The use of metabolomics for studying the effects of pre- and probiotics

Katrine Overgaard Poulsen
Title:
The use of metabolomics for studying the effects of pre- and probiotics

Danish title:
Brugen af metabolomics til at undersøge effekterne af præ- og probiotika

Purpose of the thesis (copied from bachelor’s contract):
The worldwide obesity epidemic and changes in lifestyle to less physical activity have given rise to growing numbers of people suffering from metabolic syndrome. Pre- and probiotics have been found to have potential health benefits after intake, especially in relation to obesity, diabetes and cardiovascular diseases. Based on literature, the potential health benefits associated with pre- and probiotics for human health and the role of metabolomics to investigate the effect of pre- and probiotics will be described.

Author:
Katrine Overgaard Poulsen, 201407648

Institution:
Aarhus University, Department of Agrobiology

Main supervisor:
Ulrik Kræmer Sundekilde*, Postdoc

Co-supervisor:
Jette F. Young*, Associate professor

* Aarhus University, Department of Food Science - Differentiated & Biofunctional Food

Total number of pages: 34.8 pp. (2400 counts per page, text only)
I Abstract

Pre- and probiotics have proven to modulate the colonic microbiota and activity, and to induce health effects in the host regarding metabolic syndrome, gut diseases and improvements of immune system function in vitro and in vivo. However, interactions with the host are complex and the biochemical effects on host’s metabolism remain unclear. In order to elucidate the actions of pre- and probiotics, the technique metabolomics is emerging. Thus, this thesis includes a pilot study were two participants consumed one commercial probiotic drink á day for three weeks followed by a wash-out period, and collected six urine samples individually. The metabolic profile of the urine samples were investigated by $^1$H nuclear magnetic resonance (NMR) spectroscopy based metabolomics. By the use of multivariate statistical analysis (principal components analysis) changes in urinary metabolites of the samples were examined. The study proved great inter- and intra-individual differences especially due to variations in creatinine excretion, but no clear relation to probiotic intake was found. Literature on the matter demonstrates indications of relations between pre- and/or probiotic supplementations and changes in the metabolic profile of a biological source, albeit it is mostly based on animal studies. Thus, metabolomics offers potential in exploring effects of pre- and probiotics on biochemical pathways and host’s metabolism, but there is a need of greater human intervention studies.

II Sammendrag

# Table of content

I Abstract 3
II Sammendrag 3

1 Introduction 5
   a) The human microflora 5
      a.i) Microbial activity of the human gut flora and the end-products produced 5
   b) The definition of pre- and prebiotics 7
      b. i) Introduction to potential beneficial effects of pre- and probiotics 8
   c) Exploring the human metabolome 11
      c. i) Metabolomics 12
         Chemical analysis used in metabolomics 12
         Basic principles of nuclear magnetic resonance spectroscopy (NMR) in regards to metabolomics 13
         Chemometrics 17
      c. ii) Metabolic composition of urine 18

2 Materials and methods 22
   a) Pilot study and sample collection 22
   b) Sample preparation and data acquisition by $^1$H NMR 22
   c) Identification of metabolites and multivariate analysis 23

3 Results 25
   a) $^1$H NMR spectra and some identified metabolites 25
   b) Inter-individual differences reflected by PCA plots 26
   c) Intra-individual differences reflected by PCA plots 27
      c.i) Intra-individual differences observed in subject U 27
      c.ii) Intra-individual differences observed in subject K 29

4 Discussion 32
   a) Characteristics of urine profiles 32
   b) Associations between observed discriminators and probiotic intake 34
   c) Methodological considerations 41
      c.i) Some considerations regarding NMR vs. MS 41
      c.ii) Some considerations regarding study design 42

5 Conclusion and perspectives 43

6 Acknowledgements 43

7 Literature 44

Appendix 54
1 Introduction

Studies indicating that certain species of the colonic microbiota influence the host’s health have led to the concept of probiotics and later prebiotics. Intake of pre- and probiotics is an attempt to modify the composition of the gut flora, and has shown to modulate immune response, improvements of the gut health and to exert positive effects on symptoms associated with metabolic syndrome (Ebel et al., 2014). However, the exact mechanisms leading to modifications of the host’s metabolism after intake remain elusive, but they are generally considered to be due to changes in microbial activity and metabolites thereof.

Metabolomics is “a comprehensive analysis in which all the metabolites of a biological system are identified and quantified” (Lindon & Nicholson, 2008) and offers a holistic analysis of biological samples such as tissues and urine. Hence, by the use of metabolomics, we can identify and quantify metabolites in a biological system, and by multivariate analysis explore changes induced by consuming pre- and/or probiotics.

a) The human microflora

The human microflora (the microbiome) consists of more than 400 different species and more than $10^{14}$ bacteria cells, who inhabit the gastro-intestinal system and exist in a symbiotic relationship with the host (Eckburg et al., 2005; Scott et al., 2008).

The quantitatively greatest bacterial species are of the phyla Bacteroidetes and Firmicutes (Eckburg et al., 2005), including the genera Prevotella, Eubacteria, Lactobacillus, Ruminococcus and Clostridium as well as the genera Bifidobacterium of Actinobacteria (Power et al., 2013; Tremaroli & Bäckhed, 2012). General categorisation of the human gut flora has been attempted and gave rise to different enterotypes predominated by Bacteroides, Prevotella and Ruminococcus (Arumugam et al., 2011). However, this categorisation is argued to be less discrete (Tremaroli & Bäckhed, 2012).

The composition of the gut flora is established early in life during the neonate period and is greatly determined by birth mode of delivery (i.e. vaginal or cesarean) (Murgas et al., 2011; Ravel et al., 2011), but can be remodelled by environmental factors such as lifestyle, diet and diseases. This affect homeostasis and thus the composition of gut flora through complex interactions (Nicholson et al., 2012). A distortion in the composition of the gut flora is referred to as dysbiosis and is a partial underlying mechanism in the development of metabolic diseases, inflammatory bowel diseases (IBDs) and colon cancer (Bringiotti et al., 2014; Turroni et al., 2014).

a.i) Microbial activity of the human gut flora and the end-products produced

The majority of bacteria in the proximal colon are anaerobic saccharolytic and produce hydrolytic enzymes (e.g. glycosidases, polysaccharidases) in order to degrade and ferment non-digestible carbohydrates that enter the colon (Hamer et al., 2008; Macfarlane & Macfarlane, 2012).
A variety of intermediates and end-products including CO$_2$, H$_2$ and organic acids such as lactate, succinate, acetate, propionate and butyrate are produced in the gastrointestinal system (Cummings & Macfarlane, 1997). The latter three comprise the quantitatively most prominent short-chain fatty acids produced (SCFAs) (Manas et al., n.d.).

Nitrogen constituents derived from protein typically pass through the proximal colon, since carbohydrates are the preferred source of energy (Preter et al., 2011). Hence, the majority of nitrogen compounds enter the distal colon for proteolytic fermentation, and are degraded to smaller amounts of SCFAs, branched-chain fatty acids, ammonia, hydrogen sulphide, indole and volatile amines dependent on substrate and energy availability (Cummings & Macfarlane, 1997; Hamer et al., 2008). Methanogens and other bacteria growing on intermediate products are also present in the colon (Gibson et al., 1993).

Derivatives not being assimilated by bacteria in the colon are absorbed by the colonocytes across the apical membrane (Verbeke et al., 2015). Considering the SCFAs produced, butyrate is predominantly used by the colonocytes for energy, and therefore not further transported across the basolateral membrane. Acetate and propionate are transported to the blood stream and liver (Preter et al., 2011). Resultantly, the human gut flora secures a pathway of energy salvage for the host (Tremaroli & Bäckhed, 2012).

However, some compounds produced by the gut flora are potentially toxic. Here among are phenols and indoles from dissimilatory metabolism of aromatic amino acids in the gut. These are detoxified by glucoronide or sulphate conjugation in the liver and excreted in the urine. Ammonia from deamination of amino acids is likewise excreted as urea (Cummings & Macfarlane, 1991; Smith & Macfarlane, 1996). Furthermore, some bacteria exert enzyme activities such as azoreducatase, nitroreductase, urease and β-glucorinase, which have been found to be associated with the conversion of pro-carcinogenic compounds to carcinogenic compounds (Preter et al., 2008; Ouwehand, et al., 2002).

The aforementioned divergence between the microbial activity in the proximal colon and the distal colon is partly due to differences in the energy available and the substrate entering the colon. This affect the distribution and composition of bacteria and the end-products produced (Cummings & Macfarlane, 1991). A greater availability of non-digestible carbohydrates in the proximal colon leads to greater production of SCFAs, which lowers pH in the colon (Cummings et al., 1987).

A reduction in pH changes the microbial activity from decreased proteolytic activity to increased saccharolytic activity, thus preventing production of potentially toxic derivatives of amino acid fermentation (Smith & Macfarlane, 1996).

The effect of SCFAs as bioactive molecules in regards to human biology has been widely studied.
A review by Cummings and Macfarlane (1997) emphasizes how SCFAs (especially butyrate) affect transport across the gut epithelium layer, the metabolism of the colonocytes, growth and differentiation of cells, hepatic control of lipid and carbohydrates, and finally energy salvage among other potential effects. Moreover, it is assumed that the symbiosis between microbiota and humans influences the host’s immune system and disease development (Macfarlane & Macfarlane, 2012; Martin et al., 2012). The identification that certain microorganisms have a positive impact on human health has led to the concept of pro- and prebiotics.

b) The definition of pro- and prebiotics

The human body is as before mentioned host for a variety of microorganisms, but not all exert a probiotic effect. The definition of probiotics has been redefined since its introduction the 1900s (Metchnikoff, 1908). Today, the most common definition has been established by Guarner & Schaaafsma (1998), who define probiotics as "live microorganisms, which upon ingestion in certain numbers, exert health effect on the host". They must survive passing the human gastrointestinal system in order to fulfil their potential beneficial effect in the host (FAO (Food and Agriculture Organization) & WHO, 2006).

According to Shah (2007), 56 species of the genus Lactobacillus have been identified as being probiotic, with Lactobacillus acidophilus being the most prevalent. Of the genus Bifidobacterium, 29 species are identified, here among are Bifidobacterium adolescentis, B. infantis and B. breve (Fijan, 2014; Shah, 2007). Species of Escherichia Coli, Streptococcus, Saccharomyces, Enterococcus, Leuconostoc and Pediococcus have also been identified to have probiotic effect (Fijan, 2014; Ohland & Macnaughton, 2010). It should be emphasised that probiotic genera are commonly present in the human gastrointestinal system despite lack of dietary supplements (Fontana et al., 2013). Additionally, other authors stress that the probiotics do not have to colonize the gastrointestinal tract in order to exert a beneficial effect. Non-colonizing strains thus exert effects as they pass through the colon by remodelling the gut flora (Ohland & Macnaughton, 2010). Probiotic species express β-galactosidase, fructofuranosidase and β-glucosidase activities among others, thus ferment carbohydrates to particularly lactic acid and acetate, which lowers pH in the colon and reduces proteolytic activity (Preter et al., 2008; Tabasco et al., 2014). Lactate and acetate can subsequently be converted into butyrate by other microbial species in the gut (Rajkumar et al., 2015). However, the expressions of the aforementioned enzyme activities are species and strain specific (Tabasco et al., 2014).

Consumption of certain foodstuff can selectively influence the composition of bacteria in the colon and induce the growth of these, thus leading to the concept of prebiotics (Gibson & Roberfroid, 1995).
The definition of prebiotics has been widely revisited since its introduction in 1995 (Gibson & Roberfroid, 1995). The currently accepted definition of dietary prebiotics is "a selectively fermented ingredient that results in specific changes, both in the composition and/or activity in the gastrointestinal microflora thus conferring benefits upon host health" (Gibson et al., 2010).

Prebiotics provide an attentive manipulation of the gut flora, and denote foodstuff aimed at inducing the growth of especially strains of Lactobacillus spp. and Bifidobacterium spp. in the gastrointestinal system by stimulating saccharolytic activity (Gibson & Roberfroid, 2004). They must be resistant towards hydrolysis and absorption in the upper part of the gastrointestinal system in order to exert beneficial effects. Intake of inulin-type fructans and galacto-oligosaccharides have been demonstrated to improve growth of bifidobacterium and lactobacilli in multiple studies (Kaplan & Hutkins, 2000; Roberfroid et al., 2010). A review by Gibson & Roberfroid (2004) proposes that prebiotics are non-digestible carbohydrates such as fructooligosaccharides (FOS), galacto-oligosaccharides (GOS) and lactulose.

b. i) Introduction to potential beneficial effects of pre- and probiotics

Today, various commercial probiotic products are available (Tripathi & Giri, 2014). In order to be ascribed a health claim, probiotic supplementations must provide evidence of health benefits in vitro and from animal and human studies in vivo (FAO (Food and Agriculture Organization) & WHO, 2006). Several review articles gather some demonstrated health benefits of pre- and probiotics. Here among are Preter et al. (2011), who distinguish between the effect of probiotic intake on modification of the host’s immune reactivity, and interference with colonic metabolism respectively. This is demonstrated by in vitro and in vivo studies testing the effects of selected bacterial strains and validating how certain strains are able to adhere to the mucosa and colonize the gastro-intestinal tract, resulting in a changed microbial activity towards increased saccharolytic activity (Preter et al., 2011).

