THE UNIVERSITY *of* EDINBURGH

Design and Construction of Modular Genetic Devices and the Enzymatic Hydrolysis of Lignocellulosic Biomass

Thesis by

Damian K. Barnard

Degree of Doctor of Philosophy

The University of Edinburgh

College of Science and Engineering

School Of Engineering

2012

ABSTRACT

The enzymatic deconstruction of lignocellulosic plant biomass is performed by specialist microbial species. It is a ubiquitous process within nature and central to the global recycling of carbon and energy. Lignocellulose is a complex heteropolymer, highly recalcitrant and resistant to hydrolysis due to the major polysaccharide cellulose existing as a crystalline lattice, intimately associated with a disordered sheath of hemicellulosic polysaccharides and lignin. In this thesis I aim to transfer the highly efficient cellulolytic mechanism of the bacterium *Cellulomonas fimi*, to that of a suitably amenable and genetically tractable expression host, in the hopes of better understanding the enzymatic hydrolysis of lignocellulose. Using tools and concepts from molecular biology and synthetic biology, I constructed a library of standardised genetic parts derived from *C. fimi*, each encoding a known enzymatic activity involved in the hydrolysis of cellulose, mannan or xylan; three of the major polysaccharides present in lignocellulose.

Characterization assays were performed on individual parts to confirm enzymatic activity and compare efficiencies against a range of substrates. Results then informed the rational design and construction of parts into modular devices. The resultant genetic devices were introduced into the expression hosts *Escherichia coli* and *Citrobacter freundii*, and transformed strains were assayed for the ability to utilize various forms of xylan, mannan and cellulose as a sole carbon source. Results identified devices which when expressed by either host showed growth on the respective carbon sources. Notably, devices with improved activity against amorphous cellulose, crystalline cellulose, mannan and xylan were determined. Recombinant cellulase expressing strains of *E. coli* and *C. freundii* were shown capable of both deconstruction and utilization of pure cellulose paper as a sole carbon source. Moreover, this capacity was shown to be entirely unhindered when *C. freundii* strains were cultured in saline media. These findings show promise in developing *C. freundii* for bioprocessing of biomass in sea water, so as to reduce the use of fresh water resources and improve sustainability as well as process economics. Work presented in this thesis contributes towards understanding the complementarities and synergies of the enzymes responsible for lignocellulose hydrolysis. Moreover, the research emphasizes the merits of standardizing genetic parts used within metabolic engineering projects and how adopting such design principles can expedite the research process.

DECLARATION

I hereby declare that the work presented within this thesis is my own, unless otherwise stated. In addition, this work has not been submitted for any other degree or professional qualification.

Damian Barnard

29th August 201

ACKNOWLEDGEMENTS

I would like to thank my supervisors Dr Alistair Elfick and Dr Chris French for imparting many valuable words of wisdom and for sharing their wealth of research experience, without which this thesis would not be possible. Thank you to Chris for being such an inspirational teacher and igniting that initial interest in research science. Thank you to Alistair for the giving me the opportunity to pursue a PhD and enter the exciting field of synthetic biology. I am grateful to both past and present members of the French lab with whom I shared this experience, specifically Sahreena Lakhundi, Nimisha Joshi, Steven Kane, Eugene Fletcher, Joseph White and Chao-Kuo Lui. A big thank you is reserved for Sahreena and Chao-Kuo for welcoming me into the lab and introducing many of the protocols and assays described within this thesis. Thank you to Steven Kane for the permission to use the *bxyF* and *xynF* BioBrick parts within this research. Special acknowledgements are dedicated to Natalie and Nico for their support and much needed company on the couch during the many hours spent writing this thesis. Finally, I thank my parents and brothers for their honest interest in this work and continuous support throughout my studies.

CONTENTS

LIST OF ABBREVIATIONS

Amp	Ampicillin
BMCC	Bacterial microcrystalline cellulose
Btu	British thermal unit
CAZy Database	Carbohydrate-Active Enzymes Database
CBD	Cellulose binding domain
CBM	Carbohydrate binding module
CBP	Consolidated bioprocessing
CD	Catalytic domain
GH	Glycosyl hydrolase
CMC	Carboxymethyl cellulose
Cml	Chloramphenicol
DMSO	Dimethyl sulphoxide
E.C. code	Enzyme commission code
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kan	Kanamycin
LB	Luria Broth
MUC	4-methylumbelliferyl β-D-cellobioside
MUG	4-methylumbelliferyl β-D-glucopyranoside
NCBI	National Center for Biotechnology Information
NCIMB	The National Collection of Industrial, food and Marine Bacteria
ONPC	*O*-Nitrophenyl-β-cellobioside
ONPG	*O*-Nitrophenyl-β-glucopyranoside
ONPM	*O*-Nitrophenyl-β-D-mannopyranoside
ONPX	*O*-Nitrophenyl-β-D-xylopyranoside
PCR	Polymerase chain reaction
PASC	Phosphoric acid swollen cellulose
rbs	Ribosome binding site
SEM	Scanning electron microscopy
TSS	Transformation and storage solution
X-gal	5-bromo-4-chloro-3-indoyl-β-D-galactoside

LIST OF FIGURES

Page

LIST OF TABLES

Page

Chapter 1

Introduction

Summary

In an effort to reduce carbon dioxide emissions, diversify energy resources and improve future energy security, many of the global industrialised economies have shown a resurgence of interest in the development of renewable fuels and chemicals. Drafting of international and national legislations which encourage the shift towards renewable forms of energy are intended to foster the growth of a biorenewables industry; an industry focused on the development of renewable fuels and chemicals from biomass via a engineered microbial hosts. For commercial scale production of biorenewables, significant volumes of product are required to facilitate market penetration and subsequent displacement of fossil fuels. As such, plant biomass is touted as being the most attractive feedstock for the process as it is widely available and abundant. However, due to its complex structure and recalcitrant nature, utilization of plant biomass is far from being economically viable and considerable development of microbial strains capable of utilizing the feedstock is necessary. Within this chapter the fundamentals of the enzymatic hydrolysis of lignocellulosic biomass by cellulolytic bacteria and fungi is reviewed. In particular, the concepts of cellulose hydrolysis are described with reference to the relatively well characterized cellulolytic bacterium *Cellulomonas fimi*. The various classifications of cellulases involved in the hydrolysis of cellulose are discussed with consideration for how the enzymes act in a complementary and synergistic manner to promote efficient deconstruction of plant biomass. Moreover, the mechanisms by which certain cellulases disrupt the substrate in a non-hydrolytic manner are described, and how this disruption facilitates efficient solubilisation and utilization of crystalline cellulose. Literature describing the study of optimized cellulase mixtures is also discussed, highlighting the similarities in expression profiles between cellulolytic species and shared strategies for cellulose hydrolysis. The relevance of these findings to the present study is that the design of a recombinant host capable of cellulose hydrolysis requires the expression of multiple cellulases with defined activities on the substrate. As such, a promising route for solving this problem is framed within the concepts and principles of synthetic biology. The emerging discipline is introduced with an emphasis on the tools and methodologies that can contribute towards the study of ligonocellulose hydrolysis and the development of recombinant hosts for its utilization.

1.1 Renewable chemicals and fuels

An increasing global population, diminishing natural resources, increasing fuel prices and a global economic downturn are key factors adding immense pressure upon industry to adopt leaner practices. These would need to be environmentally friendly, sustainable, cost competitive as well as socially responsible. One step towards solving these issues and contributing towards a sustainable society is the development of chemicals and fuels from renewable feedstocks. Recent advances in metabolic engineering and DNA synthesis technologies are driving the development of microbial hosts for the generation of renewable products. The following section discusses the motivations and implications of this exciting field of study.

1.1.1 The need for renewable chemicals and fuels

Global energy consumption across all sectors is predicted to rise over the coming decades. A report by the US Energy Information administration predicts an increase by 53% over less than 25 years from $5.5x10^{17}$ kJ consumed in 2008 to $7.7x10^{17}$ kJ in 2035, with fossil fuel energy sources accounting for 78% of energy consumed (Smith *et al.* 2011). These forecasts are illustrated in Figure 1.1 below.

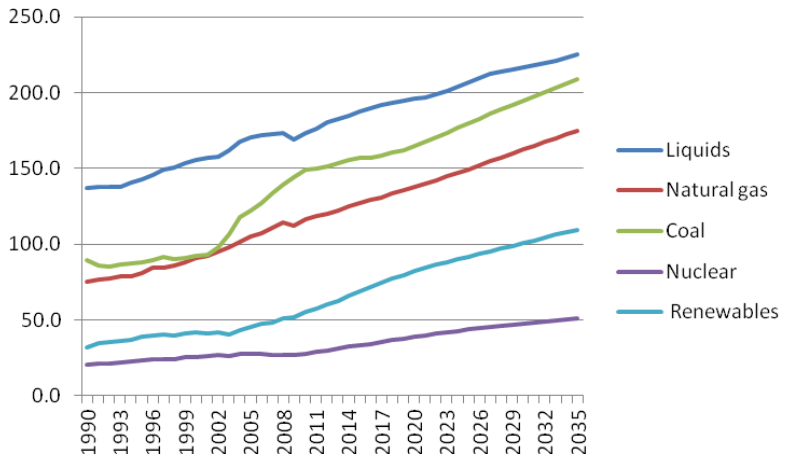

Figure 1.1: World Energy consumption by fuel, 1990-2035 ($x10^{15}$ kilo joules). Data from the EIA International Energy Outlook 2011 (Smith *et al.* 2011).

Crude oil currently supplies about 40% of the world's energy and 96% of its transportation energy (Dudley 2011). Being a key commodity, central to the activities of numerous industries, global oil consumption is projected to rise by about 60% by the year 2020. The

fastest growing oil-consuming sector will be that of transportation. The problem of increasing demand for mineral oil is compounded by the fact that 66% of global oil reserves are owned and operated by a select few countries within the Middle East, namely: Saudi Arabia (25%), Iraq (11%), UAE (9%), Kuwait (9%), Iran (8%) and Libya (2%). It is estimated that by 2020 this figure will increase to around 83% of global oil reserves (Dudley 2011). Consequently there is additional incentive for non-Middle Eastern countries to diversify fuel resources and reduce their dependency on foreign imports as a means of guarding energy security.

Diversifying energy security and moving towards renewables is also crucial in reducing human influenced climate change. A report released by the independent International Energy Agency in 2011 announced that global carbon dioxide emissions from energy use reached a new record of 30.6 billion metric tons in 2010, an increase of 5% from 2008 (Van der Hoeven 2011). In 2009 coal accounted for 43% of CO_2 emissions from fuel combustion, that of oil and gas were 37% and 20% respectively (Figure 1.2a). A larger contribution by coal is thought to be due to the high demand of the rapidly industrializing economies of China and India, whose combined populations account for a third of humanity.

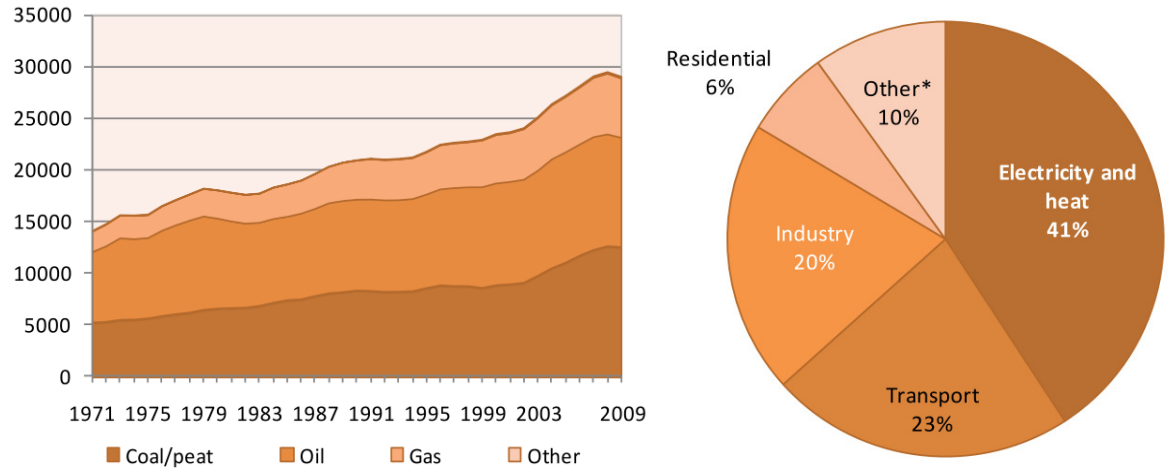

Figure 1.2: World carbon dioxide emissions by fuel, 1971-2009 (Metric tons) **(A).** World carbon dioxide emissions according to sector for the year 2009 **(B)**. Data from the IEA carbon dioxide emissions from fuel report, 2011 (Van der Hoeven 2011).

The same report describes the composition of World CO_2 emissions according to sector in 2009 (Figure 1.2b). The two largest contributors to carbon dioxide emissions are those of electricity and heat, and transport; accounting for 41% and 23% of global carbon dioxide emissions, respectively.

Several international efforts to curb and control societal carbon dioxide emissions have met with limited success, presumably due to the monumental task of coordinating and managing the international politics involved. However prospective legislation and policies at the national level show promise as individual countries and economic zones seek to reduce emissions through adoption of renewable fuels.

1.1.2 Governmental policies and encouragement

The first and most notable effort towards the adoption of renewable fuels is that of Brazil in the mid-1970s. In order to prevent soaring oil prices from impeding economic growth, the *Programa Nacional do Álcool* or the National Alcohol Program was implemented (Marris 2006). Today, Brazil is the world's second largest producer of ethanol at 16.5 billion litres per annum, deriving almost all of the feedstock required for ethanol fermentation from sugarcane. To encourage end-user adoption and market penetration, government subsidised 'flexi-fuel' cars were introduced in 2003. Brazilian motorists were able to choose between traditional petroleum, neat bioethanol or various combinations of blended bioethanol such as E85, E15 or E10 containing 85%, 15% or 10% ethanol, respectively (Marris 2006). More recently the Brazilian government announced a program for the issue of $38 billion in subsidized credit to the ethanol sector. The program is expected to encourage growth of the ethanol industry and increase exports within the international market (Lane 2012).

Seeking to emulate the success of the Brazilian National Alcohol program and achieve energy independence, The United States introduced the Energy Independence and Security Act of 2007 (Sissine 2007). The act is intended to improve vehicle fuel economy and help reduce the dependence on oil through increased production of biofuels. Accordingly, the total amount of biofuels to be used in transport fuels was required to increase from 18 billion litres in 2007 to 136 billion litres by 2022. Further specifications included that non-cornstarch (e.g. sugar or cellulose) derived biofuels should contribute 79.5 billion litres of the 2022 total (Sissine 2007). More recently and in addition to the Energy Independence Act, the Obama administration announced plans to develop a National Bioeconomy blueprint with the

intention of harnessing biological research to address several national challenges including energy. The results of this development and the implications for biorenewables are still to be seen.

Within the European Union, the Renewables Directive of 2009 was designed with similar intentions (Turmes 2009). The directive mandates levels of renewable energy within the EU, requiring member countries to produce pre-agreed proportions of energy consumption from renewable sources. The end goal of the directive is that at least 20% of total energy produced within the EU is from renewable sources by 2020 (Turmes 2009). With specific regards to biofuels, fuels generated from waste streams are counted twice, greatly encouraging the adoption and development of a biorenewables industry.

With regards to China, ambitious targets have been set to cut carbon emissions (Qiu 2009). China's State Council announced that it will reduce its carbon intensity (carbon emissions per unit of gross domestic product) by 40-45% from 2005 levels by 2020. This target is considered achievable as previously the country reduced carbon intensity by 47% between 1990 and 2005. By comparison the US has pledged to cut emissions by 17% from 2005 levels by 2020, Korea by 36% and Brazil by 30%, with the most ambitious pledge being that of the EU cutting 20% of 1990 levels by 2020 (Qiu 2009).

Evidently many economically influential countries have shown consideration and support in fostering the shift towards renewable forms of energy. Yet, in order to achieve many of the goals and aspirations set by governmental bodies, consistent and measured support of the renewables industry is required. Moreover in order for renewable fuels to make an appreciable impact upon displacing traditional petroleum products, time is required for the industry to grow and the technology to mature before a premature rush to market. Nevertheless there are many key start-up entities utilizing innovative biological research to quickly bring renewable fuels and chemicals to market.

1.1.3 A burgeoning biorenewables industry

In order to facilitate the market penetration of renewable fuels into the transport industry and realise a subsequent reduction in environmental impact, alternative candidates would need to possess several characteristics: (i) compatibility with existing distribution and processing infrastructures; (ii) compatibility with current vehicle engine technology; (iii) comparable energy content and performance to that of petroleum to encourage adoption by the end-user.

Advanced biorenewable companies are focused on the generation of fuels better suited to displacing current petroleum fossil fuels and which possess the above characteristics. In 2011 alone advanced fermentation technologies developed by companies around the world raised $687 million in private funding and equity financing (Lane 2012). Key players within the burgeoning biorenewables industry seek to develop modern genetic engineering and synthetic biology as a route for the generation of these novel substitutes. Industry leaders such as Amyris, LS9 and Solazyme are successfully engineering microbial strains to convert simple sugars to long chain alcohols, fatty-acid ethyl esters or isoprenoid hydrocarbons.

As an example, Amyris's farnesene is fermented via the mevalonate pathway for producing isoprenoids. Developed initially for the production of the antimalarial precursor artimisinic acid, the platform technology has been repurposed and modified for hydrocarbon synthesis (Martin *et al.* 2003). Engineered yeast producing farnesene from a sugar cane feedstock has been reported in the 100,000 and 200,000 litre scale capacity fermentors.

By comparison, Solazyme uses directed evolution of selected microalgae strains to ferment sugars into triglycerides and then employs a hydrotreatment technology reduces the triglycerides to fully saturated alkanes (Westfall and Gardner 2011).

LS9 has focused on direct fermentation of alkenes and alkanes via a modified fatty acid synthesis pathway in *E. coli*, thereby bypassing triglyceride formation. The technology developed is based upon multiple approaches including; the formation of alkenes via reduction or decarbonylation of fatty acyl-ACP (Schirmer *et al.* 2010), and the production of fatty alcohols via a thioesterase/fatty acid reductase pathway (Steen *et al.* 2010).

It can be noted that presently very few of these companies are developing both product formation as well as feedstock utilisation pathways within their prospective production hosts. It would seem the consensus between these start-ups is that presently a sugar based feedstock such as sugar cane is the most attractive feedstock for developing the technologies and scaling up production. Yet, in the long-term, lignocellulosic biomass is the most economical and widely available feedstock for bioprocessing and mature cellulolytic technologies available for licensing will need to be developed in the near future.

1.1.3 Generation of biorenewables from non-food crop sources

The Billion-Ton study conducted by the Oak Ridge National Laboratory and sponsored by the US Department of Energy addresses several concerns regarding the feasibility of developing renewable fuels and chemicals from plant biomass (Perlack and Stokes 2011). Crucially the study omits the inclusion of algal sources of renewable fuels, relying solely on those sourced from a lignocellulosic feedstock, yet offers positive predictions for the future of biorenewables suggesting that the 2022 target of 80 billion litres of cellulosic ethanol set by the Energy Independence Act of 2007 is attainable.

It is of paramount concern in the biorenewables industry to move away from food-crops as feedstock for the production of chemicals and fuels. The current wide-scale adoption of starch-based food crops as feedstock for the fermentation of fuel ethanol has yielded unfavourable results in many respects. According to the Food and Agriculture Organization of the United Nations Food Price Index, global food prices have reached an all time high (Diouf and Sheeran 2010). Several key factors were highlighted, namely the increase in global population, more frequent extremes of weather, and most importantly the practise of diverting food crops towards fuel production.

Examples of food-crop biofuels include those derived from corn and cassava. Traditionally used as a source of modified starch and animal feed, the root vegetable cassava is seeing new demand for biofuel production, with its price doubling between 2008 and 2011. In 2010, 98% of cassava exported by Thailand, the world's largest exporter of cassava, went to China for use in biofuel production (Diouf and Sheeran 2010). In a similar fashion, the US is the world's largest producer of fuel ethanol and for the first time in 2011, corn saw greater demand for fuel production than for use in animal feed and human consumption (Norton *et al.* 2012).

Advances in the use of non-food crops and agricultural waste have recently yielded very promising results. Work by Higashide *et al.* (2011) developed the cellulolytic bacterium *Clostridium cellulolyticum* to produce isobutanol directly from cellulose. The approach was based upon diverting 2-keto acid intermediates from amino acid biosynthesis towards alcohol biosynthesis with yields of up to 660 mg/L from untreated crystalline cellulose. Following an alternative strategy Bokinsky *et al.* (2011) engineered the genetically tractable host *Escherichia coli* to break down both xylan and cellulose, as well as synthesise one of three

advanced biofuels within the same fermentation reaction. Notably the feedstock required pre-treatment with ionic liquids and subsequent washing before use, presumably due to the host only being able to utilize pre-treated amorphous cellulose and not raw crystalline cellulose. The use of marine seaweed as a feedstock for renewable fuels has also recently met with some success. Wargacki *et al.* (2012) identified a 36 kilo-base pair fragment of DNA from *Vibrio splendidus* encoding the necessary enzymes for alginate transport and metabolism. Coupling these pathways with the ethanologenic components of *Zymomonas mobilis*, the two systems were recombined on the genome of *E. coli* to yield a strain capable of simultaneous alginate catabolism and ethanol synthesis of up to 20 g/L.

1.2 Biochemistry and structure of lignocellulosic biomass

Considering the dire need for a renewable petrochemicals industry and the early efforts underway to achieve this, the major challenge of efficiently utilising a sustainable and renewable feedstock such as plant biomass is still yet to be overcome. To ensure displacement of traditional petroleum products with renewable alternatives, plant biomass is touted as being the most cost-competitive and abundant feedstock. Whether derived from agricultural waste streams or purpose grown energy crops on marginal arable land, plant biomass offers itself as a sustainable and economically viable feedstock. However there exist several biological barriers to its breakdown and utilisation in an industrial process. Here we consider the major research hurdles and challenges to lignocellulose degradation.

1.2.1 Cellulose

Lignocellulose is the structural component of plant cell walls, offering remarkable tensile strength and rigidity. Plant cell walls are composed of a layered mesh of microfibrils, consisting of long cellulose fibres embedded in an amorphous matrix of hemicellulose and lignin (Fig. 1.3). These cellulose fibres are long chains of the order of thousands of residues with the repeating unit being cellobiose (glucose-β-1,4-glucopyranoside). Chains of cellulose are tightly packed together as insoluble hydrogen-bonded crystalline regions alternating with amorphous regions. Lignocellulose in this form typically consists of cellulose (35–50 wt. %), hemicellulose (20–35 wt. %), and lignin (5–30 wt. %) (Zhang and Lynd 2004).

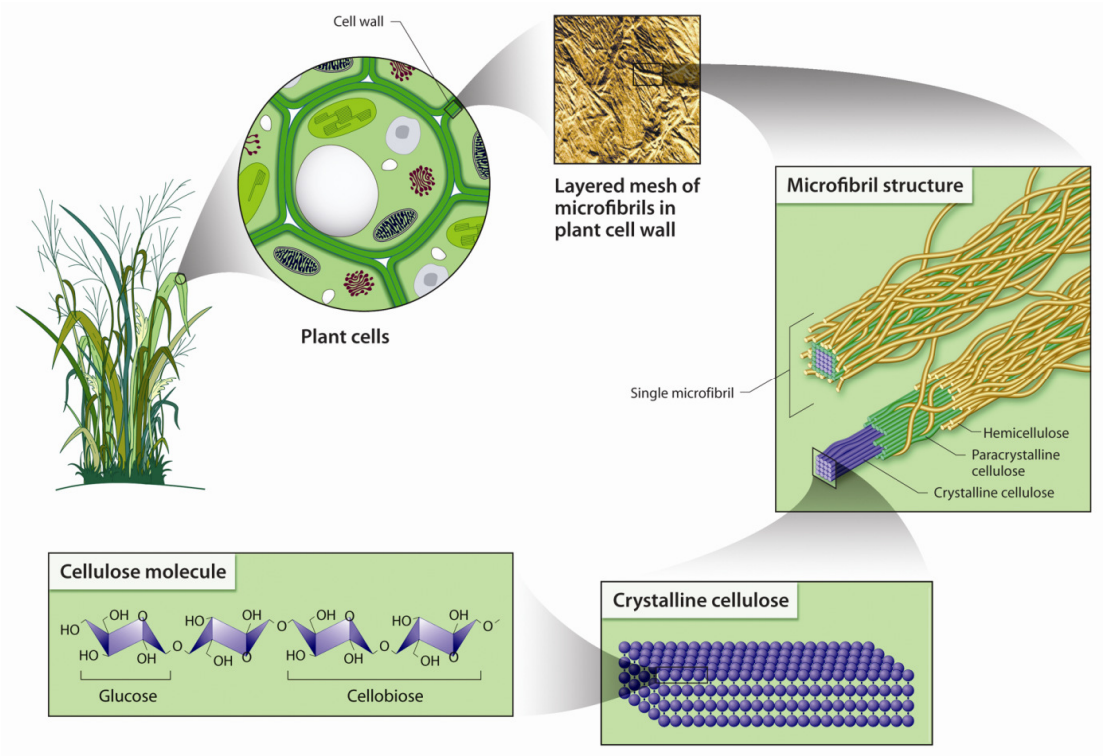

Figure 1.3: The structure of lignocellulosic plant biomass detailing microfibril structure and cellulose microstructure. Cellulose chains composed of a repeating cellobiose subunit are associated in a three dimensional crystalline-like lattice, a structure which endows significant tensile strength and rigidity to the plant cell wall. The lattice is intimately associated with hemicellulose polymers composed of a mix of five and six carbon sugars, in addition phenolic lignin acts as a supporting skeleton. The resultant macro-structure is that of a layered mesh of microfibrils forming the plant cell wall. Source: (Patrinos and Staffin 2005). Reproduced with permission of the Office of Biological and Environmental Research of the United States Department of Energy Office of Science.

In more detail, cellulose is a linear condensation polymer consisting solely of glucose (D-anhydroglucopyranose) monomers bonded by a β-1,4-glycosidic linkage (Fig. 1.4a). The length of these chains is unknown, but single glucan chains containing up to 14,000 glucose units have been observed, corresponding to a length of about 7 μm (Somerville *et al.* 2004). As a result of the β-1,4-linkage, adjacent anhydroglucose molecules are rotated 180^{o} with respect to their neighbours, resulting in anhydrocellobiose being the repeating unit of cellulose (Fig. 1.4b). Such a rotation also causes cellulose to be highly symmetrical as each side of the polymer has an equal number of hydroxyl groups (Fig. 1.4c).

