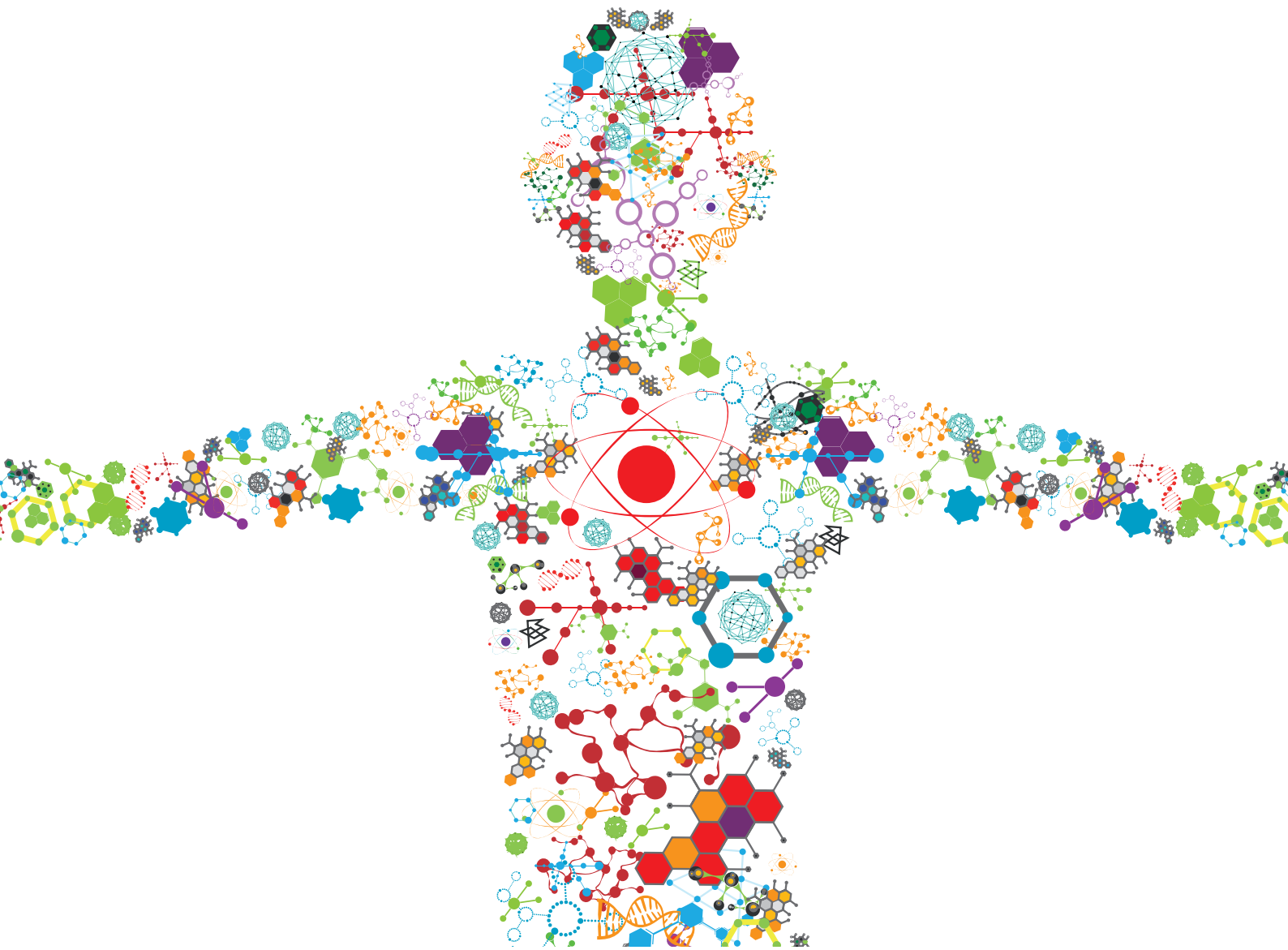
CAZYMES IN BIOREFINERY: FROM GENES TO APPLICATION

EDITED BY: Fabiano Jares Contesini, Rasmus John Normand Frandsen and André Damasio

PUBLISHED IN: Frontiers in Bioengineering and Biotechnology



frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88966-596-9
DOI 10.3389/978-2-88966-596-9

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

CAZYMES IN BIOREFINERY: FROM GENES TO APPLICATION

Topic Editors:
Fabiano Jares Contesini, Technical University of Denmark, Denmark
Rasmus John Normand Frandsen, Technical University of Denmark, Denmark
André Damasio, State University of Campinas, Brazil

Citation: Contesini, F. J., Frandsen, R. J. N., Damasio, A., eds. (2021). CAZymes in Biorefinery: From Genes to Application. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88966-596-9

Table of Contents

frontiers in Bioengineering and Biotechnology

EDITORIAL
published: 10 February 2021
doi: 10.3389/fbioe.2021.622817

Editorial: CAZymes in Biorefinery: From Genes to Application

Fabiano Jares Contesini[1]*, *Rasmus John Normand Frandsen*[1] and *André Damasio*[2]

[1] Synthetic Biology Section, Department of Biotechnology and Biomedicine, Technical University of Denmark, Kongens Lyngby, Denmark, [2] Department of Biochemistry and Tissue Biology, Institute of Biology, University of Campinas, Campinas, Brazil

Keywords: carbohydrate-active enzymes, biorefineries, transcriptomics, proteomics, feruloyl esterase, pectin lyase, carbohydrate esterase, carbohydrate-binding modules

Editorial on the Research Topic

Editorial: CAZymes in Biorefinery: From Genes to Application

OPEN ACCESS

Edited by:
Eduardo Jacob-Lopes, Federal University of Santa Maria, Brazil

Reviewed by:
Giovani Leone Zabot, Federal University of Santa Maria, Brazil
Anna Rafaela Cavalcante Braga, Federal University of São Paulo, Brazil
Luciane Maria Colla, The University of Passo Fundo, Brazil

***Correspondence:**
Fabiano Jares Contesini
fajacon@dtu.dk;
fabiano.contesini@gmail.com

Specialty section:
This article was submitted to Bioprocess Engineering, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *29 October 2020*
Accepted: *25 January 2021*
Published: *10 February 2021*

Citation:
Contesini FJ, Frandsen RJN and Damasio A (2021) Editorial: CAZymes in Biorefinery: From Genes to Application. Front. Bioeng. Biotechnol. 9:622817. doi: 10.3389/fbioe.2021.622817

Carbohydrate-Active EnZymes (CAZymes) encompasses all enzymes involved in the modification, degradation, or biosynthesis of carbohydrates and their derivatives. Especially the CAZymes acting on glycosidic bonds have proven to be crucial for the significant biotechnological advances within sectors that include bioenergy and biobased (food/feed, materials, and chemicals) industries. The concept of CAZymes and their organization into families, based on similar structurally related catalytic or functional domains, was established in the late 1990'ties, and in 1999 Lombard et al. (2014) launched the CAZy database (www.cazy.org). The CAZy database, and associated bioinformatics tools, organize all known CAZymes into the following classes; glycoside hydrolases (GHs), glycosyl transferases (GTs), polysaccharide lyases (PLs), carbohydrate esterases (CEs), and auxiliary activities (AAs) (Levasseur et al., 2013; Lombard et al., 2014).

Biorefinery has received increasing relevance in the last decades, fueled by the significant drawbacks of unsustainable fossil fuel-based production and its negative effect on the global climate (Junqueira et al., 2017; Islam et al., 2020). The biorefinery concept entails the conversion (refinement) of plant biomass into relevant products or energy (Fernandes et al., 2017; Özdenkçi et al., 2017; Chuaboon et al., 2019). The biorefinery process is highly dependent on CAZymes for the complete deconstruction of plant biomass or its transformation to high added-value compounds. Especially, CAZymes capable of degrading polysaccharide fraction of plant biomass into simple sugars have proven significant, as the resulting monomers can either be used a carbon source for fermentation-based production processes or be biotransformed into bioactive compounds, such as bioactive oligosaccharides like alginate oligosaccharides (Falkeborg et al., 2014; Rakotoarivonina et al., 2016; Oliveira et al., 2019). The non-carbohydrate fraction of plant biomass, like lignin, can also be converted into relevant products, such as ferulic acid and vanillin, by lignin-modifying enzymes (Sainsbury et al., 2013; Tian et al., 2017; Fetherolf et al., 2020).

Further advancement of the biorefinery field will depend on the prospection for new and more efficient enzymes to expand the available enzymatic toolbox. Secondly, the development of new enzymatic cocktails for efficient degradation of more complex plant biomass with diverse composition in terms of hemicellulose, pectin, and lignin content. Third, by gaining a system-level understanding of the used cell factory to allow for optimization of their functionality by genetic engineering, to increase the range of compounds that can be processed in the biorefinery context.

Although there are many papers on CAZymes in the biorefinery context (Chettri et al., 2020; Li et al., 2020; Meng et al., 2020), in this editorial, we focus on the works published in this research topic. We would like to thank all the authors that have contributed to this special feature series of articles on CAZymes in Biorefinery. The issue include six original articles and one review article,

which combined, covers different aspects of the role of CAZymes in the context of biorefineries.

Microbial genome sequencing projects routinely reveals a tremendous diversity of CAZyme encoding genes, many of which have unknown functions that potentially can be utilized in biorefineries. One way of gaining insight into the enzymatic arsenal used an organism for plant biomass deconstruction is via transcriptomics analysis of the organism grown on the given substrate. Corrêa et al. used a similar approach to analyze the lignocellulolytic enzyme secretion by *Aspergillus terreus* cultivated in sugarcane bagasse and soybean hulls. An analysis that reveals different sets of responsible genes, encoding transcription factors, transporters and enzymes, with CAZymes predominant among the latter group.

Secretomics analysis (exoproteome) is another approach that can provide valuable insights into the versatile CAZyme arsenal deployed by microorganisms during biomass degradation. In the study by Grieco et al., the secretome analysis of *Myceliophthora thermophila* cultivated on pretreated sugarcane bagasse revealed the presence of CAZymes belonging to the GH and AA families. Further biochemical analysis of two of the identified LPMOs (lytic polysaccharide monooxygenases) showed that they possessed different temperature optima and regioselectivity. The addition of both enzymes to a commercial *Trichoderma reesei* enzyme cocktail was found to boost plant biomass saccharification. Machado et al. studied the exoproteome of two white-rot fungi, *Phanerochaete chrysosporium* and *Trametes versicolor,* cultured in microcrystalline cellulose (Avicel). The most predominant enzymes in both secretomes corresponded to cellobiohydrolase I (CBHI). The enzymatic cocktails produced by both fungi were further compared to commercial lignocellulolytic cocktails and provided an alternative for enzymatic cocktail formulations.

Three original papers focus on specific CAZyme families. Two articles report on carbohydrate esterases (CE) and feruloyl esterases (FEA), which are vital accessory enzymes for hemicellulose deconstruction. Li et al. heterologously produced and characterized four new fungal enzymes belonging to the CE family, three showing acetyl xylan esterase activity and one presenting both feruloyl esterase and acetyl xylan esterase activities. The enzymes displayed promising properties, including high pH stability, thereby showing their potential for biotechnological applications. Underlin et al. characterized 14 feruloyl esterases from different subfamilies using synthetic and plant cell wall-derived substrates. The study revealed unique enzymatic profiles and diverse applicability of the various feruloyl esterases in the biorefinery context. On the other hand, Zeuner et al. report on the activity of four different pectin lyases from Aspergilli on different substrates and finds that the enzymes only displayed subtle differences in activity and product formation profiles. The highest reaction rate was found on apple pectin, while the lowest efficiency was observed for sugar beet pectin. A finding that are of high relevance for the biotechnological industries that utilize pectin lyases for food and biorefinery processes. Finally, a very interesting review concerning carbohydrate-binding modules (CBMs) and CAZymes were prepared by Sidar et al. CBMs are found in many CAZymes and are known to increase the proximity of the enzyme to its substrate, especially for insoluble ones. The review emphasizes cellulases and amylases produced by the filamentous microorganisms from the genera of *Streptomyces* and *Aspergillus* known as efficient secretors of polysaccharide degrading enzymes.

Collectively, the included articles provide a broad introduction to the major experimental strategies currently utilized to further the biorefinery field via a better understanding of CAZymes.

AUTHOR CONTRIBUTIONS

FC designed and wrote the Editorial with contributions from AD and RF. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the São Paulo Research Foundation (FAPESP 2015/50612-8; 2017/22669-0), the National Council for Scientific and Technological Development (CNPq 304816/2017-5), and the Danish National Research Foundation (DNRF137) for the Center for Microbial Secondary Metabolites.

ACKNOWLEDGMENTS

We thank the authors and reviewers for their valuable contributions to this research topic.

REFERENCES

Chettri, D., Verma, A. K., and Verma, A. K. (2020). Innovations in CAZyme gene diversity and its modification for biorefinery applications. *Biotechnol. Rep.* 28:e00525. doi: 10.1016/j.btre.2020.e00525

Chuaboon, L., Wongnate, T., Punthong, P., Kiattisewee, C., Lawan, N., Hsu, C. Y., et al. (2019). One-Pot bioconversion of l-arabinose to l-ribulose in an enzymatic Cascade. *Angew. Chem. Int. Ed.* 58, 2428–2432. doi: 10.1002/anie.201814219

Falkeborg, M., Cheong, L. Z., Gianfico, C., Sztukiel, K. M., Kristensen, K., Glasius, M., et al. (2014). Alginate oligosaccharides: enzymatic preparation and antioxidant property evaluation. *Food Chem.* 164, 185–194. doi: 10.1016/j.foodchem.2014.05.053

Fernandes, B. S., Vieira, J. P. F., Contesini, F. J., Mantelatto, P. E., Zaiat, M., Pradella, J. G., et al. (2017). High value added lipids produced by microorganisms: a potential use of sugarcane vinasse. *Crit. Rev. Biotechnol.* 37, 1048–1061. doi: 10.1080/07388551.2017.1304356

Fetherolf, M. M., Levy-Booth, D. J., Navas, L. E., Liu, J., Grigg, J. C., Wilson, A., et al. (2020). Characterization of alkylguaiacol-degrading cytochromes P450 for the biocatalytic valorization of lignin. *Proc. Natl. Acad. Sic. U.S.A.* 117 25771–25778. doi: 10.1073/pnas.1916349117

Islam, M. K., Wang, H., Rehman, S., Dong, C., Hsu, H. Y., Lin, C. S. K., et al. (2020). Sustainability metrics of pretreatment processes in a waste derived lignocellulosic biomass biorefinery. *Bioresour. Technol.* 298:122558. doi: 10.1016/j.biortech.2019.122558

Junqueira, T. L., Chagas, M. F., Gouveia, V. L. R., Rezende, M. C. A. F., Watanabe, M. D. B., Jesus, C. D. F., et al. (2017). Techno-economic analysis and climate change impacts of sugarcane biorefineries considering different time horizons. *Biotechnol. Biofuels*, 10:50. doi: 10.1186/s13068-017-0722-3

Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M., and Henrissat, B. (2013). Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol. Biofuels* 6:41. doi: 10.1186/1754-6834-6-41

Li, X., Han, C., Li, W., Chen, G., and Wang, L. (2020). Insights into the cellulose degradation mechanism of the thermophilic fungus *Chaetomium thermophilum* based on integrated functional omics. *Biotechnol. Biofuels* 13:143. doi: 10.1186/s13068-020-01783-z

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, 490–495. doi: 10.1093/nar/gkt1178

Meng, X., Ma, L., Li, T., Zhu, H., Guo, K., Liu, D., et al. (2020). The functioning of a novel protein, swollenin, in promoting the lignocellulose degradation capacity of *Trichoderma guizhouense* NJAU4742 from a proteomic perspective. *Bioresour. Technol.* 317:123992. doi: 10.1016/j.biortech.2020.123992

Oliveira, D. M., Mota, T. R., Oliva, B., Segato, F., Marchiosi, R., Ferrarese-Filho, O., et al. (2019). Feruloyl esterases: biocatalysts to overcome biomass recalcitrance and for the production of bioactive compounds. *Bioresour. Technol.* 278, 408–423. doi: 10.1016/j.biortech.2019.01.064

Özdenkçi, K., De Blasio, C., Muddassar, H. R., Melin, K., Oinas, P., Koskinen, J., et al. (2017). A novel biorefinery integration concept for lignocellulosic biomass. *Energy Convers. Manage.* 149, 974–987. doi: 10.1016/j.enconman.2017.04.034

Rakotoarivonina, H., Revol, P. V., Aubry, N., and Rémond, C. (2016). The use of thermostable bacterial hemicellulases improves the conversion of lignocellulosic biomass to valuable molecules. *Appl. Microbiol. Biotechnol.* 100, 7577–7590. doi: 10.1007/s00253-016-7562-0

Sainsbury, P. D., Hardiman, E. M., Ahmad, M., Otani, H., Seghezzi, N., Eltis, L. D., et al. (2013). Breaking down lignin to high-value chemicals: the conversion of lignocellulose to vanillin in a gene deletion mutant of rhodococcus jostii RHA1. *ACS Chem. Biol.* 8, 2151–2156. doi: 10.1021/cb400505a

Tian, D., Hu, J., Chandra, R. P., Saddler, J. N., and Lu, C. (2017). Valorizing recalcitrant cellulolytic enzyme lignin via lignin nanoparticles fabrication in an integrated biorefinery. *ACS Sust. Chem. Eng.* 5, 2702–2710. doi: 10.1021/acssuschemeng.6b03043

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 23 April 2020
doi: 10.3389/fbioe.2020.00332

Feruloyl Esterases for Biorefineries: Subfamily Classified Specificity for Natural Substrates

Emilie N. Underlin[1,2†], Matthias Frommhagen[1†], Adiphol Dilokpimol[3], Gijs van Erven[1], Ronald P. de Vries[3] and Mirjam A. Kabel[1]*

[1] Laboratory of Food Chemistry, Wageningen University & Research, Wageningen, Netherlands, [2] Department of Chemistry, Technical University of Denmark, Lyngby, Denmark, [3] Fungal Physiology, Westerdijk Fungal Biodiversity Institute and Fungal Molecular Physiology, Utrecht University, Utrecht, Netherlands

Feruloyl esterases (FAEs) have an important role in the enzymatic conversion of lignocellulosic biomass by decoupling plant cell wall polysaccharides and lignin. Moreover, FAEs release anti-oxidative hydroxycinnamic acids (HCAs) from biomass. As a plethora of FAE candidates were found in fungal genomes, FAE classification related to substrate specificity is an indispensability for selection of most suitable candidates. Hence, linking distinct substrate specificities to a FAE classification, such as the recently classified FAE subfamilies (SF), is a promising approach to improve the application of these enzymes for a variety of industrial applications. In total, 14 FAEs that are classified members of SF1, 5, 6, 7, 9, and 13 were tested in this research. All FAEs were investigated for their activity toward a variety of substrates: synthetic model substrates, plant cell wall-derived substrates, including lignin, and natural substrates. Released HCAs were determined using reverse phase-ultra high performance liquid chromatography coupled to UV detection and mass spectrometry. Based on this study, FAEs of SF5 and SF7 showed the highest release of FA, *p*CA, and diFAs over the range of substrates, while FAEs of SF6 were comparable but less pronounced for diFAs release. These results suggest that SF5 and SF7 FAEs are promising enzymes for biorefinery applications, like the production of biofuels, where a complete degradation of the plant cell wall is desired. In contrast, SF6 FAEs might be of interest for industrial applications that require a high release of only FA and *p*CA, which are needed as precursors for the production of biochemicals. In contrast, FAEs of SF1, 9 and 13 showed an overall low release of HCAs from plant cell wall-derived and natural substrates. The obtained results substantiate the previous SF classification as a useful tool to predict the substrate specificity of FAEs, which eases the selection of FAE candidates for industrial applications.

Keywords: biomass, corn stover, feruloyl esterases, hydroxycinnamic acids, lignin-carbohydrate complex, pectin, wheat straw, xylan

OPEN ACCESS

Edited by:
André Damasio, Campinas State University, Brazil

Reviewed by:
Yasser Gaber, Beni-Suef University, Egypt
Noha M. Mesbah, Suez Canal University, Egypt
Marcelo Vizoná Liberato, Brazilian Bioethanol Science and Technology Laboratory – CTBE, Brazil

***Correspondence:**
Mirjam A. Kabel
mirjam.kabel@wur.nl

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Bioprocess Engineering, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *03 December 2019*
Accepted: *25 March 2020*
Published: *23 April 2020*

Citation:
Underlin EN, Frommhagen M, Dilokpimol A, van Erven G, de Vries RP and Kabel MA (2020) Feruloyl Esterases for Biorefineries: Subfamily Classified Specificity for Natural Substrates. Front. Bioeng. Biotechnol. 8:332. doi: 10.3389/fbioe.2020.00332

Abbreviations: Ara, arabinofuranosyl residue; 8-8′-(aryl)-diFA, 8-8′-(aryl)-diferulic acid; BSA, bovine serum albumin, CFoligo, corn fiber oligomers; CS, corn stover; CSlignin, corn stover diFA diferulic acid; diFAs diferulic acids, lignin isolate; 8-*O*-4′-diFA, 8-*O*-4′-diferulic acid; 5-5′-diFA, 5-5′-diferulic acid; 8-5′-diFA, 8-5′-diferulic acid; diFA: diferulic acid, FA, ferulic acid; FAE, feruloyl esterase; 8-8′-(furan)-diFA, 8-8′-(tetrahydrofuran)-dehydroferulic acid; G, guaiacyl subunit; H, *p*-hydroxyphenyl subunit; HCA, hydroxycinnamic acid; MF, methyl ferulate; MpC, methyl *p*-coumarate; MS, mass spectrometry; NI, negative mode; *p*CA, *p*-coumaric acid; PCW, plant cell wall; S, syringyl subunit; SBP, sugar beet pectin; SF, subfamily; triFA, triferulic acid; triFAs, triferulic acids; WAX-i, wheat arabinoxylans insoluble; WS, wheat straw.

INTRODUCTION

Plant polysaccharides within lignocellulosic biomass are considered a sustainable and green resource for the production of biobased chemicals and fuels after depolymerization (Mood et al., 2013; Loqué et al., 2015). The major component of lignocellulosic biomass consists of the plant cell wall (PCW) – a network of cellulose, hemicellulose and lignin. This complex structure hinders the depolymerization of the PCW by, for example, fungal enzymes (Kabel et al., 2007; Pauly and Keegstra, 2008). Hence, enzymes that are able to increase the PCW accessibility, such as feruloyl esterases (FAEs), are widely applied in biorefinery processes for the biofuel, food, feed, pulp, and paper industries (Haruhiko et al., 2007; Shuichi et al., 2009; West et al., 2009; Hiroyuki et al., 2010; Nsereko et al., 2010a,b,c; West and William, 2010).

The PCW consists of a primary and a secondary layer, which have different compositions. The primary layer is mainly composed of cellulose, hemicelluloses such as xyloglucan, and various types of pectins (Carpita and Gibeaut, 1993; Hoffman et al., 2005; Carpita and McCann, 2008). The secondary layer is mainly composed of cellulose, which is embedded in a network consisting of hemicellulosic xylan and/or mannan, and the aromatic polymer lignin (Carpita and McCann, 2008; Harris and Stone, 2008). Beyond the complexity of the PCW polymers, the accessibility of enzymes is hindered by various cross-links, linking either two polysaccharide chains with each other or a polysaccharide chain to lignin. One of the major cross-links in the PCW is formed via one or two hydroxycinnamic acid (HCA) units, in particular ferulic acid (FA) (Harris and Stone, 2008).

In xylan-rich PCW, such as in *Gramineae* (i.e., corn fiber, various grasses), the carboxyl group of FA (**Figure 1**) is ester-linked to the *O*-5 position of α-L-arabinofuranosyl substituents of the xylan backbone (de O Buanafina, 2009; Harris and Trethewey, 2010). In pectin-rich PCWs on the other hand, e.g., the roots of various dicots (i.e., sugar beet), the carboxyl group of FA is ester linked to the *O*-6 position of β-D-galactopyranosyl residues in galactan and/or to the *O*-2 and *O*-5 positions of α-L-arabinofuranosyl residues in arabinan (Ralet et al., 1994; Levigne et al., 2004a,b). Furthermore, various diferulic acids (diFAs) and triferulic acids (triFAs) are esterified to PCW polysaccharides, of which 8-8′-(tetrahydrofuran)-diFA (8-8′-(furan)-diFA), 8-5′-diFA, 5-5′-diFA, and 8-*O*-4′-diFA are the most common in the PCW (**Figure 1**; Dilokpimol et al., 2016). A part of these mono-, di-, and tri-ferulic acid compounds can be linked to another polysaccharide chain or is etherified (or esterified) to lignin, which leads to the formation of cross-links (Ralph et al., 1994; de O Buanafina, 2009; Vismeh et al., 2013; de Oliveira et al., 2015; Dilokpimol et al., 2016; Hatfield et al., 2017). Besides FA, another esterified HCA is *p*-coumaric acid (**Figure 1**), which is mostly reported to be γ-esterified to β-*O*-4-linkages within the lignin macromolecule, in particular in the PCW of certain *Gramineae* (Ralph, 2010). Small amounts of FA could also be ester-linked to lignin for which further evidence has yet to be obtained (Regner et al., 2018).

FAEs (EC 3.1.1.73) represent a subclass of the carbohydrate esterase family 1 (CE1) from the CAZy database (CAZy.org), which release HCAs from plant biomass via the hydrolysis of the ester-linkages (Wong, 2006; Topakas et al., 2007; Faulds, 2010; Lombard et al., 2014). This hydrolysis of cross-links between polysaccharides and lignin increases the accessibility of PCW polysaccharides for other enzymes, such as glycosyl hydrolases (GH). The released HCAs are valuable precursors for the production of biochemicals, like antioxidants, that are applied in the cosmetic and pharmaceutical industry (de O Buanafina, 2009; de Oliveira et al., 2015; Dilokpimol et al., 2016).

Genomes among the fungal kingdom comprise a plethora of FAE candidates. As a result, a classification of FAEs, which is preferably coupled to their substrate specificity, is indispensable for the selection of suitable FAE candidates for industrial applications. Several classifications of FAEs have been suggested (Crepin et al., 2004; Benoit et al., 2008; Udatha et al., 2011). The most commonly used is the ABCD classification, which was introduced by Crepin et al. (2004). This classification was based on the substrate specificity of a limited number of FAEs toward the four common methylated synthetic model substrates (methyl ferulate, methyl *p*-coumarate, methyl caffeate, and methyl sinapate) and the ability of FAEs to release diFAs from isolated PCW hemicelluloses (Crepin et al., 2004). Within the past decade, it has become apparent that more and more characterized FAEs could not be categorized in this system, which indicates that the ABCD classification is limited in its ability to reflect the broad substrate specificity and diversity of the FAEs (Dilokpimol et al., 2016). Hence, a new amino acid sequence-based classification system of these enzymes has been designed, which divides (putative) FAEs into 13 subfamilies (SF) and expands the more limited SF classification that previously described seven SFs (Benoit et al., 2008; Dilokpimol et al., 2016).

This SF classification is based on phylogenetic relationships and sequence homologies of over 1000 putative fungal FAEs. Briefly, these FAE amino acid sequences were collected by BLASTP search against over 200 fungal genomes published before 2014 using 20 amino acid sequences (Dilokpimol et al., 2016). The resulting sequences were aligned using Multiple Sequence Comparison by Log-Expectation (MUSCLE). The phylogenetic relationship was analyzed by using the neighbor-joining method. The phylogenetic relationship also showed that FAEs evolved from highly divert esterase families including tannases (SF1-4), acetyl xylan esterases (SF6), lipases (SF7), and choline esterases (SF12-13) (Dilokpimol et al., 2016). In a later study, 20 representative FAEs covering 11 SFs were biochemically characterized using four common synthetic model substrates (Dilokpimol et al., 2018). However, it has not been assessed how FAEs of different SFs hydrolyze ester-linked HCAs from natural plant substrates. Therefore, in this study, we investigated the natural plant substrate specificity of 14 fungal FAEs from different SFs (SF1, 5, 6, 7, 9, and 13) (**Table 1**). Until today, only the five SF5 and three SF6 members have been classified in the CAZY database – as carbohydrate esterase family 1 members (Lombard et al., 2014). PCW-derived and natural substrates from different species were used to mimic the ability of these FAEs to release HCAs (including monomers, dimers and trimers) during biomass degradation in nature and to emphasize the potential

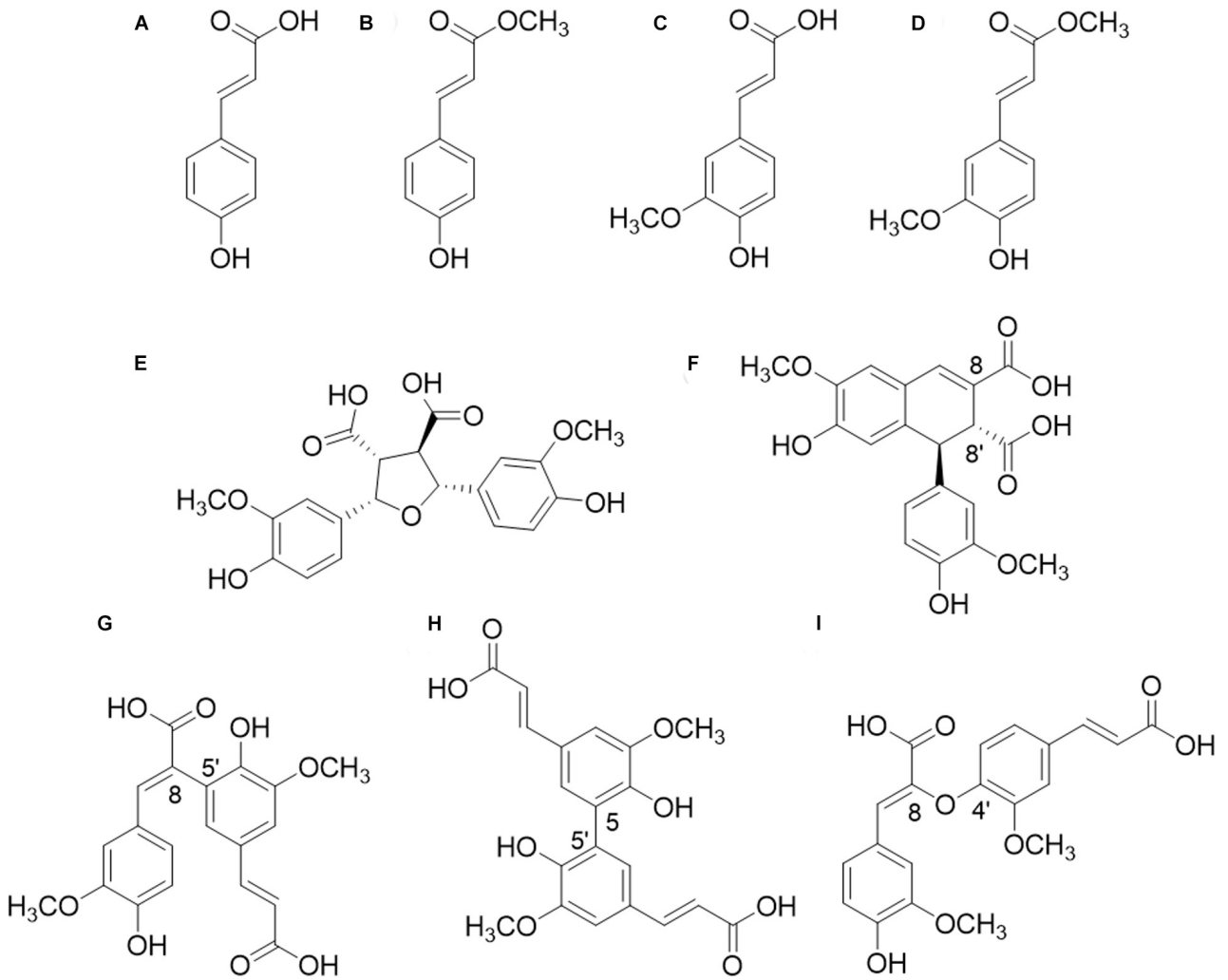


FIGURE 1 | Structures of hydroxycinnamic acids (HCAs). **(A)** p-coumaric acid (pCA); **(B)** methyl p-coumarate; **(C)** ferulic acid (FA); **(D)** methyl ferulate; **(E)** 8-8′-(furan)-diferulic acid (8-8′-diFA); **(F)** 8-8′-(aryl)-diFA; **(G)** 8-5′-diFA; **(H)** 5-5′-diFA; **(I)** 8-O-4′-diFA.

of the 14 FAEs for diverse industrial applications. In detail, the PCW-derived substrates sugar beet pectin (SBP), corn fiber oligosaccharides (CFoligo), insoluble arabinoxylans (WAX-i), lignin isolate from corn stover (CSlignin), and the natural substrates corn stover (CS) and wheat straw (WS) were tested.

Released HCAs were identified and determined by using RP-UHPLC-UV-ESI-MS/MS. The obtained results substantiate that the SF classification is a useful tool to predict the substrate specificity of FAEs, which eases the selection of FAE candidates for industrial applications.

MATERIALS AND METHODS

Substrates

Methyl ferulate, methyl *p*-coumarate, ferulic acid, and *p*-coumaric acid were obtained from Apin Chemicals Ltd. (Abingdon, United Kingdom), Carbosynth (Compton, United Kingdom), Fluka™ (Leicestershire, United Kingdom) and Sigma Aldrich (Darmstadt, Germany), respectively. Sugar beet pectin (SBP, Pectin Betapec RU301) was purchased from Herbstreith & Fox KG (Neuenbürg, Germany). Wheat arabinoxylan insoluble was obtained from (WAX-i, Megazyme, Bray, Ireland) and the corn fiber oligomers (CFoligo)-preparation was obtained as previously described (Appeldoorn et al., 2010). Corn stover (CS) was obtained from DSM (Delft, The Netherlands) and wheat straw (WS) were kindly provided by CNC Grondstoffen B.V. (Milsbeek, The Netherlands). Prior to incubation, both CS and WS were ball milled as previously described (van Erven et al., 2017), giving rise to deconstruction of cellulose crystallinity and making the substrates more susceptible to enzymatic pretreatment (Broxterman et al., 2018). The carbohydrate content and composition of all substrates used was performed as previously described (Englyst and Cummings, 1984). The analysis details are presented in the supporting information (**Supplementary Table S1**).

TABLE 1 | Overview of FAEs employed in this study including fungal origin, accession number, name identification, classification according to the SF and ABCD system.

Fungal species	Accession number	Name	SF classification[†,‡]	ABCD-classification	pH Optimum	CAZy classification[§]	References
Aspergillus niger	CAC83933.1	AnFaeB	1	B	5	Not in CAZy	de Vries et al., 2002
Aspergillus sydowii	Aspsy1_293049	AsFaeF	1	B	6	Not in CAZy	Dilokpimol et al., 2018
Aspergillus niger	XP_001395336.1	AnFaeC	5	C	6	CE1	Dilokpimol et al., 2017
Aspergillus nidulans	AN5267	AnidFAEC	5	C or D	5	CE1	Debeire et al., 2012
Aspergillus sydowii	Aspsy1_154482	AsFaeC	5	C or D	6	CE1	Dilokpimol et al., 2018
Chrysosporium lucknowense C1	JF826027	C1FaeA1	5	A	7	CE1	Kühnel et al., 2012
Chrysosporium lucknowense C1	JF826028	C1FaeA2	5	A	7.5	CE1	Kühnel et al., 2012
Fusarium oxysporum	Fusox1_8990	FoFae2	6	n/a	6	CE1	Dilokpimol et al., 2016
Aspergillus sydowii	Aspsy1_1158585	AsFaeE	6	C or D	6	CE1	Dilokpimol et al., 2018
Chrysosporium lucknowense C1	JF826029	C1FaeB2	6	B	7	CE1	Kühnel et al., 2012
Aspergillus niger	CAA70510.1	AnFaeA	7	A	5	Not in CAZy	de Vries et al., 1997
Aspergillus niger	An15g05280	AnFaeJ	9	n/a	5	Not in CAZy	Dilokpimol et al., 2016
Aspergillus sydowii	Aspsy1_160668	AsFaeI	13	B	6	Not in CAZy	Dilokpimol et al., 2018
Stereum hirsutum	Stehi1_73641	ShFae1	13	n/a	n/a	Not in CAZy	Dilokpimol et al., 2016

[†]*SF classification as according to the phylogenetic tree made in 2016 (Dilokpimol et al., 2016).* [‡]*Based on reactivity toward the four most common methyl esters according to Crepin et al. (2004).* [§]*Based on CAZy classification according to Lombard et al. (2014). n/a, not applicable.*

Washing of CS and WS

Around 10 mg of ball mill CS and WS was washed with 1 mL water in a head-over-tail shaker for 15 min at 20°C. Afterward, all samples were centrifuged at (12,000 *g*, 15 min, 10°C). The supernatant was removed and the obtained pellet was washed another two times under the same conditions. The washed pellet was freeze dried and the weight loss (water soluble CS and WS) was calculated. Experiments were performed in triplicate. Results are given in the Discussion section.

Preparation and Characterization of the Corn Stover Lignin Isolate

Corn stover lignin isolate (CSlignin) was obtained as previously described for wheat straw lignin and analyzed for residual carbohydrate content. Briefly, lignin was isolated from extractive-free planetary ball-milled corn stover by aqueous dioxane, freeze-dried and subsequently enzymatically purified (van Erven et al., 2017). Structural analysis of the CSlignin was performed by using 2D heteronuclear single quantum coherence (HSQC) NMR; (van Erven et al., 2018) see supplement for details and **Supplementary Tables S1, S2** for compositional details.

Enzymes

The expression, purification and characterization of C1FaeA1 (JF826027, subfamily 5 (SF5, Dilokpimol et al., 2016), C1FaeA2 (JF826028, SF5), and C1FaeB2 (JF826029, SF6) from *Chrysosporium lucknowense* C1 have been described previously (Kühnel et al., 2012). The expression and production of the FAEs in *Pichia pastoris* X-33 from *Aspergillus niger*: AnFaeB (CAC83933.1, SF1), AnFaeC (XP_001395336.1, SF5), AnFaeA (CAA70510.1, SF7), and AnFaeJ (An15g05280, SF9), from *Aspergillus sydowii*: AsFaeF (Aspsy1_293049, SF1), AsFaeC (Aspsy1_154482, SF5), AsFaeE (Aspsy1_1158585, SF6), and AsFaeI (Aspsy1_160668, SF13), from *Aspergillus nidulans*: AnidFAEC (AN5267, SF5), from *Fusarium oxysporum*: FoFae2 (Fusox1_8990, SF6), and *Stereum hirsutum*: ShFae1 (Stehi1_73641, SF13) has previously been described (Dilokpimol et al., 2018) (**Table 1**). The *P. pastoris* X-33 harboring the corresponding genes were grown according to Dilokpimol et al. (2017). The induction was continued for 96 h at 22°C with 0.5% (v/v) methanol supplement every 24 h. Culture supernatants were harvested (4000 × g, 4°C, 20 min), filtered (0.22 μm; Merck Millipore, Darmstadt, Germany) and stored at −20°C prior to further analysis.

Enzyme Filtering and Measurement of Protein Content

Approximately 7.5 mL of the culture filtrates containing the individual FAE and the culture filtrate without FAE (broth) were concentrated to approx. 2 mL using ultrafiltration (Amicon Ultra, molecular mass cut-off of 3 kDa, Merck Millipore, Cork, Ireland). The filtrate was removed, and the concentrate of the culture filtrates containing the individual FAE and the culture filtrate without FAE diluted with a sodium acetate buffer (50 mM, pH 5.8) to 7 mL and concentrated again using ultrafiltration. This washing procedure was performed twice. The protein content of the enzyme-containing concentrates and the broth from *P. pastoris* without FAE insertion (negative control, coded as "broth") were measured using a BCA Protein Assay Kit (Thermo Scientific, Rockford, IL, United States). The results from the broth was compared to the results obtained from a BSA Protein Assay Kit – Reducing agent compatibility (Thermo Scientific, Rockford, IL, United States), where similar protein concentration was measured as described above. For both assays bovine serum albumin (BSA) was used as a standard. After protein content determination, all enzymes were diluted to a final protein concentration of 1 mg/mL in a sodium acetate buffer (50 mM, pH 5.8).

Incubation of FAEs With Synthetic Model Substrates

Activity assays toward the model substrates methyl ferulate and methyl *p*-coumarate were performed in 200 μL reaction mixtures. 2 mM substrate was dissolved in a sodium acetate buffer (50 mM, pH 5.8) (175 μL) and 25 μg of FAE-containing filtrate (1 mg/mL) or the broth (1 mg/mL) was added. In contrast, 1 μg of the purified FAEs from *Chrysosporium lucknowense* C1 (C1FaeA1, C1FaeA2, and C1FaeB2) was used for the activity assays. The reaction mixtures were incubated at 37°C for 2 and 19 h. All reactions were performed in duplicate. At 2 and 19 h, 50 μL was taken from the reaction mixture and diluted 10 times with MilliQ. All reactions were stopped by incubating the samples at 99°C for 2 min. Samples were stored in the freezer and thawed shortly before subjection to RP-UHPLC-UV-ESI-MS/MS.

Determination of Free and Bound HCAs in PCW-Derived and Natural Substrates

The maximum release (free and bound content) of HCAs in the substrates was estimated based on the amounts released in 0.5 M KOH (saponification). In total, 5 mg of substrate (or 1 mg for CSlignin), 500 μL of acetate buffer (or 100 μL for CSlignin; 50 mM, pH 5.8) and 500 μL (or 100 μL for CSlignin) of 0.5 M KOH were mixed. Incubation were prepared in triplicate. The free amount of HCAs in the substrates was determined by mixing 5 mg of substrate (or 1 mg for CSlignin), and 500 μL of acetate buffer (or 100 μL for CSlignin; 50 mM, pH 5.8). This preparation was prepared in duplicate. All samples were incubated for 19 h in the dark at 37°C, under head-over-tail mixing. After incubation, saponified samples were acidified (pH ∼3) by addition of 6 M HCl. All samples were centrifuged (12,000 *g*, 10 min, 4°C), and appropriately (between 10 and 50 times) diluted before RP-UHPLC-UV-ESI-MS/MS analysis.

Enzyme Activity Assays With PCW-Derived and Natural Substrates

Enzyme activity assays toward PCW-derived and natural substrates were performed for all 14 FAEs and the broth. The insoluble substrates (SBP, CS, and WS) were weighed in the corresponding amounts. CFoligo was dissolved in acetate buffer (50 mM, pH 5.8) to a stock solution of 10 mg/mL. CSlignin was dissolved in EtOH/$CHCl_3$ 50:50 to 1 mg/mL and after distribution of sample-material the solvent was evaporated under a stream of nitrogen. A lower amount of CSlignin was used due to the limited substrate availability. The reactions were performed in the presence of 5 mg of substrate (or 1 mg of CSlignin) in 500 μL (or 100 μL for CSlignin) of a sodium acetate buffer (50 mM, pH 5.8). Enzyme loading was 25 μg for FAE-containing filtrates (1 mg/mL), and 2.5 μg enzyme for purified FAEs from *C. lucknowense* C1 (C1FaeA1, C1FaeA2, and C1FaeB2). All reactions were performed in duplicate. The samples were incubated for 19 h in the dark at 37°C, with head-over-tail mixing. Enzymes were inactivated at 99°C for 5 min. All samples were centrifuged (12,000 *g*, 10 min, 4°C), and appropriately diluted before RP-UHPLC-UV-ESI-MS/MS analysis. The released amount of HCAs by the enzymes was determined as the total sum of HCAs subtracted by the free amount (section Enzymes) present. Percentages are presented based on total esterified contents (section Enzymes).

RP-UHPLC-UV-ESI-MS/MS Analysis: Model Substrates Methyl Ferulate and Methyl *p*-Coumarate Incubated With FAEs

All samples were analyzed on an Accela reversed phase ultra-high performance liquid chromatography (RP-UHPLC) system, equipped with a pump, degasser, autosampler, and photodiode array (PDA) detector (Thermo Scientific, San Jose, CA, United States). Samples (5 μL) were injected onto an Acquity UPLC BEH C18 column (150 × 2.1 mm, particle size 1.7 μm) (Waters, Milford, MA, United States). To ensure the stability of the released compounds, the temperature of the autosampler was kept at 4°C during the analysis. The flow rate was 400 μL/min at 45°C. The binary mobile phases consisted of (A) water + 0.1% formic acid and (B) acetonitrile + 0.1% formic acid. The elution profile was as follows: Isocratic on 5% B; 0.0–1.5 min, B linearly from 5 to 60%; 1.5–20.0 min, B linearly from 60 to 100% B; 20.0–20.1 min, isocratic on 100% B; 20.1–25.0 min, B linearly from 100 to 5% B; 25.0–26.0 min, isocratic on 5% B; 26.0–31.0 min. UV spectra for methyl *p*-coumarate and methyl ferulate were recorded at 310 and 320 nm, respectively. No MS data was acquired for the incubation of the model substrates with the FAEs. The decrease in substrate concentration representing activity was determined from a standard curve of the substrates (0.625–50 μg/mL). Data were processed using Xcalibur 2.2 (Thermo Fisher Scientific).

RP-UHPLC-UV-ESI-MS/MS Analysis: HCA Content and Released Amounts From PCW-Derived and Natural Substrates

Samples were analyzed by using RP-UHPLC-UV as described above with a modified elution profile combined with electrospray ionization – ion trap mass spectrometry (ESI-ITMS) detection. To ensure the stability of the released compounds, the temperature of the autosampler was kept at 4°C during the analysis. The binary mobile phases consisted of (A) water + 0.1% formic acid and (B) acetonitrile + 0.1% formic acid. The elution profile was as follows: The first 2 min isocratic on 5% B; 0.0–2.0 min, B linearly from 5 to 40%; 2.0–13.0 min, B linearly from 40 to 100% B; 13.0–13.1 min, isocratic on 100% B; 13.1–18.0 min, B linearly from 100 to 5% B; 18.0–18.1 min, isocratic on 5% B; 18.1–23.0 min. The flow rate was 0.400 mL/min. UV spectra of FA and *p*CA were recorded at 310 and 285 nm, respectively. MS data was acquired by using an LTQ-Velos Pro mass spectrometer (Thermo Fisher Scientific) equipped with a heated ESI probe. Nitrogen was used as sheath gas and auxiliary gas. Data was collected over a m/z range of 120–1500 in negative (NI) mode. Data dependent MS^2 analysis was performed using collision-induced dissociation with a normalized collision energy of 35%. The ion transfer tube temperature was 300°C, source heater temperature was 250°C and the source voltage was 3.5 kV. Data were processed using Xcalibur 2.2 (Thermo Fisher Scientific). A linear correlation between the MS signal and the concentration of FA and *p*CA was found. Based on the MS molar response

factor obtained for FA, the molar response factors of diFAs and triFAs were estimated, as no standards are commercially available. Our values obtained for amounts of diFAs and triFAs based on the estimated molar response factors are expected to be close to actual absolute amounts, but remain to be confirmed. Nevertheless, the incubation of the substrates with FAEs led to a difference in released amounts of diFAs and triFAs, which allowed us to determine variations in the ability of FAEs to release HCA derivatives. For SBP and WAX-i the UV-signal was used, as low amounts of the analyzed compounds were released. The contents of diFAs and triFAs were calculated relative to FA including mass correction. FA and *p*CA standard curves showed linearity both in UV and in MS (data not shown). The MS standard curves for FA and *p*CA resulted in the R^2-values between 0.9723 and 0.9997 for all experiments.

RESULTS

Specificity of FAEs Toward Synthetic Model Substrates

Usually, the specificity of FAEs is assayed via synthetic substrate conversion. Therefore, we assayed the specificity of our 14 FAEs toward the two synthetic substrates methyl ferulate and methyl *p*-coumarate after 2 and 19 h of incubation (**Figure 2** and **Supplementary Table S3**). These two substrates were chosen from the four most commonly used model substrates (methyl ferulate, methyl *p*-coumarate, methyl caffeate, and methyl sinapate), as FA and *p*CA are the major HCAs present in PCWs (Crepin et al., 2004). Thirteen out of 14 FAEs that were tested for their activity toward methyl ferulate decreased its concentration after an incubation time of 2 h (**Figure 2A**). Only ShFae1 (SF13) did not decrease the methyl ferulate concentration within 2 h. After 19 h, the methyl ferulate concentration decreased further for the other 13 tested FAEs. Among them, the FAEs of SF5 almost completely converted methyl ferulate after 19 h, except AnFaeC (**Figure 2A**). The incubation of methyl *p*-coumarate with FAEs also resulted in a decrease in the methyl *p*-coumarate concentration (**Figure 2B**). In brief, nine FAEs showed activity toward methyl *p*-coumarate after 2 h of incubation. For those FAEs, almost all methyl *p*-coumarate had been hydrolyzed after 19 h. In contrast, the incubation of methyl *p*-coumarate with AnFaeC (SF5), AnFaeA (SF7), AnFaeJ (SF9), and ShFae1 (SF13) did not lead to a decrease in the substrate concentration, even after an incubation time of 19 h. AnFaeC and ShFAE1 showed a higher activity toward methyl *p*-coumarate within the first 2 h compared to the 19 h incubation, although both enzymes showed only a relatively low activity toward this substrate. The incubation of methyl ferulate and methyl *p*-coumarate with the broth from a *P. pastoris* strain, which did not contain any of the selected FAEs, did not alter the methyl ferulate and methyl *p*-coumarate concentration after 2 or 19 h (**Figures 2A,B**).

Identification of HCAs in PCW-Derived and Natural Samples

In total, four compounds were chosen as representatives of different classes of PCW-derived substrates: SBP is a pectin isolate, soluble CFoligo are composed of branched xylo-oligosaccharides, which are highly feruoylated (Appeldoorn et al., 2013), WAX-i is an insoluble xylan (Harris and Trethewey, 2010), and CSlignin is a HCA-rich lignin isolate (**Supplementary Files S5, S6** and **Supplementary Tables S1, S2**). In addition, the natural substrates CS and WS were used as representatives of HCA-rich lignocellulose, mainly composed of xylan, lignin and cellulose (van Dongen et al., 2011; Bakker et al., 2013). The total carbohydrate content and composition of all PCW-derived and natural substrates is shown in **Supplementary Files S5, S6** and **Supplementary Tables S1, S2**.

The presence of various types of esterified HCAs was analyzed for all PCW-derived and natural substrates through saponification. RP-UHPLC-UV-ESI-MS/MS showed the release of FA, *p*CA, eight different diFAs of which three are proposed diFAs (m/z 389, m/z $401_{9.74\ min}$, m/z $401_{10.66\ min}$), and two triFAs under alkaline conditions (0.5 M KOH, **Table 2**). Tentative structures of diFAs were proposed based on MS^2 data (**Table 2** and **Supplementary Figures S2, S3**). In addition, four diFAs (8-8′-aryl-diFA, 8-5′-diFA, 5-5′-diFA, and 8-*O*-4′-diFA) were identified based on retention time using reversed phase separation, which corresponded to previous reported data (**Table 2** and **Supplementary Figures S2C,D,F,H**; Appeldoorn et al., 2010; Kühnel et al., 2012; Vismeh et al., 2013).

We propose the identification of 8-8′-(furan)-diFA (CFoligo, 8.23 min, m/z 404; **Table 2** and **Supplementary Figure S2A**), in addition to the previously reported diFAs (8-8′-aryl-diFA, 8-5′-diFA, 5-5′-diFA, and 8-*O*-4′-diFA). The presence of the 8-*O*-4′-diFA has been reported for PCWs and in particular for CFoligo (Bunzel et al., 2006; Schatz et al., 2006; Appeldoorn et al., 2010). Interestingly, a peak (R_t 8.23 min) in the UV chromatogram was observed for saponified SBP, WAX-i, CS, WS, and CSlignin, and corresponded to the predominant MS peak comprising a m/z 389 (390 Da; **Supplementary Figure S2B**). Based on the co-elution with 8-8′-(furan)-diFA and the main MS^2 fragment (m/z 393; **Supplementary Figure S2A** and **Table 2**), it is plausible that this compound is related to 8-8′-(furan)-diFA.

Further, two structurally unknown diFAs with a m/z 401 were observed (402 Da, R_t 9.74 and 10.66 min, respectively) (**Table 2** and **Supplementary Figures S2E,G**). Although these diFAs have been previously observed, their structure has not yet been elucidated (Appeldoorn et al., 2010; Kühnel et al., 2012). Based on the mass, it is likely that both compounds are dehydrodimers composed of FA and 5-hydroxyferulic acid, which are also involved in HCA biosynthesis (Humphreys et al., 1999; Morreel et al., 2004). Furthermore, we observed that the diFAs m/z $401_{10.66\ min}$ was unstable during saponification (**Tables 2**, **3** and **Supplementary Figure S2G**). Finally, only one triFA was identified (R_t 11.35 min, **Table 2** and **Supplementary Figure S3**).

Analysis of the Free and Bound Content of HCAs

The amount of ester-linked HCAs that were released after saponification (0.5 M KOH) was set to 100%, corresponding to the highest possible release of HCAs (**Table 3**). Determination of these HCAs for all substrates was performed based on MS, except for SBP and WAX-i. The HCA-content of the latter two

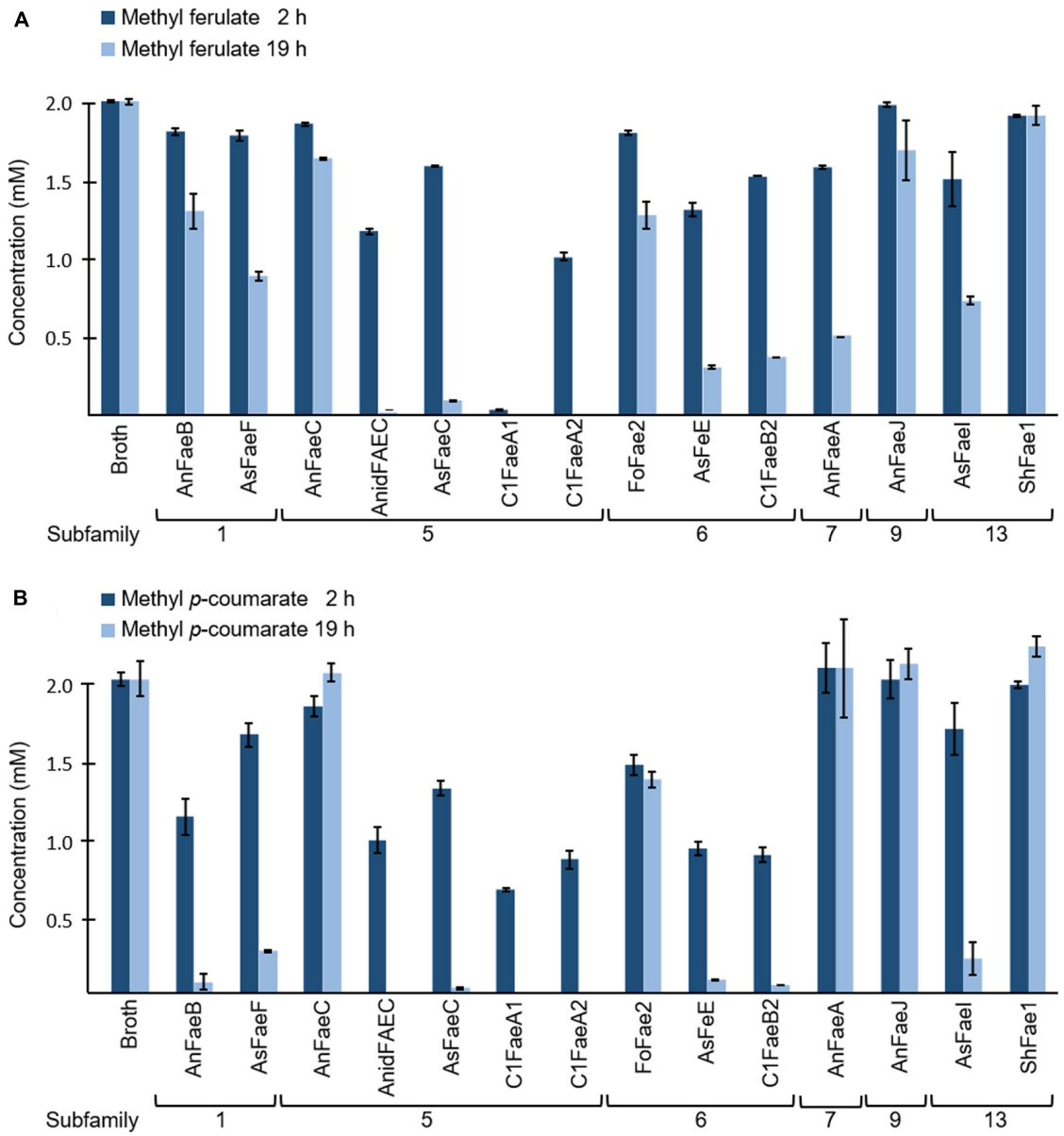


FIGURE 2 | Specificity of FAEs toward **(A)** methyl ferulate (MF) and **(B)** methyl *p*-coumarate (MpC). Both methyl ferulate (2 mM) and methyl *p*-coumarate (2 mM) were incubated with and without (broth) FAEs at 37°C for 2 and 19 h. The reduction in methyl ferulate and methyl *p*-coumarate concentration was measured in duplicate using UHPLC-UV ($n = 2$). The broth is the culture supernatant of *P. pastoris*, which was grown without FAE insertion (negative control). Error bars represent the average standard deviation based on determined absolute numbers.

were determined based on UV (see M&M for details). Since di- and tri-FA standards were not available, HCAs were estimated based on calibration with "corrected" FA and *p*CA standards (see M&M for details).

The amount of FA, *p*CA, diFAs, and triFAs that were present in the PCW-derived and natural substrates differed vastly (**Table 3**). As an example, CFoligo was highly feruloylated with FA, diFAs, and triFAs, whereas neither CS, WS, nor CSlignin contained triFAs. The amounts of esterified FA and *p*CA corresponded well with what has been described previously (Micard et al., 1996; Saulnier et al., 2007; Appeldoorn et al., 2013; Damásio et al., 2013). The estimated total amount of HCAs present in SBP, CFoligo, WAX-i, CS, WS, and CSlignin summed up to 1.7, 62.3, 2.9, 10.8, 5.4, and 17.2% w/w (sum of % free and % bound), respectively, based on dry matter (**Table 3**).

The characterization above shows that the chosen PCW-derived and natural substrates provide a wide range of HCAs that are bound to different structural moieties. Hence, these substrates were subjected to our FAEs in order to further explore the specificity of FAEs from different subfamilies.

TABLE 2 | Esterified diferulic (diFAs) and triferulic (triFAs) acids analyzed by RP-UHPLC-UV-ESI-MS/MS.

Compound	Rt (min)	Theoretical mass	Observed mass $[M-H]^{-}$ †	MS^2 product ions (relative intensity)‡
8-8′-(furan)-diFA	8.23	404.11	403	134 (7.8), 151 (7.2), 165 (7.1), 178 (7.5), 193 (100), 209 (5.0), 215 (12.9), 341 (59.9)
m/z $389_{8.22\ min}$	8.22	390.13	389	151 (5.8), 165 (7.0), 178 (6.4), 181 (7.8), 193 (100), 195 (16.0), 321 (30), 341 96.9), 343 (13.0)
8-8′-(aryl)-diFA	8.36	386.10	385	341 (100)
8-5′-diFA	8.72	386.10	385	297 (29.6), 341 (100)
m/z $401_{9.74\ min}$	9.74	402.10	401	325 (8.0), 357 (100.0)
5-5′-diFA	10.03	386.10	385	282 (6.9), 326 (11.6), 341 (100), 342 (9.7), 370 (13.5)
m/z $401_{10.66\ min}$	10.66	402.10	401	191 (25.3), 235 (9.1), 357 (30.9), 371 (34.5), 383 (100)
8-*O*-4′-diFA	11.00	386.10	385	193 (100), 249 (8.1), 313 (70.4), 317 (7.4), 341 (79.1)
triFA 1§	11.34	578.14	577	311 (6.5), 355 (34.2), 489 (5.7), 533 (100.0), 534 (6.2)

*Tentative structures have been proposed based on their retention time and MS^2 fragmentation pattern (**Figure 1** and **Supplementary Figure S2**; Appeldoorn et al., 2010; Kühnel et al., 2012; Vismeh et al., 2013). †[M-H]⁻ refer to the observed mass peak representing the negative ions; loss of 1 Da. ‡Intensity: the highest peak correspond to 100%, the other peaks are given as percentages hereof. Only peaks which were higher than 5% in the intensity are reported. §TriFA $2_{12.38\ min}$ is not shown, as no MS^2 data was obtained for this compound.*

TABLE 3 | Determined (free and bound) hydroxycinnamic acids in PCW-derived and natural substrates.

Compound	Type	SBP† (μg/mg sample)	CFoligo (μg/mg sample)	WAX-i† (μg/mg sample)	CS (μg/mg Sample)	WS (μg/mg sample)	CSlignin (μg/mg sample)
FA‡	Free	0.05 ± 0.01	74.5 ± 1.6	0.02 ± 0.00	0.13 ± 0.00	0.06 ± 0.00	0.01 ± 0.00
	Bound	15.0 ± 0.3	57.3 ± 1.2	15.0 ± 0.2	12.7 ± 0.6	5.51 ± 0.09	8.82 ± 0.16
*p*CA‡	Free	0.04 ± 0.00	5.58 ± 0.20	0.04 ± 0.00	1.13 ± 0.01	0.39 ± 0.00	0.78 ± 0.01
	Bound	0.21 ± 0.00	6.54 ± 0.12	0.98 ± 0.04	28.4 ± 1.4	6.60 ± 0.16	98.3 ± 2.8
m/z $389_{8.22\ min}$	Free	0.02 ± 0.00	1.08 ± 0.12§	0.01 ± 0.00	0.46 ± 0.05	1.01 ± 0.08	0.03 ± 0.00
	Bound	0.60 ± 0.12	20.0 ± 0.4§	0.51 ± 0.03	27.9 ± 1.0	30.3 ± 0.6	45.9 ± 2.2
8-8′-aryl-diFA	Free	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	Bound	n.d.	45.1 ± 2.3	1.13 ± 0.04	3.39 ± 0.23	0.99 ± 0.05	2.37 ± 0.17
8-5′-diFA	Free	n.d.	4.13 ± 0.15	n.d.	n.d.	n.d.	n.d.
	Bound	0.23 ± 0.01	56.6 ± 1.2	1.13 ± 0.04	13.8 ± 0.7	2.45 ± 0.09	5.45 ± 0.23
m/z $401_{9.74\ min}$	Free	n.d.	7.10 ± 0.03	n.d.	0.20 ± 0.00	n.d.	n.d.
	Bound	n.d.	19.9 ± 0.4	n.d.	5.54 ± 0.08	0.97 ± 0.02	9.31 ± 0.04
5-5′-diFA	Free	n.d.	15.6 ± 1.0	n.d.	n.d.	n.d.	n.d.
	Bound	0.21 ± 0.00	154 ± 4	2.58 ± 0.07	8.35 ± 0.27	2.66 ± 0.09	1.29 ± 0.13
m/z $401_{10.66\ min}$¶	Free	n.d.	5.83 ± 0.25	n.d.	n.d.	0.17 ± 0.00	n.d.
	Bound	0.04 ± 0.00	*n/a*	0.18 ± 0.00	0.32 ± 0.01	0.16 ± 0.02	n.d.
8-O-4′-diFA	Free	n.d.	3.55 ± 0.10	n.d.	n.d.	n.d.	n.d.
	Bound	0.54 ± 0.02	106 ± 3	2.97 ± 0.00	6.00 ± 0.16	2.54 ± 0.19	n.d.
TriFA $1_{11.35\ min}$	Free	n.d.	0.64 ± 0.10	n.d.	n.d.	n.d.	n.d.
	Bound	0.06 ± 0.00	26.8 ± 0.9	2.16 ± 0.06	n.d.	n.d.	n.d.
TriFA $2_{12.38\ min}$	Free	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	Bound	0.05 ± 0.00	13.9 ± 0.7	2.22 ± 0.06	n.d.	n.d.	n.d.
Total	Free	0.11	118	0.07	1.92	1.63	0.82
	Bound	17.0	505	28.9	106	52.2	171
Total (%) w/w	Free	0.01	11.8	0.01	0.19	0.16	0.08
	Bound	1.70	50.5	2.89	10.6	5.22	17.1

*Structural details of the analyzed HCAs (mono, di-, and triFAs) are presented above (**Figure 1**, **Supplementary Figures S2, S3**, and **Table 2**). Numbers of the quantified (FA, pCA) and estimated free and bound (total – free amount) HCAs content is presented in μg/mg sample (dry matter). †Estimated by UV. ‡Estimated by UV. Correlation to MS determination confirmed (**Supplementary Figure S1**). § The m/z 389 was not observed, however, the major mass peak with m/z 403 corresponds to the proposed 8-8′-(furan)-diFA. ¶The measured bound amount was less than what has been released when the substrate was incubated with FAEs (see later results). Consequently, this compound is rendered unstable under the basic conditions (saponification at 0.5 M KOH) used for measurement of the total amounts. See M&M for details. CFoligo, corn fiber oligomers; CS, corn stover; CSlignin, corn stover lignin isolate; diFA, diferulic acid; FA, ferulic acid; n/a, not applicable; n.d, not detected; pCA, p-coumaric acid; SBP, sugar beet pectin; triFA, triferulic acid; WAX-i, insoluble wheat arabinoxylans; WS, wheat straw.*

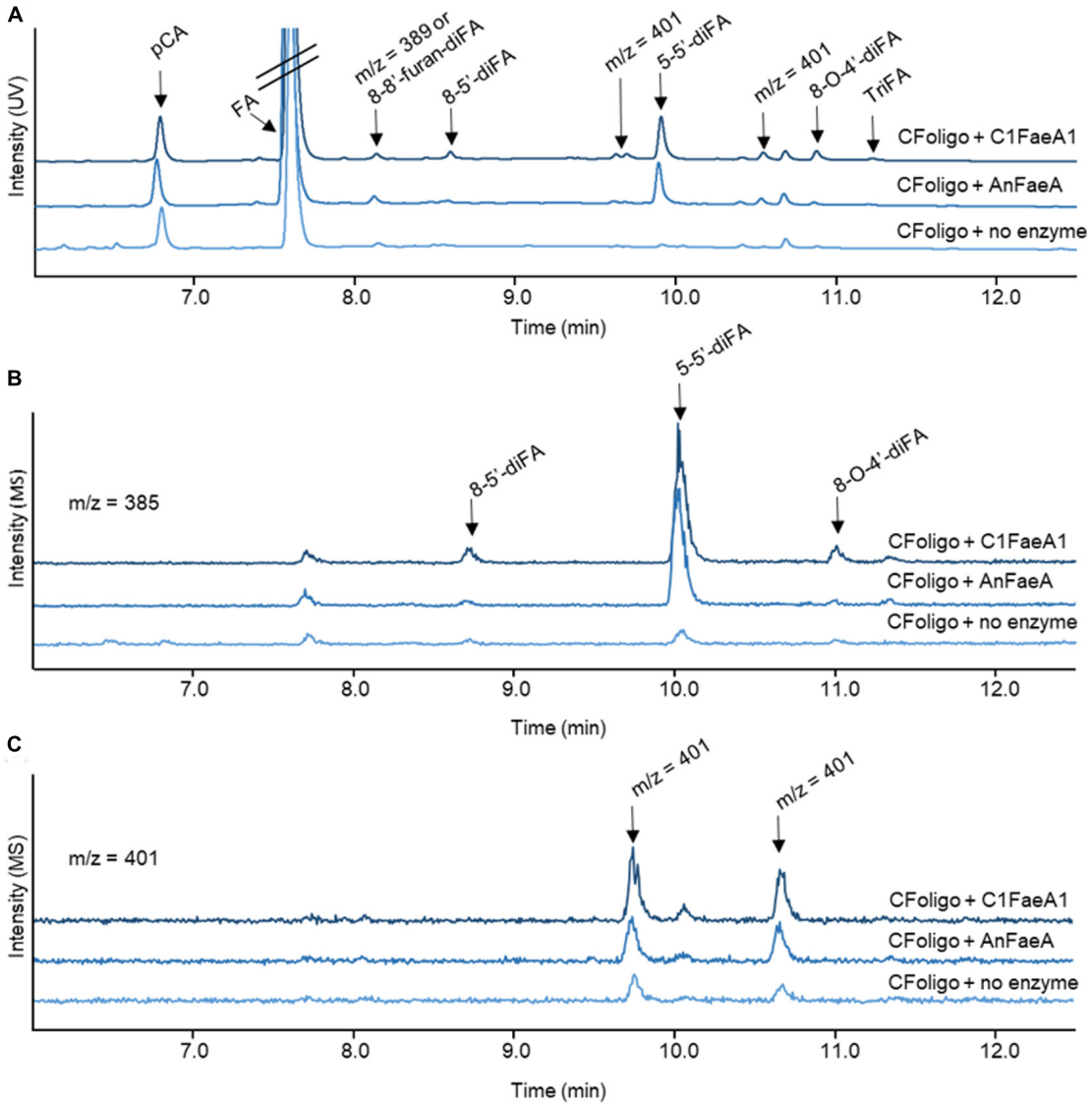


FIGURE 3 | RP-UHPLC-UV-ESI-MS/MS elution profiles for corn fiber oligomers (CFoligo) incubated with FAEs. CFoligo was incubated with either C1FaeA1 or AnFaeA, and without FAE. **(A)** Elution profile based on UV (310 nm) showing the release of ferulic acid (FA); *p*-coumaric acid (*p*CA), diferulic acids (diFAs) and triferulic acids (triFAs) from CFoligo after the incubation with FAEs. **(B,C)** Elution profile of diferulic acids (diFAs) corresponding to the m/z 385 and 401, respectively. The m/z values represent the mass loss of 1 Da (MS operation in negative mode).

Specificity of FAEs Toward PCW-Derived and Natural Substrates

The specificity of the 14 FAEs toward the PCW-derived and natural substrates was determined by using RP-UHPLC-UV-ESI-MS/MS.

As an example, the reversed-phase elution profiles of CFoligo prior to and after incubation with two different FAEs is shown in **Figure 3**. This figure shows that from CFoligo, FA was most abundantly present in the CFoligo sample before and after incubation with C1FaeA1 and AnFaeA. Furthermore, the release of several diFAs was observed, of which the most pronounced was 5-5′-diFA.

Both FA and *p*CA were released from the incubations of CFoligo, CS, WS, and CSlignin with all 14 FAEs (**Figure 4** and **Supplementary Table S4**). Based on maximum relative amounts (100%) of ester bound FA or *p*CA in the tested substrates, FA is released to a larger extent than *p*CA by all FAEs (**Supplementary Table S4**). However, the proportions of these detected compounds was rather different when absolute amounts of released FA and *p*CA were considered (**Supplementary Table S4**). For the primary aim of this work, only relative released amounts were considered. Overall, FAEs from SF5, SF6, and SF7 released higher amounts of FA and *p*CA than FAEs from SF1, SF9, and SF13 under the present conditions (**Figure 4**). As for the model substrates mentioned above, the incubation of all

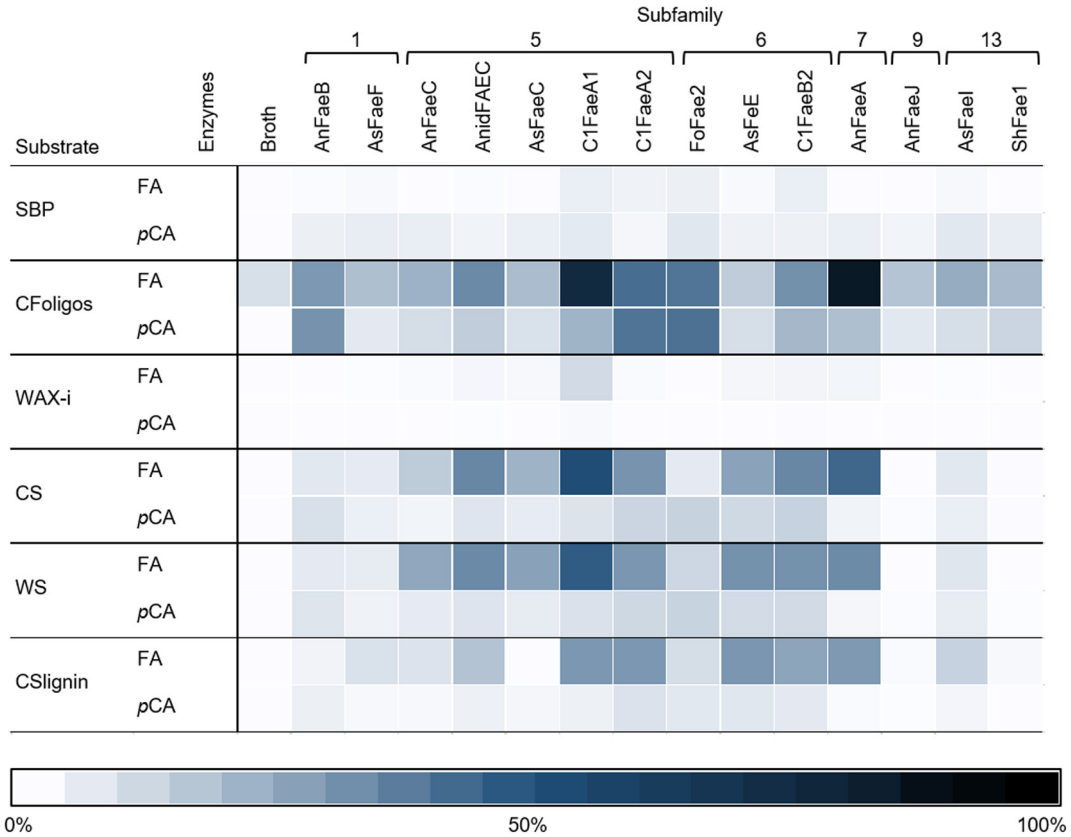


FIGURE 4 | The heatmap shows the specificity of FAEs (representing 6 SFs) toward six substrates. Corn fiber oligomers (CFoligo), corn stover (CS), corn stover lignin isolate (CSlignin), sugar beet pectin (SBP), insoluble wheat arabinoxylans (WAX-i), and wheat straw (WS) were incubated with FAEs and without (broth) at 37°C for 19 h and the release of FA and pCA was measured by RP-UHPLC-UV (n = 2). The release of FA and pCA from the substrates by the FAEs are given as percentages of the total amount of ester bound FA and pCA that have been determined by saponification (0.5 M KOH). The broth is the culture supernatant of *P. pastoris* which was grown without FAE insertion (negative control).

substrates with the control *P. pastoris* fermentation broth did not lead to a substantial release of FA or *p*CA (**Figure 4** and **Supplementary Table S4**).

In addition to FA and *p*CA, most of the FAEs released various diFAs and triFAs, especially from CFoligo, CSlignin, CS, and WS (**Figures 5**, **6** and **Supplementary Table S3**). A clear release of diFAs was only observed for the FAEs classified as SF5 and SF7. FAEs of the other SFs did not release diFAs from these substrates. In contrast, the incubation of SBP and WAX-i with the FAEs released either none or very low amounts of the various diFAs and triFAs (**Figure 5** and **Supplementary Table S5**).

The diFA composing a m/z value of $401_{10.66\ min}$ was unstable during saponification (0.5 M KOH). Therefore, the total amount of the diFA (m/z $401_{10.66\ min}$) could not be determined and the relative release of this diFA from the substrates incubated with FAEs was not calculated. As a result, the released amounts of diFA m/z $401_{10.66\ min}$ are presented separately in **Figure 7** (**Supplementary Table S6**). Apart from CFoligo, this diFA was released in similar amounts by FAEs of SF5, SF7, and C1FaeB2 of SF6, whereas this diFA was either not or barely released by FAEs of SF1, 9, and 13.

DISCUSSION

Fungal FAEs From Different SFs

In this study, we determined the specificity of 14 fungal FAEs from six different SFs via the analysis of the release of different HCAs from two model substrates and six PCW-derived and natural substrates (Dilokpimol et al., 2016). It is expected that these results provide further insights in the predictability of the SF classification for substrate specificities of FAEs, which eases the selection of suitable FAE candidates for industrial applications.

Notably, some FAE family members, especially of SF 5 and 6, have also been described to release acetic acid from acetyl-esterified model substrates (**Table 1**). In this study, the release of acetyl residues from PCW-derived and natural substrates has not been further investigated. Three FAEs (C1FaeA1, C1FaeA2, and C1FaeB2) have previously been purified and intensely studied by Kühnel et al. (2012). These FAEs were chosen because they enlarge SF5 and 6, and their activity toward a variety of substrates have been described earlier. The other expressed FAEs that were used in this study were pre-purified by filtration from the culture supernatant.

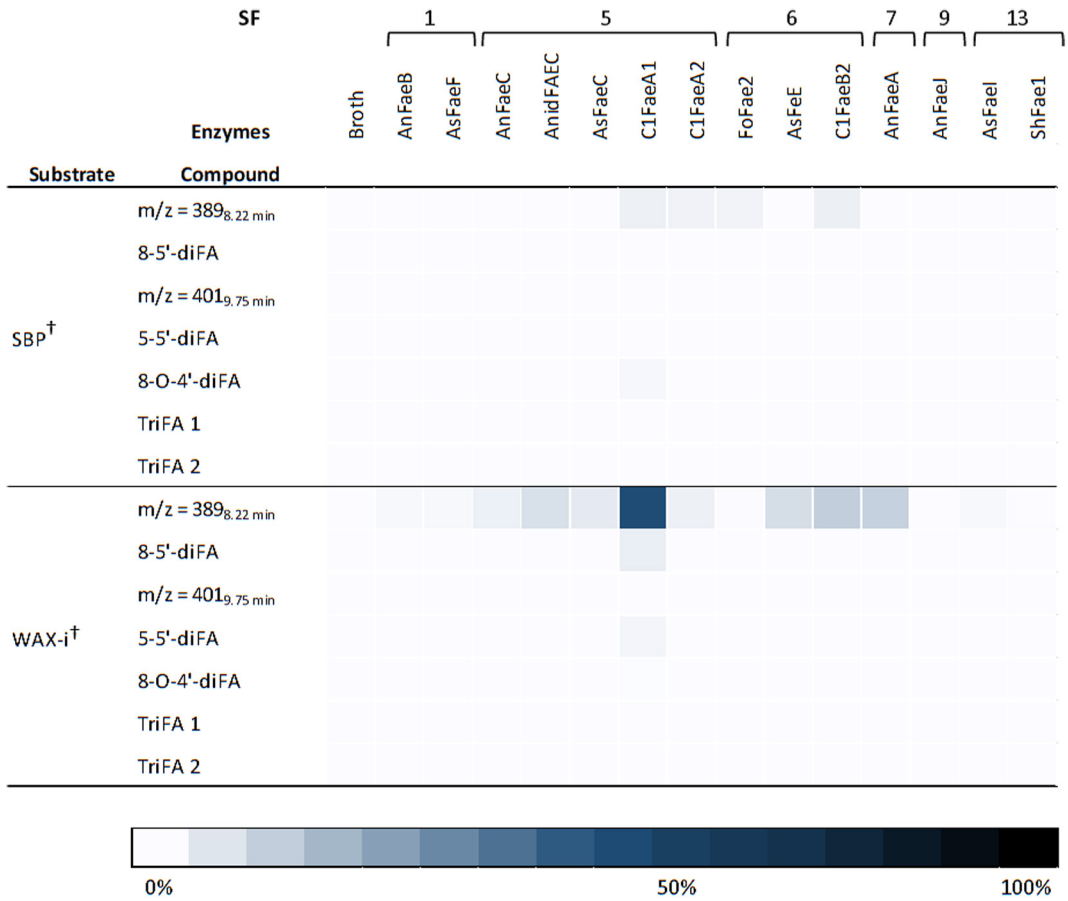


FIGURE 5 | Heatmap presenting the FAE activity toward diFAs and triFAs from incubations with SBP and WAX-i. Substrates were incubated with the 14 FAEs and the fermentation broth for 19 h at 37°C in duplicates. The results are gives as percentages of bound content as measured by RP-UHPLC-UV (n = 2). Abbreviations: diFA, diferulic acid; SBP, sugar beet pectin; triFA, triferulic acid; WAX-i, insoluble wheat arabinoxylans. The broth is the culture supernatant of Pichia pastoris which was grown without FAE insertion (negative control). Error bars represent the average standard deviation based on determined absolute numbers. †Estimated by UV. See M&M for details.

It is important to note that we mainly point out differences between the substrate specificities of FAEs from different SF toward the tested substrates, rather than specify the exact amounts of HCAs released (**Table 4**). All incubations were performed with the same protein concentration and it is likely that the amount of FAEs, which were present in the protein fractions varied due to differences in the expression efficiency. Moreover, all incubations have been performed at the same conditions (37°C and pH 5.8). These conditions were closely related to optimal pH, temperature, and stability of most of these FAEs reported in previous works (**Table 1**; Kühnel et al., 2012; Dilokpimol et al., 2016). Thus, the absolute catalytic performance of the FAEs that was determined based on the release of HCAs from various substrates might be higher at the optimal pH condition. Still, variations in the specificity of these FAEs among these substrates, like methyl ferulate and methyl *p*-coumarate, and the release of structural-different HCAs from the six PCW-derived and natural substrates can be determined. Finally, the representation of SF7 and SF9 has to be critically evaluated, since only one FAE of SF7 and one of SF9 was investigated in this study (**Table 1**). To indisputably define the substrate specificity of these SFs, more SF9 and SF13 candidates need to be studied.

Activity of FAEs Toward Synthetic Model Substrates

First, the enzyme activity of FAEs toward two synthetic FAE substrates (methyl ferulate and methyl *p*-coumarate) was compared. Among all FAEs, SF5 FAEs showed the highest activity toward both model substrates (**Table 4**). Moreover, SF1 FAEs (AnFaeB and AsFaeF) were moderately active toward methyl ferulate, but showed a strong activity toward methyl *p*-coumarate, which is in line with previously reported data (**Figure 2** and **Table 1**; Crepin et al., 2004; Dilokpimol et al., 2017, 2018; Antonopoulou et al., 2018). SF6 FAEs has previously been shown to cleave both methyl ferulate and methyl *p*-coumarate, which is agreement with our findings (**Figure 2** and **Table 1**; Kühnel et al., 2012; Dilokpimol et al., 2018). Interestingly, AnFaeA of SF7 showed a high specificity toward methyl ferulate, where 75% had been degraded after 19 h, but was inactive toward

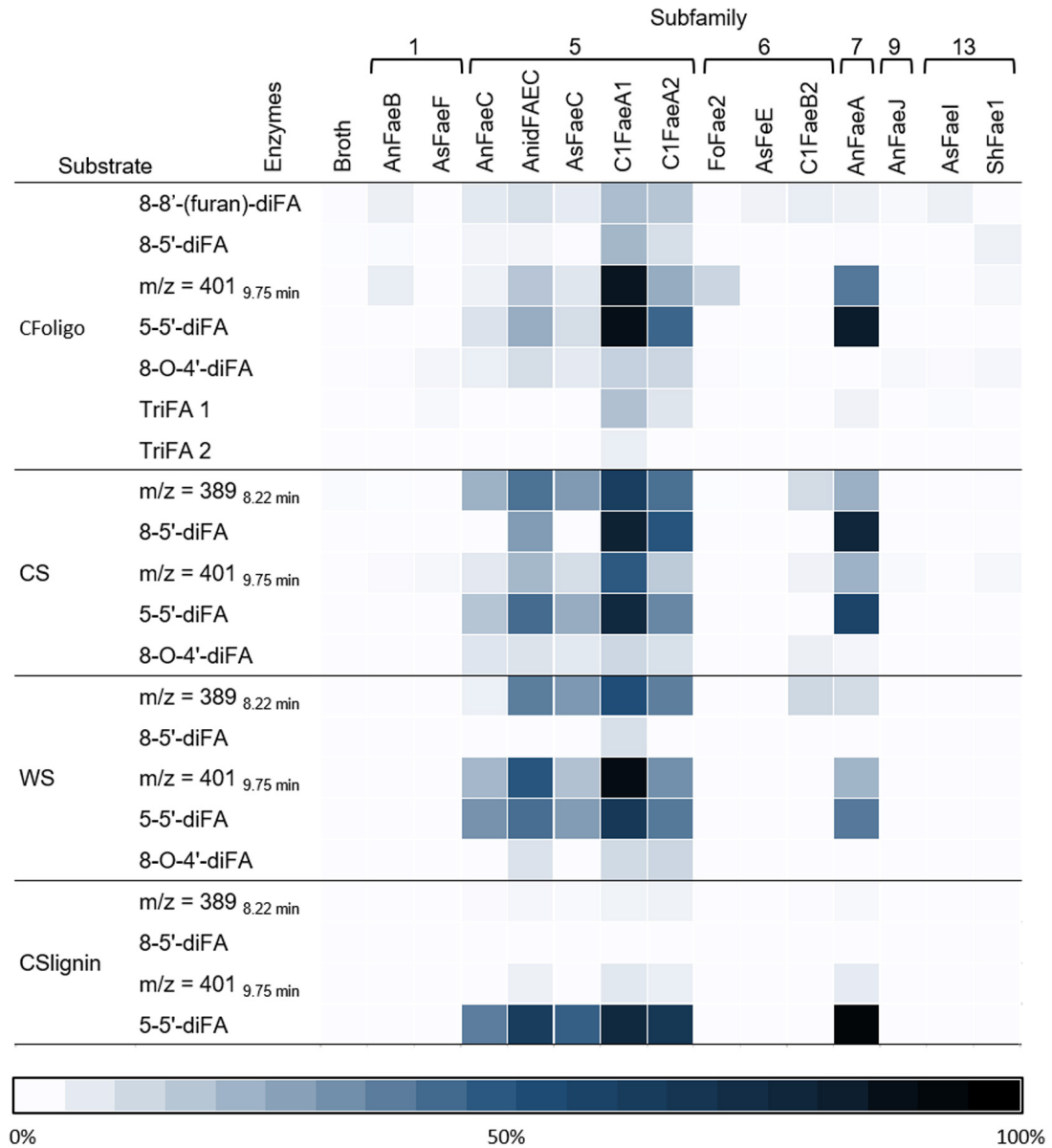


FIGURE 6 | Heatmap presenting the specificity of FAEs. This heatmap highlights the release of diFAs and triFAs by 14 FAEs (representing 6 SFs) from the PCW-derived and natural substrates corn fiber oligomers (CFoligo), corn stover (CS), wheat straw (WS), and corn stover lignin isolate (CSlignin). All substrates were incubated with FAEs and without (broth) at 37°C for 19 h in duplicate. The results are given as percentages of bound content as measured by RP-UHPLC-UV-ESI-MS/MS (n = 2). The broth is the culture supernatant of *P. pastoris* which was grown without FAE insertion (negative control).

methyl *p*-coumarate. A previous study has shown that the Tyr80 of this AnFaeA interacts with the methoxyl group at the C3 position of FA (Hermoso et al., 2004; McAuley et al., 2004; Faulds et al., 2005). An absence of this methoxyl group may reduce the substrate binding, which could possibly explain why AnFaeA did not cleave methyl *p*-coumarate. Surprisingly, AnFaeA did release *p*CA from PCW-derived and natural substrates (**Figures 2**, **4**). How the different structural properties of these substrates influence the substrate binding of AnFaeA is currently unknown. AsFaeJ (SF9) and ShFae1 (SF13) did not cleave methyl ferulate nor methyl *p*-coumarate, whereas AsFaeI (SF13) was active toward both methyl ferulate and methyl *p*-coumarate (**Figure 2**). Apparently, SF13 FAEs vary in their substrate specificity toward these model substrates, as the SF13 FAEs *Um*ChlE from *Ustilago maydis* showed activity toward both methyl ferulate and methyl *p*-coumarate, what is similar to the AsFaeI (Nieter et al., 2015).

