fermentation

MDPI

Review

Glycoside Hydrolase Family 48 Cellulase: A Key Player in Cellulolytic Bacteria for Lignocellulose Biorefinery

Cai You [1,2,3,4,5], Ya-Jun Liu [1,2,3,4,5,6], Qiu Cui [1,2,3,4,5,6,*] and Yingang Feng [1,2,3,4,5,6,*]

1 CAS Key Laboratory of Biofuels, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, 189 Songling Road, Qingdao 266101, China
2 Shandong Provincial Key Laboratory of Synthetic Biology, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, 189 Songling Road, Qingdao 266101, China
3 Shandong Engineering Laboratory of Single Cell Oil, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, 189 Songling Road, Qingdao 266101, China
4 Shandong Energy Institute, 189 Songling Road, Qingdao 266101, China
5 Qingdao New Energy Shandong Laboratory, 189 Songling Road, Qingdao 266101, China
6 University of Chinese Academy of Sciences, Beijing 100049, China
* Correspondence: cuiqiu@qibebt.ac.cn (Q.C.); fengyg@qibebt.ac.cn (Y.F.)

Abstract: Cellulases from glycoside hydrolase family 48 (GH48) are critical components of natural lignocellulose-degrading systems. GH48 cellulases are broadly distributed in cellulolytic microorganisms. With the development of genomics and metatranscriptomics, diverse GH48 genes have been identified, especially in the highly efficient cellulose-degrading ruminal system. GH48 cellulases utilize an inverting mechanism to hydrolyze cellulose in a processive mode. Although GH48 cellulases are indispensable for cellulolytic bacteria, they exhibit intrinsically low cellulolytic activity. Great efforts have been made to improve their performance. Besides, GH48 cellulases greatly synergize with the complementary endoglucanases in free cellulase systems or cellulosome systems. In this review, we summarized the studies on the diversity of GH48 cellulases, the crystal structures, the catalytic mechanism, the synergy between GH48 cellulases and endocellulases, and the strategies and progress of GH48 engineering. According to the summarized bottlenecks in GH48 research and applications, we suggest that future studies should be focused on mining and characterizing new GH48 enzymes, thoroughly understanding the progressive activity and product inhibition, engineering GH48 enzymes to improve stability, activity, and stress resistance, and designing and developing new biocatalytic system employing the synergies between GH48 and other enzymes.

Keywords: GH48 family; cellulase; lignocellulose degradation; structure; catalytic mechanism; synergy

Citation: You, C.; Liu, Y.-J.; Cui, Q.; Feng, Y. Glycoside Hydrolase Family 48 Cellulase: A Key Player in Cellulolytic Bacteria for Lignocellulose Biorefinery. *Fermentation* **2023**, *9*, 204. https://doi.org/10.3390/fermentation9030204

Academic Editor: Liqun Jiang

Received: 31 January 2023
Revised: 18 February 2023
Accepted: 20 February 2023
Published: 21 February 2023

1. Introduction

Lignocellulose—the fibrous material in the plant cell wall—consisting of 30–50% cellulose, 10–40% hemicellulose, and 5–30% lignin [1], is one of the most abundant renewable resources on the planet. Biodegradation and conversion of lignocellulose are essential for carbon recycling, agricultural development, and environmental homeostasis [2,3]. However, the dense, tough, and complex structural properties of lignocellulose make it difficult to be efficiently degraded and utilized. In nature, various microorganisms have evolved effective strategies to biodegrade lignocellulosic materials through secreting synergistic mixtures of enzymes including cellulases, hemicellulose, and lignin-degrading enzymes [4,5]. The degradation of cellulose into glucose (i.e., the saccharification) needs the synergy of three types of cellulases: endoglucanase, cellobiohydrolase/cellobiosidase, and β-glucosidase [6–8]. The hemicellulases are more diverse because hemicellulose has a complex saccharide composition and structure [9,10]. Lignin is a complex aromatic polymer and ligninolytic enzymes are mainly by oxidoreductases, such as lignin peroxidase,

manganese peroxide, laccase, etc. [11]. The efficiency of enzymatic degradation of lignin is often very low, so lignin is first isolated by a pretreatment procedure in most strategies of the current lignocellulose biorefinery [12,13]. After the pretreatment, the saccharification processes of lignocellulose biorefinery mainly depend on cellulases and hemicellulases.

Cellulases and hemicellulases belong to glycoside hydrolases (GHs). Glycoside hydrolases are a large collection of enzymes that hydrolyze glycosidic bonds including as many as 173 families. Many GH families are widely involved in lignocellulose degradation, with GH families 5, 6, 7, 8, 12, 45, and 48 containing most fungal cellulases [8,14] and GH5, GH9, GH10, GH11, GH43, and GH48 commonly found in cellulolytic bacteria [7,15]. Family GH48 is an essential member in the hydrolysis of crystalline cellulose [16]. It includes mainly, but not always, exocellulases (namely, exoglucanase or cellobiohydrolase) that catalyze cellulose hydrolysis to release cellobiose molecules [17]. Besides, GH48 cellobiohydrolases synergize greatly with endoglucanases to improve cellulose degradation efficiency and synergize with β-glucosidase to eliminate feedback inhibition [13,18]. For instance, the cellulosomal Cel48S from *Clostridium thermocellum* synergizes with GH9 endoglucanases for efficient crystalline cellulose hydrolysis [18,19], while the large multi-module CelA from *Caldicellulosiruptor bescii* containing both a GH9 module Cel9A and a GH48 module Cel48A is also known to be synergistic in the activities of the two modules [20,21]. To date, all organisms possessing at least one GH48 member are truly cellulolytic organisms, i.e., they can substantially degrade crystalline cellulose [22].

Since the first GH48 exoglucanase CelS (also known as Cel48S, Cel48A, or S8) was characterized in 1991, more than 1500 GH48 sequences have been released in GenBank according to the CAZy (Carbohydrate-Active EnZymes) database (http://www.cazy.org, accessed on 13 December 2022) [23]. GH48 enzymes are found mainly in bacteria including Firmicutes, Chloroflexi, Proteobacteria, and Actinobacteria. They also exist in some Fungi, viruses, and even Arthropoda. The GH48 genes of Actinobacteria, Firmicutes, and Chloroflexi may originate from a common ancestor, while the GH48 genes of Proteobacteria, Fungi, and Arthropoda are believed to be obtained from Actinobacia and Firmicutes through gene horizontal transfer [16,24,25].

The GH48 family is primarily classified as $(\alpha/\alpha)_6$-fold enzymes. Currently, crystal structures of 11 GH48 cellulases with or without oligosaccharides have been reported [26]. GH48 cellulases utilize an inverting mechanism with aspartate and glutamate acting as base and acid donors, respectively, to cleave the sugar chain of cellulose [26,27]. GH48 cellulases exhibit low enzyme activities mainly attributed to end-product inhibition, which limits the utilization of cellulases in lignocellulose bioconversion [17]. In view of this, mutagenesis strategies based on crystal structures and molecular dynamics simulations have been adopted to relieve end-product inhibition and enhance enzyme activity [17,28,29].

GH48 cellulases are among the most abundant glycoside hydrolases in nature. Despite the importance of GH48 cellulases, few related reviews have been published. To deeply understand the role and mechanism of GH48 in cellulose degradation, in this review, we elaborated on the advances in the distribution and evolutionary relationship, the structures, as well as the catalytic and processive mechanism of GH48 cellulases. We also summarized the engineering strategies for activity improvement and reviewed the synergistic effects of GH48 enzymes with endocellulases in various cellulolytic systems. This review will provide a reference for better utilization of GH48 in lignocellulose biorefinery.

2. Key Role of GH48 Cellulases in Cellulose Degradation

Lignocellulose-derived sugars represent the largest reserve of fermentable sugars in Nature [30]. However, lignocellulose is difficult to deconstruct and utilize due to its recalcitrant structure and diverse and complex composition, with cellulose in crystalline and fibrous forms [31,32]. Three dominant enzymatic systems are utilized by cellulolytic organisms to overcome the recalcitrant nature of lignocellulose, including the free cellulases derived from aerobic fungi and bacteria, the cellulosome system mainly produced by anaerobic bacteria, and multimodular glycoside hydrolases with carbohydrate-binding

modules usually produced by *Caldicellulosiruptor* species. One common feature of these various cellulolytic systems is the high expression of GH48 cellulases, which are considered to be the key component of cellulose degradation [33–35].

