Metagenomic Discovery and Characterization of Multi-Functional and Monomodular Processive Endoglucanases as Biocatalysts

Ming Z. Fan 1,2,*, Weijun Wang 1, Laurence Cheng 1, Jiali Chen 1, Wenyi Fan 1 and Min Wang 1

Citation: Fan, M.Z.; Wang, W.; Cheng, L.; Chen, J.; Fan, W.; Wang, M. Metagenomic Discovery and Characterization of Multi-Functional and Monomodular Processive Endoglucanases as Biocatalysts. Appl. Sci. 2021, 11, 5150. https://doi.org/10.3390/app11115150

Academic Editors: Carl Grey and Javier Linares-Pastén

Featured Application: Multifunctional and monomodular processive endoglucanases, once optimized to be efficacious, stable and inexpensive, will be potentially valuable biocatalysts for various transformative biomass processing industrial applications, including biomass degradation for manufacturing lignocellulosic biofuels such as ethanol and butanol; biomaterials; light industries such as food and fruit processing, textile and wood processing; and exogenous nutraceutical enzyme supplements for human foods, nutrition and health as well as exogenous feed enzyme additives for aquaculture, companion animal and agricultural food animal feed industrial applications.

Abstract: Biomass includes cellulose, hemicelluloses, pectin and lignin; constitutes the components of dietary fibre of plant and alge origins in animals and humans; and can potentially provide inexhaustible basic monomer compounds for developing sustainable biofuels and biomaterials for the world. Development of efficacious cellulases is the key to unlock the biomass polymer and unleash its potential applications in society. Upon reviewing the current literature of cellulase research, two characterized and/or engineered glycosyl hydrolase family-5 (GH5) cellulases have displayed unique properties of processive endoglucanases, including GH5-tCel5A1 that was engineered and was originally identified via targeted genome sequencing of the extremely thermophilic Thermotoga maritima and GH5-p4818Cel5_2A that was screened out of the porcine hindgut microbial metagenomic expression library. Both GH5-tCel5A1 and GH5-p4818Cel5_2A have been characterized as having small molecular weights with an estimated spherical diameter at or < 4.6 nm; being monomodular without a required carbohydrate-binding domain; and acting as processive $\beta-1,4$-endoglucanases.

These two unique GH5-tCel5A1 and GH5-p4818Cel5_2A processive endocellulases are active in hydrolyzing natural crystalline and pre-treated cellulosic substrates and have multi-functionality towards several hemicelluloses including $\beta$-glucans, xylan, xyloglucans, mannans, galactomannans and glucomannans. Therefore, these two multifunctional and monomodular GH5-tCel5A1 and GH5-p4818Cel5_2A endocellulases already have promising structural and functional properties for further optimization and industrial applications.

Keywords: biomass; porosity; biofuels and biomaterials; dietary fibre; cellulose and hemicelluloses; metagenomic discovery of cellulases; processive endocellulases; structure and functionality; exogenous feed & food enzymes; biocatalysts

1. Introduction

Biomass primarily of plant cell wall and alge polysaccharides, including cellulose, hemicelluloses, pectin and lignin, constitutes the major constituents of dietary fibre [1,2]. As an important dietary polymer nutrient, dietary fibre is utilized by various species of
animals [2,3]. Microbial degradation of these fibre components in the gastrointestinal tract is indispensable in maintaining healthy gut functions, gaining energy, aiding digestion of minerals and providing vitamins to the host [2,3]. Lignin, representing a large group of aromatic polymers, and native cellulose of the quasi-crystalline structure in the lignocellulose composite are mainly responsible for the recalcitrance of plant cell wall by limiting accessibility to biomass degrading microbial enzymes [4–8]. Microorganisms have evolved a multiplicity of enzymes to effectively degrade the composite structure of lignocellulose-based biomass, including laccases and peroxidases, glycoside hydrolases such as cellulases and hemicellulases, polysaccharide lyases, lytic polysaccharide mono-oxygenases (LPMO) and the carbohydrate esterases to harvest metabolic energy ATP in adapting to their diverse environments [9–14]. Thus, microbial biomass degrading cellulases are a group of important biocatalysts for biomass industrial applications.

The enzymatic degradation of biomass polysaccharides is initiated by the cleavage of glycosidic bonds in the backbone. Thus, the major scientific-technological hurdle for effective utilization of lignocellulose biomass is to develop efficacious industrial biomass-degrading biocatalysts for cost-effective degradation of lignocellulose polymer into their constituent simple sugars such as glucose and pentoses, including xylose, mannose, galactose and arabinose, as illustrated in Figure 1. Lignocellulosic matrix is covered by a protective sheath locked by heterogenous polymers such as hemicelluloses, pectin and guar gum in cross-linkages with lignin. Among a variety of glycosidic bonds, β-1,4-glycosidic linkage is the most widely distributed bond in plant cell wall polysaccharides such as in cellulose and some hemicelluloses [1,11,15]. Collectively, these two group fibre components account for approximately 60% of plant cell wall materials [16–18], making β-1,4-glycosidic bond cleavage one of the key biochemical reactions for biomass degradation [5,12,15]. As the most abundant biosynthetic polymer on the planet, native cellulose, referred to as cellulose I, has two distinct crystallite forms with Iα being dominant in bacterial and algal celluloses and Iβ being dominant in celluloses of higher plant origins [19]. It is believed that the two crystalline cellulose allomorphs of Iα and Iβ coexist in celluloses of most natural lignocellulose of biomass materials; and an all-atom coarse-grained simulation model has been developed in describing the conversion process between the cellulose Iα and Iβ allomorphs, characterizing the interface between the two crystalline forms of the allomorphs and determining their amorphous transition states under the room-temperature [20]. It has been estimated that the yearly biomass production of cellulose in the world is at about 1.5 trillion tons, making it an essentially inexhaustible resource for mankind [21]. By assuming effective pre-treatment and the use of highly efficacious cellulase enzymes and microbes capable of fermenting and processing hexose and pentose sugars in biorefineries, an integrated methodology of modelling assessment revealed that deployment of wheat straw could be economically feasible for producing liquid transportation fuel ethanol, biomaterial ethyl levulinate and electricity meanwhile gaining a positive environmental margin in terms of mitigating a significant amount of annual CO₂ emission [22,23]. Saccharification of various natural sources of cellulose is a major focus of research efforts aimed at developing renewable biorefinery-based alternatives to fossil fuels [21]. Efficient cleavage of β-1,4-glycosidic bonds in natural and pretreated lignocellulose polymer substrates by cellulases is the limiting biochemical step and the bottleneck for the bioconversion of lignocellulose to biofuels and biochemicals [5,11,24–26], food processing, textile, laundry detergent, paper and wood pulp light industries, plant and horticulture based agriculture [27], and for the improvement of efficiency of fibre feed utilization and food animal productivity [3,28–31] and enteric health [30,32,33] in animal production industries, which are other core aspects of developing a sustainable bioeconomy on the global stage.

There has been extensive genomic and gene-cataloguing research examining the microbial anaerobic fermentation of dietary fibre in the mammalian gut, particularly in humans, rodents and ruminants [34,35]. In contrast, understanding of parallel microbial processes at functional and genomic levels in the non-herbivore and non-human farm
animal monogastric hindgut is very limited [36]. Pigs, the single largest source of red meat for human consumption worldwide, are monogastric omnivores well-known for their ability to readily adapt to a variety of high-fibre feeding. Interestingly, the turnover rate of digesta throughout the porcine gut system is much faster than that of many other large frame size herbivore animals. For example, digesta retention time is approximately 2 to 6 h in the small bowel and 12 to 18 h in the large bowel in the pig [37] in comparison with cattle that require approximately 50–70 h for feed to pass through the gastrointestinal tract (GI) [38–40]. Monogastric pigs are surprisingly efficient at degrading a commercial crystalline cellulose product of Solka-Floc® that is generated from wheat straw and woodchips via physicochemical pretreatment [41]. Dietary supplemental Solka-Floc® was degraded by 62% and 83% at the distal ileum and the fecal levels, respectively [41]. These would represent estimated crystalline cellulose degradation rates of 10.4 and 4.6%/h at the distal ileum and the fecal levels, respectively, in the pig. This crystalline cellulose degradation process is dependent largely on the symbiotic microbiota in the porcine intestine [42,43]. However, ruminal degradation rate of this commercial crystalline cellulose Solka-Floc® was estimated at 1.2%/h [3]. Wang et al. (2012; 2019) identified and characterized a multifunctional processive endocellulase from the porcine hindgut microbiome through a functional screening metagenomic expression library [44,45]. These research reports suggest that a unique and efficient microbial fibre degradation system that adapts to a faster rate of food digesta passage has evolved in the pig intestinal system for the discovery of novel and efficacious fibre degradation enzymes.

Figure 1. An illustration of the major processes and biological steps in utilization and conversion of agricultural waste feedstock biomass in providing sustainable functional foods and food animal production and renewable lignocellulosic ethanol, functional soluble fiber and other biomaterials with an emphasis on the essential roles of biomass-degrading enzymes as novel industrial enzymes.

2. Metagenomic Discovery of New Microbial Cellulases

Commercial cellulase products are already widely used in some traditional industrial applications such as food processing, the textile industry, detergent products, and pulp and paper industries [4,25,27,45]. Enzyme companies Novozymes of Denmark and Monosanto of St. Louis in US have announced a joint global venture [46]; and aimed to increase crop yields while reducing inputs presumably by enhancing soil organic matter fertility via microbial enzyme products. More profound contemporary applications of new generations of highly efficient and capable cellulases in saccharification of raw and pretreated lignocellulose substrates for commercial production of cellulosic ethanol and in supplementation as exogenous feed enzyme additives in enhancing feed energy and carbon conversion efficiency in agricultural food animal feed industries are the two key directions in developing a biorefinery-based bioeconomy in the continued global economic development [11,25,26,47]. Novel and efficient cellulases as biocatalysts are the key driver for these new developments. Currently, commercial cellulase genes were obtained through the traditional cloning techniques from Trichoderma and Aspergillus fungal species and Bacillus sp. with further engineering for steadily improved enzyme functionality and protein yields and a dramatically reduced fermentation cost [4,25,48,49]. Prospectively, a commer-
cial scale of cellulosic ethanol production was projected to launch businesses on multiple sites [50]. Needless to say, there is a continued need to further improve cellulase enzyme specific activities to ensure that this newly emerging cellulosic ethanol industry will become progressively viable and economically competitive [51]. On the other hand, there is still a lack of well recognized cellulase enzyme products due to limited enzyme activity in hydrolyzing natural raw lignocelluloses of feed sources for their effective applications in animal production [29–32]. Under this context, the discovery and characterization of newly emerged novel cellulases, as described herein, have been discussed for their potential commercial applications as novel and efficient cellulase products for both the cellulosic ethanol production and the animal feed enzyme industry, among others.

One newly evolved possible approach for discovering new cellulases involves metagenomic mining, including metagenomic analysis and metagenomic expression library, for microbial cellulases in random samples of DNA fragments from material extracted directly from complex environments. This metagenomic approach circumvents the highly restrictive requirement that organisms be culturable [52]. This allows enzymes from a far richer variety of microorganisms to be sampled and characterized. For example, functional screening of microbial metagenomic expression libraries [44,45] and gene-centric based metagenomic gene cataloguing have unearthed a diversity of cellulases from the bovine rumen [52–55], the rabbit cecum [56], the human gut [57], the rodent gut [58], hindgut of wood-feeding higher termites [59], the earthworm gut [60,61], decaying wood [62], soil [63–65] and environmental samples [66]. By comparing to the major metagenomic database, Zhou et al. (2010) have concluded that animal gut metagenomes are substantially enriched with novel glycosyl hydrolases [67]. Wang et al. (2019) identified a multifunctional processive endocellulase from the porcine hindgut microbiome [45]. Therefore, metagenomic tools are powerful for the discovery of novel cellulases and the animals’ gut environment is a unique genomic resource for such discovery.

3. Limiting Steps of the Cellulose Hydrolytic Pathway

The word cellulases is a generic term referring to a group of cellulose hydrolases classified under different glycosyl hydrolase families (GH), including endo-acting endocellulases, exo-acting exocellulases and β-glucosidases that independently and synergistically degrade polymer cellulose substrates into soluble cellodextrins and glucose [15]. The soluble cellodextrins typically have a degree of polymerization (DP) of DP of 2–6 and include cellobiose (DP = 2), cellotriose (DP = 3), cellotetraose (DP = 4), cellopentaose (DP = 5) and cellohexaose (DP = 6) [19]. Classic non-processive endocellulases (EC 3.2.1.1), also termed 1,4-β-D-glucan-4-glucanohydrolases, carry out hydrolysis by inserting a water molecule in the 1,4-β-glucosidic bond randomly at internal amorphous sites in the polymer cellulose, generating cello-oligosaccharides of various lengths with new chain ends [5,15].

The first group of exocellulases (EC 3.2.1.74), specifically 1,4-β-D-glucan glucanohydrolases or cellodextrinases [15], are also known as glucan-1,4-β-glucosidases according to the Braunschweig Brenda Enzyme Database (http://www.brenda-enzyme.org, accessed on 31 May 2021) or 1,4-β-D-glucan cellobiohydrolases II (CBHII) [68]. These exocellulases primarily cut the β-1,4-glycosidic linkage in cello-oligosaccharides as well as amorphous and crystalline polymer cellulosics from non-reducing chain ends with low specific activities towards cellobiose, thus processively releasing cellobiose as the main product with very little glucose release [69,70]. The second group of exocellulases (EC 3.2.1.91), namely, 1,4-β-D-glucan cellobiohydrolases [15] or cellobiohydrolases I (CBHI) [67], are also termed cellulose 1,4-β-cellobiosidases according to the Braunschweig Brenda Enzyme Database (http://www.brenda-enzyme.org). These exocellulases work on the β-1,4-glycosidic linkage in low-DP cellodextrins from reducing chain ends with activities towards soluble cellodextrins and cellobiose, hence primarily releasing cellobiose and glucose in a processive manner [70]. The Trichoderma fungal exocellulases are so far the best elucidated processive cellulases in terms of biological mechanisms [71,72]. Thus, exocellulases are also the classic group of processive cellulases.
More recently, processive endocellulases have also been discovered and characterized to contain a GH5 or a GH9 catalytic domain (Table 1), and these are very rare endocellulases that can continue to hydrolyze their initially broken-apart cello-oligosaccharides into soluble cellodextrins (DP of 2–6) and glucose with one such type of cellulases playing both endo-acting and exo-acting hydrolytic roles [72,73]. β-Glucosidases (EC 3.2.1.21) are primarily responsible for the terminal hydrolysis of cellobiose and, to some extent, may also hydrolyze some soluble cellooligosaccharins into D-glucose [5,15]. McCleary et al. (2012) further suggested that the classical β-glucosidases hydrolyze soluble cellooligosaccharins (DP of 2–6) at similar rates whereas the rate of hydrolysis of exo-β-glucosidases would increase as the DP increases from 2 to 6 [74], further distinguishing between the classic β-glucosidase and the exo-β-glucosidase activities.

Furthermore, it is believed that endocellulase and exocellulase activities in hydrolyzing natural and pretreated cellulose substrates are the rate-limiting steps of the polymer cellulose biochemical degradation process [4,5,10,15]. This notion can be further substantiated by the following facts. For comparison, one standard enzyme activity unit (U) is defined to be 1 umol D-glucose or reducing sugar released per min. Firstly, the highest reported endocellulase activity is at about 30 U/mg protein on Avicel and the exocellulase activity is at about 246 U/mg protein on phosphoric acid swollen cellulose (PASC) [75]. Most of the well characterized microbial endocellulase and the exocellulase activities are typically below 0.200 U/mg protein as summarized by Lynd et al. (2002) [15] and compiled by the Braunschweig Brenda Enzyme Database (http://www.brenda-enzyme.org). Correspondingly, α-amylase (EC 3.2.1.1) and maltase-glucoamylase or amyloglucosidase (EC 3.2.1.3) activities in hydrolyzing α-glycosidic bond-based polysaccharides of starch and maltodextrins are very high at about 4000 U/mg and 758 U/mg proteins as compiled by the Braunschweig Brenda Enzyme Database (http://www.brenda-enzyme.org). Thus, endocellulase and the exocellulase activities in hydrolysis of β-glycosidic bond-based natural raw and pretreated lignocelluloses & β-glucans and insoluble cellodextrins are considerably lower than the α-amylase and maltase-glucoamylase or amyloglucosidase activities towards α-glycosidic bond-based starch and maltodextrins. Secondly, characterized microbial β-glucosidase activities are typically much higher than the endocellulase and the exocellulase activities in hydrolysis of their natural substrates including cellobiose. For example, β-glucosidase activity in hydrolysis of cellobiose reaches the level of about 665 U/mg protein as compiled by the Braunschweig Brenda Enzyme Database (http://www.brenda-enzyme.org). This is also likely, in part, due to the fact the natural substrates of β-glucosidases are soluble cellooligosaccharins. Thirdly, comprehensive studies demonstrated that free multimodular processive cellulases were effective and the limiting enzymatic factors in hydrolysis of pure crystalline cellulose substrates, physiochemically pretreated biomass and raw biomass materials [76,77]. Thus, discovery of highly efficient and efficacious synergistical endocellulases and exocellulases, particularly the uniquely high active processive endocellulases that hydrolyze crystalline cellulose substrates, is the key to conquer the biomass recalcitrance.

4. Diversity of Microbial Cellulases

A large diversity of endocellulases and exocellulases has been characterized in aerobic and anaerobic bacteria and fungi, protozoa and some unique animal species [12,26]. Endocellulases and exocellulases are further classified into non-complex free cellulases [12,26] and the complex cellulosomes [26] based upon how these enzymes are associated with their original organisms and what biological mechanisms that these enzymes employ to carry out cellulose degradation. The highest wild-type bacterial free cellulase activity was reported at about 2.5 U/mg protein on Avicel for a GH9 processive endoglucanase originated from Clostridium phytofermentans [78] and a GH5 bacterial endoglucanase screened from the Buffalo rumen metagenomic library [55]. Directed evolution of this Clostridium phytofermentans GH9 processive endoglucanase further enhanced its activity to about 6 U/mg protein on Avicel [79]. The highest fungal free cellulase activities were reported at 9.7 and 29.8 U/mg protein for two exocellulases on Avicel and these free cellulases
were respectively originated from ruminal Neocallimastix patriciarum [80] and the White Rot fungus Irpex lacteus [75]. The highest purified cellulosome activities were reported at 9.6–14.6 U/mg protein on Avicel from Clostridium thermocellum [81]. It is apparent that these abovementioned highest levels of cellulase specific activities reported so far for hydrolyzing crystalline cellulosic substrates are still far below the specific activities reported for some of the well-established industrial enzymes such as commercial phytases, hemicellulases and α-amylases, typically having specific activity ranging from several hundreds to thousands U/mg protein. Thus, it is imperative to continue the research efforts for the discovery and engineering of highly active endocellulases and exocellulases for their various industrial applications.

Free cellulases are typically secreted out of their originating organisms or cells and are sloughed into surrounding environments and these enzymes are further divided into mono-modular, possessing only one catalytic domain, bi-modular, having one carbohydrate-binding domain or module (CBD or CBM) and one catalytic domain, and multi-modular, containing one or more CBD and one or more catalytic domains, cellulosomes [12,26,27]. The classic view is that almost all cellulase enzymes that degrade insoluble cellulosic substrates contain a substrate-binding domain linked to the catalytic domain via a linker peptide [12]. Thongekkaew et al. (2013) demonstrated that fusion of a cellulose-binding module enhanced cellulase binding affinity and cellulolytic activity to insoluble cellulosic substrates [82]. However, possessing a CBD in cellulases does not mean that cellulases have high activity on insoluble especially crystalline cellulosic substrates, which is deemed to be an essential property for potential industrial applications. For example, two ruminal GH5 cellulases that were characterized to contain CBD could effectively hydrolyze soluble substrate CMC but did not have activities on the model crystalline cellulosic substrate Avicel [83,84]. A number of free cellulases that are reported to possess CBD have little or no activities on model crystalline cellulosic substrates as compiled in the Braunschweig Brenda Enzyme Database (http://www.brenda-enzyme.org). Alternatively, cellulose substrates as ligands can directly bind to catalytic domains to engage hydrolysis with non-processive and processive cellulases [85]. Payne et al. (2011) [86] and Taylor et al. (2013) [87] further demonstrated that cellulases bound to hydrophobic regions in catalytic domain active sites, especially the tryptophan residues in processive cellulases.