Activity of FAEs Toward Six PCW-Derived and Natural Substrates

Next, the specificity for releasing various HCAs of the 14 FAEs toward six PCW-derived and natural substrates was determined.

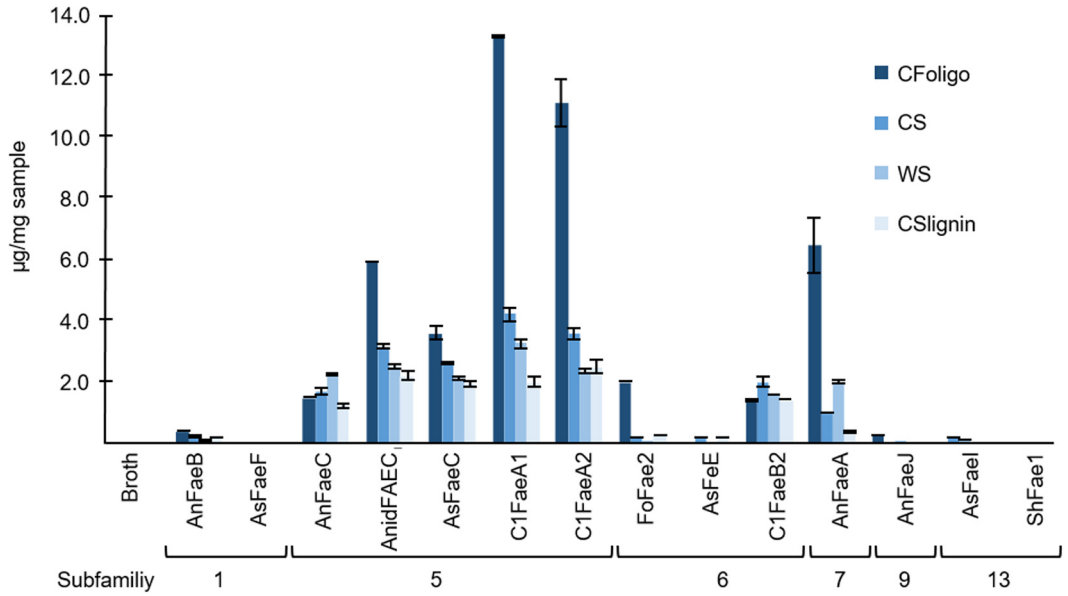


FIGURE 7 | The release of the diFA corresponding to a m/z value of $401_{10.66\ min}$. The release of this diFA (µg of compound/mg dry matter of natural substrate) from corn fiber oligomers (CFoligo), corn stover (CS), wheat straw (WS), and corn stover lignin isolate (CSlignin) that were incubated with and without (broth) FAEs at 37°C for 19 h. Samples were measured by RP-UHPLC-UV-ESI-MS/MS (n = 2). The broth is the culture supernatant of *P. pastoris* which was grown without FAE insertion (negative control). Error bars represent the average standard deviation based on determined absolute numbers.

Generally, the FAEs exhibited a low, but still detectable specificity toward SBP and WAX-i (**Figures 4**, **5** and **Table 4**). These findings match those of previous studies, which showed that FAEs comprise a low specificity toward branched and highly substituted substrates, like SBP and WAX-i. The incubation of these substrates with hydrolases, like xylanases or pectinases, will lead to a cleavage of the backbone chain and, thereby, an enhanced solubility of the substrates (Crepin et al., 2004; Kühnel et al., 2012; Kumar et al., 2013; Mäkelä et al., 2018). Some of the FAEs tested might comprise activity toward acetyl esters, which could have decreased the stearic hindrance and improved the binding toward the substrates tested (Tenkanen et al., 1991; Hashimoto et al., 2010).

Incubations of the natural substrates CS and WS with FAEs demonstrated that FAEs of SF5, SF6, and SF7 released the same types of HCAs from these substrates as observed for the fully soluble CFoligo (**Figures 4**, **6** and **Table 4**). The physical pretreatment (ball-milling) that was applied to the substrates enhanced the accessibility and also partly dissolved the cell walls of CS (26.6 ± 0.9% w/w dry matter) and WS (17.8 ± 1.1% w/w dry matter). It can, however, still be argued that FAEs were also active toward the insoluble fraction of CS and WS, since, for example, C1FaeA1 released 64–92% of various diFAs, such as 8-5′-diFA, 5-5′-diFA, and m/z $401_{9.75\ min}$ (**Figure 4** and **Supplementary Table S5**) and it is unlikely that these HCAs originated entirely from the soluble fraction. The latter is, however, not further confirmed in this work.

The two tested SF1 FAEs showed a low release (<20%) of FA and *p*CA from the xylan-type substrates (CS, WS, and CSlignin) and a moderate release of both FA and *p*CA was observed from soluble CFoligo (**Figures 4**, **6**, **Table 4**, and **Supplementary Table S4**). Moreover, SF1 FAEs set free only very low amounts of diFAs and triFAs from the xylan-type substrates, whereas higher amounts of these compounds were released from soluble CFoligo. While SF13 FAEs exhibited an almost similar substrate specificity toward these substrates as SF1 FAEs, their subdivision is justified based on their different specificity toward, in particular, methyl *p*-coumarate (**Figure 2** and **Table 4**; Nieter et al., 2015; Dilokpimol et al., 2016). The latter is also reflected in the relatively higher absolute amounts of *p*CA than FA released from CFoligo by SF1 compared to SF13 FAEs (**Supplementary Table S4**). Similar to SF1 and SF13, the SF9 FAE did not release any HCAs from the PCW-derived and natural substrates tested, except for CFoligo (**Figure 6**). Moreover, no *p*CA was released from CFoligo by this SF9 FAE, which indicates that this FAE is distinctive from SF1 and SF13 FAEs. The inactivity of this SF9 enzyme might coincide with its similarity to other SF9 tannase-like members, which are not active toward these types of ester-linked HCAs and, possibly, should not be considered true FAEs (Dilokpimol et al., 2016; Antonopoulou et al., 2018). However, possible conclusions in relation to SF9 FAEs are limited, as only one member has been investigated in this study.

FAEs of SF5 and SF6, released FA and *p*CA from most PCW-derived and natural substrates (CFoligo, CS, WS, and CSlignin) (**Figure 4** and **Table 4**). In contrast, SF7 released FA from these substrates, but was only able to release *p*CA from the soluble CFoligo (**Figure 4**). In addition, the FAEs from SF5 and SF7 released higher amounts of diFAs and triFAs, compared to all other FAEs tested. AnFaeC (SF5) released diFAs from CFoligo, CS, WS, and CSlignin, while previous research has indicated

TABLE 4 | Overview of the HCA-release from the incubation of synthetic model substrates, plant cell wall-derived substrates, including lignin, and natural substrates with FAE of various SFs.

		Subfamily					
Substrates	**HCA**	**1**	**5**	**6**	**7**	**9**	**13**
MF	MF[a]	++	±| +++	++| +++	++	±	± | ++
M*p*C	M*p*C[a]	+++	±| +++	++| +++	±	±	± | +++
SBP	FA	±	±	±	±	±	±
	*p*CA	±	±	±	±	±	±
	DiFAs	±	±	±	±	±	±
	TriFAs	±	±	±	±	±	±
CFoligos	FA	± | ++	± | ++	± | ++	+++	±	± | ++
	*p*CA	± | ++	± | ++	± | ++	±	±	±
	DiFAs	±	± | ++ | +++	±	± | ++ | +++	±	±
	TriFAs	±	±	±	±	±	±
WAX-i	FA	±	±	±	±	±	±
	*p*CA	±	±	±	±	±	±
	DiFAs	±	± | ++	±	±	±	±
	TriFAs	±	±	±	±	±	±
CS	FA	±	± | ++	± | ++	++	±	±
	*p*CA	±	±	±	±	±	±
	DiFAs	±	±| ++	±	±| ++	±	±
	TriFAs	n.a.[b]	n.a.[b]	n.a.[b]	n.a.[b]	n.a.[b]	n.a.[b]
WS	FA	±	±| ++	±| ++	++	±	±
	*p*CA	±	±	±	±	±	±
	DiFAs	±	±| ++ | +++	±	±| ++	±	±
	TriFAs	n.a.[b]	n.a.[b]	n.a.[b]	n.a.[b]	n.a.[b]	n.a.[b]
CSlignin	FA	±	±| ++	±| ++	++	±	±
	*p*CA	±	±	±	±	±	±
	DiFAs	±	± | ++	±	± | +++	±	±
	TriFAs	n.a.[b]	n.a.[b]	n.a.[b]	n.a.[b]	n.a.[b]	n.a.[b]

*The symbols represent percentages which are calculated based on the release of HCAs (FA, pCA, diFAs and triFAs) from substrates incubated with FAEs of one SF compared to the total amount of ester bond HCAs that have been determined by saponification of the substrates. The type of HCA and the ratio within these compounds released varied for different FAEs of one SF (**Figures 4–6**), which leads to the occurrence of more than one symbol. The substrate specificity is shown as low ("±"; 0–20%), medium ("++"; 20–80%) and/or high ("+++"; 20–80%) values. [a]Results represent the incubation of MF and MpC with FAEs for 19 h (**Figure 4**). [b]No release of TriFA was determined from the incubation of these substrates with FAEs.*

that this FAE did not release diFAs from PCW-derived and natural substrates, such as WAX-i and SBP (Dilokpimol et al., 2017). Most likely, the absence of released diFAs from WAX-i and SBP by FAEs relates to the low amounts of diFAs that are present in these substrates (**Table 2**). SF6 FAEs hardly released diFAs and triFAs, except for the incubation of CS with C1FaeB2 (**Figures 5**, **6**). The general inability to release diFAs from PCW-derived and natural substrates of the SF6 FAEs is in accordance with previous data, which reported SF6 FAEs as type B FAEs which are incapable to release diFAs from PCW materials (**Table 1**; Crepin et al., 2004; Dilokpimol et al., 2016). Interestingly, in comparison to the SF6 enzyme AsFaeE, the incubation of CFoligo with the SF6 enzyme FoFae2 showed the highest specificity for FA and *p*CA, while for the synthetic substrates, CS, WS, and CSlignin, AsFaeE released higher levels of FA and *p*CA compared to FoFae2 (**Figure 4**). These results indicate mostly the variation in substrate specificity amongst FAEs of the same SF, which likely results from variations of the amino acids that are present in the proximity of the substrate-binding site and the catalytic pocket of the FAEs (Hermoso et al., 2004; McAuley et al., 2004; Faulds et al., 2005; Goldstone et al., 2009; Uraji et al., 2018).

Remarkably, SF5 enzymes were ranked as different types according to the ABCD classification, such as type A, C, or C/D, which does not reflect their overall ability to release *p*CA, FA, diFAs, and triFAs from both synthetic as well as PCW-derived and natural substrates (**Table 1**; Debeire et al., 2012; Kühnel et al., 2012; Dilokpimol et al., 2017, 2018). Based on our results, the co-classification of these FAEs in SF5 shows that the SF classification covers the FAE promiscuity and reflects their common evolutionary origin to a larger extent than the ABCD classification.

In particular, FAEs of SF5, SF6, and SF7 released higher amounts of FA from PCW-derived and natural substrates, compared to *p*CA (**Figure 4**), even though rather high amounts of esterified *p*CA were present in these substrates (i.e., 28.39

and 98.25 μg/mg sample for CS and CSlignin, respectively) (Ralph, 2010; Reinoso et al., 2018).

The type of structure of diFAs released by FAEs of SF5 and SF7 from various substrates differed. Generally, the release of 5-5′-diFA was most prevalent from CFoligo, CS, and WS (up to 91% for CFoligo), followed by m/z $389_{8.22\ min}$ (up to 61% for CS) and m/z $401_{9.75\ min}$ (up to 87 and 92% for CFoligo and WS, respectively). Of the released diFAs from CSlignin by SF5 and SF6 FAEs, 5-5′-diFA was most common (up to 97%), while the m/z $389_{8.22\ min}$, which is the most predominant diFA of this substrate (45.93 μg/mg CSlignin), was hardly released. These differences might again be ascribed to the complexity of the lignin structure (Ralph, 2010). Furthermore, it is tempting to speculate that certain types of diFAs preferentially cross-link between the arabinoxylan chains rather than cross-linking with lignin, which might steer FAE specificity. The 8-*O*-4′ diFA was hardly released from any substrate, not even from CFoligo in which relatively high amounts of this structure were present (153.83 μg/mg sample for CFoligo). This suggests that the 8-*O*-4′ diFA structure is less accessible for FAEs compared to other diFAs, like the 5-5′-diFA (**Figure 1**), which might either be the result of the catalytic site structure of the FAEs or of substrate inaccessibility.

Release of FA and *p*CA From Lignin by FAEs

For the first time we show that several FAEs, such as C1FaeA1, C1FaeA2, and AnFaeA, were able to release *p*CA from lignin (isolated from CS), next to the release of FA and diFAs as mentioned above (**Tables 3**, **4**). Although *p*CA has been shown to be γ-esterified to lignin, ester linkages of FA and diFAs were not identified in the 2D-NMR spectrum nor in literature (**Supplementary File 7**). But, since FA-like HCAs (FA, diFAs and triFAs) have been shown to be esterified to arabinofuranosyl residues (Ara), we have calculated the molar ratio of FA-like HCAs to Ara. Saponification of 1 mg CSlignin led to a release of 0.81 μmole of FA, diFAs, and triFAs (FA-derived HCAs). In comparison, 0.07 μmole arabinofuranosyl residues were determined to be present in CSlignin, which corresponds to a molar ratio of approximately 12:1 (FA-derived HCAs:arabinofuranosyl residues). These results support that a part of the released FA-derived HCAs could be ester-linked to lignin in corn stover (van Erven et al., 2017). For the three FAEs (C1FaeA1, C1FaeA2, and AnFaeA) that showed the highest activity toward CSlignin, between 0.02 and 0.07 μmole of FA-derived HCAs from 1 mg of CSlignin were released resulting in a molar ratio ranging from 0.3:1.0 up to 0.75:1.0 (FA-derived HCAs:arabinofuranosyl residues) (**Table 2**). Thus, all FA and diFAs released by these FAEs could theoretically have originated from highly substituted Ara residues in CSlignin. Since the lignin isolate was thoroughly water-washed, it is expected that the remaining oligosaccharides were bound to lignin (covalent or non-covalent) (van Erven et al., 2017). Though the release of lignin-linked FA-derived HCAs by FAEs can thus not be concluded and needs to be investigated in future work, at the very least, our results signify that HCAs can be released from lignin-bound carbohydrate-HCA complexes.

CONCLUSION

FAE specificity is highly relevant for industrial applications and it is important to evaluate this specificity on PCW-derived and natural substrates rather than synthetic model substrates alone. We demonstrate that FAEs show a vast variation in their ability to hydrolyze HCAs from PCW-derived and natural substrates. Furthermore, the type of HCA and the ratio within these compounds released varied for different FAEs and substrates. In addition to FA and *p*CA, the 5-5′ diFA structure was released most by the FAEs tested, while the 8-*O*-4′- counterpart was released to a much lower extent.

Both substrate specificity and product profiles of the FAEs tested were linked to the new SF classification. This study supports the reliability of the SF classification, which recognizes a more subtle difference in FAE specificity, compared to the ABCD classification. Based on this study, FAEs of SF5 and SF7 showed the highest release of FA, *p*CA, and diFAs over the range of substrates, while FAEs of SF6 were comparable but less pronounced for diFAs release (**Table 4**). These results suggest that SF5 and SF7 FAEs are promising enzymes for biorefinery applications, like the production of biofuels, where a complete degradation of the plant cell wall is desired. In contrast, SF6 FAEs might be of interest for industrial applications that require a high release of only FA and *p*CA, which are needed as precursors for the production of biochemicals. Therefore, our work provides new insights on the selection of suitable FAE-candidates for industrial applications.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

EU, MF, RV, and MK designed the research. EU and MF performed the research and analyzed the data from FAE studies. GE performed isolation and analysis of the CSlignin substrate. AD produced the FAE filtrates. EU, MF, and MK wrote the manuscript. AD, GE, and RV critically revised the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We thank Eva Jermendi, Department of Food Chemistry, Wageningen University & Research for the sugar composition analysis of SBP.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00332/full#supplementary-material

REFERENCES

Antonopoulou, I., Dilokpimol, A., Iancu, L., Mäkelä, M. R., Varriale, S., Cerullo, G., et al., (2018). The synthetic potential of fungal feruloyl esterases: a correlation with current classification systems and predicted structural properties. *Catalysts* 8, 242–264. doi: 10.3390/catal8060242

Appeldoorn, M., de Waard, M. P., Kabel, M. A., Gruppen, H., and Schols, H. A. (2013). Enzyme resistant feruloylated xylooligomer analogues from thermochemically treated corn fiber contain large side chains, ethyl glycosides and novel sites of acetylation. *Carbohydr. Res.* 381, 33–42. doi: 10.1016/j.carres.2013.08.024

Appeldoorn, M. M., Kabel, M. A., van Eylen, D., Gruppen, H., and Schols, H. A. (2010). Characterization of oligomeric xylan structures from corn fiber resistant to pretreatment and simultaneous saccharification and fermentation. *J. Agric. Food Chem.* 58, 11294–11301. doi: 10.1021/jf102849x

Bakker, R. R. C., Elbersen, H. W., Poppens, R. P., Lesschen, J. P., Boschma, S., and Kwant, K. W. (2013). "Rice straw and wheat straw: potential feedstocks for the biobased economy," in *NL Agency Ministry of Economic Affairs* (Wageningen: Wageningen UR, Food & Biobased Research), 1–31.

Benoit, I., Danchin, E. G. J., Bleichrodt, R. J., and de Vries, R. P. (2008). Biotechnological applications and potential of fungal feruloyl esterases based on prevalence, classification and biochemical diversity. *Biotechnol. Lett.* 30, 387–396. doi: 10.1007/s10529-007-9564-6

Broxterman, S. E., van Erven, G., and Schols, H. A. (2018). The solubility of primary plant cell wall polysaccharides in LiCl-DMSO. *Carbohydr. Pol.* 200, 332–340. doi: 10.1016/j.carbpol.2018.07.083

Bunzel, M., Ralph, J., Brüning, P., and Steinhart, H. (2006). Structural identification of dehydrotriferulic and dehydrotetraferulic acids isolated from insoluble maize bran fiber. *J. Agric. Food Chem.* 54, 6409–6418. doi: 10.1021/jf061196a

Carpita, N. C., and Gibeaut, D. M. (1993). Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* 3, 1–30. doi: 10.1111/j.1365-313X.1993.tb00007.x

Carpita, N. C., and McCann, M. C. (2008). Maize and sorghum: genetic resources for bioenergy grasses. *Trends Plant Sci.* 13, 415–420. doi: 10.1016/j.tplants.2008.06.002

Crepin, V. F., Connerton, I. F., and Faulds, C. B. (2004). Identification of a type-D feruloyl esterase from *Neurospora crassa*. *Appl. Microbiol. Biotechnol.* 63, 567–570. doi: 10.1007/s00253-003-1466-5

Crepin, V. F., Faulds, C. B., and Connerton, I. F. (2004). Functional classification of the microbial feruloyl esterases. *Appl. Biochem. Biotechnol.* 63, 647–652. doi: 10.1007/s00253-003-1476-3

Damásio, A. R. L., Braga, C. M. P., Brenelli, L. B., Citadini, A. P., Mandelli, F., Cota, J., et al., (2013). Biomass-to-bio-products application of feruloyl esterase from *Aspergillus clavatus*. *Appl. Microbiol. Biotechnol.* 97, 6759–6767. doi: 10.1007/s00253-012-4548-4

de Oliveira, D. M., Finger-Teixeira, A., Mota, T. R., Salvador, V. H., Moreira-Vilar, F. C., Molinari, H. B. C., et al., (2015). Ferulic acid: a key component in grass lignocellulose recalcitrance to hydrolysis. *Plant Biotechnol. J.* 13, 1224–1232. doi: 10.1111/pbi.12292

de O Buanafina, M. M. (2009). Feruloylation in grasses: current and future perspectives. *Mol. Plant* 2, 861–872. doi: 10.1093/mp/ssp067

de Vries, R. P., Michelsen, B., Poulsen, C. H., Kroon, P. A., van den Heuvel, R. H., Faulds, C. B., et al., (1997). The FaeA genes from *Aspergillus niger* and *Aspergillus tubingensis* encode ferulic acid esterases involved in degradation of complex cell wall polysaccharides. *Appl. Environ. Microbiol* 63, 4638–4644. doi: 10.1128/aem.63.12.4638-4644.1997

de Vries, R. P., vanKuyk, P. A., Kester, H. C., and Visser, J. (2002). The Aspergillus niger FaeB gene encodes a second feruloyl esterase involved in pectin and xylan degradation and is specifically induced in the presence of aromatic compounds. *Biochem. J.* 386, 377–386. doi: 10.1042/0264-6021:3630377

Debeire, P., Khoune, P., Jeltsch, J.-M., and Phalip, V. (2012). Product patterns of a feruloyl esterase from Aspergillus nidulans on large feruloyl-arabino-xylo-oligosaccharides from wheat bran. *Bioresour. Technol.* 119, 425–428. doi: 10.1016/j.biortech.2012.01.185

Dilokpimol, A., Mäkelä, M. R., Aguilar-Pontes, M. V., Benoit-Gelber, I., Hildén, K. S., and de Vries, R. P. (2016). Diversity of fungal feruloyl esterases: updated phylogenetic classification, properties, and industrial applications. *Biotechnol. Biofuels* 9, 231–240.

Dilokpimol, A., Mäkelä, M. R., Mansouri, S., Belova, O., Waterstraat, M., Bunzel, M., et al., (2017). Expanding the feruloyl esterase gene family of *Aspergillus niger* by characterization of a feruloyl esterase, FaeC. *N. Biotechnol.* 37, 200–209. doi: 10.1016/j.nbt.2017.02.007

Dilokpimol, A., Mäkelä, M. R., Varriale, S., Zhou, M., Cerullo, G., Gidijala, L., et al., (2018). Fungal feruloyl esterases: functional validation of genome mining based enzyme discovery including uncharacterized subfamilies. *N. Biotechnol.* 41, 9–14. doi: 10.1016/j.nbt.2017.11.004

Englyst, H. N., and Cummings, J. H. (1984). Simplified method for the measurement of total non-starch polysaccharides by gas-liquid chromatography of constituent sugars as alditol acetates. *Analyst* 109, 937–942. doi: 10.1039/AN9840900937

Faulds, C. B. (2010). What can feruloyl esterases do for us? *Phytochem. Rev.* 9, 121–132. doi: 10.1007/s11101-009-9156-2

Faulds, C. B., Molina, R., Gonzalez, R., Husband, F., Juge, N., Sanz-Aparicio, J., et al., (2005). Probing the determinants of substrate specificity of a feruloyl esterase, AnFaeA, from *Aspergillus niger*. *FEBS Journal* 272, 4362–4371. doi: 10.1111/j.1742-4658.2005.04849.x

Goldstone, D. C., Villas-Bôas, S. G., Till, M., Kelly, W. J., Attwood, G. T., and Arcus, V. L. (2009). Structural and functional characterization of a promiscuous feruloyl esterase (Est1E) from the rumen bacterium *Butyrivibrio proteoclasticus*. *Proteins* 78, 1457–1469. doi: 10.1002/prot.22662

Harris, P. J., and Stone, B. A. (2008). "Chemistry and molecular organization of plant cell walls," in *Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy*, 1st Edn, ed. M. E. Himmel (Hoboken, NJ: Blackwell Publishing Ltd), 61–93. doi: 10.1002/9781444305418.ch4

Harris, P. J., and Trethewey, J. A. K. (2010). The distribution of ester-linked ferulic acid in the cell walls of angiosperms. *Phytochem. Rev.* 9, 19–33. doi: 10.1007/s11101-009-9146-4

Haruhiko, K., Ryuji, S., Shuichi, M., and Moriyasu, T. (2007). Method for producing ferulic acid ester compound with enzymatic method. Japan Patent No JP2007000010. Tokyo: Japan Patent Office.

Hashimoto, K., Kaneko, S., and Yoshida, M. (2010). Extracellular carbohydrate esterase from the basidiomycete Coprinopsis cinerea released ferulic and acetic acids from xylan. *Biosci. Biotechnol. Biochem.* 74, 1722–1724. doi: 10.1271/bbb.100299

Hatfield, R. D., Rancour, D. M., and Marita, J. M. (2017). Grass cell walls: a story of cross-linking. *Front. Plant Sci.* 7:2056. doi: 10.3389/fpls.2016.02056

Hermoso, J. A., Sanz-Aparicio, J., Molina, R., Juge, N., González, R., and Faulds, C. B. (2004). The crystal structure of feruloyl esterase a from *Aspergillus niger* suggests evolutive functional convergence in feruloyl esterase family. *J. Mol. Biol.* 338, 495–506. doi: 10.1016/j.jmb.2004.03.003

Hiroyuki, Y., Kyoko, H., Hideaki, K., Yuto, Y., and Kumiko, K. (2010). Fermented alcoholic beverage with high content of ferulic aid. JP2010148485.

Hoffman, M., Jia, Z., Peña, M. J., Cash, M., Harper, A., Blackburn, A. R., et al., (2005). Structural analysis of xyloglucans in the primary cell walls of plants in the subclass asteridae. *Carbohydr. Res.* 340, 1826–1840. doi: 10.1016/j.carres.2005.04.016

Humphreys, J. M., Hemm, M. R., and Chapple, C. (1999). New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. *Proc. Natl. Acad. Sci. U.S.A.* 96, 10045–10050. doi: 10.1073/pnas.96.18.10045

Kabel, M. A., Bos, G., Zeevalking, J., Voragen, A. G. J., and Schols, H. A. (2007). Effect of pretreatment severity on xylan solubility and enzymatic breakdown of the remaining cellulose from wheat straw. *Bioresour. Technol.* 98, 2034–2042. doi: 10.1016/j.biortech.2006.08.006

Kühnel, S., Pouvreau, L., Appeldoorn, M. M., Hinz, S. W. A., Schols, H. A., and Gruppen, H. (2012). The ferulic acid esterases of *Chrysosporium lucknowense* C1: purification, characterization and their potential application in biorefinery. *Enzyme Microb. Technol.* 50, 77–85. doi: 10.1016/j.enzmictec.2011.09.008

Kumar, C. G., Kamle, A., and Kamal, A. (2013). Purification and biochemical characterization of feruloyl esterases from *Aspergillus terreus* MTCC 11096. *Biotechnol. Prog.* 29, 924–932. doi: 10.1002/btpr.1729

Levigne, S., Ralet, M.-C., Quéméner, B., and Thibault, J.-F. (2004a). Isolation of diferulic bridges ester-linked to arabinan in sugar beet cell walls. *Carbohydr. Res.* 339, 2315–2319. doi: 10.1016/j.carres.2004.07.006

Levigne, S. V., Ralet, M.-C. J., Quéméner, B. C., Pollet, B. N.-L., Lapierre, C., and Thibault, J.-F. J. (2004b). Isolation from sugar beet cell walls of arabinan oligosaccharides esterified by two ferulic acid monomers. *Plant Physiol.* 134, 1173–1180. doi: 10.1104/pp.103.035311

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucl. Acids Res.* 42, 490–495. doi: 10.1093/nar/gkt1178

Loqué, D., Scheller, H. V., and Pauly, M. (2015). Engineering of plant cell walls for enhanced biofuel production. *Curr. Opin. Plant Biol.* 25, 151–161. doi: 10.1016/j.pbi.2015.05.018

Mäkelä, M. R., Dilokpimol, A., Koskela, S. M., Kuuskeri, J., de Vries, R. P., and Hildén, K. (2018). Characterization of a feruloyl esterase from *Aspergillus terreus* facilitates the division of fungal enzymes from carbohydrate esterase family 1 of the carbohydrate-active enzymes (CAZy) Database. *Microb. Biotechnol.* 11, 869–880. doi: 10.1111/1751-7915.13273

McAuley, K. E., Svendsen, A., Patkar, S. A., and Wilson, K. S. (2004). Structure of a feruloyl esterase from *Aspergillus niger*. *Acta Cryst. D* 60, 878–887. doi: 10.1107/S0907444904004937

Micard, V., Renard, C. M. G. C., and Thibault, J. (1996). Enzymatic saccharification of sugar-beet pulp. *Enzyme Micobial Technol.* 19, 162–170. doi: 10.1016/0141-0229(95)00224-3

Mood, S. H., Golfeshan, A. H., Tabatabaei, M., Jouzani, G. S., Najafi, G. H., Gholami, M., et al., (2013). Lignocellulosic biomass to bioethanol, a comprehensive review with a focus on pretreatment. *Renew. Sustain. Energy Rev* 27, 77–93. doi: 10.1016/j.rser.2013.06.033

Morreel, K., Ralph, J., Lu, F., Goeminne, G., Busson, R., Herdewijn, P., et al., (2004). Phenolic proling of caffeic acid O-methyltransferase-deficient poplar reveals novel benzodioxane oligolignols. *Plant Physiol.* 136, 4023–4036. doi: 10.1104/pp.104.049312.The

Nieter, A., Haase-Aschoff, P., Kelle, S., Linke, D., Krings, U., Popper, L., et al., (2015). A chlorogenic acid esterase with a unique substrate specificity from Ustilago maydis. *Appl. Environ. Microbiol.* 81, 1679–1688. doi: 10.1128/AEM.02911-14

Nsereko, V., Rutherford, W., Smiley, B. K., and Spielbauer, A. (2010a). Ferulate esterase producing strains and methods of using same. EP2186422.

Nsereko, V., Rutherford, W., Smiley, B. K., and Spielbauer, A. (2010b). Ferulate esterase producing strains and methods of using same. US20090010903.

Nsereko, V., Rutherford, W., Smiley, B. K., and Spielbauer, A. (2010c). Ferulate esterase producing strains and methods of using same. US20090011085.

Pauly, M., and Keegstra, K. (2008). Cell-wall carbohydrates and their modification as a resource for biofuels. *Plant J.* 54, 559–568. doi: 10.1111/j.1365-313X.2008.03463.x

Ralet, M. C., Thibault, J. F., Faulds, C. B., and Williamson, G. (1994). Isolation and purification of feruloylated oligosaccharides from cell walls of sugar-beet pulp. *Carbohydr. Res.* 263, 227–241. doi: 10.1016/0008-6215(94)00175-8

Ralph, J. (2010). Hydroxycinnamates in lignification. *Phytochem. Rev.* 9, 65–83. doi: 10.1007/s11101-009-9141-9

Ralph, J., Quideau, S., Grabber, J. H., and Hatfieida, R. D. (1994). Identification and synthesis of new ferulic acid dehydrodimers present in grass cell walls. *J. Chem. Soc. Perkin Trans. 1* 23, 3485–3498.

Regner, M., Bartuce, A., Padmakshan, D., Ralph, J., and Karlen, S. D. (2018). Reductive cleavage method for quantitation of monolignols and low-abundance monolignol conjugates. *ChemSusChem* 11, 1600–1605. doi: 10.1002/cssc.201800958

Reinoso, F. A. M., Rencoret, J., Gutiérrez, A., Milagres, A. M. F., del Río, J. C., and Ferraz, A. (2018). Fate of p-hydroxycinnamates and structural characteristics of residual hemicelluloses and lignin during alkaline-sulfite chemithermomechanical pretreatment of sugarcane bagasse. *Biotechnol. Biofuels* 11, 1–12. doi: 10.1186/s13068-018-1155-3

Saulnier, L., Sado, P. E., Branlard, G., Charmet, G., and Guillon, F. (2007). Wheat arabinoxylans: exploiting variation in amount and composition to develop enhanced varieties. *J. Cereal Sci.* 46, 261–281. doi: 10.1016/j.jcs.2007.06.014

Schatz, P. F., Ralph, J., Lu, F., Guzei, I. A., and Bunzel, M. (2006). Synthesis and identification of 2,5-bis-(4-hydroxy-3-methoxyphenyl)-tetrahydrofuran-3,4-dicarboxylic acid, an unanticipated ferulate 8-8-coupling product acylating cereal plant cell walls. *Org. Biomol. Chem.* 4, 2801–2806. doi: 10.1039/b605918j

Shuichi, M., Tadanobu, O., Moriyasu, T., Takuo, C., Megumi, K., and Hisa, A. (2009). Method for producing ferulic acid ester compound by enzymatic method. JP2009089689.

Tenkanen, M., Schuseil, J., Puls, J., and Poutanen, K. (1991). Production, purification and characterization of an esterase liberating phenolic acids from lignocellulosics. *J. Biotechnol.* 18, 69–83. doi: 10.1016/0168-1656(91)90236-O

Topakas, E., Vafiadi, C., and Christakopoulos, P. (2007). Microbial production, characterization and applications of feruloyl esterases. *Process Biochem.* 42, 497–509. doi: 10.1016/j.procbio.2007.01.007

Udatha, D. B. R. K. G., Kouskoumvekaki, I., Olsson, L., and Panagiotou, G. (2011). The interplay of descriptor-based computational analysis with pharmacophore modeling builds the basis for a novel classification scheme for feruloyl esterases. *Biotechnol. Adv.* 29, 94–110. doi: 10.1016/j.biotechadv.2010.09.003

Uraji, M., Tamura, H., Mizohata, E., Arima, J., Wan, K., Ogawa, K., et al., (2018). Loop of Streptomyces feruloyl esterase plays an important role on its activity of releasing ferulic acid from biomass. *Appl. Environ. Microbiol.* 84:e02300-17. doi: 10.1128/AEM.02300-17

van Dongen, F. E. M., van Eylen, D., and Kabel, M. A. (2011). Characterization of substituents in xylans from corn cobs and stover. *Carbohydr. Polym.* 86, 722–731. doi: 10.1016/j.carbpol.2011.05.007

van Erven, G., de Visser, R., Merkx, D. W. H., Strolenberg, W., de Gijsel, P., Gruppen, H., et al., (2017). Quantification of lignin and its structural features in plant biomass using 13C lignin as internal standard for pyrolysis-GC-SIM-MS. *Anal. Chem.* 89, 10907–10916. doi: 10.1021/acs.analchem.7b02632

van Erven, G., Nayan, N., Sonnenberg, A. S. M., Hendriks, W. H., Cone, J. W., and Kabel, M. A. (2018). Mechanistic insight in the selective delignification of wheat straw by three white-rot fungal species through quantitative 13C-IS Py-GC–MS and whole cell wall HSQC NMR. *Biotechnol. Biofuels* 11, 262–278. doi: 10.1186/s13068-018-1259-9

Vismeh, R., Lu, F., Chundawat, S. P. S., Humpula, J. F., Azarpira, A., Balan, V., et al., (2013). Profiling of diferulates (plant cell wall cross-linkers) using ultrahigh-performance liquid chromatography-tandem mass spectrometry. *Analyst* 138, 6683–6692. doi: 10.1039/c3an36709f

West, S. I., and William, H. G. (2010). Use of type C and D feruloyl esterases in the manufacture of biofuels. US20100256353.

West, S. I., Williams, H. G., and Biocatalysts Limited. (2009). Use of type C and D feruloyl esterases in the manufacture of biofuels. WO2009027638.

Wong, D. W. S. (2006). Feruloyl esterase: a key enzyme in biomass degradation. *Appl. Biochem. Biotechnol* 133, 87–112. doi: 10.1385/ABAB:133:2:87

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 26 June 2020
doi: 10.3389/fbioe.2020.00694

Functional Validation of Two Fungal Subfamilies in Carbohydrate Esterase Family 1 by Biochemical Characterization of Esterases From Uncharacterized Branches

Xinxin Li[1†], Kelli Griffin[1†‡], Sandra Langeveld[1], Matthias Frommhagen[2], Emilie N. Underlin[2,3], Mirjam A. Kabel[2], Ronald P. de Vries[1] and Adiphol Dilokpimol[1]*

[1] Fungal Physiology, Westerdijk Fungal Biodiversity Institute, Fungal Molecular Physiology, Utrecht University, Utrecht, Netherlands, [2] Laboratory of Food Chemistry, Wageningen University & Research, Wageningen, Netherlands, [3] Department of Chemistry, Technical University of Denmark, Lyngby, Denmark

OPEN ACCESS

Edited by:
André Damasio,
Campinas State University, Brazil

Reviewed by:
Peter Biely,
Slovak Academy of Sciences (SAS), Slovakia
Lene Lange,
Independent Researcher, Copenhagen, Denmark
Takuya Koseki,
Yamagata University, Japan

***Correspondence:**
Ronald P. de Vries
r.devries@wi.knaw.nl

[†]These authors share first authorship

[‡]**Present address:**
Kelli Griffin,
NTrans Technologies, Utrecht, Netherlands

Specialty section:
This article was submitted to Bioprocess Engineering, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: 13 March 2020
Accepted: 03 June 2020
Published: 26 June 2020

Citation:
Li X, Griffin K, Langeveld S, Frommhagen M, Underlin EN, Kabel MA, de Vries RP and Dilokpimol A (2020) Functional Validation of Two Fungal Subfamilies in Carbohydrate Esterase Family 1 by Biochemical Characterization of Esterases From Uncharacterized Branches. Front. Bioeng. Biotechnol. 8:694. doi: 10.3389/fbioe.2020.00694

The fungal members of Carbohydrate Esterase family 1 (CE1) from the CAZy database include both acetyl xylan esterases (AXEs) and feruloyl esterases (FAEs). AXEs and FAEs are essential auxiliary enzymes to unlock the full potential of feedstock. They are being used in many biotechnology applications including food and feed, pulp and paper, and biomass valorization. AXEs catalyze the hydrolysis of acetyl group from xylan, while FAEs release ferulic and other hydroxycinnamic acids from xylan and pectin. Previously, we reported a phylogenetic analysis for the fungal members of CE1, establishing five subfamilies (CE1_SF1–SF5). Currently, the characterized AXEs are in the subfamily CE1_SF1, whereas CE1_SF2 contains mainly characterized FAEs. These two subfamilies are more related to each other than to the other subfamilies and are predicted to have evolved from a common ancestor, but target substrates with a different molecular structure. In this study, four ascomycete enzymes from CE1_SF1 and SF2 were heterologously produced in *Pichia pastoris* and characterized with respect to their biochemical properties and substrate preference toward different model and plant biomass substrates. The selected enzymes from CE1_SF1 only exhibited AXE activity, whereas the one from CE1_SF2 possessed dual FAE/AXE activity. This dual activity enzyme also showed broad substrate specificity toward model substrates for FAE activity and efficiently released both acetic acid and ferulic acid (∼50%) from wheat arabinoxylan and wheat bran which was pre-treated with a commercial xylanase. These fungal AXEs and FAEs also showed promising biochemical properties, e.g., high stability over a wide pH range and retaining more than 80% of their residual activity at pH 6.0–9.0. These newly characterized fungal AXEs and FAEs from CE1 have high potential for biotechnological applications. In particular as an additional ingredient for enzyme cocktails to remove the ester-linked decorations which enables access for the backbone degrading enzymes. Among these novel enzymes, the dual FAE/AXE activity enzyme also supports the evolutionary relationship of CE1_SF1 and SF2.

Keywords: carbohydrate esterase, CAZy subfamilies, feruloyl esterase, acetyl xylan esterase, fungi, plant biomass degradation, ferulic acid

INTRODUCTION

Over 5000 million tons of agro-food industrial side streams, such as wheat straw, rice straw, corn stover, potato peelings, and sugarcane bagasse, are produced from the agro-industry annually (Chandra et al., 2012). This plant biomass contains mainly polysaccharides, i.e., cellulose, hemicellulose and pectin, and the aromatic polymer lignin, which form a network of the plant cell wall and, therefore, provide protection against pathogens and pests (Ithal et al., 2007; Liu et al., 2018). Xylan is a major component of hemicellulose from agro-food industrial side streams, accounting for 20–30% in secondary cell wall of dicots and up to 50% in commelinid monocots (Scheller and Ulvskov, 2010). Xylan composes of a β-1,4-xylosyl backbone with different substituents, e.g., L-arabinose, D-galactose, D-(methyl)glucuronic acid, ferulic acid, and acetic acid. The acetyl groups are linked to the *O*-2 and/or *O*-3 position of the xylosyl units in commelinid monocots (cereals and grasses) (Puls, 1997; Appeldoorn et al., 2010, 2013), whereas feruloyl and to a lesser degree *p*-coumaryl groups are esterified mainly at the *O*-5 position on arabinosyl residues of xylan of commelinid monocots (Smith and Hartley, 1983; Saulnier et al., 1995; Wende and Fry, 1997; Schendel et al., 2016). Feruloyl groups are also present in pectin, which esterified to the *O*-2 and/or *O*-5 positions of arabinan side chains and to the *O*-6 position of D-galactosyl residues in (arabino-)galactan of rhamnogalacturonan I (Carpita, 1996; Levigne et al., 2004; Harris, 2005; Harris and Trethewey, 2010). Acetyl xylan esterases (AXEs, EC 3.1.1.72) catalyze the hydrolysis of ester linkages between acetyl groups and xylan, whereas feruloyl esterases (FAEs, EC 3.1.1.73) hydrolyze the ester linkages between hydroxycinnamic acids and plant cell-wall polysaccharides (Christov and Prior, 1993).

The Carbohydrate Active enZyme (CAZy) database classifies carbohydrate esterases into 17 families (CE1–CE17) (Lombard et al., 2013). The characterized fungal AXEs belong to CE1-CE6 and CE16 families, with the majority of the characterized enzymes in CE1. Part of the characterized fungal FAEs are also assigned to CE1, whereas the other FAE subfamilies are not classified in this database (Dilokpimol et al., 2016). Recently, a phylogenetic classification of the fungal members of CE1 was established, which divided the fungal CE1 members into five subfamilies (CE1_SF1-SF5) (Mäkelä et al., 2018). The characterized AXEs were grouped in CE1_SF1 and the characterized FAEs were in CE1_SF2 and SF5, whereas none of the members from CE1_SF3 and SF4 has been characterized so far. Among these subfamilies, CE1_SF1 and SF2 are more related and originated from the same node, but target substrates with a different molecular structure. In this study, we aimed to explore the differences between CE1_SF1 and SF2 by selecting six ascomycete candidates from the uncharacterized brunches of the CE1_SF1 and SF2 for heterologous production in *Pichia pastoris* and biochemical characterization using different model and plant biomass substrates.

MATERIALS AND METHODS

Materials

Methyl *p*-coumarate, methyl caffeate, methyl ferulate, methyl sinapate, ethyl *p*-coumarate, ethyl ferulate and chlorogenic acid were purchased from Apin Chemicals Limited (Oxon, United Kingdom). Insoluble wheat arabinoxylan (WAX, P-WAXYI, from wheat flour), endo-α-(1→5)-arabinanase (E-EARAB, GH43 from *Aspergillus niger*), and endo-β-(1→4)-galactanase (E-EGALN, GH53 from *A. niger*) were from Megazyme (Wicklow, Ireland). Corn oligosaccharides mix (COS) was provided by Wageningen University (Appeldoorn et al., 2013). Wheat bran (WB) was from Wageningen Mill (Wageningen, Netherlands). Corn xylooligosaccharides mix (CX) was from Carl Roth GmbH + Co. KG (Karlsruhe, Germany). Sugar beet pectin (SBP, Pectin Betapec RU 301) was from Herbstreith & Fox KG (Neuenbürg, Germany). Xylanase (GH11 from *Thermomyces lanuginosus*), *p*-methylumbelliferyl acetate (MUB-acetate) and other chemicals were purchased from Sigma-Aldrich (Merck KGaA, Darmstadt, Germany).

Bioinformatics

All amino acid sequences in this study were obtained from JGI Mycocosm[1] (Grigoriev et al., 2014) and the CAZy database[2] (Lombard et al., 2013). The secretory signal peptides were detected using the SignalP 4.1[3] (Petersen et al., 2011). The amino acid sequences without predicted signal peptides were aligned using Multiple Alignment using Fast Fourier Transform (MAFFT)[4] (Katoh and Standley, 2013) and visualized using Easy Sequencing in Postscript[5] (Robert and Gouet, 2014). The phylogenetic analysis was performed using maximum likelihood (ML), neighbor-joining (NJ), and minimal evolution (ME) implemented in the Molecular Evolutionary Genetic Analysis software version 7 (MEGA7)[6] (Kumar et al., 2016) with 95% partial deletion and the Poisson correction distance of substitution rates. Statistical support for phylogenetic grouping was estimated by 500 bootstrap re-samplings. The final phylogenetic tree was shown using the ML tree with above 40% bootstrap next to the branches. *N*-glycosylation sites were predicted using NetNGlyc 1.0[7] (Blom et al., 2004). The theoretical molecular masses were calculated from the amino acid sequences without signal peptide using Sequence Manipulation Suite online tool[8] (Kearse et al., 2012).

Cloning and Transformation to *Pichia pastoris*

The selected candidate genes (JGI accession numbers are provided in **Table 1**) without predicted signal peptide

[1]https://genome.jgi.doe.gov/mycocosm/home
[2]http://www.cazy.org/
[3]http://www.cbs.dtu.dk/services/SignalP/
[4]https://mafft.cbrc.jp/alignment/server/
[5]http://espript.ibcp.fr/ESPript/ESPript/
[6]https://www.megasoftware.net/
[7]http://www.cbs.dtu.dk/services/NetNGlyc/
[8]https://www.bioinformatics.org/sms2/protein_mw.html

and introns were codon optimized and synthesized into pPicZαA plasmid for production in *P. pastoris* (Genscript Biotech, Leiden, Netherlands). The pPicZαA containing synthetic genes were transformed into *Escherichia coli* DH5α for propagation and sequencing. Then, the plasmids were extracted, linearized by *PmeI* (Promega, Madison, WI, United States), and transformed into *P. pastoris* strain X-33 (Invitrogen, Thermo Fisher Scientific, Carlsbad, CA, United States) according to the manufacturer's recommendation (Dilokpimol et al., 2017). The positive colonies were selected for the highest protein production based on colony Western Blot using anti Histidine-tag antibody conjugated with alkaline phosphatase (Thermo Fisher Scientific). The transformants were grown on a nitrocellulose membrane (0.45 μm; Whatman, GE Healthcare Life Sciences, Buckinghamshire, United Kingdom) over minimal medium (MM, 1.34% yeast nitrogen base, 4×10^{-5}% biotin, 1% v/v methanol and 1.5% agar) for 2–4 days at 30°C. Then the membrane was washed with milli-Q water, blocked with 2–5% skim milk in phosphate-buffered saline, and blotted with anti Histidine-tag antibody. The signal was detected using 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium system.

Production of Recombinant Proteins

The selected *P. pastoris* transformants were grown according to Knoch et al. (2013) in a buffered complex glycerol medium (BMGY, 1% yeast extract, 2% peptone, 0.1 M potassium phosphate buffer pH 6.0, 1% w/v glycerol, 1.34% yeast nitrogen base, 4×10^{-5}% biotin) overnight at 30°C, 250 rpm. Induction was done in buffered complex methanol medium (BMMY, BMGY without glycerol) at 22°C, 250 rpm with methanol supplemented (1% v/v) every 24 h for 96 h. Culture supernatants were harvested (4000 × *g*, 4°C, 20 min), filtered (0.22 mm; Merck), aliquoted and stored at −20°C prior further analysis. The stability of the recombinant enzymes was assessed by their activity (see below) each time after thawing. The enzyme activity remained more than 95% of their original activity while performing the experiment (data not shown).

Biochemical Properties of Recombinant Proteins

Molecular masses of the recombinant proteins were estimated by SDS-PAGE (12% w/v, sodium dodecyl sulfate-polyacrylamide gel) using Mini-PROTEAN® Tetra Cell (Bio-Rad, Hercules, CA, United States). Deglycosylation was performed by incubating 20 μL of *P. pastoris* culture supernatant with endoglycosidase H (New England Biolabs, Ipswich, MA, United States) as recommended by the manufacturer. Protein concentration was assessed from SDS-PAGE gel by densitometric method using ImageJ program (Schneider et al., 2012) and 0.5–2.0 μg Bovine Serum Albumin (Pierce, Thermo Scientific, Carlsbad, CA, United States) as a standard.

Enzyme Activity Assays

p-Methylumbelliferyl Substrate Assay

Activity toward *p*-methylumbelliferyl acetate (MUB-acetate) was performed in 100 μL reaction mixtures containing 10 μM MUB- acetate (dissolved in acetone), 100 mM phosphate buffer (pH 6.0) and 10 μL culture supernatant. The reaction was performed at 30°C. The release of umbelliferone group was spectrophotometrically quantified by following the excitation/emission at 340/520 nm up to 30 min with a 2 min interval. The activity of the enzymes was determined by quantification of the umbelliferone group using a standard curve (0.5–100 μM). One unit (U) of AXE activity was defined as the amount of enzyme which released 1 μmol of umbelliferone group from MUB-acetate per min under the assay condition. Culture filtrate from *P. pastoris* harboring pPicZαA plasmid without insertion was used as negative control. All assays were performed in triplicate.

Hydroxycinnamate Substrate Assay

Activity of the CE1_SF1 and SF2 candidates toward hydroxycinnamate substrates (methyl *p*-coumarate, methyl caffeate, methyl ferulate, methyl sinapate, ethyl *p*-coumarate, ethyl ferulate and chlorogenic acid) was performed in 250 μL reaction mixtures containing 0.12 mM substrate (dissolved in dimethylformamide), 80 mM phosphate buffer (pH 6.0) and 50 μL culture supernatant. The decrease of substrate concentration was spectrophotometrically quantified by following the absorbance at 340 nm up to 30 min with a 2 min interval. The activity of the enzymes was determined by quantification of the substrate concentration using standard curves in a range between 0.005 and 0.25 mM. One unit (U) of FAE activity was defined as the amount of enzyme which decreased 1 μmol of substrates per min under the assay condition. Culture filtrate from *P. pastoris* harboring pPicZαA plasmid without insertion was used as negative control. All assays were performed in triplicate.

pH and Temperature Profiles

p-methylumbelliferyl-acetate and methyl ferulate were used as substrates for AXE and FAE activity, respectively, based on the above assay. pH profiles of enzyme activity were determined at 30°C in 100 mM Britton-Robinson buffer (Britton and Robinson, 1931) (pH 2.0–9.0). The temperature profiles of enzyme activity were measured in 100 mM phosphate buffer pH 6.0 at 22–80°C. The pH and thermal stability of the enzymes were determined by measuring the residual enzyme activities after 2 or 16 h incubation at 30°C in 100 mM Britton-Robinson buffer pH 2.0–10.0, or after 2 h incubation at 22 to 80°C in phosphate buffer pH 6.0 (Dilokpimol et al., 2011).

Hydrolytic Activity Toward (poly-/oligo-)saccharides

The activity toward plant biomass was determined using insoluble wheat arabinoxylan (WAX) and corn oligosaccharides mix (COS) for AXE activity, and WAX, wheat bran (WB), corn xylooligosaccharides mix (CX), and sugar beet pectin (SBP) for FAE activity. For pre-treatment, 1% of WAX, COS, WB, and CX

TABLE 1 | Molecular mass and properties of characterized CE1_SF1 and SF2 and the selected enzymes in this study.

SF	Fungal species	GenBank or JGI Accession number	Enzyme Name	Calculated molecular mass (kDa)	Apparent molecular mass (kDa)[a,b]	pH optimum[a]	pH stability[a]	Temp. optimum (°C)[a]	Temp. stability (°C)[a]	References
1.1	*Aspergillus luchuensis*[c]	BAA13434.2	*Al*AXEA (*Aw*AXE)	33	30	7.0	6.0–9.0	N.A.	40	(Koseki et al., 1997; Komiya et al., 2017)
1.1	*Aspergillus niger*	CAA01634.1	AxeA	33	N.A.	N.A.	N.A.	N.A.	N.A.	(De Graaff et al., 1992)
1.1	*Aspergillus ficuum*	AAK60128.1	*Af*AXE	33	34	7.0	N.A.	37	50	(Chung et al., 2002)
1.1	*Aspergillus oryzae*	jgi| Aspor1| 10625	*Ao*AXE	34	30	6.0	6.0–7.0	45	40–50	(Koseki et al., 2006)
1.1	*Aspergillus nidulans*	jgi| Aspnid1| 4188	*An*AXE	34	N.A.	N.A.	N.A.	N.A.	N.A.	(Bauer et al., 2006)
1.1	*Thermothelomyces thermophilus*[d]	ADZ98864.1	*Mt*Axe3	35	N.A.	7.0	N.A.	35–45	N.A.	(Pouvreau et al., 2011)
1.1	*Rasamsonia emersonii*	ADX07526.1	*Te*CE1	30	33	N.A.	N.A.	N.A.	N.A.	(Neumüller et al., 2015)
1.2	***Podospora anserina* S mat+**	**jgi| Podan2| 8904**	**AxeA**	**32**	**35**	**7.0**	**6.0–9.0[f]**	**40**	**40[f]**	**This study**
1.3	***Parastagonospora nodorum* SN15[e]**	**jgi| Stano2| 8578**	**Axe1**	**35**	**35**	**7.0**	**6.0–9.0[f]**	**30**	**40[f]**	**This study**
1.3	***Podospora anserina* S mat+**	**jgi| Podan2| 6933**	**AxeB**	**45**	**70**	**6.0**	**6.0–9.0[f]**	**40**	**30[f]**	**This study**
1.4	*Volvariella volvocea*	jgi| Volvo1| 112494	*Vv*AXE1	38	45	8.0	N.A.	60.0	40–50	(Ding et al., 2007)
1.4	*Phanerochaete chrysosporium*	AEX99751.1	*Pc*Axe2	40	55	7.0	N.A.	30–35	N.A.	(Huy et al., 2013)
2.1	***Podospora anserina* S mat+**	**jgi| Podan2| 8825**	**FaeD**	**33**	**35**	**7.0**	**3.0–10.0[f]**	**50**	**40[f]**	**This study**
2.2	*Penicillium funiculosum*	CAC14144.1	FAEB	39	35	N.A.	N.A.	N.A.	N.A.	(Kroon et al., 2000)
2.2	*Aspergillus sydowii*	jgi| Aspsy1| 1158585	*As*FaeE	33	32	N.A.	N.A.	N.A.	N.A.	(Dilokpimol et al., 2018)
2.2	*Neurospora crassa*	jgi| Neucr2| 3766	*Nc*Fae-1	32	29	6.0	6.0–7.5	55	N.A.	(Crepin et al., 2003)
2.2	*Chaetomium sp.*	AFU88756.1	*Chae*Fae	33	30	7.5	4.0–10.0	60	55	(Yang et al., 2013)
2.2	*Thermothelomyces thermophilus*[d]	AEO62008.1	*Mt*Fae1a	33	30	7.0	7.0–10.0	50	55	(Topakas et al., 2012)
2.2	*Chrysosporium lucknowense*	AEP33618.1	*Cl*FaeB2	33	33	7.0	N.A.	45	N.A.	(Kühnel et al., 2012)

[a] N.A., not available. [b] Apparent mass after deglycosylation by Endoglycosidase H. [c] Formerly Aspergillus awamori. [d] Formerly Myceliophthora thermophila. [e] Formerly Stagonospora/Septoria/nodorum. [f] Retaining more than 80% residual activity after incubating at 30°C or at pH 6.0 for pH or temperature stability, respectively. Enzymes in bold are used in this study.

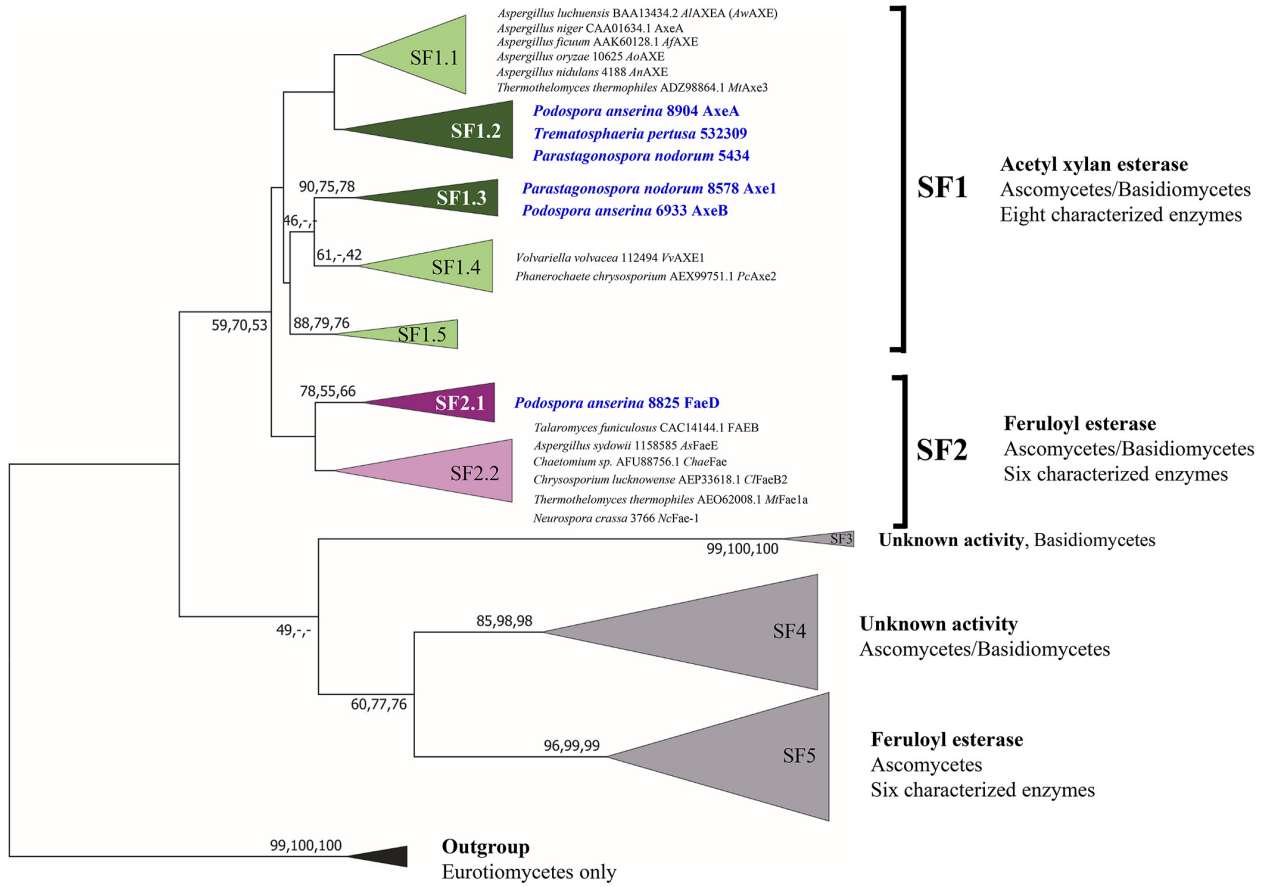


FIGURE 1 | Reconstructed phylogenetic tree of fungal CE1 members [modified from Mäkelä et al. (2018)]. The phylogenetic analysis was performed by maximum likelihood (ML) implemented in MEGA7 (Kumar et al., 2016) with 95% partial deletion of gaps and the Poisson correction distance of substitution rates. The main branches/subfamilies were collapsed. Statistical support for phylogenetic grouping was estimated by 500 bootstrap re-samplings, only the bootstrap above 40% were shown on the branches, also of a neighbor-joining (NJ) and minimal evolution (ME) tree using the same dataset (order: ML, NJ, ME). Eight FAEs from subfamily 7 (Dilokpimol et al., 2016) were used as an outgroup. Names in blue and bold indicates the selected candidates for this study. SF, subfamily. The full phylogenetic tree can be found in **Supplementary Figure 1**.

was incubated with 0.1 mg xylanase, or 0.1 mg of arabinanase and galactanase for SBP, in a 50 mM sodium acetate buffer pH 4.5 and 0.02% sodium azide at 100 rpm and 30°C for 72 h. The reaction was stopped by heat inactivation at 95°C for 10 min (Dilokpimol et al., 2017). For hydrolytic activity assay, the reaction containing 500 μL 1% (pre-treated) substrate and 100 μL culture supernatant (containing 1 μg enzyme) was incubated at 30°C, 24 h, 100 rpm. Except for COS reactions that were stopped by the addition of 50 μL 2 M HCl, the others were inactivated by heating at 95°C, 10 min and centrifuged at 14,000 rpm, 4°C for 15 min. The supernatants were used for HPLC analysis.

Acetic Acid Content Analysis

The acetic acid release was measured by using HPLC (Dionex ICS-3000 chromatography system; Thermo Scientific, Sunnyvale, CA, United States) equipped with an Aminex HPX-87H column with a guard-column (300 × 7.8 mm; Bio-Rad) and a refractive index detector (Bio-Rad). An isocratic elution of 5.0 mM sulfuric acid, with a flow rate of 0.6 mL/min at 40°C was used according to Zeppa et al. (2001). The release of acetic acid was quantified using standards in a range between 0.01 and 2.0 mg/mL.

Hydroxycinnamic Acid Content Analysis

The reaction supernatants were mixed with 100% acetonitrile (1:3, v/v), then incubated on ice for 10 min and centrifuged for 15 min at 4°C to remove precipitants prior to the analysis. The ferulic release was monitored by using HPLC (Dionex ICS-5000+ chromatography system; Thermo Scientific, Sunnyvale, CA) equipped with an Acclaim Mixed-Mode WAX-1 LC Column (3 × 150 mm; Thermo Scientific) and a UV detector (310 nm; Thermo Scientific). The chromatographic separation was carried out according to Dilokpimol et al. (2017) using an isocratic elution of 25 mM potassium phosphate buffer, 0.8 mM pyrophosphate, pH 6.0 in 50% (v/v) acetonitrile with a flow rate of 0.25 mL/min at 30°C. Ferulic and *p*-coumaric acids (0.25–50 μM) were used as standards for identification and quantification.

RESULTS

Production of Four New Fungal CE1 Enzymes

Previous phylogenetic tree for fungal CE1 was proposed which classified the members of CE1 into five subfamilies (Mäkelä et al., 2018). Using a genome-mining strategy by BLAST searching against 18 genome sequenced fungal species (Dilokpimol et al., 2018), we added 87 additional CE1 sequences to the previous CE1 tree (Mäkelä et al., 2018), and reconstructed the phylogenetic tree (**Figure 1**, **Supplementary Figure 1**, and **Supplementary Table 1**). The reconstructed CE1 tree is agreed with the previous one. CE1_SF1 could be further separated into five branches (CE1_SF1.1–SF1.5), while CE1_SF2 was divided into two branches (CE1_SF2.1–SF2.2). Fungal CE1_SF1 contains nine biochemically characterized AXEs from CE1_SF1.1 and SF1.4, while fungal CE1_SF2 contains six biochemically characterized FAEs that were all from CE1_SF2.2 (**Figure 1** and **Table 1**). It is possible that the uncharacterized branches contain enzymes with different activity. To systematically evaluate the subfamilies and compare the activities and biochemical properties, six fungal candidates from the uncharacterized branches of these two subfamilies (three from CE1_SF1.2, two from CE1_SF1.3 and one from CE1_SF2.1) were selected for recombinant protein production in *P. pastoris*.

Of these six candidates, only four candidate enzymes, i.e., *Podospora anserina* AxeA, AxeB, *Parastagonospora nodorum* Axe1, and *Po. anserina* FaeD, were detected by SDS-PAGE (**Supplementary Figure 2**). The apparent masses of *Pa. nodorum* Axe1, *Po. anserina* AxeB, and *Po. anserina* FaeD were 40, 100, and 45 kDa, respectively, whereas *Po. anserina* AxeA showed smears. After deglycosylation using endoglycosidase H, the molecular masses of *Po. anserina* AxeA, *Pa. nodorum,* Axe1, and *Po. anserina* FaeD decreased to 35, 35, and 35 kDa, respectively, which corresponded with their calculated masses based on the amino acid sequence. In contrast, the deglycosylation of *Po. anserina* AxeB resulted in a molecular mass of about 70 kDa, which was still higher than the calculated molecular mass. It is possible that the glycosylation sites of *Po. anserina* AxeB were inaccessible for endoglycosidase H or that other post-translational modifications were present.

All Four New Enzymes Have High Alkaline Tolerance

Podospora anserina AxeA, AxeB, *Pa. nodorum* Axe1 were active toward MUB-substrates, while *Po. anserina* FaeD showed a broad substrate range. It hydrolyzed all tested substrates and showed the highest activity toward methyl *p*-coumarate (141.5 ± 3.5 U/mg) (**Table 2**).

Based on MUB-acetate, *Po. anserina* AxeA, AxeB, *Pa. nodorum* Axe1 and *Po. anserina* FaeD showed the highest activity at pH 7.0, 6.0, 7.0, and 7.0, respectively, when incubated at 30°C (**Figure 2A**). They showed the highest activity at 40, 40, 50, and 30°C, respectively, when incubated at pH 6.0 (**Figure 2B**). The enzymes showed stability over a wide pH range, retaining more than 80% residual activity after incubation at pH 6.0–9.0 for 2 h (**Figure 2C**). In addition, *Pa. nodorum* Axe1 and *Po. anserina* FaeD retained more than 85% residual activity between pH 3.0 and 10.0. The stability of *Po. anserina* AxeA and *Pa. nodorum* Axe1 was higher, since both enzymes retained 55 and 90% of their residual activity at pH 2, compared to *Po. anserina* AxeB and *Po. anserina* FaeD, which did not show any residual activity after incubation for 16 h (**Figure 2E**). Most enzymes retained more than 80% residual activity after incubation at 40°C for 2 h except for *Po. anserina* AxeB that retained only 55% residual activity (**Figure 2D**).

Furthermore, *Po. anserina* FaeD showed the highest activity toward methyl ferulate at pH 7.0 and 50°C (**Figures 3A,B**). It retained more than 80% residual activity after incubation at pH 4.0–10.0 for 2 h (**Figure 3C**). *Po. anserina* FaeD also retained more than 70% residual activity after incubation at 80°C for 2 h (**Figure 3D**).

Po. anserina FaeD Is a Dual FAE/AXE Activity Enzyme

Wheat arabinoxylan and COS were used as substrates to determine the ability of the enzymes to release acetic acid from plant biomass (**Figure 4**). WAX and COS contain about 1 and 3.2 mg acetic acid per one g polysaccharide, respectively. *Po. anserina* AxeA, *Pa. nodorum* Axe1 and *Po. anserina* FaeD released acetic acid from both WAX and COS, whereas *Po. anserina* AxeB only released acetic acid from WAX. The highest acetic acid release was detected for *Po. anserina* AxeA and *Pa. nodorum* Axe1 from WAX (78%) and COS (85%), respectively. A non-CE1 FAE from *Penicillium subrubescens* FaeA (Dilokpimol et al., 2016) was used for comparing the degrading ability of CE1 enzymes. *P. subrubescens* FaeA showed no release of acetic acid from either WAX or COS.

Wheat arabinoxylan, WB, CX and SBP were used as substrates to verify the ability of the enzymes to release ferulic acid from plant biomass (**Figure 5**), as these substrates contain about 3, 1, 40, and 1.9 mg total ferulic acid per one g polysaccharide, respectively (Dilokpimol et al., 2017). *Po. anserina* FaeD released ferulic acid from all tested substrates, whereas none of the others showed any FAE activity. *Po. anserina* FaeD released the highest amount (64%) of ferulic acid from CX which was pre-treated with xylanase and no ferulic acid was detected when CX was not pre-treated with xylanase. *Po. anserina* FaeD also released a higher amount of ferulic acid from pre-treated WAX and WB with xylanase (54 and 48%, respectively) than those without xylanase pre-treatment (2 and 8%). *P. subrubescens* FaeA released a higher amount of ferulic acid from WAX, WB and CX than *Po. anserina* FaeD, and released all ferulic acid from pre-treated WAX and WB. The incubation of SBP with arabinanase and galactanase did not improve the release of ferulic acid by *Po. anserina* FaeD and *P. subrubescens* FaeA, and both enzymes only released less than 10% ferulic acid from SBP.

DISCUSSION

The CE1 is a diverse family and the fungal CE1 members can be classified into five subfamilies (CE1_SF1-SF5) (**Figure 1**

TABLE 2 | Specific activity of positive candidates toward model substrates.

Fungal species	Enzyme	Concentration	Specific activity (U/mg)[a]							
		(mg/L)	MUB-acetate	Methyl *p*-coumarate	Methyl caffeate	Methyl ferulate	Methyl sinnapate	Ethyl *p*-coumarate	Ethyl ferulate	Chlorogenic acid
Podospora anserina	AxeA	12	30.4 ± 0.4	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Parastagonospora nodorum	AxeB	21	33.4 ± 0.3	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Podospora anserina	Axe1	55	22.1 ± 0.0	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Podospora anserina	FaeD	73	3.1 ± 0.1	141.5 ± 3.5	53.0 ± 0.3	49.8 ± 0.3	50.0 ± 0.1	48.1 ± 2.1	39.2 ± 0.6	22.8 ± 1.4

[a]*N.A., not active.*

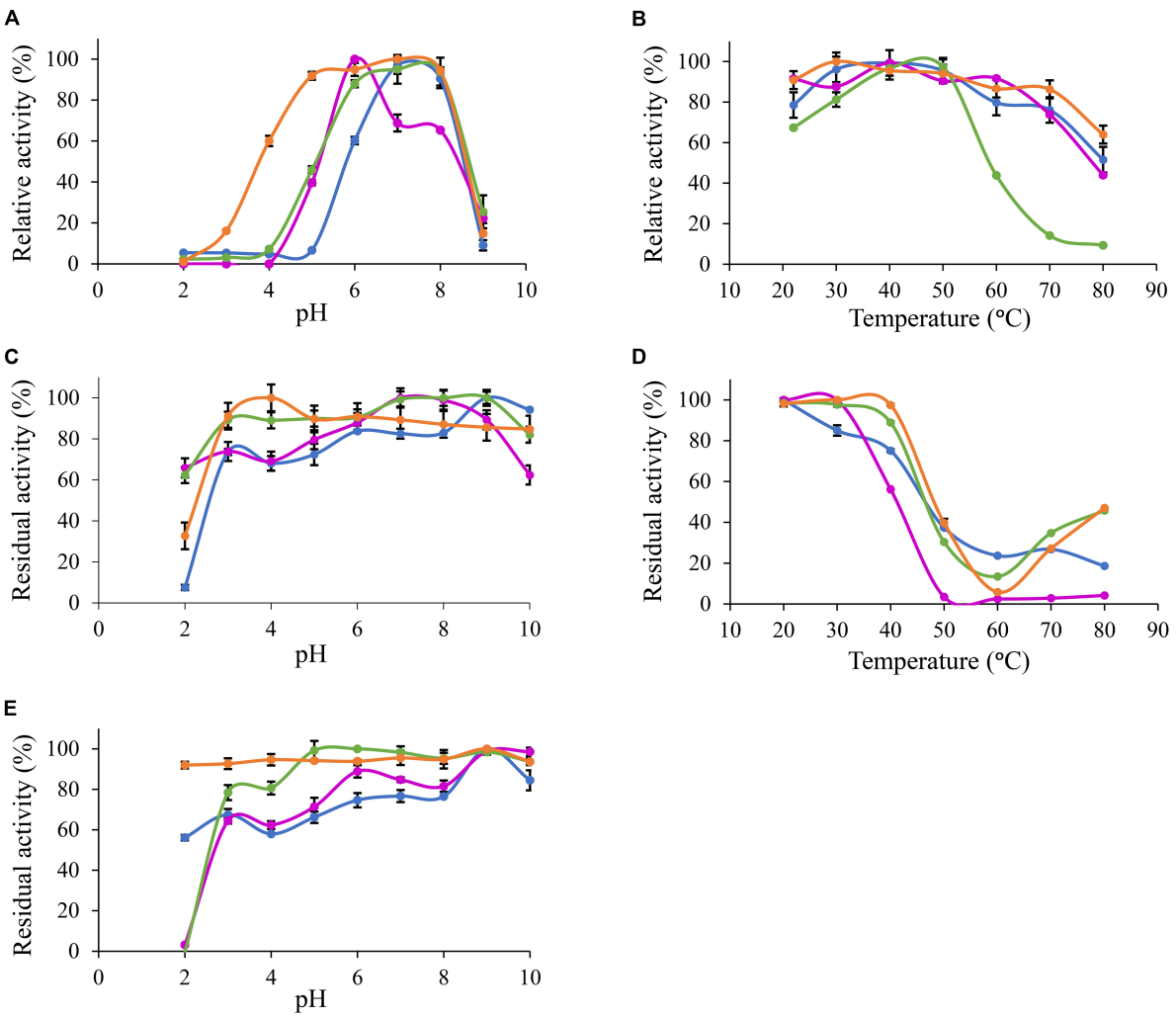


FIGURE 2 | pH and temperature profiles of the selected CE1 enzymes toward *p*-methylumbelliferyl acetate. **(A)** pH optimum, **(B)** Temperature optimum, **(C)** pH stability for 2 h incubation at 30°C, **(D)** Temperature stability for 2 h incubation at pH 6.0, **(E)** pH stability for 16 h incubation at 30°C. ● *Po. anserina* AxeA; ● *Pa. nodorum* Axe1; ● *Po. anserina* AxeB; ● *Po. anserina* FaeD. Each experiment was performed in triplicate. Standard deviations are shown as error bars.

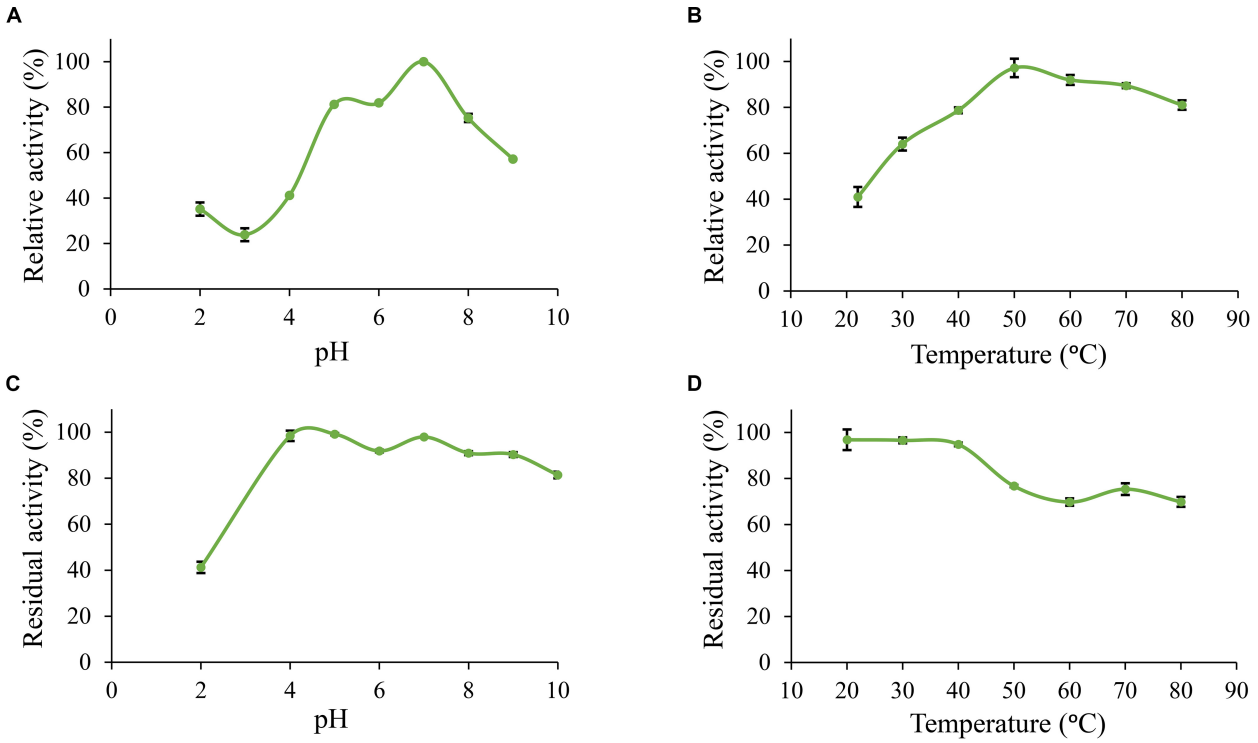


FIGURE 3 | pH and temperature profiles of *Po. anserina* FaeD toward methyl ferulate. **(A)** pH optimum, **(B)** Temperature optimum, **(C)** pH stability for 2 h incubation at 30°C, **(D)** Temperature stability for 2 h incubation at pH 6.0. Each experiment was performed in triplicate. Standard deviations are shown as error bars.

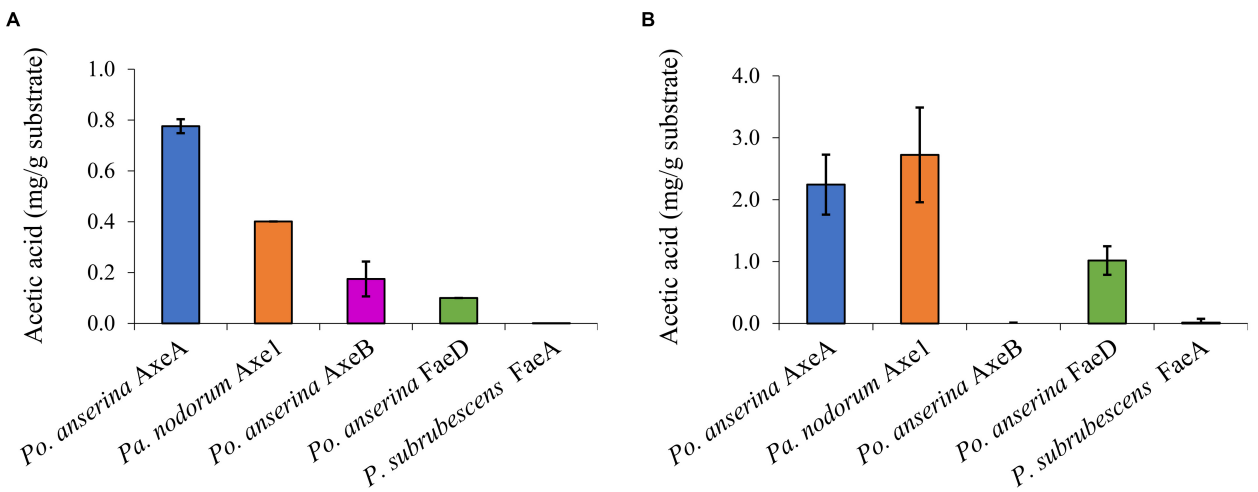


FIGURE 4 | Acetic acid release from wheat arabinoxylan and corn oligosaccharides the selected CE1 enzymes or *P. subrubescens* FaeA. **(A)** Wheat arabinoxylan, **(B)** corn oligosaccharides. Each experiment was performed in triplicate. Standard deviations are shown as error bars.

and **Supplementary Figure 1**). Multiple sequence alignment of members from CE1_SF1 and SF2 showed that these enzymes share an amino acid sequence similarity above 70%, which indicated that both CE1_SF1 and SF2 are more related to each other than to the other subfamilies. The same catalytic triad (Ser/Asp/His), the signature motif (G-X-S-X-G) and the shared common node suggested that these two subfamilies are likely to evolve from the same ancestor (Koseki et al., 1997; Komiya et al., 2017; Mäkelä et al., 2018). However, both subfamilies target substrates with a different molecular structure. So far, only a limited number of fungal CE1 enzymes have been characterized (**Table 1**), which hinders to explain the different activity and evolution relationship between CE1_SF1 and SF2.

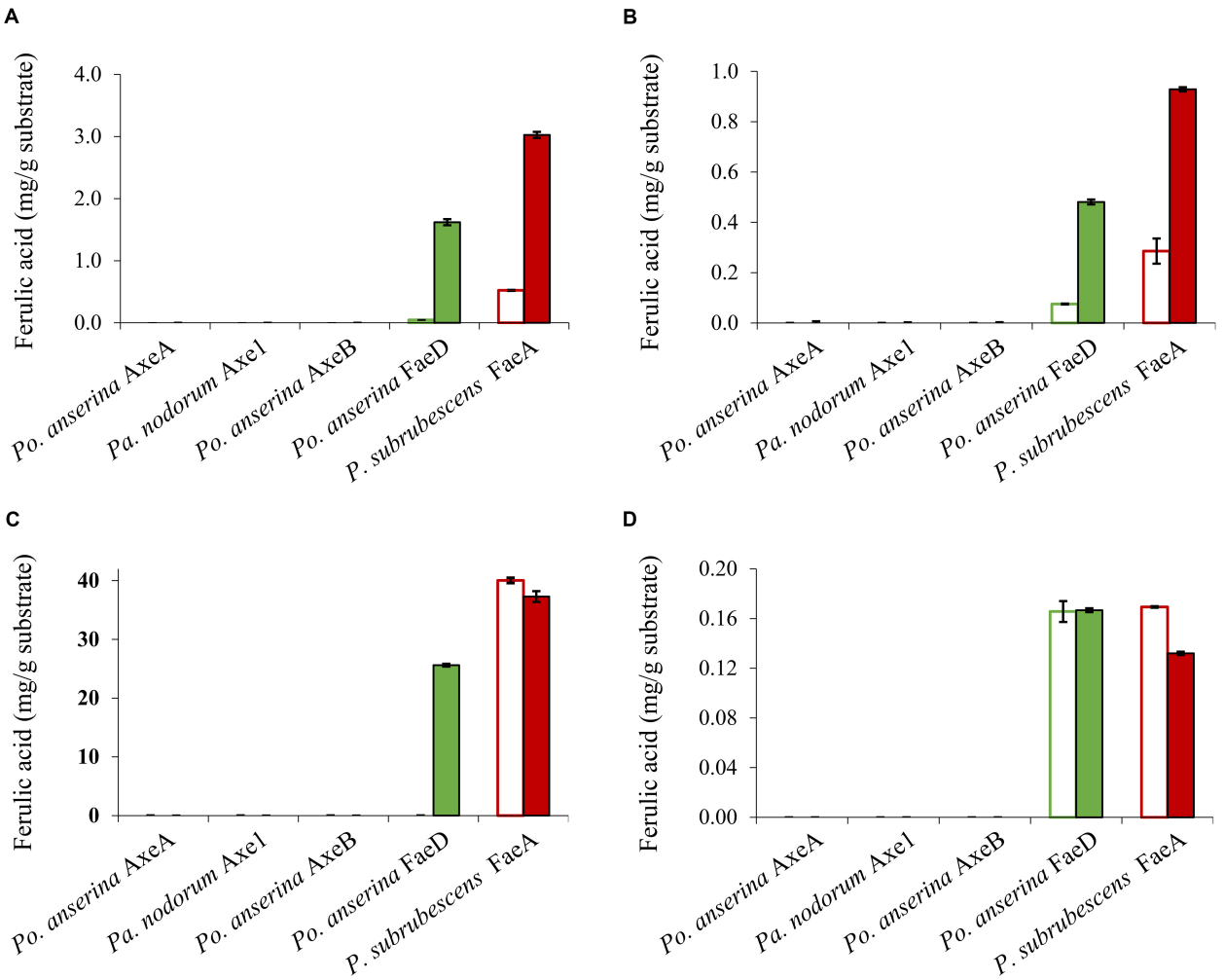


FIGURE 5 | Ferulic acid release from plant biomass using the selected CE1 enzymes or *P. subrubescens* FaeA. **(A)** Wheat arabinoxylan, **(B)** wheat bran, **(C)** corn xylooligosaccharides, **(D)** sugar beet pectin. Empty bars indicate reactions with non-treated plant biomass, filled bars indicate reactions with pre-treated biomass (with xylanase for wheat arabinoxylan, wheat bran, corn xylooligosaccharides or arabinose and galactanase for sugar beet pectin) prior to incubation with CE1 enzymes or *P. subrubescens* FaeA. Each experiment was performed in triplicate. Standard deviations are shown as error bars.

In this study, we selected six ascomycete candidates from the uncharacterized branches in CE1_SF1 and SF2 to further explore the differences between CE1_SF1 and SF2. Four of these (*Po. anserina* AxeA, *Po. anserina* AxeB, *Pa. nodorum* Axe1 and *Po. anserina* FaeD) could be successfully produced in *P. pastoris* and exhibited different properties. Both *Pa. nodorum* and *Po. anserina* belongs to the Phylum Ascomycota, Subphylum Pezizomycotina, while the first is classified in the Class Dothideomycetes, the latter is part of the Class Sordariomycetes. Based on the AXE activity assay, *Po. anserina* AxeA, *Po. anserina* AxeB and *Pa. nodorum* Axe1 from CE1_SF1 showed the highest activity toward MUB-acetate, while *Po. anserina* FaeD from CE1_SF2 showed around 10-fold less activity on this substrate (**Table 2**). Even though all four enzymes showed activity toward AXE substrates, only *Po. anserina* FaeD also de-esterified the FAE substrates with broad substrate preference. Most characterized FAEs from this subfamily showed broad substrate specificity including FaeB, AsFae, MtFae1a, and ClFaeB2, supporting that this is one of the unique features of the FAEs in CE1 (Kroon et al., 2000; Kühnel et al., 2012; Topakas et al., 2012; Dilokpimol et al., 2016, 2018).

All enzymes could release acetic acid from both WAX and COS, except for *Po. anserina* AxeB for which the release of acetic acid was only detected from WAX (**Figure 4**). This could result from the different position of the acetylation in the substrates. Acetylation in commelinid monocots including wheat generally occurs at the *O*-2 and/or *O*-3 position of the xylose residues of the xylan backbone (Puls, 1997; Appeldoorn et al., 2010). It has been shown earlier that AXEs from CE1 regioselectively cleave the substituents in the *O*-2 and *O*-3 position, and deacetylate the *O*-2 position faster than the *O*-3 position (Altaner et al., 2003). However, COS contains more than half of the acetyl group attached to the *O*-2 position of the xylose backbone, while the same xylose is also substituted with a monomeric α-L-arabinosyl residue at the *O*-3 position (Appeldoorn et al., 2013). Because

Po. anserina AxeA and *Pa. nodorum* Axe1 released over 75% of total acetic acid, our results also indicated the potential ability of these enzymes to attack the dense acetyl substitution on the xylan backbone in COS.