Shah (2007) has systematically reviewed the effect of colonization by probiotic strains in the colon and assesses some health effects of lactobacilli and bifidobacteria. These effects include an induced immune response, anti-mutagenic activity, improvements of bowel diseases and contribution to enhanced epithelial cellular barrier growth (Ewaschuk et al., 2008; Shah, 2007).

Immunological and anti-mutagenic effects are inter alia provided by the ability of some strains to increase IgA antibody levels and to reduce nitroreductase, azoreductase and β-glucuronidase activity, the latter three enzymes involved in activation of mutagens (Goldin & Gorbach, 1984; Shah, 2007).

Additionally, multiple reviews (Arora et al., 2013, Gibson & Roberfroid, 1995; Preter et al., 2011; Shah, 2007) highlight potential beneficial effects of Bifidobacteria and Lactobacillus by reducing
blood level ammonia, lowering blood cholesterol, improvement of serum lipid profile and to inhibit the growth of pathogens.

Based on in vitro studies, the cholesterol lowering effect is assumed to be either due to 1) inhibition of dietary cholesterol absorption from the small intestine by binding of bile acids, 2) assimilation of cholesterol by lactic acid bacteria or 3) excretion of bile acids after binding of them (Pereira & Gibson, 2002).

Table 1 summarizes selected studies demonstrating health benefits after administration of pre- and/or probiotics. Search criteria for table 1 were results of in vivo human studies.

The search objective was aimed at 3 categories: obesity and serum lipids, metabolism and disease markers (markers being related to inflammatory bowel’s disease (IBD), lactose intolerance, immune system) and alterations in gut flora composition owing to the statements of the aforementioned reviews.

<table>
<thead>
<tr>
<th>Health marker</th>
<th>Genus or prebiotic</th>
<th>Species (strain)</th>
<th>Duration and administration</th>
<th>Study design</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obesity and cholesterol level</td>
<td>Bifidobacterium</td>
<td>lactis (HN019)</td>
<td>FM 80mL/day 45 d</td>
<td>R, PC n = 51 (metabolic syndrome symptoms)</td>
<td>Bernini et al. (2016)</td>
</tr>
<tr>
<td>Reduced total cholesterol, reduced body-mass index, reduced LDL-lipoprotein</td>
<td>Lactobacillus</td>
<td>acidophilus lactis</td>
<td>FM 3x100 g/day 6 weeks</td>
<td>SB, PC n = 14 (moderate hypercholesterolemia)</td>
<td>Ataie-Jafari et al. (2009)</td>
</tr>
<tr>
<td>Reducing total plasma cholesterol level compared to placebo group</td>
<td>Lactobacillus</td>
<td>plantarum (CECT 7527, CECT 7528, CECT 7529)</td>
<td>1 capsule/day 12 weeks</td>
<td>R, DB, PC, parallel trial n = 60</td>
<td>Fuentes et al. (2012)</td>
</tr>
<tr>
<td>Reducing total plasma cholesterol from baseline and placebo-group</td>
<td>Lactobacillus</td>
<td>renteri (NCIMB 30242)</td>
<td>Capsule 2/day, 9-week intervention</td>
<td>R, DB, PC n = 127 (hypercholesterolemic)</td>
<td>Jones et al. (2012)</td>
</tr>
<tr>
<td>Reducing LDL cholesterol correlated with an increase in plasma de-conjugated bile acids Decreasing plasma fibrinogen</td>
<td>Lactobacillus</td>
<td>salivarius (UBL S22)</td>
<td>Capsule 1/day, 6 weeks</td>
<td>R, SB, PC pilot study n = 45 3 groups; P, w/L. salivarius, w/L. salivarius + FOS</td>
<td>Rajkumar et al. (2015)</td>
</tr>
<tr>
<td>Metabolism and disease markers</td>
<td>Lactobacillus</td>
<td>acidophilus</td>
<td>FM 2/day</td>
<td>Intervention study with multiple phases n = 11 (healthy)</td>
<td>Lidbeck et al. (1992)</td>
</tr>
<tr>
<td>Reducing in serum concentrations of CRP, IL-1β/IL-6, TNF-α</td>
<td>Lactobacillus</td>
<td>salivarius (UBL S22)</td>
<td>Capsule 1/day, 6 wk</td>
<td>R, SB, OC pilot study n = 45 3 groups; P, w/L. salivarius, w/L. salivarius + FOS</td>
<td>Rajkumar et al. (2015)</td>
</tr>
<tr>
<td>Decreased TNF-α, IL-6 cytokines</td>
<td>Bifidobacterium</td>
<td>lactis (HN019)</td>
<td>FM 80mL/day 45 d</td>
<td>R, PC n = 51 (symptoms of metabolic syndrome)</td>
<td>Bernini et al. (2016)</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Improvement of IBD symptoms Increased faecal β-galactosidase activity</td>
<td>Lactobacillus</td>
<td>acidophilus, casei, brueckii spp.</td>
<td>3/day, 20 d</td>
<td>VSL-3 (pharmaceutical probiotic) administered</td>
<td>Brigidi et al. (2001)</td>
</tr>
<tr>
<td></td>
<td>Bifidobacterium</td>
<td>longum (Y10), breve (Y8), infantis (Y1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Streptococcus</td>
<td>salivarius, spp. thermophilus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decreased β-glucuronidase activity</td>
<td>Lactulose, oligofructose-inriched inulin</td>
<td>2/day, 4 wk</td>
<td>R, C-O, PC study n = 53 (healthy)</td>
<td>Preter et al. (2008)</td>
<td></td>
</tr>
<tr>
<td>Decreased β-glucuronidase activity Increased β-glucosidase activity (B. brevis)</td>
<td>Lactobacillus</td>
<td>casei (Shirota)</td>
<td>FM 2/day, 4 wk</td>
<td>R, C-O, PC study n = 53 (healthy)</td>
<td>Preter et al. (2008)</td>
</tr>
<tr>
<td></td>
<td>Saccharomyces</td>
<td>boulardii</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bifidobacterium</td>
<td>brevis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Alterations in the intestinal microflora</strong></td>
<td>Lactobacillus</td>
<td>salivarius (UBL S22)</td>
<td>Capsule 1/day, 6 wk</td>
<td>R, SB, PC, pilot study n = 45 3 groups; P, w/L, salivarius, w/L, salivarius + FOS n = 45</td>
<td>Rajkumar et al. (2015)</td>
</tr>
<tr>
<td></td>
<td>FOS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increased counts of Bifidobacterium and Lactobacillus</td>
<td>XOS</td>
<td>Rice porridge 1/day 6 wk</td>
<td>R, PC n = 20 (healthy)</td>
<td></td>
<td>Lin et al. (2016)</td>
</tr>
<tr>
<td>No significant effect on faecal concentration of total lactobacilli or bifidobacterium</td>
<td>Lactobacillus</td>
<td>acidophilus, rhamnosus, plantarum, casei, helveticus, bulgaricus</td>
<td>2 capsules/day 1 wk</td>
<td>Short-term intervention study n = 11 (healthy) Commercial product Measurement of faecal composition of lactobactobacilli and Bifidobacterium using qPCR</td>
<td>Taverniti et al. (2014)</td>
</tr>
<tr>
<td></td>
<td>Lactococcus</td>
<td>lactis thermophilus bifidum shermani</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Streptococcus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bifidobacterium</td>
<td>Propionibacterium</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** C-O, cross-over; CRP, C-reactive protein; d, days; DB, double-blinded; FM, fermented milk; FOS, fructo-oligosaccharide; IL, interleukin; P, placebo; PC, placebo controlled; R, randomized; SB, single-blinded; TNF, tumor necrosis factor; w/, with; wk, weeks; XOS, xylo-oligosaccharide.

Indications of alterations in microbial activity and anti-mutagenic activity following pre- and/or probiotic intake have been observed by Preter et al. (2008), who found decreased β-glucuronidase activity in faeces samples from patients receiving Lactulose, *Lactobacillus Casei* (Shirota), *Bifidobacterium brevis* or combined pre- and probiotics (symbiotic). The authors assigned this decrease to have an impact on the development of potentially carcinogenic compounds.

Furthermore, Rajkumar et al. (2015) observed indications of an improved immune response due to a reduction in pro-inflammatory cytokines in the plasma. The authors assigned the effect to improvements of intestinal barrier function induced by *Lactobacillus salivarius* and/or FOS.
Despite demonstrations of improved health effects in humans following pre- and/or probiotic supplementation, results of different studies are frequently contradictory. For example, Ataie-Jafari et al. (2009) and Fuentes et al. (2012) found improved effects on plasma cholesterol following probiotic supplementation. However, a comprehensive systematic review of human studies on probiotics and lipid profile (triglyceride, total cholesterol, LDL cholesterol, HDL cholesterol and blood pressure) found that out of 22 studies with 16 different probiotic species only 8 species proved to have an effect, indicating great inconsistency (Kallioma et al., 2010). Another thorough review by Preter et al. (2011) on the effect of pre- and probiotics stresses how prebiotics generally prove to increase the production of SCFAs in human, animal and in vitro studies. However, the effects of probiotics on microbial activity in human studies are more inconsistent as compared to animal studies or in vitro studies. Furthermore, Flaherty et al. (2015) reviewed the impact of supplementation on mucosal immunity and found that the potential beneficial effector molecules produced by probiotics are highly species and strain specific. Thus, probiotics exert species and strain specific changes on the host’s health, and the effect of one strain cannot be extrapolated to other strains (Fijan, 2014; Shah, 2007).

To sum up: the human gastrointestinal tract is colonized by a wide range of microorganisms that ferment non-digested foodstuff entering the colon to a variety of end-products. Some of these microorganisms improve the host’s health and are said to be probiotic. The activities of these can be selectively induced by prebiotics. The underlying biochemical mechanisms of pre- and probiotic impacts on host’s metabolism are not clearly elucidated, but are generally assumed to be exerted through modulation of microbial activity and microbial produced metabolites (Arora et al., 2013; Preter et al., 2011; Turroni et al., 2014). These metabolites either exert beneficial effect in the colon or in the circulatory system of the host, thus influencing other biochemical reactions through complex interactions.

c) Exploring the human metabolome

Through a symbiotic relationship with the host and subsequent degradation of foodstuff, the gut flora influences the metabolome of an individual. The metabolome is a collection of metabolites (<2000 Da*) of both extrinsic and intrinsic origin found in a biological system (Wishart et al., 2013). Therefore, the metabolome of an individual is composed of endogenous compounds and xenobiotics of extrinsic origin (Holmes et al., 2008). Metabolites can be perceived as the “end-point of the bio-molecular gene-protein-metabolite cascade” (Claus & Swann, 2013). Hence, the metabolites and the metabolome expose information regarding the biochemical processes in an organism and/or a biological compartment (Alonso et al., 2015; Holmes et al., 2008). Several factors exhibit great influence on the composition of metabolites in an organism.

*The limit is not clearly defined, others refer to metabolites as being <1500Da, (Wishart, 2014).
Among others are gender, age, gut flora and diet contributors to variation (Claus & Swann, 2013; Slupsky et al., 2007), making the composition highly complex. In order to analyse the human metabolome by quantifying and identifying metabolites, the technique metabolomics is emerging (Alonso et al., 2015; Wishart, 2014).

c. i) Metabolomics
The terms metabolomics and metabonomics are frequently used interchangeably. However, metabonomics is defined as “the quantitative measurements of the time-related multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification” (Nicholson et al., 1999), whereas metabolomics can be defined as “a comprehensive analysis in which all the metabolites of a biological system are identified and quantified” (Lindon & Nicholson, 2008). The term metabolomics will be used in the current study.

Metabolomics offers analysis of endogenous and exogenous metabolites in different biological samples by the use of multitargeted analysis (Bijlsma et al., 2006; Lindon, 2004) and have been used in studies regarding the gut flora (Nicholls et al., 2003; Nicholson et al., 2005).

Metabolomics aims at exploring the full metabolome of an organism by either targeted and/or non-targeted analysis. Targeted metabolomics can be described as being biased for selected metabolites. Opposite, un-targeted metabolomics is an analysis of all detectable metabolites, where one aims at profiling the entire metabolic profile of a sample without a prior hypothesis regarding the composition (Alonso et al., 2005; Ryan et al., 2013).

The procedure of metabolomics progresses from chemical analysis to identification and quantification of compounds to a final interpretation of the data in a biological context (Roessner & Bowne, 2009), resulting in a biological integrated analysis of the organism (Holmes et al., 2008; Slupsky et al., 2007). This follows the need of analytical software and databases in order to interpret the results of the chemical analysis (Beckonert et al., 2007).

The widespread use of metabolomics (and metabonomics) in different scientific fields has given rise to the Human Metabolome Database (HMDB: http://www.hmdb.ca/), which “brings together quantitative, chemical, physical, clinical and biological data about all experimentally ‘detected’ and biologically ‘expected’ human metabolites” (Wishart et al., 2013).

Chemical analysis used in metabolomics
The following two types of chemical analysis techniques are most frequently used in metabolomics: Nuclear magnetic resonance spectroscopy (NMR) and mass spectroscopy (MS) (gas chromatography-mass spectrometry (GC-MS) or liquid chromatography-mass spectrometry (LC-MS)) (Slupsky et al., 2007; Wishart, 2014).
A description of NMR-spectroscopy is pursued in the following with emphasis on the exploitation of physical and chemical properties in regards to metabolomics and to a lesser extend on the physical and chemical principles behind NMR analysis. Finally, an introduction to interpretation and exploration of spectroscopic data by multivariate analysis is given in the section Chemometrics.

