

DATASET BRIEF

The intracellular proteome of the gut bacterium *Bacteroides thetaiotaomicron* is widely unaffected by a switch from glucose to sucrose as main carbohydrate source

Jerome Genth¹ | Patrick Kaleja¹ | Christian Treitz¹ | Kathrin Schäfer² |
 Simon Graspentner² | Jan Rupp² | Andreas Tholey¹ 

¹Systematic Proteome Research & Bioanalytics, Institute for Experimental Medicine, Christian-Albrechts-Universität zu Kiel, Kiel, Germany

²Department of Infectious Diseases and Microbiology, University of Lübeck, Lübeck, Germany

Correspondence

Andreas Tholey, Systematic Proteome Research & Bioanalytics, Institute for Experimental Medicine, Christian-Albrechts-Universität zu Kiel, 24105 Kiel, Germany.
 Email: a.tholey@iem.uni-kiel.de

Funding information

Deutsche Forschungsgemeinschaft, Grant/Award Numbers: RU 5042/miTarget, project P8, Cluster of Excellence PMI, RTF-V

Abstract

Bacteroides thetaiotaomicron is a gram negative bacterium within the human gut microbiome that metabolizes a wide range of dietary and mucosal polysaccharides. Here, we analyze the proteome response of *B. thetaiotaomicron* cultivated on two different carbon sources, glucose and sucrose. Two quantitative LC-MS based proteomics approaches, encompassing label free quantification and isobaric labeling by tandem mass tags were applied. The results obtained by both workflows were compared with respect to the number of identified and quantified proteins, peptides supporting identification and quantification, sequence coverage, and reproducibility. A total of 1719 and 1696 proteins, respectively, were quantified, covering 35 % of the predicted *B. thetaiotaomicron* proteome. The data show that *B. thetaiotaomicron* widely maintains its intracellular proteome upon change of the carbohydrates and that major changes are observed solely in the machinery necessary to make use of the carbon sources provided. With respect to the central role of carbohydrates on gut health these data contribute to the understanding of how different carbohydrates contribute to shape bacterial community in the gut microbiome. All proteomics raw data have been uploaded to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD033704.

KEYWORDS

2D-LC-MS/MS, bacterial proteomics, gut microbiome, isobaric labelling, label free quantification

The glycolytic anaerobic gram-negative bacterium *Bacteroides thetaiotaomicron* (*B. thetaiotaomicron*) is an important member of the human gut microbiome [1]. Due to its ability to produce a wide range of polysaccharide depolymerases and glycosylhydrolases [2, 3], this organism is capable to break down a variety of dietary polysaccharides.

High-sugar consumption as prevalent in western diet is regarded to be one causative or at least promoting factor for many diseases such

as inflammatory bowel diseases (IBD) [4, 5], with simple carbohydrates enhancing the susceptibility to colitis [6]. It has been shown, that an increased intake of glucose, sucrose and other simple carbohydrates inhibits gut colonization of *B. thetaiotaomicron* by silencing the ROC protein (regulator of colonization) in addition to altering the composition of other beneficial microbial species in the gut [7, 8]. In contrast, an experimental study in mice provided strong evidence that a high-sugar

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. Proteomics published by Wiley-VCH GmbH

diet promotes the relative abundance of mucolytic bacteria, such as *Bacteroides* genus, leading to the degradation of the mucus barrier, an initiating event of intestinal inflammation [4]. These contradictory findings highlight the need for studies focusing on the functional behavior of the gut microorganisms.

Previous proteomic studies on *B. thetaiotaomicron* focused mainly on the composition of the outer membrane proteome, surfacome, and outer membrane vesicles [9, 10]. To the best of our knowledge, a quantitative analysis of the cellular proteome of *B. thetaiotaomicron* has not yet been presented, especially its reaction on different carbon sources.

