Metabolic fate of $^{13}$C-labelled polydextrose and impact on the gut microbiome: A triple-phase study in a colon simulator

Santosh Lamichhane,1, 7* Christian C. Yde,1, 4 Henrik Max Jensen,4 Wesley Morovic,8 Ashley A. Hibberd,8 Arthur C. Ouwehand,3 Markku T. Saarinen,3 Sofia D. Forssten,3 Lars Wiebe,5 Jørn Marcussen,4 Kresten Bertelsen,4, 9 Sebastian Meier,6 Jette F. Young,1 and Hanne Christine Bertram1

1 Department of Food Science, Aarhus University, Kirstinebjergvej 10, 5792 Aarslev and Blichers Allé 20, 8830 Tjele, Denmark

3 Dupont, Nutrition and Health, Sokeritehtaantie 20, 02460 Kantvik, Finland

4 DuPont Nutrition Biosciences ApS, Edwin Rahrsvej 38, 8220 Brabrand, Aarhus, Denmark

5 DuPont Nutrition Biosciences ApS, Tårnvej 25, 7200 Grindsted, Denmark

6 Technical University of Denmark, Department of Chemistry, Kemitorvet, Building 207, 2800 Kgs. Lyngby, Denmark

7 Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, Turku 20520, Finland.

8 DuPont Nutrition and Health, 3329 Agriculture Drive | Madison, WI 53716, USA.

9 Vestas Wind Systems A/S, Hedeager 42, 8200 Aarhus N, Denmark

*Corresponding author. Santosh Lamichhane

Email: santosh.lamichhane@utu.fi or lamichhane17@gmail.com, Phone: +358 2299070
Abstract

The present study introduces a novel triple-phase (liquids, solids and gases) approach, which employed uniformly labelled [U-\textsuperscript{13}C] polydextrose (PDX) for the selective profiling of metabolites generated from dietary fiber fermentation in an \textit{in vitro} colon simulator using human fecal inocula. Employing \textsuperscript{13}C NMR spectroscopy, [U-\textsuperscript{13}C] PDX metabolism was observed from colonic digest samples. The major \textsuperscript{13}C-labelled metabolites generated were acetate, butyrate, propionate, and valerate. In addition to these short-chain fatty acids (SCFAs), \textsuperscript{13}C-labelled lactate, formate, succinate, and ethanol were detected in the colon simulator samples. Metabolite formation and PDX substrate degradation were examined comprehensively over time (24 and 48 hours). Correlation analysis between \textsuperscript{13}C NMR spectra and gas production confirmed the anaerobic fermentation of PDX to SCFAs. In addition, 16S rRNA gene analysis showed that the level of \textit{Erysipelotrichaceae} was influenced by PDX supplementation and \textit{Erysipelotrichaceae} level were statistically correlated with SCFA’s formation. Overall, our study demonstrates a novel approach to link substrate fermentation and microbial function directly in a simulated colonic environment.

Keywords. Gut microbiome, Dietary fiber, \textsuperscript{13}C-labelled metabolites, \textsuperscript{13}C NMR
Introduction

The human gut harbours a complex microbial ecosystem, and the gut microbial activity has been recognized as a pivotal regulator of the host metabolism.\textsuperscript{1-4} Recent evidence revealing that gut microbial activity can be influenced by dietary fiber (i.e. non-digestible food ingredients), has led to an increased demand for evaluation of the specific functions of these ingredients in the gut.\textsuperscript{5-7} Polydextrose (PDX), which is used as functional ingredient in a variety of foods, is a highly branched and complex polymer of glucose. PDX is a soluble fiber produced by thermal condensation of glucose and possesses prebiotic effects as a food ingredient.\textsuperscript{8-9} An understanding of the microbial composition and metabolism is fundamental to evaluate the effects of dietary fibers in the gut. In addition, evaluating the selective microbial products formed from the substrate is key to directly linking gut microbial activity and health benefit of any particular dietary fiber. Stable isotope labelling provides an excellent approach for tracking the conversion of the substrate and identifying resultant metabolic products, thus allowing a rapid and comprehensive profiling of selective microbial activity.\textsuperscript{10} An approach using uniformly labelled [U-\textsuperscript{13}C] substrate also offers a robust method to elucidate the metabolic pathways involved in the microbial metabolism. In fact, previous studies based on stable isotope labelling have shown the usefulness of [U-\textsuperscript{13}C] approaches. For instance, de Graaf et al. introduced \textsuperscript{13}C enrichment as an approach to evaluate the kinetics of \textsuperscript{13}C-labelled glucose fermentation in an \textit{in vitro} colon model.\textsuperscript{11-13}

