



Characterization and engineering of oligosaccharide transporters

Xiaoqian Li

► To cite this version:

Xiaoqian Li. Characterization and engineering of oligosaccharide transporters. Microbiology and Parasitology. Université de Toulouse, 2024. English. ⟨NNT : 2024TLSEI008⟩. ⟨tel-04756895⟩

HAL Id: tel-04756895

<https://theses.hal.science/tel-04756895v1>

Submitted on 28 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Doctorat de l'Université de Toulouse

préparé à l'INSA Toulouse

Caractérisation et ingénierie de transporteurs d'oligosaccharides

Thèse présentée et soutenue, le 26 septembre 2024 par

XIAOQIAN LI

École doctorale

SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries

Spécialité

Ingénieries microbienne et enzymatique

Unité de recherche

TBI - Toulouse Biotechnology Institute, Bio & Chemical Engineering

Thèse dirigée par

Gabrielle VERONESE et Aurore LABOUREL

Composition du jury

M. Roland Marmeisse, Rapporteur, CNRS, Muséum national d'Histoire naturelle, Paris

M. Henri-Pierre Fierobe, Rapporteur, CNRS, Laboratoire de Chimie Bactérienne, Marseille

M. François Thomas, Examineur, CNRS, Laboratoire de Biologie Intégrative des Modèles Marins, Roscoff

Mme Isabelle Meynial-Salles, Examinatrice, Institut National des Sciences Appliquées, Toulouse Biotechnology Institute, Toulouse

Mme Gabrielle POTOCKI-VERONESE, Directrice de thèse, Institut national de recherche pour l'agriculture, l'alimentation et l'environnement, Toulouse Biotechnology Institute, Toulouse

Mme Aurore Labourel, Co-directrice de thèse, Institut national de recherche pour l'agriculture, l'alimentation et l'environnement, Toulouse Biotechnology Institute, Toulouse

Summary

Mammals host very diverse microbial communities. Among these, the gut microbiota plays a crucial role in host health. Gut bacteria utilize various polysaccharides and oligosaccharides as carbon sources, using complex protein machineries to degrade these substrates into assimilable monosaccharides. Bacteroidota represents one of the dominant phyla in mammal gut microbiomes. They produce a series of proteins necessary for the sensing, capture, transport and degradation of polysaccharides and oligosaccharides. These complex machineries are encoded by polysaccharide utilization loci (PULs). Among the proteins encoded by PULs, cell surface glycan-binding proteins (SGBPs) are essential for the efficient capture of substrates surrounding the cells. In this thesis work, we focused on the characterization and engineering of Bacteroidota transporters, in particular their SGBP components, identified from the human and bovine gut microbiomes. The aim was to elucidate the structure-function relationships of the different protein elements of these transporters, involved in the utilization of dietary or microbial glycans in gut microbiomes. We first characterized the function and crystallographic structure of a SusD-like protein (referred to F5_SusD-like here) encoded by a xylooligosaccharide (XOS) PUL from an uncultured human gut *Bacteroides* species, which is conserved in the prominent *Bacteroides vulgatus* species. Despite its inability to bind to XOS, that are the cognate substrates of the F5_PUL, the F5_SusD-like protein is essential for the uptake functionality. Structural analysis revealed disordered loops and misaligned key residues, which could be responsible for the inability of this SGBP-like protein to bind XOS. Then, we investigated the specificity of F5_SusC/D and F5_MFS transporters towards XOS by introducing a β -mannosidase into the periplasm of the recombinant *E. coli* strains harboring these transporters. This approach aimed at determining if the transporters could uptake mannotriose, which would be degraded to monosaccharides by the β -mannosidase and support *E. coli* growth. Using F5 variants previously constructed by Tauzin *et al.*, we complemented the F5 oligosaccharide utilization system with the BT_0458 β -mannosidase from the GH2 CAZy family. Our preliminary data suggest that the F5 transporters cannot recognize and transport mannotriose, indicating a probable strict specificity towards XOS. Subsequently, we explored the potential of adaptive laboratory evolution (ALE) to engineer the recombinant F5_XOS utilization pathway in *E. coli*. After serial sub-cultures of the F5_XOS containing *E. coli* strains, the mutation rates and positions were assessed by next-generation sequencing. This preliminary study indicated that ALE could serve to improve transporter functions and provided the basis for further investigation of the structure-function relationships of oligosaccharide

transporters under evolutionary pressures. Finally, we characterized another SusD-like protein (referred to 41O1_SusD) encoded by a β -glucan utilization locus from a metagenomic clone isolated from the bovine rumen microbiome. Functional characterization showed that the 41O1_SusD-like protein exhibits binding affinity for β -1,6 branched β -1,3-glucans. Structural comparison with homolog SusD-like proteins highlighted a similar pattern of substrate recognition, involving three tryptophan residues. Our findings provided advances in the understanding of the role played by SusD-like proteins in plant cell wall and/or fungal polysaccharide utilization in the cow rumen. Overall, this thesis generated advances in the understanding of the structure-function relationships of SusD-like proteins. In the long term, the findings will contribute to the development of potential applications in synthetic biology and microbial engineering for enhanced polysaccharide utilization.

Acknowledgements

At this point, my PhD project and thesis have come to an end. The entire PhD thesis research is a long journey. Without the support, guidance and encouragement of many individuals and institutions, I would not have reached the final destination. Here, I have many things and people I want to thank. Without them, I would not have grown to where I am now.

First of all, I would like to express my deepest appreciation and sincere gratitude to my supervisor Research Director Gabrielle Potocki-Veronese and my co-supervisor Aurore Labourel. Gaby and Aurore are such enthusiastic and responsible supervisors! They never tire of reminding me, teaching me, and answering my questions from beginning to end. During the writing process of the thesis, their unwavering support, strict requirements and insightful guidance enabled me to complete this thesis. My graduation project would not have been completed so smoothly without their help. In addition to professional guidance on research, they also give me so much courage and strength like family and friends!

My sincere thanks also go to Elisabeth Laville, she was so kind and helpful when I first arrived at the laboratory, allowing me to feel the first warmth in a foreign country. Subsequently, she also supports me a lot in my project and experiment.

Then, I would like to thank my colleagues, Sophie Lajus, Laurence Tarquis, Alexandra Tauzin, Gianluca Cioci, Sophie Bozonnet, Maxant Vivier, Guy Lippens, Jérémy Esque, for their help in the experiment process and technique. I also want to thank all current and past members of the CIMEs team for their help and support in the past years.

I would like to thank my thesis committee members, Prof. Matthieu Arlat and Dr. François Thomas, for their serving as my committee members and for their fruitful discussion and meaningful suggestions for my project.

I would like to thank my jury members, Dr. Henri-Pierre Fierobe, Dr. Roland Marmeisse, Dr. François Thomas and Prof. Isabelle Meynial Salles. It is my honor to invite them to be part of my jury and take the time to judge my work.

Thanks to the financial support from the China Scholarship Council (CSC) in the frame of CSC-UT/INSA, provided me with an excellent opportunity to study in France and develop my scientific discoveries.

I would like to thank my family and friends, who give me consistent support and encouragement during this challenging period.

Especially, I would like to thank my parents, Ms. Hongxia Ding (丁洪霞) and Mr. Zhongchen Li (李忠臣), for all their help and support over the years. They've been there for me every step of the way, and I'm really grateful for all they've done. I'm so grateful to them for their full support, because without them, I wouldn't be the person I am today. Here, I want to say from the bottom of my heart: Thank you so much, I love you forever! 谢谢你们, 我永远爱你们!

Lastly, I would like to express my deepest gratitude to everyone who has supported me in any way!

Table of Contents

Acknowledgements.....	1
Table of Contents	3
List of figures and tables.....	5
Abbreviations list.....	11
General introduction.....	13
Scientific production	17
Introduction	19
Functional, structural and taxonomical diversity of glycoside transporters in bacteria.....	21
Engineering of oligosaccharide transporters	139
Thesis objectives	191
Results	193
First chapter.....	195
Second chapter.....	219
Third chapter.....	257
Fourth chapter.....	293
Conclusion and perspectives	317
Résumé de la thèse.....	327

List of figures and tables

Introduction	19
 Functional, structural and taxonomical diversity of glycoside transporters in bacteria	
Table 1. Comparison of the different methods used for functional characterization of glycoside transporters.....	26
Figure 1. Schematic representation of the starch utilization system in <i>Bacteroides thetaiotaomicron</i> 5482.....	32
Table 2. Summary of the SusD and SusC/D pairs involved in polysaccharide utilization by Bacteroidota, characterized either biochemically, or by gene deletion or mutation.	37
Figure 2. Transport of maltooligosaccharides by ABC in Gram-negative bacteria.	45
Table 3. List of the bacterial ABC transporters involved in glycoside utilization, characterized either by binding or transport assays, gene deletion and/or transcriptional approaches.	46
Table 4. Phylum and order distribution of the characterized glycoside-targeting ABC transporters in bacteria.	67
Figure 3. Schematic illustration of the process of glycoside transport and phosphorylation by the PTS.	70
Table 5. List of the bacterial PTS transporters involved in glycoside utilization, characterized either by binding or transport assays, gene deletion and/or transcriptional approaches.	71
Figure 4. Carbohydrate transport through the transmembrane PTS_EIIC domain.....	81
Table 6. Phylum and order distribution of the characterized glycoside-targeting PTS transporters in bacteria.	82
Figure 5. Schematic representation of the three types of MFS transporters.	83
Table 7. List of the bacterial MFS transporters involved in glycoside utilization, characterized either by binding or transport assays, gene deletion and/or transcriptional approaches.	86
Figure 6. Alternating access mechanism for a H ⁺ /lactose symport.....	91

Table 8. Phylum and order distribution of characterized glycoside-targeting MFS transporters in bacteria.	92
Engineering of oligosaccharide transporters	
Table 1: Summary of the strategies of oligosaccharide transporter engineering for 4 different transporter families coming from fungi and bacteria, including the biochemical and structural information of parental proteins and the new functions of engineered transporters.	141
Table 2: Phylum and order distribution of the engineered oligosaccharide transporters in microorganisms	160
Table 3: Transporter families of the engineered oligosaccharide transporters	160
Figure 1: Spontaneous mutations using physical or chemical agents.	163
Figure 2: An overview of adaptive laboratory evolution (ALE).	164
Figure 3: An overview of site-directed mutagenesis.	166
Figure 4: An overview of saturation mutagenesis.	167
Figure 5: An overview of the directed evolution process.	171
Results	193
First chapter	
Figure 1. Affinity gel electrophoresis of F5_SusD-like on a gel containing no polysaccharide (negative control) and on gels with 0.5 % (w/v) of different xylans.	199
Figure 2. Characterization of the ligand binding abilities of F5_SusD-like using STD-NMR.	200
Table 1. Data collection and refinement statistics for the crystal structure of F5_SusD-like.	201
Figure 3. F5_SusD-like crystal structure.	202
Table S1. Molecular weight comparison of F5_SusD-like with SusD-like proteins from xylan targeting PULs.	209

Table S2. List of PCR primers used in this study.....	209
Figure S1. Affinity gel electrophoresis of F5_SusD-like on a gel containing no polysaccharide (negative control) and on gels with 0.5 % (w/v) of different β -glucans. ..	210
Figure S2. Analysis of F5_SusD-like oligomerization.....	211
Figure S3. 1D proton spectrum at 298K of 500 μ l of the F5_SusD-like protein at 3.8 mg/ml (54 μ M) in Tris buffer.....	212
Figure S4. (left) 1D proton spectrum at 298K of 500 μ l of xylan at 500 μ M with F5_SusD-like at 0 (blue), 15 (red) or 27 μ M (magenta) in Tris buffer. (right) STD spectra of the different samples with saturation at 0 ppm for the isolated xylan (blue), in presence of 15 (red) or 27 μ M (magenta) of F5_SusD-like. The reference xylan spectrum is shown in black.....	213
Figure S5. (left) 1D proton spectrum of 500 μ l of xylan at 500 μ M with F5_SusD-like at 0 (blue), 15 (red) or 27 μ M (magenta) in Tris buffer $T_e = 298$ K (right). STD spectra of the different samples with saturation at 10.1 ppm for the isolated xylan (blue), in presence of 15 (red) or 27 μ M (magenta) of F5_SusD-like. The reference xylan spectrum is shown in black... ..	214
Figure S6. Close-up view of the interaction between the loop 292-315 and a symmetry related molecule.	215
Figure S7. Comparison of the oligosaccharide binding areas of SusD-like proteins with the homologous region of F5_SusD-like.	216
Figure S8. Structural comparison of SusD-like proteins alone and in interaction with cognate SusC-like proteins.	217

Second chapter

Figure 1: Chemical structure of xylotriose (A) and mannotriose (B).	224
Figure 2: Representation of the F5_XOS_PUL and three variants.	225
Figure 3: Representation of the internalization and hydrolysis processes of mannotriose in the metagenomic clones issued from truncation of the F5 locus.	226
Figure 4: Sequence analysis of the BT_0458 <i>exo</i> - β -mannosidase using SignalP 6.0.	227

Figure 5: Gene organization of the variants and their combinations with the two different recombinant GH2 constructs.....	229
Figure 6: Agarose gel electrophoresis of colony-PCR screening with T7 universal primers.....	230
Figure 7: TLC analysis of the degradation profiles of 0.2 % (w/v) mannotriose and xylotriose.....	236
Figure 8: Growth curves of the different F5 variants in M9 minimal medium containing either xylotriose (XTR) or mannotriose (MTR) at 0.5 % (w/v) as sole carbon source. The media also contain antibiotics, 12.5 mg/l of chloramphenicol and 50 mg/l of kanamycin (for the controls F5minSus and F5minMFS no kanamycin was added).....	239
Figure 9: Strategy for regulating gene expression using the pET-expression system.....	241
Third chapter	
Figure 1: Schematic representation of the F5 XOS utilization locus and the truncated versions used in the present work.....	262
Figure 2: Overview of adaptive laboratory evolution (ALE) strategy used in the present work for improvement of xylotriose metabolization by an <i>E. coli</i> strain expressing the F5minSus locus (orange).....	263
Figure 3: Growth curves of the five serial cultures of F5minSus (A) and F5minGH (B) in M9 minimal medium supplemented with 12.5 mg/l chloramphenicol and 0.5 % (w/v) xylotriose as the sole carbon source.....	264
Figure 4: Xylotriose consumption by the five serial cultures of F5minSus (A) and F5minGH (b) growing on xylotriose.....	266
Figure 5: HPAEC-PAD chromatograms of the culture supernatants analyzed at different growth times.....	271
Figure 6: Positions of the mutations detected along the F5minSus locus by NGS read analysis.....	273
Table 1: Positions, type and relative abundance of the most prominent mutations detected along the F5minSus locus in the S3 and S5 samples.	274

Figure 7: Cell growth screening in droplet-microfluidics, based on accumulation of fluorescent flavins in pico-liter droplets during cell growth (patent EP22305759.7). 283

Figure 8: Experimental workflow that will be used to characterize the evolved strains. .. 284

Fourth chapter

Figure 1: Representation of the 41O1 contigs..... 297

Figure 2: Growth curves of the clone 41O1 on β -1,3/ β -1,4-linked gluco-oligosaccharides at 0.5 % (w/v), highlighting significant growth ability on laminaritriose and a mixture of β -d-cellobiosyl-cellobiose and β -d-glucosyl-cellobiose, and slight growth is observed with cellobiose. 298

Figure 3: Cellular localization of enzymatic activity based on analysis of oligosaccharide hydrolysis reaction products by TLC, suggesting that clone 41O1 presents both membrane-bound and cytoplasmic β -1,3-glucosidase activities. 299

Figure 4: Enzymatic activities of clone 41O1, highlighting significant activity of the GH3 and/or the GH16 proteins on laminarins and barley β -glucan. 300

Figure 5: Affinity gel electrophoresis of 41O1_SusD-like on a gel containing no polysaccharide (control with H₂O) and on gels with different β -glucans, with bovine serum albumin (BSA) as a control protein..... 301

Table 1: Thermal shift analysis of 10 μ M 41O1_SusD-like protein incubated with 0.5 % (w/v) ligands. 302

Figure 6: Characterization of the ligand-binding abilities of 41O1_SusD-like using NMR spectroscopy, demonstrating the binding affinity of 41O1_SusD-like for laminaritriose.. 303

Figure 7: Structural analysis of the best AF2 model of 41O1_SusD-like mature protein. 304

Figure S1: Dot Blot of the 41O1 sequence against itself, using the nucleotide Blastn tool from the NCBI, highlighting repeated sequences in this clone..... 313

Figure S2: Restriction analysis of the 41O1 clone. 314

Figure S3: (A) 41O1_SusD-like full-length protein sequence. (B) Gel filtration of 41O1_SusD-like (57 kDa)..... 315

Figure S4: Thermal shift analysis of 10 μ M 41O1_SusD-like protein incubated with 0.5 % (w/v) ligands. Vertical lines indicate inflection temperatures (Ti). 316

Abbreviations list

ABC:	ATP-binding cassettes
AGE:	Affinity gel electrophoresis
ALE:	Adaptive laboratory evolution
CAZymes:	Carbohydrate-active enzymes
CBMs:	Carbohydrate-binding modules
CIMES:	Catalyse et Ingénierie Moléculaire Enzymatiques
DSF:	Differential scanning fluorimetry
DP:	Degree of polymerization
ECF:	Energy-coupling factor
ECF- σ :	Extracytoplasmic function sigma
EI:	Enzyme I
EII:	Enzyme II
GAGs:	Glycoaminoglycans
GHs:	Glycoside Hydrolases
FACS	Fluorescence activated cell sorting
HPAEC-PAD:	High performance anion exchange chromatography with pulsed amperometric detection
HGM:	Human gut microbiota
HPr:	Histidine protein
HTCSs:	Hybrid two-component systems
ITC:	Isothermal titration calorimetry
MFS:	Major facilitator superfamily
MSA:	Multiple sequence alignments
NBDs	Nucleotide-binding domains
PEP:	Phosphoenolpyruvate
PSI-BLAST:	Position-Specific Iterated BLAST

PTS:	Phosphotransferase systems
PUL:	Polysaccharide utilization locus
PULDB:	Polysaccharide-Utilization Loci Data Base
qPCR:	Quantitative real-time PCR
ROK:	Repressors, ORFs and Kinases
SBP:	Substrate-binding protein
SD:	Scaffold domain
SDS-PAGE:	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGBP:	Surface glycan-binding protein
SPR:	Surface plasmon resonance
STD-NMR:	Saturation transfer difference nuclear magnetic resonance
SUS:	Starch utilization system
TBDT:	TonB-dependent transporters
TCDB:	Transporter Classification Database
TCID:	Transporter Classification Identifier
TD:	Transport domain
TMD:	Transmembrane domain
TMG:	Thio- β -methyl galactopyranoside
TONPG:	β -thio-o-nitrophenylgalactoside
TPR:	Tetratricopeptide repeat
XOS:	Xylooligosaccharide

General introduction

Mammals harbor very diverse microbial communities. Gut microbiota, in particular, play a crucial role in the overall health and well-being of mammals. These intricate communities, consisting of bacteria, fungi, bacteriophages have emerged as a fascinating area of research, revealing its profound influence on various aspects of mammalian physiology (Camarillo-Guerrero et al. 2021; Chibani et al. 2022; Gregory et al. 2020; Raimondi et al. 2019; Huttenhower et al. 2012). For instance, mammal gut microbiota are involved in the food digestion and interact with the immune system, significantly influencing host development and health (Rooks and Garrett 2016). Gut bacteria mainly use polysaccharides and oligosaccharides from dietary components (such as plant storage and cell wall polysaccharides), from host glycoproteins, or of microbial origin as carbon sources. To cope with the structural diversity of these substrates, gut bacteria have evolved various and complex protein machineries that act in synergy to break down polysaccharides into oligosaccharides and, ultimately, into monosaccharides which support microbial cell growth in the gut.

Oligosaccharides are composed of less than 20 monosaccharide units. They serve as crucial energy sources and signaling molecules in various biological systems. The use of oligosaccharides as dietary prebiotics in food and feed results in complex effects on the gut bacterial communities. For example, fructo-oligosaccharides, galacto-oligosaccharides and lactulose, which are largely commercialized as functional foods, have been shown to alter the balance of the large bowel microbiota by increasing the abundance of *Bifidobacteria* and *Lactobacillus* species in the human gut (Macfarlane et al. 2006). In bacteria, the uptake of oligosaccharides and their breakdown into monosaccharides are essential processes for their survival and proliferation. The major phyla in mammal gut microbiota (Bacteroidota, Bacillota, Actinomycetota and Pseudomonadota) have developed very diverse glycan utilization systems that match their specific cell structures and nutrient environments (Briggs et al. 2021). Gram-negative species, exemplified by the widespread Bacteroidota, necessitate systems that can transport glycans and their degradation products across two membranes via the periplasm (Martens et al. 2009). Bacteria have developed specific polysaccharide utilization loci to facilitate the utilization of structurally complex glycosides. These loci encode several common functional elements, including transporters and carbohydrate-active-enzymes (CAZymes), which are necessary for the import of these substrates into the cells and their subsequent degradation into monosaccharides (Grondin et al. 2017). Four different superfamilies of

oligosaccharide transporters have been described, to date, in bacteria, each exhibiting unique mechanisms and architectures: ATP-binding cassettes (ABC), major facilitator superfamily (MFS) transporters, phosphotransferase systems (PTS), and TonB-dependent transporters (TBDT).

Over several past decades, there has been a growing interest in the study of oligosaccharide transporters, since these proteins ensure pivotal roles in microbial ecology, as well as for fermentation processes and synthetic biology, allowing the chassis strains to grow on low-cost substrates of plant origin (such as cellobiose, maltose or sucrose, for instance).

Of particular interest to our team are the SusC/D transporters, members of the TBDT family, that is specifically found in Bacteroidota. The designation 'SusC/D' is derived from the starch utilization system (Sus) initially identified by the Salyers group in the late 1980s (Tancula et al. 1992). The Sus includes the requisite proteins for binding and degrading starch at the surface of the cell, thereby facilitating subsequent transport of oligosaccharides across the outer membrane for further depolymerization to glucose in the periplasm (Foley et al. 2016). In particular, the two main components of the transporter, SusC and SusD, have been identified as the main contributors to substrate uptake. SusC is a transmembrane TonB-dependent transporter, while SusD is a cell surface glycan-binding protein (SGBP) involved in the capture of extracellular polysaccharides and oligosaccharides (Shipman et al. 2000; Reeves et al. 1997; Pollet et al. 2021).

The SusC protein physically interacts with the SusD protein to form a complex which mediates substrate uptake via a “pedal bin” mechanism (Glenwright et al. 2017). The SusC transporter constitutes the barrel of the bin, while SusD is situated on top of the barrel, opening and closing like a lid to facilitate substrate binding. The cooperation of these proteins with the other ones encoded by the Sus operon facilitates the uptake and utilization of complex carbohydrates by all the *Bacteroides* species, contributing to their overall fitness in the gut environment and influencing both microbial ecology and host-microbiota interactions. When I started my PhD work, only 11 SusC/D systems had been functionally or structurally characterized, which was very little compared to CAZymes. Furthermore, except one, all these SusC/D transporters were issued from prominent gut *Bacteroides* species, that are cultivated and of which the genomes have been sequenced. Only one study (published by the CIMEs team in which I did my PhD work) dealt with a SusC/D transport system belonging to an uncultured *Bacteroides* species retrieved from the human gut microbiome using functional metagenomics. Thus, there is a huge

lack of knowledge regarding oligosaccharide transporters from uncultured bacteria, which dominate in mammal gut microbiota.

In this context, my thesis work was focused on the functional characterization and engineering of oligosaccharide transporters that were previously identified by the CIMEs team using functional metagenomics of the human and bovine gut microbiomes. The objective was to get knowledge on the structure-function relationships of the different protein elements of the transporter systems involved in the degradation of dietary glycans.

In order to provide a comprehensive understanding of the context of my thesis work, this manuscript starts with a bibliographic introduction. Subsequently, the results obtained during this thesis work will be presented, with a particular focus on the structure-function relationships of SusD-like proteins and the engineering of SusC/D transporters.

This manuscript includes five chapters, each of which being written in the form of a review or an article:

The bibliographic introduction includes two sections. Part A is an overview of the various glycoside transporter families present in bacteria. This article is co-authored by Zhi Wang, Alexandra S. Tauzin, Nicolas Terrapon, Matthieu Arlat, Xiaoqian Li, Aurore Labourel, Gabrielle Potocki-Veronese. I have provided data on the most recently characterized oligosaccharide transporters to update this review, that will be submitted to *Biological Reviews*. Part B is a review article in preparation, which focuses on the different strategies employed in the engineering of oligosaccharide transporters.

The first results chapter is an article published in the journal *mSphere* in 2022. It focuses on the functional and structural characterization of the SusD-like protein from the F5 PUL involved in XOS uptake by an uncultured *Bacteroides* from the human gut microbiota. This article is co-authored by Alexandra S. Tauzin, Zhi Wang, Gianluca Cioci, Xiaoqian Li, Aurore Labourel, Barbara Machado, Guy Lippens and Gabrielle Potocki-Veronese. I contributed to it by producing and purifying the SusD-like protein to perform the complementation experiments, affinity gel electrophoresis analyses and differential scanning fluorimetry for the binding ability.

The second chapter is dedicated to the study of the specificity towards oligosaccharides of the SusC/D and MFS transporters from the metagenomic locus described in the previous chapter. A new strategy was proposed in which a β -mannosidase was introduced into the periplasm of

bacterial strains harboring F5_SusC/D and F5_MFS transporters. The objective was to determine if the transporters could uptake mannotriose (structurally close to xylotriose), which would then be degraded to monosaccharides by the β -mannosidase to support the growth of the *E. coli* strain.

The third chapter is dedicated to the evolution of the SusC/D transporter studied in the two previous chapters, using Adaptive Laboratory Experiment. This is a preliminary study to assess the potential of such an approach to engineer recombinant oligosaccharide transporters expressed in *E. coli* and also to perform the first SusC/D engineering study based on random mutagenesis, in order to improve oligosaccharide transport efficiency.

The fourth chapter is an article recently accepted for publication in the journal *mSphere*. It focuses on the biochemical characterization of a SusD-like protein encoded by a β -glucan utilization locus from a metagenomic clone, which was previously isolated by functional screening of the bovine rumen microbiome. This study aimed at determining the specificity of the SusD protein towards β -linked-glucooligosaccharides and highlighting the molecular determinants of its preference to β -1,6 branched β -1,3-glucans.

The conclusions and perspectives at the end of this manuscript are a summary of the advances made in this thesis work in the understanding of the structure-function relationships of SusD-like proteins and an analysis of the advantages and limitations of the different strategies that could be used for transporter characterization and engineering studies, in order to provide further insights into this intriguing aspect of bacterial physiology.

Scientific production

Research papers

- **Xiaoqian Li**, Guy Lippens, Jean-Luc Parrou, Gianluca Cioci, Jérémy Esque, Zhi Wang, Elisabeth Laville, Gabrielle Potocki-Veronese, Aurore Labourel. Biochemical characterization of a SusD-like protein involved in β -1,3-glucan utilization by an uncultured cow rumen *Bacteroides*. *mSphere*. (Accepted on 18/06/2024)

- Alexandra S. Tauzin, Zhi Wang, Gianluca Cioci, **Xiaoqian Li**, Aurore Labourel, Barbara Machado, Guy Lippens, Gabrielle Potocki-Veronese. Structural and biochemical characterization of a non-binding SusD-like protein involved in xylooligosaccharide utilization by an uncultured human gut *Bacteroides* strain. *mSphere*. (2022). 10.1128/msphere.00244-22

Review papers

- Zhi Wang, Alexandra S. Tauzin, Nicolas Terrapon, **Xiaoqian Li**, Matthieu Arlat, Aurore Labourel, Gabrielle Potocki-Veronese. Functional, structural and taxonomical diversity of glycoside transporters in bacteria. (To be submitted to *Biological Reviews*)

- **Xiaoqian Li**, Aurore Labourel, Gabrielle Potocki-Veronese. Engineering of oligosaccharide transporters. (In preparation)

Oral and poster communications

- Alexandra S. Tauzin, Zhi Wang, Gianluca Cioci, **Xiaoqian Li**, Aurore Labourel, Barbara Machado, Guy Lippens and Gabrielle Potocki-Veronese. Crystallographic structure of a non-binding SusD-like protein involved in xylooligosaccharide utilization by a human gut uncultured *Bacteroides*. 14th Carbohydrate Bioengineering Meeting (CBM). September 25th - 28th, 2022, Norefjell, Norway. (Flash presentation and poster)

- **Xiaoqian Li**, Guy Lippen, Jean-Luc Parrou, Gianluca Cioci, Jérémy Esque, Zhi Wang, Elisabeth Laville, Gabrielle Potocki-Veronese, Aurore Labourel. Biochemical and Structural characterization of a SusD-like protein involved in glucooligosaccharide utilization by a cow rumen uncultured *Bacteroides* strain. Enzyme Engineering XXVII. October 1st-6th, 2023, Singapore. (Poster presentation)

Introduction

Functional, structural and taxonomical diversity of glycoside transporters in bacteria

Zhi Wang^{1†}, Alexandra S. Tauzin¹, Nicolas Terrapon², Matthieu Arlat³, Xiaoqian Li¹, Aurore Labourel¹, Gabrielle Potocki-Veronese^{1#}

¹TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France

²Architecture et Fonction des Macromolécules Biologiques, CNRS and Aix-Marseille University, Marseille, France.

³Laboratoire des Interactions Plantes-Microbes-Environnement, INRAE, CNRS, Castanet-Tolosan, France. Université de Toulouse, Toulouse, France.

[†]Present adress: School of Chemical Engineering, Zhengzhou University, Zhengzhou, Henan, 450001, China

[#]Corresponding author

Abstract

Glycoside metabolization is a crucial function for bacteria. To feed on various glycosides with complex structures, bacteria have developed very diverse proteic machineries to recognize, transport and degrade them. Despite the crucial role of transporters, their identification and characterization remain limited due to the fact that these membrane proteins are difficult to study in recombinant form, and that the majority of bacteria are uncultured. In this review, we embodied the available methods and technologies for the identification and characterization of glycoside transporters from bacteria. We synthesized the current knowledge of experimentally proven glycoside transporters from different families, including TonB-dependent transporters, ATP-binding cassettes, the phosphotransferase system and the major facilitator superfamily, highlighting their taxonomical, functional, mechanistic and structural diversity. The limited number of characterized glycoside transporters versus the available sequence data underlines the need for development of efficient approaches to characterize their specificity and to better understand the molecular mechanisms of glycoside utilization in bacteria.

Key words: glycoside transporters, oligosaccharides, TonB-dependent transporters, ATP-binding cassettes (ABC), phosphotransferase system (PTS), major facilitator superfamily (MFS)

CONTENTS

I.	Introduction	24
II.	Techniques for functional characterization of glycoside transporters	25
	(1) Analysis of growth phenotypes	26
	(2) Determination of binding specificity	28
	(3) Determination of uptake specificity	30
III.	TonB-dependent transporters and associated surface glycan-binding proteins.....	31
	(1) The SUS paradigm.....	31
	(2) Functional diversity of SUS-like systems	32
	(3) Structure and mechanism of TonB-dependent transporters and associated surface glycan-binding proteins.....	33
	(4) Taxonomical diversity of TonB-dependent transporters	36
IV.	ABC transporters of glycosides	43
	(1) Diversity of sequences and structures	43
	(2) Functional diversity	44
	(3) Taxonomical diversity	65
	(4) Mechanism of glycoside import by ABC transporters	65
V.	PTS transporters of glycosides	68
	(1) Diversity of sequences and structures	68
	(2) Functional diversity	69
	(3) Taxonomical diversity	79
	(4) Mechanism of glycoside import by PTS transporters	79
VI.	MFS transporters of glycosides	83

(1) Diversity of sequences and structures	84
(2) Functional diversity	84
(3) Taxonomical diversity	90
(4) Mechanism of glycoside import by MFS transporters	90
VII. Conclusions	93
VIII. ACKNOWLEDGEMENTS.....	93
IX. REFERENCES	94

I. Introduction

Carbohydrates are one of the major biological macromolecules found in living systems. They seldom occur in nature as monosaccharides (Prestegard et al. 2017). Instead, glycosyl units are mostly linked among themselves, or with other non-glycosyl groups, to form glycosides (Cummings and Stephen 2007). Glycosides are structurally highly diverse because of the variety of glycosyl units occurring and the types of glycosidic linkages of which they are composed. Glycosides represent the main energy sources for microorganisms, regardless of habitat (Ladevèze et al. 2017; Grondin et al. 2017; Flint et al. 2008). To deal with the structural diversity of glycosides, bacteria, as fungi (Barbi et al. 2021), have evolved a large panel of highly specific enzymes and transporters to efficiently capture and metabolize them. Knowledge of the specificities and mechanisms of transporters is an important prerequisite in understanding how bacteria metabolize glycosides, and thus, colonize their habitats. This is an important field for human and animal nutrition and health, since the type of bacteria that selectively ferment plant and host-derived glycosides in the gut influences the well-being and health of the host (Koropatkin et al. 2012). Furthermore, from a biotechnological perspective, the robust processes that transform plant biomass into high-value products rely on the use of efficient chassis strains with engineered transporters and metabolic networks (Lee et al. 2013; Lian et al. 2014). However, we only know the function of a small fraction of glycoside transporters. Even in the model organism *Escherichia coli*, approximately 13 % of the open reading frames encode membrane transport proteins, while more than half of them lack experimental validation (Ren and Paulsen 2007; Genée et al. 2016).

The advances in the development of culturing, sequencing and computing technologies results in an increasing number of glycoside transporters sequences (Gelfand and Rodionov 2008), which are listed and classified in the Transporter Classification Database (TCDB, <http://www.tcdb.org/>) (Saier et al. 2016), together with transporters of monosaccharides (Severi et al. 2021), metals or amino acids. Based on sequence similarity, glycoside transporter sequences can indeed be assigned to one of the four different superfamilies of transmembrane proteins, which present different mechanisms and architectures: ATP-binding cassettes (ABC), major facilitator superfamily (MFS) transporters, phosphotransferase systems (PTS) and TonB-dependent transporters (TBDT). Nevertheless, their specificity towards glycosides may vary dramatically among members of the same family (Diallinas 2014). The prediction and experimental validation of transporter specificity thus remain a challenging task.

The transport of glycosides in bacteria has been reviewed several times in the last few years, each time with a specific focus on (i) the human gut (Schwalm and Groisman 2017; Cockburn and Koropatkin 2016); (ii) plant polysaccharides and prebiotics (Goh and Klaenhammer 2015); or (iii) the TBDT family of transporters (Bolam and van den Berg 2018; Grondin et al. 2017). However, none of these studies synthesizes the current knowledge of experimentally proven glycoside transporters, neither addresses the challenges of the biochemical characterization of transporter specificity. However, this is crucial to improve the quality of databases and functional annotation of transporter sequences, and, furthermore, to establish the relationships between the structure and function of glycoside transporters. Here, we present the structural, taxonomical and functional diversity of glycoside transporters in bacteria, focusing in particular on the different methods used to characterize their functions, with their advantages and their limits.

II. Techniques for functional characterization of glycoside transporters

Transporter sequences can be assigned to a specific superfamily or family thanks to the presence of specific motifs, and based on sequence similarity to previously characterized transporters.

However, as members of the same family can transport various glycosides (Wang et al. 2020), it is difficult and often impossible to infer transporter specificity from sequence analysis alone, without experimental evidence. Moreover, there might also be transporters with unassigned function because of their low sequence similarity with characterized transporters, such as the

2.A.1.85 family of MFS transporters listed in the TCDB. Thus, a set of methods and techniques have emerged to identify the correct substrate for a given predicted glycoside transporter. We have classified them into three categories: analysis of growth phenotypes, determination of binding specificity, and determination of uptake specificity (Table 1).

Table 1. Comparison of the different methods used for functional characterization of glycoside transporters

	Growth phenotype	Binding specificity	Uptake specificity
Approaches	Growth assays	AGE, ITC, DSF, SPR, STD-NMR	Uptake experiments with radiolabeled glycosides
Advantages	- easily implemented	- easily determined and/or quantified	- transport ability quantified
Drawbacks	- coupling to transcriptomics and/or genome engineering required due to functional redundancy	- limited to transport system including SBP - transport specificity not reflected precisely	- limited to available radiolabeled glycosides - recombinant production of transmembrane protein or genome engineering required due to functional redundancy in native strains

(1) Analysis of growth phenotypes

The simplest method of screening for the ability of a strain to utilize a glycoside and thus to internalize its substrate or one of its components, is to perform growth assays by monitoring the turbidity of the culture medium. Glycoside utilization can also be easily detected by using MacConkey indicator plates, which contain a solid differential media used to selectively isolate enteric Gram-negative bacteria (such as *Escherichia coli*) and differentiate lactose fermenting Gram-negative rods, as positive colonies turn red when they have been able to utilize the lactose (Peng et al. 2009).

The choice of glycosides to be tested can be rationalized using genome sequencing, to identify putative operons coding for complete machineries of glycoside sensing, binding, translocation through the cell membrane, and degradation by CAZymes (Lombard et al. 2014). The type of

substrate (for example pectins, xylans, mannans, cellulose, starch, *O*- and *N*-glycans, etc.) targeted by an operon can be predicted from the putative specificities of each of the CAZymes it encodes. Enzyme specificities are inferred from sequence homology with the numerous biochemically characterized members of the same family, which are listed in the CAZy database (<http://www.cazy.org>) (Lombard et al. 2014). Such genomic environment analyses recently enabled the prediction of the specificity of three ABC transporters from *Thermus thermophilus* HB8 towards oligosaccharides in starch and β -glucans, with the polysaccharides targeted by the CAZymes surrounding the transporter in each locus (Chandravanshi et al. 2019).

The up-regulation of the transporter-encoding gene (or of the entire operon) during growth on a particular glycoside is a good indication that this transporter is involved in the utilization of this substrate. Quantification of gene expression levels requires transcriptomic analysis, either using quantitative real-time PCR (qPCR) (Tauzin, Laville, et al. 2016) or RNA chips (Martens et al. 2011), DNA microarrays (Andersen et al. 2012) or more recently, RNA sequencing (Despres et al. 2016). With the improvements in sensitivity and the decrease in cost, large-scale transcriptomic studies can be undertaken to compare the transcriptomes of thousands of bacteria (Lowe et al. 2017), giving indications of the specificity of putative transporters towards glycosides. However, transcriptomic analysis does not provide experimental proof of transporter specificity, for three reasons. Firstly, transcript levels might not correlate well with protein levels (Liu et al. 2016). In this case, combining transcriptomics and proteomics strengthens the hypotheses regarding glycoside transporter function (Pichler et al. 2020; McNulty et al. 2013). Secondly, the glycoside detected by the transcription regulator differs from the glycoside translocated through the cell membrane (Schwalm and Groisman 2017). Finally, functional redundancy is frequent in bacteria, in which genes encoding several transporters can be upregulated and even overproduced during utilization of a complex glycoside (Kabisch et al. 2014). In any case, the involvement of a transporter in the utilization of a specific glycoside can be proven only after deletion or inactivation of the target gene (Andersen et al. 2012; Webb et al. 2008; Chen et al. 2015), to abolish or reduce the growth.

As described hereafter for each glycoside transporter family, growth assays, often coupled with transcriptomic analyses, are mostly performed with native strains. However, in order to avoid functional redundancy and facilitate the genetic manipulations for inactivating or deleting genes, transporter function can also be characterized in recombinant strains.

For example, heterologous expression of the SusC/D gene pair from the inulin-using strain *Bacteroides thetaiotaomicron* 8736 in the non-inulin-using strain *B. thetaiotaomicron* VPI-5482 resulted in inulin use by the recipient strain, proving that the SusC/D from *Bacteroides thetaiotaomicron* 8736 is involved in the utilization of inulin (Joglekar et al. 2018). Furthermore, a locus encoding an ABC transporter and CAZymes from *Streptococcus mutans* LT11 gave *E. coli* the ability to ferment melibiose (Tao et al. 1993). Most recently, we extended this approach to the study of transporters from uncultured bacteria (Tauzin, Laville, et al. 2016; Wang et al. 2020) in *Escherichia coli*. By cloning various truncated forms of entire metagenomic loci encoding several transporters and at least one glycosidase in a fosmid, Tauzin *et al.*, and further Wang *et al.*, demonstrated the specificity of each of these transporters towards oligosaccharides of various structure and length. Furthermore, Dagkesamanskaya *et al.* and further, Lajus *et al.*, developed workflows to perform recombinant *E. coli* growth assays in pico-liter droplets using millifluidics, in order to decrease by several orders of magnitude the amount of substrate required to screen for transporter functionality and specificity (Dagkesamanskaya et al. 2018; Potocki-Veronese et al. 2023).

Whether in native or recombinant strains, the ability of a bacterium to grow on glycosides and its efficiency in doing so are related to the levels of production and to the combined efficiency of several proteins for sensing, binding, transporting, breaking down and metabolizing the substrate. By analyzing the growth phenotype only, it is thus impossible to quantify the efficiency of a transporter and thus, to obtain a precise profile of its specificity towards glycosides of defined structures. This requires other methods, which are described hereafter.

(2) Determination of binding specificity

For transporters involving binding proteins, known as substrate-binding proteins (SBPs) and surface glycan-binding proteins (SGBPs) in ABC and SusC/D transporters respectively, it is easier to investigate the characteristics of the binding protein than those of the transmembrane protein. Binding proteins are not transmembrane proteins, which facilitates their recombinant production and analysis. Once the recombinant binding proteins have been successfully purified, it is easy to screen for their binding specificity towards substrates using affinity gel electrophoresis (AGE) (Mystkowska et al. 2018) or differential scanning fluorimetry (DSF) (Boucher and Noll 2011). The specificity can then be quantified with isothermal titration calorimetry (ITC) (Leth et al. 2018), surface plasmon resonance (SPR) (Theilmann et al. 2019; Pichler et al. 2020), fluorescence spectroscopy (Phansopa et al. 2014) or saturation transfer

difference nuclear magnetic resonance (STD-NMR) (Bell et al. 2019). As the principles of these techniques are different, each has its own strengths and weaknesses. Affinity gel electrophoresis is a simple, rapid method for the analysis of protein-binding affinities to soluble polysaccharides, based on the reduced mobility of the protein interacting with substrates (Tomme et al. 2000). However, it cannot detect binding with oligosaccharides which are often the target substrates of the transporter. The ITC, DSF, SPR, STD-NMR and spectroscopic approaches all rely on the detection of a change in a physicochemical property of the ligand or the protein because of the binding (Vuignier et al. 2010). While ITC, DSF, STD-NMR and spectroscopy necessitate protein in solution, SPR experiments imply surface immobilization of biomolecules. Thus, binding of the protein to the sensor chip surface may involve loss of structural integrity and function, negatively impacting the interactions between the protein and its ligand (Nguyen et al. 2015). DSF rarely provides a dissociation constant at physiological temperatures (Grøftehaug et al. 2015). STD-NMR can detect a reduction in the free ligand signal from ligand exchange between free and bound state when ligand-protein binding is present (Viegas et al. 2011). Its use for characterizing glycoside binding proteins is rare (Bell et al. 2019). ITC is one of the most common methods for examining the binding specificity of glycoside binding proteins, since it allows direct determination of the association constant by measuring the heat exchange during complex formation at a constant temperature (Bjelić and Jelesarov 2008). However, a significant issue with ITC is the large amount of soluble protein required for accurate measurements, especially for certain proteins that may be hard to prepare in large amounts. SPR, which is also widely used for the investigation of binding kinetics and strengths, consists of measuring the change in the refractive index near a sensor surface. It only requires a few micrograms of protein. The equipment and maintenance costs of these techniques might be discouraging (Jecklin et al. 2009). The steady-state tryptophan spectroscopic methods allow for measurement of binding affinity constants using relatively low amounts of protein, and less expensive equipment compared to SPR and ITC (Niesen et al. 2007). However, the spectroscopic methods are mainly for high-affinity binding sites, in the nanomolar range (Vuignier et al. 2010). Finally, all these techniques require purified proteins, and thus, optimization of recombinant protein production. It is noteworthy that the binding specificity of these transporters does not lead to direct determination of uptake specificity. For example, MalE, the SBP of the maltose ABC from *E. coli* K-12 binds α - and β - cyclodextrins, while these substrates cannot be taken up by the maltose ABC transporter (Quijcho et al. 1997; Mächtel et al. 2019).

(3) Determination of uptake specificity

The most generic method for determining transport specificity for all transporter families (including those without binding proteins) and without any bias due, for example, to CAZyme activity, is based on analyzing the accumulation of radiolabeled compounds in cells via uptake assays (Martin and Russell 1987). This quantitative method is also highly sensitive. It can indeed detect transport events in the micromolar range in less than 10 minutes. However, it requires a specific secure environment to manipulate radiolabeled compounds. In addition, this method is mainly used for monitoring monosaccharide uptake (Blanvillain et al. 2007), since most glycosides are not available in radioactive form or are prohibitively expensive (for example, USD 12,000 per gram for D-Lactose-1-¹³C monohydrate), which precludes large-scale screening campaigns. In order to circumvent this problem of radiolabeled substrate availability, a highly generic method was recently developed to quantify the uptake rates of transporters (Wang et al. 2020). This novel and simple approach is based on the use of crude cell debris and intact cells (native or recombinant), and of nonlabeled substrates of which very weak concentration variations (less than one nanomole) are monitored by high performance anion exchange chromatography coupled to pulsed amperometric detection (HPAEC-PAD).

Besides, as explained previously, in native strains, the multiplicity of carbohydrate transporters can bias the results of screening for specificity. To solve this problem, the targeted transporter can be produced in recombinant strains lacking the function of interest (Wei et al. 2012), or in membrane vesicles or proteoliposomes (Cao et al. 2011; McCoy et al. 2016). Nevertheless, these approaches can easily fail when the transport proteins are not active due to targeting to the membrane or to insertion or folding issues during the purification or reconstitution process (Majd et al. 2018).

Less generic methods also exist, focusing on particular transport mechanisms involving energy-coupling or substrate modification. Phosphoenolpyruvate (PEP-dependent) glycoside phosphorylation, which is a specific feature of PTS transporters, can be tracked by analyzing the amount of pyruvate in cell-free systems, and to deduce the amount of phosphorylated glycosides (Martin and Russell 1987). The amount of pyruvate can be monitored by quantifying the product of NADH oxidation by the lactate. For ABC transporters, glycoside uptake requires the energy of ATP binding and hydrolysis. Thus, the specificity of ABC transporters can be determined by testing growth on the target substrate when ATP synthesis is inhibited with a specific inhibitor such as iodoacetate (Kaplan and Hutkins 2003).

III. TonB-dependent transporters and associated surface glycan-binding proteins

(1) The SUS paradigm

In the late 1980s, the Salyers' group identified and characterized a locus involved in the utilization of starch in *Bacteroides thetaiotaomicron* 5482 (Tancula et al. 1992) and later named the starch utilization system (SUS) (D'Elia and Salyers 1996). In this strain, the SUS locus, named polysaccharide utilization locus (PUL) (Bjursell et al. 2006)) is composed of eight genes named *susRABCDEFG* (Figure 1). First, starch recognition and binding at the cell surface are mediated by outer membrane proteins consisting of SusD to F. Then, starch is initially degraded into maltooligosaccharides by the surface lipoprotein SusG corresponding to a GH13 α -amylase (Shipman et al. 1999). The oligosaccharides are then internalized via the TonB-dependent transporter (TBDT, SusC) (Reeves et al. 1996) before being subsequently degraded by periplasmic GHs (SusA and SusB) (Smith and Salyers 1991; D'Elia and Salyers 1996). Finally, SusR, which is an inner membrane-spanning sensor/regulator protein, recognizes periplasmic maltose and triggers the upregulation of the SUS genes (D'Elia and Salyers 1996). Generally speaking, PULs are mostly regulated by three kinds of transcription factors including the SusR-like regulators, the AraC-like hybrid two-component systems (HTCSs), and the extracytoplasmic function sigma (ECF- σ) factor/anti-sigma factors (Ravcheev et al. 2013). Both SusC and SusD have been identified as the main contributors to substrate uptake, while SusE and SusF, which are not present in all PULs, have auxiliary roles in substrate recognition and binding. The SUS-paradigm and the PUL concept have been extended in the last decades to other polysaccharide utilization pathways in Bacteroidota (formerly Bacteroidetes). Tandem *susD*-like and *susC*-like genes are considered as a hallmark of PULs (Bjursell et al. 2006), and the presence of these *susCD*-like gene pairs is used to identify PULs in Bacteroidota genomes (Terrapon et al. 2015). Nevertheless, the characterization of a non-canonical PUL and genes distal to the PUL, including a *susCD*-like pair (Ficko-Blean et al. 2017), strengthens the recent proposition that the PUL definition should not be restricted to the presence of a *susCD*-like pair (Grondin et al. 2017).

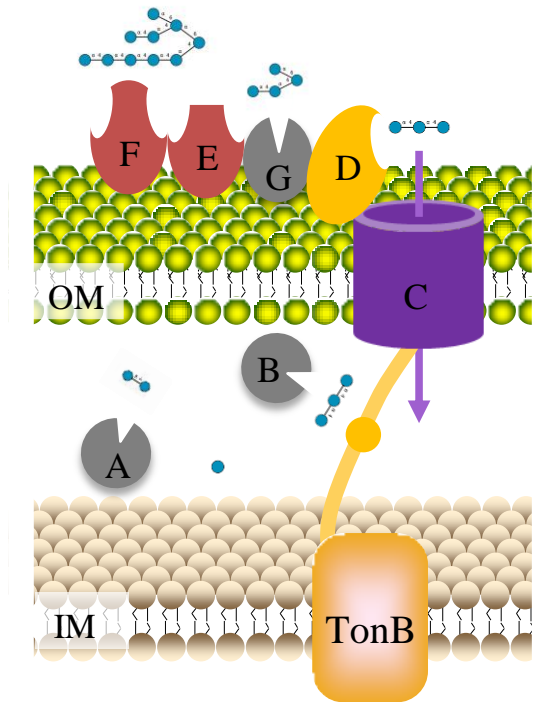


Figure 1. Schematic representation of the starch utilization system in *Bacteroides thetaiotaomicron* 5482.

The starch is bound to the surface of the cell by the starch-binding outer membrane lipoproteins SusDEF. Subsequent hydrolysis by a similarly membrane-tethered α -amylase, SusG, generates oligosaccharides small enough to transit through the TonB-dependent transporter. Once in the periplasm, SusA and SusB, a neopullulanase and α -glucosidase, respectively, process oligosaccharides into glucose. The monosaccharide is then shuttled into the cytoplasm by an unknown transporter. OM and IM represent the outer membrane and inner membrane, respectively.

(2) Functional diversity of SUS-like systems

In the past decade, 45 *Bacteroides* SUS-like systems have been deeply characterized (Table 2), targeting a large diversity of polysaccharides including xyloglucan (Foley et al. 2019; Tauzin, Kwiatkowski et al. 2016; Larsbrink et al. 2014), xylans (Rogowski et al. 2015), mannans (Cuskin et al. 2015), β -glucans (Tamura et al. 2021; Déjean et al. 2020; Tamura et al. 2019; 2017; Temple et al. 2017), chitin (Larsbrink et al. 2016), galactomannan (Bågenholm et al. 2019; 2017), arabinogalactans (Cartmell et al. 2018), pectins (Luis et al. 2018; Ndeh et al. 2017), levan (Sonnenburg et al. 2010), porphyran (Hehemann et al. 2012), carrageenan (Ficko-Blean

et al. 2017), glycoaminoglycans (GAGs) (Ndeh et al. 2020; Cartmell et al. 2017) and mucins (Luis et al. 2021; Crouch et al. 2020). Some cellulolytic activities have been highlighted in PULs of rumen uncultured Bacteroidota (Mackenzie et al. 2015; Pope et al. 2012; Naas et al. 2014), and the SusD-like proteins of these PULs were shown to bind strongly to cellulose (Naas et al. 2014; Mackenzie et al. 2012). Since growth studies and transcriptomic analysis of these metagenomic species have been impossible to perform to date, there is no proof that the native strains containing these PULs are able to utilize cellulose.

The complexity of PUL organization reflects the structural complexity of their cognate substrate, with the presence of genes encoding enzymes that are not CAZymes, such as sulfatases and peptidases, epimerase and members of the ROK (Repressors, ORFs and Kinases) family, as well as additional transporters such as ATP-binding Cassette and Major Facilitator Superfamily transporters (Terrapon et al. 2018). For example, the utilization of the yeast α -mannan by *Bacteroides ovatus* is mediated by three PULs working in synergy and encoding more than a dozen glycoside hydrolases, and a few phosphatases and glycosyltransferases (Cuskin et al. 2015). By contrast, the xyloglucan PUL from *B. ovatus* consists of ‘only’ 12 genes, including eight GHs and one unique SusC/D pair (Larsbrink et al. 2014). However, characterization of syntenic xyloglucan PULs from other Bacteroidota strains revealed slight differences in gene organization and enzyme specificity, conferring the ability to access distinct plant-specific xyloglucan side chains (Déjean et al. 2019). Up until now, the main focus of biochemical characterization of PULs has been GHs and SGBPs. The characterization of the transmembrane component of SUS systems, the TBDT, is more difficult.

(3) Structure and mechanism of TonB-dependent transporters and associated surface glycan-binding proteins

SusC is a TBDT structurally characterized by an N-terminal plug domain inserted into a 22-stranded β -barrel. TBDTs are active transport systems mediating substrate-specific transport across the outer membrane. Binding of the substrate induces an allosteric rearrangement of the plug domain containing a conserved binding motif (TonB box) that is then released into the periplasmic space. There, it interacts with the C-terminal domain of TonB, while the N-terminal of TonB forms a complex with the inner membrane proteins ExbB and ExbD (TonB complex). The active transport of nutrients into the periplasm by TBDTs requires the energy derived from the proton motive force transmitted from the TonB complex (Larsen et al. 1999).

SusD-like proteins are cell surface glycan-binding proteins (SGBPs) involved in the capture of extracellular glycans. Interestingly, SusD-like proteins share no sequence similarity with known carbohydrate-binding modules (CBMs) and are notably around twice the size (around 65 kDa) of CBMs (Foley et al. 2016). In 2008, the first crystal structure was solved for a SusD from *Bacteroides thetaiotaomicron* VPI-5482 involved in starch utilization (Koropatkin et al. 2008). Thus far, the overall structure of SusD-like proteins consists of a region of α helices organized into four tetratricopeptide repeat (TPR) units and a more variable region including many loops and short α helices. The variable region is unique for each SusD-like protein. It depends on their cognate substrate and includes an optimized platform for specific substrate binding, mostly provided by at least two tryptophan residues. SusD-like proteins are attached to the outer membrane via a lipidated N-terminal cysteine projecting the protein above the surface of the membrane to facilitate glycan capture (Koropatkin et al. 2008; Cameron et al. 2014). In rare instances, some SusD-like proteins do not bind the cognate substrate of the transporter efficiently. In such cases, substrate binding is ensured by additional SGBPs (Déjean et al. 2020; Luis et al. 2018). To our knowledge, there is only one example of a nonbinding SusD-like protein present in a PUL that does not encode for additional SGBPs (Tauzin et al. 2022, see the 1st Results Chapter). Noteworthy, this PUL is involved in xylooligosaccharide (XOS) utilization by an uncultured human gut *Bacteroides* strain and this work points out that the term “surface glycan-binding proteins” generally used for SusD-like proteins is thus not generic.

In SUS, SusC physically interacts with SusD to form a complex that represents a minimal system when complemented with SusG for starch utilization by *B. thetaiotaomicron* VPI-5482 (Cho and Salyers 2001). However, by using non-binding SusD-like mutants, substrate binding to SusD-like proteins was proven not to be a prerequisite for growth, since the expression of other SGBPs (SusE and SusF-like proteins from the same PUL) tends to supplement this loss (Cameron et al. 2014; Tauzin, Kwiatkowski, et al. 2016). Deletion of the SusD-like proteins impairs growth on the cognate substrate, implying that the proteins play an essential structural role in SusC/D pair formation in some systems where binding ability is also not a prerequisite for substrate transport (Koropatkin et al. 2008; Tauzin, Laville, et al. 2016). By contrast, it was demonstrated recently that both substrate binding and the physical presence of the SusD-like protein are crucial for efficient utilization of mixed linkage glucans by *B. ovatus* (Tamura et al. 2019). To date, only three 3D structures of three different SusC/D complexes have been solved. In Glenwright *et al.*, the authors used X-ray crystallography, molecular dynamics, and

electrophysiology studies to determine the structure of two pairs of SusC/D from *B. thetaiotaomicron* VPI-5482 (Glenwright et al. 2017). One pair (BT2263-64) is found in a four-component complex of unknown function, while the second one (BT1762-63) is part of a levan PUL (Sonnenburg et al. 2010). Glenwright *et al.* suggested a pedal bin mechanism for nutrient acquisition by SusC/D-like systems, with SusD forming the lid and SusC being the bin. In the absence of the ligand, SusD moves away from SusC in a hinge-like fashion in order to expose the substrate-binding site to the extracellular environment. After substrate binding, the SusD 'lid' swings into a closed state, resulting in a stabilized complex due to ligand interaction with both SusC and SusD. The substrate then moves into the SusC 'bin', preceding the binding of TonB to the TonB box and subsequent conformational changes for the release of the substrate into the cell. The structure of SusD is not altered in the closed state and the interaction with SusC includes the highly variable region that forms the substrate-binding site of the SusD-like proteins. According to the authors, this pedal bin mechanism is consistent with the envelopment of the whole ligand molecule that precludes the import of undigested glycans with high molecular weights. Interestingly, SusC/D is dimeric, with the dimer interface formed by SusC. Cryo-EM studies on the same complex confirmed the dimeric nature of the transporter and revealed three distinct states: open-open, open-closed, and closed-closed, based on the position of the SusD lid (Gray et al. 2021). Very recently, cryo-EM experiments enabled to solve the 3D structure of the four components (SusC, SusD, SGBP and enzyme) of the levan (BT1760-63) and dextran (BT3087-90) PULs. The SusC/Ds are still dimeric, producing octameric complexes where the positioning of the subunits is consistent with pedal bin transport (White et al. 2023). In the case of the levan PUL, the authors identified two binding sites within the SusC barrel, one at the interface of SusC and SusD and which binds a hexasaccharide, and the second one at the bottom of the SusC cavity where a tetrasaccharide is also in contact with the plug. Interestingly, at higher threshold level, diffuse density could be detected between the two binding sites, indicative of multiple conformations and the absence of interaction with SusC in this segment. These suggest that the SusC/D core is promiscuous and can accommodate relatively long oligosaccharides (probably up to DP 15). It is noteworthy that the SusC/D pair is not often linked with glycan utilization as some pairs are isolated on the genome or surrounded by proteins of unknown function (Bolam and van den Berg 2018; Lapébie et al. 2019). However, as discussed above, an isolated SusC/D pair could also be working in synergy with other PULs in order to achieve the complete degradation of a complex polysaccharide such as pectin by *B. thetaiotaomicron*, where one of the three PULs includes only a SusC/D pair (RG-II PUL2) (Ndeh et al. 2017).

(4) Taxonomical diversity of TonB-dependent transporters

The presence of PULs is a unique feature of members of the Gram-negative phylum Bacteroidota. The SusC/D pair was used as a hallmark for PUL identification and led to the establishment of the Polysaccharide-Utilization Loci DataBase (PULDB, <http://www.cazy.org/PULDB/>) (Terrapon et al. 2018; 2015). This database lists the experimentally characterized PULs (the ‘literature-derived PULs’), based on transcriptomic studies of Bacteroidota strains cultured on polysaccharides and proteoglycans, and also 68,500 predicted PULs mainly from 2,065 Bacteroidota species (January 2020 release). However, TonB-dependent transporters specific for glycosides are also present in Pseudomonadota, such as in the plant pathogen *Xanthomonas campestris* pv. *campestris* (Blanvillain et al. 2007; Déjean et al. 2013).

Interestingly, depending on the strain and the substrate of interest, different mechanisms have been evolved by gut Bacteroidota, leading to two opposing behaviors, selfish or distributive (Grondin et al. 2017). Co-culturing experiments of *B. thetaiotaomicron* with other *Bacteroides* strains revealed that extracellular hydrolysis of the substrate will only benefit that species itself, while for *B. ovatus*, the products of partial breakdown will diffuse into the extracellular environment and benefit the other species (Cuskin et al. 2015; Rogowski et al. 2015; Rakoff-Nahoum et al. 2014). These PULs thus provide a significant advantage for environmental niche colonization and thus, may play a role in modulating microbiota composition and driving community dynamics, depending on the diverse mechanisms by which nutrients are used.

Finally, recently, the PUL concept has been extended to members of the other dominant phylum Bacillota (gpPULs) (Sheridan et al. 2016), ultimately leading to the general use of the term ‘PULs’ to denote bacterial loci encoding for at least one glycoside-degrading enzyme, one transport system and one transcriptional regulator.

Table 2. Summary of the SusD and SusC/D pairs involved in polysaccharide utilization by Bacteroidota, characterized either biochemically, or by gene deletion or mutation.

Unless they were also characterized by one of these techniques, this list does not include proteins whose genes have been shown to be upregulated during polysaccharide utilization by transcriptomic analysis (listed in the PULDB).

Protein name (locus tag)	Substrate	Organism	Characterization method	Structure	Reference
SusD proteins					
BT3701	Starch Maltooligosaccharide	<i>B. thetaiotaomicron</i>	ITC Gene deletion	3CKC apo 3CKB maltotriose 3CK9 maltoheptaose 3CK8 β -cyclodextrin 3CK7 α -cyclodextrin	(Koropatkin et al. 2008)
BT1043	N-acetyllactosamine	<i>B. thetaiotaomicron</i>	ITC	3EHM apo 3EHN acetyllactosamine	N- (Koropatkin et al. 2009)
Uncultured species	Cellulose	Uncultured reindeer assigned Bacteroidota	Svalbard rumen to AGE	-	(Mackenzie et al. 2012)
BACOVA_02651 SGBP-A	Xyloglucan	<i>B. ovatus</i>	AGE ITC Gene deletion	5E75 apo 5E76 xylo- glucooligosaccharide	(Larsbrink et al. 2014; Tauzin, Kwiatkowski, et

			Point mutation		al. 2016; Foley et al. 2019)
BT3984	Mannan	<i>B. thetaiotaomicron</i>	ITC Gene deletion	-	(Cuskin et al. 2015)
BACOVA_04392	Birchwood glucuronoxylan Wheat arabinoxylan				
BACOVA_03427	Birchwood glucuronoxylan Corn glucuronoarabinoxylan	<i>B. ovatus</i>	ITC	-	(Rogowski et al. 2015)
Fjoh_4558		<i>Flavobacterium</i>	SDS-PAGE		
Fjoh_4561	Chitin	<i>johnsoniae</i> ATCC 17061	ITC Gene deletion	5J90 apo 5J5U apo	(Larsbrink et al. 2016)
F5_SusD	Xylooligosaccharides	Uncultured <i>Bacteroides</i>	Gene deletion Growth assay HPAEC-PAD SDS-PAGE AGE ITC DSC	7NEK	(Tauzin et al. 2022; Tauzin, Laville, et al. 2016)
BACOVA_02095	Galactomannan Glucomannan	<i>B. ovatus</i>	AGE Gene deletion	-	(Bågenholm et al. 2017)
BT3311	1,6-β-glucan	<i>B. thetaiotaomicron</i>	AGE	-	(Temple et al. 2017)

ITC					
BT4659	Heparin/heparan sulfate	<i>Bacteroides thetaiotaomicron</i>	ITC	3IHV	(Cartmell et al. 2017)
SAMN04488552_1347	Laminarin	<i>Gramella</i> sp. MAR_2010_102	AGE Point mutation	6GCZ apo	(Mystkowska et al. 2018)
BT4113	Homogalacturonan	<i>Bacteroides thetaiotaomicron</i> ,	Gene deletion ITC	-	(Luis et al. 2018)
BT4670	Pectic galactan				
BoManPUL SusD-like	Galactomannan, glucomannan	<i>Bacteroides ovatus</i>	MST	-	(Bågenholm et al. 2019)
BACOVA_02743	$\beta(1,3)/\beta(1,4)$ -glucans	<i>B. ovatus</i>	AGE ITC Gene deletion Point mutation	6E60 apo 6DMF cellobiose 6E61 heptasaccharide	(Tamura et al. 2019)
BT3331	Glycosaminoglycan chondroitin sulfate	<i>B. thetaiotaomicron</i>	ITC		(Ndeh et al. 2020)
BuSGBP-A BACUNI_01488		<i>Bacteroides uniformis</i>		7KV1	
BtSGBP-A	$\beta(1,3)$ -glucan	<i>Bacteroides thetaiotaomicron</i>	AGE ITC Point mutation	7KV2 7KV3 7KV4	(Tamura et al. 2021; Déjean et al. 2020)
BfSGBP-A		<i>B. fluxus</i>			

Bs04079	Fungal mannan	<i>B. salyersiae</i>	AGE ITC	-	(Buzun et al. 2022)
BuXyGUL2 SGBP-A	Xyloglucan oligosaccharides (XyGOs)	<i>Bacteroides ovatus</i> ATCC 8483	ITC	-	(Grondin et al. 2022)
FCWusD (Cpin_2192)	β -1,3-glucans	<i>Chitinophaga pinensis</i>	Pull-down assays		(Lu et al. 2023)
BX_29260 ^{SusD} , BX_29280 ^{SusD}	Xylo-oligosaccharides (XOS)	<i>Degradation of xylan by human gut Bacteroides xylanisolvens XB1A</i>	Comparative genomics analysis Western blotting		(Zhao et al. 2023)
AqSusD	Chitin	Bacteroidota strain <i>Aquimarina sp. SCSIO</i> 21287	Pull-down assay ITC Crystallization	8J0P 8JXZ	(Yang et al. 2023)
Looped SusD	Glycogen, pullulan	<i>Polaribactersp.</i> <i>Hel_I_88</i>	Sequence alignment Metatranscriptomics	-	(Beidler et al. 2023)
Open SusD	Glycogen	<i>Muricaudasp.</i> <i>MAR_2010_75</i>	Comparative genomics		
SusD_I_1 (P162_RS02310)	Beechwood xylan (BX), Palmaria palmata xylan (PPX), Caulerpa prolifera xylan (CPX), rye arabinoxylan (RAX), wheat arabinoxylan of medium	<i>Flavimarina</i> <i>sp.</i> <i>Hel_I_48</i>	Comparative genomics Proteome analysis DNS assay AGE AlphaFold prediction	-	(Dutschei et al. 2023)

	viscosity (WAX-M), insoluble wheat arabinoxylan (WAX-I), apple pectin (Pec)			
SusD_I_2 (P162_RS02355)	BX			
SusD_I_3 (P162_RS02370)	BX			
SusD_I_4	BX			
SusD_II_1 (P162_RS04065)	RAX, WRAX-M			
SusD_II_2 (P162_RS04075)	RAX, WRAX-M			
BT3089	Dextran	<i>Bacteroides thetaiotaomicron</i>	Gene deletion Binding tracking by fluorescent polymers - Cryogenic electron microscopy	(Wong et al. 2024; White et al. 2022)
FjDusD (Fjoh_4433)	Isomaltooligosaccharides	Bacteroidota <i>Flavobacterium johnsoniae</i>	AGE ITC - AlphaFold prediction	(Nakamura et al. 2023)
SusC/D complexes				

BT2261-2264	No	<i>B. thetaiotaomicron</i>	ITC Gene deletion	5FQ6	(Glenwright et al. 2017; Sonnenburg et al. 2010)
BT1762-1763	Levan	<i>B. thetaiotaomicron</i>	ITC Gene deletion	5T3R	
BT0363-0364	Arabinan				
BT4113-4114	Homogalacturonan	<i>Bacteroides</i>	Gene deletion	-	(Luis et al. 2018)
BT4164-4165	Rhamnogalacturonan-I	<i>thetaitotaomicron</i>			
BT4168-4169					
SusC ₂ D ₂ core (levan utilisome)	Levan fructo- oligosaccharides (FOS) DP 8-12	<i>Bacteroides</i>	Cryogenic electron microscopy	8AA1	(White et al. 2023; Silale and van den Berg 2023)
SusC ₂ D ₂ core (inactive levan utilisome)	FOS DP 15-25	<i>thetaitotaomicron</i>		8AA3	

IV. ABC transporters of glycosides

ATP-binding cassettes (ABCs) form one of the largest super-families of transmembrane-spanning proteins. ABCs are present in many eukaryotes and prokaryotes, and even in some large viruses (Ford and Beis 2019). These permeases are, in general, primary active transporters which may act either as importers or as exporters, depending on the direction of translocation of the cargo molecules (Locher 2016). They are able to transport various substrates in response to ATP hydrolysis (Mächtel et al. 2019). All ABC transporters have a core architecture comprising two transmembrane domains (TMDs) and two cytosolic nucleotide-binding domains (NBDs) where ATP binds and is hydrolyzed (Beis 2015). In prokaryotes, these four domains can be present in a combination of one or two multi-domain polypeptides (Wilkens 2015; Theodoulou and Kerr 2015). Glycoside importers also require the cooperation of an additional cognate partner: a substrate-binding protein, also named substrate-specific binding component (both abbreviated SBP) (Wilkens 2015; Theodoulou and Kerr 2015). However, there are exceptions in a subgroup of ABC transporters lacking SBPs, involved in the uptake of vitamins and micronutrients and referred to as energy-coupling factor (ECF) transporters (Erkens et al. 2012). Recently, a sucrose ABC importer lacking an SBP has been described (Wang et al., 2020), but its mechanism remains to be determined. A recent publication gathers numerous authors who speak in favor of a new ABC transporter classification that is based on structural homology in the TMDs (Thomas et al. 2020).

(1) Diversity of sequences and structures

Among the 34 families of ABC importers currently listed in the TCDB, carbohydrate-specific transporters are clustered within two families: carbohydrate uptake transporter-1 (CUT1 annotated with TC# 3.A.1.1) and carbohydrate uptake transporter-2 (CUT2 annotated with TC# 3.A.1.2) (Saier 2000). Members of the CUT1 family are specific to a variety of oligosaccharides with a degree of polymerization ranging from 2 to 7, as well as polyols, glycerol and glycerol-phosphate. Members of the CUT2 family only transport monosaccharides (Schneider 2001). This does not rule out the existence of exceptions, however, such as a D-xylose and L-arabinose ABC transporter in *Haloferax volcanii* H26 belonging to the CUT1 family (Johnsen et al. 2019). In addition, members of the peptide/opine/nickel uptake transporter (PepT) family (3.A.1.5) were also found to be involved in glycoside transport (Andersen et al. 2012; Connors et al. 2005; Nanavati et al. 2006). It can therefore be expected that the number of families containing glycoside transporters will increase as new transport specificities are discovered.

All the characterized glycoside transporters are SBP-dependent. Generally speaking, CUT1 ABC transporters harness two separate transmembrane proteins interacting with a protein involving two copies of NBDs (Schneider 2001). For glycoside ABC transporters of the PepT family, either the two TMDs or the two NBDs are in separate proteins.

The selectivity and specificity of carbohydrate ABC transporters are largely determined by the SBPs that are the most diverse components of the ABC systems, whereas the NBDs, and to a lesser extent, the TMDs, are clearly more similar. Indeed, SBPs have relatively low sequence similarity (less than 20 %), and phylogenetic analyses based on multiple sequence alignments did not yield a stable tree (Berntsson et al. 2010). Traditionally, SBPs have been categorized into different groups on the basis of their substrate specificity, in which glycoside ABC transporters were clustered into the CUT1 families (Tam and Saier 1993). In 2010, SBPs of ABC importers were reclassified into six clusters based on their available three-dimensional structures. SBPs binding to glycosides are spread across three different clusters (B, C and D) (Berntsson et al. 2010). Recently, a seventh cluster was suggested to extend the structural classification, as the structure of the fructooligosaccharide-binding protein of ABCs shows additional structural elements that define a larger ligand-binding cavity with respect to the canonical SBPs, and an EF hand-like calcium-binding site (Culurgioni et al. 2017).

(2) Functional diversity

Since the term ‘ABC transporters’ was defined in 1990, a huge number of proteins have been annotated as ABC transporters, as described in the following section (Holland 2019). Meanwhile, only a small set of ABC transporters have been experimentally proven to be involved in glycoside utilization by bacteria, by using one or more of the methods described in section 3.1. of this chapter (Table 3). In some cases, only the function and structure of SBPs were characterized, because these proteins are highly soluble and thus easier to characterize than TMDs. However, substrate specificity determined by binding assays does not always correspond to actual transport specificity. The MalE SBP of the maltose ABC from *E. coli* K-12, binds α - and β -cyclodextrins, while these substrates cannot be taken up by the maltose ABC transporter (Quioco et al. 1997). Moreover, in a few cases only, the ABC transporters – such as those targeting maltooligosaccharides and alginate – were deeply characterized by combining biochemical, molecular and structural approaches (Mächtel et al. 2019; Maruyama et al. 2019).

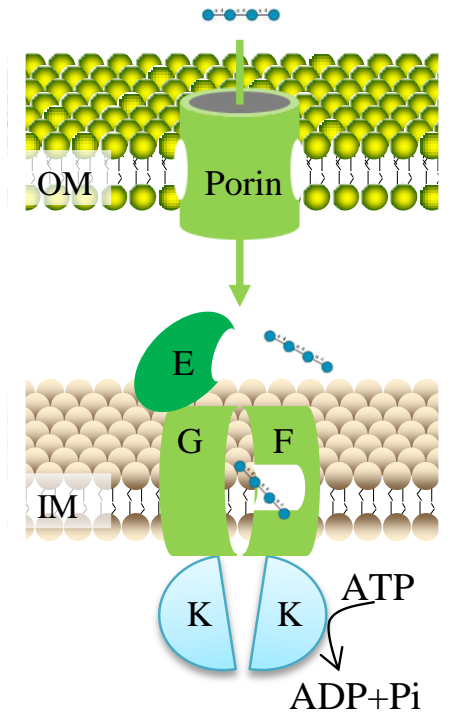


Figure 2. Transport of maltooligosaccharides by ABC in Gram-negative bacteria.

Maltooligosaccharides diffuse across the outer membrane through the porin, bind with the MalE protein in the periplasm, and are transported across the inner membrane by the ABC transporter MalFGK2. OM and IM represent the outer membrane and inner membrane, respectively.

Table 3. List of the bacterial ABC transporters involved in glycoside utilization, characterized either by binding or transport assays, gene deletion and/or transcriptional approaches.

The glycoside ABC transporters retrieved from the literature but which are not listed in the TCDB were classified based on their conserved domains by using a BLAST analysis against the NCBI's Conserved Domain Database (Marchler-Bauer et al. 2017).)

TCDB Family name	TCID	Protein name (locus tag of SBP)	Substrate	Organism	Characterization method	Reference
CUT1	3.A.1.1.1	MalEFGK (b4034)	Maltooligosaccharides (DP2-7)	<i>Escherichia coli</i> K-12	Gene deletion Binding test Transport assay Crystal structure	(Wandersman et al. 1979; Oldham et al. 2007; Khare et al. 2009; Oldham and Chen 2011; Oldham et al. 2013)
	3.A.1.1.2	YufKLM (BSU34160)	Galactooligosaccharides (DP2-4)	<i>Bacillus subtilis</i> KM0	Binding test	(Watzlawick et al. 2016)
	3.A.1.1.9	GuoEFGK	Xylooligosaccharides (DP2-6)	<i>Geobacillus stearothermophilus</i> T-6	Binding test	(Shulami et al. 1999; 2007)

	3.A.1.1.10	AlgSM1M2Q1Q2	Alginate (DP 2-7)	<i>Sphingomonas</i> sp. strain A1	Transport assay Gene deletion Crystal structure	(Momma et al. 2000; Mishima et al. 2003; Maruyama et al. 2015; Kaneko et al. 2017; Murata et al. 2022)
	3.A.1.1.11	TogABMN	Oligogalacturonides (DP2-3)	<i>Erwinia chrysanthemi</i> 3937	Gene deletion Heterologous expression in <i>E. coli</i>	(Hugouvieux-Cotte-Pattat et al. 2001; Hugouvieux-Cotte-Pattat and Reverchon 2001)
	3.A.1.1.17	ThuEFGK (CN159_00480)	Trehalose, maltose, sucrose	<i>Sinorhizobium meliloti</i> Rm1201	Gene deletion Transport assay	(Jensen et al. 2002)
	3.A.1.1.18	NacEFG	N-Acetylglucosamine, chitobiose	<i>Streptomyces olivaceoviridis</i>	Transport assay Binding test Gene deletion	(Xiao et al. 2002)
	3.A.1.1.19	PalEFGK	Platinose (Isomaltulose)	<i>Agrobacterium tumefaciens</i> MAFF301001	Gene deletion	(De Costa et al. 2003)

3.A.1.1.20	MsmEFGK (D7H66_05730)	Fructooligosaccharides (DP3-5)	<i>Lactobacillus acidophilus</i> NCFM	Transcriptional analysis Gene deletion	(Barrangou et al. 2003)
3.A.1.1.21	BxlEFGK (SCO7028)	Xylobiose, xylotriose	<i>Streptomyces thermoviolaceus</i> OPC-520	Transcriptional analysis Binding test	(Tsuji et al. 2004)
3.A.1.1.22	MalE1E2FGK (TM_1204)	Maltose, maltotriose, mannotetraose (MalE1)	<i>Thermotoga maritima</i> MSB8	Transcriptional analysis Binding test	(Nanavati et al. 2005)
3.A.1.1.22	MalE1E2FGK (TM_1839)	Maltose, maltotriose, trehalose (MalE2)	<i>Thermotoga maritima</i> MSB8	Transcriptional analysis Binding test	(Nanavati et al. 2005)
3.A.1.1.23	CebEFG	Cellobiose, cellotriose	<i>Streptomyces reticuli</i> DSM 40776	Transcriptional analysis Gene deletion Transport assay	(Schlösser et al. 1999)
3.A.1.1.25	- (TTC1627)	Trehalose, maltose, sucrose, palatinose	<i>Thermus thermophilus</i> HB27	Transport assay Binding test Crystal structure	(Silva et al. 2005; Chandravanshi et al. 2019)
3.A.1.1.26	MalEFG (BSU34610)	Maltodextrin	<i>Bacillus subtilis</i> 168	Transport assay Binding test	(Schönert et al. 2006)

					Gene deletion	
3.A.1.1.27	MalXFGK (SMU_1568)	Maltodextrin (DP2-7)	<i>Streptococcus mutans</i> ATCC 700610	Transport assay Gene deletion	(Webb et al. 2008)	
3.A.1.1.28	MsmEFGK (SMU_878)	Raffinose, stachyose	<i>Streptococcus mutans</i> ATCC 700610	Gene deletion	(Webb et al. 2008)	
3.A.1.1.31	LpqY-SugA-SugB-SugC	Trehalose	<i>Mycobacterium tuberculosis</i> H37Rv	Transport assay Gene deletion	(Kalscheuer et al. 2010; Sharma et al. 2022; Furze et al. 2021; De la Torre et al. 2021; Cassio Barreto de Oliveira and Balan 2020; Liu et al. 2020; Sabharwal et al. 2020)	
3.A.1.1.32	GgtABCD (slr0529)	Sucrose, trehalose	<i>Synechocystis</i> sp. strain PCC6803	Gene deletion Binding test	(Mikkat and Hagemann 2000)	

	3.A.1.1.33	DasABC	Chitobiose	<i>Streptomyces coelicolor</i> A3(2)	Transcriptional analysis Gene deletion	(A. Saito et al. 2008)
	3.A.1.1.34	AraNPQ	α -1,5-Arabinooligosaccharides (DP3-4)	<i>Bacillus subtilis</i> IQB496	Gene deletion	(Mário José Ferreira and De Sá-Nogueira 2010)
	3.A.1.1.41	TreEGF	Trehalose	<i>Thermotoga maritima</i> MSB8	Binding test	(Boucher and Noll 2011)
	3.A.1.1.44	MalEFG (SCO2231)	Maltose	<i>Streptomyces coelicolor</i> A3(2)	Gene deletion	(Van Wezel et al. 1997)
	3.A.1.1.45	MusEFGK (Cgl2461)	Maltose	<i>Corynebacterium glutamicum</i> ATCC 13032	Transcriptional analysis Gene deletion	(Henrich et al. 2013)
	3.A.1.1.48	MalEFGK (BLJ_1686)	Lacto-N-biose, galacto-N-biose	<i>Bifidobacterium longum</i> JCM1217	Binding test	(Suzuki et al. 2008)
	3.A.1.1.51	MalEFGK (Bd1227)	Maltooligosaccharides (DP2-7)	<i>Bdellovibrio bacteriovorus</i> ATCC 15356	Gene deletion Binding test Transport assay Crystal structure	(Licht et al. 2019)

	3.A.1.1.52	GlsQPCD (DSM107007_RS23065)	Sucrose, maltose, esculin	<i>Anabaena</i> sp. PCC 7120	Gene deletion	(Nieves-Mori3n and Flores 2018)
	3.A.1.1.53	RafEFGK (BIF_01592)	α -1,6-Glucosides and galactosides (2-7)	<i>Bifidobacterium animalis</i> subsp. <i>lactis</i> B1-04	Transcriptional analysis Binding test Crystal structure	(Morten Ejby et al. 2016)
	3.A.1.1.60	UspABC	Peptidoglycan precursors	<i>Mycobacterium tuberculosis</i> (<i>Mtb</i>) <i>M. smegmatis</i> mc2155	Gene deletion Operon analysis Topology mapping Proteomic analysis Phylogenetic analysis Crystal structure	(Karlikowska et al. 2021; De la Torre et al. 2021; Cassio Barreto de Oliveira and Balan 2020)
	3.A.1.1.61	Rv2038c-0Rv2041c	Sugar	<i>Mycobacterium tuberculosis</i> H37Rv	Phylogenetic analysis Crystal structure	(Cassio Barreto de Oliveira and Balan 2020; De la Torre et al. 2021)

	Literature	XynEFG	Xylooligosaccharides (DP2-6)	<i>Geobacillus stearothermophilus</i> T-6	Transcriptional analysis Binding test	(Shulami et al. 2007)
	Literature	CbpA	Cellotriose	<i>Clostridium thermocellum</i> DSM 1237	Binding test	(Nataf et al. 2009)
	Literature	CbpB	Cellooligosaccharides (DP2-5)	<i>Clostridium thermocellum</i> DSM 1237	Binding test	(Nataf et al. 2009)
	Literature	CbpC	Cellooligosaccharides (DP2-5)	<i>Clostridium thermocellum</i> DSM 1237	Binding test	(Nataf et al. 2009)
	Literature	CbpD	Cellooligosaccharides (DP3-5)	<i>Clostridium thermocellum</i> DSM 1237	Binding test	(Nataf et al. 2009)
	Literature	LBP	Laminaribiose	<i>Clostridium thermocellum</i> DSM 1237	Binding test	(Nataf et al. 2009)
	Literature	GacHFG	Acarbose, maltooligosaccharides (DP2-4)	<i>Streptomyces glaucescens</i> GLA.O	Binding test Crystal structure	(Vahedi-Faridi et al. 2010)

	Literature	AbnEFG	Arabinooligosaccharides (DP3-8)	<i>Geobacillus stearothermophilus</i> T-6	Binding test	(Shulami et al. 2011)
	Literature	XBP1	Xyooligosaccharides (DP2-4)	<i>Caldanaerobius polysaccharolyticus</i> ATCC BAA-17	Binding test Crystal structure	(Han et al. 2012)
	Literature	BlAXBP (Balac_0514)	Xylooligosaccharides (DP2-6), arabinoxylan-oligosaccharides (DP3-5)	<i>Bifidobacterium animalis</i> subsp. <i>lactis</i> Bl-04	Binding test Crystal structure	(Ejby et al. 2013)
	Literature	CpMnBP1	(Galacto-) Mannooligosaccharides (DP2-5)	<i>Caldanaerobius polysaccharolyticus</i> ATCC BAA-17	Binding test Crystal structure	(Chekan et al. 2014)
	Literature	- (EUR_01830)	Mannooligosaccharides (DP3-7), acarbose	<i>Eubacterium rectal</i> DSM 17629	Proteomic analysis Binding test	(Cockburn et al. 2015)
	Literature	- (EUR_31480)	Mannooligosaccharides (DP2-3)	<i>Eubacterium rectal</i> DSM 17629	Proteomic analysis Binding test	(Cockburn et al. 2015)
	Literature	- (EUR_21010)	Cellooligosaccharides (DP2, 3)	<i>Eubacterium rectal</i> DSM 17629	Proteomic analysis Binding test	(Cockburn et al. 2015)

	Literature	CsnEFG	Chitosan-derived oligosaccharides; Glucosamine oligosaccharides (GlcN) _n	<i>Streptomyces coelicolor</i> A3(2)	Gene deletion Transport assay	(Viens et al. 2015)
	Literature	xylIEFG ABC transporter	Xyloside, Arabino-xylooligosaccharides, xylooligosaccharides	<i>Corynebacterium glutamicum</i>	Gene deletion	(Watanabe et al. 2015)
	Literature	GanSPQ-MsmX	Galacto-oligosaccharides (DP3, DP4), galactotetraose, galactotriose Cyclodextrin	<i>Bacillus subtilis</i>	Surface plasmon resonance (SPR) spectroscopy Fluorescence and circular dichroism (CD) spectroscopy Gene deletion TSA	(Watzlawick et al. 2016; Kamionka and Dahl 2001)
	Literature	ManHIJ	(Galacto-) Mannooligosaccharides (DP3-5)	<i>Bacillus</i> sp. N16-5	Transcriptional analysis Binding test Gene deletion	(Song et al. 2016)
	Literature	FL-SBP	Fucosyllactose	<i>Bifidobacterium breve</i> BR-A29	Gene deletion	(Matsuki et al. 2016)

Literature	FusABC (SP_1796)	Fructooligosaccharides (DP3-5)	<i>Streptococcus pneumoniae</i> TIGR4	Transcriptional analysis Binding test Gene deletion Crystal structure	(Linke et al. 2013; Culurgioni et al. 2017)
Literature	SfuABC (SPV_1585)	Fructooligosaccharides (DP3-5)	<i>Streptococcus pneumoniae</i> D39	Binding test Gene deletion Crystal structure	(Culurgioni et al. 2017)
Literature	CpSBC4535	α -1,5- Arabinooligosaccharides (DP 3-6)	<i>Caldanaerobius polysaccharolyticus</i> ATCC BAA-17	Transcriptional analysis Binding test Gene deletion	(Wefers et al. 2017)
Literature	CuaABC (Ccel_2112)	Cellooligosaccharides (DP2-5)	<i>Ruminiclostridium cellulolyticum</i> H10 ATCC 35319	Binding test	(Fosses et al. 2017)
Literature	- (Smon_0123)	Sulfated GalNAc	<i>Streptobacillus moniliformis</i> DSM 12112	Binding test Crystal structure	(Oiki et al. 2017)
Literature	MdxEFG (HMPREF9496_01133)	Maltooligosaccharides (DP4-7)	<i>Enterococcus faecalis</i> JH2-2	Gene deletion	(Sauvageot et al. 2017)

	Literature	CycB-GanPQ (GanSPQ)	Galactan; Nonlinear arabinooligosaccharides; Galactooligosaccharides	<i>B. subtilis</i>	Transcriptional analysis Gene deletion	(Ferreira et al.2017)
		YtcQP-YteP	Polygalacturonan and rhamnogalacturonan			
		YesOPQ	Polygalacturonan and rhamnogalacturonan			
	Literature	PglK	Lipid-linked-oligosaccharide (LLO), GlcGalNAc5Bac-PP- undecaprenyl	<i>Campylobacter jejuni</i>	Crystal structure <i>In vitro</i> tLLO flipping assay <i>In vivo</i> LLO flipping assay Mass Spectrometry (MS)	(Perez et al. 2015; 2017; 2019)
	Literature	RiXBP (ROSINTL182_0819 9)	Xylooligosaccharides (DP3-6), arabinoxylotriose, arabinoxylotetraose	<i>Roseburia intestinalis</i> L1-82	Transcriptional analysis Binding test	(Leth et al. 2018)
	Literature	Wzm–Wzt	O-antigen polysaccharide	<i>Aquifex aeolicus</i>	Crystal structure	(Bi et al. 2018; Jones and George 2020)

	Literature	RiMnBP (ROSINTL182_05479)	Mannooligosaccharides (DP3-5)	<i>Roseburia intestinalis</i> L1-82	Binding test	(La Rosa et al. 2019)
	Literature	BiMnBP1 (BLAC_00780)	Mannooligosaccharides (DP3-6)	<i>Bifidobacterium animalis</i> subsp. <i>lactis</i> ATCC27673	Transcriptional analysis Binding test Crystal structure	(Ejby et al. 2019)
	Literature	BiMnBP2 (BLAC_00785)	Mannooligosaccharides (DP3-6)	<i>Bifidobacterium animalis</i> subsp. <i>lactis</i> ATCC27673	Transcriptional analysis Binding test Crystal structure	
	Literature	Bal6GBP (Balac_0483)	β -(1,3/6)-Galactosides (DP2, 4)	<i>Bifidobacterium animalis</i> subsp. <i>Lactis</i> BI-04	Binding test Crystal structure	(Theilmann et al. 2019)
	Literature	TTHA0356	Trehalose, maltose	<i>Thermus thermophilus</i> HB8	Crystal structure	(Chandravanshi et al. 2019)
		TTHA1652	Cyclodextrin, maltodextrin (4–5 pyranose units)			
		TTHB082	β -glucosides (laminaribiose, cellobiose disaccharides)			

	Literature	cphy2464-cphy2466	Cellobiose	<i>C. phytofermentans</i> <i>ISDg</i> (ATCC 700394)	Gene deletion Binding test	(Cerisy et al. 2019)
	Literature	MsmEX-AmyCD	Melibiose, raffinose and stachyose	<i>Bacillus subtilis</i>	Transcriptional analysis Gene deletion Binding test	(Heravi et al. 2019)
	Literature	GltABC	Lacto-N-biose I/galacto-N-biose, lacto-N-tetraose (LNT)	<i>B. longum</i> 105-A	Crystal structure Gene deletion Transcriptional analysis Binding test	(Sakanaka et al. 2019)
		FL1-BP (Blon_0341– 0343)	Fucosyllactose (FL) 2'-FL, 3-FL	<i>B. infantis</i> JCM 1222 ^T		
		FL2-BP (Blon_2202– 2204)	Fucosyllactose (FL) 2'-FL, 3-FL, lactodifucotetraose (LDFT) and lacto-N- fucopentaose (LNFP) I	<i>B. infantis</i> JCM 1222 ^T		
	Literature	GalCDE	Galactooligosaccharide (GOS)	<i>Bifidobacteria</i> : <i>B. breve</i> UCC2003 <i>B. infantis</i> strain DSM 20088 <i>B. breve</i> strain DSM 20091	Transport assay Transcriptional analysis	(Böger et al. 2019)

				<i>Bacteroides thetaiotaomicron</i>		
		GosDEC	Galactooligosaccharide (GOS)	<i>Bifidobacteria</i> : <i>B. lactis</i> Bl-04 <i>B. lactis</i> strains W51–W53 <i>B. breve</i> strain DSM 20091 <i>B. infantis</i> strain DSM 20088		
	Literature	BABP (BLLJ_0208)	β -L-arabinobiose	<i>Bifidobacterium longum</i>	Binding test Crystal structure	(Miyake et al. 2020)
	Literature	BIMsmE	Stachyose	<i>Bacillus licheniformis</i>	Binding test Crystal structure	(Zhang et al. 2020)
	Literature	CMMBP	Cyclic α -maltosyl-(1,6)-maltose	<i>Arthrobacter globiformis</i>	Binding test Crystal structure	(Kohno et al. 2020)
	Literature	BpAXBP1	Xylooligosaccharides (XOS) and AXOS4 (arabinoxyloligosaccharides whose nonreducing end of the	<i>Bifidobacterium pseudocatenulatum</i>	Transcriptional analysis Binding test	(Saito et al. 2020)

			xylose backbone was not decorated with arabinose)			
		BpAXBP2	Arabinoxyloligosaccharides (AXOS), including double arabinofuranosyl-decorated XOS (AXOS2)			
		BpAXBP3	AXOS2			
	Literature	α GlyBP	Disaccharide α -glycosides	<i>Thermus thermophilus HB8</i>	Crystal structure Binding test	(Chandravanshi et al. 2020)
	Literature	β GlyBP	β -glucosides	<i>Thermus thermophilus HB8</i>	Binding test Crystal structure	
	Literature	RhLNBBP	Lacto-N-biose (LNB), galacto-N-biose (GNB), lacto-N-tetraose (LNT)	<i>R. hominis</i>	Transcriptional analysis	(Pichler et al. 2020)
		RiLe ^{a/b} BP	Fucosyl-decorated Lewis ^b (Le ^b) tetraose and Lewis ^a (Le ^a) triose, lacto-N-biose (LNB), galacto-N-biose (GNB)	<i>R. inulinivorans</i>	Binding test Crystal structure	
	Literature	GL-BP	Lacto-N-biose (LNB), galacto-N-biose (GNB)	<i>B. longum JCM 1217</i>	Crystal structure	

		FL transporter-2	Fucosyllactose (FL)	<i>B. infantis</i> JCM 1222		(Fushinobu and Abou Hachem 2021)
		BLAXBP	Xylooligosaccharides (DP2-6), arabinoxylan-oligosaccharides (DP3-5)	<i>Bifidobacterium animalis</i> subsp. <i>lactis</i> B1-04		
		BIG16BP	Oligosaccharides sharing a [Gal/Glc] α 1,6Glc	<i>Bifidobacterium</i> except for <i>B. bifidum</i>		
	Literature	FpMOBP	β -mannooligosaccharides (β -MOS)	<i>Faecalibacterium prausnitzii</i> SL3/3	Transcriptional analysis Binding test	(Lindstad et al. 2021)
	Literature	NagB1 and NagB2	N-acetyl-D-glucosamine disaccharides (GlcNAc) ₂ , (GlcNAc) ₃	<i>Paenibacillus</i> sp. str. <i>FPU-7</i>	Crystal structure Binding test	(Itoh et al. 2021)
	Literature	AviABCD	Cellulose oligosaccharides (COS)	<i>Caldicellulosiruptor bescii</i>	Transcriptional analysis	(Rodionov et al. 2021)
		MosABC	Mannose oligosaccharides (MOS)			
		pecXYZ	Pectin oligosaccharides (POS)			
		hemXYZ	Pectin oligosaccharides (POS)			
		RhaXYZ	Rhamnogalacturonides			
		RglFGH	Rhamnogalacturonides			

		BxgLFG	Glucuronoxylan oligosaccharides (GXOS)			
		AxoFGE	β -xylooligosaccharide (β -XOS)			
		XloEFG	β -xylooligosaccharide (β -XOS)			
		XynUVW	xylooligosaccharide (XOS)			
		AxuABC	α -xylooligosaccharide (α -XOS)			
		AxeUVW	Arabinoxyloligosaccharides (AXOS)			
		MalEFG	Maltooligosaccharides (MOS)			
		KojEFG	α -glucoside disaccharides trehalose and kojibiose			
	Literature	inuABCDEF	Inulin (DP of > 8)	<i>Lactobacillus delbrueckii</i> JCM 1002 ^T	Transcriptional analysis TLC HPLC	(Tsujikawa et al. 2021)
	Literature	ESBP (BAD_RS06835), bglE	Copra meal β -manno-oligosaccharides (CM- β -MOS) (DP2, DP3)	<i>Bifidobacterium adolescentis</i> DSMZ 20083	Transport assay Transcriptional analysis Binding test Crystal structure	(Mary et al. 2022)

	Literature	msmEFG	Raffinose, stachyose, melibiose	<i>Bifidobacterium</i>	Transcriptional analysis	(Su et al. 2022)
	Literature	xuaABC	Arabinoxylan dextrins	<i>Ruminiclostridium cellulolyticum</i>	Transcriptional analysis Binding test	(Liu et al. 2022)
	Literature	NgcEFG	Chitooligosaccharides, GlcNAc, (GlcNAc) ₂	<i>Streptomyces sp. UH6</i>	Transcriptional analysis	(Duhsaki et al. 2022)
		DasABC				
	Literature	CbpB	Cellodextrin (G2 to G5), laminaribiose	<i>Clostridium thermocellum</i>	Binding test Crystal structure	(Yan et al. 2022)
	Literature	SBP _{MLG}	Barley β -glucan, mixed-linkage β -glucans (β -1,3/1,4)	<i>Blautia producta</i> ATCC 27340	Targeted gene deletion ITC Circular dichroism spectroscopy AlphaFold prediction	(Singh et al. 2023)
	Literature	MusEFGK ₂ I	Maltose, maltotriose	<i>Gardnerella spp.</i>	Transcriptional analysis	(Bhandari and Hill 2023)
		MalXFGK	Maltooligosaccharides and/or maltodextrins	<i>Gardnerella spp.</i>		
		RafEFGK	Glucosides and galactosides	<i>G. vaginalis</i>		

		TMSP	Trehalose, maltose, sucrose, and palatinose	<i>G. vaginalis</i>		
		MalEFG	Maltose and maltodextrins	<i>G. leopoldii</i>		
PepT	3.A.1.5.16	BglpEFGKL (TM_0031)	Cellobiose, laminaribiose	<i>Thermotoga maritima</i> MSB8	Transcriptional analysis Binding test	(Conners et al. 2005; Nanavati et al. 2005)
	3.A.1.5.32	F0TFS5-9 (LAC30SC_07285)	Stachyose	<i>Lactobacillus acidophilus</i> NCFM	Transcriptional analysis Gene deletion	(Andersen et al. 2012)

(3) Taxonomical diversity

We examined the taxonomical diversity of all the characterized bacterial ABC transporters targeting glycosides from the CUT1 and PepT families (Table 4). These 119 glycoside ABC transporters occur widely in Gram-positive and negative bacteria, including seven phyla and 21 orders from different environments, some of which cause human diseases, such as *Streptococcus mutans* and *Enterococcus faecalis* (Webb et al. 2008; Sauvageot et al. 2017). Of the 119 glycoside ABC transporters, more than 90 % were assigned to the four phyla Bacillota, Actinomycetota, Pseudomonadota and Thermotogae. However, the actual taxonomical distribution of glycoside ABC transporters is highly biased as a result of the sequencing and biochemical characterization efforts, which are unequal in their coverage of bacterial diversity. A search of the Pfam 32.0 database for the domain (PF01547) of the SBP of the maltose ABC transporter alone (El-Gebali et al. 2019) retrieved 25,735 sequences across 4,106 species, highlighting the huge diversity of ABC transporters targeting glycosides, and the fact that extremely few of them have been characterized (Table 4). Notably, there is no glycoside targeting ABC transporters from the Bacteroidota phylum in the TCDB, and we did not find any characterized in the literature. In the Pfam 32.0 database, however, 56 Bacteroidota species harbor 74 sequences containing the Pfam domain of the SBP of the maltose ABC transporter. In addition, the PULDB has recently included the annotation of ABC transporters, since they appear, with SusC/D transporters, in around 200 PULs (Terrapon et al. 2018). It is thus likely that some Bacteroidota ABC transporters play a role in glycoside transport, but this remains to be proven. More generally, the taxonomical diversity of the glycoside-targeting ABC transporters already characterized is far from an accurate representation, as most ABC transporters have still to be characterized.

