

# Cloning and characterization of different domain deletion constructs of *Thermobifida fusca* Cel9B

Kathleen Hefferon<sup>1</sup>

<sup>1</sup>Cornell University

April 28, 2020

## Abstract

Cel9B, an endocellulase produced by *Thermobifida fusca* YX, contains a number of structural domains, including carbohydrate binding modules 2 and 4 (CBM2 and CBM4), a fibronectin-like (Fn3) domain, an Eset domain (an Ig-like domain that may play a role in enzyme folding), a catalytic domain, and a fibronectin-like (Fn3) domain. To elucidate the roles of these domains with respect to Cel9B function, a series of truncation mutants were designed and examined for their binding properties and activities on different substrates. Different binding properties of CBM2 and 4 with a variety of substrates distinguish important roles for these domains and provide insight as to how distinct domains interact with each other during substrate degradation. The results of this study implicate the collective roles of the non-catalytic domains with respect to Cel9B function, and in turn, this information can be incorporated into protein engineering strategies for improved biomass conversion.

## Introduction

*Thermobifida fusca* YX is an aerobic, moderately thermophilic cellulolytic bacterium that produces and secretes a large number of proteins when it is grown on cellulose as a sole carbon source (Gomez del Pulgar and Saadeddin, 2014). These include six cellulases (Cel5A, Cel6A, Cel6B, Cel9A, Cel9B, and Cel48A), two AA10 proteins, several hemicellulases as well as a number of uncharacterized proteins (Jeoh et al., 2002, Irwin et al., 2003). All of these proteins have had their genes cloned and have been expressed in *E.coli*, purified, and characterized (Jung et al., 2002, Wilson, 2004, Posta et al., 2004).

One cellulase, Cel9B, was determined to be a  $\beta$ -1,4-endocellulase due to its ability to reduce the viscosity of a carboxymethyl cellulose (CMC) solution over time (Calza et al., 1985, Irwin et al., 1993). Cel9B has a complex domain structure conformed by a N-terminal family 4 carbohydrate binding module (CBM) linked to the protein catalytic domain, which is linked to a fibronectin-like (Fn3) domain, followed by an Eset domain (an Ig-like domain that may play a role in enzyme folding) and ending with a carbohydrate binding module family 2 (CBM2) (Kataeva et al., 2002, 2005, Posta et al., 2004). This enzyme has extremely high activity on CMC but very low activity on crystalline cellulose. This domain structure is very similar to the structure presented on *Clostridium thermocellum* cellulosomal cellulase, CBHA, that shows quite different activity compared to Cel9B (Alahuhta et al., 2010). CBHA was originally thought to be an exocellulase (Schubot et al., 2004). However, its unliganded crystal structure shows an open active site cleft that is characteristic of endocellulases (Kataeva et al., 2005). The endocellulase nature of CBHA was confirmed by several definitive assays (Kataeva et al., 2004). Although Cel9B has been cloned and its protein product purified, the role of the functional domains in the hydrolysis of biomass substrates have not been resolved.

In this paper, we have cloned, expressed, and characterized a number of Cel9B constructs, in which one or more domains were deleted in order to try to understand the roles of the individual domains in activity and substrate binding on a variety of substrates: Bacterial microcrystalline cellulose (BMCC), phosphoric acid-swollen cellulose (PASC),  $\alpha$ -chitin,  $\beta$ -chitin, xylan (birchwood), and xyloglucan. The study offers insight

as to how these functional domains might interact with each other to assist in the process of breaking down recalcitrant cellulose.

## Materials and Methods

### Constructs Used

Cel9B of *T. fusca*, similarly to cellulases from other microorganisms, is comprised of a series of recognizable domains, including a catalytic domain and the carbohydrate binding domains CBM2 and CBM4, as well as the smaller domains of unknown function Eset and Fn3. A depiction of the structure is presented in Figure 1.