**Basic principles of nuclear magnetic resonance spectroscopy (NMR) in regards to metabolomics**

**Data acquisition**

NMR is a quantitative, non-destructive technique and can be used for analysis of molecular structures in a biological sample (Beckonert et al., 2007). NMR is able to identify and quantify approximately 50-200 metabolites simultaneously with a detection limit of 5µM, and has been utilised in the past decades for characterization of human derived biofluids (Wishart, 2014).

The molecules detected by NMR are polar molecules such as amino acids, sugars and organic acids (Wishart, 2014). The technique is based on nuclei spin, energy absorption of specific nuclei subjected to a magnetic field and re-emission (Bothwell & Griffin, 2011).

Nuclei of atoms are charged and have a certain spin quantum number, \( I \), which differs between different nuclei, though always in multiples of \( \pm \frac{1}{2} \) (Claridge, 1999). Spin properties of nuclei create distinct magnetic moments when the nuclei are placed in an external magnetic field. This makes the nuclei behave like a magnet and causes them to be oriented in a certain direction (Lane, 2012).

In a magnetic field, \( B \), the individual nuclei precess around their own axis. They absorb photons of the frequency, \( \nu \), determined by the gyromagnetic ratio, \( \gamma \), of the nuclei and the strength of the magnetic field (Pavia et al., 2001), thus:

\[
\nu = \gamma \times B
\]

By applying a mathematical approach, the orientations (magnetic moments) of nuclei are expressed through a vector model by which the bulk of magnetic moments of nuclei is along the \( z \)-axis in a Cartesian co-ordinate system, parallel to the magnetic field (Claridge, 1999).

Nuclei with spin take up \( 2I+1 \) possible orientations. Therefore, protons with spin \( \frac{1}{2} \) have 2 distinct energy states: in the orientation parallel to the magnetic field (lower energy state, \( \alpha \)) or opposite to the magnetic field (higher energy state, \( \beta \)). Not all nuclei possess spin properties, and these therefore have a quantum number of 0. This makes them undetectable in NMR analysis (Claridge, 1999). The isotopes \(^1\text{H}, ^{13}\text{C} \) or \(^{31}\text{P} \) spin \( \pm \frac{1}{2} \) and are frequently used.

Individual nuclei have distinct magnetic moments and thus absorb radio waves at different frequencies. NMR exploits spin properties and magnetic moments and induces transition of nuclei to a higher energy state (Lane, 2012). The energy of photons must match the energy difference between the energy states of the nuclei.
The energy of a photon is related to Planck’s constant, \( h \), and the frequency, \( \nu \). Therefore the energy, \( E \), of a photon needed to cause transition is expressed by:

\[
E = h \cdot \gamma \cdot B \quad \text{(Claridge, 1999)}.
\]

Thus the frequency needed to excite the system is given by:

\[
\nu = \Delta E / h \quad \text{(Lane, 2012)}.
\]

However, despite some nuclei having the same spin (such as \(^1\text{H}, ^{13}\text{C} \text{ or } ^{31}\text{P}\)), they do not have the same differences in energy state. Therefore, one can apply a certain energy pulse in order to selectively induce energy transition in specific nuclei (Lane, 2012).

Resultantly, NMR spectroscopy induces a transition in energy states of nuclei with spin properties and distinct magnetic moments by applying a defined energy pulse on a sample, causing the nuclei to have resonance with the incoming wave (Pavia et al., 2001).

Nuclei are not equally distributed in the distinct energy states, but separated according to Boltzmann’s distribution. Some will be in the greater energy state and others in the lower energy state. The stronger the energy applied on nuclei, the greater the difference in population. Hence, more nuclei will be in the low energy state for excitation (Pavia et al., 2001).

Nuclei in a complex sample are present in different molecules, and subject to different local magnetic environments, which affect the resonance frequency by which they absorb. This leads to individual chemical shift, \( \partial \), of nuclei (Bothwell & Griffin, 2011). Therefore two magnetic sources affect the nuclei; namely the applied magnetic field and the local magnetic environment.

Application of a 90° pulse causes the bulk magnetization of specific nuclei to rotate from the \( z \)-axis to the \( x \)-\( y \) plane in a Cartesian co-ordinate system. When turning off the applied pulse, the nuclei relax back to equilibrium and return the magnetization to the \( z \)-axis (Lane, 2012), subsequently emitting radio waves with wavelengths dependent on the nature of the molecules and the chemical environment. The signals emitted from the nuclei and the time it takes to restore to the \( z \)-plane are characteristic for different nuclei in different molecules and measured in NMR. This time is determined by the spin-lattice relaxation time, \( T_1 \), and the spin-spin relaxation time, \( T_2 \). \( T_1 \) refers to the time it takes for the longitudinal magnetization to recover, whereas \( T_2 \) refers to the decay of transverse magnetization (Hornak, 1997). This energy emission is collected as signals called free induction decays (FIDs) by a detector (Bothwell & Griffin, 2011).

The electromagnetic energy emitted can be translated to a function of frequencies by a Fourier transformation, which converts time-domain data into a spectrum or plot of intensity vs. frequency (ppm) (Bothwell & Griffin, 2011) as depicted in figure 1.
Figure 1: Transformation of radio-waves emission into a 1D spectrum, adapted from Bothwell & Griffin (2011)

Data of NMR spectroscopy are reflected in one-dimensional (1D) spectra or two-dimensional (2D) spectra with 1D being the most prevalent. A 1D spectrum has two axes. A y-axis represents the intensity of absorption and a x-axis represents the chemical shifts (values in parts per million, ppm) expressed by the difference in resonance between a reference substance and a given substance (Blümich & Callaghan, 1995).

In order to characterize and quantify the peaks, one must consider the chemical shifts, the area of the peaks and the pattern of the peaks regarding spin-coupling.

**Spectral analysis - identification and quantification of metabolites**

Chemical shifts are essential for identification of molecules in a sample (Lane, 2012). The chemical shifts reflect the nature of nuclei and the local magnetic environment they are situated in, which causes the nuclei to absorb at different frequencies than they otherwise would in a single atom (Chemistry, n.d.). The chemical shifts are relative to a reference set at 0 ppm.

Adjacent nuclei located in different magnetic environments influence each other’s magnetic field, reflected by intramolecular interactions called scalar coupling.

Scalar coupling reflects bond order and strength between nuclei in a molecule, and is portrayed in a NMR spectrum by splitting of the resonance signal.

This separation is determined by the coupling constant, $J$. A general rule is that a nucleus situated in a molecule with mutually coupled nuclei less than three bonds away results in $n+1$ peaks (Hornak, 1997; Lane, 2012). A triplet therefore reflects a nucleus scalar coupled to two equivalent nuclei.
Through examination of the pattern of spectral peaks, one can obtain information concerning the chemical structure of the molecules, and identify the metabolites detected by NMR (Alonso et al., 2015). A table of chemical shifts in $^1$HNMR is depicted in appendix A1.

The area under a peak is proportional to the relative number of hydrogens given rise to the peak (proton NMR). Therefore, one can integrate the peaks and quantify the concentration, provided the presence of a reference and knowledge of chemical shifts (Lindon, 2004; Pavia et al., 2001; Smolinska et al., 2012). The references for chemical shift that are most frequently used are 3-trimethylsilylpropionic acid (TSP), 2,2-dimethyl-2-siapentane-5-sulfonate sodium salt (DSS), and tetramethylsilane (TMS) for organic solvents (Beckonert et al., 2007) and are set at 0.00 ppm.

A basic pulse sequence of a 1D NMR experiment demands setup of relaxation delay, pulse width and acquisition time of the time domain data. The latter can be zero-filled in order to maximize the data points used for computing the spectrum (Pauli et al., 2005). Deuterated solvents such as deuterium oxide (D$_2$O) are added in order to lock on a deuterium signal.

The aforementioned HMDB offers great assistance in identification of metabolites in a spectrum. However, owing to the complexity of a biological sample, peaks in a spectrum might overlap, making it difficult to distinguish between compounds. To overcome this, one can run another experiment with an increased concentration of one of the compounds, thereby spiking the sample (Dona et al., 2016). By subsequent spectrum analysis, the peaks can be identified. Other methods include running a heteronuclear 2D spectrum analysis. Heteronuclear 2D spectrum analysis provides increased signal dispersion and divides overlapping peaks by exploring the scalar interactions between a proton and another magnetic nucleus (Lane, 2012; Wishart, 2014). Graphically, the analysis is displayed by a spectrum of two axes, with chemical shift values of one of the nuclei represented on either of the axis and chemical shifts of the other nuclei on the other axis.

**Data pre-processing**

Data pre-processing is an intermediate step between the raw data obtained from the chemical analysis and data interpretation. This follows the need of phase correction, baseline correction, alignment, normalization, binning and scaling. Some authors distinguish between data pre-processing and data pre-treatment (Goodacre et al., 2007), however in this thesis, the terms will be considered in concert as according to Smolinska et al. (2012).

Phasing and baseline-correction are done manually or automatically by NMR spectrum software. Phasing aims at correcting asymmetrical peaks following Fourier transformation. Base-line correction aims at reducing bends across the base-lines (Goodacre et al., 2007).

In order to align the peaks in a spectrum, the signal of the internal reference is adjusted to 0 ppm.
Alignment by the method *icoshift* (interval correlated shifting) divides spectra into various segments and aligns the spectra to the corresponding segments of a reference spectrum (Tomasi et al., 2011).

Normalization of the urinary metabolite concentration is important, since the overall concentration of metabolites in urine can vary to a great extent. The peak area of creatinine is often used as a reference for normalization, since creatinine clearance in urine is considered constant (Smolinska et al., 2012), and for compensation of potential variations in urine volume, which can affect the concentrations of metabolites (Craig et al., 2006; Saude et al., 2007).

Other frequently applied methods include normalization according to urine volume, area of TSP signal, the use of probabilistic quotient normalization, the sum of the squared value of all variables for a given sample (2-Norm) and sum of value of all variables for a given sample (1-Norm) (Rasmussen et al., 2011).

Binning divides the spectrum into segments of either equal size (equidistant binning) or non-equidistant, and therefore reduces the dimensionality of the spectrum (Ross et al., 2007).

Mean-centering and scaling are frequently applied in concert. Mean-centering is done by subtracting the mean from each value, thereby differences between high-concentrated and low-concentrated metabolites are adjusted to vary around zero (Craig et al., 2006; Smolinska et al., 2012). Scaling is applied to reduce variations of data values (e.g. spectral intensities).

Scaling can be done by unit variance (UV) scaling, where one calculates the standard deviation of each variable, $s_k$, and subsequently multiply with a scaling weight $1/s_k$ (Eriksson et al., 2006) or by Pareto scaling, where each variable is divided by the square root of $s_k$ (Trygg et al., 2007).

Semiautomatic identification and/or quantification of metabolites in a sample can be obtained by commercial software such as Chenomx NMR suite, which also refers to HMDB for peak identification. This programme also allows for phase correction, baseline correction, reference adjustment and pH determination (Wishart, 2014).

**Chemometrics**

Data analysis aims at exploring the data and spectra following data pre-processing and detecting significant trends and differences (Smolinska et al., 2012). Chemometrics is a mathematical interpretation of chemical data. Both supervised and unsupervised analysis can be employed.

Principal component analysis (PCA) is an unsupervised, multivariate projection method, which aims at portraying the total variance in a data set (Eriksson et al., 2006; Trygg, 2003).

It summarizes variation in a dataset into scores (T) and different loadings (P) of these scores resulting in “principals components” (PCs). A dataset is summed in a matrix X with theoretically N observations and K variables (Eriksson et al., 2006). The observations are plotted in a K-dimensional space; therefore each observation is situated according to its measured value.
The first PC (PC1) represents the greatest variance explained, and is an approximation of data according to least square sense (Eriksson et al., 2006). Each point in the dataset is projected onto the PC1 line and has a co-ordinate named a score. The second PC (PC2) is orthogonal to PC1, and represents the second greatest variance. These two PCs now define a plane with the observations having different co-ordinates (scores) and provide insight in the maximum explained variation in two dimensions. However, multiple PCs can be computed dependent on the size of the data set with each PC explaining the observed variance in descending order. A general rule is that at least 4 observations are needed to compute one PC.

A graphical presentation of the scores is depicted from a scores plot, which presents the observations by the use of the co-ordinate values resulting from the PCs and groups the observations according to variance. Each variable influence the variance between observations to different extends and therefore have individual vectors, which are applied on the PCs. These vectors are termed loading vectors, and the resulting influence of the individual loadings on the PCs is depicted in a loading plot (Eriksson et al., 2006). Thus, the latter represents the weight by which different factors affect the observed variance.

In order to validate one’s model as plotted by PCA, one can confer with the $Q_2$ value. This value indicates the quality of the model in terms of consistency between predicted and original data. The value has a theoretical maximum of 1, and generally one can accept a model of $Q_2 \geq 0.40$ (Worley & Powers, 2013).

c. ii) Metabolic composition of urine

The metabolic composition of urine reflects metabolic processes in both liver and kidneys (Lindon et al., 2000; Wishart, 2014). The primarily role of the kidneys is to maintain homeostasis by filtering body fluids and excreting toxins and other waste products.

