



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Characterisation of novel biomass degradation enzymes from the genome of *Cellulomonas fimi*

**Citation for published version:**

Kane, SD & French, CE 2018, 'Characterisation of novel biomass degradation enzymes from the genome of *Cellulomonas fimi*', *Enzyme and Microbial Technology*, vol. 113, pp. 9-17.  
<https://doi.org/10.1016/j.enzmictec.2018.02.004>

**Digital Object Identifier (DOI):**

[10.1016/j.enzmictec.2018.02.004](https://doi.org/10.1016/j.enzmictec.2018.02.004)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Enzyme and Microbial Technology

**Publisher Rights Statement:**

© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

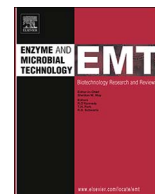
**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.





# Characterisation of novel biomass degradation enzymes from the genome of *Cellulomonas fimi*

Steven D. Kane<sup>\*,1</sup>, Christopher E. French

School of Biological Sciences, University of Edinburgh, Roger Land Building, Edinburgh EH9 3JR, United Kingdom

## ARTICLE INFO

### Keywords:

*C. fimi*  
Arabinofuranosidase  
Thermotolerant xylosidase  
Multifunctional endoxylanase

## ABSTRACT

Recent analyses of genome sequences belonging to cellulolytic bacteria have revealed many genes potentially coding for cellulosic biomass degradation enzymes. Annotation of these genes however, is based on few biochemically characterised examples. Here we present a simple strategy based on BioBricks for the rapid screening of candidate genes expressed in *Escherichia coli*. As proof of principle we identified over 70 putative biomass degrading genes from bacterium *Cellulomonas fimi*, expressing a subset of these in BioBrick format. Six novel genes showed activity in *E. coli*. Four interesting enzymes were characterised further.  $\alpha$ -L-arabinofuranosidase AfsB,  $\beta$ -xylosidases BxyF and BxyH and multi-functional  $\beta$ -cellobiosidase/xylosidase XynF were partially purified to determine their optimum pH, temperature and kinetic parameters. One of these enzymes, BxyH, was unexpectedly found to be highly active at strong alkaline pH and at temperatures as high as 100 °C. This report demonstrates a simple method of quickly screening and characterising putative genes as BioBricks.

## 1. Introduction

Cellulose, the major component of plant cell walls, is the most abundant biopolymer on earth, forming a large part of societal waste [1]. Its structure is made of glucose molecules linked by  $\beta$ -1,4-glycosidic bonds forming long fibres that interact through hydrogen bonds and van der Waal forces to produce microfibrils of crystalline and semi-crystalline cellulose [2]. It's this structure that leads to the majority of the recalcitrance associated with the hydrolysis of cellulose, but the interactions of secondary polysaccharides, termed hemicellulose, also play a major role. For full hydrolysis of cellulose to its composite glucose molecules, the hemicellulose must first be removed. While the structure of the cellulose remains largely the same for all sources, the composition of hemicellulose can vary greatly with different levels of xylan, mannans, arabinose, pectin, ferulic acids and lignin dependent on the plant material [2,3].

In nature the hydrolysis of cellulose requires the concerted effort of a range of enzymes. The accepted model for cellulosic hydrolysis requires the synergistic action of endo  $\beta$ -1,4-glucanases (E.C. 3.2.1.4), exo- $\beta$ -1,4-glucanases (E.C. 3.2.1.74), cellulose 1,4- $\beta$  cellobiosidases (E.C. 3.2.1.91) and  $\beta$ -glucosidases (E.C. 3.2.1.21). The hydrolysis of hemicellulose requires a far greater number of enzymes due to the variability of the compositions and structures. The hydrolysis of xylan for example requires endoxylanase and xylosidase activities for the

backbone, and then enzymes capable of hydrolysing the branched sugars (arabinofuranosidase, galactosidase, glucosidase) and chemical attachments (feruloyl esterase, acetyl xylan esterase). Different sugar backbones and branching schemes obviously require further enzymes (pectate lyase, chitinase, mannanase, mannosidase, arabinanase, fucosidase, etc.) [3,4]. Enzymes can be just as important for hydrolysis of chemically pre-treated biomass as for hydrolysis of raw cellulosic material [5].

These hemicellulose sugars have been largely considered as contaminants blocking the access of enzymes to the glucose stored as cellulose. Often pentose sugars, these are not readily utilised by many common fermenting species such as *Saccharomyces cerevisiae* [6,7]. The removal of these oligosaccharides is usually performed using physical and chemical pre-treatments, selectively solubilising the sugars before enzymatic treatment of the revealed cellulose [3]. This adds a lot of the cost and complexity to the conversion of biomass to fermentable sugars and often results in the production of fermentation inhibitors as by-products [8].

Consolidated bio-processors (CBPs) are organisms engineered to hydrolyse cellulose *and* ferment the released sugars, and are espoused as a solution to these hindrances [9,10]. Synthetic biology uses modular biological units, which can be combined to create artificial networks to perform specific functions [11]. The hydrolysis of cellulosic biomass by a CBP will entail the combination of genes encoding a wide variety of

\* Corresponding author.

E-mail address: [steven.kane@glasgow.ac.uk](mailto:steven.kane@glasgow.ac.uk) (S.D. Kane).

<sup>1</sup> Molecular, Cell and Systems Biology, University of Glasgow, Bower Building, University Avenue, G12 8QQ, United Kingdom.

**Table 1**

A selection of putative functions and known gene products required for cellulose degradation found in the *C. fimi* genome.

Putative Enzyme Function	GH Families	Total Novel	Already Known in <i>C. fimi</i>
Endo 1,4- $\beta$ -glucanase	5, 6, 9, 16, 64, 81	12	CenA, CenB, CenC, CenD
1,4- $\beta$ cellobiohydrolase	6, 48	1	CbhA, CbhB
Endo 1,4- $\beta$ xylanase	10, 11, 43	3	XynB, XynC, XynD, XynE
Xylan 1,4- $\beta$ xylosidase	39, 43	5	
$\alpha$ -L-arabinofuranosidase	43, 51, 62	4	AbfCelf(AfsA)
arabinanase	43	2	
$\beta$ -mannanase/mannosidase	None, 2, 26	2	ManA, ManD
$\alpha$ -amylase	None, 13	9	
Pectate Lyase	None	4	
Chitinase	None, 3, 18, 20	3	NagA, NagB
Acetyl Xylan Esterase	None	3	

polysaccharide degrading enzymes. Synergistic enzymes with multiple functions, which operate at similar optimum temperatures and pH values, from sources easy and cheap to produce should reduce the cost associated with lignocellulosic biofuel. Enzymes from the same source organism are most likely to have these qualities.

Cellulolytic, Gram-positive, mesophilic, facultative anaerobic soil bacterium *Cellulomonas fimi* has been studied for over thirty years due to its efficiency for degrading cellulosic material and ease of culturing [12,13]. Prior to the release of the genome by the US DOE [14], there were 14 cellulolytic enzymes known in *C. fimi* (Table 1). Endoglucanases CenA, CenB and xylanase XynB (Cex) were identified from screening of phage lambda libraries in *E. coli* for activity on CMC and avicel [15–17]. XynC and XynD were similarly identified by screening with oat-spelt xylan [18,19], mannanase Man26A and mannosidase Man2A were identified using azo-carob galactomannan and 4-methylumbelliferyl- $\beta$ -D-mannoside as substrates [20]. N-acetyl glucosamidases NagA and NagB were found using 4-methylumbelliferyl  $\beta$ -N-acetyl-D-glucosaminide [21]. Examination of enzymes secreted by *C. fimi* on different substrates and able to bind to sugar led to identification of CenC [22], CenD [23], CbhA [24], CbhB [25] and XynE (Cfx) [26]. Since the release of the genome five  $\beta$ -glycosidases including one bifunctional  $\beta$ -xylosidase [27] and an  $\alpha$ -L-arabinofuranosidase [28] have been discovered as a direct result of data mining the genome annotations.

The traditional methods used thus far to identify polysaccharide degrading enzymes in *C. fimi* have resulted in a relatively meagre set of enzymes when compared to those predicted for other cellulolytic organism whose genomes are available [29–32]. The further 6 enzymes latterly discovered supports this supposition that there are many useful enzymes still to be discovered in this organism. A simple and rapid means of assessing individual genes is therefore essential. In this paper we report a pilot study for the identification of a range of putative cellulolytic gene products, and the cloning, expression, functional screening and partial characterisation of a subset of these genes using the BioBrick format [11].