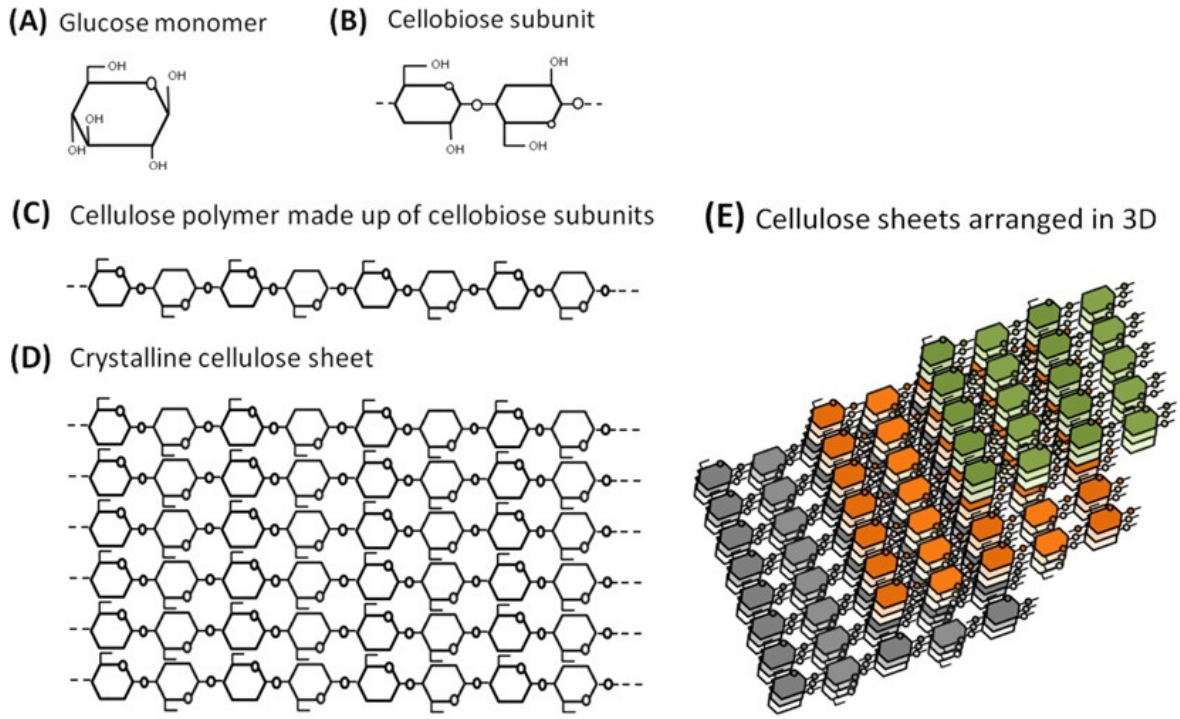

Figure 1.4: Cellulose microstructure formed by cellulose sheets composed of cellulose polymers. Crystalline and amorphous cellulose is made exclusively of a D-glucose monomer **(A)**; two glucose monomers undergo a condensation reaction between the β1 and 4 carbons forming the cellulose disaccharide subunit anhydrous cellobiose **(B)**; repeating cellobiose subunits polymerise and form a the highly regular cellulose polymer **(C)**; multiple polymers arrange in parallel and associate via strong intermolecular forces to form a regular crystalline cellulose sheet **(D)**; multiple cellulose sheets associate via hydrophobic interactions to form a crystalline like lattice **(E).**

Adjacent cellulose polymers are coupled together by intermolecular hydrogen bonds and Van der Waal's forces resulting in a regular parallel alignment and crystalline sheet structure (Fig. 1.4d). Multiple cellulose sheets associate by hydrophobic interactions along the planar face and arrange together to form a three-dimensional lattice like structure. This lattice is illustrated in Figure 1.4e as nine sheets stacked upon one another in sets of three. Consequently, the single repeating β-1,4-glycosidic bond throughout cellulose creates a deceptively simple polymer, yet the resultant crystalline form resists hydrolysis and free diffusion of water, making it intrinsically resistant to dissolution. It is this recalcitrant nature which is one of the major hurdles to the deconstruction and utilization of plant biomass.

1.2.2 Hemicellulose

Hemicelluloses are polysaccharides found in plant cell walls, intimately associated with cellulose and lignin. As these are not chemically well defined polymers, the hemicelluloses are considered a family of polysaccharides consisting of a mixture of 5- and 6-carbon sugar monomers predominantly comprising D-xylose, L-arabinose, D-mannose and D-galactose. The hemicellulose family includes xyloglucans, xylans, mannans and glucomannans, as well as mixed linkage β-glucans, all of which have β-1,4-linked backbones of glucose, mannose or xylose (Scheller and Ulvskov 2010). The composition and detailed structure of these components will vary widely according to plant species and cell type. Figure 1.5 summarises the major monosaccharides and polysaccharides of the hemicelluloses, and gives representative examples of exhibiting plant species.

Being the second most abundant component of dicotyledonous plant cell walls after cellulose, xyloglucans are a crucial target for the deconstruction and utilization of plant biomass. Xyloglucans can exist with varying degrees of branching and substitution along the backbone, with solubility decreasing with degree of branching, an aspect which correlates to function and cell type (Fig. 1.5a). For example, highly branched xyloglucans predominate in elongating cell walls, and those cells within the internodes connecting stems and branches of dicots. This distribution is widely believed to impart plasticity to the cell wall as the cross-linking of cellulose with xyloglucans helps reduce rigidity (Gilbert *et al.* 2008).

By comparison, grasses and conifers generally exhibit a relatively lower proportion of xyloglucans and instead display either glucuronoarabinoxylan or galactoglucomannan as the dominant hemicelluloses, respectively (Scheller and Ulvskov 2010). Soft-woods and hard-woods exhibit up to 25% or 3-5% of dry weight as acetylated-galactoglucomannan, respctively.

In many respects it is of importance to note these differences in hemicellulose content between the various plant taxonomic groups, especially when considering what feedstock will be used for the synthesis of renewable chemicals and fuels. As a raw material, the predefined plant biomass will exhibit predictable hemicellulose content and consequently a tailor made microbial host may be designed which specifically targets those polysaccharides present.

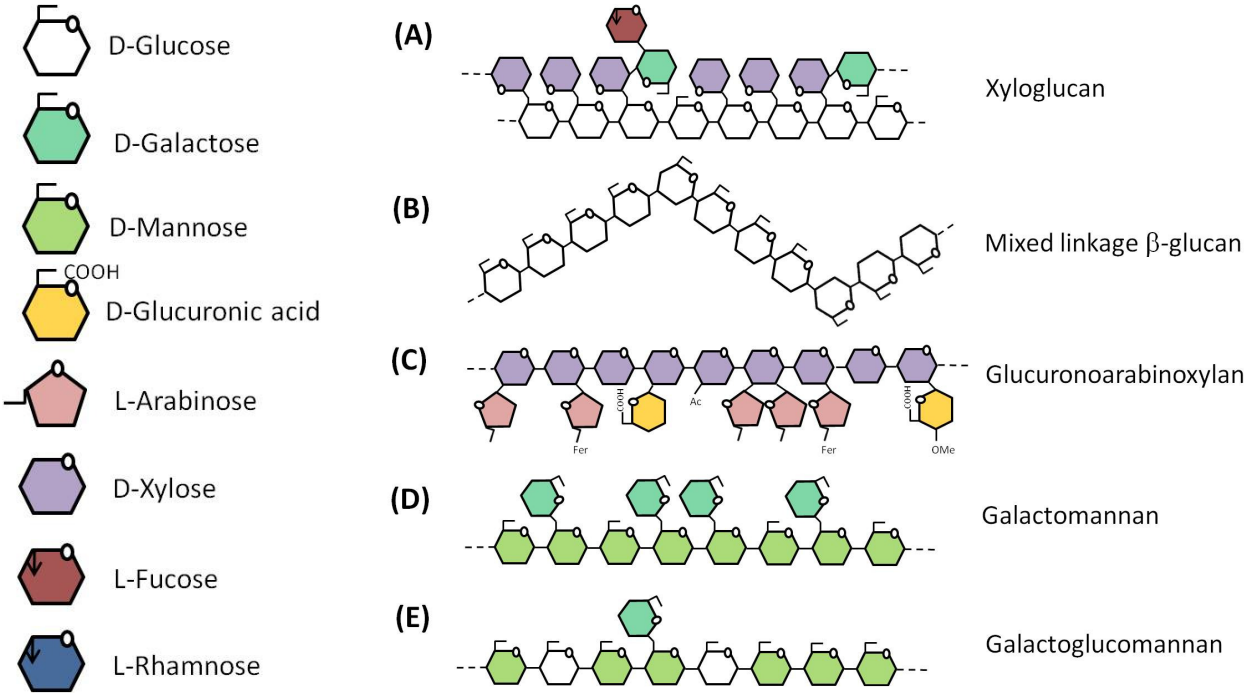

Figure 1.5: The major hemicellulose oligosaccharides found within plant cell walls are composed of distinct sugar units (shown on the left above). A xyloglucan backbone substituted with fucose and galactose is common within the model plant *Arabidopsis thaliana* and related woody shrubs **(A)**. The mixed linkage β-glucan oligosaccharide is restricted to specific grasses and a few other groups **(B)**. Palm and banana trees typically present glucuronoarabinoxylan which consists of a D-xylose backbone substituted with L-arabinose and D-glucuronic acid **(C).** Coniferous woods typically present galactomannan and galactoglucomannan as major hemicelluloses polymers **(D, E)**.

1.2.3 Lignin

Lignin is a complex phenolic heteropolymer resulting from the polymerization of various cinnamyl alcohols termed monolignols. These all share a phenylpropane structure that is a benzene ring exhibiting a tail of three carbons (Boudeti, Lapierre, and Pettenati 1995). Figure 1.6a illustrates the three major lignin monolignol monomers; paracoumaryl, coniferyl and sinapyl alcohols. These are incorporated into lignin in the form of the phenylpropanoids *p*-hydroxyphenyl (H), guaiacyl (G), and syringal (S), respectively. Though extraordinarily complex and irregular, polymerization patterns do occur within the lignin macromolecule and have been characterized to an extent, revealing the lignin GS and lignin G subunits, illustrated respectively in Figure 1.6b and c.

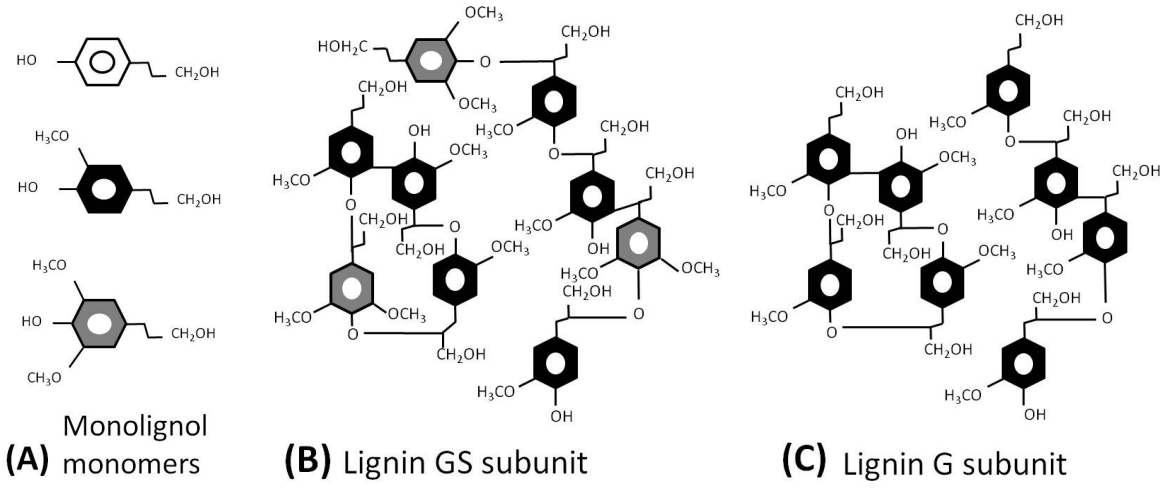

Figure 1.6: Lignin monomers and subunits. Lignin within terrestrial plant biomass is predominantly made up of three phenolic alcohol monomers; paracoumaryl (top), coniferyl (middle) and sinapyl (bottom) **(A)**. The two major subunits of lignin are the lignin GS subunit composed of both coniferyl and sinapyl monomers **(B),** and the lignin G subunit which in turn is made up of exclusively the coniferyl monomer **(C).**

The process of lignification occurs late in the development of plant cell walls, being synthesized after cellulose and hemicellulose. Consequently lignin is primarily located on the exterior of the microfibrils where it covalently bonds to the hemicelluloses (Somerville *et al.* 2004). Serving a role to impermeabilise the cell wall, lignins are concentrated around the xylem elements of vascular plants and also function as a latent defence mechanism against pathogens. With respect to cellulolysis, lignin is thought to be a competitive cellulase adsorbent reducing the effective enzyme load in contact with cellulose and lowering hydrolytic efficiency. In addition it has been suggested that lignin cross-linking has a role in physically hindering the progress of cellulases along the cellulose glucan chain and impeding enzyme accessibility to the cellulose chain (Zhang and Lynd 2004).

1.3 Characteristics of microbial cellulases

Due to the insoluble nature of the crystalline regions as well as the presence of a cross-linking hemicellulose and lignin matrix surrounding the fibre, both enzymatic and non-enzymatic hydrolysis of cellulose are made difficult. The ability to completely degrade crystalline cellulose is uncommon and restricted to specialized cellulose-degrading microorganisms displaying a battery of cellulolytic and hemicellulolutic enzymes. Classified as *O*-glycosyl hydrolases (Enzyme Commission number 3.2.1.-) these enzymes target the glycosidic bond between carbohydrates or between carbohydrates and non-carbohydrate moieties (Henrissat,

Teeri, and Warren 1998). The classifications of this group are summarised within the Carbohydrate-Active Enzymes Database or the CAZy database (Cantarel *et al.* 2009). Considerable research effort over the past 60 years has yielded a comprehensive albeit incomplete understanding of lignocellulose deconstruction by the glycosyl hydrolases. Gaps in knowledge still persist with regards to their regulation and complementary actions on the complex substrate. In this section I consider the underlying mechanisms of the enzymatic hydrolysis of lignocellulose and the promising approach of studying synergism between and within the various cellulases.

1.3.1 Microbial communities utilizing cellulose

Terrestrial plants produce approximately $1.3x10^{10}$ metric tonnes (dry weight) of biomass per year. This represents the energetic equivalent of $7x10^{9}$ metric tonnes of coal or about two-thirds of humanity's annual energy requirement (Perlack and Stokes 2011). It is exciting to consider that the recycling of this material at such a scale is performed by communities of cellulolytic microbes; whether present within the digestive tract of ruminants and insects, or within the soils of various forest and savannah habitats (Kumar *et al.* 2008). Unsurprisingly, numerous efforts have focused upon reproducing this scale and efficiency of lignocellulose utilization within a laboratory setting with the aim of developing an industrialized process.

Most notably, a stable microbial community capable of efficient raw-biomass utilization was constructed and maintained (Haruta *et al.* 2002). This work presented a mixed community of both aerobic and anaerobic species capable of degrading more than 60% of untreated rice straw within 4 days at 50^{o}C. The community was shown to consist of at least six individual species of bacteria and archaea, comprising *Clostridium* sp., *Pseudoxanthomonas* sp., *Brevibacillus* sp. and *Bordetella* sp. From these observations a second community of five individual bacterial strains was constructed and shown to retain the same cellulolytic ability and population stability (Kato *et al.* 2005). Through systemically constructing 'knockout communities' where individual strains were removed from the original community, specific roles and dynamics could be deduced. It was noted that the most efficient and stable communities were those which balanced cellulose hydrolysis and the removal of end-products such as excess acetic acid.

The above research highlights the potential in designing microbial communities for targeted biomass deconstruction as well as emphasizing the need to understand cellulose hydrolysis at

a multispecies level, which is currently lacking. To this effect, the remainder of this section will introduce the enzymatic components of cellulose hydrolysis with an emphasis on introducing these to a non-cellulolytic host.

1.3.2 Non-complexed cellulase systems

Celluloytic microorganisms typically employ one of two systems for crystalline cellulose hydrolysis; complexed or non-complexed. Microbes presenting complexed cellulase systems are commonly found in anaerobic environments, exhibiting a stable enzyme complex known as a cellulosome, protruding from the outer cell membrane. The cellulosome consists of a scaffold like architecture on to which the various cellulase components are tethered and kept in close proximity. This strategy is thought to offer a concentrated cellulolytic action close to the cell, reducing the distance by which hydrolysed products must diffuse and allowing for efficient uptake by the host (Lynd *et al.* 2002).

In contrast to this, non-complexed systems are those in which the cellulases are not cell associated, but are secreted into the medium and free to diffuse away from the host. Such a strategy is common amongst aerobic microbes where relatively higher ATP levels are available for cellulase synthesis compared to anaerobic microbes. Figure 1.7 illustrates the three constituent enzymatic activities involved in a typical non-complexed system, namely; (i) endoglucanases (EC 3.2.1.4); (ii) exoglucanases, including cellodextrinases (EC 3.2.1.74) and cellobiohydrolases (EC 3.2.1.91); and (iii) β-glucosidases (EC 3.2.1.21).

Enzymatic attack on crystalline cellulose is initiated by the action of endoglucanases predominantly targeting amorphous regions, where inconsistencies in intermolecular forces disrupt the cellulose macrostructure and expose vulnerable glycosidic bonds (Fig. 1.7a). Once an initial nick is made at these sites within the glucan chain, subsequent hydration and partial solubilisation of that chain accommodates the binding of exoglucanases (Fig. 1.7b). These processively move along the cellulose polymer from either the reducing or non-reducing end cleaving off cellobiose subunits. In turn underlying cellulose chains are exposed to further hydrolysis due to this ‘shearing action’ of the exoglucanases (Fig 1.7c). As cellobiose subunits are liberated from cellulose, β-glucosidases relieve end-product inhibition by hydrolysing the disaccharide into two glucose monomers (Fig. 1.7d). The concerted and synergistic action of the cellulases leads to extensive hydrolysis of the cellulose substrate.

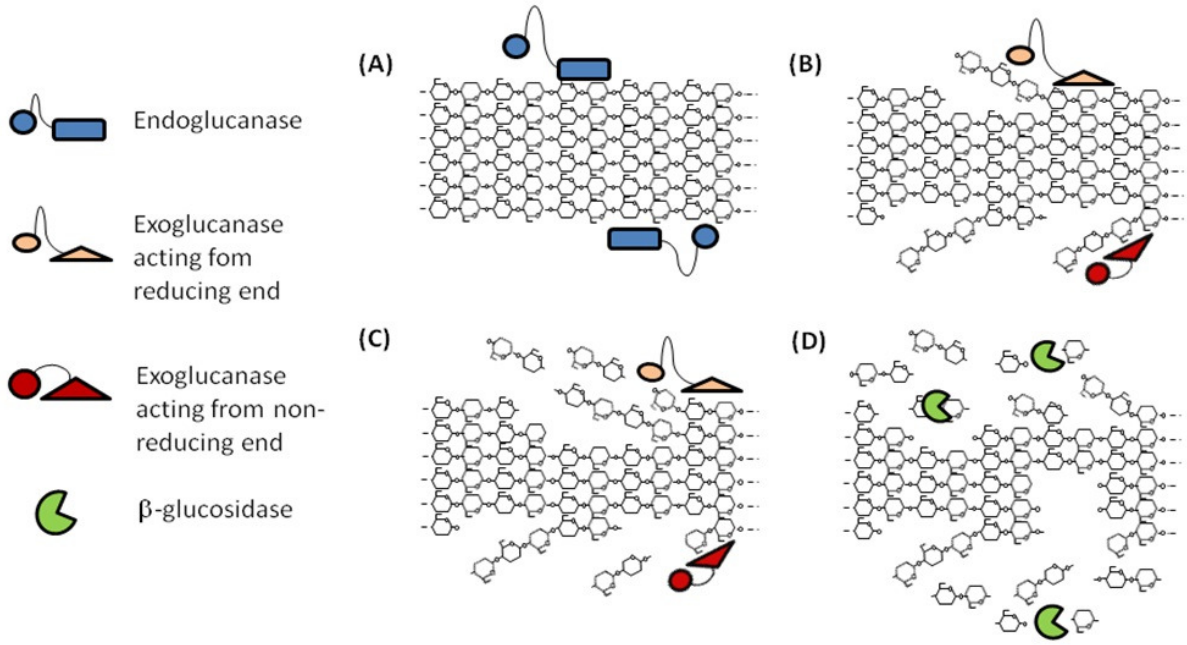

Figure 1.7: Schematic representation of the hydrolysis of crystalline cellulose by a non-complexed cellulase system. Endoglucanases initiate the enzymatic attack on cellulose, commonly targeting amorphous cellulose regions where irregularities in the structure offer access to exposed glycosidic bonds **(A)**; free glucan chains are hydrated and made accessible to the exoglucanases which processively move along the polymer cleaving off cellobiose subunits **(B)**; in turn, underlying glucan chains are exposed and made accessible to the attack of both endo- and exoglucanases **(C)**; liberated cellobiose subunits are cleaved to glucose monomers by β-glucosidases **(D)**.

It should be noted that the cellulases in non-complexed systems tend to be modular proteins in that one or more discreet carbohydrate binding modules (CBMs) are connected to a catalytic domain (CD) via a flexible linker. The roles of these individual modules are discussed further below, with a consideration for how this arrangement of protein structure affects function.

1.3.3 Role of the flexible linker region

The flexible linker region is thought to provide spatial separation between the catalytic domain and cellulose binding domain (CBD) so as to reduce steric hindrance, and allow for the independent action of each. Moreover it has been suggested that the linker acts as a buffer between the two domains, preventing a conformational change in the CBD upon adsorbing to cellulose from being translated throughout the entire enzyme (Henrissat 1994). It would seem that the role of the linker varies across the cellulases, as deletion of the linker has been shown to reduce both catalytic and CBD activities in some cases, where as in other cases no effect has been observed (Poon, Withers, and McIntosh 2007).

At the nucleotide level, sequences encoding the linker may also serve a function in domain shuffling, facilitating the rearrangement of CBDs and catalytic domains (J. Gilbert and Hazlewood 1993). In particular, metagenomic analysis of the rumen of herbivores has yielded insight into the horizontal gene transfer between anaerobic fungi and bacteria, and how a close proximity encourages genome plasticity within this niche cellulolytic environment (Sommer, Church, and Dantas 2010).

1.3.4 Cellulose binding domains

A CBM is defined as a contiguous amino acid sequence within a carbohydrate-active enzyme with a discreet fold having carbohydrate-binding activity (Cantarel *et al.* 2009). CBMs were previously classified as cellulose-binding domains (CBDs) based on those modules that bound cellulose; however the classification was extended to include those that bound other carbohydrate substrates. CBDs of cellulases within non-complexed systems have been shown to play several key roles in the deconstruction of cellulose; (i) enzyme loading onto the substrate, (ii) targeted recruitment of cellulases to sites of hydrolysis, (iii) anchoring of host cells to the substrate surface, and (iv) non-hydrolytic disruption of cellulose. These roles are discussed further.

1.3.4.1 Enzyme loading

To increase the effective concentration of the enzyme on the cellulose substrate, CBDs bind to cellulose via hydrophobic interactions between aromatic amino acid residues and the cellulose surface. Through this association the catalytic domain is brought into close contact with the substrate over a prolonged period, increasing the activity of the catalytic domain. Once bound, the adsorption and desorption of the CBD is highly dependent upon the presence of water, as the CBD will only migrate between neighbouring substrate surfaces when hydrated (Cavaco–Paulo *et al.* 1999). This mechanism is thought to prioritize and reinforce the adsorption of the CBD to non-hydrated crystalline regions whilst also enabling the CBD to dissociate from those surfaces already hydrated or compromised, allowing for the arrival of other cellulases.

The CBD is also shown to initiate mobility of the cellulase along the substrate (Jervis, Haynes, and Kilburn 1997). Using flurorescence after photobleaching analysis, Jervis *et al.* found that the CBDs of the *Cellulomonas fimi* cellulases CenA and Cex were mobile on the cellulose surface. Notably of those CBDs bound to cellulose, more than 70% were mobile,

demonstrating that packing of cellulases and their isolated CBD's is a dynamic process. This surface diffusion of the CBD and attached catalytic domain is proposed to allow the enzyme to search for accessible glycosidic linkages, improving both productivity and processivity.

1.3.4.2 Recruitment of cellulases to sites of hydrolysis

Examples of both bacterial and fungal cellulases have provided evidence for CBDs mediating the recruitment of cellulases to targeted sites of hydrolysis (Koivula *et al.* 2000). Although the crystalline structure of cellulose presents a uniform substrate seemingly devoid of distinguishing molecular landmarks, its progressive hydrolysis unmasks a varied landscape, presenting the repeated β-1,4-glycosidic bond in various contexts and providing distinguishable sites for binding. Consequently, hydrolysis of the cellulose substrate may be limited according to the availability of those sites for binding by the CBDs. Therefore a range of cellulases, specifically CBDs are required for complete hydrolysis. Such a concept has been suggested to be the reason why cellulolytic bacteria and fungi exhibit multiple endoglucanases, and why multiple CBDs are present in some examples of these cellulases (Tomme *et al.* 1996).

Moreover, different CBDs are shown to have varied affinities for either amorphous or crystalline cellulose based on structural properties (Tomme *et al.* 1998). Notably CBDs with a high affinity for crystalline cellulose are all β-proteins containing a ridge of linearly arranged and regularly spaced, solvent exposed aromatics which are involved in the binding of the CBD to cellulose. Conversely, CBDs with an affinity for soluble or amorphous cellulose exhibit a small binding cleft rather than a linear array of binding residues. Such a cleft is rich in polar residues and small hydrophobics, but has relatively few aromatic residues. Based on these described structural features, CBDs can recruit the catalytic modules to specific regions on the cellulose fibre, either crystalline or amorphous.

An additional aspect to CBDs playing a role in recruitment is that of xylanases displaying cellulose binding capabilities, even though the catalytic domain solely targets xylan for hydrolysis (Gilbert and Hazlewood 1993). The fact that some xylanases display binding affinities for both cellulose and xylan, suggests that CBDs play a role in concentrating enzymatic action towards sites where xylan and cellulose are intimately associated. Since the only constant feature of plant cell walls is cellulose, it makes an ideal scaffold on which to

append or rally hemicellulases active against the associated hemicellulose polymers (Gilbert 2010).

1.3.4.3 Cell anchoring to the cellulose surface

Evidence that CBDs can display a cell-*to*-substrate anchoring function is quite prominent amongst non-motile cellulolytic bacteria, notably amongst the *Cellulomonas* sp. (Kenyon, Esch, and Buller 2005). A recent review of the CBD's roles in cellulose deconstruction highlights two examples whereby cellulases are tethered to the bacterial host cell wall, anchoring the cognate enzyme to the cell surface which in turn associates with the cellulose surface via the CBD (Gilbert 2010). It is thought that this functionality keeps the cellulases and their respective end-products within close proximity of the bacterium, much the same way that cellulosomes function in complexed cellulase systems. A second suggestion is that the anchoring facilitates the shuttling of appended enzymes from the bacterial surface to the substrate, improving the economics of the non-complexed system.

1.3.4.4 Non-hydrolytic disruption of cellulose fibres

A rate determining step in the biodegradation of cellulose is that of the initial adsorption of cellulases onto the cellulose microfibril, after which follows the rapid digestion of individual glucan chains. Considering this, it was hypothesized early on in the study of microbial cellulases that the enzymatic degradation of crystalline cellulose requires the initial action of a non-hydrolytic component to disrupt the hydrogen-bonding network that maintains the ordered lattices (Reese, Siu, and Levinson 1950). The first direct evidence to support this hypothesis was that of the *C. fimi* CenA cellulose binding domain. The CBD was shown to 'slough off' cellulose fragments from the substrate surface as the CBDs penetrated the fibres at surface discontinuities and released non-covalently attached fragments to uncover new cellulose chain ends. Moreover, the CBD disrupted the structure of cotton fibres as it travelled within the lattice, presumably between individual cellulose sheets (Din *et al.* 1991). It was later observed that small oligosaccharide particles were also released without any detectable hydrolytic activity being present (Din *et al.* 1994). This phenomenon was consequently confirmed within multiple CBD families (Levy, Shani, and Shoseyov 2002) (Shoseyov, Shani, and Levy 2006), and likened to the process in plant cell walls whereby expansin proteins disrupt hydrogen bonding between cellulose polymers in a non-hydrolytic

manner, facilitating elongation of the plant cell during growth and expansion (Cosgrove 2000). This concept of amorphogenesis is discussed further below.

1.3.5 Amorphogenesis: Non-hydrolytic agents increase substrate susceptibility

Cellulose microfibrils present a highly ordered and tightly packed architecture which occludes access of cellulases to the internal cellulose chains and by extension occlude access to the majority of the substrate. It has been suggested that the sequential 'shaving' and 'planing' of exterior cellulose fibrils by endo- and exoglucanases does not account for the entire deconstruction of the cellulose substrate, and that inaccessible regions are disrupted or loosened via non-hydrolytic proteins. Such a mechanism was termed amorphogenesis, a term coined by Coughlan to suggest a possible means by which the dispersion, swelling or delamination of cellulose resulted in a reduction in crystallinity and an increased internal surface area of cellulose exposed to hydrolysis (Coughlan 1985).

