Outer membrane vesicles from *Fibrobacter succinogenes* S85 contain an array of Carbohydrate-Active Enzymes with versatile polysaccharide-degrading capacity

Magnus Ø. Arntzen\textsuperscript{a*}, Anikó Vármai\textsuperscript{a}, Roderick I. Mackie\textsuperscript{b}, Vincent G. H. Eijsink\textsuperscript{a}, Phillip B. Pope\textsuperscript{a}

\textsuperscript{a}. Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Ås, Norway.

\textsuperscript{b}. Institute for Genomic Biology, and Department of Animal Sciences, University of Illinois at Urbana-Champaign, Illinois, USA

*Address correspondence to Magnus Ø. Arntzen, magnus.arntzen@nmbu.no*

Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, P.O. Box 5003, N-1432 Ås, NORWAY. Tel: +47 67 23 24 46, Fax: +47 64 96 59 01

**Running title:** *F. succinogenes* outer membrane vesicles

**Keywords:** *Fibrobacter succinogenes* S85, outer membrane vesicle, proteomics, protein complex, polysaccharide degradation
Outer membrane vesicles (OMVs) are gaining increasing attention for their role in pathogenesis and microbial ecology. OMVs provide a means to increase bacterial outreach since they allow delivery of degradation-protected biomolecules to the environment, at high local concentrations. Indeed OMVs are known to exert influences on eco-systems via horizontal gene transfer, biofilm formation, intra- and interspecies communication, and biomass degradation. Here we report that OMVs produced by *Fibrobacter succinogenes* are equipped with a diverse suite of enzymes able to depolymerize most common plant polysaccharides, including cellulose. Our data indicate that OMVs assist the metabolism of the host cell by deconstructing non-essential polysaccharides that restrict access to the host’s target carbon source, cellulose. We also demonstrate that previously identified cellulose binding proteins are arranged in novel putative complexes in OMVs. Thus, *F. succinogenes* degrades biomass using means that differ fundamentally from well-known degradative machineries in Nature.
Summary

*Fibrobacter succinogenes* is an anaerobic bacterium naturally colonizing the rumen and cecum of herbivores where it utilizes an enigmatic mechanism to deconstruct cellulose into cellobiose and glucose, which serve as carbon sources for growth. Here, we illustrate that outer membrane vesicles (OMVs) released by *F. succinogenes* are enriched with carbohydrate-active enzymes and that intact OMVs were able to depolymerize a broad range of linear and branched hemicelluloses and pectin, despite the inability of *F. succinogenes* to utilize non-cellulosic (pentose) sugars for growth. We hypothesize that the degradative versatility of *F. succinogenes* OMVs is used to prime hydrolysis by destabilizing the tight networks of polysaccharides intertwining cellulose in the plant cell wall, thus increasing accessibility of the target substrate for the host cell. This is supported by observations that OMV-pretreatment of the natural complex substrate switchgrass increased the catalytic efficiency of a commercial cellulose-degrading enzyme cocktail by 2.4-fold. We also show that the OMVs contain a putative multiprotein complex, including the fibro-slime protein previously found to be important in binding to crystalline cellulose. We hypothesize that this complex has a function in plant cell wall degradation, either by catalyzing polysaccharide degradation itself, or by targeting the vesicles to plant biomass.
Introduction

Cellulose and hemicellulose are the most abundant components of plant biomass. These polysaccharides, although recalcitrant, do not accumulate on our planet due to their removal by the concerted action of highly specialized (hemi)-cellulose degrading microbes, including fungi and bacteria. These microorganisms exploit sophisticated enzyme systems to degrade plant material, and the enzymes involved in plant cell wall degradation have potential in biotechnological applications, such as in biofuel production (Himmel et al., 2010). In aerobic cellulolytic microorganisms, cellulose degradation is catalyzed by a consortium of mostly secreted enzymes including cellobiohydrolases, endoglucanases, β-glucosidases and lytic polysaccharide monooxygenases (LPMOs) (Horn et al., 2012; Mba Medie et al., 2012). The polysaccharide-degrading enzymes release soluble oligosaccharides and sugars that are transported into the cell and further metabolized. In contrast, some anaerobic cellulolytic bacteria form large multi-enzyme complexes referred to as cellulosomes, which often are bound to the outer surface of the cells (Bayer et al., 2004; Bayer et al., 2008). These complexes contain a backbone scaffoldin protein onto which several types of cellulases are docked via dockerin domains. The scaffoldin binds to cellulose primarily through family-3 carbohydrate-binding modules (CBMs), whereas substrate-affinity may be additionally tuned by CBMs attached to the cellulosomal enzymes. Recently, a third enzyme system, the Bacteroidetes-affiliated Polysaccharide Utilization Loci (PULs), has been described, which entails physically-linked genes organized around a signature SusCD-encoding gene pair (representing an outer membrane porin and a carbohydrate-binding protein, respectively). PULs seem to predominantly target soluble glycans, but PUL-based conversion of crystalline chitin has been shown (Larsbrink et al., 2016), and there are indications that uncultured rumen populations utilize PULs to degrade cellulose (Naas et al., 2014). In addition to these
strategies, there are examples of cellulolytic enzymes being attached directly to the peptidoglycan layer (such as in *Clostridium thermocellum* (Zhao et al., 2006)) or to cell surface polysaccharides (such as in *Ruminococcus albus* (Ezer et al., 2008)) of biomass-degrading bacteria.

One of the most highly specialized cellulose-degrading bacteria is *Fibrobacter succinogenes*, a strictly anaerobic, Gram-negative, rod-shaped bacterium. It is considered one of the major cellulolytic bacteria within the herbivore gut (Krause et al., 2003; Kobayashi et al., 2008) and has been the subject of extensive research due to its ability to adhere to and efficiently degrade plant cell walls. *F. succinogenes* does not produce cellulosomes, does not secrete high titers of cellulolytic enzymes, and its genome seems devoid of genes encoding known cellobiohydrolases and PULs (Suen et al., 2011). These observations suggest that *F. succinogenes* employs an alternative strategy for cellulose degradation. To understand why *F. succinogenes* is such a powerful biomass degrader, a number of endoglucanases, xylanases and cellulose-binding proteins have been cloned and characterized (see summary in (Toyoda et al., 2009)), without revealing particularly powerful enzymes. It has been suggested that outer membrane (OM) proteins are involved in cellulose degradation (Jun et al., 2007; Raut et al., 2015), but details remain ambiguous.

In 1981 it was discovered that *F. succinogenes* releases sedimentable membranous fragments into the culture fluid, which are able to hydrolyze carboxymethylcellulose (CMC) (Groleau and Forsberg, 1981). Subsequently, it was demonstrated that the membrane fragments are in fact vesicles originating from the outer membrane (OMV: outer membrane vesicle) that are produced during growth on cellulose (Forsberg et al., 1981). The OMVs showed a distinct and complex protein composition (Groleau and Forsberg, 1983) and were shown to exhibit both endoglucanase, xylanase and acetyesterase activity (Gong and Forsberg, 1993). These studies also showed that
the OMVs adhere to cellulose and are not produced during growth on glucose (Forsberg et al., 1981; Gong and Forsberg, 1993; Burnet et al., 2015). The role of these OMVs in *F. succinogenes* is currently debated; some claim that their production merely reflects aging of the cells i.e. a stationary phase phenomenon (Gaudet and Gaillard, 1987), while others speculate that they have a biological function in cellulose degradation (Forsberg et al., 1981). Interestingly, it was recently shown that OMVs from *Bacteroides fragilis* and *Bacteroides thetaiotaomicron* are equipped with hydrolytic enzymes and are important in polysaccharide degradation (Elhenawy et al., 2014).