In this study, we applied two bottom-up proteomics approaches encompassing label free quantification (LFQ) and isobaric labeling by tandem mass tags (TMT) for a total proteome analyses of *B. thetaiotaomicron*'s grown on the monosaccharide glucose and the disaccharide sucrose, aiming to identify potential changes in the cellular repertoire of the bacterium under two growth conditions known for increasing the risk of gut inflammatory diseases [4, 6]. While most proteome studies focus on the exponential growth phase (resembling the ideal growth situation for a bacterium), the stationary phase has been described to be the most common condition bacteria reside in reality [11]. Thus, retrieving the proteome from the stationary phase is resembling in vivo situation more likely than exponential growth phase and was performed in this study.

We cultivated *B. thetaiotaomicron* on both carbon sources and performed subsequent proteomics analysis in triplicates. The experimental workflows for the two quantification strategies are illustrated in Figure S1. For LFQ, an one-dimensional LC-separation was applied, for the isobaric labeling a two-dimensional separation was applied encompassing a pH10-reversed phase \times pH2-ion-pairing reversed phase separation with concatenation of first dimension fractions [12]. A Q Exactive HF quadrupole Orbitrap MS driven in data-dependent acquisition mode was employed in both experiments. Raw data were filtered prior to quantification using established quality criteria: (a) contaminating proteins (cRAP) were filtered out, (b) only proteotypic peptides were used, (c) proteins had to be identified with at least two peptides with one unique peptide, (d) quantitative data had to be collected in all three biological replicates. In addition, for isobaric labeling, only peptide spectrum matches (PSMs) with a co-isolation interference below 50% were considered for quantification.

The application of a two-dimensional chromatographic separation reduced the percentage of co-isolation interference for PSMs above 50% and below 30% interference, prone to yielding good accuracy [13], by only small margins (3.1% and 4.6%, respectively) (Figure S2B).

The numbers of identified and quantified proteins were similar between the LFQ (2066/1719) and the isobaric labeling approach (2068/1696) (Figure 1A, Table S1). Overall, the identified proteins covered 43% of the predicted *B. thetaiotaomicron* proteome, while 35% of the proteome could be quantified.

In both experiments, more than 73% of the identifications are supported by data with more than five peptides, and 72% of the quantifications by more than five unique peptides, whereas less than 2% were quantified with only one unique peptide (Figure 1C). In addition to the overall comparable number of identified and quantified proteins

between the two approaches, an overlap of 83% of the 2258 identified proteins and 86% of the 1831 quantified proteins was observed, while each individual approach uniquely identified and quantified 6% or 8% proteins, respectively (Figure 1B, Figure S3). With an average sequence coverage of 37%, LFQ performed slightly better compared to isobaric labeling (33%), considering only multiple-peptide identifications (≥ 2 peptides) (Figure S3). On average, peptides in LFQ were identified with 27 PSMS, whereas in isobaric labeling 10 PSMS were provided. Considering proteins, each identification was provided by 10 peptides on average in both datasets (Table S1). Variability analysis shows a higher degree of reproducibility for isobaric labeling results (standard deviation 0.0762) compared to the LFQ results (standard deviation 0.135) (Figure S5). To some extent, this is attributable to the simultaneous sample measurement of a set of TMT 6-plex isobaric labels in a single run [14] and the application of a two-dimensional separation scheme, which has been shown to improve accuracy by 20% [15]. In comparison to the relative quantification of TMT labeled peptides, which is performed on the peptide fragmentation spectra at the MS² level, LFQ is based on comparing the peak area or intensities of precursor ions at the MS¹ level. Therefore, the standard deviation depends on the reliable detection of identical precursor peptides in each run [14, 16]. Inconsistencies in the selection of precursor ions due to the stochastic nature of DDA measurements affect run-to-run reproducibility and lead to the so-called "missing value problem" [17]. Missing values can greatly reduce the completeness of data, impacting the validity of statistical conclusions and downstream functional analyses [18]. In the complete LFQ dataset, 20.4% of all peptides contain at least one missing value. By analyzing each sample four times and applying established quality criteria, data imputation was avoided. In comparison, isobaric labeling dataset revealed only 1.0% of all PSMS with a missing reporter ion on MS² level. In summary, both approaches achieved similar quantitative metrics such as sequence coverage and the number of unique peptides.

