Modular Organization of the Thermobifida fusca Exoglucanase Cel6B Impacts Cellulose Hydrolysis and Designer Cellulosome Efficiency

Modular organization of the *Thermobifida fusca* exoglucanase Cel6B impacts cellulose hydrolysis and designer cellulosome efficiency

Eva Setter-Lamed[1], Sarah Moraïs[1], Johanna Stern[1], Raphael Lamed[2] and Edward A. Bayer[1]*

[1]*Department of Biomolecular Sciences, The Weizmann Institute of Science, Rehovot 76100 Israel.*

[2]*Department of Molecular Microbiology and Biotechnology, Tel Aviv University, Ramat Aviv, Israel*

*Corresponding author:

Edward A. Bayer, Department of Biomolecular Sciences, The Weizmann Institute of Science, Rehovot, Israel. Tel: (+972)-8-934-2373. Fax: (+972)-8-934-4118. Email: ed.bayer@weizmann.ac.il

Keywords: Cellulase, Dockerin, Enzymatic paradigm, Multifunctional enzyme, Synergy

Abbreviations: Avicel, microcrystalline cellulose; **Coh**, cohesin; **CBM**, cellulose-binding module; **GH**, glycoside hydrolases; **PASC**, phosphoric acid-swollen cellulose or amorphous cellulose

Terms: Bifunctional enzyme, an enzyme that carries two catalytic modules within a single polypeptide chain; **Cellulosome**, a multi-enzyme complex comprising a scaffoldin subunit that integrates dockerin-containing enzymes into the complex; **Enzymatic scaffoldin**, a scaffoldin that harbors an enzyme together with a cohesin in a single polypeptide chain; **Scaffoldin**, a cohesin-bearing cellulosomal scaffolding subunit; **Pseudo-cellulosome**, a divalent enzyme complex composed of interacting cohesin- and dockerin-containing enzymes.

Abstract

Cellulose deconstruction can be achieved by three distinct enzymatic paradigms: free enzymes, multifunctional enzymes and self-assembled, multi-enzyme complexes (cellulosomes). To study their comparative efficiency, the simple and efficient cellulolytic system of the aerobic bacterium, *Thermobifida fusca,* was developed as an enzymatic model. In previous studies, most of its cellulases were successfully converted to the cellulosomal mode and exhibited high cellulolytic activities, except for Cel6B, a key exoglucanase of the *T. fusca* enzymatic system. Here, we investigated the impact of the modular organization of Cel6B on enzymatic activity. The position of the cellulose-binding module (CBM), its family and linker segment were shown to affect activity. Surprisingly, exchange of the native family-2 CBM to family-3 generated an increase in Cel6B activity on cellulosic substrates. Conversion of Cel6B to the cellulosomal mode by fusing a cohesin to the catalytic module enabled formation of divalent enzyme complexes with dockerin-bearing enzymes. The resultant pseudo-cellulosomes, containing Cel6B combined with endoglucanase Cel5A, exhibited enhanced enzymatic activity, compared to mixtures of wild-type enzymes or bifunctional enzymes, unlike similar pseudo-cellulosomes containing endoglucanase Cel6A or proccessive endoglucanase Cel9A. Insight into the different enzymatic paradigms benefits ongoing development of efficient cellulolytic systems for conversion of plant-derived biomass into valuable sugars.

Novelty statement

The protein engineering of the modular arrangement of a key exoglucanase from a highly cellulolytic bacterium, *Thermobifida fusca*, served to explore and compare three major enzymatic paradigms for cellulose degradation. This approach revealed highly active chimaeric forms of the exoglucanase that act in synergy together with a potent endoglucanase in bifunctional enzymes or divalent pseudo-cellulosome-like complexes. Such engineered enzymes could be further integrated into larger enzymatic complexes, thereby providing a significant step forward towards conversion of the entire *T. fusca* free cellulolytic system into the cellulosomal mode and the enhanced conversion of cellulosic biomass into soluble sugars.

1 Introduction

Effective deconstruction of cellulose, the most abundant biopolymer on earth, is of high interest in the quest for alternative sources of energy, and also contributes to waste management [1,2]. Due to its microcrystalline nature, cellulose is, however, extremely difficult to degrade. The biological degradation of plant cell wall cellulose requires a number of cellulases and hemicellulases that act in synergy to saccharify this biopolymer and produce pure glucose units as a valuable product.

Cellulases include endoglucanases that hydrolyze the chain internally, exoglucanases that act on chain ends; both types of enzymes mainly produce cellobiose[3,4] that can be further degraded into glucose by β-D-glucosidases. The different types of cellulases hydrolyze the crystalline cellulose fibrils of the plant cell wall in synergistic fashion. In nature, three dominant enzymatic paradigms, produced by cellulolytic microorganisms for plant cell wall deconstruction have been distinguished: free enzymes, multifunctional enzymes and multi-enzyme complexes (cellulosomes) [5].

Free enzymes contain a catalytic module and generally include a carbohydrate-binding module (CBM) that targets the enzyme to the polysaccharide substrate. Multifunctional enzymes are composed of two or more catalytic modules in the same polypeptide chain, and they generally contain one or several CBMs [6,7]. The proximity of several catalytic modules in the same polypeptide chain results in an enhanced concerted action on cellulosic substrates [5,6,8,9].

Cellulosomes, first described in the anaerobe *Clostridium thermocellum* [4], adopt various supramolecular architectures. In *C. thermocellum*, the cellulosomal complex is composed of a non-catalytic scaffoldin, which contains a CBM for substrate targeting and multiple cohesin modules that integrate dockerin-bearing enzymatic subunits into the complex. The inter-modular high affinity and specific cohesin-dockerin interaction dictates the self-assembly of the cellulosome complex. The combination of substrate targeting and proximity effects contributes to the synergistic action and efficiency of cellulosomal enzymes on the cellulosic substrate [10].

Thermobifida fusca is a highly efficient cellulolytic bacterium that possesses a confined well-characterized enzymatic system. The cellulase system of this bacterium has been used as a model to compare free and cellulosomal paradigms by converting the enzymes to the cellulosomal mode, by providing a dockerin module (usually by replacement of the native CBM) [11–15]. Two of its enzymes, the endoglucanase Cel5A and the exoglucanase Cel48A, have also been engineered

together as a multifunctional enzyme by fusing the catalytic modules of two enzymes and then comparing its action on cellulosic substrates to that of the combined free wild-type enzymes and to the enzymes assembled into cellulosomal mode [9]. These studies revealed that in most cases, the cellulosomal mode has an advantage over the other paradigms.

Out of its seven cellulases [16], *T. fusca* contains two family-6 cellulases (endoglucanase Cel6A and exoglucanase Cel6B). Family 6 glycoside hydrolases (GH6s) are not part of any known natural bacterial cellulosome-producing species; they appear only in free cellulase systems of both fungi and bacteria. A previous study examined whether the two GH6 enzymes can function in the cellulosome mode by converting them into cellulosomal enzymes [17]. On cellulosic substrates, the chimaeric dockerin-containing family 6 endoglucanase (originating from Cel6A) was an efficient cellulosomal enzyme and even more active than the parent wild-type enzyme, when inserted in the designer cellulosome [11]. On the other hand, the chimaeric family 6 exoglucanase (*t*-**6B**, originating from Cel6B), was significantly less efficient than the wild-type enzyme as a component in designer cellulosomes, indicating its probable incompatibility with the cellulosomal mode of action. In addition, in another work [18], multiple copies of the same dockerin-containing enzyme chimaera originating from Cel6B were attached to a self-assembled scaffoldin containing 12 cohesin subunits, and the complex exhibited a marked reduction in its enzymatic activity as compared to same number of free wild-type enzymes. Only when the free endoglucanase Cel5A was added to the complexed exoglucanases was the efficiency increased compared to the free enzymes.