To characterize the metabolic process of dietary fiber catabolism, an \textit{in vitro} colon simulator offers the advantage of enabling dynamic sampling and implying substantially less ethical issues compared to an \textit{in vivo} study. On that account, an \textit{in vitro} colon model could be a robust
approach to profile the complete set of $^{13}$C-labelled metabolites produced during the microbial fermentation processes. Nuclear magnetic resonance (NMR) or mass spectrometry (MS) is often used to measure the isotopically labelled metabolites in biological samples. An NMR-based approach can directly detect an isotopic label in the molecule and distinguish different isotopomers, whereas MS-based techniques requires complicated protocols with different ionizations and derivatization steps to obtain information on the isotopomer. Nevertheless, both analytical platforms provide complementary information. In this study, we employed a combination of liquid and solid state NMR to characterize the fate of uniformly labelled [U-$^{13}$C] PDX catabolism by gut microbiota in a four-stage in vitro colon simulator model. Concomitantly, gas chromatography (GC) coupled with MS was used to monitor the gas generated during PDX catabolism. In addition, 16S rRNA gene amplicon sequencing was performed to identify and compare the microbial composition within the given sample sets. Overall, the aim was to measure the selective metabolic and microbial changes occurring in the colon simulator using uniformly labelled [U-$^{13}$C] PDX as a substrate for gut microbes.
Material and methods

Human colon simulator

The in vitro simulator used for the study models the human large intestine (Figure 1A). This anaerobic in vitro colon model has been used to investigate the fermentation, and function of prebiotic and probiotic ingredients. Mäkivuokko et al. have previously described the technical specification of this in vitro model. Briefly, the simulator used for the study consists of eight separate units, each containing four semi-continuously connected glass vessels. The vessels in each unit (V1 to V4) model the ascending (V1), transverse (V2), descending (V3) and sigmoidal (V4) parts of the human colon, respectively. In the initial phase of the simulation, each unit is inoculated with pre-incubated fecal inoculum obtained from a fresh human fecal sample, which forms the microbiota of the colonic model. These fecal inoculums are prepared with three parts (w/v) of anaerobic simulator medium, incubated anaerobically for 24 hours at 37 °C (details in). In this study, three healthy Finnish volunteers provided the fecal samples for inoculum. According to Finnish law, no ethical approval was required, since there has not been any interference with a person’s physical or mental integrity. During simulation, the fecal inoculum from one volunteer was used, and the system was fed with a synthetic ileal medium in three-hour cycles to simulated colon model for 24 or 48 hours, during which transition of fermented fluids and microbes and feeding of fresh medium occurred. As test substance, 2% PDX with natural isotope abundance (Litesse Ultra; Danisco UK, Redhill, U.K.) or 2% [U-13C]-PDX (Danisco Denmark, Grindsted, Denmark) was added to the synthetic ileal medium used as feeding solution. The syntetic ileal medium without addition of any PDX was used as a control (CTRL). The synthetic ileal medium was used as a control (CTRL). The simulations were performed at DuPont, Nutrition and Health, Kantvik, Finland. To track the microbial activity
over time, the microbial slurry and gas were collected from all vessels (V1-V4) after 24 and 48 hours of simulation. In total, 72 (3 donons × 4 vessels (V1-V4) × 2 time points (24 and 48 hours) × 3 without PDX (CTRL)/2% PDX/ 2% [U-13C] - PDX) samples were collected from the in vitro colon simulator and stored at -80 °C prior to analysis.

114 **Preparation of simulated fecal water samples**

The collected fecal slurry samples from the simulator were homogenized by vortex mixing for 1 min, and then centrifuged at 5,000 g for 10 minutes at 4 °C (Eppendorf 5471, USA). The supernatants were removed carefully and subsequently lyophilized. The sample pre-treatment (adjusting the pH ≥ 7) before lyophilisation was carried out as recommended by Jacobs et al. The remaining pellets were stored at -80°C for solid-state NMR analysis. Lyophilized fecal samples were reconstituted in D2O and the samples were homogenized by vortex mixing and then centrifuged at 10,000 g for 10 minutes at 4 °C.

122 **Spectroscopic analyses**

123 **Liquid-state NMR**

A volume of 600 μL supernatant containing 0.025 mg/mL 3-trimethylsilyl propionic acid-d4 sodium salt (TSP) as reference compound was transferred to a 5 mm NMR tube. One-dimensional 13C NMR experiments were acquired using a Bruker Avance III 800 MHz spectrometer equipped with a 5-mm 1H observe TCI cryoprobe (Bruker Biospin, Rheinstetten, Germany) at 25°C. A standard pulse sequence using power-gated decoupling and 30° flip angle was used (zgpg30). Acquisition parameters for the spectra were 128 scans, a spectral width of 48077 Hz collected into 64 K data points, an acquisition time of 0.68 s and an inter-scan
relaxation delay of 5 s. The Free Induction Decay (FID) obtained was multiplied by 1 Hz of exponential line broadening before Fourier transformation. The spectra were referenced to TSP (chemical shift defined at 0 ppm), phased, and baseline corrected in Topspin 3.0 software. Assignments of $^{13}$C NMR signals were carried out according to the Human Metabolome Database. In addition, an HSQC experiment was acquired with a spectral width of 11961 Hz in the $^1$H dimension and 33276 Hz in the $^{13}$C dimension, a matrix with a size of 1024×128 complex data points, 2 transients per increment and a relaxation delay of 1.5 s. Besides 1D carbon spectra, $^{13}$C positional enrichment were confirmed by $^1$H NMR spectroscopy ($^{13}$C decoupled and $^{13}$C coupled spectra) using a Bruker Avance III 500 MHz spectrometer equipped with a 5 mm triple resonance (TXI) probe at 25°C (experiment details in Supporting information (SI)).