## 2. Materials and methods

### 2.1. Chemicals, strains and reagents

All chemicals and reagents were purchased from Sigma-Aldrich unless otherwise stated. *C. fimi* ATCC484 was used as source for all cloned genes. *Escherichia coli* JM109 was used for plasmid construction and gene expression. Primers were synthesised and purchased from Life Technologies.

### 2.2. Gene selection

A keyword search of the uploaded and *C. fimi* genome was carried out to identify putative gene sequences of interest for extraction. DNA was translated into amino acid sequence and searched against databases using BLAST and InterProScan [33] to identify domains and motifs to support the annotation given. From this pool a subset was selected, as proof-of-concept, for cloning and screening based on ease of assay method and their comparability to published and commercially available enzymes.

### 2.3. Cloning and plasmid construction

*C. fimi* genomic DNA was extracted using the Puregene Yeast/Bact. extraction kit (Qiagen) and used as template. PCR cloning of genes was performed using gene specific primers using Kod Xtreme Hot Start DNA Polymerase with GC buffer (Novagen, cat.: 71975). PCR products were purified from solution using Qiagen PCR purification kit (cat.: 28104). Purified product and plasmid Edinbrick1 (pSB1A2-BBa\_J33207, a high copy number ampicillin resistance plasmid with *lac* promoter and *lacZ $\alpha$ '*, available from the Registry of Standard Biological Parts) were digested by restriction enzymes *EcoRI*-HF and *SpeI* (NEB) following the blueprint of the BioBrick construction method RFC 10 [11] and ligated overnight using T4 DNA ligase (NEB). Ligation mixture was used to transform *E. coli* JM109 using the method of Chung et al. for cell preparation [34]. Transformants were plated on Luria agar containing 100  $\mu$ g/ml ampicillin, 40  $\mu$ g/ml X-gal and 90  $\mu$ g/ml IPTG. White colonies, in which the gene of interest had replaced the *P<sub>lac</sub>-lacZ $\alpha$ '*, cassette, were selected for extraction using Qiagen QIAprep Spin Miniprep (cat.: 27104) and plasmid insert confirmed by Sanger sequencing performed by Edinburgh Genomics. Correct constructs were digested using *EcoRI*-HF and *XbaI*, while plasmid pSK1, containing *P<sub>lac</sub>-lacZ $\alpha$ '*-RBS BBa\_B0034 was digested using *EcoRI*-HF and *SpeI*, mixed and ligated. *E. coli* JM109 was transformed and plated as previously described, with blue colonies, in which the *P<sub>lac</sub>-lacZ $\alpha$ '*-RBS, cassette had been incorporated, being selected for plasmid recovery. Sanger sequencing by Edinburgh Genomics was used to confirm the presence of promoter, RBS and gene in correct orientation. *E. coli* containing the correct construct was stored as a glycerol stock at  $-80$  °C.

### 2.4. Expression and activity screening

*E. coli* transformed with the correct final construct were grown overnight in 20 ml LB in the presence of antibiotic and IPTG. Cells were harvested by centrifugation at 7000g for 5 min at room temperature, the supernatant discarded and the cells suspended in 0.5 ml PBS (10 mM  $\text{Na}_2\text{HPO}_4$ , 1.8 mM  $\text{KH}_2\text{PO}_4$ , 137 mM NaCl, 2.7 mM KCl, adjusted to pH7.4 using HCl) and 0.5 ml 50% v/v glycerol and transferred to clean 1.5 ml microcentrifuge tubes pre-chilled on ice. The cells were lysed by sonication through 10 rounds of mixing by inversion, followed by pulsing at 10 kHz for 5 s then incubation on ice for 30 s. The lysate was centrifuged at 14,000  $\times$  g for 10 min to remove cell debris. The soluble protein fraction was transferred to a fresh microcentrifuge tube, kept on ice or stored at  $-20$  °C when not in use.

#### 2.4.1. Assay procedures

A reaction mixture of 15  $\mu$ l 50 mM sodium acetate buffer pH 5.0, 5  $\mu$ l raw cell lysate or purified protein and 5  $\mu$ l of nitrophenyl substrate was routinely used. 2-nitrophenyl  $\beta$ -D-xylopyranoside (ONPX), 2-nitrophenyl  $\beta$ -D-cellobioside (ONPC), 2-nitrophenyl  $\beta$ -D-galactopyranoside (ONPGal) and 4-nitrophenyl  $\alpha$ -L-arabinofuranoside (PNPA) were used to assay for  $\beta$ -xylosidase,  $\beta$ -cellobiosidase,  $\beta$ -galactosidase and  $\alpha$ -arabinofuranosidase activities, respectively. The twenty five microlitre reaction mixture was incubated for 1 h at 37 °C then placed on ice to effectively stop the reaction. Immediately before taking the spectrophotometric reading at 405 nm, an equal volume (25  $\mu$ l) of 1 M  $\text{Na}_2\text{CO}_3$

was added and thoroughly mixed by inversion. All readings were taken using a NanoDrop 2000, using the UV vis setting.

Lysates were assayed for  $\beta$ -glucosidase activity with 4-methylumbelliferyl  $\beta$ -D-glucopyranoside (MUG). Ten microlitres of cell lysate was added to 85  $\mu$ l of 50 mM sodium acetate, pH 5.0 and 5  $\mu$ l of MUG (5 mg/ml in H<sub>2</sub>O). The lysates were then incubated at 37 °C for 1 h and the presence of released methylumbelliferol was measured by absorbance at 348 nm using a NanoDrop 2000.

*E. coli* lysate from cells expressing plasmid with no gene insert were used as negative controls for assays, and where possible a *C. fimi* gene of known enzymatic function expressed in the same fashion as the novel genes was used as a positive control. XynB (Cex) was used as positive control for cellobiosidase and xylosidase assays, *Cytophaga hutchinsonii* BglX (CHU2268, C. K. Liu et al., manuscript in preparation) was used as positive control for  $\beta$ -glucosidase assays.

## 2.5. Protein purification

### 2.5.1. Cell cultures

*E. coli* JM109 expressing the gene of interest were grown in 50 ml volumes of LB containing antibiotic and IPTG overnight. The cells were recovered by centrifugation at 7000  $\times$  g for 5 min and the pellet suspended in 1 ml of Binding Buffer (BB – 25 mM Tris-HCl, pH 8.0, 25% v/v glycerol). The cells were lysed by sonication as previously described.

### 2.5.2. Anion exchange chromatography

All purification steps were performed at 4 °C. A 4 ml column of DEAE-Sepharose was washed, equilibrated and regenerated according to recommended operating procedures. The sonicated cell lysate was loaded on the column and washed through with 3 ml of BB. The flow through (FT) and all subsequent fractions were collected. The column was washed three times with 1 column volume (CV) of BB. The protein was eluted in a stepwise fashion, 1 CV of BB containing NaCl in 0.1, 0.2, 0.3, 0.4 and 0.5 M concentrations unless otherwise stated. Fractions were assayed for enzyme activity as previously described. Elution fractions showing activity were pooled and made up to a concentration of 1 M NaCl and used for the next purification step.

### 2.5.3. Hydrophobic interaction chromatography (HIC)

A 2 ml phenyl-agarose column was washed, equilibrated and regenerated according to recommended operating procedures. The pooled fraction from the previous purification step was loaded onto the column. The column was washed three times with one CV of BB2 (BB plus 1 M NaCl). The proteins were then eluted in a stepwise fashion, with 3\*1 ml of BB with 0.5, 0.1, 0.08, 0.06, 0.04, 0.02 and 0 M concentrations of NaCl. Fractions were assayed for enzyme activity as previously described and all fractions were stored at –20 °C.

### 2.5.4. Cation exchange chromatography

XynF was purified on CM-sepharose after being passed through DEAE-sepharose. The active DEAE fractions were loaded directly to a 1 ml CM-sepharose column. The column was washed 3 times with 1 CV of BB and eluted using 3 times 1 ml BB plus 0.05 M, 0.1 M and 1 M NaCl each. Fractions were assayed for enzyme activity as before.