In contrast, only *Po. anserina* FaeD could release ferulic acid from the feruloylated substrates (**Figure 5**). As mentioned before, feruloylation in commelinid monocots mainly occurs at the *O*-5 position on arabinosyl residues of xylan (Smith and Hartley, 1983; Saulnier et al., 1995; Wende and Fry, 1997; Schendel et al., 2016), while in pectin it mainly occurs at the *O*-2 and/or *O*-5 positions of arabinan side chains and at the *O*-6 position of D-galactosyl residues in (arabino-)galactan of rhamnogalacturonan I (Carpita, 1996; Levigne et al., 2004; Harris, 2005; Harris and Trethewey, 2010). *Po. anserina* FaeD was shown to release more ferulic acid from commercial xylanase-treated xylan substrates (WAX, WB, and CX) than non-treated ones, from which it released only small amounts of ferulic acid. In contrast, pre-treatment of SBP with arabinanase and galactanase did not improve the release of ferulic acid by *Po. anserina* FaeD and less than 10% of ferulic acid was released from intact and pre-treated pectin. Analysis of ferulic acid release from xylan and SBP substrates indicated that pre-treatment had a larger effect on *Po. anserina* FaeD activity on xylan substrates than SBP. This difference indicated that *Po. anserina* FaeD is mainly active on *O*-5 feruloylated xylooligosaccharides and much less on other substitutions. *Po. anserina* FaeD from CE1_SF2 showed a high preference for xylan substrates, which resembles FAEB from *Penicillium funiculosum* (Kroon et al., 2000) and FaeC from *Aspergillus niger* (Dilokpimol et al., 2017). *P. subrubescens* FaeA, a non-CE1 FAE control, did not release acetic acid from WAX or COS, but could efficiently release ferulic acid from WAX, WB, and CX. Earlier, an FAE from *Aspergillus oryzae* (AoFae) was shown to possess dual activity (Tenkanen et al., 1991). AoFae also belongs to CE1 based on sequence homology of the N-terminal amino acid sequence (Tenkanen et al., 2003). However, based on the hydrolysis pattern toward different monoacetylated and monoferuloylated *p*-nitrophenyl glycosides, AoFae was later suggested to be a non-specific acetyl esterase (Puchart et al., 2006). In the same study, the authors also showed that some FAEs could also release acetyl residue from the same substrates with specific positional specificity, but with at least two orders of magnitude lower (Puchart et al., 2006). This aspect should be validated further for these new enzymes to monitor the de-esterified positions.

Based on the multiple sequence alignment (**Supplementary Figure 3**) between CE1_SF1 and SF2, Trp160 was highly conserved among CE1_SF1 members, whereas the corresponding amino acids in CE1_SF2 can be an Ala, Ser, Pro, Gln or Thr (**Supplementary Table 1**). Recently, a crystal structure of an *Aspergillus luchuensis* (formerly *A. awamori*) *Al*AXEA (PDB code 5X6S) showed Trp160 controls the substrate specificity of AXE in CE1 (Komiya et al., 2017). When replacing Trp160 to Ala, Ser, or Pro, the mutants showed significant FAE activity. Trp160 in *Al*AXEA is corresponding to an Ala in *Po. anserina* FaeD supporting that the expanded substrate specificity of an AXE to FAE and the dual activity is potentially influenced by this amino acid.

To investigate the pH and temperature profile of these recombinant enzymes, we used MUB-acetate as a substrate for AXE activity. All candidates showed optimum pH in a neutral range (pH 6.0 to 7.0) and an optimum temperature between 30 and 50°C, which are similar to most other characterized CE1_AXE enzymes (Koseki et al., 1997, 2006; Chung et al., 2002; Pouvreau et al., 2011; Huy et al., 2013;**Table 1**). Surprisingly, they all showed excellent pH stability over a broad pH range, retaining more than 80% residual activity after an incubation at pH 6.0–9.0 for 2 h, which showed a similar pH stability range to *Aspergillus luchuensis Al*AXEA (Koseki et al., 1997; Komiya et al., 2017). It should be noted that at pH more than 7.0, the ester-linkages are alkali labile and tend to degrade easily. Hence, to determine the pH stability of AXEs, we used culture filtrate from *P. pastoris* harboring pPicZαA plasmid without insertion as negative control, from which the residual activity was deduced. Using methyl ferulate as a substrate for FAE activity, *Po. anserina* FaeD also showed optimum pH at 7.0 and optimum temperature at 50°C, these properties are within the range of other reported fungal FAEs, mainly with activity at pH 4–8 and temperature 30–65°C (**Table 1**; Crepin et al., 2003; Topakas et al., 2007, 2012; Kühnel et al., 2012). It also showed a high pH stability over a broad pH range from pH 3.0 to 10.0, which is more superior than most reported CE1 enzymes (**Table 1**). The CAZymes from *Pa. nodorum* (synonyms: *Stagonospora*/*Septoria*/*Phaeosphaeria* nodorum) have not been characterized in much detail, mainly because it is a major pathogen of wheat and related cereals (Downie et al., 2018). Several CAZymes from *Po. anserina* were previously reported, e.g., *Pa*Man5A and *Pa*Man26A mannanases, *Pa*Xyn11A xylanase, and *Pa*Abf51A and *Pa*Abf62A arabinofuranosidases were active in a range of pH 3–7 and 5–75°C (Couturier et al., 2011), while *Pa*Cel6A-C were active in a range of pH 4–9 and 25–65°C (Poidevin et al., 2013). The biochemical properties of different enzymes from the same strain can be vastly different also depending on the function of the enzyme, which most of the enzymes in this study are quite stable at high pH. The alkaline tolerance of the newly characterized enzymes is of interested for many bio-industrial applications particularly for the alkali pre-treated plant biomass.

CONCLUSION

CE1_SF1 and SF2 are related, even though the characterized enzymes from the first possess AXE activity and the ones from latter possess FAE activity. So far no additional activity was detected in these subfamilies except for the dual activity. A novel dual FAE/AXE activity enzyme was identified from CE1_SF2, which showed promising industrial applications because of its broad substrate specificity. To further understand the functional features and physiological role of individual enzymes, the positional specificity of these new enzymes should be further investigated. The phylogenetic analysis and multiple sequence alignment supported the evolutionary relationship between CE1_SF1 and SF2, and identified possible amino acids that control the AXE or FAE activity between these two subfamilies.

Moreover, these novel fungal AXE and FAE enzymes showed a great alkaline tolerance and can selectively release acetic acid and ferulic acid from different plant-based biomass making them attractive for various biotechnological applications.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

XL, KG, and AD conducted the main experiments, analyzed the data, and drafted the manuscript. MF, EU, and MK contributed to HPLC analysis for acetic acid. XL, KG, and SL contributed to phylogenetic analysis. AD and RV contributed to data interpretation and commented on the manuscript. RV conceived and supervised the overall project. All authors commented on the manuscript.

FUNDING

This project was partly supported by China Scholarship Council (Beijing, China), grant number: 201803250066.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00694/full#supplementary-material

REFERENCES

Altaner, C., Saake, B., Tenkanen, M., Eyzaguirre, J., Faulds, C. B., Biely, P., et al. (2003). Regioselective deacetylation of cellulose acetates by acetyl xylan esterases of different CE-families. *J. Biotechnol.* 105, 95–104. doi: 10.1016/s0168-1656(03)00187-1

Appeldoorn, M. M., de Waard, P., Kabel, M. A., Gruppen, H., and Schols, H. A. (2013). Enzyme resistant feruloylated xylooligomer analogues from thermochemically treated corn fiber contain large side chains, ethyl glycosides and novel sites of acetylation. *Carbohyd. Res.* 381, 33–42. doi: 10.1016/j.carres.2013.08.024

Appeldoorn, M. M., Kabel, M. A., Van Eylen, D., Gruppen, H., and Schols, H. A. (2010). Characterization of oligomeric xylan structures from corn fiber resistant to pretreatment and simultaneous saccharification and fermentation. *J. Agric. Food Chem.* 58, 11294–11301. doi: 10.1021/jf102849x

Bauer, S., Vasu, P., Persson, S., Mort, A. J., and Somerville, C. R. (2006). Development and application of a suite of polysaccharide-degrading enzymes for analyzing plant cell walls. *Proc. Natl. Acad. Sci. U.S.A.* 103, 11417–11422. doi: 10.1073/pnas.0604632103

Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S., and Brunak, S. (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 4, 1633–1649. doi: 10.1002/pmic.200300771

Britton, H. T. S., and Robinson, R. A. (1931). CXCVIII.—Universal buffer solutions and the dissociation constant of veronal. *J. Chem. Soc.* 2, 1456–1462. doi: 10.1039/jr9310001456

Carpita, N. C. (1996). Structure and biogenesis of the cell walls of grasses. *Annu. Rev. Plant Biol.* 47, 445–476. doi: 10.1146/annurev.arplant.47.1.445

Chandra, R., Takeuchi, H., and Hasegawa, T. (2012). Methane production from lignocellulosic agricultural crop wastes: a review in context to second generation of biofuel production. *Renew. Sust. Energ. Rev.* 16, 1462–1476. doi: 10.1016/j.rser.2011.11.035

Christov, L. P., and Prior, B. A. (1993). Esterases of xylan-degrading microorganisms: production, properties, and significance. *Enzyme Microb. Technol.* 15, 460–475. doi: 10.1016/0141-0229(93)90078-g

Chung, H. J., Park, S. M., Kim, H. R., Yang, M. S., and Kim, D. H. (2002). Cloning the gene encoding acetyl xylan esterase from *Aspergillus ficuum* and its expression in *Pichia pastoris*. *Enzyme Microb. Technol.* 31, 384–391. doi: 10.1016/s0141-0229(02)00122-9

Couturier, M., Haon, M., Coutinho, P. M., Henrissat, B., Lesage-Meessen, L., and Berrin, J. G. (2011). *Podospora anserina* hemicellulases potentiate the *Trichoderma reesei* secretome for saccharification of lignocellulosic biomass. *Appl. Environ. Microbiol.* 77, 237–246.

Crepin, V. F., Faulds, C. B., and Connerton, I. F. (2003). A non-modular type B feruloyl esterase from *Neurospora crassa* exhibits concentration-dependent substrate inhibition. *Biochem. J.* 370, 417–427. doi: 10.1042/bj20020917

De Graaff, L., Visser, J., Van den Broeck, H., Strozyk, F., Kormelink, F., and Boonman, J. (1992). *Cloning and Use of Acetyl Xylan Esterase from Fungal Origin*. Eur. Patent. No 0,507,369. Munich: E.P. Office.

Dilokpimol, A., Mäkelä, M. R., Aguilar-Pontes, M. V., Benoit-Gelber, I., Hilden, K. S., and de Vries, R. P. (2016). Diversity of fungal feruloyl esterases: updated phylogenetic classification, properties, and industrial applications. *Biotechnol. Biofuels* 9:231. doi: 10.1186/s13068-016-0651-6

Dilokpimol, A., Mäkelä, M. R., Mansouri, S., Belova, O., Waterstraat, M., Bunzel, M., et al. (2017). Expanding the feruloyl esterase gene family of *Aspergillus niger* by characterization of a feruloyl esterase, FaeC. *New Biotechnol.* 37, 200–209. doi: 10.1016/j.nbt.2017.02.007

Dilokpimol, A., Mäkelä, M. R., Varriale, S., Zhou, M., Cerullo, G., Gidijala, L., et al. (2018). Fungal feruloyl esterases: functional validation of genome mining based enzyme discovery including uncharacterized subfamilies. *New Biotechnol.* 41, 9–14. doi: 10.1016/j.nbt.2017.11.004

Dilokpimol, A., Nakai, H., Gotfredsen, C. H., Baumann, M. J., Nakai, N., Hachem, M. A., et al. (2011). Recombinant production and characterisation of two related GH5 endo-β-1,4-mannanases from *Aspergillus nidulans* FGSC A4 showing distinctly different transglycosylation capacity. *BBA Proteins Proteom.* 1814, 1720–1729. doi: 10.1016/j.bbapap.2011.08.003

Ding, S. J., Cao, J., Zhou, R., and Zheng, F. (2007). Molecular cloning, and characterization of a modular acetyl xylan esterase from the edible straw mushroom *Volvariella volvacea*. *FEMS. Microbiol. Lett.* 274, 304–310. doi: 10.1111/j.1574-6968.2007.00844.x

Downie, R. C., Bouvet, L., Furuki, E., Gosman, N., Gardner, K. A., Mackay, I. J., et al. (2018). Assessing European wheat sensitivities to *Parastagonospora nodorum* necrotrophic effectors and fine-mapping the Snn3-B1 locus conferring sensitivity to the effector SnTox3. *Front. Plant Sci.* 9:881. doi: 10.3389/fpls.2018.00881

Grigoriev, I. V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R., et al. (2014). MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res.* 42, D699–D704.

Harris, P. J. (2005). "Diversity in plant cell walls," in *Plant Diversity and Evolution: Genotypic and Phenotypic Variation in Higher Plants*, ed. R. J. Henry, (Cambridge: CABI), 201–227. doi: 10.1079/9780851999043.0201

Harris, P. J., and Trethewey, J. A. (2010). The distribution of ester-linked ferulic acid in the cell walls of angiosperms. *Phytochem. Rev.* 9, 19–33. doi: 10.1007/s11101-009-9146-4

Huy, N. D., Thiyagarajan, S., Kim, D. H., and Park, S. M. (2013). Cloning and characterization of a novel bifunctional acetyl xylan esterase with carbohydrate binding module from *Phanerochaete chrysosporium*. *J. Biosic. Bioeng.* 115, 507–513. doi: 10.1016/j.jbiosc.2012.11.018

Ithal, N., Recknor, J., Nettleton, D., Maier, T., Baum, T. J., and Mitchum, M. G. (2007). Developmental transcript profiling of cyst nematode feeding cells in soybean roots. *Mol. Plant. Microbe. Int.* 20, 510–525. doi: 10.1094/mpmi-20-5-0510

Katoh, K., and Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Knoch, E., Dilokpimol, A., Tryfona, T., Poulsen, C. P., Xiong, G., Harholt, J., et al. (2013). A β–glucuronosyltransferase from *Arabidopsis thaliana* involved in biosynthesis of type II arabinogalactan has a role in cell elongation during seedling growth. *Plant. J.* 76, 1016–1029. doi: 10.1111/tpj.12353

Komiya, D., Hori, A., Ishida, T., Igarashi, K., Samejima, M., Koseki, T., et al. (2017). Crystal structure and substrate specificity modification of acetyl xylan esterase from *Aspergillus luchuensis*. *Appl. Environ. Microbiol.* 83, 1251–1271.

Koseki, T., Furuse, S., Iwano, K., Sakai, H., and Matsuzawa, H. (1997). An *Aspergillus awamori* acetylesterase: purification of the enzyme, and cloning and sequencing of the gene. *Biochem. J.* 326, 485–490. doi: 10.1042/bj3260485

Koseki, T., Miwa, Y., Akao, T., Akita, O., and Hashizume, K. (2006). An *Aspergillus oryzae* acetyl xylan esterase: molecular cloning and characteristics of recombinant enzyme expressed in *Pichia pastoris*. *J. Biotechnol.* 121, 381–389. doi: 10.1016/j.jbiotec.2005.07.015

Kroon, P. A., Williamson, G., Fish, N. M., Archer, D. B., and Belshaw, N. J. (2000). A modular esterase from *Penicillium funiculosum* which releases ferulic acid from plant cell walls and binds crystalline cellulose contains a carbohydrate binding module. *Eur. J. Biochem.* 267, 6740–6752. doi: 10.1046/j.1432-1033.2000.01742.x

Kühnel, S., Pouvreau, L., Appeldoorn, M., Hinz, S., Schols, H., and Gruppen, H. (2012). The ferulic acid esterases of *Chrysosporium lucknowense* C1: purification, characterization and their potential application in biorefinery. *Enzyme Microb. Technol.* 50, 77–85. doi: 10.1016/j.enzmictec.2011.09.008

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Levigne, S. V., Ralet, M.-C. J., Quéméner, B. C., Pollet, B. N.-L., Lapierre, C., and Thibault, J.-F. J. (2004). Isolation from sugar beet cell walls of arabinan oligosaccharides esterified by two ferulic acid monomers. *Plant Physiol.* 134, 1173–1180. doi: 10.1104/pp.103.035311

Liu, Q., Luo, L., and Zheng, L. (2018). Lignins: biosynthesis and biological functions in plants. *Int. J. Mol. Sci.* 19:E335.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2013). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495.

Mäkelä, M. R., Dilokpimol, A., Koskela, S. M., Kuuskeri, J., de Vries, R. P., and Hildén, K. (2018). Characterization of a feruloyl esterase from *Aspergillus terreus* facilitates the division of fungal enzymes from Carbohydrate Esterase family 1 of the carbohydrate-active enzymes (CAZy) database. *Microb. Biotechnol.* 11, 869–880. doi: 10.1111/1751-7915.13273

Neumüller, K. G., de Souza, A. C., van Rijn, J. H., Streekstra, H., Gruppen, H., and Schols, H. A. (2015). Positional preferences of acetyl esterases from different CE families towards acetylated 4-O-methyl glucuronic acid-substituted xylo-oligosaccharides. *Biotechnol. Biofuels.* 8:7. doi: 10.1186/s13068-014-0187-6

Petersen, T. N., Brunak, S., Von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786. doi: 10.1038/nmeth.1701

Poidevin, L., Feliu, J., Doan, A., Berrin, J. G., Bey, M., Coutinho, P. M., et al. (2013). Insights into exo- and endoglucanase activities of family 6 glycoside hydrolases from *Podospora anserina*. *Appl. Environ. Microbiol.* 79, 4220–4229. doi: 10.1128/aem.00327-13

Pouvreau, L., Jonathan, M. C., Kabel, M. A., Hinz, S. W. A., Gruppen, H., and Schols, H. A. (2011). Characterization and mode of action of two acetyl xylan esterases from *Chrysosporium lucknowense* C1 active towards acetylated xylans. *Enzyme Microb. Technol.* 49, 312–320. doi: 10.1016/j.enzmictec.2011.05.010

Puchart, V., Vrsanská, M., Mastihubová, M., Topakas, E., Vafiadi, C., Faulds, C. B., et al. (2006). Substrate and positional specificity of feruloyl esterases for monoferuloylated and monoacetylated 4-nitrophenyl glycosides. *J. Biotechnol.* 127, 235–243. doi: 10.1016/j.jbiotec.2006.06.020

Puls, J. (1997). Chemistry and biochemistry of hemicelluloses: relationship between hemicellulose structure and enzymes required for hydrolysis. *Macromol. Symp.* 120, 183–196. doi: 10.1002/masy.19971200119

Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* 42, W320–W324.

Saulnier, L., Vigouroux, J., and Thibault, J. F. (1995). Isolation and partial characterization of feruloylated oligosaccharides from maize bran. *Carbohyd. Res.* 272, 241–253. doi: 10.1016/0008-6215(95)00053-v

Scheller, H. V., and Ulvskov, P. (2010). Hemicelluloses. *Annu. Rev. Plant Biol.* 61, 263–289.

Schendel, R. R., Puchbauer, A.-K., Britscho, N., and Bunzel, M. (2016). Feruloylated wheat bran arabinoxylans: isolation and characterization of acetylated and *O*-2-monosubstituted structures. *Cereal Chem.* 93, 493–501. doi: 10.1094/cchem-12-15-0250-r

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi: 10.1038/nmeth.2089

Smith, M. M., and Hartley, R. D. (1983). Occurrence and nature of ferulic acid substitution of cell-wall polysaccharides in graminaceous plants. *Carbohyd. Res.* 118, 65–80. doi: 10.1016/0008-6215(83)88036-7

Tenkanen, M., Eyzaguirre, J., Isoniemi, R., Faulds, C. B., and Biely, P. (2003). "Chapter 13: comparison of catalytic properties of acetyl xylan esterases from three carbohydrate esterase families," in *Applications of Enzymes to Lignocellulosics*, eds S. D. Mansfield, and J. N. Saddler, (Washington, DC: American Chemical Society ACS), 211–229. doi: 10.1021/bk-2003-0855.ch013

Tenkanen, M., Schuseil, J., Puls, J., and Poutanen, K. (1991). Production, purification and characterization of an esterase liberating phenolic acids from lignocellulosics. *J. Biotechnol.* 18, 69–83. doi: 10.1016/0168-1656(91)90236-o

Topakas, E., Moukouli, M., Dimarogona, M., and Christakopoulos, P. (2012). Expression, characterization and structural modelling of a feruloyl esterase from the thermophilic fungus *Myceliophthora thermophila*. *Appl. Microbiol. Biotechnol.* 94, 399–411. doi: 10.1007/s00253-011-3612-9

Topakas, E., Vafiadi, C., and Christakopoulos, P. (2007). Microbial production, characterization and applications of feruloyl esterases. *Process Biochem.* 42, 497–509. doi: 10.1016/j.procbio.2007.01.007

Wende, G., and Fry, S. C. (1997). O-feruloylated, O-acetylated oligosaccharides as side-chains of grass xylans. *Phytochemistry* 44, 1011–1018. doi: 10.1016/s0031-9422(96)00648-6

Yang, S.-Q., Tang, L., Yan, Q.-J., Zhou, P., Xu, H.-B., Jiang, Z.-Q., et al. (2013). Biochemical characteristics and gene cloning of a novel thermostable feruloyl esterase from *Chaetomium* sp. *J. Mol. Catal. B. Enzym.* 97, 328–336. doi: 10.1016/j.molcatb.2013.06.011

Zeppa, G., Conterno, L., and Gerbi, V. (2001). Determination of organic acids, sugars, diacetyl, and acetoin in cheese by high-performance liquid chromatography. *J. Agr. Food Chem.* 49, 2722–2726. doi: 10.1021/jf0009403

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 17 July 2020
doi: 10.3389/fbioe.2020.00826

The Secretome of *Phanerochaete chrysosporium* and *Trametes versicolor* Grown in Microcrystalline Cellulose and Use of the Enzymes for Hydrolysis of Lignocellulosic Materials

*Angela S. Machado, Fernanda Valadares, Tatiane F. Silva, Adriane M. F. Milagres, Fernando Segato and André Ferraz**

Departamento de Biotecnologia, Escola de Engenharia de Lorena, Universidade de São Paulo, Lorena, Brazil

OPEN ACCESS

Edited by:
Fabiano Jares Contesini, Technical University of Denmark, Denmark

Reviewed by:
Ana M. R. B. Xavier, University of Aveiro, Portugal
Yinbo Qu, Shandong University, China
Lene Lange, Independent Researcher, Copenhagen, Denmark

***Correspondence:**
André Ferraz
aferraz@debiq.eel.usp.br

Specialty section:
This article was submitted to Bioprocess Engineering, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *28 February 2020*
Accepted: *29 June 2020*
Published: *17 July 2020*

Citation:
Machado AS, Valadares F, Silva TF, Milagres AMF, Segato F and Ferraz A (2020) The Secretome of Phanerochaete chrysosporium and Trametes versicolor Grown in Microcrystalline Cellulose and Use of the Enzymes for Hydrolysis of Lignocellulosic Materials. Front. Bioeng. Biotechnol. 8:826. doi: 10.3389/fbioe.2020.00826

The ability of white-rot fungi to degrade polysaccharides in lignified plant cell walls makes them a suitable reservoir for CAZyme prospects. However, to date, CAZymes from these species are barely studied, which limits their use in the set of choices for biomass conversion in modern biorefineries. The current work joined secretome studies of two representative white-rot fungi, *Phanerochaete chrysosporium* and *Trametes versicolor*, with expression analysis of cellobiohydrolase (CBH) genes, and use of the secretomes to evaluate enzymatic conversion of simple and complex sugarcane-derived substrates. Avicel was used to induce secretion of high levels of CBHs in the extracellular medium. A total of 56 and 58 proteins were identified in cultures of *P. chrysosporium* and *T. versicolor*, respectively, with 78–86% of these proteins corresponding to plant cell wall degrading enzymes (cellulolytic, hemicellulolytic, pectinolytic, esterase, and auxiliary activity). CBHI predominated among the plant cell wall degrading enzymes, corresponding to 47 and 34% of the detected proteins in *P. chrysosporium* and *T. versicolor*, respectively, which confirms that Avicel is an efficient CBH inducer in white-rot fungi. The induction by Avicel of genes encoding CBHs (*cel*) was supported by high expression levels of *cel7D* and *cel7C* in *P. chrysosporium* and *T. versicolor*, respectively. Both white-rot fungi secretomes enabled hydrolysis experiments at 10 FPU/g substrate, despite the varied proportions of CBHs and other enzymes present in each case. When low recalcitrance sugarcane pith was used as a substrate, *P. chrysosporium* and *T. versicolor* secretomes performed similarly to Cellic® CTec2. However, the white-rot fungi secretomes were less efficient than Cellic® CTec2 during hydrolysis of more recalcitrant substrates, such as acid or alkaline sulfite-pretreated sugarcane bagasse, likely because Cellic® CTec2 contains an excess of CBHs compared with the white-rot fungi secretomes. General comparison of the white-rot fungi secretomes

highlighted *T. versicolor* enzymes for providing high glucan conversions, even at lower proportion of CBHs, probably because the other enzymes present in this secretome and CBHs lacking carbohydrate-binding modules compensate for problems associated with unproductive binding to lignin.

Keywords: basidiomycetes, biorefinery, cellobiohydrolases, glucoside hydrolases, secretome, sugarcane, white-rot fungi

INTRODUCTION

Hydrolytic enzymes are major constituents of commercial enzymatic cocktails used in plant biomass conversion of pretreated lignocellulosic materials (Payne et al., 2015; Adsul et al., 2020). The cellulose and hemicellulose hydrolysis promoted by these enzymatic cocktails is the basis of modern biorefineries. However, improvement of the performance of these enzymatic cocktails is still necessary to advance cost effective production of plant biomass derivatives (Ellila et al., 2017; Adsul et al., 2020). To date, most of these enzymatic cocktails are developed for polysaccharide conversion to monosaccharides and contain diverse cellulolytic and hemicellulolytic enzymes, including glucoside hydrolases (GHs) and auxiliary activity enzymes, such as lytic polysaccharide monooxygenases (LPMOs), that cause oxidative cleavage of polysaccharides (Payne et al., 2015; Adsul et al., 2020). Both enzyme groups can contain carbohydrate-binding modules (CBMs) associated with the catalytic domain that enable efficient adsorption of the enzymes on insoluble polysaccharides, improving the enzymatic catalysis efficiency (Payne et al., 2015). Enzymatic accessibility is another critical parameter when plant biomass is used as a substrate, which usually makes pretreatment of the plant biomass material necessary before the enzymatic conversion processes (Chundawat et al., 2011). Residual lignin contained in pretreated materials can also cause unproductive binding of GHs and LPMOs, especially because CBMs also adsorb on lignin surfaces (Rahikainen et al., 2013; Siqueira et al., 2017; Santos et al., 2019).

Currently, ascomycetes are the main source for production of commercial enzymatic cocktails containing GHs and LPMOs destined for plant biomass conversion, because this class of fungi is amenable to genetic engineering, presents robust machinery to secrete such enzymes and is well developed from a bioprocess-engineering point of view (Juturu and Wu, 2014; Bischof et al., 2016; Druzhinina and Kubicek, 2017; Fitz et al., 2018). In ascomycete-based enzymatic cocktails, CBHs are the chief enzymes because they are uniquely able to cleave cellulose chains in a processive manner, releasing cellobiose with high efficiency (Payne et al., 2015). In contrast to ascomycetes, basidiomycetes involved in natural wood decay have been less developed as a source of such enzymes, despite the fact that this fungal class has evolved to degrade polysaccharides in lignified substrates and thus appears to be a logical reservoir for searching for proper plant biomass conversion enzymes (Floudas et al., 2012; Bentil et al., 2018; Valadares et al., 2019). In this context, sequencing of fungal genomes has increasingly opened an opportunity to search for new enzymes in wood decay fungi, which has resulted in the discovery of hitherto unexplored enzymes with particular characteristics that enable improved plant biomass conversion (Ohm et al., 2014; Lange et al., 2019).

Lignocellulosic substrates or pure polysaccharides and their degradation products have been used to induce secretion of plant biomass degrading proteins in fungi (Szabó et al., 1996; Suto and Tomita, 2001; Ahamed and Vermette, 2009; Juturu and Wu, 2014; Payne et al., 2015; Bischof et al., 2016; Presley and Schilling, 2017; Lin et al., 2019). Certain studies of white-rot fungi have shown that Avicel (a commercial preparation of microcrystalline cellulose) and its biodegradation products act as GH inducers in the model basidiomycete *Phanerochaete chrysosporium* (Uzcategui et al., 1991; Broda et al., 1995; Szabó et al., 1996; Suzuki et al., 2010). Other white-rot species have been less explored as plant biomass degrading protein producers (Rytioja et al., 2014; Bentil et al., 2018; Valadares et al., 2019).

Available studies indicate that *P. chrysosporium* produce a more diverse group of CBHs than the model cellulase producer, the ascomycete *Hypocrea jecorina* (anamorph *Thrichoderma reesei*). Seven CBHs are recognized in *P. chrysosporium* genome (Cel7A-F/G and one Cel6), and five of them are efficiently secreted when the fungus is cultured in cellulose, cello-oligosaccharides or plant biomass substrates (Martinez et al., 2004; Ravalason et al., 2008; Suzuki et al., 2009; Adav et al., 2012). In contrast, *T. reesei* encodes and produces only one CBHI (Cel 7A) and one CBHII (Cel 6A) (Peterson and Nevalainen, 2012; Kunamneni et al., 2014).

Previous literature suggests that distinct enzymes, useful for plant biomass conversion, can be found in wood decay fungi. In this context, the current work explores CBH-enriched secretomes of two white-rot fungi, *P. chrysosporium* and *T. versicolor*. Expression analysis of *cbh* genes was performed via RT-qPCR assays, and the produced enzymes were deeply characterized and used for digestion of lignin-free polysaccharides and plant biomass substrates. Both white-rot fungi secretomes enabled efficient cellulose and hemicellulose conversion to monosaccharides.

MATERIALS AND METHODS

Cultures of *P. chrysosporium* and *T. versicolor* in Microcrystalline Cellulose (Avicel) or Glucose as a Single Carbon Source

Stock cultures of *P. chrysosporium* (RP-78, ATCC MYA-4764) and *T. versicolor* (BAFC 266, Mycological Culture Collection of the Department of Biological Sciences, University of Buenos Aires,

Argentina - Levin et al., 2007) were maintained at 4°C on 2.0% malt extract (Synth, SP, Brazil) and 0.2% yeast extract (Vetec, Rio de Janeiro, Brazil) agar slants containing a wood chip slice.

Both basidiomycetes were cultured on Norkrans medium (Eriksson and Johnsrud, 1983) containing 20 g/L Avicel (Fluka, PH101) or 20 g/L glucose as a single carbon source. Cultures were grown in 1 L Erlenmeyer flasks containing 300 mL medium and inoculated with 10^4 spores/L in the case of *P. chrysosporium* or 500 mg/L of blended mycelium (Machado and Ferraz, 2017) in the case of *T. versicolor*. Cultures were maintained at 37°C for *P. chrysosporium* and 27°C for *T. versicolor* under 150 rpm rotary shaking for up to 12 days. At least three replicate culture flasks were inoculated for each culture period.

After defined culture periods, flask contents were filtered through Miracloth® (Millipore Sigma, Burlington, MA, United States). Solids representing fungal mycelium (from glucose cultures) or fungal mycelium plus residual Avicel were frozen in liquid nitrogen and stored at −80°C. Liquid broths from at least three replicate cultures were combined and further filtered through 0.45 μm polyethersulfone membranes (Sartorius Stedim, Göttingen, Germany). Filtrated broths were then concentrated via ultrafiltration through 10 kDa cut-off polyethersulfone membranes (Biomax Millipore, Janfrey, NH, United States) up to 35 times, depending on the fungal species and culture period (see **Supplementary Figure S1** for details). Samples of the concentrated culture broths were assayed for filter paper activity (FPA) and total proteins. The remaining concentrated broths were freeze-dried and stored at −18°C.

Protein Determination and Enzymatic Assays

Concentrated culture broths or dissolved freeze-dried broths were assayed for total proteins and for filter paper activity (FPA) as described in Ghose (1987). CBH (EC 3.2.1.176 and EC 3.2.1.91, for reducing and non-reducing end activity, respectively) (Wood and Bhat, 1988), endoglucanase (EC 3.2.1.4) (Ghose, 1987), endoxylanase (EC 3.2.1.8) (Bailey et al., 1992), β-glucosidase (EC 3.2.1.21) and β-xylosidase (EC EC 3.2.1.37) (Tan et al., 1987) were determined according to the indicated methods. At least three independent cultures from each fungal species were combined before concentration and freeze-drying of culture broths. After dissolution, freeze-dried broths were assayed in three analytical determinations of each enzymatic activity. Standard deviation for analytical triplicates varied less than 5% of the reported value.

SDS-PAGE and Proteomic Analysis

Freeze-dried culture broths were dissolved, and 30 μg of total protein from each sample was assessed via SDS-PAGE using a Mini Protean Tetra Cell System (Bio-Rad, Hercules, CA, United States). Protein bands were stained with 0.2% (w/v) Coomassie Brilliant Blue G 250, 50% (v/v) ethanol, and 10% (v/v) acetic acid. Each lane from the SDS-PAGE gels was divided into five fragments to cover the entire range of protein molar masses. Gel fragments were stored in a 1:1 methanol/water solution containing 0.1% formic acid and analyzed according to the protocols developed by the Life Sciences Core Facility (LaCTAD) at University of Campinas, SP, Brazil[1]. Gel-fragments were prepared by incubation with 100 μL of 100 mM ammonium bicarbonate/acetonitrile (1:1) for 30 min. After centrifugation, the liquid fraction was discarded, and solids were suspended in 500 μL acetonitrile until the gel became white. Decanted solids were dried and treated with 100 μL of 10 mM dithiothreitol (DTT) solution in 100 mM ammonium bicarbonate at 56°C for 30 min. Liquids were discarded, and dried solids were treated with 100 μL of 55 mM iodoacetoamide at room temperature for 30 min. Liquids were discarded again, and solids were digested with 100 μL of 13 mg/mL trypsin solution in 50 mM ammonium bicarbonate. Digestion with trypsin lasted overnight at 37°C. After centrifugation, the supernatant was collected and transferred to another microcentrifuge tube. Solids were further extracted for 15 min with 100 μL of 100 M ammonium bicarbonate/acetonitrile (1:1), and liquid fractions were combined for subsequent peptide analysis via LC-MS/MS.

For peptide analysis, 5 μL of digested sample was trapped in a Symmetry C18 precolumn (5 μm × 180 μm × 20 mm – Waters, Milford, MA, United States). The trapped sample was eluted inline onto an HSS T3 column (1.8 μm × 75 μm × 100 mm – Waters, Milford, MA, United States) using a solvent gradient ranging from 7% (v/v) to 85% (v/v) acetonitrile containing 0.1% (v/v) formic acid at 0.4 μL/min with a runtime of 73 min. LC-MS/MS data acquisition was achieved in a XEVO G2 Xs QToF mass spectrometer equipped with a nanolockspray source in the positive ion mode (Waters, Corp., Milford, MA. United States). Instrument calibration was performed with the MS/MS fragment ions of GFP [Glu 1]-Fibrinopeptide B with a doubly charged ion $[M + 2H]^{2+}$ = 785,84206 *m/z* (Waters, Corp., Milford, MA, United States). Data-independent scanning (MSE) experiments were performed by switching between low (3 eV) and high (15–50 eV) collision energies applied to a trap "T-wave" cell filled with argon. A scan time of 0.5 s was used to acquire data from *m/z* 50 to 2000.

Data were processed with Progenesis QI 4.0 (Non-linear Dynamics). The processing parameters included an automatic mode for MS-TOF resolution and chromatographic peak width. The low-energy and high-energy detection thresholds were optimized by software. Protein identification was fitted to a minimum of three fragment ions matched per peptide, a minimum of five fragment ions matched per protein, a minimum of one unique peptide matched per protein, one possible trypsin missed cleavage, carbamidomethylation of cysteine as a fixed modification and oxidation of methionine as a variable modification, and a maximum false positive discovery rate (FDR) in auto mode. The FDR for peptide and protein identification was determined based on a search of a reversed database, which was generated automatically using the Progenesis QI 4.0 software. All protein hits were identified with a confidence of >95%. The databases (Filtered Models) for *P. chrysosporium* v2.2 (RP-78) and *T. versicolor* v1.0 (FP-101664 SS1) used in the protein search were obtained at the MycoCosm portal from DOE Joint Genome Institute (Floudas et al., 2012; Ohm et al., 2014). The number of sequences and the average protein length (in amino

[1]https://www.lactad.unicamp.br/en/proteomics

acids) for *P. chrysosporium* and *T. versicolor* were 13602/408 and 14530/422, respectively.

Real Time qPCR Analysis

The primers used for amplification of the cDNA of *P. chrysosporium* genes for *cel7A-E, cel6*, and *actin* were the same reported by Suzuki et al. (2009). The sequences of primers used for amplification of the cDNA fragments derived from *T. versicolor cel7A-D* and *actin* were designed based on analysis of the non-coding regions of the genes (3′ UTRs) following the method established by Suzuki et al. (2009) (**Supplementary Table S1**). In the case of the *cel6* gene from *T. versicolor*, the primers were designed within the coding region of the gene (**Supplementary Table S1**). The corresponding transcripts of the *actin* gene were used as the internal control in both cases. Relative gene expression was calculated with the $2^{-\Delta\Delta CT}$ method using the same target gene expressed in mycelia recovered from catabolic-repressed cultures grown on 20 g/L glucose medium for 5 days as the calibrator sample (Livak and Schmittgen, 2001).

Frozen fungal mycelia were ground to a fine powder using a mortar and pestle in the presence of liquid nitrogen. Total RNA was extracted with an illustra™ RNAspin Mini Isolation Kit (GE Healthcare, Chicago, IL, United States) and treated with an RNase-free DNase set (Promega, Madison, Wisconsin, United States) according to the manufacturer's instructions. First-strand cDNA was synthesized using 800 ng treated RNA and 100 μM oligo $(dT)_{24}$ primer SuperScript™ III Reverse Transcriptase (Life Technologies™, Thermo Fisher) according to the manufacturer's instructions.

Real-time qPCR was performed in a 7500 Fast Real-Time PCR System (Life Technologies™, Thermo Fischer). The reaction wells contained 2.5 μL of prepared cDNA (diluted 25 times), 5 μL of Maxima SYBR Green/Rox qPCR Master Mixes buffer (Life Technologies™, Thermo Fischer), 0.4 μL of forward primer (100 mM), 0.4 μL of reverse primer (100 mM), and 11.7 μL of H_2O. The reaction conditions were 50°C for 2 min, 95°C for 10 min and 40 repetitions of 95°C for 15 s and 60°C for 1 min. All reactions were subjected to the same analysis conditions, and the results were normalized to the signal of the passive reference dye ROX to correct for fluctuations in the reading resulting from variations in volume and evaporation throughout the reaction. Two biological replicates were assayed in three technical replicates per cDNA sample.

Sequence Analysis and Comparison of CBHI Proteins

The sequences of the CBHI from *P. chrysosporium* (JGI 137372, JGI 2971601, JGI 3024803, and JGI 2976245) and *T. versicolor* (JGI 112163, JGI 110790, and JGI 124366) identified in the secretomes were checked against the InterPro 77.0 database[2] to validate the domains and used for multiple alignments with Geneious v4.8.5 using CBHI from *T. reesei* QM6a as a model (JGI 123989). Tridimensional modeling of *T. versicolor* CBHI-Cel7C (JGI 112163) was performed using the server I-TASSER (Yang and Zhang, 2015). Cel7A from *Heterobasidium irregulare* (formely *H. annosum*)[3] (Momeni et al., 2013) was used as a search model owing to 78% similarity between both CBH sequences. The CBHs from *T. reesei* (Cel7A) and *P. chrysosporium* (Cel7D) were used to compare structural differences (PDB id 1CEL and 1GPI, respectively). The figures were generated using PyMOL Molecular Graphics System version 2.3 (Schrödinger, LLC).

Enzymatic Hydrolysis of Lignocellulosic Substrates

Three different sugarcane-derived substrates were prepared as described in previous studies and selected to represent lignified plant biomass samples with varying recalcitrance to enzymatic digestion. The selected materials corresponded to (A) the pith region of mature sugarcane stalks of a low recalcitrant hybrid denominated H89 (Costa et al., 2013, 2016); (B) alkaline-sulfite pretreated sugarcane bagasse prepared as described in Reinoso et al. (2018); and (C) dilute acid pretreated sugarcane bagasse prepared as described in Santos et al. (2018), using a 90 min reaction time at 150°C. Chemical composition of these substrates was determined according to Ferraz et al. (2000).

Pretreated materials were digested inside 2 mL microcentrifuge tubes containing 20 mg of milled sugarcane biomass sample (passing 0.84 mm screen) suspended in a 1 mL reaction volume. Reaction solutions contained the enzymatic preparations at 10 FPU/g substrate dissolved in 50 mM sodium acetate (pH 4.8) containing 0.01% sodium azide. Reaction tubes were agitated at 120 rpm and 45°C for periods up to 72 h and sampled for monomeric sugar analysis according to Valadares et al. (2016).

RESULTS

Secretomes of *P. chrysosporium* and *T. versicolor* Grown in Microcrystalline Cellulose

Secretomes produced by *P. chrysosporium* and *T. versicolor* grown in microcrystalline cellulose (Avicel) were characterized and used for plant biomass saccharification. Avicel was used as the sole carbon source to enrich the secretomes with CBHs (Tangnu et al., 1981; Szabó et al., 1996; Payne et al., 2015). Increasing FPA levels were detected with extended culture of both white-rot fungi, with *P. chrysosporium* providing FPA (242 FPU/L) four times higher than *T. versicolor* (59 FPU/L) at the longest culture period (12 days) (**Supplementary Figure S1**). The secretomes contained several cellulolytic and hemicellulolytic activities (**Table 1**), resulting from a broad protein diversity, which was revealed by LC-MS/MS (**Figure 1** and **Supplementary Tables S2**, **S3**). Comparison of the two white-rot fungi showed that *P. chrysosporium* provided the highest FPA, CBH and endoglucanase activities (**Table 1**). In contrast, *T. versicolor* showed significantly higher titers for β-glucosidase and for the xylanolytic complex. The reference cellulolytic cocktail, Cellic® Ctec2 (Cannella et al., 2012; Bischof et al., 2016), differs

[2]http://www.ebi.ac.uk/interpro/

[3]https://www.ncbi.nlm.nih.gov/pubmed/?term=23303184

TABLE 1 | Enzymatic activities determined from freeze-dried culture extracts prepared from *P. chrysosporium* and *T. versicolor* grown in 20 g/L Avicel as sole carbon source for 12 days.

Fungi	Enzymatic activity* in the freeze-dried culture extracts (IU/mg protein)					
	FPA	CBH	EG	β-gluc	endo-Xyl	β-Xyl
P. chrysosporium	0.36	0.42	22.0	2.1	180	3.0
T. versicolor	0.13	0.03	11.8	20.0	894	9.6
Cellic® Ctec2**	0.95	1.90	25	46.0	96	0.5

**FPA, filter paper activity measured with paper strips; CBH, cellobiohydrolases measured with Avicel substrate; EG, endoglucanases measured with carboxymethyl cellulose substrate; β-gluc, β-glucosidases measured with pNPG substrate; endoXyl, endoxylanases measured with birchwood xylan substrate; and β-xyl, β-xylosidases measured with pNPX substrate. **Cellic® Ctec2 is a commercial enzyme blend acquired from Sigma (SAE0020). At least three independent cultures from each fungal species were combined before concentration and freeze-drying of culture broths. After dissolution, freeze-dried broths were assayed in three analytical determinations of each enzymatic activity. Standard deviation for analytical triplicates varied less than 5% of the reported value.*

from white-rot secretomes because it contains higher CBH and β-glucosidase activities, which consistently result in higher FPA (**Table 1**).

The *P. chrysosporium* and *T. versicolor* secretomes contained 56 and 58 detectable proteins, respectively (**Supplementary Tables S2**, **S3**). Most of the identified proteins presented signal peptides (Petersen et al., 2011) or signals for non-classical secretion pathways (Bendtsen et al., 2004). Only one protein in the *P. chrysosporium* and four in the *T. versicolor* secretomes were intracellular. Plant cell wall degrading enzymes (cellulolytic, hemicellulolytic, pectinolytic, esterase, and auxiliary activity) predominated among the detected proteins, altogether representing 78 and 86% of the peptides detected by LC-MS/MS in the *P. chrysosporium* and *T. versicolor* secretomes, respectively (**Figure 1**). Minor proteins included fungal cell wall degradation proteins, amylolytic enzymes, proteases, hypothetical proteins, and proteins with other classifications.

Figure 2 highlights the relative abundance of each plant cell wall degrading protein in both secretomes. GH7-CBHIs corresponded to 47 and 34% of the total plant cell wall degrading proteins in *P. chrysosporium* and *T. versicolor*, respectively. GH6-CBHIIs were also detected in both fungi, with 1–3% relative abundance. The predominance of CBHs in both secretomes confirms that cellulose and its degradation products are useful CBH inducers in white-rot fungi (Broda et al., 1995; Szabó et al., 1996; Suzuki et al., 2010). The second most abundant enzyme group in both fungi was GH10 and GH11 endoxylanases in *P. chrysosporium* (17%) and GH10 endoxylanases in *T. versicolor* (21%). Endoglucanases appeared in the sequence, with GH5, GH12 and GH45 families in *P. chrysosporium* (14%) and GH5 and GH45 families in *T. versicolor* (10.5%) (**Figure 2**). One GH3 β-glucosidase was detected in the *T. versicolor* cultures (3.1%), but putative β-glucosidases were not detectable in the *P. chrysosporium* secretome. Several auxiliary activity (AA) enzymes were also detected, with 6 LPMOs (EC 1.14.99.54-C1-hydroxylating and EC 1.14.99.56-C4-dehydrogenating) from AA9 (4%) and one cellobiose-dehydrogenase (CDH-EC 1.1.99.18) from AA3_AA8 (2%) in *P. chrysosporium*. In contrast, only one putative LPMO from AA9 (0.6%) but one CDH from AA3_AA8 (4.7%) and one cytochrome c domain from AA8 appended to a CBM1 (1.7%) were found in *T. versicolor*. Accessory enzymes involved in cleavage of hemicellulose side groups, including GHs and carbohydrate esterases (CEs), as well as putative pectinolytic GHs, also appeared in both fungal secretomes (**Figure 2**).

CBHs are major proteins in the commercial enzymatic cocktails prepared for efficient plant biomass saccharification (**Table 1**) because they are able to decompose long cellulose chains into cellobiose via a processive mode of action (Payne et al., 2015). CBHs also formed the main enzymes in the secretomes of both white-rot fungi cultured in 20 g/L Avicel as the sole carbon source. However, the processivity of each particular CBH and the velocity at which CBHs cleave glucosidic bonds depend on the enzyme structure (Momeni et al., 2013; Taylor et al., 2018) and on the substrate characteristics (Kurasin and Väljamäe, 2011). In this context, the type of CBH present in each white-rot fungi secretome was detailed to reveal CBH variations among the well-known *P. chrysosporium* (Munoz et al., 2001; von Ossowski et al., 2003; Kurasin and Väljamäe, 2011; Tachioka et al., 2017; Taylor et al., 2018) and the understudied *T. versicolor*.

Among CBHs, CBHI-Cel7D (JGI 137372) predominated in the secretome of *P. chrysosporium* (37%), followed by CBHI-Cel7C (JGI 2971601) (6.8%), CBHII (JGI2965119) (1.7%), CBHI-Cel7E (JGI 3024803) (1.3%), and CBHI-Cel7F/G (JGI 2976245/JGI 2976248) (0.5%). In *T. versicolor* cultures, CBHI-Cel7C (JGI 112163) predominated (26%) over CBHI-Cel7B (JGI 110790) (7.0%), followed by CBHII (JGI 63826) (3.1%), and CBHI-Cel7D (JGI 124266) (1.0%). Real-time qPCR provided the pattern of *cbh* gene transcripts in the fungal mycelia grown during the cultivation periods. For both fungal species, *cbh* transcripts measured at 5, 8, and 12 days attained maximal levels at day 8, decreasing afterward (**Figure 3**). This behavior suggested an initial induction period as already demonstrated for *P. chrysosporium* K-3 strain (Suzuki et al., 2009), whereas the late decrease in *cbh* transcripts is likely associated with repression caused by glucose released from Avicel at advanced culture periods (Daly et al., 2019). However, total FPA progressively accumulated in the culture media along the 12 days culture period (**Supplementary Figure S1**), indicating a delay between gene transcription and enzyme secretion and extracellular accumulation (Midorikawa et al., 2018; Lin et al., 2019).

The CBHs detected in the secretome dovetail with the *cbh* gene transcript levels, since CBHI-Cel7D (JGI 137372) and CBHI-Cel7C (JGI 112163) predominated in the *P. chrysosporium*

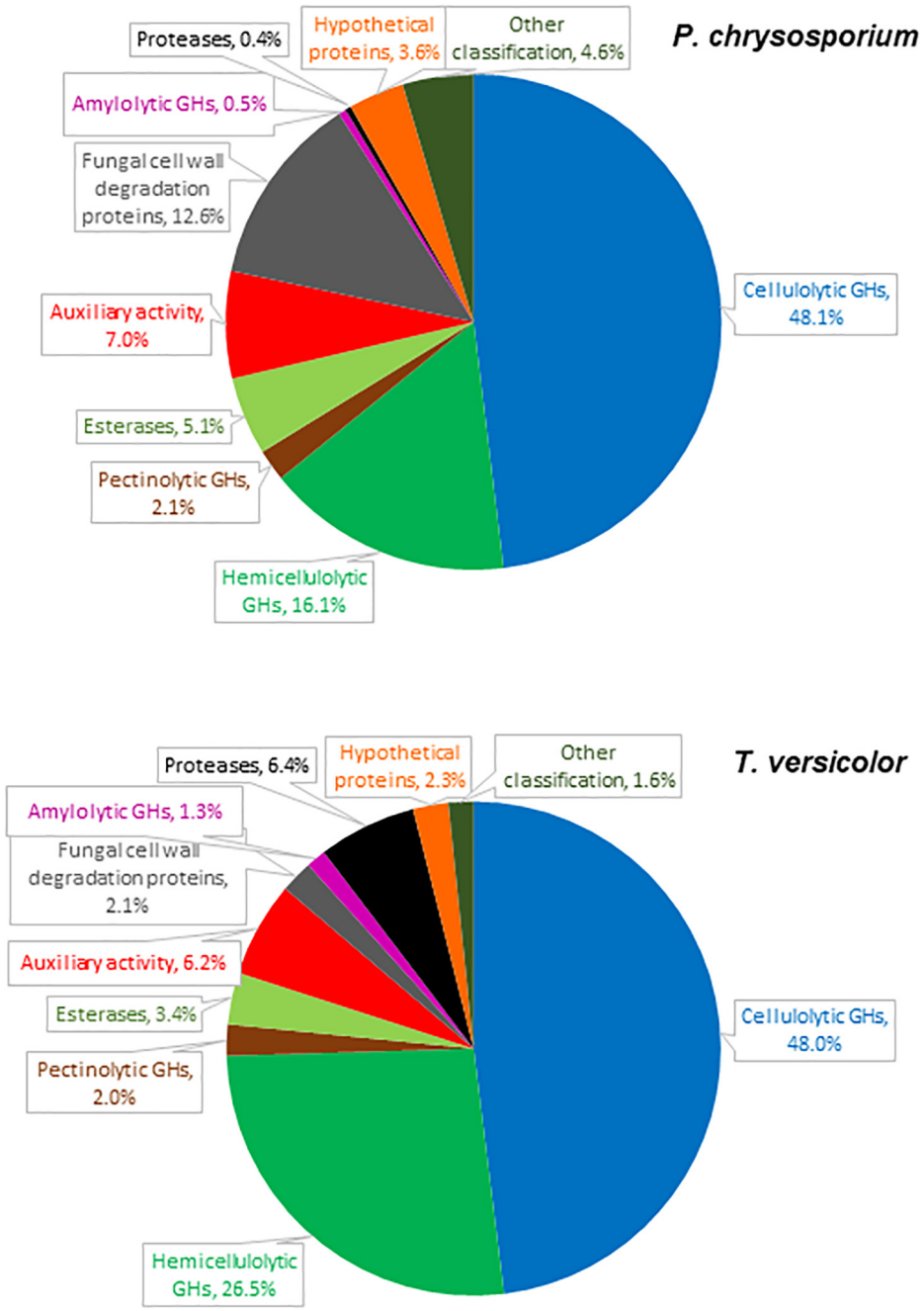


FIGURE 1 | Relative abundance of proteins detected in cultures of *P. chrysosporium* and *T. versicolor* grown for 12 days in 20 g/L Avicel as sole carbon source.

and *T. versicolor* secretomes, respectively (**Figure 2**), and the corresponding transcripts (*Pc cel7D* and *Tv cel7C*) prevailed in the studied mycelium of each species (**Figure 3**). Detection of *Pc cel6* and *Pc cel7C*, and *Tv cel6* (**Figure 3**) also agreed with the occurrence of CBHII (JGI2965119) and CBHI Cel 7C (JGI 2971601) in the *P. chrysosporium* and CBHII (JGI 63826) in the *T. versicolor* secretome, respectively (**Figure 2**). The primers used for amplification of *cel7A* and *cel7B* fragments from cDNA of *P. chrysosporium* confirmed that the genes were not being transcribed under the tested conditions, consistent with the absence of the corresponding enzymes in the secretome of this fungal species. *Pc cel7E* transcripts were only barely detectable in *P. chrysosporium* (data not shown), as was the case for CBHI-Cel7E (JGI 3024803). In *T. versicolor* cultures, *cel7A* transcripts and the corresponding protein were not detectable. However, in contrast with previous observations, CBHI-Cel7B (JGI 110790) and CBHII (JGI 63826) appeared in the *T. versicolor* secretome, whereas the primers used for amplification of *cel7B* and *cel6*

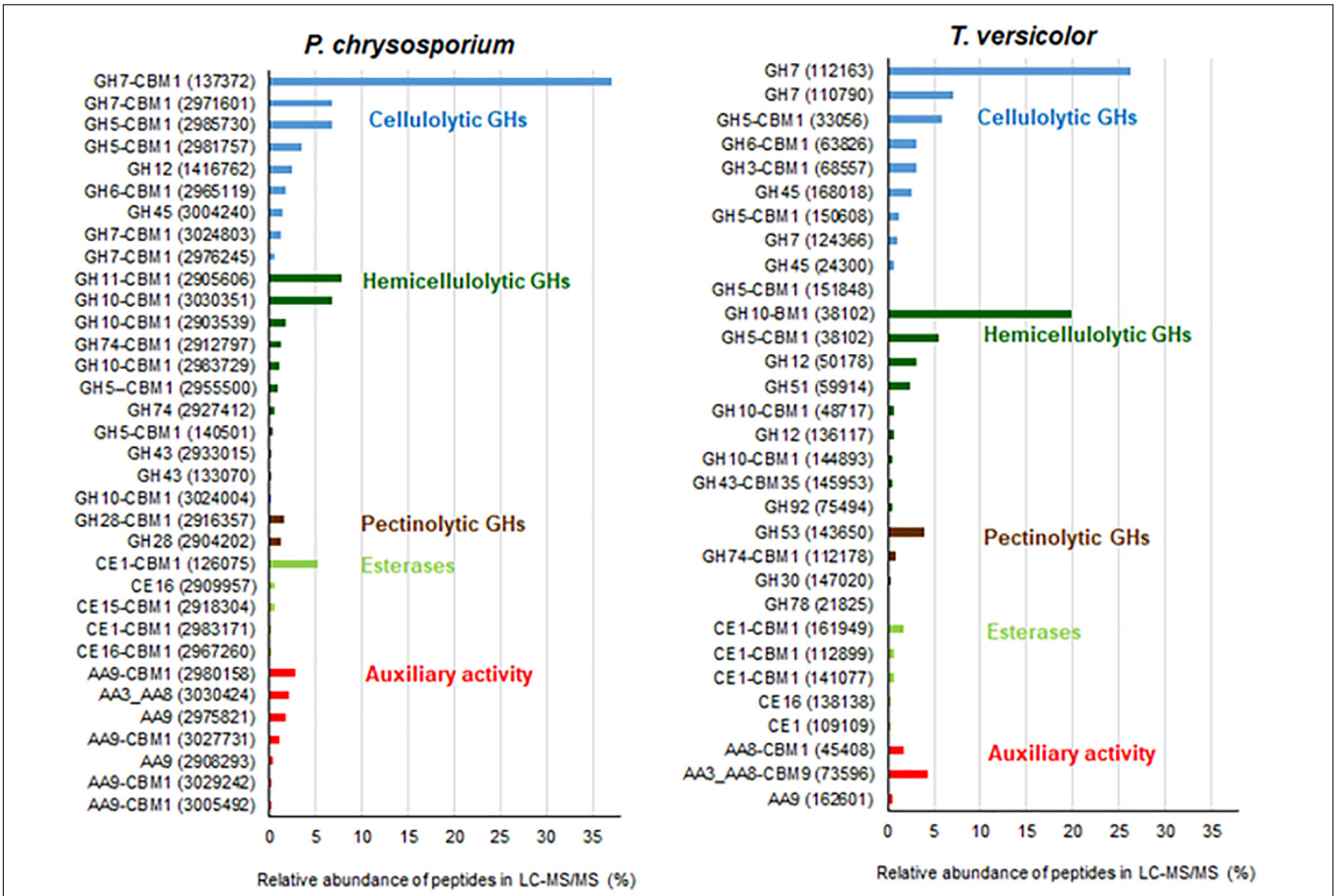


FIGURE 2 | Relative abundance of plant cell wall degrading proteins detected in cultures of *P. chrysosporium* and *T. versicolor* grown for 12 days in 20 g/L Avicel as sole carbon source. JGI identification protein numbers are indicated in parenthesis. Relative abundance is expressed considering the total plant cell wall degrading proteins as 100%. Original relative abundance of plant cell wall degrading proteins in each secretome was 78 and 86% for *P. chrysosporium* and *T. versicolor*, respectively.

fragments from cDNA (Table S1) were applied but did not amplify any fragments.

CBHs were further compared through a sequence analysis that highlighted certain structural differences among the CBHs from *P. chrysosporium* and *T. versicolor* detected in the secretome. CBHI from *T. reesei* (JGI 123989) was used as a reference (**Figure 4A**). Alignment of the catalytic modules and molecular modeling showed several common structures, such as a conserved cellulose-biding site and the catalytic triad (Glu-214, Asp-216 and Glu-209) (**Figure 4A**) (Momeni et al., 2013). However, the whole sequences confirmed that CBHIs from *T. versicolor* lack the typical CBM from family 1 found in *P. chrysosporium* and *T. reesei* (data not shown).

Structural analysis showed that the catalytic cores of CBHI were typical, with a β-sandwich surrounded by loops (**Figures 4B–D**). Compared to CBHIs from *P. chrysosporium* and *T. reesei*, the Loop A1 in CBHIs from *T. versicolor* differs in length and sequence and contains a tyrosine at position 101 (Y-101 in **Figures 4A,B**). The tunnel entrance in *T. versicolor* CBHIs is narrower than that in *P. chrysosporium* and *T. reesei* CBHIs (**Figures 4E–G**). Other differences among these CBHIs occur in Loop B2 (**Figure 4A**), which lacks two amino acid residues in *P. chrysosporium* CBHIs and is substituted in *T. versicolor* compared to the *T. reesei* enzyme that contain two asparagine at positions 197 and 198. Finally, in loop B3 (**Figure 4A**), CBHIs from *T. versicolor* and *P. chrysosporium* are shorter by two and six residues, respectively, compared to *T. reesei* CBHI.

Saccharification of Sugarcane Substrates by the Crude Secretomes of White-Rot Fungi

Three different sugarcane-derived substrates were digested with the secretomes produced by *P. chrysosporium* and *T. versicolor* with the aim of evaluating their performance on more complex, plant biomass-derived substrates (**Figure 5**). For all experiments, the protein loads were adjusted to provide 10 FPU/g substrate. Cellic® CTec2 served as an upmost reference enzymatic cocktail in the same experiments. As can be depicted from **Table 1** data, the protein load and enzymatic activities in the digestion tubes varied considerably to assure an equal load of 10 FPU/g substrate in all cases.

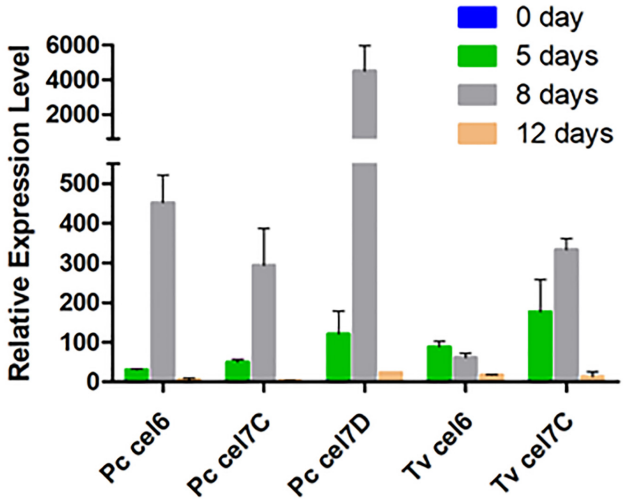


FIGURE 3 | Relative transcription levels of genes encoding cellobiohydrolase in *P. chrysosporium* (*Pc*) and *T. versicolor* (*Tv*) grown in 20 g/L Avicel as sole carbon source. Actin was used as house-keeping gene in both cases. Relative transcription levels are reported as $2^{-\Delta\Delta Ct}$ using catabolic repressed culture (mycelia recovered from 20 g/L glucose for 5 days) as the calibrator sample. Two biological replicates were assayed in analytical triplicate for qPCR.

The *T. versicolor* secretome, with low FPA and CBH activities, required a high protein load (75 mg protein/g substrate) to reach 10 FPU/g substrate. The *P. chrysosporium* secretome presented more balanced CBH:EG:β-glucosidase levels, providing 10 FPU/g substrate with 28 mg protein/g substrate, whereas Cellic® CTec2, presenting a proper CBH:EG:β-glucosidase balance, required minimal protein loads (10 mg protein/substrate) to reach the same 10 FPU/g substrate. In addition to traditional cellulases, Cellic® CTec2 also contains a sufficient number of xylanases and LPMOs (Cannella et al., 2012), enhancing the digestion efficiency for complex plant biomass substrates. Finally, owing to the low FPA and CBH and high endoxylanase and β-xylosidase activities, experiments with the *T. versicolor* secretome occurred with a disproportionally high xylanase load, corresponding to 27000 IU/g substrate compared to 1000 and 500 IU/g substrate for Cellic® Ctec2 and the *P. chrysosporium* secretome, respectively.

The three enzymatic cocktails were used for saccharification of diverse sugarcane-derived substrates: (A) sugarcane pith that does not need pretreatment (Costa et al., 2013, 2016); (B) alkaline-sulfite pretreated material (Reinoso et al., 2018); and (C) dilute acid pretreated material (Santos et al., 2018). The chemical compositions of these substrates are presented in **Supplementary Table S4**. The major characteristics of these substrates indicate that sugarcane pith is a non-pretreated plant biomass sample rich in cellulose and β-1-3/β-1-4 mixed linkage glucan, presenting a low crystallinity index, and a low lignin content (Costa et al., 2016). The xylan contained in sugarcane pith retains its original structure, presenting a high number of side groups, such as acetylated groups (Costa et al., 2016). Alkaline-sulfite pretreated sugarcane bagasse is depleted in lignin (approximately 50% of lignin is removed during pretreatment), has almost all cellulose preserved and its xylan lacks major acetyl and feruloyl side groups owing to saponification during alkaline pretreatment (Reinoso et al., 2018). Dilute acid pretreated sugarcane bagasse is enriched in crystalline cellulose and lignin because almost all amorphous cellulose and 90% of xylan are removed during pretreatment (Santos et al., 2018).

Based on the glucan conversion levels provided by the reference Cellic® Ctec2 cocktail, the three substrates presented decreasing digestibility, with sugarcane pith $\geq$ alkaline sulfite $>>$ dilute sulfuric acid (**Figure 5**). For pretreated substrates, the high CBH titers detected in Cellic® Ctec2 (**Table 1**) seem essential for high glucan conversion to glucose, since the reference cocktail performed significantly better than the *P. chrysosporium* and *T. versicolor* secretomes (**Figure 5**). In contrast, for low recalcitrance materials, such as the sugarcane pith (Costa et al., 2016), glucan conversion provided by both white-rot secretomes reached efficiency values comparable to those observed during saccharification with Cellic® Ctec2 (**Figure 5A**).

During sugarcane pith digestion, the xylan conversion levels were significantly lower than the glucan conversion levels for all three enzymatic cocktails (**Figure 5**). Selected sugarcane hybrids contain a pith region comprising a very low recalcitrant lignocellulosic material rich in parenchyma cells that do not require pretreatment for efficient enzymatic glucan conversion to glucose (Costa et al., 2013; Costa et al., 2016). However, the absence of pretreatment results in preserved xylan structures containing all side group decorations, which restrain efficient xylan hydrolysis (Varnai et al., 2014).

Saccharification of Sugarcane-Derived Substrates by the Secretomes of White-Rot Fungi Supplemented With β-Glucosidases

The lack of β-glucosidase in the *P. chrysosporium* secretome was experimentally overcome by supplementation with commercial *Aspergillus niger* β-glucosidase, which also presents some xylanase activity (Valadares et al., 2016; **Figure 6**). Supplementation with β-glucosidase enhanced the glucan conversion levels of all substrates. With β-glucosidase supplementation, the *P. chrysosporium* secretome provided higher glucan and xylan conversion than the *T. versicolor* secretome. However, in sugarcane pith, supplementation with *A. niger* β-glucosidase presented a simple additive effect, since glucan conversion to glucose was enhanced 19% in the experiment with *P. chrysosporium* secretome compared to the 17% provided by *A. niger* β-glucosidase alone (**Figure 6A**). The *T. versicolor* secretome was less reliant on β-glucosidase supplementation, probably because it presented a high original β-glucosidase activity (**Table 1**). For the pretreated substrates, supplementation with β-glucosidase provided synergic action for glucan hydrolysis, because the enhanced glucan conversion surpassed the values provided by *A. niger* β-glucosidase alone (**Figures 6B,C**). A remarkable enhancement in xylan hydrolysis

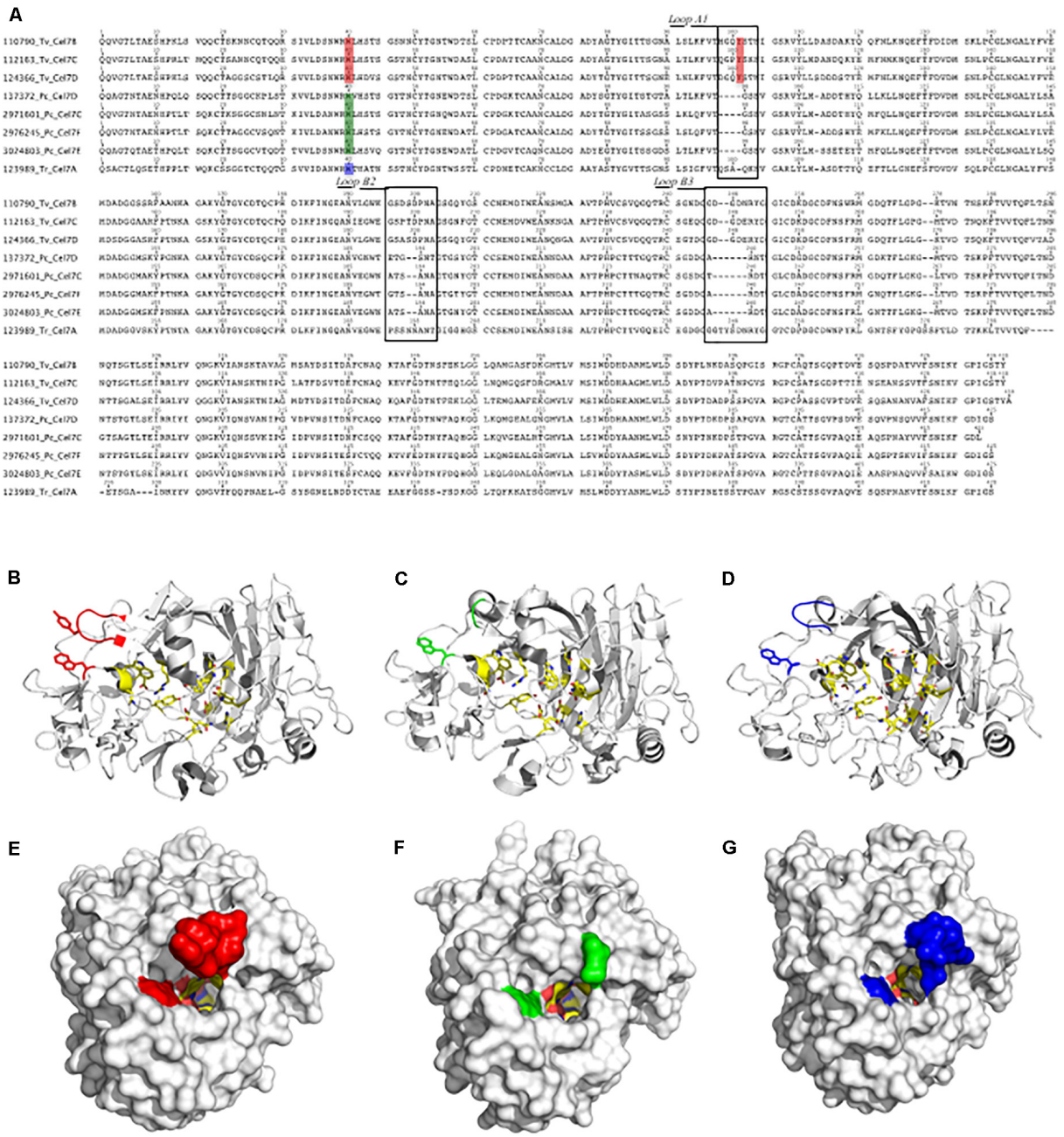


FIGURE 4 | Sequence alignment of CBHIs from *T. versicolor* (110790 Tv Cel7B, 112163 Tv Cel7C and 124366 Tv Cel7D), *P. chrysosporium* (137372 Pc Cel7D, 2971601 Pc Cel7C, 2976245 Pc Cel7F, and 3024803 Pc Cel7E) and *T. reesei* (123989 Tr Cel7A) **(A)**. Structural models for CBHI 112163 Tv Cel7C **(B,E)**, CBHI 137372 Pc Cel7D **(C,F)** and 123989 Tr Cel7A **(D,G)**. In **(A)** the squares indicate the loop A1, loop B2, and loop B3. Amino acids in the tunnel entrance are also highlighted in red, green and blue **(A)**, which can be compared at the structures **(B–D)** or in the surface **(E–G)** in the corresponding colors.

efficiency was also noted when the white-rot secretomes were supplemented with *A. niger* β-glucosidase, probably associated with the high xylanase activity described for this commercial preparation (Valadares et al., 2016). The enhanced xylan hydrolysis could also enhance glucan hydrolysis owing to progressive xylan removal from cellulose fibrils during the hydrolysis process (Varnai et al., 2014).

DISCUSSION

The enzymatic activities measured in the secretome of *P. chrysosporium* and *T. versicolor* (**Table 1**) and the relative abundance of proteins detected by LC-MS/MS (**Figure 2**) were in good agreement. For example, xylanase activity (endo-xylanase and β-xylosidase) was higher in *T. versicolor* than

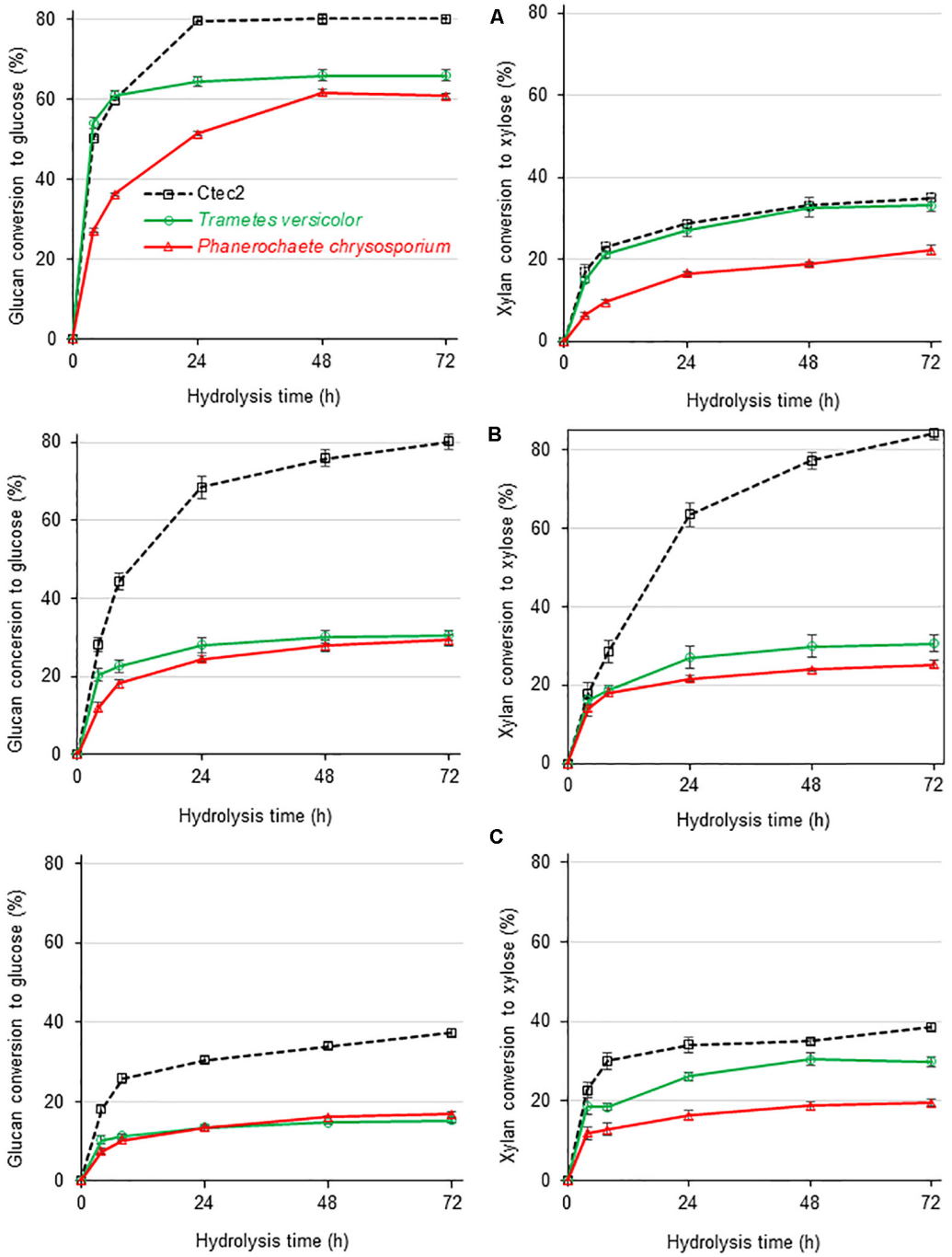


FIGURE 5 | Glucan (left) and xylan (right) conversion during enzymatic hydrolysis of three different sugarcane-derived substrates by secretomes from white-rot fungi as compared to Cellic® Ctec2 at a fixed load of 10 FPU/g of substrate but varied compositional enzymatic activities as reported in **Table 1**. **(A)** Sugar cane pith, **(B)** alkaline-sulfite pretreated sugarcane bagasse, and **(C)** dilute acid pretreated sugarcane bagasse, substrates. At least five independent cultures from each fungal species were combined before concentration and freeze-drying of culture broths. After dissolution, freeze-dried broths were used for plant biomass saccharification in three analytical determinations.

in *P. chrysosporium* cultures (**Table 1**), which agrees with the higher relative abundance of hemicellulolytic enzymes in the secretome of *T. versicolor*. The high β-glucosidase activity in *T. versicolor* cultures and the near absence of this activity in *P. chrysosporium* cultures also agrees with the detection of significant GH3 β-glucosidase in the secretome of *T. versicolor* but the absence of putative β-glucosidases in the *P. chrysosporium* secretome. Actually, *P. chrysosporium*

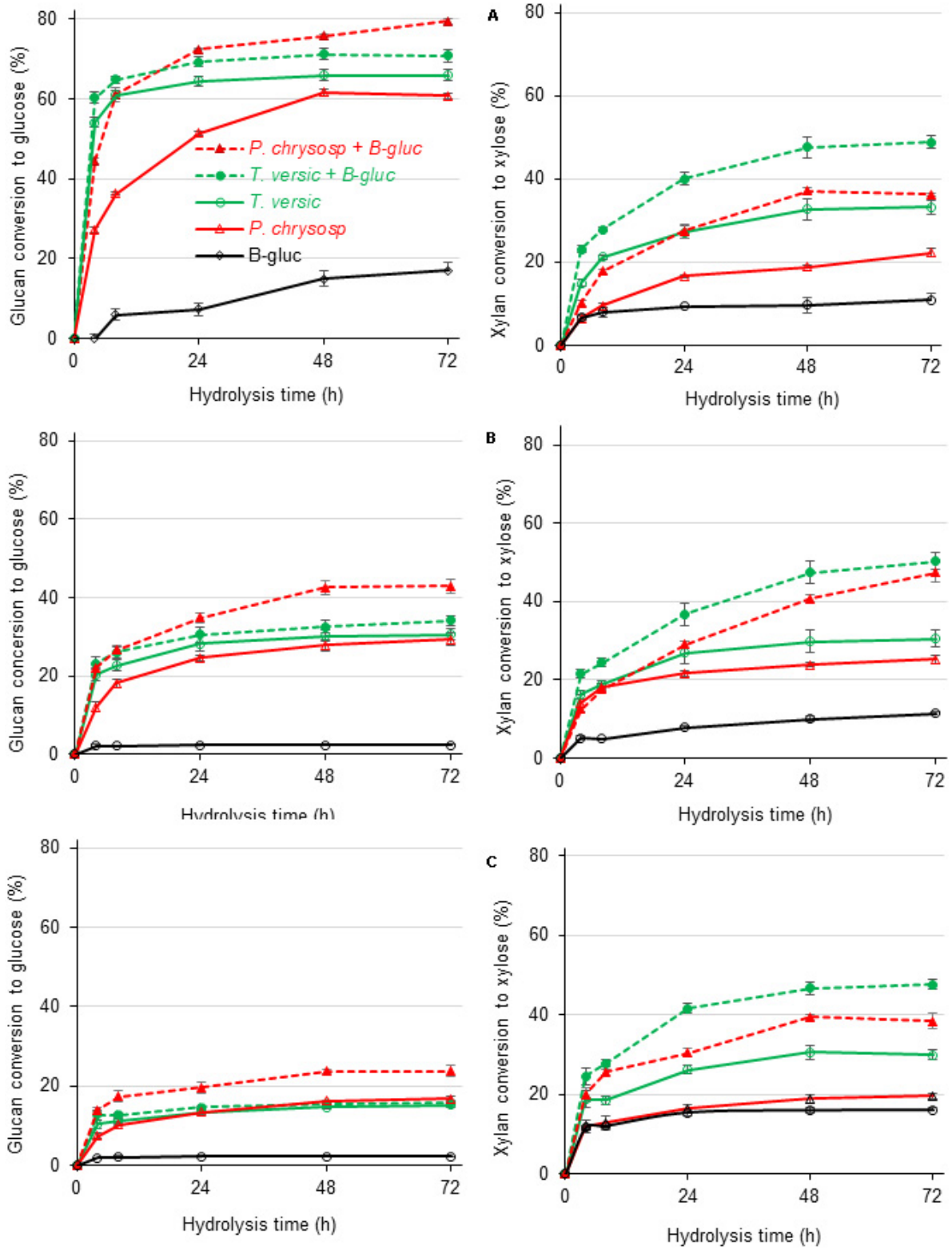


FIGURE 6 | Glucan (left) and xylan (right) conversion during enzymatic hydrolysis of three different sugarcane-derived substrates by concentrate secretomes from white-rot fungi additional 10 IU of β-glucosidase/g of substrates (labeled with + B-glic). **(A)** Sugar cane pith, **(B)** alkaline sulfite pretreated sugarcane bagasse, and **(C)** dilute acid pretreated sugarcane bagasse substrates. At least five independent cultures from each fungal species were combined before concentration and freeze-drying of culture broths. After dissolution, freeze-dried broths were used for plant biomass saccharification in three analytical determinations.

has intracellular β-glucosidases and proper mechanisms for cellobiose transportation through the cell membrane instead of extracellular β-glucosidases (Tsukada et al., 2006). Endoglucanases also appeared at higher relative abundance in *P. chrysosporium* (14%) than in *T. versicolor* (10.5%), consistent with the higher endoglucanase activity in the *P. chrysosporium* secretome (**Table 1**). In contrast to previous remarks, *T. versicolor* presented significantly lower activity on crystalline cellulose than *P. chrysosporium* (**Table 1**), although its secretome presented relatively abundant CBHs (37.5%) compared to the *P. chrysosporium* secretome (48.7%).

The lack of CBMs in *T. versicolor* CBHs seems critical in explaining the low activity of the *T. versicolor* secretome on crystalline cellulose. In contrast, in addition to higher CBH abundance, the *P. chrysosporium* secretome contained abundant LPMOs and CDH (**Figure 2** and **Supplementary Tables S2**, **S3**) that could boost hydrolysis of crystalline cellulose. FPA followed a trend similar to that of CBH activity, since the activities were significantly higher in *P. chrysosporium* than in *T. versicolor* (**Table 1**).

Further comparison of the main CBHs detected in both fungal species (CBHI-Cel7D – JGI 137372 in *P. chrysosporium* and CBHI-Cel7C – JGI 112163 in *T. versicolor*) with CBHI from *T. reesei* indicated that the basidiomycete enzymes present a more open cellulose-binding tunnel, which is related to the shorter loops (**Figure 4**). Previous studies based on computational analysis demonstrated that enzymes with shorter loops dissociate faster from substrate, showing a low inhibition rate by cellobiose and a high capacity to degrade microcrystalline cellulose because these enzymes can attach to the substrate with an endo-initiation mode of action (Kurasin and Väljamäe, 2011). In addition, the CBHIs from *T. versicolor* contain a tyrosine in the tunnel entrance, which in combination with tryptophan (W40 and Y101 highlighted in red, **Figures 4A, B,** and **E**) provides an extra binding platform for the enzyme, helping to drive the cellulose chain end into the catalytic site tunnel; this structure was already demonstrated in a similar CBHI produced by the basidiomycete *Heterobasidion irregulare* (Momeni et al., 2013).

Regarding plant biomass substrates, the hydrolytic performance of the three enzymatic cocktails varied considerably according to the characteristics of the substrate and the polysaccharide under evaluation (**Figure 5**). This varied performance primarily reflects the diverse composition of enzymes in each cocktail, although the same 10 FPU/g of substrate was maintained in all cases.

The *T. versicolor* secretome provided higher glucan conversion levels than the *P. chrysosporium* secretome, despite presenting a lower CBH proportion in the enzymatic cocktail (**Table 1** and **Figure 2**). The higher efficiency of the *T. versicolor* secretome was especially evidenced during hydrolysis of sugarcane pith (**Figure 5A**). In this substrate, the initial glucan hydrolysis rate (up to 8 h of reaction time) provided by the *T. versicolor* secretome reached the same values observed for Cellic® Ctec2. It is noteworthy that sugarcane pith presents high amounts of β-1-3/β-1-4 mixed linkage glucans and a low crystallinity index (Costa et al., 2016). Therefore, the very high β-glucosidase activity and sufficient endoglucanases in the *T. versicolor* secretome (**Table 1**) likely compensate for the lower proportion of CBH in its secretome. Excess β-glucosidase has also been reported to help in diminishing unproductive CBH adsorption on lignin, since β-glucosidase strongly adsorbs to this component (Jung et al., 2020).

The higher xylan conversion levels provided by the high xylanase activity in the *T. versicolor* secretome (similar to that provided by Cellic® Ctec 2 – **Figure 5A**) could also facilitate the access of cellulolytic enzymes to cellulose in complex substrates (Varnai et al., 2014), favoring glucan conversion by this secretome. Therefore, the high endoxylanase and β-xylosidase titers detected in the *T. versicolor* secretome (**Table 1**) and the great diversity of xylan degrading enzymes in its secretome, including acetyl-xylan esterases (**Supplementary Table S3** and **Figure 3**), seems to have favored significant xylan hydrolysis in sugarcane pith (**Figure 5A**). However, compared with the reference cocktail Cellic® Ctec2, the *T. versicolor* secretome failed to promote the same efficiency in xylan hydrolysis in alkaline-sulfite pretreated substrate (**Figure 5B**). This substrate is more recalcitrant than sugarcane pith and presents xylan structures that lack side decorations, especially acetyl groups, which are mostly removed by the alkaline pretreatment (Reinoso et al., 2018).

Another factor associated with the more efficient glucan conversion by *T. versicolor* than *P. chrysosporium* can be attributed to lack of CBMs in CBHsI (**Figure 2** and **Supplementary Table S3**). The lack of CBM in CBHsI from *T. versicolor* seems to be a limitation for hydrolysis of pure crystalline cellulose (**Table 1**) but not for digestion of lignified substrates, since less unproductive binding on lignin occurs with this type of CBH (Strobel et al., 2016; Lu et al., 2017). Indeed, some reports emphasize conflicting effects of CBMs in CBHs during hydrolysis of lignified substrates because they favor CBH attachment to cellulose chains, favoring cellulose hydrolysis efficiency; however, CBMs also adsorb unproductively to lignin, decreasing the amount of available CBH for cellulose processing (Strobel et al., 2016; Zhang et al., 2016; Lu et al., 2017).

In summary, the comparison of the white-rot fungi secretomes acting on lignified (**Figure 5**) versus lignin-free substrates (**Table 1**) suggests that *P. chrysosporium* secretome (with higher FPA and CBH activities) is more efficient to digest crystalline cellulose present in lignin-free substrates such as filter paper and Avicel because *P. chrysosporium* secretome presents higher amounts of CBHs, which contain CBMs. However, when both secretomes were applied to lignified substrates, at fixed 10 FPU/g substrate, *T. versicolor* provided higher glucan conversions, even presenting a lower proportion of CBHs, probably because the other enzymes present in this secretome and CBHs lacking CBMs compensate for problems associated with unproductive binding to lignin. Therefore, future exploring of *T. versicolor* CBHs seems relevant for preparation of new enzymatic cocktails used for plant biomass saccharification.

CONCLUSION

Integrated studies of secretome, CBH transcription and enzymatic hydrolysis of varied substrates highlighted that microcrystalline cellulose (Avicel) was useful in inducing abundant CBHs in cultures of *P. chrysosporium* and *T. versicolor* white-rot fungi, enabling production of fungal secretomes for capable digestion of complex-lignified substrates. CBHI-Cel7D-CBM1 (JGI 137372) predominated in the *P. chrysosporium* secretome, whereas CBHI-Cel7C lacking CBM (JGI 112163) was the more abundant enzyme in the *T. versicolor* secretome. On a fixed 10 FPU/g substrate, these secretomes performed similarly to a commercial enzymatic cocktail (Cellic® Ctec 2) when acting on a low-recalcitrant sugarcane pith substrate. However, efficient digestion of alkali or acid pretreated sugarcane-derived substrates required higher CBH proportion in the enzymatic cocktail, detected only in Cellic® Ctec2. Comparison of the two white-rot fungi secretomes indicated that *T. versicolor* performed more efficiently in lignified substrates, even at a relatively low CBH proportion, likely helped by its high abundance of GH3 β-glucosidase and the presence of enzymes from the xylanolytic complex, including GH10 endoxylanases and CE1 acetylxylan esterases. *T. versicolor* CBHIs also lacked CBMs, which could have contributed to lower unproductive binding of these CBHs to the lignin contained in the complex sugarcane-derived substrates. Therefore, *T. versicolor* enzymes induced by Avicel seem to be good candidates for future research on preparation of new enzymatic cocktails used for hydrolysis of lignified substrates.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

AM performed most of the experiments, the data interpretation, and discussion. FV, TS, and AMM participated in secretome and enzyme expression studies, the data interpretation, and discussion. FS participated in secretome and enzyme expression studies, protein modeling, data interpretation, and discussion. AF conceived the study, participated in data interpretation, discussion, and prepared the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by FAPESP (2014/06923-6 and 2014/18714-2) and CNPq. This study was also financed in part by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-Brazil) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Finance Code 001.

