



**CHALMERS**  
UNIVERSITY OF TECHNOLOGY

## **Proteomic Dissection of the Cellulolytic Machineries Used by Soil-Dwelling Bacteroidetes**

Downloaded from: <https://research.chalmers.se>, 2026-04-04 15:46 UTC

Citation for the original published paper (version of record):

Taillefer, M., Arntzen, M., Henrissat, B. et al (2018). Proteomic Dissection of the Cellulolytic Machineries Used by Soil-Dwelling Bacteroidetes. *mSystems*, 3(6): 1-16.  
<http://dx.doi.org/10.1128/mSystems.00240-18>

N.B. When citing this work, cite the original published paper.



# Proteomic Dissection of the Cellulolytic Machineries Used by Soil-Dwelling *Bacteroidetes*

Marcel Taillefer,<sup>a</sup>  Magnus Ø. Arntzen,<sup>b</sup> Bernard Henrissat,<sup>c</sup> Phillip B. Pope,<sup>b</sup>  Johan Larsbrink<sup>a</sup>

<sup>a</sup>Wallenberg Wood Science Center, Department of Biology and Biotechnology, Chalmers University of Technology, Gothenburg, Sweden

<sup>b</sup>Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Ås, Norway

<sup>c</sup>Architecture et Fonction des Macromolécules Biologiques, CNRS, Aix-Marseille University, Marseille, France

**ABSTRACT** Bacteria of the phylum *Bacteroidetes* are regarded as highly efficient carbohydrate metabolizers, but most species are limited to (semi)soluble glycans. The soil *Bacteroidetes* species *Cytophaga hutchinsonii* and *Sporocytophaga myxococcoides* have long been known as efficient cellulose metabolizers, but neither species conforms to known cellulolytic mechanisms. Both species require contact with their substrate but do not encode cellulosomal systems of cell surface-attached enzyme complexes or the polysaccharide utilization loci found in many other *Bacteroidetes* species. Here, we have fractionated the cellular compartments of each species from cultures growing on crystalline cellulose and pectin, respectively, and analyzed them using label-free quantitative proteomics as well as enzymatic activity assays. The combined results enabled us to highlight enzymes likely to be important for cellulose conversion and to infer their cellular localization. The combined proteomes represent a wide array of putative cellulolytic enzymes and indicate specific and yet highly redundant mechanisms for cellulose degradation. Of the putative endoglucanases, especially enzymes of hitherto-unstudied glycoside hydrolase family, 8 were abundant, indicating an overlooked important role during cellulose metabolism. Furthermore, both species generated a large number of abundant hypothetical proteins during cellulose conversion, providing a treasure trove of targets for future enzymology studies.

**IMPORTANCE** Cellulose is the most abundant renewable polymer on earth, but its recalcitrance limits highly efficient conversion methods for energy-related and material applications. Though microbial cellulose conversion has been studied for decades, recent advances showcased that large knowledge gaps still exist. Bacteria of the phylum *Bacteroidetes* are regarded as highly efficient carbohydrate metabolizers, but most species are limited to (semi)soluble glycans. A few species, including the soil bacteria *C. hutchinsonii* and *S. myxococcoides*, are regarded as cellulose specialists, but their cellulolytic mechanisms are not understood, as they do not conform to the current models for enzymatic cellulose turnover. By unraveling the proteome setups of these two bacteria during growth on both crystalline cellulose and pectin, we have taken a significant step forward in understanding their idiosyncratic mode of cellulose conversion. This report provides a plethora of new enzyme targets for improved biomass conversion.

**KEYWORDS** carbohydrate-active enzymes, cellulase, cellulose, proteomics, soil microbiology

**P**roduction of biofuels and other commodities from renewable sources has steadily increased in importance as environmentally favorable alternatives to fossil-derived products. First-generation biofuel production, i.e., biofuel derived from starch and


Received 5 October 2018 Accepted 2 November 2018 Published 20 November 2018

**Citation** Taillefer M, Arntzen MØ, Henrissat B, Pope PB, Larsbrink J. 2018. Proteomic dissection of the cellulolytic machineries used by soil-dwelling *Bacteroidetes*. *mSystems* 3:e00240-18. <https://doi.org/10.1128/mSystems.00240-18>.

**Editor** William Orsi, Woods Hole Oceanographic Institution

**Copyright** © 2018 Taillefer 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 Johan Larsbrink, [johan.larsbrink@chalmers.se](mailto:johan.larsbrink@chalmers.se).

 The cellulose-degrading enzyme setups of two bacteria that do not conform to known cellulolytic mechanisms have been mapped using label free quantitative proteomics, showing how these cellulose specialists use highly redundant enzyme systems.

sucrose, has been criticized with respect to potential socioeconomic and environmental issues due to direct competition with land usage for crop growth (1), and further development of second-generation biofuels from lignocellulosic biomass is desirable (2). Lignocellulose primarily consists of cellulose, hemicelluloses, and lignin, where cellulose represents the most abundant polymer, though its crystalline nature makes it highly recalcitrant to enzymatic degradation. Microorganisms have evolved complex systems to degrade cellulose: noncomplexed enzyme systems involving the secretion of cellulases into the environment as well as complexed systems involving (cell surface-attached) multienzyme complexes (cellulosomes) (3, 4). Both of these strategies are assumed to require both  $\beta$ -1,4-endoglucanases (EGs) (to cleave amorphous cellulose chains) and *exo*-acting cellobiohydrolases (CBHs), which attack the cellulose chains from either the reducing end or the nonreducing end (5–7). Product inhibition of EGs and CBHs by the main product, cellobiose, is alleviated by  $\beta$ -glucosidases (BGs), which cleave di- and oligosaccharides into glucose. In the Carbohydrate-active enzymes database CAZy ([www.cazy.org](http://www.cazy.org) [8]), EGs are typically found in glycoside hydrolase (GH) families GH5 and GH9; CBHs in GH6, GH7, and GH48; and BGs in GH1 and GH3. The discovery of lytic polysaccharide monooxygenases (LPMOs) that cleave glycosidic bonds via oxidation and substantially boost the action of the aforementioned hydrolytic enzymes has greatly improved our understanding of microbial turnover of lignocellulose (9–13).

*Cytophaga hutchinsonii* and *Sporocytophaga myxococcoides* are two Gram-negative, aerobic, mesophilic, cellulose-degrading bacterial species belonging to the *Bacteroidetes* phylum (14–17). Both species were isolated almost a century ago but remain poorly studied, and their mechanisms of carbon source utilization are unclear. A striking feature of *C. hutchinsonii* and *S. myxococcoides* is their ability to efficiently grow on crystalline cellulose despite an apparent lack of both CBHs and LPMOs in their respective genomes (14–16, 18–20). While both species encode several EGs, these apparently do not account for their efficient cellulose utilization (19–22). Further, their cellulolytic machineries are known to be cell associated but do not contain any cellulosome-associated dockerin or scaffoldin proteins (14, 18, 20, 23–25). In several *Bacteroidetes* species, so-called polysaccharide utilization loci (PULs) encode the required proteins and enzymes for degradation, capture, and import of targeted carbohydrates, both soluble and crystalline (26–28). In *C. hutchinsonii*, a protein pair homologous to the canonical *Bacteroides thetaiotaomicron* SusC/D proteins (sugar transport and capture; PUL identifiers), is found in the genome, though disruption of the encoding genes did not result in an observable phenotype regarding growth on cellulose (29), indicating that PULs are not used for this process.

Both *C. hutchinsonii* and *S. myxococcoides* are viewed as cellulose specialists, as they are extremely limited in their carbon source utilization profiles. Apart from cellulose, both species can utilize cellobiose and glucose, but this requires extensive adaptation periods and, likely, mutational events (30, 31). *S. myxococcoides* has been shown to disrupt the hemicellulose xylan, though it is incapable of utilizing either the polysaccharide or its main constituent xylose as the sole carbon source (30, 32). Both species are immobile in liquid cultures but exhibit rapid gliding motility on solid surfaces similar to that seen with the more extensively studied *Bacteroidetes* species *Flavobacterium johnsoniae* (14, 15, 17, 20, 33). Many genes associated with motility and polysaccharide degradation in *F. johnsoniae* are linked to its type 9 secretion system (T9SS). The T9SS in *C. hutchinsonii* is also essential in polysaccharide utilization, as mutating the T9SS *sprP* component abolishes the ability to utilize crystalline cellulose (34). Further, many genes putatively related to cellulose utilization in *C. hutchinsonii* contain carboxy-terminal domains (CTDs; TIGRFAMs TIGR04183 and TIGR04131), which direct the proteins for transport through the outer membrane by the T9SS (35, 36), and a strong link can be postulated between the T9SS and motility and cellulose utilization in both *C. hutchinsonii* and *S. myxococcoides*.

In previous investigations of *C. hutchinsonii*, individual cellulases were identified and studied through mutagenesis and biochemical characterization (22, 37–42). The mode

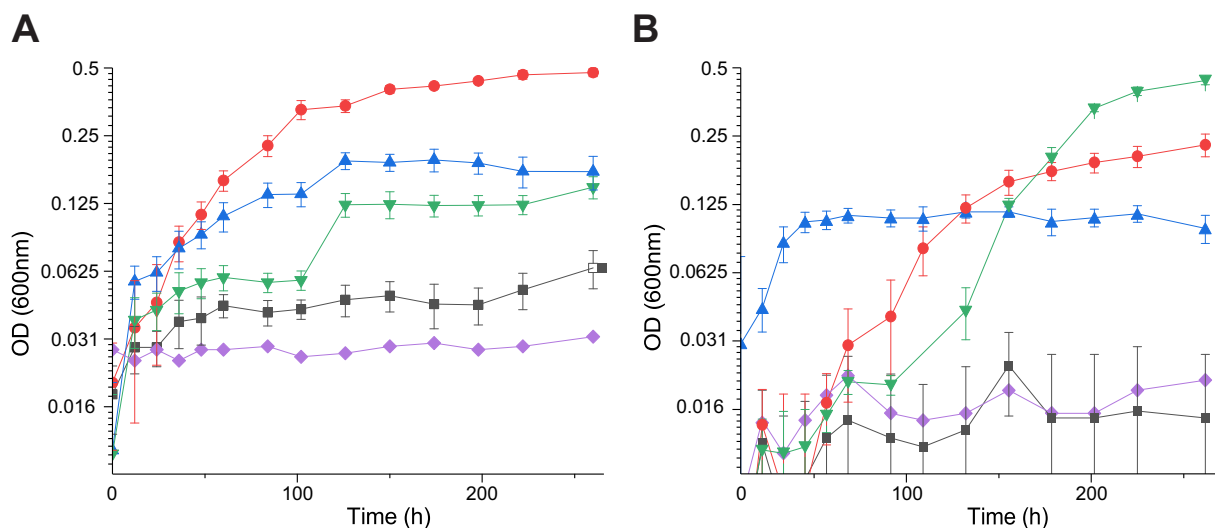
of cellulose turnover by this bacterium remains enigmatic, however, and similar studies of *S. myxococcoides* are nonexistent. As the cellulolytic systems of both species may encompass industrially relevant enzymes and/or present a new paradigm in microbial cellulose turnover, we chose to investigate the entire proteomes of both species during growth. Proteomic analyses were performed on different cellular compartments during both early and late growth on cellulose as well as pectin as a non-cellulosic/glucose-based control, and the analyses were further coupled to enzyme activity analyses.

