MR. NICHOLAS S SARAI (Orcid ID : 0000-0002-4655-0038)
DR YANNICK J BOMBLE (Orcid ID : 0000-0001-7624-8000)

Received Date : 29-Aug-2019
Revised Date : 06-Jan-2020
Accepted Date : 14-Feb-2020

Color : Fig 1-9 SuppInfo : 3 Table

Glycosylation of hyperthermostable designer cellulosome components yields enhanced stability and cellulose hydrolysis

Running Title: Glycosylation improves designer cellulosomes

Amaranta Kahn[a], Sarah Moraïs[a, b], Daehwan Chung[c], Nicholas S. Sarai[c,*], Neal N. Hengge[c], Audrey Kahn [a], Michael E. Himmel[c], Edward A. Bayer[a†], and Yannick J. Bomble[c†]

[a]*Department of Biomolecular Sciences, The Weizmann Institute of Science, Rehovot 7610001 Israel;*
[b]*Faculty of Natural Sciences, Ben-Gurion University of the Negev, Beer-Sheva 8499000 Israel;*
[c]*Biosciences Center, National Renewable Energy Laboratory, Golden, Colorado 80401, United States of America.*
[*]Present Address: *Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, United States of America*
amaranta.kahn@weizmann.ac.il; morais@post.bgu.ac.il; daewhan.chung@nrel.gov; nsarai@caltech.edu; Neal.Hengge@nrel.gov; kahn.audrey@gmail.com; mike.himmel@nrel.gov; ed.bayer@weizmann.ac.il; yannick.bomble@nrel.gov

†To whom correspondence should be addressed:
Yannick J. Bomble, Tel: 1-303-384-7729, Email: yannick.bomble@nrel.gov;

Edward A. Bayer, Tel: (+972)-8-934-2373, Email: ed.bayer@weizmann.ac.il

Key words: Cellulosome, Cellulase, thermostability, scaffoldin, glycosylation, expression host, *Caldicellulosiruptor bescii*

Abbreviations

CBM, carbohydrate-binding module; Coh, cohesin; DC, designer cellulosome; Doc, dockerin; DNS, dinitrosalicylic acid; Scaf, scaffoldin; GH, glycoside hydrolase; *lk*, linker peptide

Abstract

Biomass deconstruction remains integral for enabling second-generation biofuel production at scale. However, several steps necessary to achieve significant solubilization of biomass, notably harsh pretreatment conditions, impose economic barriers to commercialization. By employing hyperthermostable cellulase machinery, biomass deconstruction can be made more efficient, leading to milder pretreatment conditions and ultimately lower production costs. The hyperthermophilic bacterium *Caldicellulosiruptor bescii* produces extremely active hyperthermostable cellulases, including the hyperactive multifunctional cellulase *Cb*Cel9A/Cel48A. Recombinant *Cb*Cel9A/Cel48A components have been previously produced in *Escherichia coli* and integrated into synthetic hyperthermophilic designer cellulosome complexes. Since then, glycosylation has been shown to be vital for the high activity and stability of *Cb*Cel9A/Cel48A. Here, we studied the impact of glycosylation on a hyperthermostable designer cellulosome system in which two of the cellulosomal components, the scaffoldin and the GH9 domain of *Cb*Cel9A/Cel48A, were glycosylated as a consequence of employing *Ca. bescii* as an expression host. Inclusion of the glycosylated components yielded an active cellulosome system that exhibited long-term stability at 75°C. The resulting glycosylated designer cellulosomes showed significantly greater synergistic activity compared to the enzymatic components alone, as well as higher thermostability than the analogous nonglycosylated designer cellulosomes. These results indicate that glycosylation can be used as an essential engineering tool to improve the properties of designer cellulosomes. Additionally, *Ca. bescii* was shown to be an attractive candidate for production of glycosylated designer cellulosome components, which may further promote the viability of this bacterium both as a cellulase expression host and as a potential consolidated bioprocessing platform organism.

Introduction

Second generation biofuels, resulting from the deconstruction of cellulosic biomass into simple sugars followed by their fermentation to various biofuels, is among the most promising alternatives to replace fossil fuels, given that cellulosic biomass is the most abundant source of renewable organic matter on Earth [1-6]. However, its complex structure and recalcitrance to enzymatic deconstruction remain a primary barriers for conversion of lignocellulosic biomass to

biofuels [7]. Cellulosic substrates can be degraded by enzyme formulations containing three different types of complementary cellulases: endoglucanases that cleave internal glycosidic bonds of the cellulose chain, exoglucanases that hydrolyze cellobiose units "processively" from one or both ends of the cellulose chain, and β-glucosidases that hydrolyze cellobiose into two molecules of glucose [8-10]. Members of a sub-group of endoglucanases, termed processive endoglucanases, also cleave the cellulose chain processively by hydrolyzing internal cellulose chains [11].

Cellulases are found primarily in bacteria and fungi, either in a "free state", or in multi-modular complexes, called cellulosomes [12-16]. Cellulosome architecture is based on a non-catalytic subunit, the scaffoldin, that possess multiple cohesin modules to which the enzymes can bind. The enzymes themselves contain dockerin modules that bind strongly and selectively to the scaffoldin-borne cohesins. The scaffoldin can also contain a carbohydrate binding module (CBM) that mediates the binding of the entire complex to the cellulosic substrate [17].

Designer Cellulosomes (DCs) are artificial cellulosomes based on the species specificity of the cohesin-dockerin pairs. Chimeric scaffoldins are constructed by synthetic biology to contain a CBM and different cohesins of divergent specificities, usually derived from different species. Each chimeric enzyme contains an appended dockerin that matches the specificity of one of the cohesins on the chimeric scaffoldin [18]. DCs mimic the actions of native cellulosomes and allow control of the composition and position of the resident enzymes. DCs are efficient nanomachines, capable of a "tuned deconstruction" of cellulosic substrates. On the other hand, some anaerobic bacteria possess a collection of multi-modular enzymes that contain more than one catalytic subunit and several CBMs. Among them, the hyperthermophilic bacterium *Caldicellulosiruptor bescii* produces the multi-modular cellulase *Cb*Cel9A/Cel48A, which is considered to be one of the most active cellulases known to degrade biomass substrates [19-26]. *Cb*Cel9A/Cel48A, also known as CelA, is composed of an N-terminal GH9-CBM3c processive endoglucanase, two cellulose-binding CBM3b modules and a C-terminal GH48 exoglucanase. These components of *Cb*Cel9A/Cel48A are connected by Pro/Thr-rich linker peptides [27-30]. The reported hyperthermostability of the *Ca. bescii* enzymes (up to 90°C) renders these enzymes attractive candidates for industrial deconstruction of cellulosic substrates. Indeed, they are more suited to surviving the harsh conditions from chemical biomass pretreatments than most hydrolytic enzymes [31-35]. Performing enzymatic deconstruction at elevated temperatures also reduces the risk of microbial contamination and raises the kinetics and specificity of the reaction leading to a more efficient process.

In our recent study [36], the two catalytic domains of *Cb*Cel9A/Cel48A were produced separately in *Escherichia coli* with appended dockerins, and the resulting nonglycosylated recombinant enzymes were integrated into a hyperthermostable DC, together with a dockerin-bearing *Ca. bescii* GH5 endoglucanase. The resulting hyperthermostable DC exhibited activity on a microcrystalline cellulose substrate (Avicel) for 72 h at 75 °C. The three cohesin-dockerin pairs, selected in the previous work, originated from the following thermophilic or hyperthermophilic microbes: *Archeoglobus fulgidus (G,g)*, *Clostridium thermocellum* (*T,t*) and *Clostridium clariflavum* (*V,v*). The length of the linker peptide between the GH9-CBM3c and its dockerin was found to influence the overall activity of the complex, with the longer linker serving to increase the activity of the DC. When produced in its host organism and not in *E. coli*, this linker was found to be highly glycosylated [30, 37].

Glycosylation has been extensively studied and is fairly well understood in eukaryotes but has not been as extensively studied in prokaryotes. Glycosylation is known to affect protein folding, protein solubility and thermostability, protection against proteolysis and catalytic performance [30, 38-40]. Glycosylation of *Cb*Cel9A/Cel48A and its derivatives has indeed been found to enhance cellulolytic activity, to reduce nonproductive substrate binding, and to increase the thermostability of the intact enzyme and its truncated derivatives [30, 37]. The authors described this glycosylation as consisting primarily of *O*-glycosylation with an unusual homogeneous α-1,2-galactose disaccharide pattern – responsible for ~10% of the enzyme's molecular weight. Intriguingly, *O*-linked α-1,2-galactose-containing oligosaccharides were also described decorating the linkers of cellulosomal scaffoldin proteins from *Cl. thermocellum* and *Bacteroides cellulosolvens* [41-44] as well as those produced by *Clostridium cellulolyticum* [45]. However, the impact of glycosylation in these cases was not fully investigated.

In the current study, we examined the influence of glycosylation on DC complexes in terms of activity, synergy, and stability at high temperatures. To do so, *Ca. bescii* was used as an expression host to produce the linkers containing components from the most active DCs, determined in our previous study [36]. Dockerin-bearing GH9-CBM3c and GH48 from *Cb*Cel9A/Cel48A and the chimeric scaffoldin were successfully expressed in *Ca. bescii* but the dockerin bearing GH5 enzyme failed to express. The appended GH9-based dockerin was shown to actively bind to the recombinant scaffoldin, whereas the dockerin with the recombinant GH48 enzyme failed to bind. Thus, nonglycosylated dockerin-bearing versions of GH48 and GH5 were expressed in *E. coli* with minimal linkers. The latter enzymes successfully complexed, together

with the glycosylated form of the GH9 enzyme, to the glycosylated scaffoldin. The resulting partially-glycosylated complex showed enhanced cellulolytic activity on crystalline cellulose as well as enhanced thermal stability.

Results

Heterologous expression and purification of designer cellulosome (DC) components in *Ca. bescii*

The recombinant proteins used in this section are represented schematically in Figure 1. To investigate the impact of glycosylation on the activity and thermostability of the hyperthermostable DCs, we expressed and purified three proteins used previously to construct hyperthermostable DCs in *Ca. bescii* [36]: *Cb***Scaf***GTV*, *Cb***GH9**-*lk*-*v,* and *Cb***GH48**-*lk*-*t*. These three proteins possess long linkers, rich in prolines and threonines, originating either from *Ca. bescii* for the enzymes or from *Cl. thermocellum* and *Ca. bescii* for the scaffoldins. The scaffoldin was modified from the previous study [36], whereby the *Cl. thermocellum* CBM3a was replaced by an analogous CBM3b derived from *Ca. bescii Cb*Cel9A/Cel48A and positioned between a *Cl. thermocellum* (Coh*T*) and a cohesin (Coh*V*) from *Cl. clariflavum*, which mimics the wild-type CbCel9A/Cel48A architecture (Figure 1).

Three *Ca. bescii* expression vectors and expression strains were constructed, and successfully used for the expression and purification of proteins employed in assembly of hyperthermostable DCs (Figure 2A, B, Figure 3A). We hypothesized that the production of these enzymes and the scaffoldin in a proper expression host where proteins are glycosylated would provide enhanced activity and stability to the proteins at higher temperatures [30, 37, 46]. The design of the *Ca. bescii* expression vectors was based on those cloned in *E. coli,* and the corresponding genes for the desired proteins were inserted into plasmids pDCYB038 and pDCYB134 which contain an expression cassette for *Cb*Cel9A/Cel48A that was used successfully for homologous expression of *Cb*Cel9A/Cel48A in *Ca. bescii* [28, 37]. This expression cassette possesses the regulatory region of the Cbes2303, rho-independent terminator, as well as a C-terminal 6X histidine-tag and a stop codon. The plasmid DNA, which possesses a thermostable Kan^R antibiotic-resistant gene cassette for selection [47], was transformed into JWCB029 *(ΔpyrFA Δldh::ISCbe4 Δcbe1 ΔcelA*) [29]. Transformants were selected for kanamycin resistance. The existence of expression vectors in *Ca. bescii* transformants were verified by PCR amplification

using the extracted total DNA from transformants as a template. The primers used for construction of the expression vectors can be found in Table S1 and Table S2. The primers used for verification of these vectors can be found in Table S3. Note that the recombinant *Cb*GH5-*g* plasmid was constructed as well, however, the expression levels for this transformant were almost inexistent. As glycosylation is expected only on the linker peptides between folded domains and given that no linker is contained in *Cb*GH5-*g*, we decided to use the *Ec*GH5-*g* for the study instead.