Therefore, urine is apt for studying degradation or detoxification pathways (Ryan et al. 2013), and a suitable matrix for investigating biomarkers, which can be defined as “measureable changes associated with physiological and pathological processes” (Wu & Gao, 2015).

Urine sampling is non-invasive, and the complex biofluid has been largely investigated by several chemists in order to characterize and identify the human urine metabolome (Bouatra et al., 2013; Saude & Sykes, 2007).

Through review of preceding studies and literature, Bouatra et al. (2013) composed the Urine Metabolome Database (UMDB: http://www.urinemetabolome.ca) as an integrated part of HMDB. This database provides knowledge of metabolites found in human urine, their corresponding concentrations and potential disease association (Bouatra et al., 2013). Urine has a final pH of around 6.2, and the most abundant metabolites include urea, creatinine, hippurate and citrate, inorganic ions and salts, water-soluble toxins and pigments (Bouatra et al., 2013; Rose et al., 2015).
Extended information on the composition of the final urine and concentration of metabolites can be found in Bouatra et al. (2013).

A typical spectrum of urinary metabolites is visualised in figure 2 showing the complexity of urine and illustrating some quantitatively great metabolites such as hippurate, urea, creatinine and citrate measured by $^1$H NMR.

![Figure 2: Spectrum of metabolites found in urine by the use of $^1$H NMR. TMAO, trimethylamine-N-oxide; TSP, 3-trimethylsilylpropionic acid (reference substance), adapted from Beckonert et al., 2007.](image)

The metabolic composition of urine depends 	extit{inter alia} on diet, time of day and age, leading to intra- and inter-individual differences with an average concentration range of ±50% (Bouatra et al., 2013). This makes quantification, identification and analysis of the urine in regards to elucidating characteristics of phenotypes challenging (Bouatra et al., 2013; Lindon et al., 2000; Rose et al., 2015). Table 2 assembles examples of sources to variations and their effect on the metabolic composition of urine.

### Table 2: Sources of variation in the metabolic composition of human urine

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Main effects of human NMR studies (urine)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W:</td>
<td>$^{1,2,3,4}$citrate, $^{2,3}$glycine, $^2$hippurate, $^3$lactate, $^4$fumarate, $^4$creatine</td>
<td>$^1$Kochhar et al. (2006)</td>
</tr>
<tr>
<td>HL:</td>
<td>$^4$taurine, $^{1,2,3,4}$creatinine/creatinine, $^3$TMAO, $^3$3-hydroxybutyrate, $^3$carnitine, $^3$acetyl/carnitine, $^3$acetone</td>
<td>$^2$Wang et al. (2005)</td>
</tr>
<tr>
<td>M:</td>
<td>$^3$hippurate, $^4$TMAO</td>
<td>$^3$Psihogios et al. (2008)</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Older individuals:</td>
<td>$^{1,3,4}$creatinine/creatinine, $^3$3-hydroxybutyrate, $^3$3-hydroxyisovalerate, $^4$cis-aconitate, $^4$lactate, $^4$alanine, $^4$carnitine</td>
<td>$^4$Slupsky et al. (2007)</td>
</tr>
<tr>
<td>LL:</td>
<td>$^4$hippurate, $^4$TMAO, $^4$trigonelline</td>
<td>$^4$Slupsky et al. (2007)</td>
</tr>
</tbody>
</table>
As depicted in table 2, dietary differences have contributed to variations in acetate, hippurate, lactate, citrate and glycine excretions in urine (Holmes et al., 1994). Additionally, intake of protein sources also contributes to variations in the composition of urinary metabolites, especially due to changes in trimethylamine-N-oxide, creatinine and citrate excretions (Lenz et al., 2004; Rasmussen et al., 2012; Stella et al., 2006).

Furthermore, gender differences have been reflected by greater citrate and glycine excretions in females than males, whereas creatinine is excreted in higher concentrations in males as it correlates with body weight and muscle mass (Kochhar et al., 2006; Psihogios et al., 2008; Wang et al., 2005). Adding to the complexity, the gut flora of an individual also affects the metabolic profile of urine. As previous stated, the gut flora produces several compounds when fermenting foodstuff in the colon, leading to great impact on the metabolites circulating in the body and therefore also the composition of metabolites in urine. The composition of gut flora is dependent on multiple factors such as host genotype, diet, age, and can be linked to certain metabolic phenotypes and metabolic profiles (Calvani et al., 2010; Li et al., 2008).

Prior knowledge of the metabolic composition of urine and the normal physiological differences that are present between individuals is therefore of great importance for urine metabolomic analysis, since one must differentiate between normal inter- and intra-individual variations and the

<table>
<thead>
<tr>
<th>Diurnal</th>
<th>Morning:</th>
<th>8Slupsky et al. (2007)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diet</td>
<td>8Standardization proved less inter-individual variation</td>
<td>8Walsh et al. (2006)</td>
</tr>
<tr>
<td></td>
<td>8Greatest proportional variation between people in citrate, hippurate, lactate, acetate, glycine</td>
<td>8Holmes et al. (1994)</td>
</tr>
<tr>
<td></td>
<td>HL: hippurate after 8fruit and vegetables intake and 8tea consumption</td>
<td>8Lenz et al. (2004)</td>
</tr>
<tr>
<td></td>
<td>7.8.9.10 Protein diet (fish/meat/protein):</td>
<td>8Dumas et al. (2006)</td>
</tr>
<tr>
<td></td>
<td>HL: 7.8.9.10 TMAO, 8.10 creatinine, 8 carnitine, 8 taurine, 9 creatine, 7,8 and 3-methylhistidine, 10 nitrogen excretion</td>
<td>8Stella et al. (2006)</td>
</tr>
<tr>
<td></td>
<td>LL: 9 citrate</td>
<td>8Rasmussen et al. (2012)</td>
</tr>
<tr>
<td></td>
<td>12 Vegetarian diet:</td>
<td>8Xu et al. (2010)</td>
</tr>
<tr>
<td></td>
<td>HL: NAG, glycine, succinate, citrate, hippurate</td>
<td>8Tuk et al. (2012)</td>
</tr>
<tr>
<td></td>
<td>LL: taurine, formate, TMAO, phenylalanine, methylhistidine</td>
<td>8Mulder et al. (2005)</td>
</tr>
<tr>
<td>Body mass index</td>
<td>9Obese, BMI &gt;25:</td>
<td>8Calvani et al. (2010)</td>
</tr>
<tr>
<td></td>
<td>HL: 2-hydroxyisobutyrate</td>
<td>8Slupsky et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>LL: hippurate, xanthine, trigonelline</td>
<td>8Calvani et al. (2010)</td>
</tr>
<tr>
<td>Gut flora derived metabolites</td>
<td>8TMAO, 9 dimethylamine, 8.11 hippurate, 8.11.2 hydroxyisobutyrate, 9 phenolic, benzoyl, 11 phenyl derivatives, 8 derivatives of choline metabolism, 8 indole derivatives, 8 p-cresol, 11 3 aminoisobutyrate, 11 4-cresol sulphate, 11 taurine, 11 lactate</td>
<td>8Nicholson et al. (2012)</td>
</tr>
<tr>
<td></td>
<td>8 Slupsky et al. (2007)</td>
<td>8Preter et al. (2008)</td>
</tr>
<tr>
<td></td>
<td>9 Calvani et al. (2010)</td>
<td>8Li et al. (2008)</td>
</tr>
</tbody>
</table>

Abbreviations: HL, higher level; LL, lower level; M, men; NAG, N-acetyl glycoprotein; P, plasma; TMAO, trimethylamine-N-oxide; W, women; a reference and a study are related, if they have the same subscripts.
differences correlating to a physiological response of a disease state or intervention (Saude & Sykes, 2007). In regards to elucidating isolated effects of probiotic supplementation on the metabolic composition of urine, several factors such as diet, lifestyle, gender and age must be taken into account.
2 Materials and methods

a) Pilot study and sample collection

Two participants were enrolled in the study: subject K (female) and subject U (male).
Both participants were equipped with 22 Actimel Danone Strawberry yoghurts fermented with Lactobacillus Casei Danone-culture (Lactobacillus paracasei subsp. paracasei ((DN-114001/CNCM I-1518), see appendix A2 for ingredients and brand details).
The participants were asked to continue their regular diet and regular life routines though should not ingest any other products with claimed probiotic content. No controls were used in the study, and the subjects were not asked to keep track of their diet.
Urine sample number I was collected on day 0. The first yoghurt was subsequently ingested in the morning of day 0 after sample number I had been collected.
One yoghurt á day was consumed for 21 days. Last yoghurt was ingested on day 21.
The yoghurts were consumed in the morning preferable at the same time of day each day.
The samples should represent the “morning pee”, thereby should be collected before breakfast in the early morning. Complete sample homogeneity and integrity cannot be obtained, however were attempted by collecting the urine approximately at the same time of day and by storing the samples in a freezer around [-20; -18°C] immediately after collection.
The intervention period was followed by a wash-out period of 6 weeks, after which a final urine sample was collected.
Figure 3 is an outline of the course of the intervention study including number of day and the corresponding sample number. See appendix A3 for a detailed outline of the intervention.

b) Sample preparation and data acquisition by \(^1\)H NMR

The urine samples were stored in a regular freezer (-18° C) until chemical analysis.
The 12 urine samples thawed in a water bath and were centrifuged at 12000g for 5 minutes at +4 °C. The samples were kept cool in a refrigerator (+5 ° C) until pH measurements were executed.
SIII from each observation was used for pH measurements and buffer adjustments in order to verify that the samples were in the range of pH [6.5-7.5].
After verification, a Master Mix was prepared by mixing 1380 µL buffer and 1620 µL D₂O water (Sigma-Aldrich, Copenhagen, Denmark) to a total volume of 3000 µL. The buffer was composed of NaH₂PO₄ (583 mM) (Sigma-Aldrich, Copenhagen, Denmark) sodium formate (1.162 mM) with a final concentration of 0.25 mM DSS (Sigma-Aldrich, Copenhagen, Denmark) as internal chemical shift reference (see appendix A4 for ingredients list).

A duplicate was prepared for each of the 12 samples resulting in 24 samples (see appendix A5 for IDs). Therefore, 24 standard 5 mm NMR tubes were prepared by adding 130 µL Master Mix and 570 µL of a given sample to a corresponding tube. All tubes were gently mixed and allocated in a random order for analysis.

Spectral measurements were obtained by the use of a Bruker Avance 600 NMR spectrometer (Bruker Biospin, Rheinstetten, Germany) with a frequency of 600.13 MHz, and data acquisition based on Beckonert et al. (2007). Samples were analysed at 298 K and individually shimmed before each run. Samples in an aqueous matrix (e.g. urine and serum) need water suppression. This is done by running pulse experiments such as a presaturation 1D NOESY (Nuclear Overhauser Enhancement Spectroscopy) - experiment (Beckonert et al., 2007). Thus, one-dimensional spectra were acquired by the use of 1D NOESY -presat with a 90° pulse sequence and automatic water suppression during the relaxation delay (Bruker standard pulse: noesypr1). The relaxation delay was set at 2 seconds and mixing time 0.1 s. A total of 64 scans for each sample were collected. The spectral width was set at 20.03 ppm with 32K data points. Subsequent zero-filling resulted in approximately 65K data points.

The 1H spectra were processed in Topspin 3.0 (Bruker Biospin) and adjusted to the signal of the internal reference DSS at 0 ppm. The spectra were individually phased and baseline-corrected before further processing.

Prior to Fourier transformation, the FIDs were multiplied by 0.30 Hz line-broadening and the water region was eliminated from each spectra.

c) Identification of metabolites and multivariate analysis
Chenomx NMR suite 7.0 was used to identify the metabolites in the spectra. Automatic fit was applied (internal library: Chenomx NMR 600 MHz) in order to ease the identification.

The spectra were analysed in order to construct metabolic profiles of the urine samples. Afterwards, peaks that were not identified or verified by the automatic approach were analysed and evaluated by comparing the spectra with literature (Bouatra et al., 2013) and HMDB.

A targeted approach was used to examine for peaks of known microbial origin such as SCFAs.

Raw NMR data were pre-processed in Matlab. Alignment was done using Icoshift (Tomasi et al., 2011) by setting the reference peak to 0.00 ppm and subsequently overlaying every spectrum on top of each other.
The spectra were subsequently divided in intervals and aligned again in order to reduce variation in chemical shift between the spectra. Prior to normalization, the water region ranging from 4.7-4.9 ppm was removed, and likewise were the regions below 0.05 ppm and above 10.0 ppm. The resulting spectra therefore ranged from [0.05;10.0] ppm with the water region excluded. Normalization was proceeded by 1-Norm and subsequent binning of spectra by binning sizes of 0.1 reduced the dimensionality of the spectra to consist of 927 data points per spectrum. The resulting data were transferred to SIMCA 14.1 for multivariate analysis by Principal component analysis (PCA) and were scaled by Pareto scaling. The computed models were selected by conferring to the $Q_2$ value, and data were then explored. Mayor metabolites contributing to variation were quantified by normalization to creatinine.
3 Results

a) \(^1\)H NMR spectra and some identified metabolites

The main objective of this metabolomics study was to analyse changes in urinary metabolic profiles following probiotic intake. By the use of automatic fit in Chenomx NMR suite 7.0, an average of 109 out of 338 metabolites were identified per spectrum. After analysis of the individual spectra and fitting of the identified pre-fitted metabolites to each spectrum, a total of 36 different metabolites were frequently observed (appendix A6). Additionally, 12 unknown compounds (UN) were observed in the different spectra (appendix A6).

