

Article

CalkGH9T: A Glycoside Hydrolase Family 9 Enzyme from *Clostridium alkallicellulosi*

Paripok Phitsuwan ^{1,2,*}, Sengthong Lee ², Techly San ² and Khanok Ratanakhanokchai ²

¹ LigniTech-Liginin Technology Research Group, School of Bioresources and Technology, King Mongkut's University of Technology Thonburi, Bangkok 10150, Thailand

² Division of Biochemical Technology, School of Bioresources and Technology, King Mongkut's University of Technology Thonburi, Bangkok 10150, Thailand; sengthong.k55132@mail.kmutt.ac.th (S.L.); santechly123@gmail.com (T.S.); khanok.rat@kmutt.ac.th (K.R.)

* Correspondence: paripok.phii@kmutt.ac.th; Tel.: +66-247-7751

Abstract: Glycoside hydrolase family 9 (GH9) endoglucanases are important enzymes for cellulose degradation. However, their activity on cellulose is diverse. Here, we cloned and expressed one GH9 enzyme (*CalkGH9T*) from *Clostridium alkallicellulosi* in *Escherichia coli*. *CalkGH9T* has a modular structure, containing one GH9 catalytic module, two family 3 carbohydrate binding modules, and one type I dockerin domain. *CalkGH9T* exhibited maximal activity at pH 7.0–8.0 and 55 °C and was resistant to urea and NaCl. It efficiently hydrolyzed carboxymethyl cellulose (CMC) but poorly degraded regenerated amorphous cellulose (RAC). Despite strongly binding to Avicel, *CalkGH9T* lacked the ability to hydrolyze this substrate. The hydrolysis of CMC by *CalkGH9T* produced a series of cello-oligomers, with cellotetraose being preferentially released. Similar proportions of soluble and insoluble reducing ends generated by hydrolysis of RAC indicated non-processive activity. Our study extends our knowledge of the molecular mechanism of cellulose hydrolysis by GH9 family endoglucanases with industrial relevance.

Keywords: carbohydrate binding module; crystalline cellulose; endoglucanase; processive activity



Citation: Phitsuwan, P.; Lee, S.; San, T.; Ratanakhanokchai, K. *CalkGH9T*: A Glycoside Hydrolase Family 9 Enzyme from *Clostridium alkallicellulosi*. *Catalysts* **2021**, *11*, 1011. <https://doi.org/10.3390/catal11081011>

Academic Editors: Michele C. Loewen and Peter Adewale

Received: 2 July 2021
Accepted: 17 August 2021
Published: 22 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cellulose is the most abundant polysaccharide in plant cell walls. It is composed of glucose units linked by β -glucosidic bonds. Each glucose residue is rotated 180° relative to its adjacent residues, yielding cellobiose as the structural repeating unit of cellulose [1]. In addition, this structural configuration makes the glucan chain linear. This chain linearity allows the glucan chains to contact other chains, forming intra- and inter-hydrogen bonding within and between the chains. This close association between cellulose chains results in a tightly packed, crystalline polymeric structure. This rigid and strong cellulose polymer provides structural integrity to plants and recalcitrance to microbial attack [2].

Microorganisms, including fungi and bacteria, have developed enzyme systems to gain sugar from cellulose [1,3]. Cellulose-degrading enzymes, or cellulases, are collectively defined as enzymes that hydrolyze β -glucosidic linked cellulose chains [1]. Indeed, cellulases contain three classes of enzymes according to their modes of actions to cleave cellulose. β -1,4-Endoglucanases (EC3.2.1.4) randomly cut the intramolecular β -1,4-glucosidic bond of cellulose chains to create new chain ends, and cellobiohydrolases or exoglucanases (EC3.2.1.91 and EC3.2.1.176) processively attack at the ends of cellulose chains to produce soluble sugars, particularly cellobiose. β -Glucosidases (EC3.2.1.21) then cleave cellobiose to yield glucose. Therefore, the synergistic interaction of three classes of cellulases are believed to completely convert cellulose to glucose [4,5].

According to the Carbohydrate Active Enzyme Database (<http://www.cazy.org/>, accessed on 3 December 2020), cellulases are grouped into several glycoside hydrolase (GH) families based on the amino acid sequence similarity of their catalytic modules.

Cellulases are lodged into GH families: 1, 3, 5, 6, 7, 8, 9, 10, 12, 16, 44, 45, 48, 51, and 61; among these, GH9 cellulases are distinctive from cellulases in other families [6–9]. The GH9 family exclusively contains enzymes with endoglucanase activity. Interestingly, these endoglucanases can show either processive or non-processive activities [6]. The endoglucanases with processive activity, called processive endoglucanases, behave like exoglucanases during hydrolysis of insoluble cellulose. This means that the enzymes can attach to the substrate and progressively hydrolyze it before diffusion occurs [10]. However, the catalytic performance of GH9 endoglucanases varies, and it cannot be precisely predicted by sequence similarity, as individual GH9 endoglucanases have shown different substrate preferences and diverse hydrolysis activities on different forms of cellulose [6,10–12].

In addition, several GH9 endoglucanases are modular enzymes, containing accessory modules such as immunoglobulin (Ig)-like modules, fibronectin III-like domains, and carbohydrate binding modules (CBMs) [6,13]. Family 3 CBMs are involved in binding to cellulose, and three major subgroups (3a, 3b and 3c) were classified in the family 3 CBM on the basis of primary sequence similarity [8]. Subfamilies 3a and 3b bind to the surface of microcrystalline cellulose strongly and promote the cellulolytic rate by concentrating the enzyme near the cellulose surface. Unlike Subfamilies 3a and 3b, subfamily 3c binds weakly to cellulose due to the lack of several conserved aromatic residues that are important for the strong binding to cellulose [14]. However, CBM3c is believed to convey a single cellulose chain into the nearby catalytic domain and plays an essential role for processivity [8,14,15]. Therefore, the functions of these modules (immunoglobulin (Ig)-like modules, fibronectin III-like domains, and CBMs) remain unclear; however, they are believed to be helper modules that are involved in structural stability or the hydrolysis efficiency of GH9 enzymes [10,15]. The modular structure and functionality of GH9 endoglucanases require comprehensive investigation. In addition, a more complicated modular organization of GH9 endoglucanases exist in the genomes of some cellulosome-producing microorganisms, with a GH9 catalytic module connected downstream to two family 3 CBMs and one dockerin [16,17]. Unfortunately, data on the functional roles of these individual modules during cellulose hydrolysis by this GH9 theme are scarce.

Clostridium alkalicellulosi is a mesophilic bacterium collected from alkaline environments. This microorganism demonstrates high cellulose degrading activity on recalcitrant polysaccharides and algal biomass [18,19]. This efficient catalysis is believed to be linked to the cellulosome system, which is known to be an efficient enzyme complex for cellulose and polysaccharide degradation [18]. Nevertheless, a detailed investigation of cellulosomal elements in *C. alkalicellulosi* has not been performed. Recently, draft genome sequencing of *C. alkalicellulosi* was achieved, revealing many genes related to structural components essential for cellulosome assembly [20,21]. In addition, various cellulosomal GH enzymes, specifically dockerin-borne enzymes, exist in the genome. Herein, one putative cellulosomal GH9 enzyme, named *CalkGH9T*, from the *C. alkalicellulosi* genome was cloned, expressed, and biochemically characterized. This GH9 enzyme has a modular structure and demonstrated unusual enzymatic activity on cellulosic substrates, differing from previously characterized GH9 endoglucanases.

2. Results and Discussion

2.1. Bioinformatics Analysis

The genome of *C. alkalicellulosi* was sequenced, and analysis showed that it contained one gene, CloalDRAFT_2759, which encodes a protein belonging to glycoside hydrolase family 9. The gene (2841 bp) was computationally translated to a protein sequence of 947 amino acids with a molecular weight of 107.5 kDa and pI = 4.52 (<https://web.expasy.org/cgi-bin/protparam/protparam/>, accessed on 13 June 2019) (Figure S1). Accordingly, we designated this deduced protein *CalkGH9T*, and this name is used hereafter. SignalP 5.0 server analysis (<http://www.cbs.dtu.dk/services/SignalP/>, accessed on 13 June 2019) showed the presence of an N-terminal signal peptide in the

polypeptide with a cleavage site between Ala34 and Asp35, indicating that *CalkGH9T* is secreted extracellularly. BlastP server analysis (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>, accessed on 13 June 2019) revealed the modular structural organization of *CalkGH9T*, which contains, from the N-terminus, one GH9 catalytic module, two family 3 CBMs (CBM3s), and one type I dockerin domain (Figure 1A). Similarly, this modular architecture is present in the genes from *C. thermocellum*, *Acetivibrio cellulolyticus*, and *C. clariflavum* [16,17].

