu[c]$
ALATA_L		$akg[c] + ala_L[c] \rightleftharpoons pyr[c] + glu_L[c]$
GLUCYS		$atp[c] + glu_L[c] + cys_L[c] \rightarrow h[c] + adp[c] + pi[c] + glucys[c]$
ILETA		$akg[c] + ile_L[c] \rightleftharpoons glu_L[c] + 3mop[c]$
LEUTA		$akg[c] + leu_L[c] \rightleftharpoons glu_L[c] + 4mop[c]$
VALTA		$akg[c] + val_L[c] \rightleftharpoons glu_L[c] + 3mob[c]$
ARGSS		$atp[c] + asp_L[c] + citr_L[c] \rightarrow h[c] + amp[c] + ppi[c] + argsuc[c]$
ASPTA		$akg[c] + asp_L[c] \rightleftharpoons glu_L[c] + oaa[c]$
ALATA_L		$akg[c] + ala_L[c] \rightleftharpoons pyr[c] + glu_L[c]$
GHMT2r		$ser_L[c] + thf[c] \rightleftharpoons h2o[c] + gly[c] + mlthf[c]$
GLYAMDTRc		$gly[c] + arg_L[c] \rightleftharpoons orn[c] + gudac[c]$

Mechanistic model of dopaminergic neuron metabolism

Reaction abbreviation	Reaction description	Reaction formula
r0060		ser_L[c] -> nh4[c] + pyr[c]
THRD_L		thr_L[c] -> nh4[c] + 2obut[c]
GHMT2r		ser_L[c] + thf[c] <=> h2o[c] + gly[c] + mlthf[c]
GLYAMDTRc		gly[c] + arg_L[c] <=> orn[c] + gudac[c]
ARGN		h2o[c] + arg_L[c] -> orn[c] + urea[c]
r0145		2 o2[c] + 2 nadph[c] + arg_L[c] -> 2 h2o[c] + 2 nadp[c] + citr_L[c] + no[c]
THRD_L		thr_L[c] -> nh4[c] + 2obut[c]
PHETHPTOX2		o2[c] + thbpt[c] + phe_L[c] -> tyr_L[c] + thbpt4acam[c]
PROD2		fad[c] + pro_L[c] -> h[c] + fadh2[c] + 1pyr5c[c]
HMR_6728		o2[c] + thbpt[c] + tyr_L[c] -> h2o[c] + 34dhphe[c] + dhbpt[c]
HMR_6874		o2[c] + tyr_L[c] -> h2o[c] + dopaqn[c]
TYR3MO2		o2[c] + thbpt[c] + tyr_L[c] -> 34dhphe[c] + thbpt4acam[c]
TYRCBOX		h[c] + tyr_L[c] -> co2[c] + tym[c]
TYRTA		akg[c] + tyr_L[c] <=> 34hpp[c] + glu_L[c]
LYSOXp		
HISDC		h[c] + his_L[c] -> co2[c] + hista[c]
HISD		his_L[c] -> nh4[c] + urcan[c]
LEUTA		akg[c] + leu_L[c] <=> glu_L[c] + 4mop[c]
ILETA		akg[c] + ile_L[c] <=> glu_L[c] + 3mop[c]
VALTA		akg[c] + val_L[c] <=> glu_L[c] + 3mob[c]
METAT		h2o[c] + atp[c] + met_L[c] -> pi[c] + amet[c] + ppi[c]
GLUDxm	glutamate dehydrogenase (NAD) (mitochondrial)	h2o[m] + nad[m] + glu_L[m] <=> h[m] + akg[m] + nadh[m] + nh4[m]
GLUDym	glutamate dehydrogenase (NADP), mitochondrial	h2o[m] + nadp[m] + glu_L[m] <=> h[m] + nadph[m] + akg[m] + nh4[m]
ACONTm		
r0426		
r1109		
MTHFR3		
METS		
DOPAOQNOX		

Chapter 5

Table S3. List of metabolites predicted to be secreted along with their psychochemical nature and biological class.

Recon3D Names	Hydrophobicity	Chemical Class
Lipoamide	Apolar / hydrophobic	Lipoamide
17beta-hydroxy-5alpha-androstan-3-one	Apolar / hydrophobic	Steroid
Dolichol	Apolar / hydrophobic	Polyprenol
Lanosterin	Apolar / hydrophobic	Sterol
Bilirubin	Predicted informative metabolites	Bilirubin
Thromboxane A2	Apolar / hydrophobic	Eicosanoid
Retinoyl b-glucuronide	Apolar / hydrophobic	Terpene glycoside
21-Hydroxypregnenolone	Apolar / hydrophobic	Steroid
7-a,27-Dihydroxycholesterol	Apolar / hydrophobic	Bile acid
(24S)-7Alpha,24-Dihydroxycholesterol	Apolar / hydrophobic	Bile acid
Hexanoyl-CoA	Apolar / hydrophobic	Fatty acyl Coenzyme
Ceramide 1-Phosphate	Apolar / hydrophobic	Sphingolipid
Stearidonic acid	Apolar / hydrophobic	Fatty acid
12,13-DHOME	Apolar / hydrophobic	Fatty acid
24-Hydroxycholesterol	Apolar / hydrophobic	Sterol
Cholesterol	Apolar / hydrophobic	Sterol
Cholesterol ester	Apolar / hydrophobic	Sterol
Gamma-Linolenic acid	Apolar / hydrophobic	Fatty acid
Vitamin K1	Apolar / hydrophobic	Prenol Lipid
Sphinganine 1-phosphate	Apolar / hydrophobic	Sphingolipid
Cholestane-3,7,12,24,25-pentol	Apolar / hydrophobic	Sterol
7Z,10Z-Hexadecadienoic Acid	Apolar / hydrophobic	Fatty acid
Linoelaidic Acid (All Trans C18:2)	Apolar / hydrophobic	Fatty acid
13-cis-Retinoic acid	Apolar / hydrophobic	Prenol lipid
Retinal	Apolar / hydrophobic	Prenol Lipid
5-Amino-1-(5-Phospho-D-ribose)imidazole-4-carboxamide	Polar / hydrophilic	Nucleotide
5-Methylthioadenosine	Polar / hydrophilic	Nucleoside
S-Adenosyl-L-homocysteine	Polar / hydrophilic	Nucleoside
Urocanate	Polar / hydrophilic	Imidazol carboxylic acid

Mechanistic model of dopaminergic neuron metabolism

Recon3D Names	Hydrophobicity	Chemical Class
Kynurenate	Polar / hydrophilic	Quinoline carboxylic acid
2-Oxobutanoate	Polar / hydrophilic	Keto acid
2-Oxoadipate(2-)	Polar / hydrophilic	Keto acid
Leucylleucine	Polar / hydrophilic	Dipeptide
Dehydroascorbic acid	Polar / hydrophilic	Sugar acid
4-Acetamidobutanoic acid	Polar / hydrophilic	Amino acid
Lactose	Polar / hydrophilic	Sugar
D-Xylose	Polar / hydrophilic	Sugar
D-Iduronic acid	Polar / hydrophilic	Sugar acid
4-Methylpentanal	Polar / hydrophilic	Aldehyde
D-Glucuronic acid	Polar / hydrophilic	Sugar acid
Orotidylic acid	Polar / hydrophilic	Nucleotide
Galactose	Polar / hydrophilic	Sugar
Adenosine triphosphate	Polar / hydrophilic	Nucleotide
N-Acetylmannosamine	Polar / hydrophilic	Amino sugar
Isomaltose	Polar / hydrophilic	Sugar
Nicotinic acid mononucleotide	Polar / hydrophilic	Nucleotide
N-Acetyl-L-aspartic acid	Polar / hydrophilic	Amino acid
N-Acetylglutamic acid	Polar / hydrophilic	Amino acid
1-Methylnicotinamide	Polar / hydrophilic	Nicotinamide
Tetrahydrobiopterin	Polar / hydrophilic	Pterin
D-Gluconate	Polar / hydrophilic	Sugar acid
Inosine	Polar / hydrophilic	Nucleoside
Xanthine	Polar / hydrophilic	Nucleobase
Alpha-D-Glucose 1,6-bisphosphate	Polar / hydrophilic	Sugar phosphate
1,5-D-Gluconolactone	Polar / hydrophilic	Sugar lactone
2-Methyl-3-Hydroxy-Valerate	Polar / hydrophilic	Carboxylic acid
Hydroxytyrosol	Polar / hydrophilic	Phenyl alcohol

Supplementary References

1. Inglis A.S. and Teh-Yung Liu. "The Stability of Cysteine and Cystine during Acid Hydrolysis of Proteins and Peptides". In: *The Journal of Biological Chemistry* 245.1 (1970), pp. 112–116. url: <http://w/ww.jbc.org/content/245/1/112> (visited on 04/13/2018).
2. Linda Ahonen et al. "Analysis of Oxysterols and Vitamin D Metabolites in Mouse Brain and Cell Line Samples by Ultra-High-Performance Liquid Chromatography-Atmospheric Pressure Photoionization–MassSpectrometry". In: *Journal of Chromatography A* 1364(Oct.2014),pp.214– 222. doi: 10.1016/j.chroma.2014.08.088. pmid: 25204266.
3. Rodrigo D. A. M. Alvesetal. "Global Profiling of the Muscle Metabolome : Method Optimization, Validation and Application to Determine Exercise-Induced Metabolic Effects". In: *Metabolomics* 11.2 (Apr. 1, 2015), pp. 271–285. doi: 10.1007/s11306-014-0701-7.
4. Maïke K. Aurich, Ronan M. T. Fleming, and Ines Thiele. "MetaboTools: A Comprehensive Toolbox for Analysis of Genome-Scale Metabolic Models". In: *Frontiers in Physiology* 7 (2016). doi: 10.3389/fphys.2016.00327.
5. M. Banay-Schwartz et al. "Protein Content of Various Regions of Rat Brain and Adult and Aging Human Brain". In: *AGE* 15.2 (Apr. 1, 1992), pp. 51–54. doi: 10.1007/BF02435024.
6. O. Begou et al. "Hyphenated MS-Based Targeted Approaches in Metabolomics". In: *The Analyst* 2011 (2017), pp. 3079–3100. doi: 10.1039/C7AN00812K. pmid: 28792021.
7. Elizabeth Brunketal. "Recon3D Enables a Three-Dimensional View of Gene Variation in Human Metabolism". In: *Nature Biotechnology* 36 (Feb. 19, 2018), p. 272. url: <http://dx.doi.org/10.1038/nbt.4072>.
8. Tomas Cajka and Oliver Fiehn. "Comprehensive Analysis of Lipids in Biological Systems by Liquid Chromatography-Mass Spectrometry". In: *TrAC Trends in Analytical Chemistry* 61 (Oct. 2014), pp. 192–206. doi: 10.1016/j.trac.2014.04.017. pmid: 25309011.
9. Tomas Cajka, Jennifer T. Smilowitz, and Oliver Fiehn. "Validating Quantitative Untargeted LipidomicsAcrossNineLiquidChromatography–High ResolutionMassSpectrometryPlatforms". In: *Analytical Chemistry* 89.22 (Nov.

