Current models in bacterial hemicellulase-encoding gene regulation

Jessica K. Novak and Jeffrey G. Gardner[#]

Running Title
Regulation of bacterial hemicellulase-encoding genes

Keywords
Carbon Catabolite Repression, Carbohydrate Active Enzyme, Extra Cytoplasmic Function, Hemicellulose, Hybrid Two-Component Systems, Transcription Factor

Author Affiliations
Department of Biological Sciences, University of Maryland – Baltimore County
Baltimore, Maryland, USA

[#]Correspondence
Jeffrey G. Gardner
Department of Biological Sciences
University of Maryland – Baltimore County
Email: jgardner@umbc.edu
Phone: 410-455-3613
Fax: 410-455-3875

ABSTRACT

The discovery and characterization of bacterial carbohydrate active enzymes is a fundamental component of biotechnology innovation, particularly for renewable fuels and chemicals, however these studies have increasingly transitioned to exploring the complex regulation required for recalcitrant polysaccharide utilization. This pivot is largely due to the current need to engineer and optimize enzymes for maximal degradation in industrial or biomedical applications. Given the structural simplicity of a single cellulose polymer, and the relatively few enzyme classes required for complete bioconversion, the regulation of cellulases in bacteria has been thoroughly discussed in the literature. However, the diversity of hemicelluloses found in plant biomass and the multitude of carbohydrate active enzymes required for their deconstruction has resulted in a less comprehensive understanding of bacterial hemicellulase-encoding gene regulation. Here we review the mechanisms of this process and common themes found in the transcriptomic response during plant biomass utilization. By comparing regulatory systems from both Gram-negative and Gram-positive bacteria, as well as drawing parallels to cellulase regulation, our goals are to highlight the shared and distinct features of bacterial hemicellulase-encoding gene regulation and provide a set of guiding questions to improve our understanding of bacterial lignocellulose utilization.

INTRODUCTION

The decomposition of plant biomass plays a significant role in environmental and biotechnological settings (Zhang et al. 2020). As the largest source of renewable carbon on the planet, the deconstruction of its polysaccharide components are heavily studied (Von Freiesleben et al. 2018; Michalak et al. 2020; Mhatre et al. 2022). Plant cell wall polysaccharides are broadly classified as either cellulose or hemicellulose. Cellulose polymers are exclusively comprised of glucose with a single linkage type (Gardner and Blackwell 1974). Alternatively, hemicelluloses possess greater linkage and sugar varieties which can include xyloglucans, xylans, mannans, arabinans, and pectins (Hoch 2007). This diversity in linkage and sugar type contributes to the insolubility and recalcitrance of plant cell walls, making them difficult to degrade (Holland et al. 2020).

Environmental bacteria and fungi are the central decomposers of this material (Pascoal et al. 2021), and produce Carbohydrate-Active Enzymes (CAZymes) for its deconstruction (Henrissat et al. 2022). Considerable biochemical and bioinformatic research has organized CAZymes into classes and families based on amino acid sequence and are documented in the CAZy database (Drula et al. 2022). This resource has facilitated efforts to predict and sort novel CAZymes for evolutionary phylogeny studies of lignocellulose degradation (Aspeborg et al. 2012; Wu et al. 2023), as well as enzyme engineering for industrial applications (Chettri and Verma 2023; Jayachandran et al. 2023).

As bacterial lignocellulose degradation systems become more fully described, work has branched out to several new areas to include the regulation of CAZyme-encoding genes. While cellulase systems in both Gram-negative and Gram-positive

bacteria have been reviewed (Liu et al. 2021; Ziles Domingues et al. 2022), there have been much fewer for hemicellulase systems because of the large number of substrates and enzymes required, as well as the assertion that Carbon Catabolite Repression (CCR) is the dominant modulator of gene expression (Stülke and Hillen 1999). Despite these challenges, recent hemicellulase-encoding gene regulation studies have characterized novel systems that were leveraged to engineer a single bacterium capable of fully degrading and fermenting lignocellulose (Mhatre et al. 2022; Singhania et al. 2022).

The goal of this review is to consolidate previously summarized work for a single phyla (Grondin et al. 2017; Lee et al. 2020) and provide commentary on the current direction of regulation-based studies for genes encoding hemicellulases like: xyloglucanases, xylanases, mannanases, arabinanases, and pectinases in both Gram-negative and Gram-positive bacteria. Furthermore, this review discusses the breadth of knowledge regarding CAZyme-encoding gene regulatory systems to include the recent influx of transcriptomic and computational studies that predict regulons specific to hemicellulase-encoding genes. We conclude with a few open questions and offer suggestions on promising future directions for studying the regulation of hemicellulase-encoding genes that may be of environmental or industrial interest.

CANONICAL REGULATORY MECHANISMS FOR BACTERIAL HEMICELLULASE-ENCODING GENE EXPRESSION

Expression of CAZyme-encoding genes requires precise regulation to ensure efficient energy expenditure under specific nutrient conditions. Despite the multitude of mechanisms that bacteria employ to regulate gene expression, there are only three systems commonly used for CAZyme-encoding genes, specifically hybrid two-component systems, extra cytoplasmic function-σ/anti-σ systems, and carbon catabolite repression (**Fig. 1**). Given that these regulatory systems have been comprehensively reviewed previously (Liu et al. 2019; Pinto et al. 2019; Franzino et al. 2022), we will only briefly summarize each of their general functions and the current knowledge on these systems that is relevant for the expression of hemicellulase-encoding genes.

Hybrid two-component systems. Hybrid two-component systems (HTCS) in bacteria use a sensing/phosphorylation relay mechanism to up-regulate genes involved in antibiotic resistance, virulence, biofilm formation, quorum sensing, and carbohydrate metabolism (Gutu et al. 2013; Cui et al. 2018; Gellatly et al. 2018; Kampik et al. 2020). This system, which is found in both Gram-negative and Gram-positive bacteria, recognizes an external stimulus with a cytoplasmic membrane protein that initiates a phosphorylation cascade to modulate gene expression (Howell et al. 2003). As shown in **Fig. 1A**, a substrate binds the sensor domain of a transmembrane histidine kinase. Substrate binding initiates a phosphate transfer from ATP to a histidine residue on the cytoplasmic domain. The phosphorylated histidine kinase then transfers the phosphate to a response regulator which binds the transcriptional start site of interest to modulate

transcription (Buschiazzo and Trajtenberg 2019; Francis and Porter 2019). It should be noted that there are examples of much lengthier phospho-relays with additional histidine kinases and response regulators before RNA polymerase recruitment. Two specific examples can be found in *Bacteroides thetaiotaomicron* and *Bacillus cereus* for glycan utilization and stress response, respectively (Been et al. 2006; Sonnenburg et al. 2006).

Previous research on hybrid two-component systems characterized the regulation of genes encoding xylanases, glucanases, arabinanases, and esterases from a diverse set of Gram-negative and Gram-positive bacteria (Emami et al. 2009; Martens et al. 2011; Shulami et al. 2014; Kampik et al. 2020). For example, in Gram-negative *Cellvibrio japonicus, Bacteroides thetaiotaomicron*, and Gram-positive *Ruminiclostridium cellulolyticum*, it was noted that HTCS regulators induced expression for biochemically or physiologically important xylanase-, arabinosidase-, and esterase-encoding genes (Emami et al. 2009; Martens et al. 2011; Kampik et al. 2020). The characterized HTCSs associated with xylanase and arabinanase-encoding gene expression are now cataloged as response regulators belonging to the AraC/XylS family of transcriptional activators (Emami et al. 2009; Celik et al. 2013). This family has recently been reviewed and is categorized based on two characteristic helix-turn-helix DNA-binding motifs (Cortes-Avalos et al. 2021). Regulation predominantly occurs via activation when the phosphorylated regulator binds to a recognized -10 and -35 consensus sequence up-stream of the promoter for RNA polymerase recruitment (Celik et al. 2013). The sensing domains of these HTCS bind branched xylo-oligosaccharides or arabino-oligosaccharides in the periplasmic space for Gram-negative bacteria (Emami et al. 2009; Schwalm III et al. 2017) and extracellularly for Gram-positive

bacteria (Lansky et al. 2020). For the former, species like *C. japonicus* and *B. thetaiotaomicron* require an efficient mechanism to degrade extracellular hemicellulose into oligosaccharides and transport them to the periplasm where they can be sensed by the corresponding HTCS. It is therefore unsurprising that these two species possess a disproportionally high number of outer membrane transporters that can bring large complex oligosaccharides into the periplasm (Emami et al. 2009; Larsbrink et al. 2014; Blake et al. 2018; Pollet et al. 2021).

Extra Cytoplasmic Function (ECF)-σ/anti-σ systems. Similar to HTCS, Extra Cytoplasmic Function (ECF)-σ/anti-σ systems are also comprised of a membrane-spanning sensory protein with a cytoplasmic regulatory protein partner that controls gene expression, with specific roles in bacterial virulence, stress response, and carbohydrate catabolism (Sun et al. 2004; Alvarez-Martinez et al. 2007; Wang et al. 2019a). ECF-σ/anti-σ systems are found in both Gram-negative and Gram-positive bacteria, but have been most well-characterized in Actinobacteria and human gut symbionts belonging to the *Bacteroides* phylum (Martens et al. 2009; Bahari et al. 2011; Huang et al. 2015; Despres et al. 2016a; Wang et al. 2019a). The anti-σ element of the system is a protein in the cytoplasmic membrane that binds a cytoplasmic ECF-σ protein (Helmann 2002) (**Fig. 1B**). Release of the ECF-σ protein occurs upon substrate binding, which can be a glycan, metal, or chemical stressor like limonene (Pudio et al. 2015; Marcos-torres et al. 2016; Goutam et al. 2017). The freed σ-factor then binds to RNA polymerase, forming a holoenzyme, and initiates transcription after binding a recognized consensus mRNA sequence (Bae et al. 2015).

In the context of carbohydrate catabolism, ECF-σ/anti-σ systems are prominent regulators in human gut symbionts, especially for the expression of genes encoding *O*-glycan-degrading enzymes (Martens et al. 2008). ECF-σ/anti-σ systems in *Bacteroides* species also frequently regulate genes encoding TonB-dependent transporters (*e.g.* SusC/D) (Martens et al. 2009). Furthermore, Gram-negative *Cytophaga hutchinsonii* and Gram-positive *Clostridium thermocellum* also have well-characterized ECF-σ/anti-σ systems that regulate cellulase-encoding gene expression (Nataf et al. 2010; Sand et al. 2015; Wang et al. 2019a). In *C. thermocellum*, cellulosomes are assembled using at least six ECF-σ/anti-σ systems that are specific for distinct cellulolytic regulons (Ortiz de Ora et al. 2018; Ichikawa et al. 2022).

In contrast to what is known about ECF-σ/anti-σ system to control cellulase-encoding genes, the regulatory involvement of ECF-σ/anti-σ systems for hemicellulase-encoding genes is less understood. Using the best described examples from Actinobacteria, ECF-σ/anti-σ systems have been placed into families based on the regulons they control (Huang et al. 2015). For example, ECF families 52 and 53 have been computationally predicted to possess a C-terminal fusion domain comprised of the anti-sigma factor sequence coupled with a transmembrane portion of the protein (Marcos-Torres et al. 2022). More interestingly, some ECF52 and ECF53 proteins also have computationally predicted glycosyl hydrolase catalytic domains and carbohydrate-binding domains (Huang et al. 2015; Pinto et al. 2019), however experimental validation has yet to be performed. In *C. thermocellum* xylanase-encoding genes are regulated by alternative sigma factors σ^{I6} and σ^{I7} and the vegetative promoter σ^{A} (Sand et al. 2015; Ichikawa et al. 2022). It was demonstrated that the vegetative σ^{A} provided basal

expression of xylanase-encoding genes, while σ^{I6} and σ^{I7} were employed for stronger expression in the presence of xylans (Bahari et al. 2011; Sand et al. 2015). Furthermore, the characterization of *C. thermocellum* ECF-σ/anti-σ systems aided in the prediction of homologous regulators in related species like *Psuedobacteroides cellulosolvens*, specifically for a pectin-degrading regulon (Ortiz de Ora et al. 2018).