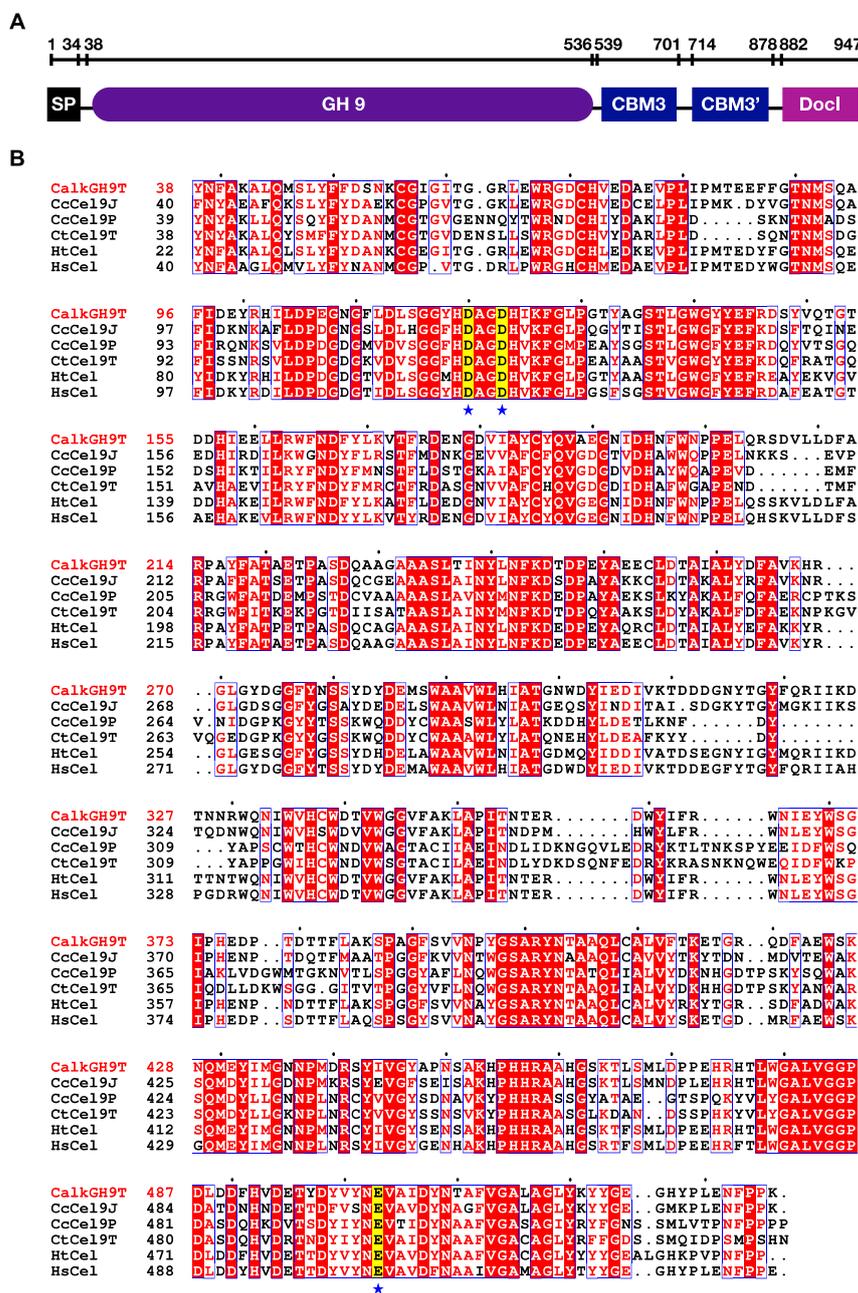


Figure 1. (A) Modular architecture of *CalkGH9T*, showing a full-length sequence (947 amino acids) starting from an N-terminal signal peptide followed by one GH9 catalytic module, two family 3 CBMs, and one dockerin. (B) Multiple sequence alignment of catalytic module of *CalkGH9T* and four selected GH9 enzymes. Identical residues are shown in white with a red background, whereas amino acid replacements (conservative changes) are shown in red with a white background. The three conserved catalytic residues are highlighted in yellow and marked with stars. The sequences were aligned by use of Clustal Omega [22] and the sequence alignment was displayed by using ESPrpt [23].

The sequence of the *CalkGH9T* catalytic module was blasted against protein sequences available in the NCBI database. The *CalkGH9T* catalytic module shared some similarity with endo-1,4- β -glucanases of *Hungateiclostridium* (formerly *Clostridium*) *thermocellum* (CAK22317.1; 81.64%), *H. saccincola* (WP_101299764.1; 82.36%), and *H. straminisloven* (WP_054847055.1; 80.64%). Unfortunately, these most closely related putative enzymes have not been biochemically characterized. To identify amino acids involved in catalysis, we selected the two closest uncharacterized sequences and three known catalytic module sequences of Cel9J and Cel9P from *C. cellulolyticum* (CcCel9J and CcCel9P) [6] and Cel9T from *C. thermocellum* (CtCel9T) [24], which shared some similarity (68%, 44%, and 46%, respectively) with *CalkGH9T*. Multiple sequence alignment of these five GH9 sequences was performed by using Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>, accessed on 13 June 2019), which revealed highly conserved active site regions (Figure 1B). The important catalytic amino acids were predicted to be Asp120, Asp123, and Glu504 [15]. Glu serves as a catalytic general acid, and the two Asp residues potentially bind water molecules and serve as catalytic general bases that carry out the nucleophilic attack on the C1 carbon of the target glucose unit, resulting in bond cleavage through an inversion-type reaction [15].

CalkGH9T contains two CBM3s, designated *CalkCBM3* and *CalkCBM3'*. *CalkCBM3* is fused to the C-terminus of GH9 catalytic module, followed by *CalkCBM3'*. The *CalkCBM3* sequence is 29.3% and 31.5% similar to the characterized CBM3s of endoglucanases Cell and CelQ from *C. thermocellum*, and the *CalkCBM3'* sequence shares 40% and 53% similarity with the known CBM3s of the *C. thermocellum* Cell and Cel9V, respectively. The Cell and CelQ CBM3s adjacent to the GH9 catalytic modules belong to the type C CBM3 (CBM3c) [15,25], whereas the additional CBM3s of the *C. thermocellum* Cell and Cel9V belongs to type B and B' CBM3 (CBM3b and b'), respectively [8,14,16]. Unlike CBM3a or b, CBM3c binds weakly to cellulose because of the absence of conserved aromatic residues that are involved in strong binding to cellulose. Nevertheless, some CBM3cs participate in catalytic function by facilitating cellulose decrystallization and supply the cellulose chain to the active site of the catalytic module [25,26]. Based on sequence similarity, *CalkCBM3* and *CalkCBM3'* possibly show similar functions as the previously characterized CBM3c and CBM3b. However, the presence of two CBM3s in a single polypeptide suggests that individual CBM3s may have more selective roles during cellulose hydrolysis. Therefore, the functions of *CalkCBM3* and *CalkCBM3'* remain to be studied.

A search for similarity of the *CalkGH9T* dockerin domain against available dockerin sequences revealed that the *CalkGH9T* dockerin is a type I dockerin, and it shares 62% similarity to the type I dockerin of CtCel5c of *C. thermocellum* [27]. Pairwise alignment showed that, like CtCel5c, *CalkGH9T* dockerin has conserved identical amino acids, Ser and Thr at positions 10 and 11, and similar amino acids, Arg and Arg at positions 17 and 18 (Lys and Arg are at the same positions in CtCel5c dockerin). However, the amino acid at position 22 of *CalkGH9T* dockerin is Glu, whereas that of CtCel5c dockerin is Arg (Figure 2). The amino acids at these positions (10, 11, 17, 18 and 22) serve for binding recognition residues to interact with cohesin modules. Dockerin sequences are highly conserved within cellulosome producing microorganisms, and some dockerin sequences are very species-specific [20]. The function of type I dockerin is to dock the enzymatic subunit to the scaffolding protein to form a cellulosome complex. Our results indicate that dockerin-borne *CalkGH9T* is a cellulosomal enzymatic subunit of the *C. alkalicellulosi* cellulosome, and it possibly shows cross-reactivity with *C. thermocellum* cellulosome based on sequence similarity [21].



Figure 2. Pairwise alignment of *CalkGH9T* and *CtCel5c* dockerin modules. The sequences were aligned by use of EMBOSS Needle. The amino acid residues involved in Ca²⁺-binding loops are highlighted in cyan, and those involved in binding recognition with cohesin are highlighted in magenta.

2.2. Protein Expression and Purification

The gene encoding *CalkGH9T*, without a signal peptide, was subcloned and expressed in competent *E. coli* BL21 (DE3) cells at 16 °C. The protein was expressed in a soluble form and purified with an Ni²⁺-NTA column, as it contained N-terminal His-tag. The size and purity of the purified *CalkGH9T* was determined by SDS-PAGE, which showed a single band with an estimated size of 107 kDa (Figure S2). The apparent molecular weight is consistent with the size deduced from the translated gene, and the full-length sequence of *CalkGH9T* was used for biochemical study.

2.3. Substrate Preference and Effects of pH and Temperature on Enzyme Activity

The enzymatic activity of *CalkGH9T* on polysaccharides was investigated (Table 1). Here, we used CMC, an ionic carboxymethyl cellulose with degree of substitution of about 0.4, as a soluble substrate for detecting endoglucanase activity. RAC and Avicel, the insoluble substrates that contain amorphous fractions and crystallinity (70%) [28], were used for detecting cellulase and exo-glucanase activities, respectively [29]. Xylan, a partially acetylated β-(1,4)-linked xylose monosaccharides, contain both soluble and insoluble forms. This substrate was used for detecting endo-xylanase activity, as this substrate shares similar β-1,4-glycosidic bonds in its main chain to cellulose. The results indicated that *CalkGH9T* had the highest activity towards CMC (Table 1). The activity on RAC was low, and the activity on Avicel was undetectable. *CalkGH9T* showed side-activity on xylan. This result suggests that *CalkGH9T* seems to hydrolyze β-1,4 linked glucan and xylan substrates. However, insoluble and crystalline forms of the substrates likely restrict the enzymatic activity. Based on this evidence, CMC was used as a substrate for further characterization.

Table 1. Substrate preference of *CalkGH9T*.

Substrate	Main Linkage	Solubility in Water	Activity (μM Product/min/μM Protein)	Relative Activity (%)
CMC	β-1,4 glucan	Soluble	1021	100
RAC	β-1,4 glucan	Insoluble	142 ¹	14
Avicel	β-1,4 glucan	Insoluble	ND ²	ND
Xylan	β-1,4 xylan	In/soluble	261	26

¹ Enzymatic assay for RAC was incubated for 3 h with enzyme load of 1 μM. ² ND, no detectable activity at the assay conditions for 24 h with enzyme load of 1 μM.