Protein groups with a Benjamini-Hochberg adjusted p-value ≤ 0.05 and with \log_2 fold changes (\log_2 FC) of ± 0.485 or ± 1.0 were considered relaxed or strict significantly differentially abundant. A comparison of \log_2 FC values showed good consistency between LFQ and isobaric labeling when quantification was based on an equivalent number of peptides (Table S2). In general, higher \log_2 FC were observed for the LFQ data; for example, for the putative fructokinase (BT1757) isobaric labeling determined a \log_2 FC of 1.66, whereas LFQ detected a \log_2 FC of 5.27. In addition to lower \log_2 FC (about three orders of magnitude smaller), lower p-values were observed for the isobaric labeling approach, which could be explained by the reduced variation (Figure S5).

In summary, 35 of those proteins with a \log_2 FC ≥ 0.485 and two proteins with a \log_2 FC ≤ -0.485 were quantified by LFQ proteomics, whereas isobaric labeling proteomics revealed 28 proteins with a \log_2 FC ≥ 0.485 and four proteins with a \log_2 FC ≤ -0.485 (Figure 2, Table S2) in response to the different carbohydrates. Combining significant IDs from LFQ and isobaric labeling datasets, a total of 54 proteins were identified, including 49 high abundant and five low abundant proteins in sucrose compared to glucose. Of these, 15 proteins were quantified in both proteomic approaches (Figure S6,

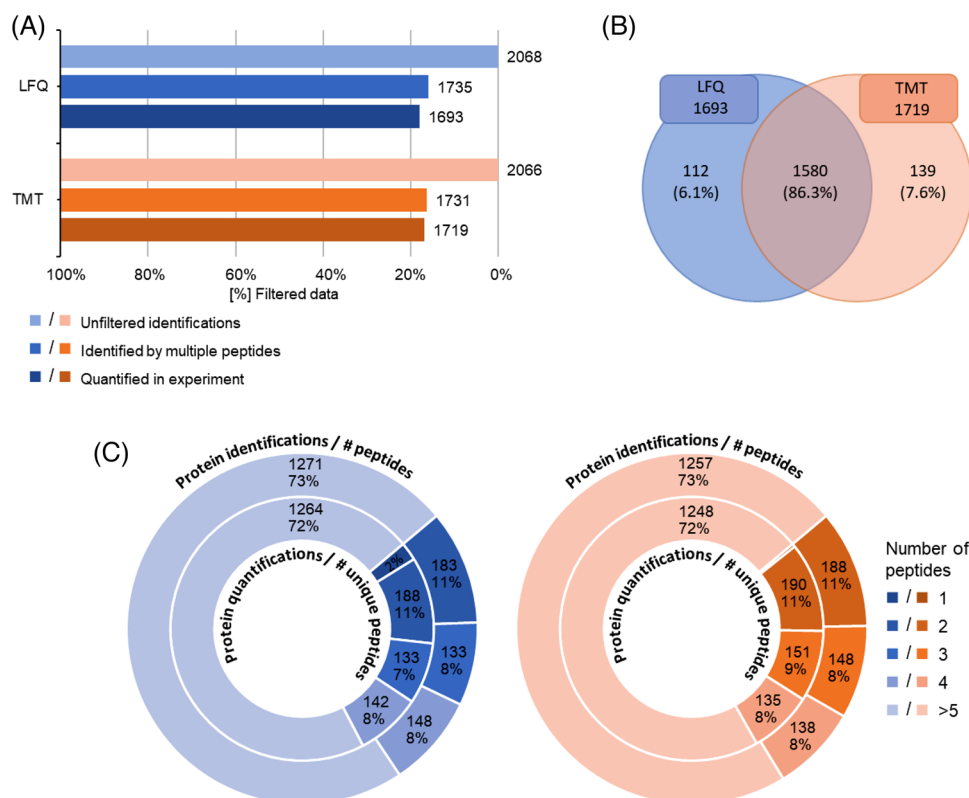


FIGURE 1 Identification and quantification metrics of proteins. Data for LFQ are displayed in blue, of isobaric labeling in orange. (A) Distribution of identified protein groups and data filtering with the applied quality criteria. The total number of protein groups in each category is shown next to columns. (B) Venn diagram illustrating overlap and complementarity of quantified proteins. (C) Number of peptides supporting protein identifications and unique peptides supporting quantitation of the filtered protein groups.