Carbon Catabolite Repression. The final canonical system, carbon catabolite repression (CCR), is widely known for controlling the preferential utilization of specific carbon sources (typically glucose) over others (Ammar et al. 2018). In contrast to HTCS and ECF systems, which work similarly in Gram-negative and Gram-positive bacteria, the CCR mechanism in Gram-negative is markedly different compared to Gram-positive bacteria (Kundig et al. 1964; Deutscher and Saier Jr 1983). In Gram-negative bacteria, a phosphotransferase system is utilized wherein glucose is imported intracellularly and simultaneously phosphorylated by a component of the transport protein (EII_A). Expression of non-glucose metabolizing genes have very low basal expression and require activation (**Fig. 1C**). A phosphorylated EI protein transfers a phosphate group to the HPr protein, which in turn phosphorylates EII_A. In the absence of glucose, there is an abundance of phosphorylated EII_A (EII_A~P), which activates adenylate cyclase (AC) via phosphorylation (Magasanik 1961; Feucht and Saier 1980). The resulting accumulation of cAMP activates the cAMP Receptor Protein (CRP) and increases the transcription of genes that encode the proteins responsible for the metabolism of non-preferred carbon sources.

In Gram-positive bacteria, expression of genes important to the metabolism of non-glucose sugars requires inactivation of the repressor catabolite control protein (CcpA) (**Fig. 1D**). This occurs in the absence of glucose wherein fructose 1,6-bisphosphate (FBP) is not generated because glycolysis is not occurring. Without FBP, histidine protein (HPr) cannot be phosphorylated and dimerize with CcpA to repress transcription of genes involved in metabolizing non-preferred carbon sources (Deutscher and Saier Jr 1983). It should be noted that CcpA can also act as a transcriptional activator for quorum sensing (*trpA*), stress response (*cidAB*), and export of excess carbon (*ackA*) in *Streptococcus pneumoniae, Streptococcus mutans,* and *Bacillus subtilis* respectively (Henkin 1996; Kim et al. 2019a). Additionally, other counter-examples of CCR in *Pseudomonas* sp. found preferential utilization of succinate, citrate, or aromatic compounds over glucose (Liu 1952; Basu et al. 2006).

One example of CCR-based regulation for hemicellulase-encoding genes is found in *Bacillus subtilis* and uses both CcpA and the repressor GmuR (Sadaie et al. 2008). Mannanase-encoding genes in *B. subtilis* are in an operon that also contains genes encoding substrate-specific transporters and metabolic enzymes. In the presence of cellobiose or mannobiose (and in the absence of glucose), expression of the mannan utilization operon occurs due to a lack of fructose 1,6-bisphosphate. This results in limited amounts of phosphorylated HPr, which is necessary for CcpA binding to the promoter region. Consequently, the mannanase-encoding genes are de-repressed. Mannanase-encoding genes are further regulated by the repressor GmuR, which requires phosphorylation via GmuA, a component protein of the phosphotransferase system (PTS) and a structural homolog to EII_A (Sadaie et al. 2008). Briefly,

glucomannan disaccharides are imported and phosphorylated via the PTS (comprised of transport proteins GmuABC). Inverse to the processes described for carbon catabolite repression, the presence of glucomannan oligosaccharides results in an abundance of unphosphorylated GmuA. Consequently, GmuR cannot be phosphorylated, which results in the transcription of mannanase-encoding genes.

Co-regulation of arabinanase and xylanase-encoding genes are found in Gram-negative and Gram-positive bacteria, with two characterized repressors being AraR and XylR (Laikova et al. 2001; Rodionov et al. 2001). Both belong to the LacI family of transcriptional regulators and work in conjunction with CCR (Book et al. 2016; Ohashi et al. 2021; Rodionov et al. 2021). Co-regulation of xylanase and arabinanase genes provides an efficient means of streamlining gene expression given the monosaccharide composition of lignocellulose, namely hexoses coming from cellulose and pentoses coming from hemicellulose (Jamander et al. 2014; Kim et al. 2015). Not surprisingly, CCR has been widely studied to characterize the regulation of lignocellulose-derived sugar metabolism in *Clostridium, Caldicellulosiruptor, Pseudomonas,* and *Escherichia* species (Gosset 2005; Vanfossen et al. 2009; Bruder et al. 2015; Liu et al. 2015).

Current Applications of Canonical Systems. The use of bacteria as lignocellulose bioprocessors has renewed interest in the three canonical regulatory mechanisms for biotechnologically relevant bacteria (Mearls et al. 2015; Taylor II et al. 2018; Elmore et al. 2020; Ling et al. 2022). Using HTCS and ECF-σ/anti-σ systems, recent studies have focused on regulation of polysaccharide utilization loci (PULs) containing hemicellulase-encoding genes, especially in *Bacteroides* sp. (Luis et al. 2018; Mackie and Cann 2018;

Pereira et al. 2021; Beidler et al. 2023). Similarly, *C. thermocellum* is commonly used to study ECF-σ/anti-σ systems due to it possessing unique σ^I factors that can be studied heterologously in *B. subtilis* (Munoz-Gutierrez et al. 2016). Comparative studies of *C. thermocellum* σ^I factors were also important to the discovery that transcriptional initiation of cellulosomal genes relied on an auto-proteolysis system for ECF upon binding to the glycan of interest (Chen et al. 2023a). Likewise, dismantling CCR-related mechanisms in biotechnologically relevant bacteria (*e.g. E. coli, C. thermocellum,* and *P. putida*) found that co-utilization of xylose and glucose is more easily achieved with intracellular cellobiose hydrolysis (Xiong et al. 2018; Wang et al. 2019b; Cabulong et al. 2021). Intracellular cellobiose hydrolysis and phosphorylation bypassed some of the inhibitory effects caused by bacterial sensing/detection of extracellular glucose. Moreover, *Pseudomonas putida* KT2440 has undergone extensive engineering to co-metabolize glucose with cellobiose, galactose, xylose, and arabinose (Dvorak and de Lorenzo 2018; Peabody V et al. 2019; Elmore et al. 2020).

TRANSCRIPTOMIC APPROACHES TO IDENTIFY HEMICELLULASE-ENCODING GENE REGULATORY PATTERNS

The use of transcriptomic data to assess global changes in CAZyme-encoding gene regulation has rapidly become a standard approach to identify critical components of polysaccharide degradation (Gruninger et al. 2018; Lillington et al. 2020; Chen et al. 2023b). This method is particularly useful for non-model bacterial systems whose regulatory mechanisms are less characterized compared to *E. coli* or *B. subtilis.* While it should be noted that CAZyme-encoding gene expression was previously known to be

290 regulated by growth rate and bacterial life cycle for *Bacteroides succinogenes* and
291 *Clostridium thermocellum* (Russell 1987; Rydzak et al. 2012), more recent reports have
292 uncovered unique differences in hemicellulase-encoding gene regulation for both Gram-
293 positive and Gram-negative bacteria. Below is a summarization of the recent
294 developments using transcriptomics to elucidate regulatory features in lignocellulose-
295 degrading bacteria.
296
297 **Hemicellulase gene expression in Gram-positive species.** Current RNAseq
298 analyses using Gram-positive bacteria grown on hemicelluloses have often revealed
299 highly specific gene expression responses (Blumer-schuette et al. 2017; La Rosa et al.
300 2019; Rodionov et al. 2021). For example, the human gut symbiont *Roseburia*
301 *intestinalis* has a substrate-specific response during growth on glucomannan and
302 galactomannan (**Fig. 2A**) (La Rosa et al. 2019). Notably, 16 up-regulated genes were
303 from two distinct mannan utilization loci that differ from PULs described in *Bacteroides*
304 by the absence of genes that encode TonB-dependent transporters. Additionally, *R.*
305 *intestinalis* growth on galactose (a component of galactomannan) did not result in up-
306 regulation of any of these genes, suggesting that mannose or manno-oligosaccharides
307 were the sole nutritional signal for mannan deconstruction (La Rosa et al. 2019).
308 Highly specific CAZyme-encoding gene regulation has been observed in *Bacillus*
309 sp. N16-5, where up-regulation of β-mannanase and α-galactosidase-encoding genes
310 was only observed when the bacterium was grown using galactomannan, but not on
311 xylan, pectin, CMC, or any tested monosaccharide (glucose, fructose, mannose,
312 galactose, arabinose, or xylose) (Song et al. 2013). Furthermore, *Bacillus* sp. N16-5

grown using xylan only up-regulated β-xylanase-encoding genes, but growth on xylan or xylose up-regulated xylulokinase and xylose-related transporter-encoding genes.

As a third example, in *Caldicellulosirupter* species like *C. bescii* and *C. saccharolyticus,* xylanase-encoding genes were strongly up-regulated during growth on xylan (**Fig. 2B**) but repressed on either xylose or cellulose (Blumer-schuette et al. 2017; Rodionov et al. 2021). Expression data of *C. bescii* when grown using xylan also identified a putative key xylanase for extracellular xylan degradation (Xyn11A-2) (Crosby et al. 2022); however, a comparison of enzymatic activity between the *C. bescii* xylanases showed relatively mediocre activity for Xyn11A-2. The authors suspect this observed difference in gene expression could be a compensatory mechanism to overcome modest activity of Xyn11A-2. The use of transcriptomic data from *C. bescii* when grown on xylan has also proven useful for pairing the important degradative loci to their likely regulators, which included XynR, XylR, AraR, BxgRS, and AxuRS (Rodionov et al. 2021). Interestingly, transcriptomic analysis of *C. saccharolyticus* grown using pectin found a much broader gene expression response than that observed on other hemicelluloses (Blumer-schuette et al. 2017). Growth of *C. saccharolyticus* using pectin elicited up-regulation of various CAZyme-encoding genes, including cellulases, mannanases, xylanases, arabinanases, pectinases, and chitinases (**Fig. 2C**).

As a final example, *Clostridium* sp. exhibited some divergence in their regulatory circuits for xylanase-encoding genes (Petit et al. 2015; Munir et al. 2016). The expression of xylanase-encoding genes possessed by *C. termitidis* were dependent on xylan, but not xylose, cellobiose, or cellulose, while those belonging to *C. phytofermentans* were up-regulated when grown on both xylan and cellulose. Alongside

the differences in hemicellulase gene expression observed between growth media, growth-rate is also a critical mediator of CAZyme gene expression in *Clostridium* sp., with several studies reporting *C. thermocellum* transcription of cellulase-encoding genes dependent upon growth phase (Dror et al. 2003; Riederer et al. 2011). One interesting exception was for a xylanase-encoding gene (*xynC*), which exhibited high expression irrespective of growth rate (Dror et al. 2005).

Hemicellulase gene expression in Gram-negative species. For Gram-negative bacterial species, transcriptomic studies have revealed much broader gene expression responses than those observed in Gram-positive bacteria (Blake et al. 2018; Chen et al. 2018; Novak and Gardner 2023). For example, in *Leeuwenhoekiella* sp. MAR_2009_132, and *Salegentibacter* sp. Hel_I_6, up-regulated α- and β-mannanase-encoding genes were identified when these bacteria were grown on both α- or β-mannan despite the selective activity of these CAZymes for each substrate (Chen et al. 2018). This suggested that these species regulate mannanase gene expression with less specificity, possibly at the level of the mannose monosaccharide given that these bacteria cannot differentiate between α- versus β-mannan.