Analysis of the effects of pH and temperature on enzymatic activity was performed using CMC as a substrate. *CalkGH9T* was active over a pH range of 5.0–11.0 at 55 °C (Figure 3A). *CalkGH9T* showed more than 80% of its maximal activity at pH 7.0–8.0, with the highest activity (100%) at pH 7.4. The activity sharply decreased when the pH increased from 8.0 to 11.0. However, the enzyme remained active at pH 10.0, with 50% residual activity. To test the stability of *CalkGH9T* in various pH environments, the enzyme was

incubated in different buffers (pH 3.0–11.0) at 30 °C for 1 h prior to incubation with CMC (Figure 3B). *CalkGH9T* showed good stability in the pH range of 7.0–9.0, and the activity gradually decreased outside this pH range. The pH profile and stability results indicate that *CalkGH9T* functions better in near neutral or slightly alkaline environments.

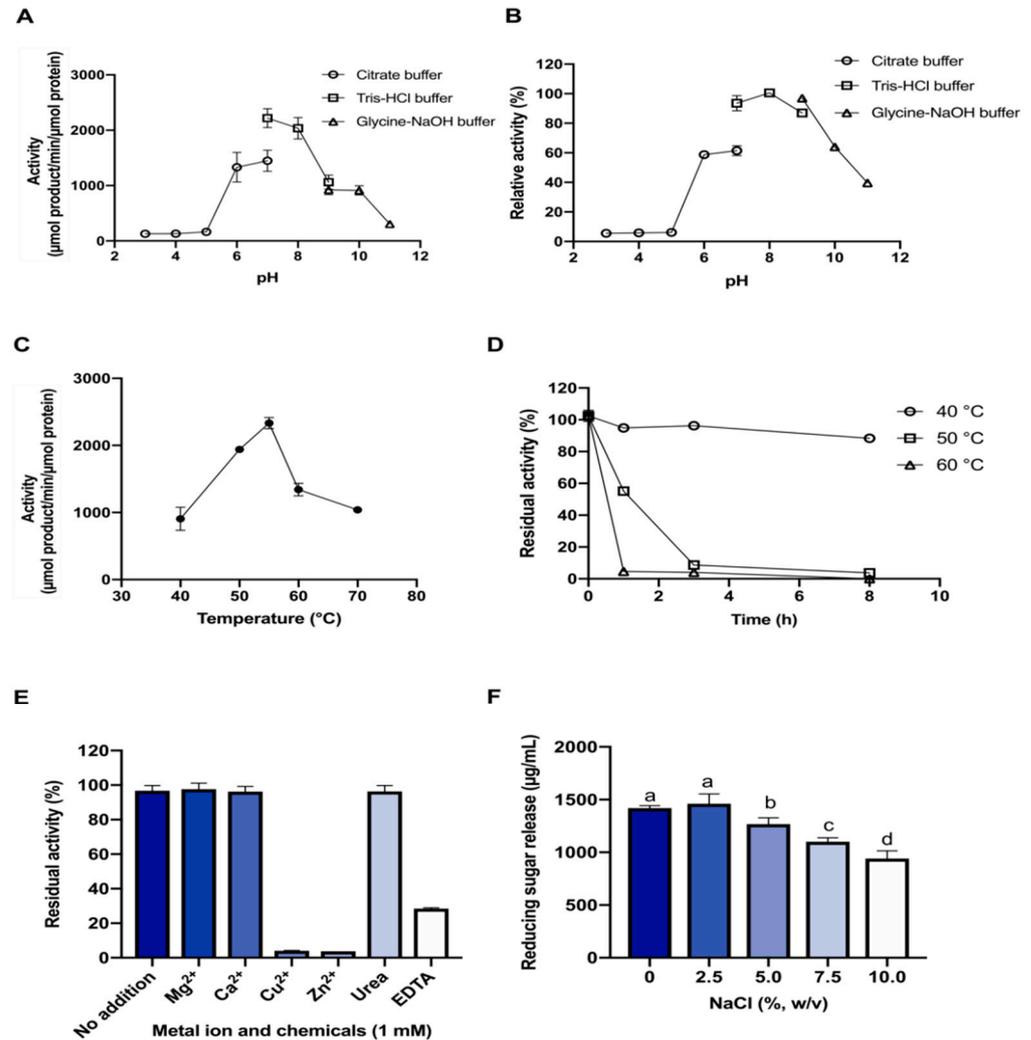


Figure 3. Biochemical properties of *CalkGH9T*. (A) Profiles of pH on activity and (B) pH stability assayed at 55 °C. (C) Influence of temperature on and (D) thermostability analysis of the activity of *CalkGH9T* assayed at pH 7.4. For pH and temperature assays, the maximum activity was considered 100%; for pH and temperature stability assays, the enzyme activity without any treatments was considered 100%. (E) Effects of metal ions and (F) NaCl on *CalkGH9T* activity. The enzymatic reaction was tested with 1% (*w/v*) CMC. Values are means and bars represent the standard deviations for three independent experiments. For NaCl study, different letters within bars indicate significant differences ($p < 0.05$) between the means of glucose concentrations as analyzed by Tukey's test of multiple comparisons after ANOVA test.

CalkGH9T was active over a temperature range of 30–70 °C at pH 7.4 (Figure 3C). The activity increased as a function of temperature, and the maximal activity (100%) occurred at 55 °C. The activity was lower with temperatures over 60 °C, but the enzyme retained approximately 40% residual activity at 70 °C. At 80 °C, *CalkGH9T* had completely lost its activity. The thermal stability was investigated because it is a key factor for enzymes used in industrial applications. *CalkGH9T* was incubated at three temperatures: 40, 50, and 60 °C (Figure 3D). *CalkGH9T* was highly stable at 40 °C; prolonged incubation up to

two days did not significantly diminish the activity (data not shown). The half-lives of *CalkGH9T* at 50 and 60 °C were 1.5 h and 42 min, respectively. Our results indicate that *CalkGH9T* functions best at a moderate temperature.

2.4. Effect of Metal Ions and Chemicals

The influence of metal ions (1 mM) on the activity of *CalkGH9T* against CMC was tested (Figure 3E). *CalkGH9T* activity was strongly inhibited by Cu^{2+} and Zn^{2+} . Mg^{2+} and Ca^{2+} did not show any inhibitory effects on *CalkGH9T* activity. Urea did not decrease the activity of *CalkGH9T*, but EDTA remarkably decreased the activity, resulting in 25% residual activity. It has been reported that Ca^{2+} is necessary for GH9 enzymes because it is involved in conformational stability and catalysis [9,30]. In addition, *CalkGH9T* contains one dockerin and two CBMs that require Ca^{2+} to stabilize their structures [8,20]. The addition of EDTA may chelate with metal ions, thus decreasing overall structural stability and activity of the enzyme.

2.5. Effect of NaCl Concentration

The activity of *CalkGH9T* on CMC in the presence of NaCl was investigated because *C. alkalicellulosa* was collected from a soda lake, which has high salinity (up to 2.4%) [19] (Figure 3F). The activity of *CalkGH9T* without added NaCl was taken as 100%. *CalkGH9T* showed great resistance to NaCl. NaCl concentration in a range of 1–5% (*w/v*) did not significantly affect the enzyme activity. At 7.5% and 10% (*w/v*) NaCl, *CalkGH9T* retained more than 80% activity. This result indicates that, unlike common cellulases, high salinity does not strongly inhibit the catalytic efficiency of *CalkGH9T* [31,32]. This suggests that *CalkGH9T* is halotolerant. The high-salt resistance may be in part explained by the fact that the catalytic domain of *CalkGH9T* is rich in acidic amino acids (Figure S1). The distribution of the excess charges on the protein surfaces might prevent aggregation and/or precipitation induced by a high concentration of salts [33].

2.6. Mode of Action

2.6.1. Activity on Polysaccharides

To characterize hydrolysis by *CalkGH9T*, the profile of CMC hydrolysis was studied (Figure 4A). The release of reducing sugar increased with hydrolysis time. The rate of reducing sugar release reached a plateau after 3 h. After that, reducing sugar release was steady. The constant release of reducing sugar may be in part explained by end-product inhibition. In addition, the activity of the enzymes was abolished when the temperature was at 50 °C or higher after 3-h incubation (Figure 3D). Therefore, the constant release of reducing sugar by *CalkGH9T* at 55 °C after 3 h could be due to the loss of enzyme activity and stability.

The reducing sugars released from CMC by *CalkGH9T* were analyzed by TLC (Figure 4B). The initial hydrolysis of CMC was considered to occur within 5 min. At 1 min, cellotetraose appeared as the initial hydrolysis product. From 1 to 5 min, cellodextrins with various degrees of polymerization arose. The series of the hydrolysis oligomers produced indicates an endo-acting mode for *CalkGH9T*. However, at different time points from 1 min to 16 h, cellotetraose (G4) existed as the major hydrolysis product throughout the course of hydrolysis, and other oligosaccharides, including cellopentaose (G5), cellotriose (G3), cellobiose (G2), and glucose (G1), appeared as minor hydrolysis products. Larger cellodextrins (>G5) were not observed. The high accumulation of G4 during CMC hydrolysis suggests that *CalkGH9T* is a cellotetraose producer with an endo-acting mode of action. On the other hand, a small amount of glucose was observed after 3 h of hydrolysis. The slight release of glucose is possibly a result of partial hydrolysis of some cellodextrins, such as cellopentaose (see results in Section 2.6.2).