Properties of the chimeric cellulosomal *Ca. bescii*-expressed proteins

Expression, purification, and validation of glycosylation of the proteins. SDS-PAGE analysis (Figure 2C and Figure 3B) demonstrated the expression, secretion, and purification of the extracellular proteins, including *Cb***Scaf***GTV*, *Cb***GH9**-*lk*-*v*, and *Cb***GH48**-*lk*-*t*, respectively. *Cb***GH9**-*lk*-*v* showed higher homogeneity than *Ec***GH9**-*lk*-*v* (lanes 5-6 respectively). *Ec***GH9**-*lk*-*v* was previously shown to display several bands on the gel [36], probably due to some protein degradation/cleavage as is often observed following the addition of a dockerin and linker to an enzyme. The improvement of protein stability during protein purification and storage might be due to glycosylation in *Cb***GH9**-*lk*-*v.* The highest band corresponds to the full-size protein (85 kDa). The *Cb***GH9**-*lk*-*v* enzyme migrated slightly higher than *Ec***GH9**-*lk*-*v,* likely due to the glycosylation. The *Cb***GH48**-*lk*-*t* construct exhibits a similar size and expression/purification profile when compared to *Ec***GH48**-*lk*-*t,* resulting in a single homologous band in SDS-PAGE gel analysis (Figure 3). The *Cb***Scaf***GTV* scaffoldin shows several bands on the gel, which likely corresponds to a mixture of impurities and cleavage of the scaffoldin. The use of a glycosylation staining kit (Figure 2 C, D and Figure 3C), confirmed that the three *Ca. bescii* expressed proteins were glycosylated, whereas no glycosylation was observed on the same proteins expressed in *E. coli*. Examination of *Cb***Scaf***GTV* revealed no glycosylation in the unbound fraction (i.e., after affinity pull down assay), which indicates that the observed bands probably consisted of impurities. In the bound fractions, two of the bands appeared to be strongly glycosylated. One band that matches the expected size of the full-length *Cb***Scaf***GTV* (71 kDa) and another smaller band that probably corresponds to the scaffoldin with a missing cohesin or portion thereof.

Functional binding of the scaffoldin to cellulose. In order to test whether *Cb***Scaf***GTV* is able to bind to cellulose, an affinity pull down experiment was conducted using the model cellulosic substrate, Avicel. After interaction with Avicel, the bound fraction (Figure 2C and D, Lane 1) was separated from the unbound fraction (Lane 2), washed and analyzed via SDS-PAGE.

We observed that the protein bands of *Cb***Scaf***GTV* bound to Avicel as they are present in the bound fraction.

Functional binding of the scaffoldin and enzymes. An ELISA approach served to demonstrate that the dockerin of the *Cb***GH9**-*lk*-*v* bound to Coh*V* (Figure 4A). Similarly, the specific binding properties of each cohesin of the *Cb***Scaf***GTV* to their matching dockerin was also shown to be functional (Figure 4C). However, the dockerin of *Cb***GH48**-*lk*-*t* did not bind to Coh*T* (Figure 4B). This lack of interaction might be explained by steric hindrance, due to possible glycosylation of *Cb***GH48**-*lk*-*t* in or around the dockerin binding site. Hence, instead of *Cb***GH48**-*lk*-*t,* we used the nonglycosylated GH48-*t* expressed in *E. coli.* In our previous study [36], *Ec***GH48**-*t* (with a short linker segment) was shown to possess a functional dockerin, to have the same impact on cellulose degradation as *Ec***GH48**-*lk*-*t* when incorporated into DCs, and to be more efficiently expressed and purified than *Ec***GH48**-*lk*-*t*. Hence, instead of *Cb***GH48**-*lk*-*t,* we used the nonglycosylated GH48-*t* expressed in *E. coli.*

Cellulosomal complex formation. DCs were constructed by self-assembly of stoichiometric amounts of the four components, i.e., the scaffoldin and the three enzymes: *Cb***Scaf***GTV*, *Cb***GH9**-*lk*-*v*, *Ec***GH48**-*t* and *Ec***GH5**-*g.* Two of these enzymes were produced in *E. coli*: *Ec***GH48**-*t,* because Doc*t* is not functional when produced in *Ca. bescii* and *Ec***GH5**-*g,* because of the absence of a long linker, implying that no glycosylation would occur on this construct (or chimeric protein). For the purposes of the current work, this DC will be referred to as *Ca. bescii* Designer Cellulosome (*Cb*DC), even though two of its enzymes were produced in *E. coli*. Conversely, *E. coli*-derived Designer Cellulosome (*Ec*DC) comprises proteins exclusively produced in *E. coli*, implying that all of these proteins, including *Ec***GH9**-*lk*-*v* and *Ec***Scaf***GTV,* are not glycosylated.

Stoichiometric molar ratios of each of the dockerin-bearing enzymes to the trivalent scaffoldin were determined by non-denaturing PAGE assay (see *Ec***GH48**-*t* : *Cb***Scaf***GTV* on Figure 5) and/or an affinity pull-down approach. Non-denaturing PAGE clearly indicated the most effective molar ratio for the glycosylated scaffoldin and the *E. coli*-produced nonglycosylated enzymes based on the range of ratios examined (see *Ec***GH48**-*t* : *Cb***Scaf***GTV* on Figure 5). Affinity pull-down experiments (Figure 6) showed that both *Cb*DC and *Ec*DC were properly assembled and bound to Avicel (Lanes 7 and 9). Much weaker bands were found in the unbound fraction (Lanes 8 and 10), indicating that near-stoichiometric molar ratio was properly determined. The free enzymes without scaffoldin were assayed as well. The majority of the free enzyme

fraction was found in the unbound state (lane 12 and 14), emphasizing the role of the scaffoldin in binding to the substrate. Lower amounts of free enzymes could also be found in the bound fractions (lane 11 and 13), probably reflecting binding of the active site of the enzymes to Avicel.

Glycosylation impacts the hydrolytic activity of the designer cellulosomes

*Cb*DC and *Ec*DC were compared for activity on a microcrystalline cellulosic substrate (Avicel) for 24 to 72 h at 75°C (Figure 7). Comparisons to the activities of the free enzymes were also made. The *Cb*DC exhibited the highest activity by degrading 1.5-fold more cellulose than *Ec*DC. The activity of *Cb*DC was improved over the free uncomplexed enzymes, whereas the *Ec*DC did not show any improvement over the free enzyme preparation, thus implying that glycosylation of the *Cb*DC components impacts its hydrolytic activity. Furthermore, without the presence of the glycosylated scaffoldin, the *Ca. bescii*-expressed free enzymes did not show significant improvement in activity over their *E. coli* expressed counterparts.

To test whether glycosylation of both the scaffoldin and the **GH9**-*lk*-*v* is needed to improve the activity of the DCs, we compared the activities of both *Cb*DC and *Ec*DC to the activities of mixed DC containing only one of the glycosylated components (Figure 8). We observed that full activity was dependent on the presence of both glycosylated scaffoldin and the **GH9**-*lk*-*v*. The presence of the glycosylated **GH9**-*lk*-*v* in a mixed complex alone provided some enhancement of activity, but the mixed complex with only the glycosylated scaffoldin failed to elicit enhanced activity.

Overall complex stability is improved by glycosylation

To test whether protein glycosylation affects the thermostability of the DC, we subjected glycosylated and nonglycosylated monovalent DCs to heat treatment at 85°C for 3 h. The monovalent DCs contained the GH9 enzyme and the scaffoldin in stoichiometric amounts and the complexes were tested for residual activity on Avicel for 16 h at 75°C (Figure 9). After 30 min of exposure to 85°C, the glycosylated complex exhibited a 2.2-fold higher activity than that of the nonglycosylated complex and 2.8-fold higher after 3 h. The *E. coli*-based complex lost more than half of its activity after 30 min of incubation, while the *Ca. bescii*-based complex still retained ~ 65% of its initial activity after 3 h of incubation. These results suggest that glycosylation of the complexed proteins dramatically improves the thermostability of the complex.

Discussion

The quest for sustainable, renewable energy sources and waste management systems is ongoing [5, 7, 48-50]. Deconstruction of cellulosic biomass for production of biofuels is one of the major avenues being studied for this purpose [51-53]. Due to its intrinsic recalcitrance, deconstruction of lignocellulosic biomass remains a rate-limiting and expensive step. We recently described the assembly of a hyperthermostable cellulosomal system that could be beneficial in terms of cost, reduced contamination risks, substrate specificity, and enzymatic performance [36]. Our first attempts at producing hyperthermostable DCs indeed resulted in enhanced stability and cellulolytic activity compared both to previously designed, thermostable DCs and to the native *Cl. thermocellum* cellulosome at 75°C. However, the components of this hyperthermostable DC were not glycosylated, as is the case for both the native enzymatic components produced by their native host (*Ca. bescii)* [36] and the scaffoldin of the native *Cl. thermocellum* cellulosome [42, 44]. Additionally, the performance of the previously described hyperthermostable DC was examined at sub-optimal temperatures, i.e., about 5°C lower than the optimal growth temperature for *Ca. bescii* [32, 54]. The glycosylation of some of the *Ca. bescii* proteins, including linkers rich in Thr/Pro, was also investigated recently and found to play a vital role in their activity and robustness [30, 37, 55].

Glycosylation has been extensively characterized in fungi for its impact on hydrolytic enzyme activity, thermostability, resistance to proteolytic cleavage, and binding affinity [37, 56]. Glycosylation of Gram-positive bacterial enzymes has received less attention, except for some early research (1989-1993) on the *Cl. thermocellum* and *B. cellulosolvens* cellulosomal scaffoldins [41-44] and pathogenic bacteria [55]. *Ca. bescii Cb*Cel9A/Cel48A was first reported to be glycosylated in 2015 and was predicted to be composed of *O*-linked glycosylation in the linker region [37]. An *O*-glycosyltransferase in the *Ca. bescii* genome was subsequently discovered [55]. Its deletion resulted in complete elimination of extracellular protein glycosylation including *Cb*Cel9A/Cel48A, and induced severe growth defects on crystalline cellulose, implying that glycosylation of *Cb*Cel9A/Cel48A and other hydrolases plays an essential role in cellulose deconstruction by *Ca. bescii* [55]. Recently, Chung and coworkers [30] conducted a study where they discovered that *Cb*Cel9A/Cel48A is uniformly *O*-glycosylated on the Thr or Ser of its three linker peptides, primarily by α-1,2 galactose disaccharides, representing approximately 10% of the enzyme's molecular weight. They observed that glycosylation was distributed similarly on its truncated N and C termini. We therefore expressed DC elements that possess long linker peptides

in *Ca. bescii* in order to investigate the impact of glycosylation on the stability and performance of the DCs.

Here, we report the successful expression of cellulosomal proteins (i.e., containing a dockerin or a cohesin) in *Ca. bescii.* These proteins are challenging to clone and express, due to their chimeric nature (e.g., the chimeric scaffoldin is constructed of modules from four different microorganisms, including an archaeon). In addition, dockerin-containing proteins and large scaffoldins have been reported to be difficult to express in heterologous hosts [63, 64]. However, we succeeded in expressing two chimeric dockerin-containing enzymes and one chimeric scaffoldin, all of which were glycosylated. *Cb***GH9**-*lk*-*v* showed a better purification profile than did *Ec***GH9**-*lk*-*v* (Figure 2C), which is consistent with previous results that demonstrated that glycosylated CbCel9A/Cel48A and its truncated derivatives, produced in *Ca. bescii,* were better protected from protease cleavage [30].

Although glycosylation interfered with the interaction of *Cb***GH48**-*lk*-*t* with Coh*T*, *Cb***GH9**-*lk*-*v* bound to both Coh*V* and both the glycosylated and non-glycosylated scaffoldins. The decision to use the GH48 enzyme produced in *E. coli* is consistent with the results of Chung and coworkers, which showed that the hydrolytic activity of the **GH48**-CBM3b enzyme (derived from *Cb*Cel9A/Cel48A) was higher in the absence of glycosylation, suggesting that glycosylation could affect to the activity of the enzymes differently, depending on its mode of action (unpublished data). We decided not to express the **GH5**-*g* in *Ca. bescii* and to produce it in *E. coli,* as it contains only a very small linker peptide (3 amino-acids) with no expected glycosylation sites.

