

https://dspace.org/

The determination of early potential biomarkers for Diabetic Nephropathy in human urine samples using ultra high-performance liquid chromatography coupled to quadrupole time of flight high resolution mass spectrometer (UHPLC - QTOF HRMS)

Mbhele, Thapelo

2023

https://hdl.handle.net/10539/35721

Downloaded from a DSpace repository. https://dspace.org/



UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

The determination of early potential biomarkers for Diabetic Nephropathy in human urine samples using ultra high-performance liquid chromatography coupled to quadrupole time of flight high-resolution mass spectrometer (UHPLC–QTOF HRMS)

By

Thapelo Mbhele

1430070

Supervisors:

Dr Maya Makatini

Dr Donald Tanyanyiwa

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg in fulfilment of the requirements for the degree of Master of Science

University of the Witwatersrand, Johannesburg

Date: 24 March 2022

Declaration

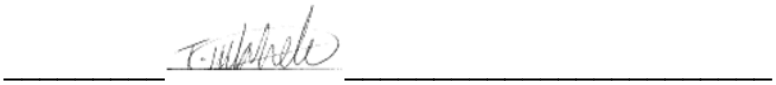
I declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any university. This work was done by me under the supervision of Dr Maya Makatini and co-supervised by Dr Donald Tanyanyiwa.



Thapelo Mbhele Date: 24 March 2022

Ethical declaration

I declare that medical practice in patient recruitment, and ethical guidelines as stipulated by the ethics committee were followed, the results obtained in this study were kept confidential, and sample codes were used to preserve anonymity. Ethical clearance was obtained from the Human and Animal Rights Chris Hani Baragwanath Academic Hospital Medical Advisory Committee and the University of the Witwatersrand Human Research Ethics Committee (reference number: M180798). The purpose of the study was communicated to the participants, and they gave written consent to allow for the collection and analysis of their urine samples and medical records. Appropriate disposal of medical samples was followed, there was no contamination to one's self or others, as well as the environment.



Thapelo Mbhele Date: 24 March 2022

Acknowledgements

First and foremost, I would like to thank my supervisor, **Dr Maya Makatini** for her professional guidance and knowledge throughout the project, assisting with both technical and scientific approach in research methodology and writing.

I would also like to thank my co-supervisor **Dr Donald Tanyanyiwa** who gave me the opportunity to work on a project that links clinical medicine with chemistry.

I give thanks to the Chris Hani Baragwanath Academic Hospital Chemical Pathology staff: **Thandeka Khumalo,** for the collection and transportation of samples.

I also give thanks to the National Institute of Occupational Health staff: **Mandile Samantha Thobela** and **Lehlohonolo Matabane** for long-term storage, sorting, and transportation of samples.

I also give special thanks to my colleagues: **Dr Eric Morifi** and **Refilwe Moepya** in the Witwatersrand mass spectrometry laboratory for keeping a good friendly and professional working environment, and for brilliant work in maintaining our analytical instruments.

To all my fellow postgraduate students from the peptide research group thanks for the support and advice. To **Dr Thabo Peme** thank you for proofreading my dissertation.

Lastly, I would like to thank my family, My Mother **Neuoe Mbhele**, brother **Katleho Mbhele** and my younger sister **Relebohile Mbhele** for their support and encouragement to keep on working and achieving my academic goals.

Finally, I would like to thank the National Research Foundation (NRF) and the School of chemistry for their financial support.

In memory of my father

Jackie Mbhele

1949-2014

Abstract

Diabetes is a chronic disease that causes serious global challenges to the health care system and continues to increase in the adult population. The lack of insulin or insulin action is the primary cause of diabetes. Diabetes has several complications including retinopathy, neuropathy, and nephropathy which affect the eyes, nerves, and kidneys respectively. Diabetic nephropathy remains the main cause of chronic kidney disease and the current diagnostic methods lack specificity, selectivity and cannot detect the disease early and/or accurately monitor the progression of the disease.

The aim of this study was to develop a quantitative ultra-high performance liquid chromatography coupled to high resolution mass spectrometry method for measuring amino acids concentration ratios in urine samples of non-diabetic and diabetic patients which can be used by South African hospitals as an early diagnostic tool of diabetic nephropathy.

A liquid chromatography mass spectrometry method with optimum separation conditions of target amino acids (arginine, citrulline, isoleucine, ornithine, phenylalanine, and tyrosine) on a C18 column was developed. The developed method was validated for linearity R^2 in the analytical range of 0.01 mg/L to 5.00 mg/L. The accuracy (%recovery) and precision (%RSD), of the developed method was determined by analysing quality control samples spiked with amino acid concentrations of 0.10, 0.30, 0.50, 3.00, and 4.00 mg/L in between runs and on two different days. The validated method was then used to analyse urine samples qualitatively and quantitatively.

The collection of spot urine samples from two groups of participants: the diabetes (glycated haemoglobin > 6.4%) group and non-diabetic patients (glycated haemoglobin ≤ 6.4%) group was conducted at Chris Hani Baragwanath Hospital, South Africa. The degree of renal function predicted by the estimated glomerular filtration (eGFRs) was used to categorise samples into 5 stages of chronic kidney disease. The qualitative mass spectrometry data obtained from the analysis of urine samples was submitted to the MetaboAnalyst data base to determine pathways affected by diabetic nephropathy at different stages of chronic kidney disease and to

also identify compounds that can be used as biomarkers for both the onset and progression of the disease.

The linearity R^2 on day one ranged from 0.960-0.999, while on day two it was 0.962-0.999 therefore the method showed good accuracy for all the amino acids over the two days analysis period. Amino acid concentration ratios of ornithine/arginine, tyrosine/phenylalanine, and citrulline/ornithine showed statistically significant differences between the control and the chronic kidney disease patient groups (stage 1 and stage 3). These amino acids ratios may have the potential to indicate both the onset and progression of the disease.

Several biomarkers were identified in the qualitative analysis, but L-arginine was the most sensitive biomarker for early chronic kidney disease as well as for disease progression. The metabolism of D-glutamine was one of the pathways that was strongly impacted by the disease and could be a predictor for the disease progression.

The study demonstrated that both the quantitative data and qualitative data obtained from the mass spectrometry, when used together or independently can assist in the biomarker discovery for diabetic nephropathy. It also showed that mass spectrometry is a promising technique for the detection and quantification of biomarkers, and the method would be revalidated to ensure that all the validation criteria are met.

List of tables

Table 3. 1: MRM conditions used to generate fragment ions of each targeted amino acid (parent ion). 47

Table 4. 1: Matrix effects determined by calculating the matrix effect factor (MEF) using average peak areas (mean). 61

Table 4. 2: Analytical method characteristics for the quantification of Arg, Cit, Ile, Orn, Phe, and Phe on (intra-day) of method validation. 64

Table 4. 3: Recoveries of QC samples for BU spiked with 0.10, 0.30, 0.50, 3.00 and 4.00 mg/L of amino acid concentrations respectively on intra-day method validation n=9. 65

Table 4. 4: Recoveries of QC samples for U19 spiked with 0.10, 0.30, 0.50, 3.00 and 4.00 mg/L of amino acid concentrations respectively on intra-day method validation n=9. 66

Table 4. 5: Analytical method characteristics for the quantification of Arg, Cit, Ile, Orn, Phe, and Phe on day 2 method validation. 67

Table 4. 6: Extraction of QC samples (BU) when spiked with 0.10, 0.30, 0.50, 3.00 and 4.00 mg/L of amino acid concentration respectively day 2 method validation n=9. 68

Table 4. 7: Extraction of QC samples (U19) when spiked with 0.10, 0.30, 0.50, 3.00 and 4.00 mg/L of amino acid concentration respectively day 2 method validation n=9. 69

Table 4. 8: The %recovery and %RSD of inter-day at spiked amino acid concentration (0.10, 0.30, 0.50, 3.00 and 4.00 mg/L) in BU n=18. 70

Table 4. 9: The %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (0.10 mg/L) in U19. 71

Table 5. 1: Pathway analysis of significant hits and or impact (diabetes group and non-diabetes). 79

Table 5. 2: Control (n=17) vs stage 1 CKD group (n=66) (eGFR ≥ 90 mL/min per 1.73 m^2). 83

Table 5. 3: Control (n=17) vs stage 2 (n=) (eGFR ≥ 60 < 90 mL/min per 1.73 m^2). 84

Table 5. 4: Control vs stage 3 (eGFR ≥ 30 < 60 mL/min per 1.73 m^2). 84

Table 5. 5: Control vs stage 4 (eGFR ≥ 15 < 30 mL/min per 1.73 m^2). 85

Table 5. 6: Control vs patients without eGFR data. 86

Table 5. 7: Potential biomarkers at different stages of diabetic nephropathy. 95

Sd Table 7. 1: Peak areas of Arg in H_2O, BU and U19 at spiked concentrations of 0.10, 0.50 and 4.00 mg/L 102

Sd Table 7. 2: Peak areas of Cit in H_2O, BU and U19 at spiked concentrations of 0.10, 0.50 and 4.00 mg/L 103

Sd Table 7. 3: Peak areas of Ile in H_2O, BU and U19 at spiked concentrations of 0.10, 0.50 and 4.00 mg/L 104

Sd Table 7. 4: Peak areas of Orn in H_2O, BU and U19 at spiked concentrations of 0.10, 0.50 and 4.00 mg/L 105

Sd Table 7. 5: Peak areas of Phe in H_2O, BU and U19 at spiked concentrations of 0.10, 0.50 and 4.00 mg/L 106

Sd Table 7. 6: Peak areas of Tyr in H_2O, BU and U19 at spiked concentrations of 0.10, 0.50 and 4.00 mg/L 107

Sd Table 7. 7: The mean average, %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (0.10 mg/L) in U19. 108

Sd Table 7. 8: The mean average, %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (0.30 mg/L) in U19. 108

Sd Table 7. 9: The mean average, %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (0.50 mg/L) in U19. 109

Sd Table 7. 10: The mean average, %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (3.00 mg/L) in U19. 109

Sd Table 7. 11: The mean average, %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (4.00 mg/L) in U19. 110

Sd Table 7. 13: The mean average, %recovery and %RSD of day 1 and day 2 spiked with 0.30 mg/L amino acid concentration and overall accuracy and precision 111

Sd Table 7. 14: The mean average, %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (0.50 mg/L) in BU 111

Sd Table 7. 15: The mean average, %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (3.00 mg/L) in BU 112

Sd Table 7. 16: The mean average, %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (4.00 mg/L) in BU 112

Sd Table 7. 17: Regression analysis of Arg and calibration curve (X variable 1line fit plot) 113

Sd Table 7. 18: Regression analysis of Cit and calibration curve (X variable 1line fit plot) 114

Sd Table 7. 19: Regression analysis of Ile and calibration curve (X variable 1line fit plot) 115

Sd Table 7. 20: Regression analysis of Orn and calibration curve (X variable 1line fit plot) 117

Sd Table 7. 21: Regression analysis of Phe and calibration curve (X variable 1line fit plot) 118

Sd Table 7. 22: Regression analysis of Tyr and calibration curve (X variable 1line fit plot) 119

Sd Table 7. 23: Regression analysis of Arg and calibration curve (X variable 1line fit plot) 121

Sd Table 7. 24: Regression analysis of Cit and calibration curve (X variable 1line fit plot) 122

Sd Table 7. 25: Regression analysis of Ile and calibration curve (X variable 1line fit plot) 123

Sd Table 7. 26: Regression analysis of Orn and calibration curve (X variable 1line fit plot) 124

Sd Table 7. 27: Regression analysis of Phe and calibration curve (X variable 1line fit plot) 126

Sd Table 7. 28: Regression analysis of Tyr and calibration curve (X variable 1line fit plot) 127

Sd Table 7. 29: Pathway analysis Control vs diabetic group 129

Sd Table 7. 30: Control vs stage 1 CKD 136

Sd Table 7. 31: Control vs stage 2 CKD 137

Sd Table 7. 32: Control vs stage 3 CKD 139

Sd Table 7. 33: Control vs stage 4 CKD 145

Sd Table 7. 34: Control vs unknown CKD 147

List of schemes

Scheme 1. 1: The process of glycation of haemoglobin (HbA1c) by glucose molecule [39]. 8

Scheme 1. 2: Metabolism of phenylalanine, phenylalanine is converted to tyrosine by the action of phenylalanine hydroxylase [82]..... 19

Scheme 1. 3: Pathway of branched chain amino acids [83]. 20

Scheme 1. 4: The metabolism of arginine [84]. 21

Scheme 1. 5: The conversion of serine to glycine by the action of hydroxymethyltranferase enzyme [85]..... 21

Scheme 1. 6: Structural characteristics (quantifier ions) of phenylalanine (top) and tyrosine (bottom) obtained by fragmentation through collision at 15.00 eV and 12.00 eV using multiple reaction monitoring (MRM) respectively [89]. 23

List of figures

Figure 1. 1: Metabolic pathways affected by diabetes as well as different organs and their biomarker molecules [24]. 4

Figure 1. 2: The anatomy of a kidney [69]. 14

Figure 1. 3: Advance glycation products (AGEs) [75]. 15

Figure 1. 4: The normal functioning of a glomerulus [69]. 17

Figure 4. 1: Schematic diagram of the workflow for calibration, QC's standards and urine samples analysed on the UHPLC-ESI-QTOF-HRMS (method B). 56

Figure 4. 2: The separation of target amino acids (1.00 mg/L standard in H_2O) on a C18 column. Chromatogram A (1=Orn, 2=Arg, 3=Cit, 4=Ile, 5=Tyr, 6=Phe), Chromatogram B [– blue line UHPLC pressure (bars) – red line multistep gradient elution 5% (0-2.5 min), 5-95% (2.5-8 min), 95% (8-9 min), 95-5% (9-9.5 min) 5% (9.5-10 min) of solvent B, – green line flow rate (0.40 mL/min)]. 58

Figure 4. 3: The separation of target amino acids (1.00 mg/L standard in H_2O) on a C18 column. Chromatogram (1=Orn, 2=Arg, 3=Cit, 4=Ile, 5=Tyr, 6=Phe), UHPLC conditions, [– blue line UHPLC pressure (bars) – red line multistep gradient elution of 5% (0-2 min), 5-40% (2-6 min), 40-95% (6-8 min), 95% (8-9 min), 95-5% (9-9.5 min), 5% (9.5-10 min) of solvent B – green line flow rate (0.40 mL/min)]. 58

Figure 4. 4: Chromatogram A top, retention times and intensities of a 5.00 mg/L standard mixture in H_2O and chromatogram B bottom, retention times and intensities of 5.00 mg/L standard mixture in blank urine (BU) matrix. 59

Figure 4. 5: Compares recovery of target amino acid in BU and U (urine sample U33) spiked with 1.00 mg/L concentrations using method A and method B. 63

Figure 4. 6: The average amino acids concentrations in the non-diabetic patients. . 72

Figure 4. 7: The average amino acids concentrations in diabetic patients in the negative stage of nephropathy based on A/C ratio. 72

Figure 4. 8: The average amino acids concentrations in diabetic patients in the microalbumin stage of nephropathy based on A/C ratio. 73

Figure 4. 9: The average amino acids concentrations in diabetic patients in the microalbumin stage of nephropathy based on A/C ratio. 73

Figure 5. 1: Control vs diabetes group. The PLS-DA multivariate was used. A (Plot of ROC curves for all biomarker model based on its average performance across all MCCV runs model-0). B (Plot of the predictive accuracy of biomarker models with an

increasing number of features model-6. C (plot of predicted class probabilities for all samples using a single biomarker model-6. The most accurate biomarker model is highlighted with a red dot), D (Plot of the most important features of a selected model-6 ranked from most to least important). 80

Figure 5. 2: Enrichment factor pathway analysis. 82

Figure 5. 3: Control vs stage 1 CKD group (eGFR ≥ 90 mL/min per 1.73 m^2). The PLS-DA multivariate was used. A (Plot of ROC curves for all biomarker model based on its average performance across all MCCV runs model-0). B (Plot of the predictive accuracy of biomarker models with an increasing number of features model-2. C (plot of predicted class probabilities for all samples using a single biomarker model-2. The most accurate biomarker model is highlighted with a red dot), D (Plot of the most important features of a selected model-2 ranked from most to least important). 87

Figure 5. 4: Control vs stage 2 CKD group (eGFR ≥ 60 < 90 mL/min per 1.73 m^2). The PLS-DA multivariate was used. B (Plot of ROC curves for all biomarker model based on its average performance across all MCCV runs model-0). A (Plot of the predictive accuracy of biomarker models with an increasing number of features model-2. C (plot of predicted class probabilities for all samples using a single biomarker model-2. The most accurate biomarker model is highlighted with a red dot), D (Plot of the most important features of a selected model-2 ranked from most to least important. 89

Figure 5. 5: Control vs stage 3 CKD group (eGFR ≥ 30 < 60 mL/min per 1.73 m^2). The PLS-DA multivariate was used. B (Plot of ROC curves for all biomarker model based on its average performance across all MCCV runs model-0). A (Plot of the predictive accuracy of biomarker models with an increasing number of features model-6). C (plot of predicted class probabilities for all samples using a single biomarker model-6. The most accurate biomarker model is highlighted with a red dot), D (Plot of the most important features of a selected model-6 ranked from most to least important). 90

Figure 5. 6: Control vs stage 4 CKD group (eGFR ≥ 15 < 30 mL/min per 1.73 m2). The PLS-DA multivariate was used. B (Plot of ROC curves for all biomarker model based on its average performance across all MCCV runs model-0). A (Plot of the predictive accuracy of biomarker models with an increasing number of features model-2). C (plot of predicted class probabilities for all samples using a single biomarker model-2. The most accurate biomarker model is highlighted with a red dot), D (Plot of the most important features of a selected model-2 ranked from most to least important). 91

Figure 5. 7: Control vs patients without eGFR data. The PLS-DA multivariate was used. B (Plot of ROC curves for all biomarker model based on its average performance across all MCCV runs model-0). A (Plot of the predictive accuracy of biomarker models with an increasing number of features model-3). C (plot of predicted class probabilities for all samples using a single biomarker model-3. The most accurate biomarker model is highlighted with a red dot), D (Plot of the most important features of a selected model-3 ranked from most to least important). 92

List of abbreviations

AA	Amino acids
ADMA	Asymmetric dimethylarginine
Arg	Arginine
BU	Blank Urine
Cit	Citrulline
CKD	Chronic kidney disease
CONC	Concentration
CVD	Cardiovascular disease
CSV	Comma-separated values
DKD	Diabetic kidney disease
DN	Diabetic nephropathy
DOH	Department of health
eGFR	Estimated glomerular filtration
ESI	Electron spray ionisation
ESRD	End stage renal disease
eV	Electron voltage
GFR	Glomerular filtration rate
HbA1c	Haemoglobin A1c
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometer
hrs	Hours
IDF	International Diabetes Federation
Ile	Isoleucine
kV	Kilovolt
L	Litre
LC	Liquid chromatography
LLOD	Lowest limit of detection
LLOQ	Lowest limit of quantification
LOD	Limit of detection
LOQ	Limit of quantification
MCCV	Monte Carlo cross-validation
mg	Milli grams
min	Minutes
mL	Millilitre
MLOQ	Mid-range limits of quantification
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
netCDF	network Common Data Form
Orn	Ornithine
Phe	Phenylalanine
PiMP	Polyomics integrated Metabolomics Pipeline
QC	Quality control
rmp	Revolutions per minute
RSD	Relative standard deviation

Sd	Supporting document
sec	Seconds
SSA	Sulfosalicylic acid
spec/sec	Spectrum per second
SDMA	Symmetric dimethylarginine
SRM	Single reaction monitoring
T1D	Type 1 diabetes
T2D	Type 2 diabetes
Tyr	Tyrosine
ULOQ	Upper limit of quantitation
UHPLC	Ultra-high performance liquid chromatography
µL	Micro litre
V	Voltage
v/v	Volume to volume
Vol	Volume
WHO	World health organisation

Table of Contents

Declaration i
Ethical declaration ii
Acknowledgements iii
Abstract iv
List of tables vi
List of schemes ix
List of figures x
List of abbreviations xii
Chapter 1 : Introduction 1
1.1 Overview 1
1.2 Diabetes 4
1.2.1 Diabetes definition 4
1.2.2 Diabetes prevalence 6
1.2.3 Diabetes diagnoses 7
1.3 Diabetic Nephropathy (DN) 10
1.3.1 Diabetic definition 10
1.3.2 Diabetic nephropathy Prevalence 11
1.3.3 Diabetic nephropathy diagnosis 12
1.3.4 Management and treatment (DN) 12
1.4 Kidneys Function 13
1.4.1 Normal kidney function 13
1.4.2 How Diabetic Nephropathy affects the kidneys 14
1.5 Diabetic nephropathy biomarkers 16
1.5.1 Biomarker definition 16
1.5.2 Trends in DN biomarkers 16
1.5.3 Examples of biomarkers (DN) 17
1.6 Amino acid metabolism 18
1.6.1 Essential and non-essential amino acids 18
1.6.2 Interconversions of amino acids 19
1.7 Liquid chromatography coupled to a mass spectrometry or tandem mass spectrometry (LC-MS or LC-MS/MS) 22
1.7.1 HPLC-MS 22

1.7.2 Sample storage 22
1.7.3 Instrumentation 22
1.8 Aims and Objectives 24
1.8.1 Aim 24
1.8.2 The main objective was: 24
1.8.3 Hypothesis and Question 25
1.9 References 25
Chapter 2 African Journal of Laboratory Medicine 34
Relationship between amino acid ratios and decline in estimated glomerular filtration rate in diabetic and non-diabetic patients in South Africa 34
Chapter 3 Chemicals and materials 44
3.1 Chemicals and materials 44
3.2 Ethical considerations 44
3.3 Data analysis 45
3.4 Calibration and QC standards preparation 45
3.4.1 Stock solutions 45
3.4.2 Calibration standards 46
3.4.3 QC standards 46
3.5 UHPLC and URHMS conditions 46
Chapter 4 Method development and validation 48
4.1 Introduction 48
4.1.1 Method validation 48
4.1.2 Method validation parameters 48
4.1.3 Linearity 49
4.1.4 Limits of detection and quantitation 49
4.1.5 Accuracy 49
4.1.6 Precision 50
4.1.7 Selectivity 50
4.1.8 Matrix effects 50
4.1.9 Carryover 51
4.2 Experimental 51
4.2.1 Method development 51
4.2.2 Method validation 53

4.3 Results and discussion 57
4.3.1 Method development 57
4.3.2 Method validation 63
4.3.3 Quantitative data 71
4.4 References 73
Chapter 5 Qualitative analysis 76
5.1 Introduction 76
5.2 Experimental 77
5.3 Results and discussion 78
5.3.1 Matching 90 features (CSV files with m/z values and peak areas of EIC (m/z) to pathways 78
5.3.2 Diabetes pathways 81
5.3.3 Pathway analysis 82
5.3.4 Biomarker analysis 86
5.3.5 Summary of results pathway analysis 93
5.3.6 Summary of results biomarker analysis 94
5.4 Conclusion 97
5.5 References 97
Chapter 6 Conclusion and future work 99
6.1 Conclusion 99
6.2 Future work 99
Chapter 7 Supplementary information 101
7.1 Appendix 101
7.1.1 Matrix effects 102
7.1.2 Recoveries 108
7.1.3 Method validation day1 (Sd Table 7.17 to 7.22) and day 2 (Sd Table 7.23 to 7.28) 113
7.2 Chapter 5 supporting documents 129
7.2.1 Pathway analysis 129
7.3 Biomarker discovery (Sd Tables 30 to 34) 136

Chapter 1: Introduction

1.1 Overview

Diabetes is a chronic disease that causes serious challenges to the health care system and continues to increase in the adult population. It is a metabolic disease that affects well-developed and developing countries [1,2]. The international diabetes federation (IDF) estimated that approximately 5 million deaths occur on an annual basis globally in the population group aged between 20-79 years [3]. South Africa has the highest prevalence of diabetes in sub-Sahara Africa and diabetes is the second highest cause of morbidity and mortality in this country [4]. Diabetes has several complications which can be grouped into two types, macrovascular complications resulting in stroke and atherosclerosis (heart), and microvascular complications resulting in nephropathy (Kidney failure), retinopathy (blindness) and neuropathy (amputations) [5].

One of the most severe complications of diabetes is diabetic nephropathy (DN) which affects the kidneys, and it is highly common among patients who have had type 2 diabetes for more than 10 years. The major challenges faced by health care systems in dealing with this complication, are due to the fact that the development and progression of DN is still not fully understood [6]. Diabetic nephropathy and/or chronic kidney disease (DN/CKD) are the major cause of end-stage renal disease (ESRD) (kidney failure), and a risk factor for cardiovascular disease (CVD) as well as stroke [7].

There are several clinical diagnostic techniques for DN, but the measurement of albumin concentration in urine (microalbuminuria) remains the most common diagnosis for DN. It is used to predict the onset and progression of DN. However, albumin as a biomarker has been shown to lack specificity and sensitivity, which at times may result in an inaccurate diagnosis. Patients with normoalbuminuria (lower concentration of albumin in urine) may show a decline in glomerular filtration rate (GFR) (kidney function), hence researchers are actively looking for better and more sensitive biomarkers [8,9].

Metabolomics research is one of many types of research which are currently being explored in the pursuit of treatment and early diagnosis of diabetes and its complications. This approach uses the metabolic information of patients to determine their health status [10]. Metabolomics aims at improving sensitivity and selectivity for diagnosis and explaining the biochemical processes involved in disease manifestation and progression [11]. There are two main types of metabolomics, the targeted and untargeted type; the targeted type focuses on the quantity of the known groups of metabolites [12]. While untargeted metabolomics is aimed at profiling all metabolites of different classes found in the biological matrix of interest (urine, blood, etc.) [13]. Both targeted and untargeted metabolomics often involves sample preparation such as extraction, pre-concentration, and derivatization of analytes before analysis [12].

Previous metabolomic studies have shown that some branched chain (isoleucine, leucine, valine), and aromatic amino acids (tyrosine, and phenylalanine) are associated with diabetes and diabetic nephropathy progression, these studies also reveal biological mechanisms of amino acids, phospholipids, and lipids [8]. Other studies have looked at genomics, transcriptomics, and proteomics but these have not revealed the progression of DN in detail. Furthermore, plasma metabolomics have shown a limited number of compounds linked to the disease progression, however, urine metabolomics offers a vast measurable range of metabolites that are directly linked to kidney dysfunction [10].

The analysis of amino acids in biological fluids is central in determining the status of a person’s nutrition, tissue damage as well as renal function. Several chromatographic methods such as ion-exchange chromatography as well as pre-and post-column derivatisation methods for the analysis of amino acids have been reported. But these techniques are lengthy and utilize expensive reagents. Furthermore, they are also not suitable for urine metabolomics procedures as they are prone to co-elution of compounds which results in poor identification of analytes [14].

Liquid chromatography-mass spectrometry (LC-MS) enables fast analysis of analytes, reduces the amount of time spent in sample preparation, and increases sample throughput and selectivity [15]. Biofluids of interest in diagnostics and

biomarker research are urine, serum, plasma, and whole blood. Urine has added advantages over other biofluids. It is easily obtained in large quantities non-invasively, compounds are fairly stable in it, and it may reflect most pathological changes in the human body hence its best suited for monitoring disease progression [16].

Newer techniques in chromatography such as high-performance liquid chromatography-electrospray ionization tandem mass spectrometry (HPLC-ESI-MS/MS), give added advantages over traditional LC-MS systems because they reduce sample preparation times while improving throughput and increasing sensitivity [17]. Other advances in mass spectrometry technology are single reaction monitoring (SRM) and multiple reaction monitoring (MRM). These advances enable the separation of target analytes ions from other ions (matrix or inter-fearing compounds), hence increasing selectivity and simplify the identification of target analytes [18]. The SRM and MRM techniques can quantify analytes in complex metrics.

In clinical or forensics studies it is important to validate the developed method used in the identification and quantification of analytes of interest. This ensures that the results are accurate and reliable and will not lead to incorrect conclusions about a patient condition or prosecution of the wrong person in legal proceedings [19]. The method to be validated can be developed in-house, obtained from literature, or from another laboratory. Parameters to be validated include but are not limited to sensitivity, selectivity, accuracy, precision, reproducibility, and stability [20]. Hence, this study focuses on the development and validation of a UHPLC-ESI-HRMS method to be used in the qualitative and quantitative determination of potential DN biomarkers.

This research was conducted in collaboration with the chemical pathology department at Chris Hani Baragwanath Hospital, Gauteng. It was noticed at several clinics that patients progressed to the late stages of kidney disease, and a considerable percentage require dialysis. This meant that the current problem needs immediate attention and exploration of other possible early biomarkers [21].

1.2 Diabetes

1.2.1 Diabetes definition

Diabetes is a metabolic disorder caused by high levels of glucose in the blood (hyperglycaemia). It affects several metabolic pathways, such as fat, carbohydrates, and protein. Diabetes results from the shortage of insulin or from its inability to regulate glucose, therefore a sustained high concentration of glucose in the blood is a characteristic marker for diabetes [22]. The role of insulin in the human body is to regulate the glucose concentration in the blood and its movement or absorption into cells. The insulin hormone is produced in the pancreas. Glucose is converted into energy, and this drives metabolic processes, hence it is an essential molecule [5].

Most cells in the human body require insulin, to store and use nutrients from the food we eat. The shortage or inability to use insulin results in functional changes in muscles and liver (Figure 1.1), such as inhibition of protein synthesis, and this results in protein degradation. Amino acid concentrations in plasma, as well as in urine are altered in patients with diabetes [23].

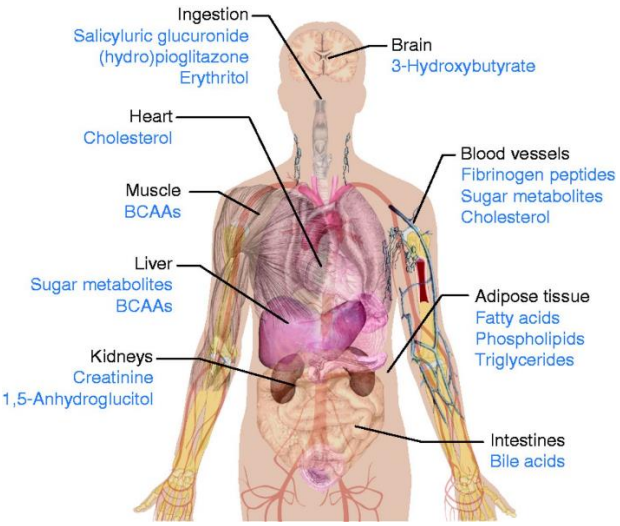


Figure 1. 1: Metabolic pathways affected by diabetes as well as different organs and their biomarker molecules [24].

The majority of cases of diabetes can be classified into two types, namely Type 1 diabetes (T1D) and Type 2 diabetes (T2D).

1.2.1.1 Type 1 diabetes T1D

Type1 diabetes (T1D) is the least prevalent form, and only accounts for 10% of all cases. It is caused by autoimmune destruction of β-cells found in the pancreas and is mostly detected in early childhood [25]. The destruction of β-cells results in insufficient or absence of insulin and in turn the development of hypoglycaemia, ketoacidosis and finally hyperglycaemia [26]. T1D can occur at any age and is no longer called insulin-dependent or juvenile diabetes because other forms can require insulin treatment. The destruction of β-cells in the pancreas can result from many factors such as virus infections, genetic susceptibility, and autoimmune. The exact cause of autoimmune destruction is not fully known, but the body can attack itself by secreting antibodies.

1.2.1.2 Type 2 diabetes T2D

Type 2 diabetes (T2D) occurs when there is insufficient insulin and /or a poor response to insulin in the tissues. Several organs are involved in the development of T2D such as the liver, pancreas, kidneys, brain, small intestine, adipose tissue, and skeletal muscle [27]. The concentration of nitric oxide (NO) which is responsible for monitoring blood vessel's diameter and function, is reduced due to hyperglycaemia or insulin resistance and this increases the risk of patients developing CVD [28]. T2D is caused by a combination of factors, such as stress, anxiety, old age, obesity, poor lifestyle and diet, this results into the body not being able to produce enough insulin to cope with high glucose levels [29].

Some patients can have both type 1 and type 2 diabetes, this is known as double diabetes. It is believed that patients with type 1 diabetes can develop type 2, due to family history, and usually become obese due to insulin resistance even if they increase their insulin dose [30].

Gestational diabetes is another type of diabetes which is found during pregnancy and results from glucose intolerance, it is tested in healthy patients, and only a few patients remain diabetic after pregnancy [31].

1.2.2 Diabetes prevalence

1.2.2.1 Global prevalence

The prevalence of diabetes is increasing globally and has continued to do so for many decades. The world health organisation (WHO) estimated that 30 million people had diabetes worldwide in the 1960s and that this number had risen to 151 million by the year 2000 and to 385 million people by 2013 [32]. In 2019 the International Diabetes Federation (IDF) estimated that there were 463 million people living with diabetes globally and the number is projected to increase to 700 million by 2045. This raises a concern because approximately 40% of both type 1 and type 2 diabetes cases result in DN [33].

Global estimates by IDF show that there is approximately 50% of undiagnosed people who are living with diabetes and in Africa the percentage (69.2%) was estimated to be is higher. Approximately 77% of deaths in Africa occur in the population group below the age of 60 years. The global and African statistics show that approximately 90% of diabetes cases are due to type 2 diabetes [34].

1.2.2.2 African prevalence

In the African continent the IDF estimates that 15.9 million people are living with diabetes, this continent has the largest number of undiagnosed people with diabetes in the world, and Africa will experience a 156% increase in diabetes cases by 2045 [35]. In South Africa the prevalence of diabetes nearly doubled in less than 10 years, increasing from 5.5% in 2000 to 9% in 2009 and approximately 73 000 diabetes-related disabilities were recorded in that year 2009. Approximately 8000 people lost their eyesight, and 2000 amputations were reported mostly linked to type 2 diabetes. Such a rapid rise in the disease results in immense pressure on the economy of the country. The national Department of Health (DOH) proposed a prevention strategy which was to tax foods with high sugar content [36].

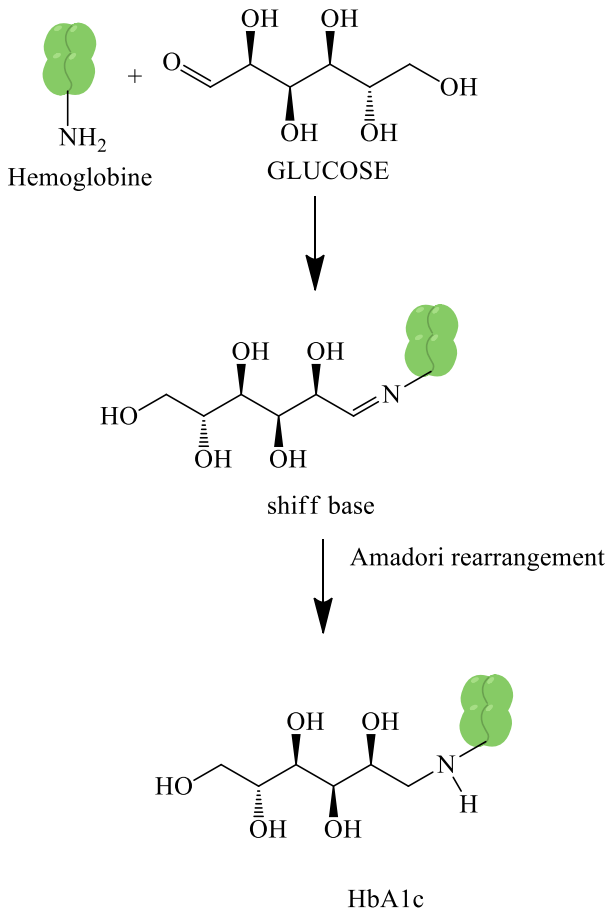
1.2.2.3 South African prevalence

The prevalence of diabetes in South Africa remains underestimated, but recent data shows that the two main contributing factors to this sharp rise are an unhealthy diet accompanied by lack of physical activity, which results in obesity. Approximately 87% of women and 39% of men are reported to be obese, the studies which show these predictions remain small in South Africa [34]. The majority of cases in Africa are more prevalent in the working class which is between the ages of 40 to 60 years. This is different from the global prevalence; diabetes seems to affect older populations >60 years [37]. Estimates by IDF show that diabetes in South Africa had moved from being the fifth cause of death to the second in just 2 years from 2013 to 2015, with 2.3 million people believed to have diabetes in 2015 [38].

1.2.3 Diabetes diagnoses

The criteria for diagnosing diabetes according to World Health Organization (WHO) involves the following parameters: a plasma glucose level of ≥7.0 mmol/L when fasting or a 2-hour plasma glucose level of ≥11.1 mmol/L, and more recently a glycated haemoglobin (HbA1c) level of ≥6.5% is recommended by WHO as the cut-off point for diabetes diagnosis [26]. HbA1c is the standard measure for diagnosing and monitoring type 2 diabetes, it provides information about the patient's average glucose levels in the blood for a period of 2 to 3 months. The binding of glucose to haemoglobin is a non-enzymatic process, whereby glucose binds to the N-terminal of the protein Scheme 1.1 [39].

The measurement of random glucose and HbA1c is usually carried out using a glucometer and DCA vantage instrument respectively. The glucometer measures glucose concentrations in the concentration range of 0.6-33 mmol/L, while DCA vantage measures HbA1c in the percentage range of 2.5-14%, these instruments are currently used in clinical settings such as in Chris Hani Baragwanath Academic Hospital, Diabetic Clinic in South Africa [40].



Scheme 1. 1: The process of glycation of haemoglobin (HbA1c) by glucose molecule [39].

Determining a patient's diabetes status using HbA1c measurement is advantageous over other diabetes prognoses (which measure glucose concentration). HbA1c measurements are not influenced by rapid changes in the concentration of glucose such as when fasting or after meals and is also not affected by the patient's medical conditions [41]. Glycation of haemoglobin is an irreversible non-enzymatic reaction and once glucose is bound HbA1c remains in the red blood cell for almost 4 months [42].

1.2.3.1 Diabetes Complications

There are two main classes of diabetes complications, they are microvascular and macrovascular complications, the most prevalent form is microvascular complications.

1.2.3.2 Macrovascular complications

Macrovascular complications can be subdivided into three diseases which are cardiovascular disease, stroke, and peripheral artery disease (diabetic foot syndrome) responsible for foot ulcers and leg amputations [43].

1.2.3.3 Microvascular complications

Retinopathy, neuropathy, and nephropathy are the result of microvascular complications. The major cause of blindness and visual complications in diabetes patients is diabetic retinopathy, but other causes need to be examined upon diagnosis [44]. Retinopathy is mostly associated with type 1 diabetes and is seen in a third of patients upon diagnosis, but only fully establishes itself in youth [45]. Cardiovascular disease is the leading cause of death among adults with and without diabetes but individuals with diabetes remain at high risk [46].

The most severe disease in microvascular complications is diabetic nephropathy (DN) also referred to as diabetic kidney disease (DKD), it was traditionally described by the presence of microalbuminuria, although recent studies show that microalbuminuria diagnoses lack consistency and can lead to misdiagnoses or proper classification of patients [47].

In both micro- and macrovascular complications, the susceptibility of gender has not been fully demonstrated. Macrovascular complications appear to affect more women than men. Micro- complications with respect to retinopathy are observed more in men than in women, similar trends are observed with neuropathy and DN, but these studies were conducted on small populations. It is important to understand the sex

discrepancies in the occurrence of these complications as this would make therapeutic interventions more effective and personalised [48].

Obesity as a risk factor for developing nephropathy, peripheral neuropathy, retinopathy cardiovascular autonomic neuropathy is actively being researched and some findings show a strong association with nephropathy and neuropathy, but not a strong link to retinopathy [49].

1.3 Diabetic Nephropathy (DN)

1.3.1 Diabetic definition

Diabetic nephropathy (DN) is a microvascular complication resulting from diabetes. The major causes for the onset and progression of DN are hyperglycaemia, hyperlipidaemia, and hypertension [50]. Kidney disease is defined as any abnormal function of the renal system resulting from inflammation, blood flow impairment (hemodynamic), fibrotic lesions, and irreversible organ damage. Kidney disease which results in the damage of the glomeruli (glomerulosclerosis) and kidney tubules (tubulointerstitial fibrosis), as a result of chronic disease, is known as chronic kidney disease (CKD). Most cases of CKD are caused by DN but can also result from other several pathological conditions [51].