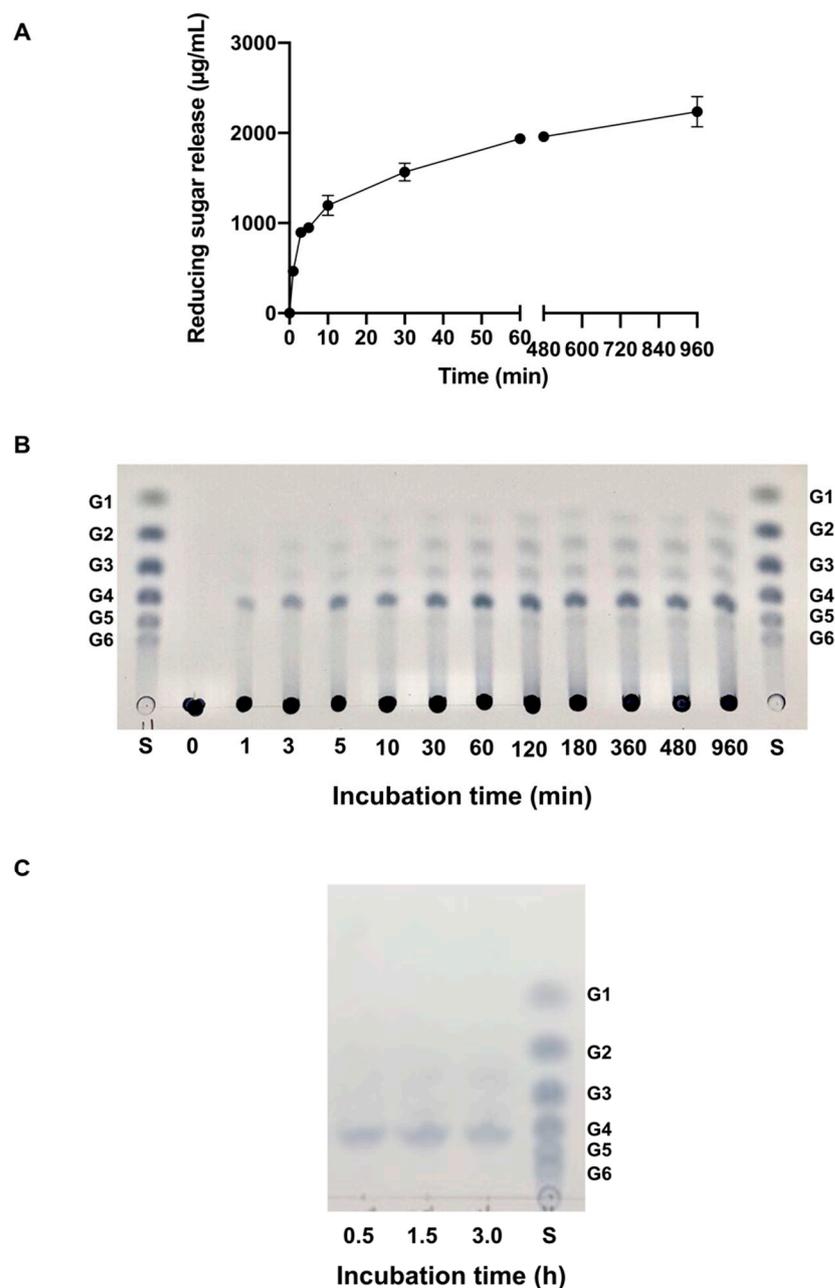


Figure 4. (A) Time course of reducing sugar release by 0.25 μM *CalkGH9T* using CMC as the substrate. The hydrolysis reactions between 60 to 480 min were analyzed; however, the X-axis break was used to enhance the readability for the chart. (B) TLC analysis of the hydrolysis products using CMC as the substrate. (C) TLC analysis of soluble sugars derived from RAC hydrolysis by 1.0 μM *CalkGH9T*. The enzymatic reactions of CMC and RAC were performed at pH 7.4 and 55 $^{\circ}\text{C}$. The samples of soluble sugars from RAC hydrolysis were concentrated by evaporation prior to being spotted on TLC. S, standards of glucose and cellodextrins with a degree of polymerization of 2–6 (G2–G6).

Unlike CMC, hydrolysis of RAC by *CalkGH9T* at 0.25 μM enzyme load from 1 to 16 h-incubation did not result in any detectable sugar release. To confirm this, the enzyme load was increased to 1.0 μM (a 4-fold increase), and the reaction was incubated for 0.5, 1.5, and 3 h. Nevertheless, the amount of sugar release from RAC was very small (0.1–0.2 mg/mL), and there was no significant difference in sugar concentrations at the three incubation times tested (Figure S3). The low activity on RAC indicates that insoluble amorphous cellulose is not a preferable substrate for *CalkGH9T*. To determine the reducing

sugar produced, the supernatant was concentrated by evaporation and analyzed via TLC. Like CMC hydrolysis, G4 appeared as the main product (Figure 4C).

2.6.2. Activity on Cellodextrins

To understand more about cleavage patterns of *CalkGH9T*, the hydrolysis of cellodextrins (G6–G2) was studied (Figure 5). Within 5 min of hydrolysis, *CalkGH9T* completely hydrolyzed G6 to G4 and G2 and partially hydrolyzed G5 to G4 and G1. However, G4, G3, and G2 were not hydrolyzed (Figure 5). This result indicates the preferential cleavage of long chain oligomers. With prolonged incubation time (16 h), the hydrolysis of G5 was incomplete, leaving G5, G4, and G1. Nevertheless, no hydrolysis of G4, G3, and G2 was observed). This result indicates that G4, G3, and G2 are potential end-products of *CalkGH9T* (Figure S4). The active sites might contain at least six binding subsites.

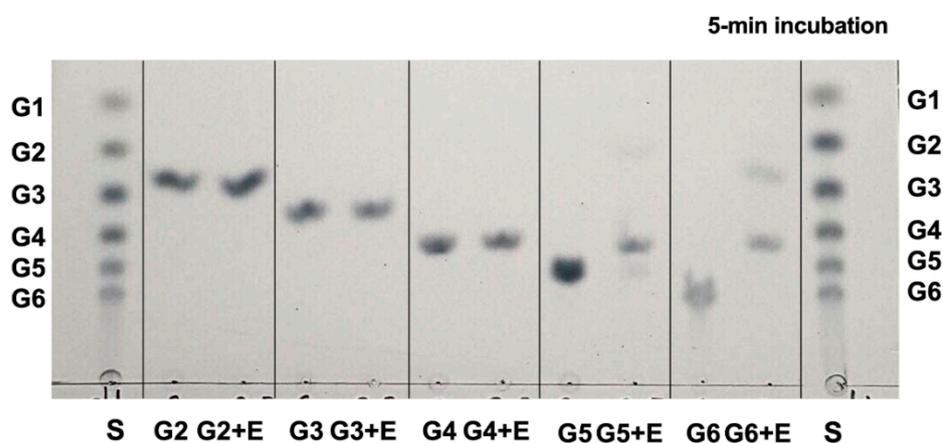


Figure 5. TLC analysis of hydrolysis products generated by 0.25 μM *CalkGH9T* using cellodextrins with a degree of polymerization of 2–6 (G2–G6) as the substrates for 5-min incubation. S, standards of glucose and cellodextrins with a degree of polymerization of 2–6 (G2–G6). E, *CalkGH9T*. G2, G3, G4, G5, and G6 represent a reaction control (a reaction mixture containing a substrate and a denatured enzyme). G2 + E, G3 + E, G4 + E, G5 + E, and G6 + E represent enzymatic reactions. The enzymatic reactions were performed at pH 7.4 and 55 $^{\circ}\text{C}$.

2.7. Cohesin-Dockerin Interaction

The unique feature of cellulosome assembly is the interaction of enzyme borne dockerins and cohesin containing scaffolds [20]. The primary scaffold, named ScaA, of *C. alkalicellusi* cellulosome contains 10 type I cohesin modules (Figure 6A) [21], and it was selected because it was a core protein that integrates the key cellulosomal enzymes [21]. To prove the type I dockerin of *CalkGH9T* is functionally active, the dockerin module and four cohesin modules of ScaA were expressed and tested for binding interaction. These four cohesin modules were selected according to their difference in sequence similarity [21]. Based on the affinity-based ELISA assay, the type I dockerin of *CalkGH9T* bound significantly to four selected cohesin modules of ScaA (Figure 6B), indicating that *CalkGH9T* is an enzymatic subunit of the *C. alkalicellusi* cellulosome. The *CalkGH9T* dockerin exhibited a relatively higher affinity to the cohesin module I (ScaA-1) than ScaA-4, 5, and 10. This difference may reflect the difference in cohesin-dockerin recognition. A similar phenomenon was observed in *Clostridium clariflavum* [34] and *Acetivibrio cellulolyticus* [35], in which the dockerins of the cellulosomal enzyme subunits bind to different cohesin modules of the primary scaffold with different binding intensities.