Since Cel6B is abundantly expressed during growth on cellobiose [19] and, as a key exoglucanase, Cel6B is important for achieving maximum activity of synergistic free enzyme mixtures [20], we re-investigated, in the present study, the potential incorporation of the Cel6B enzyme in designer cellulosomes by examining its modular organization (linker length, CBM family and position) and its impact on enzymatic activity. We selected the chimaeric Cel6B enzyme which provided the highest enzymatic activity and converted it to the cellulosomal mode, by using an alternative strategy: by fusing a cohesin module (as opposed to a dockerin module) to the catalytic module as observed in some cases in nature [21–25]. Finally we examined its synergistic interactions on microcrystalline cellulose as a component in divalent "pseudo-cellulosomes" with three different dockerin-bearing endoglucanases from *T. fusca*: **5A**-*t*, **6A**-*t* and **9A**-*t* (i.e., derivatives of Cel5A, Cel6A and Cel9A, respectively). In addition, we examined the action of two bifunctional enzymes

(i.e., the third enzymatic paradigm), engineered by fusions between Cel5A and Cel6B enzymatic modules and their expression in a single polypeptide chain. The action of the various engineered enzymes on cellulose degradation was examined as compared to other enzymatic paradigms containing the same catalytic units.

2 Materials and methods

2.1 Cloning of wild-type *T. fusca* enzymes and chimaeric proteins

Wild-type enzyme, Cel6B, recombinant enzymes, Cel6B$_{SL}$, **6B**-CBM2, CBM2$_5$-L5-**6B**, CBM2$_5$-L6-**6B**, CBM3-L3-**6B**, CBM3-L6-**6B**, chimaeric enzymes, **5A**-CBM2-**6B** , **5A**-**6B**-CBM2, **6B**-CBM2-**5A**, CBM2-**6B**-**5A** and cohesin constructs CBM3-L3-**6B**-CohT and CohT-CBM3-L3-**6B**, were assembled from catalytic modules and CBM cloned from *T. fusca* genomic DNA or previously described plasmids as described in Supplemental Tables 1 and 2. The **9A**-*t* chimaera was obtained by fusing the optimized catalytic module of Cel9A [26] to a dockerin of *C. thermocellum* (Cel48S). Chimaera **5A**-*t* and **6A**-*t* cloning was described before [11,12]. All enzyme constructs were designed to contain a His-tag at the N-terminus for the subsequent purification.

PCR reactions were performed using Phusion DNA polymerase, (Thermo Scientific, Waltham, MA, USA) and ABgene Reddymix x2 (Advanced Biotechnologies Ltd., United Kingdom), and DNA samples were purified using a HiYield^{TM} Gel/PCR Fragments Extraction Kit (Real GenomicsTaiwan). Minipreps were prepared using QIA prep Spin Mini prep kit (Qiagen, Netherlands).

2.2 Protein expression and purification

The plasmids were expressed in *Escherichia coli* BL21 (1DE3) pLysS cells and purified on a Ni-NTA column (Qiagen), as reported earlier [17]. The CBM-containing proteins were purified by affinity to cellulose as previously described [27]. Purity of the recombinant proteins was tested by SDS-PAGE on 10% acrylamide gels and the fractions containing the pure recombinant protein were pooled. Protein concentration was determined by absorbance (280 nm) based on the known amino acid composition of the protein using the Protparam tool http://www.expasy.org/tools/protparam.html). If necessary, proteins were concentrated using Vivaspin concentrators. Proteins were stored in 50% (v/v) glycerol at -20°C.

2.3 Non-denaturing gel electrophoresis

The optimal stoichiometry of the interaction was determined by non-denaturing gel electrophoresis. Equimolar amounts of the different proteins were combined in 30 µl final volume of Binding buffer, consisting of Tris buffer saline (TBS) pH 7.4, supplemented with 10 mM $CaCl_2$ and 0.05% Tween-20. The tubes were incubated 1 hr at 37°C. Sample buffer (15 µl from 3X stock, containing: 3 mL glycerol, 0.6 mL 10x running buffer, 6.4 mL H_2O and bromophenol blue), was added to 30 µl of the reaction mixture, and 25 µl was loaded onto non-denaturing gel (4.3% stacking/9% separating gels). The gel was run at 100 V for 2h and stained with Coomassie Blue. The resultant mobility patterns were assessed, and the optimal cohesin-dockerin bearing enzyme ratio, was chosen for the enzymatic activity reactions.

2.4 Enzyme activity assay

All assays were performed at least twice and in triplicate. Activity was performed on PASC (phosphoric acid swollen cellulose) and crystalline cellulose (Avicel, Sigma). Degradation of cellulosic substrates was assayed using enzyme concentrations of 1 µM (wild-type and chimaeras) in 50 mM acetate buffer pH 5.0 in a final volume of 200 µl. For PASC assays, 150 µl of 7.5 g/l PASC was added into the reaction and for Avicel assays, a volume of 100 µl 10% Avicel was used. The 1.5 ml tubes were incubated at 50°C for 1h for PASC or 21 h for Avicel under oblique agitation (400 rpm). Reactions were terminated by immersing the sample tubes in ice water, and the samples were centrifuged at maximum speed (14,000 x g to pellet the substrate. Dinitrosalicylic acid (DNS) reagent (150 µl) was added to 100 µl supernatant fluid and the reaction mixture was boiled for 10 minutes [28]. A volume of 200 µl was transferred to a 96 well plate, and optical density was measured at 540 nm by ELISA plate reader (Modular devices, Versa max microplate reader). Enzymatic activity was calculated using a glucose standard curve and was expressed in millimolar of glucose equivalent.

For designer cellulosome experiments, the cohesin-bearing enzyme interacted with dockerin-bearing enzymes in the absence of substrate, prior to the activity assay for 1 hr.

2.5 Affinity pull-down assay

Equimolar amount of the constructs: Cel6B, CBM3-L3-**6B**, $CBM2_5$-L5-**6B** and $CBM2_5$-L6-**6B**, were tested for their binding ability to cellulose as described by Morais [9]. Briefly, each enzyme construct (1000 picomoles) was incubated with 20 mg Avicel cellulose in 1xTBS containing 12 mM $CaCl_2$ .1

hr at 4°C. The tubes were then centrifuged at 16,000 x g for 2 min. The supernatants (130 µl containing unbound proteins), were carefully removed, and 40 µl were supplied with 20 µl SDS containing sample buffer to a final volume of 60 µl. The pellets (containing bound proteins) was washed twice by resuspension in 200 µl of 1xTBS containing 12 mM $CaCl_2$ and 0.05% Tween-20, to eliminate non-specific binding. It was then centrifuged at 16,000 x g for 2 min and the pellet was resuspended in 130 µl of 1xTBS with $CaCl_2$. SDS containing sample buffer (20 µl) was added to 40 µl of the pellet suspension. Both unbound and bound fractions were boiled for 10 min and then analyzed by SDS PAGE using a 10% polyacrylamide gel.

3 Results

3.1 Production of Cel6B chimaeras with various modular organizations

The recombinant proteins designed for use in this study are shown schematically in Figure 1. Four recombinant wild-type *T. fusca* enzymes were prepared for this work: Endoglucanase Cel5A and exoglucanase Cel6B have family-2 CBMs (CBM2) at their N termini, whereas endoglucanase Cel6A and processive-endoglucanase Cel9A each has a CBM2 at its C terminus. The latter enzyme also has a CBM3c fused to the GH9 catalytic module, which is thought to be essential for endoglucanase processivity[29,30].

Several variants of the wild-type Cel6B were prepared which differ in the nature (family) of their CBM, its position in the polypeptide chain and the linker that connects it to the GH6 catalytic module (Figure 1). In addition, two bifunctional enzymes were designed, in which the catalytic modules of exoglucanase Cel6B and endoglucanase Cel5A were fused with a CBM2 on either the N or C terminus of the protein. In the bifunctional enzymes, only one CBM2 (originating from Cel6B) is used to target both catalytic modules to the cellulosic substrate. Finally, a number of cellulosome-like components were prepared as will be discussed below.