Purification fractions were analysed by SDS-PAGE, and total protein concentrations estimated by Bradford assay using BSA for the standard curve. After the final purification step if more than one fraction demonstrated enzyme activity, the fraction exhibiting the highest activity:protein ratio was used for characterisation assays.

## 2.6. Enzyme characterisation

### 2.6.1. pH Optima

Enzymes were assayed as previously described but with increasing pH levels using the following buffers: pH 4.0–5.5, 50 mM sodium acetate buffer, pH 6.0–7.5, 50 mM potassium phosphate buffer, pH

8.0–9.0, 20 mM Tris-HCl buffer and pH 9.5–10.5, 50 mM glycine buffer. All assays were performed at 37 °C.

### 2.6.2. Temperature optima

The derived optimum pH for the enzyme was used for all temperature experiments. The assays were performed as previously described but at increasing temperatures from 20 °C–70 °C (or higher if needed) in 5 °C steps using a DNA Engine PCR machine (MJ Research) and 0.2 ml thin walled PCR tubes (Axygen).

### 2.6.3. Effect of supplements

The enzyme specific optimum pH and temperature were used for all subsequent assays. The assays were performed as previously described but with 14  $\mu$ l of buffer. The remaining 1  $\mu$ l was one of the following supplements at a final concentration of 1 mM – CaCl<sub>2</sub>, CoCl<sub>2</sub>, CuSO<sub>4</sub>, FeSO<sub>4</sub>, MgCl<sub>2</sub>, MnCl<sub>2</sub>, ZnSO<sub>4</sub>, DTT, EDTA, glucose, cellobiose, xylose or arabinose. FeSO<sub>4</sub> was made fresh on the day of use as precipitate would form and obfuscate the OD 405 nm readings. All incubations were performed in duplicate with enzyme negative controls.

### 2.6.4. Enzyme kinetics

Enzymes were diluted to a level where activity was still detectable under normal assay conditions. Enzymes were assayed at their optimum pH and temperature with an increasing level of substrate for 10 min. Precipitation prevented high levels of ONPX (> 20 mM) from being used. Enzyme negative controls were used to correct for background noise at high substrate concentrations. All assays were performed in duplicate.

Activity was measured as relative to the maximum activity observed for a reaction minus any enzyme negative background observed. For the purification table (Table 4), standard curves of 2- and 4-nitrophenol at 0, 50, 100, 250, 500, 750 and 1000  $\mu$ M concentrations at pH 5.0 were used to estimate molar absorption coefficients and fraction activities (1192.6 M cm<sup>-1</sup> and 9594.4 M cm<sup>-1</sup>, respectively). A unit is defined as the amount of protein required to release 1  $\mu$ mol of nitrophenol from its bound sugar per minute.

## 3. Results

### 3.1. Putative polysaccharide degrading enzymes

The uploading of the *Cellulomonas fimi* ATC484 genome by the US DOE (NC\_015514.1) has allowed the identification of potential ORFs that putatively encode polysaccharide hydrolysing enzymes further to those already identified. The annotation of the uploaded genome was purely automated and was used as a starting point, with further confirmation of identified potential ORF products being based on manual protein domain and BLAST searches, and the putative genes named after the most frequent hit in the top 20. By this method all previously known enzymes were identified along with over 80 other putative polysaccharide degrading enzymes including a further 7 endo-glucanases, 3 putative endo-xylanases, 6  $\beta$ -xylosidases, and 2 mannanases (Table 1). There appears to be a full complement of hemicellulolytic genes, necessary for the hydrolysis of the highly variable hemicellulose component of plant biomass including 5  $\alpha$ -L-arabinofuranosidases, 2 arabinanases and 11  $\beta$ -galactosidases, all functions largely unreported in *C. fimi*. Four of the five arabinofuranosidases have no predicted GH family, however they group closely to other characterised GH 51 enzymes when aligned using T-coffee [35] and AfsA (Abfcelf) is now reported as such [28]. These putative genes potentially give *C. fimi* the ability to utilise a wide variety of sugar sources from plants, fungi, insects and animals.

Many of the putative proteins are predicted to have domains for binding to simple sugars, such as the ricin-B-lectin and concanavalin-A like domains. Four carbohydrate binding proteins, with no detected catalytic domain, were also found (Celf\_0270, 0403, 1913, 2856). Ten

**Table 2**  
List of *C. fimi* genes cloned and screened in this study.

Gene Name	DOE Annotation	Genome Position	GH Family, Domains	Predicted Function by BLAST	Predicted localisation
<i>afsA</i>	$\alpha$ -L-arabinofuranosidase domain containing protein	Celf_3321	probable GH51, $\alpha$ -L-arabinofuranosidase C-terminal	arabinofuranosidase	cytoplasmic
<i>afsB</i>	$\alpha$ -L-arabinofuranosidase domain containing protein	Celf_0903	probable GH51, $\alpha$ -L-arabinofuranosidase C-terminal	arabinofuranosidase	cytoplasmic
<i>afsC</i>	$\alpha$ -L-arabinofuranosidase domain containing protein	Celf_3249	probable GH51, $\alpha$ -L-arabinofuranosidase C-terminal	arabinofuranosidase	cytoplasmic
<i>bxyC</i>	Xylan 1,4- $\beta$ xylosidase	Celf_3643	GH43, concanavalin A-like glucanase	xylosidase	unknown
<i>bxyD</i>	Glycoside Hydrolase Family 43	Celf_1745	GH43, concanavalin A-like glucanase	xylosidase	unknown
<i>bxyE</i>	Glycoside Hydrolase Family 43	Celf_0899	GH43, endo-1,5-arabinanase domain	xylosidase	unknown
<i>bxyF</i>	Glycoside Hydrolase Family 39	Celf_1744	GH39, alpha-beta hydrolase fold 3	xylosidase	cytoplasmic
<i>bxyH</i>	Xylan 1,4- $\beta$ -xylosidase	Celf_3270	GH39	xylosidase	unknown
<i>celA</i>	Cellulase	Celf_0376	GH5, DUF291, Immunoglobulin-like E-set domain, GH5 endoglucanase B, signal peptide	endoglucanase	unknown, not cytoplasmic
<i>celD</i>	Glycoside Hydrolase Family 9	Celf_0045	GH9, CBMII, Fibronectin III, six hairpin glycosidase like domain	endoglucanase	extracellular
<i>celE</i>	Glycoside Hydrolase Family 9	Celf_1705	GH9, CenC-like CBM (CBM_4_9), 6 hairpin glycosidase like, galactose binding domain like, immunoglobulin like fold, transmembrane regions, signal peptide	cellobiohydrolase/ glucanase	extracellular
<i>celF</i>	Glycoside Hydrolase Family 9	Celf_1481	GH9, 6 hairpin glycosidase like domain	endoglucanase	extracellular
<i>celN</i>	Ricin-B-lectin	Celf_2403	GH5, ricin-B-lectin domain	cellulase	extracellular
<i>xynF</i>	Endo-1,4- $\beta$ -xylanase	Celf_0088	GH10, CBMII, TAT signal peptide	xylanase	extracellular
<i>xynG</i>	Glycoside Hydrolase Family 10	Celf_1729	GH10	xylanase	extracellular

Domains were identified using Interproscan [33], extracellular localisations by PSORTB V3 [52], named based on top hits in blast: afs – arabinofuranosidase, bxy –  $\beta$ -xylosidase, cel – endoglucanase/cellulase, xyn – xylanase.

putative genes (Celf\_0006, 0651, 0718, 0720, 0735, 1024, 3189, 3746), arabinogalactan-endo- $\beta$ -galactosidase *abgA* (Celf\_3434) and lectin domain containing Celf\_3497 were found with LPXTG domains, which is known to anchor to cell surfaces. Also of note are the gene clusters and putative operons which are evident. By looking at the surrounding genes of the putative polysaccharide hydrolases, 30 of these genes are flanked with other genes of related function often less than 100 nucleotides from either the start or the end of the gene. These clusters almost always contain extracellular solute binding proteins and transmembrane transport proteins, presumably for the import of specific released sugars into the cell. These clusters have been observed in other species and may be a vital link in the construction of an efficient CBP [14,32,36,37].