On the other hand, cellulosomes are multi-enzyme based multi-functional complexes that are anchored onto the outer membrane of some anaerobic bacteria and fungi and can synergistically hydrolyize cellulosic substrates and/or hemicelluloses [26,88–90]. Cellulosomes are typically structured in such a way that multi-enzymes are linked by one or more scaffold proteins as scaffold subunits that, in one direction or end, are further extended with one or more “X” modules of unknown function and cellulose binding module(s) that are tethered to bind to cellulose or hemicellulose substrates, and in the other direction or end, are further anchored to the bacterial outer membrane via dockerin and cohesion proteins, and related stepwise actions and mechanisms have been well documented and illustrated in the literature [26,91,92]. By using synthetic biology, designer cellulosomes are being developed to optimize their catalytic efficiency [13,90–92], ideally via combining component enzyme proximity and flexibility with a reduced complex enzyme mass.

5. Biomass Porosity and Cellulase Efficacy

With this context, it should be pointed out that cellulase accessibility to hydrolytic cellulosic surface area is a major determinant of lignocellulosic biomass degradation. Biomass has both external and internal surface areas with natural lignocellulosic materials having a much greater internal surface area [19]. It is well accepted that the primary cell wall of most untreated and natural plant cells is semi-permeable with pores, permitting the simple diffusive passage of small molecules such as CO₂ and H₂O, and small proteins with size exclusion estimated to be at 30–60 kDa (Cell wall, Wikipedia). Assuming that free cellulases have the simplest shape, a sphere, these aforementioned small proteins with size exclusion estimated at 30–60 kDa for access to plant cell wall pores are calculated to have a radius
range of 2.05–2.6 nm and a diameter range of 4.1–5.2 nm, respectively, by using the formula provided by Erickson (2009) [93]. Much earlier studies by Carpita et al. (1979) reported that pore size diameters in various living plant cells ranged between 3.5 and 5.2 nm [18]. Mechanically processed (e.g., via chopping and grinding) but untreated lignocellulosic biomass materials are particulate and porous with intraparticulate pores (a diameter range of 1–10 nm) as well as interparticulate voids (> 5 µm) [19]. Several studies documented that pretreated lignocellulosic biomass materials and crystalline cellulose model substrates are associated with much larger pore diameters, ranging between 5.1–11.0 nm [94–97]. Thus, it can be concluded that natural, raw, untreated biomass materials such as plant feeds have an estimated upper cut-off pore size diameter at about 5.2 nm while physiochemical pretreated biomass materials are associated with an upper pore size diameter at about 11.0 nm.

As biocatalysts, molecular dimension sizes of enzymes are generally related to their molecular weights. Enzymes with large molecular weights are likely to bear large molecular sizes or molecular diameters in aqueous solutions. Erickson (2009) viewed that most proteins fold into globular domains and polypeptides larger than 50 kDa typically form two or more domains [91]. Cowling and Kirk (1976) summarized estimated sizes of various fungal cellulases that were projected with a sphere and/or ellipsoid shape [98]. The surveyed fungal cellulases were associated with spheric diameters ranging between 3.3 and 7.7 nm, and ellipsoid shapes measuring (width x length) between 1.8 × 10.8 and 4.2 × 25.0 nm for their molecular weights ranging between 11.4 and 76 kDa [98]. Typical free cellulases, with a large catalytic domain core being connected to a small carbohydrate-binding module via a short peptide linker or spacer, were shown to have tadpole shapes by using small-angle x-ray scattering analysis [99,100]. For example, the Trichoderma reesei exocellulase Cel7A (CBHI) and endocellulase 7B (EGI) were estimated to have a tadpole shape (catalytic domain core diameter x molecule length) measured at 4.4 × 19.6 and 5.3 × 18.0 nm, respectively [99,100]. Ding et al. (2012) showed that only a few surface microfibrils were degraded by the fungal Trichoderma reesei cellulases in the untreated raw cell wall of biomass materials [7]. It is known that microfibril surface represents a major portion of the internal surface area in biomass materials. Their results would suggest that Trichoderma reesei free cellulases with a catalytic domain core diameter ranging 4.1–5.2 nm had limited diffusion into plant cell wall pores that have an upper cut-off porosity diameter at about 5.2 nm in the untreated raw biomass materials. In the Ding et al. (2012) study, the authors further showed that the Trichoderma reesei cellulases effectively penetrated into the microfibril internal surface and engaged degradation of these microfibrils with digestion holes or pits in the pretreated delignified biomass materials, whereas purified cellulosomes from Clostridium thermocellum were shown to only peel off individual microfibrils from the cell wall surface in these pretreated biomass materials [7]. These results by Ding et al. (2012) [7] would suggest that the dimension size of purified Clostridium thermocellum cellulosomes was likely much larger than the upper pore diameter of 11 nm in the pretreated biomass materials reported by Grethlein (1985) [95] and Hui et al. (2009) [97], reflecting the fact that cellulosomes are rather large multi-enzyme complexes with molecular weights ranging from several hundreds to over 1 million kDa. Xu et al. (2013) [101] used domain engineering to exploit the advantage of intramolecular and intermolecular proximity synergies by designing mini-cellulosomes to improve cellulase specific activity as well as to enhance their biomass-penetrating feature. Other strategies such as glycosylation by overexpression designer cellulosomes in other microbial platforms could further improve designer cellulosomes’ thermostability and activity [102].

Meanwhile, Brunecky et al. (2013) demonstrated that the multimodular free cellulase CelA cloned from the thermophilic bacterium Caldicellulosiruptor bescii could effectively hydrolyze crystalline cellulose substrates via mechanisms of surface ablation and cavity excavating driven by a combined endo-acting and exo-acting processivity [77]. However, the Caldicellulosiruptor bescii Cel A had a large molecular weight of about 230 kDa [103], and an estimated effective size of 10–35 nm, thus limiting its diffusion into microfibril internal surface for biomass hydrolysis [77]. Furthermore, the large molecular Caldicel-
losiruptor bescii Cel A was associated with multiple carbohydrate-binding modules and its activity would be affected, to a greater extent, by nonproductive adsorption than much lower molecular weight cellulases [77]. Therefore, small molecular weights may be a favorable and advantageous biochemical feature in developing highly active processive free cellulases that may be effective to penetrate into microfibrillar internal surfaces for biomass hydrolysis.

Cell wall biomass recalcitrance is collectively referring to the natural resistance of cell wall materials to microbial and enzymatic deconstruction [5,9]. The biomass recalcitrance is essentially composed of two layers of resistance, including the water-excluding nature of the resistance of the crystalline cellulose bundle core within each of the elementary fibrils as well as the resistance representing the crosslinks around the cellulose bundle core formed by hemicellulose & pectin embedding each elementary fibril within each of microfibrils and the aromatic net further patched by lignin polymers on the exterior of the microfibrils within each of the macrofibril structure primarily in plant cell wall materials [19,104–106]. The first layer of biomass resistance as a measure of water-excluding nature associated with the crystalline cellulose bundle core within the elementary fibril is also reflected by the crystallinity index (CrI, ranging between 0 and 1) that is experimentally measured for a number of biofuel materials [19,107] but in very few plant feed materials for animals. The crystalline cellulose bundle core within each elementary fibril represents about 100 cellulose glucans that are tightly aggregated via extensive interchain and intrachain hydrogen bonds and van der Waals forces in forming a straight and stable supramolecular insoluble quasi-crystalline structure with a typical crystal size of 4–5 nm, being variable dependent upon biomass sources [19,105,106,108]. Each cellulose glucan is a polymer consisting of D-anhydroglucopyranose joined together by β-1,4-glycosidic bonds, with anhydrocellobiose as the repeating unit [19]. Each microfibril contains a single elementary fibril and is surrounded and embedded in a matrix of non-crystalline hemicelluloses & pectic polysaccharides with microfibril dimensions typically ranging from 2–3 nm to 20–50 nm, as affected by plant species and tissue types [19,105,106]. When microfibrils interact with one another to form aggregation via van der Waals forces, macrofibrils are formed and their dimensions vary from 50 to 250 nm in diameter [99,105,106,108]. Thus, plant cell materials typically consist of macrofibril, microfibril and elementary fibril units.

Cell wall biomass materials are porous materials. Porous materials are classified into three categories including micropores (pore diameters less than 2 nm), mesopores (pore diameters between 2–50 nm) and macropores (pore diameters larger than 50 nm) [108]. The micropores (around 1 nm in pore diameters) and the mesopores (between 2 and 5.2 nm in pore diameters) were likely to be the capillaries between adjacent microfibrils while the macropores (around 20 to 100 nm in pore diameters) were likely to be the space along the external surface between macrofibrils in plant cell wall biomass materials, as was originally proposed by Cowling and Kirk (1976) [98] and further discussed by Guo and Catchmark (2011) [107]. It is also reasonable to believe that micropores represent the majority whereas mesopores and macropores likely only represent a small proportion of the total pores in the raw biomass materials. Hence, another layer of biomass resistance is the restricted nanoporosity, which reflects the crosslinks formed by hemicellulose & pectic polysaccharides embedding each elementary fibril via non-covalent bonds and the lignin polymers surrounding the exterior of each microfibril via covalent bonds, and is related to free space across microfibrils and macrofibrils in the biomatrix [109]. Pores in biomass can be open or closed. Open pores are accessible, whereas closed pores are inaccessible, to accommodate cellulases for cellulose hydrolysis [107]. The drying process, especially drying cell wall biomass materials at higher temperature for a long duration (e.g., 150 °C for 1 h), can cause fibre hornification that is a consequence of the irreversible change of the cell wall structure due to the collapse of pores [109]. Carpita et al. (1979) reported that pore size diameters in various living plant cells ranged between 3.5 and 5.2 nm, as measured by using a solute exclusion technique [18]. Biomass material accessibility and porosity can be experimentally measured with several techniques, including a solute exclusion, water
vapor sorption and a liquid N\textsubscript{2} adsorption, and these have been conducted for a number of biofuel materials \[19,109\]. As reviewed in an earlier section, several previous studies documented that pretreated lignocellulosic biomass materials and purified crystalline cellulose model substrates such as Avicel, α-cellulose, Sigmacell (20, 50 and 100) and Soltka-Floc\textsuperscript{®} were associated with much larger pore diameters, ranging 5.1–11.0 nm \[94–97\]. Thus, it is safe to predict that fresh fibre feeds such as grasses for grazing herbivores have an upper porosity diameter at about 5.2 nm according to the study by Carpita et al. (1979) \[18\]. To the best of our knowledge, these types of porosity measurements have not been well reported in other raw plant feed materials for agricultural food animals such as corn silage, ground hay, DDGS, wheat shorts and middlings, canola meal and soybean meal, etc.

6. Properties of Fungal Biomass Degradation Hemi-Cellulases and Cellulases

Aerobic fungal species, typically including white, brown and gray wood rot fungi, are well characterized microorganisms, and they employ a multiplicity of synergistic mechanisms to effectively degrade wood-based lignocellulose biomass materials high in lignin \[110–113\]. It is equally important to compare GH5-p4818Cel5\_2A specific activity (0.216 \(\mu\)mol D-glucose/mg protein.min) in hydrolyzing the model insoluble crystalline cellulose substrate Avicel reported by Wang et al. (2019) \[45\] with specific activities that are expressed and are converted into the international unit (1 IU = \(\mu\)mol D-glucose per mg protein per min) in some of the well characterized aerobic fungal endocellulases and exocellulases reported in the literature. The white rot fungus basidomycete \textit{Irпex lacteus} grows on fallen wood and decomposes it by secreting multiple enzymes including cellulases, hemicellulases and lignin-degrading enzymes with its strain MC-2 being used as a producer of a commercial enzyme preparation called Driselase marketed by Kyowa Hakko, Co., Ltd. \[75\]. Hamada et al. \[114,115\] isolated a cellulase gene \textit{cel2} encoding for two highly active GH7 exocellulases Ex-1 (a molecular weight of about 53 kDa) and Ex-2 (a molecular weight of 56 kDa) and a cellulase gene \textit{cel3} encoding a GH7 exocellulase \textit{Cel3} protein (a predicted molecular weight of about 56 kDa), equivalent to the \textit{Trichoderma reesi} exocellulase \textit{Cel7A} (CBHI) in the basidomycete \textit{Irпex lacteus} with specific activities reported only for Ex-1 and Ex-2 (20.9 and 21.1 \(\mu\)mol D-glucose/mg protein.min) determined with Avicel. Toda et al. (2005) isolated a cellulase gene \textit{cel4} encoding for another highly active GH6 exocellulase Ex-4 (a molecular weight of about 45 kDa), equivalent to the \textit{Trichoderma reesi} exocellulase \textit{Cel6A} (CBHIII), in the basidomycete \textit{Irпex lacteus} with specific activity (29.8 \(\mu\)mol D-glucose/mg protein.min) determined with Avicel \[116\]. Toda et al. (2005) also isolated a cellulase gene \textit{cen1} encoding for a GH5 endoglucanase I termed En-1 (a molecular weight of about 52 kDa) in the basidomycete \textit{Irпex lacteus} with relatively low specific activities (7.76–25.9 \(\mu\)mol D-glucose/mg protein.min) determined with the soluble substrate CMC \[116\]. While the basidomycete \textit{Irпex lacteus} GH5 endoglucanase-I En-1 activity in hydrolyzing Avicel was not reported in the study, the researchers did show that En-1 in combination with the exocellulase Ex-1 had dramatic synergistic effects in enhancing the hydrolysis of crystalline cellulose substrates of Avicel and PASC \[116\]. Zheng and Ding \[117\] isolated a bi-modal processive GH5 endoglucanase termed EG1 (a molecular weight of about 48 kDa) in an atypical white rot basidomycete \textit{Volvariella volvacea} with the activity (0.1479 \(\mu\)mol D-glucose/mg protein.min) determined with filter paper that was lower by about 46% compared with GH5-p4818Cel5\_2A activity on Avicel from the Wang et al. (2019) study \[45\]. Clearly, while the exocellulase activities have been characterized to be the highest levels reported so far, to the best of our knowledge, the endocellulase activities are shown to be relatively low and are limiting cellulolic degradation within the white rot basidomycote fungi.

On the other hand, very few β-1,4-endoglucanase genes are shown to be present in the brown rot basidomycote genome \[112\]. The characterized brown rot basidomycote \textit{Postia placenta} β-1,4-endoglucanases are associated with very low activities. Specifically, the GH5-p4818Cel5\_2A activity on Avicel from the Wang et al. (2019) study \[45\] was about 57-fold higher than the activity (0.0045 \(\mu\)mol D-glucose/mg protein.min) determined
with Avicel for the GH5 processive endoglucanase Cel5A in the brown rot basidomycete Postia placenta [118]. Several lines of evidence suggest that extracellular reactive oxygen species (ROS), specifically hydroxyl radicals produced from the Fenton reaction, play a compensatory biodegradative role in extensively disrupting crystalline cellulose in wood biomass in concert with their relatively weak and lack of activities of processive and non-processive endocellulases in the white and brown rot basidomycete fungi [112,119].

Last but not least, the white rot and the brown rot basidomycete fungi possess unique lignin-degradation enzymes [112]. Hence, low and lack of endocellulase, especially the processive endocellulase activities, are the recognized limiting enzymatic factors, and these basidomycete fungi are using ROS as recognized compensatory mechanisms in engaging effective lignocellulosic degradation of wood biomass materials. Giving the fact that lignin polymers further encapsulate microfibril units and reduce porosity of raw lignocellulose biomass materials, possession of lignin-degradation enzymes by these fungi would facilitate their cellulases in effectively penetrating into the internal microfibrillar surface for effective hydrolysis.

The most studied aerobic cellulolytic microorganism is the fungus, Hypocrea jecorina, originally also called Trichoderma viride, and is renamed Trichoderma reesei to honor Drs. Reese and Mandels for their contributions in isolating this fungus during World War II [49]. Trichoderma reesei fungi produce the two most abundant exocellulases Cel7A (CBHI, making about 70% of the total cellulase proteins secreted) and Cel6A (CBHII, making about 10% of the total cellulase proteins secreted), as well as seven endoglucanases including Cel7B (EGI, the most abundant EG), Cel5A (EGII), Cel12A (EGIII, serving as expansin activity), Cel61A (EGIV), Cel45A (EGV), Cel5B and Cel61B [49]. Variable cellulase activities have been reported for various Trichoderma reesei likely due to different research conditions such as enzyme purity, substrates and assay temperature. The two Trichoderma reesei exocellulases Cel7A (CBHII) and Cel6A (CBHII) activities were reported to be 0.0091 and 0.0085 µmol D-glucose/mg protein.min, respectively, determined with filter paper [120]. However, these two Trichoderma reesei exocellulases Cel7A (CBHI) and Cel6A (CBHI) activities were reported to be 0.370 and 0.330 µmol D-glucose/mg protein.min, respectively, determined with Avicel [121].

Another aerobic fungus used for industrial cellulase production is Humicola insolens that produces a similar set of multiple endocellulases and exocellulases but with a much lower enzyme protein yield and specific activities [49,50]. Trichoderma reesei and Humicola insolens are not closely related, however, both organisms are brown rot fungi and do not degrade lignin [49]. This is likely due to the fact that these two brown rot fungi have lost their lignin degradation genes during their evolution process [108]. Tambor et al. [122] screened 55 fungal endocellulases out of several fungal species by comparing their specific activities with the Trichoderma reesei Cel5A (EGII) activity at 0.0290 µmol D-glucose/mg protein.min measured with filter paper. The top three selected fungal endocellulases included the Aspergillus niger ApCel5A activity at 0.120, Gloeophyllum trabeum GtCel12A at 0.030 and Sporotrichum thermophile StCel5A at 0.024 µmol D-glucose/mg protein.min by using filter paper [122]. By comparison, GH5-p4818Cel5_2A activity on Avicel by Wang
et al. [45] was about 80% higher than the *Aspergillus niger* endocellulase ApCel5A activity, six-fold higher than the *Gloeophyllum trabeum* endocellulase GtCel12A, and eight-fold higher than the *Sporotrichum thermophile* endocellulase StCel5A, respectively.

Therefore, it can be concluded from this review that while cellulases from the three main aerobic fungi *Trichoderma reesei*, *Humicola insolens* *Aspergillus niger* are useful and the predominant commercial cellulases on the current industrial enzyme markets largely for textile, detergent, wood pulp & paper polishing and food processing, these cellulases are clearly limited in their applications as exogenous feed enzyme additives for livestock feeding industries and as an industrial cellulases for producing commercial scale of lignocellulosic ethanol from raw feedstock biomass materials. This is largely due to the fact that these fungal cellulases are not processive endocellulases, need multiple enzyme synergism to carry out effective degradation, and are bi-modular, relatively large in their molecular weights and sizes, and thus less effective and efficient in breaking down raw lignocellulose composite in plant nature of feeds and feedstock.

7. Microbial Processive Cellulase Properties of Hydrolyzing Natural Crystalline Substrates

The chemically modified soluble cellulose substrates, such as CMC and HEC, are widely used for biochemically characterizing various endocellulases and exocellulases. However, cellulase hydrolytic activities measured with CMC and HEC are not necessarily related to and cannot be simply extrapolated to their potential activities on pure crystalline or amorphous cellulose substrates, pretreated lignocellulose biomass and natural raw lignocellulose biomass materials [4] (the Braunschweig Brenda Enzyme Database at http://www.brenda-enzyme.org). It is a standard research practice to further characterize and compare the hydrolytic potential of cellulases on native lignocellulose biomass by using model insoluble cellulose substrates. Major model insoluble cellulosic substrates include nearly pure celluloses such as cotton fibre, Whatman No. 1 filter paper, bacterial cellulose, bacterial microcrystalline cellulose (BMCC), microcrystalline cellulose (Avicel) and Sigmacell (20, 50 and 100) as well as impure cellulose-containing substrates such as dyed celluloses, α-cellulose and various pretreated lignocellulosic substrates (e.g., Solka-Floc®) [4,124].

There are some recognized intrinsic differences in the physiochemical properties among the major model insoluble cellulosic substrates including crystallinity index, cellulose size as marked by their DP and intraparticulate pore size [4,19]. Crystallinity index (CrI) differs among the model insoluble native cellulosic substrates and variability of the CrI values largely reflects different intramolecular and intermolecular hydrogen bonding among polymer cellulose glucan molecules within each of the elementary fibrillar bundles as well as different techniques used for measuring the CrI values [4,96,106]. There are inconclusive views regarding the relationship between hydrolysis rates and CrI values in the major model insoluble cellulosic substrates. Weimer and Weston [96] and Hall et al. [125] showed that CrI values negatively predicted hydrolysis rates of the major model insoluble cellulose substrates, whereas Fierobe et al. [126] and Zhang et al. [4] concluded that substrate recalcitrance and hydrolysis rates were not strongly associated with their CrI values but reflected their overall accessibility of reactive sites to cellulases. Intrinsic differences among cellulases involved in the studies such as mechanisms of enzyme actions might have been responsible for the differences in the role of CrI values in the afore-discussed major model insoluble cellulosic substrate hydrolysis studies.