Solid-state NMR

The pellets obtained from the in vitro colon fecal slurry were analyzed using cross-polarization magic-angle spinning (CP MAS) NMR. Lyophilized pellet samples (mean sample weight: 29.38 mg) were powdered and packed in 4-mm diameter zirconium rotors (Bruker Biospin, Rheinstetten, Germany). $^{13}$C CP MAS NMR spectra were acquired on a Bruker Avance III 600 MHz spectrometer equipped with a $^1$H/$^{13}$C CP MAS probe with gradient aligned along the magic angle. For acquisition of $^{13}$C CP MAS NMR spectra, a contact time of 0.5 ms, a proton field of approximately 45 kHz during CP and data acquisition, a relaxation delay of 2 s and a spinning speed of 14 kHz were employed. The FID was multiplied by 50 Hz of exponential line broadening before Fourier transformation. The spectra were phased, and baseline corrected using Bruker Topspin 3.0 software.

LC-MS analysis
The HP-LC was operated at a flow rate of 12 uL/min, and the mobile phase consisted of solution A (0.1% HCO$_2$H in water) and solution B (0.1% HCO$_2$H in ACN). For all analyses, 1 uL of calibrant (sodium formate) and sample were used. The samples were separated by an eclipse plus phenyl C18 column (Agilent Technologies, Waldbronn, Germany), using the following gradient: 2% B for 4 min, linear gradient of 2-60% B in 10 min, 60-90% B in 3 min, 90% B for 2 min, 90-2% B in 1 min, and 2% B for 10 min. MS was operated in both negative and positive ion mode. The mass range was set from 40 to 800 m/z, and data were acquired in profile mode at a frequency of 1 Hz with active focus mode off. The raw data were automatically calibrated using the sodium formate cluster signals and required m/z values were integrated in the data analysis software package Compass HyStar (Bruker Daltonics, Bremen, Germany).

**CO$_2$ analysis by GC-MS**

CO$_2$ analysis was performed using an Agilent Technologies 6890N/5975C GC-MS system. A total of 250 µL sample was injected (split ratio 1:150) using a 2.5 mL heated syringe (50 °C), into a 50 m (L) x 0.25 mm (ID) x 0.25 µm (df) FFAP column (Quadrex, Bethany, CT 06524 USA ), kept at 30 °C. Helium was used as carrier gas at 1 mL/min. The sampling needle was flushed with He between sample injections to avoid interference from air and from the previous sample. The MS was used in full scan mode (scan range 10 – 100 amu) with electron ionization (70 eV). The selected ions were m/z=44 for $^{12}$CO$_2$ and m/z=45 for $^{13}$CO$_2$ with retention time of 161 s. Argon (Ar) with m/z=40 was used as a reference gas for relative quantification of the gases.
Metabolite data analysis

$^1$H NMR spectra were imported into Matlab R2014b (The Mathworks, Inc., USA) and the misalignments of the spectra were corrected using the icoshift algorithm based on the correlational shifting of spectral intervals. The spectra were mean centered and pareto-scaled before principal component analysis (PCA). PCA was performed using the PLS Toolbox (Eigenvector Research, USA) in MATLAB R2014b. Relative quantification of selected $^{13}$C resonances were performed by integration of peak areas in MATLAB R2014b. Spearman correlation coefficient between the integrals of these metabolites was calculated using the statistical toolbox in Matlab R2014b. The results were illustrated by heat maps using MATLAB R2014b.

Extraction and quantification of bacterial DNA

The DNA from the simulation samples was extracted and purified with an automated MagMAX™ Sample Preparation System (Life Technologies, Halle, Belgium), using the MagMAX™ Nucleic Acid Isolation Kit. The amount of extracted DNA was determined by a Qubit® dsDNA HS Assay Kit (Thermo Fisher Scientific, Vantaa, Finland). Concurrently, non-template negative controls were processed to detect possible contamination.

16S rRNA gene amplicon sequencing and microbiome data analysis

The V4 region of the 16S rRNA gene communities was amplified using a PCR workflow as previously described. Purified amplicons were pooled and sequenced using 2 x 250 base pair (bp) Paired-End Illumina MiSeq (Pioneer, Johnston, IA). Resulting sequencing data were analyzed using the Quantitative Insights into Microbial Ecology (QIIME, v. 1.9.1) pipeline.
Sequencing reads were joined using fastq-join\textsuperscript{22} with a minimum 200 bp overlap and 5% maximum difference, de-multiplexed, and quality filtered using a Phred threshold of >20. Sequences were clustered into operational taxonomic units (OTUs) at 97% similarity with uclust\textsuperscript{23} using an open reference scheme against the Greengenes database,\textsuperscript{24} where non-matching reads were retained and clustered \textit{de novo}. Sequences were aligned with PyNAST\textsuperscript{25} and the phylogenetic tree used for downstream analysis was generated using FastTree.\textsuperscript{26} Samples were normalized for sequencing depth by rarefaction at 15,958 reads, and one sample (Simulation II, Time 48, 2\% PDX $^{13}$C, vessel 1) was removed due to low coverage. $\alpha$-diversity was assessed by the phylogenetic diversity whole tree metric\textsuperscript{27} and compared between groups using non-parametric t-tests with 999 Monte Carlo permutations. $\beta$-diversity was measured using the weighted UniFrac distance\textsuperscript{28} and groups were compared with Analysis of Similarities (ANOSIM). Genera (or higher taxonomic levels if the genus could not be confirmed) comprising >0.2\% abundance were compared between groups using Kruskal-Wallis tests and graphed using Prism 7 (GraphPad Software, La Jolla, California, USA). P values were adjusted for multiple comparisons with the Benjamini-Hochberg false discovery rate (FDR) correction and p<0.05 was considered a significant difference for all tests.
Results and Discussion