- 2017), pp. 12360–12368. doi: 10.1021/acs.analchem.7b03404. pmid: 29064229.
10. AlexiesDagnino-Subiabreetal. "Glutathione Transferase M22 Catalyzes Conjugation of Dopamine and Dopa o-Quinones". In: *Biochemical and Biophysical Research Communications* 274.1 (July 21, 2000), pp. 32–36. doi: 10.1006/bbrc.2000.3087.
11. Carola W. N. Damen et al. "Enhanced Lipid Isomer Separation in Human Plasma Using Reversed-Phase UPLC with Ion-Mobility/High-Resolution MS Detection". In: *Journal of Lipid Research* 55.8 (Aug. 2014), pp. 1772–1783. doi: 10.1194/jlr.D047795. pmid: 24891331.
12. F. Diego et al. "Automated Identification of Neuronal Activity from Calcium Imaging by Sparse Dictionary Learning". In: *2013 IEEE 10th International Symposium on Biomedical Imaging (ISBI)*. 2013 IEEE 10th International Symposium on Biomedical Imaging (ISBI). Apr. 2013, pp. 1058–1061. doi: 10.1109/ISBI.2013.6556660.
13. Arthur Edelstein et al. "Computer Control of Microscopes Using μ Manager". In: *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc., 2001. url: <http://onlinelibrary.wiley.com/doi/10.1002/0471142727.mb1420s92/abstract> (visited on 04/14/2015).
14. Graeme Eisenhofer, Irwin J. Kopin, and David S. Goldstein. "Catecholamine Metabolism: A Contemporary View with Implications for Physiology and Medicine". In: *Pharmacological Reviews* 56.3 (Jan. 9, 2004), pp. 331–349. doi: 10.1124/pr.56.3.1. pmid: 15317907.
15. Adam M Feist and Bernhard O Palsson. "The Biomass Objective Function". In: *Current Opinion in Microbiology*. Ecology and Industrial Microbiology • Special Section: Systems Biology 13.3 (June 2010), pp. 344–349. doi: 10.1016/j.mib.2010.03.003.
16. Junzeng Fu et al. "Metabolomics Profiling of the Free and Total Oxidised Lipids in Urine by LCMS/MS: Application in Patients with Rheumatoid Arthritis". In: *Analytical and Bioanalytical Chemistry* 408.23 (2016), pp. 6307–6319. doi: 10.1007/s00216-016-9742-2. pmid: 27405874.
17. Juan C. García-Cañaveras et al. "Targeted Profiling of Circulating and Hepatic Bile Acids in Human, Mouse, and Rat Using a UPLC-MRM-MS-Validated

- Method". In: *Journal of Lipid Research* 53.10 (2012), pp. 2231–2241. doi: 10.1194/jlr.D028803. pmid: 22822028.
18. Teresa A. Garrett, Ziqiang Guan, and Christian R H Raetz. "Analysis of Ubiquinones, Dolichols, and Dolichol Diphosphate-Oligosaccharides by Liquid Chromatography-Electrospray IonizationMass Spectrometry". In: *Methods in Enzymology* 432.07 (2007), pp. 117–143. doi: 10.1016/S0076-6879(07)32005-3. pmid: 17954215.
 19. Helen G. Gika et al. "Quantitative Profiling of Polar Primary Metabolites Using Hydrophilic Interaction Ultrahigh Performance Liquid Chromatography-Tandem Mass Spectrometry". In: *Journal of Chromatography A* 1259 (2012), pp. 121–127. doi: 10.1016/j.chroma.2012.02.010. pmid: 22381888.
 20. Nicola Gray et al. "High-Speed Quantitative UPLC-MS Analysis of Multiple Amines in Human PlasmaandSerumviaPrecolumnDerivatizationwith6-Aminoquinolyl-N-Hydroxysuccinimidyl Carbamate: Application to Acetaminophen-Induced Liver Failure". In: *Analytical Chemistry* 89.4 (Feb. 2017), pp. 2478–2487. doi: 10.1021/acs.analchem.6b04623. pmid: 28194962.
 21. Steinn Gudmundsson and Ines Thiele. "Computationally Efficient Flux Variability Analysis". In: *BMC Bioinformatics* 11.1 (Sept. 29, 2010), p. 489. doi: 10.1186/1471-2105-11-489.
 22. Jennifer Haggarty and Karl EV Burgess. "Recent Advances in Liquid and Gas Chromatography Methodology for Extending Coverage of the Metabolome". In: *Current Opinion in Biotechnology* 43 (2017), pp. 77–85. doi: 10.1016/j.copbio.2016.09.006. pmid: 27771607.
 23. J. A. Hanley and B. J. McNeil. "The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve". In: *Radiology* 143.1 (Apr. 1982), pp. 29–36. doi: 10.1148/radiology.143.1.7063747. pmid: 7063747.
 24. Hulda S. Haraldsdóttir and Ronan M. T. Fleming. "Identification of Conserved Moieties in Metabolic Networks by Graph Theoretical Analysis of Atom Transition Networks". In: *PLOS Computational Biology* 12.11 (Nov. 21, 2016), e1004999. doi: 10.1371/journal.pcbi.1004999.

25. Hulda S. Haraldsdóttir et al. "CHRR: Coordinate Hit-and-Run with Rounding for Uniform Sampling of Constraint-Based Models". In: *Bioinformatics* 33.11 (Jan. 6, 2017), pp. 1741–1743. doi: 10.1093/bioinformatics/btx052.
26. Laurent Heirendt et al. "Creation and Analysis of Biochemical Constraint-Based Models: The COBRA Toolbox v3.0". In: *Nature Protocols (accepted)* (2018). url: <https://arxiv.org/abs/1710.04038> (visited on 07/28/2017).
27. Katriina Itäaho et al. "Dopamine Is a Low-Affinity and High-Specificity Substrate for the Human UDP-Glucuronosyltransferase 1A10". In: *Drug Metabolism and Disposition* 37.4 (Jan. 4, 2009), pp. 768–775. doi: 10.1124/dmd.108.025692. pmid: 19116261.
28. Gregory Kapatos. "The Neurobiology of Tetrahydrobiopterin Biosynthesis: A Model for Regulation of GTP Cyclohydrolase I Gene Transcription within Nigrostriatal Dopamine Neurons". In: *IUBMB Life* 65.4 (Apr. 1, 2013), pp. 323–333. doi: 10.1002/iub.1140.
29. Hiroko Kato et al. "Widely Targeted Metabolic Profiling Analysis of Yeast Central Metabolites". In: *Journal of Bioscience and Bioengineering* 113.5 (May 2012), pp. 665–673. doi: 10.1016/j.jbiosc.2011.12.013. pmid: 22280965.
30. Daehwan Kim et al. "TopHat2: Accurate Alignment of Transcriptomes in the Presence of Insertions, Deletions and Gene Fusions". In: *Genome Biology* 14.4 (2013), R36. doi: 10.1186/gb2013-14-4-r36. pmid: 23618408.
31. Hee-Yong Kim, Bill X. Huang, and Arthur A. Spector. "Phosphatidylserine in the Brain: Metabolism and Function". In: *Progress in lipid research* 0 (Oct. 2014), pp. 1–18. doi: 10.1016/j.plipres.2014.06.002. pmid: 24992464.
32. Maud M Koek et al. "Microbial Metabolomics with Gas Chromatography/Mass Spectrometry." In: *Analytical Chemistry* 78.4 (Feb. 2006), pp. 1272–81. doi: 10.1021/ac051683+. pmid: 16478122.
33. Michael J. Kuhar. "On the Use of Protein Turnover and Half-Lives". In: *Neuropsychopharmacology* 34.5 (Oct. 15, 2008), pp. 1172–1173. doi: 10.1038/npp.2008.190.
34. Petri Kylli, Thomas Hankemeier, and Risto Kostianen. "Feasibility of Ultra-Performance Liquid Chromatography–Ion Mobility–Time-of-Flight Mass Spectrometry in Analyzing Oxysterols". In: *Journal of Chromatography A* 1487 (2017), pp. 147–152. doi: 10.1016/j.chroma.2017.01.039.

35. Ruth Landolt, Helen H. Hess, and Caroline Thalheimer. "Regional Distribution of Some Chemical Structural Components of the Human Nervous System—I". In: *Journal of Neurochemistry* 13.12 (Dec. 1, 1966), pp. 1441–1452. doi: 10.1111/j.1471-4159.1966.tb04305.x.
36. Heng Li et al. "The Sequence Alignment/Map Format and SAMtools". In: *Bioinformatics* 25.16 (Aug. 15, 2009), pp. 2078–2079. doi: 10.1093/bioinformatics/btp352. pmid: 19505943.
37. Gerhard Liebisch et al. "High Throughput Quantification of Cholesterol and Cholesteryl Ester by Electrospray Ionization Tandem Mass Spectrometry (ESI-MS/MS)". In: *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* 1761.1 (2006), pp. 121–128. doi: 10.1016/j.bbali.2005.12.007.
38. R. Mahadevan and C. H. Schilling. "The Effects of Alternate Optimal Solutions in ConstraintBased Genome-Scale Metabolic Models". In: *Metabolic Engineering* 5.4 (Oct. 2003), pp. 264–276. doi: 10.1016/j.ymben.2003.09.002.
39. Marcel Martin. "Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads". In: *EMBnet journal* 17.1 (May 2, 2011), pp. 10–12. doi: 10.14806/ej.17.1.200.
40. C. A. Mawson. "Meaning of 'Turnover' in Biochemistry". In: *Nature* 176.4476 (Aug. 13, 1955), pp. 317–317. doi: 10.1038/176317a0.
41. Johannes Meiser, Daniel Weindl, and Karsten Hiller. "Complexity of Dopamine Metabolism". In: *Cell Communication and Signaling : CCS* 11 (May 17, 2013), p. 34. doi: 10.1186/1478811X-11-34.
42. Alfred H. Merrill et al. "Sphingolipidomics: High-Throughput, Structure-Specific, and Quantitative Analysis of Sphingolipids by Liquid Chromatography Tandem Mass Spectrometry". In: *Methods* 36 (2 SPEC. ISS. 2005), pp. 207–224. doi: 10.1016/j.ymeth.2005.01.009. pmid: 15894491.
43. Patricia Muñoz et al. "Dopamine Oxidation and Autophagy". In: *Parkinson's Disease* 2012 (2012). doi: 10.1155/2012/920953.
44. A. Napolitano, P. Manini, and M. d'Ischia. "Oxidation Chemistry of Catecholamines and Neuronal Degeneration: An Update". In: *Current Medicinal Chemistry* 18.12 (2011), pp. 1832–1845. pmid: 21466469.
45. Alessandra Napolitano, Alessandro Pezzella, and Giuseppe Prota. "New Reaction Pathways of Dopamine under Oxidative Stress Conditions: Nonenzymatic Iron-Assisted Conversion to Norepinephrine and the