### 3.2. Cloning, expression and activity screening

Table 2 lists the genes cloned as BioBricks from *C. fimi* in this study. These were initially selected as the putative genes coded for industrially useful enzymes, could be easily assayed for and compared to available enzymes. Initially crude *E. coli* cell lysates containing the soluble protein fraction after IPTG induction were used to assay for cellulolytic activity, summarised in Table 3. None of the genes cloned and screened appeared to produce any  $\beta$ -endoglucanase enzymes, with all assayed lysates having levels below that of the negative control, with RBB-CMC as substrate. With RBB-xylan as substrate, XynF proved positive with an absorbance reading well above background and AfsA, AfsB and BxyF showing possible low levels of activity.

Fig. 1A and B shows the specific activities of the cell lysates on specific substrates. AfsA and AfsB showed activity above background levels with PNPA as substrate and activity on ONPX. BxyF proved highly active against ONPX, XynF displayed activity with both ONPC and ONPX as substrates and BxyH displayed activity against ONPX. These initial screens were used to rapidly detect activity on substrates of interest. The other genes showed no detectable activity. ONP-Galactoside was used to assay for  $\beta$ -galactosidase activity as an internal promoter activity control, and was found to be of near identical levels in all cell lysates, indicating that  $P_{lac}$  expression was induced to similar levels in all cases. Negative results therefore, may be due to poor translation of a given protein in *E. coli*, or due to lack of activity of the

enzyme under the parameters used.

All activities observed in the initial screens were confirmed (Fig. 1B). AfsA and AfsB have both arabinofuranosidase and xylosidase activities, with AfsA at this temperature and pH being more active with ONPX than PNPA and AfsB being roughly equal. BxyF and BxyH show strong activity with ONPX as substrate but BxyF is far more active in these assay conditions. XynF displays dual activities against ONPC and ONPX with a greater degree of activity shown towards ONPC.

### 3.3. Characterisation of enzyme activities

A subset of the enzymes were purified and characterised further (Tables 3 and 4). AfsA and AfsB showed similar activity patterns, however AfsB displayed higher activity with PNPA as substrate so was selected for purification. Enzymes were purified as described in the methods section, and their main unique activities investigated (arabinofuranosidase for AfsB, xylosidase activity for BxyF and BxyH, and cellobiohydrolase activity for XynF).

#### 3.3.1. Optimum pH

AfsB, BxyF and XynF displayed apparent maximal activity in the neutral range of pH 6–7, 6 and 6, respectively. Outwith those pH values they differ slightly in their tolerances. AfsB activity drops significantly below pH 4.5 or above pH 8, with no activity recorded for pH 4 or pH 9.5 and above, whereas BxyF still retains around 20% of activity at these values. XynF also remains fairly active with nearly 60% activity at pH 4 and 10–20% activity between pH 8.5 and 10.5. BxyH activity was greatly affected by the differing buffers making determining the optimal pH problematic. It was clear however that BxyH activity greatly increased at the alkaline limit of a buffer, and greatest activity was measured at pH 9, which was used for subsequent assays.

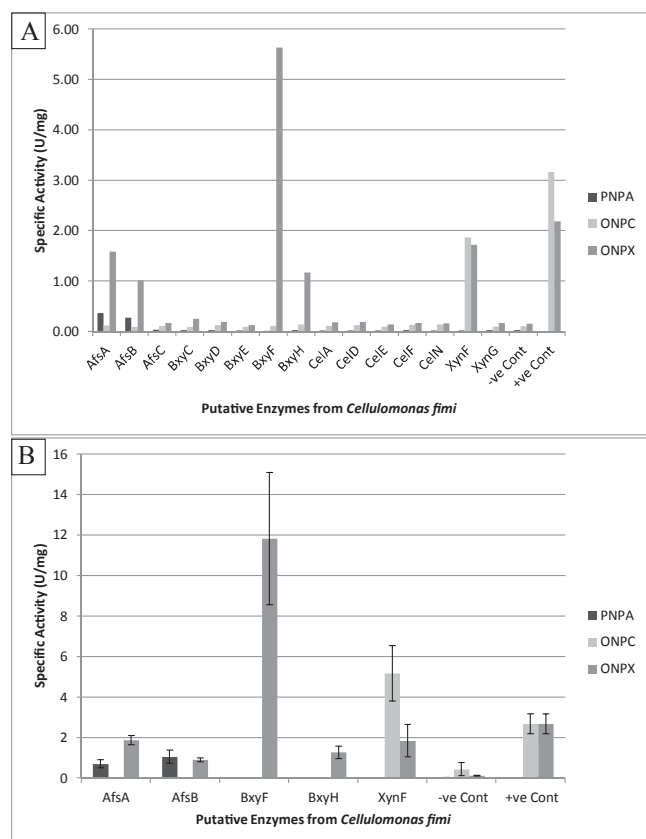
#### 3.3.2. Optimum temperature

Optimum temperatures were determined using the apparent optimum pH environments and are shown in Fig. 2. AfsB (2A) and BxyF (2B) both show 45 °C optima, in the mesophilic range, which might be expected for enzymes originating from *C. fimi*. Their operating temperature range is fairly narrow, with less than 80% of activity detected for AfsB below 30 °C and above 45 °C. BxyF shows less tolerance for

**Table 3**Activities of *E. coli* crude cell lysates and after purification (a) against substrates listed (see materials and methods).

Gene (Product putative name)	PNPA	ONPC	ONPX	RBB-CMC	RBB-Xyl	Substrate for Optimal Activity <sup>a</sup>	Temperature (°C), pH optima <sup>a</sup>	V <sub>max</sub> (U/mg), K <sub>M</sub> (mM) <sup>a</sup>
Celf_3321 (AfsA)	+	–	+	–	+			
Celf_0903 (AfsB)	+	–	+	–	–	PNPA	45, 6.5	133, 6.8
Celf_3249 (AfsC)	+	–	–	–	–			
Celf_3643 (BxyC)	–	–	–	–	–			
Celf_1745 (BxyD)	–	–	–	–	–			
Celf_0899 (BxyE)	–	–	–	–	–			
Celf_1744 (BxyF)	–	–	+	–	+	ONPX	45, 6.0	326, 1.7
Celf_3270 (BxyH)	–	–	+	–	–	ONPX	80, 9.0	NM
Celf_0376 (CelA)	–	–	–	–	–			
Celf_0045 (CelD)	–	–	–	–	–			
Celf_1705 (CelE)	–	–	–	–	–			
Celf_1481 (CelF)	–	–	–	–	–			
Celf_2403 (CelN)	–	–	–	–	–			
Celf_0088 (XynF)	–	+	+	–	+	ONPC, ONPX, RBB-xyl	60, 6.0 50, 4.5 40, 5.5	10151, 87 NM
Celf_1729 (XynG)	–	–	–	–	–			

+ = activity detected; – = no detected activity; NM = not measured.



**Fig. 1.** Initial screening data for genes expressed in *E. coli* with A) PNPA, ONPC and ONPX as substrates. Negative (–ve) control for all substrates, plasmid only; Positive (+ve) control for ONPC and ONPX only was Cex. B) Biological repeats of cell lysates displaying positive activity. Mean activities of n = 2 independent biological repeats and  $\pm$  standard error are shown.

colder temperatures, with less than 80% of activity observed below 40 °C, but more tolerance to slightly higher temperatures with 84% activity observed at 50 °C. XynF (2D) with ONPC as substrate displays an optimum temperature at around 60 °C, but with 96% still retained at 65 °C. A steep decline is then observed with as little as 17% activity at 70 °C. BxyH (2C) operates optimally at 80 °C and shows high activity beyond this with 95% of functionality still retained at even 100 °C, far outside what may be expected from a mesophilic organism. Indeed, at 45 °C and 60 °C, the optima for AfsB, BxyF and XynF, the activities of

BxyH are only 8% and 50%, respectively.

### 3.3.3. Effect of metal ions and sugars

As can be seen in Fig. 3 there are seemingly no universally beneficial ions for these enzymes with the most pronounced effects being negative.

At the 5% significance level only Ca<sup>2+</sup> and Fe<sup>2+</sup> ions had a positive effect, increasing the activity seen for BxyF by 20% and XynF by 45%, respectively. However, Fe<sup>2+</sup> had a detrimental effect on BxyF, almost completely inhibiting activity, and enough of an inhibitory effect on BxyH to be significant. The other metal ions tested had either no significant effect or were highly detrimental. Cu<sup>2+</sup> ions completely disable AfsB and BxyF and significantly reduce the activity of XynF and BxyH. Co<sup>2+</sup> ions also significantly inhibit the activities of AfsB (reduced to < 25% activity), BxyF (< 25% activity) and BxyH (< 65% activity), with XynF not being significantly affected. Zn<sup>2+</sup> is another apparent inhibitor of these enzymes with AfsB, BxyF and BxyH all being significantly inhibited, completely in the case of BxyF. BxyH was significantly inhibited by all the metal ions tested except for Ca<sup>2+</sup>.