An increase in the excretion of urinary albumin accompanied by a decline in renal function as indicated by reduced glomerular filtration rate (GFR) as well as increased plasma concentration of creatinine are characteristics of diabetic nephropathy (DN). Microalbuminuria is the historic biomarker for DN. An amount of 30-300 mg per day of albumin or 2.5-25.0 mg/mmol, 3.5-35.0 mg/mmol of albumin/creatinine ratio in males and females respectively is an indication of microalbuminuria [52]. A decline in glomerular filtration below 60 mL/min/1.73 m^2 for a sustained period of approximately three months with the presence of kidney damage is used as a biomarker for CKD. GFR is calculated by estimation using chronic kidney disease epidemiology collaboration (CK-EPI), and modification of diet in renal disease (MDRD) equations, which uses the concentration of plasma creatinine [53]. GFR or estimated glomerular filtration rate (eGFR) is used to classify CKD into five stages of renal function [54].

The development of DN is a complex process but the driving factor is hyperglycaemia which results in the production of reactive oxygen species, which in turn causes oxidative stress. This then affects several processes such as lipid metabolism, insulin signalling, and glomerular hypertension [55]. The role of oxidative stress in the onset of DN and progression to ESRD is an important factor to consider. It is linked to several metabolic disturbances, as well as affecting the normal functioning of the kidneys. It achieves this by destroying cells in the kidneys (podocytes) and preventing self-repair (fibrosis) of tubulointerstitial tissue [33].

1.3.2 Diabetic nephropathy Prevalence

The prevalence of DN is increasing rapidly due to the ever-increasing diabetes epidemic, approximately 30-50% of patients with diabetes develop the renal disease. Patients with DN have an increased chance of premature death, developing end stage renal disease (ERSD), needing a kidney transplant, and developing cardiovascular diseases. The burden of kidney diseases in Africa remains very irregular (poor data availability), and there is not enough data on the incidence of kidney disease in African populations with diabetes [56]. Approximately 60% of deaths globally are a result of chronic non-communicable diseases and 80% of these deaths occur in developing countries. CKD affects 10% of the adult population, and due to high limited financial resources in many parts of sub-Saharan Africa, not many people requiring dialysis and renal transplant receive these services [7]. South African data shows that T2D was responsible for a 38% increase in the number of deaths in a seven-year period (1999-2006), and 67% of the deaths were attributed to kidney diseases [57].

Amongst T1D patients only a third develop DKD, while the number is higher (40%) in patients with T2D. Most CVD events and deaths are associated with DKD, the highest ever recorded rapid rise of DKD deaths was between 1990-2012 with an almost 100% increase [58].

T2D patients with poorly controlled hyperglycaemia progress to later stages of DN (macroalbuminuria) within a 20-year period, and this results in approximately 20% of

ESRD cases. Due to the irreversible nature of ESRD T2D patients are encouraged to check their DN status once a year after being diagnosed with diabetes [59].

1.3.3 Diabetic nephropathy diagnosis

Diagnosis of diabetic nephropathy (DN) before the 80s was defined as the presence of proteins in urine. A value greater than 0.5 grams in a urine sample collected in a 24-hour period was characteristic of nephropathy. Later research showed that the presence of albumin in urine (microalbuminuria) was an early indicator of DN [60].

In South Africa, the clinical diagnostic technique for determining diabetic nephropathy uses the DCA vantage instrument (Status Analyser), which calculates the ratio of albumin to creatinine in urine samples, a single test using this technique is not reliable hence a patient needs to do at least 2 tests in a period of 3 to 6 months [61].

The test conducted using the DCA vantage instrument takes approximately 7 minutes, and the patient is classified as negative if the albumin to creatinine ratio is < 3.5 mg/mmol for females, and < 2.5 mg/mmol for males. If the ratio is > 3.6 mg/mmol for females, and> 2.6 mg/mmol for males, the patient is showing signs of kidney injury (microalbumin stage). A value > 36 mg/mmol for females, and > 26 mg/mmol for males shows extensive kidney injury (macroalbumin stage) [62]. In the UK and the rest of Europe, GFR is the standard measure for renal function, it is determined by measuring the clearance of insulin, but it is tedious compared to chromium-ethylenediaminetetraacetic acid (Cr-EDTA) which has been shown to give better and more accurate results [63,64].

1.3.4 Management and treatment (DN)

The management of nephropathy patients involves, high blood pressure (BP) control, a BP value of <130/80 mmHg is recommended, also a blood glucose level of < 6.5% needs to be sustained. Nutritional adjustment, lower salt intake, and weight loss for obese patients are some of the many suggested strategies to manage the disease [65].

In order to control DKD or DN, the distinction between non-diabetes and diabetes causes of the disease needs to be well understood and thoroughly investigated These can be seen if proper patient management and follow-ups are done. Recent data shows the following differences between CKD caused by diabetes and the one caused by other factors: a) rapid onset of proteinuria, b) rapid progression of CKD in a short period since the onset of diabetes and c) the absence of retinopathy if these are observed then the cause is nondiabetic [66].

1.4 Kidneys Function

1.4.1 Normal kidney function

A human kidney's (Figure 1.2) functional unit is the nephron and is made up of the glomerulus which is a system of small tubes where filtration takes place, this is achieved by filtering

blood from circulation and converting the filtrate to urine. A GFR of approximately 90L/day is deemed as a normal filtration rate, this helps the kidney to achieve its task of monitoring the concentration of dissolved substances as well as the amount of water in circulation [67]. Amino acid reabsorption by the kidneys is approximately 97-98%. The kidney is also involved in amino acid interconversions, which influence the concentration of plasm amino acid, the conversion of Phe to Tyr, Gly to Ser, and Cit to Arg, these are some of the conversions that are carried out by the kidney. Some studies show that the conversion of Phe to Tyr occurs in the kidneys as much as it does in the liver [68].

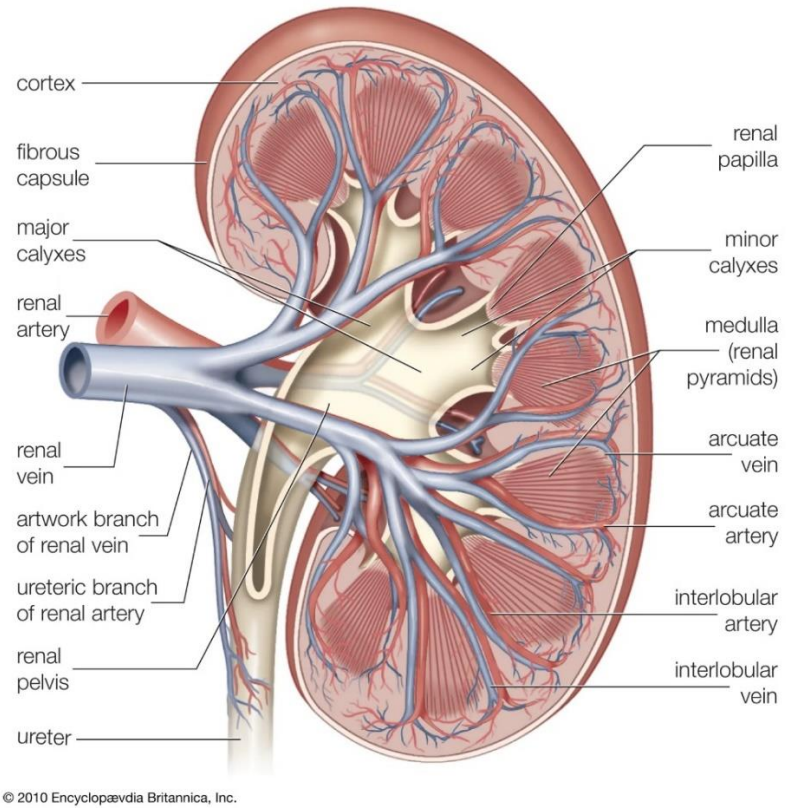


Figure 1. 2: The anatomy of a kidney [69].

1.4.2 How Diabetic Nephropathy affects the kidneys

Diabetic nephropathy is known to affect all the renal cellular elements, which include the glomerular endothelium, mesangial cells, podocytes, and tubular epithelium. The first changes that occur involve the thickening of the glomerular basement, resulting in glomerular enlargement, followed by increased glomerular filtration [70]. During DN progression podocytes can be excreted into the urine, podocytes are a specialised type of cells found in the glomerular basement membrane and are responsible for preventing leakage of proteins, this loss of podocytes is then characterised by the presence of proteinuria [71]. The renal impairments caused by both types of diabetes are the same, as DN progresses the degradation of the compartments of the filtration system of the kidney results in hyperfiltration, and microalbuminuria [72].

An understanding of the disease mechanism and its progression is critical in managing patients and improving their quality of life. Several pathways are affected

by DN, and to date, there is no effective treatment that is being used to prevent patients from progressing to the late stages of the disease [73].

Persistent hyperglycaemia in diabetic patients can result in the formation of advanced glycation end products (AGEs) (Figure 1.3), which leads to advanced glycation of long-lived proteins and short-lived proteins. AGEs are also found in the foods we eat and are responsible for kidney damage. Advanced glycation is believed to be a process by which aging or old proteins are tagged for degradation. When a person gets old, AGEs are formed due to long-term exposure of proteins to glucose, increased insulin resistance, and declining kidney function [74].

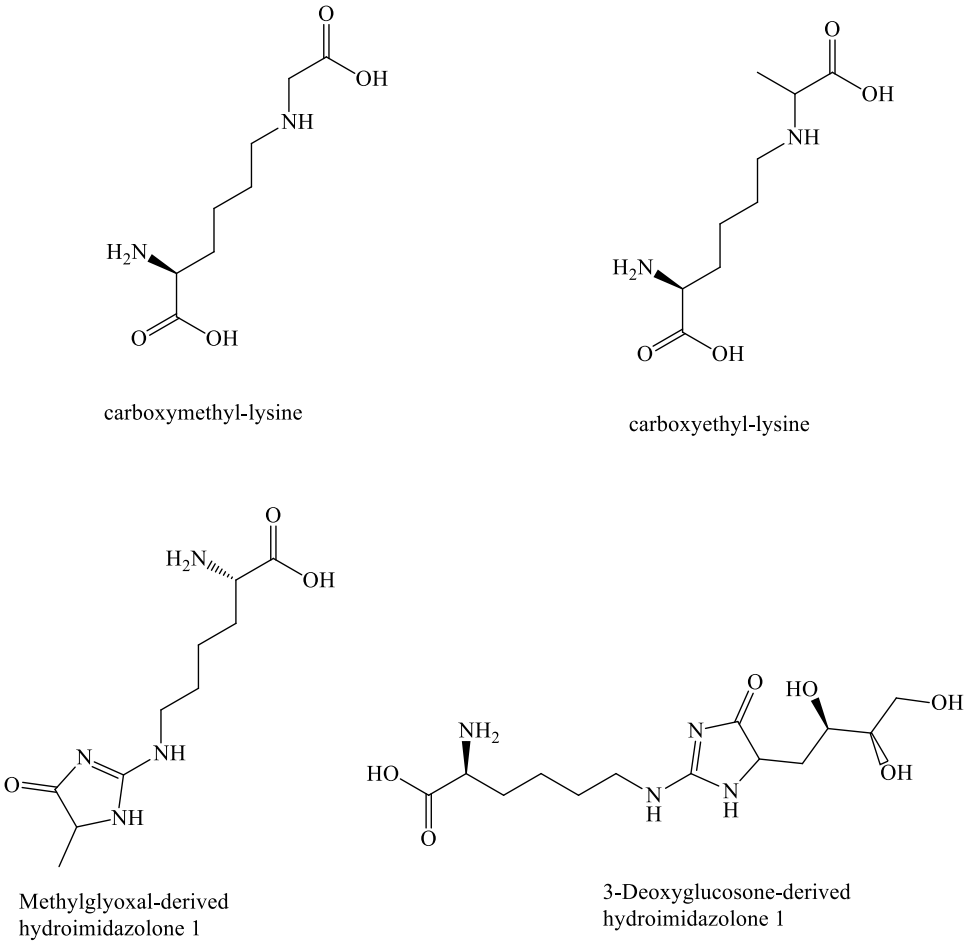


Figure 1. 3: Advance glycation products (AGEs) [75].

1.5 Diabetic nephropathy biomarkers

1.5.1 Biomarker definition

A biomarker is a measure of change in a biological system, which can differentiate between biological processes occurring due to pathogenic, environmental, or pharmaceutical interventions [73]

1.5.2 Trends in DN biomarkers

Monitoring and keeping HbA1c levels low have been shown to reduce the chances of developing micro- and macrovascular complications, but this is not necessarily always effective as in some cases keeping lower glycaemic levels for diabetic patients does not prevent the development of both micro- and macro- complications [41]

DN biomarker research is an active field that continues to grow. There are many molecular approaches, which are searching for early indicators of DN, and they are more sensitive and usable than the current microalbuminuria, one such technique looks at using micro-RNA (miRNA). These types of RNA are believed to increase cell production (fibrosis) in an event where kidney injury has occurred hence there are mostly upregulated and can act as early markers of DN[76].

Other alternative markers that are being studied include renal structures that are outside the glomerulus (tubulointerstitial) (Figure 1.4) and can indicate differences in pathogenic processes of renal injury. These structures have been studied recently and shown to associate well with renal function decline and can indicate the pathogenesis of diabetic nephropathy. These potential biomarkers might be better indicators of renal progression and prognosis in T2D [77].

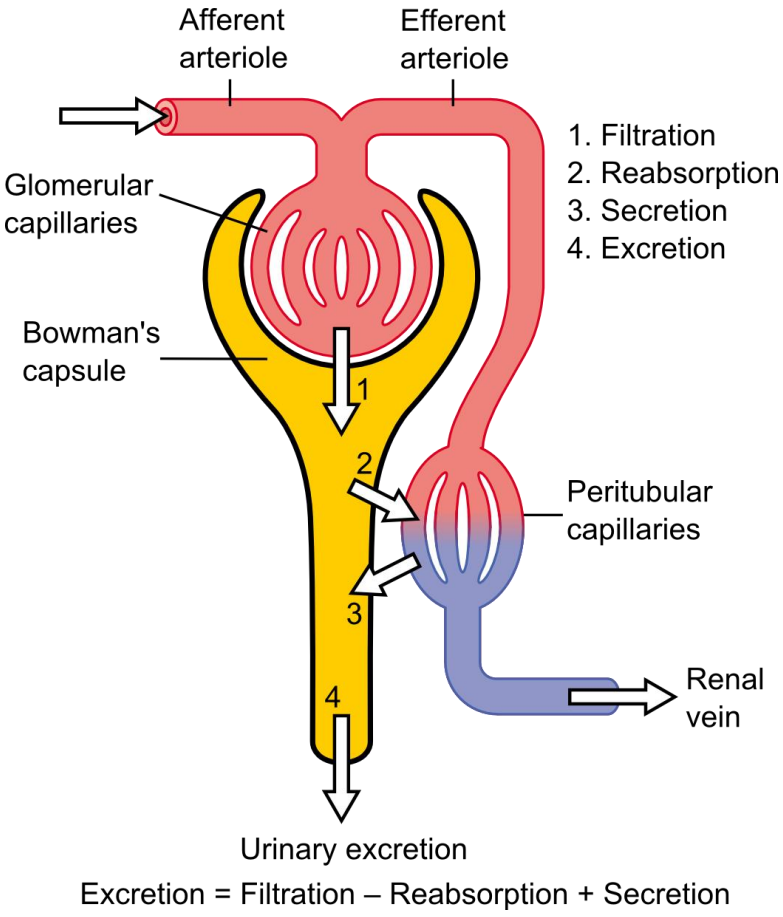


Figure 1. 4: The normal functioning of a glomerulus [69].

1.5.3 Examples of biomarkers (DN)

Arginine derivatives, asymmetric dimethylarginine (ADMA) and symmetric dimethylarginine (SDMA), are formed from the hydrolysis of proteins in an enzymatic process catalysed by arginine methyltransferases. They are promising as possible indicators of renal dysfunction. They are both associated with CVD, ADMA is also linked to other diseases such as high cholesterol, build-up of fat, diabetes, and CKD [78].

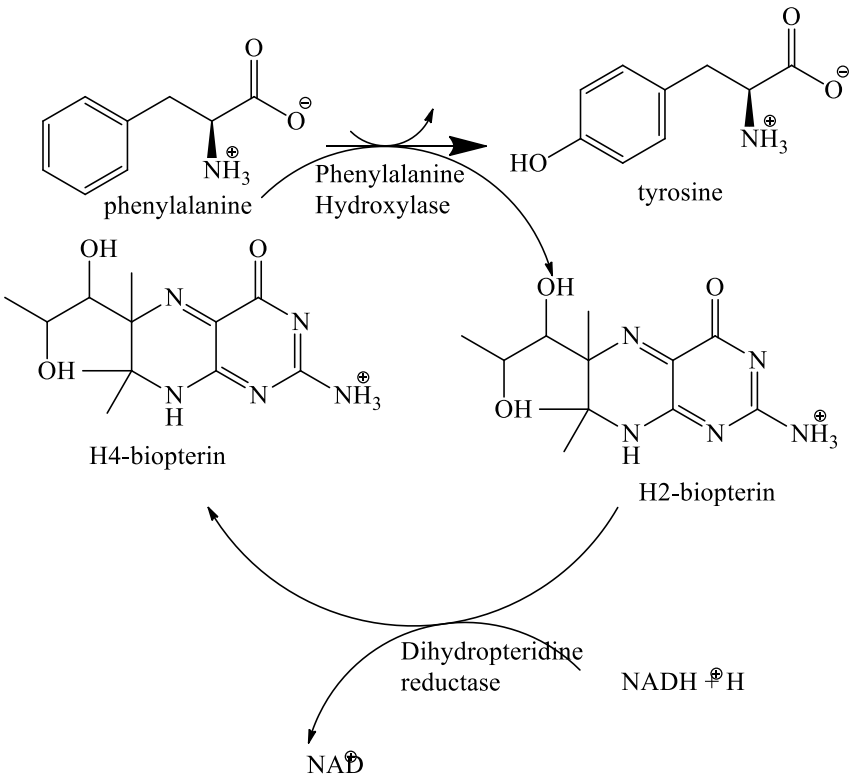
1.6 Amino acid metabolism

1.6.1 Essential and non-essential amino acids

Amino acids are found in two forms of chirality (dexter- D and laevus- L) except for the glycine amino acid. The L-isomers are the most abundant amino acids in nature or biological systems. Amino acids are important building blocks of proteins and also participate in metabolic pathways as intermediates. The D-amino acids which have been thought to play no significant role in biological systems are recently being studied in biomarker discovery [79], but L-amino acids dominate this field.

Some studies have shown that the metabolism, synthesis, and excretion of aromatic amino acids are altered in CKD progression, this is seen in the case of phenylalanine and tyrosine whereby their plasma concentration is increased and decreased respectively. This suggests that the action of the phenylalanine hydroxylase is affected by CKD. The phenylalanine hydroxylase converts phenylalanine to tyrosine, scheme 1.2. and its activity can be evaluated by monitoring the increase or decrease of the tyrosine/phenylalanine concentration ratio which indicates less or more production of tyrosine [80].

Alterations in the concentration of branched-chain amino acids has been linked to T2D (insulin resistance) and these can act as early indicators of T2D long before diagnosis and manifestation of complications. The most abundant essential amino acids are the branch chain amino acids, these include leucine, isoleucine, and valine [81].

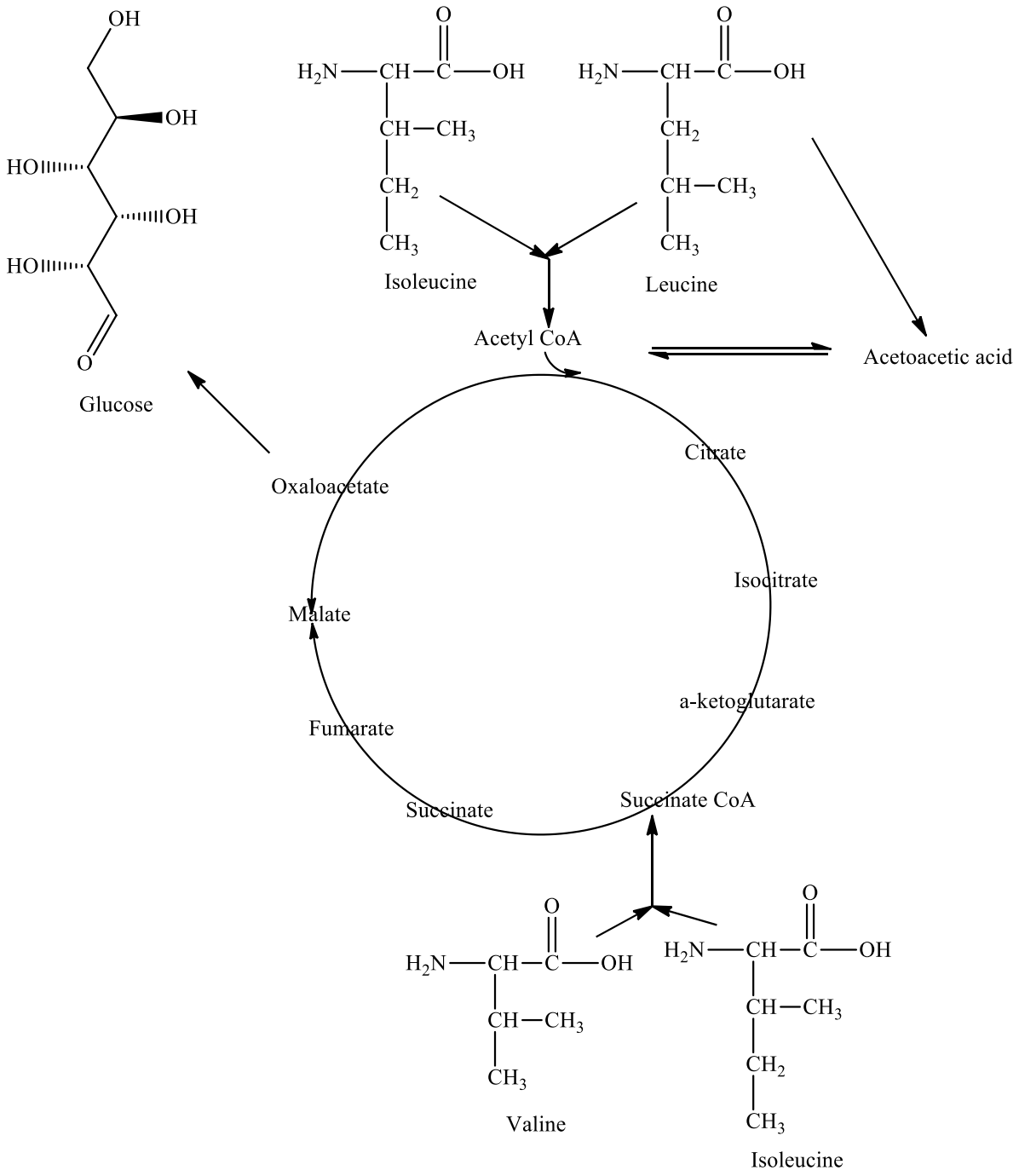


Scheme 1. 2: Metabolism of phenylalanine, phenylalanine is converted to tyrosine by the action of phenylalanine hydroxylase [82]

1.6.2 Interconversions of amino acids

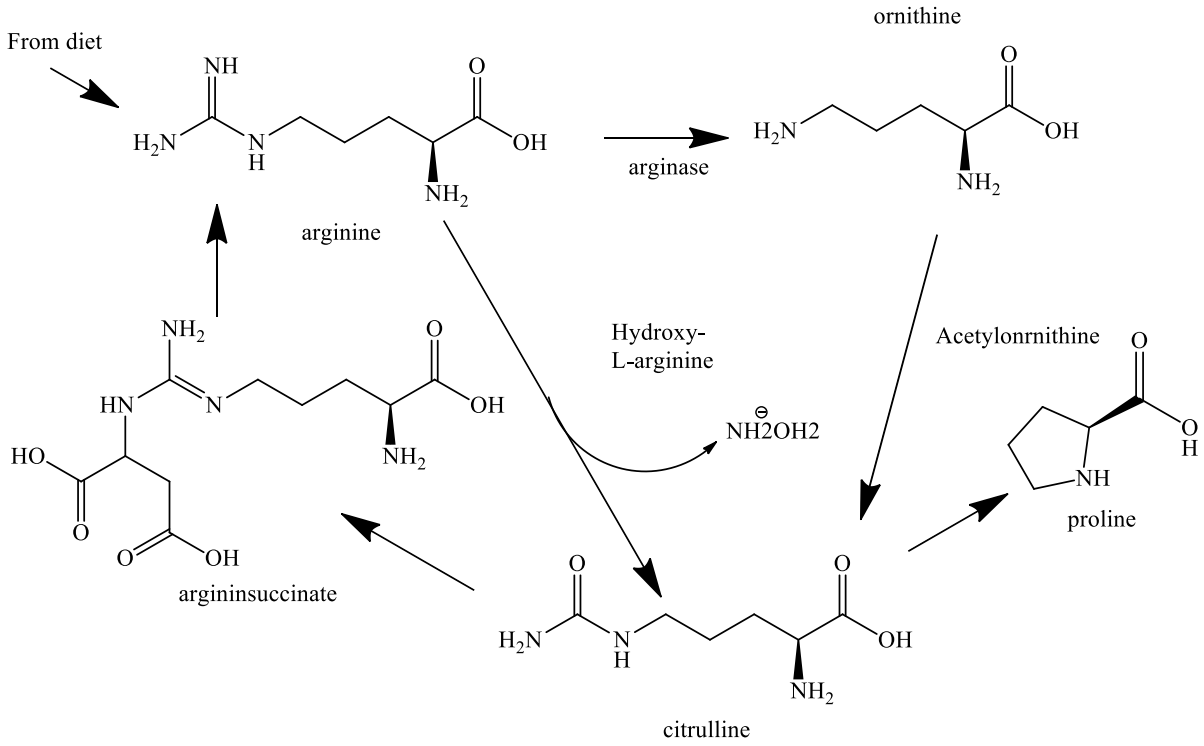
Changes in the amino acid metabolism can give insight as to which pathways are affected by DN, this could be monitored by studying the conversion concentration ratios of certain amino acids in diseased patients and normal patients.

The branched-chain amino acids (leucine, isoleucine, and valine) are essential and are only available through diet. In the TCA cycle, they are catabolised to either acetyl-CoA or succinate-CoA, (Scheme 1.3), by transamination and decarboxylation form the main steps of the reaction [83].



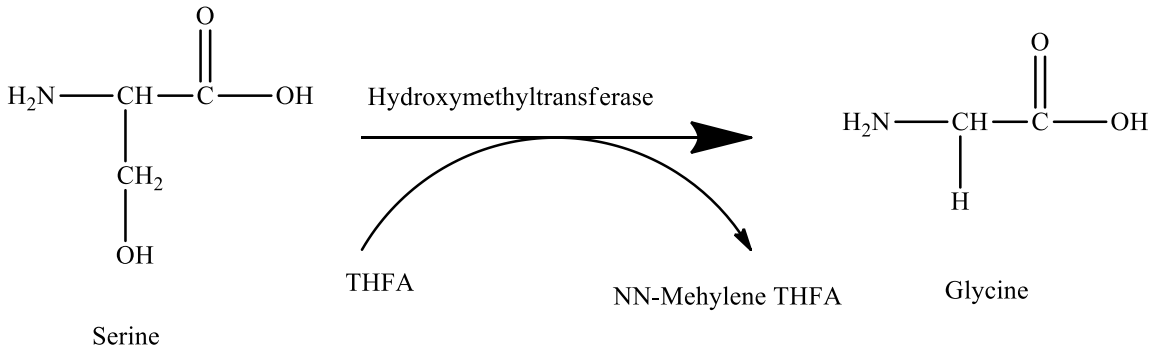
Scheme 1. 3: Pathway of branched chain amino acids [83].

During the metabolism of arginine, arginine is converted to citrulline by the action of hydroxy-L-arginine (Scheme 1.4). It is also converted to ornithine, by the action of arginase. Ornithine is converted to citrulline which gets converted back to arginine by the action of arginine succinate [84].



Scheme 1. 4: The metabolism of arginine [84].

Glycine is a non-essential amino acid, its interconversion from serine is carried out by hydroxymethyltranferase enzyme which removes a carbon from serine and adds it to the methionine metabolic pathway cycles (Scheme 1.5) [85].



Scheme 1. 5: The conversion of serine to glycine by the action of hydroxymethyltranferase enzyme [85].

The conversion concentration ratios of amino acids in diseased patients and normal patients can be monitored using liquid chromatography coupled to a mass spectrometry or tandem mass spectrometry (LC-MS or LC-MS/MS).

1.7 Liquid chromatography coupled to a mass spectrometry or tandem mass spectrometry (LC-MS or LC-MS/MS)

1.7.1 HPLC-MS

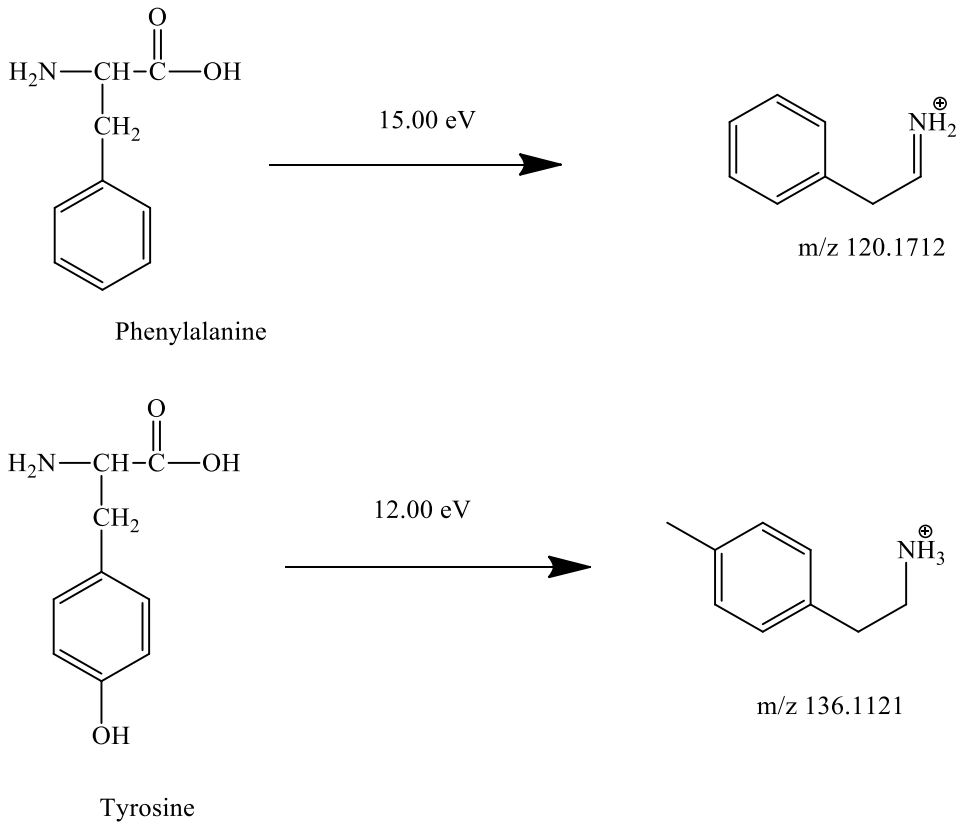
The most recent and more sensitive methods used to analyse amino acids in biological fluids, include the use of a liquid chromatography coupled to a mass spectrometry or tandem mass spectrometry (LC-MS or LC-MS/MS). Traditionally LC-MS methods for amino acid analysis involved long sample preparation techniques such as removal of ammonia in the urine samples [86]. This was followed by derivatisation of amino acids to be analysed.

1.7.2 Sample storage

Sample storage plays an important role in quantitative analysis, the concentration of certain amino acids can be altered by storage conditions such as temperature and the duration of storage. The freezing and thawing cycles can also affect the measurement of amino acid concentration, the ideal condition of storage for urine sample for metabolite analysis is ≤ -20°C and the number of thaw cycles should be kept to a minimum [87].

1.7.3 Instrumentation

Tandem mass spectrometry (MS/MS) technique is based on structural characteristics of analytes obtained by fragmentation through collision (Scheme 1.6), it is a powerful technique for quantitation. In most practices, it is coupled with high-performance liquid chromatography (HPLC), which greatly increases its analytical performance by reducing the matrix effect thereby improving ionization [88].



Scheme 1. 6: Structural characteristics (quantifier ions) of phenylalanine (top) and tyrosine (bottom) obtained by fragmentation through collision at 15.00 eV and 12.00 eV using multiple reaction monitoring (MRM) respectively [89].

The multiple reaction monitoring (MRM) technique is a powerful quantification tool in mass spectrometry and enables for many parent ions (targeted analytes) to be analysed in a single run. Prior knowledge of the masses to be separated is important such that the generated product ions (daughter ions) of each parent ion are stable (reproducible) under known mass spec conditions [88].

For screening a large pool of metabolites within a single analysis, untargeted metabolomics methods which utilize UPLC-QTOFMS are best suited because they capture all detectable metabolites in the matrix of interest (i.e., urine or plasma). In this technique analytes eluting from UHPLC are ionized in an electrospray ionization source (ESI), at atmospheric pressure. Ionization is achieved by the use of a strong electric field applied to the eluent as it passes through the nebulizer, resulting in molecular ions and their adducts, which then enter the mass spectrometer for analysis. The detected analytes are then identified based on their accurate mass,

mass spectra, and retention time by comparison to standards or using online databases [90]

In mass spectrometry (MS) there are different types of mass isolation systems called mass analysers, and they vary in their degree of sensitivity, resolution, and accuracy. They include time of flight (TOF), quadrupole, Fourier transform ion cyclotron resonance (FT-ICR) and orbitrap and can be hyphenated to enable tandem MS experiments (MS/MS, MRM etc.) these analysers are capable of separating molecules of the same nominal mass, provide accurate mass, and detect very low concentrations [91].

In this study electro spray ionization was used to ionize molecules. An eluent solvent either from HPLC or syringe pump, enters into the ESI source through a sprayer needle which is kept at high voltage (2–6 kV). This strong electric field results in highly charged droplets which are dried by N_2 gas, and then flow through a capillary into the mass analyser The advantages of using this technique over other ionisation techniques is its ability to ionize a wide variety of biological molecules such as protein and nucleic acids. It enables rapid identification of molecules based on their molecular mass [92].

1.8 Aims and Objectives

1.8.1 Aim

The aim of this study was to develop a quantitative UHPLC – QTOF HRMS method for measuring amino acid concentration ratios in urine samples of non-diabetic and diabetic patients and to analyse the relationship between the changes in amino acid ratios and known diabetes indicators in the patient clinical data.

1.8.2 The main objective was:

- To develop and validate an LC MS/MS quantitative method for the analysis of amino acids in human urine samples.

- To determine and contrast the changes in concentrations of amino acids (Phe, Tyr, Arg, Ile and Cit) which are affected by diabetes amongst diabetic and non-diabetic patient groups using the UHPLC-HRMS.
- To analyse the relationship between changes in amino acid ratios and the known patient clinical data such as estimated glomerular filtration rate (eGFR) decline in diabetic and non-diabetic patients.
- To determine other metabolites that are significantly elevated or decreased in the amino acid (Arg, Cit, Orn, Ile Phe, and Tyr) interconversion metabolic pathways of diabetic patients using the MetaboAnalyst database.

1.8.3 Hypothesis and Question

Question: Can amino acid concentration ratios acquired using the LC MS/MS from urine samples be used as early indicators of diabetic nephropathy?

Hypothesis: The significant difference observed in concentration and metabolism of some amino acids between non-diabetic patients and diabetic patients can be measured using the UHPLC-HRMS and utilized as an indicator to the early development of diabetic nephropathy.

1.9 References

[1] Tumbo JM, Kadima FN. Screening of long-term complications and glycaemic control of patients with diabetes attending Rustenburg Provincial Hospital in North West Province, South Africa. Afr J Prim Health Care Fam Med 2013;5:1–5. https://doi.org/10.4102/phcfm.v5i1.375.

[2] Guariguata L, Whiting DR, Hambleton I, Beagley J, Linnenkamp U, Shaw JE. Global estimates of diabetes prevalence for 2013 and projections for 2035. Diabetes Res Clin Pract 2014;103:137–49. https://doi.org/10.1016/j.diabres.2013.11.002.

[3] Ogurtsova K, da Rocha Fernandes JD, Huang Y, Linnenkamp U, Guariguata L, Cho NH, et al. IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. Diabetes Res Clin Pract 2017;128:40–50. https://doi.org/10.1016/j.diabres.2017.03.024.

[4] Adedokun AO, ter Goon D, Owolabi EO, Adeniyi OV, Ajayi AI. Prevalence, awareness, and determinants of type 2 diabetes mellitus among commercial taxi drivers in buffalo city metropolitan municipality South Africa: A cross-

sectional survey. Medicine 2019;98:1–6. https://doi.org/10.1097/MD.0000000000014652.

[5] Soumya D, Srilatha B. Late Stage Complications of Diabetes and Insulin Resistance. J Diabetes Metab 2011;02:1–7. https://doi.org/10.4172/2155-6156.1000167.

[6] Meng X, Ma J, Kang AN, Kang SY, Jung HW, Park YK. A Novel Approach Based on Metabolomics Coupled With Intestinal Flora Analysis and Network Pharmacology to Explain the Mechanisms of Action of Bekhogainsam Decoction in the Improvement of Symptoms of Streptozotocin-Induced Diabetic Nephropathy in Mice. Front Pharmacol 2020;11:1–19. https://doi.org/10.3389/fphar.2020.00633.

[7] Adeniyi AB, Laurence CE, Volmink JA, Davids MR. Prevalence of chronic kidney disease and association with cardiovascular risk factors among teachers in Cape Town, South Africa. Clin Kidney J 2017;10:1–7. https://doi.org/10.1093/ckj/sfw138.

[8] Wu T, Qiao S, Shi C, Wang S, Ji G. Metabolomics window into diabetic complications. J Diabetes Investig 2018;9:244–55. https://doi.org/10.1111/jdi.12723.

[9] Lim AKH. Diabetic nephropathy – Complications and treatment. Int J Nephrol Renovasc Dis 2014;7:361–81. https://doi.org/10.2147/IJNRD.S40172.

[10] Sharma K, Karl B, Mathew A v, Gangoiti JA, Wassel CL, Saito R, et al. Metabolomics Reveals Signature of Mitochondrial Dysfunction in Diabetic Kidney Disease. J Am Soc Nephrol 2013;24:1901–12. https://doi.org/10.1681/ASN.2013020126.

[11] Abbiss H, Maker GL, Trengove RD. Metabolomics approaches for the diagnosis and understanding of kidney diseases. Metabolites 2019;9:1–22. https://doi.org/10.3390/metabo9020034.

[12] Wang L, Pi Z, Liu S, Liu Z, Song F. Targeted metabolome profiling by dual-probe microdialysis sampling and treatment using Gardenia jasminoides for rats with type 2 diabetes. Sci Rep 2017;7:1–11. https://doi.org/10.1038/s41598-017-10172-w.

[13] Salihovic S, Broeckling CD, Ganna A, Prenni JE, Sundström J, Berne C, et al. Non-targeted urine metabolomics and associations with prevalent and incident type 2 diabetes. Sci Rep 2020;10:1–9. https://doi.org/10.1038/s41598-020-72456-y.

[14] Le A, Ng A, Kwan T, Cusmano-Ozog K, Cowan TM. A rapid, sensitive method for quantitative analysis of underivatized amino acids by liquid chromatography-tandem mass spectrometry (LC-MS/MS). J Chromatogr B Analyt Technol Biomed Life Sci 2014;944:166–74. https://doi.org/10.1016/j.jchromb.2013.11.017.

[15] Deventer K, Pozo OJ, Verstraete AG, van Eenoo P. Dilute-and-shoot-liquid chromatography-mass spectrometry for urine analysis in doping control and analytical toxicology. TrAC - Trends in Analytical Chemistry 2014;55:1–13. https://doi.org/10.1016/j.trac.2013.10.012.