- Neurotoxins 6-Hydroxydopamine and 6,7-Dihydroxytetrahydroisoquinoline". In: *Chemical Research in Toxicology* 12.11 (Nov. 1, 1999), pp. 1090–1097. doi: 10.1021/ tx990079p.
46. Marek J. Noga et al. "Metabolomics of Cerebrospinal Fluid Reveals Changes in the Central Nervous System Metabolism in a Rat Model of Multiple Sclerosis". In: *Metabolomics* 8.2 (Apr. 2011), pp. 253–263. doi: 10.1007/s11306-011-0306-3.
 47. Alberto Noronha et al. "The Virtual Metabolic Human Database: A Comprehensive Metabolic Resource of Human and Human Associated Microbes." In: *In preparation* (2016).
 48. William T. Norton et al. "The Lipid Composition of Isolated Brain Cells and Axons". In: *Journal of Neuroscience Research* 1.1 (Jan. 1, 1975), pp. 57–75. doi: 10.1002/jnr.490010106.
 49. John S. O'Brien and E. Lois Sampson. "Lipid Composition of the Normal Human Brain: Gray Matter, White Matter, and Myelin". In: *Journal of Lipid Research* 6.4 (Jan. 10, 1965), pp. 537– 544. pmid: 5865382. url: <http://w/www.jlr.org/content/6/4/537> (visited on 08/12/2014).
 50. G. Paglia et al. "Monitoring Metabolites Consumption and Secretion in Cultured Cells Using Ultra-Performance Liquid Chromatography Quadrupole-Time of Flight Mass Spectrometry (UPLC-Q-ToF-MS)." In: *Analytical and Bioanalytical Chemistry* 402.3 (2012), pp. 1183–98. doi: 10.1007/s00216-011-5556-4. pmid: 22159369.
 51. Leslie B. Poole. "The Basics of Thiols and Cysteines in Redox Biology and Chemistry". In: *Free Radical Biology & Medicine* 80 (Mar. 2015), pp. 148–157. doi: 10.1016/j.freeradbiomed. 2014.11.013. pmid: 25433365.
 52. German A. Preciat Gonzalez et al. "Comparative Evaluation of Atom Mapping Algorithms for Balanced Metabolic Reactions: Application to Recon3D". In: *Journal of Cheminformatics* 9 (2017), p. 39. doi: 10.1186/s13321-017-0223-1.
 53. Syed Asad Rahman et al. "Reaction Decoder Tool (RDT): Extracting Features from Chemical Reactions". In: *Bioinformatics* 32.13 (Jan. 7, 2016), pp. 2065–2066. doi: 10.1093/bioinformatics/ btw096. pmid: 27153692.
 54. Peter Reinhardt et al. "Derivation and Expansion Using Only Small Molecules of Human Neural Progenitors for Neurodegenerative Disease Modeling". In:

- PLoS ONE* 8.3 (Mar. 22, 2013). Ed. by Marcel Daadi, e59252. doi: 10.1371/journal.pone.0059252.
55. Koen Sandra et al. "Comprehensive Blood Plasma Lipidomics by Liquid Chromatography/Quadrupole Time-of-Flight Mass Spectrometry". In: *Journal of Chromatography A* 1217.25 (June 2010), pp. 4087–4099. doi: 10.1016/j.chroma.2010.02.039. pmid: 20307888.
 56. Augustin Scalbert et al. "Mass-Spectrometry-Based Metabolomics: Limitations and Recommendations for Future Progress with Particular Focus on Nutrition Research". In: *Metabolomics* 5.4 (Dec. 2009), pp. 435–458. doi: 10.1007/s11306-009-0168-0. pmid: 20046865.
 57. Jan Schellenberger et al. "Quantitative Prediction of Cellular Metabolism with Constraint-Based Models: The COBRA Toolbox v2.0". In: *Nature Protocols* 6.9 (Sept. 2011), pp. 1290–1307. doi: 10.1038/nprot.2011.308.
 58. Juan Segura-Aguilar et al. "Protective and Toxic Roles of Dopamine in Parkinson's Disease". In: *Journal of Neurochemistry* 129.6 (June 1, 2014), pp. 898–915. doi: 10.1111/jnc.12686.
 59. Kirti Sharma et al. "Cell Type- and Brain Region-Resolved Mouse Brain Proteome". In: *Nature Neuroscience* advance online publication (Nov. 2, 2015). doi: 10.1038/nn.4160.
 60. David Siegel et al. "Integrated Quantification and Identification of Aldehydes and Ketones in Biological Samples". In: *Analytical Chemistry* 86.10 (May 2014), pp. 5089–5100. doi: 10.1021/ac500810r. pmid: 24745975.
 61. Katrin Strassburger et al. "Quantitative Profiling of Oxylipins through Comprehensive LC-MS/MS Analysis: Application in Cardiac Surgery". In: *Analytical and Bioanalytical Chemistry* 404.5 (2012), pp. 1413–1426. doi: 10.1007/s00216-012-6226-x. pmid: 22814969.
 62. Lloyd W. Sumner et al. "Proposed Minimum Reporting Standards for Chemical Analysis". In: *Metabolomics* 3.3 (2007), pp. 211–221. url: <http://link.springer.com/article/10.1007/s11306-007-0082-2> (visited on 03/19/2015).
 63. Ruben T'Kindt et al. "Profiling over 1500 Lipids in Induced Lung Sputum and the Implications in Studying Lung Diseases". In: *Analytical Chemistry* 87.9 (May 2015), pp. 4957–4964. doi: 10.1021/acs.analchem.5b00732. pmid: 25884268.

64. Ines Thiele and Bernhard Ø Palsson. "A Protocol for Generating a High-Quality Genome-Scale Metabolic Reconstruction". In: *Nature Protocols* 5.1 (Jan. 2010), pp. 93–121. doi: 10.1038/nprot.2009.203.
65. Ines Thiele et al. "A Community-Driven Global Reconstruction of Human Metabolism". In: *Nature Biotechnology* 31.5 (May 2013), pp. 419–425. doi: 10.1038/nbt.2488.
66. Ines Thiele et al. "Candidate Metabolic Network States in Human Mitochondria. Impact of Diabetes, Ischemia, and Diet." In: *Journal of Biological Chemistry* 280.12 (Mar. 25, 2005), pp. 11683–11695. doi: 10.1074/jbc.M409072200.
67. Cole Trapnell et al. "Differential Gene and Transcript Expression Analysis of RNA-Seq Experiments with TopHat and Cufflinks". In: *Nature Protocols* 7.3 (Mar. 1, 2012), pp. 562–578. doi: 10.1038/nprot.2012.016. pmid: 22383036.
68. Hiroshi Tsugawa et al. "Highly Sensitive and Selective Analysis of Widely Targeted Metabolomics Using Gas Chromatography/Triple-Quadrupole Mass Spectrometry". In: *Journal of Bioscience and Bioengineering* 117.1 (Jan. 2014), pp. 122–128. doi: 10.1016/j.jbiosc.2013.06.009. pmid: 23867096.
69. Monica Villa et al. "One-Electron Reduction of 6-Hydroxydopamine Quinone Is Essential in 6-Hydroxydopamine Neurotoxicity". In: *Neurotoxicity Research* 24.1 (Feb. 6, 2013), pp. 94–101. doi: 10.1007/s12640-013-9382-7.
70. Christina Virgiliou et al. "Development and Validation of a HILIC- MS/MS Multi-Targeted Method for Metabolomics Applications." In: *Electrophoresis* 36.18 (July 2015), pp. 2215–2225. doi: 10.1002/elps.201500208. pmid: 26180020.
71. Nikos Vlassis, Maria Pires Pacheco, and Thomas Sauter. "Fast Reconstruction of Compact Context-Specific Metabolic Network Models". In: *PLoS Comput Biol* 10.1 (Jan. 16, 2014), e1003424. doi: 10.1371/journal.pcbi.1003424.
72. Sharon J. Wiback et al. "Monte Carlo Sampling Can Be Used to Determine the Size and Shape of the Steady-State Flux Space". In: *Journal of Theoretical Biology* 228.4 (June 2004), pp. 437–447. doi: 10.1016/j.jtbi.2004.02.006.
73. Jenny Marie T. Wong et al. "Benzoyl Chloride Derivatization with Liquid Chromatography/MS Mass Spectrometry for Targeted Metabolomics of Neurochemicals in Biological Samples". In: *Journal of Chromatography A*

- 1446 (2016), pp. 78–90. doi: 10.1016/j.chroma.2016.04.006. pmid: 27083258.
74. Takayuki Yamada et al. “Development of a Lipid Profiling System Using Reverse-Phase Liquid Chromatography Coupled to High-Resolution Mass Spectrometry with Rapid Polarity Switching and an Automated Lipid Identification Software”. In: *Journal of Chromatography A* 1292 (May 1537 2013), pp. 211–218. doi: 10.1016/j.chroma.2013.01.078. pmid: 23411146.
75. Min Yuan et al. “A Positive/Negative Ion-Switching, Targeted Mass Spectrometry-Based Meta1539 bolomics Platform for Bodily Fluids, Cells, and Fresh and Fixed Tissue”. In: *Nature Protocols* 1540 7.5 (Apr. 2012), pp. 872–881. doi: 10.1038/nprot.2012.024. pmid: 22498707.
76. L Zecca et al. “Substantia Nigra Neuromelanin: Structure, Synthesis, and Molecular Behaviour”. 154₂ In: *Molecular Pathology* 54.6 (Dec. 2001), pp. 414–418. pmid: 11724917. url: <http://w/www.ncbi.nlm.nih.gov/pmc/articles/PMC1187132/> (visited on 11/04/2015).
77. Jiangjiang Zhu et al. “Colorectal Cancer Detection Using Targeted Serum Metabolic Profiling”. 1545 In: *Journal of Proteome Research* 13.9 (Sept. 2014), pp. 4120–4130. doi: 10.1021/pr500494u.

Chapter 6

Concluding discussion and future prospects

Conclusions

Available therapeutic strategies that target the causal factors associated with neurodegenerative diseases are scarce. Most medications offer only symptomatic relief to the patient, such as levodopa that helps to reduce the motor symptoms of Parkinson's disease¹, donepezil that reduces the memory and learning deficits in Alzheimer's disease², and antidepressants to tackle depression in Huntington's disease³. Unfortunately, the majority of these medications have a temporary effect and, in addition, may not be effective in all patients. This is due to the multifactorial complexity of these disease mechanisms which are challenging to unravel and understand. In **chapter 1**, we highlight the prospects of applying systems biology to explore and further our understanding of specific neurodegenerative diseases, with a focus on Parkinson's disease. By combining this holistic approach with omics data, we are able to integrate valuable information for a wide variety of biological matrices. We chose metabolomics as our main omics technique due to the ability to take a snapshot of the dynamic metabolic processes in different disease states, thus promoting the opportunity to understand the disease from an alternative perspective, i.e. by specifically identifying the biological phenotype. Metabolomics has demonstrated its usefulness in biomarkers in a range of conditions ranging from cancer⁴ to neurological disorders^{5,6}. However, there is still a relatively low number of methods providing absolute quantitative data that broadly cover the metabolism. Additionally, current techniques that are available are limited by their sensitivity, speed, experimental equipment requirements and/or cost.

In this thesis, we develop and utilise new analytical strategies to capture the metabolome in a broad and absolute quantitative manner by modifying the conditions of a derivatization reagent to derivatise amine, thiol and carboxyl metabolites. Consequently, we address the several limitations experienced in the quantitative analytical workflow. With the established methods, we chose to concentrate on the central carbon and energy metabolism along with neurochemicals as these are strongly implicated in neurodegenerative diseases such as Parkinson's disease. To showcase the broadness in application, we applied the methods to several biological models, including human urine, *in vitro* cancer cell line cultures (SUIT-2 cells and HepG2) and induced pluripotent stem cell (iPSC)-derived dopaminergic neurons. Also, we used a derivatization reagent on rodent brain tissue

with hope to determine the distinct neurochemical profiles in several regions across the healthy rodent brains, thus providing information to further decipher the connectome. Furthermore, we aimed to demonstrate the application of absolute quantitative metabolomics data by the integration into a genome-scale constraint-based model that captures the functionality of dopaminergic neurons, specifically the midbrain substantia nigra dopaminergic neurons associated with Parkinson's disease.