The accession number for Cel9B is Q08166. Cel9B-1.pET26b-1, containing wild type Cel9B, was previously cloned into plasmid pET26b-1 using restriction sites *Xho* I and *Nde* I (Collmer and Wilson, 1983). A number of truncation mutant forms of Cel9B were generated with various functional domains removed or retained (Figure 1). For each truncation mutant, the catalytic domain was retained but each carbohydrate binding motif was removed to create the mutants M1 (missing CBM2) and M2 (missing CBM4). Truncation mutants E9B and Fn3 were created by removing all functional domains except for the catalytic domain and E9B or Fn3, respectively. Truncation mutants  $\Delta$ E9B and  $\Delta$ Fn3 lack the functional domains E9B and Fn3 located between the CBM and the catalytic domain, respectively. Truncation mutant cd9B contained only the catalytic domain. M1, M2, cd9B, Fn3, and E9B were PCR amplified using the primers listed in Table 1. Primers were designed according to homology to structural domains identified with other cellulases. Cel9B deletion mutants  $\Delta$ Fn3 and  $\Delta$ E9B were constructed by PCR amplification of deletion constructs M2 and M1, respectively. During primer design, a 6His tag was included on the C-terminal portion of the protein. PCR products were digested with *Xho* I and *Nde* I, then religated to linearized vector pET26 and transformed into DH5 $\alpha$  cells. Subclones containing each truncation mutant were identified by nucleotide sequence analysis and retransformed into *E. coli* BL-21 cells for protein amplification.

### Protein Purification

Purification of Cel9B and its truncation mutants was explained somewhere else (Kostylev et al., 2014). Wild type (WT) and mutants of Cel9B were expressed in *E. coli* BL21-CodonPlus (DE3)-RIPL cells (Agilent Technologies) as a starter culture in Lysogeny Broth (LB) medium overnight at 37 °C. This starter culture was diluted 33 times in fresh LB, and the cells were grown at 37 °C until an OD<sub>600</sub> 0.8 was reached (approximately 3 h). At this point, isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.8 mM, and the culture was incubated at 37 °C with shaking for an extra 16 h. The cells were harvested by centrifugation and resuspended in 10 mL of 20 mM sodium phosphate buffer (pH 8.0), 0.01 M imidazole (Sigma-Aldrich), and 0.5 M sodium chloride (solution A). The cells were lysed by sonication then placed in a 50 °C water bath for 30 min to precipitate the *E. coli* proteins. The precipitate was removed by centrifugation.

The supernatant was loaded onto a nickel Sepharose 6 Fast Flow column (GE Healthcare) and the protein was eluted by a gradient of 0.01-0.5 M imidazole in 50 mM NaH<sub>2</sub>PO<sub>4</sub> buffer, (pH 8.0) and 0.3 M NaCl. Elution fractions which contain WT Cel9B and its truncation mutants, as determined by gel electrophoresis, were loaded on a Q Sepharose (Sigma-Aldrich) column equilibrated with 0.1 M sodium chloride in 5 mM Bis-Tris buffer (pH 5.8), and 10% glycerol (v/v). The protein was eluted from a 0.1-0.7 M sodium chloride gradient in 5 mM Bis-Tris buffer (pH 5.8), and 10% glycerol (v/v). Fractions were measured by gel electrophoresis and those that contained the WT Cel9B or its truncation mutants were combined and concentrated using Millipore centrifugal filter units with a 30 kDa cutoff membrane. Buffer exchange was carried out in the same filter units by washing the concentrated protein three times with 5 mM sodium acetate buffer (pH 5.5), and 10% glycerol (v/v). The protein was passed through a 0.22  $\mu$ m filter using a syringe and stored at -20 °C. All enzyme concentrations were determined by spectroscopy using a NanoDrop 1000 spectrophotometer.

### Substrates

BMCC and PASC were prepared from Avicel powder (PH-105; FMC Corporation, Philadelphia, PA, gift

from Monsanto). To prepare BMCC, Avicel powder was suspended in deionized water with 0.04% sodium azide as described in Kruer-Zerhusen et al., (2017). PASC was prepared from Avicel powder as described in Zhang et al., (2006), and was stored in DI water with 0.04% sodium azide.  $\alpha$ -chitin,  $\beta$ -chitin, xylan (birchwood), and xyloglucan (Sigma) were prepared as 2% solutions in 0.2 mM NaOAc buffer.

### Binding Assays

For each enzyme binding assay, 0.2 mg/mL WT Cel9B / each truncation mutant was mixed separately with excess of substrate (1 mg/mL BMCC, PASC, CMC, xylan,  $\beta$ -chitin,  $\alpha$ -chitin, and xyloglucan) and 0.2 M NaOAc (pH 5.5) in DNA low bind Eppendorf tubes for a total reaction volume of 160  $\mu$ L. Samples were incubated at 4 °C for 1 hour and run in triplicate. Enzyme incubated in the absence of substrate or in buffer alone were included as controls. After incubation, the reaction mixtures were centrifuged at 4,000 rpm for 5 minutes. Supernatants containing free protein were mixed into 900  $\mu$ L Bradford reagent and the concentration was determined using an Ultraspec 10 (Amersham Biosciences) spectrophotometer measuring at OD600.

### Activity Assays

Activity assays were performed in triplicate as follows: 2-fold decreasing concentrations of WT Cel9B enzyme and truncation mutants were incubated for various times (16 hours for BMCC, PASC and xylan, 2 hours for CMC, and 30 minutes for xyloglucan) with 1 mg/ mL of the substrates at 50°C. After each incubation period, the amount of glucose released was determined by PAHBAH assay as described in Kruer-Zerhusen et al., 2017. Briefly, 40  $\mu$ L of each sample was mixed in a 96 well plate with 120  $\mu$ L PAHBAH solution (0.5% 4-hydroxy-benzhydrazide in 0.5 M HCl and 0.5 M NaOH) as The plate was then sealed and heated to 100°C for 6 minutes. Plates were cooled to 42°C and the glucose concentration was measured using a microplate spectrophotometer at 410 nm (Synergy 2, BioTek Instruments Inc., Winooski, VT and a standard curve previously generated.

### Results

Cel9B truncation mutants were designed in such a way that five different structural domains (identified in the CAZy website; (<http://www.cazy.org/>) Lombard et al., 2013) were either removed or retained with the objective of elucidating their specific function (Figure 1, Table 2). All the truncation mutants contained a His tag so that they could be easily purified on a nickel column. Although each of the seven mutants tested exhibited activity on CMC, it was observed that mutant Fn3, which retained the fibronectin-like domain located within the protein, had a reduced capacity to bind to the nickel column. This suggested that the protein structure had been altered so that the 6His tag was no longer available to bind to the column. In addition to this, the truncation mutant  $\Delta$ Fn39B grew poorly.

### Binding Assay

Binding assays were performed on the Cel9B mutants using several different substrates: CMC, PASC, BMCC, xylan, xyloglucan and  $\alpha$  and  $\beta$ -chitin (Table 3). Assays were performed by incubating a known amount of each mutant protein with each substrate at 4 °C for 1 hr. Then the samples were centrifuged at 6,000 rpm for 10 min and the concentration of unbound protein in the supernatant was determined. By difference to the protein added, the amount of bound protein could be quantitated. Generally, wild type protein and mutants showed stronger affinity to BMCC compared to the other substrates (with the exception of cd9B, and Fn3 and  $\Delta$ Fn3 that showed higher affinity to PASC and xylan, respectively). Wild type Cel9B and mutant M2, were found to have a the highest binding yields of 0.144 and 0.128 on BMCC, respectively. On the other hand, the lowest binding yields were found for mutant  $\Delta$ E9B, harboring a CBM4, for all the substrates, and cd9B for xylan and  $\alpha$ -chitin.

Mutants containing CBM2 showed significantly higher binding yield to PASC than those containing CBM4, a trend that was observed to a smaller extent for xylan. Substrate  $\alpha$ -chitin was found to bind to both wild type and M2 (CBM2) at the same yield (approximately 0.104 vs 0.106), followed by M1 (CBM4), in lower amounts. None of the mutants Fn3, E9B or  $\Delta$ E9B appeared to be associated with any of the substrate

fractions. No binding of WT Cel9B or truncation mutants was found for substrates CMC, xyloglucan or  $\beta$ -chitin (data not shown).

### Activity assay on CMC, PASC, and xylan

Cel9B truncation mutants were assayed for their relative enzyme activities on various substrates. Release of glucose (in terms of reducing sugars) from each substrate was quantitated with increasing concentrations of enzyme for wild type and truncation mutants of Cel9B. Mutant cd9B, consisting of only the catalytic domain, had the highest activity on CMC, followed by mutant M1 (containing CBM4), M2 (containing CBM2), and then wild type Cel9B. Other truncation mutants exhibited activity that was lower than that of Cel9B (Figure 2). Conversely, when PASC was used as a substrate, Cel9B and mutants M1 and M2 exhibited similar activities that were greater than cd9B (Figure 3). All other truncation mutants exhibited activities that were similar to cd9B, with mutant  $\Delta$ Fn3 displaying the lowest activity (Figure 4). A similar activity profile to that found for PASC was found for Cel9B and the truncation mutants when xylan was used as a substrate.

### Activity assay on BMCC and xyloglucan

In contrast to the results obtained from the Cel9B truncation mutants on CMC, PASC, and xylan, constructs M1 and to a lesser extent  $\Delta$ E9B, both containing CBM4, exhibited a marked increase in enzyme activity when BMCC was used as a substrate. All other truncation mutants, as well as wild type Cel9B, exhibited lower activity on BMCC (Figure 5). Similarly, both mutants M1 and  $\Delta$ E9B displayed significantly higher activities on xyloglucan than the other constructs, suggesting that CBM4 plays an unknown role for these substrates (Figure 6).

### Discussion

The vast majority of microbial cellulases and other cellulose-degrading enzymes such as xylanases are organized into a modular format which can be divided into catalytic and non-catalytic domains. One or a few carbohydrate-binding motifs (CBMs), responsible for binding to different substrates, are associated with the catalytic domain (CD) of many microbial cellulases. Each cellulase possesses its own unique assortment of CBMs. CBMs are often separated from other cellulase modules by linker molecules, which are believed to assist in the process of targeting the CD to different types of substrate (Boraston et al., 1998).

Cel9B of *T. fusca* follows this modular organization by including both CBM2 and CBM4, which are separated from the CD by a fibronectin-like domain (Fn3) and an Eset domain, respectively. The arrangement of the catalytic domain followed by a CBM2 that is found in Cel9B is also common in a number of enzymes produced by bacteria as well as other organisms, suggesting a common gene transfer event may have taken place in the distant past (Gao et al., 2010, Li et al., 2009).

Experiments carried out in this paper using truncation mutants and different substrates have indicated that mutants keeping the CBM2 domain bind substrates at higher yield than the other mutants, and increase cellulase activity against both crystalline and amorphous cellulose. This was clearly demonstrated using wild type Cel9B as well as truncation mutant M2, which lacks the CBM4. These results agree with the work performed for other cellulases which exhibit higher levels of activity in the presence of CBM2, including *Pseudomonas*, *Cellulomonas fini*, and *Cellvibrio japonicus* (Boraston et al., 1998, Tomme et al., 1996, Gilbert et al., 2013, Attia et al., 2016). CBM2 is one of the most prevalent carbohydrate binding modules and is over-represented in thermophilic microorganisms. It is possible that under thermophilic conditions, CBM2 binding to cellulose may be more stable than those of other CBMs (Moser et al., 2008). Xylan-specific binding of CBM2 has been demonstrated in the past by Simpson et al., (2000). The fact that CBM2 can bind to both amorphous as well as crystalline cellulose could imply that there are differences in CBM2 function on these vastly different substrates. It has been suggested that CBM2s may enhance cellulase activity through its binding mechanism, either by bringing the catalytic domain proximal to the substrate itself or by disrupting and displacing recalcitrant substrate so that more of it is made to be accessible to the catalytic domain (Knowles et al., 1988; Din et al., 1991, Tomme et al., 1996).

Truncation mutant  $\Delta$ Fn3, which lacks an intact Fn3 domain, displayed reduced substrate binding and cellulase activity. It is possible that this protein is misfolded. During the production strain grew slowly (approximately 8 hours to reach OD 0.6 rather than 3-4 hours for the others) and the protein bound poorly to the nickel column. Unlike the other truncation mutant proteins, all of Fn3 came out in the flowthrough and did not bind to the nickel column, as determined by activity assay (Table 2). Brunecky et al., 2012, suggested that the Fn3 domain aids in the thermostability of the protein and may in fact function as a spacer between other functional domains. It is possible, therefore, that the Fn3 domain may have some indirect stabilizing role in the protein that was disrupted upon deletion, or perhaps it aids in the folding of the native protein structure. Fn3 domains have also been thought to play roles in regulating CBM interactions with substrates (Kataeva et al., 2002, Cadena et al., 2010). Although removal of the Fn3 domain in the  $\Delta$ Fn3 mutant appears to have impeded substrate binding, as shown in Table 3, it is more likely that this is really the result of improper protein folding, as truncation mutant Fn3 (containing solely the catalytic domain and the Fn3 domain) does not bind to substrate. Other research groups have shown that Fn3 domain does not affect enzyme binding, rather it can disrupt the surface of substrate so that single cellulase chains are directed to the catalytic site, and thus play a role in regulating catalysis on specific substrates (Kataeva et al., 2002, Guillen et al., 2010).

Truncation mutants M1 and  $\Delta$ E9B, containing CBM4, exhibited pronounced activity on xyloglucan and BMCC that was missing from the other truncation mutants. Truncation mutant  $\Delta$ E9B lacks the Eset domain, an Ig-like domain that may play a role in enzyme folding (Kataeva et al., 2005). This Eset domain is retained in the M1 mutant (Figure 1). CBM4 is known to bind to crystalline cellulose as well as amorphous cellulose (Alahuhta et al., 2010). In this study, truncation mutant M1 exhibited binding to BMCC, albeit at almost half of the yield of the M2 mutant that contains the CBM2 domain. Truncation mutant  $\Delta$ E9B was not able to bind strongly to BMCC (5% binding yield) yet exhibited activity on this substrate. Furthermore, both M1 and  $\Delta$ E9B showed no ability to bind to xyloglucan (data not shown) yet displayed significant activity on this substrate. These results bring into question the role of CBM4 in Cel9B. The relatively high activity of these two mutants on xyloglucan (with respect to the wild type) and their poor ability to bind to substrates would suggest that CBM4 plays a role that differs in general from CBM2. Blake et al., (2006) have demonstrated that CBM4 binding to cellulose chains is limited but requires access to the entire width of a cellulose chain. Indeed, this different binding of CBM4 may be necessary for Cel9B to efficiently catalyze areas of the cell wall that cannot be targeted effectively by CBM2. The poor binding ability to various substrates may indicate that CBM4 plays a role at a later stage of cell wall degradation, where substrates that are available to interact with CBM4 differ from those that are found early on in the degradation process.

Truncation mutants cd9B, E9B, and Fn3 contain the catalytic domain and no CBMs. These mutants exhibited very low yield binding (below 10%) to any of the substrates tested in this study. While all were able to display catalytic activity, cd9B displayed the most, suggesting that retention of the Eset domain and fibronectin-like domain (in E9B and Fn3 mutants, respectively) may have a deleterious effect on the folding of cd9B that impacts its catalytic activity.

WT, M1, and M2 have the highest and similar activities on PASC while the other mutants have activities which are similar to each other, but somewhat lower than WT. This shows that either CBM can improve activity on this substrate. PASC is an amorphous insoluble substrate with much higher surface area than BMCC. On the other hand, the activity results on BMCC show that deleting CBM4 (M1 and  $\Delta$ E9B) causes a large increase in activity for M1 and a small increase for  $\Delta$ E9B over time with respect of the wild type. This suggests that for Cel9B, CBM4 is not important for crystalline cellulose activity. Furthermore, deleting the Eset domain decreases the activity of M1 for reasons that would require further research.

Since CMC is a soluble substrate, removing the CBMs would not be expected to affect this activity, and the removal of CBMs presents increased catalytic activity on CMC for several other cellulases. The activities on PASC and xylan are very similar to the activities on CMC and this may be because both are partially soluble. In addition, these two substrates are not as highly ordered as cellulose, so that binding domains may not be required for their hydrolysis.

## Conclusions

The purpose of this study was to further elucidate the role of the non-catalytic domains by designing a series of truncation mutants and analyzing their ability to bind to and hydrolyze a variety of substrates. The results of this study suggest that the CBMs play broadly different roles in the cellulose degradation process, both in terms of their binding specificity and enhancement of enzyme activity. CBMs are characterized as having both targeting and disruptive functions on different types of substrates. The presence and interaction of these CBM modalities with cellulose may represent the requirement of different binding properties that come into play during various states of cell wall disruption. The observation that removal of CBM4 improved Cel9B's ability to degrade recalcitrant cellulose could potentially be used to engineer this cellulase in the future with improved industrial characteristics. The information gathered from this study will thus be useful for the further understanding of how cellulases in general are able to break down recalcitrant cellulose and will improve our knowledge regarding how microbial cellulases can be utilized more effectively in the future for the development of biofuel.

**Ethics approval and consent to participate :** Not applicable  
**Consent for publication:** The authors consent to the publication of this manuscript  
**Availability of data and material:** Not applicable  
**Competing interests:** The authors declare that they have no competing interests  
**Funding:** This work was supported by the BioEnergy Science Center, a U.S. Department of Energy (DOE) center supported by the Office of Biological and Environmental Research in the DOE Office of Science.  
**Authors' contributions:** All authors contributed to the research presented in this manuscript. The manuscript was written by Kathleen Hefferon.  
**Acknowledgements:** The authors wish to acknowledge the work to the memory of Dr. David Wilson, who is now deceased.

## References

- Alahuhta M, Xu Q, Bomble YJ, Brunecky R, Adney WS, Ding SY, Himmel ME, Lunin VV. 2010. The unique binding mode of cellulosomal CBM4 from *Clostridium thermocellum* cellobiohydrolase A. *J Mol Biol.* 402(2):374-387. DOI: 10.1016/j.jmb.2010.07.028.
- Alahuhta, M., Luo, Y., Ding, S-Y., Himmel, M.E., Lunin, V.V. 2011. Structure of CBM4 from *Clostridium thermocellum* cellulase K *Acta Crystallogr Sect F Struct Biol Cryst Commun.* 67(Pt 5): 527-530. DOI: 10.1107/S1744309111003307
- Attia M, Stepper J, Davies GJ, Brumer H. 2016. Functional and structural characterization of a potent GH74 endo-xyloglucanase from the soil saprophyte *Cellvibrio japonicus* unravels the first step of xyloglucan degradation. *FEBS J.* 283(9):1701-1719. DOI: 10.1111/febs.13696
- Blake AW, McCartney L, Flint JE, Bolam DN, Boraston AB., Gilbert, H.J. Knox, J.P. 2006. Understanding the Biological Rationale for the Diversity of Cellulose-directed Carbohydrate-binding Modules in Prokaryotic Enzymes. *Journal of Biological Chemistry* 281: 29321-29329. DOI: 10.1074/jbc.M605903200
- Boraston, A., Hazelwood, G.P., Gilbert, H.J., Bolam, D.N., Ciruela, A., McQueen-Mason, S., Simpson, P., Williamson, M.P., Rixon, J.E. 1998. Pseudomonas cellulose-binding domains mediate their effects by increasing enzyme substrate proximity *Biochem. J.* 331, 775-781. DOI: 10.1042/bj3310775
- Brunecky, R., Alahuhta, M., Bomble, Y.J., Xu, Q., Baker, J.O., Ding, S.Y., Himmel, M.E., Lunin, V.V. 2012. Structure and function of the *Clostridium thermocellum* cellobiohydrolase A X1-module repeat: enhancement through stabilization of the CbhA complex *Acta Cryst D* 68, 292-299. DOI: 10.1107/S09074444912001680.
- Cadena EM, Chriac AI, Pastor FI, Diaz P, Vidal T, Torres AL. 2010. Use of cellulases and recombinant cellulose binding domains for refining TCF kraft pulp. *Biotechnol Prog.* 26(4):960-967. DOI: 10.1002/btpr.411.
- Calza, R.E., Irwin, D.C., Wilson, D.B. 1985. Purification and characterization of two beta.-1,4-endoglucanases from *Thermomonospora fusca*. *Biochemistry* 1985, 2426, 7797-7804. PMC195339
- Chriac AI, Cadena EM, Vidal T, Torres AL, Diaz P, Pastor FI. 2010. Engineering a family 9 processive

endoglucanase from *Paenibacillus barcinonensis* displaying a novel architecture. *Appl Microbiol Biotechnol.* 86(4):1125-1134. DOI: 10.1007/s00253-009-2350-8

Din, N., Gilkes, N., Tekant, B., Miller, R., Warren, R. and Kilburn, D. 1991. Nonhydrolytic disruption of cellulose fibers by the binding domain of a bacterial cellulase. *Bio/Technology* 9, 1096-1099.

Gao Z, Ruan L, Chen X, Zhang Y, Xu X. 2010. A novel salt-tolerant endo-beta-1,4-glucanase Cel5A in *Vibrio sp.* G21 isolated from mangrove soil. *Appl Microbiol Biotechnol.* 87(4):1373-1382. DOI: 10.1002/jobm.200410422

Guillén D, Sánchez S, Rodríguez-Sanoja R. 2010. Carbohydrate-binding domains: multiplicity of biological roles. *Appl Microbiol Biotechnol.* 85(5):1241-1249. DOI: 10.1007/s00253-009-2331-y

Gilbert, H.J., Knox, J.P., Boraston, A.B. 2013. Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Current Opinion in Structural Biology* 23(5); 669–677. DOI: 10.1016/j.sbi.2013.05.005

Gomez del Pulgar EM, Saadeddin A. 2014. The cellulolytic system of *Thermobifida fusca*. *Crit Rev Microbiol.* 40(3); 236-247. DOI: 10.3389/fpls.2015.01047

Collmer, A., Wilson, D.B. Cloning and expression of a *Thermomonospora X* endocellulase gene in *E. coli*. *Bio/Technology*, 1 (1983)1 :594-601

Irwin DC, Spezio M, Walker LP, Wilson DB. 1993. Activity studies of eight purified cellulases: Specificity, synergism, and binding domain effects. *Biotechnol Bioeng.* 42(8); 1002-1013. DOI: 10.1002/bit.260420811

Jeoh T, Wilson DB, Walker LP. 2002. Cooperative and competitive binding in synergistic mixtures of *Thermobifida fusca* cellulases Cel5A, Cel6B, and Cel9A. *Biotechnol Prog.* 18(4);760-769. DOI: 10.1021/bp0200402

Jung H, Wilson DB, Walker LP. 2002. Binding mechanisms for *Thermobifida fusca* Cel5A, Cel6B, and Cel48A cellulose-binding modules on bacterial microcrystalline cellulose. *Biotechnol Bioeng.* 80(4); 380-392. DOI:10.1002/bit.10375

Kataeva IA, Brewer JM, Uversky VN, Ljungdahl LG. 2005. Domain coupling in a multimodular cellobiohydrolase CbhA from *Clostridium thermocellum*. *FEBS Lett.* 579(20); 4367-4373. DOI:10.1016/j.febslet.2005.06.074

Kataeva IA, Uversky VN, Brewer JM, Schubot F, Rose JP, Wang BC, Ljungdahl LG. 2004. Interactions between immunoglobulin-like and catalytic modules in *Clostridium thermocellum* cellosomal cellobiohydrolase CbhA. *Protein Eng Des Sel.* 17(11) ; 759-769. DOI:10.1093/protein/gzh094

J.K.C. Knowles, T.T. Teeri, P. Lehtovaara, M. Penttilä, M. Saloheimo. The use of gene technology to investigate fungal cellulolytic enzymes J.- *Biotechnology Volume 17, Issue 1, January 1991, Pages 35-49.* DOI: 10.1016. 0168-1656(91)90025-Q

P. Aubert, P. Béguin, J. Millet (Eds.), *Biochemistry and Genetics of Cellulose Degradation*, Academic Press, New York (1988), pp. 153-169. DOI: 10.1007.BF00058833

Kataeva IA, Seidel RD 3rd, Shah A, West LT, Li XL, Ljungdahl LG. 2002. The fibronectin type 3-like repeat from the *Clostridium thermocellum* cellobiohydrolase CbhA promotes hydrolysis of cellulose by modifying its surface. *Appl Environ Microbiol.* 68(9); 4292-4300. DOI: 10.1128/aem.68.9.4292-4300.2002

Li Y, Yin Q, Ding M, Zhao F. 2009. Purification, characterization and molecular cloning of a novel endo-beta-1,4-glucanase AC-EG65 from the mollusc *Ampullariacrossean*. *Comp Biochem Physiol B Biochem Mol Biol.* 153(2); 149-156. DOI: 10.1016/j.cbpb.2009.02.011

Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The Carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42:D490–D495. DOI: 10.1093/nar/gkt1178

Moser F, Irwin D, Chen S, Wilson DB. 2008. Regulation and characterization of *Thermobifida fusca* carbohydrate-binding module proteins E7 and E8. *Biotechnol Bioeng.* 100(6); 1066-1077. DOI: 10.1002/bit.21856

Posta K, Béki E, Wilson DB, Kukolya J, Hornok L. 2004. Cloning, characterization and phylogenetic relationships of cel5B, a new endoglucanase encoding gene from *Thermobifida fusca*. J Basic Microbiol. 44(5); 383-399. DOI: 10.1002/jobm.200410422

Schubot FD, Kataeva IA, Chang J, Shah AK, Ljungdahl LG, Rose JP, Wang BC. 2004. Structural basis for the exocellulase activity of the cellobiohydrolase CbhA from *Clostridium thermocellum*. Biochemistry. 43(5); 1163-1170. DOI:10.1021/bi030202i

Simpson PJ, Xie H, Bolam DN, Gilbert HJ, Williamson MP. 2000. The structural basis for the ligand specificity of family 2 carbohydrate-binding modules. J Biol Chem. 275(52); 41137-41142. DOI: 10.1074/jbc.M006948200

Tomme, P., Creagh, A.L., Kilburn, D.G., Haynes, C.A. 1996. Interaction of Polysaccharides with the N-Terminal Cellulose-Binding Domain of *Cellulomonas fimi* CenC. 1. Binding Specificity and Calorimetric Analysis. Biochemistry 35 (44); 13885-13894. DOI: 10.1021/bi961185i

Wilson DB. 2004. Studies of *Thermobifida fusca* plant cell wall degrading enzymes. Chem Rec. 4(2); 72-82. DOI: 10.1002/tcr.20002

### List of Figure and Table Captions

**Figure 1.** Schematic diagram of wild type Cel9B and truncation mutants constructed for this study. Sp; signal peptide, E; Eset domain, CBM; carbohydrate binding motif, CD; catalytic domain. Numbers on the top horizontal bar represent amino acid residues.

**Figure 2.** Study depicting activity profiles of WT and truncation mutants of TfCel9B using CMC as a substrate. For each time course, increasing concentrations of TfCel9B were incubated on 1 mg/ml CMC in 0.6 mL 50 mM sodium acetate used as buffer and  $\mu$ moles of product were released. Solid lines were derived using logarithm expression. Upper portion, activity profiles of mutants Fn3 and  $\Delta$ Fn3; middle portion, mutants E9B,  $\Delta$ E9B and cd9B. Bottom portion; WT, mutants M1 and M2.

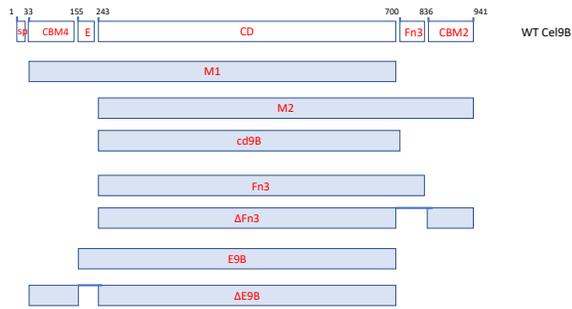
**Figure 3.** Study depicting activity profiles of WT and truncation mutants of TfCel9B using PASC as a substrate. For each time course, increasing concentrations of TfCel9B were incubated on 1 mg/ml PASC in 0.6 mL 50 mM sodium acetate used as buffer and  $\mu$ moles of product were released. Solid lines were derived using polynomial curves that best fit the data. Upper portion, activity profiles of mutants Fn3 and  $\Delta$ Fn3; middle portion, mutants E9B,  $\Delta$ E9B and cd9B. Bottom portion; WT, mutants M1 and M2.

**Figure 4.** Study depicting activity profiles of WT and truncation mutants of TfCel9B using xylan as a substrate. For each time course, increasing concentrations of TfCel9B were incubated on 1 mg/ml xylan in 0.6 mL 50 mM sodium acetate used as buffer and  $\mu$ moles of product were released. Solid lines were derived using polynomial curves that best fit the data. Upper portion, activity profiles of mutants Fn3 and  $\Delta$ Fn3; middle portion, mutants E9B,  $\Delta$ E9B and cd9B. Bottom portion; WT, mutants M1 and M2.

**Figure 5.** Activity Assay of WT and truncation mutants on BMCC. Study depicting activity profiles of WT and truncation mutants of TfCel9B using BMCC as a substrate and  $\mu$ moles of product were released. For each time course, increasing concentrations of TfCel9B were incubated on 1 mg/ml BMCC in 0.6 mL 50 mM sodium acetate used as buffer. Solid lines were derived using polynomial curves that best fit the data. Upper portion, activity profiles of mutants Fn3 and  $\Delta$ Fn3; middle portion, mutants E9B,  $\Delta$ E9B and cd9B. Bottom portion; WT, mutants M1 and M2.

**Figure 6.** Study depicting activity profiles of WT and truncation mutants of TfCel9B using xyloglucan as a substrate. For each time course, increasing concentrations of TfCel9B were incubated on 1 mg/ml xyloglucan in 0.6 mL 50 mM sodium acetate used as buffer, and micromoles of glucose were released. Solid lines were derived using polynomial curves that best fit the data. Upper portion, activity profiles of mutants Fn3 and  $\Delta$ Fn3; middle portion, mutants E9B,  $\Delta$ E9B and cd9B. Bottom portion; WT, mutants M1 and M2.

**Figure 1.** Schematic diagram of wild type Cel9B and truncation mutants constructed for this study. Sp: signal peptide, E: Eset domain, CBM: carbohydrate binding domain, CD: catalytic domain. Numbers represent amino acid residues.



## Hosted file

Cel9B Figures.pdf available at <https://authorea.com/users/314480/articles/444827-cloning-and-characterization-of-different-domain-deletion-constructs-of-thermobifida-fusca-cel9b>