Several amorphogenesis-inducing agents, including modules and domains of cellulases have been observed to display this disrupting capacity. The model cellulolytic fungus *Trichoderma reesei* has been shown to secrete a non-catalytic protein termed swollenin, comprising a CBD connected via a linker region to an expansin-like domain (Saloheimo *et al.* 2002). Swollenin recombinantly expressed in yeast disrupted the structure of cotton fibres as well as cellulose filter paper with no detectable liberation of reducing sugars.

Similarly, a novel protein termed loosenin with the same disruptive tendencies as swollenin was observed in the cellulolytic fungus *Bjerkandera adusta* (Quiroz-Castañeda *et al.* 2011). Loosenin was shown to disrupt cotton fibres as well as fibres from the *Agave tequilana* plant. It was experimentally shown that incubation of the *Agave* fibres with commercial cellulases following pre-treatment with purified loosenin, led to a 7.5 fold greater liberation of reducing sugars compared to the untreated control.

Research by Kerff *et al.* (2008) has also shown a similar strategy employed by the soil bacterium *Bacillus subtilis* in promoting root colonization. The EXLX1 protein was shown to have structural similarity to plant expansins as well as similar binding and non-hydrolytic activities against the major plant cell wall polysaccharides. When recombinantly expressed in *E. coli*, EXLX1 bound to cellulose and weakened the structure of cellulose paper. In accordance with the aforementioned work on fungal swollenin and loosenin, EXLX1 also

showed a synergistic effect when combined with commercial cellulases, increasing celluloytic activity by up to 5.9 fold (Arantes and Saddler 2010).

The above results highlight a significant route utilized by cellulolytic fungi to the economical deconstruction and utilisation of lignocellulose biomass. Moreover, the apparent synergism between cellulases and non-hydrolytic agents is shown to be valuable in improving the economics of a non-complexed cellulase system.

1.4 The glycanases of *Cellulomonas fimi*

One of the better studied of the cellulolytic bacterial species is *Cellulomonas fimi*. Having a relatively well characterized and understood cellulolytic system, *C. fimi* offers itself as an attractive source for prospecting a library of cellulases for introduction in a recombinant host to study cellulose hydrolysis. Here we review the literature surrounding *C. fimi*, to highlight current gaps in knowledge and exciting avenues for future research.

1.4.1 Introducing *Cellulomonas fimi*

Cellulomonas fimi is a mesophilic facultative anaerobe. With optimum growth at 30^{o}C it is widely distributed amongst soil environments rich in decaying plant matter (Coughlan and Mayer 1992). *C. fimi* produces extracellular enzymes that hydrolyze starch, cellulose, chitin, xylan and mannan (Stoll, Stålbrand, and Warren 2001). Cells stain Gram positive and are motile, exhibiting one or more flagella. Being one of a few truely cellulolytic species of bacteria capable of efficient crystalline cellulose digestion as well as culturing under laboratory conditions, *C. fimi* has received much attention with regards to the hydrolysis of cellulose. As such, its genome has recently been sequenced and work is underway with regards to gene annotation and characterisation of encoded products (Lucas *et al.* 2012). Prior to the sequencing of *C. fimi*, extensive work in describing and characterizing its cellulase system has yielded considerable insight into the strategies and mechanisms of lignocellulose deconstruction. The known glycanase encoding genes of *C. fimi* are listed in Table 1.1 overleaf.

Gene name	Locus tag	Encoded protein function	E.C. code	Reference
cex (xyn10A)	Celf_1271	exo-1,4-β-glucanase	3.2.1.91	(O'Neill *et al.* 1986)
		endo-1,4-β-xylanase	3.2.1.8	
cenA (cel6A)	Celf_3184	endo-1,4-β-glucanase	3.2.1.4	(Wong *et al.* 1986)
cenB (cel9A)	Celf_0019	endo-1,4-β-glucanase	3.2.1.4	(Greenberg, Warren, *et al.* 1987)
cenC (cel9B)	Celf_1537	endo-1,4-β-glucanase	3.2.1.4	(Coutinho *et al.* 1991)
cenD (cel5A)	Celf_1924	endo-1,4-β-glucanase	3.2.1.4	(Meinke *et al.* 1993)
cbhA (cel6B)	Celf_1925	β-1,4-cellobiohydrolase	3.2.1.91	(Meinke *et al.* 1994)
cbhB (cel48A)	Celf_3400	β-1,4-cellobiohydrolase	3.2.1.91	(Shen *et al.* 1995)
man26A	Celf_0862	endo-1,4-β-mannosidase	3.2.1.78	(Stoll, Stålbrand, and Warren 1999)
man2A	Celf_2770	exo-1,4-β-mannosidase	3.2.1.25	(Stoll, Stålbrand, and Warren 1999)
xynC (xyn10B)	Celf_0574	endo-1,4-β-xylanase	3.2.1.8	(Clarke *et al.* 1996)
xynD (xyn11A)	Celf_0374	endo-1,4-β-xylanase	3.2.1.8	(Millward-Sadler *et al.* 1994)
		xylan deacetylase	3.5.1.-	(Laurie *et al.* 1997)
cfx	Celf_3156	endo-1,4-β-xylanase	3.2.1.8	(Hekmat *et al.* 2005)
cfbglu	Celf_2783	β-glucosidase	3.2.1.21	(Kim and Pack 1989)
nag3A	Celf_2983	β-glucosidase	3.2.1.21	(Mayer *et al.* 2006)
		3-β-N-Acetylglucosaminidase	3.2.1.52	(Mayer *et al.* 2006)

Table 1.1: The known glycanase encoding genes of *Cellulomonas fimi*. Table identifies the gene name(s), the NCBI locus tag (gene symbol), the experimentally determined encoded protein function, the Enzyme Commission code denoting the enzymatic activity, and the founding study.

The known glycanases of *C. fimi* according to published literature to date comprise the following; four endoglucanases, three exoglucanases including one dual specificity xylanase/cellodextrinase, two mannosidases, three xylanases, and two β-glucosidases. The majority of the glycanases of *C. fimi* have had their secondary structures determined and are shown in Figure 1.8 below. It can be noted that of these, only Man2A consists as a singular

catalytic domain, whilst the rest consist of multiple modules. These modules include a flexible linker region, a carbohydrate binding domain, and in some instances fibronectin type-III like repeats and other domains of unknown function or relatedness to known domains.

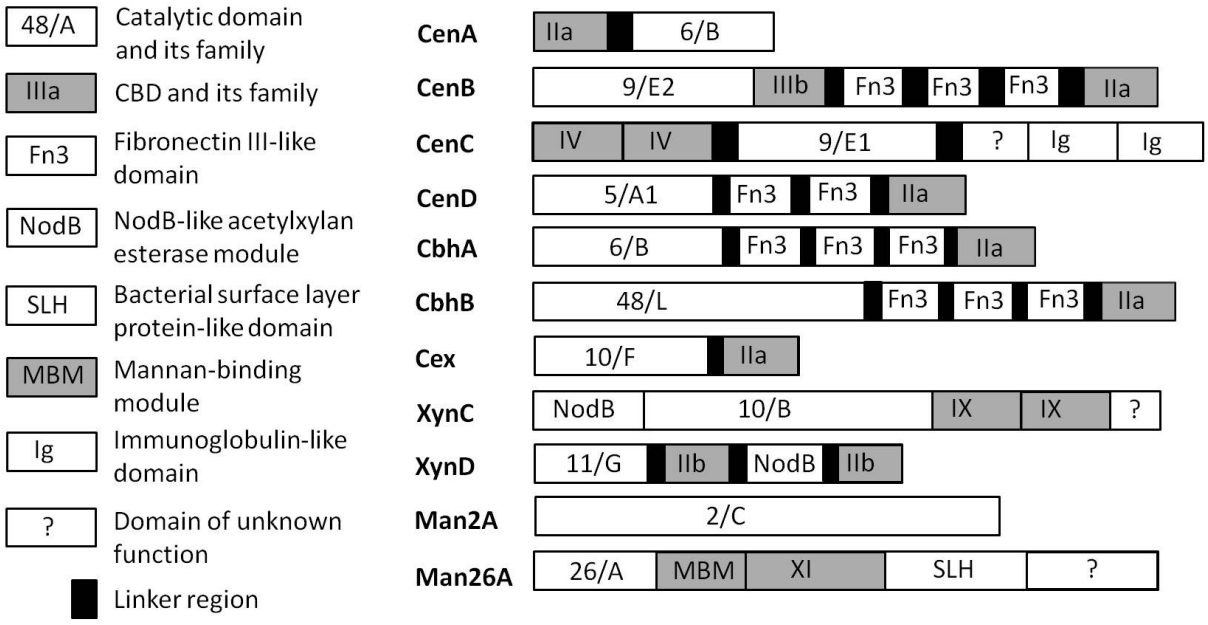

Figure 1.8: The known secondary structures of the *C. fimi* glycanases. Glycanases within non-complexed cellulases systems are typically modular, consisting of multiple domains and repeated sequences. Separate modules are linked via a flexible linker with no discernible secondary structure. Within the cellulases of *C. fimi*, the catalytic domain is typically present at the N-terminus, whilst the carbohydrate binding domain tends to be near the C-terminus. Fibronectin type-III like repeats of unknown function are seen in four cellulases; CenB, CenD, CbhA and CbhB. Data taken from: Tomme *et al.* (1998), Rabinovich, Melnick and Bolobova (2002), Le Nours *et al.* (2005), Sandercock *et al.* (1996), and Andreas Meinke *et al.* (1992).

1.4.2 The endoglucanases

The non-complexed cellulase system of *C. fimi* comprises four endoglucanases: CenA, CenB, CenC and CenD. The expression of multiple endoglucanases within a single cellulolytic system is a common feature within both bacteria and fungi capable of utilizing cellulose. By comparison, *Trichoderma reesei* as one of the most well studied cellulolytic fungi, exhibits a cellulase system comprising seven endoglucanases (Wilson 2008). It is generally agreed that multiple endoglucanases facilitate efficient cellulose deconstruction by performing subtly different yet complementary roles in both the disruption and hydrolysis of the substrate. These differences include: (i) hydrolysis of cellulose at non-overlapping sites on the substrate; (ii) varying degrees of processivity and frequency of cutting; (iii) and binding affinities for either amorphous or crystalline regions. Consequently, the endoglucanases

display varied activities against a range of substrates. These are summarised in Table 1.2 and this significance is discussed further below.

	Specific activity[a] on:					
Enzyme	**Avicel[b]**	**BMCC[b]**	**CMC[b]**	**PASC[b]**	**β-glucan[b]**	**Glucomannan[b]**
CenA	2.18	0.21	760	244	2,180	ND[c]
CenB	2.22	10.87	928	66	5,700	912.5
CenC	0.99	1.55	1,016	114	3,900	722.0
CenD	2.42	9.66	47	81	940	ND[c]
Cex	0.16	ND[c]	10	46	ND[c]	ND[c]

Table 1.2: Activities of *C. fimi* enzymes on soluble and insoluble glucans. [a] Specific activities are expressed as micromoles of reducing glucose per micromole of enzyme per minute. [b]Activity was determined with the dinitrosalicylic acid reagent after incubation of the enzyme with substrate at 30°C. [c] ND, not detected. Data taken from Tomme *et al.* (1996).

1.4.2.1 Endoglucanase CenA

CenA preferentially attacks amorphous zones within the cellulose substrate, though the CBD has been shown to bind both amorphous and crystalline cellulose (Warren 1996; Kleman-Leyer *et al.* 1994). This is experimentally shown in Table 1.2 as CenA shows the greatest activity against phosphoric acid swollen cellulose (PASC), an amorphous form of cellulose. Upon hydrolysis of cellulose chains, cellobiose is the major product liberated (Irwin *et al.* 1993). CenA has also been shown to hydrolyse the soluble oligosaccharide cellotetraose, and to a lesser extent cellotriose, to cellobiose and glucose (Kleman-Leyer *et al.* 1994; Damude *et al.* 1996). The expression of CenA is induced by glycerol, cellobiose and cellulose. No detectable expression is observed in the presence of glucose (Greenberg *et al.* 1987). Figure 1.9 illustrates the expression of the *C. fimi* cellulases in the presence of various carbon sources.

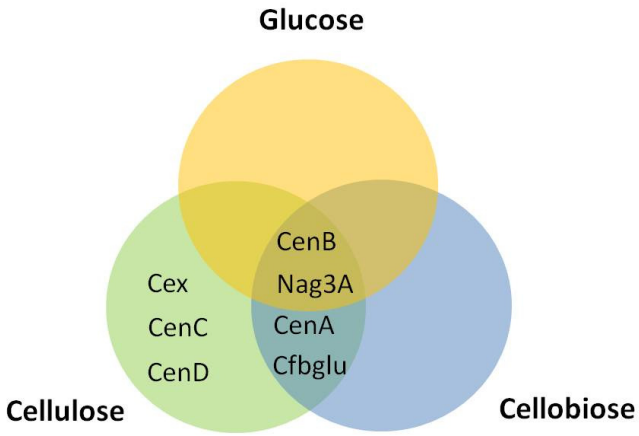

Figure 1.9: Venn-Diagram of cellulases present when *C. fimi* is grown on various carbon sources. The expression of cellulases in *C. fimi* is regulated by a combination of induction and catabolite repression. Nag3A and CenB are expressed regardless of the carbon source present, whilst Cex, CenC and CenD are only expressed in the presence of cellulose. CenA and Cfbglu are detectable when either cellulose or cellobiose is present. References include; Greenberg *et al.* (1987), Greenberg *et al.* (1987), Moser *et al.* (1989), Wakarchuk *et al.* (1984), and Hekmat *et al.* (2007).

A unique phenomenon of CenA, not observed amongst the other endoglucanases of *C. fimi*, is an intramolecular synergism between the catalytic and cellulose binding domains (Din *et al.* 1994). It was observed that the isolated CBD of CenA disrupts the structure of cellulose fibres, releasing fine cellulose particles without any detectable hydrolytic activity. Ramie fibres, a form of cellulose with a crystallinity of 70-74%, were treated with the fluorescently labelled purified CBD of CenA. Fluorescence was observed within the fibre as the CBD penetrated the crystalline lattice. Moreover, fibres stained with a fluorescent dye following pretreatment with CenA showed full penetration of the dye within the fibre, suggesting a high level of disruption compared to the untreated control. An extended incubation of up to six months of the CBD with the substrate did not lead to increased levels of reducing sugars, further suggesting a non-hydrolytic mechanism of substrate disruption (Cavaco–Paulo *et al.* 1999). Fibres pretreated with the isolated catalytic domain did not show any fluorescence, indicating that disruption of the cellulose structure was due to the presence of the CBD (Din *et al.* 1991). However the isolated catalytic domain 'polished' the surface of the fibres, hydrolysing exposed glucan chains as evidenced by the liberation of reducing sugars, and yielding fibres with a smooth exterior under scanning electron microscopy (Din *et al.* 1991).

The proposed mechanism of this synergy is that the CBD binds and penetrates at surface discontinuities to 'slough' off cellulose fragments not covalently associated with the fibre. Further penetration exfoliates the fibre releasing the cellulose chain ends and roughening the

cellulose surface, allowing for the action of the catalytic domain which hydrolyses β-1,4 bonds cleaving off vulnerable chains and smoothening the cellulose surface (Din *et al.* 1991).

An equally exciting role of the CBD of CenA that is not observed amongst the other endoglucanases is the dispersion of soluble oligosaccharides in solution. The isolated CBD as well as purified CenA were shown to prevent flocculation of BMCC and consequently disperse the substrate, presumably as a means to increase the effective surface area of the substrate and promote hydrolysis (Henrissat 1994).

1.4.2.2 Endoglucanase CenB

Unlike CenA, CenB shows high affinity for binding and hydrolysis of crystalline cellulose, releasing 50 times more reducing sugar from BMCC than CenA (Table 1.2). Purified CenB has been shown to hydrolyse up to 87% of BMCC which exists as ~76% crystalline cellulose (Meinke *et al.* 1993; Kleman-Leyer *et al.* 1994). In addition, CenB shows a low activity against glucomannan (Table 1.2), presumably hydrolysing β-1,4-glucosidic linkages as no appreciable activity is seen against either mannan or galactomannan (Tomme *et al.* 1996). Surprisingly the catalytic domain also binds cellulose, this being initiated by the 130aa sequence at the C-terminal side of the domain (Meinke *et al.* 1991).

The most striking feature of CenB however, is probably the existence of two CBDs which flank a series of three fibronectin type-III like sequences (Fig. 1.8). Fibronectin is a multifunctional, extracellular matrix and plasma protein of higher eukaryotes which is typically involved in protein-protein interactions in neural and muscle tissues. With this in mind, it has been suggested that CenB either interacts with other cellulase components exhibiting fibronectin like repeats (see Figure 1.8), or is associated to the cell surface by the Arg-Gly-Asp-Ser sequence which is known to mediate cell adhesion (Meinke *et al.* 1991).

To date, CenB is the only endoglucanase of the *C. fimi* cellulase system found to be constitutively expressed, regardless of carbon source (Fig. 1.9). Expression is under the control of two promoters: one weak constitutive promoter which is proximal to the transcription start site, and one strong inducible promoter which is distal to the transcription site. The inducible promoter was shown to be unaffected by the addition of glycerol, cellobiose or glucose, and only initiated transcription in the presence of cellulose (Greenberg *et al.* 1987).

1.4.2.3 Endoglucanase CenC

CenC is the only endoglucanase to have two CBDs at the N-terminus. These are arranged in tandem and show a binding affinity to amorphous, but not crystalline cellulose. This preference for the disordered form of cellulose rather than the regular crystalline form is attributed to the binding site within both CBDs being a groove like cleft (Brun *et al.* 2000). Binding is orchestrated through the formation of hydrogen bonds between hydroxyl groups along the substrate surface and polar amino acid side chains lining the CBD cleft. Both CBDs have a strong affinity for cellotetraose and cellopentaose (Johnson *et al.* 1996; Brun *et al.* 2000).

CenC is said to be a semi-processive endoglucanase, adsorbing to cellulose and hydrolysing sequential β-1,4,-glycosidic bonds from the reducing end of the chain, before desorbing from the substrate and initiating adsorption and attack elsewhere (Irwin *et al.* 1993; Tomme *et al.* 1996). The major product of hydrolysis of cellulose is cellobiose. Like CenB, CenC also shows low activity on glucomannan, presumably acting on glucan chains decorating the mannan backbone (Table 1.2).

CenC is known to have two C-terminal immunoglobulin-like domains of unknown function (Figure 1.8). These do not bind to avicel or any cellulosic substrates, and thought to possibly be involved in protein-protein interactions serving a function similar to the fibronectin-like repeats seen in other cellulase components (Coutinho *et al.* 1992). In relation to the other *C. fimi* cellulases, the expression of CenC is induced by cellulose, but not glucose or glycerol (Moser *et al.* 1989).

1.4.2.4 Endoglucanase CenD

CenD is the least well studied of the *C. fimi* endoglucanases. Expression of CenD is induced in the presence of cellulose, and is undetectable when *C. fimi* is grown in media containing cellobiose or glucose as the sole carbon source (Fig. 1.9). The CBD binds both amorphous and crystalline cellulose (Warren 1996). However the hydrolytic activity against crystalline substrates is far greater than that of amorphous substrates (Kleman-Leyer *et al.* 1994) (Table 1.2). By comparison, CenD is able to hydrolyse up to 85% of BMCC, a figure comparable to that of CenB which hydrolyses up to 87% of the substrate (Meinke *et al.* 1993). Like CenB, the presence of fibronectin-like repeats is also apparent in CenD (Sandercock *et al.* 1996) (Fig. 1.8).

1.4.3 The exoglucanases: cellobiohydrolases and cellodextrinases

1.4.3.1 Cellodextrinase Cex

Cex is the only known cellodextrinase of *C. fimi*. It shows activity against both cellulose and xylan, though notably it is 40 times more active on xylan (Notenboom *et al.* 1998). These two activities are discussed individually within this section.

As an exoglucanase, Cex hydrolyses cellulose and cellotetraose oligosaccharides to liberate cellobiose from the non-reducing chain ends (White *et al.* 1994). The CBD of Cex binds both amorphous and crystalline regions of cellulose (Mclean *et al.* 2000; Warren 1996; Jervis *et al.* 1996). Though its adsorption to crystalline BMCC is irreversible (Tomme *et al.* 1998; Esteghlalian *et al.* 2001), the CBD has been shown to migrate along the surface of crystalline cellulose, presumably facilitating the diffusion of the attached catalytic domain and increasing the likelihood of encountering available glycosidic linkages for attack (Jervis, Haynes, and Kilburn 1997). Moreover, evidence suggests that the binding of the CBD can be either perpendicular or parallel to the cellulose chain, further increasing the range and directionality over which the enzyme may act (Mclean *et al.* 2000).

Overexpression of Cex in *E. coli* has a toxic effect upon the host, leading to cell death (Fu *et al.* 2005). Expression under a weaker promoter however improves cell viability and mature protein expression levels, relieving toxic effects (Fu *et al.* 2006). In other work, recombinant expression of Cex in *E. coli* was performed by constructing fusion proteins of either the catalytic domain or CBD to an outer membrane protein (Francisco *et al.* 1993). Exoglucanase activity was retained by the fusion protein, with 90% of activity occurring at the cell surface. In addition to this, binding of cells to crystalline cellulose was also observed.

1.4.3.2 Cellobiohydrolases CbhA and CbhB

One of the most extensively studied cellulolytic microorganisms is *Trichoderma reesei*, a soft-rot fungus. The mechanism by which crystalline cellulose is hydrolysed by *T. reesei* is not completely understood, though it is generally agreed to involve the concerted action of two cellobiohydrolases; cellobiohydrolase I and II, representing approximately 60% and 20% of extracellular protein, respectively, attacking from opposite ends of the cellulose chain. It has been shown that a similar mechanism is also employed by some cellulolytic bacteria, including *C. fimi* (Gilkes *et al.* 1997).

The cellobiohydrolases of *C. fimi* include CbhA and CbhB. CbhA shows structural and catalytic properties similar to those of *T. reesei* CBH II, whereas CbhB is not related to either CBH I or II, and instead is classed as a family L-glucanase (Hua *et al.* 1996). Both enzymes are modular with a similar arrangement of CBD and catalytic domain; a C-terminal CBD is joined to the N-terminal catalytic domain by three fibronectin type-III like modules, each containing 95 to 98 amino acid residues (Sandercock *et al.* 1996). This arrangement is illustrated in Figure 1.8.

The CBDs of both cellobiohydrolases have been shown to bind both bacterial microcrystalline cellulose (BMCC) and carboxymethyl-cellulose (CMC), representative examples of crystalline and amorphous cellulose, respectively (Warren 1996). A preference is evident for the attack of unsubstituted residues from the terminal-ends of CMC, though the hydrolysis of both crystalline and amorphous substrates liberates cellobiose as the major product (Irwin *et al.* 1993). CbhB is able to hydrolyse cellohexaose, cellopentaose and cellotetraose, however not cellotriose. A distinct difference between CbhA and CbhB is the directionality of attack along the cellulose chain; CbhA attacks from the non-reducing end and CbhB from the reducing end (Gilkes *et al.* 1997). These complementarities are consistent with other microbial cellulase systems such as that of *T. reesei*, and are suggested to be a commonality between the cellulose hydrolysing strategies of bacteria and fungi.

1.4.4 The β-glucosidases

C. fimi is known to express two strictly cytoplasmic β-glucosidases; Nag3A and Cfbglu (Wakarchuk *et al.* 1984). Expression of Cfbglu is induced four-fold in the presence of avicel and seven-fold in the presence of cellobiose. The enzyme hydrolyses the β-1,4-glycosidic linkages in 4'-nitrophenyl β-D-glucopyranoside, 4-methylumbelliferyl-β-D-glucoside and cellobiose. Unlike Nag3A it is a true cellobiase, capable of hydrolysing cellobiose to two glucose monomers (Kim and Pack 1989).

Nag3A has dual activities as both a 3-β-N-glucosaminidase and a β-glucosidase (Mayer *et al.* 2006). The enzyme hydrolyses terminal non-reducing acetylglucosamine residues from glycoproteins as well as β-1,4-glycosidic linkages in 4'-nitrophenyl β-D-glucopyranoside and 4-methylumbelliferyl-β-D-glucoside, though not cellobiose. Nag3A is constitutively expressed, regardless of carbon source (Wakarchuk *et al.* 1984).

1.4.5 The Xylanases

Cex catalyses the hydrolysis of β-1,4-xylosidic linkages at internal sites along the xylan polymer, showing 1,500 times more activity against xylan than any of the *C. fimi* endoglucanases (Kleman-Leyer *et al.* 1994). Expression of Cex is induced in the presence of xylan or cellulose, with the highest levels of induction in the presence of both substrates (Hekmat *et al.* 2008). This expression profile highlights the enzyme's roles as a broad specificity endoxylanase and exoglucanase, active on both cellulose and xylan. Similarly, the endoxylanase Cfx is also induced by both cellulose and xylan (Hekmat *et al.* 2008). It would appear that this is an economical strategy for the deconstruction of lignocellulose, as plant biomass encountered *in situ* would contain both cellulose and xylan polymers and not those individually.

This close intimacy of cellulose to xylan means that cellulose offers itself as a consistent target to which xylanases can bind, and so suggests why xylanases commonly show cellulose binding capabilities. For example, the endoxylanase XynC has two family IX CBDs which bind cellulose, but not xylan (Clarke *et al.* 1996). XynC shows no hydrolytic activity against cellulose or mixed linkage β-glucans, but hydrolyses soluble xylan to produce predominantly xylobiose. Interestingly, maximal activity is shown to occur at 60^{o}C, with 40% of that activity retained at 70^{o}C. In addition, XynC contains a domain homologous to the nodulation protein, NodB, from the nitrogen fixing *Rhizobium* spp (Fig. 1.8).

The NodB domain, also present in XynD, has been shown to deacetylate acetylxylan (Laurie *et al.* 1997). Acetylxylan is resistant to hydrolysis by xylanases and as much as 70% of xylose residues found within hardwoods can be acetylated. Like XynC, XynD is an endoxylanase showing stability at higher temperatures, being stable up to 52^{o}C (Millward-Sadler *et al.* 1994). However XynD also shows a binding affinity for both cellulose and xylan, coordinated by two family IIb CBDs (Fig. 1.8). Binding to crystalline cellulose is co-ordinated by the C-terminal CBD (Bolam *et al.* 2001), and binding to xylan by the internal CBD (Black *et al.* 1995). The two CBDs show 70% sequence similarity and exhibit intramolecular synergy whereby the binding affinity for xylan is 20-times greater when both CBDs are incorporated into a single protein species than when expressed as distinct species (Bolam *et al.* 2001).

1.4.6 The Mannosidases

C. fimi is known to express two mannosidases: Man2A and Man26A. In combination, both enzymes are sufficient for the complete hydrolysis of ivory nut mannan, an insoluble and unsubstituted mannan polymer, to its constituent monomer of mannose (Stoll, Stålbrand, and Warren 2001).

1.4.6.1 Man2A

Man2A is an exo-1,4-β-mannosidase (E.C. code 3.2.1.25) catalysing the removal of β-D-mannose residues from the non-reducing ends of mannan oligosaccharides (Stoll, Stålbrand, and Warren 1999). It is an intracellular enzyme comprising a single catalytic domain (Figure 1.8). Man2A readily hydrolyses p-nitrophenyl-β-mannoside, and does not show hydrolytic activity against any other saccharide derivatives of p-nitrophenyl such as β-glucose, β-xylose or β-cellobiose (Stoll, Stålbrand, and Warren 1999). The complete hydrolysis of manno-oligosaccharides to mannose is observed in the presence of Man2A (Stoll, Stålbrand, and Warren 2001), with end-product inhibition also being observed (Zechel *et al.* 2003).