Usually, only one or two (rarely three) GH48 gene(s) are present in the genomes of cellulolytic organisms. For the well-known cellulolytic bacteria *C. thermocellum*, which is one of the most efficient cellulose degraders in nature, two GH48 cellulases—cellulosomal Cel48S and non-cellulosomal Cel48Y—exhibit hydrolysis activities on crystalline cellulose. These GH48 cellulases are upregulated during growth on crystalline cellulose. Meanwhile, the deletion of Cel48S and Cel48Y led to a significant decrease in performance but does not completely abolish cellulolytic activity [35]. Moreover, three mutants of the cellulolytic bacterium *Ruminococcus albus* which are impaired in the production of Cel9B and Cel48A were all reported to be defective in adhesion to and degradation of cellulose [33]. Additionally, deletion of the CelA gene in *C. bescii* seriously damaged its ability to grow on crystalline cellulose and abolished its growth on lignocellulosic biomass, with a 15-fold decrease in sugar release on crystalline cellulose compared with the parent and wild-type strains. Meanwhile, the loss of exoglucanase activity could not be compensated by other enzymes in the *C. bescii* secretome [36]. Except for the genetics perspective, the importance of GH48 cellulases for crystalline cellulose hydrolysis has also been addressed from the perspectives of metatranscriptomic [37] and genomic analyses [38,39]. GH48 cellulases are necessary for microorganisms to degrade crystalline cellulose, so GH48 genes were reported to be a suitable molecular marker for the characterization of truly cellulolytic bacteria, especially in anaerobic environments [34,40]. The detection and quantification of GH48 genes also can be used to identify cellulolytic organisms [22]. Moreover, Cel48A, the GH48 cellulase domain in CelA, exhibits catalytic promiscuity in hydrolyzing xylan with an unrevealed mechanism [41].

3. Diversity and Distribution of GH48 Family

The glycoside hydrolase family 48 (GH48) comprises reducing end cellobiohydrolases (EC 3.2.1.176), non-reducing end cellobiohydrolases (EC 3.2.1.91), endo-β-1,4-glucanases (EC 3.2.1.4), which are important sources of cellulases, as well as part of chitinase (EC 3.2.1.14), according to the CAZy database. As the key components in bacterial cellulase systems, GH48 cellulases are highly expressed in cellulolytic bacteria, such as CelS from *C. thermocellum* [42,43], CelF from *Clostridium cellulolyticum* [44], CelA from *C. bescii* [21,45,46], TfCel48A from *Thermobifida fusca* [47], and so on. With an increasing number of bacterial genome sequences and the growing availability of metagenomes, the number of GH48 sequences identified in the CAZy database increases dramatically. Till December 2022, the GH48 family contains 1595 known protein sequences, of which 95.6% are from bacterial organisms, 3.7% are from eukaryotes including Fungi and Arthropoda, and even 4 sequences from viruses. Of the more than 1500 bacterial GH48 genes currently predicted, only 20 have been characterized, indicating a vast untapped resource of GH48 enzymes (Figure 1). About 40.6% of the sequences are from uncultured and unclassified bacteria. The bacterial GH48 is known to be expressed by more than 70 genera, mainly distributed in two terrabacteria groups, Actinobacteria and Firmicutes, and also, a few from Proteobacteria, Bacteroidota, and Chloroflexi. Metagenomic data analysis indicated that the GH48-type enzymes of three closely related phyla, Actinobacteria, Firmicutes, and Chloroflexi, may have originated from a common ancestor; meanwhile, fungi, insects, and Proteobacteria may receive their GH48 genes horizontally from Firmicutes and Actinobacteria [25]. Surprisingly, GH48 family genes have been identified from several polyphagan coleopterans, especially from the two superfamilies (Chrysomeloidea and Curculionoidea) [48]. This finding confirmed the possibility of their acquisitions by horizontal gene transfer rather than simple vertical transmission from ancestral lineages of insects [48]. Besides, the GH48 family from insects often exhibits chitinase activity instead of glucanase or cellobiohydrolase activity [25,48].

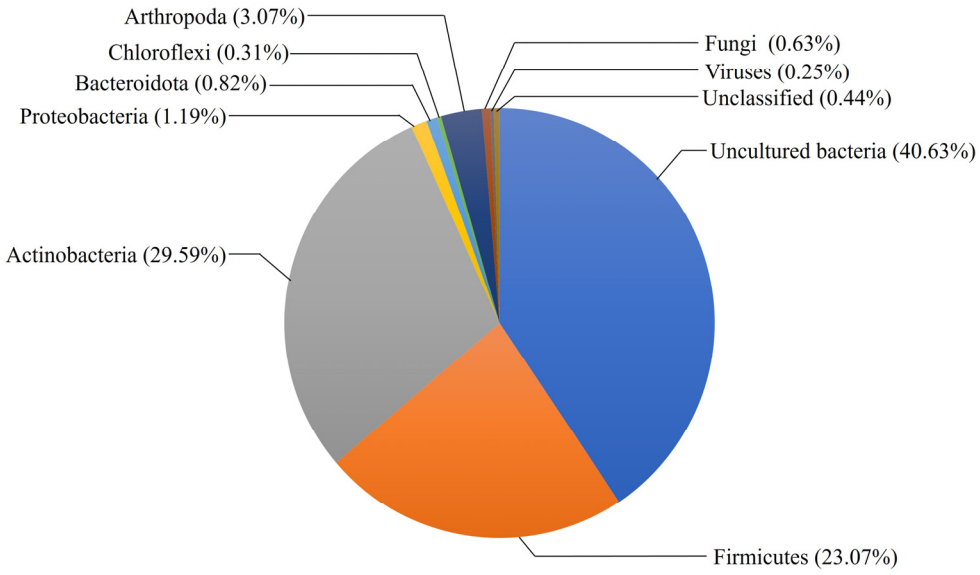

Figure 1. Distribution of GH48 sequences retrieved from the CAZy database.

The rumen is one of the most efficient cellulose degradation systems in nature, harboring abundant cellulose-degrading microorganisms that efficiently break down plant biomass to provide energy to the host animal [14]. Currently, in cultured rumen microorganisms, reported GH48-producing bacteria mainly include *R. albus*, *Ruminococcus flavefaciens*, *Cellulosilyticum ruminicola*, and *Clostridium acetobutylicum*, and anaerobic fungi include *Neocallimastix patriciarum*, *Piromyces* sp., and *Piromyces equi*. However, as about 77% of the ruminal microorganisms attached to solid fibers are considered to be uncultured, the GH48 cellulases in the rumen are underestimated [37]. Transcripts for GH48 cellulases are in a relatively high abundance in the metatranscriptomic analysis of the rumen microbiomes, accounting for 3.0% of CAZyme transcripts in a metatranscriptomic study of the fibrolytic microorganisms in the rumen of a cow fed a mixed diet [37,49]. The abundance and diversity of GH48 genes in the rumen may provide new gene resources for the efficient degradation of cellulose.

4. Structures of GH48 Cellulases

As early as 1998, Parsiegla et al. reported the first crystal structure of GH48 cellulase, the catalytic domain of processive endocellulase Cel48F from *C. cellulolyticum* in complex with a thiooligosaccharide inhibitor [50]. Subsequently, they reported structures of native and mutated Cel48F in complex with the cello-oligosaccharides, hemithiocellooligosaccharide, and thio-oligosaccharide inhibitors (Table 1) [51,52]. Currently, 26 structures of 11 unique GH48 cellulases including Cel48F, CelS from *C. thermocellum*, ExgS from *Clostridium cellulovorans*, TfCel48A from *T. fusca*, two GH48s from *Caldicellulosiruptor* genus, three GH48s from *Bacillus* genus, and HcheGH48 from *Hahella chejuensis*. Most of them are exocellulases, except for Cel48F, which is a processive endocellulase.

Table 1. Characterized and structure-determined GH48 cellulases.

GH48s	Activities	Origin	PDB Code (Ligand)	Form	Reference
CelS (Cel48S)	Exocellulase	*C. thermocellum* DSM1313	5YJ6	cellulosomal	[53]
CelS (Cel48A)	Exocellulase	*C. thermocellum* F7	1L1Y(cellobiose), 1L2A (cellobiose, cellohexaose)	cellulosomal	[54]
Cel48Y	Exocellulase	*C. thermocellum* ATCC 27405		free-cellulase	[55]
Cel48S	Exocellulase	*C. thermocellum* ATCC 27405		cellulosomal	[56]

Table 1. *Cont.*

GH48s	Activities	Origin	PDB Code (Ligand)	Form	Reference
CelY	Exocellulase	*Clostridium stercorarium*		free-cellulase	[57]
GH48	Exocellulase	*Clostridium clariflavum* DSM 19732		cellulosomal	[58]
CpCel48	Exocellulase	*Clostridium phytofermentans* ISDg		free-cellulase	[59]
ExgS	Exocellulase	*Clostridium cellulovorans* ATCC 35296	4XWL (PEG), 4XWM (cellobiose), 4XWN (cellobiose, cellopentaose)	cellulosomal	[26]
Cel48F	Processive endocellulase	*C. cellulolyticum* H10	1FCE (inhibitor IG4), 1F9D (cellotetraose), 1FBW (cellohexaose), 1FAE (cellobiose), 1FBO (cellobiitol), 1F9O (inhibitor PIPS-IG3), 1G9G (glucose), 2QNO (thiocellodecaose), 1G9J (hemithiocellooligosaccharide)	cellulosomal	[50–52]
CbCel48A	Exocellulase	*C. bescii* DSM 6725	4EL8, 4L0G (cellobiose), 4TXT (cellotriose), 4L6X	multi- module	[41]
Cdan_2053	Cellulase	*Caldicellulosiruptor danielii*	6D5D (cellobiose)	multi-module	[15]
BlCel48B	Processive cellulase	*Bacillus licheniformis* DSM 13	7KW6 (cellobiose, cellotetraose)	free-cellulase	[28]
BpCel48	Cellulase	*Bacillus pmilus* SAFR-032	5BV9 (cellobiose), 5CVY (cellobiose, cellohexaose)	free-cellulase	[17,60]
BpGH48	Cellulase	*Bacillus pmilus* SH-B9	5VMA (cellobio-derived isofagomine)	free-cellulase	To be published
TfCel48A	Exocellulase	*T. fusca* YX	4JJJ (cellobiose, cellohexaose)	free-cellulase	[61]
HcheGH48	Cellulase	*H. chejuensis* KCTC 2396	4FUS (cellobiose)	free-cellulase	[25]
CbhB	Exocellulase	*Cellulomonas fimi* ATCC 484		free-cellulase	[62]
Cel48A	Exocellulase	*C. ruminicola* H1		free-cellulase	[63]
Cel48	Exocellulase	*Myxobacter* sp. AL-1		free-cellulase	[64]
Cel48C	Exocellulase	*Paenibacillus* sp. BP-23		free-cellulase	[65]