Sensitive absolute quantitative method development

One of the key goals in the metabolomics community is to establish an alternative technique to study the human metabolome in an absolute quantitative manner in response to the common limitations that are experienced with the current approaches. This is evident when using cell culture and brain samples as the biomass and biofluids can present in low quantities. This setback can be solved by the application of mass spectrometry which has superior sensitivity at the compromise of resulting absolute concentration accuracy. The quantitative inaccuracy in MS is caused by the matrix effect in the electrospray ionisation source. The most common solution is the use of stable isotope-labelled analyte pairs that are analysed simultaneously during separation and ionisation. This technique is expensive and depends on isotope availability. Also, separation sciences coupled to MS, such as LC-MS, GC-MS and CE-MS, have other additional limitations during analysis as detailed below. One approach to solve this issue is the use of isotope-coded derivatization that has the ability to modify the physicochemical properties of metabolites to encourage improved separation and ionisation features whilst providing an identical isotope pair for each analyte of interest. This allows for absolute quantitation analysis in a cost-effective manner. With this in mind, these quantities can be integrated into systems biology models, thus progressing the successful and comprehensive modelling of a variety of biological matrices.

Chapter 2 illustrates the development and validation of a pre-column derivatization ultra-high-performance liquid chromatography mass spectrometry (UHPLC-MS) analytical method⁷ with a 10-minute acquisition time, using only positive ionization mode. In this method, we expand the reactivity of the reagent dimethylaminophenacyl bromide (DmPABr) by altering the reaction conditions previously published by Guo *et al.* (2010)⁸. The change in reaction conditions

Chapter 6

resulted in the ability of DmPABr to label primary amines, secondary amines, thiols and carboxyls, compared to the original labelling coverage which was exclusive to carboxyls. This extension vastly encourages higher coverage of the human metabolome.

The method was employed to analyse healthy human urine and rotenone-treated (at 1 nM, 10 nM and 100 nM for 3 h, 8 h and 24 h) pancreatic cancer cells (SUIT-2), yielding 64 metabolites associated with central carbon and energy-related metabolism. These include: amino acids, creatinine, *N*-acetylated amino acids, metabolites from the TCA cycle and pyruvate metabolism, acylcarnitines and medium-/long-chain fatty acids. Rotenone blocks the complex I of the electron transport chain in the mitochondria - virtually depleting function. After exposure to 100 nM rotenone, 50% of the metabolites showed significant changes. This demonstrates the ability of the method to assess the health of the mitochondria within cells. Additionally, a total of 57 metabolites were detected and quantified in the urine samples, with low intra-day and inter-day variability in the amino acids (within recommended ICH guidelines). Furthermore, creatinine was included in the method to enable in method normalisation of metabolite concentrations of urine.

This method also addresses the weaknesses associated with other commonly used quantitative analytical techniques such as HILIC-MS, GC-MS and CE-MS. These weaknesses include insufficient coverage requiring combination of multiple methods, lack of sensitivity and poor metabolite stability. Another issue is the lack of internal standard availability (heavy isotope metabolite pairs). Therefore, to enhance quantitation, isotope-coded derivatisation (ICD) was also applied using standards derivatised with an isotopically-labelled reagent (DmPABr-D₆). The presented work showcases the versatility and potential of utilising DmPABr for future metabolomics studies in a range of biological matrices. Our novel method unveiled its ability to cover a larger proportion of the metabolome in a fast, sensitive and absolute quantitative manner.

Utilising DmPABr provides a versatile method that can be further extended to other metabolites that contain the previously mentioned functional groups. The method is adaptable with minimal additional work required, creating a suitable metabolomics approach for systems biology integration. This is required due to the speed of new metabolic pathway predictions identified by computational approaches.

Concluding discussion and future prospects

Additionally, the method is suitable for the study of a range of diseases associated with energy imbalance and mitochondrial dysfunction such as Leigh's syndrome⁵ and diseases with deficiency of aminoacylase I⁹. Furthermore, we also believe that computational approaches can be created to predict the labelling and retention of metabolites, thus allowing a high-throughput analytical technique with broader coverage. The method has the potential for sensitive analysis of volume-limited samples, and this is discussed in **chapter 3**.

Chapter 3 follows on from **chapter 2**, by focusing on the lack of accurate absolute quantitation in low-volume samples experienced by the current available methods for analysis. We mention in **chapter 2** that the method is not fully optimised for volume-limited samples due to an issue with detector saturation (caused by the high concentration of metabolites in urine). In response to this, we optimised the electrospray ionisation and mass spectrometry parameters. After this alteration, we were able to validate and showcase the DmPABr derivatisation technique on the application to material-limited HepG2 cell samples (ranging from 250 cells to 1×10^5 cells) via RPLC-MS/MS¹⁰. A total of 37 metabolites were detected and quantified from 1×10^4 HepG2 cells within 7-minute elution, including: amino acids, *N*-acetylated amino acids, acylcarnitines, fatty acids and TCA cycle metabolites. Most of the amino acids had a limit of detection below 20 nM, and for the *N*-acetylated amino acids and acylcarnitines, below 5 nM. The intraday variability was within the ICH guidelines for the majority of concentrations detected in 5×10^3 HepG2 cells, and the quantification of twelve metabolites and the detection of three additional metabolites below LLOQ was achieved in 250 HepG2 cells.

As mentioned earlier, cell cultures (particularly microfluidic cell culture) and brain samples often provide relatively low volumes of sample that require highly sensitive methods to deliver accurate absolute quantitative metabolite concentrations. Although methods such as LC-MS have decent sensitivity, there is inaccuracy in the ability to produce absolute quantitative results. Chemical derivatisation is an attractive choice to not only further improve the sensitivity, but also enhance the detection of metabolites in samples at low volumes.

This proof-of-concept revealed further attainable applications for the DmPABr derivatisation technique in the form of sensitive analysis of material-limited biological samples whilst maintaining the ability to create a representative profile of

the metabolome. Moreover, we utilised the DmPABr technique on UPLC-MS/MS without modification of the sample preparation volume, solvent composition and injection volume. This resulted in very small numbers of cells on column. However, we still were able to quantify 12 metabolites from the equivalent of 0.25 HepG2 cells on column. This highlights the potential for further sensitivity when utilising this approach, which could be used for single-cell metabolomics¹¹. This may be possible by optimisation of solvent composition and coupling to CE-MS and micro/nanoLC-MS with prospects of optimisation to study various diseases on a smaller scale. Furthermore, techniques such as CE-MS often suffer in the separation of anionic metabolites¹²; DmPABr could also aid the separation by the introduction of the tertiary amine, allowing cationic separation.

Data acquisition and model integration

Systems biology has the potential to advance our understanding of human physiology and complex diseases, and identify possible therapeutic targets. This is particularly the case for genome-scale constraint-based metabolic models. The main advantage is the holistic fashion in which biological information can be linked together in an interpretable manner. This is achieved by connecting information obtained from the omics field, i.e., genomics, transcriptomics, proteomics and metabolomics. Constraint-based modelling has a strength in being able to capture the dynamic biological system by including not only the genome but the end stage phenotype of the functionality or disease. It produces this by limiting the bounds of the metabolism using quantitative metabolomics information and taking understanding further from the potential of the system (genotype) to showing the functionality of the system (phenotype)¹³. However, capturing the metabolic bounds requires the use of absolute quantitative concentrations as the use of relative quantitative information alone no longer suffices. By using the quantitative methods developed and discussed in **chapter 2** and **chapter 3**, we aimed to apply these methodologies to provide the international scientific community with a quantitative neurochemical profile of the mammalian brain *ex vivo* in **chapter 4** and demonstrate the integration of the metabolic concentrations into a genome-scale constraint-based metabolic model in **chapter 5**.

In **chapter 4**, we present a comprehensive metabolic atlas of the mammalian brain. Twenty-five regions in the brains of healthy adult male Wistar rodents were analysed

using a 20-minute sensitive neurochemical stabilisation derivatisation LC-MS/MS method following Bligh and Dyer liquid-liquid extraction and benzoyl chloride derivatisation. The brain regions investigated included: the orbitofrontal cortex, cerebral cortex, frontal lobe, ventromedial prefrontal cortex, subcortical structure and brain stem. Our findings provided us with a comprehensive profile of 43 neurochemical metabolites and highlighted the brain regions that are associated with key metabolic pathways such as the mesolimbic, limbic and nigrostriatal pathways.

Benzoyl chloride is one of the gold standard derivatization reagents for the analysis of neurochemicals within the metabolomics community. It was previously developed and validated by Song *et al.* (2012)¹⁴ and Wong *et al.* (2016)¹⁵ on a range of matrices including serum, microdialysate and tissue. We developed our method on an AbSciex QTrap 6500, which provided greater sensitivity. Then, we independently validated this method within our lab by following the ICH guidelines. This provided us with a trustworthy method supported by two independent institutes. Additionally, benzoyl chloride was chosen in preference of DmPABr (as used in **chapter 2** and **chapter 3**) because of its soft labelling conditions. This is critical in the analysis of neurochemicals such as catecholamines because they are very vulnerable to degradation outside of the cell¹⁶. Thus, this improves the quantitative reliability. Prior to the derivatization, we also had to develop and validate the liquid-liquid extraction of the neurochemicals from the brain tissue. After analysis of the brain samples across ten batches, all metabolites had an analytical RSD below 20% (except epinephrine and homoserine). Moreover, all metabolites passed the ICH guidelines linearity assessment showing that absolute quantitative analysis was possible.

When presenting healthy adults rat control data, developmental factors are important. We ensured that the samples were time-independent by measuring the rats at two time-points (17 weeks and 19.5 weeks). After this, we investigated the metabolic profile across the 25 brain regions, attempting to identify metabolic similarities and differences. The brain regions exist in a connected lattice but each region has its own distinct genome, transcriptome, proteome and metabolome, leading to the idea of the connectome¹⁷. With this comes a variation in the composition of cells such as neurons and glial cells. We understand that specific neurons such as cholinergic, dopaminergic and serotonergic are expressed

Chapter 6

differently across the brain. In this study, we wanted to map and correlate the neurochemical profile, including neurotransmitters to the specific regions associated. In a review by Ivanisevic *et al.* (2015)¹³, they highlighted the need for more metabolomics data to improve our understanding of the brain. We believe this potentially provides the scientific community with additional knowledge relating to the connectome.

To improve understanding of the mammalian brain and contribute to the connectome and study of diseases, the brain regions and metabolites included in the study need to be relevant. Within this method, we investigated brain regions that were associated with a range of neurological diseases such as Parkinson's, Alzheimer's and Huntington's disease, and psychiatric disorders such as anxiety, addictive behaviours and PTSD. Examples of these regions include the olfactory bulb, nucleus accumbens, globus pallidus, bed of the stria terminalis, substantia nigra and raphe. The metabolites covered include core metabolites such as amino acids as well as specific pathways associated with neurological illnesses. Pathways such as the tyrosine metabolite, urea cycle and polyamine metabolism were investigated. We identified significant differences in the pathways such as the tyrosine metabolism; this was seen mainly in the ratio of dopamine to epinephrine. This difference was seen with the regions that express a high density of dopaminergic neurons in comparison to adrenergic neurons. The data collected shows the turnover of neurotransmitters such as serotonin to 5-hydroxyindoleacetic acid and dopamine to homovanillic acid, DOPAC and 3-methoxytyramine. This demonstrates the extensiveness of the data presented within **chapter 4**.