1.4.6.2 Man26A

Man26A is an endo-β-1,4-mannosidase (E.C. code 3.2.1.78) and catalyses the random hydrolysis of β-1,4 mannosidic linkages within the backbones of mannans, galactomannans, and glucomannans (Stoll, Stålbrand, and Warren 1999). Complete hydrolysis of mannotetraose, mannopentaose and mannohexaose to mannose and mannobiose as major products is observed. Hydrolysis of mannotriose is comparatively slow and that of mannobiose is undetectable (Stoll, Stålbrand, and Warren 2001). Unlinke Man2A, Man26A is a secreted mannanase, with a modular structure (Figure 1.8). Proteolysis by a *C. fimi* serine protease gives active fragments displaying both independent binding and hydrolytic activities (Stoll, Stålbrand, and Warren 1999).

A unique feature of the modular Man26A is the presence of a putative SLH like domain (Fig. 1.8). These are present in cellulosome proteins from complexed cellulase systems and are implicated in the attachment and anchoring of cellulase components to the cellulosome scaffold (Le Nours *et al.* 2005). It is unclear what the role of the SLH like domain is in Man26A, though the enzyme has been shown to be transiently associated to the outer

membrane, with a large proportion of mannanase activity associated with the cell when recombinantly expressed in *E. coli* (Stoll *et al.* 1999). Further work identified that Man26A does not bind to a peptidoglycan fraction prepared from *C. fimi* and perhaps the SLH domain functions otherwise (Stoll, Stålbrand, and Warren 2001).

The mannan binding module of Man26A binds soluble mannans in the form of azo-carob galactomannan and locust bean gum, but does not bind insoluble mannan in the form of ivory nut mannan (Stoll *et al.* 2000). This is striking since the binding domain enhances the activity of the catalytic domain on insoluble substrates, but not on the soluble substrates to which it binds. Binding is shown to be reversible (Stoll, Stålbrand, and Warren 2001).

1.4.7 Modular cellulases and the role of a conserved linker region

Extensive study over the past two decades has yielded exciting insights into the role and importance of the flexible linker region present amongst the cellulases of *C. fimi*. Work by Warren *et al.* (1986) first indentified the conserved linker region between the catalytic and cellulose binding domains of the cellulases CenA and Cex. This short linker region was shown to be of about 20 amino acids in length and to contain only proline and threonine (the Pro-Thr box). The region is conserved almost perfectly in the two enzymes and suggested to have arisen due to the shuffling of three or four conserved nucleotide sequences.

Langsford *et al.* (1987) proved that glycosylation of cellulases from *C. fimi* did not significantly affect their kinetic properties, or their stabilities towards heat and pH. However, glycosylation of enzymes was necessary for protection from the attack of a secreted *C. fimi* protease when bound to cellulose. Non-glycosylated counterparts synthesized in *E. coli* were vulnerable to attack and yielded active, truncated products with reduced affinity towards crystalline cellulose.

Furthering this study, Gilkes *et al.* (1988) showed that the extracellular serine protease is secreted by *C. fimi* when grown on glycerol, xylan or cellulose. In accordance with Langsford *et al.* (1987), native glycosylated CenA and Cex were resistant to proteolysis when bound to cellulose, but susceptible when in solution (Fig. 1.10). The protease cleaved both CenA and Cex in a highly specific manner, cleaving at the conserved Pro-Thr box, making independent the catalytic and CBD modules in each protein. The independent modules retained their respective hydrolysing and cellulose binding functions. However the catalytic domains

showed reduced activity against insoluble substrates and increased activity against soluble substrates, implying a critical role of the CBD in the hydrolysis of crystalline cellulose.

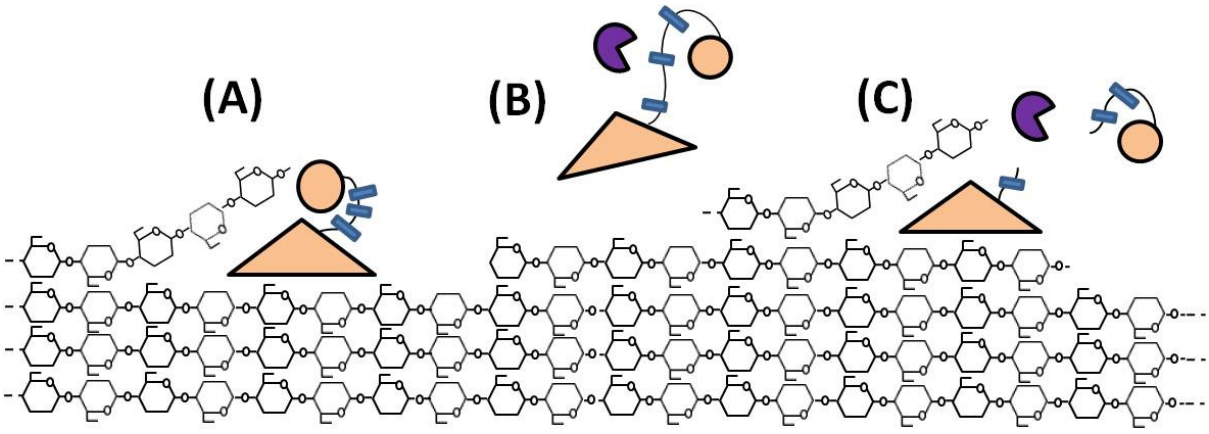

Figure 1.10: Schematic representation of the proteolytic cleavage of a hypothetical modular exoglucanase. The glycosylated linker region is protected from proteolysis when the CBD is bound to crystalline cellulose **(A)**; when in solution and hydrated, the linker region is susceptible to attack by the secreted extracellular serine protease **(B)**; once cleaved, the catalytic and cellulose binding modules are made independent and perform their respective roles of cellulose hydrolysis and binding **(C)**.

Work by Shen *et al.* (1991) showed that deletion of the linker region does not affect the enzymatic activity of CenA or its adsorption to cellulose. However deletion does affect desorption from the substrate as the modified CenA, unlike native CenA, cannot be eluted from cellulose with water. Since removal of the linker did not significantly impair enzymatic activity, Miller *et al.* (1992) substituted the linker region of CenA with the human antibody IgA1 linker region. Results showed that recombinant CenA was unaffected by the wild-type *C. fimi* serine protease, however showed susceptibility to the IgA protease from *Neisseria gonorrhoeae*. Fragments generated from this cleavage were identical to those generated from wild-type *C. fimi* protease, resulting in a fully functional catalytic domain and CBD.

Considering that glycosylation is not essential for activity, but is necessary for the protection of cellulases from proteolysis when bound to cellulose, Coughlan and Mayer (1992) suggested that the modular cellulases are slowly cleaved to yield active fragments with enhanced activity against soluble substrates which accumulate as crystalline cellulose is degraded. Evidence to support this hypothesis was later found when the same modularity present in CenA was also shown to exist in CenB, CenD, Cex, CbhA and CbhB (Sandercock *et al.* 1996). The enzymes were degraded proteolytically in supernatants of *C. fimi* cultures, with the discrete fragments generated maintaining catalytic function, displaying reduced

activity against intact or insoluble cellulose fibres and increased activity against free-fibres in solution. Sanercock *et. al* (1996) found this difference to be due to the CBD no longer anchoring the catalytic domain to the cellulose fibre.

Finally, Poon *et al.* (2007) using Nuclear Magnetic Resonance (NMR) spectroscopic analysis showed that the proline-threonine (PT) linker of Cex did not exhibit any predominant structure in either glycosylated or non-glycosylated forms. The PT linker was shown to be flexible and glycosylation slightly dampened this flexibility. Interestingly, it was also shown that there are no non-covalent interactions between the two domains of Cex or between both domains and the linker. Poon *et al.* demonstrated that the PT linker is a flexible tether, joining the structurally independent catalytic and cellulose binding domains of Cex in an ensemble of conformations. These findings support the idea that the CBD anchors Cex to the surface of cellulose, whilst the linker provides flexibility for the catalytic domain to hydrolyse nearby cellulose or xylan chains.

1.4.8 Modification of the linker for controlled cleavage within a heterologous host

Considering the above review, one can appreciate the importance of the linker region within the modular cellulases. In the presence of crystalline cellulose, the cellulases will bind cellulose via the CBD and be protected from proteolysis as the linker is occluded, however as hydrolysis of the substrate reaches completion, the enzymes will desorb and spend more time in solution where the linker is susceptible to cleavage (Cavaco–Paulo *et al.* 1999). Upon cleavage of the linker, the now independent CBD shows a tendency to disrupt the structure of the remaining insoluble substrate, whilst the catalytic domain shows greater activity against soluble oligosaccharides. This two-stage strategy would appear to improve the efficiency of a non-complexed cellulase system, since the modules perform multiple roles over the course of substrate hydrolysis and are repurposed as the substrate is hydrolysed.

The relevance of this within the present study is the potential design of a similar two step process for cellulose degradation within a recombinant host, replicating that of *C. fimi*. Using the technique by Miller *et al.* (1992) discussed earlier, substitution of the CenA linker region with an alternative sequence could afford the controlled proteolysis of the cellulase and consequently controlled activity of the independent modules. In theory, the disassociation of catalytic and cellulose binding domains could be controlled by inducing expression of an appropriate protease capable of cleaving the substituted recombinant linker region. Moreover,

to leverage the independent catalytic domain's increased activity against soluble oligosaccharides, protease expression could coincide with the depleting cellulose carbon source. To achieve this, one could use the method devised by Bokinsky *et al.* (2011) whereby transcription of recombinant genes in *E. coli* is under the control of the *wrbA, cstA* or *cspD* promoters which initiate transcription prior to stationary phase or a limited carbon supply. Ultimately, the controlled proteolysis of the modular cellulases in wildtype *C. fimi* is thought to promote the extensive hydrolysis of crystalline cellulose (Meinke *et al.* 1991; Ong *et al.* 1994), and perhaps is a necessary mechanism to duplicate if the full cellulolytic efficiency of *C. fimi* is to be realised in a heterologous host.

1.4.9 Speculated roles of the fibronectin type-III repeats within the cellulases of *C. fimi*

The fibronectin type-III (Fn3) like repeats observed within the *C. fimi* cellulases CenB, CenD, CbhA and CbhB (shown in Fig.1.8) are also exhibited within the cellulase components of other cellulolytic bacteria and fungi. The conserved Fn3 unit is shown to have a distinctive motif of seven anti-parallel β-strands arranged in two sheets, enclosing a core of highly conserved hydrophobic residues (Kataeva *et al.* 2002). This domain is reported in a number of bacterial species which degrade cellulose and/or chitin, including *Bacillus circulans, Cellulomonas flavigena, Clostridium thermocellum* and *Clostridium cellulovorans* (Hansen 1992). The acquisition of Fn3 sequences was originally suggested to have been from an animal species and propagated down the lineage through natural selection, and across different genera through horizontal gene transfer (Little, Bork, and Doolittle 1994), although a more recent finding of its presence in fungal systems gives reason to dispute this claim.

The Fn3 repeats are also present in swollenin, the expansin like protein from *T. reesei*. Interestingly, this is the first and only example of these repeats occurring within a fungal species (Saloheimo *et al.* 2002). Swollenin shows no hydrolytic activity against cellulose, but is involved in the non-hydrolytic disruption of the crystalline lattice, introduced earlier within this thesis as amorphogenesis. Considering that swollenin is not implicated in cellulose hydrolysis but associated with the disruption of the structure instead, one could assume the Fn3 repeat is also involved in this role.

One example within the cellobiohydrolase CbhA of *Clostridium thermocellum* suggests this assumption to be true as the Fn3 repeats are observed to promote cellulose hydrolysis by modifying the substrate surface (Kataeva *et al.* 2002). Using PCR cloning to prepare

truncated forms of the cellulase, Kataeva *et al.* measured the efficiency of cellulose hydrolysis and observed structural changes by scanning electron microscopy (SEM). Results demonstrated that the efficiency of cellulose hydrolysis by the truncated forms of CbhA increase in the following order; CD (lowest efficiency), CD-Fn3-Fn3 (more efficient), and CD-Fn3-Fn3-CBD (greatest efficiency). Moreover, SEM studies of filter paper treated with Fn3 domains showed that the surface of the cellulose fibres had been loosened and crenellated, displaying roughening and erosion, however no changes in the crystallinity of the fibres were observed. It was concluded that the Fn3 domain loosened neighbouring cellulose chains by exfoliation and separation, directing single chains towards the catalytic centre of the enzyme. A tighter interaction with the cellulose substrate was mediated through the CBD, as Fn3 was found to bind cellulose only weakly.

Regarding *C. fimi*, it has been demonstrated that the Fn3 repeats may play a role in the aggregation of the modular cellulase components, which arise from proteolytic cleavage of the enzyme. Meinke *et al.* (1992) identified a truncated module of the endoglucanase CenB that strongly aggregated and bound to the surface of crystalline cellulose. The module was shown to contain the C-terminal family IIa CBD and the three Fn3 repeats (Figure 1.8). Considering that the disruption and dispersion of cellulose fibres has previously been shown to be performed by the Fn3 domains (Kataeva *et al.* 2002), as well as the CBD domains (Din *et al.* 1994), one may speculate that the Fn3-CBD module arising from proteolysed CenB could also fulfil a similar role in cellulose disruption.

The role of the Fn3 domains is still however unclear. Given the absence of the Fn3 domains amongst numerous homologues of the cellulases across multiple cellulolytic species, it has been suggested that the role of the Fn3 domains is an accessory one and not essential to the function of the hydrolases (Little, Bork, and Doolittle 1994).

1.4.10 Cell anchoring and association to cellulose via extracellular *C. fimi* cellulases

As previously discussed, the CBMs within the cellulases of *C. fimi* fulfil numerous roles; (i) recruitment of enzymes to target regions, such as CenB and CenD to areas of crystalline cellulose (Tomme *et al.* 1998); (ii) mobility of the cellulases along the cellulose surface, in the case of CenA and Cex (Jervis, Haynes, and Kilburn 1997); and (iii) non-hydrolytic disruption and dispersion of cellulose fibres, exhibited by CenA (Din *et al.* 1991). In addition to these observations, there is evidence within the literature which suggests an alternative role

in that cellulases are anchored to the cell surface via the CBD, forming a rudimentary cellulosome complex.

Coughlan and Mayer (1992) suggested that *C. fimi* initiates hydrolysis of cellulose via a two step mechanism; (i) the first step is the presentation of the cellulase components on the cell surface whereby hydrolysis is mediated in a similar fashion to that of a cellulosome; (ii) the second is the dissociation of the putative cellulase complexes from the cell surface and free diffusion of the now independent agents to act on the substrate in accordance with a non-complexed cellulase system. This suggestion was based on the observation that the related *Cellulomonas uda,* which also displays a non-complexed cellulase system, presents multiple protuberant structures on the cell surface when grown in cellulose or cellobiose (Lamed *et al.* 1987). These structures were likened to the cellulosome of complexed systems, as they are absent in the presence of glucose, and are thought to facilitate binding of cells to the cellulose substrate prior to hydrolysis.

Currently no evidence within the literature describes similar protuberant structures on the surface of *C. fimi* in the presence of cellulose. However it is tempting to hypothesize that the endoglucanase CenB has a role here; the weak constitutitve expression of CenB as well as the presence of Fn3-like repeats thought to govern protein-protein interactions could suggest that the cellulase is presented on the cell surface as part of the basal cellulase system of *C. fimi*. Upon contact with cellulose, CenB could inititate the preliminary hydrolysis of cellulose and liberation of small-oligosaccharides which in turn induce the expression of the remaining cellulases. The fact that CenB expression is also under the control of a second strong inducible promoter as well as the enzyme's strong affinity for crystalline cellulose could support its role as the initiating cellulase component in cellulose deconstruction.

1.5 Synergistic combinations of cellulases for cellulose hydrolysis

Synergism within cellulase systems was first proposed by Eveleigh (1987), and has since been widely observed both within and between cellulolytic bacteria and fungi. Several types of synergism have been reported within the literature and including; (i) endo-exo synergy between endoglucanases and exoglucanases (ii) exo-exo synergy between exoglucanases processing from reducing and non-reducing ends of cellulose chains, (iii) endo-endo synergy between endoglucanases targeting different sites within the cellulose substrate, (iv) synergy between exoglucanases and β-glucosidases which relieve end-product inhibition by

cellobiose, and (v) intramolecular synergism between catalytic and cellulose binding domains (Zhang and Lynd 2004).

The study of synergism within cellulase systems has met with some difficulty, notably due to some synergistic effects only being operative in specific conditions. For example, intramolecular synergism between the catalytic and cellulose binding domains within CenA was reported on cotton fibres but not on BMCC (Din *et al.* 1994). Moreover, the quantification of synergistic effects is troublesome as the overall deconstruction of the substrate by complementary cellulases will show no overall modification in architecture. This was discussed by Mansfield & Meder (2003). The authors noted that synergistic effects of cellulase mixes showed no overall deconstruction of the substrate as the actions of individual cellulases are offset by the concurrent modification by complementing enzymes.

In addition to substrate properties, experimental conditions also affect the extent of synergy observed. It has been reported that endo- and exo-glucanase synergies increase with an increase in enzyme loading below saturation, but decrease with oversaturated enzyme loading (Irwin *et al.* 1993). This is thought to be due to competitive binding for sites on the substrate when it is oversaturated with enzyme. Furthermore, synergistic effects are noted to be greatest under conditions chosen to minimize end-product inhibition. An example of this is illustrated by Irwin *et al.* (1993), whereby the addition of β-glucosidase to a cellulase mix of endo- and exo-glucanases improved total substrate deconstruction by up to two-fold, presumably due to relief of end-product inhibition ocaused by cellobiose.

1.5.1 Combinations of *C.fimi* cellulases for efficient hydrolysis of cellulose

Utilizing knowledge gained in the study of synergy between cellulases, the deconstruction of cellulose by a recombinant host can be improved by the design of complementary combinations of cellulases. This was investigated by Mansfield and Meder (2003), using four recombinant cellulases from *C. fimi,* interrogated the respective roles and activities of the enzymes in cellulose hydrolysis. Table 1.3 below summarizes the extent of Sigmacell cellulose hydrolysis by the individual cellulases and pair-wise combinations thereof.

	Glucose (mg)	Cellobiose (mg)	Cellotriose (mg)	Total sugars (mg)	% Hydrolysis
CenA	11.35	40.68	0.00	52.03	30.61
CenD	2.64	77.51	5.42	85.57	50.34
CbhA	0.53	73.78	3.30	77.61	45.65
CbhB	0.00	22.10	8.45	30.55	17.97
CenA + CbhA	10.86	88.07	1.10	100.03	58.84
CenA + CbhB	13.97	38.22	0.68	52.87	31.10
CenD + CbhA	**0.55**	**108.33**	**5.75**	**114.63**	**67.43**
CenD + CbhB	0.00	84.10	9.98	94.08	55.34

Table 1.3: Soluble oligosaccharides liberated and the degree of saccharification of Sigmacell cellulose hydrolysed by recombinant *C. fimi* cellulases over 48 hours. Data taken from Mansfield and Meder (2003).

The most productive cellulase combination in terms of total hydrolysis of cellulose was that of CenD and CbhA with 67.43%. Notably this combination was also shown to increase the relative crystallinity of the substrate as illustrated in Table 1.4 below. This observation is most likely due to CbhA preferentially hydrolysing the less recalcitrant amorphous regions of the Sigmacell cellulose, explaining why a larger proportion of the substrate was hydrolysed. The least productive enzyme combination was that of CenA and CbhB which showed only 31.1% hydrolysis of Sigmacell, half that of CenD and CbhA, though it is of importance to note that the combined action of CenA and CbhB reduced the relative crystallinity of the substrate, suggesting that this pair of enzymes targeted the crystalline regions of the substrate.

These results are consistent with previous studies which show CbhB to have a higher activity against crystalline BMCC and PASC compared to Sigmacell cellulose, suggesting the cellobiohydrolase is effective in decrystallizing cellulose (Stålbrand *et al.* 1998), whereas CbhA is more effective in the solubilisation of cellulose, hydrolysing 45.65% of the substrate, compared to CbhB at 17.97%.

Cellulases	CenA	CenD	CbhA	CbhB
CenA				
CenD				
CbhA				
CbhB				

Change in crystallinity:

Increase

No change

Decrease

Table 1.4: The crystallinity index of Sigmacell cellulose is altered according to the combinations of cellulases present. Pair-wise combinations of the *C. fimi* cellulases CenA, CenD, CbhA and CbhB are shown to have varying effects on the crystallinity of cellulose. Combinations which increase the relative crystallinity are shown as dark green, those having no effect are shown as light green, and finally combinations decreasing the crystallinity of cellulose are shown as pale green. Data taken from (Mansfield and Meder 2003).

This study highlights the significance of synergism within the cellulase system of *C. fimi*, whereby specific combinations of cellulases act in concert to fulfil defined roles, including: the decrystallization of cellulose by CenA and CbhB, and subsequent solubilisation by CenD and CbhA. Furthermore, evidence within the literature suggests that the relative proportions of the individual enzymes within these combinations are a key determinant in their combined enzymatic efficiency. This is discussed further in the following section.

1.5.2 Optimized proportions of endo- and exo-glucanases

Evidence within the literature indicates that the optimal cellulase mix required for cellulose hydrolysis will vary greatly according to the substrate used within the study, as well as the organism from which the enzymes are derived. Using purified cellulases from *T. reesei* Baker *et al.* (1998) showed that as much as 16.5% of Sigmacell cellulose could be hydrolysed over 120 h when an optimal ternary mix was employed, consisting of cellobiohydrolase Cel7A and Cel6A and endoglucanase Cel7B in a molar ratio of 60:20:20. By comparison, Boisset *et al.* (2001) using the homologous cellobiohydrolases Cel7A, Cel6A and the endoglucanase Cel45A of *Humicola insolens* in a molar ratio of 69:30:1 showed a far greater extent of hydrolysis as 90% of BMCC could be hydrolysed over a much shorter incubation of 24 h.

A reoccurring observation within the literature is that an optimal enzyme mix will comprise a type I and type II cellobiohydrolase and a single endoglucanase in a molar ratio of approximately 70:30:1, respectively (Meyer, Rosgaard, and Sørensen 2009). It has been suggested that this ideal ratio is a direct result of how the endo- and exoglucanases bind to and subsequently hydrolyse cellulose chains; a larger proportion of the cellulase mix will

consist of the exoglucanases as these exhibit a longer residence time bound to the cellulose chain, compared to the endoglucanases which are only transiently associated long enough to hydrolyse at a single internal site before migrating to neighbouring chains (Linder and Teeri 1997).

Presently no study within the literature describes an optimal ratio of endoglucanase to exoglucanase within the cellulase system of *C. fimi*, though one can expect it to match those previously described given the extensive similarities and overlap between microbial cellulase systems.

1.6 Synthetic Biology for lignocellulose deconstruction

The study of optimized cellulase cocktails alludes to the potential of employing a refined and more precise approach to studying the deconstruction of lignocellulosic biomass. Through exploiting synergy between glycanases, and reducing reliance upon a brute-force approach of high level cellulase expression, competitive inhibition between cellulases can be reduced and overall enzymatic efficiency improved. As such, the desired result would be the design of an enzyme cocktail targeted to a pre-defined cellulosic substrate. This could potentially alleviate metabolic stress upon an already overburdened production host, which is of paramount importance when considering the design of a single chassis organism for both biomass utilisation and secondary product formation (Lynd *et al.* 2005).

The study of synergism between cellulases and therefore the understanding of cellulose degradation would benefit greatly from an extensive library of well characterized cellulases whose properties are clearly defined (French 2009). This would allow for the selection of individual enzymes for the introduction within a chassis organism to degrade a pre-defined cellulosic substrate. Here, we discuss the potential of leveraging tools and concepts from sythetic biology for developing a combinatorial method of studying the enzymatic hydrolysis of cellulose.

1.6.1 Research goals of Synthetic Biology

The emerging field seeks to utilise the design principles of electronic engineering and incorporate those within the framework of biological systems and traditional genetic engineering (Endy 2005). Namely this entails the detailed characterization and quantification of the behaviour of biological agents such as proteins, DNA, RNA and the relationships

between them. The end goal of which is the design of novel biological systems from defined parts and devices or the redesign of existing biological systems. Areas of development in this respect include the scaled construction of DNA from parts to devices, universal characterization standards, engineering of host chassis' and computer aided design (CAD) software for the design of novel biological systems (Cheng and Lu 2012). The motivation for the development of synthetic biology as an engineering discipline is most notably driven by technological innovations in nucleotide acid chemistry leading to falling costs in the synthesis and sequencing of DNA, allowing for greater control in manipulating and studying systems in molecular biology (Carlson 2009).

Moreover to realise the above goals and the reliable engineering of biology for predictable functionality, methods in bypassing emergent behaviour are needed. Unseen complexities of gene products and host interactions are undesirable for the engineering of biology in a reliable fashion. As such the design of genetic systems and components from the bottom up can contribute towards better understanding the interactions within and behaviours of the target biological system.

1.6.2 Applying engineering concepts to biological systems: a parts based approach

The principles of a parts based approach to synthetic biology research and the design of biological systems is illustrated in Figure 1.11. The hierarchy described is that of (i) DNA, (ii) parts, (iii) devices and (iv) systems. Manipulation of DNA sequences for the design of standardized genetic components or parts is achieved using methodologies and techniques of modern day molecular biology, described within Chapter 2 of this thesis. Resultant genetic parts are further characterized and studied so as to reliably predict their behaviour when composed into composite devices. Finally, systems made up of multiple devices may be designed with the intention of studying a biological phenomenon, such as the hydrolysis of lignocellulosic biomass, within a reliable and reproducible context.

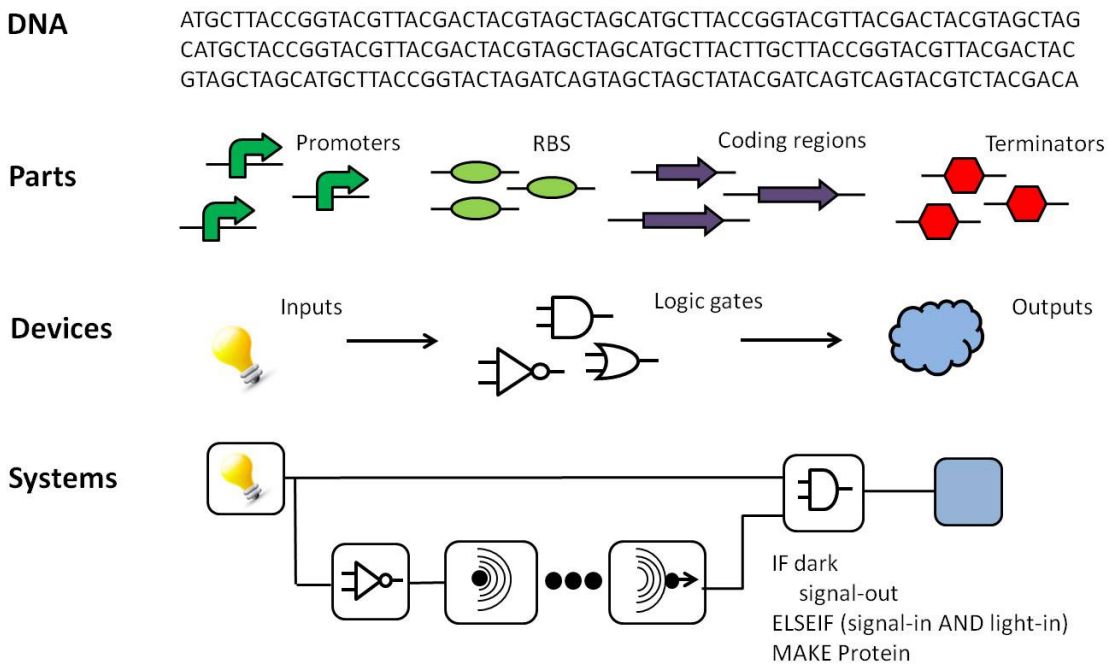

Figure 1.11: The principle design concepts within synthetic biology as illustrated according to a hierarchy of parts, devices and systems. The foundations to engineering biology via a parts based approach are centred upon the concepts of parts, devices and systems; (i) manipulation of DNA sequences facilitates the construction of standardized parts, (ii) comprehensive study and characterization of parts helps to inform the design of devices composed of multiple parts, (iii) systems or networks of devices with reliable functionality are built for a defined purpose.