## RESULTS

**Carbohydrate utilization and carbohydrate-active enzyme repertoires.** The carbohydrate-active enzyme (CAZyme) repertoire of *C. hutchinsonii* is available in the CAZy database (see Table S1 in the supplemental material) and consists of 189 unique proteins. The genome of *S. myxococcoides* was analyzed using the CAZy pipeline (8), and the presence of 289 putative CAZymes, including putative cellulases (GH5, GH8, and GH9), was revealed (Table S2). As expected, no putative CBHs or LPMOs were identified. *S. myxococcoides* encodes eight putative BGs (GH1 and GH3), which were predicted to be intracellular. The bacterium encodes nine putative GH5 enzymes (*C. hutchinsonii* encodes five); all but one of these were predicted to be extracellular enzymes. Also, genes encoding putative GH8 and GH9 enzymes were more numerous in the *S. myxococcoides* genome (11 GH8-encoding genes and 8 GH9-encoding genes) than in the *C. hutchinsonii* genome (6 GH8-encoding genes and 7 GH9-encoding genes). Despite the presence of multiple GH8 members, the possible involvement of GH8 enzymes in cellulose conversion had not been reported in previous studies of *C. hutchinsonii* (20, 43). The GH8 enzymes of both species were predicted to be either extracellular and soluble or localized in the outer membrane, whereas the GH9 enzymes were predicted to be either extracellular or periplasmic. *S. myxococcoides* also encodes multiple putative hemicellulose- and pectin-modifying enzymes (Table S2).

While both *C. hutchinsonii* and *S. myxococcoides* are regarded as cellulose specialists, their respective encoded abundances of CAZymes indicate that they possess the ability to deconstruct various polysaccharides. No comprehensive growth data on polysaccharides other than cellulose are available in the literature; thus, both species were assayed for growth on multiple mono- and polysaccharides. Both *C. hutchinsonii* and *S. myxococcoides* were able to grow on filter paper (20, 37, 43), pectin, and mannose. To our knowledge, this represents the first reported study of *C. hutchinsonii* or *S. myxococcoides* growing on any polysaccharide other than cellulose. The *C. hutchinsonii* strain used (G2, glucose adapted) was, as expected, able to grow on glucose, but *S. myxococcoides* exhibited a lag phase of several days before growing on this monosaccharide (Fig. 1) (30, 31). Interestingly, despite the ability to metabolize mannose, neither organism was able to utilize mannan polysaccharides (galactomannan and glucomannan), and both were further unable to utilize xylose, xylan, or xyloglucan as carbon sources, despite the  $\beta$ -1,4-glucan backbone of the latter. During growth on mannose, curiously, *S. myxococcoides* formed yellow spheres, which later disintegrated, leading to a sudden increase in turbidity (Fig. 1). The growth of both species on filter paper (crystalline cellulose) led to initial colonization of the substrate without an apparent increase in the turbidity of the liquid media due to the strong cell attachment to the solid cellulose filaments. Growth on the solid substrate continued until the paper strips disintegrated, causing the cells to be released into the medium in a manner analogous to previously reported data on filter paper utilization (37, 43). Regardless of the carbon source, both *C. hutchinsonii* and *S. myxococcoides* grew relatively slowly in liquid media but exhibited rapid growth on agar plates overlaid with filter paper strips.

**Analysis of the *C. hutchinsonii* proteome.** The proteome of *C. hutchinsonii* was obtained from cells grown with either filter paper (cellulose) or pectin (non-glucose-based control) as the sole carbon source, from both early and late growth. During the time period of late growth on filter paper, the cellulose had visibly disintegrated, with the bacteria being in solution rather than surface attached. In addition to the total proteome, proteins from the late growth stages were fractionated to investigate



**FIG 1** Semi-log growth curves of *C. hutchinsonii* (A) and *S. myxococcoides* (B) in MMM with cellobiose (■), glucose (●), pectin (▲), mannose (▼), or no carbon source (◆). Error bars represent standard deviations of results from triplicate experiments. OD, optical density.

protein abundances in different cellular compartments, including the abundances of extracellular soluble, periplasmic, inner membrane, and outer membrane proteins. The fractionation procedure and protein cellular localizations were verified by enrichment of homologues of the marker proteins *ompA* (outer membrane protein A), outer membrane (44); *ptsS* (phosphate ABC transporter phosphate-binding protein), periplasm (45); and *nuoL* (NADH-quinone oxidoreductase subunit L), inner membrane (46) (Table S3). The abundances of individual proteins were determined using label-free quantitative (LFQ) proteomic analyses (47). The total number of detected proteins from cells grown on filter paper (1,301 proteins) was much greater than the number from pectin-grown cells (337 proteins); similarly, the total number of CAZymes detected was much larger in filter paper-grown cells (50 proteins) than in pectin-grown cells (11 proteins) (Table S1; see also Fig. S1 in the supplemental material).

**Endoglucanases.** The five putative GH5 enzymes encoded by *C. hutchinsonii* were all identified in the proteome in the early stage of growth on filter paper; three of the five (locus tag identifiers *C. hutchinsonii* 1107 [CHU\_1107], CHU\_1842, and CHU\_2149) remained in the late stage of growth (Table 1; see also Fig. S1). Two enzymes, CHU\_1107 and CHU\_1727, were also present in both the early and late growth stages on pectin. Based on abundances in the cellular fractions, both CHU\_1107 and CHU\_1842 were predicted to be inner membrane-bound proteins despite CHU\_1107 containing a CTD. Further, CHU\_2103 and CHU\_2149 were associated with the outer membrane whereas CHU\_1727 was abundant in the extracellular soluble fraction. CHU\_2103 has been shown to be abundant in membrane enrichments and soluble when cell lysates were separated (37), which, together with the proteomic data, suggests that CHU\_2103 is an outer membrane protein facing the periplasm. CHU\_2103 mutants have shown moderate reductions in the ability to digest crystalline cellulose (22, 37, 48). Our observation that CHU\_2103 was present only when cells were attached to solid filter paper indicates a role in crystalline cellulose conversion (Table 1; see also Fig. S1).

Five of the seven GH9 enzymes of *C. hutchinsonii* were detected in the proteome (Table 1; see also Fig. S1). A double-disruption mutant of the genes encoding the GH9 enzyme CHU\_1280 and the GH5 enzyme CHU\_2103 (see above) was shown to be unable to digest cellulose (37). CHU\_1280 was detected in both the early and late stages of growth on filter paper and in the early stage of growth on pectin (Table 1; see also Fig. S1). Despite previous reports showing a periplasmic localization (37), CHU\_1280 was shown in the proteome analysis to be equally abundant in the

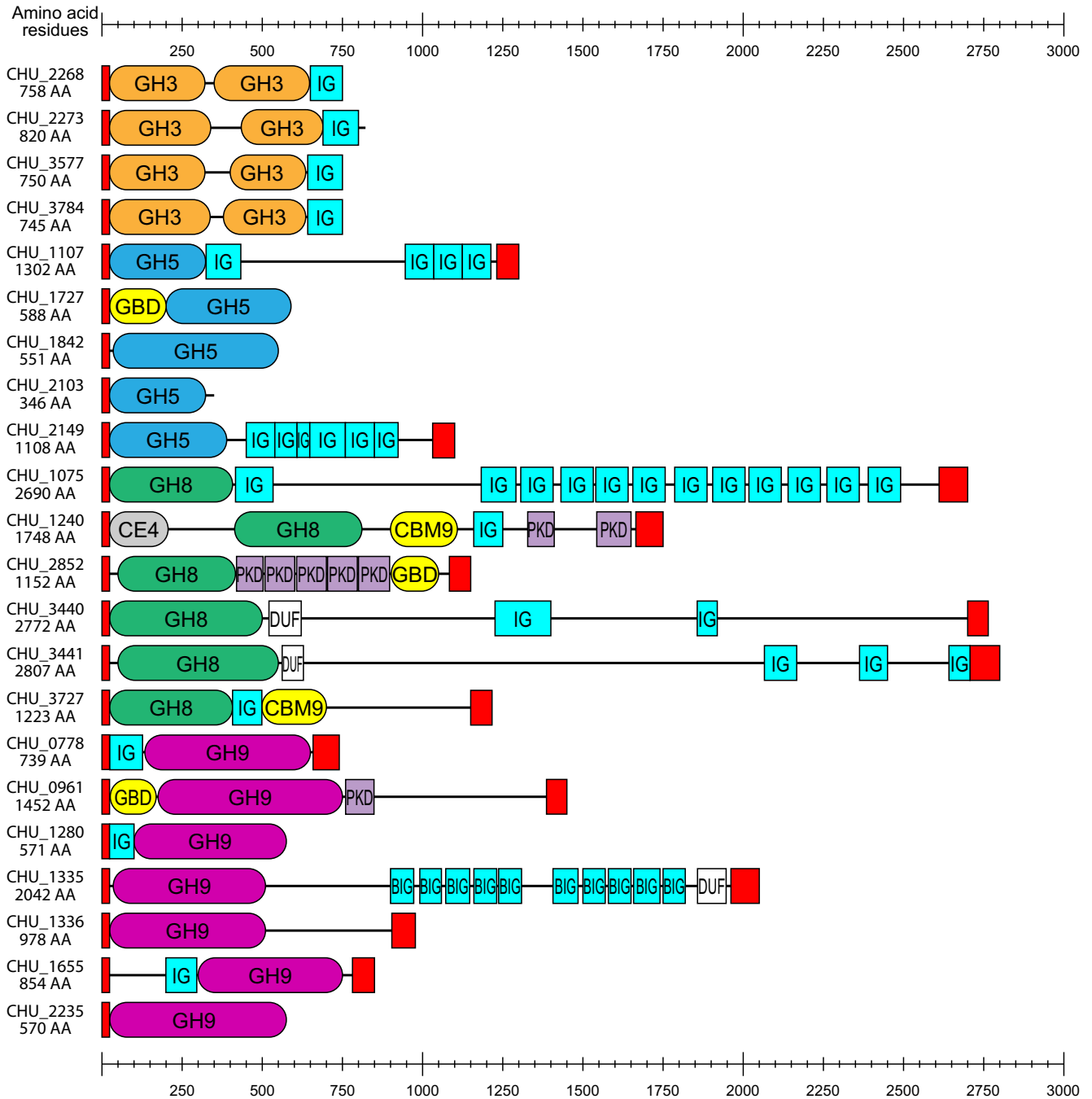
**TABLE 1** Predicted endoglucanases and  $\beta$ -glucosidases from *C. hutchinsonii* displaying  $\log_{10}$  LFQ intensity values during early and late phases of growth with filter paper or pectin as the sole carbon source<sup>a</sup>