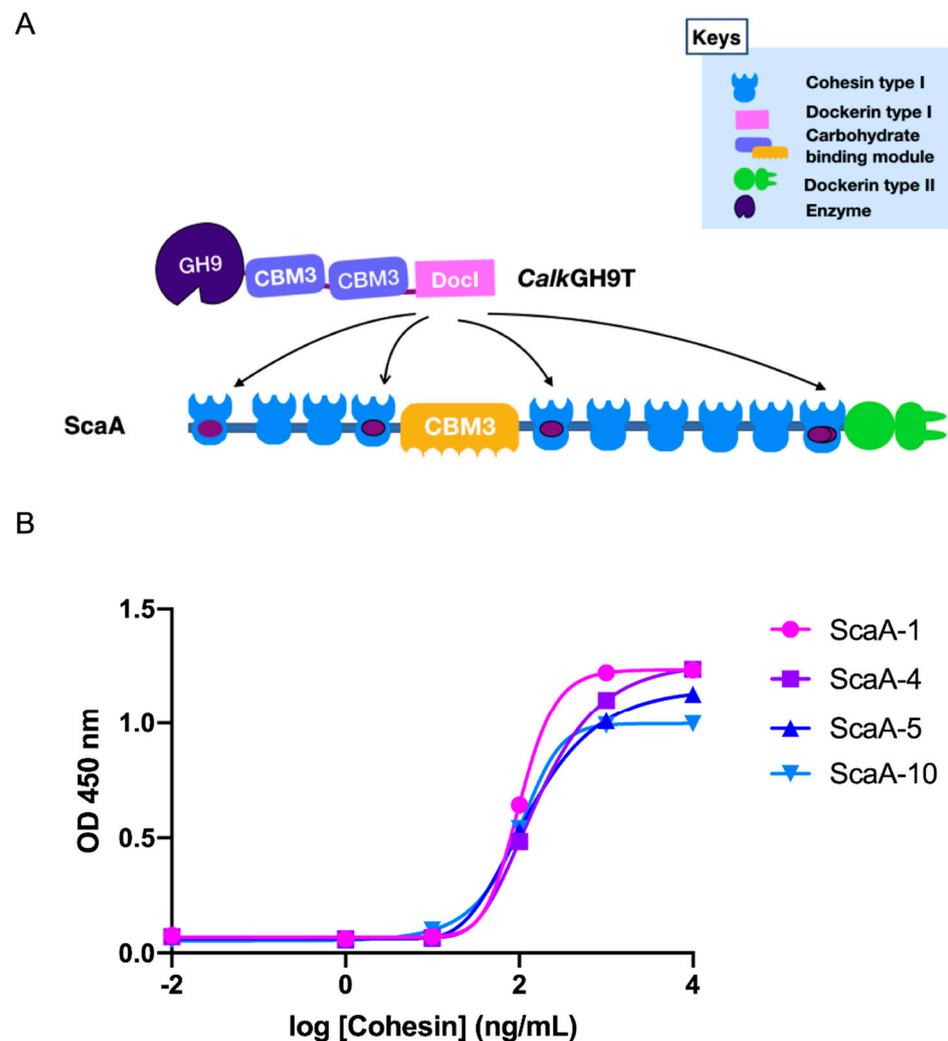


Figure 6. Binding interaction of the type I dockerin of *CalkGH9T* and the cohesins modules of the primary scaffold ScaA of *C. alkallicellulosi* cellulosome. **(A)** Schematic representation of *CalkGH9T* showing a modular architecture and ScaA containing 10 cohesin modules. Dots indicate the cohesin modules that were expressed for affinity binding test. **(B)** Affinity-based ELISA profiles the dockerin towards different ScaA cohesin modules. ScaA-1, 4, 5, and 10 refer to of cohesin modules 1, 4, 5, and 10 of ScaA, and the designated numbers of the cohesin modules begin from the cohesin at the N-terminus.

2.8. Binding Ability to Polysaccharides

We performed an in vitro binding ability test of *CalkGH9T* with amorphous insoluble cellulose RAC, crystalline cellulose Avicel, and xylan because *CalkGH9T* contains family 3 CBMs that are known to potentiate the activity of cognate catalytic modules against insoluble substrates (Figure 7A). Here, *CalkGH9T* was allowed to bind with RAC, Avicel, and xylan at 4 °C for 1 h. After binding, the samples were centrifuged to separate unbound (supernatant) and bound (associated with the substrate) fractions. Individual fractions were analyzed on SDS-PAGE. *CalkGH9T* appeared in the bound fractions of RAC (52%) and Avicel (nearly 100%), indicating the ability to bind both insoluble amorphous and crystalline cellulose. However, a higher intensity protein band occurred in the Avicel-bound fraction, reflecting the ligand recognition of the CBMs, which prefer to bind the surfaces of crystalline cellulose. This strong binding ability is linked to the presence of *CalkCBM3(s)*. In contrast, *CalkGH9T* was unable to bind xylan. It should be noted that although *CalkGH9T* contains CBM3c (*CalkCBM3*), which has been reported to facilitate

cellulose disruption and continuously feed the cellulose chain to the catalytic module, the processivity of the enzyme cannot be predicted by the presence of the CBMs.

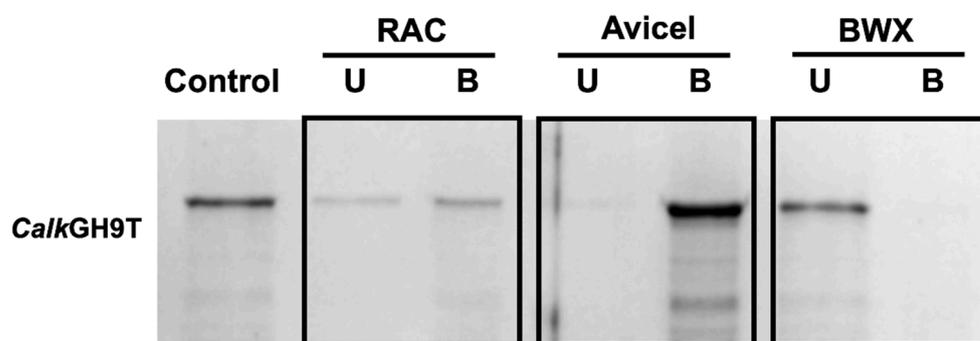


Figure 7. Binding ability of *CalkGH9T* to RAC, Avicel, and BWX. U and B refer to unbound and bound proteins, respectively.

2.9. Processive Activity

In the present study, although *CalkGH9T* bound tightly to Avicel, it lacked the ability to hydrolyze this microcrystalline cellulose. Therefore, RAC was used for the processivity assays because it is a suitable substrate to distinguish processive from non-processive activity [36].

Processive activity is when the enzyme remains bound to the substrate and continuously performs subsequent catalysis before dissociation. In addition, processive endoglucanase enzymes demonstrate both endoglucanase and exoglucanase activities. The exoglucanase activity releases soluble reducing sugars from insoluble cellulose into the solution; in contrast, the endoglucanase activity randomly cleaves the cellulose chain and leaves several reducing ends attached on the insoluble substrate. Therefore, the ratio of soluble and insoluble reducing ends is commonly used to differentiate exoglucanases from endoglucanases [13,37]. Here, we determined the distributions of reducing ends between soluble (released sugars) and insoluble (pellet) fractions of hydrolysis of RAC by *CalkGH9T* (Figure 8). The ratios of soluble to insoluble reducing ends did not increase as hydrolysis time was prolonged from 0.5 to 3.0 h. This proportion (approximately 0.8) is similar to that found for GH9 endoglucanases from *C. cellulolyticum*, with proportions ranging from 0.5 to 1.2 [6]. In addition, the percentage of reducing ends (40%) produced from the insoluble RAC is common for endoglucanase activity and it is consistent with other true endoglucanases that produce approximately 30 to 50% reducing ends from insoluble substrates [37,38]. This result indicates that *CalkGH9T* is an endoglucanase with non-processive activity.

2.10. Addition of *CalkGH9T* to a Fungal Enzyme Preparation

Based on substrate specificity and hydrolysis product generation, *CalkGH9T* appears to have different mode of action compared to common fungal endoglucanases [39] and it contains the CBM3c, which is reported to disrupt cellulose [25,26]. To investigate whether *CalkGH9T* could benefit a fungal enzyme preparation, hydrolysis of RAC by a combination of *CalkGH9T* and a commercial fungal enzyme Ctec2 was tested (Figure 9). Here, the enzymatic reaction was performed at pH 7.4 and 55 °C with a substrate loading of 1% (*w/v*) RAC and a total protein load of 100 µg. The hydrolysis of RAC by Ctec2 alone (100 µg) released reducing sugars after 15, 30, and 60 min of incubation. The supplementation of Ctec2 (90 µg) with *CalkGH9T* (10 µg) resulted in increased release of reducing sugars at all incubation times compared to Ctec2 alone. The increase in reducing sugar release indicates that *CalkGH9T* can work in concert with Ctec2. This synergistic interaction can be explained by at least two mechanisms. First, *CalkGH9T* may bind to the insoluble RAC, and loosen and cleave cellulose chains, thereby creating new free chain ends for other enzymes, i.e., exoglucanases or cellobiohydrolases, present in Ctec2 to attack. Second, despite the

limited activity on RAC, *CalkGH9T* may hydrolyze soluble cellodextrins produced by *Ctec2* to short cello-oligomers, thus increasing the concentration of reducing ends.

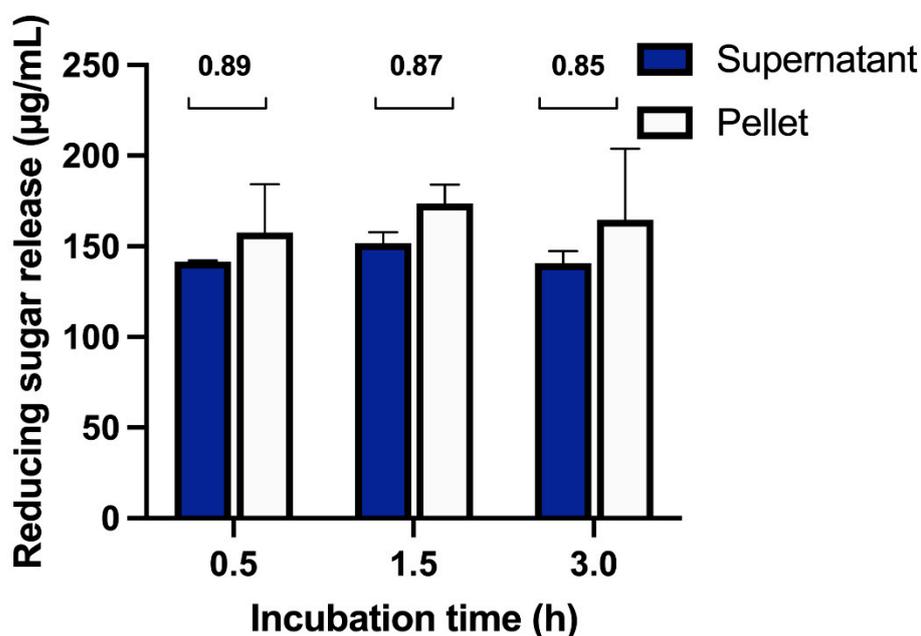


Figure 8. Processive activity test on RAC. Values are means and bars represent the standard deviations for three independent experiments.