Initially, the modular architecture of Cel6B and its effect on enzymatic activity was investigated. The CBM of this enzyme, located at the N-terminus of the protein in the native state, belongs to family 2, and has the ability to bind microcrystalline and amorphous cellulosic substrates [31]. The CBM2 was shown to be critical to proper enzymatic activity into designer cellulosomes, as its replacement with a dockerin module [11] led to a strong antagonistic effect with partner enzymes. Therefore, several modifications of the original wild-type enzyme were investigated as described below.

In order to examine whether linker length affects the enzymatic activity of Cel6B, the wild-type linker was reduced in length from 32 to 15 residues, and the resultant chimaera was termed Cel6B$_{SL}$. In order to explore whether the position of the CBM influences the enzymatic activity of the cellulase, the location of the CBM2 of the wild-type enzyme was altered from the N to the C terminus to form 6B-CBM2. Furthermore, in order to gain insight into the relationship of the Cel6B catalytic module with its original family 2 CBM, a series of chimaeras were designed in which the CBM2 of the Cel6B enzyme was replaced either by a CBM2 from the endoglucanase Cel5A from *T. fusca*, thus generating the CBM2$_5$-L5-**6B** (CBM2 and linker, both from Cel5A) and CBM2$_5$-L6-**6B** (CBM2 from Cel5A and linker from Cel6B), or by the CBM3a from the *C. thermocellum* scaffoldin CipA thus forming CBM3-L6-**6B** (CBM3a from CipA and linker from Cel6B) and CBM3-L3-**6B** (CBM3a and linker from CipA) (see linker length and composition in Table 1).

In addition, two bifunctional enzyme-chimaeras were designed, which contained permutations of the family 6 and family 5 catalytic module together with the CBM2 from the wild-type Cel6B, thus generating the **5A**-**6B**-CBM2 and CBM2-**6B**-**5A** bifunctional enzymes. Two additional bifunctional enzymes, **5A**-CBM2-**6B** and **6B**-CBM-**5A** were designed, cloned, expressed and purified on a Ni-NTA column, but the purity of those proteins tested by SDS-PAGE was not sufficient. Therefore, we only examined the activities of **5A**-**6B**-CBM2 and CBM2-**6B**-**5A** in our study. All additional purified recombinant proteins, showed a single major band on SDS-PAGE (Supplemental Figure 1), and in each case the mobility was consistent with their molecular mass.

3.2 Comparative enzymatic activity of the Cel6B chimaeras

The enzymatic activity of the recombinant enzymes was assayed on amorphous cellulose (PASC) and microcrystalline cellulose (Avicel). The recombinant enzymes and chimeras: Cel6B$_{SL}$, **6B**-CBM2, CBM2$_5$-L5-**6B**, CBM2$_5$-L6-**6B**, CBM3-L3-**6B** and CBM3-L6-**6B** were compared to the wild-type enzyme Cel6B (Figure 2A and B).

In most cases, the enzymatic activity of the recombinant enzymes exhibited similar tendencies on PASC and Avicel substrates. The enzymatic activity of Cel6B$_{SL}$ was very similar to the wild-type Cel6B on both substrates, suggesting that the length of the linker is not an important parameter for catalytic efficiency of Cel6B. In fact, the linker origin had a higher impact on the cellulolytic activity than its length. In addition, the Cel6B variants with a linker adjacent to a CBM of the same

origin (i.e., CBM3 along with L3; and $CBM2_5$ along with L5) were in all cases more active than variants bearing a linker adjacent to a CBM of different origin (i.e., CBM3 or $CBM2_5$ with L6).

Surprisingly, both displacement of the CBM from the N to the C terminus in **6B**-CBM and reduction of the linker length produced a higher cellulosic activity *versus* that of the wild-type Cel6B on both substrates. Intriguingly, CBM3-L3-**6B**, where both CBM and linker originated from the *C. thermocellum* cellulosomal scaffoldin, generated the highest cellulolytic activity on both substrates, compared to the wild-type Cel6B and all other recombinant enzymes tested. CBM3-L6-**6B** (with the Cel6B wild-type linker) exhibited enhanced activity on Avicel substrate only.

The fact that the CBM3 from the *C. thermocellum* scaffoldin together with its adjacent linker enhances the enzymatic activity of the catalytic module, could result from an enhancement of the binding affinity to cellulose. Therefore, an affinity pull-down assay was performed to demonstrate whether the binding to cellulose of the various enzyme constructs is similar, and to understand whether the difference in enzymatic activity is connected to their binding ability. Four enzymes were thus compared for their ability to bind cellulose: the wild-type Cel6B exoglucanase (containing the native CBM2), CBM3-L3-**6B** (containing CBM3), $CBM2_5$-L5-**6B** and $CBM2_5$-L6-**6B** (both containing the CBM2 from a different enzyme, i.e., the Cel5A endoglucanase). All proteins were fully present in the bound fractions, and no protein bands were detected in any of the supernatant (unbound) fraction (Supplemental Figure 2), indicating, that all of the tested enzymes exhibited strong and full binding ability to cellulose.

3.3 Enzymatic activity of the bifunctional enzymes on microcrystalline cellulose

In contrast to a previous report by Irwin et al [20], where a synergistic effect of 7.8 between E5 (Cel5A) and E3 (Cel6B) was observed on filter paper; here, on Avicel, and under our defined reaction conditions, no synergism was observed between the wild-type free enzymes Cel5A and Cel6B (Figure 3). Nevertheless, fusion of the two catalytic modules into a single polypeptide chain in the recombinant **5A**-**6B**-CBM2 and CBM2-**6B**-**5A** bifunctional enzymes resulted in an enhancement of 1.1 to 1.4 (bars 9 and 10 in Figure 3) as compared to the sum of the individual activities of the free wild-type enzymes (bar 7), CBM2-**6B**-**5A** exhibiting the highest enzymatic activity. The enzymatic activity of the combination of free enzymes **6B**-CBM + Cel5A (bar 8), which reflects the modular organization **5A**-**6B**-CBM2 (bar 10), was even higher than that of the bifunctional chimaera, and only slightly lower than that of CBM2-**6B**-**5A** (bar 9).

3.4 Conversion of free enzymes to the cellulosomal mode

The CBM3-L3-**6B** chimaera, which yielded the highest cellulolytic activity, was converted to the cellulosomal mode, by fusing a cohesin to its catalytic module either at the N- or C-terminal end, thus generating the enzymatic scaffoldins: CBM3-L3-**6B**-CohT and CohT-CBM3-L3-**6B**. The CohT (cohesin 3 from the C. *thermocellum* CipA scaffoldin) interacts specifically with the dockerin (designated *t*) from *C. thermocellum* Cel48S. This dockerin was ligated to three different *T. fusca* endoglucanases: Cel5A, Cel6A and Cel9A at their C termini to form **5A**-*t,* **6A**-*t* and **9A**-*t*. In previous reports, the CBM2 was removed from the respective **5A**-*t* and **6A**-*t* chimaeras without affecting enzyme catalytic efficiency [13], [12]. In the case of the processive endoglucanase Cel9A, our present data indicate that the CBM-restored enzyme (i.e., the **9A**-*t* chimaera with the matching monovalent scaffoldin, Supplemental Figure 3) yielded similar levels of activity on Avicel as substrate compared to that of the wild-type enzyme.

Pseudo-cellulosomes were achieved via interaction of enzyme-bearing scaffoldins CBM3-L3-**6B**-CohT or CohT-CBM3-L3-**6B** with either **5A**-*t*, **6A**-*t* or **9A**-*t* resulting in six different pseudo designer cellulosomes composed in each case of two interacting proteins (Figure 3, bars 11 and 12, and Figure 4 bars 10 through 13). Complex formation of each cohesin construct with dockerin-bearing enzymes was verified by non-denaturing gel electrophoresis. Predetermined stoichiometric mixtures of the enzymatic scaffoldins with enzyme bearing dockerins, versus single constructs, resulted in a single band with altered mobility (band with increased intensity and shifted), indicating, that complete or near complete complexation was achieved (see Supplemental Figure 4 for a representative example of complex formation).