|   | Locus Tag | CAZY domains   | Early Filter Paper | Late Filter Paper | Early Pectin | Late Pectin | Pred. Local | Local    | T9SS |
|---|-----------|----------------|--------------------|-------------------|--------------|-------------|-------------|----------|------|
| <b>GH1 (<math>\beta</math>-glucosidase)</b> |           |                |                    |                   |              |             |             |          |      |
| <b>bglA</b>                                 | CHU_3811  | GH1            | ND                 | ND                | ND           | ND          | C           | ND       | NO   |
| <b>GH3 (<math>\beta</math>-glucosidase)</b> |           |                |                    |                   |              |             |             |          |      |
| <b>bglA</b>                                 | CHU_2268  | GH3            | 5.89               | 5.84              | ND           | ND          | P, L        | IM       | NO   |
| <b>bglB</b>                                 | CHU_2273  | GH3            | 7.37               | 7.21              | 6.41         | 5.67        | P, L        | IM       | NO   |
| <b>bglC</b>                                 | CHU_3577  | GH3            | ND                 | ND                | ND           | ND          | P           | ND       | NO   |
| <b>bglD</b>                                 | CHU_3784  | GH3            | 4.68               | ND                | ND           | ND          | P           | ND       | NO   |
| <b>GH5 (endoglucanase)</b>                  |           |                |                    |                   |              |             |             |          |      |
| <b>cel5A</b>                                | CHU_1107  | GH5            | 6.52               | 6.64              | 5.67         | 5.59        | S           | IM       | YES  |
| <b>cel5C</b>                                | CHU_1727  | GH5            | 5.58               | ND                | 5.42         | 5.52        | P           | S        | NO   |
| <b>cel5D</b>                                | CHU_1842  | GH5            | 5.92               | 5.8               | ND           | ND          | OM, L       | IM       | NO   |
| <b>cel5B</b>                                | CHU_2103  | GH5            | 5.61               | ND                | ND           | ND          | S           | OM       | NO   |
| <b>cel5E</b>                                | CHU_2149  | GH5            | 5                  | 5.44              | ND           | ND          | S           | OM       | YES  |
| <b>GH8 (endoglucanase)</b>                  |           |                |                    |                   |              |             |             |          |      |
|   | CHU_1075  | GH8            | 6.94               | 6.84              | 6.3          | 6.01        | S           | S        | YES  |
|   | CHU_1240  | GH8, CE4, CBM9 | 5.88               | ND                | ND           | ND          | OM          | S        | YES  |
|   | CHU_2852  | GH8            | 5.6                | ND                | ND           | ND          | S           | S        | YES  |
|   | CHU_3440  | GH8            | 5.41               | ND                | ND           | ND          | S           | OM       | YES  |
|   | CHU_3441  | GH8            | ND                 | 5.64              | ND           | ND          | S           | OM       | YES  |
|   | CHU_3727  | GH8, CBM9      | 5.87               | ND                | ND           | ND          | OM          | S        | YES  |
| <b>GH9 (endoglucanase)</b>                  |           |                |                    |                   |              |             |             |          |      |
| <b>cel9E</b>                                | CHU_0778  | GH9            | ND                 | ND                | ND           | ND          | S           | ND       | NO   |
| <b>cel</b>                                  | CHU_0961  | GH9            | 5.41               | ND                | ND           | ND          | S           | S        | YES  |
| <b>cel9C</b>                                | CHU_1280  | GH9            | 5.96               | 5.34              | 5.44         | ND          | S           | S, P     | NO   |
| <b>cel9B</b>                                | CHU_1335  | GH9            | 5.13               | 5.09              | ND           | ND          | S           | S        | YES  |
| <b>cel9A</b>                                | CHU_1336  | GH9            | 5.72               | 5.99              | ND           | 5.27        | S           | S        | YES  |
| <b>cel9D</b>                                | CHU_1655  | GH9            | 5.65               | 5.78              | ND           | ND          | S           | IM, P, S | YES  |
| <b>cel9F</b>                                | CHU_2235  | GH9            | ND                 | ND                | ND           | ND          | S           | ND       | NO   |

<sup>a</sup>The predicted localization is based on prediction using pSORTb 3.0 (65). The proteome localizations were based on overrepresentation in a fraction and are abbreviated as follows: C, cytoplasmic; P, periplasmic; S, secreted extracellular soluble; IM, inner membrane; OM, outer membrane; L, lipoprotein. The presence of the T9SS was based on the presence of the TIGR04183 or TIGR04131 C-terminal domain. Red shading represents  $\log_{10}$  LFQ intensity levels over 6, purple levels between 5.5 and 5.99, blue levels between 5 and 5.49, and turquoise levels between 4.5 and 4.99. A lack of shading indicates that the levels of the proteins in question were below detectable limits.

periplasm and in the extracellular soluble fractions, suggesting that *C. hutchinsonii* might produce enzyme-loaded outer membrane vesicles (OMV) targeting recalcitrant biomass, similarly to other Gram-negative species such as *Fibrobacter succinogenes* (49–51). Further experiments are necessary to substantiate this hypothesis. Another GH9 enzyme, CHU\_1655, was abundant in both the early and late growth phases on filter paper and was detected in the periplasm and inner membrane as well as in the extracellular soluble fraction. Previous efforts to create a CHU\_1655 mutant to obtain a viable strain were unsuccessful (37, 43), which, together with the proteomic data, suggests that CHU\_1655 and/or its genetic neighborhood is essential for cellulose digestion or carbohydrate transport (37). The three remaining GH9 enzymes identified in the proteome (CHU\_0961, CHU\_1335, and CHU\_1336) were found in the extracellular soluble fraction during early stages of growth on filter paper. CHU\_1335 and CHU\_1336 were also present during the later growth stage.

GH8 enzymes have curiously been overlooked as potential cellulases in studies of *C. hutchinsonii* (20, 43), and yet all six encoded putative GH8 enzymes were detected in



**FIG 2** Protein architectures of *C. hutchinsonii* enzymes putatively involved in cellulose degradation. Red rectangles represent N-terminal signal peptides or C-terminal T9SS sorting domains (TIGR04183 or TIGR04131). Abbreviations: BIG, bacterial Ig-like, group 2 (IPR003343);  $\beta$ -lac, beta-lactamase-related domain (IPR001466); DUF, domain of unknown function; GBD, galactose-binding-like domain; IG, immunoglobulin-like fold; PKD, polycystic kidney disease domain.

the proteome. Among those enzymes, CHU\_1240 and CHU\_3727 contain family 9 carbohydrate binding modules (CBMs), and CHU\_1240 also comprises a carbohydrate esterase family 4 (CE4) domain, hinting at a putative role in xylan modification. The remaining GH8 enzymes, CHU\_1075, CHU\_2852, CHU\_3440, and CHU\_3441, comprised single GH8 catalytic domains, though they accounted for less than 50% of the total protein lengths (Fig. 2). All GH8 enzymes except for CHU\_3441 were detected during early growth on filter paper, though only CHU\_1075 and CHU\_3441 were detected at the late stages of growth (Table 1; see also Fig. S1). CHU\_1075 was abundant also in pectin-grown cells, indicating potential constitutive expression and a key role in

carbohydrate turnover. The GH8 domain of CHU\_1075 exhibits 32% primary sequence identity with *Hungateiclostridium thermocellum* (previously *Clostridium thermocellum* and *Ruminiclostridium thermocellum*) CelA (Cthe\_0269), the primary endoglucanase of *H. thermocellum* cellulosomes, which suggests endoglucanase activity of the CHU\_1075 GH8 domain (3, 52, 53). However, the GH8 domain itself accounts for only ~400 amino acids in a total protein length of 2,690 amino acids. The remainder of the protein contains a putative galactose-binding domain, several immunoglobulin-like folds, and a CTD (Fig. 2), explaining its abundance in the extracellular soluble fraction (Table 1; see also Fig. S1). The biological role(s) of this and other large GH8-containing proteins is currently unknown.

**$\beta$ -Glucosidases.** *C. hutchinsonii* encodes five putative BGs (one GH1 and four GH3s [20]). Three GH3 enzymes (CHU\_2268, CHU\_2273, and CHU\_3784) were detected in the proteome during early stages of growth on filter paper; CHU\_2268 and CHU\_2273 remained in the late samples and the latter was identified also in pectin-grown cultures (Table 1), which indicates a general role in carbohydrate turnover. While all of the GH3 enzymes were predicted to be periplasmic, the two most abundant members (CHU\_2268 and CHU\_2273) were abundant also in the inner membrane, and the two enzymes have previously been shown to play redundant key roles in cellobiose utilization (42). The data agree well with both predictions and with previous studies demonstrating localization and activity of BGs in the periplasm without association with the outer membrane (42, 43).

**Hemicellulases and pectinases.** Despite the apparent inability of *C. hutchinsonii* to metabolize pentoses, several putative hemicellulose-modifying enzymes were detected during the early stages of growth on filter paper (14, 20). These included putative xylan-modifying enzymes (GH10, GH11, CE1, CE3, CE4, CE6, and CE15) and putative xyloglucan-modifying enzymes (GH31 and GH74) (Table S1). However, most of the encoded hemicellulose-modifying enzymes were not present during growth on pectin or during the late stages of growth on filter paper. Thus, hemicellulases are likely produced primarily by cells in direct contact with solid substrates to facilitate access to cellulose fibrils in the plant cell wall, similarly to other cellulose specialist species unable to metabolize pentoses and hemicellulose-derived oligosaccharides (3). Surprisingly, none of the putative pectin/pectate and rhamnogalacturonan lyases (CHU\_1162 and CHU\_1157, respectively) were detected in the proteome during growth.

**Type 9 secretion system.** The T9SS is essential for crystalline cellulose degradation as it transports proteins containing conserved CTDs across the outer membrane (34). Components of the multiprotein T9SS were detected during growth on both filter paper and pectin (Table S4). Mutating the gene encoding SprP (CHU\_0170) has been shown to cause both gliding motility and cellulose utilization defects (34). CHU\_0170 was detected throughout growth and was most abundant in the inner membrane. The protein is similar to the *Porphyromonas gingivalis* PorP, which is believed to form part of the T9SS outer membrane (OM) channel (36, 54). Another putative SprP-like protein, CHU\_3434, was found in the outer membrane in both early and late stages of growth on filter paper but was not detected in pectin-grown cultures. The genomic neighborhood of CHU\_3434 contains three large proteins without functional annotations (CHU\_3435, CHU\_3437, and CHU\_3439), all detected in the proteome of cells grown on filter paper, in addition to two large GH8 domain-containing proteins (CHU\_3440 and CHU\_3441; see above). The genomic colocalization and proteomic abundance suggest an important role of the cluster consisting of CHU\_3434 to CHU\_3441 in the T9SS, gliding motility, and/or cellulose utilization.