Table S3), including a D-ribitol-5-phosphate phosphatase, which overall presented lower abundance in glucose condition, and seven proteins associated with fructan metabolism (Table S4). Genes encoding the enzymes for fructan metabolism are located on a polysaccharide utilization locus PUL 22 in the genome of *B. thetaiotaomicron*, which contains eight open reading frames. These proteins are involved in the binding, import, degradation, and catabolism of fructose-based oligosaccharides (Figure 3A) [19, 20]. For example, the SusE-like protein (BT1761) is predicted to mediate fructan binding while the SusC-/SusD-like transport system (BT1763 and BT1762) binds fructan oligomers and imports them into the periplasm, where the oligosaccharides are hydrolyzed to monosaccharides [20]. An overview of how *B. thetaiotaomicron* utilizes this pathway to process sucrose is shown in Figure 3B. Three putative glycoside hydrolases of family 32 (GH32) (BT1759, BT1760, BT1765) are involved in oligosaccharide degradation, acting as exo- or endofructanases [3]. Import of monosaccharides occurs through an inner membrane protein (BT1758) and lastly a putative fructokinase (BT1757) catalyzes entry into the glycolytic pathway. To regulate the efficient utilization of fructose-based carbohydrates, a hybrid two-component system (HTCS) (BT1754) directly adjacent to the PUL is essential [20].

Gene ontology (GO) enrichment analysis using a Fisher's exact test with a p-value threshold of 0.05, revealed chemical reactions and metabolic pathways involving carbohydrates to be significantly

enriched (Figure S7A and Figure S8A). These include molecular functions such as binding of carbohydrates and hydrolysis of their glycosidic bond. To illustrate whether the respective fold changes have a preference to be systematically larger or smaller than the global distribution for all proteins we performed a 1D enrichment analysis [21]. The calculated median \log_2FC across both datasets reflects a shift toward positive values for hydrolyzing O-glycosyl compounds, acting on glycosyl bonds and carbohydrate binding (Figure S7B and Figure S8B). Concerning the localization of differentially abundant and total identified proteins, no enrichment or depletion was observed for cytosolic, periplasmic space or membrane proteins.

Our findings were further used to investigate additional polysaccharide utilization sites (PULs), using the Polysaccharide Utilization Loci DataBase (PULDB) [22]. Notably, four other PULs were found to be expressed along with PUL22 in *B. thetaiotaomicron* in the presence of sucrose, including PUL90 and PUL91 (MAN-PUL1 and MAN-PUL2, respectively) and PUL71 and PUL72 (Table S5). A volcano plot of label-free quantification highlights these proteins associated with these PULs (Figure S9). These findings revealed a cohort of enzymes targeting α -mannosidic linkages, indicating metabolism of α -mannose-containing glycans. This may be attributable to the utilized medium, which incorporates yeast extract. A previous study reported that α -mannan, found in yeast cell walls, serves as a nutrient for *B. thetaiotaomicron*, which utilizes it through a so-called *selfish mechanism* [23].