The attachment of either CBM3-L3-6B or CBM3-L3-6B to **5A**-*t* did not result in a proximity effect between the catalytic modules (Figure 3, bars 3+6 compared with 11 and 4+ 6 compared with 12). However, the pseudo-cellulosomes comprising endoglucanase **5A**-*t* and the exoglucanase Cel6B-bearing scaffoldins were about 1.6 fold more efficient than the combination of Cel5A and Cel6B as free enzymes (Figure 3 bars 11 and 12 as compared to bar 7).

For endoglucanase Cel6A or the proccessive endoglucanase Cel9A complexed to cellulosomal forms of Cel6B, the free wild-type enzymes retained their advantage compared to the pseudo-cellulosomes (Figure 4, bars 10–13). Nevertheless, the enzymatic activities of these cellulosomes were relatively high, supporting the high functionality of the cellulosomal form of Cel6B.

4 Discussion

The critical involvement of exoglucanases in the hydrolysis of recalcitrant crystalline cellulosic substrates is a dynamic phenomenon, associated with their movement on the cellulose surface [32–36]– a process believed to be a function of the resident CBM [37–40]. This would presumably infer an intimate spatial and functional relationship between the catalytic module and CBM of the parent protein, although the precise mechanism of processivity remains poorly understood[41].

Although the enzymes of the cellulase system in *T. fusca* are all of the free-enzyme paradigm[42], several of them have been subject to conversion into the cellulosomal [12,13,17] and multi-functional enzyme modes[9]. We therefore employed this approach in the present work to explore the role of the original CBM2 and linker on the enzyme activity of *T. fusca* exoglucanase Cel6B.

Cel6B is one of the key enzymes of the *T. fusca* cellulase system, contributing to its very high cellulolytic activity [20], and the enzyme is abundantly expressed during growth on cellobiose [19]. However, conversion of the free Cel6B to the cellulosome mode by replacement of its CBM2 with a dockerin (*t*-**6B**) caused an "anti-proximity" effect, and integration of the exoglucanase into designer cellulosomes led to relatively poor levels of activity [17]. It is therefore intriguing that either displacement of the CBM2 of Cel6B from the N- to the C-terminal end of the protein or its replacement by a scaffoldin-derived CBM3 both led to an increase in enzyme activity on cellulosic substrates (Figure 2). This surprising result would infer plasticity in the relationship of the CBM and catalytic module of this exoglucanase, similar to previous studies on modular positioning in endoglucanase Cel5A [12].

According to our results (Figure 2), linker origin had a more significant impact on cellulolytic activity than linker length. Similar results were obtained earlier, where the length of the linker segment that separates the dockerin from the catalytic module appeared to have little if any effect on enzymatic activity [12]. Additional studies on linker length in designer cellulosomes [43] revealed that linker length between modules in scaffoldins does have a limited but defined impact on cellulolytic activity on crystalline cellulose and wheat straw. Another study [44] concluded similarly that the inter-cohesin linkers, whose length and composition vary considerably among bacterial scaffoldins, have limited or no impact on the synergy, proximity effect and activity of

cellulosomal cellulases. Nevertheless, long linkers in bacterial scaffoldins that mediate the anchoring of cells to the cellulose surface might be required. Indeed in intricate cellulosome systems, long linkers (up to 550 residues) may help maintain conformational flexibility, which is critical for binding to the cellulosic substrate and compensating for stringency [25].

We selected the chimaeric enzyme, which yielded the highest enzymatic activity and converted it to the cellulosomal mode. Here, we opted for the fusion of a cohesin module (and not a dockerin module) to the catalytic module as observed in some cases in nature [23,45], notably as complexed bifunctional enzymes (toxins) in certain pathogenic bacteria, e.g., *Clostridium perfringens*[46]. Such enzyme-bearing scaffoldins can interact with dockerin-containing enzymes and were used to form various types of cellulosome-like complexes of novel atypical geometry, defined as "symmetrical", "asymmetrical", "cyclic" and "polymeric" cellulosomes, as engineered previously by Mingardon and co-workers [25]. The resulting unconventional cellulosome types exhibited reduced enzymatic acitvity as compared to conventional designer cellulosomes, presumably due to the altered mobility of the enzymes.

In our present study, synergism was tested between the wild-type *T. fusca* exoglucanase Cel6B (that hydrolyzes cellulose from the non-reducing end*)*, together with endoglucanases Cel5A or Cel6A, or with processive endoglucanase Cel9A (Figures 3 and 4). According to previous studies, the Cel6B exoglucanase is important for achieving maximum activity of synergistic mixtures [20,47]. However, in our study, (Figures 3 and 4) no synergism was observed between the wild-type enzymes. This lack of synergism of Cel6B, can be due to the different substrate employed and/or different ratios of the individual enzymes [48].

As reported before [20], excess exocellulase favors synergism, and the optimal molar ratio between Cel6B and Cel5A was observed at 4:1 [20,48]. In our work, equal ratios were used between endocellulase and exocellulase to conform with the bifunctional enzymes and divalent pseudo-cellulosomes, wherein each GH module was integrated in a single copy. In our experimental conditions, increased ratio in favor of exoglucanase also favored synergism, until 10:1 ratio (Cel6B:Cel5A) as can been seen in Supplemental Figure 5.

A relationship between synergism and processivity was observed by Vuong and co-workers [48]. In the latter study, *T. fusca* Cel6B was mutated to enhance activity on selected cellulosic substrates, but the mutants were less synergistic with the wild-type Cel5A endoglucanase than the wild-type

Cel6B. These results suggested that improving activity by increasing processivity of a single enzyme is not always an effective strategy to achieve higher synergism between enzymes.

Here, fusing the catalytic modules on a single polypeptide chain into the bifunctional enzyme CBM2-**6B**-**5A**, resulted in activity enhancement as compared to the mixture of the free wild-type enzymes. Likewise, the pseudo-cellulosomes containing Cel6B and Cel5A derivatives achieved enhanced levels of degradation of the cellulosic substrate as compared to the corresponding mixture the wild-type enzymes. Moreover, pseudo-cellulosomes containing 6A-*t* and 9A-*t* attached to the cellulosomal forms of Cel6B were only slightly less effective than the combined activities of the wild-type enzymes counterparts. In light of the results provided here, the designed pseudo-cellulosomal version of Cel6B, containing a cohesin module and a CBM2 is adequate for integration of cellulosomal enzymes into pseudo designer cellulosomes. In future studies, Cel6B could thus be fused into larger scaffoldins to integrate larger numbers of enzymes thus providing a considerable step forward in the conversion of the entire *T. fusca* cellulolytic system into the cellulosomal mode.

Acknowledgements

This manuscript is dedicated to the memory of David B. Wilson (Cornell University), who passed away on April 29, 2017. We are indebted to David for establishing the cellulase system of *Thermobifida fusca* and for many years of close collaboration and friendship. David was an exceptional scientist, a kind, goodhearted and very straightforward person, and a dear colleague and friend.

This research was supported by Grant No. 1349 from the Israel Science Foundation (ISF) and by the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel. The authors appreciate the support of the European Union, Area NMP.2013.1.1-2: Self-assembly of naturally occurring nanosystems: CellulosomePlus Project number: 604530 and European Union Horizon 2020 contract: Sustainable production of next generation biofuels from waste streams: Waste2Fuels. EAB is the incumbent of The Maynard I. and Elaine Wishner Chair of Bio-organic Chemistry.

Conflict of interest

The authors declare no financial or commercial conflict of interest.

5 References

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

[2] Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol Mol Biol Rev*. 2002;*66*:506-577.

[3] Morag E, Halevy I, Bayer EA, Lamed R. Isolation and properties of a major cellobiohydrolase from the cellulosome of *Clostridium thermocellum*. *J Bacteriol*. 1991;*173*:4155-4162.

