Engineering Complex Carbohydrate Catabolism in Industrial Bacteria

Submitted by Holly Cromar to the University of Exeter as a thesis for the degree of Masters by Research in Biological Sciences, September 2017.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

(Signature) ……

Abstract

Second generation biofuels from the fermentation of lignocellulose have the potential to combat increasing problems between transport fuel use and the environment. However, current lignocellulosic degradation methods are complex and expensive requiring large amounts of energy and costly enzymes. This study aimed to improve current degradation methods, increasing product yield, making bioprocessing economically competitive. *Geobacillus thermoglucosidans* DSM2542 was engineered to express CelZ, a thermostable cellulase protein from *Clostridium stercorarium.* Optimal expression of the CelZ gene was confirmed by examining alternative promoters controlling the expression. The stability of pS797 P_{LdhA}::*CelZ* in host organism *Geobacillus thermoglucosidans* DSM2542 resulted in P_{LdhA} showing the greatest activity against cellulose. CelZ successfully degraded cellulose at 33^{o}C, 55^{o}C and 65^{o}C. In addition, CelZ exhibited lignocellulosic degradation against un-pretreated *Zea mays* ssp. *Mays L* and *Sorghum bicolor*.

Key words: *Geobacillus thermoglucosidans*, Cellulase, CelZ, Cellulose, Consolidated bioprocessing, *Zea mays* ssp. *Mays L*, *Sorghum bicolor*

Contents

Table of figures

Table of tables

Table of Abbreviations

bp	Base pairs
CMC	Carboxy-methyl cellulose
GFP	Green Fluorescence Protein
GFP Fluorescence:: OD_{600}	GFP Fluorescence in relation to Optical Density at a wavelength of 600 nm
h	Hour
HPLC	High Performance Liquid Chromatography
l	Liter
LB	Lennox broth
LBA	Lennox broth plus Ampicillin (100 µg ml^{-1})
LA	Lennox Agar plate
LAA	Lennox Agar plate plus Ampicillin (100 µg ml^{-1})
M	Mole
min	Minute
ml	Milliliter
nm	Nanometre
OD	Optical Density
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
Phusion Polymerase	*Taq polymerase fused to a Sso7d domain (NEB)*
Q5 polymerase	*Taq polymerase fused to a Sso7d domain (NEB)*
rpm	Revolutions Per Minute
s	Second
TBE	Tris Bis EDTA gel electrophoresis buffer
TSA	Tryptic Soy Agar
TSAK	Tryptic Soy Agar plus Kanamycin (12.5 µg ml^{-1})
TSB	Tryptic Soy Broth
TSBK	Tryptic Soy Broth plus Kanamycin (12.5 µg ml^{-1})

µl	Microlitre
µm	Micrometre

1. Introduction

1.1 The Energy Industry

Energy is a basic need, required for economic and cultural development (Asif & Muneer, 2007). The combined effects of industrialization, increasing prosperity and population growth mean the amount of energy required globally is increasing at a rapid rate (Adams & et. al, 2016). There are ambitious plans for universal energy access, decreasing the number of people without electricity from 1.2 billion to 800 million by 2030 (Birol & et. al, 2015). Supplying energy is predicted to help eliminate poverty and improve the standard of living for all. The International Energy Outlook 2016 reports total world energy consumption will rise by 48%, from 549 x 10^{15} British thermal units (Btu) in 2012 to 815 x 10^{15} Btu in 2040 (Adams & et. al, 2016). These developments are important because energy supply is a key component for developing countries growth, and energy conservation is thought to have a negative effect on economic growth (Lee, 2005).

Approximately 80% of energy consumed worldwide is provided by fossil fuels (Castro, 2016). According to the latest 'Total Primary Energy Supply' statistics from the International Energy Agency, in 2014, 31.3% came from oil, 21.2% from natural gas and 28.6 % came from coal (Birol & et. al, 2016). Fossil fuels form over millions of years, when biological materials become trapped in the earth. Applied heat and pressure converts the matter into concentrated energy sources (Walker, 2003). This process means fossil fuels are a finite resource, with strong growth in fossil fuel consumption predicted to decline after 2025, mainly due to Chinese coal production peaking (Mohr, et al., 2015).

In addition to supply concerns of fossil fuels, further concerns have been raised about climate change and global energy security of this energy source (Fortman, et al., 2008). Environmental concerns were first acknowledged during the Industrial Revolution. That period was a major turning point in history, bringing new manufacturing processes that transformed industry, requiring the consumption of fossil fuels. The consumption resulted in major incidents caused by air pollution. If current fossil fuel exploitation continues at similar rates as seen in the past, it has

been predicted that global atmospheric temperature will increase by more than 3.6^{o}C by 2100 compared to pre industrial levels (Correa, et al., 2017). This rise in global mean temperature could have a catastrophic effect on ecological communities and drive certain species to extinction.

Today, there is large concern surrounding anthropogenic climate change. The Renewable Energy Directive has established a policy that requires the European Union to fulfil at least 27% of its total energy with renewables by 2030 (European Comission, 2017). In addition, the most recent COP21 pledges give insight to future energy policy's, with the intention of increasing the share of non-fossil fuels from a global average of 19% today, to 25% in 2040 (Birol & et. al, 2015). Only natural gas, the least carbon intensive fossil fuel, is expected to show an increase in consumption percentage (Birol & et. al, 2015).

The energy industry is under pressure to balance economic growth with climate change. Renewable energy is beneficial for the environment but costly which can make it limiting to economic growth. This limit increases the difficulty for renewable energy sources to replace coal and gas from industry or oil from transportation (Birol & et. al, 2015). However, the advantages of renewable energy, in comparison to fossil fuel energy, include renewable energy's infinite supply of sustainable energy that is nonpolluting (Sang-Bing , et al., 2017). The five primary sources of renewable energy are solar, hydro power, wind, geothermal and biomass. Biomass can undergo a bio-chemical conversion to produce biofuels (Kaltschmitta, et al., 2003). Biofuels are reported to be a critical fuel source to combat the global energy issue, as it is thought biofuels can be both eco-friendly and cost-effective (Rodionov, et al., 2017).

1.2 Biofuels

Biofuels are a possible solution for energy independence from traditional fossil fuels. With regards to economic growth, they also have the potential to improve rural economic development (Lin, 2016). To be considered a biofuel, a fuel must be

composed of at least 80% material from living organisms that have been produced within a 10 year period (Uhlenbrook, 2007).

Biofuels can be categorized depending on the feedstock used and method by which the fuel was derived. First generation biofuels are produced from mostly edible crops (Saladini, et al., 2016). The downside of this fuel is that it uses potential food crops, competing for agricultural land. This can cause environmental degradation and has limited sustainability (Correa, et al., 2017).

Second generation biofuels are primarily produced from non-edible, agricultural crops and utilize lignocellulosic residue. Therefore, second generation biofuels do not compete with food sources, are cheap and readily available all year and can quickly be replenished (Ghosh, et al., 2017).However, they have economic limitations as the biomass requires costly and energy intensive pre-treatment.

Third generation biofuels produced from oleaginous microbes provide a technically viable solution, overcoming first and second generation limitations. Microalgae can be used to produce oil and lipids that can be used directly or easily converted into both biofuels and valuable co-products (Brennan & Owende, 2010). Theoretically, algae can produce significantly more oil than first generation biofuels per hectare (Srinophakun, et al., 2017). However, many challenges have impeded algae's development as a biofuel resource. For example, species selection balancing the requirements for biofuel production and a possible energy deficit following production (Brennan & Owende, 2010).

Fourth generation biofuels aim to resolve these issues. The biofuels are created using metabolically engineered microorganisms. The engineering aims to make lignocellulose degradation economically viable, requiring minimal downstream processing. Current retail biofuel molecules, ethanol and biodiesel, are not fully compatible with current, low emission internal combustion engines. As a result, ethanol and biodiesel need blending with petrol or diesel, require downstream processing and have a low energy density (Howard, et al., 2013). Next generation biofuels, aim to be 'drop-in-fuels' that mimic traditional petroleum products currently consisting of a complex mix of hydrocarbons. These biofuels are precise chemical

replacements that do not require a large amount of downstream processing (Peralta-Yahya, et al., 2012). Synthesis of hydrocarbon biofuels primarily fall under two categories. Either the fatty acid pathway-based alkane/alkenes or the isoprenoid biosynthetic pathway based terpenes. Successful production of hydrocarbons has been demonstrated in a number of organisms, however the titers are generally less than 1 g/L, which is too low for commercial viability (Lin, et al., 2015). As a result, significant efforts are still required to optimize the synthetic pathways, increasing the titer.

1.3 Lignocellulosic biomass as a sustainable feedstock

Lignocellulosic biomass has been identified as a suitable feedstock for fourth generation biofuels. Lignocellulosic biomass is composed primarily of three polymers: cellulose, hemicellulose and lignin (Bajpai, 2016). Optimal bioprocessing methods are selected based on the relative abundance of the polymers. Of the three polymers cellulose is the most abundant, naturally occurring polysaccharide source on the planet. Cellulose provides the structural framework in the cell wall of plants and is composed of photosynthetically fixed carbon enabling it to be utilized as a carbon neutral, renewable source of energy (Hinestroza & Netravali, 2014) (Mascal & Nikitin, 2008). Characteristically, cellulose consists of linear polymer chains of glucose molecules, linked via the β-1,4 glycosidic linkage. The chains interact with each other via hydrogen bonds that form a crystalline microfibril (Keegstra, 2010). The glycosidic bond between the molecules can be broken by hydrolysis, splitting into glucose molecules (Fukuoka & Dhepe, 2006). Biochemical conversion of lignocellulosic biomass to ethanol commonly consists of three stages: pretreatment of the biomass, followed by hydrolysis, then fermentation of the glucose molecules to ethanol (Binod & Pandey, 2015).

At present, common methods of conversion of lignocellulosic biomass to ethanol include thermochemical conversion and biochemical conversion. Thermochemical conversion is effective, however has a high energy requirement (Niziolek, et al., 2014). Biochemical conversion is advantageous, because it has the potential to utilize a variety of lignocellulosic biomass, enabling it to replace many liquid fuels

from crude oil. However, biochemical conversion is currently limited by the requirement to produce high concentrations of glucose and xylose, in order to produce high concentrations of bioethanol (Raftery & Karim, 2017).

In addition, research has shown separate pretreatment, hydrolysis and fermentation is frequently a limiting and costly factor. To overcome this limitation, a single process known as consolidated bioprocessing using a single organism or consortium organisms, is being developed (Raftery & Karim, 2017).

Cellulases are a potential solution for consolidated bioprocessing, as the majority have an optimum temperature between 50 to 55^{o}C, allowing for simultaneous cellulose hydrolysis and fermentation (Lin, et al., 2014). This can also reduce the loading of hydrolytic enzymes, reducing costs and improving efficiency.

1.4 Cellulase for lignocellulosic biomass degradation

The research laboratory for whom this research paper was commissioned selected CelZ as its cellulase of choice. CelZ is a cellulase enzyme produced by *Clostridium stercorarium* consisting of an open reading frame of 2958 bp. CelZ combines both *end*oglucanase and *exo*glucanase activities, resulting in hydrolysis of β-1,4-glycosidic linkages along of the polysaccharide cellulose chain (Riedel, et al., 1998). *Endo*nuclease activity occurs at random sites along the polysaccharide chain, producing oligosaccharides of varying length. *Exo*glucanase activity occurs on the reducing and non-reducing ends of the polysaccharide chains producing free glucose or cellobiose (Bhatia, et al., 2017). CelZ was chosen because it has been shown to be able to degrade Avicel and carboxymethyl cellulose (CMC) (Bronnenmeier & Staudenbauer, 1990), CelZ has enhanced cellulolytic activity due to the presence of multiple cellulose binding sites making it a suitable candidate for lignocellulosic biomass degradation (Jauris, et al., 1990). Additionally, CelZ is stable at 60 °C and no loss of activity was observed after 90 h of incubation suggesting it could be suitable for consolidated bioprocessing CMC (Bronnenmeier & Staudenbauer, 1990)

The research laboratory has the *CelZ* gene under the expression of the *lactate dehydrogenase A (ldhA)* promoter. *LdhA* has variable expression patterns and is sensitive to oxygen. Therefore, this research project reviewed alternative promoters. The aim of the alternative promoters was to increase the amount of cellulase protein translated and make cellulase production more predictable. The promoters were selected by Jamie Gilman, a PhD student in the laboratory, based on his phylogenetics study. Jamie confirmed that the promoters resulted in high levels of transgene expression with tight control of gene expression (J. Gillman, PhD thesis 2018).

1.5 Geobacillus thermoglucosidans DSM2542 as host organism

In addition to selecting a suitable cellulase for consolidated bioprocessing, a suitable host organism is required for cellulase production. *Geobacillus thermoglucosidans* DSM2542 (formally referred to as *Geobacillus thermoglucosidasius* (Coorevits, et al., 2012)) was selected by the industrial research lab, following studies by Ross Kent, due to its developed molecular toolkit (Kent, 2016). The complete genome information, including sequence, can be accessed through GenBank (accession number CP012712) (Chen, et al., 2015).

G.thermoglucosidans is a facultative anaerobic, thermophilic, rod-shaped, gram-positive bacteria, with an optimum growth temperature of 55^{o}C (Chen, et al., 2015), (Yu, et al., 2015). *G.thermoglucosidans* has the ability to switch anaerobic respiration if oxygen is absent and is capable of producing ethanol. Whist vegetative, *G.thermoglucosidans* sporulates producing endospores from terminal cells (Nazina & et al., 2001). *G.thermoglucosidans* is a thermophilic organism and has been discovered in a wide range of hostile environments, suggesting *G.thermoglucosidans* is suitable for industrial environments (Brumm, 2015). *G.thermoglucosidans* ability to grow in high temperatures is advantageous because it reduces energy consumption associated with product separation and fermenter cooling and minimizes the chance of contamination, as well as increasing both production rate and organism selection (Taylor, et al., 2009). *G.thermoglucosidans's advantageous traits include its rapid growth rate, amenability to genetic manipulation*

and ability to ferment a wide range of monosaccharides, cellobiose and short-chain oligosaccharides (Bartosiak-Jentys, et al., 2013)

G.thermoglucosidans has been demonstrated to be a suitable platform for biofuel production with the appropriate thermostable enzymes. *G.thermoglucosidans* has been shown to be capable of cellulase expression and hydrolysis of cellulose and capable of fermentation successfully producing istobutanol (Bartosiak-Jentys, et al., 2013) (Lin, et al., 2014). The genome of *G.thermoglucosidasius* suggests that the organism has the capability to translocate proteins across the cytoplasmic membrane using either the secretion (Sec) or twin-arginine translocation (TAT) pathway. However, protein secretion proceeds predominantly via the Sec pathway (Bartosiak-Jentys, et al., 2013). The Sec system translocates proteins across the cytoplasmic membrane unfolded and folds them extracellularly, whilst the TAT pathway transports the folded protein using a metal ion cofactor (Anné, et al., 2016)

1.6 Aim and hypothesis

- ➔ Aim:

  The aim of this research project was to develop the research laboratory's current consolidated bioprocessing. The primary focus was on optimizing cellulose degradation, using the laboratory's preselected host organism *Geobacillus thermoglucosidans* DSM2542 and cellulase, CelZ.

- ➔ Hypothesis: The heterologous expression of CelZ in *Geobacillus thermoglucosidans* DSM2542 confers cellulolytic activity to the organism

This hypothesis will be investigated and addressed through a number of experiments. CelZ is currently under the expression of the *ldhA* promoter. *LdhA* is sensitive to oxygen and has variable expression patterns, therefore it is presumed increased and constant expression of CelZ will improve cellulase degradation. To achieve this, a number of alternative promoters will be examined and compared to *ldhA*, when expressed in *Geobacillus thermoglucosidans* DSM2542. The alternative promoters were selected by Jamie Gillman following his phylogenetics study. The promoters were received upstream of the Green Fluorescence Protein (GFP) gene. Promoter expression in *Geobacillus thermoglucosidans* DSM2542 will be measured by GFP concentration. GFP expression in vented and non-vented flasks will be compared to examine if the promoter is sensitive to oxygen (Figure 1 –phase 2). The current laboratory promoter *ldhA* is sensitive to oxygen.

Next, the promoters that exhibited the greatest GFP expression will be inserted in front of the CelZ gene on the pS797 plasmid (Figure 1 – phase 3). pS797 P_{ldhA}*::CelZ* expression will then be compared to CelZ expression with alternative promoters. To achieve this, the plasmids containing the alternative promoters must first be purified. Then the alternative promoter region exchanged with *ldhA* to synthesize the alternative plasmid for CelZ expression in *Geobacillus thermoglucosidans.*

Once the plasmids containing alternative promoters have been synthesized CelZ activity will be compared using a range of activity assays. *In vitro* assays will be used to perform qualitative and quantitative analysis of cellulase activity when the

transgene is under the expression of different promoters (Figure 1 –phase 4). *In vivo* assays will be used to examine the ability of *Geobacillus thermoglucosidans* DSM2542 to grow on or using the byproducts of cellulose degradation (Figure 1 – phase 5).

Additionally, further analysis will be performed to examine the effect of biocatalyst on biomass degradation (Figure 1 –phase 6). This quantitative analysis will be performed using a*n in vitro* assay, measuring cellulase activity from P_{ldhA}*::CelZ,* under the expression of *Geobacillus thermoglucosidans* DSM2542for the degradation of un-pretreated biomass at a range of temperatures.

Process Map:

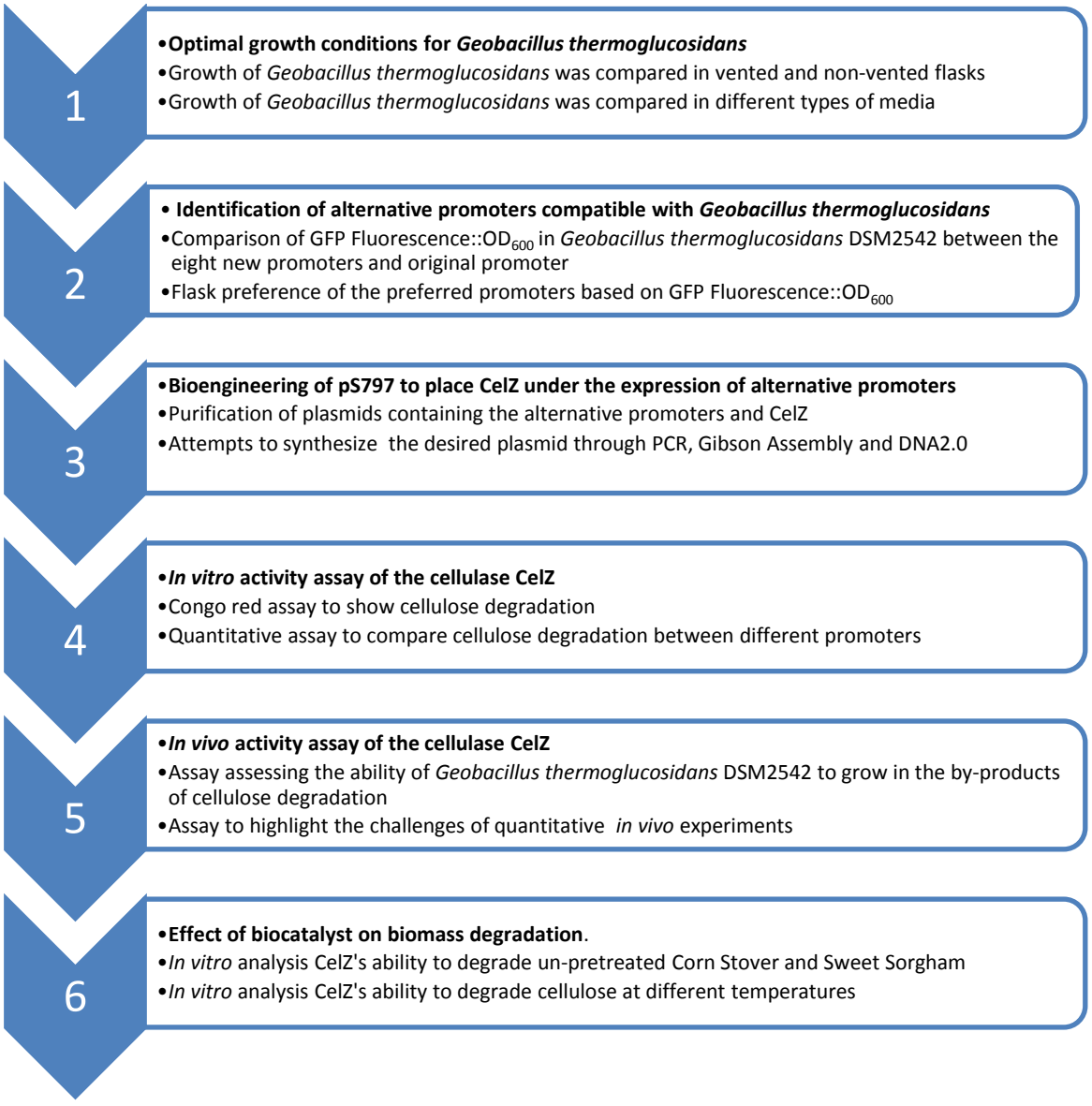


Figure 1 Process flow of experiments conducted in this study

A step by step guide of the experiments conducted in this study, with the objective of improving cellulose degradation. The focus was towards maximizing cellulase expression and understanding optimal conditions for cellulose degradation by the cellulase.

2. Material and Methods

2.1 Identification of alternative promoters compatible with Geobacillus thermoglucosidans *DSM2542*

2.1.1 Optimal growth conditions for *Geobacillus thermoglucosidans* DSM2542

Growth of wild type *Geobacillus thermoglucosidans* DSM2542 was compared in two types of flask and three types of media. The 250 ml flasks were either vented (i.e. lid had a gas permeable gauze) and baffled (i.e. bottom of flask contained baffles to promote mixing) or non-vented (i.e. rubber bung) and plain bottomed (i.e. no baffles). The three types of media were Tryptic Soy Broth (TSB) (15 g l^{-1} tryptone, 5 g l^{-1} peptone, 5 g l^{-1} sodium chloride), Lennox Broth (LB) (10 g l^{-1} tryptone, 5 g l^{-1}yeast extract, 5 g l^{-1} NaCl), and modified Lennox Broth + (LLB) (10 g l^{-1} tryptone, 5 g l^{-1} yeast extract, 10 g l^{-1} NaCl).

A new lyophilised stock of wild type *G.thermoglucosidans* from the American Type Culture was revived. The stock was revived by cracking the surrounding glass vial, and removing the insulation and cotton plug with forceps. TSB (1 ml) was used to suspend the lyophilized bacteria, and the suspension transferred into 5 ml of TSB in a 50 ml falcon tube. The tube was incubated overnight at 55^{o}C in a shaking incubator at 200 rpm. The suspension (150 µl) was spread onto a Tryptic Soy Agar plate (TSA) (15 g l^{-1} tryptone, 5 g l^{-1} peptone, 5 g l^{-1} sodium chloride, 15 g l^{-1}) and incubated at 55^{o}C , for 24- 48 h depending on the extent of growth.

From the plate, a single colony of wild type *G.thermoglucosidans* was used to inoculate 10 ml of TSB in a 50 ml falcon tube. The tube was incubated overnight, at 55^{o}C, with shaking at 200 rpm. The pre-culture was used to inoculate 40 ml of corresponding media in the 250 ml flasks, with a starting optical density (OD_{600}) of approximately 0.05. TSB was used as a negative control.

All OD readings in this research paper were measured using a Tecan plate reader (Infinite 200 Pro). Samples were measured in BD Falcon 96 Flat Bottom Transparent/Black Polyethylene Terephthalate HTS FluoroBlok's. The Tecan plate

reader was set to a temperature of 42^{o}C, with 5 s of initial shaking. 3 x 3 readings were taken per well, at a wavelength of 600 nm. The flasks were incubated at 55^{o}C, at 200 rpm for 6 h. OD readings were taken every hour in triplicate, using the Tecan plate reader. TSB was used as a negative control.

2.1.2 Screening of alternative promoters

Eight new promoter strains were received from the University of Exeter, in *G.thermoglucosidans* on agar plates (

Table 1). Each promoter was in the pS797 plasmid, placed in front of a Green Fluorescent Protein gene (pS797 P_{xxxxx}*::GFP*). The plasmid contained two origins of replication for replication and two selection markers for transformation. In *Escherichia coli,* CoIE ORI was the origin of replication and bla AmpR the selection marker in the presence of ampicillin or carbenicillin (100 µg ml^{-1}). In *Geobacillus sp.*repBST1 was the origin of replication for replication and KanR TK101 was the selection marker for transformation in the presence of kanamycin (12.5 µg ml^{-1}). The plasmid also contained the required genes and recognition sequences for conjugation.