All these GH48 structures exhibit similar overall fold, with the C^{α} RMSD values ranging from about 0.5 Å to 0.7 Å by superimposing CelS to other GH48 cellulases. They share a typical $(\alpha/\alpha)_6$ barrel consisting of an inner core of six mutually parallel α-helixes (helices with even numbers) and an outer shell of six peripheral α-helixes (helices with odd numbers), and the N-terminus of each inner helix is connected by long loops, additional helices, or sheets to the C-terminus of one outer helix (Figure 2A). The catalytic residues are located in the N-terminal region of two inner helices, while most substrate-binding residues are located on the additional elements which form a layer covering the barrel. The covering elements may play roles in modulating the function of GH48 enzymes. BpCel48 from *Bacillus pumilus* exhibits eight longer loops compared to other GH48 structures. Structural overlay revealed that all three GH48 enzymes from *Bacillus* sp. feature these extra loops (Figure 2B). Molecular dynamics simulations indicated that BpCel48 loops near the tunnel exit do not affect product inhibition. However, these loops are speculated to be responsible for the lower thermostability of BpCel48 by being more exposed to solvent [17,60]. A recent study indicates that two extra longer loops (loop2 and loop6) located at the exit of the active site in BlCel48B act as an extension of the catalytic pocket and form a platform for product anchoring at the exit from the open-cleft part of the active site [28]. Among these GH48 structures, HcheGH48 is more special as it is from Proteobacteria, whose gene is believed to be obtained by horizontal gene transfer [25]. Consistent with the gene transfer statement, the structure of HcheGH48 is almost identical to other GH48 structures. Besides, a structural element termed ω-loop located between residue Pro469 and Ala482 (as in Cel48F) in all cellulases is proposed to distinguish cellulases and non-cellulases from insects (i.e., chitinases) coded by horizontally transferred GH48 genes.

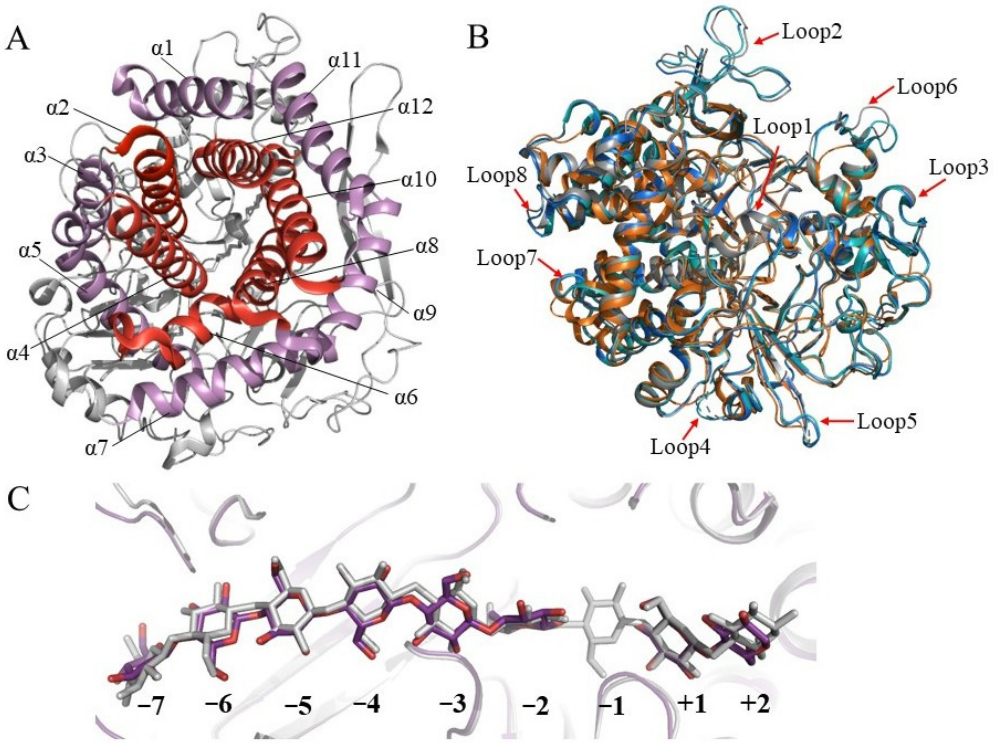

Figure 2. Structure and substrate sites of GH48 cellulase. (**A**) The overall structure of GH48 cellulase, CelS (PDB entry 5YJ6). The core $(\alpha/\alpha)_6$ barrels are colored red and violet, and the additional secondary structures are gray. The helices of the inner barrel with even numbers are in red while the helices of the outer barrel with odd numbers are shown in violet. (**B**) Superimposition of CelS (PDB entry 5YJ6, orange) and three *Bacillus* GH48 cellulases (PDB entry 5WMA, marine; PDB entry 5BV9, cyan; PDB entry 7KW6, gray). The extra loops of *Bacillus* GH48s are indicated by red arrows. (**C**) The substrate sites of GH48 cellulase. The cellobiose in subsites +1 and +2 and cellohexaose in subsites −2 to −7 in CelS-cellohexaose are shown in deep purple (PDB entry1L2A). The thiocellodecaose located at subsites +2 to −7 in CelF E55Q/thiocellodecaose complex are colored in gray (PDB entry 2QNO). The figure was prepared using PyMOL (Schrödinger).

Structures of GH48 enzymes in complex with oligosaccharides revealed the active-site topology, generally featuring a tunnel-like substrate binding part (subsites named −7 to −1) and an open-cleft product binding part (subsites named +1 and +2) (Figure 2C). In most complex structures, the tunnel-shaped active site is occupied by cello-oligosaccharides, and the cleft part is bounded by cellobiose in subsites +1 and +2 after the cleavage site. Nevertheless, another subsite +3 is surmised in the complex of the inactive mutant E55Q of CelF and cellohexaose or cellotetraose, indicating that there is sufficient sugar-binding potential at the tunnel exit [51]. Residues constituting the tunnel are quite conserved. In the statistics of the five GH48 enzymes, including BpCel48, CbCel48A, TfCel48A, CelS, and Cel48F, 27 residues of 36 that represent the tunnel walls and contact with the substrate/product are universally conserved and most of the rest are highly conserved [60]. A large content of the conserved residues is aromatic residues, as well as several charged residues, including Arg, Asp, and Glu residing along the tunnel exit [17]. The aromatic residues interact and stabilize the cellulose chain along the tunnel length by stacking interactions with the sugar moieties. These aromatic residues are supposed to serve as lubricating agents to reduce the sliding barrier in the processive action [51]. Further studies confirm their essential roles in the molecular recognition of insoluble cellulosic substrates as their mutants dramatically affect the enzyme hydrolysis rate and processivity [17,61].

5. Catalytic Mechanism

Although GH48 exocellulases were reported to be able to initiate hydrolysis from either reducing or nonreducing end of the cellulose chain, all characterized GH48s are reducing end-acting cellobiohydrolases. In general, after recognition, GH48 cellulases acquire the reducing end of a cellulose chain into its active site tunnel. Subsequently, the cellulose chain processes through the tunnel till the position for hydrolysis. The hydrolysis reaction occurs at every two glycosidic bonds, and the product of this process is cellobiose. The cleavage site is located between subsite −1 and subsite +1. Then, the produced cellobiose exits from an open shallow cleft located after the tunnel and is released into the aqueous environment [17,66]. Upon product release, the subdomain that forms the open cleft at the reducing end of the substrate is supposed to rearrange to facilitate the linear sliding of the uncleaved cellulose chain from the tunnel to subsites +1 and +2. In turn, substrate occupation of the +1 and +2 positions will trigger a reverse conformational change of the $(\alpha/\alpha)_6$–helix barrel, inducing substrate kinking and subsequent catalysis (Figure 3). This is the so-called "substrate-control" mechanism [27,52]. At some point, the enzymes dissociate from the cellulose substrate, stopping the processive cycle.