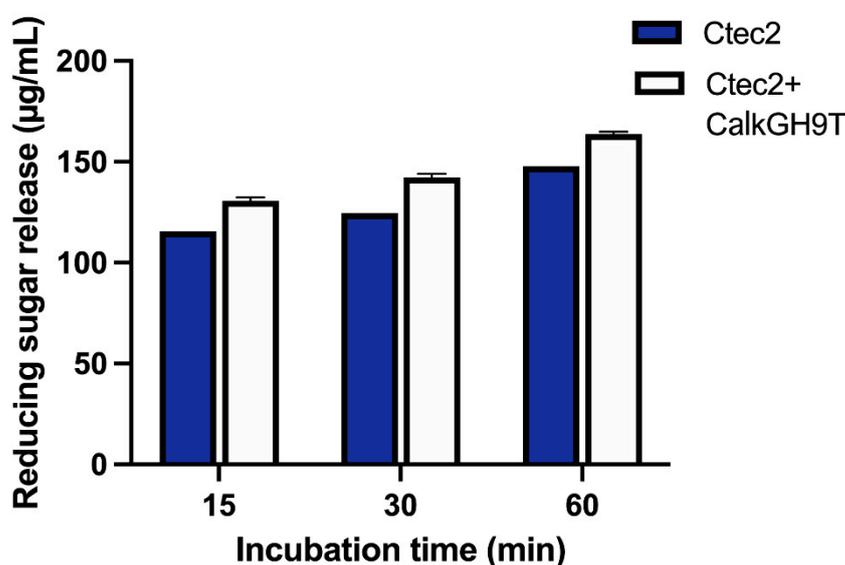


Figure 9. The effect of *CalkGH9T* addition to a commercial cellulase preparation *Ctec2* on RAC hydrolysis. The enzymatic reaction was performed at pH 7.4 and 55 °C with a substrate loading of 1% (*w/v*) RAC and a total protein load of 100 µg. The reaction with *Ctec2* alone (100 µg protein load) was used as a control. For supplementation of *Ctec2* with *CalkGH9T*, *Ctec2* was replaced with *CalkGH9T* by 10%, corresponding to a load of 90 µg *Ctec2* and 10 µg *CalkGH9T*, respectively. Values are means and bars represent the standard deviations for three independent experiments.

3. Materials and Methods

3.1. DNA, Bacterial Strains, Plasmid, and Chemicals

The bacterial genomic DNA of *C. alkalicellulosi* DSM17461^T was prepared by Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Germany.

E. coli competent cell strains NEB5 α and BL21 (DE3), used for cloning and protein expression, respectively, were purchased from New England BioLabs, Inc., (Ipswich, MA, USA). The expression vector pET28a(+) used for both cloning and protein expression and T4 DNA ligase were obtained from Novagen, (Madison, WI, USA). Restriction enzymes (i.e., *Nhe*I and *Xho*I) were obtained from ThermoFisher Scientific, (Waltham, MA, USA). 4M carboxymethylcellulose (CMC), Azo-CMC, beechwood xylan, standard glucose and cellodextrins (G2–G6), and standard xylose and xylo-oligomers (X2–X6) were purchased from Megazyme, Ireland, and Avicel PH105 was purchased from Sigma-Aldrich Chemicals Co., USA. Thin layer chromatography plates (TLC silica gel 60 F₂₅₄, 20 × 20) were procured from Merck, (Kenilworth, NJ, USA). Regenerated amorphous cellulose (RAC) was prepared as described by Zhang, et al. [40].

3.2. Gene Cloning and Recombinant DNA Techniques

The gene CloalDRAFT_2759 present in the *C. alkalicellulosi* genome was predicted to encode a *CalkGH9T* protein. This gene without a signal peptide was amplified by PCR using specific primers. The primers were forward 5'-AATTAGCTAGCGATCCAGAGTACAACCTTGC-3' and reverse 5'-AATTACTCGAGTTAACGCGGCAGTTGTGGAA-3', in which the restriction sites for *Nhe*I and *Xho*I are underlined, respectively. The PCR program was initial denaturation at 98 °C for 30 s, followed by 40 cycles of denaturation at 98 °C for 10 s, annealing at 60 °C for 40 s, elongation at 72 °C for 90 s, and final extension at 72 °C for 5 min [36]. The PCR product was checked by gel electrophoresis using 0.8% (*w/v*) agarose. The amplified product with the target size was cut from the gel and purified by using a Qiagen gel extraction kit (Qiagen, New York, NY, USA). The purified product and plasmid pET28a(+) were double digested with *Nhe*I and *Xho*I restriction enzymes and ligated with T4 DNA ligase (BioRad, Hercules, CA, USA). After ligation, the recombinant plasmid was transformed into *E. coli* NEB5 α competent cells. The transformed cells were grown on LB agar plate supplemented with kanamycin (50 μ g/mL) at 37 °C overnight. The positive recombinant clones were selected, and their recombinant DNA sequences were confirmed by means of colony PCR and DNA sequencing.

3.3. Protein Expression and Purification

The recombinant plasmid pET28a(+) containing a gene encoding *CalkGH9T* was transformed into *E. coli* BL21 (DE3) cells. The transformed cells were grown at 37 °C in LB broth supplemented with kanamycin (50 μ g/mL) and 2 mM CaCl₂ to reach mid-exponential growth phase (optical density at 600 nm (OD₆₀₀)~0.7–1.0). After that, the growth culture was supplemented with isopropyl-1-thio- β -D-galactopyranoside (IPTG) at a final concentration of 0.2 mM to induce protein synthesis and further incubated at 16 °C overnight. The cells were harvested by centrifugation at 4200 × *g* at 4 °C for 15 min, and the cell pellets were then resuspended in 20 mL Tris buffer saline (TBS, 25 mM Tris-HCl, 137 mM NaCl, 2.7 mM KCl, pH 7.4) containing 5 mM imidazole and 300 mM NaCl. The cell suspension was sonicated and centrifuged at 22,000 × *g* at 4 °C for 30 min to separate cell debris.

The crude protein supernatant was mixed with Ni-NTA resin (Qiagen, New York, NY, USA), incubated at 4 °C for 1 h with gentle rotation, and loaded into a 20-mL Econo-Pack column (BioRad, Hercules, CA, USA). By gravitational flow, the column was washed with 100 mL washing buffer (TBS containing 30 mM imidazole and 300 mM NaCl). After washing, the protein was eluted with 16 mL elution buffer (TBS containing 250 mM imidazole and 300 mM NaCl). The eluted fraction was dialyzed against TBS at 4 °C overnight, and the dialyzed protein was determined for its purity on 10% SDS-PAGE stained with InstantBlue coomassie stain (Expedeon, San Diego, CA, USA). The protein concentrations were determined by the absorbance at 280 nm. The extinction coefficient of the protein was determined by use of the ProtParam tool on the EXPASY server (<https://web.expasy.org/protparam/>, accessed on 13 June 2019).

3.4. Enzyme Activity Assay

The enzyme assay was performed in a total reaction volume of 200 μL . The reaction mixture consisted of 1% (*w/v*) substrate (i.e., CMC, RAC, BWX, and Avicel) and 0.25 μM of *CalkGH9T* in 50 mM TBS pH 7.4. The reaction mixture was incubated at 50 $^{\circ}\text{C}$ for 15 min (or 3–24 h for RAC and Avicel, respectively). The reaction was terminated by immersion in ice water and centrifuged at $13,000\times g$ at 4 $^{\circ}\text{C}$ for 10 min to separate the substrate from the soluble fraction. One hundred microliters of the soluble fraction was taken to a new 1.5-mL microcentrifuge tube and 150 μL of dinitrosalicylic acid (DNS) [41] was added. The reaction mixture was boiled for 10 min and centrifuged at high speed. A sample was taken and measured at an absorbance of 540 nm. The amount of reducing sugar released in the mixture was estimated using a glucose standard curve. One unit of enzyme was defined as the amount of enzyme that produces 1 μMol of glucose equivalent in 1 min under the assayed conditions.

3.5. Effect of pH and Temperature

The effect of pH on enzyme activity was analyzed by determining the activity at different pH values from pH 3.0–11.0 at 55 $^{\circ}\text{C}$ for 10 min. Buffers (50 mM) with different pH ranges were citrate-sodium citrate buffer for pH 3.0–7.0, Tris-HCl buffer for pH 7.0–9.0, and glycine-NaOH for pH 9.0–11.0. The maximum activity at the corresponding pH was taken as 100%. The pH stability assay was performed by incubating 5.5 μL of the 11.5 μM enzyme in 10 mM of the buffers mentioned above at 30 $^{\circ}\text{C}$ for 30 min. After that, the enzyme solution was diluted, and the residual activity was measured under optimal assay conditions.

The effect of temperature on enzyme activity was tested by incubating 0.25 μM enzyme in a 200- μL reaction mixture containing 1% (*w/v*) substrate at a temperature range of 40–70 $^{\circ}\text{C}$, pH 7.4 (TBS) for 10 min. The maximum activity at the corresponding temperature was taken as 100%. The thermal stability of the enzyme was determined by incubating 5.5 μL of the 11.5 μM enzyme at the temperatures of 40, 50, and 60 $^{\circ}\text{C}$ for 0–24 h. The enzyme solution was diluted, and the residual activity was determined under optimal assay conditions.

3.6. Effect of Metal Ions, Salts, and Chemicals

The effect of different metal ions (in the form of sulfate salts), including Ca^{2+} , Mg^{2+} , Cu^{2+} , Zn^{2+} , and chemicals, including chelating agent EDTA and urea, on enzyme activity were explored. Metal ions, EDTA, and urea were added to a final concentration of 1 mM to a 200- μL reaction mixture containing 0.25 μM enzyme and 1% (*w/v*) substrate. The reaction was carried out at 55 $^{\circ}\text{C}$, pH 7.4 (TBS) for 10 min, and the residual activity was determined by measuring the reducing sugar released in the reaction mixture as described above. Substrate blanks with added metal ions, EDTA, and urea were run in parallel and assayed. The enzyme activity measured in the absence of metal ions and chemicals was taken as 100%.