A broad gene expression response was revealed in the saprophyte *Cellvibrio japonicus* when grown on glucomannan (**Fig. 2A**) (Novak and Gardner 2023). Eight of the ten predicted mannanase-encoding genes were up-regulated, as well as an additional 46 CAZyme-encoding genes. Strong up-regulation of non-substrate specific CAZyme-encoding genes in *C. japonicus* suggests that it is likely the presence of complex polysaccharides that induce gene expression. Additionally, a previous study of

359 the *C. japonicus* transcriptomic response on cellobiose also resulted in broader up-
360 regulation of cellulases and hemicellulases (Nelson et al. 2017). Interestingly, a much
361 more specific response was elicited when *C. japonicus* was grown on oat-spelt xylan
362 (Blake et al. 2018). This report concluded that *C. japonicus* only up-regulated xylanase
363 genes during mid-exponential growth, though a comparison of the RNAseq from
364 stationary phase showed up-regulation of genes encoding xylanases, arabinanases,
365 mannanases, and cellulases. In terms of growth rate affecting CAZyme-encoding gene
366 expression in *C. japonicus*, it was observed that expression was more prominent during
367 active growth compared to stationary phase (Blake et al. 2018; Novak and Gardner
368 2023).

369 *Roseithermus sacchariphilus* exhibited a transcriptomic response quite dissimilar
370 to *C. japonicus* when it was grown on beechwood xylan (Liew et al. 2020). This
371 bacterium had up-regulation of genes encoding cellulases, mannanases, xylanases,
372 arabinanases, pectinases, and other glycosidases (**Fig. 2B**). Surprisingly, pectinase-
373 encoding genes were the most prominently up-regulated CAZyme-encoding genes
374 when *R. sacchariphilus* was grown on xylan. The authors hypothesize that the broad
375 response was due to co-expression of genes encoding various glycosidic activities by
376 the same promoter. However, they also suggested that a multi-timepoint transcriptomic
377 analysis could reveal more about the patterns of hemicellulase gene expression.

378 Finally, expression of CAZyme-encoding genes in *Bacteroides xylanisolvens* also
379 elicited a broad gene expression response on oat-spelt xylan, with up-regulation of 150
380 carbohydrate utilization-encoding genes that included all identified PULs for xylan
381 utilization and 15 PULs for starch and pectic metabolism (Despres et al. 2016a). The

authors hypothesized that the broad response was from detection of shared oligosaccharides present in both oat-spelt xylan and pectins (*i.e.* arabinoside side-chains). However, this response was very different when *B. xylanisolvens* was grown on citrus pectin and resulted in a much more specific result (**Fig. 2C**) (Despres et al. 2016b). Here, researchers were able to compare the gene expression response on two different types of pectins and discern the PULs that were most likely to be involved in the degradation of different pectic-linkages. Specifically, PUL 2 was suspected to be important to degrading type II rhamnogalacturonan, PUL 13 was likely involved in de-branching arabinose sidechains, and PULs 49 and 50 were the most up regulated on both pectins and were suspected to be involved in degrading homogalacturonan and type I rhamnogalacturonan, respectively. Additionally, *B. xylanisolvens* shared the traits observed in other bacterial species with high expression of CAZyme-encoding genes during active growth compared to stationary phase (Despres et al. 2016b).

Hemicellulase gene expression in bacterial co-culture. There has been increasing interest in the metatranscriptomic of co-cultured bacteria using complex polysaccharide-rich substrates given that environmental lignocellulose degradation is performed by a microbial community. For example, a study of the Gram-positive *Butyrivibrio hungatei* MB2003 transcriptome during mono- and co-culture with rumen gut symbiont *Butyrivibrio proteoclasticus* B316 found that in monoculture, *B. hungatei* was unable to grow on xylan or pectin despite the presence and expression of several hemicellulase-encoding genes (Palevich et al. 2019). Strikingly, when in co-culture with *B. proteoclasticus*, *B. hungatei* had a substantial increase in its growth capabilities at the

405 expense of *B. proteoclasticus* final cell density. Since *B. hungatei* acts more as a sugar
406 scavenger than a hemicellulose-degrader, its RNAseq results in monoculture
407 unsurprisingly showed marked increases in the expression of many genes important to
408 translation, signal transduction, defensive mechanisms, lipid/amino acid metabolism,
409 and cell wall biogenesis compared to its co-cultured counterpart. During co-culture with
410 *B. proteoclasticus, B. hungatei* expressed fewer genes overall but exhibited more
411 specificity in the expression of genes encoding for carbohydrate metabolism (*e.g.* ABC
412 sugar transporters). Interestingly, *B. proteoclasticus* gene expression was relatively
413 unchanged between mono- and co-culture (excluding a few CAZyme-encoding genes
414 which were up-regulated during co-culture) despite the increase in competition provided
415 by culturing with *B. hungatei*.

416 As another example, the Gram-negative gut symbionts *P. intestinalis, P. muris,*
417 and *P. rodentium* underwent comparative metatranscriptomic analysis, and the study
418 concluded that *P. intestinalis* was the most competitive strain due to its distinct up-
419 regulation of PULs encoding xylanase and pectinase-encoding genes when the rat host
420 was given a diet heavy in arabinoxylans (Galvez et al. 2020). The three most up-
421 regulated glycoside hydrolase families in all three species were from GH43, GH2, and
422 GH28. These families contain members able to hydrolyze β-glucan, β-xylan, α-arabinan,
423 and pectic linkages (Lombard et al. 2014).

424 Co-cultures containing both Gram-positive and Gram-negative species have
425 been used to investigate the bottlenecks of complete lignocellulose bioconversion in the
426 guts of rumen or humans (Leth et al. 2018; Badhan et al. 2022). A recent
427 metatranscriptomic study examined a complex consortium of Gram-positive and Gram-

negative gut symbionts in ruminant animals grown in *ex vivo* batch culture on Total Tract Indigestible Residue (TTIR). The primary goal of the study was to assess the bottlenecks in ruminant digestion to uncover mechanisms to enhance the system. Transcripts encoding xylanases were abundant when the micro-community was grown on TTIR, which indicated that heteroxylans and xyloglucans were the primary remaining polysaccharide in the TTIR. It was hypothesized that the sheer quantity of inter- and intramolecular bonds act as a hindrance to enzyme accessibility to the substrate.

Overall, there appears to be a distinguishing difference between the hemicellulose-encoding gene expression patterns in Gram-positive versus Gram-negative bacteria. Specifically, the narrowed specificity of gene expression observed in Gram-positive compared to Gram-negative species. Additionally, investigations of co-culture transcriptomics containing Gram-positive and/or Gram-negative communities on lignocellulose have focused on the interspecies relationships and competition for carbon acquisition (Palevich et al. 2019; Galvez et al. 2020; Badhan et al. 2022). The knowledge obtained from these analyses has subsequently been applied in studies on gut microbiomes and biotechnology applications, specifically for studies that successfully predicted the impact of synthetic gut microbiota on host immune response (Afrizal et al. 2022) and identified patterns in microbe abundance based on diet (Corbin et al. 2023).

Computational prediction of transcriptional regulators using compilations of transcriptomic data. In addition to the plethora of information provided by RNAseq data from a singular dataset, compilations of such data can extrapolate more

information on transcriptional regulatory systems using computational methods. For example, transcriptomic compilations with DNA-binding motif studies have predicted extensive transcriptional regulatory networks of several different bacteria (Poudel et al. 2020; Rychel et al. 2020). The known computationally predicted regulons of Gram-negative plant bioprocessors is relatively exclusive to the fermentative bioprocessing bacteria that possess few hemicellulases (Sastry et al. 2019; Lim et al. 2022). However, this approach has yielded interesting results for Gram-positive species. For example in *C. thermocellum,* a LacI transcriptional regulator (GlyR2) was computationally predicted as important for genes encoding two mannanases (*man5A* and *man26A*), a xylanase (*clo1313_2530*), and two cellulases (*clo1313_0413* and *clo1313_1425*) (Wilson et al. 2017; Hebdon et al. 2021). Previous experimental research on GlyR2 had identified it as a mannobiose-responsive transcriptional repressor with only confirmed regulatory activity on a mannanase-encoding gene (*man5A*) (Wilson et al. 2017). GlyR2 was hypothesized to have indirect effects on transcriptional regulation of certain hemicellulose-encoding genes that may require different conditions to de-repress other genes with the recognized binding motif (Hebdon et al. 2021). Additionally, a *C. bescii* genome analysis and comparison to other *Caldicellulosiruptor* species improved the organism-specific bioprocessing model through the discovery of 16 key regulators important to the degradation and metabolism of hemicellulose and pectin (Rodionov et al. 2021). It was noted that most of these regulators were involved in the expression of xylanase or pectinase-encoding genes, while genes that encoded cellulases, mannanases, and amylases generally only had one regulator for each CAZyme type. Additionally, the mechanistic regulatory actions of the predicted regulators were

overwhelmingly repressive in function with the few activators belonging to the AraC-family. Interestingly, the study found that most of these activators were involved in the regulation of pectinase-encoding genes.

FUTURE DIRECTIONS

A thorough understanding of how hemicellulase-encoding genes are regulated is essential to optimize lignocellulose bioprocessing (Chettri et al. 2020). Consequently, detailed studies that include hemicellulase-encoding gene regulation are generally conducted exclusively on well-characterized model bacteria and those already being used as chassis in biotechnology applications (Xiong et al. 2018; Rodionov et al. 2021). Given that metagenomic and metatranscriptomic data for less characterized lignocellulolytic bacteria with unoptimized systems are available (Dai et al. 2015; Kougias et al. 2018; Lopez-Mondejar et al. 2020) but beyond the scope of this review, we have endeavored to summarize and highlight the current state of hemicellulose-encoding gene regulation patterns between Gram-positive and Gram-negative bacteria. Overall, we argue there are two critical features of hemicellulase-encoding gene regulation that must be considered for optimization, which are (1) identifying the specific metabolic inducer (often an oligosaccharide), and (2) mitigating the impacts of carbon catabolite repression. Current lignocellulose bioconversion systems typically use Gram-positive species for saccharification and Gram-negative species for fermentation (Dai et al. 2015; Thapa et al. 2019). While it has been previously argued that co-culture or consortia-based bioconversion processes will improve efficiency and completeness of lignocellulose degradation (Chin et al. 2020; Kumar et al. 2023), the amount of strain

497 engineering and optimization significantly increases for each strain added to the
498 process, especially given the current trend of focusing only on improving either
499 degradative or metabolic/fermentative capabilities. Therefore, the following commentary
500 will focus exclusively on the optimization of single bacterium bioprocessing systems for
501 the complete deconstruction and utilization of lignocellulose (**Table 1**).
502
503 **Optimizing Gram-positive systems will require integration of degradative and**
504 **fermentative capabilities.** *Clostridia* and *Caldicellulosiruptor* species are highly studied
505 genera for their prolific degradation of plant polymers (Artzi et al. 2018; Brunecky et al.
506 2018; Williams-Rhaesa et al. 2018). However, neither system has been successfully
507 engineered to fully metabolize and ferment all components of lignocellulose. In the case
508 of *Clostridia* systems, this is due to an inherent inability to ferment pentoses. A previous
509 attempt to engineer a pathway for xylose fermentation in *C. thermocellum* found that
510 while xylose and Avicel could be co-utilized, xylan and Avicel could not (Xiong et al.
511 2018). It was argued that this is likely due an inhibitory effect posed by cello-
512 oligosaccharides on xylanases or unknown regulators that repress xylanase gene
513 expression in the presence of cello-oligosaccharides. More recently, efforts have
514 transitioned to develop CRISPR/Cas systems or riboswitches (Marcano-Velazquez et
515 al. 2019; Walker et al. 2020) to mediate the observed repression of xylanase gene
516 transcription in the presence of cellodextrins or cellobiose.
517 In *Caldicellulosiruptor* systems the limiting factor is that the expression and
518 degradative efficiency of heterologously expressed CAZymes is low. *C. bescii* has been
519 extensively manipulated to improve its saccharifying proficiency via heterologous

expression of xylanases (Kim et al. 2018; Crosby et al. 2022), however it has been observed that degraded oligosaccharides repress expression of secreted enzymes. Additionally, many heterologously expressed genes in *C. bescii* employ a highly active constitutive promoter, which is unoptimized for lignocellulose bioprocessing due to the energetic output required to constitutively and highly express heterologous CAZyme-encoding genes (Conway et al. 2017; Kim et al. 2017; Lee et al. 2020). Therefore, control over the expression of the heterologously expressed genes could spare the metabolic burden of their high expression levels and improve this limitation.