The successful and productive development of a parts based approach to biological study is heavily reliant upon extensive characterization and understanding of components at the 'parts' level, on which the rest of the hierarchy is supported. As such, the need to develop standard assay protocols as well as reporting and sharing of parts is of paramount importance. Within the synthetic biology community, protocols and methodologies are largely shared through an online forum known as Openwetware (http://openwetware.org). The storage, distribution and sharing of physical DNA parts is co-ordinated by a central repository known as the Parts Registry (http://partsregistry.org). Here standardized parts all conforming to the 'BioBrick' format are deposited, to allow access to characterization data and distribution to the larger research community.

1.6.3 BioBricks: idempotent biological parts for standardized assembly of genetic devices

BioBricks are modular genetic parts developed by the research community for submission to an open source 'Registry of Standard Biological Parts' which are then freely distributed for the design and construction of purpose built genetic circuits within microbial hosts. The

power of this approach lies within the ability to combine parts in any order and in any number to rapidly generate novel and complex genetic circuits with relative ease (Knight 2003). Figure 1.12 illustrates the BioBrick 1.0 assembly standard.

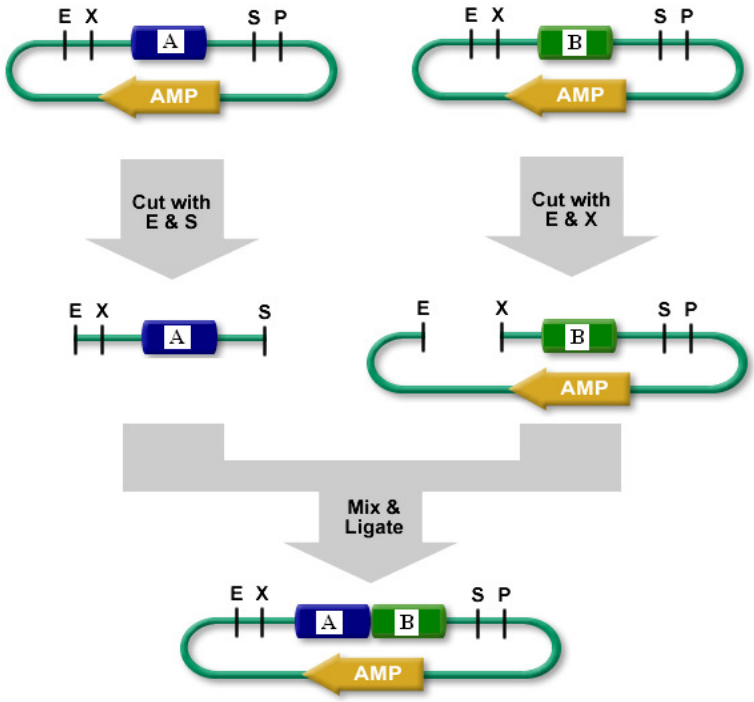

Figure 1.12: Illustration of the BioBrick 1.0 assembly standard and the protocol for combining modular parts. BioBrick parts are preceded upstream by EcoRI and XbaI restriction sites, termed the 'prefix'. The part is also succeeded downstream by SpeI and PstI restriction sites, termed the 'suffix'. BioBrick parts A or B may be assembled either upstream or downstream from one another depending on which combinations of prefix and suffix sites are cut. The example illustrated shows BioBrick part A being assembled upstream of part B on a single plasmid vector. Part A is cut with EcoRI and SpeI and part B with EcoRI and XbaI. The compatible sticky ends generated may be ligated upon mixing of the two parts. The final construct generated is also a BioBrick part and may undergo successive rounds of assembly. Antibiotic resistance encoded within the plasmid backbone offers a method for selection of the desired construct. Source: Registry of Standard Biological Parts (2012).

Once constructed, these components can be assembled in parallel so as to generate a range of configurations which may vary according to protein coding sequences, promoters, enhancers and secretion signals, amongst others (French 2009). With consideration to the present study, the BioBrick format lends itself well to the characterization and design of individual composite parts each encoding a single enzymatic activity for hydrolysis of a defined cellulosic substrate. The construction of modular devices is reliably achieved from individual

parts, and as such the assembly of a device encoding multiple enzymatic activities and the complete hydrolysis of a defined cellulosic substrate.

1.7 Overview of the remainder of the thesis

The remainder of this thesis is focused on the construction and characterization of a library of BioBrick parts, each encoding the expression of a single cellulase or hemicellulase cloned from *Cellulomonas fimi*. Parts are assayed for activity using a suite of colorimetric and fluorogenic substrates so as to confirm functional expression as well as quantify relative activities. In addition, characterization assays are performed for two promising expression hosts, namely *Escherichia coli* and *Citrobacter freundii*. These are carried out with an emphasis on the design of suitable media compositions for assaying growth on cellulosic substrates as well as suitable experimental conditions. The characterization of BioBrick parts and host chassis, allows for the informed and rational design of modular composite devices composed of multiple BioBrick parts. Devices are assembled with the purpose of encoding defined activities against one of three of the major polysaccharides present in plant biomass, namely cellulose, xylan or mannan. Following this, recombinant expression hosts transformed with the assembled constructs are assayed for growth in minimal media containing one of four well defined substrates as a sole source of carbon, including carboxymethyl cellulose, avicel, mannan from *Saccharomyces cerivisiae*, or beechwood xylan. Based on results from these growth assays inferences are made as to the suitability of each host chassis for the expression of recombinant cellulases, as well as the relative efficiencies of the enzymatic cocktail encoded by each genetic device against the respective carbon source. Further assays for those strains transformed with cellulase encoding constructs are conducted on cellulose paper. The ability of each strain to utilize the substrate is assayed for. Moreover, the extent of deconstruction of cellulose paper over the incubation period is documented so as to qualitatively compare the disruptive potential of each cellulase cocktail on the substrate. The main results from each chapter are summarized within the final discussion. Contributions to knowledge made by this work are also highlighted. Future amendments for improving experimental protocols and assays are suggested, with a particular consideration for those approaches which can potentially yield valuable insights into the mechanics of lignocellulose hydrolysis.

Chapter 2

Materials and Methods

2.1 Chemicals and reagents

Table 2.1 below summarizes the chemicals and reagents used within the present study.

Supplier	Reagent (product number)
New England Biolabs Inc.	Restriction enzymes; EcoRI, XbaI, SpeI, PstI
Melford Laboratories Ltd.	Ampicillin (A0104)
	Chloramphenicol (C0113)
	Kanamycin (K0126)
	Isopropyl β-D-1-thiogalactopyranoside (IPTG) (MB1008)
	5-bromo-4-chloro-3-indoyl-β-D-galactoside (X-gal) (MB1001)
	Agarose (MB1200)
Sigma-Aldrich Co.	Avicel (11365)
	Xylan (X4252)
	Mannan (M7504)
	Carboxymethyl cellulose (CMC) (419273)
	4-methylumbelliferyl β-D-cellobioside (MUC) (M6018)
	4-methylumbelliferyl β-D-glucopyranoside (MUG) (M3633)
	Magenta glycoside (B4527)
	Dimethyl sulphoxide (DMSO) (D5879)
	Congo Red (C6277)
	Yeast Extract (Y1625)
	Blotting paper, pure cellulose (P8046)
Thermo Scientific	Coomassie (Bradford) protein assay kit (1856209)
Invitrogen Corporation	SYBR-Safe DNA stain (S33012)

Table 2.1: Chemicals and reagents used within the present study, including names of suppliers.

2.2 Bacterial strains and expression vectors

Genomic DNA from *Cellulomonas fimi* ATCC484 was used as a source for all cellulase and hemicellulase encoding genes within this study. Subsequent DNA manipulation and cloning was carried out using *Escherichia coli* JM109. Expression hosts for growth assays include *E. coli* MG1655, *Citrobacter freundii* NCIMB11490 and *Citrobacter freundii* SBS197. *C. freundii* SBS197 is taken from the University of Edinburgh School of Biological Sciences teaching laboratory culture collection. Over expression of recombinant cellulases under the

T7 promoter system was performed in *E. coli* BL21(DE3). Construction and expression of BioBrick parts were carried out in pSB1A3, pSB1C3 and pSB1K3. In addition, low copy vectors including pSB4C5 and pSB4K5 were used for expression where stated. These are summarized below.

Plasmid vector	BioBrick Parts registry ID	Copy number per host cell	Encoded resistance
pSB1A3	pSB1A3-BBa_J04450	100-300	100 µg/ml Amp.
pSB1C3	pSB1C3-BBa_J04450	100-300	35 µg/ml CML.
pSB1K3	pSB1K3-BBa_J04450	100-300	50 µg/ml Kan.
pSB4C5	pSB4C5-BBa_J04450	~5	35 µg/ml CML.
pSB4K5	pSB4K5-BBa_J04450	~5	50 µg/ml Kan.

Table 2.2: List of plasmid vectors used for the construction and expression of BioBrick parts.

2.3 Cellulase and hemicellulase encoding genes from *Cellulomonas fimi*

Table 2.3 overleaf summarizes the lignocellulase encoding genes of *Cellulomonas fimi* which were cloned and characterized within the present study.

2.4 Media and culture conditions

Luria Broth was the chosen medium for routine culturing. M9 minimal medium supplemented with a carbon source and yeast extract (where stated) was used for growth assays. The compositions of each are shown below. All cultures were made up to a final volume of 5 ml in 1 oz glass vials and incubated at 37^{o}C on a rotary shaker at 200 rpm, unless otherwise stated.

Media	Composition
Luria Broth	10 g/l tryptone
	5 g/l yeast extract
	10 g/l NaCl
M9 minimal media	6 g/l Na_2HPO_4
	3 g/l KH_2PO_4
	0.5 g/l NaCl
	1 g/l NH_4Cl
	0.34 g/l thiamine
	493 mg/l $MgSO_4$
	14.7 mg/l $CaCl_2$

Table 2.4: Compositions of Luria Broth and M9 minimal media used for culturing of expression hosts.

In the preparation of growth assays, cultures were inoculated with a cell pellet spun down from overnight cultures in LB and resuspended in M9 so that the final OD_{600} is 0.1. All media used for growth assay experiments were inoculated in this way so as to reduce the amount of carbon carried over from the initial inoculums and maintain consistent levels of starter carbon present in M9 minimal media. Antibiotics were added according to the expression plasmid used to transform host strains (see Table 2.2), and cultures induced with 90 μg/ml of IPTG.

Gene name	Locus tag	Encoded protein function	E.C. code	Gene length (base pairs)	Protein weight (kDa)
**cex (xyn10A)*	Celf_1271	exo-1,4-β-glucanase	3.2.1.91	1,533	54.14
		endo-1,4-β-xylanase	3.2.1.8		
**cenA (cel6A)*	Celf_3184	endo-1,4-β-glucanase	3.2.1.4	1,350	46.71
cenB (cel9A)	Celf_0019	endo-1,4-β-glucanase	3.2.1.4	3,138	109.01
cenC (cel9B)	Celf_1537	endo-1,4-β-glucanase	3.2.1.4	3,306	115.23
cenD (cel5A)	Celf_1924	endo-1,4-β-glucanase	3.2.1.4	2,244	78.95
cbhA (cel6B)	Celf_1925	β-1,4-cellobiohydrolase	3.2.1.91	2,619	89.31
cbhB (cel48A)	Celf_3400	β-1,4-cellobiohydrolase	3.2.1.91	3,273	114.85
man26A	Celf_0862	endo-1,4-β-mannosidase	3.2.1.78	3,033	107.03
man2A	Celf_2770	exo-1,4-β-mannosidase	3.2.1.25	2,529	93.71
xynC (xyn10B)	Celf_0574	endo-1,4-β-xylanase	3.2.1.8	4,053	141.69
xynD (xyn11A)	Celf_0374	endo-1,4-β-xylanase	3.2.1.8	1,938	66.57
		xylan deacetylase	3.5.1.-		
cfx	Celf_3156	endo-1,4-β-xylanase	3.2.1.8	1,482	53.27
***bxyF*	Celf_1744	exo-1,4-β-xylanase	3.2.1.37	2,691	98.38
***xynF*	Celf_3155	exo-1,4-β-xylanase	3.2.1.37	1,503	53.01
cfbglu	Celf_2783	β-glucosidase	3.2.1.21	1,455	53.52
nag3A	Celf_2983	β-glucosidase	3.2.1.21	1,695	60.00
		3-β-N-Acetylglucosaminidase	3.2.1.52		

Table 2.3: Cellulase and hemicellulase encoding genes from *Cellulomonas fimi* cloned within the present study. *Denotes those genes cloned by Natasha Cain. **Denotes those genes cloned by Steven Kane, used herein with permission.

2.5 Congo Red endoglucanase activity assay

To assay for endoglucanase activity, the method described by Teather and Wood (1982) was employed, with some modifications. CMC was incorporated into molten LB agar to a final concentration of 0.2% w/v. CMC-agar was poured into plates and appropriate antibiotics and inducers added. Plates were inoculated with selected strains and incubated at 37^{o}C over 48 hours. CMC-plates were then stained with Congo Red dye by flooding with 5 ml of 0.5 mg/ml Congo Red solution for 45mins, at the end of which excess solution was removed. Excess unbound dye was washed from the plate by flooding with 5 ml 1 M NaCl solution for 45mins. Zones of clearing appeared around colonies positive for endoglucanase activity as neighbouring CMC is hydrolysed, liberating the Congo Red dye.

2.6 MUC exoglucanase activity assay

The β-1,4-glycosidic bond exhibited by 4-methylumbelliferyl β-D-cellobioside (MUC) is cleaved by exoglucanases to yield fluorescent 4-methylumbelliferyl under long-wave UV light (366nm). The assay used within the present study involves 100 µl of 5 mg/ml MUC solution spread on suitable LB agar plates and allowed to dry for 2 hours. Plates were inoculated and incubated at 37^{o}C overnight before being exposed to long-wave UV light to visualise fluorescence.

2.7 MUG β-glucosidase activity assay

Like MUC above, the β-1,4-glycosidic bond exhibited by 4-methylumbelliferyl β-D-glucopyranoside (MUG) is cleaved by β-glucosidases to yield a fluorescent product when illuminated under light at 366 nm. The MUG assay is identical to that above with the exception that plates are spread with 100 µl of 5 mg/ml MUG before inoculation and incubation at 37^{o}C.

2.8 *O*-Nitrophenyl assays

O-Nitrophenyl substrates including ONP-xylopyranoside, ONP-cellobioside and ONP-mannopyranoside were used to confirm the activity of exo-xylanases, exoglucanases and exo-mannanases, respectively. If active, the enzyme will catalyse the hydrolysis of the β-1,4 bond liberating the ONP product which yields a strong yellow colour, quantifiable at 420 nm. The following reaction mixture was set up in a microcentrifuge tube for each substrate to be tested.

Component	Volume (µl)
Cell lysate	100
25 mM ONP substrate	100
Phosphate buffered saline (pH7.4)	300
Total	500

Table 2.5: Composition of reaction mixtures used in confirming the activity of recombinant cellulases active on ONP-xyloside, ONP-cellobioside and ONP-mannoside.

After mixing, the reaction mixture is incubated at 37^{o}C for 3-24 h. The reaction was effectively stopped by placing the microcentrifuge tube on ice and a colour change is initiated by the addition of 500 µl 1 M sodium carbonate solution. The total volume was transferred to a cuvette and absorbance is then read at 420 nm and recorded. Absorbance was measured on a Pharmacia LKB Ultraspec III spectrophotometer. Any sample reading an absorbance greater than 1 unit was diluted 10-fold and re-measured for accuracy.

2.9 Coomassie protein assay

Total protein concentration of cultures was determined by a Coomassie protein assay kit. 100 µl of culture is added to 900 µl of reagent. The mixture is vortexed and incubated at 65^{o}C for 60mins. The mixture is vortexed again and left at room temperature overnight for the blue colour to develop. Absorbance is measured at 595 nm and recorded. In instances were cultures include a carbon source high in particulate matter (i.e. avicel and xylan), optical density at 530 nm is also measured and subtracted from that at 595 nm to exclude absorbance generated by turbidity. Absorbance is measured on a Pharmacia LKB Ultraspec III spectrophotometer.

2.10 DNA manipulation techniques

2.10.1 PCR conditions

In each instance KOD Hot Start DNA polymerase was used for both cloning of target genes and mutagenic PCR for the generation of silent mutations. PCR reaction mixtures were prepared according to the manufacturer's specifications and are shown in Table 2.6. All subsequent PCR cloning steps were run using a PTC-200 DNA Engine Thermal Cycler from Bio-Rad and followed manufacturer's specifications. Reaction conditions and cycles are detailed in Table 2.7 overleaf.

Reagents	Volume (µl)
Sterile water	32.0*
10x reaction buffer	5.0
25 mM MgS04	3.0
Forward primer (10 pmol/µl)	1.5
Reverse primer (10 pmol/µl)	1.5
2 mM dNTP mix	5.0
Template (cell suspension or supercoiled plasmid DNA)	1.0
KOD Hot start DNA Polymerase	1.0
Total	**50.0**

Table 2.6: Reaction mixture for PCR using KOD Hot Start DNA polymerase. *Indicates that when GC-rich DNA was used as a template, 10 µl of water was substituted with 10 µl 50% v/v glycerol as a means to reduce the formation of secondary structures and unwanted annealing events.

	Reaction conditions		
Reaction Step	**Temperature (°C)**	**Time (s)**	**Cycles**
1. Initial denaturation	95	120	1
2a. Denaturation	95	20*	30
2b. Annealing	Dependent on primer T_m	10	
2c. Extension	70	10 s/kb for <500 bp, 15 s/kb for 500-1000 bp, 20 s/kb for 1000-3000 bp, 25 s/kb for >3000 bp	
3. Final extension (chase)	70	600	1
4. Hold	4	~	1

Table 2.7: Programme used for PCR cloning and mutagenesis. *Indicates that when GC-rich DNA was used as a template, the denaturation time was extended to 60 seconds so as to ensure complete denaturation of complementary strands.

2.10.2 MABEL: Mutagenesis with Blunt End Ligation

Unwanted endonuclease restriction sites within protein coding genes cloned from genomic *C. fimi* DNA were removed by single-point base pair mutation via mutagenic PCR. In each instance, cloned genes were digested with EcoRI and SpeI endonucleases and ligated into pSB1A3, C3 or K3 plasmid vectors. Mutagenic PCR was carried out to generate silent mutations by altering the sequence [ctgcag] recognised by PstI to that of [ct**c**cag] or [ctgc**t**g]. Primers used are listed in Table 2.16. PCR products generated were then blunt-end ligated according to the protocol described in section 2.10.6 below with the alteration of adding 1 µl T4 DNA polynucleotide Kinase.

2.10.3 DNA purification

Purification of DNA from restriction digests or agarose gel was performed using glass silica beads. To 1 unit of DNA solution, 3 units of 6M NaI are added followed by 5 µl of glass bead suspension. The suspension was mixed and incubated on ice for 10 mins before spinning down of the glass beads. The supernatant is removed and discarded. The glass bead pellet is resuspended and washed in 250 µl ice-cold wash buffer (Table 2.9) three times. After a final spin, the supernatant is removed and discarded and the pellet resuspended in 15 µl elution buffer (Table 2.8) and incubated at 55^{o}C for 10mins. A final spin to pellet the glass beads is performed and the DNA containing supernatant is transferred to a clean microcentrifuge tube.

Component	Volume or mass
Tris Base	53 mg
Tris HCl	88 mg
Deionised H_20	100 ml
Total volume	**100 ml**

Table 2.8: Composition of Elution Buffer (EB) used for DNA purification and storage. Total solution is autoclaved before use and stored at room temperature.

Component	Volume or mass
Tris base	24 mg
Tris HCl	127 mg
NaCl	293 mg
Na_2EDTA	93 mg
Ethanol	50 ml
Deionised water	50 ml
Total volume	**100 ml**

Table 2.9: Composition of wash buffer used in DNA purification. Final solution is stored at -20^{o}C.

2.10.4 Agarose gel electrophoresis

All agarose gel electrophoresis steps were carried out using a Bio-Rad minisub cell using 0.8% agarose gels in 0.5x TAE buffer (Table 2.10) at 100 V and 50 mA for the duration of 40-50 mins. Gels were stained using one of three solutions in distilled water; ethidium bromide, SYBR Safe or Gel Green. Visualization of gels was under UV light. A 1 kb DNA ladder was used as a marker.

Component	Volume or mass
Tris base	48.40 g
Sodium EDTA	3.73 g
Deionised water	450 ml
Glacial acetic acid	11.40 ml
Total	**500 ml**

Table 2.10: Composition of 20x TAE buffer used for the preparation of agarose gels and running of agarose gel electrophoresis. 20x TAE is diluted to 0.5x TAE using distilled water.

2.10.5 Restriction digests

Digestion of BioBrick parts using endonucleases from New England Biolabs and Promega were carried out according to the manufacturer's instructions. The composition of the digest reaction mixtures are shown in Table 2.11. Reaction mixtures were mixed and incubated at 37^{o}C for 30-60 mins. Prior to ligation, restriction digests were either purified by glass beads DNA purification or denatured at 80^{o}C for 20mins, according to manufacturer's instructions.

Component	Volume (µl)
Supercoiled plasmid DNA or PCR product	2.0
Sterile dH_20	23.5
Buffer	3.0
Endonuclease(s)	1.0
BSA	0.5
Total	30.0

Table 2.11: Composition of restriction digest reactions used for the construction of BioBrick parts.

2.10.6 Ligations

Sticky-ended and blunt-ended ligations were carried out using T4 DNA ligase from Promega. According to manufacturer's instructions, the reaction mix was set up according to Table 2.12 below. Ligation reaction mixtures were mixed and incubated at 25^{o}C for 30 mins, followed by incubation at 16^{o}C for 15 h. Ligation mixtures were denaturated at 65^{o}C for 20 mins prior to transformation of competent cells.

Component	Volume (ul)
Total DNA solution to ligate	2.0
Sterile dH_20	20.0
T4 DNA ligase	0.5
T4 DNA ligase buffer	2.5
Total volume	**25.0**

Table 2.12: Composition of ligation reactions used for the ligation of BioBrick parts and blunt-end ligation of PCR mutagenesis products.

2.10.7 Preparation and transformation of chemically competent cells

For the preparation and transformation of chemically competent cells, the protocol presented by Chung *et al.* (1989) was followed. Competent cells were prepared by growing a cell culture in 50 ml LB to an OD_{600} of between 0.4 and 0.5. Cells were then incubated on ice for 5 mins before spinning down at 13,000 rpm for 15 mins. The supernatant was removed and the pellet resuspended in 6 ml ice-cold Transformation and storage solution (TSS). Table 2.13 below details the composition of TSS. Aliquots of 200 µl of cells are transferred to 1.5 ml microcentrifuge tubes to rest on ice for 30 mins before storage at -80^{o}C.

Component	Volume (ml)
Luria Broth	17
40% w/v PEG 3350	5
1M $MgCl_2$	1
DMSO	1
Total	24

Table 2.13: Composition of Transformation and Storage Solution (TSS) used for the preparation of competent *E. coli* or *Citrobacter freundii* for transformation.

Competent cells were transformed by the addition of 1 µl supercoiled plasmid DNA or 10 µl ligation mix to 100 µl aliquots of previously prepared cells. The mixture was incubated on ice for 30-45 mins before heat shock at 42^{o}C for 90 seconds, followed by incubation on ice for 90 seconds. Following this, 900 µl of LB (warmed to 37^{o}C) is added to the cells and incubated at 37^{o}C for 60 mins to recover. Cells transformed with either kanamycin or chloramphenicol resistance encoding plasmids are allowed to recover for 2 h. After recovery, cells centrifuged and the pellet resuspended in 200 µl LB for plating on LB agar selection plates containing appropriate antibiotics.

2.10.8 Plasmid DNA extraction: Minipreps

For the extraction of plasmid DNA from *E. coli* JM109, the alkaline lysis protocol described by Sambrook *et al.* (1989) was employed. DNA is stored in 40 µl Elution buffer at -20^{o}C.

2.10.9 DNA sequencing

DNA sequencing reactions were performed on an ABI 3730 capillary Sanger sequencer maintained and operated by the Gene Pool sequencing service within the University of Edinburgh School of Biological Sciences. Reactions were set up as shown in Table 2.14.

Component	Volume (µl)
DNA template for sequencing	1
5 pmol ssDNA primer	1
Sterile dH_20	4
Total volume	**6**

Table 2.14: Composition of sequencing reaction and total volume used for Sanger sequencing.

2.11 List of Primers

2.11.1 Primers for cloning of protein encoding genes from *C. fimi* genomic DNA

Table 2.15 details the primers used for cloning of *C. fimi* genomic DNA. Biobrick tail sequences non-complementary to the coding sequence are capitalized. Biobrick restriction sites are underlined; EcoRI and XbaI in the forward primer, SpeI and PstI in the reverse primer.

2.11.2 Primers for silent point mutations to remove PstI sites within coding sequences

Exhibiting a high GC content of 74%, genes cloned from *C. fimi* DNA frequently contain multiple PstI restriction sites. These are removed by mutagenic PCR through the generation of silent single base pair mutations. Table 2.16 details the primers used in mutagenic PCR. Single bases changed by point mutation are capitalized. In each instance the sequence [ct**g**cag] recognised by PstI is changed to [ct**c**cag] or [ctgc**t**g] in the case of *nag3A* mut2.

Primer name	Primer sequence (5’ to 3’)
cenB clonf	TCTGAATTCCTTCTAGatgctccgccaagtcccacgc
cenB clonr	TCTCTGCAGCTACTAGTATTATTAgccgcagacctcaccgtg
cenC clonf	TCTGAATTCCTTCTAGatggtttctcgcaggtcatca
cenC clonr	TCTCTGCAGCTACTAGTATTATTAgctgcgcggacgctgcac
cenD clonf	TCTGAATTCCTTCTAGatgcactccgcatcgcgcacc
cenD clonr	TCTCTGCAGCTACTAGTATTATTAgcgcgtcgtgcaggtcgc
cbhA clonf	TCTGAATTCCTTCTAGatgtccacactcggcaagcga
cbhA clonr	TCTCTGCAGCTACTAGTATTATTAgccgagcgtgcaggcgac
cbhB clonf	TCTGAATTCCTTCTAGatgtcgtcaacgacccgccgg
cbhB clonr	TCTCTGCAGCTACTAGTATTATTAcgtgcaggacgcgccgtt
man2A clonf	TCTGAATTCCTTCTAGatgatcacccaggacctc
man2A clonr	TCTCTGCAGCTACTAGTATTATTAgcgggactgctggagatc
man26A clonf	TCTGAATTCCTTCTAGatgacgaaccgcagccgt
man26A clonr	TCTCTGCAGCTACTAGTATTATTAgcgcagctcgacgtcgtc
cfx clonf	TCTGAATTCCTTCTAGatgcacacgaaactccac
cfx clonr	TCTCTGCAGCTACTAGTATTATTAggatgccgcgcacgtgac
xynC clonf	TCTGAATTCCTTCTAGatgtacctcgacggcggc
xynC clonr	TCTCTGCAGCTACTAGTATTATTAcgagaagacgggccgcac
xynD clonf	TCTGAATTCCTTCTAGatgtccgacagtttcgaa
xynD clonr	TCTCTGCAGCTACTAGTATTATTAgcccgtggcgcacgtagc
nag3A clonf	TCTGAATTCCTTCTAGatgatcgacctgaccgca
nag3A clonr	TCTCTGCAGCTACTAGTATTATTAtcacaggtgggtgtccca
cfbglu clonf	TCTGAATTCCTTCTAGatgggcgaccggttccag
cfbglu clonr	TCTCTGCAGCTACTAGTATTATTAtcagggctggtaggtcgc

Table 2.15: List of ssDNA primer sequences including BioBrick tails used for the cloning of cellulase and hemicellulase encoding genes from genomic *C. fimi* DNA.