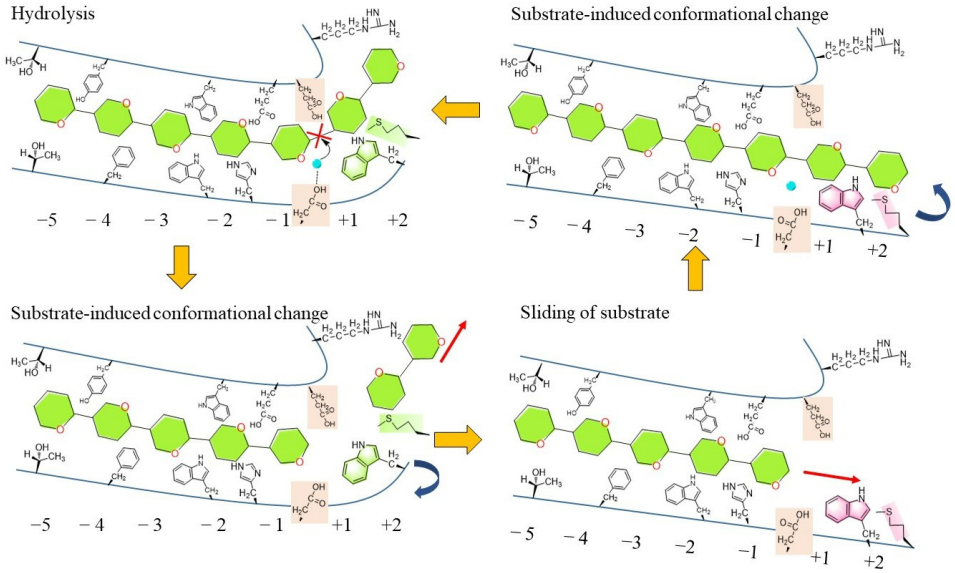

Figure 3. The hypothetical model of the processive action of GH48 cellulases [27,54]. The sugar moieties of cellulose are represented by green hexagons. The catalytic acid and base are shown in light orange shades, the nucleophilic-attacking water molecule is drawn as a blue sphere. Two important residues (Trp417 and Met414 as in Cel48F) contributing to the conformational changes are colored green (hydrolyzing conformation) and pink (sliding conformation), respectively. Some of the residues consisting of the channel are also indicated. The figure was prepared in Microsoft PowerPoint with molecular structure pictures generated by ChemDraw (PerkinElmer).

The catalytic mechanism that GH48 cellulases adopt is a concerted, inverting mechanism and a pair of carboxylic residues are required. In particular, one residue (glutamic acid) generally acts as a catalytic acid providing a proton, and the other (aspartic acid) acts as a catalytic base to activate the nucleophilic-attacking water molecule [27,67]. The average distance between the acid and the base should be approximately 10 Å [66]. The key catalytic residues have been identified in many reported GH48 cellulases. For instance, Glu87 and Asp255 act as a general acid and a general base in the catalytic function of GH48 CelS, and they are equivalent to Glu50 and Asp222 in *C. cellulovorans* exoglucanase ExgS, respectively. Specifically, in the ExgS–cellobiose structure, the O1 atom of Asp222 bridges the C1 atom of subsite −1 through one water molecule (W1), and it was proposed that after the removal of an H atom from water W1 by Asp222, the resulting hydroxyl group attacks the anomeric C atom at subsite −1, which contributes to the mechanism of hy-

drolysis. In Cel48F from *C. cellulolyticum*, Glu55 is a confirmed catalytic acid and two possible candidates Asp230 and Glu44 are proposed as the base. The results of molecular dynamics simulations indicated that Asp230 acts as the base and Glu44 also plays a crucial role to maintain the proper conformation of the substrate to ensure a successful cleavage reaction [27,66]. Besides, Met414 and Trp417 constitute the water-control system, and some hydrophobic residues are supposed to reduce the sliding energy barrier or provide a hydrophobic environment to prevent surrounding water molecules from entering the active site. Therefore, except for the key catalytic residues, other hydrophilic or hydrophobic residues including aromatic entrance residues also have significant effects on the catalytic process. Still, much remains to be understood about the molecular mechanisms underlying the progressive action of GH48 cellulases. For example, it is reported that the neighboring water molecules move into the void left by the product after the product expulsion, but how this water traffic signal is communicated to the substrate movement remains unclear [29].

6. Strategies and Progress of Engineering GH48 Cellulases

Compared with other family cellulases, GH48 cellulases exhibit relatively low specific activity on cellulose in assays in vitro. Three main factors are speculated to be responsible for the low enzymatic activities of GH48s: the inefficient acquisition of cellulose by the tunnel entrance, the slow processivity of the cellulose substrate in the tunnel, and the end-product inhibition [17]. A fundamental factor affecting the enzymatic activity of GH48 is the substrate properties. A comprehensive study of enzymatic properties of the processive BlCel48B cellulase from *B. licheniformis* indicates that the heterogeneity and structural nature of cellulose substrates, including substrate size and morphology, impact the substrate affinity, cleavage patterns, processivity, and hydrolytic efficiency of BlCel48B [28]. In addition, other cellulases are reported to be strongly influenced by the ratio between the average free path for cellulase processive dislocation after one catalytic step and its processivity, as well as by the physical and chemical structure of the substrate [28,68,69]. The end-product cellobiose has been reported to strongly inhibit the activity of several GH48 cellulases, such as *C. thermocellum* CelS [70] and *T. fusca* Cel48A [71].

Various factors are reported to influence the product binding affinity, such as the pH of the solution, the type of the product, and the enzymatic environment [29,72,73]. The product inhibitory effect of four GH48s has been quantitatively evaluated, with CelS exhibiting the highest product inhibitory level, followed by BpCel48, CelF, and CbCel48A. A series of single mutants with theoretically reduced levels of product inhibition have also been proposed [17]. For the well-studied Cel48F, a hydrogen bond rearrangement that reduces the sliding barrier and stimulates the product to move toward the exit is important for the product release progress. This provides clues and directions to the modification or the mutation of cellulase to enhance the catalytic activity [27,66].

Effects on improving the enzyme secretion and stability have also been made to enhance GH48 enzymatic activity on cellulose. A PelB signal peptide mediating post-translational secretion has been attached to the N-terminal end of CelS (P-Cel48S) and allowed catalytically active Cel48S to be successfully produced in the culture medium of recombinant *Escherichia coli* [74]. Meanwhile, recombinant Cel48S via the co-translational pathway (attached with a DsbA signal peptide) yielded a 2.2-times higher specific activity than that associated with P-Cel48S expression. A set of Cel48 chimeras created from the catalytic domains of three native Cel48 enzymes CelF, CelS, and CelY by structure-guided recombination have been evaluated and subsequent sequence-function analysis demonstrates a high degree of additivity in the sequence–stability relationship, and this will help to predict highly stable and active Cel48 enzymes [75].

In addition to product inhibition, the sliding of the substrate into the active site is another crucial step in cellulose degradation. Recently, nonequilibrium molecular dynamics simulations are carried out to investigate the energetics and mechanism of the substrate dynamics and product expulsion in CelS. The results indicate that product removal is relatively easier and faster than the sliding of the substrate to the catalytic active site [29].

Therefore, the details of the substrate passage in the processive action of GH48s will be another noteworthy entry point to the rationale design of enzymes with better yield and performance.

7. Synergistic Effects of GH48 Cellulases with Endocellulases

Synergism between different cellulases was reported as early as 1950 [76]. In the classical synergistic action, endocellulases attack the amorphous part of cellulose, creating more chain ends for exocellulases to attack. In turn, the exocellulase activity exposes new amorphous regions in the bulk substrate, thereby further stimulating subsequent endocellulase activity [77,78]. Although GH48 exocellulases exhibit relatively low hydrolysis activities, they often act in synergy with endocellulases, for example, GH9 endocellulases, which dramatically increase their performance on crystalline cellulose. According to the form of cellulases, there are three types of synergistic systems between GH48 cellulases and endocellulases (Figure 4) [79]. The most common one is the intermolecular synergy resulting from the complementary function of separate GH48 cellulases and endocellulases which exists mainly in the free cellulase system used by cellulolytic fungi and bacteria. For instance, TfCel48A has been shown to synergize well with a processive endocellulase TfCel9A and other *T. fusca* cellulases [47,80]. The addition of Cel48A to a balanced mixture of *T. fusca* endocellulase and exocellulase led to improved hydrolytic activity [71].

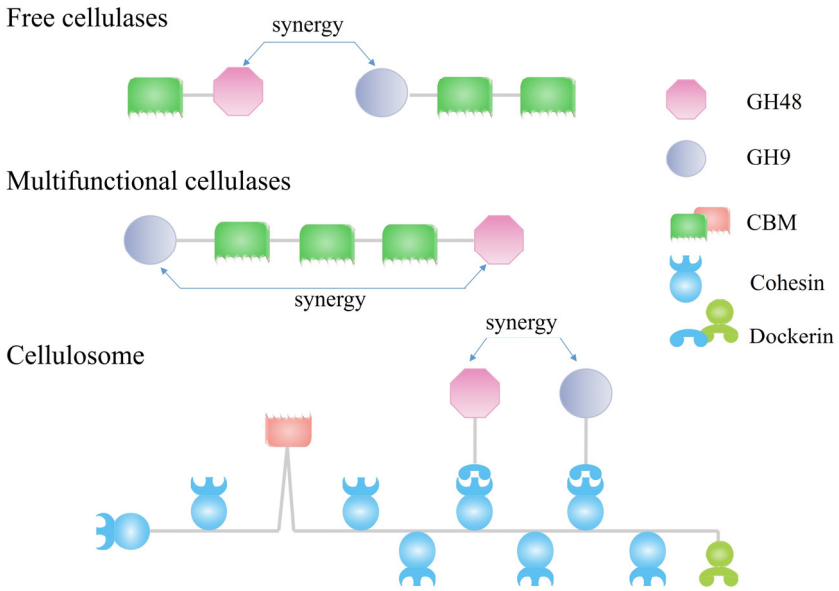

Figure 4. The three types of synergistic systems between GH48 cellulases and endocellulases. The figure was prepared in Microsoft PowerPoint.