OMVs are spherical, bi-layered, membranous structures that are released naturally from the OM of Gram-negative bacteria (Beveridge, 1999). They are typically between 10-300 nm in diameter and contain phospholipids, liposaccharides, OM proteins and proteins from the periplasmic space. OMVs have been observed in a wide range of Gram-negative species grown in different environments and under various growth conditions (see references in (Kulp and Kuehn, 2010)). They have been suggested to play wide-ranging roles in microbial ecology (e.g. horizontal gene transfer, biofilm formation, communication and biomolecule delivery) and can be numerically far more abundant than the organisms themselves (Elhenawy et al., 2014; Roier et al., 2016).

In this study, we have isolated and studied the content of OMVs produced by *F. succinogenes* during growth on crystalline cellulose. We used proteomics to identify the proteins in the OMVs and show they are enriched in polysaccharide-degrading enzymes. Importantly, we demonstrate the presence of a novel putative multiprotein complex, comprising several proteins known to be involved in interactions with cellulose, that could be a driver of polysaccharide degradation. Activity assays showed that the OMVs are able to depolymerize a broad range of hemicelluloses in addition to cellulose, and use of OMVs as pretreatment of a natural grass substrate (switchgrass)
enabled a 2.4-fold increase in downstream saccharification. The results add support to the hypothesis that *F. succinogenes* actively uses OMVs to convert biomass.

## Results

OMVs produced by *F. succinogenes* S85 vary in size and are equipped with carbohydrate-active enzymes

Similar to other Gram-negative bacteria and according to previous reports, *F. succinogenes* S85 produces OMVs (Forsberg et al., 1981; Burnet et al., 2015), but currently little is known about their specific nature and enzyme contents. To isolate OMVs, we employed a series of microfiltration and ultra-centrifugation steps and vesicles were obtained as a broad, strong band in a sucrose gradient, with an average density of 1.13 g/mL. The band broadness suggested a heterogeneous size distribution, which was confirmed by dynamic light scattering experiments that indicated a population ranging from 8-136 nm in radius, with an average of 49 nm (Figure S1A). Transmission electron microscopy (TEM) confirmed that the OMV preparation contained vesicles (Figure S1B).

Using quantitative proteomics, we detected 347 proteins in the OMVs covering a range in abundance of four orders of magnitude (Table S1) and with high reproducibility between biological replicates (Pearson correlation $R = 0.805$) (Figure S2). Using an algorithm for predicting signal peptides, lipoprotein signal peptides and transmembrane helices (LipoP; see Supplementary Text S2), 79% of the detected proteins were predicted to be associated with the extracellular milieu. In particular, 50% harbored a SpI signal peptide, 28% an SpII lipoprotein signal peptide, and 1% contained a transmembrane helix. The remaining 21% were predicted to be cytosolic.
proteins. We performed functional annotation of the complete proteome of *F. succinogenes* (2871 protein sequences) via protein searches and categorical classification using the NCBI Conserved Domain Database (NCBI Web-CD) and the database of Clusters of Orthologous Groups (COG) of proteins (see Supplementary Text S2). This analysis revealed that the OMVs showed a higher proportion of proteins in the COG-category ‘carbohydrate transport and metabolism’, which covered 12% of the OMV proteome, compared to 4% in the complete proteome (Figure S3).

Analysis of the OMV proteins using dbCAN, a specialized database for prediction of carbohydrate-active enzymes (CAZymes; (Yin et al., 2012)), showed that 21% of the OMV proteins (i.e. 74 of the 347 proteins) had predicted carbohydrate-active functions (Figure 1A). Comparing these numbers with predicted extracellular proteins in *Fibrobacter* (992 proteins with either Spl or SpII cleavage sites or containing a TMH, according to LipoP, of which 116 are CAZymes), suggests an enrichment of carbohydrate-active enzymes in the OMVs (Fisher’s Exact p-value 1.36E-5). Forty-eight were classified as glycoside hydrolases (GHs), two as glycosyl transferases (GTs), five as polysaccharide lyases (PLs) and 13 as carbohydrate esterases (CEs), while no auxiliary activities (AAs) were identified. In addition, we found six proteins that contain a carbohydrate-binding module (CBM), but lack a catalytic domain with a known carbohydrate-active function. Figure 1A shows these 74 proteins plotted against their relative abundance in the OMV proteome. The most abundant protein (FSU_2303) belongs to the GH family 8 and could be responsible for hydrolyzing the backbone of cellulose and xylan. Amongst the 50 most abundant proteins in the OMVs, there are seven CAZymes (one GH8, two GH9 and four GH5; Table S1). The six CBM-only proteins show similar abundances as the catalytic CAZymes, and one of these, a CBM11 (FSU_2007), is highly abundant. These proteins could be interesting to investigate further for the presence of hitherto unknown carbohydrate-active catalytic domains.
To look further into the enrichment of certain proteins in the OMVs, we performed an enrichment analysis using Pfam, a tool for predicting functional domains in proteins. First, we counted the occurrence of all the Pfam domains in the complete proteome of *F. succinogenes* and then compared these values to similar values for the OMVs. Using Fisher’s Exact test to calculate the significance of enrichment, we detected 18 domains to be overrepresented in the OMVs, half of which were CAZyme-domains (Table 1, Figure S4). The most frequent of the enriched Pfam domains was the family-6 CBM (PF03422), which is known to target amorphous cellulose or xylan. In the OMV proteins, this module is found associated with GH5 endoglucanases (PF00150), GH43 (PF04616) and GH30 (PF17189) hemicellulases, and a sialic acid-specific acetylesterase (PF03629), indicating involvement in degradation of both cellulose and hemicellulose. Another enriched CBM, the family-11 CBM (PF03425), is known to target amorphous cellulose and is appended to a GH51 endoglucanase domain (e.g. in FSU_0382) or to a GH5 endoglucanase domain (e.g. in FSU_2914) or occurs as a single domain protein (FSU_2007). In addition to CBMs, several endoglucanases (GH5 and GH9: PF00759) and hemicellulase (GH16: PF00722, GH30 and GH43) domains were enriched, indicating a potential role of vesicles in delivering carbohydrate-active enzymes to the substrate. Notably, the analysis of Pfam domains revealed the enrichment of several non-carbohydrate-active domains, some of which are potentially involved in carbohydrate-binding or metabolism, as discussed below.

**OMVs are active on a wide range of plant-derived substrates**

To explore the actual enzymatic activity present in the OMVs, we incubated the purified vesicles with nine different substrates: phosphoric acid swollen cellulose (PASC) made from Avicel, tamarind xyloglucan, cabbage pectin, wheat arabinoxylan, birchwood xylan, aspen xylan, ivory nut mannan, carob galactomannan and konjac glucomannan. The products formed by substrate
hydrolysis were identified by LC-MS using a library of m/z-time tags (combination of measured mass and retention time) established on a high-sensitive mass spectrometer connected to a HPLC. The dbCAN analysis (Figure 1A) predicted OMV proteins that target these abovementioned substrates, namely: endoglucanases (e.g. GH5s), xyloglucanases (e.g. GH74s), pectin lyases (e.g. PL1s), endo-xylanases (e.g. GH11s and GH43s linked to xylan-binding CBM6s) and mannanases (e.g. GH26s). In accordance with the prediction, we detected formation of oligosaccharide products from each substrate (Figure 1B: I-IX, details in Table S2), indicating that the OMVs are able to degrade the plant cell wall polysaccharides tested.