Primer name	Primer sequence (5’ to 3’)
cenB mut1f	Ccagaagtcgatgttct
cenB mut1r	agggcctcggcgtagtt
cenB mut2f	Ccagccgggccagagca
cenB mut2r	agggtggcgttccacga
cenC mut1f	Ccagacgtacgagcggg
cenC mut1r	aggagctgcccgaccgc
cenD mut1f	Ccagatcttcgagtact
cenD mut1r	aggctgttcttgccctc
cbhB mut1f	Ccagggtccaacggcg
cbhB mut1r	agccacgtgtagaactc
cbhB mut2f	Ccaggcgggcgtcgtca
cbhB mut2r	aggcccgtcgggaccgt
man2A mut1f	Ccagaccgcgggcctgtg
man2A mut1r	aggtcggggccccagtc
man2A mut2f	Ccaggacggctaccgcgtc
man2A mut2r	agcggcgacacggtcgc
man26A mut1f	Ccagctcaacgccggtggt
man26A mut1r	aggaccatcctgttgtt
man26A mut2f	Ccaggtcggctcgacctgg
man26A mut2r	agggccgccttggtgtc
cfx mut1f	Ccagcgcaccggcaacgac
cfx mut1r	aggttcgagtcgcgccg
cfx mut2f	Ccagaacttcgccgacctc
cfx mut2r	agggtcgtgtggtagttc
xynC mut1f	Ccagcccgtctcgcagatg
xynC mut1r	agcaggttgacgtggaa
xynC mut2f	Ccaggccgtcaccgagctc
xynC mut2r	agcaggccggtcgtcgc
xynC mut3f	Ccaggccaagccggccttc
xynC mut3r	aggtcgtcgtcgaacgg
xynD mut1f	Ccaggacctgcgctcgcgc
xynD mut1r	aggatgccgggcagcgc
nag3A mut1f	Ccagagcaccgccacagtc
nag3A mut1r	agcgtggtcaccgtggcg
nag3A mut2f	Tggcgtgaagaccgtca
nag3A mut2r	gcagcgatggcggcgcg
nag3A mut3f	Ccagaacggcggcagcccc
nag3A mut3r	aggtcgtagacgagcacc
cfbglu mut1f	Ccaggcctaccggttctcg
cfbglu mut1r	aggccgaggctcttcatg

Table 2.16: List of ssDNA primer sequences used for the generation of silent point mutations to remove internal PstI restriction sites within cloned *C. fimi* DNA. Multiple mutations within a single gene are listed as *mut1, 2 or 3*. Single bases modified by PCR are capitalized.

Chapter 3

Characterization of host chassis and BioBrick parts

Summary

Using the large body of literature surrounding the cellulases of *Cellulomonas fimi*, we constructed a library of standardized BioBrick parts, each encoding the expression of a single cellulase or hemicellulase with defined activities against cellulose, mannan or xylan. Further characterization studies of these parts were performed so as to confirm enzymatic activity as well as contribute towards scientific knowledge. Moreover, characterization of suitable expressions hosts namely *Escherichia coli* and *Citrobacter freundii*, aided in the design and planning of future growth assays as well as the design of a suitable M9 medium composition for growth on cellulosic substrates. Both expression hosts are enterobacteria, chosen for ease of genetic manipulation and cultivation in a laboratory environment in the case of *E. coli*, or chosen for the ability to utilize cellobiose and effectively secrete extracellular proteins in the case of *C. freundii*. It was shown within our assays that *C. freundii* is capable of growth in saline media consisting of 0.6 M sodium chloride, and as such is a promising host for the development of bioprocessing projects utilizing sea water as a readily available medium. Knowledge gained from previous published results as well as the additional characterization assays performed here, informed the design and construction of composite devices made up of multiple cellulase encoding parts. Nineteen devices in total were assembled with defined activities against mannan, xylan or cellulose. Each device is described here according to the Synthetic Biology Open Language, a graphical annotation format for representing synthetic gene constructs.

3.1 Introduction

To better understand the complementarities and synergies of the cellulases exhibited by *C. fimi* and attempt to recapitulate the cellulolytic system in a recombinant host, I built a library of standardised genetic parts. Using the BioBrick 1.0 assembly method devised by Knight (2003), fourteen parts encoding a defined activity against cellulose, mannan or xylan were constructed. These are summarized in Table 2.3 (see Chapter 2). Each part was cloned in standard BioBrick expression vectors; pSB1A3, pSB1C3 or pSB1K3. Expression of each protein encoding part was put under the control of one of two IPTG inducible promoters from the Parts Registry; $P_{spac\text{-}lacI}$ (BBa_K174004) or $P_{lac\text{-}lacZ}$ (BBa_J33207). Translation was initiated by a strong *E. coli* ribosome binding site (BBa_J15001).

Constructing each part in this manner allowed for the parallel testing and combining of multiple part arrangements into devices. This strategy lends itself well to the study of lignocellulose hydrolysis where multiple enzymatic functions are required for complete deconstruction and subsequent utilization of the substrate. This chapter discusses the considerations made in choosing a suitable expression host and the characterization of the cellulases and hemicellulases cloned from *C. fimi*.

3.2 Considerations for expression hosts

3.2.1 Growth of *Escherichia coli* and *Citrobacter freundii* on cellobiose

For the efficient deconstruction of cellulose and the utilization of liberated glucose and cellobiose, a suitable expression host would need to exhibit efficient extracellular protein secretion as well as the ability to assimilate glucose and cellobiose as a carbon source. For the purposes of this thesis we considered the use of two enterobacteriaceae; *Citrobacter freundii* and *Escherichia coli.*

Since both bacteria are closely related and use the same processing sites for transcription and translation, we were able to leverage a single library of BioBrick parts for testing across both species. Although *E. coli* is known to exhibit non-specific leakage of *C. fimi* cellulases (Guo *et al.* 1988), we inferred that this level of extracellular protein production would not be sufficient to support growth on a crystalline cellulose substrate, a hypothesis which is later tested in Chapters 4 and 5 of this thesis. As such, *Citrobacter freundii* was chosen for its type II secretion system which could afford a greater level of cellulase secretion and subsequent deconstruction of the substrate. Moreover, *Citrobacter freundii* being a common commensal

in the intestinal tracts of ruminants, the bacterium exhibits a native β-glucosidase capable of cleaving cellobiose into two glucose monomers. Figure 3.1 below shows the growth of *E. coli* strains MG1655 and JM109 as well as *Citrobacter freundii* NCIMB11490 and SBS197 in M9 media containing 1% w/v cellobiose. Growth in M9 medium containing 1% w/v glucose was used as a control in each instance. Cultures were performed in duplicate.

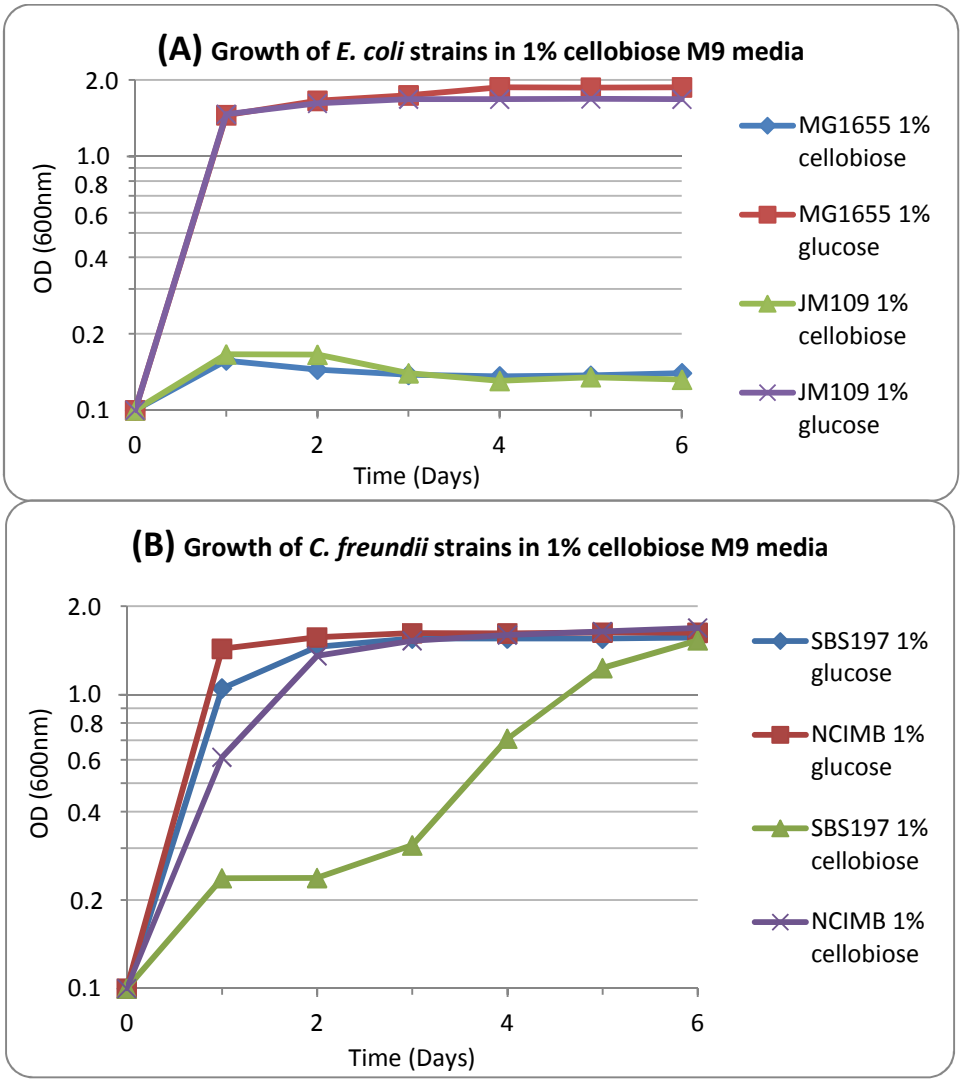

Figure 3.1: Growth of *E. coli* and *Citrobacter freundii* strains in 1% w/v cellobiose M9 medium. Growth curves of *E. coli* MG1655, *E. coli* JM109, *C. freundii* NCIMB11490 and *C. freundii* SBS197 cultured in M9 media containing 0.34 g/l thiamine and either 1% w/v cellobiose or 1% w/v glucose. *E. coli* MG1655 and JM109 are shown to be unable to grow on cellobiose. Positive controls in glucose reach an OD_{600} of ~1.75 after 6 days of culturing **(A).** *Citrobacter freundii* strains NCIMB11490 and SBS197 are able to utilize cellobiose though notably not as efficiently as glucose. *C. freundii* NCIMB11490 shows the shortest lag phase compared to that of SBS197 which only reaches stationary phase at day 6 **(B). Note:** OD_{600} results above 1 are considered unreliable as samples were not diluted before readings were recorded.

Figure 3.1 identifies wildtype *Citrobacter freundii*, notably NCIMB11490, as a promising host for the expression of recombinant cellulases as it is able to utilize cellobiose at a comparable rate to that of glucose, whereas both *E. coli* MG1655 and JM109 are entirely unable to utilize cellobiose and would require the expression of a recombinant β-glucosidase in addition to endo- and exo-glucanases for complete cellulose hydrolysis and utilization.

3.2.2 Growth of expression hosts in saline media

Another interesting and relevant aspect to modifying *Citrobacter freundii* for the deconstruction of plant biomass is that the bacterium is able to grow in media with a salinity equivalent to that of sea water.

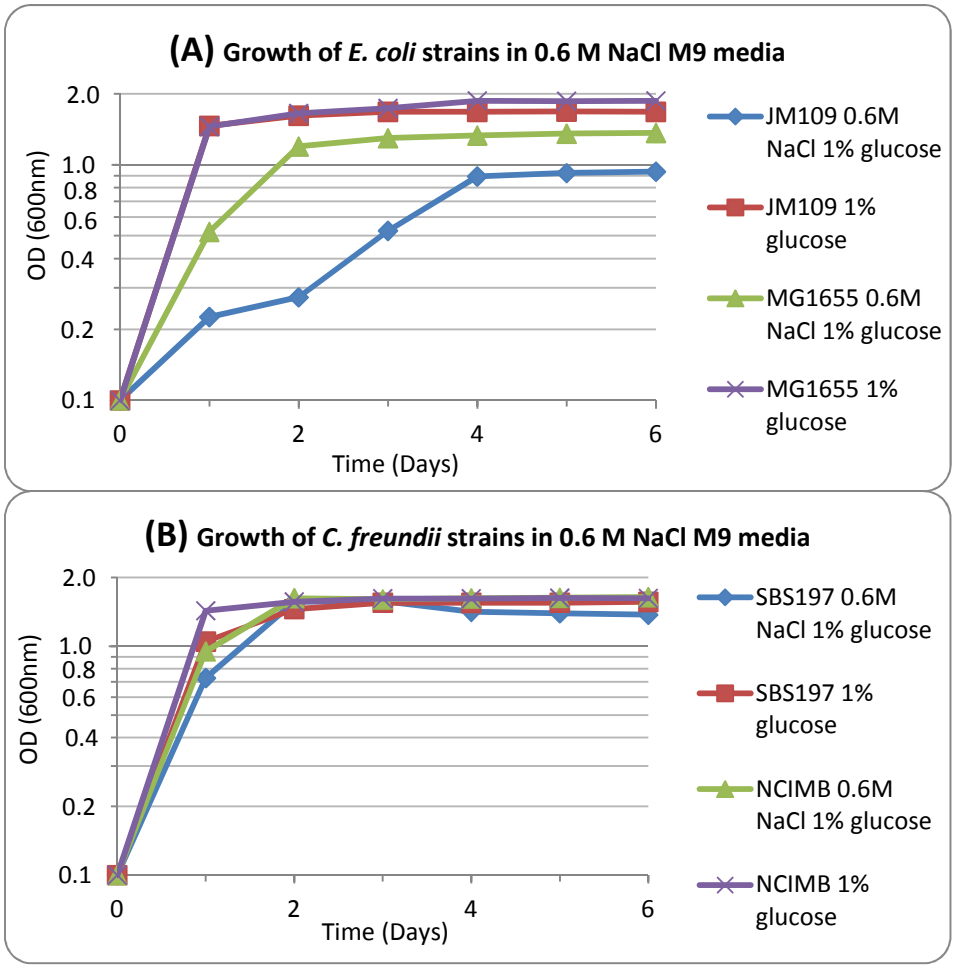

Figure 3.2: Growth of *Escherichia coli* and *Citrobacter freundii* strains in 0.6 M NaCl M9 media with 1% w/v glucose. *E. coli* strains MG1655 and JM109 grown in M9 0.6 M NaCl media show a marked reduction in cell density compared to positive controls grown in M9 media **(A).** By comparison *Citrobacter freundii* NCIMB11490 and SBS197 are able to grow in 0.6 M NaCl M9 medium at a comparable rate to that of the positive controls grown in M9 medium without NaCl **(B).** All media contained 0.34 g/l thiamine and 1% w/v glucose. **Note:** OD_{600} results above 1 are considered unreliable as samples were not diluted before readings were recorded.

To replicate the saline condition of sea water, which has a salt content of 3.5%, and test each strain for the ability to grow in saline media, strains were inoculated into M9 medium supplemented with 1% w/v glucose and 0.6 M sodium chloride. *E. coli* MG1655, JM109, and *Citrobacter freundii* NCIMB11490 and SBS197 were grown in 0.6 M NaCl M9 media over the course of 6 days. Cultures were performed in duplicate. Results in Figure 3.2A identify that both *E. coli* MG1655 and JM109 are not well suited to growth in the saline media as a marked difference in cell density is observed compared to the control. By comparison, Figure 3.2B shows that both *C. freundii* NCIMB11490 and SBS197 are able to tolerate 0.6 M NaCl as growth is comparable to that of the controls lacking sodium chloride. These results suggest that *C. freundii* is a suitable host for performing cultures in saline conditions.

3.2.3 Addition of a starter carbon source and loss of β-glucosidase activity

Before attempting any growth assays of recombinant hosts on cellulosic substrates, preliminary growth curves of *E. coli* and *C. freundii* strains in M9 media containing varying amounts of glucose as a starter carbon were performed. These results were intended to inform the design of a medium composition suitable for assaying growth where carbon is the limiting factor. As such, M9 medium containing 2% w/v cellobiose with either no added glucose or 0.1% w/v glucose or 0.2% w/v glucose was prepared. Each medium was inoculated with one of three strains; wildtype *C. freundii* NCIMB11490 or SBS197, or *E. coli* MG1655 expressing the recombinant β-glucosidase Cfbglu. Cultures were performed in triplicate. Results are shown in Figure 3.3.

Figure 3.3A shows that *E. coli* MG1655 P_{lac}-*cfbglu* reached the highest cell density with an OD of ~1.6 in M9 medium containing only cellobiose with no added glucose, compared to an OD of ~1.0 on those media supplemented with glucose. A possible explanation is that in media supplemented with glucose, there is a reduced selection pressure for the growth of those cells expressing the β-glucosidase since carbon is readily available in the form of glucose. As such cells harbouring a plasmid where *cfbglu* is excised by a recombination event outcompete those where Cfbglu is actively expressed; however the limited carbon pool (in the form of glucose) limits the maximal cell density achieved within the culture. This hypothesis could be tested by streaking out colonies and performing colony PCR to confirm the presence of the *cfbglu* gene, or plating colonies to check for CML resistance and cellobiase activity.

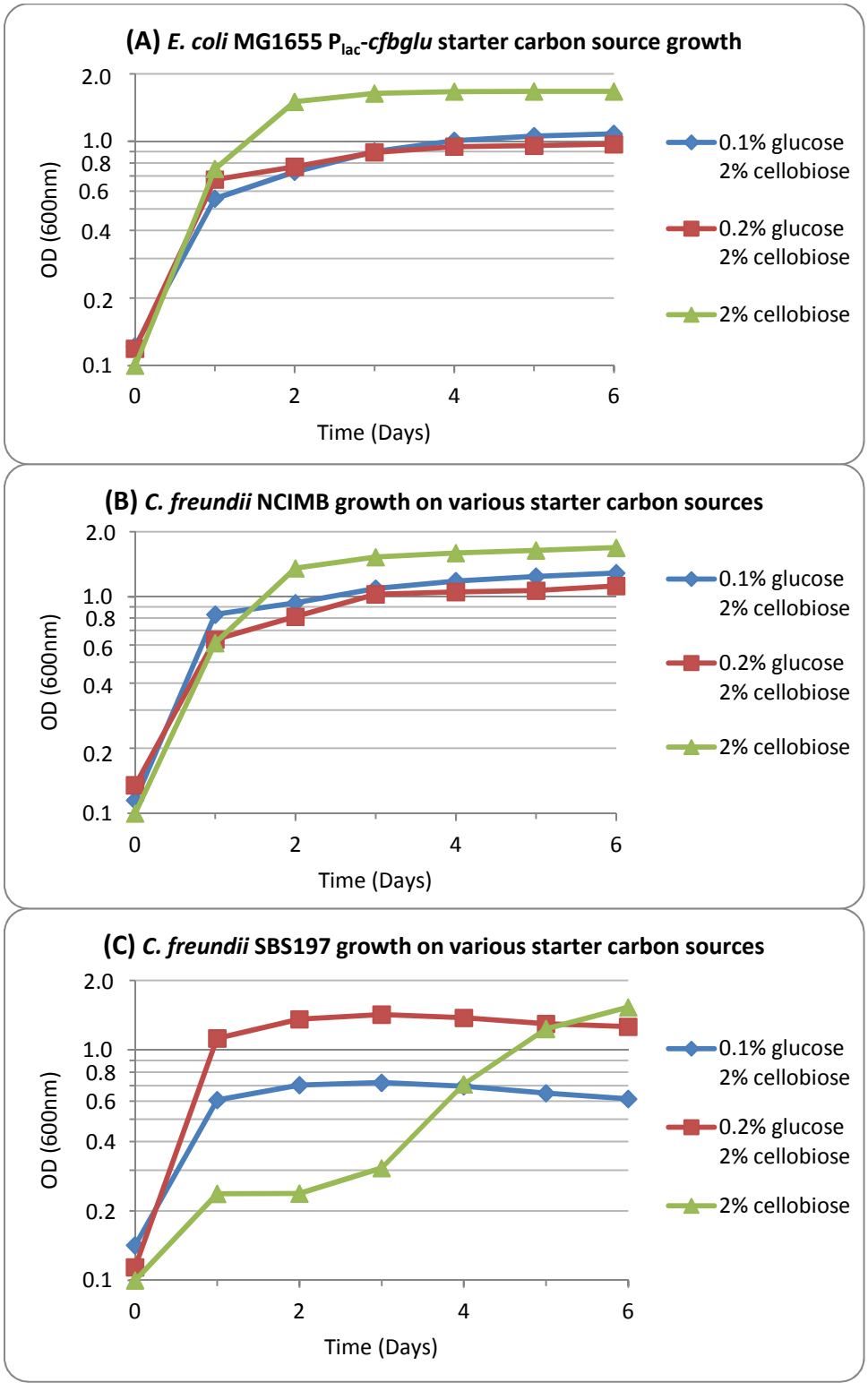

Figure 3.3: Growth of *E. coli* and *Citrobacter freundii* in M9 media containing varying amounts of glucose as a starter carbon source. Growth curves of *E. coli* MG1655 expressing a recombinant β-glucosidase (Cfbglu) grown in 2% w/v cellobiose M9 medium with no glucose or 0.1% w/v glucose or 0.2% w/v glucose **(A).** Growth curves of *C. freundii* NCIMB11490 and SBS197 grown in 2% w/v cellobiose M9 medium with no glucose or 0.1% or 0.2% w/v glucose **(B,C).** All media contained 0.34 g/l thiamine. **Note:** OD_{600} results above 1 are considered unreliable as samples were not diluted before readings were recorded.

Figure 3.3B illustrates a similar scenario. In the case of *C. freundii* NCIMB11490, which is shown to assimilate cellobiose, growth is observed to be best in media containing only cellobiose where a maximal absorbance of ~1.6 is measured compared to ~1.2 in media supplemented with glucose. There is no evidence within the literature suggesting how *C. freundii* regulates expression of β-glucosidase activity, though it would seem logical for expression to be repressed in the presence of glucose. These results are consistent with this hypothesis. Similarly, *C. freundii* SBS197 also reaches a maximal cell density in media containing only cellobiose (Figure 3.3C). Notably though, a longer lag phase is evident, suggesting that β-glucosidase expression is more tightly regulated or perhaps an additional starter carbon source is required for sufficient growth. Another point to make is that cell density drops from 1.4 at Day 3 to 1.2 at Day 6 in media containing 0.2% glucose. This observation has been noted amongst other members within the lab working with *C. freundii* SBS197 (data not shown).

Results presented here show that media best suited for measuring growth where carbon is limited should contain a minimal amount of starter carbon, in the cases of *E. coli* MG1655 and *C. freundii* NCIMB11490. Conversely, the *C. freundii* SBS197 strain requires additional starter carbon to support initial growth before utilization of cellobiose is possible. As such, a starter carbon source such as yeast extract would be suitable for these purposes since it contains a mix of small peptides, amino acids and trace elements offering itself as a good carbon source without significant repression of β-glucosidase expression within *C. freundii*.

3.3 Assaying for activity of β-glucosidases

3.3.1 Qualitative assays for β-glucosidase activity

According to the literature, *C. fimi* expresses two β-glucosidases; Cfbglu which is induced by cellobiose and cellulose, and Nag3A which is constitutively expressed regardless of carbon source (Wakarchuk *et al.* 1984; Kim and Pack 1989). Though both enzymes are able to hydrolyse the β-1,4-glycosidic bond, only Cfbglu is a true cellobiase capable of hydrolysing cellobiose into glucose. To confirm these activities, 4-methylumbelliferyl β-D-glucopyranoside (MUG) was used as a substrate; in the presence of β-glucosidases, the β-1,4 bond present in MUG is cleaved liberating methylumberliferone which is fluorescent when excited at 366 nm. Figure 3.4A shows an LB agar plate containing 1.5 mM MUG, with 30 µl of cell lysate from an *E. coli* pSB1C3 vector control strain, *E. coli* P_{lac}-*cfbglu* or *E. coli* P_{lac}-

nag3A added to wells within the plate. Figure 3.4B shows an identical plate spotted with 10 µl of overnight culture of the previous strains.

These results identify both enzymes as being able to hydrolyse the β-1,4 bond present in MUG, as fluorescence is observed for both Nag3A and Cfbglu. Notably, a larger zone of fluorescence is observed in Figure 3.4A where 30 µl of cell lysate is added due to free diffusion of the β-glucosidases which are both cytoplasmic proteins when expressed by *C. fimi*.

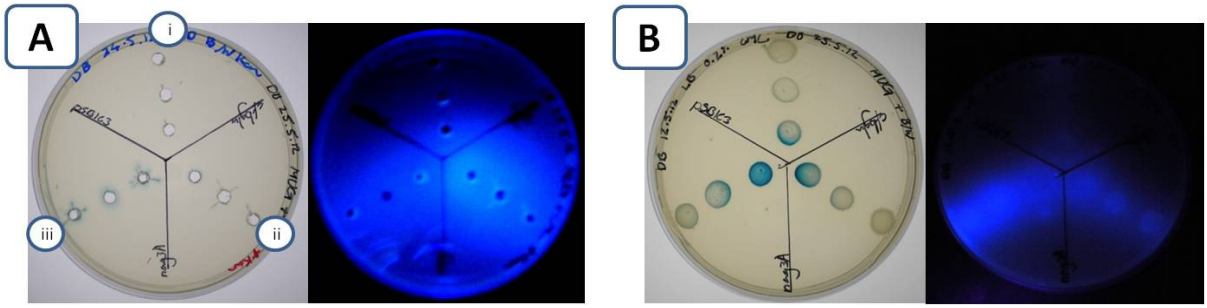

Figure 3.4: Assaying for the activity of *C. fimi* β-glucosidases expressed by recombinant *E. coli*. LB agar plates containing 1.5 mM 4-methylumbelliferyl β-D-glucopyranoside (MUG) are inoculated with *E. coli* MG1655 pSB1C3 vector control **(i),** P_{lac}-*cfbglu* **(ii)** and P_{lac}-*nag3A* **(iii)** strains. 30 µl cell lysate is added to wells within the plate **(A),** or 10 µl of overnight culture is spotted directly onto the agar surface **(B).** Fluorescence is observed when plates are illuminated at 366 nm.

To compare β-glucosidase activity of *Citrobacter freundii* NCIMB11490 to that of the above strains, 10 µl of overnight culture was spotted onto a LB agar plate containing MUG. Fluorescence is observed for each strain except the negative vector control, as expected (Figure 3.5). As a purely qualitative assay, the results confirm β-glucosidase activity for *E. coli* MG1655 P_{lac}-*cfbglu* and P_{lac}-*nag3A* strains as well as *Citrobacter freundii*.

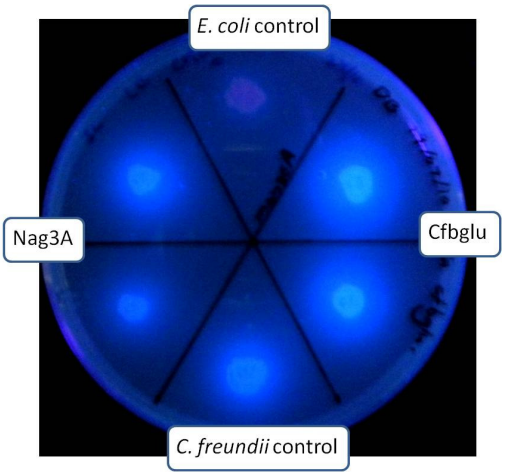

Figure 3.5: Comparing relative β-glucosidase activity of recombinant Cfbglu and Nag3A to that of *Citrobacter freundii* NCIMB11490. An LB agar plate containing 1.5 mM 4-methylumbelliferyl β-D-glucopyranoside (MUG) is inoculated with 10 µl of overnight culture of *E. coli* MG1655 pSB1C3 vector control, P_{lac}-*cfbglu*, P_{lac}-*nag3A*, and *Citrobacter freundii* NCIMB11490. Fluorescence is observed when the plate is illuminated at 366 nm.