\(^1\)H NMR spectra of sample V (K5 and U5) from each individual are portrayed in figure 4 and 5.

Figure 4: Representative spectra of subject K with some identified metabolites. Sample K5 is portrayed, collected 22 days after the start of the intervention. Horizontal axis: chemical shift (ppm), vertical axis: intensity (arbitrary units).

Figure 5: Representative spectra of subject U with some identified metabolites. Sample U5 is portrayed, collected 22 days after start of the intervention. Horizontal axis: chemical shift (ppm), vertical axis: intensity (arbitrary units).
Figure 4 and figure 5 illustrate great similarity regarding the overall urinary metabolic profile, but the intensity of some identified metabolites differs to a great extend between the subjects. Hippurate and urea are seemingly more intense in subject K than subject U, whereas peaks corresponding to creatinine and citrate are greater in subject U (figure 4 and figure 5).

b) Inter-individual differences reflected by PCA plots
In order to examine the variance in the data set, a principal component analysis of $^1$H NMR spectra from every urine sample collected was performed (including replicates). The model (PCA1) resulted in 5 principal components (PCs) with $Q^2 = 0.70$. A scores plot with PC1 vs. PC2 is depicted in figure 6.

Figure 6: PCA1 scores plot of the first and second principal component (PC1 and PC2 respectively), representing variance between individuals K (green) and U (blue). Ellipses encircle the two individuals.

Figure 7: PCA1 loading plot of variation across PC1 with identified peaks. Encircled areas comprise a given identification.
The greatest source of variance is reflected by PC1 (54.6% of total variance) and separates the samples according to individual with subject K located on the left side of the plot and subject U on the right side (figure 6). Separation by PC2 proves less variation in samples of subject U, while samples of subject K are dispersed to a greater extent (figure 6).

The peaks accountable for the variation across the PC1 line were identified in Chenomx and the identifications are depicted in figure 7. Metabolites strongly contributing to variation are creatinine, citrate, TMAO/betaine, trigonelline, urea, UN1 and hippurate. Metabolites elevated in subject U were creatinine and citrate, whereas metabolites elevated in subject K were trigonelline, urea, UN1 and hippurate (figure 7).

The concentration of citrate, TMAO (subject U only), trigonelline and hippurate were subsequently normalized to creatinine and quantified (appendix A7 and appendix A8).

PCA1 was explored in order to investigate any mutual correlation between the observations and different time points. Therefore, scores plots of different combinations of PCs were analysed. However, no noteworthy pattern was observed, since scores were always grouped according to individuals (results not shown).

Another attempt was preceded to examine any relation to the different time points. Here, the mean of sample 1 (collected at time point 0) from both individuals was subtracted from each of the resultant duplicates. Using this matrix, another PCA model was made (PCA2), in order to elucidate a generalization between the values and time point. Again, samples from the two subjects collected at the same time point did not correlate. Thus, this model was rejected (results not shown).

c) Intra-individual differences reflected by PCA plots

Due to the dominance of inter-individual differences, PCA models of intra-individual differences were computed (figure 8 and figure 11). The observations of each subject were divided into two groups: group “No” representing the samples collected before and after the intervention (i.e. U1 and U6 respectively, K1 and K6 respectively) and group “Yes” representing the samples collected during the course of the intervention.

c.i) Intra-individual differences observed in subject U

PCA3 was computed to examine intra-individual differences in subject U, with 3 PCs and $Q^2 = 0.58$.

Scores plots of different combinations of PCs were explored in order to elucidate a link between members of group “No” or “Yes”. No clear grouping was found. The combination reflecting the most noticeable grouping is depicted in figure 8. Some members of group “Yes” are located on the positive side of both PC1 and PC3, and hence distinct from U1 and U6 (green and purple respectively) (figure 8).
Figure 8: PCA3 scores plot (PC1 vs. PC3) of intra-individual differences observed in U, colour legend to the right reflects sample number according to appendix A5. Scores are assigned according to NMR numbers in appendix A5. An ellipse encircles members of group “Yes”.

Figure 9: Corresponding loading plot of PCA3 (PC1 vs. PC3). Green: the metabolites accounting for less variance between group “No” and “Yes”, Red: metabolites accounting for greater variance between group “No” and “Yes”. Loadings related are encircled.

The replicates of U2 (blue) and U3 (red) are not situated in couples as expected, thus hampering the distinction between the groups (figure 8). However, the analysis was still preceded, and the corresponding loading plot of PCA3 was computed (figure 9).

The loadings with the greatest influence on variance between “Yes” and “No” across PC1 and PC3 are coloured red on figure 9, and the corresponding peaks were identified in Chenomx and are marked on figure 10.
Combined it is apparent that metabolites accounting for discrimination between the two groups “No” and “Yes” in subject U are: TMAO/betaine, UN12/3-methylxanthine, N-phenylacetylglucose/phenyl-derivatives, citrate, urea, dimethylamine, UN1, glycolate/1,7-dimethylxanthine and a new unknown compound at 1.9 ppm (UN13) (figure 9 and figure 10). Phenyl-derivatives, hippurate, TMAO/betaine and citrate are seemingly more abundant in samples on the positive side of PC3 (figure 9), thus distinct from U1 situated on the negative side (figure 8). Identification of the majority of highlighted metabolites between [3;4] ppm was neglected due to overlapping peaks (figure 10). The presence of N-phenylacetylglucose in urine is controversial and could be mistaken for phenylacetylglutamine (Wishart et al., 2013). It will therefore also be regarded as a region of phenyls and aromatic derivatives in general terms.

A reduced excretion of trigonelline is found after normalization to creatinine (appendix A7), though this is not reflected in PCA3 (figure 9 and figure 10). Additionally, the concentrations of hippurate, citrate and TMAO/betaine are found to fluctuate during the intervention after normalization to creatinine (appendix A7). Hence, no clear association between variations of the aforementioned metabolites and the intervention is evident when combining both PCA3 and the normalized values. The discrepancies could be due to differences in normalization methods, and the fact that PC1 and PC3 only accounts for 36.6 % and 14 % of the observed variance in subject U respectively. Hence, the variance analysed in figure 8-10 accounts for a minor part of the observed variance. Furthermore, the excretion of creatinine additionally varies to a great extend (appendix A7) and accounts for some variance observed in PCA3 (figure 10).

c.ii) Intra-individual differences observed in subject K
The samples of subject K were also grouped into “No” and “Yes”. 3 PCs were computed with $Q^2 = 0.65$. 

Figure 10: $^1$H NMR spectrum of samples U1-U6. Areas in red reflect metabolites contributing to observed differences between groups “No” and “Yes”. Areas encircled by a black cover a given identification.
The model was explored using scores plots of different PC combinations in order to detect any linkage between group members of “No” or “Yes”. However, no clear association was evident. The combination reflecting the most association between members of the groups is depicted in figure 11.

Figure 11: PCA4 scores plot (PC1 vs. PC2) of intra-individual differences observed in subject K. Colour legend to the right reflects sample number according to appendix A5. Scores are assigned according to NMR number in appendix A5. Members of group “No” are encircled by a red line. Members of group “Yes” are encircled by a black line.

The replicates on figure 11 are well coupled, and group “No” and group “Yes” are somewhat separated across PC2 (accounting for 28.3 % of the total variance) with group “No” located on the negative side. Thus, a loading-line plot of PCA4 was computed and is depicted in figure 12.

The chemical shifts accountable for some of the variation across the PC2 line were identified in Chenomx, and are also denoted on figure 12.

Identified discriminators include: trigonelline, betaine, creatinine, hippurate, urea and UN1, with the latter three being more abundant in K1 and K6 (figure 12).
However, both trigonelline and hippurate normalized to creatinine concentration resulted in highly fluctuating concentrations, with the lowest quantified concentration of trigonelline measured in K6 (48.12 µM/mM creatinine) and the highest measured in K1 (146.62 µM/mM creatinine), and the lowest concentration of hippurate in K6 (435 µM/mM creatinine) and the highest in K2 (1904.28 µM/mM creatinine) (appendix A8). As with subject U, the discrepancies can be reasoned with the different normalization methods and additionally fluctuations in creatinine concentration (appendix A8), as well as the fact that PC2 of PCA4 accounts for 28.3% of the observed intra-individual variance in subject K. Hence, the identified metabolites are not the only source of variation.
4 Discussion

In the present pilot study, the effect of consuming probiotic yoghurts for three weeks was investigated by the use of $^1$H NMR metabolomics in order to elucidate the potential in metabolomics as a tool for characterizing effects of probiotics on the urinary metabolome. In the following discussion section, an assessment of the overall results through a metabolomic approach is presented. Furthermore, associations between the results and the probiotic treatment will be examined by reviewing literature regarding metabolomic studies on the effect of pre- and/or probiotics. Lastly, the discussion will bring about some methodological considerations regarding the current study.

a) Characteristics of urine profiles

A table of frequently identified peaks is found in appendix A6, here among are creatinine, urea, hippurate and citrate as expected from section c.ii Urine metabolites.

As stressed in c.ii, different food constituents can affect the metabolic composition of urine. This influence was apparent in the current study, more specifically concerning TMAO/betaine of choline metabolism, dimethylamine of choline metabolism, tartrate, phenyl derivatives, dimethyl sulfone, trigonelline and xanthine derivatives.

The identification of betaine and trimethylamine-N-oxide should be considered with caution. Betaine and TMAO both have a singlet peak at around 3.26 ppm. The peaks overlap, which makes it hard to distinguish between the two compounds. Both can be derived from dietary choline, though differ in metabolic fate as visualised in figure 13. The path for betaine production is mainly mammalian, whereas the path of TMAO formation is sym-xenobiotic, therefore both mammalian and microbiological (Xie et al., 2013). Additionally, TMAO derives directly from dietary intake of fish (Lenz et al., 2004).

![Figure 13: Metabolism of choline, adapted from Martin et al. (2009). Red ellipses encircle betaine and TMAO.](image-url)
In order to clarify the identification, one could either spike a sample with e.g. TMAO, or one could acquire a heteronuclear [¹H,¹³C] 2D NMR spectrum (see appendix A10 for 2D spectra). Various xanthine derivatives were identified, including theophylline, 3-methylxanthine and 1,3 dimethylurate (results not shown, confer to appendix A6). Following pattern was observed: when xanthine derivatives were detected by automatic fit, caffeine could also be detected, albeit with some uncertainty. The individual metabolites were analysed by HMDB (Wishart et al., 2013) and KEGG (Kanehisa et al., 2017), and it is hypothesised that the metabolites derive from the metabolism of caffeine when present in the same sample. Rothwell et al. (2014) also suggested a relation between coffee consumption and caffeine derived metabolites, with the greatest discriminators of difference being 1-methylxanthine, 1,7-dimethylerate, paraxanthine and theophylline among others, thus further supports the generalization made based on the present results. In order to verify the hypothesis, one could make a correlation of the metabolites, though this is beyond the scope of this bachelor thesis. Dimethyl sulfone originates from intestinal metabolism and foodstuff containing sulfoxides (e.g. onions) (Winning et al., 2009). These are metabolised in the liver and kidneys. Tartrate is a food additive, but also found in wine. When ingested, tartrate is metabolized by certain colonic bacteria species. A minor part is excreted in the urine (Wishart et al., 2013). Trigonelline is both involved in regeneration of S-adenosylmethionine, found in coffee and a product of niacin metabolism (Lang et al., 2011; Wishart et al., 2013). Furthermore, niacin itself derives from diet constituents and to some extent from metabolism of gut flora (Wishart et al., 2013). Additionally, as stressed in c.ii, the urinary metabolic profile is highly affected by the composition of the gut flora. In the present study, identified metabolites related to microbiota activity were: trimethylamine N-oxide, hippurate, tartrate, dimethylamine and 2-hydroxyisobutyrate. Thus greatly associated with diet-derived metabolites. The formation of hippurate occurs in 2 steps. The gut flora metabolizes dietary polyphenols and other aromatics to benzoate. Subsequently, benzoate is conjugated with glycine in the liver or kidneys, resulting in hippurate (Williams et al., 2010). Hippurate is found to correlate with Clostridia spp. (Li et al., 2008), be reduced in patients with irritable bowel diseases (Dawiskiba et al., 2014; Williams et al., 2010) and to be dependent on dietary intake of polyphenols or benzoate (Krupp et al., 2012; Mulder et al., 2005; Walsh et al., 2006). Additionally, hippurate has been consistently reported to be negatively correlated with obesity in animal and human studies, and to vary between individuals (Holmes, 1994; Holmes et al., 2011). Thus, the production of hippurate might be related to a certain metabolome or diet, and lower levels of hippurate excretion could be an indication of dysbiosis in the gut.
Combined, the urinary metabolic profiles in the present study are greatly influenced by diet and the gut flora.

Inclusion of PCA1 allowed for clear separation between individuals, suggesting how these are subject to different environmental factors. This inter-individual difference accounts for the greatest variation (reflected by PC1, figure 6). By analysis of the loading-line plot (figure 7) discriminators include creatinine, urea, hippurate, citrate, trigonelline and UN1. The sources of the majority of these metabolites have already been considered. Urea is a product of protein catabolism and produced after deamination of amino acids. Creatinine excretion correlates with body and muscle mass and found to be associated with a high-protein diet (Rasmussen et al., 2012; Stella et al., 2006).