DTT had no significant ill effects on these enzymes suggesting disulfide bonds are not crucial for enzyme activity or structure. EDTA also had no significant effect suggesting that certain metal ion co-factors are not necessary for enzymatic function. None of the sugars assayed showed any significant effect on enzyme activity. This suggests that at the 1 mM concentration level competitive inhibition of these enzymes is not occurring to a significant degree, but may still inhibit activity at higher concentrations.

### 3.3.4. Enzyme kinetics

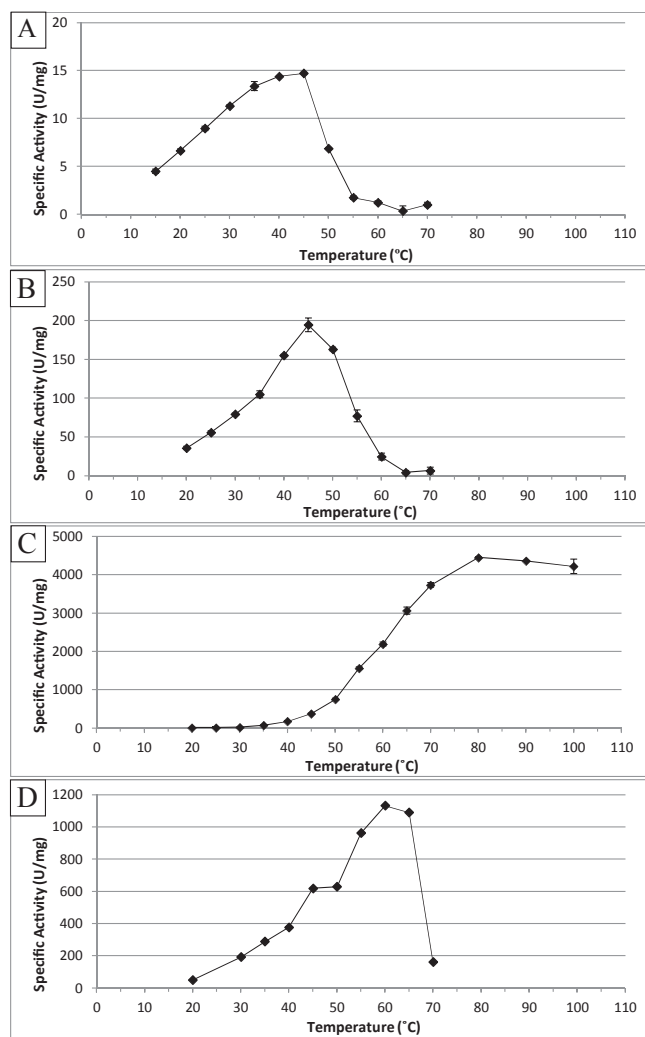
The solubility of ONPX prevented the kinetics of BxyH from being determined as velocity was still increasing at 20 mM concentration, after which ONPX would precipitate. V<sub>max</sub> and K<sub>M</sub> were determined using Eadie-Hofstee plots [38] minus outliers when linearly transformed. AfsB was determined to have a V<sub>max</sub> of 133 U/mg and K<sub>M</sub> of 6.8 mM on PNPA (Fig. 4A). For BxyF (ONPX, Fig. 4B), assays at 20 mM and above showed non-linearity, probably due to precipitation of ONPX substrate as was also observed for BxyH. V<sub>max</sub> was determined as 326 U/mg and K<sub>M</sub> 1.7 mM using the remaining data points. XynF (ONPC, Fig. 4C) was determined to have a V<sub>max</sub> of 8942 U/mg and K<sub>M</sub> of 58 mM.

## 4. Discussion

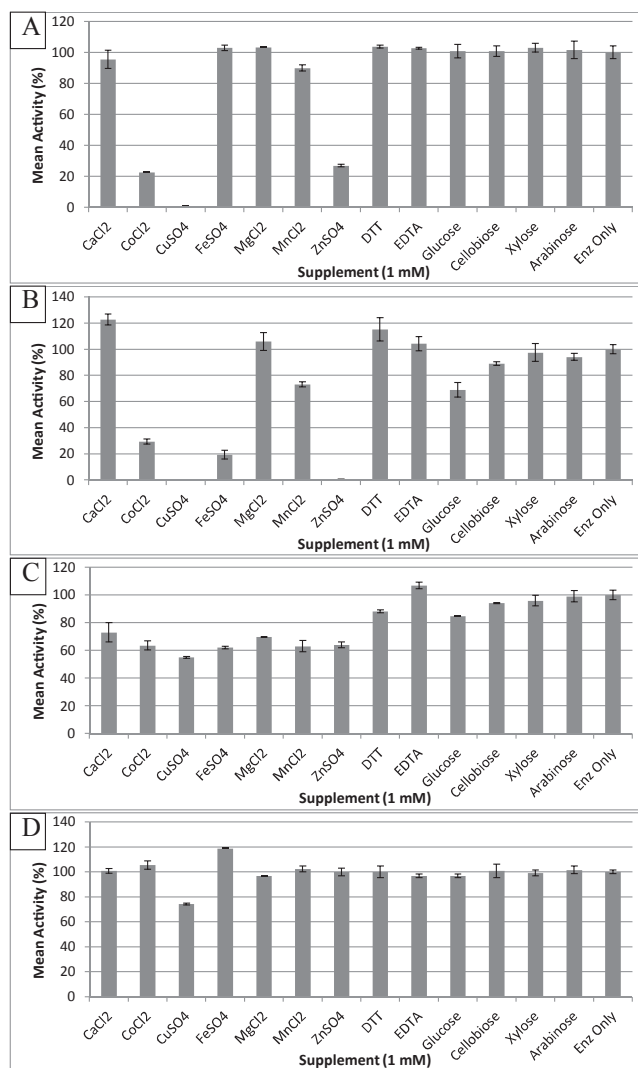
The main aim of this research was to identify novel enzymes active in *E. coli* from a list of putative genes annotated in the *C. fimi* genome.

**Table 4**Fold purifications and yields of target enzymes in their first and final purification steps from crude cell lysates of *E. coli*.

Enzyme	Elution fraction	Elution Volume (ml)	Total Protein (mg)	Total Activity (Units*ml)	Specific Activity (Units/mg)	Fold Purification
AfsB	DEAE 0.3 M NaCl	4	2	13.2	26.4	14.8
	HIC 0.1 M NaCl_1	1	0.3	4.4	17.1	9.6
BxyF	DEAE 0.4 M NaCl	6	7.5	47.6	38.1	2.8
	HIC 20 mM NaCl_3	1	0.1	12.6	167.7	12.1
XynF	DEAE FT	4	3	7.3	9.8	3.6
	CM 0.5 M NaCl	1	0.1	10.8	269.7	100.1
BxyH	DEAE 0.5–0.3 M NaCl	8	2.6	1.4	4.4	5.7
	HIC 0.1 M NaCl	2	0.2	2.4	31.7	40.9

**Fig. 2.** Temperature optima of the purified *C. fimi* enzymes AfsB (A), BxyF (B), BxyH (C), XynF (D). Mean activities ( $n = 2$ ) of independent biological repeats are shown with  $\pm$  SE displayed.

This paper has shown that the BioBrick format can be used to screen for novel gene functionality in *E. coli*, but could just as easily work for any organism of choice by using a host-specific BioBrick compatible plasmid. The standardised nature of the BioBrick format (i.e. coding sequence only) allows for the automation of both primer design and subsequent PCR in plate form up to 384 genes simultaneously, with the rate-limiting step then becoming plasmid construction and transformation. Once transformants are generated, crude cell lysates are easily obtainable and can be used to screen a wide range of substrates through automation using 96–384 well plates to quickly identify genes of interest, colorimetrically in this paper, with results being widely comparable as the only difference between constructs is the gene ORF.

**Fig. 3.** Mean activities ( $n = 2$ ) of independent biological repeats of enzymes under optimal assay conditions in the presence of varying buffer supplements, all of 1 mM concentration. Enz only is the purified enzyme under assay conditions with no supplement.  $\pm$  standard error is shown.