[16] Kemperman RFJ, Horvatovich PL, Hoekman B, Reijmers TH, Muskiet F a J, Bischoff R. Comparative Urine Analysis by Liquid Chromatography - Mass Spectrometry and Multivariate Statistics : Method Development, Evaluation, and Application to Proteinuria. Journal of Proteome 2007;6:194–206. https://doi.org/10.1021/pr060362r CCC: $37.00.

[17] Tang X, Gu Y, Nie J, Fan S, Wang C. Quantification of Amino Acids in Rat Urine By Solid-Phase Extraction and Liquid Chromatography/Electrospray Tandem Mass Spectrometry: Application To Radiation Injury Rat Model. J Liq Chromatogr Relat Technol 2014;37:951–73. https://doi.org/10.1080/10826076.2013.765451.

[18] Mani DR, Abbatiello SE, Carr SA. Statistical characterization of multiple-reaction monitoring mass spectrometry (MRM-MS) assays for quantitative proteomics. BMC Bioinformatics 2012;13 Suppl 1:1–18. https://doi.org/10.1186/1471-2105-13-S16-S9.

[19] Peters FT, Drummer OH, Musshoff F. Validation of new methods. Forensic Sci Int 2007;165:216–24. https://doi.org/10.1016/j.forsciint.2006.05.021.

[20] Kelly M, DeSilva B. Key elements of bioanalytical method validation for macromolecules. AAPS Journal 2007;9:109–14. https://doi.org/10.1208/aapsj0902017.

[21] Golebemang M, Eugine ZS, Omar B, Trevor G. The burden of chronic kidney disease is increasing, along with other chronic diseases. CME 2007;Vol.25:380–5. https://doi.org/10.10520/EJC63362.

[22] Bekele H, Asefa A, Getachew B, Belete AM. Barriers and Strategies to Lifestyle and Dietary Pattern Interventions for Prevention and Management of TYPE-2 Diabetes in Africa, Systematic Review. J Diabetes Res 2020;2020:1–14. https://doi.org/10.1155/2020/7948712.

[23] Durá-Travé T, Gallinas-Victoriano F, Cortes-Castell E, Moya-Benavent M, Teodoro Durá-Travé, Fidel Gallinas-Victoriano EC-C and MM-B. Amino Acid Plasma Concentrations and Urinary Excretion in Young Diabetics. Diabetes and Its Complications, 2018, p. 109–18. https://doi.org/10.5772/intechopen.72080.

[24] Suhre K. Metabolic profiling in diabetes. Journal of Endocrinology 2014;221:75–85. https://doi.org/10.1530/JOE-14-0024.

[25] Paschou SA, Papadopoulou-Marketou N, Chrousos GP, Kanaka-Gantenbein C. On type 1 diabetes mellitus pathogenesis. Endocr Connect 2018;7:38–46. https://doi.org/10.1530/EC-17-0347.

[26] American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care 2014;37:S81-90. https://doi.org/10.2337/dc14-S081.

[27] Galicia-Garcia U, Benito-Vicente A, Jebari S, Larrea-Sebal A, Siddiqi H, Uribe KB, et al. Molecular Sciences Pathophysiology of Type 2 Diabetes Mellitus. International Journal of Molecular Science 2020;21:1–34. https://doi.org/10.3390/ijms21176275.

[28] Hussain M, Qureshi MA, Arain AQ, Rehman H-U-. Type 2 Diabetic Patients. The Professional Medical Journal 2018;25:1689–95. https://doi.org/10.29309/tpmj/18.4812.

[29] Mukhtar Y, Galalain A, Yunusa U. A modern overview on diabetes mellitus: A chronic endocrine disorder. European Journal of Biology 2020;5:1–15. https://doi.org/10.47672/ejb.409.

[30] Cleland SJ, Fisher BM, Colhoun HM, Sattar N, Petrie JR. Insulin resistance in type 1 diabetes: what is “double diabetes” and what are the risks? Diabetologia 2013;56:1462–70. https://doi.org/10.1007/s00125-013-2904-2.

[31] Buchanan TA, Xiang A, Kjos SL, Watanabe R. What Is Gestational Diabetes? Diabetes Care 2007;30:105–11. https://doi.org/10.2337/dc07-s201.

[32] Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. Diabetes Res Clin Pract 2019;157:107843–53. https://doi.org/10.1016/j.diabres.2019.107843.

[33] Hojs NV, Bevc S, Ekart R, Hojs R. Oxidative stress markers in chronic kidney disease with emphasis on diabetic nephropathy. Antioxidants 2020;9:1–22. https://doi.org/10.3390/antiox9100925.

[34] Pheiffer C, Wyk VP van, Turawa E, Levitt N, Kengne AP, Bradshaw D. Prevalence of type 2 diabetes in South Africa: A systematic review and meta-analysis. Int J Environ Res Public Health 2021;18:1–15. https://doi.org/10.3390/ijerph18115868.

[35] Kibirige D, Lumu W, Jones AG, Smeeth L, Hattersley AT, Nyirenda MJ. Understanding the manifestation of diabetes in sub Saharan Africa to inform therapeutic approaches and preventive strategies: a narrative review. Clin Diabetes Endocrinol 2019;5:1–8. https://doi.org/10.1186/s40842-019-0077-8.

[36] Manyema M, Veerman JL, Chola L, Tugendhaft A, Labadarios D, Hofman K. Decreasing the burden of type 2 diabetes in South Africa: The impact of taxing sugar-sweetened beverages. PLoS One 2015;10:1–17. https://doi.org/10.1371/journal.pone.0143050.

[37] Bertram MY, Jaswal AVS, van Wyk VP, Levitt NS, Hofman KJ. The non-fatal disease burden caused by type 2 diabetes in South Africa, 2009. Glob Health Action 2013;6:206–12. https://doi.org/10.3402/gha.v6i0.19244.

[38] Mutyambizi C, Booysen F, Stokes A, Pavlova M, Groot W. Lifestyle and socio-economic inequalities in diabetes prevalence in South Africa: A decomposition analysis. PLoS One 2019;14:1–21. https://doi.org/10.1371/journal.pone.0211208.

[39] Sherwani SI, Khan HA, Ekhzaimy A, Masood A, Sakharkar MK. Significance of HbA1c test in diagnosis and prognosis of diabetic patients. Biomark Insights 2016;11:95–104. https://doi.org/10.4137/Bmi.s38440.

[40] Tanyanyiwa D, Dandara C, Bhana SA, Pauly B, Marule F, Ramokoka M, et al. Implementation of POCT in the diabetic clinic in a large hospital. Afr Health Sci 2015;15:902–7. https://doi.org/10.4314/ahs.v15i3.26.

[41] Lau C, Aw T. HbA1c in the diagnosis and management of diabetes mellitus: an update. Diabetes Updates 2020;6:1–4. https://doi.org/10.15761/du.1000137.

[42] Chauhan N. Laboratory Diagnosis of HbA1c: A Review. J Nanomed Res 2017;5:1–10. https://doi.org/10.15406/jnmr.2017.05.00120.

[43] Papatheodorou K, Banach M, Bekiari E, Rizzo M, Edmonds M. Complications of Diabetes 2017. J Diabetes Res 2018;20:1–5. https://doi.org/10.1155/2018/3086167.

[44] Khan A, Petropoulos IN, Ponirakis G, Malik RA. Visual complications in diabetes mellitus: beyond retinopathy. Diabet Med 2017;34:478–84. https://doi.org/10.1111/dme.13296.

[45] Martin M Nentwich, Michael W Ulbig. Diabetic retinopathy - ocular complications of diabetes mellitus. Wold Journal of Diabetes 2015;6:489–99. https://doi.org/10.4239/wjd.v6.i3.489.

[46] Harding JL, Pavkov ME, Magliano DJ, Shaw JE, Gregg EW. Global trends in diabetes complications: a review of current evidence. Diabetologia 2019;62:3–16. https://doi.org/10.1007/s00125-018-4711-2.

[47] Reidy K, Kang HM, Hostetter T, Susztak K. Molecular mechanisms of Diabetic kidney disease. Journal of Clinical Investigation 2014;124:2333–40. https://doi.org/10.1172/JCI72271.

[48] Singh SS, Roeters-van Lennep JE, Lemmers RFH, van Herpt TTW, Lieverse AG, Sijbrands EJG, et al. Sex difference in the incidence of microvascular complications in patients with type 2 diabetes mellitus: a prospective cohort study. Acta Diabetol 2020;57:725–32. https://doi.org/10.1007/s00592-020-01489-6.

[49] Callaghan BC, Reynolds EL, Banerjee M, Chant E, Villegas-Umana E, Gardner TW, et al. The Prevalence and Determinants of Cognitive Deficits and Traditional Diabetic Complications in the Severely Obese. Diabetes Care 2020;43:683–90. https://doi.org/10.2337/dc19-1642.

[50] Merchant ML, Klein JB. Proteomic Discovery of Diabetic Nephropathy Biomarkers. Adv Chronic Kidney Dis 2010;17:480–6. https://doi.org/10.1053/j.ackd.2010.09.001.

[51] Brenneman J, Hill J, Pullen S. Emerging therapeutics for the treatment of diabetic nephropathy. Bioorg Med Chem Lett 2016;26:4394–402. https://doi.org/10.1016/j.bmcl.2016.07.079.

[52] Fineberg D, Jandeleit-Dahm KAM, Cooper ME. Diabetic nephropathy: Diagnosis and treatment. Nat Rev Endocrinol 2013;9:713–23. https://doi.org/10.1038/nrendo.2013.184.

[53] Soares AA, Eyff TF, Campani RB, Ritter L, Camargo JL, Silveiro SP. Glomerular filtration rate measurement and prediction equations. Clin Chem Lab Med 2009;47:1023–32. https://doi.org/10.1515/CCLM.2009.263.

[54] Meyers AM. Significance, definition, classification and risk factors of chronic kidney disease in south africa. South African Medical Journal 2015;105:233–6. https://doi.org/10.7196/samj.9412.

[55] Zhang J, Liu J, Qin X. Advances in early biomarkers of diabetic nephropathy. Rev Assoc Med Bras 2018;64:85–92. https://doi.org/10.1590/1806-9282.64.01.85.

[56] Noubiap JJN. Diabetic nephropathy in Africa: A systematic review. World J Diabetes 2015;6:759–73. https://doi.org/10.4239/wjd.v6.i5.759.

[57] Madala ND, Thusi GP, Assounga AGH, Naicker S. Characteristics of South African patients presenting with kidney disease in rural KwaZulu-Natal: A cross sectional study. BMC Nephrol 2014;15:1–9. https://doi.org/10.1186/1471-2369-15-61.

[58] Alicic RZ, Rooney MT, Tuttle KR, §|† ‡. Diabetic Kidney Disease Challenges, Progress, and Possibilities. Clin J Am Soc Nephrol 2017;12:2032–45. https://doi.org/10.2215/CJN.11491116.

[59] Aldukhayel A. Prevalence of diabetic nephropathy among type 2 diabetic patients in some of the arab countries. International Journal of Health Science 2017;11:60–3. https://doi.org/28293155/PMC/5327670.

[60] Gross JL, de Azevedo MJ, Silveiro SP, Canani LH, Caramori ML, Zelmanovitz T. Diabetic nephropathy: Diagnosis, prevention, and treatment. Diabetes Care 2005;28:164–76. https://doi.org/10.2337/diacare.28.1.164.

[61] Ibrahim HO, Stapar D, Mash B. Is screening for microalbuminuria in patients with type 2 diabetes feasible in the Cape Town public sector primary care context? a cost and consequence study. South African Family Practice 2013;55:367–72. https://doi.org/10.1080/20786204.2013.10874378.

[62] Alhajim SA, Sherhan AA. Albumin to Creatinine Ratio Is Fairly Correlated with Albumin Excretion Rate in Chronic Kidney Disease. Am J Lab Med 2018;3:48–52. https://doi.org/10.11648/j.ajlm.20180306.11.

[63] Motta, L.A., Shephard M.D.S., Brink J., Lawson S., Rheeder P. Point-of-care testing improves diabetes managment in primary care clinic in South Africa. Primary Care Diabetes Journal 2017;9:1–6. https://doi.org/10.2215/CJN.06000613.

[64] Regnier T, Shephard M, Shephard A, Graham P, DeLeon R, Shepherd S. Results from 16 years of quality surveillance of urine albumin to creatinine ratio testing for a national indigenous point-of-care testing program. Arch Pathol Lab Med 2020;144:1199–203. https://doi.org/10.5858/arpa.2020-0106-OA.

[65] Moosa MR, van der Walt I, Naicker S, Meyers AM. Important causes of chronic kidney disease in South Africa. South African Medical Journal 2015;105:1–8. https://doi.org/10.7196/SAMJ.9535.

[66] Doshi SM, Friedman AN. Diagnosis and Management of Type 2 Diabetic Kidney Disease. Clin J Am Soc Nephrol 2017;12:1366–73. https://doi.org/10.2215/CJN.11111016.

[67] Diamond GL, Thayer WC, Klotzbach JM, Ingerman LD. Part A Current Issues Urinary cadmium clearance, its relationship to glomerular filtration rate and implications for cadmium epidemiology Urinary cadmium clearance, its relationship to glomerular filtration rate and implications for cadmium epidemiology. J Toxicol Environ Health A 2019;82:1187–98. https://doi.org/10.1080/15287394.2019.1707947.

[68] Garibotto Giacomo G, Sofia A, Saffioti S, Bonanni A, Mannucci I, Verzola D. Amino acid and protein metabolism in the human kidney and in patients with chronic kidney disease. Clinical Nutrition 2010;29:424–33. https://doi.org/10.1016/j.clnu.2010.02.005.

[69] Boundless Anatomy and Physiology. Physiology of the Kidneys n.d.:1–16. https://courses.lumenlearning.com/boundless-ap/chapter/physiology-of-the-kidneys/ (accessed March 23, 2022).

[70] Arora MK, Singh UK. Molecular mechanisms in the pathogenesis of diabetic nephropathy : An update. Vascul Pharmacol 2013;58:259–71. https://doi.org/10.1016/j.vph.2013.01.001.

[71] Jim B, Ghanta M, Qipo A, Fan Y, Chuang PY, Cohen HW, et al. Dysregulated Nephrin in Diabetic Nephropathy of Type 2 Diabetes : A Cross Sectional Study. Early Biomarker for Diabetic Nephropathy 2012;7:1–7. https://doi.org/10.1371/journal.pone.0036041.

[72] Kanwar YS, Wada J, Sun L, Xie P, Wallner EI, Chen S, et al. Diabetic nephropathy: Mechanisms of renal disease progression. Exp Biol Med 2008;233:4–11. https://doi.org/10.3181/0705-MR-134.

[73] Califf RM. Biomarker definitions and their applications. Exp Biol Med 2018;243:213–21. https://doi.org/10.1177/1535370217750088.

[74] Perrone A, Giovino A, Benny J, Martinelli F. Advanced Glycation End Products (AGEs): Biochemistry, Signaling, Analytical Methods, and Epigenetic Effects. Oxid Med Cell Longev 2020;2020:1–18. https://doi.org/10.1155/2020/3818196.

[75] Brings S, Fleming T, Freichel M, Muckenthaler MU, Herzig S, Nawroth PP. Dicarbonyls and advanced glycation end-products in the development of diabetic complications and targets for intervention. Int J Mol Sci 2017;18. https://doi.org/10.3390/ijms18050984.

[76] Abdelsalam M, Wahab AM, el Sayed Zaki M, Motawea M. MicroRNA-451 as an Early Predictor of Chronic Kidney Disease in Diabetic Nephropathy. Int J Nephrol 2020;2020:1–7. https://doi.org/10.1155/2020/8075376.

[77] Satirapoj B. Tubulointerstitial Biomarkers for Diabetic Nephropathy. J Diabetes Res 2018;2018:1–6. https://doi.org/10.1155/2018/2852398.

[78] El-Khoury JM, Bunch DR, Reineks E, Jackson R, Steinle R, Wang S. A simple and fast liquid chromatography-tandem mass spectrometry method for measurement of underivatized l-arginine, symmetric dimethylarginine, and asymmetric dimethylarginine and establishment of the reference ranges. Anal Bioanal Chem 2012;402:771–9. https://doi.org/10.1007/s00216-011-5462-9.

[79] Harada M, Karakawa S, Miyano H, Shimbo K. Simultaneous analysis of d, l-amino acids in human urine using a chirality-switchable biaryl axial tag and liquid chromatography electrospray ionization tandem mass spectrometry. Symmetry (Basel) 2020;12:1–14. https://doi.org/10.3390/SYM12060913.

[80] Kopple JD. Phenylalanine and tyrosine metabolism in chronic kidney failure. J Nutr 2007;137:1586S-1590S. https://doi.org/137/6/1586S [pii].

[81] Adeva MM, Calviño J, Souto G, Donapetry C. Insulin resistance and the metabolism of branched-chain amino acids in humans. Amino Acids 2012;43:171–81. https://doi.org/10.1007/s00726-011-1088-7.

[82] Matthews DE. An Overview of Phenylalanine and Tyrosine Kinetics in Humans 1-3. J Nutr 2007;137:1549–55. https://doi.org/0022-3166/07.

[83] Zhang S, Zeng X, Ren M, Mao X, Qiao S. Novel metabolic and physiological functions of branched chain amino acids: A review. J Anim Sci Biotechnol 2017;8:1–12. https://doi.org/10.1186/s40104-016-0139-z.

[84] Cynober L, Boucher L, Vassons M. Review Arginine metabolism in mammals. Nutritional Biochemistry 1995;6:402–13. https://doi.org/552863(95)00066-9 /10.00.

[85] Pan S, Fan M, Liu Z, Li X, Wang H. Serine, glycine and one-carbon metabolism in cancer (Review). Int J Oncol 2021;58:1–9. https://doi.org/10.3892/ijo.2020.5158.

[86] Hara K, Hijikata Y, Hiraoka E, Ogawa N, Takahashi H. Measurement of urinary amino acids using high-performance liquid chromatography equipped with a

strong cation exchange resin pre-column. Ann Clin Biochem 1999;36:202–6. https://doi.org/10.1177/000456329903600211.

[87] Wang X, Gu H, Palma-Duran SA, Fierro A, Jasbi P, Shi X, et al. Influence of storage conditions and preservatives on metabolite fingerprints in urine. Metabolites 2019;9:1–13. https://doi.org/10.3390/metabo9100203.

[88] Wei R, Li G, Seymour AB. High-throughput and multiplexed LC/MS/MRM method for targeted metabolomics. Anal Chem 2010;82:1–10. https://doi.org/10.1021/ac100331b.

[89] Mikulski L, Schäfer J, Brockmeyer K, Kraut R, Li SM. Comparative studies on similarities and differences of cyclodipeptide oxidases for installation of C–C double bonds at the diketopiperazine ring. Appl Microbiol Biotechnol 2020;104:2523–36. https://doi.org/10.1007/s00253-020-10392-7.

[90] González O, Blanco ME, Iriarte G, Bartolomé L, Maguregui MI, Alonso RM. Bioanalytical chromatographic method validation according to current regulations, with a special focus on the non-well defined parameters limit of quantification, robustness and matrix effect. J Chromatogr A 2014;1353. https://doi.org/10.1016/j.chroma.2014.03.077.

[91] Eckel-Passow JE, Oberg AL, Therneau TM, Bergen HR. An insight into high-resolution mass-spectrometry data. Biostatistics 2009;10:481–500. https://doi.org/10.1093/biostatistics/kxp006.

[92] Banerjee S, Mazumdar S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. Int J Anal Chem 2012;2012:1–41. https://doi.org/10.1155/2012/282574.

Relationship between amino acid ratios and decline in estimated glomerular filtration rate in diabetic and non-diabetic patients in South Africa

Thapelo Mbhele, Donald M. Tanyanyiwa, Refilwe J. Moepya, Sindeep Bhana, Maya M. Makatini

African Journal of Laboratory Medicine | Vol 10, No 1 | a1398, Published: 10 December 2021
DOI: **https://doi.org/10.4102/ajlm.v10i1.1398** |

Author(s)
Thapelo Mbhele, Molecular Sciences Institute, School of Chemistry, Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa.
Donald M. Tanyanyiwa, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; and, Department of Chemical Pathology, Faculty of Health Sciences, Sefako Makgatho Health Sciences University, Pretoria, South Africa
Refilwe J. Moepya, Molecular Sciences Institute, School of Chemistry, Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa.
Sindeep Bhana, School of Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
Maya M. Makatini, Molecular Sciences Institute, School of Chemistry, Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa

Background: Diabetic kidney disease is a major complication resulting from type 1 and type 2 diabetes. Currently, the microalbuminuria test is used to monitor renal function; however, it does not detect albumin until progressive loss of renal function has occurred.
Objective: This study analysed the relationship between changes in amino acid ratios and estimated glomerular filtration rate (eGFR) decline in diabetic and non-diabetic patients.
Methods: Urine samples were collected from participants between February 2019 to April 2019 and analysed from November 2020 to January 2021. Diabetic (glycated haemoglobin > 6.4%) and non-diabetic patients (glycated haemoglobin ≤ 6.4%) from Chris Hani Baragwanath Hospital, South Africa, were further categorised based on the degree of renal function predicted by the eGFRs. Amino acids were quantified using tandem mass spectrometry to determine the concentrations and ratios of tyrosine/phenylalanine, ornithine/arginine, arginine/citrulline and citrulline/ornithine at different stages of the chronic kidney disease.
Results: Among diabetic patients, the tyrosine/phenylalanine ratio showed a statistically significant increase ($p = 0.04$) as the eGFR declined from stage 1 to stage 4; the ornithine/arginine ratio showed a strong negative correlation with eGFR. The citrulline/ornithine ratio differed between the diabetic and non-diabetic patients in stage 1 of chronic kidney disease.
Conclusion: Amino acid ratios (ornithine/arginine and tyrosine/phenylalanine) are affected by the progression of diabetes and can be correlated to renal function. The citrulline/ornithine ratios differ between the studied groups in stage 1 of the disease and may be utilised to predict the onset of chronic kidney disease.
Keywords: diabetic nephropathy; albuminuria; amino acids; LC-MS/MS; chronic kidney disease; glomerular filtration rate.

Introduction

Diabetic nephropathy or chronic kidney disease (CKD), resulting from both type 1 and type 2 diabetes, is one of the leading causes of death in the world among non-communicable diseases.[1,2] Diabetic nephropathy usually develops about 10 years after the onset of diabetes.[3] Chronic kidney disease affects about half a million people globally, and people with CKD are at a high risk of developing stroke and cardiovascular disease. Chronic kidney disease patients are most likely to die from cardiovascular disease than progress to end-stage renal failure.[4]

In South Africa, diabetic nephropathy is diagnosed clinically using a digital community analyser vantage instrument that measures the amount of albumin and creatinine, as well as the ratio of albumin to creatinine (A/C) in urine. The diagnostic criteria based on the A/C ratio are as follows: normal range ≤ 3.5 mg/mmol (female) and ≤ 2.5 mg/mmol (male); microalbumin range 2.6 mg/mmol – 25.0 mg/mmol (female) and 3.6 mg/mmol – 35.0 mg/mmol (male); macroalbumin range 26.0 mg/mmol – 100.0 mg/mmol (female) and 36.0 mg/mmol – 100.0 mg/mmol (male).[5] Although the A/C ratio is a good indicator of CKD, the detected value needs to be sustained in the patient over 3 months to accurately diagnose CKD.[6] It also lacks specificity and sensitivity during a progressive decline in renal function and can give false results, where patients with CKD could appear normal.[7]The glomerular filtration rate (GFR) is an important marker in the assessment of kidney function. It can be calculated using the CKD Epidemiology Collaboration formula and the modification of diet in renal disease equation. The modification of diet in renal disease equation estimates the GFR (eGFR) based on age, sex and serum creatinine (SCr).[7,8] Renal decline in CKD is classified into five stages, which indicate the progression of the disease. These stages are based on the GFR and include stage 1 (≥ 90 mL/min per 1.73 m^2), stage 2 (60 mL/min – 89 mL/min per 1.73 m^2), stage 3 (30 mL/min – 59 mL/min per 1.73 m^2), stage 4 (15 mL/min – 29 mL/min per 1.73 m^2) and stage 5 (< 15 mL/min per 1.73 m^2).[9]

The assessment of GFR using SCr concentrations gives inconclusive results in the early stages of renal decline and lacks sensitivity because SCr only becomes detectable after renal decline has progressed significantly. The CKD Epidemiology Collaboration formula provides a good indicator of the eGFR in the early stages of CKD (> 90 mL/min per 1.73 m^2); however, it also underestimates the GFR in diabetic patients.[10] Similarly, although the modification of diet in renal disease equation provides reliable results for patients with CKD, it may give false positives for healthy patients.[8]

Measuring the concentration of amino acids in the body is important in the diagnosis and management of several metabolic abnormalities. Amino acid concentrations can be used to determine the status of patients' nutrition, renal function and tissue injury.[11] The metabolism and excretion of amino acids such as phenylalanine and tyrosine have been reported to undergo considerable changes in patients with CKD.[12] Urine is the most preferred biofluid in metabolomics due to its ease of collection in large volumes and its lesser complexity compared to other body fluids such as blood.[13,14]

Liquid chromatography tandem mass spectrometry eliminates lengthy sample preparations as amino acids can be analysed without the need for derivatisation.11 The advantages of using this technique include high sensitivity, specificity, and throughput.[15] This technique is based on the structural characteristics of analytes obtained by fragmentation through collision, and it is a powerful technique for quantitation.[16]

This study aimed to assess the relationship between urinary amino acid concentration ratios and renal function decline as predicted by the eGFR in diabetic and non-diabetic patients using liquid chromatography tandem mass spectrometry.

Methods

Ethical considerations

Ethical clearance was obtained from the Human and Animal Rights Chris Hani Baragwanath Academic Hospital Medical Advisory Committee and the University of the Witwatersrand Human Research Ethics Committee (reference number: M180798). A brief explanation of the study was communicated to the participants, who then provided written informed consent to allow the collection and analysis of their urine samples and medical records. Medical practice and ethical guidelines were followed to ensure the confidentiality of the results obtained in this study. Samples were coded to preserve anonymity.

Study design and sampling

A total of 187 spot urine specimens were collected from study participants between February 2019 and April 2019 at the Chris Hani Baragwanath Academic Hospital located in Soweto, Johannesburg, South Africa. This is the biggest hospital in Africa and serves a large community from urban to rural areas; it sees hundreds to thousands of patients annually. Data on ethnicity, age, eGFR, glycated haemoglobin (HbA1c), SCr, urinary creatinine, albumin, A/C and gender were obtained from participants' medical records.

Samples from 36 participants were excluded from the analysis because of missing data (33 medical records without eGFR data and 3 without HbA1c data). All participants with complete medical data (n = 151) were of African descent, including women (n = 104) and men (n = 47) between the ages of 18 years and 90 years.

Urine samples were divided into two groups based on HbA1c: people with diabetes (n = 129; HbA1c > 6.4%) and people without diabetes (n = 22; HbA1c ≤ 6.4%). These groups were further divided into sub-groups based on the stage of CKD as follows: 63 people with diabetes and 9 without diabetes were categorised as being in stage 1 (≥ 90 mL/min per 1.73 m^2), 43 people with diabetes and 10 without diabetes were classified as being in stage 2 (60 mL/min – 89 mL/min per 1.73 m^2), 15 people with diabetes and 3 without diabetes were classified as being in stage 3 (30 mL/min – 59 mL/min per 1.73 m^2), 7 people with diabetes were classified as being in stage 4 (15 mL/min – 29 mL/min per 1.73 m^2), and 1 participant was categorised as being in stage 5 (< 15 mL/min per 1.73 m^2).

The amino acids arginine, citrulline, isoleucine, ornithine, phenylalanine, and tyrosine were quantitatively analysed in each sample. Using the modification of diet in renal disease equation, the eGFR was calculated to assess the degree of kidney function. Urine samples were collected into 20 mL sterile polypropylene screw-top containers and placed on wet ice in polystyrene containers before transportation to the National Institute of Occupational Health biobank for long-term storage at −20 °C. Before analysis, the samples were transported to the University of the Witwatersrand School of Chemistry mass spectrometry laboratory, where they were briefly stored at −20 °C for a few days before analysis.

Sample and data analysis

Chemicals and materials

All amino acids (purity ≥ 98.0%), as well as the formic acid (purity ≥ 98.0%), Sigmatrix Urine Diluent (purity ≥ 98.0%) and HPLC-grade acetonitrile (≥ 99.9%), were obtained from Sigma-

Aldrich (Johannesburg, South Africa). Deionised water, which was used as a solvent for dilution of analytes, was purified using a Direct-Q® 3 UV Millipore system (Molsheim, France).

A stock solution of each amino acid (100 mg/L) was prepared by weighing 10 mg of the amino acid which was then dissolved to a final volume with Millipore H_2O in a 100 mL volumetric flask. From the stock solution, 10 calibration points (0.01 mg/L, 0.05 mg/L, 0.10 mg/L, 0.30 mg/L, 0.50 mg/L, 1.00 mg/L, 2.00 mg/L, 3.00 mg/L, 4.00 mg/L, and 5.00 mg/L) were prepared by diluting appropriate volumes of stock solution in a urine diluent–H_2O (1:5 ratio) mixture. The urine diluent was used as a blank matrix to mimic human urine.

Quantitative analysis

Frozen urine samples were thawed for 30 min, then 1 mL of the urine was transferred into a 2 mL centrifuge tube and diluted with 1 mL of Millipore H_2O, and centrifuged at 3000 rpm for 10 min. Afterwards, 0.2 mL of the supernatant was transferred into a 2 mL ultra-high performance liquid chromatography (UHPLC) vial and 0.8 mL of Millipore H_2O was added to make up the final volume to 1 mL. Standards and samples were analysed on a YMC-Triart C18 UHPLC column (1.9 μm beads, pore size 120 Å, 2 mm ID × 150 mm length) (Analytical Science Technology Pty Ltd, Cape Town, South Africa). All analyses were conducted in triplicate with quality control samples (lowest limit of quantitation [LLOQ]1 [0.1 mg/L], LLOQ3 [0.5 mg/L] and upper limit of quantitation [ULOQ] [4 mg/L]), which were analysed in the same way as the urine samples above and were used to assess the performance (linearity, accuracy, precision, and percentage recovery) of the method.

The Thermo Scientific™ UltiMate™ 3000 UHPLC system coupled to a Bruker Compact Qq-Time-of-Flight high-resolution mass spectrometer (Bruker Daltonics, Bremen, Germany) was used for the analysis. The samples and standards were placed in an autosampler set at 5 °C. A two-buffer system was employed, utilising formic acid as the ion-pairing agent; solvent A consisted of 0.1% formic acid in H_2O (volume per volume), and solvent B comprised 0.1%formic acid in acetonitrile (volume per volume). Twenty μL of the sample was injected into the C18 UHPLC column and eluted using a gradient elution technique, starting from an H_2O:acetonitrile ratio of 95:5% to 5:95% in 10 min; the standards and samples were injected three times.

Mass spectrometry parameters

The electrospray ionisation source was operated in the positive ionisation mode at a temperature of 250 °C. The dry gas (N_2) flow rate was set at 9.0 L/min, the nebuliser gas (N_2) pressure was set to 1.8 bars, the capillary voltage was 4500 V, the end plate offset was -500 V, and the charging voltage was 2000 V.

In the quadrupole, a multiple reaction monitoring method was applied with a mass range of 50 mass-to-charge ratio (m/z) to 1300 m/z to fragment arginine (175 m/z), citrulline (176 m/z), isoleucine (132 m/z), ornithine (133 m/z), phenylalanine (166 m/z), and tyrosine (182 m/z) at 26.00, 14.00, 22.00, 12.00, 15.00 and 12.00 eV collision energies. The time-of-flight detector voltage was set at 2242.9 V.

Data analysis

All data processing was done on the Bruker QuantAnalysis software (Bruker Daltonics, Bremen, Germany) to determine the concentration of amino acids. The dilution factor of 10 was used to determine the actual concentration of amino acids in the urine samples. The LLOQ was determined for each amino acid and found to be within the range of 0.52 mg/L to 0.58 mg/L for arginine, citrulline, and ornithine and 0.12 mg/L to 0.19 mg/L for isoleucine, phenylalanine, and tyrosine (Table 1). Statistical data analysis was carried out in Microsoft

Excel version 2101 (Microsoft Corporation, Redmond, Washington, United States). Mean and standard deviation values were obtained using descriptive statistics, significant differences in variables between diabetic and non-diabetic patients were assessed using analysis of variance, and linear correlations between biochemical and analytical data were determined using Pearson's correlation coefficient (R). A p-value 0.05 or less was considered significant.

Results

All amino acids were well separated on the C18 column and gave different fragment ions and retention times (arginine, 0.81 min; citrulline, 0.90 min; isoleucine, 2.10 min; ornithine, 0.74 min; phenylalanine, 4.44 min and tyrosine, 2.96 min), indicating good selectivity (Figure 1).

In diabetic patients, the ornithine/arginine and tyrosine/phenylalanine ratios increased as the eGFR decreased (Table 2). This relationship was confirmed by strong negative correlations (R = –0.81; p = 0.19 and R = –0.96; p = 0.04) between the amino acid concentration ratios and the eGFR. Conversely, the arginine/citrulline ratio decreased with decreasing eGFR, and this was supported by a statistically significant positive correlation (R = 0.71; p = 0.29). The concentrations of ornithine, isoleucine, tyrosine and phenylalanine decreased with decreasing renal function and this was confirmed by statistically significant correlations ([R = 0.83; p = 0.17], [R = 0.94; p = 0.06], [R = 0.94; p = 0.06] and [R = 1.00; p = 0.004]). The amino acid concentrations demonstrated statistically significant positive correlations to HbA1c coefficient (R) and correlated negatively with urinary creatinine and age. There was no significant correlation between the amino acid concentrations and A/C ratio, albumin, or SCr. Haemoglobin A1c correlated strongly with concentrations of citrulline (R = 0.76), isoleucine (R = –0.83) and ornithine (R = –0.86), while age correlated negatively with the concentrations of isoleucine (R = –0.98), ornithine (R = –0.95), phenylalanine (R = –0.83) and tyrosine (R = –0.80). The trends observed between amino acid concentrations and ratios with clinical parameters for non-diabetic patients are

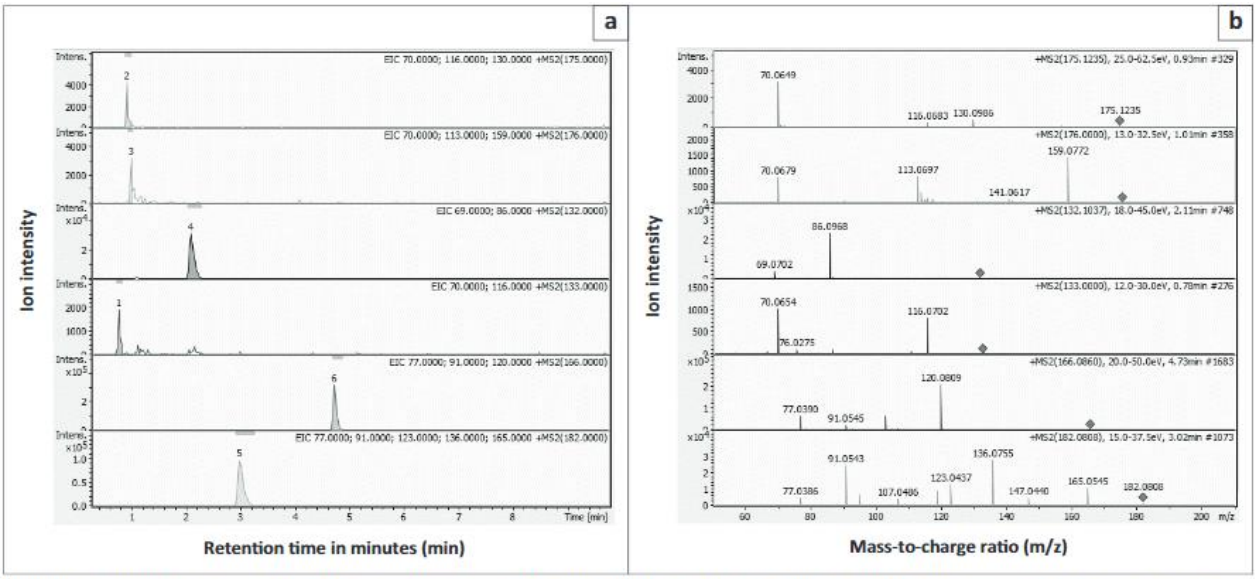


FIGURE 1: Chromatogram showing (a) the UHPLC baseline separation of arginine, citrulline, isoleucine, ornithine, phenylalanine and tyrosine; and (b) the mass spectrum with retention times and collision energy required to fragment each amino acid to give the fragment ions (quant masses). Data analysed between November 2019 and January 2021, Johannesburg, South Africa.

TABLE 1: Analytical method performance characteristics of three spiked blank matrices with amino acid standards at different concentrations, Johannesburg, South Africa, November 2019 to January 2021.

Target analyte	Linearity R^2	Concentration range (mg/L)	Retention time (min)	Lowest limit of detection (mg/L)	Lowest limit of quantification (mg/L)	% Recovery ± % relative standard deviation		
						Blank matrix spiked with (0.1 mg/L) amino acid	Blank matrix spiked with (0.5 mg/L) amino acid	Blank matrix spiked with (4 mg/L) amino acid
Arginine	0.987	0.01–5	0.81	0.17	0.52	111.16 ± 26.68	113.10 ± 16.58	101.02 ± 9.70
Citrulline	0.983	0.01–5	0.90	0.19	0.58	155.99 ± 13.03	103.66 ± 8.08	107.97 ± 12.78
Isoleucine	0.999	0.01–5	2.10	0.05	0.16	108.72 ± 19.05	110.21 ± 10.75	107.08 ± 5.00
Ornithine	0.985	0.01–5	0.74	0.18	0.55	118.96 ± 50.56	124.42 ± 18.09	108.89 ± 17.72
Phenylalanine	0.999	0.01–5	4.44	0.04	0.12	99.08 ± 8.95	125.29 ± 2.80	120.92 ± 1.45
Tyrosine	0.998	0.01–5	2.96	0.06	0.19	90.54 ± 6.45	130.53 ± 2.69	120.92 ± 1.15

Note: Linear (Pearson's correlation) regression was used to determine R^2.

within the normal range for HbA1c and eGFR, and hence do not show disease progression (Table 3). Serum creatinine correlated significantly (p = 0.45) with isoleucine concentration (R = 0.76) and citrulline/ornithine ratios (R = 0.81, p = 0.40) in non-diabetic patients. The tyrosine/phenylalanine ratios (R = 0.85), as well as the concentrations of tyrosine (R = 0.76) and ornithine (R = 0.83), decreased with decreasing eGFR, while the concentrations of citrulline (R = –0.94) and isoleucine (R = –0.91) increased with decreasing eGFR (Table 3).

The citrulline/ornithine ratios (p = 0.01), as well as the concentrations of citrulline (p < 0.001) and isoleucine (p = 0.04), were significantly higher in the diabetic group compared to the non-diabetic group in stage 1 of CKD (Table 4). Similarly, the ornithine/arginine ratios (p = 0.02) and citrulline concentrations (p = 0.03) were significantly higher in the diabetic group compared to the non-diabetic group in stage 3. In stage 4, the tyrosine/phenylalanine ratios were significantly (p = 0.03) higher, while phenylalanine concentrations were lower in the diabetic group compared to the non-diabetic group.

Discussion

The quantitative method developed for the analysis of amino acids in urine was validated and found to be reliable, selective, sensitive, and accurate. We could relate the alterations in urinary amino acid ratios at different stages of CKD to the eGFR and some clinical characteristics associated with the disease. The analysis showed that the amino acid ratios of ornithine/arginine, tyrosine/phenylalanine and arginine/citrulline are affected by the progression of diabetes and can be correlated to renal function. There were significant changes in amino acid ratios between the diabetic and non-diabetic patients in stage 1 of CKD, and this could be utilised to predict the onset of CKD.

The accuracy of the analytical method used to obtain the LLOQ was determined to be within the acceptable range for all the amino acids analysed. The linearity (R^2) ranged from 0.983 to 0.999, the percentage recovery ranged from 90% to 130%, and the precision (%RSD) was less than 30%, except for ornithine. The method was also selective and sensitive as shown by the lowest limit of detection and LLOQ, and all the analytes of interest were well separated in the liquid chromatography and accurately identified by the mass spectrometer. Hence, the method is reliable and can be used to quantify the amino acids of interest within the analytical range of 0.01 mg/L to 5 mg/L. In diabetics patients, the concentrations of citrulline, isoleucine, phenylalanine and ornithine correlated positively with HbA1c

TABLE 2: Correlation between the amino acid concentrations and ratios and clinical data of patients with diabetes, Johannesburg, South Africa, November 2019 to January 2021.