This work focused on the study of [U-\(^{13}\)C] labelled PDX as a substrate during microbial digestion and characterization of \(^{13}\)C-labelled metabolites in different compartments of an in vitro colon simulator. The \(^{13}\)C-labelled metabolites were characterized by liquid-state and solid-state NMR spectroscopy. GC-MS was performed to evaluate the gas generated during [U-\(^{13}\)C] PDX catabolism. In addition, 16S rRNA gene sequencing was performed to study the phylogeny and taxonomy of the gut microbiome from the \textit{in vitro} gut fermentation.

Solution state analysis

A significant effect of [U-\(^{13}\)C] PDX metabolism was observed in \(^{13}\)C NMR spectra obtained from colonic digest extracts. When compared to experiments using a natural abundance reference PDX reference, the spectra with [U-\(^{13}\)C] PDX clearly show the low-molecular-weight compounds generated from carbohydrate metabolism (Figure 1B). The results show that the major \(^{13}\)C-labelled metabolites included acetate, butyrate, propionate, and valerate. To map these metabolic changes, peak integrals were obtained from \(^{13}\)C NMR spectra and plotted accordingly (SI, Figure S1). Even though an inter-individual variation is apparent, the integrals indicate that the SCFA concentrations gradually increased from the ascending colon (V1) to the descending colon (V3) during the simulation. The overall metabolite concentration decreased in the sigmoid colon part (V4) during 24 hours while it remained stable at 48 hours (SI, Figures S1-S4).

Intriguingly, the pattern seen in SCFA formation is supported by the degradation pattern of PDX as the concentration of PDX decreased during simulation from the proximal colon to the sigmoid colon at 48 hours (SI, Figure S3), however, an inter-individual difference between the donors exists mainly at 24 hours. Thus, our result reveal how fiber/substrate availability affects the gut
fermentation pattern and SCFA concentrations in different segments of the gut over time. The present study corroborates that PDX fiber fermentation results in the synthesis of SCFAs as eminent metabolic end products.\textsuperscript{29} Previous reports from animal models, \textit{in vitro} models, and humans have shown increased fecal levels of butyrate and acetate in subjects after PDX intake.\textsuperscript{16, 30-31} However, none of these studies measured the microbial metabolites derived directly from PDX catabolism. We report a novel approach to determine the specific metabolites generated from PDX fermentation. Furthermore, this \textit{in vitro} study enabled us to directly monitor the dynamics of microbial metabolites formed during PDX fermentation in different parts of the colon at different time point (24 hrs and 48 hrs), while the same information would be practically and ethically impossible to obtain \textit{in vivo} in humans. The result revealed that acetate clearly is the principal SCFA produced, however butyrate is of particular interest as it has been described as a vital metabolite for colon.\textsuperscript{32-33} An increase in SCFA production, specifically butyrate formation in the colon simulator, can be considered as the primary metabolic significance of PDX digestion in relation to gut microbial activity and benefits for host. Interestingly, our data also revealed an increased level of valerate from V1 to V4 due to PDX catabolism (for inoculum from subject). Until now, studies focussing on valerate as product of carbohydrate metabolism have been sparse.\textsuperscript{34-36} Although only observed in one subject, our finding indicates a possible direct association between dietary fiber fermentation, valerate formation, and gut microbial activity. However, to validate this hypothesis, other independent studies are needed.

In addition to SCFAs, we identified $^{13}$C labelled lactate, formate, succinate, and ethanol in the fecal $^{13}$C NMR spectra. As the colon simulator genuinely allowed dynamic sampling, we comprehensively examined these biochemical changes over time (24 and 48 hours) in the vessels
The results show that succinate, lactate, and formate mainly appeared in vessel V1 and V2 (ascending and transverse colon) during the fermentation, but an inter-individual difference between the donors was observed. This inter-individual variation can be ascribed to the differential microbial composition between individuals and the result suggests that these metabolites are metabolised further and thus disappear in V3 and V4 (described below). A previous study has shown that lactate can be fermented to SCFAs (mainly butyrate) by human intestinal bacteria.\textsuperscript{37} Previous work also supports the finding that lactate accumulates at low pH (pH 5.2 likewise in V1 colon simulation) but butyrate and propionate production were initiated at pH 5.9 and 6.4, respectively, equivalent to the pH of descending and sigmoid colon (V3 and V4) in the simulation.\textsuperscript{38}

Similarly, formate is involved in acetate formation via the Wood-Liungdhal pathway and succinate is a well-known precursor for the formation of propionate through the succinate pathway (Figure 2).\textsuperscript{29} Interestingly, studies have also shown that when sufficient carbohydrate is present during bacterial metabolism, succinate is accumulated due to a reduced requirement for succinate decarboxylation.\textsuperscript{39} Figure 2 depicts the possible pathways linked to carbohydrate (PDX) catabolism and metabolites formation in the gut during the anaerobic microbial process. Accumulation of lactate, formate, and succinate (V1 and V2) indicates an imbalance between the production and consumption pathway involved in anaerobic gut metabolism. However, the finding also reflects diverse microbiota composition/activity and potential cross feeding between microbes in the gut simulator.