Therefore, we assembled a DC that contained a glycosylated GH9 and two non-glycosylated cellulases complexed on a glycosylated scaffoldin and tested its hydrolytic activity on Avicel. In addition to the 1.5-fold enhancement in activity, a synergistic effect was observed when enzymes were complexed on the scaffoldin but no synergy was observed without glycosylation (Figure 8).

The results further indicated that both proteins are required in their glycosylated form to achieve the highest activity. In this context, Chung and coworkers [30] modeled putative hydrophobic patches on the *Cb*Cel9A/Cel48A linkers, at predicted positions of glycosylation, that, in the absence of glycosylation could be more exposed and could induce aggregation. This phenomenon has been described earlier in the glycoengineering of pharmaceutical proteins [40], where protein stability can be increased by glycosylating specific residues. A study focusing on the effects of glycosylation on protein conformational transitions and stabilization suggested that

an increase in protein stability can be achieved by introducing glycans of suitable hydrophobicity at flexible peptide residues. This indicates that glycosylation could induce both stabilizing and non-stabilizing effects according to the nature of the glycosylation site and its hydrophobicity [65].

The observed increased thermostability upon glycosylation is consistent with many previous studies [30, 40, 66, 67]. Sola and Griebenow [40] reviewed studies conducted on many glycosylated pharmaceutical proteins and suggested that an increase in thermostability is one of the most fundamental alterations conferred by glycosylation on the biophysical properties of a protein. The authors reported that several researchers interpret enhanced thermostability as a decrease in overall structural dynamics, even in regions far from the glycosylation site, thus indicating a transfer of the local effect to other regions of the protein. We hypothesized that this could also be applied to our complex. In the present study, we have demonstrated that glycosylation of DC components can make a significant improvement both in the thermostability of the intact complex as well as its overall catalytic performance on recalcitrant cellulosic feedstocks. These enhanced properties may provide the improvement required to render DCs a cost-competitive component in the conversion of biomass to free sugars that could be fermented to liquid biofuels. Our results also promote the use of the hyperthermophilic bacterium, *Ca. bescii,* as a viable host for production of highly active and stable glycosylated components for assembly of hyperthermostable DCs.

Material and Methods

Cloning for heterologous protein production in *E. coli*

All *E. coli* plasmids used in this study have been described in Kahn and coworkers (2019), except for the *E. coli* plasmid (pSMAK137) used for expressing the *Cb***Scaf***GTV* protein. Cloning was conducted using the plasmid *Ec***Scaf***GTV* [68], as a template in pET9d, by inserting the CBM3b of *Ca. bescii* cloned from the genomic DNA of *Ca. bescii* (DSM 6725) and removing the CBM3a of *Cl. thermocellum* following the restriction free method [69]. Primers are listed in Table S1. PCR was performed with Q5 High-Fidelity DNA polymerase (New England BioLabs). Competent *E. coli* XL1 cells were used for plasmid maintenance and production.

Cloning for heterologous protein production in *Ca. bescii*

Each plasmid was constructed by cloning the amplified opening reading frames (or part/s thereof) from the previously constructed *E. coli* plasmids (see above section).

Polymerase chain reaction (PCR) used in this study were performed using Q5 High-Fidelity DNA polymerase (New England BioLabs). Generated linear fragments were assembled using the HiFi DNA Assembly Cloning Kit (New England BioLabs) according to the manufacturer's instruction. Restriction sites in the inserted fragments were used to verify the presence of the inserted site in the vector.

Plasmid pDCYB137 (Figure 2A. Table S1) was constructed by inserting the open reading frame (ORF) of the gene for C*b***Scaf***GTV* into modified pDCYB038 (*Ca. bescii* expression vector of CelA N-terminus (GH9-CBM3- CBM3))[30], which contains the regulatory region of the *Ca. bescii* ribosomal protein S30EA (Cbes_2105) to express the thermostable kanamycin resistance gene [47] and a Rho-independent transcription terminator. The 8.074 kb DNA backbone was amplified with primers DCB413 and DCB414. The 1.959 kb DNA fragment used as an insert was amplified with primers DCB415 and DCB417, containing an AvaI and a XhoI restriction site. These two linear fragments were assembled to construct pDCYB137. Plasmid pDCYB144 (Table S1, Figure 3) was constructed by inserting the ORF of **GH48**-*lk*-*t* into modified pDCYB038 [30], which was used for construction of pDCYB137. The 6.913 kb DNA backbone was amplified with primers DCB402 and DCB403. The 2.313 kb DNA fragment used as an insert was amplified with primers DCB432 and DCB433 containing KpnI and AvaI restriction sites. These two linear fragments were assembled to construct pDCYB144. Plasmid pDCYB146 (Figure 2B. Table S1) was constructed by inserting ORF of **GH9**-*lk*-*v* into modified pDCYB038 [30], which was used for construction of pDCYB137. The 6.913 kb DNA backbone was amplified with primers DCB402 and DCB403. The 2.274 kb DNA fragment used as an insert was amplified with primers DCB436 and DCB403 containing BamHI and EcoRV restriction sites. These two linear fragments were assembled to construct pDCYB146. Primers used for plasmid construction and sequencing are listed in Table S1 and S2. *E. coli* strain DH5α cells were transformed by electroporation in a 2-mm-gap cuvette at 2.5 kV, and transformants were selected through apramycin resistance. Sequences of all plasmids were confirmed by automatic sequencing (Genewiz, NJ USA). The

*Cb*GH5-*g* plasmid was constructed in the same way; however, no protein expression could be detected (data not shown).

Plasmids pDCYB137, pDCYB144, and pDCYB146 were electro-transformed into JWCB029 (*ΔpyrFAΔldh::ISCbe4 Δcbe1ΔcelA*) cells as described by Chung and coworkers [37]. The cultures were electro-pulsed with plasmid DNA (0.5–1.0 μg), and placed for recovery in low osmolarity complex (LOC) growth medium [37, 70] at 75°C. Recovery cultures were directly plated onto solid LOC media supplemented with kanamycin (20 μg/mL). The isolated colonies were obtained after 4 d of incubation at 75°C in anaerobic chamber. Total DNA was isolated from transformants. PCR amplification using primers DCB008 and 010 (Table S2) located outside the inserted fragment on the plasmid was used to confirm the presence of plasmid with the correct sequence for the gene of interest.

Protein expression and purification

Heterologous protein production in *E. coli.* The proteins used in this study that were produced in an *E. coli* expression system were prepared following the procedure described in our previous study [36].

Heterologous protein production in *Ca. bescii.* Each relevant strain of *Ca. bescii* was inoculated from frozen stocks into serum bottles containing 20 mL of low osmolarity defined growth media (LOD) supplemented with kanamycin (20 μg/mL). Cultures were grown overnight at 65°C to reach mid-log phase. A 1 to 2% subculture was then inoculated into 100 mL of LOD from the 20 mL culture. These cultures were grown between 18 to 24 h to mid-log phase. The entire 100 mL of seed culture were then inoculated into 10-L fermentation vessels, containing LOD media with kanamycin. Cultures were agitated at 50 RPM and sparged continuously with N_2 at a flow rate of 0.5 L per min, while maintaining a pH of 6.8 (titrating with 2 N HCl and 10% w/v sodium bicarbonate) for 24 to 48 h. A constant temperature of 65°C was maintained. Each extracellular protein (ECP) was supplemented with a His-tag at the C-terminus to facilitate purification, during construction of the plasmids. Each ECP was collected, filtered, concentrated and buffer was exchanged through a GE tangential flow filtration 30 kDa molecular weight cut off from 75 mL to 150 mL in 10 mM Tris buffer, pH 7.4. Proteins of interest were submitted to His-Tag purification using an AKTA and a pre-packed nickel-NTA column. The proteins were

subjected to a second purification step using a Superdex® 26/60 200 PG column or a Superdex® 10/300 75 GL according to their size and concentrated to approximately 1 mg/mL. All proteins were assessed for purity according to their calculated molecular weight by SDS-PAGE and presence or absence of additional bands in the preparation. Concentration of the protein was measured by tryptophan absorbance at 280 nm, based on their extinction coefficient measured by the Protparam tool [71]. The proteins were subsequently stored in 50% (v/v) glycerol at -20°C.

Detection of protein glycosylation

Glycosylated proteins were visualized by staining with a Pierce™ Glycoprotein Staining Kit (Thermo Scientific, Rockford, IL) according to the manufacturer's instructions. Two gels were prepared simultaneously, with the same samples, for each assay: one for Glycoprotein staining and the other for Coomassie staining. All the samples were supplemented with dithiothreitol, boiled for 5 min and loaded on both gels. The positive control is a glycosylated protein (horseradish peroxidase) furnished with the kit.

Analysis of cohesin-dockerin interactions

Affinity-based ELISA was performed to display the specific interaction between the matching cohesin and dockerin modules after *Ca. bescii* expression according to Barak and coworkers [72]. Matching cohesin and dockerin modules produced in *E. coli* were previously assayed [36]. The modules produced in *Ca. bescii* were assayed using the same conditions. The scaffoldin and the enzyme containing the dockerin was coated with 0.1 M sodium carbonate coating buffer, pH 9, for 1 h at 37°C onto a MaxiSorp ELISA plates (Nunc A/S, Roskilde, Denmark). The interacting modules were treated with blocking buffer. Proteins containing a dockerin fused to a xylanase or a cohesin counterpart fused to a CBM, were added for interaction (diluted from 1 to 1000 ng/mL). After three washing steps with washing buffer, the primary antibody rabbit anti-xylanase or rabbit anti-CBM was added diluted to a concentration of 1:10,000 or 1:3000 respectively in blocking buffer. Following three washes, the secondary anti-rabbit horseradish peroxidase (HRP)-labeled antibody preparation (diluted 1:10,000) was added. The plates were finally washed four times and 3,3',5,5'-tetramethylbenzidine (TMB) + substrate-Chromogen (Dako, Glostrup Denmark) was added. Color formation was terminated by adding 1 M H_2SO_4 to each plate well a few min after the addition of TMB. The absorbance was measured at 450 nm using a microplate reader. All the

steps were conducted in 100 µL/well: we first discarded the previous solution and then added the next one and put it to interact for 30 min at 37°C. (All washing steps were conducted with washing buffer (Tris-buffered saline: 3 g/L Tris , 8 g/L NaCl , and 0.2 g/L KCl adjusted to pH 7.4), TBS, containing 10 mM $CaCl_2$, and 0.05% Tween 20); all dilution steps were conducted in blocking buffer (TBS, with added 10 mM $CaCl_2$, 0.05% Tween 20, and 2% BSA).

Each dockerin-bearing enzyme was attached to its matching cohesin on *Ec***Scaf***GTV* or on C*b***Scaf***GTV* following proper experimentally determined stoichiometric ratios. The ratio of the enzymes complexed to the *E. coli* produced scaffoldin was calculated in our previous study [36].

The molecular ratio of the *Ca. bescii* and *E. coli*-produced proteins, bound to the *Cb***Scaf***GTV,* was assayed similarly by non-denaturing PAGE and the affinity pull-down technique. Determination of the effective molar ratios was performed as follows: different ratios of pure proteins were prepared (200 pmol each in 50 mM acetate buffer, pH 5.0, 12 mM $CaCl_2$, and 2 mM EDTA) and incubated for 2 h at 37°C in the presence of 10% cellobiose (Sigma-Aldrich Chemical Co., St. Louis, MO). The disaccharide cellobiose binds to the catalytic modules of the enzymes and blocks the binding and action of the enzymes on the cellulosic substrate. The fractions were then gently mixed with microcrystalline cellulose (Avicel; FMC Biopolymer, Philadelphia, PA) for 1 h at 4°C. The tubes were centrifuged at 16,000×*g* for 2 min. The supernatant fluids (containing unbound proteins) were carefully removed and supplemented with SDS-containing buffer to a final volume of 60 µL. The pellets (containing bound proteins) were washed twice by resuspension in 200 µL of 50 mM acetate buffer supplemented with 0.05% Tween 20 to eliminate nonspecific binding. The samples were then centrifuged at 16,000×*g* for 2 min and re-suspended in 60 µL of SDS-containing buffer (New England BioLabs). The resultant unbound and bound fractions were each boiled for 10 min and then analyzed by SDS-PAGE using a 10% polyacrylamide gel. For the non-denaturing PAGE assay, the solutions containing the different protein ratios were loaded directly onto the gel and run for 2 to 3 h at 100 V. The same amount of protein was used for assembly of both the *E. coli* and *Ca. bescii* complexes for further activity assay.