Target analyte	HbA1c (%)		A/C (mg/mmol)		Albumin (mg/L)		UCr (mmol/L)		SCr (mg/dL)		Age (years)		eGFR (mL/min per 1.73 m^2)	
	R	*p*	*R*	*p*	*R*	*p*	*R*	*p*	*R*	*p*	*R*	*p*	*R*	*p*
Arg	0.60	0.40	0.16	0.84	−0.68	0.32	−0.88	0.12	−0.07	0.93	−0.67	0.33	0.56	0.45
Cit	0.76	0.24	0.34	0.66	−0.28	0.72	−0.94	0.06	0.16	0.84	−0.65	0.35	0.33	0.67
Ile	0.83	0.17	−0.53	0.47	−0.39	0.61	−0.84	0.16	−0.69	0.31	−0.98	0.02	0.94	0.06
Orn	0.86	0.14	−0.28	0.72	−0.44	0.56	−0.95	0.05	−0.46	0.54	−0.95	0.05	0.83	0.17
Phe	0.54	0.46	−0.75	0.25	−0.53	0.47	−0.55	0.45	−0.88	0.12	−0.83	0.17	1.00	0.004
Tyr	0.53	0.47	−0.48	0.52	−0.74	0.26	−0.68	0.32	−0.67	0.33	−0.80	0.20	0.94	0.06
Cit/Orn	0.46	0.54	0.75	0.25	0.06	0.94	−0.61	0.39	0.66	0.34	−0.17	0.83	−0.25	0.75
Arg/Cit	−0.04	0.96	−0.90	0.10	−0.37	0.63	0.12	0.88	−0.92	0.08	−0.32	0.68	0.71	0.29
Orn/Arg	−0.12	0.88	0.48	0.52	0.88	0.12	0.32	0.68	0.66	0.34	0.49	0.51	−0.81	0.19
Tyr/Phe	−0.40	0.60	0.67	0.33	0.69	0.31	0.49	0.51	0.83	0.17	0.73	0.27	−0.96	0.04

Note: Pearson's correlation was used to calculate R and p-values. $p > 0.05$, not significant; $p \leq 0.05$, significant.
A/C, albumin creatinine ratio; Arg, arginine; Cit, citrulline; eGFR, estimated glomerular filtration rate; HbA1c, haemoglobin A1c; Ile, isoleucine; Orn, ornithine; Phe, phenylalanine; R, correlation coefficient; SCr, serum creatinine; Tyr, tyrosine; UCr, urinary creatinine.

TABLE 3: Correlation between the amino acid concentrations and ratios and clinical data of non-diabetic patients, Johannesburg, South Africa, November 2019 to January 2021.

Target analyte	HbA1c (%)		A/C (mg/mmol)		ALB (mg/L)		UCr (mmol/L)		SCr (mg/dL)		Age (years)		eGFR (mL/min per 1.73 m^2)	
	R	*p*	*R*	*p*	*R*	*p*	*R*	*p*	*R*	*p*	*R*	*p*	*R*	*p*
Arg	0.69	0.52	−0.31	0.80	−0.09	0.94	0.00	1.00	−0.16	0.90	0.48	0.68	−0.13	0.92
Cit	0.56	0.62	−0.87	0.33	−0.96	0.19	−0.98	0.13	1.00	0.04	0.76	0.45	−0.94	0.22
Ile	0.98	0.14	−0.97	0.15	−0.90	0.29	−0.85	0.35	0.76	0.45	1.00	0.03	−0.91	0.27
Orn	−0.34	0.78	0.71	0.49	0.85	0.35	0.89	0.30	−0.95	0.20	−0.57	0.61	0.83	0.38
Phe	−0.22	0.86	−0.23	0.86	−0.44	0.71	−0.52	0.66	0.64	0.56	0.04	0.97	−0.40	0.74
Tyr	−0.22	0.86	0.62	0.57	0.78	0.43	0.83	0.37	−0.91	0.27	−0.47	0.69	0.76	0.46
Cit/Orn	0.02	0.99	−0.45	0.70	−0.64	0.56	−0.71	0.50	0.81	0.40	0.28	0.82	−0.61	0.58
Arg/Cit	0.06	0.96	0.38	0.75	0.57	0.61	0.64	0.55	−0.75	0.46	−0.20	0.87	0.54	0.64
Orn/Arg	−0.20	0.87	−0.25	0.84	−0.46	0.70	−0.53	0.64	0.66	0.54	0.06	096	−0.42	0.73
Tyr/Phe	−0.38	0.76	0.74	0.47	0.87	0.33	0.91	0.27	−0.96	0.17	−0.61	0.59	0.85	0.35

Note: Pearson's correlation was used to calculate R and p-values. $p > 0.05$, not significant; $p \leq 0.05$, significant.
A/C, albumin creatinine ratio; Arg, arginine; Cit, citrulline; eGFR, estimated glomerular filtration rate; HbA1c, haemoglobin A1c; Ile, isoleucine; Orn, ornithine; Phe, phenylalanine; R, correlation coefficient; SCr, serum creatinine; Tyr, tyrosine; UCr, urinary creatinine.

([R = 0.76; p = 0.24)], [R = 0.83; p = 0.17], [R = 0.54; p = 0.46], [R = 0.86; p = 0.41]) and eGFR ([R = 0.33; p = 0.67], [R = 0.94, p = 0.06], [R = 1.00; p = 0.004], [R = 0.83; p = 0.17]). The concentrations of these amino acids decreased with eGFR decline, meaning that the disease and renal function may alter these amino acids. In non-diabetic patients, arginine, citrulline and isoleucine correlated positively with HbA1c, meaning that these amino acid concentrations fluctuate with glucose levels. Even though citrulline showed a positive correlation with HbA1c in both the diabetic and non-diabetic groups, its concentrations levels were significantly ($p < 0.001$) higher in the diabetic group.

A study conducted in France in 2014 reported that citrulline concentration and citrulline/ornithine ratio increased while the tyrosine/phenylalanine ratio decreased (in plasma samples) in the late stages of CKD.[17] Similarly, in this study, we observed (in urine samples) an increase in citrulline concentrations and citrulline/ornithine ratios in stage 1; however, tyrosine/phenylalanine ratios increased from stage 1 to stage 4 of CKD with decreasing concentrations of phenylalanine. Amino acid concentrations have been reported to decrease in diabetic groups, and this decrease is attributed to the conversion of these amino acids to glucose.[18] Our findings show that phenylalanine follows this trend, as its concentrations decreased from stage 1 to stage 4 of CKD. The observed trend can be explained by looking at the metabolic pathways of the studied amino acids that occur in the liver and kidneys. Arginine is converted to ornithine, ornithine to citrulline and citrulline to arginine, while phenylalanine is converted to tyrosine.[17] Thus, the declining kidney function may affect the proper conversion of citrulline to arginine, as shown also by the significant (p= 0.05) decrease in the ratio of arginine/citrulline in stage 3 of CKD. Citrulline is a non-essential amino acid that is mainly synthesised in the gut and then taken up by the kidneys,

wherein it Is converted to arginine.[19] The increased concentration of citrulline in diabetic patients could be attributed to the inability of the kidneys to carry out this function.

The kidneys convert 50% of the total body phenylalanine to tyrosine through hydroxylation by the action of the phenylalanine-4-hydroxylase enzyme.[19] At stage 4 of CKD, the tyrosine/phenylalanine ratios were significantly different between the diabetic and non-diabetic groups, and the phenylalanine concentrations were lower in the diabetic group. The metabolism of phenylalanine has two outcomes – the production of tyrosine, and the complete oxidation of phenylalanine to carbon dioxide and water to drive the production of glucose.[20] We observed an increase in tyrosine/phenylalanine ratios with eGFR decline, which suggests an increased production of tyrosine in the kidneys as renal function decreases and could explain the low phenylalanine concentrations in stage 4 of CKD.

Limitations

The exact cut-off values for amino acid concentrations and ratios, which indicate the onset of diabetic nephropathy or CKD, could not be determined. Also, the sample size for non-diabetic patients was below 30 (n = 22), which is the acceptable size for accurate statistical evaluation, and this may have influenced the results of comparisons between the two groups of patients. A more detailed statistical evaluation would have to be conducted to indicate early CKD using amino acid ratio.

Conclusion

The amino acid concentrations and ratios in human urine samples could provide valuable information about the onset and progression of diabetic nephropathy as there appeared to be significant correlations between the amino acid concentrations and declining eGFR. Amino acid ratios (ornithine/arginine and tyrosine/phenylalanine) are affected by decreased renal function, and citrulline/ornithine ratios may predict early CKD.

Acknowledgements

The authors would like to thank the Chris Hani Baragwanath Academic Hospital Chemical Pathology staff for the collection and transportation of samples, and the National Institute of Occupational Health staff for long-term storage, sorting, and transportation of samples. We also thank the Witwatersrand mass spectrometry team for granting access to the analytical instruments. This article is partially based on the second author's thesis[21] of the degree of Doctor of Philosophy in Division of Human Genetics, Department of Pathology, Faculty of Health Sciences at the University of Cape Town, South Africa, with supervisor Prof. Collet Dandara, received September 2018, available here: http://hdl.handle.net/11427/30132.

Competing interests

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

Authors' contributions

D.M.T. generated the hypothesis and planned the patient recruitment and sample collection logistics. T.M. ran the project, analysed the samples, carried out the statistical analysis and wrote the manuscript. M.M.M. supervised the project and provided expert analytical interpretation. R.J.M. assisted with the analytical processes. S.B. assisted with patient recruitment.

Sources of support

The work was supported by the National Research Foundation, under the Thuthuka Award (Gun 117975).

Data availability

Data sharing is not applicable to this article.

Disclaimer

The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of any affiliated agency of the authors.

References

1. Brenneman J, Hill J, Pullen S. Emerging therapeutics for the treatment of diabetic nephropathy. Bioorg Med Chem Lett. 2016;26(18):4394–4402. https://doi.org/10.1016/j.bmcl.2016.07.079

2. Kanwar YS, Wada J, Sun L, et al. Diabetic nephropathy: Mechanisms of renal disease progression. Exp Bio Med. 2008;233(1):4–11. https://doi.org/ 10.3181/ 0705-MR-134

3. Moosa MR, Van Der Walt I, Naicker S, Meyers AM. Important causes of chronic kidney disease in South Africa. South African Med J. 2015;105(4):1–8. https://doi.org/10.7196/SAMJ.9535

4. Adeniyi AB, Laurence CE, Volmink JA, Davids MR. Prevalence of chronic kidney disease and association with cardiovascular risk factors among teachers in Cape Town, South Africa. Clin Kidney J. 2017;10(3):1–7. https://doi.org/10.1093/ckj/sfw138

5. Motta LA, Shephard MDS, Brink J, Lawson S, Rheeder P. Point-of-care testing improves diabetes management in a primary care clinic in South Africa. Prim Care Diabetes. 2017;11(3):248–253. https://doi.org/10.1016/j.pcd.2016.09.008

6. Alhajim SA, Sherhan AA. Albumin to creatinine ratio Is fairly correlated with albumin excretion rate in chronic kidney disease. Am J Lab Med. 2018;3(6):48–52.

7. Colhoun HM, Marcovecchio ML. Biomarkers of diabetic kidney disease. Diabetologia. 2018;61(5):996–1011. https://doi.org/10.1007/s00125-018-4567-5

8. Lopez-Giacoman S. Biomarkers in chronic kidney disease, from kidney function to kidney damage. World J Nephrol. 2015;4(1):57–73. https://doi.org/10.5527/wjn.v4.i1.57

9. Canal C, Pellicer R, Rocha CI, et al. Tables for estimating glomerular filtration rate from plasma creatinine. Nefrologia. 2008;28(3):317–324.

10. Currie G. Biomarkers in diabetic nephropathy: Present and future. World J Diabetes. 2014;5(6):763–776. https://doi.org/10.4239/wjd.v5.i6.763

11. Le A, Ng A, Kwan T, Cusmano-Ozog K, Cowan TM. A rapid, sensitive method for quantitative analysis of underivatized amino acids by liquid chromatography-tandem mass spectrometry (LC-MS/MS). J Chromatogr B Anal Technol Biomed Life Sci. 2014;944:166–174. https://doi.org/10.1016/j.jchromb.2013.11.017

12. Kopple JD. Phenylalanine and tyrosine metabolism in chronic kidney failure. J Nutr. 2007;137(6 Suppl 1):1586S–1590S. https://doi.org/10.1093/jn/137.6.1586S

13. Kemperman RFJ, Horvatovich PL, Hoekman B, Reijmers TH, Muskiet FAJ, Bischoff R. Comparative urine analysis by liquid chromatography – Mass spectrometry and multivariate statistics: Method development, evaluation, and application to proteinuria. J Proteome. 2007;6(1):194–206. https://doi.org/10.1021/pr060362r

14. Kalantari S, Jafari A, Moradpoor R, Ghasemi E, Khalkhal E. Human urine proteomics: Analytical techniques and clinical applications in renal diseases. Int J Proteomics. 2015;2015:782798. https://doi.org/10.1155/2015/782798

15. Tang X, Gu Y, Nie J, Fan S, Wang C. Quantification of amino acids in rat urine by solid-phase extraction and liquid chromatography/electrospray tandem mass spectrometry: Application to radiation injury rat model. J Liq Chromatogr Relat Technol. 2014;37(7):951–973. https://doi.org/10.1080/10826076.2013.765451

16. Waterval WAH, Scheijen JLJM, Ortmans-Ploemen MMJC, Habets-van der Poel CD, Bierau J. Quantitative UPLC-MS/MS analysis of underivatised amino acids in body fluids is a reliable tool for the diagnosis and follow-up of patients with inborn errors of metabolism. Clin Chim Acta. 2009;407(1–2):36–42. https://doi.org/10.1016/j.cca.2009.06.023

17. Duranton F, Lundin U, Gayrard N, et al. Plasma and urinary amino acid metabolomic profiling in patients with different levels of kidney function. Clin J Am Soc Nephrol. 2014;9(1):37–45. https://doi.org/10.2215/CJN.06000613

18. Durá-Travé T, Gallinas-Victoriano F, Cortes-Castel E, Moya-Benavent M. Amino acid plasma concentrations and urinary excretion in young diabetics. In: Ahmed AG, editor. Diabetes and its complications. London: IntechOpen, 2018; p. 109–118. https://doi.org/10.5772/intechopen.72080

19. Van de Poll MGC, Soeters PB, Deutz NEP, Fearon KCH, Dejong CHC. Renal metabolism of amino acids: Its role in interorgan amino acid exchange. Am J Clin Nutr. 2004;79(1):185–197. https://doi.org/10.1093/ajcn/79.2.185

20. Williams RA, Mamotte CDS, Burnett JR. Phenylketonuria: An inborn error of phenylalanine metabolism. Clin Biochem Rev. 2008;29(1):31–41.

21. Tanyanyiwa DM. Type 2 diabetes mellitus and dyslipidaemia: Effects of genetic variation in African populations [homepage on the Internet]. Cape Town: University of Cape Town; 2018 [cited 2021 Feb 24]. Available from: http://hdl.handle.net/11427/30132

Chapter 3 Chemicals and materials

This chapter presents the reagents, conditions and materials used in the study. Detailed information about the method development and validation criteria are presented in chapter 4 and in the published article.

3.1 Chemicals and materials

All chemicals were purchased from Sigma-Aldrich (Johannesburg, South Africa). The following amino acid standards (purity ≥ 98.0%) were used in the preparation of calibration curves for arginine (Arg), citrulline (Cit), isoleucine (Ile), ornithine (Orn), phenylalanine (Phe), and tyrosine (Tyr). Formic acid was used as an ion-pairing agent and also in the preparation of the mass spectrometry calibrant solution (purity ≥ 98.0%). Sigmatrix Urine Diluent [blank urine (BU)] (purity ≥ 98.0%) was used in the analysis of the matrix effects and thereafter used for the preparation of quality control samples and preparation of calibration standards. HPLC-grade solvents, (purity ≥ 99.9%) namely acetonitrile, methanol, and isopropanol were used as an eluent solvent, a wash solvent, and as a calibration solution solvent respectively. Deionised water, which was used as a solvent for dilution and elution of analytes and as a calibration solution, was purified using a Direct-Q® 3 UV Millipore system (Molsheim, France).

3.2 Ethical considerations

This research used human subjects therefore it needed ethical clearance. This was obtained from the Human and Animal Rights Chris Hani Baragwanath Academic Hospital Medical Advisory Committee and the University of the Witwatersrand Human Research Ethics Committee (reference number: M180798). Participants were informed about the research study, what it entails as well as its significance in the health care system in general. Those who were interested in participating gave a written consent to allow for the use of their urine samples and medical records. Sample codes were to preserve the anonymity and confidentiality of the results obtained in this study.

3.3 Data analysis

All data processing was done on the Bruker QuantAnalysis software (Bruker Daltonics, Bremen, Germany) to determine the concentration of amino acids. The dilution factor of 10 was used to determine the actual concentration of amino acids in the urine samples.

Statistical data interpretation was carried out in Microsoft Excel version 2101 (Microsoft Corporation, Redmond, Washington, United States). Mean and standard deviation values were obtained using descriptive statistics, significant differences in variables between diabetic and non-diabetic patients were assessed using analysis of variance, and linear correlations between biochemical and analytical data were determined using Pearson's correlation coefficient (R). A p-value 0.05 or less was considered significant.

3.4 Calibration and QC standards preparation

3.4.1 Stock solutions

The following stock solutions were prepared, **Stock solution 1** for calibration curves: 100 mg/L of each amino acid. **Preparation of Stock solution 1**: 10 mg of each amino acid was weighed on a balance (KERN $_{ADB}$), and then dissolved in a final volume of 100 mL Milliq H_2O, in a 100 mL volumetric flask. Then a 10 mg/L and 1.00 mg/L stock solution were prepared.

Stock solution 2 for quality control standards (QC's): containing a mixture of 6 amino acids, Arginine (Arg), Phenylalanine (Phe), Tyrosine (Tyr), Citrulline (Cit), Ornithine (Orn), and Isoleucine (Ile).

Preparation of Stock solution 2: 10 mg of each amino acid was weighed and dissolved in 100 mL Milliq H_2O, in a 100 mL volumetric flask to make a 100 mg/L of stock solution 2. Then a 10 mg/L and 1.00 mg/L stock solution were prepared.

3.4.2 Calibration standards

The following calibration points [(0.01), (0.05), (0.10), (0.30), (0.50), (0.80), (1.00), (3.00), (4.00), and (5.00) mg/L] were prepared by spiking the H_2O or the blank urine (BU) matrix with an appropriate amount of the **stock solution 1** to make a final volume of 1.00 mL in UHPLC vials.

3.4.3 QC standards

Quality control standards (QC's) were prepared using **stock solution 2,** the following concentration were prepared to evaluate the behaviour of the method at lowest limit of quantitation (LLOQ1) [0.10 mg/L], (LLOQ2) [0.30 mg/L], 3x the limit of quantification (LLOQ3) [0.50 mg/L], (MLOQ) [3.00 mg/L], and upper limit of quantitation (ULOQ) [4.00 mg/L]. The QCs were further used to test for accuracy, precision, and reproducibility.

3.5 UHPLC and URHMS conditions

The developed mass spectrometry method was designed such that it is capable of doing both qualitative and quantitative analysis in a single run. The quantitative analysis part of the method was developed by running a 1 mg/L AA standard using a multiple reaction monitoring (MRM) method through direct infusion into the QTOF-HRMS. The MRM method parameters are shown in Table 3.1, the collision energy for each amino acid was varied starting from 6.00 eV and increased until a stable fragment pattern was observed. The syringe pump (KdScientific), flow rate was set at 180 µL/hrs. The instrument was run in positive ion mode with a scan range from 50-1300 mass to charge ratio (m/z). To incorporate the qualitative analysis method a row in the Hyster software acquisition table below (Table 3.1) was added with these following conditions isolation: mass 50 m/z, isolation width 0 m/z, acquisition rate 6 spec/sec, collision energy 6 eV. This enables the MRM method to run targeted and untargeted analysis in a single run.

Table 3. 1: MRM conditions used to generate fragment ions of each targeted amino acid (parent ion).

Target compound	Isolation mass (m/z)	Collision energy (eV)	fragment ions (m/z)	Mass range (m/z)	Isolation width (m/z)	Acquisition rate (spec/s)	Acquisition voltage (V)
Arg	175.0000	26.00	70, 116, 130	50-1300	± 1.00	6.00	2200
Cit	176.0000	14.00	70, 113, 159				
Ile	132.0000	22.00	69 ,86				
Orn	133.0000	12.00	70, 116,				
Phe	166.0000	15.00	77, 91, 120				
Tyr	182.0000	12.00	77, 91 95, 107, 119, 123, 136, 147, 165				
All compounds	50.0000	6.00			± 0.00		

Chapter 4 Method development and validation

4.1 Introduction

This chapter deals with method development and validation criteria followed in the quantification of target amino acids (arginine, citrulline, isoleucine, ornithine, phenylalanine, and tyrosine). These amino acids were quantified in the urine samples, of patients with diabetes, and without diabetes. Their link to diabetic nephropathy as potential biomarkers was resented in chapter 2 (published article) as well as the discussion and conclusion.

4.1.1 Method validation

Method development and validation play a key role in ensuring reliable analytical data, in routine laboratory work as well as clinical and forensic scientific studies. Unreliable data could lead to incorrect conclusions that can easily be carried out within a field of research or scientific community resulting in correct decisions about patient treatment. The quality of the data mainly depends on good analytical methods and adequate validation of the methods which ensures that the method is fit for the intended purpose [1]. The choice of method validation criteria to be tested depends on the analytical method to be used as well as the type of analysis, whether it is an identification or quantification method [2]. The handling of data is important, and this requires the use of statistical packages, such as SPSS, Minitab, and the guidelines (International Conference on Harmonization (ICH)) to be followed are to be known beforehand to ensure reliability [2].

4.1.2 Method validation parameters

In clinical and forensic studies when evaluating metabolites in biological matrices, these validation parameters should be determined: linearity, selectivity, precision, accuracy, recoveries, and the lowest limit of quantification. The analytical steps of each method being validated should be well understood because each analytical method has its own validation parameters that need to be determined and have its own limits (acceptable criteria). Furthermore, decisions on how to carry out the

validation of the parameters and which statistical test to use has to be made [3]. When working with biological samples and using separation instruments such as ultra-high-performance liquid chromatography (UHPLC) or gas chromatography (GC), with a mass spectrometer (MS) as a detector, it is essential to validate the method to ensure accurate continuous sample analysis [4].

4.1.3 Linearity

Linearity is defined as the ability of a method to show a direct proportionality of the analyte concentration to signal response in a measurable range. Linearity is tested between the instrument's response versus the analyte concentration using R^2 (coefficient of determination) which evaluates how the line of best fit drawn through a set of data points represents the association between dependent and independent variables. The closer R^2 is to one ($R^2 \approx 1$), the stronger the linear relationship [5].

4.1.4 Limits of detection and quantitation

The limit of detection (LOD) is the minimum concentration of an analyte that can be reliably detected under the experimental condition used. The limit of quantification (LOQ) is the minimum quantifiable concentration of a substance that can be determined with acceptable accuracy and precision. Both LOD and LOQ are evaluated by either visualisation, signal-to-noise ratio (s/n), standard deviation of the response (δ), and standard deviation of the slope of linearity (S) [6].

$$LOD = 3\,\delta/S \quad (1)$$

$$LOQ = 10\,\delta/S \quad (2)$$

4.1.5 Accuracy

Accuracy is the closeness of the experimental value to the true value of a measurement or reference value and is obtained by repeated measurements. Accuracy plays an important role in revealing whether a method is fit for the intended purpose and should be assessed in all analytical methods. To test for accuracy, quality control samples are used, this can be achieved by spiking with known

concentrations of analyte, it is reported as a percentage and is evaluated within (same day replicates) and between runs (at least two deferent days). The measured value should be within 15% of the nominal or reference value [7].

4.1.6 Precision

Precision refers to the repeatability of measurement in an analytical method, it is used to evaluate calibration curve performance characteristics such as the LLOQ, MLOQ and ULOQ. it should also be determined on same day (inter-precision) and also on a different day (intra-precision) [8]. For the LLOQ 20% from the nominal value is acceptable while other parameters should give 15%. It is measured as the coefficient of variation CV% [9].

4.1.7 Selectivity

Selectivity is defined as the ability for a method to identify an analyte or a group of compounds in a mixture or in the presence of interferences, a method with a perfect selectivity (100%) is deemed specific or has specificity for the analytes of interest [9].

4.1.8 Matrix effects

Matrix effect refers to the interferences that may occur when measuring an analyte concentration in a biological matrix such as plasma or urine, therefore it’s important to determine this effect when working with biological samples. If a suitable matrix reference standard is not available, then synthetic matrices can be used. The matrix is spiked with known concentrations of the analyte usually three times above the LLOQ and close to the ULOQ. To ensure that the matrix matched blank gives accurate recoveries, about six samples from donors are needed to carry out this test. It is a necessity when using mass spectrometric analysis [3].

4.1.9 Carryover

Carryover is assessed to ensure that there is no enhancement of concentration from a previous injection, it is determined after a high concentration standard by running a blank sample, and the accepted value is less or equal to 20% of the LLOQ [10].

4.2 Experimental

4.2.1 Method development

4.2.1.1 UHPLC separation method

A standard mixture with a concentration of 1.00 mg/L of target amino acids (Arg, Cit, Ile, Orn, Phe, and Tyr) in H_2O was run through a C18 column on a UHPLC system. This was used to develop the UHPLC separation method of these amino acids. The separation was optimised by changing the solvent ratio to yield baseline separation. The standard mixture was run in a 10-minute UHPLC gradient separation method as follows. A two-buffer system was employed, utilizing formic acid as the ion-pairing agent; solvent A consisted of 0.1% formic acid in H_2O volume per volume (v/v), and solvent B comprised 0.1% formic acid in acetonitrile volume per volume (v/v). The buffers were sonicated for 5 minutes to remove bubbles, then purging of the UHPLC system was done before analysis. The standard was placed in an autosampler set at 5 °C and 20 µL of the standard was injected into the C18 UHPLC column and eluted using a gradient elution technique, starting from an H_2O: acetonitrile ratio of 95:5% to 5:95% in 10 min at a flow rate of 0.4 mL/min.

4.2.1.2 Matrix effects (ME)

To determine matrix effects, the above procedure was repeated: A standard mixture with a concentration of 1.00 mg/L of target amino acids (Arg, Cit, Ile, Orn, Phe, and Tyr) in blank urine (BU) was ran through a C18 column on a UHPLC system. Matrix effects were evaluated to determine the best-suited solvent (Blank Urine (BU) vs H_2O) for the calibration curve to be used in the quantification of target amino acids.

This was done by comparing the two chromatograms [(1.00 mg/L H_2O) and (1.00 mg/L BU)].

The matrix effects were further investigated by determining the matrix effect factor (MEF). This was done by spiking H_2O, blank urine (BU) matrix and urine sample (U19) with (0.10, 0.50, and 4.00 mg/L) AA standard mixture respectively. This was done to see whether the matrix effect between H_2O and urine was large enough and also if blank urine can be used to contract calibration curves that would yield accurate data, when quantifying the target amino acids in urine.

4.2.1.3 Sample preparation methods

Method development was conducted to determine the sample preparation technique that will yield accurate results (determined by % recoveries). This is done to ensure that analyte losses during the sample preparation (clean-up) are accounted for. Two methods (method A and method B) were compared after noting that spiked urine samples gave very low % recoveries when using method A. In method A 6% sulfosalicylic acid (6% SSA) was used to clean up (protein precipitation) the urine samples before analysis. This method was adapted from literature [11] and modified to suite laboratory settings. Method A was compared to method B which does not involve the precipitation of proteins with 6% SSA. Method B gave better recoveries for all the amino acids and was then validated to ensure that it was fit for its purpose (accurate quantification of target amino acids). Method validation guidelines were adopted from the literature (European Medicines Agency) [12] and modified to suit the laboratory settings.

4.2.1.4 Method A with 6% sulfosalicylic acid

In method A, a precipitating agent, 6% sulfosalicylic acid was used to precipitate proteins and other macromolecules. Two QCs (1 mL) (blank urine and urine (U19)) were spiked with (1 mL) of 1.00 mg/L AA standard mixture in a 2 mL Eppendorf respectively. The mixture was then centrifuged at 3000 rpm, for 10 minutes, then 1 mL of the supernatant was mixed with 1 mL of 6% sulfosalicylic acid in a 2 mL Eppendorf tube. The mixture was allowed to stand for 5 minutes and then

centrifuged at 3000 rpm for 30 minutes. Thereafter 200 µL of the supernatant was transferred into a 2 mL HPLC analysis vials and ran on the UHPLC-HRMS.

4.2.1.5 Method B without 6% sulfosalicylic acid

In this method, the precipitation agent (6% sulfosalicylic acid) was not the added. Two QCs (1.00 mL) (blank urine and urine (U19)) were spiked with (1.00 mL) of 1.00 mg/L AA standard mixture in a 2.00 mL Eppendorf respectively. The mixture was then centrifuged at 3000 rpm, for 10 minutes. Thereafter 200 µL of the supernatant was transferred into a 2.00 mL HPLC analysis vials and ran on the UHPLC-HRMS.

The recovery of 1.00 mg/L for BU QC's and U19 QC's showed that the use of the precipitating agent (6% sulfosalicylic acid) resulted in peak suppression which led to poor recoveries for all the target amino acids. Better recoveries were obtained with method B when the acid is omitted.

4.2.2 Method validation

In this section, method B was evaluated for its analytical performance characteristics, and also assessed for its validity. Method validation parameters such as calibration performance, accuracy, precision, selectivity matrix effects, carry over, repeatability, reproducibility, uncertainty of measurement, were assessed. The calibration performance was assessed by linearity (R^2), LLOD, LLOQ, ULOQ, the acceptable criteria for calibration was $R^2 \geq 0.995$.

The acceptable criteria for accuracy and precision are as follows: a recovery of 80-120% and % relative standard deviation (%RSD) ≤ 10%; a matrix effect and carry over of ≤ 20% for the lowest limit of quantification, and a 15% for the mid-point or upper limit of quantification. Method validation parameters were investigated on different days (inter accuracy/precision), and also in between runs (intra accuracy/precision).

4.2.2.1 Calibration performance

Calibration performance was evaluated in this calibration range (0.01 to 5.00 mg/L) using the following calibration points [(0.01), (0.05), (0.10), (0.30), (0.50), (0.80), (1.00), (3.00), (4.00), and (5.00) mg/L] in H_2O:blank urine matrix (1:5).

4.2.2.2 Linearity (R2)

Linearity was calculated using linear regression analysis on excel.

Correlation coefficient (R^2).

$$R^2 = \frac{\sum[(x_i - x_m)(y_i - y_m)]}{\sqrt{[\sum[(x_i - x_m)^2 (y_i - y_m)^2]}} \quad (3)$$

Where x_i is the individual data points in the independent variable (concentration).

Where x_m is the average of data points in the independent variable (concentration).

Where y_i is the individual data points in the dependent variable (area).

Where y_m is the average data points in the independent variable (area).

4.2.2.3 LLOD and LLOQ

Lowest limit of detection (LLOD).

$$LLOD = b + 3\delta_m \quad (4)$$

Where b is the y intercept.

Where $3\delta_m$ is three times the standard deviation and m is slope.

Lowest limit of quantification (LLOQ).

$$LLOQ = b + 10\delta_m \quad (5)$$

Where b is the y intercept.

Where $10\delta_m$ is ten times the standard deviation and m is slope.

4.2.2.4 Accuracy

Analyte concentration recoveries were used to determine accuracy, three QCs were used, a known concentration of the AA mix standard was spiked LLOQ1 [0.10 mg/L], LLOQ3 [0.50 mg/L] and ULOQ [4.00 mg/L] into H_2O/blank urine matrix, and on a urine sample. The spiked samples were then run through sample preparation using method B, and the recoveries were determined, the samples were prepared in triplicate and each injected three times to give nine replicates (n=9), the experiment was repeated on three different days.

4.2.2.5 Recoveries

Recoveries are concentrations recovered after running a QC sample (known concentration spike) through an extraction or sample preparation method. They were determined by subtracting the final concentration from a blank or un-spiked sample and dividing by the spiked concentration (known concentration) or true value.

$$Recovery = \frac{final\ concentration - intial\ concentration}{spiked\ concentration} \quad (6)$$

An acceptable percentage recovery range between 80-120%.

$$\%Recovery = \frac{final\ concentration - intial\ concentration}{spiked\ concentration} * 100\% \quad (7)$$

4.2.2.6 Precision

Analyte concentration recoveries were used to determine precision, three QCs were used, a known concentration of the AA mix standard was spiked LLOQ1 [0.10 mg/L], LLOQ3 [0.50 mg/L] and ULOQ [4.00 mg/L] into H_2O, H_2O/blank urine matrix, and on a urine sample. The spiked samples were then run through sample preparation using method 3, and the recoveries were determined, the samples were prepared in triplicate and each injected three times to give nine replicates (n=9), the experiment was repeated on two different days.

4.2.2.7 Relative standard deviation

Relative standard deviation was used to calculate precision, a percentage relative standard deviation (%RSD) was calculated to see how close each replicate experiment is to each other and also to check instrument performance with replicate injections. A good %RSD is below 10% (≤ 10%), the acceptable range may vary from (≤ 10% to ≤ 30%).

$$\%RSD = \frac{\delta}{\bar{x}} * 100\% \quad (8)$$

Where δ is relative standard deviation and $\bar{x}$ is the mean.

After the method was validated an analysis of 187 samples was carried out as shown below (Figure 4.1).

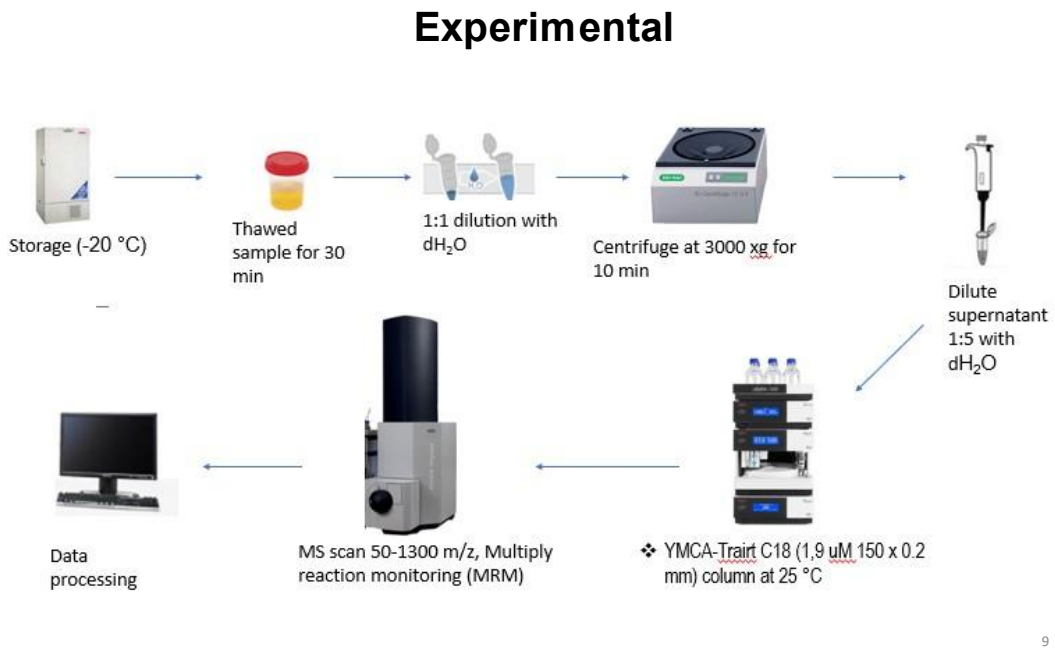


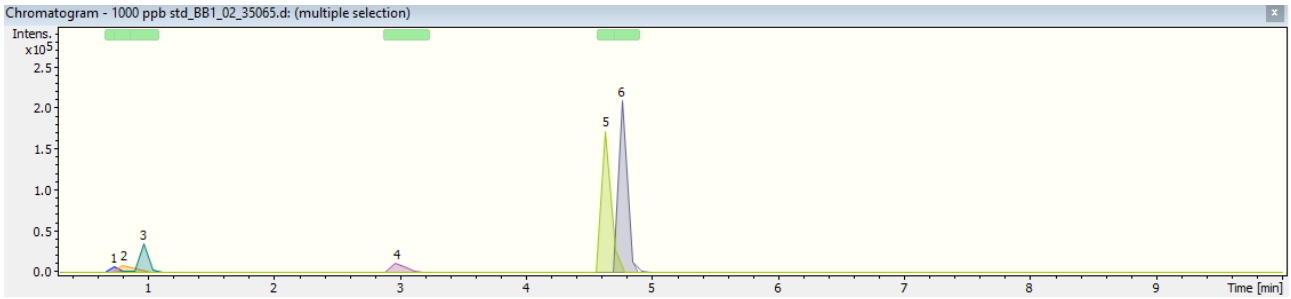
Figure 4. 1: Schematic diagram of the workflow for calibration, QC's standards and urine samples analysed on the UHPLC-ESI-QTOF-HRMS (method B).

4.3 Results and discussion

4.3.1 Method development

The analysis of a 1.00 mg/L AA standard mixture in H_2O on a C18 column using the solvent gradient from 5% ACN, to 95% ACN did not separate Phe and Tyr, Ile was well separated from other amino acids, while Arg, Cit and Orn were not retained and eluted before 1 minute (Figure 4.2 A). The graphical representation of the separation conditions is shown in Figure 4.2 B, as single-step gradient elution, constant solvent flow and column back pressure. Several gradient elution methods were investigated, and the chromatographic conditions that improved separation are shown in Figure 4.3. There was a baseline separation between Ile, Phe, and Tyr amino acids, while Arg, Cit, and Orn could not be well separated on the C18 column. Arg, Cit, and Orn are chemically similar, and this poses a challenge for C18 columns as they often struggle to separate similar molecules [13]. The poor retention observed for Arg, Cit and Orn could be due to starting the gradient method with a highly polar mobile phase ratio (5% ACN: 95%H_2O). These findings are similar to the ones reported by (Ridwan et al. 2018) where a polar mobile phase resulted in rapid elution (before 1 min) of Arg and Cit [14].

A



B

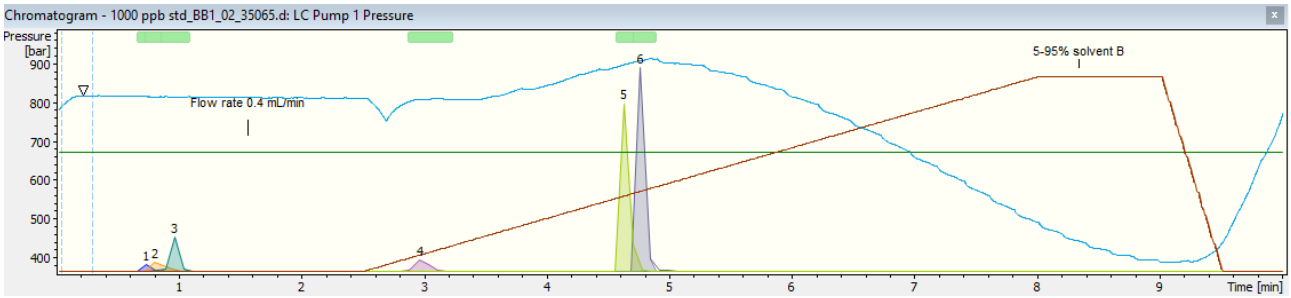


Figure 4. 2: The separation of target amino acids (1.00 mg/L standard in H_2O) on a C18 column. Chromatogram A (1=Orn, 2=Arg, 3=Cit, 4=Ile, 5=Tyr, 6=Phe), Chromatogram B [– blue line UHPLC pressure (bars) – red line multistep gradient elution 5% (0-2.5 min), 5-95% (2.5-8 min), 95% (8-9 min), 95-5% (9-9.5 min) 5% (9.5-10 min) of solvent B, – green line flow rate (0.40 mL/min)].