With the sharing of absolute quantitative neurochemical concentrations, we have captured the metabolome of the mammalian brain. This information can potentially be integrated with genome-scale constraint-based models. These findings encourage a deeper understanding of the role of the metabolome on brain function and connectivity, and create a solid foundation upon which future brain studies can build. The information can also be used as a biological validation for analytical and cell biology quality. For example, during differentiation of iPSCs into midbrain neurons, the neurons are usually identified by their genetic markers¹⁸, however, we clearly see within this chapter that most neurons are present within all regions. Using this data, we can also assess the metabolic signature to potentially allocate the area

representation more clearly. Additionally, we can assess the quality of the iPSC-derived midbrain neurons to decipher whether they mirror the metabolic functionality.

Chapter 5 presents *iNESC2DN*, a validated constraint-based metabolic model in human dopaminergic neurons created from integrating quantitative omics data with generic metabolic model *Recon3D*¹⁹ using iPSC-derived, human neuroepithelial stem cells (hNESC) differentiated into dopaminergic neurons. These neurons represent nigrostriatal dopaminergic neurons that can be used to improve the understanding of Parkinson's disease. *Recon3D* provided a foundation upon which *iNESC2DN* could be built and refined to include the active/inactive pathways and reactions involving genetic and metabolic product fluxes. We applied COntstraint-Based Reconstruction and Analysis (COBRA)^{20,21} which computationally models the integration of known biochemical data and new experimental data with the ability of generating new hypotheses. This approach is achieved mathematically and mechanistically. We hypothesised that these neurons have a genetic predisposition that makes them be vulnerable to energy imbalances, i.e., mitochondrial dysfunction. In addition, Parkinson's disease has known mitochondrial genetic mutations that make an individual predisposed to developing the condition. For this reason, energy-related and neurochemical metabolites were isolated and quantified, using LC-MS and GC-MS methods, from hNESC differentiated to midbrain-specific dopaminergic neurons.

Manual curation of metabolic literature was performed using an established protocol from *Recon2* and included in an update to *Recon3D*, which provides information about gene-protein-reaction associations. In addition, further manual curation enabled us to identify active/inactive reactions and genes, transport reactions, degradation pathways and quantitative constraints. Transcriptomics data were obtained via RNA-sequencing, with 1,202 genes mapped to metabolic genes in *Recon3D*. Metabolomics data were generated using four partially overlapping platforms from fresh and spent culture media, with 49 metabolites passing the limit of detection and integrated into *iNESC2DN*. We used AccQ-Tag derivatization (RPLC-MS) and GC-MS to quantify central carbon and energy-related metabolites, and neurochemicals. In addition, published biochemical literature (bibliomics) was manually curated to enrich the *iNESC2DN* model and validate our data findings.

Chapter 6

Evidence of activity/inactivity in 252 metabolic genes and 445 metabolic reactions were highlighted in dopaminergic neurons.

Our resulting *iNESC2DN* model offers the first, functional, genome-scale, context-specific constraint-based reconstruction of human dopaminergic neuronal metabolism. Future applications of this model include the ability to quantitatively predict the rate and route of metabolite movement in various neurodegenerative disease conditions, and the design of exometabolomic and tracer-based metabolomics experiments.

Future prospective

Method expansion and sensitivity enhancement

The metabolome contains a vast amount of biochemicals in its repertoire that may be the key to understanding disease. As discussed throughout this thesis, these metabolites vary in their physicochemical properties, providing analytical challenges for the metabolomics community. However, we have developed and validated a new derivatization technique with dimethylaminophenacyl bromide (DmPABr) that labels the amine, thiol and carboxyl metabolites which constitutes over 90% of the human metabolome. This technique provides the capacity to capture a significant proportion of the metabolome as well as the ability to study the disease in an absolute quantitative manner. The main strength of the reagent DmPABr is the flexibility to label a broad range of functional groups and introduce the isotope-coded derivatization approach, thus improving coverage and quantitation. With this, we envision that the application of the methods created in **chapter 2** and **chapter 3** can be used to broaden the quantitative coverage and detect metabolites that were previously undetectable. In addition, we hope that the sensitivity gain can also inject life into outdated mass spectrometers that previously suffered from ion suppression or low sensitivity.

During our research of neurodegenerative diseases within **chapter 4** and **chapter 5**, we highlighted key pathways that can aid the study of the diseases, with the tryptophan metabolism as the main pathway. Tryptophan metabolism, and the serotonin and kynurenine pathways²² have been associated with inflammation²³ and oxidative stress in illnesses such as alcohol use disorder (AUD)²⁴⁻²⁶, Parkinson's disease²⁷, Alzheimer's disease²⁸, Huntington's disease²⁹ and schizophrenia^{30,31}.

Unfortunately, there are very few metabolomics methods that capture tryptophan metabolism in detail in an absolute quantitative fashion. The tryptophan metabolism and related pathways have a diverse range of physicochemical properties that can provide a range of challenges, such as the separation of isomers nicotinic acid and picolinic acid. Additional challenges include the instability, structural variation and sensitivity requirements. This pathway explicitly demonstrates the potential utilisation of DmPABr to improve current approaches, provide biomarker discovery and enable absolute quantitative data integration into constraint-based metabolic models.

The above methodologies can be used to broaden the coverage of known pathways but can also be utilised in the study of single-cell metabolomics¹¹. Single-cell metabolomics is a growing field of interest that can potentially aid cancer diagnosis, the study of aging and the development of drug resistance³². Additionally, this can improve systems biology models by allowing the construction of models based on a single cell line with its distinct phenotype. The single-cell approach also allows the observation of the cell in a dynamic fashion as it changes through the aging and maturation processes; this is likely to be crucial in the understanding of slow-onset neurodegenerative diseases. To achieve this, supplementary method expansions can be made, including the coupling of the derivatization approaches to sensitive analytical equipment such as sheathless CE-MS (with stacking)³³, nanoLC-MS/MS³⁴⁻³⁶ and nanoESI³⁵. Another approach we expect is the use of quaternary amine-containing derivatization groups that not only reduce ion suppression but provide a more sensitive analysis approach^{37,38}.

Systems biology and metabolomics

The future of systems biology has the potential to improve disease understanding and provide personalised therapeutic suggestions. However, to achieve this, systems biology will need to transition past the evaluation of diseases using single compartment models and focus on the connectivity between specific regions and organs as we try to understand the whole organism, as discussed by Thiele et al. (2020)³⁹. Therefore, several steps need to occur such as the development of multi-organ cell culture devices, creation of new metabolomics assays with sensitivity and global metabolic models. However, one major limitation is the dependency on the reporting of information from omics communities. Experimental data that is difficult

Chapter 6

to reproduce or inaccuracies in reported data potentially can misdirect models. Additionally, several experimental factors potentially lead to bias in data that realistically represents the human metabolites, i.e. culturing cells in an artificial environment, variability in co-culture cell line expression, inaccuracy in brain region cell line association and cell life cycle stage.

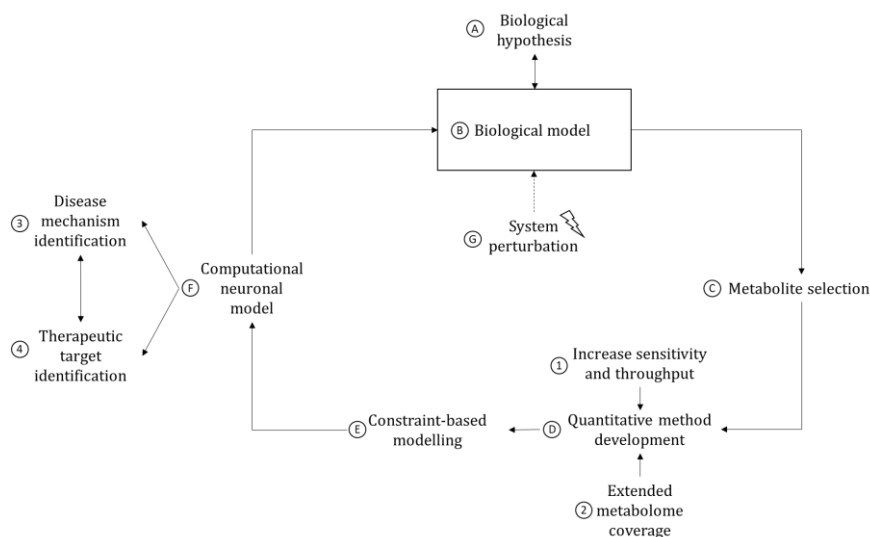


Figure 6.1. A schematic workflow of the systems biology approach with the use of metabolically constraint-based modelling. This workflow is used throughout this thesis and the future prospects are labelled 1-4.

The generation of organ or multi-organ metabolic models will improve our understanding of complex neurological disorders. For example, Parkinson's disease is not only associated with changes in the substantia nigra, but it also has been shown to exhibit changes within several brain regions, for example, the orbitofrontal cortex⁴⁰, caudate putamen, globus pallidus⁴¹, subthalamic nucleus⁴², thalamus⁴³, ventral tegmental area⁴⁴, locus coeruleus⁴⁵ and raphe⁴⁶. Once genome-scale metabolic models are extended to broader regions and organs, we can understand what causes the vulnerability of substantia nigra dopaminergic neurons compared with other neurons. Furthermore, after construction of these models, we can delve further into the study of mitochondrial genetic mutations that are associated with Parkinson's disease, i.e. PINK1 and LRRK2. This approach would also benefit from

Concluding discussion and future prospects

the regional/organ comparison as the mutations exists within all cell lines. However, the substantia nigra dopaminergic neurons seem to be the only known cells that significantly suffer from this genetic vulnerability.

Once these models are established, it offers the possibility not only to explore the causal factors associated with disease but also provides the opportunity for therapeutic target identification. By offering a complex yet comprehensive atlas of disease function from genotype to phenotype, this could emphasise the key pathways that may alter the disease symptoms or slow the progression of neurodegeneration. Using techniques such as induced pluripotent stem cell-derived neurons to evaluate the effectiveness of therapies could be possible. Using the systems biology approaches will allow us to predict the future of therapeutic responses in a dynamic and deeper fashion. Once these milestones are achieved, we will be one step closer to personalised medicines.