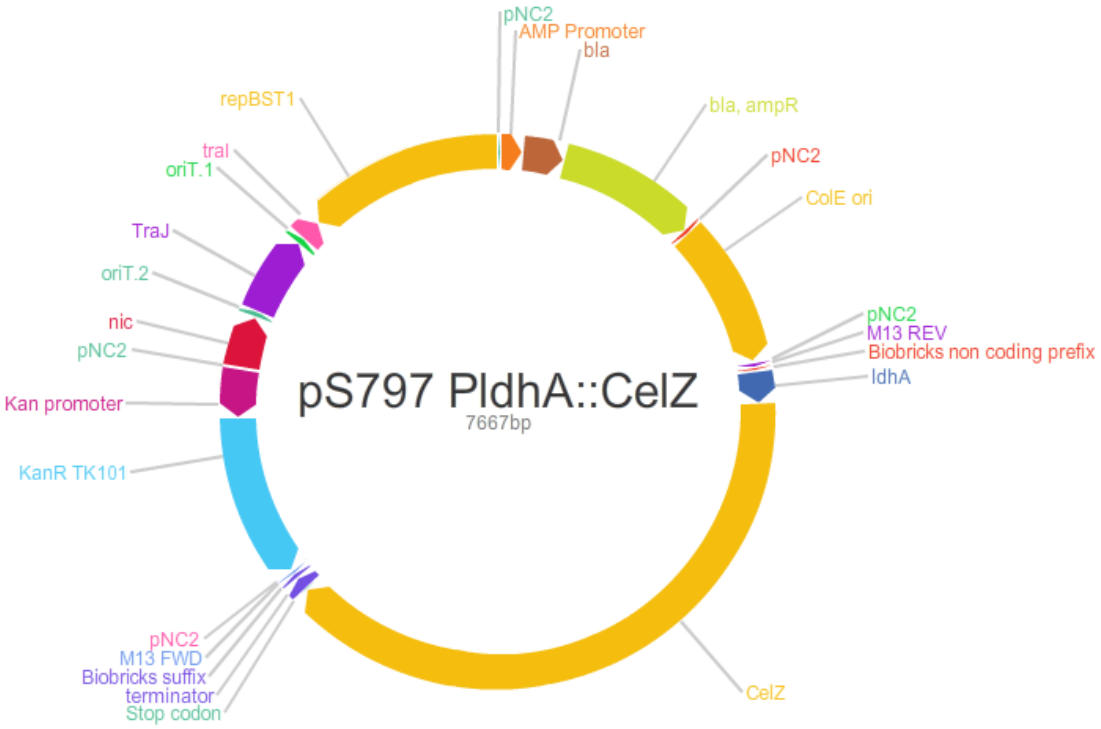


Figure 2 Plasmid map of pS797 P_{ldhA}::*CelZ*

The map highlights key regions of the plasmid, pS797 P_{ldhA}::*CelZ*. The plasmid contained two origins of replication and two selection markers. CoIE ORI was the origin of replication in *E.coli*, and sp.repBST1 was the origin of replication for replication in *Geobacillus*. bla AmpR was the selection marker in *E.coli*, making it resistant to ampicillin or carbenicillin (100 µg ml^{-1}). KanR TK101 was the selection marker for transformation in *Geobacillus* making it resistant to kanamycin (12.5 µg ml^{-1}).

Promoter Identification	Source species	Promoter Function	Reference
GTGNS_03382	*G. thermoglucosidans* DSM2542	Cell division protein ZapA	P_{03382}
GTDN_00966	*G. stearothermophilus* K1041	Hypothetical protein	P_{00966}
GSTEA_02364	*G. stearothermophilus* DSM22	Hydrogen cyanide synthase HcnC precursor	P_{02364}
GTGNS_00505	*G. thermoglucosidans* DSM2542	HTH-type transcriptional repressor GlcR, Tagatose-6-phosphate kinase	P_{00505}
GSTEA_00548	*G. stearothermophilus* DSM22	Enterobactin/ferric enterobactin esterase	P_{00548}
GSTEA_00328	*G. stearothermophilus* DSM22	GTP binding protein TypA/BipA	P_{00328}
GKAU_00878	*G. kaustophilus* DSM7263	Oxygen-independent coproporphyrinogen-III oxidase	P_{00878}
GSTEA_00241	*G. stearothermophilus* DSM22	Hypothetical protein	P_{00241}

Table 1 Promoters received from the University of Exeter in the pS797 plasmid

These promoter were selected by Jamie Gillman due to their high levels of GFP expression and low cell to cell variance of GFP expression, unlike the current promoter *ldhA* (J.Gillman, 2018, PhD thesis). Therefore they were considered more useful promoters for this investigation. Each promoter was received in the pS797 plasmid, situated upstream of the Green Fluorescence Protein gene. To determine the most promising promoters for increasing CelZ expression, they were transformed into *Geobacillus thermoglucosidans* DSM2542. Once transformed GFP Fluorescence::OD_{600} was compared using a Tecan plate reader (Infinite 200 Pro).

A colony from each bacterial strains was streaked onto a TSA plate containing 12.5 µg ml^{-1} of kanamycin (TSAK). Growth of the strain confirmed antibiotic resistance. Wild type *G.thermoglucosidans* was also streaked onto a TSA and TSAK plate as a control.

A single colony from the plate was suspended in 10 ml of TSB containing 12.5 µg ml^{-1} of kanamycin (TSBK) in a 50 ml falcon tube, and grown overnight at 55^{o}C, 200 rpm. Promoter activity was confirmed by identifying GFP expression and Optical Density using a Tecan Plate Reader (Infinite 200 Pro). GFP expression was measured by excitation wavelength (488 nm) and emission wavelength (530 nm). 3 x 3 readings were taken per well with 25 flashes. Wild type, *G.thermoglucosidans* culture was used as a control.

Plasmids were purified using a GeneJET Plasmid Miniprep Kit, from Thermo Fisher. Theculture (10ml) was centrifuged at 12,000 rpm for 10 min at 4^{o}C. The supernatant was disposed of and pellet re-suspended in the re-suspension solution provided in the kit. The provided protocol was followed, with one alteration. Instead of using the elution solution, 50 µl of 37^{o}C ultra-pure water was used to elute the plasmid. The purified plasmid was stored at -20^{o}C.

Once purified, plasmid size was identified by agarose gel electrophoresis (90 min at 140 V). The agarose gel (150 ml Tris Boric Acid EDTA, 1% Agarose) contained 25 µl SYBR Safe DNA Gel Stain, enabling the DNA bands to be viewed under Ultra Violet excitation. Plasmid samples were prepared by combining: ≥16 µl water, 2 µl 10x Fast Digest Green Buffer, 2 µl purified plasmid DNA, 1 µl of each Fast Digest enzyme as required (EcoRI, SacI and PstI); to give a total volume of 20 µl. Samples containing undigested plasmid were used as a digest control. The mixture was briefly centrifuged and incubated at 37^{o}C for 30 min. Following incubation, the tube was re-centrifuged to ensure all condensation returned to the base before being loaded into the gel. The centrifuged sample (20 µl) was loaded into wells on the gel. In addition, 2 µl of DNA weight marker (1 kb DirectLoad from Sigma Aldrich) was loaded into a well for DNA band size calculation under Ultra Violet light.

Plasmid purification and identification from host organism *E.coli* followed the same method.

2.1.3 Transformation of pS797 plasmids into *Geobacillus thermoglucosidans* DSM2542

Once both plasmid and promoter activity had been confirmed, the purified plasmids were transformed into a fresh culture of *G.thermoglucosidans*.

Initially the plasmids underwent heat shock transformation into *E.coli* S17-1. A vial of chemically competent *E.coli* S17-1 frozen stock cells was defrosted on ice, and 2 µl of purified plasmid added (Please refer to 'Preparation of competent E. coli cells' for preparation method). The vial was left on ice for 40 mins and transferred to a heat block, at 42^{o}C, for 2 min. LB (700 µl) was added to the *E.coli* and incubated at 37^{o}C for one hour. The suspension was centrifuged at 8000 rpm for 5 min, and 500 µl carefully removed without disrupting the bacteria pellet. The pellet was a dark beige, with a cloudy precipitate. The pellet was re-suspended in the remaining supernatant, and 200 µl spread onto a Lennox Agar plate (LA) (10 g l $^{-1}$ tryptone, 5 g l $^{-1}$ yeast extract, 5 g l $^{-1}$ NaCl, 12 g l $^{-1}$ agar) containing 100 µg ml^{-1} of ampicillin (LAA). The plate was incubated overnight at 37^{o}C.

Then, *E.coli* was the conjugal donor strain, transferring the plasmid into *G.thermoglucosidans* by conjugal transformation. Approximately 2 ml of biomass from the overnight plate was re-suspended in 600 µl of LB and centrifuged for five min at 5000 rpm. The supernatant was disposed of to 'wash away' any antibiotic residue. The pellet was re-suspended in 600 µl of LB and approximately 5 ml of the recipient, *G.thermoglucosidans*, was also suspended in the solution. Both strains were suspended in the solution and then 300 µl of the solution was dispensed in small droplets onto an LA plate. Once the drops had dried, the plate was transferred into an incubator at 37^{o}C for 6 h. The plates were then transferred into an incubator at 55^{o}C for one hour. Following incubation, 1 ml of LB was dispensed onto the conjugation plate and the biomass was re-suspended. The suspension (200 µl) was spread on a TSAK plate and incubated at 55^{o}C overnight. In addition, 200 µl of the suspension was used to inoculate 5 ml of TSBK and incubated at 55^{o}C, with shaking at 200 rpm overnight. The plate was used to grow subsequent cultures.

To ensure successful transformation at each stage, GFP expression was measured and an enzyme digest performed. An empty pS797 vector was used as a control.

2.1.4 Assessment of promoter expression of GFP gene when grown in vented and non-vented flasks

Multiple individual colonies from each strain were grown (6 colonies per strain). Each colony was used to inoculate 3 ml of TSBK, in a falcon tubes (10 ml). The cultures were incubated overnight, at 55ºC, 200 rpm. The pre-cultures were then used to inoculate 8 ml of TSBK in a 50 ml falcon tube, giving an initial OD_{600} of 0.05. The cultures were incubated for 24 h, at 55ºC, 200 rpm. At 24 h a Tecan Plate Reader (Infinite 200 Pro) was used to measure OD_{600} and Green Fluorescent Protein (GFP) expression of each culture. Wild type *G.thermoglucosidans* was used as a control, and grown in TSB without antibiotic. For each strain, the culture with the highest GFP Fluorescence:: OD_{600} was taken forward for further experiments and streaked onto TSAK / TSA plate as appropriate. The plates were incubated at 55ºC for 24 h.

A single colony from the plates of each promoter was used to inoculate 8 ml of TSBK in a 50 ml falcon tube. The culture was grown overnight, at 55ºC, 200 rpm. The pre-culture was used to inoculate; 40 ml TSBK in 250 ml vented, baffled bottom flask. The flasks were incubated at 55ºC, at 200 rpm for 24 h and Green Fluorescent Protein (GFP) and OD_{600} measurements were taken at 24 h. And to inoculate 25 ml TSBK, in 125 ml non-vented, plain bottom and vented, baffled bottom flasks. Green Fluorescent Protein (GFP) and OD_{600} measurements were taken after incubation at 55ºC, 200 rpm for 24 h.

2.1.5 Storage of the purified plasmids in *Escherichia coli* DH5α

The plasmids were stored in chemically competent *E.coli* DH5α cells. A vial of the *E.coli* DH5α cells was thawed on ice. Purified plasmid (2 µl) was added to the vial and incubated on ice for 40 mins. The vial was placed incubated at 42ºC for 2 mins and returned to ice. LB (700 µl) was added to the vial, and transferred to a shaking incubator, at 37ºC, 220 rpm for 1 h. The vial was centrifuged (5 min, 8000 rpm) and 500 µl of supernatant removed. The pellet was re-suspended in the remaining supernatant and 200 µl was spread onto an LAA plate. Then, the plate was incubated at 37ºC overnight.

The following day a single colony from the plate was grown in 5 ml of LBA overnight producing a fermentation broth, then stored as a glycerol stock (25% Glycerol).

To produce the glycerol stock 1 ml of fermentation broth, was combined with 1 ml of 50% glycerol. The cryogenic vial was inverted 10 times to thoroughly mix the solution and flash frozen in liquid nitrogen. The stocks were then transferred to -80^{o}C for long term storage.

2.1.6 Preparation of competent *Escherichia coli* cells

Approximately 3 µl of *E.coli* DH5α / S17-1, was streaked out onto an LB plate, straight from the -80°C freezer, and incubated overnight at 37^{o}C. Then a colony was removed from the plate and re-streaked onto a second plate and incubated overnight at 37^{o}C. A colony from the second plate was used to inoculate 5 ml LB and incubated overnight at 37^{o}C. The pre-culture (0.4 ml) was used to inoculate 40 ml of LB in a 250 ml non-vented, plain bottom flask. The culture in this flask was grown at 37^{o}C and 200 rpm until A_{650} = 0.4-0.5. The culture was harvest by centrifugation at 8,000 rpm for 8 min at 4^{o}C. The supernatant was removed and re-suspend the pellet in 8 ml of Transformation buffer, TF-1. It was placed on ice for 15 min and centrifuged at 8,000 rpm for 8 min at 4^{o}C again. The pellet was drained and re-suspended in 4 ml of TF-2. The competent cells were either kept on ice and used within a few hours or aliquoted (0.2 ml) and frozen at –70^{o}C.

Transformation buffer 1, TF-1, was composed of potassium chloride 7.4 g.1^{-1}, 1 M potassium acetate, pH 7.5 30 ml, calcium chloride dehydrate 1.5 g.1^{-1} and glycerol 150 g.1^{-1}). All were dissolved in 950 ml of deionized water; adjusted to pH 6.4 with acetic acid, then autoclaved. Once cooled 50 ml (per 950 ml) of filter sterilized 1 M manganese chloride tetrahydrate was added.

Transformation buffer 2, TF-2, was composed of potassium chloride 0.74 g.1^{-1}, calcium chloride dehydrate 11 g.1^{-1} and glycerol 150 g.1^{-1}. All were dissolved in 980 ml of deionized water and autoclaved. Once cooled 20 ml (per 980 ml) of 0.5 M

MOPS buffer (10.47 g in 100 ml), pH 6.8 (adjusted with 5 M KOH) added and the buffer was filter sterilized.

2.2 Exchange of the ldhA *promoter in the pS797* P_{ldhA}::CelZ *plasmid with alternative promoters* P_{00505}, P_{00966} and P_{03382}

Three techniques were used to exchange the lactate dehydrogenase promoter (P_{ldhA}) with the chosen promoters from the University of Exeter (P_{00505}, P_{00966}, P_{03382}).

The objective was to replace the lactate dehydrogenase promoter sequence (151 base pairs) with the following promoter sequences (100 base pairs):

- P_{00505}:5' TTTTTTTGAG TGTTTTTGAT TGAAAATGAT TGATTTTTAA ATCGTTTTCA TTTATGATAA GGGTACAAAG TAATCATTTC CATCCAAGGA GGTGGGATCG 3'
- P_{00966}: 5' TTTTTACATA ATCTTTTTTG GCCATCCTCC CCTCTGACAT CTCTCCGCTT TCCATTTTTC TTTCATTCAT GGTATGATGA ATAATGGAGG TGCCAGATTT 3'
- P_{03382}: 5' CATCGTTTTT TCTTAGGCAT ATCTAGTAGT CGTATCTTTT CATTATATCA GCATTCATGT TATGATAAAC AATAGGATTT TGAAGAATGG GGGAAATCAT 3'

The entire pS797 P_{ldhA}::*CelZ* sequence can be found in the appendix. (For the pS797 P_{ldhA}::*CelZ* plasmid map please refer to Figure 2).

The first, used Phusion High-Fidelity DNA Polymerase kit from New England Biolabs for the Polymerase Chain Reaction (PCR). The second attempt used a Q5 High-Fidelity DNA Polymerase kit from New England Biolabs. Due to industrial patent restrictions in the United States, Q5 High-Fidelity DNA Polymerase amplification needed to be carried out at the University of Exeter, in the U.K. The third attempt, had DNA2.0 synthesize the plasmid.

All three methods required purification of the pS797 P_{ldhA}::*CelZ* and the pS797 P_{00505}::*GFP* plasmids. To purify, cryostocks were streaked onto a TSAK plate if stored in *G.thermoglucosidans* or LAA plate if stored in *E.coli*. The cultures were grown for 24 h at 55^{o}C and 37^{o}C respectively, then used to inoculate 3 ml of appropriate media. The cultures were grown overnight, at 55^{o}C / 37^{o}C, 200 rpm and used to inoculate 25 ml of TSBK or LBA in a 125 ml plain bottom, non-vented flask. The new culture was grown for 24 h, and centrifuged to remove the supernatant. Next, the plasmid was purified using the GeneJET Plasmid Miniprep Kit as previously stated. Once purified, plasmid size was confirmed by agarose gel electrophoresis.

In addition, *in vitro* cellulase activity was confirmed by a Congo Red Assay. The final culture was centrifuged at 12,000 rpm for 10 min at 4^{o}C. The supernatant was retained and pH adjusted to pH 5 using 50% glacial acetic acid, then sterilized by filtration through a Nalgene Rapid-Flow Sterile Disposable Filter Unit. The supernatant was aliquoted into a well in a carboxymethyl cellulose (CMC) plate (25 ml of 0.25% CMC sodium, 0.7% agar). Wells were made by pouring molten CMC agar around sterilized HPLC vials, then removing once set. The plate was incubated at 55^{o}C for 24 h, then stained for 20 min using 0.1% congo red solution, and washed off by soaking in 1 M NaCl for 20 min. The plates were observed on a white light plate, enabling cellulose degradation to easily be identified by orange halo staining on the red stained plate. Empty vector supernatant was tested as a control.

For each of the methods it was important to be able to calculate DNA concentration. Concentration was calculated using a Qubit dsDNA Broad Range Assay Kit. Each assay was carried out at room temperature. The standard protocol provided was followed, using 2 µl of DNA for each assay. Readings were taken in triplicate and the mean concentration was calculated.

2.2.1 Phusion High-Fidelity DNA polymerase

A ‘Phusion High-Fidelity DNA Polymerase’ kit was used from New England Biolabs. Phusion DNA polymerase exhibits polymerase activity and *exo*nuclease activity, increasing reaction speed, yield and fidelity.

Four primers were designed using the New England Biolabs ‘NEBuilder Assembly Tool’, and ordered from Integrated DNA technologies (**Error! Reference source not ound.Error! Reference source not found.**Table 2). The oligonucleotide primers were designed to amplify (Part A) the pS797 P_{ldhA}::*CelZ* plasmid, excluding the lactate dehydrogenase promoter sequence, to form the main backbone of the new plasmid. In addition, primers were designed to amplify the most promising promoter (Part B): P_{00505}. These primers were designed to incorporate ‘tails’ at either end of the promoter region that were complementary to either end of the main backbone sequence.

The primers were briefly centrifuged and resuspended in nuclease-free water.

Primer ref.	Primer Sequence	Anneals	Direction	3'Tm	3'Ta
P1	CTCTAGAAGCGGCCGCGA	CelZ	Rev	68.1°C	68.0°C
P2	ATGCGCAAGTTTTGGAGCTTTG	CelZ	Fwd	65.0°C	68.0°C
P3	cggccgcttctagagTTTTTTTGAGTGTTTTTGATTGAAAATGATTG	00505	Fwd	60.5°C	63.5°C
P4	ctccaaaacttgcgcatCGATCCCACCTCCTTGGA	00505	Rev	64.3°C	63.5°C

Table 2 Primers for the amplification of the pS797 plasmids

The above primers were designed for use in the Polymerase Chain Reaction (PCR). Primer 1 (P1) and primer 2 (P2) were designed to anneal to specific regions on pS797 P_{LdhA}*::CelZ* to amplify the backbone minus the $P_{LdhA.}$ Primer 3 (P3) and primer 4 (P4) were designed to anneal to specific regions on pS797 P_{00505}*::GFP* to amplify the promoter plus complementary regions to the ends of the amplified pS797 P_{LdhA}*::CelZ* fragment. Tm refers to primer melting temperature. Ta refers to primer annealing temperature.

P1 and P2; primers designed to amplify the backbone of pS797 P_{ldhA}*::CelZ*, excluding the lactate dehydrogenase promoter sequence (Part A). P3 & P4; primers designed to amplify the promoter sequence with complementary 'tails' to the backbone of pS797 P_{00505}*::GFP* (Part B) (Figure 3).

The PCR mix (25 µl) was assembled on ice in 50 µl PCR tubes. It was composed of Nuclease water (≤25 µl), 1.25 µl of 10 µM forward primer with a final concentration of 0.5 µM, 1.25 µl of 10 µM reverse primer with a final concentration of 0.5 µM, <250 ng of template DNA and 2X Phusion Master Mix (12.5 µl). Phusion DNA Polymerase was added last to prevent primer degradation caused by 3' -> 5' *exo*nuclease activity. Once assembled, it was mixed gently and placed into a Thermocycler, pre-heated to 98ºC. The cycle started with an initial denaturation (98ºC, 30 s), followed by 32 sub-cycles of the following; denaturation (98ºC for 10 s), annealing (P1 & P2: 69ºC for 30 s, P3 & P4: 72ºC for 30 s) and extension (P1 & P2: 72ºC for 3 min 45 s, P3 & P4: 72ºC for 30 s). Once 32 sub-cycles had occurred the reaction underwent a final extension period (72ºC for 5 - 10 mins). Once complete, the temperature dropped to 4ºC.

The annealing temperature for each primer pair was calculated using the NEB Tm Calculator, taking into account the primer sequence and polymerase kit. Extension time depended on fragment size (15-30 s per kilo base of DNA being synthesized per strand).

Gradient PCR was used to improve product quality of Part A. A variety of annealing temperatures were tested (ºC); 72, 71.5, 70.5, 68.7, 68.7, 66.6, 64.8, 63.6, 63. All other variables remained constant with the above method. Separately, a variety of extension times were tried for the backbone (2 min 45 s and 5 min) and template DNA concentration varied.

Following completion of PCR, the products were purified. A GeneJET PCR Purification Kit from Thermo Fisher Scientific was used and the standard protocol provided followed. Gel electrophoresis was used to separate the DNA products in an agarose gel with a TBE buffer (120 V, 90 min).

If multiple bands were observed the one of desired size was excised and dissolved in binding buffer, provided in a GeneJET Gel Extraction kit from Thermo Fisher Scientific. The desired band from multiple lanes were pooled and weighed, to give a 1:1 volume (µl) of binding buffer to gel weight (mg). The gel mixture was incubated at 55^{o}C until the gel was completely dissolved (approximately 10 min), then vortexed for 3 s. The dissolved solution (<800 µl) was transferred to a GeneJET purification column and centrifuged four times. (1 min, 12,000 rpm). First to remove the binding buffer, then 700 µl of wash buffer was filtered through and centrifuged twice to remove residual wash buffer. Then to elute the DNA with 50 µl of elution buffer from the column membrane.

2.2.2 Gibson Assembly of the synthesized DNA strands

Gibson Assembly was attempted to assemble the DNA fragments from the Polymerase Chain Reaction (Part A and Part B) (Figure 3). A Gibson Assembly Master Mix was used from New England Biolabs. The master mix included three different enzymes; an *exo*nuclease to create a complementary single stranded 3' overhang, a polymerase to fill the gaps of the annealed fragments and DNA ligase to seal nicks in the assembled DNA.

The experiment was performed on ice. Each reaction was composed of deionized water: ≤ 10 µl, DNA strands: 0.02 – 0.5 pmoles, Gibson Assembly Master Mix 10 µl, giving a total volume of 20 µl. The NEBioCalculater used to calculate the optimal moles of DNA, depending on the backbone and insert sizes. The reaction mix was incubated at 50^{o}C for 15 min. Because the promoter fragment (Part B) was less than 200 base pairs long, it was used in 5X excess to the vector backbone (Part A).

The result was observed by gel electrophoresis, using previously described method.