The second type of synergism refers to multifunctional cellulases utilizing intramolecular synergy between nearby exo- and endocellulase catalytic modules tethered into a single gene product. This synergism type is widespread in some cellulolytic bacteria, represented by *Caldicellulosiruptor* sp. [46]. The most extensively studied multifunctional cellulase is CelA from *C. bescii* containing an N-terminal GH9 endoglucanase module and a C-terminal GH48 endoglucanase module separated by three family 3 cellulose-binding modules (CBMs) (GH9-CBM3c-CBM3b-CBM3b-GH48) [38]. CelA is one of the most highly expressed proteins in the extracellular proteome of *C. bescii* [41,45,81] and displays a high efficiency in hydrolyzing crystalline cellulose, which is attributed to the increased intramolecular synergy with the proximity of chain-end-forming endoglucanase and an efficient cellobiohydrolase in the same molecule [41]. Homologs of the CelA gene are widely found in the genomes of most cellulolytic members of the genus *Caldicellulosiruptor*. Besides, the GH48 module, *Cb*Cel48B of multimodule CelC from *C. bescii* is demonstrated to work synergically with the GH10 module *Cb*Xyn10C, a xylanase module, to hydrolyze both

cellulose and xylan [38]. The synergistic hydrolysis of xylan is dependent on the covalent linkage of the two domains in the same polypeptide, while the synergy effect can take place both intra- and intermolecularly in hydrolyzing cellulose. Furthermore, *Cb*Xyn10C can also cooperate with the GH48 domain of both CelA and CelF. The combination of GH10 and GH48 domains in a protein is also reported in five other *Caldicellulosiruptor* species that can efficiently degrade lignocellulose [39,82,83].

The third type of synergism is a tethered multi-enzyme proximity-dependent synergy observed in cellulosomes. Cellulosomes are multienzyme complexes produced by anaerobic cellulolytic bacteria for the degradation of lignocellulosic biomass [7,84]. Cellulosomes are composed of various dockerin-containing enzymatic subunits and cohesin-containing structural subunits termed scaffoldins. Multiple catalytic enzymes are integrated into the complex through the strong specific non-covalent interactions between dockerins and cohesins [85,86]. *C. thermocellum* is the first reported cellulosome producer and its cellulosome integrates more than 70 enzymatic subunits with the most abundant subunits including exoglucanases Cel48S and endoglucanases Cel9Q, Cel9R, and Cel5G [19,87]. Besides, GH48 exoglucanases displaying distinct synergies with different GH9 enzymes are observed in the cellulosomes from *C. cellulolyticum* [88] and *Ruminococcus champanellensis* [89]. Nevertheless, extensive intra- and inter-molecular synergies and the physical proximity of enzymatic components are determined to be critical for the high efficiency of cellulosomes. An artificial minicellulosome constructed with a modified Cel48S and multi-endoglucanases (GH8 and GH9) demonstrates the high conversion of crystalline cellulose at a high temperature of 60 °C [79].

8. Conclusions and Outlook

Members of the GH48 family are broadly distributed, mainly in bacteria but also in Fungi and Arthropoda [25]. Genomics and metagenomics have revealed a large number of GH48 sequences, demonstrating the prevalence of GH48 enzymes in various lignocellulolytic systems. It is agreed that GH48 cellulases are among the most abundant CAZymes and play a key role in processes associated with lignocellulose biorefinery, but the function and activity of these GH48 enzymes remain uncharacterized. The available structures of various GH48 cellulases provide key information for their catalytic mechanism, substrate recognition, product inhibition, stability, and even evolution. In light of the structures of GH48 cellulases with or without the oligosaccharides, extensive engineering efforts have been made to improve thermal stability, reduce end-product inhibition, and explore the mechanism of product expulsion. GH48 cellulases synergize greatly with the complementary endocellulase intra- or inter-molecularly to improve their performance, providing great values in lignocellulose biorefinery, but the mechanism of synergistic effects has not been fully understood.

Despite the great importance of GH48 enzymes in lignocellulose-degrading systems, the knowledge of these enzymes is still far from complete. Among a large number of GH48 sequences available in the database, only a few of them have been functionally and structurally characterized. The catalytic mechanism, particularly the progressive mechanism and product-releasing mechanism, are still not clear. The current information about the mechanism of synergistic effects of GH48 and other family GHs in different cellulolytic systems is still very limited and mostly descriptive. Compared to other GH families, the engineering studies of GH48 enzymes are much less and lack depth and width. One reason is that the recombinant expression of GH48 cellulases in *E. coli* often results in the formation of inclusion bodies [43,90]. Although both the refolding of GH48 enzymes from the inclusion bodies and direct purification of GH48 enzymes from their original species have been reported in the literature [43,53], the complexity of these methods makes the high-throughput screening of engineered mutants difficult and infeasible. The activity assays of GH48 enzymes often use crystalline cellulose as the substrate and generate cellobiose as the product, leading to difficulty in spectroscopy determination and time-consuming analysis is indispensable. The development of new methods to obtain soluble

expression of GH48 enzymes and the design of convenient assays for cellulase activity are needed. Therefore, in the future, the research and exploration of GH48s for practical utilization in lignocellulose biorefinery should focus on the following aspects: mining and characterizing new GH48s, thoroughly understanding the relationship between the structure, dynamics, and functions (particularly the progressive activity and product inhibition) of GH48 enzymes, engineering GH48 enzymes to improve the properties, and designing and developing new biocatalytic system employing the elaborate synergistic effects between GH48 and other family enzymes. Recent progresses in artificial intelligence for multi-omics analysis, protein structure prediction, protein design, and laboratory automation will accelerate the theoretical research and applicational technique development of the important but lagging GH48 enzymes, promoting their key roles in lignocellulose biorefinery for a future green circular economy.

Author Contributions: Conceptualization, Y.F. and Q.C.; writing—original draft preparation, C.Y.; writing—review and editing, Y.-J.L., Q.C. and Y.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Strategic Priority Research Program of the Chinese Academy of Sciences (grant no. XDA21060201 to Q.C.); National Natural Science Foundation of China (grant nos. 32170051 to Q.C., 32070125 to Y.F., 32200030 to C.Y., 32070028 to Y.-J.L.); Shandong Energy Institute (SEI) (grant nos. SEI S202106 to Q.C., SEI I202106 to Y.F., SEI I202142 to Y.-J.L.); Qingdao Independent Innovation Major Project (grant no. 21-1-2-23-hz to Y.-J.L.).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. McKendry, P. Energy production from biomass (part 1): Overview of biomass. *Bioresour. Technol.* **2002**, *83*, 37–46. [CrossRef] [PubMed]
2. Ogonda, L.A.; Saumonneau, A.; Dion, M.; Muge, E.K.; Wamalwa, B.M.; Mulaa, F.J.; Tellier, C. Characterization and engineering of two new GH9 and GH48 cellulases from a *Bacillus pumilus* isolated from Lake Bogoria. *Biotechnol. Lett.* **2021**, *43*, 691–700. [CrossRef] [PubMed]
3. Himmel, M.E.; Bayer, E.A. Lignocellulose conversion to biofuels: Current challenges, global perspectives. *Curr. Opin. Biotechnol.* **2009**, *20*, 316–317. [CrossRef] [PubMed]
4. Cragg, S.M.; Beckham, G.T.; Bruce, N.C.; Bugg, T.D.H.; Distel, D.L.; Dupree, P.; Etxabe, A.G.; Goodell, B.S.; Jellison, J.; McGeehan, J.E.; et al. Lignocellulose degradation mechanisms across the Tree of Life. *Curr. Opin. Chem. Biol.* **2015**, *29*, 108–119. [CrossRef] [PubMed]
5. Taha, M.; Foda, M.; Shahsavari, E.; Aburto-Medina, A.; Adetutu, E.; Ball, A. Commercial feasibility of lignocellulose biodegradation: Possibilities and challenges. *Curr. Opin. Biotechnol.* **2016**, *38*, 190–197. [CrossRef] [PubMed]
6. Alessi, A.M.; Bird, S.M.; Oates, N.C.; Li, Y.; Dowle, A.A.; Novotny, E.H.; deAzevedo, E.R.; Bennett, J.P.; Polikarpov, I.; Young, J.P.W.; et al. Defining functional diversity for lignocellulose degradation in a microbial community using multi-omics studies. *Biotechnol. Biofuels* **2018**, *11*, 166. [CrossRef]
7. Artzi, L.; Bayer, E.A.; Moraïs, S. Cellulosomes: Bacterial nanomachines for dismantling plant polysaccharides. *Nat. Rev. Microbiol.* **2017**, *15*, 83–95. [CrossRef]
8. Payne, C.M.; Knott, B.C.; Mayes, H.B.; Hansson, H.; Himmel, M.E.; Sandgren, M.; Ståhlberg, J.; Beckham, G.T. Fungal cellulases. *Chem. Rev.* **2015**, *115*, 1308–1448. [CrossRef]
9. Østby, H.; Hansen, L.D.; Horn, S.J.; Eijsink, V.G.H.; Várnai, A. Enzymatic processing of lignocellulosic biomass: Principles, recent advances and perspectives. *J. Ind. Microbiol. Biotechnol.* **2020**, *47*, 623–657. [CrossRef]
10. Shallom, D.; Shoham, Y. Microbial hemicellulases. *Curr. Opin. Microbiol.* **2003**, *6*, 219–228. [CrossRef]
11. Kamimura, N.; Sakamoto, S.; Mitsuda, N.; Masai, E.; Kajita, S. Advances in microbial lignin degradation and its applications. *Curr. Opin. Biotechnol.* **2019**, *56*, 179–186. [CrossRef]
12. Liu, Y.-J.; Li, B.; Feng, Y.; Cui, Q. Consolidated bio-saccharification: Leading lignocellulose bioconversion into the real world. *Biotechnol. Adv.* **2020**, *40*, 107535. [CrossRef]
13. Mankar, A.R.; Pandey, A.; Modak, A.; Pant, K.K. Pretreatment of lignocellulosic biomass: A review on recent advances. *Bioresour. Technol.* **2021**, *334*, 125235. [CrossRef] [PubMed]