Fresh forages, including green leaves and stems, are commonly found in the rumen of pasture fed ruminants, the natural habitat for *F. succinogenes*. These are rich in primary cell walls which are mainly composed of cellulose, xyloglucan and pectin, where the two latter polysaccharides cross-link cellulose microfibrils (Park and Cosgrove, 2015). The activity assays with tamarind xyloglucan and cabbage pectin revealed that the OMVs contain enzymes that are able to break these polymers, which theoretically would yield improved access to cellulose, the breakdown products of which serve as the main carbon source for growth of *F. succinogenes*. The OMV proteins cleaved xyloglucan not only into its repeating units (cellotetraose backbone with three xylosyl substitution, e.g. Hex$_4$Pen$_5$; Hex: hexose, Pen: pentose) but also into fragments with a shorter backbone (e.g. Hex$_1$Pen$_1$, and Hex$_2$Pen$_2$; Figure 1B: IX). The occurrence of xyloglucan oligosaccharides carrying less than three pentose units (most likely xylosyl substitutions) indicates cleavage of the xyloglucan backbone between two substituted glucosyl units. This unique cleavage pattern has only been shown for a handful of enzymes belonging to the GH74 and AA9 families so far (Desmet et al., 2007; Feng et al., 2014; Kojima et al., 2016; Nekiunaite et al., 2016) and
could potentially be attributed to FSU_2866, an OMV protein annotated as a BNR repeat protein and predicted to harbor four GH74 modules (Table S1).

Incubation of cabbage pectin, a mixture of homogalacturonan (partly methyl esterified polygalacturonic acid) and rhamnogalacturonan type I (a rhamnose-galacturonic acid copolymer substituted with arabinogalactan side chains), with the OMVs led to fragmentation of various structural elements of pectin (Figure 1B:VII). The formation of galacturonic acid oligosaccharides containing an unsaturated galacturonic acid revealed the cleavage of homogalacturonan by \( \beta \)- elimination with a pectate lyase. In the OMVs, five proteins with polysaccharide lyase domains were identified (belonging to PL families 1, 9 and 22), of which one has been identified as being potentially active on pectin (FSU_0577, putative pectate lyase) (Table S1). OMVs could also depolymerize the arabinogalactan side chains of rhamnogalacturonan moieties. While we did not identify rhamnose-containing oligosaccharides (indicative of cleavage of the rhamnogalacturonan backbone), oligosaccharides that are likely to originate from the arabinogalactan side chains were observed (\( \text{Hex}^2-4 \), \( \text{Pen}^2-14 \), \( \text{Hex}^3-5\text{Pen}^2 \), corresponding to \( \text{Gal}^2-4 \), \( \text{Ara}^2-14 \), \( \text{Gal}^3-5\text{Ara}^2 \), respectively; \( \text{Gal} \): galactose, \( \text{Ara} \): arabinose). The OMV proteins performing this action could be FSU_3024 (identified as a GH53 arabinogalactan endo-\( \beta \)-1,4-galactanase), FSU_0145 (a GH43 arabinosidase) and FSU_2288 (a GH2 \( \beta \)-1,4-galactosidase).

The OMVs were also active on cellulose releasing cellobiose, cellotriose and cellotetraose from PASC (Figure 1B:VIII). The OMVs were able to depolymerize close to 70% of the PASC within 24 hours at a reasonable enzyme loading (2.4 mg OMV proteins with predicted carbohydrate-active function was loaded per g cellulose). The initial depolymerization rate was 1.8 U/mg/min (i.e. one mg enzyme releases 1.8 \( \mu \)mol reducing end sugars during one minute incubation); using the same conditions, the activity of the commercial enzyme cocktail Celluclast by Novozymes
(Bagsvaerd, Denmark) was determined to be 2.5 U/mg/min. For more details, see Experimental Procedures.

The OMVs were also capable of depolymerizing substituted hemicelluloses. The most common hemicelluloses in grasses, commonly fed to ruminants, are branched xylans. In reactions with arabinoxylan (Figure 1B:VI), where the β-1,4-xylan backbone is 3-O-mono- or 2,3-O-disubstituted with α-L-arabinose, we observed a range of oligosaccharides with a degree of polymerization (DP) up to 10. Although elution times indicated hydrolysis products were not linear oligosaccharides, we were unable to identify arabinosylation patterns of the released oligosaccharides, because arabinose and xylose have the exact same mass and are undistinguishable by mass spectrometry. The OMVs were also active on xylans with different substituting groups that are more common in woody plant cell walls. In the reactions with birchwood xylan (Figure 1B:IV), we detected three types of xylo-oligosaccharides: linear, substituted with 4-O-methyl-glucuronic acid and substituted with glucuronic acid. Reactions with aspen xylan (Figure 1B:V) showed release of xylooligosaccharides carrying methyl-glucuronyl and/or acetyl groups.

The OMVs were also shown to target mannans with various backbone and substitution patterns. The fact that the OMVs were able to depolymerize ivory nut mannan (Figure 1B:I) (a linear mannose homopolymer) to mannoooligosaccharides shows the presence of true mannanases that can cleave β-1,4-linkages between two mannose units in the polymer backbone. Galactosylation (as in carob galactomannan) of the mannan backbone did not prevent depolymerization (Figure 1B:II) and yielded both linear (nongalactosylated) mannoooligosaccharides and galactosylated oligosaccharides. From konjac glucomannan (acetylated glucomannan), the OMVs released both
cello- and mannooligosaccharides and a range of glucomannan oligosaccharides (Figure 1B:III). In addition, mono- and diacetylated oligosaccharides were detected.

Finally, as *Fibrobacter succinogenes* grows exclusively on cellulose, which in plant cell walls is embedded in a hemicellulose and pectin network, we hypothesized that a potential function for OMVs (carrying hemicellulose- and pectin-degrading enzymes) could be to increase the accessibility of cellulose in grasses in the rumen by loosening up the pectin-hemicellulose matrix localized around the cellulose fibers. To test this, we selected a milled and washed switchgrass substrate and compared its degradability with a commercial enzyme cocktail with and without pretreatment with OMVs. Pretreatment with OMVs had a significant effect on saccharification of switchgrass by the commercial enzyme cocktail, leading to a 2.4-fold increase in the solubilized sugar yield as compared to when the commercial enzymes were acting alone (Figure 2A). MS analysis of the products formed during switchgrass degradation revealed that even though the commercial enzyme mixture was able to depolymerize both pentose and hexose-containing sugars, auxiliary enzyme activities were present in the OMVs leading to additional products formed (Figure 2B). Perhaps the most important difference is the formation of uGalA-GalA2 (GalA: galacturonic acid; u: unsaturated) and acetylated oligosaccharides containing both hexose and pentose units. The latter most likely originate from plant xyloglucan since only arabinogalactan of the other hemicelluloses contain both sugar types and arabinogalactan has not been shown to carry any acetyl groups. The appearance of these compounds suggests that the OMVs were able to hydrolyze the homogalacturonan backbone in pectin using lyase activities (hence the unsaturated galacturonic acid) as well as heavily substituted (acetylated and probably also fucosylated) xyloglucans, and that they thus likely open up the intertwined pectin-hemicellulose-cellulose network. These observations may explain why OMV-pretreatment increases the saccharification
of switchgrass by the commercial enzyme cocktail and strengthen the hypothesis that the primary role of OMVs may be to provide *F. succinogenes* better access to cellulose.