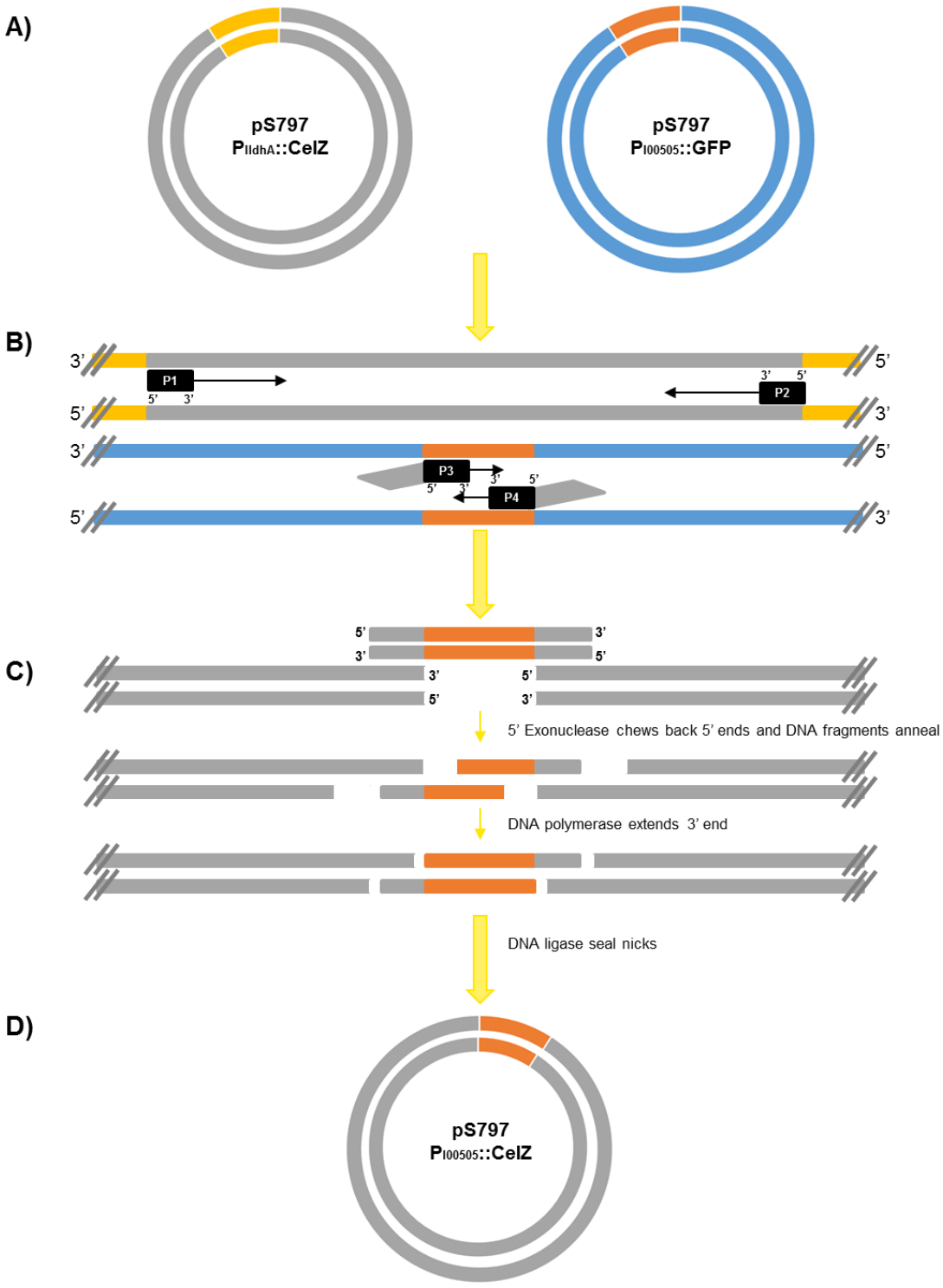


Figure 3 An overview of the method one to synthesize pS797 *P_{00505}::CelZ*

A) *P_{00505}::GFP and P_{ldhA}::CelZ* purified B) PCR: amplification of the desired regions of either plasmid (Part A & B). Primers 3 & 4 add 'tails' to the amplified promoter (Part B), complementary to Part A. C) Gibson assembly: Part A & B anneal D) Desired product pS797 *P_{00505}::CelZ*

2.2.3 Q5 High-Fidelity DNA Polymerase

Due to industrial patent restrictions in the United States, Q5 High-Fidelity DNA Polymerase amplification needed to be carried out at the University of Exeter, U.K.

The purified plasmids (pS797 P_{ldhA}*::CelZ* and pS797 P_{00505}*::GFP* were sent for sequencing at GenScript, then on to the University of Exeter. A member of the research team in the U.K, Paul James, performed the experiment, following the standard protocol provided.

2.2.4 DNA2.0 synthesis

The desired plasmid constructs (P_{00966}*::CelZ*, P_{00505}*::CelZ* and P_{03382}*::CelZ*) were ordered from DNA2.0 (Menlo Park, C.A, USA).

2.3 CelZ activity assay under the expression of P_{00966} *in comparison to the* ldhA *promoter*

2.3.1 Qualitative analysis of pS797 P_{00966}*::CelZ*

pS797 P_{00966}*::CelZ* was received from DNA2.0 as both a stab culture and a lyophilized stock. Bacteria from the stab culture was streaked onto an LAA plate and grown overnight at 37^{o}C. A single colony was taken to inoculate LBA and grown overnight 37^{o}C, 200 rpm.

Additionally, the lyophilized stock was briefly centrifuged and resuspended. pS797 P_{00966}*::CelZ* was transformed into *E.coli* DH5α and *E.coli* S17-1 by heat shock transformation. A vial of the *E.coli* DH5α and *E.coli* S17-1 cells were thawed on ice. The pS797 P_{00966}*::CelZ* solution (2 µl) was added to each vial and incubated on ice for a further 40 mins. The vials were incubated at 42^{o}C for 2 mins and returned to ice. LB (700 µl) was added to each vial, and transferred to a shaking incubator, at 37^{o}C, 220 rpm for 1 h. The vials were centrifuged (5 min, 8000 rpm) and 500 µl of supernatant removed. The pellets were re-suspended in the remaining supernatant and 200 µl was spread onto LAA plates. The plates were incubated at 37^{o}C overnight.

pS797 P_{00966}::*CelZ* was transformed into *G.thermoglucosidans* by conjugal transformation. Approximately 2 ml of biomass from the overnight plate of transformed *E.coli* S17-1 or *E.coli* DH5α was re-suspended in 600 µl of LB and centrifuged for five min at 5000 rpm. The supernatant was disposed of to 'wash away' any antibiotic residue. The pellet was re-suspended in 600 µl of LB and approximately 5 ml of the recipient, *G.thermoglucosidans*, was also suspended in the solution. Both strains were suspended in the solution and then 300 µl of the solution was dispensed in small droplets onto an LA plate. Once the drops had dried, the plates were transferred into an incubator at 37ºC for 6 h. The plates were then transferred into an incubator at 55ºC for 1 h. Following incubation, 1 ml of LB was dispensed onto a conjugation plate and the biomass was re-suspended. 200 µl of the suspension was spread on a TSAK plate and incubated at 55ºC overnight.

A single colony from the overnight plates of each transformation (heat shock and conjugal) were used to inoculate 10 ml of appropriate media (LBA if *E.coli*, TSBK if *G.thermoglucosidans*). The initial culture was incubated at 37ºC / 55ºC, with shaking at 200 rpm overnight. Following overnight growth, the initial cultures were used to inoculate 40 ml media in 250 ml vented, baffled bottom flasks giving a secondary culture.

A sample from both the initial and secondary culture underwent plasmid purification using a GeneJET Plasmid Miniprep Kit, from Thermo Fisher. Plasmid size was observed using fast digest enzymes and gel electrophoresis. In addition, the culture was centrifuged (12,000 rpm for 10 min at 4ºC), supernatant retained and adjusted to pH 5 (with 50% glacial acetic acid) and filter sterilized (Nalgene Rapid-Flow Sterile Disposable Filter Unit).Then the cellulose degrading ability of each supernatant was qualitatively examined by a congo red assay. The supernatant was aliquoted into a well in a carboxymethyl cellulose (CMC) plate (25 ml of 0.25% CMC sodium, 0.7% agar) and incubated at 55ºC for 24 h. Then the plate was stained for 20 min using 0.1% congo red solution, and washed off by soaking in 1 M NaCl for 20 min.

Empty vector supernatant and pS797 P_{ldhA}::*CelZ* were used as controls.

2.3.2 Quantitative analysis comparing promoters *00966* and *ldhA*

An *in vitro* assay was performed to quantitatively examine pS797 P_{00966}*::CelZ* activity in comparison to pS797 P_{ldhA}*::CelZ*. The plasmids were freshly transformed. pS797 P_{00966}*::CelZ* was transformed into *E.coli* DH5α using the heat shock transformation protocol described above. Following overnight incubation of the newly transformed strain on an antibiotic selective plate, individual colonies directly inoculated 40 ml of LBA in 250 ml vented, baffled flasks. The culture was incubated at 200 rpm, 37^{o}C for 22 h. pS797 P_{ldhA}*::CelZ* was transformed into *E.coli* S17-1 by heat shock transformation, then transformed into *G.thermoglucosidans* by conjugal transformation. Following overnight incubation of the newly transformed strain on an antibiotic selective plate, colonies were taken to directly inoculate 40 ml of TSBK in 250 ml vented, baffled flasks.

Following growth, the cultures were centrifuged for 15 min at 10,000 rpm and the supernatant retained, whilst the pellet was discarded. The pH, of approximately 7.3, was adjusted to pH 5 using 50% glacial acetic acid. Then the supernatant was sterilized by filtration through Nalgene Rapid-Flow Sterile Disposable Filter Units.

G.thermoglucosidans and *E.coli* DH5α containing the pS797 empty vector were used as negative controls.

When necessary the supernatant was concentrated tenfold using Amicon Ultra Centrifugal Filters. Supernatant (12 ml) was inserted into each Amicon tube. Then the centrifugal filter was aligned perpendicular to the center of the centrifuge and spun at 5000 g for 30 min. The concentrated supernatant was pooled according to plasmid and host organism.

To quantitatively examine activity of each supernatant, 1 ml of supernatant was added to a 2 ml microcentrifuge, screwcap tube, along with 0.02 g of Whatman 1 filter paper. Five repeats were done for each supernatant, both concentrated and non-concentrated. The tubes were placed on their side in the incubator, at 55^{o}C shaking at 200 rpm. Samples were taken at four time points; Initial samples at 0 hour, 5 days, 11 days and 17 days.

High Performance Liquid Chromatography (HPLC) was used to measure glucose and cellobiose production. Samples (20 µl) were injected into an Agilent Hi-Plex Na column (column size: 7.7x300 mm, guard: Hi-Plex Na 7.7x50 mm, pore size of 10 µm) of the HPLC (Dionex 3000 HPLC with a Shodex RI detector). Samples ran at 80^{o}C, 0.6 ml/min, mobile phase water/0.015 M NaOH

2.3.3 *In vivo* analysis*:* ability of *Geobacillus thermoglucosidans* DSM2542 to grow in the byproducts of cellulose degradation

Supernatant from *G.thermoglucosidans* containing pS797 P_{ldhA}*::CelZ* was adjusted to pH 5, purified and concentrated 10 fold, following the method previously stated. The concentrated supernatant (1ml) was added to 0.02 g of Whatman 1 filter paper, in a 2 ml microcentrifuge screwcap tube. Initial HPLC samples were taken at 0 h, and then the tubes were placed in a 55^{o}C incubator at 200 rpm. The CelZ was left to degrade the cellulose for 6 days. The second set of HPLC samples were taken at 6 days, before cultures were added to test tubes. Then, either *E.coli* S17-1 pS797 P_{00966}*::GFP*, *G.thermoglucosidans* DSM2542 pS797 P_{00966}*::GFP*, TSB or water (200 µl) were added to tubes. They were returned to the incubator overnight, with *E.coli* going into a 37^{o}C incubator at 200 rpm, and the remaining tubes returning to the 55^{o}C incubator, 200 rpm. After 24 h, a sample from each of the tubes was observed using a Tecan plate reader to measure Optical Density and GFP fluorescence. In addition, samples were prepared for HPLC analysis to measure glucose and cellobiose concentration. Samples (20 µl) were injected an Agilent Hi-Flex Na column (pore size of 10 µm) of the HPLC (Dionex 3000 HPLC with a Shodex RI detector).

2.4 Ability of pS797 P_{ldhA}::CelZ expressed in Geobacillus thermoglucosidans to degrade plant biomass and the effect of temperature and enzyme concentration

2.4.1 Degradation of plant biomass by *Geobacillus thermoglucosidans* expressing pS797 P_{ldhA}::CelZ

Zea mays ssp. *Mays L* (Corn stover) and *Sorghum bicolor* (Sweet Sorghum) stem sections were flash frozen in liquid nitrogen, and ground using a pestle and mortar. Then the biomass was autoclaved for sterilization and 0.1 g inserted into sterile Eppendorf tubes.

The *G.thermoglucosidans* supernatant was prepared using the standard protocol describe above. 10 fold concentrated supernatant (1 ml) was added to the each Eppendorf tube. As a control a parallel experiment was set up, with *G.thermoglucosidans* contained an empty pS797 vector. The tubes were double wrapped in parafilm and placed horizontally into an incubator at 37^{o}C, 200 rpm. Samples were measured at 0 h and after 4 days of incubation. The contents of the Eppendorf tube was transferred into a sterile, 2 ml centrifugal filter unit and centrifuged for 2 min at 12,000 rpm. The filtered solution underwent HPLC analysis to measure glucose and cellobiose concentration. Samples (20 µl) were injected an Agilent Hi-Flex Na column (pore size of 10 µm) of the HPLC (Dionex 3000 HPLC with a Shodex RI detector.

2.4.2 Effect of temperature and concentration on cellulose degradation

Supernatant was produced from *G.thermoglucosidans* expressing pS797 P_{ldhA}::*CelZ*. pS797 P_{ldhA}::*CelZ* was transformed into a fresh culture of *G.thermoglucosidans*, using heat shock transformation into *E.coli S17-1* followed by conjugal donor transformation. A colony from the selective TSAK plate was used to inoculate 5 ml of TSBK and incubated overnight at 55^{o}C, 200 rpm. This pre-culture was used to inoculate 40 ml of TSBK in a 250 ml flasks, with a starting optical density (OD_{600}) of approximately 0.05 and incubated for 22 h at 55^{o}C, 200 rpm. Following growth, the culture was centrifuged for 15 min at 10,000 rpm and the supernatant

retained, whilst the pellet was discarded. The pH was adjusted to pH 5 using 50% glacial acetic acid and supernatant sterilized by filtration through Nalgene Rapid-Flow Sterile Disposable Filter Units.

When necessary the supernatant was concentrated tenfold using Amicon Ultra Centrifugal Filters. Supernatant (12 ml) was inserted into each Amicon tube. Then the centrifugal filter was aligned perpendicular to the center of the centrifuge and spun at 5000 g for 30 min. The concentrated supernatant was pooled according to plasmid and host organism.

Cellulase activity of the supernatant was examined at three different temperatures: 33^{o}C, 55^{o}C and 65^{o}C. In addition, it was examined both concentrated and non-concentrated at each temperature.

Supernatant (1 ml) was inserted into 2 ml Eppendorf tubes, containing 0.02 g of sterilized, Whatman 1 filter paper, from Sigma Aldrich. Samples were taken at 3 or 4 time points: 0 hour, 4 days, 7 days, and 21 days (33^{o}C only). The aqueous phase of each sample was filter sterilized before being sent for HPLC analysis. HPLC analysis measured for glucose and cellobiose production, using an Agilent Hi-Flex Na column.

As a control a parallel experiment was set up with *G.thermoglucosidans* containing an empty pS797 vector.

3. Results

3.1 Identification of alternative promoters compatible with Geobacillus thermoglucosidans *DSM2542*

3.1.1 Optimal growth conditions for *Geobacillus thermoglucosidans* DSM2542

The first aim of this research project was to understand the optimal growth conditions for *Geobacillus thermoglucosidans* DSM2542*,* the research laboratory's chosen industrial host. A growth curve of the Wild Type *G.thermoglucosidans* was plotted for three types of media (TSB, LB and Modified LB), in two flask types (vented, baffled and non-vented, plain bottom).

Growth was observed in all types of media. Over a six hour time period, *G.thermoglucosidans* reached its highest optical density growing in TSB, in a vented flask (

Figure 4), however the difference between the media's was not significant.

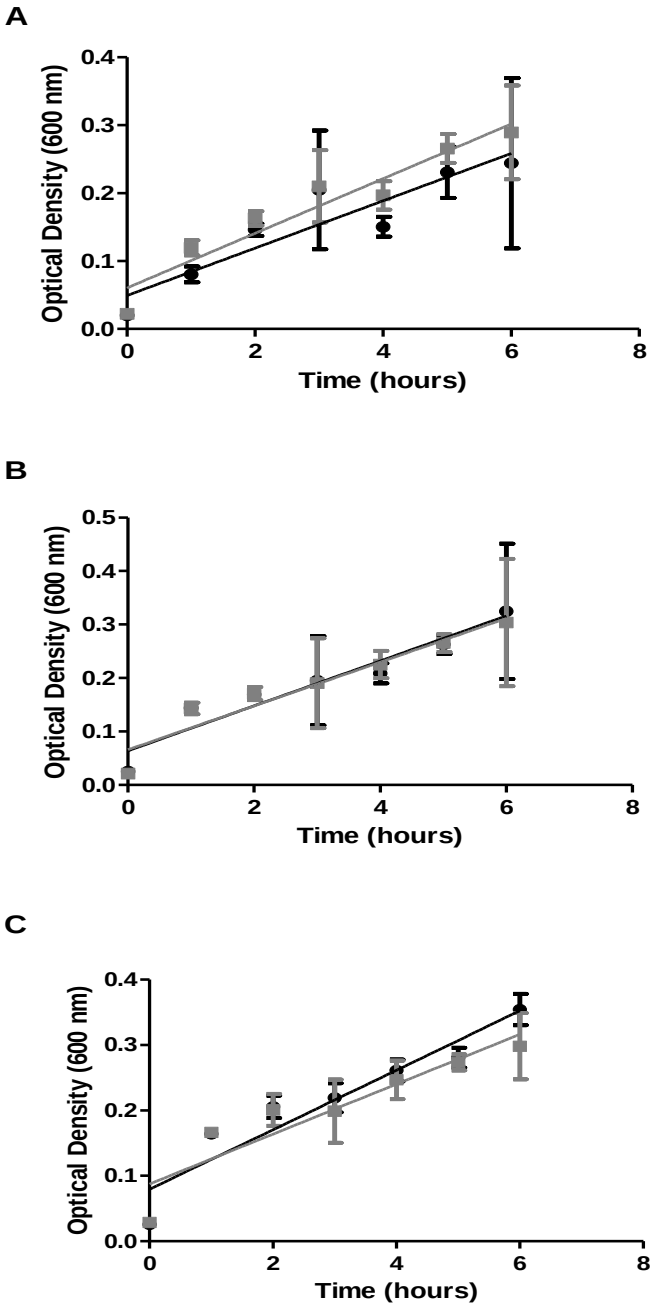


Figure 4 Preferred growth conditions of *Geobacillus thermoglucosidans*

This figure shows the optical density of *Geobacillus thermoglucosidans* DSM2542 grown over a 6 h period. Each graph shows growth in a different media; A) LB media, B) LB modified media, C) TSB media. The two growth curves per graph indicate the 250 ml flask *Geobacillus thermoglucosidans* DSM2542 was grown in. Grey curve and symbols represent the non-vented flask, black curve and symbols represent the vented flask. The square symbol and circular symbol represent the mean optical density at a particular time point. This experiment was carried out in triplicate and the bars represent the standard error of the mean. The curve represents a linear line of best fit, generated by GraphPad Prism.

3.1.2 Promoter identification and activity assessment

Once preferable growth conditions had been established, eight *G.thermoglucosidans* strains sent from the University of Exeter, were successfully revived. Each strain contained a variation of the pS797 plasmid, varying the promoter situated in front of a GFP gene.

All 8 plasmids were purified and restriction enzymes used to confirm plasmid size. Initially this was attempted from a *G.thermoglucosidans* host, however no bands were observed, despite GFP activity. The purified plasmids were then transformed into *Escherichia coli* S17-1 and purified again. Bands could clearly be observed in the gel (Figure 5). Each plasmid was digested with FastDigest enzymes SacI and PstI, giving a single and double digest respectively. Each band appeared in the expected location on the agarose gel.

The pS797 $P_{*\text{alternative promoter}*}$::GFP plasmids were each 5350 base pairs in length. The SacI Fast Digest enzyme cut at base pair 423, giving a straight chain fragment of 5350 base pairs. The EcoRI Fast Digest enzyme cut at 1706/1710 bases, giving a straight chain fragment of 5350 base pairs. The PstI Fast Digest enzyme cut in two places at 2430/ 2426 and 4926/ 4922 bases giving two straight chain fragments of approximately 2496 and 3171 base pairs. Therefore each of the bands is in the expected position (Figure 5).

All of the plasmids were successfully transformed into *E.coli* S17-1 and *G.thermoglucosidans* by conjugal transformation. This was confirmed antibiotic selection, gel electrophoresis and by measuring the optical density and green fluorescence protein expression. GFP and OD readings were taken after 24 h of growth.

The plasmids were successfully stored in a purified state and in 50% glycerol stocks.

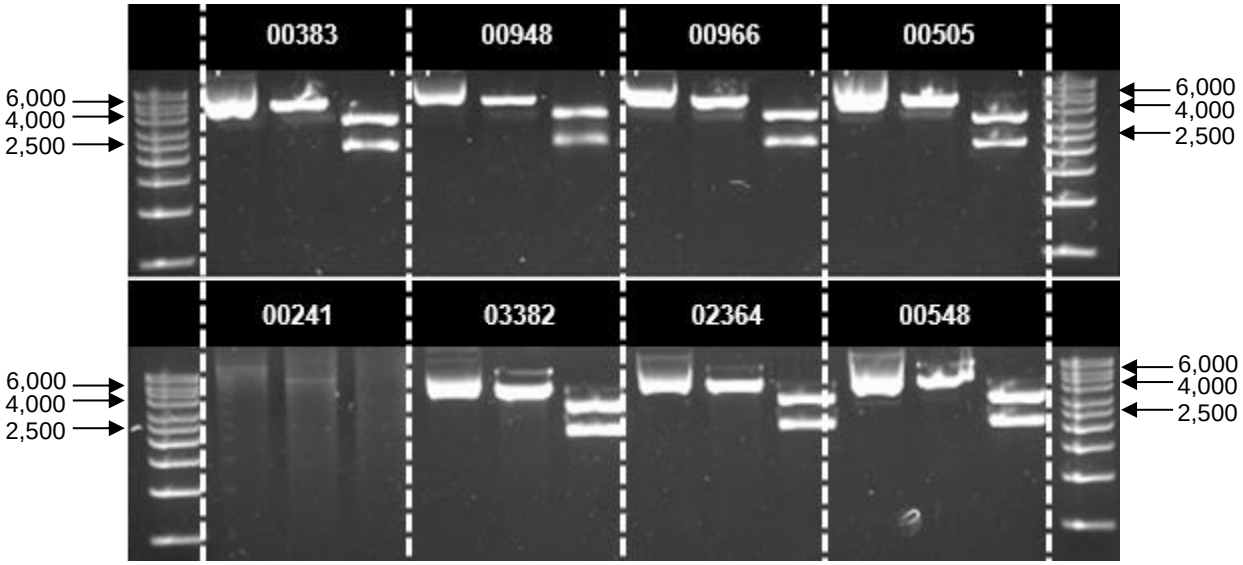


Figure 5 Identification of the eight new promoters found in the pS797 plasmid

Gel electrophoresis of purified DNA from *Escherichia coli* S17-1. The gel has been sectioned according to the variable promoter the host organism contained. It examined eight different promoters selected by the University of Exeter specifically for this research project. The first lane of each section contains the undigested plasmid, the second lane shows the result of a digest with SacI and the third lane shows the result of a digest with PstI. The only variable region between the plasmids was the promoter. All promoters were 100 bp, therefore identical digests can be observed. All SacI digests give a band of the correct size of 5350 bp. All PstI digests give two bands of approximately the correct size of 2496 and 3171 bp. Promoter 00241 is unclear in this figure but was successfully observed in a subsequent gel. The DNA marker was a Direct Load 1Kb DNA ladder from Sigma. Gel was observed under UV light.

To observe if there was expression variance between different cultures of the same strain, six cultures were grown for each of the eight promoter strains. GFP Fluorescence::OD_{600} was measured using a Tecan plate reader after 24 h of growth at 55^{o}C, 200 rpm. Large variance was seen for every promoter, with some cultures showing half the GFP Fluorescence::OD_{600} of others (Figure 6 and Appendix).

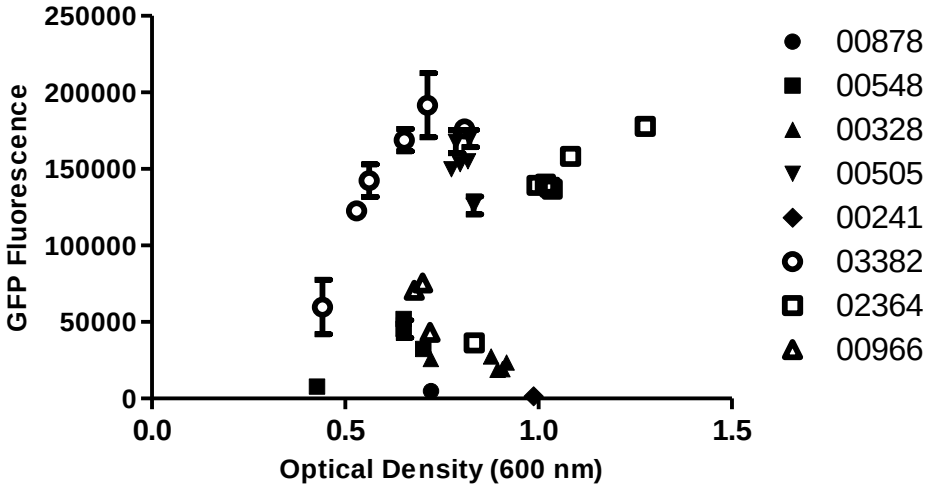


Figure 6 Variability of GFP fluorescence between colonies in *Geobacillus thermoglucosidans* DSM2542.