The modular format of the BioBrick allows genes of interest to be easily swapped into other BioBrick formatted plasmids enabling rapid construction of expression cassettes for a range of host organisms or the construction of multi-part plasmids. The BioBrick therefore has the potential to be a useful tool for the screening of large sets of comparably expressed genes in multiple host environments to increase the number of experimentally verified and correctly annotated enzymes within the databases used for automated annotation, increasing accuracy. With the rapid and increasing number of genomes available a standardised scalable format for functional screening becomes essential.

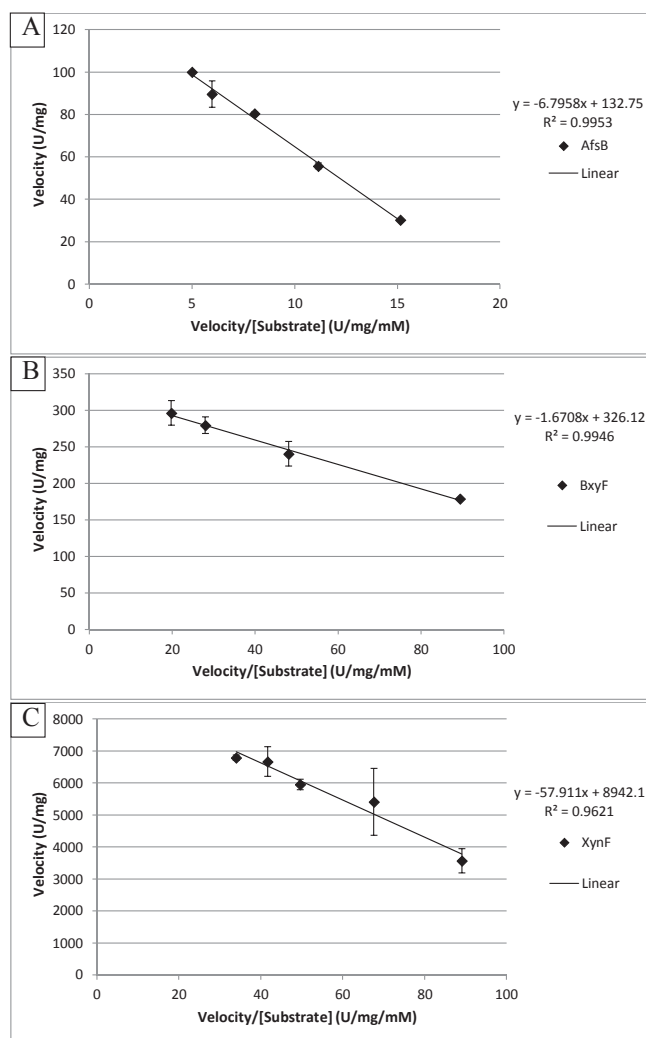


Fig. 4. Eadie-Hofstee plots of enzyme velocity (U/mg) as a factor of substrate concentration (mM). A) AfsB, PNPA; B) BxyF, ONPX; C) XynF, ONPC. Error bars of  $\pm$  SE,  $n = 2$  independent biological repeats.

To date only twenty cellulolytic enzymes have been identified in *C. fimi*. However, similar to the putative gene numbers of other sequenced cellulolytic species [29,31,32,39,40], there was found to be in excess of 90 ORFs coding for putative genes with polysaccharide degrading enzymes as products (Table 1). Further enzymes for degrading simpler sugars include 4  $\alpha$ -galactosidases, 2  $\alpha$ -glucuronidases and 1  $\alpha$ -glucosidase. Almost all have gone unidentified due to the screening methods which generally centred on identifying functionally specific enzymes secreted when grown on a specific and narrow range of simple carbon sources.

Wakarchuk et al. recently published a mass spectrometry secretome analysis of *C. fimi* when grown on CMC or xylan which showed some interesting results [41]. They generated a list of 37 proteins present in the media, excluding any they thought to be intracellular. Six of the previously characterised proteins were identified (CenA, CenD, CbhA, XylC, Cex, and Man26A). They denoted Celf\_0088 as being XynE (Cfx) (DQ146941 [26]), but translation of the DNA sequence shows significant differences in amino acid sequence for the first 178 AAs of the protein, following which the sequences are identical. This is the ORF we have named XynF in this paper to avoid confusion. Of the other hits they found AfsA (Celf\_3321) was induced with both CMC and xylan and BxyF (Celf\_1744) which was induced on xylan only. AfsC (Celf\_3249) was also found in low levels for both carbon sources. This indicates that these enzymes are biologically significant, and validates the approach

we have taken.

Analysis of the genome has given a more accurate representation of the cellulolytic ability of *C. fimi* but can say nothing of the biochemistry of these proteins, or if they are even real. Genomes are automatically annotated based on other automatically annotated genomes, with little basis in experimental evidence, resulting in a “percolation of errors” [42]. Eleven of our fifteen cloned genes had no characterised enzymes in the top 20 returned hits after a BLAST search. A rapid way to test the most interesting of the annotated genes would greatly enhance the accuracy of annotated genomes and increase the range of potentially industrially useful cellulolytic enzymes.

Of the genes cloned 6 were active against the substrates assayed for (AfsA, AfsB, AfsC, BxyF, XynF and BxyH). This low number could be due to several factors. *C. fimi* is a high G + C (75%) organism, whereas the G + C content of *E. coli* JM109 is much lower (51%). The genes that were assayed as negative may not have been functionally expressed, or expressed at low levels due to the presence of rarer codons truncating the product or causing a bottleneck in translation [43]. The number of active enzymes increased when using an alternative host species *Citrobacter freundii* (supplementary figure SF7). CelD for example was positive for activity on ONPC and ONPX, while BxyC, CelE and XynG became detectably active on ONPX. Similar results were seen by [44], suggesting *C. freundii* may be a superior test system for putative biomass degrading enzymes. The reason for this was not clear nor was it robustly explored further. Another possibility is that the gene products are active against substrates not assayed, which would suggest that the genes may have been incorrectly annotated. It may also be that the temperature and pH conditions were not ideal for these particular enzymes to function, or a mix of all of the above.

Three arabinofuranosidases were identified with AfsA and AfsB being strongly active but AfsC exhibiting very low activity, but still marginally detectable (Table 3). AfsB is a typical bacterial arabinofuranosidase in that it is optimally active at neutral pH 6, as opposed to a more acidic pH as is characteristic of fungal arabinofuranosidases [45]. AfsB's amino acid sequence most closely groups with GH51 by sequence alignment searching. Its optimal values are consistent with those of Yang et al., who report AbfCelf (AfsA) as belonging to GH51, having an optimal temperature of 40 °C, optimal pH of 6.0,  $V_{max}$  of 112.4 U/mg and  $K_M$  of 1.5 mM using PNPA as substrate [28].

The two GH39  $\beta$ -xylosidases, BxyF and BxyH, are the first of this family of enzymes to be cloned and characterised from *C. fimi*. A bifunctional  $\beta$ -xylosidase/ $\alpha$ -L-arabinofuranosidase, active on the corresponding PNP-substrates, has been reported for gene product celf\_2726, stating optima of 50 °C and pH of 7.0, but this belongs to GH family 3 [27]. BxyF most closely resembles those that have been previously published, of which there are only 10 GH39 enzymes from bacterial sources characterised according to CAZy [46]. Its optima of pH 6.0 and 45 °C are consistent with the majority of *C. fimi* enzymes identified thus far and a  $\beta$ -xylosidase from *Cellulomonas uda* [47].

Perhaps of most interest is BxyH. Its optimum temperature of 80 °C, with over 90% activity seen as high as 100 °C, would make it of particular interest for industrial applications. These values are higher than those reported for the GH39 xylosidases previously mentioned, and is more akin to temperatures for  $\beta$ -xylosidases from thermophilic bacteria such as *Thermotoga maritima* [48], 90 °C, *Bacillus thermanarcticus* [49], 70 °C and *Geobacillus pallidus* [50], 70 °C to name but a small number. This enzyme has so far been missed by all previous identification methods. Our simple screening technique was able to identify this interesting enzyme. No kinetic data was produced, though it is likely we were not using its preferred substrate, concentrating on a relatively slim list of artificial chemicals for activity screening which could easily be expanded in the future.