Enzymatic assays

In Kahn and coworkers (2019), we determined the optimal concentration of the enzymes for subsequent activity assays [36]. We found that in order to use the enzymes in their linear range, a

0.05 μM concentration of the enzyme in solution was required. The same concentration was thus used in this study for all of the enzymes. DCs containing the glycosylated *Cb***GH9**-*lk*-*v* and the *Cb***Scaf***GTV* were termed *Cb*DCs even though the other two enzymes (*Ec***GH48**-*t* and *Ec***GH5**-*g)* were derived from *E. coli*. DCs containing the nonglycosylated *Ec***GH9**-*lk*-*v* and the *Ec***Scaf***GTV* were termed *Ec*DCs, where all of the components (the three enzymes and the scaffoldin) were derived from *E. coli*. The cellulose-degrading performance of equimolar amounts of all of the DCs were assayed in the same experiment with the two groups of free enzymes (i.e., the free *Ca. bescii* and the free *E. coli* enzyme preparations) on microcrystalline cellulose (Avicel, Sigma-Aldrich, St. Louis, MO) as the substrate (4% final concentration). An additional assay was performed in a similar manner to compare *Cb*DC and *Ec*DC, with the "mixed" DCs (i.e., the nonglycosylated *E. coli*-produced proteins complexed with the glycosylated *Cb***Scaf***GTV,* termed Mixed DC 1, and the glycosylated *Cb***GH9**-*lk*-*v* enzyme complexed with the other nonglycosylated *E. coli*-produced protein, termed Mixed DC 2). The latter DCs were prepared by incubation for 2 h at 37°C of the three enzymes and the scaffoldin in equimolar amounts to a final concentration of 0.05 μM per protein in the interaction buffer (TBS with added 10 mM $CaCl_2$ and 0.05% Tween 20). The assay was carried out for 24 h to 72 h at 75°C in 50 mM (final concentration) of acetate buffer, pH 5.5 using a New Brunswick Scientific (New Jersey, United States) shaker incubator controlled at 120 RPM. The enzymatic reactions were terminated by placing the reaction tubes in an ice bath and spinning them down at 17,000×g for 2 min at room temperature. In all assays, the substrate is in excess of the enzymes. Under those conditions, the highest activity that was measured (~12 mM of released sugar) represents about 5.5% degradation of the insoluble substrate.

For each activity assay, estimation of the released soluble sugar was calculated using the dinitrosalicylic acid (DNS) assay. A volume of 150 μL of DNS was added to 100 μL of each sample followed by 10 min incubation in a boiling water bath, as described previously [73, 74]. Serial dilutions of glucose were used as a standard to determine released sugar. The absorbance of each sample (100 μL of each) was measured in micro-well plates at 540 nm, using a micro plate reader. Assays were performed three times in triplicate.

Denaturation-residual activity assays

Residual activities of nonglycosylated *Ec***GH9**-*lk-v* complexed on *Ec***Scaf***GTV* versus glycosylated *Cb***GH9**-*lk-v* complexed on the *Cb***Scaf***GTV* were examined at 75°C after incubation at 85°C on Avicel. The complexes of the *Ca. bescii*-produced protein and of the *E. coli*-produced protein were formed as described above. Each preparation was then aliquoted into PCR tubes and incubated in a Biometra (BioSciences, Germany) thermocycler block, heated at 85°C for periods of between 30 min to 3 h. The 3-h tubes were added first, followed by the 2-h, 1-h and 30-min tubes. Zero-time incubation was carried out at room temperature. All tubes (including the zero-time) were then placed on ice to terminate the heat treatment and transferred to test tubes for activity assay on Avicel. The Avicel enzymatic assay was conducted as described above, three times in triplicate.

References

1. Abo, B. O., Gao, M., Wang, Y., Wu, C., Ma, H. & Wang, Q. (2019) Lignocellulosic biomass for bioethanol: an overview on pretreatment, hydrolysis and fermentation processes, Rev Environ Health. 34, 57-68.
2. Aditiya, H. B., Chong, W. T., Mahlia, T. M., Sebayang, A. H., Berawi, M. A. & Nur, H. (2016) Second generation bioethanol potential from selected Malaysia's biodiversity biomasses: A review, Waste Manag. 47, 46-61.
3. Wilson, D. B. (2012) Processive and nonprocessive cellulases for biofuel production--lessons from bacterial genomes and structural analysis, Appl Microbiol Biotechnol. 93, 497-502.
4. Ho, D. P., Ngo, H. H. & Guo, W. (2014) A mini review on renewable sources for biofuel, Bioresour Technol. 169, 742-749.
5. Lynd, L. R., Laser, M. S., Bransby, D., Dale, B. E., Davison, B., Hamilton, R., Himmel, M., Keller, M., McMillan, J. D., Sheehan, J. & Wyman, C. E. (2008) How biotech can transform biofuels, Nat Biotechnol. 26, 169-72.
6. Bayer, E. A., Lamed, R. & Himmel, M. E. (2007) The potential of cellulases and cellulosomes for cellulosic waste management, Curr Opin Biotechnol. 18, 237-45.
7. Himmel, M. E., Ding, S. Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W. & Foust, T. D. (2007) Biomass recalcitrance: engineering plants and enzymes for biofuels production, Science. 315, 804-7.

8. Sakon, J., Irwin, D., Wilson, D. B. & Karplus, P. A. (1997) Structure and mechanism of endo/exocellulase E4 from Thermomonospora fusca, Nat Struct Biol. 4, 810-8.
9. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics, Nucleic Acids Res. 37, D233-8.
10. Bayer, E. A., Shoham, Y. & Lamed, R. (2006) Cellulose-Decomposing Bacteria and Their Enzyme Systems in Prokaryote pp. 578-617, New York: Springer-Verlag, http://link.springer.de/link/service/books/10125/index.htm;.
11. Wilson, D. B. & Kostylev, M. (2012) Cellulase processivity, Methods Mol Biol. 908, 93-9.
12. Bayer, E. A., Kenig, R. & Lamed, R. (1983) Adherence of Clostridium thermocellum to cellulose, J Bacteriol. 156, 818-27.
13. Lamed, R., Setter, E. & Bayer, E. A. (1983) Characterization of a cellulose-binding, cellulase-containing complex in Clostridium thermocellum, J Bacteriol. 156, 828-36.
14. Lamed, R., Setter-Lamed, E., Kenig, R. & Bayer, E. A. (1983) The cellulosome - a discrete cell surface organelle of Clostridium thermocellum which exhibits separate antigenic, cellulose-binding and various cellulolytic activities., biotechnol bioeng Symp. 13, 163-181.
15. Shoham, Y., Lamed, R. & Bayer, E. A. (1999) The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides, Trends Microbiol. 7, 275-81.
16. Artzi, L., Bayer, E. A. & Morais, S. (2017) Cellulosomes: bacterial nanomachines for dismantling plant polysaccharides, Nat Rev Microbiol. 15, 83-95.
17. Bayer, E. A., Belaich, J. P., Shoham, Y. & Lamed, R. (2004) The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides, Annu Rev Microbiol. 58, 521-54.
18. Gunnoo, M., Cazade, P. A., Galera-Prat, A., Nash, M. A., Czjzek, M., Cieplak, M., Alvarez, B., Aguilar, M., Karpol, A., Gaub, H., Carrion-Vazquez, M., Bayer, E. A. & Thompson, D. (2016) Nanoscale Engineering of Designer Cellulosomes, Adv Mater. 28, 5619-47.
19. Yang, S. J., Kataeva, I., Wiegel, J., Yin, Y., Dam, P., Xu, Y., Westpheling, J. & Adams, M. W. (2010) Classification of 'Anaerocellum thermophilum' strain DSM 6725 as Caldicellulosiruptor bescii sp. nov, Int J Syst Evol Microbiol. 60, 2011-5.
20. Basen, M., Rhaesa, A. M., Kataeva, I., Prybol, C. J., Scott, I. M., Poole, F. L. & Adams, M. W. (2014) Degradation of high loads of crystalline cellulose and of unpretreated plant biomass by the thermophilic bacterium Caldicellulosiruptor bescii, Bioresour Technol. 152, 384-92.

21. Kanafusa-Shinkai, S., Wakayama, J., Tsukamoto, K., Hayashi, N., Miyazaki, Y., Ohmori, H., Tajima, K. & Yokoyama, H. (2013) Degradation of microcrystalline cellulose and non-pretreated plant biomass by a cell-free extracellular cellulase/hemicellulase system from the extreme thermophilic bacterium Caldicellulosiruptor bescii, J Biosci Bioeng. 115, 64-70.

22. Luthi, E., Jasmat, N. B. & Bergquist, P. L. (1990) Xylanase from the extremely thermophilic bacterium "Caldocellum saccharolyticum": overexpression of the gene in Escherichia coli and characterization of the gene product, Appl Environ Microbiol. 56, 2677-83.

23. Rainey, F. A., Donnison, A. M., Janssen, P. H., Saul, D., Rodrigo, A., Bergquist, P. L., Daniel, R. M., Stackebrandt, E. & Morgan, H. W. (1994) Description of Caldicellulosiruptor saccharolyticus gen. nov., sp. nov: an obligately anaerobic, extremely thermophilic, cellulolytic bacterium, FEMS Microbiol Lett. 120, 263-6.

24. Saul, D. J., Williams, L. C., Grayling, R. A., Chamley, L. W., Love, D. R. & Bergquist, P. L. (1990) celB, a gene coding for a bifunctional cellulase from the extreme thermophile "Caldocellum saccharolyticum", Appl Environ Microbiol. 56, 3117-24.

25. Te'o, V. S., Saul, D. J. & Bergquist, P. L. (1995) celA, another gene coding for a multidomain cellulase from the extreme thermophile Caldocellum saccharolyticum, Appl Microbiol Biotechnol. 43, 291-6.

26. Bayer, E. A., Shoham, Y. & Lamed, R. (2013) Lignocellulose-Decomposing Bacteria and Their Enzyme Systems. in The prokaryotes-Prokaryotic Physiology and Biochemistry (E, R., EF, D., S, L., Stackebrandt, E. & F, T., eds) pp. 215-265.

27. Brunecky, R., Donohoe, B. S., Yarbrough, J. M., Mittal, A., Scott, B. R., Ding, H., Taylor Ii, L. E., Russell, J. F., Chung, D., Westpheling, J., Teter, S. A., Himmel, M. E. & Bomble, Y. J. (2017) The Multi Domain Caldicellulosiruptor bescii CelA Cellulase Excels at the Hydrolysis of Crystalline Cellulose, Sci Rep. 7, 9622.

28. Brunecky, R., Alahuhta, M., Xu, Q., Donohoe, B. S., Crowley, M. F., Kataeva, I. A., Yang, S. J., Resch, M. G., Adams, M. W., Lunin, V. V., Himmel, M. E. & Bomble, Y. J. (2013) Revealing nature's cellulase diversity: the digestion mechanism of Caldicellulosiruptor bescii CelA, Science. 342, 1513-6.

29. Young, J., Chung, D., Bomble, Y. J., Himmel, M. E. & Westpheling, J. (2014) Deletion of Caldicellulosiruptor bescii CelA reveals its crucial role in the deconstruction of lignocellulosic biomass, Biotechnol Biofuels. 7, 142.

30. Chung, D., Sarai, N. S., Knott, B. C., Hengge, N., Russell, J., Yarbrough, J. M., Brunecky, R., Young, J., Supekar, N., Vander Wall, T., Sammond, D. W., Crowley, M. F., Szymanski, C. M., Wells, L., Azadi, P., Westpheling, J., Himmel, M. & Bomble, Y. J. (2019) Glycosylation is vital for industrial performance of hyper-active cellulases, ACS Sustainable Chemistry & Engineering.

31. Yeoman, C. J., Han, Y., Dodd, D., Schroeder, C. M., Mackie, R. I. & Cann, I. K. (2010) Thermostable enzymes as biocatalysts in the biofuel industry, Adv Appl Microbiol. 70, 1-55.