(4) Mechanism of glycoside import by ABC transporters

Since the sequencing of the genes encoding the maltose ABC transporter in the 1970s (Silhavy et al. 1979), the functional and structural characterization of maltooligosaccharide transporters (especially in *E. coli*) has resulted in the establishment of a glycoside ABC transporter paradigm. Maltooligosaccharide import takes place via a classic alternating access mechanism and substrate specificity is controlled by the outer membrane maltoporin, the periplasmic protein MalE (also known as MBP, for maltose-binding protein, in maltooligosaccharide ABC transporters) and the transmembrane subunits (MalFG) (Figure 2) (Schneider 2001; Mächtel et al. 2019).

MalE consists of two globular lobes (the N-terminal and the C-terminal domains) between which a substrate-binding pocket is formed. The shape complementarity, extensive hydrogen bonding, and aromatic stacking interactions together account for the high-affinity binding (in the micromolar range) of MalE to maltooligosaccharides (Quioco et al. 1997). The conformation of MalE is predominantly open in the absence of a ligand. Once the maltooligosaccharide passes through the maltoporin, MalE will completely trap the substrate at the interface of the two lobes and then dock with the transmembrane subunits, triggering ATPase activity and initiating the transport cycle (Hospital et al. 2010; Mächtel et al. 2019). Crystal structures of the MBP-MalFGK2 transporter complex from *E. coli* K-12 bound with a maltooligosaccharide in two different conformational states revealed that i) in the pre-translocation structure, MalG forms two hydrogen bonds with maltooligosaccharide, up to four glucosyl units at the reducing end, and ii) in the outward-facing conformation, the MalF transmembrane subunit binds glucosyl units from the non-reducing end (Oldham et al. 2013; Oldham et al. 2007). Thus, here MalFG provides another level of selection of the substrate, in addition to the levels due to LamB and MalE. The substrate can thus be unidirectionally imported with TMDs switching between inward- and outward-facing conformations, and ATP hydrolysis is stimulated by both MalE, which stabilizes the semi-open configuration of MalFGK2, and open MalE, which stabilizes the closed state of MalFGK2 (Mächtel et al. 2019). The binding cavity of the pre-translocation state of the MBP-MalFGK2 complex would be large enough to accommodate seven glucosyl units; this corresponds to the maximum DP of the oligosaccharides that can be transported (Oldham et al. 2013). The maltooligosaccharide ABC transporter could serve as a smart template for understanding how other glycoside ABC transporters evolve and function, despite their structural and functional diversity.

Table 4. Phylum and order distribution of the characterized glycoside-targeting ABC transporters in bacteria.

Phylum	Order	Number of characterized ABCs	Percentage of ABCs characterized	
			Order	Phylum
Actinomycetota	Actinomycetales	5	4.2	31.9
	Corynebacteriales	1	0.8	
	Bifidobacteriales	27	22.7	
	Streptomycetales	3	2.5	
	Mycobacteriales	1	0.8	
	Micrococcales	1	0.8	
Bacillota	Bacillales	14	11.8	47.9
	Lactobacillales	8	6.7	
	Clostridiales	11	9.2	
	Thermoanaerobacterales	17	14.3	
	Eubacteriales	7	5.9	
Pseudomonadota	Enterobacterales	3	2.5	8.4
	Rhizobiales	5	4.2	
	Sphingomonadales	1	0.8	
	Bdellovibrionales	1	0.8	
Cyanobacteriota	Synechococcales	1	0.8	1.7
	Nostocales	1	0.8	
Thermotogae	Thermotogales	4	3.4	3.4
Deinococcota	Thermales	6	5.0	5.0
Campylobacterota	Campylobacterales	1	0.8	0.8
Fusobacteriota	Fusobacteriales	1	0.8	0.8

V. PTS transporters of glycosides

The phosphotransferase system (PTS) was first described in 1964 in *Escherichia coli* K235, as a novel system coupling sugar transport with sugar phosphorylation from phosphoenolpyruvate (PEP) as the phosphoryl donor, in the laboratory of Saul Roseman (Kundig et al. 1964). Since then, PTSs have been identified in various species of Gram-positive and Gram-negative bacteria as well as in archaea, transporting a wide range of monosaccharides, glycosides and polyols (Saier 2015). The PTS thus represents a major carbohydrate active-transport system in bacteria (Saier 2000). It is a multicomponent system which always involves enzymes of the plasma membrane and others in the cytoplasm (Deutscher et al. 2006). The membranous component of the PTS system, enzyme II (EII), is made up of three to four subunits: IIA, IIB, IIC and sometimes IID, which are specific for one substrate or, in a few cases, for a small group of closely related carbohydrates (Erni 2013). The cytoplasmic components are enzyme I (EI) and histidine protein (HPr), which are not carbohydrate specific. Together, they constitute a phosphorylation cascade, allowing the phosphoryl group to be successively transferred from PEP to EI, to the HPr, then to subunits A and B of EII, and finally to the carbohydrate imported by the subunit EIIC or by the EIICD complex (Figure 3) (Deutscher et al. 2014).

(1) Diversity of sequences and structures

Based on sequence similarities between the C domain of the EII complex, carbohydrate PTSs have been classified into seven families in the TCDB (Saier 2006): glucose-glucoside (4.A.1, Glc), fructose-mannitol (4.A.2, Fru), lactose-*N,N*-diacetylchitobiose (4.A.3, Lac), glucitol (4.A.4, Gut), galactitol (4.A.5, Gat), mannose-fructose-sorbose (4.A.6, Man), and L-ascorbate (4.A.7, L-Asc). The Glc family is further divided into the glucose (4.A.1.1) and glucoside (4.A.1.2) subfamilies. It is noteworthy that some families include members which are nearly as distant from each other as they are from members from other families (Jeckelmann and Erni 2019). PTS family names are not always consistent with the function of their members. For example, there are cellobiose PTS transporters in both the Glc family (Francel et al. 2010) and the Lac family (Kowalczyk et al. 2008). Based on our review of the carbohydrate PTSs in the TCDB and in the literature (Table 5), glycoside PTS transporters are only listed in the Glc, Lac and Man families.

Depending on the family, the EII complexes of the glycoside PTSs are encoded either by one or by a number of encoding genes. The Glc family contains members with either a single

multidomain protein including the A, B and C domains, or two distinct proteins, with one for the A domain and another one for the B and C domains. Some of the members of the Glc family lack their own IIA domain, and use the IIA domain of another system of the same family instead (Tchieu et al. 2001). By contrast, the members of the Lac family always have multidomain EII complexes classified into two subfamilies: one (4.A.3.1) consisting of two polypeptide chains (EIIA and EIIBC) and another (4.A.3.2) including three individual polypeptide chains (EIIA, EIIB and EIIC). The Man family is the only PTS family whose members possess an additional IID domain. The EIIA, B, C and D domains of glycoside PTSs in the Man family can be present either in four individual proteins or in three individual proteins with the A and B domains in the same protein. The Glc, Lac and Fru families belong to a phylogenetic PTS-GFL superfamily (Nguyen et al. 2006) because of the structural homogeneity of their IIC domains, with a uniform 10 transmembrane helix topology (Cao et al. 2011; McCoy et al. 2016). The members of the Man family are evolutionarily distinct from the GFL superfamily (Chen et al. 2012). In this family, the IIC domains probably consist of six or seven transmembrane helices, and the second integral membrane IID domains probably contains six transmembrane helices, based on the HMMTOP prediction (<http://www.enzim.hu/hmmtop/index.php>). However, the only solved structure of a mannose PTS transporter (ManYZ) from the Man family shows nine transmembrane helices in both the IIC (ManY) and IID (ManZ) domains (Liu et al. 2019). The structural organization and thus, the classification of glycoside PTSs, have a complex evolutionary history because of numerous horizontal gene transfer events, duplications, and non-orthologous displacements (Zúñiga et al. 2005).

(2) Functional diversity

Only a few dozen PTSs targeting glycosides have been functionally characterized using one or more of the methods listed above (Table 5). Most of them are specific for α - or β - linked glucosides, such as maltose, sucrose and cellobiose, or their derivatives such as N,N'-diacetylchitobiose and salicin, as the GFL superfamily is by far the largest (Jeckelmann and Erni 2019). Recently, transcriptomic analyses revealed that, in *Lactobacillus plantarum* strain WCFS1, the PTS1 (Panwar and Kapoor 2020; Saulnier et al. 2007) of the Glc family, and the PTS23C (Panwar and Kapoor 2020) of the Lac family, might be involved in galactomannose/galactomannobiose and mannobiose utilization, respectively, while the specificity of the two PTSs still needs to be further characterized. In addition, in the Man family, few PTSs are responsible for β -galactosides, β -mannosides, hyaluronic acid and fucosyl- α -1,3-*N*-acetylglucosamine. Glycoside PTSs are thus mainly specific for disaccharides, although

some are able to transport longer oligosaccharides with a DP of no more than four. Other PTS transporters in the list were shown to be involved in the utilization of hyaluronic acid, glucomannooligosaccharides or lichenan oligosaccharides by transcriptional analysis or deletion of one of the PTS encoding genes, but this does not mean that the PTS can transport glycosides with a DP of more than four.

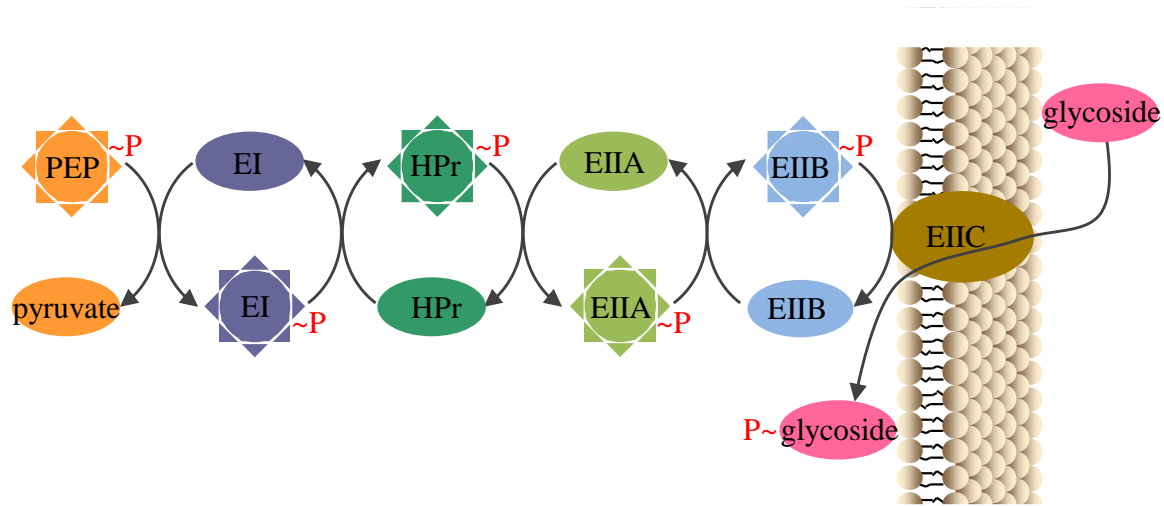


Figure 3. Schematic illustration of the process of glycoside transport and phosphorylation by the PTS.

This figure shows the serial transfer of the phosphoryl group is transferred to the EI, HPr, EIIA and EIIB, which ultimately phosphorylates the sugar imported by the transmembrane EIIIC.

Table 5. List of the bacterial PTS transporters involved in glycoside utilization, characterized either by binding or transport assays, gene deletion and/or transcriptional approaches.

The glycoside PTS transporters retrieved from the literature but which are not listed in the TCDB were classified based on their conserved domains by using a BLAST analysis against the NCBI's Conserved Domain Database (Marchler-Bauer et al. 2017).

TCDB Family name	TCID	Protein name (locus tag)	Substrate	Organism	Characterization method	Literature
Glc	4.A.1.1.3	MalX (b1621)	Maltose	<i>Escherichia coli</i> K12	Heterologous expression in <i>E. coli</i> strain	(Reidl and Boos 1991)
	4.A.1.1.8	MalP (BSU08200)	Maltose	<i>Bacillus subtilis</i> 168	Transcriptional analysis Gene deletion	(Yamamoto et al. 2001)
	4.A.1.1.10	AgIA (FYK33_RS28885)	Isomaltose, trehalulose, maltulose, leucrose, palatinose	<i>Klebsiella pneumoniae</i> ATCC 23357	Heterologous express in <i>E. coli</i> strain Transport assay	(Pikis et al. 2006)
	4.A.1.1.11	MalT (SMU_2047)	Maltose, maltotriose	<i>Streptococcus mutans</i> ATCC 700610	Gene deletion	(Webb et al. 2007)
	4.A.1.1.12	MalT (SPy_1986)	Maltose, maltotriose	Group A <i>Streptococcus</i> MGAS5005	Transcriptional analysis	(Shelburne et al. 2008)

					Gene deletion	
4.A.1.1.17	MalT (BCE33L0344)	Maltose	<i>Bacillus cereus</i> E33L	Transport assay Binding test Crystal structure	(McCoy et al. 2016)	
4.A.1.2.1	ScrA	Sucrose	<i>Klebsiella pneumoniae</i> KAY2026	Phosphorylation activity	(Titgemeyer et al. 1996)	
4.A.1.2.3	AscF (b2715)	Salicin	<i>Escherichia coli</i> K-12	Gene deletion	(Desai et al. 2010)	
4.A.1.2.4	TreB (b4240)	Trehalose, maltose, sucrose	<i>Escherichia coli</i> K-12	Heterologous expression in other <i>E. coli</i> strain Gene deletion	(Klein et al. 1995; Steen et al. 2019)	
4.A.1.2.5	BglF (cgR_2729)	Methyl-β-glucoside, salicin, arbutin	<i>Corynebacterium glutamicum</i> R	Heterologous expression in <i>C. glutamicum</i>	(Kotrba et al. 2003)	
4.A.1.2.6	BglP	Aesculin, arbutin	<i>Streptococcus mutans</i> NG8	Gene deletion Heterologous expression in <i>E. coli</i>	(Cote et al. 2000)	
4.A.1.2.9	SacP (BSU38050)	Sucrose	<i>Bacillus subtilis</i> 168	Transport assay	(Fouet et al. 1987)	

	4.A.1.2.12	PtsS (Cgl2642)	Sucrose	<i>Corynebacterium glutamicum</i> ATCC 13032	Gene deletion	(Moon et al. 2005)
	4.A.1.2.14	PTS15 (LGAS_1669)	Cellobiose	<i>Lactobacillus gasseri</i> ATCC 33323	Transcriptional analysis Gene deletion	(Franci et al. 2010)
	4.A.1.2.16	PTS20 (LGAS_1778)	Sucrose	<i>Lactobacillus gasseri</i> ATCC 33323	Transcriptional analysis Gene deletion	(Franci et al. 2010)
	Literature	scrT (SP_1722)	Sucrose	<i>Streptococcus pneumoniae</i> TIGR4	Gene deletion	(Iyer and Camilli 2007)
	Literature	PTS EIIBCA (lp_0185)	Short-chain fructooligosaccharides (scFOS)	<i>L. plantarum</i> WCFS1	Transport assay Transcriptional analysis	(Saulnier et al. 2007)
	Literature	PTS1 (LPST_C0152)	Fructooligosaccharides (DP2-5)	<i>Lactobacillus plantarum</i> ST-III	Transcriptional analysis Gene deletion	(Chen et al. 2015)
	Literature	PTS26 (LPST_C2650)	Fructooligosaccharides (DP2-5)	<i>Lactobacillus plantarum</i> ST-III	Transcriptional analysis Gene deletion	(Chen et al. 2015)
	Literature	PTS EIIBCA	Sucrose	<i>Lactobacillus plantarum</i>	Genomic analysis	(Buntin et al.

		FosABCDX PTS	Fructooligosaccharides (FOS)		Proteomic analysis	2017)
Literature	Lac_PTS, LacEF	Lactose	<i>L. casei, L. paracasei</i> W20	Transport assay Transcriptional analysis	(Böger et al. 2019)	
Literature	GAG-PTS (EIIABCD)	Unsaturated hyaluronan disaccharide	<i>Streptococcus agalactiae</i>	Transport assay Gene deletion Crystal structure	(Oiki et al. 2019)	
Literature	EIIB ^{Cel} , EIIB ^{Cel}	β-1,6-oligosaccharides	<i>Enterococcus faecalis</i>	Gene deletion Binding test Transcriptional analysis	(Grand et al. 2019)	
Literature	EIIA/B ^{Cel1} , EIIA/B ^{Cel2}	Cellobiose	<i>Listeria monocytogenes</i>	Gene deletion	(Cao et al. 2019)	
Literature	PTS_EIIBCA (CCG34776)	Kestose	Uncultured <i>Dorea</i>	Transcriptional analysis Gene deletion	(Wang et al., 2020)	
Literature	celBCD	Cellobiose, β-galactosides	<i>G. stearothermophilus</i> T-1	Transport assay Transcriptional analysis	(Shulami et al. 2020)	
Literature	bglP	β-glucoside	<i>Streptococcus pyogenes</i>	Gene deletion	(Braza et al. 2020)	

				(group A <i>Streptococcus</i> [GAS])	Transcriptional analysis	
Literature	PTS23C (lp_3010)	Un-substituted galactomannooligosaccharides (GMOS, DP2 and DP3)	<i>Lactobacillus plantarum</i> WCFS1	Transcriptional analysis	(Panwar and Kapoor 2020)	
	PTS1BCA (lp_0185)	Substituted galactomannooligosaccharides (GMOS, DP2 and DP3)				
Literature	EIIABC ^{cel2}	Cellobiose	<i>Thermoanaerobacterium aotearoense</i> SCUT27	Transcriptional analysis Gene deletion	(Qu et al. 2021)	
Literature	PTS IIC (scrA)	Fructooligosaccharides	<i>L. murine</i> (A56) <i>E. hirae</i> (BG) <i>E. faecalis</i> (E1) <i>L. casei</i> (B52) <i>E. gallinarum</i> (B41) <i>S. alactolyticus</i> (B28)	Transcriptional analysis	(Gu et al. 2022)	
	PTS EIIBCA	Fructooligosaccharides	<i>L. plantarum</i> (A64)			
	PTS IIBC and IIAB	Fructooligosaccharides	<i>L. rhamnosus</i> (B63)			

		PTS IIBCA	Fructooligosaccharides	<i>L. paracasei (B45)</i>		
Literature	- (peg.1635, peg.4961 and peg.1634)		Chitooligosaccharides, GlcNAc	<i>Streptomyces sp. UH6</i>	Transcriptional analysis	(Duhsaki et al. 2022)
Literature	PTS IIBCA		Short-chain oligofructose	<i>L. plantarum QS7T</i>	Transcriptional analysis	(Zhu et al. 2022)
Literature	ScrA		Sucrose, glucose	<i>Streptococcus thermophilus LMD-9</i>	Transcriptional analysis Gene deletion	(Gasser et al. 2022)
Literature	-		Beta-glucoside, cellobiose	<i>Gilliamella Lactobacillus Firm-5</i>	Transcriptional analysis	(Su et al. 2022)
Literature	mini PTS		Cellobiose, glucose, pentoses	<i>Ruminiclostridium cellulolyticum</i>	Transcriptional analysis Gene deletion Binding test Crystal structure Phosphorylation activity	(Xu et al. 2023)
Literatur	PTS		Arbutin, cellobiose, fructose, glucose, trehalose	<i>P. thermoglucosidasius DSM 2542</i>	Gene deletion	(Bidart et al. 2023)

Lac	4.A.3.1.2	PTS 6CB (LGAS_343)	Lactose	<i>Lactobacillus gasseri</i> ATCC33323	Transcriptional analysis Gene deletion	(Francis et al. 2012)
	4.A.3.1.3	PTS 9BC (LGAS_501)	Lactose	<i>Lactobacillus gasseri</i> ATCC33323	Transcriptional analysis Gene deletion	(Francis et al. 2012)
	4.A.3.2.1	ChbC (b1737)	N,N'-diacetylchitobiose, cellobiose	<i>Escherichia coli</i> K-12	Gene deletion	(Kachroo et al. 2007)
	4.A.3.2.2	LicC (BSU38580)	Lichenan oligosaccharides (DP unclear)	<i>Bacillus subtilis</i> 168	Gene deletion	(Tobisch et al. 1997)
	4.A.3.2.5	ChbC	N,N'-diacetylchitobiose	<i>Serratia marcescens</i> 2170	Gene deletion	(Uchiyama et al. 2003)
	4.A.3.2.6	CelB (VC_1282)	N,N'-diacetylchitobiose	<i>Vibrio cholerae</i> serotype O1 ATCC 39315	Transcriptional analysis Gene deletion	(Meiborn et al. 2004)
	4.A.3.2.7	CelC (KP1_4524)	Cellobiose	<i>Klebsiella pneumoniae</i> NTUH-K2044	Gene deletion	(Wu et al. 2012)
	4.A.3.2.8	CelB (BCE_5325)	N,N'-diacetylchitobiose	<i>Bacillus cereus</i> ATCC 10987	Transport assay Gene deletion Structure	(Cao et al. 2011)
	4.A.3.2.9	Lactose transporter	Lactose	<i>Klebsiella pneumoniae</i>	Transport assay	(Imai and Hall)

				CG43		1981)
	Literature	ChbBCA	Chitin disaccharide (GlcNAc-GlcNAc; N,N'-diacetylchitobiose) and chitin trisaccharide (GlcNAc-GlcNAc-GlcNAc; N,N',N''-triacetylchitotriose)	<i>Escherichia coli K12</i>	Gene deletion Transport assay	(Walter et al. 2021; Keyhani et al. 2000)
	Literature	GenB (OG1RF_10235)	Gentiobiose, amygdalin	<i>Enterococcus faecalis</i> OG1RF	Gene deletion Transcriptional analysis	(Grand et al. 2019)
Man	4.A.6.1.13	AlfF (LCAB_28290)	Fucosyl- α -1,3-N-acetylglucosamine	<i>Lactobacillus casei</i> BL23	Gene deletion	(Rodríguez-Díaz et al. 2012)
	4.A.6.1.14	PTS-IIC (spr0294)	Hyaluronic acid (DP unclear)	<i>Streptococcus pneumoniae</i> ATCC BAA-255	Gene deletion	(Marion et al. 2012)
	4.A.6.1.18	GnbC (LCABL_02930)	Galacto N-biose and lacto N-biose	<i>Lactobacillus casei</i> BL23	Gene deletion Transcriptional analysis	(Bidart et al. 2014)
	Literature	Mpt PTS	N-acetylglucosamine; glucosamine	<i>Enterococcus faecalis</i>	Gene deletion	(Keffeler et al. 2021)
		PTS IICBA (EF1516, NagE)				

(3) Taxonomical diversity

We examined the taxonomical diversity of all the characterized bacterial PTS transporters targeting glycosides derived from the TCDB Glc, Lac and Man families and from the literature (Table 6). These 57 glycoside PTS transporters are divided between both Gram-positive and Gram-negative bacteria, including three phyla and eleven species. More than half of the characterized glycoside-targeting PTS transporters are from species of the Bacillota phylum. The PTSs targeting glucosides are distributed between three phyla (Bacillota, Pseudomonadota and Actinomycetota), while PTSs targeting mannosides and galactosides all belong to species of the Bacillota phylum. The PTS content is consistent with the species' potential for facultative anaerobic and anaerobic growth (Jeckelmann and Erni 2019), which correlates with the fact that the Bacillota phylum is one of the dominant phyla in the mammalian gut microbiota. Thirty percent of the glycoside PTSs from Bacillota belong to the *Lactobacillus* genus; the species can utilize diverse glycosides of this genus found in the human gut, allowing them to survive in this ecosystem. This trait, and the fact that many *Lactobacillus* species impart benefits to the general health and wellness to the host, make them useful as probiotics (Franci et al. 2010). Nevertheless, PTSs are also present in some pathogenic species, such as *Klebsiella pneumoniae* NTUH-K2044, which causes nosocomial infections and opportunistic infections (Wu et al. 2012), and *Vibrio cholerae* O1, which causes a dehydrating diarrheal illness (Meiborn et al. 2004). PTSs might thus play a role in habitat colonization by these strains, and thus, in bacterial infection.

(4) Mechanism of glycoside import by PTS transporters

As described above, the transport of glycosides by PTS is accompanied by a cascade of phosphorylations, in which the phosphate moiety is transferred from PEP to the imported glycoside by a set of cytoplasmic phosphocarrier proteins (EI, HPr, EIIB and EIIC, serially). EI, HPr and IIA are all phosphorylated at a histidine residue. The IIB domains of the GFL superfamily are phosphorylated at a cysteine, while the IIB domains of the mannose family are phosphorylated at a histidine residue. The crystal structures of several EI, HPr, IIA and IIB proteins have been solved, alone or in complex with other PTS, which have been described in previous reviews (Robillard and Broos 1999; Peterkofsky et al. 2001; Deutscher et al. 2006; Erni 2013).

Little was known about how the transmembrane EIIC domain selectively imports the glycoside across the membrane until a high-resolution crystal structure of diacetylchitobiose EIIC protein

(bcChbC) from *Bacillus cereus* E33L was solved in the inward-facing conformation (Cao et al. 2011). A few years later, the maltose-specific EIIC protein (bcMalT) from the same strain was solved in the inward-facing and outward-facing conformations (McCoy et al. 2016; Ren et al. 2018). Even if bcChbC and bcMalT have different specificities, they belong to the same GFL superfamily and both proteins are homodimers with 10 transmembrane helices, forming a scaffold domain (SD) containing the large and mostly hydrophobic protomer-protomer interface and a transport domain (TD) containing the sugar binding site and translocation pathway. A so-called elevator mechanism was therefore proposed. Overall, the substrate firstly binds to the transporter (IIC) in the outward-facing open conformation, then the substrate is occluded and translocated by the rigid body motion of TD against the immobile SD across the membrane, where the EIIC isomerizes to the inward-facing open conformation (Figure 4) (Jeckelmann and Erni 2019). The isomerization rate is increased 100–1,000-fold by the activity of the phosphorylated IIB domain and rendered unidirectional by the concomitant phosphoryl transfer from the phosphorylated IIB domain to the substrate in the binding cavity and/or by the subsequent release of the phosphorylated substrate. However, it is not known whether the two protomers isomerize independently or in synchrony, and how the IIB and IIC interact in the phosphate transformation process.

PTSs in members of the Man family contain a unique additional IID transmembrane domain. The only established structure of an EIICD complex is that of the *E. coli* K-12 protein specific for mannose (Liu et al. 2019), which also suggested a possible elevator mechanism. However, the feasibility of this mechanism for the glycoside PTS from the Man family still needs further investigation.

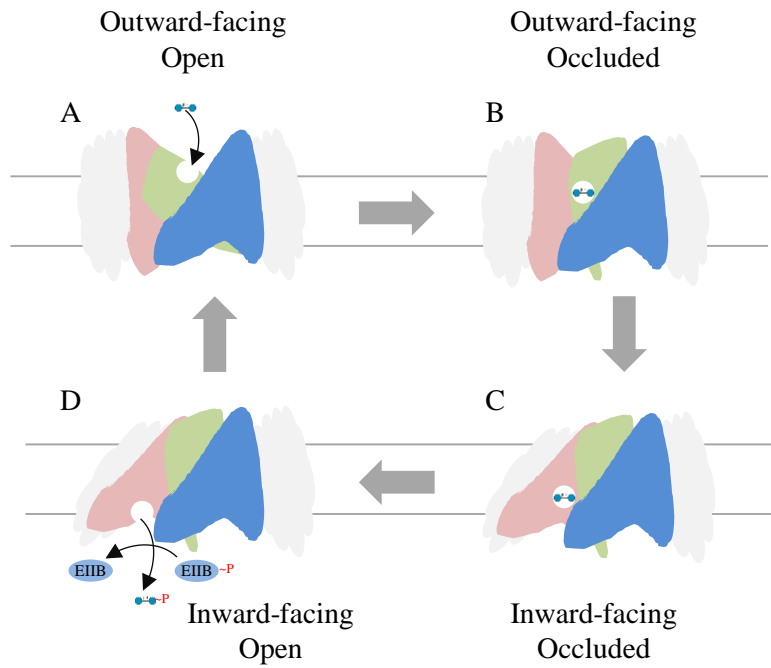


Figure 4. Carbohydrate transport through the transmembrane PTS_EIIC domain.

This figure shows a model of the glucose superfamily EIIC transport cycle. (A–D) The substrate binds to the outward-facing open structure (A), which is then converted to the outward-facing occluded state (B) by means of a movement of the transport domain. The transport domain moves toward the cytoplasmic side of the membrane to form the inward-facing occluded state (C); finally, the substrate is phosphorylated by the phosphoryl from the phosphorylated-IIB domain and is released to the cytoplasmic space (D).

Table 6. Phylum and order distribution of the characterized glycoside-targeting PTS transporters in bacteria.

Phylum	Order	Number of characterized PTSs	Percentage of PTSs characterized	
			Order	Phylum
Bacillota	Bacillales	8	14.0	75.4
	Lactobacillales	32	56.1	
	Clostridiales	1	1.8	
	Eubacteriales	1	1.8	
	Thermoanaerobacterales	1	1.8	
Pseudomonadota	Enterobacterales	10	17.5	19.3
	Vibrionales	1	1.8	
Actinomycetota	Corynebacteriales	2	3.5	5.3
	Streptomycetales	1	1.8	

VI. MFS transporters of glycosides

Along with the ABC superfamily, the major facilitator superfamily (MFS) is one of the two exceptionally large superfamilies of transporters, with millions of sequenced members found ubiquitously in bacteria, archaea, and eukaryotes (Saier 2006). MFS transporters are single-polypeptide secondary carriers capable only of transporting small solutes in response to chemiosmotic ion gradients. MFS transporters are the largest family of secondary membrane transporters. They can be classified into symporters, antiporters and, in case of a driving substance being absent, facilitators (Figure 5) (Yan 2015). In the TCDB (February 2020 release), the MFS superfamily is subdivided into 104 families, consisting of 89 families classified under the main MFS family and annotated TC# 2.A.1, and a further 15 families including distant members of MFSs. Each family contains members specific to a different set of related compounds, including monosaccharides, oligosaccharides, polyols, drugs, neurotransmitters, Krebs cycle metabolites, phosphorylated glycolytic intermediates, amino acids, peptides, organic anions, and inorganic anions (Reddy et al. 2012).

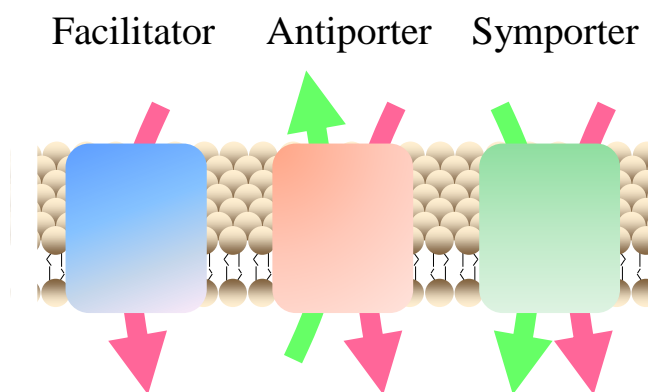


Figure 5. Schematic representation of the three types of MFS transporters.

This figure shows the difference in the transport process for facilitator, antiporter and symporter. The facilitator catalyzes substrate diffusion across the membrane down its concentration gradient. The symporter or antiporter uses the energy released from the downhill translocation of one substrate to drive the uphill translocation of another substrate either in the same direction (symporter) or in the opposite direction (antiporter).

(1) Diversity of sequences and structures

There are 104 established MFS families in the TCDB, with bacterial MFS transporters of glycosides found in three of them: the Oligosaccharide: H⁺ Symporter (OHS) family (2.A.1.5), the Fucose: H⁺ Symporter (FHS) family (2.A.1.7) and the Glycoside-Pentoside-Hexuronide (GPH): Cation Symporter family (2.A.2) (Table 7). The GPH family is one of the distant families of the MFS superfamily, since transporters of the GPH family show only marginal sequence similarities with other members of the other families. Position-Specific Iterated BLAST (PSI-BLAST) analysis suggested that these proteins all arose from a common ancestor (Saier 2000). Glycoside MFS transporters are generally about 400-550 amino acids in length, with the exception of the Gram-positive bacterial lactose permeases (LacS) which are larger because of an additional EII- like domain at the C-terminal extremity, usually found in PTS transporters. Despite their low sequence similarities, all these proteins share the same topology, with 12 transmembrane helices where the first six transmembrane helices exhibit significant sequence similarity with the last six (Saier 2000).

(2) Functional diversity

The bacterial MFS transporters characterized as responsible for the uptake of different glycosides are listed in Table 7. The transporters of the OHS family have been reported to be specific for a set of galactosides (such as lactose, raffinose and melibiose), sucrose and maltose. Among the members of the OHS family, the LacY permease involved in the uptake of lactose and other galactosides by *Escherichia coli* K-12 was the first MFS member to be investigated (Büchel et al. 1980) and is still one of the best characterized MFS transporters (Kaback and Guan 2019).

Only one member of the FHS family, the ScrT transporter specific for sucrose uptake from *Shewanella frigidimarina* NCIMB 400, has been functionally characterized as specific for oligosaccharide transport (Rodionov et al. 2010); the other characterized transporters of this family are involved in the transport of monosaccharide or its derivatives. By contrast, the members of the GPH family exhibit a broad specificity towards glycosides, including galactosides, glucosides, xylosides and fructooligosaccharides.

Some transporters, including members of the OHS family such as LacY from *Escherichia coli* K-12 (King and Wilson 1990), CscB from *Escherichia coli* EC3132 (Peng et al. 2009), MelY from *Enterobacter cloacae* ATCC 13047 (Shinnick et al. 2003; Tavoulari and Frillingos 2008)

and the only member of the GPH family: LacS from *Lactobacillus acidophilus* NCFM (Andersen et al. 2011), are able to transport different glycosides that share structural traits. For example, the LacY protein has the ability to internalize lactose, melibiose, thio- β -methyl galactopyranoside (TMG), 4-nitrophenyl- β -D-galactopyranoside and 4-nitrophenyl- α -D-galactopyranoside, because there is one sugar-binding site in LacY that selectively recognizes the galactosyl moiety (Kaback and Guan 2019).

From the listing of the substrates of the characterized glycoside MFS transporters, they appear to be mainly specific for disaccharides or their derivatives. However, MFS transporters can also import longer glycosides. Using gene deletion and growth dynamic analysis, two transporters were shown to be able to internalize oligosaccharides up to DP 4: the MFS from an uncultured *Bacteroides* strain (Tauzin, Kwiatkowski, et al. 2016) and the FosT from *Escherichia coli* BEN2908 (Schouler et al. 2009), which internalize xylotetraose and fructotetraose, respectively. Other MFs, such as TogT (Hugouvieux-Cotte-Pattat and Reverchon 2001) from *Erwinia chrysanthemi* 3937 and LacS (Lba1463) (Andersen et al. 2011) from *Lactobacillus acidophilus* NCFM, have been shown to be responsible for growth on a mixture of oligosaccharides, but the size of the oligosaccharide they were able to transport has not been characterized. It is thus possible that MFS could transport glycosides with a DP of more than 4, although further experiments are required to validate this hypothesis.

Table 7. List of the bacterial MFS transporters involved in glycoside utilization, characterized either by binding or transport assays, gene deletion and/or transcriptional approaches.

The glycoside MFS transporters retrieved from the literature but which are not listed in the TCDB were classified based on the identified conserved domain using a BLAST analysis against the NCBI's Conserved Domain Database (Marchler-Bauer et al. 2017).

TCDB Family name	TCID	Protein name (locus tag)	Substrate	Organism	Characterization method	Reference
OHS	2.A.1.5.1	LacY (b0343)	Lactose, melibiose, thio- β -methyl galactopyranoside (TMG), 4-nitrophenyl- β -D-galactopyranoside, 4-nitrophenyl- α -D-galactopyranoside	<i>Escherichia coli</i> K-12	Transport assay Point mutation Structure	(King and Wilson 1990; Abramson et al. 2003; Kaback 2015)
	2.A.1.5.2	RafB (RAFB_ECOL)	Raffinose	<i>Escherichia coli</i> K-12	Transport assay Point mutation	(Camp et al. 2007)
	2.A.1.5.3	CscB	Sucrose, maltose	<i>Escherichia coli</i> EC3132	Transport assay Heterologous expression in another <i>E. coli</i> strain	(Peng et al. 2009)

	2.A.1.5.4	MeY (ECL_05077)	Melibiose, lactose	<i>Enterobacter cloacae</i> ATCC 13047	Gene deletion Point mutation Transport assay	(Shinnick et al. 2003; Tavoulari and Frillingos 2008)
	Literature	FosT (Aec46)	Fructooligosaccharides (DP3-4)	<i>Escherichia coli</i> BEN2908	Gene deletion	(Schouler et al. 2009)
	Literature	ydjE	Fructooligosaccharides	<i>Bifidobacterium longum</i> (SB22) <i>B. pseudolongum</i> (SB12)	Transcriptional analysis	(Gu et al. 2022)
		lacY		<i>Bifidobacterium longum</i> (SB22) <i>B. pseudolongum</i> (SB12) <i>S. alactolyticus</i> (B28)		
	Literature	cscB	Copra meal β -manno- oligosaccharide (CM- β - MOS, DP \geq 3)	<i>Bifidobacterium adolescentis</i> DSMZ 20083	Transport assay Transcriptional analysis Binding test Crystal structure	(Mary et al. 2022)
	Literature	MalT (AZI09_10325)	Maltose	<i>Levilactobacillus brevis</i> TMW 1.2112	Transcriptional analysis	(Bockwoldt et al. 2022)

FHS	2.A.1.7.6	ScrT (Sfri_3989)	Sucrose	<i>Shewanella frigidimarina</i> NCIMB 400	Heterologous expression in <i>E. coli</i>	(Rodionov et al. 2010)
GPH	2.A.2.1.1	MelB (EcE24377A_4674)	Melibiose	<i>Escherichia coli</i> E24377A	Transport assay Binding test Structure	(Ethayathulla et al. 2014)
	2.A.2.2.1	LacS (DID82_ 03005)	Lactose	<i>Streptococcus thermophilus</i> A147	Heterologous expression in <i>E. coli</i> Transport assay	(Poolman et al. 1989)
	2.A.2.3.1	GusB (FAM13_RS00460)	Glucuronide	<i>Escherichia coli</i> K-12	Transcriptional analysis Transport assay	(Liang et al. 2005)
	2.A.2.3.3	XylP (XYLP_ LACP)	Isoprimeverose	<i>Lactobacillus pentosus</i> NZ9000	Transport assay	(Heuberger et al. 2001)
	2.A.2.5.1	TogT	Oligogalacturonides (DP>2)	<i>Erwinia chrysanthemi</i> 3937	Gene deletion	(Hugouvieux-Cotte-Pattat and Reverchon 2001)
	2. A.2.6.2	MalY (CC_2283)	Maltose	<i>Caulobacter crescentus</i> ATCC 19089	Gene deletion Transport assay	(Lohmiller et al. 2008)

	Literature	F5_MFS (CCG34981.1)	Xylooligosaccharides (DP2-4)	Uncultured <i>Bacteroides</i>	Gene deletion	(Tauzin, Kwiatkowski, et al. 2016)
	Literature	LacS (Lba1463)	Galactooligosaccharides (DP2-6), lactose, lactiol	<i>Lactobacillus</i> <i>acidophilus</i> NCFM	Transcriptional analysis Gene deletion	(Andersen et al. 2011)

(3) Taxonomical diversity

The 18 characterized bacterial MFS transporters of glycosides are span three phyla (Pseudomonadota, Bacillota and Bacteroidota) (Table 8). Most of them (57.9 %) belong to Pseudomonadota; 47.4 % of these are from the order Enterobacterales, including the well-known LacY from *Escherichia coli* (Table 8). The identification of LacY might have facilitated the identification of other Enterobacterales MFS, due to the fact that gene transfer or duplication occur more frequently intra-phyla than inter-phyla (Soucy et al. 2015). However, even if most of the MFS retrieved to date belong to this order, these results could be biased as the analyses are based on the well-known LacY sequences from Pseudomonadota, and it cannot be denied that the vast majority of sequences remain hitherto unknown.

Glycoside MFS transporters are present in diverse bacterial phyla from different habitats, including the human gut, food and soil microbiota, suggesting that MFSs play an important role in the survival and adaptation of bacteria to their ecological niches. The phytopathogen *Erwinia chrysanthemi* 3937 (Hugouvieux-Cotte-Pattat and Reverchon 2001) uses a MFS to transport pectin oligomers, a component of plant cell walls, thereby causing soft rot in plants. The fructooligosaccharide MFS transporter of the pathogenic *Escherichia coli* BEN2908 strain is important for colonization of the chicken intestine, indicating a potential risk related to a high level of fructooligosaccharide prebiotic usage in food and feed (Schouler et al. 2009). Nevertheless, glycoside MFSs are also found in strains that are beneficial to human health, such as in the probiotic strains *Streptococcus thermophilus* A147 (Poolman et al. 1989), *Lactobacillus acidophilus* NCFM (Andersen et al. 2011) and *Lactobacillus pentosus* NZ9000 (Heuberger et al. 2001).

(4) Mechanism of glycoside import by MFS transporters

Before any crystal structures were solved, the mutation of some MFS transporters, such as the LacY (King and Wilson 1990), RafB (Camp et al. 2007) and MelB (Granell et al. 2010) proteins, enabled the identification of amino acid residues that play an important role in the recognition and transport of substrates across the membrane.

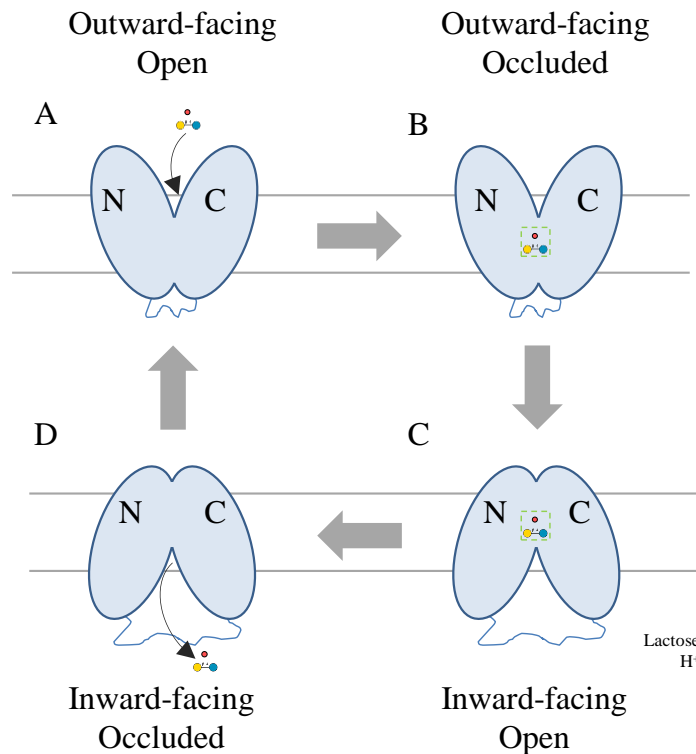


Figure 6. Alternating access mechanism for a H^+ /lactose symport.

This figure shows the transport cycle. It starts with an empty outward-facing transporter which binds the lactose and coupled H^+ from the extracellular solution. The transporter undergoes a conformational transition into an inward-facing state in which the extracellular gate can no longer open while opening of the intracellular gate is enabled. Release of the substrate and ion into the cytoplasm yields an empty inward-facing state, and isomerization back into the outward-facing state completes the cycle.

The first glycoside MFS structure solved was that of lactose permease (LacY) from *E. coli* K-12, in an inward-facing conformation with a bound lactose homolog, β -D-galactopyranosyl-1-thio- β -D-galactopyranoside (Abramson et al. 2003), and in an inward-opening conformation (Guan et al. 2007). More recently, the structure of a second glycoside MFS transporter, the melibiose-specific MelB protein from *Salmonella typhimurium* LT2, has been solved in an outward partially occluded and an outward inactive state (Ethayathulla et al. 2014). These structures enabled the general features of MFS transporters to be described (Figure 6). A canonical MFS fold comprises two domains, the N and C domains, each consisting of six consecutive transmembrane helices. A single substrate-binding cavity is located in the center of the membrane and enclosed by the N and C domains. Moreover, an alternating access

mechanism has been proposed to explain the translocation process (Zhang et al. 2015; Kaback and Guan 2019). Carbohydrate binding is the first step and is the driving force for the alternating access mechanism. The transporters adopt two main open conformations for substrate loading and release in the outward- and inward-facing open conformations, respectively, and alternate between these two conformations through occluded states to translocate substrates across the membrane.

Table 8. Phylum and order distribution of characterized glycoside-targeting MFS transporters in bacteria.

Phylum	Order	Number characterized MFSs	Percentage of MFSs characterized	
			Order	Phylum
Pseudomonadota	Enterobacterales	8	44.4	55.6
	Alteromonadales	1	5.6	
			5.6	
	Caulobacterales	1		
Bacillota	Lactobacillales	4	22.2	22.2
Bacteroidota	Bacteroidales	1	5.6	5.6
Actinomycetota	Bifidobacteriales	3	16.7	16.7

VII. Conclusions

Glycoside transporters play an essential role in bacterial metabolism, enabling the import of various glycosidic substrates required for bacterial survival and growth in diverse environments. This review focused on four transporter superfamilies, namely the TonB-dependent transporters (TBDT), the ATP-binding cassette system (ABC), the phosphotransferase system (PTS) and the major facilitator superfamily (MFS), highlighting their taxonomical, functional, mechanistic and structural diversity.

During the last decade the development of culturing and metagenomics approaches have enabled to access an exponential number of transporters sequences from both cultured and uncultured bacteria. However, experimental validation remains crucial to confirm their predicted functions and fully understand the specific interactions between transporters and substrates. This functional characterization is challenging as glycoside transporters are protein systems embedded in a membrane. They may also require the presence of other protein partners, either for substrate modification or to supply the energy necessary for the transport. As each glycoside transporters superfamilies has its own mechanistic specificities, the development of appropriate methodologies is crucial. Also, each methodology having its own limitations, the structural and functional characterization of such transporters will benefit from multidisciplinary approaches, including cell-based assays, synthetic biology, biophysics and artificial intelligence, a fast-growing field. These new-generation approaches will likely considerably improve our understanding of the transport mechanisms and ecological adaptation of bacteria via their glycoside transport systems.

Finally, glycoside transporters represent a dynamic field of research with important implications for biotechnology and medicine. A thorough understanding of these systems could lead to the development of new strategies for manipulating microbiomes and treating bacterial infections.

VIII. ACKNOWLEDGEMENTS

This research was funded by the European Union's Horizon 2020 framework program (MSCA-IF-2015_707457, CaTSYS, and LEIT-BIO-2015-685474, Metafluidics). Z. W. and X. L. are supported by INSA Toulouse and China Scholarship Council as part of the CSC-UT/INSA Program.

IX. REFERENCES

- Abramson, Jeff, Irina Smirnova, Vladimir Kasho, Gillian Verner, H. Ronald Kaback, and So Iwata. 2003. 'Structure and Mechanism of the Lactose Permease of Escherichia Coli'. *Science* 301 (5633): 610–15. <https://doi.org/10.1126/science.1088196>.
- Andersen, Joakim M., Rodolphe Barrangou, Maher Abou Hachem, Sampo Lahtinen, Yong Jun Goh, Birte Svensson, and Todd R. Klaenhammer. 2011. 'Transcriptional and Functional Analysis of Galactooligosaccharide Uptake by lacS in Lactobacillus Acidophilus'. *Proceedings of the National Academy of Sciences of the United States of America* 108 (43): 17785–90. <https://doi.org/10.1073/pnas.1114152108>.
- Andersen, Joakim Mark, Rodolphe Barrangou, Maher Abou Hachem, Sampo J. Lahtinen, Yong Jun Goh, Birte Svensson, and Todd R. Klaenhammer. 2012. 'Transcriptional Analysis of Prebiotic Uptake and Catabolism by Lactobacillus Acidophilus NCFM'. *PLoS ONE* 7 (9). <https://doi.org/10.1371/journal.pone.0044409>.
- Bågenholm, Viktoria, Sumitha K. Reddy, Hanene Bouraoui, Johan Morrill, Evelina Kulcinskaja, Constance M. Bahr, Oskar Aurelius, et al. 2017. 'Galactomannan Catabolism Conferred by a Polysaccharide Utilization Locus of Bacteroides Ovatus'. *Journal of Biological Chemistry* 292 (1): 229–43. <https://doi.org/10.1074/jbc.M116.746438>.
- Bågenholm, Viktoria, Mathias Wiemann, Sumitha K. Reddy, Abhishek Bhattacharya, Anna Rosengren, Derek T. Logan, and Henrik Stålbrand. 2019. 'A Surface-Exposed GH26 β -Mannanase from Bacteroides Ovatus: Structure, Role, and Phylogenetic Analysis of BoMan26B'. *Journal of Biological Chemistry* 294 (23): 9100–9117. <https://doi.org/10.1074/jbc.RA118.007171>.
- Barbi, Florian, Laurent Vallon, Carmen Guerrero-Galán, Sabine D. Zimmermann, Delphine Melayah, Danis Abrouk, Jeanne Doré, et al. 2021. 'Datamining and Functional Environmental Genomics Reassess the Phylogenetics and Functional Diversity of Fungal Monosaccharide Transporters'. *Applied Microbiology and Biotechnology* 105 (2): 647–60. <https://doi.org/10.1007/s00253-020-11076-y>.
- Barrangou, R., E. Altermann, R. Hutkins, R. Cano, and T. R. Klaenhammer. 2003. 'Functional and Comparative Genomic Analyses of an Operon Involved in Fructooligosaccharide

- Utilization by *Lactobacillus Acidophilus*'. *Proceedings of the National Academy of Sciences* 100 (15): 8957–62. <https://doi.org/10.1073/pnas.1332765100>.
- Beidler, Irena, Nicola Steinke, Tim Schulze, and Chandni Sidhu. 2023. 'Alpha-Glucans from Bacterial Necromass Indicate an Intra-Population Loop within the Marine Carbon Cycle'. 14 August 2023. <https://doi.org/10.21203/rs.3.rs-3205445/v1>.
- Beis, K. 2015. 'Structural Basis for the Mechanism of ABC Transporters'. *Biochemical Society Transactions* 43 (5): 889–93. <https://doi.org/10.1042/bst20150047>.
- Bell, Andrew, Jason Brunt, Emmanuelle Crost, Laura Vaux, Ridvan Nepravishta, C. David Owen, Dimitrios Latousakis, et al. 2019. 'Elucidation of a Sialic Acid Metabolism Pathway in Mucus-Foraging Ruminococcus Gnavus Unravels Mechanisms of Bacterial Adaptation to the Gut'. *Nature Microbiology* 4 (12): 2393–2404. <https://doi.org/10.1038/s41564-019-0590-7>.
- Berntsson, Ronnie P.A., Sander H.J. Smits, Lutz Schmitt, Dirk Jan Slotboom, and Bert Poolman. 2010. 'A Structural Classification of Substrate-Binding Proteins'. *FEBS Letters* 584 (12): 2606–17. <https://doi.org/10.1016/j.febslet.2010.04.043>.
- Bhandari, Pashupati, and Janet E. Hill. 2023. 'Transport and Utilization of Glycogen Breakdown Products by *Gardnerella* Spp. from the Human Vaginal Microbiome'. Edited by Jannell V. Bazurto. *Microbiology Spectrum*, March, e04435-22. <https://doi.org/10.1128/spectrum.04435-22>.
- Bi, Yunchen, Evan Mann, Chris Whitfield, and Jochen Zimmer. 2018. 'Architecture of a Channel-Forming O-Antigen Polysaccharide ABC Transporter'. *Nature* 553 (7688): 361–65. <https://doi.org/10.1038/nature25190>.
- Bidart, Gonzalo N., Hani Gharabli, and Ditte Heddam Welner. 2023. 'Functional Characterization of the Phosphotransferase System in *Parageobacillus Thermoglucosidasius*'. *Scientific Reports* 13 (May):7131. <https://doi.org/10.1038/s41598-023-33918-1>.
- Bidart, Gonzalo N., Jesús Rodríguez-Díaz, Vicente Monedero, and María J. Yebra. 2014. 'A Unique Gene Cluster for the Utilization of the Mucosal and Human Milk-Associated Glycans Galacto-N-Biose and Lacto-N-Biose in *Lactobacillus Casei*'. *Molecular Microbiology* 93 (3): 521–38. <https://doi.org/10.1111/mmi.12678>.

- Bjelić, Saša, and Ilian Jelesarov. 2008. 'A Survey of the Year 2007 Literature on Applications of Isothermal Titration Calorimetry'. *Journal of Molecular Recognition* 21 (5): 289–311. <https://doi.org/10.1002/jmr.909>.
- Bjursell, Magnus K., Eric C. Martens, and Jeffrey I. Gordon. 2006. 'Functional Genomic and Metabolic Studies of the Adaptations of a Prominent Adult Human Gut Symbiont, *Bacteroides Thetaiotaomicron*, to the Suckling Period'. *Journal of Biological Chemistry* 281 (47): 36269–79. <https://doi.org/10.1074/jbc.M606509200>.
- Blanvillain, Servane, Damien Meyer, Alice Boulanger, Martine Lautier, Catherine Guynet, Nicolas Denancé, Jacques Vasse, Emmanuelle Lauber, and Matthieu Arlat. 2007. 'Plant Carbohydrate Scavenging through TonB-Dependent Receptors: A Feature Shared by Phytopathogenic and Aquatic Bacteria'. *PLoS ONE* 2 (2): e224. <https://doi.org/10.1371/journal.pone.0000224>.
- Bockwoldt, Julia A., Chen Meng, Christina Ludwig, Michael Kupetz, and Matthias A. Ehrmann. 2022. 'Proteomic Analysis Reveals Enzymes for β -D-Glucan Formation and Degradation in *Levilactobacillus Brevis* TMW 1.2112'. *International Journal of Molecular Sciences* 23 (6): 3393. <https://doi.org/10.3390/ijms23063393>.
- Böger, Markus, Sander S. van Leeuwen, Alicia Lammerts van Bueren, and Lubbert Dijkhuizen. 2019. 'Structural Identity of Galactooligosaccharide Molecules Selectively Utilized by Single Cultures of Probiotic Bacterial Strains'. *Journal of Agricultural and Food Chemistry* 67 (50): 13969–77. <https://doi.org/10.1021/acs.jafc.9b05968>.
- Bolam, David N., and Bert van den Berg. 2018. 'TonB-Dependent Transport by the Gut Microbiota: Novel Aspects of an Old Problem'. *Current Opinion in Structural Biology* 51 (Figure 1): 35–43. <https://doi.org/10.1016/j.sbi.2018.03.001>.
- Boucher, Nathalie, and Kenneth M. Noll. 2011. 'Ligands of Thermophilic ABC Transporters Encoded in a Newly Sequenced Genomic Region of *Thermotoga Maritima* MSB8 Screened by Differential Scanning Fluorimetry'. *Applied and Environmental Microbiology* 77 (18): 6395–99. <https://doi.org/10.1128/AEM.05418-11>.
- Braza, Rezia Era, Aliyah B. Silver, Ganesh S. Sundar, and Sarah E. Davis. 2020. 'Phosphotransferase System Uptake and Metabolism of the β -Glucoside Salicin Impact

- Group A Streptococcal Bloodstream Survival and Soft Tissue Infection | Infection and Immunity'. 2020. <https://journals.asm.org/doi/10.1128/iai.00346-20>.
- Büchel, Dagmar E., Bruno Gronenborn, and Benno Müller-Hill. 1980. 'Sequence of the Lactose Permease Gene'. *Nature* 283 (5747): 541–45. <https://doi.org/10.1038/283541a0>.
- Buntin, Nirunya, Tipparat Hongpattarakere, Jarmo Ritari, François P. Douillard, Lars Paulin, Sijef Boeren, Sudarshan A. Shetty, and Willem M. De Vos. 2017. 'An Inducible Operon Is Involved in Inulin Utilization in *Lactobacillus Plantarum* Strains, as Revealed by Comparative Proteogenomics and Metabolic Profiling'. Edited by Christopher A. Elkins. *Applied and Environmental Microbiology* 83 (2): e02402-16. <https://doi.org/10.1128/AEM.02402-16>.
- Buzun, Ekaterina, Tiaan Heunis, Curtis Cottam, Carl Morland, Matthias Trost, and Elisabeth C Lowe. 2022. 'The Type 9 Secretion System Enables Sharing of Fungal Mannan by Human Gut *Bacteroides*'. Preprint. Microbiology. <https://doi.org/10.1101/2022.07.15.500217>.
- Cameron, Elizabeth A, Kurt J Kwiatkowski, Byung-hoo Lee, Bruce R Hamaker, Nicole M Koropatkin, and C Martens. 2014. 'Multifunctional Nutrient-Binding Proteins Adapt Human Symbiotic Enhanced Sensing and Catalysis'. *mBio* 5 (5): e01441-14. <https://doi.org/10.1128/mBio.01441-14>.
- Camp, Bonnie M. Van, Robert R. Crow, Yang Peng, and Manuel F. Varela. 2007. 'Amino Acids That Confer Transport of Raffinose and Maltose Sugars in the Raffinose Permease (RafB) of *Escherichia Coli* as Implicated by Spontaneous Mutations at Val-35, Ser-138, Ser-139, Gly-389 and Ile-391'. *Bone* 23 (1): 1–7. <https://doi.org/10.1038/jid.2014.371>.
- Cao, Thanh Nguyen, Philippe Joyet, Francine Moussan Désirée Aké, Eliane Milohanic, and Josef Deutscher. 2019. 'Studies of the *Listeria Monocytogenes* Cellobiose Transport Components and Their Impact on Virulence Gene Repression'. *Microbial Physiology* 29 (1–6): 10–26. <https://doi.org/10.1159/000500090>.
- Cao, Yu, Xiangshu Jin, Elena J Levin, Hua Huang, Yinong Zong, Matthias Quick, Jonathan A Javitch, and Kanagalaghatta R Rajashankar. 2011. 'Crystal Structure of a

- Phosphorylation-Coupled Saccharide Transporter'. *Nature* 473 (7345): 50–54. <https://doi.org/10.1038/nature09939>.
- Cartmell, Alan, Elisabeth C. Lowe, Arnaud Baslé, Susan J. Firbank, Didier A. Ndeh, Heath Murray, Nicolas Terrapon, et al. 2017. 'How Members of the Human Gut Microbiota Overcome the Sulfation Problem Posed by Glycosaminoglycans'. *Proceedings of the National Academy of Sciences* 114 (27): 7037–42. <https://doi.org/10.1073/pnas.1704367114>.
- Cartmell, Alan, Jose Muñoz-Muñoz, Jonathon A. Briggs, Didier A. Ndeh, Elisabeth C. Lowe, Arnaud Baslé, Nicolas Terrapon, et al. 2018. 'A Surface Endogalactanase in *Bacteroides Thetaiotaomicron* Confers Keystone Status for Arabinogalactan Degradation'. *Nature Microbiology* 3 (11): 1314–26. <https://doi.org/10.1038/s41564-018-0258-8>.
- Cassio Barreto de Oliveira, Marcelo, and Andrea Balan. 2020. 'The ATP-Binding Cassette (ABC) Transport Systems in Mycobacterium Tuberculosis: Structure, Function, and Possible Targets for Therapeutics'. *Biology* 9 (12): 443. <https://doi.org/10.3390/biology9120443>.
- Cerisy, Tristan, Alba Iglesias, William Rostain, Magali Boutard, Christine Pelle, Alain Perret, Marcel Salanoubat, Henri-Pierre Fierobe, and Andrew C. Tolonen. 2019. 'ABC Transporters Required for Hexose Uptake by *Clostridium Phytofermentans*'. Edited by William W. Metcalf. *Journal of Bacteriology* 201 (15). <https://doi.org/10.1128/JB.00241-19>.
- Chandravanshi, Monika, Reshama Samanta, and Shankar Prasad Kanaujia. 2020. 'Conformational Trapping of a β -Glucosides-Binding Protein Unveils the Selective Two-Step Ligand-Binding Mechanism of ABC Importers'. *Journal of Molecular Biology* 432 (20): 5711–34. <https://doi.org/10.1016/j.jmb.2020.08.021>.
- Chandravanshi, Monika, Anjaney Sharma, Pratik Dasgupta, Suraj Kumar Mandal, and Shankar Prasad Kanaujia. 2019. 'Identification and Characterization of ABC Transporters for Carbohydrate Uptake in *Thermus Thermophilus* HB8'. *Gene* 696 (May):135–48. <https://doi.org/10.1016/j.gene.2019.02.035>.
- Chandravanshi, Monika, Anjaney Sharma, Pratik Das Gupta, Suraj Kumar Mandal, and Shankar Prasad Kanaujia. 2019. 'Identification and Characterization of ABC

- Transporters for Carbohydrate Uptake in *Thermus Thermophilus* HB8'. *Gene* 696 (November 2018): 135–48. <https://doi.org/10.1016/j.gene.2019.02.035>.
- Chekan, Jonathan R., In Hyuk Kwon, Vinayak Agarwal, Dylan Dodd, Vanessa Revindran, Roderick I. Mackie, Isaac Cann, and Satish K. Nair. 2014. 'Structural and Biochemical Basis for Mannan Utilization by *Caldanaerobius Polysaccharolyticus* Strain ATCC BAA-17'. *Journal of Biological Chemistry* 289 (50): 34965–77. <https://doi.org/10.1074/jbc.M114.579904>.
- Chen, Chen, Guozhong Zhao, Wei Chen, and Benheng Guo. 2015. 'Metabolism of Fructooligosaccharides in *Lactobacillus Plantarum* ST-III via Differential Gene Transcription and Alteration of Cell Membrane Fluidity'. *Applied and Environmental Microbiology* 81 (22): 7697–7707. <https://doi.org/10.1128/AEM.02426-15>.
- Chen, Jonathan S., Vamsee Reddy, Joshua H. Chen, Maksim A. Shlykov, Wei Hao Zheng, Jaehoon Cho, Ming Ren Yen, and Milton H. Saier. 2012. 'Phylogenetic Characterization of Transport Protein Superfamilies: Superiority of SuperfamilyTree Programs over Those Based on Multiple Alignments'. *Journal of Molecular Microbiology and Biotechnology* 21 (3–4): 83–96. <https://doi.org/10.1159/000334611>.
- Cho, K. H., and A. A. Salyers. 2001. 'Biochemical Analysis of Interactions between Outer Membrane Proteins That Contribute to Starch Utilization by *Bacteroides Thetaiotaomicron*'. *Journal of Bacteriology* 183 (24): 7224–30. <https://doi.org/10.1128/JB.183.24.7224-7230.2001>.
- Cockburn, Darrell W., and Nicole M. Koropatkin. 2016. 'Polysaccharide Degradation by the Intestinal Microbiota and Its Influence on Human Health and Disease'. *Journal of Molecular Biology* 428 (16): 3230–52. <https://doi.org/10.1016/j.jmb.2016.06.021>.
- Cockburn, Darrell W., Nicole I. Orlovsky, Matthew H. Foley, Kurt J. Kwiatkowski, Constance M. Bahr, Mallory Maynard, Borries Demeler, and Nicole M. Koropatkin. 2015. 'Molecular Details of a Starch Utilization Pathway in the Human Gut Symbiont Eubacterium Rectale'. *Molecular Microbiology* 95 (2): 209–30. <https://doi.org/10.1111/mmi.12859>.
- Connors, S B, C I Montero, D A Comfort, K R Shockley, M R Johnson, S R Chhabra, and R M Kelly. 2005. 'Prediction of Carbohydrate Transport and Utilization Regulons in the

- Hyperthermophilic Bacterium *Thermotoga Maritima* through the Use of Carbohydrate-Specific Transcriptional Response'. *Abstracts of Papers of the American Chemical Society* 229 (21): U240–U240. <https://doi.org/10.1128/JB.187.21.7267>.
- Costa, Devika M. De, Katsunori Suzuki, and Kazuo Yoshida. 2003. 'Structural and Functional Analysis of a Putative Gene Cluster for Palatinose Transport on the Linear Chromosome of *Agrobacterium Tumefaciens* MAFF301001'. *Journal of Bacteriology* 185 (7): 2369–73. <https://doi.org/10.1128/JB.185.7.2369-2373.2003>.
- Cote, Christopher K., Dennis Cvitkovitch, Arnold S. Bleiweis, and Allen L. Honeyman. 2000. 'A Novel β -Glucoside-Specific PTS Locus from *Streptococcus Mutans* That Is Not Inhibited by Glucose'. *Microbiology* 146 (7): 1555–63. <https://doi.org/10.1099/00221287-146-7-1555>.
- Crouch, Lucy I., Marcelo V. Liberato, Paulina A. Urbanowicz, Arnaud Baslé, Christopher A. Lamb, Christopher J. Stewart, Katie Cooke, et al. 2020. 'Prominent Members of the Human Gut Microbiota Express Endo-Acting O-Glycanases to Initiate Mucin Breakdown'. *Nature Communications* 11 (1): 4017. <https://doi.org/10.1038/s41467-020-17847-5>.
- Culurgioni, Simone, Gemma Harris, Anirudh K. Singh, Samantha J. King, and Martin A. Walsh. 2017. 'Structural Basis for Regulation and Specificity of Fructooligosaccharide Import in *Streptococcus Pneumoniae*'. *Structure* 25 (1): 79–93. <https://doi.org/10.1016/j.str.2016.11.008>.
- Cummings, J. H., and A. M. Stephen. 2007. 'Carbohydrate Terminology and Classification'. *European Journal of Clinical Nutrition* 61:S5–18. <https://doi.org/10.1038/sj.ejcn.1602936>.
- Cuskin, Fiona, Elisabeth C. Lowe, Max J. Temple, Yanping Zhu, Elizabeth A. Cameron, Nicholas A. Pudlo, Nathan T. Porter, Karthik Urs, Andrew J. Thompson, Alan Cartmell, Artur Rogowski, Brian S. Hamilton, Rui Chen, Thomas J. Tolbert, Kathleen Piens, Debby Bracke, Wouter Vervecken, Zalihe Hakki, Gaetano Speciale, Jose L. Munõz-Munõz, et al. 2015. 'Human Gut Bacteroidetes Can Utilize Yeast Mannan through a Selfish Mechanism'. *Nature* 517 (7533): 165–69. <https://doi.org/10.1038/nature13995>.

- Cuskin, Fiona, Elisabeth C. Lowe, Max J. Temple, Yanping Zhu, Elizabeth A. Cameron, Nicholas A. Pudlo, Nathan T. Porter, Karthik Urs, Andrew J. Thompson, Alan Cartmell, Artur Rogowski, Brian S. Hamilton, Rui Chen, Thomas J. Tolbert, Kathleen Piens, Debby Bracke, Wouter Vervecken, Zalihe Hakki, Gaetano Speciale, Jose L. Munoz-Munoz, et al. 2015. 'Human Gut Bacteroidetes Can Utilize Yeast Mannan through a Selfish Mechanism'. *Nature* 517 (7533): 165–69. <https://doi.org/10.1038/nature13995>.
- Dagkesamanskaya, Adilya, Krzysztof Langer, Alexandra S. Tauzin, Catherine Rouzeau, Delphine Lestrade, Gabrielle Potocki-Veronese, Laurent Boitard, et al. 2018. 'Use of Photoswitchable Fluorescent Proteins for Droplet-Based Microfluidic Screening'. *Journal of Microbiological Methods* 147 (February): 59–65. <https://doi.org/10.1016/j.mimet.2018.03.001>.
- De la Torre, Lilia I., José G. Vergara Meza, Sindy Cabarca, André G. Costa-Martins, and Andrea Balan. 2021. 'Comparison of Carbohydrate ABC Importers from Mycobacterium Tuberculosis'. *BMC Genomics* 22 (1): 841. <https://doi.org/10.1186/s12864-021-07972-w>.
- Déjean, Guillaume, Servane Blanvillain-Baufumé, Alice Boulanger, Armelle Darrasse, Thomas Dugé De Bernonville, Anne Laure Girard, Sébastien Carrère, et al. 2013. 'The Xylan Utilization System of the Plant Pathogen Xanthomonas Campestris Pv Campestris Controls Epiphytic Life and Reveals Common Features with Oligotrophic Bacteria and Animal Gut Symbionts'. *New Phytologist* 198 (3): 899–915. <https://doi.org/10.1111/nph.12187>.
- Déjean, Guillaume, Kazune Tamura, Adriana Cabrera, Namrata Jain, Nicholas A. Pudlo, Gabriel Pereira, Alexander Holm Viborg, Filip Van Petegem, Eric C. Martens, and Harry Brumer. 2020. 'Synergy between Cell Surface Glycosidases and Glycan-Binding Proteins Dictates the Utilization of Specific Beta(1,3)-Glucans by Human Gut Bacteroides'. *mBio* 11 (2): 10.1128/mbio.00095-20. <https://doi.org/10.1128/mbio.00095-20>.
- Déjean, Guillaume, Alexandra S. Tauzin, Stuart W. Bennett, A. Louise Creagh, and Harry Brumer. 2019. *Adaptation of Syntenic Xyloglucan Utilization Loci of Human Gut Bacteroidetes to Polysaccharide Side Chain Diversity. Applied and Environmental Microbiology*. Vol. 85. <https://doi.org/10.1128/AEM.01491-19>.

- D'Elia, John N., and Abigail A. Salyers. 1996. 'Contribution of a Neopullulanase, a Pullulanase, and an α -Glucosidase to Growth of *Bacteroides Thetaiotaomicron* on Starch'. *Journal of Bacteriology* 178 (24): 7173–79. <https://doi.org/10.1128/jb.178.24.7173-7179.1996>.
- Desai, Stuti K., Krithi Nandimath, and S. Mahadevan. 2010. 'Diverse Pathways for Salicin Utilization in *Shigella Sonnei* and *Escherichia Coli* Carrying an Impaired Bgl Operon'. *Archives of Microbiology* 192 (10): 821–33. <https://doi.org/10.1007/s00203-010-0610-8>.
- Despres, Jordane, Evelyne Forano, Pascale Lepercq, Sophie Comtet-Marre, Gregory Jubelin, Christophe Chambon, Carl J. Yeoman, et al. 2016. 'Xylan Degradation by the Human Gut *Bacteroides Xylanisolvens* XB1AT Involves Two Distinct Gene Clusters That Are Linked at the Transcriptional Level'. *BMC Genomics* 17 (1): 1–14. <https://doi.org/10.1186/s12864-016-2680-8>.
- Deutscher, J., C. Francke, and P. W. Postma. 2006. 'How Phosphotransferase System-Related Protein Phosphorylation Regulates Carbohydrate Metabolism in Bacteria'. *Microbiology and Molecular Biology Reviews* 70 (4): 939–1031. <https://doi.org/10.1128/MMBR.00024-06>.
- Deutscher, Josef, Francine Moussan Désirée Aké, Meriem Derkaoui, Arthur Constant Zébré, Thanh Nguyen Cao, Houda Bouraoui, Takfarinas Kentache, Abdelhamid Mokhtari, Eliane Milohanic, and Philippe Joyet. 2014. 'The Bacterial Phosphoenolpyruvate:Carbohydrate Phosphotransferase System: Regulation by Protein Phosphorylation and Phosphorylation-Dependent Protein-Protein Interactions'. *Microbiology and Molecular Biology Reviews* 78 (2): 231–56. <https://doi.org/10.1128/MMBR.00001-14>.
- Diallinas, George. 2014. 'Understanding Transporter Specificity and the Discrete Appearance of Channel-like Gating Domains in Transporters'. *Frontiers in Pharmacology* 5 AUG (September): 1–17. <https://doi.org/10.3389/fphar.2014.00207>.
- Duhsaki, Lal, Saumashish Mukherjee, Tirupaati Swaroopa Rani, and Jogi Madhuprakash. 2022. 'Genome Analysis of *Streptomyces* Sp. UH6 Revealed the Presence of Potential Chitinolytic Machinery Crucial for Chitosan Production'. *Environmental Microbiology Reports* 14 (3): 431–42. <https://doi.org/10.1111/1758-2229.12986>.

- Dutschei, Theresa, Irena Beidler, Daniel Bartosik, Julia-Maria Seeßelberg, Michelle Teune, Marcus Bäumgen, Soraia Querido Ferreira, et al. 2023. 'Marine Bacteroidetes Enzymatically Digest Xylans from Terrestrial Plants'. *Environmental Microbiology* n/a (n/a). <https://doi.org/10.1111/1462-2920.16390>.
- Ejby, M., A. Guskov, M. J. Pichler, G. C. Zanten, E. Schoof, W. Saburi, D. J. Slotboom, and M. Abou Hachem. 2019. 'Two Binding Proteins of the ABC Transporter That Confers Growth of *Bifidobacterium Animalis* Subsp. Lactis ATCC27673 on β -Mannan Possess Distinct Manno-Oligosaccharide-Binding Profiles'. *Molecular Microbiology* 112 (1): 114–30. <https://doi.org/10.1111/mmi.14257>.
- Ejby, Morten, Folmer Fredslund, Joakim Mark Andersen, Andreja Vujičić Žagar, Jonas Rosager Henriksen, Thomas Lars Andersen, Birte Svensson, Dirk Jan Slotboom, and Maher Abou Hachem. 2016. 'An Atp Binding Cassette Transporter Mediates the Uptake of α -(1,6)-Linked Dietary Oligosaccharides in *Bifidobacterium* and Correlates with Competitive Growth on These Substrates'. *Journal of Biological Chemistry* 291 (38): 20220–31. <https://doi.org/10.1074/jbc.M116.746529>.
- Ejby, Morten, Folmer Fredslund, Andreja Vujicic-Zagar, Birte Svensson, Dirk Jan Slotboom, and Maher Abou Hachem. 2013. 'Structural Basis for Arabinoxylo-Oligosaccharide Capture by the Probiotic *Bifidobacterium Animalis* Subsp. lactisBI-04'. *Molecular Microbiology* 90 (5): 1100–1112. <https://doi.org/10.1111/mmi.12419>.
- El-Gebali, Sara, Jaina Mistry, Alex Bateman, Sean R. Eddy, Aurélien Luciani, Simon C. Potter, Matloob Qureshi, et al. 2019. 'The Pfam Protein Families Database in 2019'. *Nucleic Acids Research* 47 (D1): D427–32. <https://doi.org/10.1093/nar/gky995>.
- Erkens, Guus B, Maria Majsnerowska, Josy ter Beek, and Dirk Jan Slotboom. 2012. 'Energy Coupling Factor-Type ABC Transporters for Vitamin Uptake in Prokaryotes'. *Biochemistry* 51:4390–96. <https://doi.org/10.1021/bi300504v>.
- Erni, Bernhard. 2013. 'The Bacterial Phosphoenolpyruvate: Sugar Phosphotransferase System (PTS): An Interface between Energy and Signal Transduction'. *Journal of the Iranian Chemical Society* 10 (3): 593–630. <https://doi.org/10.1007/s13738-012-0185-1>.
- Ethayathulla, Abdul S., Mohammad S. Yousef, Anowarul Amin, Gérard Leblanc, H. Ronald Kaback, and Lan Guan. 2014. 'Structure-Based Mechanism for Na(+)/Melibiose

- Symport by MelB'. *Nature Communications* 5:3009. <https://doi.org/10.1038/ncomms4009>.
- Ferreira, Mário J., Aristides L. Mendes, and Isabel de Sá-Nogueira. 2017. 'The MsmX ATPase Plays a Crucial Role in Pectin Mobilization by *Bacillus Subtilis*'. Edited by Hernâni Gerós. *PLOS ONE* 12 (12): e0189483. <https://doi.org/10.1371/journal.pone.0189483>.
- Ferreira, Mário José, and Isabel De Sá-Nogueira. 2010. 'A Multitask ATPase Serving Different ABC-Type Sugar Importers in *Bacillus Subtilis*'. *Journal of Bacteriology* 192 (20): 5312–18. <https://doi.org/10.1128/JB.00832-10>.
- Ficko-Blean, Elizabeth, Aurélie Préchoux, François Thomas, Tatiana Rochat, Robert Larocque, Yongtao Zhu, Mark Stam, et al. 2017. 'Carrageenan Catabolism Is Encoded by a Complex Regulon in Marine Heterotrophic Bacteria'. *Nature Communications* 8 (1): 1685. <https://doi.org/10.1038/s41467-017-01832-6>.
- Flint, Harry J., Edward A. Bayer, Marco T. Rincon, Raphael Lamed, and Bryan A. White. 2008. 'Polysaccharide Utilization by Gut Bacteria: Potential for New Insights from Genomic Analysis'. *Nature Reviews Microbiology* 6 (2): 121–31. <https://doi.org/10.1038/nrmicro1817>.
- Foley, Matthew H., Darrell W. Cockburn, and Nicole M. Koropatkin. 2016. 'The Sus Operon: A Model System for Starch Uptake by the Human Gut Bacteroidetes'. *Cellular and Molecular Life Sciences* 73 (14): 2603–17. <https://doi.org/10.1007/s00018-016-2242-x>.
- Foley, Matthew H., Guillaume Déjean, Glyn R. Hemsworth, Gideon J. Davies, Harry Brumer, and Nicole M. Koropatkin. 2019. 'A Cell-Surface GH9 Endo-Glucanase Coordinates with Surface Glycan-Binding Proteins to Mediate Xyloglucan Uptake in the Gut Symbiont *Bacteroides Ovatus*'. *Journal of Molecular Biology* 431 (5): 981–95. <https://doi.org/10.1016/j.jmb.2019.01.008>.
- Ford, Robert C., and Konstantinos Beis. 2019. 'Learning the ABCs One at a Time: Structure and Mechanism of ABC Transporters'. *Biochemical Society Transactions* 47 (1): 23–36. <https://doi.org/10.1042/bst20180147>.
- Fosses, Aurélie, Romain Borne, Nathalie Franche, Pascale de Philip, Henri-Pierre Fierobe, Maria Maté, Stéphanie Perret, Nian Liu, and Yann Denis. 2017. 'A Seven-Gene Cluster in *Ruminiclostridium Cellulolyticum* Is Essential for Signalization, Uptake and

- Catabolism of the Degradation Products of Cellulose Hydrolysis'. *Biotechnology for Biofuels* 10 (1): 1–14. <https://doi.org/10.1186/s13068-017-0933-7>.
- Fouet, A., M. Arnaud, A. Klier, and G. Rapoport. 1987. 'Bacillus Subtilis Sucrose-Specific Enzyme II of the Phosphotransferase System: Expression in Escherichia Coli and Homology to Enzymes II from Enteric Bacteria.' *Proceedings of the National Academy of Sciences of the United States of America* 84 (24): 8773–77. <https://doi.org/10.1073/pnas.84.24.8773>.
- Francel, Alyssa L., Jennifer L. Hoeflinger, and Michael J. Miller. 2012. 'Identification of Lactose Phosphotransferase Systems in Lactobacillus Gasseri ATCC 33323 Required for Lactose Utilization'. *Microbiology* 158 (4): 944–52. <https://doi.org/10.1099/mic.0.052928-0>.
- Francel, Alyssa L., Taksawan Thongaram, and Michael J. Miller. 2010. 'The PTS Transporters of Lactobacillus Gasseri ATCC 33323'. *BMC Microbiology* 10. <https://doi.org/10.1186/1471-2180-10-77>.
- Furze, Christopher M., Ignacio Delso, Enriqueta Casal, Collette S. Guy, Chloe Seddon, Chelsea M. Brown, Hadyn L. Parker, et al. 2021. 'Structural Basis of Trehalose Recognition by the Mycobacterial LpqY-SugABC Transporter'. *Journal of Biological Chemistry* 296 (January):100307. <https://doi.org/10.1016/j.jbc.2021.100307>.
- Fushinobu, Shinya, and Maher Abou Hachem. 2021. 'Structure and Evolution of the Bifidobacterial Carbohydrate Metabolism Proteins and Enzymes'. *Biochemical Society Transactions* 49 (2): 563–78. <https://doi.org/10.1042/BST20200163>.
- Gasser, C., P. Garault, C. Chervaux, V. Monnet, J.-M. Faurie, and F. Rul. 2022. 'Co-Utilization of Saccharides in Mixtures: Moving toward a New Understanding of Carbon Metabolism in Streptococcus Thermophilus'. *Food Microbiology* 107 (October):104080. <https://doi.org/10.1016/j.fm.2022.104080>.
- Gelfand, Mikhail S., and Dmitry A. Rodionov. 2008. 'Comparative Genomics and Functional Annotation of Bacterial Transporters'. *Physics of Life Reviews* 5 (1): 22–49. <https://doi.org/10.1016/j.plrev.2007.10.003>.
- Genee, Hans J., Anne P. Bali, Søren D. Petersen, Solvej Siedler, Mads T. Bonde, Luisa S. Gronenberg, Mette Kristensen, Scott J. Harrison, and Morten O.A. Sommer. 2016.

- ‘Functional Mining of Transporters Using Synthetic Selections’. *Nature Chemical Biology* 12 (12): 1015–22. <https://doi.org/10.1038/nchembio.2189>.
- Glenwright, Amy J., Karunakar R. Pothula, Satya P. Bhamidimarri, Dror S. Chorev, Arnaud Baslé, Susan J. Firbank, Hongjun Zheng, et al. 2017. ‘Structural Basis for Nutrient Acquisition by Dominant Members of the Human Gut Microbiota’. *Nature* 541 (7637): 407–11. <https://doi.org/10.1038/nature20828>.
- Goh, Yong Jun, and Todd R. Klaenhammer. 2015. ‘Genetic Mechanisms of Prebiotic Oligosaccharide Metabolism in Probiotic Microbes’. *Annual Review of Food Science and Technology* 6 (1): 137–56. <https://doi.org/10.1146/annurev-food-022814-015706>.
- Grand, Maxime, Marion Aubourg, Andreas Pikis, John Thompson, Josef Deutscher, Axel Hartke, and Nicolas Sauvageot. 2019. ‘Characterization of the Gen Locus Involved in β -1,6-Oligosaccharide Utilization by *Enterococcus Faecalis*’. *Molecular Microbiology* 112 (6): 1744–56. <https://doi.org/10.1111/mmi.14390>.
- Granell, Meritxell, Xavier León, Gérard Leblanc, Esteve Padrós, and Víctor A. Lórenz-Fonfría. 2010. ‘Structural Insights into the Activation Mechanism of Melibiose Permease by Sodium Binding’. *Proceedings of the National Academy of Sciences of the United States of America* 107 (51): 22078–83. <https://doi.org/10.1073/pnas.1008649107>.
- Gray, Declan A., Joshua B. R. White, Abraham O. Oluwole, Parthasarathi Rath, Amy J. Glenwright, Adam Mazur, Michael Zahn, et al. 2021. ‘Insights into SusCD-Mediated Glycan Import by a Prominent Gut Symbiont’. *Nature Communications* 12 (1): 44. <https://doi.org/10.1038/s41467-020-20285-y>.
- Grøftehauge, Morten K., Nelly R. Hajizadeh, Marcus J. Swann, and Ehmke Pohl. 2015. ‘Protein-Ligand Interactions Investigated by Thermal Shift Assays (TSA) and Dual Polarization Interferometry (DPI)’. *Acta Crystallographica Section D: Biological Crystallography* 71:36–44. <https://doi.org/10.1107/S1399004714016617>.
- Grondin, Julie M., Guillaume Déjean, Filip Van Petegem, and Harry Brumer. 2022. ‘Cell Surface Xyloglucan Recognition and Hydrolysis by the Human Gut Commensal *Bacteroides Uniformis*’. *Applied and Environmental Microbiology* 88 (1): e01566-21. <https://doi.org/10.1128/AEM.01566-21>.

- Grondin, Julie M., Kazune Tamura, Guillaume Déjean, D. Wade Abbott, and Harry Brumer. 2017. 'Polysaccharide Utilization Loci: Fueling Microbial Communities'. Edited by George O'Toole. *Journal of Bacteriology* 199 (15). <https://doi.org/10.1128/JB.00860-16>.
- Gu, Jiayu, Bingyong Mao, Shumao Cui, Xin Tang, Zhenmin Liu, Jianxin Zhao, and Hao Zhang. 2022. 'Bifidobacteria Exhibited Stronger Ability to Utilize Fructooligosaccharides, Compared with Other Bacteria in the Mouse Intestine'. *Journal of the Science of Food and Agriculture* 102 (6): 2413–23. <https://doi.org/10.1002/jsfa.11580>.
- Guan, Lan, Osman Mirza, Gillian Verner, So Iwata, and H. Ronald Kaback. 2007. 'Structural Determination of Wild-Type Lactose Permease'. *Proceedings of the National Academy of Sciences of the United States of America* 104 (39): 15294–98. <https://doi.org/10.1073/pnas.0707688104>.
- Han, Yejun, Vinayak Agarwal, Dylan Dodds, Jason Kims, Brian Bae, Roderick I. Mackie, Satish K. Nair, and Isaac K O Cann. 2012. 'Biochemical and Structural Insights into Xylan Utilization by the Thermophilic Bacterium *Caldanaerobius Polysaccharolyticus*'. *Journal of Biological Chemistry* 287 (42): 34946–60. <https://doi.org/10.1074/jbc.M112.391532>.
- Hehemann, Jan-Hendrik, Amelia G. Kelly, Nicholas A. Pudlo, Eric C. Martens, and Alisdair B. Boraston. 2012. 'Bacteria of the Human Gut Microbiome Catabolize Red Seaweed Glycans with Carbohydrate-Active Enzyme Updates from Extrinsic Microbes'. *Proceedings of the National Academy of Sciences* 109 (48): 19786–91. <https://doi.org/10.1073/pnas.1211002109>.
- Henrich, Alexander, Nora Kuhlmann, Alexander W. Eck, Reinhard Krämer, and Gerd M. Seibold. 2013. 'Maltose Uptake by the Novel ABC Transport System MusEFGK2I Causes Increased Expression of ptsG in *Corynebacterium Glutamicum*'. *Journal of Bacteriology* 195 (11): 2573–84. <https://doi.org/10.1128/JB.01629-12>.
- Heuberger, Esther H.M.L., Elly Smits, and Bert Poolman. 2001. 'Xyloside Transport by XylP, a Member of the Galactoside-Pentoside-Hexuronide Family'. *Journal of Biological Chemistry* 276 (37): 34465–72. <https://doi.org/10.1074/jbc.M105460200>.

- Holland, I. Barry. 2019. 'Rise and Rise of the ABC Transporter Families'. *Research in Microbiology*, ABC systems in microorganisms, 170 (8): 304–20. <https://doi.org/10.1016/j.resmic.2019.08.004>.
- Hospital, The Heart, Rei Fidalgo, Reconstructive Surgery, and Perioperative Care. 2010. *The Maltose ABC Transporter in the 21st Century – towards a Structural-Dynamic Perspective on Its Mode of Action. Annals of Neurology*.
- Hugouvieux-Cotte-Pattat, Nicole, Nicolas Blot, and Sylvie Reverchon. 2001. 'Identification of TogMNAB, an ABC Transporter Which Mediates the Uptake of Pectic Oligomers in *Erwinia Chrysanthemi* 3937'. *Molecular Microbiology* 41 (5): 1113–23. <https://doi.org/10.1046/j.1365-2958.2001.02564.x>.
- Hugouvieux-Cotte-Pattat, Nicole, and Sylvie Reverchon. 2001. 'Two Transporters, TogT and TogMNAB, Are Responsible for Oligogalacturonide Uptake in *Erwinia Chrysanthemi* 3937'. *Molecular Microbiology* 41 (5): 1125–32. <https://doi.org/10.1046/j.1365-2958.2001.02565.x>.
- Imai, K., and B. G. Hall. 1981. 'Properties of the Lactose Transport System in *Klebsiella* Sp. Strain CT-1'. *Journal of Bacteriology* 145 (3): 1459–62.
- Itoh, Takafumi, Misaki Yaguchi, Akari Nakaichi, Moe Yoda, Takao Hibi, and Hisashi Kimoto. 2021. 'Structural Characterization of Two Solute-Binding Proteins for N,N'-Diacetylchitobiose/N,N',N''-Triacetylchitotriose of the Gram-Positive Bacterium, *Paenibacillus* Sp. Str. FPU-7'. *Journal of Structural Biology: X* 5:100049. <https://doi.org/10.1016/j.yjsbx.2021.100049>.
- Iyer, Ramkumar, and Andrew Camilli. 2007. 'Sucrose Metabolism Contributes to *in Vivo* Fitness of *Streptococcus Pneumoniae*'. *Molecular Microbiology* 66 (1): 1–13. <https://doi.org/10.1111/j.1365-2958.2007.05878.x>.
- Jeckelmann, Jean Marc, and Bernhard Erni. 2019. *Carbohydrate Transport by Group Translocation: The Bacterial Phosphoenolpyruvate: Sugar Phosphotransferase System. Subcellular Biochemistry*. Vol. 92. Springer International Publishing. https://doi.org/10.1007/978-3-030-18768-2_8.
- Jecklin, Matthias C., Stefan Schauer, Christoph E. Dumelin, and Renato Zenobi. 2009. 'Label-Free Determination of Protein-Ligand Binding Constants Using Mass Spectrometry and

- Validation Using Surface Plasmon Resonance and Isothermal Titration Calorimetry'. *Journal of Molecular Recognition* 22 (4): 319–29. <https://doi.org/10.1002/jmr.951>.
- Jensen, John Beck, N. Kent Peters, and T. V. Bhuvaneswari. 2002. 'Redundancy in Periplasmic Binding Protein-Dependent Transport Systems for Trehalose, Sucrose, and Maltose in *Sinorhizobium Meliloti*'. *Journal of Bacteriology* 184 (11): 2978–86. <https://doi.org/10.1128/JB.184.11.2978-2986.2002>.
- Joglekar, Payal, Erica D. Sonnenburg, Steven K. Higginbottom, Kristen A. Earle, Carl Morland, Sarah Shapiro-Ward, David N. Bolam, and Justin L. Sonnenburg. 2018. 'Genetic Variation of the SusC/SusD Homologs from a Polysaccharide Utilization Locus Underlies Divergent Fructan Specificities and Functional Adaptation in *Bacteroides Thetaiotaomicron* Strains'. *mSphere* 3 (3): 1–11. <https://doi.org/10.1128/mSphereDirect.00185-18>.
- Johnsen, Ulrike, Marius Ortjohann, Jan-Moritz Sutter, Sarah Geweke, and Peter Schönheit. 2019. 'Uptake of D-Xylose and L-Arabinose in *Haloferax Volcanii* Involves an ABC Transporter of the CUT1 Subfamily'. *FEMS Microbiology Letters* 366 (8): 1–19. <https://doi.org/10.1093/femsle/fnz089>.
- Jones, Peter M., and Anthony M. George. 2020. 'Is the Emperor Wearing Shorts? The Published Structures of ABC Transporters'. *FEBS Letters* 594 (23): 3790–98. <https://doi.org/10.1002/1873-3468.13941>.
- Kaback, H. Ronald. 2015. 'A Chemiosmotic Mechanism of Symport'. *Proceedings of the National Academy of Sciences of the United States of America* 112 (5): 1259–64. <https://doi.org/10.1073/pnas.1419325112>.
- Kaback, H. Ronald, and Lan Guan. 2019. 'It Takes Two to Tango: The Dance of the Permease'. *The Journal of General Physiology* 151 (7): 878–86. <https://doi.org/10.1085/jgp.201912377>.
- Kabisch, Antje, Andreas Otto, Sten König, Dörte Becher, Dirk Albrecht, Margarete Schüler, Hanno Teeling, Rudolf I. Amann, and Thomas Schweder. 2014. 'Functional Characterization of Polysaccharide Utilization Loci in the Marine Bacteroidetes "Gramella Forsetii" KT0803'. *ISME Journal* 8 (7): 1492–1502. <https://doi.org/10.1038/ismej.2014.4>.