A study by Walsh et al. (2006) found how standardization of diet reduced variation in hippurate and creatinine in urine. The authors highlight the influence of non-nutrient food components as a source of variation, thus stressing the impact of diet on the urinary metabolic profile. The observed inter-individual differences are therefore ascribed to gender differences, differences in diet and to other environmental and physiological differences.

To sum up: analysis of the urine samples by 1H NMR led to identification of some common metabolites between individuals, albeit those found to be associated with gender differences and diet predominantly attribute to great inter-individual differences. Hence, no metabolites indicate clear changes in gut flora activity thus far.

Therefore, it became relevant to explore the intra-individual differences in terms of changes in metabolite composition as preceded by PCA3 (subject U, figure 8) and PCA4 (subject K, figure 11). Identified metabolites accountable for some intra-individual differences in subject U were: TMAO/betaine, citrate, phenyl derivatives, dimethylamine, hippurate, urea and creatinine, and in subject K: betaine, trigonelline, hippurate, urea and creatinine.

The following section seeks to explore any associations between the aforementioned discriminators and probiotic intervention. The most prominent identified gut-derived metabolites contributing to variation reflected by PCA3 and PCA4 are considered.

These include: TMAO/betaine, citrate and phenyl derivatives (identified as N-phenylacetylglycine) for subject U (figure 9 and figure 10) and hippurate and trigonelline for subject K (figure 12).

b) Associations between observed discriminators and probiotic intake

A typical biomarker of changes in gut flora activity is the concentration of SCFAs measured. As stated in section b.i, probiotic activity is mainly saccharolytic, thus they ferment non-digestible carbohydrates and induce production of SCFAs. The SCFAs are absorbed across the gut epithelium layer or excreted with faeces.
In the present study, a targeted approach was applied to all of the spectra before multivariate analysis in order to assess whether probiotic intake changed the excretion of SCFAs in the urine and hence, could indicate changes in microbial activity. However, no SCFAs were identified (appendix A6).

Zheng et al. (2011) successfully detected butyrate in urine and faeces samples of rats by the use of GC-MS and LC-MS. Additionally, Verbeke et al. (2010) successfully detected $^{13}$C-SCFAs in human urine samples by the use of GC-MS after supplementation of radio-active labelled non-starch polysaccharides. In another study, Boets et al. (2016) investigated the availability of colonic derived SCFAs in healthy adults after supplementation of labelled $^{13}$C-SCFAs and detection by GC-MS, and found that only 0.05% of supplemented SCFAs were found in urine.

The aforementioned studies and the results of the current study thus suggest at least 3 following hypotheses: 1) $^1$H NMR might be too insensitive to identify and quantify SCFAs in human urine, 2) SCFAs in urine are suitable biomarkers of changes in microbial activity, but the study design was insufficient or the commercial product was not able to affect the composition of microbiota, 3) SCFAs are not suitable biomarkers of probiotic supplementation when analysing human urine samples by $^1$H NMR, and other metabolites in urine might be suitable biomarkers. The 2 former hypotheses will be considered in Methodological considerations, whereas the latter will be discussed in the following.

Studies using metabolomics in order to elucidate potential effects of pre- or probiotic treatment in humans are not plentiful, therefore the following table includes animal studies on the effect of pre- or probiotic supplementation.

### Table 3: Intervention studies using metabolomic techniques in order to elucidate the effect of pre- or probiotic supplementation

<table>
<thead>
<tr>
<th>Prebiotic or probiotic used</th>
<th>Study design</th>
<th>Overall results</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>L. rhamnosus Galactooligosaccharides Combined (synbiotic)</td>
<td>$^1$H NMR (HBM mice, urine, serum) Microbial Profiling (fecal and jejunal contents) With controls</td>
<td><strong>Urine:</strong> Probiotic: HL: α-ketoisovalerate, TMA LL: butyrate, isovalerate Prebiotic: HL: carnitine, acylcarnitine, taurine LL: α-ketoisocaproate, butyrate, lysine, α-aminoacidipate Synbiotic: Similar to prebiotic, though increase in creatine and α-ketoisovalerate</td>
<td>Martin et al. (2009)</td>
</tr>
<tr>
<td>L. paracasei or L. rhamnosus</td>
<td>Changes in metabolite profile of multiple biological compartments (HBM mice, plasma, urine, liver, fecal) Supplemented with L. rhamnosus or L. paracasei</td>
<td><strong>Urine:</strong> HL: tryptamine, indoleacetylglucine, phenylacetylglucine LL: creatine, citruline (L. rhamnosus), LL: 3-hydroxyisovalerate (L. paracasei), LL: α-ketoisocaproate, citrate</td>
<td>Martin et al. (2008a)</td>
</tr>
<tr>
<td>Probiotic</td>
<td>Prebiotics</td>
<td>Metabolomic Study</td>
<td>Biomarkers/Results</td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
<td>-------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Changes in metabolic profile and microbiome after consumption of prebiotics</td>
<td>'H NMR (human, faeces) n = 12 PC</td>
<td>Residues of PDX in faeces Positive correlation between acetate and <em>Bacteriodes</em> in faeces Negative correlation between bifidobacteria and succinate and acetate</td>
<td>Lamichhane et al. (2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Bifidobacterium longum</em> Lactobacillus acidophilus and prebiotic (FOS) (synbiotic)</td>
<td>'H NMR (human faeces) n = 16 1 month In vitro study</td>
<td>Less change in lactobacilli counts LL: Succinate, tyrosine, phenylalanine, alanine, glutamate, lysine, glycine, valine, isoleucine HL: propionate, butyrate, acetate, lactate Favor of amino acid assimilation rather than catabolism</td>
<td>Ndagijimana et al. (2009)</td>
</tr>
<tr>
<td><em>Lactobacillus salivarius</em> PS2</td>
<td>'H NMR (women with mastitis, urine) n = 23 BL-C</td>
<td>HL: hippurate, TMAO, creatine</td>
<td>Vázquez-Fresno et al. (2014)</td>
</tr>
<tr>
<td><em>Lactobacillus acidophilus</em> La5 <em>Bifidobacterium lactis</em> Bb12</td>
<td>'H NMR (mice, urine and faeces) Aged mice Adult mice</td>
<td>Adult mice: HL: N-methylnicotinamide, dimethylglycine, choline LL: sarcosine, phenylacetylglycine, nicotinate, α-ketoisocaproate, 2-oxo-4-methylvalerate</td>
<td>Brasili et al. (2013)</td>
</tr>
</tbody>
</table>
Aged mice:
HL: dimethylglycine, butyrate, 3-hydroxyisovalerate, 2-oxo-4-methylvalerate, threonine
LL: succinate, N-methylnicotinamide, citrate
Response dependent on age

<table>
<thead>
<tr>
<th>Dietary inulin-type fructans</th>
<th>1H NMR (Obese women, urine and plasma) qPCR analysis of 16s rDNA (faeces) DB, PC n = 30</th>
<th>Increased Bifidobacterium and Faecalibacterium prausnitzii No clear clustering of plasma and urine metabolic profiles between subjects Positive correlation between higher urinary hippurate and Collinsella</th>
<th>Dewulf et al. (2013)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potato fibres</td>
<td>Lactose-[15N2]ureide degradation Isotope ratio mass spectrometry (human, 15N in blood, urine, faeces) n = 14</td>
<td>LL: renal excretion by supplementation of fibres (potato or pea) vs. control Prebiotic intake lowered colonic generation and renal excretion of 15NH3</td>
<td>Wutzke et al. (2010)</td>
</tr>
<tr>
<td>Pea starch</td>
<td>GC-MS (IBD patients, serum) 8 w n = 61 Consumption of acidified milk with/without probiotics</td>
<td>Effect independent of probiotic supplementation HL: lactate, glutamine, proline, creatinine/creatinine, aspartic acid LL: glucose</td>
<td>Pedersen et al. (2011)</td>
</tr>
<tr>
<td>Lactobacillus acidophilus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lactobacillus paracasei F19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lactobacillus acidophilus LA-5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bifidobacterium lactis BB-12</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** BL-C, baseline controlled; C-O, cross-over; DB, double-blinded; HBM, human baby microflora; HL, higher level; IBD, irritable bowel disease; LL, lower level; OF-IN, oligofructose-enriched inulin; PC, placebo-controlled; Pre, prebiotic, R, randomized, w, weeks.

Table 3 demonstrates that changes in the concentration of some discriminators in the present study have previously been linked to pre- or probiotic treatment, and that other metabolites than SCFAs are identified following an intervention.

Preter et al. (2004) and Preter et al. (2007) investigated the use of labelled supplementation of lactose-[15N]ureide and [3H4]tyrosine as biomarkers of changes in microbial activity towards saccharolytic fermentation following supplementation of *Lactobacillus Casei* Shirota and Lactoluse (Preter et al., 2004) and *Lactobacillus casei* Shirota, *Bifidobacterium breve* Yakult and OF-IN (Preter et al., 2007).

The authors found reduced urinary excretion of 15N and p-cresol, thus supplying indications of other biomarkers than production of SCFAs. These findings are further substantiated by Wutzke et al. (2010) and indicated by Ross et al. (2013), who found that intake of a whole-grain diet (though not assigned prebiotic) led to changes in protein metabolism in men and lowered the urinary urea excretion. Additionally, Martin et al. (2008b) observed a decrease in phenylacetylglycine following prebiotic supplementation and related the results to indications of decreased proteolytic activity. Contradictory, Martin et al. (2008a) reported an increase in specific phenyl and indole derivatives by supplementation of probiotics alone (table 3). The authors highlight how these derivatives imply proteolytic fermentation in the gut and are linked to the metabolism of aromatic amino acids.
In the current study, the region of phenolic and aromatic derivatives was assigned to contribute to variation observed in subject U (figure 10). These peaks were initially identified as N-phenylacetylglucose, a glycine conjugate of phenylacetic acid metabolised through phenylalanine metabolism (Wishart et al., 2013). The metabolism has been correlated to Bacteroides, Clostridia and E. Coli in vitro (Smith & Macfarlane, 1996), thus demonstrating the compound’s association with the composition of gut flora. However, the peaks could also reflect other aromatic and phenolic derivatives, and therefore not only a product of protein metabolism. Lee et al. (2006) demonstrated that certain phenolic derivatives of black tea suppress the growth of pathogenic bacteria found in faeces in vitro. Albeit not completely comparable with the current study, it is interesting that an increase in the concentration of phenolic derivatives has proven to be an indication of modulated gut flora, thus the variation observed could indicate changes in microbial activity in agreement with Martin et al. (2008a) and Lee et al. (2006). However, Martin et al. (2008a) also found reduced level of 3-hydroxyisovalerate following supplementation of L. paracasei and reduced level of citrate. This is contradictory with the current study. Citrate is identified as a contributor to variance (figure 10) with increasing concentration during the course of the intervention (appendix A9). Thus, the observed changes in phenolic derivatives/N-phenylacetylglucose are more likely due to diet and not probiotic supplementation.

Combined, these findings indicate that mainly prebiotic intake alters the metabolism of the gut flora towards increased saccharolytic activity, perhaps owing to the nature of their structure. An increase in energy available for saccharolytic fermentation leads to greater demand of nitrogen sources for bacterial growth, and hence reduces the formation of potential toxic compounds and the excretion of nitrogen in the urine (Cummings & Macfarlane, 1991; Scheppach et al., 2001).

Probiotic supplementation alone has not proven to affect saccharolytic activity as much (Martin et al., 2008a; Wutzke et al., 2010).

In the present study, trigonelline was identified to contribute to intra-individual differences in subject K (figure 12).

Trigonelline has been found to be a constituent in coffee (Ashihara et al., 2014; Lang et al., 2011). A study found that 50 % of ingested trigonelline was excreted 0-8 h after coffee consumption (Lang et al., 2010). In the present study, samples were taken in the early morning, 12 h past last coffee consumption. It might also be expected that other biomarkers of coffee consumption and caffeine metabolism were contributors to discrimination, if coffee was the only factor accountable for differences. Therefore, it is relevant to explore other sources of urinary trigonelline excretion.

Trigonelline derives both from the metabolism of vitamin B3 (niacin), and from the conversion of S-adenosylmethionine (SAM) to S-adenosylhomocysteine. The latter cycle is linked to the metabolism of betaine (figure 13).
Brasili et al. (2013) found indications of modulations of betaine-homocysteine metabolism and NAD metabolism after probiotic supplementation in mice of different age by 1H NMR analysis of urine and faeces. The probiotic treatment induced urinary excretion of dimethylglycine in both adult and aged mice. Additionally, the authors linked probiotic supplementation to modulation of other pathways regulating homocysteine amounts, namely methylation of nicotinate to N-methylnicotinate (trigonelline) and N-methylnicotinamide. Combined, the results indicated changes in betaine-homocysteine and NAD metabolism. The authors also found increased concentration of 3-hydroxyisovalerate among others, though no such changes were identified in the current study, and opposing the findings of Martin et al. (2008a).