Since lignin and celluloses are essentially removed in the major model insoluble cellulosic substrates, intraparticulate porosity (radius of 1–10 nm) reflects free space or the capillary structure between each of the elementary fibrillar bundles in the major insoluble cellulosic substrates [19,96,107]. Avicel (PH101, PH102 and PH105) products are processed from wood pulp with removal of the amorphous fraction and are thus associated with intermediate CrI values (0.5–0.6) but relatively small DP values (DP of 150–500) [4]. The main Avicel (PH101, PH102 and PH105) products are further characterized with an intraparticulate pore peak radius between 2 and 3 nm [94]. Thus, model insoluble cellulosic
substrates such as filter paper (Whatman #1) and Avice are widely used to characterize and compare hydrolysis rates of endocellulases and exocellulases for their potential commercial applications while Avicel is a more suitable substrate for assessing exocellulase activity due to its much smaller chain length.

It is important to note that the GH5-tCel5A1 processive endoglucanase reported by Basit and Akhtar [127] and the GH5-p4818Cel5_2A processive endoglucanase by Wang et al. [45] had a specific activity (Table 1, 0.216 vs. 24.8 μmol D-glucose/mg protein.min) in hydrolyzing the model insoluble crystalline cellulosic substrate Avicel. Specific cellulase activities are expressed and converted into the international unit (1 IU = μmol D-glucose per mg pure enzyme protein per min) in some of the well characterized bacterial endocellulases reported in the literature. The GH5-p4818Cel5_2A endoglucanase activity on Avicel by Wang et al. [45] was about 20–215 folds higher than the range of activities (0.0010–0.0105 μmol D-glucose/mg protein.min) determined with filter paper for the six Thermobifida fusca endoglucanases and exocellulases [120]. The GH5-p4818Cel5_2A endoglucanase activity on Avicel was about 8– to 19-fold higher than the activities (0.011 and 0.024 μmol D-glucose/mg protein.min) determined with Sigmacell-20 for the termite saliva endogenous endoglucanase [128]. While the GH5-p4818Cel5_2A endoglucanase activity on Avicel was about 30% higher than the activity (0.1696 μmol D-glucose/mg protein.min) measured with Avicel for the Clostridium cellulolyticum cellulosomal GH9 processive endoglucanase CelF [129], its activity was about 36– to 119-fold higher than the activities (0.0018–0.0059 μmol D-glucose/mg protein.min) measured with Avicel for a total of 13 other Clostridium cellulolyticum cellulosomal GH9 endoglucanases [130]. The GH5-p4818Cel5_2A endoglucanase activity on Avicel was about 18% higher than the activity (0.1828 μmol D-glucose/mg protein.min) determined with Avicel for the Clostridium thermocellum noncellulosomal processive endoglucanase GH9B (Cell) [131], and was about 431-fold higher than the activity (0.0005 μmol D-glucose/mg protein.min) measured with bacterial microcrystalline cellulose (BMCC) for the Clostridium thermocellum cellulosomal non-processive endoglucanase GH9C (CbhA) [132]. The GH5-p4818Cel5_2A endoglucanase activity on Avicel in this study was about 9– to 97-fold higher than the range of activities (0.0022–0.0218 μmol D-glucose/mg protein.min) determined with Avicel for the three Saccharophagus degradans GH5 processive endoglucanases of Cel5G, Cel5H and Cel5J [133]. GH5-p4818Cel5_2A endoglucanase activity on Avicel by Wang et al. [45] was three-fold lower than the activity (0.1817 μmol D-glucose/mg protein.min) measured with Avicel (PH105) for the Bacillus subtilis GH5 endoglucanase Bscel5 [75], however, the Bacillus subtilis Bscel5 was a classic non-processive endoglucanase. The GH5-p4818_2A endoglucanase activity on Avicel in this study was about 24– and 71-fold higher than the activities (0.0030 and 0.0088 μmol D-glucose/mg protein.min) measured with Avicel for novel GH5 endoglucanases obtained from vermicompost and sugarcane soil metagenomes [61,134]. Thus, this truncated GH5-tCel5A1 reported by Basit and Akhtar [127] and the porcine gut GH5-p4818Cel5_2A endoglucanase reported by Wang et al. [45] had an excellent hydrolytic activity on pure crystalline cellulose substrates in comparison with a number of bacterial endocellulases reported in the literature.

On the other hand, we should also compare the GH5-tCel5A1 and the GH5-p4818Cel5_2A processive endoglucanase properties, as reported by Basit and Akhtar [127] and Wang et al. [45], with the following two thermophilic (i.e., 45–122 °C) processive endocellulases of bacterial origins that are highly regarded for their potential industrial applications with high endocellulase activities in hydrolyzing crystalline cellulosic substrates in the current literature. The Clostridium Phytofermentans GH9 processive endoglucanase Cel9 was cloned and characterized to have about 11-fold higher specific activity (2.488 μmol D-glucose/mg protein.min) determined with Avicel by Zhang et al. [135] than GH5-p4818Cel5_2A. Clostridium Phytofermentans is an obligately anaerobic, mesophilic (i.e., 20–45 °C), cellulosolytic Gram-positive bacterium strain (i.e., ISDgT = ATCC 700394T) isolated from forest soil by Warnick et al. [136]. Clostridium Phytofermentans is also regarded as biofuel bacterium and a model microorganism for consolidated bioprocessing (CBP), since it can utilize a broad range of carbon sources for generation of ethanol and other
metabolites in one step [78,135]. However, the Clostridium Phytofermentans GH9 processive endoglucanase Cel9 is a thermophilic cellulase, because it has an optimal peak temperature at 65 °C and its activity drops down sharply at 40 °C [135]. The researchers believed that the mesophilic Clostridium Phytofermentans might have acquired the thermophilic GH9 processive endoglucanase Cel9 via the horizontal gene transfer mechanism [78]. Liao et al. [76] further reported that the Clostridium Phytofermentans endoglucanase Cel9 only had a specific activity at 0.3464 µmol D-glucose/mg protein.min determined at 37°C with Avicel, which may limit its potential industrial applications in some areas such as exogenous enzyme additives. Furthermore, the Clostridium Phytofermentans endoglucanase Cel9 has a multi-modular structure and thus a relatively large molecular weight and dimensional size for penetration into internal microfibril surface of raw and pretreated biomass materials, which may be the additional disadvantage associated with the GH5-tCel5A1 and the GH5-p4818Cel5_2A processive endoglucanases for their industrial applications.

Zverlov et al. [103] cloned and characterized a multimodular free cellulase CelA from the extremely thermophilic cellulolytic bacterium Caldicellulosiruptor bescii (previously called Anaerocellum thermophilum) with a specific activity at 0.2391 µmol D-glucose/mg protein.min with Avicel, and 1.6174 µmol D-glucose/mg protein.min with xylan, respectively, determined at 72 °C driven by its bi-modal endo-acting and exo-acting processivity and the bi-functional hydrolytic property. The Caldicellulosiruptor bescii CelA was determined to have an optimal peak temperature at 85 °C for hydrolyzing Avicel and this activity dropped down sharply to zero at about 45 °C [103], which may limit its potential industrial applications in some areas such as serving as an exogenous feed enzyme additive in livestock feeding. The Caldicellulosiruptor bescii CelA had a large molecular weight of about 230 kDa [103], and an estimated effective size of 10–35 nm [77], thus potentially limiting its effective diffusion into microfibril internal surface for biomass hydrolysis. It was shown that the Caldicellulosiruptor bescii CelA was effective in excavating extensive cavities on the surface of cellulosic substrates, thus increasing its effective hydrolytic areas [77]. In their Science paper, Brunecky et al. [77] demonstrated that the Caldicellulosiruptor bescii CelA was effective in enzymatically converting both pretreated and raw biomass materials at 75 and 85 °C. Berlin [88] commented that hydrolytic activities of this Caldicellulosiruptor bescii CelA showing that now there were no enzymatic barriers for effective cellulose breakdown. Berlin [88] further viewed that two other advantages associated with hydrolysis of biomass through the Caldicellulosiruptor bescii CelA activities at the high temperatures of 75 and 85 °C included the inhibition of potential bacterial contamination and fluidity of fermentation broth. However, maintaining the high temperatures of 75 and 85 °C to achieve the optimal Caldicellulosiruptor bescii CelA activities will likely be too expensive at the large-scale commercial level for manufacturing lignocellulosic ethanol. On the other hand, the Caldicellulosiruptor bescii CelA cellulase was originally characterized by Zverlov et al. [103] and the Clostridium Phytofermentans bacterium expressing the endoglucanase Cel9 was originally isolated and patented by Warnick et al. [136]. Hence, considering the high activity in hydrolyzing insoluble cellulose substrates and other ideal enzymatic properties such as the mesophilic or thermophilic, optimal pH and a small molecular weight and size, these recently characterized GH5-tCel5A1 and the GH5-p4818Cel5_2A, as reported by Basit and Akhtar [127] and Wang et al. [45], are outstanding processive endoglucanases candidates for various potential industrial applications.

8. Discovered and Characterized Mono-Modular Processive Endoglucanase Properties

The multi-functional endoglucanases GH5-Cel5A and variants were originally identified via targeted genome sequencing of the extremely thermophilic Thermotoga maritima bacterium grown on glucan- and mannan-based polysaccharides [137]. Basit and Akhtar [127] further demonstrated truncation of GH5-Cel5A further enhanced its activity towards cellulosic substrates and GH5-Cel5A and its truncated variants such as GH5-tCel5A1 were monomodular processive endoglucanases. More recently, Wang et al. [45] identified and characterized a mono-modular GH5-p4818Cel5_2A processive endoglucanase in one step [78,135]. However, the Clostridium Phytofermentans GH9 processive endoglucanase Cel9 is a thermophilic cellulase, because it has an optimal peak temperature at 65 °C and its activity drops down sharply at 40 °C [135]. The researchers believed that the mesophilic Clostridium Phytofermentans might have acquired the thermophilic GH9 processive endoglucanase Cel9 via the horizontal gene transfer mechanism [78]. Liao et al. [76] further reported that the Clostridium Phytofermentans endoglucanase Cel9 only had a specific activity at 0.3464 µmol D-glucose/mg protein.min determined at 37°C with Avicel, which may limit its potential industrial applications in some areas such as exogenous enzyme additives. Furthermore, the Clostridium Phytofermentans endoglucanase Cel9 has a multi-modular structure and thus a relatively large molecular weight and dimensional size for penetration into internal microfibril surface of raw and pretreated biomass materials, which may be the additional disadvantage associated with the GH5-tCel5A1 and the GH5-p4818Cel5_2A processive endoglucanases for their industrial applications.

Zverlov et al. [103] cloned and characterized a multimodular free cellulase CelA from the extremely thermophilic cellulolytic bacterium Caldicellulosiruptor bescii (previously called Anaerocellum thermophilum) with a specific activity at 0.2391 µmol D-glucose/mg protein.min with Avicel, and 1.6174 µmol D-glucose/mg protein.min with xylan, respectively, determined at 72 °C driven by its bi-modal endo-acting and exo-acting processivity and the bi-functional hydrolytic property. The Caldicellulosiruptor bescii CelA was determined to have an optimal peak temperature at 85 °C for hydrolyzing Avicel and this activity dropped down sharply to zero at about 45 °C [103], which may limit its potential industrial applications in some areas such as serving as an exogenous feed enzyme additive in livestock feeding. The Caldicellulosiruptor bescii CelA had a large molecular weight of about 230 kDa [103], and an estimated effective size of 10–35 nm [77], thus potentially limiting its effective diffusion into microfibril internal surface for biomass hydrolysis. It was shown that the Caldicellulosiruptor bescii CelA was effective in excavating extensive cavities on the surface of cellulosic substrates, thus increasing its effective hydrolytic areas [77]. In their Science paper, Brunecky et al. [77] demonstrated that the Caldicellulosiruptor bescii CelA was effective in enzymatically converting both pretreated and raw biomass materials at 75 and 85 °C. Berlin [88] commented that hydrolytic activities of this Caldicellulosiruptor bescii CelA showing that now there were no enzymatic barriers for effective cellulose breakdown. Berlin [88] further viewed that two other advantages associated with hydrolysis of biomass through the Caldicellulosiruptor bescii CelA activities at the high temperatures of 75 and 85 °C included the inhibition of potential bacterial contamination and fluidity of fermentation broth. However, maintaining the high temperatures of 75 and 85 °C to achieve the optimal Caldicellulosiruptor bescii CelA activities will likely be too expensive at the large-scale commercial level for manufacturing lignocellulosic ethanol. On the other hand, the Caldicellulosiruptor bescii CelA cellulase was originally characterized by Zverlov et al. [103] and the Clostridium Phytofermentans bacterium expressing the endoglucanase Cel9 was originally isolated and patented by Warnick et al. [136]. Hence, considering the high activity in hydrolyzing insoluble cellulose substrates and other ideal enzymatic properties such as the mesophilic or thermophilic, optimal pH and a small molecular weight and size, these recently characterized GH5-tCel5A1 and the GH5-p4818Cel5_2A, as reported by Basit and Akhtar [127] and Wang et al. [45], are outstanding processive endoglucanases candidates for various potential industrial applications.
canase from the porcine hindgut microbiome. Soluble substrates CMC, HEC and purified barley or oat β-glucans that are also associated with a viscoelastic property are widely used to characterize an endo-β-1,4-glucanase activity [74,138]. Since some exocellulases may also have considerable hydrolytic activities to these soluble substrates, simultaneous measurements of reduction in viscosity in one of the above soluble substrate solutions over a short period of incubation time are a traditional approach of distinguishing endocellulolytic activity from exocellulolytic activity [120,132]. Substantial activities of these Cel5A and GH5-p4818Cel5_2A on CMC, HEC and purified barley β-glucan, as summarized in Table 1; and the time course of dramatic reduction in specific viscosity in the CMC solution were consistent with the data previously reported on three purified Thermobifida fusca endoglucanases of Cel5A (E5), Cel6A (E2) and Cel9A (E4) [120], the Clostridium thermocellum noncellulosomal endoglucanase GH9B (Cell) [131], the Clostridium thermocellum cellulosomal GH9C (CbhA) [132], the Saccharophagus degradans GH5 endoglucanase Cel5H [133], and the Clostridium Phytofermentans GH5 endoglucanase CpCel5B [77]. Meanwhile, hydrolytic activities of exocellulases were not effective in causing reduction in specific viscosity in the test CMC solutions as clearly demonstrated in some of the studies [120,132,133]. On the other hand, the non-detectable activities of the GH5-p4818Cel5_2A on larminarin (β-1,3-1,6-glucans) and curdlan (β-1,3-glucans) excluded the possibility that GH5-p4818Cel5_2A could have the endo-β-1,3-glucanase (i.e., laminarinases, EC 3.2.1.39) and/or the endo-β-1,3-1,4-glucanase (i.e., lichenases, EC 3.2.1.73) activities [45,139] (the Braunschweig Brenda Enzyme Database at http://www.brenda-enzyme.org). Thus, the substrate specificity data in comparison with literature reports collectively suggest that these recently characterized Cel5A1 by Basit and Akhtar [127] and GH5-p4818Cel5_2A by Wang et al. [45] are the family-5 glycosyl hydrolase (GH5) β-1,4-endoglucanases.

There are two types of processive cellulases: exocellulases and processive endocellulases [73]. Degree of processivity for exocellulases is usually determined through calculating the ratio of cellobiose to cellotriose based upon quantitative chromatographic analysis of hydrolytic products of a target exocellulase [73,140–142]. The processive nature of endocellulases can be verified by calculating the processivity value, i.e., the ratio of the amount of total soluble reducing sugars to the amount of total insoluble reducing sugar ends, which is usually measured with pure crystalline cellulose substrates such as filter paper over a number of hours incubated with target endocellulases [73,141,142]. The following substrate hydrolytic data in comparison with literature reports support the notion that the GH5-p4818Cel5_2A reported by Wang et al. [45] is a novel processive endoglucanase. Firstly, the pattern of dramatic reduction in specific viscosity in the CMC solution by Wang et al. [45] was similar to the patterns reported for the Clostridium thermocellum noncellulosomal GH9 endoglucanase GH9B (Cell) [131] and the Saccharophagus degradans GH5 endoglucanase Cel5H [133]. Both the Clostridium thermocellum noncellulosomal GH9 endoglucanase GH9B (Cell) and the Saccharophagus degradans GH5 endoglucanase Cel5H were shown to be processive endoglucanases with cellobiose as the main hydrolytic end product [131,133]. Arguably, the carboxymethyl substitution on cellulose in the CMC may prevent cellulases in hydrolyzing CMC in a processive manner, thus questioning the suitability of CMC as a substrate for processivity study [131]. Secondly, pNP-cellobioside has been recognized as a chromogenic substrate to measure exocellulase activity [74]. Kinetic activities of this cellulase GH5-p4818Cel5_2A on pNP-cellobioside, as reported by Wang et al. [45], suggest that GH5-p4818Cel5_2A, as a recognized endoglucanase, had an exocatalytic property. However, McGrath and Wilson [132] cautioned that activity on pNP-cellobioside was a function of where and how well this synthetic substrate compound could bind into the active site of an enzyme and would not predict whether an enzyme was an exocellulase or endocellulase.

Determination of the processivity value with purified cellulose substrates can help define the processive nature of endocellulases and this has been demonstrated in the past in several studies. The processivity value was determined with filter paper (Whatman #1) for the Thermobifida fusca processive endoglucanase Cel9A (E4) (at 7.0) [120], for the Clostridium
thermocellum noncellulosomal processive endoglucanase GH9B (CelI, at about 50) [131], and for the three Saccharophagus degradans GH5 processive endoglucanases of Cel5G, Cel5H and Cel5J (at 4.0–4.6) [133]. The processivity value was determined with phosphoric acid-swollen cellulose (PASC) for the Clostridium cellulolyticum cellulosomal GH9 processive endoglucanase CelF (at 19.0) [129]. The processivity value was increased from 2.4 to 3.9 when the incubation was prolonged from 0.5 to 3.5 h with the soluble amorphous cellulose substrate RAC for the Clostridium Phytofermentans GH9 processive endoglucanase Cel9 [135]. The processivity value was also seen increasing from 3.6 to 8.6 when the incubation was prolonged from 0.5 to 24 h with filter paper for the processive GH5 endoglucanase EG1 in an atypical white rot basidiomycete Volvariella volvacea [117]. Meanwhile, the processivity value was also determined with filter paper for the Thermobifida fusca non-processive endoglucanase Cel6A (E2, at 2.1 and 2.6) [120,133] and the Clostridium thermocellum cellulosomal non-processive endoglucanase GH9C (CbhA, at 1.9) [132]. Watson et al. [133] reviewed that the processivity values, measured with filter paper for the extensively characterized Thermobifida fusca processive endoglucanase Cel9A, typically ranged between 3.0 and 7.0 for processive endoglucanases. GH5-p4818Cel5_2A and GH5-p4818Cel5_2A were determined to have a processivity value of 10 and 4.6, respectively [45,127]. Thus, determination of processivity values is affected by types of substrates and duration of incubation time used.

9. Properties of Processive Endocellulases in Hydrolyzing Pre-Treated Cellulosic Substrates

While Avicel is widely used as a pure crystalline cellulose substrate with estimated CrI at 0.5–0.6, it is a microcrystalline cellulose with relatively much small molecular weights (DP of 150–500 glucose units) in comparison with other crystalline cellulose substrates such as Whatman No. 1 filter paper (DP of 750–2800 glucose units) and Solka-Floc® (DP of 750–1500 glucose units) [4]. Thus, GH5-p4818Cel5_2A activity reported by Wang et al. [45] was also determined at 0.127 μmol D-glucose/mg protein.min by using Solka-Floc® and this activity value was about 70% lower than the value measured with Avicel (Table 1). Differences in porosity and molecular weights, as reflected by their different DP values, were likely responsible for this discrepancy in the GH5-p4818Cel5_2A activities measured between these two crystalline cellulose substrates. Due to the fact that biological pretreatment of biomass materials is not well established, various physicochemical pre-treatments of biomass materials have been established and optimized to release semi-pure pretreated cellulosic substrates for potential industrial scale of saccharification and subsequent fermentation of ethanol [10,24]. Physicochemical pre-treatments of biomass feedstock are expensive and represent about 21% of the total expense of producing lignocellulosic ethanol [24]. It should be pointed out that Solka-Floc® is a semi-pure crystalline cellulose product commercially available from the International Fibre Corporation. Solka-Floc® has properties comparable to pretreated cellulosic substrates including CrI values (CrI for both 0.4–0.7) and DP (DP of 750–1500 vs. 400–1000 glucose units) [4]. Thus, the GH5-p4818Cel5_2A activity determined by using Solka-Floc® reported by Wang et al. [45] should represent specific activity of this novel enzyme in hydrolyzing pretreated crystalline cellulose substrates.