- Kachroo, Aashiq H., Aswani K. Kancherla, Nongmaithem S. Singh, Umesh Varshney, and Subramony Mahadevan. 2007. 'Mutations That Alter the Regulation of the Chb Operon of Escherichia Coli Allow Utilization of Cellobiose'. *Molecular Microbiology* 66 (6): 1382–95. <https://doi.org/10.1111/j.1365-2958.2007.05999.x>.
- Kalscheuer, Rainer, Brian Weinrick, Usha Veeraraghavan, Gurdyal S. Besra, and William R. Jacobs. 2010. 'Trehalose-Recycling ABC Transporter LpqY-SugA-SugB-SugC Is Essential for Virulence of Mycobacterium Tuberculosis'. *Proceedings of the National Academy of Sciences of the United States of America* 107 (50): 21761–66. <https://doi.org/10.1073/pnas.1014642108>.
- Kamionka, Annette, and Michael K. Dahl. 2001. 'Bacillus Subtilis Contains a Cyclodextrin-Binding Protein Which Is Part of a Putative ABC-Transporter'. *FEMS Microbiology Letters* 204 (1): 55–60. <https://doi.org/10.1111/j.1574-6968.2001.tb10862.x>.
- Kaneko, Ai, Kasumi Uenishi, Seiki Baba, Nobuhiro Mizuno, Yukie Maruyama, Kousaku Murata, Bunzo Mikami, Wataru Hashimoto, and Takashi Kumasaka. 2017. 'A Solute-Binding Protein in the Closed Conformation Induces ATP Hydrolysis in a Bacterial ATP-Binding Cassette Transporter Involved in the Import of Alginate'. *Journal of Biological Chemistry* 292 (38): 15681–90. <https://doi.org/10.1074/jbc.m117.793992>.
- Kaplan, Handan, and Robert W. Hutkins. 2003. 'Metabolism of Fructooligosaccharides by Lactobacillus Paracasei 1195'. *Applied and Environmental Microbiology* 69 (4): 2217–22. <https://doi.org/10.1128/AEM.69.4.2217-2222.2003>.
- Karlikowska, Magdalena, Albel Singh, Apoorva Bhatt, Sascha Ott, Andrew R. Bottrill, Gurdyal S. Besra, and Elizabeth Fullam. 2021. 'Biochemical and Phenotypic Characterisation of the Mycobacterium Smegmatis Transporter UspABC'. *The Cell Surface* 7 (December):100052. <https://doi.org/10.1016/j.tcs.2021.100052>.
- Keffeler, Erica C., Srivatsan Parthasarathy, Zakria H. Abdullahi, and Lynn E. Hancock. 2021. 'Metabolism of Poly-B1,4-N-Acetylglucosamine Substrates and Importation of N-Acetylglucosamine and Glucosamine by Enterococcus Faecalis'. *Journal of Bacteriology* 203 (21): e00371-21. <https://doi.org/10.1128/JB.00371-21>.
- Keyhani, N. O., K. Bacia, and S. Roseman. 2000. 'The Transport/Phosphorylation of N,N'-Diacetylchitobiose in Escherichia Coli. Characterization of Phospho-IIB(Chb) and of a

- Potential Transition State Analogue in the Phosphotransfer Reaction between the Proteins IIA(Chb) AND IIB(Chb)'. *The Journal of Biological Chemistry* 275 (42): 33102–9. <https://doi.org/10.1074/jbc.M001045200>.
- Khare, Dheeraj, Michael L. Oldham, Cedric Orelle, Amy L. Davidson, and Jue Chen. 2009. 'Alternating Access in Maltose Transporter Mediated by Rigid-Body Rotations'. *Molecular Cell* 33 (4): 528–36. <https://doi.org/10.1016/j.molcel.2009.01.035>.
- King, S. C., and T. H. Wilson. 1990. 'Identification of Valine 177 as a Mutation Altering Specificity for Transport of Sugars by the Escherichia Coli Lactose Carrier. Enhanced Specificity for Sucrose and Maltose'. *Journal of Biological Chemistry* 265 (17): 9638–44.
- Klein, W., R. Horlacher, and W. Boos. 1995. 'Molecular Analysis of treB Encoding the Escherichia Coli Enzyme II Specific for Trehalose'. *Journal of Bacteriology* 177 (14): 4043–52. <https://doi.org/10.1128/jb.177.14.4043-4052.1995>.
- Kohno, Masaki, Takatoshi Arakawa, Naoki Sunagawa, Tetsuya Mori, Kiyohiko Igarashi, Tomoyuki Nishimoto, and Shinya Fushinobu. 2020. 'Molecular Analysis of Cyclic α -Maltosyl-(1→6)-Maltose Binding Protein in the Bacterial Metabolic Pathway'. *PLOS ONE* 15 (11): e0241912. <https://doi.org/10.1371/journal.pone.0241912>.
- Koropatkin, Nicole M., Elizabeth A. Cameron, and Eric C. Martens. 2012. 'How Glycan Metabolism Shapes the Human Gut Microbiota'. *Nature Reviews Microbiology* 10 (5): 323–35. <https://doi.org/10.1038/nrmicro2746>.
- Koropatkin, Nicole M., Eric C. Martens, Jeffrey I. Gordon, and Thomas J. Smith. 2008. 'Starch Catabolism by a Prominent Human Gut Symbiont Is Directed by the Recognition of Amylose Helices'. *Structure* 16 (7): 1105–15. <https://doi.org/10.1016/j.str.2008.03.017>.
- Koropatkin, Nicole, Eric C. Martens, Jeffrey I. Gordon, and Thomas J. Smith. 2009. 'The Structure Of a SusD Homolog, BT1043, Involved in Mucin O- Glycan Utilization in a Prominent Human Gut Symbiont'. *Biochemistry* 23 (1): 1–7. <https://doi.org/10.1038/jid.2014.371>.
- Kotrba, Pavel, Masayuki Inui, and Hideaki Yukawa. 2003. 'A Single V317A or V317M Substitution in Enzyme II of a Newly Identified β -Glucoside Phosphotransferase and

- Utilization System of *Corynebacterium Glutamicum* R Extends Its Specificity towards Cellobiose'. *Microbiology* 149 (6): 1569–80. <https://doi.org/10.1099/mic.0.26053-0>.
- Kowalczyk, Magdalena, Muriel Cocaign-Bousquet, Pascal Loubiere, and Jacek Bardowski. 2008. 'Identification and Functional Characterisation of Cellobiose and Lactose Transport Systems in *Lactococcus Lactis* IL1403'. *Archives of Microbiology* 189 (3): 187–96. <https://doi.org/10.1007/s00203-007-0308-8>.
- Kundig, Werner, Sudhamoy Ghosh, and Saul Roseman. 1964. 'Phosphate Bound to Histidine in a Protein as an Intermediate in a Novel Phospho-Transferase System'. *Proceedings of the National Academy of Sciences* 52 (4): 1067–74. <https://doi.org/10.1073/pnas.52.4.1067>.
- Ladevèze, Simon, Elisabeth Laville, Jordane Despres, Pascale Mosoni, and Gabrielle Potocki-Véronèse. 2017. 'Mannoside Recognition and Degradation by Bacteria'. *Biological Reviews* 92 (4): 1969–90. <https://doi.org/10.1111/brv.12316>.
- Lapébie, Pascal, Vincent Lombard, Elodie Drula, Nicolas Terrapon, and Bernard Henrissat. 2019. 'Bacteroidetes Use Thousands of Enzyme Combinations to Break down Glycans'. *Nature Communications* 10 (1): 2043. <https://doi.org/10.1038/s41467-019-10068-5>.
- Larsbrink, Johan, Theresa E. Rogers, Glyn R. Hemsworth, Lauren S. McKee, Alexandra S. Tauzin, Oliver Spadiut, Stefan Klintner, et al. 2014. 'A Discrete Genetic Locus Confers Xyloglucan Metabolism in Select Human Gut Bacteroidetes'. *Nature* 506 (7489): 498–502. <https://doi.org/10.1038/nature12907>.
- Larsbrink, Johan, Yongtao Zhu, Sampada S. Kharade, Kurt J. Kwiatkowski, Vincent G. H. Eijssink, Nicole M. Koropatkin, Mark J. McBride, and Phillip B. Pope. 2016. 'A Polysaccharide Utilization Locus from *Flavobacterium Johnsoniae* Enables Conversion of Recalcitrant Chitin'. *Biotechnology for Biofuels* 9 (1): 260. <https://doi.org/10.1186/s13068-016-0674-z>.
- Larsen, Ray A., Michael G. Thomas, and Kathleen Postle. 1999. 'Protonmotive Force, ExbB and Ligand-Bound FepA Drive Conformational Changes in TonB'. *Molecular Microbiology* 31 (6): 1809–24. <https://doi.org/10.1046/j.1365-2958.1999.01317.x>.

- Lee, Sang Jun, Sang-Jae Lee, and Dong-Woo Lee. 2013. 'Design and Development of Synthetic Microbial Platform Cells for Bioenergy'. *Frontiers in Microbiology* 4 (April): 1–13. <https://doi.org/10.3389/fmicb.2013.00092>.
- Leth, Maria Louise, Morten Ejby, Christopher Workman, David Adrian Ewald, Signe Schultz Pedersen, Claus Sternberg, Martin Iain Bahl, et al. 2018. 'Differential Bacterial Capture and Transport Preferences Facilitate Co-Growth on Dietary Xylan in the Human Gut'. *Nature Microbiology* 3 (5): 570–80. <https://doi.org/10.1038/s41564-018-0132-8>.
- Lian, Jiazhang, Yanglin Li, Mohammad Hamedirad, and Huimin Zhao. 2014. 'Directed Evolution of a Cellodextrin Transporter for Improved Biofuel Production under Anaerobic Conditions in *Saccharomyces Cerevisiae*'. *Biotechnology and Bioengineering* 111 (8): 1521–31. <https://doi.org/10.1002/bit.25214>.
- Liang, Wei Jun, Kate J. Wilson, Hao Xie, Jan Knol, Shun'ichi Suzuki, Nicholas G. Rutherford, Peter J.F. Henderson, and Richard A. Jefferson. 2005. 'The gusBC Genes of *Escherichia Coli* Encode a Glucuronide Transport System'. *Journal of Bacteriology* 187 (7): 2377–85. <https://doi.org/10.1128/JB.187.7.2377-2385.2005>.
- Licht, Anke, Martin Bommer, Tobias Werther, Kristin Neumann, Carolin Hobe, and Erwin Schneider. 2019. 'Structural and Functional Characterization of a Maltose/Maltodextrin ABC Transporter Comprising a Single Solute Binding Domain (MalE) Fused to the Transmembrane Subunit MalF'. *Research in Microbiology* 170 (1): 1–12. <https://doi.org/10.1016/j.resmic.2018.08.006>.
- Lindstad, Lars J., Galiana Lo, Shaun Leivers, Zijia Lu, Leszek Michalak, Gabriel V. Pereira, Åsmund K. Røhr, et al. 2021. 'Human Gut Faecalibacterium Prausnitzii Deploys a Highly Efficient Conserved System To Cross-Feed on β -Mannan-Derived Oligosaccharides'. *mBio* 12 (3): 10.1128/mbio.03628-20. <https://doi.org/10.1128/mbio.03628-20>.
- Linke, Caroline M., Shireen A. Woodiga, Dustin J. Meyers, Carolyn M. Buckwalter, Hussam E. Salhi, and Samantha J. King. 2013. 'The ABC Transporter Encoded at the Pneumococcal Fructooligosaccharide Utilization Locus Determines the Ability to Utilize Long- and Short-Chain Fructooligosaccharides'. *Journal of Bacteriology* 195 (5): 1031–41. <https://doi.org/10.1128/JB.01560-12>.

- Liu, Fengjiang, Jingxi Liang, Bing Zhang, Yan Gao, Xiuna Yang, Tianyu Hu, Haitao Yang, Wenqing Xu, Luke W. Guddat, and Zihe Rao. 2020. 'Structural Basis of Trehalose Recycling by the ABC Transporter LpqY-SugABC'. *Science Advances* 6 (44): eabb9833. <https://doi.org/10.1126/sciadv.abb9833>.
- Liu, Nian, Séverine Gagnet, Yann Denis, Deborah Byrne, Craig Faulds, Henri-Pierre Fierobe, and Stéphanie Perret. 2022. 'Selfish Uptake versus Extracellular Arabinoxylan Degradation in the Primary Degradar *Ruminiclostridium Cellulolyticum*, a New String to Its Bow'. *Biotechnology for Biofuels and Bioproducts* 15 (1): 127. <https://doi.org/10.1186/s13068-022-02225-8>.
- Liu, Xueli, Jianwei Zeng, Kai Huang, and Jiawei Wang. 2019. 'Structure of the Mannose Transporter of the Bacterial Phosphotransferase System'. *Cell Research* 29 (8): 680–82. <https://doi.org/10.1038/s41422-019-0194-z>.
- Liu, Yansheng, Andreas Beyer, and Ruedi Aebersold. 2016. 'On the Dependency of Cellular Protein Levels on mRNA Abundance'. *Cell* 165 (3): 535–50. <https://doi.org/10.1016/j.cell.2016.03.014>.
- Locher, Kaspar P. 2016. 'Mechanistic Diversity in ATP-Binding Cassette (ABC) Transporters'. *Nature Structural and Molecular Biology* 23 (6): 487–93. <https://doi.org/10.1038/nsmb.3216>.
- Lohmiller, S., K. Hantke, S. I. Patzer, and Volkmar Braun. 2008. 'TonB-Dependent Maltose Transport by *Caulobacter Crescentus*'. *Microbiology* 154 (6): 1748–54. <https://doi.org/10.1099/mic.0.2008/017350-0>.
- Lombard, Vincent, Ramulu H.G., Elodie Drula, Pedro M. Coutinho, and Bernard Henrissat. 2014. 'The Carbohydrate-Active Enzymes Database (CAZy) in 2013'. *Nucleic Acids Research* 42 (D1): 490–95. <https://doi.org/10.1093/nar/gkt1178>.
- Lowe, Rohan, Neil Shirley, Mark Bleackley, Stephen Dolan, and Thomas Shafee. 2017. 'Transcriptomics Technologies'. *PLOS Computational Biology* 13 (5): e1005457. <https://doi.org/10.1371/journal.pcbi.1005457.t005>.
- Lu, Zijia, Alma Kvammen, He Li, Mengshu Hao, Annie R. Inman, Vincent Bulone, and Lauren S. McKee. 2023. 'A Polysaccharide Utilization Locus from *Chitinophaga Pinensis*

- Simultaneously Targets Chitin and β -Glucans Found in Fungal Cell Walls'. *mSphere* 8 (4): e00244-23. <https://doi.org/10.1128/msphere.00244-23>.
- Luis, Ana S., Jonathon Briggs, Xiaoyang Zhang, Benjamin Farnell, Didier Ndeh, Aurore Labourel, Arnaud Baslé, et al. 2018. 'Dietary Pectic Glycans Are Degraded by Coordinated Enzyme Pathways in Human Colonic Bacteroides'. *Nature Microbiology* 3 (2): 210–19. <https://doi.org/10.1038/s41564-017-0079-1>.
- Luis, Ana S., Chunsheng Jin, Gabriel Vasconcelos Pereira, Robert W. P. Glowacki, Sadie R. Gugel, Shaleni Singh, Dominic P. Byrne, et al. 2021. 'A Single Sulfatase Is Required to Access Colonic Mucin by a Gut Bacterium'. *Nature* 598 (7880): 332–37. <https://doi.org/10.1038/s41586-021-03967-5>.
- Mächtel, Rebecca, Alessandra Narducci, Douglas A. Griffith, Thorben Cordes, and Cédric Orelle. 2019. 'An Integrated Transport Mechanism of the Maltose ABC Importer'. *Research in Microbiology* 170. <https://doi.org/10.1016/j.resmic.2019.09.004>.
- Mackenzie, A. K., A. E. Naas, S. K. Kracun, J. Schückel, J. U. Fangel, J. W. Agger, W. G.T. Willats, V. G.H. Eijsink, and Phillip B. Pope. 2015. 'A Polysaccharide Utilization Locus from an Uncultured Bacteroidetes Phylotype Suggests Ecological Adaptation and Substrate Versatility'. *Applied and Environmental Microbiology* 81 (1): 187–95. <https://doi.org/10.1128/AEM.02858-14>.
- Mackenzie, A. K., P. B. Pope, H. L. Pedersen, R. Gupta, M. Morrison, W. G.T. Willats, and V. G.H. Eijsink. 2012. 'Two SusD-like Proteins Encoded within a Polysaccharide Utilization Locus of an Uncultured Ruminant Bacteroidetes Phylotype Bind Strongly to Cellulose'. *Applied and Environmental Microbiology* 78 (16): 5935–37. <https://doi.org/10.1128/AEM.01164-12>.
- Majd, Homa, Martin S. King, Shane M. Palmer, Anthony C. Smith, Liam D.H. Elbourne, Ian T. Paulsen, David Sharples, Peter J.F. Henderson, and Edmund R.S. Kunji. 2018. 'Screening of Candidate Substrates and Coupling Ions of Transporters by Thermostability Shift Assays'. *eLife* 7:1–17. <https://doi.org/10.7554/eLife.38821>.
- Marchler-Bauer, Aron, Yu Bo, Lianyi Han, Jane He, Christopher J. Lanczycki, Shennan Lu, Farideh Chitsaz, et al. 2017. 'CDD/SPARCLE: Functional Classification of Proteins via

- Subfamily Domain Architectures'. *Nucleic Acids Research* 45 (D1): D200–203. <https://doi.org/10.1093/nar/gkw1129>.
- Marion, Carolyn, Jason M. Stewart, Mia F. Tazi, Amanda M. Burnaugh, Caroline M. Linke, Shireen A. Woodiga, and Samantha J. King. 2012. 'Streptococcus Pneumoniae Can Utilize Multiple Sources of Hyaluronic Acid for Growth'. *Infection and Immunity* 80 (4): 1390–98. <https://doi.org/10.1128/IAI.05756-11>.
- Martens, Eric C., Elisabeth C. Lowe, Herbert Chiang, Nicholas A. Pudlo, Meng Wu, Nathan P. McNulty, D. Wade Abbott, et al. 2011. 'Recognition and Degradation of Plant Cell Wall Polysaccharides by Two Human Gut Symbionts'. *PLoS Biology* 9 (12). <https://doi.org/10.1371/journal.pbio.1001221>.
- Martin, S. A., and J. B. Russell. 1987. 'Transport and Phosphorylation of Disaccharides by the Ruminant Bacterium Streptococcus Bovis.' *Applied and Environmental Microbiology* 53 (10): 2388–93.
- Maruyama, Yukie, Wataru Hashimoto, and Kousaku Murata. 2019. 'Structural Studies on Bacterial System Used in the Recognition and Uptake of the Macromolecule Alginate'. *Bioscience, Biotechnology and Biochemistry* 83 (5): 794–802. <https://doi.org/10.1080/09168451.2019.1578642>.
- Maruyama, Yukie, Takafumi Itoh, Ai Kaneko, Yu Nishitani, Bunzo Mikami, Wataru Hashimoto, and Kousaku Murata. 2015. 'Structure of a Bacterial ABC Transporter Involved in the Import of an Acidic Polysaccharide Alginate'. *Structure* 23 (9): 1643–54. <https://doi.org/10.1016/j.str.2015.06.021>.
- Mary, Priyanka Rose, P. Monica, and Mukesh Kapoor. 2022. 'Insights into β -Mannose Oligosaccharide Uptake and Metabolism in Bifidobacterium Adolescentis DSMZ 20083 from Whole-Genome Microarray Analysis'. *bioRxiv*. <https://doi.org/10.1101/2022.03.08.483560>.
- Matsuki, Takahiro, Kana Yahagi, Hiroshi Mori, Hoshitaka Matsumoto, Taeko Hara, Saya Tajima, Eishin Ogawa, et al. 2016. 'A Key Genetic Factor for Fucosyllactose Utilization Affects Infant Gut Microbiota Development'. *Nature Communications* 7 (May): 1–12. <https://doi.org/10.1038/ncomms11939>.

- McCoy, Jason G., Zhenning Ren, Vitali Stanevich, Jumin Lee, Sharmistha Mitra, Elena J. Levin, Sebastien Poget, Matthias Quick, Wonpil Im, and Ming Zhou. 2016. 'The Structure of a Sugar Transporter of the Glucose EIIC Superfamily Provides Insight into the Elevator Mechanism of Membrane Transport'. *Physiology & Behavior* 176 (1): 139–48. <https://doi.org/10.1016/j.physbeh.2017.03.040>.
- McNulty, Nathan P., Meng Wu, Alison R. Erickson, Chongle Pan, Brian K. Erickson, Eric C. Martens, Nicholas A. Pudlo, et al. 2013. 'Effects of Diet on Resource Utilization by a Model Human Gut Microbiota Containing *Bacteroides Cellulosilyticus* WH2, a Symbiont with an Extensive Glycobiome'. *PLoS Biology* 11 (8). <https://doi.org/10.1371/journal.pbio.1001637>.
- Meiborn, Karin L., Xibing B. Li, Alex T. Nielsen, Cheng Yen Wu, Saul Roseman, and Gary K. Schoolnik. 2004. 'The *Vibrio Cholerae* Chitin Utilization Program'. *Proceedings of the National Academy of Sciences of the United States of America* 101 (8): 2524–29. <https://doi.org/10.1073/pnas.0308707101>.
- Mikkat, Stefan, and Martin Hagemann. 2000. 'Molecular Analysis of the ggtBCD Gene Cluster of *Synechocystis* Sp. Strain PCC6803 Encoding Subunits of an ABC Transporter for Osmoprotective Compounds'. *Archives of Microbiology* 174 (4): 273–82. <https://doi.org/10.1007/s002030000201>.
- Mishima, Yumiko, Keiko Momma, Wataru Hashimoto, Bunzo Mikami, and Kousaku Murata. 2003. 'Crystal Structure of AlgQ2, a Macromolecule (Alginate)-Binding Protein of *Sphingomonas* Sp. A1, Complexed with an Alginate Tetrasaccharide at 1.6-Å Resolution'. *Journal of Biological Chemistry* 278 (8): 6552–59. <https://doi.org/10.1074/jbc.M209932200>.
- Miyake, Masayuki, Tohru Terada, Michiko Shimokawa, Naohisa Sugimoto, Takatoshi Arakawa, Kentaro Shimizu, Kiyohiko Igarashi, Kiyotaka Fujita, and Shinya Fushinobu. 2020. 'Structural Analysis of β -L-Arabinobiose-Binding Protein in the Metabolic Pathway of Hydroxyproline-Rich Glycoproteins in *Bifidobacterium Longum*'. *The FEBS Journal* 287 (23): 5114–29. <https://doi.org/10.1111/febs.15315>.
- Momma, Keiko, Masako Okamoto, Yumiko Mishima, Shigetaro Mori, Wataru Hashimoto, and Kousaku Murata. 2000. 'A Novel Bacterial ATP-Binding Cassette Transporter

- System That Allows Uptake of Macromolecules’. *Journal of Bacteriology* 182 (14): 3998–4004. <https://doi.org/10.1128/JB.182.14.3998-4004.2000>.
- Moon, Min Woo, Hyo Jin Kim, Tae Kwang Oh, Chul Soo Shin, Jong Soo Lee, Seong Jun Kim, and Jung Kee Lee. 2005. ‘Analyses of Enzyme II Gene Mutants for Sugar Transport and Heterologous Expression of Fructokinase Gene in *Corynebacterium Glutamicum* ATCC 13032’. *FEMS Microbiology Letters* 244 (2): 259–66. <https://doi.org/10.1016/j.femsle.2005.01.053>.
- Morabbi Heravi, Kambiz, Hildegard Watzlawick, and Josef Altenbuchner. 2019. ‘The *melREDCA* Operon Encodes a Utilization System for the Raffinose Family of Oligosaccharides in *Bacillus Subtilis*’. Edited by Anke Becker. *Journal of Bacteriology* 201 (15). <https://doi.org/10.1128/JB.00109-19>.
- Murata, Kousaku, Shigeyuki Kawai, and Wataru Hashimoto. 2022. ‘Bacteria with a Mouth: Discovery and New Insights into Cell Surface Structure and Macromolecule Transport’. *Proceedings of the Japan Academy, Series B* 98 (10): 529–52. <https://doi.org/10.2183/pjab.98.027>.
- Mystkowska, Agata Anna, Craig Robb, Silvia Vidal-Melgosa, Chiara Vanni, Antonio Fernandez-Guerra, Matthias Höhne, and Jan Hendrik Hehemann. 2018. ‘Molecular Recognition of the Beta-Glucans Laminarin and Pustulan by a SusD-like Glycan-Binding Protein of a Marine Bacteroidetes’. *FEBS Journal* 285 (23): 4465–81. <https://doi.org/10.1111/febs.14674>.
- Naas, A. E., A. K. Mackenzie, J. Mravec, J. Schückel, W. G.T. Willats, V. G.H. Eijsink, and P. B. Pope. 2014. ‘Do Rumen Bacteroidetes Utilize an Alternative Mechanism for Cellulose Degradation?’ *mBio* 5 (4): 4–9. <https://doi.org/10.1128/mBio.01401-14>.
- Nakamura, Shuntaro, Rikuya Kurata, Takashi Tono-zuka, Kazumi Funane, Enoch Y. Park, and Takatsugu Miyazaki. 2023. ‘Bacteroidota Polysaccharide Utilization System for Branched Dextran Exopolysaccharides from Lactic Acid Bacteria’. *Journal of Biological Chemistry* 299 (7): 104885. <https://doi.org/10.1016/j.jbc.2023.104885>.
- Nanavati, Dhaval M., Tu N. Nguyen, and Kenneth M. Noll. 2005. ‘Substrate Specificities and Expression Patterns Reflect the Evolutionary Divergence of Maltose ABC Transporters

- in *Thermotoga Maritima*'. *Journal of Bacteriology* 187 (6): 2002–9. <https://doi.org/10.1128/JB.187.6.2002-2009.2005>.
- Nanavati, Dhaval M., Kamolwan Thirangoon, and Kenneth M. Noll. 2006. 'Several Archaeal Homologs of Putative Oligopeptide-Binding Proteins Encoded by *Thermotoga Maritima* Bind Sugars'. *Applied and Environmental Microbiology* 72 (2): 1336–45. <https://doi.org/10.1128/AEM.72.2.1336-1345.2006>.
- Nataf, Y., S. Yaron, E. A. Bayer, F. Stahl, T.-H. Scheper, Y. Shoham, A. L. Sonenshein, and R. Lamed. 2009. 'Cellodextrin and Laminaribiose ABC Transporters in *Clostridium Thermocellum*'. *Journal of Bacteriology* 191 (1): 203–9. <https://doi.org/10.1128/jb.01190-08>.
- Ndeh, Didier, Arnaud Baslé, Henrik Strahl, Edwin A. Yates, Urszula L. McClurg, Bernard Henrissat, Nicolas Terrapon, and Alan Cartmell. 2020. 'Metabolism of Multiple Glycosaminoglycans by *Bacteroides Thetaiotaomicron* Is Orchestrated by a Versatile Core Genetic Locus'. *Nature Communications* 11 (1): 646. <https://doi.org/10.1038/s41467-020-14509-4>.
- Ndeh, Didier, Artur Rogowski, Alan Cartmell, Ana S. Luis, Arnaud Baslé, Joseph Gray, Immacolata Venditto, et al. 2017. 'Complex Pectin Metabolism by Gut Bacteria Reveals Novel Catalytic Functions'. *Nature* 544 (7648): 65–70. <https://doi.org/10.1038/nature21725>.
- Nguyen, Hoang Hiep, Jeho Park, Sebyung Kang, and Moonil Kim. 2015. 'Surface Plasmon Resonance: A Versatile Technique for Biosensor Applications'. *Sensors (Switzerland)* 15 (5): 10481–510. <https://doi.org/10.3390/s150510481>.
- Nguyen, Thai X., Ming Ren Yen, Ravi D. Barabote, and Milton H. Saier. 2006. 'Topological Predictions for Integral Membrane Permeases of the Phosphoenolpyruvate: Sugar Phosphotransferase System'. *Journal of Molecular Microbiology and Biotechnology* 11 (6): 345–60. <https://doi.org/10.1159/000095636>.
- Niesen, Frank H., Helena Berglund, and Masoud Vedadi. 2007. 'The Use of Differential Scanning Fluorimetry to Detect Ligand Interactions That Promote Protein Stability'. *Nature Protocols* 2 (9): 2212–21. <https://doi.org/10.1038/nprot.2007.321>.

- Nieves-Mori3n, Mercedes, and Enrique Flores. 2018. ‘Multiple ABC Glucoside Transporters Mediate Sugar-Stimulated Growth in the Heterocyst-Forming Cyanobacterium *Anabaena* Sp. Strain PCC 7120’. *Environmental Microbiology Reports* 10 (1): 40–48. <https://doi.org/10.1111/1758-2229.12603>.
- Oiki, Sayoko, Reiko Kamochi, Bunzo Mikami, Kousaku Murata, and Wataru Hashimoto. 2017. ‘Alternative Substrate-Bound Conformation of Bacterial Solute-Binding Protein Involved in the Import of Mammalian Host Glycosaminoglycans’. *Scientific Reports* 7 (1): 1–13. <https://doi.org/10.1038/s41598-017-16801-8>.
- Oiki, Sayoko, Yusuke Nakamichi, Yukie Maruyama, Bunzo Mikami, Kousaku Murata, and Wataru Hashimoto. 2019. ‘Streptococcal Phosphotransferase System Imports Unsaturated Hyaluronan Disaccharide Derived from Host Extracellular Matrices’. Edited by Alberto G. Passi. *PLOS ONE* 14 (11): e0224753. <https://doi.org/10.1371/journal.pone.0224753>.
- Oldham, Michael L., and Jue Chen. 2011. ‘Snapshots of the Maltose Transporter during ATP Hydrolysis’. 2011. <https://doi.org/10.1073/pnas.1108858108>.
- Oldham, Michael L., Shanshuang Chen, and Jue Chen. 2013. ‘Structural Basis for Substrate Specificity in the *Escherichia Coli* Maltose Transport System’. *Proceedings of the National Academy of Sciences of the United States of America* 110 (45): 18132–37. <https://doi.org/10.1073/pnas.1311407110>.
- Oldham, Michael L., Dheeraj Khare, Florante A. Qui3cho, Amy L. Davidson, and Jue Chen. 2007. ‘Crystal Structure of a Catalytic Intermediate of the Maltose Transporter’. *Nature* 450 (7169): 515–21. <https://doi.org/10.1038/nature06264>.
- Panwar, Deepesh, and Mukesh Kapoor. 2020. ‘Transcriptional Analysis of Galactomannooligosaccharides Utilization by *Lactobacillus Plantarum* WCFS1’. *Food Microbiology* 86 (April):103336. <https://doi.org/10.1016/j.fm.2019.103336>.
- Peng, Yang, Sanath Kumar, Ricardo L. Hernandez, Suzanna E. Jones, Kathleen M. Cadle, Kenneth P. Smith, and Manuel F. Varela. 2009. ‘Evidence for the Transport of Maltose by the Sucrose Permease, CscB, of *Escherichia Coli*’. *Journal of Membrane Biology* 228 (2): 79–88. <https://doi.org/10.1007/s00232-009-9161-9>.

- Perez, Camilo, Sabina Gerber, J  r  my Boilevin, Monika Bucher, Tamis Darbre, Markus Aebi, Jean-Louis Reymond, and Kaspar P. Locher. 2015. ‘Structure and Mechanism of an Active Lipid-Linked Oligosaccharide Flippase’. *Nature* 524 (7566): 433–38. <https://doi.org/10.1038/nature14953>.
- Perez, Camilo, Martin K  hler, Daniel Janser, Els Pardon, Jan Steyaert, Renato Zenobi, and Kaspar P. Locher. 2017. ‘Structural Basis of Inhibition of Lipid-Linked Oligosaccharide Flippase PglK by a Conformational Nanobody’. *Scientific Reports* 7 (1): 46641. <https://doi.org/10.1038/srep46641>.
- Perez, Camilo, Ahmad Reza Mehdipour, Gerhard Hummer, and Kaspar P. Locher. 2019. ‘Structure of Outward-Facing PglK and Molecular Dynamics of Lipid-Linked Oligosaccharide Recognition and Translocation’. *Structure* 27 (4): 669-678.e5. <https://doi.org/10.1016/j.str.2019.01.013>.
- Peterkofsky, A., G. Wang, D. S. Garrett, B. R. Lee, Y. J. Seaok, and G. M. Clore. 2001. ‘Three-Dimensional Structures of Protein-Protein Complexes in the E. Coli PTS’. *Journal of Molecular Microbiology and Biotechnology* 3 (3): 347–54.
- Phansopa, Chatchawal, Sumita Roy, John B. Rafferty, C. W. Ian Douglas, Jagroop Pandhal, Phillip C. Wright, David J. Kelly, and Graham P. Stafford. 2014. ‘Structural and Functional Characterization of NanU, a Novel High-Affinity Sialic Acid-Inducible Binding Protein of Oral and Gut-Dwelling Bacteroidetes Species’. *Biochemical Journal* 458 (3): 499–511. <https://doi.org/10.1042/BJ20131415>.
- Pichler, Michael Jakob, Chihaya Yamada, Bashar Shuoker, Camila Alvarez-Silva, Aina Gotoh, Maria Louise Leth, Erwin Schoof, et al. 2020. ‘Butyrate Producing Colonic Clostridiales Metabolise Human Milk Oligosaccharides and Cross Feed on Mucin via Conserved Pathways’. *Nature Communications* 11 (1): 3285. <https://doi.org/10.1038/s41467-020-17075-x>.
- Pikis, Andreas, Sonja Hess, Ingrid Arnold, Bernhard Erni, and John Thompson. 2006. ‘Genetic Requirements for Growth of Escherichia Coli K12 on Methyl-  -D-Glucopyranoside and the Five   -D-Glucosyl-D-Fructose Isomers of Sucrose’. *Journal of Biological Chemistry* 281 (26): 17900–908. <https://doi.org/10.1074/jbc.M601183200>.

- Poolman, B., T. J. Royer, S. E. Mainzer, and B. F. Schmidt. 1989. 'Lactose Transport System of *Streptococcus Thermophilus*: A Hybrid Protein with Homology to the Melibiose Carrier and Enzyme III of Phosphoenolpyruvate-Dependent Phosphotransferase Systems.' *Journal of Bacteriology* 171 (1): 244–53. <https://doi.org/10.1128/jb.171.1.244-253.1989>.
- Pope, Phillip B., Alasdair K. Mackenzie, Ivan Gregor, Wendy Smith, Monica A. Sundset, Alice C. McHardy, Mark Morrison, and Vincent G.H. Eijsink. 2012. 'Metagenomics of the Svalbard Reindeer Rumen Microbiome Reveals Abundance of Polysaccharide Utilization Loci'. *PLoS ONE* 7 (6): 1–10. <https://doi.org/10.1371/journal.pone.0038571>.
- Potocki-Veronese, Gabrielle, Sophie LAJUS, Adilya DAGKESAMANSKAYA, Amandine DEROITE, and Delphine LESTRADE. 2023. New milli-to-micro-fluidics-based process to screen microbial growth in droplets. World Intellectual Property Organization WO2023222765A1, filed 17 May 2023, and issued 23 November 2023. [https://patents.google.com/patent/WO2023222765A1/fr?q=\(lajus+potocki\)&oq=lajus+potocki](https://patents.google.com/patent/WO2023222765A1/fr?q=(lajus+potocki)&oq=lajus+potocki).
- Prestegard, James H., Jian Liu, and Göran Widmalm. 2017. 'Oligosaccharides and Polysaccharides'. In *Essentials of Carbohydrate Chemistry*, 3rd editio, 2015–17. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press. <https://doi.org/10.1101/glycobiology.3e.003>.
- Qu, Chunyun, Yang Zhang, Kaiqun Dai, Hongxin Fu, and Jufang Wang. 2021. 'Metabolic Engineering of *Thermoanaerobacterium Aotearoense* SCUT27 for Glucose and Cellobiose Co-Utilization by Identification and Overexpression of the Endogenous Cellobiose Operon'. *Biochemical Engineering Journal* 167 (March):107922. <https://doi.org/10.1016/j.bej.2020.107922>.
- Quiocho, Florante A., John C. Spurlino, and Lynn E. Rodseth. 1997. 'Extensive Features of Tight Oligosaccharide Binding Revealed in High-Resolution Structures of the Maltodextrin Transport/Chemosensory Receptor'. *Structure* 5 (8): 997–1015. [https://doi.org/10.1016/S0969-2126\(97\)00253-0](https://doi.org/10.1016/S0969-2126(97)00253-0).
- Rakoff-Nahoum, Seth, Michael J. Coyne, and Laurie E. Comstock. 2014. 'An Ecological Network of Polysaccharide Utilization among Human Intestinal Symbionts'. *Current Biology* 24 (1): 40–49. <https://doi.org/10.1016/j.cub.2013.10.077>.

- Ravcheev, Dmitry A., Adam Godzik, Andrei L. Osterman, and Dmitry A. Rodionov. 2013. 'Polysaccharides Utilization in Human Gut Bacterium *Bacteroides Thetaiotaomicron*: Comparative Genomics Reconstruction of Metabolic and Regulatory Networks'. *BMC Genomics* 14 (1): 1–17. <https://doi.org/10.1186/1471-2164-14-873>.
- Reddy, Vamsee S, Maksim a Shlykov, Rostislav Castillo, Eric I Sun, and Milton H Saier. 2012. 'The Major Facilitator Superfamily (MFS) Revisited'. *FEBS Journal* 279 (11): 2022–35. <https://doi.org/10.1111/j.1742-4658.2012.08588.x>.
- Reeves, Andrew R., John N. D'Elia, Jorge Frias, and Abigail A. Salyers. 1996. 'A *Bacteroides Thetaiotaomicron* Outer Membrane Protein That Is Essential for Utilization of Maltooligosaccharides and Starch'. *Journal of Bacteriology* 178 (3): 823–30. <https://doi.org/10.1128/jb.178.3.823-830.1996>.
- Reidl, J., and W. Boos. 1991. 'The malX malY Operon of *Escherichia Coli* Encodes a Novel Enzyme II of the Phosphotransferase System Recognizing Glucose and Maltose and an Enzyme Abolishing the Endogenous Induction of the Maltose System'. *Journal of Bacteriology* 173 (15): 4862–76. <https://doi.org/10.1128/jb.173.15.4862-4876.1991>.
- Ren, Qinghu, and Ian T. Paulsen. 2007. 'Large-Scale Comparative Genomic Analyses of Cytoplasmic Membrane Transport Systems in Prokaryotes'. *Journal of Molecular Microbiology and Biotechnology* 12 (3–4): 165–79. <https://doi.org/10.1159/000099639>.
- Ren, Zhenning, Jumin Lee, Mahdi Muhammad Moosa, Yin Nian, Liya Hu, Zhichun Xu, Jason G McCoy, Allan Chris M Ferreon, Wonpil Im, and Ming Zhou. 2018. 'Structure of an EIIC Sugar Transporter Trapped in an Inward-Facing Conformation.' *Proceedings of the National Academy of Sciences of the United States of America* 115 (23): 5962–67. <https://doi.org/10.1073/pnas.1800647115>.
- Robillard, G. T., and J. Broos. 1999. 'Structure/Function Studies on the Bacterial Carbohydrate Transporters, Enzymes II, of the Phosphoenolpyruvate-Dependent Phosphotransferase System'. *Biochimica et Biophysica Acta - Reviews on Biomembranes* 1422 (2): 73–104. [https://doi.org/10.1016/S0304-4157\(99\)00002-7](https://doi.org/10.1016/S0304-4157(99)00002-7).
- Rodionov, Dmitry A., Irina A. Rodionova, Vladimir A. Rodionov, Aleksandr A. Arzamasov, Ke Zhang, Gabriel M. Rubinstein, Tania N. N. Tanwee, et al. 2021. 'Transcriptional Regulation of Plant Biomass Degradation and Carbohydrate Utilization Genes in the

- Extreme Thermophile *Caldicellulosiruptor Bescii*'. Edited by Zarath M. Summers. *mSystems* 6 (3): e01345-20. <https://doi.org/10.1128/mSystems.01345-20>.
- Rodionov, Dmitry A., Chen Yang, Xiaoqing Li, Irina A. Rodionova, Yanbing Wang, Anna Y. Obraztsova, Olga P. Zagnitko, et al. 2010. 'Genomic Encyclopedia of Sugar Utilization Pathways in the *Shewanella* Genus'. *BMC Genomics* 11 (1). <https://doi.org/10.1186/1471-2164-11-494>.
- Rodríguez-Díaz, Jesús, Antonio Rubio-del-Campo, and María J. Yebra. 2012. 'Lactobacillus Casei Ferments the N-Acetylglucosamine Moiety of Fucosyl- α -1,3-N-Acetylglucosamine and Excretes L-Fucose'. *Applied and Environmental Microbiology* 78 (13): 4613–19. <https://doi.org/10.1128/AEM.00474-12>.
- Rogowski, Artur, Jonathon A. Briggs, Jennifer C. Mortimer, Theodora Tryfona, Nicolas Terrapon, Elisabeth C. Lowe, Arnaud Baslé, et al. 2015. 'Glycan Complexity Dictates Microbial Resource Allocation in the Large Intestine'. *Nature Communications* 6 (1): 7481. <https://doi.org/10.1038/ncomms8481>.
- Rosa, Sabina Leanti La, Maria Louise Leth, Leszek Michalak, Morten Ejby Hansen, Nicholas A. Pudlo, Robert Glowacki, Gabriel Pereira, et al. 2019. 'The Human Gut Firmicute *Roseburia Intestinalis* Is a Primary Degradar of Dietary β -Mannans'. *Nature Communications* 10 (1): 905. <https://doi.org/10.1038/s41467-019-08812-y>.
- Sabharwal, Nishika, Khushbu Varshney, Pragyan Parimita Rath, Samudrala Gourinath, and Uddipan Das. 2020. 'Biochemical and Biophysical Characterization of Nucleotide Binding Domain of Trehalose Transporter from *Mycobacterium Tuberculosis*'. *International Journal of Biological Macromolecules* 152 (June):109–16. <https://doi.org/10.1016/j.ijbiomac.2020.02.237>.
- Saier, M. H. 2006. 'TCDB: The Transporter Classification Database for Membrane Transport Protein Analyses and Information'. *Nucleic Acids Research* 34 (90001): D181–86. <https://doi.org/10.1093/nar/gkj001>.
- Saier, Milton H. 2000. 'Families of Transmembrane Sugar Transport Proteins'. *Molecular Microbiology* 35 (4): 699–710. <https://doi.org/10.1046/j.1365-2958.2000.01759.x>.

- . 2015. ‘The Bacterial Phosphotransferase System: New Frontiers 50 Years after Its Discovery’. *Journal of Molecular Microbiology and Biotechnology* 25 (0): 73–78. <https://doi.org/10.1159/000381215>.
- Saier, Milton H., Vamsee S. Reddy, Brian V. Tsu, Muhammad Saad Ahmed, Chun Li, and Gabriel Moreno-Hagelsieb. 2016. ‘The Transporter Classification Database (TCDB): Recent Advances’. *Nucleic Acids Research* 44 (D1): D372–79. <https://doi.org/10.1093/nar/gkv1103>.
- Saito, Akihiro, Takeshi Fujii, Tomonori Shinya, Naoto Shibuya, Akikazu Ando, and Kiyotaka Miyashita. 2008. ‘The *msiK* Gene, Encoding the ATP-Hydrolysing Component of N,N’-Diacetylchitobiose ABC Transporters, Is Essential for Induction of Chitinase Production in *Streptomyces Coelicolor* A3(2)’. *Microbiology* 154 (11): 3358–65. <https://doi.org/10.1099/mic.0.2008/019612-0>.
- Saito, Yuki, Akira Shigehisa, Yohei Watanabe, Naoki Tsukuda, Kaoru Moriyama-Ohara, Taeko Hara, Satoshi Matsumoto, Hirokazu Tsuji, and Takahiro Matsuki. 2020. ‘Multiple Transporters and Glycoside Hydrolases Are Involved in Arabinoxylan-Derived Oligosaccharide Utilization in *Bifidobacterium Pseudocatenulatum*’. Edited by Ning-Yi Zhou. *Applied and Environmental Microbiology* 86 (24): e01782-20. <https://doi.org/10.1128/AEM.01782-20>.
- Sakanaka, Mikiyasu, Morten Ejby Hansen, Aina Gotoh, Toshihiko Katoh, Keisuke Yoshida, Toshitaka Odamaki, Hiroyuki Yachi, et al. 2019. ‘Evolutionary Adaptation in Fucosyllactose Uptake Systems Supports Bifidobacteria-Infant Symbiosis’. *Science Advances* 5 (8): eaaw7696. <https://doi.org/10.1126/sciadv.aaw7696>.
- Saulnier, D. M A, Douwe Molenaar, Willem M. De Vos, Glenn R. Gibson, and Sofia Kolida. 2007. ‘Identification of Prebiotic Fructooligosaccharide Metabolism in *Lactobacillus Plantarum* WCFS1 through Microarrays’. *Applied and Environmental Microbiology* 73 (6): 1753–65. <https://doi.org/10.1128/AEM.01151-06>.
- Saulnier, Delphine M. A., Douwe Molenaar, Willem M. de Vos, Glenn R. Gibson, and Sofia Kolida. 2007. ‘Identification of Prebiotic Fructooligosaccharide Metabolism in *Lactobacillus Plantarum* WCFS1 through Microarrays’. *Applied and Environmental Microbiology* 73 (6): 1753–65. <https://doi.org/10.1128/AEM.01151-06>.

- Sauvageot, Nicolas, Abdelhamid Mokhtari, Philippe Joyet, Aurélie Budin-Verneuil, Víctor S. Blancato, Guillermo D. Repizo, Céline Henry, et al. 2017. 'Enterococcus Faecalis Uses a Phosphotransferase System Permease and a Host Colonization-Related ABC Transporter for Maltodextrin Uptake'. *Journal of Bacteriology* 199 (9): 1–17. <https://doi.org/10.1128/JB.00878-16>.
- Schlösser, Andreas, Jens Jantos, Karl Hackmann, and Hildgund Schrempf. 1999. 'Characterization of the Binding Protein-Dependent Cellobiose and Cellotriose Transport System of the Cellulose Degradier Streptomyces Reticuli'. *Applied and Environmental Microbiology* 65 (6): 2636–43.
- Schneider, Erwin. 2001. 'ABC Transporters Catalyzing Carbohydrate Uptake'. *Research in Microbiology* 152 (3–4): 303–10. [https://doi.org/10.1016/S0923-2508\(01\)01201-3](https://doi.org/10.1016/S0923-2508(01)01201-3).
- Schönert, Stefan, Sabine Seitz, Holger Krafft, Eva Anne Feuerbaum, Iris Andernach, Gabriele Witz, and Michael K. Dahl. 2006. 'Maltose and Maltodextrin Utilization by Bacillus Subtilis'. *Journal of Bacteriology* 188 (11): 3911–22. <https://doi.org/10.1128/JB.00213-06>.
- Schouler, Catherine, Ahmed Taki, Iman Chouikha, Maryvonne Moulin-Schouleur, and Philippe Gilot. 2009. 'A Genomic Island of an Extraintestinal Pathogenic *Escherichia Coli* Strain Enables the Metabolism of Fructooligosaccharides, Which Improves Intestinal Colonization'. *Journal of Bacteriology* 91 (1): 388–93. <https://doi.org/10.1128/JB.01052-08>.
- Schwalm, Nathan D., and Eduardo A. Groisman. 2017. 'Navigating the Gut Buffet: Control of Polysaccharide Utilization in *Bacteroides* Spp.' *Trends in Microbiology* 25 (12): 1005–15. <https://doi.org/10.1016/j.tim.2017.06.009>.
- Severi, Emmanuele, Michelle Rudden, Andrew Bell, Tracy Palmer, Nathalie Juge, and Gavin H. Thomas. 2021. 'Multiple Evolutionary Origins Reflect the Importance of Sialic Acid Transporters in the Colonization Potential of Bacterial Pathogens and Commensals'. *Microbial Genomics* 7 (6). <https://doi.org/10.1099/mgen.0.000614>.
- Sharma, Dipika, Mandeep Singh, Punit Kaur, and Uddipan Das. 2022. 'Structural Analysis of LpqY, a Substrate-Binding Protein from the SugABC Transporter of *Mycobacterium Tuberculosis*, Provides Insights into Its Trehalose Specificity'. *Acta Crystallographica*

- Shelburne, Samuel A., David B. Keith, Michael T. Davenport, Nicola Horstmann, Richard G. Brennan, and James M. Musser. 2008. ‘Molecular Characterization of Group A Streptococcus Maltodextrin Catabolism and Its Role in Pharyngitis’. *Molecular Microbiology* 69 (2): 436–52. <https://doi.org/10.1111/j.1365-2958.2008.06290.x>.
- Sheridan, Paul O., Jennifer C Martin, Trevor D. Lawley, Hilary P. Browne, Hugh M. B. Harris, Annick Bernalier-Donadille, Sylvia H. Duncan, Paul W. O’Toole, Karen P. Scott, and Harry J. Flint. 2016. ‘Polysaccharide Utilization Loci and Nutritional Specialization in a Dominant Group of Butyrate-Producing Human Colonic *Firmicutes*’. *Microbial Genomics* 2 (2): e000043. <https://doi.org/10.1099/mgen.0.000043>.
- Shinnick, Steven G., Stephanie A. Perez, and Manuel F. Varela. 2003. ‘Altered Substrate Selection of the Melibiose Transporter (MelY) of Enterobacter Cloacae Involving Point Mutations in Leu-88, Leu-91, and Ala-182 That Confer Enhanced Maltose Transport’. *Journal of Bacteriology* 185 (12): 3672–77. <https://doi.org/10.1128/JB.185.12.3672-3677.2003>.
- Shipman, Joseph A., Kyu Hong Cho, Hilary A. Siegel, and Abigail A. Salyers. 1999. ‘Physiological Characterization of SusG, an Outer Membrane Protein Essential for Starch Utilization by Bacteroides Thetaiotaomicron’. *Journal of Bacteriology* 181 (23): 7206–11. <https://doi.org/10.1128/jb.181.23.7206-7211.1999>.
- Shulami, Smadar, Orit Gat, Abraham L. Sonenshein, and Yuval Shoham. 1999. ‘The Glucuronic Acid Utilization Gene Cluster from Bacillus Stearothermophilus T-6’. *Journal of Bacteriology* 181 (12): 3695–3704.
- Shulami, Smadar, Ayelet Raz-Pasteur, Orly Tabachnikov, Sarah Gilead-Gropper, Itzhak Shner, and Yuval Shoham. 2011. ‘The L-Arabinan Utilization System of Geobacillus Stearothermophilus’. *Journal of Bacteriology* 193 (11): 2838–50. <https://doi.org/10.1128/JB.00222-11>.
- Shulami, Smadar, Galia Zaide, Gennady Zolotnitsky, Yael Langut, Geoff Feld, Abraham L. Sonenshein, and Yuval Shoham. 2007. ‘A Two-Component System Regulates the Expression of an ABC Transporter for Xylo-Oligosaccharides in Geobacillus

- Stearothermophilus'. *Applied and Environmental Microbiology* 73 (3): 874–84. <https://doi.org/10.1128/AEM.02367-06>.
- Shulami, Smadar, Arie Zehavi, Valery Belakhov, Rachel Salama, Shifra Lansky, Timor Baasov, Gil Shoham, and Yuval Shoham. 2020. 'Cross-Utilization of β -Galactosides and Cellobiose in *Geobacillus Stearothermophilus*'. *Journal of Biological Chemistry* 295 (31): 10766–80. <https://doi.org/10.1074/jbc.RA120.014029>.
- Silale, Augustinas, and Bert van den Berg. 2023. 'TonB-Dependent Transport Across the Bacterial Outer Membrane'. *Annual Review of Microbiology* 77 (1): 67–88. <https://doi.org/10.1146/annurev-micro-032421-111116>.
- Silhavy, Thomas J., Edith Brickman, Philip J. Bassford, Malcolm J. Casadaban, Howard A. Shuman, Vicki Schwartz, Leonard Guarente, Maxime Schwartz, and Jonathan R. Beckwith. 1979. 'Structure of the malB Region in Escherichia Coli K12'. *Molecular & General Genetics* 174 (3): 261–67. <https://doi.org/10.1007/bf00267798>.
- Silva, Zélia, Maria-manuel Sampaio, Anke Henne, Alex Bo, Ruben Gutzat, Winfried Boos, Milton S Costa, and Helena Santos. 2005. 'The High-Affinity Maltose / Trehalose ABC Transporter in the Extremely Thermophilic Bacterium Thermus Thermophilus HB27 Also Recognizes Sucrose and Palatinose'. *Journal of Bacteriology* 187 (4): 1210–18. <https://doi.org/10.1128/JB.187.4.1210>.
- Singh, Ravindra Pal, Jayashree Niharika, Raksha Thakur, Ben A. Wagstaff, Gulshan Kumar, Rikuya Kurata, Dhaval Patel, Colin W. Levy, Takatsugu Miyazaki, and Robert A. Field. 2023. 'Utilization of Dietary Mixed-Linkage β -Glucans by the Firmicute Blautia Producta'. *Journal of Biological Chemistry* 299 (6): 104806. <https://doi.org/10.1016/j.jbc.2023.104806>.
- Smith, K. A., and A. A. Salyers. 1991. 'Characterization of a Neopullulanase and an α -Glucosidase from Bacteroides Thetaiotaomicron 95-1'. *Journal of Bacteriology* 173 (9): 2962–68. <https://doi.org/10.1128/jb.173.9.2962-2968.1991>.
- Song, Yajian, Jinshan Li, Shan Meng, Liang Yin, Yanfen Xue, and Yanhe Ma. 2016. 'A Novel Manno-Oligosaccharide Binding Protein Identified in Alkaliphilic Bacillus Sp. N16-5 Is Involved in Mannan Utilization'. *PLoS ONE* 11 (3): 1–12. <https://doi.org/10.1371/journal.pone.0150059>.

- Sonnenburg, Erica D., Hongjun Zheng, Payal Joglekar, Steven K. Higginbottom, Susan J. Firbank, David N. Bolam, and Justin L. Sonnenburg. 2010. 'Specificity of Polysaccharide Use in Intestinal Bacteroides Species Determines Diet-Induced Microbiota Alterations'. *Cell* 141 (7): 1241–52. <https://doi.org/10.1016/j.cell.2010.05.005>.
- Soucy, Shannon M., Jinling Huang, and Johann Peter Gogarten. 2015. 'Horizontal Gene Transfer: Building the Web of Life'. *Nature Reviews Genetics* 16 (8): 472–82. <https://doi.org/10.1038/nrg3962>.
- Steen, Andrew D., Alexander Crits-Christoph, Paul Carini, Kristen M. DeAngelis, Noah Fierer, Karen G. Lloyd, and J. Cameron Thrash. 2019. 'High Proportions of Bacteria and Archaea across Most Biomes Remain Uncultured'. *The ISME Journal* 13:3126–30. <https://doi.org/10.1038/s41396-019-0484-y>.
- Su, Qinzhi, Min Tang, Jiahui Hu, Junbo Tang, Xue Zhang, Xingan Li, Qingsheng Niu, Xuguo Zhou, Shiqi Luo, and Xin Zhou. 2022. 'Significant Compositional and Functional Variation Reveals the Patterns of Gut Microbiota Evolution among the Widespread Asian Honeybee Populations'. *Frontiers in Microbiology* 13 (September):934459. <https://doi.org/10.3389/fmicb.2022.934459>.
- Suzuki, Ryuichiro, Jun Wada, Takane Katayama, Shinya Fushinobu, Takayoshi Wakagi, Hirofumi Shoun, Hayuki Sugimoto, et al. 2008. 'Structural and Thermodynamic Analyses of Solute-Binding Protein from Bifidobacterium Longum Specific for Core 1 Disaccharide and Lacto-N-Biose'. *Journal of Biological Chemistry* 283 (19): 13165–73. <https://doi.org/10.1074/jbc.M709777200>.
- Tam, R., and M. H. Saier. 1993. 'Structural, Functional, and Evolutionary Relationships among Extracellular Solute-Binding Receptors of Bacteria'. *Microbiological Reviews* 57 (2): 320–46. <https://doi.org/10.1128/mr.57.2.320-346.1993>
- Tamura, Kazune, Guillaume Dejean, Filip Van Petegem, and Harry Brumer. 2021. 'Distinct Protein Architectures Mediate Species-Specific Beta-Glucan Binding and Metabolism in the Human Gut Microbiota'. *Journal of Biological Chemistry* 296 (January):100415. <https://doi.org/10.1016/j.jbc.2021.100415>.

- Tamura, Kazune, Matthew H. Foley, Bernd R. Gardill, Guillaume Dejean, Matthew Schnizlein, Constance M.E. Bahr, A. Louise Creagh, Filip van Petegem, Nicole M. Koropatkin, and Harry Brumer. 2019. 'Surface Glycan-Binding Proteins Are Essential for Cereal Beta-Glucan Utilization by the Human Gut Symbiont *Bacteroides Ovatus*'. *Cellular and Molecular Life Sciences*, no. 0123456789. <https://doi.org/10.1007/s00018-019-03115-3>.
- Tamura, Kazune, Glyn R. Hemsworth, Guillaume Déjean, Theresa E. Rogers, Nicholas A. Pudlo, Karthik Urs, Namrata Jain, Gideon J. Davies, Eric C. Martens, and Harry Brumer. 2017. 'Molecular Mechanism by Which Prominent Human Gut Bacteroidetes Utilize Mixed-Linkage Beta-Glucans, Major Health-Promoting Cereal Polysaccharides'. *Cell Reports* 21 (7): 2030. <https://doi.org/10.1016/j.celrep.2017.11.013>.
- Tancula, E., M. J. Feldhaus, L. A. Bedzyk, and A. A. Salyers. 1992. 'Location and Characterization of Genes Involved in Binding of Starch to the Surface of *Bacteroides Thetaiotaomicron*'. *Journal of Bacteriology* 174 (17): 5609–16. <https://doi.org/10.1128/jb.174.17.5609-5616.1992>.
- Tao, L., I. C. Sutcliffe, R. R.B. Russell, and J. J. Ferretti. 1993. 'Cloning and Expression of the Multiple Sugar Metabolism (Msm) Operon of Streptococcus Mutans in Heterologous Streptococcal Hosts'. *Infection and Immunity* 61 (3): 1121–25. <https://doi.org/10.1128/iai.61.3.1121-1125.1993>
- Tauzin, Alexandra S., Kurt J. Kwiatkowski, Nicole I. Orlovsky, Christopher J. Smith, A. Louise Creagh, Charles A. Haynes, Zdzislaw Wawrzak, Harry Brumer, and Nicole M. Koropatkin. 2016. 'Molecular Dissection of Xyloglucan Recognition in a Prominent Human Gut Symbiont'. *mBio* 7 (2): e02134-15. <https://doi.org/10.1128/mBio.02134-15>.
- Tauzin, Alexandra S., Elisabeth Laville, Yao Xiao, Sébastien Nouaille, Pascal Le Bourgeois, Stéphanie Heux, Jean Charles Portais, Pierre Monsan, Eric C. Martens, Gabrielle Potocki-Veronese, et al. 2016. 'Functional Characterization of a Gene Locus from an Uncultured Gut *Bacteroides* Conferring Xylo-Oligosaccharides Utilization to *Escherichia Coli*'. *Molecular Microbiology* 102 (4): 579–92. <https://doi.org/10.1111/mmi.13480>.
- Tauzin, Alexandra S., Elisabeth Laville, Yao Xiao, Sébastien Nouaille, Pascal Le Bourgeois, Stéphanie Heux, Jean-Charles Portais, Pierre Monsan, Eric C. Martens, Gabrielle

- Potocki-Veronese, et al. 2016. 'Functional Characterization of a Gene Locus from an Uncultured Gut *Bacteroides* Conferring Xylo-oligosaccharides Utilization to *Escherichia Coli*'. *Molecular Microbiology* 102 (4): 579–92. <https://doi.org/10.1111/mmi.13480>.
- Tauzin, Alexandra S., Zhi Wang, Gianluca Cioci, Xiaoqian Li, Aurore Labourel, Barbara Machado, Guy Lippens, and Gabrielle Potocki-Veronese. 2022. 'Structural and Biochemical Characterization of a Nonbinding SusD-Like Protein Involved in Xylooligosaccharide Utilization by an Uncultured Human Gut *Bacteroides* Strain'. Edited by Vincent B. Young. *mSphere* 7 (5): e00244-22. <https://doi.org/10.1128/msphere.00244-22>.
- Tavoulari, Sotiria, and Stathis Frillingos. 2008. 'Substrate Selectivity of the Melibiose Permease (MelY) from *Enterobacter Cloacae*'. *Journal of Molecular Biology* 376 (3): 681–93. <https://doi.org/10.1016/j.jmb.2007.12.015>.
- Tchieu, J H, V Norris, J S Edwards, and M H Saier. 2001. 'The Complete Phosphotransferase System in *Escherichia Coli*'. *Journal of Molecular Microbiology and Biotechnology* 3 (3): 329–46.
- Temple, Max J., Fiona Cuskin, Arnaud Baslé, Niall Hickey, Gaetano Speciale, Spencer J. Williams, Harry J. Gilbert, and Elisabeth C. Lowe. 2017. 'A Bacteroidetes Locus Dedicated to Fungal 1,6--Glucan Degradation: Unique Substrate Conformation Drives Specificity of the Key Endo-1,6--Glucanase'. *Journal of Biological Chemistry* 292 (25): 10639–50. <https://doi.org/10.1074/jbc.M117.787606>.
- Terrapon, Nicolas, Vincent Lombard, Élodie Drula, Pascal Lapébie, Saad Al-Masaudi, Harry J. Gilbert, and Bernard Henrissat. 2018. 'PULDB: The Expanded Database of Polysaccharide Utilization Loci'. *Nucleic Acids Research* 46 (D1): D677–83. <https://doi.org/10.1093/nar/gkx1022>.
- Terrapon, Nicolas, Vincent Lombard, Harry J. Gilbert, and Bernard Henrissat. 2015. 'Automatic Prediction of Polysaccharide Utilization Loci in Bacteroidetes Species'. *Bioinformatics* 31 (5): 647–55. <https://doi.org/10.1093/bioinformatics/btu716>.
- Theilmann, Mia Christine, Folmer Fredslund, Birte Svensson, Leila Lo Leggio, and Maher Abou Hachem. 2019. 'Substrate Preference of an ABC Importer Corresponds to

- Selective Growth on β -(1,6)-Galactosides in *Bifidobacterium Animalis* Subsp. *Lactis*'. *Journal of Biological Chemistry* 294 (31): 11701–11. <https://doi.org/10.1074/jbc.ra119.008843>.
- Theodoulou, F. L., and I. D. Kerr. 2015. 'ABC Transporter Research: Going Strong 40 Years On'. *Biochemical Society Transactions* 43 (5): 1033–40. <https://doi.org/10.1042/BST20150139>.
- Thomas, Christoph, Stephen G. Aller, Konstantinos Beis, Elisabeth P. Carpenter, Geoffrey Chang, Lei Chen, Elie Dassa, et al. 2020. 'Structural and Functional Diversity Calls for a New Classification of ABC Transporters'. *FEBS Letters* 594 (23): 3767–75. <https://doi.org/10.1002/1873-3468.13935>.
- Titgemeyer, F., K. Jahreis, R. Ebner, and J. W. Lengeler. 1996. 'Molecular Analysis of the scrA and scrB Genes from *Klebsiella Pneumoniae* and Plasmid pUR400, Which Encode the Sucrose Transport Protein Enzyme IIScr of the Phosphotransferase System and a Sucrose-6-Phosphate Invertase'. *Molecular and General Genetics* 250 (2): 197–206. <https://doi.org/10.1007/BF02174179>.
- Tobisch, Steffen, Philippe Glaser, Steffen Krüger, and Michael Hecker. 1997. 'Identification and Characterization of a New β -Glucoside Utilization System in *Bacillus Subtilis*'. *Journal of Bacteriology* 179 (2): 496–506. <https://doi.org/10.1128/jb.179.2.496-506.1997>.
- Tomme, Peter, Alisdair Boraston, Jeffery M. Kormos, R. Antony J. Warren, and Douglas G. Kilburn. 2000. 'Affinity Electrophoresis for the Identification and Characterization of Soluble Sugar Binding by Carbohydrate-Binding Modules'. *Enzyme and Microbial Technology* 27 (7): 453–58. [https://doi.org/10.1016/S0141-0229\(00\)00246-5](https://doi.org/10.1016/S0141-0229(00)00246-5).
- Tsujibo, Hiroshi, Mitsuo Kosaka, Sadao Ikenishi, Takaji Sato, Katsushiro Miyamoto, and Yoshihiko Inamori. 2004. 'Molecular Characterization of a High-Affinity Xylobiose Transporter of *Streptomyces Thermoviolaceus* OPC-520 and Its Transcriptional Regulation'. *Journal of Bacteriology* 186 (4): 1029–37. <https://doi.org/10.1128/JB.186.4.1029-1037.2004>.
- Tsujikawa, Yuji, Shu Ishikawa, Iwao Sakane, Ken-ichi Yoshida, and Ro Osawa. 2021. 'Identification of Genes Encoding a Novel ABC Transporter in *Lactobacillus*

- Delbrueckii for Inulin Polymers Uptake'. *Scientific Reports* 11 (1): 16007. <https://doi.org/10.1038/s41598-021-95356-1>.
- Uchiyama, Taku, Ryousuke Kaneko, Junko Yamaguchi, Akane Inoue, Takahiro Yanagida, Naoki Nikaidou, Miguel Regue, and Takeshi Watanabe. 2003. 'Uptake of N,N'-Diacetylchitobiose [(GlcNAc)₂] via the Phosphotransferase System Is Essential for Chitinase Production by *Serratia Marcescens* 2170'. *Journal of Bacteriology* 185 (6): 1776–82. <https://doi.org/10.1128/JB.185.6.1776-1782.2003>.
- Vahedi-Faridi, Ardeschir, Anke Licht, Haydar Bulut, Frank Scheffel, Sandro Keller, Udo F. Wehmeier, Wolfram Saenger, and Erwin Schneider. 2010. 'Crystal Structures of the Solute Receptor GacH of *Streptomyces Glaucescens* in Complex with Acarbose and an Acarbose Homolog: Comparison with the Acarbose-Loaded Maltose-Binding Protein of *Salmonella Typhimurium*'. *Journal of Molecular Biology* 397 (3): 709–23. <https://doi.org/10.1016/j.jmb.2010.01.054>.
- Viegas, Aldino, João Manso, Franklin L. Nobrega, and Eurico J. Cabrita. 2011. 'Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for Ligand Screening and Characterization of Protein Binding'. *Journal of Chemical Education* 88 (7): 990–94. <https://doi.org/10.1021/ed101169t>.
- Viens, Pascal, Marie-Pierre Dubeau, Akane Kimura, Yoshitake Desaki, Tomonori Shinya, Naoto Shibuya, Akihiro Saito, and Ryszard Brzezinski. 2015. 'Uptake of Chitosan-Derived D-Glucosamine Oligosaccharides in *Streptomyces Coelicolor* A3(2)'. *FEMS Microbiology Letters* 362 (9). <https://doi.org/10.1093/femsle/fnv048>.
- Vuignier, Karine, Julie Schappler, Jean Luc Veuthey, Pierre Alain Carrupt, and Sophie Martel. 2010. 'Drug-Protein Binding: A Critical Review of Analytical Tools'. *Analytical and Bioanalytical Chemistry* 398 (1): 53–66. <https://doi.org/10.1007/s00216-010-3737-1>.
- Walter, Axel, Simon Friz, and Christoph Mayer. 2021. 'Chitin, Chitin Oligosaccharide, and Chitin Disaccharide Metabolism of *Escherichia Coli* Revisited: Reassignment of the Roles of ChiA, ChbR, ChbF, and ChbG'. *Microbial Physiology* 31 (2): 178–94. <https://doi.org/10.1159/000515178>.

- Wandersman, C., M. Schwartz, and T. Ferenci. 1979. 'Escherichia Coli Mutants Impaired in Maltodextrin Transport'. *Journal of Bacteriology* 140 (1): 1–13. <https://doi.org/10.1128/jb.140.1.1-13.1979>
- Wang, Zhi, Alexandra S. Tauzin, Elisabeth Laville, Pietro Tedesco, Fabien Létisse, Nicolas Terrapon, Pascale Lepercq, Myriam Mercade, and Gabrielle Potocki-Veronese. 2020. 'Harvesting of Prebiotic Fructooligosaccharides by Nonbeneficial Human Gut Bacteria'. Edited by Vincent B. Young. *mSphere* 5 (1): e00771-19. <https://doi.org/10.1128/mSphere.00771-19>.
- Watanabe, Akira, Kazumi Hiraga, Masako Suda, Hideaki Yukawa, and Masayuki Inui. 2015. 'Functional Characterization of Corynebacterium Alkanolyticum β -Xylosidase and Xyloside ABC Transporter in Corynebacterium Glutamicum'. *Applied and Environmental Microbiology* 81 (12): 4173–83. <https://doi.org/10.1128/AEM.00792-15>.
- Watzlawick, Hildegard, Kambiz Morabbi Heravi, and Josef Altenbuchner. 2016. 'Role of the ganSPQAB Operon in Degradation of Galactan by Bacillus Subtilis'. *Journal of Bacteriology* 198 (20): 2887–96. <https://doi.org/10.1128/JB.00468-16>.
- Webb, Alexander J., Karen A. Homer, and Arthur H.F. Hosie. 2007. 'A Phosphoenolpyruvate-Dependent Phosphotransferase System Is the Principal Maltose Transporter in Streptococcus Mutans'. *Journal of Bacteriology* 189 (8): 3322–27. <https://doi.org/10.1128/JB.01633-06>.
- . 2008. 'Two Closely Related ABC Transporters in Streptococcus Mutans Are Involved in Disaccharide and/or Oligosaccharide Uptake'. *Journal of Bacteriology* 190 (1): 168–78. <https://doi.org/10.1128/JB.01509-07>.
- Wefers, Daniel, Jia Dong, Ahmed M. Abdel-Hamid, Hans Müller Paul, Gabriel V. Pereira, Yejun Han, Dylan Dodd, et al. 2017. 'Enzymatic Mechanism for Arabinan Degradation and Transport in the Thermophilic Bacterium Caldanaerobius Polysaccharolyticus'. *Applied and Environmental Microbiology* 83 (18). <https://doi.org/10.1128/AEM.00794-17>.
- Wei, Xiao, Yanhong Guo, Changlin Shao, Zhongke Sun, Daria Zhurina, Dawei Liu, Wei Liu, et al. 2012. 'Fructose Uptake in Bifidobacterium Longum NCC2705 Is Mediated by an

- ATP-Binding Cassette Transporter'. *Journal of Biological Chemistry* 287 (1): 357–67. <https://doi.org/10.1074/jbc.M111.266213>.
- Wezel, G. P. Van, J. White, M. J. Bibb, and P. W. Postma. 1997. 'The malEFG Gene Cluster of *Streptomyces Coelicolor* A3(2): Characterization, Disruption and Transcriptional Analysis'. *Molecular and General Genetics* 254 (5): 604–8. <https://doi.org/10.1007/s004380050458>.
- White, Joshua B. R., Augustinas Silale, Matthew Feasey, Tiaan Heunis, Yiling Zhu, Hong Zheng, Akshada Gajbhiye, et al. 2022. 'Outer Membrane Utilisomes Mediate Oligosaccharide Uptake in Gut Bacteroidetes'. Preprint. Microbiology. <https://doi.org/10.1101/2022.08.15.503959>.
- . 2023. 'Outer Membrane Utilisomes Mediate Glycan Uptake in Gut Bacteroidetes'. *Nature*, June, 1–7. <https://doi.org/10.1038/s41586-023-06146-w>.
- Wilkins, Stephan. 2015. 'Structure and Mechanism of ABC Transporters'. *F1000Prime Reports* 7 (14): 426–31. <https://doi.org/10.1016/j.sbi.2004.06.005>.
- Wong, Jeremy P. H., Noémie Chillier, Michaela Fischer-Stettler, Samuel C. Zeeman, Tom J. Battin, and Alexandre Persat. 2024. 'Bacteroides Thetaiotaomicron Metabolic Activity Decreases with Polysaccharide Molecular Weight'. *mBio* 15 (3): e02599-23. <https://doi.org/10.1128/mbio.02599-23>.
- Wu, Meng Chuan, Ying Chun Chen, Tzu Lung Lin, Pei Fang Hsieh, and Jin Town Wang. 2012. 'Cellobiose-Specific Phosphotransferase System of *Klebsiella Pneumoniae* and Its Importance in Biofilm Formation and Virulence'. *Infection and Immunity* 80 (7): 2464–72. <https://doi.org/10.1128/IAI.06247-11>.
- Xiao, X., F. Wang, A. Satio, J. Majka, A. Schlösser, and H. Schrempf. 2002. 'The Novel *Streptomyces Olivaceoviridis* ABC Transporter Ngc Mediates Uptake of N-Acetylglucosamine and N, N'-Diacetylchitobiose'. *Molecular Genetics and Genomics* 267 (4): 429–39. <https://doi.org/10.1007/s00438-002-0640-2>.
- Xu, Tao, Xuanyu Tao, Hongxi He, Megan L Kempfer, Siping Zhang, Xiaochun Liu, Jun Wang, et al. 2023. 'Functional and Structural Diversification of Incomplete Phosphotransferase System in Cellulose-Degrading Clostridia'. *The ISME Journal* 17 (6): 823–35. <https://doi.org/10.1038/s41396-023-01392-2>.

- Yamamoto, H., M. Serizawa, J. Thompson, and J. Sekiguchi. 2001. 'Regulation of the Glv Operon in *Bacillus Subtilis*: YfiA (GlvR) Is a Positive Regulator of the Operon That Is Repressed through CcpA and Cre'. *Journal of Bacteriology* 183 (17): 5110–21. <https://doi.org/10.1128/JB.183.17.5110-5121.2001>.
- Yan, Fei, Sheng Dong, Ya-Jun Liu, Xingzhe Yao, Chao Chen, Yan Xiao, Edward A. Bayer, et al. 2022. 'Deciphering Cellodextrin and Glucose Uptake in *Clostridium Thermocellum*'. Edited by Mikiyasu Sakanaka and Sang Yup Lee. *mBio* 13 (5): e01476-22. <https://doi.org/10.1128/mbio.01476-22>.
- Yan, Nieng. 2015. 'Structural Biology of the Major Facilitator Superfamily Transporters'. *Annual Review of Biophysics* 44 (1): 257–83. <https://doi.org/10.1146/annurev-biophys-060414-033901>.
- Yang, Jian, Liping Zhang, Shanshan Lin, Wei Li, Chen Liu, Jingheng Yan, Shubo Li, and Lijuan Long. 2023. 'Structural Insights of a SusD-like Protein in Marine Bacteroidetes Bacteria Reveal the Molecular Basis for Chitin Recognition and Acquisition'. *The FEBS Journal* n/a (n/a). <https://doi.org/10.1111/febs.16974>.
- Zhang, Xuejun C., Yan Zhao, Jie Heng, and Daohua Jiang. 2015. 'Energy Coupling Mechanisms of MFS Transporters'. *Protein Science* 24 (10): 1560–79. <https://doi.org/10.1002/pro.2759>.
- Zhang, Yubo, Xianfeng Zhong, Siyun Su, and Guidong Huang. 2020. 'Discovery of Novel Prebiotic Carbohydrates and Sugar Mimics of BLMsmE, a Solute-Binding Protein of the ABC Transporter from *Bacillus Licheniformis*'. *The Journal of Physical Chemistry B* 124 (45): 9996–10006. <https://doi.org/10.1021/acs.jpcc.0c05583>.
- Zhao, Tingting, Han Yue, Junfeng Peng, Yingmin Nie, Longzhen Wu, Tingting Li, Wei Niu, et al. 2023. 'Degradation of Xylan by Human Gut Bacteroides Xylanisolvens XB1A'. *Carbohydrate Polymers* 315 (September):121005. <https://doi.org/10.1016/j.carbpol.2023.121005>.
- Zhu, Yuan-ting, Shao-lan Zhang, Qian Ma, Yong Zuo, Rui-tong Li, Jian-xia Yang, Rui-yu Zhu, and Li-hong Ge. 2022. 'Global Genome and Comparative Transcriptomic Analysis Reveal the Inulin Consumption Strategy of *Lactiplantibacillus Plantarum* QS7T'. *Food*

Zúñiga, Manuel, Iñaki Comas, Raquel Linaje, Vicente Monedero, María Jesús Yebra, Carlos David Esteban, Josef Deutscher, Gaspar Pérez-Martínez, and Fernando González-Candelas. 2005. 'Horizontal Gene Transfer in the Molecular Evolution of Mannose PTS Transporters'. *Molecular Biology and Evolution* 22 (8): 1673–85.
<https://doi.org/10.1093/molbev/msi163>.

Engineering of oligosaccharide transporters

I. Introduction

As mentioned in the previous chapter, there are four different superfamilies of oligosaccharide transporters described to date, which present different mechanisms and architectures: the ATP-binding cassettes (ABC), the major facilitator superfamily (MFS) transporters, the phosphotransferase systems (PTS) and the TonB-dependent transporters (TBDT). These membrane-bound proteins play crucial roles in various biological processes, facilitating the transport of oligosaccharides across cell membranes to support cell growth and to be transformed into bioproducts of interest for biorefinery and the food, feed, health and cosmetics industries.

Transporter engineering, defined here as introducing mutations in the target proteins, has received increasing attention in the past decades, with a double objective. In some studies, the goal was to improve the transport of specific oligosaccharides and increase the level of utilization in host cells used for fermentation, or to engineer the specificity of transporters to render the strains able to grow on cheap substrates. In other studies, the goal was to get knowledge on the structure-function relationships of these proteins. Indeed, as explained in the previous chapter, only a few structural data are available for these complex systems containing transmembrane proteins and which present weak sequence homologies, although for those belonging to a same superfamily. In this review, we provide an overview of the protein engineering studies targeting oligosaccharide transporters, with a focus on the different methods used, their limitations and advantages to investigate the mechanism of action of these protein complexes at the molecular level, or to improve their functional efficiency.

II. Strategies of oligosaccharide transporter engineering

Several review articles have recently summarized the engineering works on oligosaccharide transporters (Lv et al. 2022), particularly in *E. coli* (Alva et al. 2020). However, none has summarized the studies carried out regardless of the microorganisms or family of transporters targeted. The materials of the present review chapter were retrieved by querying the Google Scholar and PubMed databases, using the following keywords: oligosaccharide transporter engineering, oligosaccharide transporter mutation, oligosaccharide transporter evolution, MFS transporter mutation, ABC transporter mutation, PTS transporter mutation, mutant SusD protein, SusD mutation, SusC mutation and TonB transporter mutation. First, we focused the search for publications from 2003. Then, we expanded the search for earlier references based

on their citations in articles published over the last two decades. The studies of interest are listed in Table 1.

Table 1: Summary of the strategies of oligosaccharide transporter engineering for 4 different transporter families coming from fungi and bacteria, including the biochemical and structural information of parental proteins and the new functions of engineered transporters.

Protein name (TCDB Accessi Number)	PDB 3D structure	Classification	Organism	Substrate	Method	Selection of Mutations	Mutations	Impact of the mutations	Reference
Yeasts									
Mal11 (P53048)	-	Major facilitator superfamily (MFS)	<i>Saccharomyces cerevisiae</i>	Maltose, sucrose and maltotriose	<i>In cellulo</i> Adaptive laboratory evolution	Mutagenic strain, spontaneous mutagenesis	Mal11 _{QAQ/A384D}	Restored proton coupled transport mechanism	(Henderson et al. 2021)
							Mal11 _{QNA/A515D}	Restored proton coupled transport mechanism	
							Mal11 _{QAQ/R504C}	Increased sucrose uptake rate	
							Mal11 _{ANA/W376S}	Fully uncoupled transport, such as sugar uniporters	
			<i>Saccharomyces cerevisiae</i>	Maltose	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on MSA and structural model	E120A, E120Q, D123N, E167A, E167Q	Decreased coupling efficiency	(Henderson and Poolman 2017)
							E120A/E167A, E120A/E167Q, E120Q/E167A, E120Q/E167Q	Internal localization	
							D123N/E167Q	No maltose uptake, peripheral localization	
							D123A/E120A, D123A/E120Q, D123N/E120A,	Reduced proton-coupled maltose transport, peripheral localization	

					D123N/E120Q, D123A/E167A, D123A/E167Q, D123N/E167A				
					D123A/E120Q/E167A	No maltose uptake			
					D123A/E120A/E167A	Internal localization			
					D123A/E120A/E167Q, D123A/E120Q/E167Q, D123N/E120A/E167A, D123N/E120A/E167Q, D123N/E120Q/E167A, D123N/E120Q/E167Q	Abolished proton-coupled maltose transport, peripheral localization			
CDT-2 (Q7SD12)	-	MFS	<i>Neurospora crassa</i> (cellobiose transporter recombinantly expressed in <i>Saccharomyces cerevisiae</i>)	Cellobiose	<i>In cellulo</i> Adaptive laboratory evolution	Mutagenic strain, spontaneous mutagenesis	N306I	Enhanced cellobiose fermentation	(Kim et al. 2018)
					<i>In cellulo</i> Adaptive laboratory evolution	Error-prone PCR, Mutagenic strain	I96N/T487A	Increased cellobiose consumption rate and ethanol productivity, improved tolerance under low pH	(Oh et al. 2017)
					<i>In vitro</i> directed evolution	Error-prone PCR	Q207H/N311H/I505T	Increased cellobiose consumption rate and ethanol productivity	(Lian et al. 2014)
					<i>In vitro</i> rational engineering	Site-directed mutagenesis based on MSA and structural model	N311H, I505T, Q207H Q207H/I505T, N311H/I505T	Increased growth rate Increased growth rate compared to that of the	

									single mutants and lower than the triple mutants
CDT-1 (Q7SCU1)	-	MFS	<i>Neurospora crassa</i> (cellobiose transporter recombinantly expressed in <i>Saccharomyces cerevisiae</i>)	Cellobiose	<i>In vitro</i> directed evolution	Error-prone PCR	D433G, C82S	Increased growth rate on cellobiose, improved cellobiose utilization and ethanol productivity	(Eriksen et al. 2013)
					<i>In vitro</i> rational engineering	Site-directed mutagenesis based on structural model	D433G, C82S	Increased cellodextrin uptake	
HXT2.4	-	MFS	<i>Scheffersomyces stipiti</i> (cellobiose transporter recombinantly expressed in <i>Saccharomyces cerevisiae</i>)	Cellobiose	<i>In cellulo</i> Adaptive laboratory evolution	Mutagenic strain, spontaneous mutagenesis	A291D	Improved cellobiose consumption rate and ethanol productivity, facilitated cellotriose and cellotetraose transport	(Ha et al. 2013)
					<i>In vitro</i> rational engineering	Saturation mutagenesis	A291E	Improved cellobiose consumption rate and ethanol productivity	
							A291S or A291G	Slowed cellobiose consumption	
							A291K, A291R	Improved cellobiose consumption	
Agt1p (P53048)	-	MFS (SP)	<i>Saccharomyces cerevisiae</i>	α -glucosides: maltotriose, maltose	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on MSA and structural model	I505T	Abolished maltotriose transport	(Smit et al. 2008)
							V549A	Decreased maltotriose transport ability	

							T557S	Abolished maltotriose transport	
Bacteria									
TonB-dependent transporter (TBDT) Superfamily									
AqSusD	8J0P; 8JXZ	TBDT superfamily	<i>Aquimarina sp.</i> <i>SCSIO 21287</i>	Chitin	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on crystallographic structure	W64A, Q67M, L72D, F87H, W102A, W274A, D276L	No binding to chitotetraose	(Yang et al. 2023)
Bt1762 SusD	6Z8I		<i>Bacteroides</i> <i>thetaiotaomicron</i>	β -2,6 fructo- oligosacchari des	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on crystallographic structure	W85A, C298A Y395A D41A/N43A/D67A/W85A/C2 98A/R368A/Y395A	Abolished levan binding Reduced affinity to levan Grew after a prolonged lag phase of ~8 h	(Gray et al. 2021)
BoSGBP _{ML} G-A	6E60; 6DMF; 6E61		<i>Bacteroides ovatus</i>	β (1,3)/ β (1,4) -glucans	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on crystallographic structure	W77A, W350A Y266A, W353A	Abolished binding Diminished binding	(Tamura et al. 2019)
GMSusD	6GCZ		<i>Bacteroidetes-</i> <i>Gramella sp.</i> <i>MAR_2010_102</i>	branched laminarin, pustulan	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on crystallographic structure	W287A, W290A, W323A	Reduced binding	(Mystkowska et al. 2018)
SGBP-A	5E75; 5E76		<i>Bacteroides ovatus</i>	Xyloglucan	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on crystallographic structure	W82A W306A W82A/W283A/W306A	Increased binding Abolished XyG binding Abolished XyG binding	(Tauzin, Kwiatkowski, et al. 2016)
Major facilitator superfamily (MFS)									

MelB _{St} (A7ZUZ0)	7L17; 7L16	MFS	<i>Salmonella</i> <i>typhimurium</i>	Melibiose, 4- nitrophenyl- α -D- galactoside, n-dodecyl- β - D- melibioside	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on crystallographic structure and MD simulations	T121C T121P, T121A, T121G T121S T121R, T121W T121D, T121N, T121L, T121M	Abolished H ⁺ -coupled and reduced Na ⁺ -coupled transport activity Abolished Na ⁺ coupled transport Reduced Na ⁺ -coupled transport activity, decreased melibiose uptake Abolished all three modes of transport, no melibiose uptake Inhibited all the three modes of transport	(Katsube et al. 2022)
						Site-directed mutagenesis based on crystallographic structure	D59C, D59A	Abolished Na ⁺ -coupled binding and transport	(Guan and Hariharan 2021)
						Saturation mutagenesis (Cysteine- scanning)	Single-Cys mutants: 476	Altered uptake	(Markham et al. 2021)
						Site-directed mutagenesis based on crystallographic structure	G117C, G117S, G117N P148L, G117C/P148L, G117S/P148L, G117N/P148L G117A	Increased melibiose transport Decreased co-substrate binding, inhibited transport Little difference in either co-substrate binding or	(Jakkula and Guan 2012) (Guan et al. 2012)

								stimulation of melibiose transport by Na ⁺ or Li ⁺	
							G117R	Reduced transport, decreased Na ⁺ binding	
							G117P, G117W	Reduced transport, decreased Na ⁺ binding	
MelB _{Ec} (A7ZUZ0)	7L17; 7L16	MF	<i>Escherichia coli</i>	Galactosides: melibiose, <i>p</i> -nitrophenyl- α -D-6- [³ H]galactopyranoside or 6- <i>O</i> - α -D- [³ H]galactopyranosyl-d-glucose, fluorescent sugar analog, methyl- β -D-galactopyranoside	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on crystallographic structure	R149C, R149K, R149Q	Bound Na ⁺ and melibiose in proteoliposomes, where the protein is disposed inside-out	(Lin et al. 2013)
							R149C	Restrained the orientation of transporter to an inward-open conformation, reduced the frequency of access of outer substrates to the binding sites, impaired active transport	
							R141C	Decreased Na ⁺ affinity and substrate transport	(León et al. 2009)
							R141C	Decreased transport rate	(Dayem et al. 2003)
							R149C	Impaired sugar binding	
							R141K	Retained ion-coupled binding and transport	
							R141Q	Abolished H ⁺ -, Na ⁺ -, or Li ⁺ - coupled transport	
							R149Q, R149K	Reduced sugar affinity	

<i>In cellulo</i> random mutagenesis	Mutagenic strain, spontaneous mutagenesis	R149C/V343G	Partial recovery of transport activity	
<i>In vitro</i> rational engineering	Site-directed mutagenesis based on crystallographic structure	K377V K377D D59K D59K/K377D	Reduced transport Reduced transport Reduced transport Restored transport	(Franco et al. 2001)
<i>In cellulo</i> random mutagenesis	Mutagenic strain, spontaneous mutagenesis	K377V/D59K, K377V/D59N	Restored transport	
<i>In vitro</i> rational engineering	Site-directed mutagenesis based on cysteine substitution (helix I residue 17-37)	G17C, K18C, D19C, Y32C, T34C, D35C F20C, A21C, I22C, G23C, I24C, V25C, Y26C, M27C, Y28C, M30C, Y31C	Abolished transport Strong transport inhibition by PCMBs (Cys containing helices facing the aqueous channel)	(Ding and Wilson 2001)
<i>In cellulo</i> random mutagenesis	Mutagenic strain, spontaneous mutagenesis	K18C/D19N (helix I), K18C/M123R (helix IV), Y31C/S234L and Y31C/L236F (helix VII).	Improved transport activity	
<i>In cellulo</i> random mutagenesis	Mutagenic strain, spontaneous mutagenesis	K377C/D59A, A383C/I22S, A383C/I22N, A383C/F20L, F385C/R441C, F385C/A388G, F385C/I387M, L391C/I22N, L391C/I22T, L391C/A152S, L391C/D19E, G395C/D19E, G395C/I22N	Improve transport activity	(Ding and Wilson 2000)

LacY (P02920)	2Y5Y; 2V8N; 2CFQ; 1PV7	MFS	<i>Escherichia coli</i>	Lactose, melibiose, 4- nitrophenyl- β-D- galactopyran oside, methyl-thio- β-D- galactopyran oside, o- nitro-phenyl-	<i>In vitro</i>	Site-directed	K377C, G379C, A383C,	Lost transport activity	(Ding and
					rational	mutagenesis based	F385C, L391C, G395C,		Wilson 2000)
					engineering	on cysteine	Y396C		
						substitution	S380C, A381C, A384C, I387C, A388C, L391C	Strong transport inhibition by PCMBs (Cys containing helices facing the aqueous channel)	
					<i>In vitro</i>	Site-directed	R52C, D55C, D59C, G63C	Abolished Na ⁺ -coupled	(Matsuzaki et
					rational	mutagenesis based		transport	al. 1999)
					engineering	on cysteine	V44C, G45C, L47C, F48C,	Inhibited transport	
						substitution	L49C, A51C, A56C	inhibition by PCMBs (Cys	
						(residue 39-66)		containing helices facing the aqueous channel)	
					<i>In cellulo</i>	Mutagenic strain,	70 single mutants in sites:	Impaired transport, altered	(Botfield and
					random	spontaneous	D15-I18, Y116-P122, V342-	Li ⁺ recognition, changed	Wilson 1988)
					mutagenesis	mutagenesis	I348, A364-G374, I61, A236	sugar recognition	
					<i>In vitro</i>	Site-directed	V352H, C353H, F354H,	Altered membrane protein	(Vitrac et al.
					rational	mutagenesis based	C355H, F356H, F357H,	organization	2020)
					engineering	on crystallographic	Q359H	Altered membrane protein	
						structure		organization	
						Site-directed	G46A/G262A, G46S/G262S	Increased affinity,	(Jiang et al.
						mutagenesis based		decreased transport	2016)
						on crystallographic	G46T/G262T, G46V/G262V,	Increased affinity,	
						structure and MD	G46L/G262L, G46N/G262N,	decreased or abolished	
						simulations	G46Q/G262Q, G46Y/G262Y,	transport	
							G46W/G262W, G46E/G262E,		
							G46K/G262K		

b-D- galactopyran oside, [¹⁴ C]TMG β,D- galactopyran osyl 1-thio- β,D- galactopyran oside	Site-directed	E126A/N245C,	Blocked sugar-induced	(Zhou et al.
	mutagenesis based	R144Q/N245C	periplasmic opening	2012)
	on crystallographic	E269A, H322N, Y236F	Decreased closing of the	
	structure and MD		cytoplasmic cavity in the	
	simulations		presence of sugar	
		E269A/H322N/Y236F	Decreased affinity,	
			abolished transport	
		G24C/C154G	Increased transport with	(Zhou et al.
			binding	2011)
		P28C P28C/V158C	Inhibited transport	
		C154G	Increased affinity,	(Jiang et al.
			decreased transport	2012; Nie et
				al. 2008)
		D68A, D68C, D68N, D68E	Abolished transportn	(Liu et al.
		K131R	Increased transport	2010)
		K131A, K131C	Reduced transport	
		D68/K131A, D68C/K131C	Rescued transport in D68	
			mutants	
		D240E	Decreased affinity and	(Guan and
			transport	Kaback 2009)
		I40C/N245C	Affected in gating the	(Zhou et al.
			periplasmic cavity	2009)
		I40C/N245C, I32C/N245C,	Abolished transport	(Zhou et al.
		T45C/N245C		2008)
		A122C	Abolished transport	(Guan et al.
		A122F, A122Y	Decreased transport	2002)
		A122S	Decreased transport	

<i>In cellulo</i> random mutagenesis	Spontaneous mutagenesis, mutagenic strain	G64S/A50T, G64S/C234F, G64S/T266I, G64S/V367E, P280L/G46C, P280L/L216Q, P280L/C333G, P280L/F354C, P280L/G370V	Increased transport	(Cain et al. 2000)
		G46C, G46S, F49L, A50T, L212Q, L216Q, S233P, C333G, F354C, G370C, G370S, G370V	Increased transport	(Pazdernik et al. 1997)
		G64C/P28S, G64S/P28S, G64V/P28S, G64C/P28L, G64S/P28L, G64V/P28L, G64V/P28T, G64S/F29S, G64S/A50T, G64C/C154G, G64C/C234F, G64S/C234F, G64S/Q241L, G64S/F261V, G64C/T266I, G64S/V367E, G64V/A369P	Restored transport	(Jessen- Marshall et al. 1997)
		D68S/T45R, D68X/C234W, D68X/C234F, D68T/Q241L, D68X/F247V, D68T/G257D, D68S/S366F, D68T/V367E, D68T/A369P, D68T/ins5 (5- Codon insertion)	Restored transport	(Jessen- Marshall and Brooker 1996)
<i>In vitro</i> rational engineering	Site-directed mutagenesis based on MSA	G64S, G64C, G64P, G64V, G64R	Reduced transport	(Jessen- Marshall et al. 1995)
		D68E, D68H, D68N, D68Y, D68T, D68S, D68A	Reduced transport	

		G71S, G71C, G71P, G71V, G71E, G71L, G71W	Reduced transport	
<i>In cellulo</i> random mutagenesis	Chemical agent (β- thio- <i>o</i> - nitrophenylgalacto- side, TONPG), mutagenic strain	M1I	Affected specificity	(Shinnick and Varela 2002)
		G370C	Affected specificity	
		R144W	Reduced transport	
		Residues 387–392 deletion	Affected specificity	
<i>In cellulo</i> random mutagenesis	Chemical agent (TONPG), mutagenic strain	M23I, W151L, G257D, A295D, G377V	Abolished transport	(Varela et al. 2000)
<i>In cellulo</i> random mutagenesis	Chemical agent (Nitrosoguanidine) , mutagenic strain	A177V	Enhanced recognition of L- arabinose	(Goswitz and Brooker 1993)
		A177V	Decreased transport	(Brooker 1991)
		A177V/K319N	Decreased transport	

<i>In cellulo</i>	Mutagenic strain,	A177T, A177V	Increased transport rates	(Brooker et
random	spontaneous			al. 1985;
mutagenesis	mutagenesis			Brooker and
				Wilson 1985)

		Y236F, Y236D, Y236S, Y236H	Abolished transport	
<i>In cellulo</i> random mutagenesis	UV irradiation, mutagenic strain	LacY ^{mal}	Generated a new transport system, increased affinity	(Shuman and Beckwith 1979)
<i>In cellulo</i> random mutagenesis	Mutagenic strain, Spontaneous mutagenesis	N272L/Q60L, N272L/V229G, N272L/V229A, N272L/F354V, Y272D, L272V, N272L/V229A, N272L/F354V T265Y/M276Y/F261L, T265Y/M276Y/Q60P, T265Y/M276Y/L70H, T265Y/M276Y/C355G	Restored transport Restored transport Partly restored transport	(Green et al. 2003)
<i>In vitro</i> rational engineering	Site-directed mutagenesis based on Gly-scanning mutagenesis	G262A/E269D T266G/M267G/E269D G262A/T266G/M267G/E269 D	Increased transport Increased transport Increased transport	(Weinglass et al. 2002)
<i>In vitro</i> rational engineering	Site-directed mutagenesis based on MSA and structural model	F261D, N272Q, N272Y, T265Q, T265Y, M276A, T265Y/M276Y, G268V	Decreased transport	(Green and Brooker 2001)
<i>In vitro</i> rational engineering	Site-directed mutagenesis based on MSA and structural model	S56Y, S56L, S56Q, Q60A S56Y/Y56C, S56L/L56P, S56Q/Q56P, S56Q/Q60P, S56Q/V229G, S56L/Q60L, S56L/V229G, S56L/V229A, Q60V/Y26H, Q60V/S53F,	Decreased transport Restored transport	(Green et al. 2000)

		Q60A/S53Y, Q60A/S56L, Q60A/Q359L		
<i>In cellulo</i>	Mutagenic strain,	S56Y/Y56C, S56L/L56P	Restored transport	
random	Spontaneous	S56Q/Q60P, S56Q/V229G,	Restored transport	
mutagenesis	mutagenesis	S56L/Q60L, S56L/V229G, S56L/V229A, Q60V/Y26H, Q60V/S53F, Q60A/S53Y, Q60A/S56L, Q60A/Q359L		
<i>In vitro</i>	Site-directed	Single-Cys mutants in sites: 2-	Decreased transport	(Frillingos et
rational	mutagenesis based	402		al. 1998)
engineering	on Cys-scanning			
	mutagenesis			
<i>In vitro</i>	Site-directed	Single-Cys mutants in sites:	Decreased transport	(Frillingos et
rational	mutagenesis based	residue 34-74		al. 1997)
engineering	on Cys-scanning	H35C	Decreased transport	
	mutagenesis	G64C, D68C	Decreased transport	
<i>In vitro</i>	Site-directed	E269D	Increased transport	(Franco and
rational	mutagenesis based	E269G, E269A, E269Q	Abolished transport	Brooker
engineering	on MSA	E325D	Increased transport	1994)
		E325G, E325A, E325S, E325Q	Abolished transport	
<i>In vitro</i>	Site-directed	E269C, E269Q	Abolished transport	(Ujwal et al.
rational	mutagenesis based	E269D	Abolished transport	1994)
engineering	on MSA and			
	structural model			

CscB (P30000)	-	MFS	<i>Escherichia coli</i>	Sucrose	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on structure model	Q353H	Altered membrane protein organization	(Vitrac et al. 2020)
					<i>In cellulo</i> random mutagenesis	Mutagenic strain, Spontaneous mutagenesis		Increased transport	(Jahreis et al. 2002)
RafB (P16552)	-	MFS	<i>Escherichia coli</i>	Raffinose, maltose	<i>In cellulo</i> random mutagenesis	Mutagenic strain, Spontaneous mutagenesis	V35A, I391S, S138D/S139L/G389A	Affected specificity	(Van Camp et al. 2007)
MelY (P96517)	-	MFS	<i>Enterobacter cloacae</i>	α -galactoside sugar melibiose, maltose	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on MSA and structural model	A182P L88P, L91P	Increased affinity Increased affinity and transport	(Shinnick et al. 2003)
ATP-binding Cassette (ABC) family									
AlgQ2 (Q9KWT6)	6JHX	ABC	<i>Sphingomonas sp.</i> <i>Al</i>	Acidic polysacchari de alginate	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on crystallographic structure	D179A/E180A	Reduced affinity for calcium	(Okumura et al. 2021)
<i>B/MnBP1</i> ,	6I5R; 6I5V; 6I5W	ABC	<i>Bifidobacterium</i> <i>malis subsp. Lactis</i> <i>ATCC27673</i>	β -manno- oligosacchari de (MnOS) (DP3-6)	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on crystallographic structure	N63G G61N	Reduced sugar affinity Increased sugar affinity	(Ejby et al. 2019)
<i>B/MnBP2</i>	6FUV								
MalE (P0AEX9)	1ANF; 1MPB; 1MPC; 1MPD	ABC	<i>Escherichia coli</i>	Maltose, maltodextrin	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on crystallographic structure and previous random	G32D I33P MalE31 G32D/I33P	No change Decreased transport Abolished transport, defective folding	(Saul et al. 2003; Betton and Hofnung 1996)

	insertion			
	mutagenesis			
	Site-directed mutagenesis based on crystallographic structure	Y210S, Y210L, Y210N, A213I, E214K	Reduced transport	(Szmelcman et al. 1997)
	Site-directed mutagenesis based on crystallographic structure	S337C	Reduced affinity	(Gilardi et al. 1994)
<i>In cellulo</i>	Chemical agent	W230R	Inhibited transport	(Hor and
random	(ethylmethane	Y210, G13D, D14Y	Increased affinity	Shuman
mutagenesis	sulfonate),	Y210C	Increased transport	1993;
	mutagenic strain			Treptow and Shuman 1988)
<i>In vitro</i>	Spontaneous	MalE:133-VPI:93-103	Increased transport	(Rodseth et
random	mutagenesis,			al. 1990)
insertion	mutagenic strain			
mutagenesis				
<i>In vitro</i>	Site-directed	W62A, W230A, W232A,	Decreased affinity and	(Martineau et
rational	mutagenesis based	W340A	transport	al. 1990)
engineering	on crystallographic structure	W158A	Reduced the chemotactic response towards maltose	
<i>In vitro</i>	Random insertion;	Class II mutants (deletion in	Affected transport	(Duplay et al.
random	<i>in vitro</i>	322, 341)		1987; Duplay
insertion	recombinations	Class III mutants (deletion in	Decreased transport	and
mutagenesis		91, 140, 333)		

							Class IV mutants (deletion in 301, 346)	Abolished transport	Szmecman 1987)
							Other mutants (Insertion in 120, 183 and 220)	Abolished binding	
MalF (P02916)	--	ABC	<i>Escherichia coli</i>	Maltose	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on MD simulations	L334W, F336L, G338R, W378C, V442A, A502V, N505I	Conformational state closed to the transition state	(Khare et al. 2009)
						Site-directed mutagenesis based on MSA	MalF540 Q99(Am) E130(Am), K275(Am), Y317(Am)	Altered specificity Reduced transport	(Merino and Shuman 1998)
				Maltose, lactose	<i>In cellulo</i> random mutagenesis	Spontaneous mutagenesis	MalF515 L334W	Altered substrate specificity, transport lactose in a saturable manner	(Merino and Shuman 1997)
				Maltose	<i>In cellulo</i> random mutagenesis	UV irradiation, mutagenic strain, spontaneous mutagenesis	L334W, F336L, G338R, W378C, V442A, A502V, N505I, N505S	Altered interaction with MBP	(Covitz et al. 1994; Treptow and Shuman 1985)
				Maltose, maltoheptaose	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on structure model	L372R	Altered specificity	(Ehrle et al. 1996)
MalG (P68183)	2R6G; 3FH6	ABC	<i>Escherichia coli</i>	Maltose	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on crystallographic structure	MalG511: L135F, L135H, I154S	High basal ATPase activity, shift its equilibrium towards the transition state	(Bajaj et al. 2018)

						Site-directed mutagenesis based on MD simulations	L135F, L135H, I154S	Conformational state closed to the transition state	(Khare et al. 2009)
						Site-directed mutagenesis based on structure model	31-codon-insertion	Affected assembly and transport	(Nelson and Traxler 1998)
					<i>In cellulo</i> random mutagenesis	UV irradiation, mutagenic strain, spontaneous mutagenesis	L135F, L135H, I154S	Altered interaction with MBP	(Covitz et al. 1994; Treptow and Shuman 1985)
MalK (P68187)	-	ABC	<i>Escherichia coli</i>	Maltose	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on structure model	N262D, L268Q, L291Q, E350Q E308Q, G340A, F355, G346S/G346D, D297G/D297N	Decreased transport	(Böhm et al. 2002)
							N262D, L268Q, L291Q, E350Q	Altered expression	
							L268Q	Altered expression	
					<i>In vitro</i> random insertion mutegenesis	Spontaneous mutagenesis, mutagenic strain	31-codon-insertion	AlteredMalFGK assembly	(Lippincott and Traxler 1997)
					<i>In cellulo</i> random mutagenesis	Chemical agent (hydroxylamine), UV irradiation, mutagenic strain	W267G, G346S	Normal transport	(Kühnau et al. 1991)
							G137A, ΔGln R140, D158N	Abolished transport	
							S282L	Altered specificity	
Sugar phosphotransferase system (PTS)									
	-	PTS					C414A	Decreased transport	

spr0276-82, spr0505			<i>Streptococcus pneumoniae</i>	β -glucosides: cellobiose, gentiobiose, arbutin, amygdalin and aesculin	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on MSA	H577A	Decreased transport	(Boianelli et al. 2012)
IIB ^{Chb} = IIB ^{Cel}	1E2B	PTS	<i>Escherichia coli</i>	Cellobiose, chitin disaccharide (GlcNAc) ₂	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on crystallographic structure	C10S	Stabilized conformation	(Keyhani et al. 2000; Ab et al. 1994)
LacS	-	PTS	<i>Streptococcus thermophilus</i>	α - and β - galactosides	<i>In vitro</i> rational engineering	Site-directed mutagenesis based on MSA	H376Q	Decreased transport and altered affinity	(Poolman et al. 1992)

Table 2: Phylum and order distribution of the engineered oligosaccharide transporters in microorganisms

Phyla	Organism	Number
Ascomycota	<i>Saccharomyces cerevisiae</i>	8
Pseudomonadota	<i>Salmonella typhimurium</i>	5
	<i>Escherichia coli</i>	55
	<i>Sphingomonas sp. AI</i>	1
	<i>Enterobacter cloacae</i>	1
Actinomycetota	<i>Bifidobacterium animalis subsp. Lactis ATCC27673</i>	1
Bacillota	<i>Streptococcus pneumoniae</i>	1
	<i>Streptococcus thermophilus</i>	1
Bacteroidota	<i>Bacteroides strain (Aquimarina sp. SCSIO 21287)</i>	1
	<i>Bacteroides ovatus</i>	3
	<i>Bacteroides thetaiotaomicron</i>	1
	<i>Bacteroidetes-Gramella sp. MAR_2010_102</i>	1

Table 3: Transporter families of the engineered oligosaccharide transporters

Transporter Family	Number
Major facilitator superfamily (MFS)	48
ATP-binding Cassette (ABC) family	23
Sugar phosphotransferase system (PTS)	3
TonB-dependent transporter (TBDT) SusC/D	5

***In cellulo* random mutagenesis**

The first strategy used to engineer transporters is based on the generation of sequence diversity by randomly introducing mutations in the genome of the target strain, or specially in their transporter encoding genes. The desired phenotype is screened (herein by selective growth on the target oligosaccharides, or by measurement of radiolabeled oligosaccharide uptake) before sequencing the best clones in order to identify the location of the mutations.