32. Yang, S. J., Kataeva, I., Hamilton-Brehm, S. D., Engle, N. L., Tschaplinski, T. J., Doeppke, C., Davis, M., Westpheling, J. & Adams, M. W. (2009) Efficient degradation of lignocellulosic plant biomass, without pretreatment, by the thermophilic anaerobe "Anaerocellum thermophilum" DSM 6725, Appl Environ Microbiol. 75, 4762-9.

33. Olson, D. G., Sparling, R. & Lynd, L. R. (2015) Ethanol production by engineered thermophiles, Curr Opin Biotechnol. 33, 130-41.

34. Turner, P., Mamo, G. & Karlsson, E. N. (2007) Potential and utilization of thermophiles and thermostable enzymes in biorefining, Microb Cell Fact. 6, 9.

35. Zeldes, B. M., Keller, M. W., Loder, A. J., Straub, C. T., Adams, M. W. & Kelly, R. M. (2015) Extremely thermophilic microorganisms as metabolic engineering platforms for production of fuels and industrial chemicals, Front Microbiol. 6, 1209.

36. Kahn, A., Morais, S., Galanopoulou, A. P., Chung, D., Sarai, N. S., Hengge, N., Hatzinikolaou, D. G., Himmel, M. E., Bomble, Y. J. & Bayer, E. A. (2019) Creation of a functional hyperthermostable designer cellulosome, Biotechnol Biofuels. 12, 44.

37. Chung, D., Young, J., Bomble, Y. J., Vander Wall, T. A., Groom, J., Himmel, M. E. & Westpheling, J. (2015) Homologous expression of the Caldicellulosiruptor bescii CelA reveals that the extracellular protein is glycosylated, PLoS One. 10, e0119508.

38. Varki, A. (2017) Biological roles of glycans, Glycobiology. 27, 3-49.

39. Nothaft, H. & Szymanski, C. M. (2010) Protein glycosylation in bacteria: sweeter than ever, Nat Rev Microbiol. 8, 765-78.

40. Sola, R. J. & Griebenow, K. (2009) Effects of glycosylation on the stability of protein pharmaceuticals, J Pharm Sci. 98, 1223-45.

41. Gerwig, G. J., Kamerling, J. P., Vliegenthart, J. F., Morag, E., Lamed, R. & Bayer, E. A. (1993) The nature of the carbohydrate-peptide linkage region in glycoproteins from the cellulosomes of Clostridium thermocellum and Bacteroides cellulosolvens, J Biol Chem. 268, 26956-60.

42. Gerwig, G. J., de Waard, P., Kamerling, J. P., Vliegenthart, J. F., Morgenstern, E., Lamed, R. & Bayer, E. A. (1989) Novel O-linked carbohydrate chains in the cellulase complex (cellulosome) of Clostridium thermocellum. 3-O-Methyl-N-acetylglucosamine as a constituent of a glycoprotein, J Biol Chem. 264, 1027-35.

43. Gerwig, G. J., Kamerling, J. P., Vliegenthart, J. F., Morag, E., Lamed, R. & Bayer, E. A. (1992) Novel oligosaccharide constituents of the cellulase complex of Bacteroides cellulosolvens, Eur J Biochem. 205, 799-808.

44. Gerwig, G. J., Kamerling, J. P., Vliegenthart, J. F., Morag, E., Lamed, R. & Bayer, E. A. (1991) Primary structure of O-linked carbohydrate chains in the cellulosome of different Clostridium thermocellum strains, Eur J Biochem. 196, 115-22.

45. Gal, L., Pages, S., Gaudin, C., Belaich, A., Reverbel-Leroy, C., Tardif, C. & Belaich, J. P. (1997) Characterization of the cellulolytic complex (cellulosome) produced by Clostridium cellulolyticum, Appl Environ Microbiol. 63, 903-9.

46. Brunecky, R., Chung, D., Sarai, N. S., Hengge, N., Russell, J. F., Young, J., Mittal, A., Pason, P., Vander Wall, T., Michener, W., Shollenberger, T., Westpheling, J., Himmel, M. E. & Bomble, Y. J. (2018) High activity CAZyme cassette for improving biomass degradation in thermophiles, Biotechnol Biofuels. 11, 22.

47. Lipscomb, G. L., Conway, J. M., Blumer-Schuette, S. E., Kelly, R. M. & Adams, M. W. W. (2016) A Highly Thermostable Kanamycin Resistance Marker Expands the Tool Kit for Genetic Manipulation of Caldicellulosiruptor bescii, Appl Environ Microbiol. 82, 4421-4428.

48. Kerr, R. A. (2007) Climate change. Global warming is changing the world, Science. 316, 188-90.

49. Kerr, R. A. (2013) Climate change. Humans fueled global warming millennia ago, Science. 342, 918.

50. Solomon, B. D. (2010) Biofuels and sustainability, Ann N Y Acad Sci. 1185, 119-34.

51. Schmer, M. R., Vogel, K. P., Mitchell, R. B. & Perrin, R. K. (2008) Net energy of cellulosic ethanol from switchgrass, Proc Natl Acad Sci U S A. 105, 464-9.

52. Gronenberg, L. S., Marcheschi, R. J. & Liao, J. C. (2013) Next generation biofuel engineering in prokaryotes, Curr Opin Chem Biol. 17, 462-71.

53. Jordan, D. B., Bowman, M. J., Braker, J. D., Dien, B. S., Hector, R. E., Lee, C. C., Mertens, J. A. & Wagschal, K. (2012) Plant cell walls to ethanol, Biochem J. 442, 241-52.

54. Blumer-Schuette, S. E., Kataeva, I., Westpheling, J., Adams, M. W. & Kelly, R. M. (2008) Extremely thermophilic microorganisms for biomass conversion: status and prospects, Curr Opin Biotechnol. 19, 210-7.
55. Russell, J., Kim, S. K., Duma, J., Nothaft, H., Himmel, M. E., Bomble, Y. J., Szymanski, C. M. & Westpheling, J. (2018) Deletion of a single glycosyltransferase in Caldicellulosiruptor bescii eliminates protein glycosylation and growth on crystalline cellulose, Biotechnol Biofuels. 11, 259.
56. Langsford, M. L., Gilkes, N. R., Singh, B., Moser, B., Miller, R. C., Jr., Warren, R. A. & Kilburn, D. G. (1987) Glycosylation of bacterial cellulases prevents proteolytic cleavage between functional domains, FEBS Lett. 225, 163-7.
57. Kim, S. K., Chung, D., Himmel, M. E., Bomble, Y. J. & Westpheling, J. (2017) Engineering the N-terminal end of CelA results in improved performance and growth of Caldicellulosiruptor bescii on crystalline cellulose, Biotechnol Bioeng. 114, 945-950.
58. Kim, S. K., Chung, D., Himmel, M. E., Bomble, Y. J. & Westpheling, J. (2019) Heterologous co-expression of two beta-glucanases and a cellobiose phosphorylase resulted in a significant increase in the cellulolytic activity of the Caldicellulosiruptor bescii exoproteome, J Ind Microbiol Biotechnol.
59. Kim, S. K., Chung, D., Himmel, M. E., Bomble, Y. J. & Westpheling, J. (2017) Heterologous expression of a beta-D-glucosidase in Caldicellulosiruptor bescii has a surprisingly modest effect on the activity of the exoproteome and growth on crystalline cellulose, J Ind Microbiol Biotechnol. 44, 1643-1651.
60. Kim, S. K., Chung, D., Himmel, M. E., Bomble, Y. J. & Westpheling, J. (2016) Heterologous expression of family 10 xylanases from Acidothermus cellulolyticus enhances the exoproteome of Caldicellulosiruptor bescii and growth on xylan substrates, Biotechnol Biofuels. 9, 176.
61. Chung, D., Young, J., Cha, M., Brunecky, R., Bomble, Y. J., Himmel, M. E. & Westpheling, J. (2015) Expression of the Acidothermus cellulolyticus E1 endoglucanase in Caldicellulosiruptor bescii enhances its ability to deconstruct crystalline cellulose, Biotechnol Biofuels. 8, 113.
62. Chung, D., Verbeke, T. J., Cross, K. L., Westpheling, J. & Elkins, J. G. (2015) Expression of a heat-stable NADPH-dependent alcohol dehydrogenase in Caldicellulosiruptor bescii results in furan aldehyde detoxification, Biotechnol Biofuels. 8, 102.
63. Stern, J., Morais, S., Lamed, R. & Bayer, E. A. (2016) Adaptor Scaffoldins: An Original Strategy for Extended Designer Cellulosomes, Inspired from Nature, MBio. 7, e00083.

64. Morais, S., Morag, E., Barak, Y., Goldman, D., Hadar, Y., Lamed, R., Shoham, Y., Wilson, D. B. & Bayer, E. A. (2012) Deconstruction of lignocellulose into soluble sugars by native and designer cellulosomes, MBio. 3.

65. Lu, D., Yang, C. & Liu, Z. (2012) How hydrophobicity and the glycosylation site of glycans affect protein folding and stability: a molecular dynamics simulation, J Phys Chem B. 116, 390-400.

66. Beckham, G. T., Dai, Z., Matthews, J. F., Momany, M., Payne, C. M., Adney, W. S., Baker, S. E. & Himmel, M. E. (2012) Harnessing glycosylation to improve cellulase activity, Curr Opin Biotechnol. 23, 338-45.

67. Chen, L., Drake, M. R., Resch, M. G., Greene, E. R., Himmel, M. E., Chaffey, P. K., Beckham, G. T. & Tan, Z. (2014) Specificity of O-glycosylation in enhancing the stability and cellulose binding affinity of Family 1 carbohydrate-binding modules, Proc Natl Acad Sci U S A. 111, 7612-7.

68. Morais, S., Stern, J., Kahn, A., Galanopoulou, A. P., Yoav, S., Shamshoum, M., Smith, M. A., Hatzinikolaou, D. G., Arnold, F. H. & Bayer, E. A. (2016) Enhancement of cellulosome-mediated deconstruction of cellulose by improving enzyme thermostability, Biotechnol Biofuels. 9, 164.

69. van den Ent, F. & Lowe, J. (2006) RF cloning: a restriction-free method for inserting target genes into plasmids, J Biochem Biophys Methods. 67, 67-74.

70. Farkas, J., Chung, D., Cha, M., Copeland, J., Grayeski, P. & Westpheling, J. (2013) Improved growth media and culture techniques for genetic analysis and assessment of biomass utilization by Caldicellulosiruptor bescii, J Ind Microbiol Biotechnol. 40, 41-9.

71. Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D. & Hochstrasser, D. F. (1999) Protein identification and analysis tools in the ExPASy server, Methods Mol Biol. 112, 531-52.

72. Barak, Y., Handelsman, T., Nakar, D., Mechaly, A., Lamed, R., Shoham, Y. & Bayer, E. A. (2005) Matching fusion protein systems for affinity analysis of two interacting families of proteins: the cohesin-dockerin interaction, J Mol Recognit. 18, 491-501.

73. Vazana, Y., Morais, S., Barak, Y., Lamed, R. & Bayer, E. A. (2012) Designer cellulosomes for enhanced hydrolysis of cellulosic substrates, Methods Enzymol. 510, 429-52.

74. Stern, J., Artzi, L., Morais, S., Fontes, C. & Bayer, E. A. (2017) Carbohydrate Depolymerization by Intricate Cellulosomal Systems, Methods Mol Biol. 1588, 93-116.

Figure legends

Figure 1. Schematic of the wild-type and chimeric proteins used in this study. The key defines the symbols and acronyms used for the protein modules, which are assembled into a cellulosomal complex. The bacterial or archaeal source of each cohesin and dockerin module is color-coded as follows: red, *Clostridium thermocellum*; purple, *Clostridium clariflavum*; brown, *Archaeoglobus fulgidus*. Upper case (*T, V* and *G*) and lower case (*t, v* and *g*) characters indicate the source (*Cl.* ***t****hermocellum*, *Cl. clarifla****v****um*, and *A. ful****g****idus,* respectively) of the source of the respective cohesin and dockerin modules. All catalytic modules originate from *Ca. bescii*, the numbers correspond to their GH family (GH9, GH48, GH5). Proteins containing glycosylation were expressed in *Caldicellulosiruptor bescii*, nonglycosylated proteins in *Escherichia coli*. Molecular weights were calculated on the basis of the amino acid sequence without taking into consideration the glycosylated moieties.