**Analysis of the *S. myxococcoides* proteome.** The proteome for *S. myxococcoides* was obtained and fractionated from cells grown on either filter paper or pectin in a manner similar to that described for *C. hutchinsonii*, and protein localizations were verified using marker proteins (Table S3). As was seen with *C. hutchinsonii*, a much greater number of proteins were detected from cells grown on filter paper (1,090

proteins total; 81 CAZymes) than from pectin-grown cells (245 proteins; 13 CAZymes) (see Table S2; see also Fig. S2).

**Endoglucanases.** As mentioned above, *S. myxococcoides* encodes multiple putative EGs (10 putative GH5, 11 GH8, and 8 GH9 enzymes). Among the GH5 enzymes, four (UniProt identifiers MYP\_458, MYP\_2666, MYP\_3206, and MYP\_3455) were detected during early stages of growth on filter paper (Table 2; see also Fig. S2). In contrast to *C. hutchinsonii*, the majority of GH5 members detected in the *S. myxococcoides* proteome were abundant in the extracellular soluble fraction rather than in the outer membrane and/or periplasm and, furthermore, were detected in both the early and late growth stages on filter paper. Only MYP\_3455 was also abundant throughout growth on pectin. MYP\_3455 and MYP\_458 harbor additional CBM6 domains, which may bind crystalline or amorphous cellulose (Fig. S3) (55). No similar EGs with appended CBM6 domains were found in *C. hutchinsonii*.

Six of the putative GH9 enzymes were detected in the proteome, and, similarly to the GH5 enzymes, the majority were most abundant in the extracellular soluble fraction (Table 2; see also Fig. S2). Based on sequence similarity and characterized *C. hutchinsonii* GH9 enzymes, processive EG activities for certain *S. myxococcoides* GH9 enzymes can be inferred (48). Four of the GH9 enzymes were detected throughout growth on filter paper; two (MYP\_327 and MYP\_4968) are putatively processive EGs and one (MYP\_1185) a putatively nonprocessive EG. MYP\_327 was present in all growth samples, and thus was possibly constitutively expressed, and was abundant in both the extracellular soluble and periplasmic fractions despite lacking a CTD. As with *C. hutchinsonii*, this points to the possibility of OMVs being utilized also by *S. myxococcoides* (49–51). MYP\_327 shares 61% amino acid identity with CHU\_1280, and their similar abundance profiles in the respective proteomes indicate an equivalently important role of MYP\_327 in *S. myxococcoides*.

In contrast to the *C. hutchinsonii* proteome, the majority of the *S. myxococcoides* GH8 enzymes were not detected in the proteomic analyses (Table 2; see also Fig. S2). The periplasmic MYP\_2156 enzyme was the most abundant GH8 enzyme, detected under all growth conditions. MYP\_2156 is not highly similar to any *C. hutchinsonii* enzyme (36% and 34% amino acid identities with CHU\_3440 and CHU\_3441, respectively) and differently located compared to the most abundant GH8 enzyme of *C. hutchinsonii* (CHU\_1075; extracellular soluble). Three other GH8 enzymes (MYP\_51, MYP\_1794, and MYP\_2112) were associated mainly with growth on filter paper. Six of the GH8 enzymes (MYP\_51, MYP\_1794, MYP\_2156, MYP\_2756, MYP\_3325, and MYP\_3786) exhibited architectures similar to that of *C. hutchinsonii* CHU\_1075, with very large protein regions of unknown function in addition to the GH8 catalytic domain (Fig. S3).

**$\beta$ -Glucosidases.** Six putative GH3  $\beta$ -glucosidases were identified in early and late stages of *S. myxococcoides* growth on filter paper (Table 2; see also Fig. S2). One, MYP\_4245, was also abundant during growth on pectin, indicating possible constitutive expression. Interestingly, this enzyme differed from the *C. hutchinsonii*  $\beta$ -glucosidases in that it was found in the extracellular soluble fraction, in contrast to its predicted location in the periplasm. The other putative  $\beta$ -glucosidases were all present in the periplasm at abundance levels similar to those seen with the BGs of *C. hutchinsonii*. *S. myxococcoides* has previously not been shown to grow on cellobiose without extensive adaptation, which is curious considering the extracellular location of MYP\_4245 and the BG activity detected in the secreted soluble fraction (see below and Fig. 3; see also Table S2) (30).

**Hemicellulases, pectinases, and the type 9 secretion system.** In contrast to *C. hutchinsonii*, the vast majority of putative hemicellulose- and pectin-modifying enzymes of *S. myxococcoides* were not detected in the proteomes, and these are likely differently regulated in the two species. Putative essential T9SS components of *S. myxococcoides* were detected under all growth conditions, which indicates constitutive expression of the T9SS (Table S5). While these proteins have not been confirmed to play

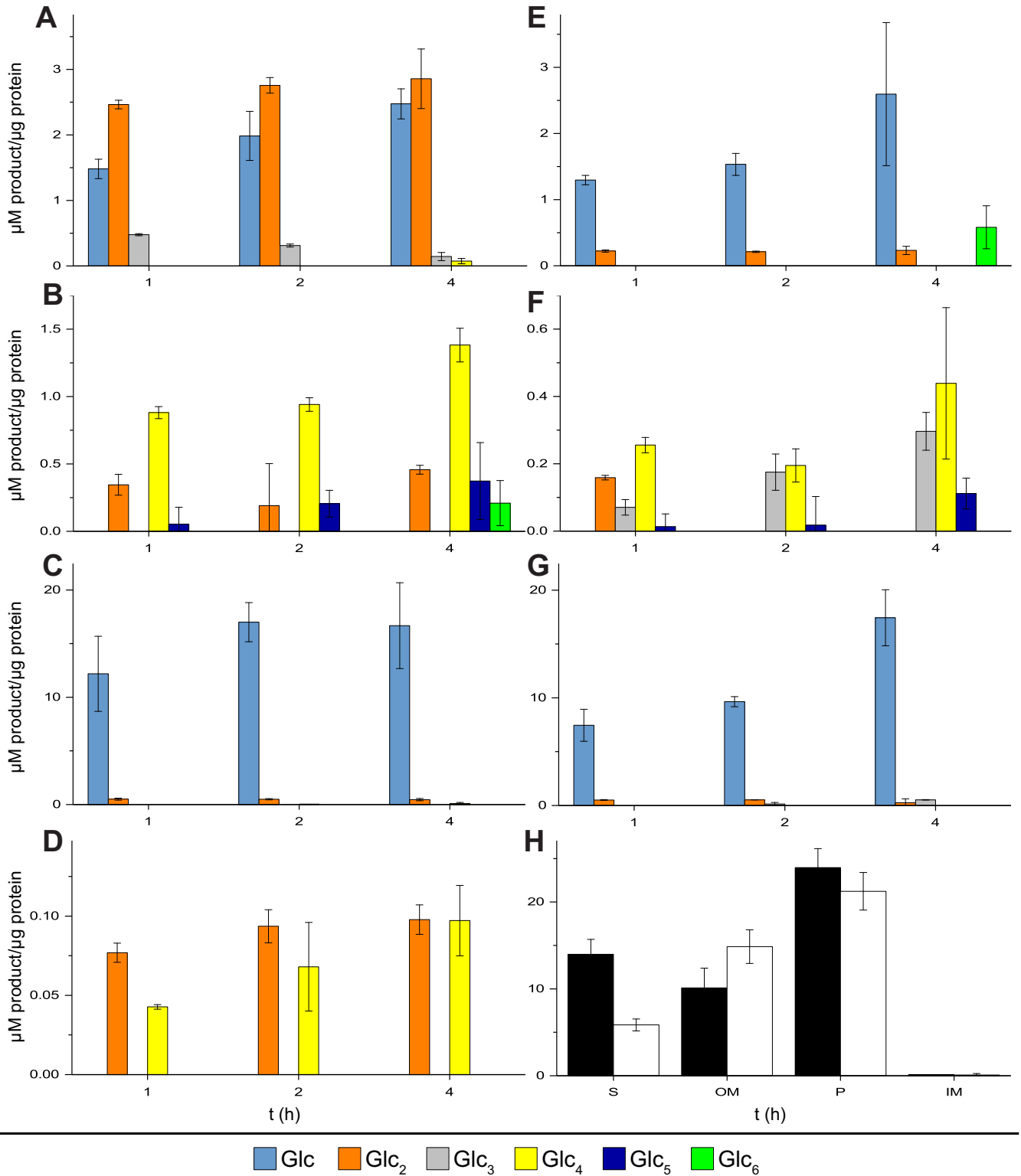
**TABLE 2** Predicted endoglucanases and  $\beta$ -glucosidases from *S. myxococcoides* displaying  $\log_{10}$  LFQ intensity values during early and late phases of growth with filter paper or pectin as the sole carbon source<sup>a</sup>