Interestingly, by multicompartmental metabolic profiling, Martin et al. (2009) found modulation of transmethylation pathways (betaine-homocysteine metabolism) after supplementation of pre- and probiotics in HBM mice linked to the regeneration of SAM. They argued that probiotic supplementation induced the conversion of homocysteine to methionine, "crucial for maintaining methionine homeostasis, detoxifying homocysteine and producing S-adenosylmethionine" (Martin et al., 2009). Additionally, other studies have linked reduced trigonelline excretion with depletion of SAM and increased oxidative stress (Calvani et al., 2010; Sun et al., 2008). In the present study, it is not possible to clearly relate the observed differences in trigonelline to modulations of betaine-homocysteine metabolism, since there is no indication of induced production of other derivatives, which might be expected considering the aforementioned studies, but it is interesting that trigonelline - inter alia produced from demethylation of SAM to S-adenosylcysteine, is found to contribute to variation (figure 12).

1-methylnicotinamide (another derivative of nicotinate metabolism) was identified in the current study (appendix A6), but not assigned to contribute to observed variation in the present PCA model (figure 12). The variation in trigonelline concentration could also be an indication of modulation of NAD metabolism, and it can be speculated, whether probiotic supplementation results in modulations of homocysteine and NAD metabolism as observed in the aforementioned studies, and whether this modulation has an impact on health. However, after quantification of trigonelline, no clear pattern of variation is observed in subject K (appendix A8).

Contradictory, subject U excreted less trigonelline during the course of the intervention (appendix A7), but the differences were not identified in PCA3 (figure 10).

Considering the potential discriminator TMAO, Martin et al. (2008a) found increased level of TMA and TMAO in the liver of mice supplemented with Lactobacillus rhamnosus, indicating modulation in methylamine and choline metabolism. The authors referred to Allison & Macfarlane (1989) and argued that supplementation of L. rhamnosus induced production of methylamines by Bacterioides and Clostridium perfringens, thus modulating the gut flora.
*L. paracasei* was found to have different metabolic effect on choline and not related to increased TMAO production. As stressed previously, TMAO originates from more than one source, and the compound could be mistaken for betaine in the current study. Despite an observed variation (figure 10), evidence does not support a relation between the changes in TMAO/betaine concentration and probiotic intake in the current study.

Citrate was also identified as a potential discriminator in subject U (figure 10). Citrate is an intermediate in the TCA cycle and derives from dietary sources. Additionally, citrate is found to positively correlate with milk intake (Zheng et al., 2015). After quantification of citrate no clear association to the intervention became evident, since the quantitatively greatest concentration of citrate was measured in the post-intervention sample (U6, 411.09 µM/mM creatinine, appendix A7). Citrate was also found in increasing but diminishing amounts in subject K (appendix A8). The reason for a seemingly elevated excretion of citrate could be due to an increased consumption of carbohydrate or from the intake of fermented yoghurts, since no other TCA cycle intermediates were identified as contributors of variation.

In the present study, the level of hippurate led to intra-individual differences in subject K (figure 12). The excretion ranged from 814.52 µM/mM creatinine to 1904.28 µM/mM creatinine, with a lower level of hippurate during the course of the intervention (appendix A8). This is contradictory with other studies, which have found increased level of hippurate after pre- or probiotic intake and further associated this increase with an improvement of gut-health (Dewulf et al., 2013; Vázquez-Fresno et al., 2014). The referred studies are not fully comparable, since one concerns obese women and the other concerns women with mastitis, but it remains interesting that the level of a human-gut co-metabolite is found in decreasing amounts during the current intervention.

This could indicate either changes in the composition and activities of the gut flora owing to the aforementioned correlation to certain species, or simply changes due to diet as observed in other studies (Krupp et al., 2012; Mulder et al., 2005; Walsh et al., 2006). Since the excretion of hippurate varies to a great extend between K1 and K2 (representing day to day changes, appendix A8), it remains more likely that the variation is due to diet, as likewise suggested in a study by Walsh et al. (2006).

To sum up: the present pilot study did not allow for any clear elucidation of the effect of probiotics by the use of $^1$H NMR based metabolomics. The metabolomic approach revealed great inter- and intra-individual differences, especially concerning gut flora derived metabolites as hippurate, TMAO and phenyls and diet related metabolites. A relation to intake of a probiotic drink cannot be verified, since the results were not unanimously related to specific changes in urinary metabolic profile and could not be fully substantiated by literature. Additionally, by normalizing the results of the current study to creatinine excretion, it is not possible to elucidate clear associations.
It was therefore not possible to verify hypothesis 3, but it remains interesting that the use of labelled compounds has proven to be a suitable biomarker of changes in microbial activity following mainly prebiotic supplementation.

The most obvious metabolites contributing to differences were discussed. These differences only account for 14 % and 28.3 % of intra-individual differences in subject U and K respectively, indicating a relative minor impact on the total variance. It cannot be rejected that other metabolites contributing to variation are linked to probiotic supplementation.

Literature indicates a relation to specific metabolites after pre- or probiotic ingestion, but there is a great lack of human intervention studies. Based on table 3 it is additionally hard to decipher consistency between the outcomes of studies using metabolomics, thus stressing a significance of study design and the prebiotic or probiotic strain investigated.

Furthermore, the microbial activity in the colon is highly influenced by the nature of substrate available as stressed in section a.i. As already discussed and indicated by literature, induction of saccharolytic fermentation is dependent on appropriate carbohydrate sources e.g. in terms of prebiotics. Thus, the intake of a synbiotic combination could lead to greater impact on microbial activity.

c) Methodological considerations

c.i) Some considerations regarding NMR vs. MS

Several review papers give insight in factors, which can be considered when comparing NMR spectroscopy and MS, here among are Wishart (2014) and Lindon & Nicholson (2008). Factors of importance in regards to the outcome and procedure of a metabolomic analysis include sample size, preservation of sample integrity and sensitivity. Samples for NMR require relative little preparation, however, there is a need of greater sample size (0.5μL) than GC-MS and LC-MS (50μL and 10μL respectively) (Wishart, 2014). Thus, some biological samples are harder to analyse by NMR (e.g. urine of mice) due to sample size. Though, NMR based metabolomic analysis is widely used owing to a relatively easy data acquisition and the non-destructive nature, thereby maintaining sample integrity.

Opposite, samples for GC-MS and LC-MS are most often chemically modified and derivatized to gas phase and liquid phase respectively before analysis (Lindon & Nicholson, 2008).

Whereas NMR can detect 50-75 compounds in a human biofluid, GC-MS is assigned 50-100 compounds and LC-MS “hundreds to thousands of features” (Wishart, 2014). GC-MS and LC-MS are therefore more sensitive techniques than NMR and therefore hypothesis 1 (NMR might be too insensitive to detect SCFAs in urine) cannot be excluded.
However, an advance concerning NMR is the ability to detect molecular structures rather than ion masses. In MS analysis, the output corresponds to ion masses, which might fit multiple compounds and therefore be harder to identify and separate (Claus & Swann, 2013).

Other noteworthy factors relevant for both NMR and MS include differences in pH, temperature and salt, since these can influence the chemical analysis (Wishart, 2008).

c.ii) Some considerations regarding study design

Referring hypothesis 2, the present study design has multiple drawbacks, and only some aspects will be elaborated. First: the pilot study includes only two participants, uses no controls and only spans the course of 21 days. The design is therefore not statistically reasoned and the findings cannot be generalized. Additionally, a lack of criteria for the participants in regards to factors such as lifestyle and diet presents a challenge when wanting to elucidate biases. However, the producer of the probiotic drink involved claims that the drink is beneficial “as part of a regular healthy diet” (appendix A2), therefore screening and criteria assessment should prove to be negligible, since potential health benefits are be mutual between healthy subjects. However, neither the current study nor the producer has clearly defined term “healthy”, but this thesis has sought to elucidate different health effects observed by other studies.

As indicated by the former discussion on the identified discriminators in the present study, it is difficult to distinguish whether a change in metabolite composition is due to changes in gut flora activity induced by intake of probiotics, or induced by differences in diet. Hence, the study would benefit from a standardized diet or at least a dietary record in order to identify potential biases. Additionally, factors as age, weight, gender, etc. must be taken into account, as it is evident that these influence the metabolic profile of urine. A cross-over, placebo controlled study with a greater number of participants and approval of the Ethical Comitee in accordance with the Declaration of Helsinki would provide more statistical power and a validation for an intervention.

Multiple chemical techniques and multiple biological sources can be used in order to interpret the outcome of an intervention, as demonstrated by studies in table 3. In the current study, this notion is relevant in regards to hypothesis 2, since the use of one biological source and one chemical technique is evidently insufficient in order to achieve the ideal goal of elucidating an effect of an intervention on the entire metabolome (Xie et al., 2013).
5 Conclusion and perspectives
The present pilot study demonstrates the use of $^1$H NMR urine metabolomics to investigate the effect of probiotic intake on urinary metabolic profile. The study found variations in creatinine, citrate, hippurate and trigonelline excretions during the course of the intervention. However, the variations could not reasonably be related to intake of probiotics. Multiple methodological factors must be improved in order to fully confirm these findings. Other metabolomics studies prove significant but inconsistent results dependent on the supplement consumed and study design. However, several of these studies have assigned certain metabolites as biomarkers of changed microbial activity in the gut, and indicators of modulations in biochemical pathways in the host following an intervention. Hence, metabolomics has proved promising potential in improving our understanding of the biochemical mechanisms behind pre- and/or probiotic activities. Nevertheless, there is a need of greater human intervention trials and a potential in holistic integration of multiple biological compartments, when wanting to understand the complete biochemical effects of pre- and probiotic action on host’s health.

6 Acknowledgements
I will like to thank my supervisor Ulrik Sundekilde and co-supervisor Jette F. Young for taken the time to help me during the making of this bachelor thesis. Special thanks to Ulrik Sundekilde for given great guidance and for volunteering as a participant in the pilot-study. Thanks to Christian Clement Yde for initially setting up the project.
7 Literature


Megavariate Data Analysis (pp. 43–70).


Manas, M; Martinez de Victoria, E; Gil, A; Yago, M; Mathers, J. (n.d.). Nutrition and Metabolism. In H. M. Gibney, Michael J; Macdonald, Ian A; Roche (Ed.) (pp. 218–222).


https://doi.org/10.1021/tr0340293


Nutrition, 62(2), 225–231. https://doi.org/10.1038/sj.ejcn.1602706


Metabolomics investigation to shed light on cheese as a possible piece in the French paradox puzzle. *Journal of Agricultural and Food Chemistry*, 63(10), 2830–2839. https://doi.org/10.1021/jf505878a
Appendix

A1:
Table 1.A: Chemical shifts of protons in different molecules, adapted from Lancashire (2003).

<table>
<thead>
<tr>
<th>Type of proton</th>
<th>Chemical shift (δ ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alkyl, RCH₃</td>
<td>0.8-1.0</td>
</tr>
<tr>
<td>Alkyl, R₂CH₃</td>
<td>1.2-1.4</td>
</tr>
<tr>
<td>Alkyl, R₃CH</td>
<td>1.4-1.7</td>
</tr>
<tr>
<td>Allylic, R₂C=CRCH₃</td>
<td>1.6-1.9</td>
</tr>
<tr>
<td>Benzyllic, ArCH₃</td>
<td>2.2-2.5</td>
</tr>
<tr>
<td>Alkyl chloride, RCH₂Cl</td>
<td>3.6-3.8</td>
</tr>
<tr>
<td>Alkyl bromide, RCH₂Br</td>
<td>3.4-3.6</td>
</tr>
<tr>
<td>Alkyl iodide, RCH₂I</td>
<td>3.1-3.3</td>
</tr>
<tr>
<td>Ether, ROCH₂R</td>
<td>3.3-3.9</td>
</tr>
<tr>
<td>Alcohol, HOCH₂R</td>
<td>3.3-4.0</td>
</tr>
<tr>
<td>Ketone, ROCOH₂</td>
<td>2.1-2.6</td>
</tr>
<tr>
<td>Aldehyde, RCOH</td>
<td>9.5-9.6</td>
</tr>
<tr>
<td>Vinylec, R₂C=CR₂H</td>
<td>4.6-5.0</td>
</tr>
<tr>
<td>Vinylec, R₂C=CRH</td>
<td>5.2-5.7</td>
</tr>
<tr>
<td>Aromatic, ArH</td>
<td>6.0-9.5</td>
</tr>
<tr>
<td>Acetylenic, RC=CH₂</td>
<td>2.5-3.1</td>
</tr>
<tr>
<td>Alcohol hydroxy, ROH</td>
<td>0.5-6.0</td>
</tr>
<tr>
<td>Carboxylic, RCOOH</td>
<td>10-13*</td>
</tr>
<tr>
<td>Phenolic, ArOH</td>
<td>4.5-7.7*</td>
</tr>
<tr>
<td>Amino, R-NH₂</td>
<td>1.0-5.0*</td>
</tr>
</tbody>
</table>

A2: Ingredients and brand:

Photo 1.A: Actimel Danone strawberry yoghurt drink, ingredients and brand, (Poulsen, 2017)

Composition of yoghurt:
Translated ingredients list from Danish to English based upon photo 1.A:

Ingredients: Pasteurised milk, sugar, strawberries 2.1 %, glucose, milkminerals, modified tapioca starch, natural aroma, acidity regulator (sodium citrate), yoghurtculture, L. Casei Danone-culture (Lactobacillus paracasei subsp. Paracasei (DN-114001/CNCM I-1518)), vitamins (B6, D)
Photo 2.A: Recommended daily intake of Actimel, Danone strawberry yoghurt drink (Poulsen, 2017)