3.3.2 Comparing growth of recombinant *E. coli* JM109 and MG1655 on cellobiose

To effectively determine and confirm the activity of Cfbglu as a true cellobiase and Nag3A as a non-cellobiose active β-glucosidase, *E. coli* MG1655 and JM109 strains expressing either Cfbglu or Nag3A were grown in M9 medium containing 1% w/v cellobiose as a sole carbon source. In each instance the positive control included a strain harbouring the pSB1C3 vector grown in 1% w/v glucose medium, and the negative control included the same pSB1C3 vector strain grown in 1% w/v cellobiose medium. Results are shown in Figure 3.6 below.

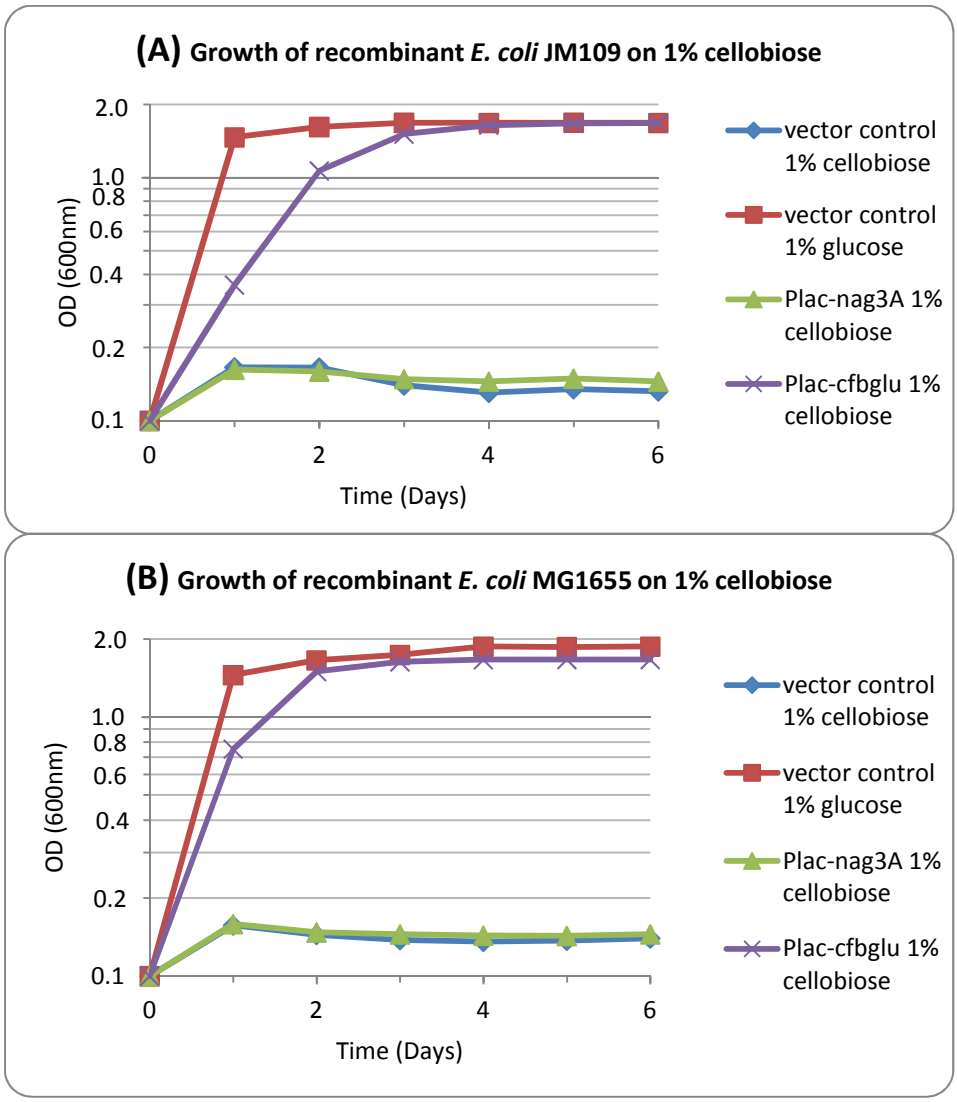

Figure 3.6: Growth of recombinant *E. coli* MG1655 and JM109 strains on cellobiose. *E. coli* JM109 **(A)** and MG1655 **(B)** strains expressing either Cfbglu or Nag3A β-glucosidases are used to inoculate M9 media containing 0.34 g/l thiamine and 1% w/v cellobiose as the sole carbon source to assay for growth. In each instance a negative vector control in 1% w/v cellobiose is used as well as a positive vector control in 1% w/v glucose. **Note:** OD_{600} results above 1 are considered unreliable as samples were not diluted before readings were recorded.

Figure 3.6A shows that *E. coli* JM109 expressing Cfbglu was able to utilize cellobiose as a source of carbon to achieve a comparable cell density to that of the positive control grown on 1% w/v glucose, although evidently a longer lag phase exists when cellobiose is the sole carbon source. Similar results can be seen for *E. coli* MG1655 expressing Cfbglu when grown on cellobiose, with the exception of a shorter lag phase. Results for the negative control and P_{lac}-*nag3A* in both MG1655 and JM109 strains show no growth on cellobiose media, indicating that Nag3A is not a true cellobiase. These observations are interesting as although Nag3A gives positive β-glucosidase activity in the MUG assay, it is unable to support growth of a recombinant host on cellobiose.

3.4 Assaying for activity of endoglucanases

3.4.1 Confirming the activity of endoglucanases on amorphous cellulose

A qualitative assay for confirming endoglucanase activity was performed for *E. coli* JM109 strains expressing CenA, CenB, CenC or CenD (Figure 3.7). 10 µl spots of overnight culture of *E. coli* P_{lac}-*cenA*, P_{lac}-*cenB*, P_{lac}-*cenC* and P_{lac}-*cenD* were pipetted onto LB agar plates containing 0.2% carboxymethyl cellulose (CMC). These were induced with IPTG and incubated at 37°C overnight, following which plates were flooded with 5 ml of 500 µg/ml Congo Red solution and left to rest at 25°C for 45 mins. Excess dye was then removed and plates flooded with 5 ml of 1 M NaCl solution and left to rest for a further 45 mins to remove any unbound dye. Zones of clearing are evident around colonies positive for endoglucanase activity as the Congo Red dye does not bind to hydrolysed CMC around those colonies. See Chapter 2 of this thesis for full methods.

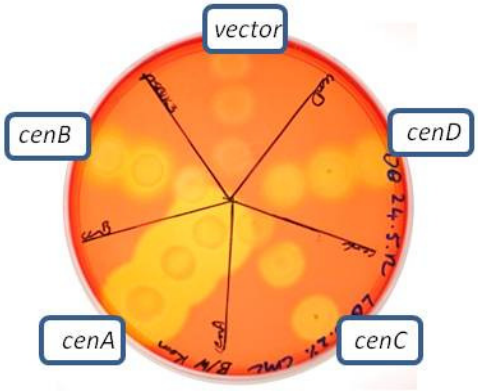

Figure 3.7: Assaying for the activity of endoglucanases active on carboxymethyl cellulose using Congo Red dye. An LB agar plate containing 0.2% carboxymethyl cellulose (CMC) shows zones of clearing around colonies positive for endoglucanase activity when dyed with Congo Red dye. The figure shows 10 ml of overnight culture of five *E. coli* JM109 strains spotted onto LB CMC agar plates.

Figure 3.7 shows that the assay is a purely qualitative one, though does give some indication of which endoglucanases are more active on the amorphous form of cellulose, CMC. The largest zone of clearing generated is that of P_{lac}-*cenA*, followed by that of P_{lac}-*cenB,* P_{lac}-*cenC and* P_{lac}-*cenD*, each of which shows comparable zones of clearing. Notably a slight clearing is also seen for the negative vector control, though this is commonly observed as the presence of the bacterial lawn on the surface of the agar impedes the binding of Congo Red dye to CMC immediately beneath it.

3.4.2 Endoglucanase activity on dyed carboxymethyl cellulose

As a means of generating more quantitative data for the assay of endoglucanase activity on amorphous cellulose, the following colormetric assay measures the absorbance generated by dye liberated from a hydrolysed CMC substrate. CMC dyed with Remazol Brilliant Blue R dye is treated with cell lysates from one of the *E. coli* JM109 endoglucanase expressing strains; P_{lac}-*cenA*, P_{lac}-*cenB*, P_{lac}-*cenC* or P_{lac}-*cenD*. Over the course of 16 hours, samples were taken at regular intervals and a precipitating solution added to terminate the reaction. High-molecular weight material is removed by centrifugation and the absorbance of the supernatant was measured at 590 nm. Results are shown in Figure 3.8 below.

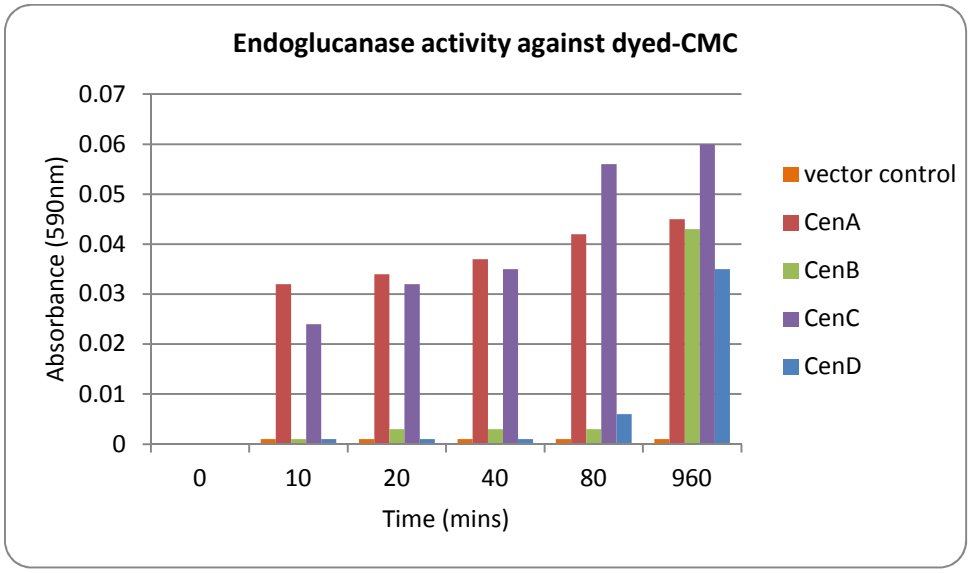

Figure 3.8: Measuring endoglucanase activity on carboxymethyl cellulose dyed with Remazol Brilliant Blue R dye. Cell lysates from *E. coli* JM109 endoglucanase expressing strains are used to treat dyed carboxymethyl cellulose and measure endoglucanase activity as a function of absorbance generated by dye liberated from the cellulose substrate. Results show absorbance measured at 590 nm of dye liberated from CMC treated with the vector control, CenA, CenB, CenC or CenD over the course of 16 h. Reactions are incubated at 37°C and performed in PBS solution at pH 7.4.

Results from Figure 3.8 show the relative activities of the endoglucanases on dyed CMC. In accordance with the published literature (Tomme *et al.* 1996), both CenA and CenC show the greatest activity against amorphous cellulose over the course of 80mins, with CenC liberating 19 times more dye than CenB and 9 times more dye than CenD. However, over the course of 16 h the extent of substrate hydrolysis is more or less comparable between the four endoglucanases as CenC shows only 1.4 times greater activity than CenB and 1.7 times greater activity than CenD. This observation is most likely due to the limited availability of dyed cellulose substrate present within the test samples, leading to a maximal absorbance observed as the substrate is hydrolysed and the dye liberated. In addition, although CenB and CenD show a preference for binding to crystalline cellulose, the cellulases are still able to hydrolyse amorphous regions given a long enough incubation period, although less effectively compared to CenA and CenC (Tomme *et al.* 1996).

3.4.3 Amorphogenesis of cellulose paper by CenA

An interesting feature of endoglucanase CenA not observed amongst the other endoglucanases of *C. fimi*, is the non-hydrolytic disruption of Ramie cellulose fibres by the CBD (Din *et al.* 1994). It is thought that this disruption of the fibre allows for enhanced activity of the cellulases by increasing substrate surface area and liberating low molecular weight cellulose chains weakly associated to the fibre. Currently, no evidence exists within the literature for this phenomenon occurring on other cellulosic substrates other than that of Ramie fibres. As such, I chose to assay CenA for non-hydrolytic disruption of pure cellulose paper. This particular substrate was chosen as its deconstruction can be easily observed during future growth assays of recombinant cellulase expressing hosts (see Chapter 5).

To observe non-hydrolytic disruption of cellulose paper by CenA, a single 10 mg cellulose paper square measuring 1 mm x 10 mm x 5 mm was added to 100 µl PBS solution with 100 µl cell lysate from either *E. coli* JM109 P_{lac}-*cenA,* P_{lac}-*cenB* or pSB1C3 vector control strains. Samples were mixed by vortexing before incubation at 37^{o}C for 48 h. Following this, cellulose paper squares were then washed three times in distilled water and dried at 65^{o}C over 16 h. Samples were coated with 4-6 nm gold by an EmScope sputter coater before subsequent visualization using a Hitachi 4700 II Cold Field-Emission Scanning Electron Microscope.

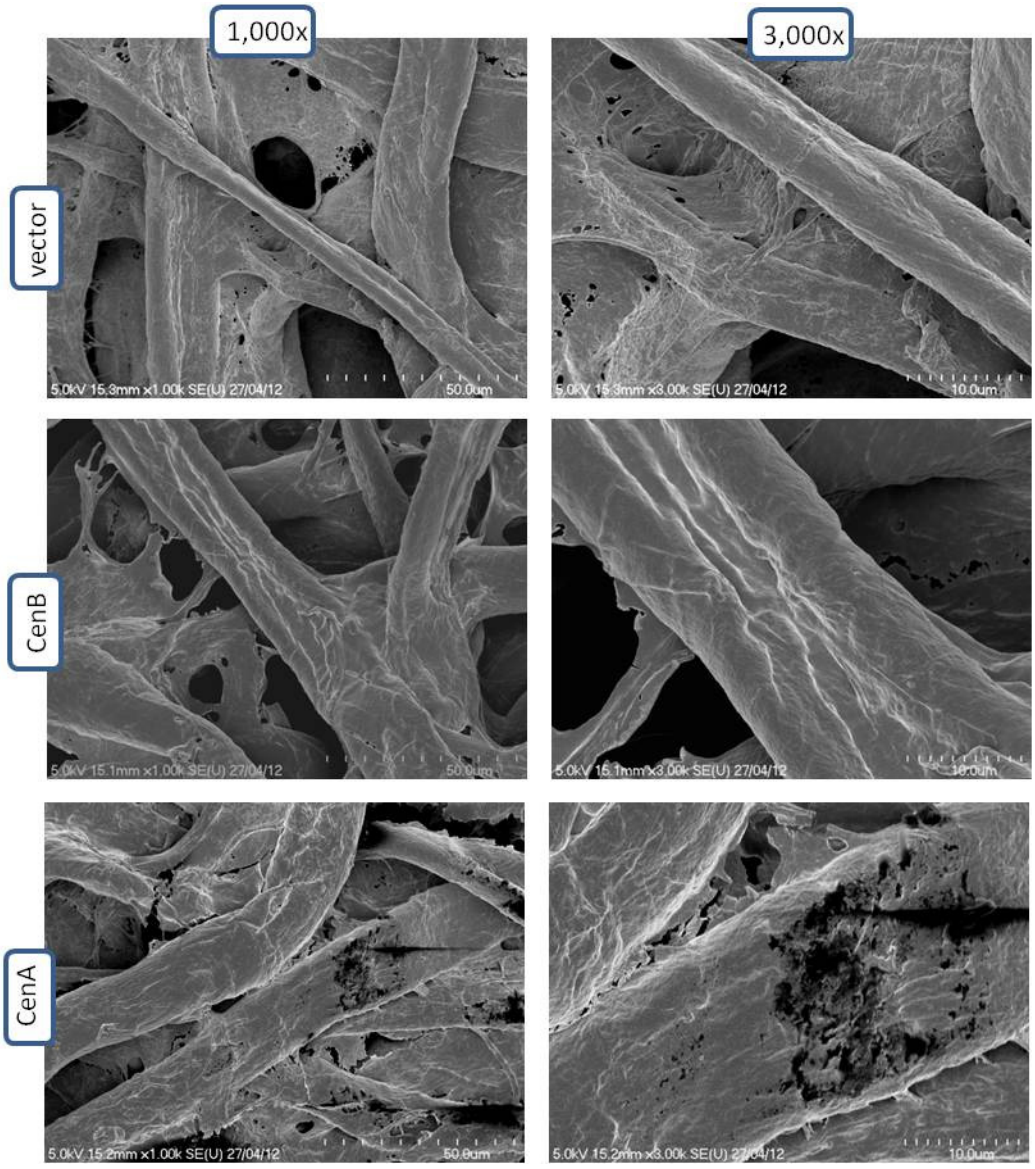

Figure 3.9: Scanning electron micrographs of observed amorphogenesis of cellulose paper by CenA. Scanning electron microscopy images of cellulose paper treated with cell lysates of *E. coli* JM109 vector control, P_{lac}-*cenB*, or P_{lac}-*cenA* strains taken at 1000x and 3000x magnification.

SEM images of cellulose paper fibres treated with cell lysate from the vector control and P_{lac}-*cenB* strains show a smooth and relatively uniform surface at 1,000 times magnification, with little structural disruption (Fig. 3.9). That of fibres treated with cell lysate from P_{lac}-*cenA* exhibit a crenulated surface which is rough and scaled. At 3,000 times magnification those differences are more pronounced with fibres treated with CenA showing pits and grooves as dark areas. This 'shadowing effect' is noted to occur when differences exist in the depth of field between two points on the surface coated with the gold conductive film, a consequence of the disrupted cellulose fibre surface generated by CenA activity.

3.5 Assaying for activity of exoglucanases

3.5.1 Qualitative assays for exoglucanase activity

The ability of exoglucanases to cleave terminal β-1,4-glycosidic bonds and liberate cellobiose can be assayed using the substrate 4-methylumbelliferyl β-D-cellobioside (MUC). Like the MUG assay previously described (see 3.3.1 above), the β-1,4 bond between 4-methylumbelliferyl and cellobioside can be hydrolysed to release methylumberlliferone, which will generate fluorescence under 366 nm light. Figure 3.10A below shows an LB agar plate containing 1.5 mM MUC, with 30 µl of cell lysate from an *E. coli* negative control, *E. coli* P_{lac}-*cex*, *E. coli* P_{lac}-*cbhA* or *E. coli* P_{lac}-*cbhB* added to wells within the plate. Figure 3.10B shows an identical plate spotted with 10 µl of overnight culture of the same strains.

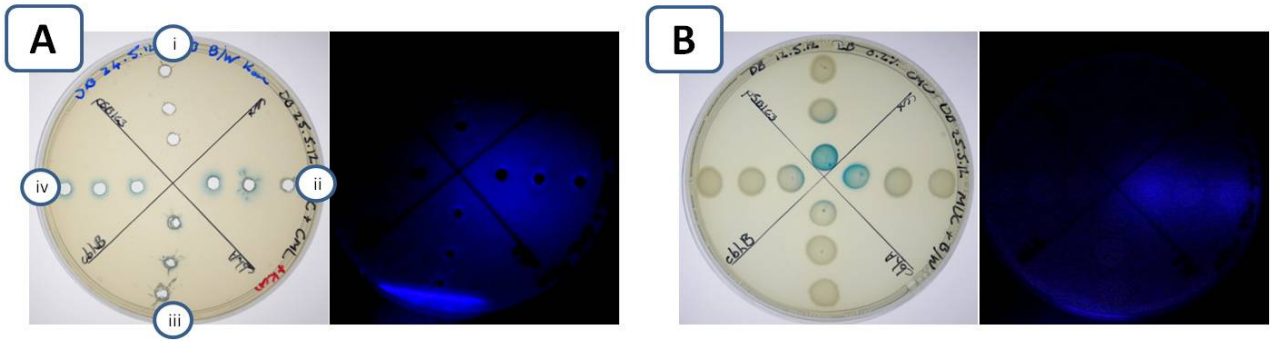

Figure 3.10: Activity of recombinant *C. fimi* exoglucanases on methylumbelliferyl cellobioside. LB agar plates containing 1.5 mM 4-methylumbelliferyl β-D-cellobioside (MUC) were inoculated with *E. coli* MG1655 pSB1C3 vector control **(i),** P_{lac}-*cex* **(ii),** P_{lac}-*cbhA* **(iii),** and P_{lac}-*cbhB***(iv)** strains. 30 µl cell lysate was added to wells within the plate **(A)** or 10 µl of overnight culture was spotted directly onto the agar surface **(B).** Fluorescence was observed when plates were illuminated at 366 nm.

Results from the MUC assay clearly show that both cell lysate and cultures of P_{lac}-*cex* show positive exoglucanase activity, due to fluorescence observed in each instance. Surprisingly, neither CbhA nor CbhB showed any fluorescence under the same conditions. There is no evidence within the published literature which describes either CbhA or CbhB having activity against MUC, as the activity of these cellulases has in the past been routinely measured as a function of reducing sugar liberated from cellulose (Gilkes *et al.* 1997). Attempts were made to replicate methods for measuring reducing sugars described by Gilkes *et al.* though results were not significant and showed no activity for even the positive control Cex (data not shown). The reason for this failure is not known, though a possibility is that a higher purity of cellulase enzyme is needed to generate significant levels of reducing sugar for detection.

3.5.2 Exoglucanase activity against *O*-Nitrophenyl cellobioside

In order to confirm the activities of both CbhA and CbhB, *O*-Nitrophenyl cellobioside was chosen as a substrate. The β-1,4-glycosidic bond when cleaved will release *O*-nitrophenol which produces a vibrant yellow colour quantifiable at 420 nm. Cell lysates of *E. coli* JM109 strains P_{lac}-*cex*, P_{lac}-*cbhA,* P_{lac}-*cbhB* or pSB1C3 vector control were added to 25 mM ONP-cellobioside solution in PBS buffer with a pH of 7.4. Reaction mixes were vortexed and incubated at 37°C. Samples were taken after 3 h and 24 h, with 1 M sodium carbonate solution added to terminate the reaction before absorbance was immediately measured at 420 nm. Measurements showing an initial absorbance of greater than 1unit were diluted 10-fold and measured again, then multiplied by 10 to improve technical accuracy. Results are shown in Figure 3.11 below.

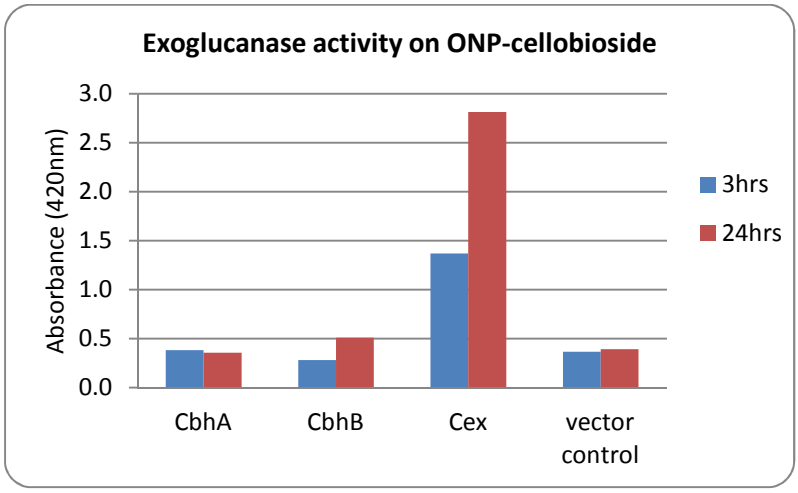

Figure 3.11: Activity of recombinant *C. fimi* exoglucanases on *O*-Nitrophenyl cellobioside. 100 µl of cell lysate from *E. coli* JM109 strains P_{lac}-*cex,* P_{lac}-*cbhA,* P_{lac}-*cbhB* or pSB1C3 vector control are added to 100 µl 25 mM *O*-Nitrophenyl cellobioside in PBS solution in a final volume of 500 µl. Reaction mixtures are incubated at 37°C over 24 h. The reaction was terminated by addition of 500 µl 1 M sodium carbonate. Absorbance measured at 420 nm is proportional to the extent of *O*-Nitrophenyl cellobioside cleaved by exoglucanase active enzymes.

Consistent with results observed from the previous MUC assay, neither CbhA nor CbhB show any appreciable exoglucanase activity compared to the vector control. However that of Cex shows an absorbance greater than ~6-fold that of CbhA and CbhB after 24 h. Although these results show that both CbhA and CbhB lack exoglucanase activity, growth assays of recombinant hosts expressing both CbhA and CbhB suggest that the presence of these genes contribute towards the utilization of cellulose paper (see Chapter 5). The significance of these

observed negative results for measuring exoglucanase activity of CbhA and CbhB on MUC and ONP-cellobioside highlight the difficulties in characterization of the cellulases, and how the choice of substrate and reaction conditions can greatly affect experimental outcomes.

3.6 Assaying for activity of mannanases and xylanases

3.6.1 Exo-mannosidase activity against ONP-mannopyranoside

Using the same protocol as that of the ONP-cellobioside assay, exo-mannosidase activity can be determined using *O*-Nitrophenyl mannopyranoside. Reaction mixes containing 100 µl cell lysate of *E. coli* JM109 P_{lac}-*man2A*, P_{lac}-*man26A* or pSB1C3 vector control strains were added to 25 mM ONP-mannopyranoside solution in PBS buffer. Absorbance was measured at 420 nm after 3 h and 24 h incubation at 37°C. Results are shown in Figure 3.12 below.

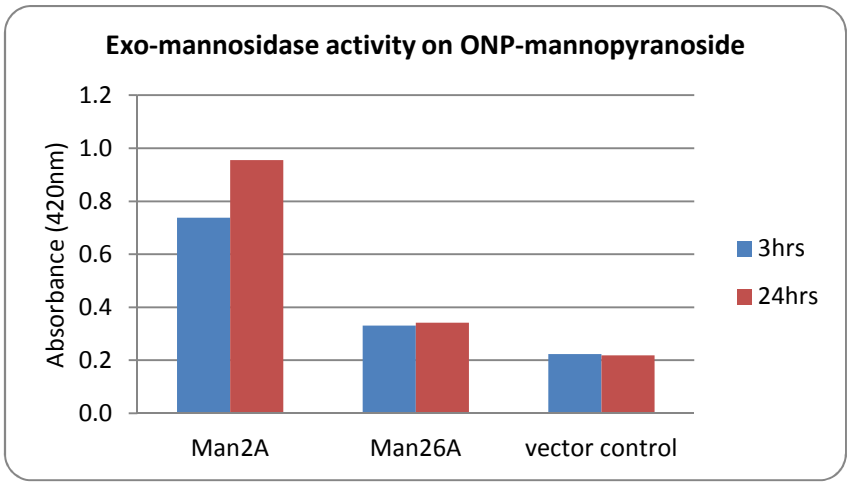

Figure 3.12: Activity of recombinant *C. fimi* mannanases on *O*-Nitrophenyl mannopyranoside. 100 µl of cell lysate from *E. coli* JM109 strains P_{lac}-*man2A*, P_{lac}-*man26A* or pSB1C3 vector control are added to 100 µl 25 mM *O*-Nitrophenyl mannopyranoside in PBS in a final volume of 500 µl. Reaction mixtures are incubated at 37°C for 24 h. The reaction is terminated upon addition of 500 µl 1 M sodium carbonate. Absorbance measured at 420 nm is proportional to the extent of *O*-Nitrophenyl mannopyranoside cleaved by exo-mannosidase active enzymes.

Figure 3.12 shows recombinant Man2A to be an active exo-mannosidase, results which are consistent with the literature (Stoll, Stålbrand, and Warren 1999). Man26A though classed as an endo-mannanase also shows some exo-mannosidase activity as absorbance after 24 h is ~1.5 times that of the vector control.

3.6.2 Exo-xylosidase activity against ONP-xylopyranoside

To assay for exo-xylosidase activity, *O*-Nitrophenyl xylopyranoside was used as a substrate. Following the *O*-Nitrophenyl protocol, lysates of the *E. coli* JM109 strains P_{lac}-*cex*, P_{lac}-*xynD,* P_{lac}-*bxyF*, P_{lac}-*xynF* and pSB1C3 vector control were added to ONP-xylopyranoside. Absorbance was measured at 420 nm after 3 h and 24 h incubation at 37°C. Measurements showing an initial absorbance of greater than 1 were diluted 10-fold and measured again, then multiplied by 10 to improve technical accuracy. Results are shown in Figure 3.13 below.