**OMVs also contain putative multiprotein complexes**

To detect potential protein complexes amongst the OMV proteins, we utilized high-resolution clear native electrophoresis (hrCNE). This technique, which takes advantage of mixed micelles to stabilize proteins and convey a negative net charge, has proven to separate equally well compared to blue native electrophoresis, while being superior for downstream catalytic activity assays (Wittig et al., 2007). In combination with SDS-PAGE, it is possible to generate two-dimensional gels in which the protein complexes separated in the first dimension (hrCNE) are separated into single protein spots in the second dimension. Proteins originating from the same complex will fall on a straight vertical line. Figure 3A shows such a 2D-hrCN-SDS-PAGE separation of 40 µg OMV proteins using a 6.5% native gel and 10% SDS-gel. 15 protein spots were selected for proteomics analysis (Table 2). Three putative protein complexes can be seen: complex C1: spot number 3, 4 and 5, complex C2: spot number 6, 7, 8 and 9, and complex C3: spot number 11, 12, 13 and 14.

The C2 and C3 putative complexes seem to contain at least some identical proteins (Figure 3) as was indeed confirmed by the proteomic analysis (Table 2). The main difference between the putative complexes is the lack of spot number 9 in C3. This may indicate that the complex could exist in two variants, with or without the protein(s) in spot 9, or that a part of the complex was lost during sample preparation. Considering only the most abundant proteins in each spot, complex C1 consists of two proteins with no predicted functional domains (FSU_1029, FSU_2008) and one OmpA family protein (FSU_2078) harboring a C-terminal OmpA-like domain and five thrombospondin type 3-like repeats, which are known to bind calcium (Kvansakul et al., 2004). Complex C2 consists of four proteins, two OmpA family proteins (FSU_2396, FSU_2078), a
tetradicopeptide repeat (TPR) domain protein (FSU_2397) and a fibro-slime domain protein (FSU_2502). Spot 9 was broad and dense, and found to contain many proteins (Table S3), including several endoglucanases. It is not possible to judge whether all these spot 9 proteins are part of the C2 complex. Regardless, the emPAI values clearly show that the fibro-slime domain protein is the dominating protein in spot 9. Notably, we have consistently observed spot 9 to co-occur with spots 6, 7 and 8, independent of the acrylamide percentage in the first dimension (data not shown); this indicates a true association of the proteins in these spots. Complex C3 seems to be a fragment of C2, containing only two of the proteins, the OmpA family protein (FSU_2396) and the TPR domain protein (FSU_2397). Strikingly, these two proteins, which are partners in both complex C2 and C3, are neighboring genes located in an operon, according to the Database of prokaryotic operons (DOOR; (Mao et al., 2009)), and show co-expression with high abundance in the OMV total data set (Table S1).

To assess the carbohydrate degrading capabilities of these putative complexes, we used another lane from the native gel, identical to the one used for the SDS-PAGE separation, and divided it into seven fractions as indicated on the top of Figure 3A. The gel pieces were ground using a pestle and mortar and then incubated with PASC for detection of enzyme activity. The products were analyzed using PGC-MS, and the amounts of the different oligosaccharide products were determined (Figure 3B). All fractions, except fraction VII gave release of cello- and xylooligosaccharides from PASC. The first two fractions (I and II) released oligosaccharides to a low extent, suggesting that complex C1 has a limited role in cellulose degradation. Fractions III - VI, including complexes C2 and C3, all produced high amounts of oligosaccharides. Notably, separation is not optimal due to horizontal streaking in the first dimension, meaning that it is impossible to assign activities to particular protein complexes or individual proteins. It is
interesting to note that Fraction V, lacking the fibro-slime protein seems less active on cellulose. No products were detected in fraction VII, indicating that this protein, *F. succinogenes* major paralogous domain protein (FSU_2794), is not able to degrade PASC under these conditions.

**Discussion**

OMVs are formed by membrane blebbing, followed by release of spherical outer membrane vesicles, which enclose a fraction of the periplasmic space. Vesiculation appears to be a common phenomenon for Gram-negative bacteria (Beveridge, 1999; Roier et al., 2016), suggesting an important physiological role for this process (Kulp and Kuehn, 2010). OMVs represent a confined transportable environment where enzymes, virulence factors or other molecules are protected from inhibitors and can be present in high concentrations (Biller et al., 2014). Recent reports have shown that OMVs from *B. fragilis* and *B. thetaiotaomicron* are selectively packaged with acidic hydrolases and proteases compared with the outer membrane, which contains more alkaline proteins (Elhenawy et al., 2014). This suggests a sorting mechanism that could be pI related. Interestingly, the vast majority (79%) of the 347 OMV proteins detected in *F. succinogenes* were also acidic (pI < 7; for the whole proteome this fraction was 64%). Further, we used Pfam-based analysis to detect domains that were overrepresented in the OMVs. In this analysis, we compared the domains present in the OMV proteome to the whole cell’s proteome and found that half of the domains that were statistically enriched (Fisher’s Exact p-value < 0.05) were CAZyme-domains. It is important to note however, that comparing the OMV proteome to the whole cell’s proteome is not necessarily a fair comparison as OMV proteins are expected to be biased to the outer membrane, and enriching for extracellular proteins *de facto* does select for CAZymes (as enzymes taking part in polysaccharide-degradation are almost exclusively extracellular). To account for this bias, we compared the fraction of CAZymes present in the OMVs (21%) with that in the predicted
extracellular proteome of *F. succinogenes* (12%). This supported our hypothesis that CAZymes are enriched in the OMVs (Fisher’s Exact p-value 1.36E-5) and suggests a selective packaging of carbohydrate-active enzymes into OMVs.

Reports have shown that *F. succinogenes* releases OMVs during growth on cellulose, but not during growth on glucose (Forsberg et al., 1981; Burnet et al., 2015). Producing OMVs is an energy-demanding task for the bacteria, and given the ubiquitous presence of OMVs across Gram-negative species, the selective sorting of acidic proteins and the enrichment of carbohydrate-metabolizing proteins observed here, it is reasonable to assume that the OMVs have important biological functions.

*F. succinogenes* is widely known to efficiently hydrolyze the variety of plant polysaccharides it encounters with in the rumen. The current data shows that OMVs produced by *F. succinogenes* are equipped with enzymes targeting these polysaccharides, in accordance with previous observations (Gong and Forsberg, 1993). OMVs were capable of hydrolyzing nine different isolated plant polysaccharides, but also showed activity on a more complex and natural substrate, switchgrass, whereby OMV-pretreatment increased the efficiency of a commercial cellulase cocktail 2.4-fold. We speculate this is due to complementary enzyme activities present in the OMVs that enhance cellulose accessibility. The promiscuous activity of the OMVs towards plant polysaccharides that are embedded with cellulose (the sole carbon source of the host), suggest that a primary role of OMVs could be to provide *F. succinogenes* better access to cellulose.

An analysis of the most enriched protein families in the OMVs revealed several without a CAZyme annotation, yet with high abundance in the OMVs (Table 1). Some of these domains have properties that suggest potential involvement in carbohydrate binding or metabolism. This includes the PA14 domain, a hypothesized carbohydrate-binding module found in a wide variety of
enzymes including glycosidases, and the sulfatase-modifying factor enzyme, which belongs to the lectin-like superfamily. Furthermore, type IV pilin proteins and cadherins were highly abundant in the OMV proteome. A detailed discussion on these domains and their potential contribution to carbohydrate binding or metabolism is provided in Supplementary Text S1.