Figure 2. *Ca. bescii* expression system for glycosylated *CbScafGTV* and *Cb*GH9-*lk*-*v*.
Plasmid maps of **A**, *Cb***Scaf***GTV* (Coh*G*-linker-Coh*T*-linker-CBM3b-linker-Coh*V*), and **B**, *Cb*GH9-*lk*-*v* (*Ca. bescii* **GH9**-CBM3c-linker-Doc*v*), expressed in *Ca. bescii*. The multi-modular proteins were expressed with a signal peptide as well as a C- terminal 6X His-tag. As in the previous paper [37], the plasmid contained a Rho independent terminator, with a pBAS2 sequence for replication in *Ca. bescii* that is under the control of the *Ca. bescii* S-layer protein (Cbes_2303), and a thermostable Kan^R antibiotic resistant gene cassette was now included for selection. In addition, pSC101, the low-copy replication origin in *E. coli,* and *repA*, a plasmid-encoded gene required for pSC101 replication, are represented on the map. Panels **C** and **D** represent the characterization of the purified proteins by SDS-PAGE, stained either by Coomassie (**C**) or for glycoproteins (**D**). The scaffoldin and GH9 enzyme were expressed in either *Ca. bescii* or *E. coli*. Lane 1 represents the part of the original *Cb***Scaf***GTV* preparation (Lane 0) that binds to Avicel versus Lane 2 that shows proteins that failed to bind to Avicel. The positive control was a glycosylated protein (horseradish peroxidase) furnished with the glycoprotein-staining kit. Panel C and D were each replicated three times.

Figure 3. Expression of glycosylated CbGH48-lk-t in Ca. bescii.

A. Plasmid map of CbGH48-lk-t (GH48-linker-Doct) expressed in Ca. bescii expression system. This multi-modular protein was expressed using the same expression vector backbone, which was used for the construction of pDCYB137 and pDCYB146 (Figure 2). Panels B and C represent the characterization of the purified proteins, as determined by SDS-PAGE, stained either by Coomassie staining (B) or for glycoproteins (C). The positive and negative control are proteins supplied in the glycoprotein-staining kit. Panel B and C were conducted three times and showed the same results.

Figure 4. ELISA-based assay to confirm the dockerin-binding functionality of the chimaeric enzyme (A, B) and the cohesin-binding specificity of CbScafGTV (C). The procedure of Barak, et al. (2005) was followed for these experiments. In A and B, plates were coated with the target dockerin-bearing enzyme, and a monovalent scaffoldin (CtCBM3a linked to the indicated cohesin) was then introduced. In C, plates were coated with the scaffoldin, and xylanase fused to the indicated dockerin was then allowed to interact with the cohesins of CbScafGTV. Primary anti-rabbit anti-CBM or anti-xylanase antibody was then applied, followed by secondary goat anti-rabbit antibody conjugated to an horseradish peroxidase (HRP). Each reaction was conducted in triplicate, three times (n = 3); the data represent the mean ± SD.

Figure 5. Non-denaturing PAGE assay reveals the proper stoichiometric ratio of interaction between the chimaeric enzyme and its specific monovalent scaffoldin. Titration of the dockerin-bearing *Ec*GH48-lk-t with the *Cb*ScafGTV, using the estimated molar ratios of components, yielded precise experimental data that indicated that the functional stoichiometric ratio was effectively between 0.8 and 1.0 of the original *Cb*ScafGTV: *Ec*GH48-t preparations. The functional ratios of the other components of the system were determined in like fashion. Each stoichiometric ratio assay was conducted in triplicate, with the same result.

Figure 6. Affinity pull-down of *Cb*DC, *Ec*DC and the free enzymes. Equimolar concentrations of the three chimaeric enzymes and their matching scaffoldin (glycosylated and not glycosylated) were mixed then bound to Avicel substrate for 1 h. Parallel samples containing the same chimaeric enzymes without the scaffoldin were treated in like fashion. The bound (Lanes 7, 9, 11, 13) and

unbound (Lanes 8, 10, 12, 14) fractions were then boiled in sample buffer, centrifuged and subjected to denaturing SDS-PAGE, together with each of the protein sampless alone as controls (lane 1 to 6) without the Avicel binding step. The DC complexes (Lanes 7 to 10) and free enzymes (Lanes 11 to 14) were supplemented with excess cellobiose to avoid unspecific Avicel binding of the catalytic site of the enzymes. This affinity pull-down assay has been replicated three times and showed the same result.

Figure 7. Comparison of the hydrolytic activity of the glycosylated and nonglycosylated designer cellulosomes (*Cb*DC and *Ec*DC, respectively). The relative activities of the glycosylated (red bars, *Cb*DC) versus nonglycosylated (purple bars, *Ec*DC) complexes, and the corresponding uncomplexed glycosylated (orange) and nonglycosylated (light pink) enzymes. Enzymatic activity is defined as mM of released sugars after 24-, 48-, and 72-h incubation with 4% of Avicel substrate at 75°C. Each reaction was conducted in triplicate, three times (n = 3); the data represent the mean ± SD, where two asterisks (**) indicate $p<0.01$ (two-tailed *t*-test).

Figure 8. Impact of the different glycosylated modules on the activity of *Cb*DC. The relative activities of the glycosylated (*Cb*DC), nonglycosylated (*Ec*DC), and mixed DCs, were determined at different time intervals. Mixed DC 1 contained only the glycosylated scaffoldin and all three enzymes nonglycosylated, and Mixed DC 2 contained the glycosylated GH9 enzyme with a nonglycosylated form of the scaffoldin. Enzymatic activity is defined as mM of released sugars after 24-, 48-, and 72-h incubation of the enzymes with 4% of Avicel substrate at 75°C. Each reaction was conducted in triplicate, three times (n = 3); the data represent the mean ± SD, where one asterisk (*) indicates $p<0.05$ and two asterisks (**) indicate $p<0.01$ (two-tailed *t*-test): each group at each time point was compared to *Cb*DC.

Figure 9. Thermostability of glycosylated and nonglycosylated monovalent DCs following heat treatment at 85°C. The glycosylated *Cb***GH9**-*lk*-*v* complexed to the glycosylated *Cb***Scaf***GTV* (*Cb* complex) and the nonglycosylated *Ec***GH9**-*lk*-*v* complexed to nonglycosylated *Ec***Scaf***GTV* were incubated at 85°C for 3 h, prior to incubation with the Avicel substrate. Enzymatic activity is defined as mM of released sugars after incubation of the enzymes with 4%

(w/v) Avicel substrate at 75°C. Each reaction was conducted in triplicate, three times (n = 3); the data represent the mean ± SD.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Primers used for plasmid cloning in *E. coli*.

Table S2. Primers used for plasmid cloning in *Ca. bescii*.

Table S3. Primers used for *Ca. bescii* plasmid verification.

Acknowledgments

AmK greatly appreciates scholarships received from the Ministry of Immigrant Absorption, Jerusalem, Israel and from the ministry of Foreign Affairs, Paris, France. AmK is a Sustainability and Energy Weizmann Fellow. E.A.B. is the incumbent of The Maynard I. and Elaine Wishner Chair of Bio-organic Chemistry. This work was authored in part by Alliance for Sustainable Energy, LLC, the manager and operator of the National Renewable Energy Laboratory for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. Funding was also provided by the Center for Bioenergy Innovation (CBI), a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. This research was also supported by the United States–Israel Binational Science Foundation (BSF grant No. 2013284), Jerusalem, Israel; the Israel Science Foundation (ISF grant no. 1349/13); the European Union NMP.2013.1.1-2: CellulosomePlus Project number 604530. The funding sources were not involved in the design of the study, data interpretation, report writing or decision in submitting the article for publication.

Authors' contributions

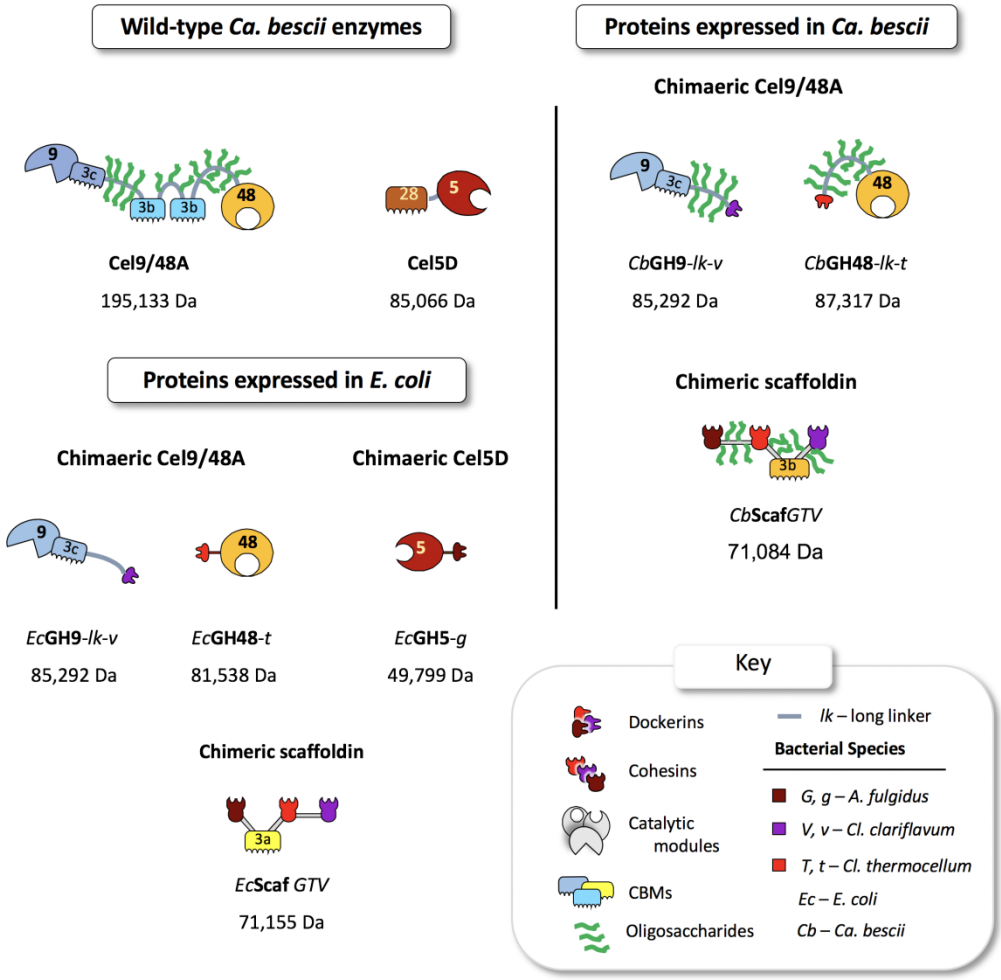
AmK designed the research, performed the experiments and wrote the manuscript. AmK, SM and DC designed and performed the experiments for the cloning and transformation of the plasmid in *Ca. bescii*. AmK, NH and NSS produced and purified the proteins. AmK designed and conducted the binding assay. AmK, AuK and YJB designed and conducted the activity assay. AmK and SM designed and conducted the thermostability assay. All authors analyzed the results. AK, SM, EAB, MEH, and YJB wrote the manuscript. All authors read and approved the manuscript.