The graph shows the variability in GFP Fluorescence::OD_{600} resulting from each of the new promoters. No correlation was observed between OD and GFP expression. Each point on the graph represents the mean GFP Fluorescence::OD_{600} from three replicates of three independent transformant cultures per colony Six individual colonies were selected per promoter, however not every colony grew successfully. Colonies were taken from a single transformation plate of *Geobacillus thermoglucosidans* DSM2542. Cultures were grown in TSBK. The error bars represent the standard error of the mean. The instrument's setting were constant.

For each promoter strain, the colony with the greatest GFP Fluorescence::OD_{600} was taken forward for subsequent experiments. Including comparing mean GFP Fluorescence::OD_{600} of each strain following 24 h of growth in a vented flask, in TSBK. The results show the strains with the greatest GFP Fluorescence::OD_{600} were: 1st -P_{00505}, 2nd - P_{03382} and 3rd – P_{00966}. This experiment was repeated 3 times, with 3 samples per experiment (Figure 7). The most successful plasmids were chosen as potential replacements for the P_{ldhA} promoter.

GFP Fluorescence::OD_{600} was compared in a non-vented flask and a vented flask for $P_{ldhA,}$ $P_{00505,}$ P_{03382} and P_{00966}. Expression levels were higher when cells were grown in vented flasks, irrespective of the type of plasmid used. P_{ldhA} was the only exception, with expression higher in a non-vented flask. (Figure 8).

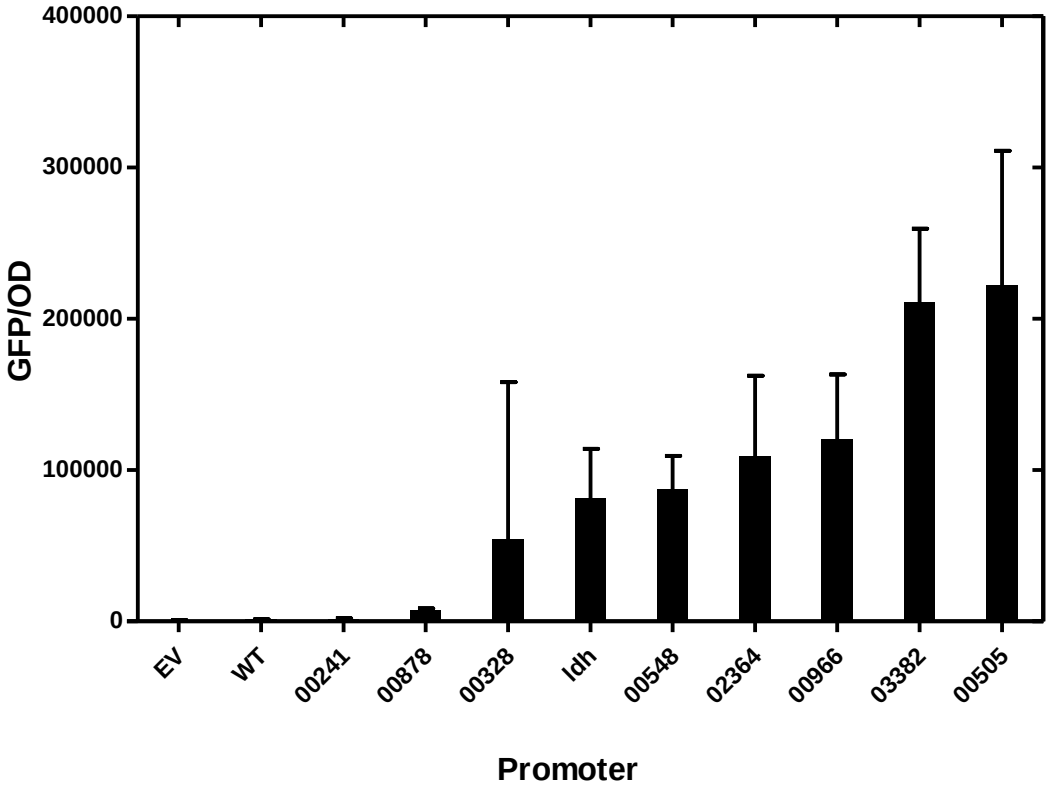


Figure 7 Resulting GFP fluorescence from each promoter

The graph shows the mean GFP Fluorescence::OD_{600} of each of the new promoters, when expressed in *Geobacillus thermoglucosidans* DSM2542. The black bars represent mean GFP Fluorescence::OD_{600} from three replicates of four independent cultures, after 24 hours growth. Cultures were grown in TSBK, in a vented flask. This figure shows promoters ranked in order of increasing GFP Fluorescence::OD_{600}, with P_{00505} showing the greatest levels of fluorescence. P_{ldhA} was measured as a comparison and pS797 EV and WT used as controls. The error bars represent the standard error of the mean. The instrument's setting were constant.

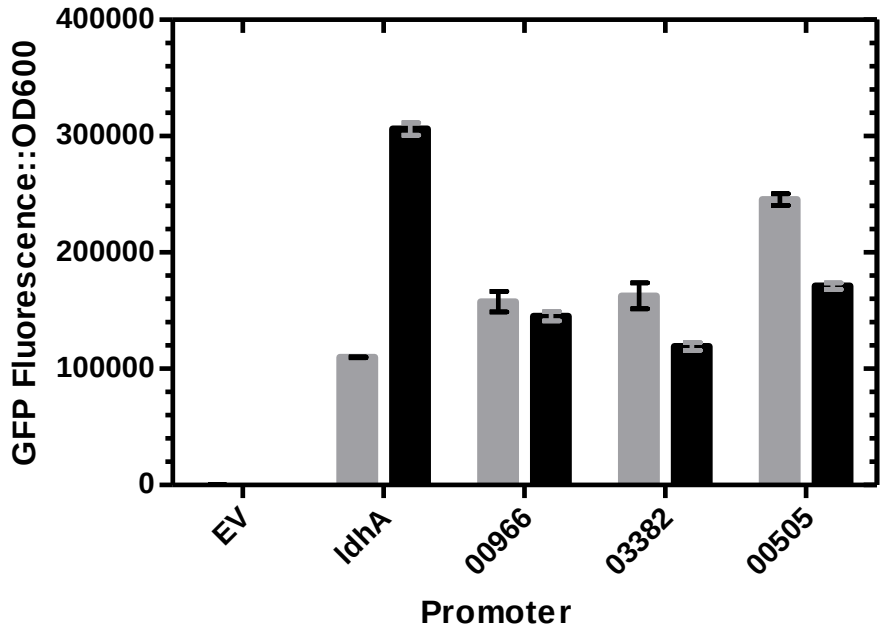


Figure 8 GFP expression in *Geobacillus thermoglucosidans* DSM2542 grown in a vented compared to a non-vented flask

This figure shows the GFP Fluorescence::OD_{600} of each of the chosen promoters (*ldhA*, 00966, 03382 and 00505) expressed in *Geobacillus thermoglucosidans* DSM2542 grown over a 24 hour period. The cultures were grown in TSBK media in both a vented and non-vented flask. The grey bar indicates mean GFP Fluorescence::OD_{600} in a vented flask and the black bar indicates mean GFP Fluorescence::OD_{600} in a non-vented flask, from nine replicates. The error bars show the standard error of the mean. The three new promoters had greater GFP Fluorescence::OD_{600} in the vented flask, compared to the original promoter, *ldhA*, which had greater expression in the non-vented flask. An empty vector was used as a control.

3.2 Modification of pS797 P_{ldhA}::CelZ

Greatest GFP expression in relation to OD was observed in the samples containing plasmid P_{00505}.Therefore, P_{00505} was first selected to be inserted in front of the cellulose gene.

3.2.1 Plasmid purification

Both plasmids needed purifying for PCR. Successful purification and activity of pS797 P_{00505}::*GFP* can be seen in (Figure 5).

However, purification of pS797 P_{ldhA}::*CelZ* took multiple attempts. The research laboratory's glycerol stocks were successfully revived in appropriate media, with antibiotic selectivity. Then the plasmids were purified and gel electrophoresis used to determine plasmid size, with fast digest enzymes used to digest at specific locations. pS797 P_{ldhA}::*CelZ* was 7667 base pairs in length. Therefore, Fast Digest enzyme SacI should have cut at base pair 423, giving a straight chain fragment of 7667 base pairs. Fast Digest enzyme EcoRI should have cut at 1706/1710 bases, giving a straight chain fragment of 7667 base pairs. Fast Digest enzyme PstI should have cut in two places at 2430/ 2426 and 4926/ 4922 bases giving two straight chain fragments of approximately 2496 and 3171 base pairs.

All pS797 P_{ldhA}::*CelZ* bands from *E.coli* sp. appeared in the incorrect position on the agarose gel (

Figure 9). For *E.coli* TOP10 the undigested plasmid appeared below the 4000 base pair marker. The single digest bands were approximately level with the 5000 base pair marker. In addition, PstI should have resulted in a double digest, but only one band appeared level with that of the single digest.

For *E.coli* S17-1 the single digest also appeared level with the 5000 base pair marker. A second attempt at a double digest combined SacI and EcoRI as they both perform single digests. The result produced two bands. This problem was observed in another report (R. Kent, Screening and Engineering of *G.thermoglucosidans spp*. For Consolidated Bioprocessing of Lignocellulosic Biomass, p.68), suggesting the PstI site may have been removed or altered.

As with the previous experiment, when *G.thermoglucosidans* was used to amplify the selected plasmids, the culture grew successfully suggesting antibiotic resistance. However, plasmid bands were not observed on the agarose gel electrophoresis.

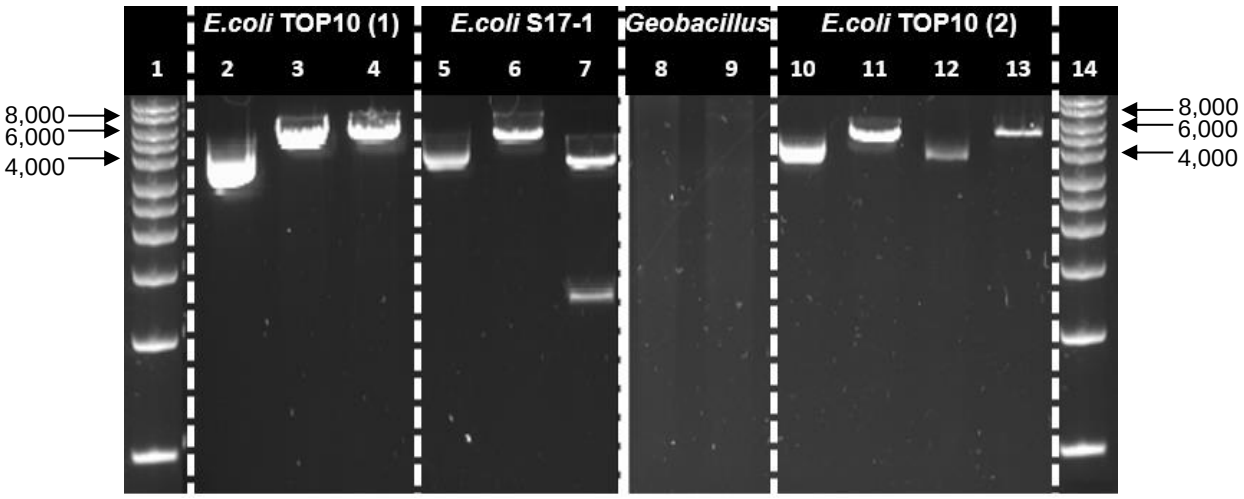


Figure 9 Gel electrophoresis results for Fast Digest of pS797 P_{ldhA}::*CelZ*

Gel electrophoresis of purified DNA from *Escherichia coli* S17-1, *Escherichia coli* TOP10 and Geobacillus thermoglucosidans DSM2542. The gel has been sectioned according to the glycerol stock used for inoculation and plasmid purification. It examined glycerol stocks containing the pS797 P_{ldhA}::*CelZ* plasmid. The plasmid from each glycerol stock should have been identical. *pS797* P_{ldhA}::*CelZ* is 7667 bp in length. Therefore a digest with SacI should produce a single band of 7667 bp. A digest with PstI should produce two bands of approximately 2500 and 3200 bp in length. Lane 2, 5, 8, 10 and 12 contain the undigested plasmid. Lane 2, 9 and 11 contain the plasmid digested with SacI. Lane 4 and 13 contain the plasmid digested with PstI. Lane 7 contains the plasmid digested with EcoRI & PstI. All bands with a SacI digest appear too low in comparison to the DNA marker. None of the PstI digests produced two bands as expected. The DNA marker was a Direct Load 1Kb DNA ladder from Sigma. Gel was observed under UV light.

The CelZ plasmid band appeared smaller than expected on the agarose gel so a congo red assay was performed to check for cellulase activity. Activity would have been identified by an orange halo indicating hydrolysis of CMC in 0.7% agar caused by the cellulase proteins (Carder, 1986). Due to the reduced binding capability of the dye to the beta-1, 4 glyosidic bonds in polysaccharide polymers of the CMC (Teather, 1982). A large variety of glycerol stocks were tested in hope of finding a successful candidate, however no activity was observed, despite the successful growth under antibiotic selectivity. CelA supernatant was used as a positive control. CelA is another of the laboratory's cellulase enzymes that has shown to be active against CMC. *E. coli BL21* containing CelA was grown under the same conditions as *E. coli* samples containing CelZ. CelA activity was observed (Figure 10).

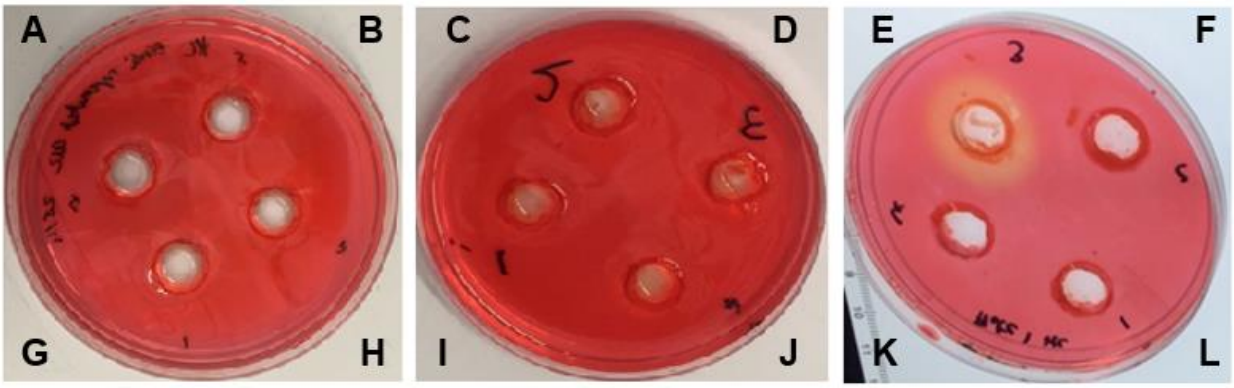


Figure 10 Cellulase activity of the laboratory's glycerol stocks

The images represent cellulase activity of the supernatant from the laboratory's glycerol stocks. Glycerol stocks from a number of species of bacteria, containing pS797 P_{ldhA}*::CelZ* were examined: *Escherichia coli* TOP10 (Wells A, C), *Escherichia coli* S17-1 (Wells B, D), *Geobacillus thermoglucosidans* DSM2542 (Wells G, I) and *Escherichia coli* BL21 (Well F). The supernatant was aliquoted into the well on a CMC plate and incubated at 55^{o}C for 24 h, then stained with congo red. An orange halo surrounding the well indicated cellulase activity, and cannot be seen for any of the glycerol stocks. Supernatant from *E. coli* BL21 containing pS797 P_{ldhA}*::CelA* was used as a positive control (Well E) and the remaining wells contained supernatant from wild type, host bacteria as a negative control.

As a control, glycerol stocks containing pS797 P_{ldhA}*::CelA* were grown at the same time as pS797 P_{ldhA}*::CelZ.* pS797 P_{ldhA}*::CelA* expressed in *G.thermoglucosidans* exhibited cellulase activity in a congo red assay (Figure 10). In addition, following electrophoresis the bands for pS797 P_{ldhA}*::CelA* appeared in the correct position on the gel (Figure 11). Fast digest enzymes were used to digest at specific locations. pS797 P_{ldhA}*::CelZ* should have shown the same digests as described previously, but the bands once again appeared in the wrong location and it did not undergo a double digest. pS797 P_{ldhA}*::CelA* was 9986 base pairs in length. Therefore, Fast Digest enzyme SacI should have cut at base pair 4609/4605, giving a straight chain fragment of 9986 base pairs. Fast Digest enzyme EcoRI should have cut at 1706/1710 bases, giving a straight chain fragment of 9986 base pairs. Fast Digest enzyme PstI should have cut in four places at 2510/2506, 4316/4312, 6347/6343, 7245/7241 giving four straight chain fragments of approximately 1806, 2031, 898 and 5251 base pairs.

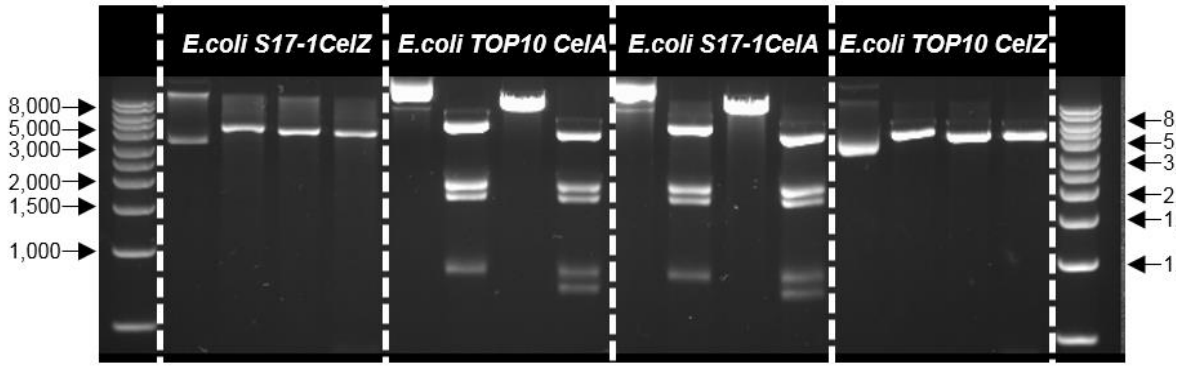


Figure 11 Gel electrophoresis results for Fast Digest of pS797 P_{ldhA}::CelZ and pS797 P_{ldhA}::CelA
Gel electrophoresis of purified DNA from *Escherichia coli* S17-1, *Escherichia coli* TOP10 and *Geobacillus thermoglucosidans* DSM2542. The gel has been sectioned according to the glycerol stock used for inoculation and plasmid purification. It examined glycerol stocks containing the pS797 *P_{ldhA}::CelZ* plasmid or pS797 *P_{ldhA}::CelA* plasmid. The first lane of each section shows the undigested plasmid, the second lane contains the plasmid digested with PstI, the third lane contains the plasmids digested with EcoRI and the forth lane contains the plasmid digested with EcoRI and PstI. None of the bands for pS797 *P_{ldhA}::CelZ* appear in the correct location and are not correctly digested by PstI. The bands for pS797 *P_{ldhA}::CelA* are in the correct location and each digest is as expected. The DNA marker was a Direct Load 1Kb DNA ladder from Sigma. Gel was observed under UV light.

Due to the lack of cellulose activity, an alternative *G.thermoglucosidans* culture expressing CelZ was obtained from Kun Zhang. The plasmid was purified from the sample and freshly transformed into *E.Coli* S17-1 and *G.thermoglucosidans*. The freshly transformed strain of *G.thermoglucosidans* containing pS797 P_{ldhA}*::CelZ* exhibited cellulase activity in a congo red assay (Figure 12).

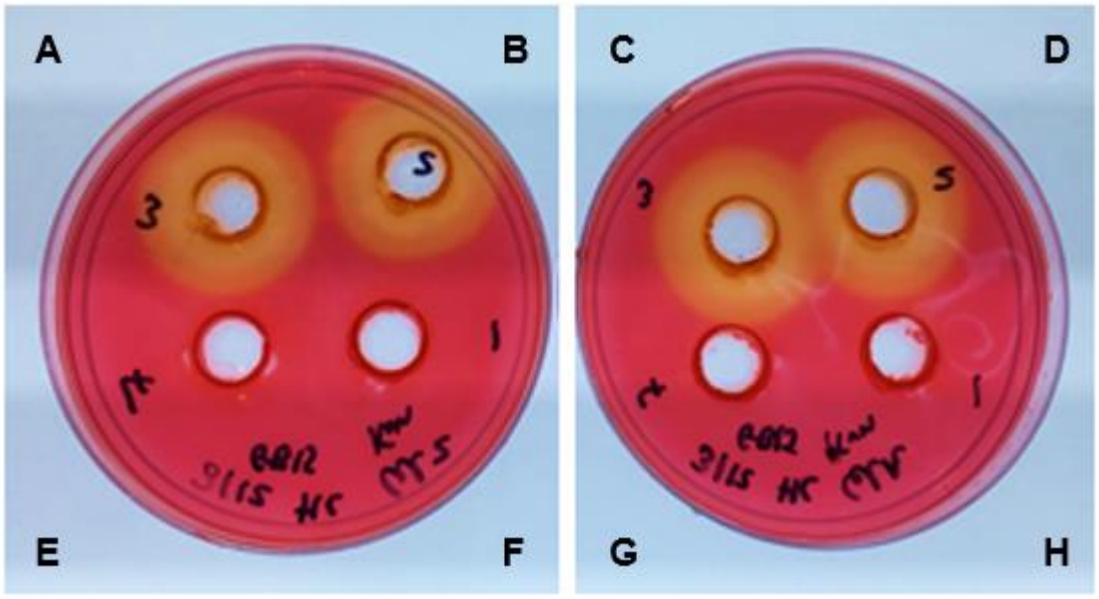


Figure 12 Cellulase activity of CelZ and CelA in *Geobacillus thermoglucosidans* DSM2542
The images represent cellulase activity of supernatant from *Geobacillus thermoglucosidans* DSM2542, with varying plasmids. Wells A and B contained pS797 P_{ldhA}*::CelZ,* wells C and D contained pS797 P_{ldhA}::CelA and wells E, F, G and H were controls containing the empty vector pS797. The supernatant was aliquoted into the well on a CMC plate and incubated at 55^{o}C for 24 h, then stained with congo red. An orange halo surrounding the well indicated cellulase activity, and can be seen in wells: A, B, C and D. It suggests both CelA and CelZ proteins were expressed and present in the supernatant of *Geobacillus thermoglucosidans* DSM2542.

3.2.2 Amplification of P_{00505} and the pS797 plasmid backbone

Polymerase Chain Reaction (PCR) was used to amplify the promoter sequence of the purified pS797 *P_{00505}::GFP* (Part B) and the backbone of the pS797 *P_{ldhA}::CelZ* minus the *P_{ldhA}* promoter (Part A). For part B, primers had been specifically designed to amplify the 100 base pair promoter sequence and add a complementary base pairs to the end sequences of part A.

Amplification of part A resulted in multiple bands when only one was expected of 7516 bp (Figure 13). Each band represented a different sized fragment. This could have been the result of the primer attaching to multiple sites on the plasmid backbone or the polymerase not reading the entire frame. A pure backbone sample was needed for Gibson Assembly, so this product could not go forward.

Gradient PCR was attempted, from 63^{o}C – 72^{o}C, and different annealing times tried. An improved result was seen the higher the temperature. However, none of the annealing temperatures produced a single band as desired (

Figure 14). At 72^{o}C a strongly defined band was present on the gel, in the approximate position of the desired band (7516), however an additional prominent band was also observed below the 500 bp marker. Gel extraction was attempted by carefully cutting the bands out of the gel and using a gel extraction kit. The concentration of the band was low and it did not appear when rerun on an agarose gel.

Amplification of part B, with the addition of complementary tails to the plasmid backbone, was more successful. A single band of 132 bp was expected and the product was in approximately the correct position (Figure 13).

Despite the low concentration of the CelZ PCR product, Gibson Assembly was still attempted. No product appeared when run on an agarose gel (

Figure 15).