[4] Lamed R, Setter E, Bayer EA. Characterization of a cellulose-binding, cellulase-containing complex in *Clostridium thermocellum*. *J Bacteriol*. 1983;*156*:828-836.

[5] Himmel ME, Xu Q, Luo Y, Ding S-Y, Lamed R, Bayer EA. Microbial enzyme systems for biomass conversion: emerging paradigms. *Biofuels*. 2010;*1*:323-341.

[6] Xu Q, Luo Y, Ding S, Himmel ME. Multifunctional Enzyme Systems for Plant Cell Wall Degradation. *Compr Biotechnol*. 2011;*3*:15-25.

[7] Xu Q, Ding S-Y, Brunecky R, Bomble YJ, Himmel ME, Baker JO. Improving activity of minicellulosomes by integration of intra- and intermolecular synergies. *Biotechnol Biofuels*. 2013;*6*:126.

[8] Walker JA, Takasuka TE, Deng K, et al. Multifunctional cellulase catalysis targeted by fusion to different carbohydrate-binding modules. *Biotechnol Biofuels*. 2015;*8*:220.

[9] Moraïs S, Barak Y, Lamed R, et al. Paradigmatic status of an endo- and exoglucanase and its effect on crystalline cellulose degradation. *Biotechnol Biofuels*. 2012;*5*:78.

[10] Fierobe H, Mingardon F, Mechaly A, et al. Action of Designer Cellulosomes on Homogeneous Versus Comlex substrates. 2005;*280*:16325-16334.

[11] Caspi J, Barak Y, Haimovitz R, et al. *Thermobifida fusca* exoglucanase Cel6B is incompatible with the cellulosomal mode in contrast to endoglucanase Cel6A. *Syst Synth Biol*. 2010;*4*:193-201.

[12] Caspi J, Barak Y, Haimovitz R, et al. Effect of linker length and dockerin position on conversion of a *Thermobifida fusca* endoglucanase to the cellulosomal mode. *Appl Environ Microbiol*. 2009;*75*:7335-7342.

[13] Caspi J, Irwin D, Lamed R, et al. Conversion of *Thermobifida fusca* free exoglucanases into cellulosomal components: Comparative impact on cellulose-degrading activity. *J Biotechnol*. 2008;*135*:351-357.

[14] Moraïs S, Morag E, Barak Y, et al. Deconstruction of lignocellulose into soluble sugars by native and designer cellulosomes. *MBio*. 2012;*3*:1-11.

[15] Stern J, Moraïs S, Lamed R, Bayer EA. Adaptor scaffoldins: An original strategy for extended

designer cellulosomes, inspired from nature. *MBio*. 2016:*7*:e00083-16.

[16] Posta K, Béki E, Wilson DB, Kukolya J, Hornok L. Cloning, characterization and phylogenetic relationships of cel5B, a new endoglucanase encoding gene from *Thermobifida fusca*. *J Basic Microbiol*. 2004;*44*:383-399.

[17] Caspi J, Irwin D, Lamed R, et al. *Thermobifida fusca* family-6 cellulases as potential designer cellulosome components. *Biocatal Biotransformation*. 2006;*24*:3-12.

[18] Moraïs S, Heyman A, Barak Y, et al. Enhanced cellulose degradation by nano-complexed enzymes: Synergism between a scaffold-linked exoglucanase and a free endoglucanase. *J Biotechnol*. 2010;*147*:205-211.

[19] Chen S, Wilson DB. Proteomic and transcriptomic analysis of extracellular proteins and mRNA levels in *Thermobifida fusca* grown on cellobiose and glucose. *J Bacteriol*. 2007;*189*:6260-6265.

[20] Irwin DC, Spezio M, Walker LP, Wilson DB. Activity studies of eight purified cellulases: Specificity, synergism, and binding domain effects. *Biotechnol Bioeng*. 1993;*42*:1002-1013.

[21] Dassa B, Borovok I, Lamed R, et al. Genome-wide analysis of *Acetivibrio cellulolyticus* provides a blueprint of an elaborate cellulosome system. *BMC Genomics*. 2012;*13*:210.

[22] Artzi L, Dassa B, Borovok I, Shamshoum M, Lamed R, Bayer EA. Cellulosomics of the cellulolytic thermophile *Clostridium clariflavum*. *Biotechnol Biofuels*. 2014;*7*:100.

[23] Ben David Y, Dassa B, Borovok I, et al. Ruminococcal cellulosome systems from rumen to human. *Environ Microbiol*. 2015;*17*:3407-3426.

[24] Jindou S, Xu Q, Kenig R, Shulman M, Shoham Y, Bayer EA. in cellulosomal enzymes of *Acetivibrio cellulolyticus* and *Clostridium thermocellum*. 2006;*254*:308-316.

[25] Mingardon F, Tardif C, Bayer EA, Fierobe H. Exploration of new geometries in cellulosome-like chimeras. 2007;*73*:7138-7149.

[26] Mingardon F, Bagert JD, Maisonnier C, Trudeau DL, Arnold FH. Comparison of family 9 cellulases from mesophilic and thermophilic bacteria. *Appl Environ Microbiol*. 2011;*77*:1436-1442.

[27] Moraïs S, David Y Ben, Bensoussan L, et al. 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.

[28] Miller GL. Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal Biochem*. 1959;*31*:426-428.

[29] Wilson DB, Kostylev M. Cellulase Processivity. *Methods Mol Biol*. 2012;*908*:93-99.

[30] Sakon J, Irwin D, Wilson D, Karplus P. Structure and mechanism of endo/exocellulase E4

from *Thermomonospora fusca*. *Nat Struct Biol*. 1997;*4*:810-818.

[31] Irwin DC, Zhang S, Wilson DB. Cloning, expression and characterization of a family 48 exocellulase, Cel48A, from *Thermobifida fusca*. *Eur J Biochem*. 2000;*267*:4988-4997.

[32] Kostylev M, Wilson D. Synergistic interactions in cellulose hydrolysis. *Biofuels*. 2012;*3*:61-70.

[33] Wilson DB. Microbial diversity of cellulose hydrolysis. *Curr Opin Microbiol*. 2011;*14*:1-5.

[34] Igarashi K, Uchihashi T, Koivula A, et al. Traffic Jams Reduce Hydrolytic Efficiency of Cellulase on Cellulose Surface. *Science*. 2011;*333*:1279-1282.

[35] Igarashi K, Koivula A, Wada M, Kimura S, Penttila M, Samejima M. High speed atomic force microscopy visualizes processive movement of *Trichoderma reesei* cellobiohydrolase I on crystalline cellulose. *J Biol Chem*. 2009;284:36186-36190.

[36] Jung J, Sethi A, Gaiotto T, Han JJ, Jeoh T, Goodwin PM. Binding and Movement of Individual Cel7A Cellobiohydrolase s on Crystalline Cellulose Surfaces Revealed by Single -Molecule Fluorescence Imaging. *JBC*. 2013;*288*:24164-24172.

[37] Jung H, Wilson DB, Walker LP. Binding mechanisms for *Thermobifida fusca* Cel5A, Cel6B, and Cel48A cellulose-binding modules on bacterial microcrystalline cellulose. *Biotechnol Bioeng*. 2002;*80*:380-392.

[38] Boraston AB, Bolam DN, Gilbert HJ, Davies GJ. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J*. 2004;*382*:769-781.

[39] Kont R, Kari J, Borch K, Westh P, Väljamäe P. Inter-domain Synergism Is Required for Efficient Feeding of Cellulose Chain into Active Site of Cellobiohydrolase Cel7A. *J Biol Chem* . 2016;*291*:26013-26023.

[40] Liu Y-S, Zeng Y, Luo Y, et al. Does the cellulose-binding module move on the cellulose surface? *Cellulose*. 2009;*16*:587-597.