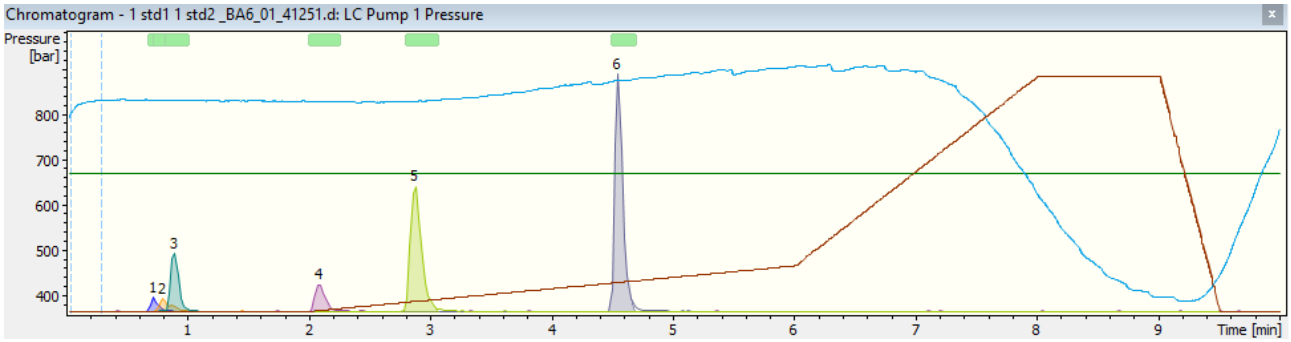


Figure 4. 3: The separation of target amino acids (1.00 mg/L standard in H_2O) on a C18 column. Chromatogram (1=Orn, 2=Arg, 3=Cit, 4=Ile, 5=Tyr, 6=Phe), UHPLC conditions, [– blue line UHPLC pressure (bars) – red line multistep gradient elution of 5% (0-2 min), 5-40% (2-6 min), 40-95% (6-8 min), 95% (8-9 min), 95-5% (9-9.5 min), 5% (9.5-10 min) of solvent B – green line flow rate (0.40 mL/min)].

4.3.1.1 Matrix effects

The developed separation method above (Figure 4.3) was further tested for matrix effect. The figure below (Figure 4.4), shows a visually assessment of any matrix effects on the retention and on the sensitivity of each of the analytes of interest. The non-matrix adjusted H_2O 5.00 mg/L QC standard showed good separation and high intensities (intens. $X10^6$), whereas the matrix adjusted BU 5.00 mg/L standard also showed good separation, but the intensities (intens. $X10^5$) of all compounds decreased. The causes of matrix effects are not fully elucidated but can be linked to coeluting peaks from solvent impurities or from competing analytes. They are believed to reduce the ionization efficiency of target compounds, and this is common in electrospray ionization [15].

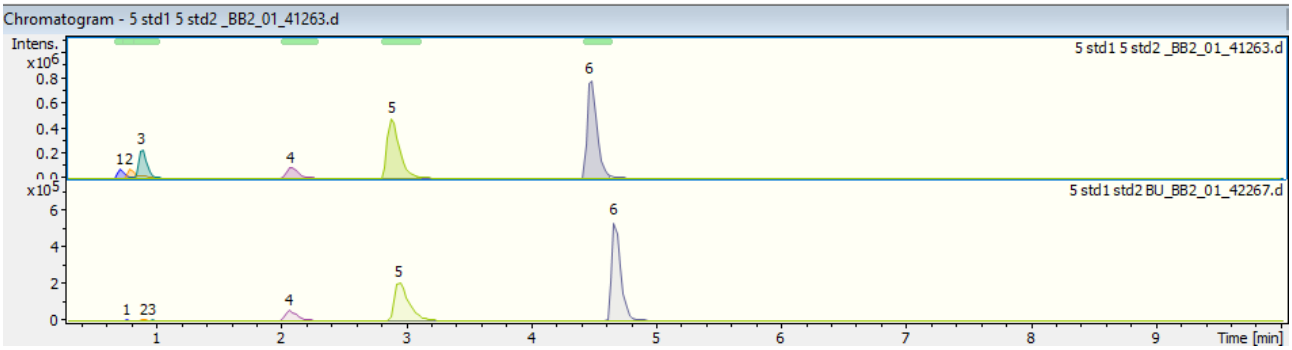


Figure 4. 4: Chromatogram A top, retention times and intensities of a 5.00 mg/L standard mixture in H_2O and chromatogram B bottom, retention times and intensities of 5.00 mg/L standard mixture in blank urine (BU) matrix.

The average peak areas presented in the appendix Sd Table 7.1 -7.16, for each amino acid at different concentrations were used to calculate the matrix effect factor (MEF) shown in Table 4.1 below. The matrix effects in BU, when compared to H_2O, show a peak suppression in all the analytes at three spiked concentrations. Similarly, when U19 (from patient 19) is compared to H_2O all the analytes were suppressed. This showed that matrix effects are playing a huge role in the sensitivity of all the analytes. When comparing BU and U19, surprisingly Arg, Cit, and Orn showed peak enhancement. This could be due to the fact that these analytes are not baseline separated in the C18 column and contribute some similar ions to each other's peak areas. The similar ions that were generated by the fragmentation of each analyte are

(70, and 116 m/z) (Table 3.1 methodology section). Coeluting matrix can result in two phenomena either peak suppression or peak enhancement, this is common in electrospray ionisation (ESI) and can be corrected by the use of internal standard [16].

Table 4. 1: Matrix effects determined by calculating the matrix effect factor (MEF) using average peak areas (mean).

Target analyte	MEF (H_2O-BU/H_2O) x 100%			MEF (H_2O-U19/H_2O) x 100%			MEF (BU-U19/BU) x 100%		
	0.10 mg/L	0.50 mg/L	4.00 mg/L	0.10 mg/L	0.50 mg/L	4.00 mg/L	0.10 mg/L	0.50 mg/L	4.00 mg/L
Arg	96.17	96.28	96.04	92.74	94.48	92.63	-89.52	-48.51	-86.27
Cit	99.29	99.17	98.58	98.06	97.89	95.44	-172.32	-153.85	-220.24
Ile	76.80	55.25	31.53	73.29	49.61	38.10	-15.15	-12.59	9.60
Orn	93.31	95.15	95.96	77.82	92.01	92.25	-231.29	-64.62	-91.98
Phe	59.09	58.01	42.48	85.67	62.27	49.07	64.98	10.16	11.46
Tyr	70.74	68.20	59.33	85.56	63.56	60.37	50.65	-14.57	2.55

The results from Table 4.1, show that there are considerable matrix effects in the urine and blank urine matrix (BU) than in water (H_2O), hence calibration curves will have to be matrix adjusted to ensure accurate data when quantifying analytes. The samples to be used in the calibrations and quality control were prepared in BU matrix to account for matrix effects that could alter the accuracy of the quantified concentrations from real urine samples. The acceptable range for matrix effects is 20% for spiked concentration near the lowest limit of quantification (LLOQ) and 15% for concentration in the mid-range of the calibration curve or in the upper limit of quantification (ULOQ) [12]. The value greater than zero indicates peak suppression and a value below zero shows peak enhancement while closer to zero means no matrix effect [15]. In this case, no matrix effect would be an MEF value below [(±20% of 0.10 mg/L) and (±15% of 0.50 and 4.00 mg/L)]. Only Ile had an acceptable matrix effect at all spiked concentrations (0.10 mg/L to 4.00 mg/L). Phe and Tyr showed acceptable matrix effects at (0.50 mg/L and 4.00 mg/L spiked concentration).

4.3.1.2 Sample preparation method optimization

The evaluation of sample preparation methods performances using method A and method B, is shown in Figure 4.5 below. In method A, 6% sulfosalicylic acid was added to precipitate proteins, this resulted in lower recoveries when compared to method B where the addition of the acid was omitted. The range considered for acceptable recoveries was from (80% to 120%) of the spiked concentration (1.00 mg/L). Low recoveries [below 80% (0.80 mg/L) of the spiked concentration (1.00 mg/L)] were obtained for all the amino acids (Arg, Cit, Ile, Orn, Phe, and Tyr). All amino acids showed significant improvements [recoveries within (80% to 120%) of (1.00 mg/L)] when the acid was not added in both matrices (BU and U33 (from patient 33)).

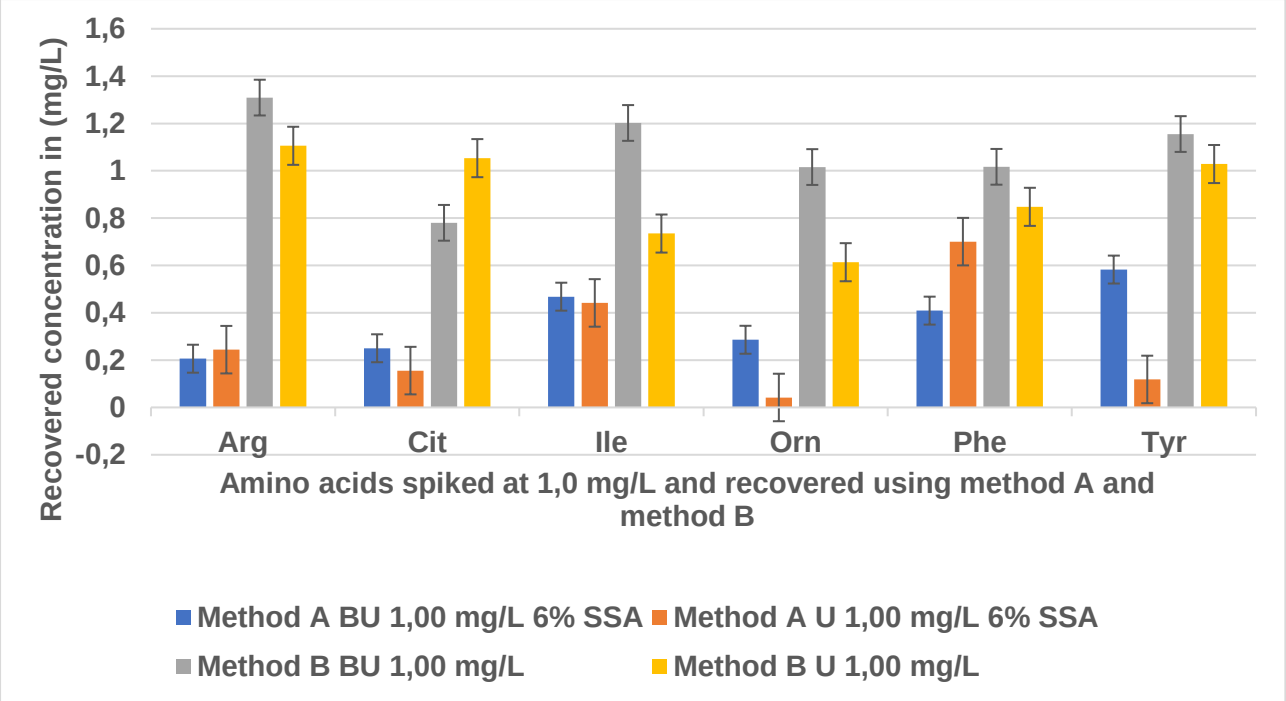


Figure 4. 5: Compares recovery of target amino acid in BU and U (urine sample U33) spiked with 1.00 mg/L concentrations using method A and method B.

4.3.2 Method validation

Method validation was carried out as shown in the upcoming sections. The accuracy and precision of the method were determined on the intra-day and on inter-day.

4.3.2.1 Intra-day (same day) accuracy (% recovery) and precision (%RSD)

A 10-point calibration curve of each amino acid was used to evaluate linearity (R^2), the lowest limit of detection (LLOD), and the lowest limit of quantification (LLOQ) as shown in Table 4.2 and Table 4.5. Calibration curves and regression analysis for day 1 (Sd Table 17-22) and day 2 (Sd Table 23-28) is given in the appendix chapter 7.

The linearity of the target amino acids ranged from 0.960 to 0.999, on day 1 of method validation test as shown in Table 4.2. The analytical concentration range was from 0.01 to 5.00 mg/L. The linearity for Isoleucine (Ile), phenylalanine (Phe), and Tyrosine was acceptable whereas arginine (Arg), citrulline (Cit), and ornithine (Orn) were below the cut-off value (0.995) recommended by the European Medicines Agency [12]. A low R^2 value may result in an inaccurate determination of the

concentration of these amino acids in real urine samples leading to incorrect conclusions.

Table 4. 2: Analytical method characteristics for the quantification of Arg, Cit, Ile, Orn, Phe, and Phe on (intra-day) of method validation.

Target analyte	Linearity (R^2)	Concentration range (mg/L)	Retention time (min)	LLOD (mg/L)	LLOQ (mg/L)
Arg	0.992	0.01 – 5.00	0.81	0.19	0.57
Cit	0.960	0.01 – 5.00	0.90	0.40	1.29
Ile	0.999	0.01 – 5.00	2.11	0.05	0.16
Orn	0.986	0.01 – 5.00	0.74	0.25	0.75
Phe	0.999	0.01 – 5.00	4.44	0.03	0.12
Tyr	0.999	0.01 – 5.00	2.96	0.06	0.19

Intra-day accuracy and precision were determined by %Recovery and %RSD respectively as shown in Tables 4.3 and Table 4.4. Recoveries for 0.1, 0.3, 0.5, 3.0, and 4.0 mg/L spiked concentrations range from 52% to 169%, 113% to 131%, 111% to 136%, 113% to 125% and 109% to 125% respectively for intra-day of method validation.

Table 4. 3: Recoveries of QC samples for BU spiked with 0.10, 0.30, 0.50, 3.00 and 4.00 mg/L of amino acid concentrations respectively on intra-day method validation n=9.

Spiked	**0,10 mg/L**		**0,30 mg/L**		**0,50 mg/L**		**3.00 mg/L**		**4.00 mg/L**	
Analyte	%Recovery	%RSD	%Recovery	%RSD	%Recovery	%RSD	%Recovery	%RSD	%Recovery	%RSD
Arg	138.91	39.05	113.71	30.91	117.95	26.28	113.44	17.40	109.74	14.18
Cit	169.91	26.86	131.07	15.88	136.42	15.06	121.31	15.12	125.15	16.46
Ile	52.19	105.52	115.87	15.55	127.91	16.69	125.30	3.61	122.22	1.78
Orn	109.76	35.35	100.46	24.30	111.84	18.07	131.13	15.80	120.13	17.96
Phe	99.08	8.95	120.35	2.68	125.29	2.80	121.86	1.47	120.92	1.45
Tyr	90.54	6.45	123.90	2.54	130.53	2.69	121.78	0.90	120.92	1.15

Table 4. 4: Recoveries of QC samples for U19 spiked with 0.10, 0.30, 0.50, 3.00 and 4.00 mg/L of amino acid concentrations respectively on intra-day method validation n=9.

Spiked	U19	0,10 mg/L		0,30 mg/L		0,50 mg/L		3.00 mg/L		4.00 mg/L	
Analyte		%Recovery	%RSD	%Recovery	%RSD	%Recovery	%RSD	%Recovery	%RSD	%Recovery	%RSD
Arg	1.47	15.76	14.02	101.13	10.24	97.99	11.40	102.30	8.84	94.37	5.90
Cit	0.87	88.58	19.66	90.82	16.45	112.06	7.39	91.00	10.32	88.08	13.21
Ile	0.14	66.54	16.45	93.41	16.55	97.60	10.90	98.25	9.76	101.96	8.21
Orn	0.70	66.68	25.75	97.21	9.17	96.24	6.73	85.16	19.40	98.06	11.88
Phe	0.31	113.73	6.94	113.20	5.82	99.26	7.58	110.63	6.52	111.86	5.16
Tyr	0.38	65.48	6.26	109.36	6.16	109.66	7.19	116.54	6.54	106.41	6.38

The linearity of the target amino acids ranges from 0.962 to 0.999, on day 2 of method validation test. The linearity of the target amino acids ranged from 0.960 to 0.999, on day 1 of method validation test as shown in Table 4.5. The analytical concentration range was from 0.01 to 5.00 mg/L. The linearity for citrulline (Cit), Isoleucine (Ile), phenylalanine (Phe), and Tyrosine were above 0.995 R^2 value whereas arginine (Arg), and ornithine (Orn) were below the cut-off value (0.995) recommended by the European Medicines Agency [12].

Table 4. 5: Analytical method characteristics for the quantification of Arg, Cit, Ile, Orn, Phe, and Phe on day 2 method validation.

Target analyte	Linearity (R^2)	Concentration range (mg/L)	Retention time (min)	LLOD (mg/L)	LLOQ (mg/L)
Arg	0.962	0.01 - 5	0.81	0.41	1.26
Cit	0.997	0.01 - 5	0.90	0.10	0.33
Ile	0.998	0.01 - 5	2.11	0.08	0.23
Orn	0.979	0.01 - 5	0.74	0.30	0.92
Phe	0.999	0.01 - 5	4.44	0.06	0.17
Tyr	0.999	0.01 - 5	2.96	0.04	0.13

Recoveries for 0.10, 0.30, 0.50, 3.00, and 4.00 mg/L spiked concentration range from 87% to 150%, 103% to 126%, 102% to 136%, 103% to 117% (Table 4.6) and 109% to 115% respectively, day 2 method validation (Table 4.7).

Table 4. 6: Extraction of QC samples (BU) when spiked with 0.10, 0.30, 0.50, 3.00 and 4.00 mg/L of amino acid concentration respectively day 2 method validation n=9.

Spiked	**0,10 mg/L**		**0,30 mg/L**		**0,50 mg/L**		**3.00 mg/L**		**4.00 mg/L**	
Analyte	%Recovery	%RSD	%Recovery	%RSD	%Recovery	%RSD	%Recovery	%RSD	%Recovery	%RSD
Arg	100.65	34.81	104.51	29.38	100.17	27.82	112.84	15.56	109.21	15.40
Cit	120.24	10.98	103.88	19.36	102.39	17.51	103.79	15.21	114.32	7.42
Ile	96.55	41.22	120.83	8.03	132.27	12.80	117.15	3.91	113.95	2.00
Orn	150.10	80.18	111.22	37.66	114.43	25.41	115.09	14.65	109.16	7.76
Phe	87.00	10.07	112.81	4.53	116.65	3.29	115.44	1.76	112.62	1.70
Tyr	114.07	9.43	126.31	4.17	128.97	3.02	117.05	1.91	115.86	1.57

Table 4. 7: Extraction of QC samples (U19) when spiked with 0.10, 0.30, 0.50, 3.00 and 4.00 mg/L of amino acid concentration respectively day 2 method validation n=9.

Spiked	U19	0,10 mg/L		0,30 mg/L		0,50 mg/L		3.00 mg/L		4.00 mg/L	
Analyte		%Recovery	%RSD	%Recovery	%RSD	%Recovery	%RSD	%Recovery	%RSD	%Recovery	%RSD
Arg	1.496	-11.881	18.889	79.377	14.478	175.454	10.014	100.592	5.791	96.929	11.312
Cit	1.028	18.262	19.510	1.461	10.649	90.609	15.272	82.925	27.619	89.983	17.266
Ile	0.124	74.043	26.344	123.669	8.703	119.977	12.299	99.319	9.520	104.706	8.746
Orn	0.799	77.107	16.719	109.584	10.364	102.313	8.968	95.865	9.524	94.959	11.278
Phe	0.444	89.330	8.793	116.779	1.587	104.361	11.287	94.700	10.569	101.227	6.854
Tyr	0.646	129.428	8.931	145.347	2.463	123.826	10.601	108.258	10.615	108.170	7.343

4.3.2.2 Inter-day (two different days) Accuracy (% recovery) and precision (%RSD)

The inter-day accuracy and precision for the extraction of 0.10, 0.30, 0.50, 3.00 and 4.00 mg/L spiked concentrations of each target amino acid in BU and U19 matrices is shown below in Sd Tables 7. 8 through to Sd Table 7.19. BU samples spiked with 0.10 mg/L

Table 4. 8: The %recovery and %RSD of inter-day at spiked amino acid concentration (0.10, 0.30, 0.50, 3.00 and 4.00 mg/L) in BU n=18.

BU	Arg	Cit	Ile	Orn	Phe	Tyr	CONC mg/L
% Recovery	119.78	145.07	74.37	129.93	93.04	102.31	0.10
%RSD	40.47	28.55	69.78	68.65	11.36	14.4	
% Recovery	109.11	117.47	118.35	105.84	116.58	125.1	0.30
%RSD	29.66	20.67	12.06	31.86	4.87	3.51	
% Recovery	109.06	119.41	130.09	113.14	120.97	129.75	0.50
%RSD	27.53	21.46	14.47	21.51	4.71	2.84	
% Recovery	113.14	112.55	121.23	123.11	118.65	119.41	3.00
%RSD	16.02	16.78	5.02	16.32	3.2	2.49	
% Recovery	109.47	119.73	118.09	114.65	116.77	118.39	4.00
%RSD	14.36	13.59	4.04	14.71	3.96	2.57	

The recoveries were within (80% -120%) at this concentration level (0.10 mg/L) for Ile (103.49%), Phe (115.79%), Tyr (99.24%), and the other amino acids were above the 20% acceptable range (around LLOD or LLOQ). The precision was bad for all the amino acids. The recoveries were good at this concentration level (0.30 mg/L) for Ile (106.26%), the other amino acids were above the 20% acceptable range while Orn (76.82%), was below the 20%. The precision was bad for all the amino acids except Phe (14.75%), and Tyr (20.28%). The recoveries were good at this concentration level (0.50 mg/L) for Phe (113.85%), Tyr (99.24%), the other amino acids were above the 20% acceptable range (for LLOD and LLOQ). The precision was bad for all the amino acids except for Phe (16.85%), and Tyr (18.08%). The recoveries were good at this concentration level (0.10 mg/L) for Ile (107.53%), the

other amino acids were above the 15% acceptable range. The precision was bad for other amino acids except Phe (9.56%), and Tyr (12.17). The recoveries were good at this concentration level (4.00 mg/L) for all amino acids except Arg (121.08%) which was above the 15% acceptable range. The precision was good for Ile (13.22%), Phe (5.29%), and Tyr (8.77%).

Table 4. 9: The %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (0.10 mg/L) in U19.

U19	Arg	Cit	Ile	Orn	Phe	Tyr	CONC mg/L
% Recovery	166.73	167.39	103.49	185.66	115.79	99.24	0.10
%RSD	260.19	174.75	68.44	138.5	21.25	42.16	
% Recovery	300.26	106.26	129.38	76.82	126.59	130.7	0.30
%RSD	79.34	64.35	31.76	83.48	14.75	20.28	
% Recovery	262.22	129.92	124.98	124.9	113.85	128.48	0.50
%RSD	73.75	26.4	25.47	33.34	16.85	18.08	
% Recovery	139.09	107.53	117.34	117.38	119.26	130.13	3.00
%RSD	29.39	22.46	18.08	38.25	9.56	12.17	
% Recovery	121.08	99	112.65	106.53	114.37	114.08	4.00
%RSD	25.38	20.82	13.22	32.49	5.29	8.77	

4.3.3 Quantitative data

There were two main sets of quantitative data, in chapter 2 patient groups are arranged according to the stages of kidney function as predicted by estimated glomerular filtration rate (eGFR). The data below show patient groups that are arranged according to the albumin creatinine ratio (A/C ratio).

The figures below Figures 4.6 to 4.9, show amino acid concentration at different stages of diabetic nephropathy (DN), given by A/C ratio where < 3.5 mg/mmol indicates negative stages provided the patient has shown this over two clinical visits in a period of 3 to 6 months. Microalbumin stages are indicated by A/C ratios between 3.5 through to 30 mg/mmol and stages above A/C ration of 30 mg/mmol are

categorized as macroalbumin. The A/C ratio is not accurate enough to predict the progression of DN [17], and only the eGFR was included in the publication.

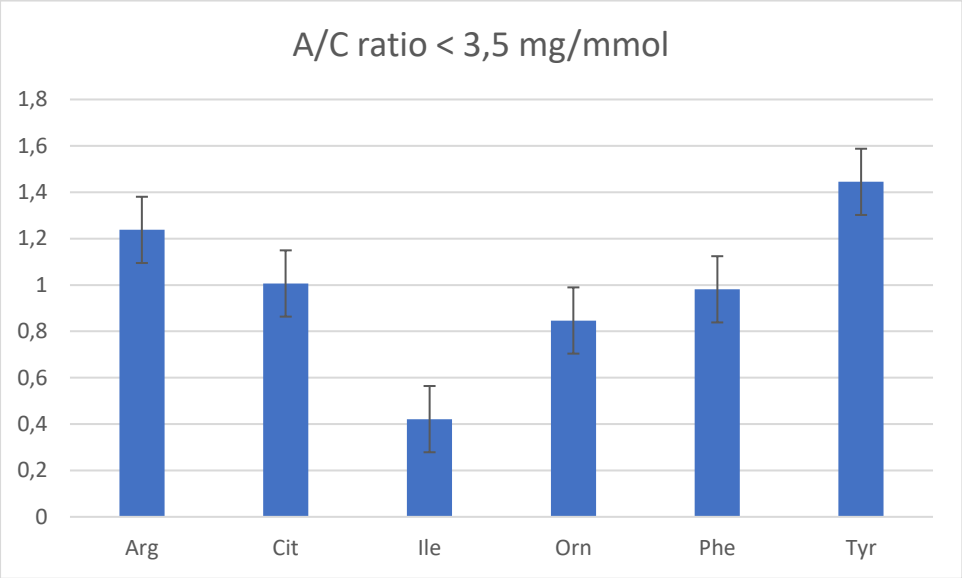


Figure 4. 6: The average amino acids concentrations in the non-diabetic patients.

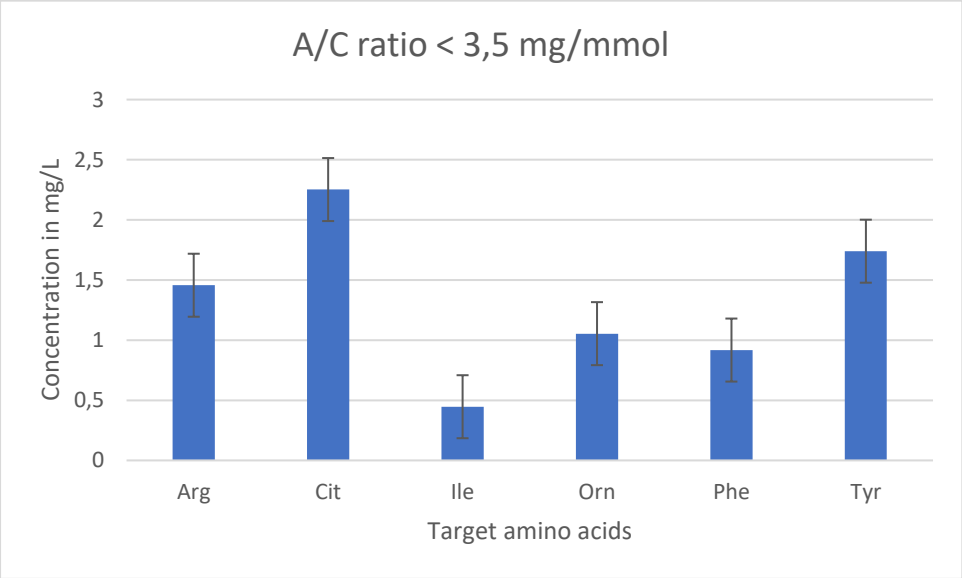


Figure 4. 7: The average amino acids concentrations in diabetic patients in the negative stage of nephropathy based on A/C ratio.

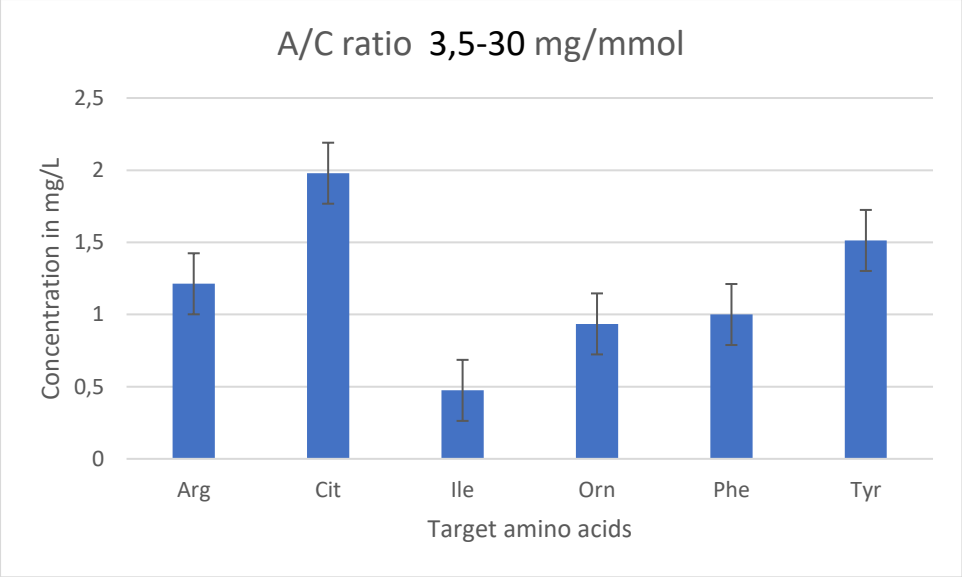


Figure 4. 8: The average amino acids concentrations in diabetic patients in the microalbumin stage of nephropathy based on A/C ratio.

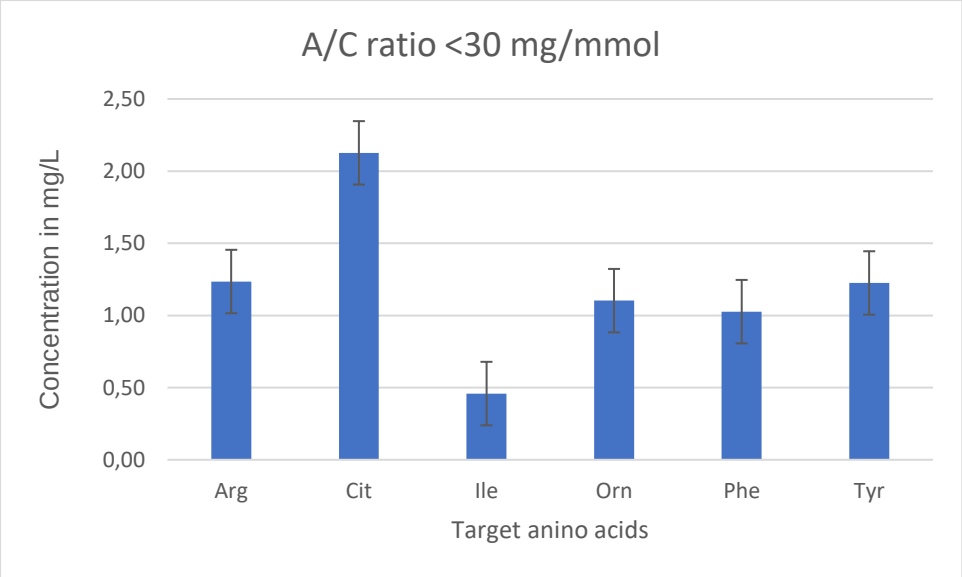


Figure 4. 9: The average amino acids concentrations in diabetic patients in the microalbumin stage of nephropathy based on A/C ratio.

Discussion and Conclusion provided in chapter 2.

4.4 References

[1] Tiwari G, Tiwari R. Bioanalytical method validation: An updated review. Pharm Methods 2010;1:25–38. https://doi.org/10.4103/2229-4708.72226.

[2] Belouafa S, Habti F, Benhar S, Belafkih B, Tayane S, Hamdouch S, et al. Statistical tools and approaches to validate analytical methods: Methodology

and practical examples. International Journal of Metrology and Quality Engineering 2017;8:1–10. https://doi.org/10.1051/ijmqe/2016030.

[3] Duggan JX, Vazvaei F, Jenkins R. Bioanalytical method validation considerations for LC-MS/MS assays of therapeutic proteins. Bioanalysis 2015;7:1389–95. https://doi.org/10.4155/bio.15.69.

[4] Kelly M, DeSilva B. Key elements of bioanalytical method validation for macromolecules. AAPS Journal 2007;9:109–14. https://doi.org/10.1208/aapsj0902017.

[5] Tijare LK, Rangari NT, Mahajan UN. A review on bioanalytical method development and validation. Asian Journal of Pharmaceutical and Clinical Research 2016;9:6–10. https://doi.org/10.22159/ajpcr.2016.v9s3.14321.

[6] Sulyok M, Stadler D, Steiner D, Krska R. Validation of an LC-MS/MS-based dilute-and-shoot approach for the quantification of > 500 mycotoxins and other secondary metabolites in food crops: challenges and solutions. Anal Bioanal Chem 2020;412:2607–20. https://doi.org/10.1007/s00216-020-02489-9.

[7] González-Medina S, Hyde C, Lovera I, Richard |, Piercy J, Piercy RJ. Detection of hypoglycin A and MCPA-carnitine in equine serum and muscle tissue: Optimisation and validation of a LC-MS-based method without derivatisation. Equine Vet J 2021;53:558–68. https://doi.org/10.1111/evj.13303.

[8] González O, Blanco ME, Iriarte G, Bartolomé L, Maguregui MI, Alonso RM. Bioanalytical chromatographic method validation according to current regulations, with a special focus on the non-well defined parameters limit of quantification, robustness and matrix effect. J Chromatogr A 2014;1353:10–27. https://doi.org/10.1016/j.chroma.2014.03.077.

[9] Wijngaard R, Perramón M, Parra-Robert M, Hidalgo S, Butrico G, Morales-Ruiz M, et al. Validation of a gas chromatography-mass spectrometry method for the measurement of the redox state metabolic ratios lactate/pyruvate and β-hydroxybutyrate/acetoacetate in biological samples. Int J Mol Sci 2021;22:1–14. https://doi.org/10.3390/ijms22094752.

[10] Ghosh C. Relative matrix effects: A step forward using standard line slopes and ANOVA analysis. Arabian Journal of Chemistry 2019;12:1–8. https://doi.org/10.1016/j.arabjc.2014.11.019.

[11] Le A, Ng A, Kwan T, Cusmano-Ozog K, Cowan TM. A rapid, sensitive method for quantitative analysis of underivatized amino acids by liquid chromatography-tandem mass spectrometry (LC-MS/MS). J Chromatogr B Analyt Technol Biomed Life Sci 2014;944:166–74. https://doi.org/10.1016/j.jchromb.2013.11.017.

[12] Committee for Medicinal Products for Human Use. Guideline on bioanalytical method validation. European Medicines Agency 2012;44:1–23. https://doi.org/EMEA/CHMP/EWP/192217/2009.

[13] Finšgar M, Perva-Uzunalić A, Behr H, Ledinek N, Knez Ž, Novak Z. An Improved Reversed-Phase High-Performance Liquid Chromatography Method for the Analysis of Related Substances of Prednisolone in Active Ingredient. ACS Omega 2020;5:1–11. https://doi.org/10.1021/acsomega.0c00037.

[14] Ridwan R, Abdul Razak HR, Adenan MI, Md Saad WM. Development of Isocratic RP-HPLC Method for Separation and Quantification of L-Citrulline and L-Arginine in Watermelons. Int J Anal Chem 2018;2018:1–9. https://doi.org/10.1155/2018/4798530.

[15] Zhou W, Yang S, Wang PG. Matrix effects and application of matrix effect factor. Bioanalysis 2017;9:1839–44. https://doi.org/10.4155/bio-2017-0214.

[16] Powley CR, George SW, Ryan TW, Buck RC. Matrix effect-free analytical methods for determination of perfluorinated carboxylic acids in environmental matrixes. Anal Chem 2005;77:6353–8. https://doi.org/10.1021/ac0508090.

[17] Mbhele T, Tanyanyiwa DM, Moepya RJ, Bhana S, Makatini MM. Relationship between amino acid ratios and decline in estimated glomerular filtration rate in diabetic and non-diabetic patients in South Africa. Afr J Lab Med 2021;10:1–7. https://doi.org/10.4102/ajlm.v10i1.1398.

Chapter 5 Qualitative analysis

5.1 Introduction

In this chapter untargeted metabolomics was used to investigate other urine metabolites that may be linked to the development of diabetic nephropathy (DN) and disease progression. MetaboAnalyst was chosen as an online profiling database, to search, identify metabolites, and also to link the metabolites to metabolic pathways altered by DN.

Metabolomics looks at molecules smaller than 1500 Daltons (Da), it is a relatively new approach, and has advantages over other omics. It provides information about possible drug targets by reviling metabolic pathways affected by the onset and progression of the disease [1].

Metabolomics is an important addition to the field of genomics; it is easily linked to the organism's genotype through elucidation of biochemical pathways and gene regulatory networks. This technique requires sample collection from biological specimens. Followed by extraction of metabolites, and then the separation to enable ease of identification and quantification using instruments such as gas chromatography or high-performance liquid chromatography coupled to mass spectrometer (GC/HPLC-MS) [2].

There are several software and online databases for interpreting metabolomics data, the one used in this research study was MetaboAnalyst which is available free online. It is the most preferred tool by researchers in the field of metabolomics, it was founded in 2009 and has over 60 000 users [3]. It has undergone several upgrades over the years and can be utilised to predict biological pathways, biomarker discovery, and statistical analysis from mass spectrometry data. In MetaboAnalyst metabolomic data can be linked to gene expression and this information can be used for elucidating a pathway or several pathways that may be linked to disease pathogenesis [4].

The data from the GC/LC-MS can be quite challenging to interpret accurately due to the vast majority of metabolites that can be detected from biological samples using these techniques. Several data formats can be generated from raw mass spectrometry data such as mzML, mzXML, and netCDF, these are usually submitted to online databases (MetaboAnalyst, PiMP, MZmine and many others) for analysis [5].

Statistical tools are required to make sense of complex metabolic data. The analyst might be interested in showing similarities or differences amongst certain groups of study subjects or populations. Multivariate data analysis can be employed to look at trends, and correlations between data groups, and can also discriminate (discriminant analysis) by looking at deference's in data sets [6]. Metaboanalyst has built in statistical tools which assist with data sorting, and interpretation and is user-friendly.

5.2 Experimental

The mass spectrometry method [MS conditions in chapter 2 (Table 2.1)] used in this study, allowed for the collection of two MS data sets. The first data set was the MRM data that was used in the quantification of target amino acids in chapters 2 and 4. The second data set was the [full scan (50-1300 m/z) MS data], which gave all detectable metabolites in the analysis of 187 urine samples. The MS data were processed using the MetaboAnalyst database.

The raw mass spectrometry data from each urine sample was converted to mzXML using the RawConveter software which is freely available online. The mzXML data files were submitted to MetaboAnalyst database together with a plain text (.tbxt) file generated and sorted using excel.

Step 1: Raw data (mzXML with text (.txt) file) was submitted into the database using the LC-MS processing function which only generates extracted ion chromatogram (EIC) of significant intensity peaks. The processed data returned 90 features (m/z) as CSV files with m/z values and peak areas of EIC.

Step 2: The CSV files were then sorted and submitted to the HMBD library to identify m/z values. The library matched m/z value were then sorted according to compound name/HMBD ID.

Step 3: The data was then resubmitted to the MetaboAnalyst pathway analysis using CSV files.

5.3 Results and discussion

5.3.1 Matching 90 features (CSV files with m/z values and peak areas of EIC (m/z) to pathways

After submitting to pathway analysis tool in MetaboAnalyst only 50 features were linked to pathways as shown below in Table 5. 1 The table shows the number of compounds present in each pathway (Total Cmpd), and the number of hits refers to the number of compounds matched to that pathway from the two groups being compared. The significance of a hit is given by raw p-value. A higher significance ($p \leq 0.05$) and impact value closer to 1, means that the difference in the hit compounds between the two groups affects the pathway. In the table below (Table 5. 1), two compounds in the phenylalanine, tyrosine and tryptophan biosynthesis pathway have a reasonable impact (0.50). Other hit compounds have very low impacts on their pathways.

Table 5. 1: Pathway analysis of significant hits and or impact (diabetes group and non-diabetes).