14. Hagen, L.H.; Brooke, C.G.; Shaw, C.A.; Norbeck, A.D.; Piao, H.; Arntzen, M.Ø.; Olson, H.M.; Copeland, A.; Isern, N.; Shukla, A.; et al. Proteome specialization of anaerobic fungi during ruminal degradation of recalcitrant plant fiber. *ISME J.* **2021**, *15*, 421–434. [CrossRef]
15. Conway, J.M.; Crosby, J.R.; Hren, A.P.; Southerland, R.T.; Lee, L.L.; Lunin, V.V.; Alahuhta, P.; Himmel, M.E.; Bomble, Y.J.; Adams, M.W.W.; et al. Novel multidomain, multifunctional glycoside hydrolases from highly lignocellulolytic *Caldicellulosiruptor* species. *AIChE J.* **2018**, *64*, 4218–4228. [CrossRef]
16. Li, J.; Su, X.; Tian, Y.; Dong, Z.; Hu, S.; Huang, L.; Dai, X. Gene diversity of the bacterial 48 family glycoside hydrolase (GH48) in rumen environment. *Acta Microbiol. Sin.* **2014**, *54*, 53–61.
17. Chen, M.; Bu, L.; Alahuhta, M.; Brunecky, R.; Xu, Q.; Lunin, V.V.; Brady, J.W.; Crowley, M.F.; Himmel, M.E.; Bomble, Y.J. Strategies to reduce end-product inhibition in family 48 glycoside hydrolases. *Proteins Struct. Funct. Bioinform.* **2016**, *84*, 295–304. [CrossRef] [PubMed]
18. Vazana, Y.; Moraïs, S.; Barak, Y.; Lamed, R.; Bayer, E.A. Interplay between *Clostridium thermocellum* family 48 and family 9 cellulases in cellulosomal versus noncellulosomal states. *Appl. Environ. Microbiol.* **2010**, *76*, 3236–3243. [CrossRef]
19. Bayer, E.A.; Lamed, R.; White, B.A.; Flint, H.J. From cellulosomes to cellulosomics. *Chem. Rec.* **2008**, *8*, 364–377. [CrossRef]
20. Kim, S.-K.; Chung, D.; Himmel, M.E.; Bomble, Y.J.; Westpheling, J. Engineering the N-terminal end of CelA results in improved performance and growth of *Caldicellulosiruptor bescii* on crystalline cellulose. *Biotechnol. Bioeng.* **2017**, *114*, 945–950. [CrossRef]
21. Yi, Z.; Su, X.; Revindran, V.; Mackie, R.I.; Cann, I. Molecular and biochemical analyses of CbCel9A/Cel48A, a highly secreted multi-modular cellulase by *Caldicellulosiruptor bescii* during growth on crystalline cellulose. *PLoS ONE* **2013**, *8*, e84172. [CrossRef] [PubMed]
22. Rettenmaier, R.; Lo, Y.K.; Schmidt, L.; Munk, B.; Lagkouvardos, I.; Neuhaus, K.; Schwarz, W.; Liebl, W.; Zverlov, V. A novel primer mixture for GH48 genes: Quantification and identification of truly cellulolytic bacteria in biogas fermenters. *Microorganisms* **2020**, *8*, 1297. [CrossRef] [PubMed]
23. Drula, E.; Garron, M.-L.; Dogan, S.; Lombard, V.; Henrissat, B.; Terrapon, N. The carbohydrate-active enzyme database: Functions and literature. *Nucleic Acids Res.* **2022**, *50*, 571–577. [CrossRef] [PubMed]
24. Brás, J.L.A.; Cartmell, A.; Carvalho, A.L.M.; Verzé, G.; Bayer, E.A.; Vazana, Y.; Correia, M.A.S.; Prates, J.A.M.; Ratnaparkhe, S.; Boraston, A.B.; et al. Structural insights into a unique cellulase fold and mechanism of cellulose hydrolysis. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 5237–5242. [CrossRef]
25. Sukharnikov, L.O.; Alahuhta, M.; Brunecky, R.; Upadhyay, A.; Himmel, M.E.; Lunin, V.V.; Zhulin, I.B. Sequence, structure, and evolution of cellulases in glycoside hydrolase family 48. *J. Biol. Chem.* **2012**, *287*, 41068–41077. [CrossRef] [PubMed]
26. Tsai, L.C.; Amiraslanov, I.; Chen, H.R.; Chen, Y.W.; Lee, H.L.; Liang, P.H.; Liaw, Y.C. Structures of exoglucanase from *Clostridium cellulovorans*: Cellotetraose binding and cleavage. *Acta Crystallogr. F Struct. Biol. Commun.* **2015**, *71*, 1264–1272. [CrossRef]
27. Zhang, H.; Zhang, J.-L.; Sun, L.; Niu, X.-D.; Wang, S.; Shan, Y.-M. Molecular dynamics simulation of the processive endocellulase Cel48F from *Clostridium cellulolyticum*: A novel "water-control mechanism" in enzymatic hydrolysis of cellulose. *J. Mol. Recognit.* **2014**, *27*, 438–447. [CrossRef]
28. Araújo, E.A.; Dias, A.H.S.; Kadowaki, M.A.S.; Piyadov, V.; Pellegrini, V.O.A.; Urio, M.B.; Ramos, L.P.; Skaf, M.S.; Polikarpov, I. Impact of cellulose properties on enzymatic degradation by bacterial GH48 enzymes: Structural and mechanistic insights from processive *Bacillus licheniformis* Cel48B cellulase. *Carbohydr. Polym.* **2021**, *264*, 118059. [CrossRef]
29. Penneru, S.K.; Saharay, M.; Krishnan, M. CelS-catalyzed processive cellulose degradation and cellobiose extraction for the production of bioethanol. *J. Chem. Inf. Model.* **2022**, *62*, 6628–6638. [CrossRef]
30. 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]
31. Usmani, Z.; Sharma, M.; Awasthi, A.K.; Lukk, T.; Tuohy, M.G.; Gong, L.; Nguyen-Tri, P.; Goddard, A.D.; Bill, R.M.; Nayak, S.C.; et al. Lignocellulosic biorefineries: The current state of challenges and strategies for efficient commercialization. *Renew. Sustain. Energy Rev.* **2021**, *148*, 111258. [CrossRef]
32. Liu, Y.; Cruz-Morales, P.; Zargar, A.; Belcher, M.S.; Pang, B.; Englund, E.; Dan, Q.; Yin, K.; Keasling, J.D. Biofuels for a sustainable future. *Cell* **2021**, *184*, 1636–1647. [CrossRef] [PubMed]
33. Devillard, E.; Goodheart, D.B.; Karnati, S.K.R.; Bayer, E.A.; Lamed, R.; Miron, J.; Nelson, K.E.; Morrison, M. *Ruminococcus albus* 8 mutants defective in cellulose degradation are deficient in two processive endocellulases, Cel48A and Cel9B, both of which possess a novel modular architecture. *J. Bacteriol.* **2004**, *186*, 136–145. [CrossRef] [PubMed]
34. Izquierdo, J.A.; Sizova, M.V.; Lynd, L.R. Diversity of bacteria and glycosyl hydrolase family 48 genes in cellulolytic consortia enriched from thermophilic biocompost. *Appl. Environ. Microbiol.* **2010**, *76*, 3545–3553. [CrossRef]
35. Olson, D.G.; Tripathi, S.A.; Giannone, R.J.; Lo, J.; Caiazza, N.C.; Hogsett, D.A.; Hettich, R.L.; Guss, A.M.; Dubrovsky, G.; Lynd, L.R. Deletion of the Cel48S cellulase from *Clostridium thermocellum*. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 17727–17732. [CrossRef]
36. Young, J.; Chung, D.; Bomble, Y.J.; Himmel, M.E.; Westpheling, J. Deletion of *Caldicellulosiruptor bescii* CelA reveals its crucial role in the deconstruction of lignocellulosic biomass. *Biotechnol. Biofuels* **2014**, *7*, 142. [CrossRef]
37. Dai, X.; Tian, Y.; Li, J.; Luo, Y.; Liu, D.; Zheng, H.; Wang, J.; Dong, Z.; Hu, S.; Huang, L. Metatranscriptomic analyses of plant cell wall polysaccharide degradation by microorganisms in the cow rumen. *Appl. Environ. Microbiol.* **2015**, *81*, 1375–1386. [CrossRef]