Translated recommendations from Danish to English based upon the highlighted section of photo 2.A:

“1 bottle a day as part of a balanced diet and healthy lifestyle”

Below:

“Actimel contains vitamin B6 and D, which help maintaining normal function of the immune system”

A3: Set-up of pilot study

Table 2.A: Set-up of pilot study

<table>
<thead>
<tr>
<th>Date</th>
<th>Time point in number of days</th>
<th>Sample, S</th>
</tr>
</thead>
<tbody>
<tr>
<td>17/2</td>
<td>0</td>
<td>I</td>
</tr>
<tr>
<td>18/2</td>
<td>1</td>
<td>II</td>
</tr>
<tr>
<td>19/2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>20/2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>21/2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>22/2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>23/2</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>24/2</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>25/2</td>
<td>8</td>
<td>III</td>
</tr>
<tr>
<td>26/2</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>27/2</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>28/2</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1/3</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>2/3</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>3/3</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>4/3</td>
<td>15</td>
<td>IV</td>
</tr>
<tr>
<td>5/3</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>6/3</td>
<td>17</td>
<td></td>
</tr>
</tbody>
</table>
### A4: Ingredients of Master Mix with corresponding concentrations

Table 3.A: Composition of stock buffer used for chemical analysis of urine samples

<table>
<thead>
<tr>
<th>Ingredients</th>
<th>11.7x stock (100mL)</th>
<th>Concentration</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na₂HPO₄</td>
<td>10.377 g</td>
<td>583 mM</td>
<td>50 mM</td>
</tr>
<tr>
<td>DSS</td>
<td>0.065 g</td>
<td>2.9 mM</td>
<td>0.25 mM</td>
</tr>
<tr>
<td>Sodium formate</td>
<td>0.0079 g</td>
<td>1.162 mM</td>
<td>100 µM</td>
</tr>
</tbody>
</table>

### A5: Sample IDs

Table 4.A: ID of samples for NMR analysis, K, Katrine; U, Ulrik

<table>
<thead>
<tr>
<th>Sign on NMR tube</th>
<th>ID</th>
<th>Corresponding sample, S</th>
<th>Sample number</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>U1</td>
<td>I</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td>U2</td>
<td>II</td>
<td>2</td>
</tr>
<tr>
<td>III</td>
<td>U3</td>
<td>III</td>
<td>3</td>
</tr>
<tr>
<td>IV</td>
<td>U4</td>
<td>IV</td>
<td>4</td>
</tr>
<tr>
<td>V</td>
<td>U5</td>
<td>V</td>
<td>5</td>
</tr>
<tr>
<td>VI</td>
<td>U1</td>
<td>I</td>
<td>1</td>
</tr>
<tr>
<td>VII</td>
<td>U2</td>
<td>II</td>
<td>2</td>
</tr>
<tr>
<td>VIII</td>
<td>U3</td>
<td>III</td>
<td>3</td>
</tr>
<tr>
<td>IX</td>
<td>U4</td>
<td>IV</td>
<td>4</td>
</tr>
<tr>
<td>X</td>
<td>U5</td>
<td>V</td>
<td>5</td>
</tr>
<tr>
<td>U1</td>
<td>U6</td>
<td>VI</td>
<td>6</td>
</tr>
<tr>
<td>U2</td>
<td>U6</td>
<td>VI</td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td>K1</td>
<td>I</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>K2</td>
<td>II</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>K3</td>
<td>III</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>K4</td>
<td>IV</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>K5</td>
<td>V</td>
<td>5</td>
</tr>
<tr>
<td>K1</td>
<td>K6</td>
<td>VI</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>K1</td>
<td>I</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>K2</td>
<td>II</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>K3</td>
<td>III</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>K4</td>
<td>IV</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>K5</td>
<td>V</td>
<td>5</td>
</tr>
<tr>
<td>K2</td>
<td>K6</td>
<td>VI</td>
<td>6</td>
</tr>
</tbody>
</table>
### A6: Identification of metabolites

#### Table 5.A: Frequently identified metabolites from the 24 spectra

<table>
<thead>
<tr>
<th>Identified metabolite</th>
<th>Chemical shifts (multiplicity)</th>
<th>Assignments</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-hydroxyisovalerate</td>
<td>2.36(s), 1.3(s)</td>
<td>CH₂, CH₃</td>
</tr>
<tr>
<td>UN1</td>
<td>2.33(s)</td>
<td>-</td>
</tr>
<tr>
<td>Citrate</td>
<td>2.5(d), 2.7(d)</td>
<td>CH₂</td>
</tr>
<tr>
<td>Dimethylamine</td>
<td>2.72 (s)</td>
<td>CH₃</td>
</tr>
<tr>
<td>Succinate</td>
<td>2.4 (s)</td>
<td>CH₃</td>
</tr>
<tr>
<td>Alanine</td>
<td>1.48(d), 3.8 (q)</td>
<td>CH₃, CH₆</td>
</tr>
<tr>
<td>2-hydroxyisobutyrate</td>
<td>1.35(s)</td>
<td>CH₃</td>
</tr>
<tr>
<td>UN2</td>
<td>2.28-2.25</td>
<td>-</td>
</tr>
<tr>
<td>Methylamine</td>
<td>2.60(s)</td>
<td>CH₃</td>
</tr>
<tr>
<td>Creatinine</td>
<td>3.04(s), 4.1(s)</td>
<td>CH₂, CH₃</td>
</tr>
<tr>
<td>Malonate</td>
<td>3.11(s)</td>
<td>CH₂</td>
</tr>
<tr>
<td>Dimethylsulfone</td>
<td>3.13(s)</td>
<td>CH₃</td>
</tr>
<tr>
<td>UN3</td>
<td>3.19(s)</td>
<td>-</td>
</tr>
<tr>
<td>Betaine</td>
<td>3.25(s), 3.9(s)</td>
<td>CH₂, CH₂</td>
</tr>
<tr>
<td>Trimethylamine N-oxide (TMAO)</td>
<td>3.26(s)</td>
<td>CH₃</td>
</tr>
<tr>
<td>1,3 dimethylurate</td>
<td>3.31(s), 3.44(s)</td>
<td>CH₂, CH₃</td>
</tr>
<tr>
<td>UN4</td>
<td>3.35(s)</td>
<td>-</td>
</tr>
<tr>
<td>4-hydroxyphenylacetate</td>
<td>3.44(s), 6.9 (d), 7.2(d)</td>
<td>CH₂, ArH, ArH</td>
</tr>
<tr>
<td>3-methylxanthine</td>
<td>3.5(s), 8.0 (s)</td>
<td>CH₂, CH</td>
</tr>
<tr>
<td>Glycine</td>
<td>3.56(s)</td>
<td>CH₂</td>
</tr>
<tr>
<td>Glycolate</td>
<td>3.95(s)</td>
<td>CH₂</td>
</tr>
<tr>
<td>Hippurate</td>
<td>3.96(d), 7.54(-), 7.63(tt), 7.82(dd), 8.5(s)</td>
<td>CH₂, ArH, ArH, ArH, NH</td>
</tr>
<tr>
<td>Tartrate</td>
<td>4.34(s)</td>
<td>CH</td>
</tr>
<tr>
<td>Trigonelline</td>
<td>4.43(s), 8.07(t), 8.82(d), 8.84(d), 9.0 (s)</td>
<td>CH₂, ArH, ArH, ArH, ArH</td>
</tr>
<tr>
<td>Urea</td>
<td>5.8(s)</td>
<td>NH₂</td>
</tr>
<tr>
<td>UN5</td>
<td>7.19(s)</td>
<td>-</td>
</tr>
<tr>
<td>UN6</td>
<td>7.20(s)</td>
<td>-</td>
</tr>
<tr>
<td>UN7</td>
<td>7.48(s)</td>
<td>-</td>
</tr>
<tr>
<td>Hypoxanthine</td>
<td>8.17(s), 8.20(s)</td>
<td>CH, CH</td>
</tr>
<tr>
<td>UN8</td>
<td>8.19(s) or (d)</td>
<td>-</td>
</tr>
<tr>
<td>Formate</td>
<td>8.45(s)</td>
<td>H-COOH</td>
</tr>
<tr>
<td>1-methylnicotinamide</td>
<td>4.5(s), 8.2 (t) 8.88 (d), 8.95(d)</td>
<td>CH₂, ArH, ArH, ArH,ArH</td>
</tr>
<tr>
<td>N-phenylacetylglucose</td>
<td>8.0 (s), 7.4(t), 7.35(t), 7.35(d), 3.7(d), 3.7(s)</td>
<td>NH, ArH, ArH, ArH, CH₂, CH₃</td>
</tr>
<tr>
<td>UN9</td>
<td>8.10 (s)</td>
<td>-</td>
</tr>
<tr>
<td>UN10</td>
<td>5.00-5.02</td>
<td>-</td>
</tr>
<tr>
<td>UN11</td>
<td>1.22(d)</td>
<td>-</td>
</tr>
<tr>
<td>UN12</td>
<td>7.97 (-)</td>
<td>-</td>
</tr>
</tbody>
</table>
A7: Quantification of selected metabolites (U), normalized to creatinine

**Trigonelline, µM/mM creatinine**

- U1: 35.09
- U2: 24.46
- U3: 26.73
- U4: 24.75
- U5: 30.78
- U6: 40.67

**Hippurate, µM/mM creatinine**

- U1: 253.13
- U2: 352.94
- U3: 264.98
- U4: 244.50
- U5: 321.43

**Citrate, µM/mM creatinine**

- U1: 72.29
- U2: 114.36
- U3: 173.96
- U4: 295.85
- U5: 177.21
- U6: 411.09
Figures 1.A: Quantified concentration of selected metabolites normalized to creatinine (subject U)

A8: Quantification of selected metabolites (subject K), normalized to creatinine
Figures 2.A Quantified concentration of selected metabolites normalized to creatinine (subject K)
A9: Average concentration of selected metabolites

Table 6.A: Average concentrations of selected metabolites (µM/mM creatinine) from each observation and corresponding standard deviation (Std)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Trigonelline (µM/mM creatinine)</th>
<th>Hippurate (µM/mM creatinine)</th>
<th>Citrate (µM/mM creatinine)</th>
<th>Creatinine(mM)</th>
<th>TMAO (µM/mM creatinine)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U1</td>
<td>35.09</td>
<td>211.20</td>
<td>72.29</td>
<td>10263.25</td>
<td>2.91</td>
</tr>
<tr>
<td>U2</td>
<td>24.46</td>
<td>253.13</td>
<td>114.36</td>
<td>20616.45</td>
<td>65.53</td>
</tr>
<tr>
<td>U3</td>
<td>26.73</td>
<td>352.94</td>
<td>173.96</td>
<td>10880.31</td>
<td>27.35</td>
</tr>
<tr>
<td>U4</td>
<td>24.75</td>
<td>264.98</td>
<td>295.85</td>
<td>12885.90</td>
<td>27.01</td>
</tr>
<tr>
<td>U5</td>
<td>30.78</td>
<td>244.50</td>
<td>177.21</td>
<td>10977.55</td>
<td>0.68</td>
</tr>
<tr>
<td>U6</td>
<td>40.67</td>
<td>321.43</td>
<td>411.09</td>
<td>9524.80</td>
<td>23.96</td>
</tr>
<tr>
<td>Average</td>
<td>28.76</td>
<td>259.06</td>
<td>175.59</td>
<td>10928.93</td>
<td>25.48</td>
</tr>
<tr>
<td>Std(average)</td>
<td>6.45</td>
<td>52.55</td>
<td>125.10</td>
<td>4118.69</td>
<td>23.38</td>
</tr>
<tr>
<td>K1</td>
<td>146.62</td>
<td>814.52</td>
<td>0.00</td>
<td>2435.65</td>
<td></td>
</tr>
<tr>
<td>K2</td>
<td>139.48</td>
<td>1904.28</td>
<td>0.00</td>
<td>2132.65</td>
<td></td>
</tr>
<tr>
<td>K3</td>
<td>95.99</td>
<td>451.83</td>
<td>20.10</td>
<td>4065.05</td>
<td></td>
</tr>
<tr>
<td>K4</td>
<td>127.82</td>
<td>357.81</td>
<td>21.82</td>
<td>4700.50</td>
<td></td>
</tr>
<tr>
<td>K5</td>
<td>119.89</td>
<td>1589.25</td>
<td>24.44</td>
<td>2988.75</td>
<td></td>
</tr>
<tr>
<td>K6</td>
<td>48.12</td>
<td>435.00</td>
<td>52.40</td>
<td>3116.25</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>123.85</td>
<td>633.17</td>
<td>20.96</td>
<td>3052.50</td>
<td></td>
</tr>
<tr>
<td>Std(average)</td>
<td>36.32</td>
<td>663.04</td>
<td>19.35</td>
<td>976.26</td>
<td></td>
</tr>
</tbody>
</table>

A10: Heteronuclear $^1$H $^{13}$C 2D spectra of trimethylamine N-oxide and betaine

![Trimethylamine oxide spectrum](image)

Figure 3.A: Heteronuclear $^1$H $^{13}$C 2D spectrum of trimethylamine N-oxide, adapted from HMDB.com
Figure 4.A: Heteronuclear [$^1$H, $^{13}$C] 2D spectrum of betaine, adapted from HMDB.com