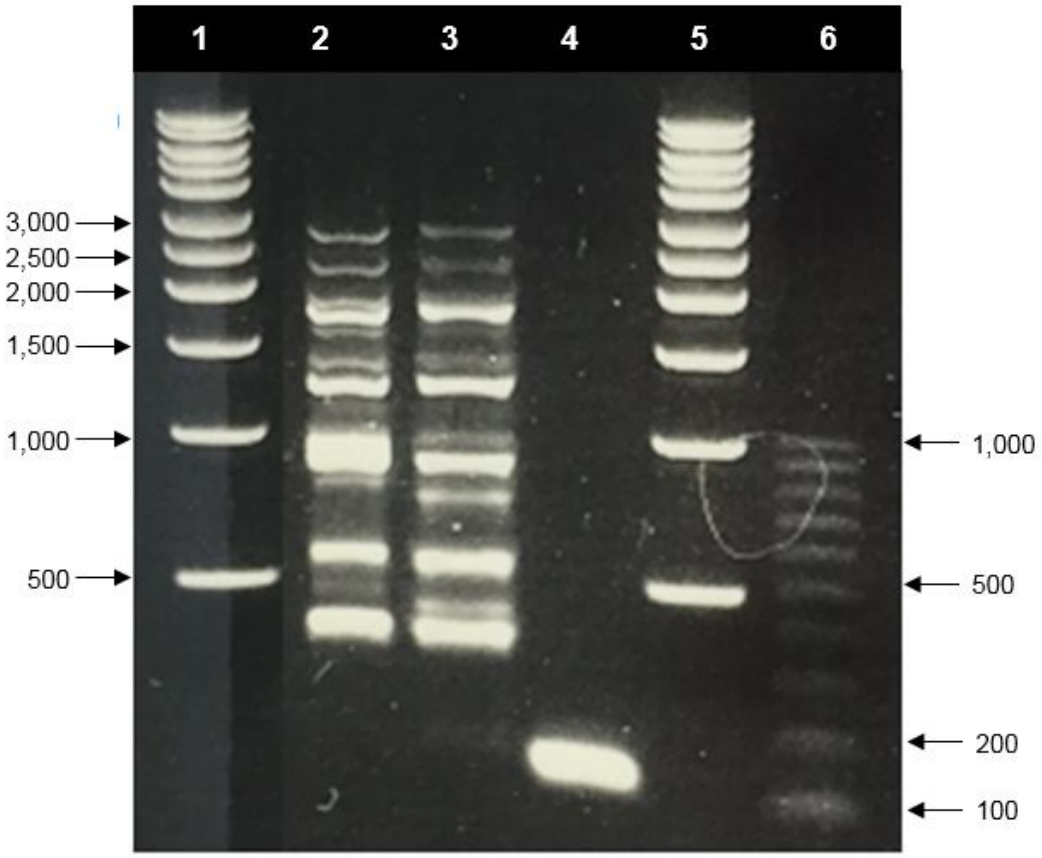


Figure 13 Electrophoresis gel of amplified P_{00505} and pS797 plasmid backbone following PCR
Gel electrophoresis of purified DNA following PCR. Lane 2 and 3 should contain amplified pS797 P_{ldhA}*::CelZ* minus the *ldhA* gene (Part A) and be approximately 7516 bp in length. Lane 3 shows the product digested with PstI and therefore three bands of approximately 552 bp, 2496 bp and 5468 bp were expected. Far more bands of the incorrect size were observed in both columns. Lane 4 should contain amplified P_{00505} plus 'tails' (Part B) and be approximately 132 bp. Part B appears in approximately the correct location. The DNA marker in lanes 1 and 5 was a Direct Load 1Kb DNA ladder from Sigma. The DNA marker in lane 6 was the PCR 100 bp Low Ladder from Sigma Aldrich. Gel was observed under UV light.

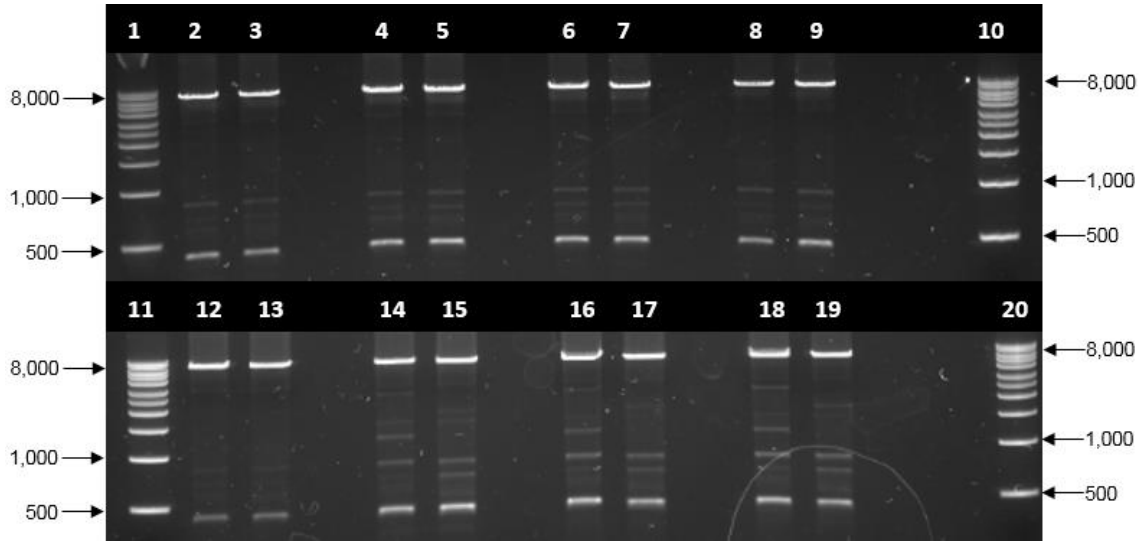


Figure 14 Electrophoresis gel of amplified pS797 plasmid backbone following gradient PCR
Purified PCR product of amplified pS797 CelZ at different annealing temperatures, *(°C): Lane 2&3; 72, 4&5; 71.5, 6&7; 70.5, 8&9; 68.7, 12&13; 66.6, 14&15; 64.8, 16&17; 63.6, 18&19; 63*. The left lane of each pair contained the undigested PCR fragment and the right lane contains the purified PCR fragment digested by EcoRI. Expected band size: 7516 bp – the uppermost band in each lane appears in the appropriate position. Outer lanes (1, 10, 11 and 20); DNA marker was a Direct Load 1Kb DNA ladder from Sigma. Gel was observed under UV light.

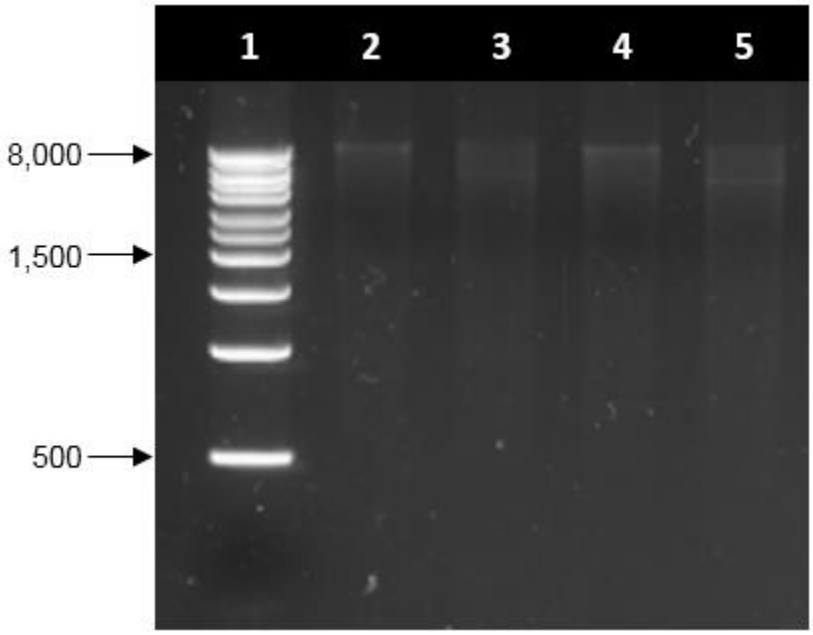


Figure 15 Electrophoresis gel of the product following Gibson Assembly of amplified P_{00505} and pS797 plasmid backbone

Gel electrophoresis of purified DNA following Gibson Assembly of Part A and Part B following PCR. The desired product was pS797 *P_{00505}::CelZ*, of approximately 7616 bp .Lanes 2-5 contained the purified product of Gibson Assembly, however no bands were observed. Lane 1 contained the DNA marker was a Direct Load 1Kb DNA ladder from Sigma. Gel was observed under UV light.

3.2.3 Production of pS797 P_{ldhA}::*CelZ* with alternative promoters by DNA2.0

Following the lack of success, the desired plasmids were ordered from DNA 2.0. An order was placed for pS797 P_{00505}::*CelZ*, pS797 P_{03382}::*CelZ* and pS797 P_{00966}::*CelZ*. However, DNA 2.0 were only able to synthesis one plasmid: pS797 P_{00966}::*CelZ*. They reported that pS797 P_{00505}::*CelZ* and pS797 P_{03382}::*CelZ* were 'probably toxic' and that they were unable 'to get anything to clone without new deletions or chunks of the 5' completely missing'.

pS797 P_{00966}::*CelZ* was taken forward and transformed into *G.thermoglucosidans, E.coli* DH5α and *E.coli S17-1,* as well as being successfully grown and purified from the culture stab.

pS797 P_{00966}::*CelZ* was 7616 bp. SacI should have cut at 423, producing a single band of 7616 bp. EcoRI should have cut at 1706, producing a single band of 7616 bp. PstI should have cut at 2379 and 4875, producing two bands of approximately 2497 and 5119 bp's.

As with previous experiments, bands did not show for *G.thermoglucosidans* (

Figure 16). The bands from the culture stab and *E.coli* DH5α were as expected. However, additional unexpected bands were observed from the *E.coli* S17-1 sample. Lane 11 and 12 clearly show 2 bands when only 1 is expected. The higher band appears in the correct position, however the unexpected lower band is located between the 3000 and 2500 marker. Lane 13 shows 3 bands, none of which appear in the correct position.

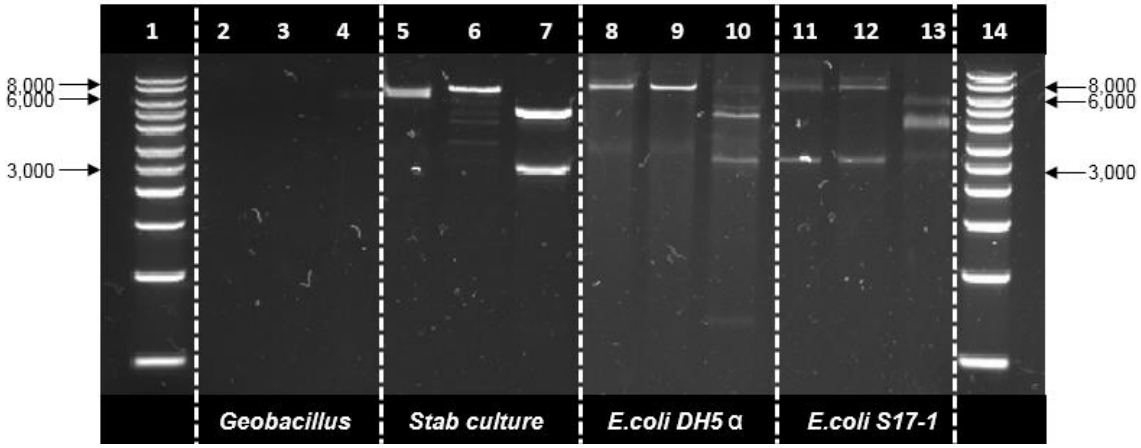


Figure 16 Electrophoresis gel of plasmid pS797 *P_{00966}::CelZ, expressed in a range of hosts,* following synthesis by DNA2.0

Gel electrophoresis of purified pS797 *P_{00966}::CelZ* from a number of host organisms following transformation. Lane 2-4: *Geobacillus thermoglucosidans* DSM2542, Lane 5-7: Stab culture, Lane 8-10: *Escherichia coli* DH5α, Lane11-13: *Escherichia coli* S17-1. The first lane of each section contains the undigested plasmid, the second lane contains an EcoRI digest, and the third lane contains a PstI digest. The expected product, pS797 *P_{00966}::CelZ*, is 7616 bp in length. When digested with SacI a single digest is expected resulting in a single band of 7616 bp. When digested with EcoRI a double digest is expect resulting in two bands of approximately 2497 and 5119 bps. Lane 1 & 14: DNA marker- Direct Load 1Kb DNA ladder from Sigma. Gel was observed under UV light.

3.3 CelZ activity assay

3.3.1 *In Vitro* assay

A congo red assay was conducted with concentrated supernatant of each host. The aim of the assay was to observe if the supernatant had cellulose activity. If activity was observed it would suggest the agarose gel sample may have been contaminated or the plasmid altered in a way that did not stop activity. However, no activity was observed from the main cultures of *G.thermoglucosidans* DSM2542 supernatant or the *E.coli* DH5α and S17-1 supernatant.

The congo red assays were repeated, with freshly transformed cultures, using the pre-culture of *E.coli* DH5α. and S17-1. Activity was observed from the pre-culture *E.coli* DH5α and S17-1 supernatant. However, no activity was observed from the *G.thermoglucosidans* DSM2542

A)

1: pS797 P_{00966}::CelZ
2: Stab culture
3: 2nd generation
4. Non concentrated

B)

1. pS797 P_{ldhA}::CelZ
2.*G.thermoglucosidans*
3.2nd generation
4. Non concentrated

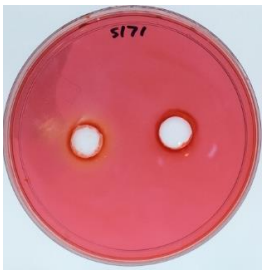
C)

1.pS797 P_{00966}::CelZ
2.*G.thermoglucosidans*
3.2nd generation
4. Non concentrated

D)

1.pS797 P_{00966}::CelZ
2.*E.coli* DH5α
3.2nd generation
4. Non concentrated

E)

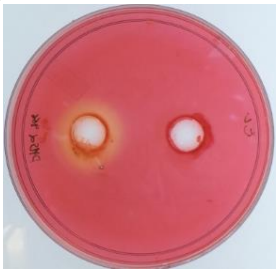


1.pS797 P_{00966}::CelZ
2.E.coli S17-1
3.1st generation
4. Non concentrated

F)

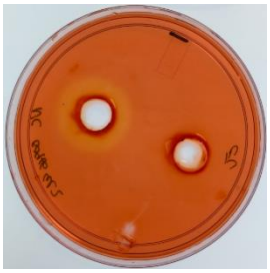
1.pS797 P_{00966}::CelZ
2.*G.thermoglucosidans*
3.1st generation
4. Non concentrated

G)



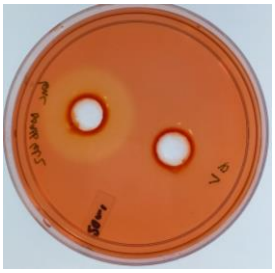
1.pS797 P_{00966}::CelZ
2.*E.coli* DH5α
3.1st generation
4. Non concentrated

H)



1.pS797 P_{00966}::CelZ
2.*E.coli* DH5α
3.1st generation
4. Non concentrated

I)



1.pS797 P_{00966}::CelZ
2.*E.coli* DH5α
3.1st generation
4. Concentrated

Figure 17 Congo red assays to observe cellulase activity

Each image represents cellulase activity of supernatant from a host organism. The left well contains observed culture, the right well contains the control. The four statements under each image refer to the observed culture. 1. Refers to the plasmid expressed: pS797 *P_{00966}::CelZ* or pS797 *P_{ldhA}::CelZ*. 2. Refers to the host organismsthat expressed the plasmid; *Escherichia coli* DH5α*, Escherichia coli* S17-1 *or Geobacillus thermoglucosidans* DSM2542. 3. Refers to either: 1st generation growth - the initial culture following transformation or 2nd generation growth - media that was inoculated with the 1st first generation culture. 4. Refers to if the supernatant was concentrated or not. The supernatant was aliquoted into a well on a CMC plate and incubated at 55°C for 24 h, then stained with congo red. An orange halo surrounding the well indicated cellulase activity, and can be seen in images: A, B, G, H and I. Wild type supernatant of the bacteria was used as a control.

Due to time restraints of the project, *E.coli* DH5α containing pS797 P_{00966}*::CelZ* was taken forward to examine its ability to degrade Whatman 1 filter paper. It was compared to the original laboratory method, using *G.thermoglucosidans* DSM2542 containing pS797 P_{ldhA}*::CelZ*. Concentrated and non-concentrated samples were tested for each host, over 17 days. *G.thermoglucosidans* DSM2542 containing pS797 P_{ldhA}*::CelZ* produced the greatest amount of cellobiose, followed by *E.coli* DH5α containing pS797 P_{00966}*::CelZ.* The *E.coli* host wasn't expected to be as successful as the *G.thermoglucosidans* DSM2542 host at producing cellulose, after an experiment comparing pS797 P_{ldhA}*::GFP* expression in both hosts.

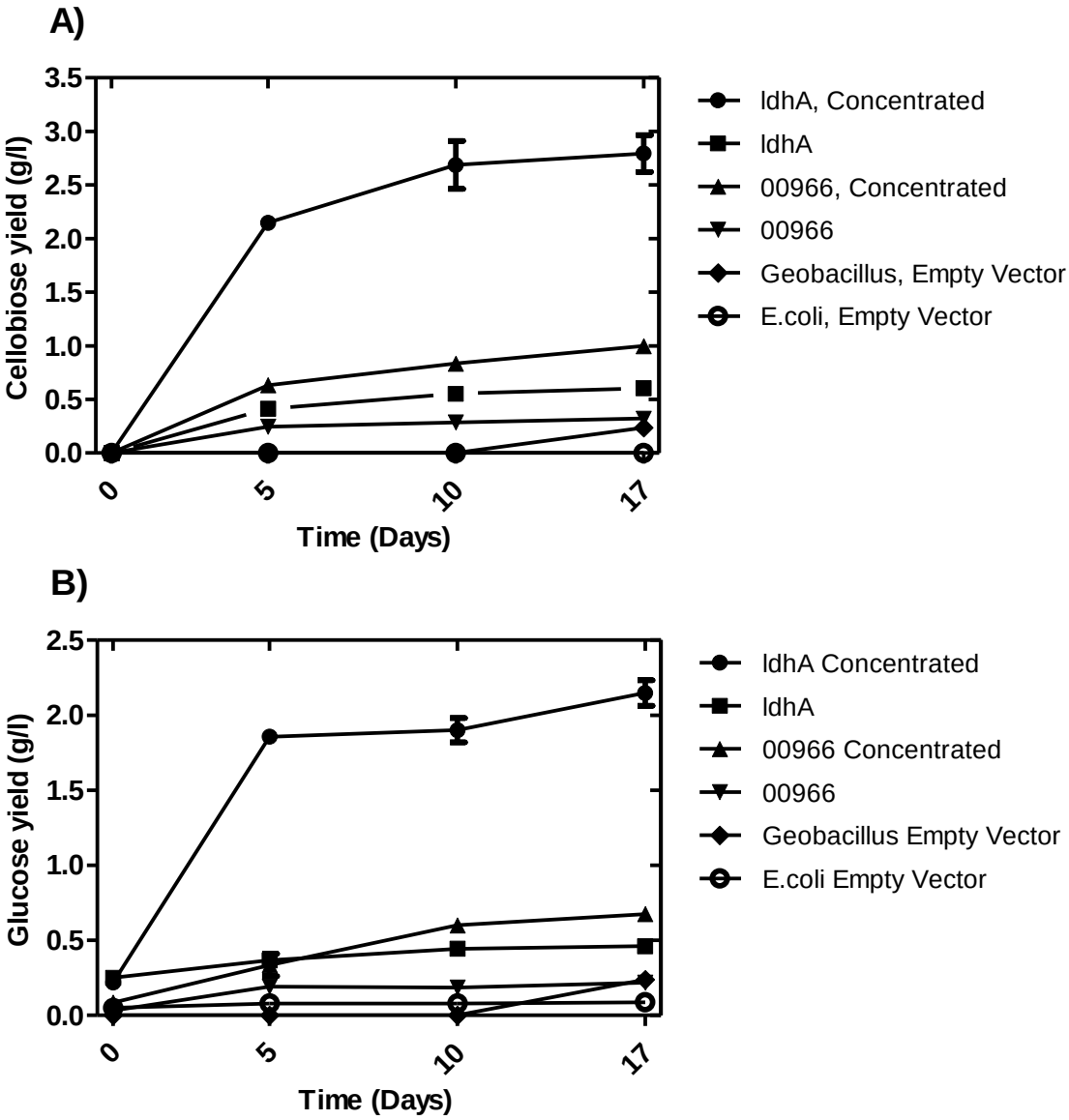


Figure 18 Cellobiose and Glucoses production of 00966-CelZ compared to *ldhA*-CelZ

This figure shows the cellobiose and glucose yield following the digestion of cellulose. Cellulose was digested by the cellulase CelZ in supernatant from *Geobacillus thermoglucosidans* DSM2542 containing pS797 P_{ldhA}*::CelZ* or *Escherichia coli* DH5α containing pS797 P_{00966}*::CelZ*. The supernatant was either used undiluted or 10 fold concentrated. The graphs show yield over a 17 day period with multiple time points. This experiment was carried out with five samples per time point and three repeats. The bars represent the standard error of the mean.

3.3.2 *In vivo* assay

To observe if *G.thermoglucosidans* DSM2542 could grow in the byproducts of cellulose degradation, an experiment was set up as above. Supernatant from *G.thermoglucosidans* DSM2542 containing pS797 P_{ldhA}*::CelZ*, successfully degraded cellulose (Whatman 1 filter paper) over a 7 day period. Then a colony from *G.thermoglucosidans* DSM2542 or *E.coli*, both containing pS797 P_{00966}*::GFP* were added to the culture. Growth of both organisms was successfully observed by GFP Fluorescence. However, this experiment highlighted an issue *in vitro* quantitative analysis. The glucose produced was used to the organisms, meaning glucose yield could not be measured as an indication of cellulose degradation.

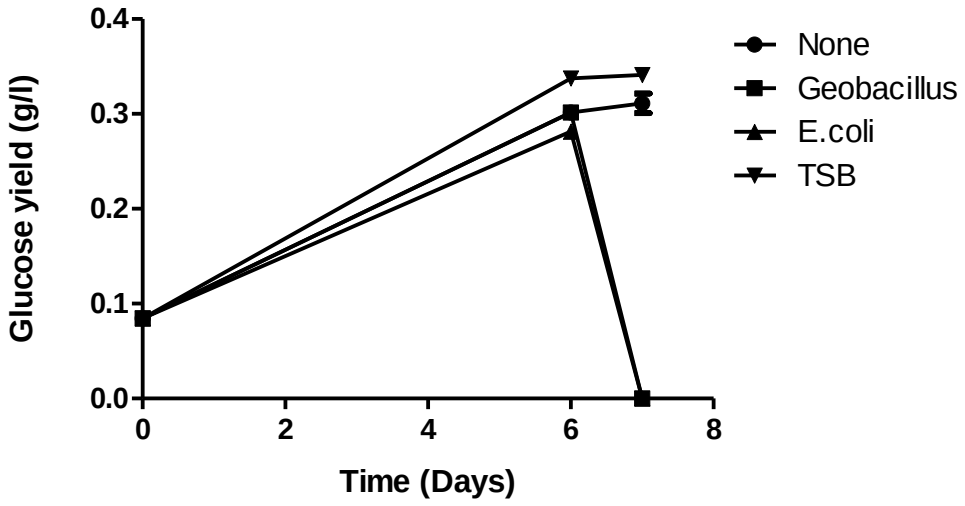


Figure 19 Glucose levels obtained from *in vivo* analysis of *Geobacillus thermoglucosidans as a function of time* in the presence of cellulose

This graph shows the results effect of cellulose degradation on *Geobacillus thermoglucosidans* DSM2542 and the effect of *Geobacillus thermoglucosidans* DSM2542 on glucose production. Degradation was initially observed in all tubes when supernatant from *Geobacillus thermoglucosidans* DSM2542 containing pS797 P_{ldhA}*::CelZ* was incubated with the cellulose (Whatman 1 filter paper) for 6 days. Then a colony from *Geobacillus thermoglucosidans* DSM2542 or *Escherichia coli,* both containing pS797 P_{00966}*::GFP,* were added to two of the tubes. Another tube had TSB added and a further not touched as controls. The tubes with the organisms showed GFP Fluorescence suggesting growth. In addition, the glucose concentration fell dramatically. The controls were not affected. Glucose concentration was measured using HPLC with an Agilent Hi-Flex Na column with a pore size of 10 µm. A 20 µl sample was injected into the HPLC

3.4 Effect of biocatalyst on biomass degradation

In addition, experiments were carried to examine CelZ's ability to degrade cellulose at different temperatures. Three temperatures were examined to see the effect on cellulase activity. The greater the temperature, the greater amount of glucose produced. However, a greater amount of cellobiose was produced at 55^{o}C compared to 65^{o}C.

At 55^{o}C and 65^{o}C the difference in the amount of both cellobiose and glucose produced at 4 and 7 days was similar. But there was a more noticeable change at 33^{o}C, particularly the samples that were left for 21 days.