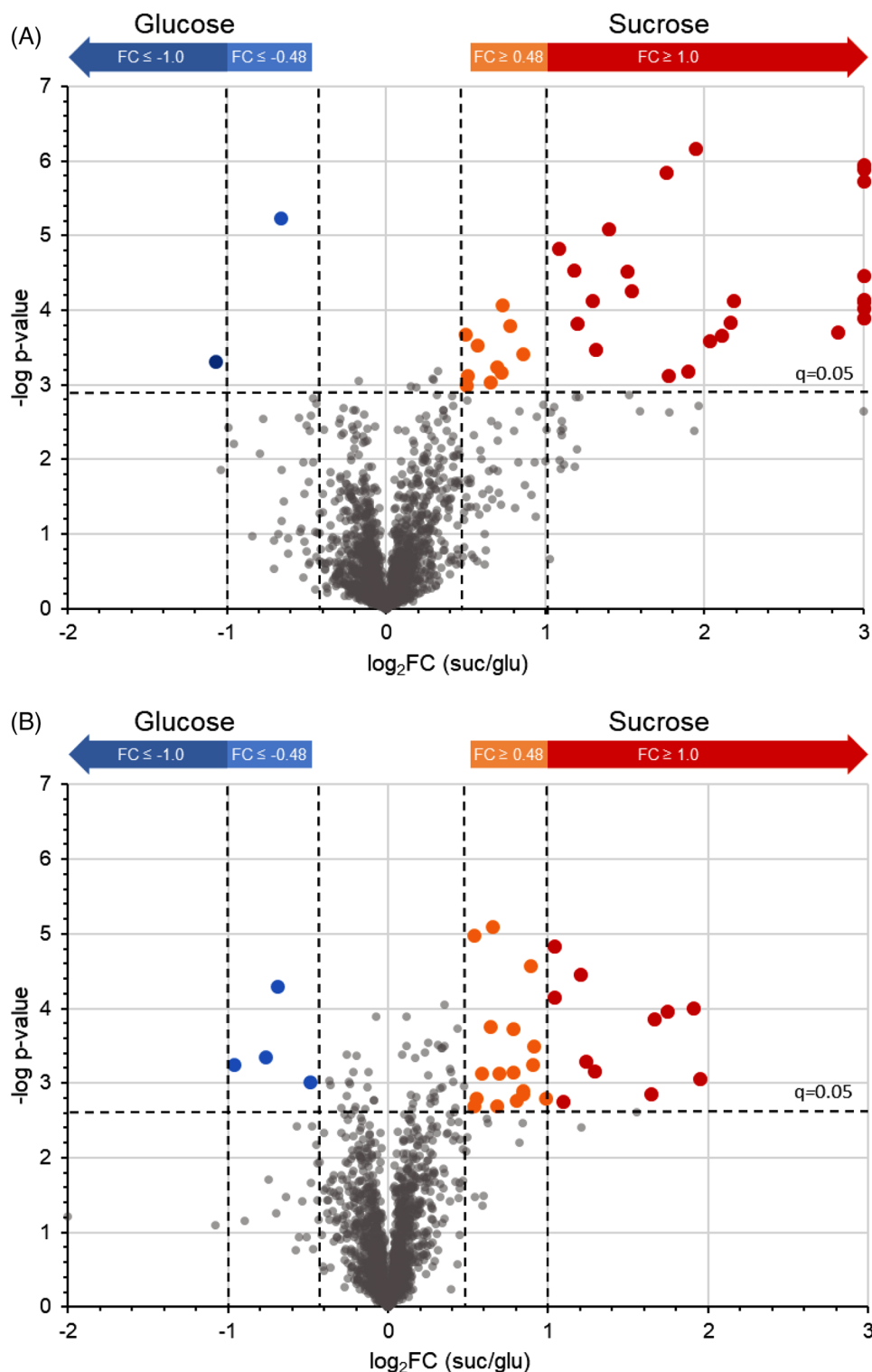


FIGURE 2 Analysis of sucrose associated proteins compared to the LFQ and TMT-dataset of quantified proteins. Volcano plot displaying the log₂FC (sucrose/glucose) versus significance (-log p-value) for all quantified proteins. The dashed vertical lines represent log₂FC cutoffs and the dashed horizontal line represent a q-value of 0.05. Proteins with log₂FC of ± 0.485 or ± 1.0 and a q-value of < 0.05 are considered to be differentially abundant and indicated by colored dots. (A) Volcano plot of label-free-quantification. (B) Volcano plot of isobaric-labeling quantification. Two-sided Welch t-test, $\alpha = 0.05$, Benjamini-Hochberg correction.