Amorphous disordered insoluble celluloscs and dissolved soluble celluloscs are different entities, which is different from crystalline celluloscs in terms of physicochemical properties and rates of hydrolysis by celluloscs for industrial applications [124]. While needing to be further optimized for potential industrial scale of applications, both amorphous insoluble cellulose and soluble cellulose substrates have been developed and widely used in cellulase activity tests [4,143]. Acid swollen cellulose, primarily in the form of phosphoric acid swollen cellulose (PASC), is a typical insoluble amorphous disordered cellulose with CrI at 0.00–0.04 and DP of 100–1000 glucose units [19,124]. Regenerated amorphous celluloscs (RAC) are soluble amorphous cellulose substrates and have been generated by using concentrated phosphoric acid and cellulose solvents [143], which represents
further treated readily degradable next generation cellulosic substrates for commercial ethanol production [76,143]. The tCel5A1 processive endoglucanase reported at 24.8 µmol D-glucose/mg protein.min by using RAC at 70 °C but not at the physiological pH by Basit and Akhtar [127]. The GH5-p4818Cel5_2A activity reported by Wang et al. [45] was determined at 1.39 µmol D-glucose/mg protein.min by using RAC at 50 °C and this activity was about 5.4-fold higher than the value measured with Avicel in this study (Table 1). This result of GH5-p4818Cel5_2A activity on RAC was consistent with previous studies in showing that endocellulases with activities on crystalline celluloses would have much amplified hydrolytic activities in hydrolyzing amorphous cellulose substrates frequently demonstrated with PASC or RAC [118,120,131,133]. Hence, these tCel5A1 and the GH5-p4818Cel5_2A endoglucanase activities determined by RAC, as reported by Basit and Akhtar [127] and Wang et al. [45], would suggest that this novel enzyme is very effective in degradation of amorphous cellulosic substrates that are derived from further advanced physicochemical processing in the bioethanol industry.

10. Potential of Processive Endoellulases as Newly Emerging Exogenous Fibre Enzymes for Food Animal Nutrition

To reveal the potential of these tCel5A1 and the GH5-p4818Cel5_2A processive endoglucanase, as reported by Basit and Akhtar [127] and Wang et al. [45], as potentially emerging exogenous feed enzyme additives for livestock feeding especially in cattle, it is imperative to review current understanding of ruminal cellulolytic enzymology and cellulase activities. Two lines of evidence suggest ruminal microbial degradation of lignocellulosic feeds is not very efficient in ruminant livestock species, especially in the highly productive Holstein dairy cows and double muscle beef cattle, and this has been a major sustainable development concern on the global stage. Firstly, about 1 billion tonnes of grains, including wheat, corn, barley, oats, rye, millet and sorghum, accounting for an estimated 70% of grains produced by developed countries and one-third or more of the world’s cereal grains produced, are used in feeding livestock annually for global animal production, which could be directly used to feed some 3.5 billion humans to resolve the world hunger crisis [144]. Furthermore, 40% of feed grains or 400 million tonnes of grains are used annually in ruminants—mainly in cattle [144]. Biologically speaking, this is due to the fact that ruminal microbial degradation of lignocellulosic feeds is not fast enough in ruminants, especially Holstein dairy cows and beef cattle that have been selected and bred for high milk yield and lean growth. Thus, large proportions of processed grains need to be included in rations for these cattle [145,146]. Secondly, it has been well documented that fermentation of feed cellulose in rumen is not optimal as supported by the fact that a significant amount of fermentable fibre is recovered in feces of ruminants [30]. This less optimal digestive utilization of lignocellulosic fibre feeds in combination with grain inclusion in rations & diets has resulted in poor efficiency of carbon and energy utilization in livestock production sectors, accounting for 14.5% of the human-induced greenhouse gas emissions and exceeding that from transportation [144], and becoming a major sustainability concern [34,147]. Selinger et al. [148] reviewed that exogenous feed enzymes including cellulases produced by recombinant DNA technology would be powerful for enhancing efficiency and cutting costs in livestock production. Adeola and Cowieson [32] reviewed that the global exogenous enzyme market value was estimated at about $600 million with about 60% of this market share being dedicated to the single exogenous enzyme product phytase in monogastric feeding and nutrition. Hemicellulases such as β-gucanase and xylanases were recognized as some of the other exogenous feed enzyme products while cellulases had very negligible market share [32]. Fan [34] and Meale et al. [33] concluded that current exogenous feed cellulase additive products were not efficacious in showing their synergy and sufficient activity in breaking down cellulose in vivo, suggesting new cellulase enzyme products need to be developed for the potentially huge global market.

Enzymatically, several lines of evidence have pointed out that ruminal microbial cellulases, especially in bovine species, are not likely to be self-sufficient to carry out fast, efficient and compete degradation of natural crystalline lignocelluloses associated with
fibre feeds. Firstly, Denman et al. [80] and Wang et al. [149] isolated and characterized a GH6 endocellulase celA with very high activities at 9.7 and 1.1 μmol D-glucose/mg protein.min, respectively, measured by using Avicel from the ruminal anaerobic fungus Neocallimastix patriciarum. Liu et al. [150] isolated and characterized high GH5 endocellulase celA activity at 1.4 μmol D-glucose/mg protein.min by using Avicel from the ruminal anaerobic fungus Piromyces rhizinflata. Indeed, these reported ruminal anaerobic fungal endocellulase activities were several folds higher than our GH5-p4818Cel5_2A endoglucanase activity in hydrolyzing Avicel. However, anaerobic fungi only represent a small population (5–10%) of the total microbes in rumen [151]. Secondly, these ruminal anaerobic fungal endocellulases are not processive in nature, thus cannot lead to efficient degradation of crystalline cellulose into soluble cellodextrin in vivo. Krause et al. (2003) were the first to recognize and propose that the lack of expression of GH6 and GH7 exo-acting cellulases was likely the primary enzymatic factor limiting efficient degradation of crystalline cellulose in rumen [30]. Intriguingly, more recent work on microbial metagenomic sequencing and the GH gene cataloguing of cow’s ruminal microbiome by Hess et al. (2011) [41] confirmed the original view by Krause et al. [30]. No GH6 and only one GH7 cellulase gene were identified out of a total of 2365 cow’s ruminal microbial cellulase genes catalogued in the work conducted by Hess et al. [41]. Also characterized with the cellulase systems in the aerobic fungal Trichoderma reesei and Humicola insolens, the type-I and type-II exo-acting cellulases, i.e., cellobiohydrolases, are largely classified to have GH6 and GH7 catalytic domains, and belong to GH6 and GH7 families [49]. Thus, the metagenomic analysis-based gene cataloguing by Hess et al. [41] would support the notion that there are essentially very little or no exo-acting microbial cellulase genes and their activities in the bovine rumen. On the other hand, a large number of GH5 and GH9 cellulase genes were identified and catalogued from cow’s ruminal microbiome by Hess et al. [41]. It is known that some of the well-characterized processive endocellulases belong to the GH5 and GH9 cellulase families. A large number of studies have been conducted to isolate and characterize ruminal microbial cellulases by using the classic chromatography-based biochemical protein purification [152,153], gene cloning and expression [69,154], organism genome sequencing based cloning and expression [82,151,155], metagenomic library based functional screening [53,156], and gene-centric metagenomic sequencing, cataloguing and expression [41] during the past three decades. Palackal et al. [156] and Duan et al. [157] screened and characterized some ruminal GH5 endocellulases with non-detectable activities on crystalline cellulose substrates. Ferrer et al. [53] characterized a number of novel cellulases with activities on several tested crystalline cellulose substrates; however, they did not classify catalytic domain GH families on these enzymes, and their enzymes were not purified and assayed by using the international unit for direct comparison with other studies. Fibrobacter succinogenes and Eubacterium cellulosolvens are some of the main fibrolytic bacteria in rumen [158]. Yoda et al. [83] cloned and characterized a large bifunctional Eubacterium cellulosolvens GH5 endocellulase cel5A (127 kDa) showing activities on soluble substrate CMC and hemicellulose xylan but did not report activity on crystalline cellulose substrates. Cavicchioli and Watson [155] cloned and characterized a Fibrobacter succinogenes endocellulase with activities on several tested crystalline cellulose substrates; however, they did not classify a catalytic domain GH family on this enzyme, and their enzyme was not purified and assayed by using the international unit for direct comparison with other studies. Qi et al. [154] cloned and characterized two main bifunctional Fibrobacter succinogenes GH5 and GH9 endocellulases of rcCel5H and rcCel9B with activities on soluble substrates CMC and HEC, amorphous substrate Ball-milled cellulose and hemicellulose xylan. While the Fibrobacter succinogenes GH5 endocellulase rcCel5H had non-detectable activity on the crystalline cellulose substrates Avicel and Sigmacel 100, the Fibrobacter succinogenes GH9 endocellulase rcCel9B (67.3 kDa) had very low activities on the crystalline cellulose substrates Avicel (0.05 μmol D-glucose/mg protein.min) and Sigmacel 100 (0.07 μmol D-glucose/mg protein.min) [154]. By comparison, the GH5-p4818Cel5_2A activity on Avicel from this study was about two-fold higher than the GH9 endocellulase rcCel9B
activity from the *Fibrobacter succinogenes* in the study by Qi et al. [154]. Qi et al. [69] further cloned an atypical GH9 gene termed Cel9D (79.4 kDa) in expressing the exocellulase (E.C. 3.2.1.74) activity, and characterized its synergism with other four endocellulases of Cel8B (81.4 kDa), Cel9B (67.3 kDa), Cel45C (37.7 kDa) and Cel51A (119 kDa) that were all cloned from the ruminal *Fibrobacter succinogenes*. By comparison, the GH5-p4818Cel5_2A activity (1 U = µmol D-glucose/min) on Avicel from this study was about 431-fold higher than the atypical GH9 exocellulase Cel9D activity (0.0005 U/mg protein), 307-fold higher than the endocellulase Cel8B activity (0.0007 U/mg protein), 215 fold higher than the endocellulase Cel9B activity (0.0010 U/mg protein), 1079-fold higher than the endocellulase Cel45C activity (0.0002 U/mg protein) and 359 fold higher than the endocellulase Cel51A activity (0.0006 U/mg protein) from the ruminal *Fibrobacter succinogenes* in the study by Qi et al. [69]. Thus, as compiled in Table 1, in comparisons with the GH5-tCel5A1 and the GH5-p4818Cel5_2A endoglucanase activities as reported by Basit and Akhtar [127] and Wang et al. [45], cattle ruminal bacterial endocellulase and exocellulase activities reported so far in the literature are considerably much lower.

Arguably, highly active processive endocellulases may still exist in cow’s ruminal microbes till all the catalogued 2365 cows’ ruminal cellulase genes with large numbers of the GH5 and GH9 genes are fully expressed and characterized. Aspeborg et al. [159] reviewed that GH5 genes represent one of the largest glycoside hydrolase family genes within the 113 GH families. As summarized by Aspeborg et al. [159], while a few of the GH5 subfamilies did not encode for endocellulase enzyme activities, a number of the GH5 genes subfamilies have encoded GH proteins that have lost their catalytic machinery, indicating evolution towards other novel functions. Interestingly, studies by Duan et al. [157] and Liu et al. [160] screened several novel GH5 endocellulases with little or none detectable activities on crystalline cellulose substrates from the buffalo rumen, whereas Bao et al. [54] screened a novel GH5 endocellulase with high activity (0.20 U/mg protein) on Avicel from the yak rumen, and Nguyen et al. [55] screened a novel GH5 endocellulase with very high activity (2.5 U/mg protein) on crystalline cellulose Avicel from the buffalo rumen metagenome. Last but not least, the presence of lignin polymers that wrap around the surface of individual microfibrillar units encapsulates the microfibrillar contents of hemicelluloses and celluloses from effective enzyme penetration and attack. The negative effect of lignin on fibre degradation has been well documented in animal nutrition [145,146,158,161].

Nevertheless, it can be concluded that the overall ruminal microbial cellulase activities, particularly processive endocellulase and exocellulase activities towards crystalline celluloses, are low and are limited to allow fast and efficient lignocellulolytic degradation. The present commercial cellulase additives that are primarily derived from the aerobic fungal *Trichoderma reesei*, *Humicola insolens* and *Aspergillus niger* have been viewed to be less efficacious as exogenous enzyme additives likely due to the fact these enzymes have to work in groups and proper ratios for optimal synergism and activities, which prevent them from effectively penetrating into internal microfibrillar surfaces for their actions in fibre feeds high in lignin. Thus, commercialization of GH5-tCel5A1 and the GH5-p4818Cel5_2A as emerging exogenous endocellulase additives will have the potential to greatly improve efficiency of plant cell wall based fibre feed utilization in livestock feeding sectors, particularly in the cattle industry.

### 11. Multifunctionality of the Newly Characterized Processive Endocellulases

The enzyme activity results summarized in Table 1 have shown that apart from hydrolysis on crystalline, amorphous and soluble celluloses, tCel5A1 and the GH5-p4818_2A could effectively hydrolyze four other hemicelluloses & hemicellulose complex substrates. The GH5-p4818Cel5_2A activity in hydrolyzing xyloglucans was 57%, much higher than its endocellulase activity in hydrolyzing Avicel has important industrial applications [45]. Firstly, as one of the main branched structural and principal storage hemicellulose polysaccharides, xyloglucans are present in all land plants and comprise between 20 and 50% of the plant cell wall or seed dry mass [162]. Secondly, xyloglucans are β-(1,4)-linked D-glucose
backbone heavily substituted with α-(1,6)-xylose-linked side-chain polymers. Xyloglucans are not linear polymers but branched-structure polymers. Xyloglucan-specific hydrolytic activities are carried out by three enzyme activities of xyloglucan-specific β-(1,4)-endo-xyloglucanase activity (EC 3.2.1.151), β-(1,4)-exo-xyloglucanase activity (EC 3.2.1.155) and oligoxyloglucan reducing-end-specific cellbiohydrolase activity (EC 3.2.1.150) [162]. The GH5-p4818Cel5_2A activity on xyloglucans by measuring reducing sugar yield would at least suggest that this endocellulase has significant EC 3.2.1.151 activity. Further studies are needed to determine the hydrolytic end-products of the tCel5A1 and GH5-p4818Cel5_2A activity on xyloglucans and to further characterize if this endocellulase also possesses the exo-acting mode of EC 3.2.1.155 and EC 3.2.1.150 activities following the procedures of Eklöf et al. [162]. Xyloglucan-specific hydrolytic activities have been demonstrated in several members of GH families including GH5, 7, 9, 12, 44 and 74 [162]. Vincken et al. [163] demonstrated that Trichoderma viride but not Trichoderma reesei cellulases had hydrolytic activities towards Xyloglucans. The previously reported xyloglucan-specific hydrolytic activities in two GH5 enzymes were a GH5_4 subfamily with xyloglucan-specific β-(1,4)-endo-xyloglucanase activity (EC 3.2.1.151), and/or β-(1,4)-exo-xyloglucanase activity (EC 3.2.1.155) with the enzyme genes being cloned and characterized from Paenibacillus sp. [159,164]. These Paenibacillus sp. GH5 xyloglucanases were only active towards xyloglucans and not towards Avicel, CMC, barley glucan or xylan [160]. However, Cel5A1 and variants did not show detectable xyloglucan-specific hydrolytic activity [137]. Thus, GH5-p4818Cel5_2A is likely among the first reported multifunctional monomodular processive endcellulases being associated with xyloglucan-specific hydrolytic activity.

Furthermore, GH5-p4818Cel5_2A activities on the purified birch wood xylan as well as Kojac glucomannan and locust bean gum, i.e., galactomannan that was also a branched structure polymer, whereas GH5-Cel5A1 variants had activities on Kojac glucomannan and locust bean gum, i.e., galactomannan; these findings would suggest that the GH5-p4818Cel5_2A endocellulase have the β-(1,4)-endo-xylanase activity (EC 3.2.1.21) and the β-(1,4)-endo-mannanase activity (EC 3.2.1.78) [45], and GH5-Cel5A1 variants have the β-(1,4)-endo-mannanase activity [137]. Given the fact that both GH5-Cel5A1 variants and GH5-p4818Cel5_2A also have exo-acting activity, further measurements need to be conducted to examine if GH5-p4818Cel5_2A has xylan β-(1,4)-xylosidase activity (EC 3.2.1.37) and β-(1,4)-mannosidase activity (EC 3.2.1.25) and if GH5-Cel5A1 variants also have β-(1,4)-mannosidase activity.

On the other hand, the lack of activity by GH5-Cel5A1 variants and GH5-p4818Cel5_2A on colloidal chitin would suggest that these endocellulases had no chitosanase activity (EC 3.2.1.149) [45,137]. Moreover, the lack of GH5-p4818Cel5_2A activity on the purified arabinobinoxylans would indicate that this branched structure polymer substrate could not be acted on by this enzyme, while the lack of GH5-p4818Cel5_2A activity on the purified arabinan from this study would also suggest that GH5-p4818Cel5_2A had no arabinan endo-(1,5)-α-L-arabinase activity (EC 3.2.1.99) and no α-L-arabinofuranosidase activity (EC 3.2.1.55) [45].

Some GH5 endocellulases with dual substrate specificity such as concurrent activities on glucan-mannan or glucan-xylan have been well documented [159,165]. Pereira et al. [166] reported that the hyperthermophilic Thermotoga maritima GH5 endoglucanase Cel5A had activities on soluble CMC, barley β-glucan, xyloglucan, lichenin, galactomannan and glucomannan; however, it had no activities on important plant cell wall substrates of crystalline celluloses and xylan. Vlasenko et al. [167] suggested that the GH5 multifunctional properties likely descended from an ancestral enzyme gene through bioinformatic phylogenetic analyses. Cohen et al. [118] and Qi et al. [69] reported that GH5 processive endoglucanase Cel5A in the brown rot basidomycete Postia placenta and the ruminal Fibrobacter succinogenes GH9D had concurrent activities on crystalline cellulose and xylan. The bi-modal endo-acting and exo-acting processive and bi-functional thermophilic Caldicellulosiruptor bescii endocellulase CelA were shown to hydrolyze both crystalline cellulose and xylan [103], which feature was highly regarded as an outstanding potential
Appl. Sci. 2021, 11, 5150

industrial cellulase for saccharification without required physicochemical pretreatment of feedstock [77, 88]. However, the *Clostridium thermocellum* noncellulosomal GH9, the *Clostridium Phytofermentans* GH9 and the *Saccharophagus degradans* GH5 processive endoglucanases were all characterized to only hydrolyze cellulose substrates [78, 131, 133].

The current commercial cellulase enzyme products that are primarily derived from the aerobic fungal species of *Trichoderma reesei*, *Humicola insolens*, *Aspergillus niger* and basidomycete *Irpen lacteus* do not possess hydrolytic properties beyond crystalline celluloses, CMC, barley glucans and/or xylan. To the best of our knowledge, the GH5-Cel5A1 variants and particularly GH5-p4818Cel5_2A have been the only processive endocellulase characterized so far to have multiple substrate specificity in hydrolyzing multiple plant cell wall substrates including crystalline cellulose, pretreated cellulose and β-glucan, as well as hemicelluloses of xylan, xyloglucan, glucomannan and galactomannan.

