The handle [https://hdl.handle.net/1887/3209244](https://hdl.handle.net/1887/3209244) holds various files of this Leiden University dissertation.

**Author:** Willacey, C.C.W.

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

**Issue Date:** 2021-09-08
Chapter 6

Concluding discussion and future prospects
Chapter 6

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 disease\(^1\), donepezil that reduces the memory and learning deficits in Alzheimer’s disease\(^2\), and antidepressants to tackle depression in Huntington’s disease\(^3\). 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 cancer\(^4\) 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
Concluding discussion and future prospects

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 method 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). The change in reaction conditions
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 syndrome\(^5\) and diseases with deficiency of aminoacylase \(^1\). 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 \(1 \times 10^5\) cells) via RPLC-MS/MS\(^10\). A total of 37 metabolites were detected and quantified from \(1 \times 10^4\) 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 \(5 \times 10^3\) 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 metabolomics. 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 metabolites; 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). 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
Concluding discussion and future prospects

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)\textsuperscript{14} and Wong et al. (2016)\textsuperscript{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 cell\textsuperscript{16}. 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 connectome\textsuperscript{17}. 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
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), 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 markers, 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* 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) 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.
Evidence of activity/inactivity in 252 metabolic genes and 445 metabolic reactions were highlighted in dopaminergic neurons.

Our resulting *iNES2DN* 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 pathways have been associated with inflammation and oxidative stress in illnesses such as alcohol use disorder (AUD), Parkinson’s disease, Alzheimer’s disease, Huntington’s disease and schizophrenia.
Concluding discussion and future prospects

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 metabolomics. Single-cell metabolomics is a growing field of interest that can potentially aid cancer diagnosis, the study of aging and the development of drug resistance. 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), nanoLC-MS/MS and nanoESI. 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.

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). 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
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 cortex, caudate putamen, globus pallidus, subthalamic nucleus, thalamus, ventral tegmental area, locus coeruleus and raphe. 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
the regional/organ comparison as the mutations exist 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


Concluding discussion and future prospects


Concluding discussion and future prospects


Chapter 6