Conflict of interests

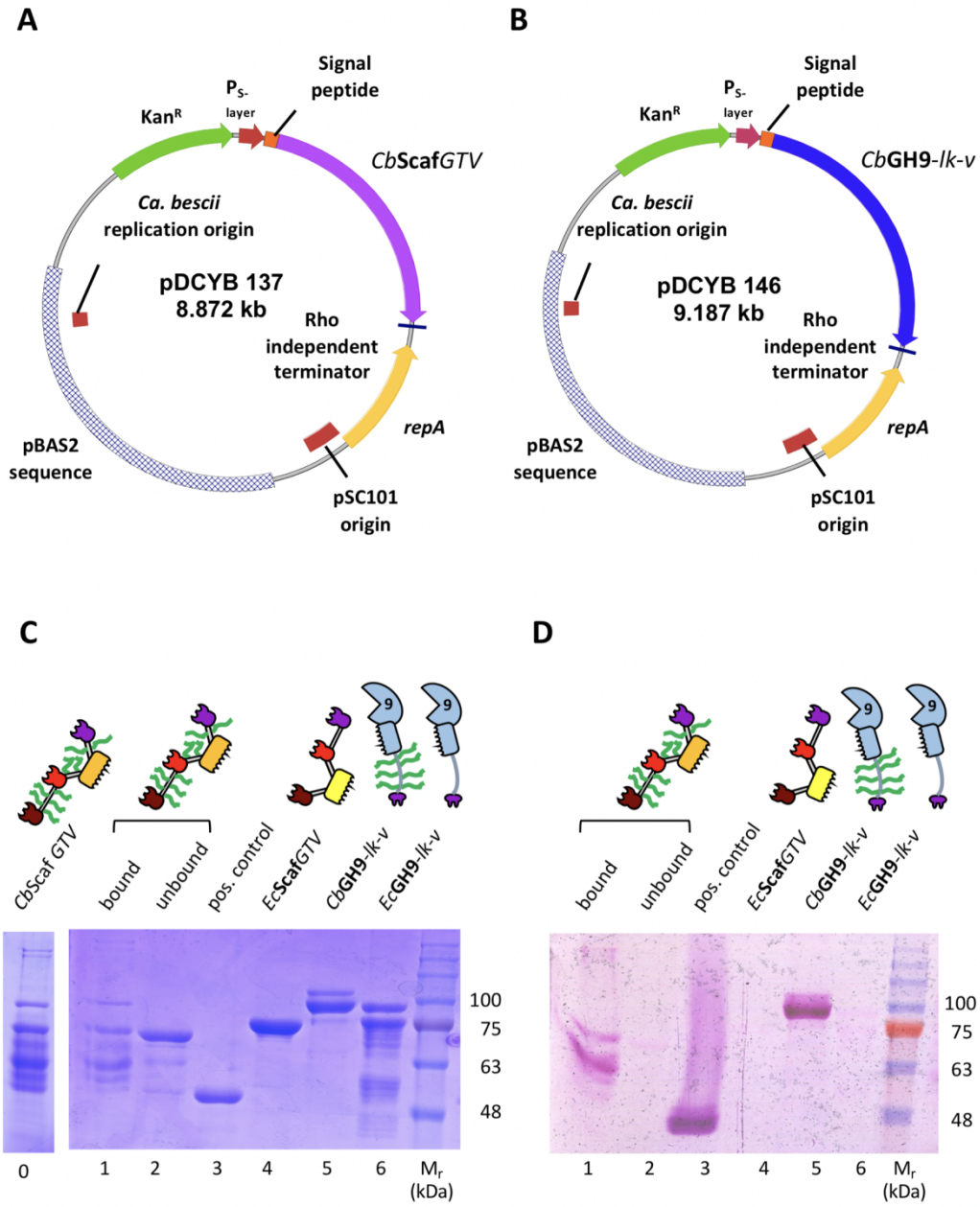
The authors declare that they have no competing interests.

Authors' information

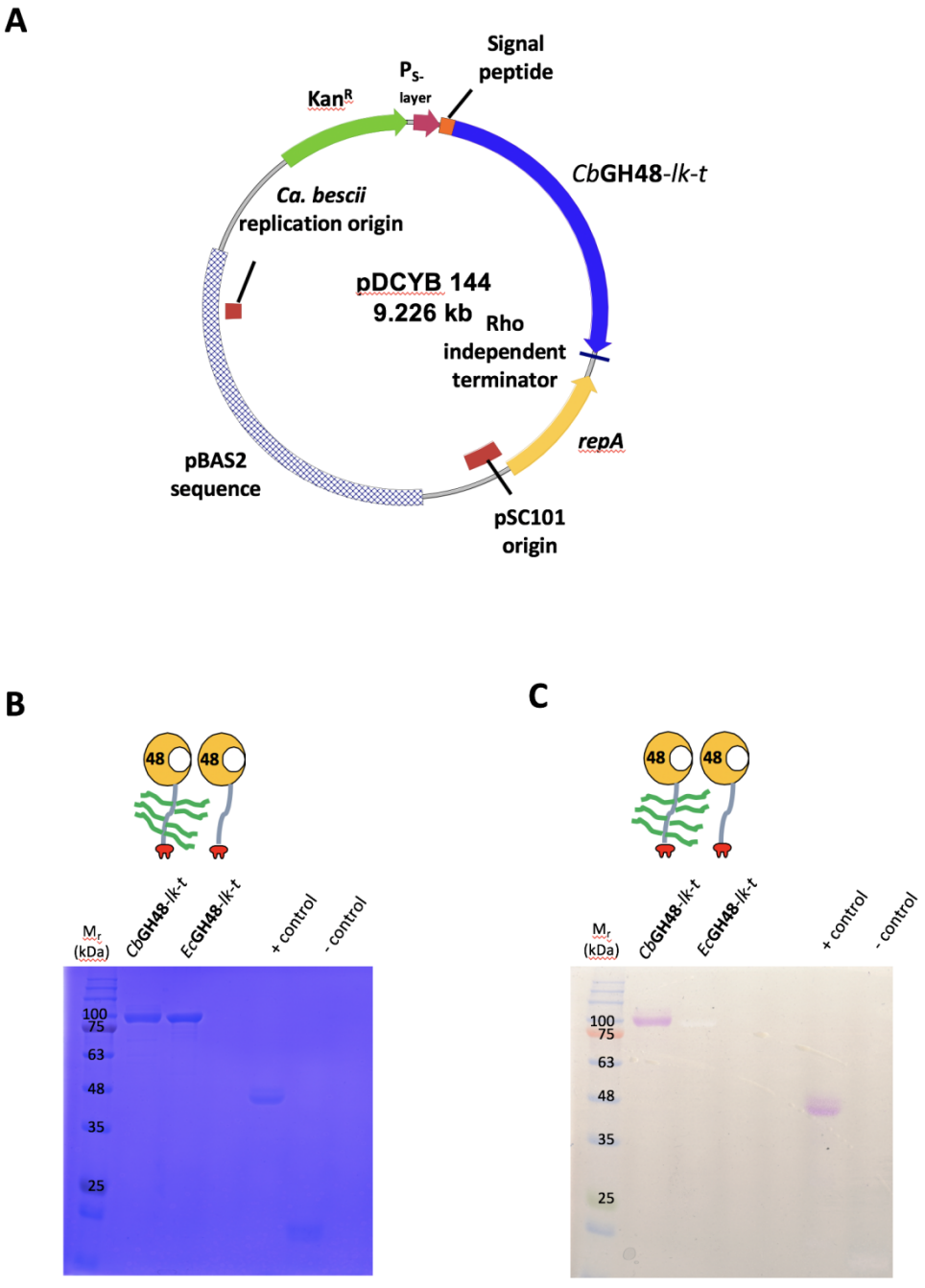
Amaranta Kahn, Email: amaranta.kahn@weizmann.ac.il
Sarah Moraïs, Email: morais@post.bgu.ac.il
Daehwan Chung, Email: Daehwan.Chung@nrel.gov
Nicholas S. Sarai, Email: nsarai@caltech.edu
Neal N. Hengge, Email: Neal.Hengge@nrel.gov
Audrey Kahn, Email: kahn.audrey@gmail.com
Michael E. Himmel, Email: mike.himmel@nrel.gov
Edward A. Bayer, Email: Ed.Bayer@weizmann.ac.il
Yannick J. Bomble, Email: Yannick.Bomble@nrel.gov