Pathway	Total Cmpd	Hits	Raw p	Impact
Aminoacyl-tRNA biosynthesis	48.00	3.00	0.001	0.00
Tyrosine metabolism	42.00	1.00	0.002	0.14
Ubiquinone and other terpenoid-quinone biosynthesis	9.00	1.00	0.002	0.00
Valine, leucine and isoleucine degradation	40.00	1.00	0.009	0.00
Valine, leucine and isoleucine degradation	8.00	1.00	0.009	0.00
Pantothenate and CoA biosynthesis	19.00	1.00	0.009	0.00
Phenylalanine, tyrosine and tryptophan biosynthesis	4.00	2.00	0.009	0.50
Phenylalanine metabolism	10.00	2.00	0.009	0.26

The data in Table 5.1 showed lower impact values meaning that even though there were significant ($p < 0.05$) hit compounds in some pathways their effect or influence, to course considerable changes to the pathway is minimal. This could be caused by a low number of features searched in the two population groups.

The ROC curves also showed poor predictions and there was no discriminant deference between the groups using the PLS-DA.

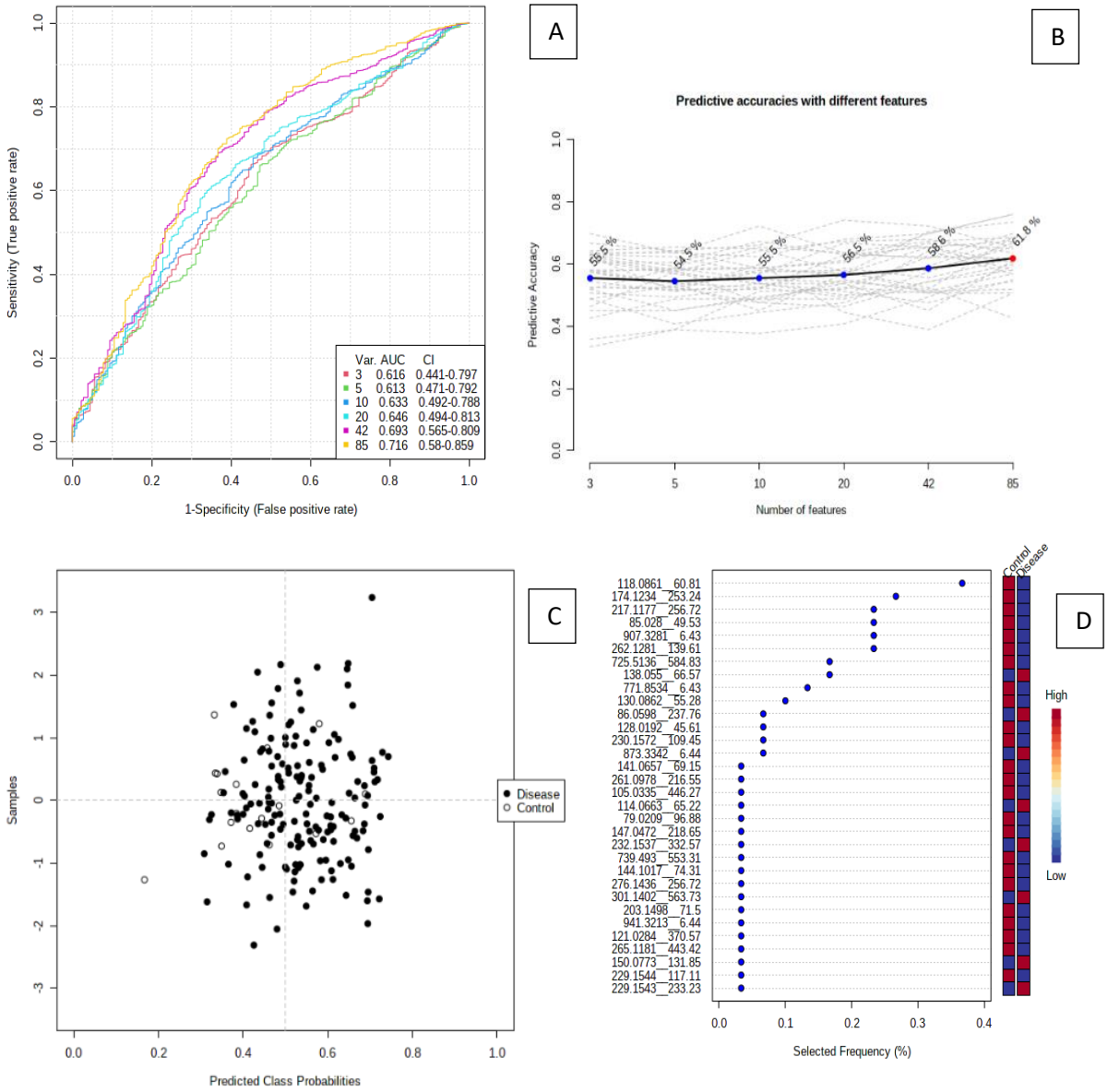


Figure 5. 1: Control vs diabetes group. The PLS-DA multivariate was used. A (Plot of ROC curves for all biomarker model based on its average performance across all MCCV runs model-0). B (Plot of the predictive accuracy of biomarker models with an increasing number of features model-6. C (plot of predicted class probabilities for all samples using a single biomarker model-6. The most accurate biomarker model is highlighted with a red dot), D (Plot of the most important features of a selected model-6 ranked from most to least important).

In an attempt to get more meaningful data, the TIC data obtained from each sample using the Bruker DataAnalysis software was sorted and submitted to the HMBD database to identify urine metabolites. The searches returned with features (m/z) linked to the KEGG ID [(KEGG is a collection of databases dealing with genomes, biological pathways, diseases, drugs, and chemical substances) (KEGG ID is a

unique identifier for each KEGG object, which is also the database entry identifier)]. The features were then used in QuantAnalysis to obtain EIC peak areas. This generated approximately 400 features (m/z) which were higher than the 90 features (m/z) from mzXML files.

5.3.2 Diabetes pathways

This data was then sorted on excel and submitted to Mummichog Pathway Analysis in MetaboAnalyst database. The data returned more pathway hits together with compound identification (KEGG ID) for 353 significant features (m/z).

The figure below (Figure 5.2) shows a summary of pathway analysis which displays all matched pathways as circles. The colour and size of each circle corresponds to its p-value and enrichment factor, respectively. The enrichment factor of a pathway is calculated as the ratio between the number of significant pathway hits and the expected number of compound hits within the pathway.

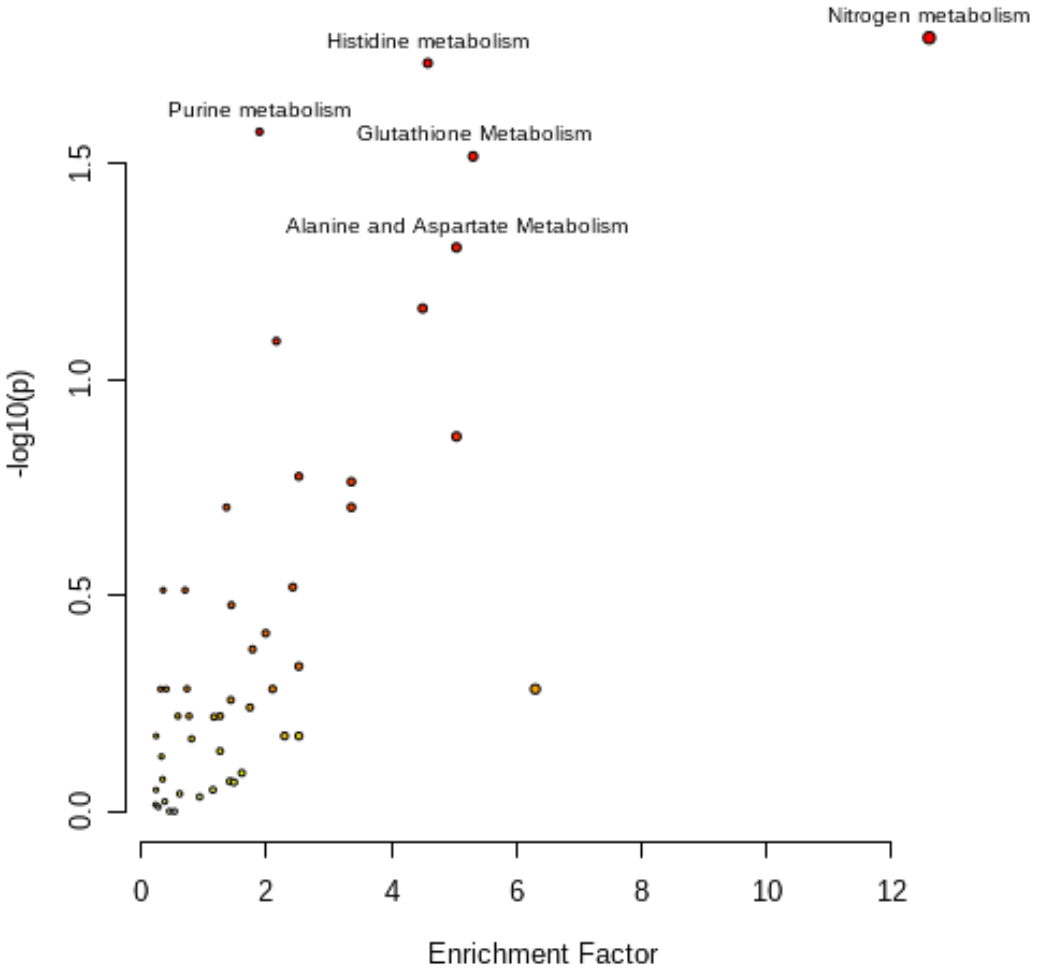


Figure 5. 2: Enrichment factor pathway analysis.

The significant features generated above using Mummichog Pathway Analysis in MetaboAnalyst database were used to determine pathways affected at each stage of CKD as shown in the following tables (Tables 5.2 to 5.7).

5.3.3 Pathway analysis

The table below (Table 5. 2) shows only pathways with a raw p value ≤ 0.05 other pathways with p > 0.05 returned by the database are not shown. The data set that was imputed on pathway analysis was that of the control group and diabetic group with (eGFR ≥ 90 mL/min per 1.73 m^2). Impact values ranged from 0.0 to 0.5. D-Glutamine and D-glutamate metabolism have 6 compounds in its pathway and 3 hit compounds have been linked to this pathway from the submitted data between the two groups and had a 50% impact (0.5). A single hit compound in the phenylalanine

tyrosine and tryptophan biosynthesis (4 total compounds) also had a 50% impact (0.5).

Interestingly regarding D-glutamine Atsushi Hesaka et al in 2019 had presented a study of D-Amino acids, as potential biomarkers, especially for kidney diseases. The study showed that D-glutamine, D-threonine, D-methionine, D-valine and D-isoleucine, which have never been detected in human body, were observed in the urine [7]. Their findings collaborate the concept that D-amino acids are present in the human body and are excreted via the kidney [8]. Hence the observed impact value of 0.5 for the D-Glutamine and D-glutamate metabolism should be further investigated in the future studies.

Table 5. 2: Control (n=17) vs stage 1 CKD group (n=66) (eGFR ≥ 90 mL/min per 1.73 m^2).

Pathway	Total Cmpd	Hits	Raw p	Impact
Arginine biosynthesis	14	7	<0.001	0.482
Glycine serine and threonine metabolism	33	5	<0.001	0.169
Arginine and proline metabolism	38	9	<0.001	0.274
D-Glutamine and D-glutamate metabolism	6	3	<0.001	0.500
Valine leucine and isoleucine degradation	40	2	<0.001	0.011
Valine leucine and isoleucine biosynthesis	8	2	<0.001	0.000
Tryptophan metabolism	41	12	<0.001	0.415
Taurine and hypo-taurine metabolism	8	1	0.001	0.000
Histidine metabolism	16	7	0.001	0.197
Glutathione metabolism	28	5	0.012	0.090
beta-Alanine metabolism	21	3	0.047	0.104

In the following table (Table 5. 3) the control group was compared to stage 2 CKD. The pathway analysis retuned eight amino acid pathways which had an impact value ranging from 0.0 to 0.5. The arginine biosynthesis pathway had 7 hit compounds (impact 0.482), while tryptophan had 14 hits with an impact of 0.415. The pathway

with the highest impact (0.5) and 3 hit compounds was D-Glutamine and D-glutamate metabolism.

Table 5. 3: Control (n=17) vs stage 2 (n=) (eGFR ≥ 60 < 90 mL/min per 1.73 m^2).

Pathway	Total Cmpd	Hits	Raw p	Impact
Arginine biosynthesis	14	7	<0.001	0.482
Glycine serine and threonine metabolism	33	5	<0.001	0.169
Arginine and proline metabolism	38	9	<0.001	0.274
D-Glutamine and D-glutamate metabolism	6	3	<0.001	0.500
Valine leucine and isoleucine degradation	40	2	<0.001	0.011
Valine leucine and isoleucine biosynthesis	8	2	<0.001	0.000
Tryptophan metabolism	41	12	<0.001	0.415
Taurine and hypo-taurine metabolism	8	1	0.001	0.000
Histidine metabolism	16	7	0.001	0.197
Glutathione metabolism	28	5	0.012	0.090
beta-Alanine metabolism	21	3	0.047	0.104

Table 5. 4: Control vs stage 3 (eGFR ≥ 30 < 60 mL/min per 1.73 m^2).

Pathway	Total Cmpd	Hits	Raw p	Impact
Arginine biosynthesis	14	7	0.042	0.482
Tyrosine metabolism	42	7	0.044	0.317
Glycerophospholipid metabolism	36	1	0.049	0.048
Ether lipid metabolism	20	1	0.049	0.000
Phosphonate and phosphinate metabolism	6	1	0.053	0.000

Nine amino acid metabolic pathways are shown in Table 5. 5 below the impact ranges from 0.0 to 0.5, D-Glutamine and D-glutamate metabolism had the highest impact value of 0.5 generated from 3 hit compounds.

Table 5. 5: Control vs stage 4 (eGFR ≥ 15 < 30 mL/min per 1.73 m^2).

Pathway	Total Cmpd	Hits	Raw p	Impact
D-Glutamine and D-glutamate metabolism	6	3	<0.001	0.500
Arginine biosynthesis	14	7	<0.001	0.482
beta-Alanine metabolism	21	3	0.002	0.104
Alanine aspartate and glutamate metabolism	28	7	0.002	0.455
Arginine and proline metabolism	38	9	0.003	0.274
Histidine metabolism	16	7	0.011	0.197
Tryptophan metabolism	41	12	0.020	0.415
Tyrosine metabolism	42	7	0.024	0.317
Glycine serine and threonine metabolism	33	5	0.031	0.169
Cysteine and methionine metabolism	33	10	0.034	0.386

Table 5. 6: below shows ten amino acid metabolic pathways, the most impacted is D-Glutamine and D-glutamate metabolism with an impact value of 0.5.

Table 5. 6: Control vs patients without eGFR data.

Pathway	Total Cmpd	Hits	Raw p	Impact
Arginine biosynthesis	14	7	<0.001	0.482
Arginine and proline metabolism	38	9	<0.001	0.274
Glycine serine and threonine metabolism	33	5	<0.001	0.169
Taurine and hypotaurine metabolism	8	1	<0.001	0.000
D-Glutamine and D-glutamate metabolism	6	3	<0.001	0.500
Valine leucine and isoleucine degradation	40	2	<0.001	0.011
Valine leucine and isoleucine biosynthesis	8	2	<0.001	0.000
Tryptophan metabolism	41	12	<0.001	0.415
Cysteine and methionine metabolism	33	10	<0.001	0.386
Histidine metabolism	16	7	0.003	0.197
beta-Alanine metabolism	21	3	0.008	0.104
Alanine aspartate and glutamate metabolism	28	7	0.054	0.455

5.3.4 Biomarker analysis

The multivariate ROC curve analysis was used to evaluate the performance of each biomarker model created through automated important feature identification. One multivariate algorithm was selected for ROC curve analysis [partial least squares discriminant analysis (PLS-DA)]. The top-ranking features (max top 100) in terms of importance were used to build the biomarker classification models. In the PLS-DA method, the number of latent variables (LV) used was 2.

In the plots below, (Figure 5. 3), all 6 biomarker models presented in the ROC plot were used ranging from model-0 to models 1-6. Model 0 = all biomarker models 1-6, Model-1 = 5 features, AUC 0.992, 95% CI 0.956 and Model-6 = AUC 0.985, 95% CI 0.938-1, (see ROC curve for model 2-5).

In the B ROC plot all models are shown from the best model-1 (5 features shown as a red square, AUC 0.992, 95% CI 0.956-1) to the last model-6 (AUC 0.985, 95% CI 0.938-1). The first 10 features (model-2) gave the best accuracy of 98.3% (A plot).

The C plot gives the top 15 biomarker compounds and L-Arginine has the highest selected frequency (100%) according to model-2. The D plot indicates how well the model used is able to discriminate between the two groups, a dotted line at 0.5 separates the two populations, there are no overlaps observed. This graph is also plotted according to model- 2.

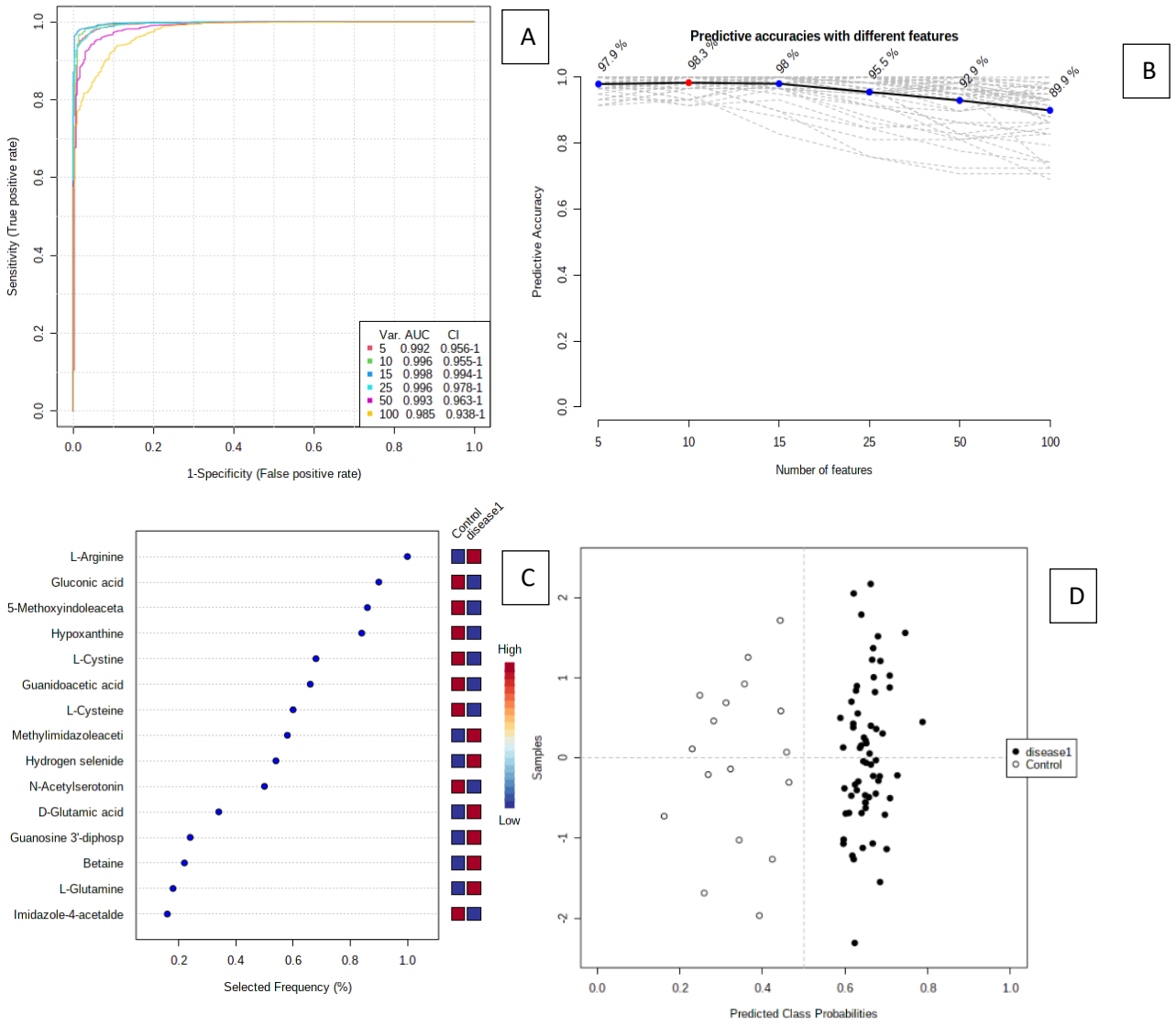


Figure 5. 3: Control vs stage 1 CKD group (eGFR ≥ 90 mL/min per 1.73 m^2). The PLS-DA multivariate was used. A (Plot of ROC curves for all biomarker model based on its average performance across all MCCV runs model-0). B (Plot of the predictive accuracy of biomarker models with an increasing number of features model-2. C (plot of predicted class probabilities for all samples using a single biomarker model-2. The most accurate biomarker model is highlighted with a red dot), D (Plot of the most important features of a selected model-2 ranked from most to least important).

The biomarker analysis of Control vs stage 2 CKD group (eGFR ≥ 60 < 90 mL/min per 1.73 m^2) is shown in the figure below (Figure 5. 4). In Figure 5. 4 the B plot indicates how well the model used is able to discriminate between the two groups a dotted line at 0.5 separates the two populations, there are no overlaps observed model-2. In the A plot all models are shown from the best (model-1 5 features are shown red square AUC 0.991, 95% CI 0.889-1) to the last model-6 (AUC 0.976, 95% CI 0.894-1) model-0. The first 10 features (model-2) gave the best accuracy of 97.4% (C plot). The D plot gives the top 15 biomarker compounds, L-Arginine has the highest selected frequency (100%), and D-Glutamic acid has a frequency of approximately 70%.

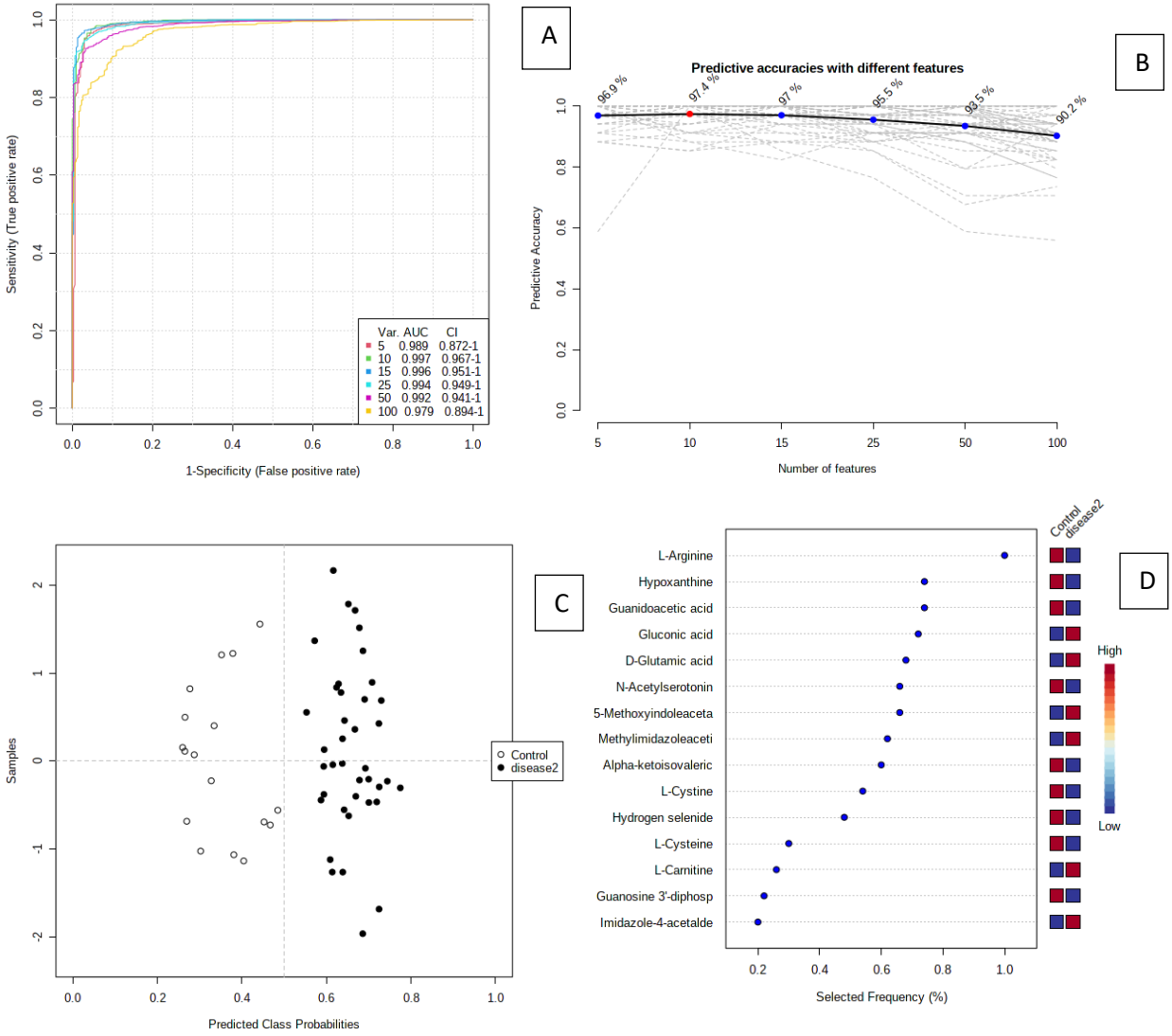


Figure 5. 4: Control vs stage 2 CKD group (eGFR ≥ 60 < 90 mL/min per 1.73 m^2). The PLS-DA multivariate was used. B (Plot of ROC curves for all biomarker model based on its average performance across all MCCV runs model-0). A (Plot of the predictive accuracy of biomarker models with an increasing number of features model-2. C (plot of predicted class probabilities for all samples using a single biomarker model-2. The most accurate biomarker model is highlighted with a red dot), D (Plot of the most important features of a selected model-2 ranked from most to least important.

In Figure 5.5 the B plot indicates how well the model used is able to discriminate between the two groups a dotted line at 0.5 separates the two populations, there is an overlap between the groups. In the A plot all models are shown from the best (model-1 5 features are shown red square AUC 0.606, 95% CI 0.226-0.906) to the last model-6 (AUC 0.654, 95% CI 0.351-0.906). The first 100 features (model-6)

gave the best accuracy of 62.5% (C plot). The D plot gives the top 15 biomarker compounds ornithine has the highest selected frequency (100%).

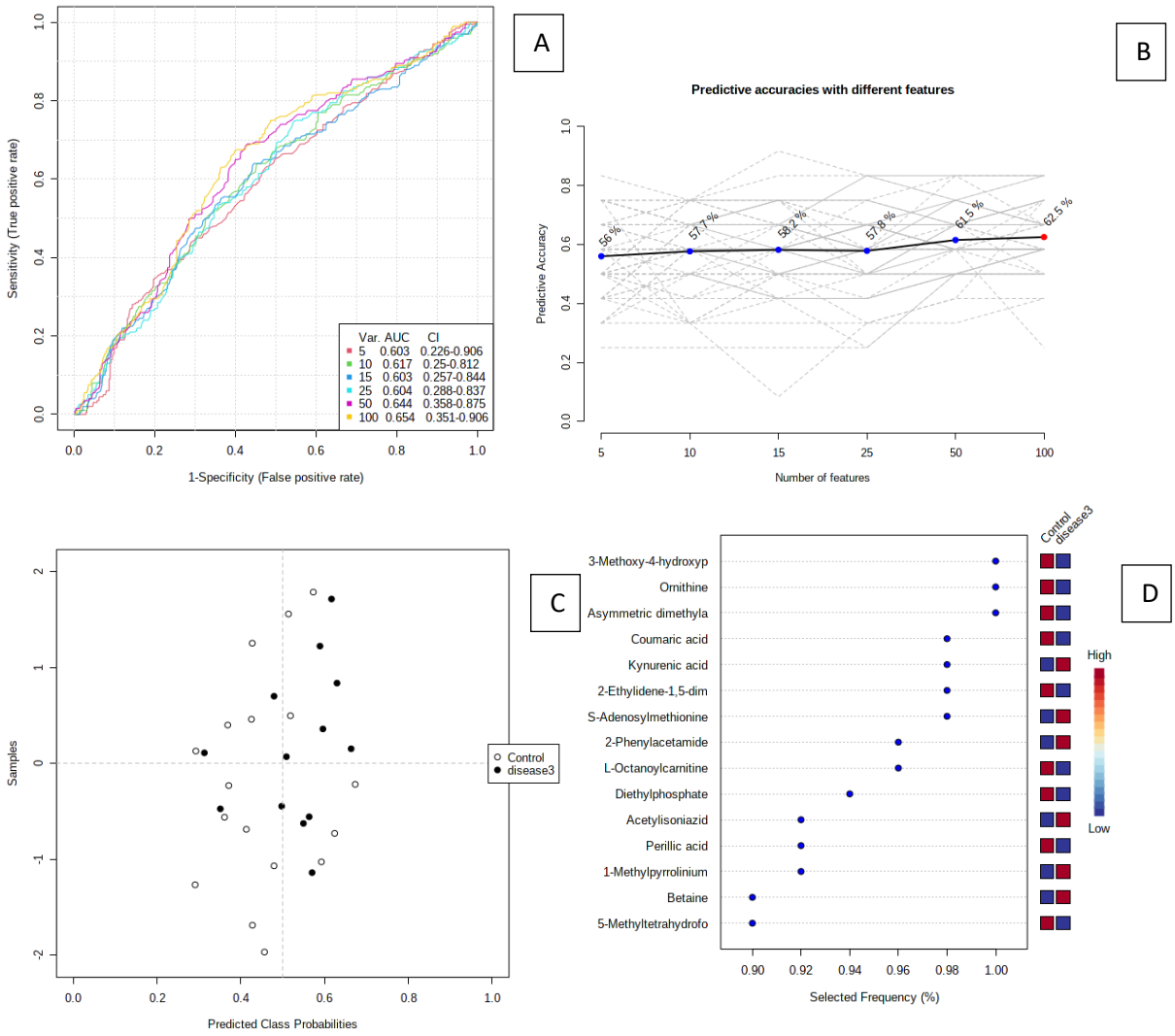


Figure 5. 5: Control vs stage 3 CKD group (eGFR ≥ 30 < 60 mL/min per 1.73 m^2). The PLS-DA multivariate was used. B (Plot of ROC curves for all biomarker model based on its average performance across all MCCV runs model-0). A (Plot of the predictive accuracy of biomarker models with an increasing number of features model-6). C (plot of predicted class probabilities for all samples using a single biomarker model-6. The most accurate biomarker model is highlighted with a red dot), D (Plot of the most important features of a selected model-6 ranked from most to least important).

In Figure 5. 6 the B plot indicates how well the model used is able to discriminate between the two groups a dotted line at 0.5 separates the two populations, there is no overlap observed between the groups. In the A plot all models are shown from the best (model-1 5 features are shown red square AUC 0.973, 95% CI 0.727-1) to

the last model-6 (AUC 0.973, 95% CI 0.792-1). The first 10 features (model-2) gave the best accuracy of 95.3% (C plot). The D plot gives the top 15 biomarker compounds D-Glutamic acid had the highest selected frequency followed by L-lysine, and L-arginine with a frequency above 80%.

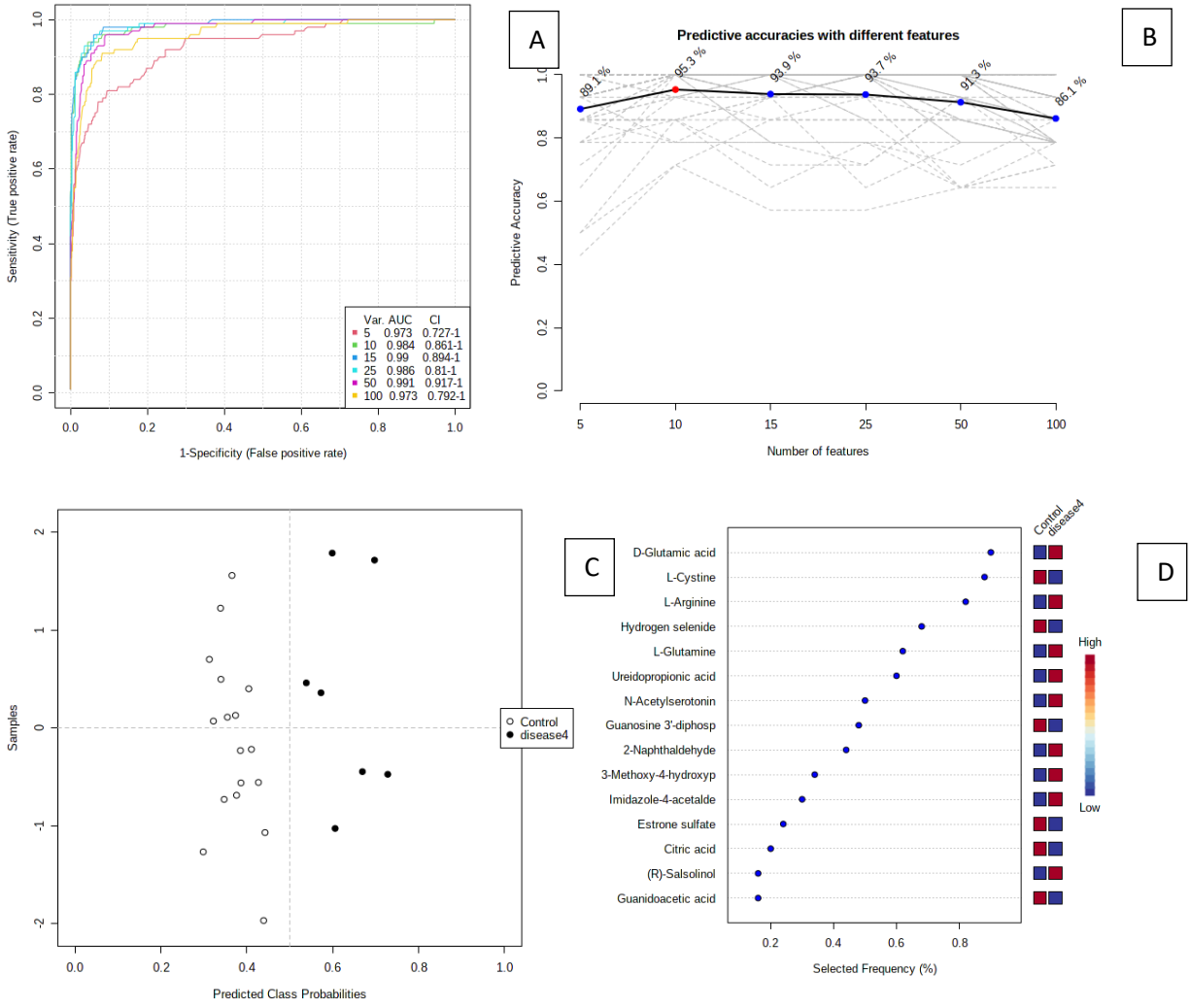


Figure 5. 6: Control vs stage 4 CKD group (eGFR ≥ 15 < 30 mL/min per 1.73 m2). The PLS-DA multivariate was used. B (Plot of ROC curves for all biomarker model based on its average performance across all MCCV runs model-0). A (Plot of the predictive accuracy of biomarker models with an increasing number of features model-2). C (plot of predicted class probabilities for all samples using a single biomarker model-2. The most accurate biomarker model is highlighted with a red dot), D (Plot of the most important features of a selected model-2 ranked from most to least important).

In Figure 5. 7 the B plot indicates how well the model used is able to discriminate between the two groups a dotted line at 0.5 separates the two populations, there is no overlap observed between the groups. In the A plot all models are shown from

the best (model-1 5 features are shown red square AUC 0.990, 95% CI 0.941-1) to the last model-6 (AUC 0.972, 95% CI 0.903-1). The first 15 features (model-3) gave the best accuracy of 95.8% (C plot). The D plot gives the top 15 biomarker compounds L-Arginine) has the highest selected frequency (100%) and D-glutamic acid had a frequency around 80%

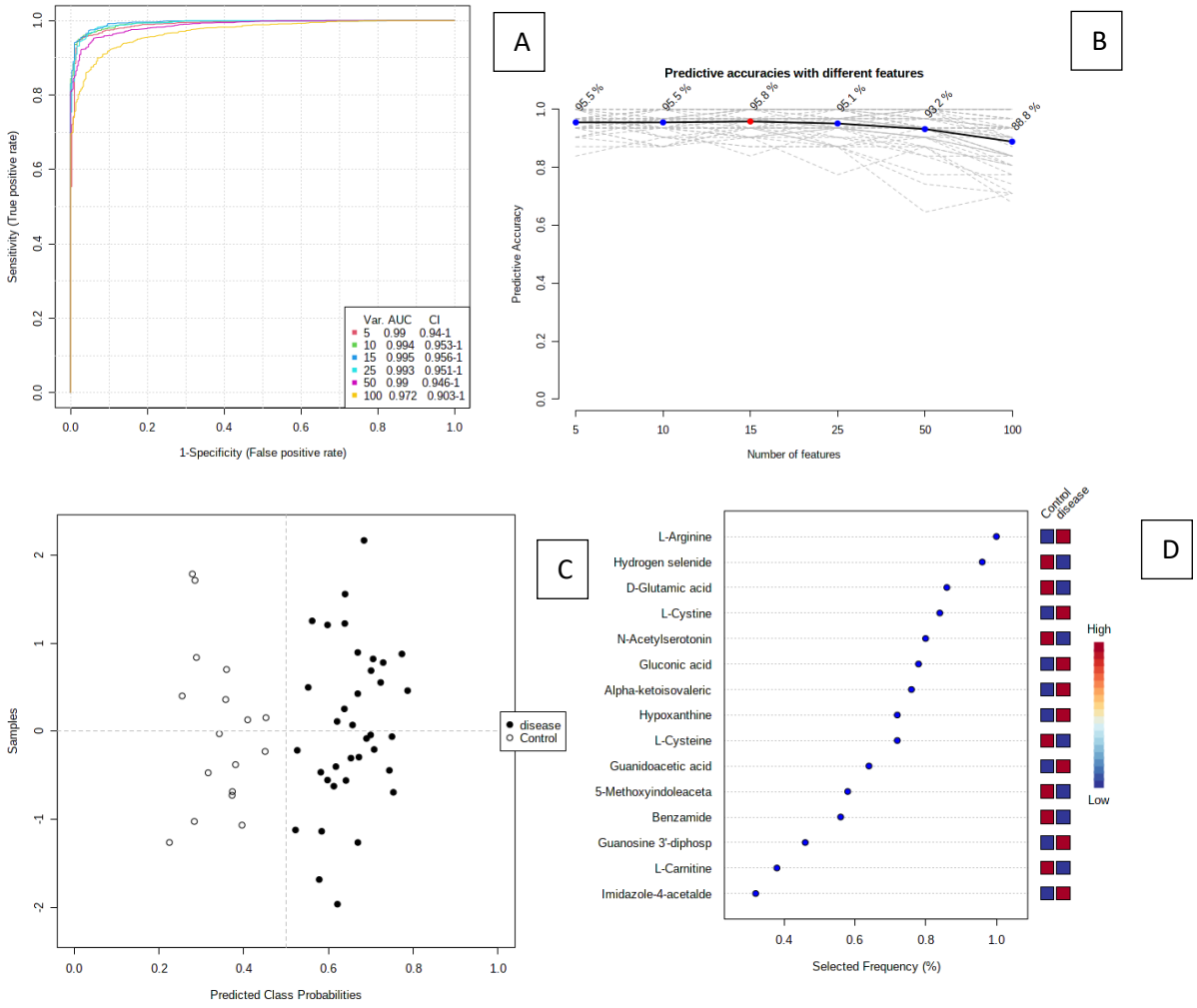


Figure 5. 7: Control vs patients without eGFR data. The PLS-DA multivariate was used. B (Plot of ROC curves for all biomarker model based on its average performance across all MCCV runs model-0). A (Plot of the predictive accuracy of biomarker models with an increasing number of features model-3). C (plot of predicted class probabilities for all samples using a single biomarker model-3. The most accurate biomarker model is highlighted with a red dot), D (Plot of the most important features of a selected model-3 ranked from most to least important).