The effect of NaCl was investigated by adding different concentrations of NaCl (0–10%, *w/v*) to a 200- μL reaction mixture containing 0.25 μM enzyme and 1% (*w/v*) substrate. The reaction was carried out at 55 $^{\circ}\text{C}$, pH 7.4 (TBS) for 10 min, and the residual activity was determined by measuring the reducing sugar released in the reaction mixture as described earlier.

3.7. Analysis of the Hydrolysis Product

The reaction mixture (total volume of 200 μL) containing 0.25 μM enzyme and 1% (*w/v*) substrate (i.e., CMC and RAC) was assayed under optimal conditions with varying times from 15 min to 16 h. The reaction was stopped at specific time intervals by boiling for 5 min and centrifuging at $13,000\times g$ at 4 $^{\circ}\text{C}$ for 10 min to remove the undigested substrate. The soluble fraction was transferred to a new tube, and 4 μL of the soluble fraction was mixed with 4 μL of absolute ethanol. The mixed sample (8 μL) was spotted

on TLC plates. The plates were immersed in a TLC chamber containing *n*-butanol:acetic acid:water at a ratio of 2:1:1 as the mobile phase. After sufficient migration of the mobile phase, the TLC plate was removed from the chamber, dried in a hood, and sprayed with visualization solution containing 1 mL aniline, 1 g α -diphenylamine, 50 mL acetone, and 7.5 mL phosphoric acid. The TLC plate was heated at 90 °C in a hot air oven for 5 min. Five microliters of a series of sugars (G1–G6, 0.2 mg/mL each) was run in parallel and used as a standard. Hydrolysis of cellodextrins (G2–G6) was performed as with the hydrolysis of polysaccharides described above, except that the final concentration of the substrate was 0.2 mg/mL and the incubation time was 15 min.

3.8. Cohesin-Dockerin Interaction Test Using Enzyme-Linked Immunosorbent Assay (ELISA)

Affinity-based ELISA was performed as described previously by Phitsuwan, et al. [21]. In brief, the 96-well plates (NuNc, A/S, Roskilde, Denmark) were coated with the fusion protein Xyn-Doc I of *CalkGH9T* at a concentration of 1 μ g/mL, and different concentrations of CBM-cohesins of ScaA ranging from 0.001 to 1000 ng/mL were applied on the coated plates to detect cohesin-dockerin interactions. The interactions were examined immunochemically by using anti-CBM primary antibody and horseradish peroxidase (HRP)-labeled secondary antibody in the presence of chromogenic substrate TMB (3,3',5,5'-tetramethylbenzidine, Dako, Glostrup Municipality of Glostrup, Denmark). The anti-CBM primary antibody was prepared from rabbit and diluted 1:10000 in blocking buffer [42]. The reaction product (color formation) was measured at the absorbance of 450 nm.

3.9. Polysaccharide Binding Assay

Enzyme (0.25 μ M) was mixed with 1% (*w/v*) polysaccharides in TBS in a final volume of 200 μ L in a 1.5-mL microcentrifuge tube. The reaction tube was gently rotated, and the incubation temperature was maintained at 4 °C for 1 h. After binding, the reaction mixture was centrifuged at 13,000 \times *g* at 4 °C for 10 min. All of the soluble fraction was transferred to a new 1.5-mL microcentrifuge tube, and this fraction was considered “unbound protein”. The substrate pellet was washed with 500 μ L TBS by vortex mixer and then centrifuged at 13,000 \times *g* at 4 °C for 10 min. This washing step was repeated three times. The washed pellet was then mixed with 4 \times SDS sample buffer (Bio-Rad, Hercules, CA, USA), to a final volume of 60 μ L and boiled for 10 min. The protein associated with the pellet was considered “bound protein.” For unbound protein, 45 μ L of the soluble fraction was mixed with 15 μ L of 4 \times SDS sample buffer and boiled for 10 min. Thirty microliters of both unbound and bound proteins were taken and analyzed on 10% SDS-PAGE. The protein intensity in SDS-PAGE was calculated using ImageJ gel analysis [43].

3.10. Processive Activity Assay

The content of reducing ends in the soluble and insoluble fractions was determined using RAC as a substrate. Different concentrations of enzyme were added to a 1.5-mL microcentrifuge tube containing 1% (*w/v*) RAC in TBS (final volume of 200 μ L). The reaction mixture was incubated at 55 °C for different incubation times. After incubation, the reaction tube was centrifuged at 13,000 \times *g* at 4 °C for 10 min, and the supernatant was removed, of which 100 μ L was taken for measurement of reducing ends (soluble fraction) by the DNS method. The pellet was washed three times with 500 μ L TBS by centrifugation at 13,000 \times *g* at 4 °C for 5 min. After washing, the pellet was resuspended in 100 μ L TBS and taken for measurement of reducing ends (insoluble fraction) by the DNS method.

4. Conclusions

CalkGH9T of *C. alkallicellulosi* is a GH9 endoglucanase with a modular architecture. The presence of a dockerin sequence and the affinity interaction between the dockerin and the cohesin modules of the primary scaffold prove *CalkGH9T* is a cellulosomal enzyme. The enzyme actively hydrolyzed CMC but had low activity on RAC and lacked activity on Avicel. However, the existence of the CBM3 allowed *CalkGH9T* to bind tightly to

Avicel. The production of 40% reducing ends on insoluble RAC by *CalkGH9T* indicates its non-processive activity. Cellotetraose was the major hydrolysis product of CMC and RAC. The addition of *CalkGH9T* to commercial cellulase Ctec2 increased reducing sugar yield during RAC hydrolysis. Our study indicates an important role of *CalkGH9T* in cellulose degradation.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/catal11081011/s1>, Figure S1: translated amino acid sequence and a calculated molecular weight and a *pI* value of *CalkGH9T* analysed by ProtParam online tool, Figure S2: 10% SDS-PAGE analysis of *CalkGH9T* expression and purification. Lane: M, molecular weight marker (PageRuler™ prestained protein ladder, 10 to 180 kDa). Lane: 1–7, the eluted proteins from 2-mL collected fractions. Each well was loaded with 20 µL of individual eluted protein fractions, and Figure S3: Reducing sugar release from RAC by *CalkGH9T*. Values are means and bars represent the standard deviations for three independent experiments, Figure S4: TLC analysis of hydrolysis products generated by 0.25 µM *CalkGH9T* using cellodextrins with a degree of polymerization of 2–6 (G2–G6) as the substrates for 16-h incubations. S, standards of glucose and cellodextrins with a degree of polymerization of 2–6 (G2–G6). E, *CalkGH9T*. G2, G3, G4, G5, and G6 represent a reaction control (a reaction mixture containing a substrate and a denatured enzyme). G2 + E, G3 + E, G4 + E, G5 + E, and G6 + E represent enzymatic reactions. The enzymatic reactions were performed at pH 7.4 and 55 °C.

Author Contributions: Conceptualization, P.P.; methodology, P.P.; validation, P.P., S.L., and T.S.; formal analysis, S.L. and T.S.; investigation, P.P., S.L., and T.S.; resources, P.P. and K.R.; data curation, P.P.; writing—original draft preparation, review and editing, P.P.; funding acquisition, P.P. and K.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Thailand Research Fund, grant number MRG6180076, and the Biodiversity-Based Economy Development Office (Public Organization), Thailand. The APC was funded by Thailand Science Research and Innovation (TSRI), Basic Research Fund: Fiscal year 2021 under project number 64A306000039.

Data Availability Statement: All data are provided by the manuscript and the Supporting Information. Relevant data are available from the authors on request.

Acknowledgments: We gratefully thank Edward A. Bayer (Weizmann Institute of Science, Israel) for providing genomic information of *C. alkalicellulosi*.

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

References

1. Hamid, S.B.A.; Islam, M.M.; Das, R. Cellulase biocatalysis: Key influencing factors and mode of action. *Cellulose* **2015**, *22*, 2157–2182. [CrossRef]
2. Phitsuwan, P.; Sakka, K.; Ratanakhanokchai, K. Improvement of lignocellulosic biomass in planta: a review of feedstocks, biomass recalcitrance, and strategic manipulation of ideal plants designed for ethanol production and processability. *Biomass Bioenergy* **2013**, *58*, 390–405. [CrossRef]
3. Gupta, V.K.; Steindorff, A.S.; de Paula, R.G.; Silva-Rocha, R.; Mach-Aigner, A.R.; Mach, R.L.; Silva, R.N. The post-genomic era of *Trichoderma reesei*: What's next? *Trends Biotechnol.* **2016**, *34*, 970–982. [CrossRef]
4. Akram, F.; Haq, I.u.; Imran, W.; Mukhtar, H. Insight perspectives of thermostable endoglucanases for bioethanol production: A review. *Renew. Energy* **2018**, *122*, 225–238. [CrossRef]
5. Beckham, G.T.; Ståhlberg, J.; Knott, B.C.; Himmel, M.E.; Crowley, M.F.; Sandgren, M.; Sørlie, M.; Payne, C.M. Towards a molecular-level theory of carbohydrate processivity in glycoside hydrolases. *Curr. Opin. Biotechnol.* **2014**, *27*, 96–106. [CrossRef] [PubMed]
6. 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] [PubMed]
7. Arai, T.; Kosugi, A.; Chan, H.; Koukiekolo, R.; Yukawa, H.; Inui, M.; Doi, R.H. Properties of cellulosomal family 9 cellulases from *Clostridium cellulovorans*. *Appl. Microbiol. Biotechnol.* **2006**, *71*, 654–660. [CrossRef]
8. Petkun, S.; Jindou, S.; Shimon, L.J.W.; Rosenheck, S.; Bayer, E.A.; Lamed, R.; Frolow, F. Structure of a family 3b' carbohydrate-binding module from the Cel9V glycoside hydrolase from *Clostridium thermocellum*: Structural diversity and implications for carbohydrate binding. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 33–43. [CrossRef]
9. Schröder, C.; Burkhardt, C.; Busch, P.; Schirmacher, G.; Claren, J.; Antranikian, G. Characterization of a theme C glycoside hydrolase family 9 endo-β-Glucanase from a biogas reactor metagenome. *Protein J.* **2018**, *37*, 454–460. [CrossRef]