[41] Cruys-bagger N, Elmerdahl J, Praestgaard E, Borch K, Westh P. A steady-state theory for processive cellulases. *FEBS J*. 2013;*280*:3952-3961.

[42] Wilson DB. Studies of *Thermobifida fusca* plant cell wall degrading enzymes. *Chem Rec*. 2004;*4*:72-82.

[43] Vazana Y, Barak Y, Unger T, et al. A synthetic biology approach for evaluating the functional contribution of designer cellulosome components to deconstruction of cellulosic substrates. *Biotechnol Biofuels*. 2013;*6*:182.

[44] Molinier A-L, Nouailler M, Valette O, Tardif C, Receveur-Bréchot V, Fierobe H-P. Synergy, Structure and Conformational Flexibility of Hybrid Cellulosomes Displaying Various Inter-cohesins Linkers. *J Mol Biol*. 2011;*405*:143-157.

[45] Hamberg Y, Ruimy-israeli V, Dassa B, et al. Elaborate cellulosome architecture of *Acetivibrio cellulolyticus* revealed by selective screening of cohesin – dockerin interactions. *PeerJ*. 2014;2:e636.

[46] Adams JJ, Gregg K, Bayer EA, Boraston AB, Smith SP. Structural basis of *Clostridium perfringens* toxin complex formation. *PNAS*. 2008;*105*:12194-12199.

[47] Jeoh T, Wilson DB, Walker LP. Cooperative and Competitive Binding in Synergistic Mixtures of *Thermobifida fusca* Cellulases Cel5A, Cel6B, and Cel9A. *Biotechnol Prog*. 2002;*18*:760-769.

[48] Vuong T V, Wilson DB. Processivity , synergism , and substrate specificity of *Thermobifida fusca* Cel6B. *Appl Env Microbiol*. 2009;*75*:6655-6661.

Table 1: Linker length and composition of the Cel6B chimaeras

Enzyme	***Linker origin***	***Linker composition***	***Linker length***
Cel6B	Cel6B	SDDPDPEPSPSPSPSPSPTDPDEPGGPTNPPT	32
$Cel6B_{SL}$	Cel6B	(TS)DPDEPGGPTNPPT	15
6B-CBM2	Cel6B (N terminus)	(TS)APANAAG	9
$CBM2_5$-L5-**6B**	Cel5A	DEGSEPGGPGGPGTPSPDPGTQPGTGT	29
$CBM2_5$-L6-**6B**	Cel6B	(TS)SDDPDPEPSPSPSPSPSPTDPDEPGGPTNPPT	34
CBM3-L3-**6B**	CipA *C. thermocellum*	GGSVVPSTQPVTTPPATTKPPATTIPPSDDPNA	33
CBM3-L6-**6B**	Cel6B	TSDPDEPGGPTNPPT	15

(TS) represents the restriction site of SpeI

Figures

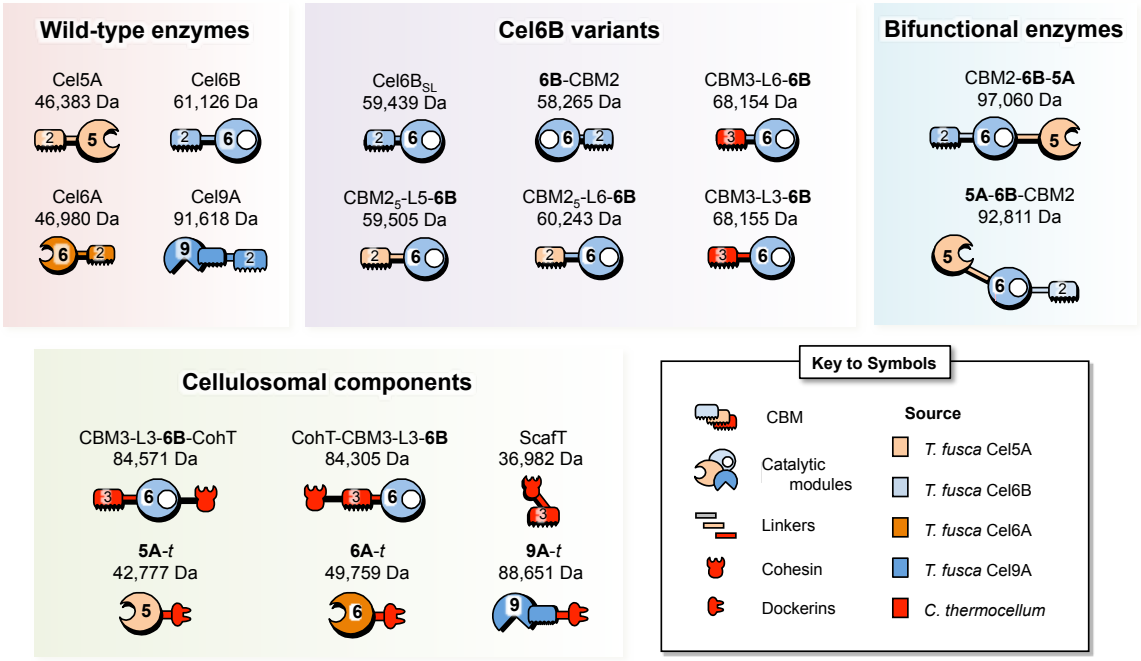


Figure 1. Schematic representations of the proteins used in this study. The origin of the catalytic module and cellulosomal elements is shown color-coded in the key. The family number of the catalytic module and CBM is shown within the pictogram. The predicted molecular weight of each protein is given.