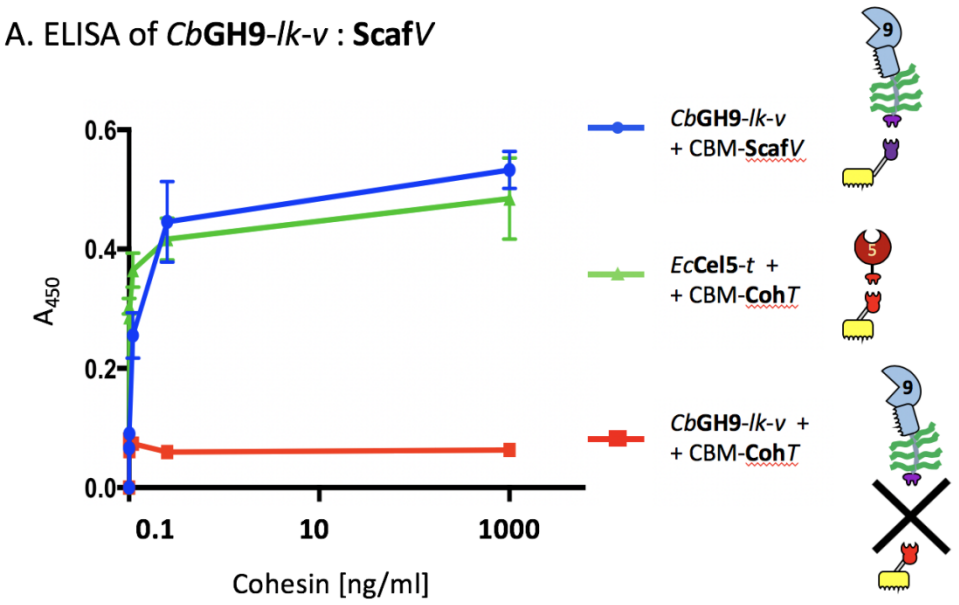
febs_15251_f1.tiff

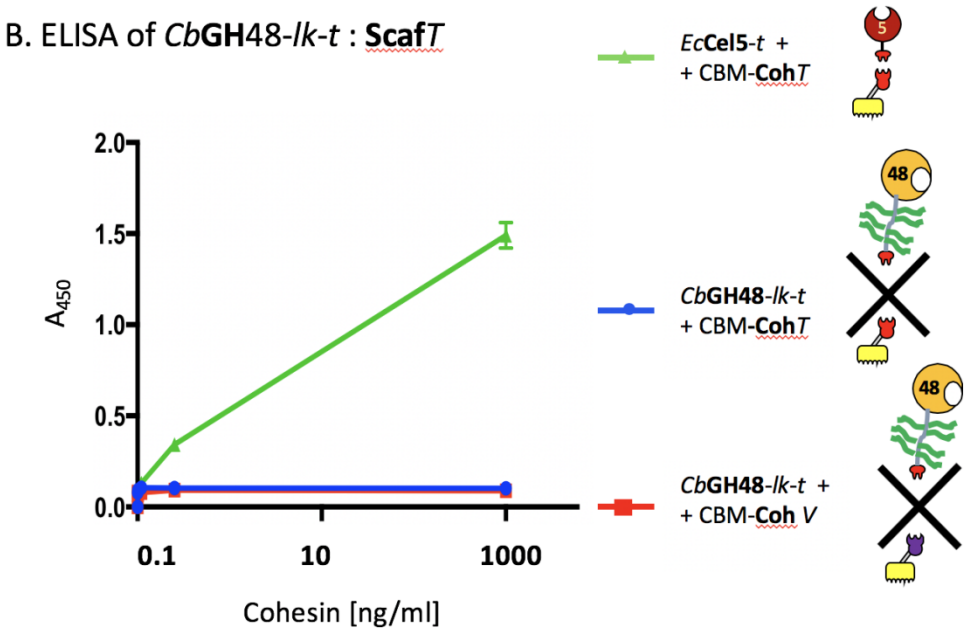


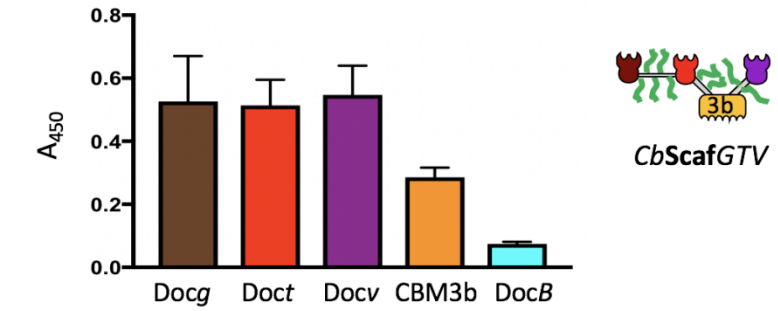
febs_15251_f2.tiff

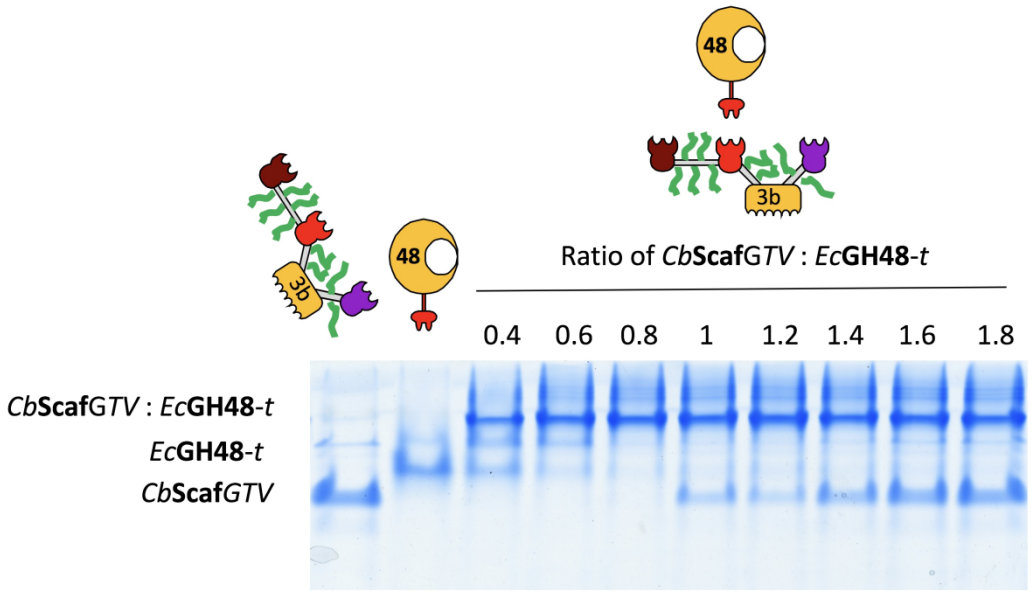


febs_15251_f3.tiff

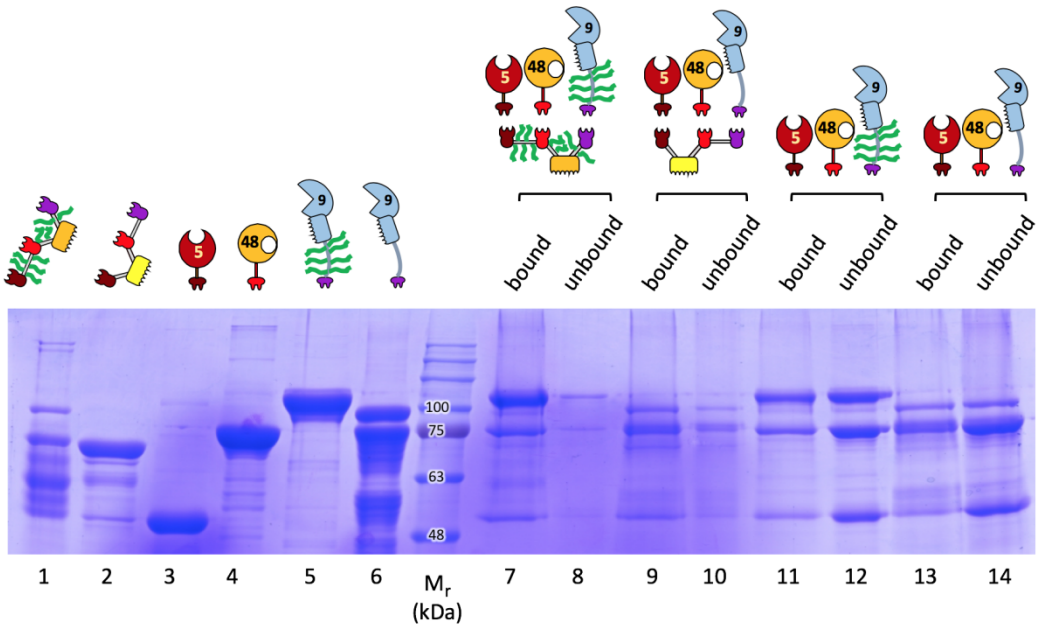




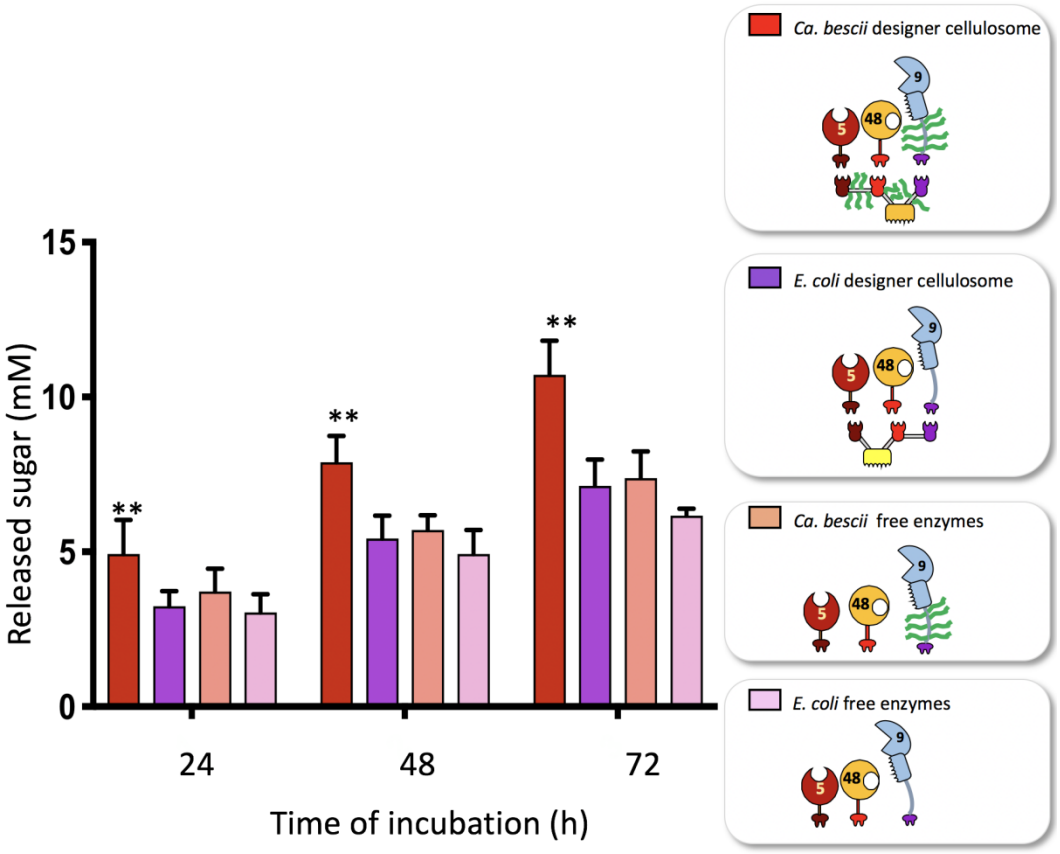




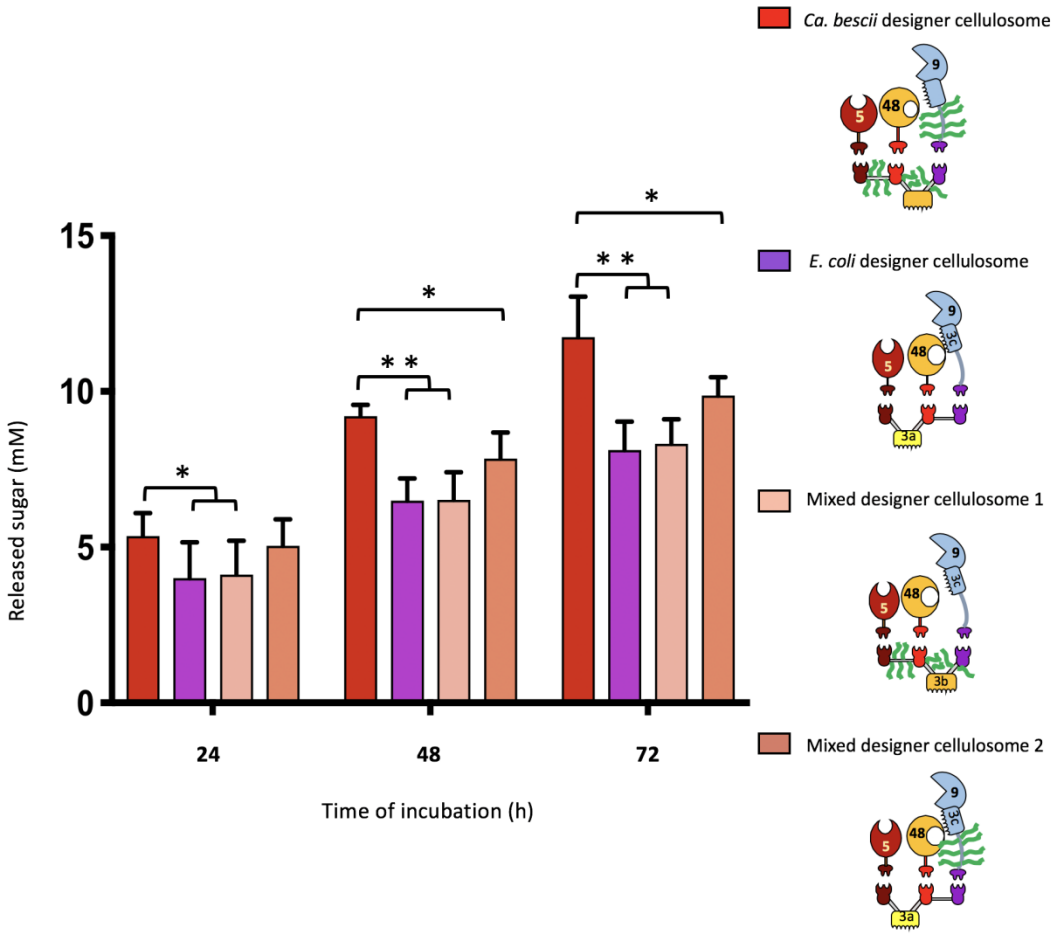
febs_15251_f5.tiff



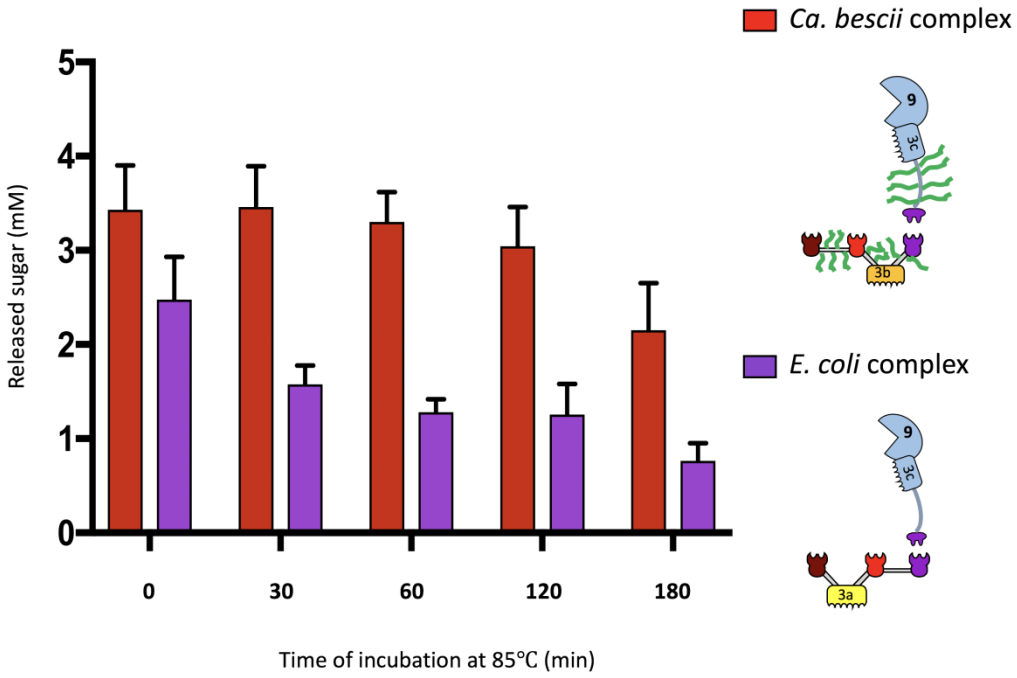
febs_15251_f6.tiff



febs_15251_f7.tiff



febs_15251_f8.tiff



febs_15251_f9.tiff