ACKNOWLEDGMENTS

We authors would like to thank J. M. Silva for technical assistance. We thank the staff of the Life Sciences Core Facility (LaCTAD) from State University of Campinas (UNICAMP), for the proteomic analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00826/full#supplementary-material

REFERENCES

Adav, S. S., Ravindran, A., and Sze, S. K. (2012). Quantitative proteomic analysis of lignocellulolytic enzymes by *Phanerochaete chrysosporium* on different lignocellulosic biomass. *J. Proteomics* 75, 1493–1504. doi: 10.1016/j.jprot.2011.11.020

Adsul, M., Sandhu, S. K., Singhania, R. R., Gupta, R., Puri, S. K., and Mathur, A. (2020). Designing a cellulolytic enzyme cocktail for the efficient and economical conversion of lignocellulosic biomass to biofuels. *Enzyme Microb. Technol.* 133:109442. doi: 10.1016/j.enzmictec.2019.109442

Ahamed, A., and Vermette, P. (2009). Effect of culture medium composition on *Trichoderma reesei's* morphology and cellulase production. *Bioresource Technol.* 100, 5979–5987. doi: 10.1016/j.biortech.2009.02.070

Bailey, M. J., Biely, P., and Poutanen, K. (1992). Interlaboratory testing of methods for assay of xylanase activity. *J. Biotechnol.* 23, 257–270. doi: 10.1016/0168-1656(92)90074-j

Bendtsen, J. D., Jensen, L. J., Blom, N., von Heijne, G., and Brunak, S. (2004). Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng. Des. Sel.* 17, 349–356. doi: 10.1093/protein/gzh037

Bentil, J. A., Thygesen, A., Mensah, M., Lange, L., and Meyer, A. S. (2018). Cellulase production by white-rot basidiomycetous fungi: solid-state versus submerged cultivation. *Appl. Microbiol. Biotechnol.* 102, 5827–5839. doi: 10.1007/s00253-018-9072-8

Bischof, R. H., Ramoni, J., and Seiboth, B. (2016). Cellulases and beyond: the first 70 years of the enzyme producer *Trichoderma reesei*. *Microb. Cell. Fact.* 15, 106–118.

Broda, P., Birch, P. R., Brooks, P. R., and Sims, P. F. (1995). PCR-mediated analysis of lignocellulolytic gene transcription by *Phanerochaete chrysosporium*: substrate-dependent differential expression within gene families. *Appl. Environ. Microbiol.* 6, 2358–2364. doi: 10.1128/aem.61.6.2358-2364.1995

Cannella, D., Hsieh, C.-W. C., Felby, C., and Jørgensen, H. (2012). Production and effect of aldonic acids during enzymatic hydrolysis of lignocellulose at high dry matter content. *Biotechnol. Biofuels* 5:26. doi: 10.1186/1754-6834-5-26

Chundawat, S. P. S., Beckham, G. T., Himmel, M. E., and Dale, B. E. (2011). Deconstruction of lignocellulosic biomass to fuels and chemicals. *Ann. Rev. Chem Biomolecular Eng.* 2, 121–145.

Costa, T. H. F., Masarin, F., Bonifácio, T. O., Milagres, A. M. F., and Ferraz, A. (2013). The enzymatic recalcitrance of internodes of sugar cane hybrids with contrasting lignin contents. *Ind. Crop. Prod.* 51, 202–211. doi: 10.1016/j.indcrop.2013.08.078

Costa, T. H. F., Vega-Sanchez, M. E., Milagres, A. M. F., Scheller, H. V., and Ferraz, A. (2016). Tissue-specific distribution of hemicelluloses in six different sugarcane hybrids as related to cell wall recalcitrance. *Biotechnol. Biofuels* 9:99. doi: 10.1186/s13068-016-0513-512

Daly, P., Peng, M., Di Falco, M., Lipzen, A., Wang, M., Ng, V., et al. (2019). Glucose-mediated repression of plant biomass utilization in the white-rot

fungus *Dichomitus squalens*. *Appl. Environ. Microbiol.* 85:e01828-19. doi: 10.1128/AEM.01828-1819

Druzhinina, I. S., and Kubicek, C. P. (2017). Genetic engineering of *Trichoderma reesei* cellulases and their production. *Microbial Biotechnol.* 10, 1485–1499. doi: 10.1111/1751-7915.12726

Ellila, S., Fonseca, L., Uchima, C., Cota, J., Goldman, G. H., Saloheimo, M., et al. (2017). Development of a low-cost cellulase production process using *Trichoderma reesei* for Brazilian biorefineries. *Biotechnol Biofuels* 10:30. doi: 10.1186/s13068-017-0717-710

Eriksson, K. E., and Johnsrud, S. C. (1983). Mutants of the white-rot fungus *Sporotrichum pulverulentum* with increased cellulase and beta-D-glucosidase production. *Enzyme Microbial Technol.* 5, 425–429. doi: 10.1016/0141-0229(83)90024-8

Ferraz, A., Baeza, J., Rodriguez, J., and Freer, J. (2000). Estimating the chemical composition of biodegraded pine and eucalyptus wood by DRIFT spectroscopy and multivariate analysis. *Bioresour. Technol.* 74, 201–212. doi: 10.1016/s0960-8524(00)00024-9

Fitz, E., Wanka, F., and Seiboth, B. (2018). Promoter toolbox for recombinant gene expression in *Trichoderma reesei*. *Front. Bioeng. Biotechnol.* 6:135. doi: 10.3389/fbioe.2018.00135

Floudas, D., Binder, M., Riley, R., Barry, K., Blanchette, R. A., Henrissat, B., et al. (2012). The paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* 336, 1715–1719.

Ghose, T. K. (1987). Measurement of cellulase activities. *Pure Appl. Chem.* 59, 257–268. doi: 10.1351/pac198759020257

Jung, W., Sharma-Shivappa, R., Park, S., and Kolar, P. (2020). Effect of cellulolytic enzyme binding on lignin isolated from alkali and acid pretreated switchgrass on enzymatic hydrolysis. *3 Biotech* 10:1. doi: 10.1007/s13205-019-1978-z

Juturu, V., and Wu, J. C. (2014). Microbial cellulases: engineering, production and applications. *Ren. Sust. Ener. Rev.* 33, 188–203. doi: 10.1016/j.rser.2014.01.077

Kunamneni, A., Plou, F. J., Alcalde, M., and Ballesteros, A. (2014). "*Trichoderma Enzymes* for food industries," in *Biotechnology and Biology of Trichoderma*, Vol. 24, eds V. K. Gupta, M. Schmoll, A. Herrera-Estrella, R. S. Upadhyay, I. Druzhinina, and M. G. Tuohy (Oxford: Newnes), 339–344. doi: 10.1016/b978-0-444-59576-8.00024-2

Kurasin, M., and Väljamäe, P. (2011). Processivity of cellobiohydrolases is limited by the substrate. *J. Biol. Chem.* 286, 169–177. doi: 10.1074/jbc.m110.161059

Lange, L., Pilgaard, B., Herbst, F. A., Busk, P. K., Gleason, F., and Pedersen, A. G. (2019). Origin of fungal biomass degrading enzymes: evolution, diversity and function of enzymes of early lineage fungi. *Fungal Biol. Rev.* 33, 82–97. doi: 10.1016/j.fbr.2018.09.001

Levin, L., Villalba, L., Da Re, V., Forchiassin, F., and Papinutti, L. (2007). Comparative studies of loblolly pine biodegradation and enzyme production by Argentinean white rot fungi focused on biopulping processes. *Process Biochem.* 42, 995–1002. doi: 10.1016/j.procbio.2007.03.008

Lin, L., Wang, S., Li, X., He, Q., Benz, J. P., and Tian, C. (2019). STK-12 acts as a transcriptional brake to control the expression of cellulase-encoding genes in *Neurospora crassa*. *PLoS Genet.* 15:e1008510. doi: 10.1371/journal.pgen.1008510

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-((Ct method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Lu, X., Wang, C., Li, X., Zhao, J., and Zhao, X. (2017). Studying nonproductive adsorption ability and binding approach of cellobiohydrolase to lignin during bioconversion of lignocellulose. *Ener. Fuels* 31, 14393–14400. doi: 10.1021/acs.energyfuels.7b02427

Machado, A. S., and Ferraz, A. (2017). Biological pretreatment of sugarcane bagasse with basidiomycetes producing varied patterns of biodegradation. *Bioresource Technol.* 225, 17–22. doi: 10.1016/j.biortech.2016.11.053

Martinez, D., Larrondo, L. F., Putnam, N., Gelpke, M. D., Huang, K., Chapman, J., et al. (2004). Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. *Nat. Biotechnol.* 22, 695–700. doi: 10.1038/nbt967

Midorikawa, G. E. O., Correa, C. L., Noronha, E. F., Ferreira, E. X., Togawa, R. C., Costa, M. M. D., et al. (2018). Analysis of the transcriptome in *Aspergillus tamarii* during enzymatic degradation of sugarcane bagasse. *Front. Bioeng. Biotechnol.* 6:123. doi: 10.3389/fbioe.2018.00123

Momeni, M. H., Payne, C. M., Hansson, H., Mikkelsen, N. E., Svedberg, J., Engstrom, A., et al. (2013). Structural, biochemical, and computational characterization of the glycoside hydrolase family 7 cellobiohydrolase of the tree-killing fungus *Heterobasidion irregulare*. *J. Biol Chem.* 288, 5861–5872. doi: 10.1074/jbc.m112.440891

Munoz, I. G., Ubhayasekera, W., Henriksson, H., Szabo, I., Pettersson, G., Johansson, G., et al. (2001). Family 7 cellobiohydrolases from *Phanerochaete chrysosporium*: crystal structure of the catalytic module of Cel7D (CBH58) at 1.32 Å resolution and homology models of the isozymes. *J. Mol. Biol.* 314, 1097–1111. doi: 10.1006/jmbi.2000.5180

Ohm, R. A., Riley, R., Salamov, A., Min, B., Choi, I. G., and Grigoriev, I. V. (2014). Genomics of wood-degrading fungi. *Fungal Genet. Biol.* 72, 82–90. doi: 10.1016/j.fgb.2014.05.001

Payne, C. M., Knott, B. C., Mayes, H. B., Hansson, H., Himmel, M. E., Sandgren, M., et al. (2015). Fungal cellulases. *Chem. Rev.* 115, 1308–1448.

Petersen, P. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods.* 8, 785–786. doi: 10.1038/nmeth.1701

Peterson, R., and Nevalainen, H. (2012). *Trichoderma reesei* RUT-C30-thirty years of strain improvement. *Microbiol. SGM* 158, 58–68. doi: 10.1099/mic.0.054031-0

Presley, G. N., and Schilling, J. S. (2017). Distinct growth and secretome strategies for two taxonomically divergent brown rot fungi. *Appl. Environ. Microbiol.* 83:e02987-16. doi: 10.1128/AEM.02987-2916

Rahikainen, J. L., Martin-Sampedro, R., Heikkinen, H., Rovio, S., Marjamaa, K., Tamminen, T., et al. (2013). Inhibitory effect of lignin during cellulose bioconversion: the effect of lignin chemistry on non-productive enzyme adsorption. *Bioresour. Technol.* 133, 270–278. doi: 10.1016/j.biortech.2013.01.075

Ravalason, H., Jan, G., Molle, D., Pasco, M., Coutinho, P. M., Lapierre, C., et al. (2008). Secretome analysis of *Phanerochaete chrysosporium* strain CIRM-BRFM41 grown on softwood. *Appl. Microbiol. Biotechnol.* 80, 719–733.

Reinoso, F. A. M., Rencoret, J., Gutiérrez, A., Milagres, A. M. F., del Río, J. C., and Ferraz, A. (2018). Fate of p-hydroxycinnamates and structural characteristics of residual hemicelluloses and lignin during alkaline-sulfite chemithermomechanical pretreatment of sugarcane bagasse. *Biotechnol. Biofuels* 11:153. doi: 10.1186/s13068-018-1155-1153

Rytioja, J., Hilden, K., Yuzon, J., Hatakka, A., de Vries, R. P., and Makela, M. R. (2014). Plant-polysaccharide-degrading enzymes from basidiomycetes. *Microbiol. Mol. Biol. Rev.* 78, 614–649. doi: 10.1128/mmbr.00035-14

Santos, A. C., Ximenes, E., Kim, Y., and Ladisch, M. R. (2019). Lignin-enzyme interactions in the hydrolysis of lignocellulosic biomass. *Trends Biotechnol.* 37, 518–531. doi: 10.1016/j.tibtech.2018.10.010

Santos, V. T. O., Siqueira, G., Milagres, A. M. F., and Ferraz, A. (2018). Role of hemicellulose removal during dilute acid pretreatment on the cellulose accessibility and enzymatic hydrolysis of compositionally diverse sugarcane hybrids. *Ind. Crop. Prod.* 111, 722–730. doi: 10.1016/j.indcrop.2017.11.053

Siqueira, G., Arantes, V., Saddler, J. N., Ferraz, A., and Milagres, A. M. F. (2017). Limitation of cellulose accessibility and unproductive binding of cellulases by pretreated sugarcane bagasse lignin. *Biotechnol. Biofuels* 10:176. doi: 10.1186/s13068-017-0860-867

Strobel, K. L., Pfeiffer, K. A., Blanch, H. W., and Clark, D. S. (2016). Engineering Cel7A carbohydrate binding module and linker for reduced lignin inhibition. *Biotechnol. Bioeng.* 113, 1369–1374. doi: 10.1002/bit.25889

Suto, M., and Tomita, F. (2001). Induction and catabolite repression mechanisms of cellulase in fungi. *J. Biosci. Bioeng.* 92, 305–311. doi: 10.1263/jbb.92.305

Suzuki, H., Igarashi, K., and Samejima, M. (2009). Quantitative transcriptional analysis of the genes encoding glycoside hydrolase family 7 cellulase isozymes in the basidiomycete *Phanerochaete chrysosporium*. *FEMS Microbiol. Lett.* 299, 159–165. doi: 10.1111/j.1574-6968.2009.01753.x

Suzuki, H., Igarashi, K., and Samejima, M. (2010). Cellotriose and cellotetraose as inducers of the genes encoding cellobiohydrolases in the basidiomycete *Phanerochaete chrysosporium*. *Appl. Environ. Microbiol.* 76, 6164–6170. doi: 10.1128/aem.00724-10

Szabó, I. J., Johansson, G., and Pettersson, G. (1996). Optimized cellulase production by *Phanerochaete chrysosporium*: control of catabolite repression by

fed-batch cultivation. *J. Biotechnol.* 48, 221–230. doi: 10.1016/0168-1656(96)01512-x

Tachioka, M., Nakamura, A., Ishida, T., Igarashi, K., and Samejima, M. (2017). Crystal structure of a family 6 cellobiohydrolase from the basidiomycete *Phanerochaete chrysosporium*. *Acta Crystallogr. F: Struct. Biol. Commun.* 73, 398–403.

Tan, L. U. L., Mayers, P., and Saddler, J. N. (1987). Purification and characterization of a thermostable xylanase from a thermophilic fungus *Thermoascus aurantiacus*. *Can. J. Microbiol.* 33, 689–692. doi: 10.1139/m87-120

Tangnu, S. K., Blanch, H. W., and Wilke, C. R. (1981). Enhanced production of cellulase, hemicellulase, and beta-glucosidase by *Trichoderma-reesei* (RUT C-30). *Biotechnol. Bioeng.* 23, 1837–1849. doi: 10.1002/bit.260230811

Taylor, L. E., Knott, B. C., Baker, J. O., Alahuhta, P. M., Hobdey, S. E., Linger, J. G., et al. (2018). Engineering enhanced cellobiohydrolase activity. *Nat. Comm.* 9:1186. doi: 10.1038/s41467-018-03501-3508

Tsukada, T., Igarashi, K., Yoshida, M., and Samejima, M. (2006). Molecular cloning and characterization of two intracellular beta-glucosidases belonging to glycoside hydrolase family 1 from the basidiomycete *Phanerochaete chrysosporium*. *Appl. Microbiol. Biotechnol.* 73, 807–814. doi: 10.1007/s00253-006-0526-z

Uzcategui, E., Raices, M., Montesino, R., Johansson, G., Pettersson, G., and Eriksson, K. E. (1991). Pilot-scale production and purification of the cellulolytic enzyme-system from the white-rot fungus *Phanerochaete-chrysosporium*. *Biotechnol. Appl. Biochem.* 13, 323–334.

Valadares, F., Goncalves, T. A., Damasio, A., Milagres, A. M. F., Squina, F. M., Segato, F., et al. (2019). The secretome of two representative lignocellulose-decay basidiomycetes growing on sugarcane bagasse solid-state cultures. *Enzyme Microbial Technol* 130:UNS109370. doi: 10.1016/j.enzmictec.2019.109370

Valadares, F., Goncalves, T. A., Goncalves, D. S. P. O., Segato, F., Romanel, E., Milagres, A. M. F., et al. (2016). Exploring glycoside hydrolases and accessory proteins from wood decay fungi to enhance sugarcane bagasse saccharification. *Biotechnol. Biofuels.* 9:100. doi: 10.1186/s13068-016-0525-y

Varnai, A., Costa, T. H. F., Faulds, C. B., Milagres, A. M. F., Siika-Aho, M., and Ferraz, A. (2014). Effects of enzymatic removal of plant cell wall acylation (acetylation, p-coumaroylation, and feruloylation) on accessibility of cellulose and xylan in natural (non-pretreated) sugar cane fractions. *Biotechnol. Biofuels* 7:153. doi: 10.1186/s13068-014-0153-153

von Ossowski, I., Stahlberg, J., Koivula, A., Piens, K., Becker, D., Boer, H., et al. (2003). Engineering the exo-loop of *Trichoderma reesei* cellobiohydrolase, Cel7A. A comparison with *Phanerochaete chrysosporium Cel*7D. *J. Mol. Biol.* 333, 817–829. doi: 10.1016/s0022-2836(03)00881-7

Wood, T. M., and Bhat, K. M. (1988). Methods for measuring cellulase activities. *Meth. Enzymol.* 160, 87–112. doi: 10.1016/0076-6879(88)60109-1

Yang, J. Y., and Zhang, Y. (2015). I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res.* 43, W174–W181.

Zhang, L., Zhang, L. M., Zhou, T., Wu, Y. Y., and Xu, F. (2016). The dual effects of lignin content on enzymatic hydrolysis using film composed of cellulose and lignin as a structure model. *Bioresour. Technol.* 200, 761–769. doi: 10.1016/j.biortech.2015.10.048

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 30 July 2020
doi: 10.3389/fbioe.2020.00873

Comparative Characterization of *Aspergillus* Pectin Lyases by Discriminative Substrate Degradation Profiling

Birgitte Zeuner[1], Thore Bach Thomsen[1], Mary Ann Stringer[2], Kristian B. R. M. Krogh[2], Anne S. Meyer[1] and Jesper Holck[1]*

[1] Protein Chemistry and Enzyme Technology, Department of Biotechnology and Biomedicine, Technical University of Denmark, Lyngby, Denmark, [2] Novozymes A/S, Lyngby, Denmark

Fungal genomes often contain several copies of genes that encode carbohydrate active enzymes having similar activity. The copies usually have slight sequence variability, and it has been suggested that the multigenecity represents distinct reaction optima versions of the enzyme. Whether the copies represent differences in substrate attack proficiencies of the enzyme have rarely been considered. The genomes of *Aspergillus* species encode several pectin lyases (EC 4.2.2.10), which all belong to polysaccharide lyase subfamily PL1_4 in the CAZy database. The enzymes differ in terms of sequence identity and phylogeny, and exhibit structural differences near the active site in their homology models. These enzymes catalyze pectin degradation via eliminative cleavage of the α-(1,4) glycosidic linkages in homogalacturonan with a preference for linkages between methyl-esterified galacturonate residues. This study examines four different pectin lyases (PelB, PelC, PelD, and PelF) encoded by the same *Aspergillus* sp. (namely *A. luchuensis*), and further compares two PelA pectin lyases from two related *Aspergillus* spp. (*A. aculeatus* and *A. tubingensis*). We report the phylogeny, enzyme kinetics, and enzymatic degradation profiles of the enzymes' action on apple pectin, citrus pectin, and sugar beet pectin. All the pectin lyases exerted highest reaction rate on apple pectin [degree of methoxylation (DM) 69%, degree of acetylation (DAc) 2%] and lowest reaction rate on sugar beet pectin (DM 56%, DAc 19%). Activity comparison at pH 5–5.5 produced the following ranking: PelB > PelA > PelD > PelF > PelC. The evolution of homogalacturonan-oligomer product profiles during reaction was analyzed by liquid chromatography with mass spectrometry (LC-MS) detection. This analyses revealed subtle differences in the product profiles indicating distinct substrate degradation preferences amongst the enzymes, notably with regard to acetyl substitutions. The LC-MS product profiling analysis thus disclosed that the multigenecity appears to provide the fungus with additional substrate degradation versatility. This product profiling furthermore represents a novel approach to functionally compare pectin-degrading enzymes, which can help explain structure-function relations and reaction properties of disparate copies of carbohydrate active enzymes. A better understanding of the product profiles generated by pectin modifying enzymes has significant implications for targeted pectin modification in food and biorefinery processes.

Keywords: pectin lyase, multigenecity, sugar beet pectin, apple pectin, citrus pectin, product profiling

OPEN ACCESS

Edited by:
André Damasio,
State University of Campinas, Brazil

Reviewed by:
Noppol Leksawasdi,
Chiang Mai University, Thailand
Ashokkumar Balasubramaniem,
Madurai Kamaraj University, India
Breeanna Urbanowicz,
University of Georgia, United States

****Correspondence:***
Anne S. Meyer
asme@dtu.dk

Specialty section:
This article was submitted to Bioprocess Engineering, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *11 April 2020*
Accepted: *07 July 2020*
Published: *30 July 2020*

Citation:
Zeuner B, Thomsen TB, Stringer MA, Krogh KBRM, Meyer AS and Holck J (2020) Comparative Characterization of Aspergillus Pectin Lyases by Discriminative Substrate Degradation Profiling. Front. Bioeng. Biotechnol. 8:873. doi: 10.3389/fbioe.2020.00873

INTRODUCTION

Pectin lyases (EC 4.2.2.10) catalyze cleavage of the α-(1,4) glycosidic linkages between methyl-esterified galacturonic acid (GalA) units in the homogalacturonan backbone of pectin through a β-elimination reaction (**Figure 1**). The β-elimination leads to formation of a 4,5-unsaturated 6-*O*-methylated galacturonide molecule in the non-reducing end of one of the cleavage products, which has absorbance maximum at 235 nm. Pectin lyases thus prefer linkages between methyl-esterified (methoxylated) GalA units, but also catalyze bond cleavage at sites where only the GalA moiety in the +1 subsite is methoxylated, albeit with lower specific activity (Mutenda et al., 2002; Van Alebeek et al., 2002). Accordingly, pectin lyase activity is heavily dependent on the degree of methoxylation (DM) of the substrate, and also on the methyl ester distribution: A blockwise distribution of methoxylated GalA units results in higher pectin lyase activity than a random distribution (Mutenda et al., 2002; Van Alebeek et al., 2002). Detailed studies on *Aspergillus niger* pectin lyase A have further revealed requirement for methoxylated GalA in subsites +1 and +3, whereas subsites +2, +4, and -1 to -4 appear to allow accommodation of unesterified GalA units (Van Alebeek et al., 2002).

Important industrial applications of pectin lyases [usually employed in mixtures with other pectin depolymerizing enzymes, notably endo-polygalacturonases (EC 3.2.1.15)] include wine and juice pre-press maceration as well as juice clarification (Mantovani et al., 2005; Kassara et al., 2019). Uses of enzymes, including pectinase blends and pectin lyases, in food and beverage processing constitute about 30% of the total enzymes market, and a conservative estimate is that the global market value for pectin modifying enzymes is above US$ 100 million and growing. Pectin lyases are also actively investigated for special new uses such as bio-degumming and cleaning of bast fibers (Kohli and Gupta, 2019), and for production of distinct pectic oligosaccharides having prebiotic or anti-inflammatory properties (Chung et al., 2017; Tan et al., 2018). Most commercial pectinases originate from filamentous fungi, notably several derive from the *Aspergillus* section *Nigri* such as *Aspergillus niger* and *A. aculeatus* (Sandri et al., 2011). These fungi are specialized in secreting a broad spectrum of enzymes, including pectin-degrading enzymes, which can degrade and utilize the surrounding biomass (Benoit et al., 2012; de Vries et al., 2017; Kowalczyk et al., 2017).

Genome sequencing of the industrial enzyme-producing strain *A. niger* CBS 513.88 has revealed that this fungus contains multiple genes encoding for the same enzymatic activity (Pel et al., 2007; Andersen et al., 2012). Of the 131 genes encoding secreted polysaccharide-active enzymes, five (3.8%) encode pectin lyases; these five are denominated *pelA*, *pelB*, *pelC*, *pelD*, and *pelF* (Pel et al., 2007). Expression profiling of strains derived from *A. niger* CBS 120.49 has shown that the genes are active at different stages of pectin degradation, suggesting different regulatory mechanisms either related to extent of substrate degradation or to the increased acidification of the medium during pectinase action (de Vries et al., 2002). From studies on closely related *Aspergillus* species, it has furthermore become evident that the different species have different enzymatic strategies for efficient degradation of the same substrate (Benoit et al., 2015; Barrett et al., 2020b). Characterization of the individual gene products can therefore be used to efficiently select enzymes for a specific application, e.g., optimal degradation of a given substrate under a certain set of reaction conditions.

We hypothesized that the multiple copies (multigenecity) of genes encoding for pectin lyases in *Aspergilli* might be related to small differences in substrate attack preferences of the enzymes. Such differences could reflect adaptation to particular substitution patterns on the pectin backbone and materialize as minor variations in the active site region of the enzymes. This study was undertaken to examine PelB, PelC, PelD, and PelF from *A. luchuensis* in terms of substrate specificity, oligomer product profiles, pH-temperature optima, kinetic constants, active site and thermal stability. Two PelA pectin lyases from *A. aculeatus* and *A. tubingensis* were included for comparison. *A. tubingensis* and *A. luchuensis* both belong to the *A. niger* clade (Varga et al., 2011; Hong et al., 2013) and the pectin lyases studied here show > 90% sequence identity to the corresponding enzyme proteins from *A. niger* CBS 513.88 (**Table 1**). *A. aculeatus* is less closely related, yet all three *Aspergillus spp*. belong to section *Nigri* (Varga et al., 2011; Hong et al., 2013).

MATERIALS AND METHODS

Chemicals

Sugar beet pectin was provided by CP Kelco ApS (Lille Skensved, Denmark). Pectin from apple with 70–75% esterification (apple pectin), pectin from citrus peel (citrus pectin), and polygalacturonic acid and all other chemicals were purchased from Sigma-Aldrich (Merck) (Darmstadt, Germany).

Enzyme Expression and Purification

Genes encoding the full length native pectin lyases (**Table 1**) were PCR amplified using gene specific primers from either a cDNA pool from *Aspergillus tubingensis* or genomic DNA from *Aspergillus luchuensis* and *Aspergillus aculeatus*. The amplified full-length coding sequences were cloned into *Aspergillus* expression vectors and recombinantly expressed in *Aspergillus niger* for *At*PelA and *Aspergillus oryzae* for the remaining pectin lyases. The enzymes were purified to SDS-PAGE electrophoretic purity (**Supplementary Figure S1**) using hydrophobic interaction chromatography followed by ion exchange chromatography.

All pectin lyases are classified in the CAZy polysaccharide lyase family 1 (PL1). PL1 comprises pectin lyases (EC 4.2.2.10), pectate lyases (EC 4.2.2.2), and exo-pectate lyases (EC 4.2.2.9). Based on phylogenetic analysis of the catalytic domain, PL1 is currently divided into 13 subfamilies, which to some extent reflects substrate specificity (Lombard et al., 2010). Following CAZy nomenclature, PL1 subfamily X is written PL1_X. Characterized pectin lyases are found in PL1_4 and PL1_8, both of which are monospecific subfamilies comprising pectin lyases only. The six *Aspergillus* pectin lyases studied here all belong to PL1_4.

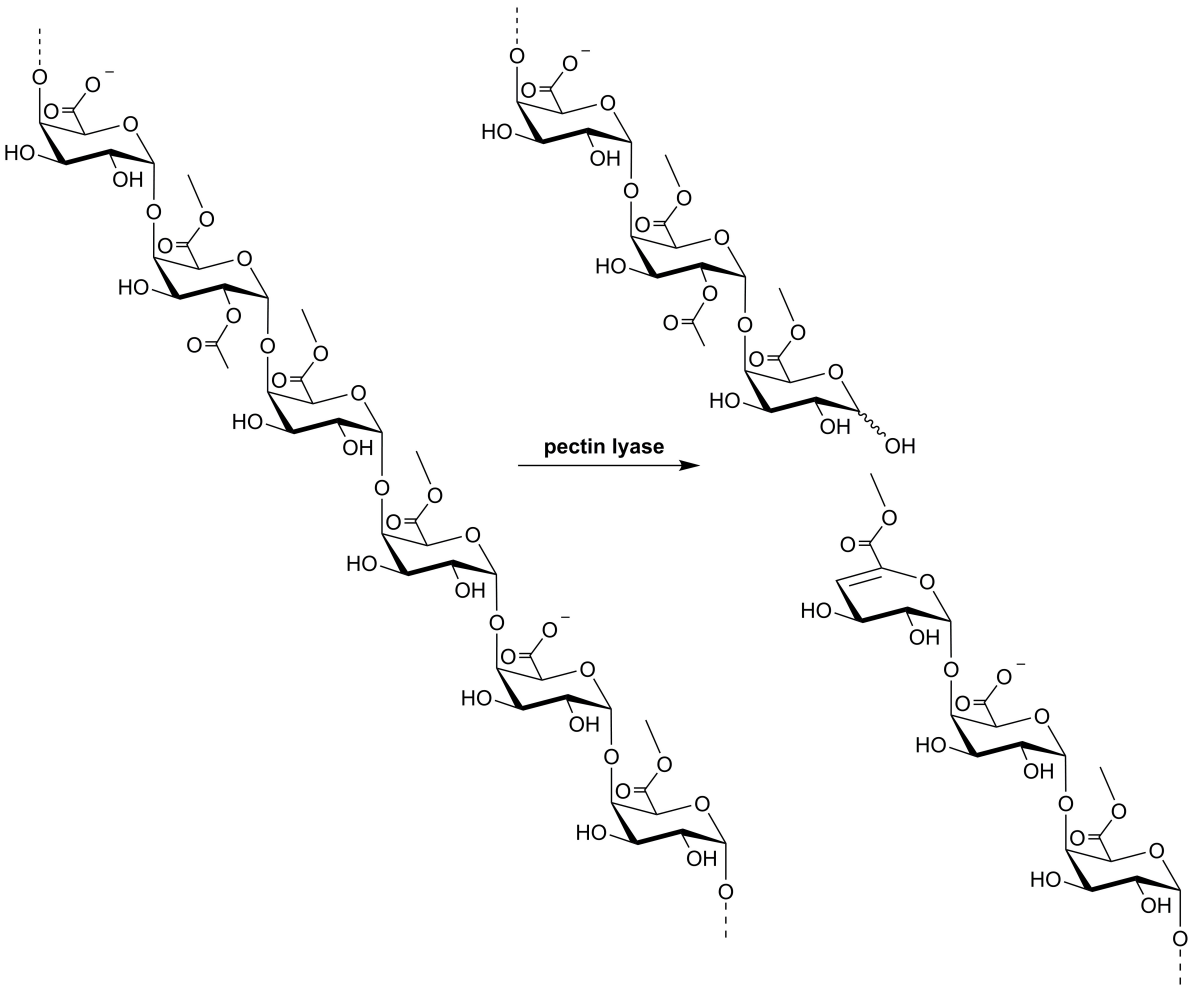


FIGURE 1 | Stylized illustration of a pectin lyase-catalyzed β-elimination reaction on pectin. GalA can be methoxylated on *O*-6 and acetylated on *O*-2 and/or *O*-3. The illustration depicts a pectin fragment substrate with a degree of methoxylation (DM) of 67% and a degree of acetylation (DAc) of 17%, whereas the true substitution degrees vary between pectins of different sources (**Table 2**). Note the preference for cleavage between methylesterified GalA moieties as outlined in more detail in the text.

TABLE 1 | Name, origin, and publicly available sequence entries of the *Aspergillus* sp. pectin lyases used in the current work.

Name	Origin[a]	Public sequence	Comparable to *A. niger* CBS 513.88	Seq. id.	MW (kDa)
*Aa*PelA	*A. aculeatus*	–	PelA (A2R3I1)	75%[b]	37.8
*At*PelA	*A. tubingensis*	Swiss-Prot A0A100IK89	PelA (A2R3I1)	98%	37.9
*Al*PelB	*A. luchuensis*	Swiss-Prot G7Y0I4	PelB (A2QFN7)	99%	37.9
*Al*PelC	*A. luchuensis*	Swiss-Prot G7XAF0	PelC (A2QW65)	91%	37.4
*Al*PelD	*A. luchuensis*	Swiss-Prot G7Y107	PelD (A2RBL2)	97%	37.2
*Al*PelF	*A. luchuensis*	Swiss-Prot G7XZ48	PelF (A2R6A1)	99%	47.3

The amino acid sequence of AaPelA is given in ***Supplementary Table S3****. Comparison and sequence identity of the pectin lyases to the commonly known Aspergillus niger CBS 513.88 pectin lyases (reviewed Swiss-Prot entry IDs indicated) are given. Calculated molecular weight (MW) after signal peptide removal as predicted by SignalP-5.0 server is also indicated. AtPelA was produced recombinantly in A. niger, the rest in A. oryzae.* [a]*Swiss-Prot entries G7Y0I4, G7XAF0, G7Y107, and G7XZ48 are listed with Aspergillus kawachii NBRC 4308 as origin. but it has been established that A. kawachii should be named A. luchuensis (Hong et al., 2013) and it is therefore named as such here. Swiss-Prot entry A0A100IK89 is listed as Aspergillus niger in UniProt, but was recently reclassified as A. tubingensis.* [b]*BLAST against Swiss-Prot gives highest sequence identity (83–84%) to PelAs from A. fischeri, A. fumigatus, and A. oryzae.*

In silico Studies

The sequences of the six pectin lyases studied in the current work were compared to the reviewed Swiss-Prot entries by BLAST analysis from UniProt. Phylogenetic trees were made in CLC Main Workbench (QIAGEN, Hilden, Germany) using all reviewed Swiss-Prot entries for pectin lyases and the sequences

of the pectin lyases studied in the current work. Homology models of the enzymes were generated using SWISS-MODEL (Waterhouse et al., 2018). Conserved amino acids in the substrate binding site were identified and colored using PyMOL (Schrödinger, New York, NY, United States). Homology models and crystal structures were aligned in PyMOL using the *super* function. Multiple sequence alignments were generated with ClustalW in MUSCLE (Madeira et al., 2019). Signal peptides were predicted using the SignalP 5.0 web server (Almagro Armenteros et al., 2019), and the theoretical molecular weight of the secreted enzymes were calculated from the amino acid sequence in CLC Main Workbench. Secondary structure prediction was performed with PSIPRED (Buchan and Jones, 2019). The CUPP webserver for functional annotation was accessed at cupp.info (Barrett and Lange, 2019; Barrett et al., 2020a).

Substrate Characterization

Apple pectin, citrus pectin, and sugar beet pectin substrates were subjected to acid hydrolysis with 4% sulfuric acid following the NREL method (Sluiter et al., 2012). The resulting monosaccharide composition was determined by HPAEC-PAD using a CarboPac PA1 column (Dionex, Sunnyvale, CA, United States) as described previously (Zeuner et al., 2018).

In order to determine degree of methoxylation (DM) and degree of acetylation (DAc) of the pectin substrates, methyl and acetyl substitutions were released by saponification as described previously (Voragen et al., 1986). The released methanol and acetate were quantified by HPLC (Shimadzu, Kyoto, Japan) essentially as described previously (Agger et al., 2010) using an Aminex HPX-87H ion exclusion column (BioRad, Hercules, CA, United States). Elution was performed over 40 min with 4 mM sulfuric acid at a flow rate of 0.6 mL/min and 63°C. Compounds were detected by refractive index detection using methanol and acetate as external standards. The DM is defined as the percentage of methoxylated GalA moieties out of the total amount of GalA moieties. Similarly, the DAc is defined as the percentage of acetyl substituted GalA units.

Substrate Screening and Enzyme Activity

The initial reaction rates of the pectin lyases were determined on four different pectin substrates: apple pectin, citrus pectin, sugar beet pectin, and polygalacturonic acid using a substrate concentration of 1 g/L in 100 mM sodium acetate buffer at pH 5. The enzyme dosage was 0.4% (w/w) E/S for *Aa*PelA, *At*PelA, and *Al*PelB, 12.5% (w/w) E/S for *Al*PelC, 0.625% (w/w) E/S for *Al*PelD, and 1.25% (w/w) E/S for *Al*PelF. The pectin lyase activity was monitored at 25°C for 10 min by measuring the change in absorbance at 235 nm using an Infinite 200Pro plate reader (Tecan, Männedorf, Switzerland). This change was converted to product formation using an extinction coefficient of 5500 M^{-1} cm^{-1} for the 4,5-unsaturated galacturonide product (Kester and Visser, 1994). One unit (U) is defined as the amount of enzyme which catalyzes the formation of one μmol of 4,5-unsaturated galacturonide product per min at pH 5.0 and 25°C. In turn, the reported specific activity is given as μmol product released per min per mg enzyme.

Determination of pH-Temperature Optima and Kinetic Constants

The pH-temperature optima of *Aa*PelA, *At*PelA, *Al*PelB, and *Al*PelD were determined by response surface methodology (RSM) using a two-factor face-centered central composite design of experiments (CCF) with three center points. Each sample was analyzed in duplicate. Statistical analysis, model generation, and graphic illustration was made in JMP 14 Pro (SAS, Cary, NC, United States). All reactions were performed with pectin solutions of 2 g/L where pH was adjusted by NaOH and HCl. For *At*PelA and *Al*PelD a pH range of 3.5–5.5 was employed for all substrates. For *Aa*PelA and *Al*PelB the pH range was 5.5–7.5, except for *Aa*PelA on sugar beet pectin, where it was 4.5–6.5. The tested temperature range was 45–65°C for *Aa*PelA, 50–70°C for *At*PelA, 40–60°C for *Al*PelB, and 40–70°C for *Al*PelD. Enzyme dosages ranged from 0.0125 to 0.05% E/S depending on the initial substrate screening results. Reactions took place in preheated substrate solution for 5 min. The reaction was stopped by heating at 95°C. Finally, the extent of reaction was determined from the difference in absorbance at 235 nm when compared to an enzyme-free solution of the same substrate.

Kinetic constants were determined using the same reaction setup at the pH-temperature optima predicted for each enzyme on each substrate based on the RSM results; for *Aa*PelA on sugar beet pectin where no model was generated, center point values (pH 5.5, 55°C) were used instead. Eight different pectin concentrations in the range from 0.25 to 10 g/L were used. The substrate concentration was converted to molar concentration of bonds between GalA moieties based on the monosaccharide composition (**Table 2**). Reactions were run in triplicates. Kinetic constants were determined from linear regression in a Hanes-Woolf plot of the data.

Thermal Stability Assessment

A thermal shift assay (TSA) (Lo et al., 2004) was used to determine the melting point (T_m) of the pectin lyases at pH 5. Purified pectin lyase samples were prepared for TSA by diluting to a standard concentration of 0.24 mg/mL in Milli-Q water. SYPRO Orange dye (S6650; Life Technologies, Carlsbad, CA, United States) was diluted 1:200 in assay buffer. The assay buffer was 0.1 M succinic acid, 0.1 M HEPES, 0.1 M CHES, 0.1 M CAPS, 0.15 M KCl, 1 mM $CaCl_2$, and 0.01 % Triton X100, adjusted to pH 5. For measurement, 10 μl of diluted enzyme sample was combined with 20 μl of diluted dye in the well of a TSA assay plate (Roche LightCycler 480 Multiwell plate 96, white; Roche, Basel, Switzerland) and the plate was covered with optic sealing foil. Thermal ramping and fluorescence measurements were run in a Roche Lightcycler 480 II machine (Roche, Basel, Switzerland) running with a ramp from 25 to 99°C at a rate of 200°C/h. The data collected was analyzed by Roche LightCycler 480 software (release 1.5.0 SP4). All samples were analyzed in duplicate and averaged.

For *Al*PelF, a different analysis method was used for determining T_m, namely nano-differential scanning fluorimetry (nDSF). The enzyme was diluted to 0.5 mg/mL and buffer exchanged to 50 mM acetic acid (pH 5.0) using a PD

TABLE 2 | Average monosaccharide composition in μmol/g pectin dry matter, degree of acetylation (DAc), and degree of methoxylation (DM) of pectin from apple, citrus, and sugar beet.

Pectin source	Monosaccharide composition (μmol/g pectin dry matter)									DAc	DM
	Fuc	Ara	Rha	Gal	Glc	Xyl	Man	GalA	GlcA		
Apple	6 ± 0.2^{a}	157 ± 3^{b}	110 ± 1^{c}	281 ± 6^{c}	237 ± 5^{a}	51 ± 1^{a}	5 ± 0.3^{a}	3875 ± 64^{b}	8 ± 1^{b}	2 ± 0.1^{b}	69 ± 3^{a}
Citrus	4 ± 0.2^{c}	98 ± 3^{c}	152 ± 5^{b}	478 ± 16^{b}	71 ± 2^{b}	15 ± 1^{b}	4 ± 0.3^{ab}	4195 ± 15^{a}	5 ± 1^{b}	1 ± 0.1^{b}	53 ± 1^{b}
Sugar beet	5 ± 0.1^{b}	512 ± 11^{a}	264 ± 5^{a}	532 ± 9^{a}	30 ± 1^{c}	11 ± 0.5^{c}	3 ± 0.4^{b}	2940 ± 8^{c}	33 ± 1^{a}	19 ± 1^{a}	56 ± 5^{b}

Standard deviations of triplicates are indicated. Superscript letters indicate significant difference ($p < 0.05$) across the substrates for each monosaccharide or degree of substitution.

SpinTrap G-25 column (GE Healthcare, Uppsala, Sweden). The nDSF experiments were done utilizing a Prometheus NT.48 (NanoTemper, Munich, Germany). The experiments were conducted from 20 to 95°C with a temperature gradient of 200°C/h. The transition temperatures (T_m values) were obtained from peak values derived from the first-derivative of the signal trace (350/330 nm fluorescence ratio or 330 nm fluorescence) using PR.ThermControl software (NanoTemper, Munich, Germany).

Product Profiling by LC-ESI-MS and SEC

The pectin degradation product profiles were studied for all six pectin lyases using 10 g/L pectin (apple, citrus, or sugar beet) in acetate buffer at pH 5.5 and 40°C. An equimolar enzyme dosage of 60 nM (corresponding to 0.023% E/S for *At*PelA) was used for all enzymes, but due to low activity enzyme dosages of 600 nM were also included for *Al*PelC and *Al*PelF. After 5 and 20 min, and 2 and 24 h, samples were taken out and the enzymes were inactivated by heating at 95°C for 10 min. The samples were centrifuged (14,000 *g*, 5 min) to remove insoluble particles.

For size exclusion chromatography (SEC) analysis of the pectin degradation profiles, the samples were diluted 4 times in the eluent buffer and filtered (0.22 μm). The SEC analysis was performed by injecting on a TSKgel G3000PW column (300 mm × 7.5 mm) equipped with a TSKgel PWH guard column (7.5 mm × 7.5 mm) (Tosoh Bioscience, Tokyo, Japan) using an UltiMate iso-3100 SD pump and an RI-101 refractive index detector. Elution took place at 40°C using 0.1 M $NaNO_3$ with 0.02% NaN_3 as eluent and a flow rate of 1 mL/min. Pullulan standards (180 Da to 110 kDa) were used as reference.

Identification and relative quantification of pectic oligosaccharides was performed by liquid chromatography electrospray ionization mass spectrometry (LC-ESI-MS) on an Amazon SL iontrap (Bruker Daltonics, Bremen Germany) coupled to an UltiMate 3000 UHPLC from Dionex (Sunnyvale, CA, United States). Samples of 10 μL were injected on a porous graphitized carbon column (Hypercarb PGC, 150 mm × 2.1 mm, 3 μm; Thermo Fisher Scientific, Waltham, MA, United States). The chromatography was performed at 0.4 mL/min at 70 °C on a three-eluent system with eluent A (water), eluent B (acetonitrile), and eluent C (100 mM ammonium acetate pH 5). The elution profile was as follows: 0–1 min, 0% B; 1–30 min, linear gradient to 30% B; 30–35 min, isocratic 30% B; 35–40 min, isocratic 0% B. In addition, a constant level of 10% C was maintained throughout the elution. For every four samples, a cleaning procedure of 0–15 min, isocratic 50% B and C; 15–40 min, isocratic 10% C was performed to avoid build-up of polymeric pectin. The electrospray was operated in negative or positive mode with UltraScan mode and a scan range from 100 to 2000 *m*/*z*, smart parameter setting of 500 *m*/*z*, capillary voltage at 4.5 kV, end plate off-set 0.5 kV, nebulizer pressure at 3.0 bar, dry gas flow at 12.0 L/min, and dry gas temperature at 280°C. Identification of observed compounds by *m*/*z* and MS^2 fragmentation pattern was performed in DataAnalysis 4.2 and relative quantification based on intensity peaks in positive mode was performed in Compass QuantAnalysis 2.2 (Bruker Daltonics, Bremen Germany).

All samples were investigated with respect to their content of all confirmed and putative compounds. The intensity of each compound was defined as the area under the curve of a given extracted ion chromatogram. In order to investigate the product profile, or fingerprint, of each enzyme rather than their overall activity/productivity, all intensities were normalized with respect to the most abundant compound in a given enzyme-substrate-time combination. Data was clustered by hierarchical clustering with complete linkage on the euclidian distance matrix and visualized in the pheatmap-package using R version 3.6.0.

A representative sample (*At*PelA, sugar beet pectin, 24 h) was selected for validation of the method. Three consecutive injections were made for technical replicates. The average coefficient of variation (CV) of the normalized values for all compounds was 9.0% (±5.5%) with no observed linearity between CV% and compound average (data not shown).

Statistics

One-way ANOVA for determination of statistical significance was performed with JMP 14 Pro (SAS, Cary, NC, United States). Statistical significance was established at $p < 0.05$.

RESULTS AND DISCUSSION

Pectin Lyase Identity, Phylogeny, and Homology Modeling

The multigenecity of pectinolytic enzymes from *Aspergilli* has been the subject of several studies over the last three decades (Harmsen et al., 1990; Kusters-van Someren et al., 1992; de Vries et al., 2002; Martens-Uzunova and Schaap, 2009; Benoit et al., 2012; Kowalczyk et al., 2017; He et al., 2018). Differences in pH optima (**Supplementary Table S1**) and expression inducers may partially explain the large gene families observed for several

different classes of pectinolytic enzymes (Martens-Uzunova and Schaap, 2009; Andersen et al., 2012), but there is a need for rigorous biochemical characterization of the gene products including specificity toward pectins of different origin and in particular determination of product profiles in order to fully understand this multigenecity. In fact, several cases exist where one of the genes in a pectinolytic gene family is present in one *Aspergillus niger* strain, but absent from another (Martens-Uzunova and Schaap, 2009). For instance, the genome of *A. niger* ATCC 1015 harbors a putatively pectin lyase-encoding *pelE* gene, but there is no evidence of such a gene in the genome of *A. niger* CBS 513.88 (Martens-Uzunova and Schaap, 2009). The presence of *pelE* in *A. niger* CBS 120.49 was described 30 years ago, although some doubt was raised about its actual function (Harmsen et al., 1990). The genomes of the *A. tubingensis* and *A. luchuensis* strains from which pectin lyases were sourced in the current work also contain *pelE* genes (UniProt entries A0A100I604 and G7XZW9, respectively), but no reports on PelE expression or activity of the pectin lyases from these strains exist. In a recent study, all five pectin lyase genes of *A. niger* CBS 513.88 – *pelA*, *pelB*, *pelC*, *pelD*, and *pelF* – were cloned and overexpressed in *A. niger* SH-2 (He et al., 2018). Of these, the *pelA*-recombinant strain exhibited the highest pectin lyase activity on apple pectin (DM 50–75%), and thus only PelA was selected for characterization.

In the current work, we aimed to express and systematically characterize representatives of *Aspergillus* sp. PelA, PelB, PelC, PelD, and PelF. The PL1_4 pectin lyases selected for the current work originated from three different strains which all belong to the *Aspergillus* section *Nigri*: *A. aculeatus, A. tubingensis,* and *A. luchuensis* (**Table 1**). The two latter strains were originally both assigned as *A. niger*, but subsequently determined via sequence analysis to in fact be the very closely related species *A. tubingensis* and *A. luchuensis*. Indeed, *A. tubingensis* and *A. luchuensis* both belong to the *A. niger* clade (Varga et al., 2011; Hong et al., 2013). We attempted to express *Al*PelA recombinantly in *A. oryzae*, but did not succeed. The PelA from *A. tubingensis* (which we named *At*PelA for the current work) is 99% identical to the PelA of *A. luchuensis* (*Al*PelA; UniProt G7XV55) and is therefore useful as a PelA representative to compare to PelB, PelC, PelD, and PelF from *A. luchuensis*. The UniProt entries list different *Aspergillus* species as origins of the employed pectin lyases, e.g., *A. kawachii* which has since been replaced by *A. luchuensis* (Hong et al., 2013). Together, these shifting taxonomic designations indicate that assigning exact *Aspergillus* species classification is a complex matter. In the current work, we have named the pectin lyases according to their most recent species classification as well as to their identity with the *A. niger* CBS 513.88 pectin lyases named PelA, PelB, PelC, PelD, and PelF (**Table 1**).

In order to allow comparison of the current work to literature, the six PL1_4 pectin lyases were compared to Swiss-Prot pectin lyase sequences in terms of sequence identity (**Table 1**) and phylogeny (**Figure 2**). A clear distinction between bacterial and fungal pectin lyases is observed (**Figure 2**). *At*PelA, *Aa*PelA, *Al*PelD, and *Al*PelB cluster in the same fungal cluster (FC1), whereas *Al*PelC and *Al*PelF cluster in each their fungal cluster (FC2 and FC3, respectively, **Figure 2**). This division is also evident from their pairwise sequence comparisons: *Al*PelC and *Al*PelF are clearly different from the members of FC1 and also different from each other (sequence identities $\leq$ 50%; **Supplementary Figure S2**). The putative PelE from *A. niger* CBS 120.49 (SwissProt B3GQR3) clusters at the edge of FC2 containing the PelC representatives, which have not been deeply characterized either (**Figure 2** and **Supplementary Table S1**). The phylogenetic tree emphasizes the close relatedness of the pectin lyases used in the current work to those of *A. niger* CBS 513.88 and CBS 120.49 **Figure 2**); this is also apparent from their high sequence identity (**Table 1**). Furthermore, the high similarity between *At*PelA and PelA from *A. luchuensis* (*Al*PelA) is also clear from the phylogenetic analysis (**Figure 2**). As expected, *Aa*PelA clusters with *At*PelA and PelA (**Figure 2**, **Table 1**, and **Supplementary Figure S2**). The decision to include *Aa*PelA and no other *A. aculeatus* PL1_4 enzymes in the study reflects the fact that PelA is the most studied enzyme of the five, possibly due to its high activity and expression levels (He et al., 2018).

Two PL1_4 subfamily members have resolved crystal structures: *A. niger* PelA (PDB IDs 1IDK and 1IDJ; Mayans et al., 1997), which has 99% sequence identity to *At*PelA, and *A. niger* PelB (PDB ID 1QCX; Vitali et al., 1998), which has 98% sequence identity to *Al*PelB. Structurally, the two pectin lyases appear highly similar. Neither contain ligand structures, but several crystal structures of bacterial PL1 pectate lyases from other subfamilies do. Together, this enables homology modeling and identification of the active site in the PL1_4 pectin lyases studied here. From the homology models it is evident that the three-dimensional structures of *Aa*PelA, *At*PelA, *Al*PelB, and *Al*PelD are indeed very similar (**Figure 3**). In contrast, *Al*PelC and *Al*PelF appear to exhibit a more open conformation around the active site and the homology models suggest that *Al*PelC and *Al*PelF have a more disordered and open loop in this position instead of the short α-helix observed in the crystal structures of PelA and PelB as well as in the models of *Aa*PelA, *At*PelA, *Al*PelB, and *Al*PelD (**Figure 3**). From the multiple sequence alignment it is evident that *Al*PelC and *Al*PelF have shorter loops in this region (marked in blue in **Supplementary Figure S3**). This loop contains a conserved substrate-interacting Trp residue corresponding to Trp81 in *A. niger* PelA and PelB (Mayans et al., 1997; Vitali et al., 1998; Herron et al., 2002). In the *Al*PelC homology model this Trp is completely aligned with that of *At*PelA, whereas the Trp side chain points in the opposite direction in the *Al*PelF model (**Figure 3**). For a neighboring loop above the active site (marked in gray in **Supplementary Figure S3**), *Al*PelC is similar to the other pectin lyases, whereas *Al*PelF exhibits a longer, disordered loop in this position (**Figure 3**). Importantly, this loop contains a Trp residue (Trp66 in *A. niger* PelA and PelB), which interacts with the methoxyl group of GalA in subsite +3 (Mayans et al., 1997; Herron et al., 2002). This Trp residue is generally conserved, but absent from *Al*PelF (**Figure 3** and **Supplementary Figure S3**). Furthermore, *Al*PelF is larger than the other five pectin lyases, which are almost identical in predicted size (glycosylations excluded; **Table 1**). This larger size is due to a C-terminal stretch found in *Al*PelF (**Supplementary Figure S3**), which according to a BLAST analysis appears to be unique to PelF from various *Aspergillus* species. The function

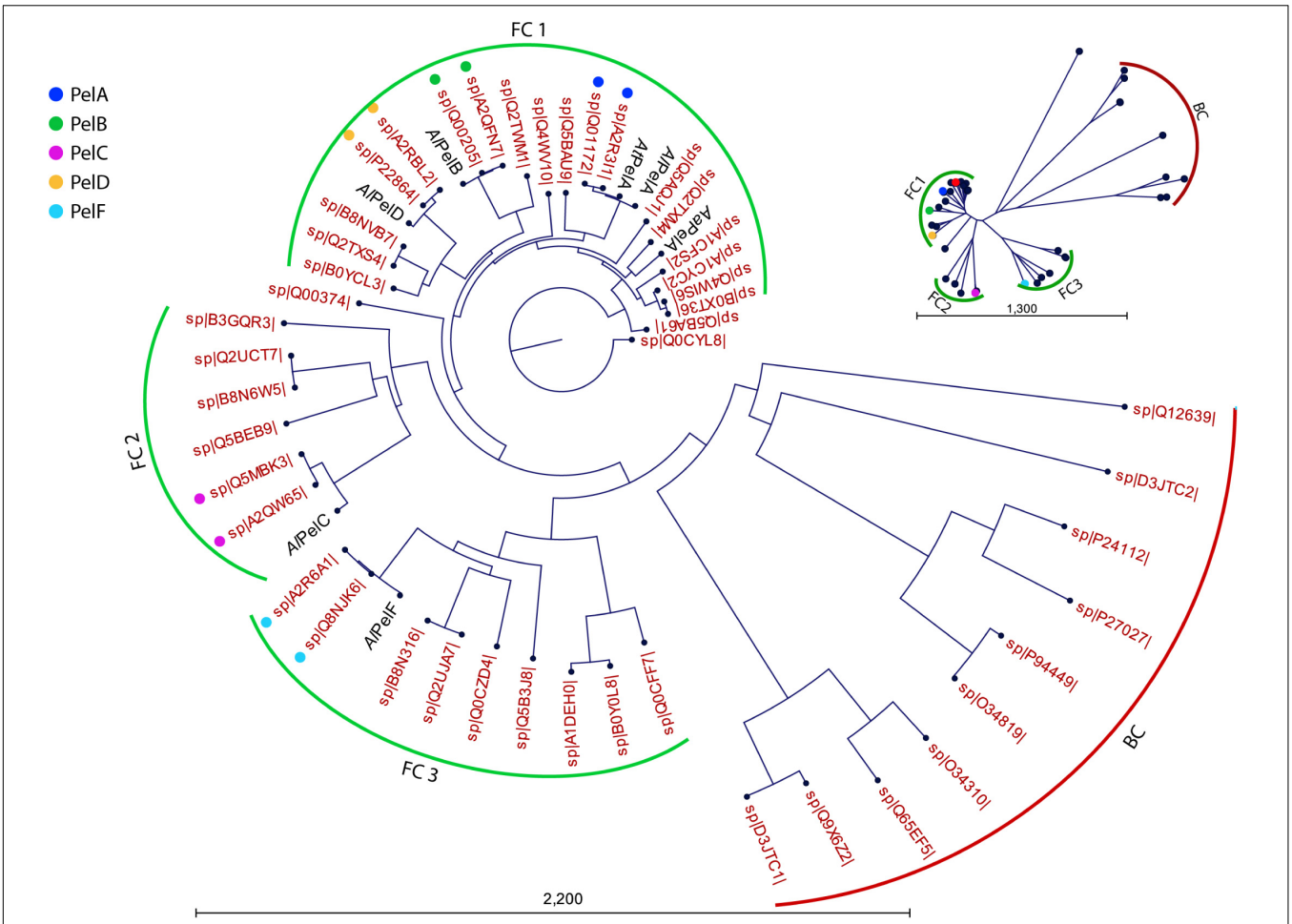


FIGURE 2 | Phylogenetic tree of pectin lyases included in the current study shown versus all reviewed Swiss-Prot entries (red labels indicating Swiss-Prot accession numbers). From the radial tree (top right), three clusters were identified: a bacterial cluster (BC, red) and three fungal clusters (FC1-FC3, green). Pectin lyases PelA, PelB, PelC, PelD, and PelF from *A. niger* strains CBS 513.88 and CBS 120.49 are indicated with colored dots. The UniProt database does not indicate CBS 120.49 in these entries, but tracking the submissions they all originate from *A. niger* CBS 120.49. *Al*PelA from the same *A. luchuensis* strain as *Al*PelB, *Al*PelC, *Al*PelD, and *Al*PelF is indicated to visualize its close relatedness to *At*PelA studied here.

of this C-terminal stretch is unknown; it could not be modeled by homology modeling, and secondary structure prediction suggested the entire structure to be a random coil. In conclusion, *Al*PelF differs most from the generally similar structures of the other pectin lyases. This assessment of the six pectin lyases show that there is good agreement between the phylogenetic analysis and the three-dimensional enzyme structure (**Figures 2**, **3**).

Using the CUPP webserver for functional annotation (Barrett and Lange, 2019; Barrett et al., 2020a) confirmed that all the included enzymes are predicted to belong to PL1_4. *Aa*PelA, *At*PelA, *Al*PelB, *Al*PelD, and *Al*PelF were all assigned EC 4.2.2.10 and CUPP group PL1:12.1[1] (Barrett et al., 2020a). In contrast, *Al*PelC could not be assigned to a CUPP group due to a low score. Although *Al*PelF differs most in terms of structure (**Figure 3**), CUPP singled out *AlPelC* as being more different. The detailed grouping provided by CUPP relies on an unsupervised peptide-based clustering algorithm that can provide systematic grouping and functional annotation of carbohydrate-active enzymes (CAZymes) based on comparison of peptide motifs in the enzyme amino acid sequences (Barrett and Lange, 2019). This emphasizes that *Al*PelC may be functionally different from the other pectin lyases studied here.

Characterization of Pectin Substrates

Quantitatively, the major constituent of pectin is homogalacturonan followed by rhamnogalacturonan I (RG-I). Homogalacturonan consists of an unbranched chain of α-(1,4)-linked GalA units, which may be methoxylated and/or acetylated. Homogalacturonan can also be further substituted to form other less abundant pectin components, namely xylogalacturonan, apiogalacturonan, and rhamnogalacturonan II (RG-II). RG-I is made up of a backbone of alternating α-(1,2)-linked rhamnose and α-(1,4)-linked GalA units, where the rhamnose units are substituted with neutral sidechains of

[1]https://cupp.info/cupp/PL1:12.1

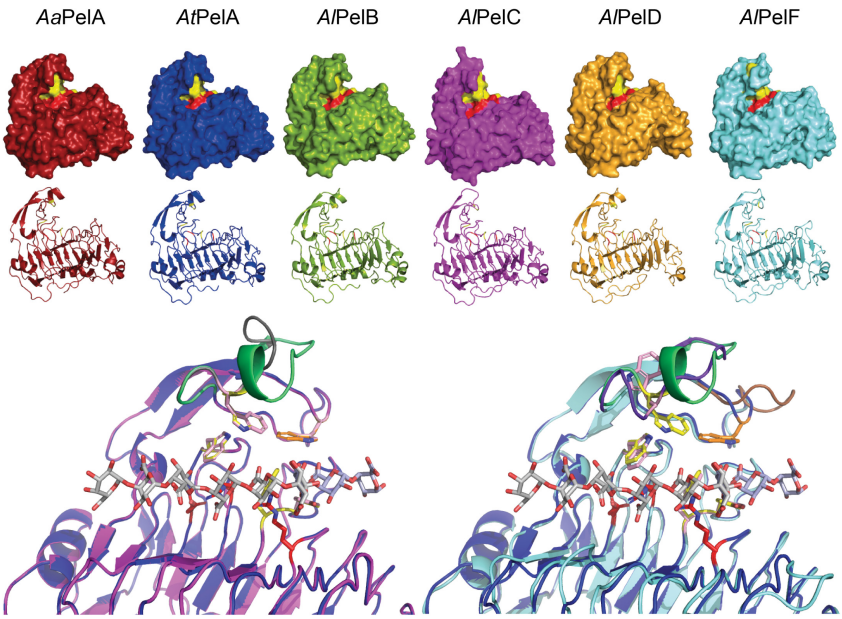


FIGURE 3 | *Top:* Homology models of the six PL1_4 pectin lyases studied here. Three charged residues believed to hold the catalytic function (one Asp and two Arg residues) are indicated in red, while aromatic residues which interact with the substrate and maintain the active site architecture are indicated in yellow (Mayans et al., 1997; Herron et al., 2002; Van Alebeek et al., 2002). The C-terminals of *Al*PelD and *Al*PelF could not be modeled and are thus not visible. *Bottom:* Active site zoom in structural alignments of homology models of *At*PelA (blue) and *Al*PelC (magenta) and of *At*PelA and *Al*PelF (cyan), respectively. The pectate oligomer ligands (gray and light blue) are derived from alignment with *Bacillus subtilis* pectate lyase (PDB IDs: 3KRG and 5AMV) (Seyedarabi et al., 2010; Ali et al., 2015). Assumed catalytic residues (red) are indicated alongside substrate interacting residues (yellow in *At*PelA, light pink in *Al*PelC and *Al*PelF). The substrate-interacting Trp86 in *At*PelA (corresponding to Trp66 in crystal structures of *A. niger* PelA and PelB (Mayans et al., 1997; Vitali et al., 1998; Herron et al., 2002), which were numbered without the 20-amino acid signal peptide) is shown in orange to highlight its absence in *Al*PelF, where this differing loop region is shown in brown (**Supplementary Figure S3**). The loop above the active site, which is four amino acids shorter in *Al*PelC and *Al*PelF and contains the substrate-interacting Trp residue corresponding to Trp81 in *A. niger* PelA and PelB (Mayans et al., 1997; Vitali et al., 1998; Herron et al., 2002), is highlighted in green in *At*PelA, gray in *Al*PelC, and purple in *Al*PelF. Alignment of *Al*PelF and *Al*PelC with *Al*PelB is highly similar to that with *At*PelA (not shown).

galactan, arabinan, and/or different arabino-galactan sidechains (Bonnin et al., 2014).

Compared to apple and citrus pectins, sugar beet pectin is known to have shorter homogalacturonan stretches, more RG-I, and a higher degree of acetylation (DAc) (Buchholt et al., 2004). This was reflected in the monosaccharide composition observed in the current work: sugar beet pectin contained more rhamnose, arabinose, and galactose (indicators of RG-I) and less GalA than apple and citrus pectins (**Table 2**). The DAc was 19% for sugar beet pectin and only 1% and 2% for citrus and apple pectins, respectively (**Table 2**). Comparing apple and citrus pectins, citrus pectin had slightly more GalA. Since pectin lyase prefers methylesterified GalA, the degree of methoxylation (DM) is an important parameter: Apple pectin had the highest DM of 69%, whereas there was no significant difference between sugar beet pectin (56%) and citrus pectin (53%; **Table 2**). This is in agreement with previous reports on DM for these pectin sources (Müller-Maatsch et al., 2016). Not accounting for other factors such as methyl group distribution, the pectin lyases were thus expected to have the highest affinity for apple pectin due to the high DM. Similar DMs were observed for citrus pectin and sugar beet pectin, but due to the significantly higher DAc observed in sugar beet pectin, it was expected that the pectin lyases would exhibit the lowest affinity for sugar beet pectin.

Enzyme Activity on Different Pectins

The specific activity of the six pectin lyases was determined as initial rates obtained at 25°C in 1 g/L solutions of apple pectin (DM 69%, DAc 2%), citrus pectin (DM 53%, DAc 1%), sugar beet pectin (DM 56%, DAc 19%), and polygalacturonic acid at pH 5.0 (**Table 3**). The highest initial rate was obtained with *Al*PelB on apple pectin. Indeed, all the pectin lyases exhibited highest activity on apple pectin, which is likely linked to the high DM of this pectin substrate. In general, the following substrate preference was observed: apple pectin > citrus pectin > sugar beet pectin > polygalacturonic acid, which reflects both DM and DAc of the substrates (**Table 3**). As expected, the difference in DM appeared to determine the difference in preference between apple and citrus pectins (**Tables 2**, 3), whilst the markedly higher DAc of sugar beet pectin perhaps in combination with a higher amount of RG-I (**Table 2**) impaired pectin lyase activity on sugar beet pectin. Finally, the consistently low activities of all pectin lyases on polygalacturonic acid (**Table 3**) confirm the requirement for methoxylation to achieve pectin lyase activity. The same trend was recently reported for PelA from *A. niger* CBS 513.88, which had highest activity on citrus pectin (DM ≥ 85%), a relative activity of 59% compared to citrus pectin on a sample of apple pectin with a DM of 50–75%, and almost no activity on polygalacturonic acid (He et al., 2018).

TABLE 3 | Specific enzyme activity (μmol product released per min per mg enzyme) determined from initial rates of the six pectin lyases on four different pectin substrates at pH 5 and 25°C: Apple pectin, citrus pectin, sugar beet pectin, and polygalacturonic acid; c.f. **Table 2** for composition of apple, citrus, and sugar beet pectins.

	Apple pectin	Citrus pectin	Sugar beet pectin	Polygalacturonic acid
*Aa*PelA	3.9 ± 0.2^{d}	2.5 ± 0.1^{e}	1.1 ± 0.05^{f}	0.06 ± 0.03^{i}
*At*PelA	8.7 ± 0.3^{b}	2.4 ± 0.1^{e}	0.7 ± 0.04^{g}	0.07 ± 0.02^{i}
*Al*PelB	9.1 ± 0.2^{a}	5.2 ± 0.2^{c}	0.8 ± 0.06^{g}	0.06 ± 0.06^{i}
*Al*PelC	$\leq 0.0005^{i}$	$\leq 0.0005^{i}$	$\leq 0.0005^{i}$	$\leq 0.0005^{i}$
*Al*PelD	0.34 ± 0.06^{h}	0.08 ± 0.01^{i}	0.008 ± 0.007^{i}	0.03 ± 0.03^{i}
*Al*PelF	0.06 ± 0.01^{i}	0.02 ± 0.01^{i}	0.04 ± 0.01^{i}	0.02 ± 0.01^{i}

Superscript letters indicate significant difference ($p < 0.05$) across the entire dataset.

Among the six pectin lyases, *Al*PelB had the highest initial rate on both apple pectin and citrus pectin, whereas *Aa*PelA had highest activity on sugar beet pectin (**Table 3**). *At*PelA had higher activity on apple pectin than *Aa*PelA, but the opposite was observed for sugar beet pectin (**Table 3**), and the two PelA pectin lyases exhibited similar activity on citrus pectin and polygalacturonic acid. For apple and citrus pectins, the following general ranking in specific activity was observed: PelB > PelA > > PelD > PelF > PelC (**Table 3**). The activity of *Al*PelC was at the detection limit despite using a high enzyme dosage. For polygalacturonic acid there was no significant difference between the very low specific activities, indicating that the enzymes do not have pectate lyase activity.

In a recent study, the fermentation broths of *A. niger* SH-2 used for recombinant expression of *A. niger* CBS 513.88 PelA, PelB, PelC, PelD, and PelF were tested on apple pectin (DM 50–75%) at pH 5.2. The enzyme activities in U/mL was PelA > PelD > > PelF > PelB, while PelC could not be expressed. However, these activities also reflect the expression levels, and especially PelB was expressed in lower amounts than the other pectin lyases (He et al., 2018). It was previously reported that *A. niger* CBS 120.49 PelB was more prone to protease degradation than PelA of the same strain when expressed homologously (Kusters-van Someren et al., 1992). As a consequence, it may not be possible to detect the high specific activity of PelB without controlling the enzyme dosage.

pH-Temperature Optima and Thermal Stability

Data on pH and temperature optima of pectin lyases from *Aspergillus* sp. are scarce (**Supplementary Table S1**). Recently, the optima for PelA from *A. niger* CBS 513.88 were determined separately on citrus pectin (DM ≥ 85%) to pH 4.5 and 50°C (He et al., 2018). Previously, PelA from *A. niger* (identified as PLII in the commercial Ultrazym preparation) was found to have a slightly substrate concentration-dependent optimum around pH 5.5–6.5 when determined at 25°C on apple pectin with a DM of 94% (van Houdenhoven, 1975; Kester and Visser, 1994). The same was true for Ultrazym PLI, which was found to be PelD from *A. niger* (van Houdenhoven, 1975; Gysler et al., 1990; Kester and Visser, 1994). PelF from *A. niger* ZJF also exhibited an acidic pH optimum of 5.0 and a temperature optimum of 43°C on citrus pectin (Xu et al., 2015). In contrast, PelB from *A. niger* CBS 120.49 had an alkaline optimum at pH 8.5 when determined at 25°C (Kester and Visser, 1994).

As pH optima are temperature dependent, we set out to estimate combined pH-temperature optima for the four most active pectin lyase in the current study using response surface methodology. Based on preliminary studies, acidic pH ranges (pH 3.5–5.5) were selected for *At*PelA and *Al*PelD, whereas a range from slightly acidic to neutral pH (pH 5.5–7.5) was selected for *Aa*PelA and *Al*PelB. Temperatures ranged from 40°C to 70°C (**Table 4**) and the reaction was monitored for 5 min. All the generated models had R^2 above 0.92, and most of them above 0.98, indicating that the modeled response surfaces described the data well (**Supplementary Table S2** and **Supplementary Figure S4**). For most models, the lack of fit was statistically significant (**Supplementary Table S2**). This indicates that the generated models are not sufficient for predicting activity at the identified optima, and is a result of the fact that the model assumes that the effect of pH and temperature can be described by quadratic polynomials. This is, however, not always the case for pH and temperature optima, which often follow non-symmetrical bell-shaped curves. For a few of the models, one of the main factors (T or pH) was not significant, but they were kept in the model in order to predict a pH-temperature optimum for the four pectin lyases on three different pectins (**Table 4**, **Supplementary Table S2**, and **Supplementary Figure S4**). For *Aa*PelA on sugar beet pectin, no model was obtained as the differences between the data points in the design were not statistically significant within the chosen ranges (**Supplementary Table S2** and **Supplementary Figure S4**).

In general, *At*PelA and *Al*PelD exhibited pH-temperature optima, which were a combination of high temperature and low pH: pH 4.4–4.9 and 57–59°C for *At*PelA and pH 4.3–4.4 and 59–60°C for *Al*PelD (**Table 4**). However, for *Al*PelD on citrus and sugar beet pectins, T was not a significant factor, indicating low variability across the range from 40 to 70°C (**Supplementary Figure S4**). The same was true for pH for AtPelA (**Table 4** and **Supplementary Table S2**). For *Aa*PelA and *At*PelB, the pH-temperature optima were a combination of lower temperature and higher pH: pH 6.1–6.4 and 54–55°C for *Aa*PelA and pH 6.4–6.8 and ≤ 40–49°C for *Al*PelB (**Table 4**). These trends are fairly well in line with previously reported optima for *A. niger* pectin lyases (**Supplementary Table S1**), and can partially explain the pectin lyase multigenecity in *Aspergilli*. For *Al*PelB we observe a lower (yet high in comparison with the other pectin lyases) pH optimum than found for PelB from *A. niger* CBS 120.49 (Kester and Visser, 1994). This is most likely tied to the fact that we have included an elevated temperature (40–60°C), whereas the previously established pH optimum of 8.5 was determined at 25°C, where PelB is more stable. Remarkably, the fact that the two PelA pectin lyases included here differ in their pH-temperature optima emphasize that results obtained with a given pectin lyase from one *Aspergillus* strain is not directly transferable to one of another strain. However, this effect may not be as dramatic within the *A. niger* clade where sequence similarities above 90% are

TABLE 4 | Combined pH-temperature optima on apple pectin, citrus pectin, and sugar beet pectin predicted from the regression models generated by the two-factor face-centered central composite experimental design (**Supplementary Table S2**).

	T_{opt} [°C]				pH_{opt}			
	Test interval	**Apple pectin**	**Citrus pectin**	**Sugar beet pectin**	**Test interval**	**Apple pectin**	**Citrus pectin**	**Sugar beet pectin**
*Aa*PelA	45–65°C	55	54*	–	pH 5.5–7.5	6.4*	6.1*	–
*At*PelA	50–70°C	57*	58*	59*	pH 3.5–5.5	4.9*	4.5	4.4
*Al*PelB	40–60°C	40*	49*	48	pH 5.5–7.5	6.8*	6.4*	6.7*
*Al*PelD	40–70°C	59*	60	60	pH 3.5–5.5	4.9*	4.3*	4.4*

*An asterisk (*) indicates that the main factor (T or pH) was statistically significant ($p < 0.05$) in this model. For AaPelA on sugar beet pectin, no model was obtained as the differences between the data points in the design were not statistically significant within the chosen ranges (**Supplementary Table S2** and **Supplementary Figure S4**).*

observed for the enzymes discussed here, whereas the sequence similarity between *Aa*PelA and *At*PelA is only 75% (**Table 1** and **Supplementary Figure S2**).

Generally, pH values of the optima were higher on apple pectin than on citrus or sugar beet pectins (**Table 4**). The pK_a of pectin decreases with increasing pH (Plaschina et al., 1978). Possibly, the higher DM of the apple pectin entails that a higher pH is required to reach the net charge preferred by the pectin lyases as compared to the lower-DM pectins.

Thermal stability of the pectin lyases was assessed by determining the melting temperatures (T_m) at pH 5. The highest T_m was found for *Aa*PelA (75°C), followed by *At*PelA (68°C), *Al*PelD (67°C), and *Al*PelB (65°C). The lowest T_m values were found for *Al*PelF (61°C) and *Al*PelC (57°C). Significant difference ($p < 0.05$) was observed between all values, indicating that the PelA pectin lyases were the most thermally stably, whereas the enzymes of low specific activity, i.e., *Al*PelF and *Al*PelC, also had the lowest thermal stability.

Kinetic Parameters

The kinetic parameters of the four most active pectin lyases were determined at the estimated pH-temperature optima (**Table 4**) on apple pectin, citrus pectin, and sugar beet pectin (**Table 5**). The enzymes differed in terms of K_m values on the different substrates except for *Aa*PelA where no significant difference in K_m was observed across the three pectins (11–14 mM GalA). For *At*PelA the K_m was significantly lower on apple pectin (8 mM GalA) than on citrus pectin (19 mM GalA), and highest on sugar beet pectin (25 mM GalA). For both *Al*PelB and *Al*PelD, no significant difference in K_m was observed between citrus and sugar beet pectins, indicating that acetylations had little effect on affinity for these enzymes (**Table 5**). The effect of methoxylation differed between the two: while K_m decreased with increasing DM for *Al*PelB, an increase was observed for *Al*PelD (**Table 5**). The lowest K_m values were observed for *At*PelA on apple pectin and for *Al*PelD on citrus and sugar beet pectins (**Table 5**). The latter were at the same level as that observed previously for *A. niger* PelD on high-methoxylated apple pectin (**Supplementary Table S1**). However, for *A. niger* PelA and PelB, previously reported K_m values are lower than the values observed here; this could be linked to the use of chemically esterified pectins in previous studies, where the pectins were esterified to a much higher DM than we used here, namely a very high DM of around 95% (**Supplementary Table S1**) (van Houdenhoven, 1975; Kusters-van Someren et al., 1991; Kester and Visser, 1994).

For all enzymes, k_{cat} was significantly lower on sugar beet pectin than on the other pectin substrates (**Table 5**). For *Aa*PelA and *Al*PelD, k_{cat} was highest on apple pectin, whereas there was no significant difference between k_{cat} values on apple and citrus pectins for *At*PelA and *Al*PelB (**Table 5**). For *Aa*PelA, *At*PelA, and *Al*PelB k_{cat}/K_m was significantly higher on apple pectin. For *Al*PelD the difference to citrus pectin was not statistically significant, but the value for apple pectin was higher. For all enzymes, k_{cat}/K_m was significantly lower on sugar beet pectin (**Table 5**). Again, this directly reflects the preference of pectin lyases for a highly methoxylated substrate and suggests a preference for substrates with low DAc. Regardless of the substrate, large differences in k_{cat}/K_m were observed between the enzymes, which ranked in the following order with respect to k_{cat}/K_m: *Al*PelB > *At*PelA > *Al*PelA > *Al*PelD (**Table 5**). The finding that *Al*PelB was the most efficient enzyme is in good agreement with literature data on *A. niger* pectin lyases, where the highest k_{cat}/K_m has indeed been observed for PelB (**Supplementary Table S1**). In that previous study (Kester and Visser, 1994), the k_{cat}/K_m was higher than observed in the current work, which could be due to the use of a highly methylesterified substrate, and that the reaction took place at pH 8.5. However, such a high pH is close to the pH where pectin spontaneously decomposes by β-elimination (Renard and Thibault, 1996). Analogously, significantly lower k_{cat}/K_m values were obtained for *Aa*PelA and *Al*PelD as compared to the *A. niger* PelA and PelD, which for *Aa*PelA was linked to an eight times higher K_m and for *Al*PelD to a five times higher k_{cat} (**Table 5** and **Supplementary Table S1**). For *At*PelA the k_{cat}/K_m of 17 mM^{-1} s^{-1} on apple pectin was comparable to that of the *A. niger* PelA (18–22 mM^{-1} s^{-1}); the use of a higher reaction temperature in the current work to increase k_{cat} balanced out the negative effect of the lower DM on K_m in this particular case (**Table 5** and **Supplementary Table S1**).

Pectin Degradation Product Profiling by LC-ESI-MS and SEC

To compare the pectin degradation profiles of all six pectin lyases, they were dosed in equimolar ratios on apple pectin, citrus pectin, and sugar beet pectin. For simplicity, a single set of reaction conditions was selected, reflecting the best

TABLE 5 | Kinetic parameters K_m (in mM GalA) and k_{cat} (in mol product released per mol enzyme per second) for *Aa*PelA, *At*PelA, *Al*PelB, and *Al*PelD on apple pectin (DM 69 ± 3%, DAc 2 ± 0.1%), citrus pectin (DM 53 ± 1%, DAc 1 ± 0.1%), and sugar beet pectin (DM 56 ± 5%, DAc 19 ± 1%).

Enzyme	Substrate	pH	*T*	R^2	K_m	k_{cat}	k_{cat}/K_m
			°C		mM	s^{-1}	mM^{-1} s^{-1}
*Aa*PelA	Apple pectin	6.4	55	1.00	$11 \pm 0.4^{x,a}$	$87 \pm 2.8^{x,c}$	$7.6 \pm 0.01^{x,c}$
	Citrus pectin	6.1	54	0.99	$14 \pm 0.8^{x,c}$	$54 \pm 1.6^{y,c}$	$3.8 \pm 0.1^{y,c}$
	Sugar beet pectin	5.5	55	0.92	$13 \pm 2.5^{x,b}$	$17 \pm 3.1^{z,c}$	$1.3 \pm 0.1^{z,c}$
*At*PelA	Apple pectin	4.9	57	0.99	$7.9 \pm 1.4^{z,b}$	$136 \pm 11^{x,b}$	$17 \pm 2^{x,b}$
	Citrus pectin	4.5	58	0.99	$19 \pm 1.8^{y,b}$	$131 \pm 8.5^{x,b}$	$7.0 \pm 0.2^{y,b}$
	Sugar beet pectin	4.4	59	0.96	$25 \pm 4.0^{x,a}$	$48 \pm 8.4^{y,b}$	$1.9 \pm 0.1^{z,b}$
*Al*PelB	Apple pectin	6.8	40	0.98	$13 \pm 2.7^{y,a}$	$274 \pm 35^{x,a}$	$22 \pm 2.7^{x,a}$
	Citrus pectin	6.4	49	1.00	$23 \pm 1.8^{x,a}$	$294 \pm 18^{x,a}$	$13 \pm 0.2^{y,a}$
	Sugar beet pectin	6.7	48	0.91	$21 \pm 1.6^{x,a}$	$96 \pm 3.6^{y,a}$	$4.5 \pm 0.3^{z,a}$
*Al*PelD	Apple pectin	4.9	59	0.98	$13 \pm 0.9^{x,a}$	$14 \pm 0.3^{x,d}$	$1.1 \pm 0.1^{x,d}$
	Citrus pectin	4.3	60	0.99	$8.9 \pm 0.7^{y,d}$	$8.7 \pm 0.2^{y,d}$	$1.0 \pm 0.1^{x,d}$
	Sugar beet pectin	4.4	60	0.90	$8 \pm 1.5^{y,c}$	$2.6 \pm 0.1^{z,d}$	$0.3 \pm 0.1^{y,d}$

*Reactions were performed at the estimated pH-temperature optima (**Table 4**). The R^2 values for the linear regression on the Hanes-Woolf plots are indicated, and reactions were run in triplicates. Superscript letters a-d indicate significant difference ($p < 0.05$) between the values of each parameter for the four enzymes when compared on the same substrate, whereas superscript letters x-z indicate significant difference ($p < 0.05$) between the three substrates for the same enzyme.*

compromise based on the obtained pH-temperature optima (**Table 4**): pH 5.5 and 40°C. The reactions were monitored over 24 h and analyzed by size exclusion chromatography (SEC) and liquid chromatography coupled to mass spectrometry with electrospray ionization (LS-ESI-MS). From the SEC chromatograms, the concentration of pectin molecules below ˜100 kDa could be monitored over time (**Figure 4**). The SEC analysis of pectin degradation products reflect the enzyme activity and substrate specificity data well: Apple pectin is degraded faster than citrus pectin, which is in turn degraded faster than sugar beet pectin (**Figure 4**). In a comparison of the six pectin lyases, the specific activities (**Table 3**) are directly reflected in the development of the SEC profiles over time: PelB > PelA > PelD > PelF > PelC (**Figure 4** and **Supplementary Figure S5**). Comparing the two PelA pectin lyases, they were again found to be equally fast on citrus pectin, while *At*PelA was faster on apple pectin, and *Aa*PelA was faster on sugar beet pectin (**Figure 4** and **Table 3**). As a consequence of their low activity, *Al*PelC and *Al*PelF were dosed 10 times higher than the other pectin lyases in order to better capture their product profiles in the LC-MS-based product profiling. For *Al*PelF the increased dosage had a large effect on pectin degradation (**Supplementary Figure S5**). For *Al*PelC the very low activity was only detectable after 24 h, with a noticeable difference between the two enzyme dosages (**Figure 4** and **Supplementary Figure S5**).

The product profiles obtained by LC-ESI-MS were expected to contain both partly and fully methoxylated structures. Partly methoxylated pectin-derived oligosaccharides can easily be identified by MS^2 in negative mode (Quéméner et al., 2003a; Ralet et al., 2005; Leijdekkers et al., 2011; Remoroza et al., 2012), whereas fully methoxylated pectin derived oligosaccharides can be identified in positive mode (Van Alebeek et al., 2000; Mutenda et al., 2002; Quéméner et al., 2003b). Manual inspection of the chromatograms obtained in negative mode revealed 28 unique compounds and 5 isomers all present in at least one enzyme-substrate-time combination (**Table 6**). The nomenclature of the identified compounds are $gA_xm_ya_z$, meaning x GalA residues, y methyl substitutions, and z acetyl substitutions, e.g., a trimer with three GalA, one methyl substitution, and one acetyl substitution is denoted $gA_3m_1a_1$. All compounds observed in negative mode appeared as [M-H], except gA_3m_3 and gA_4m_4 which appeared as formate $[M+FA]^-$ adducts and with a disproportionally low intensity. For all compounds observed in single charge state, the molecular formula could be confirmed by MS^2 fragmentation with almost exclusively C_i (or Z_j due to isomeric masses) and $^{0,2}A_i$ fragments observed, according to the nomenclature of Domon and Costello (Domon and Costello, 1988), including the occasional loss of 32 Da that can be attributed to methoxylation at the galacturonic acid residue in the reducing end (**Supplementary Figure S6**; Quéméner et al., 2003b). For larger structures, e.g., molecular masses larger than ˜1100 Da, a tendency toward double charging was observed (**Table 6**), making full identification from the fragmentation pattern difficult. Hence, for these compounds it was necessary to rely on the *m*/*z* values only. All compounds carried a 4,5 unsaturation in the non-reducing end due to the lyase reaction. For some masses, two distinct peaks were observed with two different retention times, indicating two different molecular structures (denoted #1 or #2; **Table 6**). The different retention times of the compounds are assumed to derive from different positions of the methyl and/or acetyl substitutions. The porous graphite column in use has previously shown the capability to separate different isomers of compounds with identical molecular composition, but with different linkages and/or planarity (Jamek et al., 2018; Mosbech et al., 2018; Zeuner et al., 2018).