References

1. LeWitt PA. Levodopa for the treatment of Parkinson's disease. *New England Journal of Medicine*. 2008;359(23):2468-2476.
2. Burns A, Rossor M, Hecker J, et al. The Effects of Donepezil in Alzheimer's Disease—Results from a Multinational Trial1. *Dementia and geriatric cognitive disorders*. 1999;10(3):237-244.
3. Moulton CD, Hopkins C, Bevan-Jones WR. Systematic review of pharmacological treatments for depressive symptoms in Huntington's disease. *Movement Disorders*. 2014;29(12):1556-1561.
4. Karu N, Wilson R, Hamede R, et al. Discovery of Biomarkers for Tasmanian Devil Cancer (DFTD) by Metabolic Profiling of Serum. *J Proteome Res*. 2016;15(10):3827-3840.
5. Thompson Legault J, Strittmatter L, Tardif J, et al. A Metabolic Signature of Mitochondrial Dysfunction Revealed through a Monogenic Form of Leigh Syndrome. *Cell Rep*. 2015;13(5):981-989.
6. Lewitt PA, Li J, Lu M, Beach TG, Adler CH, Guo L. 3-hydroxykynurenine and other Parkinson's disease biomarkers discovered by metabolomic analysis. *Mov Disord*. 2013;28(12):1653-1660.
7. Willacey CCW, Naaktgeboren M, Lucumi Moreno E, et al. LC-MS/MS analysis of the central energy and carbon metabolites in biological samples following derivatization by dimethylaminophenacyl bromide. *Journal of Chromatography A*. 2019;460413.
8. Guo K, Li L. High-Performance Isotope Labeling for Profiling Carboxylic Acid-Containing Metabolites in Biofluids by Mass Spectrometry. *Analytical Chemistry*. 2010;82(21):8789-8793.
9. Gerlo E, Van Coster R, Lissens W, Winckelmans G, De Meirleir L, Wevers R. Gas chromatographic-mass spectrometric analysis of N-acetylated amino acids: the first case of aminoacylase I deficiency. *Anal Chim Acta*. 2006;571(2):191-199.
10. Willacey CC, Karu N, Harms AC, Hankemeier T. Metabolic profiling of material-limited cell samples by dimethylaminophenacyl bromide derivatization with UPLC-MS/MS analysis. *Microchemical Journal*. 2020:105445.

11. Ali A, Abouleila Y, Shimizu Y, et al. Single-cell metabolomics by mass spectrometry: Advances, challenges, and future applications. *TrAC Trends in Analytical Chemistry*. 2019;120:115436.
12. van Mever M, Hankemeier T, Ramautar R. CE-MS for anionic metabolic profiling: An overview of methodological developments. *Electrophoresis*. 2019;40(18-19):2349-2359.
13. Ivanisevic J, Siuzdak G. The role of metabolomics in brain metabolism research. *Journal of Neuroimmune Pharmacology*. 2015;10(3):391-395.
14. Song P, Mabrouk OS, Hershey ND, Kennedy RT. In vivo neurochemical monitoring using benzoyl chloride derivatization and liquid chromatography-mass spectrometry. *Anal Chem*. 2012;84(1):412-419.
15. Wong JM, Malec PA, Mabrouk OS, Ro J, Dus M, Kennedy RT. Benzoyl chloride derivatization with liquid chromatography-mass spectrometry for targeted metabolomics of neurochemicals in biological samples. *J Chromatogr A*. 2016;1446:78-90.
16. Meiser J, Weindl D, Hiller K. Complexity of dopamine metabolism. *Cell Communication and Signaling : CCS*. 2013;11:34-34.
17. Sporns O. The human connectome: a complex network. *Annals of the New York Academy of Sciences*. 2011;1224(1):109-125.
18. Moreno EL, Hachi S, Hemmer K, et al. Differentiation of neuroepithelial stem cells into functional dopaminergic neurons in 3D microfluidic cell culture. *Lab Chip*. 2015;15(11):2419-2428.
19. Brunk E, Sahoo S, Zielinski DC, et al. Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nature biotechnology*. 2018;36(3):272.
20. Mao L, Nicolae A, Oliveira MA, He F, Hachi S, Fleming RM. A constraint-based modelling approach to metabolic dysfunction in Parkinson's disease. *Computational and structural biotechnology journal*. 2015;13:484-491.
21. Bordbar A, Monk JM, King ZA, Palsson BO. Constraint-based models predict metabolic and associated cellular functions. *Nature Reviews Genetics*. 2014;15(2):107-120.
22. Stone TW, Darlington LG. The kynurenine pathway as a therapeutic target in cognitive and neurodegenerative disorders. *British journal of pharmacology*. 2013;169(6):1211-1227.

23. Karu N, McKercher C, Nichols DS, et al. Tryptophan metabolism, its relation to inflammation and stress markers and association with psychological and cognitive functioning: Tasmanian Chronic Kidney Disease pilot study. *BMC nephrology*. 2016;17(1):171.
24. Neupane SP, Lien L, Martinez P, Hestad K, Bramness JG. The relationship of alcohol use disorders and depressive symptoms to tryptophan metabolism: cross-sectional data from a Nepalese alcohol treatment sample. *Alcoholism, clinical and experimental research*. 2015;39(3):514-521.
25. Dougherty DM, Mullen J, Hill-Kapturczak N, et al. Effects of tryptophan depletion and a simulated alcohol binge on impulsivity. *Experimental and clinical psychopharmacology*. 2015;23(2):109-121.
26. Underwood MD, Kassir SA, Bakalian MJ, et al. Serotonin receptors and suicide, major depression, alcohol use disorder and reported early life adversity. *Translational psychiatry*. 2018;8(1):279.
27. Havelund JF, Heegaard NHH, Faergeman NJK, Gramsbergen JB. Biomarker Research in Parkinson's Disease Using Metabolite Profiling. *Metabolites*. 2017;7(3).
28. Lovelace MD, Varney B, Sundaram G, et al. Recent evidence for an expanded role of the kynurenine pathway of tryptophan metabolism in neurological diseases. *Neuropharmacology*. 2017;112(Pt B):373-388.
29. Dehhaghi M, Kazemi Shariat Panahi H, Guillemin GJ. Microorganisms, Tryptophan Metabolism, and Kynurenine Pathway: A Complex Interconnected Loop Influencing Human Health Status. *International journal of tryptophan research : IJTR*. 2019;12:1178646919852996.
30. Chiappelli J, Postolache TT, Kochunov P, et al. Tryptophan Metabolism and White Matter Integrity in Schizophrenia. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*. 2016;41(10):2587-2595.
31. Stahl SM. Beyond the dopamine hypothesis of schizophrenia to three neural networks of psychosis: dopamine, serotonin, and glutamate. *CNS spectrums*. 2018;23(3):187-191.
32. Zenobi R. Single-Cell Metabolomics: Analytical and Biological Perspectives. *Science*. 2013;342(6163):1243259.

33. Zhang W, Guled F, Hankemeier T, Ramautar R. Utility of sheathless capillary electrophoresis-mass spectrometry for metabolic profiling of limited sample amounts. *Journal of Chromatography B*. 2019;1105:10-14.
34. Chetwynd AJ, David A. A review of nanoscale LC-ESI for metabolomics and its potential to enhance the metabolome coverage. *Talanta*. 2018;182:380-390.
35. Chetwynd AJ, David A, Hill EM, Abdul-Sada A. Evaluation of analytical performance and reliability of direct nanoLC-nanoESI-high resolution mass spectrometry for profiling the (xeno) metabolome. *Journal of Mass Spectrometry*. 2014;49(10):1063-1069.
36. Luo X, Li L. Metabolomics of Small Numbers of Cells: Metabolomic Profiling of 100, 1000, and 10000 Human Breast Cancer Cells. *Analytical Chemistry*. 2017;89(21):11664-11671.
37. Denver N, Khan S, Stasinopoulos I, et al. Derivatization enhances analysis of estrogens and their bioactive metabolites in human plasma by liquid chromatography tandem mass spectrometry. *Analytica chimica acta*. 2019;1054:84-94.
38. Yang WC, Regnier FE, Adamec J. Comparative metabolite profiling of carboxylic acids in rat urine by CE-ESI MS/MS through positively pre-charged and 2H-coded derivatization. *Electrophoresis*. 2008;29(22):4549-4560.
39. Thiele I, Sahoo S, Heinken A, et al. Personalized whole-body models integrate metabolism, physiology, and the gut microbiome. *Molecular systems biology*. 2020;16(5):e8982-e8982.
40. Attems J, Walker L, Jellinger KA. Olfactory bulb involvement in neurodegenerative diseases. *Acta Neuropathologica*. 2014;127(4):459-475.
41. Ramirez-Zamora A, Ostrem JL. Globus Pallidus Interna or Subthalamic Nucleus Deep Brain Stimulation for Parkinson Disease: A Review. *JAMA Neurology*. 2018;75(3):367-372.
42. Bonnevie T, Zaghloul KA. The Subthalamic Nucleus: Unravelling New Roles and Mechanisms in the Control of Action. *The Neuroscientist*. 2019;25(1):48-64.
43. Gent TC, Bassetti CLA, Adamantidis AR. Sleep-wake control and the thalamus. *Curr Opin Neurobiol*. 2018;52:188-197.

Chapter 6

44. Morales M, Margolis EB. Ventral tegmental area: cellular heterogeneity, connectivity and behaviour. *Nature Reviews Neuroscience*. 2017;18(2):73.
45. Betts MJ, Kirilina E, Otaduy MCG, et al. Locus coeruleus imaging as a biomarker for noradrenergic dysfunction in neurodegenerative diseases. *Brain*. 2019;142(9):2558-2571.
46. Huang KW, Ochandarena NE, Philson AC, et al. Molecular and anatomical organization of the dorsal raphe nucleus. *Elife*. 2019;8:e46464.

Addendum

Summary

Nederlandse samenvatting

Curriculum vitae

List of publications

Dankwoord

Summary

Neurodegenerative diseases, including Parkinson's disease (PD), are increasing in prevalence due to the aging population. Despite extensive study, these diseases are still not fully understood and the lack of personalised treatment options that can target the cause of the diseases, rather than the symptoms, has led to a greater demand for improved disease understanding, therapies and diagnostic procedures. In this thesis, we use systems biology approaches to construct disease-specific models intended for biomarker discovery, therapeutic treatment strategy identification and drug repurposing in PD. Systems biology is a mathematical field of research that analyses biological systems via construction of a computational model using experimental data. This is achieved by integration of omics data, including genomics, proteomics, transcriptomics and metabolomics. A specific approach used to identify the physico- and biochemical bounds within a biological system is constraint-based modelling, which requires the input of absolute quantitative metabolomics data. To improve our absolute quantitative coverage of the metabolome, we developed and improved new quantitative metabolomics methods using a targeted MS workflow to obtain data intended to be integrated into constraint-based metabolic models for the study of PD. A subset of PD is associated with mitochondrial dysfunction for a range of genetic mutations, consistent with the strong link between mitochondrial function and PD, with many of the associated metabolites involved in the TCA cycle and energy metabolism. Therefore, metabolomics methods were used to absolutely quantify metabolites of the central carbon and energy metabolism, and specific neurochemicals. Due to the vastness of the metabolome and the difficulty in achieving a high coverage of metabolites within a biological matrix, we employed chemical derivatization to enhance the detection and quantitation of a larger proportion of metabolites from human urine, SUIT-2 cells, HepG2 cells, rat brain and induced pluripotent stem cell (iPSC)-derived dopaminergic neurons. To evaluate the realistic biological environment, the use of 3D cell culture microfluidic devices are becoming more popular. However, this leads to a reduced sample size. Material-limited samples are providing challenges for the metabolomics community due to the required sensitivity. Chemical derivatization can aid detection by increasing sensitivity and offering other benefits, such as greater selectivity and separation. There is also a lack of existing absolute quantitative metabolite reference values (with specific focus on the mammalian brain) which is

essential for integration into metabolic models. In addition, metabolite profiles change in relation to the cell type, brain region and function. For these reasons, we also pursue the absolute quantitation of neurochemicals found in rat brain, i.e., the mammalian brain. Finally, we integrate multi-omics data from the iPSC-derived dopaminergic neurons into a genome scale constraint-based reconstruction and analysis model that can be used to understand the complexities of PD.