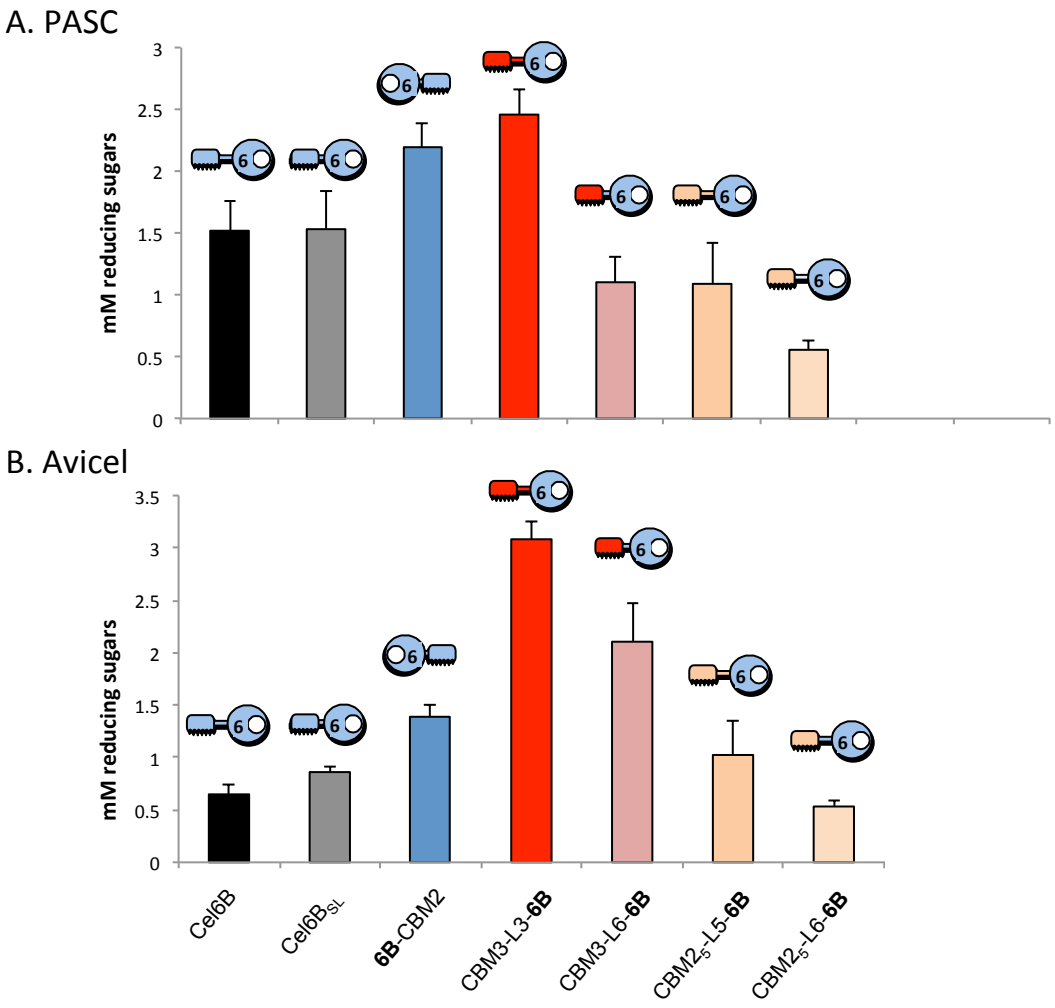


Figure 2. Comparative enzymatic activities of the Cel6B variants on (A) amorphous cellulose (PASC) and (B) microcrystalline cellulose (Avicel). Enzymatic activity is defined as mM total reducing sugars following a 1-h or 21-h reaction period (PASC and Avicel, respectively) at 50°C. Each reaction was performed in triplicate, and standard deviations are indicated.

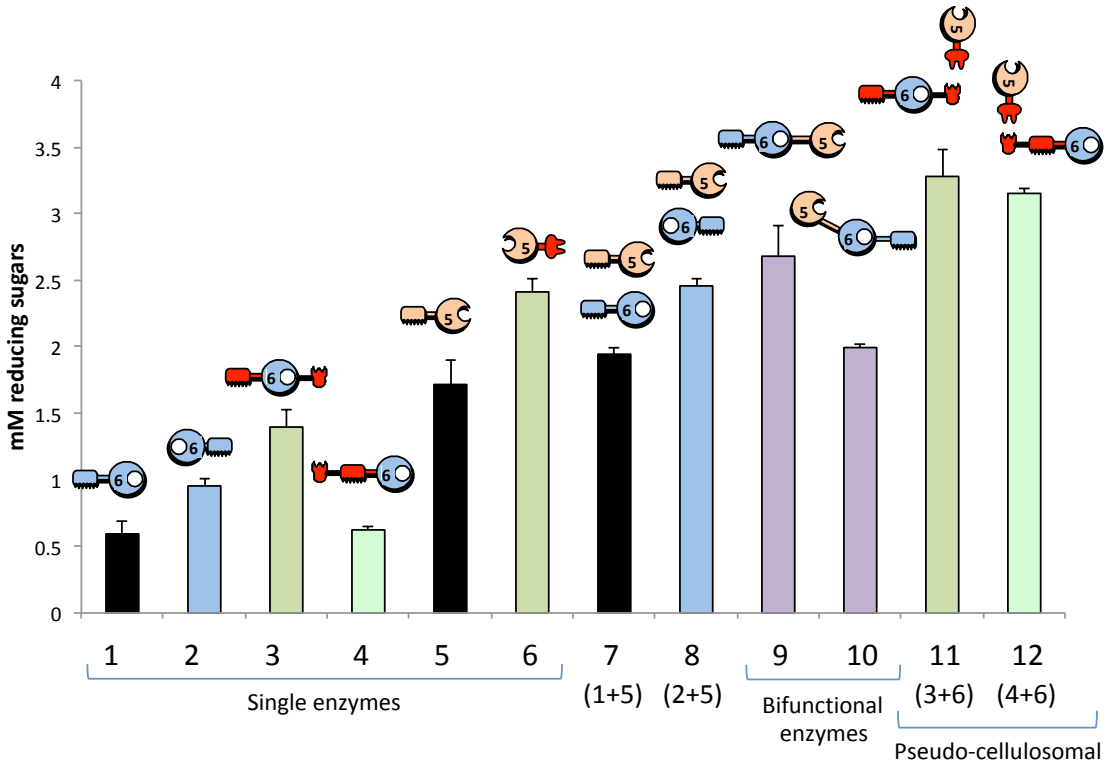


Figure 3. Comparative enzymatic activities on Avicel of the wild-type Cel6B exoglucanase and variants with the Cel5A endoglucanase, compared to those of bifunctional enzymes and pseudo-cellulosomes. Lanes: (1) wild-type Cel6B, (2) **6B**-CBM2, (3) CBM3-L3-**6B**-CohT (4) CohT-CBM3-L3-**6B** (5) wild-type Cel5A (6) **5A**-*t* (7) Cel5A + Cel6B (wild-type enzymes), (8) wild-type Cel5A + **6B**-CBM2 (9) bifunctional CBM2-**6B**-**5A** (10) bifunctional **5A**-**6B**-CBM2, (11) pseudo-cellulosome: **5A**-*t* bound to CBM3-L3-**6B**-CohT, and (12) pseudo-cellulosome: **5A**-*t* bound to CohT-CBM3-L3-**6B**. Enzymatic activity is defined as mM total reducing sugars released from Avicel following a 21-h reaction period at 50°C. Each reaction was performed in triplicate, and standard deviations are indicated.