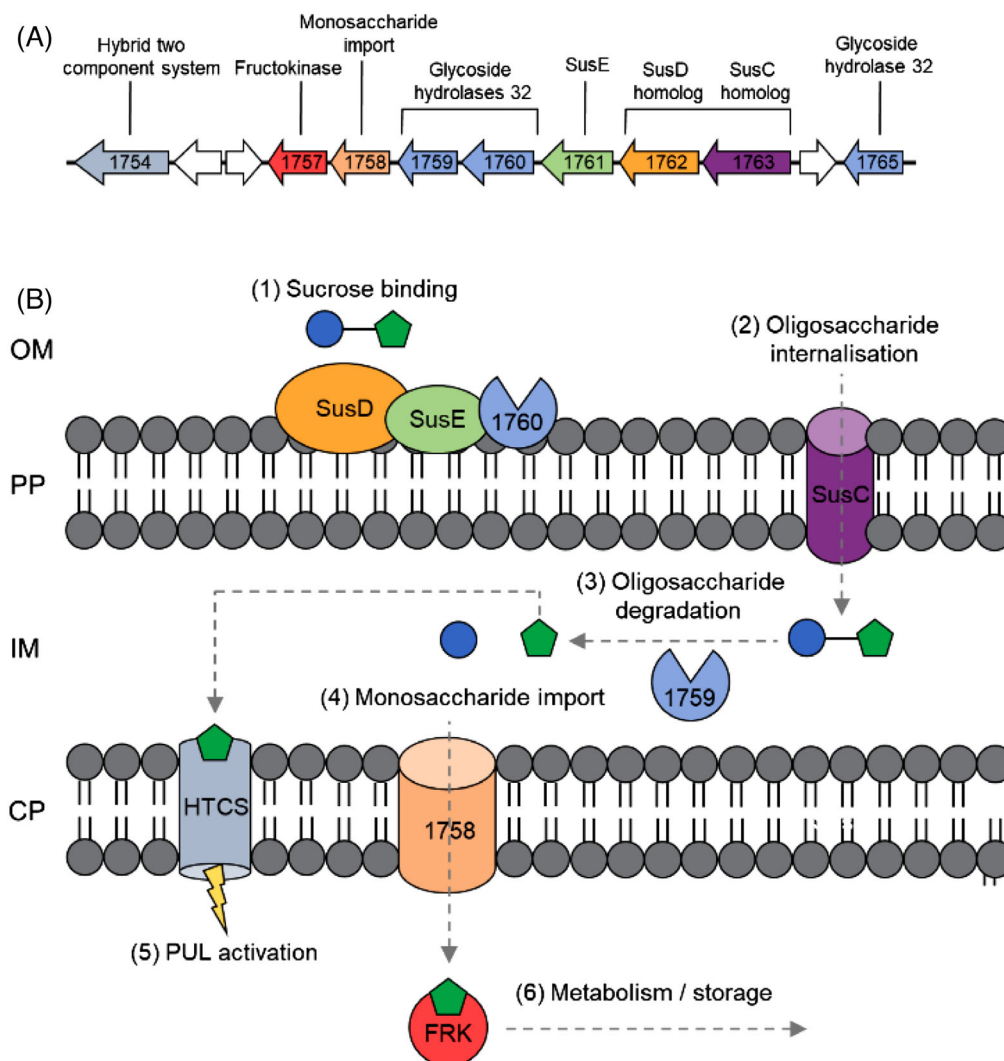


FIGURE 3 Fructan sequestration system in *B. thetaiotaomicron*. (A) Schematic representation of Fructan polysaccharide utilization locus (PUL) as encoded by the *B. thetaiotaomicron* VPI-5482 genome. (B) The predicted mechanism for uptake and degradation of fructose-based oligo-saccharides. CP, cytoplasm; FRK, Fructokinase; HTCS, Hybrid two component system; IM, inner membrane; OM, outer membrane; PP, periplasm; SusC, SusC-like TonB-dependent transporter; SusD, SusD-like outer membrane-binding protein.