Investigation of the chromatograms obtained in positive mode revealed another 14 unique compounds and 1 isomer all present in at least one enzyme-substrate-time combination (**Table 6**), with the majority being fully methoxylated

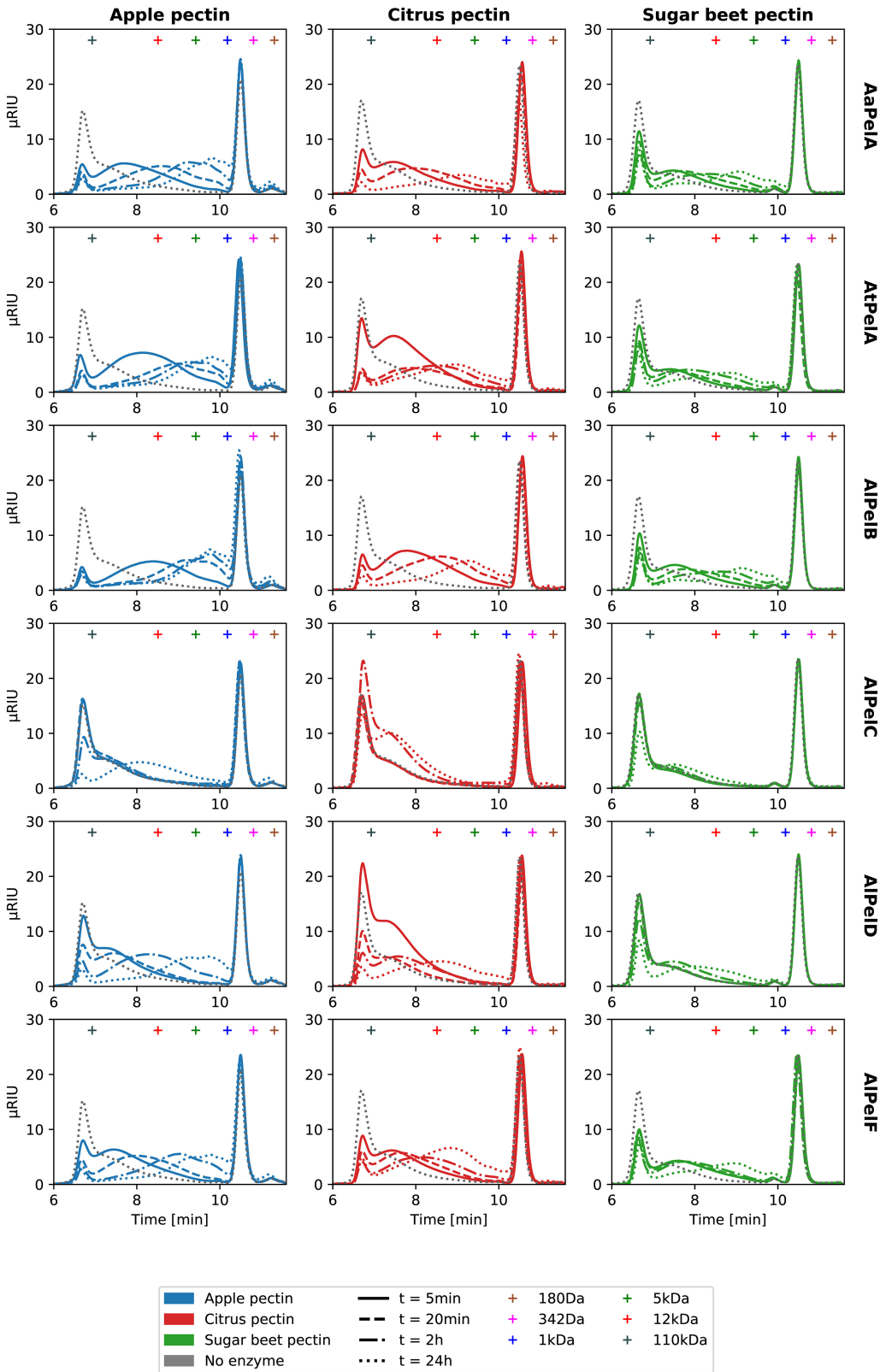


FIGURE 4 | Size exclusion chromatography (SEC) chromatograms for degradation of 10 g/L apple pectin (blue), citrus pectin (red), and sugar beet pectin (green) by 60 nM *Aa*PelA, *At*PelA, *Al*PelB, or *Al*PelD (corresponding to 0.023% E/S for *At*PelA) and 600 nM *Al*PelC or *Al*PelF at pH 5.5 and 40°C. For chromatograms of *Al*PelC and *Al*PelF at 60 nM, refer to **Supplementary Figure S5**. Retention times of pullulan standard (180 Da to 110 kDa) are indicated by crosses.

TABLE 6 | Compound ID, molecular weight (MW), number of GalA residues, number of methyl- and acetyl substitutions, charge state, and observed *m/z* in positive mode of all identified pectic oligosaccharides.

Compound ID	MW	GalA	Methyl	Acetyl	Charge *z*	Observed *m/z*
gA_2m_1	366	2	1		+1	384
$gA_2m_1a_1$	408	2	1	1	+1	426
$gA_2m_2a_1$*	422	2	2	1	+1	440
gA_3m_1 #1	542	3	1		+1	560
gA_3m_1 #2	542	3	1		+1	560
gA_3m_2 #1	556	3	2		+1	574
gA_3m_2 #2	556	3	2		+1	574
$gA_3m_2a_1$ #1	598	3	2	1	+1	616
$gA_3m_2a_1$ #2	598	3	2	1	+1	616
$gA_3m_2a_2$	640	3	2	2	+1	658
gA_3m_3	570	3	3		+1	588
$gA_3m_3a_1$*	612	3	3	1	+1	630
$gA_3m_3a_2$*	654	3	3	2	+1	672
gA_4m_2 #1	732	4	2		+1	750
gA_4m_2 #2	732	4	2		+1	750
$gA_4m_2a_1$	774	4	2	1	+1	792
$gA_4m_2a_2$	816	4	2	2	+1	834
gA_4m_3	746	4	3		+1	746
$gA_4m_3a_1$ #1	788	4	3	1	+1	806
$gA_4m_3a_1$ #2	788	4	3	1	+1	806
$gA_4m_3a_2$	830	4	3	2	+1	848
gA_4m_4	760	4	4		+1	778
$gA_4m_4a_1$ #1*	802	4	4	1	+1	820
$gA_4m_4a_1$ #2*	802	4	4	1	+1	820
$gA_4m_4a_2$*	844	4	4	2	+1	862
gA_5m_3	922	5	3		+2	479
$gA_5m_3a_1$	964	5	3	1	+1/+2	982/500
$gA_5m_3a_2$	1006	5	3	2	+2	521
gA_5m_4	936	5	4		+1/+2	954/486
$gA_5m_4a_1$	978	5	4	1	+2	507
$gA_5m_4a_2$	1020	5	4	2	+2	528
gA_5m_5*	950	5	5		+1/+2	968/493
$gA_5m_5a_1$*	992	5	5	1	+2	514
gA_6m_3	1098	6	3		+2	567
gA_6m_4	1112	6	4		+2	574
$gA_6m_4a_1$	1154	6	4	1	+2	595
$gA_6m_4a_2$*	1196	6	4	2	+2	616
gA_6m_5	1126	6	5		+2	581
$gA_6m_5a_1$*	1168	6	5	1	+2	602
$gA_6m_5a_2$*	1210	6	5	2	+2	623
gA_6m_6*	1140	6	6		+2	588
gA_7m_4	1288	7	4		+2	662
gA_7m_5	1302	7	5		+2	669
gA_7m_6*	1335	7	6		+2	676
$gA_7m_6a_1$	1358	7	6	1	+2	697
gA_7m_7*	1330	7	7		+2	683
gA_8m_6	1492	8	6		+2	764
gA_8m_7*	1506	9	6		+2	771

An asterisk (*) indicates compounds only observed in positive mode.

compounds. All identified compounds were observed as single charge $[M+NH_4]^+$ or double charge $[^1/_2M + NH_4]^{2+}$ ammonium adducts due to the constant presence of ammonium formate in the eluent. The structural verification by MS^2 fragmentation was somewhat complex due to an unexpected fragmentation pattern. From literature, fully methoxylated pectin derived oligosaccharides are expected to form C_i/Z_j ions upon fragmentation in positive mode (Mutenda et al., 2002), although a mixed pattern including all possible ion types can be observed for pectic oligosaccharides in general (Körner et al., 1999). In the present case, assuming B_i and C_i fragments and the NH_4^+ ion retained on all fragments, all observed fragments had and *m/z* value 7 Da lower than expected. For example, this pattern was observed in positive mode for gA_4m_3, for which the structure could be verified in negative mode (**Supplementary Figures S6, S7A**). A number of putative, fully methoxylated structures were observed in positive mode, including gA_3m_3, gA_4m_4, and gA_5m_5. They all shared the "B_i/C_i -7 Da" pattern (**Supplementary Figures S7B–D**). The compound gA_5m_5 was observed both as a single charge and a double charge ion, with the fragmentation pattern of the latter giving more information (**Supplementary Figure S7E**). Despite not being the most dominant fragments, true C_i ions were observed. Since this pattern is observed for compounds where additional structural verification in negative mode is available, it is assumed that the proposed structures with full methoxylation are correct.

Contrary to fragmentation in negative mode, where only C_i type (and isomeric Z_j type) ions are observed, fragmentation in positive mode revealed additional information on acetylation patterns for different isomers sharing the same mass such as $gA_4m_4a_1$ #1 and #2. The extracted ion chromatogram with the corresponding mass 820 *m/z* clearly indicates at least two different compounds with baseline separation (**Figure 5A**). The fragmentation pattern of $gA_4m_4a_1$ #1 (**Figure 5B**) shows B_3 and C_3 fragments that have lost the acetyl group, indicating that the acetyl substitution is on the reducing end galacturonic acid unit. In contrast, the acetyl group is retained on B_3 and C_3 fragments and lost on B_2 and C_2 fragments in the fragmentation pattern of $gA_4m_4a_1$ #2 (**Figure 5C**), indicating that the acetyl group is on the third galacturonic acid moiety when counting from the non-reducing end. No cross ring fragmentation was observed in either cases, leaving the position (O2 or O3) of the acetyl substitution unsolved.

The observed peak intensities showed that relatively few compounds, primarily gA_3m_2 and gA_3m_3, and to some extend gA_2m_1, gA_4m_3, gA_4m_4, and gA_6m_4, were the dominant products for the *Aspergillus* sp. pectin lyases acting on apple and citrus pectins, with a more disperse profile when acting on sugar beet pectin (**Supplementary Figure S8**). An increased level of most of the identified compounds was observed over time, although with different rates dependent on type of compound, substrate and enzyme. The levels of larger structures such as gA_6m_5, gA_6m_6, gA_7m_7, and gA_8m_7 exhibited transient maxima during the 24 h reaction, indicating that these pectic oligomers were substrates for the pectin lyases (**Supplementary Figure S8**). The large variations in production rates among the enzymes are coherent with activity data (**Table 3**, **Figure 4**, and **Supplementary Figure S8**). Because of the large variation in activity among the enzymes, it conceals the details of the profiles for enzymes with lower

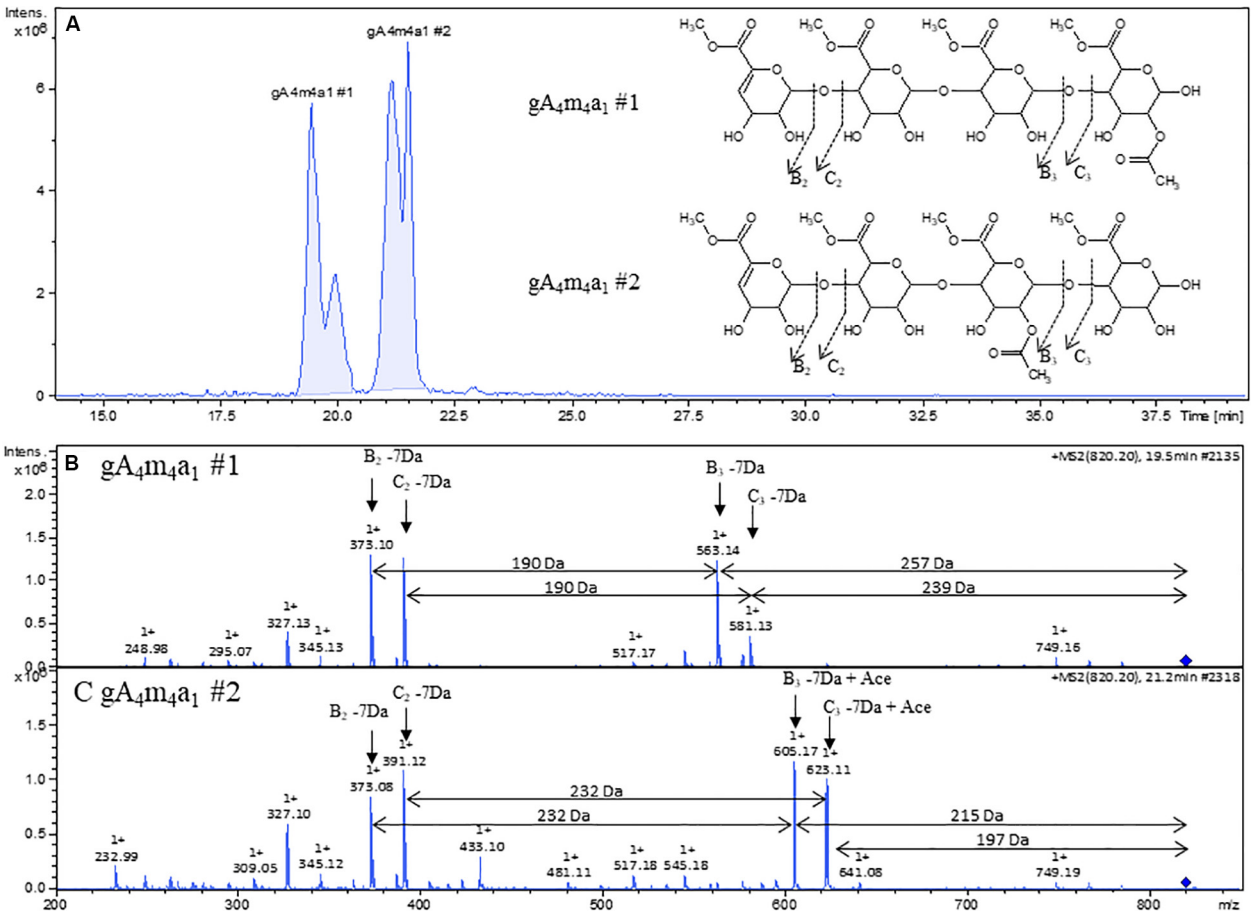


FIGURE 5 | (A) Extracted ion chromatogram of 820 *m/z* with proposed structures inserted. **(B)** Positive ion mode ESI-MS^2 of $gA_4m_4a_1$ #1. **(B)** Positive ion mode of ESI-MS^2 of $gA_4m_4a_1$.

activity, such as *Al*PelC and *Al*PelD. Hence, the data was normalized in order to facilitate an in-depth investigation and comparison of the product profiles.

For all enzymes acting on apple pectin, gA_3m_3 was generally the most dominant product throughout the time course (**Figure 6A**). In the early stages (5, 20 min) some of the longer compounds were more highly represented (degree of polymerization (DP) of 5–7, especially gA_6m_5), but the relative abundance of these went down in favor of shorter compound like gA_3m_2 during the extended reaction (120 min, 24 h). The relative levels of gA_4m_3 and gA_4m_4 were steady throughout the time course, and there was a tendency to build-up of gA_6m_4 (**Figure 6A**). The shift from longer to shorter compounds was most predominantly observed for *Al*PelB, which has the highest specific activity (**Table 3**) and the highest k_{cat}/K_m on apple pectin (**Table 5**). The product profilse after 24 h were quite similar for *Aa*PelA, *At*PelA, *Al*PelB, and *Al*PelF with gA_3m_2 and gA_3m_3 as the main products. An even narrower product profile was observed for *Al*PelD, clearly favoring gA_3m_3 and indicating a preference for fully methoxylated pectin. In contrast, a more dispersed profile was observed for *Al*PelC (**Figure 6A**). This more dispersed profile could be caused by the low activity of *Al*PelC (**Table 3** and **Figure 4**), since the profile after 24 h looked more like the profiles of *Al*PelD and *Al*PelF after 20 min. However, it may also reflect a more selective substrate specificity, where the enzyme failed to degrade the longer compounds. Indeed, a vast increase in gA_6m_5 levels was observed for *Al*PelC during 24 h, whereas slight decreases in the observed peak intensities were observed for the other pectin lyases over the course of the reaction (**Supplementary Figure S8**). The clustering analysis of the product profiles was most clear after 24 h, with the profiles of *Al*PelB and *Al*PelA clustering closest together, and the profiles of *Al*PelD and *Al*PelC clustering second-most and most apart from the other pectin lyases, respectively.

In general, the product profile for all pectin lyases acting on citrus pectin looked similar to the profile on apple pectin (**Figure 6B**). This was expected due to the similarities of the two substrates in terms of fairly high methylation, low acetylation, and a low content of RGI. Similar product profile patterns were observed with gA_3m_2 and gA_3m_3 as the main end products for *Aa*PelA, *At*PelA, *Al*PelB and *Al*PelF, almost solely gA_3m_3 for *Al*PelD, and a more dispersed profile for *Al*PelC (**Figure 6B**). A notable difference compared to the profiles of enzymes acting on apple pectin was the lower relative levels of DP6-7

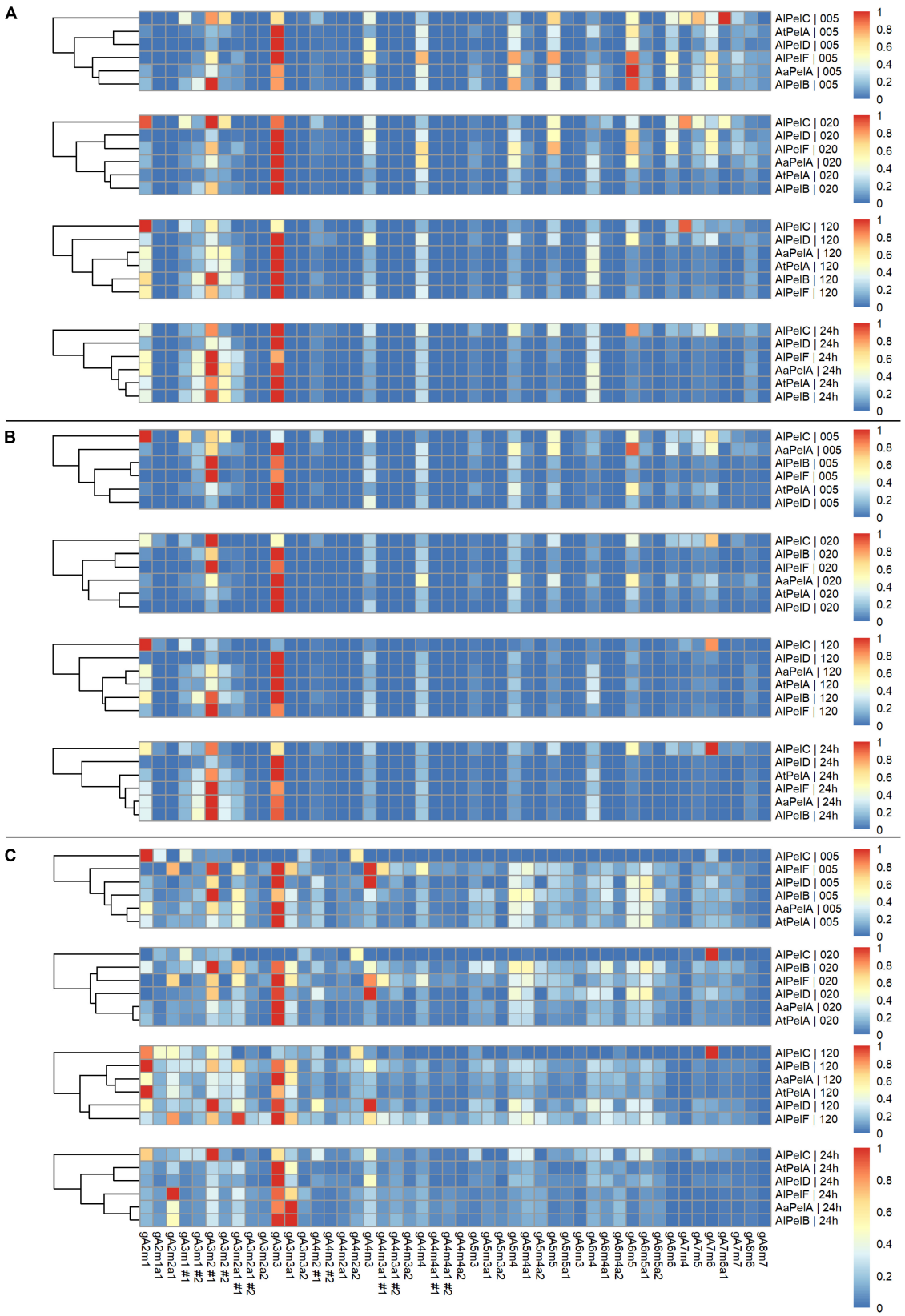


FIGURE 6 | Heat maps of normalized relative intensities for pectin lyases acting on apple pectin **(A)**, citrus pectin **(B)**, and sugar beet pectin **(C)**. Panels top to bottom: 5, 20, 120 min, and 24 h of reaction. For each substrate the pectin lyases are clustered according to product profiles within each time point.

observed for citrus pectin. For apple pectin, especially gA_6m_5 was major contributor to the profiles particularly in the early stages of the reaction; a similarly, important contribution from this compound to the product profile was not observed on citrus pectin (**Figures 6A,B**). This might be due to the different distributions of methyl substitutions in the two substrates. The clustering analysis was almost identical to that observed for apple pectin, with the product profiles of *Al*PelB and *Aa*PelA clustering most, and the product profiles of *Al*PelD and *Al*PelC being the most different from the other enzymes after 24 h (**Figure 6B**).

From activity measurements (**Tables 3, 5** and **Figure** 4) as well as the peak intensities of the individual compounds (**Supplementary Figure S8**), it is evident that sugar beet pectin is an inferior substrate compared to apple and citrus pectins, but decent degradation was nevertheless observed. The product profile of pectin lyases acting on sugar beet pectin was more differentiated than the profiles of apple and citrus pectins, due to the relatively higher presence of acetyl substituted compounds (**Figure 6C**). For all enzymes, except *Al*PelC, the major compound throughout the time course was gA_3m_3, with relatively high levels of DP5-6. The low levels of gA_6m_6 and longer compounds could be due to the nature of the sugar beet pectin substrate, where the structures necessary for releasing these products are present in limited amounts. The tendency of the longer compounds (DP5-6) to be outnumbered by the shorter compounds over time was also evident for the enzymes acting on sugar beet pectin. Remarkably, for *Aa*PelA and *Al*PelB one of the most dominating compounds was the fully methoxylated $gA_3m_3a_1$, a compound that features a galacturonic acid residue with both a methyl and an acetyl substitution (**Figure 5**). These double substitutions has previously been reported as rare in sugar beet pectin (Ralet et al., 2005). The relative contribution of $gA_3m_3a_1$ to the product profile compared to gA_3m_3 was most pronounced after 24 h, indicating that while especially *Aa*PelA, *Al*PelB, and *Al*PelF accommodate the acetylated, fully methoxylated substrate, they release the non-acetylated variant first (**Figure 6C** and Figure **Supplementary S8**). Again, a more dispersed profile was observed for *Al*PelC, with gA_3m_2 being the main product. The cluster analyses after 24 h were similar to the ones of apple and citrus pectins with the product profiles of *Al*PelB and *Aa*PelA clustering most, and *Al*PelC being most different (**Figure 6C**).

Examining the product profiles in a broader perspective with the intensities summed up according to DP with and without acetyl substitutions it was evident that DP3 is the most preferred product throughout the reactions (**Figure 7**). *Al*PelC showed a more differentiated profile compared to the other enzymes during the early phases of the reaction, but after 2 h of reaction DP3 was also dominant for this enzyme (**Figure 7**). In general, only negligible amounts of acetyl substituted compounds were observed in reactions on apple and citrus pectins, which was highly expected due to the low DAc of these substrates (**Table 2**), whereas slightly more differentiated profiles were observed for sugar beet reactions due to substrate acetylation. Hence, the division is more applicable for the profiles of enzymes acting on sugar beet pectin and the segregation revealed a division of the enzymes in two groups. *Aa*PelA, *Al*PelB, and *Al*PelF showed equal preference/tolerance for DP3 with and without acetyl substitutions as a product, and even a higher preference for releasing acetyl substituted DP4-6 compared to non-substituted DP4-6 (**Figure 7**). In contrast, *At*PelA, *Al*PelC, and *Al*PelD showed a notable preference for releasing non-acetylated DP3 compared to acetylated DP3. However, while *Al*PelC specifically preferred release of non-acetylated pectic oligosaccharides regardless of DP, the result was more blurred for *At*PelA and *Al*PelD, where the tolerance to acetylation changed across the DP range (**Figure 7**).

A broad product profile was observed after 5 and 20 min for all enzymes, with a shift toward the narrow profile of mainly DP3 over the 24 h of reaction. Even though a limited decrease in the levels of the longer DPs was observed (**Supplementary Figure S8**), this shift in the profiles can be explained by the production rates of the different compounds. The apparent production rate of DP3 is the sum of products originating from degradation of longer oligomers/polymers not accounted for in the profile analysis. It seems likely that DP3 does not serve as a substrate for any of the pectin lyases, which is in agreement with previous studies (Van Alebeek et al., 2002). In contrast, the apparent formation rate of longer DP, e.g., DP6, is the net sum of the positive production rate and the negative degradation rate. If the latter is lower than the former, the net result is a positive apparent production rate of DP6, albeit lower than the production rate of DP3. In all cases, a shift in the profiles from higher DP (DP5-7) to DP3 was observed, with the exception of *Al*PelC that had a more differentiated profile throughout the reactions, regardless of substrate (**Figures 6**, 7). For the cluster analyses on citrus and sugar beet pectins after 24 h reaction, *Aa*PelA and *Al*PelB clustered closest together, whereas *At*PelA and *Al*PelB clustered closest together in reactions on apple pectin (**Figures 6**, 7). Furthermore, *Al*PelD and *Al*PelC had the second and most different product profile, when compared to the rest of the enzymes after 24 h. Phylogenetically, *Al*PelF clustes far from PelA and PelB (**Figure 2**) and has a different profile than these enzymes at short reaction times (**Figures 6**, 7). However, at extended reaction times (and high enzyme dosage), the *Al*PelF profile became increasingly similar to that of the PelAs and *Al*PelB. Thus, the differences observed for *Al*PelF could merely be a result of low activity, which could in turn be explained by lack of an otherwise conserved substrate-interacting Trp residue (**Figure 3**).

A Basis for Understanding Pectin Lyase Multigenecity in *Aspergilli*

Major differences in specific activity were observed among the different pectin lyases (**Table 3**), and through systematic experiments, we also observed differences in pH-temperature optima for PelA, PelB, and PelD (**Table 4**). Clearly, from a competitive growth perspective, the ability of the fungus to produce a battery of similar enzymes having optimal activity at different pH and temperature represents a major advantage, and this fitness theory provides an explanation for the multigenecity phenomenon also for the pectin lyases. Interestingly, the ranking of k_{cat}/K_m across the three pectin substrates was the same for

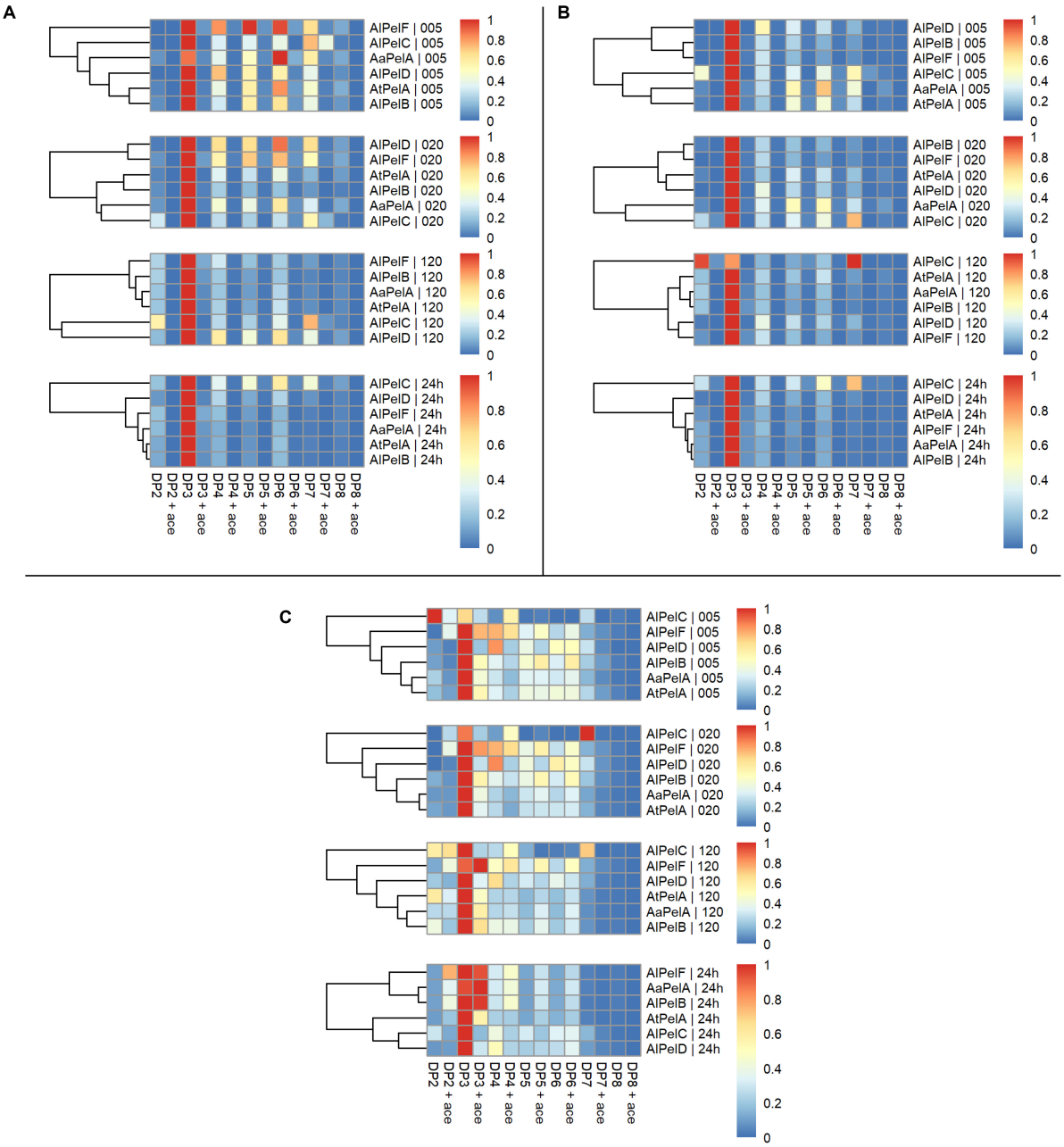


FIGURE 7 | Heat maps of normalized summed relative intensities for individual degrees of polymerization divided non-acetyl substituted (DP columns) and acetyl substituted (DP ace columns) of pectin lyases acting on apple pectin **(A)**, citrus pectin **(B)**, and sugar beet pectin **(C)**. Panels top to bottom: 5, 20, 120 min, and 24 h. For each substrate the pectin lyases are clustered according to product profiles within each time point.

all of the four most active enzymes, but some discrepancies were observed for the K_m values, suggesting different propensity toward methyl and acetyl substitutions among the pectin lyases (**Table 4**). This interpretation was further substantiated by the detailed LC-MS product profiling, where differences in substrate degradation proficiency as well as in the end product profiles were observed between the enzymes (**Figures 6**, 7). In particular, the discriminative substrate degradation pattern analysis revealed a significant differential activity sensibility to acetyl substitutions.

A. niger regulates the expression of carbohydrate-degrading enzymes at the transcriptional level. At least three different transcriptional activators are linked to pectin degradation, namely GaaR, RhaR, and AraR, which control the production of enzymes responsible for degradation of homogalacturonan, RG-I, and RG-I arabinan and arabinogalactan side chains, respectively (Kowalczyk et al., 2017). Moreover, the GalA-dependent transcriptional repressor GaaX and the general carbon catabolite repressor CreA also take part in controlling expression

of pectinolytic enzymes in *A. niger* including the pectin lyases (Kowalczyk et al., 2017; Niu et al., 2017). When grown on sugar beet pectin, PelA was found to be regulated by GaaR, whereas PelD was regulated by both GaaR, RhaR, and possibly also AraR (Alazi et al., 2016; Kowalczyk et al., 2017). These transcription regulators are proteins which are activated by the presence and concentration of certain monosaccharides or a catabolic product thereof (Kowalczyk et al., 2017). Indeed, GaaR is induced by 2-keto-3-deoxy-L-galactonate, an intermediate of the GalA catabolic pathway (Alazi et al., 2016, 2017). Signal molecules such as GalA may be released from the complex carbohydrates by enzymes which are constitutively expressed at low levels. Indeed, *A. niger* PelB, PelC, and PelF and several other pectinolytic enzymes were found not to be regulated by GaaR. Data indicated that they may instead be regulated by the general carbon catabolite repressor CreA and could therefore play a role in release of GalA to induce expression of PelA and PelD (de Vries et al., 2002; Niu et al., 2015; Kowalczyk et al., 2017). For PelE from *A. niger* CBS 120.49 transcription levels were very low, and it is therefore hard to conclude whether PelE is also regulated by the GalA-induced GaaX-GaaR system (Alazi et al., 2018). Expression profiling and transcriptomics on *A. niger* grown on sugar beet pectin or GalA revealed that PelA and PelD (both regulated by GaaR) were predominantly expressed/transcribed after short incubation times. In contrast, levels of PelB and PelF (mRNA) increased throughout 24 h of incubation (de Vries et al., 2002; Alazi et al., 2016). The mRNA levels of PelC were very low, again questioning its role in pectin degradation (Alazi et al., 2016, 2018), in agreement with the very low activity observed in the current work. Comparing mRNA levels a general tendency was observed: PelA > PelF > PelD > PelB, and PelC and PelE were hardly transcribed (Martens-Uzunova and Schaap, 2009; Alazi et al., 2016, 2018; Kowalczyk et al., 2017).

Based on this, PelA appears to remain the major pectin lyase responsible for pectin degradation by *Aspergilli* in nature with high activity and high expression levels. However, the fact that *Al*PelB was observed to be more efficient enzyme in the current work, indicates a possible important role of this enzyme in releasing GalA to induce the transcription of PelA despite low transcription levels, especially at elevated pH. Indeed, varying expression levels may be the reason for the presence of the much less efficient *Al*PelF to release GalA for GaaR activation. From an industrial perspective, where enzymes are produced recombinantly, *Al*PelB thus appears to be promising candidate for an efficient enzyme with a broad pH range. In contrast, the reason for maintaining *pelC* and *pelE* in the chromosomes is more elusive, given the inferior activity of AlPelC and low transcription levels of both enzymes. The fact that *pelE* could not be identified in *A. niger* CBS 513.88 (Pel et al., 2007; Martens-Uzunova and Schaap, 2009) may suggest evolutionary pressure to lose such genes from the multigene family.

CONCLUSION

In order to study the multigenecity of pectin lyases in *Aspergilli*, representatives of PelA, PelB, PelC, PelD, and PelF were recombinantly expressed and characterized. A comparison of the pectin lyases in terms of activity at acidic pH, which is relevant for e.g., juice processing applications, the following ranking materialized: PelB > PelA > PelD > PelF > PelC. Interestingly, PelB was consistently the most efficient pectin lyase even at acidic conditions, despite having a higher pH optimum than PelA. In general, the highest extent of pectin degradation was observed for PelA and PelB. A systematic assessment of pH-temperature optima revealed significant differences in reaction optima among the different types of the pectin lyases, and homology modeling of the different versions of the pectin lyases supported that the slight sequence disparity amongst the copies, fosters structural differences in the active site region of the enzymes. Together with differences in thermal stability amongst the pectin lyases, these differences and the activity optimum variation easily supports the prevailing theory that this multigenecity is related to competitive fitness evolution providing the fungus with optimal tools for survival in different environments.

By providing a compendium of the degradation product profiles for the available *Aspergillus* pectin lyases on three different types of pectin (from apple, citrus, and sugar beet), which varied in degrees of methoxylation and acetylation, the present work added a new dimension to pectin lyase diversity characterization, and moreover provided a new angle to understand why fungi possess multiple copies of similar CAZymes. The discriminative LC-MS based substrate degradation pattern analysis provides support to the idea that the multigenecity is a prerequisite for the sophisticated biological "plasticity" that grants the fungus the ability to degrade an array of almost similar structural moieties, in this case pectin having different methoxylation and acetylation patterns. This interpretation agrees with the idea that the evolution of the CAZymes-secretome is an integral part of fungal speciation (Benoit et al., 2015; Barrett et al., 2020b). The product profiling method presented here may furthermore be useful as a novel tool in enzyme characterization as it enables discrimination between different enzymes with identical catalytic activity, and thus provides a new approach to continued exploration of the action of CAZymes on complex biomass substrates.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

JH, AM, MS, and KK conceptualized and supervised the study. MS and KK expressed and purified the enzymes and determined their melting temperatures. TT performed experiments to determine substrate monosaccharide composition, enzyme pH-temperature optima, and kinetic constants. BZ determined initial rates on different pectin substrates and performed SDS-PAGE analysis. BZ, TT, and MS performed all *in silico* analyses.

JH performed experiments and LC-MS and SEC analyses for product profiling and degradative pattern display. JH, AM, BZ, and TT analyzed all data. BZ, TT, and JH drafted the manuscript. AM, JH, MS, KK, and BZ revised the final manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was part of the BioValue SPIR Industrial Inovation Platform funded Innovation Fund Denmark, case no. 0603-00522B.

ACKNOWLEDGMENTS

Mikael Lenz Strube, Department of Biotechnology and Biomedicine, Technical University of Denmark, is thanked for assistance in data handling and data analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00873/full#supplementary-material

REFERENCES

Agger, J., Viksø-Nielsen, A., and Meyer, A. S. (2010). Enzymatic xylose release from pretreated corn bran arabinoxylan: differential effects of deacetylation and deferuloylation on insoluble and soluble substrate fractions. *J. Agric. Food Chem.* 58, 6141–6148. doi: 10.1021/jf100633f

Alazi, E., Khosravi, C., Homan, T. G., du Pré, S., Arentshorst, M., Di Falco, M., et al. (2017). The pathway intermediate 2-keto-3-deoxy-L-galactonate mediates the induction of genes involved in D-galacturonic acid utilization in *Aspergillus niger*. *FEBS Lett.* 591, 1408–1418. doi: 10.1002/1873-3468.12654

Alazi, E., Knetsch, T., Di Falco, M., Reid, I. D., Arentshorst, M., Visser, J., et al. (2018). Inducer-independent production of pectinases in *Aspergillus niger* by overexpression of the D-galacturonic acid-responsive transcription factor gaaR. *Appl. Microbiol. Biotechnol.* 102, 2723–2736. doi: 10.1007/s00253-018-8753-7

Alazi, E., Niu, J., Kowalczyk, J. E., Peng, M., Aguilar Pontes, M. V., van Kan, J. A. L., et al. (2016). The transcriptional activator GaaR of *Aspergillus niger* is required for release and utilization of d-galacturonic acid from pectin. *FEBS Lett.* 590, 1804–1815. doi: 10.1002/1873-3468.12211

Ali, S., Søndergaard, C. R., Teixeira, S., and Pickersgill, R. W. (2015). Structural insights into the loss of catalytic competence in pectate lyase activity at low pH. *FEBS Lett.* 589, 3242–3246. doi: 10.1016/j.febslet.2015.09.014

Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37, 420–423. doi: 10.1038/s41587-019-0036-z

Andersen, M. R., Giese, M., de Vries, R. P., and Nielsen, J. (2012). Mapping the polysaccharide degradation potential of *Aspergillus niger*. *BMC Genomics* 13:313. doi: 10.1186/1471-2164-13-313

Barrett, K., Hunt, C. J., Lange, L., and Meyer, A. S. (2020a). Conserved unique peptide patterns (CUPP) online platform: peptide-based functional annotation of carbohydrate active enzymes. *Nucleic Acids Res.* 48, W110–W115. doi: 10.1093/nar/gkaa375

Barrett, K., Jensen, K., Meyer, A. S., Frisvad, J. C., and Lange, L. (2020b). Fungal secretome profile categorization of CAZymes by function and family corresponds to fungal phylogeny and taxonomy: example *Aspergillus* and *Penicillium*. *Sci. Rep.* 10:5158. doi: 10.1038/s41598-020-61907-1

Barrett, K., and Lange, L. (2019). Peptide-based functional annotation of carbohydrate-active enzymes by conserved unique peptide patterns (CUPP). *Biotechnol. Biofuels* 12, 1–21. doi: 10.1186/s13068-019-1436-5

Benoit, I., Coutinho, P. M., Schols, H. A., Gerlach, J. P., Henrissat, B., and de Vries, R. P. (2012). Degradation of different pectins by fungi: correlations and contrasts between the pectinolytic enzyme sets identified in genomes and the growth on pectins of different origin. *BMC Genomics* 13:321. doi: 10.1186/1471-2164-13-321

Benoit, I., Culleton, H., Zhou, M., DiFalco, M., Aguilar-Osorio, G., Battaglia, E., et al. (2015). Closely related fungi employ diverse enzymatic strategies to degrade plant biomass. *Biotechnol. Biofuels* 8:107. doi: 10.1186/s13068-015-0285-0

Bonnin, E., Garnier, C., and Ralet, M. C. (2014). Pectin-modifying enzymes and pectin-derived materials: applications and impacts. *Appl. Microbiol. Biotechnol.* 98, 519–532. doi: 10.1007/s00253-013-5388-6

Buchan, D. W. A., and Jones, D. T. (2019). The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Res.* 47, W402–W407. doi: 10.1093/nar/gkz297

Buchholt, H. C., Christensen, T. M. I. E., Fallesen, B., Ralet, M. C., and Thibault, J. F. (2004). Preparation and properties of enzymatically and chemically modified sugar beet pectins. *Carbohydr. Polym.* 58, 149–161. doi: 10.1016/j.carbpol.2004.06.043

Chung, W. S. F., Meijerink, M., Zeuner, B., Holck, J., Louis, P., Meyer, A. S., et al. (2017). Prebiotic potential of pectin and pectic oligosaccharides to promote anti-inflammatory commensal bacteria in the human colon. *FEMS Microbiol. Ecol.* 93:fix127. doi: 10.1093/femsec/fix127

de Vries, R. P., Jansen, J., Aguilar, G., Paøenicová, L., Joosten, V., Wülfert, F., et al. (2002). Expression profiling of pectinolytic genes from *Aspergillus niger*. *FEBS Lett.* 530, 41–47. doi: 10.1016/S0014-5793(02)03391-4

de Vries, R. P., Riley, R., Wiebenga, A., Aguilar-Osorio, G., Amillis, S., Uchima, C. A., et al. (2017). Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus *Aspergillus*. *Genome Biol.* 18:28. doi: 10.1186/s13059-017-1151-0

Domon, B., and Costello, C. E. (1988). A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconj. J.* 5, 397–409. doi: 10.1007/BF01049915

Gysler, C., Harmsen, J. A. M., Kester, H. C. M., Visser, J., and Heim, J. (1990). Isolation and structure of the pectin lyase D-encoding gene from *Aspergillus niger*. *Gene* 89, 101–108. doi: 10.1016/0378-1119(90)90211-9

Harmsen, J. A. M., Kusters-van Someren, M. A., and Visser, J. (1990). Cloning and expression of a second *Aspergillus niger* pectin lyase gene (pelA): indications of a pectin lyase gene family in A. niger. *Curr. Genet.* 18, 161–166. doi: 10.1007/BF00312604

He, Y., Pan, L., and Wang, B. (2018). Efficient over-expression and application of high-performance pectin lyase by screening *Aspergillus niger* pectin lyase gene family. *Biotechnol. Bioprocess Eng.* 23, 662–669. doi: 10.1007/s12257-018-0387-1

Herron, S. R., Benen, J. A. E., Scavetta, R. D., Visser, J., and Jurnak, F. (2002). Structure and function of pectic enzymes: virulence factors of plant pathogens. *Proc. Natl. Acad. Sci.U.S.A.* 97, 8762–8769. doi: 10.1073/pnas.97.16.8762

Hong, S. B., Lee, M., Kim, D. H., Varga, J., Frisvad, J. C., Perrone, G., et al. (2013). *Aspergillus luchuensis*, an industrially important black *Aspergillus* in East Asia. *PLoS One* 8:e0063769. doi: 10.1371/journal.pone.0063769

Jamek, S. B., Mikkelsen, J. D., Busk, P. K., Meyer, A. S., Holck, J., Zeuner, B., et al. (2018). Loop protein engineering for improved transglycosylation activity of a β-N-Acetylhexosaminidase. *ChemBioChem* 19, 1858–1865. doi: 10.1002/cbic.201800181

Kassara, S., Li, S., Smith, P., Blando, F., and Bindon, K. (2019). Pectolytic enzyme reduces the concentration of colloidal particles in wine due to changes in polysaccharide structure and aggregation properties. *Int. J. Biol. Macromol.* 140, 546–555. doi: 10.1016/j.ijbiomac.2019.08.043

Kester, H. C. M., and Visser, J. (1994). Purification and characterization of pectin lyase B, a novel pectinolytic enzyme from *Aspergillus niger*. *FEMS Microbiol. Lett.* 120, 63–68. doi: 10.1016/0378-1097(94)00176-6

Kohli, P., and Gupta, R. (2019). Application of calcium alginate immobilized and crude pectin lyase from *Bacillus cereus* in degumming of plant fibres. *Biocatal. Biotransformation* 37, 341–348. doi: 10.1080/10242422.2018.1564745

Körner, R., Limberg, G., Christensen, T. M. I. E., Mikkelsen, J. D., and Roepstorff, P. (1999). Sequencing of partially methyl-esterified oligogalacturonates by tandem mass spectrometry and its use to determine pectinase specificities. *Anal. Chem.* 71, 1421–1427. doi: 10.1021/ac981240o

Kowalczyk, J. E., Lubbers, R. J. M., Peng, M., Battaglia, E., Visser, J., and De Vries, R. P. (2017). Combinatorial control of gene expression in *Aspergillus niger* grown on sugar beet pectin. *Sci. Rep.* 7, 1–12. doi: 10.1038/s41598-017-12362-y

Kusters-van Someren, M., Flipphi, M., de Graaff, L., van den Broeck, H., Kester, H., Hinnen, A., et al. (1992). Characterization of the *Aspergillus niger* pelB gene: structure and regulation of expression. *MGG Mol. Gen. Genet.* 234, 113–120. doi: 10.1007/BF00272352

Kusters-van Someren, M. A., Harmsen, J. A. M., Kester, H. C. M., and Visser, J. (1991). Structure of the *Aspergillus niger* pelA gene and its expression in *Aspergillus niger* and *Aspergillus nidulans*. *Curr. Genet.* 20, 293–299. doi: 10.1007/BF00318518

Leijdekkers, A. G. M., Sanders, M. G., Schols, H. A., and Gruppen, H. (2011). Characterizing plant cell wall derived oligosaccharides using hydrophilic interaction chromatography with mass spectrometry detection. *J. Chromatogr. A* 1218, 9227–9235. doi: 10.1016/j.chroma.2011.10.068

Lo, M. C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M., et al. (2004). Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal. Biochem.* 332, 153–159. doi: 10.1016/j.ab.2004.04.031

Lombard, V., Bernard, T., Rancurel, C., Brumer, H., Coutinho, P. M., and Henrissat, B. (2010). A hierarchical classification of polysaccharide lyases for glycogenomics. *Biochem. J.* 432, 437–444. doi: 10.1042/bj20101185

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 47, W636–W641. doi: 10.1093/nar/gkz268

Mantovani, C. F., Geimba, M. P., and Brandelli, A. (2005). Enzymatic clarification of fruit juices by fungal pectin lyase. *Food Biotechnol.* 19, 173–181. doi: 10.1080/08905430500316284

Martens-Uzunova, E. S., and Schaap, P. J. (2009). Assessment of the pectin degrading enzyme network of *Aspergillus niger* by functional genomics. *Fungal Genet. Biol.* 46(Suppl. 1), S170–S179. doi: 10.1016/j.fgb.2008.07.021

Mayans, O., Scott, M., Connerton, I., Gravesen, T., Benen, J., Visser, J., et al. (1997). Two crystal structures of pectin lyase A from *Aspergillus* reveal a pH driven conformational change and striking divergence in the substrate-binding clefts of pectin and pectate lyases. *Structure* 5, 677–689. doi: 10.1016/S0969-2126(97)00222-0

Mosbech, C., Holck, J., Meyer, A. S., and Agger, J. W. (2018). The natural catalytic function of CuGE glucuronoyl esterase in hydrolysis of genuine lignin-carbohydrate complexes from birch. *Biotechnol. Biofuels* 11:71. doi: 10.1186/s13068-018-1075-2

Müller-Maatsch, J., Bencivenni, M., Caligiani, A., Tedeschi, T., Bruggeman, G., Bosch, M., et al. (2016). Pectin content and composition from different food waste streams in memory of Anna Surribas, scientist and friend. *Food Chem.* 201, 37–45. doi: 10.1016/j.foodchem.2016.01.012

Mutenda, K. E., Körner, R., Christensen, T. M. I. E., Mikkelsen, J., and Roepstorff, P. (2002). Application of mass spectrometry to determine the activity and specificity of pectin lyase A. *Carbohydr. Res.* 337, 1217–1227. doi: 10.1016/s0008-6215(02)00127-1

Niu, J., Alazi, E., Reid, I. D., Arentshorst, M., Punt, P. J., Visser, J., et al. (2017). An evolutionarily conserved transcriptional activator-repressor module controls expression of genes for D-galacturonic acid utilization in *Aspergillus niger*. *Genetics* 205, 169–183. doi: 10.1534/genetics.116.194050

Niu, J., Homan, T. G., Arentshorst, M., de Vries, R. P., Visser, J., and Ram, A. F. J. (2015). The interaction of induction and repression mechanisms in the regulation of galacturonic acid-induced genes in *Aspergillus niger*. *Fungal Genet. Biol.* 82, 32–42. doi: 10.1016/j.fgb.2015.06.006

Pel, H. J., de Winde, J. H., Archer, D. B., Dyer, P. S., Hofmann, G., Schaap, P. J., et al. (2007). Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. *Nat. Biotechnol.* 25, 221–231. doi: 10.1038/nbt1282

Plaschina, I. G., Braudo, E. E., and Tolstoguzov, V. B. (1978). Circular dichroism studies of pectin solutions. *Carbohydr. Res.* 60, 1–8. doi: 10.1016/S0008-6215(00)83459-X

Quéméner, B., Cabrera Pino, J. C., Ralet, M. C., Bonnin, E., and Thibault, J. F. (2003a). Assignment of acetyl groups to O-2 and/or O-3 of pectic oligogalacturonides using negative electrospray ionization ion trap mass spectrometry. *J. Mass Spectrom.* 38, 641–648. doi: 10.1002/jms.478

Quéméner, B., Désiré, C., Debrauwer, L., and Rathahao, E. (2003b). Structural characterization by both positive and negative electrospray ion trap mass spectrometry of oligogalacturonates purified by high-performance anion-exchange chromatography. *J. Chromatogr. A* 984, 185–194. doi: 10.1016/S0021-9673(02)01729-6

Ralet, M. C., Cabrera, J. C., Bonnin, E., Quéméner, B., Hellìn, P., and Thibault, J. F. (2005). Mapping sugar beet pectin acetylation pattern. *Phytochemistry* 66, 1832–1843. doi: 10.1016/j.phytochem.2005.06.003

Remoroza, C., Cord-Landwehr, S., Leijdekkers, A. G. M., Moerschbacher, B. M., Schols, H. A., and Gruppen, H. (2012). Combined HILIC-ELSD/ESI-MS n enables the separation, identification and quantification of sugar beet pectin derived oligomers. *Carbohydr. Polym.* 90, 41–48. doi: 10.1016/j.carbpol.2012.04.058

Renard, C. M. G. C., and Thibault, J.-F. (1996). "Pectins in mild alkaline conditions: beta-elimination and kinetics of demethylation," in *Pectins and Pectinases. Book Series: Progress in Biotechnology*, eds J. Visser, and A. G. J. Voragen, (Amsterdam: Elsevier), 603–608. doi: 10.1016/s0921-0423(96)80292-9

Sandri, I. G., Fontana, R. C., Barfknecht, D. M., and da Silveira, M. M. (2011). Clarification of fruit juices by fungal pectinases. *LWT Food Sci. Technol.* 44, 2217–2222. doi: 10.1016/j.lwt.2011.02.008

Seyedarabi, A., To, T. T., Ali, S., Hussain, S., Fries, M., Madsen, R., et al. (2010). Structural insights into substrate specificity and the anti β-elimination mechanism of pectate lyase. *Biochemistry* 49, 539–546. doi: 10.1021/bi901503g

Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D., et al. (2012). *Laboratory Analytical Procedure (LAP): Determination of Structural Carbohydrates and Lignin in Biomass*. Golden, CO: National Renewable Energy Laboratory.

Tan, H., Chen, W., Liu, Q., Yang, G., and Li, K. (2018). Pectin oligosaccharides ameliorate colon cancer by regulating oxidative stress- and inflammation-activated signaling pathways. *Front. Immunol.* 9:1504. doi: 10.3389/fimmu.2018.01504

Van Alebeek, G. J. W. M., Christensen, T. M. I. E., Schols, H. A., Mikkelsen, J. D., and Voragen, A. G. J. (2002). Mode of action of pectin lyase A of *Aspergillus niger* on differently C6-substituted oligogalacturonides. *J. Biol. Chem.* 277, 25929–25936. doi: 10.1074/jbc.M202250200

Van Alebeek, G. J. W. M., Zabotina, O., Beldman, G., Schols, H. A., and Voragen, A. G. J. (2000). Structural analysis of (methyl-esterified) oligogalacturonides using post-source decay matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J. Mass Spectrom.* 35, 831–840. doi: 10.1002/1096-9888(200007)35:7<831::aid-jms7>3.0.co;2-4

van Houdenhoven, F. E. A. (1975). *Studies on Pectin Lyase*. PhD thesis, Agricultural University, Wageningen.

Varga, J., Frisvad, J. C., Kocsubé, S., Brankovics, B., Tóth, B., Szigeti, G., et al. (2011). New and revisited species in *Aspergillus* section Nigri. *Stud. Mycol.* 69, 1–17. doi: 10.3114/sim.2011.69.01

Vitali, J., Schick, B., Kester, H. C. M., Visser, J., and Jurnak, F. (1998). The three-dimensional structure of *Aspergillus niger* pectin lyase B at 1.7-Å resolution. *Plant Physiol.* 116, 69–80. doi: 10.1104/pp.116.1.69

Voragen, A. G. J., Schols, H. A., and Pilnik, W. (1986). Determination of the degree of methylation and acetylation of pectins by h.p.l.c. *Top. Catal.* 1, 65–70. doi: 10.1016/S0268-005X(86)80008-X

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303. doi: 10.1093/nar/gky427

Xu, S. X., Qin, X., Liu, B., Zhang, D. Q., Zhang, W., Wu, K., et al. (2015). An acidic pectin lyase from *Aspergillus niger* with favourable efficiency in fruit juice clarification. *Lett. Appl. Microbiol.* 60, 181–187. doi: 10.1111/lam.12357

Zeuner, B., Muschiol, J., Holck, J., Lezyk, M., Gedde, M. R., Jers, C., et al. (2018). Substrate specificity and transfucosylation activity of GH29 α-L-fucosidases for enzymatic production of human milk oligosaccharides. *N. Biotechnol.* 41, 34–45. doi: 10.1016/j.nbt.2017.12.002

Conflict of Interest: MS and KK are employed by the company Novozymes A/S.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

REVIEW
published: 31 July 2020
doi: 10.3389/fbioe.2020.00871

Carbohydrate Binding Modules: Diversity of Domain Architecture in Amylases and Cellulases From Filamentous Microorganisms

Andika Sidar[1,2], Erica D. Albuquerque[1,3], Gerben P. Voshol[1,4], Arthur F. J. Ram[1], Erik Vijgenboom[1]* and Peter J. Punt[1,4]*

[1] Department of Microbial Biotechnology, Institute of Biology Leiden, Leiden, Netherlands, [2] Department of Food Science and Agricultural Product Technology, Faculty of Agricultural Technology, Universitas Gadjah Mada, Yogyakarta, Indonesia, [3] Sun Pharmaceutical Industries Europe BV., Hoofddorp, Netherlands, [4] Dutch DNA Biotech B.V., Utrecht, Netherlands

OPEN ACCESS

Edited by:
André Damasio, State University of Campinas, Brazil

Reviewed by:
Shiyou Ding, Michigan State University, United States
Jochen Schmid, Norwegian University of Science and Technology, Norway
Rasmus John Normand Frandsen, Technical University of Denmark, Denmark

***Correspondence:**
Andika Sidar
a.sidar@biology.leidenuniv.nl
Erik Vijgenboom
vijgenbo@biology.leidenuniv.nl

Specialty section:
This article was submitted to Bioprocess Engineering, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *20 April 2020*
Accepted: *07 July 2020*
Published: *31 July 2020*

Citation:
Sidar A, Albuquerque ED, Voshol GP, Ram AFJ, Vijgenboom E and Punt PJ (2020) Carbohydrate Binding Modules: Diversity of Domain Architecture in Amylases and Cellulases From Filamentous Microorganisms. Front. Bioeng. Biotechnol. 8:871. doi: 10.3389/fbioe.2020.00871

Enzymatic degradation of abundant renewable polysaccharides such as cellulose and starch is a field that has the attention of both the industrial and scientific community. Most of the polysaccharide degrading enzymes are classified into several glycoside hydrolase families. They are often organized in a modular manner which includes a catalytic domain connected to one or more carbohydrate-binding modules. The carbohydrate-binding modules (CBM) have been shown to increase the proximity of the enzyme to its substrate, especially for insoluble substrates. Therefore, these modules are considered to enhance enzymatic hydrolysis. These properties have played an important role in many biotechnological applications with the aim to improve the efficiency of polysaccharide degradation. The domain organization of glycoside hydrolases (GHs) equipped with one or more CBM does vary within organisms. This review comprehensively highlights the presence of CBM as ancillary modules and explores the diversity of GHs carrying one or more of these modules that actively act either on cellulose or starch. Special emphasis is given to the cellulase and amylase distribution within the filamentous microorganisms from the genera of *Streptomyces* and *Aspergillus* that are well known to have a great capacity for secreting a wide range of these polysaccharide degrading enzyme. The potential of the CBM and other ancillary domains for the design of improved polysaccharide decomposing enzymes is discussed.

Keywords: carbohydrate-binding module, cellulase, amylase, CAZymes diversity, domain architecture, *Aspergillus*, *Streptomyces*

INTRODUCTION

Plant biomass contains lignocellulose and starch which represent the most abundant carbohydrate biopolymers in nature. Lignocellulose is a complex polymer mix composed of lignin, an aromatic polymer and two carbohydrate polymers, cellulose and hemicellulose. The main component is cellulose consisting of β-1,4-linked D-glucose. In plant cell walls, cellulose chains interact with each other through hydrogen bonds to form microfibrils (Pérez et al., 2002; Somerville, 2006).

Like cellulose, starch is polymeric carbohydrate consisting of solely glucose molecules linked via α-glycosidic bonds, instead of β-glycosidic bonds. Starch is composed of a mixture of two polymers, amylose and amylopectin. Amylose is a linear polysaccharide composed of α-1,4-linked D-glucose. Amylopectin is also a polysaccharide composed of primarily linear α-1,4-linked D-glucose as the backbone with branches that are α-1,6-linked to the backbone (Buleon et al., 1998).

The degradation of the renewable biopolymers cellulose or starch into sugar monomers by industry has a great economic value for its use as a feedstock for the production of various value-added products, such as fuels, chemicals and foods (Zhu et al., 2016). In general, degradation of the polysaccharides into sugar monomers requires the synergistic action of several classes of carbohydrate-active enzymes (CAZymes). Amongst the CAZymes, the glycoside hydrolases (GHs) are the major enzyme family involved in the degradation of polysaccharides such as starch and cellulose (Janeček et al., 2014; Berlemont and Martiny, 2016; Nguyen et al., 2018). However, both cellulose and starch contain a fraction that resists hydrolysis and is indicated as the recalcitrant fraction (Jeoh et al., 2006). Full decomposition of the recalcitrant fraction cannot be obtained by available enzyme cocktails and thus improvement of enzyme performance is necessary (Leggio et al., 2015). One of the options to consider in improving GH characteristics toward the decomposition of the recalcitrant fraction is the engineering of chimeric GH by the addition of one or more carbohydrate-binding modules (CBM).

In the CAZy classification system, the CBMs are a large group of protein domains that are frequently found attached to GH enzymes (Lombard et al., 2014). The CBMs exist as single or multiple (duplicated) domains attached to the C- and/or N-terminus of the catalytic domain. They consist of a relatively small number of amino acids, ranging from approximately 30–200 amino acids. CBMs do not have catalytic activity but function as substrate binding modules (Boraston et al., 2004; Lombard et al., 2014). It has been suggested that the binding characteristic of a CBM improves catalytic function of the CAZymes through targeting the enzyme to the substrate and increasing substrate-enzyme proximity as well as disrupting the crystallinity of the insoluble substrate fraction (Arantes and Saddler, 2010; Reyes-Ortiz et al., 2013; Bernardes et al., 2019; Chalak et al., 2019). As a consequence, the removal of the CBM from the enzyme results in a decreased enzymatic activity (Arai et al., 2003; Cockburn et al., 2017) and a reduced enzyme stability (Bissaro et al., 2017; Courtade et al., 2018; Teo et al., 2019).

Originally, many CBMs were classified as cellulose binding domains (CBDs) because of their affinity for cellulose. The first of these domains was found appended to the cellobiohydrolase CBHI and cellobiohydrolase CBHII in the fungus *Trichoderma reesei* (Van Tilbeurgh et al., 1986; Tomme et al., 1988). Thereafter, similar CBDs but with different amino acid sequences were also found attached to the cellulases from *Cellulomonas fimi* (Gilkes et al., 1988). As several modules showed affinity to substrates other than cellulose but still met the criteria as carbohydrate binding domain, the nomenclature CBM was introduced to cover the large variety of carbohydrate binding domains. For example, CBM1 was used for the first discovered fungal CBDs and thereafter also for modules with related amino acid sequences. The CBM2 family contains the CBDs showing high similarity with the bacterial CBDs as found in *C. fimi* (Gilkes et al., 1988) and other CBM families followed based on their particular ligand specificity, amino acid similarity and structural characteristics (see review Boraston et al., 2004). Like cellulose binding domains, starch binding domains (SBDs) were initially found in fungi as a C-terminal domain of *Aspergillus niger* glucoamylase (Hayashida et al., 1982; Svensson et al., 1982) and were collected in the CBM20 family (Lombard et al., 2014). As of April 2020, 86 CBM families are listed in the CAZy database (Lombard et al., 2014).

Although it is already very well known that CBM domains are often associated with GHs and are suggested to have a role in enhancing the enzymatic degradation, only limited information is available related to how widely the CBMs are distributed within the GH classes and how these domains are organized in the GHs produced by a given microorganism. Multiple CBMs from different families can also be present in GHs at either the C- or N-terminus. It has been reported that the distribution of the polysaccharide degrading enzymes among the CAZy classes is highly variable within genera, including the GH families (Benoit et al., 2015; Berlemont, 2017). In that regard, mapping the CBM distribution within the GHs may provide novel insight for designing new enzyme architecture with the potential to improve polysaccharide degradation. One of the approaches is to design a chimeric enzyme containing one or more CBMs (Punt et al., 2011; Duan et al., 2017).

This review provides a comprehensive overview on the variation and distribution of CBM domains present in the GH families of cellulases and amylases from filamentous microorganism, particularly *Aspergillus* and *Streptomyces.* Filamentous microorganisms, such as *Streptomyces* (Book et al., 2016; Montella et al., 2017) and *Aspergillus* (Punt et al., 2002; Yuan et al., 2008; Benoit et al., 2015; Gruben et al., 2017) have a great potential to produce large numbers of CAZymes. Both genera are also well-known for the production of industrially important extracellular enzymes (Fleißner and Dersch, 2010; Sevillano et al., 2016). Many Aspergilli species, mainly the black *Aspergillus* strains such as *Aspergillus niger* and *Aspergillus oryzae* are used for the production of a variety of industrial enzymes such as amylases, pectinases, proteases, β-galactosidase, glucoamylase, lipase, phytase, protease, hemicellulase, and cellulase (Fiedler et al., 2013; Polizeli et al., 2016). *Streptomyces* strains are also used as production hosts for enzymes. *Streptomyces lividans* and *Streptomyces coelicolor* as the examples, have been applied for recombinant extracellular protein production, including hydrolases, proteases/peptidases, chitinases/chitosanases, cellulases/endoglucanases, amylases, and pectate lyases (Anné et al., 2014; Hamed et al., 2018; Ahmed et al., 2020). The striking variety of domain architectures in these organisms calls for a new approach toward enzyme engineering. Examples are discussed where the attention is focused on designing chimeric enzymes with *Streptomyces* and *Aspergillus* inspired domain organization for improved cellulose or starch degradation.

ENZYME FAMILIES INVOLVED IN STARCH AND CELLULOSE DEGRADATION

Starch-Degrading Enzymes (α-Amylase, Gluco-Amylase, β-Amylase)

All enzymes capable of degrading starch are collectively indicated as amylolytic enzymes. In this review, we focused on the main endo- and exo-amylolytic enzymes which are involved in saccharification of starch into glucose (**Figure 1A**). Endo-amylases, mainly α-amylases (1,4-α-D-glucan glucanohydrolase, EC 3.2.1.1) hydrolyse randomly the internal α-1,4-glucan chains to yield small oligosaccharides units (Gupta et al., 2003). Subsequently, gluco-amylases (1,4-α-D-glucan glucohydrolase, EC 3.2.1.3) as the main exo-amylolytic enzymes can hydrolyse linear α-(1,4) as well as the branching sites with an α-(1,6)-linkage at the non-reducing end of starch to release glucose (Sauer et al., 2000). Furthermore, β-amylases are another important class of exo-acting enzymes producing mainly maltose by cleaving starch molecules at the non-reducing end of α-1,4-glucan chains (Ray and Nanda, 1996). Their industrial application is essentially for maltose-rich syrup production (Niu et al., 2018). Overall, amylolytic enzymes have been largely used in a broad range of industrial applications, such as foods, detergents, pharmaceuticals, and the paper as well as textile industries (Ray and Nanda, 1996; de Souza and de Oliveira Magalhães, 2010).

In general, α-amylases are composed of three domains. Domain A has a TIM-barrel fold containing the active site residues and a chloride ion-binding site. Domain B is a long loop region that contains a calcium-binding site, and domain C is the C-terminal β-sheet domain that shows variability in sequence and length amongst amylases. Several amylases have at least one conserved calcium-binding site, as calcium is essential for enzyme stability, while the chloride binding site present in the active site region is important for the catalytic function. The active center of α-amylases, in general, involves mainly asparagine (Asp) and glutamic acid (Glu) residues (Pujadas and Palau, 2001; Mehta and Satyanarayana, 2016). For glucoamylase, the catalytic domain folds as a twisted (α/α)6-barrel, with the active site in a pocket shape at the N-terminal side of barrels. Two conserved glutamic acid residues are involved in the catalytic mechanism (Sauer et al., 2001; Aleshin et al., 2003). The β-amylase structure exhibits a well conserved (β/α) 8-barrel fold in the core domain with the active site as a deep cleft within the barrel where two glutamic acid (Glu) residues are involved in the hydrolysis (Mikami et al., 1994; Oyama et al., 2003).

In the sequence-based classification of GHs in the CAZy database, α-amylases are represented by four different families, GH13, GH57, GH119, GH126. The main α-amylase family, GH13, is found in a broad range of organisms, including bacteria, archaea, eukaryota, and even viruses. GH57 is found in bacteria, archaea, and very limited in eukaryotes. The other two families, GH119 and GH126 have been found in bacteria only (see review Janeček et al., 2014; Lombard et al., 2014). The glucoamylase family, GH15, does widely occur in prokaryotic and eukaryotic microorganisms including fungi (Aleshin et al., 2003), while the GH97 glucoamylases are predominantly found in bacteria (Lombard et al., 2014). Furthermore, the β-amylase family GH14 is in majority present in plants and bacteria (Lombard et al., 2014).

Cellulose Degrading Enzymes (Exo-Glucanase, Endo-Glucanase)

Almost 50% of lignocellulose is composed of cellulose (Somerville, 2006; Isikgor and Becer, 2015). The degradation of cellulose is accomplished by multiple GHs which are typically acting together as a cocktail with complementary and synergistic modes of action (Payne et al., 2015). For example, endoglucanase (EC 3.2.1.4) cleaves randomly the internal β-glycosidic bonds of cellulose, while exo-glucanases/cellobiohydrolases cleave at both

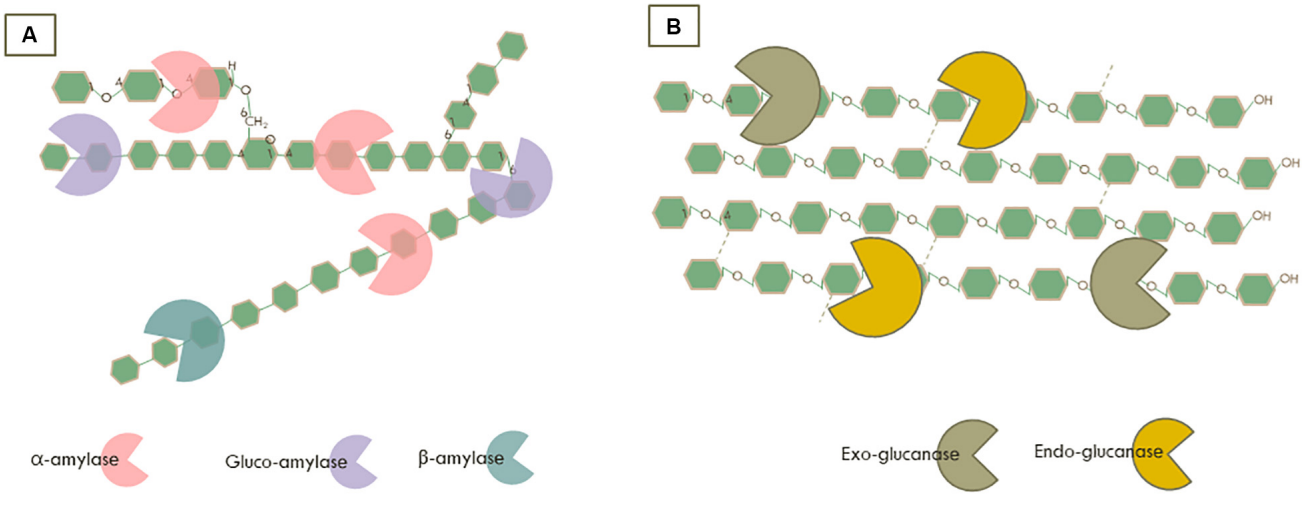


FIGURE 1 | Synergistic action of the enzymatic degradation by **(A)** Amylases on starch and **(B)** Cellulases on cellulose.

the non-reducing (EC 3.2.1.91) and reducing (EC 3.2.1.176) end of the cellulose chain to release soluble cellobiose (**Figure 1B**). These enzymes are used on a large scale for different industrial applications and are mainly produced by microorganisms, especially from the fungal and bacterial kingdoms.

Based on the active site topology, the catalytic machinery of cellobiohydrolases is set up in a tunnel-like conformation, enabling the single cellulose chain to be cleaved at the reducing or non-reducing terminus (Parkkinen et al., 2008). In contrast, the endoglucanase catalytic topology is shaped as an open cleft, allowing a linear amorphous cellulose chain to be degraded randomly at any part of the chain (Davies and Henrissat, 1995). Despite the fact that some enzymes share structural similarity, they may exhibit variations in the active site topology as well as the surface loops resulting in different substrate specificities. Based on structural modeling and enzymatic assays of GH6s from *Podospora anserina,* the exoglucanase PaCel6A has an active-site tunnel topology with the loop corresponding to amino acids 415–429 leading to the formation of the tunnel shape. In contrast the endoglucanase PaCel6B lacks the short 15 amino acid loop at the C-terminus and is predicted to contain a binding cleft rather than a tunnel structure. Those differences in the surface loop influence the mechanism of action toward substrates. Therefore, GH6 PaCel6B showed higher activity on carboxymethylcellulose (CMC) which is a highly specific substrate for endo-acting cellulase. CMC is decrystallized cellulose and thus contains more amorphous sites that are ideal for access to the cellulose chain by endoglucanases that cleave internally. Whereas, the exoglucanase GH6 PaCel6A showed higher activity on the insoluble microcrystalline cellulose Avicel than on CMC, as measured by the liberation of cellobiose from the non-reducing end of the cellulose chain (Poidevin et al., 2013). In addition, deletion of an exo loop in *C. fimi* cellobiohydrolase allows this enzyme to hydrolyze internal β-1,4-glucosidic bonds, altering its exolytic activity into endolytic activity (Meinke et al., 1995). Moreover, variations in the surface loop conformation were shown to determine the substrate specificity in GH5, such as exo- and endo-mannanase activity (Dias et al., 2004; Kumagai et al., 2015) as well as endoglucanase activity (Tseng et al., 2011). Therefore, in the CAZy classification enzymes could be clustered together as one GH family, although they are different in substrate specificity For example in members of the GH5 family a wide range of cellulase activities have been demonstrated: endo-β-1,4-glucanase, exo-β-1,4-glucanase, as well as non-cellulase activities: endo-β-1,4-xylanase, β-mannosidase, β-glucosylceramidase, glucan β-1,3-glucosidase, glucan endo-1,6-β-glucosidase, mannan endo-β-1,4-mannosidase, cellulose β-1,4-cellobiosidase, chitosanase, xyloglucan-specific endo-β-1,4-glucanase, endo-β-1,6-galactanase, β-1,3-mannanase, endo-β-1,3-glucanase/laminarinase, chitosanase, α-L-arabinofuranosidase (Lombard et al., 2014). In the CAZy database the endoglucanases are classified in 15 different families, GH5, GH6, GH7, GH8, GH9, GH10, GH12, GH26, GH44, GH45, GH48, GH51, GH74, GH124, and GH148. The exoglucanases are divided over 5 GH families, GH5, GH6, GH7, GH9, and GH48. Fungal cellulose degrading enzymes in these families commonly consist of a simple domain arrangement of only a catalytic domain or a catalytic domain with one accessory domain (Berlemont, 2017). Unlike the fungal cellulases, the bacterial cellulases often have a more complex domain structure which can also be involved in assembling multi-enzyme complexes, called cellulosomes (Bayer et al., 2004).

MULTI DOMAIN ARCHITECTURE AND DISTRIBUTION OF CBMS IN STARCH AND CELLULOSE SPECIFIC GLYCOSIDE HYDROLASES FROM THE GENUS *Streptomyces* AND *Aspergillus*

Many proteins within the various GH families show a distinct variation in the domain architecture. As discussed above, in particular the CBM domains are frequently associated with the GH catalytic domain. The abundance of GHs for cellulose degradation varies across bacterial phyla, of which the actinobacteria showed the most extensive variation in GH-cellulase content and are representatives of potential cellulose degraders (Berlemont and Martiny, 2015). For mapping the CBM and the GH diversity, datasets of cellulases and amylases carrying CBMs from *Aspergillus* and *Streptomyces* were retrieved from the UniProtKB and the CAZy databases. Subsequently, these datasets were grouped based on the type of CBMs and curated to avoid the inclusion of duplicated sequences by running protein alignment for each GH family in the same species by the Clustal Omega alignment program. Identical sequences from the same species were removed. Eventually, the final datasets were grouped based on the type of CBM associated with selected GHs. In addition, for constructing the graphical domain representation as presented in **Figures 2**, **3**, signal peptides and transmembrane segments were predicted using Phobius (Käll et al., 2004). Subsequently, Pfam domains were determined using the HMMER program (version 3.2.1) against the Pfam database version 31.0 (Eddy, 2011; El-Gebali et al., 2019). Finally, the predictions were programmatically extracted and converted into detailed graphical representations of the various domain organizations, as was previously done for the Auxiliary Activities CAZymes, the Lytic Polysaccharide Mono Oxygenases (LPMO) (Voshol et al., 2017, 2019).

In general, filamentous fungi represent a richer reservoir of CAZymes compared to bacteria (Berlemont, 2017; Voshol et al., 2019). However, the variability of enzymes associated with CBMs or other accessory domains is much less in fungi than in bacteria (Talamantes et al., 2016; Berlemont, 2017), as exemplified here by the comparison of the genera *Aspergillus* and *Streptomyces*. In total 2540 different amylase and cellulase amino acid sequences with CBMs from *Aspergillus* and *Streptomyces* were retrieved from the databases, representing 385 different *Streptomyces* species and 60 different *Aspergillus* species. Amongst the sequences, approximately 90% of the GHs cellulase and amylase with CBM were identified from *Streptomyces*. In contrast, GH cellulases without CBM are more widespread in fungi representing about 80% of these cases. Thus, as a first difference between these two genera, many

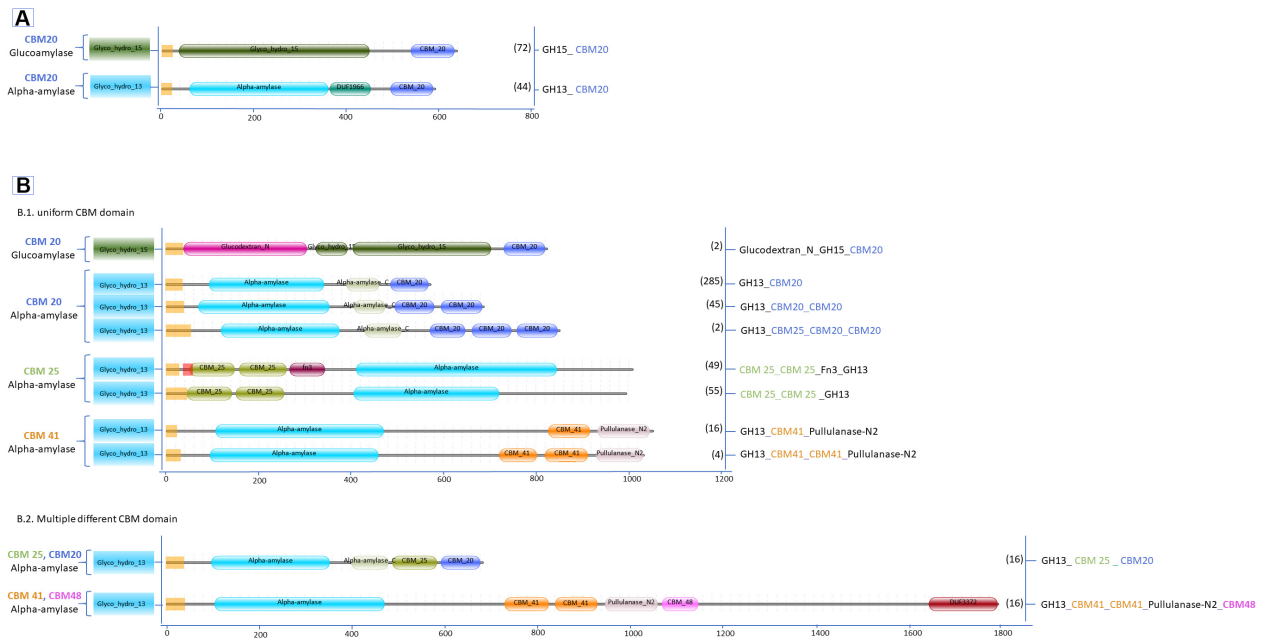


FIGURE 2 | Domain organization of amylases carrying CBM **(A)** *Aspergillus*; **(B.1)** *Streptomyces* single CBM and **(B.2)** *Streptomyces* multiple CBM domains. Between brackets is the number of protein sequences in the various *Aspergillus* and *Streptomyces* species retrieved from the UniProtKB and the CAZy databases.

cellulases and amylases from *Aspergillus* lack any CBM, while *Streptomyces* enzymes contain many CBM types as well as a very diverse enzyme architecture associated with different CBMs (**Figures 2**, **3**).

Amylases With CBM From *Aspergillus* and *Streptomyces*

Among the starch degrading enzymes only two GH families are associated with CBMs in these two genera, the GH13 α-amylases and the GH15 glucoamylases. Neither *Aspergillus* nor *Streptomyces* contain genes encoding GH14 β-amylases with a CBM (**Figure 2**). As reported by the CAZy database and several studies, β-amylases are secreted mainly by plants, followed by several bacteria, and by a very limited number in fungi (Ray, 2004; Derde et al., 2012; Thalmann et al., 2019). Two CBM families were identified in bacterial β-amylases, i.e., CBM20 from species of *Bacillus* and *Clostridium* as well as CBM25 from *Paenibacillus,* and no CBMs were found in the fungal GH14 β-amylases (see review Janeček et al., 2019).

In total, 606 protein sequences with starch binding CBMs were identified in *Aspergillus* and *Streptomyces* amylases. Among these sequences, more than half (466 sequences) are associated with CBM20. In *Aspergillus* amylases, CBM20 was found as the only starch binding domain associated with GH13 α-amylases and GH15 glucoamylases (**Figure 2**). More CBM families were found in conjunction with amylases from *Streptomyces*. CBM20 was found as the most frequently domain (68%), followed by CBM25 (21%), leaving the rest (around 4% each) to CBM41 and amylases with a combination of CBMs: CBM20&CBM25 and CBM41&CBM48. Thus, in amylases from these genera, CBM20 is the most dominant CBM (**Figure 4**).

In term of diversity of domain architecture, *Aspergillus* species tend to have a simple domain architecture consisting of only a single CBM20 module located at the C-terminus (**Figure 2A**). On the other hand, not only higher in the variability of CBM families, but also a more divers domain architecture of amylases with CBMs were identified in *Streptomyces*. For example, CBM20 and CBM41 are present as a single module, or multiple repeated domains from the same family. Moreover, some of the *Streptomyces* amylases have combinations of different CBMs attached: CBM20&CBM25 as well as CBM41&CBM48 (**Figures 2B1,B2**). In general, the binding modules of the CBM20 family are found in a single copy, while only in a few special cases a CBM20 is present together with other CBMs, such as CBM25, CBM34, and CBM48 (Janeček et al., 2019).

Based on the domain position, the CBM20s in α-amylases from *Aspergillus* and *Streptomyces* are always located at the C-terminal end (**Figure 2**). This is also true for other starch active enzymes, such as in GH14 (Oyama et al., 1999), GH15 (Bott et al., 2008; Roth et al., 2018), LPMO AA13 (Vu et al., 2014), and in most of the GH13 enzymes (see review Janeček et al., 2019). Based on the CAZy database, only in a few GH13s, the CBM20 is located at the N-terminal end of the catalytic domain, especially the GH13s from green algae (Lombard et al., 2014). Moreover, in the *Streptomyces*α-amylases, the CBM25 can be present either N- or C-terminally. Two copies of a CBM25 domain can be present at the N-termini, while co-occurrence of CBM25 with CBM20 at the C-terminus occurs. The other starch binding domains, CBM41 and the combination of CBM41 & CBM48, were found exclusively C-terminally.

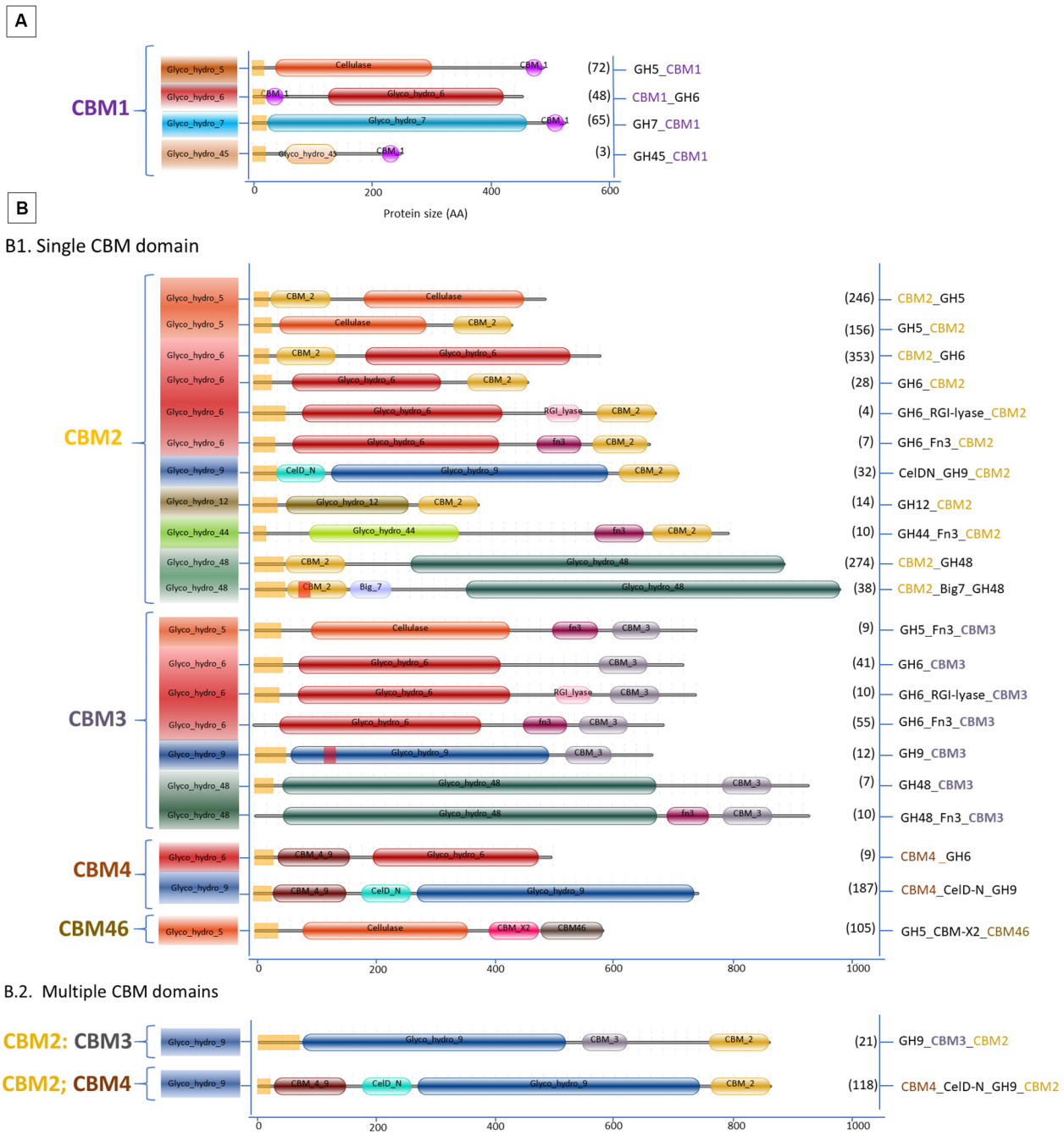


FIGURE 3 | Domain organization of cellulases carrying CBM. **(A)** *Aspergillus*; **(B.1)** *Streptomyces* single CBM and **(B.2)** *Streptomyces* multiple CBM domains. Between brackets is the number of protein sequences in the various *Aspergillus* and *Streptomyces* species retrieved from the UniProtKB and the CAZy databases.