A large difference was seen between the concentrated and non-concentrated supernatants. More than double the amount of cellobiose was produced and nearly double the amount of glucose was produced by the concentrated supernatant.

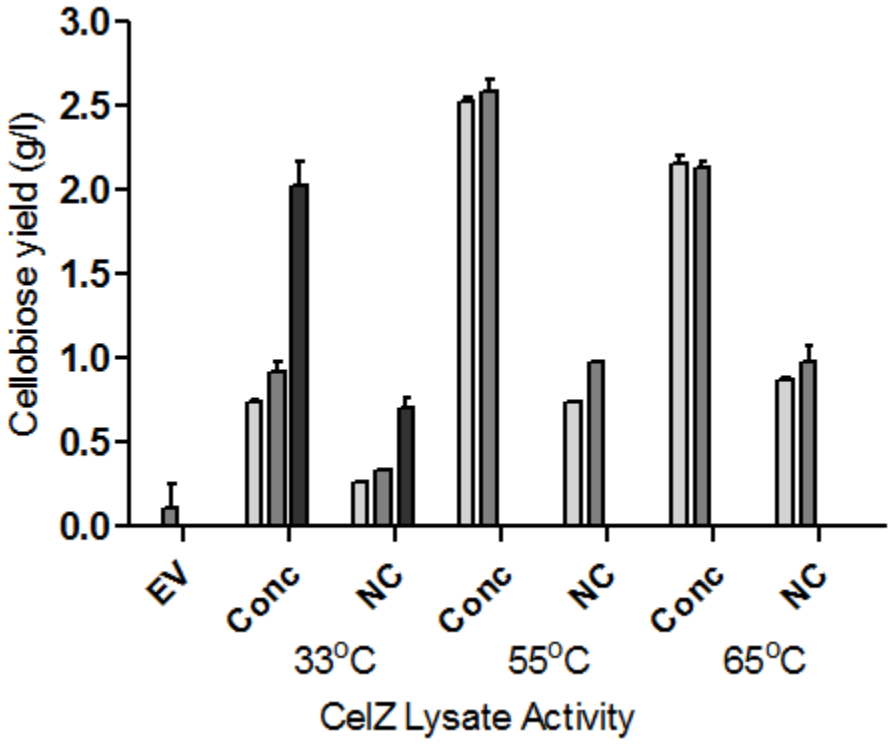


Figure 20 Cellobiose yield following cellulose degradation at 33°C, 55°C and 65°C

This graph shows the effect of different temperatures on Cellobiose production. Cellobiose production results from cellulose degradation by CelZ, from the supernatant of *Geobacillus thermoglucosidans* DSM2542 containing pS797 P_{ldhA}*::CelZ.* Samples were taken for 33°C, 55°C and 65°C. Readings were taken at 4 days (light grey bars) and 7 days (medium grey bars) for all temperatures, then again at 21 days (dark grey bars) for 33°C. 21 day samples could not be taken at all temperatures due to evaporation. This experiment was carried out in triplicate. The error bars represent the standard error of the mean. Cellobiose concentration was measured using HPLC with an Agilent Hi-Flex Na column with a pore size of 10 µm. A 20 µl sample was injected into the HPLC. Concentrated empty vector, pS797, was used as a control at 55°C. Readings were taken at 4, 7 and 21 days, although cellobiose yield was too low to give a reading at day 4 and 21.

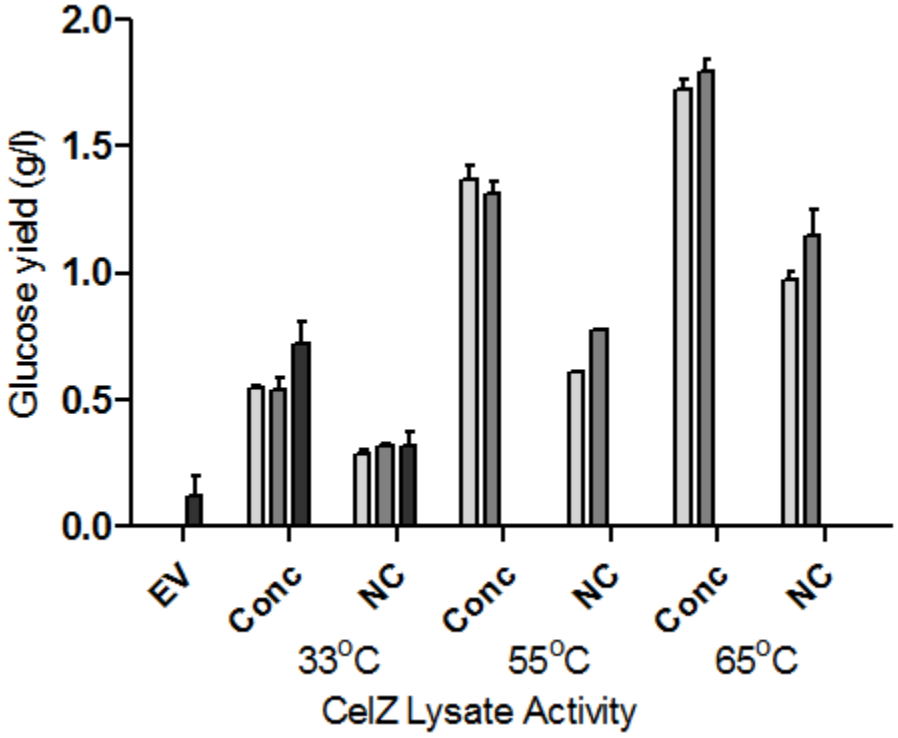


Figure 21 Glucose yield following cellulose degradation at 33°C, 55°C and 65°C

This graph shows the effect of different temperatures on glucose production. Glucose production results from cellulose degradation by CelZ, from the supernatant of *Geobacillus thermoglucosidans* DSM2542 containing pS797 P_{ldhA}*::CelZ.* Samples were taken for 33°C, 55°C and 65°C. Readings were taken at 4 days (light grey bars) and 7 days (medium grey bars) for all temperatures, then again at 21 days (dark grey bars) for 33°C. 21 day samples could not be taken at all temperatures due to evaporation. This experiment was carried out in triplicate. The error bars represent the standard error of the mean. Glucose concentration was measured using HPLC with an Agilent Hi-Flex Na column with a pore size of 10 µm. A 20 µl sample was injected into the HPLC. Concentrated empty vector, pS797, was used as a control at 55°C. Readings were taken at 4, 7 and 21 days, although glucose yield was too low to give a reading at day 4 and 7.

In addition, CelZ's ability to degrade the research laboratories chosen feedstock was examined. Over 4 days at 37^{o}C the concentrated cellulase successfully broke down cellulose in both *Zea mays* ssp. *Mays L* (Corn stover) and *Sorghum bicolor* (Sweet Sorghum).

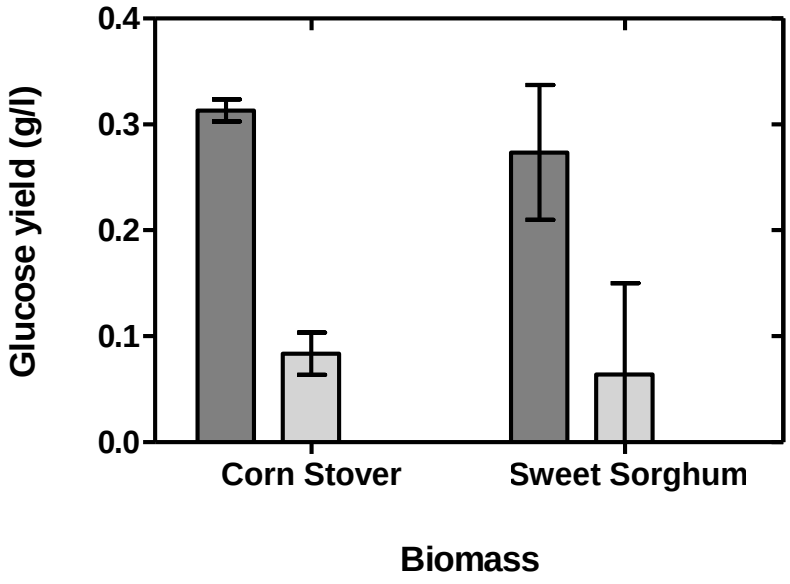


Figure 22 Glucose production by CelZ supernatant

This chart shows the ability of CelZ to produce glucose by the degradation of un-pretreated biomass; *Zea mays* ssp. *Mays L* (Corn stover) and *Sorghum bicolor* (Sweet Sorghum). Glucose production results from cellulose degradation by CelZ, from the supernatant of *Geobacillus thermoglucosidans* DSM2542 containing pS797 P_{ldhA}*::CelZ.* Samples were taken after 4 days incubation at 37ºC, 200 rpm. The dark grey bars represent CelZ supernatant, the light grey bars represent empty vector supernatant. This experiment was carried out in triplicate. The error bars represent the standard error of the mean. Glucose concentration was measured using HPLC with an Agilent Hi-Flex Na column with a pore size of 10 µm. A 20 µl sample was injected into the HPLC.

4. Discussion

4.1 Selection of host, promoter and cellulase for consolidated bioprocessing

Geobacillus thermoglucosidans DSM2542 was the chosen host organism for this project because of favorable traits of the *G.thermoglucosidans spp.* For example *G.thermoglucosidans* DSM2542 can utilize a wide range of carbon sources and has the ability to degrade CMC and xylan making it a suitable organism for cellulase degradation (Ross Kent, MRes, Exeter).Additonallly, *G.thermoglucosidans* is a fast growing organism, capable of growth between 40^{o}C and 70^{o}C and is amenable to bioengineering making it attractive for industrial use (Yu, et al., 2015), (Chen, et al., 2015). As a result, a large amount of research at the University of Exeter Microbial Biofuels Group has gone into developing *G.thermoglucosidans* DSM2542*'s* advanced molecular toolkit.

Initially this report analysed the growth of *G.thermoglucosidans* in three types of media with baffled, vented and non-baffled, non-vented flasks, to test if they affected the organism's growth rate. A slight preference for growth in TSB media in a baffled, vented flask was observed, however it was not significant enough to cause concern if a different media was used (Figure 4). The purpose of using baffled, vented and non-baffled, non-vented was because *ldhA* is not active in baffled flasks. Therefore, the aim was to confirm *G.thermoglucosidans* growth was not the limiting factor. The main focus was expression because *ldhA* is responsive to micro levels of oxygen, hence not active in baffled, vented flasks. Better growth in the presence of oxygen is supported by the data presented in this report, showing wild type *G.thermoglucosidans* DSM2542 favored growth in a baffled vented flask, over a plain bottom non-vented flask.

Cellulase gene expression in *G.thermoglucosidans* DSM2542 was originally controlled by the P_{ldhA} promoter from *G.thermoglucosidans* DSM 2542. The promoter, P_{ldhA} is responsive to oxygen limitation, inducing expression under microaerobic conditions (Bartosiak-Jentys, et al., 2012). Because the host organism is a facultative anaerobe, there may have been conflict in expression conditions between the promoter and host organism. To examine if P_{ldhA} was a limiting factor

for cellulose degradation and promoter expression, it was compared against eight alternative promoters in *G.thermoglucosidans* DSM2542 (Table 1).

GFP gene expression was successfully driven by all eight promoters received from the University of Exeter in host organism, *G.thermoglucosidans* DSM2542. The promoters, including P_{ldhA}, were situated upstream of a GFP gene, allowing for expression to be compared by GFP fluorescence. Of the eight promoters, five outperformed P_{ldhA} with respect to GFP fluorescence::$OD_{600,}$ when grown in a vented flask (Figure 7). The three top performing promoters were:

1) P_{00505} GTGNS_00505 promoter for HTH-type transcriptional repressor GlcR, Tagatose-6-phosphate kinase in *G. thermoglucosidans* DSM 2542.
2) P_{03382} GTGNS_03382 the promoter for cell division protein ZapA from *G. thermoglucosidans* DSM 2542,
3) P_{00966} GTDN_00966 the promoter in front of a hypothetical protein in *G. thermodenitrificans* K1041

All source organisms of the best promoters are from the same bacterial genus, with P_{03382} and P_{00505} from the same species as the research laboratory's chosen host (Table 1). In addition, P_{ldhA} is the lactate dehydrogenase promoter taken from *G.thermoglucosidans* DSM2542, the research laboratory's chosen host. The closer the source species to the host species, the greater the likelihood of the promoter being functional in the host organism, as gene expression is regulated by a variety of genetic and environmental factors (Škrlj, et al., 2010).

The three chosen promoters underwent further analysis against $P_{ldhA,}$ with expression in both vented and non-vented flasks compared. P_{ldhA} had the greatest GFP fluorescence::OD_{600} when grown in a non-vented flask, exceeding the best promoter in a vented flask (P_{00505}) (Figure 8). However, the performance of the three chosen promoters in a vented flask in comparison to a non-vented flask showed less variance than $P_{ldhA,}$ suggesting they are not as sensitive to oxygen concentration as $P_{ldhA.}$ Oxygen sensitivity can be advantageous when controlled as it can enable promoter expression to be controlled (Cronin, et al., 2001). When plasmid expression is suppressed the presence of oxygen, can improve the growth rate of

an organism (Baez & Shiloach, 2014). The organism focuses its energy and resources on growth instead of expression of the promoter. However, controlling oxygen levels increases the complexity of the consolidated bioprocessing, which in turn can increase costs and reduce profitability.

Two functional cellulases (CelA from *C. Bescii* and CelZ from *Clostridium stercorarium*) had been selected by the research laboratory for their promising activity in *G.thermoglucosidans* DSM2542*.* Experimental data shows they are able to digest cellulose (Kun Zhang, Shell Technology Centre Houston). Both proteins come from thermophilic bacterium and are stable at high temperatures. CelZ was chosen for this experiment due to its prior success in the laboratory and its smaller size. An assumption was made that CelZ would be easier to insert into the pS797 plasmid. This assumption was based off reports showing a decreasing recombination efficiency with the increased length of the integrated non homologous DNA fragment (Kung, et al., 2013). CelZ has both *endo*glucanase and *exo*glucanase activities allowing it to work effectively on crystalline cellulosic substrates (Bronnenmeier & Staudenbauer, 1990)**.** CelZ consists of an N-terminal catalytic domain and a cellulase binding fragment including tow putative cellulose binding domains (C and C'). The binding domains each have a different function (Riedel, et al., 1998). Domain C is recognized as a true CBD that anchors the CelZ enzyme to cellulose. However, C' does not have the aromatic residues that contribute to high – affinity cellulose binding. Instead it is recognized as having a thermo-stabilizing effect on the domain. These features were observed in this research paper.

In this project, CelZ exhibited cellulase activity at 33^{o}C, 55^{o}C and 65^{o}C, showing thermo-stability (Figure 21). CelZ's proven stability over a wide range of temperatures make it a suitable candidate for consolidated bioprocessing. Its ability to function at different temperatures reduces the cost involved with heating and cooling processes, removing the need for multiple reactors (Bhalla, et al., 2013). High temperatures allow for product distillation and reduce the risk of contamination. Furthermore, its ability to work at lower temperatures, enables it to function alongside less heat tolerant organisms.

Additionally, the *endo*nuclease and *exo*nuclease activity of CelZ was highlighted by its ability to degrade cellulose. It successfully degraded CMC and Whatman1 filter paper, both sources of pure cellulose (Figure 17 & Figure 18). The higher the temperature the greater glucose yield and the faster glucose was produced (temperature range: 33^{o}C - 65^{o}C) (Figure 21). More significantly, CelZ also exhibited cellulase activity against un-pretreated lignocellulosic biomass from *Zea mays* ssp. *Mays L* (Corn stover) and *Sorghum bicolor* (Sweet Sorghum) (Figure 22). Currently in industry, pretreatment is a critical step of bioconversion to overcome the recalcitrance of lignocellulose. However it is a costly process, and depending on the method of pretreatment can produce toxic or inhibitory byproducts (Parisutham, et al., 2014). Removing the necessity of pretreatment could circumvent these issues.

4.2 Development of molecular toolkit of Geobacillus thermoglucosidans DSM2542

To further enhance the molecular toolkit of *G.thermoglucosidans* DSM2542, an alternative promoter to P_{ldhA} was placed in front of the cellulase gene (Figure 16). The purpose of this experiment was to have a promoter that demonstrates consistent activity with variable oxygen concentrations, whilst also competitively expressing CelZ. $P_{00505,}$ P_{03382} and P_{00966} were proven to be more stable in the presence of oxygen than P_{ldhA} and competitively expressed the GFP protein (Figure 8). To confirm if GFP expression corresponded to the new promoters' ability to produce cellulase, an attempt was made to exchange $P_{00505,}$ P_{03382} and P_{00966} with $P_{ldhA,}$ in pS797 P_{ldhA}*::CelZ*.

Complications occurred with the purification of pS797 P_{ldhA}*::CelZ*. None of the laboratory's glycerol stocks had the functioning plasmid. When purified and observed in a gel the band size was repeatedly too small and one of the PstI digest sites appeared to be missing (Figure 11). This result is supported by previous work by a fellow laboratory student, Ross Kent. The removal of a PstI site is thought to be due to homologous recombination of the KanR into the genome, possibly as a result of repeated freeze-thaw cycles or a limited shelf life due to plasmid instability.

Once an active pS797 P_{ldhA}*::CelZ* was found of the correct size and digest pattern, PCR was attempted to amplify the pS797 P_{ldhA}*::CelZ* backbone, minus P_{ldhA} (Part A) and the promoter sequence, plus complementary tails from pS797 P_{00505}*::GFP* (Part B), to create pS797 P_{00505}*::CelZ.* Initially Phusion high-Fidelity DNA Polymerase was used in the PCR process. Phusion DNA polymerase has an error rate less than 50-fold of that of *Taq* DNA polymerase. It has a processivity-enhancing domain, increasing fidelity and speed. Part B was successfully amplified. However, Part A produced a large number of bands, indicating DNA amplicons of different lengths when only one should have been observed (Figure 13). This could have been due to the primers annealing at multiple locations commonly caused by an incorrect annealing temperature. As a result, gradient PCR was performed, covering a range of increasing annealing temperatures. However this did not improve the purity of the product and multiple bands were still observed (Figure 14). This was attempted multiple times, with a new samples of the CelZ plasmid used each time. Both positive and negative controls yielded the anticipated results. On the gel, one band appeared in approximately the expected positon of the desired fragment. Excision was attempted with a GeneJET gel Extraction kit, however the product was of low concentration and could not be amplified. Gibson assembly was attempted but did not produce a product.

An alternative DNA polymerase was used at the University of Exeter, due to industrial patent restriction's in the USA. However, results were no more successful with Q5 High-Fidelity DNA Polymerase.

Finally, synthesis by DNA2.0, a synthetic biology company that had been involved in the design and production of pS797 P_{ldhA}*::CelZ* attempted to synthesize all three plasmids; pS797 P_{00505}*::CelZ,* pS797 P_{003382}*::CelZ,* and pS797 P_{00966}*::CelZ*. They were only able to synthesize P_{00966}*::CelZ*, reporting the other two plasmids were too unstable.

There was an inability to clone pS797 P_{00505}*::CelZ* and pS797 P_{003382}*::CelZ* without deletions or segments of the 5' end missing. This suggests toxicity problems of the plasmid, which may lead to future limitations of pS797.

The synthesized plasmid was successfully grown in its stab culture host and transformed into *E.coli* DH5α and *E.coli* S17-1. In the stab culture and *E.coli* DH5α the plasmid appeared in the expected position on the gel following electrophoresis (Figure 16). However, in *E.coli* S17-1 the plasmid appeared to produce two multiple bands when only one was expected. The upper band appeared in the correct location, however the second band appeared level with the 4,000 bp marker, suggesting the CelZ gene may have been deleted. In addition, cellulase activity was only observed in 1st generation growth of *Escherichia coli* DH5α and not at all in *E.coli* S17-1 indicating the plasmid was not stable over time (Figure 17). As a result, activity was not observed in *G.thermoglucosidans*, as the plasmid would not have been stable for enough for transformation.

Therefore, for the comparison between P_{00966} and $P_{ldhA,}$ pS797 P_{00966}*::CelZ* was expressed in *E.coli* S17-1 and pS797 P_{ldhA}*::CelZ* in *G.thermoglucosidans*. Glucose and cellobiose yield for P_{00966} was significantly lower than that of P_{ldhA} (Figure 18). This result was expected because *E.coli* DH5α is designed for producing high amounts of plasmid DNA and maintaining plasmid stability. It is a stable plasmid host, not designed for maximizing protein expression. These characteristics indicate how unsuitable pS797 P_{00966}*::CelZ* is for future development. It is too unstable for industrial action where long term population stability is required and pS797 P_{ldhA}*::CelZ* was far superior at expressing CelZ and degrading cellulose.

4.3 Conclusion and further work

In conclusion, expression of CelZ does confers cellulolytic activity to *G.thermoglucosidans* DSM2542. An increase in CelZ concentration significantly increased the cellulolytic activity observed. However, current glucose and cellobiose yield is too low to be industrially viable.

To increase expression and hence increase yield, pS797 P_{ldhA}*::CelZ* could be optimized. The laboratory is looking to generate a codon optimized CelZ to increase the abundance of the CelZ protein. Concentrating the supernatant increased cellulose degradation, therefore increasing CelZ abundance should also have a positive effect of glucose yield. Unfortunately, pS797 isn't flexible to alteration. It

could be enhanced by the advancement of interchangeable biological DNA building-blocks, for example BioBricks, improving the ease and speed of specific gene transfer (Lewens, 2013). Additionally, novel approaches for pathway optimization have improved the efficiency and precision of commercial DNA synthesis, at an affordable price. Both methods can help with high throughput screening of alternative promoters and cellulase genes.

Optimization of CelZ expression involves developing a balanced pathway with a high flux towards cellulase expression. Development of the laboratory's consolidated bioprocessing process could include fine-tuning the pathway by exploring gene dosage, translational engineering and transcriptional engineering (Jeschek, et al., 2017). Experiments could be developed to better reflect industrial reaction conditions. For example, using bioreactor-based fermentations, using a fed-batch feeding strategy to enhance cell growth and reactor productivity. These experiments should develop our understanding of the interaction between growth, expression and oxygen tension.

Furthermore, this research project only looked at the degradation of cellulose. Lignocellulosic biomass is composed of three polymers and for maximum efficient and productivity hemicellulose and xylan would also be degraded. The polymers currently block cellulose degradation, so this development would enhance yield do to utilization of new biomass and improved degradation of cellulose.

The ability of *G.thermoglucosidans* DSM2542, expressing pS797 P_{ldhA}*::CelZ*, to degrade un-pretreated lignocellulosic biomass is an exciting development with the laboratory's host organism. It shows progress and potential towards the desired, economically-viable consolidated bioprocessing, with increased knowledge and research.

5. Acknowledgements

I would like to thank my supervisor Professor John Love for his guidance both in the laboratory and with the write up on this thesis. Secondly, to Professor Robert Lee and Professor David Parker for hosting me in the Biodomain Lab at Shell Technology Centre Houston. It was a good experience and taught me a lot.

In addition, I would like to thank Dr. Chloe Singleton and Dr. James Winkler for their continued enthusiasm and encouragement. And Dr. Trevor Zurroff and Dr. Kun Zhang for their advice and insight.

Thank you.

6. Bibliography

Adams, G. & et. al, 2016. *International Energy Outlook 2016,* Washington: U.S. Energy Information Administration.

Asif, M. & Muneer, T., 2007. Energy supply, its demand and security issues for developed and emerging economies. *Renewable and Sustainable Energy Reviews,* pp. 1388-1414.

Baez, A. & Shiloach, J., 2014. *Effect of elevated oxygen concentration on bacteria, yeasts, and cells propagated for production of biological compounds,* s.l.: Microbial Cell factories.

Bajpai, P., 2016. Pretreatment of Lignocellulosic Biomass for Biofule Production. In: *Green Chemistry for Sustainability.* s.l.:Springer, pp. 7-8.

Bartosiak-Jentys , J., Eley, K. & Leak, D. J., 2012. Application of pheB as a Reporter Gene for Geobacillus spp., Enabling Qualitative Colony Screening and Quantitative Analysis of Promoter Strength. *Applied and Environmental Microbiology,* pp. 5945-5947.

Bartosiak-Jentys, J., Eley, K. & Leak, D. J., 2012. Application of pheB as a Reporter Gene for Geobacillus spp., Enabling Qualitative Colony Screening and Quantitative Analysis of Promoter Strength. *Applied and Environmental Microbilogu,* 78(16), pp. 5945-5947.

Bartosiak-Jentys, J., Hussein, A. H., Lewis, C. J. & Leak, D. J., 2013. Modular system for assessment of glycosyl hydrolase secretion in Geobacillus thermoglucosidasius. *Microbiology,* Volume 159, pp. 1267-1275.

Bhalla, A. et al., 2013. Improved lignocellulose conversion to biofuels with thermophilic bacteria and thermostable enzymes. *Bioresource Technology,* Volume 128, pp. 751-759.