The multifunctionality of XynF and being of GH family 10, allows for comparison with the enzymes XynB and XynE from *C. fimi*. XynB and XynE have similar activity ranges to XynF with optimal pH values of 6 and 7, respectively, and optimum temperatures of 40 °C [26,51].



The enzyme kinetics are not comparable as they were determined using different substrates, but the closeness of general characteristics given XynE (Cfx) and XynF maybe the same enzyme further validates this methodology.

## 5. Conclusions

We describe the characterisation of 4 novel enzymes from *C. fimi* cloned as BioBricks: an arabinofuranosidase (AfsB, optima pH 6.5, 45 °C,  $V_{max}$  133 U/mg,  $K_M$  6.8 mM), two beta-xylosidases (BxyF, pH 6.0, 45 °C,  $V_{max}$  326 U/mg,  $K_M$  1.7 mM; BxyH pH 9.0, 80 °C, undetermined) and an endoxylanase (XynF pH 6.0, 60 °C,  $V_{max}$  8942 U/mg,  $K_M$  58 mM). The BioBrick format proves itself a quick method of cloning and testing putative genes with the added benefit of parts being easily assembled into larger networks, necessary for the construction of an organism capable of true consolidated bioprocessing.

## Conflicts of interest

None.

## Acknowledgements

This work was supported by the Biotechnology and Biological Sciences Research Council (Grant numbers BB/D526210/1 and BB/F017073/1). BglX CHU2268 was kindly donated by Chao-Kuo Liu.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.enzmictec.2018.02.004>.

## References

- L.P. Wackett, Biomass to fuels via microbial transformations, *Curr. Opin. Chem. Biol.* 12 (2) (2008) 187–193.
- P. Harris, B. Stone, Chapter 4: chemistry and molecular organisation of plant cell walls, in: M. Himmel (Ed.), *Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy*, Blackwell Publishing, 2008, pp. 61–93.
- F.M. Girio, C. Fonseca, F. Carvalho, L.C. Duarte, S. Marques, R. Bogel-Lukasik, Hemicelluloses for fuel ethanol: a review, *Bioresour. Technol.* 101 (13) (2010) 4775–4800.
- R. Carapito, C. Carapito, J.M. Jeltsch, V. Phalip, Efficient hydrolysis of hemicellulose by a *Fusarium graminearum* xylanase blend produced at high levels in *Escherichia coli*, *Bioresour. Technol.* 100 (2) (2009) 845–850.
- A.S. Meyer, L. Rosgaard, H.R. Sørensen, The minimal enzyme cocktail concept for biomass processing, *J. Cereal Sci.* 50 (3) (2009) 337–344.
- B.C. Saha, Hemicellulose bioconversion, *J. Ind. Microbiol. Biotechnol.* 30 (5) (2003) 279–291.
- D.C. la Grange, R. den Haan, W.H. van Zyl, Engineering cellulolytic ability into bioprocessing organisms, *Appl. Microbiol. Biotechnol.* 87 (4) (2010) 1195–1208.
- H.B. Klinken, A.B. Thomsen, B.K. Ahring, Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass, *Appl. Microbiol. Biotechnol.* 66 (1) (2004) 10–26.
- L.R. Lynd, W.H. van Zyl, J.E. McBride, M. Laser, Consolidated bioprocessing of cellulosic biomass: an update, *Curr. Opin. Biotechnol.* 16 (5) (2005) 577–583.
- C.E. French, Synthetic biology and biomass conversion: a match made in heaven? *J. R. Soc. Interface* 6 (2009) 12.
- T.F.J. Knight, Indempotent Vector Design for Standard Assembly of Biobricks, (2003) <http://hdl.handle.net/1721.1/21168>.
- A. Balow, Chapter 59, in: A. Balow, H. Truper, M. Dworkin, W. Harder, K.-H. Schleifer (Eds.), *The Prokaryotes*, Springer, Verlag, 1992, pp. 1323–1329.
- O.M. Poulsen, L.W. Petersen, Growth and enzyme-production of *Cellulomonas* sp ATCC-21399 on microcrystalline cellulose – effects of increasing concentration of a mineral medium, *Appl. Microbiol. Biotechnol.* 30 (5) (1989) 535–539.
- M.R. Christopherson, G. Suen, S. Bramhacharya, K.A. Jewell, F.O. Aylward, D. Mead, P.J. Brumm, 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* 8 (1) (2013).
- W.K.R. Wong, B. Gerhard, Z.M. Guo, D.G. Kilburn, R. Anthony, J. Warren, R.C. Miller, Characterization and structure of an endoglucanase gene CenA of *Cellulomonas-fimi*, *Gene* 44 (2–3) (1986) 315–324.
- N.R. Gilkes, M.L. Langsford, D.G. Kilburn, R.C. Miller, R.A.J. Warren, Mode of action and substrate specificities of cellulases from cloned bacterial genes, *J. Biol. Chem.* 259 (16) (1984) 455–459.
- B. Moser, N.R. Gilkes, D.G. Kilburn, R.A.J. Warren, R.C. Miller, Purification and characterization of endoglucanase-C of *Cellulomonas-fimi*, cloning of the gene, and analysis of in vivo transcripts of the gene, *Appl. Environ. Microbiol.* 55 (10) (1989) 2480–2487.
- J.H. Clarke, K. Davidson, H.J. Gilbert, C. Fontes, G.P. Hazlewood, A modular xylanase from mesophilic *Cellulomonas fimi* contains the same cellulose-binding and thermostabilizing domains as xylanases from thermophilic bacteria, *FEMS Microbiol. Lett.* 139 (1) (1996) 27–35.
- J.H. Clarke, J.I. Laurie, H.J. Gilbert, G.P. Hazlewood, Multiple xylanases of *Cellulomonas fimi* are encoded by distinct genes, *FEMS Microbiol. Lett.* 83 (3) (1991) 305–310.
- D. Stoll, H. Stalbrand, R.A.J. Warren, Mannan-degrading enzymes from *Cellulomonas fimi*, *Appl. Environ. Microbiol.* 65 (6) (1999) 2598–2605.
- C. Mayer, D.J. Vocadlo, M. Mah, K. Rupitz, D. Stoll, R.A.J. Warren, S.G. Withers, Characterization of a beta-N-acetylhexosaminidase and a beta-N-acetylglucosaminidase/beta-glucosidase from *Cellulomonas fimi*, *FEBS J.* 273 (13) (2006) 2929–2941.
- J.B. Coutinho, B. Moser, D.G. Kilburn, R.A.J. Warren, R.C. Miller, Nucleotide-sequence of the endoglucanase-C gene (CenC) of *Cellulomonas fimi*, its high-level expression in *Escherichia coli*, and characterization of its products, *Mol. Microbiol.* 5 (5) (1991) 1221–1233.
- A. Meinke, N.R. Gilkes, D.G. Kilburn, R.C. Miller, R.A.J. Warren, Cellulose-binding polypeptides from *Cellulomonas-fimi* – endoglucanase-D (CenD), a family beta-1,4-glucanase, *J. Bacteriol.* 175 (7) (1993) 1910–1918.
- A. Meinke, N.R. Gilkes, E. Kwan, D.G. Kilburn, R.A.J. Warren, R.C. Miller, Cellobiohydrolase-A (CbhA) from the cellulolytic bacterium *Cellulomonas-fimi* is a beta-1,4-exocellobiohydrolase analogous to *Trichoderma-reesei* Cbh-II, *Mol. Microbiol.* 12 (3) (1994) 413–422.
- H. Shen, N.R. Gilkes, D.G. Kilburn, R.C. Miller, R.A.J. Warren, Cellobiohydrolase-B, a 2nd exo-cellobiohydrolase from the cellulolytic bacterium *Cellulomonas-fimi*, *Biochem. J.* 311 (1995) 67–74.
- O. Hekmat, Y.W. Kim, S.J. Williams, S.M. He, S.G. Withers, Active-site peptide fingerprinting of glycosidases in complex mixtures by mass spectrometry – discovery of a novel retaining beta-1,4-glycanase in *Cellulomonas fimi*, *J. Biol. Chem.* 280 (42) (2005) 35126–35135.
- J. Gao, W. Wakarchuk, Characterization of five beta-glycosidase hydrolases from *Cellulomonas fimi* ATCC 484, *J. Bacteriol.* 196 (23) (2014) 4103–4110.
- Y. Yang, L.J. Zhang, M.R. Guo, J.Q. Sun, S. Matsukawa, J.L. Xie, D.Z. Wei, Novel alpha-L-arabinofuranosidase from *Cellulomonas fimi* ATCC 484 and its substrate-specificity analysis with the aid of computer, *J. Agric. Food Chem.* 63 (14) (2015) 3725–3733.
- D. Martinez, R.M. Berka, B. Henrissat, M. Saloheimo, M. Arvas, S.E. Baker, J. Chapman, O. Chertkov, P.M. Coutinho, D. Chullen, E.G.J. Danchin, I.V. Grigoriev, P. Harris, M. Jackson, C.P. Kubicek, C.S. Han, I. Ho, L.F. Larrondo, A.L. de Leon, J.K. Magnuson, S. Merino, M. Misra, B. Nelson, N. Putnam, B. Robbertse, A.A. Salamov, M. Schmoll, A. Terry, N. Thayer, A. Westerholm-Parvinen, C.L. Schoch, J. Yao, R. Barbote, M.A. Nelson, C. Detter, D. Bruce, C.R. Kuske, G. Xie, P. Richardson, D.S. Rokhsar, S.M. Lucas, E.M. Rubin, N. Dunn-Coleman, M. Ward, T.S. Brettin, Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*), *Nat. Biotechnol.* 26 (5) (2008) 553–560.
- P.C.B. Raman, G.B. Hurst, M. Rodriguez Jr, C.K. McKeown, P.K. Lankford, N.F. Samatova, J.R. Mielenz, Impact of pretreated Switchgrass and biomass carbohydrates on *Clostridium thermocellum* ATCC 27405 cellulosome composition: a quantitative proteomic analysis, *PLoS One* 4 (2009).
- B. Abt, B. Foster, A. Lapidus, A. Clum, H. Sun, R. Pukall, S. Lucas, T.G. Del Rio, M. Nolan, H. Tice, J.F. Cheng, S. Pitluck, K. Liolios, N. Ivanova, K. Mavromatis, G. Ovchinnikova, A. Pati, L. Goodwin, A. Chen, K. Palaniappan, M. Land, L. Hauser, Y.J. Chang, C.D. Jeffries, M. Rohde, M. Goker, T. Woyke, J. Bristow, J.A. Eisen, V. Markowitz, P. Hugenholtz, N.C. Kyrpides, H.P. Klenk, Complete genome sequence of *Cellulomonas flavigena* type strain (134(T)), *Stand. Genomic Sci.* 3 (1) (2010) 15–25.
- A. Lykidis, K. Mavromatis, N. Ivanova, I. Anderson, M. Land, G. DiBartolo, M. Martinez, A. Lapidus, S. Lucas, A. Copeland, P. Richardson, D.B. Wilson, N. Kyrpides, Genome sequence and analysis of the soil cellulolytic actinomycete *Thermobifida fusca* YX, *J. Bacteriol.* 189 (6) (2007) 2477–2486.
- E. Quevillon, V. Silventoinen, S. Pillai, N. Harte, N. Mulder, R. Apweiler, R. Lopez, InterProScan: protein domains identifier, *Nucleic Acids Res.* 33 (2005) W116–W120.
- C.T. Chung, S.L. Niemela, R.H. Miller, One-step preparation of competent *Escherichia coli* – transformation and storage of bacterial-cells in the same solution, *Proc. Natl. Acad. Sci. U. S. A.* 86 (7) (1989) 2172–2175.
- C. Notredame, D.G. Higgins, J. Heringa, T-coffee: a novel method for fast and accurate multiple sequence alignment, *J. Mol. Biol.* 302 (1) (2000) 205–217.
- M.E.B. Miller, D.A. Antonopoulos, M.T. Rincon, M. Band, A. Bari, T. Akraiko, A. Hernandez, J. Thimmapuram, B. Henrissat, P.M. Coutinho, I. Borovok, S. Jindou, R. Lamed, H.J. Flint, E.A. Bayer, B.A. White, Diversity and strain specificity of plant cell wall degrading enzymes revealed by the draft genome of *Ruminococcus flavefaciens* FD-1, *PLoS One* 4 (8) (2009).
- I. Anderson, B. Abt, A. Lykidis, H.P. Klenk, N. Kyrpides, N. Ivanova, Genomics of aerobic cellulose utilization systems in Actinobacteria, *PLoS One* 7 (6) (2012).
- G.S. Eadie, The inhibition of cholinesterase by physostigmine and prostigmine, *J. Biol. Chem.* 146 (1) (1942) 85–93.
- N.D. Gold, V.J.J. Martin, Global view of the *Clostridium thermocellum* cellulosome revealed by quantitative proteomic analysis, *J. Bacteriol.* 189 (19) (2007)