38. Chu, Y.; Hao, Z.; Wang, K.; Tu, T.; Huang, H.; Wang, Y.; Bai, Y.G.; Wang, Y.; Luo, H.; Yao, B.; et al. The GH10 and GH48 dual-functional catalytic domains from a multimodular glycoside hydrolase synergize in hydrolyzing both cellulose and xylan. *Biotechnol. Biofuels* **2019**, *12*, 279. [CrossRef]
39. Lee, L.L.; Blumer-Schuette, S.E.; Izquierdo, J.A.; Zurawski, J.V.; Loder, A.J.; Conway, J.M.; Elkins, J.G.; Podar, M.; Clum, A.; Jones, P.C.; et al. Genus-wide assessment of lignocellulose utilization in the extremely thermophilic genus *Caldicellulosiruptor* by genomic, pangenomic, and metagenomic analyses. *Appl. Environ. Microbiol.* **2018**, *84*, e02694-17. [CrossRef]
40. Pereyra, L.P.; Hiibel, S.R.; Prieto Riquelme, M.V.; Reardon, K.F.; Pruden, A. Detection and quantification of functional genes of cellulose- degrading, fermentative, and sulfate-reducing bacteria and methanogenic archaea. *Appl. Environ. Microbiol.* **2010**, *76*, 2192–2202. [CrossRef]
41. Brunecky, R.; Alahuhta, M.; Xu, Q.; Donohoe, B.S.; Crowley, M.F.; Kataeva, I.A.; Yang, S.-J.; Resch, M.G.; Adams, M.W.W.; Lunin, V.V.; et al. Revealing nature's cellulase diversity: The digestion mechanism of *Caldicellulosiruptor bescii* CelA. *Science* **2013**, *342*, 1513–1516. [CrossRef] [PubMed]
42. Wang, W.K.; Kruus, K.; Wu, J.H.D. Cloning and DNA sequence of the gene coding for *Clostridium thermocellum* cellulase Ss (CelS), a major cellulosome component. *J. Bacteriol.* **1993**, *175*, 1293–1302. [CrossRef] [PubMed]
43. Wang, W.K.; Kruus, K.; Wu, J.H.D. Cloning and expression of the *Clostridium thermocellum* celS gene in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **1994**, *42*, 346–352. [CrossRef] [PubMed]
44. Bagnara-Tardif, C.; Gaudin, C.; Belaich, A.; Hoest, P.; Citard, T.; Belaich, J.-P. Sequence analysis of a gene cluster encoding cellulases from *Clostridium cellulolyticum*. *Gene* **1992**, *119*, 17–28. [CrossRef] [PubMed]
45. Conway, J.M.; McKinley, B.S.; Seals, N.L.; Hernandez, D.; Khatibi, P.A.; Poudel, S.; Giannone, R.J.; Hettich, R.L.; Williams-Rhaesa, A.M.; Lipscomb, G.L.; et al. Functional analysis of the glucan degradation locus in *Caldicellulosiruptor bescii* reveals essential roles of component glycoside hydrolases in plant biomass deconstruction. *Appl. Environ. Microbiol.* **2017**, *83*, e01828-17. [CrossRef] [PubMed]
46. Zverlov, V.; Mahr, 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* **1998**, *144*, 457–465. [CrossRef] [PubMed]
47. Kostylev, M.; Wilson, D. A distinct model of synergism between a processive endocellulase (TfCel9A) and an exocellulase (TfCel48A) from *Thermobifida fusca*. *Appl. Environ. Microbiol.* **2014**, *80*, 339–344. [CrossRef]
48. Eyun, S.-I.; Wang, H.; Pauchet, Y.; Ffrench-Constant, R.H.; Benson, A.K.; Valencia-Jiménez, A.; Moriyama, E.N.; Siegfried, B.D. Molecular evolution of glycoside hydrolase genes in the Western corn rootworm (*Diabrotica virgifera virgifera*). *PLoS ONE* **2014**, *9*, e94052. [CrossRef]
49. Comtet-Marre, S.; Parisot, N.; Lepercq, P.; Chaucheyras-Durand, F.; Mosoni, P.; Peyretaillade, E.; Bayat, A.R.; Shingfield, K.J.; Peyret, P.; Forano, E. Metatranscriptomics reveals the active bacterial and eukaryotic fibrolytic communities in the rumen of dairy cow fed a mixed diet. *Front. Microbiol.* **2017**, *8*, 67. [CrossRef]
50. Parsiegla, G.; Juy, M.; Reverbel-Leroy, C.; Tardif, C.; Belaïch, J.P.; Driguez, H.; Haser, R. The crystal structure of the processive endocellulase CelF of *Clostridium cellulolyticum* in complex with a thiooligosaccharide inhibitor at 2.0 Å resolution. *EMBO J.* **1998**, *17*, 5551–5562. [CrossRef]
51. Parsiegla, G.; Reverbel-Leroy, C.; Tardif, C.; Belaich, J.P.; Driguez, H.; Haser, R. Crystal structures of the cellulase Cel48F in complex with inhibitors and substrates give insights into its processive action. *Biochemistry* **2000**, *39*, 11238–11246. [CrossRef] [PubMed]
52. Parsiegla, G.; Reverbel, C.; Tardif, C.; Driguez, H.; Haser, R. Structures of mutants of cellulase Cel48F of *Clostridium cellulolyticum* in complex with long hemithiocellooligosaccharides give rise to a new view of the substrate pathway during processive action. *J. Mol. Biol.* **2008**, *375*, 499–510. [CrossRef] [PubMed]
53. Liu, Y.-J.; Liu, S.; Dong, S.; Li, R.; Feng, Y.; Cui, Q. Determination of the native features of the exoglucanase Cel48S from *Clostridium thermocellum*. *Biotechnol. Biofuels* **2018**, *11*, 6. [CrossRef] [PubMed]
54. Guimarães, B.G.; Souchon, H.; Lytle, B.L.; David Wu, J.H.; Alzari, P.M. The crystal structure and catalytic mechanism of cellobiohydrolase CelS, the major enzymatic component of the *Clostridium thermocellum* cellulosome. *J. Mol. Biol.* **2002**, *320*, 587–596. [CrossRef] [PubMed]
55. Berger, E.; Zhang, D.; Zverlov, V.V.; Schwarz, W.H. Two noncellulosomal cellulases of *Clostridium thermocellum*, Cel9I and Cel48Y, hydrolyse crystalline cellulose synergistically. *FEMS Microbiol. Lett.* **2007**, *268*, 194–201. [CrossRef]
56. Kruus, K.; Wang, W.K.; Ching, J.; Wu, J.H.D. Exoglucanase activities of the recombinant *Clostridium thermocellum* CelS, a major cellulosome component. *J. Bacteriol.* **1995**, *177*, 1641–1644. [CrossRef]
57. Bronnenmeier, K.; Kundt, K.; Riedel, K.; Schwarz, W.H.; Staudenbauer, W.L. Structure of the *Clostridium stercorarium* gene *celY* encoding the exo-1,4-β-glucanase Avicelase II. *Microbiology* **1997**, *143*, 891–898. [CrossRef]
58. Artzi, L.; Morag, E.; Shamshoum, M.; Bayer, E.A. Cellulosomal expansin: Functionality and incorporation into the complex. *Biotechnol. Biofuels* **2016**, *9*, 61. [CrossRef]
59. Zhang, X.-Z.; Zhang, Z.; Zhu, Z.; Sathitsuksanoh, N.; Yang, Y.; Zhang, Y.-H.P. The noncellulosomal family 48 cellobiohydrolase from *Clostridium phytofermentans* ISDg: Heterologous expression, characterization, and processivity. *Appl. Microbiol. Biotechnol.* **2010**, *86*, 525–533. [CrossRef]