Bhatia, S. K., Kim, S.-H., Yoon, J.-J. & Yang, Y.-H., 2017. Current status and strategies for second generation biofuel production using microbial systems. *Energy Conversion and Management,* Volume 148, pp. 1142-1156.

Binod , P. & Pandey, A., 2015. Chapter 1 - Introduction. In: P. Binod, A. Pandey & C. Larroche, eds. *Pretreatment of Biomass .* Amsterdam: Elsevier, pp. 3-6.

Birol, F. & et. al, 2015. *World Energy Outlook 2015,* Paris: International Energy Agency.

Birol, F. & et. al, 2016. *Key World Energy Statistics,* Paris: International Energy Agency.

Brennan, L. & Owende, P., 2010. Biofuels from microalgae-A review of technologies for production, processing and extractions of biofuels and co-products. *Renerable and Sustainable Energy Reviews,* 14(2), pp. 557-577.

Bronnenmeier, K. & Staudenbauer, W. L., 1990. Cellulose hydrolysis by a highly thermostable endo-l,4-beta-glucanase (Avicelase I) from Clostridium stercorarlum. *Enzyme and Microbial Technology,* pp. 431-436.

Brumm, 2015. Complete genome sequence of Geobacillus thermoglucosidasius C56-YS93, a novel biomass degrader isolated from obsidian hot spring in Yellowstone National Park. *Standards in Genomic Sciences.*

Carder, J., 1986. Detectio and Quantitation of Cellulase by Congo Red Staining of Substrates in a Cup-Plate Diffusion Assay. *Analytical Biochemistry,* pp. 75-79.

Castro, 2016. Global Energy Demand and Its Geopolitical and Socioeconomic Implications: Which Role Would Shale Resources Have?. *The Open Petroleum Engineering Journal,* pp. 47-54.

Chen, J., Zhang, Z., Zhang, C. & Yu, B., 2015. Genome sequence of Geobacillus thermoglucosidasius DSM2542, a platform hosts for biotechnological applications with industrial potential. *Journal of Biotechnology,* Volume 216, pp. 98-99.

Correa, D. F. et al., 2017. Biodiversity impacts of bioenergy production: Microalgae vs. first generation biofuels. *Biodiversity impacts of bioenergy production: Microalgae vs. first generation biofuels,* Volume 74, pp. 1131-1146.

Cronin, C. A., Gluba, W. & Scrable, H., 2001. The lac operator-repressor system is functional in the mouse. *Genes & Development,* 15(12), pp. 1505-1517.

European Comission, 2017. *Energy Strategy and Energy Union.* [Online] Available at: https://ec.europa.eu/energy/en/topics/energy-strategy-and-energy-union [Accessed 06 09 2017].

Fong, J. C. N., 2006. Isolation and characterization of two novel ethanol-tolerant facultative-anaerobic thermophilic bacteria strains from waste compost. *Extremophiles,* pp. 363-372.

Fortman, J. L. et al., 2008. Biofuel alternatives to ethanol: pumping the microbial well.. *Trends in biotechnology,* 26(7), pp. 375-381.

Fukuoka, A. & Dhepe, P. L., 2006. Catalytic Conversion of Cellulose into Sugar Alcohols. *Angewandte,* 118(31), pp. 5285-5287.

Ghosh, S., Chowdhury, R. & Bhattacharya, P., 2017. Sustainability of cereal straws for the fermentative production of second generation biofuels: A review of the efficiency and economics of biochemical pretreatment processes. *Applied Energy,* Volume 198, pp. 284-298.

Hinestroza , J. & Netravali, A. N., 2014. Chapter 1: Cellulose Nanofibre Based Material. In: *Cellulose Based Composites: New Green Nanomaterials.* Weinheim: Wiley-VCH, p. 3.

Howard, T. P. et al., 2013. Synthesis of customized petroleum-replica fuel molecules by targeted modification of free fatty acid pool in Escherichia coli. *PNAS,* 110(19), pp. 7636-7641.

Jeschek, M., Gerngross, D. & Panke, S., 2017. Combinatorial pathway optimization for streamlined metabolic engineering. *Current Opinion in Biotechnology,* Volume 47, pp. 142-151.

Kaltschmitta, M., Thran, D. & Smith, K. R., 2003. Renewable Energy from Biomass. *Encyclopedia of Physical Science and Technology ,* Volume 3, p. Pages 203–228.

Keegstra, K., 2010. Plant Cell Walls. *Future Perspectives in Plant Biology,* Volume 154, pp. 483-486.

Kent, R., 2016. *Screening and Engineering of Geobacillus spp. for Consolidated Bioprocessing of Lignocellulosic Biomass,* Exeter: University of Exeter.

Kung, S. H., Retchless, A. C., Kwan, J. Y. & Almeida, R. P. P., 2013. Effects of DNA Size on Transformation and Recombination Efficiencies in Xylella fastidiosa. *Applied and Environmental Microbiology,* 79(5), pp. 1712-1717.

Lee, C., 2005. Energy consumption and GDP in developing countries: A cointegrated panel analysis. *Energy Economics ,* p. 415–427.

Lewens, T., 2013. From bricolage to BioBricks™: Synthetic biology and rational design. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences,* 44(4), pp. 641-648.

Lin, F.-M., Marsh, N. G. & Lin, X. N., 2015. Recent progress in hydrocarbon biofuel synthesis: Pathways and enzymes. *Chinese Chemical Letters,* 26(4), pp. 431-434.

Lin, P. P. et al., 2014. Isobutanol production at elevated temperatures in thermophilic Geobacillus thermoglucosidasius. *Metabolic Engineering,* Volume 24, pp. 1-8.

Lin, T. R. L. D. S. K. M. S. Y. G. T. L. S. a. T. K., 2016. Biomass feedstock preprocessing and long-distance transportation logistics. *GCB Bioenergy,* pp. 160-170.

Mascal, M. & Nikitin, E. B., 2008. Direct, High-Yield Conversion of Cellulose into Biofuel. *Angewandte Chemie,* 120(41), pp. 8042-8044.

Mohr, S. A. et al., 2015. Projection of world fossil fuels by country. *Fuel,* pp. 120 - 135.

Nazina , T. N. & et al., 2001. Taxonomic study of aerobic thermophilic bacilli: descriptions of Geobacillus subterraneus gen. nov., sp. nov. and Geobacillus uzenensis sp. nov. from petroleum reservoirs and transfer of Bacillus

stearothermophilus, Bacillus thermocatenulatus, Bacillus th. *International journal of systematic and evolutionary microbiology,* pp. 433 - 446.

Nazina, 2001. Taxonomic study of aerobic thermophilic bacilli: descriptions of Geobacillus subterraneus gen. nov., sp. nov. and Geobacillus uzenensis sp. nov. from petroleum reservoirs and transfer of Bacillus stearothermophilus, Bacillus thermocatenulatus, Bacillus th. *International Journal of Systematic and Evolutionary Microbiology ,* pp. 433-46.

Niziolek, A. M. et al., 2014. Coal and Biomass to Liquid Transportation Fuels: Process Synthesis and Global Optimization Strategies. *Industrial & Engineering Chemistry Research,* 53(44), pp. 17002-17025.

Parisutham, V., Kim, T. H. & Lee, S. K., 2014. Feasibilities of consolidated bioprocessing microbes: From pretreatment to biofuel production. *Bioresource Technology,* Volume 161, pp. 431-440.

Peralta-Yahya, P. P., Zhang, F., del Cardayre, S. B. & Keasling, J. D., 2012. Microbial engineering for the production of advanced biofuels. *Nature,* 888(7411), pp. 320-328.

Raftery, J. P. & Karim, M. N., 2017. Economic viability of consolidated bioprocessing utilizing multiple biomass substrates for commercial-scale cellulosic bioethanol production. *Biomass and Bioenergy,* Volume 103, pp. 35-46.

Rastogi, 2010. Characterization of thermostable cellulases produced by Bacillus and Geobacillus strains. *Bioresource Technology,* pp. 8798-8806.

Riedel, K., Ritter, J., Bauer, S. & Bronnen, K., 1998. The modular cellulase CelZ of the thermophilic bacterium Clostridium stercorarium contains a thermostabilizing domain. *FEMS Microbiology Letters,* 164(2), pp. 261-267.

Riedel, K., Ritter, J., Bauer, S. & Bronnenmeier, K., 1998. The modular cellulase CelZ of the thermophilic bacterium Clostridium stercorarium contains a thermostabilizing domain. *FEMS Microbiol Letter,* pp. 164: 261-267.

Rodionov, M. V. et al., 2017. Biofuel production: Challenges and opportunities. *International Journal of Hydrogen Energy,* 42(12), pp. 8450-8461.

Saladini, F. et al., 2016. Guidlines for emerging evaluation of first, second and third generation biofuels. *Renewable and Sustainable Energy Reviews,* Volume 66, pp. 221-227.

Sang-Bing , T. et al., 2017. Models for forecasting growth trends in renewable energy. *Renewable and Sustainable Energy Reviews,* Volume 77, pp. 1169-1178.

Škrlj, N., Vidrih, Z. & Dolinar, M., 2010. A universal approach for promoter strength evaluation supported by the web-based tool PromCal. *Analytical Biochemistry,* 396(1), pp. 83-90.

Srinophakun, P. et al., 2017. Feedstock production for third generation biofules through cultivation of Arthrobacter AK19 under stress conditions. *Journal of Cleaner Production,* Volume 142, pp. 1259-1266.

Taylor, M. P. et al., 2009. Thermophilic ethanologenesis: future prospects for second-generation bioethanol production. *Trends in Biotechnology,* 27(7), pp. 398-405.

Teather, R. M. a. W. P., 1982. Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. *Applied Environmental Microbiology,* pp. 777-780.

Uhlenbrook, S., 2007. Biofuel and water cycle dynamics: what are the related challenges for hydrological processes research?. *Hydrological Processes,* 21(26), pp. 3647-3650.

Walker, J. T., 2003. Fossil Fuel Energy. In: *Encyclopedia of Agricultural, Food, and Biological Engineering.* Columbus: Marcel Dekker, p. 382.

Appendix

pS797 *Pxxxxx CelZ* sequence

Below is the pS797 P_{ldhA} *CelZ* sequence. The current cellulase promoter (*ldhA)* is highlighted in red and is 151 bp long. The aim was to replace this sequence with the following promoters (100 bp long):

- **00505:**

  TTTTTTTGAGTGTTTTTGATTGAAAATGATTGATTTTTAAATCGTTTTCAT
  TTATGATAAGGGTACAAAGTAATCATTTCCATCCAAGGAGGTGGGATCG

- **00966:**
  TTTTTACATAATCTTTTTTGGCCATCCTCCCCTCTGACATCTCTCCGCTT
  TCCATTTTTCTTTCATTCATGGTATGATGAATAATGGAGGTGCCAGATTT

- **03382:**
  CATCGTTTTTTCTTAGGCATATCTAGTAGTCGTATCTTTTCATTATATCA
  GCATTCATGTTATGATAAACAATAGGATTTTGAAGAATGGGGGAAATCA
  T

pS797 P_{ldhA} *CelZ* sequence

>pNC2

AAGCTTTCGAG

>AMP Promoter

CGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTC
ATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGT

>bla

ATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGC
CTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGA
TCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAG
ATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATG

>bla, ampR

ATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGC
CGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTT
GAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAG
AATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTT
CTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGG
GGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCAT
ACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTG

CGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAAT
AGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTT
CCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTC
GCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGT
TATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATC
GCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAA

>pNC2

CTCGAGCTGTCAATAATC

>ColE ori

TCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCC
GTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGC
TGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCA
AGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATAC
CAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCT
GTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTG
CCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACC
GGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCA
GCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATG
AGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAG
CGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACG
CCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGA
TTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACG
CGGCC

>pNC2

TTTTTA

>M13 REV

CAGGAAACAGCTATGACCATG

>Biobricks non coding prefix

GAATTCGCGGCCGCTTCTAGAG

>ldhA.2

CTGCCTCGTCCATTTTTTTGCTTAATGGAGGTTGTCATGAAAATGACAAACAAC
GTCCAAACAATTGCCATAATCGTTTACGCATAGTTTCGATTTCATCGCGTAAAA
TAATTTGTGAATGTATTCACAATAATAAGAAGGGAGAATAGTG

>C. stercorarium CelZ.2.1.1.1

ATGCGCAAGTTTTGGAGCTTTGCGATTATCATTAGCCTGCTGGTGACGGGTCT
GTTCATTCATACCCCGAAGGCCGAGGCAGCCGGCTACAATTACGGTGAGGCT
CTGCAAAAAGCCATTATGTTCTACGAGTTCCAGCGTTCTGGCAAATTGCCTGA
GAACAAGCGTGACAACTGGCGCGGTGATTCTGGTTTGAATGACGGTGCAGAC
GTCGGCCTTGATCTGACCGGTGGCTGGTATGATGCGGGCGATCACGTTAAGT
TTAATCTGCCGATGGCGTACAGCCAAACCATGTTGGCTTGGGCGGCATACGA

AGCAGAAGAAGCGCTGGAAAGATCCGGCCAGATGGGTTATCTGCTGGACGC
AATCAAATGGGTTAGCGATTACCTGATTAAGTGTCACCCAAGCCCGAATGTGT
TCTACTATCAGGTCGGCGACGGTCACCTGGATCACTCCTGGTGGGGTCCGGC
TGAGGTTATGCAGATGGATCGTCCGGCGTACAAAGTTGATCTGGCGAACCCG
GGTAGCACCGTGGTCGCGGAAGCTGCTGCAGCTCTGGCAAGCGCAGCCGTC
GTGTTTGCGGATCGTGATCCGGCCTATGCGGCGACGTGTATTCAGCATGCGA
AAGAACTGTACAACTTTGCGGAGATTACCAAGAGCGACAGCGGTTACACCGC
AGCCAGCGGCTTCTATGACTCTCATAGCGGTTTCTACGATGAGCTGAGCTGG
GCTGGCGTGTGGCTGTACCTGGCGACCGGTGACGAAACCTACTTAAACAAAG
CAGAGCAGTATGTTGCGTACTGGGGCACCGAACCGCAGACGAACATCATCAG
CTACAAATGGGCGCACTGCTGGGATGACGTACACTATGGCGCCTGCTTGCTG
CTGGCGAAGATCACCGGTAAACAAATCTATAAAGAAGCGATTGAGCGTCACTT
GGACTACTGGAGCGTTGGTTACAATGGCGAGCGTGTGCACTATACGCCGAAA
GGTCTGGCGTGGCTGGACAGCTGGGGTTCGCTGCGTTATGCCACGACCACT
GCATTTCTGGCCTCCGTTTACGCGGACTGGGAAGGTTGCTCGCGTGAGAAAG
CGGCGATTTACAACGATTTCGCGAAGCAGCAAATCGACTACGCTCTGGGTTC
AAGCGGCCGTTCTTACGTTGTTGGTTTCGGCGTCAACCCGCCGAAGCGTCCG
CATCACCGCACCGCGCATTCCAGCTGGGCGGATTCTATGAGCGTGCCGGACT
ACCATCGCCATGTTCTGATCGGTGCTCTGGTCGGTGGTCCTGGCAAAGATGA
TAGCTATACCGACGACATCAATAACTATATCAATAACGAAGTCGCGTGTGATT
ACAACGCAGGCTTTGTCGGTGCACTGGCCAAAATGTATGAAGATTATGGTGG
CAGCCCGATTCCAGACCTGAACGCATTTGAAGAAATCACTAATGACGAGTTCT
TTGTTATGGCGGGTATTAACGCTAGCGGTCAGAATTTCATCGAAATCAAGGCC
CTGCTGCACAATCAGAGCGGTTGGCCGGCACGTGTTGCAGACAAACTGAGCT
TTCGTTACTTTGTTGATCTCACCGAGCTGATTGAGGCAGGCTACAGCGCAAGC
GATGTGACCATTACCACTAACTACAACGCGGGTGCGAAGGTTACGGGTCTGC
ACCCTTGGAATGAAGCCGAGAACATCTACTATGTCAACGTGGATTTCACGGGT
ACCAAGATTTACCCGGGTGGTCAATCTGCCTATCGCAAAGAAGTCCAGTTCC
GCATCGCCGCACCGCAAAACACGAACTTTTGGAATAACGACAATGACTACTCC
TTCCGTGACATCAAAGGTGTGACGAGCGGCAATACGGTCAAAACCGTCTATAT
CCCGGTGTATGACGATGGCGTTCTGGTCTTTGGTGTGGAGCCAGAGGGTGG
CAGCGGTGAGAATAACAGCTCCATCAGCATTACCAATGCGACCTTTGACAAGA
ATCCGGCGAAACAAGAGAATATTCAAGTGGTCATGAACCTGAACGGTAACAC
GTTGAACGGCATTAAGTATGGTAACACCTATCTGCGTGAGGGTACCGACTATA
CGGTGAGCGGCGACACCGTGACGATCTTGAAGTCTTTTCTGAACAGCTTTGA
TACCAGCACCGTTCAACTGATTTTTGATTTCAGCGCGGGTCGTGACCCGGTTC
TGACCGTTAACATTATTGACACCACCACGAGCGCCTCCATCGTCCCGACTACT
GCGGATTTCGATAAGAATCCTGATGCAAGCCGCGATGTCAAGGTGAAACTGG
TTCCGAATGGCAACACGCTGCTGGCAGTGAAGAAAGACGGCGAGGCGCTGG
TGTTGGGCCGTGATTACAGCATCGATGGCGATGAAGTTACCATTTTCCGTGAG
TATTTGGCGGACCAGCCAGTGGGTCGTGTTACCCTGACCTTCGACTTCGACC
GTGGCACGGATCCGGTCCTGACCATCAACATCACCGACAGCCGCCAGGTTGA
AACGGGTGTGATTCAAATTCAGATGTTTAATGGTAATACGAGCGATAAAACCA
ATGGTATCATGCCGCGCTACCGTCTGACGAATACCGGTACCACCCCGATTCG
CCTGTCCGATGTGAAGATCCGCTATTACTACACTATTGACGGTGAAAAGGACC
AGAATTTCTGGTGCGACTGGTCCTCAGTCGGCTCGAACAACATTACGGGCAC
CTTCGTAAAGATGGCGGAGCCGAAAGAGGGCGCAGACTATTACCTGGAGACT
GGTTTTACGGACGGTGCGGGCTATCTGCAACCGAATCAAAGCATTGAAGTCC

AGAATCGCTTCAGCAAGGCCGATTGGACGGACTACATCCAAACGAATGACTA
CAGCTTCAGCACCAATACGAGCTATGGCAGCAACGACCGTATTACCGTTTACA
TTAGCGGTGTGTTGGTGAGCGGTATCGAGCCG

>Stop

TAATAA

>terminator

AAATGGAAATGATTCACACCCGATGTGGACTTGCCTCATCGGGTGTTTTTTTG
CTATCAAAAC

>Biobricks suffix

TACTAGTAGCGGCCGCCTGCAG

>M13 FWD

ACTGGCCGTCGTTTTACA

>pNC2

AGATCC

>KanR TK101

TCAAAATGGTATGCGTTTTGACACATCCACTATATATCCGTGTCGTTCTGTCCA
CTCCTGAATCCCATTCCAGAAATTCTCTAGCGATTCCAGAAGTTTCTCAGAGT
CGGAAAGTTGACCAGACATTACGAACTGGCACAGATGGTCATAACCTGAAGG
AAGATCTGATTGCTTAACTGCTTCAGTTAAGACCGAAGCGCTCGTCGTATAAC
AGATGCGATGATGCAGACCAATCAACATGGCACCTGCCATTGCTACCTGTACA
GTCAAGGATGGTAGAAATGTTGTCGGTCCTTGCACACGAATATTACGCCATTT
GCCTGCATATTCAAACAGCTCTTCTACGATAAGGGCACAAATCGCATCGTGGA
ACTTTTGGGCTTCTACCGATTTAGCAGTTTGATACACTTTCTCTAAGTATCCAC
CTGAATCATAAATCGGCAAAATAGAGAAAAATTGACCATGTGTAAGCGGCCAA
TCTGATTCCACCTGAGATGCATAATCTAGTAGAATCTCTTCGCTATAAAAATTC
ACTTCCACCTTCCACTCACCGGTTGTCCATTCATGGCTGAACTCTGCTTCCTC
TGTTGACATGACACACATCATCTCAATATCCGAATAGGGCCCATCAGTCTGAC
GACCAAGAGAGCCATAAACACCAATAGCCTTAACATCATCCCCATATTTATCC
AATATTCGTTCCTTAATTTCATGAACAATCTTCATTCTTTCTTCTCTAGTCATTA
TTATTGGTCCATTCACTATTCTCAT

>Kan promoter

TCCCTTTTCAGATAATTTTAGATTTGCTTTTCTAAATAAGAATATTTGGAGAGCA
CCGTTCTTATTCAGCTATTAATAACTCGTCTTCCTAAGCATCCTTCAATCCTTTT
AATAACAATTATAGCATCTAATCTTCAACAAACTGGCCCGTTTGTTGAACTACT
CTTTAATAAAATAATTTTTCCGTTCCCAATTCCACATTGCAATAATAGAAAATCC
ATCTTCATCGGC

>pNC2

AGATC

>nic

CTTTTCCGCTGCATAACCCTGCTTCGGGGTCATTATAGCGATTTTTTCGGTATA
TCCATCCTTTTTCGCACGATATACAGGATTTTGCCAAAGGGTTCGTGTAGACT
TTCCTTGGTGTATCCAACGGCGTCAGCCGGGCAGGATAGGTGAAGTAGGCCC
ACCCGCGAGCGGGTGTTCCTTCTTCACTGTCCCTTATTCGCACCTGGCGGTG
CTCAACGGGAATCCTGCTCTGCGAGGCTGGCCGG

>oriT.2

CTACCGCCGGCGTAACAGATGAGGGCAAGCGG

>TraJ

ATGGCTGATGAAACCAAGCCAACCAGGAAGGGCAGCCCACCTATCAAGGTGT
ACTGCCTTCCAGACGAACGAAGAGCGATTGAGGAAAAGGCGGCGGCGGCCG
GCATGAGCCTGTCGGCCTACCTGCTGGCCGTCGGCCAGGGCTACAAAATCA
CGGGCGTCGTGGACTATGAGCACGTCCGCGAGCTGGCCCGCATCAATGGCG
ACCTGGGCCGCCTGGGCGGCCTGCTGAAACTCTGGCTCACCGACGACCCGC
GCACGGCGCGGTTCGGTGATGCCACGATCCTCGCCCTGCTGGCGAAGATCG
AAGAGAAGCAGGACGAGCTTGGCAAGGTCATGATGGGCGTGGTCCGCCCGA
GGGCAGAGCCATGA

>oriT.1

CTTTTTTAGCCGCTAAAACGGCCGGGGGGTGCGC

>traI

GTGATTGCCAAGCACGTCCCCATGCGCTCCATCAAGAAGAGCGACTTCGCGG
AGCTGGTGAAGTACATCACCGACGAGCAAGGCAAGACCGCCTAGG

>repBST1

CCTGTTAGACGCCTTAAAACGAGTGTGTCACACCTCAACGCCCTTAATTCGTG
CAACAGCCTGCAATAACGGCGCAAACATGTCTTCGCTGTCAAACGCTGGCTT
GCGCTGGCTTGTCGGCTGGCTGGCGCTTGTGTCCGCCGTGAAGTCGTACCA
AATTAAGCCGGTCGCGGTCGGTTGATAGTCGGCTTGTTTGCGTTGTTGGGCG
CGAGCTTTTTCGGCTTTGACCTTCGCTTTTGCACGTTGCATTAGGTCGTCGAG
TGTTTTGTTTATGTATGCGGATAGATATCGAATGCCGTAAGGATGACGCTTCT
GCACTTCGTTCACTTTAATAGTGACGTATTTTACAATTTTATCCAACGGATGTG
TATTGTCTATTAATTTATTAATATCGTCATGTTTTTCAACTGGTGAACTAATGGT
AGCGTCCGTTTTACGCGCTAAATCGCTTGATTCCGTAGGCTTTTCGGCGTTTT
CTACTTTATCAGCTTCGTTGTCACGCAAGTTTAAAACGGCTGCAATACGCTCG
AAATTTGCATGGCTCCGGAAAATGAATACAGGCGTTTTAGGTCCGTTACTATT
CGGATTCTCACGATAGCAAACGACGACTTCGCCCGAGTCTTTGAGCATGCGA
GTAGCACGATCAACAGTCGGCAAGGACACGCCTAACTTTTTGGCAAGCGTTT
CGCGCCCTGGGAACGAAAAACCGCGGTCTGTTGCGAGCCAAACAATTTCGTC
GATGACGCGCGTTGTTTTTTGCGCTTTGCGGAACTTGCGGCCTAATGCGTCG
GCGATACGCTCCATCAGCTCATTTTTAATCTTACGGCGTTTATAAACGAAATCT
GTGACTCTCTCGATTTCTTTCAGTTGCGTGAATTGTTCGCTAGTGATTATGTAC
ATTTCCGTTCTCCTCCCGTTTCTCCCGAATT

pS797 *P_{ldhA}::CelA*

AAGCTTTCGAGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATA
TGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAA
GGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCG
GCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGA
TGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAAC
AGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAG
CACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGC
AAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTAC
TCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATG
CAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAA
CGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCA
TGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAC
GACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAAC
TATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGG
ATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTG
GCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTAT
CATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTAC
ACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGA
TAGGTGCCTCACTGATTAAGCATTGGTAACTCGAGCTGTCAATAATCTCATGA
CCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAA
AAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTG
CAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGC
TACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAAT
ACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGC
ACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGT
GGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATA
AGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTG
GAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAA
GCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCA
GGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGG
TATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTT
GTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGC
CTTTTTACAGGAAACAGCTATGACCATGGAATTCGCGGCCGCTTCTAGAGCTG
CCTCGTCCATTTTTTTGCTTAATGGAGGTTGTCATGAAAATGACAAACAACGTC
CAAACAATTGCCATAATCGTTTACGCATAGTTTCGATTTCATCGCGTAAAATAA
TTTGTGAATGTATTCACAATAATAAGAAGGGAGAATAGTGATGAAGCGTTATC
GTCGCATCATCGCTATGGTCGTGACGTTCATCTTCATTCTGGGTGTGGTGTAC
GGCGTTAAACCGTGGCAAGAGGTTCGTGCTGGCAGCTTCAATTACGGTGAAG
CGCTCCAGAAAGCGATCATGTTCTACGAGTTTCAGATGAGCGGTAAGCTGCC