The structure of the randomly formed PDX is rather complex and is composed of both linear and branched \(\alpha\) and \(\beta\) 1-2, 1-3, 1-4 and 1-6 linkages, dominated by 1-6.\textsuperscript{40,41} The $^{13}$C spectra
indicated the presence of partly undigested $^{13}$C-labelled PDX especially in vessels V1 and V2 alongside metabolic intermediates, while sharper carbohydrate signals emerge especially in vessels V3 and V4. The degradation pattern of residual PDX suggests that vessels V1 and V2 (ascending to transverse colon) had similar metabolic patterns, as did vessels V3 and V4 (descending to sigmoid colon), respectively (Figure 3). In particular, the results indicate that vessels V1 and V2 had residual polysaccharide with broadened signals near 111.5 ppm ($\beta$-glucofuranosyl), 106 ppm (predominantly $\beta$-gluco-pyranosyl), 104 ppm (levoglucosan ($\beta$-1,6 anhydroglucose) and 101 ppm ($\alpha$-gluco-pyranosyl). Increasingly sharper signals in these regions emerged in vessels V1 to V4 due to polysaccharide degradation to smaller units, which could be ascribed to enzymatic or acid hydrolysis. The glucofuranosyl units are increasingly represented in vessels 3 and 4, possibly due to a lack of efficient enzymes degrading glucofuranosyl linkages. The results show that free glucose also varies in different compartments of the colon simulator. Our finding is in agreement with a former study where it was reported that glucofuranosyl moiety is accumulated in the final vessel indicating a lesser ability of the microbiota to degrade the glucofuranosyl structure.

**Solid-state analysis**

The residual pellets obtained from the microbial slurry were examined using CP MAS NMR spectroscopy. Figure 4A shows representative $^{13}$C CP MAS spectra obtained from the pellet of the anaerobic microbial degradation of PDX during simulation. To our knowledge, this paper presents the first work examining the triple phases in fecal samples (*in vitro* simulation), in particular CP MAS from the fecal slurry (*in vitro* samples) has not been reported before. The CP MAS spectra showed characteristic peaks that can be assigned to the carbohydrate, carbonyl,
CH$_2$, CH$_3$, anomeric and aromatic resonances.$^{44}$ The $^{13}$C MAS NMR spectrum is a sum of [U-$^{13}$C] PDX, $^{13}$C metabolites, and natural abundance (1%) of $^{13}$C from major undigested components such as cellulose, ligno-cellulose, undigested protein and fat, and bacterial cell debris. Thus, collectively data provide valuable information about the heterogeneity of microbial activity from proximal to distal section of the gut. As shown in the score plot depicted in Figure 4B, overall the intensity of signals from carbohydrate and anomeric resonances were decreased from vessel V1 to V4 while carbonyl, CH$_2$, CH$_3$ resonances were higher in V1 to V3 (ascending to descending colon) and did not noticeably change in V4 (sigmoid colon) indicating that substrate availability affects the gut fermentation pattern. This information is unique and adds to the finding obtained from solution-state NMR.

**Gas-state analysis**

In general, CO$_2$ and CH$_4$ are the products of carbohydrate fermentation during anaerobic microbial metabolism. We detected $^{12}$CO$_2$ (m/z 44) and $^{13}$CO$_2$ (m/z 45), while CH$_4$ was not detected. This finding may be explained by the fact that CH$_4$ is produced only in 30% -50% of healthy adults.$^{45}$ As expected, the ratio $^{13}$CO$_2$/$^{12}$CO$_2$ shows a marked difference between the fecal extract with [U-$^{13}$C] PDX or a natural abundance PDX reference (Figure 5). In addition, the heat map (Figure 6) depicts positive correlations between $^{13}$CO$_2$/$^{12}$CO$_2$ ratio and intensity of SCFA signals in the liquid-state $^{13}$C NMR spectra. In addition, the $^{13}$CO$_2$/$^{12}$CO$_2$ ratio shows negative correlation to certain regions of the PDX signals in $^{13}$C NMR spectra. GC-MS results are thus consistent with NMR analyses and provide an additional angle to [U-$^{13}$C] PDX catabolism and microbial metabolism in the gut simulator.
Microbiome analysis