It has been well documented that *F. succinogenes* does not utilize any of the known polysaccharide-degrading assemblages (i.e. cellulosomes or PULs) (Suen et al., 2011). In this study, we observed high levels of TPR domain proteins in the OMVs, a protein class also observed by others in the outer membrane (Jun et al., 2007; Raut et al., 2015). TPR proteins are commonly found in protein complexes, where multiple TPR domains (three in FSU_2397) have been shown to form a super-helix exposing several binding surfaces that promote formation of multiprotein complexes (Zeytuni and Zarivach, 2012). TPR proteins are consequently believed to act as scaffold proteins (Blatch and Lassle, 1999). This led us to investigate if multiprotein complexes were present in the OMVs. Our analyses revealed the presence of at least three putative multiprotein complexes in the OMVs, two of which, C2 and C3, seemingly degraded PASC. The four main components of these two complexes (FSU_2078, FSU_2502, FSU_2396 and FSU_2397) are all predicted to be secreted. Both putative complexes lack known glycoside hydrolases among their main “highly-detectable” components, although proteomic analysis detected hydrolytic enzymes in the samples, either as “contaminations” or as less abundant parts of the complexes. Interestingly, all four main proteins identified in these putative complexes have previously been detected on the outer membrane of *F. succinogenes*, and accumulating data indicate that they play a role cellulose binding (Gong et al., 1996; Jun et al., 2007; Raut et al., 2015). The abundantly present fibro-slime domain protein (FSU_2502), previously referred to as the 180-kDa cellulose-binding protein, is known to have an important role in cellulose binding (Gong et al., 1996; Suen et al., 2011). Hence,
it is likely that this protein helps targeting the vesicles to plant biomass. Notably, the four main proteins in C2 and C3 together contain hypothetical regions summing up to approximately 3000 amino acids with unknown functions, which could include hitherto unknown hydrolytic enzymes. In particular, the FSU_2396 OmpA protein contains a beta-helix domain similar to that seen in pectate lyases. In *F. succinogenes*, this domain (Pfam PF13229) is found in only one other protein (FSU_2273), a pectate lyase with a family-6 CBM, also detected in the OMVs. Figure 4A shows the domain organizations for the four proteins involved in complexes C2 and C3, whereas Figure 4B depicts an artist impression of a putative OMV-associated complex acting on the substrate.

Interestingly, in 2009, Toyoda and colleagues identified cellulose-binding proteins in rumen fluid from sheep through enrichment with crystalline cellulose (Toyoda et al., 2009). The authors detected four proteins belonging to *F. succinogenes*: a TPR domain protein (FSU_2397), a fibrilslime domain protein (FSU_2502), an OmpA family protein (FSU_2396) and cellulose binding protein (FSU_0382). Except from the latter (which we did detect in the OMVs), these proteins are part of complex C2. These observations considered collectively with earlier reports of the importance of these proteins for cellulose binding (Gong et al., 1996; Jun et al., 2007) and the operon structure of the genes encoding FSU_2396 and FSU_2397 indicate that C2 is a real complex with an important role in biomass conversion.

In conclusion, *F. succinogenes* is equipped with a surprisingly high diversity of polysaccharide-degrading enzymes and abilities, considering that the bacterium only utilizes one such polysaccharide, cellulose, as a carbon source. The observation that *F. succinogenes* packages many of these enzymes into OMVs that are released as “degrading drones” makes the bacterium even more peculiar. The exact role of OMV formation needs to be further explored for several key purposes, including the identification of signals that trigger OMV biogenesis, to understand the
impact of OMV formation on *F. succinogenes* fitness, and to explore syntrophic OMV interactions with other microbes. To this end, it is interesting to note previously observed OMV-related syntrophic interactions in the human intestine, specifically between OMV-producing polysaccharide-degrading bacteria and bacteria unable to grow on the specific polysaccharide alone (Rakoff-Nahoum et al., 2014). We hypothesize that *F. succinogenes* secretes the OMVs when grown on cellulose to degrade the surrounding hemicellulose, thus making cellulose more accessible. In electron micrographs presented by Burnet and colleagues (Figure 7D in (Burnet et al., 2015)), the OMVs are not found between the cells and the cellulose fibers, but rather distant from the cells, supporting the idea that the vesicles are paving the way for the bacterium. This idea is strengthened by our observation that the OMVs make switchgrass more susceptible to degradation by a commercial cellulose cocktail. Perhaps OMVs are useful tools for industrial biomass saccharification and/or as agents in mild biological biomass pretreatment methods. We predict that further studies on the OMVs from *F. succinogenes* will improve our understanding of the lifestyle of this enigmatic microbe, including its ability to efficiently degrade plant cell walls.

**Experimental Procedures**

**Culture conditions and isolation of OMVs**

*Fibrobacter succinogenes* S85 (ATCC 19169) cultures were grown statically at 37°C under anaerobic conditions, in the medium recommended by ATCC (ATCC medium 1943). Details of the medium can be found in Supplementary Text S2.

For isolation of OMVs, 800 mL cultures were grown. After 24 hours, the cultures were harvested by centrifugation at 9,000 × g for 15 minutes. The supernatant was filtered (0.45µm) and concentrated to 100 mL using a Vivaflow 200 cartridge with 10 kDa cut off (Sartorius AG,
Goettingen, Germany) and further down to 4 mL using a centrifugal concentrator with 100 kDa
cut off (Pall Life Sciences, Ann Arbor, MI, USA). The retentate was centrifuged at 16,600 × g for
20 minutes to remove any debris. After a second filtration (0.45 µm), the supernatant was layered
on top of a sucrose gradient and centrifuged at 200,000 × g for 3 hours. The brown, strong band,
containing the OMVs, was extracted using a needle and syringe, diluted to 12 mL with 10 mM
sodium acetate buffer (pH 6.0) containing 100 mM NaCl and re-centrifuged at 100,000 × g for 1
hour. The supernatant was discarded and the pellet (containing the OMVs) was collected and
resuspended in 10 mM sodium acetate buffer (pH 6.0). The protein concentration in the OMV
preparation was measured using Bradford protein assay and the OMVs were analyzed for size and
purity using dynamic light scattering (DLS) and transmission electron microscopy (TEM). For
details, see Supplementary Text S2.

Native and SDS-PAGE gels

For native gel electrophoresis, we prepared a 6.5% resolving (Tris/HCl pH 8.8, polyacrylamide)
gel with a 5% stacking (Tris/HCl pH 6.8, polyacrylamide) gel. The anode buffer consisted of 25
mM Tris/HCl buffer (pH 8.3) containing 192 mM glycine, while the cathode buffer contained in
addition 0.02% n-dodecyl-β-D-maltoside (DDM) and 0.05% sodium deoxycholate (DOC). The
mixed micelles formed by the non-ionic detergent DDM and the anionic detergent DOC has been
shown to stabilize membrane proteins while also providing a negative charge on the proteins
(hence the anionic detergent), resulting in high-resolution clear native electrophoresis at pH 8.3,
even for alkaline proteins (Wittig et al., 2007). Samples were prepared in a sample buffer (pH 8.3)
containing 10% glycerol, 0.001% ponceau S, 50 mM NaCl, 25 mM Tris/HCl, and 40 µg OMV
proteins were loaded per lane. Electrophoresis was performed at 4 °C and 200 V for 50 minutes.
For 2D-hrCN-SDS-PAGE, a homemade 10% resolving SDS-gel were prepared and a lane already
separated under native conditions (above) were excised and placed 10 mm above the SDS-gel. A 5% stacking gel were poured around the native lane so this would be embedded into the stack. Electrophoresis was performed at 240 V for 20 minutes and the gels were then stained with Coomassie Brilliant Blue R250. 15 spots (gel pieces; see Figure 3A) were excised and destained twice using 25 mM ammonium bicarbonate in 50% acetonitrile. The proteins entrapped in the gel pieces were reduced and carbamidomethylated using 10 mM DTT and 55 mM iodacetamide, respectively, prior to in-gel digestion with trypsin as described previously (Arntzen et al., 2015).