Under this context, it is known that hemicelluloses are important components of plant cell wall materials embedding elemental fibril crystalline cellulose core for each of microfibrillar units, reducing porosity, and hampering cellulosomes from further penetration into microfibrillar internal surfaces for effective hydrolysis. Physicochemical pretreatment of plant cell wall biomass materials for removal of hemicellulosic accounts for up to 22% of the total projected cost of producing commercial scale of lignocellulosic ethanol [24]. Harris et al. [168] reviewed that the current trend in the developing second-generation lignocellulose and hemicellulose ethanol industry is to use of pretreatment of lower severity by combining biochemical conversion technology in order to lower the capital cost and also to minimize waste generation. For example, the first commercial scale second-generation fuel ethanol plant, Beta Renewables plant in Crescentino, Italy, utilizes only steam in its pretreatment without chemical addition [168]. Uses of hemicellulases following this steam pretreatment would further increase porosity, constituent soluble sugar yields and fluidity of the pretreated feedstock slurry, since β-glucans, xyloglucans, xylan and arabinoxylans are very viscous [168]. Vincken et al. [169] showed that endocellulases, exocellulases and xyloglucanases together could exert considerable synergy in degradation of raw plant cell wall materials. Thus, these novel GH5-tCel5A1 and GH5-p4818Cel5_2A endocellulases with unique multi-functional hemicellulase activities of hydrolyzing crystalline celluloses and hemicelluloses will have an excellent potential as an exogenous feed enzyme additive for livestock feeding and as a highly efficient biocatalyst to enable competitive lignocellulosic ethanol production. Further laboratory scale of tests and field experiments need to be conducted to examine the efficacy of these novel GH5-tCel5A1 and GH5-p4818Cel5_2A endocellulases in hydrolysis of livestock feed and biofuel feedstock samples.

12. Optimal Biochemical and Physiological Properties of the Newly Characterized Processive Endocellulases

These novel GH5-tCel5A1 and GH5-p4818Cel5_2A endocellulases had pH optimum in a light acidic to neutral pH environment of 5.0–7.0 with the majority (about 60%) of its activity retained for a wide range of pH, ranging from very acidic pH of 4 to alkaline pH of 8.5 (Table 1) [45, 127]. These results were consistent with the pH optimum and effects of pH on endocellulase activity response patterns previously reported in three novel GH5 cellulases screened and characterized out of bovine rumen metagenomes [54, 55, 157] and rabbit cecum [56]. These GH5 endocellulase pH optimum and effects of pH on the endocellulase activities would reflect adaptive evolution of these endocellulases and their host bacteria to the ruminal, gastric, small and large intestinal luminal pH environment and the more alkaline pH environment in the middle and the distal regions of the small intestine in porcine, monogastric herbivore and bovine species. The light acidic to neutral pH optimum at 6–7 and the tolerance to a wide range pH environment for these novel GH5-tCel5A1 and GH5-p4818Cel5_2A are valuable biochemical properties which enable its potential wide range of industrial applications as an exogenous feed enzyme additive, saccharification of feedstock for production of lignocellulosic ethanol as well as a number of other applications in light industries such as textile, detergent, food processing and wood and paper pulp polishing.
The GH5-p4818Cel5_2A endocellulase had optimal temperature at 50 °C with majority (above 75%) of its endocellulase activity retained between 37 and 50 °C, which represented a marginal overlap between mesophilic (20–45 °C) and thermophilic (50–75 °C) temperature zones, whereas tCel5A1 endocellulase activity was shown as optimal at the thermophilic (50–75 °C) temperature zone [45,127]. Feng et al. [56] reported several novel cellulases with optimal temperature between 40 and 55 °C screened and characterized out of rabbit cecal metagenome. Bao et al. [54] reported a novel GH5 cellulase with an optimal temperature at 60 °C screened and characterized out of yak rumen metagenome, and Nguyen et al. [55] reported a novel GH5 cellulase with an optimal temperature at 50 °C screened and characterized out of buffalo rumen metagenome. It is known that normal mammalian body core temperature is at about 37 °C and ruminal digesta core temperature is at about 39 °C [54]. The mesophilic forest soil bacterium *Clostridium Phytofermentans* GH9 processive endocellulase was a thermophilic enzyme with an optimal temperature at 65 °C [78]. Horizontal gene transfer might have played a role for the discrepancy of the optimal temperatures and the inhabiting environment temperatures among these endocellulases to which the aforementioned hosting bacteria adapted. These novel GH5-tCel5A1 and GH5-p4818Cel5_2A with an optimal temperature at 50–70 °C and their major endocellulase activities retained between 37 and 50 °C will be suitable as an exogenous feed enzyme additive in a wide range of livestock feeding, also including pelleted monogastric fish, poultry and swine diets.

For non-herbivore monogastric animal species including swine, poultry and fish, a much higher thermostability is usually required for exogenous enzymes such as phytases and endocellulases because diets for these animals are commonly pelleted at high temperature (60–80 °C) [170]. Thus, a biochemical property of thermostability, as measured by residual activity after being heated at 80 °C for 10 min, will be essentially ideal for GH5-p4818Cel5_2A endocellulase to be potentially used as an exogenous enzyme additive for non-herbivore monogastric animal species [170]. On the other hand, while saccharification of feedstock with cellulases following pretreatment has been tested and developed for potential lignocellulosic ethanol production at various temperatures such as 37, 45 and 50 °C [10,76,171], a higher thermophilic temperature ranging at 60–65 °C is considered to be advantageous for commercial lignocellulosic ethanol production [88,172,173]. Thus, thermostable cellulase such as the GH5-tCel5A1, as reported by Basit and Akhtar [127], reflecting this required hyperthermophilic temperature, would be desirable for potential commercial lignocellulosic ethanol production.

Two approaches are typically used to explore these desired thermophilic cellulases, including (i) the screening of novel and highly active cellulases from thermophilic (50–75 °C) and hyperthermophilic (80–120 °C) through targeted genomic sequencing the extremely thermophilic microorganisms such as *Thermotoga maritima* for the Cel5A1 and its variants [137,174]; and (ii) improving of selected mesophilic cellulases through mutagenesis-based enzyme protein engineering [175,176] and/or glycosylation modification-based enzyme property modification [177]. The first approach, represented by the Cel5A1, may be limited due to the fact that the selected thermostable cellulases that are expressed and screened out of thermophilic & hyperthermophilic microbes will most likely have optimal temperature within the thermophilic temperature zone, which will not be suitable as exogenous enzyme additives in livestock feeding sectors. For example, the *Clostridium Phytofermentans* GH9 thermophilic processive endocellulase with an optimal temperature at 65 °C, as reported by Zhang et al. [78], is not ideal to serve as an exogenous enzyme additive in livestock feeding, since this enzyme will lose about 60–70% of its maximal activity once fed through the animals' digestive tract system at a temperature ranging from 37 °C in mammals to 41 °C in poultry. The hyperthermophilic *Caldicellulosiruptor bescii* bi-functional processive endocellulase CelA was effective in enzymatically converting pretreated and raw biomass materials at 75–85 °C, as demonstrated by Brunecky et al. [77], however, this enzyme would not be very active in animals' digestive system as a good exogenous enzyme additive in livestock feeding. Thus, upon further optimization for improving its
thermostability, reaching 75–85 °C through effective strategies of enzyme protein engineering and glycosylation modifications, this GH5-p4818Cel5_2A processive endocellulase, reported by Wang et al. [45], has the potential to be used for various industrial applications at mesophilic, thermophilic and hyperthermophilic temperature conditions. Further research needs to be conducted to examine the efficacy of this hyperthermophilic processive endocellulase GH5-tCel5A1 and its variants, as reported by Basit and Akhtar [127], as exogenous cellulases for food animal feeding.

The property of protease resistance of cellulases is generally not an immediate concern for some industrial applications such as lignocellulosic ethanol production, textile and wood and paper pulp bioprocessing. In first-generation grain ethanol production, proteases are used to maximize starch hydrolysis and sugar yields by dismantling starch-gluten complexes [168]. Perspectively, the property of protease resistance of cellulases will also likely become important for commercial lignocellulosic ethanol production, since inclusion of proteases in biochemical based pretreatment and subsequent enzymatic saccharification of feedstocks would potentially reduce interferences of nitrogenous compounds during cellulose and hemicellulose breaking down associated with plant cell wall materials. The protease resistance feature is important to consider for novel cellulases that are to be used for animal feed industrial applications and other food processing applications associated with the presence of significant levels of proteases. As reported by Wang et al. [45], GH5-p4818Cel5_2A endocellulase activities on CMC demonstrated a remarkable resistance to the proteolytic activities (5000 U/mL test media) of trypsin, whereas the GH5-p4818Cel5_2A endocellulase activities on CMC were dramatically decreased after 30 min of incubations in the presence of chymotrypsin activity (200 U/mL test media). Trypsin and chymotrypsin activities in small intestinal fluids were reported to range from 43 to 45 U/mL, and from 6.1 to 7.9 U/mL, respectively [178]. Trypsin and chymotrypsin activities used in the experiments by Wang et al. [45] were much higher than these protease activities reported in growing pigs in the literature [178]. Thus, it may be concluded that GH5-p4818Cel5_2A would be somewhat resistant to trypsin activity. Further research needs to be pursued to investigate the gastrointestinal stability of promising novel cellulases for their potential development as efficacious exogenous feed fibre enzymes in the agricultural food animal industry.

Both trypsin and chymotrypsin are endopeptidases. Trypsin specifically cleaves Arg and Lys residual based peptide bonds, while chymotrypsin preferentially hydrolyzes the aromatic amino acid residue Trp based peptide bonds as well as Tyr, Phe, Leu and Met based peptide bonds [38]. With this context, it should be pointed out that glycoside hydrolases such as endocellulases, which are ubiquitous and typically exhibit a tunnel or a cleft catalytic site, are lined with aromatic Trp residues for processing carbohydrates [86]. GH5-p4818Cel5_2A was of bacterial origins and was expressed in competent non-pathogenic E. coli cells, thus this GH5-p4818Cel5_2A endocellulase enzyme protein was not likely glycosylated. Thus, this GH5-p4818Cel5_2A endocellulase with Trp residues in its catalytic sites without glycosylation are, in principle, very much vulnerable to proteolytic activity by chymotrypsin. On a separate note, a much higher level of chymotrypsin activity load (by about 111-fold) in the study by Wang et al. [45] in comparison with the levels reported in the Fang et al. [178] study was also likely responsible for the dramatic decline pattern of the GH5-p4818Cel5_2A endocellulase activities shown by Wang et al. [45]. On the other hand, it was reported that in cecal digesta, a high residual trypsin activity was observed, while negligible residual chymotrypsin activity was seen [179]. Thus, the lack of resistance to chymotrypsin activity by this GH5-p4818Cel5_2A endocellulase reported by Wang et al. [45] was also likely due to the fact the bacteria that expressed this cellulase had evolved to adapt to the cecal environment.

Akiba et al. [180] reported that a fungal Aspergillus niger endocellulase was highly resistant to proteases. Morgavi et al. [181] shown that several fungal cellulases were resistant to proteolytic activities that were of bovine ruminal microbial origins and addition of non-enzyme sources of proteins could further dilute the ruminal proteolytic effects on
the fungal endocellulase activities. Bai et al. [182] reported a novel GH9 beta-glucanase that was cloned and characterized from thermoadidophilic *Alicyclobacillus* sp. and was highly resistant to various proteases when this enzyme was expressed in the yeast *Pichia pastoris* likely in a glycosylated form. Nguyen et al. [55] reported that a novel GH5 cellulase from buffalo rumen metagenome was highly resistant to proteases. Beckham et al. [177] viewed that O-glycosylation imparts protease resistance for the fact that glycans presumably block proteases access to the protein backbone. However, pepsin is an important animal gastric protease and its effects on cellulase activities are not well reported in the literature. Woyengo et al. [183] reported that porcine gastric pepsin activity ranged from 143 to 266 PU (pepsin activity unit)/mL digesta fluid. Further efforts should be made to examine effects of pepsin, trypsin and chymotrypsin activities on both GH5-tCel5A1 and GH5-p4818Cel5_2A endocellulase activity in the presence feed proteins under close to normal physiological conditions such as at the porcine physiological body temperature of 37°C, a typical gastric pH for testing the pepsin activity effect and a typical small intestinal luminal pH for testing the trypsin and chymotrypsin activity effects. The expression of both GH5-tCel5A1 and GH5-p4818Cel5_2A endocellulases of the bacterial origins in the major commercial fungal protein production platforms such as *Trichoderma reesei*, *Humicola insolens*, *Aspergillus niger* and yeast *Pichia pastoris* and the subsequent expression effects on GH5-tCel5A1 and GH5-p4818Cel5_2A endocellulase glycosylation and protease resistance should be further investigated for future commercial implications.

13. Small Molecular Size and Monomodular Structure of the Newly Characterized Processive Endocellulases

These novel GH5-tCel5A1 and GH5-p4818Cel5_2A processive cellulases were characterized to be a relatively small monomodular enzyme protein with a molecular weight of about 37 and 42 kDa [45,127]. Erickson [93] viewed that most single modular proteins would fold into one globular domain and polypeptides larger than 50 kDa typically form two or more domains. Thus, small molecular weight monomodular cellulases likely form various spheric shapes with a different radius size while bi-modular, larger molecular weight cellulases tend to show ellipsoid and tadpole shapes or other irregular shapes [97,98]. As shown by Wang et al. (2019), the predicted GH5-p4818Cel5_2A endocellulase structure resembles a globular shape. Wang et al. [45] reported that GH5-p4818Cel5_2A was estimated to have a diameter of 4.6 nm calculated according to the model by Erickson [93]. Thus, GH5-tCel5A1, as reported by Basit and Akhtar [127], would have even a much smaller diameter compared with the GH5-p4818Cel5_2A. Carpita et al. [18] reported that pore size diameters in various plant cell wall materials ranged between 3.5 and 5.2 nm. Several studies documented that physiochemically pretreated lignocellulosic biomass materials and crystalline cellulose model substrates were associated with much larger pore diameters, ranging between 5.1–11.0 nm [94–97]. Thus, these GH5-tCel5A1 and GH5-p4818Cel5_2A processive cellulases should be effective in penetrating into microfibrillar internal surfaces of raw biomass materials such as fibre feeds and feedstocks because of its small molecular weight and size.

Very few endocellulases have been reported in the current literature with the combined biochemical features of having a single module, a small molecular weight and size and being highly active in hydrolyzing crystalline celluloses and/or hemicelluloses. Cohen et al. [118] reported a novel single modular processive GH5 cellulase Cel5A with a molecular weight at 42 kDa from the brown rot basidiomycete *Postia placenta*, which had about 57-fold lower activity on Avicel in comparison with the novel GH5-p4818Cel5_2A endocellulase reported by Wang et al. [45]. Watson et al. [133] reported that the *Saccharophagus degradans* GH5 processive endoglucanase Cel5H had a molecular weight of 68 kDa and was bi-modular. Bao et al. [34] reported a novel single modular processive GH5 cellulase with a predicted molecular weight of 40 kDa from yak rumen metagenome with activity (0.200 U/mg protein) on Avicel as well as hydrolysis of β-glucans, which is, in our review, the only reported endocellulase with a small molecular weight and activity in hydrolyzing crystalline celluloses comparable to this novel GH5-p4818Cel5_2A. However,
it should be pointed out that this yak ruminal processive GH5 cellulase only had activities on β-1,3/1,4-glucans including soluble and insoluble celluloses and had no activities towards other hemicelluloses [54]. Nguyen et al. [55] reported a novel GH5 cellulase with a molecular weight of 60 kDa from buffalo rumen metagenome without specifying if this enzyme had a distinct second CBM. Several other well characterized GH9 cellulases were also bi-modular, including the Thermobifida fusca processive endoglucanase Cel9A with a molecular weight of 90.2 kDa [120], the Clostridium thermocellum noncellulosomal GH9 processive endocellulase with a molecular weight of 93 kDa [131], the ruminal Fibrobacter succinogenes exocellulase GH9D with a molecular weight of 76 kDa [69], and the Clostridium Phytofermentans GH9 thermophilic processive endocellulase with a molecular weight of 104.9 kDa [135]. Moreover, a bi-modular, bi-functional endo-acting and exo-acting processive thermophilic Caldicellulosiruptor bescii endocellulase CelA was characterized with a large molecular weight of 230 kDa by Zverlov et al. [103]. This thermophilic Caldicellulosiruptor bescii endocellulase CelA was further characterized to have an estimated effective size of 10–35 nm, and was highly regarded as an outstanding potential industrial cellulase for digesting raw feedstocks because of its unique mechanism of cellulose degradation [77].

Therefore, to the best of our knowledge, these novel GH5-tCel5A1 and GH5-p4818Cel5_2A processive endocellulases are among the only reported processive endocellulases documented so far that are featured to be a monomodular and a small molecular weight endocellulase with high activity on crystalline celluloses and multi-functional hydrolytic activities towards hemicelluloses, and should be highly penetrating in effectively digesting raw biomass materials such as fibre feeds and feedstocks as ideal industrial biocatalysts.

14. Structure and Functionality of the Newly Characterized Processive Endocellulases

Unique functionality of enzyme proteins is supported by their unique structures. It has been well documented that substrate specificity and the mode of action of various glycosyl hydrolases are governed by exquisite details of their three-dimensional structures rather than their global fold with three active-site topologies of (i) a pocket or crater; and (ii) a cleft or groove; and (iii) a tunnel [184]. These unique GH5-tCel5A1 and GH5-p4818Cel5_2A processive cellulases have three major unique biochemical properties within one enzyme molecule in comparison with other endocellulases reported in the literature. Firstly, GH5-tCel5A1 and GH5-p4818Cel5_2A endocellulase are small monomodular enzyme proteins with a molecular weight (below 50 kDa) and relatively high activity on crystalline celluloses [45,127]. Only two other reported studies in the literature have so far shown that an endocellulase with a single module and such a comparable small molecular weight and size could hydrolyze crystalline cellulose substrates, including the study by Cohen et al. (2005) [118] on the Cel5A from the brown rot basidomycete Postia placenta; the study by Bao et al. (2011) [54] on a yak ruminal GH5 cellulase. Secondly, both GH5-tCel5A1 and GH5-p4818Cel5_2A endocellulases have exocellulase activity on crystalline celluloses thus are a processive endocellulases. Only two other reported studies in the literature have shown that a processive endocellulase without a CBM could hydrolyze crystalline cellulose substrates, including the study by Cohen et al. [118] on the Cel5A from the brown rot basidomycete Postia placenta, and the study by Bao et al. [54] on a yak ruminal GH5 cellulase. Classic processive endocellulases are all primarily GH9 cellulases and they all contain a CBM, including the Clostridium thermocellum noncellulosomal endoglucanase GH9B cellulase GH9B [131], the Thermobifida fusca processive endoglucanase GH9A (E4) [185], and the Clostridium Phytofermentans processive endoglucanase Cel9 [135]. Structural analyses by Sakon et al. [185] show that the presence of a CBM was essential for the processive nature of these GH9 endocellulases. However, Watson et al. [133] reported that the presence of CBM was not essential for the processive property in the Saccharophagus degradans GH5 processive endoglucanases. Third, apart from hydrolyzing crystalline celluloses, both GH5-tCel5A1 and GH5-p4818Cel5_2A endocellulases have been shown to have activities on other hemicelluloses, i.e., β-glucans, xylan, xyloglucans, mannan, glucomannan and...
galactomannan. To the best of our knowledge, no other cellulases have been reported with such wide substrate specificity in the literature.