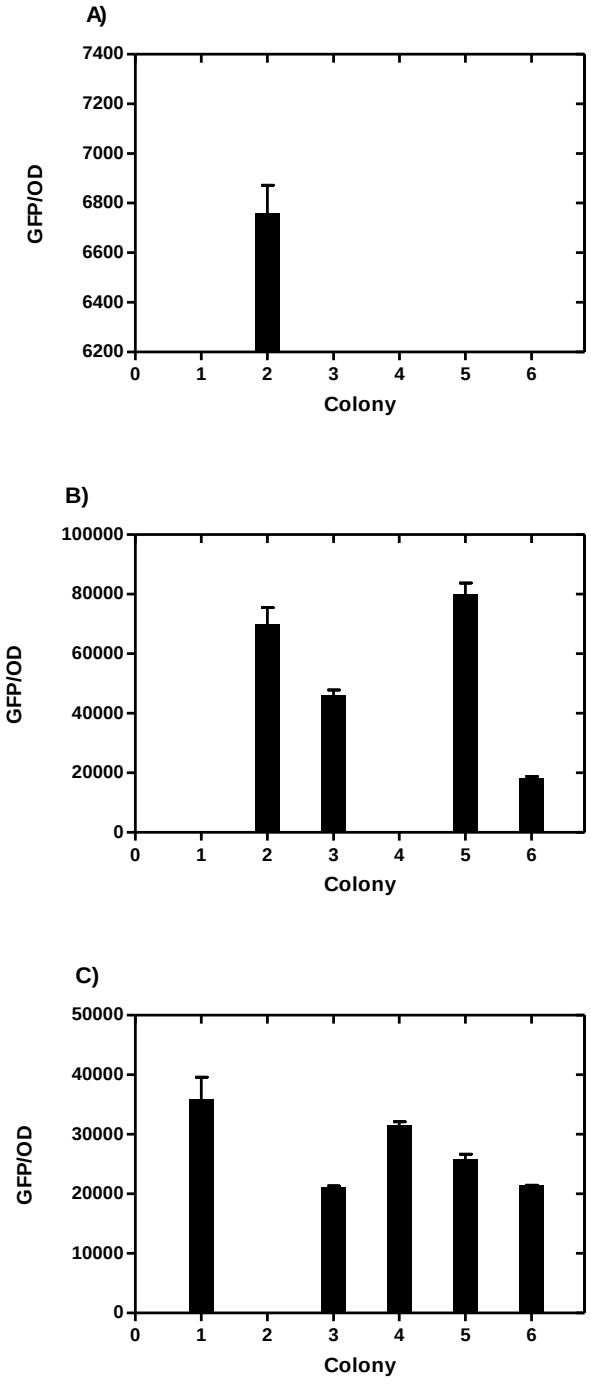
GAACTGGGTCCGCAATAACTGGCGTGGTGACTCTGCGCTGAAAGACGGTCAA
GACAACGGTCTGGACCTGACGGGTGGTTGGTTTGACGCTGGCGATCACGTG
AAGTTCAATCTGCCGATGTCTTATACCGGCACCATGCTGTCTTGGGCGGTATA
CGAGTATAAAGACGCGTTTGTTAAGAGCGGCCAACTGGAACACATTCTGAATC
AAATCGAGTGGGTTAATGACTATTTCGTCAAGTGTCACCCGAGCAAATACGTG
TACTATTATCAGGTCGGTGATGGTAGCAAAGATCATGCGTGGTGGGGTCCGG
CCGAGGTCATGCAGATGGAACGTCCGTCGTTTAAGGTCACGCAGTCTAGCCC
GGGTAGCACGGTTGTGGCGGAGACTGCCGCAAGCCTGGCTGCGGCCTCCAT
TGTTCTGAAAGATCGTAATCCAACCAAAGCCGCGACCTATCTGCAGCACGCG
AAAGAATTGTACGAGTTTGCAGAGGTGACCAAGTCCGACGCAGGCTACACCG
CGGCAAATGGCTATTACAATAGCTGGAGCGGTTTCTATGACGAACTGAGCTG
GGCGGCAGTGTGGCTGTACCTGGCAACGAATGACAGCACCTACCTGACCAAA
GCGGAGAGCTACGTCCAAAATTGGCCGAAGATCTCCGGTAGCAACACCATTG
ATTACAAATGGGCACACTGCTGGGACGATGTCCATAATGGCGCAGCGCTTCT
GTTGGCAAAAATCACCGGTAAAGACATCTACAAACAGATTATTGAGAGCCACC
TGGACTACTGGACCACCGGTTACAATGGTGAACGCATCAAGTATACCCCGAA
AGGCCTGGCGTGGCTGGATCAATGGGGCAGCTTACGTTATGCAACCACGACC
GCTTTTCTGGCCTTCGTCTACAGCGACTGGGTTGGTTGTCCGAGCACCAAGA
AAGAGATTTATCGCAAATTCGGTGAAAGCCAGATCGACTACGCGCTGGGTTC
AGCCGGTCGTAGCTTTGTTGTCGGTTTTGGCACGAACCCGCCGAAACGTCCG
CACCACCGTACTGCGCACAGCAGCTGGGCCGATTCCCAATCTATTCCGAGCT
ATCACCGTCATACCCTGTACGGTGCTCTGGTGGGTGGTCCGGGTAGCGACGA
CAGCTATACCGATGACATTAGCAATTATGTTAACAATGAAGTGGCGTGCGACT
ACAACGCTGGTTTCGTTGGTGCCCTGGCGAAGATGTACCAGCTGTATGGCGG
CAACCCGATTCCGGACTTTAAAGCAATCGAAACCCCGACGAACGATGAGTTCT
TCGTCGAGGCAGGCATCAACGCCAGCGGTACCAACTTTATCGAGATCAAGGC
CATTGTCAATAATCAGAGCGGCTGGCCAGCGCGTGCGACCGATAAACTGAAA
TTTCGTTATTTCGTGGACTTGAGCGAGCTGATTAAAGCTGGCTACAGCCCGAA
TCAACTGACCCTGTCTACCAATTACAATCAGGGCGCAAAAGTTTCTGGCCCGT
ACGTTTGGGACGCCAGCAAGAATATCTACTATATCCTGGTTGACTTTACGGGT
ACCTTGATCTACCCGGGTGGCCAGGATAAGTACAAGAAAGAGGTCCAGTTCC
GTATTGCAGCACCGCAAAACGTTCAGTGGGACAATAGCAACGACTATAGCTTT
CAAGACATCAAAGGCGTGAGCAGCGGTAGCGTCGTTAAGACCAAGTACATTC
CATTGTATGATGGTGATGTCAAGGTGTGGGGTGAGGAACCGGGTACGAGCG
GTGCAACCCCGACCCCGACGGCAACTGCGACCCCTACGCCGACCCCTACCG
TTACTCCGACCCCGACTCCGACGCCGACGTCCACGGCCACGCCAACGCCTA
CCCCAACTCCGACCGTAACGCCGACCCCGACCCCGACTCCGACCGCTACCC
CGACGGCAACGCCGACGCCGACCTCTACCCCGAGCAGCACCCCTGTTGCGG
GTGGTCAGATTAAGGTACTGTACGCAAACAAAGAAACCAACAGCACCACCAA
CACGATTCGCCCGTGGTTGAAAGTCGTTAACACCGGTAGCTCCTCGATTGAT
CTGAGCCGCGTTACCATCCGCTACTGGTACACCGTCGATGGCGATAAGGCAC
AGAGCGCGATCAGCGATTGGGCCCAGATTGGTGCAAGCAATGTCACCTTTAA

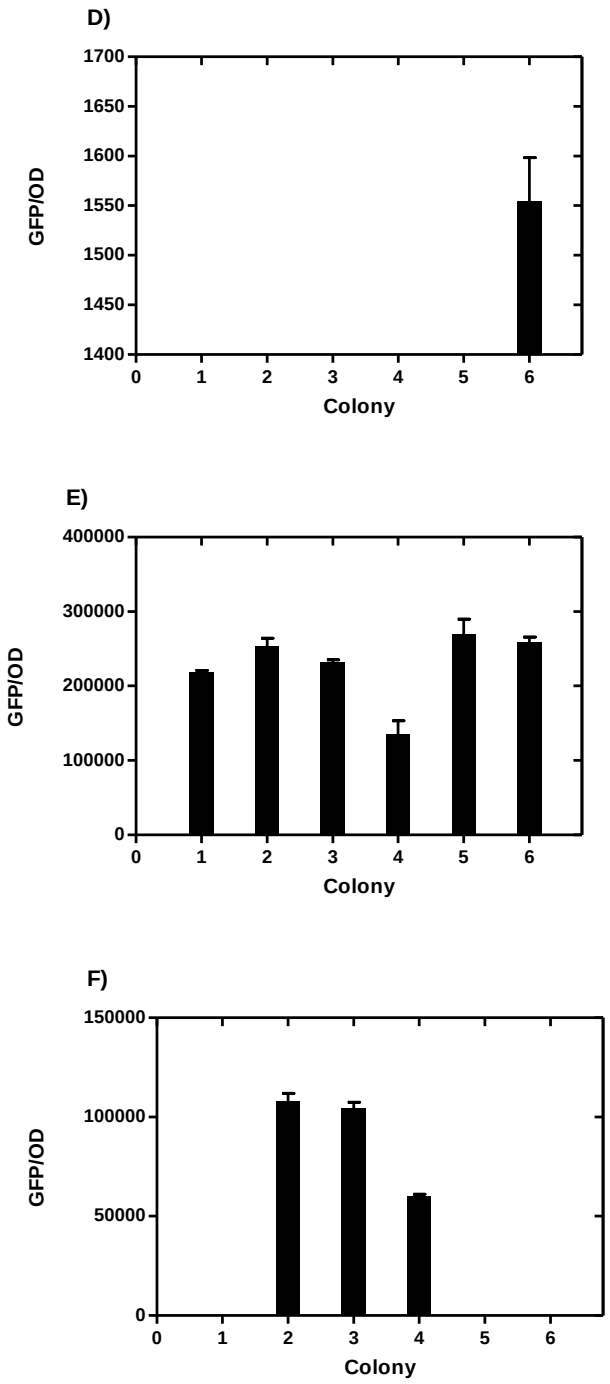
GTTCGTGAAGCTGTCATCCAGCGTTTCTGGCGCGGACTATTATCTGGAAATTG
GTTTCAAATCCGGTGCGGGCCAACTGCAGGCTGGTAAGGATACCGGTGAAAT
CCAGATTCGCTTTAACAAAAGCGATTGGAGCAACTATAACCAAGGTAACGATT
GGTCCTGGATGCAAAGCATGACCAACTACGGCGAGAATGTCAAGGTGACCGC
CTACATTGATGGTGTCTTAGTCTGGGGTCAAGAGCCGTCCGGTGCGACGCCT
ACCCCGACGGCGACGCCGGCACCGACCGTTACCCCTACGCCGACTCCGACT
CCGACGAGCACCCCGACCGCCACGCCTACCGCAACCCCGACCCCGACCCCA
ACGCCGAGCTCCACCCCGGTTGCCGGTGGTCAGATTAAAGTGCTGTACGCGA
ACAAGGAAACCAACAGCACCACCAATACGATTCGTCCGTGGCTGAAAGTTGT
GAACACGGGCAGCAGCAGCATTGACCTGTCTCGTGTGACCATTCGTTATTGG
TATACCGTGGATGGCGATAAAGCGCAGTCTGCAATCAGCGATTGGGCTCAGA
TCGGCGCTTCAAACGTTACTTTTAAGTTCGTTAAACTGAGCAGCAGCGTTTCC
GGTGCAGACTACTATCTGGAAATCGGTTTCAAGAGCGGTGCCGGCCAATTGC
AAGCGGGCAAGGATACGGGCGAGATCCAGATCCGCTTCAATAAGAGCGACT
GGAGCAACTATAATCAGGGCAACGACTGGAGCTGGATGCAATCCATGACGAA
TTACGGCGAGAACGTGAAGGTCACGGCGTATATCGATGGTGTTCTGGTGTGG
GGCCAAGAGCCGAGCGGTGCCACCCCGACCCCAACGGCGACCCCTGCGCC
GACGGTCACCCCGACCCCAACGCCTGCACCAACCCCAACCCCGACGCCGAC
TCCGACTGCGACTCCTACCCCTACTCCGACGCCAACCCCGACGGCCACGCC
GACCGTGACGGCAACCCCGACCCCGACGCCGTCCAGCACGCCGAGCGTTTT
GGGTGAATACGGTCAGCGCTTCATGTGGCTGTGGAATAAGATCCACGACCCG
GCGAATGGTTATTTCAACCAAGATGGCATTCCGTACCACTCGGTTGAAACCCT
GATCTGCGAGGCTCCGGATTACGGCCATCTGACGACCTCCGAAGCATTCAGC
TATTACGTGTGGCTGGAAGCGGTTTACGGCAAGCTGACGGGCGATTGGAGCA
AGTTTAAGACCGCTTGGGACACTCTGGAGAAGTACATGATTCCGAGCGCTGA
AGATCAACCGATGCGTTCCTACGACCCGAACAAGCCGGCTACCTACGCGGGT
GAATGGGAAACGCCGGACAAATATCCGAGCCCGCTGGAGTTCAATGTGCCG
GTTGGCAAAGATCCGCTGCACAATGAACTGGTGAGCACGTATGGCAGCACCC
TGATGTATGGTATGCACTGGCTGATGGATGTCGATAATTGGTATGGTTACGGT
AAGCGCGGTGACGGTGTGAGCCGCGCATCTTTCATTAACACCTTTCAGCGTG
GTCCGGAAGAGAGCGTGTGGGAAACCGTCCCTCACCCGAGCTGGGAAGAGT
TTAAGTGGGGTGGCCCGAACGGTTTCTTGGATCTGTTTATCAAAGATCAGAAC
TATTCGAAGCAATGGCGTTACACCGATGCGCCGGACGCGGACGCTCGTGCC
ATCCAGGCGACCTATTGGGCCAAAGTGTGGGCGAAAGAGCAAGGTAAGTTTA
ATGAAATCAGCAGCTATGTGGCGAAAGCGGCGAAGATGGGCGACTATCTGCG
TTATGCGATGTTTGACAAATACTTTAAGCCATTGGGCTGTCAAGATAAGAACG
CTGCTGGCGGCACCGGCTATGATAGCGCGCACTATCTGCTGAGCTGGTATTA
CGCGTGGGGTGGTGCGCTGGACGGTGCGTGGTCGTGGAAGATTGGTAGCTC
TCATGTGCATTTTGGTTACCAAAACCCGATGGCGGCCTGGGCCCTGGCTAAT
GATAGCGACATGAAGCCGAAGTCGCCGAATGGTGCGTCGGATTGGGCGAAG
TCCTTGAAACGTCAAATTGAGTTCTACAGATGGCTGCAGAGCGCAGAGGGTG
CGATCGCGGGTGGTGCCACCAACTCTTGGAACGGTCGTTATGAGAAATATCC

GGCAGGCACGGCAACCTTTTACGGTATGGCGTACGAGCCGAATCCGGTGTAC
CATGATCCGGGCAGCAACACCTGGTTCGGTTTCCAAGCGTGGAGCATGCAAC
GCGTTGCAGAGTACTACTACGTCACGGGTGACAAAGACGCAGGCGCGCTGC
TGGAGAAATGGGTGAGCTGGGTTAAGAGCGTGGTGAAGCTGAACAGCGACG
GTACGTTCGCAATCCCGAGCACCCTGGACTGGTCGGGCCAGCCGGATACGT
GGAACGGTGCGTATACCGGTAACAGCAATCTGCATGTCAAAGTCGTCGATTAT
GGTACTGACCTGGGTATTACCGCGAGCCTGGCCAATGCACTGTTGTACTACA
GCGCGGGTACCAAAAAGTATGGCGTTTTTGATGAAGGCGCAAAGAACCTGGC
GAAAGAACTGCTGGACCGTATGTGGAAACTGTACCGTGACGAGAAAGGCCTG
TCTGCACCAGAGAAACGTGCGGACTACAAGCGTTTCTTCGAGCAAGAAGTCT
ACATCCCGGCTGGTTGGATTGGCAAGATGCCGAACGGTGATGTGATCAAATC
TGGCGTCAAATTCATCGACATTCGCAGCAAATACAAACAGGACCCTGACTGG
CCGAAACTGGAAGCGGCCTACAAAAGCGGCCAAGCGCCGGAGTTTCGCTAC
CATCGTTTCTGGGCGCAGTGCGATATTGCCATTGCGAATGCGACTTATGAGAT
CTTGTTTGGCAATCAGTAATAAAAATGGAAATGATTCACACCCGATGTGGACT
TGCCTCATCGGGTGTTTTTTTGCTATCAAAACTACTAGTAGCGGCCGCCTGCA
GACTGGCCGTCGTTTTACAAGATCCTCAAAATGGTATGCGTTTTGACACATCC
ACTATATATCCGTGTCGTTCTGTCCACTCCTGAATCCCATTCCAGAAATTCTCT
AGCGATTCCAGAAGTTTCTCAGAGTCGGAAAGTTGACCAGACATTACGAACTG
GCACAGATGGTCATAACCTGAAGGAAGATCTGATTGCTTAACTGCTTCAGTTA
AGACCGAAGCGCTCGTCGTATAACAGATGCGATGATGCAGACCAATCAACAT
GGCACCTGCCATTGCTACCTGTACAGTCAAGGATGGTAGAAATGTTGTCGGT
CCTTGCACACGAATATTACGCCATTTGCCTGCATATTCAAACAGCTCTTCTAC
GATAAGGGCACAAATCGCATCGTGGAACTTTTGGGCTTCTACCGATTTAGCAG
TTTGATACACTTTCTCTAAGTATCCACCTGAATCATAAATCGGCAAAATAGAGA
AAAATTGACCATGTGTAAGCGGCCAATCTGATTCCACCTGAGATGCATAATCT
AGTAGAATCTCTTCGCTATAAAAATTCACTTCCACCTTCCACTCACCGGTTGTC
CATTCATGGCTGAACTCTGCTTCCTCTGTTGACATGACACACATCATCTCAATA
TCCGAATAGGGCCCATCAGTCTGACGACCAAGAGAGCCATAAACACCAATAG
CCTTAACATCATCCCCATATTTATCCAATATTCGTTCCTTAATTTCATGAACAAT
CTTCATTCTTTCTTCTCTAGTCATTATTATTGGTCCATTCACTATTCTCATTCCC
TTTTCAGATAATTTTAGATTTGCTTTTCTAAATAAGAATATTTGGAGAGCACCGT
TCTTATTCAGCTATTAATAACTCGTCTTCCTAAGCATCCTTCAATCCTTTTAATA
ACAATTATAGCATCTAATCTTCAACAAACTGGCCCGTTTGTTGAACTACTCTTT
AATAAAATAATTTTTCCGTTCCCAATTCCACATTGCAATAATAGAAAATCCATCT
TCATCGGCAGATCCTTTTCCGCTGCATAACCCTGCTTCGGGGTCATTATAGCG
ATTTTTTCGGTATATCCATCCTTTTTCGCACGATATACAGGATTTTGCCAAAGG
GTTCGTGTAGACTTTCCTTGGTGTATCCAACGGCGTCAGCCGGGCAGGATAG
GTGAAGTAGGCCCACCCGCGAGCGGGTGTTCCTTCTTCACTGTCCCTTATTC
GCACCTGGCGGTGCTCAACGGGAATCCTGCTCTGCGAGGCTGGCCGGCTAC
CGCCGGCGTAACAGATGAGGGCAAGCGGATGGCTGATGAAACCAAGCCAAC
CAGGAAGGGCAGCCCACCTATCAAGGTGTACTGCCTTCCAGACGAACGAAGA

GCGATTGAGGAAAAGGCGGCGGCGGCCGGCATGAGCCTGTCGGCCTACCTG
CTGGCCGTCGGCCAGGGCTACAAAATCACGGGCGTCGTGGACTATGAGCAC
GTCCGCGAGCTGGCCCGCATCAATGGCGACCTGGGCCGCCTGGGCGGCCT
GCTGAAACTCTGGCTCACCGACGACCCGCGCACGGCGCGGTTCGGTGATGC
CACGATCCTCGCCCTGCTGGCGAAGATCGAAGAGAAGCAGGACGAGCTTGG
CAAGGTCATGATGGGCGTGGTCCGCCCGAGGGCAGAGCCATGACTTTTTTAG
CCGCTAAAACGGCCGGGGGGTGCGCGTGATTGCCAAGCACGTCCCCATGCG
CTCCATCAAGAAGAGCGACTTCGCGGAGCTGGTGAAGTACATCACCGACGAG
CAAGGCAAGACCGCCTAGGCCTGTTAGACGCCTTAAAACGAGTGTGTCACAC
CTCAACGCCCTTAATTCGTGCAACAGCCTGCAATAACGGCGCAAACATGTCTT
CGCTGTCAAACGCTGGCTTGCGCTGGCTTGTCGGCTGGCTGGCGCTTGTGTC
CGCCGTGAAGTCGTACCAAATTAAGCCGGTCGCGGTCGGTTGATAGTCGGCT
TGTTTGCGTTGTTGGGCGCGAGCTTTTTCGGCTTTGACCTTCGCTTTTGCACG
TTGCATTAGGTCGTCGAGTGTTTTGTTTATGTATGCGGATAGATATCGAATGC
CGTAAGGATGACGCTTCTGCACTTCGTTCACTTTAATAGTGACGTATTTTACAA
TTTTATCCAACGGATGTGTATTGTCTATTAATTTATTAATATCGTCATGTTTTTC
AACTGGTGAACTAATGGTAGCGTCCGTTTTACGCGCTAAATCGCTTGATTCCG
TAGGCTTTTCGGCGTTTTCTACTTTATCAGCTTCGTTGTCACGCAAGTTTAAAA
CGGCTGCAATACGCTCGAAATTTGCATGGCTCCGGAAAATGAATACAGGCGT
TTTAGGTCCGTTACTATTCGGATTCTCACGATAGCAAACGACGACTTCGCCCG
AGTCTTTGAGCATGCGAGTAGCACGATCAACAGTCGGCAAGGACACGCCTAA
CTTTTTGGCAAGCGTTTCGCGCCCTGGGAACGAAAAACCGCGGTCTGTTGCG
AGCCAAACAATTTCGTCGATGACGCGCGTTGTTTTTTGCGCTTTGCGGAACTT
GCGGCCTAATGCGTCGGCGATACGCTCCATCAGCTCATTTTTAATCTTACGGC
GTTTATAAACGAAATCTGTGACTCTCTCGATTTCTTTCAGTTGCGTGAATTGTT
CGCTAGTGATTATGTACATTTCCGTTCTCCTCCCGTTTCTCCCGAATT

Promoter Variability





Each graph shows the variability in GFP Fluorescence::OD_{600} resulting from one of the new promoters. The black bars represent mean GFP Fluorescence::OD_{600} from three replicates of three independent transformant cultures, derived from six individual colonies per promoter. Where no bar exists the culture did not grow. Colonies were taken from a single transformation plate of *Geobacillus thermoglucosidans* DSM2542. Cultures were grown in TSBK. This figure shows; A- 00878, B- 00548, C- 00328, D- 00241, E- 03382, F- 00966. Colony 6 has the highest GFP Fluorescence::OD_{600} for promoter 00505. Colony 3 has the highest GFP Fluorescence::OD_{600} for promoter 02364. The error bars represent the standard error of the mean. The instrument's setting were constant.