| Locus Tag                                   | CAZY domains   | Early Filter Paper | Late Filter Paper | Early Pectin | Late Pectin | Pred. Local | Local | T9SS |
|---|----------------|--------------------|-------------------|--------------|-------------|-------------|-------|------|
| <b>GH1 (<math>\beta</math>-glucosidase)</b> |                |                    |                   |              |             |             |       |      |
| MYP_2015                                    | GH1            | ND                 | ND                | ND           | ND          | C           | ND    | NO   |
| <b>GH3 (<math>\beta</math>-glucosidase)</b> |                |                    |                   |              |             |             |       |      |
| <b>bgIX</b> MYP_984                         | GH3            | 7.54               | 7.83              | ND           | ND          | C           | P, OM | NO   |
| <b>bgIX</b> MYP_1909                        | GH3            | 7.17               | 7.29              | ND           | ND          | C           | P     | NO   |
| <b>bgIX</b> MYP_1602                        | GH3            | 7.1                | 7.66              | ND           | ND          | C           | P     | NO   |
| <b>bgIX</b> MYP_2588                        | GH3            | 7.29               | 7.16              | ND           | ND          | P           | P     | NO   |
| <b>bgIX</b> MYP_3067                        | GH3            | 6.84               | 6.64              | ND           | ND          | C           | P     | NO   |
| <b>bgIX</b> MYP_4245                        | GH3            | 9                  | 9.33              | 8.09         | 7.67        | P           | S     | NO   |
| <b>bgIX</b> MYP_4961                        | GH3            | ND                 | ND                | ND           | ND          | C           | ND    | NO   |
| <b>GH5 (endoglucanase)</b>                  |                |                    |                   |              |             |             |       |      |
| cel MYP_458                                 | GH5, CBM6      | 6.88               | 6.63              | ND           | ND          | S           | S     | YES  |
| cel MYP_1753                                | GH5            | ND                 | 6.71              | ND           | ND          | S           | P     | YES  |
| cel MYP_2157                                | GH5            | ND                 | ND                | ND           | ND          | P           | ND    | NO   |
| cel MYP_2390                                | GH5            | ND                 | ND                | ND           | ND          | S           | ND    | YES  |
| cel MYP_2462                                | GH5, CBM6      | ND                 | ND                | ND           | ND          | S           | ND    | YES  |
| cel MYP_2666                                | GH5            | 7.81               | 7.77              | ND           | ND          | S           | S     | NO   |
| cel MYP_3206                                | GH5            | 7.79               | 7.01              | ND           | ND          | S           | S     | YES  |
| cel MYP_3455                                | GH5, CBM6      | 7.46               | 5.86              | 6.62         | 6.62        | S           | S     | YES  |
| cel MYP_3567                                | GH5            | ND                 | ND                | ND           | ND          | S           | ND    | NO   |
| <b>GH8 (endoglucanase)</b>                  |                |                    |                   |              |             |             |       |      |
| MYP_51                                      | GH8            | 6.98               | 7.28              | ND           | ND          | S           | P, OM | YES  |
| MYP_1583                                    | GH8, CBM9      | ND                 | ND                | ND           | 7.93        | S           | S     | YES  |
| MYP_1794                                    | GH8            | 7.66               | 7.99              | ND           | 7.12        | S           | S     | YES  |
| MYP_2112                                    | GH8, CBM9      | 6.51               | 7.14              | ND           | ND          | OM          | S     | YES  |
| MYP_2156                                    | GH8            | 8.22               | 8.56              | 7.96         | 8.07        | S           | P     | YES  |
| MYP_2756                                    | GH8            | ND                 | ND                | ND           | ND          | S           | ND    | NO   |
| MYP_3325                                    | GH8            | ND                 | ND                | ND           | ND          | S           | ND    | NO   |
| MYP_3614                                    | GH8, CE4, CBM9 | ND                 | ND                | ND           | ND          | OM          | ND    | YES  |
| MYP_3786                                    | GH8            | ND                 | ND                | ND           | ND          | S           | ND    | YES  |
| MYP_4354                                    | GH8            | ND                 | 7.25              | ND           | ND          | S           | OM    | YES  |
| MYP_4347                                    | GH8, CBM4      | ND                 | ND                | ND           | ND          | S           | ND    | YES  |
| <b>GH9 (endoglucanase)</b>                  |                |                    |                   |              |             |             |       |      |
| cel MYP_327                                 | GH9            | 8.31               | 8.35              | 7.73         | 8.03        | P           | S     | NO   |
| cel MYP_1185                                | GH9            | 7.78               | 7.92              | ND           | ND          | S           | S     | YES  |
| cel MYP_1436                                | GH9            | ND                 | ND                | ND           | ND          | P           | ND    | YES  |
| cel MYP_1796                                | GH9            | ND                 | ND                | ND           | ND          | S           | ND    | YES  |
| cel MYP_2099                                | GH9            | ND                 | ND                | ND           | ND          | S           | ND    | YES  |
| cel MYP_3429                                | GH9            | ND                 | 6.58              | ND           | ND          | S           | S     | YES  |
| cel MYP_4968                                | GH9            | 8.32               | 8.13              | ND           | 7.66        | P           | P, S  | NO   |

<sup>a</sup>The predicted localization is based on prediction using pSORTb 3.0 (65). The proteome localizations were based on overrepresentation in a fraction and are abbreviated as follows: C, cytoplasmic; P, periplasmic; S, secreted extracellular soluble; IM, inner membrane; OM, outer membrane; L, lipoprotein. The presence of the T9SS was based on the presence of the TIGR04183 or TIGR04131 C-terminal domain. Red shading represents  $\log_{10}$  LFQ intensity levels over 8.5, purple levels between 7.75 and 8.49, blue levels between 7 and 7.74, turquoise levels between 6.5 and 6.99, and green levels between 5.5 and 6.49. A lack of shading indicates that the levels of the proteins in question were below detectable limits.

a role in gliding motility or cellulose turnover, a level of importance of the system in *S. myxococcoides* similar to that in *C. hutchinsonii* can be postulated.

**Enzymatic cellulose hydrolysis.** To potentially correlate the proteomic profiles with enzymatic activity, the different protein fractions from the two species, except



**FIG 3** Enzyme activity on cellulose in different protein fractions. Colored bars represent released glucose (light blue), cellobiose (orange), cellotriose (gray), cellotetraose (yellow), cellopentaose (dark blue), and cellohexaose (green) after 1-, 2-, and 4-h reactions. (A to D) Products from cellulose hydrolysis by *C. hutchinsonii* proteins: extracellular soluble (A), outer membrane (B), periplasm (C), and inner membrane (D). (E to G) Products from cellulose hydrolysis by *S. myxococcoides* proteins: extracellular soluble (E), outer membrane (F), and periplasmic (G) (no cellulolytic activity was detected in the *S. myxococcoides* inner membrane fraction). Panel H shows activity on pNP-β-glucoside in extracellular soluble (S), outer membrane (OM), periplasmic (P), and inner membrane (IM) fractions from *C. hutchinsonii* (black) and *S. myxococcoides* (white) after 1 h of incubation. Error bars represent standard deviations of results from triplicate experiments.

cytosolic proteins, were assayed for cellulase and  $\beta$ -glucosidase activities using crystalline cellulose (Avicel) and 4-nitrophenyl- $\beta$ -glucoside (*p*NP- $\beta$ -Glc) as substrates (Fig. 3). For cellulose hydrolysis, the activity levels and product profiles differed between both species and cellular fractions. While the extracellular fraction of *C. hutchinsonii* produced almost equal amounts of glucose and cellobiose, glucose was the main product from *S. myxococcoides* (Fig. 3A and E). The latter result was unexpected, as *S. myxococcoides* does not readily grow on glucose, but the activity may be attributable to the main extracellular  $\beta$ -glucosidase, MYP\_4245, as mentioned previously, and correlates with the activity on *p*NP- $\beta$ -Glc in this protein fraction (Fig. 3H). No glucose was detected in reactions using outer membrane proteins (Fig. 3B and F); instead, cellotetraose was the predominant product for *C. hutchinsonii* and cellotetraose and cellotriose were the predominant products for *S. myxococcoides*, though longer oligosaccharides were also detected. The reactions performed with periplasmic proteins yielded only glucose (Fig. 3C and G), correlating well with the presence of periplasmic  $\beta$ -glucosidases in both species and with the corresponding highest activities on *p*NP- $\beta$ -Glc (Fig. 3H). The inner membrane fraction of *C. hutchinsonii* unexpectedly exhibited cellulase activity (Fig. 3D), which may be attributable to the to-date-uncharacterized inner membrane-located GH5 enzymes CHU\_1107 and CHU\_1842. *S. myxococcoides* lacks similarly located corresponding enzymes and consequently exhibited no cellulolytic activity in this cellular compartment.

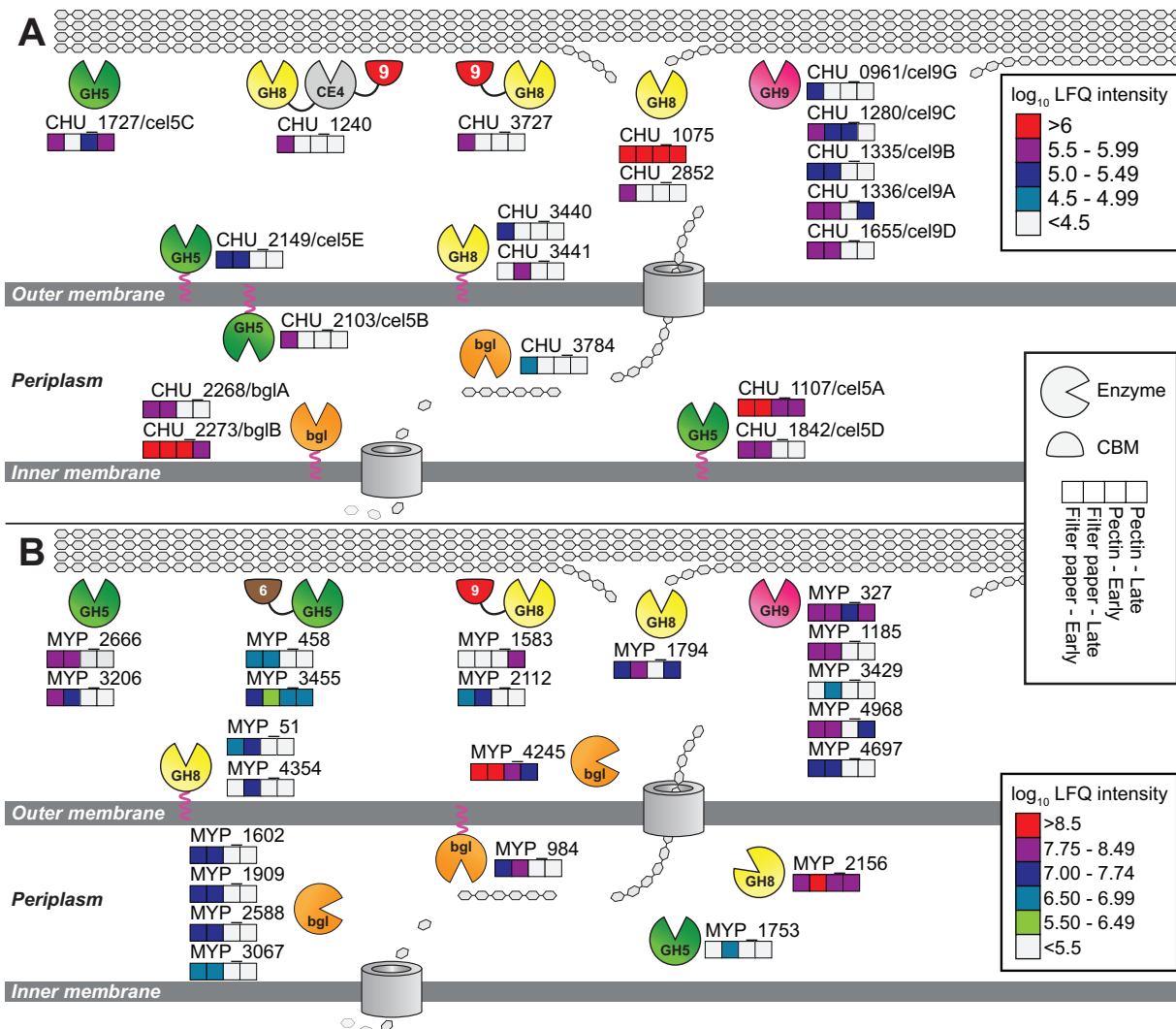
## DISCUSSION

The cellulose specialist species *C. hutchinsonii* and *S. myxococcoides* possess multiple putative cellulose-degrading enzymes, but as neither bacterium encodes known or putative exoglucanases or LPMOs, their mechanisms allowing rapid conversion of cellulose remain enigmatic. Some GH5 and GH9 *C. hutchinsonii* endoglucanases have been characterized and physiological functions postulated and evaluated through mutational studies (21, 22, 37, 48, 56), but these studies have not been sufficient to explain how crystalline cellulose is converted by this bacterium. Corresponding studies of *S. myxococcoides* are lacking, despite the likelihood that it uses a highly similar cellulolytic system.

In our proteomic analyses, all GH5 enzymes and most GH9 enzymes from *C. hutchinsonii* were detected during growth on solid filter paper but only 50% to 60% of the GH5 and GH9 enzymes of *S. myxococcoides*. Interestingly, *C. hutchinsonii* GH5 enzyme CHU\_2103, which has previously been identified as highly important in cellulose conversion (37, 43, 48), was found only in the early phase of growth on filter paper when the cells were attached to the solid substrate. In contrast, the key GH9 enzyme CHU\_1280 was found throughout growth on cellulose. Our data corroborate previous results indicating the importance of these enzymes; further, demonstrating CHU\_2103 being located in the outer membrane may explain the strong correlation between cellulase activity and attachment of *C. hutchinsonii* cells to solid substrates (39).

In the proteomes of both organisms, a large number of GH8 enzymes were abundant during growth on solid filter paper, with CHU\_1075 and MYP\_2156 being particularly abundant in *C. hutchinsonii* and *S. myxococcoides*, respectively. Interestingly, the majority of GH8 enzymes in both species displayed protein architectures where the GH8 domains accounted for relatively small portions of the overall large polypeptides (Fig. 2; see also Fig. S3 in the supplemental material). The main portions of these proteins have no known functions related to cellulose degradation, and their biological role(s) remains speculative. Further, the differences in the cellular locations of the GH8 enzymes in the two bacteria are noteworthy and may point to different roles in cellulose turnover. While all GH8 enzymes of *C. hutchinsonii* were annotated as extracellular soluble or facing the external environment, the main GH8 enzyme of *S. myxococcoides*, MYP\_2156, was found in the periplasm.

The exact mechanisms of cellulose hydrolysis in *C. hutchinsonii* and *S. myxococcoides* remain unresolved, but we can infer from our proteomic and enzyme activity data that initial hydrolysis of the cellulose backbone is undertaken by a plethora of putative EGs



**FIG 4** Putative cellulose degradation mechanism in (A) *C. hutchesonii* and (B) *S. myxococcoides* with protein localizations based on abundance in the proteome. Log<sub>10</sub> LFQ intensity is used to indicate abundance, where numbers below 4.5 and 5.5, respectively, are considered below detectable limits.

that are either secreted extracellularly or attached to the outer membrane (Fig. 4). The number of putative EGs located extracellularly in both species creates highly redundant systems which may safeguard against loss of function from single mutation events. Supporting this hypothesis, singular EG disruption mutants in *C. hutchesonii* have typically yielded strains without pronounced phenotypes (22, 37, 43, 57). It is highly interesting that additional cellulolytic enzymes are found intracellularly in the periplasm in both species and, for *C. hutchesonii*, also in the inner membrane. The role(s) of these intracellular cellulases is puzzling, as transport of cellulose filaments into the periplasm is unlikely, but the bacteria may be able to capture and import longer individual glucan chains for degradation within the periplasm, as has previously been hypothesized (43). The identities of the proteins involved in capture and transport of cellooligosaccharides across the cellular membranes are not known. Possibly, some of the proteins found enriched in the outer membrane fractions, which today lack annotations of functional domains, are important for these processes (see Table S6 in the supplemental material). The majority of proteins of unknown function enriched in the extracellular soluble and outer membrane fractions of both organisms did not contain any previously described features, such as Pfam domains or DUFs (domains of

unknown function). Thus, further research is necessary to shed light on their possible biological roles in carbohydrate turnover.

Both *C. hutchinsonii* and *S. myxococcoides* have been shown to require direct contact with the substrate for efficient cellulose hydrolysis to occur (14). However, we observed multiple proteins, including putative EGs, being secreted extracellularly in the proteomes in large amounts. During growth, both species produced significant amounts of an extracellular viscous “slime” which has not been structurally characterized to date. Possibly, the secreted proteins may be confined in the extracellular slime matrix in close proximity of the cells, minimizing the amount of solubilized sugars being lost to competing species. Compared to freely secreted enzymes or complexed cellulolytic systems, the semifluid nature of the slime might provide an advantage by allowing penetration of cellulolytic proteins into the crystalline substrate while simultaneously trapping released oligosaccharides. Further investigation is required to evaluate this hypothesis, however.

In summary, by mapping the growth and proteomes of both *C. hutchinsonii* and *S. myxococcoides* on various carbohydrates, we provide insight into the carbon source utilization capabilities of both of these cellulose-specialist species and into which proteins are utilized to metabolize crystalline cellulose and soluble pectin—polysaccharides exhibiting different monosaccharide composition, glycosidic linkage, charge, and solubility characteristics. In addition, we pinpoint the localization of different proteins within the cells, which is essential to explain the biology of these two poorly understood cellulolytic species. The plethora of putative cellulose degradation proteins, including many previously unstudied GH8 enzymes, that were abundant throughout the growth process displays the robustness and apparent redundancy of the cellulolytic systems in both *C. hutchinsonii* and *S. myxococcoides*. The combined results and the list of proteins of unknown function with putative functions in cellulose turnover (Table S6) provide a number of highly valuable targets for further analyses to identify novel key enzymes for improved cellulose degradation.

## MATERIALS AND METHODS

**Bacterial strains and growth conditions.** *C. hutchinsonii* G2 (a genome-sequenced glucose-metabolizing strain; referred to as *C. hutchinsonii* in the text) (20) and *S. myxococcoides* DSM11118 were grown in McBride minimal medium (MMM) (5.8 mM NaNO<sub>3</sub>, 2 mM MgSO<sub>4</sub>, 6.7 mM KCl, 36 μM FeSO<sub>4</sub>, 0.6 mM NaHCO<sub>3</sub>, 1 mM KH<sub>2</sub>PO<sub>4</sub>; pH 7.3) (58). *C. hutchinsonii* and *S. myxococcoides* were grown at 30°C and 25°C, respectively. The MMM was supplemented with 2 g/liter of 0.25-cm<sup>2</sup> squares of sterile Whatman filter paper or with 2 g/liter of soluble carbohydrates (pectin, xyloglucan, xylan, xylose, mannose, glucose, cellobiose, galactomannan, or glucomannan) as a carbon source. MMM agar was prepared by addition of 6 g/liter agar. Pieces of sterile Whatman filter paper were added on the solidifying agar as the sole carbon source, and the organisms were inoculated in the center of the filter paper strips.

**Protein extraction.** The organisms were grown on MMM-filter paper agar plates for 5 days, and cells were detached from the plates by gentle pipetting of MMM and used to inoculate flasks containing liquid MMM. The cultures were grown for approximately 6 days (early time point) on filter paper until the substrate was coated with yellow cells or for 14 days (late time point) when the medium became yellow due to filter paper disintegration. The cultures were centrifuged at 5,000 × *g* for 15 min, and the resulting supernatant liquids were collected and soluble secreted proteins concentrated using 4 volumes of acetone and –80°C incubation overnight. The cell pellets were resuspended in 15 ml of lysis buffer (1% SDS, 10 mM dithiothreitol [DTT], 100 mM ammonium bicarbonate, pH 8.0) and sonicated for 2 min (3-s pulses, 2-s pauses), followed by centrifugation at 16,000 × *g* for 15 min. The supernatant liquid was precipitated by addition of 4 volumes of acetone and incubation at –80°C overnight followed by centrifugation at 16,000 × *g* for 30 min at 4°C. The pellets were washed with acetone and centrifuged as described above. The acetone was removed, and the protein pellets were resuspended in a mixture containing 200 mM Tris-HCl, 40 mM β-mercaptoethanol, 8% SDS, 0.4% bromophenol blue, and 40% glycerol (pH 6.8), separated by SDS-PAGE using TGX-unstained Mini-Protean gel (Bio-Laboratories, Hercules, CA, USA), and stained using Coomassie brilliant blue R250.

**Cellular fractionation.** The cellular fractions were obtained as previously described (59) from cells grown as described above. At the late time point, the cells were collected by centrifugation at 5,000 × *g* for 15 min. The supernatant liquid was collected as the extracellular soluble fraction. The pelleted cells were washed with 25 ml 30 mM Tris-HCl (pH 8.0) and centrifuged at 5,000 × *g* for 15 min. The cell pellet was resuspended in a reaction mixture containing 25 ml 30 mM Tris-HCl (pH 8.0), 20% sucrose, and 1 mM EDTA and was incubated 10 min at room temperature, followed by centrifugation at 5,000 × *g* for 15 min. The pellet was resuspended in 25 ml of ice-cold 5 mM MgSO<sub>4</sub>, incubated on ice for 10 min, and centrifuged at 5,000 × *g* for 20 min. The supernatant liquid was collected as the periplasmic fraction. The resulting pellet was resuspended in 20 ml of 50 mM Tris-HCl (pH 8.0) and lysed by sonication followed

by centrifugation at  $16,000 \times g$  for 30 min. The supernatant liquid was centrifuged at  $100,000 \times g$  for 1 h at 4°C. Resulting supernatant liquid was collected as the cytoplasmic fraction. The pellet containing the total membranes was dissolved in a reaction mixture containing 50 mM Tris-HCl (pH 7.5), 10 mM  $MgCl_2$ , and 2% Triton X-100 and incubated at room temperature with shaking for 1 h, followed by centrifugation at  $100,000 \times g$  for 1 h at 4°C. The supernatant liquid was collected as the inner membrane fraction and the pellet as the outer membrane fraction. Outer membrane protein A (ompA; CHU\_1710 and MYP\_505), periplasmic phosphate binding protein (pstS; CHU\_3818 and MYP\_603), and NADH dehydrogenase subunit L (nuoL; CHU\_1371 and MYP\_729) were used as markers to assess the fractionation of the outer membrane, periplasm, and inner membrane, respectively (see Table S3 in the supplemental material).