- 6787–6795.
- [40] G. Xie, D.C. Bruce, J.F. Challacombe, O. Chertkov, J.C. Detter, P. Gilna, C.S. Han, S. Lucas, M. Misra, G.L. Myers, P. Richardson, R. Tapia, N. Thayer, L.S. Thompson, T.S. Brettin, B. Henrissat, D.B. Wilson, M.J. McBride, Genome sequence of the cellulolytic gliding bacterium *Cytophaga hutchinsonii*, *Appl. Environ. Microbiol.* 73 (11) (2007) 3536–3546.
- [41] W.W. Wakarchuk, D. Brochu, S. Foote, A. Robotham, H. Saxena, T. Erak, J. Kelly, Proteomic analysis of the secretome of *Cellulomonas fimi* ATCC 484 and *Cellulomonas flavigena* ATCC 482, *PLoS One* 11 (3) (2016).
- [42] W.R. Gilks, B. Audit, D. de Angelis, S. Tsoka, C.A. Ouzounis, Percolation of annotation errors through hierarchically structured protein sequence databases, *Math. Biosci.* 193 (2) (2005) 223–234.
- [43] J.F. Kane, Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*, *Curr. Opin. Biotechnol.* 6 (5) (1995) 494–500.
- [44] S.S. Lakhundi, K.O. Duedu, N. Cain, R. Nagy, J. Krakowiak, C.E. French, *Citrobacter freundii* as a test platform for recombinant cellulose degradation systems, *Lett. Appl. Microbiol.* 64 (2016) 35–42.
- [45] A. Margolles, C.G. de los Reyes-Gavilan, Purification and functional characterization of a novel alpha-L-arabinofuranosidase from *Bifidobacterium longum* B667, *Appl. Environ. Microbiol.* 69 (9) (2003) 5096–5103.
- [46] V. Lombard, H.G. Ramulu, E. Drula, P.M. Coutinho, B. Henrissat, The carbohydrate-active enzymes database (CAZy) in 2013, *Nucleic Acids Res.* 42 (D1) (2014) D490–D495.
- [47] P. Rapp, F. Wagner, Production and properties of xylan-degrading enzymes from *Cellulomonas-uda*, *Appl. Environ. Microbiol.* 51 (4) (1986) 746–752.
- [48] Y.M. Xue, W.L. Shao, Expression and characterization of a thermostable beta-xylosidase from the hyperthermophile, *Thermotoga maritima*, *Biotechnol. Lett.* 26 (19) (2004) 1511–1515.
- [49] L. Lama, V. Calandrelli, A. Gambacorta, B. Nicolaus, Purification and characterization of thermostable xylanase and beta-xylosidase by the thermophilic bacterium *Bacillus thermantarcticus*, *Res. Microbiol.* 155 (4) (2004) 283–289.
- [50] D. Quintero, Z. Velasco, E. Hurtado-Gomez, J.L. Neira, L.M. Contreras, Isolation and characterization of a thermostable beta-xylosidase in the thermophilic bacterium *Geobacillus pallidus*, *BBA Proteins Proteomics* 1774 (4) (2007) 510–518.
- [51] Q. Wang, T. Xia, Enhancement of the thermostability and hydrolytic activity of GH10 xylanase by module shuffling between *Cellulomonas fimi* Cex and *Thermomonospora alba* XylA, *World J. Microbiol. Biotechnol.* 23 (8) (2007) 1047–1055.
- [52] N.Y. Yu, J.R. Wagner, M.R. Laird, G. Melli, S. Rey, R. Lo, P. Dao, S.C. Sahinalp, M. Ester, L.J. Foster, F.S.L. Brinkman, PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes, *Bioinformatics* 26 (13) (2010) 1608–1615.