Prior to mass spectrometry, peptides were desalted using C_{18} ZipTips (Merck Millipore, Darmstadt, Germany), according to manufacturer’s instructions.

For proteomic analysis of total OMVs, two biological replicates were used. 50 µg of protein were dissolved in SDS sample buffer, separated by SDS-PAGE using an AnyKD Mini-PROTEAN gel (Bio-Rad Laboratories, Hercules, CA, USA) and stained using Coomassie Brilliant Blue R250. The gel was cut into eight slices and the slices were processed as described above.

**Proteomics and bioinformatics analysis**

Peptides were analyzed using a nanoLC-MS/MS system (Dionex Ultimate 3000 UHPLC; Thermo Scientific, Bremen, Germany) connected to a Q-Exactive mass spectrometer (Thermo Scientific, Bremen, Germany) and operated in data-dependent mode to switch automatically between orbitrap-MS and higher-energy collisional dissociation (HCD) orbitrap-MS/MS acquisition. MS raw files were analyzed using MaxQuant (Cox and Mann, 2008) and identifications were filtered in order to achieve a protein false discovery rate (FDR) of 1%. Only proteins identified in both biological replicates were considered true OMV proteins. For analysis of gel spots, we used the Mascot search engine (Perkins et al., 1999) to provide protein identifications. For further details
on the proteomics methods and for bioinformatics (LipoP prediction, COG and Pfam analysis), see Supplementary Text S2.

**Enzymatic assays and PGC-MS analysis**

To estimate the efficiency of depolymerization of the OMV preparation, 10 mg OMV proteins (corresponding to 2.4 mg carbohydrate-active enzymes based on the proteomics abundance measurements) were loaded per g of PASC. The reaction was carried out in 50 mM sodium acetate buffer, pH 6.0, for 48 hours in triplicates; samples were taken after 1, 4, 24 and 48 hours. After sampling, H$_2$SO$_4$ was added (4% final concentration), and the samples were autoclaved for 60 min at 121 °C to hydrolyze the oligosaccharides to monosugars (Sluiter et al., 2006). The sugar yield was measured as reducing sugars using 3,5-dinitrosalicylic acid (Miller, 1959). The activity (i.e. initial rate) was calculated based on the total reducing sugars at 1 hour and expressed as U/mg/min.

Enzymatic assays with OMVs were done using 20 µg OMV proteins and 1% (w/v) substrate in 10 mM sodium acetate buffer (pH 6.0). Nine different substrates were used: phosphoric acid swollen cellulose (PASC), birchwood xylan, wheat arabinoxylan, aspen xylan, ivory nut mannan, carob galactomannan, konjac glucomannan, tamarind xyloglucan and pectin. PASC from Avicel, pectin from white cabbage and aspen xylan (isolated under mild conditions to avoid autohydrolysis of the acetyl groups during the isolation process (Biely et al., 2013)) were prepared as described earlier (Wood, 1988; Westereng et al., 2009; Biely et al., 2013), birchwood xylan was purchased from Roth (Karlsruhe, Germany) and all other substrates were purchased from Megazyme (Wicklow, Ireland). Enzyme reactions were performed overnight at 40 °C and supernatants containing soluble products were collected by centrifugation at 16,600 × g for three minutes. Control reactions showed that no substrate depolymerization occurred upon incubation of the substrates in buffer, at 40 °C, in the absence of OMVs, except for ivory nut mannan, carob
galactomannan and konjac glucomannan. In these cases, the signal obtained in control reactions were used for background subtraction of the samples.

The products were analyzed using a HPLC system (Dionex Ultimate 3000RS UHPLC; Thermo Scientific, Bremen, Germany) equipped with a porous graphitic carbon (PGC) column (Hypercarb) and connected to an LTQ-Velos Pro ion trap mass spectrometer (Thermo Scientific, Bremen, Germany). Product identification was achieved using \( m/z \)-values provided by the Velos Pro mass spectrometer or, in ambiguous cases, a mixture of retention time and \( m/z \)-values. For details on the analysis of products, see Supplementary Text S2.

**Assessment of OMVs for biomass pretreatment**

Switchgrass, obtained from The Noble Foundation, Ardmore OK, USA, was ball milled at 350 rpm in consecutive series of 10 minutes on and 15 minutes off to keep the temperature <50 °C. After 1.5 hours, the ball milled switchgrass was washed two times with water to remove background color and any soluble sugars prior to usage. Pretreatment assays were done using this ball milled, washed switchgrass at 0.2% (w/v) with 20 µg OMVs for 17.5 hours in 50 mM sodium acetate buffer (pH 6.0). Celluclast (mainly cellulase activity) and Novozym 188 (mainly \( \beta \)-glucosidase activity), both purchased from Novozymes (Bagsvaerd, Denmark), were prepared as a mixture in the ratio 4:1 (w/w) for enzymatic degradation of switchgrass. Enzyme reactions (after pretreatment) were performed by adding 20 µg enzyme cocktail to the above conditions and further incubate for four hours at 40 °C. Supernatants containing soluble products were collected by centrifugation at 16,600 × g for three minutes. The products were analyzed as reducing sugars using 3,5-dinitrosalicylic acid as reagent (Miller, 1959).

**Funding information**
MOA and PBP were supported by the European Research Council through grant 336355 ("MicroDE").

Acknowledgements

The authors would like to thank Dr. Roger Scherrers at Wyatt (Dernbach, Germany) for analysis of vesicles using dynamic light scattering and Dr. Bjørge Westereng (NMBU, Norway) for helpful discussions. The imaging was performed at the Imaging Centre Campus Ås, Department of Plant Sciences, NMBU, Norway. The proteomics data has been deposited to the ProteomeXchange consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (Vizcaíno et al., 2013) with the dataset identifier PXD005442.
References


Desmet, T., Cantaert, T., Gualfetti, P., Nerinckx, W., Gross, L., Mitchinson, C., and Piens, K.


**Table 1. Enriched Pfam domains in OMVs.** The table shows Pfam domains, which were found to be overrepresented (Fisher’s Exact p < 0.05) in the OMVs compared to the complete genome. A more detailed display of these domains, and the proteins harboring them, can be found in Figure S4.