5.3.5 Summary of results pathway analysis

In the all the pathway analysis results shown from Table 5. 2 through to 5. 4, the pathways that appear almost in all the groups are the arginine biosynthesis and the D-Glutamine and D-glutamate metabolism similarly the biomarker with the highest, frequency% is arginine. In chapter 4 the amino acid ratios of ornithine/arginine were observed to be different across the different stages of the disease.

However, the phenylalanine, tyrosine and tryptophan biosynthesis is highly impacted during the stage 1 of the CKD, this is an advantage because as indicated in chapter 24 the tyrosine/phenylalanine concentration ratio can be used as an early indicator of CKD.

The pathway analysis showed that the biosynthesis of arginine is affected at stages 1,2,3, and 4 of CKD and also in the unclassified unknown group, with the same impact (impact 0.482). Arginine metabolism may be affected by its own metabolites that may participate in the pathogenesis of renal disease and in the manipulation of the arginine biosynthesis pathway [9].

Glycine, serine, and threonine metabolism (impact 0.169), Arginine and proline metabolism (impact 0.274) appear in stages 1, 2 and 4 of CKD as well as in the unclassified unknown group. Similarly, the D-Glutamine and D-glutamate metabolism pathway (impact 0.500), cysteine and methionine metabolism (impact 0.386), Histidine metabolism (0.197), beta-Alanine metabolism (impact 0.104), and Tryptophan metabolism (impact 0.415) are impacted in these stages, stage 1,2,4, of CKD and in the unclassified unknown group. In a study conducted by Zeng et al., 2017 the decrease in the amount of tryptophan was attributed to a number of catabolic pathways that are upregulated in CKD [10].

Alanine aspartate and glutamate metabolism (impact 0.455) is impacted in the first and the fourth stage of CKD including the unclassified group. Phenylalanine, tyrosine, and tryptophan biosynthesis (impact 0.500) is only impacted in the first stage of CKD and tyrosine metabolism in stage 3 (impact 0.140), and in stage 4 (impact 0.317).

The pathway impact value is an estimate of the importance of a given pathway relative to a global metabolic network. The cut-off impact value for a given pathway needs to be ≥ 0.1 to be considered useful [4]. Raw data table for pathway analysis is shown in the appendix (Sd Table 29).

5.3.6 Summary of results biomarker analysis

In the first stage of CKD the best predictor model was model-2 (98.3%) (10 features m/z). It returned the following potential biomarkers for stage 1 CKD (frequency 60-100% listed from high frequency to low). In the second stage of CKD the best predictor model was model-2 (97.4%) (10 features m/z) and it returned the following potential biomarkers for stage 2 CKD (frequency 60-100%, listed from high frequency to low). In the third stage of CKD the best predictor model was model-6 (62.5%) (100 features m/z) and it returned the following potential biomarkers for stage 3 CKD (frequency 1% listed from high frequency to low) In the fourth stage of CKD the best predictor model was model-2 (95.3%) (10 features m/z) and it returned the following potential biomarkers for stage 4 CKD (frequency 60-100%) from high frequency to low In the unclassified stage CKD the best predictor model was model-3 (95.8%) (10 features m/z) and it returned the following potential biomarkers for unknown stage of CKD (frequency 60-100%) from high frequency to low. Below (Table 5.7) is a summary of biomarkers, raw data tables are shown in the appendix (Sd Tables 30-34).

Table 5. 7: Potential biomarkers at different stages of diabetic nephropathy.

stage 1 CKD	stage 2 CKD	stage 3 CKD	stage 4 CKD	Unknown CKD stage
L-arginine	L-arginine	3-methoxy-4- hydroxy-phenyl glycolaldehyde	D-glutamine	L-Arginine
Gluconic acid	Hypoxanthine	Ornithine	L-cystine	Hydrogen selenide
5-methoxyindoleaceta	Guanidoacetic acid	Asymmetric dimethylarginine	L-arginine	D-Glutamic acid
Hypoxanthine	Gluconic acid		Hydrogen selenide	L-Cystine
L-cystine	D-glutamic acid		L-glutamine	N-Acetylserotonine
Guanidinoacetic acid	N-acetylceretonine		Ureidopropionic acid	Gluconic acid
Methylimidazole acetic acid	5-methoxyindoleacetic acid			Alpha-ketoisovaleric acid

Hydrogen selenide	Alpha ketoisovaleric acid			Hypoxanthine
	L-cystine			Guanidoacetic acid
	Hydrogen selenide			5-methoxyindoleacetic acid

5.4 Conclusion

The data in chapter 5 shows Arg as the most sensitive marker for stage 1 through to stage 4 of CKD. Ornithine is also peaked as a biomarker in stage 3 CKD. Other sensitive amino acid biomarkers are cysteine, glutamic acid, and D-glutamine.

5.5 References

[1] Ganti S, Weiss RH. Urine metabolomics for kidney cancer detection and biomarker discovery. Urologic Oncology: Seminars and Original Investigations 2011;29:551–7. https://doi.org/10.1016/j.urolonc.2011.05.013.

[2] Li S, Park Y, Duraisingham S, Strobel FH, Khan N, Soltow QA, et al. Predicting Network Activity from High Throughput Metabolomics. PLoS Comput Biol 2013;9:1–11. https://doi.org/10.1371/journal.pcbi.1003123.

[3] Chong J, Soufan O, Li C, Caraus I, Li S, Bourque G, et al. MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis. Nucleic Acids Res 2018;46:1–9. https://doi.org/10.1093/nar/gky310.

[4] Pang Z, Zhou G, Chong J, Xia J. Comprehensive meta-analysis of covid-19 global metabolomics datasets. Metabolites 2021;11:1–14. https://doi.org/10.3390/metabo11010044.

[5] Neuweger H, Albaum SP, Dondrup M, Persicke M, Watt T, Niehaus K, et al. MeltDB: A software platform for the analysis and integration of metabolomics experiment data. Bioinformatics 2008;24:2726–32. https://doi.org/10.1093/bioinformatics/btn452.

[6] Liang Q, Liu H, Wang C, Li B. Phenotypic Characterization Analysis of Human Hepatocarcinoma by Urine Metabolomics Approach. Sci Rep 2016;6:1–8. https://doi.org/10.1038/srep19763.

[7] Hesaka A, Yasuda K, Sakai S, Yonishi H, Namba-Hamano T, Takahashi A, et al. Dynamics of D-serine reflected the recovery course of a patient with rapidly progressive glomerulonephritis. CEN Case Rep 2019;8:297–300. https://doi.org/10.1007/s13730-019-00411-6.

[8] Sasabe J, Suzuki M. Emerging role of D-Amino acid metabolism in the innate defense. Front Microbiol 2018;9:1–7. https://doi.org/10.3389/fmicb.2018.00933.

[9] Cherla G, Jaimes EA. Arginine Metabolism: Enzymology, Nutrition, and Clinical Significance Role of L-Arginine in the Pathogenesis and Treatment of Renal Disease 1,2. American Society for Nutritional Sciences 2004;134:2801–6. https://doi.org/134/10/2801S/4688510.

[10] Zeng L, Yu Y, Cai XC, Xie SB, Chen JA, Zhong LA, et al. Differences in Serum Amino Acid Phenotypes Among Patients with Diabetic Nephropathy,

Hypertensive Nephropathy, and Chronic Nephritis. Med Sci Monit 2019;25:7235–42. https://doi.org/10.12659/MSM.915735.

Chapter 6 Conclusion and future work

6.1 Conclusion

The study demonstrated that the techniques used to acquire quantitative (chapter 2) and qualitative (chapter 5) data, independently and when used together are good techniques to investigate biomarkers for diabetic nephropathy.

The quantitative study identified ornithine/arginine and tyrosine/phenylalanine as potential biomarkers of CKD in stage 3 and citrulline/ornithine ratios and as potential biomarkers of CKD stage 1.

Several biomarkers were identified in the qualitative analysis, but L-arginine was the most sensitive biomarker for early CKD and for the disease progression.

The study demonstrated that both quantitative data (chapter 3) and qualitative data (chapter 5), independently and when used together are good techniques to investigate biomarker discovery for diabetic nephropathy. The quantitative data identified ornithine/arginine and tyrosine/phenylalanine, and citrulline/ornithine ratios as potential biomarkers of CKD in stage 3 and stage 1 respectively. Several biomarkers were identified in the qualitative analysis, but L-arginine was the most sensitive biomarker for early CKD as well as for disease progression. The metabolism of D-Glutamine strongly suggests that this amino acid could also be a predictor for CKD progression. The method validation showed that our analytical method did not fulfil all the requirements for full method validation, all the analytes being quantified need to have an $R2 \geq 0.995$ on the intra-day and on inter-day. The methods would have to be revalidated until all the criteria are met.

6.2 Future work

The use of different columns would be investigated to ensure that optimum separation (base line peak resolution) is achieved. Several solvent ratios would be investigated as well as ion pairing agents together with suitable pH for amino acid analysis on the UHPLC-HRMS that would yield high sensitivity with minimal matrix effect. To simplify data processing, other user-friendly statistical tools that are less erroneous than excel spreadsheets would have to be used. The method for the

quantification of amino acid would be further studied by analysing all amino acids found in the human body. This will easily support the qualitative data which attempts to profile all potential metabolites linked to a disease or progression. The qualitative data also showed that the more samples analysed, the more pathway patterns emerge. In order for all the data to be confirmed accurately, this needs to be followed up by quantitative data that looks at potential markers in terms of concentration cut-offs. The concentration differences of metabolites at different stages of the disease can then assist to develop an easier diagnostic kit, like biosensor devices.

Chapter 7 Supplementary information

7.1 Appendix

In this chapter supporting documents are provided for chapter 2, chapter 4, and chapter 5.

Tables provide additional raw data information, also attached below are approval to use the article in the dissertation, ethical clearance certificate, and patient consent form.

7.1.1 Matrix effects

Sd Table 7. 1: Peak areas of Arg in H_2O, BU and U19 at spiked concentrations of 0.10, 0.50 and 4.00 mg/L

Arg	0.10 mg/L			0.50 mg/L			4.00 mg/L		
	H_2O	**BU**	**U19**	**H_2O**	**BU**	**U19**	**H_2O**	**BU**	**U19**
Area	17176.76	552.39	-1208.31	54320.81	2503.15	1885.71	274855.59	8185.71	14768.02
	13114.33	797.43	1699.83	51314.32	1289.58	2759.86	228257.38	11942.00	18164.82
	15805.64	416.60	2856.15	50063.07	1995.54	3950.86	201241.47	7757.98	19011.19
Mean	15365.58	588.81	1115.89	51899.40	1929.42	2865.48	234784.81	9295.23	17314.68
SD	2066.65	193.01	2094.21	2188.34	609.48	1036.62	37238.63	2302.12	2245.70
RSD	0.13	0.33	1.88	0.04	0.32	0.36	0.16	0.25	0.13
%RSD	13.45	32.78	187.67	4.22	31.59	36.18	15.86	24.77	12.97

Sd Table 7. 2: Peak areas of Cit in H_2O, BU and U19 at spiked concentrations of 0.10, 0.50 and 4.00 mg/L

Cit	0.10 mg/L			0.50 mg/L			4.00 mg/L		
	H_2O	BU	U19	H_2O	BU	U19	H_2O	BU	U19
Area	55936.42	276.84	1268.34	183749.28	1711.94	2671.17	879211.63	12984.79	43295.56
	55572.73	337.60	675.54	181839.31	1827.72	5455.84	893488.25	10883.33	42512.59
	53520.65	563.35	1263.42	190284.22	1077.38	3593.22	894579.19	14080.39	35716.55
Mean	55009.93	392.60	1069.10	185290.94	1539.01	3906.74	889093.02	12649.50	40508.23
SD	1302.51	150.96	340.84	4428.50	403.95	1418.56	8574.91	1624.69	4168.15
RSD	0.02	0.38	0.32	0.02	0.26	0.36	0.01	0.13	0.10
%RSD	2.37	38.45	31.88	2.39	26.25	36.31	0.96	12.84	10.29

Sd Table 7. 3: Peak areas of Ile in H_2O, BU and U19 at spiked concentrations of 0.10, 0.50 and 4.00 mg/L

Ile	0.10 mg/L			0.50 mg/L			4.00 mg/L		
	H_2O	BU	U19	H_2O	BU	U19	H_2O	BU	U19
Area	22675.03	3746.56	6190.63	83673.66	36200.32	37857.88	428935.56	276208.38	251163.38
	23281.01	5149.35	5465.39	81001.18	39207.63	44762.02	398287.16	285344.31	249590.52
	21482.97	6746.80	6356.86	89585.76	38385.85	45504.01	406585.16	283201.97	262914.85
Mean	22479.67	5214.24	6004.29	84753.53	37931.27	42707.97	411269.29	281584.89	254556.25
SD	914.80	1501.17	474.04	4392.99	1554.34	4216.65	15852.04	4777.82	7281.36
RSD	0.04	0.29	0.08	0.05	0.04	0.10	0.04	0.02	0.03
%RSD	4.07	28.79	7.90	5.18	4.10	9.87	3.85	1.70	2.86

Sd Table 7. 4: Peak areas of Orn in H_2O, BU and U19 at spiked concentrations of 0.10, 0.50 and 4.00 mg/L

Orn	0.10 mg/L			0.50 mg/L			4.00 mg/L		
	H_2O	BU	U19	H_2O	BU	U19	H_2O	BU	U19
Area	7696.21	338.22	462.73	37964.19	1279.75	2590.43	235262.22	7453.24	18027.89
	5125.02	863.79	2235.34	31423.44	2212.28	3135.71	227461.25	11401.32	18133.66
	6700.39	104.95	1631.71	30425.37	1353.17	2249.94	222439.22	8805.94	16941.75
Mean	6507.21	435.65	1443.26	33271.00	1615.07	2658.69	228387.56	9220.17	17701.10
SD	1296.44	388.69	901.21	4094.94	518.50	446.81	6461.49	2006.37	659.74
RSD	0.20	0.89	0.62	0.12	0.32	0.17	0.03	0.22	0.04
%RSD	19.92	89.22	62.44	12.31	32.10	16.81	2.83	21.76	3.73

Sd Table 7. 5: Peak areas of Phe in H_2O, BU and U19 at spiked concentrations of 0.10, 0.50 and 4.00 mg/L

Phe	0.10 mg/L			0.50 mg/L			4.00 mg/L		
	H_2O	BU	U19	H_2O	BU	U19	H_2O	BU	U19
Area	161290.14	64461.29	-2876.31	703782.06	292340.31	213421.66	3641046.50	2085446.50	1721893.16
	161969.73	66760.87	17530.94	691467.56	298579.53	283658.59	3614098.75	2098469.00	1882110.09
	157563.02	65478.62	54235.11	701668.25	289624.84	294037.77	3633510.00	2079308.38	1941363.58
Mean	160274.30	65566.93	22963.24	698972.63	293514.90	263706.01	3629551.75	2087741.29	1848455.61
SD	2372.50	1152.33	28940.65	6584.95	4591.44	43855.66	13903.10	9784.27	113539.79
RSD	0.01	0.02	1.26	0.01	0.02	0.17	0.00	0.00	0.06
%RSD	1.48	1.76	126.03	0.94	1.56	16.63	0.38	0.47	6.14

Sd Table 7. 6: Peak areas of Tyr in H_2O, BU and U19 at spiked concentrations of 0.10, 0.50 and 4.00 mg/L

Tyr	0.10 mg/L			0.50 mg/L			4.00 mg/L		
	H_2O	BU	U19	H_2O	BU	U19	H_2O	BU	U19
Area	116650.92	30008.96	10025.92	517627.31	171045.33	153513.56	2931739.50	1205901.75	1067263.22
	115038.20	29995.92	13223.95	526396.88	162591.31	198822.52	2933628.00	1174806.13	1186869.95
	118515.99	42466.38	27320.19	527122.69	166061.97	220189.81	2941289.50	1200795.38	1235892.72
Mean	116735.04	34157.09	16856.69	523715.63	166566.20	190841.96	2935552.33	1193834.42	1163341.96
SD	1740.42	7196.06	9201.66	5285.11	4249.50	34046.99	5057.46	16675.60	86741.87
RSD	0.01	0.21	0.55	0.01	0.03	0.18	0.00	0.01	0.07
%RSD	1.49	21.07	54.59	1.01	2.55	17.84	0.17	1.40	7.46

7.1.2 Recoveries

Sd Table 7. 7: The mean average, %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (0.10 mg/L) in U19.

U19	Arg	Cit	Ile	Orn	Phe	Tyr	
Mean	0.17	0.17	0.10	0.19	0.12	0.10	
Standard Deviation	0.43	0.29	0.07	0.26	0.02	0.04	0.10 mg/L
% Recovery	166.73	167.39	103.49	185.66	115.79	99.24	
%RSD	260.19	174.75	68.44	138.50	21.25	42.16	

Sd Table 7. 8: The mean average, %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (0.30 mg/L) in U19.

U19	Arg	Cit	Ile	Orn	Phe	Tyr	
Mean	0.90	0.32	0.39	0.23	0.38	0.39	
Standard Deviation	0.71	0.21	0.12	0.19	0.06	0.08	0.30 mg/L
% Recovery	300.26	106.26	129.38	76.82	126.59	130.70	
%RSD	79.34	64.35	31.76	83.48	14.75	20.28	

Sd Table 7. 9: The mean average, %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (0.50 mg/L) in U19.

U19	Arg	Cit	Ile	Orn	Phe	Tyr	
Mean	1.31	0.65	0.62	0.62	0.57	0.64	
Standard Deviation	0.97	0.17	0.16	0.21	0.10	0.12	0.50 mg/L
% Recovery	262.22	129.92	124.98	124.90	113.85	128.48	
%RSD	73.75	26.40	25.47	33.34	16.85	18.08	

Sd Table 7. 10: The mean average, %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (3.00 mg/L) in U19.

U19	Arg	Cit	Ile	Orn	Phe	Tyr	
Mean	4.17	3.23	3.52	3.52	3.58	3.90	
Standard Deviation	1.23	0.72	0.64	1.35	0.34	0.48	3.00 mg/L
% Recovery	139.09	107.53	117.34	117.38	119.26	130.13	
%RSD	29.39	22.46	18.08	38.25	9.56	12.17	

Sd Table 7. 11: The mean average, %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (4.00 mg/L) in U19.

U19	Arg	Cit	Ile	Orn	Phe	Tyr	
Mean	4.84	3.96	4.51	4.26	4.57	4.56	
Standard Deviation	1.23	0.82	0.60	1.38	0.24	0.40	4.00 mg/L
% Recovery	121.08	99.00	112.65	106.53	114.37	114.08	
%RSD	25.38	20.82	13.22	32.49	5.29	8.77	

Sd Table 7. 12: The mean average, %recovery and %RSD of day 1 and day 2 spiked with 0.30 mg/L amino acid concentration and overall accuracy and precision

BU	Arg	Cit	Ile	Orn	Phe	Tyr	
Mean	0.12	0.15	0.07	0.13	0.09	0.10	
Standard Deviation	0.05	0.04	0.05	0.09	0.01	0.01	0.10 mg/L
% Recovery	119.78	145.07	74.37	129.93	93.04	102.31	
%RSD	40.47	28.55	69.78	68.65	11.36	14.40	

Sd Table 7. 12: The mean average, %recovery and %RSD of day 1 and day 2 spiked with 0.30 mg/L amino acid concentration and overall accuracy and precision

BU	Arg	Cit	Ile	Orn	Phe	Tyr	
Mean	0.33	0.35	0.36	0.32	0.35	0.38	0.30 mg/L
Standard Deviation	0.10	0.07	0.04	0.10	0.02	0.01	
% Recovery	109.11	117.47	118.35	105.84	116.58	125.10	
%RSD	29.66	20.67	12.06	31.86	4.87	3.51	

Sd Table 7. 13: The mean average, %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (0.50 mg/L) in BU

BU	Arg	Cit	Ile	Orn	Phe	Tyr	
Mean	0.55	0.60	0.65	0.57	0.60	0.65	0.50 mg/L
Standard Deviation	0.15	0.13	0.09	0.12	0.03	0.02	
% Recovery	109.06	119.41	130.09	113.14	120.97	129.75	
%RSD	27.53	21.46	14.47	21.51	4.71	2.84	

Sd Table 7. 14: The mean average, %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (3.00 mg/L) in BU

BU	Arg	Cit	Ile	Orn	Phe	Tyr	
Mean	3.39	3.38	3.64	3.69	3.56	3.58	3.00 mg/L
Standard Deviation	0.54	0.57	0.18	0.60	0.11	0.09	
% Recovery	113.14	112.55	121.23	123.11	118.65	119.41	
%RSD	16.02	16.78	5.02	16.32	3.20	2.49	

Sd Table 7. 15: The mean average, %recovery and %RSD of day 1 and day 2 spiked amino acid concentration (4.00 mg/L) in BU

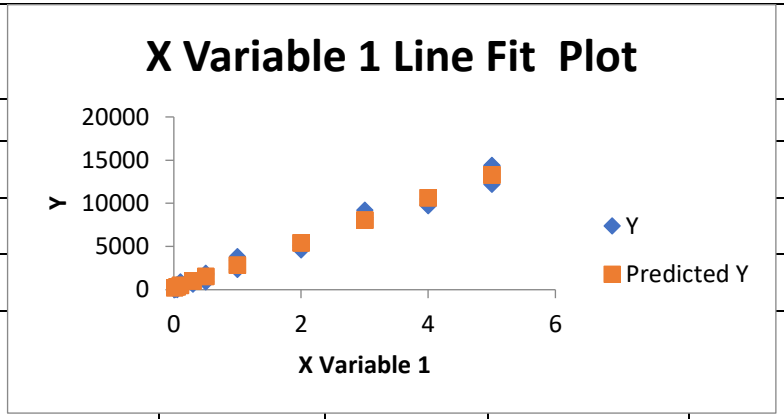
BU	Arg	Cit	Ile	Orn	Phe	Tyr	
Mean	4.38	4.79	4.72	4.59	4.67	4.74	4.00 mg/L
Standard Deviation	0.63	0.65	0.19	0.67	0.18	0.12	
% Recovery	109.47	119.73	118.09	114.65	116.77	118.39	
%RSD	14.36	13.59	4.04	14.71	3.96	2.57	

7.1.3 Method validation day1 (Sd Table 7.17 to 7.22) and day 2 (Sd Table 7.23 to 7.28)

7.1.3.1 Day 1

Sd Table 7. 16: Regression analysis of Arg and calibration curve (X variable 1line fit plot)

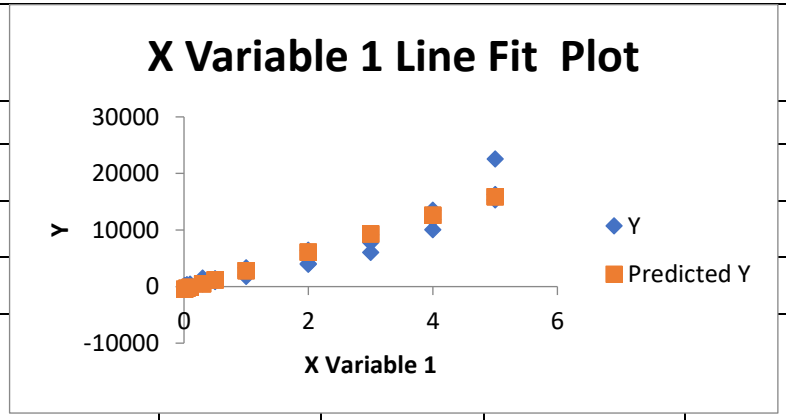
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.99176							
R Square	0.983588							
Adjusted R Square	0.983002							
Standard Error	603.8293							
Observations	30							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	6.12E+08	6.12E+08	1678.085	1.55E-26			
Residual	28	10209074	364609.8					



Total	29	6.22E+08						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	214.831	150.0476	1.431752	0.163284	-92.5276	522.1896	-92.5276	522.1896
X Variable 1	2612.564	63.7764	40.96443	1.55E-26	2481.924	2743.204	2481.924	2743.204
LLOD	0.189529	mg/mL						
LLOQ	0.574331	mg/mL						

Sd Table 7. 17: Regression analysis of Cit and calibration curve (X variable 1line fit plot)

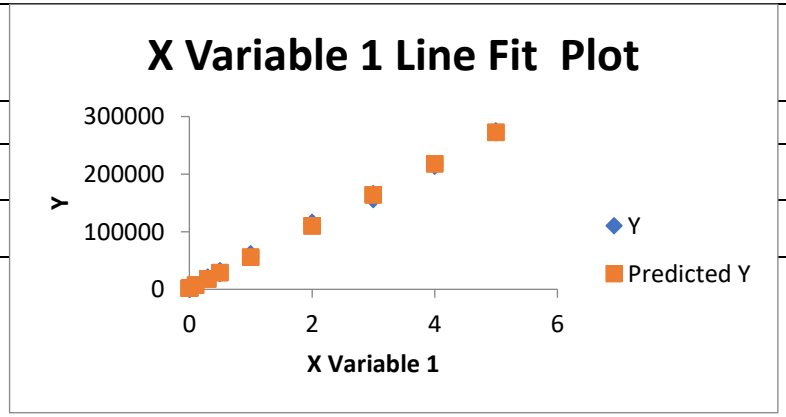
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.960641							
R Square	0.922831							
Adjusted R Square	0.920075							
Standard Error	1693.09							
Observations	30							



ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	9.6E+08	9.6E+08	334.8399	4.12E-17			
Residual	28	80263542	2866555					
Total	29	1.04E+09						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-473.532	420.7218	-1.12552	0.269923	-1335.34	388.2773	-1335.34	388.2773
X Variable 1	3272.236	178.8241	18.29863	4.12E-17	2905.931	3638.54	2905.931	3638.54
LLOD	0.424292	mg/mL						
LLOQ	1.285732	mg/mL						

Sd Table 7. 18: Regression analysis of Ile and calibration curve (X variable 1line fit plot)

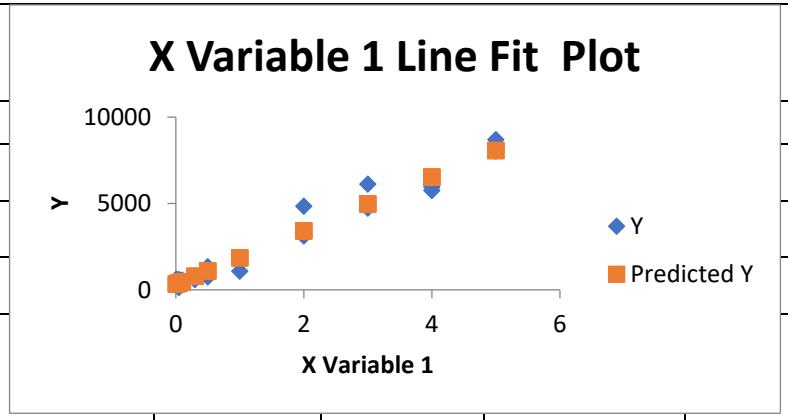
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.999358							



R Square	0.998715							
Adjusted R Square	0.99867							
Standard Error	3479.444							
Observations	30							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	2.64E+11	2.64E+11	21769.82	4.98E-42			
Residual	28	3.39E+08	12106530					
Total	29	2.64E+11						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	2105.534	862.339	2.441654	0.021195	339.1125	3871.955	339.1125	3871.955
X Variable 1	54081.19	366.5379	147.546	4.98E-42	53330.37	54832.01	53330.37	54832.01
LLOD	0.052619	mg/mL						
LLOQ	0.159453	mg/mL						

Sd Table 7. 19: Regression analysis of Orn and calibration curve (X variable 1line fit plot)

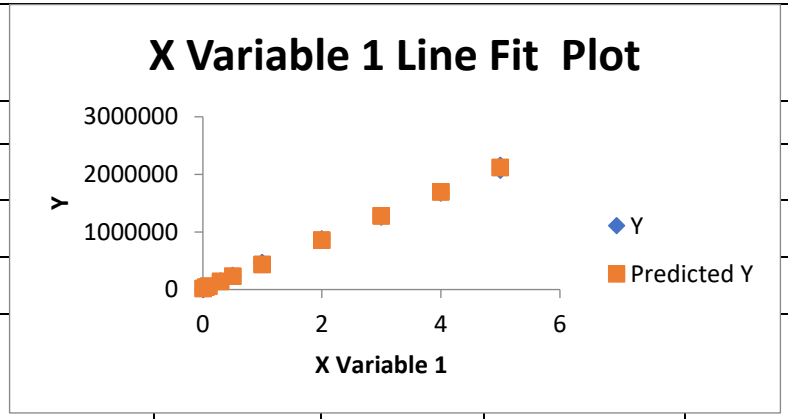
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.98609							
R Square	0.972374							
Adjusted R Square	0.971387							
Standard Error	467.8485							
Observations	30							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	2.16E+08	2.16E+08	985.5385	2.28E-23			
Residual	28	6128703	218882.2					
Total	29	2.22E+08						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%



Intercept	325.9658	116.2573	2.803831	0.009071	87.82354	564.108	87.82354	564.108
X Variable 1	1551.272	49.41413	31.39329	2.28E-23	1450.052	1652.492	1450.052	1652.492
LLOD	0.247313	mg/mL						
LLOQ	0.749432	mg/mL						

Sd Table 7. 20: Regression analysis of Phe and calibration curve (X variable 1line fit plot)

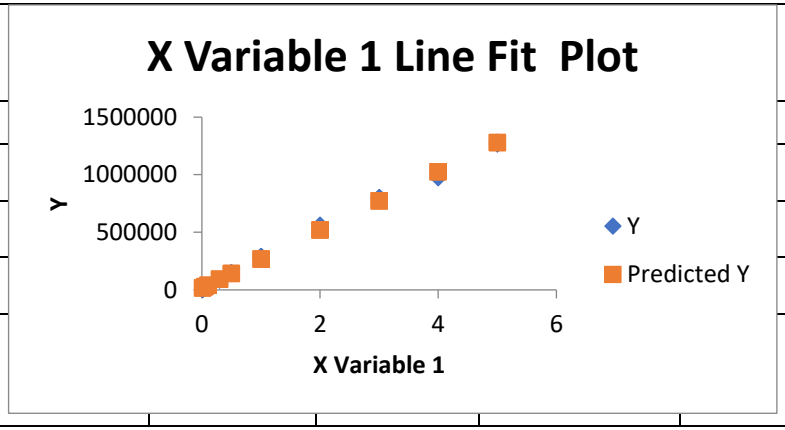
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.99966							
R Square	0.99932							
Adjusted R Square	0.999296							
Standard Error	19628.04							
Observations	30							
ANOVA								
	df	SS	MS	F	Significance F			



Regression	1	1.59E+13	1.59E+13	41177.89	6.69E-46			
Residual	28	1.08E+10	3.85E+08					
Total	29	1.59E+13						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	19044.59	4877.438	3.904629	0.000542	9053.608	29035.57	9053.608	29035.57
X Variable 1	420682.8	2073.112	202.9234	6.69E-46	416436.3	424929.4	416436.3	424929.4
LLOD	0.038261	mg/mL						
LLOQ	0.115941	mg/mL						

Sd Table 7. 21: Regression analysis of Tyr and calibration curve (X variable 1line fit plot)

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.999077							
R Square	0.998154							
Adjusted R Square	0.998088							

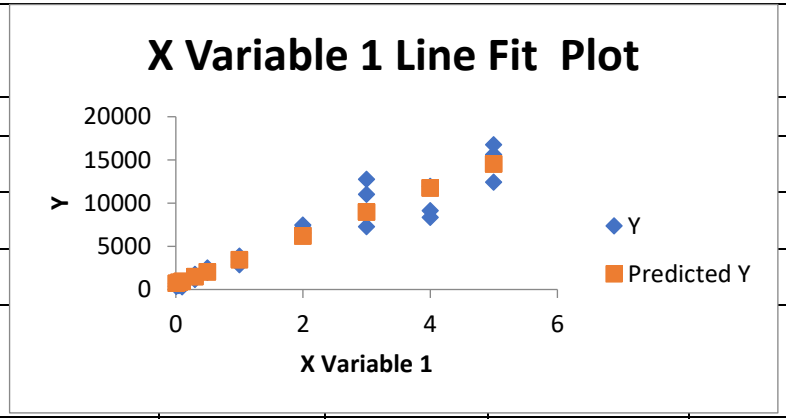


Standard Error	19440.58							
Observations	30							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	5.72E+12	5.72E+12	15142.75	7.96E-40			
Residual	28	1.06E+10	3.78E+08					
Total	29	5.73E+12						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	15531.03	4830.856	3.214965	0.003278	5635.471	25426.59	5635.471	25426.59
X Variable 1	252672.2	2053.313	123.0559	7.96E-40	248466.1	256878.2	248466.1	256878.2
LLOD	0.063093	mg/mL						
LLOQ	0.191191	mg/mL						

7.1.3.2 Day 2

Sd Table 7. 22: Regression analysis of Arg and calibration curve (X variable 1line fit plot)

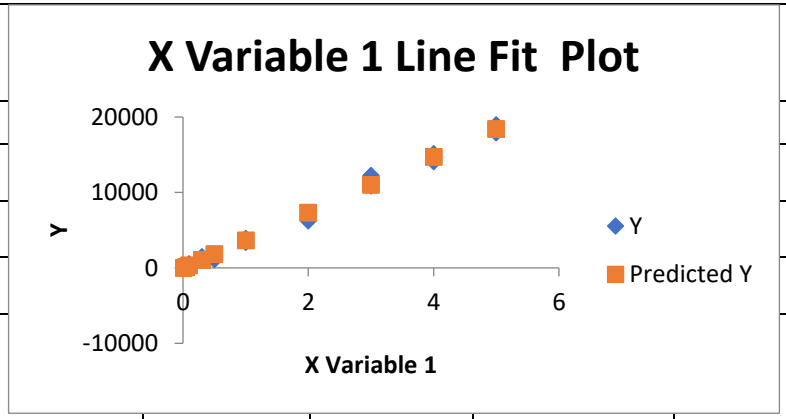
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.962361							
R Square	0.926138							
Adjusted R Square	0.9235							
Standard Error	1400.04							
Observations	30							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	6.88E+08	6.88E+08	351.0855	2.23E-17			
Residual	28	54883097	1960111					
Total	29	7.43E+08						



	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	687.551	347.9006	1.976286	0.058049	-25.0911	1400.193	-25.0911	1400.193
X Variable 1	2770.72	147.8721	18.73727	2.23E-17	2467.817	3073.622	2467.817	3073.622
LLOD	0.414359	mg/mL						
LLOQ	1.255633	mg/mL						

Sd Table 7. 23: Regression analysis of Cit and calibration curve (X variable 1line fit plot)

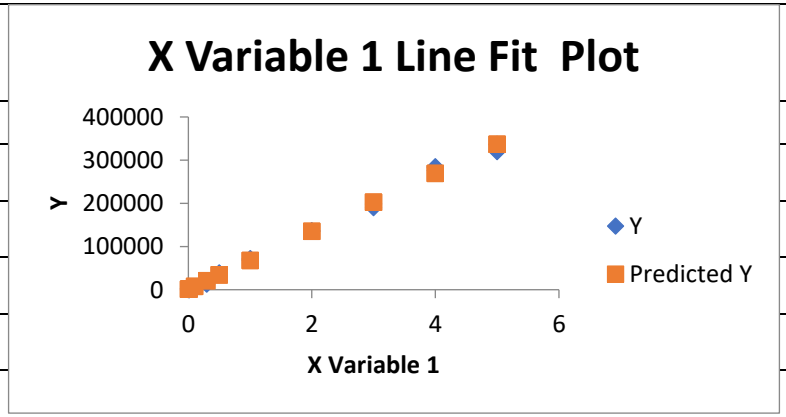
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.997267							
R Square	0.994542							
Adjusted R Square	0.994347							
Standard Error	488.9281							
Observations	30							
ANOVA								



	df	SS	MS	F	Significance F			
Regression	1	1.22E+09	1.22E+09	5101.638	3.12E-33			
Residual	28	6693420	239050.7					
Total	29	1.23E+09						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-44.2405	121.4954	-0.36413	0.718494	-293.113	204.6316	-293.113	204.6316
X Variable 1	3688.466	51.64055	71.42575	3.12E-33	3582.685	3794.246	3582.685	3794.246
LLOD	0.1087	mg/mL						
LLOQ	0.329393	mg/mL						

Sd Table 7. 24: Regression analysis of Ile and calibration curve (X variable 1line fit plot)

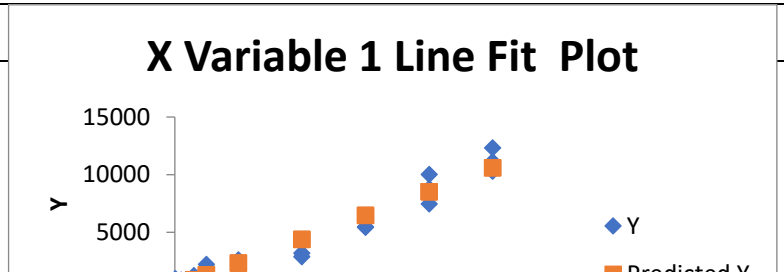
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.998625							
R Square	0.997252							
Adjusted R	0.997154							



Square								
Standard Error	6328.553							
Observations	30							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	4.07E+11	4.07E+11	10160.18	2.1E-37			
Residual	28	1.12E+09	40050587					
Total	29	4.08E+11						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	1071.909	1568.457	0.683416	0.49996	-2140.93	4284.748	-2140.93	4284.748
X Variable 1	67199.21	666.6739	100.7977	2.1E-37	65833.59	68564.83	65833.59	68564.83
LLOD	0.077023	mg/mL						
LLOQ	0.233404	mg/mL						

Sd Table 7. 25: Regression analysis of Orn and calibration curve (X variable 1line fit plot)

SUMMARY OUTPUT								

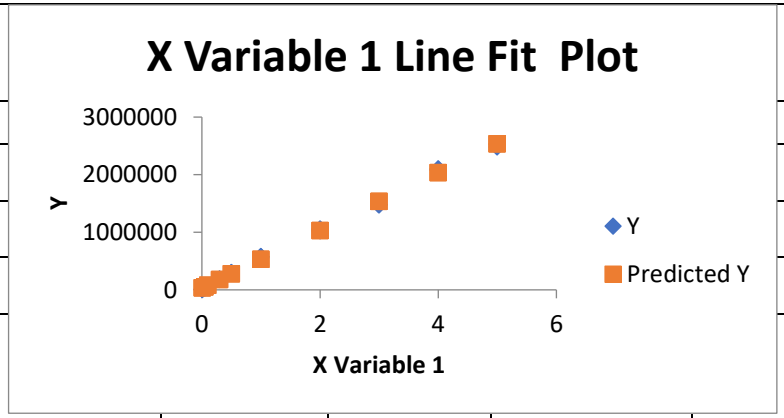


Regression Statistics								
Multiple R	0.979354							
R Square	0.959134							
Adjusted R Square	0.957675							
Standard Error	763.9046							
Observations	30							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	3.83E+08	3.83E+08	657.174	5.52E-21			
Residual	28	16339407	583550.3					
Total	29	4E+08						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	251.4707	189.8253	1.324748	0.195965	-137.369	640.3101	-137.369	640.3101
X Variable 1	2068.356	80.68355	25.63541	5.52E-21	1903.083	2233.628	1903.083	2233.628

LLOD	0.302861	mg/mL						
LLOQ	0.917759	mg/mL						

Sd Table 7. 26: Regression analysis of Phe and calibration curve (X variable 1line fit plot)

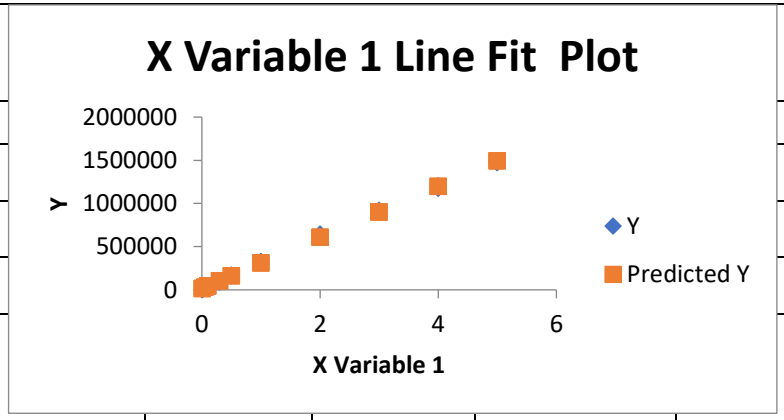
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.999297							
R Square	0.998594							
Adjusted R Square	0.998544							
Standard Error	33688.65							
Observations	30							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	2.26E+13	2.26E+13	19893.25	1.76E-41			
Residual	28	3.18E+10	1.13E+09					



Total	29	2.26E+13						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	25885.67	8371.407	3.092154	0.004467	8737.625	43033.72	8737.625	43033.72
X Variable 1	501859.6	3558.192	141.0434	1.76E-41	494571	509148.3	494571	509148.3
LLOD	0.055047	mg/mL						
LLOQ	0.166808	mg/mL						

Sd Table 7. 27: Regression analysis of Tyr and calibration curve (X variable 1line fit plot)

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.99958							
R Square	0.99916							
Adjusted R Square	0.99913							
Standard Error	15369.14							
Observations	30							



ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	7.86E+12	7.86E+12	33294.5	1.31E-44			
Residual	28	6.61E+09	2.36E+08					
Total	29	7.87E+12						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	14716.85	3819.129	3.853457	0.000622	6893.72	22539.98	6893.72	22539.98
X Variable 1	296197.6	1623.287	182.4678	1.31E-44	292872.5	299522.8	292872.5	299522.8
LLOD	0.04255	mg/mL						
LLOQ	0.128939	mg/mL						

7.2 Chapter 5 supporting documents

7.2.1 Pathway analysis

Sd Table 7. 28: Pathway analysis Control vs diabetic group

Pathway	Pathway total	Hits total	Hits sig	Expected	FET	EASE	Gamma	Emp Hits	Empirical	Pathway NO:	Cmpd Hits
Histidine metabolism	33	14	6	1.3086	0.018501	0.066331	0.033306	0	0	P1	C00025
Purine metabolism	80	15	6	3.1723	0.026678	0.086672	0.034018	0	0	P2	C00025
Nitrogen metabolism	6	4	3	0.23792	0.01617	0.1308	0.035638	0	0	P3	C00025
Glutathione Metabolism	19	8	4	0.75342	0.030431	0.13286	0.035717	0	0	P4	C00025
Alanine and Aspartate Metabolism	30	17	6	1.1896	0.049459	0.13583	0.03583	0	0	P5	C00025
Vitamin B3 (nicotinate and nicotinamide)	28	14	5	1.1103	0.068428	0.19052	0.038018	0	0	P6	C00025

metabolism											
Pyrimidine metabolism	70	19	6	2.7757	0.081473	0.19473	0.038195	0	0	P7	C01801
Arginine and Proline Metabolism	45	23	6	1.7844	0.1724	0.33147	0.044681	0	0	P8	C00025
Glutamate metabolism	15	8	3	0.5948	0.1355	0.39422	0.048251	0	0	P9	C00025
Lysine metabolism	52	22	5	2.062	0.30235	0.51105	0.056285	1	0.01	P10	C00025
Pathway	**Pathway total**	**Hits total**	**Hits sig**	**Expected**	**FET**	**EASE**	**Gamma**	**Emp Hits**	**Empirical**	**Pathway NO:**	**Cmpd Hits**
Aspartate and asparagine metabolism	114	47	9	4.5205	0.38673	0.53447	0.058173	0	0	P11	C00025
Pentose and Glucuronate Interconversions	15	5	2	0.5948	0.19761	0.59812	0.063915	0	0	P12	C00532
Pentose phosphate pathway	37	5	2	1.4672	0.19761	0.59812	0.063915	3	0.03	P13	C01801

Pentose and Glucuronate Interconversions	15	5	2	0.5948	0.19761	0.59812	0.063915	0	0	P12	C00532
Pentose phosphate pathway	37	5	2	1.4672	0.19761	0.59812	0.063915	3	0.03	P13	C01801
Urea cycle/amino group metabolism	85	31	6	3.3705	0.42126	0.60785	0.064884	1	0.01	P14	C00025
Selenoamino acid metabolism	35	7	2	1.3879	0.33285	0.72147	0.0787	0	0	P15	C01528
Glycine alanine and threonine metabolism	88	29	5	3.4895	0.55191	0.73909	0.081389	0	0	P16	C00025
Beta-Alanine metabolism	20	9	2	0.79307	0.46129	0.80718	0.093976	5	0.05	P17	C00025
Valine leucine and isoleucine degradation	65	18	3	2.5775	0.60414	0.82952	0.09915	1	0.01	P18	C00025
Aminosugars metabolism	69	10	2	2.7361	0.52001	0.83964	0.10173	2	0.02	P19	C00025
Pathway	**Pathwa**	**Hits**	**Hit**	**Expecte**	**FET**	**EASE**	**Gamma**	**Em**	**Empirica**	**Pathwa**	**Cmpd**

	y total	total	s sig	d				p Hits	l	y NO:	Hits
Ascorbate (Vitamin C) and Aldarate Metabolism	29	11	2	1.15	0.57435	0.86667	0.10953	1	0.01	P20	C00800
Methionine and cysteine metabolism	94	20	3	3.7274	0.67894	0.87213	0.1113	2	0.02	P21	C00025
Tryptophan metabolism	94	46	6	3.7274	0.81462	0.9072	0.12488	2	0.02	P22	C00025
Tyrosine metabolism	160	69	9	6.3446	0.85167	0.91659	0.12939	0	0	P23	C00025
Glycerophospholipid metabolism	156	15	2	6.1859	0.74654	0.93647	0.14088	10	0.1	P24	C14748
Butanoate metabolism	34	19	2	1.3482	0.85578	0.96987	0.17198	3	0.03	P25	C00025
C21-steroid hormone biosynthesis and metabolism	112	48	2	4.4412	0.99877	0.99988	0.37381	38	0.38	P26	C05487
C21-steroid	112	48	2	4.4412	0.99877	0.99988	0.37381	38	0.38	P26	C0548

hormone biosynthesis and metabolism											7
Arachidonic acid metabolism	95	53	2	3.7671	0.9995	0.99996	0.40414	9	0.09	P27	C00025
Sialic acid metabolism	107	12	1	4.2429	0.89143	1	1	15	0.15	P28	C05394
Pathway	**Pathway total**	**Hits total**	**Hits sig**	**Expected**	**FET**	**EASE**	**Gamma**	**Emp Hits**	**Empirical**	**Pathway NO:**	**Cmpd Hits**
Putative anti-Inflammatory metabolites formation from EPA	27	14	1	1.0706	0.92531	1	1	7	0.07	P29	C00025
Hexose phosphorylation	20	7	1	0.79307	0.72477	1	1	9	0.09	P30	C05838
Galactose metabolism	41	13	1	1.6258	0.90994	1	1	9	0.09	P31	C05394
Carnitine shuttle	72	2	1	2.855	0.30732	1	1	2	0.02	P32	C0031

											8
Vitamin B1 (thiamin) metabolism	20	5	1	0.79307	0.60152	1	1	5	0.05	P33	C00322
Vitamin B5 - CoA biosynthesis from pantothenate	12	4	1	0.47584	0.52074	1	1	7	0.07	P34	C00097
Bile acid biosynthesis	82	4	1	3.2516	0.52074	1	1	2	0.02	P35	C00037
Biopterin metabolism	22	12	1	0.87238	0.89143	1	1	44	0.44	P36	C00082
Caffeine metabolism	11	6	1	0.43619	0.66878	1	1	2	0.02	P37	C16359
Vitamin B9 (folate) metabolism	33	5	1	1.3086	0.60152	1	1	4	0.04	P38	C00025
De novo fatty acid biosynthesis	106	6	1	4.2033	0.66878	1	1	16	0.16	P39	C06428
Vitamin A (retinol) metabolism	67	16	1	2.6568	0.94868	1	1	9	0.09	P40	C16681
Fatty Acid Metabolism	63	4	1	2.4982	0.52074	1	1	12	0.12	P41	C00318
Pathway	**Pathwa**	**Hits**	**Hit**	**Expecte**	**FET**	**EASE**	**Gamma**	**Em**	**Empirica**	**Pathwa**	**Cmpd**

	y total	**total**	**s sig**	**d**				**p Hits**	**l**	**y NO:**	**Hits**
Xenobiotics metabolism	110	18	1	4.3619	0.96478	1	1	12	0.12	P42	C05851
Saturated fatty acids beta-oxidation	36	2	1	1.4275	0.30732	1	1	2	0.02	P43	C00318
Porphyrin metabolism	43	5	1	1.7051	0.60152	1	1	7	0.07	P44	C00037
Leukotriene metabolism	92	20	1	3.6481	0.97586	1	1	16	0.16	P45	C00037
Fatty acid activation	74	10	1	2.9344	0.84235	1	1	14	0.14	P46	eicostet
Carbon fixation	10	1	1	0.39653	0.16761	1	1	5	0.05	P47	C00049
Ubiquinone Biosynthesis	10	6	1	0.39653	0.66878	1	1	9	0.09	P48	C00811
1-and2-Methylnaphthalene degradation	4	4	1	0.15861	0.52074	1	1	8	0.08	P49	C14089

7.3 Biomarker discovery (Sd Tables 30 to 34)

Sd Table 7. 29: Control vs stage 1 CKD

1 CKD	Rank Freq.	Importance	Control	disease1
L-Arginine	1.00	3.70	Low	High
Gluconic acid	0.90	2.46	High	Low
5-Methoxyindoleacetate	0.86	2.35	High	Low
Hypoxanthine	0.84	2.48	High	Low
L-Cystine	0.68	2.19	High	Low
Guanidoacetic acid	0.66	2.13	High	Low
L-Cysteine	0.60	1.95	High	Low
Methylimidazoleacetic acid	0.58	2.03	Low	High
Hydrogen selenide	0.54	2.01	Low	High
N-Acetylserotonin	0.50	1.85	High	Low
D-Glutamic acid	0.34	1.82	Low	High
Guanosine 3'-diphosphate 5'-triphosphate	0.24	1.63	Low	High
Betaine	0.22	1.45	Low	High
L-Glutamine	0.18	1.50	Low	High
Imidazole-4-acetaldehyde	0.16	1.43	High	Low
L-Carnitine	0.16	1.43	High	Low
Alpha-ketoisovaleric acid	0.14	1.47	Low	High
L-Octanoylcarnitine	0.14	1.15	Low	High
3-Methoxy-4-hydroxyphenylglycolaldehyde	0.12	1.27	High	Low
Ureidopropionic acid	0.10	1.33	High	Low
L-Leucine	0.06	0.98	Low	High
Acetylisoniazid	0.06	0.95	Low	High
4a-Carbinolamine tetrahydrobiopterin	0.06	0.85	Low	High
L-Aspartic acid	0.04	0.97	High	Low

Oxoglutaric acid	0.04	0.96	High	Low
Cysteinylglycine	0.04	0.93	High	Low
Glucosamine	0.04	0.86	Low	High
L-Methionine	0.04	0.80	High	Low
3-Dehydroxycarnitine	0.04	0.67	Low	High
Linoleic acid	0.04	0.57	Low	High
Ciliatine	0.04	0.55	High	Low
L-Tyrosine	0.02	1.03	Low	High
N-Glycolylneuraminic acid	0.02	0.96	Low	High
D-1-Piperideine-2-carboxylic acid	0.02	0.90	Low	High
1 CKD	**Rank Freq.**	**Importance**	**Control**	**disease1**
(R)-Salsolinol	0.02	0.86	High	Low
1-Methylxanthine	0.02	0.84	Low	High
Glycerophosphocholine	0.02	0.82	Low	High
Glutathione episulfonium ion	0.02	0.78	Low	High
Caprylic acid	0.02	0.77	Low	High
Chloral hydrate	0.02	0.65	Low	High
Nonadeca-10(Z)-enoic acid	0.02	0.61	Low	High
Testosterone glucuronide	0.02	0.58	Low	High
Felbamate	0.02	0.57	Low	High
Chitobiose	0.02	0.56	High	Low
2-Hydroxyfelbamate	0.02	0.56	High	Low
N6-Acetyl-L-lysine	0.02	0.55	Low	High
Polylimonene	0.02	0.50	High	Low
N2-Succinyl-L-ornithine	0.02	0.47	High	Low
6-(alpha-D-Glucosaminyl)-1D-myo-inositol	0.02	0.38	High	Low

Sd Table 7. 30: Control vs stage 2 CKD

2 CKD	**Rank Freq.**	**Importance**	**Control**	**disease2**

L-Arginine	1.00	3.27	High	Low
Hypoxanthine	0.74	2.20	High	Low
Guanidoacetic acid	0.74	2.19	High	Low
Gluconic acid	0.72	2.09	Low	High
D-Glutamic acid	0.68	2.07	Low	High
N-Acetylserotonin	0.66	2.07	High	Low
5-Methoxyindoleacetate	0.66	2.06	Low	High
Methylimidazoleacetic acid	0.62	2.00	Low	High
Alpha-ketoisovaleric acid	0.60	2.02	High	Low
L-Cystine	0.54	2.03	High	Low
Hydrogen selenide	0.48	1.87	High	Low
L-Cysteine	0.30	1.66	High	Low
L-Carnitine	0.26	1.65	Low	High
Guanosine 3'-diphosphate 5'-triphosphate	0.22	1.59	High	Low
Imidazole-4-acetaldehyde	0.20	1.64	Low	High
3-Methoxy-4-hydroxyphenylglycolaldehyde	0.18	1.42	High	Low
Polylimonene	0.12	1.10	High	Low
Chondroitin	0.10	1.43	High	Low
2-Aminobenzoic acid	0.10	1.15	High	Low
L-Glutamine	0.10	1.12	High	Low
2 CKD	**Rank Freq.**	**Importance**	**Control**	**disease2**
5-Methyltetrahydrofolic acid	0.08	1.10	High	Low
Betaine	0.06	1.30	High	Low
Riboflavin reduced	0.06	1.22	High	Low
Glycerophosphocholine	0.06	1.06	Low	High
Tetrahydrocortisol	0.06	0.98	Low	High
Glutathione episulfonium ion	0.06	0.96	High	Low
N-Glycolylneuraminic acid	0.04	1.18	High	Low
N-Didesmethyl-tamoxifen	0.04	1.13	Low	High

Cocaine	0.04	1.10	High	Low
Citrulline	0.04	1.02	Low	High
L-Aspartic acid	0.04	0.97	Low	High
Acetylisoniazid	0.04	0.89	High	Low
15-Keto-prostaglandin E2	0.04	0.86	High	Low
Cysteinylglycine	0.02	1.09	Low	High
N-Acetylputrescine	0.02	1.08	Low	High
Melibiitol	0.02	1.03	Low	High
Ureidopropionic acid	0.02	0.99	Low	High
3-Hydroxylidocaine	0.02	0.97	High	Low
cis-beta-D-Glucosyl-2-hydroxycinnamate	0.02	0.92	Low	High
D-1-Piperideine-2-carboxylic acid	0.02	0.89	Low	High
N-Acetyl-L-methionine	0.02	0.83	High	Low
5-Hydroxykynurenamine	0.02	0.83	High	Low
1-Methylxanthine	0.02	0.66	Low	High
3-Methoxy-4-hydroxyphenylglycol glucuronide	0.02	0.65	High	Low
Testosterone glucuronide	0.02	0.61	High	Low
3-Dehydroxycarnitine	0.02	0.59	High	Low
Oxoglutaric acid	0.02	0.57	High	Low

Sd Table 7. 31: Control vs stage 3 CKD

3 CKD	**Rank Freq.**	**Importance**	**Control**	**disease3**
3-Methoxy-4-hydroxyphenylglycolaldehyde	1.00	1.53	High	Low
Ornithine	1.00	1.37	High	Low
Asymmetric dimethylarginine	1.00	1.27	High	Low
Coumaric acid	0.98	1.89	High	Low
Kynurenic acid	0.98	1.37	Low	High

S-Adenosylmethionine	0.98	1.24	Low	High
2-Phenylacetamide	0.96	1.61	Low	High
L-Octanoylcarnitine	0.96	1.39	High	Low
3 CKD	**Rank Freq.**	**Importance**	**Control**	**disease3**
Diethylphosphate	0.94	1.23	High	Low
Acetylisoniazid	0.92	1.21	Low	High
Perillic acid	0.92	1.19	High	Low
1-Methylpyrrolinium	0.92	1.11	Low	High
Betaine	0.90	1.47	Low	High
5-Methyltetrahydrofolic acid	0.90	1.36	High	Low
Pyridoxal	0.90	1.26	High	Low
Chenodeoxyglycocholic acid	0.90	1.26	High	Low
N-Glycolylneuraminic acid	0.90	1.17	Low	High
Cysteinylglycine	0.90	1.12	Low	High
N-Didesmethyl-tamoxifen	0.88	1.25	Low	High
Glucosamine	0.88	1.25	High	Low
7alpha-Hydroxypregnenolone	0.88	1.22	Low	High
N-Acetylgalactosamine	0.88	1.19	Low	High
L-Glutamine	0.88	1.17	High	Low
Estriol-16-Glucuronide	0.88	1.10	High	Low
Ciliatine	0.86	1.40	Low	High
Citrulline	0.86	1.26	High	Low
Glycerophosphocholine	0.86	1.25	Low	High
L-3-Hydroxykynurenine	0.84	1.23	High	Low
Testosterone glucuronide	0.84	1.15	High	Low
5-Hydroxy-L-tryptophan	0.84	1.14	Low	High
2-Hydroxynevirapine	0.84	1.11	High	Low
Dihydrothymine	0.84	1.02	High	Low
Alpha-ketoisovaleric acid	0.82	1.33	Low	High
Ecgonine methyl ester	0.82	1.29	High	Low
L-Gulonolactone	0.82	1.19	High	Low

L-Aspartic acid	0.82	1.10	High	Low
cis-beta-D-Glucosyl-2-hydroxycinnamate	0.80	1.01	Low	High
1-Methylhistidine	0.78	1.09	Low	High
1-Hydroxymethylnaphthalene	0.78	1.07	Low	High
Ureidopropionic acid	0.78	1.01	High	Low
Creatinine	0.78	0.95	High	Low
Estrone glucuronide	0.74	1.02	Low	High
Hydrogen selenide	0.72	1.24	Low	High
Chloral hydrate	0.72	1.09	Low	High
5-Acetamidovalerate	0.70	1.07	High	Low
3-Mercaptolactic acid	0.70	1.06	Low	High
Codeine	0.70	1.01	High	Low
Citric acid	0.70	0.99	High	Low
N-Acetylserotonin	0.68	1.22	High	Low
Riboflavin reduced	0.68	1.12	Low	High
3 CKD	**Rank Freq.**	**Importance**	**Control**	**disease3**
5-Hydroxyindoleacetic acid	0.68	1.05	Low	High
N-Acetylputrescine	0.68	0.87	High	Low
2-Aminobenzoic acid	0.66	1.12	Low	High
Perillyl alcohol	0.64	1.08	High	Low
Polylimonene	0.64	1.02	Low	High
13-L-Hydroperoxylinoleic acid	0.64	0.96	Low	High
Melibiitol	0.64	0.95	High	Low
L-Methionine	0.64	0.94	Low	High
Carglumic acid	0.64	0.89	High	Low
L-Lysine	0.64	0.88	High	Low
Chondroitin	0.64	0.85	High	Low
Acrylamide	0.64	0.84	High	Low
N-Methyltryptamine	0.64	0.83	Low	High
4-Aminobutyraldehyde	0.62	1.03	Low	High

Phenylacetic acid	0.62	0.99	High	Low
Cocaine	0.62	0.89	Low	High
Nonadeca-10(Z)-enoic acid	0.62	0.74	High	Low
(N-acetylneuraminosyl(a2-6)lactosamine)	0.60	0.89	Low	High
L-Alpha-aminobutyric acid	0.60	0.89	High	Low
Glutathione episulfonium ion	0.58	0.95	Low	High
Cortexolone	0.58	0.95	Low	High
Hypoxanthine	0.58	0.84	High	Low
N-Formyl-L-glutamic acid	0.58	0.81	Low	High
procollagen 5-hydroxy-L-lysine	0.58	0.75	High	Low
Chitobiose	0.58	0.75	High	Low
Estrone	0.58	0.71	High	Low
Normetanephrine	0.56	0.94	High	Low
Hordenine	0.56	0.83	Low	High
Methylimidazole acetaldehyde	0.56	0.82	High	Low
Sarcosine	0.56	0.70	High	Low
15-Keto-prostaglandin E2	0.54	0.87	High	Low
Tetrahydrobiopterin	0.54	0.84	High	Low
Adenosine monophosphate	0.54	0.82	High	Low
L-Cysteine	0.54	0.77	High	Low
Formiminoglutamic acid	0.54	0.77	Low	High
2-Aminoacrylic acid	0.52	0.87	High	Low
(R)-Salsolinol	0.52	0.85	High	Low
(S)-3-Hydroxy-N-methylcoclaurine	0.52	0.78	Low	High
gamma-Glutamylalanine	0.52	0.70	High	Low
Anserine	0.50	0.84	High	Low
Gluconic acid	0.50	0.82	High	Low
2-Keto-6-aminocaproate	0.50	0.80	Low	High
3 CKD	**Rank Freq.**	**Importance**	**Control**	**disease3**
Butyric acid	0.50	0.79	High	Low

Ureidosuccinic acid	0.48	0.85	High	Low
Argininosuccinic acid	0.48	0.85	High	Low
Oxoadipic acid	0.46	0.63	High	Low
Alpha-N-Phenylacetyl-L-glutamine	0.44	0.79	High	Low
N'-Formylkynurenine	0.44	0.77	High	Low
3-Hydroxylidocaine	0.44	0.73	High	Low
Aniline	0.44	0.73	Low	High
2-Oxo-4-methylthiobutanoic acid	0.44	0.55	High	Low
18-Hydroxycorticosterone	0.42	0.77	High	Low
1H-Indole-3-acetamide	0.42	0.76	Low	High
N-(o)-Hydroxyarginine	0.42	0.76	High	Low
Imidazole-4-acetaldehyde	0.40	0.87	High	Low
L-Leucine	0.40	0.79	Low	High
6-(alpha-D-Glucosaminyl)-1D-myo-inositol	0.40	0.70	Low	High
2-Naphthaldehyde	0.38	0.78	Low	High
Thiamine	0.38	0.72	High	Low
D-Maltose	0.38	0.70	High	Low
Pyroglutamic acid	0.38	0.62	Low	High
Dopamine	0.34	0.71	Low	High
Guanosine 3'-diphosphate 5'-triphosphate	0.34	0.70	Low	High
Dihydrotestosterone	0.34	0.66	High	Low
3-Dehydroxycarnitine	0.34	0.63	Low	High
Coniferyl alcohol	0.32	0.74	Low	High
3-Methylthiopropionic acid	0.32	0.71	Low	High
Naphthalene	0.32	0.69	Low	High
Indoleacetic acid	0.32	0.65	High	Low
L-Cystine	0.32	0.65	High	Low
Felbamate	0.32	0.62	High	Low
2-Hydroxyfelbamate	0.32	0.62	High	Low
S-Adenosylhomocysteine	0.32	0.55	Low	High

L-Tryptophan	0.30	0.72	High	Low
N2-Succinyl-L-ornithine	0.30	0.67	High	Low
3-Methoxy-4-hydroxyphenylglycol glucuronide	0.30	0.67	High	Low
2-Arachidonylglycerol	0.30	0.66	High	Low
D-1-Piperideine-2-carboxylic acid	0.30	0.66	High	Low
Tetrahydrocortisol	0.30	0.57	High	Low
L-Carnitine	0.30	0.57	High	Low
4a-Hydroxytetrahydrobiopterin	0.30	0.55	Low	High
Creatine	0.28	0.69	High	Low
3 CKD	**Rank Freq.**	**Importance**	**Control**	**disease3**
Estrone sulfate	0.28	0.66	High	Low
Pyridoxamine	0.28	0.60	High	Low
Oxoglutaric acid	0.26	0.65	High	Low
1-Methylxanthine	0.26	0.60	High	Low
L-Tyrosine	0.24	0.65	High	Low
D-Pantothenoyl-L-cysteine	0.24	0.54	High	Low
5-Methoxyindoleacetate	0.22	0.67	Low	High
D-Glutamic acid	0.22	0.62	High	Low
Saccharopine	0.22	0.56	High	Low
Phytosphingosine	0.22	0.55	High	Low
O-Phospho-4-hydroxy-L-threonine	0.20	0.61	High	Low
Leukotriene A4	0.20	0.61	Low	High
4a-Carbinolamine tetrahydrobiopterin	0.20	0.59	High	Low
N6-Acetyl-L-lysine	0.20	0.53	High	Low
Guanidoacetic acid	0.20	0.51	High	Low
Succinic acid semialdehyde	0.20	0.50	High	Low
Aminoadipic acid	0.18	0.63	High	Low
Vanillylmandelic acid	0.18	0.60	Low	High
Methylimidazoleacetic acid	0.18	0.58	High	Low
4-Guanidinobutanoic acid	0.16	0.58	High	Low

Linoleic acid	0.16	0.50	High	Low
L-Arginine	0.16	0.47	High	Low
N-Acetylneuraminate	0.14	0.59	High	Low
Caprylic acid	0.14	0.54	High	Low
5-HETE	0.14	0.51	High	Low
Adenosine	0.14	0.49	High	Low
13S-hydroxyoctadecadienoic acid	0.14	0.47	High	Low
6-Hydroxyhexanoic acid	0.14	0.45	High	Low
2-Oxoarginine	0.12	0.50	High	Low
Benzamide	0.12	0.46	High	Low

Sd Table 7. 32: Control vs stage 4 CKD

4 CKD	**Rank Freq.**	**Importance**	**Control**	**disease4**
D-Glutamic acid	0.9	2.223589	Low	High
L-Cystine	0.88	2.175518	High	Low
L-Arginine	0.82	2.209085	Low	High
Hydrogen selenide	0.68	2.011935	High	Low
L-Glutamine	0.62	1.963721	Low	High
Ureidopropionic acid	0.6	1.964464	Low	High
N-Acetylserotonin	0.5	1.890433	Low	High
Guanosine 3'-diphosphate 5'-triphosphate	0.48	1.843676	High	Low
4 CKD	**Rank Freq.**	**Importance**	**Control**	**disease4**
2-Naphthaldehyde	0.44	1.822429	Low	High
3-Methoxy-4-hydroxyphenylglycolaldehyde	0.34	1.616176	Low	High
Imidazole-4-acetaldehyde	0.3	1.601219	Low	High
Estrone sulfate	0.24	1.475653	High	Low
Citric acid	0.2	1.475303	High	Low
(R)-Salsolinol	0.16	1.399644	Low	High

Guanidoacetic acid	0.16	1.376794	High	Low
Coumaric acid	0.16	1.33407	High	Low
Hypoxanthine	0.16	1.215648	High	Low
L-Cysteine	0.16	1.198632	High	Low
Glycerophosphocholine	0.14	1.236709	High	Low
Methylimidazoleacetic acid	0.12	1.382246	High	Low
2-Phenylacetamide	0.1	1.247519	High	Low
Gluconic acid	0.1	1.13544	High	Low
Ciliatine	0.08	0.983867	Low	High
5-Methyltetrahydrofolic acid	0.08	0.940347	High	Low
Estriol-16-Glucuronide	0.06	1.096865	High	Low
2-Hydroxynevirapine	0.06	1.091891	High	Low
Polylimonene	0.06	1.070782	Low	High
18-Hydroxycorticosterone	0.06	1.041318	High	Low
2-Oxoarginine	0.06	1.023637	High	Low
Testosterone glucuronide	0.06	0.879018	High	Low
Coniferyl alcohol	0.06	0.834125	High	Low
Ecgonine methyl ester	0.04	1.143317	High	Low
L-Gulonolactone	0.04	1.133928	High	Low
N-Didesmethyl-tamoxifen	0.04	1.082286	High	Low
N-Glycolylneuraminic acid	0.04	1.038324	High	Low
Asymmetric dimethylarginine	0.04	0.988203	Low	High
Citrulline	0.04	0.967968	High	Low
L-3-Hydroxykynurenine	0.04	0.90934	High	Low
Vanillylmandelic acid	0.04	0.897084	High	Low
3-Hydroxylidocaine	0.04	0.740177	High	Low
Aniline	0.04	0.718843	High	Low
Leukotriene A4	0.04	0.715969	High	Low
2-Aminoacrylic acid	0.04	0.655704	High	Low
1-Methylhistidine	0.04	0.592722	High	Low
5-Methoxyindoleacetate	0.02	1.023511	High	Low
Indoleacetic acid	0.02	0.970686	Low	High

Tetrahydrobiopterin	0.02	0.958142	High	Low
4-Guanidinobutanoic acid	0.02	0.93374	High	Low
L-Octanoylcarnitine	0.02	0.930072	High	Low
Chondroitin	0.02	0.926819	High	Low
4 CKD	**Rank Freq.**	**Importance**	**Control**	**disease4**
2-Keto-6-aminocaproate	0.02	0.919324	High	Low
1-Methyluric acid	0.02	0.916114	High	Low
Phenylacetic acid	0.02	0.911867	High	Low
Riboflavin reduced	0.02	0.875014	High	Low
Cysteinylglycine	0.02	0.839757	High	Low
Anserine	0.02	0.771728	Low	High
N-Acetylputrescine	0.02	0.746875	High	Low
Chloral hydrate	0.02	0.706886	Low	High
Diethylphosphate	0.02	0.687958	Low	High
Succinic acid semialdehyde	0.02	0.631458	High	Low
6-Hydroxyhexanoic acid	0.02	0.630982	Low	High
Felbamate	0.02	0.627856	Low	High
N'-Formylkynurenine	0.02	0.612704	Low	High
2-Hydroxyfelbamate	0.02	0.558357	High	Low
N6-Acetyl-L-lysine	0.02	0.551539	Low	High

Sd Table 7. 33: Control vs unknown CKD

Unknown CKD	**Rank Freq.**	**Importance**	**Control**	**disease**
L-Arginine	1.00	3.25	Low	High
Hydrogen selenide	0.96	2.39	High	Low
D-Glutamic acid	0.86	2.13	High	Low
L-Cystine	0.84	2.08	Low	High
N-Acetylserotonin	0.80	2.06	High	Low
Gluconic acid	0.78	1.91	Low	High
Alpha-ketoisovaleric acid	0.76	1.96	Low	High

Hypoxanthine	0.72	2.06	Low	High
L-Cysteine	0.72	1.89	High	Low
Guanidoacetic acid	0.64	1.87	Low	High
5-Methoxyindoleacetate	0.58	1.78	High	Low
Benzamide	0.56	1.74	High	Low
Guanosine 3'-diphosphate 5'-triphosphate	0.46	1.66	Low	High
L-Carnitine	0.38	1.49	High	Low
Imidazole-4-acetaldehyde	0.32	1.53	Low	High
3-Methoxy-4-hydroxyphenylglycolaldehyde	0.32	1.51	Low	High
Methylimidazoleacetic acid	0.32	1.40	High	Low
Ureidopropionic acid	0.28	1.38	Low	High
Betaine	0.26	1.35	Low	High
Chondroitin	0.20	1.33	Low	High
L-Glutamine	0.18	1.29	High	Low
Unknown CKD	**Rank Freq.**	**Importance**	**Control**	**disease**
Oxoadipic acid	0.18	1.17	Low	High
15-Keto-prostaglandin E2	0.14	0.96	Low	High
Oxoglutaric acid	0.12	1.13	High	Low
N-(o)-Hydroxyarginine	0.12	1.10	Low	High
L-Octanoylcarnitine	0.12	1.00	Low	High
Riboflavin reduced	0.10	1.10	High	Low
1H-Indole-3-acetamide	0.10	1.08	High	Low
Polylimonene	0.08	1.02	High	Low
D-1-Piperideine-2-carboxylic acid	0.08	0.95	Low	High
4a-Carbinolamine tetrahydrobiopterin	0.08	0.82	Low	High
Indoleacetic acid	0.08	0.71	Low	High
L-Leucine	0.06	1.13	High	Low
Cysteinylglycine	0.06	1.03	High	Low
Acetylisoniazid	0.06	0.97	High	Low

L-Tyrosine	0.06	0.96	High	Low
Coumaric acid	0.06	0.95	Low	High
(S)-3-Hydroxy-N-methylcoclaurine	0.06	0.81	High	Low
Asymmetric dimethylarginine	0.06	0.81	High	Low
6-Hydroxyhexanoic acid	0.06	0.59	High	Low
L-Aspartic acid	0.04	1.00	High	Low
cis-beta-D-Glucosyl-2-hydroxycinnamate	0.04	0.98	Low	High
Perillic acid	0.04	0.95	High	Low
Adenosine	0.04	0.92	High	Low
Formiminoglutamic acid	0.04	0.86	Low	High
Perillyl alcohol	0.04	0.83	Low	High
5-Methyltetrahydrofolic acid	0.04	0.70	High	Low
Aniline	0.04	0.64	Low	High
Tetrahydrocortisol	0.04	0.61	Low	High
Diethylphosphate	0.02	1.09	High	Low
3-Methoxy-4-hydroxyphenylglycol glucuronide	0.02	0.94	High	Low
Pyridoxal	0.02	0.93	High	Low
5-Hydroxykynurenamine	0.02	0.87	High	Low
N-Acetyl-L-methionine	0.02	0.86	High	Low
N-Glycolylneuraminic acid	0.02	0.85	Low	High
Glucosamine	0.02	0.81	High	Low
2-Hydroxynevirapine	0.02	0.80	High	Low
5-Hydroxy-L-tryptophan	0.02	0.78	Low	High
13-L-Hydroperoxylinoleic acid	0.02	0.77	High	Low
Estriol-16-Glucuronide	0.02	0.76	Low	High
2-Phenylacetamide	0.02	0.75	Low	High
4-Guanidinobutanoic acid	0.02	0.75	High	Low
Unknown CKD	**Rank Freq.**	**Importance**	**Control**	**disease**
L-Lysine	0.02	0.75	Low	High

L-Gulonolactone	0.02	0.75	Low	High
13S-hydroxyoctadecadienoic acid	0.02	0.73	Low	High
Glutathione episulfonium ion	0.02	0.68	Low	High
N6-Acetyl-L-lysine	0.02	0.65	High	Low
2-Oxoarginine	0.02	0.63	High	Low
Dihydrotestosterone	0.02	0.62	High	Low
Adenosine monophosphate	0.02	0.61	High	Low
N2-Succinyl-L-ornithine	0.02	0.60	High	Low
3-Dehydroxycarnitine	0.02	0.55	High	Low
18-Hydroxycorticosterone	0.02	0.53	Low	High
Phytosphingosine	0.02	0.52	Low	High
Estrone sulfate	0.02	0.48	Low	High
5-HETE	0.02	0.48	Low	High
1-Methyluric acid	0.02	0.45	High	Low



AJLM Manuscripts
to me

Good day,

Thank you for your email.

Your may use the article, however you need to ensure that you cite the article.

Kind regards
Madeleine Coetzee

African Journal of Laboratory Medicine

www.ajlmonline.org | Follow: @AfrJLabMed | Receive email press releases and call for papers: Subscribe

Submission and Review Editorial Coordinator: Tanien Botes, Telephone Ext: 506 | submissions@ajlmonline.org
Finalisation and Publishing Production Coordinator: Madeleine Coetzee,Telephone Ext: 504 | publishing@ajlmonline.org
Compliment or complaint: Hannelie Diedericks, Telephone Ext: 502 | publishing@aosis.co.za
Tel: +27 21 975 2602 | Fax to email: 086 685 1577 | Fax: +27 21 975 4635

15 Oxford Street, Durbanville, Cape Town, 7550 | Get directions
Published by AOSIS (Pty) Ltd: aosis.co.za, Facebook, Twitter, LinkedIn, YouTube, SoundCloud, and mailing list subscription



UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

R14/49 Dr Donald M Tanyanyiwa

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M180798

NAME: (Principal Investigator)	Dr Donald M Tanyanyiwa
DEPARTMENT:	Department of Chemical Pathology Chris Hani Baragwanath Academic Hospital
PROJECT TITLE:	Investigation of Urinary Amino Acid Patterns in Diabetes
DATE CONSIDERED:	22/02/2013
DECISION:	Approved unconditionally
CONDITIONS:	Renewal for 5 Years Valid for the Period 01 July 2018 - 30 June 2023 Previously M120213
SUPERVISOR:	
APPROVED BY:	Professor CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	13/08/2018

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary on the Third Floor, Faculty of Health Sciences, Philip Tobias Building, 29 Princess of Wales Terrace, Parktown, 2193, University of the Witwatersrand. I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. **I agree to submit a yearly progress report**. The date for annual re-certification will be one year after the date of convened meeting where the study was initially reviewed. In this case, the study was initially reviewed in July and will therefore be due in the month of July each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature

Date 24 Sept 2018

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES



INFORMATION DOCUMENT

Study title: INVESTIGATION OF URINARY AMINO ACID PATTERNS IN DIABETES

Greeting: Good day, my name is Donald M Tanyanyiwa. I work in Chemical Pathology here at Chris Hani Baragwanath Academic Hospital

Introduction:

We are doing some work to find out if one of the complications of diabetes (kidneys damage) can be detected early. It is known that at some stage people with diabetes develop renal failure. At the present moment this development can be detected when protein appear in urine. When the kidneys are health, proteins will not be able leave the body and so it is all kept within the body. It is understood that glucose can cause some changes in the kidney resulting in kidney disease resulting in protein appearance in urine. Kidney disease caused by diabetes is known to be one of the highest causes of renal disease. Since the cause is known (diabetes) researchers have been looking at developing tests that can detect the early onset of the disease because once it develops, besides being very expensive to treat, it causes immobility and also known to cause complications in the cardiovascular system. The study will require the urine which you already submitted when you were tested. We will also need the blood results for HbA1c and Glucose which were tested today. There will be no immediate benefits to you but depending on the results benefits may be realised later for other patients in future especially those who are seen in district hospitals. Patients who do not participate or withdraw from the study will continue to receive the same recommended standard intervention as the participants. It is hoped that the findings and advice given to the patients will go a long way in improving the quality of life

Invitation to participate:

We saw from your results that you are one of those diabetic people with protein in the urine. We are requesting your permission to keep and test your urine for amino acids. Amino acids are the small molecules that build protein.

What is involved in the study – We will not need any more blood or urine from you. However, we will require your permission to go and look at your hospital/clinic file to get information from the time it was known that you had diabetes and the measurements like weight and height. The medication you are taking will also be documented. You will not be expected to make any special visit to Chris Hani Baragwanath Academic Hospital for this study.

Risks: There is no risk at all in this study because the tests result already dobtained here and the those from the urine you have already given will be sent to the laboratory for amino acid analysis.

Benefits - The benefits of this study will only be realized if a certain pattern of amino acids is identified and I will also present it for as a MMed Thesis.

Participation is voluntary – It is your right to accept or refuses to take part in this study and whatever decision you make nothing will change regarding your treatment. It is also your right to tell us to destroy the blood that is already in the laboratory.

CONSENT FOR STYDY ON INVESTIGATION OF URINARY AMINO ACID PATTERNS IN DIABETES

Name----------

Address ----------

Analysis and storage of Biologic Samples: Random Urine Sample

I hereby consent to the removal, storage processing, and analysis of the above material from my own body for the purpose of diagnosis and research into disorders of diabetic nephropathy. After due explanation I understand that

1) Conventional procedures and techniques are employed and that there is no health risk.

2) The material and results of the investigations remain strictly confidential according to medical practice and such ethical guidelines that govern research at the universities and country at large. To preserve anonymity, the samples are coded by numbers and my written consent is required for release of identifiable information to another party

3) Precise diagnosis may not always be possible because the defect(s) may not yet be known or there is inadequate information on roles played by factors.

4) The stored material may be used anonymously in future to derive information or for research purposes. Such future use may be of no direct benefit to the subject.

5) Permission to participate in the study may be withdrawn at any time and any stored biological material will also be destroyed. The withdrawal will not affect the subject's future medical care.

6) I confirm that the purpose of this study was explained in the language that I preferred and I fully understood it. I have not received any financial benefit and I do not expect to receive any in future and undertake not to make any future financial claims following the outcome and publication of the finding.

Singed at ----------day ----------month----------year 20___

Print Name ---------- Signature----------

Witness (Subject's Choice)

Print Name ---------- Signature----------

Witness (Health Personnel)

Print Name ---------- Signature----------