To measure the \textit{in vitro} colonic microbial composition, $\alpha$-diversity and $\beta$-diversity were compared between the groups. The $\alpha$-diversity measures the OTU diversity within the samples whereas $\beta$-diversity measures dissimilarity between each sample pairing at the OTU level. The results showed that $\alpha$-diversity was only significantly different between simulations performed with inoculum obtained from the donors I/II and I/III (Figure 7A, $p = 0.006$). These groups were compared with a non-parametric t-test using 999 Monte Carlo permutations and the p-value was adjusted using the Benjamini-Hochberg FDR. However, the $\alpha$-diversity from other simulation/treatment and simulation/time comparisons did not show any significant differences. In contrast, the results depicted that $\beta$-diversity was significantly different between simulation times (24 hours and 48 hours) and simulations performed from different donors I, II and III (Figure 7B). The principal coordinate analysis (PCoA) for the weighted UniFrac distance matrix showed significant differences between the simulation from three donors ($p=0.001$) and to a lesser degree simulation time ($p=0.01$), whereas the treatment effect was insignificant ($p=0.1029$).

An analysis of genera and higher-level taxonomy between different categories showed few bacterial lineages shifted in relative abundances after the PDX supplementation, with increasing levels of \textit{Erysipelotrichaceae} being most pronounced. The \textit{Erysipelotrichaceae} family was found to be significantly different in control and treatment for all simulations (Figure 8). In addition, the PDX resulted in an increased level of \textit{Bifidobacterium} and in a decreased level of \textit{Lactobacillus}, but these changes were not as prominent as \textit{Erysipelotrichaceae}. Recently, \textit{Erysipelotrichaceae} has gained substantial attention in gut microbial research due to its potential role in host physiology and/or disease conditions. Evidence support an association...
between *Erysipelotrichaceae* and metabolic disorders\(^47\) and inflammation-related disorders of the gastrointestinal tract.\(^48\)\(^-\)\(^49\) In addition, studies have shown that diet influences the relative abundance of *Erysipelotrichaceae* family in the gut.\(^47\)\(^,\)\(^50\)\(^-\)\(^52\) Previous studies have shown positive correlations between *Erysipelotrichaceae* occurrence and complex carbohydrate consumption.\(^51\)\(^-\)\(^52\) The *Erysipelotrichaceae* family is designated as SCFA producers, and some species within this family are capable of producing butyrate.\(^53\)\(^-\)\(^54\) Meanwhile, other studies have shown this bacterial family to be enhanced from high-fat diets.\(^55\)\(^-\)\(^56\) This is not unexpected, as identical gut microbial genera can have both specific and shared systemic effects.\(^57\)\(^-\)\(^58\) Moreover, our results substantiate an enrichment of *Erysipelotrichaceae* family during complex carbohydrate (PDX) fermentation. Specifically, our results suggest that associations between PDX fermentation, *Erysipelotrichaceae* and SCFA production exist (supported by correlation analysis, heat map Figure 9), albeit using a limited set of donors. Independent studies with larger samples sets are required to disclose a possible link, and further analysis at strain levels within *Erysipelotrichaceae* may expand our current knowledge.\(^59\)

\(^{13}\)C incorporation and statistical correlation between metabolites and microbiome

LC-MS analysis enabled us to determine the level of isotope incorporation in the SCFAs during *in vitro* colon simulation. The results indicate that there was clear difference in the \(^{13}\)C isotope incorporation among different subjects and vessels over time. In particular, butyrate showed high variability when compared to acetate and propionate (SI, Figures S5-S8). Meanwhile, these data provide an opportunity to explore if any statistical relation between metabolites and microbes in the colon simulator exists at isotope level. The correlation analysis (Figure 9) between OTUs (included OUT with zero values reported in ≤ 1 sample) and integrals of the metabolites (m/z...
from LC-MS analysis) suggests a statistical association between bacterial family *Erysipelotrichaceae, Faecalibacterium* and SCFAs. The correlation was found to be significant with *Erysipelotrichaceae* (p value < 0.05) for most of the $^{13}$C isotopes from SCFAs. Intriguingly, we did also see that *Erysipelotrichaceae* were increased in PDX treated samples. In fact our finding, is also in agreement with previous findings relating *Erysipelotrichaceae* and *Faecalibacterium* family to SCFA producers. In addition, *Blautia* was positively correlated to butyrate. The statistical correlation indicates that both of these bacterial family may be actively involved in the PDX fermentation and butyrate formation in the colon simulator. In fact, previous studies have reported *Blautia* and *Clostridiales* to be a common bacterial genus in the human gut belonging to a butyrate-producing bacterial species. However, we saw negative correlation between *Clostridiales* and butyrate. The discrepancy might be linked to higher abundance of *Erysipelotrichaceae* which possibly dominated the butyrate production in the *in-vitro* colon.

Likewise, $^{13}$C incorporation of propionate was positively correlated to *Parabacteroides* and *Bacteroidetes*. The phylum *Bacteroidetes* has been reported to be major propionate produce, even though it could also produce acetate and butyrate. Furthermore, we saw negative correlation between *Bifidobacterium* and $^{13}$C labelled butyrate (Figure 9). *Bifidobacterium* are well known carbohydrate degraders but they are antioxidants, polyphenols, and conjugated linoleic acids producers. The negative association seen here may be indication of cross feeding interactions between *Bifidobacterium* and butyrate-producing colon bacteria. However, the underlying relation is not yet clear. These findings may be corroborated from future studies using more donors. Taken together, our findings substantiate the association of various microbes...
with fermented gut microbial products. Here, we provide a novel methodological approach to study gut microbial function using stable isotope tracking. This approach could enable in-depth functional microbial analyses by directly linking substrate fermentation and microbial function in the gut system.