Optimizing Gram-negative systems will require bolstering the potency of lignocellulolytic capabilities. Gram-negative species elicit a much broader hemicellulase-encoding gene regulatory response than Gram-positive bacteria. We argue that this diversification of CAZyme gene expression is an underutilized resource to optimize lignocellulose bioconversion in single bacterium systems. Biotechnology-relevant model systems like *E. coli* and *P. putida* have been largely focused on improving co-utilization of hexoses and pentoses by overcoming the effects of CCR (Kim et al. 2019b; Peabody V et al. 2019; Elmore et al. 2020; Cabulong et al. 2021). However, these systems are limited as they are unable to innately degrade lignocellulose. The necessary step needed to drive either model into a fully self-sufficient system is the inclusion of lignocellulolytic machinery. This approach has several obstacles, most pressingly, identifying the minimally sufficient set of CAZymes that can completely depolymerize plant biomass and engineering an efficient export system for these CAZymes from the heterologous host.

In contrast, the genes/proteins needed to ferment plant sugars or produce other bioproducts are known and could be integrated into lignocellulolytic Gram-negative species. One example of a system not yet tapped for industrial use but has to potential to do so is *Cellvibrio japonicus*, a Gram-negative saprophyte that can fully degrade lignocellulose (Deboy et al. 2008; Gardner et al. 2014; Larsbrink et al. 2014; Blake et al. 2018). *C. japonicus* has also been shown to make ethanol and rhamnolipids as targeted products from lignocellulose bioconversion on a proof-of-concept scale (Gardner and Keating 2010; Horlamus et al. 2018). Another Gram-negative model is *Saccharophagus degradans* which also possesses a large number of CAZymes capable of degrading polysaccharides including cellulose, xylan, and pectin (Ensor et al. 1999). Engineering efforts using *S. degradans* have successfully generated strains capable of producing polyhydroxyalkanoate (PHAs) from cellulose, xylan, and agarose (Esteban Alva Munoz and Riley 2008; Sawant et al. 2017). However, *S. degradans* cannot generate ethanol and still relies on co-culture with other microbes for its production (Takagi et al. 2016). While both *C. japonicus* and *S. degradans* show promise with their degradative ability, improvements to their genetic systems are still needed to heterologously express the necessary metabolic pathways to produce high-value products.

Concluding statement. This review discussed mechanisms that regulate hemicellulase-encoding gene expression in Gram-positive versus Gram-negative bacteria. Experimental studies that characterize the molecular mechanisms of hemicellulase gene expression are useful to identify relevant activators or repressors for each regulon, and we argue that such research is essential for the field to significantly

advance. Given the discussed limitations of the reviewed models, the field should prioritize efforts that predict transcriptional regulatory networks and engineer the requisite enzymes for plant sugar bioconversion in species innately capable of prolific lignocellulose degradation.

DECLARATIONS

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

N/A

CONSENT FOR PUBLICATION

N/A

COMPETING INTERESTS

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. The authors also declare that they have no conflict of interest for this publication.

FUNDING

This work was supported by the National Science Foundation, Division of Environmental Biology under Award Number 2038304.

AUTHORS' CONTRIBUTIONS

JKN led the manuscript writing.

JGG contributed to writing the manuscript. All authors read and approved the final submitted version of the manuscript.

ACKNOWLEDGEMENTS

The authors wish to thank Gardner laboratory members for helpful discussions during the writing of this review.

AVAILABILITY OF DATA AND MATERIAL

N/A

REFERENCES

Afrizal A, Jennings SA V, Hitch TCA, Riedel T, Basic M, Panyot A, Treichel N, Hager FT, Wong EO-Y, Wolter B, Viehof A, von Strempel A, Eberl C, Buhl EM, Abt B, Bleich A, Tolba R, Blank LM, Navarre WW, Kiessling F, Horz H-P, Torow N, Cerovic V, Stecher B, Strowig T, Overmann J, Clavel T (2022) Enhanced cultured diversity of the mouse gut microbiota enables custom-made synthetic communities. Cell Host Microbe 30:1630–1645. https://doi.org/10.1016/j.chom.2022.09.011

Alvarez-Martinez CE, Lourenço RF, Baldini RL, Laub MT, Gomes SL (2007) The ECF sigma factor σT is involved in osmotic and oxidative stress responses in *Caulobacter crescentus*. Mol Biol 66:1240–1255. https://doi.org/10.1111/j.1365-2958.2007.06005.x

Ammar EM, Wang X, Rao C V (2018) Regulation of metabolism in *Escherichia coli* during growth on mixtures of the non-glucose sugars: arabinose, lactose, and xylose. Nat Sci Reports 8:1–11. https://doi.org/10.1038/s41598-017-18704-0

Artzi L, Dadosh T, Milrot E, Levin-zaidman S, Morag E, Bayer EA (2018) Colocalization and disposition of cellulosomes in *Clostridium clariflavum* as revealed by correlative superresolution imaging. MBio 9:e00012-18

Aspeborg H, Coutinho PM, Wang Y, Brumer H, Henrissat B (2012) Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). BMC Evol Biol 12. https://doi.org/10.1186/1471-2148-12-186

Badhan A, Low KE, Jones DR, Xing X, Milani MRM, Polo RO, Klassen L, Venketachalam S, Hahn MG, Abbott DW, McAllister TA (2022) Mechanistic insights into the digestion of complex dietary fibre by the rumen microbiota using

combinatorial high-resolution glycomics and transcriptomic analyses. Comput Struct Biotechnol J 20:148–164

Bae B, Feklistov A, Lass-Napiorkowska A, Landick R, Darst SA (2015) Structure of a bacterial RNA polymerase holoenzyme open promoter complex. Elife 4:1–23. https://doi.org/10.7554/eLife.08504

Bahari L, Gilad Y, Borovok I, Bareket HK, Yakir D, Shoham Y, Lamed R, Bayer EA (2011) Glycoside hydrolases as components of putative carbohydrate biosensor proteins in *Clostridium thermocellum*. J Ind Microbiol Biotechnol 38:825–832. https://doi.org/10.1007/s10295-010-0848-9

Basu A, Apte SK, Phale PS (2006) Preferential Utilization of Aromatic Compounds over Glucose by *Pseudomonas putida* CSV86. Appl Environ Microbiol 72:2226–2230. https://doi.org/10.1128/AEM.72.3.2226

Been M De, Francke C, Moezelaar R, Abee T, Siezen RJ (2006) Comparative analysis of two-component signal transduction systems of *Bacillus cereus*, *Bacillus thuringiensis* and *Bacillus anthracis*. Microbiology 152:3035–3048. https://doi.org/10.1099/mic.0.29137-0

Beidler I, Robb CS, Vidal-Melgosa S, Zühlke M-K, Bartosik D, Solanki V, Markert S, Becher D, Schweder T, Hehemann J-H (2023) Marine Bacteroidetes use a conserved enzymatic cascade to digest diatom β-mannan. ISME 17:276–285. https://doi.org/10.1038/s41396-022-01342-4

Blake AD, Beri NR, Guttman HS, Cheng R, Gardner JG (2018) The complex physiology of *Cellvibrio japonicus* xylan degradation relies on a single cytoplasmic β-xylosidase for xylo-oligosaccharide utilization. Mol Microbiol 107:610–622.

https://doi.org/10.1111/mmi.13903

Blumer-schuette SE, Zurawski J V, Conway JM, Khatibi P, Lewis DL, Li Q, Chiang VL, Kelly RM (2017) *Caldicellulosiruptor saccharolyticus* transcriptomes reveal consequences of chemical pretreatment and genetic modification of lignocellulose. Microb Biotechnol 10:1546–1557. https://doi.org/10.1111/1751-7915.12494

Book AJ, Lewin GR, Mcdonald BR, Takasuka TE, Fox G, Currie CR (2016) Evolution of High Cellulolytic Activity in Symbiotic *Streptomyces* through Selection of Expanded Gene Content and Coordinated Gene Expression. PLoS Biol 1–21. https://doi.org/10.1371/journal.pbio.1002475

Bruder M, Moo-young M, Chung DA, Chou CP (2015) Elimination of carbon catabolite repression in *Clostridium acetobutylicum* — a journey toward simultaneous use of xylose and glucose. Appl Microbiol Biotechnol 99:7579–7588. https://doi.org/10.1007/s00253-015-6611-4

Brunecky R, Chung D, Sarai NS, Hengge N, Russell JF, Young J, Mittal A, Pason P, Wall T Vander, Michener W, Shollenberger T, Westpheling J, Himmel ME, Bomble YJ (2018) Biotechnology for Biofuels High activity CAZyme cassette for improving biomass degradation in thermophiles. Biotechnol Biofuels 11:1–12. https://doi.org/10.1186/s13068-018-1014-2

Buschiazzo A, Trajtenberg F (2019) Two-component sensing and regulation: How do histidine kinases talk with response regulators at the molecular level? Annu Rev Microbiol 73

Cabulong RB, Bañares AB, Nisola GM, Lee W-K, Chung W-J (2021) Enhanced glycolic acid yield through xylose and cellobiose utilization by metabolically engineered

Escherichia coli. Bioprocess Biosyst Eng 44:1081–1091. https://doi.org/10.1007/s00449-020-02502-6

Celik H, Blouzard J-C, Voigt B, Becher D, Trotter V, Fierobe H-P, Tardif C, Pages S, Philip P De (2013) A Two-Component System (XydS/R) Controls the Expression of Genes Encoding CBM6-Containing Proteins in Response to Straw in *Clostridium cellulolyticum*. PLoS One 8:e56063. https://doi.org/10.1371/journal.pone.0056063

Chen C, Dong S, Yu Z, Qiao Y, Li J, Ding X, Li R, Lin J, Bayer EA, Liu Y, Cui Q, Feng Y (2023a) Essential autoproteolysis of bacterial anti-σ factor RsgI for transmembrane signal transduction. Sci Adv 14. https://doi.org/10.1126/sciadv.adg4846

Chen J, Robb CS, Unfried F, Kappelmann L, Markert S, Song T, Harder J, Avc B, Becher D, Xie P, Amann RI, Hehemann J, Schweder T (2018) Alpha- and beta-mannan utilization by marine Bacteroidetes. Environ Microbiol 20:4127–4140. https://doi.org/10.1111/1462-2920.14414

Chen L, Qu Z, Yu W, Zheng L, Qiao H, Wang D, Wei B, Zhao Z (2023b) Comparative genomic and transcriptome analysis of *Bacillus velezensis* CL-4 fermented corn germ meal. AMB Express 13:1–12. https://doi.org/10.1186/s13568-023-01510-5

Chettri D, Verma AK (2023) Biological significance of carbohydrate active enzymes and searching their inhibitors for therapeutic applications. Carbohydr Res 529:108853. https://doi.org/10.1016/j.carres.2023.108853

Chettri D, Verma AK, Verma AK (2020) Innovations in CAZyme gene diversity and its modification for biorefinery applications. Biotechnol Reports 28:1–17. https://doi.org/10.1016/j.btre.2020.e00525

Chin DWK, Lim S, Pang YL, Lam MK (2020) Fundamental review of organosolv

694 pretreatment and its challenges in emerging consolidated bioprocessing. Biofuels
695 Bioprod Biorefining 14:808–829. https://doi.org/10.1002/bbb.2096
696 Conway JM, McKinley BS, Seals NL, Hernandez D, Khatibi PA, Poudel S, Giannone
697 RJ, Hettich RL, Williams-Rhaesa AM, Lipscomb GL, Adams MWW, Kelly RM
698 (2017) Functional analysis of the glucan degradation locus in *Caldicellulosiruptor*
699 *bescii* reveals essential roles of component glycoside hydrolases in plant biomass
700 deconstruction. Appl Environ Microbiol 83:e01828-17.
701 https://doi.org/10.1128/AEM.01828-17
702 Corbin KD, Carnero EA, Dirks B, Igudesman D, Yi F, Marcus A, Davis TL, Pratley RE,
703 Rittmann BE, Krajmalnik-Brown R, Smith SR (2023) Host-diet-gut microbiome
704 interactions in fluence human energy balance: a randomized clinical trial. Nat
705 Commun 14:3161. https://doi.org/10.1038/s41467-023-38778-x
706 Cortes-Avalos D, Martinez-Perez N, Ortiz-Moncada MA, Juarez-Gonzalez A, Vanos-
707 Vargas AA, Estrada-de los Santos P, Perez-Rueda E, Ibarra JA (2021) An update
708 of the unceasingly growing and diverse AraC/XylS family of transcriptional
709 activators. FEMS Microbiol Lett 45:1–13
710 Crosby JR, Laemthong T, Bing RG, Zhang K, Tanwee TNN, Lipscomb GL, Rodionov
711 DA, Zhang Y, Adams MWW, Kelly RM (2022) Biochemical and Regulatory
712 Analyses of Xylanolytic Regulons in *Caldicellulosiruptor bescii* Reveal Genus-Wide
713 Features of Hemicellulose Utilization. Appl Environ Microbiol 88:e01302-22
714 Cui C, Yang C, Song S, Fu S, Sun X, Yang L, He F, Zhang L-H, Zhang Y, Deng Y
715 (2018) A novel two-component system modulates quorum sensing and
716 pathogenicity in *Burkholderia cenocepacia*. Mol Microbiol 108:32–44.

https://doi.org/10.1111/mmi.13915

Dai X, Tian Y, Li J, Su X, Wang X, Zhao S, Liu L, Luo Y, Liu D, Zheng H, Wang J, Dong Z, Hu S, Huang L (2015) Metatranscriptomic Analyses of Plant Cell Wall Polysaccharide Degradation by Microorganisms in the Cow Rumen. Appl Environ Microbiol 81:1375–1386. https://doi.org/10.1128/AEM.03682-14

Deboy RT, Mongodin EF, Fouts DE, Tailford LE, Khouri H, Emerson JB, Mohamoud Y, Watkins K, Henrissat B, Gilbert HJ, Nelson KE (2008) Insights into Plant Cell Wall Degradation from the Genome Sequence of the Soil Bacterium *Cellvibrio japonicus*. J Bacteriol 190:5455–5463. https://doi.org/10.1128/JB.01701-07

Despres J, Forano E, Lepercq P, Comtet-marre S, Jubelin G, Chambon C, Yeoman CJ, Miller MEB, Fields CJ, Martens E, Terrapon N, Henrissat B, White BA, Mosoni P (2016a) Xylan degradation by the human gut *Bacteroides xylanisolvens* XB1A involves two distinct gene clusters that are linked at the transcriptional level. BMC Genomics 17:1–14. https://doi.org/10.1186/s12864-016-2680-8

Despres J, Forano E, Lepercq P, Comtet-Marre S, Jubelin G, Yeoman CJ, Miller MEB, Fields CJ, Terrapon N, Bourvellec C Le, Renard CMGC, Henrissat B, White BA, Mosoni P (2016b) Unraveling the pectinolytic function of *Bacteroides xylanisolvens* using a RNA-seq approach and mutagenesis. BMC Genomics 17:1–14. https://doi.org/10.1186/s12864-016-2472-1

Deutscher J, Saier Jr MH (1983) ATP-dependent protein kinase-catalyzed phosphorylation of a seryl residue in HPr, a phosphate carrier protein of the phosphotransferase system in *Streptococcus pyogenes*. Proc Natl Acad Sci U S A 80:6790–6794. https://doi.org/10.1073/pnas.80.22.6790

740 Dror TW, Morag E, Rolider A, Bayer EA, Lamed R, Shoham Y (2003) Regulation of the
741 Cellulosomal *celS* (*cel48A*) Gene of *Clostridium thermocellum* Is Growth Rate
742 Dependent. J Bacteriol 185:3042–3048. https://doi.org/10.1128/JB.185.10.3042
743 Dror TW, Rolider A, Bayer EA, Lamed R, Shoham Y (2005) Regulation of Major
744 Cellulosomal Endoglucanases of *Clostridium thermocellum* Differs from That of a
745 Prominent Cellulosomal Xylanase. J Bacteriol 187:2261–2266.
746 https://doi.org/10.1128/JB.187.7.2261
747 Drula E, Garron M, Dogan S, Lombard V, Henrissat B, Terrapon N (2022) The
748 carbohydrate-active enzyme database: functions and literature. Nucleic Acids Res
749 50:571–577
750 Dvorak P, de Lorenzo V (2018) Refactoring the upper sugar metabolism of
751 *Pseudomonas putida* for co-utilization of cellobiose, xylose, and glucose. Metab
752 Eng 48:94–108
753 Elmore JR, Dexter GN, Salvachúa D, Brien MO, Klingeman DM, Gorday K, Michener
754 JK, Peterson DJ, Beckham GT, Guss AM (2020) Engineered *Pseudomonas putida*
755 simultaneously catabolizes five major components of corn stover lignocellulose :
756 Glucose , xylose , arabinose , p-coumaric acid , and acetic acid. Metab Eng 62:62–
757 71. https://doi.org/10.1016/j.ymben.2020.08.001
758 Emami K, Topakas E, Nagy T, Henshaw J, Jackson KA, Nelson KE, Mongodin EF,
759 Murray JW, Lewis RJ, Gilbert HJ (2009) Regulation of the Xylan-degrading
760 Apparatus of *Cellvibrio japonicus* by a Novel Two-component System. J Biol Chem
761 284:1086–1096. https://doi.org/10.1074/jbc.M805100200
762 Ensor LA, Stosz SK, Weiner RM (1999) Expression of multiple complex polysaccharide-

763 degrading enzyme systems by marine bacterium strain 2-40. J Ind Microbiol
764 Biotechnol 23:123–126
765 Esteban Alva Munoz L, Riley MR (2008) Utilization of Cellulosic Waste From Tequila
766 Bagasse and Production of Polyhydroxyalkanoate (PHA) Bioplastics by
767 *Saccharophagus degradans*. Agric Biosyst Eng 100:882–888.
768 https://doi.org/10.1002/bit.21854
769 Feucht BU, Saier MH (1980) Fine control of adenylate cyclase by the
770 phosphoenolpyruvate:sugar phosphotransferase systems in *Escherichia coli* and
771 *Salmonella typhimurium*. J Bacteriol 141:603–610.
772 https://doi.org/10.1128/jb.141.2.603-610.1980
773 Francis VI, Porter SL (2019) Multikinase Networks : Two-Component Signaling
774 Networks Integrating Multiple Stimuli. Annu Rev Microbiol 13:1–25
775 Franzino T, Boubakri H, Cernava T, Abrouk D, Achouak W, Reverchon S, Nasser W,
776 Haichar F el Z (2022) Implications of carbon catabolite repression for plant-microbe
777 interactions. Plant Commun 3:1–21. https://doi.org/10.1016/j.xplc.2021.100272
778 Galvez EJ, Iljazovic A, Amend L, Lesker TR, Renault T, Thiemann S, Hao L, Roy U,
779 Gronow A, Charpentier E, Strowig T (2020) Distinct polysaccharide utliization
780 determines interspeccies competition between intestinal *Prevotella* spp. Cell Host
781 Microbe 28:838–852
782 Gardner J, Keating DH (2010) Requirement of the Type II Secretion System for
783 Utilization of Cellulosic Substrates by *Cellvibrio japonicus*. Appl Environ Microbiol
784 76:5079–5087. https://doi.org/10.1128/AEM.00454-10
785 Gardner JG, Crouch L, Labourel A, Forsberg Z, Bukhman Y V, Vaaje-kolstad G, Gilbert

786 HJ, Keating DH (2014) Systems biology defines the biological significance of redox-
787 active proteins during cellulose degradation in an aerobic bacterium. 94:1121–
788 1133. https://doi.org/10.1111/mmi.12821
789 Gardner KH, Blackwell J (1974) The Structure of Native Cellulose. Biopolymers
790 13:1975–2001
791 Gellatly SL, Bains M, Breidenstein EBM, Strehmel J, Reffuveille F, Taylor PK, Yeung
792 ATY, Overhage J, Hancock REW (2018) Novel roles for two-component regulatory
793 systems in cytotoxicity and virulence-related properties in *Pseudomonas*
794 *aeruginosa*. AIMS Microbiol 4:173–191.
795 https://doi.org/10.3934/microbiol.2018.1.173
796 Gosset G (2005) Improvement of *Escherichia coli* production strains by modification of
797 the phosphoenolpyruvate:sugar phosphotransferase system. Microb Cell Fact 4:1–
798 11. https://doi.org/10.1186/1475-2859-4-14
799 Goutam K, Gupta AK, Gopal B (2017) The fused SnoaL 2 domain in the *Mycobacterium*
800 *tuberculosis* sigma factor oJ modulates promoter recognition. Nucleic Acids Res
801 45:9760–9772. https://doi.org/10.1093/nar/gkx609
802 Grondin JM, Tamura K, Déjean G, Abbott DW, Brumer H (2017) Polysaccharide
803 Utilization Loci: Fueling Microbial Communities. J Bacteriol 199:1–15
804 Gruninger RJ, Nguyen TT, Reid ID, Yanke JL, Wang P, Abbott DW, Tsang A, Mcallister
805 T (2018) Application of Transcriptomics to Compare the Carbohydrate Active
806 Enzymes That Are Expressed by Diverse Genera of Anaerobic Fungi to Degrade
807 Plant Cell Wall Carbohydrates. Front Microbiol 9:1–15.
808 https://doi.org/10.3389/fmicb.2018.01581

809 Gutu AD, Sgambati N, Strasbourger P, Brannon MK, Jacobs MA, Haugen E, Kaul RK,
810 Johansen HK, Hoiby N, Moskowitz SM (2013) Polymyxin Resistance of
811 *Pseudomonas aeruginosa phoQ* Mutants Is Dependent on Additional Two-
812 Component Regulatory Systems. Antimicrob Agents Chemother 57:2204–2215.
813 https://doi.org/10.1128/AAC.02353-12
814 Hebdon SD, Gerritsen AT, Chen Y, Marcano JG, Chou KJ (2021) Genome-Wide
815 Transcription Factor DNA Binding Sites and Gene Regulatory Networks in
816 *Clostridium thermocellum*. Front Microbiol 12:1–21.
817 https://doi.org/10.3389/fmicb.2021.695517
818 Helmann JD (2002) The Extracytoplasmic Function (ECF) Sigma Factors. Adv Microb
819 Physiol 46:47–110
820 Henkin TM (1996) The role of the CcpA transcriptional regulator in carbon metabolism
821 in *Bacillus subtilis*. FEMS Microbiol Lett 135:9–15
822 Henrissat B, Terrapon N, Coutinho PM, Lombard V, Drula E, Garron M-L, Hornung B
823 (2022) Carbonydrate-Active enZYmes Database. http://www.cazy.org/
824 Hoch G (2007) Cell wall hemicelluloses as mobile carbon stores in non-reproductive
825 plant tissues. Funct Ecol 21:823–834. https://doi.org/10.1111/j.1365-
826 2435.2007.01305.x
827 Holland C, Ryden P, Edwards CH, Grundy MML (2020) Plant Cell Walls : Impact on
828 Nutrient Bioaccessibility. MDPI Foods 9:1–16
829 Horlamus F, Wittgens A, Noll P, Michler J, Muller I, Weggenmann F, Oellig C, Rosenau
830 F, Henkel M, Hausmann R (2018) One - step bioconversion of hemicellulose
831 polymers to rhamnolipids with *Cellvibrio japonicus*: A proof-of-concept for a

potential host strain in future bioeconomy. Glob Chang Biol Bioenergy 11:260–268. https://doi.org/10.1111/gcbb.12542

Howell A, Dubrac S, Krogh K, Noone D, Fert J, Msadek T, Devine K (2003) Genes controlled by the essential YycG / YycF two- component system of *Bacillus subtilis* revealed through a novel hybrid regulator approach. Mol Microbiol 49:1639–1655. https://doi.org/10.1046/j.1365-2958.2003.03661.x

Huang X, Pinto D, Fritz G, Mascher T (2015) Environmental Sensing in Actinobacteria: a Comprehensive Survey on the Signaling Capacity of This Phylum. J Bacteriol 197:2517–2535. https://doi.org/10.1128/JB.00176-15

Ichikawa S, Ito D, Asaoka S, Abe R, Katsuo N, Ito T, Ito D, Karita S (2022) The expression of alternative sigma-I7 factor induces the transcription of cellulosomal genes in the cellulolytic bacterium *Clostridium thermocellum*. Enzyme Microb Technol 156:110002. https://doi.org/10.1016/j.enzmictec.2022.110002

Jamander J, Hallstrom BM, Larsson G (2014) Simultaneous Uptake of Lignocellulose-Based Monosaccharides by *Escherichia coli*. Biotechnol Bioeng 11:1108–1115. https://doi.org/10.1002/bit.25182

Jayachandran D, Smith P, Irfan M, Sun J, Yarborough JM, Bomble YJ, Lam E, Chundawat SPS (2023) Engineering and characterization of carbohydrate-binding modules for imaging cellulose fibrils biosynthesis in plant protoplasts. Biotechnol Bioeng 120:2253–2268. https://doi.org/10.1002/bit.28484

Kampik C, Denis Y, Pages S, Perret S, Tardif C, Fierobe H-P, de Philip P (2020) A novel two-component system, XygS/XygR, positively regulates xyloglucan degradation, import, and catabolism in *Ruminiclostridium cellulolyticum*. Appl

855 Environ Microbiol 86:e01357-20
856 Kim H, Waters A, Turner ME, Rice KC, Ahn S (2019a) Regulation of cid and lrg
857 expression by CcpA in *Streptococcus mutans*. Microbiology 165:113–123.
858 https://doi.org/10.1099/mic.0.000744
859 Kim J, Tremaine M, Grass JA, Purdy HM, Landick R, Kiley PJ, Reed JL (2019b)
860 Systems metabolic engineering of *Escherichia coli* improves co-conversion of
861 lignocellulose-derived sugars. Biotechnol J 14.
862 https://doi.org/10.1002/biot.201800441
863 Kim S-K, Himmel ME, Bomble YJ, Westpheling J (2018) Expression of a cellobiose
864 phosphorylase from *Thermotoga maritima* in *Caldicellulosiruptor bescii* improves
865 the phosphorolytic pathway and results in a dramatic increase in cellulolytic activity.
866 Appl Environ Microbiol 84:e02348-17. https://doi.org/10.1128/AEM.02348-17
867 Kim SK, Chung D, Himmel ME, Bomble YJ, Westpheling J (2017) Heterologous
868 expression of a β-d-glucosidase in *Caldicellulosiruptor bescii* has a surprisingly
869 modest effect on the activity of the exoproteome and growth on crystalline
870 cellulose. J Ind Microbiol Biotechnol 44:1643–1651. https://doi.org/10.1007/s10295-
871 017-1982-4
872 Kim SM, Choi BY, Ryu YS, Jung SH, Park JM, Kim G-H, Lee SK (2015) Simultaneous
873 utilization of glucose and xylose via novel mechanisms in engineered *Escherichia*
874 *coli*. Metab Eng 30:141–148. https://doi.org/10.1016/j.ymben.2015.05.002
875 Kougias PG, Campanaro S, Treu L, Tsapekos P, Armani A, Angelidaki I (2018) Spatial
876 Distribution and Diverse Metabolic Functions of Lignocellulose-Degrading
877 Uncultured Bacteria as Revealed by Genomic-Centric Metagenomics. Appl Environ

878 Microbiol 84:e01244-18
879 Kumar V, Fox BG, Takasuka TE (2023) Consolidated bioprocessing of plant biomass to
880 polyhydroxyalkanoate by co-culture of *Streptomyces* sp . SirexAA-E and Priestia
881 megaterium. Bioresour Technol 376:128934.
882 https://doi.org/10.1016/j.biortech.2023.128934
883 Kundig W, Ghosh S, Roseman S (1964) Phosphate bound to histidine in a protein as an
884 intermediate in a novel phospho-transferase system. Proc Natl Acad Sci U S A
885 52:1067–1074. https://doi.org/10.1073/pnas.52.4.1067
886 La Rosa SL, Leth ML, Michalak L, Hansen ME, Pudlo NA, Glowacki R, Pereira G,
887 Workman CT, Arntzen MØ, Pope PB, Martens EC, Hachem MA, Westereng B
888 (2019) The human gut Firmicute *Roseburia intestinalis* is a primary degrader of
889 dietary β -mannans. Nat Commun 10:1–14. https://doi.org/10.1038/s41467-019-
890 08812-y
891 Laikova ON, Mironov AA, Gelfand MS (2001) Computational analysis of the
892 transcriptional regulation of pentose utilization systems in the gamma subdivision of
893 Proteobacteria. FEMS Microbiol Lett 205:315–322
894 Lansky S, Salama R, Shulami S, Lavid N, Sen S, Schapiro I, Shoham Y, Shoham G
895 (2020) Carbohydrate-Binding Capability and Functional Conformational Changes of
896 AbnE, an Arabino-oligosaccharide Binding Protein. J Mol Biol 432:2099–2120.
897 https://doi.org/10.1016/j.jmb.2020.01.041
898 Larsbrink J, Rogers TE, Hemsworth GR, Mckee LS, Tauzin AS, Spadiut O, Klinter S,
899 Pudlo NA, Urs K, Koropatkin NM, Creagh AL, Haynes CA, Kelly AG, Cederholm
900 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

Lee LL, Crosby JR, Rubinstein GM, Laemthong T, Bing RG, Straub CT, Adams MWW, Kelly RM (2020) The biology and biotechnology of the genus *Caldicellulosiruptor*: recent developments in 'Caldi World.' Extremophiles 24:1–15. https://doi.org/10.1007/s00792-019-01116-5

Leth ML, Ejby M, Workman C, Ewald DA, Pedersen SS, Sternberg C, Bahl MI, Licht TR, Aachmann FL, Westereng B, Hachem MA (2018) Differential bacterial capture and transport preferences facilitate co-growth on dietary xylan in the human gut. Nat Microbiol 3:570–580. https://doi.org/10.1038/s41564-018-0132-8

Liew KJ, Bruce NC, Sani RK, Chong CS, Yaakop AS, Shamsir MS, Goh KM (2020) Global transcriptomic responses of *Roseithermus sacchariphilus* strain RA in media supplemented with beechwood xylan. Microorganisms 8:1–22. https://doi.org/10.3390/microorganisms8070976

Lillington SP, Leggieri PA, Heom KA, Malley MAO (2020) ScienceDirect Nature ' s recyclers : anaerobic microbial communities drive crude biomass deconstruction. Curr Opin Biotechnol 62:38–47. https://doi.org/10.1016/j.copbio.2019.08.015

Lim HG, Rychel K, Sastry A V, Bentley GJ, Mueller J, Schindel HS, Larsen PE, Laible PD, Guss AM, Niu W, Johnson CW, Beckham GT, Feist AM, Palsson BO (2022) Machine-learning from *Pseudomonas putida* KT2440 transcriptomes reveals its transcriptional regulatory network. Metab Eng 72:297–310. https://doi.org/10.1016/j.ymben.2022.04.004

Ling C, Peabody GL, Salvachúa D, Kim Y, Kneucker CM, Calvey CH, Monninger MA,

924 Munoz NM, Poirier BC, Ramirez KJ, John PCS, Woodworth SP, Magnuson JK,
925 Burnum-johnson KE, Guss AM, Johnson CW, Beckham GT (2022) Muconic acid
926 production from glucose and xylose in *Pseudomonas putida* via evolution and
927 metabolic engineering. Nat Commun 13:1–14. https://doi.org/10.1038/s41467-022-
928 32296-y
929 Liu C, Sun D, Zhu J, Liu W (2019) Two-Component Signal Transduction Systems : A
930 Major Strategy for Connecting Input Stimuli to Biofilm Formation. Front Microbiol 9.
931 https://doi.org/10.3389/fmicb.2018.03279
932 Liu L, Huang W-C, Liu Y, Li M (2021) Diversity of cellulolytic microorganisms and
933 microbial cellulases. Int Biodeterior Biodegradation 163:105277.
934 https://doi.org/10.1016/j.ibiod.2021.105277
935 Liu P (1952) Utilization of carbohydrates by *Pseudomonas aeruginosa*. J Bacteriol
936 64:773–781
937 Liu Y, Rainey PB, Zhang X-X (2015) Molecular mechanisms of xylose utilization by
938 *Pseudomonas fluorescens*: overlapping genetic responses to xylose, xylulose,
939 ribose and mannitol. Mol Microbiol 98:553–570. https://doi.org/10.1111/mmi.13142
940 Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-
941 active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:490–495.
942 https://doi.org/10.1093/nar/gkt1178
943 Lopez-Mondejar R, Tlaskal V, Vetrovsy T, Stursova M, Toscan R, Nunes da Rocha U,
944 Baldrian P (2020) Metagenomics and stable isotope probing reveal the
945 complementary contribution of fungal and bacterial communities in the recycling of
946 dead biomass in forest soil. Soil Biol Biochem 148:1–11.

947 https://doi.org/10.1016/j.soilbio.2020.107875
948 Luis AS, Briggs J, Zhang X, Farnell B, Ndeh D, Labourel A, Baslé A, Cartmell A,
949 Terrapon N, Stott K, Lowe EC, Mclean R, Shearer K, Schückel J, Venditto I, Ralet
950 M-C, Henrissat B, Martens EC, Mosimann SC, Abbott DW, Gilbert HJ (2018)
951 Dietary pectic glycans are degraded by coordinated enzyme pathways in human
952 colonic *Bacteroides*. Nat Microbiol 3:210–219. https://doi.org/10.1038/s41564-017-
953 0079-1
954 Mackie RI, Cann I (2018) Let them eat fruit. Nat Microbiol 3:127–129.
955 https://doi.org/10.1038/s41564-018-0108-8
956 Magasanik B (1961) Catabolite Repression. Cold Spring Harb Symp Quant Biol 26:249–
957 256. https://doi.org/10.1101/SQB.1961.026.01.031
958 Marcano-Velazquez JG, Lo J, Nag A, Maness PC, Chou KJ (2019) Developing
959 Riboswitch-Mediated Gene Regulatory Controls in Thermophilic Bacteria. ACS
960 Synth Biol 8:633–640. https://doi.org/10.1021/acssynbio.8b00487
961 Marcos-Torres FJ, Moraleda-Muñoz A, Contreras-Moreno FJ, Muñoz-Dorado J, Pérez J
962 (2022) Mechanisms of Action of Non-Canonical ECF Sigma Factors. Int J Mol Sci
963 23:3601. https://doi.org/10.3390/ijms23073601
964 Marcos-torres FJ, Perez J, Gomez-Santos N, Moraleda-Munoz A, Munoz-Dorado J
965 (2016) In depth analysis of the mechanism of action of metal-dependent sigma
966 factors: characterization of CorE2 from *Myxococcus xanthus*. Nucleic Acids Res
967 44:5571–5584. https://doi.org/10.1093/nar/gkw150
968 Martens EC, Chiang HC, Gordon JI (2008) Mucosal Glycan Foraging Enhances Fitness
969 and Transmission of a Saccharolytic Human Gut Bacterial Symbiont. Cell Host

970 Microbe 4:447–457. https://doi.org/10.1016/j.chom.2008.09.007
971 Martens EC, Lowe EC, Chiang H, Pudlo NA, Wu M, Nathan P, Abbott DW, Henrissat B,
972 Gilbert HJ, Bolam DN, Gordon JI (2011) Recognition and Degradation of Plant Cell
973 Wall Polysaccharides by Two Human Gut Symbionts. PLoS Biol 9:e1001221.
974 https://doi.org/10.1371/journal.pbio.1001221
975 Martens EC, Roth R, Heuser JE, Gordon JI (2009) Coordinate Regulation of Glycan
976 Degradation and Polysaccharide Capsule Biosynthesis by a Prominent Human Gut
977 Symbiont. J Biol Chem 284:18445–18457. https://doi.org/10.1074/jbc.M109.008094
978 Mearls EB, Olson DG, Herring CD, Lynd LR (2015) Development of a regulatable
979 plasmid-based gene expression system for *Clostridium thermocellum*. Appl
980 Microbiol Biotechnol 99:7589–7599. https://doi.org/10.1007/s00253-015-6610-5
981 Mhatre A, Kalscheur B, Mckeown H, Bhakta K, Sarnaik AP, Flores A, Nielsen DR,
982 Wang X, Soundappan T, Varman AM (2022) Consolidated bioprocessing of
983 hemicellulose to fuels and chemicals through an engineered *Bacillus subtilis*-
984 *Escherichia coli* consortium. Renew Energy 193:288–298.
985 https://doi.org/10.1016/j.renene.2022.04.124
986 Michalak L, La Rosa SL, Leivers S, Lindstad LJ, Røhr ÅK, Aachmann FL, Westereng B
987 (2020) A pair of esterases from a commensal gut bacterium remove acetylations
988 from all positions on complex β-mannans. Proc Natl Acad Sci U S A 117:7122–
989 7130. https://doi.org/10.1073/pnas.1915376117
990 Munir RI, Spicer V, Krokhin O V., Shamshurin D, Zhang X, Taillefer M, Blunt W, Cicek
991 N, Sparling R, Levin DB (2016) Transcriptomic and proteomic analyses of core
992 metabolism in *Clostridium termitidis* CT1112 during growth on α-cellulose, xylan,

993 cellobiose and xylose. BMC Microbiol 16:1–21. https://doi.org/10.1186/s12866-016-
994 0711-x
995 Munoz-Gutierrez I, Ortiz de Ora L, Grinberg IR, Garty Y, Bayer EA, Shoham Y, Lamed
996 R, Borovok I (2016) Decoding Biomass-Sensing Regulons of *Clostridium*
997 *thermocellum* Alternative Sigma-I Factors in a Heterologous *Bacillus subtilis* Host
998 System. PLoS One 11:e0146316. https://doi.org/10.1371/journal.pone.0146316
999 Nataf Y, Bahari L, Kahel-Raifer H, Borovok I, Lamed R, Bayer EA, Sonenshein AL,
1000 Shoham Y (2010) *Clostridium thermocellum* cellulosomal genes are regulated by
1001 extracytoplasmic polysaccharides via alternative sigma factors. Proc Natl Acad Sci
1002 U S A 107:18646–18651. https://doi.org/10.1073/pnas.1012175107
1003 Nelson CE, Rogowski A, Morland C, Wilhide JA, Gilbert HJ, Gardner JG (2017)
1004 Systems analysis in *Cellvibrio japonicus* resolves predicted redundancy of b -
1005 glucosidases and determines essential physiological functions. Mol Microbiol
1006 104:294–305. https://doi.org/10.1111/mmi.13625
1007 Novak JK, Gardner JG (2023) Galactomannan utilization by *Cellvibrio japonicus* relies
1008 on a single essential α-galactosidase encoded by the *aga27A* gene. Mol Microbiol
1009 119:312–325. https://doi.org/10.1111/mmi.15024
1010 Ohashi K, Hataya S, Nakata A, Matsumoto K, Kato N, Sato W, Carlos-Shanley C,
1011 Beebe ET, Currie CR, Fox BG, Takasuka TE (2021) Mannose- and Mannobiose-
1012 Specific Responses of the Insect-Associated Cellulolytic Bacterium *Streptomyces*
1013 sp. Strain SirexAA-E. Appl Environ Microbiol 87:e02719-20
1014 Ortiz de Ora L, Lamed R, Liu YJ, Xu J, Cui Q, Feng Y, Shoham Y, Bayer EA, Muñoz-
1015 Gutiérrez I (2018) Regulation of biomass degradation by alternative σ factors in

1016 cellulolytic clostridia. Sci Rep 8:1–11. https://doi.org/10.1038/s41598-018-29245-5
1017 Palevich N, Kelly WJ, Ganesh S, Rakonjac J, Attwood GT (2019) *Butyrivibrio hungatei*
1018 MB2003 competes effectively for soluble sugars released by *Butyrivibrio*
1019 *proteoclasticus* B316 during growth on xylan or pectin. Appl Environ Microbiol
1020 85:e02056-18
1021 Pascoal C, Fernandes I, Seena S, Danger M, Ferreira V, Cássio F (2021) Linking
1022 Microbial Decomposer Diversity to Plant Litter Decomposition and Associated
1023 Processes in Streams. In: Swan CM, Boyero L, Canhoto C (eds) The Ecology of
1024 Plant Litter Decomposition in Stream Ecosystems. Springer International
1025 Publishing, Cham, pp 163–192
1026 Peabody V GL, Elmore JR, Martinez-Baird J, Guss AM (2019) Engineered
1027 *Pseudomonas putida* KT2440 co-utilizes galactose and glucose. Biotechnol
1028 Biofuels 12:1–7. https://doi.org/10.1186/s13068-019-1627-0
1029 Pereira G V, Abdel-Hamid AM, Dutta S, D'Alessandro-Gabazza CN, Wefers D, Farris
1030 JA, Bajaj S, Wawrzak Z, Atomi H, Mackie RI, Gabazza EC, Shukla D, Koropatkin
1031 NM, Cann I (2021) Degradation of complex arabinoxylans by human colonic
1032 Bacteroidetes. Nat Commun 12:1–21. https://doi.org/10.1038/s41467-020-20737-5
1033 Petit E, Coppi M V, Hayes JC, Tolonen AC, Warnick T, Latouf WG, Amisano D, Biddle
1034 A, Mukherjee S, Ivanova N, Lykidis A, Land M, Hauser L, Kyrpides N, Henrissat B,
1035 Lau J, Schnell DJ, Church GM, Leschine SB, Blanchard JL (2015) Genome and
1036 Transcriptome of *Clostridium phytofermentans*, Catalyst for the Direct Conversion
1037 of Plant Feedstocks to Fuels. PLoS One 10:1–18.
1038 https://doi.org/10.1371/journal.pone.0118285

1039 Pinto D, Liu Q, Mascher T (2019) ECF σ factors with regulatory extensions: the one-
1040 component systems of the σ universe. Mol Microbiol 112:399–409.
1041 https://doi.org/10.1111/mmi.14323
1042 Pollet RM, Martin LM, Koropatkin NM (2021) TonB-dependent transporters in the
1043 Bacteroidetes: Unique domain structures and potential functions. Mol Microbiol
1044 115:490–501. https://doi.org/10.1111/mmi.14683
1045 Poudel S, Tsunemoto H, Seif Y, Sastry A V., Szubin R, Xu S, Machado H, Olson CA,
1046 Anand A, Pogliano J, Nizet V, Palsson BO (2020) Revealing 29 sets of
1047 independently modulated genes in *Staphylococcus aureus*, their regulators, and
1048 role in key physiological response. Proc Natl Acad Sci U S A 117:17228–17239.
1049 https://doi.org/10.1073/pnas.2008413117
1050 Pudio NA, Urs K, Kumar SS, German JB, Mills DA, Martens EC (2015) Symbiotic
1051 Human Gut Bacteria with Variable Metabolic Priorities for Host Mucosal Glycans.
1052 MBio 6:e01282-15. https://doi.org/10.1128/mBio.01282-15.Editor
1053 Riederer A, Takasuka TE, Makino S, Stevenson DM, Bukhman Y V, Elsen NL, Fox BG
1054 (2011) Global Gene Expression Patterns in *Clostridium thermocellum* as
1055 Determined by Microarray Analysis of Chemostat Cultures on Cellulose or
1056 Cellobiose. Appl Environ Microbiol 77:1243–1253.
1057 https://doi.org/10.1128/AEM.02008-10
1058 Rodionov DA, Mironov AA, Gelfand MS (2001) Transcriptional regulation of pentose
1059 utilisation systems in the *Bacillus* / *Clostridium* group of bacteria. FEMS Microbiol
1060 Lett 205:305–314
1061 Rodionov DA, Rodionova IA, Rodionov VA, Arzamasov AA, Zhang K, Rubinstein GM,

1062 Tanwee TNN, Bing RG, Crosby JR, Nookaew I, Basen M, Brown SD, Wilson CM,
1063 Klingeman DM, Poole II FL, Zhang Y, Kelly RM, Adams MWW (2021)
1064 Transcriptional Regulation of Plant Biomass Degradation and Carbohydrate
1065 Utilization Genes in the Extreme Thermophile *Caldicellulosiruptor bescii*. mSystems
1066 6
1067 Russell JB (1987) Effect of Extracellular pH on Growth and Proton Motive Force of
1068 *Bacteroides succinogenes*, a Cellulolytic Ruminal Bacterium. Appl Environ
1069 Microbiol 53:2379–2383
1070 Rychel K, Sastry A V, Palsson BO (2020) Machine learning uncovers independently
1071 regulated modules in the *Bacillus subtilis* transcriptome. Nat Commun 11:1–10.
1072 https://doi.org/10.1038/s41467-020-20153-9
1073 Rydzak T, McQueen PD, Krokhin O V., Spicer V, Ezzati P, Dwivedi RC, Shamshurin D,
1074 Levin DB, Wilkins JA, Sparling R (2012) Proteomic analysis of *Clostridium*
1075 *thermocellum* core metabolism: Relative protein expression profiles and growth
1076 phase-dependent changes in protein expression. BMC Microbiol 12:1–18.
1077 https://doi.org/10.1186/1471-2180-12-214
1078 Sadaie Y, Nakadate H, Fukui R, Yee LM, Asai K (2008) Glucomannan utilization operon
1079 of *Bacillus subtilis*. FEMS Microbiol Lett 279:103–109.
1080 https://doi.org/10.1111/j.1574-6968.2007.01018.x
1081 Sand A, Holwerda EK, Ruppertsberger NM, Maloney M, Olson DG, Nataf Y, Borovok I,
1082 Sonenshein AL, Bayer EA, Lamed R, Lynd LR, Shoham Y (2015) Three
1083 cellulosomal xylanase genes in *Clostridium thermocellum* are regulated by both
1084 vegetative SigA and alternative SigI6 factors. FEBS Lett 589:3133–3140.

https://doi.org/10.1016/j.febslet.2015.08.026

Sastry A V, Gao Y, Szubin R, Hefner Y, Xu S, Kim D, Choudhary KS, Yang L, King ZA, Palsson BO (2019) The *Escherichia coli* transcriptome mostly consists of independently regulated modules. Nat Commun 10:1–14. https://doi.org/10.1038/s41467-019-13483-w

Sawant SS, Salunke BK, Taylor LE, Kim BS (2017) Enhanced Agarose and Xylan Degradation for Production of Polyhydroxyalkanoates by Co-Culture of Marine Bacterium, *Saccharophagus degradans* and Its Contaminant, *Bacillus cereus*. MDPI Appl Sci 7:225. https://doi.org/10.3390/app7030225

Schwalm III ND, Townsed II GE, Groisman EA (2017) Prioritization of polysaccharide utilization and control of regulator activation in *Bacteroides thetaiotaomicron*. Mol Microbiol 104:32–45. https://doi.org/10.1111/mmi.13609

Shulami S, Shenker O, Langut Y, Lavid N, Gat O, Zaide G, Zehavi A, Sonenshein AL, Shoham Y (2014) Multiple Regulatory Mechanisms Control the Expression of the *Geobacillus stearothermophilus* Gene for Extracellular Xylanase. J Biol Chem 289:25957–25975. https://doi.org/10.1074/jbc.M114.592873

Singhania RR, Patel AK, Singh A, Haldar D, Soam S, Chen C-W, Tsai M-L, Dong C-D (2022) Consolidated bioprocessing of lignocellulosic biomass: Technological advances and challenges. Bioresour Technol 354:1–9. https://doi.org/10.1016/j.biortech.2022.127153

Song Y, Xue Y, Ma Y (2013) Global Microarray Analysis of Carbohydrate Use in Alkaliphilic Hemicellulolytic Bacterium *Bacillus* sp. N16-5. PLoS One 8:e54090. https://doi.org/10.1371/journal.pone.0054090

1108 Sonnenburg ED, Sonnenburg JL, Manchester JK, Hansen EE, Chiang HC, Gordon JI
1109 (2006) A hybrid two-component system protein of a prominent human gut symbiont
1110 couples glycan sensing *in vivo* to carbohydrate metabolism. Proc Natl Acad Sci U S
1111 A. https://doi.org/10.1073/pnas.0603249103
1112 Stülke J, Hillen W (1999) Carbon catabolite repression in bacteria. Curr Opin Microbiol
1113 2:195–201. https://doi.org/10.1016/S1369-5274(99)80034-4
1114 Sun R, Converse PJ, Ko C, Tyagi S, Morrison NE, Bishai WR (2004) *Mycobacterium*
1115 *tuberculosis* ECF sigma factor *sigC* is required for lethality in mice and for the
1116 conditional expression of a defined gene set. Mol Microbiol 52:25–38.
1117 https://doi.org/10.1111/j.1365-2958.2003.03958.x
1118 Takagi T, Morisaka H, Aburaya S, Tatsukami Y, Kuroda K, Ueda M (2016) Putative
1119 Alginate Assimilation Process of the Marine Bacterium *Saccharophagus degradans*
1120 2-40 Based on Quantitative Proteomic Analysis. Mar Biotechnol 18:15–23.
1121 https://doi.org/10.1007/s10126-015-9667-3
1122 Taylor II LE, Knott BC, Baker JO, Alahuhta PM, Hobdey SE, Linger JG, Lunin V V,
1123 Amore A, Subramanian V, Podkaminer K, Xu Q, VanderWall TA, Schuster LA,
1124 Chaudhari YB, Adney WS, Crowley MF, Himmel ME, Decker SR, Beckham GT
1125 (2018) Engineering enhanced cellobiohydrolase activity. Nat Commun 9:1–10.
1126 https://doi.org/10.1038/s41467-018-03501-8
1127 Thapa SP, Ii EWD, Lyu Q, Weisberg AJ, Stevens DM, Clarke CR, Coaker G, Chang JH
1128 (2019) The Evolution, Ecology, and Mechanisms of Infection by Gram-Positive,
1129 Plant-Associated Bacteria. Annu Rev Phytopathol 57:341–65
1130 Vanfossen AL, Verhaart MRA, Kengen SMW, Kelly RM (2009) Carbohydrate Utilization

Patterns for the Extremely Thermophilic Bacterium *Caldicellulosiruptor saccharolyticus* Reveal Broad Growth Substrate Preferences. Appl Environ Microbiol 75:7718–7724. https://doi.org/10.1128/AEM.01959-09

Von Freiesleben P, Spodsberg N, Stenbæk A, Stålbrand H, Krogh KBRM, Meyer AS (2018) Boosting of enzymatic softwood saccharification by fungal GH5 and GH26 endomannanases. Biotechnol Biofuels 11:1–15. https://doi.org/10.1186/s13068-018-1184-y

Walker JE, Lanahan AA, Zheng T, Toruno C, Lynd LR, Cameron JC, Olson DG, Eckert CA (2020) Development of both type I–B and type II CRISPR/Cas genome editing systems in the cellulolytic bacterium *Clostridium thermocellum*. Metab Eng Commun 10:e00116. https://doi.org/10.1016/j.mec.2019.e00116

Wang X, Zhang W, Zhou H, Chen G, Liu W (2019a) An Extracytoplasmic Function Sigma Factor Controls Cellulose Membrane Protein in *Cytophaga hutchinsonii*. Appl Environ Microbiol 85:e02606-18

Wang Y, Horlamus F, Henkel M, Kovacic F, Schläfle S, Hausmann R, Wittgens A, Rosenau F (2019b) Growth of engineered *Pseudomonas putida* KT2440 on glucose, xylose, and arabinose: Hemicellulose hydrolysates and their major sugars as sustainable carbon sources. Glob Chang Biol Bioenergy 11:249–259. https://doi.org/10.1111/gcbb.12590

Williams-Rhaesa AM, Awuku NK, Lipscomb GL, Poole FL, Rubinstein GM, Conway JM, Kelly RM, Adams MW (2018) Native xylose-inducible promoter expands the genetic tools for the biomass-degrading, extremely thermophilic bacterium *Caldicellulosiruptor bescii*. Extremophiles 22:629–638.

https://doi.org/10.1007/s00792-018-1023-x

Wilson CM, Klingeman DM, Schlachter C, Syed MH, Wu C, Guss AM, Brown SD (2017) LacI Transcriptional Regulatory Networks in *Clostridium thermocellum* DSM1313. Appl Environ Microbiol 83:e02751-16

Wu X, Bei S, Zhou X, Luo Y, He Z, Song C, Yuan H, Pivato B (2023) Metagenomic insights into genetic factors driving bacterial niche differentiation between bulk and rhizosphere soils. Sci Total Environ 891:164221

Xiong W, Reyes LH, Michener WE, Maness P-C, Chou KJ (2018) Engineering cellulolytic bacterium *Clostridium thermocellum* to co-ferment cellulose- and hemicellulose-derived sugars simultaneously. Biotechnol Bioeng 115:1755–1763. https://doi.org/10.1002/bit.26590

Zhang B, Gao Y, Zhang L, Zhou Y (2020) The plant cell wall: Biosynthesis, construction, and functions. J Integr Plant Biol 63:251–272. https://doi.org/10.1111/jipb.13055

Ziles Domingues S, Timmers LFSM, Eichelberger Granada C (2022) Cellulase production by bacteria is a strain-specific characteristic with a high biotechnological potential. A review of cellulosome of highly studied strains. Cellulose 29:8065–8083. https://doi.org/10.1007/s10570-022-04790-5

Figure 1. Common regulatory systems for Carbohydrate-Active Enzyme (CAZyme) encoding genes in Gram-positive and Gram-negative bacteria. (**A)** Hybrid two component system in *Bacteroides thetaiotaomicron*. Upon sensing of arabinoxylan from the transmembrane domain, the intracellular histidine kinase (HK) phosphorylates the associated response regulator (RR) which recruits RNA polymerase for gene transcription. (**B)** ECF-σ/anti-σ factor system in *Bacteroides xylanisolvens*. Binding of arabinoxylan to the carbohydrate domain of the transmembrane ECF protein releases the intracellular σ factor from the membrane-attached anti-σ factor which aids RNA polymerase in gene transcription. (**C)** Carbon catabolite repression in Gram-negative *Escherichia coli*. In the absence of glucose, phosphorylated EII_A accumulates and activates adenylate cyclase (AC) via phosphorylation, which generates high cAMP levels. The cAMP subsequently binds to the cAMP receptor protein (CRP) and initiates transcription of hemicellulase-encoding genes. (**D)** Carbon catabolite repression in Gram-positive *Bacillus subtilis*. In the absence of glucose, fructose 1,6-biphosphate is not generated because glycolysis does not occur. Without fructose 1,6-biphosphate, histidine protein (HPr) does not get phosphorylated and therefore cannot dimerize with the carbon catabolite control protein (CcpA). Without this dimerization, the coupled protein cannot inhibit transcription. For all panels, phosphate is shown as a gold circle with a ‘P’, arabinoxylan is shown with orange stars for xylose and the green stars for arabinose, fructose is shown as a green pentagon. *Model generated with BioRender.com*.

Figure 2. Differences in up-regulation of CAZyme-encoding genes from selected Gram-positive and Gram-negative bacteria when grown using hemicelluloses. (A) CAZyme-encoding gene expression of Gram-positive *Roseburia intestinalis* and Gram-negative *Cellvibrio japonicus* on glucomannan. (**B)** CAZyme-encoding gene expression response of Gram-positive *Caldicellulosiruptor bescii* and Gram-negative *Roseithermus sacchariphilus* on xylan. **(C)** CAZyme-encoding gene expression response of Gram-positive *Caldicellulosiruptor saccharolyticus* and Gram-negative *Bacteroides xylanisolvens* on pectin.

Table 1. Current limitations of select bacterial bioprocessors and suggested research approaches.

	Bacterium	Current limitations	Suggested approach
Gram-positive	*Clostridium thermocellum*	• Requires synthetic biology to utilize non-cellulose derived sugars • Engineered fermentation pathways for plant sugars repressed by plant oligosaccharides	Improve genetic tools to control regulation of heterologously expressed genes
	Caldicellulosiruptor bescii	• Low expression and degradative efficiency of heterologously expressed CAZyme-encoding genes	Improve transcriptional control over heterologously expressed genes
Gram-negative	*Cellvibrio japonicus*	• No high-yielding, stable plasmid system for gene expression • Does not produce any current high-value metabolite in abundance	Develop a stably replicating plasmid for gene expression
	Saccharophagus degradans	• Poor genetic system • Cannot natively ferment sugars to fuels and/or renewable chemicals	Develop genetic tools to engineer a commodity product-producing strain