To begin the journey to produce a comprehensive PD model, **chapter 2** focused on the issues met with achieving broad coverage and absolute quantitation of the metabolome by creating a high-throughput, reliable, single, pre-column derivatization RPLC-MS/MS analysis with a 10-minute acquisition time using only positive ionization mode. Using the chemical derivatization reagent dimethylaminophenacyl bromide (DmPABr), we were able to simultaneously label carboxylic acids, thiols and amines of which only few published methods have successfully targeted together. These groups are abundant in the metabolites of the central carbon and energy metabolism, thus enabling us to quantify the concentrations of these metabolites. To further enhance quantitation, we also applied isotope-coded derivatization (ICD) by using internal standards with an isotopically labelled reagent (DmPABr-D₆). From human urine and SUIT-2 cells, we detected and quantified 64 central carbon and energy-related metabolites, including amino acids, N-acetylated amino acids, metabolites from the TCA cycle and pyruvate metabolism, acylcarnitines and medium-/long-chain fatty acids. We demonstrated the ability of the method to identify mitochondrial dysfunction by exposure of SUIT-2 cells to 100 nM rotenone. Following treatment, 50% of the metabolites detected showed significant alterations.

Another common problem faced by scientists in metabolite detection and quantitation is sensitivity, particularly with low volumes or concentrations of samples. **Chapter 3** applies the derivatization method and ICD described in **chapter 2** to material-limited cell samples, in turn improving chromatographic separation and enhancing MS ionization. After fast and accessible derivatization with DmPABr, we applied our novel RPLC-MS/MS method to HepG2 cells, ranging from 250 cells to 1×10^5 cells. In sub-10,000 cells, we were able to detect and quantify 37 metabolites, and a further 11 metabolites were detected below LLOQ.

Addendum

In **chapter 4**, we address the current lack of available absolute quantitative mammalian brain metabolite reference values by quantifying neurochemicals across 25 regions isolated from rat brain to produce a comprehensive metabolic atlas of the mammalian brain. The brain regions associated with PD and other neurological disorders are focused on, including the orbitofrontal cortex, cerebral cortex, frontal lobe, ventromedial prefrontal cortex, subcortical structure and brain stem. We optimized LLE extraction before applying the derivatization LC-MS/MS technique. In this case, we used the benzoyl chloride derivatization reagent because, unlike DmPABr, it has the ability to stabilise vulnerable catecholamines and capture neuroactive metabolites. With our method, we were able to create a concentration profile of 43 metabolites including important neurotransmitters, such as dopamine, epinephrine, norepinephrine, GABA and serotonin, and key metabolites involved in specific pathways associated with PD, such as the urea cycle, and polyamine and tyrosine metabolism.

The integration of omics data into a model can vastly improve the understanding of the mechanism and function behind complex neurological conditions. **Chapter 5** introduces *iNESC2DN*, a validated constraint-based metabolic model in human dopaminergic neurons designed to aid the understanding of PD. We conducted multi-omics analysis on iPSC-derived, human neuroepithelial stem cells (hNESC) differentiated into midbrain-specific dopaminergic neurons and integrated the obtained data into a genome scale constraint-based reconstruction and analysis model that focused on the generic human metabolome, *Recon3D*, as a basis to generate stoichiometrically and flux consistent constraint-based model of dopaminergic neuron metabolism. AccQ-Tag derivatization (RPLC-MS) and GC-MS were applied to assess the central carbon and energy metabolism, and capture the neurochemical profile. In addition, GC-MS quantified sugars. Manual literature curation, transcriptomics and the metabolomics input were used to constrain the metabolism.

The final piece in the jigsaw of understanding disease may be in the metabolome. We developed and validated a new derivatization technique which has potential to absolutely quantify over 90% of the human metabolome. Within this thesis, we utilize targeted LC-MS/MS approaches and this limits the metabolic coverage required to truly piece together the human metabolism. From the data presented in

this thesis, with slight modifications, the methodology can potentially be expanded to untargeted derivatization and coupled to sensitive analytical equipment (i.e., nanospray ESI-MS) which would bring single-cell metabolomics within its reach.

With this thesis, we have demonstrated the applicability of systems biology to enhance the understanding of PD. We must advance from single compartment models onto multiple matrices/organs in order to form the whole picture of an organism. Successful execution of organ or multi-organ metabolic models will enhance the understanding of PD; this aligns with the disease's association with the metabolic changes that are exhibited not only in the substantia nigra, but also other brain regions. This is strongly dependent on access to accurate and reproducible omics data, and the ability to remove bias caused from data obtained from experiments that do not realistically represent the human metabolome, which include issues such as inaccuracy in brain region cell line association and cell life cycle stage, variability in co-culture cell line expression and culturing cells in an artificial environment. In our study, we focused on dopaminergic neurons in the substantia nigra. Further future prospects of these computational models include the ability to explore the causal factors of PD and other complex neurological diseases, and pinpoint potential therapeutic targets. The combination of cell models, high quality measurements, a set of reference values in animal models and computational modelling could bridge the gap in reaching personalised medicines.

Samenvatting

Neurodegeneratieve ziekten, waaronder de ziekte van Parkinson (PD), komen steeds vaker voor als gevolg van de vergrijzing van de bevolking. Ondanks uitgebreid onderzoek zijn deze ziekten nog steeds niet volledig begrepen en het gebrek aan gepersonaliseerde behandelingsopties die zich kunnen richten op de oorzaak van de ziekten, in plaats van op de symptomen, heeft geleid tot een grotere vraag naar beter ziekte-inzicht, therapieën en diagnostische procedures. In dit proefschrift gebruiken we systeembioologische benaderingen om ziektespecifieke modellen te construeren die bedoeld zijn voor de ontdekking van biomarkers, de identificatie van therapeutische behandelingsstrategieën en de herbestemming van geneesmiddelen voor PD. Systeembioologie is een wiskundig onderzoeksgebied dat biologische systemen analyseert via de constructie van een computationeel model met behulp van experimentele gegevens. Dit wordt bereikt door integratie van omics-gegevens, waaronder genomics, proteomics, transcriptomics en metabolomics. Een specifieke aanpak die wordt gebruikt om de fysisch- en biochemische grenzen binnen een biologisch systeem te bepalen is constraint-based modelling, waarvoor de input van absolute kwantitatieve metabolomics-gegevens vereist is. Om onze absolute kwantitatieve dekking van het metaboolom te verbeteren, hebben we nieuwe kwantitatieve metabolomics methoden ontwikkeld en verbeterd met behulp van een gerichte MS workflow om gegevens te verkrijgen die bedoeld zijn om te worden geïntegreerd in constraint-based metabole modellen voor de studie van PD. Een subset van PD wordt geassocieerd met mitochondriale disfunctie voor een reeks van genetische mutaties, consistent met het sterke verband tussen mitochondriale functie en PD, met veel van de geassocieerde metabolieten die betrokken zijn bij de TCA-cyclus en energiemetabolisme. Daarom werden metabolomics-methoden gebruikt om metabolieten van het centrale koolstof- en energiemetabolisme en specifieke neurochemicaliën absoluut te kwantificeren. Door de uitgestrektheid van het metaboolom en de moeilijkheid bij het bereiken van een hoge dekking van metabolieten binnen een biologische matrix, hebben we chemische derivatisering gebruikt om de detectie en kwantificering van een groter deel van de metabolieten in humane urine, SUIT-2-cellen, HepG2-cellen, rattenhersenen en geïnduceerde pluripotente stamcel (iPSC)-afgeleide dopaminerge neuronen te verbeteren. Om de realistische biologische omgeving te evalueren, wordt het gebruik van 3D-celcultuur microfluidische apparaten steeds populairder. Dit leidt echter tot een kleinere

steekproefgrootte. Materiaal-beperkte stalen vormen een uitdaging voor de metabolomics gemeenschap omwille van de vereiste gevoeligheid. Chemische derivatisering kan de detectie bevorderen door de gevoeligheid te verhogen en andere voordelen te bieden, zoals een grotere selectiviteit en scheiding. Er is ook een gebrek aan bestaande absolute kwantitatieve metabolietreferentiewaarden (met specifieke aandacht voor de hersenen van zoogdieren), wat essentieel is voor integratie in metabolische modellen. Bovendien veranderen metabolietprofielen naargelang het celtype, de hersenregio en de functie. Om deze redenen streven we ook naar de absolute kwantificering van neurochemicaliën die worden aangetroffen in rattenhersenen, d.w.z. het zoogdierenbrein. Tenslotte integreren we multi-omics data van de iPSC-afgeleide dopaminerge neuronen in een op genoomschaal constraint-based reconstructie- en analysemodel dat gebruikt kan worden om de complexiteit van PD te begrijpen.

Om de reis naar een uitgebreid PD model te beginnen, richtte **hoofdstuk 2** zich op de problemen bij het bereiken van een brede dekking en absolute kwantificering van het metaboloom door het creëren van een high-throughput, betrouwbare, enkele, pre-kolom derivatisering RPLC-MS/MS analyse met een 10-minuten acquisitietijd met behulp van alleen positieve ionisatie-modus. Met behulp van de chemische derivatisering reagens dimethylaminophenacyl bromide (DmPABr), waren we in staat om gelijktijdig label carbonzuren, thiolen en aminen waarvan slechts enkele gepubliceerde methoden met succes hebben gericht samen. Deze groepen zijn overvloedig aanwezig in de metabolieten van het centrale koolstof- en energiemetabolisme, zodat wij de concentraties van deze metabolieten konden kwantificeren. Om de kwantificering verder te verbeteren, hebben we ook isotoop-gecodeerde derivatisering (ICD) toegepast door het gebruik van interne standaarden met een isotopisch gelabeld reagens (DmPABr-D₆). In humane urine en SUIT-2 cellen, hebben we 64 centrale koolstof en energie-gerelateerde metabolieten, waaronder aminozuren, N-geacetylerde aminozuren, metabolieten uit de TCA-cyclus en pyruvaatmetabolisme, acylcarnitines en middellange/ lange-keten vetzuren, gedetecteerd en gekwantificeerd. Wij hebben het vermogen van de methode om mitochondriale disfunctie te identificeren aangetoond door blootstelling van SUIT-2 cellen aan 100 nM rotenon. Na de behandeling vertoonden 50% van de gedetecteerde metabolieten significante veranderingen.

Addendum

Een ander veelvoorkomend probleem voor wetenschappers bij metabolietdetectie en -kwantificatie is de gevoeligheid, vooral bij lage volumes of concentraties van monsters. **Hoofdstuk 3** past de derivatiseringsmethode en ICD zoals beschreven in **hoofdstuk 2** op materiaal-gelimiteerde cel monsters toe op het verbeteren van de chromatografische scheiding en het verbeteren van MS ionisatie. Na snelle en toegankelijke derivatisering met DmPABr, pasten we onze nieuwe RPLC-MS/MS-methode toe op HepG2 cellen, variërend van 250 cellen tot 1×10^5 cellen. In sub-10.000 cellen waren we in staat om 37 metabolieten te detecteren en te kwantificeren, en nog eens 11 metabolieten werden gedetecteerd onder de LLOQ.