Conclusion

Overall, the present study showed that the application of a triple-phase study (solid, liquid, and gas) combined with $^{13}$C labelling monitors the overall and specific metabolites derived directly from in vitro colonic PDX catabolism. The correlation between $^{13}$C NMR spectra and $^{13}$C CO$_2$ production substantiates the anaerobic fermentation of PDX to SCFAs. Furthermore, the 16S rRNA gene analysis depicted that Erysipelotrichaceae, a butyrate-producing bacterial lineage changed in relative abundance after the PDX supplementation in the in vitro colon simulator. Further, statistical correlation analysis showed that Erysipelotrichaceae and Blautia is positively correlated to $^{13}$C SCFAs while the Parabacteroides is linked to $^{13}$C propionate formation from PDX metabolism in an in vitro colon simulator. Overall, the study proposes a novel approach to link substrate fermentation and microbial function directly in a simulated gut system.

ASSOCIATE CONTENT

SUPPORTING INFORMATION

The following associated information is available free of charge on ACS publications website Supplementary methods on $^1$H NMR experiments.

Figure S1-S4: Integrals of $^{13}$C labelled acetate, $^{13}$C labelled butyrate, $^{13}$C labelled propionate, $^{13}$C labelled succinate, $^{13}$C PDX region 1 (104.5-103.9ppm), $^{13}$C labelled lactate, $^{13}$C valerate and $^{13}$C
labelled formate obtained from each vessel (V1 to V4) representing different areas of the colon i.e. proximal to the distal part in the colon simulator.

Figure S5 showing the representative MS spectra showing the m/z ratio for acetate, propionate and butyrate. Figure S6-S8 illustrating the $^{13}$C incorporation at isotopic level in the butyrate, acetate and propionate for three donors at 24 and 48 hours.

FUNDING SOURCES

The present study is part of the Ph.D. work of Santosh Lamichhane and was financially supported by DuPont, Food Future Innovation (FFI), and Aarhus University. “Advances in food and nutrition research through implementation of metabolomics technologies” (Grant No. 274- 09-0107).

ACKNOWLEDGEMENT

The 800 MHz spectra NMR spectra were obtained at the Danish Instrument Center for NMR spectroscopy of Biological Macromolecules. Mr. Kim Blauenfeldt Gosmers from DuPont, Aarhus, and Dr. Morten Rahr Clausen from Aarhus University are thanked for help during GC and LC-MS analysis, respectively. Ms. Kirsi Stenström (DuPont Nutrition and Health, Finland) is acknowledged for help with the colon simulations and DNA extraction. Ms. Paige Roos at DuPont Pioneer is acknowledged for performing amplicon sequencing.

AUTHOR INFORMATION

Santosh Lamichhane, Email: santosh.lamichhane@utu.fi or lamichhane17@gmail.com. Phone: +358 2299070
**Figure Captions**

**Figure 1.** The *in vitro* colon simulator and representative $^{13}$C NMR spectra obtained from the colon simulation. A) Schematic depiction of the *in vitro* colon simulator. Each vessel (V1–V4) in colon simulator models the different compartments of the human colon from ascending (V1), transverse (V2), descending (V3) and sigmoidal (V4) parts, each having a different pH and flow rate. B) Representative $^{13}$C NMR spectrum from the simulated faecal sample with [U-$^{13}$C] labelled PDX. Here, acetate (A), butyrate (B), propionate (P), and valerate (V) were the main metabolites produced during *in vitro* fermentation. C) Representative $^{13}$C NMR spectrum of the simulated faecal sample with non-labelled PDX. D) The NMR spectra obtained from the control sample without PDX.

**Figure 2.** Schematic illustration of compounds produced from [U-$^{13}$C] PDX catabolism. The compound illustrated with dark green stars (★) are the major products of fermentation, while light green stars (★★) show the intermediates or the less abundant metabolites in the NMR spectra. The red stars (★) represent the gaseous metabolites detected by GC-MS.

**Figure 3.** The 60-110 ppm region in representative $^{13}$C NMR spectra. Black spectra represent different vessels in the colon simulator while the red spectrum represents a spectrum obtained from pure and undigested substrate. The sharp signal at 104 ppm in PDX (black star) is tentatively assigned to levoglucosan, which is a common product of PDX formation (through acid-catalysed dehydration). The signal at 107 ppm (green star) most likely is in the same molecule as those at 83, 82, 80.5, 69 and 66 ppm (chemical shifts of a furanose-form). The broad signals at 106-ppm (red star) and 101 ppm (see for instance PDX substrate spectrum) are assigned to β-glucopyranosyl and α-glucopyranosyl units, respectively.

**Figure 4.** $^{13}$CP MAS NMR of simulated fecal pellets. A) Representative $^{13}$CP MAS NMR spectrum of simulated fecal pellets. B) PCA score plot and C) Loading plot. The PCA score and loading plot reflects the time/vessels dependent changes in the direction of PC3. The signals from carbohydrate and anomeric resonances decreased from V1 to V4 while carbonyl, CH$_2$, CH$_3$ resonances increased from V1 to V2.

**Figure 5.** $^{13}$CO$_2$/$^{12}$CO$_2$ ratio as determined with GC-MS. The representative plot shows the ratio for a simulation obtained from a single donor.

**Figure 6.** Correlation coefficients illustrated by heat map. The Pearson correlation coefficient (R) value revealed positive correlation between $^{13}$CO$_2$ /$^{12}$CO$_2$ ratio and integrals of SCFA signals. Meanwhile the $^{13}$CO$_2$ /$^{12}$CO$_2$ ratio shows negative correlation to certain region of PDX signals in $^{13}$C NMR spectra (PDX1 (104.5-103.9ppm) PDX2 (106.9 – 106.5 ppm) PDX3 (80.8 – 80.5 ppm) PDX4 (69.4-69.2 ppm)). The color-coding in the figure represents the R-values. (× indicates significant correlation p values < 0.05).
**Figure 7.** Microbial diversity between the three simulations. A) Boxplots show overall α-diversity (diversity within samples) in each simulation using the PD whole tree metric. Groups were compared with a non-parametric t-test using 999 Monte Carlo permutations and the p-value was adjusted using the Benjamini-Hochberg FDR method. Significantly different p values are denoted. B) The β-diversity (diversity between samples) of each simulation is visualized by principal coordinate analysis by weighted UniFrac distance matrix. Colors denote the different simulation numbers: red squares, I; blue circles, II, orange triangles, III. The results were obtained using the ANOSIM statistical method.

**Figure 8.** Relative abundances of 16S rDNA in all samples at the genus, family (f) or order (o) levels. Samples are ordered by simulation number, simulation time, and the various treatment types: I (inoculum), CTRL (control, without PDX), PDX (2% PDX) and [U-13C] PDX (13C 2% PDX). The 35 most abundant taxa are shown, with “all other” signifying all other organisms that are <0.2% average abundance. Significant differences are denoted by asterisk (* p<0.05) in order of treatment (CTRL: PDX, CTRL: [U-13C] PDX), and by simulation number (I/II/ III) following taxa names.

**Figure 9.** The spearman correlation coefficient (R). The value illustrated by heat map indicate blue is positive correlation while the red gradient indicates negative correlation (× indicates significant correlation, p values < 0.05). Here, N represents the natural abundance of given metabolites and the label +1, +2, +3, +4 represents the shifts of the mass of the given metabolite by one, two, three or four units due to 13C label, respectively.
References


Lahtinen, S. J.; Jakobsen, M.; Svensson, B.; Jespersen, L., The effect of selected synbiotics on
microbial composition and short-chain fatty acid production in a model system of the human

16. Makivuokko, H.; Nurmi, J.; Nurminen, P.; Stowell, J.; Rautonen, N., In vitro effects on
polydextrose by colonic bacteria and caco-2 cell cyclooxygenase gene expression. *Nutr. Cancer*
**2005**, *52* (1), 94-104.

Vaughan, E. V.; van Duynhoven, J., 1H NMR metabolite profiling of feces as a tool to assess the

HMDB 3.0--The Human Metabolome Database in 2013. *Nucleic Acids Res* **2013**, *41* (Database
issue), 801-807.

19. Savorani, F.; Tomasi, G.; Engelsen, S. B., icoshift: A versatile tool for the rapid

and High-Throughput Sequencing of Commercial Probiotic Products Reveals Composition

Koenig, J. E.; Ley, R. E.; Lozupone, C. A.; McDonald, D.; Muegge, B. D.; Pirlung, M.;
Reeder, J.; Selinsky, J. R.; Turnbaugh, P. J.; Walters, W. A.; Widmann, J.; Yatsunenko, T.;
Zaneveld, J.; Knight, R., QIIME allows analysis of high-throughput community sequencing data.


23. Edgar, R. C., Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*
**2010**, *26* (19), 2460-1.

T.; Dalevi, D.; Hu, P.; Andersen, G. L., Greengenes, a Chimera-Checked 16S rRNA Gene
Database and Workbench Compatible with ARB. *Appl. Environ. Microbiol.* **2006**, *72* (7), 5069-
5072.

25. Caporaso, J. G.; Bittinger, K.; Bushman, F. D.; DeSantis, T. Z.; Andersen, G. L.; Knight,
R., PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics*


27. Faith, D. P.; Baker, A. M., Phylogenetic diversity (PD) and biodiversity conservation:


29. Koh, A.; De Vadder, F.; Kovatcheva-Datchary, P.; Bäckhed, F., From Dietary Fiber to
Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites. *Cell* **2016**, *165* (6),
1332-1345.


FIGURE 1.

A) 

B) 

C) PDX without \( ^{13}C \) label

D) CTR without PDX
FIGURE 2.
FIGURE 3.
FIGURE 5.
FIGURE 6.
FIGURE 7.
FIGURE 9.