60. Brunecky, R.; Alahuhta, M.; Sammond, D.W.; Xu, Q.; Chen, M.; Wilson, D.B.; Brady, J.W.; Himmel, M.E.; Bomble, Y.J.; Lunin, V.V. Natural diversity of glycoside hydrolase family 48 exoglucanases: Insights from structure. *Biotechnol. Biofuels* **2017**, *10*, 274. [CrossRef]
61. Kostylev, M.; Alahuhta, M.; Chen, M.; Brunecky, R.; Himmel, M.E.; Lunin, V.V.; Brady, J.; Wilson, D.B. Cel48A from *Thermobifida fusca*: Structure and site directed mutagenesis of key residues. *Biotechnol. Bioeng.* **2014**, *111*, 664–673. [CrossRef]
62. Shen, H.; Gilkes, N.R.; Kilburn, D.G.; Miller Jr, R.C.; Warren, R.A. Cellobiohydrolase B, a second exo-cellobiohydrolase from the cellulolytic bacterium *Cellulomonas fimi*. *Biochem. J.* **1995**, *311*, 67–74. [CrossRef]
63. Cai, S.; Li, J.; Hu, F.Z.; Zhang, K.; Luo, Y.; Janto, B.; Boissy, R.; Ehrlich, G.; Dong, X. *Cellulosilyticum ruminicola*, a newly described rumen bacterium that possesses redundant fibrolytic-protein-encoding genes and degrades lignocellulose with multiple carbohydrate- borne fibrolytic enzymes. *Appl. Environ. Microbiol.* **2010**, *76*, 3818–3824. [CrossRef] [PubMed]
64. Ramírez-Ramírez, N.; Romero-García, E.R.; Calderón, V.C.; Avitia, C.I.; Téllez-Valencia, A.; Pedraza-Reyes, M. Expression, characterization and synergistic interactions of *Myxobacter* Sp. AL-1 Cel9 and Cel48 glycosyl hydrolases. *Int. J. Mol. Sci.* **2008**, *9*, 247–257. [CrossRef]
65. Sánchez, M.M.; Pastor, F.I.J.; Diaz, P. Exo-mode of action of cellobiohydrolase Cel48C from *Paenibacillus* sp. BP-23. A unique type of cellulase among Bacillales. *Eur. J. Biochem.* **2003**, *270*, 2913–2919. [CrossRef] [PubMed]
66. Qian, M.; Guan, S.; Shan, Y.; Zhang, H.; Wang, S. Structural and molecular basis of cellulase Cel48F by computational modeling: Insight into catalytic and product release mechanism. *J. Struct. Biol.* **2016**, *194*, 347–356. [CrossRef] [PubMed]
67. Saharay, M.; Guo, H.; Smith, J.C. Catalytic mechanism of cellulose degradation by a cellobiohydrolase, CelS. *PLoS ONE* **2010**, *5*, e12947. [CrossRef]
68. Kurasin, M.; Väljamäe, P. Processivity of cellobiohydrolases is limited by the substrate. *J. Biol. Chem.* **2011**, *286*, 169–177. [CrossRef]
69. Pellegrini, V.O.A.; Bernardes, A.; Rezende, C.A.; Polikarpov, I. Cellulose fiber size defines efficiency of enzymatic hydrolysis and impacts degree of synergy between endo- and exoglucanases. *Cellulose* **2018**, *25*, 1865–1881. [CrossRef]
70. Morag, E.; Halevy, I.; Bayer, E.A.; Lamed, R. Isolation and properties of a major cellobiohydrolase from the cellulosome of *Clostridium thermocellum*. *J. Bacteriol.* **1991**, *173*, 4155–4162. [CrossRef]
71. Irwin, D.C.; Zhang, S.; Wilson, D.B. Cloning, expression and characterization of a family 48 exocellulase, Cel48A, from *Thermobifida fusca*. *Eur. J. Biochem.* **2000**, *267*, 4988–4997. [CrossRef] [PubMed]
72. Bu, L.; Crowley, M.F.; Himmel, M.E.; Beckham, G.T. Computational investigation of the pH dependence of loop flexibility and catalytic function in glycoside hydrolases. *J. Biol. Chem.* **2013**, *288*, 12175–12186. [CrossRef]
73. Bu, L.; Nimlos, M.R.; Shirts, M.R.; Ståhlberg, J.; Himmel, M.E.; Crowley, M.F.; Beckham, G.T. Product binding varies dramatically between processive and nonprocessive cellulase enzymes. *J. Biol. Chem.* **2012**, *287*, 24807–24813. [CrossRef] [PubMed]
74. Lee, H.-J.; Kang, T.-G.; Kim, Y.-W.; Lee, H.-S.; Kim, S.-K. Functional expression and extracellular secretion of *Clostridium thermocellum* Cel48S cellulase in *Escherichia coli* via the signal recognition particle-dependent translocation pathway. *Enzyme Microb. Technol.* **2021**, *151*, 109918. [CrossRef]
75. Smith, M.A.; Rentmeister, A.; Snow, C.D.; Wu, T.; Farrow, M.F.; Mingardon, F.; Arnold, F.H. A diverse set of family 48 bacterial glycoside hydrolase cellulases created by structure-guided recombination. *FEBS J.* **2012**, *279*, 4453–4465. [CrossRef] [PubMed]
76. Reese, E.T.; Siu, R.G.H.; Levinson, H.S. The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. *J. Bacteriol.* **1950**, *59*, 485–497. [CrossRef] [PubMed]
77. Wood, T.M.; McCrae, S.I. Synergism Between Enzymes Involved in the Solubilization of Native Cellulose. In *Hydrolysis of Cellulose: Mechanisms of Enzymatic and Acid Catalysis*; Brown, R.D., Jurasek, L., Eds.; American Chemical Society: Washington, DC, USA, 1979; pp. 181–209. ISBN 9780841204607.
78. Kostylev, M.; Wilson, D. Synergistic interactions in cellulose hydrolysis. *Biofuels* **2012**, *3*, 61–70. [CrossRef]
79. 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]
80. 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]
81. Lochner, A.; Giannone, R.J.; Rodriguez, M.; Shah, M.B.; Mielenz, J.R.; Keller, M.; Antranikian, G.; Graham, D.E.; Hettich, R.L. Use of label-free quantitative proteomics to distinguish the secreted cellulolytic systems of *Caldicellulosiruptor bescii* and *Caldicellulosiruptor obsidiansis*. *Appl. Environ. Microbiol.* **2011**, *77*, 4042–4054. [CrossRef]
82. Bing, W.; Wang, H.; Zheng, B.; Zhang, F.; Zhu, G.; Feng, Y.; Zhang, Z. *Caldicellulosiruptor changbaiensis* sp. nov., a cellulolytic and hydrogen-producing bacterium from a hot spring. *Int. J. Syst. Evol. Microbiol.* **2015**, *65*, 293–297. [CrossRef] [PubMed]
83. Miroshnichenko, M.L.; Kublanov, I.V.; Kostrikina, N.A.; Tourova, T.P.; Kolganova, T.V.; Birkeland, N.-K.; Bonch-Osmolovskaya, E.A. *Caldicellulosiruptor kronotskyensis* sp. nov. and *Caldicellulosiruptor hydrothermalis* sp. nov., two extremely thermophilic, cellulolytic, anaerobic bacteria from Kamchatka thermal springs. *Int. J. Syst. Evol. Microbiol.* **2008**, *58*, 1492–1496. [CrossRef] [PubMed]
84. Feng, Y.; Liu, Y.; Cui, Q. Research progress in cellulosomes and their applications in synthetic biology. *Synth. Biol. J.* **2022**, *3*, 138–154. [CrossRef]
85. Yao, X.; Chen, C.; Wang, Y.; Dong, S.; Liu, Y.-J.; Li, Y.; Cui, Z.; Gong, W.; Perrett, S.; Yao, L.; et al. Discovery and mechanism of a pH-dependent dual-binding-site switch in the interaction of a pair of protein modules. *Sci. Adv.* **2020**, *6*, eabd7182. [CrossRef]

86. Bule, P.; Pires, V.; Fontes, C.M.; Alves, V.D. Cellulosome assembly: Paradigms are meant to be broken! *Curr. Opin. Struct. Biol.* **2018**, *49*, 154–161. [CrossRef]
87. Yoav, S.; Barak, Y.; Shamshoum, M.; Borovok, I.; Lamed, R.; Dassa, B.; Hadar, Y.; Morag, E.; Bayer, E.A. How does cellulosome composition influence deconstruction of lignocellulosic substrates in *Clostridium* (*Ruminiclostridium*) *thermocellum* DSM 1313? *Biotechnol. Biofuels* **2017**, *10*, 222. [CrossRef]
88. Ravachol, J.; Borne, R.; Tardif, C.; de Philip, P.; Fierobe, H.-P. Characterization of all family-9 glycoside hydrolases synthesized by the cellulosome-producing bacterium *Clostridium cellulolyticum*. *J. Biol. Chem.* **2014**, *289*, 7335–7348. [CrossRef]
89. Moraïs, S.; Ben David, Y.; Bensoussan, L.; Duncan, S.H.; Koropatkin, N.M.; Martens, E.C.; Flint, H.J.; Bayer, E.A. Enzymatic profiling of cellulosomal enzymes from the human gut bacterium, *Ruminococcus champanellensis*, reveals a fine-tuned system for cohesin-dockerin recognition. *Environ. Microbiol.* **2016**, *18*, 542–556. [CrossRef] [PubMed]
90. Reverbel-Leroy, C.; Belaich, A.; Bernadac, A.; Gaudin, C.; Belaich, J.-P.; Tardif, C. Molecular study and overexpression of the *Clostridium cellulolyticum celF* cellulase gene in *Escherichia coli*. *Microbiology* **1996**, *142*, 1013–1023. [CrossRef]