<table>
<thead>
<tr>
<th>Pfam Accession</th>
<th>Pfam Name</th>
<th>Pfam Description</th>
<th>CAZy Family</th>
<th>Genome sequences</th>
<th>OMV sequences</th>
<th>Enrichment</th>
<th>Average Log2(LFQ intensity)</th>
<th>Fisher's Exact p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF03422</td>
<td>CBM_6</td>
<td>Carbohydrate binding module (family 6)</td>
<td>CBM6</td>
<td>24</td>
<td>16</td>
<td>67%</td>
<td>23.8</td>
<td>0.000001</td>
</tr>
<tr>
<td>PF03781</td>
<td>FGE-sulfatase</td>
<td>Sulfatase-modifying factor enzyme 1</td>
<td></td>
<td>20</td>
<td>11</td>
<td>55%</td>
<td>24.1</td>
<td>0.000231</td>
</tr>
<tr>
<td>PF00150</td>
<td>Cellulase</td>
<td>Cellulase (glycosyl hydrolase family 5)</td>
<td>GH5</td>
<td>12</td>
<td>8</td>
<td>67%</td>
<td>26.5</td>
<td>0.000673</td>
</tr>
<tr>
<td>PF03425</td>
<td>CBM_11</td>
<td>Carbohydrate binding domain (family 11)</td>
<td>CBM11</td>
<td>5</td>
<td>5</td>
<td>100%</td>
<td>27.8</td>
<td>0.002282</td>
</tr>
<tr>
<td>PF04616</td>
<td>Glyco_hydro_43</td>
<td>Glycosyl hydrolases family 43</td>
<td>GH43</td>
<td>13</td>
<td>7</td>
<td>54%</td>
<td>24.0</td>
<td>0.003610</td>
</tr>
<tr>
<td>PF00759</td>
<td>Glyco_hydro_9</td>
<td>Glycosyl hydrolase family 9</td>
<td>GH9</td>
<td>7</td>
<td>5</td>
<td>71%</td>
<td>26.2</td>
<td>0.005971</td>
</tr>
<tr>
<td>PF09479</td>
<td>Flg_new</td>
<td>Listeria-Bacteroides repeat domain (List_Bact_rpt)</td>
<td></td>
<td>4</td>
<td>4</td>
<td>100%</td>
<td>24.5</td>
<td>0.006596</td>
</tr>
<tr>
<td>PF00691</td>
<td>OmpA</td>
<td>OmpA family</td>
<td></td>
<td>13</td>
<td>6</td>
<td>46%</td>
<td>28.0</td>
<td>0.012202</td>
</tr>
<tr>
<td>PF07691</td>
<td>PA14</td>
<td>PA14 domain</td>
<td></td>
<td>6</td>
<td>4</td>
<td>67%</td>
<td>25.4</td>
<td>0.016598</td>
</tr>
<tr>
<td>PF17189</td>
<td>Glyco_hydro_30C</td>
<td>Glycosyl hydrolase family 30 beta sandwich domain</td>
<td>GH30</td>
<td>4</td>
<td>3</td>
<td>75%</td>
<td>24.8</td>
<td>0.031468</td>
</tr>
<tr>
<td>PF00722</td>
<td>Glyco_hydro_16</td>
<td>Glycosyl hydrolases family 16</td>
<td>GH16</td>
<td>4</td>
<td>3</td>
<td>75%</td>
<td>22.2</td>
<td>0.031468</td>
</tr>
<tr>
<td>PF02321</td>
<td>OEP</td>
<td>Outer membrane efflux protein</td>
<td></td>
<td>4</td>
<td>3</td>
<td>75%</td>
<td>23.9</td>
<td>0.031468</td>
</tr>
<tr>
<td>PF09603</td>
<td>Fib_succ_major</td>
<td>Fibrobacter succinogenes major domain</td>
<td></td>
<td>52</td>
<td>12</td>
<td>23%</td>
<td>25.5</td>
<td>0.038435</td>
</tr>
<tr>
<td>PF13568</td>
<td>OMP_b-brl_2</td>
<td>Outer membrane protein beta-barrel domain</td>
<td></td>
<td>5</td>
<td>3</td>
<td>60%</td>
<td>27.9</td>
<td>0.046392</td>
</tr>
<tr>
<td>PF13174</td>
<td>TPR_6</td>
<td>Tetratricopeptide repeat</td>
<td></td>
<td>5</td>
<td>3</td>
<td>60%</td>
<td>26.0</td>
<td>0.046392</td>
</tr>
<tr>
<td>PF02927</td>
<td>CelD_Na</td>
<td>Cellulase N-terminal ig-like domain</td>
<td>GH9</td>
<td>5</td>
<td>3</td>
<td>60%</td>
<td>26.4</td>
<td>0.046392</td>
</tr>
<tr>
<td>PF03629</td>
<td>SASA</td>
<td>Carbohydrate esterase, sialic acid-specific acetylesterase</td>
<td>CE</td>
<td>5</td>
<td>3</td>
<td>60%</td>
<td>22.8</td>
<td>0.046392</td>
</tr>
<tr>
<td>PF01103</td>
<td>Bac_surface_Ag</td>
<td>Surface antigen</td>
<td></td>
<td>5</td>
<td>3</td>
<td>60%</td>
<td>25.2</td>
<td>0.046392</td>
</tr>
</tbody>
</table>

*aThe N-terminal ig-like domain of GH9
Table 2. Identification of proteins from the 2D-hrCN-SDS-PAGE gel. The table shows the proteomic detection of proteins in spots numbered 1 to 15 in Figure 3A. Only the most abundant protein in each spot is shown. More details, including the proteins present at lower abundance, can be found in Table S3.

<table>
<thead>
<tr>
<th>Complex</th>
<th>Spot no.</th>
<th>UniProt</th>
<th>Locus</th>
<th>Name</th>
<th>Mascot Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A7UG45</td>
<td>FSU_2404</td>
<td>Membrane protein</td>
<td>7320</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>A7UG46</td>
<td>FSU_1029</td>
<td>Membrane protein</td>
<td>2730</td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>3</td>
<td>A7UG46</td>
<td>FSU_1029</td>
<td>Membrane protein</td>
<td>8773</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>A7UG37</td>
<td>FSU_2078</td>
<td>OmpA family protein</td>
<td>5372</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>C9RRD7</td>
<td>FSU_2008</td>
<td>Uncharacterized protein</td>
<td>3832</td>
</tr>
<tr>
<td>C2</td>
<td>6</td>
<td>A7UG61</td>
<td>FSU_2396</td>
<td>OmpA family protein</td>
<td>4747</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>A7UG37</td>
<td>FSU_2078</td>
<td>OmpA family protein</td>
<td>4389</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>A7UG62</td>
<td>FSU_2397</td>
<td>TPR domain protein</td>
<td>5100</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>A7UG66</td>
<td>FSU_2502</td>
<td>Fibro-slime domain protein</td>
<td>22351</td>
</tr>
<tr>
<td>C3</td>
<td>10</td>
<td>C9RL47</td>
<td>FSU_0797</td>
<td>Uncharacterized protein</td>
<td>3784</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>A7UG61</td>
<td>FSU_2396</td>
<td>OmpA family protein</td>
<td>6853</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>A7UG61</td>
<td>FSU_2396</td>
<td>OmpA family protein</td>
<td>4207</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>A7UG62</td>
<td>FSU_2397</td>
<td>TPR domain protein</td>
<td>6181</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>A7UG62</td>
<td>FSU_2397</td>
<td>TPR domain protein</td>
<td>4282</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>C9RKA2</td>
<td>FSU_2794</td>
<td>F. succinogenes major paralogous domain protein</td>
<td>5056</td>
</tr>
</tbody>
</table>

Figure 1: CAZymes detected in the OMVs and their predicted activities on various plant-derived substrates. A) The figure shows the 74 proteins that could be annotated as carbohydrate-active enzymes (CAZymes) plotted against their relative abundances in the OMVs. Proteins carrying a carbohydrate-binding module (CBM) are colored as indicated in the figure. B) The figure shows nine different substrates susceptible to hydrolysis by the OMVs. Each substrate is annotated with CAZymes found in the OMVs, predicted cleavage sites that are based on literature and the detected products. The products detected by PGC-MS after overnight incubation of intact OMVs with the substrates are illustrated in miniature beneath the substrates, with numbers indicating the degree of polymerization (DP). A detailed list of products can be found in Table S2. Signals obtained by incubation of the substrates in buffer at 40 °C in the absence of OMVs, were used for background subtraction. The shown substrates are: I: Ivory nut mannan, II: Carob galactomannan, III: Konjac glucomannan, IV: Birchwood xylan, V: Aspen xylan, VI: Wheat arabinoxylan, VII: Cabbage pectin (including RG, rhamnogalacturonan), VIII: PASC from Avicel cellulose, IX: Tamarind xyloglucan.