Sequence alignments of several different GH5 processive cellulases, including both GH5-tCel5A1 and GH5-p4818Cel5_2A, had been previously illustrated in the supplemental Figure-S5 in the study by Wang et al. (2019) [45]. Attempts have been made to predict the structure of this GH5-p4818Cel5_2A endocellulase with the SWISS-MODEL online server by Wang et al. [45]. This predicted TIM barrel fold of this cellulase GH5-p4818Cel5_2A is consistent with the TIM barrel motif superfamily previously reported in the GH5 endocellulase CelA from *Clostridium cellulolyticum* [186]. A cleft or groove structure was apparently seen in this GH5-p4818Cel5_2A endocellulase as was shown in the *Clostridium cellulolyticum* GH5 endocellulase CelA by Ducros et al. [186] and reviewed by Davies and Henrissat [184]. High activity of GH5-p4818Cel5_2A on the branched substrate xyloglucan and locust bean gum, as summarized in Table 1, would suggest the presence of a cleft active site in this cellulase structure, since branched substrates such as xyloglucans and locust bean gum could not effectively fit into a tunnel-like or a pocket active site topology [162]. Yaoi et al. [187] further demonstrated that a cleft active site was also essential to have exo-mode hydrolysis of xyloglucans by an endocellulase to engage the hydrolytic action. The double Glu residues as the proton donor in the putative catalytic acid/base and nucleophile are at Glu136, and are consistent with the strictly conserved residues within GH5 catalytic domain functions [186]. Intriguingly, a long tunnel-like active site with multiple Trp residues is clearly visualized along the enzyme protein surface and is located from the C-terminus of the β-strands or the TIM barrel [45]. This predicted long tunnel-like active site topology for this novel GH5-p4818Cel5_2A has provided additional evidence that this novel GH5-p4818Cel5_2A has exo-acting mode, and thus is a processive GH5 endocellulase [45]. All exocellulases are processive [73]. Exocellulases have a tunnel-like active site topology [72,188,189]. However, Sakon et al. [185] shown that the processive GH9A (E4) endocellulase *Thermobifida fusca* had a cleft and not a tunnel-like active site topology, and a CBM was essential for this enzyme’s activity on crystalline celluloses. McGrath and Wilson [132] discussed that substrate binding or the formation of cellulase-substrate complexes was essential to characterize cellulase structure and catalytic relationship. Giving the multi-functional properties of this novel GH5-p4818Cel5_2A endocellulase, crystallographic ensembles to be derived from multiple substrate binding states should be generated and used to study its functional structural variations [190]. The structure and functionality of have been revealed in details for GH5-tCel5A1 [127,191]. Thus, it is very likely that GH5-p4818Cel5_2A will be unique with a combined cleft and a tunnel-like active site topology in its structure to support its superb multifunctionality, which needs to be resolved in our future studies. Further crystallization and structural characterization of this GH5-p4818Cel5_2A processive cellulase will facilitate further rationale design-based mutagenesis to improve functionality of these enzymes, such as resistance to mammalian proteases.

To facilitate further effective adoption of these two unique GH5-tCel5A1 and GH5-p4818Cel5_2A processive endocellulases for industrial applications, the following three aspects of research need to be pursued. Firstly, further enhancement of these wild-type of both GH5-tCel5A1 and GH5-p4818Cel5_2A endocellulase specific activities on crystalline celluloses and hemicelluloses will need to be conducted. Harris et al. [168] articulated that innovations at a small-scale through enzyme discovery in the laboratory could exert large-scale impacts when rolled out in industrial applications as demonstrated in commercial ethanol production. Lynd et al. [24] reviewed that enzyme and pretreatment expenses represent about up to 48% of the total cost of commercial lignocellulosic ethanol production. Directed evolution is a very powerful protein-engineering tool for improving enzyme properties such as specific activity and thermostability without in-depth immediate understanding of enzyme protein structure and enzyme-substrate interactions [78]. Given the fact that the GH9 *Clostridium Phytofermentans* processive endocellulase had a large molecular weight (over 100 kDa) and was multi-modular, Ahmad et al. [79] demonstrated this by
using *B. subtilis* as a host and applying directed evolution for improving the processive 
GH9 endocellulase activity by about two-fold from the *Clostridium Phytofermentans* on the 
solid cellulosic substrate. Efforts ought to be made to further enhance both GH5-tCel5A1 and 
GH5-p4818Cel5_2A endocellulase activity by directed evolution and using *B. subtilis* 
as a host.

Upon resolving detailed structural and functional relationships for GH5-p4818Cel5_2A in 
future, a rational design approach can also be undertaken to further optimize GH5-
tCel5A1 and GH5-p4818Cel5_2A endocellulase properties [175,177,192], considering their 
small molecular size and monomodular features. Secondly, these unique GH5-Cel5A1 and 
GH5-p4818Cel5_2A endocellulases are currently over-expressed in non-pathogenic *E. coli* 
competent cells for laboratory characterization, and advantages of these competent cells for 
characterization of novel enzymes have been reviewed by Uchiyama and Miyazaki [193]. 
Various engineered *E. coli* cells, including an endotoxin-free *E. coli* strain, referred to as 
the ClearColi® BL21 (DE3) as developed by Mamat et al. (2015), are widely used for 
commercial scale of pharmacological polypeptide & enzyme product production [194]. Use 
of *E. coli* competent cells for commercial scale of industrial biomass degradation enzyme 
production is being promoted [195]; however, governmental regulatory body approval 
would be needed if resulting biomass enzymes are to be used as exogenous fibre enzyme 
feed supplements. Several microbial platforms have been developed and used for 
industrial and biorefinery based commercial enzyme production via the solid state fermentation 
(SSF) or submerged fermentation (SmF) types of microbial cultivation, including *Tricho-
derma reesei* and/or *Trichoderma longibrachiatum* by Novozymes, Genencor-Danisco-Dupont, 
Dyadic International, AB Vista Enzymes, Biocatalysts Limited, and the formerly Iogen Bio-
products (purchased by Novozymes in 2013); *Aspergillus niger* by Novozymes and Amano 
Enzymes Inc.; and *Bacillus sp.* by Maps Limited, and Specialty Enzymes & Biotechnologies 
Co. (SEB) [50]. A logic next step of research would be to engage biological engineering 
and expressing of these two unique GH5-tCel5A1 and GH5-p4818Cel5_2A endocellulases 
into one of the above commercial platforms of microbes, i.e., *Trichoderma reesei, Trichoderma 
longibrachiatum, Aspergillus niger* or *Bacillus sp.*, by working with one or more of the major 
industrial enzyme companies for their potential commercial production.

The concept of consolidated bio-processing (CBP) of de-polymerization of structural 
carbohydrates into their monomeric constituents and the fermentation of hexose and 
pentose sugars into ethanol in a single step in one platform microorganism has been 
proposed and applied for production of lignocellulosic ethanol and other biochemicals. 
Lynd et al. [24] reviewed that an effective CBP alone could account for about 41% of the total 
cost of commercial lignocellulosic ethanol production. Several microbial platforms have 
also been researched and developed for potential industrial applications, including the 
gineered CBP organism *Clostridium Phytofermentans* (ATCC 700394) [196], the engineered 
CBP organism *Bacillus sp.* [76] and the ethanol fermenting yeasts such as *Pichia stipites* and 
Saccharomyces cerevisiae [50,197]. Thus, direct engineering of optimized versions of these two 
unique GH5-tCel5A1 and GH5-p4818Cel5_2A endocellulases along with a highly active β-
glucosidase gene into the CBP organism *Bacillus sp.* or the *Pichia stipites* or the *Saccharomyces 
cerevisiae* system would potentially allow the development of a cost-effective CBP system 
for commercial production of lignocellulosic ethanol from raw feedstock materials.
### Table 1. The summary and comparison of enzyme activity for some of the characterized processive endoglucanases towards their polymeric substrates.

| Enzyme                  | Source                                      | GH Family | Substrate, Assay Conditions and End Product(s)                  | Specific Activity (µmol D-glucose Equivalent • mg Protein⁻¹ min⁻¹) | Reference               |
|-------------------------|---------------------------------------------|-----------|-----------------------------------------------------------------|------------------------------------------------------------------|-------------------------|
| Characterized GH5 processive endoglucanases with the activity towards crystalline cellulose |                                             | GH5_2     | Avicel (1.0%, w/v), 50 °C, pH 6.0                               | 0.217                 | Wang et al. [45]        |
|                         |                                             |           | RAC (1.0%, w/v), 50 °C, pH 6.0 (G2 and G3)                       | 1.38                  |                         |
|                         |                                             |           | β-Glucan from barley (1.0%, w/v), 50 °C, pH 6.0                 | 56.2                  |                         |
|                         |                                             |           | CMC (1.0%, w/v), 50 °C, pH 6.0                                 | 41.8                  |                         |
|                         |                                             |           | Not tested                                                      |                      |                         |
|                         |                                             |           | CMC (1.0%, w/v), 50 °C, pH 6.5                                 | 9.61                  | Watson et al. [133]     |
|                         |                                             |           | RAC (1.0%, w/v), 50 °C, pH 6.5 (G2 and G3)                      | 11.8                  | Wu et al. [198]         |
|                         |                                             |           | β-Glucan from barley                                            | 117.5                 | Wu et al. [199], Zheng and Ding [117] |
|                         |                                             |           | CMC (2.5%, w/v), 45 °C, pH 6.5                                 | 115.4                 |                         |
|                         |                                             |           | Avicel (1.0%, w/v), 50 °C, pH 6.5                              | 0.62                  |                         |
|                         |                                             |           | RAC (2.5%, w/v), 45 °C, pH 6.5 (G2 and G3)                      | 2.4                   |                         |
|                         |                                             |           | β-Glucan from barley (2.5%, w/v), 45 °C, pH 6.5                | 171.5                 |                         |
|                         |                                             |           | CMC (2.5%, w/v), 45 °C, pH 6.5                                 | 115.4                 |                         |
|                         |                                             |           | Avicel (1.0%, w/v), 50 °C, pH 6.5                              | 0.62                  |                         |
|                         |                                             |           | RAC (1.0%, w/v), 50 °C, pH 7.5 (G2, G3 and G4)                 | 13.1                  |                         |
|                         |                                             |           | β-Glucan from barley                                            | 2.91                  |                         |
|                         |                                             |           | CMC (2.0%, w/v), 50 °C, pH 7.5                                 | 34.9                  |                         |
|                         |                                             |           | Avicel (1.0%, w/v), 65 °C, pH 6.5                              | 0.03                  |                         |
|                         |                                             |           | RAC, (1.0%, w/v), 65 °C, pH 6.5 (G2, G3)                       | 13.02                 |                         |
|                         |                                             |           | β-Glucan from barley                                            | 93.12                 |                         |
|                         |                                             |           | CMC (1.0%, w/v), 65 °C, pH 6.5                                 | 77.19                 |                         |
|                         |                                             |           | Avicel (1.0%, w/v), 50 °C, pH 4.8                              | 0.0045                |                         |
|                         |                                             |           | RAC, (1.0%, w/v), 50 °C, pH 4.8                                | 0.054                 |                         |
|                         |                                             |           | β-Glucan from barley                                            | Not tested            |                         |
|                         |                                             |           | CMC (1.0%, w/v), 50 °C, pH 4.8                                 | 120                   |                         |
|                         |                                             |           | Avicel (0.5%, w/v), 30 °C, pH 6.8                              | Not reported          |                         |
|                         |                                             |           | RAC (0.5%, w/v), 30 °C, pH 6.8 (G2, G3)                        | Not reported          |                         |
|                         |                                             |           | β-Glucan from barley                                            | Not tested            |                         |
|                         |                                             |           | CMC (0.5%, w/v), 30 °C, pH 6.8                                 | 25.8                  |                         |

**GH5: GH5 family, CMC: Carboxymethyl cellulose, RAC: Raw Alover leaf extract, Substrate: Cellulose, Assay Conditions: 50 °C, pH 6.0**
| Enzyme | Source | GH Family | Substrate, Assay Conditions and End Product(s) | Specific Activity (µmol D-glucose Equivalent • mg⁻¹ Protein • min⁻¹) | Reference |
|--------|--------|-----------|-----------------------------------------------|---------------------------------------------------------------|-----------|
| Cel5A (TmaCel5A) (37.6 kD, mono-modular) | Thermotoga maritima (PCR amplification) | GH5 | Avicel (2.0%, w/v), 70 °C, pH 6.0 | 24.8 | Basit and Akhtar [127] |
| | | | RAC (2.0%, w/v), 70 °C, pH 6.0 (G2, G3) | 22.5 | |
| | | | β-Glucan from barley | Not tested | |
| | | | CMC (2.0%, w/v), 70 °C, pH 6.0 | 11.0 | |
| | | | filter paper, β-glucan from barley | Not tested | |
| CpCel9 (104.9 kDa, GH9-CBM3-CBM3) | Clostridium phytofermentans (Amplified from genomic DNA) | GH9 | Avicel (1.0%, w/v), 50 °C, pH 6.0 | 2.48 | Zhang et al. [135] |
| | | | RAC (1.0%, w/v), 50 °C, pH 6.0 | 4.59 | |
| | | | β-Glucan from barley | Not tested | |
| | | | CMC (1.0%, w/v), 50 °C, pH 6.0 | 16.65 | |
| | | | filter paper, β-glucan from barley, RAC (0.25% w/v), 50 °C, pH 5.5 | 0.01 | |
| | | | RAC (0.2%, w/v), 50 °C, pH 5.5 | 2.23 | Irwin et al. [202] |
| | | | β-Glucan from barley | Not tested | |
| | | | CMC (1.0%, w/v), 50 °C, pH 5.5 | 5.25 | |
| | | | filter paper, β-glucan from barley, CMC, 37 °C, pH 6.5 | 0.001 | |
| | | | RAC (0.25%, w/v), 37 °C, pH 6.5 | 0.058 | Qi et al. [154] |
| | | | β-Glucan from barley, 37 °C, pH 6.5 | 0.703 | |

Other GH family characterized typical processive endoglucanases with the activity towards crystalline cellulose

Abbreviations used including: RAC, regenerated amorphous cellulose; G2, cellobiose; G3, cellotriose; and G5, cellopentose; CMC, carboxymethyl cellulose; GH, glycoside hydrolase; and CBM, carbohydrate binding module. The following equation was used for the interconversion of the specific activities between different units: \( A = (B \times MW)/1000 \); \( B = (A \times 1000)/MW \); where \( A \) represents the specific activity in the unit of \( \mu \text{mol glucose equivalent} \times \mu \text{mol}^{-1} \text{protein} \times \text{min}^{-1} \); and \( B \) refers to the specific activity in the unit of \( \mu \text{mol glucose equivalent} \times \text{mg}^{-1} \text{protein} \times \text{min}^{-1} \); and MW represents molecular weight of a concerned enzyme in Dalton.
In summary, while the gene-cloning and species-targeted genome sequencing techniques have been traditionally applied for identifying and characterizing cellulase genes, the metagenomic approaches, including the gene-centric metagenomic analysis and the metagenomic expression library screening, have further widened the opportunities for the discovery and characterization of novel microbial cellulase genes from various environmental resources. Upon reviewing the present literature of cellulase discovery and characterization, two recently characterized unique GH5-tCel5A1 and GH5-p4818Cel5_2A processive endocellulases are a small molecular weight, monomodular and processive β-1,4-endoglucanase with an estimated spherical diameter at 4.6 nm or smaller. These two unique GH5-tCel5A1 and GH5-p4818Cel5_2A processive endocellulases are highly active on crystalline and pre-treated celluloses and have multifunctionality towards several hemicelluloses including β-glucans, xylan, xyloglucans, mannans, galactomannans and glucomannans. Therefore, these two unique GH5-tCel5A1 and GH5-p4818Cel5_2A processive endocellulases have novel structural and functional properties for potential important industrial applications.

Author Contributions: M.Z.F. obtained research funds for support of his research team and their activities and wrote this review paper; W.W. compiled the Table 1; and W.W., L.C., J.C., W.F. and M.W. have contributed to the reviewed processive endocellulase concepts and industrial applications. All authors have read and agreed to the published version of the manuscript.

Funding: Some of the reviewed research and research team members are supported by projects from the Natural Science and Engineering Research Council (NSERC) of Canada, the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA) New Directions & Alternative Renewable Fuels Research Program; the University of Guelph-OMAFRA Partnership Research Program, the University of Guelph Gryphon’s LAAIR (Leading to Accelerated Adoption of Innovative Research) program, Agriculture, Agri-Food Canada (AAFC) Swine Innovation Porc (SIP) Swine Cluster Program, Ontario Pork and the Metagen Enzyme Corporation (to M.Z.F.). W.W. was supported by the grant-in-aid research funding at the University of Guelph from the Metagen Enzyme Corporation. J.C. is also supported by a major Ph.D. scholarship from the China Scholarship Council (CSC) and the University of Guelph Graduate Study International Student Tuition Waiver Program.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: Ming Z. Fan (M.Z.F.) is the co-founder and board chairman of the Metagen Enzyme Corporation, and has an active grant-in-aid research funding at the University of Guelph from the Metagen Enzyme Corporation. Weijun Wang (W.W.) is a board member of the Metagen Enzyme Corporation. The other authors declare that no competing interest exist.

References
1. Scheller, H.V.; Ulvskov, P. Hemicelluloses. Annu. Rev. Plant Biol. 2010, 61, 263–289. [CrossRef] [PubMed]
2. Flint, H.J. Polysaccharide breakdown by anaerobic microorganisms inhabiting the mammalian gut. Adv. Appl. Microbiol. 2004, 56, 89–120. [PubMed]
3. Fan, M. Concepts of Animal Nutrition, 2nd ed.; Pearson: Boston, MA, USA, 2014.
4. Zhang, P.Y.H.; Himmel, M.E.; Mielenz, J.R. Outlook for cellulase improvement: Screening and selection strategies. Biotechnol. Adv. 2006, 24, 452–481. [CrossRef] [PubMed]
5. Himmel, M.E.; Ding, S.-Y.; Johnson, D.K.; Adney, W.S.; Nimlos, M.R.; Brady, J.W.; Foust, T.D. Biomass recalcitrance: Engineering plants and enzymes for biofuels production. Science 2007, 315, 804–807. [CrossRef]
6. Vanholme, R.; Demedts, B.; Morreel, K.; Ralph, J.; Boerjan, W. Lignin biosynthesis and structure. Plant Physiol. 2010, 153, 895–905. [CrossRef]
7. Ding, S.Y.; Liu, Y.S.; Zeng, Y.; Himmel, M.E.; Baker, J.O.; Bayer, E.A. How does plant cell wall nanoscale architecture correlate with enzymatic digestibility? Science 2012, 338, 1055–1060. [CrossRef]
8. Foston, M.; Ragauskas, A.J. Biomass characterization: Recent progress in understanding biomass recalcitrance. Ind. Biotechnol. 2010, 8, 191–208. [CrossRef]
9. Himmel, M.E. Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy; Blackwell Pub.: Oxford, UK, 2008.
10. Mousdale, D.M. Biofuels: Biotechnology, Chemistry, and Sustainable Development; CRC Press: Boca Raton, FL, USA, 2008.
11. Chundawat, S.P.; Beckham, G.T.; Himmel, M.E.; Dale, B.E. Deconstruction of lignocellulosic biomass to fuels and chemicals. *Ann. Rev. Chem. Biomol. Eng.* 2011, 2, 121–145. [CrossRef]

12. Wilson, D.B. Microbial diversity of cellulose hydrolysis. *Curr. Opin. Microbiol.* 2011, 14, 259–263. [CrossRef]

13. Morais, S.; Morag, E.; Barak, Y.; Goldman, D.; Hadar, Y.; Lamed, R.; Shoham, Y.; Wilson, D.B.; Bayer, E.A. Deconstruction of lignocellulose into soluble sugars by native and designer cellulases. *MBio* 2012, 3, e00508-12. [CrossRef]

14. Bissaro, B.; Kommedal, E.; Rahr, Å.K.; Eijssink, V.G.H. Controlled depolymerization of cellulose by light-driven lytic polysaccharide oxygenases. *Nat. Commun.* 2020, 11, 890. [CrossRef]

15. Lynd, L.R.; Weimer, P.; Van Zyl, W.H.; Pretorius, I.S. Microbial cellulose utilization: Fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* 2002, 66, 506–577. [CrossRef]

16. Keeegstra, K. Plant cell walls. *Plant Physiol.* 2010, 154, 483–486. [CrossRef] [PubMed]

17. Pauly, M.; Keegstra, K. Plant cell wall polymers as precursors for biofuels. *Curr. Opin. Plant Biol.* 2010, 13, 304–311. [CrossRef]

18. Carpita, N.; Sabularse, D.; Montezinos, D.; Delmer, D.P. Determination of the pore size of cell walls of living plant cells. *Science* 1979, 205, 1144–1147. [CrossRef]

19. Zhang, Y.-H.P.; Lynd, L.R. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: Noncomplexed cellulase systems. *Biotechnol. Bioeng.* 2004, 88, 797–824. [CrossRef]

20. Poma, A.B.; Chwastyk, M.; Cieplak, M. Coarse-grained model of the native cellulose Iα and the transformation pathways to the Ib allomorph. *Cellulose* 2016, 23, 1573–1591. [CrossRef]

21. Kumar, R.; Singh, S.; Singh, O.V. Bioconversion of lignocellulosic biomass: Biochemical and molecular perspectives. *J. Ind. Microbiol. Biotechnol.* 2008, 35, 377–391. [CrossRef]

22. Galanopoulos, C.; Giuliani, A.; Barletta, D.; Zondervan, E. An integrated methodology for the economic and environmental assessment of a biorefinery supply chain. *Chem. Eng. Res. Des.* 2020, 160, 199–215. [CrossRef]

23. Giuliani, A.; Bari De, I.; Motola, V.; Pierro, N.; Giocoli, A.; Barletta, D. Techno-environmental assessment of two biorefinery systems to valorize the residual lignocellulosic biomass of the Basilicata region. *Math. Model Eng. Probl.* 2019, 6, 317–323. [CrossRef]

24. Lynd, L.R.; Laser, M.S.; Bransby, D.; Dale, B.E.; Davison, B.; Hamilton, R.; Himmel, M.; Keller, M.; McMillan, J.D.; Sheehan, J.; et al. How biotech can transform biofuels. *Nat. Biotechnol.* 2008, 26, 169–172. [CrossRef] [PubMed]

25. Rubin, E.M. Genomics of cellulose biofuels. *Nature* 2008, 454, 841–845. [CrossRef] [PubMed]

26. Wilson, D.B. Cellulases and biofuels. *Curr. Opin. Biotechnol.* 2009, 20, 295–299. [CrossRef]

27. Himmel, M.E.; Xu, Q.; Luo, Y.; Ding, S.-Y.; Lamed, R.; Bayer, E.A. Microbial enzyme systems for biomass conversion: Emerging paradigms. *Biofuels* 2010, 1, 323–341. [CrossRef]

28. Bhat, M.K. Cellulases and related enzymes in biotechnology. *Biotechnol. Adv.* 2000, 18, 355–383. [CrossRef]

29. Beauchemin, K.A.; Colombo, D.; Morgavi, D.P.; Yang, W.Z. Use of exogenous fibrolytic enzymes to improve feed utilization by ruminants. *J. Anim. Sci.* 2003, 81, E37–E47.