10. Zhang, K.D.; Li, W.; Wang, Y.F.; Zheng, Y.L.; Tan, F.C.; Ma, X.Q.; Yao, L.S.; Bayer, E.A.; Wang, L.S.; Li, F.L. Processive degradation of crystalline cellulose by a multimodular endoglucanase via a wirewalking mode. *Biomacromolecules* **2018**, *19*, 1686–1696. [[CrossRef](#)]
11. Mingardon, F.; Bagert, J.D.; Maisonnier, C.; Trudeau, D.L.; Arnold, F.H. Comparison of family 9 cellulases from mesophilic and thermophilic bacteria. *Appl. Environ. Microbiol.* **2011**, *77*, 1436–1442. [[CrossRef](#)]
12. Qi, M.; Jun, H.S.; Forsberg, C.W. Cel9D, an atypical 1,4- β -D-glucan glucohydrolase from *Fibrobacter succinogenes*: Characteristics, catalytic residues, and synergistic interactions with other cellulases. *J. Bacteriol.* **2008**, *190*, 1976–1984. [[CrossRef](#)]
13. Li, Y.; Irwin, D.C.; Wilson, D.B. Processivity, substrate binding, and mechanism of cellulose hydrolysis by *Thermobifida fusca* Cel9A. *Appl. Environ. Microbiol.* **2007**, *73*, 3165–3172. [[CrossRef](#)]
14. Gilad, R.; Rabinovich, L.; Yaron, S.; Bayer, E.A.; Lamed, R.; Gilbert, H.J.; Shoham, 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](#)] [[PubMed](#)]
15. Jeng, W.Y.; Liu, C.I.; Lu, T.J.; Lin, H.J.; Wang, N.C.; Wang, A.H.J. Crystal structures of the C-terminally truncated endoglucanase Cel9Q from *Clostridium thermocellum* complexed with cellodextrins and Tris. *ChemBioChem* **2019**, *20*, 295–307. [[CrossRef](#)] [[PubMed](#)]
16. Jindou, S.; Xu, Q.; Kenig, R.; Shulman, M.; Shoham, Y.; Bayer, E.A.; Lamed, R. Novel architecture of family-9 glycoside hydrolases identified in cellulosomal enzymes of *Acetivibrio cellulolyticus* and *Clostridium thermocellum*. *FEMS Microbiol. Lett.* **2006**, *254*, 308–316. [[CrossRef](#)]
17. Artzi, L.; Morag, E.; Barak, Y.; Lamed, R.; Bayer, E.A. *Clostridium clariflavum*: Key cellulosome players are revealed by proteomic analysis. *mBio* **2015**, *6*, e00411–e00415. [[CrossRef](#)]
18. Zvereva, E.A.; Fedorova, T.V.; Kevbrin, V.V.; Zhilina, T.N.; Rabinovich, M.L. Cellulase activity of a haloalkaliphilic anaerobic bacterium, strain Z-7026. *Extremophiles* **2006**, *10*, 53–60. [[CrossRef](#)] [[PubMed](#)]
19. Zhilina, T.N.; Kevbrin, V.V.; Tourova, T.P.; Lysenko, A.M.; Kostrikina, N.A.; Zavarzin, G.A. *Clostridium alkalicellum* sp. nov., an obligately alkaliphilic cellulolytic bacterium from a soda lake in the Baikal region. *Microbiology* **2005**, *74*, 557–566. [[CrossRef](#)]
20. Artzi, L.; Bayer, E.A.; Morais, S. Cellulosomes: Bacterial nanomachines for dismantling plant polysaccharides. *Nat. Rev. Microbiol.* **2017**, *15*, 83–95. [[CrossRef](#)] [[PubMed](#)]
21. Phitsuwan, P.; Morais, S.; Dassa, B.; Henrissat, B.; Bayer, E.A. The cellulosome paradigm in an extreme alkaline environment. *Microorganisms* **2019**, *7*, 347. [[CrossRef](#)]
22. Madeira, F.; Park, Y.M.; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A.R.N.; Potter, S.C.; Finn, R.D.; et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **2019**, *47*, W636–W641. [[CrossRef](#)]
23. Robert, X.; Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **2014**, *42*, W320–W324. [[CrossRef](#)] [[PubMed](#)]
24. Kurokawa, J.; Hemjinda, E.; Arai, T.; Kimura, T.; Sakka, K.; Ohmiya, K. *Clostridium thermocellum* cellulase CelT, a family 9 endoglucanase without an Ig-like domain or family 3c carbohydrate-binding module. *Appl. Microbiol. Biotechnol.* **2002**, *59*, 455–461. [[CrossRef](#)] [[PubMed](#)]
25. Petkun, S.; Rozman Grinberg, I.; Lamed, R.; Jindou, S.; Burstein, T.; Yaniv, O.; Shoham, Y.; Shimon, L.J.; Bayer, E.A.; Frolow, F. Reassembly and co-crystallization of a family 9 processive endoglucanase from its component parts: Structural and functional significance of the intermodular linker. *PeerJ* **2015**, *3*, e1126. [[CrossRef](#)] [[PubMed](#)]
26. Kim, S.J.; Kim, S.H.; Shin, S.K.; Hyeon, J.E.; Han, S.O. Mutation of a conserved tryptophan residue in the CBM3c of a GH9 endoglucanase inhibits activity. *Int. J. Biol. Macromol.* **2016**, *92*, 159–166. [[CrossRef](#)]
27. Montanier, C.; Money, V.A.; Pires, V.M.; Flint, J.E.; Pinheiro, B.A.; Goyal, A.; Prates, J.A.; Izumi, A.; Stalbrand, H.; Morland, C.; et al. The active site of a carbohydrate esterase displays divergent catalytic and noncatalytic binding functions. *PLoS Biol.* **2009**, *7*, e71. [[CrossRef](#)]
28. Gao, Z.; Ruan, L.; Chen, X.; Zhang, Y.; Xu, X. A novel salt-tolerant endo- β -1,4-glucanase Cel5A in *Vibrio* sp. G21 isolated from mangrove soil. *Appl. Microbiol. Biotechnol.* **2010**, *87*, 1373–1382. [[CrossRef](#)]
29. Percival Zhang, Y.H.; Himmel, M.E.; Mielenz, J.R. Outlook for cellulase improvement: Screening and selection strategies. *Biotechnol. Adv.* **2006**, *24*, 452–481. [[CrossRef](#)]
30. Eckert, K.; Ernst, H.A.; Schneider, E.; Larsen, S.; Lo Leggio, L. Crystallization and preliminary X-ray analysis of *Alicyclobacillus acidocaldarius* endoglucanase CelA. *Acta Crystallogr. D Biol. Crystallogr.* **2003**, *59*, 139–141. [[CrossRef](#)]
31. Yu, H.Y.; Li, X. Alkali-stable cellulase from a halophilic isolate, *Gracilibacillus* sp. SK1 and its application in lignocellulosic saccharification for ethanol production. *Biomass Bioenergy* **2015**, *81*, 19–25. [[CrossRef](#)]
32. Xue, D.S.; Liang, L.Y.; Lin, D.Q.; Gong, C.J.; Yao, S.J. Halostable catalytic properties of exoglucanase from a marine *Aspergillus niger* and secondary structure change caused by high salinities. *Process Biochem.* **2017**, *58*, 85–91. [[CrossRef](#)]
33. Danson, M.J.; Hough, D.W. The structural basis of protein halophilicity. *Comp. Biochem. Physiol. A Physiol.* **1997**, *117*, 307–312. [[CrossRef](#)]
34. Artzi, L.; Dassa, B.; Borovok, I.; Shamshoum, M.; Lamed, R.; Bayer, E.A. Cellulosomics of the cellulolytic thermophile *Clostridium clariflavum*. *Biotechnol. Biofuels* **2014**, *7*, 100. [[CrossRef](#)] [[PubMed](#)]
35. Hamberg, Y.; Ruimy-Israeli, V.; Dassa, B.; Barak, Y.; Lamed, R.; Cameron, K.; Fontes, C.M.; Bayer, E.A.; Fried, D.B. Elaborate cellulosome architecture of *Acetivibrio cellulolyticus* revealed by selective screening of cohesin-dockerin interactions. *PeerJ* **2014**, *2*, e636. [[CrossRef](#)]

36. Zhang, X.Z.; Sathitsuksanoh, N.; Zhang, Y.H.P. 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](#)] [[PubMed](#)]
37. 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](#)]
38. 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](#)]
39. 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](#)]
40. Zhang, Y.H.P.; 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](#)] [[PubMed](#)]
41. Miller, G.L. Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* **1959**, *31*, 426–428. [[CrossRef](#)]
42. Barak, Y.; Handelsman, T.; Nakar, D.; Mechaly, A.; Lamed, R.; Shoham, Y.; Bayer, E.A. Matching fusion protein systems for affinity analysis of two interacting families of proteins: The cohesin-dockerin interaction. *J. Mol. Recognit.* **2005**, *18*, 491–501. [[CrossRef](#)] [[PubMed](#)]
43. Schneider, C.A.; Rasband, W.S.; Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **2012**, *9*, 671–675. [[CrossRef](#)] [[PubMed](#)]