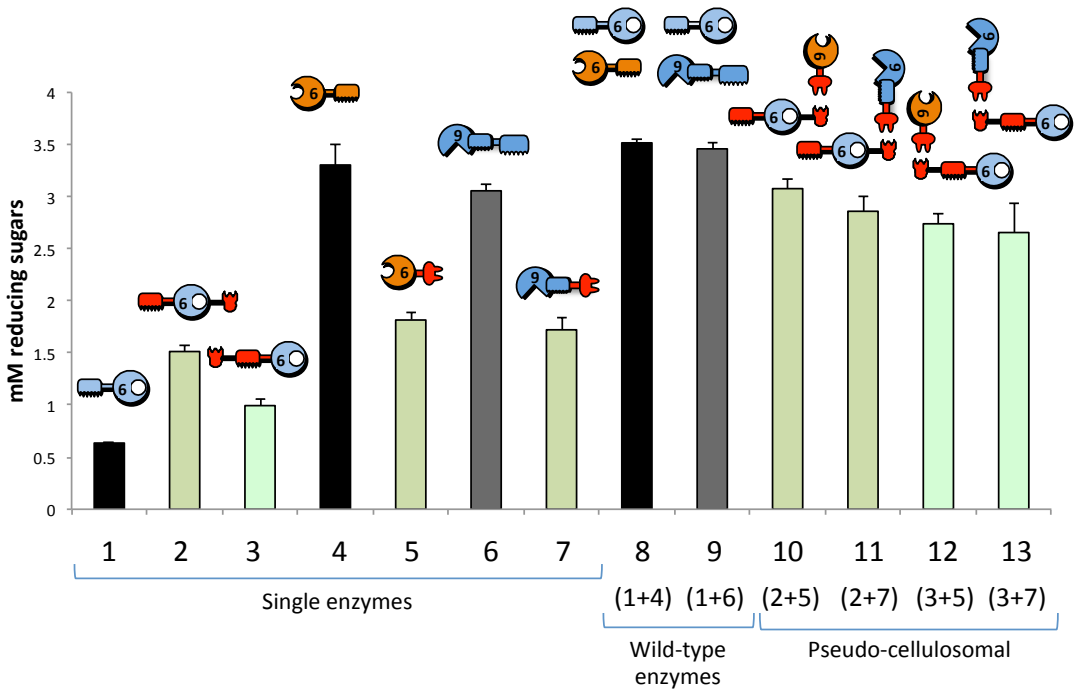
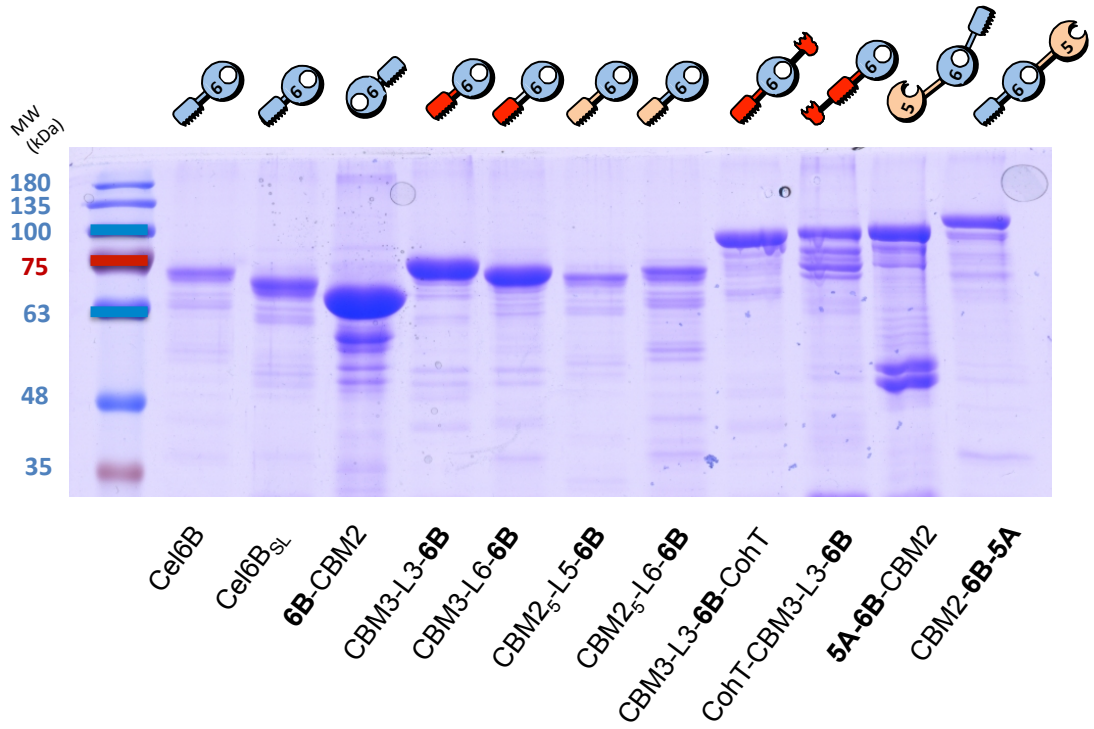
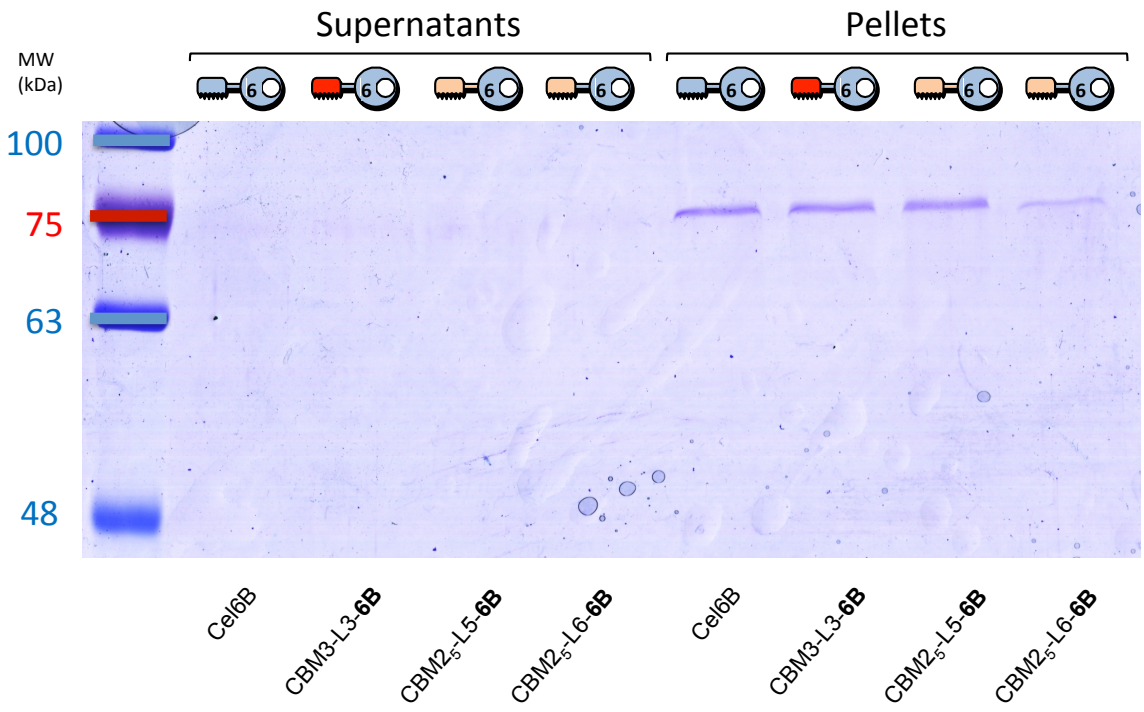


Figure 4. Comparative enzymatic activities on Avicel of wild-type Cel6B exoglucanase with wild-type endoglucanases Cel9A or Cel6A, compared to pseudo-cellulosomes. Lanes: (1) wild-type Cel6B, (2) CBM3-L3-**6B**-CohT, (3) CohT-CBM3-L3-**6B**, (4) wild-type Cel6A (5) **6A**-*t*, (6) wild-type Cel9A, (7) **9A**-*t*, (8) Cel6A + Cel6B (wild-type enzymes), (9) Cel6B + Cel9A (wild-type enzymes), (10) pseudo-cellulosome: **6A**-*t* bound to CBM3-L3-**6B**-CohT, (11) pseudo-cellulosome: **9A**-*t* bound to CBM3-L3-**6B**-CohT, (12) pseudo-cellulosome: **6A**-*t* bound to CohT-CBM3-L3-**6B** and (13) pseudo-cellulosome: **9A**-*t* bound to CohT-CBM3-L3-**6B**. Enzymatic activity is defined as mM total reducing sugars released from Avicel following A 21-h reaction period at 50°C. Each reaction was performed in triplicate, and standard deviations are indicated.

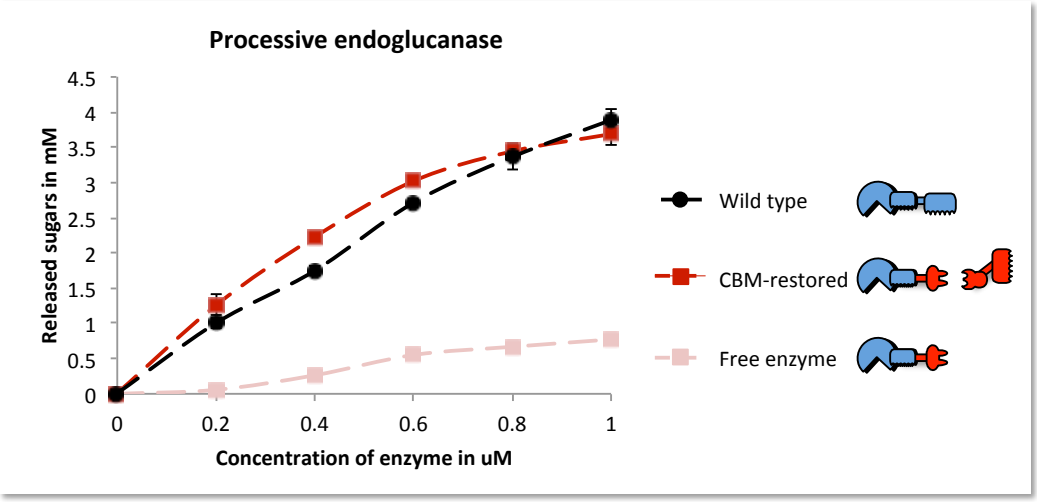
Supplemental Figure 1



Supplemental Figure 2



Supplemental Figure 3



Supplemental Figure 4