In **hoofdstuk 4** pakken we het huidige gebrek aan beschikbare absolute kwantitatieve zoogdierhersenen metaboliet referentiewaarden aan door het kwantificeren van neurochemicaliën in 25 geïsoleerde regio's van rattenhersenen om een uitgebreide metabole atlas van de zoogdierhersenen te produceren. De nadruk ligt op de hersengebieden die in verband worden gebracht met PD en andere neurologische aandoeningen, waaronder de orbitofrontale cortex, de hersenschors, de frontale kwab, de ventromediale prefrontale cortex, de subcorticale structuur en de hersenstam. Wij hebben de LLE-extractie geoptimaliseerd alvorens de LC-MS/MS-techniek voor derivatisering toe te passen. In dit geval gebruikten wij het benzoylchloride derivatiseringsreagens omdat dit, in tegenstelling tot DmPABr, het vermogen heeft kwetsbare catecholamines te stabiliseren en neuroactieve metabolieten te vangen. Met onze methode konden wij een concentratieprofiel opstellen van 43 metabolieten, waaronder belangrijke neurotransmitters zoals dopamine, epinefrine, norepinefrine, GABA en serotonine, en belangrijke metabolieten die betrokken zijn bij specifieke routes die in verband worden gebracht met PD, zoals de ureumcyclus, en het polyamine- en tyrosinemetabolisme.

De integratie van omics gegevens in een model kan het begrip van het mechanisme en de functie achter complexe neurologische aandoeningen aanzienlijk verbeteren. **Hoofdstuk 5** introduceert *iNESC2DN*, een gevalideerd constraint-gebaseerd metabool model in menselijke dopaminerge neuronen, ontworpen om het begrip van PD te bevorderen. We hebben multi-omics analyses uitgevoerd op iPSC-afgeleide, humane neuroepitheliale stamcellen (hNESC), gedifferentieerd tot midbrain-specifieke dopaminerge neuronen en de verkregen data geïntegreerd in een genoom-schaal constraint-based reconstructie en analyse model dat zich richt op het

generieke humane metaboloom, *Recon3D*, als basis om een stoichiometrisch en flux consistent constraint-based model van dopaminerge neuron metabolisme te genereren. AccQ-Tag derivatisering (RPLC-MS) en GC-MS werden toegepast om het centrale koolstof- en energiemetabolisme te beoordelen, en het neurochemische profiel vast te leggen. Bovendien kon GC-MS ingezet worden voor de kwantificatie van suikers. Handmatige literatuur curatie, transcriptomics en de metabolomics input werden gebruikt om het metabolisme te beperken.

Het laatste stukje in de puzzel van het begrijpen van ziekten kan in het metaboloom liggen. We ontwikkelden en valideerden een nieuwe derivatiseringstechniek die de potentie heeft om meer dan 90% van het menselijk metaboloom absoluut te kwantificeren. In dit proefschrift maken we gebruik van gerichte LC-MS/MS benaderingen en dit beperkt de metabole dekking die nodig is om het menselijk metabolisme echt samen te stellen. Op basis van de gegevens gepresenteerd in dit proefschrift, kan de ontwikkelde methodologie verder worden uitgebreid voor ongerichte derivatisering en indien gekoppeld aan gevoelige analytische apparatuur (dat wil zeggen, nanospray ESI-MS) dan liggen de mogelijkheden voor single-cell metabolomics binnen handbereik.

Met dit proefschrift hebben we de toepasbaarheid aangetoond van systeembioologie om het begrip van PD te vergroten. We moeten van modellen met één compartiment overgaan op modellen met meerdere matrices/organen om het hele plaatje van een organisme te kunnen vormen. Succesvolle uitvoering van orgaan- of multi-orgaan metabolische modellen zal het begrip van PD vergroten; dit sluit aan bij de associatie van de ziekte met de metabolische veranderingen die niet alleen in de substantia nigra, maar ook in andere hersengebieden worden vertoond. Dit is sterk afhankelijk van de toegang tot nauwkeurige en reproduceerbare omics-gegevens, en de mogelijkheid om bias te verwijderen die wordt veroorzaakt door gegevens die zijn verkregen uit experimenten die geen realistische weergave zijn van het menselijk metaboloom, waaronder kwesties zoals onnauwkeurigheid in hersengebied cellijn associatie en cel levenscyclus fase, variabiliteit in co-cultuur cellijn expressie en het kweken van cellen in een kunstmatige omgeving. In onze studie hebben we ons gericht op dopaminerge neuronen in de substantia nigra. Verdere toekomstperspectieven van deze computationele modellen zijn de mogelijkheid om de oorzakelijke factoren van PD en andere complexe neurologische ziekten te

Addendum

onderzoeken, en potentiële therapeutische doelen aan te wijzen. De combinatie van celmodellen, metingen van hoge kwaliteit, een reeks referentiewaarden in diermodellen en computationele modellering zou de kloof kunnen overbruggen in de richting van gepersonaliseerde geneesmiddelen.

Curriculum Vitae

Cornelius Carmichael William Willacey, was born on 25th June 1992 in the city of Nottingham, United Kingdom. He obtained his high school education and A levels in the study of biology, chemistry, maths and geography at Nottingham Bluecoat School in 2011. Before entering university, he spent time volunteering with the Bluecoat charity, in Pietermaritzburg, KwaZulu-Natal, South Africa.

Following this, he enrolled onto a master program in pharmacy at the University of Bradford, United Kingdom, during which he completed jobs and internships alongside his studies in community practice and industry. His community experience was conducted at Averroes Pharmacy, Nottingham. His pharmaceutical industrial experience included training at Boots Contract Manufacturing, Nottingham, and AstraZeneca, Macclesfield, focusing on pharmaceutical regulations and formulation science, respectively. For his research dissertation, Cornelius worked under Dr Sriharsha Kantamneni on a project titled "*Biochemical analysis of ubiquitination effects on N-methyl-D-aspartate receptor subunit GluN3A*". After developing a passion for neurological research, Cornelius continued working in the lab of Dr Sriharsha Kantamneni and completed a summer research project titled "*Kinesin-I motor complex mediated trafficking of GABA_B receptors in neurons*".

In 2015, Cornelius was awarded his Master of Pharmacy with Distinction. To complete his pharmacy registration for UK practice, he worked as a pre-registration pharmacist at Barnsley Hospital, Barnsley, United Kingdom. In 2016, Cornelius qualified as a registered UK pharmacist with the General Pharmaceutical Council (GPhC), and worked as a locum community and hospital pharmacist.

Also in 2016, Cornelius initiated his PhD in the group of Professor dr Thomas Hankemeier. Under the supervision of his co-promotors, Dr Amy Harms and Dr Ronan Fleming, he focused on the biochemical analysis of diseases coupled to systems biology and the metabolic profiling of Parkinson's disease and alcohol addiction. He developed a successful derivatisation method to improve the detection of metabolites in biological samples which resulted in publications in Journal of Chromatography A and Microchemical Journal. In 2019, he presented his results at the annual American Society for Mass Spectrometry (ASMS) conference in Atlanta, Georgia, United States.

Addendum

After his PhD, Cornelius works for Charles River Laboratories based at their bioanalytical facility in Schaijk, The Netherlands. As a Study Director, he is responsible for project management of bioanalytical sciences working with commercial partners.

List of publications

[1] **Willacey CCW**, Naaktgeboren M, Lucumi Moreno E, et al. LC-MS/MS analysis of the central energy and carbon metabolites in biological samples following derivatization by dimethylaminophenacyl bromide.

Journal of Chromatography A. 2019:460413.

[2] **Willacey CCW**, Karu N, Harms AC, Hankemeier T. Metabolic profiling of material-limited cell samples by dimethylaminophenacyl bromide derivatization with UPLC-MS/MS analysis. *Microchemical Journal*. 2020:105445.

[3] **Willacey CCW**, Secci ME, Meinhardt MW, Schilperoord T, Kohler I, Harms AC, Karu N[†], Sommer WH[†], Hankemeier T[†]. A quantitative atlas of metabolites across regions of the rat brain.

Submitted

[†] Authors contributed equally

[4] Preciat G*, Modamio J*, **Willacey CCW***, Wegrzyn AB, Lucumi Moreno E, Monteiro FL, El Assal D, Oliveira MAP, Gulersonmez MC, Heirendt L, Zhang Z, Cousins B, Haraldsdóttir HS, Zach S, Hachi S, Harms AC, Vempala S, Hengerer B, Schwamborn JC, Glaab E, Thiele I, Hankemeier T[†], Fleming RMT[†]. Mechanistic model-driven exometabolomic characterisation of human 2 dopaminergic neuronal metabolism

In preparation

*[†] Authors contributed equally

[5] van Mever M*, **Willacey CCW***, Zhang W, Drouin N, Christina AE, van Veldhoven JPD, van der Es D, Hankemeier T[†], Ramautar R[†]. Profiling acidic metabolites by capillary electrophoresis-mass spectrometry in low numbers of mammalian cells using a novel chemical derivatization approach

Submitted

*[†] Authors contributed equally

Addendum

Dankwoord

I would like to start by saying that I am inspired by Thomas for his unwavering support and enthusiasm. You helped me to strengthen my passion for science and push myself to become a better scientist. I only hope that I can maintain a youthful energy like yourself in my approach to new challenges and hopeful successes in the future.

Amy, you kept me grounded and were always there to lend a helping hand when I needed it. Your commitment to making sure everyone achieves their dreams is something to marvel at. Ronan, your support in my scientific and personal development will always be remembered. Naama, I honestly don't think I would have made it to the end if it wasn't for you! You offered your expertise in all the right places - your brutal honesty gave me the kick up the bum and helped me to keep aiming higher. For this, I am forever grateful.

Throughout the PhD journey, I have been supported by fantastic scientists that shared their knowledge and passions in the field that I love. I would like to thank Rawi, Edinson, Alireza and Liesbeth for their motivation.

I would also like to thank my students, Maik, Rianne, Martijn, Tom, Noelle and Alissa, for their help in contributing to my PhD research and giving me the chance to work with fresh, developing minds. Supervising your work was very enjoyable for me, and helped me to further enhance my research and supervision skills.

I enjoyed the collaboration with both the consortium of SysMedPD and SybilAA. Both projects had ambitious aims and I believe that metabolomics was a fundamental part in contributing towards the end goal. I would like to thank all members that assisted in the development of my scientific skills and also for the pleasant conversations along the way.

I consider myself extremely lucky to have had the opportunity to have worked alongside (and partied with!) many amazing individuals throughout my period at Leiden University. I was a traveller through many offices during my time at Leiden – there are too many names and too many enjoyable conversations to mention!

My wholehearted special thanks to my wife, Alisa. You unconditionally supported me through all the good and bad times during my PhD. You even managed to maintain

your energy whilst carrying our beautiful daughter, Aubrey, with the additional stress of the pandemic. Not many people could handle so much!

Finally, a special thanks to Aubrey. You were a big (but tiny) distraction towards to end of my PhD, but you are the most beautiful distraction a dad could ever ask for. Hopefully, you grow up to be as critical as me and point out all of the flaws in this thesis ☺.