Figure 2: Effect of OMV treatment on the enzymatic digestibility of switchgrass. A) Ball milled and washed switchgrass at 0.2% (w/v) was incubated without (OMV –) or with 20 µg OMVs (OMV +) for 17.5 hours at 40 °C, then subjected to enzymatic degradation by 20 µg 4:1 (w/w) mixture of Celluclast and Novozym 188. B) Products detected by PGC-MS after overnight incubation of switchgrass either with a mixture of Celluclast and Novozym 188 or with intact OMVs are shown with numbers indicating the degree of polymerization (DP). An explanation of
symbols can be found in Figure 1B and a detailed list of products can be found in Table S2. The sugars were assigned as hexoses and pentoses when it was not possible to obtain a clear identification using the combination of retention time and m/z-value.

Figure 3: 2-dimensional electrophoresis to analyze protein complexes. A) 2D-hrCN-SDS-PAGE gel; protein complexes are thought to migrate as an intact complex in the first dimension. In the second dimension, the complex disintegrates due to SDS, and the subunits are separated and appear on a vertical line. The gel suggests the presence of three major complexes, C1, C2 and C3. Note that the region of the gel marked by I, is the stacking region of the 1st dimension and will not separate proteins. Spots 1 and 2, which also show divergent electrophoretic elution pattern, are therefore not likely in a complex. B) Enzymatic activity on PASC. Another native gel lane, identical to the one used in panel A, was divided into seven pieces (Fractions I – VII) and grinded prior to measurement of enzyme activity. Linear cello- and xylo-oligosaccharides were separated, identified and quantified by PGC-MS, and the amounts of the different products are displayed in the form of a heat map. Note that the amounts are relative per compound due to differences in the efficiency of ionization among the compounds. The scale corresponds to the integrated peak area as reported by the Xcalibur software. Glc: glucose, Xyl: xylose.

Figure 4: Visualization of the possible mode of action of complex C2. A) The four proteins identified as complex C2, two OmpA family proteins (FSU_2078, FSU_2396), a TPR domain protein (FSU_2397) and a fibro-slime protein (FSU_2502), likely form a putative OMV-associated complex (B), where the fibro-slime protein mediates binding to cellulose (Gong et al., 1996) and
the TPR acts as a protein scaffold for complex assembly (Blatch and Lassle, 1999). The beta helix
domain in FSU_2396 OmpA, similar to that seen in pectate lyases, could be responsible for
hydrolyzing the substrate.
**SUPPLEMENTARY MATERIAL LEGENDS**

**Figure S1:** Outer membrane vesicles isolated from *F. succinogenes* after growth on Avicel cellulose. A) Data from dynamic light scattering, revealing a broad population of OMVs with radii ranging from 8-136 nm, with an average of 49 nm. The small peak observed at 3.4 nm represents micelles formed by the *n*-dodecyl-β-D-maltoside detergent present in the sample at 0.05% (w/v). B) Visualization of OMVs by negatively stained transmission electron microscopy; the picture shows vesicles with a diameter of around 100 nm. Scale bar: each white/black subsection is 40 nm.

**Figure S2:** Proteomics reproducibility. The figure shows the comparison of the two replicates analysed by quantitative proteomics, showing high reproducibility with Pearson correlation R=0.805. The axes represent log2(LFQ) values obtained in each replicate and the colors represent the number of peptides associated with each protein.

**Figure S3:** Clusters of orthologous groups (COG) analysis. The figure shows the mapping of COG functions to proteins in the complete proteome of *F. succinogenes* as well as to the proteins detected in the OMVs. The figure shows a higher proportion of carbohydrate-active enzymes in OMVs compared to the complete proteome.

**Figure S4:** Pfam enrichment analysis. The figure shows the proteins associated with each of the 18 enriched Pfam modules listed in Table 2, and the relative abundance of these proteins.
[log2(LFQ) values ranging from 20.6 (low abundant; light green) to 33.4 (high abundant; light red)]. The data is shown as a hierarchically clustered heat map where shades of red indicate the presence of a given Pfam module in a protein (light red: one occurrence, dark red: four occurrences; note that there are no proteins with three occurrences). Grey indicates the absence of listed Pfam modules.

Table S1. Proteins identified in F. succinogenes OMVs. The table shows the MaxQuant output for identified proteins and the LFQ intensities used for quantification. Protein annotations were done using the LipoP server, the peptidase database MEROPS and Pfam, while CAZy predictions were done using dbCAN. PEP: Posterior error probability.

Table S2. Products detected after OMV-mediated hydrolysis of various substrates. The table shows oligosaccharides identified by PGC-MS after overnight incubation of intact OMVs with nine different plant-derived substrates and one natural substrate (switchgrass). Product assignments were based on m/z values from the mass spectrometer and retention times on the PGC column. The compositions of the substrates are also provided. Hex: hexose, Pen: pentose, Gal: galactose, GalA: galacturonic acid, uGalA: unsaturated galacturonic acid, Xyl: xylose, Glc: glucose, Man: mannose, Ara: arabinose, GlcA: glucuronic acid, Me: methylated, Ac: acetylated, DP: degree of polymerization.

Table S3. Proteins identified in the 2D-hrCN-SDS-PAGE gel. The table shows the proteomic detection of proteins in spots numbered 1 to 15 in Figure 3A, providing additional information
next to the list of most abundant proteins in Table 2. emPAI values provided by the Mascot search engine were used to obtain quantitative estimates within each protein spot, and were used to filter out low abundant hits. Only hits with emPAI > 5 are shown; the most abundant protein in each spot is printed in bold face. Note that emPAI values are not normalized, meaning that abundance levels can only be compared within single spots.

Supplementary Text S1: Enriched non-CAZy domains and other OMV proteins with potential involvement in carbohydrate binding or metabolism. This text extends the Discussion by highlighting how the PA14 domain and the sulfatase-modifying factor enzyme could be involved in carbohydrate binding and/or metabolism. Type IV pilin proteins and cadherins are also discussed.

Supplementary Text S2. Methodological details. This text extends the Experimental Procedures section with details regarding the growth medium, dynamic light scattering, transmission electron microscopy, the proteomics and bioinformatics analysis as well as the analysis of products from enzymatic assays. In particular, the use of HPLC and mass spectrometry setups are explained in detail.
**A**

Sugar equivalents (mM)

- **OMV -**
  - n=4-6
  - n=3-4
  - n=1-2
  - n=1-2
  - n=0-1

- **OMV +**
  - n=2-6
  - n=1,3
  - n=3-4
  - n=2-4

**B**

- **Commercial mixture**
  - Blue circles: Hexose
  - Blue triangles: Pentose

- **OMVs**
  - Blue circles: Hexose
  - Blue triangles: Pentose
  - White circles: Hexose
  - White triangles: Pentose

- Blue circles: Hexose
- Blue triangles: Pentose
- White circles: Hexose
- White triangles: Pentose

OMV - or OMV +