The natural ability of bacteria to mutate can be used to randomly introduce mutations in genomes, or even better, in multicopy plasmids to facilitate DNA sequencing. Under selection pressure on solid plates, the *E. coli* mutation frequency is indeed between 10^{-8} and 10^{-9} (Brooker and Wilson 1985). The spontaneous mutating ability of *E. coli* was used in several studies to change the specificity of *E. coli* oligosaccharide transporters, conferring them the ability to transport another substrate than the native one. For examples, in order to change the specificity of any transporter to maltose, the used *E. coli* strain must have undergone mutations or deletions in the *malB* locus so that the uptake of maltose via the normal maltose transport system is blocked. By applying a selection pressure on maltose over the course of weeks in solid media, Van Camp *et al.* succeeded in changing the specificity of the raffinose transporter of *E. coli* (RafB, from the MFS superfamily) to maltose (Van Camp *et al.* 2007). The five residues involved in substrate specificity were identified. Other studies resulted in the change of substrate specificity of the β -galactoside (including lactose and melibiose) transport system (LacY, from the MFS superfamily) to maltose. By applying a selection pressure on maltose, Shuman and Beckwith and further Brooker & Wilson (1985) changed the specificity of LacY to maltose, and identified the mutations involved in substrate recognition (Shuman and Beckwith 1979; Brooker and Wilson 1985; Brooker *et al.* 1985; Brooker 1991). Besides, very smartly, mutants of the same β -galactoside transport system were isolated by their ability to grow on minimal plates in the presence of the toxic lactose analogue β -thio-*o*-nitrophenylgalactoside (TONPG). The TONPG-resistant mutants, affected in substrate recognition, were then identified (Varela *et al.* 2000; Shinnick and Varela 2002). Another strategy was used by Jessen-Marshall *et al.* and then by Pazdernik *et al.* to identify the residues involved in galactoside transport and further characterized them (Jessen-Marshall *et al.* 1995; Jessen-Marshall and Brooker 1996; Jessen-Marshall *et al.* 1997; Pazdernik *et al.* 1997; Cain *et al.* 2000). These authors used, as parental strains, previously identified LacY mutants with strongly diminished lactose transport activity, to isolate, after several days of selection on melibiose, variants with restored transport function, thus containing mutations that suppressed

the effects of initial deleterious mutations. Finally, in the Merino and Shuman's study, the previously identified LacY mutants affected in their ability to transport lactose were used to select mutants of the maltose transport system that were able to grow on lactose, and to decipher the mechanisms involved in triggering the transport process by the complex multiprotein maltose transport system (Merino and Shuman 1997; Merino and Shuman 1998).

In order to increase the spontaneous mutation frequency of *E. coli*, UV irradiation or chemical agents can be used (Figure 1). The introduction of random mutations can be achieved by targeting the transporter gene only, cloned into a plasmid exposed to mutagens, or by transforming a mutagenic *E. coli* strain (a strain with defects in DNA repair machinery, leading to more spontaneous mutations) with the plasmid of interest, or even by using linker-insertion mutagenesis. This kind of mutagenic strains increases spontaneous mutation rates over 1000-fold by affecting the activity of the editing 3'-exonuclease (Degnen and Cox 1974; DiFrancesco et al. 1984). Several pioneering engineering studies were performed using these techniques to investigate the mechanism of action of the maltose permease of *E. coli*. Regions of MBP (MalE) that may be important for substrate binding, transport activity and chemotaxis have been identified by linker-insertion mutagenesis (Duplay and Szmecman 1987; Duplay et al. 1987) and by screening for mutations that retain transport activity. Kühnau *et al.* used three different techniques (UV, hydroxylamine and a mutagenic strain) to randomly generate mutations in the maltose permease encoding gene and identify the mutations involved in the ability of *E. coli* to transport maltose (Kühnau et al. 1991). Besides, Treptow and Shuman first used UV to restore the ability of *E. coli* to transport maltose in the absence of MBP (Treptow and Shuman 1985). The revertants were obtained thanks to mutations in the inner-membrane components of the maltose transport system MalF and MalG. Then, the same authors used strain treatment by ethylmethane sulfonate to identify the residues involved in the interactions between the maltose-binding protein MalE and MalG (Treptow and Shuman 1988). Hor and Shuman further used a mutant strain affected in maltodextrin transport and a mutagenic strain to obtain revertants and to restore the growth phenotype (Hor and Shuman 1993). The effect of one of these mutations was further deeply investigated by Shilton *et al.*, based on crystallography and SAXS analysis (Shilton et al. 1996).

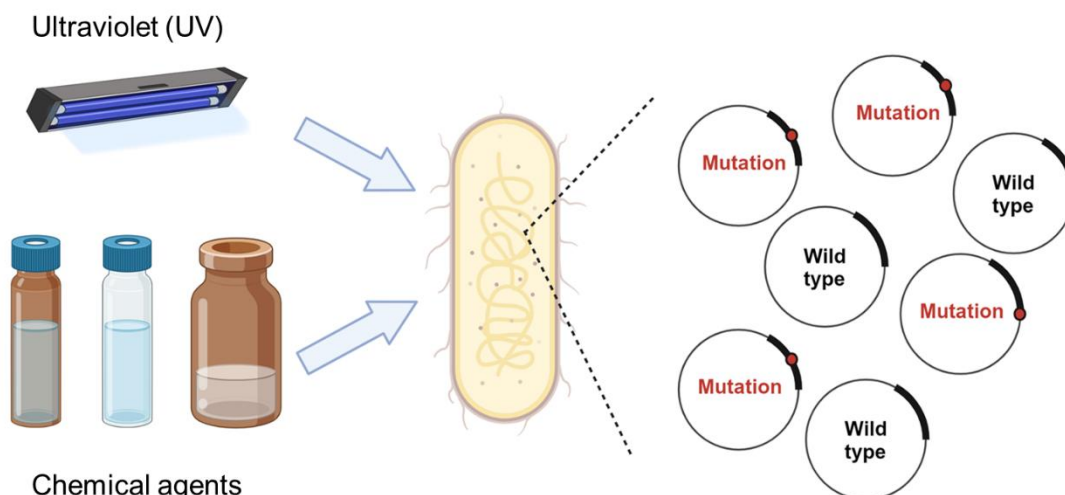


Figure 1: Spontaneous mutations using physical or chemical agents.

Besides, Botfield and Wilson identified amino acids involved in sugar specificity of the melibiose carrier of *E. coli*, belonging to the MFS superfamily (Botfield and Wilson 1988). In this study, the authors transformed a mutagenic *E. coli* strain with the melibiose permease expression plasmid. The sequencing of 70 mutants allowed the authors to identify residues involved in sugar positioning within the binding site and in cation coordination clusters, while the melibiose transporter structure was unknown, since this protein had never been purified to homogeneity. In another study, they used random mutagenesis in *E. coli* to restore the ability of the melibiose carrier after having been mutated by cysteine substitution mutagenesis (Ding and Wilson 2000).

Finally, increasing the mutation rate of *E. coli* by the use of nitrosoguanidine was also used to change the specificity of the β -galactoside transport system LacY to arabinose (Goswitz and Brooker 1993). In another study focused on LacY, the isolation of suppressor mutations was performed, starting from mutants previously obtained by rational engineering (Green and Brooker 2001), which were markedly defective for lactose transport velocity (Green et al. 2003).

***In cellulo* Adaptive Laboratory Evolution**

The spontaneous ability of living cells to generate mutations can be exploited to accelerate the obtention of mutants with desired properties. The natural evolutionary process can indeed be forced and accelerated in Adaptive Laboratory Evolution (ALE) experiments, in which a microorganism is cultivated under defined conditions for prolonged periods of time, in the range

of weeks to years, which allows the selection of improved phenotypes (Dragosits and Mattanovich 2013) (Figure 2).

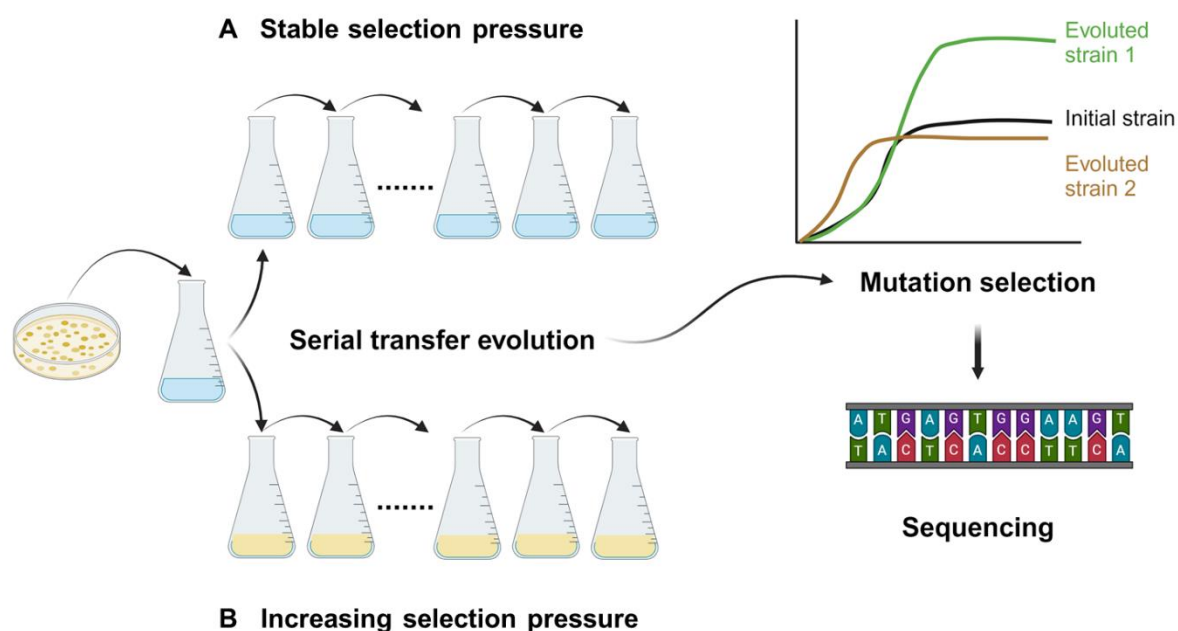


Figure 2: An overview of adaptive laboratory evolution (ALE).

(A) Serial cultures in a desired growth environment allowing natural selection to enrich the culture in mutant strains. (B) Serial cultures using increasing selection pressure.

The first ALE experiments were already performed about a hundred years ago and many other ones in the last 25 years, with *E. coli* and *Saccharomyces cerevisiae* as the most prominent organisms under investigation. If we did not find any example of successful transporter engineering in *E. coli* using ALE, this approach has been used in a few cases to evolve recombinant *S. cerevisiae* strains expressing on plasmids genes encoding recombinant hexose transporters conferring to the host yeast the ability to ferment cellobiose (Ha et al. 2013; Kim et al. 2018). Serial subcultures of the initial strain (which harbored a poor cellobiose fermentation ability) on cellobiose as a sole carbon source were performed to increase the rate of cellobiose fermentation (Oh et al. 2016). The specific rates of biomass increase on cellobiose increased drastically during serial transfers, due to single mutations in the recombinant cellobiose transporters. In a last and very smart ALE study, Oh *et al.* first generated sequence diversity in the recombinant cellobiose transporter by error-prone PCR (Oh et al. 2017). Then, from the mutated genes, they constructed a GFP-transporter fusion gene library, that was screened using FACS (Fluorescence-activated-cell-sorting) for the mutants expressing at high

level the mutated transporters. In a third step, the FACS-sorted mutants were used in serial subcultures in liquid medium containing cellobiose, in order to select the best variant, that contained 2 mutations conferring to the host strain an increased ability to ferment cellobiose in acidic conditions.

ALE was also used to engineer the native transporter of α -glucosides in *S. cerevisiae*. Starting from triple mutants previously obtained (Henderson and Poolman 2017) that were unable to grow on sucrose, sequential batch cultivation was performed in shake flasks with sucrose as the sole carbon source. Within three months, all six evolution cultures were able to grow on sucrose, enabling the authors to propose a mechanistic model of proton-coupled sugar transport (Henderson et al. 2021).

***In vitro* rational engineering**

Besides the *in cellulo* random engineering approaches, protein engineering can also be performed *in vitro*, using site-directed mutagenesis (Figure 3) or saturation mutagenesis (Figure 4), which consists in replacing a target amino acid by the 19 others. In this case, the amino acids to target are identified by analysing multiple sequence alignments (MSA), structural models established based on crystallization and X-ray diffraction, or by structure modeling, and even after performing molecular dynamics simulations. The properties of the mutants are then assessed *in cellulo*, after transforming a recombinant host with a plasmid expressing the variants and testing the growth ability of the resulting strains. Mutant properties can also be assessed *in vitro*, after purifying the soluble components of complex transporter systems usually produced in *E. coli*, using oligosaccharide binding assays such as differential scanning calorimetry, isothermal titration calorimetry, affinity gel electrophoresis and/or STD NMR.

Rational engineering has been extensively used to investigate the mode of action of oligosaccharide transporters and to identify the molecular determinants of specificity towards the target substrate.

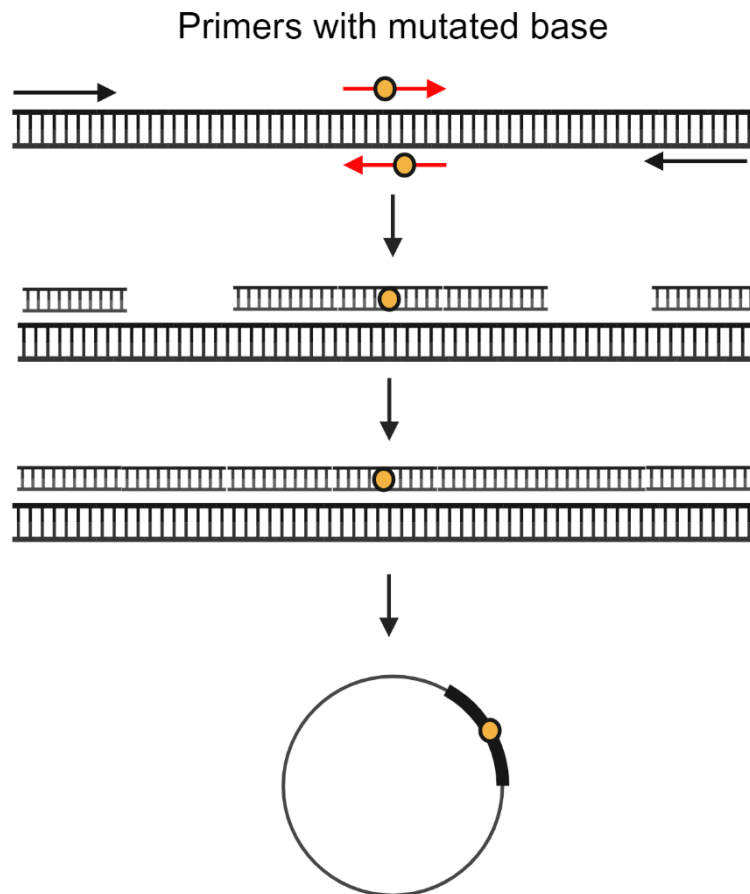


Figure 3: An overview of site-directed mutagenesis.

The red arrows indicate the designed primers which include the mutation site. The yellow dot represents the targeted mutated base.

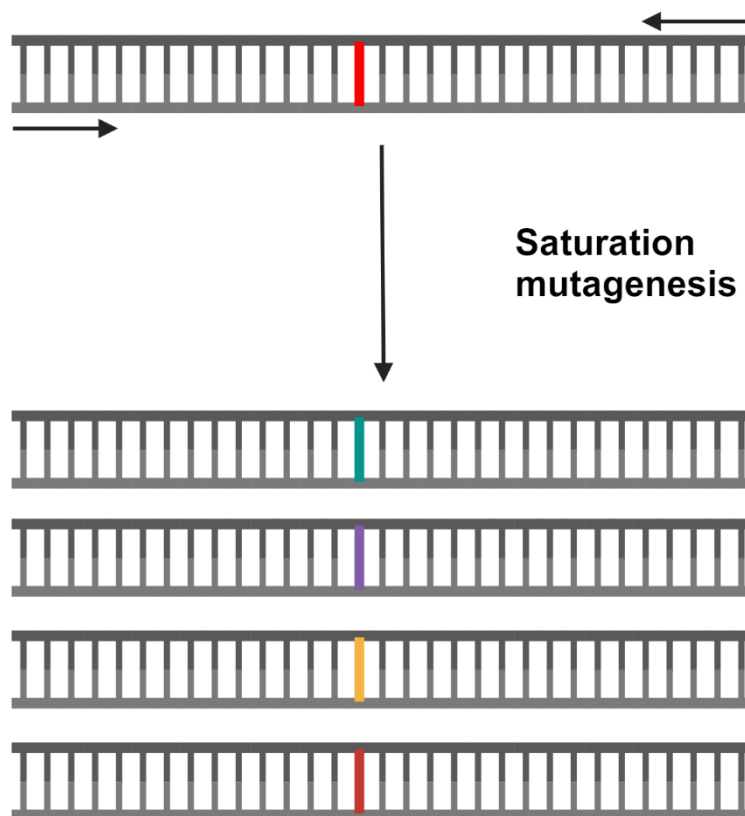


Figure 4: An overview of saturation mutagenesis.

The red line segment indicates the targeted base. The other colored line segments represent substituted bases.

We found only three studies targeting yeast transporters. Ha *et al.* used saturation mutagenesis to analyze the effect of mutations previously obtained in a recombinant MFS cellobiose transporter expressed in *S. cerevisiae* using ALE, as described before (Ha et al. 2013). Besides, Smit *et al.* and Henderson & Poolman used MSA and structural modeling to identify the residues of the native *S. cerevisiae* maltose transporter involved in the energy coupling mechanism of transport (Henderson and Poolman 2017; Smit et al. 2008).

Many other examples of rational engineering were performed to analyze the relationships between structure, specificity and uptake efficiency of bacterial transporters, especially in *E. coli*. Regarding MFS transporters, the mechanism of melibiose transporter involving MelB, in particular its interactions with Na^+ and the sugar, was extensively studied by rational engineering based on structural analysis (Lin et al. 2013; León et al. 2009; Dayem et al. 2003; Franco et al. 2001), or using cysteine substitution mutagenesis (Ding and Wilson 2000; 2000;

2001; Matsuzaki et al. 1999). The LacY β -galactoside transporter was also extensively studied, based on MSA (Jessen-Marshall et al. 1995), protein structure and molecular dynamics analyses (Jiang et al. 2016; Zhou et al. 2012; Jiang et al. 2012; Zhou et al. 2011; Liu et al. 2010; Guan and Kaback 2009; Zhou et al. 2009; Zhou et al. 2008; Green et al. 2003; Guan et al. 2002), or using Gly (Weinglass et al. 2002) or Cys (Frillingos et al. 1998) scanning mutagenesis, targeting secondary structure elements shown to be important for functional conformation of the protein. Furthermore, Vitrac *et al.* showed that the changes in the transporter structure caused by modifications in the *E. coli* membrane lipid composition can be altered by single-point mutations in LacY and CscB (Vitrac et al. 2020). In this study, the mutated residues were chosen by LacY structure analysis and CscB structural modeling based on its homolog, LacY. This work highlighted the adaptability of MFS transporters to their environment.

The functional role of the different components of the ABC transporter (MalE, F, G, K) of *E. coli* involved in maltose and maltodextrin uptake were also extensively studied using rational engineering. Prior to the solving of its X-ray structure, the maltose-binding protein (MBP, MalE) was first engineered by Martineau *et al.* (Martineau et al. 1990), by targeting mutations based on previous random insertion mutagenesis results (Betton and Hofnung 1996; Duplay and Szmelcman 1987; Duplay et al. 1987). Then, rational engineering was based on crystallographic protein structure analysis (Saul et al. 2003; Szmelcman et al. 1997; Gilardi et al. 1994). The two inner membrane proteins of the transporter (MalF and MalG) were also subjected to intensive site-directed mutagenesis studies, based on MSA (Merino and Shuman 1998), protein structure predictions (Nelson and Traxler 1998; Ehrle et al. 1996), prior the solving of the X-ray structures for these proteins, or based on X-ray structure analysis (Bajaj et al. 2018) and molecular dynamic simulations (Khare et al. 2009) after the structural resolution of these proteins. Structural modeling from the X-ray structure of an homolog protein of *Thermococcus litoralis* (PDB ID 1G29) was also used to choose the site-directed mutagenesis targets in the last component of the system, the cytoplasmic protein MalK (Böhm et al. 2002).

Finally, the IIB component of the PTS involved in cellobiose and chitobiose utilization by *E. coli* was mutated before any NMR (Ab et al. 1994) or X-ray (Van Montfort et al. 1997) crystallography based structural analyses, in order to stabilize this soluble photocarrier. To this aim, the only Cys residue of the sequence was mutated in Ser, in order to avoid any oxidation of this active site residue. The effect of the mutation was further investigated by Keyhani *et al.*, who showed that the single-point Cys mutation stabilizes the PTS by forming a stable complex between IIB^{Chb} and phospho-IIA^{Chb} (Keyhani et al. 2000).

In other Pseudomonadata, a few site-directed mutagenesis studies of the melibiose MFS transporter of *Salmonella typhimurium* were performed, this protein being used as a model for homolog transporters of human cells, acting as lung tumor suppressor. Most of the mutated residues were identified by X-ray structure (Guan and Hariharan 2021; Guan et al. 2012; Jakkula and Guan 2012) and molecular dynamic (Katsube et al. 2022) analyses, but Cys-scanning (Markham et al. 2021) was also performed, in order to identify or confirm the role of the Cys residues in cation binding, conformational stability and translocation.

Besides, the solute binding proteins of the ABC transporters involved in the transport of alignate oligosaccharides by an *Sphingomonas* species, were engineered by site-directed mutagenesis based on structural analysis to identify the residues involved in calcium binding (Okumura et al. 2021).

Finally, the specificity of the melibiose MFS transporter of an *Enterobacter* species has been engineered using site-directed mutagenesis targeting residues identified by MSA and structural modeling, showing that substrate specificity of MFS transporters function is determined by subtle structural differences (Shinnick et al. 2003).

Several site-directed mutagenesis studies were also quite extensively performed on Bacteroidota transporters, in order to investigate the relationships between the structure of SusD-like components of the TonB-dependant transporter family and their polysaccharide / oligosaccharide binding abilities. These studies were initiated after the resolution of the X-ray structures of these proteins in complex with their ligands. The authors characterized the effects of the mutation either *in vitro*, after protein purification, or *in cellulose* in the native *Bacteroides* species, after replacing the wild-type SusD by their non-binding mutant (Yang et al. 2023; Gray et al. 2021; Tamura et al. 2019; Mystkowska et al. 2018; Tauzin, Kwiatkowski, et al. 2016). These studies allowed the authors to identify the importance of tryptophan residues, in particular, for binding, as well as the crucial role of SusD proteins, even in their non-binding forms, for the transport functionality of SusC, through protein-protein interactions necessary to the pedal bin mechanism described in the previous chapter.

Regarding other bacterial phyla, very few site-directed mutagenesis studies have been published. In Bacillota, the PTS cellobiose and α -/ β -galactosides transporters of two *Streptococcus* species were engineered in order to understand the mechanism of oligosaccharide uptake, while no tridimensional data are available. In these cases, MSA were used to target conserved residues in the predicted phosphorylation sites of the PTS components EIIA and EIIB (Boianelli et al. 2012) and in the carrier domain of EIIA (Poolman et al. 1992). Finally, in Actinomycetota, we found a study focused on the engineering of the solute binding protein of the ABC transporter

involved in the transport of mannooligosaccharides by the human gut *Bifidobacterium animalis* species, in which a structural analysis was used to identify the residues involved in substrate binding (Ejby et al. 2019).

***In vitro* Directed Evolution**

Despite the success of the rational approach for understanding the structure-function relationships of oligosaccharide transporters and their mechanism of action, some key amino acids for transport efficiency or protein stability are impossible to identify by this approach, because they are located in scattered areas, far from the oligosaccharide binding site. Furthermore, rational engineering, in general, suffers from the lack of available structural data, which is the case, in particular, for transmembrane transporters whose structural resolution is particularly difficult. The random approach performed *in vitro*, or combinatorial engineering, is therefore complementary to rational engineering, in so far as it allows, on the one hand, to quickly obtain variants with improved properties, and on the other hand, to provide information on the nature of crucial amino acids, even in the case of proteins whose three-dimensional structure is not known. Directed molecular evolution consists of i) the construction of a library of variants of the gene coding for the protein of interest by mutagenesis and/or random recombination; ii) screening of the obtained library to isolate the improved variant(s) which could serve as a basis for a new campaign of mutations/recombinations/screening until a mutant with optimal properties is obtained (Figure 5).

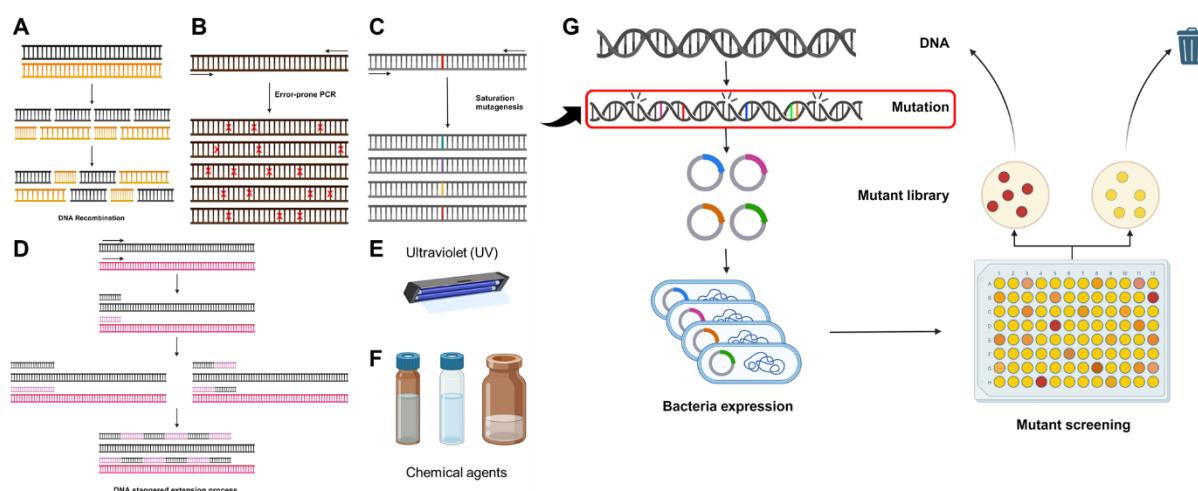


Figure 5: An overview of the directed evolution process.

The left panel displays several random mutagenesis approaches: (A) DNA shuffling; (B) Error-prone PCR; (C) Saturation mutagenesis; (D) DNA staggered extension process (StEP); (E) Physical mutagenesis; (F) Chemical mutagenesis; (G) An overview of the screening of mutant libraries.

In most cases, the generation of molecular diversity is carried out by error-prone PCR (Figure 5B) (Chen and Arnold 1993), using an infidel DNA polymerase, or by playing with concentration of DNA polymerase, time polymerization, concentration of divalent cations (Mg^{2+} , Mn^{2+}) or concentration of dNTPs (Fromant et al. 1995; Cadwell and Joyce 1992). The first *in vitro* recombination technique, known as “DNA shuffling” (Figure 5A), was developed by the Pim Stemmer's team in 1994 (Stemmer 1994). It is based on random fragmentation with DNase I of previously mutated genes, for example by error-prone PCR, and on the reassembly of the fragments by PCR. The other method, called “DNA staggered extension process (StEP)” (Figure 5D) was developed in 1996 by the Frances Arnold's team (Zhao et al. 1998). It is based on PCR cycles with interrupted extension, so that each PCR product from one variant hybridizes with that of another variant, so as to obtain real chimeric libraries. StEP and DNA shuffling are, to date, the most commonly used methods for combinatorial protein engineering. Given the size of the variant libraries generated (from 10^4 to 10^8 mutants), directed molecular evolution requires the development of high-throughput screening techniques for the isolation of variants improved for a given property. With regard to oligosaccharide transporters, library screening could be easy, since it is based on the selection of mutants capable of growing on a selective medium containing the target oligosaccharide as the sole carbon source.

The sole directed evolution studies focused on oligosaccharide transporters were performed in *S. cerevisiae*, and did not include any DNA shuffling or StEP. The generation of sequence diversity was performed by error-prone PCR. At each evolution round, a novel cycle of error-prone PCR was undertaken from the variants obtained in the preceding cycle, before screening the library. Eriksen *et al.* and Lian *et al.* evolved the recombinant cellobiose transporters CDT1 and CDT2 from *N. crassa*, that confer to the host yeast the ability to ferment cellobiose, with the intention to introduce one amino acid substitution per protein on average (Lian *et al.* 2014; Eriksen *et al.* 2013). Due to limitations of the transformation efficiency into the yeast host, the library sizes were only of $\sim 10^4$ mutants, screened on cellobiose containing plates. The maximal cellobiose uptake rate (increased by about 2 times) was obtained after only two (CDT1) or three (CDT2) rounds of directed evolution, resulting in the combination of two (CDT1) or three (CDT2) mutations in the transporter encoding gene.

Only one study was performed to randomly introduce, *in vitro*, insertions and deletions in the gene encoding the MBP protein (MalE) (Rodseth *et al.* 1990; Duplay and Szmecman 1987; Duplay *et al.* 1987) and the MalK component (Lippincott and Traxler 1997) of the *E. coli* maltose transport system. However, only one round of mutagenesis was performed, which is not directed evolution. Nevertheless, these studies allowed the authors to identify determinants of proper protein folding (Betton and Hofnung 1996).

III. Summary and future challenges in oligosaccharide transporter engineering

- The immense majority of oligosaccharide transporter engineering studies have been focused on *E. coli* substrate uptake systems, and, to a lesser extent, on those of *S. cerevisiae* (Tables 2 and 3). These two species are indeed the most used chassis in fermentation processes, as well as for synthetic biology.
- In most studies, the objectives were to improve the understanding of the relationships between the structure of the transporters, their oligosaccharide binding specificity, uptake efficiency and interactions between the different components of the multiprotein complexes. In some cases, the goal was to improve the *E. coli* or *S. cerevisiae* growth on low cost substrates such as plant derived oligosaccharides (e.g. cellobiose, maltose or sucrose) for biorefinery and the production of bulk compounds where the carbon feedstock can make up 70 % of the total variable costs of industrial processes. Many studies were also focused on the lactose (a cheap substrate from cow milk) transporter

of *E. coli*, which is the preferred platform for the production of human milk oligosaccharides for human infant nutrition. Other studies targeting prominent human gut bacteria from the Bacteroidota and Actinomycetota phyla, or human pathogens like *Salmonella* (a primary enteric pathogen issued from contaminated food or water) or *Streptococcus* (promoting dental caries) species were rather oriented towards molecular ecology, with no possible direct application due to regulation regarding the use of GMOs in human health control.

- Rational engineering based on site-directed mutagenesis is the most used approach for the study of bacterial oligosaccharide transporters. The studies are mainly focused on the soluble components of the ABC, PTS and SusC/D systems, that are easily produced as recombinant proteins in *E. coli*, the preferred chassis strain for mutagenesis studies thanks to the genetic tools available. Studies performed on the transmembrane proteins involved in substrate channelling across cell membranes requires the use of allelic exchange in the native host strains, which requires particular expertise and is restricted to only a few species for which genetic tools are available, such as Bacteroidota. Nevertheless, the CIMEs team recently demonstrated (Wang et al. 2022; 2020; Tauzin, Laville, et al. 2016) that recombinant transmembrane oligosaccharide transporters from both Gram negative and Gram positive bacteria could be produced as functional forms in *E. coli*, as single-protein transporters (MFS) or multi-proteic systems (SusC/D, PTS, ABC). Despite the fact that co-expression in *E. coli* of the different PULs-encoded proteins involved in oligosaccharide utilization is spurious, yielding a success rate of around 15 % (Wang et al. 2022), this opens new perspectives in specifically engineering transmembrane domains by keeping the soluble components of the transporters in their wild-type forms, or in simultaneous engineering of the different proteins of recombinant oligosaccharide utilization pathways. However, the rational engineering strategy is limited by the structural data available. Just analysing MSA is not sufficient in most cases, since sequence identity is low (e.g. lower than 35 % for SusD-like proteins) (Bolam and Van Den Berg 2018) among oligosaccharide transporters, even for those from the same superfamily and targeting the same oligosaccharides.
- Combinatorial random mutagenesis approaches, performed *in vitro* (directed evolution) or *in cellulo* (adaptive laboratory evolution), have been used in a limited number of studies to circumvent the limitations of rational engineering for mutagenesis target selection. We found only a few studies in *S. cerevisiae*, while *E. coli* is usually the host of choice for directed evolution, thanks to its high transformation efficiency which is

required to generate large mutant libraries, or to evolve anabolic pathways using ALE. The lack of combinatorial random mutagenesis studies in *E. coli* might be due to the difficulty of co-expression of the different components of multi-proteic transport systems, as explained above. Nevertheless, the proof of concept established regarding the successful co-expression in *E. coli* of multiple genes of metagenomic PULs opens the way to random mutagenesis studies targeting multi-proteic transporter systems.

- Oligosaccharide transporter engineering requires to be able to screen the desired phenotype, which is herein either oligosaccharide binding or uptake efficiency. Binding assays based on differential scanning calorimetry, isothermal titration calorimetry and affinity electrophoresis gels require protein purification, which significantly restrains the number of mutants to analyze and limits the use of random approaches requiring high-throughput screening approaches. STD-NMR is compatible with the use of crude-cell extracts but NMR data acquisition and spectra analysis in themselves are low throughput techniques. High-throughput pull-down assays using crude cell extracts and sugar quantification in the supernatant could be developed, on the basis of the low-throughput, quantitative assay previously developed by Wang *et al.* (Wang et al. 2020). Nevertheless, oligosaccharide uptake ability and efficiency are much easier to quantify at high-throughput than binding, since these traits might be directly correlated to the growth ability on a selective medium containing the target oligosaccharide as sole carbon source. Positive selection is easy to handle in solid plates or micro-plates, but screening of a 10^5 clones library requires grams of oligosaccharides. This is not problematic if the target substrate is a cheap, plant or milk derived oligosaccharide such as cellobiose, sucrose or lactose. Nevertheless, the cost of most structurally defined oligosaccharides, purified to be deprived of monosaccharides, usually exceeds hundreds of euros per gram, and reaches thousands of euros per gram for those of which the synthesis is difficult, such as oligosaccharides mimicking human glycans (e.g. mucin-like oligosaccharides, gangliosides, blood antigens). Consequently, random mutagenesis approaches targeting transporters of such expensive oligosaccharides cannot be used since positive selection of hit clones cannot be performed using conventional assays. Pico-liter droplet-microfluidics approaches, such as recently developed in the CIMEs team (Potocki-Veronese, et al. 2023; Potocki-Veronese, et al. 2023) allow positive selection of 10^6 strains in only one hour and using 1 mg of substrate. Such technology thus represents a great opportunity for directed evolution of transporters, either to investigate further their mechanisms and their structure-

specificity-efficiency relationships, to increase their substrate uptake potential to boost synthetic biology and fermentation processes, or to explore the potential of oligosaccharide utilization by human-associated microorganisms. Ultimately, the objective could be to control the functioning and composition of the human microbiota by the use of oligosaccharides as nutraceuticals, that could be specifically utilized by beneficial species, or by developing specific transporter inhibitors as drugs to prevent the growth of pathogens.

IV. REFERENCES

- Ab, Eiso, Gea K. Schuurman-Wolters, Milton H. Saier, Jonathan Reizer, Michel Jacuinod, Peter Roepstorff, Klaas Dijkstra, Ruud M. Scheek, and George T. Robillard. 1994. 'Enzyme IIB Bcellobiose of the Phosphoenol-Pyruvate-Dependent Phosphotransferase System of Escherichia Coli: Backbone Assignment and Secondary Structure Determined by Three-Dimensional NMR Spectroscopy'. *Protein Science* 3 (2): 282–90. <https://doi.org/10.1002/pro.5560030212>.
- Alva, Alma, Andrea Sabido-Ramos, Adelfo Escalante, and Francisco Bolívar. 2020. 'New Insights into Transport Capability of Sugars and Its Impact on Growth from Novel Mutants of Escherichia Coli'. *Applied Microbiology and Biotechnology* 104 (4): 1463–79. <https://doi.org/10.1007/s00253-019-10335-x>.
- Bajaj, Ruchika, Mariana I. Park, Cynthia V. Stauffacher, and Amy L. Davidson. 2018. 'Conformational Dynamics in Binding-Protein-Independent Mutant of the Escherichia Coli Maltose Transporter, MalG511 and Its Interaction with Maltose Binding Protein'. *Biochemistry* 57 (20): 3003–15. <https://doi.org/10.1021/acs.biochem.8b00266>.
- Betton, Jean-Michel, and Maurice Hofnung. 1996. 'Folding of a Mutant Maltose-Binding Protein of Escherichia Coli Which Forms Inclusion Bodies (*)'. *Journal of Biological Chemistry* 271 (14): 8046–52. <https://doi.org/10.1074/jbc.271.14.8046>.
- Böhm, Alex, Joachim Diez, Kay Diederichs, Wolfram Welte, and Winfried Boos. 2002. 'Structural Model of MalK, the ABC Subunit of the Maltose Transporter of Escherichia Coli: IMPLICATIONS FOR Mal GENE REGULATION, INDUCER EXCLUSION, AND SUBUNIT ASSEMBLY*'. *Journal of Biological Chemistry* 277 (5): 3708–17. <https://doi.org/10.1074/jbc.M107905200>.
- Boianelli, Alessandro, Alessandro Bidossi, Luciana Gualdi, Laura Mulas, Chiara Mocenni, Gianni Pozzi, Antonio Vicino, and Marco R. Oggioni. 2012. 'A Non-Linear Deterministic Model for Regulation of Diauxic Lag on Cellobiose by the Pneumococcal Multidomain Transcriptional Regulator CelR'. Edited by Valerie De Crécy-Lagard. *PLoS ONE* 7 (10): e47393. <https://doi.org/10.1371/journal.pone.0047393>.

- Bolam, David N, and Bert Van Den Berg. 2018. 'TonB-Dependent Transport by the Gut Microbiota: Novel Aspects of an Old Problem'. *Current Opinion in Structural Biology* 51 (August):35–43. <https://doi.org/10.1016/j.sbi.2018.03.001>.
- Botfield, M C, and T H Wilson. 1988. 'Mutations That Simultaneously Alter Both Sugar and Cation Specificity in the Melibiose Carrier of Escherichia Coli.' *Journal of Biological Chemistry* 263 (26): 12909–15. [https://doi.org/10.1016/S0021-9258\(18\)37647-6](https://doi.org/10.1016/S0021-9258(18)37647-6).
- Brooker, R J. 1991. 'An Analysis of Lactose Permease "Sugar Specificity" Mutations Which Also Affect the Coupling between Proton and Lactose Transport'. *Journal of Biological Chemistry* 266 (7): 4131–38. [https://doi.org/10.1016/S0021-9258\(20\)64296-X](https://doi.org/10.1016/S0021-9258(20)64296-X).
- Brooker, R J, K Fiebig, and T H Wilson. 1985. 'Characterization of Lactose Carrier Mutants Which Transport Maltose.' *Journal of Biological Chemistry* 260 (30): 16181–86. [https://doi.org/10.1016/S0021-9258\(17\)36218-X](https://doi.org/10.1016/S0021-9258(17)36218-X).
- Brooker, R J, and T H Wilson. 1985. 'Isolation and Nucleotide Sequencing of Lactose Carrier Mutants That Transport Maltose.' *Proceedings of the National Academy of Sciences of the United States of America* 82 (12): 3959–63. <https://doi.org/10.1073/pnas.82.12.3959>.
- Cadwell, R. C., and G. F. Joyce. 1992. 'Randomization of Genes by PCR Mutagenesis.' *Genome Research* 2 (1): 28–33. <https://doi.org/10.1101/gr.2.1.28>.
- Cain, S.M., E.A. Matzke, and R.J. Brooker. 2000. 'The Conserved Motif in Hydrophilic Loop 2/3 and Loop 8/9 of the Lactose Permease of Escherichia Coli. Analysis of Suppressor Mutations'. *The Journal of Membrane Biology* 176 (2): 159–68. <https://doi.org/10.1007/s00232001085>.
- Chen, K, and F H Arnold. 1993. 'Tuning the Activity of an Enzyme for Unusual Environments: Sequential Random Mutagenesis of Subtilisin E for Catalysis in Dimethylformamide.' *Proceedings of the National Academy of Sciences* 90 (12): 5618–22. <https://doi.org/10.1073/pnas.90.12.5618>.
- Covitz, K M, C H Panagiotidis, L I Hor, M Reyes, N A Treptow, and H A Shuman. 1994. 'Mutations That Alter the Transmembrane Signalling Pathway in an ATP Binding Cassette (ABC) Transporter.' *The EMBO Journal* 13 (7): 1752–59. <https://doi.org/10.1002/j.1460-2075.1994.tb06439.x>.

- Dayem, Manal Abdel, Cécile Basquin, Thierry Pourcher, Emmanuelle Cordat, and Gérard Leblanc. 2003. 'Cytoplasmic Loop Connecting Helices IV and V of the Melibiose Permease from Escherichia Coli Is Involved in the Process of Na⁺-Coupled Sugar Translocation *'. *Journal of Biological Chemistry* 278 (3): 1518–24. <https://doi.org/10.1074/jbc.M210053200>.
- Degnen, Gerald E., and Edward C. Cox. 1974. 'Conditional Mutator Gene in Escherichia Coli: Isolation, Mapping, and Effector Studies'. *Journal of Bacteriology* 117 (2): 477–87. <https://doi.org/10.1128/jb.117.2.477-487.1974>.
- DiFrancesco, R., S. K. Bhatnagar, A. Brown, and M. J. Bessman. 1984. 'The Interaction of DNA Polymerase III and the Product of the Escherichia Coli Mutator Gene, mutD.' *Journal of Biological Chemistry* 259 (9): 5567–73. [https://doi.org/10.1016/S0021-9258\(18\)91051-3](https://doi.org/10.1016/S0021-9258(18)91051-3).
- Ding, Ping Z., and T. Hastings Wilson. 2000. 'Physiological Evidence for an Interaction between Helix XI and Helices I, II, and V in the Melibiose Carrier of Escherichia Coli'. *Biochemical and Biophysical Research Communications* 268 (2): 409–13. <https://doi.org/10.1006/bbrc.2000.2149>.
- . 2000. 'The Melibiose Carrier of Escherichia Coli: Cysteine Substitutions for Individual Residues in Helix XI'. *The Journal of Membrane Biology* 174 (2): 135–40. <https://doi.org/10.1007/s002320001038>.
- . 2001. 'Cysteine Mutagenesis of the Amino Acid Residues of Transmembrane Helix I in the Melibiose Carrier of Escherichia Coli'. *Biochemistry* 40 (18): 5506–10. <https://doi.org/10.1021/bi002761k>.
- Dragosits, Martin, and Diethard Mattanovich. 2013. 'Adaptive Laboratory Evolution – Principles and Applications for Biotechnology'. *Microbial Cell Factories* 12 (1): 64. <https://doi.org/10.1186/1475-2859-12-64>.
- Duplay, Pascale, and Sevec Szmelcman. 1987. 'Silent and Functional Changes in the Periplasmic Maltose-Binding Protein of Escherichia Coli K12'. [https://doi.org/10.1016/0022-2836\(87\)90244-0](https://doi.org/10.1016/0022-2836(87)90244-0).
- Duplay, Pascale, Sevec Szmelcman, Hugues Bedouelle, and Maurice Hofnung. 1987. 'Silent and Functional Changes in the Periplasmic Maltose-Binding Protein of *Escherichia Coli*

- K12'. *Journal of Molecular Biology* 194 (4): 663–73. [https://doi.org/10.1016/0022-2836\(87\)90243-9](https://doi.org/10.1016/0022-2836(87)90243-9).
- Ehrle, R, C Pick, R Ulrich, E Hofmann, and M Ehrmann. 1996. 'Characterization of Transmembrane Domains 6, 7, and 8 of MalF by Mutational Analysis.' *Journal of Bacteriology* 178 (8): 2255–62. <https://doi.org/10.1128/jb.178.8.2255-2262.1996>.
- Ejby, M., A. Guskov, M. J. Pichler, G. C. Zanten, E. Schoof, W. Saburi, D. J. Slotboom, and M. Abou Hachem. 2019. 'Two Binding Proteins of the ABC Transporter That Confers Growth of *Bifidobacterium Animalis* Subsp. *Lactis* ATCC27673 on B-mannan Possess Distinct Manno-oligosaccharide-binding Profiles'. *Molecular Microbiology* 112 (1): 114–30. <https://doi.org/10.1111/mmi.14257>.
- Eriksen, Dawn T, Pei Chiun Helen Hsieh, Patrick Lynn, and Huimin Zhao. 2013. 'Directed Evolution of a Cellobiose Utilization Pathway in *Saccharomyces Cerevisiae* by Simultaneously Engineering Multiple Proteins'. *Microbial Cell Factories* 12 (1): 61. <https://doi.org/10.1186/1475-2859-12-61>.
- Franco, P. J., and R. J. Brooker. 1994. 'Functional Roles of Glu-269 and Glu-325 within the Lactose Permease of *Escherichia Coli*.' *Journal of Biological Chemistry* 269 (10): 7379–86. [https://doi.org/10.1016/S0021-9258\(17\)37295-2](https://doi.org/10.1016/S0021-9258(17)37295-2).
- Franco, Peter J, Anupam B Jena, and T. Hastings Wilson. 2001. 'Physiological Evidence for an Interaction between Helices II and XI in the Melibiose Carrier of *Escherichia Coli*'. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1510 (1): 231–42. [https://doi.org/10.1016/S0005-2736\(00\)00353-9](https://doi.org/10.1016/S0005-2736(00)00353-9).
- Frillingos, Stathis, Miklós Sahin-Tóth, Jianhua Wu, and H. Ronald Kaback. 1998. 'Cys-Scanning Mutagenesis: A Novel Approach to Structure—Function Relationships in Polytopic Membrane Proteins'. *The FASEB Journal* 12 (13): 1281–99. <https://doi.org/10.1096/fasebj.12.13.1281>.
- Frillingos, Stathis, Jianzhong Sun, Alberto Gonzalez, and H. Ronald Kaback. 1997. 'Cysteine-Scanning Mutagenesis of Helix II and Flanking Hydrophilic Domains in the Lactose Permease of *Escherichia Coli*'. *Biochemistry* 36 (1): 269–73. <https://doi.org/10.1021/bi9618629>.

- Fromant, M., S. Blanquet, and P. Plateau. 1995. 'Direct Random Mutagenesis of Gene-Sized DNA Fragments Using Polymerase Chain Reaction'. *Analytical Biochemistry* 224 (1): 347–53. <https://doi.org/10.1006/abio.1995.1050>.
- Gilardi, Gianfranco., Li Qing. Zhou, Linda. Hibbert, and Anthony E. G. Cass. 1994. 'Engineering the Maltose Binding Protein for Reagentless Fluorescence Sensing'. *Analytical Chemistry* 66 (21): 3840–47. <https://doi.org/10.1021/ac00093a047>.
- Goswitz, Visala Chepuri, and Robert J. Brooker. 1993. 'Isolation of Lactose Permease Mutants Which Recognize Arabinose'. *Membrane Biochemistry* 10 (1): 61–70. <https://doi.org/10.3109/09687689309150253>.
- Gray, Declan A., Joshua B. R. White, Abraham O. Oluwole, Parthasarathi Rath, Amy J. Glenwright, Adam Mazur, Michael Zahn, et al. 2021. 'Insights into SusCD-Mediated Glycan Import by a Prominent Gut Symbiont'. *Nature Communications* 12 (1): 44. <https://doi.org/10.1038/s41467-020-20285-y>.
- Green, Aileen L., Ethan J. Anderson, and Robert J. Brooker. 2000. 'A Revised Model for the Structure and Function of the Lactose Permease: EVIDENCE THAT A FACE ON TRANSMEMBRANE SEGMENT 2 IS IMPORTANT FOR CONFORMATIONAL CHANGES *'. *Journal of Biological Chemistry* 275 (30): 23240–46. <https://doi.org/10.1074/jbc.M909202199>.
- Green, Aileen L., and Robert J. Brooker. 2001. 'A Face on Transmembrane Segment 8 of the Lactose Permease Is Important for Transport Activity'. *Biochemistry* 40 (40): 12220–29. <https://doi.org/10.1021/bi0109055>.
- Green, Aileen L., Heather A. Hrodey, and Robert J. Brooker. 2003. 'Evidence for Structural Symmetry and Functional Asymmetry in the Lactose Permease of Escherichia Coli'. *Biochemistry* 42 (38): 11226–33. <https://doi.org/10.1021/bi034810+>.
- Guan, Lan, and Parameswaran Hariharan. 2021. 'X-Ray Crystallography Reveals Molecular Recognition Mechanism for Sugar Binding in a Melibiose Transporter MelB'. *Communications Biology* 4 (1): 931. <https://doi.org/10.1038/s42003-021-02462-x>.
- Guan, Lan, S. Vivek Jakkula, Alexey A. Hodkoff, and Yue Su. 2012. 'Role of Gly117 in the Cation/Melibiose Symport of MelB of Salmonella Typhimurium'. *Biochemistry* 51 (13): 2950–57. <https://doi.org/10.1021/bi300230h>.

- Guan, Lan, and H. Ronald Kaback. 2009. 'Properties of a LacY Efflux Mutant'. *Biochemistry* 48 (39): 9250–55. <https://doi.org/10.1021/bi9014067>.
- Guan, Lan, Miklós Sahin-Tóth, and H. Ronald Kaback. 2002. 'Changing the Lactose Permease of Escherichia Coli into a Galactose-Specific Symporter'. *Proceedings of the National Academy of Sciences of the United States of America* 99 (10): 6613–18. <https://doi.org/10.1073/pnas.102178299>.
- Ha, Suk-Jin, Heejin Kim, Yuping Lin, Myoung-Uoon Jang, Jonathan M. Galazka, Tae-Jip Kim, Jamie H. D. Cate, and Yong-Su Jin. 2013. 'Single Amino Acid Substitutions in HXT2.4 from Scheffersomyces Stipitis Lead to Improved Cellobiose Fermentation by Engineered Saccharomyces Cerevisiae'. *Applied and Environmental Microbiology* 79 (5): 1500–1507. <https://doi.org/10.1128/AEM.03253-12>.
- Henderson, Ryan K., Sophie C. De Valk, Bert Poolman, and Robert Mans. 2021. 'Energy Coupling of Membrane Transport and Efficiency of Sucrose Dissimilation in Yeast'. *Metabolic Engineering* 65 (May):243–54. <https://doi.org/10.1016/j.ymben.2020.11.014>.
- Henderson, Ryan, and Bert Poolman. 2017. 'Proton-Solute Coupling Mechanism of the Maltose Transporter from Saccharomyces Cerevisiae'. *Scientific Reports* 7 (1): 14375. <https://doi.org/10.1038/s41598-017-14438-1>.
- Hor, Lien-I, and Howard A. Shuman. 1993. 'Genetic Analysis of Periplasmic Binding Protein Dependent Transport in Escherichia Coli: Each Lobe of Maltose-Binding Protein Interacts with a Different Subunit of the MalFGK2 Membrane Transport Complex'. *Journal of Molecular Biology* 233 (4): 659–70. <https://doi.org/10.1006/jmbi.1993.1543>.
- Jahreis, Knut, Lars Bentler, Jürgen Bockmann, Stephan Hans, Astrid Meyer, Jörg Siepelmeier, and Joseph W. Lengeler. 2002. 'Adaptation of Sucrose Metabolism in the Escherichia Coli Wild-Type Strain EC3132†'. *Journal of Bacteriology* 184 (19): 5307–16. <https://doi.org/10.1128/JB.184.19.5307-5316.2002>.
- Jakkula, S. Vivek, and Lan Guan. 2012. 'Reduced Na⁺ Affinity Increases Turnover of Salmonella Enterica Serovar Typhimurium MelB'. *Journal of Bacteriology* 194 (20): 5538–44. <https://doi.org/10.1128/JB.01206-12>.
- Jessen-Marshall, A E, N J Parker, and R J Brooker. 1997. 'Suppressor Analysis of Mutations in the Loop 2-3 Motif of Lactose Permease: Evidence That Glycine-64 Is an Important

- Residue for Conformational Changes'. *Journal of Bacteriology* 179 (8): 2616–22. <https://doi.org/10.1128/jb.179.8.2616-2622.1997>.
- Jessen-Marshall, Amy E., and Robert J. Brooker. 1996. 'Evidence That Transmembrane Segment 2 of the Lactose Permease Is Part of a Conformationally Sensitive Interface between the Two Halves of the Protein (*)'. *Journal of Biological Chemistry* 271 (3): 1400–1404. <https://doi.org/10.1074/jbc.271.3.1400>.
- Jessen-Marshall, Amy E., Nanette J. Paul, and Robert J. Brooker. 1995. 'The Conserved Motif, GXXX(D/E)(R/K)XG[X](R/K)(R/K), in Hydrophilic Loop 2/3 of the Lactose Permease *'. *Journal of Biological Chemistry* 270 (27): 16251–57. <https://doi.org/10.1074/jbc.270.27.16251>.
- Jiang, Xiaoxu, Magnus Andersson, Bryan T. Chau, Larissa Y. Wong, Maria Katerina R. Villafuerte, and H. Ronald Kaback. 2016. 'Role of Conserved Gly-Gly Pairs on the Periplasmic Side of LacY'. *Biochemistry* 55 (31): 4326–32. <https://doi.org/10.1021/acs.biochem.6b00666>.
- Jiang, Xiaoxu, Lan Guan, Yonggang Zhou, Wen-Xu Hong, Qinghai Zhang, and H. Ronald Kaback. 2012. 'Evidence for an Intermediate Conformational State of LacY'. *Proceedings of the National Academy of Sciences of the United States of America* 109 (12): E698–704. <https://doi.org/10.1073/pnas.1201107109>.
- Katsube, Satoshi, Ruibin Liang, Anowarul Amin, Parameswaran Hariharan, and Lan Guan. 2022. 'Molecular Basis for the Cation Selectivity of Salmonella Typhimurium Melibiose Permease'. *Journal of Molecular Biology* 434 (12): 167598. <https://doi.org/10.1016/j.jmb.2022.167598>.
- Keyhani, N. O., K. Bacia, and S. Roseman. 2000. 'The Transport/Phosphorylation of N,N'-Diacetylchitobiose in Escherichia Coli. Characterization of Phospho-IIB(Chb) and of a Potential Transition State Analogue in the Phosphotransfer Reaction between the Proteins IIA(Chb) AND IIB(Chb)'. *The Journal of Biological Chemistry* 275 (42): 33102–9. <https://doi.org/10.1074/jbc.M001045200>.
- Khare, Dheeraj, Michael L. Oldham, Cedric Orelle, Amy L. Davidson, and Jue Chen. 2009. 'Alternating Access in Maltose Transporter Mediated by Rigid-Body Rotations'. *Molecular Cell* 33 (4): 528–36. <https://doi.org/10.1016/j.molcel.2009.01.035>.

- Kim, Heejin, Eun Joong Oh, Stephan Thomas Lane, Won-Heong Lee, Jamie H.D. Cate, and Yong-Su Jin. 2018. 'Enhanced Cellobiose Fermentation by Engineered *Saccharomyces Cerevisiae* Expressing a Mutant Cellodextrin Facilitator and Cellobiose Phosphorylase'. *Journal of Biotechnology* 275 (June):53–59. <https://doi.org/10.1016/j.jbiotec.2018.04.008>.
- Kühnau, S, M Reyes, A Sievertsen, H A Shuman, and W Boos. 1991. 'The Activities of the *Escherichia Coli* MalK Protein in Maltose Transport, Regulation, and Inducer Exclusion Can Be Separated by Mutations.' *Journal of Bacteriology* 173 (7): 2180–86. <https://doi.org/10.1128/jb.173.7.2180-2186.1991>.
- León, Xavier, Gérard Leblanc, and Esteve Padrós. 2009. 'Alteration of Sugar-Induced Conformational Changes of the Melibiose Permease by Mutating Arg141 in Loop 4-5'. *Biophysical Journal* 96 (12): 4877–86. <https://doi.org/10.1016/j.bpj.2009.03.025>.
- Lian, Jiazhang, Yanglin Li, Mohammad Hamedirad, and Huimin Zhao. 2014. 'Directed Evolution of a Cellodextrin Transporter for Improved Biofuel Production under Anaerobic Conditions in *Saccharomyces Cerevisiae*: CDT Engineering Via Directed Evolution'. *Biotechnology and Bioengineering* 111 (8): 1521–31. <https://doi.org/10.1002/bit.25214>.
- Lin, Yibin, Oliver Fuerst, and Meritxell Granell. 2013. 'The Substitution of Arg149 with Cys Fixes the Melibiose Transporter in an Inward-Open Conformation'. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1828 (8): 1690–99. <https://doi.org/10.1016/j.bbamem.2013.03.003>.
- Lippincott, J, and B Traxler. 1997. 'MalFGK Complex Assembly and Transport and Regulatory Characteristics of MalK Insertion Mutants.' *Journal of Bacteriology* 179 (4): 1337–43. <https://doi.org/10.1128/jb.179.4.1337-1343.1997>.
- Liu, Zhenyu, M. Gregor Madej, and H. Ronald Kaback. 2010. 'Helix Dynamics in LacY: Helices II/IV'. *Journal of Molecular Biology* 396 (3): 617–26. <https://doi.org/10.1016/j.jmb.2009.12.044>.
- Lv, Xiaodong, Haijie Xue, Lei Qin, and Chun Li. 2022. 'Transporter Engineering in Microbial Cell Factory Boosts Biomanufacturing Capacity'. *BioDesign Research* 2022 (January):2022/9871087. <https://doi.org/10.34133/2022/9871087>.

- Markham, Kelsey J., Elena B. Tikhonova, Aaron C. Scarpa, Parameswaran Hariharan, Satoshi Katsube, and Lan Guan. 2021. 'Complete Cysteine-Scanning Mutagenesis of the Salmonella Typhimurium Melibiose Permease'. *Journal of Biological Chemistry* 297 (3): 101090. <https://doi.org/10.1016/j.jbc.2021.101090>.
- Martineau, Pierre, Sevec Szmelcman, John C. Spurlino, Florente A. Quiocho, and Maurice Hofnung. 1990. 'Genetic Approach to the Role of Tryptophan Residues in the Activities and Fluorescence of a Bacterial Periplasmic Maltose-Binding Protein'. *Journal of Molecular Biology* 214 (1): 337–50. [https://doi.org/10.1016/0022-2836\(90\)90165-I](https://doi.org/10.1016/0022-2836(90)90165-I).
- Matsuzaki, Shigenobu, Audrey C. Weissborn, Eiji Tamai, Tomofusa Tsuchiya, and T. Hastings Wilson. 1999. 'Melibiose Carrier of Escherichia Coli: Use of Cysteine Mutagenesis to Identify the Amino Acids on the Hydrophilic Face of Transmembrane Helix 2'. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1420 (1): 63–72. [https://doi.org/10.1016/S0005-2736\(99\)00087-5](https://doi.org/10.1016/S0005-2736(99)00087-5).
- Merino, G, and H A Shuman. 1997. 'Unliganded Maltose-Binding Protein Triggers Lactose Transport in an Escherichia Coli Mutant with an Alteration in the Maltose Transport System.' *Journal of Bacteriology* 179 (24): 7687–94. <https://doi.org/10.1128/jb.179.24.7687-7694.1997>.
- Merino, Gonzalo, and Howard A. Shuman. 1998. 'Truncation of MalF Results in Lactose Transport via the Maltose Transport System of Escherichia Coli *'. *Journal of Biological Chemistry* 273 (4): 2435–44. <https://doi.org/10.1074/jbc.273.4.2435>.
- Mystkowska, Agata Anna, Craig Robb, Silvia Vidal-Melgosa, Chiara Vanni, Antonio Fernandez-Guerra, Matthias Höhne, and Jan-Hendrik Hehemann. 2018. 'Molecular Recognition of the Beta-glucans Laminarin and Pustulan by a SusD-like Glycan-binding Protein of a Marine *Bacteroidetes*'. *The FEBS Journal* 285 (23): 4465–81. <https://doi.org/10.1111/febs.14674>.
- Nelson, Bryn D., and Beth Traxler. 1998. 'Exploring the Role of Integral Membrane Proteins in ATP-Binding Cassette Transporters: Analysis of a Collection of MalG Insertion Mutants'. *Journal of Bacteriology* 180 (9): 2507–14. <https://doi.org/10.1128/jb.180.9.2507-2514.1998>.

- Nie, Yiling, Frances E. Sabetfard, and H. Ronald Kaback. 2008. 'The Cys154→Gly Mutation in LacY Causes Constitutive Opening of the Hydrophilic Periplasmic Pathway'. *Journal of Molecular Biology* 379 (4): 695–703. <https://doi.org/10.1016/j.jmb.2008.04.015>.
- Oh, Eun Joong, Suryang Kwak, Heejin Kim, and Yong-Su Jin. 2017. 'Transporter Engineering for Cellobiose Fermentation under Lower pH Conditions by Engineered *Saccharomyces Cerevisiae*'. *Bioresource Technology* 245 (December):1469–75. <https://doi.org/10.1016/j.biortech.2017.05.138>.
- Oh, Eun Joong, Jeffrey M. Skerker, Soo Rin Kim, Na Wei, Timothy L. Turner, Matthew J. Maurer, Adam P. Arkin, and Yong-Su Jin. 2016. 'Gene Amplification on Demand Accelerates Cellobiose Utilization in Engineered *Saccharomyces Cerevisiae*'. Edited by D. Cullen. *Applied and Environmental Microbiology* 82 (12): 3631–39. <https://doi.org/10.1128/AEM.00410-16>.
- Okumura, Kenji, Yukie Maruyama, Ryuichi Takase, Bunzo Mikami, Kousaku Murata, and Wataru Hashimoto. 2021. 'The Role of Calcium Binding to the EF-Hand-like Motif in Bacterial Solute-Binding Protein for Alginate Import'. *Bioscience, Biotechnology, and Biochemistry* 85 (12): 2410–19. <https://doi.org/10.1093/bbb/zbab170>.
- Pazdernik, Nanette J., Shane M. Cain, and Robert J. Brooker. 1997. 'An Analysis of Suppressor Mutations Suggests That the Two Halves of the Lactose Permease Function in a Symmetrical Manner*'. *Journal of Biological Chemistry* 272 (42): 26110–16. <https://doi.org/10.1074/jbc.272.42.26110>.
- Poolman, B, R Modderman, and J Reizer. 1992. 'Lactose Transport System of *Streptococcus Thermophilus*. The Role of Histidine Residues.' *Journal of Biological Chemistry* 267 (13): 9150–57. [https://doi.org/10.1016/S0021-9258\(19\)50402-1](https://doi.org/10.1016/S0021-9258(19)50402-1).
- Potocki-Veronese, Gabrielle, Sophie LAJUS, Adilya DAGKESAMANSKAYA, Amandine DEROITE, and Delphine LESTRADE. 2023. New milli-to-micro-fluidics-based process to screen microbial growth in droplets. World Intellectual Property Organization WO2023222765A1, filed 17 May 2023, and issued 23 November 2023. [https://patents.google.com/patent/WO2023222765A1/fr?q=\(lajus+potocki\)&oq=lajus+potocki](https://patents.google.com/patent/WO2023222765A1/fr?q=(lajus+potocki)&oq=lajus+potocki).

- Potocki-Veronese, Gabrielle, Sophie LAJUS, Adilya DAGKESAMANSKAYA, Amandine DEROITE, Delphine LESTRADE, and Rémy FLORES-FLORES. 2023. New droplet micro-to-milli-fluidics-based process to screen for phenotypes or biological processes. World Intellectual Property Organization WO2023222784A1, filed 17 May 2023, and issued 23 November 2023. [https://patents.google.com/patent/WO2023222784A1/fr?q=\(Sophie+Lajus\)&oq=Sophie+Lajus](https://patents.google.com/patent/WO2023222784A1/fr?q=(Sophie+Lajus)&oq=Sophie+Lajus).
- Rodseth, Lynn E., Pierre Martineau, Pascale Duplay, Maurice Hofnung, and Florante A. Quioco. 1990. 'Crystallization of Genetically Engineered Active Maltose-Binding Proteins, Including an Immunogenic Viral Epitope Insertion'. *Journal of Molecular Biology* 213 (4): 607–11. [https://doi.org/10.1016/S0022-2836\(05\)80246-3](https://doi.org/10.1016/S0022-2836(05)80246-3).
- Saul, Frederick A., Michaël Mourez, Brigitte Vulliez-le Normand, Nathalie Sassoon, Graham A. Bentley, and Jean-Michel Betton. 2003. 'Crystal Structure of a Defective Folding Protein'. *Protein Science: A Publication of the Protein Society* 12 (3): 577–85. <https://doi.org/10.1110/ps.0235103>.
- Shilton, Brian H., Howard A. Shuman, and Sherry L. Mowbray. 1996. 'Crystal Structures and Solution Conformations of a Dominant-Negative Mutant of Escherichia coli Maltose-Binding Protein'. *Journal of Molecular Biology* 264 (2): 364–76. <https://doi.org/10.1006/jmbi.1996.0646>.
- Shinnick, S.G., and M.F. Varela. 2002. 'Altered Sugar Selection and Transport Conferred by Spontaneous Point and Deletion Mutations in the Lactose Carrier of Escherichia Coli'. *The Journal of Membrane Biology* 189 (3): 191–99. <https://doi.org/10.1007/s00232-002-1013-9>.
- Shinnick, Steven G., Stephanie A. Perez, and Manuel F. Varela. 2003. 'Altered Substrate Selection of the Melibiose Transporter (MelY) of Enterobacter Cloacae Involving Point Mutations in Leu-88, Leu-91, and Ala-182 That Confer Enhanced Maltose Transport'. *Journal of Bacteriology* 185 (12): 3672–77. <https://doi.org/10.1128/JB.185.12.3672-3677.2003>.
- Shuman, H A, and J Beckwith. 1979. 'Escherichia Coli K-12 Mutants That Allow Transport of Maltose via the Beta-Galactoside Transport System.' *Journal of Bacteriology* 137 (1): 365–73. <https://doi.org/10.1128/jb.137.1.365-373.1979>.

- Smit, A., S.G. Moses, I.S. Pretorius, and R.R. Cordero Otero. 2008. 'The Thr505 and Ser557 Residues of the AGT1-encoded A-glucoside Transporter Are Critical for Maltotriose Transport in *Saccharomyces Cerevisiae*'. *Journal of Applied Microbiology* 104 (4): 1103–11. <https://doi.org/10.1111/j.1365-2672.2007.03671.x>.
- Stemmer, Willem P. C. 1994. 'Rapid Evolution of a Protein in Vitro by DNA Shuffling'. *Nature* 370 (6488): 389–91. <https://doi.org/10.1038/370389a0>.
- Szmelcman, S., N. Sassoon, and M. Hofnung. 1997. 'Residues in the Alpha Helix 7 of the Bacterial Maltose Binding Protein Which Are Important in Interactions with the Mal FGK2 Complex.' *Protein Science : A Publication of the Protein Society* 6 (3): 628–36. <https://doi.org/10.1002/pro.5560060312>.
- Tamura, Kazune, Matthew H. Foley, Bernd R. Gardill, Guillaume Dejean, Matthew Schnizlein, Constance M. E. Bahr, A. Louise Creagh, Filip van Petegem, Nicole M. Koropatkin, and Harry Brumer. 2019. 'Surface Glycan-Binding Proteins Are Essential for Cereal Beta-Glucan Utilization by the Human Gut Symbiont *Bacteroides Ovatus*'. *Cellular and Molecular Life Sciences* 76 (21): 4319–40. <https://doi.org/10.1007/s00018-019-03115-3>.
- Tauzin, Alexandra S., Kurt J. Kwiatkowski, Nicole I. Orlovsky, Christopher J. Smith, A. Louise Creagh, Charles A. Haynes, Zdzislaw Wawrzak, Harry Brumer, and Nicole M. Koropatkin. 2016. 'Molecular Dissection of Xyloglucan Recognition in a Prominent Human Gut Symbiont'. Edited by Justin Sonnenburg and Margaret J. McFall-Ngai. *mBio* 7 (2): e02134-15. <https://doi.org/10.1128/mBio.02134-15>.
- Tauzin, Alexandra S., Elisabeth Laville, Yao Xiao, Sébastien Nouaille, Pascal Le Bourgeois, Stéphanie Heux, Jean-Charles Portais, et al. 2016. 'Functional Characterization of a Gene Locus from an Uncultured Gut *Bacteroides* Conferring Xylo-oligosaccharides Utilization to *Escherichia Coli*'. *Molecular Microbiology* 102 (4): 579–92. <https://doi.org/10.1111/mmi.13480>.
- Treptow, N A, and H A Shuman. 1985. 'Genetic Evidence for Substrate and Periplasmic-Binding-Protein Recognition by the MalF and MalG Proteins, Cytoplasmic Membrane Components of the *Escherichia Coli* Maltose Transport System.' *Journal of Bacteriology* 163 (2): 654–60. <https://doi.org/10.1128/jb.163.2.654-660.1985>.

- Treptow, Nancy A., and Howard A. Shuman. 1988. 'Allele-Specific malE Mutations That Restore Interactions between Maltose-Binding Protein and the Inner-Membrane Components of the Maltose Transport System'. *Journal of Molecular Biology* 202 (4): 809–22. [https://doi.org/10.1016/0022-2836\(88\)90560-8](https://doi.org/10.1016/0022-2836(88)90560-8).
- Ujwal, M. L., Miklós Sahin-Tóth, Bengt Persson, and H. Ronald Kaback. 1994. 'Role of Glutamate-269 in the Lactose Permease of Escherichia Coli'. *Molecular Membrane Biology* 11 (1): 9–16. <https://doi.org/10.3109/09687689409161024>.
- Van Camp, Bonnie M., Robert R. Crow, Yang Peng, and Manuel F. Varela. 2007. 'Amino Acids That Confer Transport of Raffinose and Maltose Sugars in the Raffinose Permease (RafB) of Escherichia Coli as Implicated by Spontaneous Mutations at Val-35, Ser-138, Ser-139, Gly-389 and Ile-391'. *The Journal of Membrane Biology* 220 (1–3): 87–95. <https://doi.org/10.1007/s00232-007-9077-1>.
- Van Montfort, Rob Lm, Tjaard Pijning, Kor H Kalk, Jonathan Reizer, Milton H Saier, Marjolein Mgm Thunnissen, George T Robillard, and Bauke W Dijkstra. 1997. 'The Structure of an Energy-Coupling Protein from Bacteria, IIBcellobiose, Reveals Similarity to Eukaryotic Protein Tyrosine Phosphatases'. *Structure* 5 (2): 217–25. [https://doi.org/10.1016/S0969-2126\(97\)00180-9](https://doi.org/10.1016/S0969-2126(97)00180-9).
- Varela, M.F., T.H. Wilson, V. Rodon-Rivera, S. Shepherd, T.A. Dehne, and A.C. Rector. 2000. 'Mutants of the Lactose Carrier of Escherichia Coli Which Show Altered Sugar Recognition Plus a Severe Defect in Sugar Accumulation'. *The Journal of Membrane Biology* 174 (3): 199–205. <https://doi.org/10.1007/s002320001044>.
- Vitrac, Heidi, Venkata K. P. S. Mallampalli, Stavros Azinas, and William Dowhan. 2020. 'Structural and Functional Adaptability of Sucrose and Lactose Permeases from Escherichia Coli to the Membrane Lipid Composition'. *Biochemistry* 59 (19): 1854–68. <https://doi.org/10.1021/acs.biochem.0c00174>.
- Wang, Zhi, Alexandra S. Tauzin, Elisabeth Laville, and Gabrielle Potocki-Veronese. 2022. 'Identification of Glycoside Transporters From the Human Gut Microbiome'. *Frontiers in Microbiology* 13. <https://doi.org/10.3389/fmicb.2022.816462>.
- Wang, Zhi, Alexandra S. Tauzin, Elisabeth Laville, Pietro Tedesco, Fabien Létisse, Nicolas Terrapon, Pascale Lepercq, Myriam Mercade, and Gabrielle Potocki-Veronese. 2020.

- ‘Harvesting of Prebiotic Fructooligosaccharides by Nonbeneficial Human Gut Bacteria’. Edited by Vincent B. Young. *mSphere* 5 (1): e00771-19. <https://doi.org/10.1128/mSphere.00771-19>.
- Weinglass, Adam B., Melissa Sondej, and H. Ronald Kaback. 2002. ‘Manipulating Conformational Equilibria in the Lactose Permease of *Escherichia coli*’. Edited by G. von Heijne. *Journal of Molecular Biology* 315 (4): 561–71. <https://doi.org/10.1006/jmbi.2001.5289>.
- Yang, Jian, Liping Zhang, Shanshan Lin, Wei Li, Chen Liu, Jingheng Yan, Shubo Li, and Lijuan Long. 2023. ‘Structural Insights of a SusD-like Protein in Marine Bacteroidetes Bacteria Reveal the Molecular Basis for Chitin Recognition and Acquisition’. *The FEBS Journal* n/a (n/a). <https://doi.org/10.1111/febs.16974>.
- Zhao, Huimin, Lori Giver, Zhixin Shao, Joseph A. Affholter, and Frances H. Arnold. 1998. ‘Molecular Evolution by Staggered Extension Process (StEP) in Vitro Recombination’. *Nature Biotechnology* 16 (3): 258–61. <https://doi.org/10.1038/nbt0398-258>.
- Zhou, Yonggang, Lan Guan, J. Alfredo Freites, and H. Ronald Kaback. 2008. ‘Opening and Closing of the Periplasmic Gate in Lactose Permease’. *Proceedings of the National Academy of Sciences of the United States of America* 105 (10): 3774–78. <https://doi.org/10.1073/pnas.0800825105>.
- Zhou, Yonggang, Xiaoxu Jiang, and H. Ronald Kaback. 2012. ‘Role of the Irreplaceable Residues in the LacY Alternating Access Mechanism’. *Proceedings of the National Academy of Sciences of the United States of America* 109 (31): 12438–42. <https://doi.org/10.1073/pnas.1210684109>.
- Zhou, Yonggang, M. Gregor Madej, Lan Guan, Yiling Nie, and H. Ronald Kaback. 2011. ‘An Early Event in the Transport Mechanism of LacY Protein’. *The Journal of Biological Chemistry* 286 (35): 30415–22. <https://doi.org/10.1074/jbc.M111.268433>.
- Zhou, Yonggang, Yiling Nie, and H. Ronald Kaback. 2009. ‘Residues Gating the Periplasmic Pathway of LacY’. *Journal of Molecular Biology* 394 (2): 219–25. <https://doi.org/10.1016/j.jmb.2009.09.043>.

Thesis objectives

As explained in the bibliographic introduction chapter, there are various oligosaccharide transporter families in bacteria. These transporters are important determinants of the metabolic ability of bacteria and they represent key elements in fermentation processes and synthetic biology. The study of oligosaccharide transporters attracted increasing attention from researchers in the past two decades, with more and more studies focusing on functional and structural characterization of these proteins, as well as their engineering. Nevertheless, the knowledge about oligosaccharide transporters is still limited, particularly in comparison to the CAZymes, which constitute the catalytic parts of the multi-proteic systems encoded by polysaccharide and oligosaccharide utilization operons. If CAZymes are relatively easy to characterize, the specificity of transporters is much more difficult to decipher. This is due to their transmembrane location, the multiplicity of transport systems in native strains, and the lack of genetic tools for many species, especially the non-cultured organisms which make up the major part of microbial ecosystems. At the beginning of my thesis work, only a few dozen transporters involved in the utilization of dietary fibers by mammal gut bacteria had been functionally or structurally characterized. Furthermore, transporters share low sequence similarities and there are not enough functional data that could be used to predict, from multiple sequence alignments and phylogenetic trees, their transport specificity towards oligosaccharides, while it is essential to understand how bacteria colonise their habitat. Finally, the lack of structural data for these proteins, formed either by a single transmembrane protein or by multiple components, limits the understanding of their structure-function relationships in general.

In this context, in the past few years, the CIMEs team has demonstrated that molecular characterization of transporters derived from uncultured bacteria (including their transmembrane components) can be performed in *E. coli*. Furthermore, the team has developed several new technologies to screen and characterize their specificity. My thesis work aimed at characterizing and engineering the specificity of oligosaccharide transporters, in particular from the SusC/D superfamily, for which the lack of functional and structural data is even more prominent than for the other superfamilies. The transporters targeted for this project have previously been identified by the team through functional metagenomics of the human and bovine gut microbiomes. My thesis work was based on the expertise of the team in structural biology and protein engineering, as well as in ultra-high throughput functional screening. It

targets various long-term applications for synthetic biology and for the control of microbial ecosystems functioning. In this context, the objectives of my PhD project were the following:

- to characterize the function and structure of key proteic elements (SusD-like protein) involved in the capture and utilization of polysaccharides and oligosaccharides by uncultured human and bovine gut Bacteroidota species
- to investigate the specificity of transporters towards oligosaccharides of various structures
- to develop an evolution approach for transporters, that could be attempted to improve their transport efficiency or to change their specificity to target oligosaccharides of complex structures.

Results

First chapter

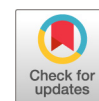
Humans rely on their microbiota to break down dietary fibers, which are composed of oligosaccharides and polysaccharides that cannot be digested by host CAZymes. In the human gut microbiota, Bacteroidota break down dietary and host-derived oligosaccharides through highly specific polysaccharide utilization loci (PULs). PULs encode a variety of proteins involved in the sensing, binding, transport and degradation of the target substrates into monosaccharides. According to the starch utilization system (Sus) paradigm, surface glycan-binding proteins (SusD-like SGBPs) are essential for the efficient capture at the cell surface of the polysaccharides and oligosaccharides, providing Bacteroidota with a competitive advantage in colonizing their habitats.

An *E. coli* metagenomic clone (F5) containing a polysaccharide utilization locus (PUL) specific for xylooligosaccharides (XOS) was previously identified and characterized in our team. The metagenomic DNA fragment was issued from an uncultured gut *Bacteroides* strain (Cecchini et al. 2013; Tauzin et al. 2016). The F5_PUL encodes a SusC/D system and a MFS transporter, which are dedicated to XOS uptake up to a degree of polymerization 3 and 4, respectively. The deletion of the *susD* gene from a truncated variant of clone F5 (F5minSus) resulted in the inability of the recombinant *E. coli* cells to grow on XOS, which indicates that the presence of SusD is essential for the functionality of the SusC transporter and for the uptake of XOS. In other characterized PUL-encoded systems, the physical presence of the SusD protein was also found to be essential for strain growth on the polysaccharides targeted by the PUL. Furthermore, the F5_SusD-like protein shares very low sequence similarities with those from characterized PULs, and its binding specificity has never been studied. In this context, we aimed at answering the following questions:

- How does F5_SusD-like contribute to the utilization of XOS?
- What are the relationships between the tridimensional structure of the F5_SusD and its function?

To do so, we successfully expressed the recombinant F5_SusD-like protein in *E. coli* and subsequently characterized its functional and structural properties. Various *in vitro* analyses, including saturation transfer difference nuclear magnetic resonance (STD-NMR) spectroscopy, isothermal titration calorimetry (ITC), differential scanning fluorimetry (DSF), affinity gel electrophoresis (AGE), were used to investigate the binding ability of F5_SusD-like protein.

Analysis of the crystal structure of the F5_SusD-like protein was also used to elucidate the structure-function relations. In this collaborative work, I performed size exclusion chromatography, SDS-PAGE, AGE and differential scanning fluorimetry analyses. I also performed complementation experiments of the F5minSus Δ SusD strain with purified SusD, but the results were not conclusive and we decided to not include them in this paper, published in *mSphere*.



Structural and Biochemical Characterization of a Nonbinding SusD-Like Protein Involved in Xylooligosaccharide Utilization by an Uncultured Human Gut *Bacteroides* Strain

Alexandra S. Tauzin,^a Zhi Wang,^{a*} Gianluca Cioci,^a Xiaoqian Li,^a Aurore Labourel,^a Barbara Machado,^b Guy Lippens,^a  Gabrielle Potocki-Veronese^a

^aTBI, CNRS, INRAE, INSAT, Université de Toulouse, Toulouse, France

^bALBA Synchrotron Light Source, Barcelona, Spain

Alexandra S. Tauzin and Zhi Wang contributed equally to this work. Author order was determined both alphabetically and in order of increasing seniority.

ABSTRACT In the human gut microbiota, *Bacteroidetes* break down dietary and endogenous glycosides through highly specific polysaccharide utilization loci (PULs). PULs encode a variety of sensor regulators, binding proteins, transporters, and carbohydrate-active enzymes (CAZymes). Surface glycan-binding proteins (SGBPs) are essential for the efficient capture of the glycosides present on the cell surface, providing *Bacteroidetes* with a competitive advantage in colonizing their habitats. Here, we present the functional and structural characterization of a SusD-like protein encoded by a xylooligosaccharide (XOS) PUL from an uncultured human gut *Bacteroides* strain. This locus is also conserved in *Bacteroides vulgatus*, thereby providing new mechanistic insights into the role of SGBPs in the metabolism of dietary fiber of importance for gut health. Various *in vitro* analyses, including saturation transfer difference nuclear magnetic resonance (STD-NMR) spectroscopy, revealed that the SusD-like protein cannot bind to the cognate substrate of the XOS PUL, although its presence is essential for the PUL to function. Analysis of the crystal structure of the SusD-like protein reveals an unfolded binding surface and the absence or inappropriate orientation of several key residues compared with other known SusD-like structures. These results highlight the critical role of the SusD-like protein in the transport of oligosaccharides and provide fundamental knowledge about the structure-function of SusC/D-like transporters, revealing that the binding specificity of SusD-like SGBPs does not necessarily reflect the uptake specificity of the transporter.

IMPORTANCE The metabolization of dietary fiber is a crucial function for many gut bacteria, especially *Bacteroidetes*, which are particularly well adapted for recognizing, binding, transporting, and degrading glycosides. In this study, we report the functional and structural characterization of a SusD-like protein involved in xylooligosaccharide utilization by an uncultured gut *Bacteroides* strain. We demonstrate that while this protein is structurally similar to many canonical *Bacteroidetes* surface glycan-binding proteins, it cannot bind the substrate taken up by the cognate SusC-like transporter. This lack of binding might be explained by the absence of several key residues known to be involved in oligosaccharide binding and/or the possible necessity of the SusC-like protein to be present to create a cooperative binding site. The term “surface glycan-binding proteins” generally used for SusD-like proteins is thus not generic. Overall, this study allowed us to revisit the concept of glycoside utilization by *Bacteroidetes*, in particular those strains that feed on the short fibers naturally present in some dietary compounds or on the leftovers of other microbes.

KEYWORDS xylooligosaccharides, SusC/D, *Bacteroidetes*, uncultured gut bacterium

Editor Vincent B. Young, University of Michigan—Ann Arbor

Copyright © 2022 Tauzin et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Gabrielle Potocki-Veronese, veronese@insa-toulouse.fr.

*Present address: Zhi Wang, School of Chemical Engineering, Zhengzhou University, Zhengzhou, Henan, China.

The authors declare no conflict of interest.

Received 24 May 2022

Accepted 13 July 2022

Published 31 August 2022

Humans rely on their microbiota to break down dietary fiber, which is composed of oligosaccharides and polysaccharides that cannot be digested by host carbohydrate-active enzymes (CAZymes) (1). To achieve this, *Bacteroidetes* harbor polysaccharide utilization loci (PULs) that encode proteins involved in sensing, binding, transporting, and degrading the target glycosides into monosaccharides (2, 3). According to the starch utilization system (Sus) paradigm (4–6), polysaccharides are first degraded by enzymes attached to the cell surface. Next, SusD-like substrate-binding lipoproteins, sometimes together with the other surface glycan-binding proteins (SGBPs), SusE and/or -F, selectively recognize and capture the oligosaccharides that will be transported into the cell through the SusC-like protein.

An in-depth characterization of the first starch utilization locus of *Bacteroides thetaiotaomicron* to be described showed that the SusD protein has a critical role in addition to its binding ability (6). The deletion of SusD actually resulted in a loss of growth, suggesting that the presence of SusD is a prerequisite for efficient transport (7–10). SusD and SusC interact to form a complex (11). Recent findings concerning the structures of SusC/D-like complexes provided new insights into the molecular determinants of the protein-protein and glycan-protein interactions involved (12, 13). The SusC/D-like pair is well conserved among *Bacteroidetes*, highlighting its crucial role in glycan uptake, and the presence of genes encoding SusD-like proteins is crucial for predicting PULs (14).

Dozens of *Bacteroidetes* strains have been shown to use xylan as their main source of carbon (15–25). Xylan is a major constituent of hemicelluloses, an abundant component in plant cell walls, particularly in cereals, fruits, and vegetables (26). The conserved backbone of xylan is composed of a linear β -(1-4)-D-xylopyranosyl chain, which may be decorated with arabinofuranosyl, glucopyranosyl, or uronic acid derivatives (26). Xylooligosaccharides (XOSs) are hydrolyzed from xylan or are found naturally in fruits, vegetable bamboo, and honey (27). XOSs have a range of health benefits, even at a lower dose than fructooligosaccharides, and are thus considered prebiotic candidates (27–30). To date, only two PULs specific for xylan utilization by *Bacteroides ovatus* ATCC 8483 (PUL-XylS and PUL-XylL) have been fully biochemically characterized. Functionally, they show different specificities depending on the structural complexity of the xylan (21). In addition, in 2016, we identified an XOS-specific PUL from an uncultured gut *Bacteroides* strain and characterized its protein components through recombinant expression in *Escherichia coli* (31, 32). This locus encodes a SusC/D-like pair and a major facilitator superfamily (MFS) transport system capable of XOS uptake up to a degree of polymerization of 3 (DP3) and DP4, respectively. We also demonstrated that the deletion of the SusD-like protein abolished the growth ability of the recombinant *E. coli* cells on XOS, consistent with the results obtained with other SusD-like proteins. Furthermore, this PUL shares 99% DNA sequence identity with a *Bacteroides vulgatus* ATCC 8482 PUL (BVU_0037 to BVU_0043). We previously proved that this PUL is involved in linear XOS utilization by the *Bacteroides vulgatus* ATCC 8482 strain (32). On XOS, its growth is very robust, while on arabinoxylooligosaccharides (AXOSs), it is delayed and takes place at a lower rate. In addition, this strain is unable to utilize complex heavily decorated xylans or arabinoxylan, although rarely, it may utilize simpler xylans from beechwood and birchwood (32).

In the present study, we investigated how the SusD-like protein from this XOS-targeting PUL contributes to the utilization of its cognate substrate. By using various *in vitro* techniques to investigate the binding ability of this SusD-like protein and analyzing its crystallographic structure, we provide new mechanistic insights into XOS uptake by *Bacteroidetes*.

RESULTS

F5_SusD-like cannot bind to XOS. (i) Determination of the binding ability by AGE, ITC, and DSF. The XOS locus encodes a SusD-like protein (referred to as F5_SusD-like here) sharing 99.4% identity with the SusD-like protein (BVU_0037) from *Bacteroides*

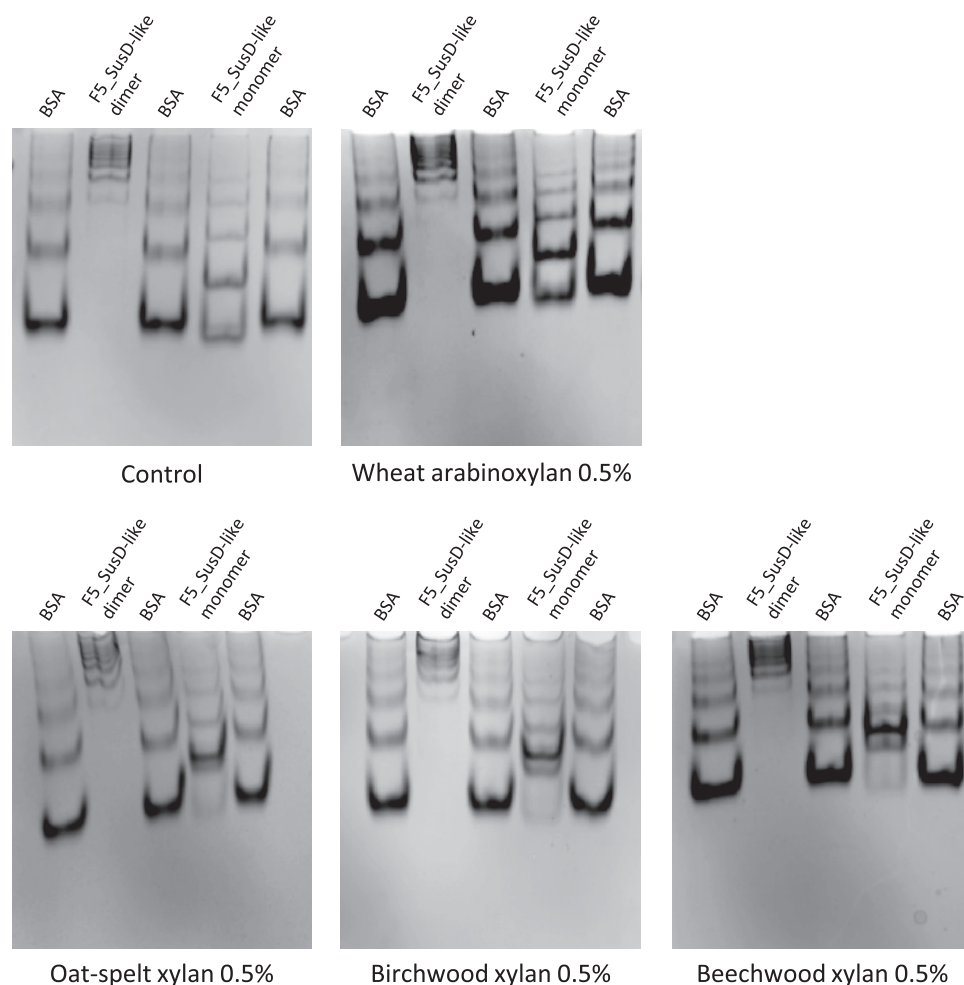


FIG 1 Affinity gel electrophoresis of F5_SusD-like on a gel containing no polysaccharide (negative control) and on gels with 0.5% (wt/vol) different xylans.

vulgatus ATCC 8482 (with four differences at positions N249, D444, A446, and A516 corresponding to D, G, P, and V residues in BVU_0037, respectively).