Conclusively, our data show that *B. thetaiotaomicron* widely maintains its intracellular proteome upon change of the carbohydrates used for feeding and that major changes are observed solely in the machinery necessary to make use of the carbon sources provided. While the limited variation of the proteome upon change from glucose to sucrose has been shown in another, non-human related case [24], our results emphasize that there may be no fundamental functional difference of *B. thetaiotaomicron* within the gut microbiome in relation to both carbohydrate sources. By integration of further bacterial stressors and/or variation of nutrient supply the plasticity of bacterial ecology can be further exploited in single bacterial isolates or in more complex co-cultures. With respect to the central role of carbohydrates on gut health this approach will foster our understanding of different saccharides on shaping bacterial communities with *B. thetaiotaomicron* being to some degree resilient upon a dietary switch.

ACKNOWLEDGMENTS

This study was supported by the Deutsche Forschungsgemeinschaft (DFG) within the research unit RU 5042 “miTarget”, project P8, and the Cluster of Excellence “Precision Medicine in Inflammation (PMI)”, RTF-V.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data are available on a public repository accession data are stated in manuscript.

ORCID

Andreas Tholey  <https://orcid.org/0000-0002-8687-6817>

REFERENCES

- Eckburg, P. B., Bik, E. M., Bernstein, C. N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S. R., Nelson, K. E., & Relman, D. A. (2005). Diversity of the human intestinal microbial flora. *Science*, 308, 1635–1638.
- Xu, J., & Gordon, J. I. (2003). Honor thy symbionts. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 10452–10459.
- Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., & Henrissat, B. (2009). The carbohydrate-active enzymes database (CAZY): An expert resource for glycogenomics. *Nucleic Acids Research*, 37, D233–D238.
- Khan, S., Waliullah, S., Godfrey, V., Khan, M. A. W., Ramachandran, R. A., Cantarel, B. L., Behrendt, C., Peng, L., Hooper, L. V., & Zaki, H. (2020). Dietary simple sugars alter microbial ecology in the gut and promote colitis in mice. *Science Translational Medicine*, 12.
- Silkoff, K., Hallak, A., Yegena, L., Rozen, P., Mayberry, J. F., Rhodes, J., & Newcombe, R. G. (1980). Consumption of refined carbohydrate by patients with Crohn's disease in Tel-Aviv-Yafo. *Tel-Aviv-Yafo. Postgraduate Medical Journal*, 56, 842.
- Laffin, M., Fedorak, R., Zalasky, A., Park, H., Gill, A., Agrawal, A., Keshteli, A., Hotte, N., & Madsen, K. L. (2019). A high-sugar diet rapidly enhances susceptibility to colitis via depletion of luminal short-chain fatty acids in mice. *Scientific Reports*, 9, 1–11.
- Jones, R. B., Alderete, T. L., Kim, J. S., Millstein, J., Gilliland, F. D., & Goran, M. I. (2019). High intake of dietary fructose in overweight/obese teenagers associated with depletion of Eubacterium and Streptococcus in gut microbiome. *Gut Microbes*, 10, 712–719.
- Townsend, G. E., Han, W., Schwalm, N. D., Raghavan, V., Barry, N. A., Goodman, A. L., & Groisman, E. A. (2019). Dietary sugar silences a colonization factor in a mammalian gut symbiont. *Proceedings of the National Academy of Sciences of the United States of America*, 116, 233–238.
- Elhenawy, W., Debelyy, M. O., & Feldman, M. F. (2014). Preferential packing of acidic glycosidases and proteases into Bacteroides outer membrane vesicles. *mBio*, 5, 1–12.
- Cartmell, A., Muñoz-Muñoz, J., Briggs, J. A., Ndeh, D. A., Lowe, E. C., Baslé, A., Terrapon, N., Stott, K., Heunis, T., Gray, J., Yu, L., Dupree, P., Fernandes, P. Z., Shah, S., Williams, S. J., Labourel, A., Trost, M., Henrissat, B., & Gilbert, H. J. (2018). A surface endogalactanase in Bacteroides thetaiotaomicron confers keystone status for arabinogalactan degradation. *Nature Microbiology*, 3, 1314–1326.