30. Krause, D.O.; Denman, S.E.; Mackie, R.I.; Morrison, M.; Rae, A.L.; Attwood, G.T.; McSweeney, C.S. Opportunities to improve fiber degradation in the rumen: Microbiology, ecology, and genomics. *FEMS Microbiol. Rev.* 2013, 27, 663–693.

31. Choct, M. Enzymes for the feed industry: Past, present and future. *World Poult. Sci. J.* 1999, 55, 121–131. [CrossRef]

32. Adeola, O.; Cowieson, A.J. BOARD-INVITED REVIEW: Opportunities and challenges in using exogenous enzymes to improve nonruminant animal production. *J. Anim. Sci.* 2011, 89, 3189–3218. [CrossRef]

33. Meale, S.J.; Beauchemin, K.A.; Hristov, A.N.; Chaves, A.V.; McAllister, T.A. BOARD-INVITED REVIEW: Opportunities and challenges in using exogenous enzymes to improve ruminant production. *J. Anim. Sci.* 2014, 92, 427–442. [CrossRef]

34. Fan, M. Swine nutrition and environment. In *Sustainable Swine Nutrition*; Chiba, L., Ed.; John Wiley & Sons Inc.: Hoboken, NJ, USA, 2013; pp. 365–411.

35. Flint, H.J.; Bayer, E.A.; Rincon, M.T.; Lamed, R.; White, B.A. Polysaccharide utilization by gut bacteria: Potential for new insights from genomic analysis. *Nat. Rev. Microbiol.* 2008, 6, 121–131. [CrossRef]

36. Morrison, M.; Pope, P.B.; Denman, S.E.; McSweeney, C.S. Plant biomass degradation by gut microbiomes: More of the same or something new? *Curr. Opin. Biotechnol.* 2009, 20, 358–363. [CrossRef] [PubMed]

37. Morrison, M. Looking large, to make more, out of gut metagenomics. *Curr. Opin. Microbiol.* 2013, 16, 630–635. [CrossRef] [PubMed]

38. Kidder, D.E.; Manners, M.J. Passage of Food. In *Digestion in the Pig*; Scientechnica Bristol: Bristol, UK, 1978.

39. Balch, C.C. Factors affecting the utilization of food by dairy cows. *Br. J. Nutr.* 1950, 4, 361–388. [CrossRef] [PubMed]

40. Bhatti, S.A.; Bowman, J.G.P.; Firkins, J.L.; Grove, A.V.; Hunt, C.W. Effect of intake level and alfalfa substitution for grass hay on ruminal kinetics of fiber digestion and particle passage in beef cattle. *J. Anim. Sci.* 2008, 86, 134–145. [CrossRef]

41. Hess, M.; Szyrba, A.; Egan, R.; Kim, T.W.; Chokhawala, H.; Schroth, G.; Luo, S.; Clark, D.S.; Chen, F.; Zhang, T.; et al. Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science* 2011, 331, 463–467. [CrossRef]

42. Rideout, T.C.; Liu, Q.; Wood, P.; Fan, M.Z. Nutrient utilisation and intestinal fermentation are differentially affected by the consumption of resistant starch varieties and conventional fibres in pigs. *Br. J. Nutr.* 2008, 99, 984–992. [CrossRef]

43. Lamendella, R.; Domingo, J.W.; Ghosh, S.; Martinson, J.; Oerther, D.B. Comparative fecal metagenomics unveils unique functional capacity of the swine gut. *BMC Microbiol.* 2011, 11, 103. [CrossRef]
71. Jürgensen, J.; Ilmberger, N.; Streit, W.R. Screening for cellulases with industrial value and their use in biomass conversion. Methods Mol. Biol. 2012, 834, 1–16.
72. Wilson, D.B. Processive and nonprocessive cellulases for biofuel production—lessons from bacterial genomes and structural analysis. Appl. Microbiol. Biotechnol. 2012, 93, 497–502. [CrossRef]
73. Wilson, D.B.; Kostylev, M. Cellulase processivity. Methods Mol. Biol. 2012, 908, 93–99.
74. McCleary, B.V.; McKie, V.; Draga, A. Measurement of endo-1,4-β-glucanase. Methods Enzymol. 2012, 510, 1–17.
75. Toda, H.; Nagahata, N.; Amano, Y.; Nozaki, K.; Kanda, T.; Okazaki, M.; Shimosaka, M. Gene cloning of cellobiohydrolase II from the white rot fungus Irpex lacteus MC-2 and its expression in Pichia pastoris. Biosci. Biotechnol. Biochem. 2008, 72, 3142–3147. [CrossRef]
76. Liao, H.; Zhang, X.Z.; Rollin, J.A.; Zhang, Y.H. A minimal set of bacterial cellulases for consolidated bioprocessing of lignocellulose. Biotechnol. J. 2011, 6, 1409–1418. [CrossRef]
77. Brunecky, R.; Alahuhta, M.; Xu, Q.; Donohoe, B.S.; Crowley, M.F.; Kataeva, I.A.; Yang, S.J.; Resch, M.G.; Adams, M.W.; Lunin, V.V.; et al. Revealing nature’s cellulase diversity the digestion mechanism of Caldicellulosiruptor bescii CelA. Science 2013, 342, 1513–1516.
78. Zhang, X.Z.; Zhang, Z.; Zhu, Z.; Sathitsuksanoh, N.; Yang, Y.; Zhang, Y.H. The noncellulosomal family 48 cellobiohydrolase from Clostridium phytofermentans ISDg heterologous expression, characterization and processivity. Appl. Microbiol. Biotechnol. 2010, 86, 525–533.
79. Ahmad, S.; Ma, H.; Akhtar, M.W.; Zhang, Y.-H.P.; Zhang, X.-Z. Directed Evolution of Clostridium phytofermentans glycoside hydrolase family 9 endoglucanase for enhanced specific activity on solid cellulosic substrate. BioEnergy Res. 2013, 7, 381–388. [CrossRef]
80. Denman, S.; Xue, G.P.; Patel, B. Characterization of a Neocallimastix patriciarum cellulase cDNA (celA) homologous to Trichoderma reesei cellobiohydrolase II. Appl. Environ. Microbiol. 1996, 62, 1889–1896. [CrossRef]
81. Bayer, E.A.; Lamed, R.; Morag (Morgenstern), E. Affinity digestion for the near-total recovery of purified cellulosome from Clostridium thermocellum. Enzyme Microb. Technol. 1992, 14, 289–292.
82. Thongekkaew, J.; Ikeda, H.; Masaki, K.; Iefuji, H. Fusion of cellulose binding domain from Trichoderma reesei CBHI to Cryptococcus sp. S-2 cellulase enhances its binding affinity and its cellulolytic activity to insoluble cellulosic substrates. Enzyme Microb. Technol. 2013, 52, 241–246. [CrossRef] [PubMed]
83. Yoda, K.; Toyoda, A.; Mukoyama, Y.; Nakamura, Y.; Minato, H. Cloning, sequencing, and expression of a Eubacterium cellulosolvens S gene encoding and endoglucanase (Cel5A) with novel carbohydrate-binding modules, and properties of Cel5A. Appl. Environ. Microbiol. 2005, 71, 5793–5797. [CrossRef] [PubMed]
84. Duan, C.J.; Feng, J.X. Mining metagenomes for novel cellulase genes. Biotechnol. Lett. 2010, 32, 1765–1775. [CrossRef] [PubMed]
85. Jung, H.; Wilson, D.B.; Walker, L.P. Binding and reversibility of Thermofluida fusca Cel5A, Cel6B, and Cel48A and their respective catalytic domains to bacterial microcrystalline cellulose. Biotechnol. Bioeng. 2003, 84, 151–159. [CrossRef]
86. Payne, C.M.; Bomble, Y.J.; Taylor, C.B.; McCabe, C.; Himmel, M.E.; Crowley, M.F.; Beckham, G.T. Multiple functions of aromatic-carbohydrate interactions in a processive cellulase examined with molecular simulation. J. Biol. Chem. 2011, 286, 41028–41035. [CrossRef]
87. Taylor, C.B.; Payne, C.M.; Himmel, M.E.; Crowley, M.F.; McCabe, C.; Beckham, G.T. Binding site dynamics and aromatic-carbohydrate interactions in processive and non-processive family 7 glycoside hydrolases. J. Phys. Chem. B 2013, 117, 4924–4933. [CrossRef]
88. Berlin, A. Microbiology. No barriers to cellulose breakdown. Science 2013, 342, 1454–1456. [CrossRef]
89. Smith, S.P.; Bayer, E.A. Insights into cellulose assembly and dynamics: From dissection to reconstruction of the supramolecular enzyme complex. Curr. Opin. Struct. Biol. 2013, 23, 686–694. [CrossRef]
90. Vazana, Y.; Morais, S.; Barak, Y.; Lamed, R.; Bayer, E.A. Designer cellulosomes for enhanced hydrolysis of cellulosic substrates. Methods Enzymol. 2012, 510, 429–452.
91. Bayer, E.A.; Chanzy, H.; Lamed, R.; Shoham, Y. Cellulose, cellulases and cellulosomes. Curr. Opin. Struct. Biol. 1998, 8, 548–557. [CrossRef]
92. Artzi, L.; Bayer, E.A.; Morais, S. Cellulosomes: Bacterial nanomachines for dismantling plant polysaccharides. Nat. Rev. Microbiol. 2017, 15, 83–95. [CrossRef]
93. Erickson, H.P. Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. Biol. Proced. Online 2009, 11, 32–51. [CrossRef]
94. Marshall, K.; Sixsmith, D. Some physical characteristics of microcrystalline cellulose 1. Powders for pharmaceutical use. Drug Dev. Ind. Pharm. 1974, 1, 51–71. [CrossRef]
95. Grethlein, H.E. The effect of pore size distribution on the rate of enzymatic hydrolysis of cellulosic substrates. Nat. Biotechnol. 1985, 3, 155–160. [CrossRef]
96. Weimer, P.J.; Weston, W.M. Relationship between the fine structure of native cellulose and cellulose degradability by the cellulase complexes of Trichoderma reesei and Clostridium thermocellum. Biotechnol. Bioeng. 1985, 27, 1540–1547. [CrossRef]
97. Hui, L.; Liu, Z.; Ni, Y. Characterization of high-yield pulp (HYP) by the solute exclusion technique. Bioreour. Technol. 2009, 100, 6630–6634. [CrossRef]
98. Cowling, E.B.; Kirk, T.K. Properties of cellulose and lignocellulosic materials as substrates for enzymatic conversion processes. *Biotechnol. Bioeng. Symp.* 1976, 6, 95–123.

99. Abuja, P.M.; Pilz, I.; Claeyssens, M.; Tomme, P. Domain structure of celllobiohydrolase II as studied by small angle X-ray scattering: Close resemblance to celllobiohydrolase I. *Biochem. Biophys. Res. Commun.* 1988, 156, 180–185. [CrossRef]

100. Abuja, P.; Hayn, M.; Chen, H.; Esterbauer, H. The structure of endoglucanase I (Trichoderma reesei) in solution. *Colloid Interface Sci.* 1975, 93, 181.

101. Xu, Q.; Ding, S.-Y.; Brunecky, R.; Bomble, Y.J.; Himmel, M.E.; Baker, J.O. Improving activity of minicellulosomes by integration of intra- and intermolecular synergies. *Biotechnol. Biofuels* 2013, 6, 126. [CrossRef]

102. Kahn, A.; Morais, S.; Chung, D.; Sarai, N.S.; Hengge, N.N.; Kahn, A.; Himmel, M.E.; Bayer, E.A.; Bomble, Y.J. Glycosylation of hyperthermostable designer cellulosome components yields enhanced stability and cellulose hydrolysis. *FEBS J.* 2020, 287, 4370–4388. [CrossRef] [PubMed]

103. Zverlov, V.; Mehr, S.; Riedel, K.; Bronnenmeier, K. Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) from the extreme thermophile ‘Anaerocellum thermophilum’ with separate glycosyl hydrolase family 9 and 48 catalytic domains. *Microbiology* 1988, 144, 457–465. [CrossRef] [PubMed]

104. Ding, S.; Himmel, M.E. Anatomy and Ultrastructure of Maize Cell Walls: An Example of Energy Plants. In *Biomass Recalcitrance*; Himmel, M.E., Ed.; Blackwell Publishing: Oxford, UK, 2008; pp. 38–60.

105. Yarbrough, J.M.; Himmel, M.E.; Ding, S.Y. Plant cell wall characterization using scanning probe microscopy techniques. *Biotechnol. Biofuels* 2009, 2, 17. [CrossRef] [PubMed]

106. Park, S.; Baker, J.O.; Himmel, M.E.; Parilla, P.A.; Johnson, D.K. Cellulose crystallinity index: Measurement techniques and their impact on interpreting cellulase performance. *Biotechnol. Biofuels* 2010, 3, 10. [CrossRef]

107. Guo, J.; Catchmark, J.M. Surface area and porosity characterization of nano-scaled cellulose from different acid treatments. In *Plant Cell Walls: Basics of Structure, Chemistry, Accessibility and the Influence on Conversion*. *Aqueous Pretreatment of Plant Biomass for Biological and Chemical Conversion to Fuels and Chemicals*; Wyman, C.E., Ed.; John Wiley & Sons: Oxford, UK, 2013; pp. 23–38.

108. Rouquerol, J.; Avnir, D.; Fairbridge, C.W.; Everett, D.H.; Haynes, J.M.; Pernicone, N.; Ramsay, J.; Sing, K.S.W.; Ünger, K.K. Recommendations for the characterization of porous solids (Technical Report). *Pure Appl. Chem.* 1994, 66, 1739–1758. [CrossRef]

109. Davison, B.H.; Parks, J.; Davis, M.F.; Donohoe, B.S. Plant Cell Walls: Basics of Structure, Chemistry, Accessibility and the Influence on Conversion. In *Biomass Recalcitrance*; Himmel, M.E.; *Aqueous Pretreatment of Plant Biomass for Biological and Chemical Conversion to Fuels and Chemicals*; Wyman, C.E., Ed.; John Wiley & Sons: Oxford, UK, 2013; pp. 23–38.

110. Wang, T.Y.; Chen, H.L.; Lu, M.Y.; Chen, Y.C.; Sung, H.M.; Mao, C.T.; Cho, H.Y.; Ke, H.M.; Hwa, T.Y.; Ruan, S.K.; et al. Functional characterization of cellulases identified from the cow rumen fungus *Neocallimastix patriciarum* W5 by transcriptomic and secretomic analyses. *Biotechnol. Biofuels* 2011, 4, 24. [CrossRef]

111. Chrapkowska, K.J. Cellulolytic activity of white, brown and gray wood rot fungi. *Acta Microbiol. Polym.* 1984, 33, 137–145.

112. Martinez, D.; Challacombe, J.; Morgenstern, I.; Hibbett, D.; Schmoll, M.; Kubicek, C.P.; Ferreira, P.; Ruiz-Duenas, F.J.; Martinez, A.T.; Kersten, P.; et al. Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. *Proc. Natl. Acad. Sci. USA* 2009, 106, 1954–1959. [CrossRef]

113. Floudas, C.A.; Elia, J.A.; Baliban, R.C. Hybrid and single feedstock energy processes for liquid transportation fuels: A critical review. *Comput. Chem. Eng.* 2012, 41, 24–51. [CrossRef]

114. Hamada, N.; Fuse, N.; Shimosaka, M.; Kodaira, R.; Amano, Y.; Kanda, T.; Okazaki, M. Cloning and characterization of a new exo-cellulase gene, cel3, in *Irpex lacteus*. *FEBS Lett.* 1999, 472, 231–237. [CrossRef]

115. Hamada, N.; Ishikawa, K.; Fuse, N.; Kodaira, R.; Shimosaka, M.; Amano, Y.; Kanda, T.; Okazaki, M. Purification, characterization, and gene analysis of exo-cellulase II (Ex-2) from the white rot *basidiomycete Irpex lacteus*. *J. Biosci. Bioeng.* 1999, 87, 442–451. [CrossRef]

116. Toda, H.; Takada, S.; Oda, M.; Amano, Y.; Kanda, T.; Okazaki, M.; Shimosaka, M. Gene cloning of an endoglucanase from the *basidiomycete Irpex lacteus* and its cDNA expression in *Saccharomyces cerevisiae*. *Biosci. Biotechnol. Biochem.* 2005, 69, 1262–1269. [CrossRef]

117. Zheng, F.; Ding, S. Procesivity and enzymatic mode of a glycoside hydrolase family 5 endoglucanase from *Volvariella volvacea*. *Appl. Environ. Microbiol.* 2013, 79, 989–996. [CrossRef]

118. Cohen, R.; Suzuki, M.R.; Hammel, K.E. Processive endoglucanase active in crystalline cellulose hydrolysis by the brown rot *basidiomycete Gloeophyllum trabeum*. *Appl. Environ. Microbiol.* 2005, 71, 2421–2417. [CrossRef]

119. Cohen, R.; Jensen, K.A.; Houtman, C.J.; Hammel, K.E. Significant levels of extracellular reactive oxygen species produced by brown rot *basidiomycetes* on cellulose. *FEBS Lett.* 2002, 531, 483–488. [CrossRef]

120. Irwin, D.C.; Spezio, M.; Walker, L.P.; Wilson, D.B. Activity studies of eight purified cellulases: Specificity, synergism, and binding domain effects. *Biotechnol. Bioeng.* 1993, 42, 1002–1013. [CrossRef]

121. Aylward, J.H.; Gebius, K.S.; Xue, G.-P.; Simpson, G.D.; Dalrymple, B. The *Neocallimastix patriciarium* cellulase, CelD, contains three almost identical catalytic domains with high specific activities on Avicel. *Enzyme Microb. Technol.* 1999, 24, 609–614. [CrossRef]

122. Tambor, J.H.; Ren, H.; Ushinsky, S.; Zheng, Y.; Riemenus, A.; St-Francois, C.; Tsang, A.; Powelljsk, J.; Storms, R. Recombinant expression, activity screening and functional characterization identifies three novel endo-1,4-β-glucanases that efficiently hydrolyse cellulosic substrates. *Appl. Microbiol. Biotechnol.* 2012, 93, 203–214. [CrossRef]

123. Yu, A.H.C.; Saddler, H.N. Identification of essential cellulase components in the hydrolysis of a steam-exploded birch substrate. *Biotechnol. Appl. Biochem.* 1995, 21, 185–202. [CrossRef]
124. Zhang, Y.H.; Cui, J.; Lynd, L.R.; Kuang, L.R. A transition from cellulose swelling to cellulose dissolution by o-phosphoric acid: Evidence from enzymatic hydrolysis and supramolecular structure. *Biomacromolecules* **2006**, *7*, 644–648. [CrossRef]

125. Hall, M.; Bansal, P.; Lee, J.H.; Realf, M.J.; Bommarius, A.S. Cellulose crystallinity—A key predictor of the enzymatic hydrolysis rate. *FEBS J.* **2010**, *277*, 1571–1582. [CrossRef] [PubMed]

126. Fierobe, H.P.; Bayer, E.A.; Tardif, C.; Czjzek, M.; Mechaly, A.; Bélaich, A.; Lamed, R.; Shoahm, Y.; Bélaich, J.P. Degradation of cellulose substrates by cellulolytic chimeras. Substrate targeting versus proximity of enzyme components. *J. Biol. Chem.* **2002**, *277*, 49621–49630. [CrossRef] [PubMed]

127. Basit, A.; Akhtar, M.W. Truncation of the processive Cel5A of *Thermotoga maritima* results in soluble expression and several fold increase in activity. *Biotechnol. Bioeng.* **2018**, *24*, 609–614. [CrossRef] [PubMed]

128. Watanabe, H.; Nakamura, M.; Tokuda, G.; Yamaoka, I.; Scrivener, A.M.; Noda, H. Site of secretion and properties of endogenous endo-beta-1,4-glucanase components from *Reticulitermes speratus* (Kolbe), a Japanese subterranean termite. *Insect Biochem. Mol. Biol.* **1997**, *27*, 305–313. [CrossRef]

129. Reverbel-Leroy, C.; Pages, S.; Bélaich, A.; Tardif, C. The processive endocellulase CelF, a major component of the *Clostridium cellulolyticum* cellulose: Purification and characterization of the recombinant form. *J. Bacteriol.* **1997**, *179*, 46–52. [CrossRef] [PubMed]

130. Ravachol, J.; Borne, R.; Tardif, C.; de Philip, P.; Fierobe, H.P. Characterization of all Family-9 glycoside hydrolase synthesized by the cellulose-producing bacterium *Clostridium cellulolyticum*. *J. Biol. Chem.*** **2014**, *289*, 7335–7348. [CrossRef]

131. Gilad, R.; Rabinovich, L.; Yaron, S.; Bayer, E.A.; Lamed, R.; Gilbert, H.J.; Shoahm, Y. Cell, a noncellulosomal family 9 enzyme from *Clostridium thermocellum*, is a processive endoglucanase that degrades crystalline cellulose. *J. Bacteriol.* **2003**, *185*, 391–398. [CrossRef]

132. McGrath, C.E.; Wilson, D.B. Endocellulolytic activity of the *Clostridium thermocellum* cel9C (formally CbhA) catalytic domain. *Ind. Biotechnol.* **2008**, *4*, 99–104. [CrossRef]

133. Watson, B.J.; Zhang, H.; Longmire, A.G.; Moon, Y.H.; Hutcheson, S.W. Processive endoglucanases mediate degradation of cellulose by *Saccharophagus degradans*. *J. Bacteriol.* **2009**, *191*, 5697–5705. [CrossRef]

134. Alvarez, T.M.; Paiva, J.H.; Ruiz, D.M.; Cairo, J.P.; Pereira, I.O.; Paixão, D.A.; de Almeida, R.F.; Tonoli, C.C.; Ruller, R.; Santos, C.R. Structure and function of a novel cellulase 5 from sugarcane soil metagenome. *PLoS ONE* **2013**, *8*, e83635. [CrossRef]

135. Zhang, X.Z.; Sathitsuksanoh, N.; Zhang, Y.H. Glycoside hydrolase family 9 processive endoglucanase from *Clostridium phytofermentans*: Heterologous expression, characterization, and synergy with family 48 cellobiohydrolase. *Bioresour. Technol.* **2010**, *101*, 5534–5538. [CrossRef]

136. Warnick, T.A.; Methé, B.A.; Leschine, S.B. *Clostridium phytofermentans* sp. nov., a cellulolytic mesophile from forest soil. *Int. J. Syst. Evol. Microbiol.* **2002**, *52*, 1155–1160. [CrossRef]

137. Chhabra, S.R.; Shockley, K.R.; Ward, D.E.; Kelly, R.M. Regulation of endo-acting glycosyl hydrolases in the hyperthermophilic bacterium *Thermotoga maritima* grown on glucon- and mannan-based polysaccharides. *Appl. Environ. Microbiol.* **2002**, *68*, 545–554. [CrossRef]

138. Zhang, Y.H.; Hong, J.; Ye, X. Cellulase assays. *Methods Mol. Biol.* **2009**, *581*, 213–231.

139. Luo, H.; Yang, J.; Yang, P.; Li, J.; Huang, H.; Shi, P.; Bai, Y.; Wang, Y.; Fan, Y.; Yao, B. Gene cloning and expression of a new acidic family 7 endo-beta-1,3–1,4-glucanase from the acidophilic fungus *Bispora* sp. MEY-1. *Appl. Microbiol. Biotechnol.* **2010**, *85*, 1015–1023. [CrossRef]

140. Vuong, T.V.; Wilson, D.B. Processivity, synergism, and substrate specificity of *Thermobifida fusca* Cel6B. *Biotechnol. Biofuels* **2012**, *5*, 1–13. [CrossRef]

141. Horn, S.J.; Vaaje-Kolstad, G.; Westereng, B.; Eijsink, V.G. Measuring processivity. *Environ. Microbiol.* **2009**, *11*, 263–284. [CrossRef] [PubMed]

142. Horn, S.J.; Sørlie, M.; Vårum, K.M.; Väljamäe, P.; Eijsink, V.G. Novel enzymes for the degradation of cellulose. *Biotechnol. Biofuels* **2012**, *5*, 32–34. [CrossRef]

143. Sathitsuksanoh, N.; Zhu, Z.; Zhang, Y.H. Cellulose solvent- and organic solvent-based lignocellulose fractionation enabled efficient sugar release from a variety of lignocellulosic feedstocks. *Bioresour. Technol.* **2010**, *101*, 1571–1582. [CrossRef] [PubMed]

144. Eisler, M.C.; Lee, M.R.; Tarlton, J.F.; Martin, G.B.; Beddington, J.; Dungait, J.A.; Greathead, H.; Liu, J.; Mathew, S.; Miller, H.; et al. Agriculture: Steps to sustainable livestock. *Nature* **2014**, *507*, 32–34. [CrossRef]

145. National Research Council. *Nutrient Requirements of Beef Cattle*; National Academies Press: Washington, DC, USA, 2000.

146. National Research Council. *Nutrient Requirements of Dairy Cattle*; The National Academies Press: Washington, DC, USA, 2001.

147. Fan, M.Z.; Kim, S.W.; Applegate, T.J.; Cervantes, M. Nonruminant Nutrition symposium: Understanding protein synthesis and nitrogen metabolism in nonruminants. *Appl. Sci.* **2021**, *11*, 5150.
152. Forsberg, C.W.; Beveridge, T.J.; Hellstrom, A. Cellulase and Xylanase Release from Bacteroides succinogenes and Its Importance in the Rumen Environment. Appl. Environ. Microbiol. 1981, 42, 886–896. [CrossRef]
153. McGavin, M.; Forsberg, C.W. Isolation and characterization of endoglucanases 1 and 2 from Bacteroides succinogenes S85. J. Bacteriol. 1988, 170, 2914–2922. [CrossRef]
154. Qi, M.; Jun, H.S.; Forsberg, C.W. Characterization and synergistic interactions of Fibrobacter succinogenes glycoside hydrolases. Appl. Environ. Microbiol. 2007, 73, 6098–6105. [CrossRef]
155. Aspeborg, H.; Coutinho, P.M.; Wang, Y.; Brumer, H., 3rd; Henrissat, B. Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). BMC Evol. Biol. 2012, 12, 186. [CrossRef] [PubMed]
156. Liu, L.; Feng, Y.; Duan, C.-J.; Hao, P. Isolation of a gene encoding endoglucanase activity from uncultured microorganisms in buffalo rumen. World J. Microbiol. Biotechnol. 2009, 25, 1035–1042. [CrossRef]
157. National Research Council. Nutrient Requirements of Swine; National Academies Press: Washington, DC, USA, 2012.
158. Koike, S.; Kobayashi, Y. Fibrolytic rumen bacteria: Their ecology and functions. Asian Aust. J. Anim. Sci. 2009, 22, 131–138. [CrossRef]
159. Fang, L.; Zhang, H.; Zhao, F.; Wang, J.; Li, D. Determination of digestive enzyme activity in the digesta from the small intestinal epithelium of pigs. Carbohydr. Res. 1997, 298, 278–285. [CrossRef]
160. Harris, P.V.; Xu, F.; Kreeke, N.E.; Kang, C.; Fukuyama, S. New enzyme insights drive advances in commercial ethanol production. Curr. Opin. Chem. Biol. 2014, 19C, 162–170. [CrossRef]
161. Vincken, J.P.; Baldman, G.; Voragen, A.G. Substrate specificity of endoglucanases: What determines endoglucanase activity? Carbohydr. Res. 1997, 298, 278–285. [CrossRef]
162. Yaoi, K.; Nakai, T.; Kameda, Y.; Hiyoshi, A.; Mitsushi, Y. Cloning and characterization of two xyloglucanases from Paeunibacillus sp. strain KM21. Appl. Environ. Microbiol. 2005, 71, 7670–7678. [CrossRef]
163. Chen, Z.; Friedland, G.D.; Pereira, J.H.; Revecq, S.A.; Chan, R.; Park, J.I.; Thelen, M.P.; Adams, P.D.; Arkin, A.P.; Keasling, J.D.; et al. Tracing determinants of dual substrate specificity in glycoside hydrolase family 5. J. Biol. Chem. 2012, 287, 25335–25343. [CrossRef]
164. Pereira, J.H.; Chen, Z.; McAndrew, R.P.; Sapra, R.; Chhabra, S.R.; Sale, K.L.; Simmons, B.A.; Adams, P.D. Biochemical characterization and crystal structure of endoglucanase Cel5A from the hyperthermophilic Thermotoga maritima. J. Struct. Biol. 2010, 172, 372–379. [CrossRef]
165. Vlasenko, E.; Schuilen, M.; Cherry, J.; Xu, F. Substrate specificity of family 5, 6, 7, 9, 12, and 45 endoglucanases. Bioresour. Technol. 2010, 101, 2405–2411. [CrossRef]
166. Yeoman, C.J.; Han, Y.; Dodd, D.; Schroeder, C.M.; Mackie, R.I.; Cann, I.K. Thermostable enzymes as biocatalysts in the biofuel industry. Adv. Appl. Microbiol. 2010, 70, 1–55. [CrossRef]
167. Blumer-Schuitte, S.E.; Kataeva, I.; Westpheling, J.; Adams, M.W.; Kelly, R.M. Extremely thermophilic microorganisms for biomass conversion: Status and prospects. Curr. Opin. Biotechnol. 2008, 19, 210–217. [CrossRef]
168. Anbar, M.; Bayer, E.A. Approaches for improving thermostability characteristics in cellulosomes. Methods Enzymol. 2012, 510, 261–271.
169. Anbar, M.; Gal, O.; Lamed, R.; Sezerman, U.O.; Bayer, E.A. Improved thermostability of Clostridium thermocellum endoglucanase Cel8A by using consensus-guided mutagenesis. Appl. Environ. Microbiol. 2012, 78, 3458–3464. [CrossRef]
170. Beckham, G.T.; Dai, Z.; Matthews, J.F.; Momany, M.; Payne, C.M.; Adney, W.S.; Baker, S.E.; Himmel, M.E. Harnessing glycosylation to improve cellulase activity. Curr. Opin. Biotechnol. 2012, 23, 338–345. [CrossRef]
171. Fang, L.; Zhang, H.; Zhao, F.; Wang, J.; Li, D. Determination of digestive enzyme activity in the digesta from the small intestinal epithelium of growing pigs and development of in vitro evaluation system for feed bioavailability using artificial small intestinal juice. J. Anim. Vet. Adv. 2012, 11, 1819–1826.
172. Banwell, J.G.; Leonard, P.J.; Lobo, R.M. Measurement of trypsin and chymotrypsin activity in stools to detect chronic pancreatic disease. Gut 1965, 6, 143–145. [CrossRef]
180. Akiba, S.; Kimura, Y.; Yamamoto, K.; Kumagai, H. Purification and characterization of a protease-resistant cellulase from *Aspergillus niger*. *J. Ferment. Bioeng.* 1995, 79, 125–130. [CrossRef]

181. Morgavi, D.P.; Newbold, C.J.; Beever, D.E.; Wallace, R.J. Stability and stabilization of potential feed additive enzymes in rumen fluid. *Enzyme Microb. Technol.* 2000, 26, 171–177. [CrossRef]

182. Bai, Y.; Wang, J.; Zhang, Z.; Shi, P.; Luo, H.; Huang, H.; Luo, C.; Yao, B. A novel family 9 beta-1,3(4)-glucanase from *thermoacidophilic Alicyclobacillus* sp. A4 with potential applications in the brewing industry. *Appl. Microbiol. Biotechnol.* 2010, 87, 251–259. [CrossRef] [PubMed]

183. Woyengo, T.A.; Adeola, O.; Udenigwe, C.C.; Nyachoti, C.M. Gastro-intestinal digesta pH, pepsin activity and soluble mineral concentration responses to supplemental phytic acid and phytase in piglets. *Livestock Sci.* 2010, 134, 91–93. [CrossRef]

184. Davies, G.; Henriissat, B. Structures and mechanisms of glycosyl hydrolases. *Structure* 1995, 3, 853–859. [CrossRef]

185. Morgavi, D.P.; Newbold, C.J.; Beever, D.E.; Wallace, R.J. Stability and stabilization of potential feed additive enzymes in rumen fluid. *Enzyme Microb. Technol.* 2000, 26, 171–177. [CrossRef]

186. Rouvinen, J.; Bergfors, T.; Teeri, T.; Knowles, J.K.; Jones, T.A. Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. *Nat. Struct. Biol.* 1997, 4, 810–818. [CrossRef]

187. Wu, T.H.; Huang, C.H.; Ko, T.P.; Lai, H.L.; Ma, Y.; Chen, C.C.; Cheng, Y.S.; Liu, J.R.; Guo, R.T. Diverse substrate recognition by fission yeast S. cerevisiae cellobiohydrolase I (CBHI) with potential applications in the brewing industry. *Appl. Microbiol. Biotechnol.* 2010, 87, 251–259. [CrossRef] [PubMed]

188. Woyengo, T.A.; Adeola, O.; Udenigwe, C.C.; Nyachoti, C.M. Gastro-intestinal digesta pH, pepsin activity and soluble mineral concentration responses to supplemental phytic acid and phytase in piglets. *Livestock Sci.* 2010, 134, 91–93. [CrossRef]

189. Divne, C.; Ståhlberg, J.; Reinikainen, T.; Ruohonen, L.; Pettersson, G.; Knowles, J.K.; Teeri, T.T.; Jones, T.A. The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Thermomonospora fusca*. *J. Mol. Biol.* 2007, 370, 53–62. [CrossRef]

190. Kohn, J.E.; Afonine, P.V.; Ruscio, J.Z.; Adams, P.D.; Head-Gordon, T. Evidence of functional protein dynamics from X-ray crystallographic ensembles. *PLoS Comput. Biol.* 2010, 6, e1000911. [CrossRef]

191. Wu, B.; Guddat, L.W.; Chang, S.; He, B.; Schenk, G. Processivity and enzymatic mechanism of a multifunctional family 5 endoglucanase from *Clostridium phytofermentans* for the production of recombinant endoglucanase from *Clostridium thermocellum*. *Microb. Cell Fact.* 2015, 14, 57. [CrossRef]

192. Rouvinen, J.; Bergfors, T.; Teeri, T.; Knowles, J.K.; Jones, T.A. Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. *Science* 1990, 249, 380–386. [CrossRef]

193. Mamat, U.; Wilke, K.; Bramhill, D.; Schromm, A.B.; Lindner, B.; Kohl, T.A.; Corchero, J.L.; Villaverde, A.; Schaffer, L.; Head, S.R.; et al. Detoxifying Escherichia coli for endotoxin-free production of recombinant proteins. *Curr. Opin. Biotechnol.* 2020, 61, 1–15. [CrossRef] [PubMed]

194. Elkins, J.G.; Raman, B.; Keller, M. Engineered microbial systems for enhanced conversion of lignocellulosic biomass. *Curr. Opin. Biotechnol.* 2011, 22, 231–238. [CrossRef]

195. Shahzadi, I.; Al-Ghamdi, M.A.; Nadeem, M.S.; Sajjad, M.; Ali, A.; Khan, J.A.; Kazmi, I. Scale-up fermentation of Escherichia coli for the production of recombinant endoglucanase from *Clostridium thermocellum*. *Sci. Rep.* 2021, 11, 7145. [CrossRef]

196. Jin, M.; Balan, V.; Gunawan, C.; Dale, B.E. Consolidated bioprocessing (CBP) performance of *Clostridium phytofermentans* on AFEX-treated corn stover for ethanol production. *Biotechnol. Bioeng.* 2011, 108, 1290–1297. [CrossRef]

197. Rouvinen, J.; Bergfors, T.; Teeri, T.; Knowles, J.K.; Jones, T.A. Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. *Science* 1990, 249, 380–386. [CrossRef]

198. Divne, C.; Ståhlberg, J.; Reinikainen, T.; Ruohonen, L.; Pettersson, G.; Knowles, J.K.; Teeri, T.T.; Jones, T.A. The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. *Science* 1994, 265, 524–528. [CrossRef]

199. Kohn, J.E.; Afonine, P.V.; Ruscio, J.Z.; Adams, P.D.; Head-Gordon, T. Evidence of functional protein dynamics from X-ray crystallographic ensembles. *PLoS Comput. Biol.* 2010, 6, e1000911. [CrossRef]

200. Wu, B.; Guddat, L.W.; Chang, S.; He, B.; Schenk, G. Processivity and enzymatic mechanism of a multifunctional family 5 endoglucanase from *Bacillus subtilis* BS-5 with potential applications in the saccharification of cellulose substrates. *Biotechnol. Biofuels* 2018, 11, 1–15. [CrossRef]

201. Wu, B.; Guddat, L.W.; Chang, S.; He, B.; Schenk, G. Processivity and enzymatic mechanism of a multifunctional family 5 endoglucanase from *Bacillus subtilis* BS-5 with potential applications in the saccharification of cellulose substrates. *Biotechnol. Biofuels* 2018, 11, 1–15. [CrossRef]

202. Wu, S.; Ding, S.; Zhou, R.; Li, Z. Comparative characterization of a recombinant *Volvox aureus* endoglucanase I (EG1) with its truncated catalytic core (EG1-5) and their impact on the bio-treatment of cellulose-based fabrics. *J. Biotechnol.* 2007, 130, 364–369. [CrossRef]

203. Uchiyama, T.; Miyazaki, K. Functional metagenomics for enzyme discovery: Challenges to efficient screening. *Curr. Opin. Biotechnol.* 2011, 22, 231–238. [CrossRef] [PubMed]

204. Shahzadi, I.; Al-Ghamdi, M.A.; Nadeem, M.S.; Sajjad, M.; Ali, A.; Khan, J.A.; Kazmi, I. Scale-up fermentation of Escherichia coli for the production of recombinant endoglucanase from *Clostridium thermocellum*. *Sci. Rep.* 2021, 11, 7145. [CrossRef]

205. Jin, M.; Balan, V.; Gunawan, C.; Dale, B.E. Consolidated bioprocessing (CBP) performance of *Clostridium phytofermentans* on AFEX-treated corn stover for ethanol production. *Biotechnol. Bioeng.* 2011, 108, 1290–1297. [CrossRef]

206. Elkins, J.G.; Raman, B.; Keller, M. Engineered microbial systems for enhanced conversion of lignocellulosic biomass. *Curr. Opin. Biotechnol.* 2010, 21, 657–662. [CrossRef] [PubMed]

207. Wu, B.; Guddat, L.W.; Chang, S.; He, B.; Schenk, G. Processivity and enzymatic mechanism of a multifunctional family 5 endoglucanase from *Bacillus subtilis* BS-5 with potential applications in the saccharification of cellulose substrates. *Biotechnol. Biofuels* 2018, 11, 1–15. [CrossRef] [PubMed]

208. Woyengo, T.A.; Adeola, O.; Udenigwe, C.C.; Nyachoti, C.M. Gastro-intestinal digesta pH, pepsin activity and soluble mineral concentration responses to supplemental phytic acid and phytase in piglets. *Livestock Sci.* 2010, 134, 91–93. [CrossRef]

209. Rouvinen, J.; Bergfors, T.; Teeri, T.; Knowles, J.K.; Jones, T.A. Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. *Science* 1990, 249, 380–386. [CrossRef]

210. Divne, C.; Ståhlberg, J.; Reinikainen, T.; Ruohonen, L.; Pettersson, G.; Knowles, J.K.; Teeri, T.T.; Jones, T.A. The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. *Science* 1994, 265, 524–528. [CrossRef]