Single or two copies of CBM41 can be found in several bifunctional enzymes, such as amylopullulanases (bifunctional with one catalytic center) and bifunctional α-amylase-pullulanases with two catalytic domains, one from α -amylase and one from pullulanase (**Figure 2B**). According to substrate binding studies, the CBM41 was demonstrated to be able to bind substrate with both 1,4- and 1,6 α-glucosidic linkages (Mikami et al., 2006; Van Bueren and Boraston, 2007). The CBM41s were largely identified in bacterial amylases and in only five members of eukaryotes, including four types of algae and a single species of fungi called *Ostreococcus tauri* (Lombard et al., 2014). CBM41s are often located at the N-terminus of the protein and present in two copies, while in a few cases, CBM41 is positioned C-terminally (see review Janeček et al., 2019). Interestingly, in bifunctional amylases from *Streptomyces,* CBM41 is present either as a single module or two repeated modules in between

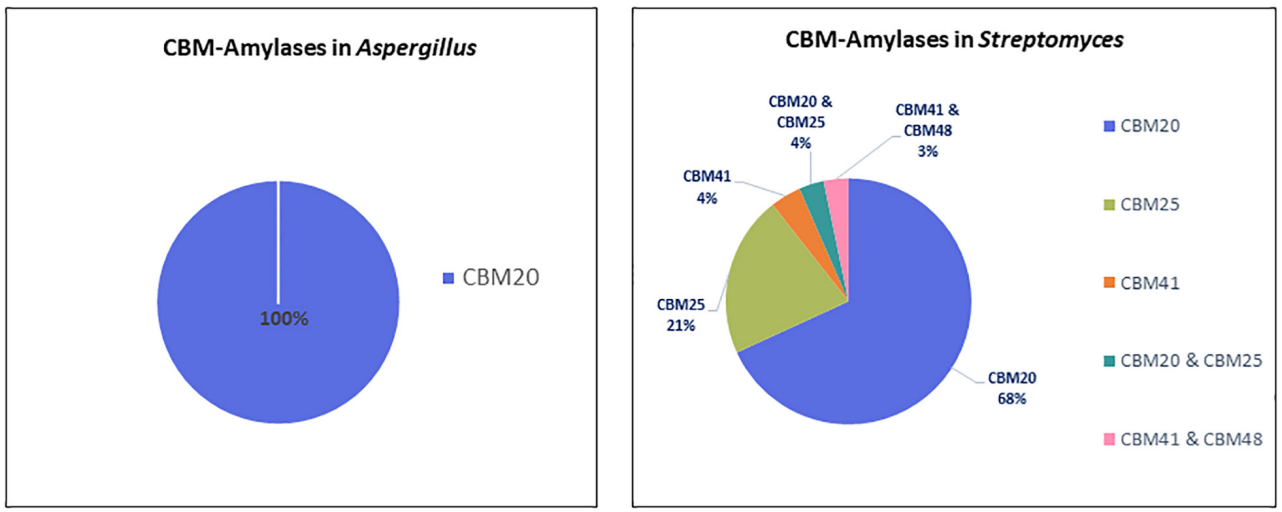


FIGURE 4 | CBMs associated with amylases from *Aspergillus* (left) and *Streptomyces* (right).

the amylase catalytic domain and the pullulanase domain (**Figure 2B1**), and even in combination with aCBM48 located C-terminal of the pullanase domain (**Figure 2B2**).

Other accessory domains present in the amylase architecture were annotated as Fn3, Glucodextran-N, DUF1966 and DUF 3372 (**Figures 2A,B**). In *Streptomyces*, the Fn3 domain is found in combination with CBM25. The CBM25-Fn3 domain combination is also present in *Microbacterium aurum* α-amylase and suggested to improve binding on starch indicated by the alteration of starch granule morphology (Huang et al., 2017). Furthermore, gradual domain truncations were made in the amylopullulanase from *Thermoanaerobacterium saccharolyticum* NTOU which consists of two Fn3 modules in between the N-terminal catalytic domain and C-terminal CBM20. The amylopullulanase remains active after removal of theCBM20 and the 2nd Fn3 module but further deletion of 1st Fn3 module that is connected with the catalytic module resulted in completely loss of activity (Lin et al., 2011). Generally, very little information is available on the function of the Fn3 modules in amylases. However, there is an indication that Fn3 is involved in binding between the enzyme and the polysaccharide substrates since its removal results in decreased or even loss of enzymatic activity. Another accessory domain, annotated as glucodextran_N in Pfam (El-Gebali et al., 2019) is especially found in bacterial and archaeal glucoamylases and glucodextranases (Mizuno et al., 2004). The domain consists of 17 antiparallel β-strands (Mizuno et al., 2004) which appear to stabilize its structure (Aleshin et al., 2003). This domain is absent in fungal glucoamylases. It was suggested that the eukaryotic glucoamylases may have evolved from prokaryotic glucoamylases together with substitution of the N-terminal glucodextran_N domain with the so-called peripheral subdomain and by the addition of a C-terminal starch-binding domain (Aleshin et al., 2003). The peripheral subdomain is an extra α-helix that is smaller than the other 12 α-helices in the catalytic domain structure and is located between α-helices αH10 and αH11 in the catalytic domain of glucoamylases *A. niger* (Aleshin et al., 1992, 2003; Lee and Paetzel, 2011). Lastly, several domains of unknown function, so-called DUFs, are present in the GH13 α-amylases of *Aspergillus* and the bifunctional α-amylase-pullulanase in *Streptomyces*, as DUF1966 and DUF3372, respectively. This domain folds in a β-barrel structure, however, the exact function has not been determined yet. Moreover, the truncation of this domain in the amylase from *Sclerotinia sclerotiorum*, ScAmy43, had no effect on the biochemical properties of the enzyme (Abdelmalek et al., 2009).

Cellulases With CBM in *Aspergillus* and *Streptomyces*

Based on the datasets retrieved, various cellulase (GH) families do contain a CBM domain in *Aspergillus* and *Streptomyces* (**Figure 3**).

Interestingly, members of two GH families (GH7 and GH45) containing a CBM in *Aspergillus* are absent in *Streptomyces*. According to Segato et al. (2012), the genome of almost every *Aspergillus* encodes two GH7-cellobiohydrolases which can be present with or without a CBM. Meanwhile, three families, GH9, GH44, and GH48, containing a CBM present in *Streptomyces* are completely lacking in *Aspergillus*. The GH12 family proteins are present in both genera, but only in *Streptomyces* the GH12 have a CBM. The GH5 and GH6 cellulases with a CBM are found in both genera.

In the *Aspergillus* and *Streptomyces* genomes, 1946 sequences of cellulases with CBM domain were identified. Among these sequences 188 were annotated as cellulase with a single CBM in *Aspergillus*, 1256 as cellulase with a single CBM in *Streptomyces* and 502 as cellulases with multiple CBMs in *Streptomyces*. **Figure 3** shows that *Streptomyces* cellulases are not only higher in their diversity of the CBM type, but also more varied in the domain organization compared to *Aspergillus*. In *Streptomyces,* four different CBMs associated with the GH cellulase catalytic domain were identified, including CBM2,

CBM3, CBM4, and CBM46. Moreover, each CBM is not only present as a single module, but also appears in combination with other accessory domains including another CBM (CBM2&CBM4 and CBM2&CBM3). CBM2 is the most frequently found CBM (67%) among the four different CBMs in *Streptomyces* cellulases (**Figure 5**). In contrast to this diversity, in *Aspergillus,* CBM1 is the only module associated with GH cellulases and always present in a simple arrangement of one catalytic domain plus one CBM. Unlike observed for the amylases, the *Aspergillus* and *Streptomyces* cellulases do not share a common CBM.

The CBM1s are found almost exclusively in fungi, while the CBM2s are mainly found in cellulases from bacterial origin. Based on the domain structure and function, CBM1s are relatively small, consisting of approximately 36 amino acids with 4 conserved cysteine residues involved in the formation of two disulphide bridges required for protein folding (Mattinen et al., 1997). The fold of CBM1 is called a cysteine knot (Boraston et al., 2004). This knotted arrangement is usually associated with a β-sheet structure and highly conducive to protein stability (Cheek et al., 2006). In contrast, the bacterial cellulose binding modules like CBM2 are usually much larger in size, about 110 amino acid residues with a β-sandwich fold providing the ligand recognition site (Boraston et al., 2004). In general, a regular plane surface with three aromatic residues such as tyrosine, as well as polar residues glutamine and asparagine are involved in CBM1-cellulose binding (Beckham et al., 2010). Likewise, the CBM2 β-sandwich fold presents a planar surface containing conserved aromatic residues, usually tryptophan and tyrosine, involved in substrate binding. Although the CBMsin *Aspergillus* and *Streptomyces* are different in protein structure, creating a chimeric enzyme by fusing a bacterial CBM with a fungal cellulase has potential to create an active enzyme. As reported by Voutilainen et al. (2014), fusing a bacterial CBM3 to a GH7 cellobiohydrolase from the thermophilic fungus *Talaromyces emersonii* which in fungi is often associated with a CBM1 domain, resulted in an active chimeric GH7 with higher thermal stability.

CBM1s are mainly located at the C-termini of GH5, GH7, and GH45. Only in the GH6 of *Aspergillus*, the CBM1 is positioned at the N-terminus (**Figure 3A**). Similar domain structures are also present in other fungi such as *Trichoderma* (Rahikainen et al., 2019). In a few special cases multiple CBM1 domains are present in a single protein. For example, five different CBM1 modules are arranged in tandem at the N-terminus of GH45 endoglucanase from the yeast, *Pichia pastoris* GS115 (Couturier et al., 2011). Furthermore, in *Streptomyces* the CBM2 is present either at the N- or C-terminus of GH cellulases. Among those, 76% were detected at the N-terminus of the cellulase catalytic domain, such as in CBM2_GH5, CBM2_GH6, and CBM2_GH48. The remaining 16 and 8% of the CBM2 were found at the C-terminus, connecting one catalytic domain with other accessory domains or different CBMs (**Figures 3B1,B2**, respectively). A CBM2 is widely occurring in cellulases from bacteria but much rarer in eukaryotes with only a few examples from gastropods and nematodes.

In the cellulase domain organization, CBM3 is always located at the C-terminus of the catalytic domain and represents the second most frequently found CBM in the domain architecture of cellulases (**Figure 3B**). Besides its presence as a single module, CBM3 is also commonly found in a multidomain architecture. In *Streptomyces,* CBM3s were mostly associated with GH6 in co-occurrence with Fn3 domains and linked with CBM2 in the GH9 multidomain architecture. A CBM3-Fn3-Fn3-CBM2 architecture was also found in the *Thermobifida cellulosilytica* TB100 GH9-endoglucanase (Tóth et al., 2017). A CBM3 domain consists of 150 amino acids and has a β-sandwich fold with nine strands and one calcium binding site (Tormo et al., 1996; Shimon et al., 2000).

In contrast to CBM3, a single CBM4 domain is always present at the N-terminus in the GH6 and GH9 family from *Streptomyces* (**Figure 3B1**). In the GH9 multidomain architecture, CBM4 can also be present together with CBM2 and another accessory domain, CelD-N. In general, the CBM4 modules are found in

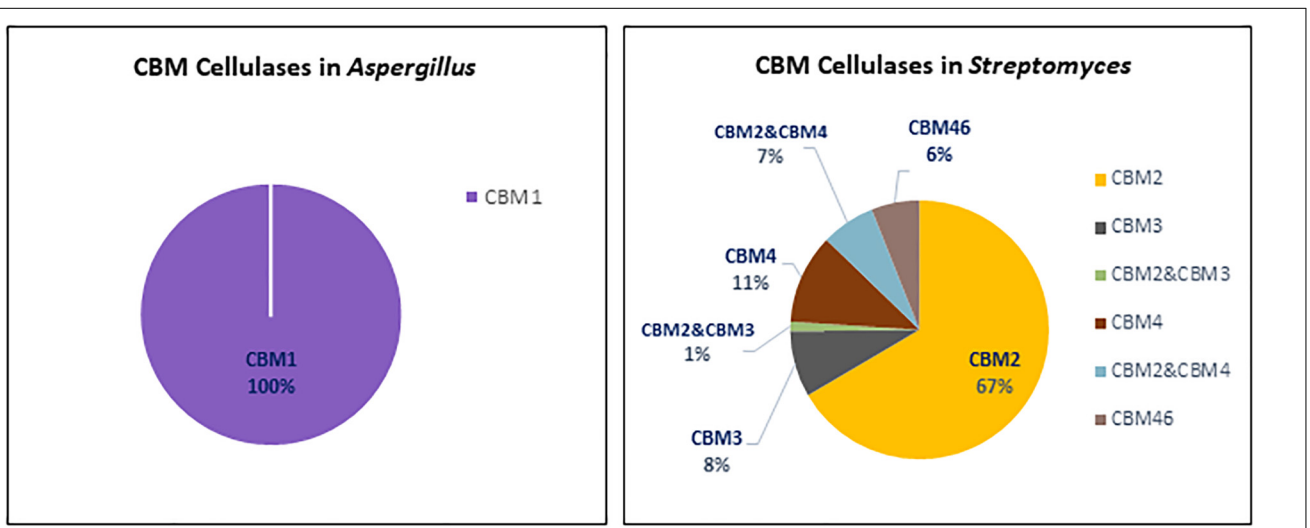


FIGURE 5 | CBMs associated with cellulases from *Aspergillus* (left) and *Streptomyces* (right).

xylanases and endoglucanases in agreement with their binding affinity toward either xylan or β-glucan (Abou-Hachem et al., 2000; Boraston et al., 2002; Simpson et al., 2002). Several binding studies with CBM4 showed differences in substrate binding preference. The differences in substrate binding affinity of CBM4 is determined by the length and topography of the surface loops. The long shallow groove conformation in CBM4 *Cellulomonas fimi* provides better binding to 1,4-β-glucan than 1,3-β-glucan (Coutinho et al., 1992).,The prominent U-shape with high-sided walls formed by loops between the β-strands from *Thermotoga maritima* CBM4 contribute to the higher affinity to 1,3-β-glucan (Zverlov et al., 2001). The conserved aromatic clusters in CBM4 containg tyrosine and/or tryptophan contribute to the surface binding site and are suggested to be involved in binding specificity (Boraston et al., 2002). Furthermore, replacing the CBM1 of a fungal acetyl xylan esterase with the CBM4 from the thermophilic bacterium *Rhodothermus marinus* Xyn10A enhanced the specific activity and the thermostability of the acetyl xylan esterase (Liu and Ding, 2016). Indicating that creating a chimeric enzyme between a mesophilic carbohydrate esterase and a CBM4 from a thermophile does result in improved thermophilic properties.

Another CBM present in *Streptomyces* cellulases is CBM46. According to the CAZy database, CBM46 is exclusively found in bacterial cellulases. In GH cellulase family 5, the CBM46 is present in combination with an accessory module called CBM_X2 (PF03442). A similar architecture was also found in GH5s of *Bacillus halodurans* which are associated with CBM46 and CBM_X2 (Venditto et al., 2015). The CBM_X2 or generally also called the X2 module is the immunoglobulin (Ig)-like domain with capability to bind cellulose and bacterial cell walls (Mosbah et al., 2000; Kosugi et al., 2004; Liberato et al., 2016). Moreover, the catalytic activity of GH5 from *Bacillus licheniformis* is entirely dependent on its CBM46 and X2 modules (Liberato et al., 2016). Besides in GH5, the X2 module is also present in GH74 endoglucanase from *Paenibacillus polymyxa* A18 with the arrangement GH74_X2_CBM3. Based on the enzyme activity studies, endoglucanase with X2-CBM3 domains have a several fold increase in activity toward insoluble substrates. In substrate binding studies, the module X2 showed a higher affinity toward phosphoric acid swollen cellulose, whereas CBM3 showed a higher affinity toward crystalline cellulose (Avicel) (Pasari et al., 2017).

Other Accessory Domains: The Fn3 Domain

Other types of accessory modules found in *Streptomyces* GH cellulases are the Fn3 domain and the variants of the X2 module CelD-N and Big7. According to the Pfam database (El-Gebali et al., 2019), domains annotated as X2, CelD_N and Big7 are a group of bacterial Ig-like domain which are often associated with bacterial GHs cellulases and suggested to be involved in assembling multienzyme complexes called cellulosomes. The Fn3 domain is a domain composed of a seven-stranded β-sandwich, usually occurring in multiple copies in both intracellular and extracellular proteins.

Further investigation of the presence and role of a Fn3(-like) domain in GH cellulases, was performed by querying the genomes of both *A. niger* and *S. leeuwenhoekii*, as representatives of the two genera we have focused or research on and mapping the Fn3 domains present in these and other proteins involved in polysaccharide degradation (**Table 1**). Indeed, besides cellulases there are several proteins with Fn3 domains, including GH3 β-glucosidase, GH3 β-xylosidase, GH5 mannosidase, GH18 chitinase, and chitin biosynthesis proteins with several types of domain organization. This wide occurrence of Fn3 domains, suggests that the Fn3 domain could be of functional significance.

As can be seen from **Table 1**, Fn3 is often present between the catalytic domain and a CBM in *Streptomyces*, in correspondence to the suggestion that it plays a role as stable linker between the enzymatic and substrate binding domain (Jee et al., 2002; Valk et al., 2015). However, the Fn3 domain in *Aspergillus* GH3 is positioned at the C-terminal end, and not followed by a CBM. Several other fungal β-glucosidases carry Fn3 domains, such as those from *Aspergillus aculeatus* and *Trichoderma reesei* (Suzuki et al., 2013), as well as the thermophilic fungus *Rasamsonia emersonii* (Gudmundsson et al., 2016).

Although Fn3 is suggested as a stable linker (Jee et al., 2002), many studies showed that the Fn3 domain also has relevance to a ligand binding function and is involved in protein-protein interaction (Hansen, 1992; Koide et al., 1998). From scanning electron microscope studies, it is concluded that Fn3 present in a bacterial cellobiohydrolase contributes to cellulose surface disruption, and therefore the presence of Fn3 increases the efficiency of degradation (Kataeva et al., 2002). Similarly, it has also been demonstrated that an isolated CBM2 from *Cellulomonas fimi* cellulase disrupted the surfaces of cellulose fibers (Din et al., 1991), indicating a potential functional similarity between Fn3 and a genuine CBM. Moreover, removal of the Fn3 domain can result in decreased or loss of enzymatic activity, as we found in a study on the GH3 function (**Figure 8**). As other examples, removal of the Fn3 domain of a xylanse dramatically decreased xylanolytic activity from *Cellulosimicrobium* sp. strain HY-13 (Kim et al., 2009), *Flavobacterium johnsoniae* (Chen et al., 2013), a *Marinifilaceae* bacterium strain SPP2 (Han et al., 2019) and cellulase activity for the GH6 cellobiohydrolase in another bacterial species (Cerda-Mejía et al., 2017). Deletion of the Fn3 domains also resulted in reduced enzymatic activity of GH55 endoglucanase (Conway et al., 2016) and GH9 endoglucanase (Zhou et al., 2004). Likewise, a CBM-truncated enzyme, in some cases, showed a significant loss of enzymatic activity, such as in GH9 (Burstein et al., 2009) and GH5 (Zheng and Ding, 2012). Surprisingly, our own results suggested that designing a GH3 chimeric enzyme by replacing the Fn3 domain with a *Streptomyces* CBM2 domain does not result in a functional GH3 enzyme (**Figure 8**, Sidar et al., in preparation). Thus, a function of the Fn3-like domain may be both to potentiate the activity of the GHs and to increase the substrate proximity to the catalytic domain in a similar fashion as the action of a CBM. Similar as found for CBM1 and CBM2 (see above) the (C-terminus of the) Fn3 domain consists of clusters of aromatic

residues, indicating that it could indeed play a role in ligand binding interaction, instead of solely existing as a linker (Lima et al., 2013). Another similarity between Fn3 and CBM is that some CBMs have a binding site for calcium which plays a significant role in the substrate interaction (Tormo et al., 1996; Montanier et al., 2010; Yaniv et al., 2012). Kataeva et al. (2002) also experimentally showed that each Fn3 domain in cellobiohydrolase CbhA binds calcium. The literature evidence for the Fn3 domain performance so far suggest a putative dual role of this domain. We hypothesize that Fn3/ Fn3-like domains play a biological role both in potentiation the functionality of the catalytic center and as a functional domain, rather than merely as a simple spacer, and may even represent a novel CBM-like domain.

DESIGNING NEW CHIMERIC ENZYMES INSPIRED BY THE DOMAIN ARCHITECTURE

In the field of enzyme engineering, the design of chimeric enzymes has been developed as one of the promising approaches to obtain an enzyme with the desired characteristics like improved hydrolysis efficiency. A chimeric enzyme is commonly generated by fusing the GH catalytic domain obtained from one species with another protein domain such as a CBM from another species. In several studies, chimeric enzymes have been shown to improve the catalytic efficiency, thermostability, as well as substrate specificity (Oliveira et al., 2015; Saadat, 2017; Christensen et al., 2020; Gilmore et al., 2020).

As shown above *Streptomyces species* display a much higher diversity in cellulase domain architecture than *Aspergillus*. This observation led us to review the *Streptomyces* domain architecture in more detail as the basis for design of novel chimeric enzymes with cellulolytic activity. As an example, we selected *Streptomyces leeuwenhoekii* C34 (Busarakam et al., 2014) and *Aspergillus niger*.

The variation in the domain organization of cellulases tethered to a CBM in *S. leeuwenhoekii* is presented in **Figure 6**. The cellulases of the GH5, GH6, GH9, and GH48 families are all associated with CBM2. When expanding our genome mining efforts in *S. leeuwenhoekii*, to proteins associated with a CBM2, other than cellulases, a considerable group of enzymes such as xylanases, chitinases, LPMOs and an esterase showed up from the genome sequence, in addition to a GH3 family β-glucosidase (**Figure 7**). The same mapping of enzymes containing a CBM1 in *A. niger*, resulted in only two other proteins, both related to polysaccharide degradation, being a cellulose active LMPO and a GH74 xyloglucanase (**Figure 7**).

In fungal cellulolytic enzyme cocktails, the β-glucosidases are considered to be the limiting factor in obtaining full conversion of cellulose to glucose. The GH3 family β-glucosidase in filamentous fungi do not have native CBM domains (see **Figure 3**). Therefore, we suggest that based on the *Streptomyces* domain architecture, a new concept to design chimeric enzymes could be studied by fusing a CBM2 domain from a *Streptomyces* GH3 β-glucosidase to a well-known *A. niger* GH3 β-glucosidase. However, it should be kept in mind that the design of new chimeric enzymes is not as simple as replacing one domain for another as shown by the example of replacing a Fn3 domain by CBM2 (**Figure 8** and section "Other Accessory Domains: The Fn3 Domain").

TABLE 1 | Protein domain associated with Fn3 in *A. niger* and *S. leeuwenhoekii.*

Organism	Domain associated with Fn3	Main activity	Accession number (NCBI)
A. niger	GH3_Fn3-like	β-glucosidase	CAK48740.1
	GH3_Fn3-like	β-xylosidase	CAK37179.1
	RhgB__Fn3_CBM-like	Rhamnogalacturonat lyase	XP_001400741.2
	Chs5n_Fn3_BRCT-CHS5like	Chitin biosynthesis protein	XP_001395839.2
S. leeuwenhoekii	GH5_Fn3_CBM2	Mannosidase	CQR60862.1
	CBM2_Fn3_GH18	Chitinase	CQR61740.1
	CBM4_Fn3_GH18	Chitinase	WP_029384933.1

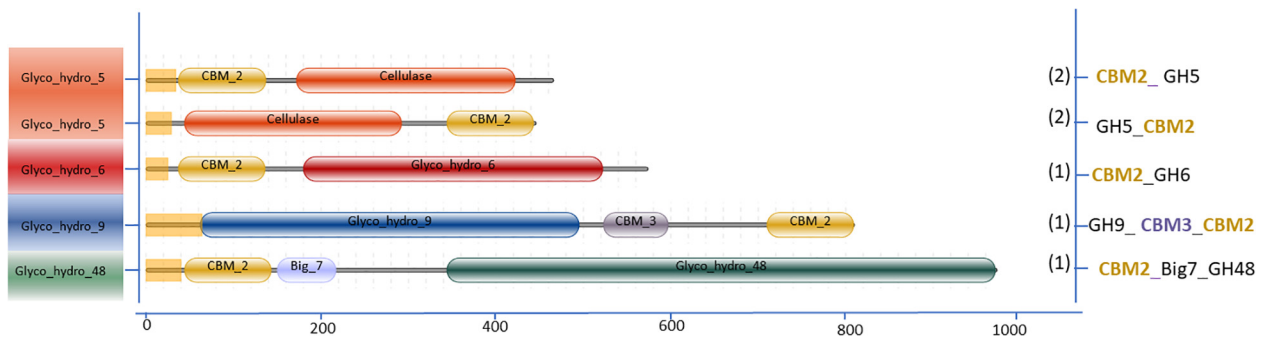


FIGURE 6 | Cellulases associated with CBM2 in *S. leeuwenhoekii* C34.

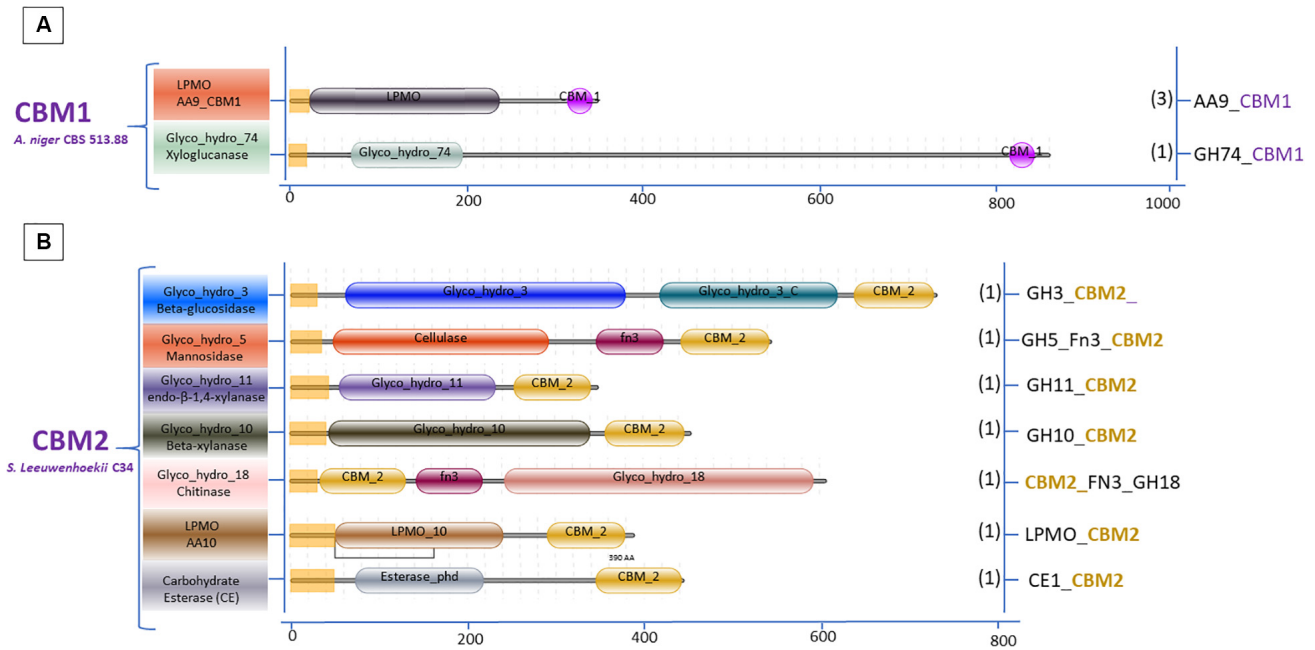


FIGURE 7 | Domain organization of polysaccharide degrading enzymes other than cellulases associated with two dominant CBMs: **(A)** CBM1 *A. niger* 513.88 and **(B)** CBM2 *S. leeuwenhoekii* C34.

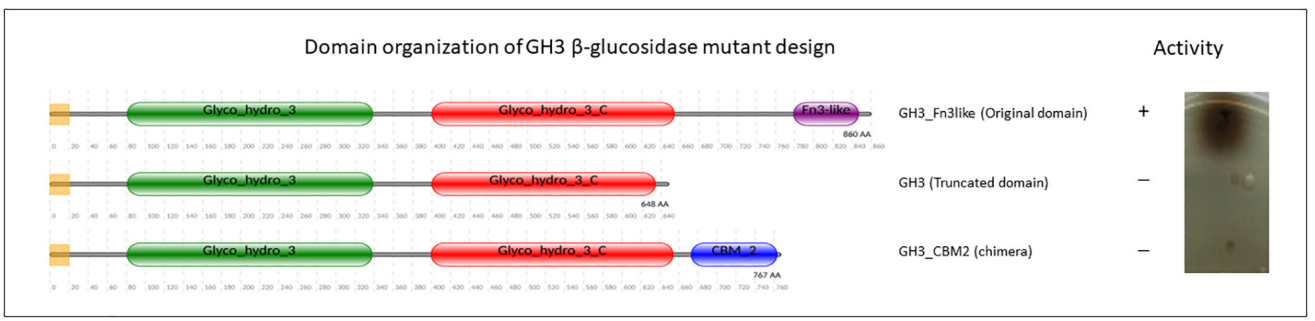


FIGURE 8 | GH3 β-glucosidase activity determined using an agar plate assay with esculin as substrate (Pointing, 1999).

FUTURE PERSPECTIVE

As functional domains, the CBMs have attractive characteristics like promoting substrate binding and thus supporting the catalytic function of enzyme. So far, CBMs have been classified in more than 80 families and more CBM families will undoubtedly be discovered. CBMs appear in association with a wide range of proteins in particular the CAZymes and are assembled in various domain organizations in numerous microorganisms. As an example, the diversity in the domain organizations in *Aspergillus* and *Streptomyces* is presented in this review. The *Streptomyces* species harbor cellulases and amylases with a more diverse domain organization, including CBMs and other accessory domains, than found in *Aspergillus*. This remarkable feature provides inspiration to design chimeric enzymes with a potential for substrate degradation improvement.

Fusions of enzyme domains have been developed in many scientific studies to create improved chimeric enzymes. In one of our studies, the *Aspergillus* GH3 enzyme that does not have a CBM but does have a C-terminal Fn3-like domain, was redesigned to replace the Fn3 with a CBM2 of *Streptomyces*. To our surprise these studies revealed that the Fn3-like domain of the fungal GH3 seems to play a pivotal role in the enzymatic activity (**Figure 8**), This observation might indicate that in addition to the well-known CBMs more domains could play an important role in the activity of CAZymes. Clearly, creating chimeric enzymes by modifying domain arrangement could provide advantages for enhancing enzymatic activity. Overall, the variability in the CBM domain organization of cellulases and amylases, as shown in this review, offers new opportunities to rethink the strategy for designing multiple-domain enzymes. This approach is not limited to chimeric cellulases and amylase

to improve degradation of recalcitrant substrate but also holds promise for other CAZymes such as xylanases, pectinases, and oxidative enzymes.

AUTHOR CONTRIBUTIONS

AS analyzed the data, wrote the review, performed the literature data mining, and designed the figures. EA conducted the cellulases analysis of *Streptomyces*. GV performed the bioinformatic analysis for the graphical domain annotation. AR presented ideas and edited the review. EV and PP involved in the review outline, ideas, planning, supervision, writing, and editing. All authors reviewed and agreed on the final manuscript.

FUNDING

This study was supported by a scholarship of the Indonesia Endowment Fund for Education (LPDP) from the Ministry of Finance, Indonesia to AS and by a grant of the Dutch National Organization for Scientific Research NWO, in the framework of an ERA-IB project FilaZyme (053.80.721/EIB.14.021) to EV, GV, and PP.

REFERENCES

Abdelmalek, I. B., Urdaci, M. C., Ali, M. B., Denayrolles, M., Chaignepain, S., Limam, F., et al. (2009). Structural investigation and homology modeling studies of native and truncated forms of α-amylases from *Sclerotinia sclerotio*rum. *J. Microbiol. Biotechnol.* 19, 1306–1318. doi: 10.4014/jmb.0903.03013

Abou-Hachem, M., Karlsson, E. N., Bartonek-Roxa, E., Raghothama, S., Simpson, P. J., Gilbert, H. J., et al. (2000). Carbohydrate-binding modules from a thermostable *Rhodothermus marinus* xylanase: cloning, expression and binding studies. *Biochem. J.* 345, 53–60. doi: 10.1042/bj3450053

Ahmed, Y., Rebets, Y., Estévez, M. R., Zapp, J., Myronovskyi, M., and Luzhetskyy, A. (2020). Engineering of *Streptomyces lividans* for heterologous expression of secondary metabolite gene clusters. *Microb. Cell Fact.* 19:5. doi: 10.1186/s12934-020-1277-8

Aleshin, A., Golubev, A., Firsov, L. M., and Honzatko, R. B. (1992). Crystal structure of glucoamylase from *Aspergillus awamori* var. X100 to 2.2-Å resolution. *J. Biol. Chem.* 267, 19291–19298. doi: 10.2210/pdb1gly/pdb

Aleshin, A. E., Feng, P.-H., Honzatko, R. B., and Reilly, P. J. (2003). Crystal structure and evolution of a prokaryotic glucoamylase. *J. Mol. Biol.* 327, 61–73. doi: 10.1016/s0022-2836(03)00084-6

Anné, J., Vrancken, K., Van Mellaert, L., Van Impe, J., and Bernaerts, K. (2014). Protein secretion biotechnology in Gram-positive bacteria with special emphasis on *Streptomyces lividans*. *Biochim. Biophys. Acta Mol. Cell Res.* 1843, 1750–1761. doi: 10.1016/j.bbamcr.2013.12.023

Arai, T., Araki, R., Tanaka, A., Karita, S., Kimura, T., Sakka, K., et al. (2003). Characterization of a cellulase containing a family 30 carbohydrate-binding module (CBM) derived from *Clostridium thermocellum* CelJ: importance of the CBM to Cellulose Hydrolysis. *J. Bacteriol.* 185, 504–512. doi: 10.1128/jb.185.2.504-512.2003

Arantes, V., and Saddler, J. N. (2010). Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. *Biotechnol. Biofuels* 3:4. doi: 10.1186/1754-6834-3-4

Bayer, E. A., Belaich, J. P., Shoham, Y., and Lamed, R. (2004). The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annu. Rev. Microbiol.* 58, 521–554. doi: 10.1146/annurev.micro.57.030502.091022

Beckham, G. T., Matthews, J. F., Bomble, Y. J., Bu, L., Adney, W. S., Himmel, M. E., et al. (2010). Identification of amino acids responsible for processivity in a Family 1 carbohydrate-binding module from a fungal cellulase. *J. Phys. Chem. B* 114, 1447–1453. doi: 10.1021/jp908810a

Benoit, I., Culleton, H., Zhou, M., DiFalco, M., Aguilar-Osorio, G., Battaglia, E., et al. (2015). Closely related fungi employ diverse enzymatic strategies to degrade plant biomass. *Biotechnol. Biofuels* 8:107. doi: 10.1186/s13068-015-0285-0

Berlemont, R. (2017). Distribution and diversity of enzymes for polysaccharide degradation in fungi. *Sci. Rep.* 7:222. doi: 10.1038/s41598-017-00258-w

Berlemont, R., and Martiny, A. C. (2015). Genomic potential for polysaccharides deconstruction in bacteria. *Appl. Environ. Microbiol.* 81, 1513–1519. doi: 10.1128/AEM.03718-14

Berlemont, R., and Martiny, A. C. (2016). Glycoside hydrolases across environmental microbial communities. *PLoS Comput. Biol.* 12:e1005300. doi: 10.1371/journal.pcbi.1005300

Bernardes, A., Pellegrini, V. O. A., Curtolo, F., Camilo, C. M., Mello, B. L., Johns, M. A., et al. (2019). Carbohydrate binding modules enhance cellulose enzymatic hydrolysis by increasing access of cellulases to the substrate. *Carbohydr. Polym.* 211, 57–68. doi: 10.1016/j.carbpol.2019.01.108

Bissaro, B., Rohr, A. K., Müller, G., Chylenski, P., Skaugen, M., Forsberg, Z., et al. (2017). Oxidative cleavage of polysaccharides by monocopper enzymes depends on H2O2. *Nat. Chem. Biol.* 13, 1123–1128. doi: 10.1038/nchembio.2470

Book, A. J., Lewin, G. R., McDonald, B. R., Takasuka, T. E., Wendt-Pienkowski, E., Doering, D. T., et al. (2016). Evolution of high cellulolytic activity in symbiotic *Streptomyces* through selection of expanded gene content and coordinated gene expression. *PLoS Biol.* 14:e1002475. doi: 10.1371/journal.pbio.1002475

Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004). Carbohydrate binding modules: fine-tuning polysaccharide recognition. *Biochem. J.* 382, 769–781. doi: 10.1042/bj20040892

Boraston, A. B., Nurizzo, D., Notenboom, V., Ducros, V., Rose, D. R., Kilburn, D. G., et al. (2002). Differential oligosaccharide recognition by evolutionarily-related β-1,4 and β-1,3 glucan-binding modules. *J. Mol. Biol.* 319, 1143–1156. doi: 10.1016/s0022-2836(02)00374-1

Bott, R., Saldajeno, M., Cuevas, W., Ward, D., Scheffers, M., Aehle, W., et al. (2008). Three-dimensional structure of an intact glycoside hydrolase family 15 glucoamylase from *Hypocrea jecorina*. *Biochemistry* 47, 5746–5754. doi: 10.1021/bi702413k

Buleon, A., Colonna, P., Planchot, V., and Ball, S. (1998). Starch granules: structure and biosynthesis. *Int. J. Biol. Macromol.* 23, 85–112. doi: 10.1016/S0141-8130(98)00040-3

Burstein, T., Shulman, M., Jindou, S., Petkun, S., Frolow, F., Shoham, Y., et al. (2009). Physical association of the catalytic and helper modules of a family-9 glycoside hydrolase is essential for activity. *FEBS Lett.* 583, 879–884. doi: 10.1016/j.febslet.2009.02.013

Busarakam, K., Bull, A. T., Girard, G., Labeda, D. P., van Wezel, G. P., and Goodfellow, M. (2014). Streptomyces leeuwenhoekii sp. nov., the producer of chaxalactins and chaxamycins, forms a distinct branch in Streptomyces gene trees. *Antonie Van Leeuwenhoek* 105, 849–861. doi: 10.1007/s10482-014-0139-y

Cerda-Mejía, L., Valenzuela, S. V., Frías, C., Diaz, P., and Pastor, F. I. J. (2017). A bacterial GH6 cellobiohydrolase with a novel modular structure. *Appl. Microbiol. Biotechnol.* 101, 2943–2952. doi: 10.1007/s00253-017-8129-4

Chalak, A., Villares, A., Moreau, C., Haon, M., Grisel, S., d' Orlando, A., et al. (2019). Influence of the carbohydrate-binding module on the activity of a fungal AA9 lytic polysaccharide monooxygenase on cellulosic substrates. *Biotechnol. Biofuels* 12:206. doi: 10.1186/s13068-019-1548-y

Cheek, S., Krishna, S. S., and Grishin, N. V. (2006). Structural classification of small, disulfide-rich protein domains. *J. Mol. Biol.* 359, 215–237. doi: 10.1016/j.jmb.2006.03.017

Chen, S., Kaufman, M. G., Miazgowicz, K. L., Bagdasarian, M., and Walker, E. D. (2013). Molecular characterization of a cold-active recombinant xylanase from *Flavobacterium johnsoniae* and its applicability in xylan hydrolysis. *Bioresour. Technol.* 128, 145–155. doi: 10.1016/j.biortech.2012.10.087

Christensen, S. J., Badino, S. F., Cavaleiro, A. M., Borch, K., and Westh, P. (2020). Functional analysis of chimeric TrCel6A enzymes with different carbohydrate binding modules. *Protein Eng. Des. Select.* 32, 401–409. doi: 10.1093/protein/gzaa003

Cockburn, D. W., Suh, C., Medina, K. P., Duvall, R. M., Wawrzak, Z., Henrissat, B., et al. (2017). Novel carbohydrate binding modules in the surface anchored α-amylase of Eubacterium rectale provide a molecular rationale for the range of starches used by this organism in the human gut. *Mol. Microbiol.* 107, 249–264. doi: 10.1111/mmi.13881

Conway, J. M., Pierce, W. S., Le, J. H., Harper, G. W., Wright, J. H., Tucker, A. L., et al. (2016). Multidomain, surface layer-associated glycoside hydrolases contribute to plant polysaccharide degradation by *Caldicellulosiruptor* species. *J. Biol. Chem.* 291, 6732–6747. doi: 10.1074/jbc.m115.707810

Courtade, G., Forsberg, Z., Heggset, E. B., Eijsink, V. G. H., and Aachmann, F. L. (2018). The carbohydrate-binding module and linker of a modular lytic polysaccharide monooxygenase promote localized cellulose oxidation. *J. Biol. Chem.* 292, 13006–13015. doi: 10.1074/jbc.ra118.004269

Coutinho, J. B., Gilkes, N. R., Warren, R. A. J., Kilburn, D. G., and Miller, R. C. Jr. (1992). The binding of Cellulomonas fimi endoglucanase C (CenC) to cellulose and sephadex is mediated by the N-terminal repeats. *Mol. Microbiol.* 6, 1243–1252. doi: 10.1111/j.1365-2958.1992.tb01563.x

Couturier, M., Feliu, J., Haon, M., Navarro, D., Lesage-Meessen, L., Coutinho, P. M., et al. (2011). A thermostable GH45 endoglucanase from yeast: impact of its atypical multimodularity on activity. *Microb. Cell Fact.* 10:103. doi: 10.1186/1475-2859-10-103

Davies, G., and Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure* 3, 853–859. doi: 10.1016/s0969-2126(01)00220-9

de Souza, P. M., and de Oliveira Magalhães, P. (2010). Application of microbial α-amylase in industry - A review. *Braz. J. Microbiol.* 41, 850–861. doi: 10.1590/s1517-83822010000400004

Derde, L. J., Gomand, S. V., Courtin, C. M., and Delcour, J. A. (2012). Characterisation of three starch degrading enzymes: thermostable β-amylase, maltotetraogenic and maltogenic α-amylases. *Food Chem.* 135, 713–721. doi: 10.1016/j.foodchem.2012.05.031

Dias, F. M. V., Vincent, F., Pell, G., Prates, J. A. M., Centeno, M. S. J., Tailford, L. E., et al. (2004). Insights into the molecular determinants of substrate specificity in glycoside hydrolase family 5 revealed by the crystal structure and kinetics of *Cellvibrio mixtus* mannosidase 5A. *J. Biol. Chem.* 279, 25517–25526. doi: 10.1074/jbc.m401647200

Din, N., Gilkes, N., Tekant, B., Robert, C. M., Anthony, J. R. W., and Kilburn, D. G. (1991). Non–hydrolytic disruption of cellulose fibres by the binding domain of a bacterial cellulase. *Nat. Biotechnol.* 9, 1096–1099. doi: 10.1038/nbt1191-1096

Duan, C. J., Huang, M. Y., Pang, H., Zhao, J., Wu, C. X., and Feng, J. X. (2017). Characterization of a novel theme C glycoside hydrolase family 9 cellulase and its CBM-chimeric enzymes. *Appl. Microbiol. Biotechnol.* 101, 5723–5737. doi: 10.1007/s00253-017-8320-7

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Comput. Biol.* 7:e1002195. doi: 10.1371/journal.pcbi.1002195

El-Gebali, J., Mistry, A., Bateman, S. R., Eddy, A., Luciani, S. C., Potter, M., et al. (2019). Finn. The Pfam protein families database in 2019. *Nucleic Acids Res.* 47, D427–D432. doi: 10.1093/nar/gky995

Fiedler, M., Nitsche, B., Wanka, F., and Meyer, V. (2013). "*Aspergillus:* a cell factory with unlimited prospect," in *Applications of Microbial Engineering*, eds V. K. Gupta, M. Schmoll, M. A. Mazutti, M. Maki, and M. G. Tuohy (Boca Raton: CRC Press), 1–51. doi: 10.1201/b15250-2

Fleißner, A., and Dersch, P. (2010). Expression and export: recombinant protein production systems for *Aspergillus*. *Appl. Microbiol. Biotechnol.* 87, 1255–1270. doi: 10.1007/s00253-010-2672-6

Gilkes, N. R., Warren, R. A., Miller, R. J., and Kilburn, D. G. (1988). Precise excision of the cellulose binding domains from two *Cellulomonas fimi* cellulases by a homologous protease and the effect on catalysis. *J. Biol. Chem.* 263, 10401–10407.

Gilmore, S. P., Lillington, S. P., Haitjema, C. H., de Groot, R., and O'Malley, M. A. (2020). Designing chimeric enzymes inspired by fungal cellulosomes. *Synth. Syst. Biotechnol.* 5, 23–32. doi: 10.1016/j.synbio.2020.01.003

Gruben, B. S., Mäkelä, M. R., Kowalczyk, J. E., Zhou, M., Benoit-Gelber, I., and De Vries, R. (2017). Expression-based clustering of CAZyme-encoding genes of *Aspergillus niger*. *BMC Genomics* 18:900. doi: 10.1186/s12864-017-4164-x

Gudmundsson, M., Hansson, H., Karkehabadi, S., Larsson, A., Stals, I., Kim, S., et al. (2016). Structural and functional studies of the glycoside hydrolase family 3 β-glucosidase Cel3A from the moderately thermophilic fungus *Rasamsonia emersonii*. *Acta Crystallogr. D Struct. Biol.* 72(Pt 7), 860–870. doi: 10.1107/S2059798316008482

Gupta, R., Gigras, P., Mohapatra, H., Goswami, V. K., and Chauhan, B. (2003). Microbial α-amylases: a biotechnological perspective. *Process Biochem.* 38, 1599–1616. doi: 10.1016/S0032-9592(03)00053-0

Hamed, M. B., Anné, J., Karamanou, S., and Economou, A. (2018). *Streptomyces* protein secretion and its application in biotechnology. *FEMS Microbiol. Lett.* 365:fny250. doi: 10.1093/femsle/fny250

Han, Z., Shang-guan, F., and Yang, J. (2019). Molecular and biochemical characterization of a bimodular xylanase from *Marinifilaceae* bacterium strain SPP2. *Front. Microbiol.* 10:1507. doi: 10.3389/fmicb.2019.01507

Hansen, C. K. (1992). Fibronectin type III-like sequences and a new domain type in prokaryotic depolymerases with insoluble substrates. *FEBS Lett.* 305, 91–96. doi: 10.1016/0014-5793(92)80871-d

Hayashida, S., Kunisaki, S., Nakao, M., and Flor, P. Q. (1982). Evidence for raw starch -affinity site on *Aspergillus awamori* glucoamylase. *I. Agric. Biol. Chem.* 46, 83–89. doi: 10.1080/00021369.1982.10865024

Huang, X. F., Nazarian, F., Vincken, J. P., Visser, R. G. F., and Trindade, L. M. (2017). A tandem CBM25 domain of α-amylase from *Microbacterium aurum* as potential tool for targeting proteins to starch granules during starch biosynthesis. *BMC Biotechnol.* 17:86. doi: 10.1186/s12896-017-0406-x

Isikgor, F. H., and Becer, C. R. (2015). Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. *Polym. Chem.* 6:4497. doi: 10.1039/C5PY00263J

Janeček, Š, Mareček, F., MacGregor, E. A., and Svensson, B. (2019). Starch-binding domains as CBM families–history, occurrence, structure, function and evolution. *Biotechnol. Adv.* 37:107451. doi: 10.1016/j.biotechadv.2019.107451

Janeček, Š, Svensson, B., and MacGregor, E. A. (2014). α -Amylase – an enzyme specificity found in various families of glycoside hydrolases. *Cell. Mol. Life Sci.* 71, 1149–1170. doi: 10.1007/s00018-013-1388-z

Jee, J. G., Ikegami, T., Hashimoto, M., Kawabata, T., Ikeguchi, M., Watanabe, T., et al. (2002). Solution structure of the fibronectin type III domain from *Bacillus circulans* WL-12 chitinase A1. *J. Biol. Chem.* 277, 1388–1397. doi: 10.1074/jbc.M109726200

Jeoh, T., Wilson, D. B., and Walker, L. P. (2006). Effect of cellulase mole fraction and cellulose recalcitrance on synergism in cellulose hydrolysis and binding. *Biotechnol. Prog.* 22, 270–277. doi: 10.1021/bp050266f

Käll, L., Krogh, A., and Sonnhammer, E. L. L. (2004). A combined transmembrane topology and signal peptide prediction method. *J. Mol. Biol.* 338, 1027–1036. doi: 10.1016/j.jmb.2004.03.016

Kataeva, I. A., Seidel, R. D., Shah, A., West, L. T., Li, X. L., and Ljungdahl, L. G. (2002). The fibronectin type 3-Like repeat from the *Clostridium thermocellum* cellobiohydrolase CbhA promotes hydrolysis of cellulose by modifying its surface. *Appl. Environ. Microbiol.* 68, 4292–4300. doi: 10.1128/aem.68.9.4292-4300.2002

Kim, D. Y., Han, M. K., Park, D. S., Lee, J. S., Oh, H. W., Shin, D. H., et al. (2009). Novel GH10 xylanase, with a fibronectin type 3 domain, from *Cellulosimicrobium* sp. Strain HY-13, a bacterium in the gut of *Eisenia fetida*. *Appl. Environ. Microbiol.* 75, 7275–7279. doi: 10.1128/aem.01075-09

Koide, A., Bailey, C. W., Huang, X., and Koide, S. (1998). The fibronectin type III domain as a scaffold for novel binding proteins. *J. Mol. Biol.* 284, 1141–1151. doi: 10.1006/jmbi.1998.2238

Kosugi, A., Amano, Y., Murashima, K., and Doi, R. H. (2004). Hydrophilic domains of scaffolding protein CbpA promote glycosyl hydrolase activity and localization of cellulosomes to the cell surface of *Clostridium cellulovorans*. *J. Bacteriol.* 186, 6351–6359. doi: 10.1128/JB.186.19.6351-6359.2004

Kumagai, Y., Yamashita, K., Tagami, T., Uraji, M., Wan, K., Okuyama, M., Yao, M., Kimura, A, Hatanaka, T. (2015). The loop structure of *Actinomycete* glycoside hydrolase family 5 mannanases governs substrate recognition. *FEBS J.* 282, 4001–4014. doi: 10.1111/febs.13401

Lee, J., and Paetzel, M. (2011). Structure of the catalytic domain of glucoamylase from *Aspergillus niger*. *Acta Crystallogr.* 67, 188–192. doi: 10.1107/s1744309110049390

Leggio, L. L., Simmons, T. J., Poulsen, J. C. N., Frandsen, K. E. H., Hemsworth, G. R., Stringer, M. A., et al. (2015). Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase. *Nat. Commun.* 6. doi: 10.1038/ncomms6961

Liberato, M. V., Silveira, R. L., Prates, E. T., De Araujo, E. A., Pellegrini, V. O. A., Camilo, C. M., et al. (2016). Molecular characterization of a family 5 glycoside hydrolase suggests an induced-fit enzymatic mechanism. *Sci. Rep.* 6, 1–16. doi: 10.1038/srep23473

Lima, M. A., Oliveira-Neto, M., Kadowaki, M. A., Rosseto, F. R., Prates, E. T., Squina, F. M., et al. (2013). *Aspergillus niger* β-glucosidase has a cellulase-like tadpole molecular shape: insights into glycoside hydrolase family 3 (GH3) β-glucosidase structure and function. *J. Biol. Chem.* 288, 32991–33005. doi: 10.1074/jbc.M113.479279

Lin, F. P., Ma, H. Y., Lin, H. J., Liu, S. M., and Tzou, W. S. (2011). Biochemical characterization of two truncated forms of amylopullulanase from *Thermoanaerobacterium saccharolyticum* NTOU1 to identify its enzymatically active region. *Appl. Biochem. Biotechnol.* 165, 1047–1056. doi: 10.1007/s12010-011-9319-7

Liu, S., and Ding, S. (2016). Replacement of carbohydrate binding modules improves acetyl xylan esterase activity and its synergistic hydrolysis of different substrates with xylanase. *BMC Biotechnol.* 16:73. doi: 10.1186/s12896-016-0305-6

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495. doi: 10.1093/nar/gkt1178

Mattinen, M. L., Linder, M., Teleman, A., and Annila, A. (1997). Interaction between cellohexaose and cellulose binding domains from *Trichoderma reesei* cellulases. *FEBS Lett.* 407, 291–296. doi: 10.1016/s0014-5793(97)00356-6

Mehta, D., and Satyanarayana, T. (2016). Bacterial and archaeal α-amylases: diversity and amelioration of the desirable characteristics for industrial applications. *Front. Microbiol.* 7:1129. doi: 10.3389/fmicb.2016.01129

Meinke, A., Damude, H. G., Tomme, P., Kwan, E., Kilburn, D. G., Miller, R. C., et al. (1995). Enhancement of the endo-β-1, 4-glucanase activity of an exocellobiohydrolase by deletion of a surface loop. *J. Biol. Chem.* 270, 4383–4386. doi: 10.1074/jbc.270.9.4383

Mikami, B., Degano, M., Hehre, E. J., and Sacchettini, J. C. (1994). Crystal structures of soybean .beta.-amylase reacted with .beta.-maltose and maltal: active site components and their apparent roles in catalysis. *Biochemistry* 33, 7779–7787. doi: 10.1021/bi00191a005

Mikami, B., Iwamoto, H., Malle, D., Yoon, H.-J., Demirkan-Sarikaya, E., Mezaki, Y., et al. (2006). Crystal structure of pullulanase: evidence for parallel binding of oligosaccharides in the active site. *J. Mol. Biol.* 359, 690–707. doi: 10.1016/j.jmb.2006.03.058

Mizuno, M., Tonozuka, T., Suzuki, S., Uotsu-Tomita, R., Kamitori, S., Nishikawa, A., et al. (2004). Structural insights into substrate specificity and function of glucodextranase. *J. Biol. Chem.* 279, 10575–10583. doi: 10.1074/jbc.M310771200

Montanier, C., Flint, J. E., Bolam, D. N., Xie, H., Liu, Z., Rogowski, A., et al. (2010). Circular permutation provides an evolutionary link between two families of calcium-dependent carbohydrate binding modules. *J. Biol. Chem.* 285, 31742–31754. doi: 10.1074/jbc.m110.142133

Montella, S., Ventorino, V., Lombard, V., Henrissat, B., Pepe, O., and Faraco, V. (2017). Discovery of genes coding for carbohydrate-active enzyme by metagenomic analysis of lignocellulosic biomasses. *Sci. Rep.* 7:42623. doi: 10.1038/srep42623

Mosbah, A., Belaich, A., Bornet, O., Belaich, J. P., Henrissat, B., and Darbon, H. (2000). Solution structure of the module X2 1 of unknown function of the cellulosomal scaffolding protein CipC of *Clostridium cellulolyticum*. *J. Mol. Biol.* 304, 201–217. doi: 10.1006/jmbi.2000.4192

Nguyen, S. T. C., Freund, H. L., Kasanjian, J., and Berlemont, R. (2018). Function, distribution, and annotation of characterized cellulases, xylanases, and chitinases from CAZy. *Appl. Microbiol. Biotechnol.* 102, 1629–1637. doi: 10.1007/s00253-018-8778-y

Niu, C., Zheng, F., Li, Y., Liu, C., and Li, Q. (2018). Process optimization of the extraction condition of β-amylase from brewer's malt and its application in the maltose syrup production. *Biotechnol. Appl. Biochem.* 65, 639–647. doi: 10.1002/bab.1650

Oliveira, C., Carvalho, V., Domingues, L., and Gama, F. M. (2015). Recombinant CBM-fusion technology. Appl.ications overview. *Biotechnol. Adv.* 33, 358–369. doi: 10.1016/j.biotechadv.2015.02.006

Oyama, T., Kusunoki, M., Kishimoto, Y., Takasaki, Y., and Nitta, Y. (1999). Crystal structure of β-amylase from Bacillus cereus var. mycoides at 2.2 Å resolution. *J. Biochem.* 125, 1120–1130. doi: 10.1093/oxfordjournals.jbchem.a022394

Oyama, T., Miyake, H., Kusunoki, M., and Nitta, Y. (2003). Crystal structures of beta-amylase from *Bacillus cereus* var mycoides in complexes with substrate analogs and affinity-labeling reagents. *J. Biochem.* 133, 467–474. doi: 10.1093/jb/mvg061

Parkkinen, T., Koivula, A., Vehmaanperä, J., and Rouvinen, J. (2008). Crystal structures of *Melanocarpus albomyces* cellobiohydrolase Cel7B in complex with cello-oligomers show high flexibility in the substrate binding. *Protein Sci.* 17, 1383–1394. doi: 10.1110/ps.034488.108

Pasari, N., Adlakha, N., Gupta, M., Bashir, Z., Rajacharya, G. H., Verma, G., et al. (2017). Impact of module-X2 and carbohydrate binding module-3 on the catalytic activity of associated glycoside hydrolases towards plant biomass. *Sci. Rep.* 7:3700. doi: 10.1038/s41598-017-03927-y

Payne, C. M., Knott, B. C., Mayes, H. B., Hansson, H., Himmel, M. E., Sandgren, M., et al. (2015). Fungal cellulases. *Chem. Rev.* 115, 1308–1448. doi: 10.1021/cr500351c

Pérez, J., Muñoz-Dorado, J., de la Rubia, T., and Martínez, J. (2002). Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview. *Int. Microbiol.* 5, 53–63. doi: 10.1007/s10123-002-0062-3

Poidevin, L., Feliu, J., Doan, A., Berrin, J. G., Bey, M., Coutinho, P. M., et al. (2013). Insights into exo- and endoglucanase activities of family 6 glycoside hydrolases from *Podospora anserina*. *Appl. Environ. Microbiol.* 79, 4220–4229. doi: 10.1128/aem.00327-13

Pointing, S. B. (1999). Qualitative methods for the determination of lignocellulolytic enzyme production by tropical fungi. *Fungal Divers.* 2, 17–33.

Polizeli, M. L. T. M., Vici, A. C., Scarcella, A. S. A., Cereia, M., and Pereira, M. G. (2016). "Enzyme system from *Aspergillus* in current industrial uses and future applications in the production of second-generation ethanol," in *New and Future Developments in Microbial Biotechnology and Bioengineering*, ed. V. K. Gupta (Amsterdam: Elsevier), 127–140. doi: 10.1016/b978-0-444-63505-1.00009-9

Pujadas, G., and Palau, J. (2001). Evolution of α-amylases: architectural features and key residues in the stabilization of the (β/α)8 scaffold. *Mol. Biol. Evol.* 18, 38–54. doi: 10.1093/oxfordjournals.molbev.a003718

Punt, P. J., Levasseur, A., Visser, H., Wery, J., and Record, E. (2011). Fungal protein production: design and production of chimeric proteins. *Annu. Rev. Microbiol.* 65, 57–69. doi: 10.1146/annurev.micro.112408.134009

Punt, P. J., van Biezen, N., Conesa, A., Albers, A., Mangnus, J., and van den Hondel, C. (2002). Filamentous fungi as cell factories for heterologous protein production. *Trends Biotechnol.* 20, 200–206. doi: 10.1016/s0167-7799(02)01933-9

Rahikainen, J., Ceccherini, S., Molinier, M., Holopainen-Mantila, U., Reza, M., Väisänen, S., et al. (2019). Effect of cellulase family and structure on modification of wood fibres at high consistency. *Cellulose* 26, 5085–5103. doi: 10.1007/s10570-019-02424-x

Ray, R. R. (2004). β-amylases from various fungal strains. *Acta Microbiol. Immunol. Hung.* 51, 85–95. doi: 10.1556/amicr.51.2004.1-2.6

Ray, R. R., and Nanda, G. (1996). Microbial β-amylases: biosynthesis, characteristics, and industrial applications. *Crit. Rev. Microbiol* 22, 181–199. doi: 10.3109/10408419609106459

Reyes-Ortiz, V., Heins, R. A., Cheng, G., Kim, E. Y., Vernon, B. C., Elandt, R. B., et al. (2013). Addition of a carbohydrate-binding module enhances cellulase penetration into cellulose substrates. *Biotechnol. Biofuels* 6:93. doi: 10.1186/1754-6834-6-93

Roth, C., Moroz, O. V., Ariza, A., Skov, L. K., Ayabe, K., Davies, G. J., et al. (2018). Structural insight into industrially relevant glucoamylases: flexible positions of starch -binding domains. *Acta Crystallogr. D Struct. Biol.* 74, 463–470. doi: 10.1107/S2059798318004989

Saadat, F. (2017). A review on chimeric xylanases: methods and conditions. *3 Biotech* 7:67. doi: 10.1007/s13205-017-0660-6

Sauer, J., Christensen, T., Frandsen, T. P., Mirgorodskaya, E., McGuire, K. A., Driguez, H., et al. (2001). Stability and Function of Interdomain Linker Variants of Glucoamylase 1 from *Aspergillus niger*. *Biochemistry* 40, 9336–9346. doi: 10.1021/bi010515i

Sauer, J., Sigurskjold, B. W., Christensen, U., Frandsen, T. P., Mirgorodskaya, E., Harrison, M., et al. (2000). Glucoamylase: structure/function relationships, and protein engineering. *Biochim. Biophys. Acta Protein Struct. Mol. Enzymol.* 1543, 275–293. doi: 10.1016/s0167-4838(00)00232-6

Segato, F., Damasio, A. R., Gonçalves, T. A., Murakami, M. T., Squina, F. M., Polizeli, M., et al. (2012). Two structurally discrete GH7-cellobiohydrolases compete for the same cellulosic substrate fiber. *Biotechnol. Biofuels* 5:21. doi: 10.1186/1754-6834-5-21

Sevillano, L., Vijgenboom, E., van Wezel, G. P., Díaz, M., and Santamaría, R. I. (2016). New approaches to achieve high level enzyme production in *Streptomyces lividans*. *Microb. Cell Fact.* 15:28. doi: 10.1186/s12934-016-0425-7

Shimon, L. J. W., Pagès, S., Belaich, A., Belaich, J.-P., Bayer, E. A., Lamed, R., et al. (2000). Structure of a family IIIa scaffoldin CBD from the cellulosome of *Clostridium cellulolyticum* at 2.2 Å resolution. *Acta Crystallogr. Sec. D Biol. Crystallogr.* 56, 1560–1568. doi: 10.1107/s0907444900012889

Simpson, P. J., Jamieson, S. J., Abou-Hachem, M., Karlsson, E. N., Gilbert, H. J., Holst, O., et al. (2002). The solution structure of the CBM4-2 carbohydrate binding module from a thermostable *Rhodothermus marinus* xylanase. *Biochemistry* 41, 5712–5719. doi: 10.1021/bi012093i

Somerville, C. R. (2006). Cellulose synthesis in higher plants. *Annu. Rev. Cell Dev. Biol.* 22, 53–78. doi: 10.1146/annurev.cellbio.22.022206.160206

Suzuki, K., Sumitani, J.-I., Nam, Y.-W., Nishimaki, T., Tani, S., Wakagi, T., et al. (2013). Crystal structures of glycoside hydrolase family 3 β-glucosidase 1 from *Aspergillus aculeatus*. *Biochem. J.* 452, 211–221. doi: 10.1042/bj20130054

Svensson, B., Pedersen, T. G., Svendsen, I., Sakai, T., and Ottesen, M. (1982). Characterization of two forms of glucoamylase from *Aspergillus niger*. *Carlsberg Res. Commun.* 47, 55–69. doi: 10.1007/BF02907797

Talamantes, D., Biabini, N., Dang, H., Abdoun, K., and Berlemont, R. (2016). Natural diversity of cellulases, xylanases, and chitinases in bacteria. *Biotechnol. Biofuels* 9:133. doi: 10.1186/s13068-016-0538-6

Teo, S. C., Liew, K. J., Shamsir, M. S., Chong, C. S., Bruce, N. C., Chan, K.-G., et al. (2019). Characterizing a halo-tolerant GH10 Xylanase from *Roseithermus sacchariphilus* strain RA and its CBM-truncated variant. *Int. J. Mol. Sci.* 20:2284. doi: 10.3390/ijms20092284

Thalmann, M., Coiro, M., Meier, T., Wicker, T., Zeeman, S. C., and Santelia, D. (2019). The evolution of functional complexity within the β-amylase gene family in land plants. *BMC Evol. Biol.* 19:66. doi: 10.1186/s12862-019-1395-2

Tomme, P., Van Tilbeurgh, H., Pettersson, G., Van Damme, J., Vandekerckhove, J., Knowles, J., et al. (1988). Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. *Eur. J. Biochem.* 170, 575–581. doi: 10.1111/j.1432-1033.1988.tb13736.x

Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y., et al. (1996). Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose. *EMBO J.* 15, 5739–5751. doi: 10.1002/j.1460-2075.1996.tb00960.x

Tóth, Á, Baka, E., Luzics, S., Bata-Vidács, I., Nagy, I., Bálint, B., et al. (2017). Plant polysaccharide degrading enzyme system of *Thermobifida cellulosilytica* TB100T revealed by de novo genome project data. *Acta Alimentaria* 46, 323–335. doi: 10.1556/066.2016.0014

Tseng, C. W., Tzu, P. K., Rey, T. G., Jian, W. H., Hao, C. W., Chun, H. H., et al. (2011). Substrate binding of a GH5 endoglucanase from the ruminal fungus *Piromyces rhizinflata*. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* 67, 1189–1194. doi: 10.1107/S1744309111032428

Valk, V., Eeuwema, W., Sarian, F. D., van der Kaaij, R. M., and Dijkhuizen, L. (2015). Degradation of granular starch by the bacterium *Microbacterium aurum* strain B8.A involves a modular α-amylase enzyme system with FNIII and CBM25 domains. *Appl. Environ. Microbiol.* 81, 6610–6620. doi: 10.1128/AEM.01029-15

Van Bueren, A. L., and Boraston, A. B. (2007). The structural basis of α-glucan recognition by a family 41 carbohydrate-binding module from *Thermotoga maritima*. *J. Mol. Biol.* 365, 555–560. doi: 10.1016/j.jmb.2006.10.018

Van Tilbeurgh, H., Tomme, P., Claeyssens, M., Bhikhabhai, R., and Pettersson, G. (1986). Limited proteolysis of the cellobiohydrolase I from *Trichoderma reesei*. *FEBS Lett.* 204, 223–227. doi: 10.1016/0014-5793(86)80816-X

Venditto, I., Najmudin, S., Luís, A. S., Ferreira, L. M. A., Sakka, K., Knox, J. P., et al. (2015). Family 46 carbohydrate-binding modules contribute to the enzymatic hydrolysis of xyloglucan and β-1,3–1,4-glucans through distinct mechanisms. *J. Biol. Chem.* 290, 10572–10586. doi: 10.1074/jbc.m115.637827

Voshol, G. P., Punt, P. J., and Vijgenboom, E. (2019). Profile Comparer Extended: phylogeny of lytic polysaccharide monooxygenase families using profile hidden Markov model alignments. *F1000Research* 8:1834. doi: 10.12688/f1000research.21104.1

Voshol, G. P., Vijgenboom, E., and Punt, P. J. (2017). The discovery of novel LPMO families with a new hidden markov model. *BMC Res. Notes* 10:105. doi: 10.1186/s13104-017-2429-8

Voutilainen, S. P., Nurmi-Rantala, S., Penttilä, M., and Koivula, A. (2014). Engineering chimeric thermostable GH7 cellobiohydrolases in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* 98, 2991–3001. doi: 10.1007/s00253-013-5177-2

Vu, V. V., Beeson, W. T., Span, E. A., Farquhar, E. R., and Marletta, M. A. (2014). A family of starch -active polysaccharide monooxygenases. *Proc. Natl. Acad. Sci. U.S.A.* 111, 13822–13827. doi: 10.1073/pnas.1408090111

Yaniv, O., Petkun, S., Shimon, L. J. W., Bayer, E. A., Lamed, R., and Frolow, F. (2012). A single mutation reforms the binding activity of an adhesion-deficient family 3 carbohydrate-binding module. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 68, 819–828. doi: 10.1107/s0907444912013133

Yuan, X. L., van der Kaaij, R. M., van den Hondel, C. A. M. J. J., Punt, P. J., van der Maarel, M. J. E. C., Dijkhuizen, L., et al. (2008). *Aspergillus niger*, genome-wide analysis reveals a large number of novel alpha-glucan acting enzymes with unexpected expression profiles. *Mol. Genet. Genomics* 279, 545–561. doi: 10.1007/s00438-008-0332-7

Zheng, F., and Ding, S. (2012). Processivity and enzymatic mode of a glycoside hydrolase family 5 endoglucanase from *Volvariella volvacea*. *Appl. Environ. Microbiol.* 79, 989–996. doi: 10.1128/aem.02725-12

Zhou, W., Irwin, D. C., Escovar-Kousen, J., and Wilson, D. B. (2004). Kinetic studies of thermobifida fuscaCel9A active site mutant enzymes. *Biochemistry* 43, 9655–9663. doi: 10.1021/bi049394n

Zhu, Y., Romain, C., and Williams, C. K. (2016). Sustainable polymers from renewable resources. *Nature* 540, 354–362. doi: 10.1038/nature21001

Zverlov, V. V., Volkov, I. Y., Velikodvorskaya, G. A., and Schwarz, W. H. (2001). The binding pattern of two carbohydrate-binding modules of laminarinase Lam16A from *Thermotoga neapolitana*: differences in beta-glucan binding within family CBM4. *Microbiology* 147, 621–629. doi: 10.1099/00221287-147-3-621

Conflict of Interest: EA was employed by company Sun Pharmaceutical Industries Europe BV. GV and PP were employed by company Dutch DNA Biotech B.V.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 25 August 2020
doi: 10.3389/fbioe.2020.01028

Evaluation of the Enzymatic Arsenal Secreted by *Myceliophthora thermophila* During Growth on Sugarcane Bagasse With a Focus on LPMOs

*Maria Angela B. Grieco[1,2], Mireille Haon[2], Sacha Grisel[2], Ana Lucia de Oliveira-Carvalho[3], Augusto Vieira Magalhães[3], Russolina B. Zingali[3], Nei Pereira Jr.[1] and Jean-Guy Berrin[2]**

[1] Laboratório de Desenvolvimento de Bioprocessos, Departamento de Engenharia Bioquímica, Escola de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil, [2] INRAE, Faculté des Sciences de Luminy, Aix Marseille Université, UMR 1163 Biodiversité et Biotechnologie Fongiques, Polytech Marseille, Marseille, France, [3] Unidade de Espectrometria de Massas e Proteômica, Instituto de Bioquímica Médica Leopoldo de Meis, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

OPEN ACCESS

Edited by:
Rasmus John Normand Frandsen, Technical University of Denmark, Denmark

Reviewed by:
Sunil Khare, Indian Institute of Technology Delhi, India
Jiaxing Xu, Huaiyin Normal University, China

***Correspondence:**
Jean-Guy Berrin
jean-guy.berrin@inrae.fr;
jean-guy.berrin@univ-amu.fr

Specialty section:
This article was submitted to Bioprocess Engineering, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: 20 May 2020
Accepted: 06 August 2020
Published: 25 August 2020

Citation:
Grieco MAB, Haon M, Grisel S, de Oliveira-Carvalho AL, Magalhães AV, Zingali RB, Pereira N Jr and Berrin J-G (2020) Evaluation of the Enzymatic Arsenal Secreted by Myceliophthora thermophila During Growth on Sugarcane Bagasse With a Focus on LPMOs. Front. Bioeng. Biotechnol. 8:1028. doi: 10.3389/fbioe.2020.01028

The high demand for energy and the increase of the greenhouse effect propel the necessity to develop new technologies to efficiently deconstruct the lignocellulosic materials into sugars monomers. Sugarcane bagasse is a rich polysaccharide residue from sugar and alcohol industries. The thermophilic fungus *Myceliophthora thermophila* (syn. *Sporotrichum thermophilum*) is an interesting model to study the enzymatic degradation of biomass. The genome of *M. thermophila* encodes an extensive repertoire of cellulolytic enzymes including 23 lytic polysaccharide monooxygenases (LPMOs) from the Auxiliary Activity family 9 (AA9), which are known to oxidatively cleave the β-1,4 bonds and boost the cellulose conversion in a biorefinery context. To achieve a deeper understanding of the enzymatic capabilities of *M. thermophila* on sugarcane bagasse, we pretreated this lignocellulosic residue with different methods leading to solids with various cellulose/hemicellulose/lignin proportions and grew *M. thermophila* on these substrates. The secreted proteins were analyzed using proteomics taking advantage of two mass spectrometry methodologies. This approach unraveled the secretion of many CAZymes belonging to the Glycosyl Hydrolase (GH) and AA classes including several LPMOs that may contribute to the biomass degradation observed during fungal growth. Two AA9 LPMOs, called *Mt*LPMO9B and *Mt*LPMO9H, were selected from secretomic data and enzymatically characterized. Although *Mt*LPMO9B and *Mt*LPMO9H were both active on cellulose, they differed in terms of optimum temperatures and regioselectivity releasing either C1 or C1-C4 oxidized oligosaccharides, respectively. LPMO activities were also measured on sugarcane bagasse substrates with different levels of complexity. The boosting effect of these LPMOs on bagasse sugarcane saccharification

by a *Trichoderma reesei* commercial cocktail was also observed. The partially delignified bagasse was the best substrate considering the oxidized oligosaccharides released and the acid treated bagasse was the best one in terms of saccharification boost.

Keywords: filamentous fungi, enzyme, LPMO, biomass, biofuels, biorefinery

INTRODUCTION

The substitution of the fossil fuels is an important strategy to decrease the environmental impacts caused by the high levels of CO_2 emission and greenhouse gases. The development of renewable energies is a promising strategy to reduce global warming damages and climate changes (Wyman, 2003; Farrell et al., 2006). Sugarcane is a well-established feedstock to produce first-generation ethanol (1G ethanol) (Smithers, 2014; Clauser et al., 2016). In 2019, more than 700 million tons were produced in the Brazilian territory. Although sugarcane as feedstock to 1G ethanol production is a successful model of the sugar-alcohol industry, its coproduct, sugarcane bagasse is still a challenging feedstock for the production of fuels and chemicals due to its recalcitrance to degradation (Eggert and Greaker, 2014; Khattab and Watanabe, 2019).

In nature, the most powerful natural decomposers of lignocellulosic biomass are filamentous fungi (wood-decaying fungi). They secrete a powerful group of enzymes called CAZymes[1] (Lombard et al., 2014) that act synergistically to cleave the β1,4-β1,3 linkages of plant polysaccharides releasing as end products a range of monosaccharides that provide energy to the microorganisms intermingled in their environmental. *In vitro*, pretreatments of lignocellulosic biomass are required to render the substrate more accessible to the enzymes. Overall, additional cost reduction is desirable to achieve competitiveness of the second-generation (2G) ethanol production (Loqué et al., 2015).

Ten years ago, the discovery of a new class of copper-enzymes, the lytic polysaccharide monooxygenase (LPMO, EC 1.14.99.53-56), cleaving with an oxidative mechanism recalcitrant polysaccharides, revolutionized the concept of plant cell wall breakdown (Vaaje-Kolstad et al., 2010; Quinlan et al., 2011). These mono-copper enzymes are classified in Auxiliary Activity families AA9-AA11 and AA13-AA16 in the CAZy database and contain a characteristic histidine brace copper binding site. Due to their impressive boosting effect on the cellulose breakdown, LPMOs have rapidly become key components of commercial enzymatic cocktails (Johansen, 2016). Filamentous fungi, especially fungal saprotrophs, which degrade complex and recalcitrant plant polymers, are proficient secretors of LPMOs and their redox partners (Berrin et al., 2017). They secrete LPMOs that differ in substrate specificities, AA9 members target cellulose and hemicelluloses (Tandrup et al., 2018), AA11 LPMOs cleave chitin (Hemsworth et al., 2014), AA13 target starch (Vu et al., 2014b; Lo Leggio et al., 2015), AA14 recalcitrant xylans (Couturier et al., 2018) and AA16 cellulose (Filiatrault-Chastel et al., 2019). AA9 LPMOs occur either as isolated modules or appended to cellulose-binding modules (CBM1) at the C-terminal in agreement with the critical role of the N-terminal histidine for enzyme activity. The remarkable expansion of the AA9 family (more than 30 candidate proteins in some genomes) raises the question of its functional relevance at the organismal level. While all AA9 LPMOs characterized to date cleave cellulose, subtle differences have been observed in their regioselectivities acting either on the C1, C4 or C1 and C4 carbon of the glucose unit.

Myceliophthora thermophila syn. *Sporotrichum thermophilum* is a ubiquitous thermophilic fungus with a strong ability to hydrolyze all major polysaccharides found in biomass due to the secretion of many CAZymes (cellulases, xylanases, pectinases and some other miscellaneous enzymes) employed in various biotechnological applications (Karnaouri et al., 2014; Singh, 2016). Characterization of the biomass-hydrolyzing activity of wild and recombinant enzymes suggests that this mold is highly efficient in biomass decomposition at both moderate and high temperatures. The genome of *M. thermophila* encodes 23 AA9 LPMOs (Berka et al., 2011) of which nine have been characterized relative to their substrate specificity and/or regioselectivity (Vu et al., 2014a; Frommhagen et al., 2015, 2016, 2018; Karnaouri et al., 2017; Hangasky et al., 2018).

In this study, we investigated the enzymatic arsenal secreted by *M. thermophila* upon growth on sugarcane bagasse using proteomics. A large range of CAZymes targeting lignocellulose was observed in the secretomes and due to the importance of LPMOs, the activity of the two AA9 LPMOs was studied in different conditions to depict their contribution to sugarcane bagasse degradation.

MATERIALS AND METHODS

Sugarcane Bagasse Substrates

The sugarcane bagasse (*Saccharum* spp.) was obtained from Boa Vista sugar-alcohol industry (Goiás, Br) and stored at room temperature. To use as substrate, the sugarcane bagasse not treated (SCBNT) was washed several times in water, dried at 45°C in an air oven for 3 days and milled (willed mill at Tecnal, TE-680, BR) in a sieve with 5 mm particle size. After, 1 kg was subjected to acid treatment using a 1.09% (v/v) H_2SO_4 solution with a solid-liquid proportion of 1:2.8 (w/v) for 30 min at 121°C (Betancur and Pereira, 2010). The acid pretreatment provided a liquid comprised mostly by hemicelluloses chemically hydrolyzed and a solid fraction composed mainly by cellulose and lignin (Khattab and Watanabe, 2019) that was separated by a pressing filtration at 10 kgf.cm^{-2}. The solid phase was subsequently washed exhaustively until pH reached 5.0 and dried at 45°C in air oven for 3 days. The resulting feedstock is called the sugarcane bagasse acid treated (SCBAT). Subsequently, half of the SCBAT

[1]http://www.cazy.org

was submitted to alkaline pre-treatment conducted using sodium hydroxide 4% (w/v) at 121°C for 30 min with a solid-liquid-ratio of 1:20 (Vasquez et al., 2007). The acid and alkaline treated bagasse called sugarcane bagasse partially delignified (SCBPD) was washed several times until pH reached 5.0 and dried in an air oven at 45°C for 3 days.

Sugarcane Bagasse Composition

The determination of the structural carbohydrates (cellulose and hemicellulose), lignin and ashes in SCBNT, SCBAT and SCBPD were derived through chemical hydrolysis with 72% (v/w) H_2SO_4 as described (Sluiter et al., 2008; **Supplementary Table S1**). The sugars released were measured by glucose oxidase assay (GOD) and 3–5, Dinitrosalycilic acid method (DNS) (Miller, 1959) using standard solution of glucose with the highest grade of purification to establish the sugars concentration. The hemicellulose portion was calculated subtracting DNS and glucose oxidase assay (GOD) values. All the analyses were carried out twice in triplicate on dried samples.

Microorganism and Cultivation

The *Myceliophthora thermophila* (M7.7 equivalent to FGSC 26436) strain used in this study was maintained in glycerol 20% (v/v) in microtubes at 4°C. The suspension was inoculated onto fresh potato dextrose agar medium (PDA) plates and incubated at 50°C to grow for 4 days. Liquid culture media enriched with 15 g L^{-1} dry matter (d.m.) of either SCBAT or SCBPD were prepared as follow: 0.2 g L^{-1} yeast extract; 1.0 g L^{-1} peptone; 0.3 g L^{-1} urea; 0.4 g L^{-1} $CaCl_2.2H_2O$; 0.2 g L^{-1} $MgSO_4.7H_2O$; 2.0 g L^{-1} KH_2PO_4; and microelements solutions formed by 5.0 mg L^{-1} $FeSO_4.7H_2O$; 1.6 mg L^{-1} $MnSO_4.4H_2O$; 1.4 mg L^{-1} $ZnSO_4.7H_2O$, and 2.0 mg L^{-1} $CuSO_4.6H_2O$ (modified from Mandels and Weber, 1969) and autoclaved. The ascospores were scraped from five plates with an inoculation loop and transferred to a 50 mL falcon tube containing 10 mL sterilized liquid culture medium. The spores were enumerated with a Malassez cell counting chamber under a microscope. Three culture media (pH 5.3) were inoculated with 106 spores mL^{-1}. The fungal growth was carried out in 1 L flask with 100 mL of medium at 45°C for 2 days in an orbital shaker (New Brunswick Excella, São Paulo, Brazil). All experiments were performed in duplicate.

Extract Production for Proteomic Analysis

The solid and liquid fractions were separated by a vacuum filtration using a glass wool (Merck, Rio de Janeiro, Brazil). The corresponding culture broths (secretomes) were first filtered using a 0.22 mm cut-off column and then concentrated using a 5 kDa cut-off column, both in a hollow fiber system (QuixStand, GE, Healthcare, São Paulo, Brazil). The resulted protein extract was further concentrated by precipitation or lyophilization. For protein precipitation the same volume of TCA/Acetone 20% v/v were added to 3 mL of samples and led at −20°C overnight. Supernatant was discarded after centrifugation at 12,000 *g* for 30 min and the precipitate was washed three times with 3 mL of cold acetone and centrifugated as previously described. Protein pellet were resuspended in 200 μL of sample buffer for electrophoresis (2% SDS and 0.5 M tris-HCl 10% glycerol) and 15 μL of each sample were loaded onto a 12.5% SDS-PAGE. In parallel, 15 mL of the corresponding protein concentrated were lyophilized (Freeze Dryer, LS6000, Terroni) resuspended in 1.0 mL citrate buffer pH 5.0, filtered in 0.22 μm pore size membrane (Syringe-driven Filters, Biofilm). After filtration, 20 μg of each sample were loaded onto a preparative SDS-PAGE gel 12% which was halted after 1 cm migration into the separation gel in order to obtain a single gel slice for each sample.

Proteomic Analysis

Gel slices, from analytical and preparative SDS-PAGE were submitted to a tryptic digestion after a treatment with 5 mM DTT (dithiothreitol), followed by 15 mM iodoacetamide and trypsin in a 1:50 mass ratio to protein. The extracted peptides from the gel spots were analyzed by LC/MS-MS.

Samples from analytical SDS-PAGE, one slice each time, were loaded on a Waters Nano acquity system (Waters, Milford, MA, United States). The peptides were desalted on-line using a Waters Symmetry C18 180 μm × 20 mm, 5 μm trap column. The sample injection volume was typically 7.5 μL, and the LC was performed by using BEH 130 C18 100 μm × 100 mm, 1.7 μm column (Waters, Milford, MA) and eluting (0.5 μL min^{-1}) with a linear gradient (10–40%) of acetonitrile containing 0.1% formic acid using 70 min running time. Samples were injected on-line into a Micro Q-Tof spectrometer (Waters, Milford, MA). The ESI voltage was set at 3500 V, the source temperature was 80°C and the cone voltage was 30 V. The instrument control and data acquisition were conducted by a MassLynx data system (Version 4.1, Waters), and experiments were performed by scanning from a mass-to-charge ratio (m/z) of 400–2000 using a scan time of 1 s, applied during the whole chromatographic process. The exact mass was determined automatically using the Q-Tof's LockSpray™ (Waters, Milford, MA). Data-dependent MS/MS acquisitions were performed on precursors with charge states of 2, 3 or 4 over a range of 50–2000 m/z and under a 2 m/z window. A maximum of three ions were selected for MS/MS from a single MS survey. Collision-induced dissociation (CID) MS/MS spectra were obtained using argon as the collision gas at a pressure of 40 psi, and the collision voltage was varied between 18 and 90 V depending on the mass and charge of the precursor.

The results from this method are referred to as 1-DE and data were processed using the ProteinLynx Global server (version 2.5, Waters). The processing automatically lock mass corrected the m/z scale of both the MS and MS/MS data utilizing the lock spray reference ion. Generated peak list files were submitted to Mascot-Matrix science search engine (Version 2.2) protein and peptide tolerance were set as 0.1 Da. All proteins with score enough to assure a p-value lower than 0.05 was considered as a positive identification.

Tryptic peptides, obtained from the preparative gel were loaded in quintuplicate on a Waters Nano acquity system (Waters, Milford, MA, United States). The peptides were desalted on-line using a Waters Symmetry C18 180 μm × 20 mm, 5 μm trap column. The sample injection volume was 2 μL according to previous injection used to normalize total ion intensity in

all samples. Glycogen phosphorylase standard (Waters, Milford, MA, United States) was added to each sample so that 100 fmol was inject in each and all samples. The LC was performed by using a 1.7 μm HSS T3 130 C18 (150 μm × 75 mm) column (Waters, Milford, MA, United States) and eluting (0.5 μL/min) with a linear gradient (10–40%) of acetonitrile containing 0.1% formic acid using 150 min running time. Electrospray tandem mass spectra were obtained using a Waters SynaptTM G1 HD/MS High Definition Mass Spectrometer (Waters, Manchester, United Kingdom) interfaced to the Nano acquity system capillary chromatograph. The ESI voltage was set at 3500 V, the source temperature was 80°C and the cone voltage was 30 V. Instrument control and data acquisition were conducted by a MassLynx data system (Version 4.1, Waters), and experiments were performed by scanning from a mass-to-charge ratio (m/z) of 50–2000 using a scan time of 0.8 s.

The results from this method are referred to as label-free LC-MS/MS and data were processed using the Progenesis QI for proteomics version 2.0 software platform (Non-linear Dynamics, Waters, Manchester, United Kingdom), the exact masses was determined automatically using the Q-Tof's LockSpray^{TM}. For all identification analysis, raw data were search against *M. thermophila* Uniprot non-reviewed protein Database, including Human Keratin proteins and Sus scorfa Trypsin as possible contaminants. Cysteine carbamidomethylation was set as fixed modification, peptide N-terminal carbamidomethylation, methionine oxidation and asparagine deamidation were set as variable modification. Only peptide ions with PLGS score above 5.0 were considered as an identification match, and only proteins quantifications with ANOVA *p*-value below 0.05 were reported. Functional protein annotation was then performed using Uniprot and the CAZy database[1].

Cloning and Heterologous Expression of Two AA9 LPMOs in *Pichia pastoris*

The two target genes AA9 LPMOs, *Mt*LPMO9B (GenBank ID 11509292) and *Mt*LPMO9H (GenBank ID 11510592), in frame adding a C-terminal $(His)_6$-tag to the recombinant proteins, were codon optimized for expression in *P. pastoris* Easy Select Expression System (Genewiz Inc., United States). The *Pme*I-linearized pPICZαA recombinant plasmids were inserted into *P. pastoris* competent cells by electroporation as previously described (Bennati-Granier et al., 2015; Haon et al., 2015). Zeocin-resistant transformants were then screened for protein production.

