



TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Analytische Lebensmittelchemie

Urine discovery metabolomics of insulin resistant individuals – a holistic approach utilizing hydrophilic interaction liquid chromatography and mass spectrometry

Wendelin F. G. Kranz

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender Prof. Dr. Erwin Grill

Prüfer der Dissertation:

1. apl. Prof. Dr. Philippe Schmitt-Kopplin
2. Prof. Dr. Wilfried Schwab

Die Dissertation wurde am 24.04.2019 bei der Technischen Universität eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 18.11.2019 angenommen.

Dedications

To my daughter Fenja and my wife Juliette.

To my mother Ramona and my father Günther.

To my family.

Acknowledgements

I would like to specially thank Prof. Dr. Philippe Schmitt-Kopplin for his guidance and support.

A special thank also goes to Prof. Dr. Rainer Lehmann for his support, providing of samples and scientific discussions.

Furthermore I would like to thank my colleagues for the pleasant atmosphere and for all the good times we spent in the laboratories as well as in the offices and the know-how and support. I would also like to thank my mother and my family, a special thanks and loving memoires to my father who passed away years ago in 1993.

Finally, I want to thank my my wife Juliette for all her patience, understanding and support during the past years.

The dissertation has been prepared at the research unit Analytical BioGeoChemistry at the Helmholtz Zentrum München – German Research Centre for Enviromental Health in the frame of the Competence Network for Diabetes mellitus funded by the German Federal Ministry of Education and Research (BMBF, FKZ 01GI0804 and 01GI1104A and B).

Summary

A lifestyle intervention program called nutritional prevention of type 2 diabetes mellitus (NUPREDM) study was initiated to elucidate the prospects of type II diabetes prevention by means of a lifestyle intervention. During the program the participants had to maintain a mild calorie deficit (1800 Kcal/day) for 6 month. Additionally, the patient cohort was randomly separated into three different groups i) no red meat, ii) increased fibre consumption and control. Baseline and follow up urine samples of the NUPREDM study were analyzed by aid of a comprehensive non-targeted metabolomics approach including DIA-ICR-FT-MS and a HILIC-MS method, which was newly developed for this purpose. Even though, the lifestyle intervention induced a reduction of average body weight and enhanced insulin sensitivity, no significant differences in the metabolic patterns before and after intervention or between the groups could be identified. During data evaluation of the samples it became abundantly clear that the analysis of urine samples is heavily biased due to the naturally occurring urine concentration variation. This bias could not be compensated via traditional creatinine normalization. Therefore, a poly parametric urine normalization factor (PPUNF) was created, which makes use of creatinine, urea, specific gravity, and UV absorbance values to provide a more robust measure of urine concentration. Additionally, a pre acquisition urine normalization strategy was developed. Briefly, urine concentration was normalized by dilution, according to the PPUNF, before analysis of the sample. Simultaneously, MS response of small and very polar metabolites could be significantly enhance via post column infusion of the ESI-dopant 2-(2-methoxyethoxy)ethanol (2-MEE). The NUPREDM baseline samples were reanalyzed under consideration of the methodological optimizations. Data analysis revealed discriminant urinary metabolic patterns for insulin sensitive and insulin resistant patients. The results obtained during this work demonstrated that non-targeted metabolomics of urine sample is a challenging task. Yet, strategies were developed which can help to vanquish the pitfalls and substantially improve results of analysis in terms of soundness and plausibility.

Zusammenfassung

Ein Lebensstil Interventionsprogramm genannt “nutritional prevention of type 2 diabetes mellitus” (NUPREDM) wurde initiiert um die Erfolgsaussichten einer Typ II Diabetes Prävention mittels Lebensstil Intervention zu untersuchen. Im Studienverlauf sollten die Probanden für sechs Monate ein moderates Kaloriendefizit einhalten (1800 Kcal/Tag). Zusätzlich wurden die Probanden per Zufallsverfahren in drei Gruppen i) kein rotes Fleisch, ii) vermehrte Ballaststoffaufnahme und iii) Kontrolle aufgeteilt. Urinproben, die vor und nach der Intervention gesammelt wurden, sind mittels breit angelegter und nicht zielgerichteten „Metabolomics“ Analytik untersucht worden. Hierfür wurde direktinfusion Analytik mittels Fourier-Transform Ionenzyklotronresonanz Massenspektrometrie (DIA-ICR-FT-MS) sowie eine neu entwickelte hydrophilic interaction liquid chromatography gekoppelt mit MS (HILIC-MS) Methode eingesetzt. Obwohl durch die Interventionsstudie eine Reduktion des Körpergewichts als auch eine verbesserte Insulinempfindlichkeit in den Probanden erreicht wurde, konnten in den Metabolitmustern keine signifikanten Unterschiede zwischen den Gruppen bzw. vor und nach der Intervention identifiziert werden. Während der Datenauswertung wurde zunehmend klar, dass die Analyse von Urinproben stark durch natürlich auftretende Konzentrationsschwankungen des Urins verzerrt wird. Diese Verzerrung konnte nicht durch herkömmliche Normierung auf Kreatinin kompensiert werden. Um über ein robusteres Maß der Urin Konzentration zu verfügen wurde ein polyparametrischer Urin Normalisierungsfaktor (PPUNF) konzipiert welcher auf Kreatinin, Harnstoff, spezifische Gravität und UV Absorption basiert. Darüber hinaus wurde eine Urin Normalisierungsstrategie entwickelt, welche vor der eigentlichen Analyse stattfindet. Diese Strategie sieht vor die Urinkonzentration durch Verdünnung vor der Analyse zu Normalisieren. Parallel dazu konnte die MS Empfindlichkeit für kleine und sehr polare Metabolite durch nach-Säulen Infusion des „ESI-dopants“ 2-(2-methoxyethoxy)ethanol (2-MEE) signifikant erhöht werden. Die Studienproben welche vor der Intervention erhoben wurden, sind unter Berücksichtigung der optimierten Methoden erneut analysiert worden. Hierbei konnten Metabolitmuster identifiziert

werden welche Insulinresistente und Insulinempfindliche Patienten Unterscheiden. Die Ergebnisse dieser Arbeit zeigten deutlich, dass die nicht zielgerichtete „Metabolomics“ Analytik von Urinproben eine Herausforderung darstellt. Jedoch konnten Strategien entwickelt werden um diese Fallstricke zu überwinden und dadurch die Resultate der Analytik in den Bereichen Zuverlässigkeit und Plausibilität wesentlich zu verbessern.

Table of Contents

Acknowledgements	IV
Summary	V
Zusammenfassung	VI
Table of Contents	IX
List of Figures	XII
List of Tables	XVII
List of abbreviations	XVIII
1 Introduction	1
1.1 Diabetes mellitus type II	1
1.1.1 Origins, symptoms, progression of disease and treatment	1
1.1.2 Diabetes early detection and prevention – prospects of life style interventions	3
1.2 Metabolomics	5
1.2.1 Definition and application possibilities.....	5
1.2.2 Standardization, harmonization and commercialization	7
1.3 Tools in metabolomics.....	9
1.3.1 Instrumentation.....	9
1.3.2 Data analysis - pretreatment and statistical evaluation	21
1.4 Scope of the thesis	23
1.5 Outline of the thesis.....	23
2 Nutritional Prevention of Type 2 Diabetes Mellitus (NUPREDM) study	26
2.1 Objectives of the study and study design	26

2.1.1	Samples	27
2.2	Method development for ICR-FT-MS and HILIC-MS analysis.....	27
2.2.1	Materials.....	27
2.2.2	ICR-FT-MS.....	27
2.2.3	HILIC-MS.....	35
2.3	Results.....	53
2.3.1	Results ICR-FT-MS	53
2.3.2	Results HILIC-MS.....	55
2.4	Conclusion.....	56
3	Optimization of HILIC-MS method.....	58
3.1	Normalization of urine concentration via dilution using a Poly Parametric Urine Normalization Factor (PPUNF).....	58
3.1.1	Introduction.....	58
3.1.2	Material and methods.....	64
3.1.3	Results and discussion.....	68
3.2	Enhancing of ESI efficiency using 2-(2-methoxy)ethoxyethanol (2-MEE) as ESI dopant	81
3.2.1	Introduction.....	81
3.2.2	Material and methods.....	82
3.2.3	Results and discussion.....	86
4	Effect of insulin sensitivity on urinary metabolome.....	100
4.1	Analysis of NUPREDM baseline cohort.....	100
4.1.1	Data handling	100

4.1.2	Results and discussion.....	104
5	Concluding remarks	118
6	Appendix	121
7	Literature	158

List of Figures

Figure 1: The most relevant factors in the onset of T2D summarized as the ominous octet of diabetes with permission of (4) (DeFronzo 2009)	3
Figure 2: Characteristics of life connected with and reflected in metabolism	5
Figure 3: Influence of mobile phase flow rate on theoretical plate height (Van deemter plot)	9
Figure 4: Illustration of HILIC retention mechanism with permission from (39).....	12
Figure 5: Schematic illustration of Electrospray ionization in POS mode with permission from (57).....	15
Figure 6: Illustration of different postulated ESI mechanisms i) ion evaporative model (IEM), ii) charged residue model (CRM) and chain ejection model (CEM) with permission from (58)	16
Figure 7: Schematic illustration of FT-ICR MS cell with permission of (65)	17
Figure 8: Cross section of ICR-cell and indicated ion movement after excitation (red) with permission of (65)	19
Figure 9: Intersection of quadrupole-time of flight mass spectrometer Waters Synapt G1 ...	20
Figure 10: Outline of this thesis	24
Figure 11 Comparison between different SPEs and protein precipitation (PPE) in FT-ICR-MS sample preparation The total (ESI+ and ESI-) detected m/z less the ones in the blanks of each method is shown. The amount of m/z's with decimals in the range of x.50 to x.99 which are especially in the lower mass range multicharged, or cluster ions are marked in red.....	28
Figure 12: SPE cartridge comparison of Waters HLB, Agilent PPL and Thermo Sola in FT-ICR-MS sample preparation. The total (ESI+ and ESI-) detected m/z less the ones in the blanks of each method is shown. The amount of m/z's with decimals in the range of x.50 to x.99 which are especially in the lower mass range multicharged, or cluster ions are marked in light grey.	29
Figure 13: Flow-chart of liquid handling robot for SPE sample cleanup	30

Figure 14: Successive calibration of FT-MS spectra with different mass error cutoffs: I: ~20 ppm cutoff, II: 5 ppm cutoff, III: 1 ppm cutoff and IV: 0.4 ppm cutoff	32
Figure 15: Maximum peak intensities of reference standards obtained during HILIC column comparison tests performed in A) ESI-POS mode and B) ESI-NEG mode	38
Figure 16: ESI spray shield after 40 analytical runs of blank (water) injections on Grace Vision HT HILIC column	39
Figure 17: Mass signals caused by column bleeding, detected in blank of Grace Vision HT blank analysis	40
Figure 18: ESI spray shield after 40 analytical runs, including 10 waters blanks and 30 blood plasma samples, injections on Waters BEH HILIC column	41
Figure 19: QC-urine samples analyzed in: POS-mode using urine buffer 2 (I), POS-mode using urine buffer 1 (II), NEG-mode using urine buffer 2 (III), NEG-mode using urine buffer 1 (IV). Additional peaks or peak shape problems are indicated with arrows.	43
Figure 20: Set-up of analytical sequence for stability study	44
Figure 21: Total ion current of standards analyzed via HILIC over seven days of consecutive measurement. (•) first injection of triplicate, (■) second injection of triplicate, (▲) third injection of triplicate. Dotted lines indicate trends.	44
Figure 22: Total ion current of standards analyzed via HILIC-MS over 84 hours of consecutive measurement sample storage in well-plate (•) versus total recovery vial (■).....	45
Figure 23: BPC overlay of five QC urine analyses with HILIC-MS after three blank and three standard injections	47
Figure 24: Average ppm errors determined with , (•) Genedata 6 point lock-mass correction and (✕) Masslynx single point lock-mass correction for creatinine, D3-acetylcarnintine, D3-palmitoyl-carnitine, D4-cholic acid, reserpin and erythromycin.....	49
Figure 25: Flow-chart of sample work-up used for randomization of samples during the study	51
Figure 26: Separation of the NUPREDM follow up individuals in groups 1 (green) and 3 (blue) in the OPLS-DA analysis.	54

Figure 27: Distribution of important mass signal resulting from OSC-OPLS-DA analysis of NUPREDM follow up group 1 vs 3 analysis in chemical categories.	55
Figure 28: Creatinine concentrations of NUPREDM study baseline samples; a potential outlier is marked with a circle.....	59
Figure 29: Retention time of [D10]-adipic acid in I. spiked blank, II. low concentrated urine sample and III. high concentrated urine sample	60
Figure 30: Exemplary Base peak chromatograms (BPC) of high and low concentrated urine samples, peak deteriorations putatively caused by high urine concentration is indicated with black arrows.....	60
Figure 31: Schematic illustration of PPUNF calculation.....	63
Figure 32: Representative HILIC-UV chromatograms from NUPREDM baseline cohort samples before smoothing 1A & 1B, 2A & 2B corresponding HILIC-UV chromatograms after smoothing and integration.....	68
Figure 33: Correlation analysis of HILIC-UV area under the curve and urinary creatinine ..	69
Figure 34: Scatterplots of potential parameters (urinary creatinine, urea, UV-AUC, LC-MS-TUS, NRM PQN and specific gravity) for urine normalization determined in NUPREDM baseline samples	70
Figure 35: Correlation matrix of potential parameters for urine normalization determined in NUPREDM baseline samples: Upper right corner indicates Pearson correlation values, and p values (bonferoni corrected) are displayed on the lower left side	71
Figure 36: Method comparison plots of validation study samples. A-1 and B-1 are Passing–Bablok regression analysis of ACR and PPUNF data, respectively. A2- and B-2 are Bland–Altman plots of ACR and PPUNF data, respectively	75
Figure 37: Scatterplott difference in % of albumin PPUNF to 24-h-UAE over PPUNF inconsistency in validation study samples.	76
Figure 38: Boxplots of internal standard retention time stability in HILIC-MS analysos of NUPREDM baselin samples with and without pre-acquisition normalization	79

Figure 39: OPLS-DA score plot of urinary HILIC-MS gender study: A) with prAN samples, B) with poAN samples	80
Figure 40: Dilution scheme for direct infusion experiment during 2-MEE dopant evaluation	85
Figure 41: ESI-responses (peak height in counts) to various buffer types and concentrations of different standards, a) L-leucine, b) D10-adipic acid, c) D6-sulfadimethoxine, d) D4-cholic acid.	87
Figure 42: Obtained Dopant factors from HILIC-MS analyses using different ammonia acetate concentrations of standard mixture A	88
Figure 43: Dopant factor over different physiochemical parameters a) molecular weight, b) logarithmic water-octanol partition coefficient of unionized solutes (logP), c) isoelectrical point, d) molecular volume, e) molecular polarizability, f) logarithmic water-octanol partition coefficient at pH6. y-Axes are ln-scaled.....	89
Figure 44: Dopant factor over HMDB metabolite taxonomy class. Individual dots depict the dopant factor of mass annotated peaks and their corresponding HMDB metabolite taxonomy class resulting from non-targeted HILIC-MS analysis of human urine samples.....	91
Figure 45: Influence of buffer additives and 2-MEE on signal abundances (with A= acetonitrile:water (2:1v/v)).....	93
Figure 46: Comparison of stable isotope labeled amino acids peak areas in a) blank and b) and c) biofluids with and without post column infusion of 2-MEE.	95
Figure 47: Dopant factor in spiked blank, urine and plasma showing differences in the matrix effect but systematic increase in sensitivity	96
Figure 48: Obtained peak areas of short chain fatty acids with different concentrations of post-column 2-MEE infusion	97
Figure 49: A) 3D and B) 2D plot of PCA results created from NUPREDM baseline HILIC-MS data without batch normalization	102
Figure 50: A) 3D and B) 2D plot of PCA results created from NUPREDM baseline HILIC-MS data with batch normalization	102

Figure 51: Score plot of OPLS model from DIA-FT-ICR-MS data for ISI-Matsuda colored by classes high (red) and low (blue) 104

Figure 52: Whisker-Box-plots for selected exemplary metabolites discriminant for low and high insulin sensitive subjects a) tryptophan, b) glycyglycylhydroxyproline, c) suberylglycine and d) cortolone-3-glucoronide 106

Figure 53: Score plot of OPLS model from HILIC-MS data for ISI-Matsuda colored by classes high, intermediate and low 107

Figure 54: Pathway analysis of significantly altered metabolites from OPLS-DA analysis . 108

List of Tables

Table 1: Urine SPE elution profile.....	30
Table 2: Tabulated overview of conducted group comparisons.....	34
Table 3: HILIC column tested for method optimization	36
Table 4: Retention time stability during 3 days of consecutive measuring	47
Table 5: Results of multivariate statistic models	53
Table 6: Comparison of minimum normalized parameters of extreme subjects	72
Table 7: Linear regression analysis of minimum normalized urine concentration estimators	73
Table 8: Observed minimum references in comparison with literature data	77
Table 9: Urinary metabolites reflecting ISI-matsuda changes identified by OPLS-DA.....	105
Table 10: Sialic acids identified as significant in patients with low ISI-Matsuda	115

List of abbreviations

2-MEE	2-(2-methoxyethoxy)ethanol
24-h-UAE	24-h urinary albumin excretion
ACR	albumin-to-creatinine ratio
NH ₄ OAc	ammonium acetate
AUC	area under the curve
BGC	Bio-GeoChemistry
BPC	base peak chromatograms
BMI	body mass index
CEM	chain ejection model
CRM	charged residue model
CoA	coenzyme A
CRO	contract research organization
Da	Dalton
DIA	direct injection analysis
ESI	electrospray ionization
FT	fourier transformation
FWHM	full width half maximum
GDM	gestational diabetes mellitus
HMDB	human metabolom database
HILIC	hydrophilic interaction chromatography
IDF	international diabetes federation
ICR	ion cyclotron resonance
IEM	ion evaporative model
LC-	
MSTUS	LC-MS total useful signals
LC	liquid chromatography
MS	Mass spectrometry
MW	megaword
MCA	metabolic control analysis
NAFLD	non-alcoholic fatty liver disease
NMR	nuclear magnetic resonance
NUPREDM	nutritional prevention of type 2 diabetes mellitus
	orthogonal partial least squares – discriminant
OPLS-DA	analysis
PPE	protein precipitation
ppm	parts per milion
PPUNF	poly parametric urine normalization factor
PCA	prinicipal component analysis
PQN	probabilistic quotient normalization
PLS	projection to latent structures
Q-TOF	quadrupole time of flight
QC	quality control
RF	radio frequency
RP	reversed phase

SFCA	short chain fatty acids
S/N	signal to noise ratio
SPE	solid phase extraction
SG	specific gravity
TOF	time of flight
TIC	total ion current
TFA	trifluoro acetic acid
TULIP	Tuebingen University life style intervention program
T2D	type 2 diabetes
UPLC	ultra pressure liquid chromatography
UV	ultraviolet
VIP	variable influence on projection
WHO	World Health Organization

1 Introduction

1.1 Diabetes mellitus type II

1.1.1 Origins, symptoms, progression of disease and treatment

The earliest written records on diabetes dates back to more than 3000 years ago, when Egypt and Indian physicians describe symptoms of diabetes like extreme thirst, excessive urination and sweet tasting urine. The name “Diabetes” goes back on Apollonius of Memphis, diabetes means “to pass through” (dia – through, betes – to go) since he and his coevals mistakenly thought diabetes was a disease of the kidneys. Today, Diabetes mellitus is one of the major health threats, which our modern culture has to manage. In fact the World Health Organization (WHO) estimates that diabetes will be the 7th leading cause of death in 2030 (1) and considered high blood glucose the third highest risk factor for premature mortality already in 2004 (2). According to the WHO: diabetes is a chronic disease that occurs either when the pancreas does not produce enough insulin or when the body cannot effectively use the insulin it produces.

Dependent on the origin of the disease diabetes is subdivided into three main types:

- type 1 diabetes
- type 2 diabetes (T2D)
- gestational diabetes

But also other, rarely occurring, types of diabetes are known.

Diabetes type 1 is allocated in the area autoimmune diseases. The immune system of a patient, suffering from diabetes type 1, directs itself against the insulin-producing beta cells in the pancreas. In the final stages of diabetes type 1, the ability of the body to produce insulin, and thereby the ability of blood glucose regulation, is complete lost.

Gestational diabetes mellitus (GDM) is a state in which pregnant women develop an insulin resistance. Gestational diabetes is a result of an altered hormone balance during pregnancy, usually symptoms disappear after giving birth.

Diabetes type 2 is the most common form of diabetes. It is estimated by the international diabetes federation (IDF), that in 2015 415 million people suffer from T2D, and T2D was hold responsible for 5.0 million deaths. (3) In contrast to the other major forms of diabetes, type 2 diabetes is the consequence of a lifestyle involving excessive calorie intake and low physical activity. In brief, T2D results from a permanent state of over nutrition and a lack of physical exercise. This combination leads to a constant state of hyperglycemia. Since high levels of blood glucose can have severe adverse health effects i.e. on kidney or arterial vascular function, the body tries to lower blood glucose via excretion of insulin which stimulates the uptake of glucose into tissues. A permanently increased insulin release induces an insulin resistance in the glucose up taking tissues, meaning that the same amount of insulin stimulates a lower glucose uptake into tissues. Over time, hyper-secretion of insulin wears out the β -cells which finally results in complete β -cells failure.

This is just a perfunctory explanation, detailed mechanism of T2D onset and manifestation appears to be a complex interaction of impaired glucose homeostasis driven by different organs, hormonal feedback, hormonal disbalance, fatty acid and lipid metabolism, lifestyle and genetic disposition. (4-6) DeFronzo summarized the most relevant factors entitled as the "ominous octet" brain, liver, kidney, pancreas, islet α -cells, gastrointestinal tissues, fat cells and muscle cells (see Figure 1).(4)

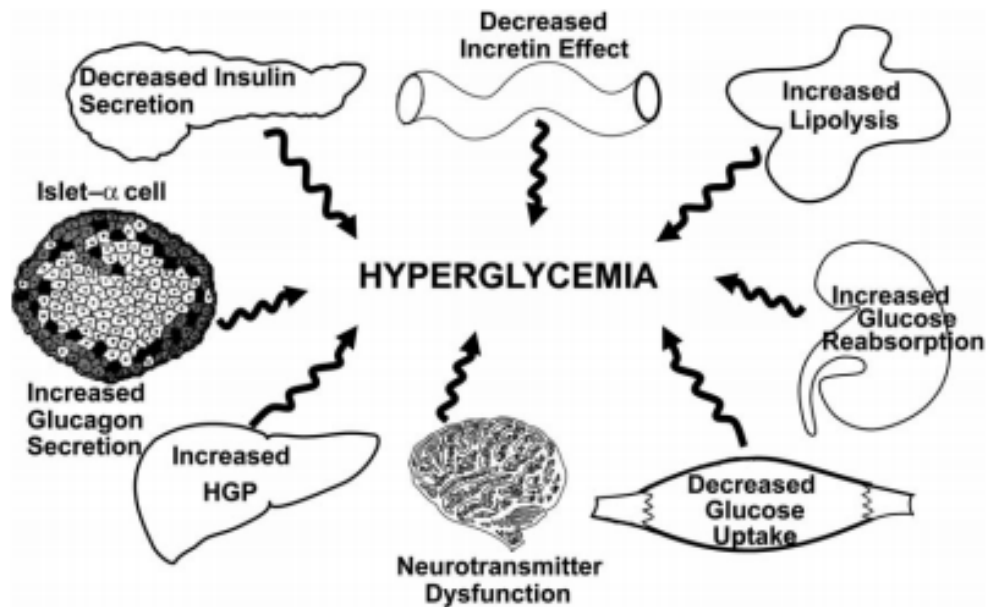


Figure 1: The most relevant factors in the onset of T2D summarized as the ominous octet of diabetes with permission of (4) (DeFronzo 2009)

1.1.2 Diabetes early detection and prevention – prospects of life style interventions

A promising approach to counteract pathogenesis of T2D, is to address one of the root causes. Ultimately, the root causes of T2D can be reduced to genetic and environmental factors, whereby the possibilities of exerting an influence on the genetic factors are hardly given. The environmental factors offer a variety of possible approaches. The most common ones: administration of pharmaceuticals (e.g. insulin-sensitizer like metformin or glitazones), bariatric surgery, change of diet, dietary restriction and physical exercise, proofed to be effective in preventing onset or manifestation of T2D. (7-9) Pharmaceuticals and bariatric surgery have proven to be effective. However, pharmaceuticals exhibit unwanted side effects (10) and bariatric surgery is a massive surgical intervention and complications occur regularly (10-17%) (11).

A change of diet, dietary restriction and physical exercise do not show adverse reactions and are therefore the less invasive measure to prevent diabetes. Recent epidemiological studies could demonstrate that a dietary change can have positive effects on preventing onset or

manifestation of T2D. Two pivotal elements of nutrition draw particular attention of the researchers; dietary fibre and red meat consumption.

1.1.2.1 Dietary fibre

Foods rich in dietary fibre, like whole grains, were associated with reduced risk of T2D. (12, 13) Also Weickert et al. observed a positive effect of dietary fibre on T2D; furthermore a positive effect on insulin sensitivity, postprandial glucose response and total LDL cholesterol was discovered. (14) But not all dietary fibre is the same, dietary fibre can be divided into soluble and insoluble fibre. Most of the protective effects on T2D is attributed to insoluble fibre. (15, 16)

There are several reasons for the positive effects of dietary fibre. Food rich in fibre have a larger volume of food, which causes longer chewing and satiety occurs earlier. Also, stomach drainage is delayed and uptake of makro nutrient is slower. (15) Besides these aspects, dietary fibre is converted by gut microbiota to short chain fatty acids (SFCA) which in turn have a positive effect on regulation of glucose metabolism. (17)

1.1.2.2 Red Meat

Besides the most obvious foods known to increase the risk to develop T2D e.g. sugar, refined starch and fat also red meat gained attention of researchers. There is a plethora of publications associating an increased red meat consumption with an increased risk to develop T2D. (18-21). Especially the consumption of processed red meat (e.g. smoked, dry aged or cured red meat) is attributed to increase the risk of T2D. (22, 23)

1.2 Metabolomics

1.2.1 Definition and application possibilities

Metabolomics is a neologism assembled from the greek words “metabolism” (μεταβολή = change) and the suffix “-omics” which is used in science (*genomics*, *proteomics*) to emphasize that the preceding compound class is studies in a quantitative, qualitative comprehensive and holistic manner. Metabolism is one of the attributes that separates living from non-living objects. It is the process that provides the necessary energy via chemical conversion of substrates to maintain the other essential characteristics of life Figure 2



Figure 2: Characteristics of life connected with and reflected in metabolism

The first studies on metabolism were conducted by Santorio Sanctorius, who is considered the designated “Founding father of Metabolic Balance Studies”. (24) He performed studies on

a phenomenon he called insensible perspiration, (25) meaning the difference in weight of ingested food and excreted feces.

Analog to the concepts of genome or proteome, the term metabolom was introduced by Oliver et al., when he and his group came up with the idea of a metabolic control analysis (MCA) - “the measurement of the change in the relative concentrations of metabolites as a result of the deletion or overexpression of a gene.” (26)

Nicholson and co-workers are considered as the initiators of modern metabolomics but calling it at that time metabonomics “the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification” (27). Later the term metabolomics (“a comprehensive analysis, in which all the metabolites of a biological system are identified and quantified”) which was introduced by Fiehn prevailed. (28).

So far, the actual size of the human metabolom is unknown but the number of identified metabolites is steadily increasing. For example, the human metabolom database (HMDB) contained 2180 metabolites, when launched in 2007.(29) After several updates (e.g. v2.0 in 2009 6408 metabolites and v3.0 in 2013 40,000 metabolites) it currently lists 114,101 metabolites in v4.0 published in 2017.(30, 31)

Metabolites are molecules with a molecular weight below 1500 Da and intermediate or end products of metabolic reactions. But metabolites do also exert signaling properties for example in the class of lipid metabolites polyunsaturated fatty acids, such as arachidic acid or phosphatidylinositols are the most prominent candidates.

Over time, different nomenclatures (metabonomics, metabolomics, metabolic fingerprinting, metabolic profiling) and subcategories of metabolomics (lipidomics, glycomics or peptidomics) were emerging. In general the “type” of metabolomic study can be differentiated by the approach, meaning hypothesis or data driven, the type of metabolites and the number of metabolites accounted for.

1.2.2 Standardization, harmonization and commercialization

More and more efforts are spent in the community of metabolomics to reach a higher level of standardization and harmonization between different laboratories. A higher degree of standardization and harmonization is deemed necessary to increase general quality and reproducibility of results and to allow inter-lab comparability and validation of results.

Standardization and harmonization concern several particular aspects of metabolomics:

- Sampling
- Sample preparation
- Chromatographic system (if applicable)
- Data treatment
- Data analysis

One of the mature bottle-necks in metabolomics is the annotation and identification of peaks. Metabolomic datasets often contain thousands to tens of thousands of peaks defined by retention time-m/z pairs. Linking these peaks to actual metabolites with help of publicly available metabolite databases like HMDB (30, 31), Metlin (32, 33) or KEGG (34) can be tedious. Therefore many groups started to set-up their own home-made metabolomics retention time databases. To increase reliability of identification, these databases often get augmented by additional MSⁿ data.

The striving for harmonization and standardization and the complexity of metabolomic studies paved the way for commercialization of metabolomics. Today many metabolomic studies are carried out with help of kits or contract research organizations (CRO) like Metabolom or Biocrates. The use of commercial kits or hiring of a CRO facilitates simplification of analysis and produces reliable and comparable results, yet these benefits are obtained by the cost of an in depth-understanding of LC-MS instrumentation and the method itself plus the limitation to the metabolites covered by the underlying database meaning that no other metabolites can be determined.

In contrast to the other major –omics sciences the metabolome is not coded by a relatively limited set of building blocks like the four nucleobases in genomics or the 20 amino acids in proteomics. Metabolites are such heterogeneous in their physiochemical properties that a single method for characterization of all metabolites does not exist. Exclusive application of highly standardized methods may have the consequence to miss out on novel findings or metabolites, therefore “thinking out of the box” and application of non-targeted and therefore non-biased screening approaches have a legitimate place in metabolomics sciences.

1.3 Tools in metabolomics

The sheer number of metabolites and the chemical diversity presents a major challenge to analytical chemistry, the methodologies used in the course of this work are briefly described in the following chapter.

1.3.1 Instrumentation

1.3.1.1 Hydrophilic interaction chromatography (HILIC)

Chromatography is the separation of a mixture dissolved in a mobile phase into its individual components. Therefore the mobile phase is passed by a stationary phase. Differential partitioning coefficients of the components into mobile and stationary phase causes different traveling speeds along the stationary phase resulting in a separation of the mixture longitudinal to the stationary phase. The separation efficiency and underlying principles are described in the van Deemter equation (see Figure 3). (35) The separation efficiency is typically described in theoretical plate height (the smaller the better) and dependant on mobile phase flow rate. (35)

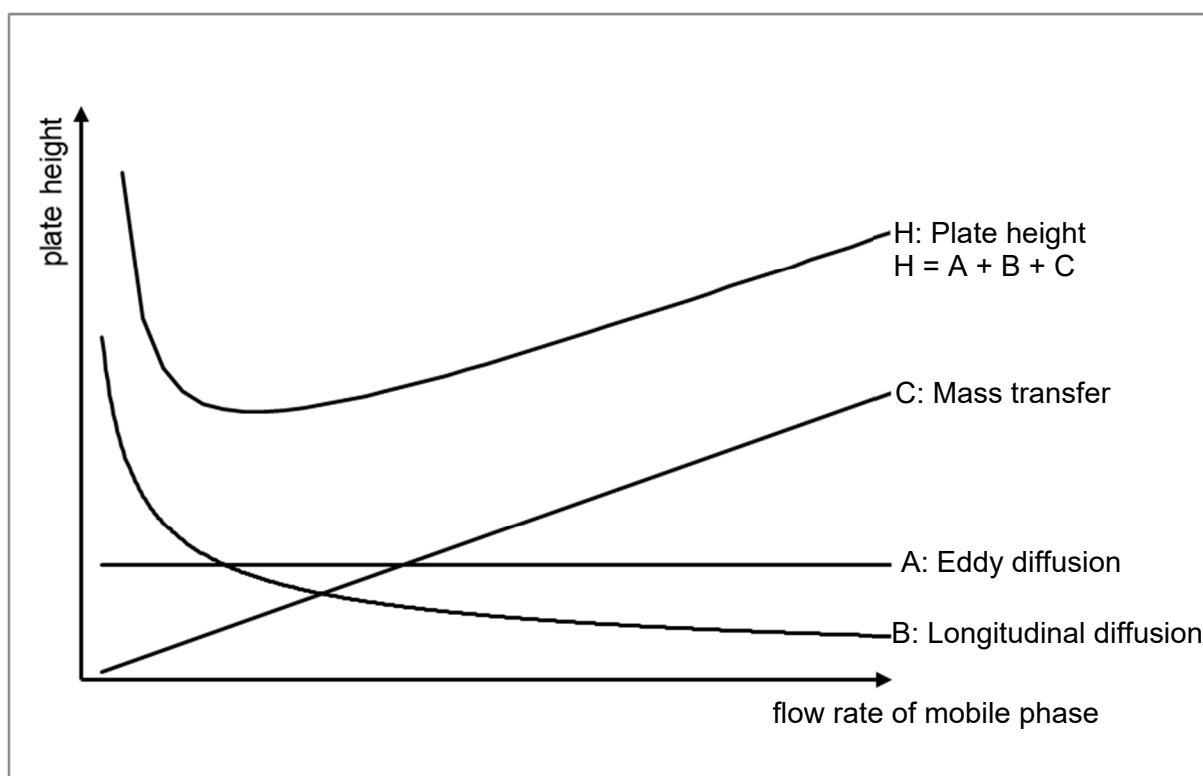


Figure 3: Influence of mobile phase flow rate on theoretical plate height of chromatography (Van Deemter plot)

Chromatography is composed of two greek words “chroma” which means color and “graphein” which means writing and Mikhail Tswett is considered as the father of modern chromatography, yet there are indications that already before him scientists used the technique. (36) Mikhail Tswett was separating plant pigments using calcium carbonate as stationary phase and ether-ethanol mixtures as mobile phase. In modern science chromatography is separated in two main chromatography types i) gas chromatography and ii) liquid chromatography. Both chromatography types can be further sub-classified especially in liquid chromatography many different forms exist, the most important ones for metabolomic studies are i) reversed phase (RP) chromatography and ii) hydrophilic interaction chromatography (HILIC). Reversed phase chromatography is typically performed with a porous silica based carrier material derivatized on the surface with octadecyl ligands. Due to octadecyl chains attached to the surface the polarity of the silica base material is reversed from hydrophilic to hydrophobic, hence it is called reversed phase (RP) chromatography. RP chromatography is frequently used to separate semi-polar to apolar analytes.

Yet, target of the study described within this manuscript were urine samples of pre-diabetic patients. Unlike in plasma, most of the urinary metabolites are highly polar. (37) For highly polar molecules hydrophilic interaction chromatography (HILIC) is better suited. The term HILIC was firstly introduced in 1990 by Andrew J. Alpert. He described: “ When a hydrophilic chromatography column is eluted with a hydrophobic (mostly organic) mobile phase, retention increases with hydrophilicity of solutes.” (38) In HILIC the elution order is opposite to RP chromatography, making it the method of choice to analyze small and very polar analytes like amino acids or mono- and oligosaccharides.

The retention mechanism behind HILIC retention is not yet fully understood, but a consensus in the scientific community exists on some of the major retention mechanisms. It is hypothesized that HILIC retention mechanism is a mixture of different interaction mechanisms the most important are:

- Partitioning
- Electrostatic interactions
- Adsorption

The partitioning mechanism is caused by a layer of water molecules that adsorbs onto the hydrophilic stationary phase material. Polar solutes partition between this immobilized water rich layer and the organic solvent rich mobile phase (see Figure 4 (39)), hence retention of a molecule increases with its polarity. (38, 40) To establish a stable water rich layer the mobile phase has to remain at least 3% of water. (41) Electrostatic interactions are a result of partially charged residues from the column bed material, i.e. silanols, amines. These electrostatic interactions can exhibit retentive but also repulsive effects. For example organic acids analyzed in HILIC mode under neutral pH conditions on a bare silica column can elute even before the dead volume of the column, since they are repelled from the also negatively charged silanol groups. (41, 42) Another example is the retention of amines which can be completely eliminated if mobile phase is acidified with TFA. Addition of TFA lowers the mobile phase pH which protonates the silanol groups and forms neutral ion pairs with the amine. (41-43) Besides the previous described ion-exchange interactions also hydrogen bonding plays an important role in HILIC. (40)

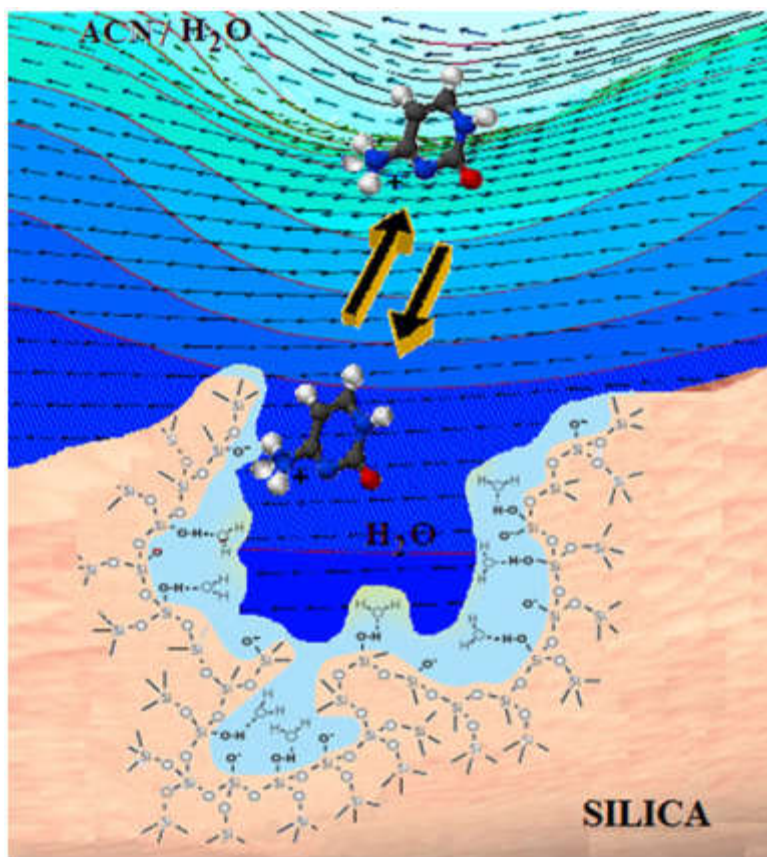


Figure 4: Illustration of HILIC retention mechanism with permission from (39)

Today, a plethora of different HILIC column types is commercially available, the most popular ones are based on:

- bare silica
- amine
- amide
- diol or
- zwitterionic (sulfobetaine, phosphocholine)

ligands, but also other HILIC column types exist like the “chocolate” HILIC phases (44). Each of the column types possesses individual surface characteristics and therefore a unique selectivity. Bare silica columns for instance exhibit weak cation exchange interactions capabilities and are therefore well suited for analysis of amines. (45) Amide or amine bonded

columns are often used to separate saccharides and zwitterionic phases are popular for separation of anions. (41)

Hydrophilic interaction chromatography also found its way into metabolomics studies for analysis of urine (46-48), plasma (47, 49) or other biological samples like amniotic fluid (47), human sweat (50) or cell lysates (51). To cover a broader range of metabolites from highly polar ones like amino acids to highly hydrophobic ones like lipids there are also attempts to combine both methods via multidimensional chromatography.(52-56)

1.3.1.2 Mass spectrometry

Mass spectrometry (MS) is a versatile tool used in many disciplines of the natural sciences. It allows researchers to separate ionized molecules according to their mass-to-charge ratio (m/z), hence giving information on the molecular sum formula of the analyzed molecule. Together with other analytical tools like liquid chromatography and a priori knowledge about the sample, conclusions on molecular structure and composition of a sample can be drawn. Thus, high resolution MS helps to resolve complex biological samples like plasma or urine and supports scientists to gain a detailed inside into the biological context of the sample. Performance of a mass spectrometer is measured in a set of performance parameters like

Mass resolution is measured in full width half maximum (FWHM) defined by Equation 1

$$R = \frac{M}{\Delta M}$$

Equation 1: Definition of mass resolution. R = resolution, M = mass, ΔM = peak width at 50% height of the peak

Mass accuracy accuracy is the discrepancy of a measured m/z signal to the nominal m/z value and is typically expressed as relative error in part per million (ppm) for high resolution MS (see Equation 2).

$$\text{relative mass error} = \frac{m/z \text{ measured} - m/z \text{ nominal}}{m/z \text{ nominal}} * 10^6$$

Equation 2: Definition of mass accuracy

1.3.1.3 Electrospray ionization (ESI)

Most of the biological samples of interest are in the liquid form (urine, plasma, cell lysate...). To allow mass spectrometric analysis of the molecules dissolved in the sample, the solvent needs to be evaporated and the molecules need to be equipped with a positive or negative charge. In the past, many ionization techniques have been developed, yet over time electrospray ionization evolved to the most popular and most frequently technique among the so called soft ionization techniques. (57) In ESI the liquid sample is feed through a thin (typically metal) capillary. The capillary is charged with a potential of up to 4 kV and positioned close to a counter electrode. The solvent that is ejected at the end of the capillary is evaporated with help of a heated inert gas (nitrogen). The liquid forms a tip, the so called taylor cone, which gets thinner with increasing distance from the capillary until small droplets are formed. The droplets are charged on the surface and shrink further due to ongoing process of solvent evaporation until the surface charge rises over a certain level and coulumb explosion occurs and even smaller droplets are formed (see Figure 5).

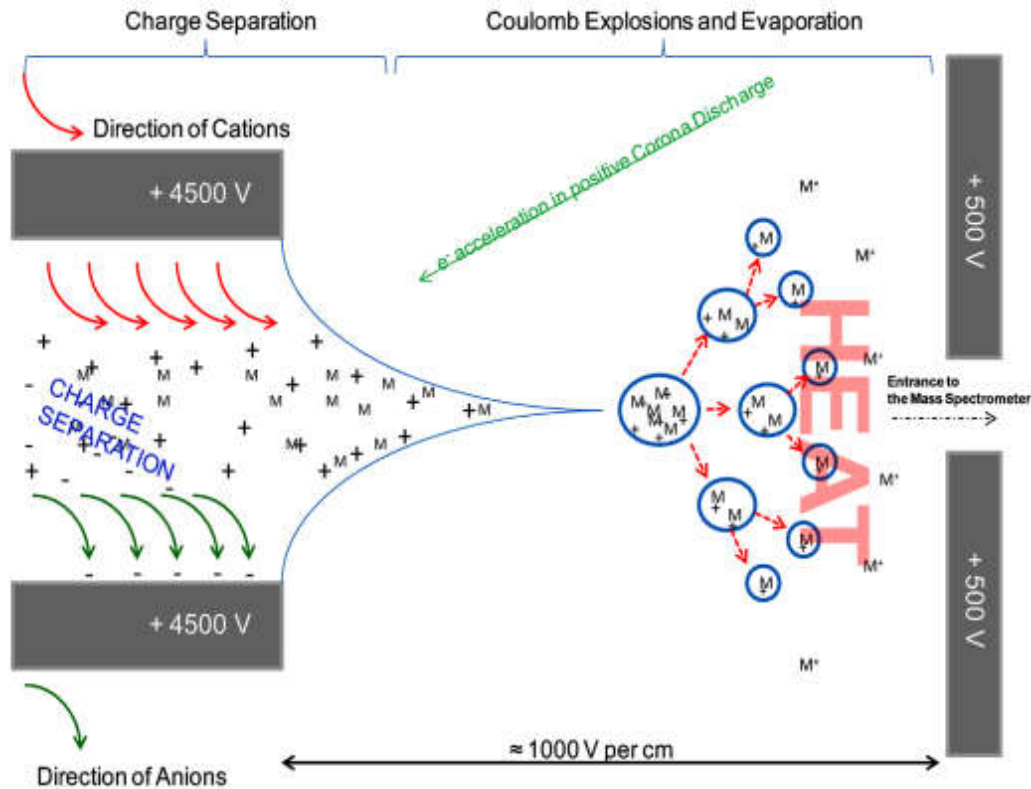


Figure 5: Schematic illustration of Electrospray ionization in POS mode with permission from (57)

These processes of droplet shrinking and coulomb explosions finally lead to charged molecules in the gas phase. Currently, there are different models that explain how the transfer from droplet to ionized molecule actually works called i) ion evaporative model (IEM) for low molecular weight molecules, ii) charged residue model (CRM) large molecules and chain ejection model (CEM) for polymers (see Figure 6). (58)

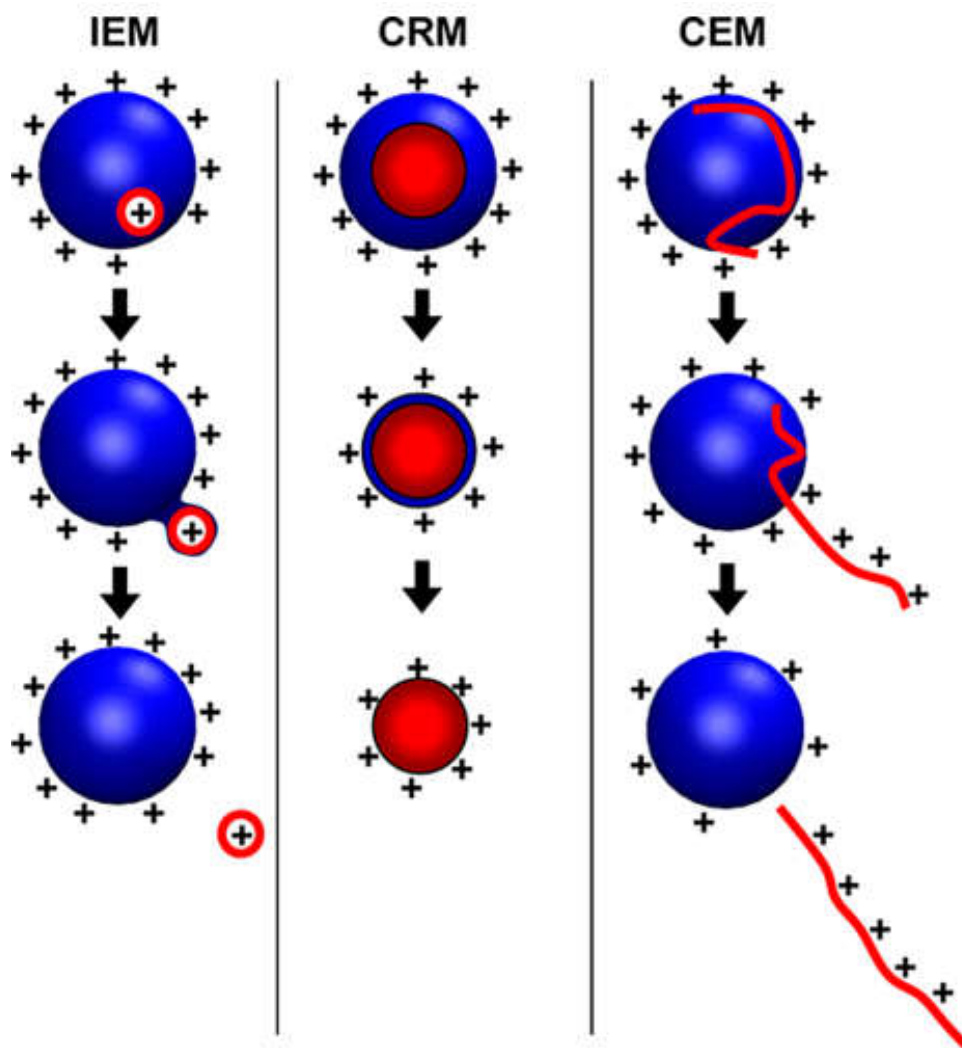


Figure 6: Illustration of different postulated ESI mechanisms i) ion evaporative model (IEM), ii) charged residue model (CRM) and chain ejection model (CEM) with permission from (58)

The applicable ESI liquid flow rate ranges from few nL/min (nano-ESI) to up to 1 mL/min which makes ESI also a powerful technique to hyphen the world of liquid chromatography with mass spectrometry. (57, 58)

1.3.1.4 Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS)

The development and history of FT-ICR MS was nicely reviewed by Marshal in 2014. (59) Theory and working principles of FT-ICR MS are described in detailed in literature. (59-64) Yet, a brief introduction shall be given in this chapter. Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) offers unsurpassed mass accuracy (<0.100 ppm) and mass resolution (up to 1.000.000 FWHM). The high mass accuracy and resolution is a result of the detection principle.

Ionized molecules are generated, for instance by ESI, focused and transferred into the FT-ICR analyzing cell. Typically, a quadrupole located in the ion optics between ESI chamber and ICR cell additionally allows for isolation of specific ions. The generated ions are then trapped in the analyzer cell. The FT-ICR analyzer cell is a cylinder segmented in 4 parts (see Figure 7 and Figure 8) placed in a special uniform magnetic field of several Tesla. The magnetic field is typically established with help of a superconducting magnet.

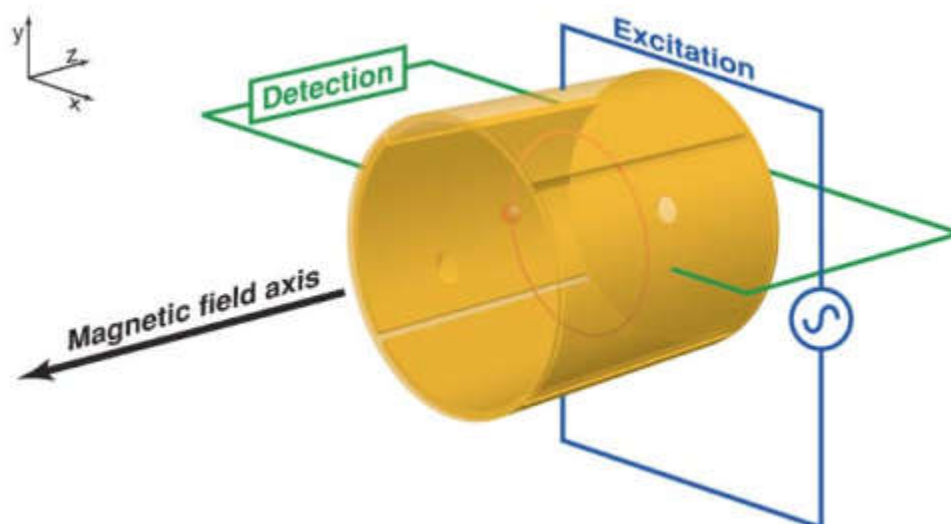


Figure 7: Schematic illustration of FT-ICR MS cell with permission of (65)

The ions, once transferred into the ICR cell, will start to rotate around the center of the magnetic field also known as “cyclotron motion”. The cyclotron motion is a product of the Lorentz force (see Equation 3),

$$F = q * v * B$$

Equation 3: Definition of the Lorentz force F , q = charge of the molecule, v = velocity of the molecule and B = magnetic field

The cyclotron frequency is a function of the mass-to-charge ratio of a molecule and the magnetic field strength and described in Equation 4

$$\omega = \frac{q * B}{m} \quad f = \frac{q * B}{2\pi m}$$

Equation 4: Definition of cyclotron frequency, m = mass, w = angular velocity and f = frequency

After introduction of the ions into the cell two of the four segments of the FT-cell are used to introduce a radio frequency (RF) which excites the charged molecules, hence these segments are called excitation plates. The excitation radio frequency forces the molecule ions to rotate in higher trajectories, this brings them close enough to the other two segments call detection plates to be detected (see Figure 8).

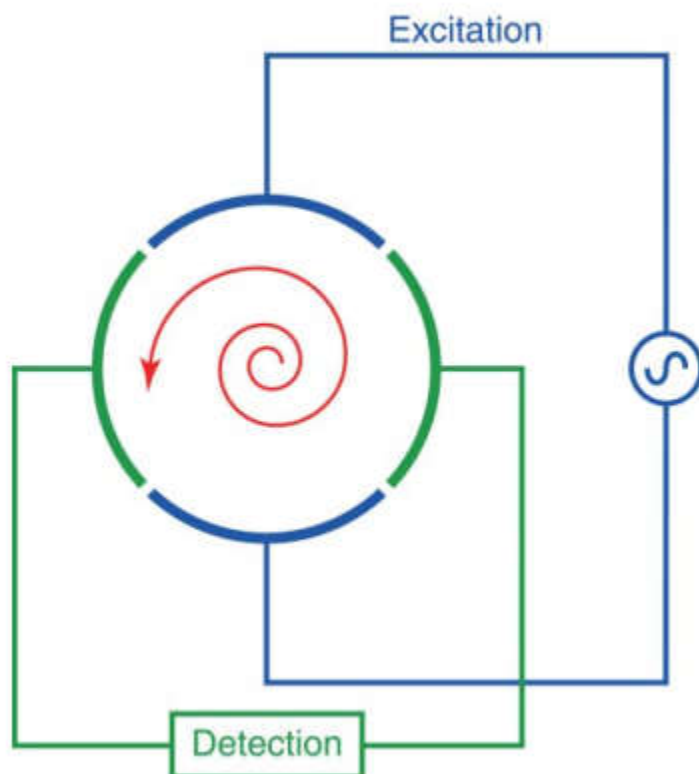


Figure 8: Cross section of ICR-cell and indicated ion movement after excitation (red) with permission of (65)

Each ion rotates at a certain frequency which is proportional to its mass-to-charge ratio. Excitation of several molecules of different mass-to-charge ratio results in an overlay of different frequencies. These frequencies are detected by the detection plates as an interferogram. The interferogram is subsequently deconvoluted by Fourier transformation to a mass spectrogram.

Despite the outstanding mass accuracy and resolution of FT-ICR MS, the method is unable to resolve chemical isomers (like all other MS techniques) which often exist in biological samples (i.e. leucine-isoleucine). Additionally, if very complex samples are analyzed ion suppression or space charge effects can occur. Both problems are usually solved via coupling of a preceding liquid chromatography step. For FT-ICR MS LC coupling is not possible since the scanning time required to achieve such high mass resolution would not allow to sufficiently resolve chromatographic peaks. Therefore, workarounds have been developed. For example

Li. et al. used liquid chromatography with fraction collection. The collected fractions were subsequently analyzed with direct-infusion (DI) FT-IC-MS. (66)

1.3.1.5 Time-of-flight mass spectrometry

Time-of-flight mass spectrometers (TOF-MS) were continuously improved since the first TOF-MS build in 1953 (67, 68). Modern commercially available TOF-MS instruments offer a mass resolution of more than 50,000 and mass accuracy of less than 1 ppm. Compared to FT-ICR MS (see 1.3.1.4) these parameters appear relatively modest, yet modern TOF-MS instruments are considered as high resolution (HR) MS. The setup of the Synapt Q-TOF used in this work is shown in Figure 9.

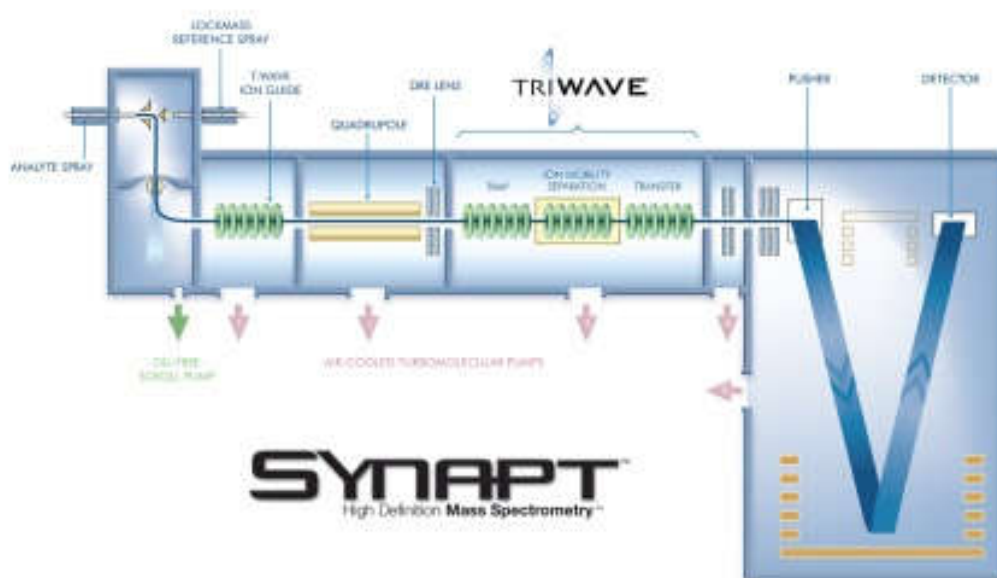


Figure 9: Intersection of quadrupole-time of flight mass spectrometer Waters Synapt G1

The underlying TOF principle is rather simple, ionized molecules get accelerated towards a drift path of defined length. The time that the ions need to travel past the drift path is proportional to its m/z . The time an ion needs to pass the travel path is defined by Equation 5. (64)

$$TOF = d * \sqrt{1/2 zeV} * \sqrt{m/z}$$

Equation 5: relation between time-of-flight and mass-to-charge ratio; d = distance; zeV = kinetic energy in electron volts

The advantage of time-of-flight mass analyzers over other instruments are i) the high scanning speed of up to 20 Hz which allows hyphenation UPLC and TOF-MS and ii) the high mass resolution and accuracy compared to other fast scanning MS techniques like quadrupole MS. As the Synapt G1, most of the TOF-MS commercially available are actually a combination of different mass analyzers. A quadrupole and a collision cell in the ion-path of the MS allows for isolation of specific m/z ions followed by fragmentation. This technique called Q-TOF facilitates a deeper structural evaluation of analyte ions via MS-MS experiments.

1.3.2 Data analysis - pretreatment and statistical evaluation

Data driven metabolomics screenings with aid of LC-MS instrumentations generates tremendous amounts of data. On top of that, obtained signal intensities may vary on very high to very low levels. To analyze such data the retention time and m/z information are detected, summarized to peaks and aligned in a matrix in which every detected metabolite is cataloged by retention, m/z value and signal intensity for the respective sample. (69, 70) In the follow up, the data is usually pre-treated: filtering, data centering, scaling or transformation are the most commonly used techniques to avoid bias in statistical tests. To extract meaningful information multivariate statistics have proven to be a valuable tool. These techniques are able to detect patterns within the data and visualize the results. (69, 71) The most commonly applied approaches can be divided in two approaches i) unsupervised and ii) supervised methods. (69-71)

Prinicipal component analysis (PCA) is probably one of the most frequently applied unsupervised multivariate statistic approach. In brief, the PCA algorithm identifies covariation or correlation between variables in the data-set and groups them together into principal

components. This way PCA reduces the dimensionality of the dataset and recognizes the major trends in the dataset. Principal component analysis is often used as an exploratory way to become a better understanding of the data, yet only the most distinct trends, which are not necessarily the ones of interest, are reflected in PCA.

Projection to latent structures (PLS) and related techniques like orthogonal-PLS (OPLS) (72) probably belongs to the most frequently applied supervised multivariate statistic approach. Projection to latent structures combines aspects of PCA and multiple linear regression. (73)

In PLS an independent variable is added which could be for example a clinical marker like insulin sensitivity. The PLS algorithm tries to find a principal component in the multivariate dataset (i.e. LC-MS data from urine) that shows a maximum in covariance between the independent variable and the principal component. This way, the PLS algorithm extracts solely variables which are predictive or discriminative for the independent variable, in our example, LC-MS peaks that discriminate insulin sensitive from insulin resistance patients.

1.4 Scope of the thesis

Scope of the thesis was the development of an analytical platform which enables the analysis of urine samples towards identification of metabolites and metabolomic patterns differentiating healthy and pre-diabetic patients. To reach this goal UPLC-MS and ultra high resolving FT-ICR-MS were used. The research was conducted in the framework of the Competence Network Diabetes and samples were collected from the Tuebingen University life style intervention program (TULIP).

1.5 Outline of the thesis

In the present work ultra high resolution analytical instrumentation as well as multivariate statistical approaches were utilized to shed a light on metabolism of pre-diabetic individuals. To provide an overview the thesis is summarized in the following section (see also Figure 10).

Chapter 1: A brief overview and introduction on background theory of diabetes and metabolomics as well as on the applied analytical and statistical techniques is given.

Chapter 2: The scope and objectives of the nutritional prevention of type 2 Diabetes Mellitus (NUPRDEM) study are described. Further, method development is detailed and results of the study sample analysis are discussed. Unfortunately, no clear result could be obtained from the analysis of the NUPRDEM study therefore further experiments were conducted with the samples from the NUPRDEM study cohort and considered separately in chapter 5.

Chapter 3: During analysis of the NUPRDEM study samples it became clear that urine is a challenging sample matrix for HILIC-MS analyses. Two method optimizations (i.e. i) urine pre-acquisition concentration normalization and ii) ESI-dopant 2-(2-methoxyethoxy)ethanol) aiming towards an improved robustness and sensitivity of the HILIC-MS analysis were established, tested and performance improvement was demonstrated.

Chapter 4: Since no detectable impact on the metabolism of the individuals subjected to the lifestyle intervention could be observed. Samples from the NUPRDEM study sample were

investigated with a different approach. Insulin sensitivity of the individual subjects was successfully correlated to their corresponding urinary metabolom and affected metabolic pathways were identified.

Chapter 5: Concluding remarks

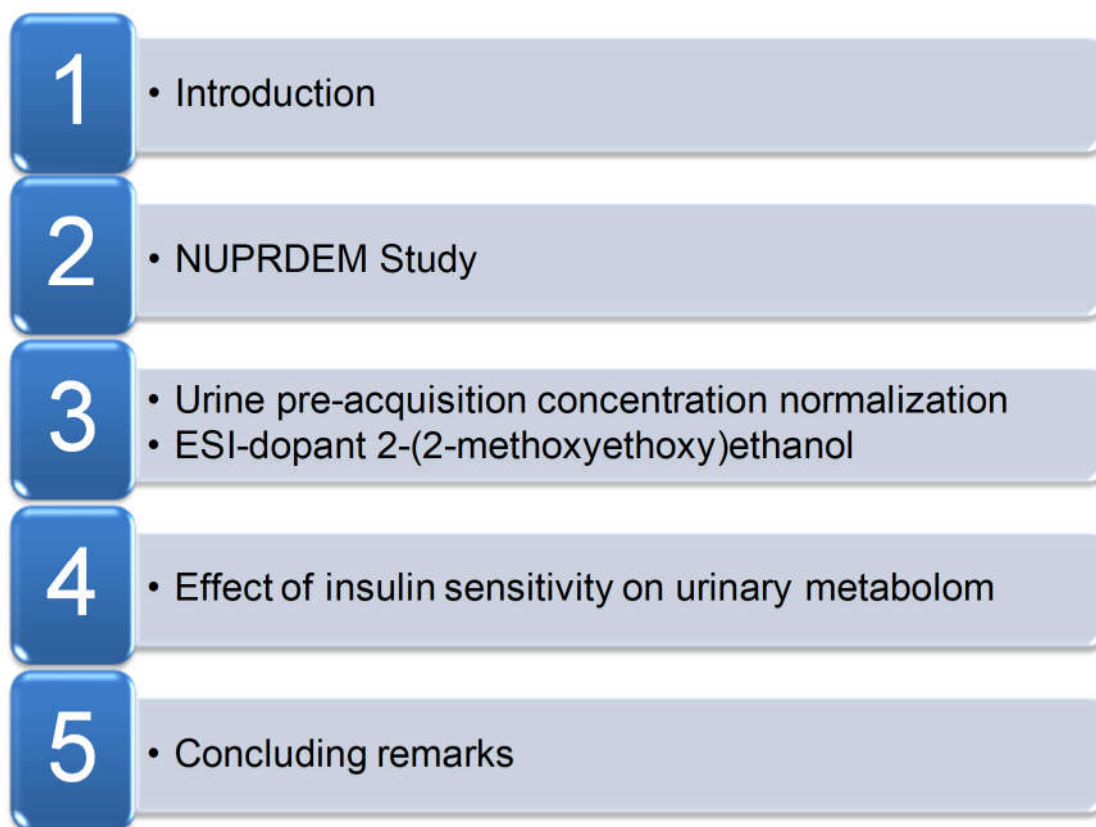


Figure 10: Outline of this thesis

2 Nutritional Prevention of Type 2 Diabetes Mellitus (NUPREDM) study

2.1 Objectives of the study and study design

Objective of the study was to elucidate the prospects of type II diabetes prevention by means of a lifestyle intervention and the effects on the metabolism of the body. For this reason, a lifestyle intervention program called nutritional prevention of type 2 diabetes mellitus (NUPREDM) was initiated in the framework of the Tuebingen University life style intervention program (TULIP). Scope of the nutritional lifestyle intervention was to setup an intervention scenario that realistically can be maintained by a majority of the target patient group. Therefore the participants received nutrition counseling and had to maintain a relatively mild calorie deficit (1800 Kcal/day) for 6 month. Inclusion criteria for patients were:

- Relatives with type 2 diabetes
- Body mass index (BMI) > 25 kg/m²
- Diagnosed impaired glucose tolerance
- Gestational diabetes in medical history.

In addition to that, the patient group was randomly separated into three different groups

- no red meat (fish, poultry allowed)
- increased fibre consumption (i.e. achieved via extra slice of whole-grain bread)
- control (no constrains but 1800 Kcal/day limit)

Several clinical parameters (BMI, oral glucose tolerance, waist and hip circumference, HBA1C, plasma and urine creatinine, fasting blood sugar and insulin) were determined and plasma and urine was sampled at baseline and at 6 month follow up time point. Samples of the NUPREDM study were analyzed using a comprehensive non-targeted metabolomics approach.

2.1.1 Samples

Spot urine samples from 118 volunteers were used. The cohort includes 79 female and 39 male participants with median age of 43 years (ranging from 22 to 71). Informed written consent was obtained from all subjects and the local medical ethics committee approved the protocol.

2.2 Method development for ICR-FT-MS and HILIC-MS analysis

2.2.1 Materials

All standards and LC-MS grade acetonitrile and methanol were obtained from Sigma Aldrich. LC-MS grade water was obtained from a MilliQ water purification system (Millipore, Molsheim, France).

2.2.2 ICR-FT-MS

Acquisition of ultrahigh resolution mass spectra was performed using a solarix™ ion cyclotron resonance Fourier transform mass spectrometer (Bruker Daltonix GmbH, Bremen Germany) in combination with an Apollo II ESI source (Bruker Daltonix GmbH, Bremen Germany) and a 12 Tesla super conducting magnet (Magnex Scientific Inc., Yarnton, GB).

2.2.2.1 *Automatization of sample work-up*

For the analysis of urine samples a method, previously established by K. Wörmann, was adapted to allow an automatized and thus highly reproducible sample workup. The former method protocol utilized 1 mL Varian C18 solid phase extraction (SPE) cartridges. These cartridges lose chromatographic retentivity if the solid phase “runs dry” during the work-up process. In an automatized approach drying of the column bed is inevitable due to variations in sample composition and SPE-cartridge column bed packing. Excessive air-purging of the SPE-cartridges after each elution step is actually necessary otherwise incomplete removal of mobile phase after each SPE step (i.e. conditioning, washing, elution) could result in a

compromised sample preparation. Because of the previously mentioned limitations the 1 mL Varian C18 cartridge needed to be exchanged to a polymer based sorbents which is resistant to column drying. K. Wörmann performed benchmarks on basis of amount of detected masses with decimals in the range of x.0 to x.5 (protonized singly charged ions) and x.5 to x.99 (multiply charged or salt adduct ions). The benchmarks demonstrated only minor drawbacks of Waters HLB cartridges performance if compared to Varian C18 cartridges (see Figure 11), thus the general suitability of polymer based SPE material was already approved.

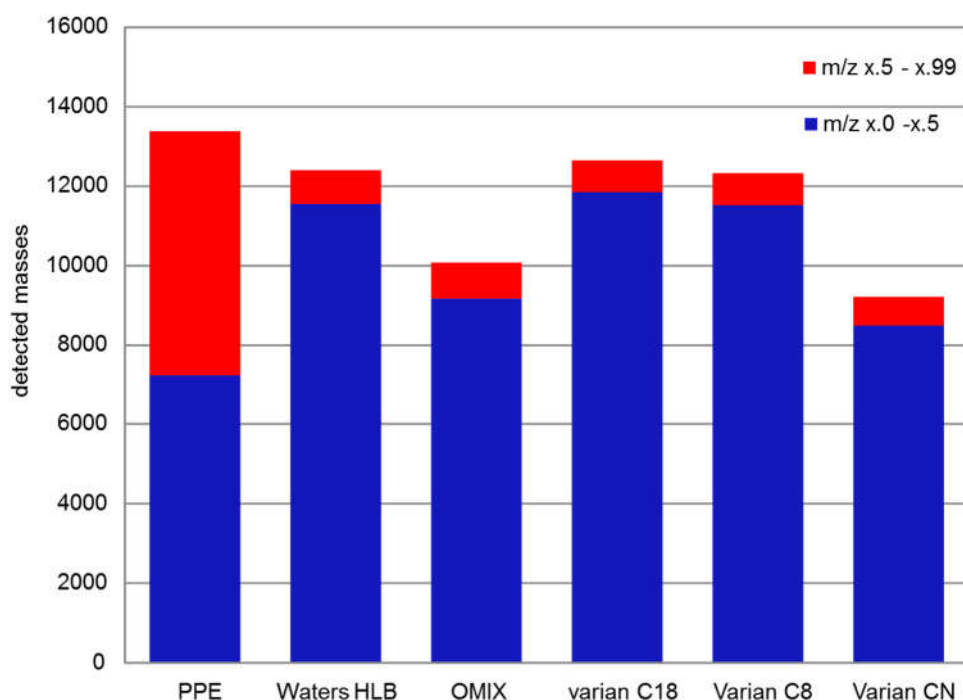


Figure 11 Comparison between different SPEs and protein precipitation (PPE) in FT-ICR-MS sample preparation. The total (ESI+ and ESI-) detected m/z less the ones in the blanks of each method is shown. The amount of m/z's with decimals in the range of x.50 to x.99 which are especially in the lower mass range multicharged, or cluster ions are marked in red.

In a further brief survey, a pooled QC-Urine sample was used to test polymer based 1 mL SPE cartridges of three vendors: Waters HLB, Agilent PPL and Thermo Sola (Figure 12).

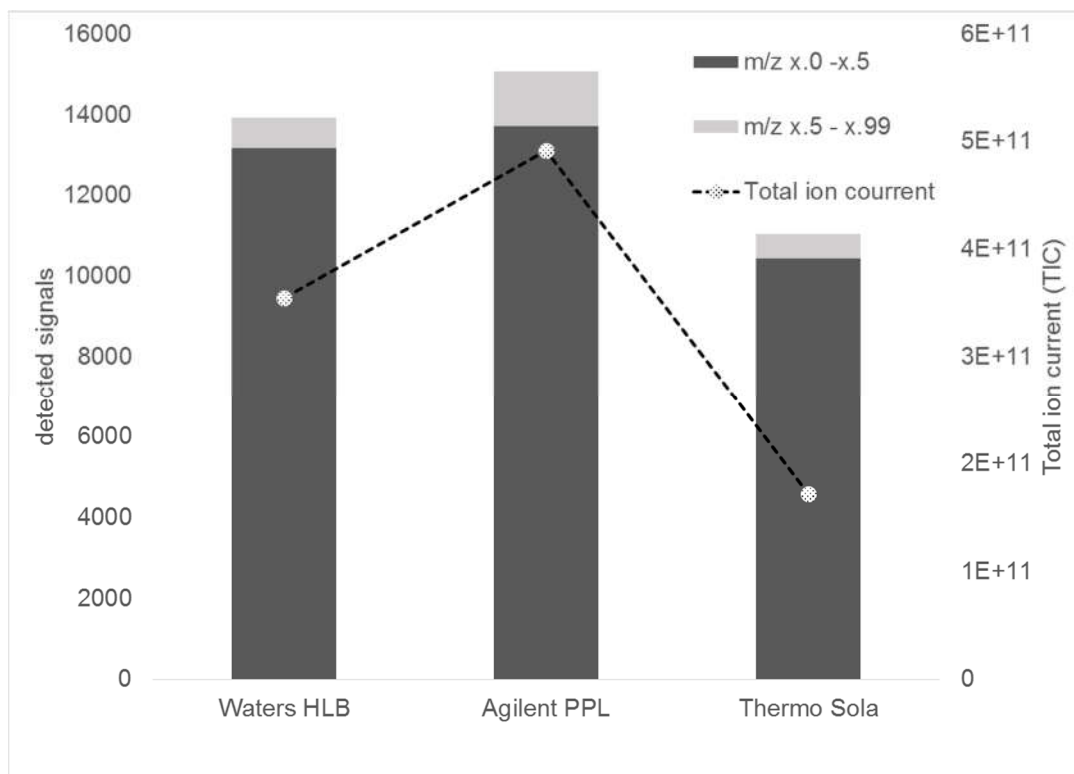


Figure 12: SPE cartridge comparison of Waters HLB, Agilent PPL and Thermo Sola in FT-ICR-MS sample preparation. The total (ESI+ and ESI-) detected m/z less the ones in the blanks of each method is shown. The amount of m/z 's with decimals in the range of $x.50$ to $x.99$ which are especially in the lower mass range multicharged, or cluster ions are marked in light grey.

The Agilent PPL cartridge provided best performance with regard to useful detected signals with mass defect between 0 and 0.5 and total ion current. For the automatized sample workup the Varian C18 1 mL was exchanged to the Agilent PPL 1 mL cartridge.

Before application to SPE cartridges, urine samples were centrifuged for 5 min at 10.000 rpm at 4°C and diluted. For dilution an aliquot of 60 μL urine sample was mixed with 540 μL of H_2O + 0.1%FA. The SPE protocol was adopted from K. Wörmann and slightly modified, the adapted SPE protocol was established as shown in Table 1. The corresponding flowchart of the liquid handling robot is shown in Figure 13.

Table 1: Urine SPE elution profile

Steps	Solvent	Volume
Wash	95% methanol, 5% water, acidified by 0.1% FA	2 x 1 mL
Equilibration	95% water, 5% methanol, acidified by 0.1% FA	2 x 1 mL
Load	centrifuged and diluted urine	2 x 250 µL
Wash	95% water, 5% methanol, acidified by 0.1% FA	2 x 500 µL
Elution	95% methanol, 5% water, acidified by 0.1% FA	2 x 250 µL
Final dilution (3x)	80% methanol, 20% water, acidified by 0.1% FA	150 µL eluate + 600 µL dilution solvent

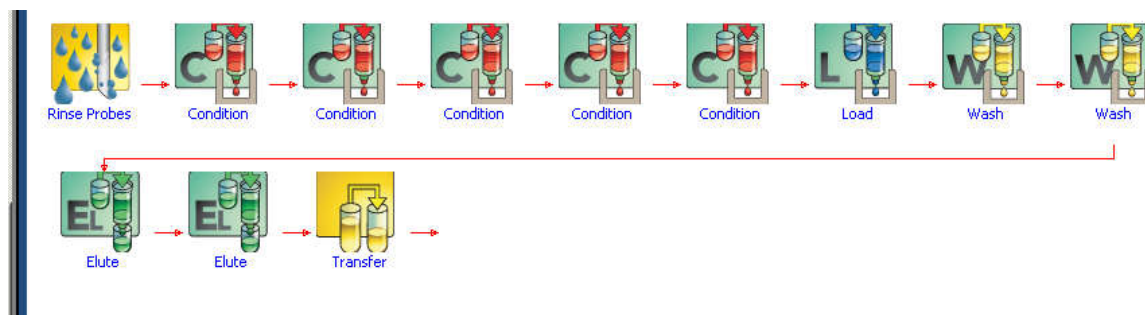


Figure 13: Flow-chart of liquid handling robot for SPE sample cleanup

The final dilution step, factor 5, was performed three times in order to obtain three equal aliquots for analysis in electrospray POS- and NEG-mode and one backup. With help of the adapted method and the transfer to the liquid handling robot, the sample preparation was conducted completely automated.

2.2.2.2 ICR-FT-MS analysis

Sample measurement was performed with help of a Gilson autosampler (sample changer 223, Gilson Inc., Middleton, USA). During analysis, samples were kept at 5 °C. Samples were injected at a flow rate of 2.0 $\mu\text{L}/\text{min}$. The analysis sequence was split to 6 batches and has been randomized, analog to sample work-up randomization (Figure 25), to avoid batch effects. Electrospray and method parameters were optimized for low masses ranging from 120 to 600 Da in broad band detection mode and with a time domain transient of 4 MW (megaword). For each spectrum 300 scans have been acquired. Detailed method parameters can be retrieved in Supplementary Table 1. The obtained mass resolution power of the obtained spectra was >400.000 at m/z . Prior to each batch, a 1 ppm arginine solution was used to calibrate the instrument to a mass error below 100 ppm.

2.2.2.3 ICR-FT-MS data handling

Post-acquisition calibration and peak picking were performed with the aid of DataAnalysis 4.1 SR1 (Bruker Daltonix GmbH, Bremen Germany). All obtained spectra were internally calibrated. The calibration list was generated via download of the Human Metabolom Data Base (HMDB) 3.5. All sum formulas were extracted. After removal of duplicates protonated respectively deprotonated mass to charge ratios were calculated and used for spectral calibration. Spectra were iteratively calibrated until mass error falls below 400 ppb and error distribution appears balanced after visual control (Figure 14).

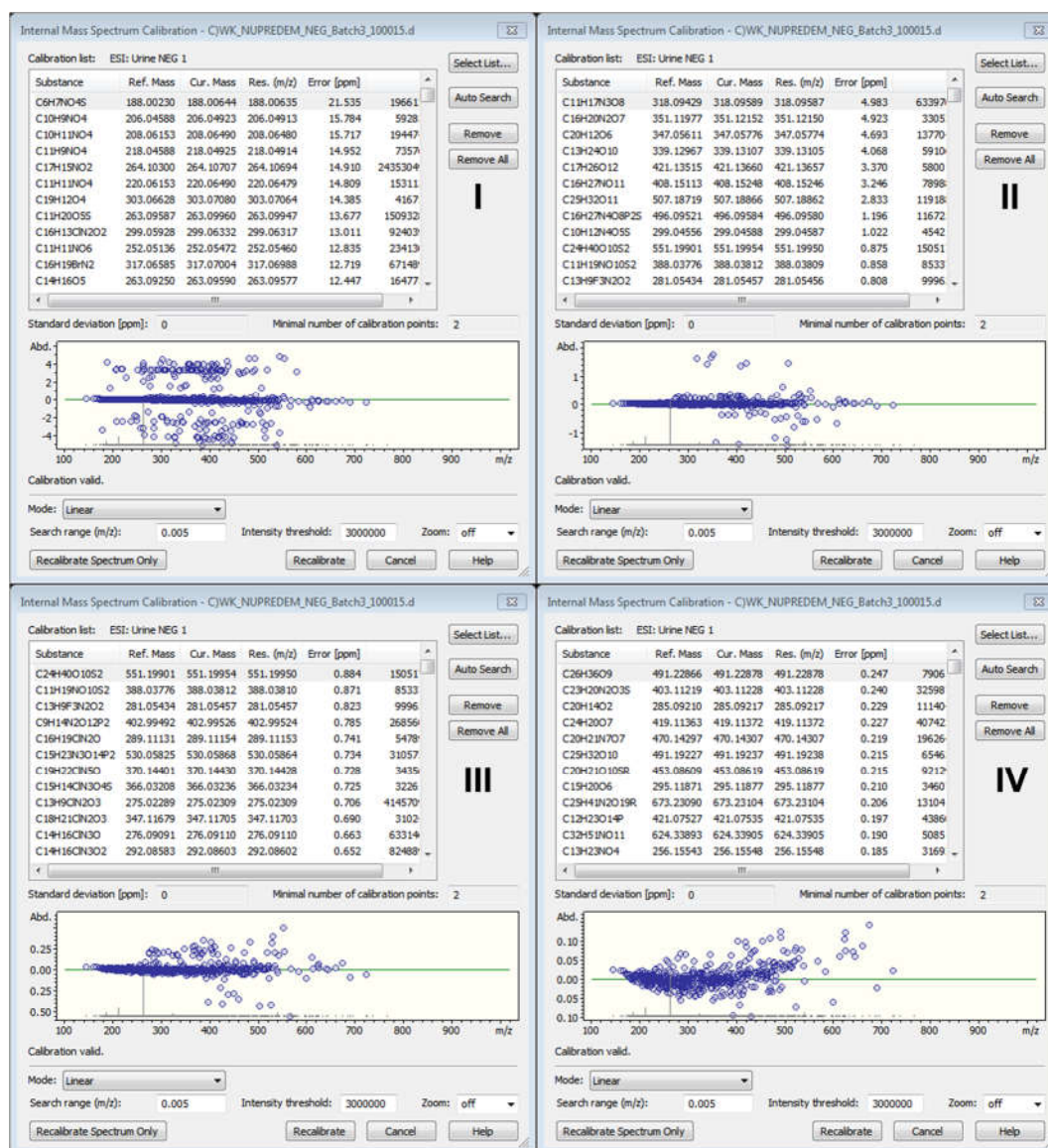


Figure 14: Successive calibration of FT-MS spectra with different mass error cutoffs: I: ~20 ppm cutoff, II: 5 ppm cutoff, III: 1 ppm cutoff and IV: 0.4 ppm cutoff

Peaks with a signal to noise ratio (S/N) less than four or an intensity of less than 2.000.000 counts were excluded and the resulting peak lists were exported to ASCII file format using the AutomationEngine 4.1 (Bruker Daltonix GmbH, Bremen Germany). Peak lists were aligned, applying a 1 ppm mass window, and data matrix was generated using Matrix generator 0.4 (developed in house, M. Lucio). Sum formulas of detected peaks have been calculated using Netcalc (74), an in house written Matlab program. Peaks remaining without Netcalc annotation have been excluded. Finally, the Netcalc filtered and annotated Matrix was imported to excel 2010. Peaks appearing in less than 10% of the samples were removed. This way, the number

of peaks could be reduced from 68886 to 16909 peaks, which is close to the limit of technical feasibility regarding multivariate statistical data evaluation limited by computational power. In addition, to Netcalc sum formula annotation, MassTrix 3 annotation was performed (<http://masstrix3.helmholtz-muenchen.de/masstrix3/>) using a 1 ppm mass window.

2.2.2.4 ICR-FT-MS data analysis

The following inter- and in-group comparisons were investigated (see Table 2). Several data pretreatment methods like scaling, normalization and transformation were evaluated but univariate, unsupervised multivariate or supervised multivariate approaches, using only a single step of “orthogonalization”, did not yield in sound results. For multivariate statistics SIMCA-P 13.0 (Umetrics, Umea, Sweden) was used by Dr. Marianna Lucio. Data evaluation was performed in two steps. In the first step the signal intensities were orthogonal signal corrected (OSC). In the second step the OSC data were analyzed using orthogonal partial least squares – discriminant analysis (OPLS-DA). The resulting models were evaluated via seven-fold cross-validation. Mass signals assigned with a variable importance on projection (VIP)-value > 1 were considered relevant.

Table 2: Tabulated overview of conducted group comparisons

In-group comparisons			
Intervention	Time	Vs.	Time
Diet 1	Baseline	Vs.	6 months
Diet 2	Baseline	Vs.	6 months
Diet 3	Baseline	Vs.	6 months
Inter-group comparisons			
Time	Intervention	Vs.	Intervention
6 months	Diet 1	Vs.	Diet 2
6 months	Diet 1	Vs.	Diet 3
6 months	Diet 2	Vs.	Diet 3

2.2.3 HILIC-MS

2.2.3.1 Materials

LC-MS grade acetonitrile and methanol were obtained from Sigma-Aldrich, (Taufkirchen, Germany). LC-MS grade water was obtained from a MilliQ water purification system (Millipore, Molsheim, France). LC-MS grade ammonium formate, ammonium acetate and formic acid were obtained from Sigma Aldrich (Taufkirchen, Germany), LC-MS grade acetic acid was obtained from Biosolve (Valkenswaard, Netherland). The standard compounds used for method development are listed in Supplementary Table 2 and Supplementary Table 3. Evaluated columns are listed in Supplementary Table 4.

2.2.3.2 Electrospray ionization

LC-MS experiments were performed using an Acquity-UPLC system (Waters Milford, MA, USA) coupled to a SYNAPT G1 qTOF HD mass spectrometer (Waters Micromass, Manchester, UK). Lock-mass infusion was performed via series 200 lc pump (perkin elmer, Shelton, USA). The following ESI parameters were applied:

- POS-mode
 - capillary voltage 3.3 kV, sample cone 24 V, extraction cone 4.0 V, source temperature 120 °C, desolvation temperature 350 °C, cone gas Flow 0 L/h, desolvation gas flow 800L/h. Scan time was set to 300 ms and m/z range from 50-1000 in V-mode (average 9000 resolution).
- NEG-mode
 - capillary voltage 2.3 kV, sample cone 24 V, extraction cone 4.0 V, source temperature 120 °C, desolvation temperature 350 °C, cone gas Flow 20.0 L/h, desolvation gas flow 800 L/h. Scan time was set to 300 ms in and m/z range from 50-1000 in V-mode (average 9000 resolution).

2.2.3.3 Column testing

The columns which have been evaluated are listed in Table 3. Primary reason for the selection of these specific columns was, the availability of HILIC UPLC particle size (> 2 µm) which was not given for other columns.

Table 3: HILIC column tested for method optimization

Name	Particle size	Ligand type
Waters BEH HILIC 150x2.1 mm	1.7 µm	Bare silica
Waters BEH Amide 150.x2.1 mm	1.7 µm	Amide group
Grace Vision HT HILIC 150x2.0 mm	1.5 µm	Bare silica
Grace Vision HT amino 150x2.0 mm	1.5 µm	Amino group

Columns were tested by analysis of a mixture of standards (Supplementary Table 3) using the following parameters.

Inject Vol	5 ul		
Column temp	30°C		
Eluent A	5 mM NH ₄ AcO + 0.1%(v/v) AcOH		
Eluent B	95% Acetonitrile 5 mM NH ₄ AcO + 0.1%(v/v) AcOH		
Time (min)	%B	Flow (mL/min)	
0.0	99.0	0.25	
2.0	99.0	0.25	
15.0	50.0	0.25	
19.5	50.0	0.25	
20.5	99.0	0.25	
25.5	99.0	0.25	

In ESI negative mode a brief pre-test gave no indication that flow-splitting will improve or worsen MS response (see Supplementary Figure 1). To prevent the ESI orifice from premature clogging a 50% split (source – waste) was applied in ESI negative mode in all following experiments. The standard mix was analyzed in positive and negative ESI mode. Resulting chromatograms were integrated using Genedata software version 7.5 (Basel, Switzerland). Details for peak detection parameters can be retrieved from Supplementary Table 5 and Supplementary Table 6. Maximum peak intensities for the standard compounds were extracted and visualized in Figure 15. The acidic amino acid aspartic acid and the dicarboxylic acid adipic acid did not elute as a peak (Figure 15 B) or even retained on the Grace Vision HT amino column. Presumably, this phenomenon is caused by attractive electrostatic forces between the cationic ligand and the anionic analytes. For this reason the Grace Vision HT amino column was excluded. Also, the Waters BEH amide column was excluded, since it showed unfavorable performance in ESI-POS (Figure 15 A) mode and fructose did not elute as a peak, for unknown reasons. In an overall comparison the Grace Vision HT HILIC column proved best performance followed by the Waters BEH HILIC column but differences were marginal.

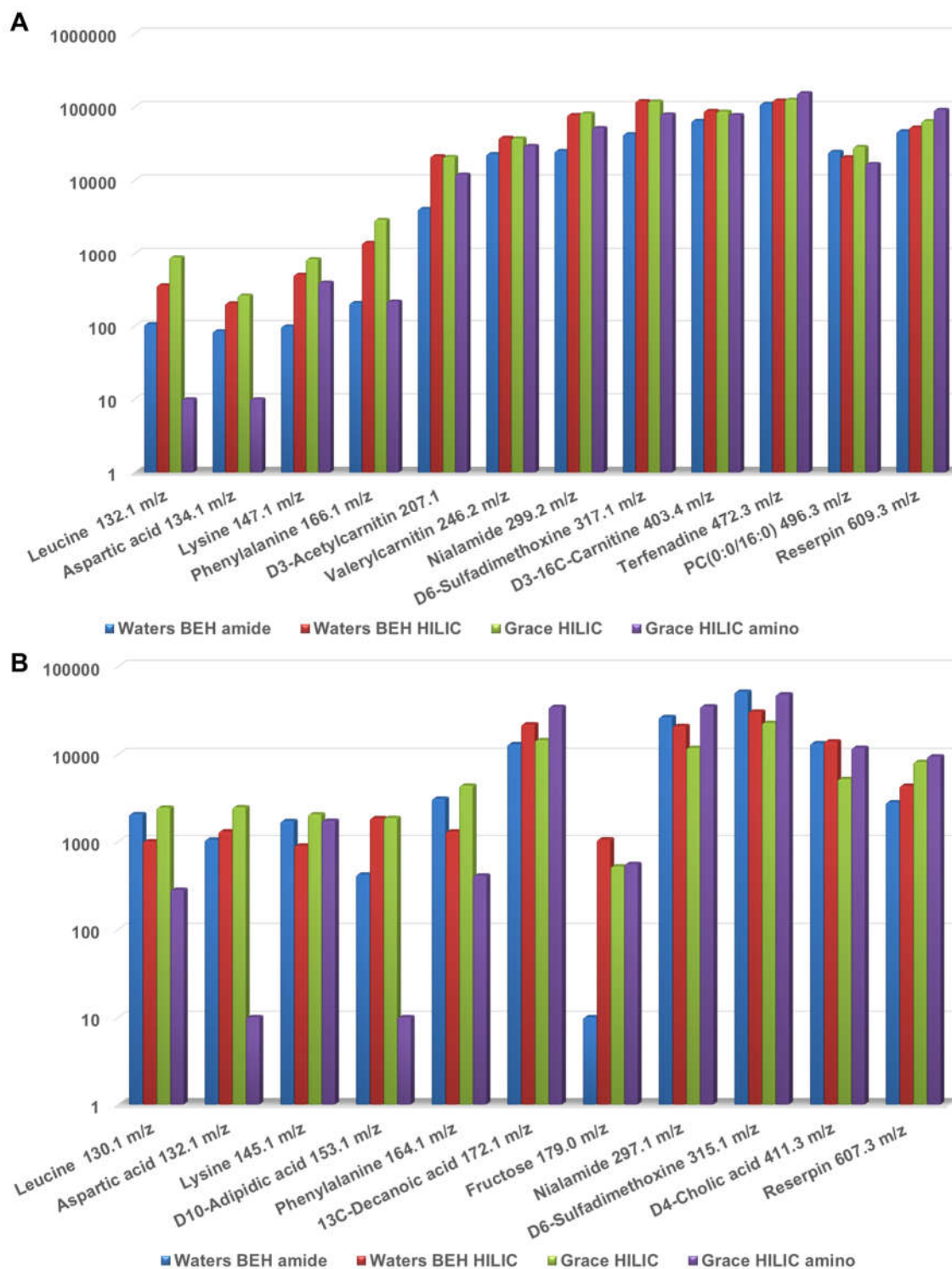


Figure 15: Maximum peak intensities of reference standards obtained during HILIC column comparison tests performed in A) ESI-POS mode and B) ESI-NEG mode

In further optimization tests the Grace HILIC column exhibited poor durability. After approximately 400 injections retention capacity decreased and column back pressure

increased dramatically, both signs of severe column deterioration. Additionally, the sensitivity of the MS instrument decreased significantly during an analysis batch of several consecutive runs (data not shown). Examination of the ESI chamber revealed, that the ESI orifice and spray shield was clogged with a white crystalline residue (Figure 16).



Figure 16: ESI spray shield after 40 analytical runs of blank (water) injections on Grace Vision HT HILIC column

After thorough cleaning of the ESI spray shield and orifice, MS sensitivity could be restored. It was hypothesized that the Grace Vision HT HILIC column, made from bare silica, suffers from strong column bleeding. Examination of mass signals, resulting from a blank analysis run (water injection) using a Grace Vision HT HILIC column corroborated this hypothesis. Clusters, that suggest presence of silicic acid in the column eluate, were observed in the blank analysis run see Figure 17. Similar results were observed for other silica based HILIC columns featuring sulfobetaine- or diol-ligands (Supplementary Figure 2 and Supplementary Figure 3). All observed phenomena can be interpreted as indications of severe hydrolysis of column material also known as column bleeding.

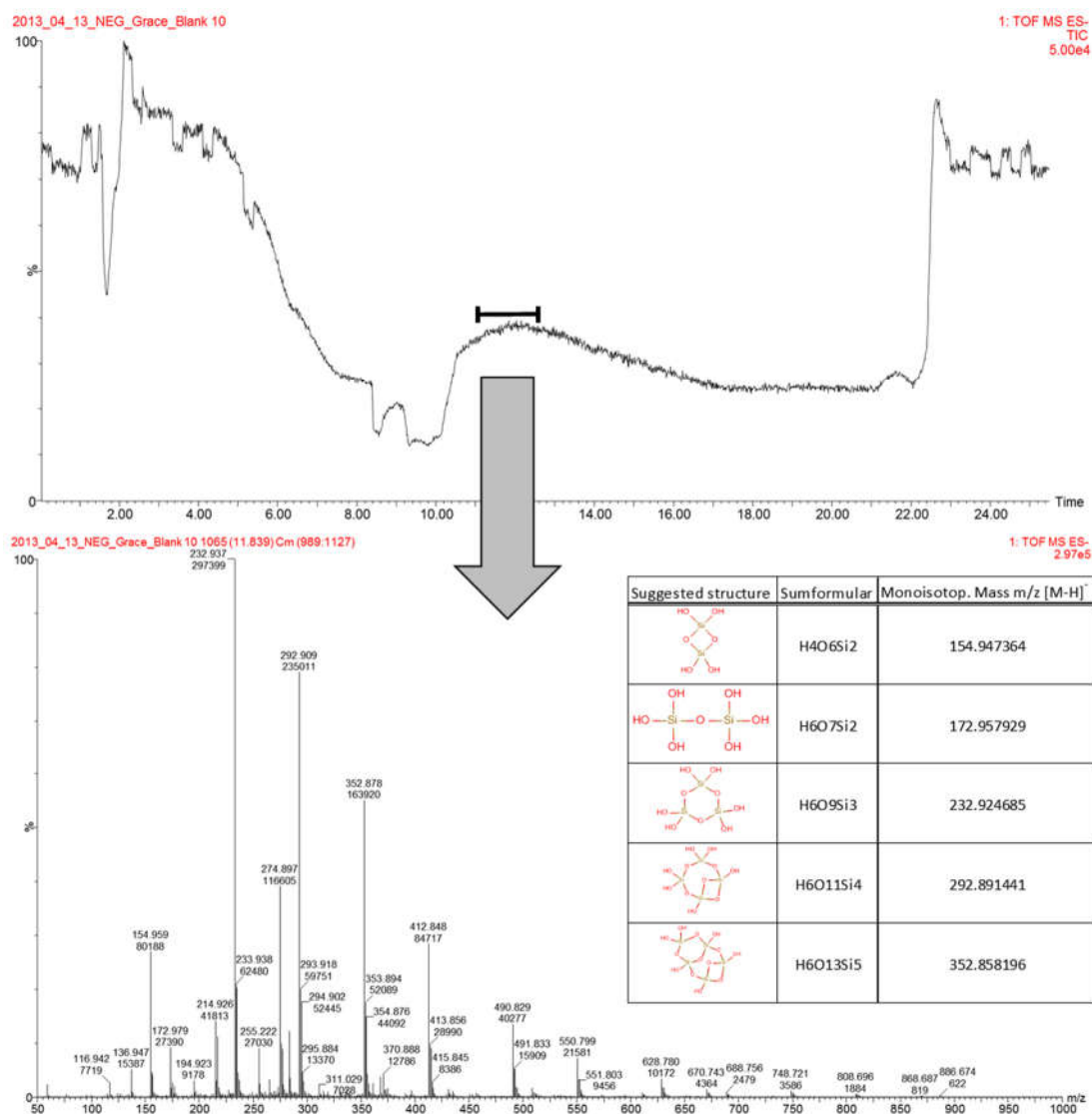


Figure 17: Mass signals caused by column bleeding, detected in blank of Grace Vision HT blank analysis

Only the HILIC column based on a silica-polymer-hybrid material show sustained durability. In an overall comparison, the Waters BEH HILIC column proofed best overall performance providing long-term stability (see Figure 18) and chromatographic efficiency. Hence the Waters BEH HILIC was chosen for all further analyses.



Figure 18: ESI spray shield after 40 analytical runs, including 10 waters blanks and 30 blood plasma samples, injections on Waters BEH HILIC column

2.2.3.4 Sample work-up

The sample work-up procedure has to meet several requirements such as: compatibility to a high throughput scenario, broad and quantitative coverage of metabolites, removal of undesired matrix components, minimum chemical alteration of the sample and facilitate the compatibility of the sample to the analytical system.

In the case of urine analysis via HILIC the following challenges have been identified. HILIC is susceptible to high water concentration in the sample, on that account a major hurdle lied in the aqueous nature of the urine sample. Protein concentration in urine samples is usually considered negligible if samples originate from healthy individuals. But, it cannot be excluded that in the course of the NUPREDM study pre-diabetic subject which already suffer from kidney damage have been sampled.

A simple and easy to perform protein precipitation procedure, using ice cooled acetonitrile and a drying step employing a vacuum centrifugation, was considered the most straight forward approach to tackle all previously mentioned issues.

Ice cooled acetonitrile (400 μL) was added to an aliquot of sample (100 μL). The resulting mixture was vortex mixed for 10 seconds and centrifuged at 4 °C 20817 g for 5 minutes. The obtained supernatants were dried in a vacuum centrifuge and stored at -80 °C until analysis.

Prior to analysis, the frozen evaporated residues were thawed and extracted with 100 μL extraction solvent which contained internal standards (D10-adipic acid 5 $\mu\text{g}/\text{mL}$, D4-cholic acid 3 $\mu\text{g}/\text{mL}$, Reserpin 5 $\mu\text{g}/\text{mL}$, D3-palmitoyl-L-carnitine 1 $\mu\text{g}/\text{mL}$, D3-acetyl-L-carnitine 1 $\mu\text{g}/\text{mL}$). The samples were subjected to supersonic bath (5 seconds) vortex mixed (10 seconds) and centrifuged at 4 °C 20817 g for 5 minutes. Supernatants were transferred to total recovery UPLC vials and kept at 4 °C in the UPLC autosampler for analysis.

To evaluate the effect of the extraction solvent on HILIC separation two different mixtures: urine buffer 1 (acetonitrile/water=2/1(v/v) buffered with 10 mM NH_4AcO + 0.1% AcOH) and urine buffer 2 suggested by Wang et. Al. 2008 (75) (acetonitrile/methanol=3/1(v/v) and 0.2% formic acid) were tested. A pooled QC-Urine sample from healthy volunteers was analyzed in ESI-POS and -NEG mode used for testing. The resulting chromatograms were visually evaluated (see Figure 19). In negative mode, the differences between urine buffer 1 and 2 are not very pronounced. When urine buffer 1 was used peak fronting was observed (Figure 19 IV) at 5.5 minutes retention time, which indicates a perturbed chromatography, probably due to high water content in the sample. In contrast, when urine buffer 2 was used an additional peak at 2.4 minutes was observable (Figure 19 III). In positive mode the differences, especially in the early eluting region of the chromatogram up to 5 minutes retention time, become more noticeable. Distinct peaks can be observed if the anhydrous urine buffer 2 is used (Figure 19-I). Urine buffer 2 was chosen for further experiments since it provided clearly better

performance in terms of less interference with HILIC methodology in the early eluting region of the chromatogram.

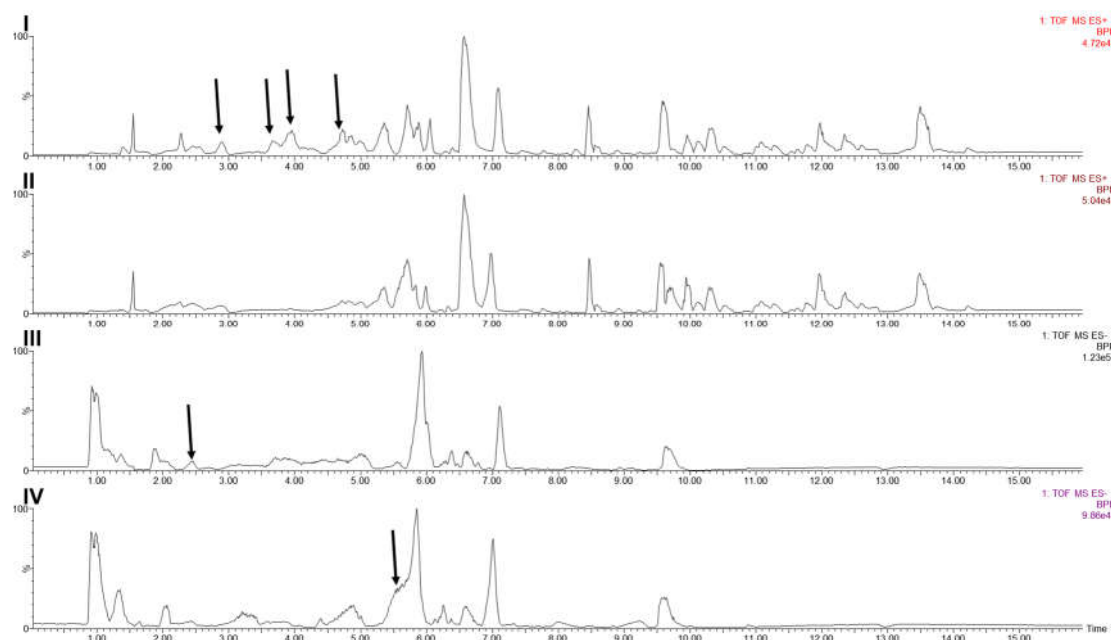


Figure 19: QC-urine samples analyzed in: POS-mode using urine buffer 2 (I), POS-mode using urine buffer 1 (II), NEG-mode using urine buffer 2 (III), NEG-mode using urine buffer 1 (IV). Additional peaks or peak shape problems are indicated with arrows.

In further gradient development and long-term stability tests it was noticed that during measuring campaigns over several days of consecutive injections, using aliquots of the same sample, peak intensities were increasing with progression of analysis time. This observation was not expected, indeed exactly the opposite was anticipated. It was prognosticated that, clogging of the ESI-orifice and sample degradation in the auto sampler over time would lead to decreasing signals intensities.

To test the stability of the analytical system an analytical sequence over seven days was set up. Sample queue was arranged in sections, one sections was composed as follows: triplicate injections of standards (Supplementary Table 3) followed by injections of QC urine and one blank injection (see Figure 20). All injections of one iteration step were sampled from a single vial for standards and urine sample respectively.

Sample	Std	Std	Std	Ur	Ur	Ur	Ur	Ur	Ur	Ur	Blank	Ur	Ur	Ur	Ur	Ur	Ur	Ur	
Vial	n+1	n+1	n+1	m+1	m+1	m+1	m+1	m+1	m+1	m+1	48	m+1	m+1	m+1	m+1	m+1	m+1	m+1	m+1

Figure 20: Set-up of analytical sequence for stability study

Per run 5 µL were injected, to supply sufficient volume vials containing standards were filled with 50 µL, vials containing urine sample were filled with 100 µL respectively. Only the observed total ion current of the standards were evaluated (Figure 21) to simplify evaluation.

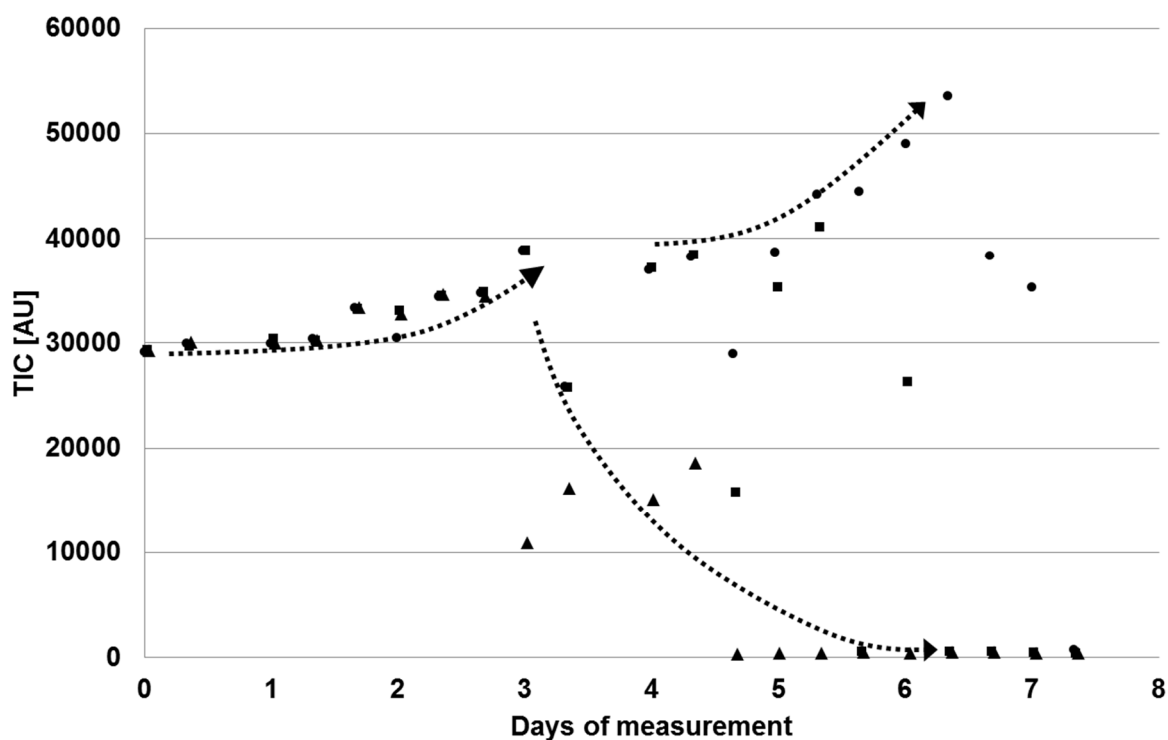


Figure 21: Total ion current of standards analyzed via HILIC over seven days of consecutive measurement. (•) first injection of triplicate, (■) second injection of triplicate, (▲) third injection of triplicate. Dotted lines indicate trends.

During the first three days a monotonic increase of TIC, encompassing all three injections of the triplicate, was observed. After three days the TIC of the third injection of the triplicates started decreasing until, after five days, no signal was detectable. At the same time the TIC of the first and second injection of the triplicates continued to increase. After six and seven days also second and first injection of the triplicates decreased. This phenomenon could be explained via successive solvent evaporation from the pre-slitted vial caps of the total recovery vials.

To validate this hypothesis a vial sealed with a pre-slit cap containing 50 μL the previously established urine buffer 2 (acetonitrile/methanol=3/1(v/v) and 0.2% formic acid) was placed in the sample manager at 4°C for several days without piercing the vial cap. And indeed, filling level of the vial successively decreased over time. Apparently, pre-slitted caps cannot prevent solvent evaporation from the vial. Unfortunately, pre-slitted vial caps are a prerequisite for the UPLC-sample manager and high organic percentage in the sample reconstitution solvent is necessary for HILIC. To prevent this phenomenon, several measures were taken: batches of maximum 3 days of consecutive analyses were allowed, instead of total recovery vials, 96-well plates sealed via peelable heat sealing (aluminum) foil were used and multiple sampling from the same sample vial/well will no longer be performed.

The new well plate set-up was tested against the total recovery vial set-up in an additional stability study, performed as explained previously (Figure 22).

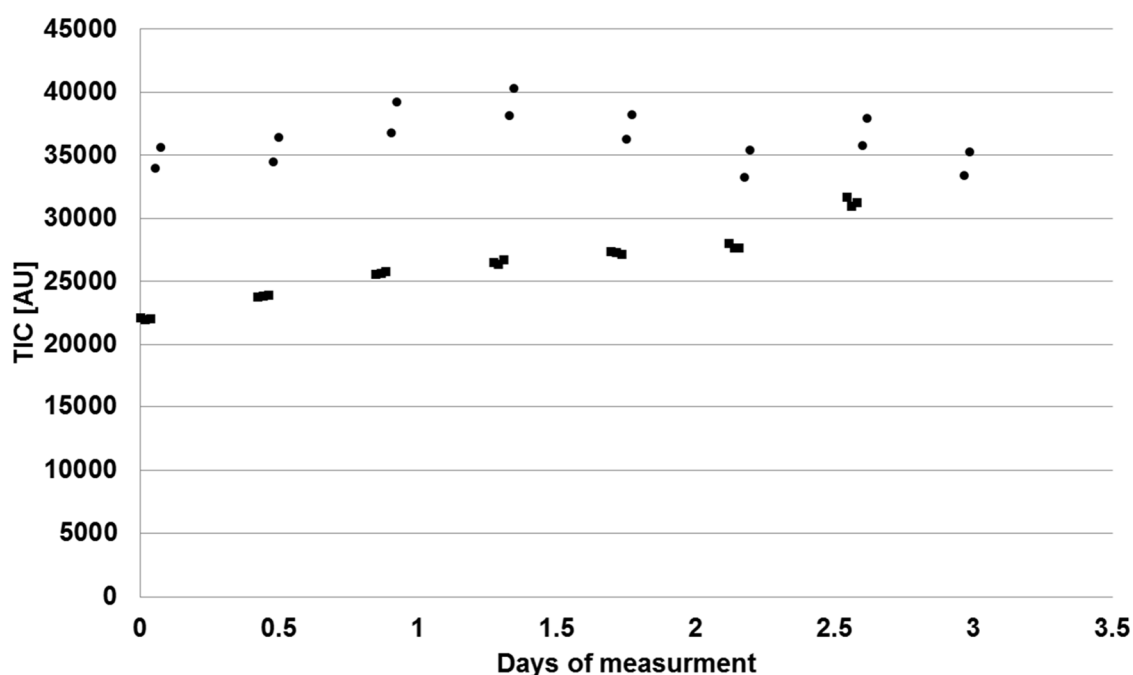


Figure 22: Total ion current of standards analyzed via HILIC-MS over 84 hours of consecutive measurement sample storage in well-plate (•) versus total recovery vial (■).

Well-plate injections did not show a trend of increasing TIC's over time in contrast to injections from pre-slit vials. Although multiple injections from the same well were precluded previously,

duplicate injections were performed. The discrepancies, which were visible in the well-plate duplicate injections (Figure 22) demonstrate the substantial impact of solvent evaporation on sample concentration if one and the same well was sampled multiple times.

2.2.3.5 System conditioning stability and accuracy

To evaluate the stability and accuracy of the system, retention time stability and mass accuracy during 3 days of consecutive measurements of urine samples has been evaluated. The chromatographic conditions were slightly adapted compared to the method described in chapter 2.2.3.3. To simplify the mobile phase preparation and to avoid retention time shifts caused by variations during mobile phase preparation mobile phase B is changed from 95% Acetonitrile 5 mM NH₄AcO + 0.1% (v/v) AcOH to pure acetonitrile. In return, buffer strength of mobile phase is A is increased from 5 mM NH₄AcO + 0.1% (v/v) AcOH to 10 mM NH₄AcO + 0.2% (v/v) AcOH. The gradient elution profile was adapted accordingly to keep the water: acetonitrile proportion similar.

Time (min)	%B	Flow (mL/min)	
0.0	95.0	0.3	
2.0	95.0	0.3	
15.0	50.0	0.3	
19.5	50.0	0.3	
20.5	95.0	0.3	
25.5	95.0	0.3	

In the course of the development of the comprehensive metabolomics platform a method harmonization was conducted. In this regard the flow rate was increased from 0.25 to 0.3 mL/m and column temperature was increased from 30°C to 40°C.

In reversed phase metabolomics approaches it is common to perform a column conditioning step in forehand of analysis. This is conducted via multiple injections (~10) of QC samples. In

the present HILIC method, the column is conditioned via three injections of acetonitrile blanks followed by three injections of HILIC method development standards. An overlay of the resulting base peak chromatograms (BPC) of the initial five QC urine injections is illustrated in Figure 23. Visual examination suggests a high degree of congruency in the quintuplicate injections and it is assumed that column conditioning is sufficient after five QC urine injections.

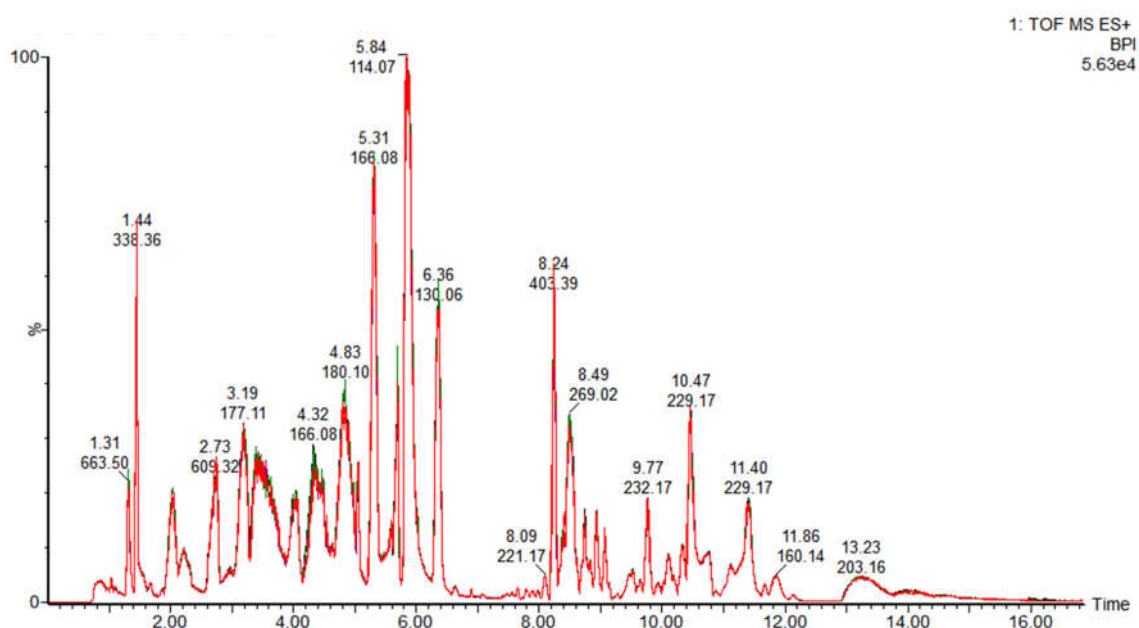


Figure 23: BPC overlay of five QC urine analyses with HILIC-MS after three blank and three standard injections

Evaluation of the retention times of carnitine, a regular urine component and the internal standards spiked to the QC urine proved a retention time stability of better than 2% relative standard deviation (see Table 4).

Table 4: Retention time stability during 3 days of consecutive measuring

Standard	Average ret. Time [min]	%SD
creatinine [M+H] ⁺ 114.066 m/z	5.9490	0.34

D3-acetylcarnitine [M+H] ⁺ 207.141 m/z	10.3481	0.04
D3-palmitoyl-carnitine [M+H] ⁺ 403.360 m/z	8.2394	0.05
D4-cholic acid [M+NH ₄] ⁺ 430.346 m/z	4.2347	1.72
reserpin [M+H] ⁺ 609.280 m/z	2.7308	1.94

Lock-mass correction was performed via Genedata with a 6 point correction approach and via Masslynx with a single point correction (leucine-enkephalin). For lock-mass standard composition refer to Supplementary Table 7.

Results of mass accuracy obtained via Genedata and Masslynx lock-mass correction are illustrated in Figure 24. Visual examination suggests that the error distribution resulting from Genedata lock-mass correction is more centered than Masslynx correction, especially in the region of $m/z > 600$. Nevertheless, in the region of 50 - 600 m/z , the region where most urinary metabolites are expected, Genedata and Masslynx lock-mass correction show equivalent performance. For the measurements of the samples the Genedata lock-mass correction was chosen, Genedata lock-mass corrections presupposes a deactivated Masslynx lock-mass correction.

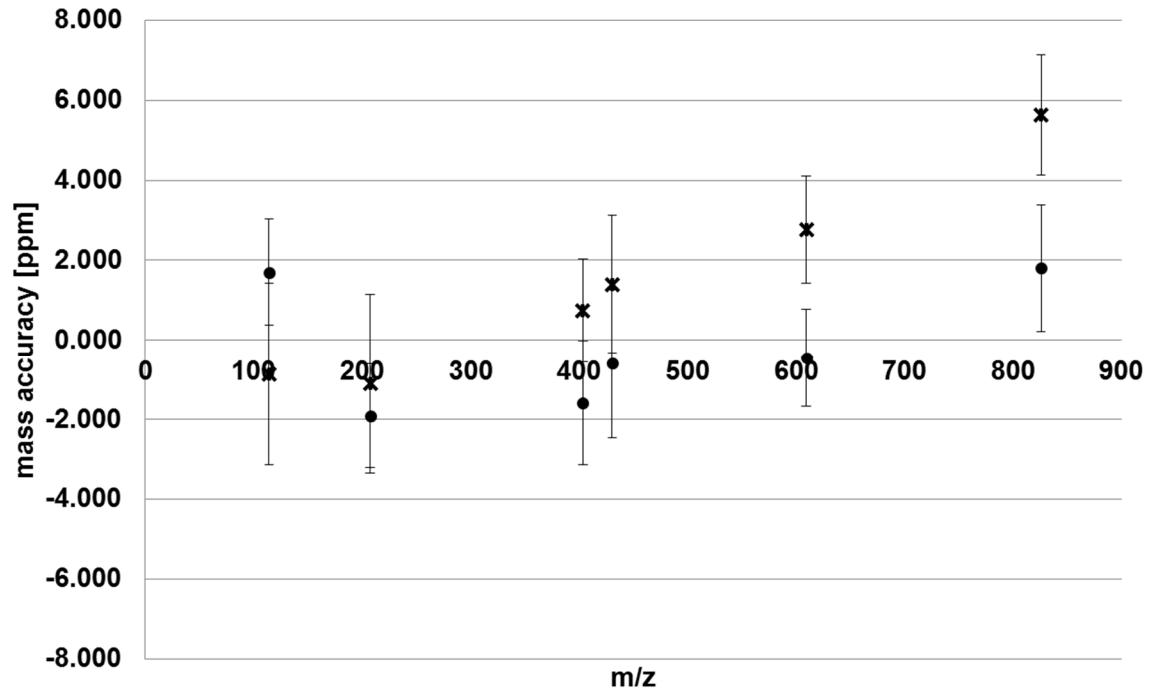


Figure 24: Average ppm errors determined with , (*) Genedata 6 point lock-mass correction and (x) Masslynx single point lock-mass correction for creatinine, D3-acetylcarnitine, D3-palmitoyl-carnitine, D4-cholic acid, reserpin and erythromycin.

2.2.3.6 Study analysis with HILIC-MS

LC-MS experiments were performed using an Acquity-UPLC system (Waters Milford, MA, USA) coupled to a SYNAPT G1 qTOF HD mass spectrometer (Waters Micromass, Manchester, UK). Lock-mass infusion was performed via series 200 Ic pump (perkin elmer, Shelton, USA). The following ESI parameters were applied: capillary voltage 2.3 kV, sample cone 24 V, extraction cone 4.0 V, source temperature 120°C, desolvation temperature, 350°C, cone gas Flow 20.0 L/h, desolvation gas flow 800L/h. Scan time was set to 300 ms and m/z range from 50-1000 in V-mode (average 9000 resolution). The mass spectrometer was calibrated with sodium formate clusters.

HILIC separations were performed on an Acquity UPLC BEH HILIC (150x2.1 mm 1.7µm, Waters) column in gradient mode. The column temperature was set to 40°C. Mobile phase A consisted of 10 mM ammonium acetate and 0.2 v/v% acetic acid in water, mobile phase B consisted of acetonitrile. The flow rate was set to 300 µL/min, the injection volume was set to 5.0 µL. Gradient elution was performed with a linear gradient: 0 - 2 min: 5 % A, 15 - 20 min 50 % A, 20.5 - 25.5 min 5 % A. A source: waste (1:2) split was applied. To ensure mass stability a lock-mass standard solution, water : methanol = 1 : 1 acidified by addition of 0.1% (v/v) formic acid for composition please refer to Supplementary Table 7, was infused via orthogonal sprayer at a flow rate of 50 µL/min. Lock-mass spectra were acquired in 90 seconds intervals. To avoid batch effects samples were randomized, flow-chart see Figure 25 and analyzed in technical duplicates.

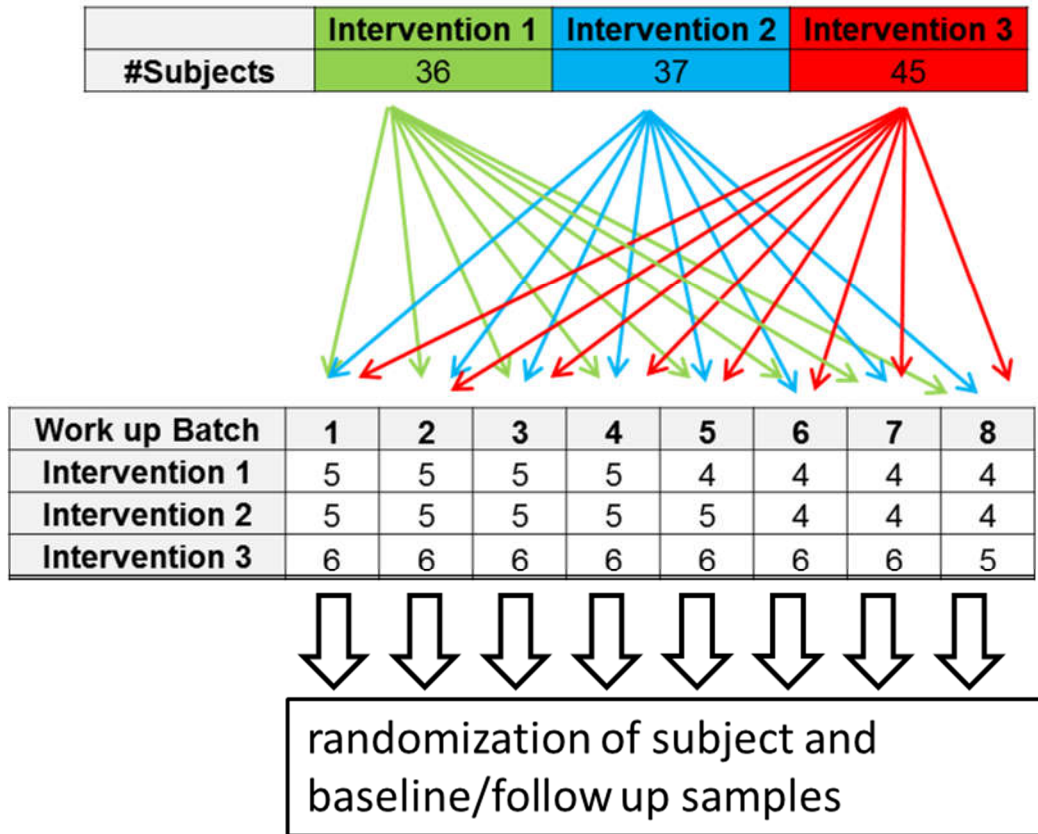


Figure 25: Flow-chart of sample work-up used for randomization of samples during the study

2.2.3.7 HILIC-MS data handling

Chromatogram alignment and peak picking for non-targeted HILIC-UPLC-QTOF analysis was performed with help of Genedata Expressionist MS Refiner 8.0 and Analyst 8.0 (Genedata, Munich, Germany). Mass annotation of detected peaks was performed using Genedata in combination with HMDB Human Metabolite Database (HMDB) version 3.5 (29-31) with a mass annotation error of ± 7 ppm.

The resulting matrix was filtered with help of Microsoft Office Excel 2010. Peaks were discarded if

- they were detected in less than 30% of the sample set,
- they were not detected in both of the respective technical duplicates or
- deviation between the signal intensity of a single peak and the arithmetic mean of the respective duplicate exceeded 30%.

Remaining peaks were averaged with using arithmetic mean.

2.2.3.8 HILIC-MS data analysis

The same inter- and in-group comparisons as described in 2.2.2.4 (see Table 2) were investigated. Several data pretreatment methods like scaling, normalization and transformation were evaluated but univariate, unsupervised multivariate or supervised multivariate approaches did not yield in sound results. For multivariate statistics SIMCA-P 13.0 (Umetrics, Umea, Sweden) was used by Dr. Marianna Lucio.

2.3 Results

2.3.1 Results ICR-FT-MS

In order to retrieve metabolites that are describing the differences between the groups, orthogonal signal corrected data were calculated and OPLS-DA was performed, with the orthogonal signal corrected data. This approach is known to be prone to overfitting the results and is not common practice, yet the only way that yields in discriminative multivariate models. An overview of successful inter- and in-group comparisons is given in Table 5.

Table 5: Results of multivariate statistic models

In-group comparisons				
Intervention	Time	Vs.	Time	OSC-OPLS-DA model
Diet 1	Baseline	Vs.	6 months	Valid
Diet 2	Baseline	Vs.	6 months	Valid
Diet 3	Baseline	Vs.	6 months	Not significant
Inter-group comparisons				
Time	Intervention	Vs.	Intervention	OSC-OPLS-DA model
6 months	Diet 1	Vs.	Diet 2	Valid
6 months	Diet 1	Vs.	Diet 3	Valid
6 months	Diet 2	Vs.	Diet 3	Valid

The results will be evaluated only exemplary on base of the inter-group comparison of the follow ups of diet 1 versus diet 3. This case was chosen since it showed best separation of the different groups in the analysis (see Figure 26) and the results vary only in minor details.

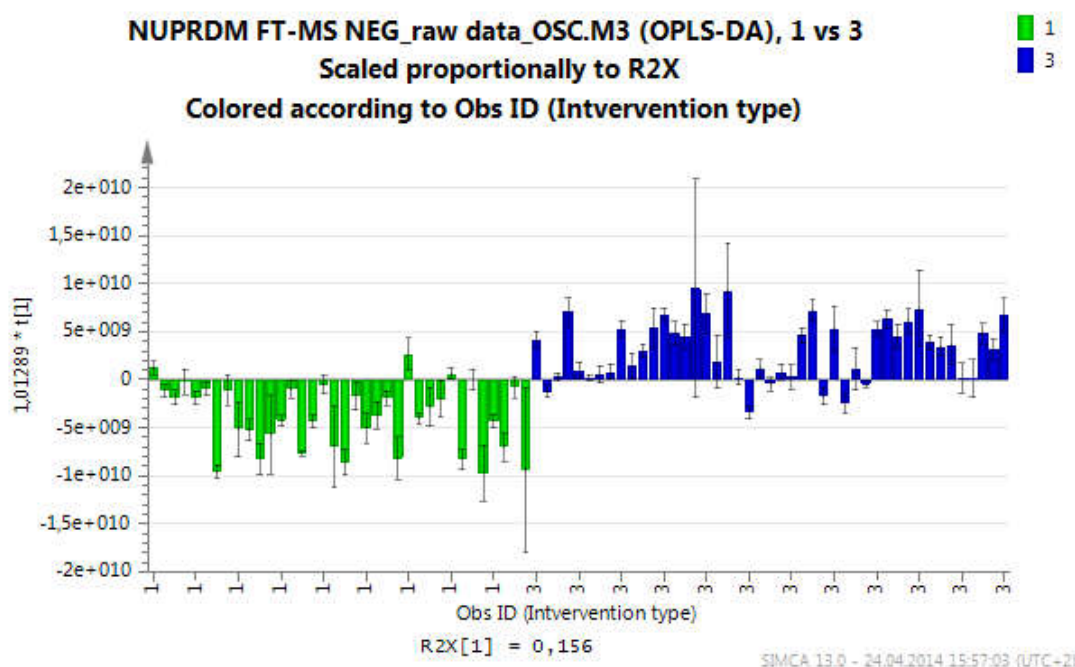


Figure 26: Separation of the NUPREDM follow up individuals in groups 1 (green) and 3 (blue) in the OPLS-DA analysis.

After OSC-OPLS-DA analysis, 119 mass signals were identified with a vip-factor >1 . The detected signals were assigned to molecular features via exact mass annotation, facilitated with MassTriX 3, and educated guess. In the next step the molecular features were classified to the chemical categories: unknown, nutritional exogenous, metabolite, drug related, FT-MS background, gut microbiome and hormones. The resulting distribution of molecular features to the chemical categories is illustrated in a pie chart see Figure 27. Only about 8% of the mass signal emerging from the analysis can be ascribed to endogenous metabolism. The vast majority of the annotated peaks are either unknowns, nutritional/exogenous, drug related or FT-MS background ions. The obtained distribution of chemical categories in the VIP's of the multivariate analysis as the absolute number of VIP's, namely 119, raise reasonable doubts regarding the informative value of the results. Also, the occurrence of well-known FT-MS

background ions as well as drug related metabolites in the VIP list of the multivariate analysis undermine the integrity of the results.

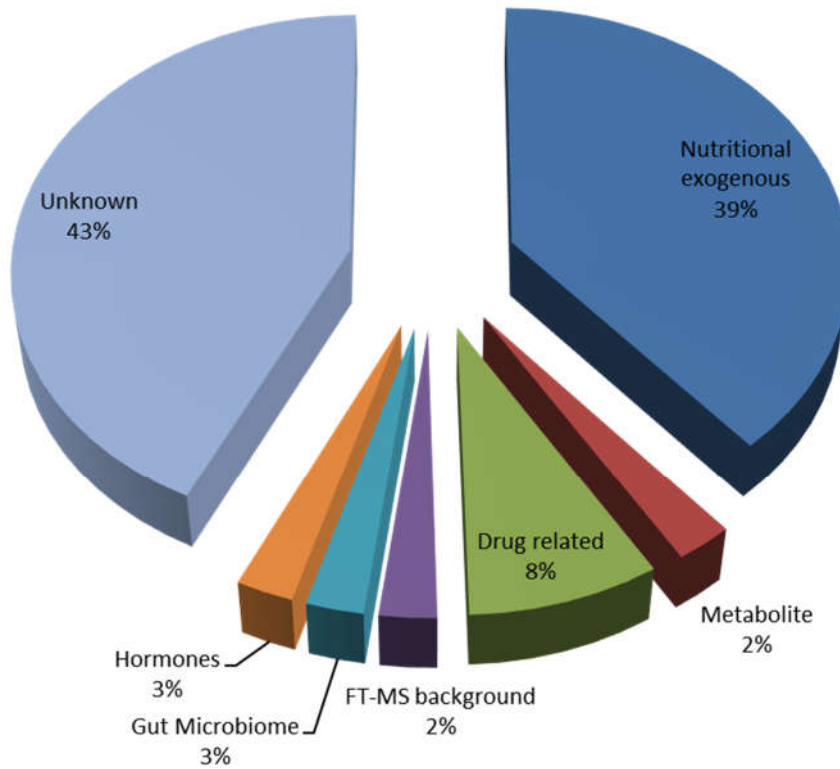


Figure 27: Distribution of important mass signal resulting from OSC-OPLS-DA analysis of NUPREDM follow up group 1 vs 3 analysis in chemical categories.

2.3.2 Results HILIC-MS

For the HILIC-MS approach no valid multivariate statistic model could be established, also univariate statistics could not reveal significant results.

2.4 Conclusion

The analysis of the samples that originated from the NUPREDM study did not show any statistical sound results. Also other attempts of metabolomics screening approaches which incorporated RP-LC-MS and FT-ICR-MS analyses of plasma sample from the same study participants and 800 MHz ¹H-NMR analysis of the same urine samples, data not shown, showed no changes in the metabolism of the study subjects. Therefore, more extensive investigation towards the initial nutritional intervention were protracted.

It appeared that the original idea, of testing a realistic – in terms of practicability – not too drastic nutritional lifestyle intervention has too little impact/effects on the metabolome to be detected with at present available techniques. Even though no changes in the metabolome of the participants could be observed, the study can be considered as success since a weight loss and a gain in insulin sensitivity in average was observed.

3 Optimization of HILIC-MS method

3.1 Normalization of urine concentration via dilution using a Poly Parametric Urine Normalization Factor (PPUNF)

Parts of this chapter were published in

Koch W, Forcisi S, Lehmann R, Schmitt-Kopplin P: Sensitivity improvement in hydrophilic interaction chromatography negative mode electrospray ionization mass spectrometry using 2-(2-methoxyethoxy)ethanol as a post-column modifier for non-targeted metabolomics
Journal/J Chromatogr A, 2014, 1361, 209-16

3.1.1 Introduction

In preliminary statistical screenings based on the NUPREDM study baseline data, discussed in chapter 2.2.2, the FT-MS data indicated that the approach could yield in promising results. In contrast, no valid multivariate models could be established with the HILIC-MS data. During investigation of the HILIC-MS data it appeared that retention time shifts impaired the peak alignment, although retention time stability was validated (see 2.2.3.5). The validation was performed with a QC sample. Of course, a QC sample does not vary in its chemical composition. But, real urine samples are not a product of homeostatic control like blood plasma. Thus, the chemical composition in particular the overall concentration varies up to 25-fold mainly depending on the hydration status of the body (76). A fact which was well known but the effect on the HILIC separation was greatly underestimated. Indeed, problems with retention time alignment have been observed in the data (discussed in chapter 2.3.2) but the cause – variation in urinary overall concentration has not been recognized.

Creatinine concentration is often measured as a surrogate for total urinary concentration in clinical diagnostics or general biologic monitoring studies based on urine samples (76). In the given sample set of the NUPREDM study baseline cohort the creatinine concentration varied,

apart from the potential outlier marked with a circle, about factor 13.5 from lowest to highest creatinine concentration Figure 28.

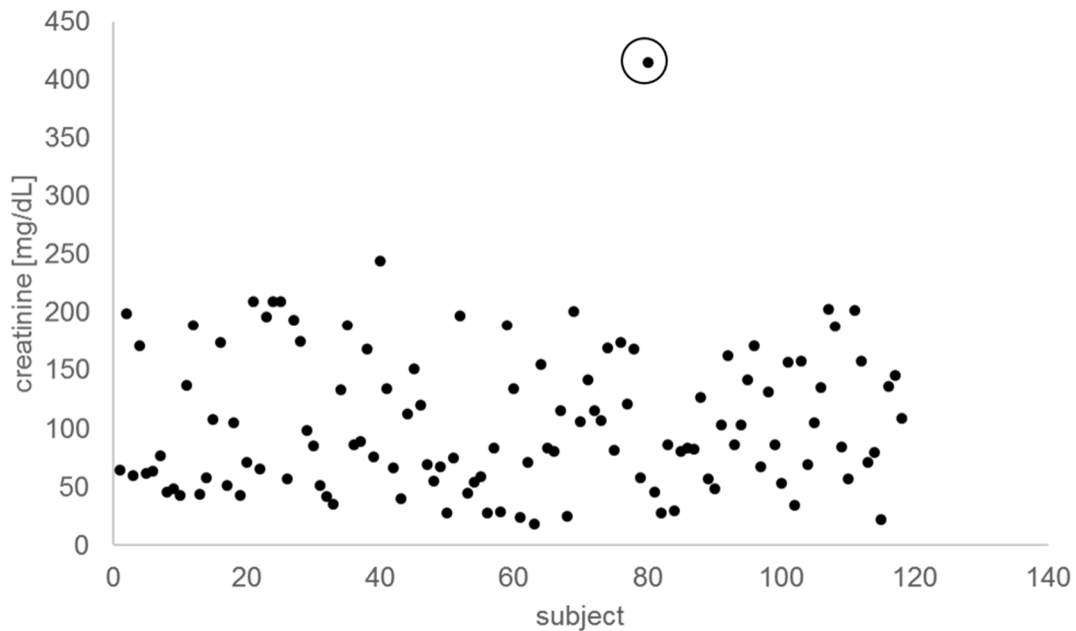


Figure 28: Creatinine concentrations of NUPREDM study baseline samples; a potential outlier is marked with a circle.

Evaluation of internal standards retention time showed a clear dependency of retention time and urinary concentration (see Figure 29).

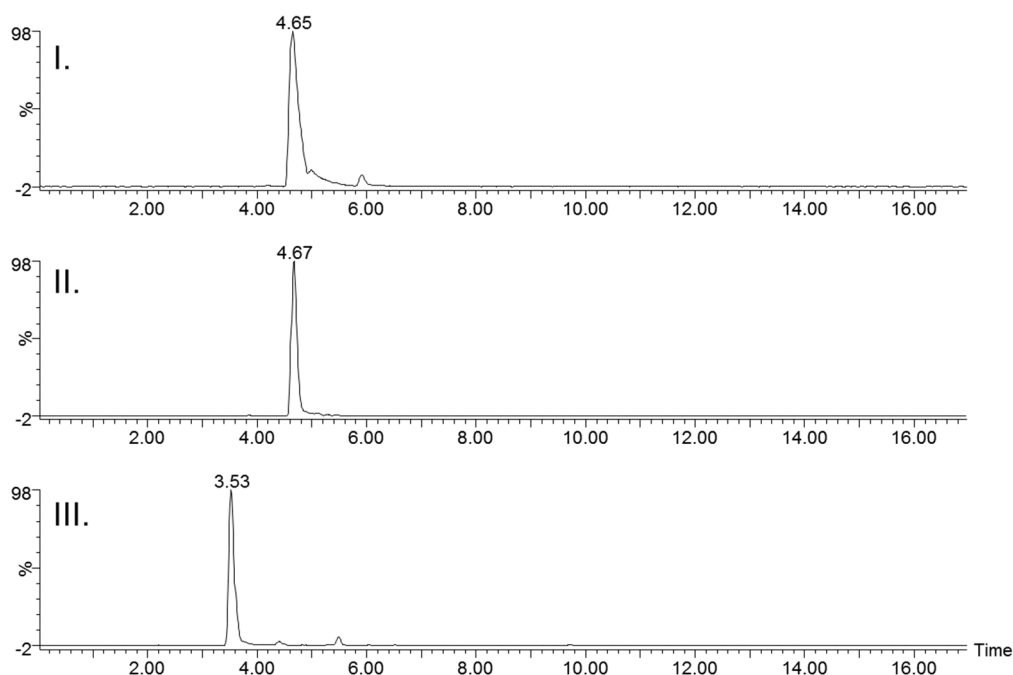


Figure 29: Retention time of [D10]-adipic acid in I. spiked blank, II. low concentrated urine sample and III. high concentrated urine sample

Besides the retention time shifts, also signs of column overloading in particular peak deteriorations like fronting, tailing, peak splitting and peaks with plateaus can be noticed in chromatograms of high concentrated samples (see Figure 30).

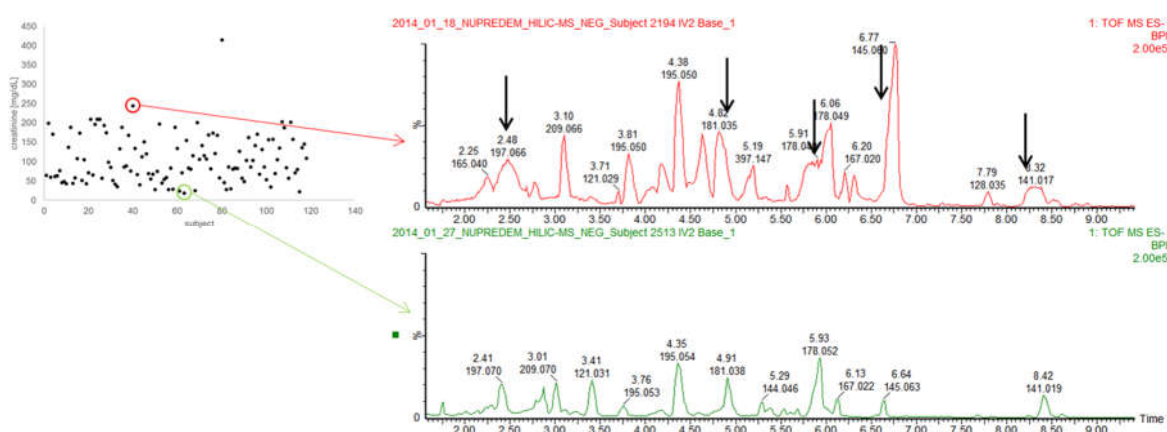


Figure 30: Exemplary Base peak chromatograms (BPC) of high and low concentrated urine samples, peak deteriorations putatively caused by high urine concentration is indicated with black arrows.

The determination of urinary creatinine to estimate the overall concentration level is probably the most frequently applied technique when targeted analysis (77) e.g. drug abuse testing (78,

79) are conducted, but also in untargeted metabolomics approaches this strategy is used. Here, the measured analyte concentration is normalized after the analysis with the corresponding creatinine value.

The presumably most accurate normalization method is the acquisition of 24-h collection urine or normalization to the urinary excretion rate (UER) (80, 81). Nevertheless, both methods are considered to be impractical if applied in large cohorts and are error prone, e.g., in regard to completeness of collected voids (82). Besides normalization to creatinine, other measures such as liquid chromatography coupled to mass spectrometry (LC-MS) total ion current, LC-MS total useful signals (LC-MSTUS) (83), probabilistic quotient normalization (PQN) (84), median fold change (85), osmolality (86-88), urinary volume or specific gravity (SG) (89), are applied.

In clinical, forensic and general scientific environments, urine normalization to creatinine or SG attained general acceptance since these methods offer a tenable balance between applicability and accuracy. Regardless of the specific reference used, all urine normalization procedures are limited to a single reference parameter – a fact that is accompanied by some major drawbacks. Each normalization parameter is error prone and the nature of the bias differs strongly between parameters. Creatinine, for example, not only reflects the overall urinary metabolite concentration but it also can be affected by physical activity, renal impairment, diseases in general, nutrition, sex, or age (76, 90). In addition, specificity, accuracy and robustness against interfering substances of currently applied methods for creatinine measurements such as enzymatic or Jaffe procedures are in discussion (91-93). Whereas SG can be distorted by diabetes, starvation, or nephrotic syndrome (89). Reduction of such normalization bias will increase confidence in data normalization and further reduce the necessity of 24-h urine collection, especially for diseased and spot-urine samples.

To tackle these issues, a concept established and successfully applied in non-targeted metabolomics is picked up and transferred to the current issue. Instead of relying on a single

value for urine normalization – multiple parameters are considered. The hypothesis established by the author was, that a poly parametric urine normalization factor (PPUNF) would deliver more accurate results for urine normalization than normalization to a single parameter in particular creatinine. The hypothesis is based on the assumption that the reason that caused a bias in one parameter will not affect the other parameters. This way an outlier in one parameter will be balanced out by the other parameters and the multifactorial nature of the PPUNF will result in a more accurate and precise urine normalization.

To identify suitable markers for urine concentration estimation and PPUNF calculation the NUPREDM baseline cohort was screened with help of different analytical measures such as urinary creatinine and urea concentration, which were already determined at the University Hospital Tübingen. In addition, also other parameters like LC-MSTUS, SG, NMR PQN and UV absorption were determined. These parameters were benchmarked against each other in a cross-validation like approach to evaluate the suitability as a urine normalization factor individually, but also to estimate accuracy and precision of the individual parameters towards urine normalization. The 4 most appropriate parameters were selected for PPUNF calculation in particular creatinine and urea concentration, SG and UV absorption area under the curve (see Figure 31). The selection criteria aimed towards a reasonable balance between maximized accuracy and precision and minimized analytical effort.

To demonstrate the advantages, the PPUNF normalization was compared to a traditional single parameter creatinine normalization in a proof of concept study which was performed in a routine clinical scenario (94-96). Forty-eight 24 h urine collections were sampled and the albumin-to-creatinine ratio (ACR) and PPUNF-normalized albumin concentrations were benchmarked against 24-h urinary albumin excretion (24-h-UAE).

Finally, the PPUNF is applied to dilution normalize the samples (pre-acquisition normalization) of the NUPREDM baseline the cohort. The pre-acquisition cohort was then reanalyzed with the HILIC-MS setup.

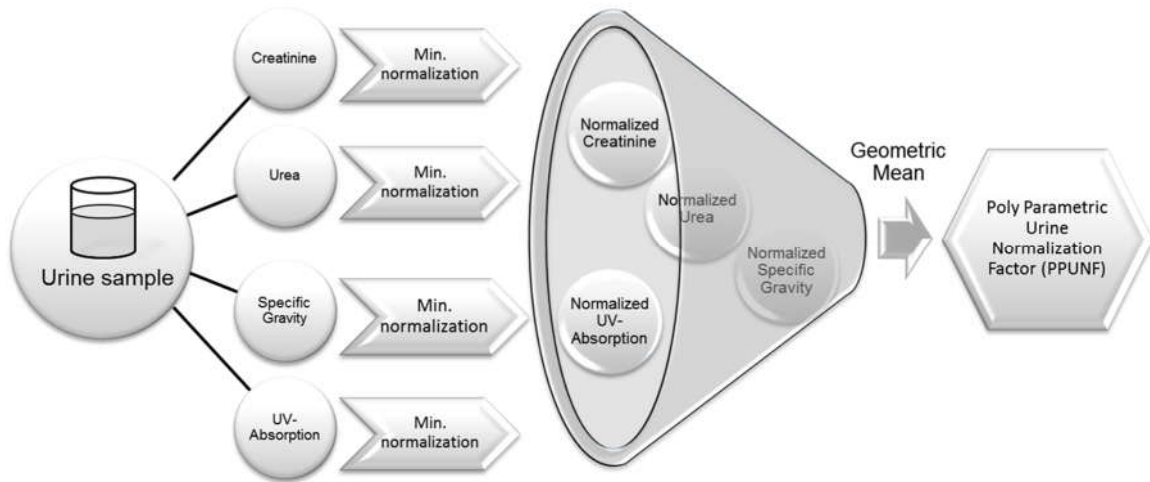


Figure 31: Schematic illustration of PPUNF calculation

3.1.2 Material and methods

Samples and study design

Spot-urine samples from the NUPREDM base line study including 118 healthy individuals were used. The cohort included 79 female and 39 male participants with a median age of 43 years (ranging from 22 to 71 years). For a proof of concept study, 48 randomly selected 24-h urine sample collections from a clinical chemical routine laboratory were used. All urine samples were obtained at the University Hospital Tübingen. Informed written consent was obtained from all subjects and the local medical ethics committee approved the protocol.

Standards

All standards and LC–MS grade acetonitrile and methanol were obtained from Sigma-Aldrich, (Taufkirchen, Germany). LC-MS grade water was obtained from a MilliQ water purification system (Millipore, Molsheim, France).

SG

SG measurements were performed with a digital clinical handheld refractometer (RD. 5712 Euromex, Arnhem, Netherlands). Before measurements the refractometer was calibrated with LC–MS grade water. A 120 μ L aliquot sample was placed on the lens, and the SG was determined. The lens aperture was rinsed extensively after each sample with LC-MS grade water.

Urinary urea, creatinine and albumin quantification

Urinary urea and creatinine were measured enzymatically on the ADVIA 1800 clinical chemistry analyzer (Siemens Healthcare Diagnostics, Eschborn, Germany). Urinary albumin concentrations were determined using a BN ProSpec nephelometer (Siemens Healthcare Diagnostics, Eschborn, Germany).

NMR measurements

An aliquot of each urine sample was analyzed with ^1H NMR analysis. 120 μL of urine was mixed with 60 μL of NMR buffer (90 % D_2O 500 mM PO_4 buffer with 0.1 % TSP, pH 7.4), centrifuged and the supernatant transferred to 3 mm outer diameter NMR Bruker Match tubes (Hilgenberg GmbH, Malsfeld, Germany). Samples were analyzed in randomized order. NMR spectra were acquired on a Bruker 800 MHz spectrometer (Bruker Biospin, Rheinstetten, Germany) operating at 800.35 MHz with a quadrupole cryogenic inverse probe. For an overview of all molecules present in the sample, a standard 1-dimensional (1D) pulse sequence [recycle delay (RD)- 90° - t_1 - 90° - t_m - 90° -acquire FID] was acquired, with water suppression irradiation during RD of 2 s, mixing time (t_m) set on 200 ms and a 90° pulse set to 10.13 μs , collecting 512 scans into 64 K data points with a spectral width of 12 $\mu\text{g}/\text{mL}$. Processing of the spectra was carried out in TopSpin 3.2 (Bruker BioSpin, Rheinstetten, Germany). FIDs were multiplied by an exponential decaying function corresponding to a line broadening of 0.3 Hz before Fourier transformation. All spectra were manually phased, baseline corrected and calibrated to TSP (δ 0.00).

UPLC-QTOF-MS measurements

Sample preparation and analytical method for HILIC-MS measurements which served as a basis for LC-MSTUS calculation are described in 2.2.3.

Ultraviolet spectroscopy

Ice-cooled acetonitrile (400 μL), acidified by addition of 0.1% (v/v) formic acid, was added to an aliquot of urine (100 μL). The resulting mixture was vortexed mixed for 10 s and centrifuged at 4 $^\circ\text{C}$ and 20,817 g for 5 min. The obtained supernatants were diluted by a factor of 20 with

a solution consisting of acetonitrile: water 4:1(v/v) acidified by addition of 0.1% formic acid (dilution solvent).

The diluted samples were transferred to 96-well sample plates sealed via a peelable heat-sealing foil. UV-absorption measurements were conducted using the photo diode array (PDA) detector of an Acquity-UPLC system bypassing the chromatographic column via 5 m (127 μm inner diameter) of peek tubing. For analysis, 10 μL of the sample (full loop) were injected and UV absorption between 215 and 400 nm was recorded with a resolution of 2.4 nm and a 10 Hz sampling rate. As carrier solvent, a solution similar to the dilution solvent, was pumped through the system with a flow rate of 300 $\mu\text{L}/\text{min}$. The resulting peak was integrated. Measurements were performed in triplicates and the resulting peak areas were averaged using arithmetic mean.

Data processing

Chromatogram alignment and peak picking for non-targeted LC-MS analysis were performed using Genedata Expressionist MS Refiner 8.0 and Analyst 8.0 (Genedata, Munich, Germany). Peaks were discarded if they were not detected in both duplicate measurements. Peak areas were averaged between duplicates. The obtained peak matrix was filtered for peaks occurring in 99% or more of the samples. For calculation of LC-MSTUS, peak areas of the remaining peaks were summed up.

NMR data were imported into Matlab (Mathworks, Massachusetts, USA) and further processed, i.e., water region removed (δ 4.70–5.21). The NMR PQN factor was calculated according to Dieterle *et. al* (84).

Pearson correlation analysis was performed using SigmaPlot 12.0 (Systat Software GmbH, Erkrath, Germany). R programming language (R Project for Statistical Computing, <http://www.r-project.org/>) was used for Passing–Bablok regression (Method Comparison

Regression (mcr) package) and ROC curve analysis (Display and Analyze ROC Curves package (pROC) (97)).

Minimum normalization

Minimum normalization for creatinine, urea, UV-AUC, LC-MSTUS, and the NMR PQN factor was calculated as follows:

$$\text{min normalized value } x = \frac{\text{value } x}{\text{lowest observed value}}$$

Minimum normalization for urine SG was calculated as described:

$$\text{min normalized specific grav.} = \frac{(\text{specific gravity}-1)}{(\text{lowest observed specific gravity}-1)}$$

Calculation PPUNF

To create the PPUNF, creatinine concentration, urea concentration, UV-AUC, and SG were minimum normalized. Via minimum normalization, all determined values were expressed as a fold change in relation to the lowest determined value. The relative standard deviation of the minimum normalized values is defined as PPUNF inconsistency. The geometric mean of individually scaled creatinine, urea, UV-AUC and SG values represents the PPUNF.

Calculation of PPUNF-normalized albumin concentrations in a proof of concept study

In the first stage Passing–Bablok regression analysis was performed with PPUNF-normalized albumin concentrations. The obtained slope of the regression equation differed significantly from 1. To enable direct comparison and to simplify interpretation of the results, the values were scaled via inverse application of the established Passing–Bablok regression equation.

3.1.3 Results and discussion

Evaluation of UV absorption as a descriptive factor for urine concentration

To evaluate whether UV absorption is an appropriate parameter to assess urinary overall metabolite concentration LC-UV chromatograms from the baseline cohort in the range of 260-300 nm were extracted (see 2.2.3.6. The resulting chromatograms were smoothed (two times moving average 99 points) and integrated see Figure 32. The resulting value is the UV area under curve (UV-AUC)

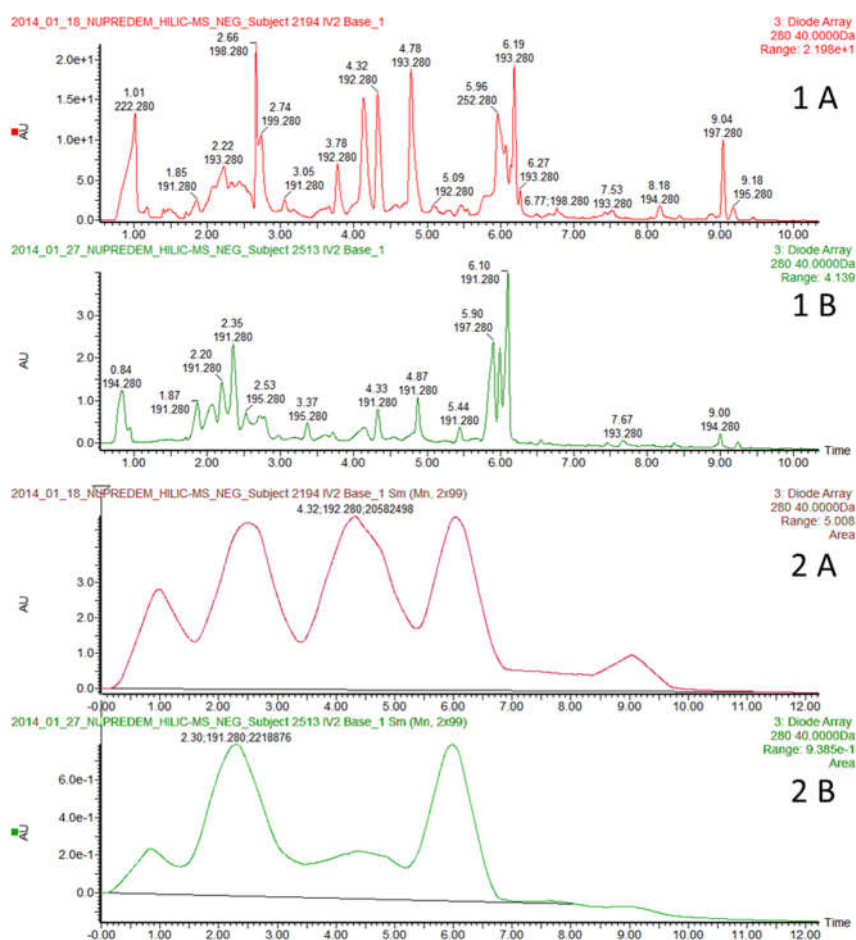


Figure 32: Representative HILIC-UV chromatograms from NUPREDM baseline cohort samples before smoothing 1A & 1B, 2A & 2B corresponding HILIC-UV chromatograms after smoothing and integration

Correlation between UV-AUC and the creatinine values is 0.79 and the visual examination of the scatterplot (see Figure 33) suggests a linear relation (the outlier observed in Figure 28 was neglected).

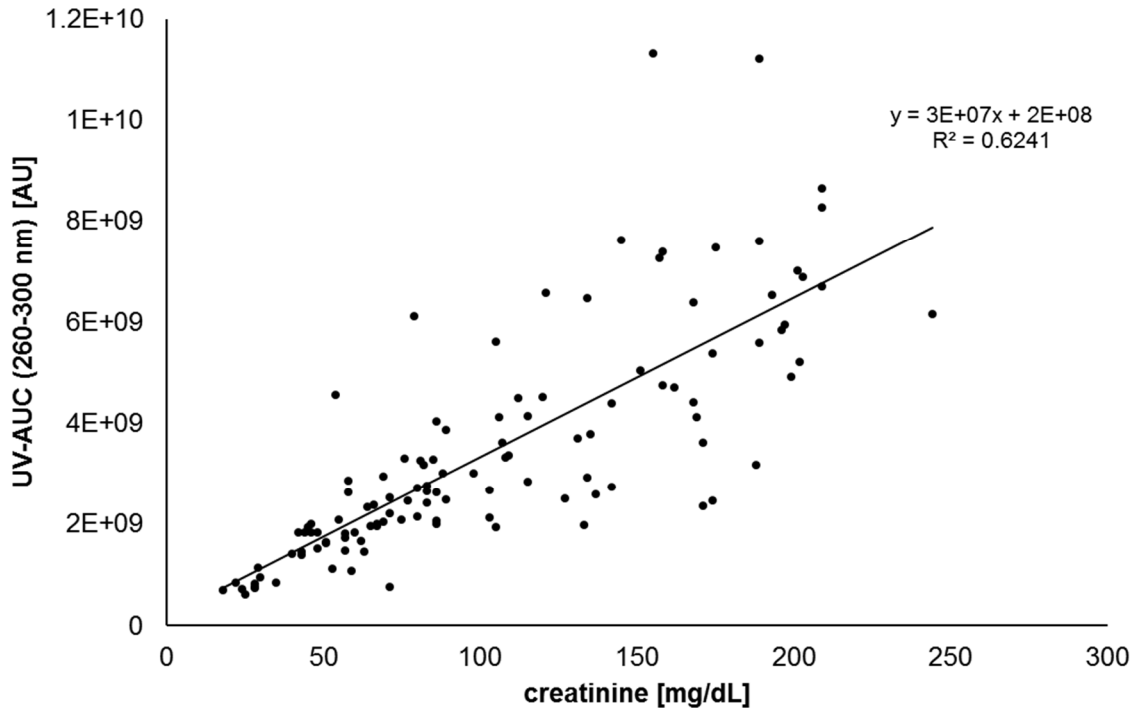


Figure 33: Correlation analysis of HILIC-UV area under the curve and urinary creatinine

Taking into account the high variability of urine samples, the correlation value of 0.79 and the relatively low scattering in the plot are strong indications that UV absorption can be used as a parameter for urine concentration estimation.

On basis of these encouraging results, a further test was performed to check if the UV-AUC determination can be simplified by omitting the HILIC separation. Also, the range of the recorded UV wavelengths were increased from 260-300 nm to 215-400 nm. This way the covered metabolomics range should be more comprehensive. The resulting correlation of 0.82 between creatinine concentration and UV-AUC is even higher compared to the initially determined value.

Screening of data

To provide a comprehensive overview, the values for the determined parameters are illustrated in a scatter plot in Figure 34. Visual examination of Figure 34 suggests direct and

linear proportionality off all parameters when compared against each other. But also the putative outlier in the creatinine values is well recognizable.

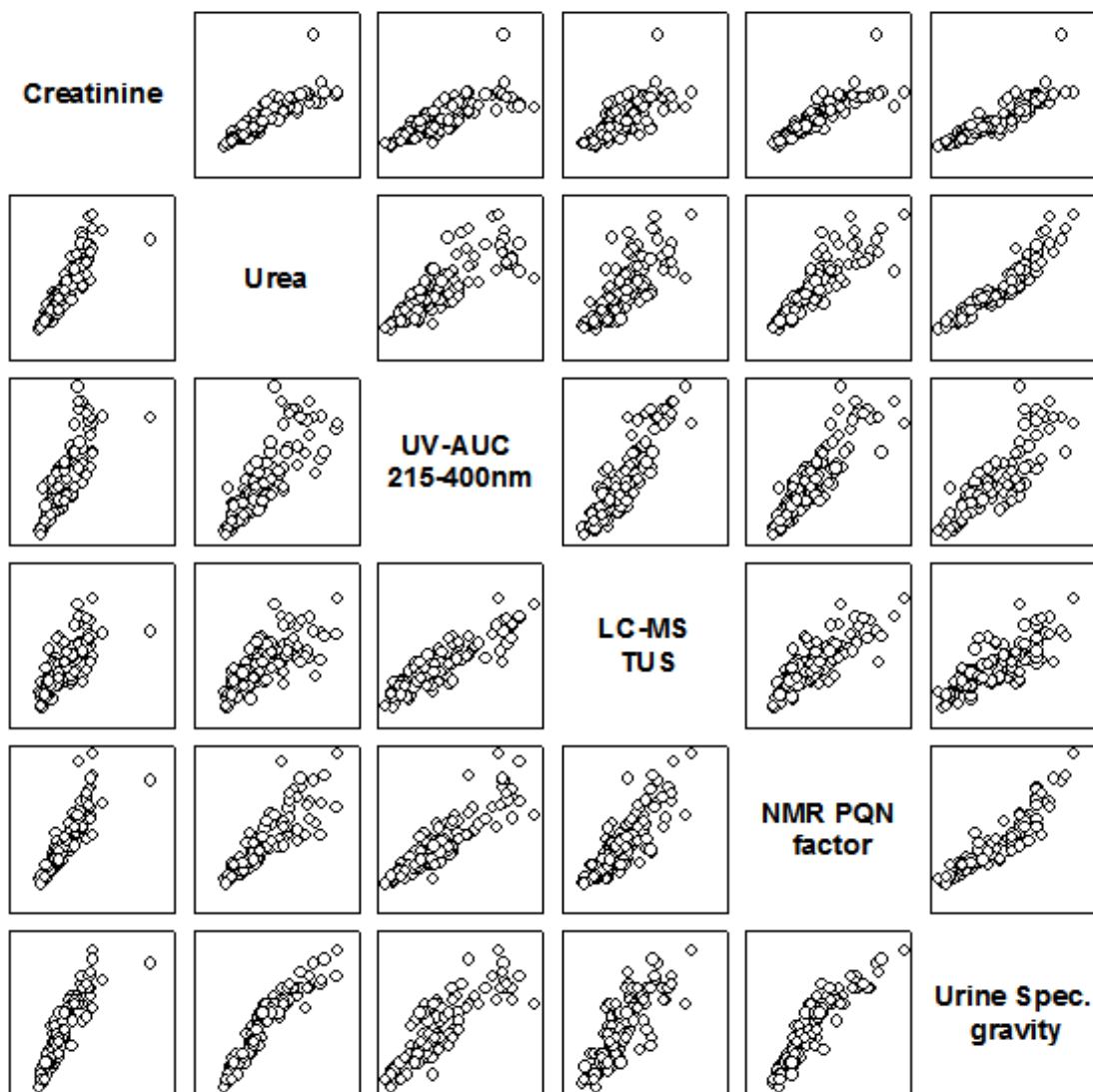


Figure 34: Scatterplots of potential parameters (urinary creatinine, urea, UV-AUC, LC-MS-TUS, NRM PQN and specific gravity) for urine normalization determined in NUPREDM baseline samples

A correlation matrix was established from the obtained data (see Figure 35), to perform a systematic evaluation. Creatinine concentration is regarded as a reference due to high acceptance of the creatinine normalization in the clinical chemistry community. All studied parameters showed significant correlation to each other. In conclusion, all considered parameter might be suitable for urine concentration normalization except for LC-MSTUS which demonstrated lowest performance.

	Creatinine	Urea	UV-AUC 215-400 nm	LC-MSTUS	NMR PQN factor	Urine Spec. gravity
Creatinine		0.882	0.815	0.674	0.880	0.893
Urea	5.09E-39		0.821	0.821	0.869	0.945
UV-AUC 215-400 nm	5.88E-28	6.86E-20		0.843	0.888	0.844
LC-MSTUS	6.12E-19	5.00E-21	8.53E-38		0.789	0.762
NMR PQN factor	1.70E-38	1.83E-36	1.79E-39	1.51E-29		0.921
Urine Spec. gravity	1.62E-37	5.80E-52	7.33E-29	2.24E-22	5.42E-44	

Figure 35: Correlation matrix of potential parameters for urine normalization determined in NUPREDM baseline samples: Upper right corner indicates Pearson correlation values, and p values (bonferoni corrected) are displayed on the lower left side

Generation and evaluation of PPUNF

First screening of the data showed sufficient correlation of all tested parameters except for LC-MSTUS, to be utilized for urine concentration normalization. Nevertheless, an outlier or biased reference can only be identified and compensated when multiple parameters per sample are obtained. To counteract the problems of outliers, a new poly parametric approach was developed to take advantage of multiple parameters. For the creation of the PPUNF the parameters creatinine, urea, UV-AUC, and SG were considered. NMR PQN and LC-MSTUS were excluded. Even though, NMR-PQN demonstrated high correlation to creatinine values NMR measurements were considered to be too expensive, time consuming and cumbersome especially with regard to a routine clinical diagnostics environment. The same applies for the LC-MSTUS determination.

Table 6: Comparison of minimum normalized parameters of extreme subjects

Subject	Creatinine	Urea	UV-AUC 215-400 nm	Urine Spec. Gravity	LC-MSTUS*	NMR PQN Factor*	PPUNF
A	9.4	18.3	9.9	13.5	3.4	8.5	11.9
B	23.1	17.1	14.8	16.0	4.8	17.0	18.0
C	8.6	10.1	18.5	11.0	6.1	11.5	11.7

*not used for PPUNF calculation

To allow a direct comparison of the different parameters, the data were scaled with minimum normalization, provided that all parameters shared the same norm. This way the obtained quantitative data are scaled to a dimensionless factor expressed as fold-changes relative to the lowest identified level of each respective factor in the cohort (see Table 6). Although, NMR

PQN and LC-MSTUS were excluded from PPUNF creation, the scaled data were calculated to allow further investigations. To create the PPUNF value the geometric mean, which is the n -th root of the product of n variables, was used. The geometric mean was preferred since it is less prone to outliers, in comparison to the arithmetic mean, and does not disregard outliers, like median does. As such the geometric mean preserves the central tendency of the data.

In Table 6 the scaled values of three extreme subjects and their corresponding PPUNF's are presented. In each of the chosen subjects one of the parameters appeared to be biased if compared the residual ones. In the calculated PPUNF, putative outliers appeared to be compensated (see Table 6). Besides the putative outliers, it was noticed that the scaled LC-MSTUS values were systematically lower than the residual values in the selected examples. To investigate this observation, a linear regression analysis of the minimum normalized values against the newly introduced PPUNF was performed and the results are shown in Table 7.

Table 7: Linear regression analysis of minimum normalized urine concentration estimators

Linear regression analysis	slope	intersect	R2
Crea – PPUNF	0.87	-0.28	0.895
Urea - PPUNF	1.20	-0.66	0.925
UV-AUC - PPUNF	1.00	0.26	0.842
LC-MSTUS - PPUNF	0.24	1.70	0.620
NMR PQN Factor - PPUNF	1.07	-0.64	0.873
Urine Specific Gravity - PPUNF	0.99	0.92	0.937

The linear regression analysis confirmed that the normalized LC-MSTUS values were systematically lower than the other values. The calculated slope is about 4 times lower compared to the other slopes. All other tested parameters exhibited a slope close to one and

an intercept close to zero which supports the assumption that the tested parameters share a similar norm.

PPUNF validation study

To demonstrate the validity and performance of PPUNF, a proof of principle study was performed. Forty-eight 24-h urine sample collections were randomly selected in a clinical routine laboratory and analyzed for albumin, creatinine, urea, UV-AUC, and SG. Twenty-four hour urinary albumin excretion (24-h-UAE) was determined by nephelometry and regarded as the reference. PPUNF and PPUNF inconsistency were calculated as mentioned above.

The Shapiro–Wilk-test for normality suggested non-normal population of the data set (W -statistic = 0.487, $P < 0.001$). Passing–Bablok regression, which does not require normal distributed data (98, 99), was applied for method comparison supported by Bland–Altman plots and Pearson correlation coefficients (see Figure 36). Passing–Bablok regression analyses showed that 95% confidence interval (CIs) of the regression equations were improved if PPUNF normalization was applied (slope CI -20.1%, intercept CI -27.09%, see Figure 36 A-1 and B-1), and Pearson correlation was significantly increased from 0.806 to 0.932 (Figure 36 A-1 and B-1). Zou's CIs for comparison of dependent overlapping correlations (100) was used for significance testing.

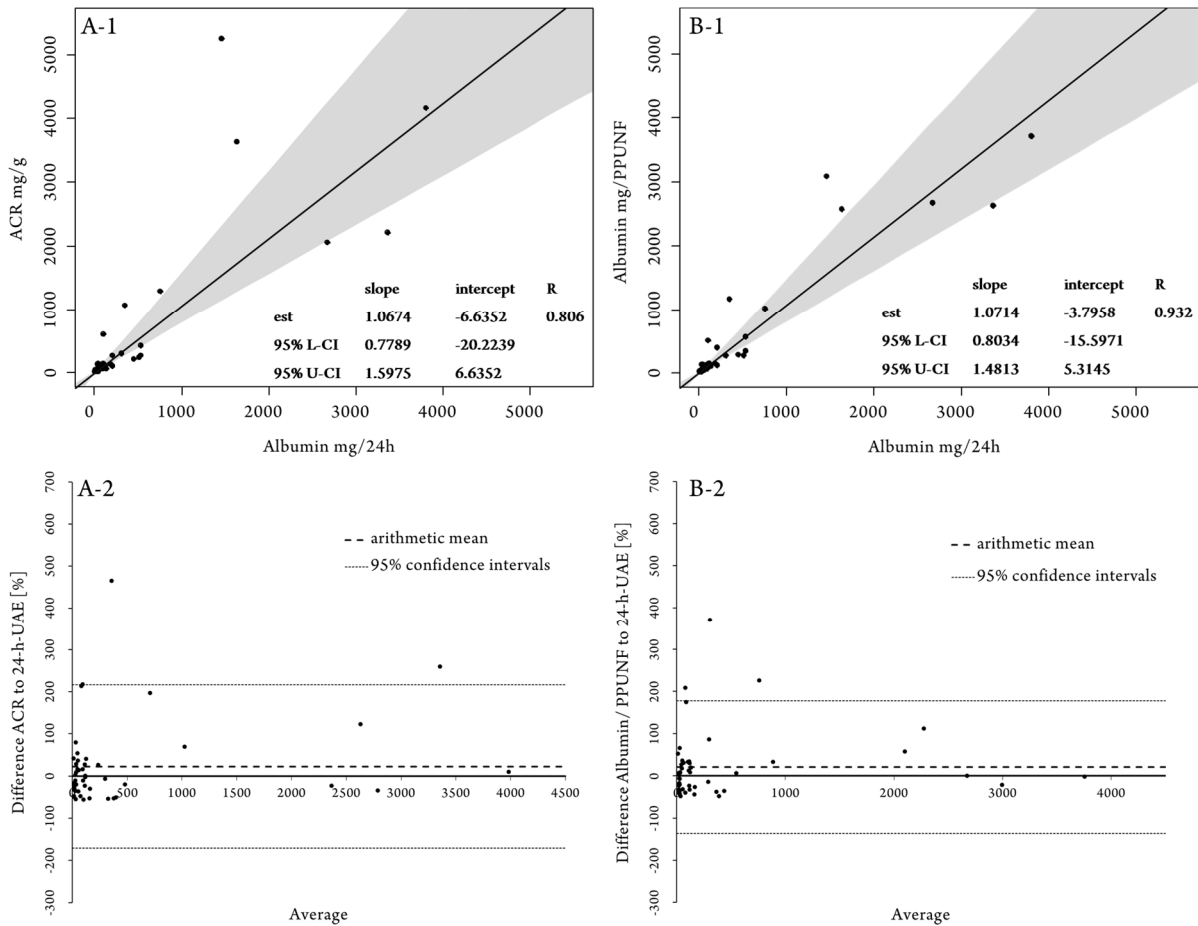


Figure 36: Method comparison plots of validation study samples. A-1 and B-1 are Passing–Bablok regression analysis of ACR and PPUNF data, respectively. A2- and B-2 are Bland–Altman plots of ACR and PPUNF data, respectively

The related correlation coefficient (creatinine to PPUNF) was 0.955, and the difference between the regression coefficients calculated via creatinine and PPUNF normalization was 0.126 with a 95% CI of 0.0862–0.2287. The difference in the correlation coefficients was considered significant since the lower CI is greater than zero. Visual examination of the Bland–Altman plots (Figure 36 A-2 and B-2) affirmed the finding of a more reliable prediction of albumin excretion attributed to PPUNF normalization. The span of Bland–Altman plot CIs was reduced if data were PPUNF normalized.

As seen in Figure 36 B-2, three subjects exceeded the upper CI. Two of them also exhibited high PPUNF inconsistencies. To illustrate this finding, the difference (albumin PPUNF normalized to 24-h UAE) in percent was plotted over the PPUNF inconsistency in Figure 37.

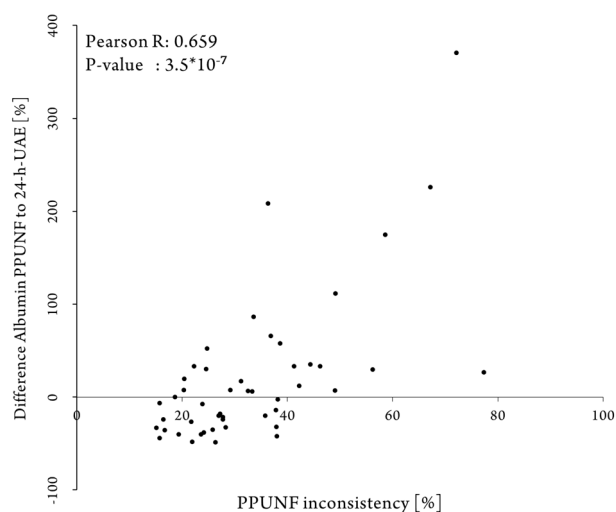


Figure 37: Scatterplot difference in % of albumin PPUNF to 24-h-UAE over PPUNF inconsistency in validation study samples.

Visual examination of Figure 37 supported the assumption that a high deviation of PPUNF-normalized albumin concentrations from the reference values coincided with high PPUNF inconsistencies. The hypothesis was supported by Pearson correlation analysis which showed a moderate but significant correlation of 0.659. This outcome led us to conclude that the PPUNF inconsistency might indicate a biased normalization or could be a hint for an incorrect urine sample, possibly manipulated by the patient. Especially with regard to drug abuse testing (78, 79) or doping screening this finding is highly interesting, but needs further investigation which is beyond the scope of our current study.

In summary application of PPUNF normalization significantly improves data normalization and helps to reduce normalization bias and outliers, but cannot fully prevent them. In the case an outlier occurs, although PPUNF normalization was applied, calculation of the PPUNF inconsistency factor enables its identification and recommends further evaluations.

Determination of minimum references to allow inter-lab transferability and comparability of PPUNF

In a poly parametric urine normalization approach, every parameter needs a minimum reference to allow for minimum scaling of the variables. In the course of PPUNF generation, an array of minimum reference values was ascertained (see Table 8).

Table 8: Observed minimum references in comparison with literature data

	Current Study	Ref (78)	Ref (101)	Ref (102)
creatinine	18.0 mg/dL	33.0 mg/dL	20.0 mg/dL	16.0 mg/dL
urea	190.0 mg/dL	~ 200.0 mg/dL	N/A	N/A
specific gravity	1.002	N/A	1.001	1.002
UV-AUC	308668 AU	N/A	N/A	N/A

With regard to creatinine, urea, and SG, these suggested minimum reference values are in good accordance to existing literature and forensic and clinical official guidelines (see Table 3). To allow application of PPUNF normalization in any analytical laboratory, a minimum reference for UV-AUC needs to be established as well. There are several hurdles to manage if a transferable minimum reference for urine UV-AUC has to be established. The most important issues are as follows: a lack of literature data concerning UV- absorption of urine, vendor-specific arbitrary units, and the variety of UV- spectrometer equipment per se utilizing different path lengths (flow-through cells, cuvettes, plate readers). To address these issues, we propose an independent standard. The minimum reference of 308668 UV-AUC that we

observed in our study equaled to a 12.4 µg/mL uric acid solution. The path length of the flow-through cell in our ultra pressure liquid chromatography (UPLC)-PDA detector is equivalent to 1 cm. Urine samples had to be diluted by a factor of 100 to ensure that UV absorbance varied within the linear range of our detector. If other path lengths are utilized, e.g. a plate reader is used, the dilution factor needs to be adapted accordingly. After adaption of the urine work-up procedure to the respective UV spectroscopy equipment and analysis of the 12.4 µg/mL uric acid standard, it is possible to use PPUNF in any laboratory that is capable of measuring creatinine, urea, SG, and UV absorbance. In routine applications, urinary creatinine and urea are typical screening parameters. Urine SG is a parameter that is also often assessed, but not as prevalent as creatinine determination. Because typical creatinine or urea concentration assays are based on colorimetric tests, UV/VIS spectrophotometers are available in most analytical laboratories. Overall, the application of a PPUNF would be feasible in most analytical laboratory settings at minimum consumption of additional analysis time and without the necessity of acquiring new analytical equipment.

Impact of pre-acquisition normalization (prAN) on HILIC chromatography

In section 3.1.1 the aspect of urine concentration variation and its detrimental effects on robustness of HILIC retention time precision and peak shape deterioration, MS detector saturation and ESI suppression due to highly concentrated urine was discussed. These effects can tremendously hamper the data analysis and multivariate evaluation of such samples in metabolomics screening approaches. In an attempt to counteract these effects, samples were reanalyzed analog the protocol described in section 2.2.3; this time, however, the samples have been diluted with LC-MS grade water according to their PPUNF before sample workup. The resulting retention times of the internal standards were determined and compared in boxplots (see Figure 38).

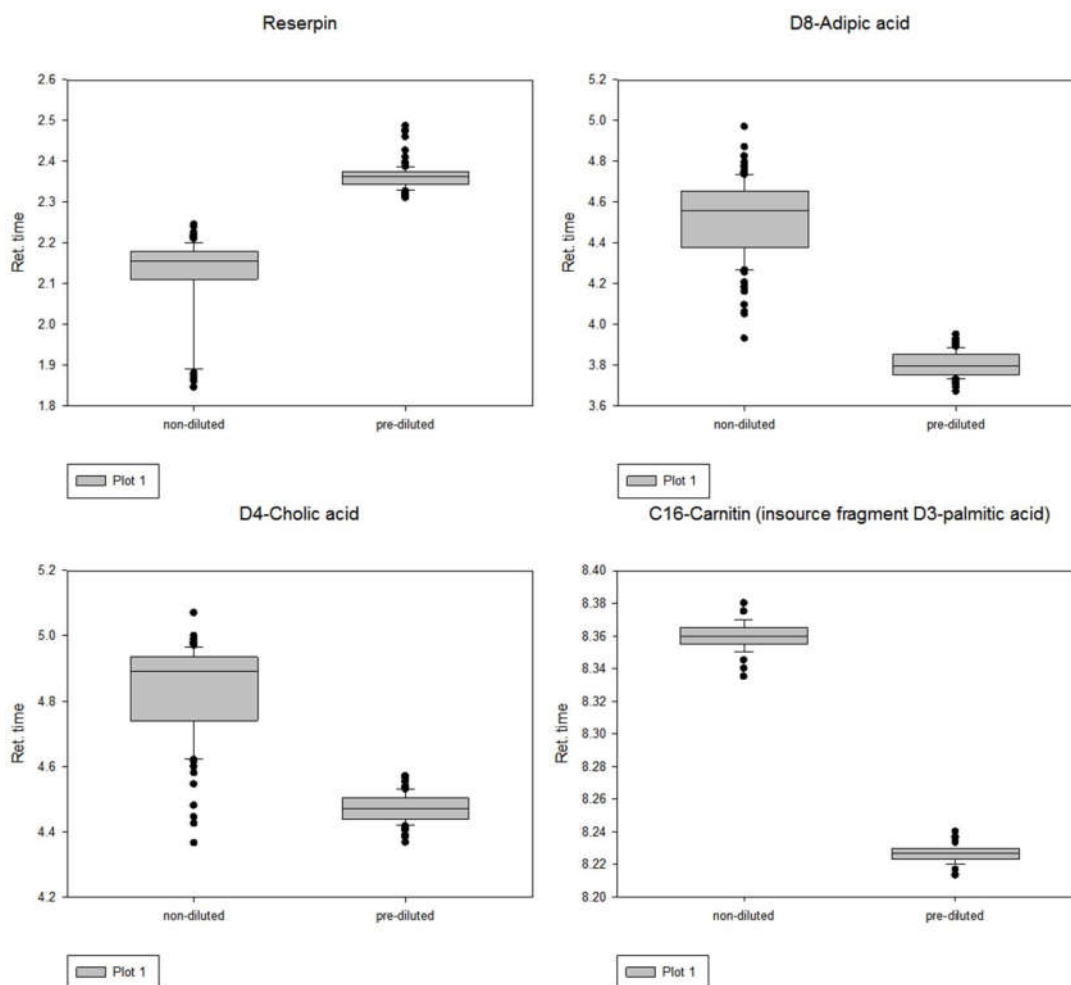


Figure 38: Boxplots of internal standard retention time stability in HILIC-MS analysis of NUPREDM baseline samples with and without pre-acquisition normalization

As the box plots in Figure 38 illustrate, retention time is heavily influenced by sample concentration. Also, the stability of the retention time is greatly increased if the urine samples were dilution normalized before analysis. If a urine sample cohort is analyzed which was treated with pre-acquisition normalization (prAN) it should be more easy to retrieve urinary metabolomic patterns and corresponding multivariate models should be more informative compared to a post-acquisition normalized (poAN) sample set. This originates from the fact that maximum stability of the analytical system is essential (103, 104). Peak alignment will be improved due to more stable retention times and less peak deterioration and the semi-quantitative MS data is less affected from ESI suppression and detector saturation.

In a small gender study the capability to separate female from male urine with a metabolomics screening approach was tested with the prAN and poAN (creatinine normalization) sample

set. Orthogonal projection to latent structures discriminant analysis (OPLS-DA) was carried out using Extended Statistics from Markerlynx package Masslynx 4.1.

The predictive power of OPLS-DA analysis increased, $Q^2(\text{cum})$ from 0.18 to 0.25 when prAN was applied, also scattering of individuals in OPLS-DA score plots is improved (see Figure 39). A brief examination of putative mass annotation of top 20 VIP-score variables suggests amongst others the glucuronidated forms of oxo-androsterone and androstenediol in the prAN dataset (see Supplementary Table 8 and Supplementary Table 9). In contrast, these well-known derivatives of the sexual hormone testosterone do not appear in the top 20 VIP-scores of to the poAN data sets.

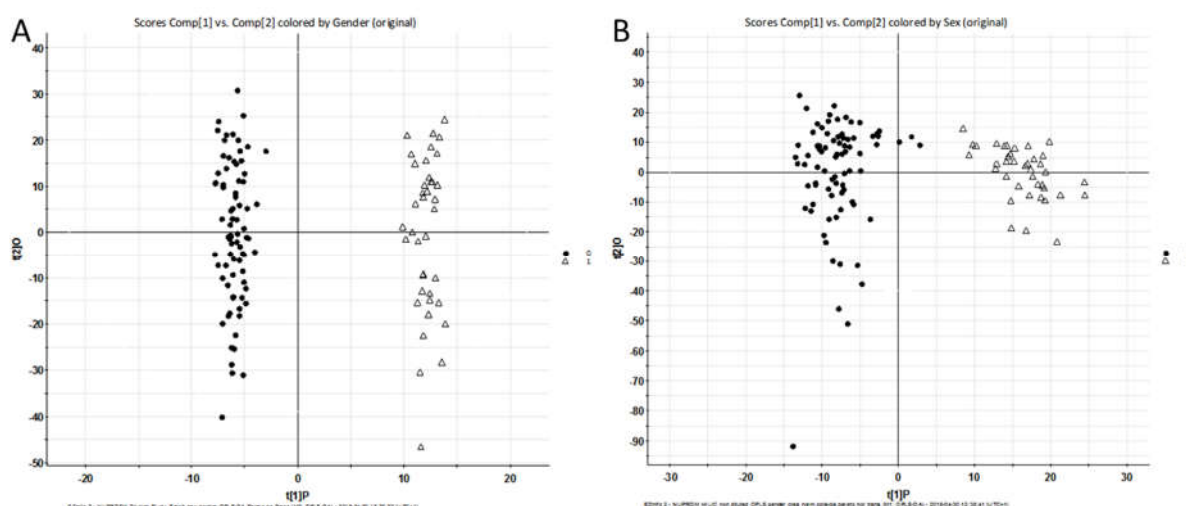


Figure 39: OPLS-DA score plot of urinary HILIC-MS gender study: A) with prAN samples, B) with poAN samples

This example suggests that metabolomics data generated via application of PPUNF prAN yield in more valid results, which further emphasizes the necessity of prAN in metabolomics studies.

3.2 Enhancing of ESI efficiency using 2-(2-methoxy)ethoxyethanol (2-MEE) as ESI dopant

3.2.1 Introduction

Pre-acquisition dilution normalization performed with help of the (in 3.1 mentioned) PPUNF showed clear advantages over post-acquisition normalization in metabolomics screening approaches. A drawback of this approach was of course, that samples needed to be diluted up to a magnitude and more. To encounter a loss in signal intensity or even a complete loss of a feature due to dilution an ESI dopant was searched for.

In 1999 Yamaguchi et al. (105) introduced 2-(2-methoxyethoxy)ethanol (2-MEE) as an ESI dopant. 2-MEE is able to compensate the ion suppression effect of ammonium acetate (NH_4OAc) buffer in negative ESI. Investigations on ibuprofen and its metabolites via ESI-MS, reported an increase in ESI response (up to 100 fold) using ammonium acetate buffer, as mobile phase and 2-MEE as post-column modifier. These results contrast the ones obtained in absence of 2-MEE.

In hydrophilic interaction chromatography (HILIC) ammonium acetate and formiate buffers are quite popular mobile phases (106), since a sufficient cation and anion concentration in the mobile phase has to be provided (39) in order to suppress, minimize or adjust secondary interactions such as ion exchange. Accordingly, the HILIC method developed within this work also utilized an ammonium acetate buffered mobile phase.

In the present case, the application of 2-MEE is ideal since:

- the ESI suppression of ammonia acetate can be reduced
- the dilution effect on the signal intensity due to prAN can be compensated.

The investigation of the beneficial effect of post-column application of 2-MEE during HILIC-MS separations was subdivided into three stages.

In the first stage metabolite standards were analyzed to elucidate the 2-MEE dopant mechanism. An extended/enhanced mechanism, compared to the one provided by Yamaguchi (105), was proposed which is responsible for the dopant effect.

In the second stage the matrix effects of biological samples on the proposed HILIC-MS 2-MEE post-column infusion system was evaluated. For this purpose stable isotope labeled amino acids were spiked into biological samples (human urine and plasma) and analyzed with and without post-column infusion of 2-MEE.

In the third stage the concentration of 2-MEE in the post-column infusion was optimized.

3.2.2 Material and methods

Chemicals

All standards were obtained from (Sigma-Aldrich, Taufkirchen, Germany). Each standard was dissolved in water/acetonitrile/methanol (4:3:1) at a final concentration of 500 mg/l. Each solution was acidified with 0.2% formic acid. 2-(2-Methoxyethoxy)ethanol (Methyl Carbitol) was purchased from Sigma Aldrich. LC-MS grade acetonitrile and methanol were obtained from Sigma Aldrich. LC-MS grade water was obtained from a MilliQ water purification system (Millipore, Molsheim, France).

Biological samples

Pooled samples of human urine and plasma from healthy volunteers were used. Ice cooled acetonitrile (400 μ L) was added to an aliquot of sample (100 μ L). The resulting mixture was vortexed for 10 seconds and centrifuged at 4°C 20817 g for 5 minutes. The obtained supernatants were dried in a vacuum centrifuge.

The evaporated residues were reconstituted in a solution of acetonitrile/methanol (3:1) and 0.2% formic acid (reconstitution solvent). The reconstituted samples were subjected to supersonic bath (5 seconds) vortexed (10 seconds) and centrifuged at 4°C 20817 g for 5 minutes. The supernatants were transferred to UPLC-vials for the analysis. All plasma and urine samples were prepared according to this procedure.

UPLC-QTOF-MS measurements

LC-MS experiments were performed with help of an Acquity-UPLC system (Waters Milford, MA, USA) coupled to a SYNAPT G1 qTOF HD mass spectrometer (Waters Micromass, Manchester, UK). In post column infusion experiment 2-MEE flow was facilitated via an Ultimate Plus separation system (Dionex/LC Packings). The following ESI parameters were applied: capillary voltage 2.3 kV, sample cone 24 V, extraction cone 4.0 V, source temperature 120°C, desolvation temperature, 350°C, cone gas Flow 20.0 L/h, desolvation gas flow 800L/h. Scan time was set to 300 ms in and m/z range from 50-1000 in V-mode (average 9000 resolution).

HILIC separations were performed on an Acquity UPLC BEH HILIC (150x2.1 mm 1.7µm, Waters) column in isocratic and gradient mode. The column temperature was set to 40°C. Mobile phase A consists of 10 mM ammonium acetate and 0.2 v/v% acetic acid in water, mobile phase B consists of acetonitrile. The flow rate was set to 300 µL/min, the injection volume was set to 5.0 µL. Isocratic elution was performed with 33% mobile phase A and 67% mobile phase B for 10 minutes.

Gradient elution was performed with a linear gradient: 0-2 min: 5%A, 15 - 20 min 50% A, 20.5 - 25.5 min 5%A. For spiking experiments stable isotope labeled standards were pooled and diluted in reconstitution solvent to a final concentration of 20 mg/L each. For all other HILIC-MS experiments standards were diluted in reconstitution solvent to a final concentration of 20 mg/L each. 2-MEE was post-column infused with a flow rate of 150 µL/min. A source: waste

(1:1) split was applied in non-post column infusion experiments. In post-column infusion experiments a source: waste (1:2) split was applied in order to maintain constant flow rate of ~ 100 $\mu\text{L}/\text{min}$ to the ESI-source.

A mixed mode separation experiment was performed on an Scherzo SM-C18 (150x2.0mm 3.0 μm , Imtakt) column. The column temperature was set to 35°C. Mobile phase A: water 2.5 mM ammonium acetate + 0.05% v/v acetic acid. Mobile phase B: 50% acetonitrile, 50% Isopropanol, 25 mM ammonium acetate + 0.5% acetic acid. The flow rate was set to 250 $\mu\text{L}/\text{min}$, injection volume was 5 μL . Gradient elution was performed with a linear gradient: 0-2 min: 3%B, 18-22 min 99%B, 22.1-25.5 min 3%B. A standard containing 50 g/L butyric and 50 g/L valeric acid dissolved in water was injected. Post column infusion was performed at 250 $\mu\text{L}/\text{min}$ with pure 2-MEE or mixtures of 2-MEE and isopropanol to reach a final 2-MEE concentration of: 50%, 20%, 10%, 4%, 2%, 1%, 0.4%, 0.2%, 0.1% and 0%. A source: waste split (1:2) was applied resulting in 166.6 $\mu\text{L}/\text{min}$ being delivered to the ESI source.

Direct infusion experiments

Experiment A:

Standard mixture A

Standard mixture A (see Supplementary Table 10) was diluted by factor 100 in acetonitrile/water (4:1 v/v). The resulting solution was split into 7 aliquots and spiked with increasing ammonia acetate concentrations. Each aliquot was again split into 3 sub-aliquots. The first sub-aliquots were not further modified. The second sub-aliquots were diluted factor 1.5 in acetonitrile/water (4:1 v/v). The third sub-aliquots were diluted by factor 1.5 in 2-MEE. The aliquots were analyzed in triplicates in direct infusion mode (200 $\mu\text{L}/\text{min}$ for undiluted samples, 300 $\mu\text{L}/\text{min}$ for diluted samples). Additionally, a dilution row of standard mixture A was prepared and analyzed (triplicates) to ensure applied concentrations are within linear or

dynamic range of the instrument data not shown (applied dilution factors: 50, 75, 100, 200 and 500). The dilution scheme is illustrated in Figure 40 for simplification.

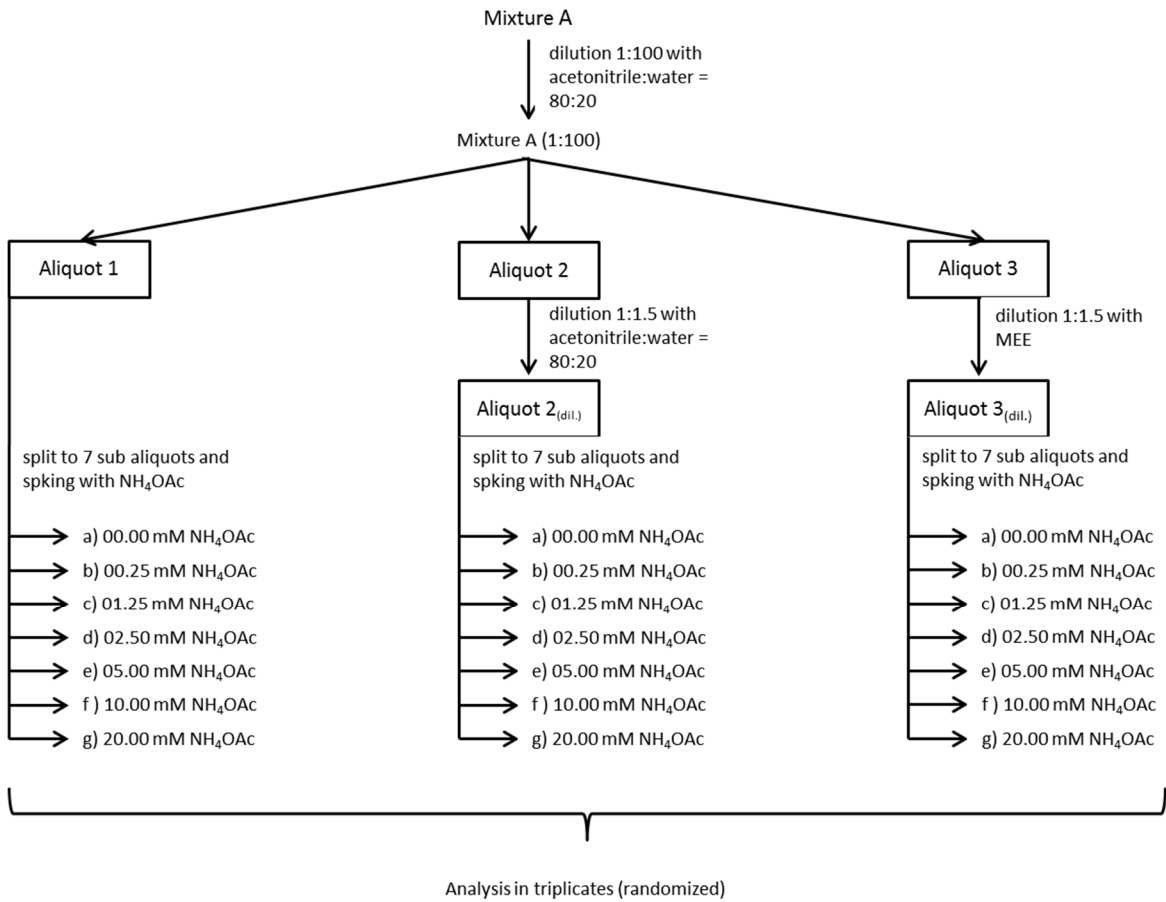


Figure 40: Dilution scheme for direct infusion experiment during 2-MEE dopant evaluation

Experiment B:

Standard mixture B

Standard mixture B containing aspartic acid (133.3 mg/L), lysine (166.7 mg/L) and phenylalanine (166.7 mg/L) were diluted by factor 100 in acetonitrile/water (2:1 v/v). The resulting solution was split into 7 aliquots. The aliquots were spiked with different ammonia acetate and 2-MEE concentrations. All samples were analyzed in triplicates at a flow rate of 100 µL/min.

Standards applied in isocratic HILIC method

L-Leucine, L-phenylalanine, L-tyrosine, L-serine, L-lysine, L-alanine, L-glutamine, L-glutamic acid, glycine, L-isoleucine, D/L-aspartic acid, L-asparagine, L-arginine, L-valine, L-tryptophan, N-acetyl-L-tryptophan, L-proline, L-threonine, D/L-cystein, L-methionine, alanyl-Glutamine, Arg-Gly-Asp-Ser, hippurate, p-hydroxyphenylacetic acid, gallic acid, decanoic acid, lauric acid, palmitic acid, stearic acid, norleucine.

Software and data processing

Physiochemical parameters of applied standards were predicted using Jchem for excel 6.1.0.688 2013 ChemAxon (<http://www.chemaxon.com>).

Chromatogram alignment and peak picking for non-targeted HILIC-U(H)PLC-QTOF analysis was performed using Genedata MS Refiner 8.0. Mass annotation of detected peaks was facilitated via Genedata in combination with HMDB Human Metabolite Database (HMDB) version 3.5(29-31) with a mass annotation error of ± 7 ppm.

Quantification of targeted HILIC-U(H)PLC-QTOF analysis was performed using Targetlynx software (included in Masslynx 4.1 (SCN639)).

UPLC-MS system was controlled via MassLynx 4.1 (SCN 639) software.

3.2.3 Results and discussion

Impact of 2-MEE on ESI response in HILIC mobile phase condition

The beneficial effects of post-column infusion of 2-MEE on ESI efficiency during RP separations carried out in buffered (ammonia acetate) solvents has been shown previously (105). ESI response of ibuprofen and related metabolites was significantly enhanced when post-column infusion of 2-MEE was applied (105). In RP separations polar analytes elute at

relatively low concentrations of organic solvents, such as acetonitrile or methanol. In contrast to RP, HILIC necessitate the use of high percentage of organic solvent (e.g. acetonitrile content 95-50%) which can also be buffered (i.e. NH_4OAc). A set of direct infusion experiments (experiment A see material and method) was performed to estimate the dopant effect of 2-MEE on ESI efficiency in HILIC solvent conditions. Standard mixtures containing different NH_4AcO concentrations were analyzed with and without the addition of 2-MEE. Each experiment was performed in triplicates.

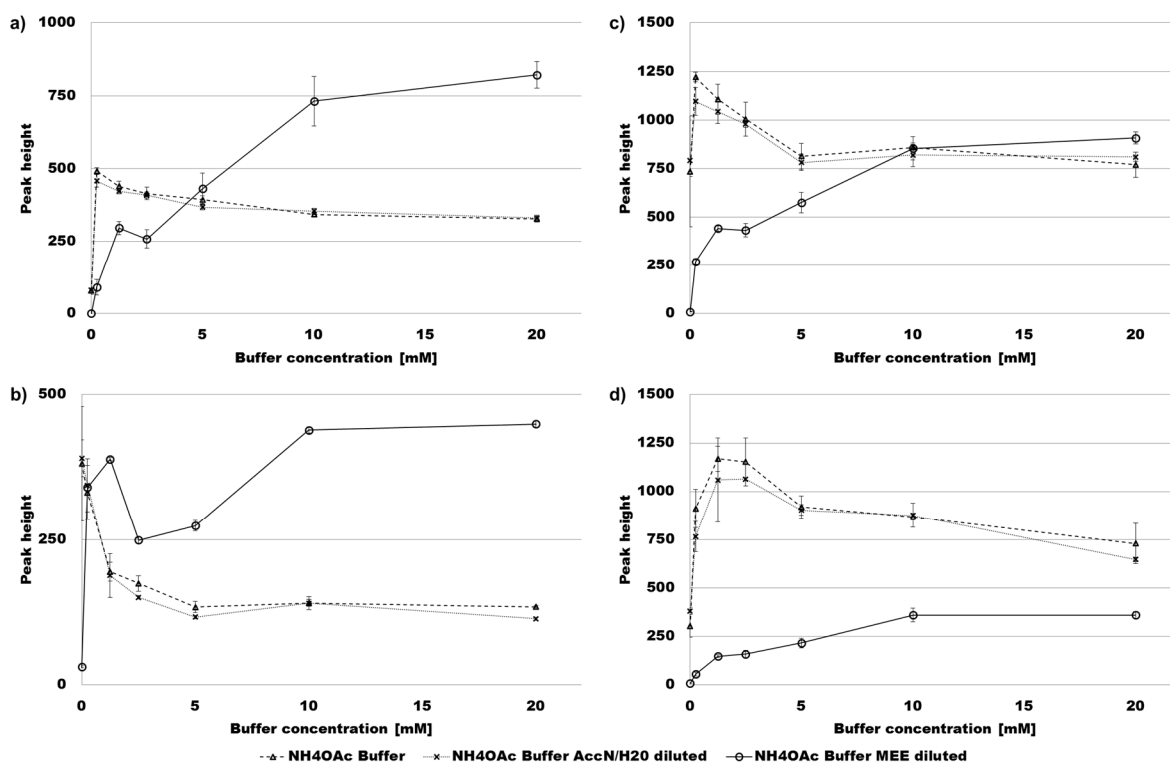


Figure 41: ESI-responses (peak height in counts) to various buffer types and concentrations of different standards, a) L-leucine, b) D10-adipic acid, c) D6-sulfadimethoxine, d) D4-cholic acid.

Figure 41 exemplarily shows the dopant effect of 2-MEE in direct infusion analysis (figures of all analyzed standards can be found in Supplementary Figure 4). Especially highly polar analytes ($m/z < 200$) seem to benefit. As expected a suppressing effect of the buffer arises. Besides a trend can be observed, when 2-MEE is applied increasing buffer concentrations

lead to better ionization efficiencies. Similar effects could be observed when standards were analyzed via HILIC-LC-MS (see Figure 42).

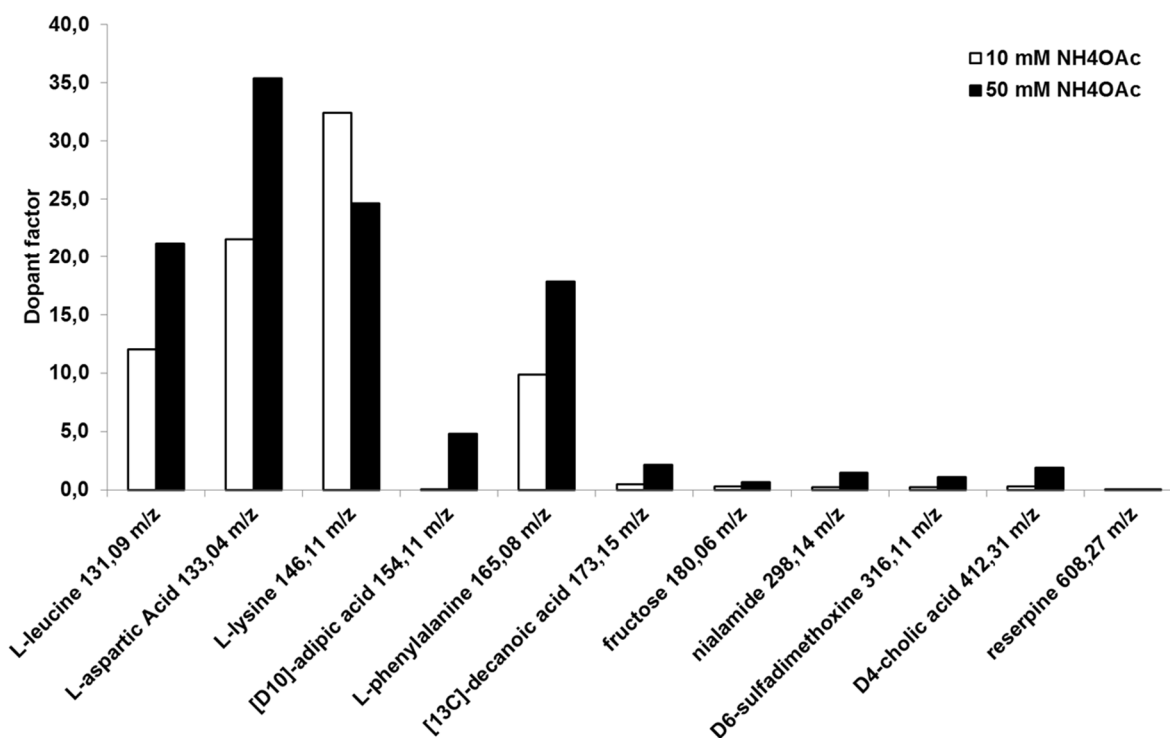


Figure 42: Obtained Dopant factors from HILIC-MS analyses using different ammonia acetate concentrations of standard mixture A

Impact of physicochemical parameters on 2-MEE dopant effect

In order to understand better the dopant mechanism, an extended array of standards (mostly amino acids) were analyzed with and without 2-MEE. Each experiment was performed in triplicates (isocratic HILIC separation method). The resulting peak areas were analyzed and a dopant factor was generated. The dopant factor is the ratio of the peak area under the influence of 2-MEE to the peak area in absence of 2-MEE. Plots of the dopant factor over different physicochemical parameters (via Jchem predicted) are shown in Figure 43. In addition, a plot of the dopant factor based on signal to noise ratios over molecular weight is provided in Supplementary Figure 5. For detailed information see Supplementary Table 11.

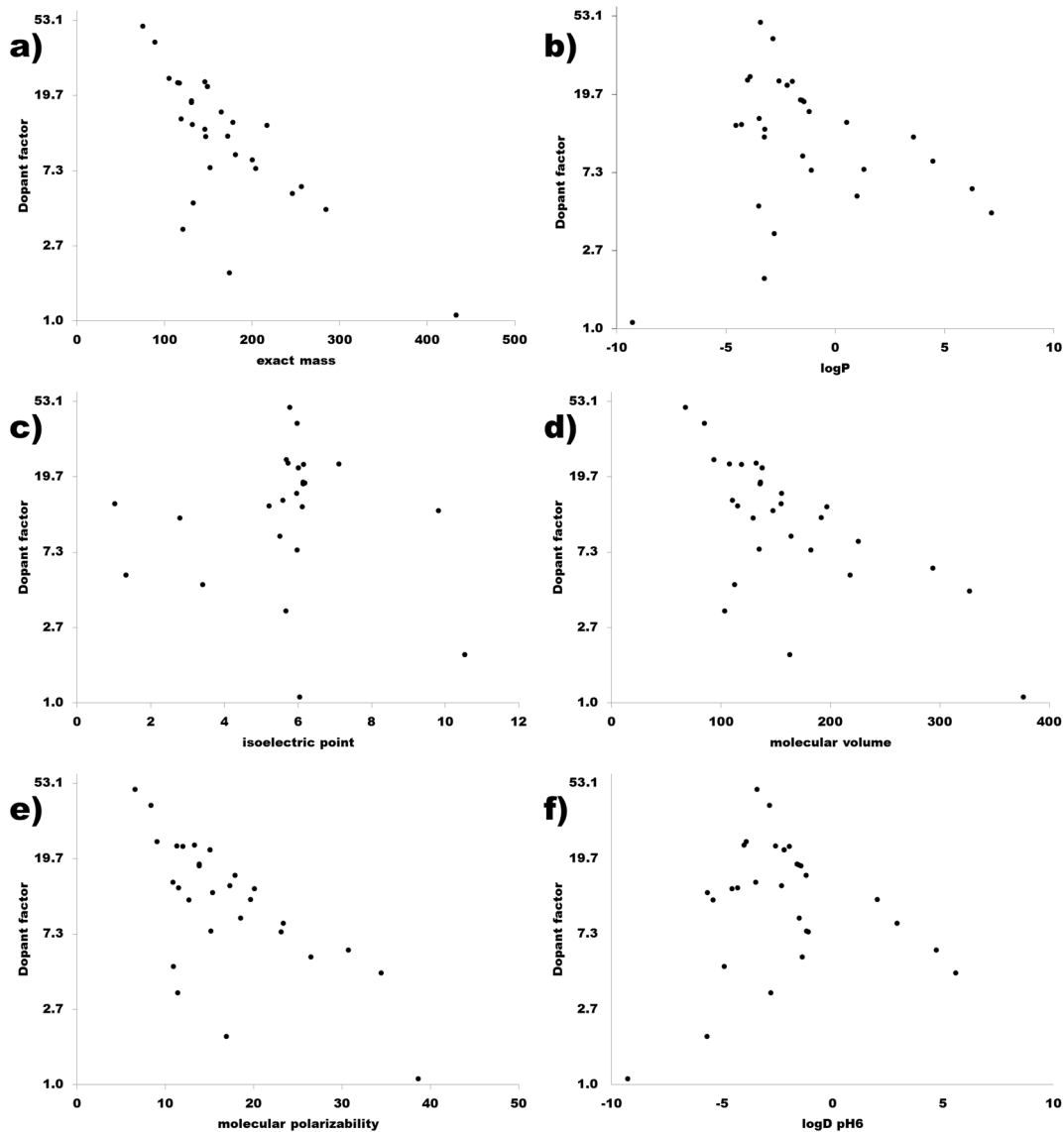


Figure 43: Dopant factor over different physicochemical parameters a) molecular weight, b) logarithmic water-octanol partition coefficient of unionized solutes ($\log P$), c) isoelectrical point, d) molecular volume, e) molecular polarizability, f) logarithmic water-octanol partition coefficient at pH6. y-Axes are ln-scaled

It is possible to observe a logarithmic correlation between dopant factor and: molecular weight, molecular volume, molecular polarizability and $\log D$. These 4 physicochemical properties are confounding factors within the set of standards. The influence of strong acidic or basic moieties such as guanidino-, tert-amino-, sulfo- or phosphorous-groups cannot be examined with the applied standards.

Effect of dopant 2-MEE on metabolites classes

A non-targeted investigation of pooled urine sample from healthy volunteers was conducted via HILIC-U(H)PLC-QTOF-MS with and without post-column infusion of 2-MEE, in order to study the effects of 2-MEE in a more holistic way. Experiments were performed in quadruplicates and detected peaks were filtered out if i) no mass annotation via Genedata was possible, ii) a peak was missing in one of the replicates or iii) a peak was only present in one condition (with or without post-column infusion). The peaks remaining after filtering were classified according to HMDB taxonomy classes and the corresponding dopant factor was calculated. The result is illustrated in Figure 44. In the plot, a trend was observed, molecules which exhibit a small molecular weight, are highly polar and or of amphoteric character benefit most from the 2-MEE dopant effect. Molecules containing strong acidic functional groups such as sulfone – or phosphate groups exhibit only a small or even no dopant effect at all.

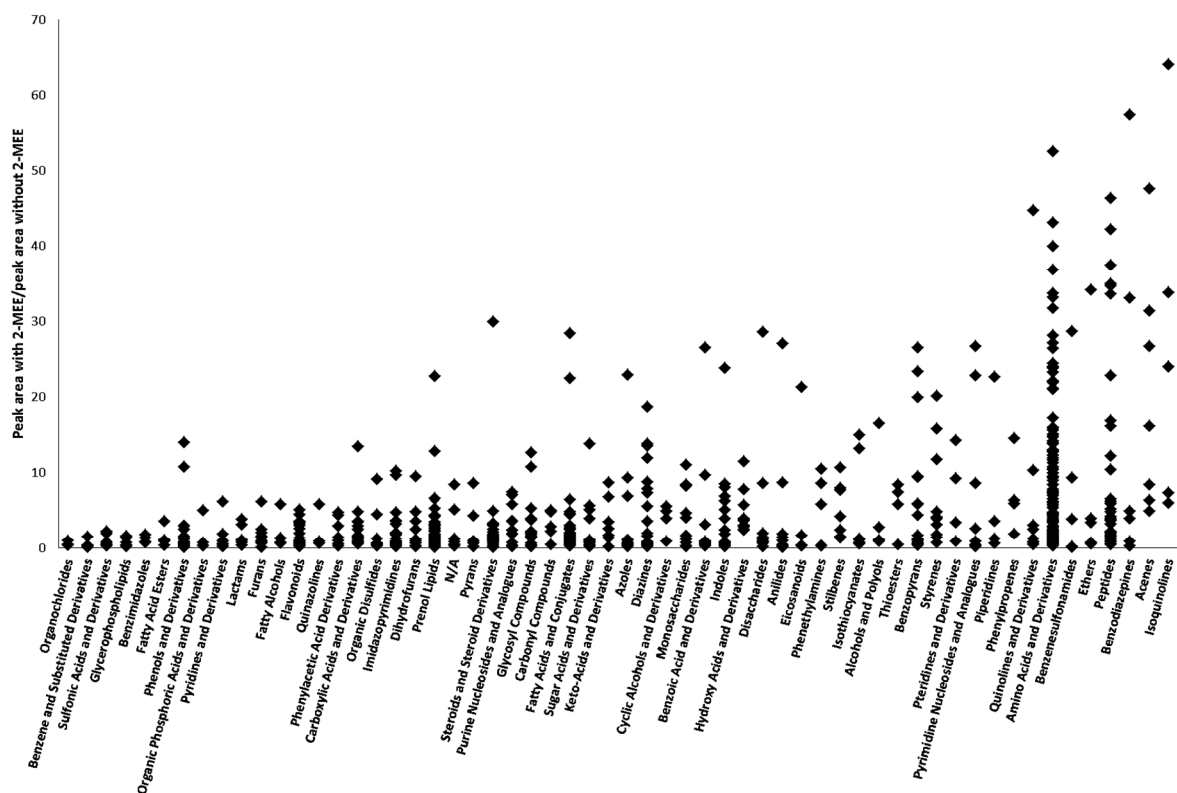


Figure 44: Dopant factor over HMDB metabolite taxonomy class. Individual dots depict the dopant factor of mass annotated peaks and their corresponding HMDB metabolite taxonomy class resulting from non-targeted HILIC-MS analysis of human urine samples.

Synergistic effect of ammonium acetate buffer and 2-MEE application

A mechanism was proposed by Yamaguchi (105) to explain the increase of ESI efficiency of 2-MEE. In summary: a buffer additive (i.e. ammonia acetate), which is used to achieve reasonable chromatographic performance, leads to excessive formation of acetate anions on the surface of the droplets created in during electrospray ionization. Thus, the analyte molecules eluting from the column are in direct competition for negative charges with the acetic acid. This circumstance results into suppression of the analytes to a greater or lesser extent, depending on the physicochemical characteristics of the analyte molecules. Addition of a solvent with a high boiling point ($2\text{-MEE}_{\text{bp}} \sim 194^{\circ}\text{C}$) results in the selective evaporation of neutral acetic acid molecules ($\text{acetic acid}_{\text{bp}} \sim 118^{\circ}\text{C}$) from the ESI droplets.

An additional experiment was performed (direct infusion experiment B) to further elucidate this hypothesis. A mixture containing 3 amino acids (i.e. Asp, Lys and Phe) was spiked with different ammonia acetate and 2-MEE concentrations respectively. The obtained combinations were analyzed via direct infusion ESI-QTOF-MS.

Considering the hypothesis of Yamaguchi (105) the following results were anticipated for the respective experiment:

1. analysis without any modifier results in highest signal intensity, considered as reference
2. application of ammonia acetate results in significant electrospray suppression depending on ammonia acetate concentration applied
3. addition of 2-MEE will result in signal intensities similar or only slightly lower compared to the signal intensities obtained in 1.
4. addition of ammonia acetate and 2-MEE will result in signal intensities similar or only slightly lower compared to the signal intensities obtained in 1.

The actually obtained results are illustrated in Figure 45. As expected, each of the additives alone produced a suppressing effect on signal intensities. Surprisingly, the obtained signal intensities of the amino acid mixture in combination with the addition of ammonia acetate and 2-MEE are much higher compared to the analysis without any additive. It was also observed, that the amount of applied solvent to 2-MEE ratio can be reduced from 2:1 to 4:1, without sacrificing the beneficial effects of 2-MEE.

The observed findings indicate that 2-MEE not only exhibits a compensating effect for acetic acid ESI suppression. It is concluded, that a mechanism, incorporating multiple steps, which increased electrospray ionization efficiency is prevalent.

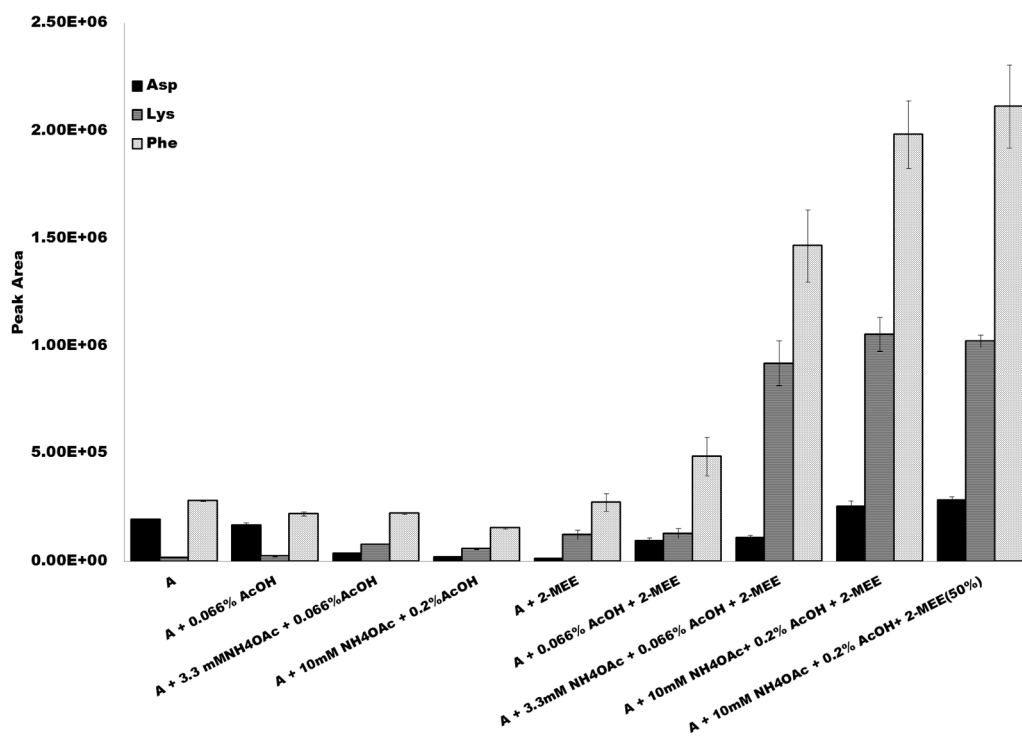


Figure 45: Influence of buffer additives and 2-MEE on signal abundances (with A= acetonitrile:water (2:1v/v))

It is assumed that acetic acid evaporates from the ESI-droplets due to its lower boiling point compared to 2-MEE, this leads to less competition for negative charge and hence to less ion suppression as postulated by Yamaguchi (105). On top of that, we assume that evaporation of acetate anions has three additional effects:

1. Acetate anions draw protons from the ESI droplets during the transfer from a charged acetate anions state in liquid phase to neutral acetic acid molecules in the gas phase. Due to this proton stripping process analyte molecules deprotonation is enhanced and more negatively charged analyte molecules are generated.
2. When acetate anions concentration is decreased, also the surface charge of the formed ESI droplets is decreased, this allows formation of smaller droplets since Rayleigh limit is reached at smaller droplet size. Smaller ESI droplets result in a better ESI efficiency (107, 108).
3. Due to the high boiling point of 2-MEE, 2-MEE is assumed to be evaporated at lower ratio compared to acetonitrile, water and acetic acid. Thus, 2-MEE is getting concentrated in the ESI droplets. 2-MEE also exhibits a relative high viscosity (~3.48

mPa*s) compared to the other used solvent and additives (waters: 1.0 mPa*s). A high viscosity also stabilizes ESI droplets and allows formation of smaller ESI droplets.

The suggested mechanism corresponds to the finding that especially small, less acidic and very poly compounds benefit from the 2-MEE dopant effect and is also in accordance to the charge residue process theory (109). In this theory, it is assumed that small and polar molecules tend to stay in the inner part of the ESI droplets. This tendency is even enhanced if lipophilic or surface active molecules are also present in the droplets. As a result, the probability of getting ionized is lower for small and polar compounds in ESI. If ESI droplets are getting smaller, the relative surface area increases in return and the ESI efficiency increases especially for small and polar compounds.

To verify this hypothesis ESI droplet size could be measured via phase Doppler anemometry (PDA) (110, 111). Additionally there is possibility to indirectly follow the droplet size via concentration determination of a fluorescence resonance energy transfer (FRET) dye pair. Reducing ESI droplet size due to solvent evaporation leads to an increase of the dissolved dye pair concentration until FRET occurs (111).

2-MEE benchmarking of dopant effect under realistic conditions

To evaluate the performance of the dopant 2-MEE under realistic conditions, in particular analysis of a complex matrix, stable isotope labeled amino acids were spiked into pooled urine and plasma samples of healthy volunteers. The types of samples were chosen to evaluate the 2-MEE dopant effect in presence of matrix effects originating from diverse biological sources. The two samples types were chosen to induce matrix effects from either highly polar and small molecules originating from urine or hydrophobic compounds like lipids (112) or peptides as a result of incomplete protein precipitation (113) from plasma.

HILIC-LC-MS analyses were performed with and without post column infusion of 2-MEE. The obtained peak areas are depicted in Figure 46.

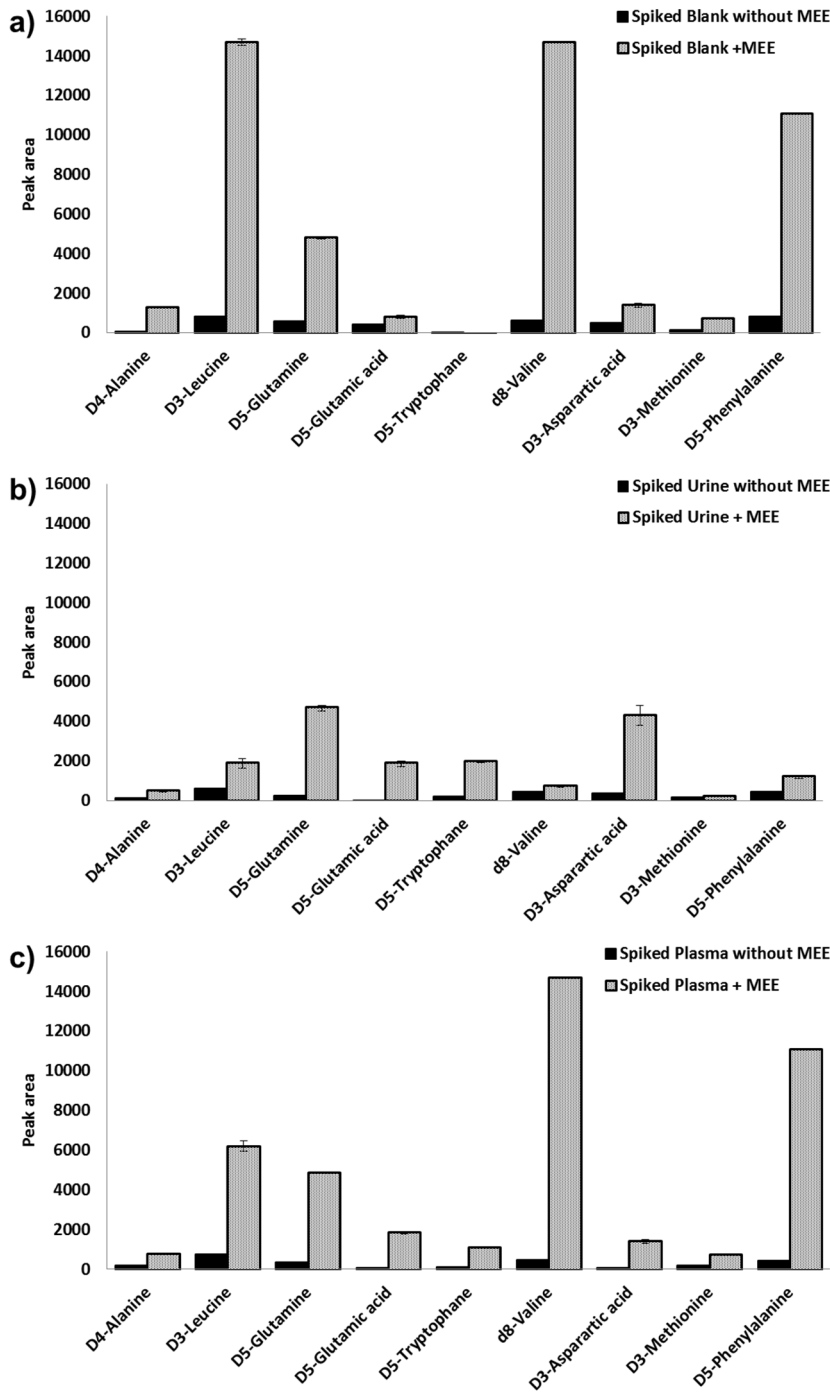


Figure 46: Comparison of stable isotope labeled amino acids peak areas in a) blank and b) and c) biofluids with and without post column infusion of 2-MEE.

Comparison of the obtained peak areas for the respective amino acids reveals, that in the urine sample the lowest peak areas were detected. This result was already expected to a

certain extent. The HILIC stationary phase retains polar compounds, which are highly abundant in urine, much better than lipophilic compounds which are commonly found in plasma and elute in the void volume of the chromatogram. Nevertheless, a dopant effect of 2-MEE was observed despite the presence of ESI suppression caused by sample matrix effect (see Figure 47).

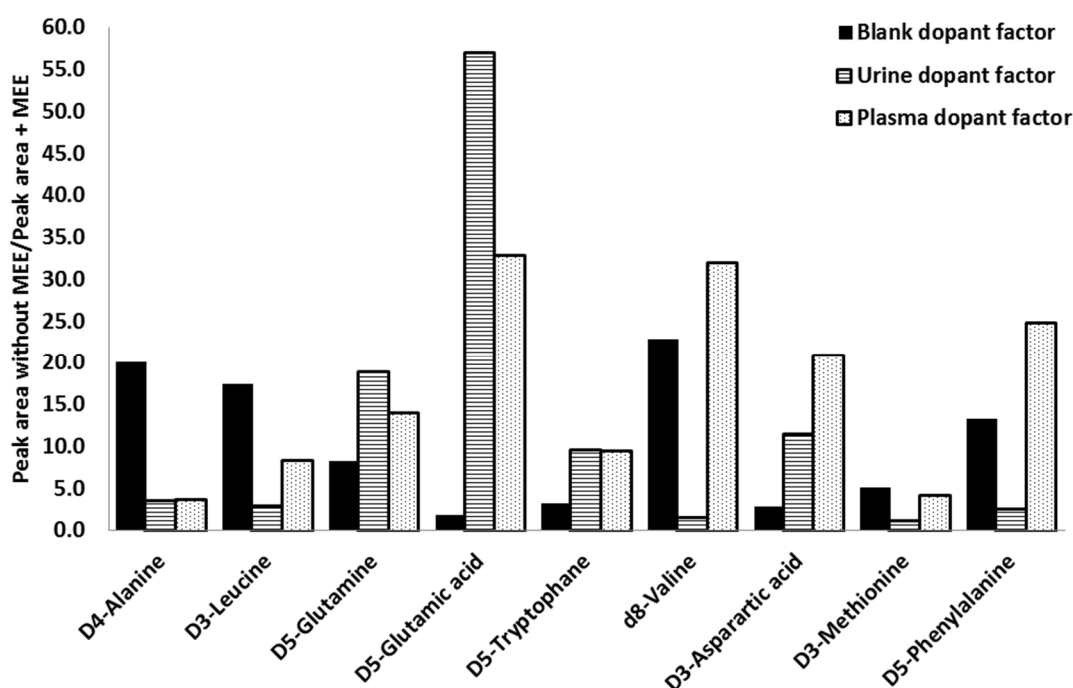


Figure 47: Dopant factor in spiked blank, urine and plasma showing differences in the matrix effect but systematic increase in sensitivity

Removal of 2-MEE and optimization of 2-MEE proportion in post-column infusion

Despite the positive effects of 2-MEE in NEG-mode ESI, the operator has to be aware that 2-MEE is not available in LC-MS grade quality. During application of 2-MEE additional mass spectrometric peaks were observed (data not shown) which were considered as 2-MEE impurities. In ESI-NEG mode these peaks did not have any further effects whereas in ESI-POS mode these impurities were causing severe ESI suppression effects (data not shown).

This effect is probably based on the PEG-like nature of the 2-MEE impurities. For this reason, a cleaning procedure was established in order to restore initial ESI performance.

Peek tubing and ESI capillary need to be flushed with water:acetonitrile (1:1 v/v) for about 20 minutes and ESI chamber needs to be cleaned according to vendors specification (tested with Bruker and Waters mass spectrometers, data not shown). Even after several days of 2-MEE infusion no persistent negative side effects could be observed when following these cleaning procedures.

Even though, application of the cleaning procedure works perfect it was tested how the system reacts if less 2-MEE is applied. A mixed mode HPLC method was used to separate butyric and valeric acid.

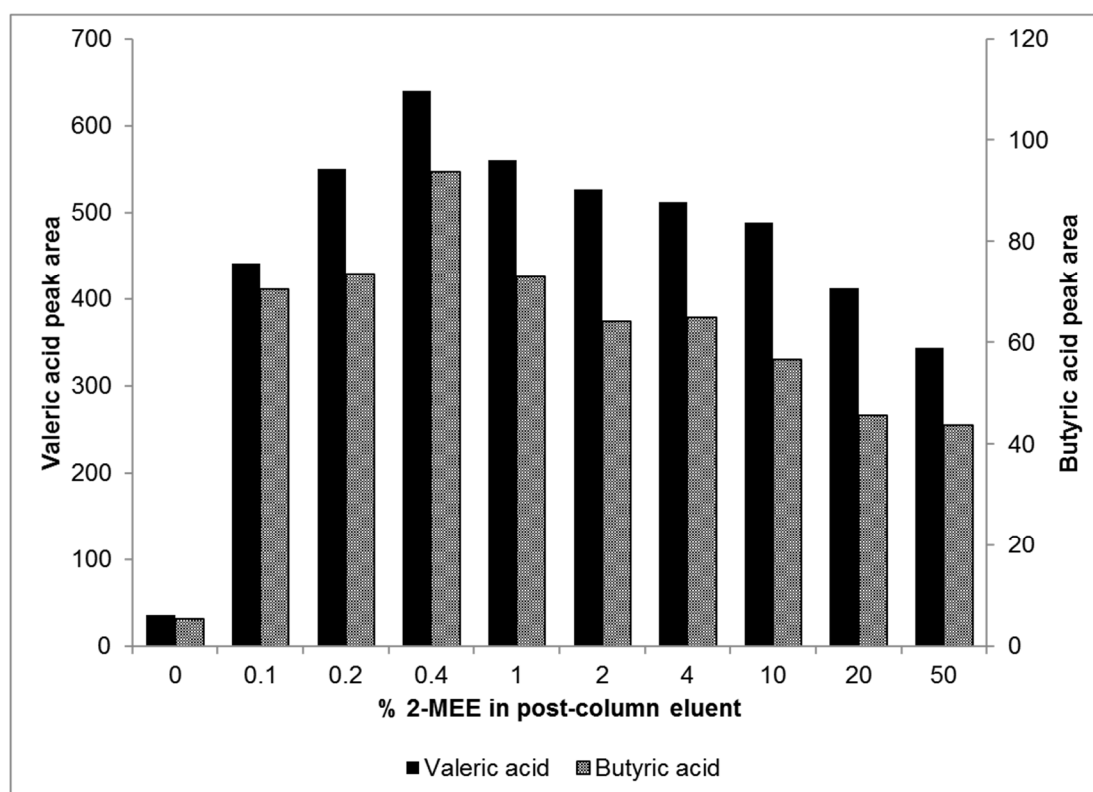


Figure 48: Obtained peak areas of short chain fatty acids with different concentrations of post-column 2-MEE infusion

These Short chain fatty acids usually are known to exhibit a poor ionization efficiency in electrospray ionization, making them an ideal marker for 2-MEE induced ESI dopant factor. Butyric and valeric acid were injected at a constant concentration whereas 2-MEE

concentration was varied. The experiment revealed an optimal 2-MEE concentration of 0.4% 2-MEE in the post-column mobile phase see Figure 48. These result could also be confirmed with the HILIC separation method (data not shown). This 0.4% 2-MEE is a much lower concentration of 2-MEE than the initially suggested 30% 2-MEE by Yamaguchi (105) and offers three advantages:

1. Consumption of 2-MEE is much lower
2. Dopant effect of 2-MEE is improved
3. Baseline noise during 2-MEE application is considerably reduced.

4 Effect of insulin sensitivity on urinary metabolome

Since the original idea of testing whether a nutritional intervention causes changes in the metabolome of the subjects failed to show significant results, another hypothesis was tested. The subjects of the NUPREDM study showed a wide range of different insulin sensitivities measured according to the method described in by Matsuda and DeFronzo (114). Already in 2010 Lucio et. Al. (115) showed that insulin sensitivity is reflected by characteristic fingerprints in the plasma derived metabolome. In conclusion, it appeared reasonable to search for characteristic fingerprints in the urinary metabolome that reflect the insulin sensitivity of the subject. To avoid multiple sampling of the same person and interferences due to the nutritional intervention, which both could bias the outcome, only the samples collected at the baseline time point were used for this approach.

4.1 Analysis of NUPREDM baseline cohort

Considering the method optimization described in chapter 3.1 and 3.2 urine dilution normalization and ESI dopant 2-MEE, the baseline samples of the NUPREDM study introduced in 2.1.1 were reanalyzed with the HILIC-MS method. For the FT-ICR-MS measurements sample preparation was repeated with the dilution normalized urine samples and were reanalyzed with the method described in chapter 2.2.2.2.

4.1.1 Data handling

4.1.1.1 FT-ICR-MS

Peak picking, mass calibration, matrix generation and mass annotation were performed as described in chapter 2.2.2.3. In the final filtering step, peaks were filtered for 30% frequency instead of 10% as described in 2.2.2.3. This way a matrix containing 5712 mass peaks was obtained.

For further statistical processing the dataset was unit variance scaled. Samples were grouped into three ISI-Matsuda levels: low (ISI-Matsuda <8.5 ; 50 subjects), intermediate ($8.5 \leq$ ISI-Matsuda <15 ; 38 subjects) and high (ISI-Matsuda ≥ 15 ; 21 subjects). Based on the groups high and low, a RELIEF (Package FSelector, Rstudio Inc., Version 0.99.896) algorithm was used by Dr. Marianna Lucio to select features which are discriminative for the comparison (116). This technique identified ultimately 3811 discriminative features. Finally, OPLS-discriminant analysis was performed by Dr. Mariann Lucio on the ISI-Matsuda classes high and low of the study subjects using SIMCA-P 13.0. Data were validated with a 7-fold one leave out cross-validation.

4.1.1.2 HILIC-MS

Chromatogram alignment and peak picking for non-targeted HILIC-U(H)PLC-QTOF analysis was performed with help of Genedata Expressionist MS Refiner 8.2 and Analyst 8.2 (Genedata, Munich, Germany) (see Appendix). Mass annotation of detected peaks was performed with help of Genedata in combination with HMDB Human Metabolite Database (HMDB) version 3.5 (29-31) with a mass annotation error of ± 10 ppm.

The resulting data was normalized via application of the batch normalizer protocol published by Wang et. Al. 2003 (117). This method removes batch and injection order drift effects and is integrated as a function in Genedata Analyst 8.2. The positive effect of the batch normalizer was demonstrated with help of an unsupervised principal component analysis. Without batch normalization, the first three principal components mainly describe a batch effect (see Figure 49 A) also the QC's are not centered in the 2D principal component plot (blue dots in Figure 49 B).

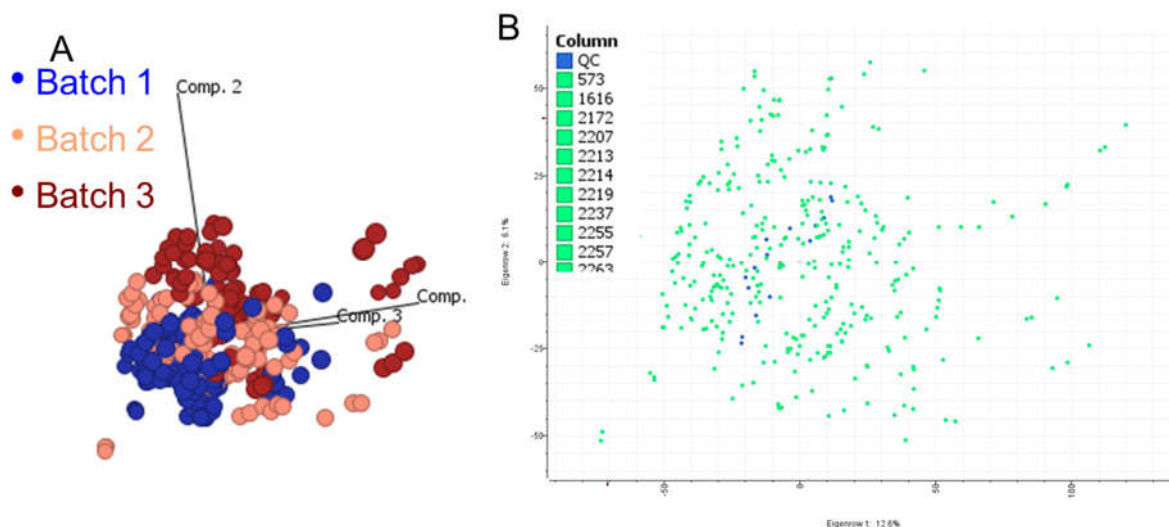


Figure 49: A) 3D and B) 2D plot of PCA results created from NUPREDM baseline HILIC-MS data without batch normalization

If batch normalization was applied the first three principal components were not dominated by the batch effects any more (see Figure 50 A) and the QC's were centered in the 2D principal component plot (see Figure 50 B) as they are supposed to be.

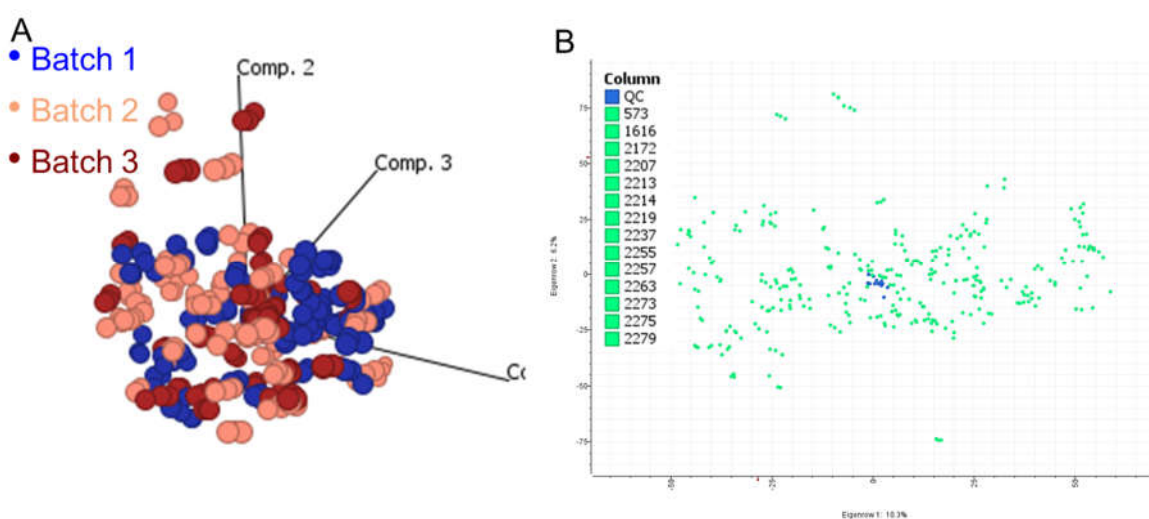


Figure 50: A) 3D and B) 2D plot of PCA results created from NUPREDM baseline HILIC-MS data with batch normalization

The resulting matrix was further processed and filtered with help of Microsoft Office Excel 2010. Peaks were discarded if the relative standard deviation of the peak areas within triplicates exceeded 30%. The triplicates were then averaged using the geometric mean. The

resulting matrix was filtered for peaks that appear in less than 30% of the samples. This way a matrix of 8255 peaks was created.

For further statistical processing the dataset was unit variance scaled. Samples were grouped into three ISI-Matsuda levels: low (ISI-Matsuda < 8.5 , 50 subjects), intermediate ($8.5 \leq$ ISI-Matsuda < 15 , 38 subjects) and high (ISI-Matsuda ≥ 15 , 21 subjects). Based on the groups high, intermediate and low, a RELIEF (Package FSelector, Rstudio Inc., Version 0.99.896) algorithm was used by Dr. Marianna Lucio to select features which are discriminative for the comparison (116). This technique identified ultimately 4930 discriminative features. Finally, OPLS analysis was performed by Dr. Mariann Lucio on the ISI-Matsuda values of the study subjects using SIMCA-P 13.0. Data were validated with a 7-fold one leave out cross-validation.

4.1.2 Results and discussion

4.1.2.1 DIA-FT-ICR-MS OPLS-DA analysis

The dataset had to be reduced to the ISI-Matsuda classes high and low, otherwise no valid OPLS-DA model could be established. The cumulated R^2 of the model was 0.7 the cumulated Q^2 was 0.45 (see Figure 53).

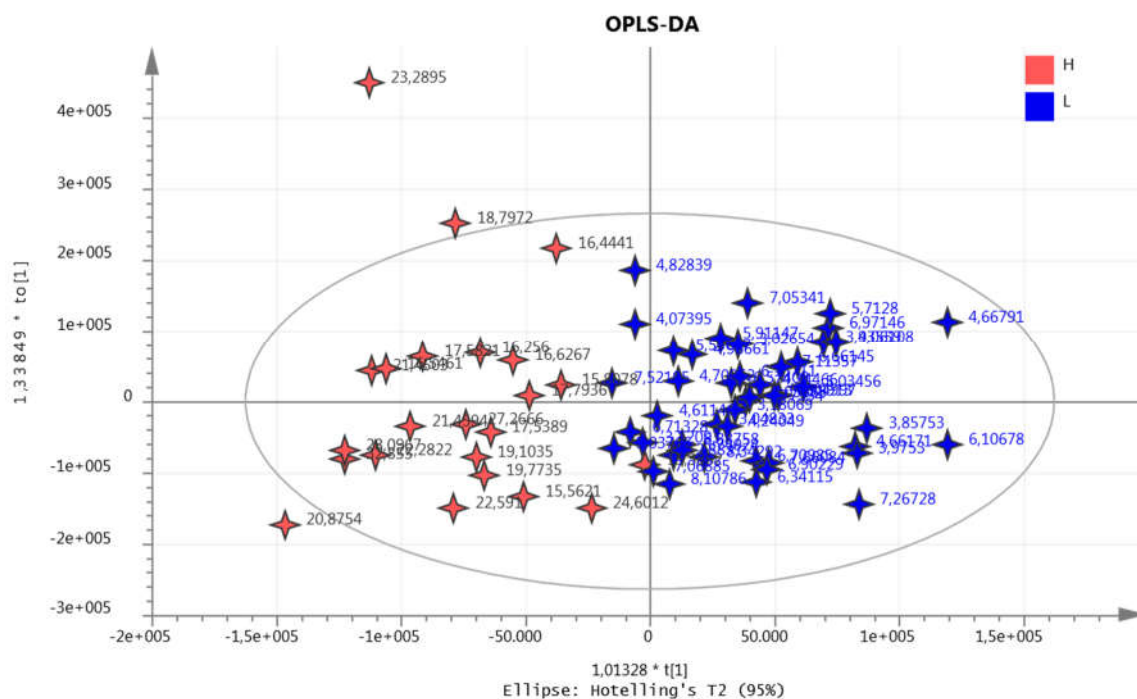


Figure 51: Score plot of OPLS model from DIA-FT-ICR-MS data for ISI-Matsuda colored by classes high (red) and low (blue)

Only the limitation to the extreme ends of the data-set (high and low ISI-Matsuda) allowed the creation of a valid OPLS-DA model. This shows that the urinary metabolome of the intermediate subjects did not allow an unambiguous group allocation.

A pathway analysis using Metaboanalyst did not result in significant results, which is probably also due to the low number of features (147) with a VIP-Scores > 1. From these 147 features, only 94 could be annotated with MassTrix4. After filtering for multiple entries of the same metabolite due to formation of adducts like $[2M-H]^-$ or $[M+FA]^-$, 28 metabolites could be identified with help of educated guess. Of these 28 metabolites 16 could be integrated into a (pre-) diabetic context (see Table 9).

Table 9: Urinary metabolites reflecting ISI-matsuda changes identified by OPLS-DA

HMDB ID	HMDB name	ISI-Matsuda correlation	Metabolite Classification
HMDB11686	p-Cresol glucuronide	Positive	
HMDB00714	Hippuric acid	Negative	
HMDB60015	Phenol sulfate	Negative	
HMDB13189	3-Indole carboxylic acid glucuronide	Negative	Gut microbiome metabolism of aromatic amino acids
HMDB00671	Indolelactic acid	Positive	
HMDB60001	Indole-3-acetic-acid-O-glucuronide	Negative	
HMDB13240	Indoleacetyl glutamine	Negative	
HMDB00929	Tryptophane	Negative	
HMDB11171	L-gamma-glutamyl-L-leucine	Negative	Peptides
HMDB02171	Glycylprolylhydroxyproline	Negative	
HMDB00350	3-Hydroxysebacic acid	Negative	
HMDB00953	Suberylglycine	Negative	Energy metabolism
HMDB00210	Pantothenic acid	Negative	
HMDB00386	3b,16a-Dihydroxyandrostene sulfate	Negative	
HMDB10357	Tetrahydroaldosterone-3-glucuronide	Negative	Cortisol metabolism
HMDB10320	Cortolone-3-glucuronide	Negative	

In a set of box-plots created from selected metabolites, one exemplary metabolite for each metabolite class, the separation between ISI-Matsuda high and low group was visualized (see Figure 52).

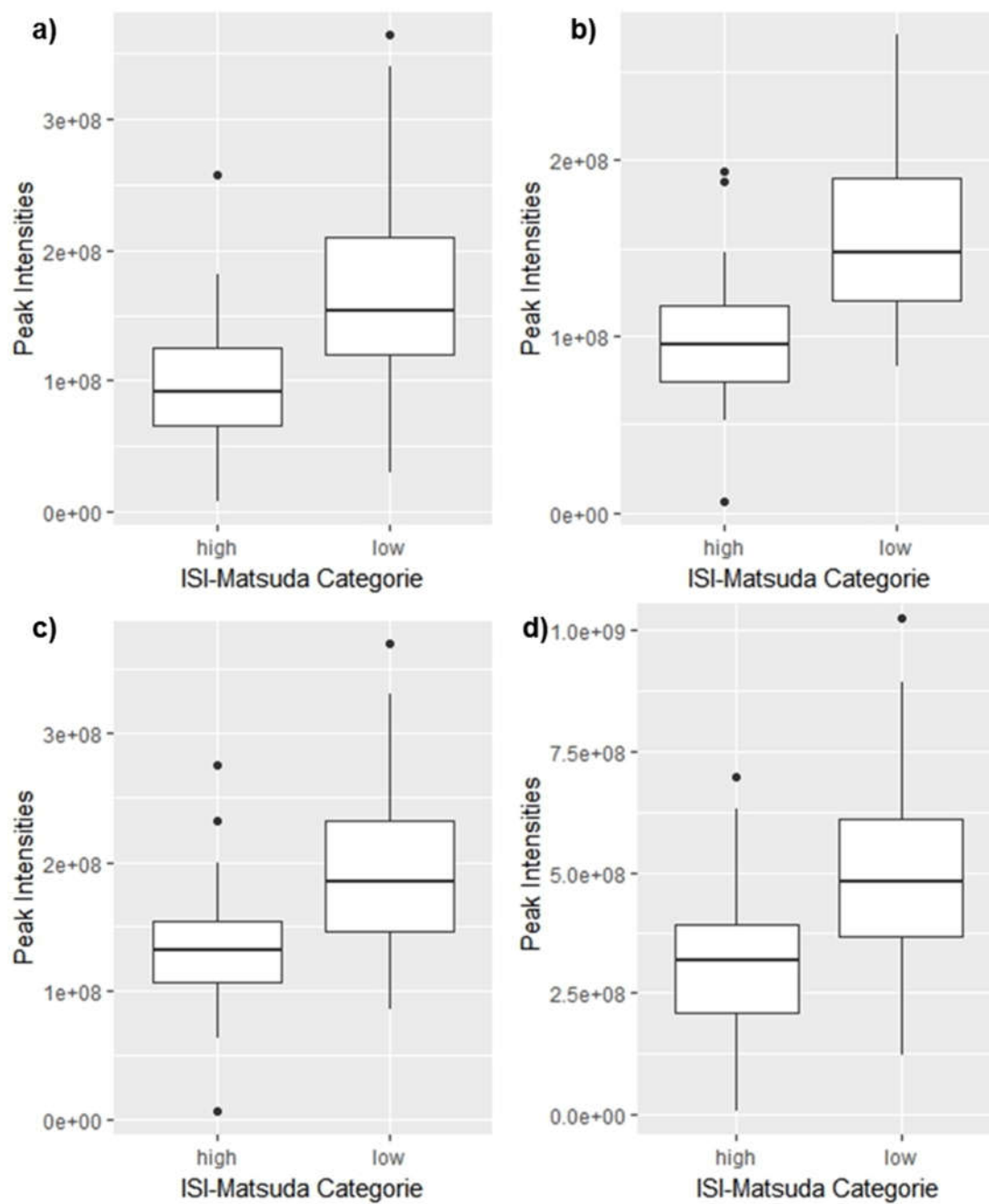


Figure 52: Whisker-Box-plots for selected exemplary metabolites discriminant for low and high insulin sensitive subjects a) tryptophan, b) glycylglycylhydroxyproline, c) suberylglycine and d) cortolone-3-glucoronide

4.1.2.2 HILIC-MS

OPLS-DA analysis

A valid OPLS model could be established which described the ISI-Matsuda levels based on the urinary metabolome in the analyzed samples. The cumulated R^2 of the model was 0.8 the cumulated Q^2 was 0.3 (see Figure 53).

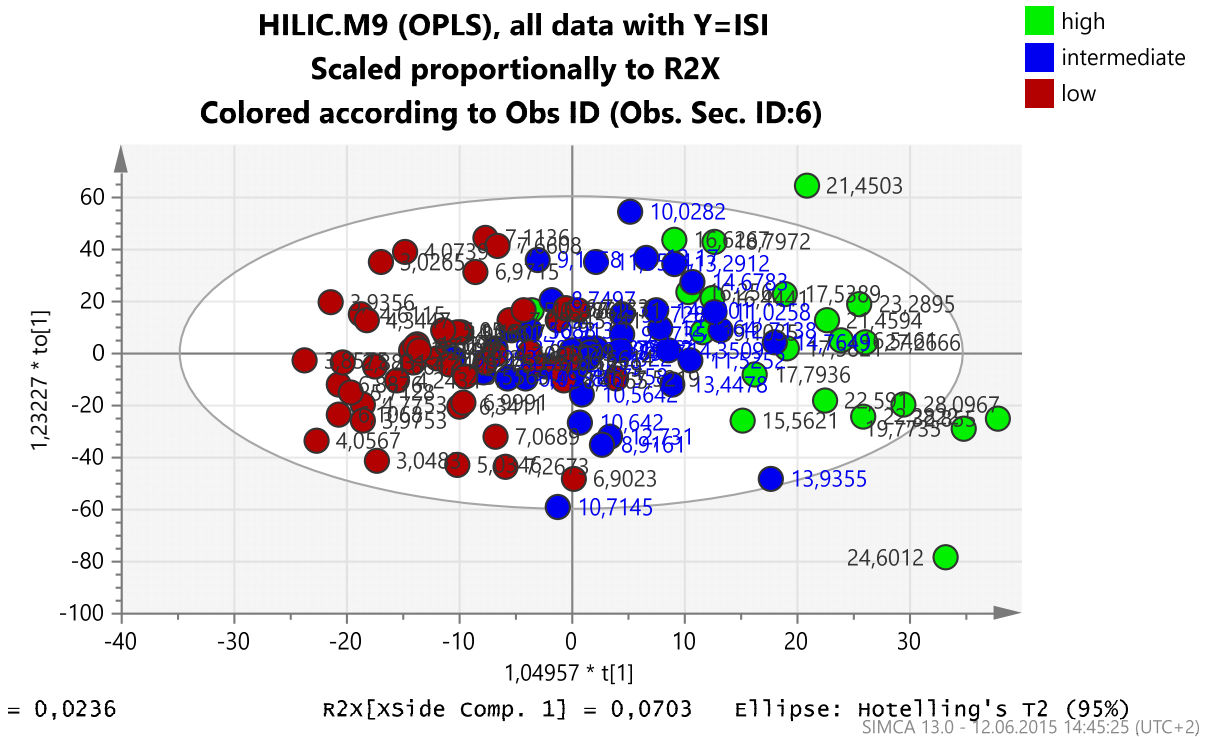


Figure 53: Score plot of OPLS model from HILIC-MS data for ISI-Matsuda colored by classes high, intermediate and low

Evaluation of the 446 peaks showing a VIP-score > 2 showed that the majority, 72.4% (323 peaks) of the peaks, were discriminative for low ISI-Matsuda values. Whereas, only 27.6% (123 peaks) of the peaks were discriminative for high ISI-Matsuda values. From these 323 peaks 77.4% (250 peaks) could be assigned with mass annotation, whereas from the residual 123 peaks only 64.2% (79 peaks) could be annotated. From these results, it was concluded that the urinary metabolome from healthy patients is very heterogeneous and only little can be concluded from it, or in other words it is more difficult to identify biomarkers for healthy insulin sensitive persons. The urinary metabolome of insulin resistant persons in comparison is more informative.

The HMDB ID's of all peaks with a VIP-score > 1.5 (see Supplementary Table 12) were uploaded to Metaboanalyst to perform a pathway analysis.

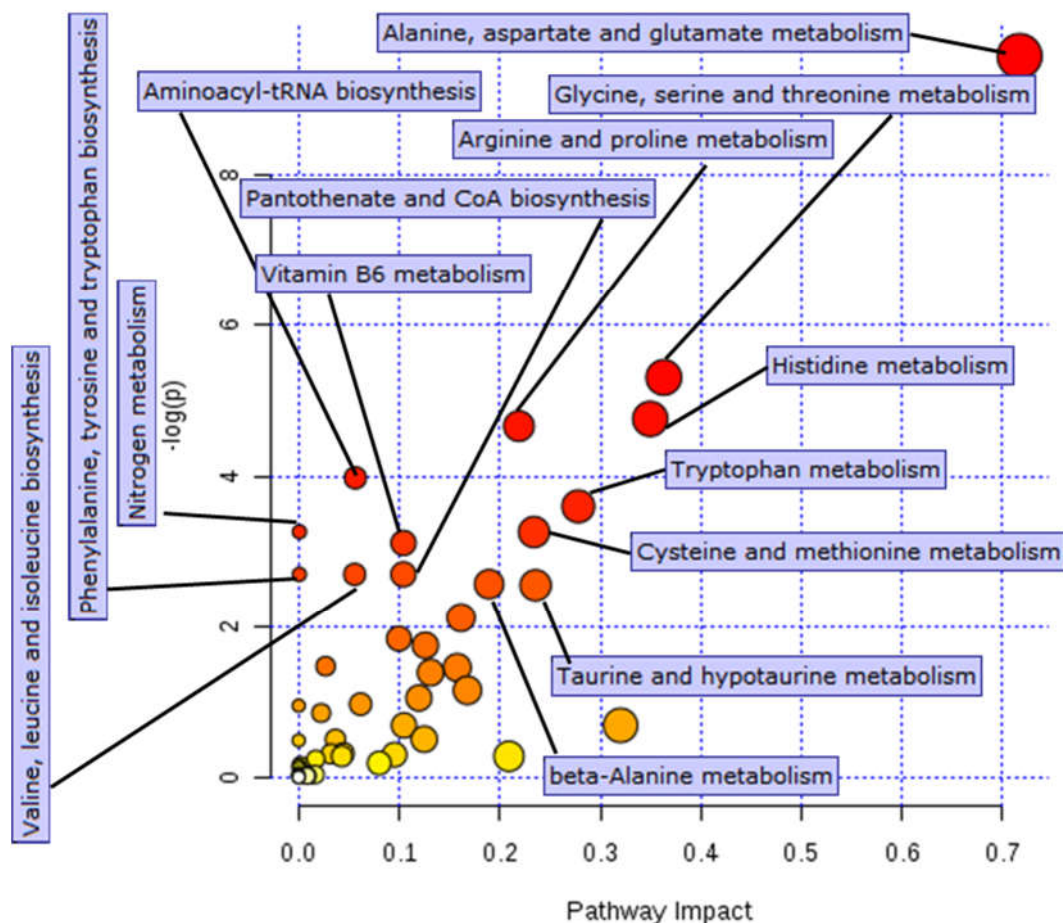


Figure 54: Pathway analysis of significantly altered metabolites from OPLS-DA analysis

In the HILIC-MS metabolomic dataset an overlap of results with the ICR-FT-MS results was observed, this intersection between HILIC and ICR-FT-MS approach was expected. In the following metabolite classes and pathways with similar results are listed

- Gut microbiome and metabolism of aromatic amino acids,
- Vitamine B5 metabolism,
- Corticosteroid hormone metabolism (non significant in HILI-MS pathway analysis) and

- Peptides/hydroxyprolins and Collagen metabolism (non significant in HILIC-MS pathway analysis).

The intersecting results of FT-ICR-MS and HILIC-MS approach show metabolite classes and pathways which were not rated significant in the pathway analysis resulting from the HILIC-MS data. An absence in the pathway analysis or a non-significant result does not necessarily mean that the observed metabolites are not relevant for the ISI-Matsuda prediction or group differentiation. For this reason, additional noteworthy metabolites and pathways besides the ones identified by pathway analysis (like the sialic acids and lipid degradants) will be discussed in the next chapter.

4.1.2.3 Pathways and metabolites differentiating for ISI-Matsuda

Gut microbiome and metabolism of aromatic amino acids

Indoles are a product of gut microbiome tryptophan metabolism (118, 119), phenols and cresols are a result of gut microbiome tyrosine metabolism (118) and hippuric acid originates from gut microbiome phenylalanine metabolism (120).

The observed changes in the urinary concentrations of these metabolites are putatively attributed to an altered gut microbiome composition – a result of a high fat and carbohydrate diet also referred to as “western diet” (121).

Hydrophilic metabolites of p-Cresol e.g. glucuronides or sulfates were found to be increased in Male Wistar rats after Roux-e-n Y gastric bypass (122). This observation was made concomitantly to other beneficial effects like weight loss and an increase in gut Proteobacteria and a reduction in Firmicutes (including the Clostridial family) and Bacteroides. Urinary cresols were therefore assumed to be an indicator for a beneficial gut microbiome composition. Furthermore, Lustgarten et. Al. 2014 observed a positive correlation of urinary cresols and indoles with physical constitution in functionally-limited older adults.

In contrary, urinary phenols, a product of gut microbiome phenylalanine metabolism, were negatively correlated to physical conditions (123). It was assumed that the root cause was a higher ratio of *Clostridium difficile* and *sporogenes* in the gut microbiome.

Also, Dong et al. observed in 2017 a negative correlation of indoxyls, indoles and caffeic acids to non-alcoholic fatty liver disease (co-morbidity of metabolic syndrome) (124)

But, also contradictory results have been published, e.g. a study conducted in 2016 (119) reported elevated urinary levels of cresols and indoles after obesity and decreased glucose tolerance was induced in rats with a high fat diet.

In summary, the highly complex host-gut microbiome interactions are not yet fully understood, the inconsistent literature data just underlined this fact.

Peptides and Hydroxyprolins (Collagen metabolism)

Diabetes is connected to a higher risk to develop osteoporosis, the mechanisms behind are not fully understood, but a perturbed collagen metabolism is assumed to be a root cause (125). Hydroxyproline conjugates like the glycylylprolylhydroxyproline observed in this study are also observed in i) chronic kidney disease – mineral and bone disorder patients (CKD-MBD) (126) and ii) diabetes mellitus and diabetic nephropathy (127) patients. Increased hydroxyproline levels are also connected to alterations in the parathormone balance

L-gamma-glutamyl-L-leucine was mentioned in a patent for biomarkers related to insulin resistance (128), without further explanations. Yet, increased serum gamma-glutamyl transferase concentrations, which might be responsible for increased gamma-glutamyl-leucine, are connected to diabetes (129) and other comorbidities of diabetes such as increased cardiovascular risk or hepatic inflammation (130).

Taurine metabolism

It could be demonstrated that, in animal models, taurine supplementation has a positive effect on insulin sensitivity (131, 132). Different mechanisms were suggested how taurine affects:

- Antioxidative effect (133)
- Modulation of diabetes-related genes (134)
- Improved Ca^{2+} handling of pancreatic beta-cells (135)
- Increased nitric oxide production (136)

Though, no positive effects on insulin sensitivity in humans could be observed after oral supplementation of taurine (137).

Energy metabolism β -oxidation

3-Hydroxysebacic acid and suberylglycine were correlated to medium chain acyl-CoA dehydrogenase deficiency (MCAD) (138, 139). 3-Hydroxysebacic acid is also associated with oxidation Acyl-CoA dehydrogenase 9 (ACAD9) deficiency and was found in urine of fasting children with glycogen storage disease V and is connected to peroxisomal diseases like Zellweger syndrome or general impaired β -oxidation in mitochondria (140, 141)

Corticosteroid hormone metabolism

The corticosteroid hormone metabolism is closely connected to bone and skeletal physiology (142). The altered levels of corticosteroids observed in the present study aligns well with the observation of altered glycyprolylhydroxyproline, both findings point to an impaired bone growth and remodeling (osteoporosis) as a comorbidity of diabetes. Urinary androsteron glucuronide and cortisol sulfate levels were detected to be higher in CKD-MBD patients (126).

Cortolone-3-glucuronide and tetrahydroaldosteron-3-glucuronide were identified as markers for metabolic syndrome in 2014 by Zhi-rui Yu et. al. (143). Furthermore, increased urinary levels of Cortolone-3-glucuronide and tetrahydroaldosteron-3-glucuronide were observed after ingestion of citrus juice and connected to bone remodeling processes and over activation

of steroid biosynthesis pathway (144). Tetrahydroaldosteron-3-glucuronide was also found to decrease as a result of weight loss and reduction in blood pressure in 1985 by Marks et. Al. (145).

Urinary concentrations of 3b,16a-Dihydroxyandrostenone sulfate were found to be reduced after ingestion of apple-products (146). 3b,16a-Dihydroxyandrostenone sulfate was also associated to an increased expression of cytochrome P450, persistent hepatotoxicity, bile acid homeostasis dysregulation and oxidative stress due to dioxin exposure (147). And is also increased in patients suffering from unstable angina (148)

In general, altered corticosteroid levels were also observed in NAFLD (124), Type 1 Diabetes (149), metabolic syndrome (150) and insulin resistance (151) . Additionally, Metabolic syndrome may also increase cortisol production caused by 11b-hydroxysteroid dehydrogenase 1 (11b-HSD1) in liver and adipose tissue and by HPA axis dysregulation (151). Corticosteroids may also play a key role in the Type II diabetes downward spiral by inducing hepatic glucose production and decreased glucose uptake into peripheral tissues (152).

Amino acids

The alanine, aspartate and glutamate metabolism was found to have the highest pathway impact in the pathway analysis, which matches the observations and hypothesis made 2009 by Newgard et al. (153). They hypothesized that overnutrition leads to high branched chain amino acids (BCAA) plasma levels which in turn results in increased glutamate concentrations. Accumulation of glutamate may increase transamination of pyruvate to alanine which is a highly gluconeogenic amino acid. Cheng et al. 2012 as Walfrod et al. 2016, observed a strong association of insulin resistance or an impaired glucose tolerance with the glutamine-to-glutamate ratio. (154, 155)

Sookoian also linked a modulated glutamate metabolism to aminotransferase reactions and suggested to consider liver transaminases as a central player of NAFLD and MetS in general since they play a key role in conversion of glutamine-to-glutamate. (156)

Coenzyme A synthesis and B-Vitamins (B5 and B6)

Increased concentrations of pantothenic acid were observed in patients with low ISI-Matsuda. Pantothenic acid is a precursor of coenzyme A (CoA) and therefore a ubiquitously present metabolite involved in the central energy metabolism of any cell. To find increased levels of one of the key building blocks of CoA appears plausible considering that indications of a perturbed CoA and β -oxidation metabolism (3-Hydroxysebacic acid and suberylglycine) were observed.

CoA synthesis is much better known than its degradation pathways (157). CoA degradation might also be inhibited by diabetes or CoA degradation products might be shuttled back into the biosynthetic pathway. (158) Thus, it remains unclear whether the observed increased urinary pantothenic acid levels are a

- result of CoA breakdown due to a perturbed β -oxidation (acyl-CoA dehydrogenase deficiency) or
- consequence of excess pantothenic acid due to impaired CoA biosynthesis which is cleared renally.

An altered Coa homeostasis is also known from diverse diseases like medium chain acyl Coa dehydrogenase deficiency, starvation, diabetes, hypertension and vitamin b12 deficiency (159). Increased levels of insulin, glucose, fatty acids, pyruvate and ketone bodies inhibit CoA biosynthesis, glucocorticoids and glucagon increase CoA concentration in tissues (159).

Vitamin B6 is known as a co-factor catalyzing many reactions like transamination, racemization, decarboxylation, glycogen breakdown and α,β -elimination reactions and is therefore important for a well-functioning lipid metabolism. (160) Interestingly, Gregory et al. 2013 demonstrated that a vitamin B6 restriction increases glutamine/glutamate ratio. (161) As discussed in the preceding chapter, a high glutamine-to-glutamate ratio is positively connotated in the context of diabetes. On the other side, Vitamin B6 protects from oxidative stress and irreversible protein glycation. (162) Vitamin B6 also holds an anti inflammatory

effect by modulating the glucocorticoid and steroid receptor mediated signal transduction pathways, in turn vitamin B6 deficiency leads to more pronounced sensitivity to glucocorticoids. (163, 164)

Sialic acids

A number of sialic acids and deaminated neuraminic acids (and probably isomers of deaminated neuraminic acid) were found to be highly discriminative between Insulin sensitive and resistant patients. Since these metabolites do not appear in the pathway analysis they are summarized in Table 10.

Sialic acids were already 1981 identified as an indicator for rheumatoid arthritis by Maury et. Al. (165) Serum sialic acids were also reported as a risk factor for cardiovascular disease, a biomarker for acute phase response and is elevated in type 1 and 2 diabetes mellitus. (166) Serum Sialic acids concentrations raise with age and in diabetes mellitus and was used as biomarker for benchmarking the antidiabetic drugs rosiglitazone and metformin. (167) Also Gopinathan Natir et. Al report in 2017 that type 2 diabetes mellitus subjects had a higher total sialic acid level compared to controls, and sialic acid correlated significantly to fasting blood sugar and cholesterol, triglycerides, systolic and diastolic blood pressure. (168)

Table 10: Sialic acids identified as significant in patients with low ISI-Matsuda

Peak ID	m/z	Ret. Time	HMDB ID	HMDB Name	VIP	ISI-Matsuda correlation
Peak_09889	267.073	7.00	HMDB00425	3-Deoxy-D-glycero-D-galacto-2-nonulosonic acid	3.60	Negative
Peak_09888	267.073	6.69	HMDB00425	3-Deoxy-D-glycero-D-galacto-2-nonulosonic acid	3.47	Negative
Peak_19636	630.187	7.51	HMDB06626	Sialyl ketolactic acid	3.38	Negative
Peak_09912	267.097	7.08	HMDB28848	3-Deoxy-D-glycero-D-galacto-2-nonulosonic acid	3.367	Negative
Peak_19983	673.23	8.64	HMDB06581	3-Sialyl-N-acetyllactosamine	2.45	Negative
Peak_19649	632.199	7.49	HMDB00825	3'-Sialyllactose	2.26	Negative
Peak_19651	632.202	8.13	HMDB06569	6'-Sialyllactose	2.09	Negative

Lipid degradants

In the urine of insulin resistant patients elevated levels of glycerophosphoinositol (GroPIns) and glycerophosphoglycerol were observed, both metabolites represent the polar head group of certain lipid species i.e. phosphatidylinositols and cardiolipins or phosphatidylglycerols respectively.

Phosphatidylinositols like phosphatidylinositol-3,4,5-Triphosphate (PtdIns(3,4,5)P3) are known to be important messenger lipids and an impaired PtdIns(3,4,5)P3 signaling is associated with metabolic diseases like MetS and type 2 diabetes. (169) Glycerophosphoinositol (GroPIns) is formed through two sequential deacylation reactions catalyzed by phospholipase A2 (PLA2). (169, 170) An increased muscle PLA2 activity was observed in diabetic rats (171) and an elevated PLA2 secretion was linked to an increased risk of early atherosclerosis in patients with metabolic syndrome. (172) In a prospective study the PLA2 activity was found to be positively associated with insulin resistance and the risk to develop type 2 diabetes mellitus was significantly higher for patients showing a high PLA2 activity. (173) To the best of the authors knowledge it is the first time, that urinary glycerophosphoinositol is connected to a diabetic or pre-diabetic phenotype.

Cardiolipin is a central building block of mitochondrial membranes. It has been demonstrated that in type 2 diabetes impaired mitochondrial function and β -cell failure go hand in hand. (174) Degradation of mitochondrial membranes due to increased oxidative stress, (175) caused by type 2 diabetes may induce a cardiolipin breakdown which is ultimately results in increased glycerophospholglycerol levels in the urine of diabetic and pre-diabetic patients.

5 Concluding remarks

Analysis of blood plasma and urine samples from the lifestyle intervention study NUPREDM by non-targeted LC-MS, DIA-FT-ICR-MS and NMR did, unfortunately, not result in a significant measurable change in the patient's metabolism. One of the targets of the NUPREDM study was to setup an intervention scenario that can be maintained realistically by a majority of the target patient group. It was concluded that these relatively mild intervention did not induce an effect on metabolism big enough to be distinguished from the huge variability from individual to individual. Yet, the study can be considered as success since a weight loss and a gain in insulin sensitivity in average was observed.

During analysis of the NUPREDM samples it became abundantly clear, that metabolomic screening of urine samples is a challenging task. The fact that urine is a biological waste material, not controlled by homeostasis, culminates in an inherent concentration variation of urine samples and an extreme chemical variability dependant on individual nutrition habits, intake of medication and exposure to exogenous compounds. For these reasons it is tricky to determine metabolites that discriminate between healthy and diseased patients. Traditionally, obtained signal intensities (or concentration values) of metabolites from urine analysis are normalized via creatinine normalization, yet in non-targeted metabolomic approaches the applicability of creatinine normalization reaches its limits. To successfully apply creatinine normalization to a urinary metabolite certain prerequisites exist:

- the linear range,
- limit of quantification and
- onset of detector saturation

of the method need to be known in forehand. Without this prior knowledge, as is usually the case in non-targeted metabolomics, creatinine normalization may create misleading result.

To illustrate the issue two extreme possible scenarios i) high concentration urine sample and ii) low concentration urine sample will be discussed. In a very high concentrated urine sample a hypothetical metabolite X induces detector saturation. The determined metabolite concentration or signal intensity does not reflect the actual metabolite concentration. Creatinine normalization of these figures will result in values that are primarily subordinated by the creatinine concentration. Thus, conclusions or biological interpretation based on the normalized metabolite concentration or signal intensity would be distorted. In a very dilute urine sample a hypothetical metabolite Y is very close or even below the limit of quantification. The obtained signal intensity or determined concentration would be either not available or mainly noise driven. Creatinine normalization would just multiply the noise and not compensate for this issue. In addition to that, normalization to only one parameter, namely creatinine, is likely to bias the results. Creatinine concentration can also vary or be biased, independent from overall urine concentration due to reasons like disease, physical exercise or an outlier due to a measurement error.

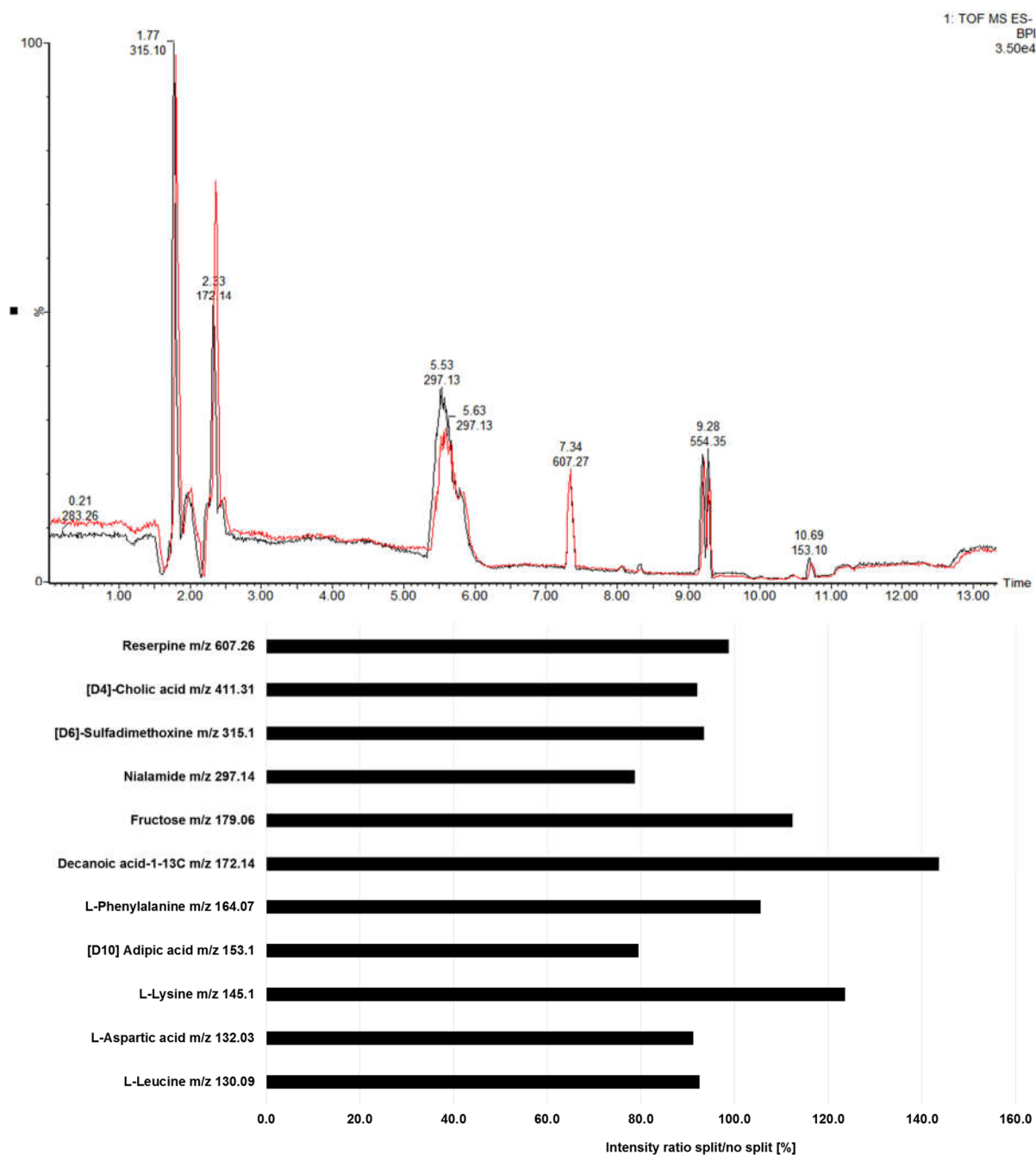
A better way to handle urinary concentration variability, with regard to metabolomics screening approaches, is to normalize the urinary concentration via dilution normalization before analysis. Therefore a pre-acquisition normalization strategy was developed. In addition, the reliability of the normalization was improved by setting up a poly parametric urine normalization factor based on creatinine, urea, specific gravity, and UV absorbance.

Pre acquisition normalization of the urine samples resulted in generally lower concentrated urine samples. This may have the effect that metabolites which are slightly above the detection limit in the original sample fall below the detection limit due to dilution. To a certain degree, this circumstance could be compensated via post-column infusion of the ESI-dopant 2-(2-methoxyethoxy)ethanol (2-MEE). Application of 2-MEE proofed to significantly increase MS sensitivity of small and polar metabolites.

With the objective to further evaluate and make more use out of the samples from the NUPREDM study, the urine baseline samples (before intervention) were reanalyzed. Combination of pre acquisition normalization and sensitivity enhancement via 2-MEE allowed to find discriminant urinary metabolite patterns for insulin sensitive and insulin resistant patients, analog to Lucio et al 2010 in plasma samples (115). The obtained markers were reflecting the onset type II diabetes showing a perturbed gut microbiome metabolism and alteration in the energy metabolism especially with regard to β -oxidation. Also changes in the corticoid steroid metabolism were observed, likely a response to a generally elevated level of inflammation in obese and pre-diabetic patients.

Additionally, urinary glycerophosphoinositol (GroPIs) and glycerophosphoglycerol were observed to be predictive markers to discriminate insulin sensitive from insulin resistant individuals. Both metabolites represent the polar head group of certain lipids i.e. phosphatidylinositols for GroPIs and cardiolipins for glycerophosphoglycerol. Cardiolipins for example are an integral part of mitochondrial cell membranes, misdirected mitochondrial metabolism in pre-diabetic patients and resulting oxidative stress were assumed to induce cardiolipin breakdown. The remnants are subsequently renally cleared and are therefore present at higher levels in the urine of insulin resistant individuals. Increased urinary GroPIs concentration was assumed to be a result of an increased PLA2 activity, the enzyme activity is linked to metabolic syndrome and diabetes and hydrolyses phosphatidylinositols to GrPIs. Both metabolites ((GroPIs) and glycerophosphoglycerol) were, to the best of the author's knowledge, reported for the first time in the context of diabetic or pre-diabetic phenotypes.

6 Appendix



Supplementary Figure 1: Comparison of HILIC Standards with and without flow-splitting

Supplementary Table 1: FT-ICR-MS method parameters

Parameter	ESI -
Number of scans	300
Transient time domain [MW]	4
Source accumulation [s]	0.010
Ion accumulation time [s]	0.3
Ion cooling time [s]	0.050
Dwell time [s]	0.100
Capillary voltage [kV]	3.6
Spray Shield [V]	-500.0
Dry gas flow [l/min]	4.0
Dry gas temperature [°C]	200
Nebulizer gas [bar]	2.0
Side kick [V]	-5.0
Side kick off set [V]	3.0
Front plate voltage [V]	-0.4
Back plate voltage [V]	-0.5
Analyzer entrance [V]	10.0
Broadband high mass [Da]	1000
Broadband low mass [Da]	98.3

Supplementary Table 2: Standard compounds of direct infusion standard mix for ESI-parameter optimization

Direct infusion standard mix					
	Name	Supplier	Monoisotopical mass [M+H] ⁺	Monoisotopical mass [M-H] ⁻	Concentration [µg/mL]
1	L-Leucine	Sigma-Aldrich, Taufkirchen, Germany	132.101905	130.086255	3
2	L-Aspartic acid	Sigma-Aldrich, Taufkirchen, Germany	134.044784	132.029134	3
3	L-Lysine	Sigma-Aldrich, Taufkirchen, Germany	147.112804	145.097154	3
4	[D10] Adipic acid	Sigma-Aldrich, Taufkirchen, Germany		153.108846	3
5	L-Phenylalanine	Sigma-Aldrich, Taufkirchen, Germany	166.086255	164.070605	3
6	Decanoic acid-1-13C	Sigma-Aldrich, Taufkirchen, Germany		172.142403	3
7	Fructose	Sigma-Aldrich, Taufkirchen, Germany		179.055015	3
8	[D3]-Acetyl-L-Carnitine	Dr. Herman J. ten Brink, VU medical center, Netherlands	207.141865		3
9	Valeryl-L-carnitine	Dr. Herman J. ten Brink, VU medical center, Netherlands	246.18815		3
10	Nialamide	Sigma-Aldrich, Taufkirchen, Germany	299.15052	297.135699	3
11	[D6]-Sulfadimethoxine	Sigma-Aldrich, Taufkirchen, Germany	317.118513	315.10396	3
12	[D3]-Hexadecanoyl-L-carnitine	Dr. Herman J. ten Brink, VU medical center, Netherlands	403.360966		3
13	Terfenadine	Sigma-Aldrich, Taufkirchen, Germany	472.321006		3
14	1-Palmitoyl-sn-glycero-3-phosphocholine	Sigma-Aldrich, Taufkirchen, Germany	496.339766		3
15	Reserpine	Sigma-Aldrich, Taufkirchen, Germany	609.280657	607.266104	3

Supplementary Table 3: Standard compounds of HILIC method development standard mix for chromatography optimization

HILIC method development standard mix					
	Name	Supplier	Monoisotopical mass [M+H] ⁺	Monoisotopical mass [M-H] ⁻	Concentration [µg/mL]
1	L-Leucine	Sigma-Aldrich, Taufkirchen, Germany	132.101905	130.087352	20
2	L-Aspartic acid	Sigma-Aldrich, Taufkirchen, Germany	134.044784	132.030231	20
3	L-Lysine	Sigma-Aldrich, Taufkirchen, Germany	147.112804	145.098251	20
4	[D10] Adipic acid	Sigma-Aldrich, Taufkirchen, Germany		153.108846	3
5	L-Phenylalanine	Sigma-Aldrich, Taufkirchen, Germany	166.086255	164.071702	3
6	Decanoic acid-1-13C	Sigma-Aldrich, Taufkirchen, Germany		172.142408	3
7	Fructose	Sigma-Aldrich, Taufkirchen, Germany		179.056112	20
8	[D3]-Acetyl-L-Carnitine	Dr. Herman J. ten Brink, VU medical center, Netherlands	207.141865		3
9	Valeryl-L-carnitine	Dr. Herman J. ten Brink, VU medical center, Netherlands	246.18815		3
10	Nialamide	Sigma-Aldrich, Taufkirchen, Germany	299.15052	297.135699	3
11	[D6]-Sulfadimethoxine	Sigma-Aldrich, Taufkirchen, Germany	317.118513	315.10396	3
12	[D3]-Hexadecanoyl-L-carnitine	Dr. Herman J. ten Brink, VU medical center, Netherlands	403.360966		3
13	[D4]-Cholic acid	Sigma-Aldrich, Taufkirchen, Germany		411.305405	3
14	Terfenadine	Sigma-Aldrich, Taufkirchen, Germany	472.321006		3
15	1-Palmitoyl-sn-glycero-3-phosphocholine	Sigma-Aldrich, Taufkirchen, Germany	496.339766		3
16	Reserpine	Sigma-Aldrich, Taufkirchen, Germany	609.280657	607.266104	3

Supplementary Table 4: Columns tested during method optimization

Name	Ligand type	Supplier
Waters BEH HILIC 150x2.1 mm 1.7 µm	Bare silica	Waters GmbH, Eschborn, Germany
Waters BEH Amide 150.x2.1 mm 1.7 µm	Amide group	Waters GmbH, Eschborn, Germany
Grace Vision HT HILIC 150x2.0 mm 1.5 µm	Bare silica	Alltech Chrom GmbH, Worms, Germany
Grace Vision HT amino 150x2.0 mm 1.5 µm	Amino group	Alltech Chrom GmbH, Worms, Germany
Nucleodur HILIC 150x2.0 mm 3.0 µm	Sulfobetaine group	Macherey-Nagel GmbH & Co. KG, Düren, Germany
YMC-Triart Diol HILIC 150x2.0 mm 3.0 µm	Diol group	YMC Europe GmbH, Dinslaken, Germany

Supplementary Table 5: Genedata parameters for column testing evaluation, ESI-NEG mode

Activity	Settings
Load from File	Name: Test from Synapt Format: waters_raw_data_parser Data Source: _func001 Files/Folders: /Home/HILIC Test: 2012_10_17_UPLCBEHAmide2Standardneg.raw, 2012_10_17_UPLCBEHHILIC2Standardneg.raw, 2012_10_17_UPLCGraceHILIC3Standardneg.raw, 2012_10_17_UPLCGraceHILICaminoStandardneg.raw, 2012_10_19_UPLCBEHAmide2Standardneg.raw, 2012_10_19_UPLCBEHHILIC2Standardneg.raw, 2012_10_19_UPLCGraceHILIC3Standardneg.raw, 2012_10_19_UPLCGraceHILICaminoStandardneg.raw, 2012_10_26_UPLCBEHAmide2Standardneg.raw, 2012_10_26_UPLCGraceHILIC3Standardneg.raw, 2012_10_29_UPLCBEHAmide2Standardneg.raw, 2012_10_29_UPLCBEHAmide2StandardnegNH3.raw,

	2012_10_29_UPLCBEHHILIC2Standardneg.raw, 2012_10_29_UPLCGraceHILIC3Standardneg.raw, 2012_10_30_UPLCBEHAmide2Standardneg.raw, 2012_10_30_UPLCBEHHILIC2Standardneg.raw, 2012_10_30_UPLCGraceHILIC3Standardneg.raw, 2012_11_02_UPLCBEHHILIC2Standardneg.raw Profile Data Cutoff: 1.0 [Intensity] Centroid Data Cutoff: 0.0 [Intensity] Restrict Range: true m/z Minimum: 124 Da m/z Maximum: 625 Da RT Minimum: 1 [Time] RT Maximum: 17 [Time]
Converter Centroids To Profile	Peak Width: 110 ppm Grid Spacing: 5 Points / Sigma Cutoff: 3 Sigma Copy Centroid Data false
Chromatogram Lock Mass	Lock Masses: /Home/HILIC Test: HILICNEG.txt m/z Dependent Corrections true RT Dependent Corrections false Use Lockmass Chromatograms false
Chromatogram Chemical Noise Subtraction	Chemical Noise Subtraction: true RT Window: 1261 Scans Quantile: 50 Threshold: 0.0 [Intensity] RT Structure Removal: true Minimum RT Length: 6 Scans m/z Structure Removal: true Minimum m/z Length: 3 Points
m/z Track Filter	m/z Values: 130.086 132.029 145.097 153.100 164.070 172.142 179.055 297.136 315.104 411.311 607.266 Window Width: 0.1 Da Invert Selection: false

Chromatogram Summed Peak Detection	Summation Window: 5.0 [Time] Overlap: 50 Minimum Peak Size: 12 Scans Maximum Merge Distance: 5 Points Use Peak RT Splitting: true Intensity Profiling: Maximum Gap/Peak Ratio: 41 Use Smoothing: true Smoothing Window: 5 Points Ascent-based Peak Detection Apply Isolation Filter: true Isolation Threshold: 3 Export Summed Chromatogram false
Chromatogram Valid Peak Filter	Validity Threshold: 85 Max Intensity Present at Least in: 1 % of Chromatograms
Excel File	Export: Peaks Observable: Max. Intensity Export Experiment Annotation true Export Item Annotation: true Widths and Heights true Bounding Boxes false Keep duplicates false Custom File Name: true Name: HILIC NEG Add MH+ Column false Add Orig. RT Column false Export Annotations and Data as a Single File true

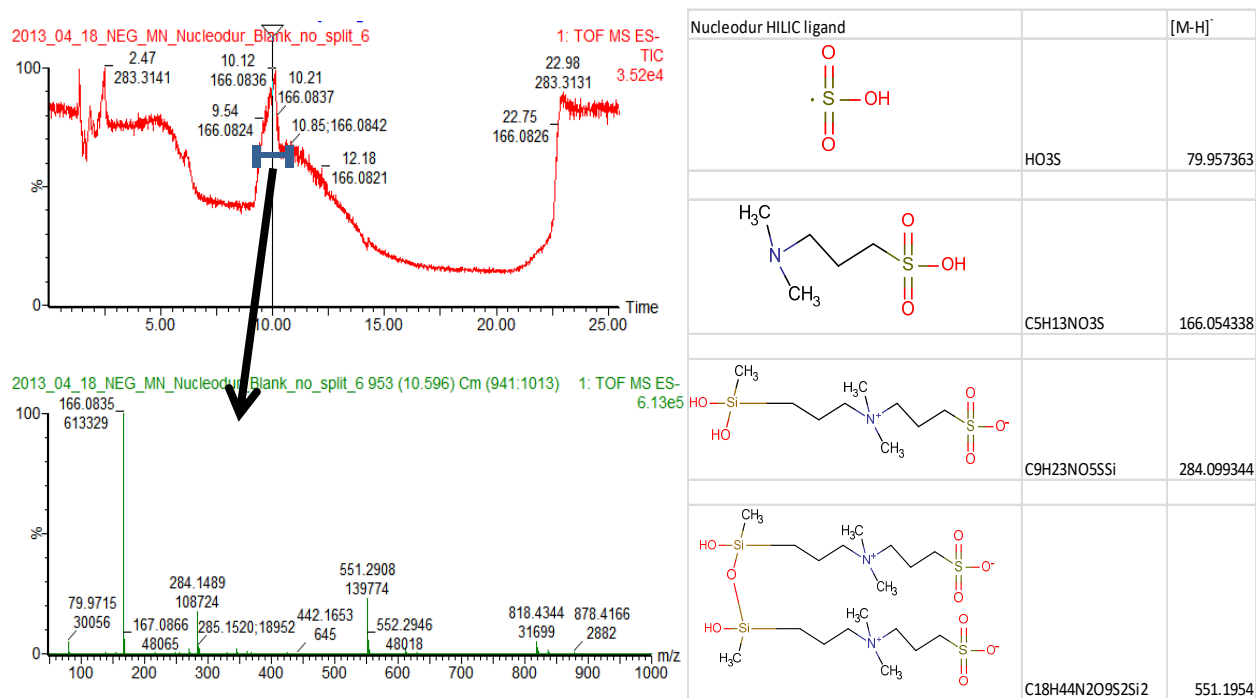
Supplementary Table 6: Genedata parameters for column testing evaluation, ESI-POS mode

Activity	Settings
Load from File	Name: Test from Synapt Format: waters_raw_data_parser Data Source: _func001 Files/Folders: /Home/HILIC Test: 2012_10_17_UPLCBEHAmide2Standardpos.raw, 2012_10_17_UPLCBEHHILIC2Standardpos.raw, 2012_10_17_UPLCGraceHILIC3Standardpos.raw, 2012_10_17_UPLCGraceHILICaminoStandardpos.raw, 2012_10_19_UPLCBEHAmide2Standardpos.raw, 2012_10_19_UPLCBEHHILIC2Standardpos.raw,

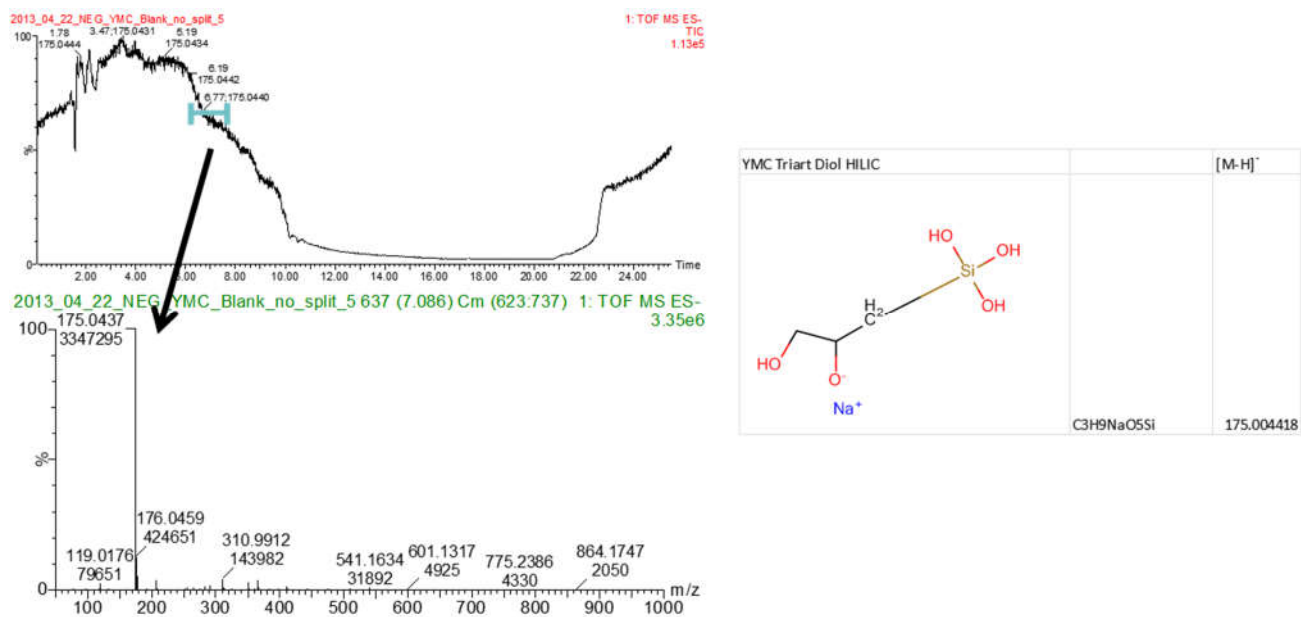
	2012_10_19_UPLCGraceHILIC3Standardpos.raw, 2012_10_19_UPLCGraceHILICaminoStandardpos.raw, 2012_10_26_UPLCBEHAMide2Standardpos.raw, 2012_10_26_UPLCGraceHILIC3Standardpos.raw, 2012_10_29_UPLCBEHAMide2Standardpos.raw, 2012_10_29_UPLCBEHAMide2StandardposNH3.raw, 2012_10_29_UPLCBEHHILIC2Standardpos.raw, 2012_10_29_UPLCGraceHILIC3Standardpos.raw, 2012_10_30_UPLCBEHAMide2Standardpos.raw, 2012_10_30_UPLCBEHHILIC2Standardpos.raw, 2012_10_30_UPLCGraceHILIC3Standardpos.raw, 2012_11_02_UPLCBEHHILIC2Standardpos.raw Profile Data Cutoff: 1.0 [Intensity] Centroid Data Cutoff: 0.0 [Intensity] Restrict Range: true m/z Minimum: 125 Da m/z Maximum: 625 Da RT Minimum: 1 [Time] RT Maximum: 17 [Time]
Converter Centroids To Profile	Peak Width: 110 ppm Grid Spacing: 5 Points / Sigma Cutoff: 3 Sigma Copy Centroid Data false
Chromatogram Lock Mass	Lock Masses: /Home/HILIC Test: HILICPOS.txt m/z Dependent Corrections true RT Dependent Corrections false Use Lockmass Chromatograms false
Chromatogram Chemical Noise Subtraction	Chemical Noise Subtraction: true RT Window: 1261 Scans Quantile: 50 Threshold: 0.0 [Intensity] RT Structure Removal: true Minimum RT Length: 6 Scans m/z Structure Removal: true Minimum m/z Length: 3 Points
m/z Track Filter	m/z Values: 132.102 134.044 147.113 166.086 207.141 246.188 299.150

	317.118
	403.361
	472.321
	496.339
	609.281
	Window Width: 0.1 Da
	Invert Selection: false
Chromatogram Summed Peak Detection	Summation Window: 5.0 [Time] Overlap: 50 Minimum Peak Size: 12 Scans Maximum Merge Distance: 5 Points Use Peak RT Splitting: true Intensity Profiling: Integral Gap/Peak Ratio: 30 Use Smoothing: true Smoothing Window: 3 Points Ascent-based Peak Detection Apply Isolation Filter: true Isolation Threshold: 3 Export Summed Chromatogram false
Chromatogram Valid Peak Filter Excel File	Validity Threshold: 85 Max Intensity Present at Least in: 1 % of Chromatograms Export: Peaks Observable: Max. Intensity Export Experiment Annotation true Export Item Annotation: true Widths and Heights true Bounding Boxes false Keep duplicates false Custom File Name: true Name: HILIC POS Add MH+ Column false Add Orig. RT Column false Export Annotations and Data as a Single File true

Appendix



Supplementary Figure 2: Mass signals observed in blank analysis run performed on sulfobetaine HILIC column



Supplementary Figure 3: Mass signals observed in blank analysis run performed on diol HILIC column

Supplementary Table 7: Lock-mass standard composition. Dissolved in methanol:water = 4:1

	Standard	Sumformular	Conc [$\mu\text{g/mL}$]	m/z $[\text{M}+\text{H}]^+$	m/z $[\text{M}-\text{H}]^-$
1	Erythromycin	$\text{C}_{37}\text{H}_{67}\text{NO}_{13}$	0.05	734.468518	732.453965
2	Leucine- enkephalin	$\text{C}_{28}\text{H}_{37}\text{N}_5\text{O}_7$	1	556.276575	554.262022
3	Arg-Gly-Asp-Ser	$\text{C}_{15}\text{H}_{27}\text{N}_7\text{O}_8$	0.3	434.199387	432.184834
4	Sulfadimethoxine	$\text{C}_{12}\text{H}_{14}\text{N}_4\text{O}_4\text{S}$	0.1	311.080852	309.066299
5	Alanine- Glutamine	$\text{C}_8\text{H}_{15}\text{N}_3\text{O}_4$	1	218.113532	216.09898
6	Leucine	$\text{C}_6\text{H}_{13}\text{NO}_2$	5	132.101905	130.087352

Appendix

Supplementary Table 8:VIP-Score list of urinary metabolomics gender study with prAN

m/z [M-H] ⁻	RT [min]	VIP-score	HMDB ID	HMDB Name
374.1125	7.11	4.96		
479.2252	5.17	3.63	HMDB10338 : HMDB33654 : HMDB36161 : HMDB38559	11-Oxo-androsterone glucuronide : 11- Deacetylvaltrate 11-(3-hydroxy-3- methylbutanoate) : 4-Propanoyl-HT2 toxin : 8- Hexanoylneosolaniol
467.2643	5.67	3.61	HMDB10321 : HMDB10339 : HMDB10359	3,17-Androstenediol glucuronide : 3-alpha- Androstenediol glucuronide : 17- Hydroxyandrostane-3-glucuronide
233.0788	3.58	3.57	HMDB11169 : HMDB28763 : HMDB29057	L-beta-aspartyl-L-threonine : Aspartyl- Threonine : Threoninyl-Aspartate :
143.0448	3.58	3.42	HMDB41946	N-Nitrosoproline
144.0324	6.20	3.35		
539.2482	5.56	3.35	HMDB10357 :	Tetrahydroaldosterone-3-glucuronide
540.252	5.56	3.34		
144.0315	6.08	3.29	HMDB01552	2-Keto-glutaramic acid
254.9837	0.82	3.26	HMDB60649	Ascorbic acid-2-sulfate
521.1758	5.64	3.24	HMDB15702	Ticagrelor
881.3654	5.55	3.23		
882.368	5.55	3.15		
115.0053	0.82	3.14		
203.1294	5.62	3.09	HMDB33217 : HMDB33642 : HMDB35133 : HMDB39053 : HMDB39974	(2xi,6xi)-7-Methyl-3-methylene-1,2,6,7- octanetetrol : (2xi,3xi,6E)-3,7-Dimethyl-6- octene-1,2,3,8-tetrol : 3,7-Dimethyl-3-octene- 1,2,6,7-tetrol : (1alpha,2alpha,4betaH,6alpha,8R)-p- Menthane-2,6,8,9-tetrol : (1S,2S,4R,8R)-p- Menthane-1,2,8,9-tetrol
74.02617	5.91	3.07		
198.1135	5.78	2.99	HMDB06406 : HMDB31820	Ecgonine methyl ester : Neotussilagine
520.1656	5.65	2.97		
465.3034	0.89	2.96	HMDB00653	Cholesterol sulfate
450.2557	4.98	2.95		

Supplementary Table 9: VIP-Score list of urinary metabolomics gender study with poAN

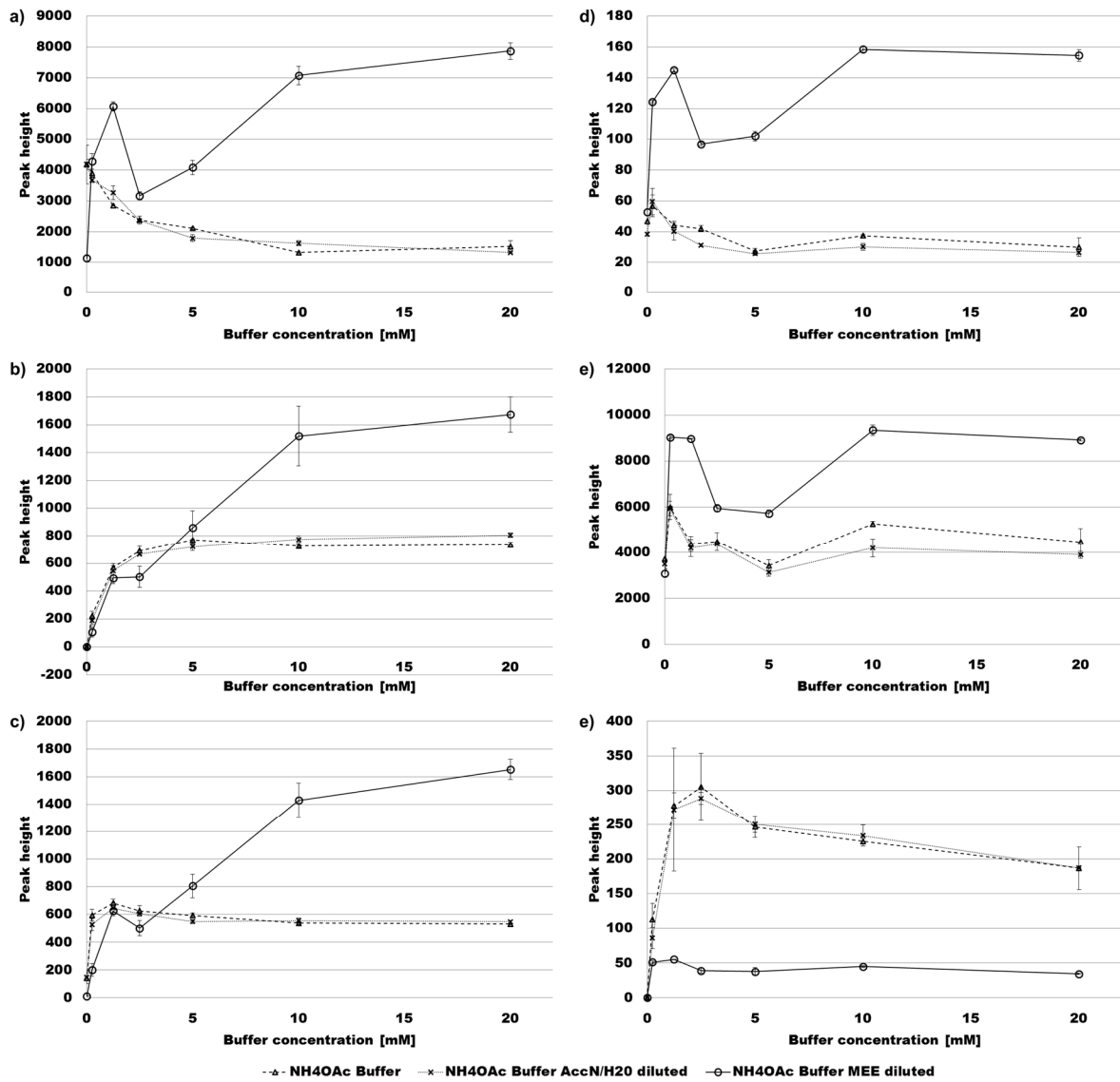
m/z [M-H] ⁻	RT [min]	VIP-score	HMDB ID	HMDB Name
328.0446	7.24	3.03	HMDB31800 : HMDB11616 : HMDB00058 : HMDB14492 HMDB02234 : HMDB01369 : HMDB00805 :	Mecarbam : Adenosine 2',3'-cyclic phosphate : Cyclic AMP : Nitisinone
128.0356	7.36	3.01	HMDB01843 : HMDB61093 : HMDB00267 : HMDB60262	1-Pyrroline-4-hydroxy-2-carboxylate : Pyrroline hydroxycarboxylic acid : Pyrrolidonecarboxylic acid : N-Acryloylglycine: dimethadione : Pyroglutamic acid : 5-Oxoprolinate
245.0433	7.35	2.83	HMDB34312 : HMDB 41525	Isopimpinellin : Coriandrone C
74.02431	6.11	2.82	HMDB00123 : HMDB14691 : HMDB 31239 HMDB00123 :	Glycine : Acetohydroxamic Acid : Ethyl nitrite
74.02441	6.64	2.80	HMDB14691 : HMDB 31239 HMDB00271 :	Glycine : Acetohydroxamic Acid : Ethyl nitrite
88.03971	5.53	2.67	HMDB00161 : HMDB31219 : HMDB00056 : HMDB01310 HMDB00271 :	Sarcosine : L-Alanine : Ethyl carbamate : Beta- Alanine : D-Alanine
88.03977	6.25	2.64	HMDB00161 : HMDB31219 : HMDB00056 : HMDB01310 HMDB00123 :	Sarcosine : L-Alanine : Ethyl carbamate : Beta- Alanine : D-Alanine
74.02436	6.90	2.63	HMDB14691 : HMDB 31239	Glycine : Acetohydroxamic Acid : Ethyl nitrite
218.104	6.26	2.61	HMDB00210 : HMDB12204	Pantothenic acid : Cis-zeatin
127.0505	7.52	2.61	HMDB29874 : HMDB31547 : HMDB00079	Squamolone : L-Cyclo(alanylglycyl) : Dihydrothymine
411.1252	7.51	2.60	HMDB41623 : HMDB13842	N6-Carbamoyl-L-threonyl-adenosine : O- Deethylated candesartan
118.0491	7.51	2.59		
245.0434	7.02	2.57	HMDB34312 : HMDB 41525 HMDB06833 : HMDB00622 : HMDB06855 :	Isopimpinellin : Coriandrone C
131.0344	4.41	2.56	HMDB00576 : HMDB02001 : HMDB01844 : HMDB33958 : HMDB29884 : HMDB00661 : HMDB40531 HMDB12249 :	2-Acetolactate : Ethylmalonic acid : (S)-2- Acetolactate : Monoethyl malonic acid : Dimethylmalonic acid : Methylsuccinic acid : 2- Deoxy-L-ribose-1,4-lactone : 2-C-Methyl-1,4- erythrone-D-lactone : Glutaric acid : 2-Hydroxy-4- oxopentanoic acid
116.0348	6.60	2.54	HMDB06454 : HMDB00532	L-Aspartate-semialdehyde : L-2-Amino-3- oxobutanoic acid : Acetylglycine
141.0668	7.29	2.52		
89.02342	5.16	2.51	HMDB00190	L-Lactic acid

Appendix

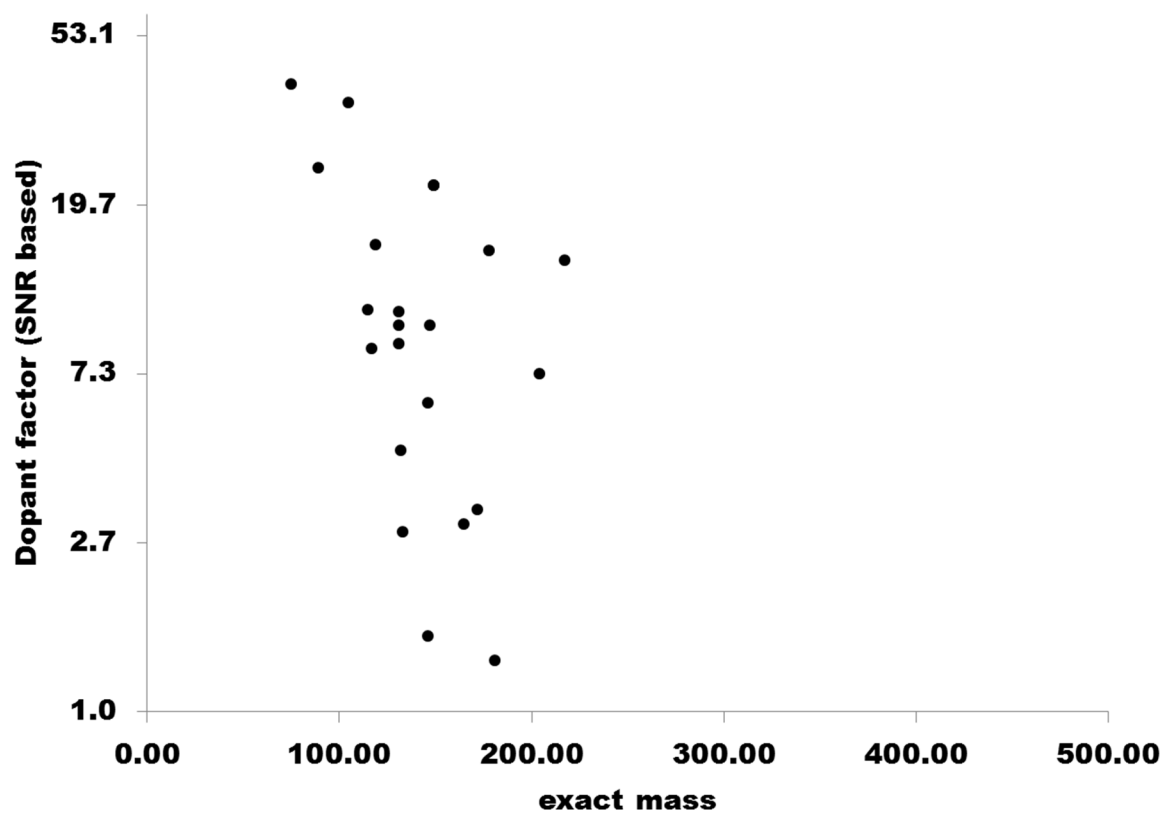
128.035	7.05	2.51	HMDB02234 : HMDB01369 : HMDB00805 : HMDB01843 : HMDB61093 : HMDB00267 : HMDB60262 : HMDB03831 : HMDB02184 : HMDB01119 : HMDB06050 : HMDB00158 : HMDB59720 : HMDB60280 :	1-Pyrroline-4-hydroxy-2-carboxylate : Pyrroline hydroxycarboxylic acid : Pyrrolidonecarboxylic acid : N-Acryloylglycine: dimethadione : Pyroglutamic acid : 5-Oxoprolinate Beta-Tyrosine : L-Threo-3-Phenylserine : 4- Hydroxy-4-(3-pyridyl)-butanoic acid : o-Tyrosine : L-Tyrosine : Meta-Tyrosine : 4,6,7-Trihydroxy- 1,2,3,4-tetrahydroisoquinoline
180.0654	6.28	2.50		
152.9957	6.58	2.49		

Supplementary Table 10: Standard mixture A

	Standard	Conc [ppm]
1	Leucine	63.83
2	Aspartic acid	212.77
3	Lysine	212.77
4	[d10] Adipidic acid	21.28
5	Phenylalanine	63.83
6	¹³ C-Decaonic acid	21.28
7	Fructose	212.77
8	Nialamide	21.28
9	[d6] Sulfadimethoxine	21.28
10	[d4]-Cholic acid	21.28
11	reserpine	21.28



Supplementary Figure 4: ESI-responses (peak height in counts) to various buffer types and concentrations of different standards, a) L-aspartic acid, b) L-lysine, c) L-phenylalanine, d) fructose, e) nialamide, f) reserpine



Supplementary Figure 5: Dopant factor based on signal to noise ratios over molecular weight

Supplementary Table 11: Physicochemical parameters and obtained dopant factor of analyzed standards

	Name	exact mass	logP	logD pH4.5	logD pH6.5	pka	Polarizability	pl	Mol Vol	Dopant factor (signal intensity based)	Dopant factor (SNR based)
1	L-Leucine	131.09	-1.59	-1.59	-1.59	2.79	13.84	6.14	136.06	18.3	9.7
2	L-Phenylalanine	165.08	-1.18	-1.19	-1.18	2.47	17.89	5.96	155.64	15.8	3.0
3	L-Tyrosine	181.07	-1.49	-1.49	-1.49	2.00	18.51	5.51	164.02	9.0	1.4
4	L-Serine	105.04	-3.89	-3.89	-3.89	2.03	9.06	5.69	93.55	24.6	35.9
5	L-Lysine	146.11	-3.21	-6.01	-5.66	2.74	15.37	9.82	147.80	12.6	1.6
6	L-Alanine	89.05	-2.84	-2.84	-2.84	2.47	8.36	5.98	84.92	39.8	24.4
7	L-Glutamine	146.07	-4.00	-4.00	-4.00	2.15	13.32	5.73	132.18	23.5	6.1
8	L-Glutamic acid	147.05	-3.24	-3.62	-5.40	1.88	12.69	2.80	129.68	11.4	9.7
9	Glycine	75.03	-3.41	-3.41	-3.41	2.31	6.55	5.78	67.84	48.9	39.8
10	L-Isoleucine	131.09	-1.51	-1.52	-1.51	2.79	13.84	6.19	136.20	18.1	8.7
11	L-Asparagine	132.05	-4.29	-4.29	-4.29	2.00	11.51	5.21	115.25	13.3	4.6
12	L-Valine	117.08	-1.95	-1.96	-1.95	2.72	12.00	6.15	119.06	23.1	8.4
13	L-Tryptophan	204.09	-1.09	-1.09	-1.09	2.54	23.09	5.97	182.29	7.5	7.3
14	N-Acetyl-L-tryptophan	246.10	1.00	0.46	-1.36	4.12	26.45	1.33	218.21	5.4	1.0
15	L-Proline	115.06	-2.57	-2.57	-2.57	1.94	11.29	7.11	107.85	23.2	10.6
16	L-Threonine	119.06	-3.47	-3.47	-3.47	2.21	10.87	5.60	110.52	14.4	15.6
17	L-Methionine	149.05	-2.19	-2.19	-2.19	2.53	15.07	6.01	137.72	22.0	22.1

18	Alanyl-Glutamine	217.11	-4.54	-4.58	-4.54	3.57	20.09	6.12	196.98	13.3	14.3
19	Arg-Gly-Asp-Ser	433.19	-9.27	-9.30	-9.27	-3.18	38.61	6.05	376.57	1.1	0.8
20	Hippurate	178.05	0.53	-0.43	-2.29	3.59	17.29	1.03	155.12	13.7	15.1
21	p-Hydroxyphenylacetic acid	152.05	1.31	0.69	-1.16	4.00	15.15		135.10	7.6	1.0
22	Decanoic acid	172.15	3.59	3.46	2.03	4.95	19.66		191.73	11.4	3.3
23	Lauric acid	200.18	4.48	4.35	2.92	4.95	23.34		225.71	8.4	0.4
24	Palmitic acid	256.24	6.26	6.13	4.70	4.95	30.71		293.75	5.9	0.6
25	Stearic acid	284.27	7.15	7.01	5.59	4.95	34.40		327.37	4.4	0.4
26	Norleucine	131.09	-1.43	-1.44	-1.43	2.79	13.84	6.15	135.88	17.8	10.5
27	D/L-Aspartic acid	133.04	-3.50	-3.58	-4.88	2.41	10.90	3.42	112.66	4.74	2.9
28	L-Arginine	174.11	-3.24	-6.21	-5.67	2.35	16.90	10.53	163.07	1.89	0.2
29	D/L-Cystein	121.02	-2.79	-2.79	-2.79	2.53	11.42	5.67	103.48	3.34	0.8
30	L-Methionine	149.05	-2.19	-2.19	-2.19		15.07	6.01	137.78	22.03	22.1

Genedata parameters for NUPREDM baseline analysis

Load from File

Properties

Name: NUPREDM HILIC-MS NEG

Format: Waters (*.raw)

Data Source: All

Files/Folders: /Shared/share_synapt_0/NUPREDM.PRO/Data:

2014_09_29_NUPREDEM_QC_01.raw, 2014_09_29_NUPREDEM_QC_02.raw, 2014_09_29_NUPREDEM_QC_03.raw,
2014_09_29_NUPREDEM_QC_04.raw, 2014_09_29_NUPREDEM_QC_05.raw, 2014_09_29_Subject_1254_1.raw,
2014_09_29_Subject_1254_2.raw, 2014_09_29_Subject_1254_3.raw, 2014_09_29_Subject_1913_1.raw, 2014_09_29_Subject_1913_2.raw,
2014_09_29_Subject_1913_3.raw, 2014_09_29_Subject_2042_1.raw, 2014_09_29_Subject_2042_2.raw, 2014_09_29_Subject_2042_3.raw,
2014_09_29_Subject_2051_1.raw, 2014_09_29_Subject_2051_2.raw, 2014_09_29_Subject_2051_3.raw, 2014_09_29_Subject_2190_1.raw,
2014_09_29_Subject_2190_2.raw, 2014_09_29_Subject_2190_3.raw, 2014_09_29_Subject_2218_1.raw, 2014_09_29_Subject_2218_2.raw,
2014_09_29_Subject_2218_3.raw, 2014_09_29_Subject_2224_1.raw, 2014_09_29_Subject_2224_2.raw, 2014_09_29_Subject_2224_3.raw,
2014_09_29_Subject_2258_1.raw, 2014_09_29_Subject_2258_2.raw, 2014_09_29_Subject_2258_3.raw, 2014_09_29_Subject_2265_1.raw,
2014_09_29_Subject_2265_2.raw, 2014_09_29_Subject_2265_3.raw, 2014_09_29_Subject_2266_1.raw, 2014_09_29_Subject_2266_2.raw,
2014_09_29_Subject_2266_3.raw, 2014_09_29_Subject_2285_1.raw, 2014_09_29_Subject_2285_2.raw, 2014_09_29_Subject_2285_3.raw,
2014_09_29_Subject_2331_1.raw, 2014_09_29_Subject_2331_2.raw, 2014_09_29_Subject_2331_3.raw, 2014_09_29_Subject_2341_1.raw,
2014_09_29_Subject_2341_2.raw, 2014_09_29_Subject_2341_3.raw, 2014_09_29_Subject_2350_1.raw, 2014_09_29_Subject_2350_2.raw,
2014_09_29_Subject_2350_3.raw, 2014_09_29_Subject_2437_1.raw, 2014_09_29_Subject_2437_2.raw, 2014_09_29_Subject_2437_3.raw,
2014_09_29_Subject_2438_1.raw, 2014_09_29_Subject_2438_2.raw, 2014_09_29_Subject_2438_3.raw, 2014_09_29_Subject_2464_1.raw,
2014_09_29_Subject_2464_2.raw, 2014_09_29_Subject_2464_3.raw, 2014_09_29_Subject_2468_1.raw, 2014_09_29_Subject_2468_2.raw,

Appendix

2014_09_26_Subject_2456_2.raw, 2014_09_26_Subject_2456_3.raw, 2014_09_26_Subject_2458_1.raw, 2014_09_26_Subject_2458_2.raw,
2014_09_26_Subject_2458_3.raw, 2014_09_26_Subject_2466_1.raw, 2014_09_26_Subject_2466_2.raw, 2014_09_26_Subject_2466_3.raw,
2014_09_26_Subject_2482_1.raw, 2014_09_26_Subject_2482_2.raw, 2014_09_26_Subject_2482_3.raw, 2014_09_26_Subject_2491_1.raw,
2014_09_26_Subject_2491_2.raw, 2014_09_26_Subject_2491_3.raw, 2014_09_26_Subject_2492_1.raw, 2014_09_26_Subject_2492_2.raw,
2014_09_26_Subject_2492_3.raw, 2014_09_26_Subject_2498_1.raw, 2014_09_26_Subject_2498_2.raw, 2014_09_26_Subject_2498_3.raw,
2014_09_26_Subject_2503_1.raw, 2014_09_26_Subject_2503_2.raw, 2014_09_26_Subject_2503_3.raw, 2014_09_26_Subject_2506_1.raw,
2014_09_26_Subject_2506_2.raw, 2014_09_26_Subject_2506_3.raw, 2014_09_26_Subject_2509_1.raw, 2014_09_26_Subject_2509_2.raw,
2014_09_26_Subject_2509_3.raw, 2014_09_26_Subject_2510_1.raw, 2014_09_26_Subject_2510_2.raw, 2014_09_26_Subject_2510_3.raw,
2014_09_26_Subject_2525_1.raw, 2014_09_26_Subject_2525_2.raw, 2014_09_26_Subject_2525_3.raw, 2014_09_26_Subject_2535_1.raw,
2014_09_26_Subject_2535_2.raw, 2014_09_26_Subject_2535_3.raw, 2014_09_26_Subject_2538_1.raw, 2014_09_26_Subject_2538_2.raw,
2014_09_26_Subject_2538_3.raw, 2014_09_26_Subject_2558_1.raw, 2014_09_26_Subject_2558_2.raw, 2014_09_26_Subject_2558_3.raw,
2014_09_26_Subject_2577_1.raw, 2014_09_26_Subject_2577_2.raw, 2014_09_26_Subject_2577_3.raw, 2014_09_26_Subject_2596_1.raw,
2014_09_26_Subject_2596_2.raw, 2014_09_26_Subject_2596_3.raw, 2014_09_26_Subject_2637_1.raw, 2014_09_26_Subject_2637_2.raw,
2014_09_26_Subject_2637_3.raw, 2014_09_26_Subject_2654_1.raw, 2014_09_26_Subject_2654_2.raw, 2014_09_26_Subject_2654_3.raw,
2014_09_26_Subject_2657_1.raw, 2014_09_26_Subject_2657_2.raw, 2014_09_26_Subject_2657_3.raw, 2014_09_26_Subject_2659_1.raw,
2014_09_26_Subject_2659_2.raw, 2014_09_26_Subject_2659_3.raw, 2014_09_23_NUPREDEM_QC_01.raw,
2014_09_23_NUPREDEM_QC_02.raw, 2014_09_23_NUPREDEM_QC_03.raw, 2014_09_23_NUPREDEM_QC_04.raw,
2014_09_23_NUPREDEM_QC_05.raw, 2014_09_23_Subject_573_1.raw, 2014_09_23_Subject_573_2.raw, 2014_09_23_Subject_573_3.raw,
2014_09_23_Subject_1616_1.raw, 2014_09_23_Subject_1616_2.raw, 2014_09_23_Subject_1616_3.raw, 2014_09_23_Subject_2172_1.raw,
2014_09_23_Subject_2172_2.raw, 2014_09_23_Subject_2172_3.raw, 2014_09_23_Subject_2207_1.raw, 2014_09_23_Subject_2207_2.raw,
2014_09_23_Subject_2207_3.raw, 2014_09_23_Subject_2213_1.raw, 2014_09_23_Subject_2213_2.raw, 2014_09_23_Subject_2213_3.raw,
2014_09_23_Subject_2214_1.raw, 2014_09_23_Subject_2214_2.raw, 2014_09_23_Subject_2214_3.raw, 2014_09_23_Subject_2219_1.raw,
2014_09_23_Subject_2219_2.raw, 2014_09_23_Subject_2219_3.raw, 2014_09_23_Subject_2237_1.raw, 2014_09_23_Subject_2237_2.raw,
2014_09_23_Subject_2237_3.raw, 2014_09_23_Subject_2255_1.raw, 2014_09_23_Subject_2255_2.raw, 2014_09_23_Subject_2255_3.raw,
2014_09_23_Subject_2257_1.raw, 2014_09_23_Subject_2257_2.raw, 2014_09_23_Subject_2257_3.raw, 2014_09_23_Subject_2263_1.raw,
2014_09_23_Subject_2263_2.raw, 2014_09_23_Subject_2263_3.raw, 2014_09_23_Subject_2273_1.raw, 2014_09_23_Subject_2273_2.raw,
2014_09_23_Subject_2273_3.raw, 2014_09_23_Subject_2275_1.raw, 2014_09_23_Subject_2275_2.raw, 2014_09_23_Subject_2275_3.raw,
2014_09_23_Subject_2279_1.raw, 2014_09_23_Subject_2279_2.raw, 2014_09_23_Subject_2279_3.raw, 2014_09_23_Subject_2294_1.raw,

2014_09_23_Subject_2294_2.raw, 2014_09_23_Subject_2294_3.raw, 2014_09_23_Subject_2295_1.raw, 2014_09_23_Subject_2295_2.raw,
2014_09_23_Subject_2295_3.raw, 2014_09_23_Subject_2299_1.raw, 2014_09_23_Subject_2299_2.raw, 2014_09_23_Subject_2299_3.raw,
2014_09_23_Subject_2436_1.raw, 2014_09_23_Subject_2436_2.raw, 2014_09_23_Subject_2436_3.raw, 2014_09_23_Subject_2460_1.raw,
2014_09_23_Subject_2460_2.raw, 2014_09_23_Subject_2460_3.raw, 2014_09_23_Subject_2461_1.raw, 2014_09_23_Subject_2461_2.raw,
2014_09_23_Subject_2461_3.raw, 2014_09_23_Subject_2469_1.raw, 2014_09_23_Subject_2469_2.raw, 2014_09_23_Subject_2469_3.raw,
2014_09_23_Subject_2481_1.raw, 2014_09_23_Subject_2481_2.raw, 2014_09_23_Subject_2481_3.raw, 2014_09_23_Subject_2488_1.raw,
2014_09_23_Subject_2488_2.raw, 2014_09_23_Subject_2488_3.raw, 2014_09_23_Subject_2489_1.raw, 2014_09_23_Subject_2489_2.raw,
2014_09_23_Subject_2489_3.raw, 2014_09_23_Subject_2518_1.raw, 2014_09_23_Subject_2518_2.raw, 2014_09_23_Subject_2518_3.raw,
2014_09_23_Subject_2533_1.raw, 2014_09_23_Subject_2533_2.raw, 2014_09_23_Subject_2533_3.raw, 2014_09_23_Subject_2549_1.raw,
2014_09_23_Subject_2549_2.raw, 2014_09_23_Subject_2549_3.raw, 2014_09_23_Subject_2554_1.raw, 2014_09_23_Subject_2554_2.raw,
2014_09_23_Subject_2554_3.raw, 2014_09_23_Subject_2572_1.raw, 2014_09_23_Subject_2572_2.raw, 2014_09_23_Subject_2572_3.raw,
2014_09_23_Subject_2582_1.raw, 2014_09_23_Subject_2582_2.raw, 2014_09_23_Subject_2582_3.raw, 2014_09_23_Subject_2590_1.raw,
2014_09_23_Subject_2590_2.raw, 2014_09_23_Subject_2590_3.raw, 2014_09_23_Subject_2626_1.raw, 2014_09_23_Subject_2626_2.raw,
2014_09_23_Subject_2626_3.raw, 2014_09_23_Subject_2648_1.raw, 2014_09_23_Subject_2648_2.raw, 2014_09_23_Subject_2648_3.raw,
2014_09_23_Subject_2649_1.raw, 2014_09_23_Subject_2649_2.raw, 2014_09_23_Subject_2649_3.raw, 2014_09_23_Subject_2653_1.raw,
2014_09_23_Subject_2653_2.raw, 2014_09_23_Subject_2653_3.raw, 2014_09_23_Subject_2680_1.raw, 2014_09_23_Subject_2680_2.raw,
2014_09_23_Subject_2680_3.raw, 2014_09_23_Subject_2715_1.raw, 2014_09_23_Subject_2715_2.raw, 2014_09_23_Subject_2715_3.raw

Advanced

Profile Data Cutoff: 1.0
[Intensity]

Centroid Data Cutoff: 5
[Intensity]

Restrict Range: true

m/z Minimum:	Da
m/z Maximum:	Da
RT Minimum:	0 Minutes
RT Maximum:	17 Minutes

Converter Centroids to Profile

Properties

Peak Sigma:	47 ppm
Grid Spacing:	2 Points / Sigma
Cutoff:	6 Sigma
Min. Intensity:	1
Copy Centroid Data	false

Chromatogramm Lock Mass

Properties

Lock Masses:	File
--------------	------

/Shared/share_synapt_0/HILIC METH.PRO/Sodium
potasium acetate NEG until 700mz.txt

m/z Dependence: Linear

RT Dependence: None

Use Lockmass false

Chromatograms

Meta Data Import

Properties

Files: /Shared/share_synapt_0/NUPREDM.PRO/NUPREDM
Dil Study MEE Meta Data.txt

RT structure removal

Properties

Minimum RT Length: 9 Scans

m/z Structure Removal

Properties

Minimum m/z Length: 8 Points

Chromatogram Chemical Noise Subtraction

Properties

Chromatogram Smoothing:	true
RT Window:	7 Scans
Estimator:	Moving Average
Chemical Noise Subtraction:	true
RT Window:	601 Scans
Quantile:	30%
Threshold:	0.0 [Intensity]

Chromatogram RT alignment

Properties

Alignment Scheme: Reference Chromatogram

Reference Index: 110

RT Search Interval: 179 Scans

Advanced

m/z Window: 7 Points

RT Window: 7 Scans

Gap Penalty: 0.5

Chromatogramm peak detection

Properties

Summation Window:	14 Scans
Overlap:	50%
Minimum Peak Size:	8 Scans
Maximum Merge Distance:	110 ppm
Use Peak RT Splitting:	true
Intensity Profiling:	Maximum
Gap/Peak Ratio:	30%

Peak detection

Method:	Resolution-based Peak Detection
Detection:	Aggressive (more peaks)

Chromatogram valid peak filter

Properties

Validity Threshold:	50 Max. Intensity
Present at Least in:	25 % of Chromatograms

Library MS search

Properties

Library:	Tab-Separated File
File:	/Shared/share_synapt_0/HILIC METH.PRO/HMDB 3.5 Library with adducts.txt
Ionization:	Deprotonation
Mass Tolerance:	10 ppm

RT/RI Tolerance: 20 Minutes

Filter by Cluster Formula false

Limit to Best Match false

Export Excel file

Properties

Type: Peaks

Observable: Summed Intensity

Include Experiment Annotations true

Include Item Annotations true

Custom Destination: true

Export File: /Home/HILIC Test/NUPREDM HILIC-MS/NUPREDM DiI
Study MEE Deisotoped.xlsx

Advanced

Add Orig. RT Column false

Export Annotations and Data as
a Single File true

Add MH+ Column false

Keep Duplicated Item Annotations false

Supplementary Table 12: Annotated features correlating with *ISI-matsuda* in OPLS with $VIP > 1.5$

PeakID	(m/z)	Retention time (min)	Var ID (Adduct)	HMDB ID	CoeffCS (ISI)	VIP
2294	129,0190	1,30	[M-H+]-	HMDB00620	-0,00271	3,736
2295	129,0200	1,85	[M-H+]-	HMDB00634	-0,00314	3,718
616	85,0289	1,30	[M-H+]-	HMDB00549	-0,00264	3,618
9889	267,0730	7,00	[M-H+]-	HMDB00425	-0,00286	3,602
9888	267,0730	6,69	[M-H+]-	HMDB00425	-0,00296	3,468
12919	335,0760	6,80	[M-H+]-	HMDB33563	-0,00300	3,448
2255	128,0350	7,12	[M-H+]-	HMDB02234	-0,00234	3,432
12832	333,0640	6,29	[M-H+]-	HMDB11649	-0,00257	3,426
19636	630,1870	7,51	[M-H+]-	HMDB06626	-0,00300	3,382
9912	267,0970	7,08	[M-H+]-	HMDB28848	-0,00274	3,365
1833	119,0370	5,83	[M-H+]-	HMDB01366	-0,00210	3,327
1215	102,0560	5,96	[M-H+]-	HMDB00112	-0,00265	3,310
12833	333,0610	6,64	[M-H+]-	HMDB11649	-0,00277	3,248
4014	159,0270	4,96	[M-H+]-	HMDB00225	-0,00258	3,192
2906	140,0730	6,84	[M-H+]-	HMDB29427	-0,00224	3,168
12935	335,1220	6,79	[M-H+]-	HMDB33107	-0,00238	3,163
3155	144,1030	5,88	[M-H+]-	HMDB06831	-0,00328	3,148
18574	539,2510	5,06	[M-H+]-	HMDB10357	-0,00313	3,123
1450	111,0090	1,26	[M-H+]-	HMDB00444	-0,00232	3,122
18602	541,2650	6,47	[M-H+]-	HMDB10320	-0,00296	3,109
3124	144,0300	6,95	[M-H+]-	HMDB30276	-0,00219	3,089
13698	355,1290	7,86	[M+Cl]-	HMDB39475	-0,00259	3,084
1541	112,0410	8,04	[M-H+]-	HMDB01301	-0,00237	3,084
1809	117,0700	5,68	[M-H+]-	HMDB02362	0,00343	3,081
3838	155,0840	7,13	[M-H+]-	HMDB59764	-0,00212	3,069
7324	217,0850	7,12	[M-H+]-	HMDB28872	-0,00233	3,063
17188	470,1520	7,49	[M-H+]-	HMDB06485	-0,00234	3,061
8791	245,1240	9,30	[M-H+]-	HMDB61115	-0,00208	3,060
6665	204,0700	5,68	[M-H+]-	HMDB00671	0,00284	2,954
4748	173,0910	7,13	[M-H+]-	HMDB03357	-0,00230	2,951
11084	291,0980	6,97	[M-H+]-	HMDB12197	-0,00202	2,924
5608	186,0880	8,06	[M-H+]-	HMDB40012	-0,00229	2,922
364	74,0253	5,57	[M-H+]-	HMDB00123	0,00287	2,907
3674	153,0540	5,19	[M-H+]-	HMDB05784	-0,00249	2,850
18601	541,2650	6,17	[M-H+]-	HMDB10320	-0,00290	2,844
17509	484,2670	6,17	[M-H+]-	HMDB41421	0,00369	2,823
2006	123,0570	6,71	[M-H+]-	HMDB04181	-0,00253	2,820
19207	589,1530	4,80	[M-H+]-	HMDB37420	0,00244	2,815
1274	104,0370	5,58	[M-H+]-	HMDB00187	0,00236	2,807
16192	429,2010	4,77	[M-H+]-	HMDB31673	-0,00206	2,801
17487	483,2570	5,80	[M-H+]-	HMDB10359	0,00317	2,758
12848	333,0800	6,26	[M-H+]-	HMDB11649	-0,00209	2,758
4813	174,0780	7,60	[M-H+]-	HMDB02201	-0,00204	2,729
2384	130,0640	5,67	[M-H+]-	HMDB00064	0,00259	2,722
12177	317,0290	8,60	[M-H+]-	C04478	-0,00248	2,718
3125	144,0300	7,32	[M-H+]-	HMDB01552	-0,00102	2,710
2099	125,0240	5,51	[M-H+]-	HMDB34355	-0,00215	2,700
17489	483,2570	6,20	[M-H+]-	HMDB10359	0,00373	2,667
7804	227,0660	7,58	[M-H+]-	HMDB00012	-0,00188	2,662
12321	321,1300	11,05	[M-H+]-	HMDB33106	-0,00214	2,657
3008	142,0510	4,12	[M-H+]-	HMDB00894	-0,00273	2,656
11620	303,0360	5,68	[M-H+]-	HMDB01570	-0,00221	2,653
1451	111,0090	1,52	[M-H+]-	HMDB00444	-0,00202	2,630
18938	565,3010	6,68	[M-H+]-	HMDB41367	-0,00249	2,628
2810	138,9800	6,98	[M-H+]-	HMDB01494	-0,00168	2,618
6613	203,0680	6,96	[M-H+]-	HMDB11667	-0,00148	2,603

PeakID	(m/z)	Retention time (min)	Var ID (Adduct)	HMDB ID	CoeffCS (ISI)	VIP
617	85,0289	1,76	[M-H+]-	HMDB00549	-0,00194	2,597
18621	543,2700	6,17	[M-H+]-	LMST0101018	-0,00258	2,590
12433	324,0730	6,07	[M-H+]-	HMDB59997	0,00211	2,584
1756	116,0710	5,88	[M-H+]-	HMDB00883	-0,00231	2,577
3416	148,0410	6,06	[M-H+]-	HMDB04058	0,00209	2,568
1658	114,0570	7,66	[M-H+]-	HMDB00162	-0,00143	2,567
12952	336,0700	4,87	[M-H+]-	HMDB13189	-0,00191	2,549
2403	130,0700	5,66	[M-H+]-	HMDB00466	0,00238	2,539
5797	189,0390	5,44	[M-H+]-	HMDB59927	-0,00185	2,535
6388	199,0760	7,15	[M-H+]-	HMDB30646	-0,00187	2,533
1149	100,0770	8,05	[M-H+]-	HMDB12815	-0,00218	2,527
11396	299,0360	5,73	[M-H+]-	HMDB04078	-0,00179	2,517
3208	145,0620	7,21	[M-H+]-	HMDB00641	0,00305	2,508
6297	197,1290	7,03	[M-H+]-	HMDB36648	-0,00159	2,505
20673	781,2180	7,50	[M-H+]-	HMDB40541	-0,00227	2,505
5603	186,0770	6,92	[M-H+]-	HMDB12150	-0,00148	2,503
15436	405,1730	5,06	[M-H+]-	HMDB33611	-0,00138	2,500
15227	398,1210	7,03	[M-H+]-	HMDB40552	-0,00186	2,500
12430	324,0740	5,57	[M-H+]-	HMDB59997	0,00210	2,487
6590	202,1090	5,78	[M-H+]-	HMDB00201	-0,00239	2,482
5606	186,0790	9,69	[M-H+]-	HMDB12150	-0,00185	2,481
12468	325,0780	6,08	[M-H+]-	HMDB39723	0,00207	2,480
4673	172,0970	5,85	[M-H+]-	HMDB00701	0,00240	2,479
8615	242,1220	7,04	[M-H+]-	HMDB15636	-0,00133	2,479
2340	129,0670	9,86	[M-H+]-	HMDB29396	-0,00168	2,478
9399	258,0760	7,90	[M-H+]-	HMDB41934	-0,00160	2,468
12320	321,1270	9,93	[M-H+]-	HMDB33105	-0,00143	2,463
6352	198,1140	5,48	[M-H+]-	HMDB31820	0,00289	2,463
1440	110,0720	9,89	[M-H+]-	HMDB00870	-0,00170	2,461
19983	673,2300	8,64	[M-H+]-	HMDB06581	-0,00178	2,445
4071	160,0390	4,87	[M-H+]-	HMDB03320	-0,00203	2,440
4264	164,0710	5,78	[M-H+]-	HMDB00159	0,00331	2,440
3881	156,0660	5,98	[M-H+]-	HMDB00459	0,00274	2,437
17488	483,2580	5,97	[M-H+]-	HMDB10359	0,00332	2,435
18599	541,2650	5,64	[M-H+]-	HMDB10320	-0,00215	2,426
7161	214,1100	6,99	[M-H+]-	HMDB13124	0,00297	2,407
3415	148,0420	5,79	[M-H+]-	HMDB00696	0,00206	2,399
13693	355,0880	6,93	[M-H+]-	HMDB39722	-0,00183	2,397
14131	366,0810	5,80	[M-H+]-	HMDB38844	-0,00245	2,391
9707	264,1070	6,48	[M-H+]-	HMDB00235	0,00183	2,385
1320	107,0500	4,82	[M-H+]-	HMDB02048	0,00194	2,374
5955	192,0650	6,55	[M-H+]-	HMDB00859	0,00320	2,364
6442	200,0580	7,47	[M-H+]-	HMDB01280	-0,00133	2,355
18537	537,2260	4,84	[M-H+]-	HMDB34333	-0,00212	2,354
9663	263,1030	6,48	[M-H+]-	HMDB06344	0,00180	2,354
2218	127,0500	7,12	[M-H+]-	HMDB00079	-0,00167	2,353
1471	111,0080	7,94	[M-H+]-	HMDB00617	-0,00110	2,345
1044	98,0612	6,18	[M-H+]-	HMDB11749	0,00304	2,341
961	96,9675	6,98	[M-H+]-	HMDB02142	-0,00165	2,341
7724	225,1120	10,67	[M-H+]-	HMDB59729	-0,00085	2,338
3719	154,0140	7,87	[M-H+]-	HMDB02916	-0,00111	2,334
10733	284,0860	4,81	[M-H+]-	HMDB05923	0,00193	2,334
6793	208,0610	5,76	[M-H+]-	HMDB00735	0,00320	2,330
10660	283,0810	4,82	[M-H+]-	HMDB11686	0,00193	2,329
7330	217,1190	9,19	[M-H+]-	HMDB28938	-0,00148	2,328
4918	177,0240	5,19	[M-H+]-	HMDB31162	-0,00115	2,323
11088	291,0920	7,66	[M-H+]-	HMDB12197	-0,00088	2,316
5667	187,0950	6,89	[M-H+]-	HMDB00784	-0,00114	2,315
1139	100,0400	6,92	[M-H+]-	HMDB29387	-0,00142	2,302

Appendix

PeakID	(m/z)	Retention time (min)	Var ID (Adduct)	HMDB ID	CoeffCS (ISI)	VIP
5172	180,0310	4,76	[M-H+]-	HMDB33733	-0,00293	2,299
3747	154,0610	9,87	[M-H+]-	HMDB00177	-0,00150	2,291
4820	174,0780	6,64	[M-H+]-	HMDB34252	0,00273	2,290
17508	484,2660	5,97	[M-H+]-	HMDB41421	0,00309	2,290
11254	296,0720	6,07	[M-H+]-	HMDB37297	0,00244	2,289
369	74,0250	6,99	[M-H+]-	HMDB00123	0,00322	2,289
2108	125,0570	4,27	[M-H+]-	HMDB41922	-0,00141	2,287
10049	270,1130	7,39	[M-H+]-	HMDB33103	-0,00127	2,282
3408	148,0400	0,91	[M-H+]-	HMDB04058	0,00191	2,280
17690	495,0860	5,21	[M-H+]-	HMDB29678	-0,00242	2,279
3007	142,0510	3,80	[M-H+]-	HMDB00894	-0,00206	2,279
3207	145,0600	6,48	[M-H+]-	HMDB02031	0,00185	2,278
10554	281,1120	10,22	[M-H+]-	HMDB29115	-0,00200	2,271
2215	127,0510	6,48	[M-H+]-	HMDB00079	0,00200	2,271
4718	173,0560	6,68	[M-H+]-	HMDB06028	-0,00156	2,265
19528	617,2140	7,74	[M-H+]-	HMDB40379	-0,00137	2,265
19649	632,1990	7,49	[M-H+]-	HMDB00825	-0,00194	2,262
9892	267,1100	9,87	[M-H+]-	HMDB28886	-0,00138	2,248
13694	355,0880	7,80	[M-H+]-	HMDB29938	-0,00170	2,236
9776	265,1090	6,48	[M-H+]-	HMDB40891	0,00178	2,236
18578	539,2500	6,30	[M-H+]-	HMDB10357	-0,00238	2,236
15405	404,1660	10,45	[M-H+]-	HMDB15594	-0,00128	2,236
755	89,0244	7,42	[M-H+]-	HMDB01882	-0,00128	2,235
15807	415,1980	4,08	[M-H+]-	HMDB36340	-0,00158	2,231
2997	142,0520	6,18	[M-H+]-	HMDB00894	0,00267	2,224
1741	116,0330	6,98	[M-H+]-	HMDB12249	0,00314	2,220
13109	339,0920	6,38	[M-H+]-	HMDB01030	-0,00149	2,215
440	79,9578	1,32	[M-H+]-	HMDB01033	0,00235	2,203
4256	164,0370	0,93	[M-H+]-	HMDB03454	0,00237	2,203
1279	105,0380	6,41	[M-H+]-	HMDB31716	-0,00125	2,198
4118	161,0470	4,57	[M-H+]-	HMDB00345	-0,00149	2,188
4350	166,0210	5,98	[M-H+]-	HMDB14496	-0,00098	2,186
7985	229,1470	6,31	[M-H+]-	HMDB00623	-0,00149	2,179
11715	305,0790	5,88	[M-H+]-	HMDB31870	-0,00202	2,172
6796	208,0720	9,86	[M-H+]-	HMDB60384	-0,00135	2,164
7331	217,1170	9,57	[M-H+]-	HMDB29043	-0,00156	2,161
783	91,0402	4,25	[M-H+]-	HMDB00131	-0,00166	2,154
4079	160,0610	9,99	[M-H+]-	HMDB00510	-0,00258	2,154
13763	357,0460	4,71	[M-H+]-	HMDB37203	-0,00190	2,150
16346	435,1280	6,82	[M-H+]-	HMDB36634	-0,00126	2,150
10806	285,0950	4,81	[M-H+]-	HMDB38331	0,00173	2,148
684	87,0081	7,68	[M-H+]-	HMDB00243	-0,00084	2,147
1610	113,0360	6,05	[M-H+]-	HMDB00076	0,00179	2,144
11589	302,0700	7,65	[M-H+]-	HMDB14436	-0,00116	2,140
1310	107,0500	0,86	[M-H+]-	HMDB02055	0,00153	2,139
1278	105,0350	0,83	[M-H+]-	HMDB06115	0,00152	2,136
2253	128,0360	6,47	[M-H+]-	HMDB00805	0,00191	2,135
10418	279,0630	8,88	[M-H+]-	HMDB31863	-0,00098	2,123
13112	339,0900	7,92	[M-H+]-	HMDB01030	-0,00100	2,122
7962	229,1180	6,13	[M-H+]-	HMDB28876	-0,00112	2,116
9664	263,1060	7,76	[M-H+]-	HMDB06344	0,00203	2,116
3009	142,0510	4,37	[M-H+]-	HMDB61705	-0,00171	2,115
480	81,0348	5,52	[M-H+]-	HMDB13749	-0,00160	2,114
2904	140,0730	4,53	[M-H+]-	HMDB29427	-0,00162	2,114
4450	167,9980	7,38	[M-H+]-	HMDB02757	-0,00077	2,110
15813	415,1970	1,06	[M-H+]-	HMDB36340	0,00127	2,105
269	71,0508	5,17	[M-H+]-	HMDB03543	-0,00120	2,097
5628	187,0060	0,86	[M-H+]-	HMDB11635	0,00148	2,097
1399	109,0660	5,20	[M-H+]-	HMDB40278	-0,00177	2,096

PeakID	(m/z)	Retention time (min)	Var ID (Adduct)	HMDB ID	CoeffCS (ISI)	VIP
657	86,0254	9,77	[M-H+]-	HMDB03609	-0,00102	2,093
19651	632,2020	8,13	[M-H+]-	HMDB06569	-0,00158	2,086
756	89,0241	7,79	[M-H+]-	HMDB00190	-0,00114	2,086
6619	203,0820	7,73	[M-H+]-	HMDB00929	-0,00136	2,086
11786	307,1160	9,85	[M-H+]-	HMDB37772	-0,00098	2,076
11567	301,1780	6,84	[M-H+]-	HMDB03959	-0,00113	2,068
1794	117,0190	11,67	[M-H+]-	HMDB00254	-0,00107	2,068
365	74,0252	5,90	[M-H+]-	HMDB00123	0,00171	2,064
7027	213,0180	6,35	[M-H+]-	HMDB01031	-0,00106	2,063
1832	119,0360	4,64	[M-H+]-	HMDB01366	-0,00085	2,057
6670	204,0870	7,73	[M-H+]-	HMDB36394	-0,00136	2,053
7815	227,1030	7,10	[M-H+]-	HMDB29018	-0,00094	2,051
3716	154,0140	7,14	[M-H+]-	HMDB02916	-0,00064	2,049
1536	112,0400	4,01	[M-H+]-	HMDB06875	-0,00179	2,046
295	72,0462	6,98	[M-H+]-	HMDB02134	0,00285	2,045
11565	301,1780	6,31	[M-H+]-	HMDB06771	-0,00139	2,041
1385	109,0400	6,48	[M-H+]-	HMDB03905	0,00159	2,034
10750	284,1300	7,08	[M-H+]-	HMDB29377	-0,00055	2,032
3384	147,0650	6,49	[M-H+]-	HMDB00227	0,00172	2,029
7578	222,0420	5,65	[M-H+]-	HMDB04083	0,00282	2,027
4494	168,9820	1,14	[M-H+]-	HMDB60176	-0,00120	2,025
7271	216,1260	5,65	[M-H+]-	HMDB00824	-0,00150	2,025
11622	303,0750	5,24	[M-H+]-	HMDB61135	-0,00118	2,024
2007	123,0570	7,09	[M-H+]-	HMDB04181	-0,00174	2,019
1410	110,0250	7,87	[M-H+]-	HMDB04230	-0,00056	2,014
781	91,0223	6,64	[M-H+]-	HMDB30299	-0,00219	2,011
88	59,0146	4,60	[M-H+]-	HMDB03344	-0,00104	2,010
12619	328,0810	6,56	[M-H+]-	HMDB38575	-0,00148	2,009
377	75,0092	10,50	[M-H+]-	HMDB00115	-0,00092	2,007
2803	138,0560	7,66	[M-H+]-	HMDB12153	-0,00058	1,992
3409	148,0410	1,31	[M-H+]-	HMDB04058	0,00172	1,990
18087	513,2530	5,73	[M-H+]-	HMDB36435	-0,00152	1,976
18575	539,2510	5,54	[M-H+]-	HMDB10357	-0,00194	1,974
10591	282,0780	6,92	[M-H+]-	HMDB38943	0,00280	1,974
4044	159,0920	7,73	[M-H+]-	HMDB00303	-0,00117	1,970
8782	245,0970	5,92	[M-H+]-	HMDB28981	0,00182	1,967
1133	100,0410	4,74	[M-H+]-	HMDB29387	-0,00148	1,966
945	96,0461	7,70	[M-H+]-	HMDB32969	-0,00082	1,965
14438	377,0330	1,32	[M-H+]-	HMDB60414	0,00192	1,965
6416	199,1690	3,64	[M-H+]-	HMDB00638	0,00264	1,964
2316	129,0200	7,87	[M-H+]-	HMDB59743	-0,00106	1,964
5083	179,0360	4,74	[M-H+]-	HMDB11663	-0,00116	1,962
6621	203,0840	6,62	[M-H+]-	HMDB13609	-0,00180	1,959
16184	429,1540	4,84	[M-H+]-	HMDB33277	-0,00225	1,951
16906	457,1910	6,59	[M-H+]-	HMDB34464	-0,00133	1,948
5733	188,0920	5,88	[M-H+]-	HMDB36604	-0,00177	1,947
19368	603,1640	7,66	[M-H+]-	HMDB12947	-0,00101	1,946
8795	245,1150	10,35	[M-H+]-	HMDB11172	-0,00077	1,942
233	70,0305	9,87	[M-H+]-	HMDB04296	-0,00149	1,942
2301	129,0190	4,80	[M-H+]-	HMDB13233	0,00152	1,941
1746	116,0500	7,73	[M-H+]-	HMDB00738	-0,00119	1,937
6353	198,1130	5,76	[M-H+]-	HMDB31820	0,00223	1,937
13111	339,0920	7,62	[M-H+]-	HMDB01030	-0,00133	1,935
555	83,0141	7,96	[M-H+]-	HMDB32330	-0,00076	1,933
4586	171,0120	7,23	[M-H+]-	HMDB59933	-0,00125	1,931
16468	439,1860	4,70	[M-H+]-	HMDB33468	-0,00102	1,927
3616	152,0350	3,65	[M-H+]-	HMDB01476	-0,00206	1,925
7524	221,0620	7,22	[M-H+]-	HMDB00099	-0,00060	1,918
789	91,0552	7,90	[M-H+]-	HMDB34168	-0,00149	1,912

Appendix

PeakID	(m/z)	Retention time (min)	Var ID (Adduct)	HMDB ID	CoeffCS (ISI)	VIP
6398	199,1130	10,00	[M-H+]-	HMDB35076	-0,00088	1,911
18134	516,1570	7,67	[M-H+]-	HMDB14464	-0,00125	1,901
13820	358,0840	7,42	[M-H+]-	HMDB30384	-0,00159	1,899
5013	178,0870	4,93	[M-H+]-	HMDB42012	-0,00097	1,899
14272	371,1450	4,37	[M-H+]-	HMDB02873	-0,00145	1,898
13389	345,1560	3,74	[M-H+]-	HMDB35822	-0,00152	1,896
3791	155,0000	7,92	[M-H+]-	HMDB04812	-0,00071	1,892
6269	197,0460	5,23	[M-H+]-	HMDB03503	-0,00087	1,891
11769	307,0620	5,16	[M-H+]-	HMDB41988	-0,00165	1,891
9904	267,1260	5,72	[M-H+]-	HMDB61643	-0,00116	1,891
1309	107,0360	8,93	[M-H+]-	HMDB34260	-0,00103	1,886
720	88,0399	8,68	[M-H+]-	HMDB00161	-0,00158	1,885
14547	379,1610	6,84	[M-H+]-	HMDB31954	-0,00103	1,883
8717	244,0030	0,98	[M-H+]-	HMDB60555	0,00193	1,874
1816	118,0510	7,33	[M-H+]-	HMDB00167	-0,00027	1,873
9107	253,1300	4,77	[M-H+]-	HMDB05767	-0,00165	1,870
4740	173,0820	4,69	[M-H+]-	HMDB00893	-0,00189	1,866
16995	462,1770	5,66	[M-H+]-	HMDB60821	-0,00122	1,859
4936	177,0390	1,89	[M-H+]-	HMDB01385	-0,00138	1,857
366	74,0253	6,17	[M-H+]-	HMDB00123	0,00188	1,857
1768	116,0720	9,92	[M-H+]-	HMDB00883	-0,00204	1,854
17946	507,1560	4,67	[M-H+]-	HMDB61229	-0,00062	1,854
10963	288,0710	2,04	[M-H+]-	HMDB12266	-0,00056	1,852
3389	147,0450	7,90	[M-H+]-	HMDB00567	-0,00134	1,850
14855	387,1640	5,10	[M-H+]-	HMDB41552	-0,00130	1,848
15438	405,1830	5,64	[M-H+]-	HMDB15267	-0,00098	1,845
14988	391,1000	7,73	[M-H+]-	HMDB38066	-0,00097	1,843
10293	276,0180	5,15	[M-H+]-	HMDB02028	-0,00143	1,842
1814	118,0510	6,80	[M-H+]-	HMDB61877	-0,00092	1,832
6925	211,0760	2,31	[M-H+]-	HMDB14814	-0,00205	1,832
13146	340,0680	5,89	[M-H+]-	HMDB61089	0,00133	1,831
2745	137,0350	9,86	[M-H+]-	HMDB00301	-0,00124	1,826
19462	611,1950	9,11	[M-H+]-	HMDB30542	-0,00138	1,825
371	74,0253	7,76	[M-H+]-	HMDB14691	-0,00105	1,825
1664	114,0930	5,64	[M-H+]-	HMDB62041	-0,00135	1,817
3821	155,0630	3,87	[M-H+]-	HMDB06524	-0,00195	1,815
4269	164,0710	7,91	[M-H+]-	HMDB62402	-0,00132	1,813
1786	117,0200	4,82	[M-H+]-	HMDB00202	0,00148	1,811
7076	213,1240	11,28	[M-H+]-	HMDB29030	-0,00107	1,811
5543	185,0570	7,32	[M-H+]-	HMDB61890	-0,00025	1,810
7822	227,1070	6,32	[M-H+]-	HMDB32133	-0,00092	1,808
1079	99,0094	6,84	[M-H+]-	HMDB32523	-0,00107	1,804
4207	163,0390	0,98	[M-H+]-	HMDB00205	0,00185	1,801
1535	112,0410	3,79	[M-H+]-	HMDB01301	-0,00165	1,800
171	67,0300	6,47	[M-H+]-	HMDB01525	0,00146	1,797
4035	159,0660	10,88	[M-H+]-	HMDB00857	0,00268	1,796
14932	389,1800	6,19	[M-H+]-	HMDB14875	-0,00125	1,795
7150	214,0490	9,31	[M-H+]-	HMDB00114	-0,00043	1,792
11963	312,0820	5,38	[M-H+]-	HMDB15510	0,00140	1,791
3420	148,0630	8,94	[M-H+]-	HMDB11600	-0,00093	1,791
5551	185,0960	9,62	[M-H+]-	HMDB30931	-0,00087	1,783
2254	128,0350	6,88	[M-H+]-	HMDB01369	-0,00052	1,783
2907	140,0730	7,02	[M-H+]-	HMDB30352	-0,00144	1,772
3494	150,0560	2,80	[M-H+]-	HMDB12219	0,00158	1,772
5102	179,0560	4,73	[M-H+]-	HMDB12326	-0,00192	1,766
600	84,0454	6,51	[M-H+]-	HMDB02039	0,00137	1,763
10492	280,1060	8,93	[M-H+]-	HMDB04326	-0,00071	1,762
5647	187,0780	6,95	[M-H+]-	HMDB32061	-0,00102	1,761
835	93,0343	5,83	[M-H+]-	HMDB00228	0,00186	1,760

PeakID	(m/z)	Retention time (min)	Var ID (Adduct)	HMDB ID	CoeffCS (ISI)	VIP
16902	457,1760	5,21	[M-H+]-	HMDB29909	-0,00156	1,760
14666	382,1010	7,22	[M-H+]-	HMDB00912	-0,00056	1,759
6386	199,1000	6,13	[M-H+]-	HMDB60606	-0,00087	1,756
4607	171,0670	4,23	[M-H+]-	HMDB00341	-0,00155	1,756
6818	209,0680	0,82	[M-H+]-	HMDB02123	-0,00172	1,753
2086	125,0350	2,70	[M-H+]-	HMDB02024	-0,00163	1,748
1758	116,0710	6,62	[M-H+]-	HMDB03355	-0,00155	1,747
13391	345,1700	4,81	[M-H+]-	HMDB41795	-0,00156	1,745
1018	98,0245	6,36	[M-H+]-	HMDB30276	-0,00110	1,744
6025	194,0450	5,87	[M-H+]-	HMDB01229	0,00161	1,740
18378	529,2670	4,99	[M-H+]-	HMDB34679	-0,00106	1,737
14866	387,1800	1,03	[M-H+]-	HMDB36587	0,00298	1,736
10057	270,1570	6,55	[M-H+]-	HMDB28717	-0,00080	1,734
3717	154,0140	7,33	[M-H+]-	HMDB02916	-0,00104	1,734
12738	331,1190	4,06	[M-H+]-	HMDB29082	-0,00172	1,732
2808	138,9710	1,11	[M-H+]-	HMDB38891	-0,00064	1,731
5512	184,1150	7,10	[M-H+]-	HMDB32201	-0,00191	1,729
11738	306,0280	6,06	[M-H+]-	HMDB14805	0,00265	1,728
2405	130,0870	6,25	[M-H+]-	HMDB00687	-0,00149	1,727
15100	393,2640	4,30	[M-H+]-	HMDB12530	-0,00133	1,725
8148	233,0820	6,99	[M-H+]-	HMDB41006	-0,00088	1,725
10659	283,0750	4,50	[M-H+]-	HMDB28787	-0,00114	1,724
4860	175,0240	6,14	[M-H+]-	HMDB00044	0,00162	1,724
15003	391,1950	6,27	[M-H+]-	HMDB15364	-0,00103	1,720
16608	446,1020	3,03	[M-H+]-	HMDB29250	0,00181	1,718
1590	113,0240	6,13	[M-H+]-	HMDB60461	0,00162	1,711
14859	387,1690	6,77	[M-H+]-	HMDB40706	-0,00148	1,708
14244	370,0960	0,88	[M-H+]-	HMDB14854	-0,00115	1,706
5169	180,0290	3,64	[M-H+]-	HMDB06954	-0,00193	1,705
6333	198,0400	2,03	[M-H+]-	HMDB14904	-0,00060	1,705
16570	444,1820	1,00	[M-H+]-	HMDB60848	-0,00092	1,703
4438	167,0840	7,08	[M-H+]-	HMDB01431	-0,00061	1,702
16096	425,1420	7,01	[M-H+]-	HMDB03556	-0,00043	1,698
13788	357,0870	6,54	[M-H+]-	HMDB01416	-0,00128	1,698
3630	152,0730	4,73	[M-H+]-	HMDB00073	-0,00159	1,696
9616	262,0880	6,59	[M-H+]-	HMDB28968	-0,00141	1,696
1073	99,0095	4,31	[M-H+]-	HMDB32523	-0,00110	1,695
7984	229,1330	6,13	[M-H+]-	HMDB15614	-0,00089	1,693
4499	169,0140	1,03	[M-H+]-	HMDB05807	-0,00130	1,692
16345	435,1180	6,51	[M-H+]-	HMDB42038	-0,00084	1,690
5568	185,1180	4,82	[M-H+]-	HMDB41376	-0,00066	1,688
11451	299,2580	4,99	[M-H+]-	HMDB61661	0,00138	1,687
9331	257,0780	8,87	[M-H+]-	HMDB02331	-0,00082	1,687
1652	114,0560	10,74	[M-H+]-	HMDB12880	-0,00016	1,686
375	75,0088	5,83	[M-H+]-	HMDB00115	-0,00046	1,683
3186	145,0340	5,17	[M-H+]-	HMDB40237	-0,00148	1,682
597	84,0459	5,75	[M-H+]-	HMDB02039	0,00266	1,681
14350	374,0200	8,34	[M-H+]-	HMDB62513	-0,00152	1,677
13196	341,0340	0,83	[M-H+]-	HMDB15034	0,00079	1,675
1146	100,0760	6,17	[M-H+]-	HMDB12815	0,00162	1,673
6078	195,0520	3,65	[M-H+]-	HMDB01982	-0,00188	1,669
19997	674,4060	4,45	[M-H+]-	HMDB12342	-0,00158	1,664
4325	165,0900	9,86	[M-H+]-	HMDB29862	-0,00126	1,663
1074	99,0088	4,81	[M-H+]-	HMDB32523	0,00132	1,663
4313	165,0400	10,67	[M-H+]-	HMDB00539	-0,00068	1,662
681	87,0087	5,59	[M-H+]-	HMDB11111	-0,00025	1,661
10514	281,0570	2,28	[M-H+]-	HMDB60868	-0,00059	1,660
18579	539,2510	6,69	[M-H+]-	HMDB10357	-0,00167	1,658
11082	291,1060	5,98	[M-H+]-	HMDB41269	-0,00206	1,658

Appendix

PeakID	(m/z)	Retention time (min)	Var ID (Adduct)	HMDB ID	CoeffCS (ISI)	VIP
9385	258,0380	0,95	[M-H+]-	HMDB01254	0,00186	1,657
2256	128,0350	7,51	[M-H+]-	HMDB00267	-0,00043	1,654
13462	347,2000	4,43	[M-H+]-	HMDB14452	-0,00182	1,654
1431	110,0610	4,38	[M-H+]-	HMDB37859	-0,00115	1,651
4549	169,1230	3,81	[M-H+]-	HMDB12183	-0,00133	1,650
11774	307,0740	6,06	[M-H+]-	HMDB14548	0,00136	1,650
17445	482,1530	9,31	[M-H+]-	HMDB31943	-0,00125	1,649
8050	231,0990	6,90	[M-H+]-	HMDB01199	-0,00044	1,648
5567	185,1180	4,28	[M-H+]-	HMDB41376	-0,00080	1,648
12887	334,1070	6,88	[M-H+]-	HMDB39504	-0,00053	1,647
7639	223,1090	7,04	[M-H+]-	HMDB14386	-0,00087	1,646
13856	359,1040	7,43	[M-H+]-	HMDB15173	-0,00060	1,644
11335	298,0730	3,85	[M-H+]-	HMDB38577	-0,00094	1,642
14368	374,1530	4,75	[M-H+]-	HMDB15178	-0,00071	1,642
7063	213,1240	10,08	[M-H+]-	HMDB29135	-0,00077	1,637
1020	98,0243	6,86	[M-H+]-	HMDB30276	-0,00130	1,635
2466	132,0300	9,62	[M-H+]-	HMDB00191	0,00237	1,629
1373	109,0300	3,18	[M-H+]-	HMDB00957	0,00147	1,625
4101	160,0770	9,92	[M-H+]-	HMDB03447	-0,00201	1,625
4949	177,0790	4,24	[M-H+]-	HMDB40348	-0,00130	1,623
19681	634,4030	4,44	[M-H+]-	ECMDB23449	-0,00111	1,623
1805	117,0560	7,72	[M-H+]-	HMDB61927	-0,00080	1,621
6383	199,0990	5,13	[M-H+]-	HMDB00603	-0,00078	1,616
6920	211,0600	5,13	[M-H+]-	HMDB00913	-0,00108	1,615
7963	229,1180	6,34	[M-H+]-	HMDB28876	-0,00068	1,612
12889	334,1250	10,44	[M-H+]-	HMDB00489	-0,00116	1,608
15420	405,0760	3,11	[M-H+]-	HMDB33069	0,00136	1,607
3187	145,0160	5,55	[M-H+]-	HMDB33966	-0,00058	1,606
15209	397,2170	4,25	[M-H+]-	HMDB41984	-0,00126	1,599
13046	338,0880	5,82	[M-H+]-	HMDB10363	0,00095	1,598
10865	286,0740	7,28	[M-H+]-	HMDB40683	-0,00118	1,597
4178	162,0410	5,83	[M-H+]-	HMDB00802	0,00147	1,594
12934	335,1380	6,09	[M-H+]-	HMDB41519	-0,00134	1,594
2408	130,0870	7,97	[M-H+]-	HMDB01901	-0,00133	1,591
18348	528,2650	5,06	[M-H+]-	HMDB02496	-0,00063	1,591
16867	455,1900	4,76	[M-H+]-	HMDB14843	-0,00024	1,589
12403	323,1050	5,35	[M-H+]-	HMDB14558	0,00178	1,587
9184	255,0480	5,11	[M-H+]-	HMDB33121	-0,00101	1,587
11224	295,1360	5,52	[M-H+]-	HMDB33950	0,00257	1,586
438	79,9578	0,83	[M-H+]-	HMDB01033	0,00099	1,583
1216	102,0570	6,22	[M-H+]-	HMDB02166	-0,00089	1,582
18086	513,2520	5,58	[M-H+]-	HMDB36435	-0,00105	1,578
4173	162,0420	1,13	[M-H+]-	HMDB00802	0,00120	1,578
1857	120,0460	0,89	[M-H+]-	HMDB04461	0,00128	1,575
7729	225,1050	4,98	[M-H+]-	HMDB15146	-0,00110	1,572
1164	101,0240	4,89	[M-H+]-	HMDB01259	-0,00106	1,566
6093	195,0530	1,73	[M-H+]-	HMDB11103	-0,00160	1,565
13979	363,1050	5,22	[M-H+]-	HMDB15024	0,00194	1,565
15000	391,1620	5,39	[M-H+]-	HMDB15339	-0,00135	1,558
16039	423,0910	3,12	[M-H+]-	HMDB32645	0,00093	1,558
13108	339,0870	5,84	[M-H+]-	HMDB30591	0,00099	1,558
7046	213,0610	8,77	[M-H+]-	HMDB40055	0,00180	1,558
1656	114,0570	6,75	[M-H+]-	HMDB30409	-0,00019	1,555
6995	212,1400	10,44	[M-H+]-	HMDB14371	-0,00063	1,553
9261	256,0600	7,75	[M-H+]-	HMDB41468	-0,00127	1,553
3123	144,0310	6,78	[M-H+]-	HMDB01552	-0,00043	1,551
5719	188,0340	5,26	[M-H+]-	HMDB00715	-0,00108	1,550
839	93,0452	13,84	[M-H+]-	HMDB33112	-0,00186	1,547
6518	201,0700	6,58	[M-H+]-	HMDB42002	-0,00084	1,547

PeakID	(m/z)	Retention time (min)	Var ID (Adduct)	HMDB ID	CoeffCS (ISI)	VIP
9848	266,1130	6,23	[M-H+]-	HMDB33758	-0,00114	1,547
17793	499,1930	4,66	[M-H+]-	HMDB34463	-0,00104	1,545
5109	179,0570	6,84	[M-H+]-	HMDB62538	-0,00167	1,542
8986	250,0720	7,05	[M-H+]-	HMDB59747	-0,00114	1,541
18725	550,1980	6,47	[M-H+]-	HMDB35475	0,00079	1,540
7960	229,0960	10,40	[M-H+]-	HMDB37942	-0,00010	1,537
2261	128,0350	10,50	[M-H+]-	HMDB02234	-0,00051	1,535
6694	205,0740	5,73	[M-H+]-	HMDB38791	0,00098	1,534
16295	433,2090	5,19	[M-H+]-	HMDB61126	-0,00017	1,531
10539	281,1030	6,90	[M-H+]-	HMDB13914	-0,00076	1,529
16643	447,0940	3,13	[M-H+]-	HMDB33751	0,00146	1,529
9775	265,1070	6,23	[M-H+]-	HMDB14383	-0,00109	1,527
2320	129,0380	7,51	[M-H+]-	HMDB40238	-0,00046	1,522
13212	341,1180	7,68	[M-H+]-	HMDB60900	-0,00067	1,522
7714	225,0970	5,17	[M-H+]-	HMDB37158	-0,00073	1,519
14034	364,1700	6,82	[M-H+]-	HMDB30380	-0,00131	1,518
5072	179,0350	0,87	[M-H+]-	HMDB06915	-0,00158	1,517
16010	422,0010	0,84	[M-H+]-	HMDB60639	0,00052	1,517
13110	339,0860	7,27	[M-H+]-	HMDB29841	-0,00088	1,516
2319	129,0390	6,48	[M-H+]-	HMDB40238	0,00119	1,514
6888	210,0770	7,03	[M-H+]-	HMDB11754	-0,00131	1,513
1408	110,0250	7,32	[M-H+]-	HMDB04230	-0,00087	1,512
10812	285,1470	5,83	[M-H+]-	HMDB00313	0,00128	1,512
2476	132,0450	5,66	[M-H+]-	HMDB61918	0,00063	1,505
14324	373,1130	7,40	[M-H+]-	HMDB34942	-0,00163	1,503
268	71,0509	4,47	[M-H+]-	HMDB03543	-0,00045	1,503
10340	277,0770	7,08	[M-H+]-	HMDB15522	-0,00136	1,502
2850	139,0510	8,98	[M-H+]-	HMDB02820	-0,00134	1,501
6092	195,0510	1,52	[M-H+]-	HMDB02271	-0,00176	1,501
10669	283,0960	5,11	[M-H+]-	HMDB03290	0,00090	1,500
9738	264,6250	5,02	[M-H+]-	HMDB38778	-0,00093	1,500

7 Literature

1. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med.* 2006;3(11):e442. doi: 10.1371/journal.pmed.0030442.
2. Global health risks: mortality and burden of disease attributable to selected major risks. In: Organization WH, editor. Geneva: World Health Organization; 2009.
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. doi: 10.1016/j.diabres.2017.03.024.
4. DeFronzo RA. Banting Lecture. From the triumvirate to the ominous octet: a new paradigm for the treatment of type 2 diabetes mellitus. *Diabetes.* 2009;58(4):773-95. doi: 10.2337/db09-9028.
5. Stumvoll M, Goldstein BJ, van Haeften TW. Type 2 diabetes: principles of pathogenesis and therapy. *The Lancet.* 2005;365(9467):1333-46. doi: 10.1016/s0140-6736(05)61032-x.
6. Chatterjee S, Khunti K, Davies MJ. Type 2 diabetes. *The Lancet.* 2017;389(10085):2239-51. doi: 10.1016/s0140-6736(17)30058-2.
7. Orozco LJ, Buchleitner AM, Gimenez-Perez G, Roque IFM, Richter B, Mauricio D. Exercise or exercise and diet for preventing type 2 diabetes mellitus. *The Cochrane database of systematic reviews.* 2008;(3):CD003054. doi: 10.1002/14651858.CD003054.pub3.
8. Merlotti C, Morabito A, Pontiroli AE. Prevention of type 2 diabetes; a systematic review and meta-analysis of different intervention strategies. *Diabetes, obesity & metabolism.* 2014;16(8):719-27. doi: 10.1111/dom.12270.
9. Long-term effects of lifestyle intervention or metformin on diabetes development and microvascular complications over 15-year follow-up: the Diabetes Prevention Program Outcomes Study. *The Lancet Diabetes & Endocrinology.* 2015;3(11):866-75. doi: 10.1016/s2213-8587(15)00291-0.
10. Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *The New England journal of medicine.* 2007;356(24):2457-71. doi: 10.1056/NEJMoa072761.
11. Chang SH, Stoll CR, Song J, Varela JE, Eagon CJ, Colditz GA. The effectiveness and risks of bariatric surgery: an updated systematic review and meta-analysis, 2003-2012. *JAMA surgery.* 2014;149(3):275-87. doi: 10.1001/jamasurg.2013.3654.