- Gefen, O., Fridman, O., Ronin, I., & Balaban, N. Q. (2014). Direct observation of single stationary-phase bacteria reveals a surprisingly long period of constant protein production activity. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 556–561.
- Treitz, C., Cassidy, L., Höckendorf, A., Leippe, M., & Tholey, A. (2015). Quantitative proteome analysis of Caenorhabditis elegans upon exposure to nematicidal Bacillus thuringiensis. *Journal of Proteomics*, 113, 337–350.
- Sandberg, A., Branca, R. M. M., Lehtiö, J., & Forshed, J. (2014). Quantitative accuracy in mass spectrometry based proteomics of complex samples: The impact of labeling and precursor interference. *Journal of Proteomics*, 96, 133–144.
- Rauniyar, N., & Yates, J. R. (2014). Isobaric labeling-based relative quantification in shotgun proteomics. *Journal of Proteome Research*, 13, 5293–5309.
- Ow, S. Y., Salim, M., Noirel, J., Evans, C., & Wright, P. C. (2011). Minimising iTRAQ ratio compression through understanding LC-MS elution dependence and high-resolution HILIC fractionation. *Proteomics*, 11, 2341–2346.
- Pappireddi, N., Martin, L., & Wühr, M. (2019). A review on quantitative multiplexed proteomics. *ChemBioChem*, 20, 1210–1224.
- Webb-Robertson, B. J. M., Wiberg, H. K., Matzke, M. M., Brown, J. N., Wang, J., McDermott, J. E., Smith, R. D., Rodland, K. D., Metz, T. O., Pounds, J. G., & Waters, K. M. (2015). Review, evaluation, and discussion of the challenges of missing value imputation for mass spectrometry-based label-free global proteomics. *Journal of Proteome Research*, 14, 1993–2001.
- Aittokallio, T. (2009). Dealing with missing values in large-scale studies: Microarray data imputation and beyond. *Briefings in Bioinformatics*, 11, 253–264.
- Martens, E. C., Chiang, H. C., & Gordon, J. I. (2008). Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host and Microbe*, 4, 447–457.
- Sonnenburg, E. D., Zheng, H., Joglekar, P., Higginbottom, S. K., Firbank, S. J., Bolam, D. N., & Sonnenburg, J. L. (2010). Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. *Cell*, 141, 1241–1252.
- Cox, J., & Mann, M. (2012). 1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data. *BMC bioinformatics*, 13((Suppl) 1), S12.
- Terrapon, N., Lombard, V., Gilbert, H. J., & Henrissat, B. (2015). Automatic prediction of polysaccharide utilization loci in Bacteroidetes species. *Bioinformatics*, 31, 647–655.
- Cuskin, F., Lowe, E. C., Temple, M. J., Zhu, Y., Cameron, E. A., Pudlo, N. A., Porter, N. T., Urs, K., Thompson, A. J., Cartmell, A., Rogowski, A., Hamilton, B. S., Chen, R., Tolbert, T. J., Piens, K., Bracke, D., Verweken, W., Hakki, Z., Speciale, G., ... Gilbert, H. J. (2015). Human gut Bacteroidetes can utilize yeast mannan through a selfish mechanism. *Nature*, 517, 165–169.
- Precht, R. M., Janßen, D., Behr, J., Ludwig, C., Küster, B., Vogel, R. F., & Jakob, F. (2018). Sucrose-induced proteomic response and carbohydrate utilization of lactobacillus sakei TMW 1.411 during dextran formation. *Frontiers in Microbiology*, 9, 1–13.

SUPPORTING INFORMATION

Additional supporting information may be found online <https://doi.org/10.1002/pmic.202200189> in the Supporting Information section at the end of the article.

How to cite this article: Genth, J., Kaleja, P., Treitz, C., Schäfer, K., Graspeutner, S., Rupp, J., & Tholey, A. (2022). The intracellular proteome of the gut bacterium *Bacteroides thetaiotaomicron* is widely unaffected by a switch from glucose to sucrose as main carbohydrate source. *Proteomics*, 22, e2200189. <https://doi.org/10.1002/pmic.202200189>