The best producing transformants were grown in 1 L of BMGY containing 1 mL L^{-1} of *Pichia* trace minerals 4 (PTM_4) salts (2 g L^{-1} $CuSO_4.5H_2O$; 3 g L^{-1} $MnSO_4.H_2O$; 0.2 g L^{-1} $Na_2MoO_4.2H_2O$; 0.02 g L^{-1} H_3BO_3; 0.5 g L^{-1} $CaSO_4.2H_2O$; 0.5 g L^{-1} $CoCl_2$; 12.5 g L^{-1} $ZnSO_4.7H_2O$; 22 g L^{-1} $FeSO_4.7H_2O$; 0.08 g L^{-1} NaI; biotin and 1 mL L^{-1} concentrated H_2SO_4) in shaken flasks at 30°C in an orbital shaker (200 rpm) for 16 h to an OD600 of 2–6. Expression was induced by transferring the cells into 200 mL of BMMY containing 1 mL L^{-1} of PTM_4 salts at 20°C in an orbital shaker (200 rpm) for another 3 days. Each day the medium was supplemented with 3% (v/v) methanol.

Recombinants AA9 *Mt*LPMOs Purification

The supernatants were collected after harvesting cells by centrifugation at 3,500 *g* for 5 min at 4°C. After adjusting the pH to 7.8, the supernatants were filtered on 0.45 μm filters (Millipore, Molsheim, France) and loaded onto a 5 mL HisTrap HP columns (GE healthcare, Buc, France) connected to an Akta Xpress system (GE healthcare). Prior to loading, the columns were equilibrated with buffer A Tris-HCl 50 mM pH 7.8, NaCl 150 mM and imidazole 10 mM. The (His)6-tagged recombinant enzymes were eluted with buffer B Tris-HCl 50 mM pH 7.8, NaCl 150 mM and imidazole 500 mM. The fractions eluted containing the purified proteins were pooled, concentrated with a 10 kDa vivaspin concentrator unit (Sartorius, Plaiseau, France) and dialyzed against 50 mM sodium acetate buffer pH 5.2. The concentrated proteins were incubated with a double equimolar equivalent of $CuSO_4$ in a cold room and buffer exchanged in 50 mM sodium acetate buffer pH 5.2 using extensive washing in a 10-kDa ultrafiltration to remove traces of $CuSO_4$.

Protein concentrations were determined using a NanoDrop ND-2000 spectrophotometer (Thermo Fisher Scientific, IL, United States) by adsorption at 280 nm with theoretical molecular masses and molar extinction coefficients calculated from protein sequences using Expasy tools. An aliquot of 10 μg of each protein was loaded onto 10% Tris-glycine precast SDS-PAGE stain-free gel (Bio-rad, Marnes-la- Coquette, France) to check protein purity and integrity (**Supplementary Figure S1**). The molecular mass under denaturing conditions was determined with PageRuler Prestained Protein Ladder (Thermo Fisher Scientific).

Activity of AA9 *Mt*LPMOs on Cellulosic Substrates

The activities of the AA9 *Mt*LPMOs were carried out by addition of 1 μM *Mt*LPMO9B or *Mt*LPMO9H in 50 mM sodium acetate pH 5.2, 1 mM of ascorbate as electron donor using 0.1% (w/v) phosphoric acid swollen cellulose (PASC) or Avicel as substrates. The cleavage assays occurred in a final reaction volume of 300 μL in a 2 mL tubes in a rotatory shaker at 850 rpm (Infors AG, Switzerland). PASC was prepared from Avicel as described by Wood (1988).

The enzymatic reactions were performed at 40°C and interrupted at 30 min, 1, 2, 6 and 24 h by boiling the samples for 10 min. The samples were then centrifuged at 10,000 *g* for 10 min and the supernatants were stocked in a cold room until analysis. To get insights about temperature dependence, the experiment was performed using PASC as substrate, in water baths adjusted for the desirable temperatures, all at the same time, from 30° to 80°C, except the room temperature experiment, which was done at 22°C on the bench. After 2 h, the samples were boiled for 10 min to stop the reaction and centrifuged as described above. The supernatants containing oligosaccharides and their corresponding aldonic acid and C4-gemdiol forms generated after PASC and Avicel cleavage were analyzed by high-performance anion exchange chromatography coupled with amperometric detection (HPAEC-PAD) as described by

Westereng et al. (2013). Non-oxidized cello-oligosaccharides standards from DP2 to DP6 were purchased from Megazyme (Wicklow, Ireland) and C1-oxidized standards were produced from non-oxidized cello-oligosaccharides using a cellobiose dehydrogenase as described in Bennati-Granier et al. (2015). The experiments were performed in triplicates.

Activities of AA9 *Mt*LPMOs on Sugarcane Bagasse and Saccharifcation Assays

The enzymatic reactions were performed in 2 mL tubes containing 20 g L^{-1} of either SCBAT or SCBPD in 50 mM sodium acetate pH 5.2, supplemented with 0.15 mg mL^{-1} of tetracycline as antibiotic, cycloheximide 0.04 mg mL^{-1} as antifungal agent, and 1 mM ascorbate. Reaction were performed at 30°C or 50°C in a rotatory shaker at 850 rpm and started by addition of 1 μM *Mt*LPMO9B or 1 μM *Mt*LPMO9H in 1 mL final volume in 2 mL tubes. After 16 h incubation, the samples were centrifuged at 5,000 *g* for 5 min, and aliquots of 0.25 mL from each sample were taken and boiled for 10 min, and oxidized products were analyzed on HPAEC-PAD. The experiments were performed in triplicates. The reminiscent 0.75 mL volume were used for saccharification assays. For this, 0.25 mL of *Trichoderma reesei* cocktail (Celluclast, Sigma, Reims, Fr) containing 0.7 filter paper Unit. G^{-1} d.m and 60 $U.g^{-1}$ d.m of β-glucosidase (Novozymes 188) were added to the 2 mL tubes. The reaction was conducted for 4 h at 40°C and 850 rpm in a rotatory shaker (Infors AG, Switzerland). The glucose yield was measured by HPAEC-PAD. Control conditions were carried out without LPMO and without commercial *T. reesei* cocktail.

RESULTS

Carbohydrate-Active Enzymes Secreted by *M. thermophila*

The strain *M. thermophila* M7.7 was previously isolated from sugarcane bagasse piles (Moretti et al., 2012). We first pretreated sugarcane bagasse using different methods to induce the secretion of enzymes by *M. thermophila*. Composition analyses showed that both pretreated sugarcane bagasse samples, i.e., acid treated (SCBAT) and partially delignified (SCBPD), differed from the sugarcane bagasse not treated (SCBNT) in terms of cellulose, hemicellulose and lignin composition (**Supplementary Table S1**). The proteins secreted by *M. thermophila* upon growth on pretreated sugarcane bagasse feedstocks were collected at the initial stage of growth and analyzed by proteomics using two different methodologies (1-DE and label-free LC-MS/MS, see "Materials and Methods" section). Data analyzed using mass matching against predicted proteins inferred from genomic and transcriptomic sequence data of the *M. thermophila* ATCC 42464 (Berka et al., 2011) reported a total of 54 proteins, 21 identified by both methodologies, 16 only by label-free LC-MS/MS and 17 only by 1-DE as showed in Additional file (**Supplementary Table S2**). Out of the 54 proteins, 33 assigned as CAZymes (**Figure 1A**). Interestingly 10 of these CAZymes harbored at least one CBM module (from family CBM1, CBM6, CBM20, CBM35, CBM43) attached to their catalytic module. The analyses from the two methodologies used in this work showed that among the 54 proteins matched, 24 enzymes were found in both secretomes among which 15 CAZymes predicted to act on cellulose (GH7 cellobiohydrolase and GH131 endoglucanase) and hemicelluloses (GH11 endo-xylanase and GH26-CBM6 endo-mannanase). Of note, a family CE16 enzyme known to act on a wide range of carbohydrate acetyl esters (Puchart et al., 2016) could help the hemicellulases to deconstruct mannans and xylans. To complete this set of GHs, two AA3 oxidoreductases have been identified in both secretomes. Regarding the secretomes produced using SCBNT or SCBAT as inducer, several GHs were specifically identified in these secretomes among which a GH3 β-glucosidase and several endoglucanases among which a GH131 broad specificity endoglucanase bearing a CBM1 module (Lafond et al., 2012). One laccase (AA1) and one AA9 LPMO bearing a CBM1 were also identified only in the SCBAT secretome. Interestingly, the secretome produced on SCBPD displayed another AA9-CBM1 LPMO together with a LPMO from the AA16 family (Filiatrault-Chastel et al., 2019) and a cellobiose dehydrogenase (CDH) enzyme (AA3_1-AA8) that may serve as electron provider (Bey et al., 2013) for AA9 LPMOs.

Sequence Analysis and Recombinant Production of Two *Mt*LPMO9s

The genes encoding two AA9 LPMOs were selected from the proteomic dataset. These two AA9 LPMOs were named according to previous work (Berka et al., 2011; Frommhagen et al., 2016; Karnaouri et al., 2017), i.e., *Mt*LPMO9B (GenBank ID AON76800.1) and *Mt*LPMO9H (GenBank ID AEO56542.1). The sequence of *Mt*LPMO9B is encoded by 323 amino-acids (aa), including the main catalytic domain (aa position at 19–241) appended to a carbohydrate binding module at the C-terminus (CBM1) (aa position at 285–323) and *Mt*LPMO9H is encoded by 342 aa including main catalytic domain (aa position at 20–251) carrying a CBM1 C-terminal extension (aa position at 306–342) (**Figure 1B**). In both enzymes, the AA9 catalytic module start with the canonical conserved histidine residue common to all LPMOs. Despite of the proximity in sequence length, sequence alignment of these AA9 LPMOs reported considerable variation with only 37% of identity (**Supplementary Figure S2**). It is indeed well known that in AA9 LPMOs, only the N-terminal part of the proteins are conserved, the rest of the sequence being highly variable (Lenfant et al., 2017). The recombinant production of these LPMOs was designed using their native signal sequence since it is important to ensure the correct processing of the signal peptide and therefore the correct binding of the catalytic copper ion to the histidine brace (Daly and Hearn, 2005; Ahmad et al., 2014; Ladevèze et al., 2017). Both enzymes were successfully expressed in *P. pastoris* and the best transformants reached a production yield above 100 mg per liter of culture. After purification, SDS-PAGE analysis revealed a unique band for each protein with apparent molecular

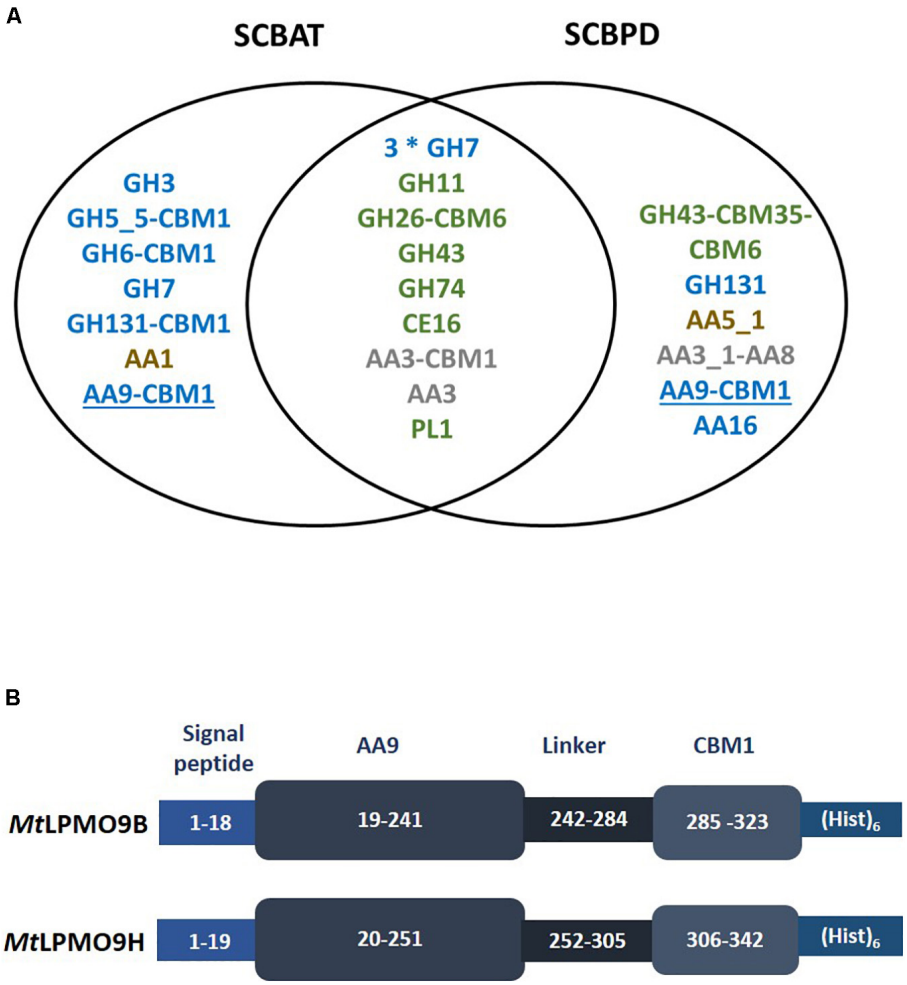


FIGURE 1 | Proteomic analysis of *Myceliophthora thermophila* secretomes leading to the identification of two AA9 LPMOs. **(A)** List of CAZymes identified in the secretomes using proteomics. Only the CAZymes targeting plant cell wall polysaccharides are indicated. Putative cellulases in blue, hemicellulases and pectinases in green and lignin-active enzymes in brown. AA3 enzymes that may act in synergy with LPMOs are indicated in gray and AA9 LPMOs are underlined. **(B)** Schematic representation of the recombinant *Mt*LPMO9s expressed in this study. The numbers represent the amino-acid numbering in the primary structure of the polypeptide formed by signal peptide, AA9 catalytic domain, carbohydrate binding domain (CBM) and the C-terminal (His)6-tag.

weights higher than the ones expected from the sequences (**Supplementary Figure S1**) suggesting some glycosylations most probably in the linker regions.

Activities of *Mt*LPMO9s on Cellulosic Substrates

To analyze the oxidative cleavage of cellulosic substrates with different levels of recalcitrance, *Mt*LPMO9B and *Mt*LPMO9H were assayed on Avicel and PASC, using ascorbate as an electron donor and soluble products were analyzed using ionic chromatography (HPAEC-PAD). Overall, we observed that PASC was a better substrate than Avicel for both LPMOs (**Figure 2**). On PASC, *Mt*LPMO9B released small peaks of non-oxidized cello-oligosaccharides (DP2–DP6) and several peaks attributed to C1-oxidized products based on standards, mostly DP4ox followed by DP2ox (retention time between 17.5 and 22 min). Regarding *Mt*LPMO9H, a different pattern of soluble products was observed. We detected peaks corresponding to non-oxidized cello-oligosaccharides (DP2–DP6) and both C1- and C4-oxidized oligosaccharides. C1-oxidized products were mostly represented by DP4ox, followed by DP5ox and DP6ox and C4-oxidized products by species eluting between 26 and 29 min (**Figure 2**). *Mt*LPMO9H also released C1-C4 double oxidized products with late retention time between 35 and 40 min. These peaks increase significantly along the course of the reaction from 30 min to 6 h (**Figure 2**). The identification of C4-oxidized and C1-C4 oxidized products are based on previous analyses (Bennati-Granier et al., 2015; Ladevèze et al., 2017; Westereng et al., 2017). On Avicel, products were much less abundant and only significantly observed with *Mt*LPMO9H (**Figure 2**). From these data, it can be concluded that these two *Mt*LPMO9s display different regioselectivities and a preference for the amorphous cellulose.

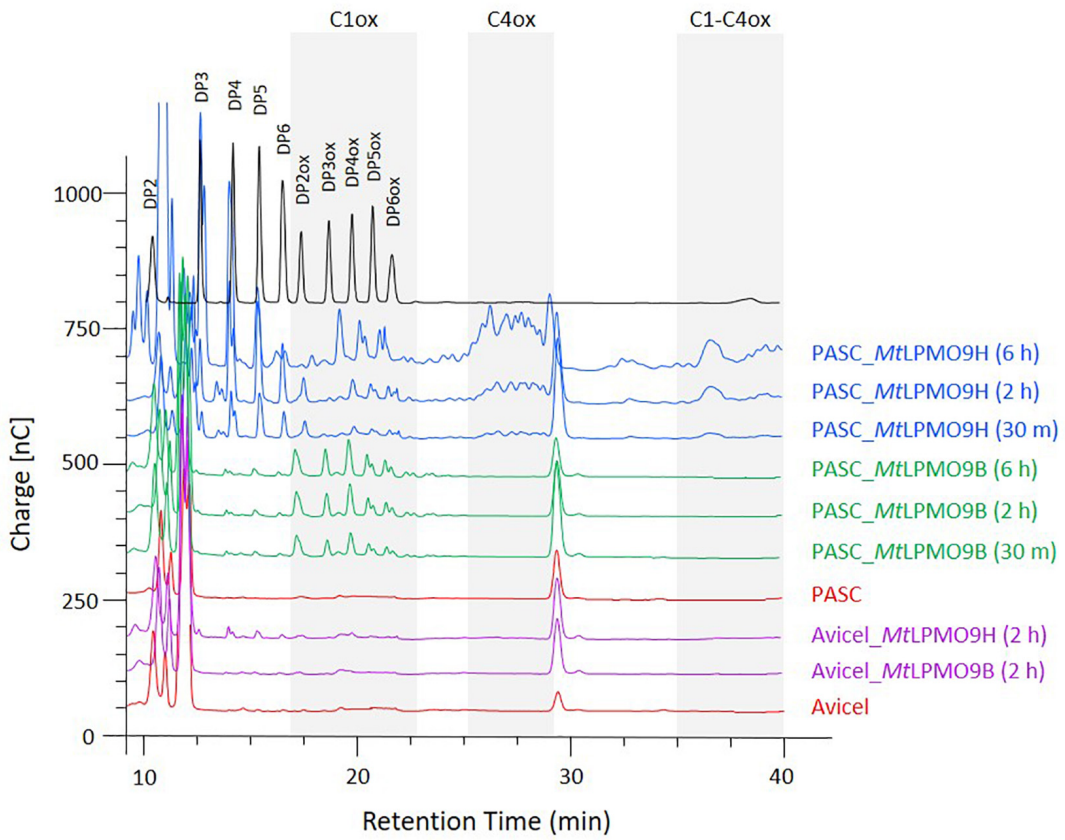


FIGURE 2 | Time course analysis of the oligosaccharides released from PASC and Avicel by *Mt*LPMO9B and *Mt*LPMO9H. HPAEC chromatograms show the enzymes activity time dependence, in the presence of 1 mM Ascorbate. Standards are shown in black and control reactions in orange.

Effect of Temperature on *Mt*LPMOs Activity

We further analyzed the activities of *Mt*LPMO9B and *Mt*LPMO9H in a biorefinery context. As *M. thermophila* is a thermophilic fungus, we hypothesized that *Mt*LPMO9s could present an interesting potential at elevated temperatures. Therefore, PASC cleavage was evaluated at different temperatures ranging from 22 to 80°C. From the chromatograms profile of the soluble non-oxidized and oxidized oligosaccharides produced, we observed that the *Mt*LPMO9s studied display distinct temperature preferences (**Figure 3**). The activity of *Mt*LPMO9B was maximum around 22–30°C and slowly decreased while temperature increased (**Figure 3C**). However, the activity of *Mt*LPMO9H displayed a typical bell shape with an optimum temperature around 50°C (**Figure 3C**). Although *Mt*LPMO9B and *Mt*LPMO9H showed optimal activity at 22–30°C and 50°C, respectively, both enzymes were able to release a broad range of cello-oligosaccharides with distinct intensity under a temperatures ranging from 22 to 70°C (**Figures 3A,B**). Interestingly, significant LPMO activity was still detected for both enzymes at 80°C (**Figures 3A,B**). These data illustrate the potential of these *Mt*LPMOs for biomass conversion at high temperatures.

Contribution of *Mt*LPMO9s on Sugarcane Bagasse Saccharification

To investigate the activities of *Mt*LPMO9s on biomass, both enzymes were tested on SCBAT and SCBPD pretreated bagasse feedstocks at different temperatures in a sequential reaction with *Mt*LPMO9s first before addition of a commercial *T. reesei* cocktail (see section "Materials and Methods"). Before addition of the *T. reesei* cocktail, we analyzed the action of *Mt*LPMO9s to make sure they were able to access and cleave cellulose present in pretreated bagasse feedstocks. The HPAEC-PAD chromatograms revealed the release of oxidized products from SCBAT and SCBPD by each of the *Mt*LPMO9s tested alone (**Figure 4A**). As expected, from the SCBPD substrate, *Mt*LPMO9B releases C1-oxidized cello-oligosaccharides, mainly from DP2ox to DP6ox while *Mt*LPMO9H releases a mixture of C1, C4, and C1/C4 oxidized products (**Figure 4A**). When assayed on SCBAT, a biomass with a lower content of cellulose and a higher level of recalcitrance than SCBPD, the *Mt*LPMO9 enzymes were still able to access cellulose and release a range of oxidized cello-oligosaccharides (**Figure 4A**). The strongest effect in term of soluble products released (both oxidized and non-oxidized) was observed using *Mt*LPMO9H at 50°C on SCBPD (**Figure 4A**). After addition of the commercial *T. reesei* cellulase cocktail, we

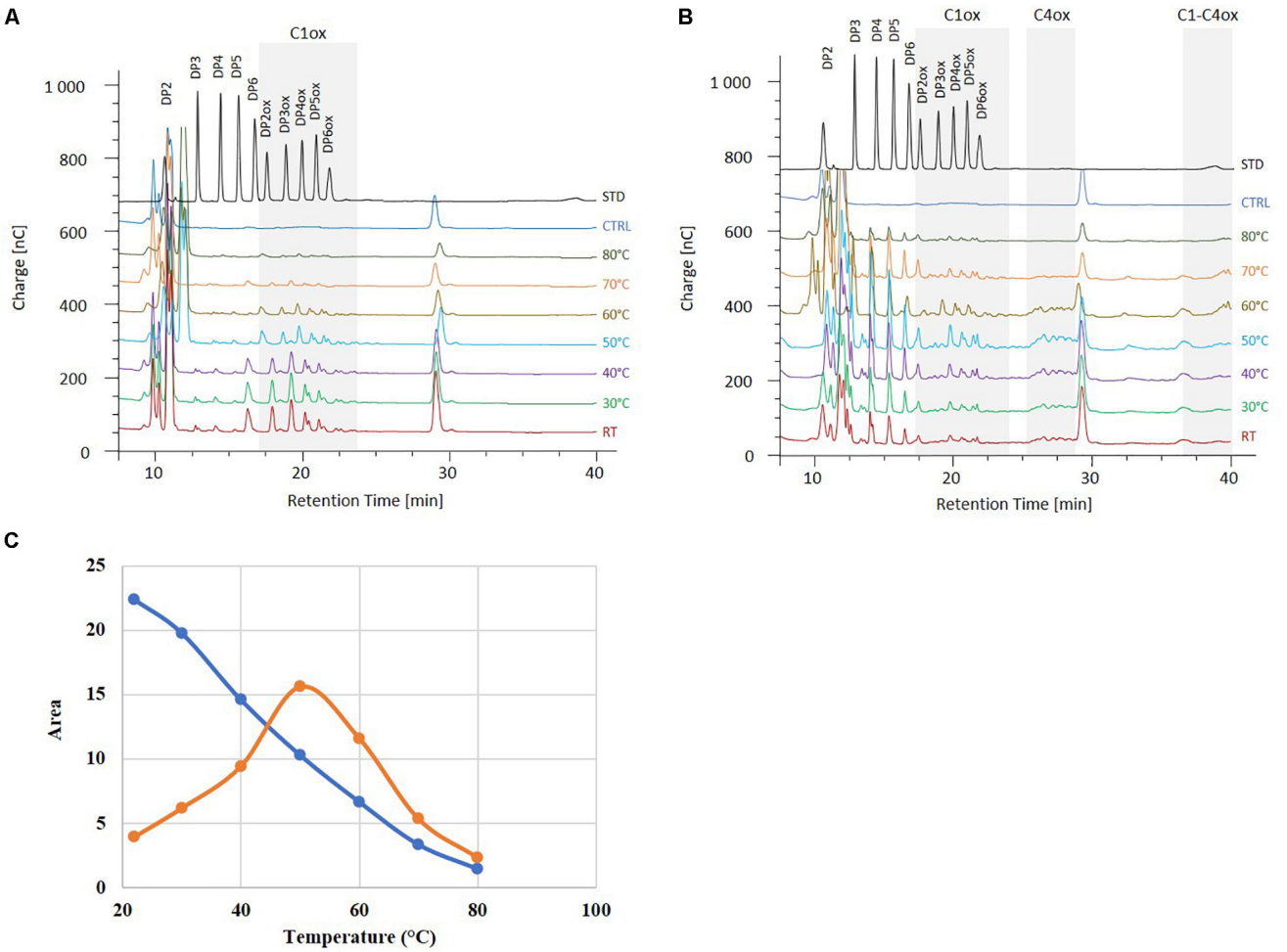


FIGURE 3 | Effect of the temperature on the *Mt*LPMO9s activities. The oligosaccharides released by *Mt*LPMO9B **(A)** and 1 μM *Mt*LPMO9H **(B)** from PASC were analyzed by HPAEC after 2 h in the presence of 1 mM ascorbate. Standards (STD) and control reaction (CTRL) are indicated in black and blue, respectively. **(C)** Effect of the temperature on the activity of *Mt*LPMO9B (blue) and *Mt*LPMO9H (orange) represented by the release of C1-oxidized DP4 (DP4ox).

observed a significant improvement of the glucose yield (up to 16% improvement) for each *Mt*LPMO9s with both SCBAT and SCBPD bagasse samples (**Figures 4B,C**). The most striking effects were observed at the optimal temperature of each enzyme. Overall, the partially delignified bagasse (SCBPD) was the best substrate considering the oxidized oligosaccharides released and the acid treated bagasse (SCBAT) was the best one in terms of saccharification boost. However, it should be noted that the SCBPD pretreatment was much more efficient as it allowed the release of three times higher yield of glucose compared to the SCBAT pretreatment (**Figures 4B,C**).

DISCUSSION

In this study, we focused on the enzymatic capabilities of *M. thermophila* on sugarcane bagasse, which is a relevant 2G feedstock in Brazil (Betancur and Pereira, 2010). After pretreatment of bagasse with different methods leading to feedstocks with various cellulose/hemicellulose/lignin proportions, we analyzed the enzymes secreted by this fungus at an early stage of bagasse degradation. This approach unraveled the secretion of several AA enzymes, including AA9 LPMOs that may act in synergy with cellulases. This early stage secretion of LPMO enzymes was also observed in the fungus *Laetisaria arvalis*, which is highly efficient on cellulose and known to secrete LPMOs before cellulases (Navarro et al., 2014). Another interesting observation is that the two AA9 LPMOs identified in the secretomes both display a CBM1 module. The preferential secretion of CBM-containing LPMOs by filamentous fungi is a trend already observed in *Podospora anserina* (Poidevin et al., 2014) and other fungal saprotrophs (Berrin et al., 2017) and could indicate a preferential role of these enzymes in lignocellulose degradation. Recent studies showed the functional relevance of CBM1 modules to promote LPMO binding and improve activity (Courtade et al., 2018; Chalak et al., 2019). For all these reasons, we decided to further study these

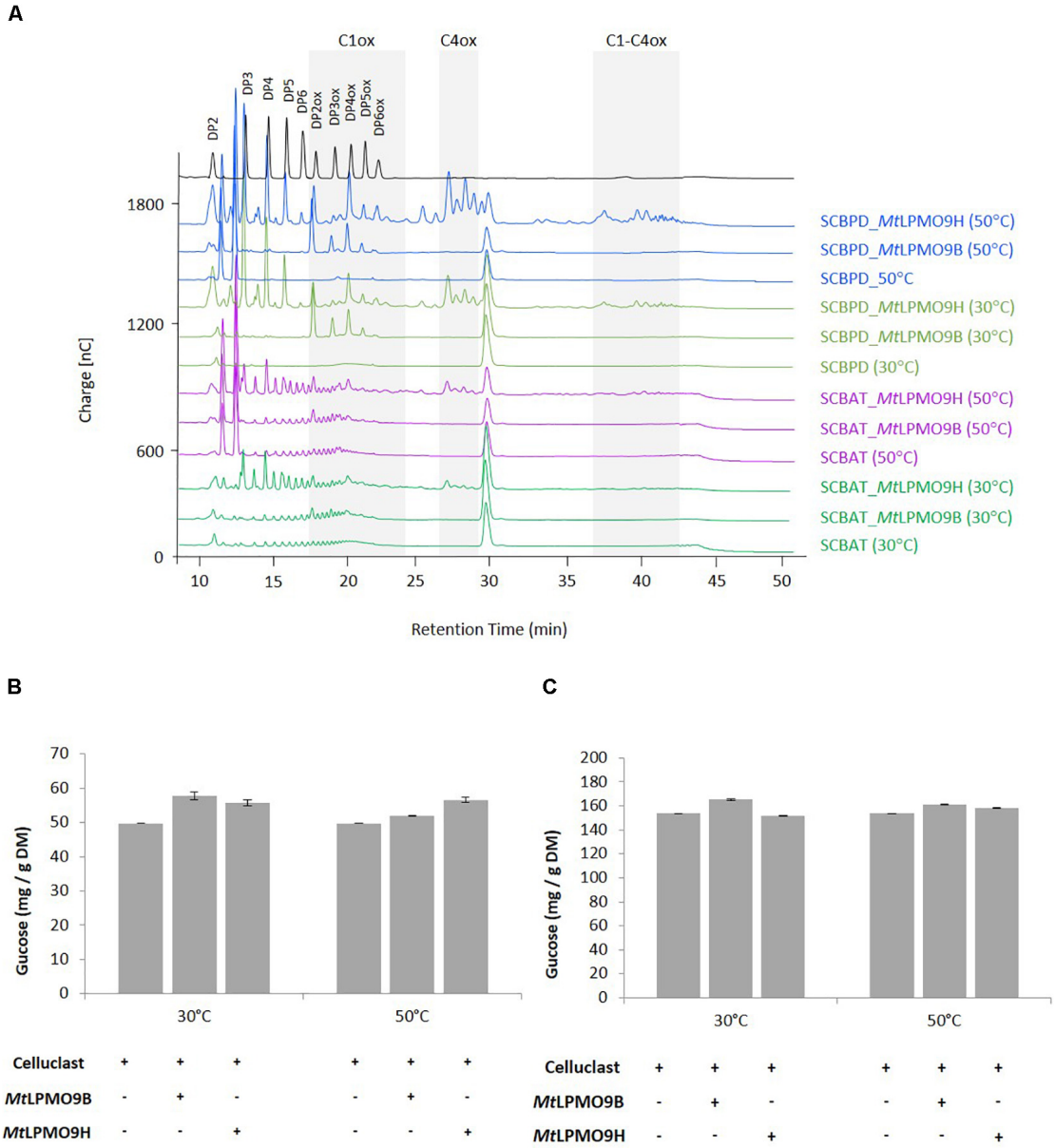


FIGURE 4 | Contribution of *Mt*LPMO9s to the saccharification of pretreated bagasse. **(A)** Analysis of oligosaccharides released from sugarcane bagasse acid treated (SCBAT) and partially delignified (SCBPD) by *Mt*LPMO9B and *Mt*LPMO9H at 30 and 50°C for 16 h before the addition of the *T. reesei* cocktail. All assays were run in the presence of 1mM ascorbate as electron donor. **(B,C)** Saccharification yield of the sugarcane bagasse **(B)** acid treated (SCBAT) and **(C)** acid/alkaline treated (SCBPD). The glucose released was quantified by HPAEC. Errors bars indicate standard deviations from triplicate samples.

AA9 LPMOs identified in *M. thermophila* secretomes. The *Mt*LPMO9B was previously studied to assess AA9 LPMO electron donor dependence (Frommhagen et al., 2016) and the *Mt*LPMO9H was studied in synergism with the GH5 endoglucanase from *M. thermophila* on PASC and pretreated wheat straw (Karnaouri et al., 2017). We comparatively characterized these *Mt*LPMO9s to identify interesting properties for the degradation of sugarcane bagasse. Our comparative assays using cellulose as model substrate revealed differences in regioselectivity and in optimum temperature cleavage. As compared to *Mt*LPMO9B, which was only active on amorphous cellulose, *Mt*LPMO9H was able to target both amorphous cellulose and crystalline cellulose. In terms of stability to temperature, *Mt*LPMO9H displayed a much better thermostability compared to *Mt*LPMO9B with an optimum temperature of 50°C but interestingly some activity was still detected for both enzymes at elevated temperatures. This observation confirms the strong binding capacity of copper to the histidine brace of LPMOs, even at elevated temperature. A strong binding of these LPMOs to cellulose could also

protect the enzyme from auto-inactivation as demonstrate by Kracher et al. (2018). Indeed, a stabilizing effect of the substrate was demonstrated on the apparent transition midpoint temperature of the reduced, catalytically active LPMO enzyme. This stability at elevated temperature is of great interest for saccharification of biomass. In our study, both LPMOs degrade pretreated sugarcane bagasse composed by distinct cellulose and lignin ratio. The pretreatment methods used in this study significantly affected the overall glucose yield but also the boosting effect observed for each LPMO suggesting that multiple factors are playing a role in this process. The regioselectivity of each LPMO and synergy with the cellobiohydrolases present in the *T. reesei* cocktail could be of importance as well as some inhibitors embedded in the sugarcane bagasse after pretreatment. Moreover, although it was not the focus of the present study, inhibitors molecules released during pretreatment can also affect the final desired product yield during the fermentation step. Another important criterion to consider when the objective is to maximize saccharification is to find the good balance between LPMO oxidation vs. cellulases hydrolysis. Indeed, too much oxidation at the surface of cellulose might not benefit the overall cellulose turnover to glucose. This could explain the reason why the partially delignified bagasse was the best substrate when considering the oxidized oligosaccharides released after LPMO action and the acid treated bagasse, which was the best one in terms of saccharification boost.

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

MG, NP, and J-GB conceived the study. MG, RZ, NP, and J-GB planned the experiments. MG, MH, SG, AO-C, and AM carried out the experiments. MG produced the secretomes. AO-C and AM carried out the proteomic analyses. MG and MH produced recombinant proteins. MG and SG performed substrate degradation assays and HPAEC–PAD analyses. MG, MH, SG, RZ, NP, and J-GB contributed to the interpretation of the results. NP and J-GB supervised the work. MG and J-GB wrote the manuscript. All authors read and approved the final version of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.01028/full#supplementary-material

TABLE S1 | Compositional analysis of sugarcane bagasses used in this study.

TABLE S2 | Proteins identified in the *M. thermophila* secretomes using proteomics.

FIGURE S1 | SDS PAGE of purified recombinant *Mt*LPMO9B (1-2) and *Mt*LPMO9H (3-4). A total of 5 μg (1 and 3) and 10 μg (2 and 4) of purified enzyme were loaded onto the gel. MW standards (kDa) and indicated in (5).

FIGURE S2 | Structure-based sequence alignment of the *Mt*LPMO9s studied. The conserved residues are shown in dark blue and the residues with similar physico-chemical properties are indicated in red. The residues forming the histidine brace are shown in light blue. The loops regions and disulfide bridges are also indicated. The sequence numbering start at the N-terminal Histidine (H1) and neither the linker nor the CBM are shown. The alignment was made using EMBL-EBI services and Pairwise Sequence Alignment (PSA).

REFERENCES

Ahmad, M., Hirz, M., Pichler, H., and Schwab, H. (2014). Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Appl. Microbiol. Biotechnol.* 98, 5301–5317. doi: 10.1007/s00253-014-5732-5

Bennati-Granier, C., Garajova, S., Champion, C., Grisel, S., Haon, M., Zhou, S., et al. (2015). Substrate specificity and regioselectivity of fungal AA9 lytic polysaccharide monooxygenases secreted by *Podospora anserina*. *Biotechnol Biofuels.* 8:90. doi: 10.1186/s13068-015-0274-3

Berka, R. M., Grigoriev, I. V., Otillar, R., Salamov, A., Grimwood, J., Reid, I., et al. (2011). Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. *Nat. Biotech.* 29, 922–927. doi: 10.1038/nbt.1976

Berrin, J. G., Rosso, M. N., and Hachem, M. A. (2017). Fungal secretomics to probe the biological functions of lytic polysaccharide monooxygenases. *Carbohydr. Res.* 7, 155–160. doi: 10.1016/j.carres.2017.05.010

Betancur, G. J., and Pereira, N. Jr. (2010). Sugar cane bagasse as feedstock for second generation ethanol production: part I: diluted acid pretreatment optimization. *Elect. J. Biotechnol.* 13, 10–11. doi: 10.2225/vol13-issue3-fulltext-3

Bey, M., Zhou, S., Poidevin, L., Henrissat, B., Coutinho, P. M., Berrin, J. G., et al. (2013). Cello-oligosaccharide oxidation reveals differences between two lytic polysaccharide monooxygenases (family GH61) from *Podospora anserina*. *Appl. Environ. Microbiol.* 2, 488–496. doi: 10.1128/AEM.02942-12

Chalak, A., Villares, A., Moreau, C., Haon, M., Grisel, S., Orlando, A., et al. (2019). Influence of the carbohydrate-binding module on the activity of a fungal AA9 lytic polysaccharide monooxygenase on cellulosic substrates. *Biotechnol. Biofuels* 12:206. doi: 10.1186/s13068-019-1548-y

Clauser, N. M., Gutiérrez, S., Area, M. C., Felissia, F. E., and Vallejos, M. E. (2016). Small-sized biorefineries as strategy to add value to sugarcane bagasse. *Chem. Eng. Res. Des.* 107, 137–146. doi: 10.1016/j.cherd.2015.10.050

Courtade, G., Forsberg, Z., Heggset, E. B., Eijsink, V. G. H., and Aachmann, F. L. (2018). The carbohydrate-binding module and linker of a modular lytic polysaccharide monooxygenase promote localized cellulose oxidation. *J. Biol. Chem.* 293, 13006–13015. doi: 10.1074/jbc.RA118.004269

Couturier, M., Ladevèze, S., Sulzenbacher, G., Ciano, L., Fanuel, M., Moreau, C., et al. (2018). Lytic xylan oxidases from wood-decay fungi unlock biomass degradation. *Nat. Chem. Biol.* 14, 306–310. doi: 10.1038/nchembio.2558

Daly, R., and Hearn, M. T. W. (2005). Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. *J. Mol. Recognit.* 18, 119–138. doi: 10.1002/jmr.687

Eggert, H., and Greaker, M. (2014). Promoting second generation biofuels: does the first generation pave the road? *Energies* 7, 4430–4435. doi: 10.3390/en7074430

Farrell, A. E., Plevin, R. J., Turner, B. T., Jones, A. D., O'Hare, M., and Kammen, D. M. (2006). Ethanol can contribute to energy and environmental goals. *Science* 311, 506–508. doi: 10.1126/science.1121416

Filiatrault-Chastel, C., Navarro, D., Haon, M., Grisel, S., Herpoël-Gimbert, I., Chevret, D., et al. (2019). AA16, a new lytic polysaccharide monooxygenase family identified in fungal secretomes. *Biotechnol. Biofuels* 12:55. doi: 10.1186/s13068-019-1394-y

Frommhagen, M., Koetsier, M. J., Westphal, A. H., Visser, J., Hinz, S. W., Vincken, J. P., et al. (2016). Lytic polysaccharide monooxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnol. Biofuels* 9, 1–17. doi: 10.1186/s13068-016-0594-y

Frommhagen, M., Sforza, S., Westphal, A. H., Visser, J., Hinz, S. W. A., Koetsier, M. J., et al. (2015). Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. *Biotechnol. Biofuels* 8:101. doi: 10.1186/s13068-015-0284-1

Frommhagen, M., Westphal, A. H., van Berkel, W. J. H., and Kabel, M. A. (2018). Distinct substrate specificities and electron-donating systems of fungal lytic polysaccharide monooxygenases. *Front. Microbiol.* 9:1080. doi: 10.3389/fmicb.2018.01080

Hangasky, J. A., Iavarone, A. T., and Marletta, M. A. (2018). Reactivity of O2 versus H2O2 with polysaccharide monooxygenases. *Proc. Natl. Acad. Sci. U.S.A.* 115, 4915–4920. doi: 10.1073/pnas.1801153115

Haon, M., Grisel, S., Navarro, D., Gruet, A., Berrin, J. G., and Bignon, C. (2015). Recombinant protein production facility for fungal biomass-degrading enzymes using the yeast *Pichia pastoris*. *Front. Microbiol.* 6:1002. doi: 10.3389/fmicb.2015.01002

Hemsworth, G. R., Henrissat, B., Davies, G. J., and Walton, P. H. (2014). Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nat. Chem. Biol.* 10, 122–126. doi: 10.1038/nchembio.1417

Johansen, K. S. (2016). Discovery and industrial applications of lytic polysaccharide monooxygenases. *Biochem. Soc. Trans.* 44, 143–149. doi: 10.1042/BST20150204

Karnaouri, A., Muraleedharan, M. N., Dimarogona, M., Topakas, E., Rova, U., Sandgren, M., et al. (2017). Recombinant expression of thermostable processive *Mt*EG5 endoglucanase and its synergism with *Mt*LPMO from *Myceliophthora thermophila* during the hydrolysis of lignocellulosic substrates. *Biotechnol. Biofuels* 10:126. doi: 10.1186/s13068-017-0813-1

Karnaouri, A., Topakas, E., Antonopoulou, I., and Christakopoulos, P. (2014). Genomic insights into the fungal lignocellulolytic system of *Myceliophthora thermophila*. *Front. Microbiol.* 5:281. doi: 10.3389/fmicb.2014.00281

Khattab, S. M. R., and Watanabe, T. (2019). "Sugarcane bagasse: status and perspectives," in *Bioethanol Production. From Food Crops*, eds R. C. Ray and S. Ramachandran (Cambridge, MA: Academic Press).

Kracher, D., Andlar, M., Furtmüller, P. G., and Ludwig, R. (2018). Active-site copper reduction promotes substrate binding of fungal lytic polysaccharide monooxygenase and reduces stability. *J. Biol. Chem.* 5, 1676–1687. doi: 10.1074/jbc.RA117.000109

Ladevèze, S., Haon, M., Villares, A., Cathala, B., Grisel, S., Herpoël-Gimbert, I., et al. (2017). The yeast *Geotrichum candidum* encodes functional lytic polysaccharide monooxygenases. *Biotechnol. Biofuels* 10:215. doi: 10.1186/s13068-017-0903-0

Lafond, M., Navarro, D., Haon, M., Couturier, M., and Berrin, J. G. (2012). Characterization of a broad-specificity β-glucanase acting on β-(1,3)-, β-(1,4)-, and β-(1,6)-glucans that defines a new glycoside hydrolase family. *Appl. Environ. Microbiol.* 24, 8540–8546. doi: 10.1128/AEM.02572-12

Lenfant, N., Hainaut, M., Terrapon, N., Drula, E., Lombard, V., and Henrissat, B. (2017). A bioinformatics analysis of 3400 lytic polysaccharide oxidases from family AA9. *Carbohydr. Res.* 448, 166–174. doi: 10.1016/j.carres.2017.04.012

Lo Leggio, L., Simmons, T. J., Poulsen, J. C., Frandsen, K. E., Hemsworth, G. R., Stringer, M. A., et al. (2015). Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase. *Nat. Commun.* 6:5961. doi: 10.1038/ncomms6961

Lombard, V., Golaconda, R. H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, 490–495. doi: 10.1093/nar/gkt1178

Loqué, D., Scheller, H. V., and Pauly, M. (2015). Engineering of plant cell walls for enhanced biofuel production. *Curr. Opin. Plant Biol.* 25, 151–161. doi: 10.1016/j.pbi.2015.05.018

Mandels, M., and Weber, J. (1969). The production of cellulases. *Adv. Chem.* 95, 391–413. doi: 10.1021/ba-1969-0095.ch023

Miller, G. L. (1959). Use of dinitrosalicilic acid for determination of reducing sugars. *Anal. Chem.* 3, 426–428. doi: 10.1021/ac60147a030

Moretti, M. M. S., Bocchini-Martins, D. A., Da Silva, R., Rodrigues, A., Sette, L. D., and Gomes, E. (2012). Selection of thermophilic and thermotolerant fungi for the production of cellulases and xylanases under solid-state fermentation. *Braz. J. Microbiol.* 43, 1062–1071. doi: 10.1590/S1517-83822012000300032

Navarro, D., Rosso, M.-N., Haon, M., Olivé, C., Bonnin, E., Lesage-Meessen, L., et al. (2014). Fast solubilization of recalcitrant cellulosic biomass by the basidiomycete fungus *Laetisaria arvalis* involves successive secretion of oxidative and hydrolytic enzymes. *Biotechnol. Biofuels.* 7:143. doi: 10.1186/s13068-014-0143-5

Poidevin, L., Berrin, J. G., Bennati-Granier, C., Levasseur, A., Herpoël-Gimbert, I., Chevret, D., et al. (2014). Comparative analyses of *Podospora anserina* secretomes reveal a large array of lignocellulose-active enzymes. *Appl. Microbiol. Biotechnol.* 17, 7457–7469. doi: 10.1007/s00253-014-5698-3

Puchart, V., Agger, J. W., Berrin, J. G., Várnai, A., Westereng, B., and Biely, P. (2016). Comparison of fungal carbohydrate esterases of family CE16 on artificial and natural substrates. *J. Biotechnol.* 233, 228–236. doi: 10.1016/j.jbiotec.2016.07.003

Quinlan, R. J., Sweeney, M. D., Lo-Leggio, L., Otten, H., Poulsen, J. C. N., Johansen, K. S., et al. (2011). Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc. Natl. Acad. Sci. U.S.A.* 108, 15079–15084. doi: 10.1073/pnas.1105776108

Singh, B. (2016). *Myceliophthora thermophila* syn. Sporotrichum thermophile: a thermophilic mould of biotechnological potential. *Crit. Rev. Biotechnol.* 36, 59–69. doi: 10.3109/07388551.2014.923985

Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., and Templeton, D. (2008). *Determination of Structural Carbohydrates and Lignin in Biomass*. Laboratory Analytical Procedure (LAP). Technical Report NREL/TP-510-42618. Golden, CO: National Renewable Energy Laboratory (NREL), 1–18.

Smithers, J. (2014). Review of sugarcane trash recovery systems for energy cogeneration in South Africa. *Renew. Sustain. Energy Rev.* 32, 915–925. doi: 10.1016/j.rser.2014.01.042

Tandrup, T., Frandsen, K. E. H., Johansen, K. S., Berrin, J. G., and Lo Leggio, L. (2018). Recent insights into lytic polysaccharide monooxygenases (LPMOs). *Biochem. Soc. Trans.* 6, 1431–1447. doi: 10.1042/BST20170549

Vaaje-Kolstad, G., Westereng, B., Horn, S. J., Liu, Z., Zhai, H., Sorlie, M., et al. (2010). An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* 330, 219–222. doi: 10.1126/science.1192231

Vasquez, M. P., Silva, J. N. C., Souza, M. B., and Pereira, N. Jr. (2007). Enzymatic hydrolysis optimization to ethanol production by simultaneous saccharification and fermentation. *Appl. Biochem. Biotechnol.* 137, 141–153. doi: 10.1007/s12010-007-9046-2

Vu, V. V., Beeson, W. T., Phillips, C. M., Cate, J. H., and Marletta, M. A. (2014a). Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J. Am. Chem. Soc.* 136, 562–565. doi: 10.1021/ja409384b

Vu, V. V., Beeson, W. T., Span, E. A., Farquhar, E. R., and Marletta, M. A. (2014b). A family of starch-active polysaccharide monooxygenases. *Proc. Natl. Acad. Sci. U.S.A.* 111, 13822–13827. doi: 10.1073/pnas.1408090111

Westereng, B., Agger, J. W., Horn, S. J., Vaaje-Kolstad, G., Aachmann, F. L., Stenstrøm, Y. H., et al. (2013). Efficient separation of oxidized cello-oligosaccharides generated by cellulose degrading lytic polysaccharide monooxygenases. *J. Chromatogr. A* 1271, 144–152. doi: 10.1016/j.chroma.2012.11.048

Westereng, B., Arntzen, M. O., Agger, J. W., Vaaje-Kolstad, G., and Eijsink, V. G. H. (2017). Analyzing activities of lytic polysaccharide monooxygenases by liquid chromatography and mass spectrometry. *Methods Mol. Biol.* 1588, 71–92. doi: 10.1007/978-1-4939-6899-2_7

Wood, T. M. (1988). Preparation of crystalline, amorphous, and dyed cellulase substrates. *Methods Enzymol.* 160, 19–25. doi: 10.1016/0076-6879(88)60103-0

Wyman, C. E. (2003). Potential synergies and challenges in refining cellulosic biomass to fuels, chemicals, and power. *Biotechnol. Prog.* 19, 254–262. doi: 10.1021/bp025654l

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 06 October 2020
doi: 10.3389/fbioe.2020.564527

Transcriptome Profiling-Based Analysis of Carbohydrate-Active Enzymes in *Aspergillus terreus* Involved in Plant Biomass Degradation

*Camila L. Corrêa[1], Glaucia E. O. Midorikawa[1], Edivaldo Ximenes Ferreira Filho[1], Eliane Ferreira Noronha[1], Gabriel S. C. Alves[1], Roberto Coiti Togawa[2], Orzenil Bonfim Silva-Junior[2], Marcos Mota do Carmo Costa[2], Priscila Grynberg[2] and Robert N. G. Miller[1]**

[1] Departamento de Biologia Celular, Universidade de Brasília, Campus Universitário Darcy Ribeiro, Brasília, Brazil, [2] Embrapa Recursos Genéticos e Biotecnologia, Parque Estação Biológica – PqEB, Brasília, Brazil

OPEN ACCESS

Edited by:
Fabiano Jares Contesini, Technical University of Denmark, Denmark

Reviewed by:
Jorge Barriuso, Spanish National Research Council, Spain
Rosana Goldbeck, State University of Campinas, Brazil
João Paulo Lourenço Franco Cairo, State University of Campinas, Brazil

***Correspondence:**
Robert N. G. Miller
robertmiller@unb.br

Specialty section:
This article was submitted to Bioprocess Engineering, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: 21 May 2020
Accepted: 16 September 2020
Published: 06 October 2020

Citation:
Corrêa CL, Midorikawa GEO, Filho EXF, Noronha EF, Alves GSC, Togawa RC, Silva-Junior OB, Costa MMC, Grynberg P and Miller RNG (2020) Transcriptome Profiling-Based Analysis of Carbohydrate-Active Enzymes in Aspergillus terreus Involved in Plant Biomass Degradation. Front. Bioeng. Biotechnol. 8:564527. doi: 10.3389/fbioe.2020.564527

Given the global abundance of plant biomass residues, potential exists in biorefinery-based applications with lignocellulolytic fungi. Frequently isolated from agricultural cellulosic materials, *Aspergillus terreus* is a fungus efficient in secretion of commercial enzymes such as cellulases, xylanases and phytases. In the context of biomass saccharification, lignocellulolytic enzyme secretion was analyzed in a strain of *A. terreus* following liquid culture with sugarcane bagasse (SB) (1% *w/v*) and soybean hulls (SH) (1% *w/v*) as sole carbon source, in comparison to glucose (G) (1% *w/v*). Analysis of the fungal secretome revealed a maximum of 1.017 UI.mL^{-1} xylanases after growth in minimal medium with SB, and 1.019 UI.mL^{-1} after incubation with SH as carbon source. The fungal transcriptome was characterized on SB and SH, with gene expression examined in comparison to equivalent growth on G as carbon source. Over 8000 genes were identified, including numerous encoding enzymes and transcription factors involved in the degradation of the plant cell wall, with significant expression modulation according to carbon source. Eighty-nine carbohydrate-active enzyme (CAZyme)-encoding genes were identified following growth on SB, of which 77 were differentially expressed. These comprised 78% glycoside hydrolases, 8% carbohydrate esterases, 2.5% polysaccharide lyases, and 11.5% auxiliary activities. Analysis of the glycoside hydrolase family revealed significant up-regulation for genes encoding 25 different GH family proteins, with predominance for families GH3, 5, 7, 10, and 43. For SH, from a total of 91 CAZyme-encoding genes, 83 were also significantly up-regulated in comparison to G. These comprised 80% glycoside hydrolases, 7% carbohydrate esterases, 5% polysaccharide lyases, 7% auxiliary activities (AA), and 1% glycosyltransferases. Similarly, within the glycoside hydrolases, significant up-regulation was observed for genes encoding 26 different GH family proteins, with predominance again for families GH3, 5, 10, 31, and 43. *A. terreus* is a promising species for production

of enzymes involved in the degradation of plant biomass. Given that this fungus is also able to produce thermophilic enzymes, this first global analysis of the transcriptome following cultivation on lignocellulosic carbon sources offers considerable potential for the application of candidate genes in biorefinery applications.

Keywords: *Aspergillus terreus*, lignocellulosic biomass, biorefinery, transcriptome, carbohydrate-active enzymes

INTRODUCTION

Biorefining can be defined as the sustainable processing of biomass into a spectrum of bio-products and bioenergy through a holistic approach that considers sustainability at the environmental, social and economic level (Jungmeier et al., 2013, 2014). In lignocellulosic feedstock biorefining, plant biomass is pre-treated to release cellulose, hemicellulose and lignin intermediates, with the first two of these polymers then bio-converted into fermentable pentose and hexoses, which can subsequently be further processed through sugar fermentation into bioenergy or other value-added bio-products. Lignin can also be employed in the production of chemical commodities and polymers (Windeisen and Wegener, 2012; de Jong and Gosselink, 2014). Given the global abundance of agricultural plant biomass residues, there is considerable potential today in such biorefinery-based applications in the production, amongst others, of bioethanol, industrial enzymes, organic acids, amino acids, secondary metabolites, pharmaceutical products, food, and animal feed (Silva et al., 2017).

The plant cell wall is recalcitrant to breakdown due to the complexity and heterogeneity of the diverse polymer components, each with different physical, chemical and functional properties (Alvira et al., 2010; Buckeridge and Souza, 2014; Zeng et al., 2017). Different macromolecular polymers make up plant lignocellulosic biomass, mainly comprising a complex matrix of interlinked cellulose, hemicelluloses, lignin and pectin. Although the amount of each polymer will vary according to each plant species, cellulose is typically the most abundant homopolymer in plant cell walls (40–55%), composed of glucose monosaccharides in β-1, 4 format that form a final amorphous or crystalline microfiber structure (Kimura et al., 1999; Mckendry, 2002). Hemicelluloses, while less abundant than cellulose (15–25%), still represent the second most abundant polymer on earth (Saha, 2003). The heteropolymers comprise xylans, xyloglucans, mannans, glucomannans, galactomannans, glucuronoarabinoxylans, glucuronoxylans, amongst other polysaccharides, interacting directly with cellulose fibers (Mohnen and Hahn, 1993). Lignin, whilst less abundant than cellulose (10–20%), provides structural integrity and overall stiffness to stems and root tissues (Albersheim et al., 2010). This chemically complex recalcitrant non-carbohydrate aromatic heteropolymer is made up of different syringyl, guaiacyl, and p-hydroxyphenyl monomer units (Boerjan et al., 2003). Pectins, which are also less abundant in the plant cell wall (up to 30%), are composed of highly variable polysaccharides, with a backbone composed principally of galacturonic acid with α-(1,4) bonds, together with different pectic domains of the structural classes homogalacturonan, xylogalacturonan, apiogalacturonan, arabinan, galactan, arabinogalactan I and II, and rhamnogalacturonan I and II (Ridley et al., 2001; Caffall and Mohnen, 2009; Gawkowska et al., 2018).

The enzymatic hydrolysis of plant lignocellulosic biomass is currently regarded as the most efficient biomass conversion method (Silva et al., 2017). Here, enzyme consortia with different specificities are required for lignocellulose breakdown, including cellulases, hemi-cellulases, ligninases, lytic polysaccharide mono-oxygenases (LPMOs) and cellobiose dehydrogenases (CDHs), together with the action of proteins such as swollenins and expansins (Van Dyk and Pletschke, 2012). Whilst such hydrolysis of plant cell wall polysaccharides to simple monomers has been a subject of considerable study, improved cost efficiency of enzymatic hydrolysis is essential for enzyme-based lignocellulosic biorefineries (Maijala et al., 2012; Silva et al., 2017).

Carbohydrate-active enzymes (CAZy enzymes/CAZymes) are classified as enzymes involved in the assembly, modification and breakdown of polysaccharides through their action on glycosidic bonds. As such, many are involved in the biosynthesis and degradation of the plant cell wall. These are today classified into several hundred enzyme protein superfamilies, families, and subfamilies, based on protein sequence similarity and predicted protein folding structure (Levasseur et al., 2013; Lombard et al., 2013). These broadly comprise the superfamilies of Glycosyl transferases (GT), Glycoside hydrolases (GH), Carbohydrate esterases (CE), Polysaccharide lyases (PL), and Auxiliary Activities (AA). Only GTs are involved in the formation of glycosidic bonds and assembly of complex carbohydrates, with all other superfamilies responsible for carbohydrate breakdown or modification. The GHs are responsible for the hydrolysis or modification of glycosidic bonds, the CEs for hydrolysis of carbohydrate esters, and the PLs for non-hydrolytic cleavage of glycosidic bonds. The broad AA superfamily groups redox enzymes and lytic polysaccharide mono-oxygenases (LPMOs) with different catalytic reaction mechanisms. These CAZymes are involved in both lignin breakdown and in facilitating access to plant cell wall carbohydrates for GH, CEs, and PLs (Cantarel et al., 2009; Levasseur et al., 2013). Different families of non-catalytic carbohydrate-binding modules (CBMs) can also occur within larger CAZyme protein sequences, or occasionally independently, where specific amino acid sequence domain folds enable carbohydrate-binding activities that serve to enhance the activity of the enzymatic modules.

Microbial enzymes can be employed in the bioconversion of complex lignocellulosic biomass, in hydrolysis and saccharification steps for release of simple sugars, and during subsequent fermentation into further industrial products. Bioconversion processes can be separate or can occur simultaneously, with consolidated bioprocessing

technologies offering complete enzyme production and secretion, hydrolysis and fermentation bioprocesses performed by a single microorganism.

Many different prokaryotic (Palazzolo and Kurina-Sanz, 2016) and eukaryotic microorganisms are important reservoirs of plant cell wall degrading enzymes appropriate for biorefinery applications. Of these, saprophytic filamentous fungi have been more widely studied as biodegraders of lignocellulosic biomass. Today a limited number of Ascomycete species within genera such as *Trichoderma, Aspergillus, Penicillium, Neurospora, Acremonium,* and *Chaetomium*, amongst others, play important roles as sources of commercial CAZymes, preferentially for cellulose and hemicellulose degradation (Martinez et al., 2008; Ribeiro et al., 2012; Glass et al., 2013; Horta et al., 2014; Borin et al., 2015, 2017; Kameshwar et al., 2019). Basidiomycete white-rot fungi can also serve as sources of enzymes to degrade lignin, through the action of oxidative ligninolytic enzymes (Ruiz-Dueñas et al., 2013; Kameshwar and Qin, 2018). Despite the considerable advances in enzyme characterization in recent years, further bioprospection is required for microbial enzymes with increased activity, resistant to inhibitors, and appropriate for universal enzyme cocktails and enzymatic consortia.

Whilst there has been considerable investigation into secreted fungal enzymes involved in the degradation of lignocellulose, the identification of the responsible genes encoding such CAZymes and their regulation mechanisms has been more limited, despite the importance of such information in optimization of biotechnological processes (Horta et al., 2014). Whole genome analysis has progressed considerably, with the MycoCosm database now housing thousands of complete fungal genomes, with annotation data providing catalogs of genes that enable filtering for gene repertoires involved in lignocellulose degradation (Grigoriev et al., 2014). As fungal enzyme production is subject to gene induction and repression based on the carbon source (Cantarel et al., 2009), the employment of different lignocellulosic materials can serve as a promising strategy for the induction of complex cellulase and hemicellulase enzymes in filamentous fungi (Sorensen et al., 2011). Although gene expression regulation mechanisms in response to specific polysaccharides has been closely studied in model fungi (e.g., Amore et al., 2013), analyses of fungal transcriptomes on different lignocellulosic carbon sources can also reveal important information on genes encoding CAZymes, their regulation, and ultimately, the molecular mechanisms of enzymatic breakdown of natural complex lignocellulosic biomasses (Hori et al., 2014; Cragg et al., 2015; Rosnow et al., 2017).

The genus *Aspergillus* comprises more than 300 recognized species, with a number of species of importance in a variety of areas of biotechnology and food production. A number of these species are also efficient in secretion of hydrolytic enzymatic of importance for enzyme-based lignocellulosic biorefineries, with transcriptome analyses in recent years providing important information on gene expression and regulation mechanisms for genes encoding different cellulases, endoxylanases, β-xylosidases and pectinases during degradation of lignocellulosic biomass. *Aspergillus terreus* is frequently isolated from soil rhizospheres and decaying agricultural cellulosic materials in tropical and subtropical regions (Wijeratne et al., 2003; Palonen et al., 2017). In addition to producing different statins that have been widely used in the treatment of heart disease, this species also produces important commercial enzymes employed in bioprocesses that include cellulases, xylanases, lipases and phytases (Yadav et al., 1998; Varga et al., 2005; Chantasingh et al., 2006; Narra et al., 2012; Sharma et al., 2014; Pierce et al., 2017). The reference genome sequence of *A. terreus* NIH2624 shows an estimated size 29.3 Mb, with 10406 predicted gene models (Arnaud et al., 2012). Of these, a total of 477 are predicted to be related to polysaccharide hydrolysis, when compared against the Carbohydrate-Active Enzymes Database[1]. Subsequent resequencing of different strains has reported approximately 500 genes as predicted to be involved in complex polysaccharide degradation (Dadheech et al., 2019).

Analysis of gene expression in lignocellulolytic fungal species cultivated on different plant biomass substrates as sole carbon source enables identification of conserved and specific mechanisms employed in the microbial degradation of diverse plant cell wall residues (Roche et al., 2014). In this study, we examine differential expression of CAZyme-encoding genes through transcriptome profiling of *A. terreus* strain BLU24 after growth on sugarcane bagasse and soybean hulls, which are both considered promising lignocellulosic substrates for biorefinery applications, in terms of the high concentration of carbohydrates that can be converted into fermentable sugar residues (Cheng and Zhu, 2013). This fungal strain has previously been reported to efficiently secrete xylanases on lignocellulosic biomass residues (de Souza Moreira et al., 2013). Increased understanding of plant cell wall degradation mechanisms will contribute to downstream rational engineering of improved microorganisms for efficient enzymatic degradation of agricultural plant biomass residues.

MATERIALS AND METHODS

Fungal Strain Identification

Aspergillus terreus strain BLU24, belonging to the fungal culture collection of the Enzymology Laboratory at the University of Brasilia (genetic heritage number 010237/2015-1), was originally isolated into pure culture from natural cotton fiber composting waste (Siqueira et al., 2009). For molecular identification confirmation, the fungus was grown on liquid Czapek Yeast Autolyzate medium (CYA) (Pitt and Hocking, 2009) for 3 days at 28°C with agitation at 120 rpm. Mycelium was washed with sterile distilled water, vacuum filtrated and lyophilized. Total DNA was extracted from 50 mg of macerated mycelium according to Raeder and Broda (1985). Quantification of total DNA was performed by comparison with a Low DNA Mass Ladder® (Invitrogen), following electrophoresis in 1% agarose gels containing ethidium bromide (1.0 $\mu g\ mL^{-1}$) and visualization under UV at 254 nm. The ribosomal DNA (rDNA) internal transcribed spacer (ITS) regions 1 and 2, together with regions of the β-tubulin and calmodulin genes, were PCR-amplified, respectively, using universal primers ITS5 and

[1]http://www.cazy.org

ITS4 (White et al., 1990), Bt2a and Bt2b (Glass and Donaldson, 1995), and Cmd5 and Cmd6 (Hong et al., 2006). Each PCR reaction contained 15 ng of template DNA, 0.4 μM of each primer, 200 μM dNTPs, 1.5 mM $MgCl_2$, 1.0U Taq DNA polymerase and 1× IB Taq polymerase buffer (Phoneutria, Belo Horizonte, Brazil). Temperature cycling followed a program of denaturation at 95°C for 4 min, 30 cycles of denaturation at 94°C for 1 min, primer annealing at 50°C (ITS) or 60°C (β-tubulin and calmodulin genes) for 1 min, and extension at 72°C for 1 min. A final elongation period at 72°C for 5 min was included. PCR products were purified using ExoSAP-IT® (USB, Cleveland, OH, United States) and forward and reverse- sequenced using the Big Dye® Terminator v3.1 Cycle Sequencing kit (Applied Biosystems, Foster City, CA, United States). Products were separated on an ABI 3130xl DNA sequencer (Applied Biosystems, Foster City, CA, United States). Generated sequences were quality filtered and clustered into contigs using the program Sequencher, version 4.8 (Gene codes Corporation, Ann Arbor, MI, United States). Molecular-based identification was performed on the basis of sequence identity using the program BLASTn (Altschul et al., 1997), against the NCBI nucleotide nr database. The ribosomal DNA ITS, β-tubulin and calmodulin gene consensus sequences confirming species level identification were deposited in GenBank under the accession number MT461408, MT472459, and MT472460, respectively.

Experimental Design

Aspergillus terreus BLU24 was grown in liquid minimal medium [7 g KH_2PO_4, 2g K_2HPO_4, 0.5 g $MgSO_4$, 1.6 g de $(NH_4)_2$ SO_4, pH 7.0, per L of distilled water], supplemented with sugarcane bagasse (SB), soybean hulls (SH) or glucose (G) (Sigma Aldrich) as sole carbon source (each at 1% w/v). Each plant biomass residue-derived carbon source, obtained from local sources, was pretreated through the addition of a 1:1 volume of water, autoclaving for 2 h at 120 psi, washing with distilled water, drying at 60°C for 48 h and material milling. For elimination of reducing sugars, these carbon sources were washed with deionized water, with reducing sugar elimination detected using the colorimetric dinitrosalicylic acid (DNS) assay (Miller, 1959). Fungal spores were harvested from a fresh culture of *A. terreus* BLU24 grown on Potato Dextrose Agar medium (Thermo Fisher Scientific, Paisley, United Kingdom), adjusted to a concentration of 1×10^8 conidiospores mL^{-1} and employed as inoculum. Liquid cultures were grown in 100 mL of each growth media in 250 mL Erlenmeyer flasks at 28°C and 128 rpm over a 9-day period. Culture arrangement followed a randomized block design, with three replicates for each growth treatment and incubation time point. Treatments were labeled according to the following codes: sugarcane carbon source, 36 h incubation (SB36); soybean hull carbon source, 36 h incubation (SH36); glucose carbon source, 36 h incubation (G36); sugarcane carbon source, 48 h incubation (SB48); soybean hull carbon source, 48 h incubation (SH48); glucose carbon source, 48 h incubation (G48).

Enzyme Analysis

Hydrolytic enzyme activities were evaluated during 9 days at 24 h intervals. Xylanase, CMCase, pectinase, and FPase assays were conducted using the DNS assay at pH 5.0 (Miller, 1959). Each assay, conducted on ELISA microtiter plates, comprised 10 μl of the fungal secretome, together with xylan (1% wt/vol), carboxy methylcellulose (1% wt/vol), or pectin (1% wt/vol) as substrate. Reactions were incubated for 30 min at 50°C. Quantification of total cellulases was performed with an assay for FPase activity according to Ghose (1987). 10 μl of each secretome was added to Whatman® filter paper No. 1.0 × 0.6 cm followed by incubation for 1 h at 50°C. All assays were conducted in triplicate. Reducing sugars from each assay were quantified on a spectrophotometer at a wavelength of 540 nm (Spectramax plus 384, Molecular Devices), with calibration using standard curves of glucose, xylose, mannose and galacturonic acid. Absorbance values obtained were expressed as international units (UI). The Tukey's test was performed at a $P < 0.05$ significance level for comparison of datasets on different carbon sources.

Scanning Electron Microscopy

The degradation of each lignocellulosic carbon source was examined by scanning electron microscopy (SEM) following growth of *A. terreus* BLU24 for 36 and 48 h in liquid minimal medium supplemented with SB or SH. Culture filtrates were washed using Karnovsky buffer (0.05M; pH 7.2) and material fixed for 4 h in 0.05M cacodylate buffer at pH 7.4. Samples were dehydrated using acetone and postfixed with 1% osmium tetroxide for 1 h. Following washes in liquid CO_2 at 4°C, exposure to critical point drying (Emitech K850, Kent, United Kingdom) and mounting on copper stubs for sputter coating with 20 nm gold particles, samples were observed at different magnifications using a Zeiss DSM 962 scanning electron microscope.

Total RNA Extraction

Conidiospores of *A. terreus* BLU24 were inoculated as described earlier into 100 mL minimal medium containing 1% glucose (w/v) or 1% of each crop-derived carbon source. Cultures were incubated on an orbital shaker at 128 rpm at 28°C for 36 or 48 h. All treatments (carbon source + growth period) were performed with repetition, with three biological replicate samples per treatment and duplicate bioassays. Culture mycelia were harvested after each treatment growth period and immediately frozen in liquid nitrogen. Isolation of total RNA from frozen mycelia was performed according to Brasileiro and Carneiro (2015), with DNA removed using Ambion® DNA-free™ DNase Treatment and Removal Reagents (Thermo Fisher Scientific, Waltham, MA, United States). RNA concentration and integrity were estimated with a NanoDrop 2000c spectrophotometer (Thermo Scientific, Waltham, MA, United States) with a quality cut-off value of 1.8 employed for the A260:280 ratio. Integrity of RNA was also measured using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, United States).

Construction of cDNA Libraries and RNA-Seq Transcriptome Profiling

A total of 12 RNA samples, representing all treatments and biological replicates, were shipped in RNAstable (Biomatrica) according to the manufacturer's instructions. Messenger RNA

isolation, cDNA library preparation and Illumina RNA-Seq were performed by Eurofins MWG Operon (Alabama, United States) from 10 μg of each total RNA sample. cDNA libraries were prepared using the TruSeq Sample RNA preparation kit v3 Kit according to the manufacturer's specifications and paired-end-sequenced (2x 100 base pairs) in two flowcell lanes on an Illumina Hiseq2500 sequencer (Illumina Inc., San Diego, CA, United States).

Sequence Processing and Bioinformatic Analyses

Sequence reads from each cDNA library were adapter-trimmed using the program Trimmomatic (Bolger et al., 2014) and quality assessment conducted using ea-utils (Aronesty, 2011). High quality reads (Fastq QC > 30) were mapped to the reference genome sequence for *A. terreus* strain NIH 2624[2] using the program Novoalign/Useq[3]. Unmapped sequence reads were analyzed through alignment to the NCBI non-redundant protein sequence database (nr) using Diamond (Buchfink et al., 2015), with an *E*-value cut-off at 10^{-9}. Reads accurately mapped to individual genes in the reference genome were counted using HTseq-count (Anders et al., 2015), with differences in gene expression between evaluated treatments determined using DESeq (Anders and Huber, 2010). Statistically significant differentially expressed genes (DEGs) were considered if a log2 fold change (FC) was at least ≥2-fold and at a probability level of $p \leq 0.01$.

Sequences were functionally annotated according to Gene Ontology (GO) terms (Ashburner et al., 2000), with over- and under-representation of DEGs assigned to GO terms analyzed using the program FUNC (Prüfer et al., 2007) and category term redundancy eliminated using the program REVIGO[4].

Genes encoding hydrolytic enzymes glycoside hydrolases, glycosyltransferases, carbohydrate-binding modules and carbohydrate esterases were identified through gene sequence alignment against the updated Carbohydrate-Active Enzymes (CAZymes) database (see text footnote 1) (Cantarel et al., 2009). Annotation validation at the family and subfamily level was conducted using CUPP for analysis of predicted conserved unique octamer peptide signatures (Barrett and Lange, 2019; Barrett et al., 2020).

Quantitative Real-Time PCR Validation of RNA-Seq-Derived Gene Expression Data

Validation of expression of 8 selected DEGs identified based on *in silico* transcriptome data was conducted by quantitative real-time PCR (RT-qPCR) analysis. Thermocycling was conducted on a Step One Plus Real Time PCR System (Applied Biosystems) using a Platinum® SYBR® Green qPCR Super Mix-UDG w/ROX kit (Invitrogen, Carlsbad, CA, United States) according to the manufacturer's guidelines and 1 μL of template cDNA. Specific primers were designed with Primer Express® (Applied Biosystems) and evaluated for specificity at OligoAnalyzer 3.1 – IDT[5]. All cDNA libraries were synthesized using original RNA samples that were employed for RNA-Seq analysis. Three independent biological replicates were analyzed for each treatment and three technical replicates included per amplification. Total RNA was initially treated with 2 U of Amplification Grade DNase I (Invitrogen, Carlsbad, CA, United States) and cDNA then synthesized using Oligo(dT)20 primers (Invitrogen, Carlsbad, CA, United States) and SuperScript® II Reverse Transcriptase (Invitrogen, Carlsbad, CA, United States). Thermal cycling was conducted by 40 cycles of denaturation at 95°C for 15 s, followed by primer annealing and amplicon extension at 60°C for 30 s. *A. terreus* Actin and rDNA 18S genes were included as stable reference genes for normalization. Cycle threshold (Ct) values were calculated using the program SDS 2.2.2 (Applied Biosystems, Foster City, United States), with PCR product specificity determined according to the Tm (dissociation) of the amplification product. Individual amplification efficiencies were calculated with the program LinRegPCR, version 2013.0 and gene expression values calculated according to the 2-ΔΔCT method (Livak and Schmittgen, 2001).

RESULTS

Scanning Electron Microscopy Analysis

Scanning electron microscopy observations following growth of *A. terreus* BLU24 in liquid minimal medium supplemented with the lignocellulosic carbon sources SB and SH showed parenchyma surface disruption and fungal colonization on each after 36 h, with considerable hyphal growth and conidiospore production visible after 48 h (**Figure 1**).

Enzyme Activity Profiles on Lignocellulosic Biomass

Enzymatic profiles of *A. terreus* BLU24 were evaluated over a 9-day period following growth on liquid minimal medium culture supplemented with SB or SH as sole lignocellulosic carbon source. Enzymatic secretions were compared to those observed following growth on the same minimal medium with G as sole carbon source, employed as a control treatment. Growth on both lignocellulosic carbon sources resulted in a rapid increase in xylanases during the first 48 h [(0.839 ± 0.007; 0.903 ± 0.010 UI mL^{-1}), respectively for SB and SH], remaining fairly constant for the remainder of the investigated time period (**Figure 2**). This was in contrast to growth on G, where xylanase activities remained low at 48 h (0.038 ± 0.040X UI mL^{-1}). Activities for CMCases, pectinases and FPases were comparatively lower than observed for xylanases after 48 h (0.124 ± 0.001, 0.072 ± 0.002; 0.055 ± 0.009, 0.388 ± 0.025; 0.304 ± 0.048, 0.234 ± 0.098 UI mL^{-1}), respectively on SB and SH, although again considerably raised in comparison to values following growth on G as carbon source at 48 h. Once again, activities remained fairly constant for these enzymes throughout the remaining time period of growth

[2]http://www.broadinstitute.org

[3]http://www.novocraft.com/documentation/novoalign-2/novoalign-user-guide/rnaseq-analysis-mrna-and-the-spliceosome/

[4]http://revigo.irb.hr/

[5]https://www.idtdna.com/calc/analyzer

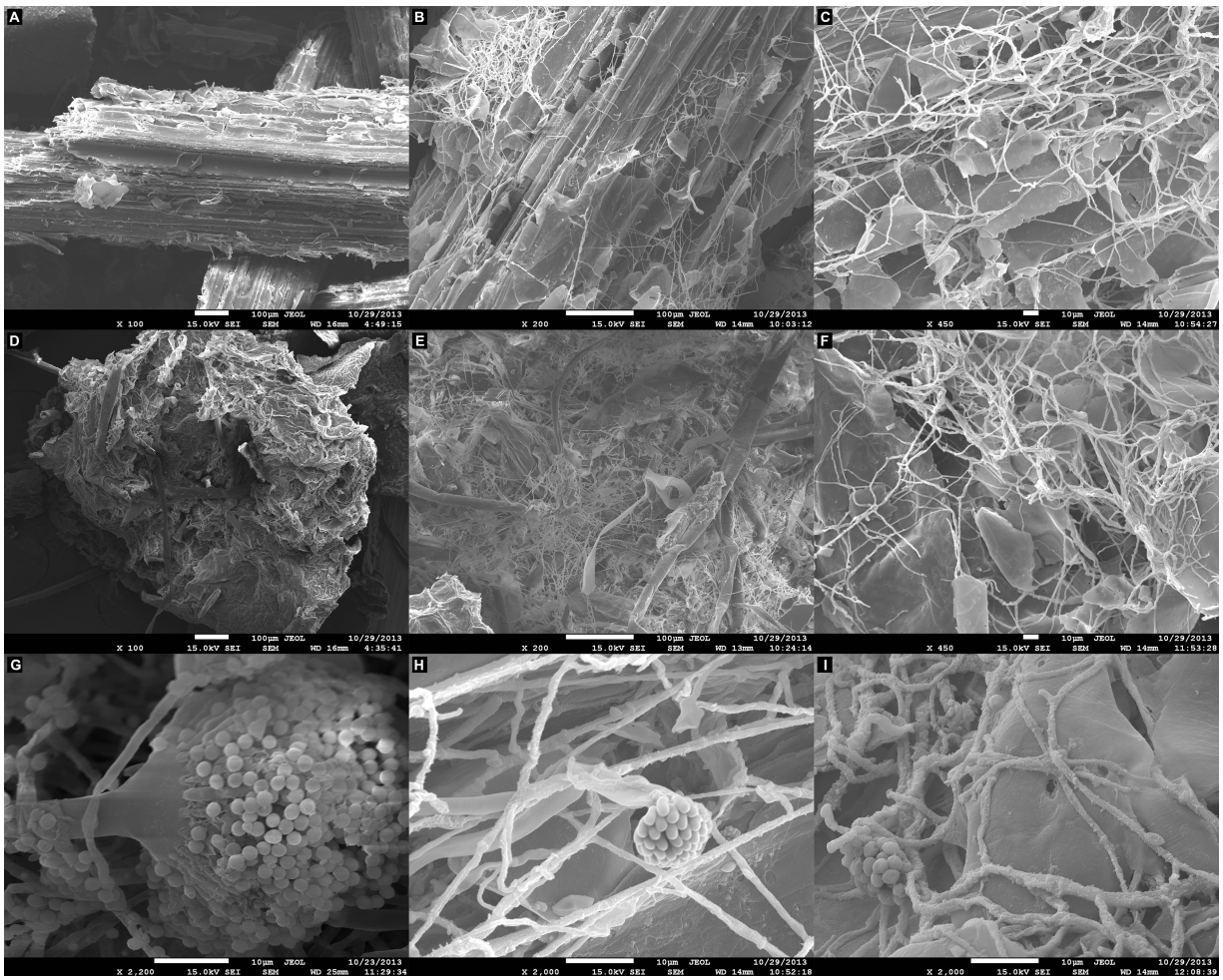


FIGURE 1 | Scanning electron microscopy (SEM) analysis of the degradation of lignocellulosic carbon sources following growth of *Aspergillus terreus* BLU24 in liquid minimal medium supplemented with each carbon source. **(A)** Non-inoculated pre-treated sugarcane bagasse parenchyma in liquid minimal medium culture. **(B)** Inoculated pre-treated sugarcane bagasse in liquid minimal medium culture after 36 h incubation, 200X magnification. **(C)** Inoculated pre-treated sugarcane bagasse in liquid minimal medium culture after 48 h incubation, 450X magnification. **(D)** Non-inoculated pre-treated soybean hull parenchyma in liquid minimal medium culture. **(E)** Inoculated pre-treated soybean hull in liquid minimal medium culture after 36 h incubation, 200X magnification. **(F)** Inoculated pre-treated soybean hull in liquid minimal medium culture after 48 h incubation, 450X magnification. **(G)** *A. terreus* BLU24 conidiophore and conidiospore development, 2200X magnification. **(H)** *A. terreus* BLU24 conidiospores and hyphae on pre-treated sugarcane bagasse in liquid minimal medium culture after 48 h incubation, 2000X magnification. **(I)** *A. terreus* BLU24 conidiospores and hyphae on pre-treated soybean hull in liquid minimal medium culture after 48 h incubation, 2000X magnification.

investigated. Considering the kinetics observed for xylanases on SB and SH, and likely evidence of carbon catabolic repression during the first 48 h of growth on G, 36 and 48 h were selected as time points for RNA extraction and subsequent analysis of the transcriptome of *A. terreus* BLU24 following growth on SB or SH lignocellulosic carbon sources in comparison to G.

Illumina Hiseq 2500 Statistics

Illumina Hiseq 2500 sequencing of cDNA libraries yielded, for bioassay 1, after adapter trimming, a total of 317 million reads for the multiplexed data, totaling 31.73 Gb of data. Similar numbers were observed for bioassay 2, with 335 million reads, and 33.59 Gb of data. Percentages of high-quality sequences (Fastq QC > 30) were similar for all treatments and repeat bioassays, averaging 81.42% after adapter trimming (**Supplementary Table 1**). All Illumina RNA-Seq datasets can be found in the NCBI Sequence Read Archive (SRA) database (BioProject ID PRJNA631730).

Differential Gene Expression

The reference genome for *A. terreus* (NIH 2624, BioProject ID PRJNA15631) contains a total of 10406 gene models. Alignment of the sequence reads generated for each cDNA library revealed accurate mapping to over 8000 gene models throughout, indicative of high-quality mRNA for all the biological treatments and replicates. Annotation data for unmapped sequence reads aligned to the NCBI nr database are summarized in **Supplementary Table 2**. Specifically, for mapped reads, a total

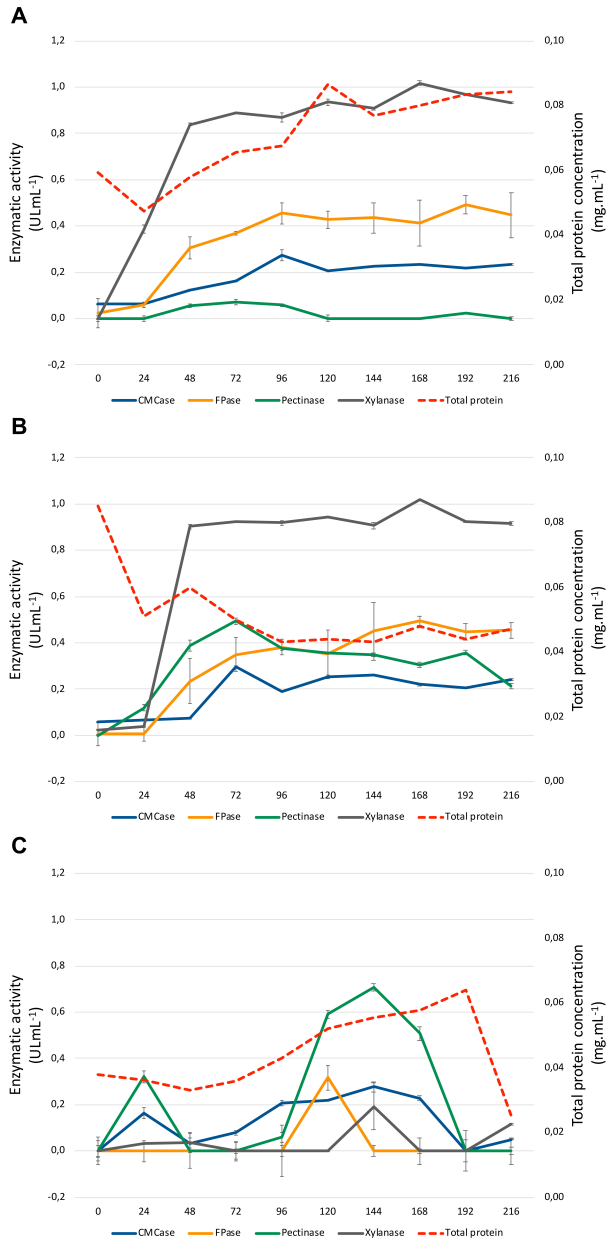


FIGURE 2 | Enzymatic activities of *Aspergillus terreus* BLU24 following growth in liquid minimal medium supplemented with sugarcane (1% w/v) **(A)**, soybean hull (1% w/v) **(B)**, or glucose (1% w/v) **(C)**, as sole carbon source. Enzyme activities in response to xylan, carboxymethylcellulose, pectin, and filter paper substrate were determined through a DNS assay at pH 5.0.

8812 genes were expressed in *A. terreus* BLU24 during growth on G as carbon source, with 8837 expressed following growth on SB and 8799 on SH.

Differential expression in genes was calculated based upon comparison of read counts mapped to a particular reference genome gene model between treatments. Datasets were compared between growth on each lignocellulosic carbon source (SB or SH) versus growth on equivalent glucose controls. Considering analysis parameters of log2 FC at least ≥2-fold and at a probability level of $p \leq 0.01$, considerable modulation of gene expression in the fungus was apparent, with statistically significant DEGs identified according to growth on each lignocellulosic carbon source, as well as growth period. Global analysis revealed similarities in gene expression modulation patterns according to growth period for both SB and SH, with numerous CAZyme-, oxidoreductase-, permease-, transcription factor- and transporter-encoding genes amongst the DEGs (**Figure 3**). In total, 642 DEGs were identified following growth on SB for 36 h, with 372 up-regulated, and 270 down-regulated. At 48 h, 750 DEGs were observed, with 414 up-regulated and 336 down-regulated. In the case of SH as sole carbon source, at 36 h a total of 648 DEGs were identified, with 300 up-regulated and 348 genes down-regulated. At 48 h, 856 DEGs were observed, with 479 up-regulated and 377 down-regulated.

Gene Ontology-Based Representation of Differentially Expressed Genes

Gene ontology categories were assigned to DEGs using FUNC to identify over or under-representation of GO terms, with enrichment values compared with those for all genes present in the *A. terreus* NIH 2624 reference genome. With a total of 13 different enriched categories, considerable enrichment was observed for terms related to lignocellulose hydrolysis for all treatments for *A. terreus* BLU24 on SB and SH when compared to expected distribution in the reference genome (**Supplementary Figure 1**). Significant enriched categories related to the focus of the study where genes were up-regulated on the lignocellulosic carbon sources in comparison with glucose comprised carbohydrate binding, starch binding, catalytic activity, oxidoreductase activity, carboxylic ester hydrolase activity, hydrolase activity, hydrolyzing O-glycosyl compounds, and carbon-sulfur lyase activity. In the case of SB, enrichment was only observed after 48 h, for the categories of carbohydrate binding, starch binding, oxidoreductase activity and carboxylic ester hydrolase activity. SH displayed fold-enrichment for a greater number of the relevant terms than SB, with exclusive gene enrichment for upregulated genes in the categories of catalytic activity, hydrolase activity, hydrolyzing O-glycosyl compounds, and carbon-sulfur lyase activity. On this carbon source, exclusive enrichment was also observed at 36 h for the categories of carbohydrate binding and starch binding. Enriched GO terms for DEGs up-regulated during growth on G as sole carbon source, by contrast, comprised heme binding, cofactor binding, catalytic activity, tryptophan synthase activity and transmembrane transport.