**In-gel trypsinization and mass-spectrometric and bioinformatic analyses.** Peptides were prepared as described by Arntzen et al. (60) and subjected to in-gel trypsinization as described by Shevchenko et al. (61). The dried peptides were analyzed by the use of NanoLC-Orbitrap tandem mass spectrometry (MS/MS) as described previously (58). MS raw files were analyzed using MaxQuant v1.5.8.3 (62) for label-free quantification (LFQ) of proteins and the MaxQuant label-free algorithm (47) with both unique and razor peptides for protein quantification. Protein identification was done using the Andromeda database search engine (63) for searches against the UniProt proteome set for *C. hutchinsonii* (UP000001822; 3,771 sequences) and *S. myxococcoides* (UP000030185; 5,042 sequences). Identifications were filtered to achieve a protein false-discovery rate (FDR) of 1%. Proteins identified in two of the three biological replicates were considered valid. Overrepresentation of the proteins in the fractions was calculated as a percentage of the protein LFQ intensity value corresponding to the sum of all LFQ intensities in the fraction.

**Enzyme assays.** The cellular fractions were separated and collected as previously described. The outer membrane fraction was resuspended in 50 mM Tris-HCl (pH 8.0). Proteins were concentrated by centrifugation performed with Pierce protein concentrators (Thermo Scientific, Bremen, Germany) (3 K molecular weight cutoff [MWCO]). Avicel degradation activity was determined in 200- $\mu$ l reaction mixtures containing 2 g/liter Avicel PH-101 (Sigma-Aldrich, Stockholm, Sweden), 50 mM sodium citrate (pH 7.0), 5 mM  $CaCl_2$ , and 10  $\mu$ g protein (determined by the Bradford assay). The assays were incubated with shaking in round-bottomed tubes at 30°C and were stopped by the addition of 50  $\mu$ l 5 M NaOH. Products were determined using a Dionex ICS-5000 system equipped with a CarboPac PA-200 analytical column (Dionex, Sunnyvale, CA, USA) (3 by 250 mm) and a flow rate of 0.5 ml/min. Eluent A consisted of water, eluent B of 300 mM NaOH, and eluent C of 1 M sodium acetate. The method used was as follows: 0 to 5 min, 15% eluent B; 5 to 15 min, a linear gradient of 15% to 33% eluent B; 15 to 25 min, 33% eluent B and a linear gradient of 0% to 10% C; 25 to 30 min, 33% eluent B and 66% eluent C. Products were detected by pulsed amperometric detection (PAD) using Chromeleon software v.7 (Thermo Scientific, Bremen, Germany).  $\beta$ -Glucosidase activity was assayed using pNP- $\beta$ -Glc in 96-well plates and 200- $\mu$ l reaction mixtures containing 25 mM sodium citrate (pH 7.0), 5 mM  $CaCl_2$ , 10 mM pNP- $\beta$ -Glc, and 1 to 10  $\mu$ g protein, and the plates were incubated at 30°C for 1 h. The reactions were stopped by addition of 100  $\mu$ l 0.4 M glycine (pH 10.8), and the products were quantified at 405 nm using a FluoStar Omega plate reader (BMG Labtech, Ortenberg, Germany) and a p-nitrophenol standard curve.

**Data availability.** The proteomics data have been deposited in the ProteomeXchange consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository (64) with the data set identifier PXD010836.

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mSystems.00240-18>.

**FIG S1**, EPS file, 1.3 MB.

**FIG S2**, EPS file, 1.2 MB.

**FIG S3**, EPS file, 1.2 MB.

**TABLE S1**, DOCX file, 0.04 MB.

**TABLE S2**, DOCX file, 0.04 MB.

**TABLE S3**, DOCX file, 0.01 MB.

**TABLE S4**, DOCX file, 0.02 MB.

**TABLE S5**, DOCX file, 0.01 MB.

**TABLE S6**, DOCX file, 0.02 MB.

## ACKNOWLEDGMENTS

We thank Mark McBride, the University of Wisconsin—Milwaukee, for constructive discussions.

M.T. and J.L. were supported by a grant from the Swedish Research Council and the Swedish Energy Agency through the Energy-Oriented Basic Research program (Dnr 2015-009561) and by the Knut and Alice Wallenberg Foundation through the Wallenberg Wood Science Center. P.B.P. and M.Ø.A. were supported by The Research Council of Norway (FRIPRO program 250479) and the European Research Council (336355, MicroDE).

M.T. planned and performed the experiments and data analysis and wrote the manuscript. J.L. designed and coordinated the overall study, planned the experiments, and wrote the manuscript. P.B.P. and M.Ø.A. advised with the bioinformatic analysis and mass spectroscopy. B.H. contributed to the annotation of the carbohydrate-active proteins and analyzed the *S. myxococcoides* genome. All of us read and approved the final manuscript.

## REFERENCES

- Fargione J, Hill J, Tilman D, Polasky S, Hawthorne P. 2008. Land clearing and the biofuel carbon debt. *Science* 319:1235–1238. <https://doi.org/10.1126/science.1152747>.
- Eisentraut A. 2010. Sustainable production of second-generation biofuels: potential and perspectives in major economies and developing countries. Renewable Energy Division, International Energy Agency (IEA), Paris, France.
- Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS. 2002. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol Mol Biol Rev* 66:506–577. <https://doi.org/10.1128/MMBR.66.3.506-577.2002>.
- Haitjema CH, Gilmore SP, Henske JK, Solomon KV, de Groot R, Kuo A, Mondo SJ, Salamov AA, LaButti K, Zhao Z, Chiniquy J, Barry K, Brewer HM, Purvine SO, Wright AT, Hainaut M, Boxma B, van Alen T, Hackstein JHP, Henrissat B, Baker SE, Grigoriev IV, O'Malley MA. 2017. A parts list for fungal cellulosomes revealed by comparative genomics. *Nat Microbiol* 2:17087. <https://doi.org/10.1038/nmicrobiol.2017.87>.
- Merino ST, Cherry J. 2007. Progress and challenges in enzyme development for biomass utilization. *Adv Biochem Eng Biotechnol* 108:95–120. [https://doi.org/10.1007/10\\_2007\\_066](https://doi.org/10.1007/10_2007_066).
- Watson BJ, Zhang H, Longmire AG, Moon YH, Hutcheson SW. 2009. Processive endoglucanases mediate degradation of cellulose by *Saccharophagus degradans*. *J Bacteriol* 191:5697–5705. <https://doi.org/10.1128/JB.00481-09>.
- Li Y, Irwin DC, Wilson DB. 2007. Processivity, substrate binding, and mechanism of cellulose hydrolysis by *Thermobifida fusca* Cel9A. *Appl Environ Microbiol* 73:3165–3172. <https://doi.org/10.1128/AEM.02960-06>.
- 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>.
- Vaae-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sorlie M, Eijsink VG. 2010. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* 330:219–222. <https://doi.org/10.1126/science.1192231>.
- Langston JA, Shaghisi T, Abbate E, Xu F, Vlasenko E, Sweeney MD. 2011. Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Appl Environ Microbiol* 77:7007–7015. <https://doi.org/10.1128/AEM.05815-11>.
- Harris PV, Welner D, McFarland KC, Re E, Navarro Poulsen JC, Brown K, Salbo R, Ding H, Vlasenko E, Merino S, Xu F, Cherry J, Larsen S, Lo Leggio L. 2010. Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry* 49:3305–3316. <https://doi.org/10.1021/bi100009p>.
- Forsberg Z, Mackenzie AK, Sorlie M, Rohr AK, Helland R, Arvai AS, Vaae-Kolstad G, Eijsink VG. 2014. Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. *Proc Natl Acad Sci U S A* 111:8446–8451. <https://doi.org/10.1073/pnas.1402771111>.
- Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. 2013. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol Biofuels* 6:41. <https://doi.org/10.1186/1754-6834-6-41>.
- Stanier RY. 1942. The *Cytophaga* group: a contribution to the biology of *Myxobacteria*. *Bacteriol Rev* 6:143–196.
- Winogradsky S. 1929. Études sur la microbiologie du sol sur la dégradation de la cellulose dans le sol. *Ann Inst Pasteur* 43:549–633.
- McBride MJ, Liu W, Lu X, Zhu Y, Zhang W. 2014. The family Cytophagaceae, p 577–593. In *Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (ed), The prokaryotes*. Springer, Berlin, Germany.
- Hutchinson HB, Clayton J. 1919. On the decomposition of cellulose by an aerobic organism (*Spirochaeta cytophaga*, n sp.). *J Agric Sci* 9:143–173. <https://doi.org/10.1017/S0021859600004755>.
- Zhu Y, Li H, Zhou H, Chen G, Liu W. 2010. Cellulose and cellodextrin utilization by the cellulolytic bacterium *Cytophaga hutchinsonii*. *Bioresour Technol* 101:6432–6437. <https://doi.org/10.1016/j.biortech.2010.03.041>.
- Liu L, Gao P, Chen G, Wang L. 2014. Draft genome sequence of cellulose-digesting bacterium *Sporocytophaga myxococcoides* PG-01. *Genome Announc* 2:e01154-14. <https://doi.org/10.1128/genomeA.01154-14>.
- Xie G, Bruce DC, Challacombe JF, Chertkov O, Detter JC, Gilna P, Han CS, Lucas S, Misra M, Myers GL, Richardson P, Tapia R, Thayer N, Thompson LS, Brettin TS, Henrissat B, Wilson DB, McBride MJ. 2007. Genome sequence of the cellulolytic gliding bacterium *Cytophaga hutchinsonii*. *Appl Environ Microbiol* 73:3536–3546. <https://doi.org/10.1128/AEM.00225-07>.
- Zhu Y, Zhou H, Bi Y, Zhang W, Chen G, Liu W. 2013. Characterization of a family 5 glycoside hydrolase isolated from the outer membrane of cellulolytic *Cytophaga hutchinsonii*. *Appl Microbiol Biotechnol* 97:3925–3937. <https://doi.org/10.1007/s00253-012-4259-x>.
- Zhang C, Zhang W, Lu X. 2015. Expression and characteristics of a Ca<sup>2+</sup>-dependent endoglucanase from *Cytophaga hutchinsonii*. *Appl Microbiol Biotechnol* 99:9617–9623. <https://doi.org/10.1007/s00253-015-6746-3>.
- Bayer EA, Belaich JP, Shoham Y, Lamed R. 2004. The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annu Rev Microbiol* 58:521–554. <https://doi.org/10.1146/annurev.micro.57.030502.091022>.
- Wilson DB. 2008. Three microbial strategies for plant cell wall degradation. *Ann N Y Acad Sci* 1125:289–297. <https://doi.org/10.1196/annals.1419.026>.
- Wilson DB. 2009. Evidence for a novel mechanism of microbial cellulose degradation. *Cellulose* 16:723–727. <https://doi.org/10.1007/s10570-009-9326-9>.
- 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>.
- Larsbrink J, Rogers TE, Hemsworth GR, McKee LS, Tazuin AS, Spadiut O, Klintner 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>.
- Larsbrink J, Zhu Y, Kharade SS, Kwiatkowski KJ, Eijsink VG, Koropatkin NM, McBride MJ, Pope PB. 2016. A polysaccharide utilization locus from *Flavobacterium johnsoniae* enables conversion of recalcitrant chitin. *Biotechnol Biofuels* 9:260. <https://doi.org/10.1186/s13068-016-0674-z>.
- Zhu Y, Kwiatkowski KJ, Yang T, Kharade SS, Bahr CM, Koropatkin NM, Liu W, McBride MJ. 2015. Outer membrane proteins related to SusC and SusD are not required for *Cytophaga hutchinsonii* cellulose utilization. *Appl Microbiol Biotechnol* 99:6339–6350. <https://doi.org/10.1007/s00253-015-6555-8>.
- Sijpesteijn AK, Fahraeus G. 1949. Adaptation of *Sporocytophaga myxococcoides* to sugars. *J Gen Microbiol* 3:224–235. <https://doi.org/10.1099/00221287-3-2-224>.
- Holt SC, Leadbetter ER. 1967. Fine structure of *Sporocytophaga myxococcoides*. *Arch Mikrobiol* 57:199–213. <https://doi.org/10.1007/BF00405947>.
- Sorensen LH. 1956. Decomposition of xylan by *Sporocytophaga myxococcoides*. *Nature* 177:845. <https://doi.org/10.1038/177845a0>.
- McBride MJ, Zhu Y. 2013. Gliding motility and Por secretion system genes are widespread among members of the phylum bacteroidetes. *J Bacteriol* 195:270–278. <https://doi.org/10.1128/JB.01962-12>.
- Zhu Y, McBride MJ. 2014. Deletion of the *Cytophaga hutchinsonii* type IX secretion system gene sprP results in defects in gliding motility and cellulose utilization. *Appl Microbiol Biotechnol* 98:763–775. <https://doi.org/10.1007/s00253-013-5355-2>.

35. Veith PD, Nor Muhammad NA, Dashper SG, Likic VA, Gorasia DG, Chen D, Byrne SJ, Catmull DV, Reynolds EC. 2013. Protein substrates of a novel secretion system are numerous in the Bacteroidetes phylum and have in common a cleavable C-terminal secretion signal, extensive post-translational modification, and cell-surface attachment. *J Proteome Res* 12:4449–4461. <https://doi.org/10.1021/pr400487b>.
36. Lasica AM, Ksiazek M, Madej M, Potempa J. 2017. The type IX secretion system (T9SS): highlights and recent insights into its structure and function. *Front Cell Infect Microbiol* 7:215. <https://doi.org/10.3389/fcimb.2017.00215>.
37. Zhu Y, Han L, Hefferon KL, Silvaggi NR, Wilson DB, McBride MJ. 2016. Periplasmic *Cytophaga hutchinsonii* endoglucanases are required for use of crystalline cellulose as the sole source of carbon and energy. *Appl Environ Microbiol* 82:4835–4845. <https://doi.org/10.1128/AEM.01298-16>.
38. Wang S, Zhao D, Bai X, Zhang W, Lu X. 1 January 2017. Identification and characterization of a large protein essential for degradation of the crystalline region of cellulose by *Cytophaga hutchinsonii*. *Appl Environ Microbiol* <https://doi.org/10.1128/AEM.02270-16>.
39. Yang T, Bu X, Han Q, Wang X, Zhou H, Chen G, Zhang W, Liu W. 2016. A small periplasmic protein essential for *Cytophaga hutchinsonii* cellulose digestion. *Appl Microbiol Biotechnol* 100:1935–1944. <https://doi.org/10.1007/s00253-015-7204-y>.
40. Zhang C, Wang X, Zhang W, Zhao Y, Lu X. 2017. Expression and characterization of a glucose-tolerant beta-1,4-glucosidase with wide substrate specificity from *Cytophaga hutchinsonii*. *Appl Microbiol Biotechnol* 101:1919–1926. <https://doi.org/10.1007/s00253-016-7927-4>.
41. Ji X, Wang Y, Zhang C, Bai X, Zhang W, Lu X. 2014. Novel outer membrane protein involved in cellulose and cellooligosaccharide degradation by *Cytophaga hutchinsonii*. *Appl Environ Microbiol* 80:4511–4518. <https://doi.org/10.1128/AEM.00687-14>.
42. Bai X, Wang X, Wang S, Ji X, Guan Z, Zhang W, Lu X. 2017. Functional studies of beta-glucosidases of *Cytophaga hutchinsonii* and their effects on cellulose degradation. *Front Microbiol* 8:140. <https://doi.org/10.3389/fmicb.2017.00140>.
43. Zhu Y, McBride MJ. 2017. The unusual cellulose utilization system of the aerobic soil bacterium *Cytophaga hutchinsonii*. *Appl Microbiol Biotechnol* 101:7113–7127. <https://doi.org/10.1007/s00253-017-8467-2>.
44. Smith SG, Mahon V, Lambert MA, Fagan RP. 2007. A molecular Swiss army knife: OmpA structure, function and expression. *FEMS Microbiol Lett* 273:1–11. <https://doi.org/10.1111/j.1574-6968.2007.00778.x>.
45. Agüena M, Ferreira GM, Spira B. 2009. Stability of the *pstS* transcript of *Escherichia coli*. *Arch Microbiol* 191:105–112. <https://doi.org/10.1007/s00203-008-0433-z>.
46. Leif H, Sled VD, Ohnishi T, Weiss H, Friedrich T. 1995. Isolation and characterization of the proton-translocating NADH: ubiquinone oxidoreductase from *Escherichia coli*. *Eur J Biochem* 230:538–548. <https://doi.org/10.1111/j.1432-1033.1995.tb20594.x>.
47. Cox J, Hein MY, Lubner CA, Paron I, Nagaraj N, Mann M. 2014. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* 13:2513–2526. <https://doi.org/10.1074/mcp.M113.031591>.
48. Zhang C, Wang Y, Li Z, Zhou X, Zhang W, Zhao Y, Lu X. 2014. Characterization of a multi-function processive endoglucanase CHU\_2103 from *Cytophaga hutchinsonii*. *Appl Microbiol Biotechnol* 98:6679–6687. <https://doi.org/10.1007/s00253-014-5640-8>.
49. Forsberg CW, Beveridge TJ, Hellstrom A. 1981. Cellulase and xylanase release from *Bacteroides succinogenes* and its importance in the rumen environment. *Appl Environ Microbiol* 42:886–896.
50. Burnet MC, Dohnalkova AC, Neumann AP, Lipton MS, Smith RD, Suen G, Callister SJ. 2015. Evaluating models of cellulose degradation by *Fibrobacter succinogenes* S85. *PLoS One* 10:e0143809. <https://doi.org/10.1371/journal.pone.0143809>.
51. Arntzen MO, Varnai A, Mackie RI, Eijsink VGH, Pope PB. 2017. Outer membrane vesicles from *Fibrobacter succinogenes* S85 contain an array of carbohydrate-active enzymes with versatile polysaccharide-degrading capacity. *Environ Microbiol* 19:2701–2714. <https://doi.org/10.1111/1462-2920.13770>.
52. Schwarz WH. 2001. The cellulosome and cellulose degradation by anaerobic bacteria. *Appl Microbiol Biotechnol* 56:634–649. <https://doi.org/10.1007/s002530100710>.
53. Alzari PM, Souchon H, Dominguez R. 1996. The crystal structure of endoglucanase CelA, a family 8 glycosyl hydrolase from *Clostridium thermocellum*. *Structure* 4:265–275. [https://doi.org/10.1016/S0969-2126\(96\)00031-7](https://doi.org/10.1016/S0969-2126(96)00031-7).
54. Vincent MS, Canestrari MJ, Leone P, Stathopoulos J, Ize B, Zoued A, Cambillau C, Kellenberger C, Roussel A, Cascales E. 2017. Characterization of the *Porphyromonas gingivalis* type IX secretion trans-envelope PorKLMNP core complex. *J Biol Chem* 292:3252–3261. <https://doi.org/10.1074/jbc.M116.765081>.
55. Michel G, Barbeyron T, Kloareg B, Czjzek M. 2009. The family 6 carbohydrate-binding modules have coevolved with their appended catalytic modules toward similar substrate specificity. *Glycobiology* 19:615–623. <https://doi.org/10.1093/glycob/cwp028>.
56. Loume C, Abazingo M, Johnson E, Latinwo L, Ikediobi C, Clark AM. 2007. Molecular cloning and biochemical characterization of a family-9 endoglucanase with an unusual structure from the gliding bacteria *Cytophaga hutchinsonii*. *Appl Biochem Biotechnol* 141:127–138. <https://doi.org/10.1007/s12010-007-9215-3>.
57. Wang X, Wang Z, Bai X, Zhao Y, Zhang W, Lu X. 2018. Deletion of a gene encoding a putative peptidoglycan-associated lipoprotein prevents degradation of the crystalline region of cellulose in *Cytophaga hutchinsonii*. *Front Microbiol* 9:632. <https://doi.org/10.3389/fmicb.2018.00632>.
58. McBride MJ, Baker SA. 1996. Development of techniques to genetically manipulate members of the genera *Cytophaga*, *Flavobacterium*, *Flexibacter*, and *Sporocytophaga*. *Appl Environ Microbiol* 62:3017–3022.
59. Thein M, Sauer G, Paramasivam N, Grin I, Linke D. 2010. Efficient sub-fractionation of Gram-negative bacteria for proteomics studies. *J Proteome Res* 9:6135–6147. <https://doi.org/10.1021/pr1002438>.
60. Arntzen MO, Karlskas IL, Skaugen M, Eijsink VG, Mathiesen G. 2015. Proteomic investigation of the response of *Enterococcus faecalis* V583 when cultivated in urine. *PLoS One* 10:e0126694. <https://doi.org/10.1371/journal.pone.0126694>.
61. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. 2006. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* 1:2856–2860. <https://doi.org/10.1038/nprot.2006.468>.
62. Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 26:1367–1372. <https://doi.org/10.1038/nbt.1511>.
63. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. 2011. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* 10:1794–1805. <https://doi.org/10.1021/pr101065j>.
64. Vizcaíno JA, Côté RG, Csordas A, Dianes JA, Fabregat A, Foster JM, Griss J, Alpi E, Birim M, Contell J, O’Kelly G, Schoenegger A, Ovelheiro D, Pérez-Riverol Y, Reisinger F, Ríos D, Wang R, Hermjakob H. 2012. The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res* 41:D1063–D1069. <https://doi.org/10.1093/nar/gks1262>.
65. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ, Brinkman FSL. 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 26:1608–1615. <https://doi.org/10.1093/bioinformatics/btq249>.