The recombinant protein was expressed in *E. coli* without the predicted signal peptide and N-terminal lipidation site. The ability of F5_SusD-like to bind several soluble glycans, including xylans and arabinoxylan, was assessed by affinity gel electrophoresis (AGE) (Fig. 1; see also Fig. S1 in the supplemental material). As shown by size exclusion chromatography (Fig. S2), F5_SusD-like forms monomers and dimers in solution. The dimer showed no difference in migration in the presence or absence of glycan. In contrast, monomer migration was slightly delayed by 27, 28, and 31% in the gels containing beechwood, birchwood, and oat spelt xylans, respectively. No migration delay was observed with wheat arabinoxylan. The observed smears suggest that the binding of the monomer to xylans could be unstable due to a weak affinity. F5_SusD-like is retained in the gel containing xylans to a much smaller degree than what was observed for all other characterized SusD-like proteins that efficiently bind polysaccharides (9, 10). These results indicate a weak affinity of F5_SusD-like for xylan, impaired by the presence of side chains on the xylan backbone. Isothermal titration calorimetry (ITC) was attempted using XOS as a ligand, but no binding could be detected (data not shown). Finally, differential scanning fluorimetry (DSF) was performed, but no significant change in the intrinsic fluorescence of F5_SusD-like was observed in the presence of birchwood, beechwood, or oat spelt xylans or barley β -glucan, carboxymethylcellulose, laminarin, wheat arabinoxylan, or xyloglucan (data not shown). These results indicate that the thermal stability of F5_SusD-

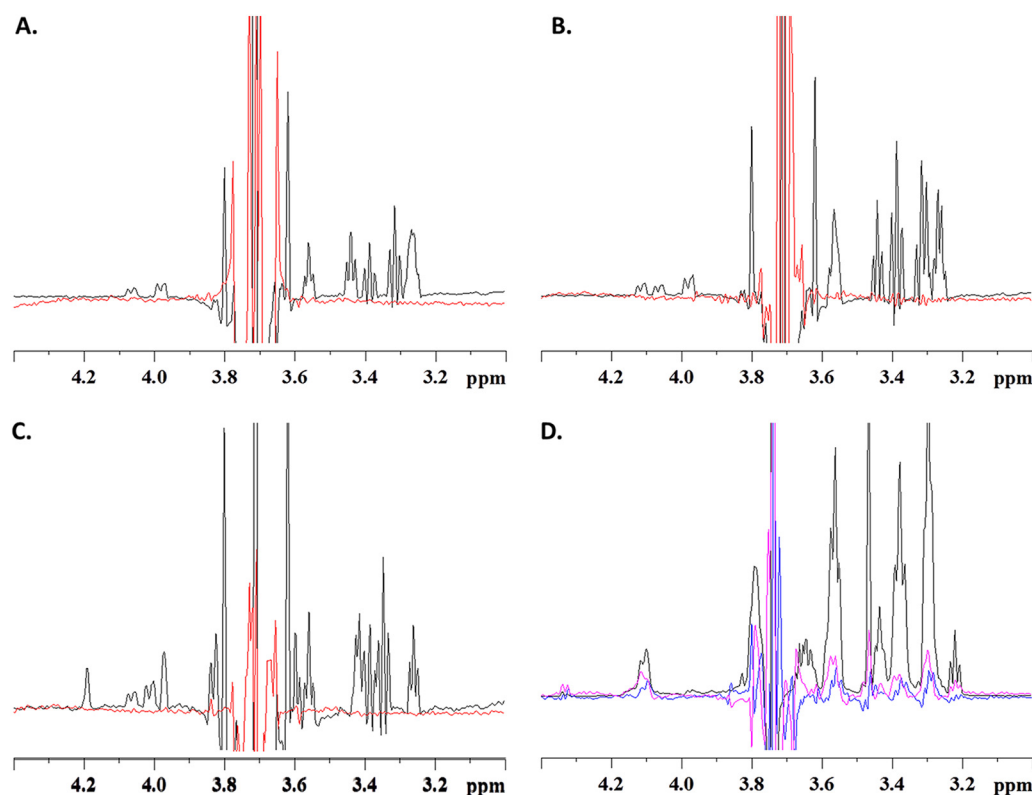


FIG 2 Characterization of the ligand-binding abilities of F5_SusD-like using STD-NMR spectroscopy. (A) Zoomed-in view of the 1D xylobiose proton spectrum centered on the signals at 3.7 ppm (black) and the STD signal of the ligand upon saturation of the protein signals at 0 ppm (red). (B) Same as panel A but for the xylotriose. (C) Same as panel A but for the branched arabinoxylooligosaccharides (AX2 [3²-α-L-arabinofuranosyl-xylobiose]). (D) Zoomed-in view of the 1D xylan spectrum (black), the STD signal upon saturation at 10.1 ppm of the isolated xylan (blue), or the xylan in the presence of F5_SusD-like (pink). The increased STD effect in the presence of F5_SusD-like indicates a molecular interaction.

like was not affected by significant ligand binding, contrary to what could be observed using DSF for genuine polysaccharide-binding proteins (33).

(ii) Determination of the binding ability by NMR spectroscopy. The value of saturation transfer difference nuclear magnetic resonance (STD-NMR) spectroscopy for the direct analysis of protein-ligand interactions with even weak affinities has been shown previously (34). However, to our knowledge, this approach has been used only rarely to analyze the binding ability of carbohydrate-binding proteins associated with CAZymes or with carbohydrate transporters (35, 36). Here, we used this method in order to further characterize the (in)ability of F5_SusD-like to bind to XOS. We first confirmed by one-dimensional (1D) NMR that the F5_SusD-like protein was correctly folded in solution (Fig. S3). The results of the STD experiments with xylobiose, xylotriose, or 3²-α-L-arabinofuranosyl-xylobiose (AX2) on the F5_SusD-like protein were negative (Fig. 2A to C), thereby confirming that small xylooligosaccharides do not bind to the protein. A beechwood xylan sample was then prepared, but its opacity indicated that even at 500 μM, the xylan chains assemble into macroscopic structures. The NMR experiment in the presence of the F5_SusD-like protein gave a robust STD signal, but this was also the case when we probed the same xylan sample in the absence of the protein (Fig. S4). Hence, we were unable to distinguish whether binding occurred between xylan and the transporter or between the macroscopic xylan superstructures. To reduce signal saturation due to these macroscopic xylan structures, we saturated the protein indole protons (at 10.1 ppm), where xylan signals could be expected to be reduced. The STD effect with xylan alone was indeed reduced (Fig. S5) and was not only more pronounced in the presence of F5_SusD-like but also increased when we added more protein (Fig. 2D). We

TABLE 1 Data collection and refinement statistics for the crystal structure of the F5_SusD-like protein^a

	Value	
Parameter	Native	S-SAD
Data collection statistics		
Wavelength (Å)	0.979	2.066
Space group	I4 ₁ 22	I4 ₁ 22
No. of molecules per asymmetric unit	1	1
Cell constants <i>a</i> , <i>b</i> , <i>c</i> (Å)	179.09, 179.09, 182.23	179.84, 179.84, 183.10
Resolution (Å) (range)	50.00–2.65	50.00–3.30
No. of measured reflections	307,172	2,114,678
No. of unique reflections	43,113	22,897
Data completeness (%)	100.0	100.0
Multiplicity (%)	7.1	92.4
<i>R</i> _{merge}	0.062 (0.736)	0.10 (0.41)
$\langle I/\sigma(I) \rangle$	8.4 (1.0)	5.2 (1.5)
CC _{1/2}	0.99 (0.86)	1.00 (0.99)
Wilson B-factor (Å ²)	65.3	73.7
Refinement statistics		
<i>R</i> _{work} / <i>R</i> _{free}	0.185/0.212	
RMSD bond length (Å)	0.070	
RMSD bond angle (°)	1.113	
Ramachandran plot favored/allowed/ outlier regions (%)	97/2/0	
No. of atoms		
Protein	4,849	
Water	16	
Other	6	
Avg B-factor (Å ²)	87.3	
Cruickshank's DPI	0.193	
MolProbity clashscore (percentile)	4 (100th)	

^aValues in parentheses are for the highest-resolution shell. DPI, diffraction precision index; S-SAD, sulfur single-wavelength anomalous diffraction phasing; CC_{1/2}, Pearson's correlation coefficient.

thus concluded that F5_SusD-like does not bind short xylooligosaccharides but binds weakly to xylan.

Crystallographic structure of F5_SusD-like. In order to better understand why F5_SusD-like does not bind the XOS substrate targeted by the PUL, we solved its three-dimensional structure by X-ray crystallography. The most similar sequence listed in the Protein Data Bank (PDB) (<https://www.rcsb.org/>) (37) is that of a SusD-like protein from *Parabacteroides distasonis* ATCC 8503 (PDB accession number 3OTN). The identity is limited, however, with only 34% identity on a short stretch of the sequence (128 amino acids), which made the structure solution by MR (molecular replacement) a challenging task. We therefore solved it using the MR-SAD (single-wavelength anomalous diffraction) technique (38) and obtained a refined structure at a 2.6-Å resolution. Data collection and refinement statistics are shown in Table 1. Despite the very low sequence similarity with structurally characterized SusD-like proteins, the F5_SusD-like protein displays a canonical “RagB/SusD” fold, which includes the conserved tetratricopeptide repeat (TPR) units. Overall, the structure of F5_SusD-like is similar to those of the 24 other SusD-like proteins that have been listed in the PDB since 2008, when the crystallographic structure of SusD was solved by Koropatkin et al. (7). F5_SusD-like presents a convex surface created by the α -helical bundle and a flat surface on the other side, where the putative sugar binding platform should be located. According to the PDB, no xylan-binding SusD-like protein has been structurally characterized so far. The closest structural homolog identified using the Dali server (<http://ekhidna2.biocenter.helsinki.fi/dali/>) (39) is the laminarin-binding protein GM_SusD from *Gramella* sp. strain MAR_2010_102 (PDB accession number 6GCZ) (40), with a root mean square deviation (RMSD) of about 2.9 Å. The superimposition of both structures highlights the presence of a large region (A328 to I340 and G382 to G444) specific for F5_SusD-like, which is composed

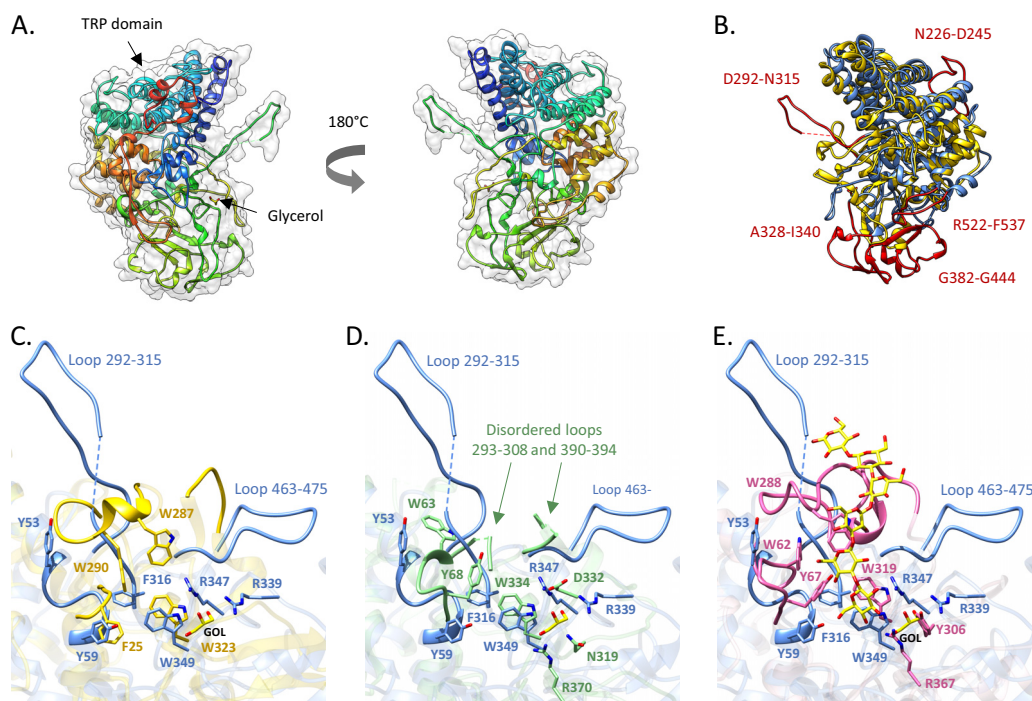


FIG 3 F5_SusD-like crystal structure. (A) Overall structure of F5_SusD-like. (B) Secondary-structure-matching (SSM) superimposition of F5_SusD-like (blue) with GM_SusD (PDB accession number 6GCZ) (the closest structural homolog) (yellow) shown to bind to branched laminarin and linear pustulan. The loops specific for F5_SusD-like are shown in red. (C) Zoomed-in view of the superimposition of F5_SusD-like with GM_SusD. (D) Superimposition of F5_SusD-like with the nonbinding SusD-like protein BuSGBP-A (PDB accession number 7KV1) (green) present in a PUL-degrading β (1,3)-glucan and a mixed-linkage β -glucan. (E) Superimposition of F5_SusD-like with BtSGBP-A (PDB accession number 7KV3) (pink) in complex with laminarihexaose. The glycerol (GOL) present in panels C to E belongs to the F5_SusD-like structure. The figures were created with Chimera (60).

of random coils and some β -sheets (Fig. 3). Another interesting feature is the presence of two loops (e.g., N226 to D245 and R522 to F537) that are inserted between the conserved secondary structural elements. Furthermore, the loop from D292 to N315, which protrudes away from the folded domain, is likely to be disordered in solution, although it is partially stabilized by interactions with the neighboring SusD-like molecules in the crystal (Fig. S6). Comparison with other structurally characterized members of the SusD family shows that these features make F5_SusD-like slightly larger albeit in the same size range as those of other SusD-like proteins from xylan-targeting PULs (Table S1).

We attempted to cocrystallize and soak F5_SusD-like crystals with XOSs ranging from DP2 to DP4, but no complex structure was obtained, confirming the inability of F5_SusD-like to bind to short-chain XOSs. Nevertheless, a glycerol molecule under crystallization conditions is found on the flat surface where the canonical binding site is located. The glycerol molecule is coordinated by two arginine residues (R339 and R347), and W349 makes a hydrophobic stacking platform (Fig. 3). Comparison with GM_SusD shows that two of the three tryptophan residues (W287 and W290) involved in substrate binding are missing in F5_SusD-like (Fig. 3C and Fig. S7A). Comparison to the Bacova_02651 SusD-like protein from *Bacteroides ovatus* (PDB accession number 5E76) (10), which displays a high glycan-binding ability toward xyloglucan (Fig. S7B), shows that an essential aromatic residue involved in substrate binding (W283) is missing in F5_SusD-like. Moreover, W75, which corresponds to W82 in Bacova_02651, displays a radically different side chain orientation.

Attempting to identify the molecular determinants that make F5_SusD-like a non-glycan-binding protein is particularly difficult, especially since no xylan-binding SusD-like protein is available in the PDB. Nevertheless, a striking feature of F5_SusD-like is the long partially disordered loop (D292 to N315) located close to the canonical

binding site. It is noteworthy that some of the tryptophan residues involved in substrate binding in the laminarin-binding SusD-like protein GM_SusD from *Gramella* sp. MAR_2010_102 (PDB accession number 6GCZ) (substrate-binding residues W287 and W290) (40) and in the xyloglucan-binding SGBP-A protein from *Bacteroides ovatus* (PDB accession number 5E76) (substrate-binding residue W283) (10), which are missing in F5_SusD-like, are carried by a folded loop located precisely in the area from D292 to N315. The presence of a disordered loop close to the binding site is reminiscent of the structure of the SGBP-A protein from *Bacteroides uniformis* (BuSGBP-A) (PDB accession number 7KV1) (RMSD, 3 Å) (41) (Fig. 3D). This SusD-like protein does not bind to $\beta(1,3)$ -glucan, while it displays the key binding residues present in the SGBP-A protein from *Bacteroides thetaiotaomicron* (BtSGBP-A) (PDB accession number 7KV3) (41), which were shown to be involved in substrate binding (Fig. 3E). Tamura et al. correlated this lack of $\beta(1,3)$ -glucan binding in BuSGBP-A with structural disorder (41). Indeed, three loops critical for shaping the binding site are structurally disordered or deformed in BuSGBP-A (positions 60 to 70, 293 to 308, and 390 to 394, corresponding to positions 54 to 58, 292 to 315, and 463 to 475 in F5_SusD-like).

Finally, the four residues N249, D444, A446, and A516 of F5_SusD-like, corresponding to D, G, P, and V residues in the nearly identical SusD-like protein BVU_0037 from *Bacteroides vulgatus* ATCC 8482, respectively, are not located in the binding-site area. The oligosaccharide-binding regions of canonical SusD-like proteins are thus strictly identical in F5_SusD-like and BVU_0037, indicating that BVU_0037, like F5_SusD-like, probably does not bind XOSs.

DISCUSSION

Despite the low sequence identity among SusD-like proteins, the SusC/D-like pair is used as a reference for PUL identification in genomes (14, 42). Indeed, substrate recognition and capture are key steps for glycoside utilization and for bacteria to colonize their habitat. So far, most of the SusD-like proteins that have been biochemically characterized are able to bind the polysaccharide targeted by the PUL but also, to some extent, oligosaccharides resulting from polysaccharide hydrolysis by cell-exposed enzymes. Only in rare cases have SusD-like proteins been shown to be unable to bind glycans, but their function has been investigated using only AGE or ITC (41, 43, 44). Here, we fully characterized a SusD homolog (F5_SusD-like) from an XOS PUL isolated from an uncultured bacterium, which shares 99.4% sequence identity with the *B. vulgatus* SusD-like protein BVU_0037. Previously, we demonstrated that the PUL containing BVU_0037 was involved in the ability of *B. vulgatus* to grow on XOS and, to a lesser extent, on branched arabinoxylooligosaccharides, while it utilizes xylans poorly (32). In the present study, we investigated the binding ability of the F5_SusD-like protein. Surprisingly, F5_SusD-like was not able to bind XOSs or AXOSs, while it exhibited a very low affinity for xylan. In contrast, the glycan-binding SusD-like proteins characterized to date display a strong binding affinity in the millimolar-to-micromolar range, depending on the length of the tested substrate. In most cases, the values obtained for complex polysaccharides are in the same range as the values obtained with the oligosaccharides resulting from polysaccharide hydrolysis (10, 45). Rogowski et al. characterized two SusD-like proteins involved in two distinct xylan PULs from *B. ovatus* (Bacova_04392 and Bacova_03427) (21). In contrast to the F5_SusD-like protein, which belongs to a PUL specific for XOS, these SusD-like proteins showed strong abilities to bind xylan, measurable by ITC, but only Bacova_04392 binds xylooligosaccharides no smaller than DP6 (K_D [equilibrium dissociation constant] = $8.0 \times 10^2 \text{ M}^{-1}$ for xylohexaose) (21).

Several studies previously showed that SusD-like proteins are essential for substrate uptake via the SusC-like transporter (7, 9, 10). This is also the case for the F5_SusC/D-like system since when the F5_SusD-like gene is deleted, the F5min_Sus Δ SusD strain can no longer grow on XOS (32). The F5_SusC/D-like pair is thus able to transport XOS, although the F5_SusD-like protein does not bind them. If the presence of the SusD-like

proteins is crucial for glycoside transport, in *Bacteroidetes*, their binding ability is not always required, for the reasons explained here. Both the presence and binding ability of the SusD homolog from the *B. ovatus* transport system are essential for growth on mixed-linkage β -glucans (9). In contrast, genetic complementation using nonbinding SusD-like mutants is sufficient to restore the functionality of the starch transport system in *B. thetaiotaomicron* and the xyloglucan transport system in *B. ovatus* (7, 10).

As explained previously by Foley and coworkers regarding the Sus (46) and recently by Gray et al. regarding the fructooligosaccharide transport system in *B. thetaiotaomicron* (12), it might be because in *Bacteroidetes*, the cell surface glycan-binding proteins functionally complement each other in the capture of the polysaccharide. Indeed, Foley et al. (46) showed that when the SusD, -E, and -F proteins of the *B. thetaiotaomicron* Sus were mutated to nonbinding alleles, growth still occurred on maltooligosaccharides, even in a SusG deletion background. In that study, those authors demonstrated that SusE, independent of its starch-binding sites, can restore glycan uptake by the SusC/D system when the SusD starch-binding ability is abolished. In this context, the most probable hypothesis is that SusE interacts with the SusC/D complex, allowing SusD to open for import. Besides, in the rare examples of nonbinding SusD-like proteins that have been identified so far from loci involved in pectin and β -1,3-glucan utilization, there are additional SGBP- or SusD-like protein-encoding genes in the same locus, suggesting that they compensated the system for glycan binding (43, 44). However, two other nonbinding SusD-like proteins (BT2625 and BT3855) have been identified in the α -mannan utilization loci MAN-PUL1 and MAN-PUL3 of *B. thetaiotaomicron*, which do not contain any other SusD-like protein or SGBP (47). One cannot exclude that mannan-binding outer surface proteins encoded by MAN-PUL2, which is coexpressed with MAN-PUL1 and MAN-PUL3 during α -mannan utilization, could compensate for the nonbinding SusD-like proteins BT2625 and BT3855 to capture the targeted polysaccharide. Nevertheless, regarding SusD-like proteins that naturally do not bind glycans as isolated proteins, another hypothesis is that their interaction with glycans requires the presence of the cognate SusC-like protein. This hypothesis is in accordance with the data reported previously by Tauzin et al. (10), who showed by reverse genetic analysis that the presence of the SusD-like SGBP-A protein is more critical than its carbohydrate-binding ability for growth on xyloglucan, while the SGBP-B protein is not essential, although it supports the efficient capture of xyloglucooligosaccharides. In the present paper, we confirm that a SusC/D-like transport system should be functional for the uptake of the substrate targeted by the PUL without any binding ability of SusD-like or any other SGBP, supporting a model of glycan import whereby at least the SusC-like and SusD-like proteins must be associated to support glycan uptake.

In order to better understand the molecular determinants of the very weak binding ability of F5_SusD-like, we determined its crystal structure. Although F5_SusD-like could not be cocrystallized or soaked with XOS, a glycerol molecule was found at the canonical oligosaccharide-binding site, bound by R339, R347, and W349, which makes a hydrophobic stacking platform. This stacking platform might be involved in the very weak binding of xylans by F5_SusD-like. However, by comparing it with solved SusD-like structures, we hypothesized that the lack of genuine binding affinity could be due to some missing aromatic residues as well as an unfolded binding surface with the loop spanning D292 to N315 being partially disordered. Also, SusD-like affinity is driven by the surface complementation between the protein and the substrate rather than by the affinity of individual chemical groups for the substrate, as proposed in the very first description of the structure of SusD-like (7) and confirmed subsequently (9, 13, 40, 48–50). Thus, the role of “nonbinding” F5_SusD-like appears to be to effectively complement the SusC-like transporter and activate its functionality. This was first proposed by Glenwright et al. (13), who solved the crystal structure of a quaternary complex from *B. thetaiotaomicron* BT2261–BT2264, including the SusD-like (BT2263) and SusC-like (BT2264) transporters. Using a variety of techniques, those authors proposed a “pedal bin” mechanism in which SusD-like moves away from SusC-like in a hinge-like fashion in the absence of a ligand to expose the substrate-binding site to the extracellular milieu. According to this model, which was further confirmed by Gray et al.

(12), SusC-like performs its function only when in complex with SusD-like, which acts as a lid that can open and close the transporter. Because SusC-like and SusD-like are associated proteins *in vivo*, the most probable hypothesis for SusD-like proteins that are not able to bind glycans as isolated proteins is that binding requires the presence of the cognate SusC-like protein. Indeed, the structural analysis of SusD-like BT1762 alone and in a complex with SusC-like BT1763 shows that a disordered loop located in the same area of positions 292 to 315 in F5_SusD-like is stabilized in the complex by the long hinge 2 loop of BT1763 (see Fig. S8 in the supplemental material). The hinge 2 loop is likely to be important for lid opening and is responsible, with hinge 1, for the majority of the interactions between BT1762 and BT1763 in the open state (12). Moreover, hinge 2 carries a phenylalanine residue (F649) making a hydrophobic interaction with the substrate. The proximity of this connection area to the SusD-like canonical binding site suggests that glycan recognition requires the presence of the cognate SusC-like protein. This could also be a hypothesis to explain why a single substitution that abolishes levan binding to BT1762 (W85A) *in vitro* leads to no growth defect *in vivo*. As discussed by Gray et al. (12), the context of the intact transporter ensures that the effects of SusD-like point mutations are much less dramatic *in vivo*. Among the amino acids composing the partially disordered loop at positions 292 to 315 in F5_SusD-like are 1 histidine, 3 tyrosine, and 2 glutamate residues. It would be interesting to know whether these residues are involved in the interaction with F5_SusC-like and/or whether the loop is folded in a way that enables some of these residues to form an extended binding site. Obtaining the three-dimensional (3D) structure of F5_SusC/D-like in complex with XOS would allow us to answer this question.

Finally, nonbinding or very weakly binding SusD-like proteins can be found in PULs, including (i) a SusC/D-like pair and an additional SGBP (43), (ii) two SusC/D-like pairs and an SGBP (44), or (iii) only one SusC/D-like pair (47; this study). The presence of an additional SGBP or a SusC/D-like pair in the same locus as that of a nonbinding SusD-like protein suggests compensation for the binding function lacking in the related SusD-like protein. In contrast, as previously hypothesized (47), the PULs containing only a SusC/D-like pair for which SusD-like, alone, is a nonbinding protein and no other SGBP might be involved in oligosaccharide or polysaccharide utilization that relies on the ability of the SusC/D-like complex to recognize and transport its cognate substrate.

MATERIALS AND METHODS

Heterologous protein expression and purification. The gene fragment encoding the F5_SusD-like protein was PCR amplified from the metagenomic DNA of clone F5 (GenBank accession number [HE717017](#)), previously isolated from a metagenomic library derived from a human fecal sample (31). The PCR forward primer F5_SusD_Cloning_F (see Table S2 in the supplemental material) includes the NdeI restriction site, and the reverse primer F5_SusD_Cloning_R includes the XhoI site. The first 30 amino acids of F5_SusD-like corresponding to the lipoprotein signal peptide predicted by the SignalP 4.1 and LipoP 1.0 servers were not included in the PCR amplicon (51, 52). The gene product was cloned into the NdeI and XhoI restriction sites of the pET-28a(+) vector (Novagen, Darmstadt, Germany) preceding an N-terminal 6×His tag for affinity purification and transformed into *E. coli* DH5α (Invitrogen). The pET-28a(+) vector expressing F5_SusD-like was fully sequenced and transformed into *E. coli* BL21(DE3) for protein production. A single bacterial colony was inoculated into 5 mL of LB medium containing kanamycin (50 μg/mL) at 37°C. A preculture grown overnight was used to inoculate 200 mL of LB medium containing 50 μg/mL of kanamycin at an initial optical density at 600 nm (OD₆₀₀) of 0.05. The culture was induced by the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) in the mid-exponential phase (OD₆₀₀ of ~0.6) and then incubated for 4 h at 37°C.

For the purification of the recombinant F5_SusD-like protein, cells were harvested by centrifugation (15 min at 6,000 × *g*) and sonicated in Tris buffer (20 mM Tris-HCl, 300 mM NaCl [pH 8.0]). Bacterial debris was cleared by centrifugation at 11,000 × *g* for 30 min at 4°C, and the lysates were then passed through a column of 2 mL of Talon metal affinity resin (Clontech, USA). The F5_SusD-like protein was eluted using a gradient of 50 to 200 mM imidazole and then resuspended in Tris buffer (20 mM Tris-HCl, 150 mM NaCl [pH 7.5]) using an Amicon Ultra filter (Sigma-Aldrich) to eliminate the imidazole. Protein purity was confirmed via SDS-PAGE, and the concentration was determined by the absorbance at 280 nm using the Thermo Scientific Nanodrop 2000 instrument with an extinction coefficient of 153,350 M⁻¹ cm⁻¹. The F5_SusD-like protein was kept at 4°C for later analysis. The protein used for crystallization experiments was purified using an ÄKTAexpress system (GE Healthcare) with an affinity step (Talon crude, 1 mL; GE Healthcare) using the same loading buffer (20 mM Tris-HCl, 300 mM NaCl [pH 8.0]) and eluted with a 250 mM step of imidazole, followed by a gel filtration step (HiPrep 16/60 Sephacryl S-200 HR) in a solution containing 20 mM Tris-HCl and 0.15 M NaCl (pH 7.5). The fractions containing the monomeric F5_SusD-like protein were pooled, concentrated, and stored at 4°C.

Glycosides. The mixture of XOSs contains chains of DP2 to -7 (Wako Chemicals and IRO TAIHE). Simple xylans, with sparsely decorated structures, were purchased from Sigma for beechwood xylan, Biochemika Fluka for oat spelt xylan, and Megazyme for birchwood xylan and wheat arabinoxylan. Xylobiose, xylotriose, and 3²- α -L-arabinofuranosyl-xylobiose (AX2) were purchased from Megazyme. Barley β -glucan and xyloglucan from tamarind seed were also purchased from Megazyme. Laminarin from *Laminaria digitata* was purchased from Sigma.

Isothermal titration calorimetry. Isothermal titration calorimetry (ITC) analysis of glycan binding by F5_SusD-like was performed using the ITC200 calorimeter (Malvern), calibrated to 25°C. Proteins (40 μ M) were prepared in buffer containing 20 mM HEPES–100 mM NaCl (pH 7.0), and oligosaccharides were prepared using the same buffer. The F5_SusD-like protein was placed into the sample cell, and the syringe was loaded with 2 to 10 mM XOS.

Affinity gel electrophoresis. The ability of F5_SusD-like to bind polysaccharides was assayed by affinity gel electrophoresis. Continuous native polyacrylamide gels were prepared, consisting of 10% (wt/vol) acrylamide in buffer containing 25 mM Tris and 250 mM glycine (pH 8.3). The substrates were added prior to polymerization at a final concentration of 0.5% (wt/vol). A total of 2.5 μ g of purified F5_SusD-like was loaded onto the gels. Electrophoresis was carried out for 90 min on ice. Bovine serum albumin (BSA) was used as a noninteracting negative-control protein. The percentage of retention of F5_SusD-like in the gel containing glycans was calculated as $(D_{\text{wog}} - D_{\text{wg}})/D_{\text{wog}} \times 100$, where D_{wog} is equal to $D_{\text{SusD-like_wog}}/D_{\text{BSA_wog}}$ (the distance of migration of the lowest and most intense F5_SusD-like band normalized to the distance of migration of BSA in the gel without glycan) and D_{wg} is equal to $D_{\text{SusD-like_wg}}/D_{\text{BSA_wg}}$ (the distance of migration of the lowest and most intense F5_SusD-like band normalized to the distance of migration of BSA in the gel with glycan).

NMR spectroscopy. All experiments were performed on a Bruker Avance II 800-MHz NMR spectrometer equipped with a QCPD cryogenically cooled probe head. Spectra were recorded at 298 K, and all samples were prepared in Tris buffer (20 mM Tris and 150 mM NaCl at pH 7.5). 1D proton spectra were acquired using Watergate water suppression (53), with 16 scans for XG_SusD-like (at 327 μ M) and 256 scans for F5_SusD-like (at 54 μ M). STD experiments were performed with saturation of the protein resonances at 0 ppm through a train of 5-ms Gaussian 180° pulses (34, 54). Typical experiments were run with 256 scans, 2,048 acquisition points, and a 5-s relaxation delay (including the 3-s presaturation train). STD experiments were run with 256 or 512 scans. Spectra were transformed after one level of zero filling and apodization with a $\pi/3$ -shifted square sine bell.

Crystallization and data collection. The purified F5_SusD-like protein was concentrated to 20 mg/mL, and initial crystallization conditions were screened using a Mosquito robot (TPP Labtech) and JCSG I to IV commercial screens (Qiagen), from which one starting condition was identified (1.26 M sodium citrate, 0.09 M HEPES-HCl, 10% glycerol [pH 7.5]). After manual condition optimization, F5_SusD-like crystals were grown using the hanging-drop method by mixing 1 μ L of the protein solution with 1 μ L of the precipitant solution (1.7 M Na₃-citrate, 10% glycerol, 0.1 M HEPES [pH 7.5]) and incubating the solution on a 24-well crystallization plate at 12°C. After harvesting, the crystals were soaked in cryoprotectant buffer (1.7 M Na₃-citrate, 15% glycerol, 0.1 M HEPES [pH 7.5]), with or without the presence of 100 mM xylotriose, and flash-frozen in liquid N₂. Data were collected at the XALOC beamline of the ALBA Synchrotron Light Source (Barcelona, Spain). A long-wavelength data set was also collected at a wavelength of 2.0 Å, using a special XALOC setup with continuous helium flow around the crystal to reduce absorption from the air (helium cone).

Structure resolution and refinement. Using the native data set only, the structure was solved using the Morda Web server (55), which identified a low-homology molecular replacement solution. This starting model was combined with the long-wavelength data set as implemented in the Phaser program (56) and used to calculate improved phases. This also enabled the correct positioning of all of the sulfur atoms in the sequence. The structure was determined via several rounds of modeling using the Arp/Warp Web server (57), alternated with manual rebuilding using Coot (58), and passed to Refmac for restrained refinement (59). The final model was validated with MolProbity (60) and WhatIf (61).

Data availability. The sequence of F5_SusD-like is available in the GenBank database under accession number [CCG34975.1](#). The crystal structure of F5_SusD-like has been deposited in the Protein Data Bank under accession number [7NEK](#).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, TIF file, 0.6 MB.

FIG S2, TIF file, 0.4 MB.

FIG S3, TIFF file, 0.01 MB.

FIG S4, TIF file, 0.2 MB.

FIG S5, TIF file, 0.2 MB.

FIG S6, TIF file, 0.3 MB.

FIG S7, TIF file, 0.8 MB.

FIG S8, TIF file, 1.2 MB.

TABLE S1, DOCX file, 0.01 MB.

TABLE S2, DOCX file, 0.01 MB.

ACKNOWLEDGMENTS

This work was carried out using the equipment at the PICT-ICEO facility and at MetaToul (Metabolomics and Fluxomics Facilities, Toulouse, France [www.metatoul.fr]). PICT-ICEO is supported by grants from the Région Midi-Pyrénées, the European Regional Development Fund, and the Institut National de la Recherche Agronomique et de l'Environnement (INRAE). MetaToul is part of the MetaboHUB-ANR-11-INBS-0010 national infrastructure (www.metabohub.fr). This research was funded by the European Union's Horizon 2020 Framework Programme (MSCA-IF-2015_707457 [CaTSYS] and LEIT-BIO-2015-685474 [Metafluidics]). Z.W. and X.L. are supported by INSA Toulouse and the China Scholarship Council under the CSC-UT/INSA Program.

We acknowledge PICT-IPBS (Toulouse, France) and the ALBA Synchrotron Light Source (Barcelona, Spain) for access to crystallization and data collection facilities, respectively.

REFERENCES

- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42:D490–D495. <https://doi.org/10.1093/nar/gkt1178>.
- Terrapon N, Lombard V, Gilbert HJ, Henrissat B. 2015. Automatic prediction of polysaccharide utilization loci in Bacteroidetes species. *Bioinformatics* 31:647–655. <https://doi.org/10.1093/bioinformatics/btu716>.
- Grondin JM, Tamura K, Déjean G, Abbott DW, Brumer H. 2017. Polysaccharide utilization loci: fueling microbial communities. *J Bacteriol* 199:e00860-16. <https://doi.org/10.1128/JB.00860-16>.
- Martens EC, Koropatkin NM, Smith TJ, Gordon JL. 2009. Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. *J Biol Chem* 284:24673–24677. <https://doi.org/10.1074/jbc.R109.022848>.
- Brown HA, Koropatkin NM. 2021. Host glycan utilization within the Bacteroidetes Sus-like paradigm. *Glycobiology* 31:697–706. <https://doi.org/10.1093/glycob/cwaa054>.
- Foley MH, Cockburn DW, Koropatkin NM. 2016. The Sus operon: a model system for starch uptake by the human gut Bacteroidetes. *Cell Mol Life Sci* 73:2603–2617. <https://doi.org/10.1007/s00018-016-2242-x>.
- Koropatkin NM, Martens EC, Gordon JL, Smith TJ. 2008. Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. *Structure* 16:1105–1115. <https://doi.org/10.1016/j.str.2008.03.017>.
- Sonnenburg ED, Zheng H, Joglekar P, Higginbottom SK, Firbank SJ, Bolam DN, Sonnenburg JL. 2010. Specificity of polysaccharide use in intestinal Bacteroides species determines diet-induced microbiota alterations. *Cell* 141:1241–1252. <https://doi.org/10.1016/j.cell.2010.05.005>.
- Tamura K, Foley MH, Gardill BR, Dejean G, Schnitzlein M, Bahr CME, Creagh AL, van Petegem F, Koropatkin NM, Brumer H. 2019. Surface glycan-binding proteins are essential for cereal beta-glucan utilization by the human gut symbiont Bacteroides ovatus. *Cell Mol Life Sci* 76:4319–4340. <https://doi.org/10.1007/s00018-019-03115-3>.
- Tauzin AS, Kwiatkowski KJ, Orlovsky NI, Smith CJ, Creagh AL, Haynes CA, Wawrzak Z, Brumer H, Koropatkin NM. 2016. Molecular dissection of xyloglucan recognition in a prominent human gut symbiont. *mBio* 7:e02134-15. <https://doi.org/10.1128/mBio.02134-15>.
- Cho KH, Salyers AA. 2001. Biochemical analysis of interactions between outer membrane proteins that contribute to starch utilization by Bacteroides thetaiotaomicron. *J Bacteriol* 183:7224–7230. <https://doi.org/10.1128/JB.183.24.7224-7230.2001>.
- Gray DA, White JBR, Oluwole AO, Rath P, Glenwright AJ, Mazur A, Zahn M, Baslé A, Morland C, Evans SL, Cartmell A, Robinson CV, Hiller S, Ranson NA, Bolam DN, van den Berg B. 2021. Insights into SusCD-mediated glycan import by a prominent gut symbiont. *Nat Commun* 12:44. <https://doi.org/10.1038/s41467-020-20285-y>.
- Glenwright AJ, Pothula KR, Bhamidimarri SP, Chorev DS, Baslé A, Firbank SJ, Zheng H, Robinson CV, Winterhalter M, Kleinekathöfer U, Bolam DN, Van den Berg B. 2017. Structural basis for nutrient acquisition by dominant members of the human gut microbiota. *Nature* 541:407–411. <https://doi.org/10.1038/nature20828>.
- Terrapon N, Lombard V, Drula E, Lapébie P, Al-Masaudi S, Gilbert HJ, Henrissat B. 2018. PULDB: the expanded database of polysaccharide utilization loci. *Nucleic Acids Res* 46:D677–D683. <https://doi.org/10.1093/nar/gkx1022>.
- Despres J, Forano E, Lepercq P, Comtet-Marre S, Jubelin G, Chambon C, Yeoman CJ, Berg Miller ME, Fields CJ, Martens E, Terrapon N, Henrissat B, White BA, Mosoni P. 2016. Xylan degradation by the human gut Bacteroides xylanisolvens XB1AT involves two distinct gene clusters that are linked at the transcriptional level. *BMC Genomics* 17:326. <https://doi.org/10.1186/s12864-016-2680-8>.
- Dodd D, Moon Y-H, Swaminathan K, Mackie RI, Cann IKO. 2010. Transcriptomic analyses of xylan degradation by Prevotella bryantii and insights into energy acquisition by xylanolytic Bacteroidetes. *J Biol Chem* 285:30261–30273. <https://doi.org/10.1074/jbc.M110.141788>.
- Martens EC, Lowe EC, Chiang H, Pudlo NA, Wu M, McNulty NP, Abbott DW, Henrissat B, Gilbert HJ, Bolam DN, Gordon JL. 2011. Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. *PLoS Biol* 9:e1001221. <https://doi.org/10.1371/journal.pbio.1001221>.
- McNulty NP, Wu M, Erickson AR, Pan C, Erickson BK, Martens EC, Pudlo NA, Muegge BD, Henrissat B, Hettich RL, Gordon JL. 2013. Effects of diet on resource utilization by a model human gut microbiota containing Bacteroides cellulosilyticus WH2, a symbiont with an extensive glycobiome. *PLoS Biol* 11:e1001637. <https://doi.org/10.1371/journal.pbio.1001637>.
- Mendis M, Martens EC, Simsek S. 2018. How fine structural differences of xylooligosaccharides and arabinoxylooligosaccharides regulate differential growth of Bacteroides species. *J Agric Food Chem* 66:8398–8405. <https://doi.org/10.1021/acs.jafc.8b01263>.
- Mirande C, Kadlecikova E, Matulova M, Capek P, Bernalier-Donadille A, Forano E, Béra-Maillet C. 2010. Dietary fibre degradation and fermentation by two xylanolytic bacteria Bacteroides xylanisolvens XB1AT and Roseburia intestinalis XB6B4 from the human intestine. *J Appl Microbiol* 109:451–460. <https://doi.org/10.1111/j.1365-2672.2010.04671.x>.
- Rogowski A, Briggs JA, Mortimer JC, Tryfona T, Terrapon N, Lowe EC, Baslé A, Morland C, Day AM, Zheng H, Rogers TE, Thompson P, Hawkins AR, Yadav MP, Henrissat B, Martens EC, Dupree P, Gilbert HJ, Bolam DN. 2015. Glycan complexity dictates microbial resource allocation in the large intestine. *Nat Commun* 6:7481. <https://doi.org/10.1038/ncomms8481>.
- Salyers AA, Gherardini F, O'Brien M. 1981. Utilization of xylan by two species of human colonic Bacteroides. *Appl Environ Microbiol* 41:1065–1068. <https://doi.org/10.1128/aem.41.4.1065-1068.1981>.
- Zhang M, Lin Y, Han Y, Jiao N. 2017. Characterization of potential polysaccharide utilization systems in the marine Bacteroidetes Gramella flava JLT2011 using a multi-omics approach. *Front Microbiol* 8:220. <https://doi.org/10.3389/fmicb.2017.00220>.
- Zhang M, Chekan JR, Dodd D, Hong P-Y, Radlinsk L, Revindran V, Nair SK, Mackie RI, Cann I. 2014. Xylan utilization in human gut commensal bacteria is orchestrated by unique modular organization of polysaccharide-degrading enzymes. *Proc Natl Acad Sci U S A* 111:E3708–E3717. <https://doi.org/10.1073/pnas.1406156111>.
- Leth ML, Ejby M, Workman C, Ewald DA, Pedersen SS, Sternberg C, Bahl MI, Licht TR, Aachmann FL, Westereng B, Hachem MA. 2018. Differential bacterial capture and transport preferences facilitate co-growth on dietary xylan in the human gut. *Nat Microbiol* 3:570–580. <https://doi.org/10.1038/s41564-018-0132-8>.
- Selvendran RR. 1984. The plant cell wall as a source of dietary fiber: chemistry and structure. *Am J Clin Nutr* 39:320–337. <https://doi.org/10.1093/ajcn/39.2.320>.
- Singh RD, Banerjee J, Arora A. 2015. Prebiotic potential of oligosaccharides: a focus on xylan derived oligosaccharides. *Bioact Carbohydr Diet Fibre* 5:19–30. <https://doi.org/10.1016/j.bcdf.2014.11.003>.

28. Amorim C, Silvério SC, Prather KLJ, Rodrigues LR. 2019. From lignocellulosic residues to market: production and commercial potential of xylooligosaccharides. *Biotechnol Adv* 37:107397. <https://doi.org/10.1016/j.biotechadv.2019.05.003>.
29. Gibson GR, Hutkins R, Sanders ME, Prescott SL, Reimer RA, Salminen SJ, Scott K, Stanton C, Swanson KS, Cani PD, Verbeke K, Reid G. 2017. Expert consensus document: the International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics. *Nat Rev Gastroenterol Hepatol* 14:491–502. <https://doi.org/10.1038/nrgastro.2017.75>.
30. Goh YJ, Klaenhammer TR. 2015. Genetic mechanisms of prebiotic oligosaccharide metabolism in probiotic microbes. *Annu Rev Food Sci Technol* 6:137–156. <https://doi.org/10.1146/annurev-food-022814-015706>.
31. Cecchini DA, Laville E, Laguerre S, Robe P, Leclerc M, Doré J, Henrissat B, Remaud-Siméon M, Monsan P, Potocki-Véronèse G. 2013. Functional metagenomics reveals novel pathways of prebiotic breakdown by human gut bacteria. *PLoS One* 8:e72766. <https://doi.org/10.1371/journal.pone.0072766>.
32. Tauzin AS, Laville E, Xiao Y, Nouaille S, Le Bourgeois P, Heux S, Portais JC, Monsan P, Martens EC, Potocki-Veronese G, Bordes F. 2016. Functional characterization of a gene locus from an uncultured gut Bacteroides conferring xylo-oligosaccharides utilization to *Escherichia coli*. *Mol Microbiol* 102:579–592. <https://doi.org/10.1111/mmi.13480>.
33. Sun C, Liu M, Sun P, Yang M, Yates EA, Guo Z, Fernig DG. 2019. Sulfated polysaccharides interact with fibroblast growth factors and protect from denaturation. *FEBS Open Bio* 9:1477–1487. <https://doi.org/10.1002/2211-5463.12696>.
34. Mayer M, Meyer B. 1999. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew Chem Int Ed Engl* 38:1784–1788. [https://doi.org/10.1002/\(SICI\)1521-3773\(19990614\)38:12%3C1784::AID-ANIE1784%3E3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1521-3773(19990614)38:12%3C1784::AID-ANIE1784%3E3.0.CO;2-Q).
35. Bell A, Brunt J, Crost E, Vaux L, Nepravishta R, Owen CD, Latousakis D, Xiao A, Li W, Chen X, Walsh MA, Claesen J, Angulo J, Thomas GH, Juge N. 2019. Elucidation of a sialic acid metabolism pathway in mucus-foraging *Ruminococcus gnavus* unravels mechanisms of bacterial adaptation to the gut. *Nat Microbiol* 4:2393–2404. <https://doi.org/10.1038/s41564-019-0590-7>.
36. Haselhorst T, Lamerz A-C, von Itzstein M. 2009. Saturation transfer difference NMR spectroscopy as a technique to investigate protein-carbohydrate interactions in solution. *Methods Mol Biol* 534:375–386. https://doi.org/10.1007/978-1-59745-022-5_26.
37. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardecki C. 2002. The Protein Data Bank. *Acta Crystallogr D Biol Crystallogr* 58:899–907. <https://doi.org/10.1107/s0907444902003451>.
38. Skubák P, Araç D, Bowler MW, Correia AR, Hoelz A, Larsen S, Leonard GA, McCarthy AA, McSweeney S, Mueller-Dieckmann C, Otten H, Salzmann G, Pannu NS. 2018. A new MR-SAD algorithm for the automatic building of protein models from low-resolution X-ray data and a poor starting model. *IUCr* 5:166–171. <https://doi.org/10.1107/S2052252517017961>.
39. Holm L. 2019. Benchmarking fold detection by DALI Lite v.5. *Bioinformatics* 35:5326–5327. <https://doi.org/10.1093/bioinformatics/btz536>.
40. Mystkowska AA, Robb C, Vidal-Melgosa S, Vanni C, Fernandez-Guerra A, Höhne M, Hehemann J-H. 2018. Molecular recognition of the beta-glucans laminarin and pustulan by a SusD-like glycan-binding protein of a marine Bacteroidetes. *FEBS J* 285:4465–4481. <https://doi.org/10.1111/febs.14674>.
41. Tamura K, Dejean G, Van Petegem F, Brumer H. 2021. Distinct protein architectures mediate species-specific beta-glucan binding and metabolism in the human gut microbiota. *J Biol Chem* 296:100415. <https://doi.org/10.1016/j.jbc.2021.100415>.
42. Martens EC, Chiang HC, Gordon JI. 2008. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host Microbe* 4:447–457. <https://doi.org/10.1016/j.chom.2008.09.007>.
43. Dejean G, Tamura K, Cabrera A, Jain N, Pudlo NA, Pereira G, Viborg AH, Van Petegem F, Martens EC, Brumer H. 2020. Synergy between cell surface glycosidases and glycan-binding proteins dictates the utilization of specific beta(1,3)-glucans by human gut Bacteroides. *mBio* 11:e00095-20. <https://doi.org/10.1128/mBio.00095-20>.
44. Luis AS, Briggs J, Zhang X, Farnell B, Ndeh D, Labourel A, Baslé A, Cartmell A, Terrapon N, Stott K, Lowe EC, McLean R, Shearer K, Schückel J, Venditto I, Ralet MC, Henrissat B, Martens EC, Mosimann SC, Abbott DW, Gilbert HJ. 2018. Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic Bacteroides. *Nat Microbiol* 3:210–219. <https://doi.org/10.1038/s41564-017-0079-1>.
45. Temple MJ, Cuskin F, Baslé A, Hickey N, Speciale G, Williams SJ, Gilbert HJ, Lowe EC. 2017. A Bacteroidetes locus dedicated to fungal 1,6-glucan degradation: unique substrate conformation drives specificity of the key endo-1,6-glucanase. *J Biol Chem* 292:10639–10650. <https://doi.org/10.1074/jbc.M117.787606>.
46. Foley MH, Martens EC, Koropatkin NM. 2018. SusE facilitates starch uptake independent of starch binding in *B. theta* taomicon. *Mol Microbiol* 108:551–566. <https://doi.org/10.1111/mmi.13949>.
47. Cuskin F, Lowe EC, Temple MJ, Zhu Y, Cameron E, Pudlo NA, Porter NT, Urs K, Thompson AJ, Cartmell A, Rogowski A, Hamilton BS, Chen R, Tolbert TJ, Piens K, Bracke D, Vervecken W, Hakki Z, Speciale G, Munöz-Munöz JL, Day A, Peña MJ, McLean R, Suits MD, Boraston AB, Atherly T, Ziemer CJ, Williams SJ, Davies GJ, Abbott DW, Martens EC, Gilbert HJ. 2015. Human gut Bacteroidetes can utilize yeast mannans through a selfish mechanism. *Nature* 517:165–169. <https://doi.org/10.1038/nature13995>.
48. Larsbrink J, Zhu Y, Kharade SS, Kwiatkowski KJ, Eijsink VGH, Koropatkin NM, McBride MJ, Pope PB. 2016. A polysaccharide utilization locus from *Flavobacterium johnsoniae* enables conversion of recalcitrant chitin. *Bio-technol Biofuels* 9:260. <https://doi.org/10.1186/s13068-016-0674-z>.
49. Phansopa C, Roy S, Rafferty JB, Douglas CWI, Pandhal J, Wright PC, Kelly DJ, Stafford GP. 2014. Structural and functional characterization of NanU, a novel high-affinity sialic acid-inducible binding protein of oral and gut-dwelling Bacteroidetes species. *Biochem J* 458:499–511. <https://doi.org/10.1042/BJ20131415>.
50. Koropatkin N, Martens EC, Gordon JI, Smith TJ. 2009. Structure of a SusD homologue, BT1043, involved in mucin O-glycan utilization in a prominent human gut symbiont. *Biochemistry* 48:1532–1542. <https://doi.org/10.1021/bi801942a>.
51. Juncker AS, Willenbrock H, von Heijne G, Brunak S, Nielsen H, Krogh A. 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci* 12:1652–1662. <https://doi.org/10.1110/ps.0303703>.
52. Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8:785–786. <https://doi.org/10.1038/nmeth.1701>.
53. Piotto M, Saudek V, Sklenár V. 1992. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR* 2:661–665. <https://doi.org/10.1007/BF02192855>.
54. Ley NB, Rowe ML, Williamson RA, Howard MJ. 2014. Optimising selective excitation pulses to maximise saturation transfer difference NMR spectroscopy. *RSC Adv* 4:7347–7351. <https://doi.org/10.1039/C3RA46246C>.
55. Hungler A, Momin A, Diederichs K, Arold ST. 2016. ContaMiner and ContaBase: a webserver and database for early identification of unwantedly crystallized protein contaminants. *J Appl Crystallogr* 49:2252–2258. <https://doi.org/10.1107/S1600576716014965>.
56. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J Appl Crystallogr* 40:658–674. <https://doi.org/10.1107/S0021889807021206>.
57. Perrakis A, Sixma TK, Wilson KS, Lamzin VS. 1997. wARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models. *Acta Crystallogr D Biol Crystallogr* 53:448–455. <https://doi.org/10.1107/S0907444997005696>.
58. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486–501. <https://doi.org/10.1107/S0907444910007493>.
59. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA. 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67:355–367. <https://doi.org/10.1107/S0907444911001314>.
60. Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA, Hintze BJ, Chen VB, Jain S, Lewis SM, Arendall WB, III, Snoeyink J, Adams PD, Lovell SC, Richardson JS, Richardson DC. 2018. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci* 27:293–315. <https://doi.org/10.1002/pro.3330>.
61. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612. <https://doi.org/10.1002/jcc.20084>.

Supplemental data

Table S1. Molecular weight comparison of F5_SusD-like with SusD-like proteins from xylan targeting PULs.

Protein name	% identity with F5_SusD-like (mature proteins)	Molecular weight (kDa) (mature proteins)
F5_SusD-like	100	69.2
BACDOR_00535	99.68	69
BVU_0037	99.36	69.1
BACINT_01037⁽¹⁾	26.01	62.4
HMPREF9447_02533⁽¹⁾	25.34	65
*BACOVA_04392⁽²⁾	23.89	59.8
BACCELL_02148⁽¹⁾	22.15	63.5
BACOVA_03429⁽²⁾	21.94	65.2
BXY_29260⁽³⁾	21.39	69.5
BXY_29280⁽³⁾	20.76	63.9
*BACOVA_03427⁽²⁾	19.43	69

*Experimentally shown to bind xylan. BACOVA_04392 binds to relatively simple xyans, xylohexaose but not shorter xylooligosaccharides. BACOVA_03427 binds to highly decorated corn glucuronoarabinoxylan but not linear xylooligosaccharides.

⁽¹⁾ From Pereira *et al.*, Nature Communications, 2021 (DOI 10.1038/s41467-020-20737-5)

⁽²⁾ From Rogowski *et al.*, Nature, 2015 (DOI 10.1038/ncomms8481)

⁽³⁾ From Despres *et al.*, BMC Genomics, 2016 (DOI 10.1186/s12864-016-2680-8)

Table S2. List of PCR primers used in this study.

Name	Sequence 5'→3'
F5_SusD_Cloning_F	CCGCGCGGCAGCCATATGGACGAGCAGCCGCGCAGCAGTTATG
F5_SusD_Cloning_R	GTGGTGGTGGTGCTCGAGTCAATCTCTGTATCCCGGATTCTG

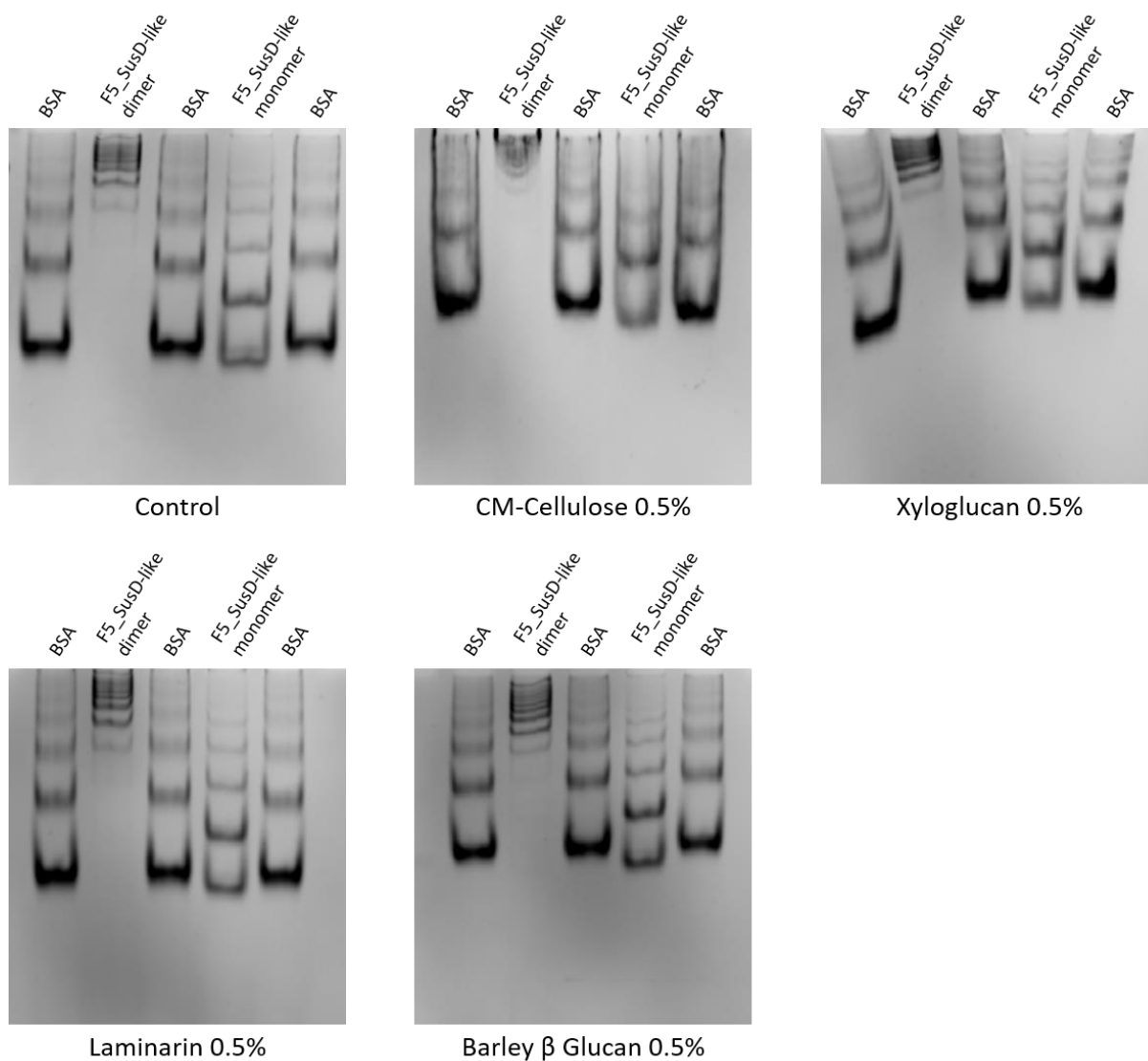


Figure S1. Affinity gel electrophoresis of F5_SusD-like on a gel containing no polysaccharide (negative control) and on gels with 0.5% (w/v) of different β -glucans.

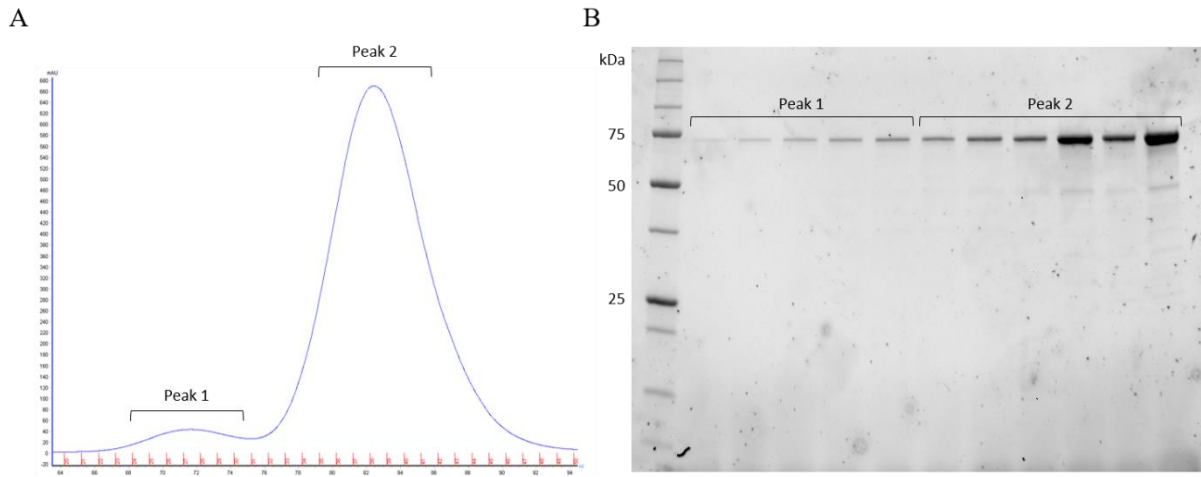


Figure S2. Analysis of F5_SusD-like oligomerization.

(A) Elution profile of the protein when injected on a Superdex 200 16/600 in 20 mM Tris-HCl, 150 mM NaCl, pH 7.5 buffer. (B) SDS-PAGE analysis of the fractions collected in A.

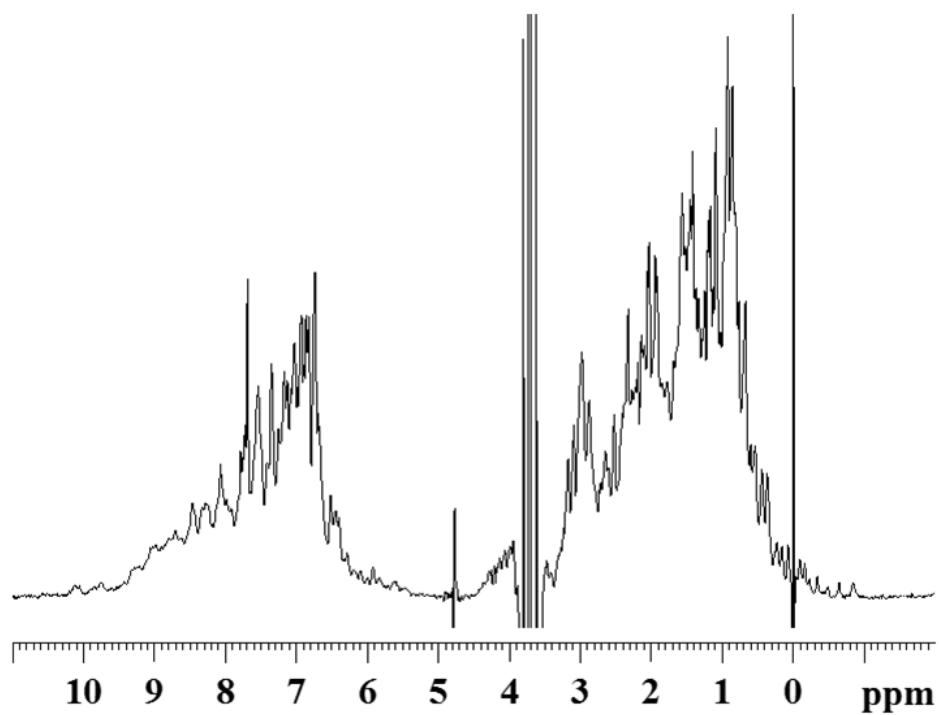


Figure S3. 1D proton spectrum at 298K of 500 μ l of the F5_SusD-like protein at 3.8 mg/ml (54 μ M) in Tris buffer.

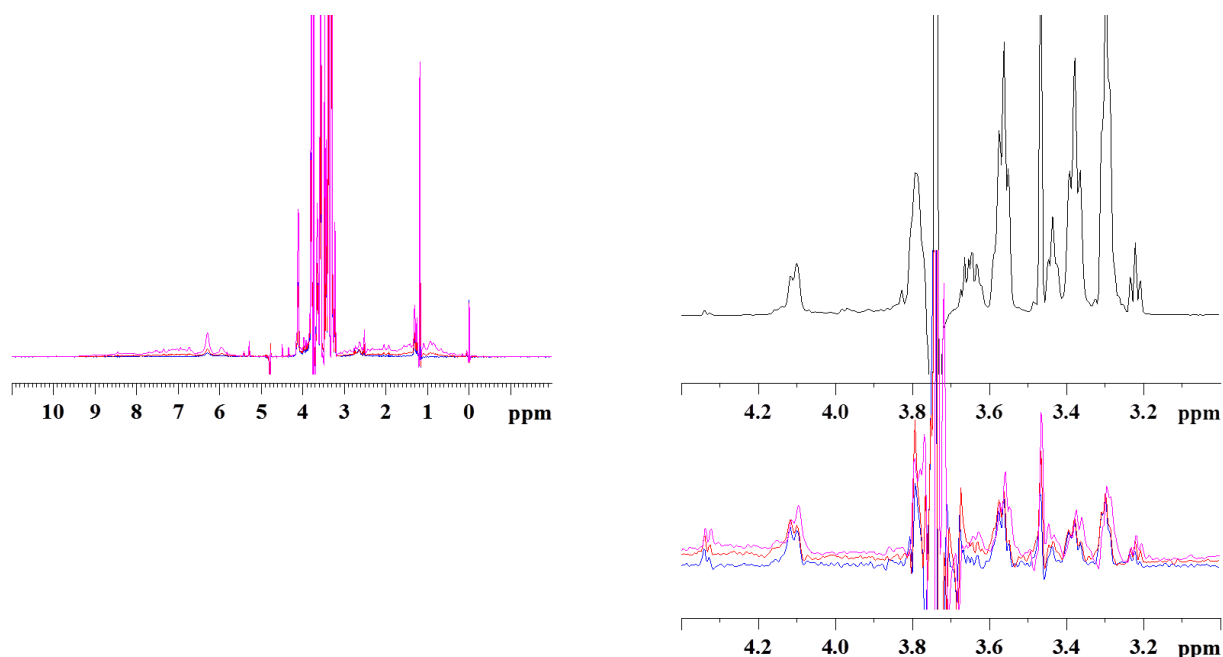


Figure S4. (left) 1D proton spectrum at 298K of 500 μ l of xylan at 500 μ M with F5_SusD-like at 0 (blue), 15 (red) or 27 μ M (magenta) in Tris buffer. (right) STD spectra of the different samples with saturation at 0 ppm for the isolated xylan (blue), in presence of 15 (red) or 27 μ M (magenta) of F5_SusD-like. The reference xylan spectrum is shown in black.

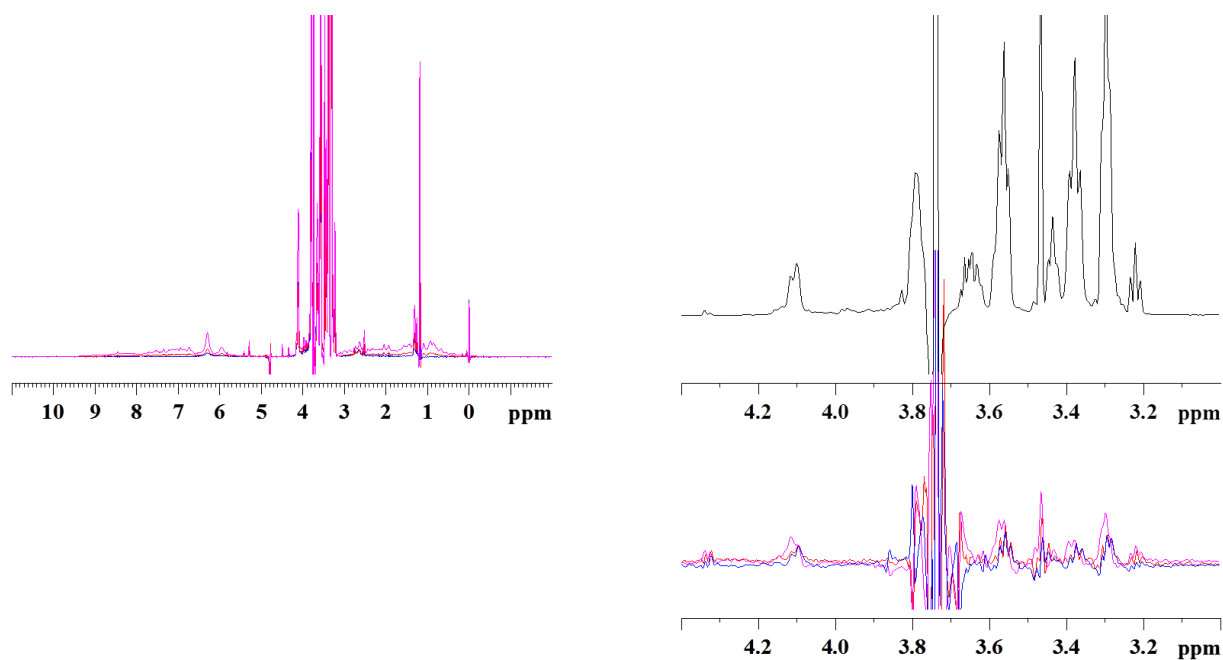


Figure S5. (left) 1D proton spectrum of 500 μ l of xylan at 500 μ M with F5_SusD-like at 0 (blue), 15 (red) or 27 μ M (magenta) in Tris buffer T_e = 298K (right). STD spectra of the different samples with saturation at 10.1 ppm for the isolated xylan (blue), in presence of 15 (red) or 27 μ M (magenta) of F5_SusD-like. The reference xylan spectrum is shown in black.

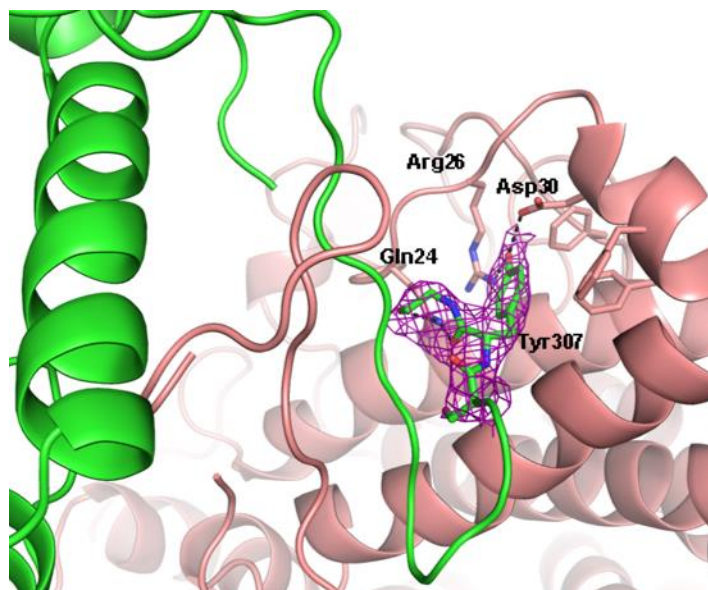
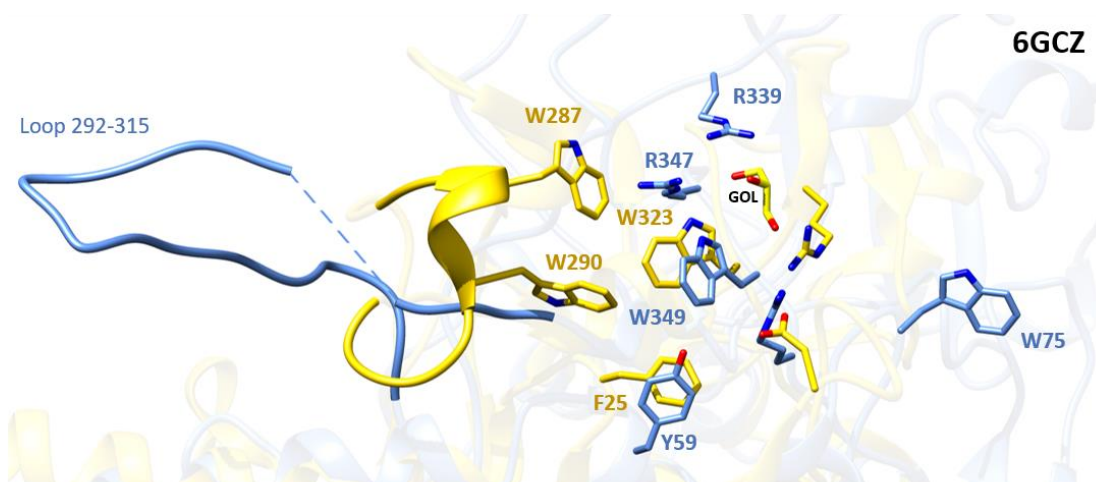


Figure S6. Close-up view of the interaction between the loop 292-315 and a symmetry related molecule.

Tyr307 is bound in a small amphipathic cavity and stabilized by hydrogen bonds with Asp30, Arg26 and Gln24. The electron density map 2Fo-Fc (contoured at 1.0σ) around the Tyr307 is shown.

A.



B.

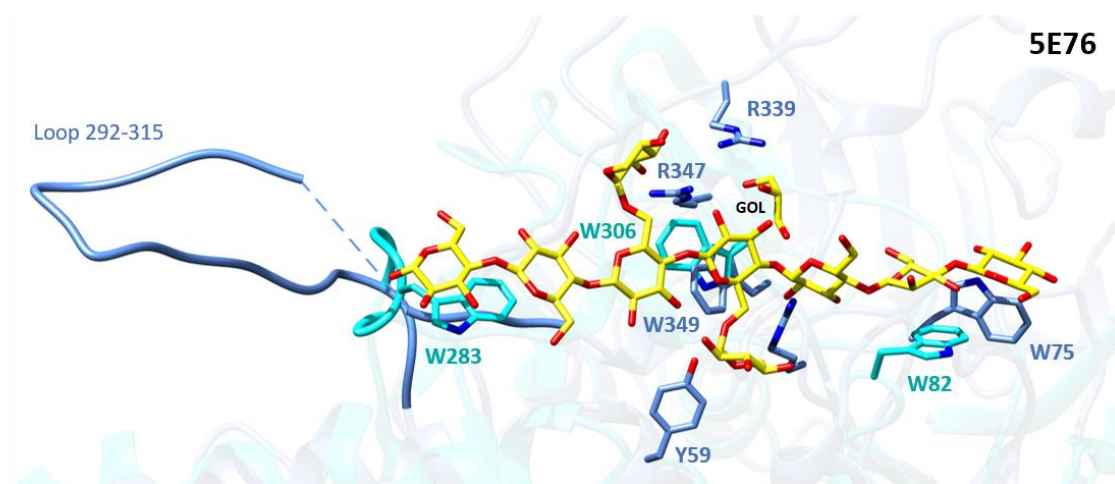


Figure S7. Comparison of the oligosaccharide binding areas of SusD-like proteins with the homologous region of F5_SusD-like.

In each panel F5_SusD-like is in blue. (A) Comparison of F5_SusD-like with the laminarin binding GM_SusD (6GCZ, yellow) from *Gramella* sp. MAR_2010_102. (B) Comparison of F5_SusD-like with Bacova_02651 (5E76, cyan) bound to xyloglucan. The stacking platforms identified in GM_SusD and Bacova_02651 are showed with the corresponding residues in F5_SusD-like which are non-effective to bind XOS. The glycerol (GOL) present in panels (A) and (B) belongs to the F5_SusD-like structure. The figures were created with Chimera.

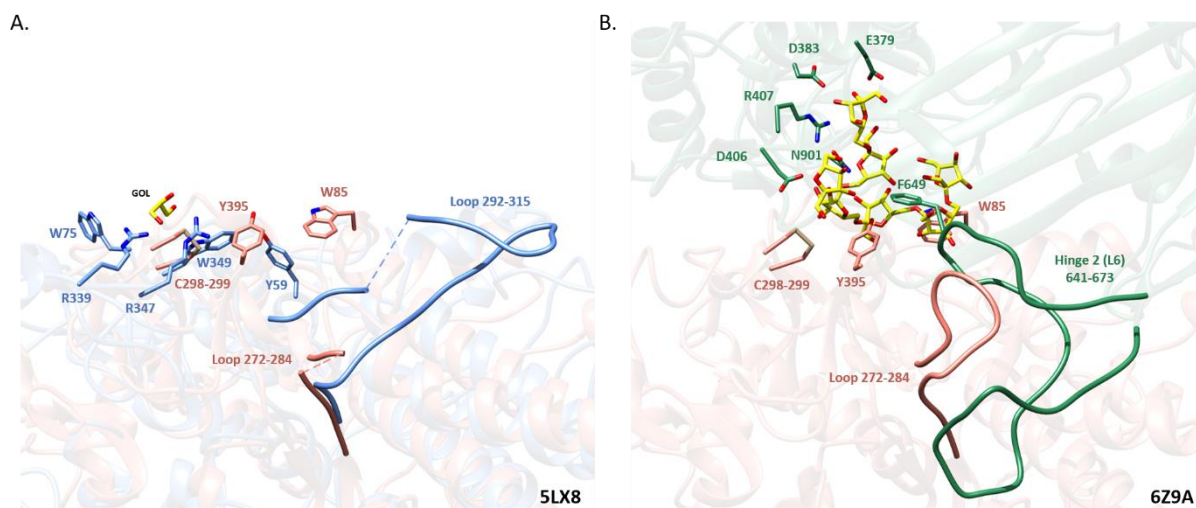


Figure S8. Structural comparison of SusD-like proteins alone and in interaction with cognate SusC-like proteins.

(A) Comparison of F5_SusD-like (blue) with BT1762 crystallized without BT1763 (PDB-ID: 6GCZ, salmon). (B) The structure of the BT1762-BT1763 complex (PDB-ID: 6Z9A, SusD-like in salmon and SusC-like in green) shows the molecular determinants for fructo-oligosaccharide binding. The figures were created with Chimera.

Second chapter

As explained in the previous chapter, the *E. coli* metagenomic clone harboring the F5_PUL is able to metabolize XOS. The XOS internalization is mediated by the two transporters of the SusC/D and MFS superfamilies, followed by their subsequent intracellular hydrolysis into xylose. F5_MFS and F5_SusC are known as membrane proteins and F5_SusD is predicted to be anchored to the outer membrane of *Bacteroides* species. The specificity of each of the two transporters encoded by the F5 XOS_PUL towards XOS of various degrees of polymerization was previously characterized using selective growth of the F5 metagenomic clone on mineral medium containing XOS with DP ranging from 2 to 5 as the sole carbon sources. Thereinto, the F5_SusC/D and F5_MFS transporters are specific for the internalization of linear xylo-oligosaccharides with a polymerization degree of up to 3 and 4, respectively. Nevertheless, the ability of the F5 transporters to internalize oligosaccharides other than XOS into the cells has never been studied. Based on this, a novel question is proposed:

- Are the SusC/D and MFS transporters promiscuous against oligosaccharides?

Previous results indicated that both F5_SusC/D and F5_MFS transporters were capable of recognizing and uptaking XOS of at least DP 3. Xylotriose is a trisaccharide comprising three xylosyl residues linked by β -1,4-glycosidic bonds. Mannotriose is composed of three mannosyl residues also linked by β -1,4-glycosidic bonds. Thus, it exhibits structural similarity to xylotriose. This subtle distinction was precisely employed to investigate the promiscuity of F5_SusC/D and F5_MFS transporters towards oligosaccharides of various structures. To address this question of substrate promiscuity, a novel strategy was proposed, in which a β -mannosidase was introduced into the periplasm of bacterial strains harboring the F5_MFS and F5_SusC/D transporters. The cultivation of these genetically modified cells on mannotriose, in comparison to a control set fed with xylotriose, provided a platform to examine the capability of these cells to recognize and internalize mannotriose, that could be hydrolyzed to mannose by the periplasmic β -mannosidase.

Assessment of the F5_SusC/D and F5_MFS transporters specificity

Introduction

The majority of the human microbiota is located in the intestinal tract and plays a fundamental role in the well-being of the host. The healthy human gut microbiota consists of approximately 50.10^{12} microorganisms per person, with an estimated biomass of 1 kg. It is mainly composed of anaerobic bacteria but also contains archaea, fungi, protozoa, bacteriophages, and viruses (Chibani et al. 2022; Camarillo-Guerrero et al. 2021; Gregory et al. 2020; Raimondi et al. 2019; Huttenhower et al. 2012). Complex dietary fibers evade the limited repertoire of human digestive enzymes and hence constitute the primary carbon source for gut bacteria (Hamaker and Tuncil 2014; Koropatkin et al. 2012). To address the structural diversity of glycosides, gut bacteria have evolved various and complex protein machineries that act in synergy to break down the substrates into monosaccharides, which are assimilated by the cells.

The specificity of oligosaccharide transporters, integral membrane proteins responsible for the translocation of oligosaccharides across bacterial membranes, plays a fundamental role in determining the metabolic capabilities and ecological niches of bacterial species. By selectively importing particular oligosaccharides from their environment, bacteria can adapt to diverse ecological niches and exploit specific carbon sources available in their surroundings. Furthermore, the ability of pathogenic and commensal bacteria to recognize host-derived oligosaccharides through specialized transporters can influence host-microbe interactions and contribute to microbial virulence. Recent studies have made significant strides in elucidating the molecular mechanisms underlying the specificity of oligosaccharide transporters in bacteria.

For instance, the promiscuity of transporters from different families has been investigated and reported. Shulami *et al.* discovered that *Geobacillus stearothermophilus* cross-utilizes β -galactosides and cellobiose (Shulami et al. 2020). Indeed, the cellobiose-PTS system, one of three different transport systems possessed by *G. stearothermophilus* T-1, can also recognize and utilize lactose and galactosyl-glycerol. Regarding the ABC transporters, several characterizations were performed on solute-binding proteins that determines the specificity of the ABC transporters. In the probiotic *Bifidobacterium animalis* subsp. *lactis* it has been shown that Bal6GBP is able to recognize both β -(1,6)- and β -(1,4)-galactobiose, but has a 1630-fold higher selectivity for the former. This uptake preference is a crucial factor that governs the onset

and extent of growth (Theilmann et al. 2019). In the same organism, the solute-binding protein *BIG16BP* was shown to bind α -(1,6)-linked glucosides and galactosides of varying sizes, displaying dual specificity for galactosides and glucosides that share a common α -(1,6)-diglycoside motif (Ejby et al. 2016). In the pathogenic *Streptobacillus moniliformis*, an ABC transporter involved in the import of both sulfated and non-sulfated glycosaminoglycan (GAG) disaccharides has been identified (Oiki, Mikami, et al. 2017; Oiki, Kamochi, et al. 2017). The solute-binding protein *Smon0123* exhibits affinity for disaccharides generated from hyaluronan and chondroitin sulfate. Particularly, the authors showed that *Smon0123* shows less affinity for the unsaturated chondroitin disaccharide sulfated at the C-4 and C-6 positions of *N*-acetyl-D-galactosamine residue than for non-sulfated and monosulfated substrates. *Streptococcus mutans* has two related ABC transporters that transport distinct disaccharides and/or oligosaccharides. The examination of carbohydrate uptake specificity by these transporters indicates that *MalXFGK* primarily transport maltodextrins, while *MsmEFGK* is mainly involved in the uptake of melibiose, raffinose, stachyose, isomaltose, and isomaltotriose (Webb et al. 2008). Lactose permease (*LacS*) is a member of the Glycoside-Pentoside-Hexuronide (GPH):Cation Symporter Family. It has a broad specificity with regard to the size and composition of the various recognized β -galactosides in probiotic lactobacilli (Abou Hachem et al. 2013; Andersen et al. 2011).

Furthermore, the promiscuity of *SusC/D* transporters was frequently mentioned and studied especially focusing on surface glycan-binding *SusD*-like proteins (SGBP) (Tamura et al. 2021; Déjean et al. 2020; Tamura et al. 2019; Rogowski et al. 2015). Rogowski and his colleagues noted the promiscuity of the *SusC/D* transporter when characterizing the xylan utilization system. *B. ovatus* strain ATCC 8483 possesses two PULs (*PUL_XylL* and *PUL_XylS*) that are activated when the organism is grown on wheat arabinoxylan (WAX). Meanwhile, *PUL-XylS* and *PUL-XylL* encode one (*BACOVA_04392*) and two (*BACOVA_03427* and *BACOVA_03429*) *SusD*-like proteins, respectively. *BACOVA_04392* binds to relatively simple xylans such as wheat arabinoxylan (WAX) and birchwood glucuronoxylan (BGX) but does not interact with highly decorated corn xylan (CX). It then binds to xylohexaose instead of shorter xylooligosaccharides. The *SusD*-like protein derived from *PUL-XylL*, *BACOVA_03427*, exhibits a strong binding affinity for BGX and CX but not for WAX or linear xylooligosaccharides (Rogowski et al. 2015). In addition, Tamura *et al.* identified a *SusD* homolog encoded by the *Bacteroides ovatus* mixed-linkage β -glucan utilization locus (MLGUL), *BoSGBP_{MLG-A}*, which presents a high specificity for mixed-linkage

$\beta(1,3)/\beta(1,4)$ -glucans. In particular, the crystallographic complexes of SGBP-A with oligosaccharides have provided insight regarding the molecular determinants of specificity. Structural analysis revealed that specificity for the kinked MLG backbone, as opposed to linear β -glucans, is underpinned by the unique shape complementarity of binding platforms (Tamura et al. 2019). Subsequently, a suite of SGBPs-A SusD-like proteins from three syntenic $\beta(1,3)$ -glucan utilization loci (1,3GULs) from *Bacteroides thetaiotaomicron* (*Bt*), *Bacteroides uniformis* (*Bu*), and *B. fluxus* (*Bf*) were shown to exhibit varying specificities for distinct β -glucans (Déjean et al. 2020; Tamura et al. 2021). The *Bt*SGBP-A could recognize and bind the branched $\beta(1,3)$ -glucans from *Laminaria digitata* (*LdLam*), which contains single $\beta(1,6)$ -linked glucosyl branches, and yeast β -glucan (*yBG*), which contains longer $\beta(1,6)$ -linked glucosyl branches extended by $\beta(1,3)$ -linked glucose units. A sequence alignment combined with structural superposition revealed that many key binding site residues are conserved in *Bf*SGBP-A compared to *Bt*SGBP-A. Therefore, it can be inferred that *Bf*SGBP-A binds $\beta(1,3)$ -glucan (*LdLam* and *yBG*) similarly to *Bt*SGBP-A. In particular, although *Bacteroides uniformis* (*Bu*) possessed a PUL to access *LdLam*, *yBG*, and barley mixed-linkage $\beta(1,3)/\beta(1,4)$ -glucan (*bMLG*), SusD-like protein *Bu*SGBP-A did not bind to the cognate substrates. The structural analysis of this protein revealed the presence of disordered loops in the vicinity of the binding site, likely hampering the substrate binding. On the other hand, *Bu*SGBP-B could bind to these substrates, thereby offsetting the lack of binding from *Bu*SGBP-A.

Nevertheless, these studies focused on the soluble proteins that are part of a more complex system. As mentioned in the first review, a binding-site has been identified at the bottom of the SusC-like cavity, so that it also plays a role in substrate recognition. A way to test the promiscuity of an overall transporter can be the utilization of previously characterized metagenomic clones.

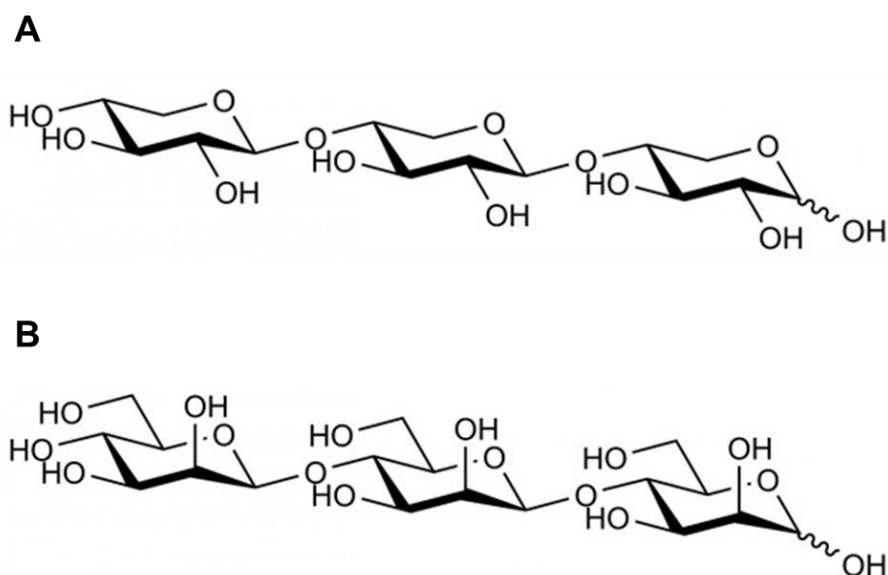


Figure 1: Chemical structure of β -1,4-xylotriose (A) and β -1,4-mannotriose (B).

As explained in the previous chapter, Tauzin *et al.* identified the F5 XOS_PUL from an uncultured gut *Bacteroides* strain, which encodes three glycoside-hydrolases (two GH43 and one GH10), a SusC/D tandem homolog and a MFS transporter (Tauzin *et al.*, 2016). The *E. coli* metagenomic clone harboring the F5 XOS_PUL is able to metabolize XOS due their internalization mediated by the F5_SusC/D and F5_MFS transporters, followed by their subsequent intracellular hydrolysis into xylose. The specificity of each of the two transporters encoded by the F5 XOS_PUL towards XOS of various degrees of polymerization was previously characterized using selective growth of the F5 metagenomic clone on mineral medium containing XOS with DP ranging from 2 to 5 as the sole carbon sources. Thereinto, the F5_SusC/D and F5_MFS transporters are specific for the internalization of linear XOS with a polymerization degree of up to 3 and 4, respectively. Nevertheless, the ability of the F5 transporters to internalize oligosaccharides other than XOS into the cells has never been studied. Specifically, the present research aims to investigate whether F5_MFS and F5_SusC/D transporters could promiscuously recognize and transport oligosaccharides with close but different structures. To do so, we decided to work with mannotriose for two reasons. First, mannosyl residues are also linked by β -1,4-glycosidic bonds but they are hexaoses while xylosyl residues are pentoses (Fig. 1). This subtle distinction is precisely employed to investigate the promiscuity of F5_SusC/D and F5_MFS transporters. Second, an *exo*- β -mannosidase from *Bacteroides thetaiotaomicron* VPI-5482 (BT_0458) belonging to the GH2 CAZy family had previously been structurally and functionally characterized and shown to

efficiently degrade mannotriose (Tailford et al. 2007). We propose a novel strategy in which this *exo*- β -mannosidase will be introduced into the periplasm of the bacterial strains harboring the F5_MFS and the F5_SusC/D transporters. If the transporters are able to uptake mannotriose, it will be subsequently degraded to monosaccharides by the mannosidase, supporting thereby the growth of the *E. coli* strains which can be easily monitored (Fig. 2).

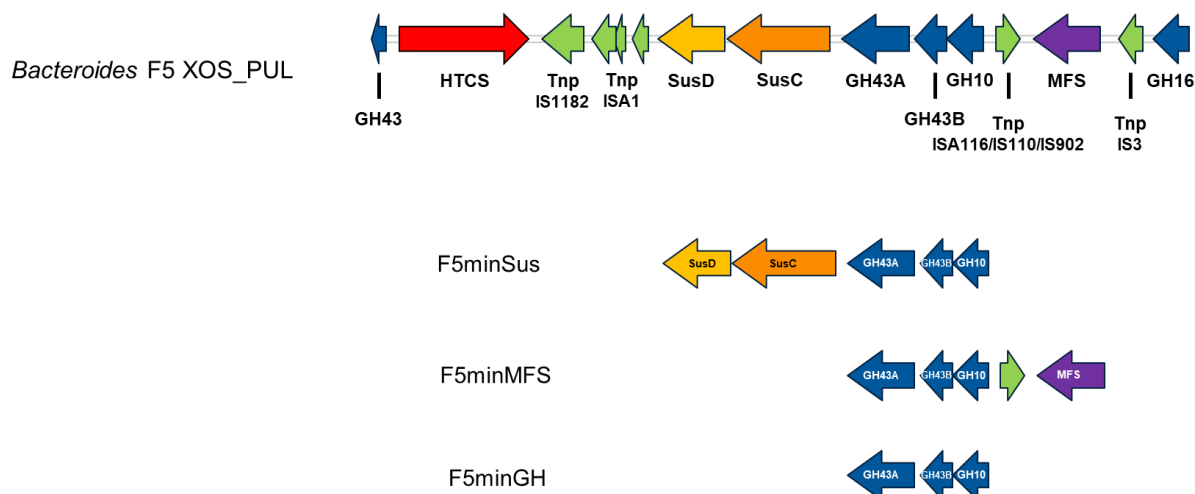


Figure 2: Representation of the F5 XOS_PUL and three variants.

Genes encoding known and predicted functionalities are color-coded: glycoside hydrolase (GH) in blue; SusD-like in yellow; SusC-like in orange; MFS in purple.

In order to characterize the specificity of the F5 transporters towards the structure of oligosaccharides, we used, as chassis strains, three variants previously constructed by Tauzin *et al.* using the In-Fusion HD Plus cloning kit (Clontech) (Tauzin et al. 2016). The first variant, named F5minSus, harbors the three GHs (one GH10 and two GH43) and the F5_SusC/D. The second variant, named F5minMFS, harbors the three GHs and the F5_MFS transporter. The third variant, named F5minGH, only harbors the three GHs without any transporter (Fig. 3). In the present study, we complemented the F5 xylo-oligosaccharide utilization system by adding to these chassis strains a plasmid harboring the BT_0458 *exo*- β -mannosidase. Normally, in *E. coli* the lipoproteins are anchored either to the inner or to the outer membrane, and oriented towards the periplasmic space (Tokuda and Matsuyama 2004). Although BT_0458 is likely a lipoprotein anchored at the outer-membrane in the native *B. thetaiotaomicron* VPI-5482 strain (according to BT_0458 signal peptide analysis by SignalP 6.0 (Fig. 4)), its precise subcellular

localization when expressed in *E. coli* cannot be predicted with certainty. Therefore, we did two constructs of BT_0458 where native and heterologous signal peptides were tested. Indeed, to increase the chance to have BT_0458 localized in the periplasm, we also replaced the native BT_0458 signal peptide with the peptide signal coming from the *E. coli* MglB protein (Uniport P0AEE5). MglB is a D-galactose/methyl-galactoside binding periplasmic protein that is part of an ABC transporter (Harayama et al. 1983).

In this chapter, we present the experimental setup as well as the preliminary data we obtained.

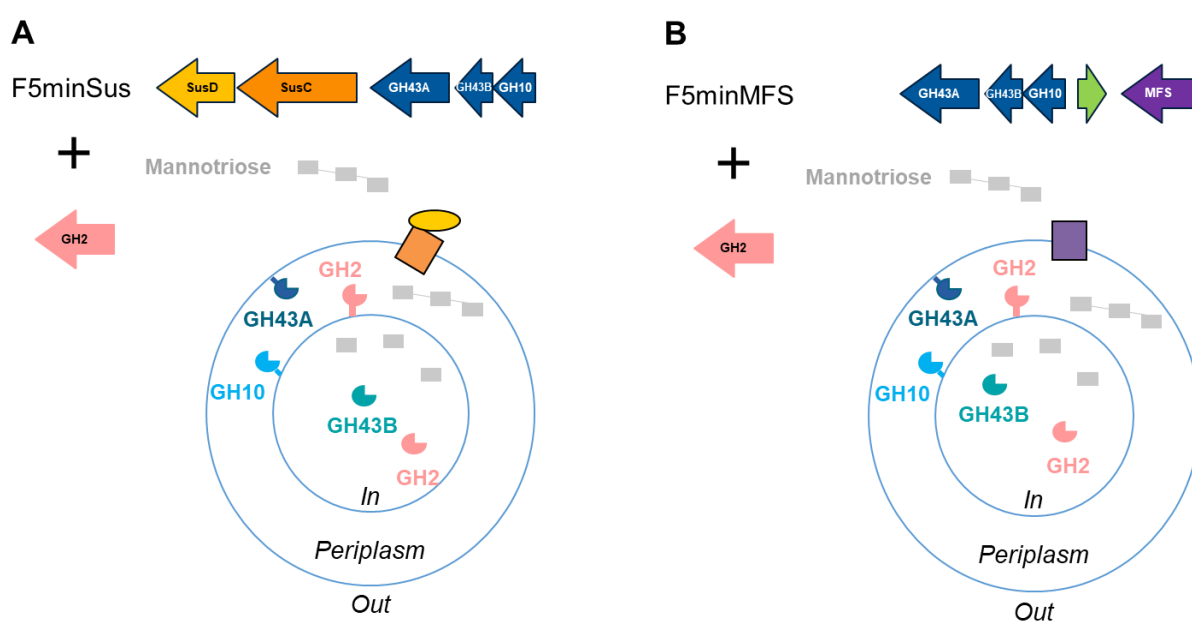
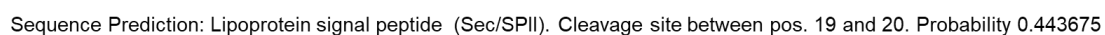


Figure 3: Representation of the internalization and hydrolysis processes of mannotriose in the metagenomic clones issued from truncation of the F5 locus.

(A) F5minSus variant with BT_0458 *exo*- β -mannosidase (GH2); (B) F5minMFS variant with BT_0458 *exo*- β -mannosidase (GH2).



227

Results

Introduction of the recombinant GH2 into the variants

The genes coding for the recombinant BT_0458 *exo*- β -mannosidase, along with the *Bt* and *E. coli* signal peptides, were synthesized and cloned into the pET-28a(+) vector using the XbaI and XhoI restriction sites. These plasmids are named p*Bt*GH2 and p*Bt*GH2_periplasm (p*Bt*GH2P), respectively. Six strains were obtained by transforming the plasmids into different F5 variants, namely F5minSus-*Bt*GH2 (Sus-GH2), F5minSus-*Bt*GH2_periplasm (Sus-GH2P), F5minMFS-*Bt*GH2 (MFS-GH2), F5minMFS-*Bt*GH2_periplasm (MFS-GH2P), F5minGH-*Bt*GH2 (F5minGH-GH2), and F5minGH-*Bt*GH2_periplasm (F5minGH-GH2P) (Fig. 5). The screening for double resistance to the kanamycin (pET-28a(+)) and chloramphenicol (pEPIFOS-5 fosmid) indicated that the recombinant plasmids p*Bt*GH2 and p*Bt*GH2P were properly introduced in the host strains. Despite the strains demonstrated ability to grow on both antibiotics, a further verification of the presence of the pET28a(+) plasmid was conducted through polymerase chain reaction (PCR) on the colonies. The colony PCR using the usual T7 primers identified the DNA fragments of *Bt*GH2 and *Bt*GH2P (Fig. 6), which matched the theoretical value. This further confirmed the presence of the recombinant plasmids.

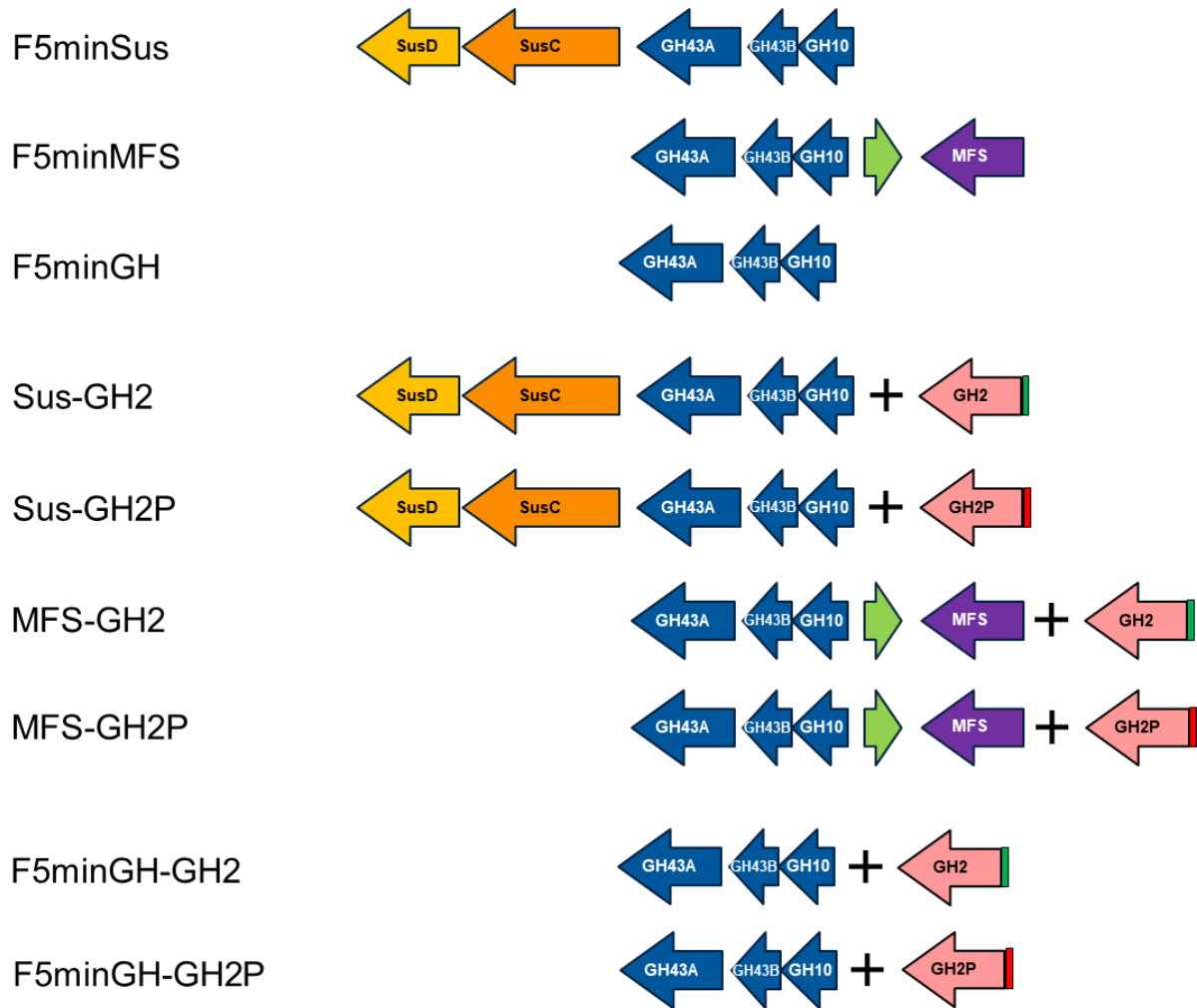


Figure 5: Gene organization of the variants and their combinations with the two different recombinant GH2 constructs.

Genes encoding known and predicted functionalities are color-coded: glycoside hydrolase (GH) from F5 XOS_PUL in blue; SusD-like in yellow; SusC-like in orange; MFS in purple. Recombinant GH2 constructs in pink, one of them contains its native signal peptide (green), the other one contains a signal peptide (red) from an *E. coli* periplasmic protein (UniProt P0AEE5).



Figure 6: Agarose gel electrophoresis of colony-PCR screening with T7 universal primers.
L1 - L8: PCR products of *BtGH2* (2774 bp); L9 - 16: PCR products of *BtGH2P* (2768 bp).

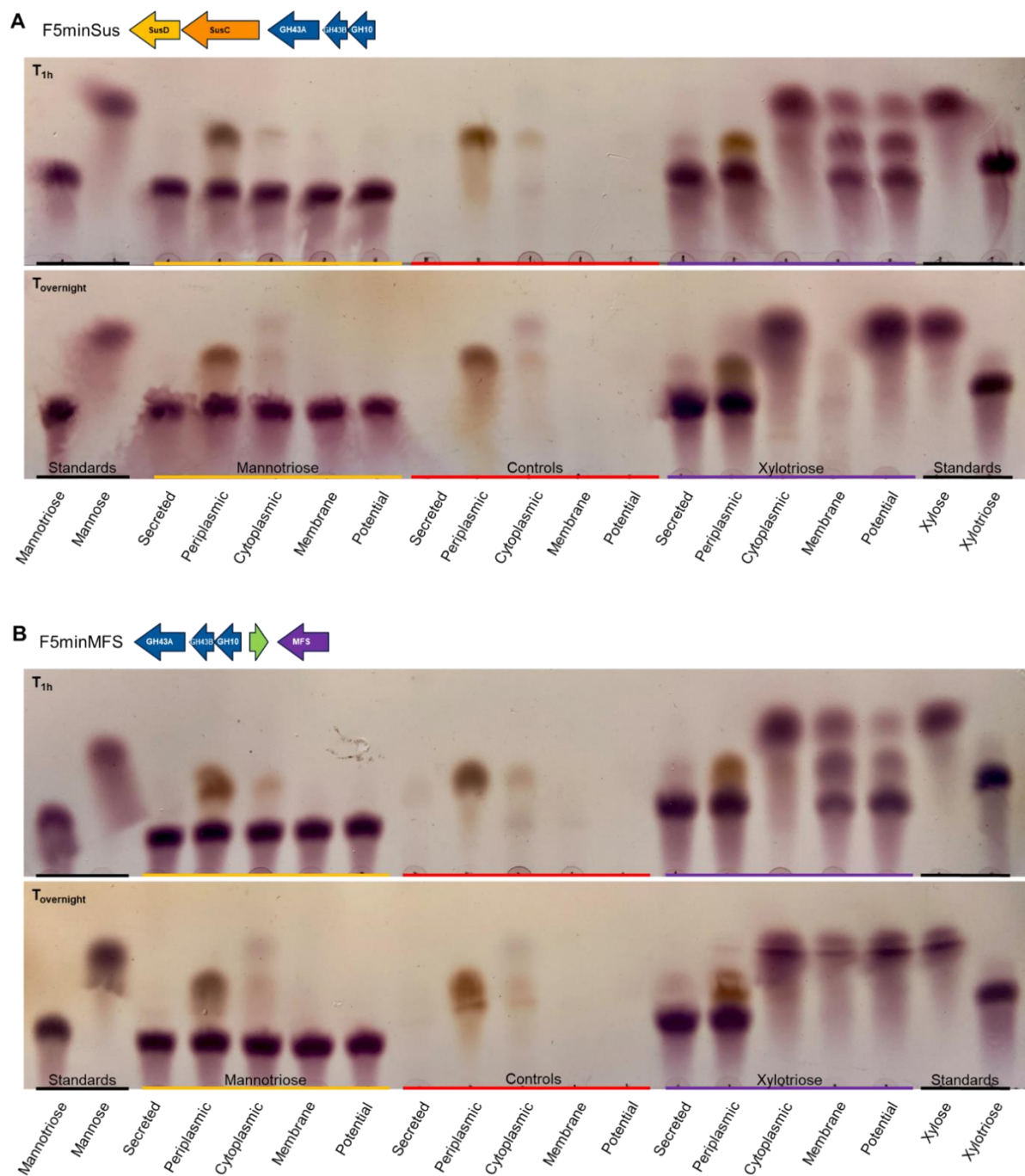
Cellular localization of the *exo*- β -mannosidase GH2

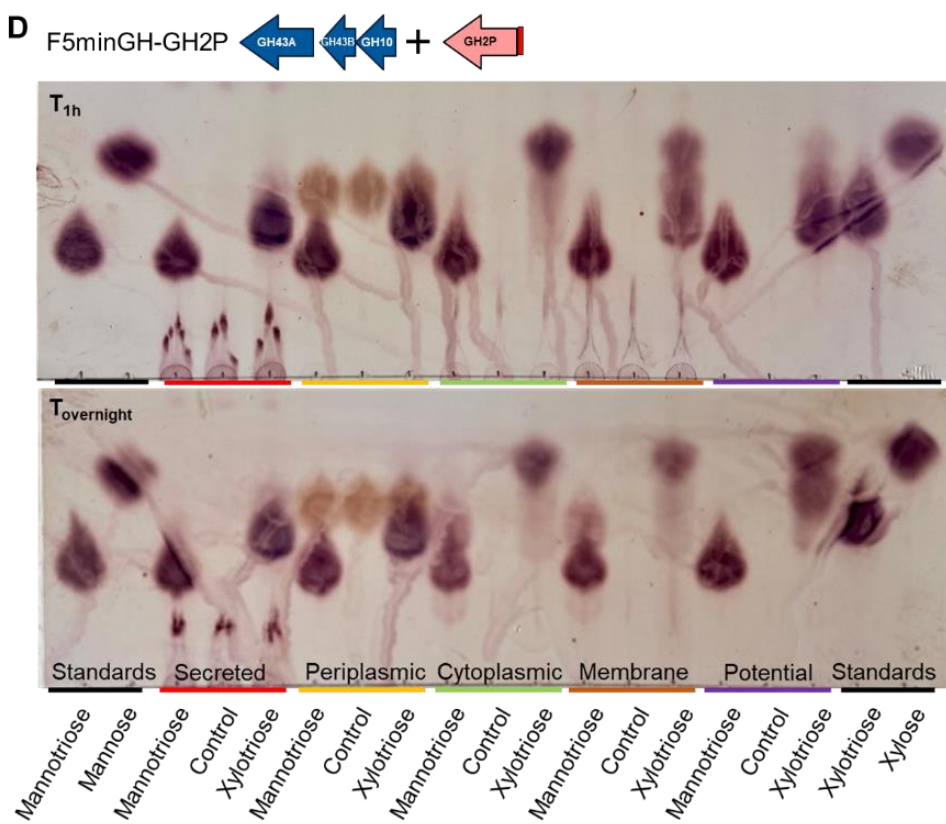
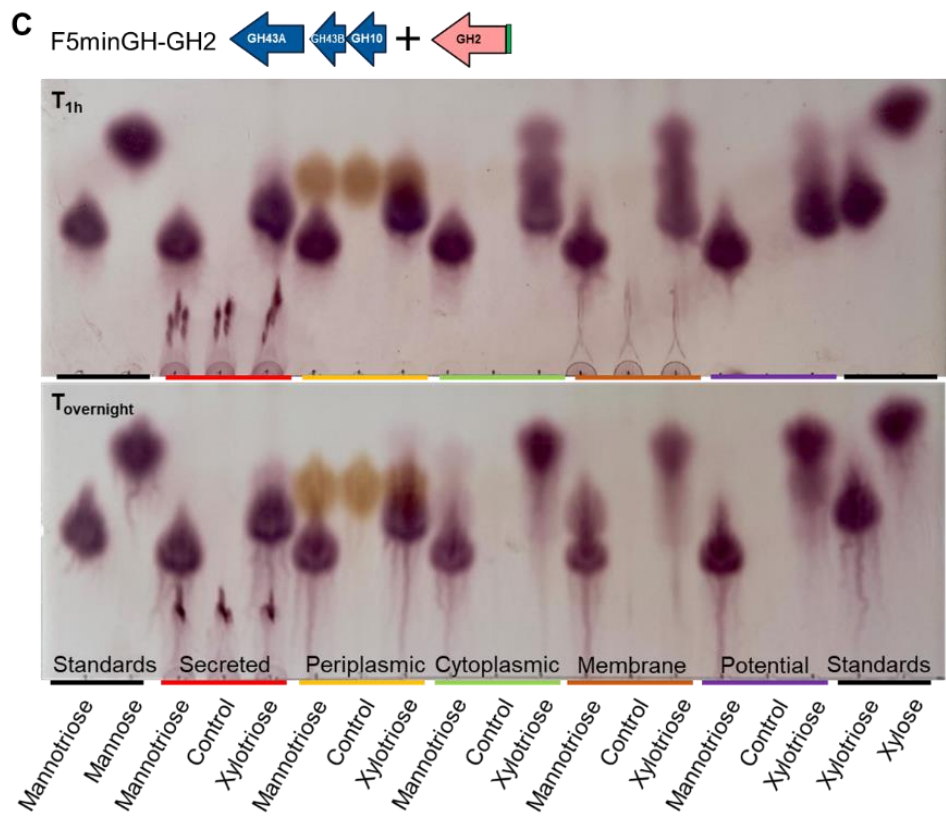
The initial assay was to investigate the cellular localization of the *exo*- β -mannosidase GH2 within the different constructs. To do so, I analyzed the hydrolytic activity of the different cellular fractions by thin layer chromatography (TLC). As a preliminary step, the initial F5minSus and F5minMFS variants without any recombinant plasmids were tested, in order to confirm that they do not display any mannosidase activity. It was also a way to double check that I could reproduce the results from Tauzin *et al.* regarding the cellular localization of the GH10 and GH43 enzymes (Tauzin et al. 2016). As shown in Fig. 7A and Fig. 7B, xylotriose was completely degraded to xylose in the cytoplasmic and membrane fractions, as well as in the fraction containing the potentially trapped soluble and/or weakly membrane-associated proteins. In contrast, mannotriose was not degraded in any of the fractions. Therefore, the results show that the F5minSus and F5minMFS variants do not have any mannosidase activity. Furthermore, the cellular localization of the xylanases is consistent with the findings of Tauzin *et al.*. Additionally, it was observed that xylose, which could be observed in the F5minSus membrane lane at T_{1h}, cannot be detected at T_{overnight} (Fig. 7A). A possible explanation is that residual living cells are present in the membrane fraction and were able to metabolize xylose. Another explanation could be that xylose is trapped by the transporter during the reaction. Indeed, in the reaction tube containing xylotriose and debris of membranes with the F5_SusC/D transporter and the GHs, xylotriose would be degraded by the enzymes and the remaining xylose would be bound to the transporter. Since only the supernatant of the reaction tube was analyzed by TLC, it could explain why we cannot detect xylose. To test this hypothesis, we conducted the same assay for F5minGH-GH2 and F5minGH2P variants that do not contain any F5_transporter. Using these variants, xylose could be detected (Fig. 7C and 7D), suggesting that the hypothesis of the xylose trapped by the transporter could be valid. Nevertheless, replicates and complementary experiments should be performed to validate this hypothesis (see discussion).

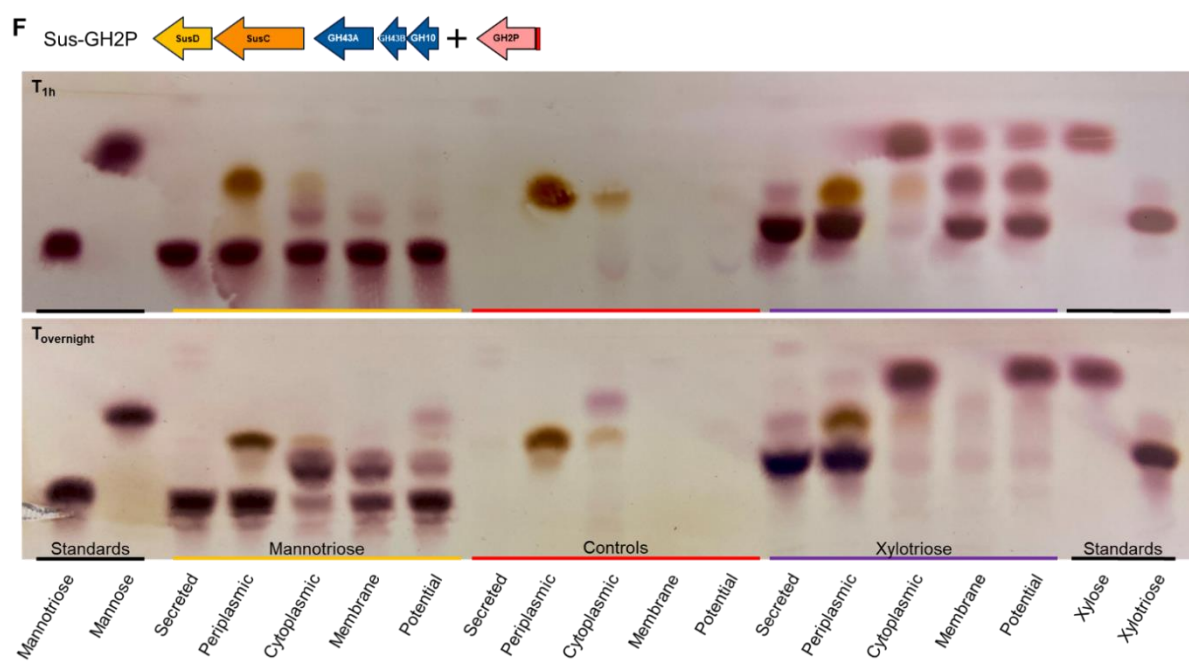
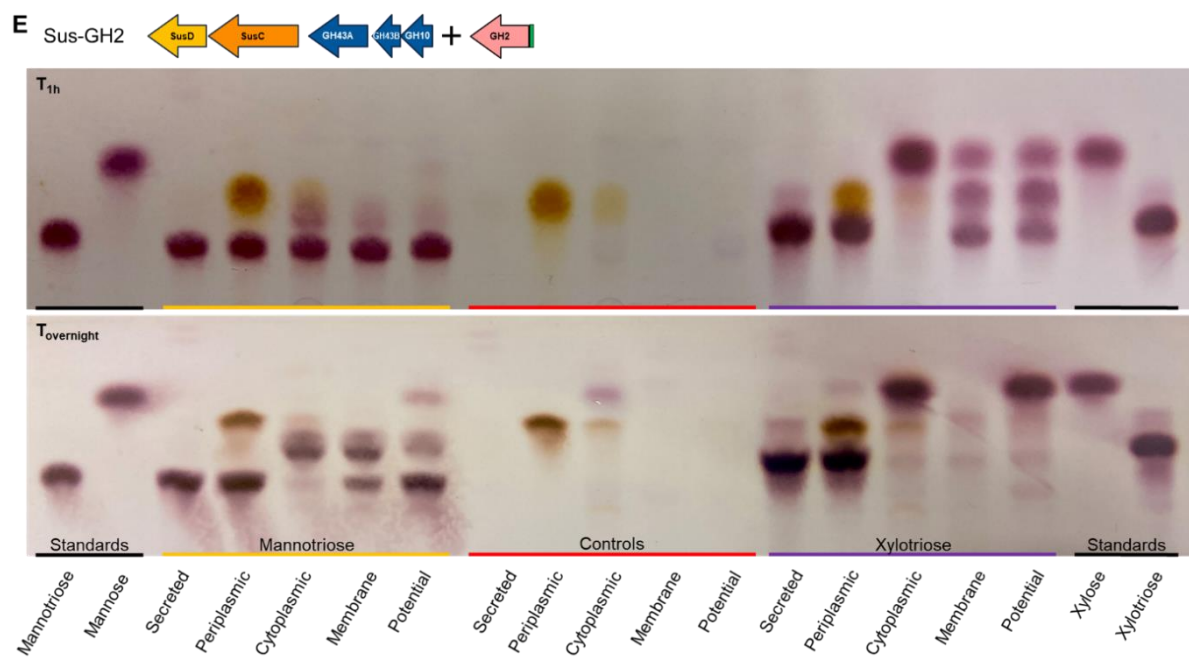
As for mannotriose, degradation profiles could be observed with the variants containing a transporter (F5_SusC/D or F5_MFS), the GHs and the *exo*- β -mannosidase, whatever its signal peptide was (Fig. 7E-H). More precisely, degradation of mannotriose can be observed with the cytoplasmic and the membrane fractions, as well as with the “potential” fraction, corresponding to the potential trapped soluble proteins and/or weakly membrane-associated proteins. Contrary to the cytoplasmic and the membrane fractions where the degradation profile only shows mannobiose, the “potential” fraction also displays some mannose at T_{overnight} (except for MFS-

GH2P). Surprisingly, no mannotriose degradation could be observed with the periplasmic fractions, not even the ones coming from the variants containing the GH2 harboring the signal peptide from the periplasmic protein from *E. coli* (Fig. 7F and 7H). Finally, no degradation is observed with the secreted fractions, showing that the GH2 is not secreted in the extracellular environment.

These preliminary results validate that the *exo*- β -mannosidase GH2 is expressed and active. It further suggests that the protein is found in the cytoplasm and associated to the membrane (probably facing the periplasm, to be investigated), whatever the signal peptide used. Also, the protein is not secreted. All together, these data allow us to conduct growth assays to test the promiscuity of the F5 transporters.







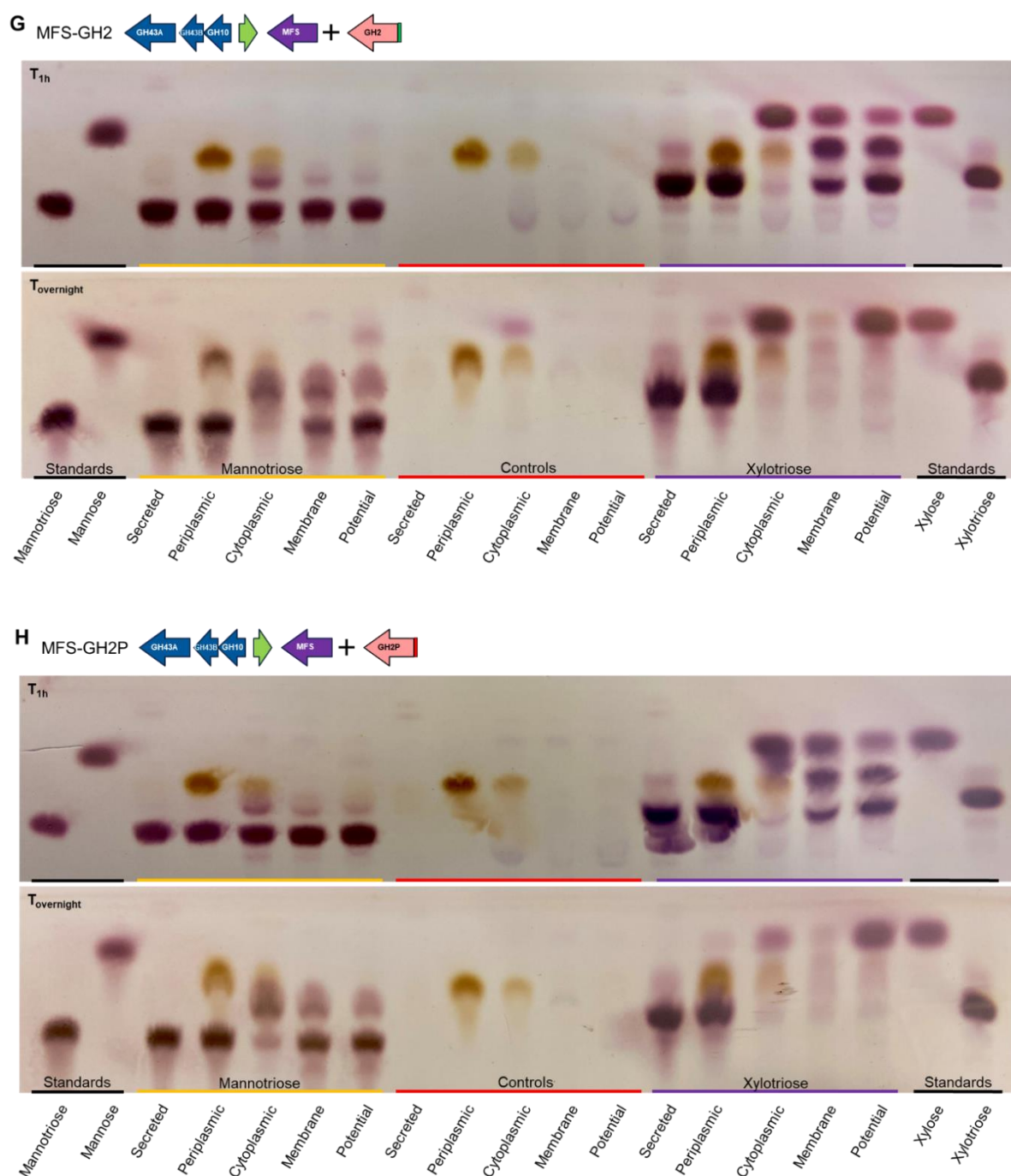


Figure 7: TLC analysis of the degradation profiles of 0.2 % (w/v) mannotriose and xylotriose.

- (A) Degradation profile of mannotriose and xylotriose produced by the different cellular fractions of the F5minSus variant (Control);
- (B) Same as panel A but with the fractions from the F5minMFS variant (Control);
- (C) Same as panel A but with the fractions from the F5minGH-GH2 variant (Control);
- (D) Same as panel A but with the fractions from the F5minGH-GH2P variant (Control).
- (E) Same as panel A but with the fractions from the Sus-GH2 variant;

- (F) Same as panel A but with the fractions from the Sus-GH2P variant;
- (G) Same as panel A but with the fractions from the MFS-GH2P variant;
- (H) Same as panel A but with the fractions from the MFS-GH2P variant;

Growth assays

To test the promiscuity of the F5_transporters, we performed several growth assays using the different F5 variants in M9 minimum medium complemented with either xylotriose or mannatriose and the appropriate antibiotics (Fig. 8). The control variants, F5minSus and F5minMFS, both without the GH2, were able to display a significant growth in presence of xylotriose, with a final OD_{600nm} being >0.5 and >1 respectively. As expected, these variants were not able to grow in presence of mannatriose. The same experiments were carried out with the F5 variants harboring either the GH2 or the GH2P. Similar growth curves to the controls were obtained. Indeed, the variants were able to grow in the xylotriose condition but not in the mannatriose condition, whatever they are harboring the GH2 or the GH2P. These results suggest that neither the SusC/D transporter nor the MFS transporter encoded by the F5_XOS_PUL could transport mannatriose, regardless of its structural similarity to xylotriose. Based on this, we could conclude that the F5 transporters are specific to XOS. However, supplemental controls are required, especially regarding the effective expression of the GH2 in M9 condition.

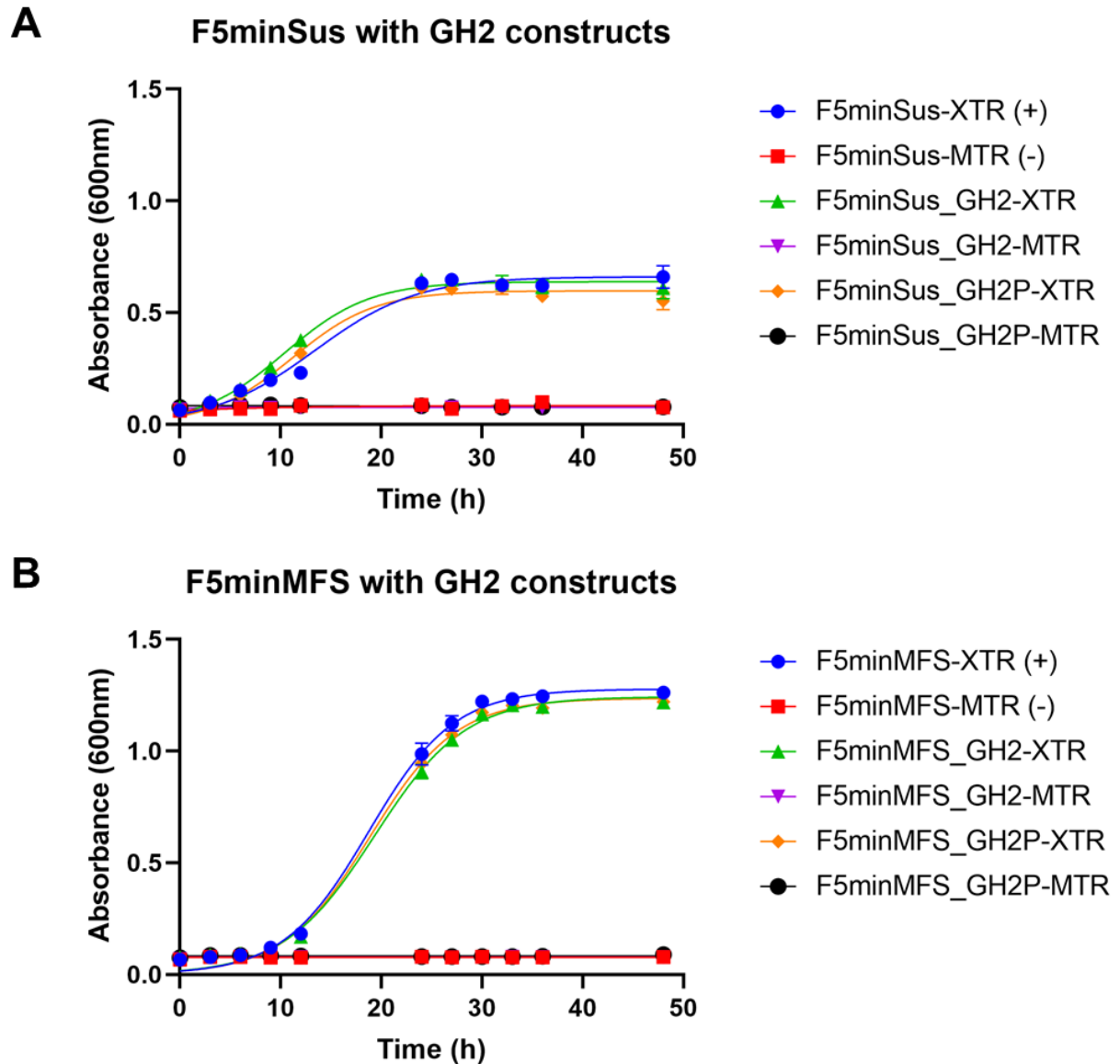


Figure 8: Growth curves of the different F5 variants in M9 minimal medium containing either xylotriose (XTR) or mannotriose (MTR) at 0.5 % (w/v) as sole carbon source. The media also contain antibiotics, 12.5 mg/l of chloramphenicol and 50 mg/l of kanamycin (for the controls F5minSus and F5minMFS no kanamycin was added).

(A) F5minSus with the GH2 recombinant constructs; (B) F5minMFS with the GH2 recombinant constructs. The data represent the average of at least triplicates.

Discussion

The pET expression system in the F5 variants

The modified metagenomic clones (the F5 variants) are *E. coli* Stellar strains containing the pEPIFOS-5 fosmid with the different F5 XOS_PUL DNA constructs. To investigate to which extent the F5_SusC/D and F5_MFS transporters are promiscuous, we decided to test their ability to uptake mannotriose. To do so, we transformed these variants with the pET-28a(+) vector containing the BT_0458 gene located directly downstream the T7 promoter. The pET system is a powerful system developed for the cloning and expression of recombinant proteins in *E. coli*. This system normally requires a λ DE3 lysogen *E. coli* strain such as BL21(DE3), which contains the T7 bacteriophage gene I, encoding the T7 RNA polymerase under the control of the *lacUV5* promoter. T7 RNA polymerase elongates chains about five times faster than does *E. coli* RNA polymerase (Chamberlin and Ring 1973; Golomb and Chamberlin 1974). The T7 RNA polymerase is expressed when the inducer IPTG is added to the growing culture. The enzyme then transcribes the target gene cloned into the pET vector (Fig. 9). Nevertheless, even in the absence of IPTG, there is some expression of T7 RNA polymerase from the *lacUV5* promoter in λ DE3 lysogens and therefore, basal expression of the target protein. As the recombinant proteins may present toxicity on the host cell and the plasmid stability, one approach to control basal expression is to use vectors (for instance pET vectors) that contain what is termed a T7 *lac* promoter (Dubendorf and Studier 1991). These plasmids contain a *lac* operator sequence just downstream of the T7 promoter. When this type of vector is used in DE3 lysogens, the *lac* repressor acts both at the *lacUV5* promoter in the host chromosome to repress transcription of the T7 RNA polymerase gene by the host polymerase and at the T7*lac* promoter in the vector to block transcription of the target gene by any T7 RNA polymerase that is made. Despite this “double repression”, basal expression can still occur (Dubendorf and Studier 1991). Additional target sequence stability can be achieved by expression in host strains containing a compatible chloramphenicol-resistant plasmid that provides a small amount of T7 lysozyme, a natural inhibitor of T7 RNA polymerase (Moffatt and Studier 1987; Studier 1991). The T7 lysozyme is provided by a clone of the T7 lysozyme gene in the BamH I site of pACYC184 which thus is referred to as pLysE.

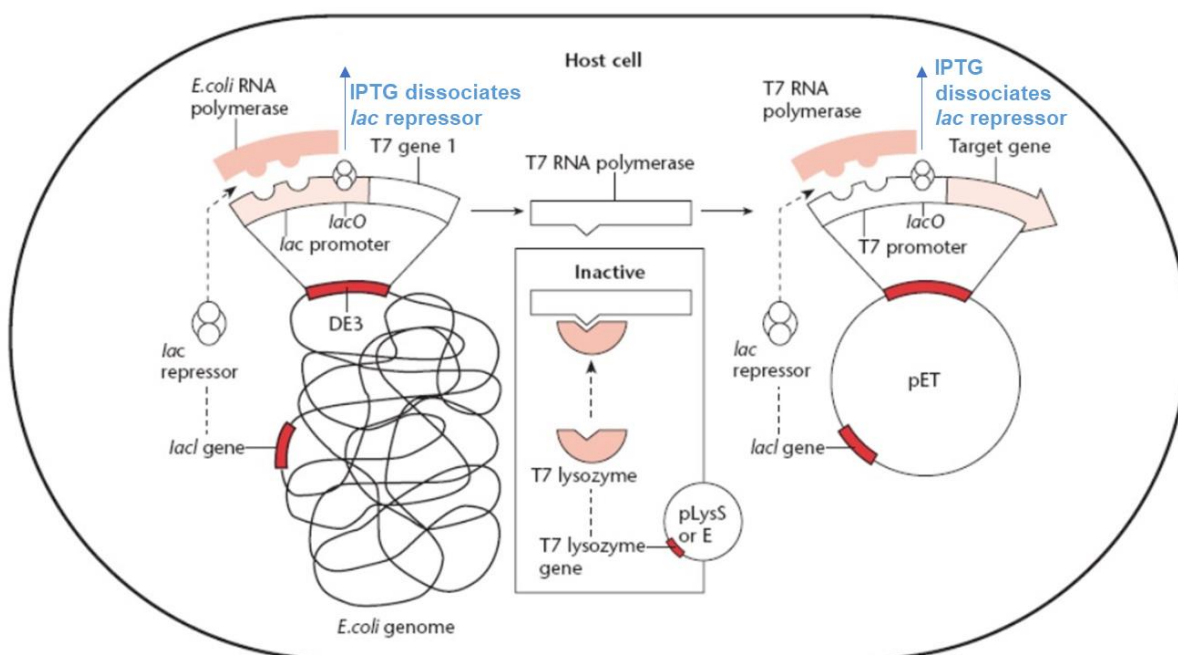


Figure 9: Strategy for regulating gene expression using the pET-expression system.

The gene encoding the T7 RNA polymerase (gene1) is inserted into the chromosome of *E. coli* and transcribed from the *lac* promoter. Therefore, it is expressed only if the inducer IPTG is added to the growing culture. The T7 RNA polymerase then transcribes the target gene cloned into the pET vector. The presence of a *lac* operator between T7 promoter and the cloned gene reduces transcription of the cloned gene in the absence of the inducer IPTG (adapted from Novagen Catalogue, Madison USA, 1995).

In our case, while the *E. coli* Stellar strain does not contain the gene encoding the T7 RNA polymerase, we were still able to detect the GH2 mannosidase activity, showing that the corresponding gene on the pET-28a(+) vector was transcribed. This can be explained by the fact that the *E. coli* RNA polymerase is able to utilize DNA from bacteriophage T7 as an effective template for RNA synthesis (Lee et al. 1979). Also, the addition of IPTG in the culture medium enables to remove the effect of the *lac* repressor. Of course, the GH2 is certainly not overexpressed, but sufficiently expressed to detect its activity, which is the key point of our experiments. It is also something to take into consideration for the complementary experiments that we have to perform to check the effective expression of the GH2 in M9 medium. Indeed, the protein is likely expressed in small amounts which might not be detectable by SDS-gel stained with Coomassie blue. It will be possible to perform a Western blot using an anti-His-

tag antibody as the BT_0458 gene was cloned so that the recombinant protein carries a C-terminal His-tag.

Subcellular localization of the recombinant *exo-β*-mannosidase GH2

To test whether the F5_transporters are promiscuous or highly specific, the key point of our strategy is the periplasmic localization of the GH2. According to the SignalP6.0 analysis, the protein is predicted to be a lipoprotein anchored at the outer-membrane in *Bt* (Fig. 4). Also, Tailford *et al.* analyzed the biophysical properties of the GH2 enzyme and revealed that it displays typical characteristics of extracellular glycoside hydrolases (Tailford et al. 2007), including complete resistance to proteolytic attack (Fontes et al. 1995), which is consistent with the highly proteolytic environment of the large bowel where numerous anaerobic bacteria secrete a range of different proteinases. Normally, in *E. coli* the lipoproteins are anchored either to the inner or to the outer membrane, and oriented towards the periplasmic space (Tokuda and Matsuyama 2004). Nevertheless, the periplasmic localization of the GH2 protein in our F5 variants is not assured. To increase our chance to have the GH2 enzyme addressed to the periplasm, we also made a gene construct where the native signal peptide was replaced by the peptide signal of an *E. coli* periplasmic protein. Our experiments showed that with both gene constructs, the recombinant proteins (GH2 and GH2P) were found in the “cytoplasm”, “membrane” and “potential” fractions. The cytoplasmic localization is surprising. It is possible that the osmotic shock did not work properly and that the periplasmic and cytoplasmic fractions are actually gathered. The GH2 activity detected in the “membrane” and “potential” fractions support that both recombinant proteins are also found associated to the membrane. Nevertheless, especially in the case of GH2 (lipoprotein signal peptide), we still have to investigate whether it faces the periplasm or the extracellular environment. Indeed, as said above, it is widely believed that in *E. coli* nearly all lipoproteins are either retained in the inner membrane or transferred to the inner leaflet of the outer membrane (Tokuda and Matsuyama 2004), but the free-form of the lipoprotein Lpp, the most numerically abundant protein in *E. coli*, has been shown to be surface-exposed (Cowles et al. 2011) and, generally speaking, the prevalence of the surface-exposed lipoproteins in Gram-negative bacteria has been underestimated (Wilson and Bernstein 2016). To test this hypothesis, we could perform some whole cells assays. The idea would be to grow the Sus-GH2, Sus-GH2P, MFS-GH2 and the MFS-GH2P F5 variants in LB medium, then to wash the cells and to incubate them in sodium phosphate buffer and mannitol and to analyze by TLC whether the degradation of mannitol occurs. This experiment will enable us to select the suitable GH2 gene construct to finalize the overall study.

In addition, there is a noteworthy phenomenon to discuss. The TLC analysis indicated that effective degradation of mannotriose occurs in the “cytoplasmic”, “membrane” and “potential” fractions, but mannose was only detected in some of the “potential” fractions (Fig. 7E-H). The absence of mannose in the “membrane” fractions could be explained by the possible presence of remaining living cells which would have metabolized it. Also, mannose could possibly be trapped by some *E. coli* mannose transporters. The “cytoplasm” fraction had been filtered at 0.22 μm , excluding the possibility of remaining living cells in this case. Also, in Tailford *et al.* (Tailford et al. 2007), they showed that the GH2 does not display any transglycosylating activity. Thus, we do not have a clear explanation for this observation so far.

Investigation of the binding ability of the F5_transporters

The TLC analysis of the xylotriase degradation by the F5 variants containing the F5_transporters and the F5_GHs showed that we can barely detect degradation products at $T_{\text{overnight}}$ in the “membrane” fractions (Fig. 7E-H). We hypothesized that in these fractions the final reaction product xylose could be bound to the F5_transporters present in the membrane debris. To further investigate this hypothesis, we will use the methodology previously developed in our team (Wang et al. 2020). Basically, membrane debris are incubated with oligosaccharides and their unbound amount is quantified over time by HPAEC-PAD, a highly sensitive detection method. First, we would use the In-Fusion cloning method to create F5 variants where the F5_GHs are removed so that we only have either the F5_SusC/D or the F5_MFS. It will prevent XOS degradation by any residual F5_GHs that would be present in the membrane debris. These new variants and the *E. coli* Stellar strain containing the empty fosmid (used as control) will be grown in LB, centrifuged and lysed. The membrane debris will be washed with sodium phosphate buffer and incubated with xylose, but we will also try DP2 XOS, DP3 XOS and DP4 XOS (DP4 XOS only for the F5_MFS). If an effective decrease of the unbound amount in the soluble fraction is measured by HPAEC-PAD, it will support our initial hypothesis.

Materials and Methods

Cloning

The BT_0458 gene (GenBank accession number AAO75565.1) coding for the mature part of the GH2 *exo*- β -mannosidase from *B. thetaiotaomicron* VPI-5482, either with its native signal peptide or with the signal peptide from the *E. coli* D-galactose/methyl-galactoside binding

periplasmic protein (UniPort P0AEE5), was optimized for *E. coli* and synthesized by the Genecust company. The gene was cloned between the XbaI/XhoI restriction sites in the pET-28a(+) expression vector (Novagen, Darmstadt, Germany) (see supplementary data 1 for details). We removed the stop codon from the gene so that the recombinant protein contains a C-terminal His-tag. The resulting plasmids were named p*Bt*GH2 and p*Bt*GH2_periplasm (p*Bt*GH2P), respectively. These plasmids were used to transform *E. coli* DH5 α by following heat shock transformation protocol. Briefly, after a 30 minutes incubation on ice, a mixture of chemically competent bacteria (50 μ l) and plasmid (1 μ l) was placed at 42°C for 45 seconds and then placed back on ice for 5 minutes. 950 μ l of SOC media was added and the transformed cells were incubated at 37°C for 1 hour with orbital shaking at 200 rpm. Transformed bacteria were diluted and plated on Luria-Bertani (LB) agar plates supplemented with 50 mg/l kanamycin. Then, these colonies were cultured in LB medium for plasmids extraction using QIAprep® Spin Miniprep Kit (QIAGEN, Hilden, Germany).

The minimal variants of clone F5 were previously constructed by our team using the In-Fusion HD Plus cloning kit (Clontech) (Tauzin et al. 2016). Briefly, the In-Fusion Enzyme provided by this kit fuses DNA fragments (e.g., PCR-generated inserts and linearized vectors) efficiently and precisely by recognizing 15-bp overlaps at their ends. These overlaps can be engineered by designing primers to amplify the desired sequences. Therefore, the different F5 truncated fragments were cloned into the pEPIFOS-5 fosmid by using this kit, and then transformed into Stellar competent cells. Subsequently, the F5minSus, F5minMFS and F5minGH variants were cultivated in 10 ml of LB media until the OD reach about 0.5, and harvested by a 10 min centrifugation at 4°C and 2500g. Then, the pelleted cells were gently resuspended in 100 μ l of cold TSS medium (Tryptone 10 g/l; Yeast extract 5 g/l; NaCl 10 g/l; PEG (M=3350g/mol) 100 g/l; DMSO 5 % (v/v); MgCl₂ 50 mM) and kept on ice. Following this, the p*Bt*GH2 and p*Bt*GH2P plasmids were transformed into competent F5 variants by heat shock transformation as mentioned above. The selected colonies were grown on LB agar medium containing 50 mg/l kanamycin and 12.5 mg/l chloramphenicol.

Colony PCR

Colony PCR was used to further validate the presence of the p*Bt*GH2 and p*Bt*GH2P plasmids in the F5minSus, F5minMFS and F5minGH metagenomic clones by T7 universal primers which apply to pET-28a(+) vector. The sequences of T7 universal primers are as follows: T7 forward primer 5'-GCTAGTTATTGCTCAGCGG-3'; T7 reverse primer 5'-

TAATACGACTCACTATAGGG-3'. The optimized colony PCR reaction mixture contained 1X PCR PrimeSTAR Max Premix, 0.2 μ M of each primer, and colony sample in 20 μ l final volume. PrimeSTAR Max Premix was purchased from TaKaRa and composed of PrimeSTAR Max DNA Polymerase and reaction buffer, as well as 2 mM Mg^{2+} and 0.4 mM of each dNTP. Colonies with approximately 1 mm in diameter were picked up with a sterilized inoculating loop and transferred to 10 μ l of sterilized H₂O and incubated in water bath for 10 min at 95°C. 1 μ l aliquot was directly added to the PCR tube as DNA templates. The thermal cycle program, run on a T100TM Thermal Cycler (BIO-RAD) consisted of one cycle of 98°C for 30 s, and 30 cycles of 98°C for 10 s, 55°C for 5 s, 72°C for 15 s, and then incubation at 72°C for 5 min, and kept at 4°C.

PCR-amplified DNA fragments were observed by agarose gel electrophoresis in 0.8 % 1X TAE agarose gels. 10 μ l of each amplification mixture containing 1x purple gel loading dye (B7025S, New England Biolabs, USA) and 5 μ l of Quick Load® 1-kb DNA ladder (N0468, New England Biolabs, USA) were loaded in the gel. Migration occurred at 135 V for 20 min. After spending 15 min in the ethidium bromide bath, the amplified DNA fragments were visualized by UV illumination.

Production of the different cellular fractions

The *E. coli* strain Stellar with the constructs F5minSus, F5minMFS, F5minGH-GH2, F5minGH-GH2P, Sus-GH2, Sus-GH2P, MFS-GH2 and MFS-GH2P were grown in 250 ml of LB media containing 12.5 mg/l chloramphenicol and 50 mg/l kanamycin (for the controls F5minSus and F5minMFS, only supplement 12.5 mg/l chloramphenicol) at 37°C with orbital shaking at 200 rpm, with an initial optical density at 600nm (OD_{600nm}) of 0.05. Next, the cultures of F5 variants harboring the pBtGH2 and pBtGH2P plasmids were induced by the addition of 0.5 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) in the mid-exponential phase (OD_{600nm}~0.6) and then incubated for 4 h at 37°C. The cells were collected by centrifugation at 4°C for 10 min at 5000 rpm, and the supernatant was filtered (0.22 μ m) and this fraction was called “secreted fraction”.

The other protein fractions were obtained from the different treatments of the pellet as described by Larsbrink *et al.* (Larsbrink et al. 2011). The pellet was gently resuspended and washed with 10 ml of 50 mM Tris-HCl (pH7.7), and collected by centrifugation at 4400 g for 10 min at 4°C. Next, the pellet was resuspended in 50 ml of 30 mM Tris-HCl, 20 % (w/v) sucrose and 1 mM EDTA (pH8.0), and the cells were incubated at room temperature for 10 min strictly. They were

centrifuged at 4400 g for 15 min at 4°C and the pellet was resuspended in 50 ml ice-cold 5mM MgSO₄ and incubated on ice for 10 min. Then the cells were collected by centrifugation at 14000 g for 10 min at 4°C. Finally, the supernatant was filtered (0.22 µm) and called “periplasmic fraction” for the hydrolysis test. The pellet was resuspended in specific volume of 50 mM sodium phosphate buffer (pH7.4) to ensure final OD_{600nm}=80 for cellular fractions extraction. The resuspension was broken by sonication, 5 cycles of 20 seconds, with the probe at 30 % of the maximal power, separated by 4 min on ice. The lysate was centrifuged at 5000 g for 10 min at 4°C to remove the cell debris. Using an ultracentrifuge, the supernatant was centrifuged at 100000 g for 1 h at 4°C and filtered (0.22 µm). The resulting solution is called “cytoplasmic fraction”. Subsequently, the pellet of the lysate was resuspended in 15 ml of 100 mM sodium carbonate buffer (pH9.0) and centrifuged at 100000 g for 1 h at 4°C. The supernatant from this step was filtered (0.22 µm) and contained the potential trapped soluble proteins and/or weakly membrane-associated proteins. This fraction is called “potential fraction”. Finally, the pellet was resuspended in 1 ml of 50 mM sodium phosphate buffer (pH7.4) and called “membrane fraction”.

Based on this method, we collected secreted fraction, periplasmic fraction, cytoplasmic fraction, membrane fraction, and potential fraction.

Enzymatic assays

The different fractions cited above were used to perform the oligosaccharides hydrolysis assays. The enzymatic reaction media -200 µl in total -were prepared in 50 mM potassium phosphate pH7.0 using 80 µl of the different fractions supplemented by mannotriose or xylotriase at a final concentration of 0.2 % (w/v). The mixtures were incubated at 37°C and 200 rpm. 50 µl of reactions were collected at T_{1h} and T_{overnight} (about 20 hours), and then boiled at 95°C for 10min. The samples were analyzed by Thin Layer Chromatography (TLC).

Thin layer chromatography

TLC experiments were performed using the plates from Sigma (Product 53356, silica gel on TLC Al foil, 20 cm X 20 cm, no fluorescent indicator). 3 µl of each sample were loaded on the TLC plates and dried using a hairdryer. The plates were inserted in a TLC tank containing the migration buffer (butan-1-ol/acetic acid/H₂O (2:1:1, v/v/v)). When the solvent migration front was near the top of the plates (0.5 cm), they were dried and put into the revelation buffer (1 g/l

orcinol, 75 % (v/v) ethanol, and 3 % (v/v) sulfuric acid) for 1 minute and shaken gently. The sugars were revealed at high temperature (>60°C) using a hairdryer.

Growth study

The F5minSus, F5minMFS, Sus-GH2, Sus-GH2P, MFS-GH2, MFS-GH2P strains stored in glycerol at -80°C were recovered on LB agar plates supplemented with the 12.5 mg/l chloramphenicol and 50mg/l kanamycin (for the controls F5minSus and F5minMFS, only 12.5 mg/l chloramphenicol). Precultures of these strains were made from three isolated colonies added to liquid LB medium supplemented with appropriate antibiotics and incubated overnight at 37°C with orbital shaking at 200 rpm. These precultures were used to inoculate 500 µl of M9 medium containing 0.5 % (w/v) xylotriase or mannotriase at an initial OD_{600 nm} of 0.05 into a 48-well microplate. For the F5 variants harboring the p*Bt*GH2 and p*Bt*GH2P plasmids, the M9 medium was supplemented with IPTG at a final concentration of 0.2 mM to induce the expression of the GH2 gene. The M9 medium contained (Na₂HPO₄·12H₂O, 17.4 g/l; KH₂PO₄, 3.03 g/l; NaCl, 0.51 g/l; NH₄Cl, 2.04 g/l; MgSO₄, 0.49 g/l; CaCl₂, 4.38 mg/l; Na₂EDTA·2H₂O, 15 mg/l; ZnSO₄·7H₂O, 4.5 mg/l; CoCl₂·6H₂O, 0.3 mg/l; MnCl₂·4H₂O, 1 mg/l; H₃BO₃, 1 mg/l; Na₂MoO₄·2H₂O, 0.4 mg/l; FeSO₄·7H₂O, 3 mg/l; CuSO₄·5H₂O, 0.3 mg/l; thiamine, 0.1 g/l; and leucine, 0.02 g/l). Cell growth was followed by measuring the OD_{600nm} over 48 h at 37°C and 200 rpm.

References

- Abou Hachem, Maher, Joakim M. Andersen, Rodolphe Barrangou, Marie S. Møller, Folmer Fredslund, Avishek Majumder, Morten Ejby, et al. 2013. 'Recent Insight into Oligosaccharide Uptake and Metabolism in Probiotic Bacteria'. *Biocatalysis and Biotransformation* 31 (4): 226–35. <https://doi.org/10.3109/10242422.2013.828048>.
- Andersen, Joakim M., Rodolphe Barrangou, Maher Abou Hachem, Sampo Lahtinen, Yong Jun Goh, Birte Svensson, and Todd R. Klaenhammer. 2011. 'Transcriptional and Functional Analysis of Galactooligosaccharide Uptake by lacS in *Lactobacillus Acidophilus*'. *Proceedings of the National Academy of Sciences of the United States of America* 108 (43): 17785–90. <https://doi.org/10.1073/pnas.1114152108>.
- Camarillo-Guerrero, Luis F., Alexandre Almeida, Guillermo Rangel-Pineros, Robert D. Finn, and Trevor D. Lawley. 2021. 'Massive Expansion of Human Gut Bacteriophage Diversity'. *Cell* 184 (4): 1098–1109.e9. <https://doi.org/10.1016/j.cell.2021.01.029>.
- Chamberlin, Michael, and Janet Ring. 1973. 'Characterization of T7-Specific Ribonucleic Acid Polymerase: I. GENERAL PROPERTIES OF THE ENZYMATIC REACTION AND THE TEMPLATE SPECIFICITY OF THE ENZYME'. *Journal of Biological Chemistry* 248 (6): 2235–44. [https://doi.org/10.1016/S0021-9258\(19\)44211-7](https://doi.org/10.1016/S0021-9258(19)44211-7).
- Chibani, Cynthia Maria, Alexander Mahnert, Guillaume Borrel, Alexandre Almeida, Almut Werner, Jean-François Brugère, Simonetta Gribaldo, Robert D. Finn, Ruth A. Schmitz, and Christine Moissl-Eichinger. 2022. 'A Catalogue of 1,167 Genomes from the Human Gut Archaeome'. *Nature Microbiology* 7 (1): 48–61. <https://doi.org/10.1038/s41564-021-01020-9>.
- Cowles, Charles E., Yongfeng Li, Martin F. Semmelhack, Ileana M. Cristea, and Thomas J. Silhavy. 2011. 'The Free and Bound Forms of Lpp Occupy Distinct Subcellular Locations in *Escherichia Coli*'. *Molecular Microbiology* 79 (5): 1168–81. <https://doi.org/10.1111/j.1365-2958.2011.07539.x>.
- Déjean, Guillaume, Kazune Tamura, Adriana Cabrera, Namrata Jain, Nicholas A. Pudlo, Gabriel Pereira, Alexander Holm Viborg, Filip Van Petegem, Eric C. Martens, and Harry Brumer. 2020. 'Synergy between Cell Surface Glycosidases and Glycan-Binding Proteins Dictates the Utilization of Specific Beta(1,3)-Glucans by Human Gut

- Bacteroides'. *mBio* 11 (2): 10.1128/mbio.00095-20.
<https://doi.org/10.1128/mbio.00095-20>.
- Dubendorf, John W., and F. William Studier. 1991. 'Controlling Basal Expression in an Inducible T7 Expression System by Blocking the Target T7 Promoter with *Lac* Repressor'. *Journal of Molecular Biology* 219 (1): 45–59.
[https://doi.org/10.1016/0022-2836\(91\)90856-2](https://doi.org/10.1016/0022-2836(91)90856-2).
- Ejby, Morten, Folmer Fredslund, Joakim Mark Andersen, Andreja Vujičić Žagar, Jonas Rosager Henriksen, Thomas Lars Andersen, Birte Svensson, Dirk Jan Slotboom, and Maher Abou Hachem. 2016. 'An ATP Binding Cassette Transporter Mediates the Uptake of α -(1,6)-Linked Dietary Oligosaccharides in Bifidobacterium and Correlates with Competitive Growth on These Substrates'. *Journal of Biological Chemistry* 291 (38): 20220–31. <https://doi.org/10.1074/jbc.M116.746529>.
- Fontes, C. M. G. A., J. Hall, B. H. Hirst, G. P. Hazlewood, and H. J. Gilbert. 1995. 'The Resistance of Cellulases and Xylanases to Proteolytic Inactivation'. *Applied Microbiology and Biotechnology* 43 (1): 52–57. <https://doi.org/10.1007/BF00170622>.
- Golomb, Miriam, and Michael Chamberlin. 1974. 'Characterization of T7-Specific Ribonucleic Acid Polymerase: IV. RESOLUTION OF THE MAJOR IN VITRO TRANSCRIPTS BY GEL ELECTROPHORESIS'. *Journal of Biological Chemistry* 249 (9): 2858–63. [https://doi.org/10.1016/S0021-9258\(19\)42709-9](https://doi.org/10.1016/S0021-9258(19)42709-9).
- Gregory, Ann C., Olivier Zablocki, Ahmed A. Zayed, Allison Howell, Benjamin Bolduc, and Matthew B. Sullivan. 2020. 'The Gut Virome Database Reveals Age-Dependent Patterns of Virome Diversity in the Human Gut'. *Cell Host & Microbe* 28 (5): 724-740.e8. <https://doi.org/10.1016/j.chom.2020.08.003>.
- Hamaker, Bruce R., and Yunus E. Tuncil. 2014. 'A Perspective on the Complexity of Dietary Fiber Structures and Their Potential Effect on the Gut Microbiota'. *Journal of Molecular Biology, Insights Into Molecular Mechanisms of Microbiota*, 426 (23): 3838–50. <https://doi.org/10.1016/j.jmb.2014.07.028>.
- Harayama, S, J Bollinger, T Iino, and G L Hazelbauer. 1983. 'Characterization of the Mgl Operon of Escherichia Coli by Transposon Mutagenesis and Molecular Cloning'.

- Journal of Bacteriology* 153 (1): 408–15. <https://doi.org/10.1128/jb.153.1.408-415.1983>.
- Huttenhower, Curtis, Dirk Gevers, Rob Knight, Sahar Abubucker, Jonathan H. Badger, Asif T. Chinwalla, Heather H. Creasy, et al. 2012. ‘Structure, Function and Diversity of the Healthy Human Microbiome’. *Nature* 486 (7402): 207–14. <https://doi.org/10.1038/nature11234>.
- Janey Lee, Wiggs, John W. Bush, and Michael J. Chamberlin. 1979. ‘Utilization of Promoter and Terminator Sites on Bacteriophage T7 DNA by RNA Polymerases from a Variety of Bacterial Orders’. *Cell* 16 (1): 97–109. [https://doi.org/10.1016/0092-8674\(79\)90191-0](https://doi.org/10.1016/0092-8674(79)90191-0).
- Koropatkin, Nicole M., Elizabeth A. Cameron, and Eric C. Martens. 2012. ‘How Glycan Metabolism Shapes the Human Gut Microbiota’. *Nature Reviews Microbiology* 10 (5): 323–35. <https://doi.org/10.1038/nrmicro2746>.
- Larsbrink, Johan, Atsushi Izumi, Farid M. Ibatullin, Azadeh Nakhai, Harry J. Gilbert, Gideon J. Davies, and Harry Brumer. 2011. ‘Structural and Enzymatic Characterization of a Glycoside Hydrolase Family 31 α -Xylosidase from *Cellvibrio Japonicus* Involved in Xyloglucan Saccharification’. *Biochemical Journal* 436 (3): 567–80. <https://doi.org/10.1042/BJ20110299>.
- Moffatt, Barbara A., and F. William Studier. 1987. ‘T7 Lysozyme Inhibits Transcription by T7 RNA Polymerase’. *Cell* 49 (2): 221–27. [https://doi.org/10.1016/0092-8674\(87\)90563-0](https://doi.org/10.1016/0092-8674(87)90563-0).
- Oiki, Sayoko, Reiko Kamochi, Bunzo Mikami, Kousaku Murata, and Wataru Hashimoto. 2017. ‘Alternative Substrate-Bound Conformation of Bacterial Solute-Binding Protein Involved in the Import of Mammalian Host Glycosaminoglycans’. *Scientific Reports* 7 (1): 17005. <https://doi.org/10.1038/s41598-017-16801-8>.
- Oiki, Sayoko, Bunzo Mikami, Yukie Maruyama, Kousaku Murata, and Wataru Hashimoto. 2017. ‘A Bacterial ABC Transporter Enables Import of Mammalian Host Glycosaminoglycans’. *Scientific Reports* 7 (1): 1069. <https://doi.org/10.1038/s41598-017-00917-y>.

- Raimondi, Stefano, Alberto Amaretti, Caterina Gozzoli, Marta Simone, Lucia Righini, Francesco Candeliere, Paola Brun, et al. 2019. 'Longitudinal Survey of Fungi in the Human Gut: ITS Profiling, Phenotyping, and Colonization'. *Frontiers in Microbiology* 10 (July):1575. <https://doi.org/10.3389/fmicb.2019.01575>.
- Rogowski, Artur, Jonathon A. Briggs, Jennifer C. Mortimer, Theodora Tryfona, Nicolas Terrapon, Elisabeth C. Lowe, Arnaud Baslé, et al. 2015. 'Glycan Complexity Dictates Microbial Resource Allocation in the Large Intestine'. *Nature Communications* 6 (1): 7481. <https://doi.org/10.1038/ncomms8481>.
- Shulami, Smadar, Arie Zehavi, Valery Belakhov, Rachel Salama, Shifra Lansky, Timor Baasov, Gil Shoham, and Yuval Shoham. 2020. 'Cross-Utilization of β -Galactosides and Cellobiose in *Geobacillus Stearothermophilus*'. *Journal of Biological Chemistry* 295 (31): 10766–80. <https://doi.org/10.1074/jbc.RA120.014029>.
- Studier, F. William. 1991. 'Use of Bacteriophage T7 Lysozyme to Improve an Inducible T7 Expression System'. *Journal of Molecular Biology* 219 (1): 37–44. [https://doi.org/10.1016/0022-2836\(91\)90855-Z](https://doi.org/10.1016/0022-2836(91)90855-Z).
- Tailford, Louise E., Victoria A. Money, Nicola L. Smith, Claire Dumon, Gideon J. Davies, and Harry J. Gilbert. 2007. 'Mannose Foraging by *Bacteroides Thetaiotaomicron*'. *Journal of Biological Chemistry* 282 (15): 11291–99. <https://doi.org/10.1074/jbc.M610964200>.
- Tamura, Kazune, Guillaume Dejean, Filip Van Petegem, and Harry Brumer. 2021. 'Distinct Protein Architectures Mediate Species-Specific Beta-Glucan Binding and Metabolism in the Human Gut Microbiota'. *Journal of Biological Chemistry* 296 (January):100415. <https://doi.org/10.1016/j.jbc.2021.100415>.
- Tamura, Kazune, Matthew H. Foley, Bernd R. Gardill, Guillaume Dejean, Matthew Schnizlein, Constance M. E. Bahr, A. Louise Creagh, Filip van Petegem, Nicole M. Koropatkin, and Harry Brumer. 2019. 'Surface Glycan-Binding Proteins Are Essential for Cereal Beta-Glucan Utilization by the Human Gut Symbiont *Bacteroides Ovatus*'. *Cellular and Molecular Life Sciences* 76 (21): 4319–40. <https://doi.org/10.1007/s00018-019-03115-3>.

- Tauzin, Alexandra S., Elisabeth Laville, Yao Xiao, Sébastien Nouaille, Pascal Le Bourgeois, Stéphanie Heux, Jean-Charles Portais, et al. 2016. 'Functional Characterization of a Gene Locus from an Uncultured Gut *Bacteroides* Conferring Xylo-oligosaccharides Utilization to *Escherichia Coli*'. *Molecular Microbiology* 102 (4): 579–92. <https://doi.org/10.1111/mmi.13480>.
- Theilmann, Mia Christine, Folmer Fredslund, Birte Svensson, Leila Lo Leggio, and Maher Abou Hachem. 2019. 'Substrate Preference of an ABC Importer Corresponds to Selective Growth on β -(1,6)-Galactosides in *Bifidobacterium Animalis* Subsp. *Lactis*'. *Journal of Biological Chemistry* 294 (31): 11701–11. <https://doi.org/10.1074/jbc.RA119.008843>.
- Tokuda, Hajime, and Shin-ichi Matsuyama. 2004. 'Sorting of Lipoproteins to the Outer Membrane in *E. Coli*'. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1693 (1): 5–13. <https://doi.org/10.1016/j.bbamcr.2004.02.005>.
- Wang, Zhi, Alexandra S. Tauzin, Elisabeth Laville, Pietro Tedesco, Fabien Létisse, Nicolas Terrapon, Pascale Lepercq, Myriam Mercade, and Gabrielle Potocki-Veronese. 2020. 'Harvesting of Prebiotic Fructooligosaccharides by Nonbeneficial Human Gut Bacteria'. Edited by Vincent B. Young. *mSphere* 5 (1): e00771-19. <https://doi.org/10.1128/mSphere.00771-19>.
- Webb, Alexander J., Karen A. Homer, and Arthur H. F. Hosie. 2008. 'Two Closely Related ABC Transporters in *Streptococcus Mutans* Are Involved in Disaccharide and/or Oligosaccharide Uptake'. *Journal of Bacteriology* 190 (1): 168–78. <https://doi.org/10.1128/jb.01509-07>.
- Wilson, Marlana M., and Harris D. Bernstein. 2016. 'Surface-Exposed Lipoproteins: An Emerging Secretion Phenomenon in Gram-Negative Bacteria'. *Trends in Microbiology* 24 (3): 198–208. <https://doi.org/10.1016/j.tim.2015.11.006>.

Supplemental data

Cloning of the *BtGH2* with its native peptide signal into pET-28a(+)

TCTAGA XbaI

vector and RBS

CTCGAG XhoI

atgat..native peptide signal

TCTAGAaataattttgtttaactttaagaaggagatataccatgatcttaaatttgacaggaaagatagccccgatagcatgtgg
attgctatgttgctgtagtatggtttatgctcaaggtaatgacacttcggaagtcagctactggataccggtgggaattttctcaatca
gggactgaaaagtggatgccagccacgggtcctgggtacgggtcatcaggacctgatcagtcataaactgcccgaatccttttatggg
atgaatgaaaagaaaattcaatgggtggagaatgaggactgggaatatagaacctcctttatagtctccgaagagcaattgaatcgtgac
ggaatccaacttattttgaaggcttgatacctatgcagacgtatatctcaacgggtcattactgctgaaggctgataatatgtttgtgggtt
atacacttctgtaaaatctgtgttcggaagggcgagaatcacttatatattttattccattcacctattcggcaaacggtgccacaatatgc
ttccaatggcttcaattatccggcgataacgatcatcacgagaacacatttagtgtgttcagccgtaaggctccgtatagttacgggttg
gactggggcatccgaatggtgacaagtgggtgtctggcgcccggttactcttggtttacgacatagctaccatatccgactactatgtca
gacaattatcgctgacagatgaaaacgctcgtctgtctaatgagctgattgtaaaccagatcgttcctcagaaaattccggctgaagtaag
ggtcaatgtatcgttgatggaactaccgttacagaagtaaacagcaggtgacactgcaaccgggataaatcatatcactcttctgc
tgaagtacgaaccctgtacgatggatgccaaatggttggggaactccgactctgtatgatttctcggtcagatagcttggtgagatcg
gattgttgcggaacaatctcatcggttggtttacgaacgattcgtgtggttaaatgaaaagataaggacggagaatctttttatttgaagt
caatggcataccgatgtttgccaaaggagctaattatattcctcaggatgcactgctgcctaattgtactacggaacgctatcagactttgt
ccgtgacatgaaggaggccaatatgaatatggttcgtatatgggggtggtgggacatatgagaataatcttttctatgacctgccgatgag
aacgggatactgtctggcaggattttatgtttgcgtgtaccccttatccatccgatcctactttcctgaaacgggtagaggcggaggctgt
ttacaatattcgtcgtcttcggaatcatgcctctttggcgatgtggtgtggttaataatgaaatactggaagctctgaaatattggggttttgag
aagaagtttactccggagggtttatcaaggattaatgcatggttatgataaactgttccgtgagctgttgccttcaacgggttaagaatttgatt
cggatcgttttatgtgcacagttcgccttatctcgtactggggcgctccggagtcattgggggacaggagacagtcataattgggga
gtctggtacggaaagaagccatttgagcttttagataccgatttgccgcgttttatgagtgaatttggttccagtcatttccggagatgaaa
actattgctgcttttgcgtcctgaagattatcagattgagctgaagtgatgaatgcgcacacagaagagtagatttggaattcgttgatcc
gtacttatatggaacgtgattatattcccgaaagtttcgaggattttgtgtatgtcggattagtcttgcaaggacaggggatgcggcat
ggactggaagctcacagacgcaatcgtccttattgtatggggacattgtattggcagttgaacgacagttggccggtagtatcctggtca
agtattgattattatggttaattggaaggcattacattatcaggcaaagcgtgcttttgcaccgggtcttgatcaatcctatacagcagaatgac

agtttgagtgtctatctgatctcagaccggcttgatacaatggagcaaatgacattggaaatgaaagtcgtggattttgatggaaaaacgc
tgggtaagaagatacaagtacattcgttggaagttccggcaataacttcaaaatgtgtttatcgtgcaaaactggatggctggctgactcc
cgaagactgccgacgtagcttcttgaattgatactcaaagataaatcgggacatcaggtggctgagagcgtccatttttcagaaagac
gaaagatttgcaacttctccgacatcagatcctatcaaatgaagcagacggatggtaaattggaattgactctctttcatcaatgttggc
aaaagatattttattgagactccattgcaaggagcccgttacagtgacaacttcttgatctcttaccggagaacgtaagaaagtaattat
cacttctccccggataaagaagggagaagaattacgtgcaatatcaagcatattcgagaaacctataaa **CTCGAG**

Cloning of the *BtGH2* with the *E. coli* D-galactose/methyl-galactoside binding periplasmic protein peptide signal into pET-28a(+)

TCTGAG XbaI

vector and RBS

CTCGAG XhoI

atgaa..peptide signal from the *E. coli* D-galactose/methyl-galactoside binding periplasmic protein (UniPort P0AEE5)

TCTAGAaataattttgtttaactttaagaaggagatataccatgaataagaagggtttaaccctgtctgctgctgatggccagcatg
ttattcgggtgccgctgcacacgctcaaggtaatgacacttcggaagtcactgctactggataccgggttgggaattttctcaatcagggact
gaaaagtggatgccagccacggttcctggtacggttcacaggacctgatcagtcgaattactgccgaatccttttatgggatgaatg
aaaagaaaattcaatgggtggagaatgaggactgggaatatagaacctcctttatagctcctgaagagcaattgaatcgtgacggaatc
caactattttgaaggcttgatcacctatgcagacgtatatctcaacggctcattactgctgaaggctgataatatgtttgtgggttatact
tctgtaaaatctgtgttcggaaggcgagaatcacttatatattttccattcacctattcggcaaacgttgccacaatatgcttccaat
ggcttcaattatccggcggataacgatcatcacgagaacatcttagtggttcagccgtaaggctccgtatagttacgggttgggactgg
ggcatccgaatggtgacaagtgggtgtctggcggccggttacttctgttttacgacatagctaccatatccgactactatgtcagacaatt
atcgtgacagatgaaaacgctcgtctgtctaatgagctgattgtaaacagatcgttcctcagaaaattccggctgaagtaagggtcaat
gtatcgttgatggaactaccgttacagaagtaaacagcaggtagactgcaacccgggataaatcatatcacttctcgtgaagtg
acgaacctgtacgatggatgccaaatggttggggaactccgactctgtatgatttctcggctcagatagcttgggagatcggattgtg
cggacaatctcatcggattggtttacgaacgattcgtgtgtaaatgaaaaagataaggacggagaatctttttatttgaagtaaatggc
ataccgatgtttgccaaaggagctaattatattcctcaggatgcactgctgcctaattgtactacggaacgctatcagactttgtccgtgac
atgaaggaggccaatatgaatatggttcgtatatggggtggtgggacatatgagaataatctttctatgaccttccgatgagaacggg
atactgtctggcaggattttatgtttcgtgtaccccttatccatccgatcctacttctgaaacgggtagaggcggaggctgtttacaat

attcgtcgtcttcggaatcatgcctctttggcgatgtggtgtggaataatgaaatactggaagctctgaaatattggggtttgagaagaa
gtttactccggaggtttatcaaggattaatgcatggttatgataaactgtccgtgagctgttgcctcaacggtaagaatttgattcggat
cgttttatgtgcacagttcgcttatctcgctaactgggggcgtccggagtcacggggacaggagacagtcataattggggagtctgg
tacggaaagaagccatttgagctttagataccgatttgccgcgtttatgagtgaatttggttccagtcattccggagatgaaaactattg
ctgcttttctgctcctgaagattatcagattgagctgaagtgaatgcgcatacagaagagtagtattgtaattcgttgatccgtactta
tatggaacgtgattatattcccgaagtttcgaggattttgtgtatgtcggattagcttgcaaggacaggggatcgggcatggactg
gaagctcacagacgcaatcgtccttattgtatggggacattgtattggcagttgaacgacagttggccggtagtatcctggtcaagtattg
attattatggaattggaaggcattacattatcaggcaaagcgtgctttgcaccggtcttgatcaatcctatacagcagaatgacagtttga
gtgtctatctgatctcagaccggcttgatacaatggagcaaatacattggaatgaaagtcgtggattttgatggaaaaacgctgggta
agaagatacaagtagcttgaagtccggcaaatacttcaaaatgtgtttatcgtgcaaaactggatggctggctgactcccgaag
actgccgacgtagcttcttgaaattgatactcaaagataaatcgggacatcaggtggctgagagcgtccatttttcagaaagacgaaag
atttgcaacttctccgacatcagatcctatcaaatgaagcagacggatggtaaatgtgaattgactctctttcatcaatgttggcaaaag
atattttattgagactccattgcaaggagcccgttacagtgacaacttctttgatctcttaccggagaacgtaagaaagtaattatcacttc
tccccgataaagaagggagaagaattacctgtcaatatcaagcatattcgagaaacctataaaCTCGAG

Third chapter

Oligosaccharide transport systems often have limitations in terms of substrate specificity and transport efficiency. Various approaches have been used in the past decades to engineer oligosaccharide transporters in order to enhance the transport of specific glycosides across cell membranes and increase substrate utilization to supply cell growth. Among them, the ALE method presents a particular interest. This approach consists of the cultivation of a microorganism under defined conditions for prolonged periods of time, in the range of weeks to years, which allows the spontaneous occurrence of mutations and the selection of improved phenotypes. Moreover, ALE has been demonstrated to be an effective approach for evolving microbial transport systems. However, to date, there is no example of ALE study used to engineer SusC/D transporters. Furthermore, despite the simplicity of the ALE's principle and of its use in *E. coli*, this approach has never been used to engineer recombinant transporters produced in *E. coli*, whatever the origin of those proteins. Here, we performed a preliminary study to assess the potential of ALE to engineer the F5 XOS utilization pathway expressed in *E. coli*. This study was initiated based on an idea of Matthieu Arlat during one of my thesis committees, and performed in collaboration with Isabelle Meynial-Salles, who leads the team "Engineering and Evolution of Metabolic Pathways in Prokaryotes" at the Toulouse Biotechnology Institute.

Adaptive laboratory evolution of a recombinant xylooligosaccharide utilization locus expressed in *Escherichia coli*

Introduction

As presented in the introduction of this thesis manuscript, transporters specialized in the uptake of oligosaccharides are important components of microbial physiology. They facilitate the acquisition of sugars from the extracellular environment and contribute to microbial fitness. However, these transport systems often have limitations in terms of substrate specificity, transport efficiency, and regulatory control. Previous studies have shown that natural transporter proteins for oligosaccharides typically have low expression and poor transport capacity, leading to suboptimal transport processes and substrate utilization (Zhu et al. 2020; Guirimand et al. 2019; Kim et al. 2018; Farwick et al. 2014; Andersen et al. 2013; Ha et al. 2013). To enhance the transport of specific glycosides across cell membranes and increase substrate utilization to supply cell growth and, in fine, to increase the production of fermentation products in synthetic biology or biorefinery processes, there has been a growing interest in engineering oligosaccharide transporters.

As presented in the second section of the bibliographic introduction, several approaches have been used in the past decades to engineer oligosaccharide transporters. Among them, adaptive laboratory evolution (ALE) presents a particular interest.

The ALE approach consists in cultivation of a microorganism under defined conditions for prolonged periods of time, in the range of weeks to years, which allows the spontaneous occurrence of mutations and the selection of improved phenotypes. ALE requires microbial strains that have i) simple nutrient requirements for an easy cultivation in the laboratory and ii) high specific rates of biomass increase, in the range of $\mu = 0.05$ to 1.0 h^{-1} (Dragosits and Mattanovich 2013). ALE is thus currently used to investigate the evolutionary dynamics of microorganisms and to rapidly improve a specific property by applying a selective pressure, mimicking what naturally happens in the evolutionary process originally highlighted by Darwin's seminal work 'On the Origin of Species' in 1859. Since microorganisms tend to evolve rapidly, they display strong evolutionary adaptations in a short time, which has been widely exploited in ALE for metabolic engineering (Van den Bergh et al. 2018; Portnoy et al.

2011) targeting various applications, in particular in *Saccharomyces cerevisiae* (Wisselink et al. 2007; Kuyper et al. 2005; Cakar et al. 2005; Van Maris et al. 2004; Sonderegger and Sauer 2003) and *E. coli* (Alva et al. 2020; Meynial Salles et al. 2007; Fong et al. 2005; Yomano et al. 1998).

However, ALE has been used only in a small number of studies to improve oligosaccharide uptake across cellular membranes. Let's mention the improvement of cellobiose fermentation rate in *S. cerevisiae* (Kim et al. 2018; Ha et al. 2013), and the study of the evolution of sucrose transport in *E. coli* (Mohamed et al. 2019). Each of those studies emphasized the importance of mutations in specific transporters (MFS in *S. cerevisiae*, sucrose permease in *E. coli*) or related regulatory elements. ALE has thus been demonstrated to be an effective approach to evolve microbial transport systems, thereby demonstrating the potential of this approach to uncover novel genetic modifications that enhance microbial performance under selective growth conditions. However, to date, there is no example of ALE use to engineer Bacteroidota SusC/D transporters. Furthermore, despite the simplicity of the ALE's principle and of its use in *E. coli*, this approach has never been used to engineer recombinant transporters produced in *E. coli*, whatever the origin of those proteins.

In this chapter, we performed a preliminary study to assess the potential of ALE to engineer a recombinant xylooligosaccharide utilization pathway expressed in *E. coli*. Our idea was to exploit two current opportunities: the first one was related to the availability in the CIMES team of fosmidic *E. coli* clones expressing various recombinant oligosaccharide utilization pathways issued from uncultured bacteria, those pathways including glycoside-hydrolases and transporters that were proved to be functional in *E. coli*; the second one was related to the continuous decline of next generation sequencing costs to due to technological advancements and market competition, allowing us to assess, at low cost, the mutation rates and mutation positions during ALE experiments. As a model case, we chose a minimal version of the F5 XOS_PUL, originating from an uncultured Bacteroidota issued from the human gut microbiota (Cecchini et al. 2013), which shares 99 % DNA sequence identity with a *Bacteroides vulgatus* ATCC 8482 PUL (BVU_0037 to BVU_0043). As described in Tauzin *et al.*, this PUL confers to *E. coli* the ability to grow on xylooligosaccharides (XOS) (Tauzin et al. 2016). The function and structure of its glycoside-hydrolase and transporters components were studied in previous studies, in particular in this thesis work (Tauzin et al. 2016; 2022), focusing in particular on the SusC/D transport system. The objective of the present work was not only to assess the interest of ALE to engineer recombinant oligosaccharide transporters in *E. coli*, but also to perform the

first PUL engineering study based on random introduction of mutations, identification of potent *E. coli* mutant strains affected in their ability to utilize oligosaccharides, and ultimately, sequencing and characterization of a few mutants in order to analyze the structure-function relationships of the SusC/D transporter system, if subjected to mutations during growth.

Results

Rate of biomass increase evolution on xylotriase during ALE

The F5 XOS_PUL encodes three glycoside hydrolases (GH) from family 43 (GH43A and GH43B) and family 10 (GH10), as well as two transport systems: a SusC/D pair and an MFS transporter. Tauzin *et al.* constructed a series of F5 variants to investigate enzymatic activities and transport abilities (Fig. 1). One variant, F5minGH, contains only the three GH encoding genes, while another, F5minSus, harbors the three GH encoding genes and those encoding the SusC and SusD components of the SusC/D transport system dedicated to XOS uptake up to a degree of polymerization 3 (Tauzin et al. 2016). F5minGH, which does not have the genes encoding transporters, is unable to grow on XOS. Therefore, in order to affect XOS uptake ability using ALE, we used the *E. coli* strain transformed by the F5minSus construct as the initial monoclonal strain to evolve, after clone isolation on solid plate. The *E. coli* strain transformed by the F5minGH construct was used as negative control. The ALE experiment consisted in five serial transfers of 10 ml batch cultures in M9 minimal medium supplemented with 5 g/l of xylotriase as the sole carbon source (Fig. 2). Given the high price of xylotriase (3280 €/g, representing 1640 € for 5 cultures of 10 ml, performed for the strain to evolve and for the negative control), we did not perform culture replicates.

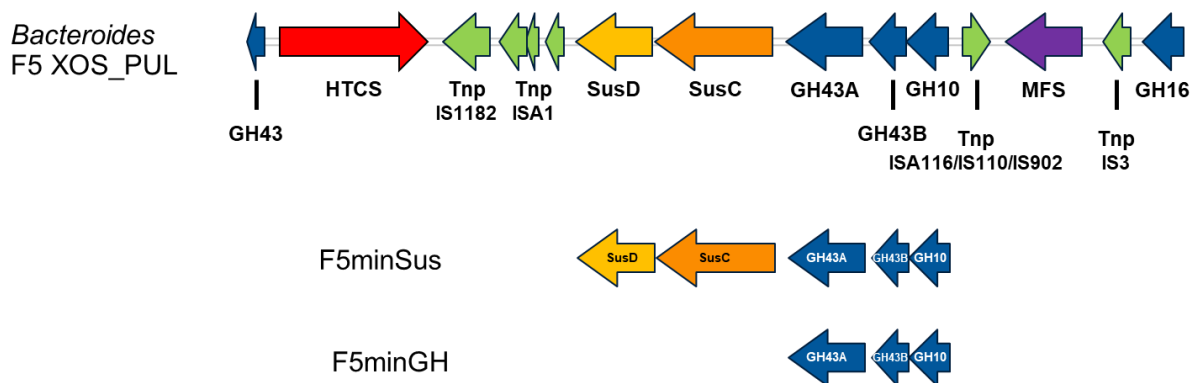


Figure 1: Schematic representation of the F5 XOS utilization locus and the truncated versions used in the present work.

Genes involved in xylooligosaccharide uptake and hydrolysis are color-coded: glycoside hydrolases (GH) with family number in blue; hybrid two component system (HTCS) in red; SusC in orange; SusD in yellow; transporter of the major facilitator superfamily (MFS) in purple.

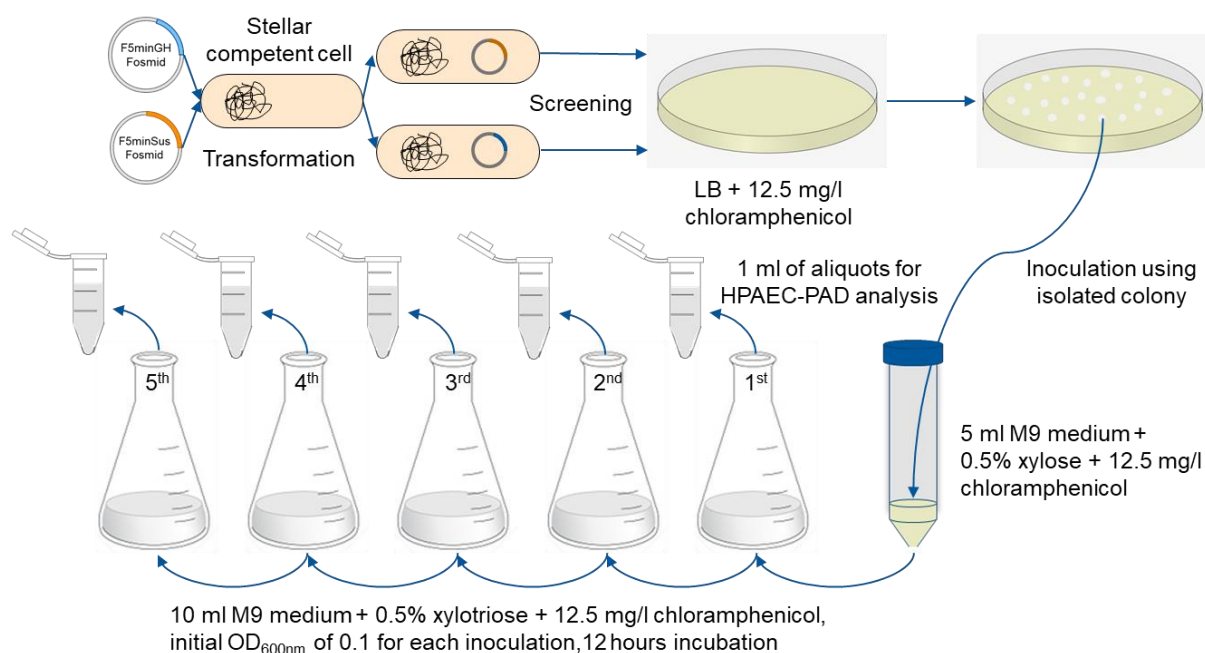


Figure 2: Overview of adaptive laboratory evolution (ALE) strategy used in the present work for improvement of xylotriose metabolization by an *E. coli* strain expressing the F5minSus locus (orange).

Strain expressing the truncated F5minGH locus (blue) was used as negative control. Selective medium was composed of M9 medium supplemented with 0.5 % xylotriose and 12.5 mg/l chloramphenicol. Successive cultures were performed during 12 hours at 37°C.

To ensure cells were in the exponential phase for serial transfers, we chose a culture time of 12 hours for all cultures, based on the growth curve of the initial F5minSus strain on XOS (Fig. 2) (Tauzin et al. 2016). Except for the 2nd and 4th cultures that were sampled at initial and 12 hours only, for each culture, a 1 ml culture aliquot was taken every 3 hours to measure OD_{600nm} and to quantify the concentration of remaining xylotriose. The growth curves in Fig. 3A demonstrated differences in the growth profiles of the various F5minSus generations. In comparison to the 1st culture at 6 hours, the 3rd and 5th generations of F5minSus exhibited a rate of biomass increase increased by 3.25 and 1.88 times, respectively. As anticipated, the negative control F5minGH did not significantly grow on xylotriose, since this strain does not harbor any xylooligosaccharide transporter required for substrate uptake and its subsequent hydrolysis in the intracellular space by the GHs. Indeed, OD_{600nm} only reached 0.18 after 12 hours of culture

(Fig. 3B). This slight OD_{600nm} increase might be due to cell lysis that naturally occurred during the culture, releasing active GHs in the culture medium, that may have hydrolyzed small amounts of xylotriose to xylobiose and xylose. Besides, no significant pH fluctuations were observed during ALE, the pH of the all culture supernatants remaining within the range of 7 - 8.

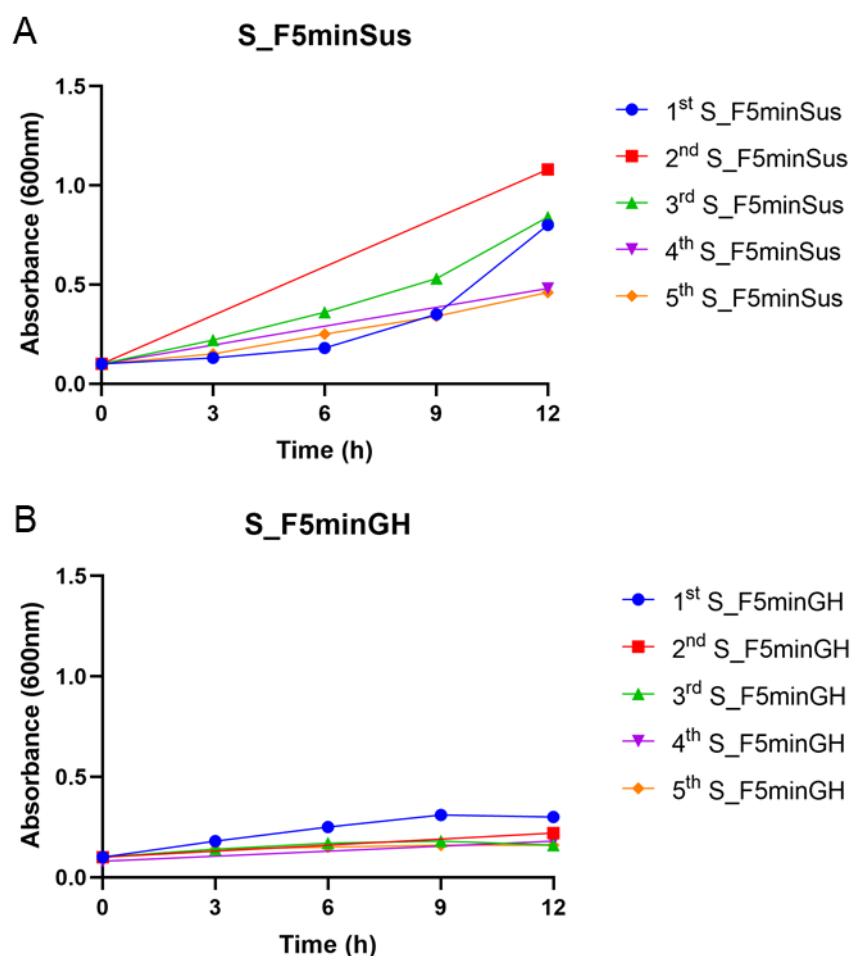


Figure 3: Growth curves of the five serial cultures of F5minSus (A) and F5minGH (B) in M9 minimal medium supplemented with 12.5 mg/l chloramphenicol and 0.5 % (w/v) xylotriose as the sole carbon source.

Kinetics of xylotriose utilization

Xylotriose concentration in the culture supernatants was quantified using HPAEC-PAD analysis. As shown in Fig. 4A, the xylotriose consumption rate was obviously altered during ALE. In the 1st F5minSus culture, only 2.1 % of xylotriose was consumed after 6 hours of growth, corresponding to the beginning of the exponential phase, against 14.5 % for the 3rd culture. At the end of the 1st, 2nd and 3rd cultures, xylotriose consumption was similar, comprised between 16.0 and 18.5 %. In contrast, the ability to uptake xylotriose was negatively affected in the 4th and 5th generations of F5minSus. The low amount of xylose observed between the initial time and 6 hours of the 1st F5minSus culture (between 0.11 and 0.04 mg/ml) is due to the content of the inoculum from the preculture, which contained 0.5 % xylose (Fig. 5A). Surprisingly, in the 2nd and 3rd cultures, the xylobiose concentration increased from 0.03 to 0.46 mg/ml between the initial and final growth times (Fig. 5B, 5C). This slight increase might be due to a slight cell lysis, which released the intracellular GHs. Nevertheless, no xylose was observed in these samples, this monosaccharide having been used by the growing cells.

Regarding the negative control, 5.5 % of xylotriose, corresponding to 0.27 mg/ml, was consumed at the end of the 1st F5minGH culture (Fig. 4B). Due to cell lysis, which released the intracellular GHs, this small amount of xylotriose might have been hydrolyzed to xylose, which entered into the central metabolism of *E. coli*, explaining why no xylose was visible in the HPAEC-PAD chromatogram at the end of the culture (Fig. 5F-J).

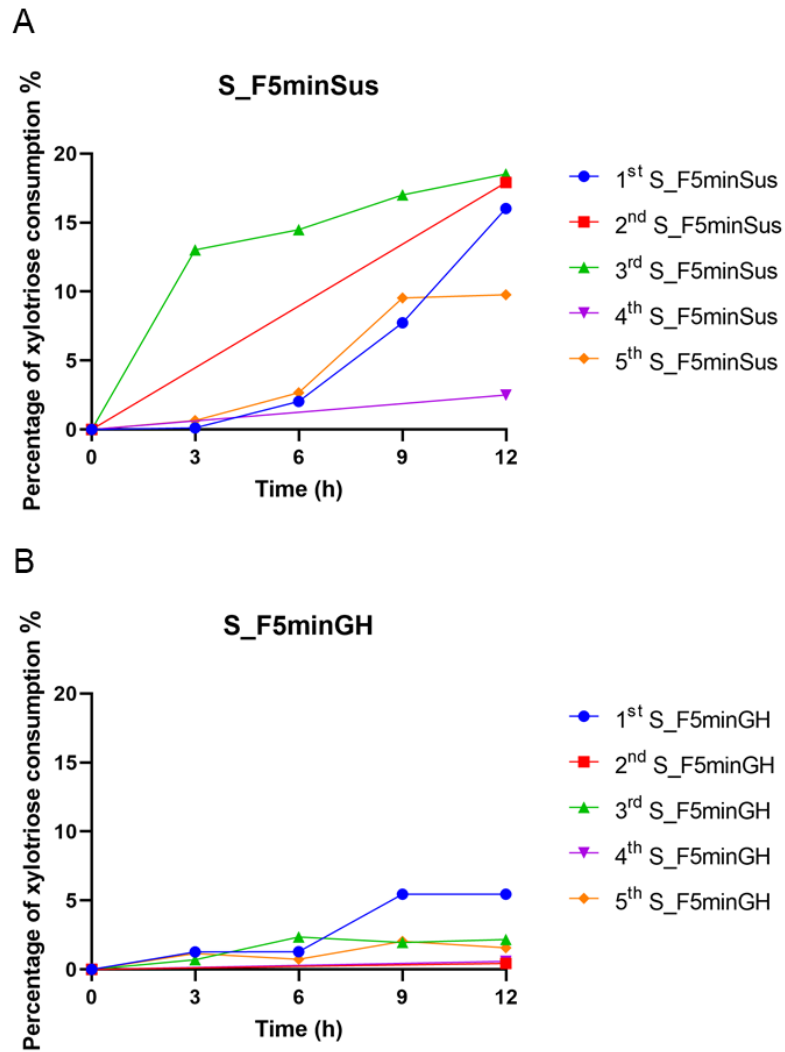
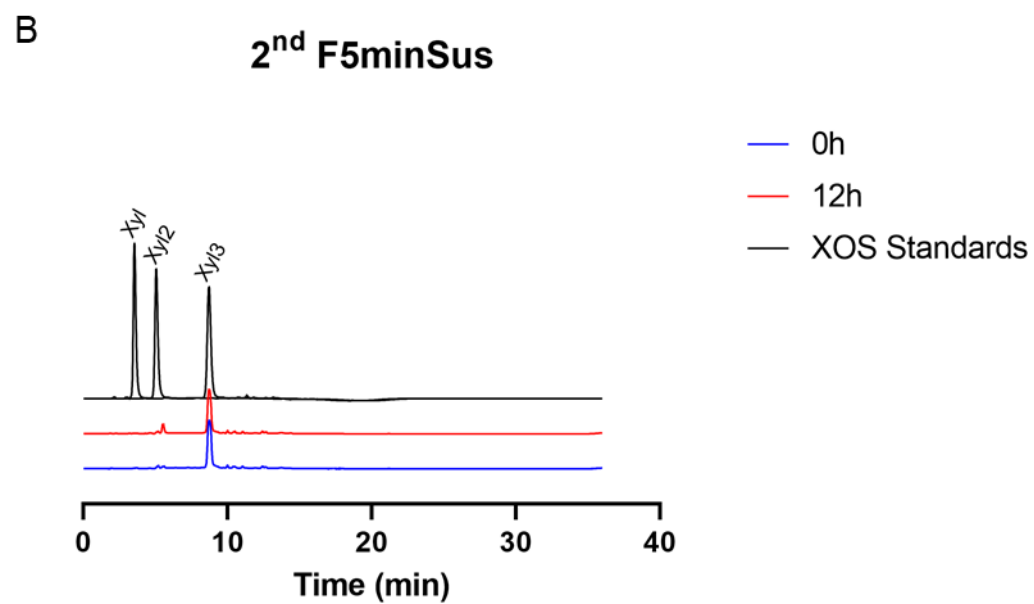
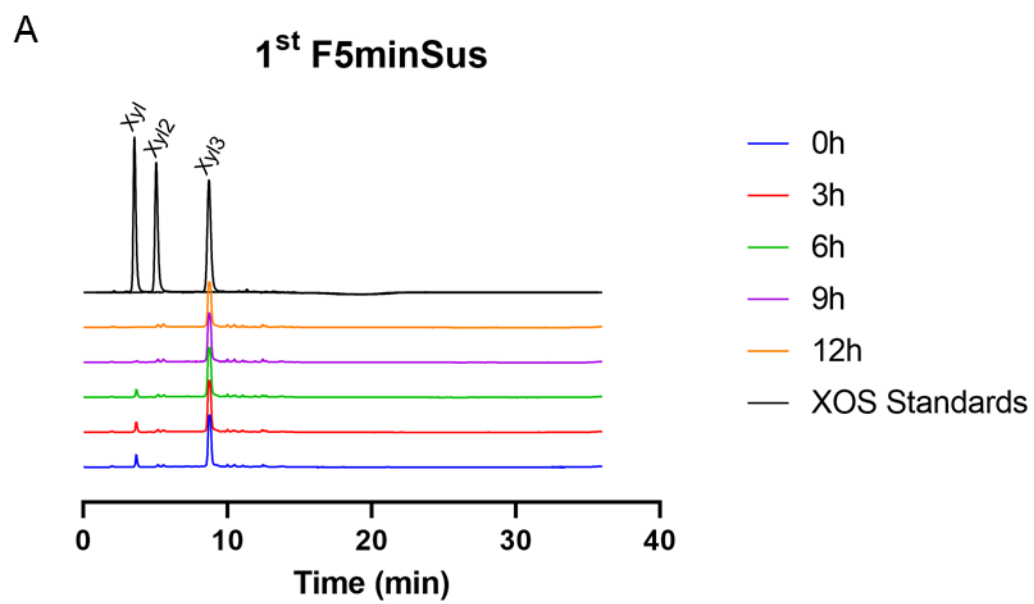
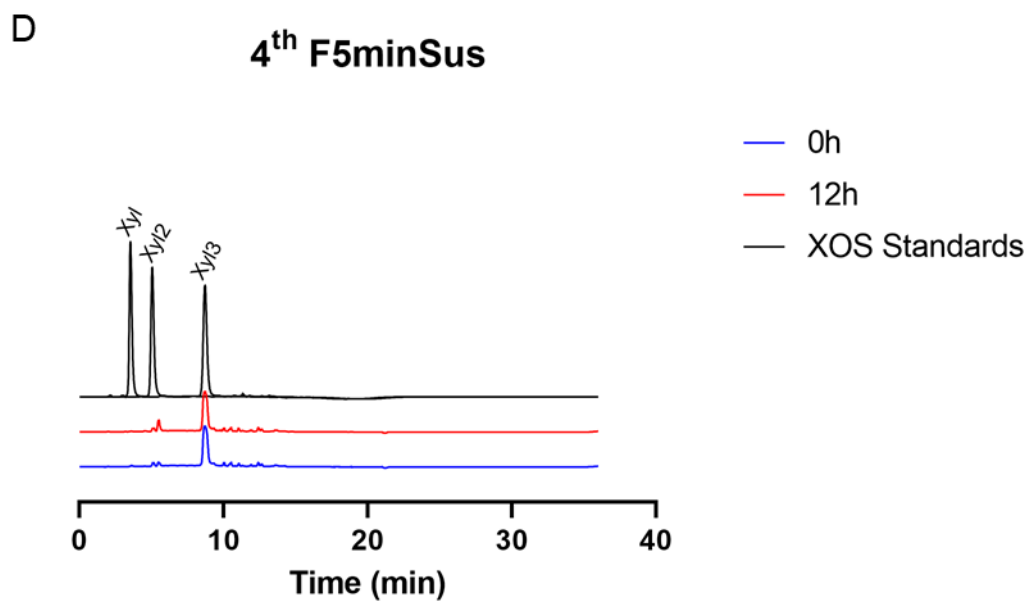
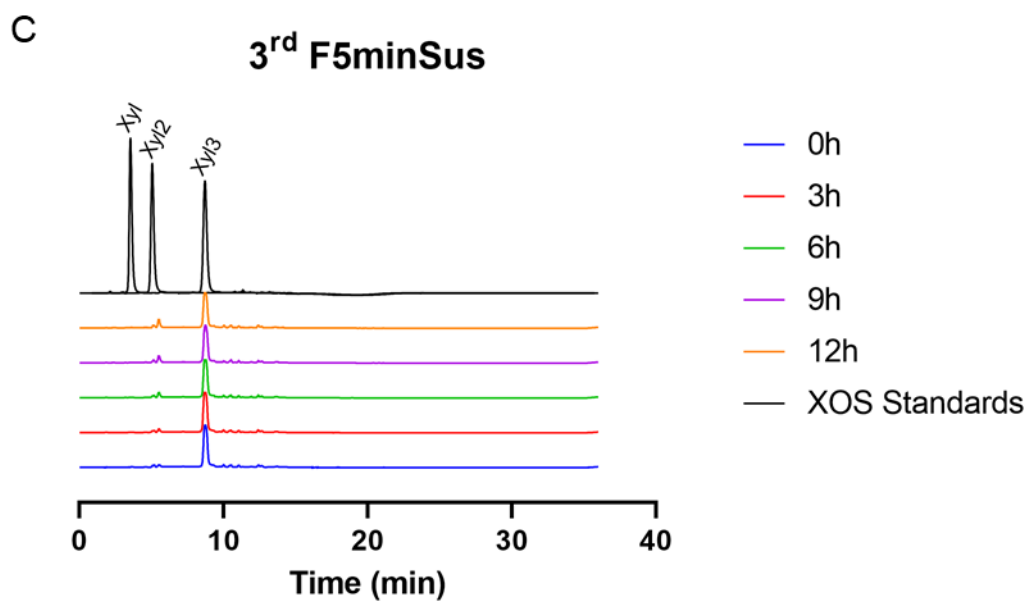
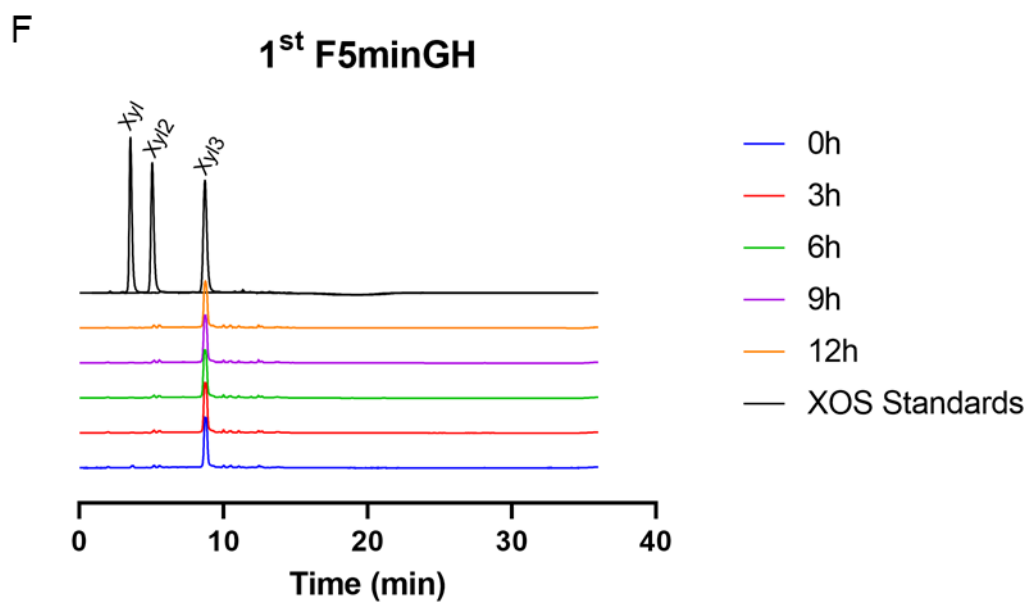
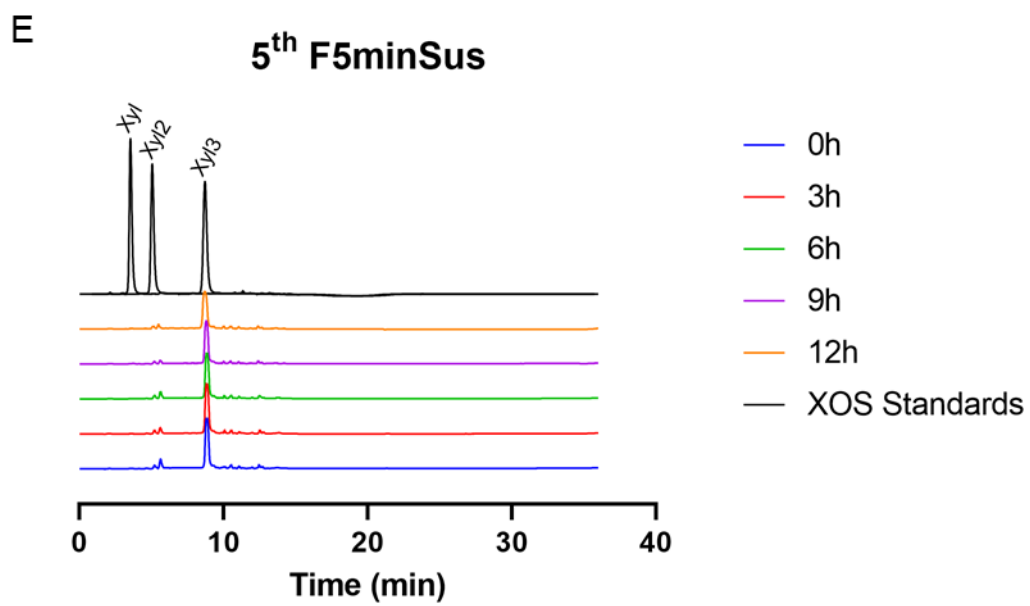


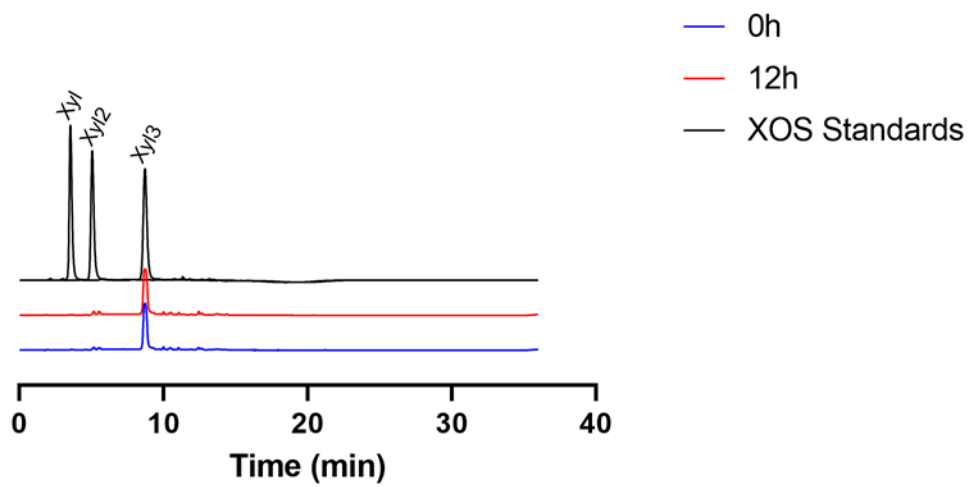
Figure 4: Xylotriose consumption by the five serial cultures of F5minSus (A) and F5minGH (b) growing on xylotriose.



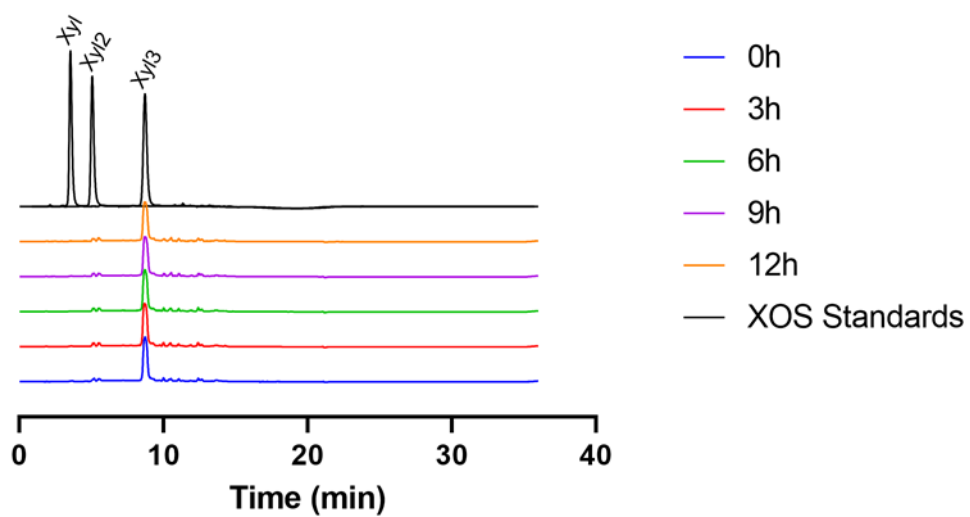




G

2nd F5minGH

H

3rd F5minGH

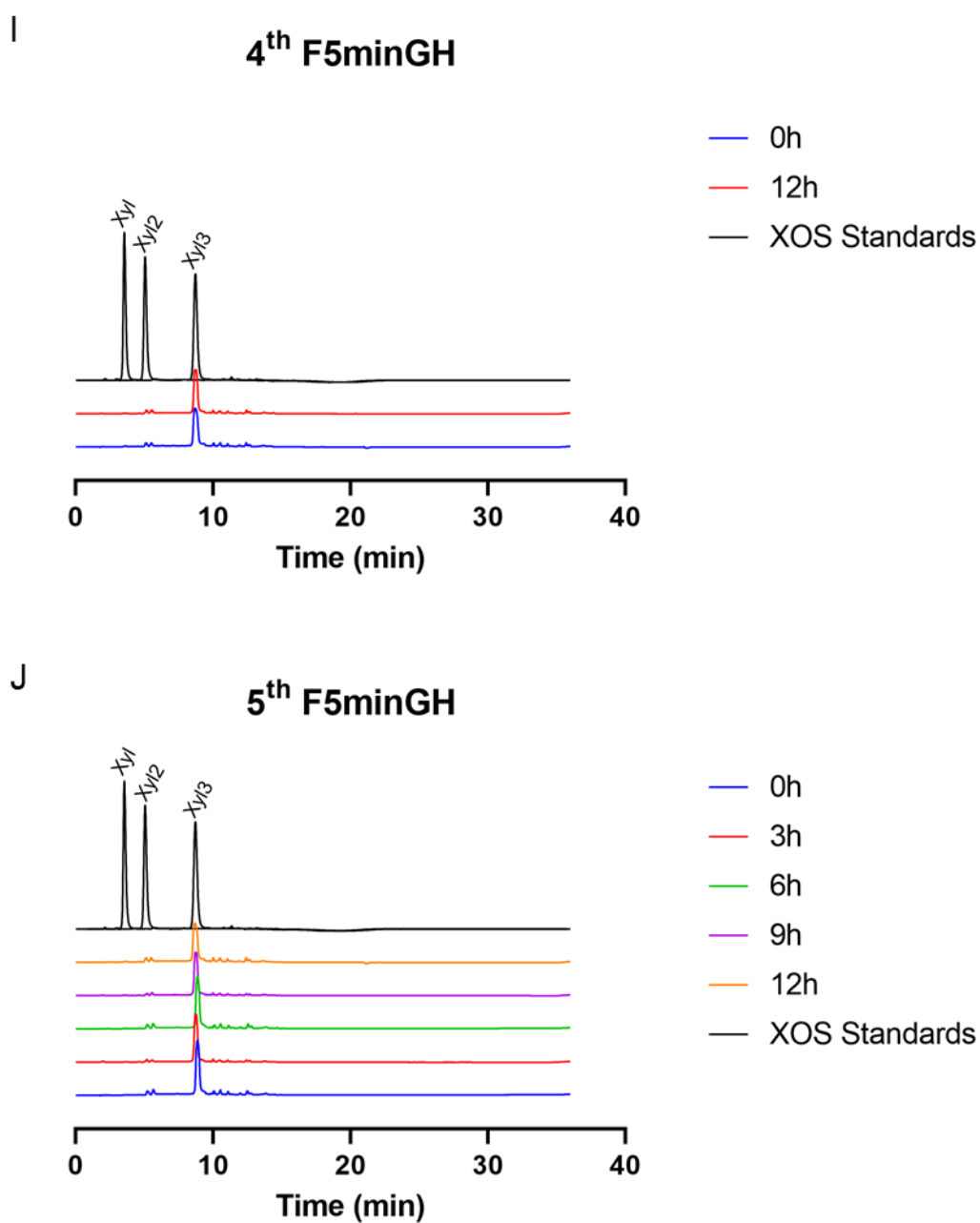


Figure 5: HPAEC-PAD chromatograms of the culture supernatants analyzed at different growth times.

A to E, F5minSus cultures; F to J, F5minGH cultures.

Identification of mutations using NGS

In order to determine if the differences in the growing and xylotriase consumption phenotypes of the 3rd and 5th samples (named hereafter S3 and S5) were correlated to mutations in the F5minSus locus, we sequenced these two samples using the NGS Ion Torrent S5 technology, which generates short reads of about 200 nt. The read sequences were aligned to the sequence of the initial F5minSus locus by using BWA and SamTools softwares. The sequence of the F5minSus locus was obtained in 2013 using the 454 GS FLX System pyrosequencing technology, with a sequencing coverage of minimum 45 (Cecchini et al. 2013). The GATK software was used to highlight the positions that were the most frequently mutated (named hot points) in the mixture of evolved strains contained in the S3 and S5 samples, based on a statistical analysis considering several criteria (coverage, read quality, distinction between sequencing errors & variant sequences). In the F5minSus locus, which sizes 10,416 nucleotides, we detected 62 hot points in the S3 reads, and 60 in the S5 reads (Table 1, Fig. 6). These mutations, which are exclusively deletions or insertions, are present at ratios varying between 9 and 41 % of the reads, which is significant. Moreover, 50 of these mutations are shared between the S3 and S5 samples. The remaining mutations, although not considered as prominent variants in the S3 or S5 samples at these positions, are found in between 6 and 16 % of the reads. For example, the C deletion at position 5542 was considered as a prominent variant in the S5 sample, but not in S3. Nevertheless, this deletion is found in 16 % of the S3 reads. Even more interestingly, 52 of the mutations are located (both for S3 and S5) in the GH, SusD or SusC encoding genes. Two other mutations are just before those genes, between 3 and 16 nucleotides before the start codon. Based on these data, we concluded that mutations in functional genes accumulated at hot points in the S3 and S5 samples.

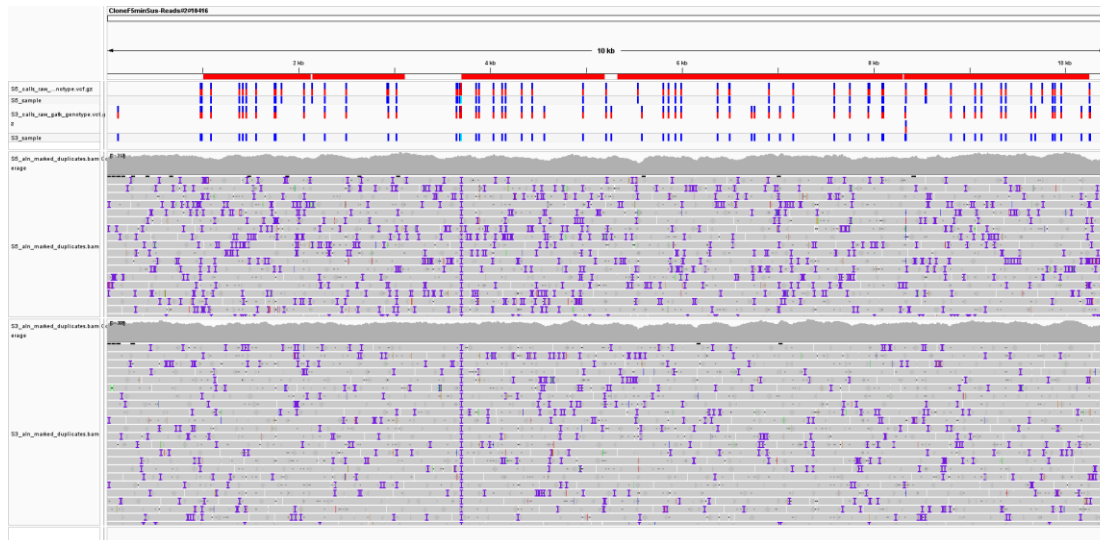


Figure 6: Positions of the mutations detected along the F5minSus locus by NGS read analysis.

Screenshot of the IGV analysis showing the gene locations (red) and the most prominent variants in the S3 and S5 samples. Each bar represents a significant variant at the position, red/blue for mutation contained in 9 to 41 % of the reads, all red bar for mutation contained in 100 % of the reads (position 3697) and insertions (multiple purple bars) in the reads (grey).

Table 1: Positions, type and relative abundance of the most prominent mutations detected along the F5minSus locus in the S3 and S5 samples.

The genes affected by the mutations or locations upstream the genes are indicated with the number of nucleotides upstream the start codon. In yellow, mutations detected only in the S3 or S5 samples.

S3					
Position	Gene	Reference	Mutation	% of reads containing the mutation	Repeats surrounding the indel
117		CG	C	15,69	C G GG
978		CA	C	12,58	C A AAAA
993		T	TA	12,93	T A AAAAA
1089	GH10	TC	T	29,37	GGGT C CTTC
1382	GH10	GA	G	10,10	G A AAAA
1418	GH10	TA	T	26,90	T A AAAAAA
1455	GH10	AG	A	14,79	AA A GGG
1561	GH10	AT	A	25,90	A T TTTT
1753	GH10	GC	G	12,14	GAAG C C
1763	GH10	TG	T	30,65	T G GCATTT
2057	GH10	CT	C	39,41	GGATTACCC T TTATT GTT
2276	GH43B	AT	A	31,31	A T TTT
2499	GH43B	AG	A	14,62	GAAG G GGCCG
2940	GH43B	GC	G	11,70	G C CTGTGGT
3023	GH43B	C	CG	34,50	CCCTTC G GGGGGG
3651		CG	C	14,29	TCC G GGAGA
3684	GH43A -16 nt	CG	C	26,52	GCC G GG
3697	GH43A -3 nt	C	CG	100,00	TCC G ATCCCA
3854	GH43A	AG	A	9,83	A G GGG
3883	GH43A	GC	G	17,53	TTTTTG C CCCCGGCC
4038	GH43A	CT	C	15,53	TTCT T TTTT
4124	GH43A	TC	T	41,09	CCAT C CGTG

4163	GH43A	TG	T	10,61	GGTGGGG
4331	GH43A	CG	C	19,44	CGGG
4438	GH43A	AC	A	32,46	AAACCCGGGAAGGG G
4568	GH43A	AG	A	12,33	AGGGGG
4971	GH43A	AG	A	17,61	AGGCGATGGGGAAA
5208		AT	A	36,75	TTTGATTTTTT
5266		TG	T	11,05	GGCCGGTGGGG
5583	SusC	GC	G	14,84	GCGCCCGCA
5809	SusC	TG	T	23,95	CCCCTTTATGGGTA ATTT
5863	SusC	AG	A	22,35	TTTTGAAGGA
5933	SusC	AG	A	21,82	AAAGGGG
5993	SusC	AT	A	11,11	ATTTTCCAAA
6376	SusC	AC	A	13,50	ACCCC
6503	SusC	CG	C	14,48	CTTCGGTAC
6727	SusC	GA	G	9,81	TTTTGGAAA
6762	SusC	GC	G	13,68	TCTGCCGAA
6913	SusC	TG	T	15,79	GGATGGG
7027	SusC	CG	C	12,86	TATCGGCAG
7161	SusC	AC	A	15,23	CAACCC
7593	SusC	GA	G	19,66	GAAAAA
7756	SusC	GC	G	19,13	GCTTGAACC
7945	SusC	AG	A	19,76	CAAGGGGGCCAAG
8095	SusC	TG	T	11,45	TGGG
8107	SusC	GC	G	9,50	AGAGCCTGC
8335	SusD	AT	A	20,83	ATTTT
8340	SusD	CA	C	23,04	CAAAAA
8808	SusD	TG	T	26,80	CCTGGGGTCC
8946	SusD	CG	C	10,29	CGGTGG
9063	SusD	AC	A	15,94	GGACCCAACGG
9132	SusD	AC	A	17,73	GAACCCTGG
9333	SusD	CG	C	31,03	TTCGGCTGA

9384	SusD	GC	G	12,50	GGCCG ^C CCTTGGAC CC
9649	SusD	GC	G	20,48	GTTG ^C CC
9688	SusD	AG	A	11,54	GGTAG ^G GTCG
9869	SusD	GC	G	10,81	TCTG ^C CCGTG
9896	SusD	GC	G	15,38	GTG ^C CCGTGCCGG
9962	SusD	GA	G	15,69	G ^A AATGA
10174	SusD	TC	T	12,38	^T CCTATCCC
10258		AT	A	17,48	A ^T TTTTT
10265		GT	G	11,11	G ^T TTTTT

S5					
Position	Gene	Reference	Mutation	% of reads containing the mutation	Repeats surrounding the indel
978		CA	C	11,29	C ^A AAAA
993		T	TA	13,82	T ^A AAAAA
1089	GH10	TC	T	26,87	GGGT ^C CTTC
1382	GH10	GA	G	19,31	G ^A AAAA
1418	GH10	TA	T	19,05	T ^A AAAAAA
1455	GH10	AG	A	15,57	AAAG ^G GG
1561	GH10	AT	A	35,86	A ^T TTTT
1753	GH10	GC	G	19,08	GAAG ^C C
1763	GH10	TG	T	30,37	T ^G GCATTT
1821	GH10	GC	G	9,80	TTG ^C CCAC
2057	GH10	CT	C	35,29	GGATTACC ^C TTAT TGT
2143	GH43B	GA	G	11,38	TGG ^A AAAGAAA
2276	GH43B	AT	A	23,56	A ^T TTTT
2499	GH43B	AG	A	14,05	GAAG ^G GGCCG
2925	GH43B	AG	A	10,67	AA ^G GGG

2940	GH43B	GC	G	10,90	G C CTGTGGT
3023	GH43B	C	CG	23,67	CCCTTC G GGGGGG
3651		CG	C	20,78	TCC G GGAGA
3663		AT	A	13,29	A T TTTT
3684	GH43A-16nt	CG	C	25,50	GCC G GG
3697	GH43A-3nt	C	CG	100,00	TCC G ATCCCA
3854	GH43A	AG	A	11,29	A G GGG
3883	GH43A	GC	G	20,00	TTTTTG C CCCGGCC
4038	GH43A	CT	C	13,22	TTTC T TTTT
4124	GH43A	TC	T	40,00	CCAT C CGTG
4163	GH43A	TG	T	10,76	GGT G GGG
4331	GH43A	CG	C	25,61	C GGG
4438	GH43A	AC	A	28,99	AA A CCGGGAAGG GG
4971	GH43A	AG	A	20,45	A G GGGG
5208		AT	A	34,71	A G GCGATGGGGAA A
5541	TonB	TC	T	19,72	TTTGAT T TTTTT
5809	TonB	TG	T	26,95	CCCCTTTAT G GGTA ATTT
5863	TonB	AG	A	18,40	TTTTGA A GGA
5933	TonB	AG	A	35,77	AA A GGGG
5993	TonB	AT	A	23,08	A T TTTTCCAAA
6376	TonB	AC	A	22,31	A C CCC
6492	TonB	AC	A	14,96	AAG A CCGAA
6503	TonB	CG	C	17,50	CTTC G GTAC
6913	TonB	TG	T	19,66	GGAT G GG
7161	TonB	AC	A	16,39	CA A CCC
7593	TonB	GA	G	14,63	G A AAAA
7756	TonB	GC	G	17,53	G C TTGGAACC
7945	TonB	AG	A	24,64	CA A GGGGGCCAAG
7957	TonB	AC	A	10,95	A C CCGAAA
8095	TonB	TG	T	11,27	T G GG

8107	TonB	GC	G	9,33	AGAGCCTGC
8340	SusD	CA	C	32,61	CAAAAA
8543	SusD	GGA	G	11,59	GGGAACC
8553	SusD	AG	A	9,86	AGGCAACC
8808	SusD	TG	T	26,39	CCTGGGGTCC
9063	SusD	AC	A	17,19	GGACCCAACGG
9132	SusD	AC	A	15,54	GAACCCTGG
9333	SusD	CG	C	30,68	TTCGGCTGA
9384	SusD	GC	G	10,61	GGCCGCCCTTGGAC CC
9649	SusD	GC	G	20,13	GTTGCC
9760	SusD	GC	G	11,54	GCCCAA
9869	SusD	GC	G	9,64	TCTGCCCGTG
9896	SusD	GC	G	9,55	GTGCCCGTGCCGG
9962	SusD	GA	G	14,35	GAAATGA
10258		AT	A	15,18	ATTTTTT

Discussion

In this chapter, we explored the potential of ALE in order to engineer a recombinant PUL conferring to *E. coli* the ability to grow on xylotriose. In the past decades, a series of cases have been documented in which mutations in transporter encoding genes or in their regulator elements have been induced by ALE, used for metabolic engineering of various microorganisms. For example, Pereira *et al.* used ALE to increase the tolerance of *Saccharomyces cerevisiae* to aromatic (coumaric and ferulic) (Pereira et al. 2020) and dicarboxylic acids (glutaric, adipic, and pimelic) (Pereira et al. 2019). The most significant adaptations involved the upregulation of transporter encoding genes, including a monocarboxylate permease and multidrug resistance transporters, which enhanced efflux and reduced intracellular acid concentrations. In bacteria, ALE has been used to identify and engineer *E. coli* amino acid transporters (Radi et al. 2022), to improve isoprenol production in *E. coli* co-cultures (Kawai et al. 2022), and to increase methanol tolerance in *Corynebacterium glutamicum* (Wang et al. 2020). In these three studies were identified mutations in transporters that enhanced the efficiency of the targeted bioconversion processes. Regarding specifically carbohydrate transporters, let's cite the study of Alva *et al.* who performed a series of ALE experiments in *E. coli* to restore ability of PTS⁻ mutants to efficiently grow on glucose (Alva et al. 2020), or that of Sarkar et al. who evolved *Zymomonas mobilis* to enhance its ability to co-utilize xylose and glucose, an important trait for the production of lignocellulosic biofuels (Sarkar et al. 2020).

However, as mentioned in the introduction of this chapter, to our knowledge, no study has been performed to date to evolve recombinant oligosaccharide transporters produced in *E. coli*. To test the potential of ALE to engineer recombinant PULs expressed in *E. coli*, we chose as model the F5minSus locus cloned into the pEPIFOS-5 fosmid, in which five functional genes involved in xylotriose transport and hydrolysis to xylose were previously shown to be successfully expressed in *E. coli* (Tauzin et al. 2016). We chose this minimal construct of the F5_XOS_PUL since the ability of the F5minSus strain to grow on XOS was significantly reduced compared to that of the F5 (containing the full-length F5_XOS_PUL) and F5minMFS strains. The potential of improvement of the F5minSus strain for XOS utilization was thus much higher than that of the F5minMFS strain, which already presents a very good, if not maximal, metabolic efficiency.

In our experimental design, we chose to use highly purified commercial xylotriase as substrate, since the F5_SusC/D transporter was previously shown to internalize this oligosaccharide (Tauzin et al. 2016). In addition, despite its high cost, this substrate presents the advantage to contain less than 0.24 % of xylose, while XOS commercial mixtures have to be purified using preparative chromatography to remove xylose (Cecchini et al. 2013). Since xylose is efficiently naturally metabolized by *E. coli*, a high amount of xylose in the substrate used for our ALE experiment would have highly biased the evolution process, or even prevented any evolution since *E. coli* would have already presented a maximal growth efficiency on xylose. Using an expensive substrate such as xylotriase nevertheless constrained us to not replicate the ALE cultures. However, this preliminary study was sufficient to observe significant changes in the growth phenotype between the F5minSus serial cultures. Indeed, in 6 hours of culture, the OD_{600nm} of the 3rd F5minSus sample reached a value 2 times higher than that of the original strain. This growth ability improvement was consistent with that in xylotriase utilization rate, which reached an optimum in the 3rd culture, after 12 hours of culture. In our experimental setup, this culture duration was chosen to ensure that the cells could grow to the exponential phase, which is required for serial transfers. However, using such short cultures, we could not, at this stage of study, assess the maximal growth ability of the samples. Characterization of the rate of biomass increase until the stationary phase will be performed for the monoclonal evolved strains that will be selected for full-length genome and fosmid sequencing, as described further in this discussion section.

The improvements in growth and xylotriase utilization rates observed for the 2nd and 3rd cultures tended to be lost during the 4th and 5th cultures. A possible hypothesis regarding the observed decline in growth and xylotriase utilization rates is that multiple mutations have been generated to affect genes encoding the F5_SusC/D transporter. The function of the F5_SusC/D transporter or expression of its encoding genes might be negatively affected due to insertions/deletions. ALE of *Thermotoga maritima* revealed mutations in evolved cultures related to glucose-responsive ABC transporters (Latif et al. 2015). The gene expression of the glucose ABC transporter GluEFK was enhanced. Conversely, BglEFGKL, a transporter of β -glucosides, is substantially downregulated due to a nonsense mutation to the solute binding protein or due to a deletion of the upstream promoter. Therefore, improved growth of glucose is achieved through mutations favoring GluEFK expression over BglEFGKL. Radi *et al.* used ALE to identify membrane transporters for a set of four amino acids as well as specific mutations that modulate the activities of these transporters. In particular, *E. coli* was adaptively evolved under

increasing concentrations of the L-threonine revealing key mutations affecting inner membrane amino acid transporter *brnQ* (Radi et al. 2022). Four different intragenic mutations involving *brnQ* were independently detected across all different replicates of the threonine ALE, these mutations were examined to result in a loss of function due to the generation of premature stop codons. The loss of function of *brnQ* may imply a role in the import of threonine, and its inactivation could reduce the intracellular levels of threonine.

Since significant phenotypic changes were observed between the initial, the 3rd and the 5th cultures, we sequenced the 3rd and 5th samples. NGS was used to confirm that, in accordance with our initial hypothesis and the observed phenotypic changes, the 3rd and 5th samples represent mixtures of several evolved strains. Indeed, read analysis and comparison to the sequence of the initial strain highlighted 72 positions at which mutations occurred in significant ratios, varying between 9 and 40 % of the reads.

We estimated that at the end of the 3rd cultures, we reached 9 generations of *E. coli*, with a doubling time of around 271 min, which is very low. The number of generations reached in our study is much lower than that of a typical ALE experiment, which lasts between 100 and 2000 generations, according to a summary of literature on ALE in bacteria and yeasts (Dragosits and Mattanovich 2013). However, the xylotriase utilization and rates of biomass increase could not be improved after 3 cultures. Besides, since our 3rd sample contains 62 prominent mutations, the mutation rate was of 6.9, meaning that one mutation occurred every 0.15 generations. In other terms, we obtained 62 mutations in 3 cultures of 12 hours, that is 1,72 mutation/hour. Without any selection pressure, the *E. coli* mutation rate is of 0.0022 mutation/hour. The mutation rate could increase by more than 3 orders of magnitude for certain mutants forced to survive, reaching 5.6 mutations/hour (Robert et al. 2018), a value in the same range as the one obtained in the present study. Besides, in our study, only deletions or insertions of bases (indels) were identified as prominent mutations. This was surprising, since in *E. coli*, these events occur at only 1/10th the rate of base-pair substitutions, although the rate of small (≤ 4 nt) indels is high at repeat sequences (Lee et al. 2012). We thus checked if repeat sequences were frequently found at the mutation positions in our targeted locus. It is indeed the case, since 71 % and 38 % of the indels are surrounded by (and/or included in) at least three or four repeats, respectively.

Furthermore, the Ion Torrent S5 technology is known to generate indels as sequencing errors in repeated regions. Nevertheless, if those indels had been due to sequencing errors, we would have identified many prominent mutations in the intergenic regions of the F5minSus locus,

which is not the case, since 52 of the prominent 62 mutations (84 %) are found in the functional genes. However, 115 repeats of more than 3 successive identical nucleotides were counted in the intergenic regions totalizing 1911 nt, which represents one repeat every 0.06 nt, versus 386 repeats in the functional genes totalizing 8505 nt, which represents one repeat every 0.05 nt. Anyway, the high occurrence of indels in the S3 and S5 reads questions us as to the functionality of the encoded F5_GHs and SusC/D proteins required to metabolize xylotriose. The Ion Torrent S5 technology generating short reads of about 200 nucleotides, it was not possible to deduce from our sequencing data what are the mutations accumulated by each individual strain contained in the S3 and S5 samples, and to deduce the amino acid sequences of the encoded GH and SusC/D proteins.

Finally, by calculating the mutation rate, we focused only on the locus cloned into the pEpifos5 fosmid. Thus, we did not consider the mutations which potentially occurred in the *E. coli* genome, since we only sequenced the pools of fosmids extracted from the 3rd and 5th cultures. One cannot exclude that during ALE, mutations occurred in native *E. coli* transporters, allowing internalization of xylotriose. Indeed, there are examples in literature, in which ALE was used to restore fast growth of mutant *E. coli* strains constructed to be strongly affected in their growth ability. In the study of Alva *et al.*, PTS⁻ derivative mutants exhibiting fast growth on glucose were selected (Alva *et al.* 2020). They contain mutations resulting in the derepression of a gene encoding an alternative glucose transporter. Also, Choe *et al.* used ALE to re-optimize the growth performance of a genome-reduced strain (Choe *et al.* 2019). The best variant contained a large deletion and 117 other mutations including single-nucleotide variations, multi-nucleotide variations as well as multiple indels. However, in the present study, mutations in the *E. coli* genome are unlikely the only determinants of the increase in xylotriose utilization observed during ALE. Indeed, ALE did not allow the control F5minGH strain to evolve to acquire ability to grow on xylotriose, although the F5 glycoside hydrolases that are able to breakdown xylotriose to xylose were produced in the initial strain. We thus conclude that i) the F5_SusC/D transporter was functional in the evolved strains of the S3 and S5 samples that were able to grow on xylotriose, ii) the mutations detected in the F5minSus locus and/or mutations in the *E. coli* genome are responsible for the increase of xylotriose utilization. Sequencing of both the fosmid and genomic DNA of monoclonal evolved strains will be required in order to conclude on the potential presence of mutations in the *E. coli* genome and on their impact on the observed phenotypes.

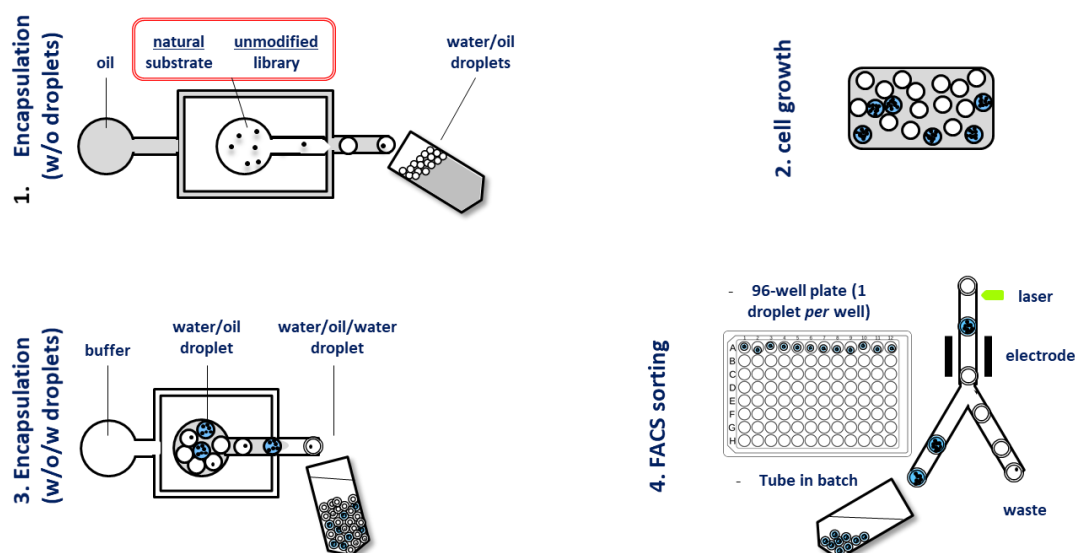


Figure 7: Cell growth screening in droplet-microfluidics, based on accumulation of fluorescent flavins in pico-liter droplets during cell growth (patent EP22305759.7).

To this aim, complementary analyses should be performed to characterize the S3 and S5 samples and to finalize this study. The next experiments will first consist in the assessment of the functional diversity, based on growth ability on xylotriose, of the strains contained in the S3 and S5 samples. To do so, we will use a recent microfluidics-based technology patented by the CIMEs team in 2022 (Potocki-Veronese et al. 2023), to quantify the growth ability of each of the strains of the S3 and S5 samples, after single cell encapsulation and selective growth on xylotriose in pico-liter droplets. The principle of this technology is based on the concentration in droplets of fluorescent compounds (flavins) naturally produced by *E. coli* during cell growth on a selective medium. This microfluidics- based workflow (Fig. 7) thus does not require any fluorogenic or chromogenic substrate, neither fluorescent strains to quantify cell growth, which is linearly correlated to droplet fluorescence. Cell growth can be quantified for each droplet, thanks to FACS (fluorescence-activated-cell-sorting) sorting of each fluorescent droplet in individual wells of micro-plates (Fig. 8A). In a second step of the characterization workflow (Fig. 8B), we will sequence the fosmidic and genomic DNA of 10 monoclonal strains that will be sorted based on their high efficiency to grow on xylotriose. At this stage, we will use a long-read sequencing technology such as the MinIon (Nanopore), which generates reads of about 10,000 nucleotides, to directly access the sequence of the evolved F5minSus loci, of which the size is 10,416 nucleotides. This workflow will allow us, ultimately, to establish the relationships between mutations in the F5minSus locus, the structure of the encoded GHs and SusC/D

proteins and their performance for xylotriase hydrolysis and internalization in *E. coli*, respectively.

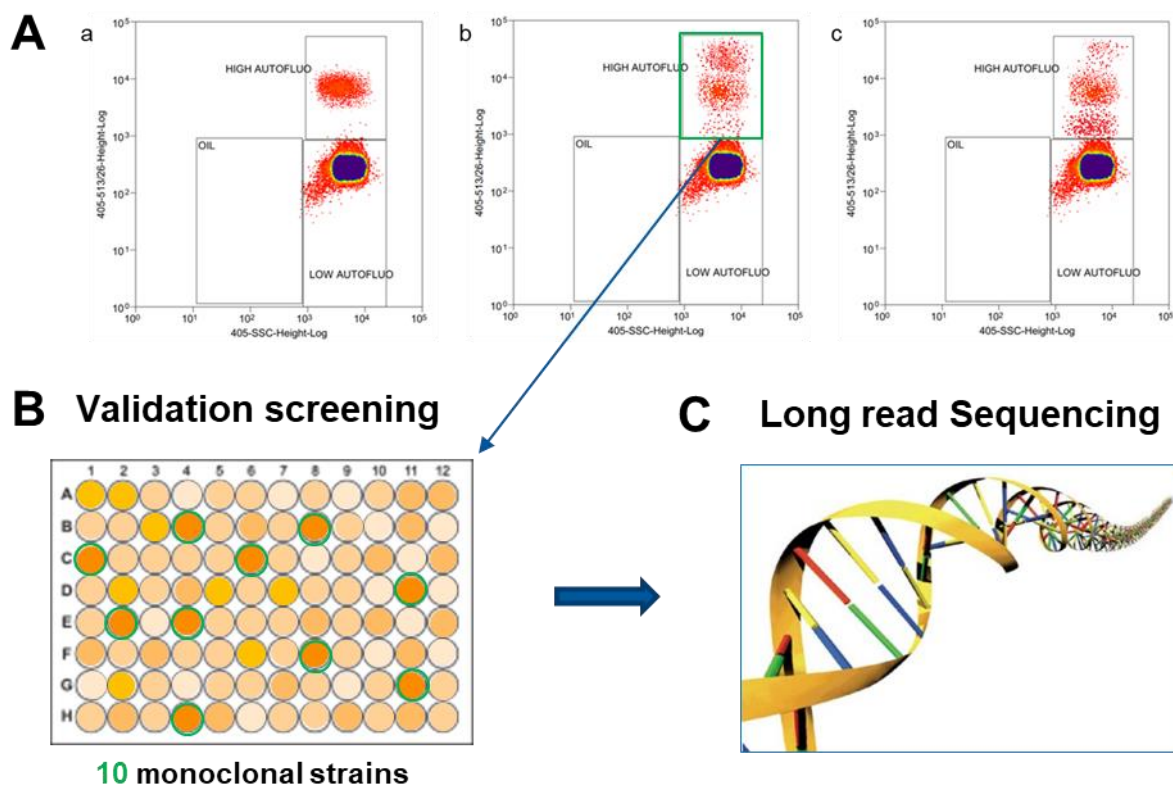


Figure 8: Experimental workflow that will be used to characterize the evolved strains.

A, schematic representation of the theoretical FACS plots that might be obtained after single cell encapsulation and growth on xylotriase of the F5minSus initial strain (a), S3 (b) and S5 (c) samples. The ‘high autofluorescence’ squares represent the droplets in which cell growth occurred. The level of fluorescence is correlated to the level of cell growth in the droplets. The ‘low autofluorescence’ squares represent the empty droplets or those in which the encapsulated cell did not replicate. B, Validation of the growth potential on xylotriase of monoclonal strains issued from the S3 sample, in micro-plates. C, sequencing of the 10 clones with the highest growth ability on xylotriase.

Materials and Methods

Cloning

The minimal variants of clone F5 were constructed by Tauzin *et al.* using the In-Fusion HD Plus cloning kit (Clontech) following the instructions in the user manual previously. Briefly, the In-Fusion Enzyme provided by this kit fuses DNA fragments (e.g., PCR-generated inserts and linearized vectors) efficiently and precisely by recognizing 15-bp overlaps at their ends. These overlaps can be engineered by designing primers to amplify the desired sequences. Therefore, the F5 truncated gene was cloned into the pEPIFOS-5 fosmid by this kit, and then transformed into Stellar competent cells. F5minSus is a truncated variant containing the glycoside hydrolases (GHs) and the SusC/D homologs. F5minGH, used as a negative control, is a truncated variant only containing the three GHs genes (Fig. 1).

Culture medium

The variants F5minSus and F5minGH were grown on minimal medium (M9) supplemented with 5 g/l xylotriase and 12.5 mg/l chloramphenicol. The M9 medium contained Na₂HPO₄·12H₂O, 17.4 g/l; KH₂PO₄, 3.03 g/l; NaCl, 0.51 g/l; NH₄Cl, 2.04 g/l; MgSO₄, 0.49 g/l; CaCl₂, 4.38 mg/l; Na₂EDTA·2H₂O, 15 mg/l; ZnSO₄·7H₂O, 4.5 mg/l; CoCl₂·6H₂O, 0.3 mg/l; MnCl₂·4H₂O, 1 mg/l; H₃BO₃, 1 mg/l; Na₂MoO₄·2H₂O, 0.4 mg/l; FeSO₄·7H₂O, 3 mg/l; CuSO₄·5H₂O, 0.3 mg/l; thiamine, 0.1 g/l; and leucine, 0.02 g/l.

Growth conditions

First, the pEPIFOS-5 fosmid containing truncated F5minSus and F5minGH were transformed into stellar competent cells by heat shock transformation method. After a 30-minutes incubation in ice, a mixture of chemically competent bacteria and fosmid was placed at 42°C for 45 seconds (heat shock) and then placed back in ice about 5 minutes. SOC media was added and the transformed cells are incubated at 37°C for 1 hour with orbital shaking at 200 rpm. Transformed bacteria were diluted and plated on LB agar plates supplemented by 12.5 mg/l chloramphenicol.

Then, 5 ml precultures of M9 minimal medium supplemented by 5 g/l of xylose and 12.5 mg/l of chloramphenicol were inoculated using isolated colonies of the F5minSus and F5minGH strains. For the ALE cultures (Fig. 2), Erlenmeyer flasks containing 10 ml of M9 minimal medium supplemented by 5 g/l of xylotriase as the sole carbon source and 12.5 mg/l of

chloramphenicol were inoculated from the precultures at an initial OD_{600nm} of 0.1. Cultures were performed during 12 hours at 37°C, using orbital shaking at 200 rpm. In total, 5 cultures were performed by serial transfers, using the preceding culture for inoculation at an initial OD_{600nm} of 0.1. Aliquots of samples were taken every 3 hours to follow cell growth by measuring the OD_{600nm}. The aliquots were stored at -20°C for further HPAEC-PAD analysis and pH measurement using pH test strips (Dominique DUTSCHER SAS). The culture samples retrieved at 12 hours were also stored in 20 % glycerol at -80°C.

The number of generations (N) and the doubling time (G) were calculated based on the *E. coli* cell numbers. Based on an *E. coli* cell counting experiment performed by Sophie Lajus (member of the CIMEs team), we considered that 1 unit of OD_{600nm} corresponds to 2.23×10^9 cells/mL of culture. The number of generations N and the doubling times G were calculated according to the following formula: $N = (\lg X_2 - \lg X_1) / \lg 2 = (\lg X_2 - \lg X_1) * 3.322$ and $G = \Delta t / N = (t_2 - t_1) * \lg 2 / (\lg X_2 - \lg X_1)$, with X1 and X2 being the cell numbers at times t2 and t1.

Xylooligosaccharide quantification

To analyze the uptake of xylotriose during the different cultures, we quantified the concentration of residual xylotriose, xylobiose and xylose in the culture media. After culture sample centrifugation, the supernatants were filtered at 0.20 µm and analyzed by HPAEC-PAD on a Dionex ICS-3000 system using a CarboPac PA100 4 250 column. Carbohydrates were eluted at 30°C, at a flow rate of 1 ml/min with the following multistep gradient: 0-30 min (0-60 % eluent B), 30-31 min (60-0 % eluent B) and 31-36 min (0 % eluent B). Solvents were 150 mM NaOH (eluent A) and 150 mM NaOH, 500 mM CH₃COONa (eluent B).

DNA extraction

10 µl of the S3 and S5 samples stored at -80°C were used to inoculate 10 ml flasks of LB medium supplemented with 12.5 mg/l of chloramphenicol. Cells were harvested after 24-hours of culture (OD_{600nm} 2 to 2.5) through centrifuging at 4°C 5000 g for 15 min. The fosmids were extracted using the Machery-Nagel NucleoBond 96 Xtra EF Kit following the instructions in the user manual. Quantity and quality of the extracted DNA was analyzed using the Thermo Scientific™ Nanodrop 2000. The concentration of the S3 fosmid DNA mix was 59.0 ng/µl, that of the S5 fosmid DNA mix was 80.7 ng/µl.

Sequencing and bioinformatic analysis

Fosmid DNA samples were sequenced by the GeT-Biopuces platform (<https://get.genotoul.fr/>, <https://get-biopuces.insa-toulouse.fr/>, Genotoul, Toulouse, France) using the NGS Ion Torrent S5 technology.

The fosmid samples were assayed with Qubit (Qubit™ dsDNA HS Assay Kit). The libraries were created from 10 ng of DNA using the NEBNext Fast DNA Fragmentation and Library prep and Ion Xpress™ Barcode Adapters according to the manufacturer's recommendations (NEB and ThermoFisher). The libraries obtained were quantified using the Qubit™ dsDNA HS Assay Kit and the bioanalyzer 2100 (Agilent-High Sensitivity DNA Assay). 5 pg/μl of each library were pooled together, amplified on beads by emulsion PCR and enriched on the IonChef System (ThermoFisher) using the Ion 540 kit. The template was sequenced (200 bp oriented single end) on the Ion Torrent S5 using an Ion 540 chip according to the manufacturer's instructions.

Raw sequences in fastq format were aligned on the reference sequence by using bwa-mem algorithm from bwa software (v0.7.17). The Samtools software (v1.12-34) was used to sort and index the alignment file. Then a bam file per sample was obtained, that was used for the variant calling with GATK suite (v4.3.0.0). We used first the MarkDuplicates module to identify duplicate reads and then HaplotypeCaller and GenotypeGVCFs to identify positions with significant variants. Final vcf files were compiled with alignment files and the reference sequence to visualize them with the IGV software.

References

- Alva, Alma, Andrea Sabido-Ramos, Adelfo Escalante, and Francisco Bolívar. 2020. 'New Insights into Transport Capability of Sugars and Its Impact on Growth from Novel Mutants of Escherichia Coli'. *Applied Microbiology and Biotechnology* 104 (4): 1463–79. <https://doi.org/10.1007/s00253-019-10335-x>.
- Andersen, Joakim M., Rodolphe Barrangou, Maher Abou Hachem, Sampo J. Lahtinen, Yong Jun Goh, Birte Svensson, and Todd R. Klaenhammer. 2013. 'Transcriptional Analysis of Oligosaccharide Utilization by Bifidobacterium lactis B1-04'. *BMC Genomics* 14 (1): 312. <https://doi.org/10.1186/1471-2164-14-312>.
- Cakar, Z, U Seker, C Tamerler, M Sonderegger, and U Sauer. 2005. 'Evolutionary Engineering of Multiple-Stress Resistant'. *FEMS Yeast Research* 5 (6–7): 569–78. <https://doi.org/10.1016/j.femsyr.2004.10.010>.
- Cecchini, Davide A., Elisabeth Laville, Sandrine Laguerre, Patrick Robe, Marion Leclerc, Joël Doré, Bernard Henrissat, Magali Remaud-Siméon, Pierre Monsan, and Gabrielle Potocki-Véronèse. 2013. 'Functional Metagenomics Reveals Novel Pathways of Prebiotic Breakdown by Human Gut Bacteria'. *PLOS ONE* 8 (9): e72766. <https://doi.org/10.1371/journal.pone.0072766>.
- Choe, Donghui, Jun Hyoung Lee, Minseob Yoo, Soonkyu Hwang, Bong Hyun Sung, Suhjung Cho, Bernhard Palsson, Sun Chang Kim, and Byung-Kwan Cho. 2019. 'Adaptive Laboratory Evolution of a Genome-Reduced Escherichia Coli'. *Nature Communications* 10 (1): 935. <https://doi.org/10.1038/s41467-019-08888-6>.
- Dragosits, Martin, and Diethard Mattanovich. 2013. 'Adaptive Laboratory Evolution – Principles and Applications for Biotechnology'. *Microbial Cell Factories* 12 (1): 64. <https://doi.org/10.1186/1475-2859-12-64>.
- Farwick, Alexander, Stefan Bruder, Virginia Schadeweg, Mislav Oreb, and Eckhard Boles. 2014. 'Engineering of Yeast Hexose Transporters to Transport D-Xylose without Inhibition by d-Glucose'. *Proceedings of the National Academy of Sciences* 111 (14): 5159–64. <https://doi.org/10.1073/pnas.1323464111>.
- Fong, Stephen S., Anthony P. Burgard, Christopher D. Herring, Eric M. Knight, Frederick R. Blattner, Costas D. Maranas, and Bernhard O. Palsson. 2005. 'In Silico Design and

- Adaptive Evolution of Escherichia Coli for Production of Lactic Acid'. *Biotechnology and Bioengineering* 91 (5): 643–48. <https://doi.org/10.1002/bit.20542>.
- Guirimand, Gregory G. Y., Takahiro Bamba, Mami Matsuda, Kentaro Inokuma, Kenta Morita, Yuki Kitada, Yuma Kobayashi, et al. 2019. 'Combined Cell Surface Display of β -d-Glucosidase (BGL), Maltose Transporter (MAL11), and Overexpression of Cytosolic Xylose Reductase (XR) in *Saccharomyces Cerevisiae* Enhance Cellobiose/Xylose Coutilization for Xylitol Bioproduction from Lignocellulosic Biomass'. *Biotechnology Journal* 14 (9): 1800704. <https://doi.org/10.1002/biot.201800704>.
- Ha, Suk-Jin, Heejin Kim, Yuping Lin, Myoung-Uoon Jang, Jonathan M. Galazka, Tae-Jip Kim, Jamie H. D. Cate, and Yong-Su Jin. 2013. 'Single Amino Acid Substitutions in HXT2.4 from *Scheffersomyces Stipitis* Lead to Improved Cellobiose Fermentation by Engineered *Saccharomyces Cerevisiae*'. *Applied and Environmental Microbiology* 79 (5): 1500–1507. <https://doi.org/10.1128/AEM.03253-12>.
- Kawai, Ryutaro, Yoshihiro Toya, Kenta Miyoshi, Manami Murakami, Teppei Niide, Takaaki Horinouchi, Tomoya Maeda, Atsushi Shibai, Chikara Furusawa, and Hiroshi Shimizu. 2022. 'Acceleration of Target Production in Co-Culture by Enhancing Intermediate Consumption through Adaptive Laboratory Evolution'. *Biotechnology and Bioengineering* 119 (3): 936–45. <https://doi.org/10.1002/bit.28007>.
- Kim, Heejin, Eun Joong Oh, Stephan Thomas Lane, Won-Heong Lee, Jamie H.D. Cate, and Yong-Su Jin. 2018. 'Enhanced Cellobiose Fermentation by Engineered *Saccharomyces Cerevisiae* Expressing a Mutant Cellodextrin Facilitator and Cellobiose Phosphorylase'. *Journal of Biotechnology* 275 (June):53–59. <https://doi.org/10.1016/j.jbiotec.2018.04.008>.
- Kuyper, Marko, Maurice J. Toirkens, Jasper A. Diderich, Aaron A. Winkler, Johannes P. Van Dijken, and Jack T. Pronk. 2005. 'Evolutionary Engineering of Mixed-Sugar Utilization by a Xylose-Fermenting *Saccharomyces Cerevisiae* Strain'. *FEMS Yeast Research* 5 (10): 925–34. <https://doi.org/10.1016/j.femsyr.2005.04.004>.
- Latif, Haythem, Merve Sahin, Janna Tarasova, Yekaterina Tarasova, Vasiliy A. Portnoy, Juan Nogales, and Karsten Zengler. 2015. 'Adaptive Evolution of *Thermotoga Maritima* Reveals Plasticity of the ABC Transporter Network'. *Applied and Environmental Microbiology* 81 (16): 5477–85. <https://doi.org/10.1128/AEM.01365-15>.

- Lee, Heewook, Ellen Popodi, Haixu Tang, and Patricia L. Foster. 2012. ‘Rate and Molecular Spectrum of Spontaneous Mutations in the Bacterium *Escherichia Coli* as Determined by Whole-Genome Sequencing’. *Proceedings of the National Academy of Sciences of the United States of America* 109 (41): E2774–83. <https://doi.org/10.1073/pnas.1210309109>.
- Meynial Salles, Isabelle, Nynne Forchhammer, Christian Croux, Laurence Girbal, and Philippe Soucaille. 2007. ‘Evolution of a *Saccharomyces Cerevisiae* Metabolic Pathway in *Escherichia Coli*’. *Metabolic Engineering* 9 (2): 152–59. <https://doi.org/10.1016/j.ymben.2006.09.002>.
- Mohamed, Elsayed T., Hemanshu Mundhada, Jenny Landberg, Isaac Cann, Roderick I. Mackie, Alex Toftgaard Nielsen, Markus J. Herrgård, and Adam M. Feist. 2019. ‘Generation of an E. Coli Platform Strain for Improved Sucrose Utilization Using Adaptive Laboratory Evolution’. *Microbial Cell Factories* 18 (1): 116. <https://doi.org/10.1186/s12934-019-1165-2>.
- Pereira, Rui, Elsayed T. Mohamed, Mohammad S. Radi, Markus J. Herrgård, Adam M. Feist, Jens Nielsen, and Yun Chen. 2020. ‘Elucidating Aromatic Acid Tolerance at Low pH in *Saccharomyces Cerevisiae* Using Adaptive Laboratory Evolution’. *Proceedings of the National Academy of Sciences* 117 (45): 27954–61. <https://doi.org/10.1073/pnas.2013044117>.
- Pereira, Rui, Yongjun Wei, Elsayed Mohamed, Mohammad Radi, Carl Malina, Markus J. Herrgård, Adam M. Feist, Jens Nielsen, and Yun Chen. 2019. ‘Adaptive Laboratory Evolution of Tolerance to Dicarboxylic Acids in *Saccharomyces Cerevisiae*’. *Metabolic Engineering* 56 (December):130–41. <https://doi.org/10.1016/j.ymben.2019.09.008>.
- Portnoy, Vasilij A, Daniela Bezdan, and Karsten Zengler. 2011. ‘Adaptive Laboratory Evolution—Harnessing the Power of Biology for Metabolic Engineering’. *Current Opinion in Biotechnology, Nanobiotechnology and Systems Biology*, 22 (4): 590–94. <https://doi.org/10.1016/j.copbio.2011.03.007>.
- Potocki-Veronese, Gabrielle, Sophie LAJUS, Adilya DAGKESAMANSKAYA, Amandine DEROITE, and Delphine LESTRADE. 2023. New milli-to-micro-fluidics-based process to screen microbial growth in droplets. World Intellectual Property Organization WO2023222765A1, filed 17 May 2023, and issued 23 November 2023.

[https://patents.google.com/patent/WO2023222765A1/fr?q=\(lajus+potocki\)&oq=lajus+potocki](https://patents.google.com/patent/WO2023222765A1/fr?q=(lajus+potocki)&oq=lajus+potocki).

- Radi, Mohammad S., Jesus E. SalcedoSora, Se Hyeuk Kim, Suresh Sudarsan, Anand V. Sastry, Douglas B. Kell, Markus J. Herrgård, and Adam M. Feist. 2022. ‘Membrane Transporter Identification and Modulation via Adaptive Laboratory Evolution’. *Metabolic Engineering* 72 (July):376–90. <https://doi.org/10.1016/j.ymben.2022.05.004>.
- Robert, Lydia, Jean Ollion, Jérôme Robert, Xiaohu Song, Ivan Matic, and Marina Elez. 2018. ‘Mutation Dynamics and Fitness Effects Followed in Single Cells’. *Science* 359 (6381): 1283. <https://doi.org/10.1126/science.aan0797>.
- Sarkar, Payel, Mayurketan Mukherjee, Gargi Goswami, and Debasish Das. 2020. ‘Adaptive Laboratory Evolution Induced Novel Mutations in *Zymomonas Mobilis* ATCC ZW658: A Potential Platform for Co-Utilization of Glucose and Xylose’. *Journal of Industrial Microbiology and Biotechnology* 47 (3): 329–41. <https://doi.org/10.1007/s10295-020-02270-y>.
- Sonderegger, Marco, and Uwe Sauer. 2003. ‘Evolutionary Engineering of *Saccharomyces Cerevisiae* for Anaerobic Growth on Xylose’. *Applied and Environmental Microbiology* 69 (4): 1990–98. <https://doi.org/10.1128/AEM.69.4.1990-1998.2003>.
- Tauzin, Alexandra S., Elisabeth Laville, Yao Xiao, Sébastien Nouaille, Pascal Le Bourgeois, Stéphanie Heux, Jean-Charles Portais, et al. 2016. ‘Functional Characterization of a Gene Locus from an Uncultured Gut *Bacteroides* Conferring Xylo-oligosaccharides Utilization to *Escherichia Coli*’. *Molecular Microbiology* 102 (4): 579–92. <https://doi.org/10.1111/mmi.13480>.
- Tauzin, Alexandra S., Zhi Wang, Gianluca Cioci, Xiaoqian Li, Aurore Labourel, Barbara Machado, Guy Lippens, and Gabrielle Potocki-Veronese. 2022. ‘Structural and Biochemical Characterization of a Nonbinding SusD-Like Protein Involved in Xylooligosaccharide Utilization by an Uncultured Human Gut *Bacteroides* Strain’. Edited by Vincent B. Young. *mSphere* 7 (5): e00244-22. <https://doi.org/10.1128/msphere.00244-22>.
- Van den Bergh, Bram, Toon Swings, Maarten Fauvart, and Jan Michiels. 2018. ‘Experimental Design, Population Dynamics, and Diversity in Microbial Experimental Evolution’.

Microbiology and Molecular Biology Reviews 82 (3): 10.1128/mmbr.00008-18.
<https://doi.org/10.1128/mmbr.00008-18>.

Van Maris, Antonius JA, Jan-Maarten A. Geertman, Alexander Vermeulen, Matthijs K. Groothuizen, Aaron A. Winkler, Matthew DW Piper, Johannes P. Van Dijken, and Jack T. Pronk. 2004. 'Directed Evolution of Pyruvate Decarboxylase-Negative *Saccharomyces Cerevisiae*, Yielding a C2-Independent, Glucose-Tolerant, and Pyruvate-Hyperproducing Yeast'. *Applied and Environmental Microbiology* 70 (1): 159–66. <https://doi.org/10.1128/AEM.70.1.159-166.2004>.

Wang, Yu, Liwen Fan, Philibert Tuyishime, Jiao Liu, Kun Zhang, Ning Gao, Zhihui Zhang, et al. 2020. 'Adaptive Laboratory Evolution Enhances Methanol Tolerance and Conversion in Engineered *Corynebacterium Glutamicum*'. *Communications Biology* 3 (1): 1–15. <https://doi.org/10.1038/s42003-020-0954-9>.

Wisselink, H. Wouter, Maurice J. Toirkens, M. del Rosario Franco Berriel, Aaron A. Winkler, Johannes P. van Dijken, Jack T. Pronk, and Antonius J. A. van Maris. 2007. 'Engineering of *Saccharomyces Cerevisiae* for Efficient Anaerobic Alcoholic Fermentation of L-Arabinose'. *Applied and Environmental Microbiology* 73 (15): 4881–91. <https://doi.org/10.1128/AEM.00177-07>.

Yomano, LP, SW York, and LO Ingram. 1998. 'Isolation and Characterization of Ethanol-Tolerant Mutants of *Escherichia Coli* KO11 for Fuel Ethanol Production'. <https://doi.org/10.1038/sj.jim.2900496>.

Zhu, Ying, Chen Zhou, Ying Wang, and Chun Li. 2020. 'Transporter Engineering for Microbial Manufacturing'. *Biotechnology Journal* 15 (9): 1900494. <https://doi.org/10.1002/biot.201900494>.

Fourth chapter

The bovine rumen microbiome is the most diverse gut ecosystem in the animal kingdom. Bacillota and Bacteroidota are dominant phyla in the rumen microbiota. As previously explained in this manuscript, Bacteroidota are able to metabolize a wide variety of complex polysaccharides using an arsenal of proteins encoded by PULs. A hallmark of canonical PULs is the presence of at least one sequential pair of SusC and SusD homologs. SusD-like proteins, sometimes together with the other SGBPs, selectively recognize and capture the oligosaccharides that will be transported into the cell through the SusC-like TonB-dependent transporter. Despite the critical role of SusD-like proteins for efficient substrate transport, they remain largely unexplored, the vast majority of SusD-like proteins characterized to date originating from cultured bacteria. Here, we present the biochemical and structural characterization of a SusD-like protein that is encoded by a β -glucan utilization locus from a metagenomic clone (41O1) previously isolated from an *in vitro* enriched (IVTE, 20,352 clones) cow rumen metagenomic library. In this chapter, we initially verified the growth ability and enzyme activity of the clone 41O1. Subsequently, the recombinant 41O1_SusD-like protein was successfully expressed and purified in *E. coli* for further functional and structural analyses. We implemented various *in vitro* assays, as described in the first ‘Results’ chapter of this manuscript, to investigate the binding ability of the 41O1_SusD-like protein towards a series of oligosaccharides and polysaccharides. Furthermore, to face failure in crystallization assays, an AlphaFold2 model of 41O1_SusD-like was constructed and compared with its closest structural homologs, in order to identify the molecular determinants of specificity. This paper was accepted for publication in *mSphere* in June 2024.

Biochemical characterization of a SusD-like protein involved in β -1,3-glucan utilization by an uncultured cow rumen *Bacteroides*

Xiaoqian Li,¹ Guy Lippens,¹ Jean-Luc Parrou,¹ Gianluca Cioci,¹ J  r  my Esque,¹ Zhi Wang,¹ Elisabeth Laville,¹ Gabrielle Potocki-Veronese,¹ Aurore Labourel¹

AUTHOR AFFILIATION See affiliation list on p. 16.

ABSTRACT In ruminants, the rumen is a specialized stomach that is adapted to the breakdown of plant-derived complex polysaccharides through the coordinated activities of a diverse microbial community. Bacteroidota is a major phylum in this bovine rumen microbiota. They contain several clusters of genes called polysaccharide utilization loci (PULs) that encode proteins working in concert to capture, degrade, and transport polysaccharides. Despite the critical role of SusD-like proteins for efficient substrate transport, they remain largely unexplored. Here, we present the biochemical characterization of a SusD-like protein encoded by a β -glucan utilization locus from an *Escherichia coli* metagenomic clone previously isolated by functional screening of the bovine rumen microbiome. In this study, we show that clone 41O1 can grow on laminaritrise, cellotriose, and a mixture of cellobiosyl-cellobiose and glucosyl-cellotriose as sole carbon sources. Based on this, we used various *in vitro* analyses to investigate the binding ability of 41O1_SusD-like towards these oligosaccharides and the corresponding polysaccharides. We observed a clear binding affinity for β -1,6 branched β -1,3-glucans (laminarins, yeast β -glucan) and laminaritrise. Comparison of the AlphaFold2 model of 41O1_SusD-like with its closest structural homologs highlights a similar pattern of substrate recognition. In particular, three tryptophan residues are shown to be crucial for laminarin recognition. In the context of the cow rumen, we discuss the possible substrates targeted by the 41O1_PUL, such as the (1,3;1,4)- β -D-glucans present in cereal grains or the β -1,3- and (1,3;1,6)- β -D-glucans that are components of the cell wall of ruminal yeasts.

IMPORTANCE The rumen microbiota can majorly impact overall animal health, feed efficiency, and release of harmful substances into the environment. This microbiota is involved in the fermentation of organic matter to provide the host with valuable and assimilable nutrients. Bacteroidota efficiently captures, breaks down, and imports complex polysaccharides through the concerted action of proteins encoded by polysaccharide utilization loci (PULs). Within this system, SusD-like protein has proven necessary for the active internalization of the substrate. Nevertheless, the vast majority of SusD-like proteins characterized to date originate from cultured bacteria. With regard to the diversity and importance of uncultured bacteria in the rumen, further studies are required to better understand the role of polysaccharide utilization loci in ruminal polysaccharide degradation. Our detailed characterization of the 41O1_SusD-like therefore contributes to a better understanding of the carbohydrate metabolism of an uncultured *Bacteroides* from the cow rumen.

KEYWORDS β -1,3-glucan, SusC/D transporter, uncultured *Bacteroides*, metagenomics, cow rumen

Editor Garret Suen, University of Wisconsin-Madison, Madison, Wisconsin, USA

Address correspondence to Aurore Labourel, labourel@insa-toulouse.fr.

The authors declare no conflict of interest.

See the funding table on p. 16.

Received 5 April 2024

Accepted 18 June 2024

Copyright   2024 Li et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

The rumen, a specialized compartment within the digestive system of ruminant animals, such as cows, serves as a fermentation chamber for digesting food. It facilitates the breakdown of complex plant fibers and polysaccharides through the coordinated activities of a diverse microbial community. These microorganisms produce short-chain fatty acids, microbial proteins and vitamins, which provide the nutrients that the host needs for maintenance and growth (1–3). The composition and activity of the rumen microbiota are largely determined by diet, among other factors (4), and are crucial for overall animal health, improving feed efficiency, and minimizing the release of harmful substances into the environment (5–7).

With the advent of large-scale anaerobic culture-based methods (8) and the boom in various “omics” approaches, it is now possible to map the phylogenetic diversity of the rumen microbiota. This is a multi-kingdom ecosystem containing bacteria (9), ciliate protozoa (10), methanogens (*Archaea*) (11), anaerobic fungi (12) and viruses (13). Because bacteria are active players in plant biomass degradation, many recent studies have investigated their carbohydrate-active enzyme (CAZyme) content (9, 14–18). Bacteroidota, one of the dominant phyla of the rumen microbiota, possesses a particularly vast arsenal of CAZyme genes, which are colocalized in polysaccharide utilization loci (PULs) (19). The first canonical PUL was uncovered through pioneering studies of dietary starch utilization by the human gut symbiont *Bacteroides thetaiotaomicron*. According to the “starch utilization system” (Sus) paradigm (20–22), polysaccharides are first degraded by a cell envelope-associated multiprotein system (the “Sus”) that enables the bacterium to bind and degrade carbohydrates efficiently. Derivatives of this prototypic system (“Sus-like systems”) are highly represented in the genome of many other saccharolytic Bacteroidota. Next, SusD-like substrate-binding proteins, sometimes together with the other surface glycan-binding proteins (SGBPs), selectively recognize and capture the oligosaccharides that will be transported into the cell through the SusC-like TonB-dependent transporter. A hallmark of canonical PULs is the presence of at least one sequential pair of SusC and SusD homologs (23). These proteins were shown to form complexes responsible for substrate uptake through a “pedal bin” mechanism (24–26). The SusC-like transporter acts as the barrel of the bin, whereas the SusD-like protein sits atop the barrel, opening and closing like a lid to facilitate substrate binding. SusD-like proteins are of utmost importance, as their deletion, whether or not it binds the cognate substrate, leads to a loss of growth (27–29).

The majority of the SusD-like proteins characterized so far come from cultivable *Bacteroides*, i.e. *B. thetaiotaomicron* (27, 30–33), *B. ovatus* (34–36), *B. fluxus* (37) or *B. uniformis* (37). To our knowledge, only one study characterized two SusD-like proteins that have been identified from an uncultured cow rumen bacterium (38). With regard to the diversity and importance of uncultured bacteria in the rumen, further studies are required to better understand the role of polysaccharide utilization loci in ruminal fiber degradation.

In a recent publication (39), activity-based metagenomics was used to identify the key enzymes involved in the breakdown of lignocellulosic substrates by two bacterial consortia derived from the bovine rumen. In this study, we further investigate the functional abilities of the clone 41O1 identified from the cow rumen microbiome after *in vitro* enrichment (IVTE) on wheat straw.

RESULTS

Functional annotation of the 41O1 metagenomic DNA sequence reveals a PUL likely involved in β -glucan utilization

In the study by Ufarté *et al.* (39), sequencing of the 41O1 clone using the MiSeq technology revealed two contigs that could not be assembled, as was also the case for 12 of the 26 sequenced clones in the metagenomic library. In this study, we thus performed long-read sequencing of 41O1 using the MinION Oxford Nanopore Technologies to get the full-length, circular sequence. Assembly results yielded a fosmid sequence of 47,880 bp with the presence of two inserts, yet separated by two full-length pEPIFOS-5

backbones (Fig. S1). These data were confirmed by enzymatic digestion of the fosmid (Fig. S2). Functional annotation of the contig_1 (22,594 bp) revealed a PUL containing genes encoding a glycoside hydrolase of family 3 (GH3), a member of glycoside hydrolase family 16 (GH16), a surface glycan-binding protein (SusD-like protein), and a TonB-dependent transporter (SusC-like protein) that is artificially truncated by the cloning (noted at one end of the DNA insert) (Fig. 1). The contig_2 (10,243 pb) contains genes encoding an L-rhamnose-proton symporter, a rhamnulose-1-phosphate aldolase, a member of the glycoside hydrolase family 106 (GH106) and two members of the glycoside hydrolase family 78 (GH78). The activities described in the CAZy database (<http://www.cazy.org/>) for both GH78 and GH106 are rhamnogalacturonan α -L-rhamnohydrolase and α -L-rhamnosidase.

The 41O1_PUL confers the ability to grow on β -1,3/1,4- gluco-oligosaccharides to *E. coli*

According to the CAZy database, GH3 and GH16 are polyspecific families with more than 15 activities listed so far, including β -glucosidase, endo- β -1,3(4)-glucanase, xylanase and β -1,4-xylosidase activities. To screen the specificity of the 41O1_PUL CAZymes and SusD-like protein towards oligosaccharides of various structures, we performed growth assays in minimal medium containing, as the sole carbon sources, either β -gluco-oligosaccharides or xylo-oligosaccharides. Growth curves (Fig. 2) show that the 41O1 metagenomic clone exhibits significant growth ability when supplied with laminaritrise (β -1,3 linkages only) and a mixture of β -D-cellobiosyl-cellobiose and β -D-glucosyl-cellobiose (linear gluco-oligosaccharides containing both β -1,4 and β -1,3 linkages). Slight growth was observed with cellotriose (β -1,4 linkages only). No growth was observed with xylo-oligosaccharides (data not shown), highlighting the specificity of 41O1_PUL towards β -1,3 and β -1,4-linked gluco-oligosaccharides and its preference for β -1,3-linkages.

Cellular localization of the clone activity

The growth phenotype indicates that the clone 41O1 produces at least one enzyme able to degrade gluco-oligosaccharides containing β -1,3 and β -1,4 linkages. Using SignalP 6.0 (40), we investigated the putative subcellular localization of the GH3 and GH16 enzymes encoded by the 41O1_PUL. Notably, a lipoprotein signal peptide (Sec/SPII) located at the N-terminal was predicted for both proteins, indicating their potential attachment to the bacterial membrane. In the native bacterium, these proteins are likely anchored to the outer membrane and surface-exposed, as several proteins from the Sus system have been shown to face the outside environment (41–43). To experimentally investigate the subcellular localization of the GH3 and GH16 enzymes produced by the clone 41O1, enzymatic assays on laminaritrise were performed using the secreted, whole cells, soluble intracellular and membrane protein fractions. The hydrolysis reactions involving these oligosaccharides were subjected to an incubation period of 2 h, after which

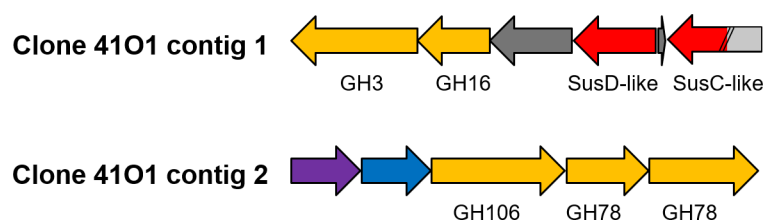


FIG 1 Representation of the 41O1 contigs. Genes encoding known and predicted functionalities are color-coded: glycoside hydrolase (GH) with family number in yellow; *SusD*-like and *SusC*-like genes in red, and genes with unknown function in dark gray; L-rhamnose-proton symporter in purple; rhamnulose-1-phosphate aldolase in blue. The *SusC*-like gene is artificially truncated (light gray) by cloning of the metagenomic DNA fragment in the pEPIFOS-5 fosmid.

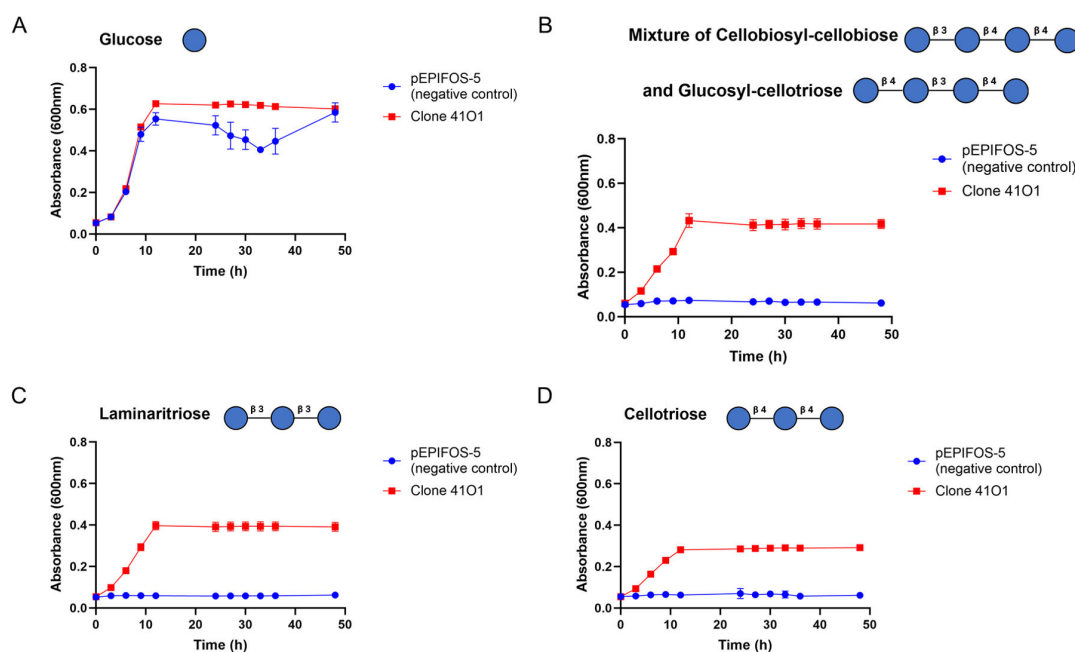


FIG 2 Growth curves of the clone 41O1 on β -1,3/ β -1,4-linked gluco-oligosaccharides at 0.5% (w/v), highlighting significant growth ability on laminaritriose and a mixture of β -D-cellobiosyl-cellobiose and β -D-glucosyl-cellobiose, and slight growth is observed with cellotriose. pEPIFOS-5 (negative control) corresponds to *E. coli* transformed with the empty pEPIFOS-5 fosmid. The data represent the average of at least biological triplicates (three different single colonies).

the resultant products were resolved on thin-layer chromatography (TLC) plates. The outcomes obtained from TLC analysis most notably indicated significant breakdown of laminaritriose (β -1,3) to laminaribiose and glucose by both the soluble intracellular and the membrane fractions (Fig. 3). After 2 h, significant amounts of laminaribiose also appeared with whole cells, but there was no clear presence of glucose, probably because it has been metabolized by the living cells. Although the experimental setup does not preclude cross-contaminations between the different cellular fractions, the overall data suggest that clone 41O1 presents both membrane-bound and cytoplasmic β -1,3-glucosidase activities.

The 41O1_PUL CAZymes prefer β -1,3-glucosyl linkages over β -1,4 and β -1,6 linkages

To investigate the preference of the 41O1_PUL CAZymes for β -glucans, we evaluated the hydrolytic activity of an intracellular extract of the metagenomic clone on polysaccharides using the 3,5-dinitrosalicylic acid reducing sugar (DNS) assay. The structure of the different polysaccharides is shown in Fig. 4A. As shown in Fig. 4B, significant hydrolytic activity was observed with the laminarins [*Ldlaminarin*, from *Laminaria digitata*, *Eblaminarin*, from *Eisenia bicyclis*, linear β -1,3-linked glucans with some 6-O-branching in the main chain and some β -(1,6) intrachain links are also present (44)], and barley β -glucan (a linear chain of β -1,4-linked cellotriosyl and cellotetraosyl units linked by β -1,3 bonds) (45). Albeit to a lesser extent, enzymatic activity is also detected with yeast β -glucan [β -1,3-glucan containing β -1,3-glucan branches extending from β -1,6-linked branch points (46)] and CM-Curdan (linear β -1,3). Conversely, no significant activity was observed with CM-Cellulose (β -1,4) and pustulan (β -1,6). This experiment supports the hypothesis that the GH3 and/or the GH16 proteins from the 41O1_PUL are involved in the growth phenotype on β -1,3/1,4-linked gluco-oligosaccharides.

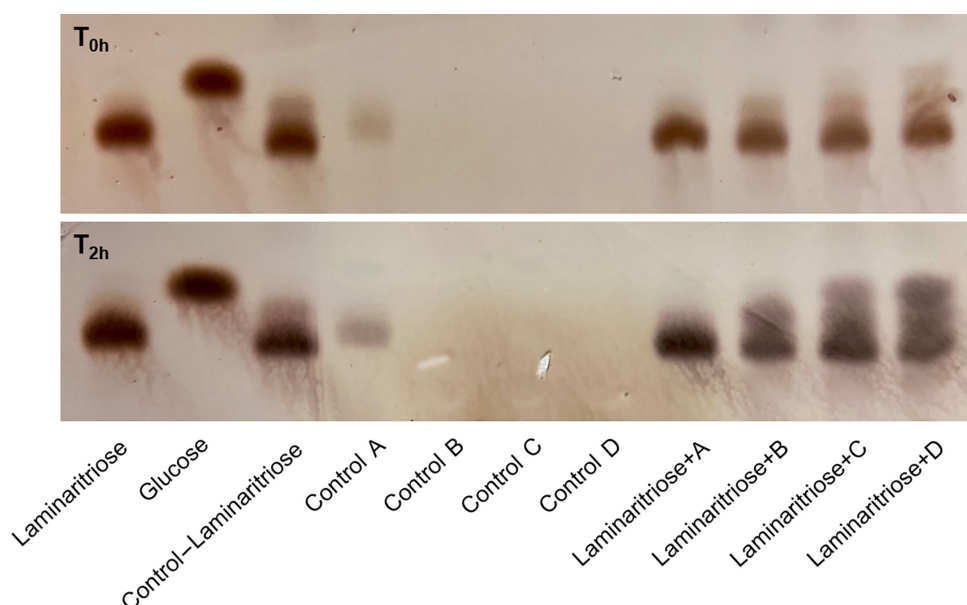


FIG 3 Cellular localization of enzymatic activity based on analysis of oligosaccharide hydrolysis reaction products by TLC, suggesting that clone 41O1 presents both membrane-bound and cytoplasmic β -1,3-glucosidase activities. Laminaritriose (β -1,3) hydrolysis reaction products were analyzed in four fractions: A – secreted fraction; B – whole cells; C – soluble intracellular fraction; D – membrane fraction. The control line “control-laminaritriose” corresponds to the substrate solution (0.2% (w/v) final concentration) and 50 mM potassium phosphate buffer, incubated overnight (about 20 h). Controls A–D correspond to the different cellular fractions, without any substrate, incubated overnight (about 20 h).

41O1_SusD-like protein has greater binding affinity for β -1,6 branched β -1,3-glucans

Because the growth phenotypes and enzymatic activity profile of clone 41O1 are likely due to the 41O1_PUL, we wanted to investigate the substrate specificity of the associated SusC/D transporter. With the 41O1_SusC-like gene being truncated, we focused characterization on the 41O1_SusD-like protein (protein sequence available in Fig. S3A). Indeed, prior investigations have provided compelling evidence highlighting the integral role of SusD-like proteins within the SusC/D transport system, particularly in the context of substrate uptake (28, 35, 47). The recombinant protein of 57 kDa was produced in soluble form in *E. coli* BL21(DE3) strain. The protein was purified to electrophoretic homogeneity by nickel affinity chromatography. The size exclusion chromatography analysis suggested that 41O1_SusD-like forms monomers and dimers in solution (Fig. S3B). We then investigated the binding abilities of the monomers towards different β -glucans.

We first tested the binding ability of the 41O1_SusD-like protein towards soluble polysaccharides at 0.2% (w/v) (otherwise stated) using affinity gel electrophoresis (AGE). As depicted in Fig. 5, the migration of the 41O1_SusD-like protein was notably impeded in the presence of different β -1,3-glucans containing occasional β -1,6-linked glucose branches. Specifically, the gels containing *L*dlaminarin, *E*blaminarin, and yeast β -glucan exhibited percentages of ~27%, 68%, and 45% retention of the 41O1_SusD-like protein, respectively. The molecular weights of the *L*dlaminarin and the *E*blaminarin are in the same range (~6 kDa) (48), so that this cannot explain the shift difference on the gels. Nevertheless, the structure of these polysaccharides can vary according to the season, and a study identified an *E*blaminarin of high molecular weight (up to 27 kDa) when *E. bicyclis* is harvested in July (49). The season in which the substrate is prepared by the supplier could affect the migration profile. It is worth noting that *E*blaminarin contains more β -1,6-linked side chains compared with *L*dlaminarin, and these branches can display a degree of polymerization up to three glucose units linked in β -1,6 (49, 50). To

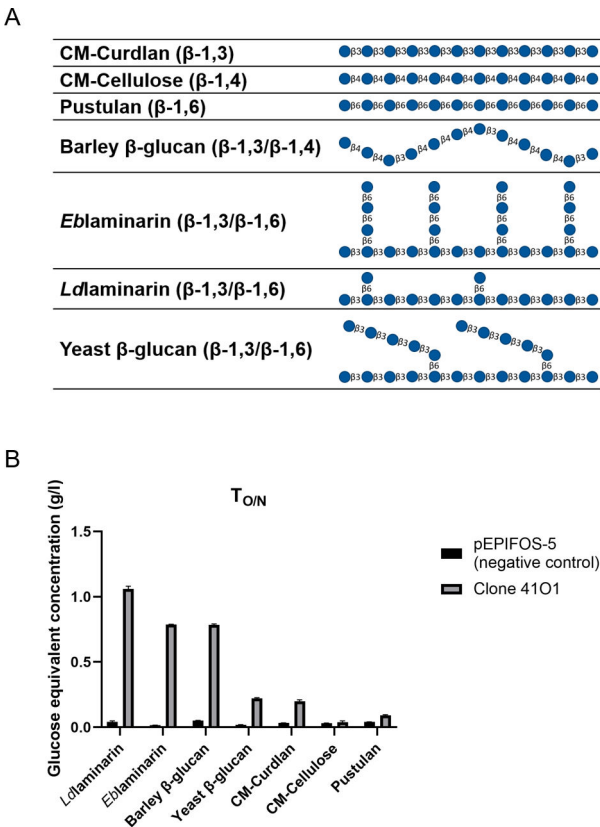


FIG 4 Enzymatic activities of clone 41O1, highlighting significant activity of the GH3 and/or the GH16 proteins on laminarins and barley β -glucan. (A) Structure of the different β -glucans used in this study. (B) Enzymatic activities of the soluble intracellular extract on various 0.2% (w/v) β -glucans were determined by DNS assays. $n = 3$, error bars represent the standard deviation of the means.

test whether 41O1_SusD-like could target β -1,6 linkages, we performed AGE using pustulan, a linear β -1,6-glucan. The results in Fig. 5 show a retention of $\sim 12\%$, indicating that it is unlikely to be the primary substrate motif targeted by 41O1_SusD-like. As mentioned earlier, the yeast β -1,3-glucan contains β -1,3-glucan branches extended from β -1,6-linked branch points (46). 41O1_SusD-like displays a significant shift on this polysaccharide ($\sim 45\%$), whereas it is much less marked ($\sim 20\%$) on the high-molecular-weight linear β -1,3-glucan CM-Curdlan (1,650 kDa; Megazyme). We cannot exclude the fact that the weaker shift on CM-Curdlan is due to the presence of the carboxymethyl groups that enable the substrate to be soluble. Because these substitutions might interfere with the protein-carbohydrate binding interactions, we could not conclude on the 41O1-SusD-like ability to bind unmodified curdlan using this experimental setup. In addition, there is no significant retention on barley (1,3;1,4)- β -glucan ($\sim 6\%$) and CM-Cellulose ($\sim 0\%$) at 0.1% (w/v). These data suggest that 41O1_SusD-like has a greater affinity for β -1,6-branched β -1,3-glucans.

We also assessed the binding ability of 41O1_SusD-like towards soluble polysaccharides by thermal shift analysis (TSA). In this experiment, we monitored the impact of the various ligands on temperature-induced protein unfolding. An increased thermal stability in the presence of a ligand was detected by a discernible shift of the unfolding profile (brightness ratio) towards a higher inflection temperature (T_i) compared with the profile of the protein without ligand. Our results are presented in Table 1; Fig. S4. Specifically, the T_i of 41O1_SusD-like is higher in the presence of *Ldlaminarin*, *Eblaminarin*, and yeast β -glucan with increments of $+2.6^\circ\text{C}$, $+8.9^\circ\text{C}$, and $+6.5^\circ\text{C}$, respectively. Minor impacts on the thermal stability can be observed in presence of CM-Curdlan ($+1.2^\circ\text{C}$) and pustulan ($+1.6^\circ\text{C}$), whereas almost no significant changes are recorded in

the presence of barley β -glucan (+0.7°C) and CM-Cellulose (+0.4°C). These results are in line with the AGE experiments and corroborate the hypothesis that 41O1_SusD-like targets β -1,6-branched β -1,3-glucans.

TSA was also used to investigate the binding ability of the protein towards the β -gluco-oligosaccharides that were used in the growth assays. Laminaritrise and the mixture of glucosyl-cellobiose and cellobiosyl-cellobiose display the highest shift in the T_i , with an increment of +4.7°C and +2.6°C, respectively. The presence of cellotriose only produces a minor impact on the T_i , with an increment of +1.2°C. These data suggest that 41O1_SusD-like has a slight ability to interact with β -1,4-linked oligosaccharides but shows a stronger affinity for the oligosaccharides containing β -1,3-linkages. Several studies on the canonical SusD protein (27) and other SusD-like proteins (36, 37) have shown that the affinity increases with degree of polymerization of the cognate oligosaccharides. Therefore, additional experiments using oligosaccharides with degrees of polymerization higher than four would enable to further investigate the binding preference of 41O1_SusD-like.

Finally, we corroborated these results using NMR spectroscopy. The 1D NMR spectra of the different putative ligands in the absence (black) and presence (red) of 41O1_SusD-like are shown in Fig. 6. The 1D proton spectrum corresponding to laminaritrise exhibits a pronounced line, broadening in the presence of 41O1_SusD-like (Fig. 6A). This is indicative not only of a discernible molecular interaction between both molecules but also of the time scale of this interaction. Indeed, to obtain line broadening despite the 100:1 excess of ligand, exchange between bound and free form cannot be too rapid. Saturation transfer difference (STD) NMR spectra (51) of the same samples were not

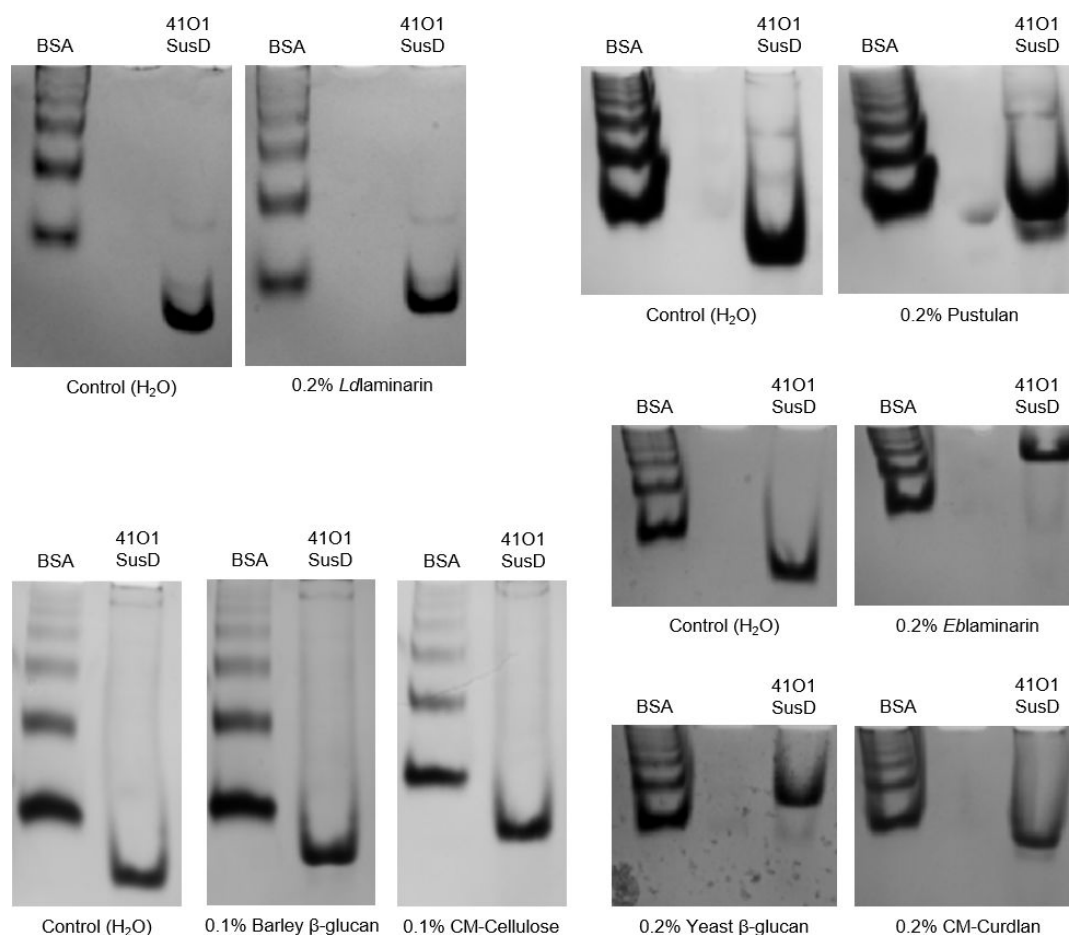


FIG 5 Affinity gel electrophoresis of 41O1_SusD-like on a gel containing no polysaccharide (control with H₂O) and on gels with different β -glucans, with bovine serum albumin (BSA) as a control protein. The gel profiles suggest that 41O1_SusD-like has a greater affinity for β -1,6-branched β -1,3-glucans.

TABLE 1 Thermal shift analysis of 10 μ M 41O1_SusD-like protein incubated with 0.5% (w/v) ligands

Substrates	$\Delta T^{\circ}\text{C}$
Laminaritriose	4.7
Cellotriose	1.2
Mixture of cellobiosyl-cellobiose and glucosyl-cellotriose	2.6
Ldlaminarin	2.6
Eblaminarin	8.9
Yeast- β -glucan	6.5
Barley- β -glucan	0.7
CM-Curdlan	1.2
CM-Cellulose	0.4
Pustulan	1.6

successful, probably due to the slow exchange. Extending the analysis to the other putative ligands, we observed no obvious influence on the 1D proton peaks of cellotriose in the presence of 41O1_SusD-like, and the STD spectrum was equally flat (Fig. 6B). In the case of the mixture of cellobiosyl-cellobiose and glucosyl-cellotriose (Fig. 6C), the major peaks were not influenced by the presence of the protein, although some minor components with signals between 3.35 and 3.40 ppm did show selective broadening.

In summary, our investigations applying different methodologies have demonstrated the clear binding affinity of 41O1_SusD-like for β -1,6 branched β -1,3-glucans and laminaritriose.

3D structure analysis of 41O1_SusD-like

To investigate the molecular determinants of substrate recognition by 41O1_SusD-like, we used crystallography in an attempt to determine its structure by X-ray crystallography. Unfortunately, we were unsuccessful in obtaining crystals, so its 3D structure was modeled using the AlphaFold2 (AF2) algorithm from LocalColabFold (52, 53). All the AF2 models showed high confidence scores with pLDDT values ranging from 95.4 to 95.8 for the mature protein, on average. All the statistics and metrics for the full length AF2 models, such as pLDDT, PAE and coverage, are given following this link <https://doi.org/10.5281/zenodo.10931522>. The 3D structure of the best AF2 model of 41O1_SusD-like mature protein is shown in Fig. 7A. To go further in the understanding of the 3D model, a BLASTp search was performed against the Protein Data Bank (PDB) (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>). The closest homolog was found to be a protein of RagB/SusD family from *B. thetaiotaomicron* (BtSGBP-A) (PDB ID: 7KV2) (37). Interestingly, the template sequence (7KV2) covered 100% of 41O1_SusD-like length, with an e-value of 0.00 and 70.99% of sequence identity. As expected, this sequence was used as one template among others by AF2. To compare both 41O1_SusD-like and 7KV2 structurally, the two 3D structures were aligned using the TM-align server (<https://zhanggroup.org/TM-align/>) (54). The RMSD and TM-score values were of 0.47 Å and 0.9, respectively, indicating that 41O1_SusD-like has the same fold as 7KV2, which corresponds to the canonical “RagB/SusD” fold. This latter is composed of one SusD-like domain (PFAM PF14322) and one SusD-RagB domain (PFAM PF07980). Based on this information, the conserved tetratricopeptide repeat (TPR) units (24, 27) known for this family were determined by comparing 41O1_SusD-like with the X-ray structure of SusD from *B. thetaiotaomicron* (PDB ID: 3CKC) (27); this is shown in Fig. 7B. Eight α -helices were found to pack together and form a right-handed superhelix along one side of the structure. These were clustered in TPR domains, such as TPR1: α 1 (residues 44–60) and α 4 (residues 110–133); TPR2: α 5 (residues 139–164) and α 6 (residues 185–202); TPR3: α 7 (residues 219–233) and α 8 (residues 236–252); and TPR4: α 16 (residues 395–408) and α 17 (residues 415–426).

To investigate the binding site region of 41O1_SusD-like, X-ray structures of SusD-like proteins were searched in complex with carbohydrates. Interestingly, the structure of BtSGBP-A was solved in a bound-form with laminarihexaose (PDB ID: 7KV3), pointing

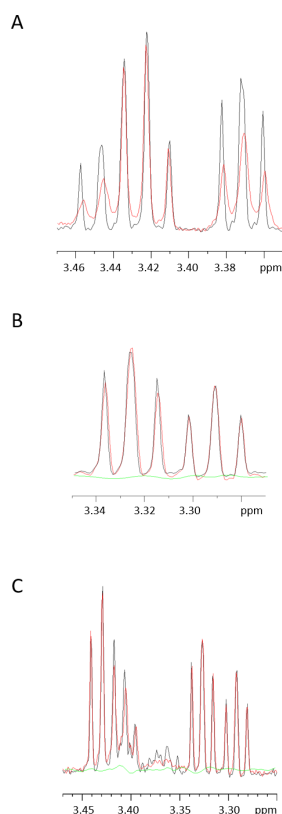


FIG 6 Characterization of the ligand-binding abilities of 41O1_SusD-like using NMR spectroscopy, demonstrating the binding affinity of 41O1_SusD-like for laminaritrifose. (A) 1D NMR spectrum of the laminaritrifose centered on the signals at 3.46–3.36 ppm: Black – Reference 1D proton spectrum of 1 mM laminaritrifose; Red - 1D proton spectrum of the same laminaritrifose in the presence of 10 μ M 41O1_SusD-like. (B) Same as panel A but for the cellotriose, proton spectrum centered on the signals at 3.34–3.28 ppm. The green spectrum is the STD signal of the same cellotriose with and without protein saturation. (C) Same as panel A but for the mixture of glucosyl-cellotriose and cellobiosyl-cellobiose, proton spectrum centered on the signals at 3.46–3.26 ppm. The green spectrum is the STD difference spectrum of the mixture with and without protein saturation.

to binding at the non-reducing end in an exo-mode (37). Moreover, *BtSGBP-A* has been shown to bind to *Ldlaminarin*, *Eblaminarin* and yeast β -glucan (55). Therefore, by superposing the 41O1_SusD-like model and the *BtSGBP-A* X-ray structure (PDB ID: 7KV3), it can be seen that the key residues involved in hydrogen bonds at the non-reducing side of the oligosaccharide are conserved (Y68, D91 and R357 for 41O1_SusD-like) (Fig. 7C). More differences occur at the reducing side. S390 (*BtSGBP-A*) is replaced by an alanine in 41O1_SusD-like (A380), leading to a loss of hydrogen bond by the serine hydroxyl group. Interestingly, a swap occurs in which D287 (*BtSGBP-A*) is replaced by S289 (41O1_SusD-like), and S391 (*BtSGBP-A*) is replaced by D381 (41O1_SusD-like). Finally, the glutamate E393 in *BtSGBP-A* is replaced by an aspartate D383 (41O1_SusD-like), both of which could form an indirect hydrogen bond with the ligand via a water molecule. With regard to hydrophobic interactions, three tryptophans (W63, W290, and W321 for 41O1_SusD-like) seem to be key to binding (Fig. 7D). These are conserved in both sequence and structure when compared with *BtSGBP-A* (W62/W288/W319) (Fig. 7D). To take things further, we compared 41O1_SusD-like with a SusD from the marine Bacteroidota member *Gramella* sp. MAR_2010_102 (GMSusD) (PDB ID: 6GCZ), which is known to bind to branched laminarin (50). Both sequences share 29% sequence identity. Surprisingly, only two of three tryptophans aligned in terms of primary sequence and 3D structures, namely W290/W287 and W321/W323 for 41O1_SusD-like and GMSusD

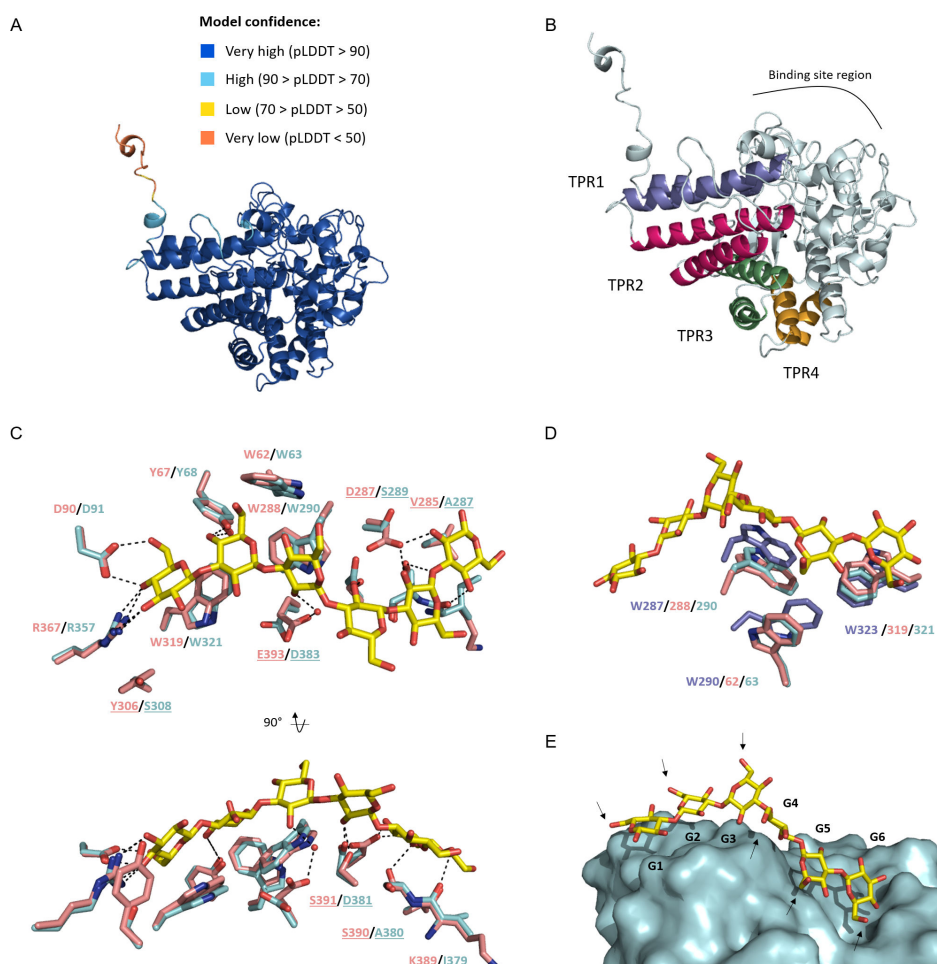


FIG 7 Structural analysis of the best AF2 model of 41O1_SusD-like mature protein. (A) Quality assessment of the best AF2 model of 41O1_SusD-like. The protein is shown in cartoon model and colored according to pLDDT confidence score. (B) 41O1_SusD-like displays the canonical “RagB/SusD” fold. TPR (tetratricopeptide repeats) domain analysis. TPR domains are highlighted in colors and labeled onto the overall 3D structure in cartoon model. (C) Structural comparison of the binding site. 41O1_SusD-like (palecyan sticks) was superimposed onto *BtSGBP-A* (light pink sticks) in complex with laminarihexaose (yellow sticks) (PDB ID: 7KV3). The residue names are labeled and colored according to the sequence. The residues that differ between the two structures are underlined. Two views rotated at 90° are displayed. Hydrogen bonds between residues and sugars are shown in dashed black lines. (D) Zoomed view on the three key tryptophans binding the laminarihexaose (yellow sticks). Three 3D structures were superimposed with 41O1_SusD-like in palecyan sticks, with *BtSGBP-A* in light pink sticks and *GM_SusD* in slate sticks (PDB ID: 6GCZ). (E) Superimposition of the laminarihexaose from 7KV3 on the surface of 41O1_SusD-like. Black arrows point to the C6-OH of each monosaccharide and represent space where β -1,6-linked branches can be accommodated. The residue numbering corresponds to the full-length protein sequences found in the UniProt database (*BtSGBP-A*: A0A173VPY3, *GM_SusD*: A0A1H1MMX0).

respectively. The third one, W290 for *GMSusD*, is close spatially to W63 for 41O1_SusD-like, but is not located in the same region in the primary sequence. Interestingly, site-directed mutagenesis of the tryptophan side chains from *GMSusD* (W287, W290 and W323) plays a critical role in substrate recognition (50), indicating the importance of this triad at the binding site. Finally, superimposing 41O1_SusD-like and *BtSGBP-A* in complex with laminarihexaose revealed a free space around the C6-OH group at each glucose binding site (Fig. 7E), thereby providing an explanation for its substrate versatility, especially towards β -1,6-branched β -1,3-glucans.

DISCUSSION

In this study, we have investigated the functional abilities of metagenomic 41O1_PUL that confers on *E. coli* the ability to grow on β -1,3/1,4-gluco-oligosaccharides. The functional annotation revealed that 41O1_SusC-like is 40% truncated on the N-terminal side when compared with SusC homologs, ruling out the 41O1_SusC-like transporter being involved in the substrate uptake. Nevertheless, the investigation of the cellular localization of the enzymatic activity revealed the presence of at least one enzyme at the cell surface. We showed that laminaritriose is broken down into laminaribiose by the whole cells, which likely metabolize the released glucose. The presence of enzymatic activities at the cell surface therefore explains the ability of the clone 41O1 to grow on the oligosaccharides, whereas its SusC-like transporter is truncated. In *E. coli*, it is widely believed that nearly all lipoproteins are either retained in the inner membrane or transferred to the inner leaflet of the outer membrane (56). Nevertheless, the free-form of the lipoprotein Lpp, the most numerically abundant protein in *E. coli*, has been shown to be surface-exposed (57) and, generally speaking, the prevalence of the surface-exposed lipoproteins in Gram-negative bacteria has been underestimated (58).

The detection of enzymatic activities in the soluble intracellular fraction (containing both the cytoplasm and the periplasm) is rather intriguing as both GH3 and GH16 proteins display a predicted lipoprotein signal peptide. In our context, it is possible that these native peptide signals are not well recognized by *E. coli*, so translocation to the membrane does not occur.

In a previous study (28), we showed that the genes are spuriously transcribed in *E. coli* due to the random presence of *E. coli* *rpoD/r70* promoter sequences on metagenomic DNA inserts. In this paper, our experiments do not allow us to know whether the 41O1-SusD-like gene is transcribed and the corresponding protein located at the cell surface. Therefore, further experiments are required to evaluate its expression and its possible role in tethering of glycans to the cell surface.

Superimposing the 3D structures of 41O1_SusD-like, BtSGBP-A and GMSusD highlights a similar pattern of β -1,3-glucans recognition, although they come from different environmental niches. Examples of gene transfer between such different environments have previously been described with the transfer of carbohydrate utilization systems from marine to human gut bacteria (59–61). Nevertheless, 41O1_SusD-like (from the cow rumen) and BtSGBP-A (from the human gut) share a high percentage of sequence identity (71% ID), although they come from environments that are not in contact with one another. Furthermore, a structural alignment showed that the three tryptophans are located at similar sequence positions within 41O1_SusD-like (W63, W290, and W321) and BtSGBP-A (W62, W288, and W319), whereas in GMSusD (from the marine environment) (W290, W287, and W323), one of the tryptophans is located at a completely different sequence position. This is made possible due to a ‘swap’ between residues in contact (i.e. that are close spatially), to offset the changes. In this instance, the residue pair W63/P293 in 41O1_SusD-like is replaced by G63/W290 in GMSusD. Overall, these results suggest that this substrate binding site topology, most notably an aromatic triad, is efficient for β -1,3-glucans recognition and might have appeared several times during evolution.

Deciphering which substrate is targeted by a PUL is always a laborious task that involves the use of several approaches. Although there are no metatranscriptomic data available that could guide us to identify in which substrates conditions our PUL is upregulated, we can still rely on biochemical data to draw hypotheses on the substrate targeted by 41O1_PUL. The 41O1 CAZymes and SusD-like protein exhibit clear enzymatic activity and binding respectively, towards the Eblaminarin (1,3;1,6)- β -D-glucan. In addition, the AlphaFold2 model of 41O1_SusD-like provides an explanation for its affinity towards Eblaminarin, with the C6-OH groups of the glucosyl residues oriented to the solvent. However, in the context of the cow rumen, this polysaccharide from brown algae is unlikely the targeted substrate. A more probable substrate is a (1,3;1,4)- β -D-glucan that is present in cereals grains (62), which are part of the cow diet (63).

A clear enzymatic activity is detectable using barley β -glucan, but we showed that 41O1_SusD does not bind to this polysaccharide and only slightly to the corresponding oligosaccharides. Nevertheless, this does not mean that it is not the substrate taken up by the 41O1_SusC/D transporter, as exemplified by the F5_SusD from a xylo-oligosaccharide (XOS) targeting PUL: this has been shown not to be able to bind to XOS or xylan (29). Indeed, regarding the recent structural insights showing that SusC-like and SusD-like proteins are closely associated *in vivo* (24, 26, 64), it is reasonable to hypothesize that SusD-like proteins that cannot bind glycans as isolated proteins can actually do it in the presence of the cognate SusC-like protein. Another hypothesis would be that the 41O1_PUL actually targets the (1,3;1,6)- β -D-glucans from the cell wall of the yeasts naturally present in the rumen, such as *Candida* species (65, 66), which are facultative anaerobes belonging to the Ascomycota phylum, or from dietary yeasts. In this study, the metagenomic DNA was indeed sampled from the rumen of cows of which the diet contained corn silage (39), which itself contains *Saccharomycetales* (67). The targeted β -glucans might also come from the strict anaerobic fungi that are part of the rumen microbiota. Presently, these anaerobic fungi all belong to the early-branching phylum Neocallimastigomycota (68), and they are attracting increasing attention for their functional contribution to the digestion of recalcitrant plant fiber (12). Nevertheless, to our knowledge, no study has been conducted on the presence of β -glucan in their cell wall and the only study we were able to find concerned the identification of chitin (69).

Determining the substrate targeted by the 41O1_PUL will require fine biochemical characterization of the GH3 and GH16 proteins on well-defined substrates. In addition, just downstream of 41O1_SusD-like, there is a protein of unknown function. Its sequence analysis reveals the presence of a lipoprotein signal peptide, suggesting that the mature protein is found anchored to the outer membrane and facing the environment. Its gene locus inside the PUL, as well as the presence of a lipoprotein signal peptide, indicate that this protein is likely an SGBP-B, with 41O1_SusD-like being the SGBP-A. Although the majority of the GHs and related SGBPs were shown to display concordant specificities (27, 32, 34–36, 70–72), recent studies have highlighted that the different specificities within GHs and SGBPs collectively dictate the overall specificity of the PUL (37, 55). Further investigation of the 41O1_PUL will therefore be needed, but our characterization of the 41O1_SusD-like paves the way for a better understanding of the carbohydrate metabolism by an uncultured bacterium from the cow rumen.

MATERIALS AND METHODS

Substrates

Barley β -glucan, carboxymethyl cellulose (4M), CM-Curdlan, and Yeast β -glucan were purchased from Megazyme (Bray, Ireland). Laminarin from *Laminaria digitata* (Ldlaminarin) and laminarin from *Eisenia bicyclis* (Eblaminarin) were purchased from Sigma-Aldrich (St. Louis, MO, USA) and Biosynth Carbosynth (Berkshire, UK), respectively. Pustulan polysaccharide was purchased from Elicityl (Grenoble, France). Laminaritriose (G3G3G), mixture of 3²- β -D-Cellobiosyl-cellobiose (G3G4G) and 3³- β -D-Glucosyl-cellobiose (G4G3G), and cellobiose (G4G4G) were purchased from Megazyme (Bray, Ireland).

Sequence of the metagenomic clone 41O1

The metagenomic clone 41O1 was issued from an *in vitro* enriched cow rumen metagenomic library (20,352 clones, *E. coli* EPI100, pEPIFOS-5 fosmid) (Ufarté *et al.*, doi.org/10.1101/2024.03.15.585145). The long-read sequencing using the MinION Oxford Nanopore Technologies was performed by the GenoToul GeT-BioPUCE platform in Toulouse. The nucleotide sequence of the clone 41O1 is available in the GenBank database at <http://www.ebi.ac.uk/ena/data/view/OZ022632-OZ022684>, under accession number OZ022674.1. The 41O1_SusD sequence is available under accession number CAK9712120.1.

Growth study

The clone 41O1 stored in glycerol at -80°C was recovered on Luria-Bertani (LB) agar plates supplemented with 12.5 mg/l chloramphenicol. Precultures of clone 41O1 were made from three isolated colonies added to liquid LB medium supplemented with 12.5 mg/l chloramphenicol and incubated overnight at 37°C with orbital shaking at 200 rpm. These precultures were used to inoculate 500 μl of M9 medium containing 0.5% (w/v) glucose, laminaritriose, cellotrios,e or a mixture of cellosyl-cellobiose and glucosyl-celotriose at an initial OD_{600} of 0.05 into a 48-well microplate. The M9 medium contained ($\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$, 17.4 g/l; KH_2PO_4 , 3.03 g/l; NaCl, 0.51 g/l; NH_4Cl , 2.04 g/l; MgSO_4 , 0.49 g/l; CaCl_2 , 4.38 mg/l; $\text{Na}_2\text{EDTA} \cdot 2\text{H}_2\text{O}$, 15 mg/l; $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 4.5 mg/l; $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$, 0.3 mg/l; $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$, 1 mg/l; H_3BO_3 , 1 mg/l; $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, 0.4 mg/l; $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 3 mg/l; $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 0.3 mg/l; thiamine, 0.1 g/l; and leucine, 0.02 g/l). Cell growth was followed by measuring the OD_{600} over 48 h at 37°C and 200 rpm.

Cellular localization of the clone activity

To investigate the cellular localization of the glycoside hydrolases (GHs) encoded by the contig_1 of the clone 41O1, laminaritriose hydrolysis assays were performed using the different cellular fractions, as well as the extracellular fraction. The clone 41O1 was grown in 20 mL of LB at 37°C until OD_{600} reached 0.7–0.8. The cells were collected by centrifugation at 4°C for 10 min at 5,000 rpm. The supernatant was filtered (0.22 μm), and this fraction was called “secreted fraction”. The pellet was gently resuspended in 2 mL of 50 mM potassium phosphate buffer (PBS) (pH 7.0). 1 ml of this suspension was kept and called “whole cells” for the activity test. The remaining 1 ml was sonicated to lyse the cells. The lysate was centrifuged at 4°C for 10 min at 10,000 rpm. The supernatant was called “soluble intracellular fraction”. The pellet was resuspended in 1 ml 50 mM of PBS and was called “membrane fraction”. The reaction media, 400 μl in total, were prepared in 50 mM PBS pH7.0 using 200 μl of the different fractions supplemented by laminaritriose at a final concentration of 0.2% (w/v). The mixtures were incubated for 2 h at 37°C and 200 rpm.

Thin layer chromatography (TLC) experiments were performed using the plates from Sigma (Product 53356, silica gel on TLC Al foil, 20 cm X 20 cm, no fluorescent indicator). Subsequently, 3 μl of each sample was loaded on the TLC plates and dried using a hairdryer. The plates were inserted in a TLC tank containing the migration buffer [butan-1-ol/acetic acid/ H_2O (2:1:1, v/v/v)]. When the solvent migration front was near the top of the plates (0.5 cm), they were dried and put into the revelation buffer (1 g/l orcinol, 75% (v/v) ethanol, and 3% (v/v) sulfuric acid) for 1 min and shaken gently. The sugars were revealed at high temperature ($>60^{\circ}\text{C}$) using a hairdryer.

Enzymatic assays

A preculture of clone 41O1 was made from an isolated colony added to liquid LB medium supplemented with 12.5 mg/L chloramphenicol. It was incubated overnight at 37°C with orbital shaking at 200 rpm. The OD_{600} was measured, and the cells were harvested by centrifugation at 4°C for 10 min at 10,000 g. The pellet was re-suspended in 50 mM potassium phosphate buffer (pH 7.0) to obtain a final OD_{600} of 80. The cells were broken by sonication (Fisher Scientific sonicator) with 5 cycles of 20 s, separated by 4 min on ice using the probe at 30% of the maximal power. The samples were centrifuged at 20,000 g for 10 min at 4°C , and the supernatant was filtered with a 0.22 μm filter (Minisart) to remove cell debris. The solution obtained was used to perform the activity tests.

To evaluate substrate specificity, the release of reducing sugar was measured using the 3,5-dinitrosalicylic acid reducing sugar (DNS) assay (DNS solution: 10 g/l DNS, 300 g/l potassium sodium tartrate, 16 g/l NaOH in deionized water) (73). The assays were carried out at 37°C , 200 rpm, in 50 mM potassium phosphate buffer pH 7.0, containing 0.2% (w/v) of substrate and 160 μl of cellular extract. The total volume of the reaction medium

was 400 μ l. Standard curves were made with glucose concentrations ranging from 0 to 2 g/l. 120 μ l of reaction was taken at T_{0h} and $T_{overnight}$ and added to an equal volume of DNS reagent. The samples were boiled for 5 min and centrifuged for 1 min at 20,000 g. Subsequently, 200 μ l of the supernatants was transferred into the wells of a polystyrene microplate to read the OD at 540 nm in an Optima (TECAN) plate reader. The experiments were carried out in triplicates.

Recombinant protein production and purification

The DNA sequence corresponding to the 41O1_SusD-like protein was optimized for *E. coli* and synthesized by GeneCust. The lipoprotein peptide signal predicted by LipoP 1.0 (74) was not included in the sequence. The gene was cloned between the NdeI and XhoI restriction sites in the pET-28a(+) expression vector (Novagen, Darmstadt, Germany) that enables the fusion of a 6-His tag at the N-terminal for affinity purification.

The pET-28a (+) plasmid containing the 41O1_SusD-like gene was transformed into *E. coli* BL21(DE3) for protein production. A single bacterial colony was inoculated in 5 ml of LB medium containing kanamycin (50 μ g/ml) and incubated at 37°C, 200 rpm. The overnight preculture was used to inoculate 200 ml of LB containing 50 μ g/ml of kanamycin at an initial OD₆₀₀ of 0.05. The overexpression of 41O1_SusD-like protein was induced by adding isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM in the mid-exponential phase (OD₆₀₀ \approx 0.6), and the culture was further incubated for 4 h at 37°C, 200 rpm. The cells were harvested by centrifugation at 5,000 rpm for 10 min at 4°C. The pellet was resuspended in Talon Buffer (20 mM Tris-HCl, 300 mM NaCl, pH 7.4) and disrupted by sonication. Cells debris were cleared by centrifugation at 11,000 g for 30 min at 4°C. The supernatant was further filtered using a 0.22 μ m filter (Minisart) and passed through a column of 2 ml of Talon metal affinity resin (Clontech, USA). A washing step was performed with buffer containing 10 mM imidazole, 20 mM Tris-HCl, 300 mM NaCl, pH 7.4. The 41O1_SusD-like protein was eluted with 250 mM imidazole, 20 mM Tris-HCl, 300 mM NaCl, pH 7.4. The imidazole was further eliminated using a PD-10 Desalting Column (Cytiva), and Tris buffer (20 mM Tris-HCl, 150 mM NaCl, pH 7.4). Protein purity was confirmed via SDS-PAGE and the concentration was determined by the absorbance at 280 nm using the Thermo Scientific Nanodrop 2000 with an extinction coefficient of 143,390 M⁻¹cm⁻¹. The protein was then further polished by gel filtration using an AKTA System (HiPrep 16/60 Sephacryl S-200 HR) (GE-Healthcare, Uppsala, Sweden) in 20 mM Tris-HCl, 150 mM NaCl, pH 7.4. The fractions containing the monomeric and dimeric 41O1_SusD-like protein were pooled respectively, concentrated and stored at 4°C. The monomeric form of 41O1_SusD-like was used for analysis.

Crystallization assays

Crystallization assays were attempted with 41O1_SusD-like monomers concentrated at 9.63 mg/ml. Initial crystallization conditions were screened using a Mosquito robot (TTP Labtech) and the commercial screens Classic Suite, JCSG+ and PACT from Qiagen. We also attempted to co-crystallize 41O1_SusD-like with 10 mM of laminaritrise using the same conditions.

Affinity gel electrophoresis

To investigate the ability of 41O1_SusD-like to bind polysaccharides, affinity gel electrophoresis (AGE) was performed. Continuous native polyacrylamide gels were prepared consisting of 8% (w/v) acrylamide in 25 mM Tris, 250 mM glycine buffer, pH 8.3. Various polysaccharides were added before polymerization at a final concentration of 0.5% (w/v). Subsequently, 2.5 μ g of the tested 41O1_SusD-like, along with bovine serum albumin (BSA) used as non-interacting negative control protein, was loaded on the gels. Electrophoresis was carried out at 90V on ice for at least 2 hours. The percentage of retention of 41O1_SusD-like in the gel containing glycans was calculated as described

by Tauzin et al. (29): $(D_{\text{wog}} - D_{\text{wg}})/D_{\text{wog}} \times 100$, where $D_{\text{wog}} = D_{\text{SusD-like_wog}}/D_{\text{BSA_wog}}$ (distance of migration of the lowest and most intense 41O1_SusD-like band normalized to the distance of migration of BSA in the gel without glycan), and $D_{\text{wg}} = D_{\text{SusD-like_wg}}/D_{\text{BSA_wg}}$ (distance of migration of the lowest and most intense 41O1_SusD-like band normalized to the distance of migration of BSA in the gel with glycan).

Thermal shift analysis

Thermal shift analysis (TSA) is a rapid and inexpensive label-free screening method to identify low-molecular-weight ligands that bind and stabilize purified proteins. The structural stability of 41O1_SusD-like in the presence or absence of potential ligands was investigated by TSA using the Tycho NT.6 (NanoTemper Technologies, Munich, Germany), according to the user manual. Subsequently, 10 μl of samples containing 10 μM of 41O1_SusD-like in Tris buffer (20 mM Tris-HCl, 150 mM NaCl, pH 7.4), with or without ligands, at a final concentration of 0.5% (w/v) was used to fill the capillary tubes. The intrinsic protein fluorescence was recorded at 330 nm and 350 nm while the sample was heated from 35°C to 95°C over a 3-min period.

NMR spectroscopy

The interaction of 41O1_SusD-like with ligands was assayed by Saturation transfer difference NMR (STD NMR) spectroscopy. All experiments were performed at 298K on a Bruker Avance II 800 MHz NMR spectrometer equipped with a QCPY cryogenically cooled probe head. The samples contained 50 μM 41O1_SusD-like protein with or without ligands (0.1 mM, 1 mM) in a Tris buffer (Tris 20 mM, NaCl 150 mM at pH 7.4).

1D proton spectra were acquired using Watergate water suppression (75), with 256 scans for 41O1_SusD-like. STD experiments were performed with saturation of the protein resonances at 0 ppm through a train of 5 ms Gaussian 180° pulses. Typical experiments were run with 256 scans, 2,048 acquisition points and a 5-s relaxation delay (including the 3-s pre-saturation train). STD experiments were run with 256 or 512 scans. Spectra were transformed after one level of zero-filling and apodization with a $\pi/3$ shifted square sine bell.

AlphaFold2 model of 41O1_SusD-like

Monomeric structures of 41O1_SusD-like were modeled using LocalColabFold 1.5.2, run on 17 April 2023 (53). Thus, the sequence search was performed using MMseqs (76) on a distant server. The targeted sequence (41O1_SusD-like) as aligned against protein sequences from several databases: (i) Uniref30 (77), (ii) a combination of Big Fantastic Database and MGnify, (iii) PDB70 database, and (iv) ColabFoldDB (53). Localcolabfold parameters were set to the default values, with three prediction recycles and the generation of five models. All 3D models were relaxed with default parameters of AF2 using AMBER Force Field. For structural comparisons, the residue numbering corresponds to the full-length protein sequences found in the UniProt database (*Bt*SGBP-A: A0A173VPY3, GM_SusD: A0A1H1MMX0).

ACKNOWLEDGMENTS

Protein purification and TSA analyses were performed on equipment which is part of the PICT-ICEO facility dedicated to enzyme screening and discovery, part of the Integrated Screening Platform of Toulouse (PICT, IBISA). PICT-ICEO is a member of IBISBA-FR (<https://doi.org/10.15454/08BX-VJ91>), the French node of the European research infrastructure, IBISBA (www.ibisba.eu). NMR spectroscopy was performed on the spectrometer of MetaToul (Toulouse Metabolomics & Fluxomics facilities, www.metatoul.fr). Metatoul is part of the French National Infrastructure for Metabolomics and Fluxomics MetaboHUB-AR-11-INBS-0010 (www.metabohub.fr), and is supported by the Région Midi-Pyrénées, the ERDF, the Sicoval, and the French Minister of Education & Research, who are all gratefully acknowledged.

This research was funded by the European Union's Horizon 2020 Framework Programme LEIT-BIO-2015–685474 Metafluidics. X.L. is supported by INSA Toulouse and the China Scholarship Council under the CSC-UT/INSA Program.

AUTHOR AFFILIATION

¹TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France

PRESENT ADDRESS

Zhi Wang, School of Chemical Engineering, Zhengzhou University, Zhengzhou, Henan, China

AUTHOR ORCIDs

Gabrielle Potocki-Veronese  <https://orcid.org/0000-0003-4232-230X>

Aurore Labourel  <http://orcid.org/0000-0002-4762-5804>

FUNDING

Funder	Grant(s)	Author(s)
EC Horizon 2020 Framework Programme (H2020)	LEIT-BIO-2015- 685474 Metafluidics	Gabrielle Potocki-Veronese
China Scholarship Council (CSC)		Xiaoqian Li

ADDITIONAL FILES

The following material is available [online](#).

Supplemental Material

Supplemental Figures (mSphere00278-24-s0001.pdf). Fig. S1 to S4.

REFERENCES

- Bergman EN. 1990. Energy contributions of volatile fatty acids from the gastrointestinal tract in various species. *Physiol Rev* 70:567–590. <https://doi.org/10.1152/physrev.1990.70.2.567>
- Franco-Lopez J, Duplessis M, Bui A, Reymond C, Poisson W, Blais L, Chong J, Gervais R, Rico DE, Cue RI, Girard CL, Ronholm J. 2020. Correlations between the composition of the bovine microbiota and Vitamin B₁₂ abundance. *mSystems* 5. <https://doi.org/10.1128/mSystems.00107-20>
- Lima J, Ingabire W, Roehe R, Dewhurst RJ. 2023. Estimating microbial protein synthesis in the rumen—can ‘omics’ methods provide new insights into a long-standing question? *12 Vet Sci* 10:679. <https://doi.org/10.3390/vetsci10120679>
- Henderson G, Cox F, Ganesh S, Jonker A, Young W, Global Rumen Census Collaborators, Janssen PH. 2015. Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. *Sci Rep* 5:14567. <https://doi.org/10.1038/srep14567>
- McCann JC, Wickersham TA, Looor JJ. 2014. High-throughput methods redefine the rumen microbiome and its relationship with nutrition and metabolism. *Bioinform Biol Insights* 8:109–125. <https://doi.org/10.4137/BBI.S15389>
- Mizrahi I, Jami E. 2018. Review: the compositional variation of the rumen microbiome and its effect on host performance and methane emission. *Animal* 12:s220–s232. <https://doi.org/10.1017/S1751731118001957>
- Mizrahi I, Wallace RJ, Morais S. 2021. The rumen microbiome: balancing food security and environmental impacts. *Nat Rev Microbiol* 19:553–566. <https://doi.org/10.1038/s41579-021-00543-6>
- Seshadri R, Leahy SC, Attwood GT, Teh KH, Lambie SC, Cookson AL, Eloe-Fadros EA, Pavlopoulos GA, Hadjithomas M, Varghese NJ, Paez-Espino D, Hungate1000 project collaborators, Perry R, Henderson G, Creevey CJ, Terrapon N, Lapebie P, Drula E, Lombard V, Rubin E, Kyrpides NC, Henrissat B, Woyke T, Ivanova NN, Kelly WJ. 2018. Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection. *Nat Biotechnol* 36:359–367. <https://doi.org/10.1038/nbt.4110>
- Stewart RD, Auffret MD, Warr A, Wiser AH, Press MO, Langford KW, Liachko I, Snelling TJ, Dewhurst RJ, Walker AW, Roehe R, Watson M. 2018. Assembly of 913 microbial genomes from metagenomic sequencing of the cow rumen. *Nat Commun* 9:870. <https://doi.org/10.1038/s41467-018-03317-6>
- Andersen TO, Altshuler I, Vera-Ponce de León A, Walter JM, McGovern E, Keogh K, Martin C, Bernard L, Morgavi DP, Park T, Li Z, Jiang Y, Firkins JL, Yu Z, Hvidsten TR, Waters SM, Popova M, Arntzen MØ, Hagen LH, Pope PB. 2023. Metabolic influence of core ciliates within the rumen microbiome. *ISME J* 17:1128–1140. <https://doi.org/10.1038/s41396-023-01407-y>
- Janssen PH, Kirs M. 2008. Structure of the archaeal community of the rumen. *Appl Environ Microbiol* 74:3619–3625. <https://doi.org/10.1128/AEM.02812-07>
- Hagen LH, Brooke CG, Shaw CA, Norbeck AD, Piao H, Arntzen MØ, Olson HM, Copeland A, Isern N, Shukla A, Roux S, Lombard V, Henrissat B, O'Malley MA, Grigoriev IV, Tringe SG, Mackie RI, Pasa-Tolic L, Pope PB, Hess M. 2021. Proteome specialization of anaerobic fungi during ruminal degradation of recalcitrant plant fiber. *ISME J* 15:421–434. <https://doi.org/10.1038/s41396-020-00769-x>
- Yan M, Pratama AA, Somasundaram S, Li Z, Jiang Y, Sullivan MB, Yu Z. 2023. Interrogating the viral dark matter of the rumen ecosystem with a global virome database. *Nat Commun* 14:5254. <https://doi.org/10.1038/s41467-023-41075-2>
- Stewart RD, Auffret MD, Warr A, Walker AW, Roehe R, Watson M. 2019. Compendium of 4,941 rumen metagenome-assembled genomes for

- rumen microbiome biology and enzyme discovery. *Nat Biotechnol* 37:953–961. <https://doi.org/10.1038/s41587-019-0202-3>
15. Tomazetto G, Pimentel AC, Wibberg D, Dixon N, Squina FM. 2020. Multi-omic directed discovery of cellulosomes, polysaccharide utilization loci, and lignocellulases from an enriched rumen anaerobic consortium. *Appl Environ Microbiol* 86:e00199–20. <https://doi.org/10.1128/AEM.00199-20>
 16. Gharechahi J, Vahidi MF, Bahram M, Han J-L, Ding X-Z, Salekdeh GH. 2021. Metagenomic analysis reveals a dynamic microbiome with diversified adaptive functions to utilize high lignocellulosic forages in the cattle rumen. *ISME J* 15:1108–1120. <https://doi.org/10.1038/s41396-020-00837-2>
 17. Bohra V, Tikariha H, Purohit HJ, Dafale NA. 2022. Unique pool of carbohydrate-degrading enzymes in novel bacteria assembled from cow and buffalo rumen metagenomes. *Appl Microbiol Biotechnol* 106:4643–4654. <https://doi.org/10.1007/s00253-022-12020-y>
 18. Auer E, Lazuka A, Huguenin-Bizot B, Jehmlich N, Déjean S, Lombard V, Henrissat B, O'Donohue M, Hernandez-Raquet G. 2023. Horizontal metaproteomics and CAZymes analysis of lignocellulolytic microbial consortia selectively enriched from cow rumen and termite gut. *ISME Communications* 3:134. <https://doi.org/10.1038/s43705-023-00339-0>
 19. Grondin JM, Tamura K, Déjean G, Abbott DW, Brumer H. 2017. Polysaccharide utilization loci: fueling microbial communities. *J Bacteriol* 199:e00860–16. <https://doi.org/10.1128/JB.00860-16>
 20. Martens EC, Koropatkin NM, Smith TJ, Gordon JL. 2009. Complex glycan catabolism by the human gut microbiota: the bacteroidetes sus-like paradigm. *J Biol Chem* 284:24673–24677. <https://doi.org/10.1074/jbc.R109.022848>
 21. Foley MH, Cockburn DW, Koropatkin NM. 2016. The Sus operon: a model system for starch uptake by the human gut Bacteroidetes. *Cell Mol Life Sci* 73:2603–2617. <https://doi.org/10.1007/s00018-016-2242-x>
 22. Brown HA, Koropatkin NM. 2021. Host glycan utilization within the Bacteroidetes Sus-like paradigm. *Glycobiology* 31:697–706. <https://doi.org/10.1093/glycob/cwaa054>
 23. Terrapon N, Lombard V, Drula É, Lapébie P, Al-Masaudi S, Gilbert HJ, Henrissat B. 2018. PULDB: the expanded database of polysaccharide utilization loci. *Nucleic Acids Res.* 46:D677–D683. <https://doi.org/10.1093/nar/gkx1022>
 24. Glenwright AJ, Pothula KR, Bhamidimarri SP, Chorev DS, Baslé A, Firbank SJ, Zheng H, Robinson CV, Winterhalter M, Kleinekathöfer U, Bolam DN, van den Berg B. 2017. Structural basis for nutrient acquisition by dominant members of the human gut microbiota. *Nature* 541:407–411. <https://doi.org/10.1038/nature20828>
 25. Bolam DN, van den Berg B. 2018. TonB-dependent transport by the gut microbiota: novel aspects of an old problem. *Curr Opin Struct Biol* 51:35–43. <https://doi.org/10.1016/j.sbi.2018.03.001>
 26. Gray DA, White JBR, Oluwale AO, Rath P, Glenwright AJ, Mazur A, Zahn M, Baslé A, Morland C, Evans SL, Cartmell A, Robinson CV, Hiller S, Ranson NA, Bolam DN, van den Berg B. 2021. Insights into SusCD-mediated glycan import by a prominent gut symbiont. *Nat Commun* 12:44. <https://doi.org/10.1038/s41467-020-20285-y>
 27. Koropatkin NM, Martens EC, Gordon JL, Smith TJ. 2008. Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. *Structure* 16:1105–1115. <https://doi.org/10.1016/j.str.2008.03.017>
 28. Tauzin AS, Laville E, Xiao Y, Nouaille S, Le Bourgeois P, Heux S, Portais J-C, Monsan P, Martens EC, Potocki-Veronese G, Bordes F. 2016. Functional characterization of a gene locus from an uncultured gut *Bacteroides* conferring xylo-oligosaccharides utilization to *Escherichia coli*. *Mol Microbiol* 102:579–592. <https://doi.org/10.1111/mmi.13480>
 29. Tauzin A.S, Wang Z, Cioci G, Li X, Labourel A, Machado B, Lippens G, Potocki-Veronese G. 2022. Structural and biochemical characterization of a nonbinding SusD-like protein involved in xylooligosaccharide utilization by an uncultured human gut *Bacteroides* strain. *mSphere* 7:e00244–22. <https://doi.org/10.1128/msphere.00244-22>
 30. Cuskin F, Lowe EC, Temple MJ, Zhu Y, Cameron EA, Pudlo NA, Porter NT, Urs K, Thompson AJ, Cartmell A, et al. 2015. Human gut Bacteroidetes can utilize yeast mannan through a selfish mechanism. *Nature* 517:165–169. <https://doi.org/10.1038/nature13995>
 31. Temple MJ, Cuskin F, Baslé A, Hickey N, Speciale G, Williams SJ, Gilbert HJ, Lowe EC. 2017. A Bacteroidetes locus dedicated to fungal 1,6-β-glucan degradation: unique substrate conformation drives specificity of the key endo-1,6-β-glucanase. *J Biol Chem* 292:10639–10650. <https://doi.org/10.1074/jbc.M117.787606>
 32. Luis AS, Briggs J, Zhang X, Farnell B, Ndeh D, Labourel A, Baslé A, Cartmell A, Terrapon N, Stott K, Lowe EC, McLean R, Shearer K, Schückel J, Venditto I, Ralet M-C, Henrissat B, Martens EC, Mosimann SC, Abbott DW, Gilbert HJ. 2018. Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic Bacteroides. *Nat Microbiol* 3:210–219. <https://doi.org/10.1038/s41564-017-0079-1>
 33. Tamura K, Brumer H. 2021. Glycan utilization systems in the human gut microbiota: a gold mine for structural discoveries. *Curr Opin Struct Biol* 68:26–40. <https://doi.org/10.1016/j.sbi.2020.11.001>
 34. Rogowski A, Briggs JA, Mortimer JC, Tryfona T, Terrapon N, Lowe EC, Baslé A, Morland C, Day AM, Zheng H, Rogers TE, Thompson P, Hawkins AR, Yadav MP, Henrissat B, Martens EC, Dupree P, Gilbert HJ, Bolam DN. 2015. Glycan complexity dictates microbial resource allocation in the large intestine. 1. *Nat Commun* 6:7481. <https://doi.org/10.1038/ncomms8481>
 35. Tauzin AS, Kwiatkowski KJ, Orlovsky NI, Smith CJ, Creagh AL, Haynes CA, Wawrzak Z, Brumer H, Koropatkin NM. 2016. Molecular dissection of xyloglucan recognition in a prominent human gut symbiont. *mBio* 7:e02134–15. <https://doi.org/10.1128/mBio.02134-15>
 36. Tamura K, Foley MH, Gardill BR, Dejean G, Schnitzlein M, Bahr CME, Louise Creagh A, van Petegem F, Koropatkin NM, Brumer H. 2019. Surface glycan-binding proteins are essential for cereal beta-glucan utilization by the human gut symbiont *Bacteroides ovatus*. *Cell Mol Life Sci* 76:4319–4340. <https://doi.org/10.1007/s00018-019-03115-3>
 37. Tamura K, Dejean G, Van Petegem F, Brumer H. 2021. Distinct protein architectures mediate species-specific beta-glucan binding and metabolism in the human gut microbiota. *J Biol Chem* 296:100415. <https://doi.org/10.1016/j.jbc.2021.100415>
 38. Mackenzie AK, Pope PB, Pedersen HL, Gupta R, Morrison M, Willats WGT, Eijsink VGH. 2012. Two SusD-like proteins encoded within a polysaccharide utilization locus of an uncultured ruminant *Bacteroidetes* phylotype bind strongly to cellulose. *Appl Environ Microbiol* 78:5935–5937. <https://doi.org/10.1128/AEM.01164-12>
 39. Ufarté L, Laville E, Lazuka A, Lajus S, Bouhajja E, Cecchini D, Rizzo A, Amblard E, Drula E, Lombard V, Terrapon N, Henrissat B, Cleret M, Morgavi DP, Dumon C, Robe P, Klopp C, Bozonnet S, Hernandez-Raquet G, Potocki-Veronese G. 2024. Functional exploration of *in vivo* and *in vitro* lignocellulose-fed rumen bacterial microbiomes reveals novel enzymes involved in polysaccharide breakdown. *bioRxiv*. <https://doi.org/10.1101/2024.03.15.585145>
 40. Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsigos KD, Winther O, Brunak S, von Heijne G, Nielsen H. 2022. SignalP 6.0 predicts all five types of signal peptides using protein language models. 7. *Nat Biotechnol* 40:1023–1025. <https://doi.org/10.1038/s41587-021-01156-3>
 41. Reeves AR, Wang GR, Salyers AA. 1997. Characterization of four outer membrane proteins that play a role in utilization of starch by *Bacteroides thetaiotaomicron*. *J Bacteriol* 179:643–649. <https://doi.org/10.1128/jb.179.3.643-649.1997>
 42. Shipman JA, Cho KH, Siegel HA, Salyers AA. 1999. Physiological characterization of Susg, an outer membrane protein essential for starch utilization by *Bacteroides thetaiotaomicron*. *J Bacteriol* 181:7206–7211. <https://doi.org/10.1128/JB.181.23.7206-7211.1999>
 43. Shipman JA, Berleman JE, Salyers AA. 2000. Characterization of four outer membrane proteins involved in binding starch to the cell surface of *Bacteroides thetaiotaomicron*. *J Bacteriol* 182:5365–5372. <https://doi.org/10.1128/JB.182.19.5365-5372.2000>
 44. Kadam SU, Tiwari BK, O'Donnell CP. 2015. Extraction, structure and biofunctional activities of laminarin from brown algae. *Int J of Food Sci Tech* 50:24–31. <https://doi.org/10.1111/ijfs.12692>
 45. Izydorczyk MS, Dexter JE. 2008. Barley β-glucans and arabinoxylans: molecular structure, physicochemical properties, and uses in food products—a review. *Food Res Int* 41:850–868. <https://doi.org/10.1016/j.foodres.2008.04.001>
 46. Manners DJ, Masson AJ, Patterson JC, Björndal H, Lindberg B. 1973. The structure of a β-(1→3)-d-glucan from yeast cell walls. *Biochem J* 135:31–36. <https://doi.org/10.1042/bj1350031>
 47. Cameron EA, Kwiatkowski KJ, Lee B-H, Hamaker BR, Koropatkin NM, Martens EC. 2014. Multifunctional nutrient-binding proteins adapt

- human symbiotic bacteria for glycan competition in the gut by separately promoting enhanced sensing and catalysis. *mBio* 5. <https://doi.org/10.1128/mBio.01441-14>
48. Pang Z, Otaka K, Maoka T, Hidaka K, Ishijima S, Oda M, Ohnishi M. 2005. Structure of β -glucan oligomer from laminarin and its effect on human monocytes to inhibit the proliferation of U937 cells. *Biosci Biotechnol Biochem* 69:553–558. <https://doi.org/10.1271/bbb.69.553>
 49. Menshova RV, Ermakova SP, Anastyuk SD, Isakov VV, Dubrovskaya YV, Kusaykin MI, Um B-H, Zvyagintseva TN. 2014. Structure, enzymatic transformation and anticancer activity of branched high molecular weight laminarin from brown alga *Eisenia bicyclis*. *Carbohydrate Polymers* 99:101–109. <https://doi.org/10.1016/j.carbpol.2013.08.037>
 50. Mystkowska AA, Robb C, Vidal-Melgosa S, Vanni C, Fernandez-Guerra A, Höhne M, Hehemann J-H. 2018. Molecular recognition of the beta - glucans laminarin and pustulan by a SusD - like glycan - binding protein of a marine *Bacteroidetes*. *FEBS J* 285:4465–4481. <https://doi.org/10.1111/febs.14674>
 51. Mayer M, Meyer B. 1999. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew Chem Int Ed Engl* 38:1784–1788. [https://doi.org/10.1002/\(SICI\)1521-3773\(19990614\)38:12<1784::AID-ANIE1784>3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1521-3773(19990614)38:12<1784::AID-ANIE1784>3.0.CO;2-Q)
 52. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Židek A, Potapenko A, et al. 2021. Highly accurate protein structure prediction with AlphaFold. 7873. *Nature* 596:583–589. <https://doi.org/10.1038/s41586-021-03819-2>
 53. Mirdita M, Schütze K, Moriawaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022. ColabFold: making protein folding accessible to all. *Nat Methods* 19:679–682. <https://doi.org/10.1038/s41592-022-01488-1>
 54. Zhang Y, Skolnick J. 2005. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res* 33:2302–2309. <https://doi.org/10.1093/nar/gki524>
 55. Déjean G, Tamura K, Cabrera A, Jain N, Pudlo NA, Pereira G, Viborg AH, Van Petegem F, Martens EC, Brumer H. 2020. Synergy between cell surface glycosidases and glycan-binding proteins dictates the utilization of specific Beta(1,3)-glucans by human gut bacteroides. *mBio* 11:e00095-20. <https://doi.org/10.1128/mBio.00095-20>
 56. Tokuda H, Matsuyama S. 2004. Sorting of lipoproteins to the outer membrane in *E. coli*. *Biochim Biophys Acta Mol Cell Res* 1693:5–13. <https://doi.org/10.1016/j.bbamer.2004.02.005>
 57. Cowles CE, Li Y, Semmelhack MF, Cristea IM, Silhavy TJ. 2011. The free and bound forms of Lpp occupy distinct subcellular locations in *Escherichia coli*. *Mol Microbiol* 79:1168–1181. <https://doi.org/10.1111/j.1365-2958.2011.07539.x>
 58. Wilson MM, Bernstein HD. 2016. Surface-exposed lipoproteins: an emerging secretion phenomenon in Gram-negative bacteria. *Trends Microbiol* 24:198–208. <https://doi.org/10.1016/j.tim.2015.11.006>
 59. Hehemann J-H, Correc G, Barbeyron T, Helbert W, Czek M, Michel G. 2010. Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* 464:908–912. <https://doi.org/10.1038/nature08937>
 60. Hehemann J-H, Kelly AG, Pudlo NA, Martens EC, Boraston AB. 2012. Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. *Proc Natl Acad Sci U S A* 109:19786–19791. <https://doi.org/10.1073/pnas.1211002109>
 61. Pudlo NA, Pereira GV, Parnami J, Cid M, Markert S, Tingley JP, Unfried F, Ali A, Varghese NJ, Kim KS, Campbell A, Urs K, Xiao Y, Adams R, Martin D, Bolam DN, Becher D, Eloje-Fadrosch EA, Schmidt TM, Abbott DW, Schweder T, Hehemann JH, Martens EC. 2022. Diverse events have transferred genes for edible seaweed digestion from marine to human gut bacteria. *Cell Host Microbe* 30:314–328. <https://doi.org/10.1016/j.chom.2022.02.001>
 62. Genç H, Özdemir M, Demirbaş A. 2001. Analysis of mixed-linked (1→3), (1→4)- β -d-glucans in cereal grains from Turkey. *Food Chem* 73:221–224. [https://doi.org/10.1016/S0308-8146\(00\)00290-9](https://doi.org/10.1016/S0308-8146(00)00290-9)
 63. Humer E, Zebeli Q. 2017. Grains in ruminant feeding and potentials to enhance their nutritive and health value by chemical processing. *Anim Feed Sci Technol* 226:133–151. <https://doi.org/10.1016/j.anifeedsci.2017.02.005>
 64. White JBR, Silale A, Feasey M, Heunis T, Zhu Y, Zheng H, Gajbiye A, Firbank S, Baslé A, Trost M, Bolam DN, van den Berg B, Ranson NA. 2023. Outer membrane utilosomes mediate glycan uptake in gut *Bacteroidetes*. *Nature* 618:583–589. <https://doi.org/10.1038/s41586-023-06146-w>
 65. Lund A. 2000. Yeasts and moulds in the bovine rumen. *Microbiology* 81:453–462. <https://doi.org/10.1099/00221287-81-2-453>
 66. Lowman DW, Greene RR, Bearden DW, Kruppa MD, Pottier M, Monteiro MA, Soldatov DV, Ensley HE, Cheng S-C, Netea MG, Williams DL. 2014. Novel structural features in *Candida albicans* hyphal glucan provide a basis for differential innate immune recognition of hyphae versus yeast. *J Biol Chem* 289:3432–3443. <https://doi.org/10.1074/jbc.M113.529131>
 67. Benjamim da Silva É, Liu X, Mellinger C, Gressley TF, Stypinski JD, Moyer NA, Kung L Jr. 2022. Effect of dry matter content on the microbial community and on the effectiveness of a microbial inoculant to improve the aerobic stability of corn silage. *J Dairy Sci* 105:5024–5043. <https://doi.org/10.3168/jds.2021-21515>
 68. Paul SS, Bu D, Xu J, Hyde KD, Yu Z. 2018. A phylogenetic census of global diversity of gut anaerobic fungi and a new taxonomic framework. *Fungal Divers* 89:253–266. <https://doi.org/10.1007/s13225-018-0396-6>
 69. Orpin CG. 1977. The occurrence of Chitin in the cell walls of the rumen organisms *Neocallimastix frontalis*, *Pyromonas communis* and *Sphaeromonas communis*. *J Gen Microbiol* 99:215–218. <https://doi.org/10.1099/00221287-99-1-215>
 70. Larsbrink J, Rogers TE, Hemsworth GR, McKee LS, Tauzin AS, Spadiut O, Klinger S, Pudlo NA, Urs K, Koropatkin NM, Creagh AL, Haynes CA, Kelly AG, Cederholm SN, Davies GJ, Martens EC, Brumer H. 2014. A discrete genetic locus confers xyloglucan metabolism in select human gut *Bacteroidetes*. *Nature* 506:498–502. <https://doi.org/10.1038/nature12907>
 71. Tamura K, Hemsworth GR, Déjean G, Rogers TE, Pudlo NA, Urs K, Jain N, Davies GJ, Martens EC, Brumer H. 2017. Molecular mechanism by which prominent human gut *Bacteroidetes* utilize mixed-linkage beta-glucans, major health-promoting cereal polysaccharides. *Cell Rep* 21:417–430. <https://doi.org/10.1016/j.celrep.2017.09.049>
 72. Ndeh D, Rogowski A, Cartmell A, Luis AS, Baslé A, Gray J, Venditto I, Briggs J, Zhang X, Labourel A, Terrapon N, Buffetto F, Nepogodiev S, Xiao Y, Field RA, Zhu Y, O'Neil MA, Urbanowicz BR, York WS, Davies GJ, Abbott DW, Ralet M-C, Martens EC, Henrissat B, Gilbert HJ. 2017. Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature* 544:65–70. <https://doi.org/10.1038/nature21725>
 73. Miller GL. 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal Chem* 31:426–428. <https://doi.org/10.1021/ac60147a030>
 74. Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H, Krogh A. 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci* 12:1652–1662. <https://doi.org/10.1110/ps.0303703>
 75. Piotto M, Saudek V, Sklenár V. 1992. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR* 2:661–665. <https://doi.org/10.1007/BF02192855>
 76. Steinegger M, Söding J. 2017. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol* 35:1026–1028. <https://doi.org/10.1038/nbt.3988>
 77. Suzek BE, Wang Y, Huang H, McGarvey PB, Wu CH, UniProt Consortium. 2015. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* 31:926–932. <https://doi.org/10.1093/bioinformatics/btu739>

The figure displays a genomic map of the 4101 region. The main map shows a 47,880 bp region with two pEPIFOS-5 copies and two inserts. Zoomed-in views show the pEPIFOS-5 locus (6 kb) and the sopC gene (7 kb). A detailed view of the sopC gene shows a 17 kb to 22 kb region with a blue arrow pointing to a specific site.

The duplication of the pEPIFOS-5 backbone appears as two inverted 7,518 pb sequences, each of them showing the repeated segments within the sopC feature. Another singular feature from this 41O1 clone is the huge, repeated area of almost 4 kb in the middle of the insert 1 sequence, which is made up of the succession of a perfectly conserved 47 bp motif separated by sequences of 20 to 30 bp.

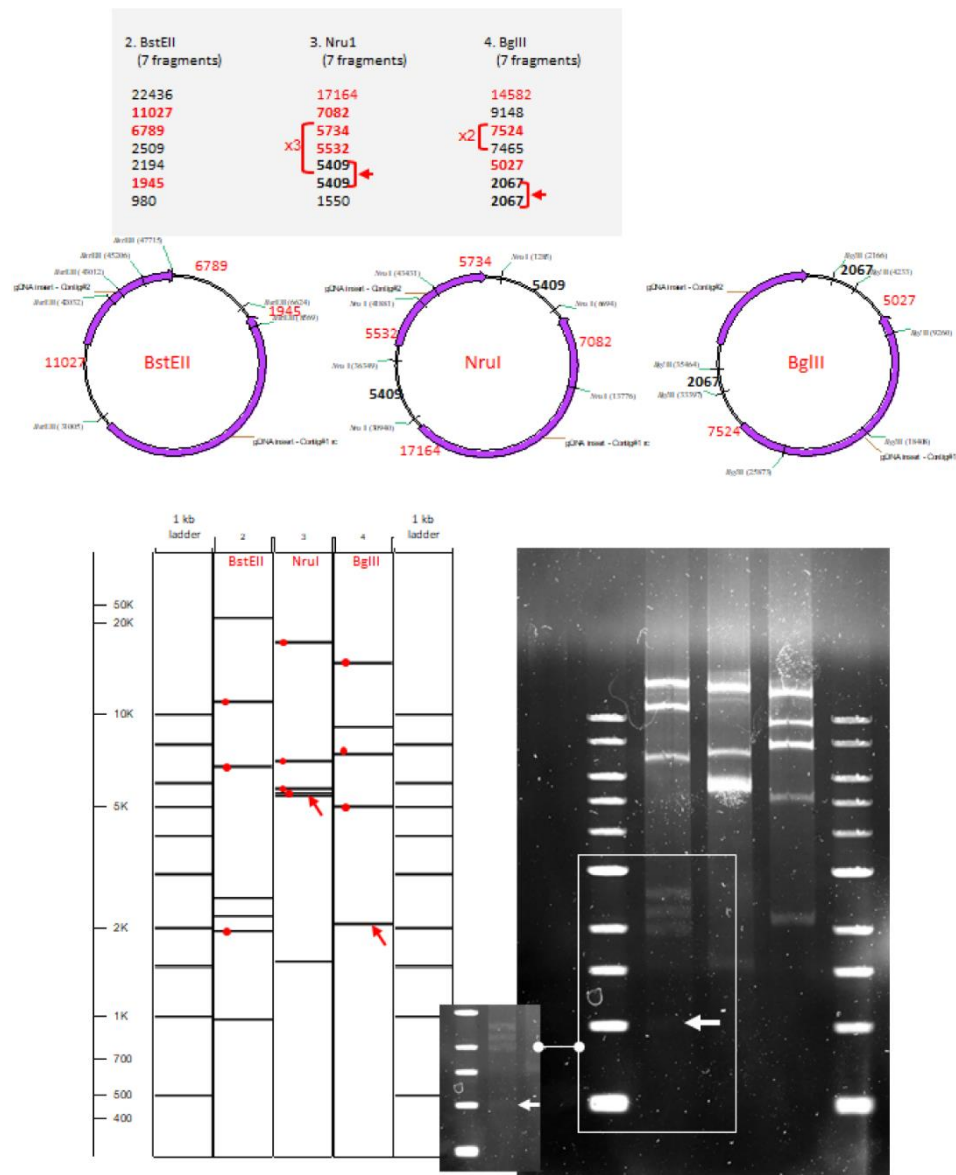


Figure S2: Restriction analysis of the 41O1 clone.

Three restriction enzymes were selected, (i.e. BstE, NruI, BglII), to provide characteristic restriction patterns highlighting the presence of two independent inserts and backbones. In red, characteristic bands at the backbone/insert junctions. Left: Theoretical migration gel prepared with the VectorNTI software; Red dots correspond to highlighted, characteristic bands at the junctions. Red arrow points doublets arising from the presence of two backbones (5409 pb and 2067 pb for NruI and BglII restriction, respectively). Right: Agarose migration gel, 0.8 % in TAE. The white arrow targets the 980 pb faint band from BstEII digestion.

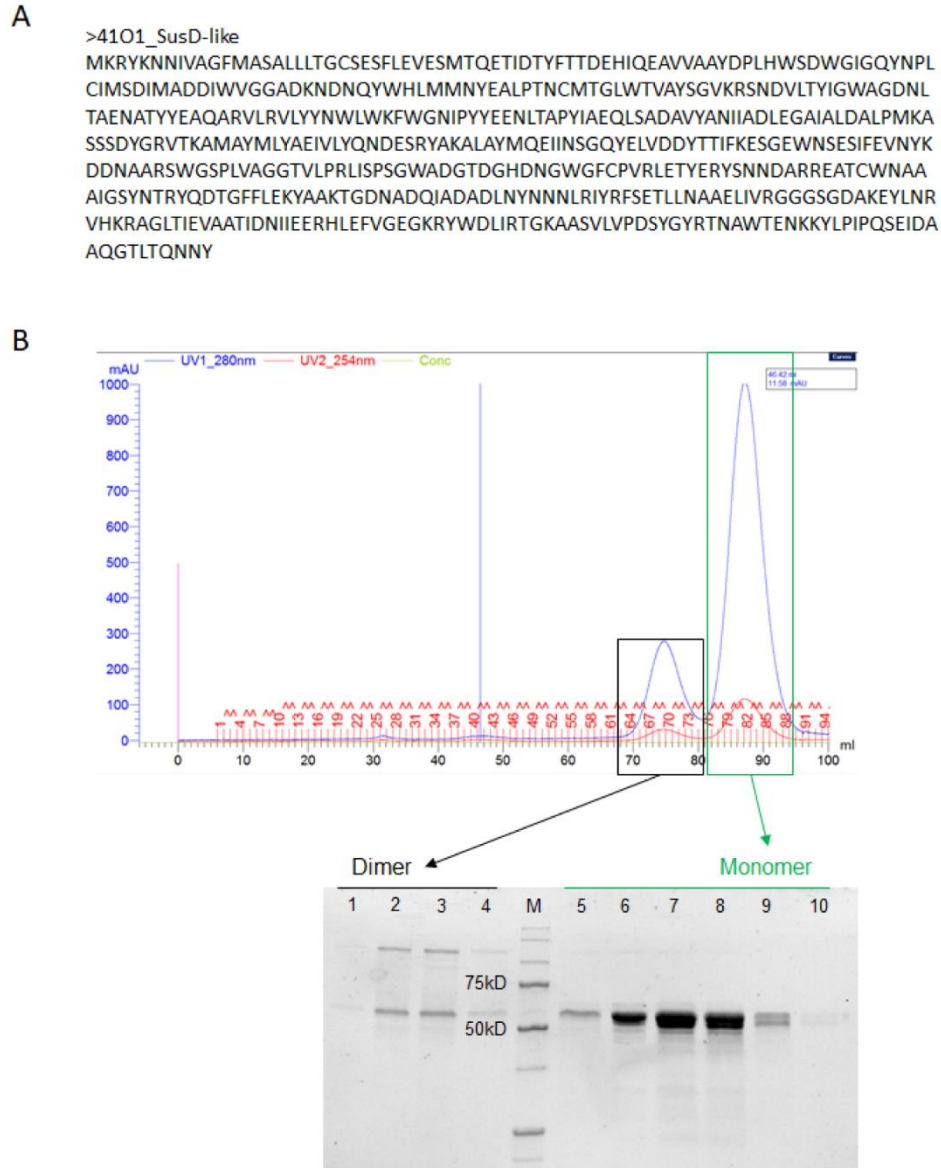


Figure S3: (A) 41O1_SusD-like full-length protein sequence. (B) Gel filtration of 41O1_SusD-like (57 kDa).

The protein was polished by gel filtration using an AKTA System (HiPrep 16/60 Sephacryl S-200 HR) (GE-Healthcare, Uppsala, Sweden) in 20 mM Tris - HCl, 150 mM NaCl, pH 7.4. SDS-PAGE gel photographed with BioRad GS-800 calibrated densitometer system. Line M: Precision Plus Protein Unstained Standards marker; Line 1-4: Fractions of dimeric 41O1_SusD-like; Line 5-10: Fractions of monomeric 41O1_SusD-like.

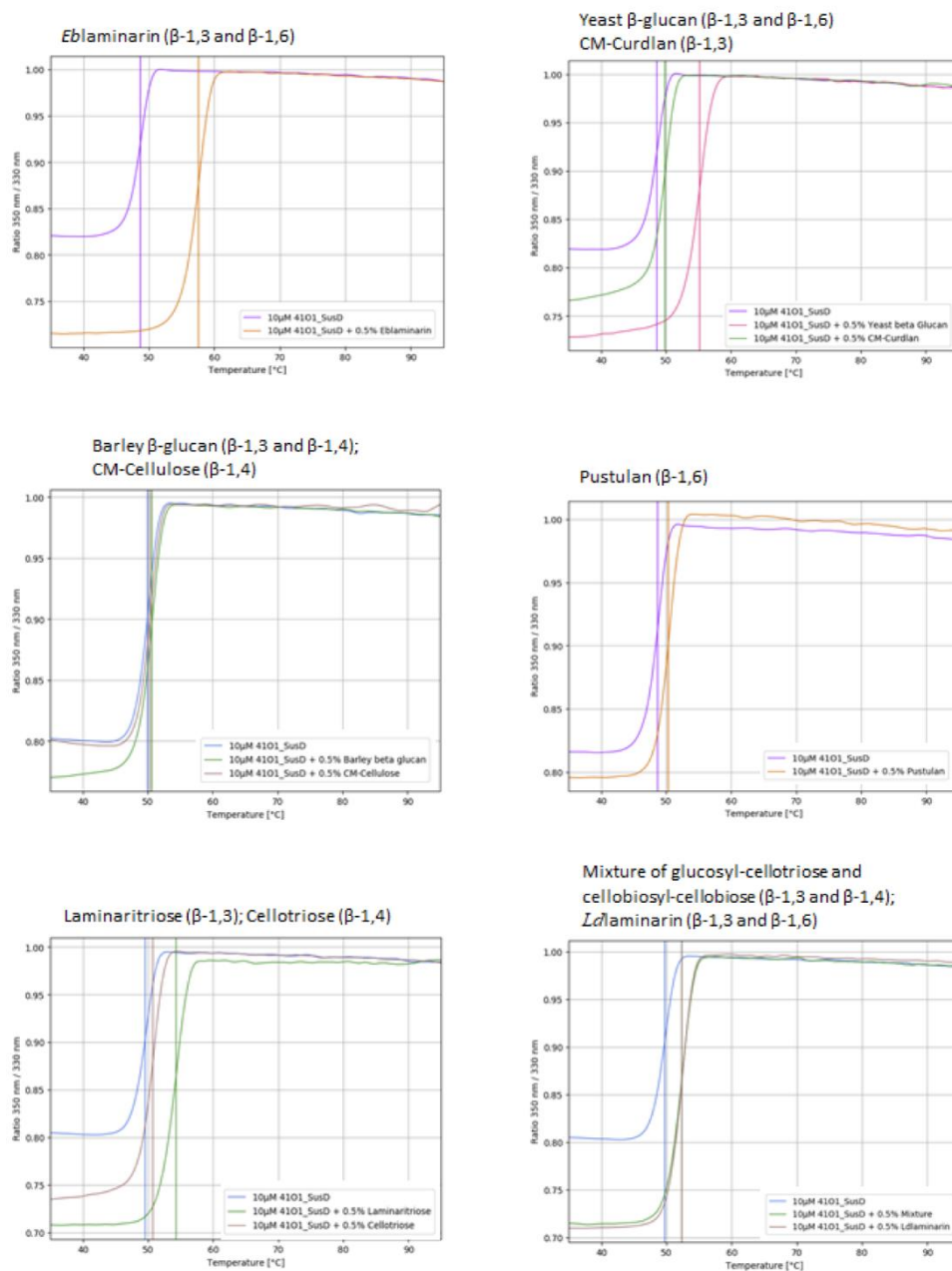


Figure S4: Thermal shift analysis of 10 μM 41O1_SusD-like protein incubated with 0.5 % (w/v) ligands. Vertical lines indicate inflection temperatures (Ti).

Conclusion and perspectives

This thesis work was mostly focused on the functional and structural characterization of oligosaccharide transporters, in particular those from the SusC/D superfamily. Besides, a first attempt of *in cellulo* random engineering was performed, with the aim to increase oligosaccharide uptake efficiency.

Study of the relationships between the structure of SusD-like proteins, their binding ability and the specificity of the SusC/D oligosaccharide transporter systems

The first part of this thesis work was dedicated to the functional and structural characterization of a SusD-like protein (F5_SusD-like) from the F5 XOS-targeting PUL involved in XOS uptake by an uncultured *Bacteroides*. This PUL is conserved in *Bacteroides vulgatus*. Previously, the F5 XOS_PUL was identified by the CIMEs team using activity-based functional metagenomics, based on positive selection on solid plates containing XOS as sole carbon sources (Cecchini et al. 2013). This PUL involved in linear XOS utilization encodes a SusC/D-like pair and a major facilitator superfamily (MFS) transport system capable of XOS uptake up to a degree of polymerization of 3 (DP3) and DP4, respectively (Tauzin et al. 2016). The deletion of the SusD-like protein abolished the growth ability of the recombinant *E. coli* cells on XOS, which demonstrates that F5_SusD-like protein is necessary for XOS transport. Based on this, we investigated the biochemical and structural properties of F5_SusD-like protein, after gene over-expression in *E. coli* and protein purification based on affinity chromatography and gel filtration, which revealed that the F5_SusD-like protein exists in both monomer and dimer forms in solution. The ability of the F5_SusD-like protein to bind several soluble polysaccharides was examined by using AGE, ITC, DSF, and STD-NMR. The results of AGE indicated that the F5_SusD-like dimer does not bind to any polysaccharide. Nevertheless, the monomer migration was delayed by around 30 % in the gels containing beechwood, birchwood, and oat spelt xylans. These results suggested a weak affinity of F5_SusD-like for xylan. Furthermore, the interaction of F5_SusD-like with XOS was investigated using ITC, DSF and STD-NMR, allowing us to conclude that this SGBP-like protein does not bind to XOS, although it weakly binds xylan. Moreover, the three-dimensional structure of F5_SusD-like was solved by X-ray crystallography, highlighting a canonical “RagB/SusD” fold, including the conserved tetratricopeptide repeat (TPR) units. The failure of co-crystallization experiments with XOS further confirmed the inability of F5_SusD-like to bind to short-chain XOS. Additionally, a structural comparison with other SGBPs revealed that F5_SusD-like displays a long partially disordered loop (D292 to N315) located close to the canonical binding site. Several key tryptophan residues involved in substrate binding are absent in F5_SusD-like and are instead

carried by a folded loop located precisely in the area from D292 to N315. Consequently, we hypothesized that the lack of binding affinity of F5_SusD-like to XOS could be due to the absence of aromatic residues as well as an unfolded binding surface with the loop spanning D292 to N315 being partially disordered. Although the F5_SusD-like protein does not bind XOS, its presence is essential for XOS transport by the SusC/D complex, suggesting that the binding specificity of SGBPs does not necessarily correlate with transporter uptake specificity. These results refine the concept of glycoside utilization by *Bacteroidetes*, highlighting the complexity of the mechanisms involved in substrate binding and uptake.

Based on these findings, we wondered if the fact that the F5_SusD-like protein could not bind the substrate that is efficiently transported by the SusC/D complex could be due to a certain flexibility of this transporter towards oligosaccharides of various structures. Tauzin *et al.* previously showed that both F5_SusC/D and F5_MFS transporters are capable of uptaking XOS with DP 3. Xylosyl units are linked by β -1,4-glycosidic bonds in xylotriose. Similarly, mannotriose is composed of three mannosyl units linked by β -1,4-glycosidic bonds, thus sharing structural similarity with xylotriose. Thus, a novel strategy was set up to investigate the specificity of recombinant transporters towards oligosaccharides. We expressed a gene encoding a periplasmic β -mannosidase (BT_0458, from the GH2 CAZy family) in the *E. coli* F5 clone in order to complement its XOS utilization system involving the F5_SusC/D and F5_MFS transporters. The BT_0458 β -mannosidase derived from *Bacteroides thetaiotaomicron* VPI-5482 had previously been structurally and functionally characterized, and shown to hydrolyze β 1,4-mannosides (Tailford et al. 2007). The aim of this strategy was to determine whether the F5 transporters are capable of uptaking mannotriose, which would then be degraded to monosaccharides by the β -mannosidase, thus supporting the growth of the *E. coli* strain. In order to characterize the specificity of each of the F5 transporters towards the structure of oligosaccharides, three F5 minor truncated variants were selected as chassis strains: F5minSus, F5minMFS, F5minGH. These variants contain the three GHs and the SusC/D transporter, the GHs and the MFS transporter, or only the GHs, respectively. Subsequently, two genes for the recombinant BT_0458 β -mannosidase, along with the *Bt* and *E. coli* signal peptides, were synthesized and introduced in three F5 variants, respectively: Sus-GH2 and Sus-GH2P, MFS-GH2 and MFS-GH2P, F5minGH-GH2 and F5minGH2P. Both the screening of the double resistance (kanamycin and chloramphenicol) and colony PCR confirmed that the recombinant constructs contained both the fosmid and plasmid. The results of TLC analysis of the mannotriose hydrolysis products by the different cellular fractions of the constructed strains,

indicated that the recombinant GH2 enzyme was successfully produced under an active form, both associated to the membrane and in the cytoplasm, rather than being secreted. No growth was observed for the GH2 expressing strains on minimal medium containing mannotriose as sole carbon source, indicating that neither the SusC/D transporter nor the MFS transporter encoded by F5_XOS_PUL could transport mannotriose. This preliminary conclusion on the strict specificity of the F5_SusC/D and MFS transporters will have to be further confirmed by checking that the GH2 enzyme is properly produced in minimum medium.

Another part of this thesis was dedicated to the biochemical characterization of a SusD protein encoded by a β -glucan utilization locus from the metagenomic clone 41O1, issued from the cow rumen microbiome after *in vitro* enrichment on wheat straw. This clone was first identified by screening a metagenomic library for endo-acting glycoside-hydrolases able to breakdown plant cell wall polysaccharides. Then, a preliminary study performed by Zhi Wang highlighted the potential of clone 41O1 to grow on laminaritriose (β -1,3-gluco-oligosaccharides) and a mixture of β -D-cellobiosyl-cellobiose and β -D-glucosyl-celotriose (linear gluco-oligosaccharides containing both β -1,4 and β -1,3-linkages). A slight growth was also observed with cellotriose (β -1,4-linkages only). The growth phenotype thus indicated that clone 41O1 produces at least one enzyme able to degrade gluco-oligosaccharides containing β -1,3- and β -1,4-linkages. In this thesis work, we thus investigated the subcellular localization of the GH3 and GH16 enzymes encoded by the 41O1_PUL. Enzymatic assays on laminaritriose were performed using the secreted, soluble intracellular and membrane protein fractions. The results indicated that clone 41O1 presents both membrane-bound and cytoplasmic β -1,3-glucosidase activities. Moreover, the hydrolytic activity of an intracellular extract of clone 41O1 on various polysaccharides was assessed by a DNS assay. Significant hydrolytic activity was observed with laminarins (*Ldl*laminarin, *Ebl*laminarin, linear β -1,3-linked glucans with some 6-O-branching in the main chain and some β -(1,6)-intrachain links) and barley β -glucan (a linear chain of β -1,4-linked cellotriosyl and cellotetraosyl units linked by β -1,3 bonds). A weak enzymatic activity was also detected with yeast β -glucan (β -1,3-glucan containing β -1,3-glucan branches extended from β -1,6-linked branch points) and CM-Curdlan (linear β -1,3-glucan). In contrast, no significant activity was observed with CM-Cellulose (linear β -1,4-glucan) and pustulan (linear β -1,6-glucan). These results thus indicated that the GH3 and/or the GH16 encoded by the 41O1_PUL are involved in the growth phenotype on β -1,3/1,4-linked gluco-oligosaccharides and highlighted their preference towards linear β -1,3-linked glucosidic chains. Furthermore, the recombinant 41O1_SusD protein was produced in *E. coli* and purified to

investigate its ligand specificity. Binding assays indicated that the 41O1_SusD protein is able to bind β -1,6 branched β -1,3-glucans and laminaritrifose. Subsequently, the 3D structure of 41O1_SusD protein was modeled using the AlphaFold2. 41O1_SusD exhibits a canonical “RagB/SusD” fold. Additionally, by superposing the 41O1_SusD model and the X-ray structure of *BtSGBP-A* (a SusD protein from *B. thetaiotaomicron* involved in the binding of *Ldlaminarin*, *Eblaminarin* and yeast- β -glucan) (PDB ID: 7KV3), we highlighted the structural conservation in 41O1_SusD of the three tryptophans (W63, W290, W321, 41O1_SusD-like numbering) involved in hydrophobic interactions with the ligands. The superimposition of the 41O1_SusD model and of the structure of *BtSGBP-A* in complex with laminarihexaose revealed the presence of a free space around the C6-OH group at each glucose binding site, thereby providing an explanation for the substrate versatility of 41O1_SusD-like, particularly towards branched β -1,3-glucans.

Overall, the results of these three chapters contributed to the understanding of the relationships between the structure of SusD proteins, their ligand binding functions, and the role that they play in oligosaccharide uptake through TonB dependent transporters.

However, there are still open questions regarding the oligosaccharide/polysaccharide binding ability and specificity of SusD proteins:

- Are all SusC-like proteins specific for the uptake of certain oligosaccharides, like the F5_SusC protein seems to be?

This question could be addressed by studying other SusC/D systems using the same strategy as that we used to assess the specificity of the F5_SusC/D system in the second results chapter of this manuscript, that is to say by complementing the system by a recombinant periplasmic glycosidase acting on a different oligosaccharide than that naturally internalized by the transporter.

- Do the SusC/D systems encoded by a same PUL strictly function in tandem or could SusD proteins interact with SusC proteins encoded by other PULs?

In order to address this question, combinatorial studies could be performed by replacing native SusD encoding genes by other ones (e.g. SusD1 from PUL1 to put in tandem with SusC2 from PUL2) in PULs that could be recombinantly expressed in *E. coli*. This work will require i) to construct libraries of SusD cassettes to plug in PULs previously proven to encode functional SusC/D systems in *E. coli* (ex the F5 PUL); ii) to screen the growth of the *E. coli* recombinant strains on the oligosaccharide targeted by the

initial SusC protein (e.g. on xylooligosaccharides if the 41O1_SusD /F5_SusC combination is tested).

- Are non-binding SusD proteins (such as the F5_SusD protein) frequent in Bacteroidota? It is indeed difficult to assess their frequency using multiple sequence alignments (MSA) to study the conservation of the residues usually involved in ligand binding, because of the low homology shared by SusD protein sequences, which affects the MSA quality. The systematic use of Alpha-Fold to model the structure of SusD proteins could help assessing the structural conservation of the ligand binding residues, and to predict binding ability.
- How does the SusC/D complex (or SusC alone) recognize the oligosaccharide to be internalized if the SusD is not able to bind oligosaccharides? The recombinant expression of PULs in *E. coli* is attractive, since it allows us to exclude the role played in binding by other SGBPs or CAZymes produced by the native strain. Nevertheless, to date, the lack of structural and functional studies targeting transmembrane SusC proteins alone, or in complex with SusD (in particular with non-binding SusD proteins), limits the identification of the molecular determinants of binding. Solving the crystallographic structure of a SusC/non-binding SusD complex would be highly interesting to identify the residues involved in protein-protein interactions, in oligosaccharide recognition and in uptake. Such study being very challenging, another approach based on i) random mutagenesis of the SusC and SusD encoding sequences (individually or in combination) and ii) screening the binding and uptake phenotypes of the mutants, could be envisioned.

Engineering of oligosaccharide transporters

As presented in the second section of the bibliographic introduction, and according to the perspectives listed just above regarding the investigation of structure-function relationships of transporters, several engineering approaches could be used to improve the transport of specific glycosides across cell membranes of bacteria and yeast and to investigate the mechanisms and specificity of transport.

In this thesis work, we carried out a preliminary study to test the potential of adaptive laboratory evolution (ALE) to randomly engineer the F5_SusC/D system expressed in *E. coli*. The minimal truncated variants of the F5_PUL (F5minSus, F5minGH as a negative control) served as a model case to assess the impact of mutations on the ability of *E. coli* to grow on XOS. The ALE

experiment consisted in five serial transfers of batch cultures in M9 minimal medium supplemented with 5 g/l of xylotriose as the sole carbon source. Each transfer was incubated for 12 hours to ensure that cells were in the exponential phase. The growth curves demonstrated differences in the growth profiles of the various F5minSus generations, with a rate of biomass increase added by 3.25 and 1.88 times, respectively, for the 3rd and 5th generations. Furthermore, the 3rd F5minSus culture consumed 14.5 % of xylotriose after 6 hours of growth corresponding to the beginning of the exponential phase, which is 6.9 times higher than for the 1st F5minSus culture. NGS was then used to determine if the differences in the growing and xylotriose consumption phenotypes of the S3 and S5 samples were correlated to mutations in the F5minSus locus. In the F5minSus locus, which sizes 10,416 nucleotides, we detected 62 hot points in the S3 reads, and 60 in the S5 reads. These mutations, which are exclusively deletions or insertions, are present at ratios varying between 9 and 41 % of the reads, which is significant. Moreover, 52 of the most significant mutations are located (both for S3 and S5) in the GH, SusD or SusC encoding genes. Thereby, we concluded that mutations in the functional genes of the F5 locus probably affected the growing phenotype on XOS. However, we still have to demonstrate that the observed mutations are not due to sequencing errors, and to isolate the strains with the best growth phenotypes from the S3 or S5 strain mixtures. Sequencing of both the fosmid and genomic DNA of monoclonal evolved strains will be required in order to identify the mutations that might have occurred in the functional genes of the F5 locus, and also to conclude on the potential presence of mutations in the *E. coli* genome and on their impact on the observed phenotypes. Complementary analyses will thus be performed, to quantify the growth ability of each of the strains of the S3 and S5 samples by a microfluidics-based technology developed in the CIMEs team. It will be followed by the sequencing of fosmidic and genomic DNA of 10 monoclonal strains that will be sorted based on their high efficiency to grow on xylotriose.

Besides, in order to engineer oligosaccharide transporters, other strategies were attempted during this thesis work. Initially, we tried to generate random mutations in the F5minSus locus by exposing liquid and solid cultures to Ultraviolet (UV) radiation. The samples were exposed to UV for varying durations and incubated. Regrettably, all of the UV-exposed cells were killed directly, even with the shortest exposure time being only 2 seconds. Then, we attempted to randomly mutate single-cells of the *E. coli* strain harboring the F5minSus locus by irradiating pico-liter droplets in which single-cells of the initial F5minSus strain were encapsulated, in order to assess the growing phenotype of each generated mutants, without recombination.

Unfortunately, irradiation of the droplets killed all cells. In addition, we attempted to generate sequence diversity specifically in the F5_SusD and SusC encoding genes using Error-prone PCR. For this assay, the F5minSus template was selected to amplify mutated SusC and SusD fragments. At the same time, the vectors corresponding to mutant genes were linearized, resulting in truncated F5minSus vectors that lacked either the SusC (F5minSus Δ SusC) or SusD (F5minSus Δ SusD) gene fractions, in order to exclude wild-type constructs from the generated mutant libraries. These linearized vectors were guaranteed to contain one of the SusC or SusD genes along with all GHs genes, and were amplified by PCR. Although we managed to amplify all mutant fragments and linearize F5minSus vectors, we were unable to complete the ligation process using the In-Fusion HD Plus cloning kit (Clontech) following the instructions in the user manual. We made multiple attempts, including optimizing PCR thermocycling conditions to obtain the most stable fragments, redesigning different primers and playing with the ratio of the In-Fusion cloning reaction components. Unfortunately, all attempts failed without exception. No rational explanation for this result can be provided, and the error-prone PCR strategy had to be abandoned.

Although the probability of significantly improving uptake efficiency by mutating transporter encoding genes is low, mutating randomly or rationally (targeting the ligand binding site and the SusC/D interacting zones) their sequences will allow us to gain insight into the relationships between structure, specificity, binding ability, and internalization efficiency of these transporters. This strategy will contribute to the development of applications of native or engineered oligosaccharide transporters in synthetic biology and in microbial engineering for enhanced polysaccharide utilization, and to the understanding of the evolutionary dynamics and of the ecological factors driving the diversification of oligosaccharide utilization strategies in bacteria.

Résumé de la thèse

Les mammifères hébergent des communautés microbiennes très diverses. Parmi celles-ci, le microbiote intestinal joue un rôle crucial dans la santé de l'hôte. Les bactéries intestinales utilisent des polysaccharides et oligosaccharides variés comme principales sources de carbone, en utilisant des systèmes protéiques complexes pour dégrader ces substrats en monosaccharides assimilables. Les bactéries du phylum Bacteroidota sont parmi les plus abondantes dans les microbiomes intestinaux des mammifères. Elles produisent une série de protéines nécessaires à la détection, à la capture, au transport et à la dégradation des polysaccharides et oligosaccharides. Ces systèmes complexes sont codés par des loci dédiés à l'utilisation des polysaccharides (PUL). A la surface cellulaire, les protéines de liaison aux glycanes (SGBPs) jouent un rôle majeur dans l'efficacité de capture des polysaccharides et des oligosaccharides entourant les cellules, offrant ainsi aux Bacteroidota un avantage compétitif dans la colonisation de leurs habitats. Dans ce travail de thèse, nous nous sommes concentrés sur la caractérisation et l'ingénierie des transporteurs de bactéries du phylum Bacteroidota, en particulier leurs composants SGBPs, issues des microbiomes digestifs humains et bovins. L'objectif était d'élucider les relations structure-fonction des différents éléments protéiques de ces transporteurs impliqués dans l'utilisation des polysaccharides d'origine alimentaire ou microbiennne. Nous avons tout d'abord caractérisé la fonction et la structure cristallographique d'une protéine de type SusD (appelée ici F5_SusD-like) codée par un PUL provenant d'un *Bacteroides* non cultivé de l'intestin humain, et qui est conservé chez *Bacteroides vulgatus*, une espèce dominante dans le microbiote intestinal. Malgré son incapacité à se lier aux xylooligosaccharides (XOS), qui sont pourtant les substrats ciblés par le F5_PUL, F5_SusD est essentiel à la fonctionnalité du transport. L'analyse structurale a révélé des boucles désordonnées et des résidus clés mal alignés, qui pourraient être responsables de l'incapacité de cette protéine de type SGBP à se lier aux XOS. Dans un second temps, nous avons étudié la spécificité des transporteurs F5_SusC/D et F5_MFS vis-à-vis des XOS en introduisant une β -mannosidase dans le périplasme des souches recombinantes d'*E. coli* produisant ces transporteurs. En utilisant des variants du locus F5 précédemment construits par Tauzin *et al.*, nous avons complété le système d'utilisation des oligosaccharides de F5 avec l'enzyme BT_0458 de la famille CAZy GH2. Nos données préliminaires suggèrent que les transporteurs du clone métagénomique F5 ne peuvent pas internaliser le mannotriose, indiquant une probable stricte spécificité de ces systèmes vis-à-vis des XOS. Par la suite, nous avons exploré le potentiel de l'évolution adaptative en laboratoire (ALE) pour l'ingénierie du système

recombinant d'utilisation des XOS chez *E. coli*. Cette étude préliminaire indique que l'évolution *in cellulo* pourrait servir à améliorer les fonctions de transporteurs recombinants chez *E. coli* et à approfondir l'étude de leurs relations structure-fonction sous pression évolutive. Enfin, nous avons caractérisé une autre protéine de type SusD (appelée 41O1_SusD) codée par un locus d'utilisation de β -glucanes et provenant d'un clone métagénomique isolé du microbiome du rumen bovin. La caractérisation fonctionnelle de cette protéine a montré que 41O1_SusD présente une affinité de liaison pour les β -1,3-glucanes ramifiés par des liaisons β -1,6. Une comparaison structurale avec des protéines homologues a permis de mettre en évidence des motifs similaires de reconnaissance du substrat, impliquant trois résidus tryptophane. Ces résultats ont permis de progresser dans la compréhension du rôle joué par les protéines de type SusD dans l'utilisation des substrats issus de la paroi cellulaire végétale et/ou d'origine fongique dans le rumen bovin.

Titre : Caractérisation et ingénierie de transporteurs d'oligosaccharides

Mots clés : Transporteurs, Oligosaccharides, Ingénierie des protéines, SusC/D , Relations structure-Fonction, Enzymes actives sur les sucres

Résumé : Les mammifères hébergent des communautés microbiennes très diverses. Parmi celles-ci, le microbiote intestinal joue un rôle crucial dans la santé de l'hôte. Les bactéries intestinales utilisent des polysaccharides et oligosaccharides variés comme principales sources de carbone, en utilisant des systèmes protéiques complexes pour dégrader ces substrats en monosaccharides assimilables. Les bactéries du phylum Bacteroidota sont parmi les plus abondantes dans les microbiomes intestinaux des mammifères. Elles produisent une série de protéines nécessaires à la détection, à la capture, au transport et à la dégradation des polysaccharides et oligosaccharides. Ces systèmes complexes sont codés par des loci dédiés à l'utilisation des polysaccharides (PUL). A la surface cellulaire, les protéines de liaison aux glycanes (SGBPs) jouent un rôle majeur dans l'efficacité de capture des polysaccharides et des oligosaccharides entourant les cellules, offrant ainsi aux Bacteroidota un avantage compétitif dans la colonisation de leurs habitats. Dans ce travail de thèse, nous nous sommes concentrés sur la caractérisation et l'ingénierie des transporteurs de bactéries du phylum Bacteroidota, en particulier leurs composants SGBPs, issues des microbiomes digestifs humains et bovins. L'objectif était d'élucider les relations structure-fonction des différents éléments protéiques de ces transporteurs impliqués dans l'utilisation des polysaccharides d'origine alimentaire ou microbienne. Nous avons tout d'abord caractérisé la fonction et la structure cristallographique d'une protéine de type SusD (appelée ici F5_SusD-like) codée par un PUL provenant d'un Bacteroides non cultivé de l'intestin humain, et qui est conservée chez Bacteroides vulgatus, une espèce dominante dans le microbiote intestinal. Malgré son incapacité à se lier aux xylooligosaccharides (XOS), qui sont pourtant les substrats ciblés par le F5_PUL, F5_SusD est essentiel à la fonctionnalité du transport. L'analyse structurale a révélé des boucles désordonnées et des résidus clés mal alignés, qui pourraient être responsables de l'incapacité de cette protéine de type SGBP à se lier aux XOS. Dans un second temps, nous avons étudié la spécificité des transporteurs F5_SusC/D et F5_MFS vis-à-vis des XOS en introduisant un β -mannosidase dans le périplasme des souches recombinantes d'E. coli produisant ces transporteurs. En utilisant des variants du locus F5 précédemment construits par Tauzin et al., nous avons complété le système d'utilisation des oligosaccharides de F5 avec l'enzyme BT_0458 de la famille CAZy GH2. Nos données préliminaires suggèrent que les transporteurs du clone métagénomique F5 ne peuvent pas internaliser le mannitriose, indiquant une probable stricte spécificité de ces systèmes vis-à-vis des XOS. Par la suite, nous avons exploré le potentiel de l'évolution adaptative en laboratoire (ALE) pour l'ingénierie du système recombinant d'utilisation des XOS chez E. coli. Cette étude préliminaire indique que l'évolution in cellulo pourrait servir à améliorer les fonctions de transporteurs recombinants chez E. coli et à approfondir l'étude de leurs relations structure-fonction sous pression évolutive. Enfin, nous avons caractérisé une autre protéine de type SusD (appelée 41O1_SusD) codée par un locus d'utilisation de β -glucanes et provenant d'un clone métagénomique isolé du microbiome du rumen bovin. La caractérisation fonctionnelle de cette protéine a montré que 41O1_SusD présente une affinité de liaison pour les β -1,3-glucanes ramifiés par des liaisons β -1,6. Une comparaison structurale avec des protéines homologues a permis de mettre en évidence des motifs similaires de reconnaissance du substrat, impliquant trois résidus tryptophane. Ces résultats ont permis de progresser dans la compréhension du rôle joué par les protéines de type SusD dans l'utilisation des substrats issus de la paroi cellulaire végétale et/ou d'origine fongique dans le rumen bovin.

Title: Characterization and engineering of oligosaccharide transporters

Key words: Transporters, Oligosaccharides, Protein Engineering, SusC/D, Structure-function relationships , Carbohydrate-active enzymes

Abstract: Mammals host very diverse microbial communities. Among these, the gut microbiota plays a crucial role in host health. Gut bacteria utilize various polysaccharides and oligosaccharides as carbon sources, using complex protein machineries to degrade these substrates into assimilable monosaccharides. Bacteroidota represents one of the dominant phyla in mammal gut microbiomes. They produce a series of proteins necessary for the sensing, capture, transport and degradation of polysaccharides and oligosaccharides. These complex machineries are encoded by polysaccharide utilization loci (PULs). Among the proteins encoded by PULs, cell surface glycan-binding proteins (SGBPs) are essential for the efficient capture of substrates surrounding the cells. In this thesis work, we focused on the characterization and engineering of Bacteroidota transporters, in particular their SGBP components, identified from the human and bovine gut microbiomes. The aim was to elucidate the structure-function relationships of the different protein elements of these transporters, involved in the utilization of dietary or microbial glycans in gut microbiomes. We first characterized the function and crystallographic structure of a SusD-like protein (referred to F5_SusD-like here) encoded by a xylooligosaccharide (XOS) PUL from an uncultured human gut Bacteroides species, which is conserved in the prominent Bacteroides vulgatus species. Despite its inability to bind to XOS, that are the cognate substrates of the F5_PUL, the F5_SusD-like protein is essential for the uptake functionality. Structural analysis revealed disordered loops and misaligned key residues, which could be responsible for the inability of this SGBP-like protein to bind XOS. Then, we investigated the specificity of F5_SusC/D and F5_MFS transporters towards XOS by introducing α -mannosidase into the periplasm of the recombinant E. coli strains harboring these transporters. This approach aimed at determining if the transporters could uptake mannitriose, which would be degraded to monosaccharides by the β -mannosidase and support E. coli growth. Using F5 variants previously constructed by Tauzin et al., we complemented the F5 oligosaccharide utilization system with the BT_0458 β -mannosidase from the GH2 CAZy family. Our preliminary data suggest that the F5 transporters cannot recognize and transport mannitriose, indicating a probable strict specificity towards XOS. Subsequently, we explored the potential of adaptive laboratory evolution (ALE) to engineer the recombinant F5_XOS utilization pathway in E. coli. After serial subcultures of the F5_XOS containing E. coli strains, the mutation rates and positions were assessed by next-generation sequencing. This preliminary study indicated that ALE could serve to improve transporter functions and provided the basis for further investigation of the structure-function relationships of oligosaccharide transporters under evolutionary pressures. Finally, we characterized another SusD-like protein (referred to 41O1_SusD) encoded by a β -glucan utilization locus from a metagenomic clone isolated from the bovine rumen microbiome. Functional characterization showed that the 41O1_SusD-like protein exhibits binding affinity for β -1,6 branched β -1,3-glucans. Structural comparison with homolog SusD-like proteins highlighted a similar pattern of substrate recognition, involving three tryptophan residues. Our findings provided advances in the understanding of the role played by SusD-like proteins in plant cell wall and/or fungal polysaccharide utilization in the cow rumen. Overall, this thesis generated advances in the understanding of the structure-function relationships of SusD-like proteins. In the long term, the findings will contribute to the development of potential applications in synthetic biology and microbial engineering for enhanced polysaccharide utilization.