Expression of Cazyme-Encoding Genes

The reference *A. terreus* NIH 2624 reference genome contains a total of 580 predicted CAZyme genes, classified in different families, with 475 potentially involved in carbohydrate hydrolysis. In this study, a total of 89 expressed CAZyme-encoding genes were identified in the *A. terreus* BLU24 transcriptome following growth on SB. A total of 29 genes encoding glycoside hydrolases were expressed and likely

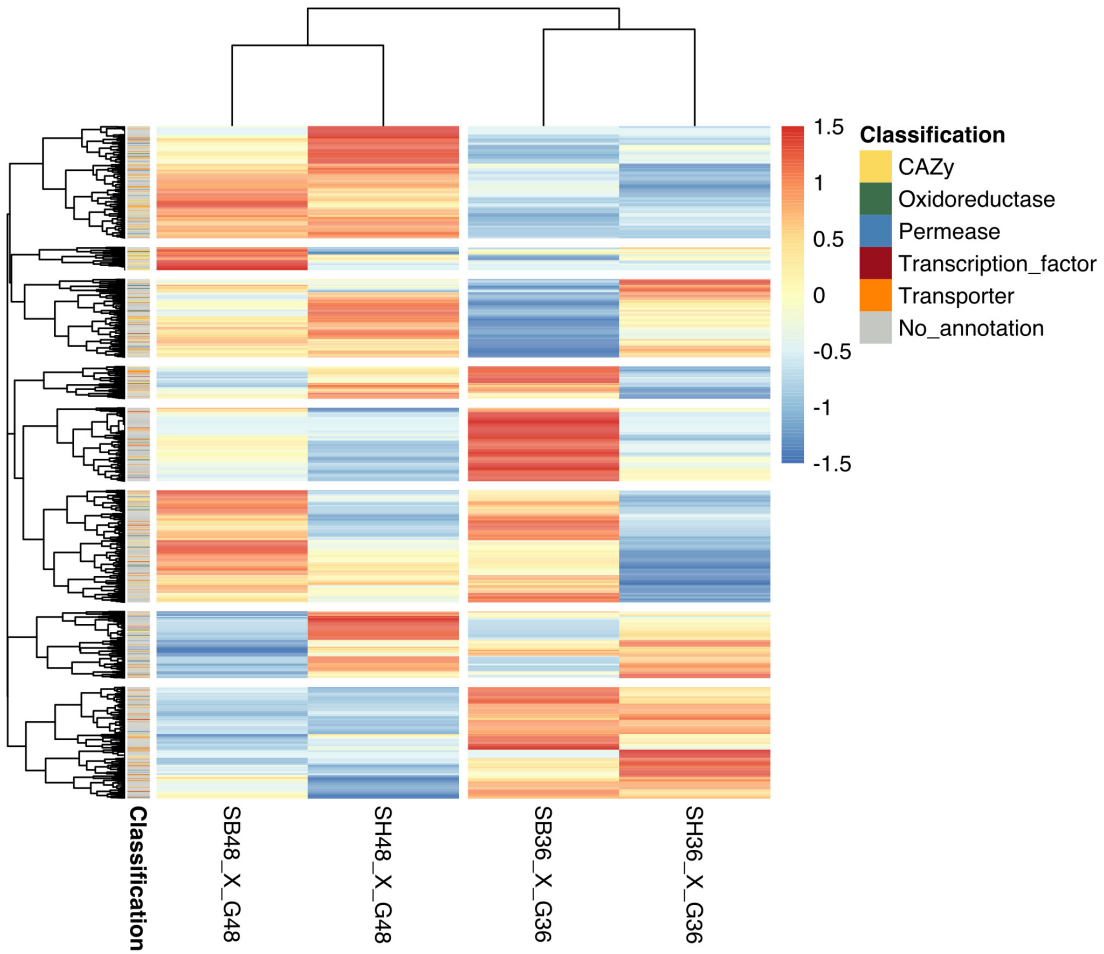


FIGURE 3 | Heatmap depiction of hierarchically clustered groups of *Aspergillus terreus* BLU24 genes according to gene expression modulation following growth on different carbon sources. Gene expression modulation was compared between the growth treatments SB36 and G36, SB48 and G48, SH36 and G36, and SH48 and G48. Statistically significant differentially expressed genes were considered if a log2 fold change (FC) was at least ≥2-fold and at a probability level of $p \leq 0.01$. All FC values below –6 or above 6 were considered as minimum or maximum values, respectively.

involved in the depolymerization of cellulose, together with a further 23 glycoside hydrolase genes likely involved in the degradation of hemicellulose. A further three genes encoding CE family proteins, two genes encoding CBM family proteins and five genes encoding the LPMO AA9 family proteins were also expressed during growth on SB (**Supplementary Table 2**). In the case of SH, for 91 CAZyme-encoding genes, 35 were expressed that encode glycoside hydrolases potentially involved in the depolymerization of cellulose, together with 28 glycoside hydrolase genes likely involved in hemicellulose degradation. Additionally, genes encoding one CE family protein, two PL family proteins, two CBM family proteins, two GT family proteins, six LPMO AA9 family proteins, and three genes encoding GH28, GH77, and GH88 proteins associated with pectin degradation, were also expressed on this carbon source (**Supplementary Table 2**).

Differential expression of CAZyme-encoding genes in treatments on SB or SH in comparison with equivalent growth periods on glucose as carbon source (log2 fold change (FC) at least ≥2-fold and at a probability level of $p \leq 0.01$) is depicted in **Figure 4** and **Supplementary Figure 2**. A total of 77 CUPP-validated CAZyme-encoding genes were observed to be significantly differentially up-regulated on SB, across the two growth time points investigated (**Table 1** and **Figure 5**). These comprised approximately 78% glycoside hydrolases, 8% carbohydrate esterases, 2.5% polysaccharide lyases, and 11.5% auxiliary activities. Analysis of the glycoside hydrolase family revealed significant up-regulation for genes encoding 25 different GH family proteins, with predominance for families GH3, 5, 7, 10, and 43. For SH, a total of 83 CAZyme-encoding genes were also significantly up-regulated in comparison to G. These comprised approximately 80% glycoside hydrolases, 7% carbohydrate esterases, 5% polysaccharide lyases, 7% auxiliary activities (AA), and 1% glycosyltransferases. Similarly, within the glycoside hydrolases, significant up-regulation was observed for genes encoding 26 different GH family proteins, with predominance, as for SB, for families GH3, 5, 10, 31, and 43. Venn analysis of CAZyme-encoding DEGs across treatments

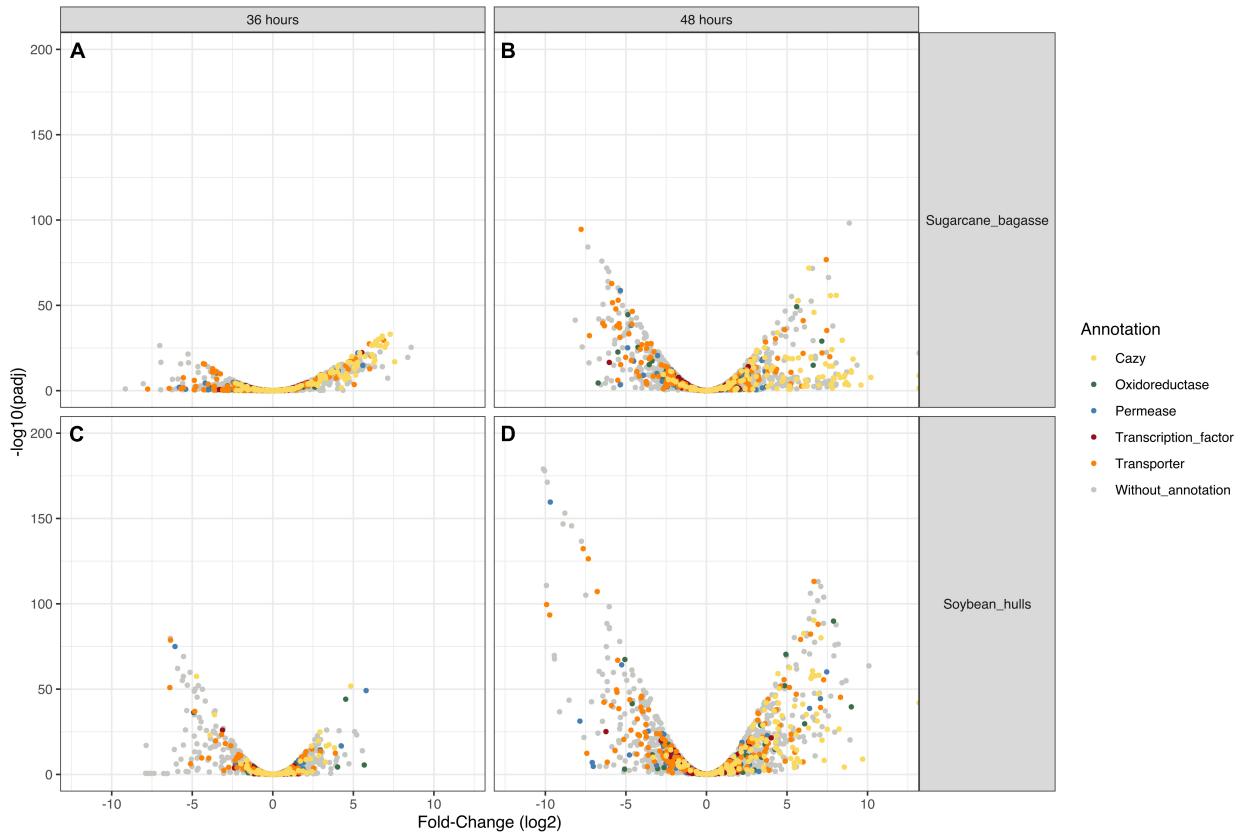


FIGURE 4 | Volcano plot analysis of differential gene expression in *Aspergillus terreus* BLU24 according to carbon source and incubation time period. **(A)** SB36 vs. G36; **(B)** SB48 vs. G48; **(C)** SH36 vs. G36; **(D)** SH48 vs. G48. Genes that were highly modulated, in relation to equivalent treatments with glucose as sole carbon source, appear further to the left and right sides of the plot, with highly significant changes in expression located higher on the plot. Yellow dots: CAZyme-encoding genes (all families); Red dots: genes encoding transcription factors; Green dots: genes encoding permeases; Blue dots: genes encoding oxidoreductases; Gray dots: genes of other function.

SB36, SB48, SH36, and SH48 revealed DEGs exclusively up-regulated on each carbon source (**Figure 6** and **Table 1**), as well as 41 up-regulated DEGs common to all four treatments, likely indicative of conserved function in hydrolysis of lignocellulose in *A. terreus* (**Figure 7**).

Up-regulated CAZyme-encoding DEGs coding for GHs involved in the depolymerization of cellulose comprised genes belonging to the GH families 1, 3, 5, 6, 7, 12, 15, and 71, with exclusive expression for GH5 members in SB (ATEG 04390), and similarly in SH (ATEG 10292). Seven up-regulated genes belonging to the GH families, 7, 12, 15, and 71 were common across carbon sources and time points. Considerable up-regulation (>6-fold) in specific treatments on SB was also observed for genes belonging to GH1 (ATEG 00687), GH5 (ATEG 05003), and GH6 (ATEG 00193, 07493). With regard to hemicellulose degradation, up-regulation was observed in genes in the GH families 3, 5, 10, 11, 27, 31, 35, 38, 39, 43, 51, 62, 67, 74, and 76, as well as CE families 1 and 5. Within this category, CAZyme-encoding DEGs associated with degradation of mannan polysaccharides were observed (GH5: ATEG 02669, 08654; GH38: ATEG 01650; GH76: ATEG 09048), with the latter up-regulated exclusively on SB. DEGs related to Xylan hydrolysis included, for both carbon sources, β-xylosidases (EC 3.2.1.37) (GH3, GH43) (ATEG 09052, 06199, 10193, 00093), xylanases (EC 3.2.1.8) (GH10, 11, 43) (ATEG 08906, 00809, 03410, 04943, 07461, 10072), α-L-arabinofuranosidase (EC 3.2.1.55) (GH43) (ATEG 01292, 07868, 10071, 10379), as well as other xylan-debranching enzymes. Twenty two genes belonging to the GH families 3, 5, 10, 11, 27, 31, 39, 43, 62, 67, and 74, as well as CE families 1 and 5, were common across carbon sources and time points. DEGs related to pectin degradation were observed within the GH families 28, 78, and 105, CE families 8 and 12, and PL families 1 and 4. Exclusive up-regulated genes were observed on SH, namely for a gene encoding an exopolygalacturonase GH28 member (ATEG 07152), for an alpha-L-rhamnosidase GH78 member (ATEG 02922 and a Rhamnogalacturonate lyase B (PL4 family) (ATEG 10327). Two genes belonging to the CE families 8 and 12 were common across carbon sources and time points. Lytic polysaccharide monooxygenases (LPMOs) were also observed amongst the DEGs, with one AA3 gene and five AA9 genes up-regulated on both SB and SH. Exclusive up-regulated AA3-encoding (ATEG 09993), AA7 (ATEG 07698) and AA9-encoding genes (ATEG 10194) were observed on SB.

TABLE 1 | Summary of CAZyme-encoding genes in *Aspergillus terreus* BLU24 with significant increased expression after growth on sugarcane bagasse and soybean hulls in comparison with growth on glucose as carbon source.

					log2 Fold Change ≥2; padj ≤ 0.01			
Class	**Gene ID**	**Enzyme description**	**CAZy Family**	**CAZy subfamily**	**SB36**	**SB48**	**SH36**	**SH48**
Cellulases	ATEG_10194	Endoglucanase	GH61/AA9			5,05		
	ATEG_00687	Beta-glucosidase 1	GH1		6,54	8,05		2,82
	ATEG_03047	Beta-glucosidase	GH3		3,24	6,35		5,59
	ATEG_07419	Beta-D-glucoside glucohydrolase	GH3		2,18	4,41		4,26
	ATEG_09314	Beta-glucosidase	GH3			4,69	3,91	4,5
	ATEG_05003	Exoglucanase I precursor	GH5	GH5_5	6,13	8,97		4,38
	ATEG_04390	Endoglucanase 3 precursor	GH5	GH5_5		4		
	ATEG_10292	Cellulase	GH5	GH5_27			5,97	
	ATEG_00193	Exoglucanase II precursor	GH6		4,62	7,38		2,7
	ATEG_07493	Exoglucanase 2 precursor	GH6		6,59	8,54		3,55
	ATEG_05002	Exoglucanase I precursor	GH7		6,33	8,59		3,58
	ATEG_08705	Endoglucanase EG-1 precursor	GH7		5,4	7,72		
	ATEG_03727	Exoglucanase I precursor	GH7		2,9	5,2	2,43	
	ATEG_08700	Endoglucanase EG-1 precursor	GH7		5,39	7,82	2,75	4,41
	ATEG_07420	Endoglucanase I precursor	GH12		7,28	8,42	4,72	6,03
	ATEG_04375	Glucoamylase precursor	GH15		2,78	4,01	3,26	5,46
	ATEG_07752	Extracellular alpha-1,3-glucanase	GH71		5,71	6,25	4,74	6,55
	ATEG_07790	Endoglucanase B	AA9		4,01	6,98	4,72	5,63
	ATEG_08942	Endoglucanase-4	AA9		5,26	7,76	4,35	5,23
	ATEG_04210	Endo-1,4-beta-glucanase	AA9		6,21	8,18	4,36	6,21
Hemicellulases	ATEG_09052	Exo-1,4-beta-xylosidase	GH3		5,14	6,49	6,32	6,95
	ATEG_09991	Endo-1,4-beta-mannanase	GH5	GH5_7	3,66	7,39	4,35	6,97
	ATEG_02669	Mannanase	GH5	GH5_7	6,27	*	6,27	*
	ATEG_08654	Endo-beta-mannanase	GH5	GH5_7	5,25	10,2	5,07	9,69
	ATEG_08906	Endo-1,4-beta-xylanase	GH10		6,78	7,88	5,03	5,79
	ATEG_00809	Endo-1,4-beta-xylanase A precursor	GH10			9,58	6,96	
	ATEG_03410	Endo-1,4-beta-xylanase precursor	GH10		6,43	8,68	6,66	6,43
	ATEG_04943	Endo-1,4-beta-xylanase B precursor	GH11		4,73	6,04	5,4	5,46
	ATEG_07461	Endo-1,4-beta-xylanase A precursor	GH11		6,79	7,61	6,1	6,74
	ATEG_03427	Alpha-galactosidase	GH27		4,55	6,14	4,23	4,27
	ATEG_08278	Alpha-glucosidase	GH31				3,33	
	ATEG_02528	Glycoside hydrolase	GH31		3,07	3,08	3,25	2,92
	ATEG_02966	Alpha-glucosidase	GH31				3,18	5,56
	ATEG_05177	Alpha-glucosidase	GH31					2,85
	ATEG_06730	Hypothetical protein similar to alpha-glucosidase (593 aa)	GH31			2,77		2,16
	ATEG_00616	Beta-galactosidase	GH35			2,84	2,63	3,58
	ATEG_01650	Mannosidase II	GH38		2,39	3,13		3,06
	ATEG_06199	Xylan 1,4-beta-xylosidase	GH39		3,77	5,33	3,91	4,69
	ATEG_01292	Arabinofuranosidase	GH43	GH43_14	4,01	4,13	4,79	4,2
	ATEG_07817	Endo-arabinase	GH43		3,68	2,99		3,57
	ATEG_10193	Xylosidase	GH43	GH43_36	4,28	5,7	4,73	5,34
	ATEG_03688	Endo-1,5-alpha-L-arabinosidase	GH43	GH43_6	6,39	6,98	6,56	7,1
	ATEG_05083	Xylosidase/arabinosidase	GH43	GH43_14	4,8	6,67	5,54	6,58
	ATEG_10072	Endo-1,4-beta-xylanase XynD	GH43	GH43_29	4,25	6,97	4,94	4,25
	ATEG_00093	Beta-xylosidase	GH43/GH117	GH43_1	3,14	4,76	4,09	4,44
	ATEG_07868	Alpha-L-arabinofuranosidase	GH51			3,06	4,39	4,3
	ATEG_10071	Alpha-L-arabinofuranosidase	GH62		6,21	7,92	7,23	7,17
	ATEG_10379	Alpha-L-arabinofuranosidase precursor	GH62		6,49	7,78	6,24	6,33
	ATEG_06085	Alpha-glucuronidase	GH67		2,54	5,16	4,57	5,51
	ATEG_04708	Endoglucanase C	GH74		6,96	9,1	5,4	6,48

(Continued)

TABLE 1 | Continued

					log2 Fold Change ≥2; padj ≤ 0.01			
Class	**Gene ID**	**Enzyme description**	**CAZy Family**	**CAZy subfamily**	**SB36**	**SB48**	**SH36**	**SH48**
	ATEG_09048	Endo mannanase	GH76		2,25			
	ATEG_09843	Acetyl xylan esterase	CE1		5,09	5,58	5,16	4,59
	ATEG_04709	Acetyl xylan esterase precursor	CE5		7,06	8,29	7,22	6,87
	ATEG_05416	Endoglucanase IV precursor	AA9		6,37	8,81		3,42
Pectinases	ATEG_07152	Exopolygalacturonase	GH28					2,02
	ATEG_04991	Endopolygalacturonase	GH28			7,82	8,45	8,55
	ATEG_02922	Alpha-L-rhamnosidase	GH78				2,05	4,75
	ATEG_07907	Unsaturated rhamnogalacturonan hydrolase	GH105				3,57	5,44
	ATEG_01704	Pectinesterase	CE8		2,41	3,08	3	3,71
	ATEG_03511	Rhamnogalacturonan acetylesterase	CE12		2,75	3,32	2,67	2,82
	ATEG_02193	Rhamnogalacturonate lyase A	PL1	PL1_4		5,15	4,15	5,22
	ATEG_07593	Pectin lyase A	PL1	PL1_4		3,03	3,22	3,15
	ATEG_10327	Rhamnogalacturonate lyase B	PL4	PL4_3			4,91	7,01
Others	ATEG_09566	Conserved hypothetical protein (633 aa)	GH76			2,72		2,73
	ATEG_04135	Hypothetical protein similar to BGLC protein (488 aa)	GH1		2,95		3,94	4,47
	ATEG_04784	Beta-D-galactosidase	GH2		2,48	3,72		2,62
	ATEG_09745	Glycoside hydrolase	GH2			2,39		2,5
	ATEG_08684	conserved hypothetical protein (1120 aa)	GH2			2,17		
	ATEG_04963	Glycoside hydrolase	GH3				5,73	Inf
	ATEG_02724	Predicted protein	GH3		5,08	6,49	5,14	6,19
	ATEG_04069	Hypothetical protein similar to Cel3d (1422 aa)	GH3			2,5		2,16
	ATEG_06617	Hypothetical protein similar to beta-glucosidase (868 aa)	GH3		4,06	3,62	3,45	3,71
	ATEG_07383	Predicted protein	GH3		7,56	8,59	8,45	8,23
	ATEG_09329	Conserved hypothetical protein (840 aa)	GH3		4,9	5,82		
	ATEG_06369	Conserved hypothetical protein (470 aa)	GH5	GH5_22	3,57	5,65	3,88	5,16
	ATEG_07190	Predicted protein	GH10		4,17	7,22	7,03	7,5
	ATEG_00838	Alpha-amylase	GH13	GH13_5	4,49	3,8	3,08	4,43
	ATEG_08279	Alpha-amylase	GH13	GH13_1			2,64	
	ATEG_02515	Hypothetical protein similar to taka-amylase (471 aa)	GH13	GH13_1			5,82	6,52
	ATEG_10103	Alpha-amylase precursor (608 aa)	GH13	GH13_1		3,74		
	ATEG_02160	Conserved hypothetical protein (642 aa)	GH27				4,35	
	ATEG_07446	Hypothetical protein similar to beta-galactosidase (907 aa)	GH35				3,21	4,26
	ATEG_07929	Alpha-galactosidase C precursor	GH36		2,63			
	ATEG_07778	Glycoside hydrolase	GH71		4,54		2,79	4,54
	ATEG_02892	Cell wall glycoside hydrolase YteR	GH105					2,41
	ATEG_01914	Hypothetical protein similar to ferulic acid esterase A	CE1		6,7	7,69	6,18	6,82
	ATEG_03676	Cellulose-binding GDSL lipase/acylhydrolase	CE16		6,15	8,99	6,04	7,03
	ATEG_08610	Predicted protein (454 aa)	PL4	PL4_5			3,63	5,97
	ATEG_09491	GMC oxidoreductase, putative	AA3	AA3_2	2,33	3,44	3,29	2,68
	ATEG_09993	Cellobiose dehydrogenase	AA3	AA3_1	5,37	7,49		
	ATEG_07698	6-hydroxy-D-nicotine oxidase	AA7		2,24	2,12		
	ATEG_05081	Fungal cellulose binding domain	AA9		4,65	6,21	5,18	5,46
	ATEG_00990	Glycosyl transferase	GT32					3,54

**Calculation not possible due to read absence in RNAseq data for corresponding treatment with glucose as carbon source.*

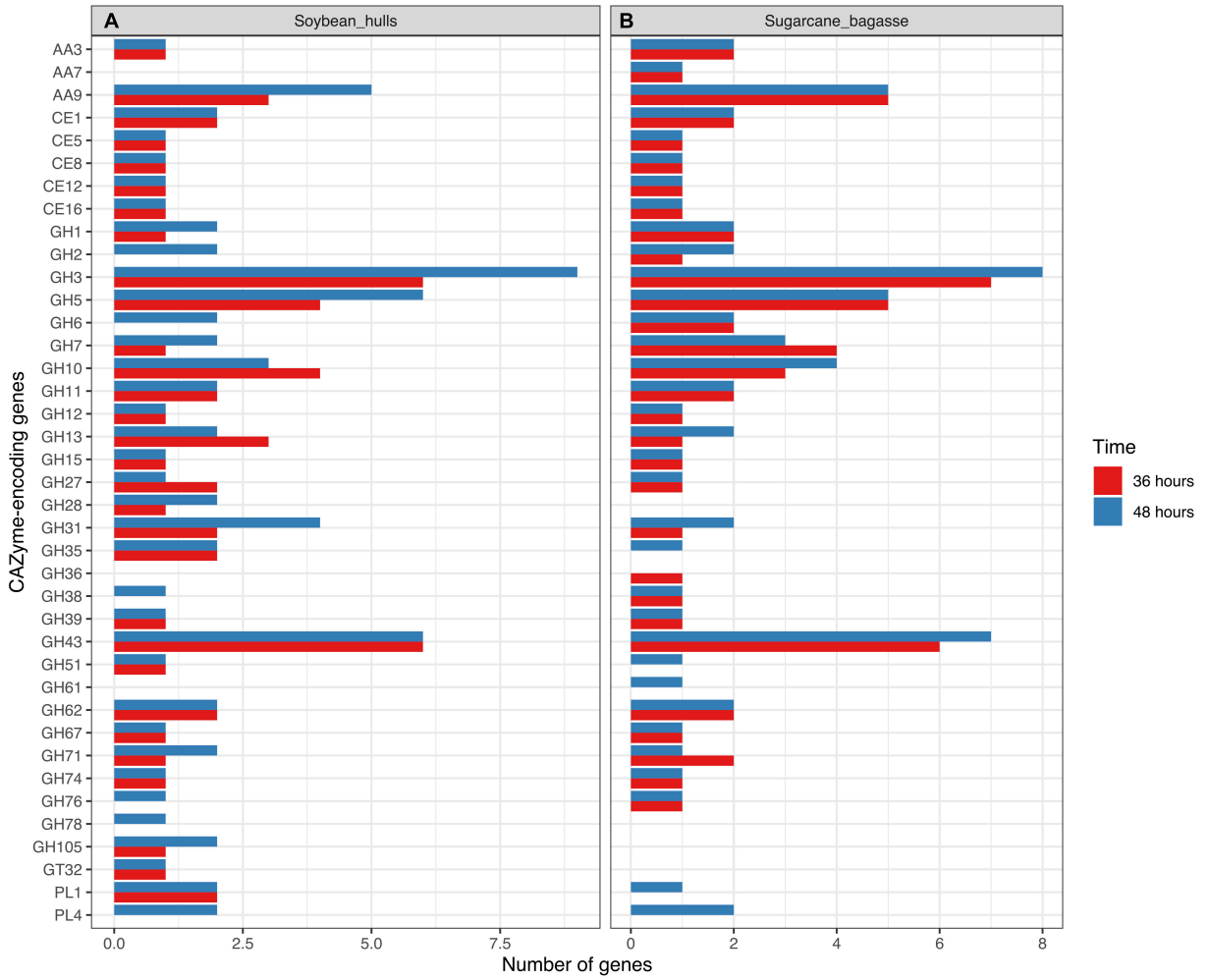


FIGURE 5 | Barplot representation of genes related to cellulose, hemicellulose and pectin depolymerization in transcriptome data for *Aspergillus terreus* BLU24 following 36 and 48 hours growth on soybean hull **(A)** and sugarcane bagasse **(B)** as sole carbon source.

Expression of Transcription Factors

Out of a total of 205 transcription factor-encoding genes expressed during growth on SB and SH, a total of 15 were differentially expressed on SB or SH in comparison with equivalent growth periods on glucose as carbon source (log2 fold change (FC) at least ≥2-fold and at a probability level of $p \leq 0.01$) (**Supplementary Table 3**). On the basis of genome annotation and FUNC-based analysis of Pfam and InterPro entries, these comprised gene orthologs encoding Zn_2Cys_6 transcription factors, C6 transcription factors (including AmyR), CP2 transcription factors, a bZIP transcription factor (AtfA), and a zinc finger Cys_2His_2 transcription factor. Where log2 fold change (FC) was <2-fold and at a probability level of $p \leq 0.01$, a total of 84 displayed modulation during at least one of the four treatments on the two lignocellulosic carbon sources and growth time points. These comprised transcription factors annotated as Zn_2Cys_6 (Btf3, Hfs, tamA, SAGA, jumonji family, GATA-type, FilB, acuM, TPA, RfeB, bZip), zinc finger Cys_2His_2, C6, APSES / Homeobox KN domain, MADS box, heat shock, PHD, CP2, Ran-specific GTPase-activating protein, TFIIH subunit Tfb4, cellobiose response regulator, HLH, bZIP, GATA, homeobox, HTH/APSES/RfeD, CCAAT-binding HapE, TFIIIB complex subunit Brf1, TFIIH basal transcription factor complex p47 subunit, and hypothetical or unnamed protein transcription factor genes.

Expression of Transporters

Analysis of expression of transporter genes across the four treatments during growth on SB and SH, in comparison to glucose, revealed considerable modulation according to carbon source (**Supplementary Table 4** and **Supplementary Figure 3**). From a total of 124 transporter genes with significant differential expression (log2 fold change (FC) at least ≥2-fold and at a probability level of $p \leq 0.01$), 60 were annotated as orthologs belonging to the protein major facilitator superfamily (MFS), in which the majority of sugar transporters are classified. Within these, candidate sugar transporter genes were classified

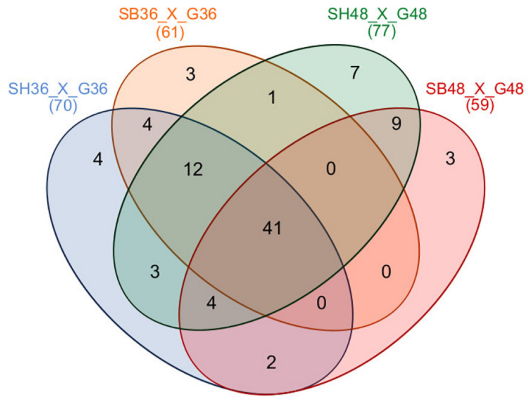


FIGURE 6 | Venn diagram summary of differentially expressed CAZyme-encoding genes for *Aspergillus terreus* BLU24 among treatments SB36, SB48, SH36, and SH48. Differential gene expression was considered significant if relative expression in comparison to equivalent treatments with glucose as sole carbon source showed a log2 fold change (FC) of at least ≥2-fold, at a probability level of $p \leq 0.01$. Overlapping regions of the diagram represent DEGs common to different growth treatments.

as MFS glucose transporters, MFS hexose transporters, MFS monosaccharide transporters, and MFS sugar transporters. A total of 16 potential sugar transporters were significantly up-regulated on both SB and SH during the 48 h incubation period, namely ATEG_07791, ATEG_03475, ATEG_07105, ATEG_04988, ATEG_01556, ATEG_07124, ATEG_08653, ATEG_07144, ATEG_03527, ATEG_02489, ATEG_07114, ATEG_04070, ATEG_07053, ATEG_04137, ATEG_03190, and ATEG_05008.

RT-qPCR Validation of *in silico* RNA-Seq Transcriptome Data

Expression profiles for selected *A. terreus* CAZyme-encoding and other genes significantly up-regulated on SB and SH in comparison to growth on G were examined by RT-qPCR in order to validate expression data based on Illumina RNA-Seq. These genes comprised endo-1,4-beta-xylanases (ATEG_07461, ATEG_03410, and ATEG_00809), an endoglucanase (ATEG_07420), a polygalacturonase (ATEG_04991), a membrane protein (ATEG_02687), and hypothetical or predicted proteins (ATEG_07383, ATEG_04652). Analysis generally revealed similar expression tendencies for both RT-qPCR and RNA-Seq data for each of the genes across the investigated growth treatments. **Supplementary Figure 4** shows log2 Fold Change representations of expression modulation on each lignocellulosic carbon source in comparison to glucose, at each investigated time point. Differences in $\log_2$ FC values may occur due to the different algorithms applied for estimation of

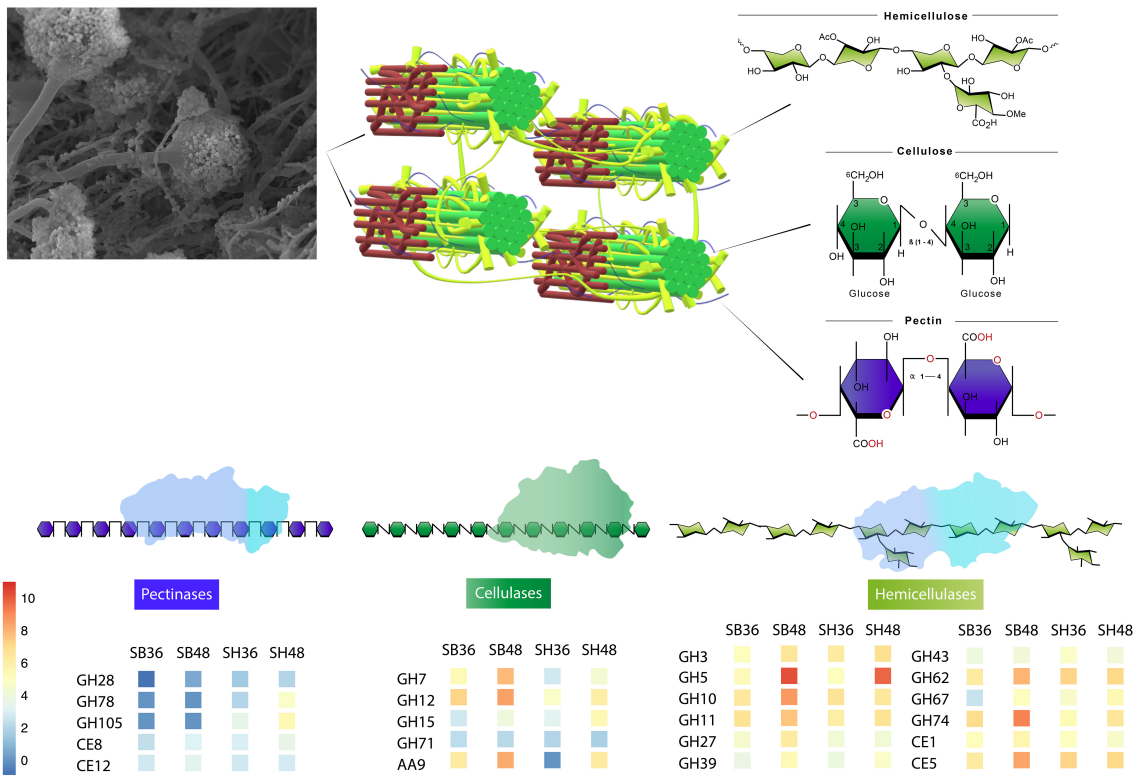


FIGURE 7 | Representative model of conserved CAZyme families in *Aspergillus terreus* BLU24 involved in degradation of lignocellulosic biomasses sugarcane bagasse and soybean hull, based on genes with significant increased expression after growth on both carbon sources and at each evaluated time point, in comparison with equivalent growth on glucose.

fold change with each approach. The primer sequences employed for RT-qPCR analysis of expression of each of the target genes are provided in **Supplementary Table 5**.

DISCUSSION

The potential of biorefinery systems in the 21st Century is considerable, not only in terms of the contribution to global sustainable energy supply in the transition to greenhouse gas neutrality, but also in terms of other bio-products (e.g., chemicals, amino acids, polymers) in a circular economy format that favors socio-economic development (King et al., 2010; Jungmeier et al., 2013; Silva et al., 2017). Countries with substantial agricultural industries, such as Brazil, produce large quantities of lignocellulosic waste, making them strategically placed to contribute to biorefining, either through the export of biomass materials or through in-country development of biorefinery hubs within agricultural zones. Whilst Brazil is already a global example for bioethanol and bioplastic production, optimization of enzymatic conversion of lignocellulose, as in all countries devoted to such white biotechnology, remains an obstacle to economic viability and sustainability, accounting, for example, for almost 30% of production costs for cellulosic bioethanol (Johnson, 2016).

Given the ability of certain filamentous Ascomycete fungi to secrete cellulases and other hydrolytic enzymes into growth media, prospection for efficient producing species has been extensive. In this context, investigation has been considerable into furthering understanding of the molecular mechanisms involved in lignocellulose biodegradation by filamentous fungi. In the case of the genus *Aspergillus*, lignocellulose biodegradation potential has been investigated at the transcriptome level in species such as *A. fumigatus* (Miao et al., 2015; Damasio et al., 2017; de Gouvêa et al., 2018), *A. niger* (de Souza et al., 2011; Delmas et al., 2012; Pullan et al., 2014; van Munster et al., 2014; Brown et al., 2016), *A. nidulans* (de Assis et al., 2015), *A. sydowii* (Cong et al., 2017) and *A. tamarii* (Midorikawa et al., 2018). Typically isolated from soil and decomposing lignocellulosic plant biomass (Reddy and Singh, 2002; He et al., 2004; Chantasingh et al., 2006), ligninolytic enzyme activities have been investigated in *A. terreus*, across diverse biorefinery applications, from bioethanol (Sethi et al., 2017), to lovastatin (Jahromi et al., 2012) and organic acid production (Jiménez-Quero et al., 2017). Analysis of the transcriptome of this fungus has, however, focused only on secondary metabolism related to biogenesis of lovastatin (Palonen et al., 2017), with no investigation, prior to this study, of global gene expression during hydrolysis of lignocellulosic plant biomass.

This analysis of the secretome and transcriptome of the fungus *A. terreus* BLU24 following growth on two important abundant agricultural lignocellulosic biomass residues in Brazil provides important information on CAZyme-encoding genes, transcription factors and transporter genes that are activated during the saccharification of these complex carbon sources. The significant upregulation of glycoside hydrolases, carbohydrate esterases, polysaccharide lyases, and auxiliary activities involved in cellulose, hemicellulose and pectin degradation observed following growth on SB and SH highlight the potential in this strain for lignocellulosic biomass degradation. Given the observed presence of additional gene sequences in *A. terreus* BLU24 that did not map to the reference genome, including, for example, those with identities to genes encoding exo- and endo-1,4-beta-xylanases, beta-glucosidases, 1,4-beta-D-glucan cellobiohydrolases and endo-1,4-beta-glucanases, further characterization of this strain at the genome level is also warranted.

Previous bromatological analyses of different lignocellulosic biomass residues as carbon sourced for induction of hydrolytic enzymes in filamentous fungi reported SH to be composed of 44.4% cellulose, 16.7% hemicellulose and 5.2% lignin. Values were somewhat reduced in SB, with 34.6% cellulose, 16.2% hemicellulose and 5.1% lignin (Siqueira, 2010). As shown in **Figure 1**, SEM observations of the cell walls of SB and SH revealed rapid colonization of these lignocellulosic materials and consistent with the ability of the fungus to secrete appropriate hydrolytic enzymes for biodegradation.

Xylanases (endo-β-1,4-xylanase; EC 3.2.1.8) are responsible for the random hydrolysis of B-(1,4) glycosidic bonds in xylan in plant cell walls. Xylans can comprise up to 35% of the dry weight of annual plants, with these enzymes widely applicable in the saccharification of biomass. Despite the higher cellulose content in the evaluated substrates, xylanase activities on SB and SH were considerably higher than those for CMCases, pectinases and FPases, both initially during the first 48 h and for the duration of the 9-day evaluated time period on each carbon source. In the plant cell wall, the interaction of xylan with cellulose microfibrils increases recalcitrance to cellulose active enzymes. The earlier production of xylanases observed here likely contributes to xylan degradation and subsequent increased substrate accessibility for CMCases and FPases. In previous studies in *A. niger* and *A. oryzae,* a clear relationship was observed between the presence of D-xylose and increased expression of cellulose-encoding genes, regulated by the transcriptional factor XlnR (de Souza et al., 2011; Tani et al., 2014). Here, our results may indicate a similar mechanism occuring in *A. terreus*, where an initial release of D-xylose following activity of xylanases results in subsequent increased secretion of CMCases and FPases for continued cellulose hydrolysis.

Given the previous reports on lignocellulolytic enzyme production in *A. terreus,* secretion of enzymes covering different CAZyme families and subfamilies was expected on the two complex lignocellulosic carbon sources employed in this study. In total, 89 CAZyme-encoding genes were expressed on SB and 91 on SH. These numbers highlight the wide spectrum of hydrolytic enzymes likely secreted and necessary for the hydrolysis of these complex carbon sources.

Global analysis of gene expression on the two lignocellulosic carbon sourced through GO enrichment analysis revealed enrichment for terms associated with carbohydrate degradation. The greater number of terms related to lignocellulose hydrolysis observed on SH rather than on SB suggests a larger arsenal of enzymes necessary for breakdown of this carbon source.

Considerable variation in CAZyme-encoding gene content can be observed across annotated reference genomes for *Aspergillus* species, with species such as *A. aculeatus*, *A. brasiliensis*, *A. carbonarius*, *A. luchuensis*, *A. nidulans*, *A. sydowii*, *A. tubingensis*, *A. versicolor*, and *A. wentii* containing 117, 435, 376, 410, 478, 501, 239, 539, and 387 predicted CAZyme-encoding genes involved in carbohydrate hydrolysis, respectively (Nordberg et al., 2014). In the case of *A. terreus*, from a total of 580 CAZyme-encoding genes predicted in the NIH 2624 reference genome, 475 encode families involved in hydrolysis. These include GHs (281), CBMs (91), AAs (57), CEs (27) and PLs (15).

The numbers of up-regulated CAZyme-encoding DEGs on SB and SH according to family reflect, to some degree, in terms of proportions, the numbers of such genes in the fungal genome. For both SB and SH, GHs were the most common up-regulated CAZyme-encoding DEGs, followed by CEs and AAs. The majority of the expected CAZyme families for a member-species of the genus *Aspergillus* were represented in the differentially expressed genes on both carbon sources (Segato et al., 2014). Whilst many of the CAZyme-encoding DEGs were up-regulated on both carbon sources, clear differences between SB and SH were also apparent, with generally greater numbers of cellulose-encoding genes significantly up-regulated on SB and greater numbers related to pectinase activities on SH.

Many genes encoding different GH family proteins were up-regulated on both SB and SH, with the most abundant within the families GH3, 5, 10, 31, and 43. Similar up-regulation was previously observed in *A. tamarii*, where transcriptome analysis on SB in comparison to glucose revealed CAZyme-encoding DEGs within GH families 3, 5, 31, 43, and 92 (Midorikawa et al., 2018). de Gouvêa et al. (2018) also reported highly expressed genes in *A. fumigatus* on SB belonging to most of these GH families, when compared to growth on fructose. Genome-based prediction of fungal secretomes has also revealed predicted conserved enzyme profiles for families GH 3, 5, 10, and 43 across up to 14 species in taxonomic sections Aspergillus, Candidi, and Flavi for the genus *Aspergillus*, indicating likely conservation of such carbohydrate-active enzymes during speciation (Barrett et al., 2020).

Independently of the biomass employed for cultivation and growth period evaluated, a conserved core set of 41 CAZyme-encoding DEGs was also observed. In **Figure 7**, we present a model of the principal CAZyme families represented in *A. terreus* BLU24 in the transcriptome data during degradation of SB and SH. With regard to cellulose, cellobiohydrolases, endoglucanases, glucosidases, LPMOs, and cellobiose dehydrogenases are required for degradation. Within the up-regulated CAZyme-encoding DEGs, all such cellulases were differentially up-regulated on SB or SH at least at one of the two investigated time points. Of these, genes belonging to the GH families 7, 12, 15, and 71 were up-regulated during all four treatments, indicating potential conservation in function in lignocellulose hydrolysis in this fungus. Previously, gene models for cellulose-degrading CAZymes within the GH family 7 have been described for the *Aspergillus* species *A. clavatus*, *A. nidulans*, *A. niger*, *A. terreus*, *A. fumigatus*, *A. flavus* and *A. oryzae*, and for GH12 for *A. clavatus*, *A. niger*, *A. terreus*, *A. fumigatus*, *A. flavus,* and *A. oryzae* (Segato et al., 2014 and references therein). Genome-based prediction has also been made for the families GH 7 and 12 in 15 species in *Aspergillus* sections Aspergillus, Candidi, and Flavi (Barrett et al., 2020). Greatest numbers of CAZyme-encoding DEGs were observed for those involved in hemicellulose degradation, corroborating the high xylanase enzyme activities observed in this fungus on both SB and SH. Hemicellulose represents a more complex structure than cellulose, with the presence of different monomers (pentoses and hexoses) and ramifications. As a consequence, greater numbers of enzymes are required for hemicellulose hydrolysis, comprising xylanases, arabinofuranosidases, alpha-glucuronidases, xyloglucanases, arabinanases, mannosidases or mannanases, acetyl xylan esterases, feruloyl esterases, xylosidases. Of the up-regulated CAZyme-encoding DEGs, all such enzymes, with the exception of feruloyl esterases, were differentially up-regulated on either SB or SH at least during one of the time points. Genes belonging to the GH families 3, 5, 10, 11, 27, 31, 39, 43, 62, 67, and 74, and CE families 1 and 5 were up-regulated during all four treatments, again indicating potential conservation in lignocellulose hydrolysis. Of these, genome-predicted enzyme profiles have been described previously for xylan-degrading enzymes in GH families 3, 10, 11, 43, 62 and 67, and in CE families 1 and 5, again, in up to 15 species in *Aspergillus* sections Aspergillus, Candidi and Flavi (Barrett et al., 2020). Mannan polysaccharides are the second most abundant hemicelluloses after xylan, with their conversion of considerable interest for the biorefinery industry (Soni and Kango, 2013). With previous purification and characterization in *A. terreus* (Soni et al., 2016), here we observed four CAZyme-encoding DEGs associated with degradation of these polysaccharides. Whilst the pectin content in SB and SH is considerably lower than for hollocellulose (de Souza Moreira et al., 2013; Li et al., 2017), evidence was also observed for pectin degradation on both, with considerable exclusive up-regulation of genes on SH likely reflecting differences in pectin content between the different carbon sources. Up-regulation of genes related to pectinolytic activity have also been reported previously on SB for *A. niger* (PL4 family), *A. fumigatus* (PL4 family) and for *A. tamarii* (PL1, PL9, GH28 families) (Borin et al., 2017; de Gouvêa et al., 2018; Midorikawa et al., 2018). Genome-predicted enzyme profiles in *Aspergillus* sections Aspergillus, Candidi and Flavi indicate the presence of pectin-degrading enzymes in the GH families 28, 43, 53, 78, and 142, in the CE family 8, and in the PL families 1, 3, 4, and 26, again with considerable conservation in up to 15 species (Barrett et al., 2020). Here, up-regulated CAZyme-encoding DEGs were observed for members of GH families 28, 78, and 105, for PL families 1 and 4, and for CE families 8 and 12, with the latter two up-regulated across all four growth treatments. As pectin also acts as a barrier that limits access of cellulase to their substrates, the secretion of pectin active enzymes might guarantee more efficient substrate degradation by *A. terreus*.

Lytic polysaccharide mono-oxygenases, which are classified into CAZy auxiliary activities families AA9-AA11 and AA13-AA16, are copper-dependent enzymes that also perform important roles in lignocellulose degradation, reducing

recalcitrance of less accessible regions of the plant cell wall through oxidative cleavage of glycosidic bonds, facilitating cellulose recognition by cellulases and boosting efficiency in hydrolysis of biomass (Monclaro and Filho, 2017). These enzymes can also depolymerize hemicellulosic polymers (Couturier et al., 2018), thus providing an overall enhancement to the polysaccharide conversion process, working in synergy with glycoside hydrolases. With AA9, AA11, AA13, AA14, and AA16 exclusive to fungal genomes, multiple genes encoding LPMOs appear to be common in fungal genomes, particularly in Ascomycetes and Basidiomycetes (Kracher et al., 2016). With previous reports of LPMOs in transcriptome and/or secretomes of *A. niger*, *A. fumigatus, and A. tamarii* following growth on sugarcane bagasse (Borin et al., 2015, 2017; de Gouvêa et al., 2018; Midorikawa et al., 2018), and recent characterization of seven AA9s in the secretome of *A. terreus* on soybean spent flakes (Pierce et al., 2017), investigation here also highlights their involvement in the degradation of the SB and SH, with one AA3 gene and five LPMO AA9 genes up-regulated on both SB and SH lignocellulosic carbon sources, as well as exclusive AA3, AA7 and AA9-encoding genes on SB. Such data highlights potential conserved functional roles in biomass degradation, as reported in other fungi (Isaken et al., 2014), as well as evidence for substrate specific activity, as documented recently for AA9 LPMOs (Jagadeeswaran et al., 2016; Hüttner et al., 2019). These advances in our knowledge of the LPMO enzyme repertoire in *A. terreus* are relevant for improving efficiency in lignocellulosic biomass degradation.

Transcription factors play major roles in the coordination of gene expression in the cell. To date, numerous transcription factors in different fungi have been reported to be involved in plant biomass degradation, modulating the expression of plant cell wall degrading enzymes (Wu et al., 2020). Activation of such transcription factors likely occurs following hyphal contact with low molecular weight mono-or disaccharide compound derivatives of complex polymers (Benocci et al., 2017). The majority that have been characterized to date belong to the zinc binuclear cluster family, with most positive expression regulators belonging to the Zn_2Cys_6 class (Todd et al., 2014). In *Aspergillus* spp. these include XlnR, a key regulator of (hemi)cellulose degradation, together with AraR, involved in L-Arabinose utilization, Clr-A and Clr-B (cellulose degradation), AmyR (starch degradation), MalR (maltose utilization), ClbR (cellobiose utilization), RhaR (pectin deconstruction / L-Rhamnose utilization), GaaR (pectin deconstruction/galacturonic acid utilization), InuR (inulin utilization), and GalX and GalR (D-galactose utilization). Other important transcription factor classes controlling gene expression in *Aspergillus* spp. related to biomass degradation comprise CreA (carbon catabolite repression), within the Cys2His2 class and the HAP complex (Cazy regulation), within the CBF class (Benocci et al., 2017 and references therein). Whilst mechanisms for regulation of expression of sugar transporter genes during sugar utilization in fungi is only partially understood at present, a number of the above transcription factors, such as XlnR, CreA, ClrA/ClrB, GalX and AraR, are also known to regulate sugar transporter genes in model *Aspergillus* species (e. g. Vankuyk et al., 2004; Andersen et al., 2008; de Souza Moreira et al., 2013). Our data revealed considerable numbers of transcription factor genes with significant differential expression on SB and/or SH at a probability level of $p \leq 0.01$, with 15 at log2 fold change (FC) at least $\geq$2-fold and 84 at log2 fold change (FC) <2-fold. Those with significant differential expression and previously reported as involved in plant biomass utilization in ascomycete fungi included AmyR (starch degradation), Ctf1B (cutinase induction / cellulase, xylanase production) and AreB (nitrogen metabolite repression). With many classified, based on the available genome annotation data, as members of the important Zn_2Cys_6 class, further work is warranted into accurate prediction, annotation and function validation in *A. terreus* BLU24.

The majority of sugar transporters belong to the sugar porter family within the major facilitator superfamily. Sugar transporters are essential proteins in lignocellulolytic fungi that enable the uptake of mono- and short oligosaccharide products of extracellular enzymatic digestion of lignocellulose into the fungal cell for growth and metabolism. Although likely to be abundant in fungal genomes, especially in those species adapted to lignocellulose decomposition, their comprehensive investigation has been restricted mostly to model organisms such as *Saccharomyces cerevisiae* (Wieczorke et al., 1999). Within the genus *Aspergillus*, sugar transporter genes have recently been characterized in species such as *A. niger*, with 86 putative genes identified *in silico*, based on proteome and transcriptome analysis (Peng et al., 2018), as well as *A. nidulans*, with 357 proteins in the major facilitator superfamily predicted in the annotated genome, albeit without validation of roles in sugar transport for the majority (de Gouvêa et al., 2018). Lignocellulose contains both hexose (glucose) and pentose (mostly xylose) sugars which can be released by the enzymatic action of ligninolytic fungi. Biorefinery of lignocellulose is limited in part by the inability of many fermenting microorganisms to metabolize sugars other than glucose (Young et al., 2010). The characterization of additional sugar transporters, therefore, such as those for xylose and cellobiose transport, which are modulated on lignocellulose, may therefore enable internalization of these additional sugars to overcome such limitations. Transformation of *S. cerevisiae* with cellobiose and xylose transporters from *A. nidulans*, for example, has highlighted the potential in the use of such genes from *Aspergillus* sp. for genetic engineering of yeasts for complete fermentation of all sugars in lignocellulose for conversion to bioethanol (dos Reis et al., 2016). Prior to this study, there have been no reports on sugar transporter modulation in *A. terreus* during lignocellulose breakdown. Here we identified 16 potential sugar transporters within the major facilitator superfamily that were up-regulated on both SB and SH during the 48 h incubation period, indicating the likely release of sufficient fermentable sugars such as glucose, xylose or cellobiose to result in induction of their expression during hydrolysis of these complex carbon sources. Such genes offer potential in genetic engineering to optimize non-glucose pentose sugar transport in fermenting organisms.

Lignocellulosic plant biomass saccharification in biorefinery applications using thermostable enzymes offers advantages in terms of efficiency, with increased catalytic activity in hydrolysis,

solubilization of lignocelluloses, lower contamination, lower cooling requirements and increased flexibility in fermentation process design (Sonnleitner and Fiechter, 1983; Sharma et al., 2014). Although numerous fungal genera contain species with the ability to secrete enzymes stable at temperatures above 60 °C, such as *Sporotrichum*, *Chaetomium*, and *Humicola*, amongst others, growth rates of these mesophilic fungi on lignocellulosic substrates are typically low, with cellulases produced in only low titers. Investigation of thermophilic fungi, by contrast, has reported faster growth rates and more rapid cellulose degradation, even at lower pH values (Turner et al., 2007). Cellulase activity in *A. terreus* has been shown to be acidothermophilic (e.g., Gao et al., 2008), with optimal activities also reported between 50 and 60°C, at pH 4.0–6.0 (Sharma et al., 2014). Similarly, high xylanolytic activity has been observed over a wide pH range from 3.0 to 10.0 (Chantasingh et al., 2006), appropriate for employment in bio-bleaching for paper pulp (Filho, 1998). With LPMOs also reported to be stable under low pH conditions (Agrawal et al., 2020), as well as mannanases (Soni et al., 2016), such thermostable properties of enzymes from *A. terreus* are therefore advantageous for industrial biorefinery applications of lignocellulose saccharification.

This first investigation of the transcriptome of *A. terreus* following growth on two abundant lignocellulosic biomasses from abundant commodity crops provides important functional genomics information on the network of CAZyme encoding genes, transcription factors and sugar transporter genes involved in the enzyme hydrolysis of these complex carbon sources. This advance in our understanding of the biology of the fungus is applicable in the genetic improvement of both promising fungi for lignocellulosic carbon hydrolysis and yeasts involved in hexose and pentose sugar fermentation steps in biorefineries.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

RM, EN, and EF planned the experiments. GM and CC performed the bioassays, enzyme analyses, RNA and cDNA preparation, and sequence data analysis. RT, MC, GA, OS-J, and PG participated in sequence data analyses and in editing of the manuscript. RM conceived the study, participated in bioassays, RNA preparation for cDNA library construction, sequence data analysis, and drafted the manuscript. All authors contributed to read and approved the final version of the manuscript.

FUNDING

This work was funded by the Fundação de Amparo à Pesquisa do Distrito Federal (FAPDF) (projects 193.000.584/2009 and 193.001195/2016). CC, GM, and GA were supported by scholarships from CAPES (project finance code 001). EF, EN, and RM were supported by fellowships from the CNPq.

ACKNOWLEDGMENTS

The authors thank the reviewers for their useful comments on the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.564527/full#supplementary-material

Supplementary Figure 1 | Selected gene ontology (GO) enrichment analysis data for differentially expressed genes in *Aspergillus terreus* BLU24 after 36 and 48 h growth on sugarcane bagasse (SB) and soybean hulls (SH) as sole carbon source compared with glucose-supplemented treatments. Significant differentially expressed genes were considered if a log2 fold change (FC) was at least ≥2-fold and at a probability level of $p \leq 0.01$. Enrichment values were compared with those occurring in the *A. terreus* NIH 2624 reference genome sequence.

Supplementary Figure 2 | Heatmap depiction of hierarchically clustered groups of *Aspergillus terreus* BLU24 CAZY genes according to gene expression modulation following growth on different carbon sources. Gene expression modulation was compared between the growth treatments SB36 and G36, SB48 and G48, SH36 and G36, and SH48 and G48. Statistically significant differentially expressed genes were considered if a log2 fold change (FC) was at least ≥2-fold and at a probability level of $p \leq 0.01$. All FC values below −6 or above 6 were considered as minimum or maximum values, respectively.

Supplementary Figure 3 | Heatmap depiction of hierarchically clustered groups of *Aspergillus terreus* BLU24 transporter genes according to gene expression modulation following growth on different carbon sources. Gene expression modulation was compared between the growth treatments SB36 and G36, SB48 and G48, SH36 and G36, and SH48 and G48. Statistically significant differentially expressed genes were considered if a log2 fold change (FC) was at least ≥2-fold and at a probability level of $p \leq 0.01$. All FC values below −6 or above 6 were considered as minimum or maximum values, respectively.

Supplementary Figure 4 | RT-qPCR validation of differential gene expression profiles based on RNA-Seq for selected genes in *Aspergillus terreus* BLU24. Standard error values (black bars) were calculated based on data generated from analysis of three biological replicates per treatment and three technical amplification replicates.

Supplementary Table 1 | Illumina RNA-Seq data sequence statistics for each *Aspergillus terreus* BLU24 cDNA library.

Supplementary Table 2 | Global gene expression data for each *Aspergillus terreus* BLU24 cDNA library following growth on SB and SH in comparison to equivalent growth on glucose as carbon source, together with annotation data for unmapped sequence reads aligned to the NCBI nr database.

Supplementary Table 3 | Transcription factor genes in *Aspergillus terreus* BLU24 differentially expressed on SB and SH in comparison to equivalent growth treatments with glucose as carbon source.

Supplementary Table 4 | Transporter genes in *Aspergillus terreus* BLU24 differentially expressed on SB and SH in comparison to equivalent growth treatments with glucose as carbon source.

Supplementary Table 5 | Primer sequences for RT-qPCR validation of differential gene expression of selected target genes.

REFERENCES

Agrawal, D., Kaur, B., Kaur Brar, K., and Chadha, B. S. (2020). An innovative approach of priming lignocellulosics with lytic polysaccharide mono-oxygenases prior to saccharification with glycosyl hydrolases can economize second generation ethanol process. *Bioresour. Technol.* 308:123257. doi: 10.1016/j.biortech.2020.123257

Albersheim, P., Darvill, A., Roberts, K., Sederoff, R., and Staehelin, A. (2010). *Plant Cell Walls*. New York, NY: Garland Science.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389

Alvira, P., Tomás-Pejó, E., Ballesteros, M., and Negro, M. J. (2010). Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: a review. *Bioresour. Technol.* 101, 4851–4861. doi: 10.1016/j.biortech.2009.11.093

Amore, A., Giacobbe, S., and Faraco, V. (2013). Regulation of cellulase and hemicellulase gene expression in fungi. *Curr. Genomics* 14, 230–249. doi: 10.2174/1389202911314040002

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* 11:R106. doi: 10.1186/gb-2010-11-10-r106

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq - a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166661. doi: 10.1093/bioinformatics/btu638

Andersen, M. R., Vongsangnak, W., Panagiotou, G., Salazar, M. P., Lehmann, L., and Nielsen, J. (2008). A trispecies *Aspergillus* microarray: comparative transcriptomics of three *Aspergillus* species. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4387–4392. doi: 10.1073/pnas.0709964105

Arnaud, M. B., Cerqueira, G. C., Inglis, D. O., Skrzypek, M. S., Binkley, J., Chibucos, M. C., et al. (2012). The Aspergillus Genome Database (AspGD): recent developments in comprehensive multispecies curation, comparative genomics and community resources. *Nucleic Acids Res.* 40, D653–D659. doi: 10.1093/nar/gkr875

Aronesty, E. (2011). *ea-utils: Command-line Tools for Processing Biological Sequencing Data*. Available online at: https://github.com/ExpressionAnalysis/ea-utils (accessed February 3, 2020).

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. *Gene Ontol. Consort. Nat. Genet.* 25, 25–29.

Barrett, K., Hunt, C. J., Lange, L., and Meyer, A. S. (2020). Conserved unique peptide patterns (CUPP) online platform: peptide-based functional annotation of carbohydrate active enzymes. *Nucleic Acids Res.* 48, W110–W115. doi: 10.1093/nar/gkaa375

Barrett, K., and Lange, L. (2019). Peptide-based classification and functional annotation of carbohydrate-active enzymes by conserved unique peptide patterns (CUPP). *Biotechnol. Biofuels* 12, 102–123.

Benocci, T., Aguilar-Pontes, M. V., Zhou, M., Seiboth, B., and de Vries, R. P. (2017). Regulators of plant biomass degradation in ascomycetous fungi. *Biotechnol. Biofuels* 10:152. doi: 10.1186/s13068-017-0841-x

Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. *Annu. Rev. Plant Biol.* 54, 519–546.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Borin, G. P., Sanchez, C. C., de Souza, A. P., de Santana, E. S., de Souza, A. T., Leme, A. F. P., et al. (2015). Comparative secretome analysis of *Trichoderma reesei* and *Aspergillus niger* during growth on sugarcane biomass. *PLoS One* 10:e0129275. doi: 10.1371/journal.pone.0129275

Borin, G. P., Sanchez, C. C., Santana, E. S., Zanini, G. K., dos Santos, R. A. C., Pontes, A. O., et al. (2017). Comparative transcriptome analysis reveals different strategies for degradation of steam-exploded sugarcane bagasse by *Aspergillus niger* and *Trichoderma reesei*. *BMC Genomics* 18:501. doi: 10.1186/s12864-017-3857-5

Brasileiro, A. C. M., and Carneiro, V. T. D. C. (2015). *Manual de Transformação Genética de Plantas*, 2nd Edn. Brasília: Embrapa.

Brown, N. A., Ries, L. N., Reis, T. F., Rajendran, R., Correa Dos, Santos, R. A., et al. (2016). RNAseq reveals hydrophobins that are involved in the adaptation of *Aspergillus nidulans* to lignocellulose. *Biotechnol. Biofuels* 9:145. doi: 10.1186/s13068-016-0558-2

Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176

Buckeridge, M. S., and Souza, A. P. (2014). Breaking the "Glycomic Code" of cell wall polysaccharides may improve second-generation bioenergy production from biomass. *Bioenerg. Res.* 7, 1065–1073. doi: 10.1007/s12155-014-9460-6

Caffall, K. H., and Mohnen, D. (2009). The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydr. Res.* 344, 1879–1900. doi: 10.1016/j.carres.2009.05.021

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009). The carbohydrate-active enzymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* 37, D233–D238. doi: 10.1093/nar/gkn663

Chantasingh, D., Pootanakit, K., Champreda, V., Kanokratana, P., and Eurwilaichitr, L. (2006). Cloning, expression, and characterization of a xylanase 10 from *Aspergillus terreus* (BCC129) in *Pichia pastoris*. *Protein Expr. Purif.* 46, 143–149. doi: 10.1016/j.pep.2005.09.013

Cheng, J., and Zhu, M. (2013). A novel anaerobic co-culture system for bio-hydrogen production from sugarcane bagasse. *Bioresour. Technol.* 144, 623–631. doi: 10.1016/j.biortech.2013.07.018

Cong, B., Wang, N., Liu, S., Liu, F., Yin, X., and Shen, J. (2017). Isolation, characterization and transcriptome analysis of a novel Antarctic *Aspergillus sydowii* strain MS-19 as a potential lignocellulosic enzyme source. *BMC Microbiol.* 17:129. doi: 10.1186/s12866-017-1028-0

Couturier, M., Ladevèze, S., Sulzenbacher, G., Ciano, L., Fanuel, M., et al. (2018). Lytic xylan oxidases from wood-decay fungi unlock biomass degradation. *Nat. Chem. Biol.* 14, 306–310. doi: 10.1038/nchembio.2558

Cragg, S. M., Beckham, G. T., Bruce, N. C., Bugg, T. D., Distel, D. L., Dupree, P., et al. (2015). Lignocellulose degradation mechanisms across the tree of life. *Curr. Opin. Chem. Biol.* 29, 108–119. doi: 10.1016/j.cbpa.2015.10.018

Dadheech, T., Jakhesara, S., Chauhan, P. S., Pandit, R., Hinsu, A., Kunjadiya, A., et al. (2019). Draft genome analysis of lignocellulolytic enzymes producing *Aspergillus terreus* with structural insight of β-glucosidases through molecular docking approach. *Int. J. Biol. Macromol.* 125, 181–190. doi: 10.1016/j.ijbiomac.2018.12.020

Damasio, A. R., Rubio, M. V., Gonçalves, T. A., Persinoti, G. F., Segato, F., Prade, R. A., et al. (2017). Xyloglucan breakdown by endo-xyloglucanase family 74 from *Aspergillus fumigatus*. *Appl. Microbiol. Biotechnol.* 101, 2893–2903. doi: 10.1007/s00253-016-8014-6

de Assis, L. J., Ries, L. N., Savoldi, M., Dos Reis, T. F., Brown, N. A., and Goldman, G. H. (2015). *Aspergillus nidulans* protein kinase A plays an important role in cellulase production. *Biotechnol. Biofuels* 8:213. doi: 10.1186/s13068-015-0401-1

de Gouvêa, P. F., Bernardi, A. V., Gerolamo, L. E., de Souza Santos, E., Riaño-Pachón, D. M., Uyemura, S. A., et al. (2018). Transcriptome and secretome analysis of *Aspergillus fumigatus* in the presence of sugarcane bagasse. *BMC Genomics* 19:232. doi: 10.1186/s12864-018-4627-8

de Jong, E., and Gosselink, R. J. A. (2014). "Chapter 17 - Lignocellulose-based chemical products," in *Bioenergy Research: Advances and Applications*, eds V. K. Gupta, C. P. Kubicek, J. Saddler, F. Xu, and M. G. Tuohy (Amsterdam: Elsevier), 277–313. doi: 10.1016/b978-0-444-59561-4.00017-6

de Souza, W. R., de Gouvêa, P. F., Savoldi, M., Malavazi, I., de Souza Bernardes, L. A., Goldman, M. H. S., et al. (2011). Transcriptome analysis of *Aspergillus niger* grown on sugarcane bagasse. *Biotechnol. Biofuels* 4:40. doi: 10.1186/1754-6834-4-40

de Souza Moreira, L. R., de Carvalho Campos, M., de Siqueira, P. H., Silva, L. P., Ricart, C. A., Martins, P. A., et al. (2013). Two β-xylanases from *Aspergillus terreus*: characterization and influence of phenolic compounds on xylanase activity. *Fungal Genet. Biol.* 60, 46–52. doi: 10.1016/j.fgb.2013.07.006

Delmas, S., Pullan, S. T., Gaddipati, S., Kokolski, M., Malla, S., Blythe, M. J., et al. (2012). Uncovering the genome-wide transcriptional responses of the filamentous fungus *Aspergillus niger* to lignocellulose using RNA sequencing. *PLoS Genet.* 8:e1002875. doi: 10.1371/journal.pgen.1002875

dos Reis, T. F., de Lima, P. B., Parachin, N. S., Mingossi, F. B., Oliveira, J. V. C., Ries, L. N. et al. (2016). Identification and characterization of putative xylose and cellobiose transporters in *Aspergillus nidulans*. *Biotechnol. Biofuels* 9:204. doi: 10.1186/s13068-016-0611-1

Filho, E. X. F. (1998). "Hemicellulase and biotechnology," in *Recent Research Development in Microbiology*, ed. S. G. Pandalai (Trivandrum: Research Signpost), 165–176.

Gao, J., Weng, H., Xi, Y., Zhu, D., and Han, S. (2008). Purification and characterization of a novel endo-*b*-1,4-glucanase from the thermoacidophilic *Aspergillus terreus*. *Biotechnol. Lett.* 30, 323–327. doi: 10.1007/s10529-007-9536-x

Gawkowska, D., Cybulska, J., and Zdunek, A. (2018). Structure-related gelling of pectins and linking with other natural compounds: a review. *Polymers* 10:762. doi: 10.3390/polym10070762

Ghose, T. K. (1987). Measurement of cellulase activities. *Pure Appl. Chem.* 59, 257–268. doi: 10.1351/pac198759020257

Glass, N. L., and Donaldson, G. C. (1995). Development of primer sets designed for use with the PCR to amplify conserved genes from filamentous ascomycetes. *Appl. Environ. Microbiol.* 61, 1323–1330. doi: 10.1128/aem.61.4.1323-1330.1995

Glass, N. L., Schmoll, M., Cate, J. H. D., and Coradetti, S. (2013). Plant cell wall deconstruction by ascomycete fungi. *Annu. Rev. Microbiol.* 67, 477–498. doi: 10.1146/annurev-micro-092611-150044

Grigoriev, I. V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R., et al. (2014). MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res.* 42, D699–D704. doi: 10.1093/nar/gkt1183

He, J., Wijeratne, E. M. K., Bashyal, B. P., Zhan, J., Seliga, C. J., Liu, M. X., et al. (2004). Cytotoxic and other metabolites of *Aspergillus* inhabiting the rhizosphere of Sonoran Desert plants. *J. Nat. Prod.* 67, 1985–1991. doi: 10.1021/np040139d

Hong, S. B., Cho, H. S., Shin, H. D., Frisvad, J. C., and Samson, R. A. (2006). Novel Neosartorya species isolated from soil in Korea. *Int. J. Syst. Evol. Microbiol.* 56, 477–486. doi: 10.1099/ijs.0.63980-0

Hori, C., Gaskell, J., Igarashi, K., Kersten, P., Mozuch, M., Samejima, M., et al. (2014). Temporal alterations in the secretome of the selective ligninolytic fungus *Ceriporiopsis subvermispora* during growth on aspen wood reveal this organism's strategy for degrading lignocellulose. *Appl. Environ. Microbiol.* 80, 2062–2070. doi: 10.1128/AEM.03652-13

Horta, M. A. C., Vicentini, R., Delabona, P. D. S., Laborda, P., Crucello, A., Freitas, S., et al. (2014). Transcriptome profile of *Trichoderma harzianum* IOC-3844 induced by sugarcane bagasse. *PLoS One* 9:e88689. doi: 10.1371/journal.pone.0088689

Hüttner, S., Várnai, A., Petroviæ, D. M., Bach, C. X., Kim Anh, D. T., Thanh, V. N., et al. (2019). Specific xylan activity revealed for AA9 lytic polysaccharide monooxygenases of the thermophilic fungus *Malbranchea cinnamomea* by functional characterization. *Appl. Environ. Microbiol.* 85, e01408–e01419.

Isaken, T., Westereng, B., Aachmann, F. L., Agger, J. W., Kracher, D., Kittle, R., et al. (2014). A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J. Biol. Chem.* 289, 2632–2642. doi: 10.1074/jbc.m113.530196

Jagadeeswaran, G., Gainey, L., Prade, R., and Mort, A. J. (2016). A family of AA9 lytic polysaccharide monooxygenases in *Aspergillus nidulans* is differentially regulated by multiple substrates and at least one is active on cellulose and xyloglucan. *Appl. Microbiol. Biotechnol.* 100, 4535–4547. doi: 10.1007/s00253-016-7505-9

Jahromi, M. F., Liang, J. B., Ho, Y. W., Mohamad, R., Goh, Y. M., and Shokryazdan, P. (2012). Lovastatin production by *Aspergillus terreus* using agro-biomass as substrate in solid state fermentation. *J. Biomed. Biotechnol.* 2012:196264. doi: 10.1155/2012/196264

Jiménez-Quero, A., Pollet, E., Zhao, M., Marchioni, E., Averous, L., and Phalip, V. (2017). Fungal fermentation of lignocellulosic biomass for itaconic and fumaric acid production. *J. Microbiol. Biotechnol.* 27, 1–8. doi: 10.4014/jmb.1607.07057

Johnson, E. (2016). Integrated enzyme production lowers the cost of cellulosic ethanol. *Biofuels Bioprod. Bioref.* 10, 164–174. doi: 10.1002/bbb.1634

Jungmeier, G., Hingsamer, M., and van Ree, R. (2013). *Biofuel-Driven Biorefineries. A Selection of the Most Promising Biorefinery Concepts to Produce Large Volumes of Road Transportation Fuels by 2025*. Paris: IEA.

Jungmeier, G., Stichnothe, H., de Bari, I., Jørgensen, H., Van Ree, R., de Jong, E., et al. (2014). *A Biorefinery Fact Sheet for the Sustainability Assessment of Energy Driven Biorefineries – Efforts of IEA Bioenergy Task 42nd 'Biorefining'*. Wageningen: IEA Research Corporation.

Kameshwar, A. K. S., and Qin, W. (2018). Comparative study of genome-wide plant biomass-degrading CAZymes in white rot, brown rot and soft rot fungi. *Mycology* 9, 93–105. doi: 10.1080/21501203.2017.1419296

Kameshwar, A. K. S., Ramos, L. P., and Qin, W. (2019). CAZymes-based ranking of fungi (CBRF): an interactive web database for identifying fungi with extrinsic plant biomass degrading abilities. *Bioresour. Bioprocess.* 6:51. doi: 10.1186/s40643-019-0286-0

Kimura, S., Laosinchai, W., Itoh, T., Cui, X., Linder, C. R., and Brown, R. M. (1999). Immunogold labeling of rosette terminal cellulose-synthesizing complexes in the vascular plant *Vigna angularis*. *Plant Cell* 11, 2075–2085. doi: 10.2307/3871010

King, D., Hagan, A., Löfter, K., Gillman, N., Weihe, U., and Oertel, S. (2010). *The Future of Industrial Biorefineries*. Geneva: World Economic Forum.

Kracher, D., Scheiblbrandner, S., Felice, A. K. G., Breslmayr, E., Preims, M., Ludwicka, K., et al. (2016). Extracellular electron transfer systems fuel cellulose oxidative degradation. *Science* 352, 1098–1101. doi: 10.1126/science.aaf3165

Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M., and Henrissat, B. (2013). Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol. Biofuels* 6:41. doi: 10.1186/1754-6834-6-41

Li, Q., Loman, A. A., Coffman, A. M., and Ju, L. K. (2017). Soybean hull induced production of carbohydrases and protease among *Aspergillus* and their effectiveness in soy flour carbohydrate and protein separation. *J. Biotechnol.* 248, 35–42. doi: 10.1016/j.jbiotec.2017.03.013

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2013). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, 490–495.

Maijala, P., Kango, N., Szijarto, N., and Viikari, L. (2012). Characterization of hemicellulases from thermophilic fungi. *Springer Sci.* 101, 905–917. doi: 10.1007/s10482-012-9706-2

Martinez, D., Berka, R. M., Henrissat, B., Saloheimo, M., Arvas, M., Baker, S. E., et al. (2008). Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nat. Biotechnol.* 26, 553–560.

Mckendry, P. (2002). Energy production from biomass (part 1): overview of biomass. *Bioresour. Technol.* 83, 37–46. doi: 10.1016/SO960-8524(01)00118-3

Miao, Y., Dongyang, L., Guangqi, L., Pan, L., Yangchun, X., Qirong, S., et al. (2015). Genome-wide transcriptomic analysis of a superior biomass-degrading strain of A. fumigatus revealed active lignocellulose-degrading genes. *BMC Genomics* 16:459. doi: 10.1186/s12864-015-1658-2

Midorikawa, G. E. O., Correa, C. L., Noronha, E. F., Filho, E. X. F., Togawa, R. C., Costa, M. M. D. C., et al. (2018). Analysis of the transcriptome in *Aspergillus tamarii* during enzymatic degradation of sugarcane bagasse. *Front. Bioeng. Biotechnol.* 6:123. doi: 10.3389/fbioe.2018.00123

Miller, G. L. (1959). Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* 31, 426–428. doi: 10.1021/ac60147a030

Mohnen, D., and Hahn, M. G. (eds) (1993). *Cell Wall Carbohydrates as Signals in Plants. Seminars in Cell Biology*. Amsterdam: Elsevier.

Monclaro, A. V., and Filho, E. X. F. (2017). Fungal lytic polysaccharide monooxygenases from family AA9: recent developments and application in lignocellulose breakdown. *Int. J. Biol. Macromol.* 102, 771–778. doi: 10.1016/j.ijbiomac.2017.04.077

Narra, M., Dixit, G., Divecha, J., Madamwar, D., and Shah, A. R. (2012). Production of cellulases by solid state fermentation with *Aspergillus terreus* and enzymatic hydrolysis of mild alkali-treated rice straw. *Bioresour. Technol.* 121, 355–361. doi: 10.1016/j.biortech.2012.05.140

Nordberg, H., Cantor, M., Dusheyko, S., Hua, S., Poliakov, A., Shabalov, I., et al. (2014). I. The genome portal of the department of energy joint genome institute: 2014 updates. *Nucleic Acids Res.* 42, D26–D31.

Palazzolo, M., and Kurina-Sanz, M. (2016). Microbial utilization of lignin: available biotechnologies for its degradation and valorization. *World J. Microbiol. Biotechnol.* 32, 1–9.

Palonen, E. K., Raina, S., Brandt, A., Meriluoto, J., Keshavarz, T., and Soini, J. T. (2017). Transcriptomic complexity of *Aspergillus terreus* velvet gene family under the influence of butyrolactone I. *Microorganisms* 5:12. doi: 10.3390/microorganisms5010012

Peng, M., Aguilar-Pontes, M. V., de Vries, R. P., and Miia, M. R. (2018). *In Silico* analysis of putative sugar transporter genes in *Aspergillus niger* using phylogeny

and comparative transcriptomics. *Front. Microbiol.* 9:1045. doi: 10.3389/fmicb.2018.01045

Pierce, B. C., Agger, J. W., Zhang, Z., Wichmann, J., and Meyer, A. S. (2017). A comparative study on the activity of fungal lytic polysaccharide monooxygenases for the depolymerization of cellulose in soybean spent flakes. *Carbohydr. Res.* 449, 85–94. doi: 10.1016/j.carres.2017.07.004

Pitt, J. I., and Hocking, A. D. (2009). *Fungi and Food Spoilage*, 3rd Edn. New York, NY: Springer.

Prüfer, K., Muetzel, B., Do, H. H., Weiss, G., Khaitovich, P., Rahm, E., et al. (2007). FUNC: a package for detecting significant associations between gene sets and ontological annotations. *BMC Bioinform.* 8:41. doi: 10.1186/1471-2105-8-41

Pullan, S. T., Daly, P., Delmas, S., Ibbett, R., Kokolski, M., Neiteler, A., et al. (2014). RNAsequencing reveals the complexities of the transcriptional response to lignocellulosic biofuel substrates in *Aspergillus niger*. *Fungal Biol. Biotechnol.* 1, 1–14. doi: 10.1186/s40694-014-0003-x

Raeder, U., and Broda, P. (1985). Rapid preparation of DNA from filamentous fungi. *Lett. Appl. Microbiol.* 1, 17–20. doi: 10.1111/j.1472-765x.1985.tb01479.x

Reddy, C., and Singh, R. (2002). Enhanced production of itaconic acid from corn starch and market refuse fruits by genetically manipulated *Aspergillus terreus* SKR10. *Bioresour. Technol.* 85, 69–71. doi: 10.1016/s0960-8524(02)00075-5

Ribeiro, D. A., Cota, J., Alvarez, T. M., Bruchli, F., Bragato, J., Pereira, B. M., et al. (2012). The *Penicillium echinulatum* secretome on sugar cane bagasse. *PLoS One* 7:e50571. doi: 10.1371/journal.pone.0050571

Ridley, B. L., O'Neill, M. A., and Mohnen, D. (2001). Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry* 57, 929–967. doi: 10.1016/s0031-9422(01)00113-3

Roche, C. M., Glass, N. L., Blanch, H. W., and Clark, D. S. (2014). Engineering the filamentous fungus *Neurospora crassa* for lipid production from lignocellulosic biomass. *Biotechnol. Bioeng.* 111, 1097–1107. doi: 10.1002/bit.25211

Rosnow, J. J., Anderson, L. N., Nair, R. N., Baker, E. S., and Wright, A. T. (2017). Profiling microbial lignocellulose degradation and utilization by emergent omics technologies. *Crit. Rev. Biotechnol.* 37, 626–640. doi: 10.1080/07388551.2016.1209158

Ruiz-Dueñas, F. J., Lundell, T., Floudas, D., Nagy, L. G., Barrasa, J. M., Hibbett, et al. (2013). Lignin-degrading peroxidases in polyporales: an evolutionary survey based on 10 sequenced genomes. *Mycologia* 105, 1428–1444. doi: 10.3852/13-059

Saha, B. C. (2003). Hemicellulose bioconversion. *J. Ind. Microbiol. Biotechnol.* 30, 279–291. doi: 10.1007/s10295-003-0049-x

Segato, F., Damasio, A. R., de Lucas, R. C., Squina, F. M., and Prade, R. A. (2014). Genomics review of holocellulose deconstruction by *Aspergilli*. *Microbiol. Mol. Biol. Rev.* 78, 588–613. doi: 10.1128/mmbr.00019-14

Sethi, B. K., Das, A. S., Satpathy, A., and Behera, B. C. (2017). Ethanol production by a cellulolytic fungus *Aspergillus terreus* NCFT 4269.*10* using agro-waste as a substrate. *Biofuels* 8, 207–213. doi: 10.1080/17597269.2016.1221296

Sharma, R., Kocher, G. S., Bhogal, R. S., and Oberoi, H. S. (2014). Cellulolytic and xylanolytic enzymes from thermophilic *Aspergillus terreus* RWY. *J. Basic Microbiol.* 54, 1–11.

Silva, C. O. G., Vaz, R. P., and Filho, E. X. F. (2017). Bringing plant cell wall-degrading enzymes into the lignocellulosic biorefinery concept. *Biofuels Bioprod. Bioref.* 12, 277–289. doi: 10.1002/bbb.1832

Siqueira, F. G. (2010). *Resíduos Agroindustriais Com Potencial Para a Produção de Holocelulases de Origem Fúngica e Aplicações Biotecnológicas de Hidrolases*. Ph D. Thesis. Brasilia: Universidade de Brasília.

Siqueira, F. G., Siqueira, E. G., Jaramillo, P. M. D., Silveira, M. H. L., Andreaus, J., Couto, F. A., et al. (2009). The potential of agro-industrial residues for production of holocellulase from filamentous fungi. *Int. Biodeterior Biodegradation* 64, 20–26. doi: 10.1016/j.ibiod.2009.10.002

Soni, H., and Kango, N. (2013). "Microbial mannanases: properties and applications," in *Advances in Enzyme Biotechnology*, eds P. Shukla and B. I. Pletschke (New York, NY: Springer), 41–56. doi: 10.1007/978-81-322-1094-8_4

Soni, H., Rawat, H. K., Pletschke, B. I., and Kango, N. (2016). Purification and characterization of β-mannanase from *Aspergillus terreus* and its applicability in depolymerization of mannans and saccharification of lignocellulosic biomass. *3 Biotech* 6: 136. doi: 10.1007/s13205-016-0454-2

Sonnleitner, B., and Fiechter, A. (1983). Advantages of using thermophiles in biotechnological processes: expectations and reality. *Trends Biotechnol.* 1, 74–80. doi: 10.1016/0167-7799(83)90056-2

Sorensen, A., Teller, P. J., Lubeck, P. S., and Ahring, B. K. (2011). Onsite enzyme production during bioethanol production from biomass: screening for suitable fungal strains. *Appl. Biochem. Biotechnol.* 164, 1058–1070. doi: 10.1007/s12010-011-9194-2

Tani, S., Kawaguchi, T., and Kobayashi, T. (2014). Complex regulation of hydrolytic enzyme genes for cellulosic biomass degradation in filamentous fungi. *Appl. Microbiol. Biotechnol.* 98, 4829–4837. doi: 10.1007/s00253-014-5707-6

Todd, R. B., Zhou, M., Ohm, R. A., Leeggangers, H. A., Visser, L., and de Vries, R. P. (2014). Prevalence of transcription factors in ascomycete and basidiomycete fungi. *BMC Genomics* 15:214. doi: 10.1186/1471-2164-15-214

Turner, P., Mamo, G., and Karlsson, E. N. (2007). Potential and utilization of thermophiles and thermostable enzymes in biorefining. *Microb. Cell Fact.* 6, 9–12. doi: 10.1186/1475-2859-6-9

Van Dyk, J. S., and Pletschke, B. I. (2012). A review of lignocellulose bioconversion using enzymatic hydrolysis and synergistic cooperation between enzymes-Factor affecting enzymes, conversion and synergy. *Biotechnol. Adv.* 30, 1458–1480. doi: 10.1016/j.biotechadv.2012.03.002

van Munster, J. M., Daly, P., Delmas, S., Pullan, S. T., Blythe, M. J., Malla, S., et al. (2014). The role of carbon starvation in the induction of enzymes that degrade plant-derived carbohydrates in *Aspergillus niger*. *Fungal Genet. Biol.* 72, 34–47. doi: 10.1016/j.fgb.2014.04.006

Vankuyk, P. A., Diderich, J. A., Maccabe, A. P., Hererro, O., Ruijter, G. J., and Visser, J. (2004). *Aspergillus niger* mstA encodes a high-affinity sugar/H+symporter which is regulated in response to extracellular pH. *Biochem. J.* 379, 375–383. doi: 10.1042/BJ20030624

Varga, J., Tóth, B., Kocsubé, S., Farkas, B., Szakács, G., Téren, J., et al. (2005). Evolutionary relationships among *Aspergillus terreus* isolates and their relatives. *Anton. Leeuw. Int. J. G.* 88, 141–150. doi: 10.1007/s10482-005-3870-6

White, T. J., Bruns, T., Lee, S., and Taylor, J. (1990). "Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics," in *PCR Protocols: A Guide to Methods and Applications*, eds M. A. Innis, D. H. Gelgard, J. J. Sninsky, and T. J. White (New York, NY: Academic Press), 315–322. doi: 10.1016/b978-0-12-372180-8.50042-1

Wieczorke, R., Krampe, S., Weierstall, T., Freidel, K., Hollenberg, C. P., and Boles, E. (1999). Concurrent knock-out of at least 20 transporter genes is required to block uptake of hexoses in *Saccharomyces cerevisiae*. *FEBS Lett.* 464, 123–128. doi: 10.1016/S0014-5793(99)01698-1

Wijeratne, E. M. K., Turbyville, T. J., Zhang, Z., Bigelow, D., Pierson, L. S., VanEtten, H. D., et al. (2003). Cytotoxic constituents of *Aspergillus terreus* from the rhizosphere of *Opuntia versicolor* of the Sonoran Desert. *J. Nat. Prod.* 66, 1567–1573.

Windeisen, E., and Wegener, G. (2012). "Lignin as building unit for polymers," in *Polymer Science: A Comprehensive Reference*, eds K. Matyjaszewski and M. Miller (Amsterdam: Elsevier), 255–265. doi: 10.1016/b978-0-444-53349-4.00263-6

Wu, V. W., Thieme, N., Huberman, L. B., Dietschmann, A., Kowbel, D. J., Lee, J., et al. (2020). The regulatory and transcriptional landscape associated with carbon utilization in a filamentous fungus. *Proc. Natl. Acad. Sci. U.S.A.* 117, 6003–6013. doi: 10.1073/pnas.1915611117

Yadav, R. P., Saxena, R. K., Gupta, R., and Davidson, W. S. (1998). Purification and characterization of a regiospecific lipase from *Aspergillus terreus*. *Biotechnol. Appl. Biochem.* 28, 243–249.

Young, E., Lee, S.-M., and Alper, H. (2010). Optimizing pentose utilization in yeast: the need for novel tools and approaches. *Biotechnol. Biofuels*. 3:24. doi: 10.1186/1754-6834-3-24

Zeng, Y., Himmel, M. E., and Ding, S.-Y. (2017). Visualizing chemical functionality in plant cell walls. *Biotechnol. Biofuels* 10:63. doi: 10.1186/s13068-017-0953-3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership