THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Characterization of a Cellulomonas fimi exoglucanase/xylanase-endoglucanase gene fusion which improves microbial degradation of cellulosic biomass

OPEN ACCESS

Accepted Manuscript

Title: Characterization of a *Cellulomonas fimi* exoglucanase/xylanase-endoglucanase gene fusion which improves microbial degradation of cellulosic biomass

Author: Kwabena O. Duedu Christopher E. French

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Characterization of a *Cellulomonas fimi* exoglucanase/xylanase-endoglucanase gene fusion which improves microbial degradation of cellulosic biomass

Kwabena O. Duedu[a,b,*] and Christopher E. French[a]

[a]*Institute of Quantitative Biology, Biochemistry and Biotechnology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3FF, UK*
[b]*Department of Biomedical Sciences, School of Basic & Biomedical Sciences, University of Health & Allied Sciences, Ho, Ghana*

*Corresponding author

Email addresses:
koduedu@uhas.edu.gh (K.O.Duedu),
c.french@ed.ac.uk (C.E.French)

Highlights

- A novel endoglucanase-exoglucanase fusion with hybrid carbohydrate-binding module (CBM) was constructed.
- The fusion protein improves degradation of cellulose filter paper compared to the mixture of unfused endoglucanase and exoglucanase.
- Degradation of cellulose filter paper results in release of cellobiose and cellotriose.
- The hybrid CBM retains the ability to bind to microcrystalline cellulose.
- *Citrobacter freundii* expressing the fusion protein is able to grow at the expense of cellulose filter paper, to a greater extent than when expressing the two unfused proteins.

Abstract

Effective degradation of cellulose requires multiple classes of enzyme working together. However, naturally occurring cellulases with multiple catalytic domains seem to be rather rare in known cellulose-degrading organisms. A fusion protein made from *Cellulomonas fimi* exo- and endo- glucanases, Cex and CenA which improves breakdown of cellulose is described. A homologous carbohydrate binding module (CBM-2) present in both glucanases was fused to give a fusion protein CxnA. CxnA or unfused constructs (Cex+CenA, Cex, or CenA) were expressed in *Escherichia coli* and *Citrobacter freundii*. The latter recombinant strains were cultured at the expense of cellulose filter paper. The expressed CxnA had both exo- and endo- glucanase activities. It was also exported to the supernatant as were the non-fused proteins. In addition, the hybrid CBM from the fusion could bind to microcrystalline cellulose. Growth of *C. freundii* expressing CxnA was superior to that of cells expressing the unfused proteins. Physical degradation of filter paper was also faster with the cells expressing fusion protein than the other constructs. Our results show that fusion proteins with multiple catalytic domains can improve the efficiency of cellulose degradation. Such fusion proteins could potentially substitute cloning of multiple enzymes as well as improving product yields.

1. Introduction

About 75% of the total biomass of lignocellulosic materials is made of cellulose and hemicellulose, making cellulose the most suitable feedstock for production of biofuels and renewable feedstock chemicals [1-3]. Cellulose can be degraded naturally and used as a source of energy by various bacteria and fungi. Such organisms produce endo-β-1,4-glucanases (EC 3.2.1.4), exo-β-1,4-glucanases or cellobiohydrolases (EC 3.2.1.91) and β-glucosidases (EC 3.2.1.21) [4, 5]. The combined action of these enzymes results in saccharification of cellulose. Endoglucanases and cellobiohydrolases break the long cellulose into short-chain units (cellobiose and other short oligonucleotidesaccharides) which are hydrolysed by β-glucosidases into glucose [4, 6]. The endoglucanases and cellobiohydrolases usually consist of two functional units, the catalytic domain (CD) which hydrolyses the β-1,4-glycosidic bonds in cellulose and a non-catalytic CBM which binds to cellulose. In addition to their substrate binding function, CBMs have been shown to improve the enzymatic activity of their cellulases. CBMs may loosen individual cellulose chains from the cellulose surface prior to hydrolysis or enhance the solubilisation of individual glucan chains of the cellulose surface. In addition, some CBMs also serve as thermostabilizing domains [7-10].

The cellulases of *Cellulomonas fimi* are well studied, and the genome has been sequenced [11]. Two of the best studied cellulases of *C. fimi* are the endoglucanase CenA and the bifunctional exoglucanase/xylanase Cex. The N-terminal family-2 CBM of CenA, unlike those of other endoglucanases, has been shown to disrupt the structure of cellulose fibres, resulting in the release of fine cellulose particles without any detectable hydrolytic activity [12]. Cex possesses a homologous CBM at its C-terminus. The N-terminal GH10 catalytic domain of Cex hydrolyses cellulose and cellotetraose to release cellobiose from non-reducing chain ends [13, 14], and has also been shown to possess endoglucanase activity as well as activity against xylan [15-17]. In both CenA and Cex, the CBM is linked to the respective catalytic domain by a proline-threonine (PT) linker. Whist the PT linker acts as a flexible tether, the CBM anchors the catalytic domains onto amorphous, semi-crystalline and crystalline cellulose chains enabling the catalytic domains to hydrolyse nearby cellulose chains [18-22]. In addition the CBM initiates mobility and facilitates the diffusion of the catalytic domain [23].

Gene fusion is an essential tool in systems and synthetic biology, and has been widely used in the production of synthetic bifunctional enzymes [24-27]. In cellulose hydrolysis, gene fusion has been used in studying functions and effectiveness of various CBMs [28-30] as well as to create enzymes with multiple activities [31-33]. Warren and colleagues described a fusion of Cex and CenA made by the use of restriction enzymes. Although the resulting fusion had both endoglucanase and exoglucanase activities, it lacked the flexible PT linkers of both enzymes and all of the CBM of Cex, and most of the CBM of CenA. It did not bind to cellulose, but, like native CenA and Cex, was translocated to the periplasm when expressed *in E. coli* [34]. A xylanase-endoglucanase gene fusion of *Clostridium thermocellum xynX* gene and an enhanced *Erwinia chrysanthemi PY35 cel5Z*::Ω gene has been constructed using restriction independent overlapping [32].

Although various fusions of cellulases have been reported, their effectiveness for cellulose degradation *in vivo* has not been examined. It has been previously demonstrated that recombinant *Citrobacter freundii* (a close relative of *E. coli* possessing the native ability to assimilate cellobiose) expressing CenA and Cex is able to grow at the expense of crystalline cellulose [35, 36]. Here, the preparation of a Cex-CenA fusion protein retaining all the essential properties for effective cellulose degradation is described.

2. Methods

2.1 Bacteria strains and growth conditions

Escherichia coli JM109 was used as host for initial cloning, activity testing and secretion experiments. *Citrobacter freundii* NCIMB 11490 (ATCC 8090) was used as host strain for cellulose utilization experiments. We chose *C. freundii* because it is a close relative of *E. coli* and has natural ability to utilize cellobiose and hence eliminating the requirement for the supply of β-glucosidase to cultures. Luria agar (LA) [37] supplemented with chloramphenicol was used for all plate cultures unless otherwise stated. Our initial experiments showed that 40 µg/ml chloramphenicol slowed down the growth of *C. freundii* constructs. A lower concentration of 15 µg/ml was found to be optimal and was used for *C. freundii* cultures whereas 40 µg/ml was used for *E. coli* cultures. For liquid cultures, Luria broth (LB) [37] or M9 minimal medium [38] supplemented with 0.1% w/v yeast extract was used. All organisms were grown at 37°C. In addition, liquid cultures were incubated with shaking at 200 rpm. Cellulose

degradation experiments were conducted using M9 minimal medium supplemented with Ford's Gold medal blotting paper (2 cm squares, approx. 50 mg each) or Whatman GB003 pure cellulose blotting paper (1 cm squares, approx 30 mg each). Cultures containing filter paper were agitated vigorously for 1 minute every 24 hours on a vortex mixer. Aliquots of the culture were taken from time to time, serially diluted with phosphate buffered saline (PBS) and plated on LA without antibiotics. Colonies on the dilution plates were counted, averaged and plotted as growth curves. The total growth of cells was estimated by calculating the area under the growth curves.

2.2 Molecular biology procedures

Design and construction of genetic parts were done in accordance with the BioBrick assembly standard [39]. The strains, plasmids and oligonucleotides used in this study are described in table 1. PCR was performed using Kod Polymerase (Novogen) according to manufacturer's protocol. 'For amplification of GC rich genes, glycerol (10% v/v) was included in reaction mixtures, and the initial denaturation step in each cycle was extended to 60 s. The PCRs were performed in 50 µL of reaction with 1 µL of the supercoiled plasmid DNA and 1x reaction buffer with 1 mM $MgCl_2$, 0.2 mM each of dNTPs, 1 U of KOD DNA polymerase and 0.4 µM each of forward and reverse primers. Thermal cycling was performed at 95°C for 1 min; followed by 30 cycles of 95°C for 20s, 65°C for 10s, 72°C for 20s/kb; followed by 72 °C for 1 min. Expected product sizes and specific extension times have been provided in the supplementary material (sup. table 1). DNA sequencing was performed by the Sanger sequencing method using the BigDye® Terminator v3.1 cycle sequencing kit (Life Technologies) according to the manufacturer's protocol.

2.3 Construction of the fusion protein

The gene encoding the fusion protein was constructed from previously prepared pSB1C3 plasmids containing BBa_K523016 (P_{lac}+*lacZ'α*+RBS+*cex*) and BBa_K523015 (P_{lac} + *lacZ'α* + RBS + *cenA*). The *lac* promoter was BBa_J33207 and ribosome binding site was BBa_J15001 [40]. An NcoI restriction site was introduced into the CBM of the two constructs by PCR such that the amino acid sequences were not changed. The *cenA* construct was amplified using the forward primer cenAfNcoI and reverse primer pSBNX3insr2. Likewise the *cex* construct was amplified using the forward primer pSBNX3insf2 and reverse primer cexrNcoI. Primer sequences are provided in table 1. PCR amplifications were performed using KOD hot start DNA polymerase (Novagen) using the MJ Research PTC-200 DNA Engine Thermal Cycler. The amplified PCR products were each purified using the QIAquick PCR purification

kit (QIAGEN GmbH, Germany). The purified DNA products were then digested with NcoI (New England Biolabs, Inc.) for one hour and purified. The purified products were ligated overnight at 16°C in a 30 μL reaction volume containing 10 μL of each digested product, 3 μL T4 DNA ligase buffer (with ATP), 1 μL T4 ligase, 1 μL T4 polynucleotide kinase (PNK) and 5 μL nuclease free water. The ligation product was used as a template for PCR with primers pSBNX3insf2 and pSBNX3insr2 using KOD polymerase. PCR products were electrophoresed on 0.8% agarose gels in 1X TAE buffer. Electrophoresis was performed at 100 V, 50 mA for 40 minutes using the Bio-Rad mini-sub cell horizontal electrophoresis system. The gel was post-stained with SYBR-Safe nucleic acid stain (1 μL 10 000 x stock in 50 ml deionized water) for 20 minutes and visualized on the Life Technologies Safe Imager 2.0. Bands were excised and purified using the QIAquick gel extraction kit (QIAGEN GmbH, Germany). The purified products were digested with EcoRI and PstI at 37°C for one hour. The plasmid pSB1C3 was also digested with EcoRI/PstI. Plasmid digests and PCR product digests were ligated together as described above and cloned into *E. coli* JM109. White clones were tested for Cex and CenA activities as described below. Positive clones were grown overnight in 5ml Luria broth supplemented with 40 μg/ml chloramphenicol. Plasmid DNA was isolated using the QIAprep spin miniprep kit (QIAGEN GmbH, Germany) and tested for insert size. Constructs with the right sizes were confirmed by Sanger sequencing performed by Edinburgh Genomics, University of Edinburgh, UK. All procedures were performed according the respective manufacturers' protocols.

2.4 Enzyme activity assays

Enzyme activity within cells harbouring cloned genes were tested qualitatively using 4-methylumbelliferyl β-D-cellobioside (MUC) [20] and carboxymethyl cellulose (CMC) [41] substrates. To detect endoglucanase activity in *E. coli* or *C. freundii* harbouring *cenA* or *cxnA* genes, bacterial colonies were grown on LA with appropriate antibiotics, 0.38 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (Melford) and 0.2% w/v Carboxymethyl cellulose (Sigma Aldrich). The plates were incubated at 37°C. After 24 hours of incubation the plates were stained and washed with 0.1% w/v Congo red solution and 1M NaCl respectively. Staining and washing was done by flooding the plates for 15 minutes during each step. Positive recombinants with endoglucanase activity were surrounded by a cleared zone against a red background. For exoglucanase activity in *E. coli* or *C. freundii* harbouring *cex* or *cxnA* genes, 4-Methylumbelliferyl β-D-cellobioside (MUC) substrate was used. LA plates supplemented with appropriate antibiotics, 0.38mM IPTG were spread with 100 μl of

10 mM MUC and allowed to dry. Bacterial colonies were streaked onto the plates and incubated at 37°C overnight. Following incubation, long wavelength UV (364 nm) lamp was shone onto the plates to reveal fluorescent methlylumbelliferone (4-MU) released when exoglucanase hydrolyses MUC.

Quantitative assays were performed on cell lysate or supernatants using the MUC assay (exoglucanase activity) and the Azo CM-cellulose assay (Megazyme, Ireland). Exoglucanase activity was tested by adding 50 µl of cell lysate to a UV grade cuvette (Fisher Scientific, CXA-110-005J). PBS (1.450 ml) containing MUC (final concentration 0.2mM) was added. The cuvette was covered with parafilm and inverted about 6 times to mix. Fluorescence was measured using the UV filter in the Modulus™ Single Tube Multimode Reader (Turner BioSystems, Inc.), after which the reaction was incubated at 37°C for 60 minutes and fluorescence measured again. A blank solution was prepared like samples but with distilled water. The value of blanks was subtracted from samples. Dilutions of 4-MU were prepared and measured as a standard curve to determine the concentration of 4-MU released in samples (supplementary figure 1a).

For quantitative determination of endoglucanase activity, the Azo CM-cellulose assay was performed according to manufacturer's protocol with modifications. Cell lysate (0.4 ml) or supernatant (0.4ml) to be tested was added to 0.4 ml of Azo-CM cellulose reagent. The mixture was mixed by vortexing after which 300 µl was collected into a 1.5 ml tube (time 0) and the reaction stopped by the addition of 750 µl of absolute ethanol. The remainder was incubated at room temperature for 60 minutes. The stopped reaction was vortexed and centrifuged at 15,000xg for 10 minutes. The supernatant containing depolymerized low molecular weight substrate dyed fragment was collected into a cuvette and the absorbance determined at 600 nm using distilled water as blank. Dilutions of Remazol Brilliant Blue R were prepared and measured as a standard curve to determine the amount of dye released (supplementary figure 1b).

2.5 Protein expression and electrophoresis

E. coli JM109/pSB1C3-BBa_K523015, *E. coli* JM109/pSB1C3-BBa_K523016, *E. coli* JM109/pSB1C3-BBa_K523025, *E. coli* JM109/pSB1C3-BBa_J15509 and *E. coli* JM109/pSB1C3-BBa_J33207 were grown overnight in 5ml LB cultures with 40 µg/ml chloramphenicol. Following incubation, the cultures were used to inoculate 50 ml fresh LB pre-warmed to 37°C in a 1:100 dilution. Cultures were grown to OD_{600} of about 0.6 and induced by adding IPTG to a final concentration of 0.38 mM. To make cell lysates,

cells were harvested after 16 hours by centrifugation (11 300 x g, 20 min) and washed twice with phosphate buffered saline (PBS), pH7 (1.5 mM KH_2PO_4, 5.1 mM Na_2HPO_4, 154 mM NaCl). The culture supernatant was stored at -80°C and assayed for enzyme activity. The cell pellet was washed, resuspended in 5 ml PBS and sonicated for 5 pulses at 10 μm (amplitude) for 20 seconds on ice using the Soniprep 150 ultrasonic disintegrator (MSE, UK). The sonicated samples were centrifuged as before and the supernatant was collected as cell-free lysates. Samples were either analysed immediately or stored at -80°C overnight.

The cell extracts were analysed by electrophoresis using Mini-PROTEAN® TGX™ precast gels under non-denaturing conditions. The gels were run in 1x Tris-Glycine buffer (0.25M Tris, 1.92M Glycine; pH 8.3) at 40 V for 3 hours. For endoglucanase activity, gels were hand-prepared according to the manufacturer's instructions for Mini-PROTEAN system with modifications. CMC (0.2% w/v) was incorporated into the gels. After electrophoresis, the gels were rinsed once in distilled water. They were then flooded with 0.1% w/v Congo red solution and incubated at room temperature on a rotary shaker (80 rpm, 45 min). This was followed by washing in 1M NaCl with shaking. Gels to be assayed for exoglucanase activity were immersed in 10 mM MUC and incubated at 37°C for 1 hour. Due to poor migration in native-PAGE gels (described further in results) lysates were also analysed by agarose gel electrophoresis on 5% w/v agarose gels using Tris-borate buffer (90 mM Tris base, 90 mM Boric acid, pH 8.44).

2.6 Preparation of periplasmic fractions

To investigate the location of expressed proteins, cell fractions (cytoplasmic and periplasmic) were prepared as described in [42]. Briefly, the cells were centrifuged and resuspended in osmotic shock buffer (100 mM Tris–HCl, pH 8.0, 0.5 mM ethylenediaminetetraacetic acid (EDTA), 0.5 M sucrose, and phenylmethylsulfonyl fluoride (PMSF) (20 μg/mL)). The suspension was incubated on ice and centrifuged. Pellets were warmed to room temperature and resuspended in ice-cold water for the osmotic shock. The cell were then centrifuged and the supernatant collected as the periplasmic fraction. The cell pellets were resuspended in PBS and sonicated, and prepared as described earlier for the cytoplasmic fraction. Expression of the recombinant proteins from *E. coli* JM109 cells was performed as described above in M9 medium [38] supplemented with 0.02% yeast extract and 40 μg/ml chloramphenicol.

2.7 Cellulose degradation experiments

C. fimi NCIMB11490 (ATCC8090) was transformed with pSB1C3-BBa_K523025 or pSB1C3-BBa_J15509 and selected on plates containing chloramphenicol. Initial experiments showed that 40 μg/ml chloramphenicol slowed growth, hence 15 μg/ml was used. Enzyme activity was confirmed for all constructs on agar plates. M9 minimal medium supplemented with 1 g/l yeast extract and a main carbon source was used for cellulose degradation experiments. Cellulose degradation experiments were conducted using M9 minimal medium [38] supplemented with filter/blotting paper (Whatman® GB003, 0.8 mm, 300 gsm or Ford 428 mill 0.2 mm, 148gsm) as the main carbon source. For each construct, 7 ml cultures were set up in 30 ml McCartney bottles with either two 1.5 cm squares of Ford 428 paper (approx. 66 mg) or 1 cm squares of Whatman's paper (approx.. 60 mg). Inocula were prepared from overnight cultures grown in LB. The cultures were centrifuged at 6000 g for 10 minutes at 4°C. The supernatant was discarded and pellets were re-suspended in M9 medium to make a 10x concentrate. Each tube was inoculated to a density of 0.1 OD_{600} after which chloramphenicol (15 μg/ml) and IPTG (0.38 mM) were added. In a separate experiment using conical flasks, four 1.5 cm squares of Ford 428 paper were added to 20 ml culture. All other culture constituents were proportional to those prepared for the bottles. The bottles and flasks were grown at 37°C on a rotary shaker (200 rpm). Positive (with glucose) and negative (no carbon) controls were also set up. In addition, a vector control (*C. freundii* NCIMB11490/pSB1C3-BBa_J033207) was included in all experiments. Previous experiments indicated that vigorous agitation for one minute per day was helpful in the breakdown of bulk insoluble substrates such as filter paper (manuscript submitted for publication). Hence, cultures containing filter paper were agitated vigorously (maximum speed) for 1 minute every 24 hours on the Genie 2 vortex mixer (Scientific Industries, Inc.). Colony counts were prepared by serial dilution in sterile PBS and subsequent plating on LA without antibiotic.

2.8 Thin layer chromatography

One millilitre aliquots were collected from the cellulose utilization cultures and analysed by thin layer chromatography (TLC) [43]. The aliquots were concentrated using a DNA Speed Vac (DNA 110, Savant Instruments, Inc. USA) using the low drying rate overnight. The concentrate was resuspended in 100 μl 5% v/v chlorobutanol. To assess the sugar, three microliters of the resuspended samples were loaded as spots

onto 20 cm x 20 cm TLC Silica Gel 60 plates (Merck KGaA, Germany) alongside a master mix standard (MMS) containing the seven main cell wall monosaccharides (galacturonic acid (GalA), galactose glucose, mannose, arabinose xylose and rhamnose). The plates were allowed to dry at room temperature and run in a TLC tank with 5:2:2:1:1 EPPAW buffer (ethlyacetate, propanol, pyridine, acetic acid and water). The TLC was run for 3 hours and air dried overnight in a fume hood. The plates were developed by dipping into thymol solution (5 mg thymol, 95 ml ethanol, 5 ml H_2SO_4). Colour development was achieved by heating plates in an oven (105°C, 5 min).

3. Results

3.1 Generation of the fusion protein.

By analysis of the CBM sequences of Cex and CenA, a consensus sequence PWNGSIP was found. It was determined that an NcoI site could be introduced into both without changing the amino acid sequence (supplementary figure 2). The two plasmids containing each of the fusion partners were amplified by the respective primers, digested with NcoI and ligated together. The ligation product was used as a template for PCR which yielded two products; a larger around 3.1 kb and a smaller around 2.1 kb. The larger was presumed to be correct and was purified from gel. The smaller was presumed to be a mixture of Plac-lacZα-cenA and Plac-lacZα-cex from intact template which had made it through the purification. The 3.1 kb fragment was gel purified and digested with EcoRI/PstI and ligated into pSB1C3 (replacing mRFP, BBa_J04450) pre-digested with the same enzymes (figure 1). *E. coli* JM109 competent cells were transformed and white colonies selected from plates containing chloramphenicol. The resulting fusion protein had the catalytic domains of Cex at the N-terminus and CenA at the C-terminus. In between these was a new chimeric CBM (CBM CxnA) generated as a result of the fusion (supplementary figure 2).

3.2   Cellulase activity of the expressed gene-fusion

E. coli JM109/pSB1C3-BBa_K523025, showed fluorescence on MUC plates as well as zones of clearing on CMC plates indicating Cex and CenA activities respectively (figure 2; A & B). Both activities were also detected in cell lysates on agar plates (figure 2; C & D) and also by electrophoresis of cell lysates on non-denaturing (native) PAGE gels. On native-PAGE gels (4% stacking, 12% resolving) however, the proteins did not migrate through the gels (precast or in-house prepared), but remained at the junction

between the stacking and resolving gels. Similarly, no migration was observed in 4 - 15 % precast gradient gels with proteins remaining in the wells. Electrophoresis was therefore repeated using 5% agarose gels in Tris-borate buffer (pH 8.44) (Figure 2; E & F). Unexpectedly, the active proteins migrated towards the cathode, rather than the anode, although the predicted isoelectric points (pI) of the proteins (Cex, pI 6.18; CenA, 7.56; CxnA 6.19) were lower than the pH of the buffer. Activity spots due to CenA and Cex migrated further than those due to CxnA, indicating that both activities were associated with a protein of a different size from either CenA or Cex, as expected (figure 2; E& F).

3.3 Investigation of the ability of the fused CBM to bind to cellulose

To determine whether the fused CBM of CxnA could bind to cellulosic substrate, the crude protein extracts were incubated with and without microcrystalline cellulose (avicel) and the reduction in activity due to binding to cellulose was determined. Binding of the enzymes to 20 mg/ml and 200 mg/ml of avicel in PBS was also tested. There was no significant difference between the enzyme bound to 20 mg/ml and 200 mg/ml of avicel. Unbound enzyme activity was 4.5% and 38% for non-fused and fused (CxnA) enzymes (figure 3). The remaining activity was recovered when the avicel pellet was resuspended in PBS and assayed. No significant difference ($p > 0.33$) was found between the activity obtained before addition of avicel and the activity recovered from the resuspended avicel. There was however significant differences ($p < 0.01$) between the activity recovered from the supernatant after binding and that obtained prior to addition of avicel.

3.4 Investigation of extracellular secretion/leakage of expressed proteins

Secretion of cellulase is important for accessibility and hydrolysis of substrate. To determine whether CxnA was secreted into the culture supernatant, the enzyme activity levels in culture medium (extracellular), periplasmic and cytoplasmic protein extracts of recombinant *C. freundii* were compared. About 20% of the exoglucanase activity was found in the extracellular fraction of both CxnA and non-fused enzymes (figure 4). A similar result was obtained with *E. coli* (data not shown).The majority of enzyme activity was however located in the periplasmic fraction.

3.5 Growth of Citrobacter freundii utilizing cellulosic substrates

Growth of *C. freundii* NCIMB11490/pSB1C3-BBa_K523025 (fused) was higher than *C. freundii* NCIMB11490/pSB1C3-BBa_J15509 (non-fused) (figure 5; a, b & c). This observation was consistent irrespective of the filter paper used. Physical destruction of filter paper was faster with cells expressing the fused enzyme as opposed to cells expressing the non-fused. In shake flasks (100 ml conical flasks, 62 mm base diameter, foam stopper), significant destruction of filter paper (Ford 428) was seen at day 7 for the cells expressing the fused construct. In McCartney bottles (30 ml capacity, 28 mm base diameter, and plastic screw cap) however physical destruction was observed earlier (day 3; supplementary figure 4). Physical destruction of filter paper was faster for cells expressing the fused enzyme than cells expressing the non-fused enzymes. By day 5 however, filter paper was broken down in cells expressing both fused and non-fused enzymes whereas those of the vector control were not destroyed.

3.6 Release of sugars from cellulose filter paper

To determine whether the filter paper was being hydrolysed by cells expressing the recombinant cellulases, aliquots of the cultures were taken at various time points and analysed using thin layer chromatography. Spots corresponding to cellobiose and cellotriose were detected indicating that the expression of the gene fusion led to the hydrolysis of cellulose in the filter paper. On day 17 only the cells expressing the fusion protein showed extra spots (figure 6, supplementary figure 3).

4. Discussion

In order to test synergistic effects of co-localization of different enzyme domains for microbial cellulose hydrolysis, there was a need to design and test simple enzyme complexes with multiple enzyme activities. Several cellulases have been isolated and characterized. However, finding a combination of cellulases that will enable strains to utilize crystalline cellulose as a carbon source is challenging. The expression of *C. fimi* exoglucanase/xylanase Cex and endoglucanase CenA in *C. freundii* which allows the recombinants to grow on cellulosic substrates has been described earlier [35, 36, 40]. Fusion enzymes catalysing the same multistep sequential reaction have been shown to demonstrate superiority over the individual native enzymes [44]. Cellulases with co-localized domains are exemplified by the multienzyme machines known as cellulosomes. Co-localization improves synergy between domains resulting in effective degradation of cellulose [45-47].

The fusion protein CxnA was found to bind cellulose. Binding however was not as great as that of the native CBM. The reason for this is unclear but could be due some small changes in the conformation of the chimeric CBM. Further engineering of the CBM could potentially improve binding as well as help to understand requirements and process of binding. In addition to binding, it has been suggested that CBMs enhance the cellulose degradation by loosening individual chains from the cellulose surface [10]. The ability of the chimeric CBM of CxnA to do this has not been determined.

Enzyme activity was detected in all fractions (cytoplasmic, periplasmic and culture media). The mechanism by which *E. coli* and *C. freundii* secrete these proteins from the periplasm is not clear. Secretion however indicates that the signal peptides of the cellulases are recognized by the *E. coli* and *C. freundii* Sec pathway. The ability of cellulolytic organisms to effectively secrete cellulases into the medium is necessary to enhance degradation of cellulosic substrates. There have been previous reports of Cex and CenA secretion in *E. coli* [48].

Synergy in other cellulase fusions has been reported [24, 30, 32]. Fusion of Cex and CenA was reported earlier [34]. Though this fusion contained both exoglucanase and endoglucanase activities, it was unable to bind to cellulose because it lacked an intact CBM. Furthermore, it was not tested whether the fusion could hydrolyse cellulosic substrates. To check whether the new fusion protein could hydrolyse cellulose better than non-fused ones, *Citrobacter freundii* was transformed with the fusion as well as with the individual cellulases. The ability of the transformants to utilize filter paper for growth was tested. *C. freundii* was chosen because it is a close relative of *E. coli* with native ability to assimilate cellobiose. Furthermore, it is not considered pathogenic (Advisory Committee on Dangerous Pathogens, UK) although it is capable of causing nosocomial infections in the immunocompromised. By using *C. freundii,* there was no need to include a β-glucosidase to the cultures or clone it into the constructs. From previous experiments, it was determined that *C. freundii* required a small amount (0.1% w/v) of yeast extract to help it grow and reach sufficient numbers to produce the initial enzymes to break down cellulose. It was observed that all the cultures (including negative controls) grew similarly within the first 48 hours. However, numbers of live cells within the negative control cultures dropped from this point forward whereas the number of live cells within the cultures expressing the cellulases continued to increase. Cultures expressing the fused enzyme increased about 50% more than the cultures

expressing the non-fused enzymes. This was confirmed when the total growth was expressed as the total area covered by the growth curves.

In order to assess the hydrolysis of the filter paper and production of cellobiose, aliquots of the cultures were analysed over time. It was observed that the cellulases enabled cells expressing them to produce cellobiose and cellotriose in culture from day 2. Production of these products was observed for the cultures until day 17 when the cultures expressing the non-fused proteins no longer showed free cellobiose. Furthermore, the colony counts of the non-fused cultures had also reduced by about 90%. Physical degradation of filter paper by the fusion constructs started earlier than the non-fused constructs. These results are evidence that the fused protein improves degradation of the filter paper releasing more cellobiose such that there is still a detectable quantity to support growth at a time when the non-fused constructs no longer show detectable cellobiose. The fused constructs may perhaps be more stable than the unfused enzymes making it stay active for a longer period. The stability of the enzymes was not determined. Here we have described improved cellulose degradation by our fused construct *in vivo* using cell growth and physical degradation. Fusion proteins can lead to increased effectiveness due to substrate channelling even if the measured activities of the two individual partners decreases. This was reported earlier from our lab [49] as well as is shown in this study. Presently assays to measure quantitative cellulose degradation *in vivo* are not available. We are in the process of evaluating some assays we have developed to be able to better quantify cellulose degradation. Follow-on studies on liberation of sugars from cellulosic biomass by purified enzyme are being undertaken. These studies will help improve our understanding of how fusion proteins can be used to improve biocatalysis.

5. Conclusions

This work has demonstrated that fusing homologous domains of two different enzymes can give a new gene fusion which when expressed is functional and may be more effective than the native proteins. Co-localization of the catalytic domains of cellulases may enhance the overall catalytic activity of both enzymes. Such fusions could potentially replace cloning of multiple enzymes for improved biomass conversion and downstream processing. The apparent leakage of the *C. fimi* cellulases from *E. coli* is an interesting finding. Further investigations to understand the mechanism of secretion will provide useful insights into how to enhance extracellular secretion of cellulases and other proteins.

- 15 -

Acknowledgements

References

[1] Bayer EA, Lamed R, Himmel ME. The potential of cellulases and cellulosomes for cellulosic waste management. Curr Opin Biotechnol. 2007;18:237-45.

[2] Carroll A, Somerville C. Cellulosic biofuels. Annu Rev Plant Biol. 2009;60:165-82.

[3] Gomez LD, Steele-King CG, McQueen-Mason SJ. Sustainable liquid biofuels from biomass: the writing's on the walls. New Phytol. 2008;178:473-85.

[4] Van Dyk JS, Pletschke BI. A review of lignocellulose bioconversion using enzymatic hydrolysis and synergistic cooperation between enzymes--factors affecting enzymes, conversion and synergy. Biotechnol Adv. 2012;30:1458-80.

[5] Knowles JKC, Lehtovaara P, Teeri TT. Cellulase families and their genes. TIBTECH. 1987;5:255–61.

[6] Xiao Z, Zhang X, Gregg DJ, Saddler JN. Effects of sugar inhibition on cellulases and beta-glucosidase during enzymatic hydrolysis of softwood substrates. Appl Biochem Biotechnol. 2004;113-116:1115-26.

[7] Hall M, Rubin J, Behrens SH, Bommarius AS. The cellulose-binding domain of cellobiohydrolase Cel7A from *Trichoderma reesei* is also a thermostabilizing domain. J Biotechnol. 2011;155:370-6.

[8] Linder M, Teeri TT. The cellulose-binding domain of the major cellobiohydrolase of Trichoderma reesei exhibits true reversibility and a high exchange rate on crystalline cellulose. Proceedings of the National Academy of Sciences of the United States of America. 1996;93:12251-5.

[9] Teeri TT, Penttila M, Keranen S, Nevalainen H, Knowles JK. Structure, function, and genetics of cellulases. Biotechnol. 1992;21:417-45.

[10] Teeri TT, Reinikainen T, Ruohonen L, Jones TA, Knowles JKC. Domain function in *Trichoderma reesei* cellobiohydrolases. J Biotechnol. 1992;24:169-76.

[11] Christopherson MR, Suen G, Bramhacharya S, Jewell KA, Aylward FO, Mead D, et al. The genome sequences of *Cellulomonas fimi* and "*Cellvibrio gilvus*" reveal the cellulolytic strategies of two facultative anaerobes, transfer of "*Cellvibrio gilvus*" to the genus *Cellulomonas*, and proposal of *Cellulomonas gilvus sp. nov*. PloS one. 2013;8:e53954.

[12] Din N, Damude HG, Gilkes NR, Miller RC, Jr., Warren RA, Kilburn DG. C1-Cx revisited: intramolecular synergism in a cellulase. Proceedings of the National Academy of Sciences of the United States of America. 1994;91:11383-7.

[13] Kleman-Leyer KM, Gilkes NR, Miller RC, Jr., Kirk TK. Changes in the molecular-size distribution of insoluble celluloses by the action of recombinant *Cellulomonas fimi* cellulases. Biochem J. 1994;302 (Pt 2):463-9.

[14] White A, Withers SG, Gilkes NR, Rose DR. Crystal structure of the catalytic domain of the β-1,4-glycanase Cex from *Cellulomonas fimi*. Biochem. 1994;33:12546-52.

[15] Notenboom V, Birsan C, Warren RA, Withers SG, Rose DR. Exploring the cellulose/xylan specificity of the β-1,4-glycanase cex from *Cellulomonas fimi* through crystallography and mutation. Biochem. 1998;37:4751-8.

[16] Notenboom V, Williams SJ, Hoos R, Withers SG, Rose DR. Detailed structural analysis of glycosidase/inhibitor interactions: complexes of Cex from *Cellulomonas fimi* with xylobiose-derived aza-sugars. Biochem. 2000;39:11553-63.

[17] MacLeod AM, Lindhorst T, Withers SG, Warren RA. The acid/base catalyst in the exoglucanase/xylanase from *Cellulomonas fimi* is glutamic acid 127: evidence from detailed kinetic studies of mutants. Biochem. 1994;33:6371-6.

[18] Poon DKY, Withers SG, McIntosh LP. Direct demonstration of the flexibility of the glycosylated proline-threonine linker in the *Cellulomonas fimi* xylanase Cex through NMR spectroscopic analysis. J Biol Chem. 2007;282:2091-100.

[19] Gilkes NR, Henrissat B, Kilburn DG, Miller RC, Jr., Warren RA. Domains in microbial β-1,4-glycanases: sequence conservation, function, and enzyme families. Microbiol Rev. 1991;55:303-15.

[20] Ong E, Gilkes NR, Miller RC, Jr., Warren RA, Kilburn DG. The cellulose-binding domain (CBD$_{Cex}$) of an exoglucanase from *Cellulomonas fimi*: production in *Escherichia coli* and characterization of the polypeptide. Biotechnol Bioeng. 1993;42:401-9.

[21] Creagh AL, Ong E, Jervis E, Kilburn DG, Haynes CA. Binding of the cellulose-binding domain of exoglucanase Cex from *Cellulomonas fimi* to insoluble microcrystalline cellulose is entropically driven. Proceedings of the National Academy of Sciences of the United States of America. 1996;93:12229-34.

[22] McLean BW, Bray MR, Boraston AB, Gilkes NR, Haynes CA, Kilburn DG. Analysis of binding of the family 2a carbohydrate-binding module from *Cellulomonas fimi* xylanase 10A to cellulose: specificity and identification of functionally important amino acid residues. Protein Eng. 2000;13:801-9.

[23] Jervis EJ, Haynes CA, Kilburn DG. Surface diffusion of cellulases and their isolated binding domains on cellulose. J Biol Chem. 1997;272:24016-23.

[24] Bulow L. Characterization of an artificial bifunctional enzyme, beta-galactosidase/galactokinase, prepared by gene fusion. Eur J Biochem. 1987;163:443-8.

[25] Ljungcrantz P, Carlsson H, Mansson MO, Buckel P, Mosbach K, Bulow L. Construction of an artificial bifunctional enzyme, beta-galactosidase/galactose dehydrogenase, exhibiting efficient galactose channeling. Biochem. 1989;28:8786-92.

[26] Seo HS, Koo YJ, Lim JY, Song JT, Kim CH, Kim JK, et al. Characterization of a bifunctional enzyme fusion of trehalose-6-phosphate synthetase and trehalose-6-phosphate phosphatase of *Escherichia coli*. Appl Environ Microbiol. 2000;66:2484-90.

[27] Tamada Y, Swanson BA, Arabshahi A, Frey PA. Preparation and characterization of a bifunctional fusion enzyme composed of UDP-galactose 4-epimerase and galactose-1-P uridylyltransferase. Bioconjug Chem. 1994;5:660-5.

[28] Zhao L, Pang Q, Xie J, Pei J, Wang F, Fan S. Enzymatic properties of *Thermoanaerobacterium thermosaccharolyticum* beta-glucosidase fused to *Clostridium cellulovorans* cellulose binding domain and its application in hydrolysis of microcrystalline cellulose. BMC Biotechnol. 2013;13:101.

[29] Murashima K, Kosugi A, Doi RH. Solubilization of cellulosomal cellulases by fusion with cellulose-binding domain of noncellulosomal cellulase engd from Clostridium cellulovorans. Proteins. 2003;50:620-8.

[30] Thongekkaew J, Ikeda H, Masaki K, Iefuji H. Fusion of cellulose binding domain from *Trichoderma reesei* CBHI to *Cryptococcus sp.* S-2 cellulase enhances its binding affinity and its cellulolytic activity to insoluble cellulosic substrates. Enzyme Microb Technol. 2013;52:241-6.

[31] Lu P, Feng MG, Li WF, Hu CX. Construction and characterization of a bifunctional fusion enzyme of Bacillus-sourced beta-glucanase and xylanase expressed in Escherichia coli. FEMS microbiology letters. 2006;261:224-30.

[32] An JM, Kim YK, Lim WJ, Hong SY, An CL, Shin EC, et al. Evaluation of a novel bifunctional xylanase–cellulase constructed by gene fusion. Enzyme Microb Technol. 2005;36:989-95.

[33] Bower B, Larenas E, Mitchinson C. Exo-endo cellulase fusion protein. Google Patents; 2006.

[34] Warren RA, Gerhard B, Gilkes NR, Owolabi JB, Kilburn DG, Miller RC, Jr. A bifunctional exoglucanase-endoglucanase fusion protein. Gene. 1987;61:421-7.

[35] French CE, Horsfall L, Barnard DK, Duedu K, Fletcher E, Joshi N, et al. Beyond Genetic Engineering: Technical Capabilities in the Application Fields of Biocatalysis and Biosensors. In: Giese B, Pade C, Wigger H, von Gleich A, editors. Synthetic Biology: Springer International Publishing; 2015. p. 113-37.

[36] Lakhundi SS, Duedu KO, Cain N, Nagy R, Krakowiak J, French CE. *Citrobacter freundii* as a test platform for recombinant cellulose degradation systems. *Manuscript submitted*.

[37] Bertani G. Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. J Bacteriol. 2004;186:595-600.

[38] Cold Spring Harb. Recipe: M9 minimal medium (modified). Cold Spring Harb Protoc. 2010;doi:10.1101/pdb.rec12296.

[39] Knight T. Idempotent vector design for standard assembly of BioBricks. MIT Artificial Intelligence Laboratory; MIT Synthetic Biology Working Group; 2003.

[40] Lakhundi S. A synthetic biology approach to cellulose degradation [PhD Thesis]. Edinburgh: University of Edinburgh; 2011.

[41] Teather RM, Wood PJ. Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. Appl Environ Microbiol. 1982;43:777-80.

[42] Rutter C, Mao Z, Chen R. Periplasmic expression of a *Saccharophagus cellodextrinase* enables *Escherichia coli* to ferment cellodextrin. Appl Microbiol Biotechnol. 2013;97:8129-38.

[43] Fry SC. The growing plant cell wall: chemical and metabolic analysis. Harlow, Essex, England, New York: Longman Scientific & Technical; Wiley; 1988.

[44] Bulow L, Mosbach K. Multienzyme systems obtained by gene fusion. Trends Biotechnol. 1991;9:226-31.

[45] Bayer EA, Belaich JP, Shoham Y, Lamed R. The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. Annu Rev Microbiol. 2004;58:521-54.

[46] Bayer EA, Chanzy H, Lamed R, Shoham Y. Cellulose, cellulases and cellulosomes. Curr Opin Struct Biol. 1998;8:548-57.

[47] Bayer EA, Shimon LJ, Shoham Y, Lamed R. Cellulosomes-structure and ultrastructure. J Struct Biol. 1998;124:221-34.

[48] Guo Z, Arfman N, Ong E, Gilkes NR, Kilburn DG, Warren RAJ, et al. Leakage of *Cellulomonas fimi* cellulases from *Escherichia coli*. FEMS Microbiol Lett. 1988;49:279-83.

[49] Lewicka AJ, Lyczakowski JJ, Blackhurst G, Pashkuleva C, Rothschild-Mancinelli K, Tautvaisas D, et al. Fusion of pyruvate decarboxylase and alcohol dehydrogenase increases ethanol production in Escherichia coli. ACS Synth Biol. 2014;3:976-8.

[50] Breed RS, Dotterrer WD. The number of colonies allowable on satisfactory agar plates. Journal of bacteriology. 1916;1:321-31.

Figures

Figure 1: Illustration of the design and construction of the exoglucanase-endoglucanase fusion

(A) Fusion parts were amplified from BioBrick plasmid pSB1C3 containing BBa_K523016 (P$_{lac}$+*lacZα*+RBS+*cex*) and BBa_K523015 (Plac + lacZα + RBS + *cenA*) using primers PSBNX3insf2/cexrNcoI and cenANcoI/ PSBNX3insr2 respectively. (B) The two PCR products were purified and digested with NcoI restriction enzyme in a single tube. The digests were purified and ligated with T4 DNA ligase. The ligation was purified and used as template for PCR with primers pSB1NX2insf2 and pSB1NX3insr2. (C) PCR was performed, the product was purified, digested with EcoRI/PstI restriction enzymes and cloned into pSB1C3 pre-digested with the same enzymes. *E. coli* JM109 competent cells were transformed and positive clones selected from plates containing Chloramphenicol.

Figure 2: Enzyme activity in cells expressing cex, cenA, cxnA and non-fused cex+cenA

Exoglucanase activity (A, C & E) is detected by fluorescence of 4-MU following incubation of samples (whole cells or lysates) with MUC. Endoglucanase activity (B, D & F) is detected by the formation of a halo following incubation of samples (whole cells or lysates) with CMC followed by Congo Red stain and destaining with NaCl. A & B shows activity in whole cells; C & D shows activity in cell lysates on agar plates; E & F shows activity from cell lysates following electrophoresis on agarose gels. Enzymes were expressed by cells growing on plates and could also be detected from crude cell lysates run on a 5% horizontal agarose gel. Exoglucanase and endoglucanase activities determined by MUC and Congo Red assays respectively. Vector control (VC) is pSB1C3-BBa_J33207.

Figure 3: The gene fusion retained its property of binding to cellulose

The ability of the chimeric CBM to bind to cellulose was tested by incubating 50 µl of crude lysate with 20 mg/ml avicel at 0°C for 1 hour. The avicel was separated by

centrifugation. MUC assays were performed on aliquots of samples taken before addition of avicel (no avicel), supernatant after binding test and avicel is separated (unbound) and then in a resuspended pellet (bound). The Azo-CMC endoglucanase activity did not work in the presence of avicel. Three biological replicates were performed.

Figure 4: Enzymes were secreted/leaked into the culture supernatant

Secretion/leakage of proteins from the periplasm into the supernatant was investigated by comparing enzyme activities obtained from culture supernatant, cytoplasm and periplasmic fractions. Cytoplasmic and periplasmic fractions were resuspended in equal volumes and diluted to 1× equivalent of extracellular fraction. The activities from all fractions were statistically different ($p<0.01$) for all of the constructs. Three biological replicates were used. Error bars indicate standard error of the mean.

Figure 5: C. freundii expressing the gene fusion grew better and physical destroyed filter paper faster than the non-fused genes

C. freundii with pSB1C3-BBa_J15509 (Cex+CenA1.0) or pSB1C3-BBa_K2523025 (CxnA1.0) or pSB1C3-BBa_J33207 (VC) cultured with the Ford 428 (A: conical flasks & B: McCartney bottles) and the GB003 (C) blotting papers. The cultures were grown with shaking at 200 rpm and monitored daily. In addition, cultures were vortexed for one minute (highest speed on Genie 2 Vortex Mixer, (SLS, UK)) daily. Growth was monitored by performing viable counts on serial dilutions. Three biological replicates were used.

Figure 6: Spots of hydrolysed sugars correspond to cellobiose and cellotriose

Aliquots (1 ml) of the growing cultures were taken at various time intervals and concentrated to about 100 µl in a speedvac. The concentrate was resuspended in 100 µl chlorobutanol after which 5 µl was analysed on silica gel TLC plates. Light arrows point to spots of cellobiose whereas dark arrows point to spots of cellotriose.
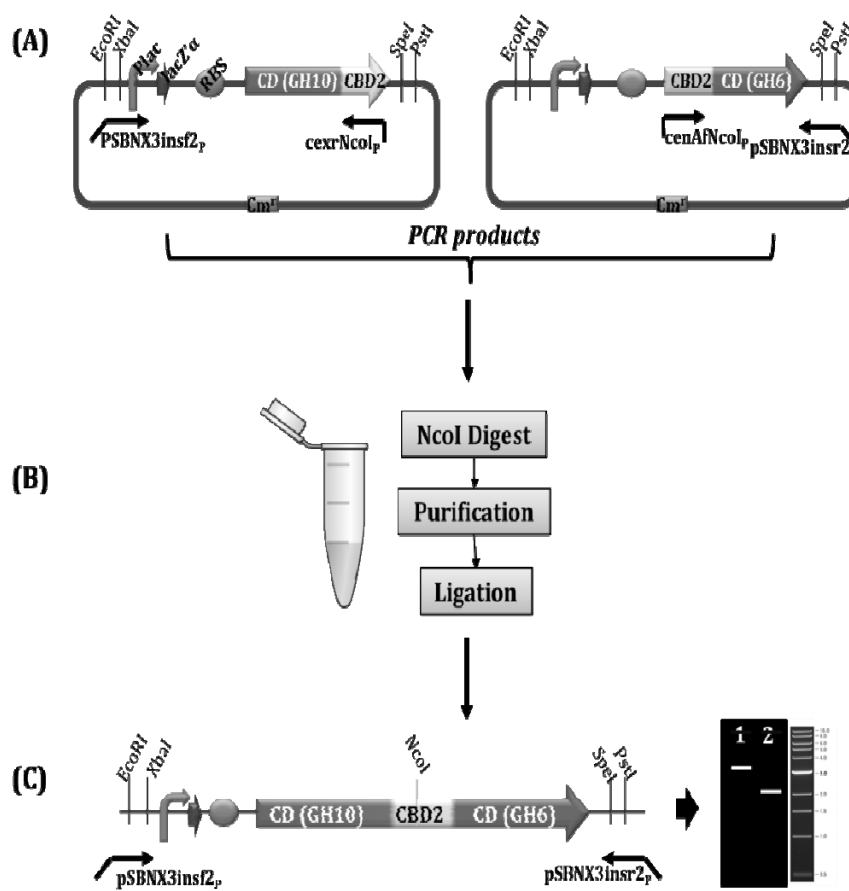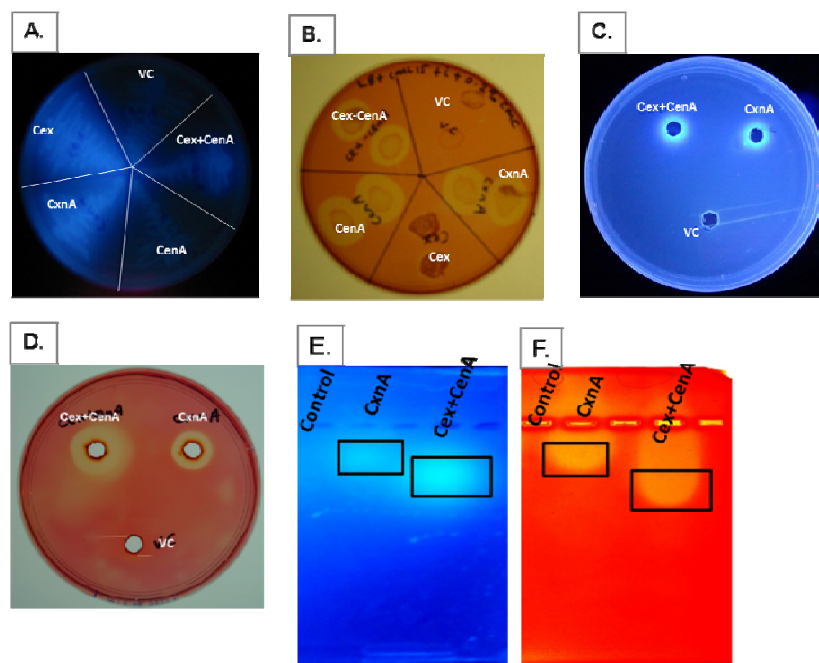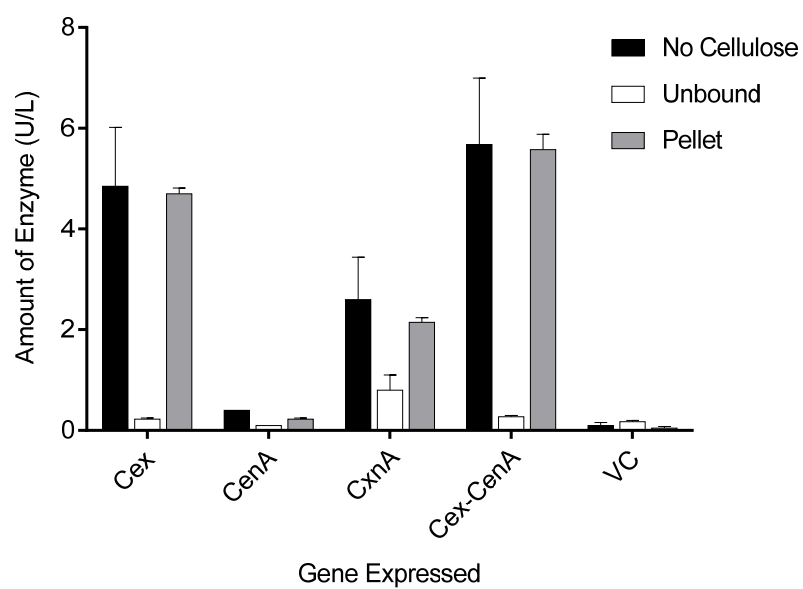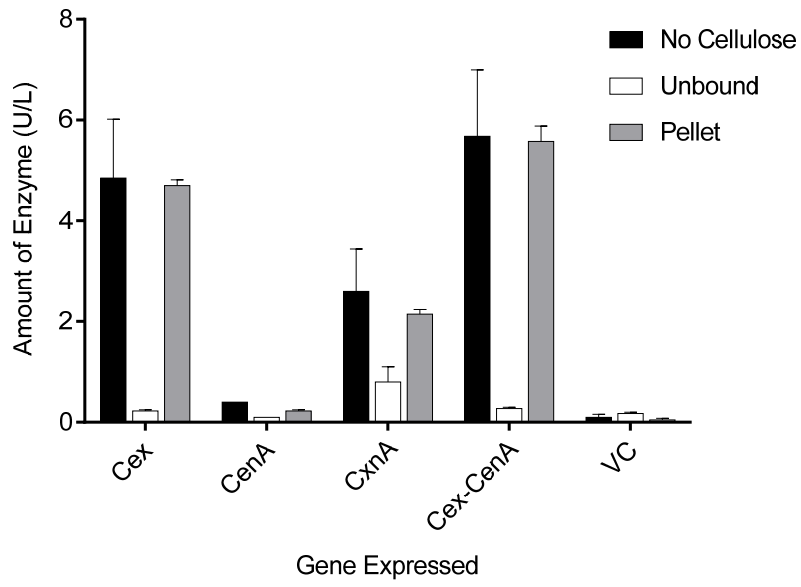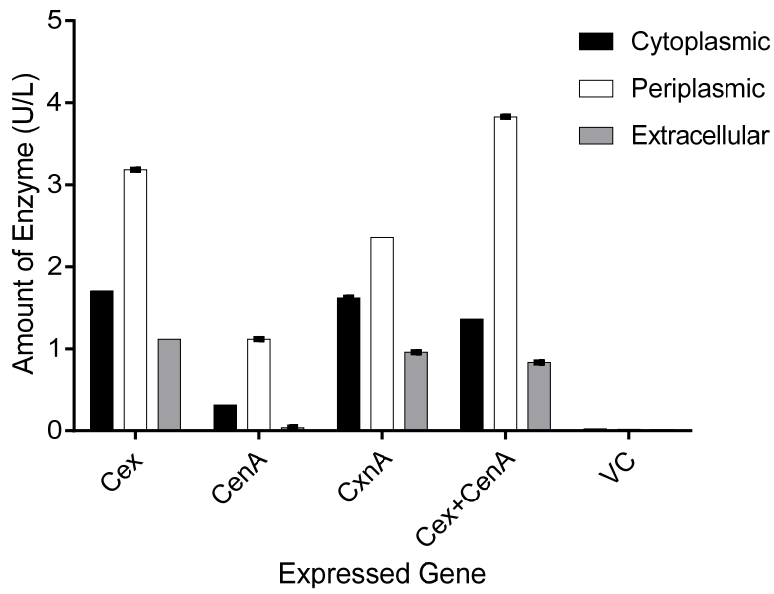
Fig 1