12. Fung TT, Hu FB, Pereira MA, Liu S, Stampfer MJ, Colditz GA, et al. Whole-grain intake and the risk of type 2 diabetes: a prospective study in men. *The American journal of clinical nutrition*. 2002;76(3):535-40.
13. Murtaugh MA, Jacobs DR, Jacob B, Steffen LM, Marquart L. Epidemiological support for the protection of whole grains against diabetes. *Proceedings of the Nutrition Society*. 2003;62(01):143-9. doi: 10.1079/pns2002223.
14. Weickert MO, Pfeiffer AF. Metabolic effects of dietary fiber consumption and prevention of diabetes. *The Journal of nutrition*. 2008;138(3):439-42.
15. Kaline K, Bornstein SR, Bergmann A, Hauner H, Schwarz PE. The importance and effect of dietary fiber in diabetes prevention with particular consideration of whole grain products. *Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme*. 2007;39(9):687-93. doi: 10.1055/s-2007-985811.
16. Galisteo M, Duarte J, Zarzuelo A. Effects of dietary fibers on disturbances clustered in the metabolic syndrome. *The Journal of nutritional biochemistry*. 2008;19(2):71-84. doi: 10.1016/j.jnutbio.2007.02.009.
17. den Besten G, van Eunen K, Groen AK, Venema K, Reijngoud DJ, Bakker BM. The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *Journal of lipid research*. 2013;54(9):2325-40. doi: 10.1194/jlr.R036012.
18. Fung TT, Schulze M, Manson JE, Willett WC, Hu FB. Dietary patterns, meat intake, and the risk of type 2 diabetes in women. *Archives of internal medicine*. 2004;164(20):2235-40. doi: 10.1001/archinte.164.20.2235.
19. Montonen J, Boeing H, Fritsche A, Schleicher E, Joost HG, Schulze MB, et al. Consumption of red meat and whole-grain bread in relation to biomarkers of obesity, inflammation, glucose metabolism and oxidative stress. *European journal of nutrition*. 2013;52(1):337-45. doi: 10.1007/s00394-012-0340-6.
20. Pan A, Sun Q, Bernstein AM, Manson JE, Willett WC, Hu FB. Changes in red meat consumption and subsequent risk of type 2 diabetes mellitus: three cohorts of US men and women. *JAMA internal medicine*. 2013;173(14):1328-35. doi: 10.1001/jamainternmed.2013.6633.
21. Barnard N, Levin S, Trapp C. Meat consumption as a risk factor for type 2 diabetes. *Nutrients*. 2014;6(2):897-910. doi: 10.3390/nu6020897.
22. Song Y, Manson JE, Buring JE, Liu S. A Prospective Study of Red Meat Consumption and Type 2 Diabetes in Middle-Aged and Elderly Women: The Women's Health Study. *Diabetes Care*. 2004;27(9):2108-15. doi: 10.2337/diacare.27.9.2108.

23. Vang A, Singh PN, Lee JW, Haddad EH, Brinegar CH. Meats, processed meats, obesity, weight gain and occurrence of diabetes among adults: findings from Adventist Health Studies. *Annals of nutrition & metabolism*. 2008;52(2):96-104. doi: 10.1159/000121365.
24. Eknoyan G. Santorio Sanctorius (1561–1636) – Founding Father of Metabolic Balance Studies. *American Journal of Nephrology*. 1999;19(2):226-33. doi: 10.1159/000013455.
25. Holmes E, Wilson ID, Nicholson JK. Metabolic phenotyping in health and disease. *Cell*. 2008;134(5):714-7. doi: 10.1016/j.cell.2008.08.026.
26. Oliver S. Systematic functional analysis of the yeast genome. *Trends in Biotechnology*. 1998;16(9):373-8. doi: 10.1016/s0167-7799(98)01214-1.
27. Nicholson JK, Lindon JC, Holmes E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica; the fate of foreign compounds in biological systems*. 1999;29(11):1181-9. doi: 10.1080/004982599238047.
28. Fiehn O. Metabolomics - The link between genotypes and phenotypes. *Plant Molecular Biology*. 2002;48(1/2):155-71. doi: 10.1023/a:1013713905833.
29. Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, Young N, et al. HMDB: the Human Metabolome Database. *Nucleic acids research*. 2007;35(Database issue):D521-6. doi: 10.1093/nar/gkl923.
30. Wishart DS, Knox C, Guo AC, Eisner R, Young N, Gautam B, et al. HMDB: a knowledgebase for the human metabolome. *Nucleic acids research*. 2009;37(Database issue):D603-10. doi: 10.1093/nar/gkn810.
31. Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, et al. HMDB 3.0--The Human Metabolome Database in 2013. *Nucleic acids research*. 2013;41(Database issue):D801-7. doi: 10.1093/nar/gks1065.
32. Smith CA, O'Maille G, Want EJ, Qin C, Trauger SA, Brandon TR, et al. METLIN: a metabolite mass spectral database. *Therapeutic drug monitoring*. 2005;27:747-51.
33. Guijas C, Montenegro-Burke JR, Domingo-Almenara X, Palermo A, Warth B, Hermann G, et al. METLIN: A Technology Platform for Identifying Knowns and Unknowns. *Analytical chemistry*. 2018;90(5):3156-64. doi: 10.1021/acs.analchem.7b04424.
34. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*. 2000;28(1):27-30.

35. van Deemter JJ, Zuiderweg FJ, Klinkenberg A. Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography. *Chemical Engineering Science*. 1956;5(6):271-89. doi: 10.1016/0009-2509(56)80003-1.
36. Abraham MH. 100 years of chromatography—or is it 171? *Journal of Chromatography A*. 2004;1061(1):113-4. doi: 10.1016/j.chroma.2004.10.060.
37. Bouatra S, Aziat F, Mandal R, Guo AC, Wilson MR, Knox C, et al. The human urine metabolome. *PloS one*. 2013;8(9):e73076. doi: 10.1371/journal.pone.0073076.
38. Alpert AJ. Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. *Journal of Chromatography A*. 1990;499:177-96. doi: [https://doi.org/10.1016/S0021-9673\(00\)96972-3](https://doi.org/10.1016/S0021-9673(00)96972-3).
39. Buszewski B, Noga S. Hydrophilic interaction liquid chromatography (HILIC)—a powerful separation technique. *Analytical and bioanalytical chemistry*. 2012;402(1):231-47. doi: 10.1007/s00216-011-5308-5.
40. Hemström P, Irgum K. Hydrophilic interaction chromatography. *Journal of Separation Science*. 2006;29(12):1784-821. doi: 10.1002/jssc.200600199.
41. McCalley DV. Understanding and manipulating the separation in hydrophilic interaction liquid chromatography. *Journal of chromatography A*. 2017;1523:49-71. doi: 10.1016/j.chroma.2017.06.026.
42. McCalley DV. Effect of mobile phase additives on solute retention at low aqueous pH in hydrophilic interaction liquid chromatography. *Journal of Chromatography A*. 2017;1483:71-9. doi: 10.1016/j.chroma.2016.12.035.
43. McCalley DV. Study of retention and peak shape in hydrophilic interaction chromatography over a wide pH range. *Journal of chromatography A*. 2015;1411:41-9. doi: 10.1016/j.chroma.2015.07.092.
44. Schuster G, Lindner W. Chocolate HILIC phases: development and characterization of novel saccharide-based stationary phases by applying non-enzymatic browning (Maillard reaction) on amino-modified silica surfaces. *Analytical and bioanalytical chemistry*. 2011;400(8):2539-54. doi: 10.1007/s00216-011-4745-5.
45. Eric S. Grumbach* DMW-D, Jeffrey R. Mazzeo, Bonnie Alden, and Pamela C. Iraneta. HILIC using silica columns for the retention of polar analytes and enhanced esi-ms sensitivity. *LCGC North America*. 2004;22(10):1010-23.
46. Trivedi DK, Iles RK. HILIC-MS-based shotgun metabolomic profiling of maternal urine at 9-23 weeks of gestation - establishing the baseline changes in the maternal metabolome. *Biomedical chromatography : BMC*. 2014. doi: 10.1002/bmc.3266.

47. Virgiliou C, Sampsonidis I, Gika HG, Raikos N, Theodoridis GA. Development and validation of a HILIC- MS/MS multi-targeted method for metabolomics applications. *Electrophoresis*. 2015. doi: 10.1002/elps.201500208.
48. Zhang T, Watson DG, Wang L, Abbas M, Murdoch L, Bashford L, et al. Application of Holistic Liquid Chromatography-High Resolution Mass Spectrometry Based Urinary Metabolomics for Prostate Cancer Detection and Biomarker Discovery. *PLoS one*. 2013;8(6):e65880. doi: 10.1371/journal.pone.0065880.
49. Contrepolis K, Jiang L, Snyder M. Optimized Analytical Procedures for the Untargeted Metabolomic Profiling of Human Urine and Plasma by Combining Hydrophilic Interaction (HILIC) and Reverse-Phase Liquid Chromatography (RPLC)-Mass Spectrometry. *Molecular & cellular proteomics : MCP*. 2015;14(6):1684-95. doi: 10.1074/mcp.M114.046508.
50. Calderon-Santiago M, Priego-Capote F, Jurado-Gamez B, Luque de Castro MD. Optimization study for metabolomics analysis of human sweat by liquid chromatography-tandem mass spectrometry in high resolution mode. *Journal of chromatography A*. 2014;1333:70-8. doi: 10.1016/j.chroma.2014.01.071.
51. Fei F, Bowdish DM, McCarry BE. Comprehensive and simultaneous coverage of lipid and polar metabolites for endogenous cellular metabolomics using HILIC-TOF-MS. *Analytical and bioanalytical chemistry*. 2014. doi: 10.1007/s00216-014-7797-5.
52. Holcapek M, Ovcacikova M, Lisa M, Cifkova E, Hajek T. Continuous comprehensive two-dimensional liquid chromatography-electrospray ionization mass spectrometry of complex lipidomic samples. *Analytical and bioanalytical chemistry*. 2015;407(17):5033-43. doi: 10.1007/s00216-015-8528-2.
53. Wang Y, Lehmann R, Lu X, Zhao X, Xu G. Novel, fully automatic hydrophilic interaction/reversed-phase column-switching high-performance liquid chromatographic system for the complementary analysis of polar and apolar compounds in complex samples. *Journal of chromatography A*. 2008;1204(1):28-34. doi: 10.1016/j.chroma.2008.07.010.
54. D'Attoma A, Grivel C, Heinisch S. On-line comprehensive two-dimensional separations of charged compounds using reversed-phase high performance liquid chromatography and hydrophilic interaction chromatography. Part I: orthogonality and practical peak capacity considerations. *Journal of chromatography A*. 2012;1262:148-59. doi: 10.1016/j.chroma.2012.09.028.
55. Klavins K, Drexler H, Hann S, Koellensperger G. Quantitative metabolite profiling utilizing parallel column analysis for simultaneous reversed-phase and hydrophilic interaction liquid chromatography separations combined with tandem mass spectrometry. *Analytical chemistry*. 2014;86(9):4145-50. doi: 10.1021/ac5003454.
56. Hemmler D, Heinzmann SS, Wöhr K, Schmitt-Kopplin P, Witting M. Tandem HILIC-RP liquid chromatography for increased polarity coverage in food analysis. *Electrophoresis*. 2018. doi: 10.1002/elps.201800038.

57. Forcisi S, Moritz F, Kanawati B, Tziotis D, Lehmann R, Schmitt-Kopplin P. Liquid chromatography-mass spectrometry in metabolomics research: mass analyzers in ultra high pressure liquid chromatography coupling. *Journal of chromatography A*. 2013;1292:51 - 65. doi: 10.1016/j.chroma.2013.04.017.
58. Konermann L, Ahadi E, Rodriguez AD, Vahidi S. Unraveling the mechanism of electrospray ionization. *Analytical chemistry*. 2013;85(1):2-9. doi: 10.1021/ac302789c.
59. Marshall AG, Chen T. 40 years of Fourier transform ion cyclotron resonance mass spectrometry. *International Journal of Mass Spectrometry*. 2015;377:410-20. doi: 10.1016/j.ijms.2014.06.034.
60. Marshall AG, Hendrickson CL, Jackson GS. Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass spectrometry reviews*. 1998;17(1):1-35. doi: 10.1002/(sici)1098-2787(1998)17:1<1::aid-mas1>3.0.co;2-k.
61. Huang N, Siegel MM, Kruppa GH, Laukien FH. Automation of a Fourier transform ion cyclotron resonance mass spectrometer for acquisition, analysis, and e-mailing of high-resolution exact-mass electrospray ionization mass spectral data. *Journal of the American Society for Mass Spectrometry*. 1999;10(11):1166-73. doi: 10.1016/s1044-0305(99)00089-6.
62. Marshall AG. Milestones in fourier transform ion cyclotron resonance mass spectrometry technique development. *International Journal of Mass Spectrometry*. 2000;200(1-3):331-56. doi: 10.1016/s1387-3806(00)00324-9.
63. Kim S, Rodgers RP, Marshall AG. Truly "exact" mass: Elemental composition can be determined uniquely from molecular mass measurement at ~0.1mDa accuracy for molecules up to ~500Da. *International Journal of Mass Spectrometry*. 2006;251(2-3):260-5. doi: 10.1016/j.ijms.2006.02.001.
64. Marshall AG, Hendrickson CL. High-resolution mass spectrometers. *Annual review of analytical chemistry*. 2008;1:579-99. doi: 10.1146/annurev.anchem.1.031207.112945.
65. Barrow MP, Burkitt WI, Derrick PJ. Principles of Fourier transform ion cyclotron resonance mass spectrometry and its application in structural biology. *The Analyst*. 2005;130(1):18-28. doi: 10.1039/b403880k.
66. Li X, Fekete A, Englmann M, Frommberger M, Lv S, Chen G, et al. At-line coupling of UPLC to chip-electrospray-FTICR-MS. *Analytical and bioanalytical chemistry*. 2007;389(5):1439-46. doi: 10.1007/s00216-007-1524-4.
67. Wolff MM, Stephens WE. A Pulsed Mass Spectrometer with Time Dispersion. *Review of Scientific Instruments*. 1953;24(8):616-7. doi: 10.1063/1.1770801.

68. Anonymous. Proceedings of the American Physical Society. *Physical Review*. 1946;69(11-12):674-. doi: 10.1103/PhysRev.69.674.
69. Marianna L. *Datamining metabolomics: the convergence point of non-target approach and statistical investigation*: Technische Universität München; 2008.
70. Liland KH. Multivariate methods in metabolomics – from pre-processing to dimension reduction and statistical analysis. *TrAC Trends in Analytical Chemistry*. 2011;30(6):827-41. doi: 10.1016/j.trac.2011.02.007.
71. Boccard J, Veuthey JL, Rudaz S. Knowledge discovery in metabolomics: an overview of MS data handling. *Journal of separation science*. 2010;33(3):290-304. doi: 10.1002/jssc.200900609.
72. Trygg J, Wold S. Orthogonal projections to latent structures (O-PLS). *J Chemometr*. 2002;16(3):119-28. doi: 10.1002/cem.695.
73. Abdi H. Partial least squares regression and projection on latent structure regression (PLS Regression). *Wiley Interdisciplinary Reviews: Computational Statistics*. 2010;2(1):97-106. doi: 10.1002/wics.51.
74. Tziotis D, Hertkorn N, Schmitt-Kopplin P. Kendrick-analogous network visualisation of ion cyclotron resonance Fourier transform mass spectra: improved options for the assignment of elemental compositions and the classification of organic molecular complexity. *Eur J Mass Spectrom (Chichester, Eng)*. 2011;17(4):415-21. doi: 10.1255/ejms.1135.
75. Wang Y, Lu X, Xu G. Simultaneous separation of hydrophilic and hydrophobic compounds by using an online HILIC-RPLC system with two detectors. *Journal of separation science*. 2008;31(9):1564-72. doi: 10.1002/jssc.200700663.
76. Barr DB, Wilder LC, Caudill SP, Gonzalez AJ, Needham LL, Pirkle JL. Urinary Creatinine Concentrations in the U.S. Population: Implications for Urinary Biologic Monitoring Measurements. *Environmental Health Perspectives*. 2004;113(2):192-200. doi: 10.1289/ehp.7337.
77. Henneges C, Bullinger D, Fux R, Friese N, Seeger H, Neubauer H, et al. Prediction of breast cancer by profiling of urinary RNA metabolites using Support Vector Machine-based feature selection. *BMC cancer*. 2009;9:104. doi: 10.1186/1471-2407-9-104.
78. Hartleb HEJ. Evaluation of sources of low and high creatinine concentrations in drug-screening urine samples. *Toxichem Krimtech*. 2011;78(378).

79. Cone EJ, Caplan YH, Moser F, Robert T, Shelby MK, Black DL. Normalization of urinary drug concentrations with specific gravity and creatinine. *Journal of analytical toxicology*. 2009;33(1):1-7.
80. Rigas ML, Okino MS, Quackenboss JJ. Use of a pharmacokinetic model to assess chlorpyrifos exposure and dose in children, based on urinary biomarker measurements. *Toxicological sciences : an official journal of the Society of Toxicology*. 2001;61(2):374-81.
81. Nicolli A, Chiara F, Gambalunga A, Carrieri M, Bartolucci GB, Trevisan A. Reliability of urinary excretion rate adjustment in measurements of hippuric acid in urine. *International journal of environmental research and public health*. 2014;11(7):7036-44. doi: 10.3390/ijerph110707036.
82. Babazono T, Takahashi C, Iwamoto Y. Definition of microalbuminuria in first-morning and random spot urine in diabetic patients. *Diabetes Care*. 2004;27(7):1838-9.
83. Warrack BM, Hnatyshyn S, Ott KH, Reily MD, Sanders M, Zhang H, et al. Normalization strategies for metabonomic analysis of urine samples. *Journal of chromatography B, Analytical technologies in the biomedical and life sciences*. 2009;877(5-6):547-52. doi: 10.1016/j.jchromb.2009.01.007.
84. Dieterle F, Ross A, Schlotterbeck G, Senn H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in ¹H NMR metabonomics. *Analytical chemistry*. 2006;78(13):4281-90. doi: 10.1021/ac051632c.
85. Veselkov KA, Vingara LK, Masson P, Robinette SL, Want E, Li JV, et al. Optimized preprocessing of ultra-performance liquid chromatography/mass spectrometry urinary metabolic profiles for improved information recovery. *Analytical chemistry*. 2011;83(15):5864-72. doi: 10.1021/ac201065j.
86. Popowski LA, Oppliger RA, Patrick Lambert G, Johnson RF, Kim Johnson A, Gisolf CV. Blood and urinary measures of hydration status during progressive acute dehydration. *Med Sci Sports Exerc*. 2001;33(5):747-53. doi: 10.1097/00005768-200105000-00011.
87. Riesenhuber A, Boehm M, Posch M, Aufricht C. Diuretic potential of energy drinks. *Amino acids*. 2006;31(1):81-3. doi: 10.1007/s00726-006-0363-5.
88. Oppliger RA, Magnes SA, Popowski LA, Gisolfi CV. Accuracy of urine specific gravity and osmolality as indicators of hydration status. *Int J Sport Nutr Exerc Metab*. 2005;15(3):236-51.
89. Miller RC, Brindle E, Holman DJ, Shofer J, Klein NA, Soules MR, et al. Comparison of specific gravity and creatinine for normalizing urinary reproductive hormone concentrations. *Clin Chem*. 2004;50(5):924-32. doi: 10.1373/clinchem.2004.032292.

90. Auray-Blais C, Boutin M, Gagnon R, Dupont FO, Lavoie P, Clarke JT. Urinary globotriaosylsphingosine-related biomarkers for Fabry disease targeted by metabolomics. *Analytical chemistry*. 2012;84(6):2745-53. doi: 10.1021/ac203433e.
91. Greenberg N, Roberts WL, Bachmann LM, Wright EC, Dalton RN, Zakowski JJ, et al. Specificity characteristics of 7 commercial creatinine measurement procedures by enzymatic and Jaffe method principles. *Clin Chem*. 2012;58(2):391-401. doi: 10.1373/clinchem.2011.172288.
92. Myers GL, Miller WG, Coresh J, Fleming J, Greenberg N, Greene T, et al. Recommendations for improving serum creatinine measurement: a report from the Laboratory Working Group of the National Kidney Disease Education Program. *Clin Chem*. 2006;52(1):5-18. doi: 10.1373/clinchem.2005.0525144.
93. Hoste L, Deiteren K, Pottel H, Callewaert N, Martens F. Routine serum creatinine measurements: how well do we perform? *BMC Nephrol*. 2015;16(1):21. doi: 10.1186/s12882-015-0012-x.
94. Witte EC, Lambers Heerspink HJ, de Zeeuw D, Bakker SJ, de Jong PE, Gansevoort R. First morning voids are more reliable than spot urine samples to assess microalbuminuria. *Journal of the American Society of Nephrology : JASN*. 2009;20(2):436-43. doi: 10.1681/ASN.2008030292.
95. Gansevoort RT, Verhave JC, Hillege HL, Burgerhof JG, Bakker SJ, de Zeeuw D, et al. The validity of screening based on spot morning urine samples to detect subjects with microalbuminuria in the general population. *Kidney Int Suppl*. 2005;67(94):28-35. doi: 10.1111/j.1523-1755.2005.09408.x.
96. Saydah SH, Pavkov ME, Zhang C, Lacher DA, Eberhardt MS, Burrows NR, et al. Albuminuria prevalence in first morning void compared with previous random urine from adults in the National Health and Nutrition Examination Survey, 2009-2010. *Clin Chem*. 2013;59(4):675-83. doi: 10.1373/clinchem.2012.195644.
97. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC bioinformatics*. 2011;12:77. doi: 10.1186/1471-2105-12-77.
98. Passing H, Bablok. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I. *J Clin Chem Clin Biochem*. 1983;21(11):709-20.
99. Passing H, Bablok W. Comparison of several regression procedures for method comparison studies and determination of sample sizes. Application of linear regression procedures for method comparison studies in Clinical Chemistry, Part II. *J Clin Chem Clin Biochem*. 1984;22(6):431-45.

100. Zou GY. Toward using confidence intervals to compare correlations. *Psychol Methods*. 2007;12(4):399-413. doi: 10.1037/1082-989X.12.4.399.
101. Bush DM. The U.S. Mandatory Guidelines for Federal Workplace Drug Testing Programs: current status and future considerations. *Forensic science international*. 2008;174(2-3):111-9. doi: 10.1016/j.forsciint.2007.03.008.
102. Carrieri M, Trevisan A, Bartolucci GB. Adjustment to concentration-dilution of spot urine samples: correlation between specific gravity and creatinine. *International Archives of Occupational and Environmental Health*. 2000;74(1):63-7. doi: 10.1007/s004200000190.
103. Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. *Mass spectrometry reviews*. 2007;26(1):51-78. doi: 10.1002/mas.20108.
104. Gika HG, Theodoridis GA, Plumb RS, Wilson ID. Current practice of liquid chromatography-mass spectrometry in metabolomics and metabonomics. *Journal of pharmaceutical and biomedical analysis*. 2014;87:12-25. doi: 10.1016/j.jpba.2013.06.032.
105. Jun-ichi Yamaguchi MO, Shigeji Jingu, Naoyoshi Ogawa, and Shohei Higuchi. Utility of postcolumn addition of 2-(2-methoxyethoxy)ethanol, a signal-enhancing modifier, for metabolite screening with liquid chromatography and negative ion electrospray ionization mass spectrometry. *Anal Chem*. 1999;71:5386-90.
106. Guo Y, Gaiki S. Retention behavior of small polar compounds on polar stationary phases in hydrophilic interaction chromatography. *Journal of Chromatography A*. 2005;1074(1-2):71-80. doi: 10.1016/j.chroma.2005.03.058.
107. Cech NB, Enke CG. Practical implications of some recent studies in electrospray ionization fundamentals. *Mass spectrometry reviews*. 2001;20(6):362-87. doi: 10.1002/mas.10008.
108. Cech NB, Enke CG. Effect of Affinity for Droplet Surfaces on the Fraction of Analyte Molecules Charged during Electrospray Droplet Fission. *Analytical chemistry*. 2001;73(19):4632-9. doi: 10.1021/ac001267j.
109. Wilm M. Principles of electrospray ionization. *Molecular & cellular proteomics : MCP*. 2011;10(7):M111 009407. doi: 10.1074/mcp.M111.009407.
110. Grimm RL, Beauchamp JL. Evaporation and discharge dynamics of highly charged multicomponent droplets generated by electrospray ionization. *The journal of physical chemistry A*. 2010;114(3):1411-9. doi: 10.1021/jp907162w.
111. Wortmann A, Kistler-Momotova A, Zenobi R, Heine MC, Wilhelm O, Pratsinis SE. Shrinking droplets in electrospray ionization and their influence on chemical equilibria. *Journal of the American Society for Mass Spectrometry*. 2007;18(3):385-93. doi: 10.1016/j.jasms.2006.10.010.

112. Ekdahl A, Johansson MC, Ahnoff M. Tracing and separating plasma components causing matrix effects in hydrophilic interaction chromatography-electrospray ionization mass spectrometry. *Journal of chromatography B, Analytical technologies in the biomedical and life sciences*. 2013;923-924C:83-91. doi: 10.1016/j.jchromb.2013.02.013.
113. Polson C, Sarkar P, Incledon B, Raguvaran V, Grant R. Optimization of protein precipitation based upon effectiveness of protein removal and ionization effect in liquid chromatography–tandem mass spectrometry. *Journal of Chromatography B*. 2003;785(2):263-75. doi: 10.1016/s1570-0232(02)00914-5.
114. Matsuda M, DeFronzo RA. Insulin sensitivity indices obtained from oral glucose tolerance testing - Comparison with the euglycemic insulin clamp. *Diabetes Care*. 1999;22(9):1462-70. doi: DOI 10.2337/diacare.22.9.1462.
115. Lucio M, Fekete A, Weigert C, Wagele B, Zhao X, Chen J, et al. Insulin sensitivity is reflected by characteristic metabolic fingerprints--a Fourier transform mass spectrometric non-targeted metabolomics approach. *PloS one*. 2010;5(10):e13317. doi: 10.1371/journal.pone.0013317.
116. Adrian M, Lucio M, Roullier-Gall C, Heloir MC, Trouvelot S, Daire X, et al. Metabolic Fingerprint of PS3-Induced Resistance of Grapevine Leaves against *Plasmopara viticola* Revealed Differences in Elicitor-Triggered Defenses. *Front Plant Sci*. 2017;8:101. doi: 10.3389/fpls.2017.00101.
117. Wang S-Y, Kuo C-H, Tseng YJ. Batch Normalizer: A Fast Total Abundance Regression Calibration Method to Simultaneously Adjust Batch and Injection Order Effects in Liquid Chromatography/Time-of-Flight Mass Spectrometry-Based Metabolomics Data and Comparison with Current Calibration Methods. *Analytical chemistry*. 2013;85(2):1037-46. doi: 10.1021/ac302877x.
118. Wikoff WR, Anfora AT, Liu J, Schultz PG, Lesley SA, Peters EC, et al. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proceedings of the National Academy of Sciences of the United States of America*. 2009;106(10):3698-703. doi: 10.1073/pnas.0812874106.
119. Men L, Pi Z, Zhou Y, Wei M, Liu Y, Song F, et al. Urine metabolomics of high-fat diet induced obesity using UHPLC-Q-TOF-MS. *Journal of pharmaceutical and biomedical analysis*. 2017;132:258-66. doi: 10.1016/j.jpba.2016.10.012.
120. Iavarone AT, Williams ER. Supercharging in electrospray ionization: effects on signal and charge. *International Journal of Mass Spectrometry*. 2002;219(1):63-72. doi: 10.1016/s1387-3806(02)00587-0.
121. Musso G, Gambino R, Cassader M. Obesity, diabetes, and gut microbiota: the hygiene hypothesis expanded? *Diabetes Care*. 2010;33(10):2277-84. doi: 10.2337/dc10-0556.

122. Li JV, Ashrafian H, Bueter M, Kinross J, Sands C, le Roux CW, et al. Metabolic surgery profoundly influences gut microbial-host metabolic cross-talk. *Gut*. 2011;60(9):1214-23. doi: 10.1136/gut.2010.234708.
123. Lustgarten MS, Price LL, Chale A, Fielding RA. Metabolites related to gut bacterial metabolism, peroxisome proliferator-activated receptor-alpha activation, and insulin sensitivity are associated with physical function in functionally-limited older adults. *Aging Cell*. 2014;13(5):918-25. doi: 10.1111/accel.12251.
124. Dong S, Zhan ZY, Cao HY, Wu C, Bian YQ, Li JY, et al. Urinary metabolomics analysis identifies key biomarkers of different stages of nonalcoholic fatty liver disease. *World journal of gastroenterology : WJG*. 2017;23(15):2771-84. doi: 10.3748/wjg.v23.i15.2771.
125. Baerts L, Glorie L, Maho W, Eelen A, Verhulst A, D'Haese P, et al. Potential impact of sitagliptin on collagen-derived dipeptides in diabetic osteoporosis. *Pharmacol Res*. 2015;100:336-40. doi: 10.1016/j.phrs.2015.08.023.
126. Wu Q, Lai X, Zhu Z, Hong Z, Dong X, Wang T, et al. Evidence for Chronic Kidney Disease-Mineral and Bone Disorder Associated With Metabolic Pathway Changes. *Medicine (Baltimore)*. 2015;94(32):e1273. doi: 10.1097/MD.0000000000001273.
127. Chen H, Li X, Yue R, Ren X, Zhang X, Ni A. The effects of diabetes mellitus and diabetic nephropathy on bone and mineral metabolism in T2DM patients. *Diabetes Res Clin Pract*. 2013;100(2):272-6. doi: 10.1016/j.diabres.2013.03.007.
128. Hu YF, Chirila C, Alexander D, Milburn M, Mitchell MW, Gall W, et al. Biomarkers Related to Insulin Resistance and Methods using the Same. Google Patents; 2012.
129. Kawamoto R, Tabara Y, Kohara K, Miki T, Ohtsuka N, Kusunoki T, et al. Serum gamma-glutamyl transferase within its normal concentration range is related to the presence of impaired fasting glucose and diabetes among Japanese community-dwelling persons. *Endocr Res*. 2011;36(2):64-73. doi: 10.3109/07435800.2010.534756.
130. Mason JE, Starke RD, Van Kirk JE. Gamma-glutamyl transferase: a novel cardiovascular risk biomarker. *Prev Cardiol*. 2010;13(1):36-41. doi: 10.1111/j.1751-7141.2009.00054.x.
131. Anuradha CV, Balakrishnan SD. Taurine attenuates hypertension and improves insulin sensitivity in the fructose-fed rat, an animal model of insulin resistance. *Can J Physiol Pharmacol*. 1999;77(10):749-54.
132. Nandhini AT, Thirunavukkarasu V, Anuradha CV. Taurine modifies insulin signaling enzymes in the fructose-fed insulin resistant rats. *Diabetes Metab*. 2005;31(4 Pt 1):337-44.
133. Schaffer SW, Azuma J, Mozaffari M. Role of antioxidant activity of taurine in diabetes. *Can J Physiol Pharmacol*. 2009;87(2):91-9. doi: 10.1139/Y08-110.

134. Zheng Y, Ceglarek U, Huang T, Wang T, Heianza Y, Ma W, et al. Plasma Taurine, Diabetes Genetic Predisposition, and Changes of Insulin Sensitivity in Response to Weight-Loss Diets. *J Clin Endocrinol Metab.* 2016;101(10):3820-6. doi: 10.1210/jc.2016-1760.
135. Ribeiro RA, Bonfleur ML, Amaral AG, Vanzela EC, Rocco SA, Boschero AC, et al. Taurine supplementation enhances nutrient-induced insulin secretion in pancreatic mice islets. *Diabetes Metab Res Rev.* 2009;25(4):370-9. doi: 10.1002/dmrr.959.
136. Nakaya Y, Minami A, Harada N, Sakamoto S, Niwa Y, Ohnaka M. Taurine improves insulin sensitivity in the Otsuka Long-Evans Tokushima Fatty rat, a model of spontaneous type 2 diabetes. *The American journal of clinical nutrition.* 2000;71(1):54-8.
137. Brons C, Spohr C, Storgaard H, Dyerberg J, Vaag A. Effect of taurine treatment on insulin secretion and action, and on serum lipid levels in overweight men with a genetic predisposition for type II diabetes mellitus. *Eur J Clin Nutr.* 2004;58(9):1239-47. doi: 10.1038/sj.ejcn.1601955.
138. He M, Rutledge SL, Kelly DR, Palmer CA, Murdoch G, Majumder N, et al. A new genetic disorder in mitochondrial fatty acid beta-oxidation: ACAD9 deficiency. *Am J Hum Genet.* 2007;81(1):87-103. doi: 10.1086/519219.
139. Truscott RJW, Hick L, Pullin C, Halpern B, Wilcken B, Griffiths H, et al. Dicarboxylic aciduria: The response to fasting. *Clinica Chimica Acta.* 1979;94(1):31-9. doi: 10.1016/0009-8981(79)90183-9.
140. Muth A, Jung J, Bilke S, Scharrer A, Mosandl A, Sewell AC, et al. Simultaneous enantioselective analysis of chiral urinary metabolites in patients with Zellweger syndrome. *Journal of Chromatography B.* 2003;792(2):269-77. doi: 10.1016/s1570-0232(03)00285-x.
141. Bennett MJ, Weinberger MJ, Kobori JA, Rinaldo P, Burlina AB. Mitochondrial Short-Chain L-3-Hydroxyacyl-Coenzyme A Dehydrogenase Deficiency: A New Defect of Fatty Acid Oxidation. *Pediatric Research.* 1996;39(1):185-8. doi: 10.1203/00006450-199601000-00031.
142. Clarke BL, Khosla S. Androgens and bone. *Steroids.* 2009;74(3):296-305. doi: 10.1016/j.steroids.2008.10.003.
143. Yu ZR, Ning Y, Yu H, Tang NJ. A HPLC-Q-TOF-MS-based urinary metabolomic approach to identification of potential biomarkers of metabolic syndrome. *Journal of Huazhong University of Science and Technology Medical sciences = Hua zhong ke ji da xue xue bao Yi xue Ying De wen ban = Huazhong keji daxue xuebao Yixue Yingdewen ban.* 2014;34(2):276-83. doi: DOI 10.1007/s11596-014-1271-7.
144. Medina S, Ferreres F, Garcia-Viguera C, Horcajada MN, Orduna J, Saviron M, et al. Non-targeted metabolomic approach reveals urinary metabolites linked to steroid biosynthesis pathway after ingestion of citrus juice. *Food chemistry.* 2013;136(2):938-46. doi: 10.1016/j.foodchem.2012.09.004.

145. Marks P, Wilson B, Delassalle A. Aldosterone Studies in Obese Patients with Hypertension. *The American Journal of the Medical Sciences*. 1985;289(6):224-8. doi: 10.1097/00000441-198506000-00003.
146. Rago D, Gürdeniz G, Ravn-Haren G, Dragsted LO. An explorative study of the effect of apple and apple products on the human plasma metabolome investigated by LC–MS profiling. *Metabolomics : Official journal of the Metabolomic Society*. 2014;11(1):27-39. doi: 10.1007/s11306-014-0668-4.
147. Jeanneret F, Boccard J, Badoud F, Sorg O, Tonoli D, Pelclova D, et al. Human urinary biomarkers of dioxin exposure: analysis by metabolomics and biologically driven data dimensionality reduction. *Toxicol Lett*. 2014;230(2):234-43. doi: 10.1016/j.toxlet.2013.10.031.
148. Sun M, Gao X, Zhang D, Ke C, Hou Y, Fan L, et al. Identification of biomarkers for unstable angina by plasma metabolomic profiling. *Molecular bioSystems*. 2013;9(12):3059-67. doi: 10.1039/c3mb70216b.
149. Galderisi A, Pirillo P, Moret V, Stocchero M, Gucciardi A, Perilongo G, et al. Metabolomics reveals new metabolic perturbations in children with type 1 diabetes. *Pediatr Diabetes*. 2017. doi: 10.1111/pedi.12524.
150. Musani SK, Vasan RS, Bidulescu A, Liu J, Xanthakis V, Sims M, et al. Aldosterone, C-reactive protein, and plasma B-type natriuretic peptide are associated with the development of metabolic syndrome and longitudinal changes in metabolic syndrome components: findings from the Jackson Heart Study. *Diabetes Care*. 2013;36(10):3084-92. doi: 10.2337/dc12-2562.
151. Baudrand R, Campino C, Carvajal CA, Olivieri O, Guidi G, Faccini G, et al. Increased urinary glucocorticoid metabolites are associated with metabolic syndrome, hypoadiponectinemia, insulin resistance and beta cell dysfunction. *Steroids*. 2011;76(14):1575-81. doi: 10.1016/j.steroids.2011.09.010.
152. Giordano R, Guaraldi F, Berardelli R, Karamouzis I, D'Angelo V, Marinazzo E, et al. Glucose metabolism in patients with subclinical Cushing's syndrome. *Endocrine*. 2012;41(3):415-23. doi: 10.1007/s12020-012-9628-9.
153. Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, Lien LF, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell metabolism*. 2009;9(4):311-26. doi: 10.1016/j.cmet.2009.02.002.
154. Ho JE, Larson MG, Vasan RS, Ghorbani A, Cheng S, Rhee EP, et al. Metabolite profiles during oral glucose challenge. *Diabetes*. 2013;62(8):2689-98. doi: 10.2337/db12-0754.

155. Walford GA, Ma Y, Clish C, Florez JC, Wang TJ, Gerszten RE, et al. Metabolite Profiles of Diabetes Incidence and Intervention Response in the Diabetes Prevention Program. *Diabetes*. 2016;65(5):1424-33. doi: 10.2337/db15-1063.
156. Sookoian S, Pirola CJ. Alanine and aspartate aminotransferase and glutamine-cycling pathway: their roles in pathogenesis of metabolic syndrome. *World journal of gastroenterology : WJG*. 2012;18(29):3775-81. doi: 10.3748/wjg.v18.i29.3775.
157. Leonardi R, Zhang YM, Rock CO, Jackowski S. Coenzyme A: back in action. *Prog Lipid Res*. 2005;44(2-3):125-53. doi: 10.1016/j.plipres.2005.04.001.
158. Tahiliani AG, Beinlich CJ. Pantothenic Acid in Health and Disease. 1991;46:165-228. doi: 10.1016/s0083-6729(08)60684-6.
159. Zhyvoloup A, Nemazanyy I, Babich A, Panasyuk G, Pobigailo N, Vudmaska M, et al. Molecular cloning of CoA Synthase. The missing link in CoA biosynthesis. *J Biol Chem*. 2002;277(25):22107-10. doi: 10.1074/jbc.C200195200.
160. Mooney S, Leuendorf JE, Hendrickson C, Hellmann H. Vitamin B6: a long known compound of surprising complexity. *Molecules*. 2009;14(1):329-51. doi: 10.3390/molecules14010329.
161. Gregory JF, 3rd, Park Y, Lamers Y, Bandyopadhyay N, Chi YY, Lee K, et al. Metabolomic analysis reveals extended metabolic consequences of marginal vitamin B-6 deficiency in healthy human subjects. *PloS one*. 2013;8(6):e63544. doi: 10.1371/journal.pone.0063544.
162. Muellenbach EA, Diehl CJ, Teachey MK, Lindborg KA, Archuleta TL, Harrell NB, et al. Interactions of the advanced glycation end product inhibitor pyridoxamine and the antioxidant alpha-lipoic acid on insulin resistance in the obese Zucker rat. *Metabolism*. 2008;57(10):1465-72. doi: 10.1016/j.metabol.2008.05.018.
163. Tully DB, Allgood VE, Cidlowski JA. Modulation of steroid receptor-mediated gene expression by vitamin B6. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*. 1994;8(3):343-9.
164. Allgood VE, Powell-Oliver FE, Cidlowski JA. Vitamin B6 influences glucocorticoid receptor-dependent gene expression. *J Biol Chem*. 1990;265(21):12424-33.
165. Maury CP, Wegelius O. Urinary sialyloligosaccharide excretion as an indicator of disease activity in rheumatoid arthritis. *Rheumatol Int*. 1981;1(1):7-10.
166. Crook MA, Tutt P, Simpson H, Pickup JC. Serum sialic acid and acute phase proteins in type 1 and type 2 diabetes mellitus. *Clinica chimica acta; international journal of clinical chemistry*. 1993;219(1-2):131-8.

167. Rahman IU, Malik SA, Bashir M, Khan RU, Idrees M. Serum sialic acid changes in type 2 diabetic patients on metformin or rosiglitazone treatment. *J Clin Pharm Ther.* 2010;35(6):685-90. doi: 10.1111/j.1365-2710.2009.01145.x.
168. Gopinathan Nair B, Beevi Osman F. A Study of Serum Sialic Acid Levels in Patients with Type 2 Diabetes Mellitus and Its Relation to Lipid Parameters. *Journal of Evolution of Medical and Dental Sciences.* 2017;6(09):668-72. doi: 10.14260/Jemds/2017/144.
169. Manna P, Jain SK. Phosphatidylinositol-3,4,5-triphosphate and cellular signaling: implications for obesity and diabetes. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology.* 2015;35(4):1253-75. doi: 10.1159/000373949.
170. Corda D, Zizza P, Varone A, Filippi BM, Mariggio S. The glycerophosphoinositols: cellular metabolism and biological functions. *Cellular and molecular life sciences : CMLS.* 2009;66(21):3449-67. doi: 10.1007/s00018-009-0113-4.
171. Sun CQ, Zhong CY, Sun WW, Xiao H, Zhu P, Lin YZ, et al. Elevated Type II Secretory Phospholipase A2 Increases the Risk of Early Atherosclerosis in Patients with Newly Diagnosed Metabolic Syndrome. *Scientific reports.* 2016;6:34929. doi: 10.1038/srep34929.
172. Berrie CP, Iurisci C, Corda D. Membrane transport and in vitro metabolism of the Ras cascade messenger, glycerophosphoinositol 4-phosphate. *European journal of biochemistry.* 1999;266(2):413-9.
173. Abdella N, Akanji AO, Mojiminiyi OA, Al Assoussi A, Moussa M. Relation of serum total sialic acid concentrations with diabetic complications and cardiovascular risk factors in Kuwaiti Type 2 diabetic patients. *Diabetes Res Clin Pract.* 2000;50(1):65-72.
174. Ma ZA, Zhao Z, Turk J. Mitochondrial dysfunction and beta-cell failure in type 2 diabetes mellitus. *Experimental diabetes research.* 2012;2012:703538. doi: 10.1155/2012/703538.
175. Yan LJ. Pathogenesis of chronic hyperglycemia: from reductive stress to oxidative stress. *Journal of diabetes research.* 2014;2014:137919. doi: 10.1155/2014/137919.

Curriculum Vitae

Wendelin Kranz née Koch

Personal information

Place and date of birth	06.11.1983, Stuttgart-Bad Cannstatt
Marital status	married
Nationality	German

Work experience

Apr 2018 - now	Scientist LC-MS at Coriolis pharma research GmbH
Apr 2016 – Mar 2018	Postdoctoral scientist at Coriolis pharma research GmbH: characterization of polysorbates and their degradation products in biopharmaceutical formulations.
Aug 2012 – Dec 2015	PhD student at Helmholtz Zentrum München. Department: Analytical BioGeochemistry (BGC). Dissertation: "Urine <i>discovery metabolomics</i> of insulin resistant individuals – an holistic approach utilizing hydrophilic interaction liquid chromatography and mass spectrometry".
Mar 2009 – Jul 2012	Analytical Chemist at Ascendis Pharma GmbH Heidelberg. Occupation: Development of new analytical methods for preclinical drug candidates, identification of impurities, responsibility for (U)HPLC-MS systems, induction training of new employees.
Sep 2008 - Dec 2008	Engineer at Merck KGaA Darmstadt. Department: CAS/AC10 process analytics

Education

- Mar 2007 - Aug 2008 Degree course Master Bio- und Process analytics at Reutlingen University with degree Master Of Science.
- Oct 2003 - Feb 2007 Degree course Bachelor Chemistry with Marketing (area of specialization: Instrumental analytics) at Reutlingen University with degree Bachelor Of Science.
- Sep 2000 - Jul 2003 Commercial high school (Wirtschaftsgymnasium) Göppingen with degree A levels

Publications

Koch W, Forcisi S, Lehmann R, Schmitt-Kopplin P: Sensitivity improvement in hydrophilic interaction chromatography negative mode electrospray ionization mass spectrometry using 2-(2-methoxyethoxy)ethanol as a post-column modifier for non-targeted metabolomics *Journal/J Chromatogr A*, 2014, 1361, 209-16

Sampsonidis I, Witting M, Koch W, Virgiliou C, Gika H G, Schmitt-Kopplin P, Theodoridis G A: Computational analysis and ratiometric comparison approaches aimed to assist column selection in hydrophilic interaction liquid chromatography-tandem mass spectrometry targeted metabolomics *Journal/J Chromatogr A*, 2015, 1406, 145-55

Martos A, Koch W, Jiskoot W, Wuchner K, Winter G, Friess W, Hawe A: Trends on Analytical Characterization of Polysorbates and Their Degradation Products in Biopharmaceutical Formulations *Journal/J Pharm Sci*, 2017, 106, 1722-1735

Sillner N, Walker A, Koch W, Witting M, Schmitt-Kopplin P. Metformin impacts cecal bile acid profiles in mice. *Journal of chromatography B, Analytical technologies in the biomedical and life sciences*. 2018;1083:35-43.

Kranz W, Wuchner K, Corradini E, Berger M, Hawe A. Factors influencing polysorbate's sensitivity against enzymatic hydrolysis and oxidative degradation. *Journal of pharmaceutical sciences*. (article in press)