Fig 2

Fig 3

Fig 3

Fig 4
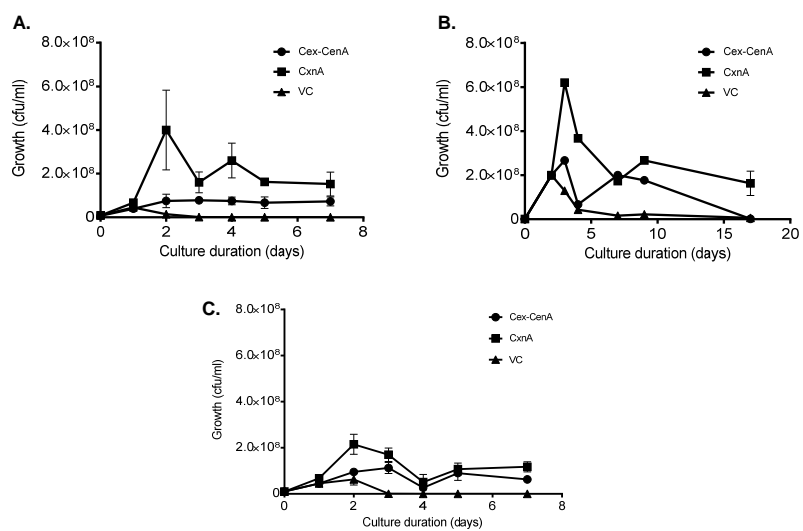
Fig 5
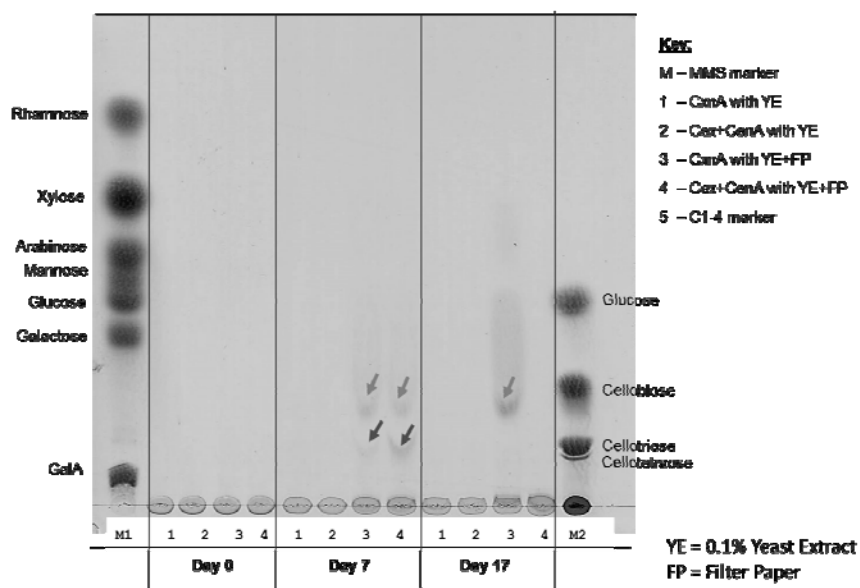
Fig 6

Tables

Table 1: Plasmids, DNA parts and primers used

Plasmids / Parts	Relevant characteristics	Reference
pSB1C3	High copy number BioBrick plasmid, pMB1 origin, Cm^R	[50]
BBa_J33207	*lac* promoter plus sequence encoding N-terminal 77 amino acids of *LacZ*	[50]
BBa_J15001	Strong synthetic ribosome binding site	[50]
BBa_K523015	*Cellulomonas fimi* endoglucanase A (CenA), BBa_J33207, BBa_J15001	[50]
BBa_K523016	*Cellulomonas fimi* exoglucanase (Cex), BBa_J33207, BBa_J15001	[50]
BBa_K523025	fused *cex* and *cenA* with BBa_J33207 and BBa_J15001	This study
BBa_J15509	not-fused *cex* and *cenA* with BBa_J33207 and BBa_J15001	[40]
Primers		
cenAfNcoI	5'-cctgccatggaacggcagcatcccgacc-3'	This study
cexrNcoI	5'-cgttccatggggcgttgcggaccgtcacg-3'	This study
pSB1NX3insf2	5'-aaataggcgtatcacgaggc-3'	This study
pSB1NX3insr2	5'-cagtgagcgaggaagcctgc-3'	This study