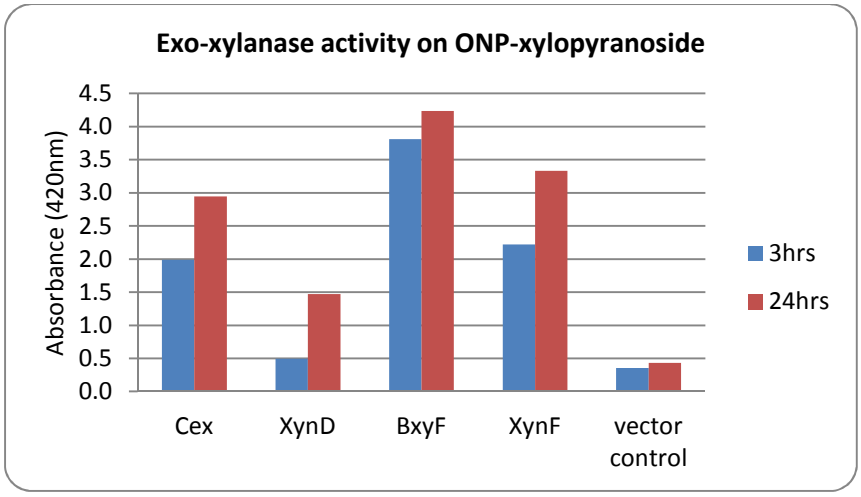

Figure 3.13: Activity of recombinant *C. fimi* xylanases on *O*-Nitrophenyl xylopyranoside. 100 µl of cell lysate from *E. coli* JM109 strains P_{lac}-*cex*, P_{lac}-*xynD,* P_{lac}-*bxyF*, P_{lac}-*xynF* or pSB1C3 vector control were added to 100 µl 25 mM *O*-Nitrophenyl xylopyranoside in PBS solution in a final volume of 500 µl. Reaction mixes were incubated at 37°C for 24 h. The reaction was terminated by addition of 500 µl 1 M sodium carbonate. Absorbance measured at 420 nm is proportional to the extent of *O*-Nitrophenyl xylopyranoside cleaved by exo-xylosidase active enzymes.

Of the four xylanases assayed here, only Cex and XynD have been documented within the literature and both are described as endo-xylanases. BxyF and XynF xylanases are previously uncharacterized *C. fimi* glycanases (cloned by Steven Kane, a fellow lab member, and included here with permission). Results in Figure 3.13 show all four xylanases to have exo-xylosidase activity after 24 h, the greatest activity being that of BxyF with an absorbance of greater than 4 units. It is noteworthy that xynD shows an absorbance comparable to that of the vector control after 3 hours incubation, though this increases 3 fold over the course of 24 hours. Compared to that of BxyF which reaches a near maximal absorbance of ~3.75 units after 3 hours, it may be likely that BxyF is a true exo-xylosidase whilst XynD only shows slight exo- activity. These results highlight the somewhat promiscuous nature of the xylanase

enzymes which are known to display an indiscriminate activity against a range of β-1,4-substrates given a sufficient incubation period (Biely and Puchart 2006).

3.7 Overexpression of recombinant *C. fimi* cellulases

According to Fu *et al.* (2005) overexpression of Cex in *E. coli* has a toxic effect upon the host, leading to cell death. Expression of Cex under a weaker promoter was later shown to improve cell viability and relieve this toxicity (Fu *et al.* 2006). Observations such as these are of great value when considering the rational design of a non-native cellulolytic host, though unfortunately such data is currently lacking within the literature surrounding the *C. fimi* cellulases. As such, those cellulase encoding genes of interest to the current study were cloned in the expression vector pT7-7 and placed under the control of a bacteriophage T7 promoter which is recognized by a highly active T7 RNA polymerase encoded on a second plasmid with expression induced in the presence of IPTG (Tabor 2001). Transformed *E. coli* BL21(DE3) strains were cultured in LB with selective antibiotics and cellulase expression induced after 1 hour by addition of IPTG. Optical density was measured over the course of 6 hours, the results of which are shown in Figure 3.14 below.

Our results confirm the work published by Fu *et al.*, showing that overexpression of Cex reduces cell viability. Moreover a similar trend is seen in that of the strain expressing endoglucanase CenB, though a sharper decline in cell density is seen between 2 and 4 hours, with recovery after 4 hours. Strains expressing the remaining *C. fimi* cellulases show growth broadly similar to that of the vector control strain, with the exception of CenC and CenD which reach stationary phase at 4 hours with an OD of ~1.2. These results suggest that a toxic effect is not observed from overexpression of these cellulases. A possible reason for this is that the genomes of *C. fimi* and *E. coli* BL21(DE3) have a very dissimilar GC content; 74.7% compared to 50.8%. Consequently, these bacteria have drastically different codon usage profiles and the expression of recombinant *C. fimi* cellulases by *E. coli* will be limited by the presence of rare tRNAs such as CGG and CCC of which *E. coli* only exhibits a small pool of compared to *C. fimi.* Ultimately it may be that overexpression of recombinant cellulases in *E. coli* is limited at a transcriptional level, reducing the levels of mature protein present and reducing any potential toxic effects observed by those cellulases.

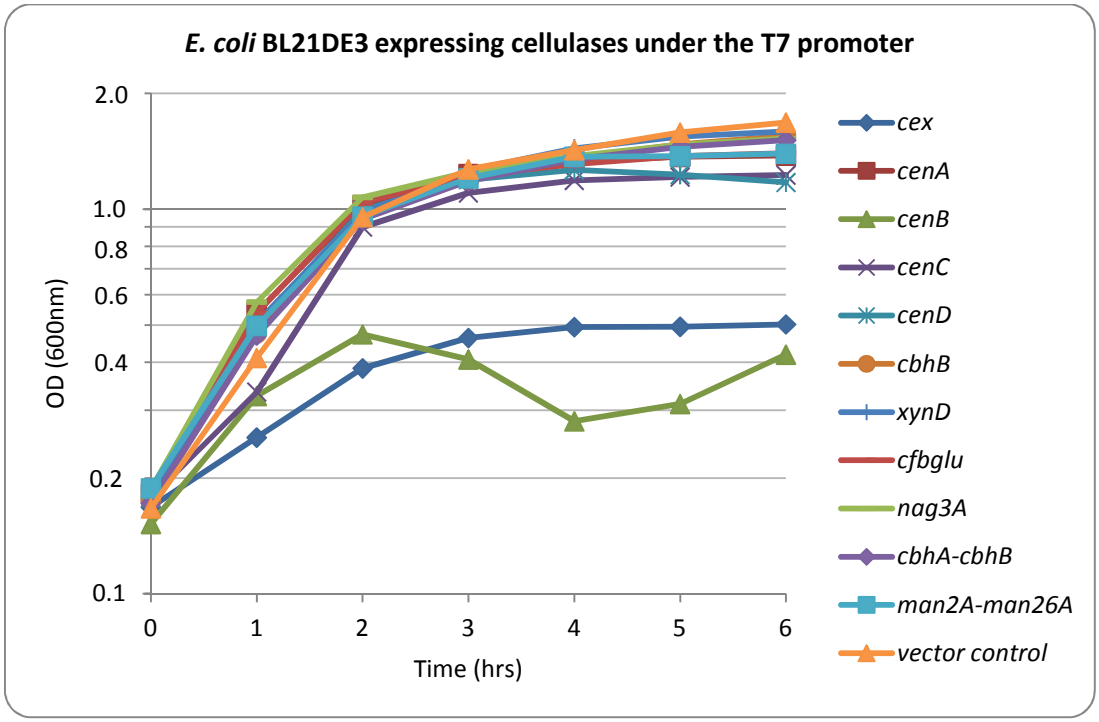

Figure 3.14: Growth curves of *E. coli* BL21(DE3) strains overexpressing recombinant *C. fimi* cellulases under the T7 promoter. *E. coli* BL21(DE3) strains expressing *C. fimi* cellulases under transcriptional control of the highly active bacteriophage T7 RNA polymerase. Cultures were grown in LB media containing 100 µg/ml ampicillin and 40 µg/ml chloramphenicol. Expression is induced by the addition of 90 µg/ml IPTG at 1 hour, and absorbance measured at 600 nm over 6 hours. **Note:** OD_{600} results above 1 are considered unreliable as samples were not diluted before readings were recorded.

3.8 Design and construction of composite parts

Previous work in characterization of expression hosts and cellulase encoding BioBrick parts was used to inform the design and assembly of parts into composite devices. These constructs were designed with the intention of encoding defined activities against a target substrate such as cellulose, mannan or xylan. Here we discuss the rationale in their design and construction with specific emphasis on the intended expression host.

Constructs described here are illustrated according to the Synthetic Biology Open Language (SBOL) visual standard version 1.0 (www.sbolstandard.org). The purpose of this graphical annotation is to standardise the format in which parts and devices are described, and to facilitate their design *in silico*. Figure 3.15 below illustrates the legend used for representing the various parts used within this study according the SBOL standard. Parts described include protein coding sequences, ribosome binding sites, promoter sequences and BioBrick prefixes and suffixes.

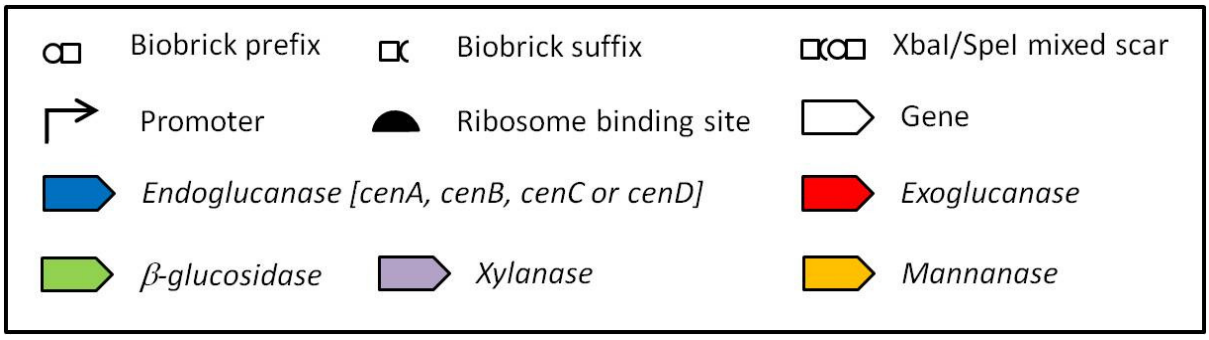

Figure 3.15: Synthetic Biology Open Language (SBOL) version 1.0 graphical annotations used in describing the components of BioBrick constructs assembled. The legend illustrates the BioBrick prefix and suffix, as well as the mixed scar site generated from the ligation of cut XbaI and SpeI restriction sites. Transcriptional and translational control elements depicted here include promoters and ribosome binding sites. Protein coding sequences relevant to this study are colour coded according to the encoded enzymatic activity. Each part used is represented according to the Synthetic Biology Open Language visual standard version 1.0 (www.sbolstandard.org).

3.8.1 Assembly of cellulase encoding constructs for *Citrobacter freundii*

As noted within the literature, the endoglucanases of *C. fimi* show variable preferences in the binding of and activity against different forms of cellulose (Tomme *et al.* 1996; Koivula *et al.* 2000). Consequently, constructs were designed to encode a single endoglucanase which is expressed in concert with one or more exoglucanases. This was intended to confirm whether or not recombinant hosts transformed with a single construct and expressing a single endoglucanase will show enhanced growth on a particular cellulose substrate. Moreover, co-cultures of multiple strains each expressing a single endoglucanase can easily be performed so as to study any synergistic effects between these various strains due to complementarities between the endoglucanases present.

Figure 3.16 below illustrates the design of two construct formats encoding cellulase activity intended for expression by *Citrobacter freundii*. The first format encodes a single endoglucanase (CenA, CenB or CenD) and the exoglucanase Cex, with expression under the control of a single inducible P_{lac} promoter. Unfortunately the assembly of a construct encoding both CenC and Cex was not achieved for reasons unknown, perhaps a toxic phenotype when the two cellulases are expressed in tandem. The second format illustrated in Figure 3.16 is the same as the first, with the addition of CbhA and CbhB expression under the control of the IPTG inducible P_{spac} promoter. This format was designed so as to determine whether or not the two cellobiohydrolases contributed towards deconstruction of cellulose as

activity assays discussed previously were uninformative. In total, six constructs encoding cellulase activity for expression in *C. freundii* were assembled.

A notable advantage of assembling all parts according to a standardized assembly method such as that of BioBricks, is that any devices composed of multiple parts can be easily transferred to a range of expression vectors without the need to modify the original construct. As such, constructs were placed in both high-copy and low-copy plasmid vectors (pSB1K3 and pSB4K5 respectively), as a means to vary enzymatic expression and metabolic load upon the expression host. Results are discussed further in Chapters 4 and 5.

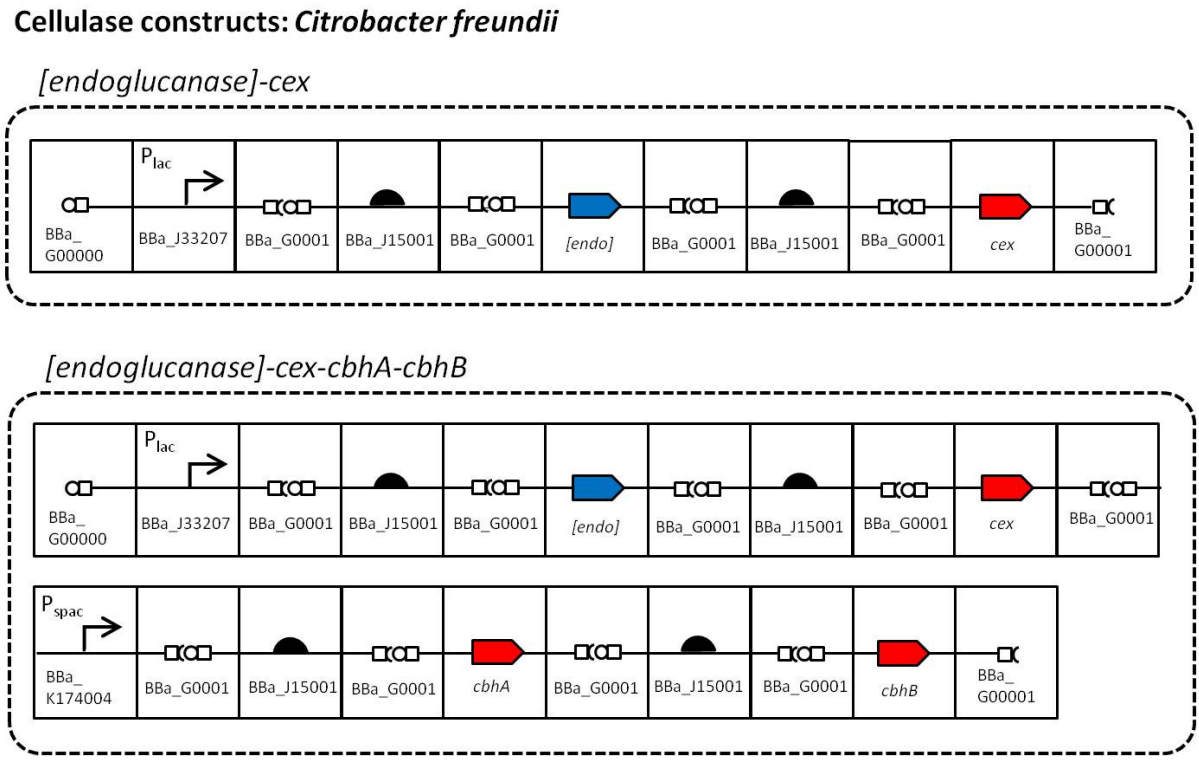

Figure 3.16: Cellulase encoding constructs assembled for expression in *Citrobacter freundii* described according to the Synthetic Biology Open Language v.1.0. Constructs described as *[endoglucanase]-cex* are those which endcode the expression of a single endoglucanase (CenA, CenB or CenD) as well as the exoglucanase Cex under the control of the IPTG inducible P_{lac} promoter within a single mRNA transcript. Constructs described as *[endoglucanase]-cex-cbhA-cbhB* are those which encode the expression of a single endoglucanase (CenA, CenB or CenD) and the exoglucanase Cex under the control of P_{lac}, as well as the expression of the cellobiohydrolases CbhA and CbhB under the control of the IPTG inducible P_{spac} promoter. Parts are listed according to their BioBrick Foundation Parts Registry part number where applicable (www.partsregistry.org).

Successful assembly of composite devices and host cell transformation was determined by: (i) PCR cloning, (ii) restriction digest analysis of plasmid DNA (see Appendix) (iii) as well as activity assays for the encoded enzymes. Figure 3.17 below shows results for Congo Red endoglucanase assays for the *[endoglucanase]-cex* constructs (Fig. 3.17A) and the *[endoglucanase]-cex-cbhA-cbhB* constructs (Fig. 3.17B). Results confirm the activities of each endoglucanase encoded by the respective construct. Interestingly, varying expression of *cenA-cex-cbhA-cbhB* in either low or high copy vectors does not appear to have an appreciable effect on endoglucanase activity measured in this assay (Fig. 3.17B). Further study of low and high-copy expression of this construct is presented in Chapters 4 and 5.

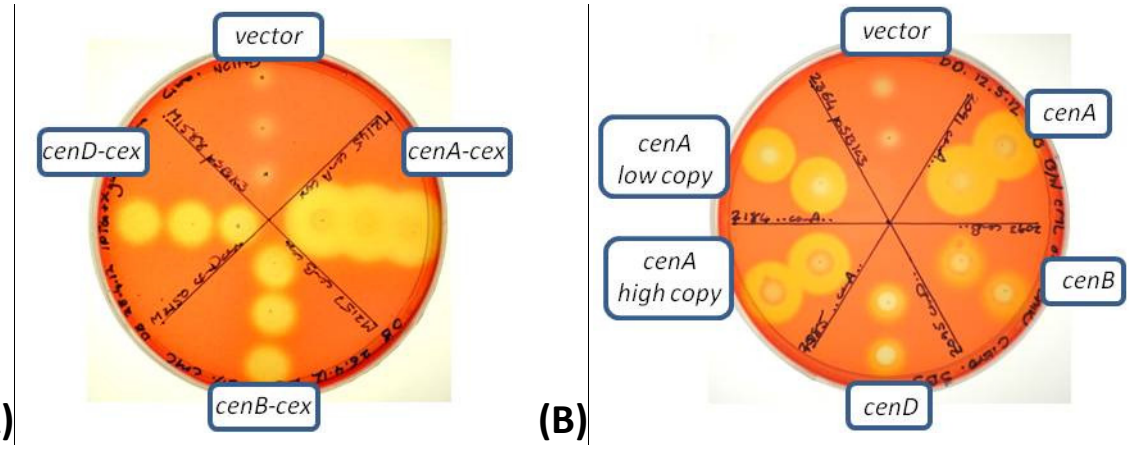

Figure 3.17: Congo Red endoglucanase activity assays confirming the activity of cellulase constructs assembled for expression in *Citrobacter freundii*. *[endoglucanase]-cex* constructs encoding CenA, CenB or CenD show endoglucanase activity on CMC agar plates **(A).** *[endoglucanase]-cex-cbhA-cbhB* constructs encoding CenA, CenB or CenD show endoglucanase activity on CMC agar plates **(B).** Results for *cenA-cex-cbhA-cbhB* expression in both low copy and high copy vectors are shown **(B).**

3.8.2 Assembly of cellulase encoding constructs for *Escherichia coli*

Following the same design principles as that of the constructs assembled for expression in *C. freundii*, cellulase encoding constructs assembled for *E. coli* expression also included a single endoglucanase encoding gene part. Figure 3.18 illustrates three formats of construct design for cellulase expression in *E. coli*.

Cellulase constructs: *E. coli*

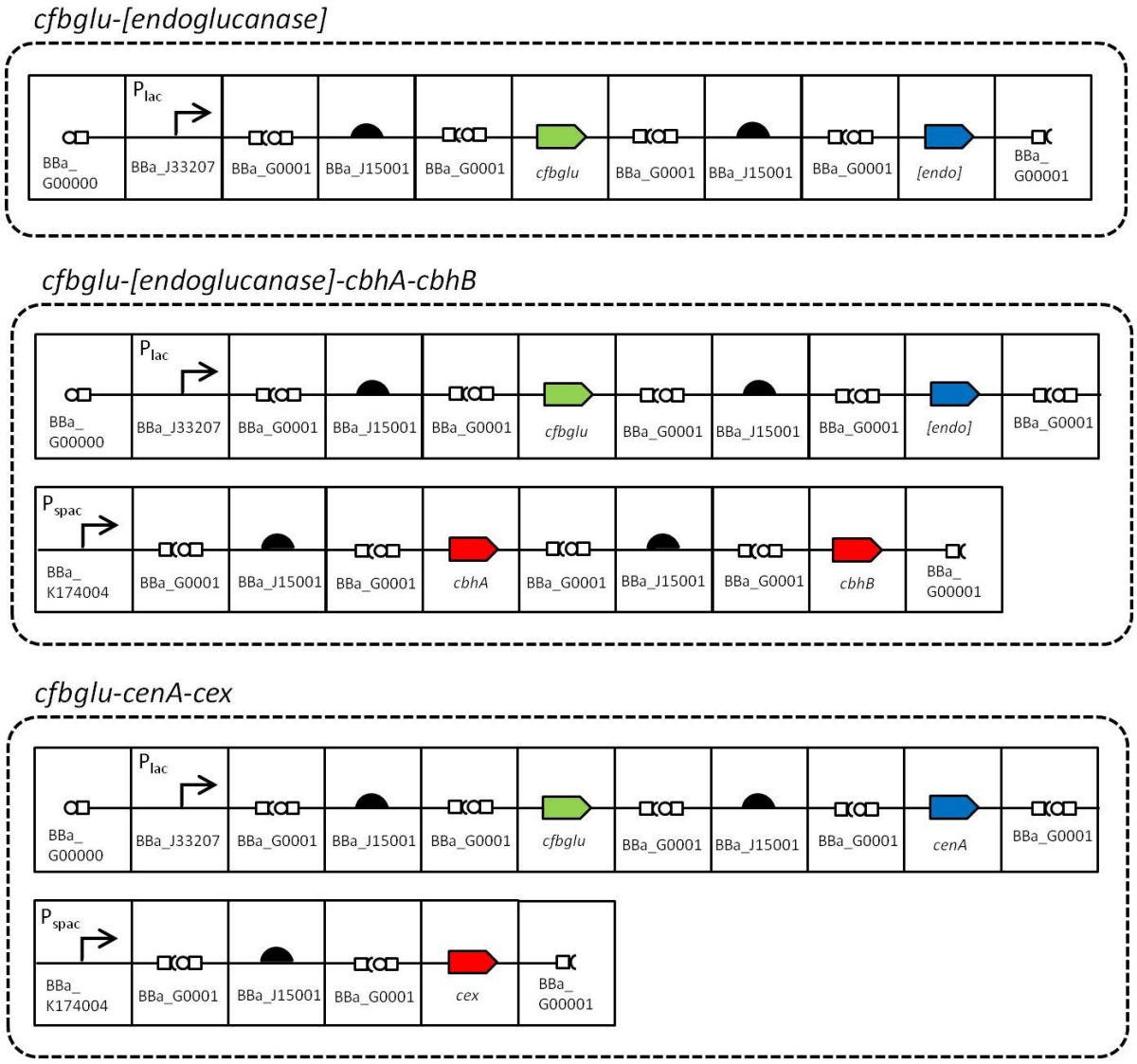

Figure 3.18: Cellulase encoding constructs assembled for expression in *Escherichia coli* described according to the Synthetic Biology Open Language v.1.0.0. Constructs described as *cfbglu-[endoglucanase]* are those which encode the expression of the β-glucosidase Cfbglu and a single endoglucanase (CenA, CenB, CenC or CenD) under the control of the IPTG inducible P_{lac} promoter within a single mRNA transcript. Constructs described as *cfbglu-[endoglucanase]-cbhA-cbhB* are those which encode the expression of the β-glucosidase Cfbglu and a single endoglucanase (CenA, CenB, CenC or CenD) under the control of P_{lac}, as well as the expression of the cellobiohydrolases CbhA and CbhB under the control of the IPTG inducible P_{spac} promoter. Parts are listed according to their BioBrick Foundation Parts Registry part number where applicable (www.partsregistry.org).

The first format is *cfbglu-[endoglucanase]* where the β-glucosidase Cfbglu is expressed in addition to one of the endoglucanases; CenA, CenB, CenC or CenD. This format was designed to investigate whether or not the expression of a single β-glucosidase and endoglucanase was sufficient to support growth on cellulose, since literature surrounding the *C. fimi* endoglucanases describes the liberation of glucose and cellobiose from cellulosic substrates treated with CenA, CenB, CenC or CenD (Tomme *et al.* 1996). The second format is that of *cfbglu-[endoglucanase]-cbhA-cbhB* which includes in addition to the above, the expression of the two cellobiohydrolases CbhA and CbhB. The third format is that of *cfbglu-cenA-cex*, designed for the purpose of determining whether multiple exoglucanases are required to support growth of recombinant *E.coli* on cellulose.

In total, nine constructs encoding cellulase activity were assembled for expression in *E. coli*. Successful assembly of composite devices and host cell transformation was determined by: (i) PCR cloning, (ii) restriction digest analysis of plasmid DNA (see Appendix) (iii) as well as activity assays for the encoded enzymes. Figure 3.19 below shows results for Congo Red endoglucanase assays of strains transformed with the *cfbglu-[endoglucanase]-cbhA-cbhB* constructs (Fig. 3.19A). Results confirm the activities of each endoglucanase encoded by the respective construct. Figure 3.19B shows strains that are transformed with *cfbglu-[endoglucanase]-cbhA-cbhB* are positive for β-glucosidase activity. *C. freundii* is used as a positive control for β-glucosidase activity.

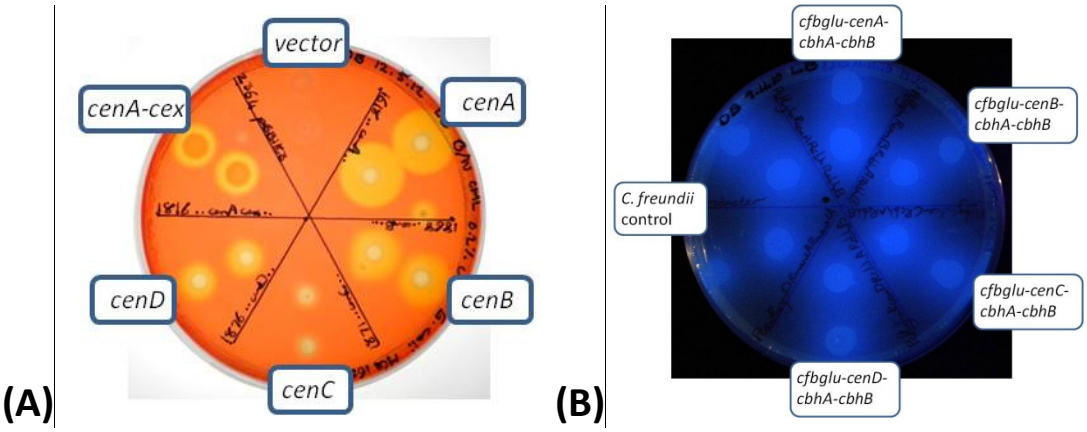

Figure 3.19: Congo Red endoglucanase activity assays and MUG β-glucosidase activity assays confirming the activity of cellulase constructs assembled for expression in *Escherichia coli*. *cfbglu-[endoglucanase]-cbhA-cbhB* constructs encoding CenA, CenB, CenC or CenD show endoglucanase activity on CMC agar plates, as well as *cfbglu-cenA-cex* showing endoglucanase activity **(A).** Strains transformed with *cfbglu-[endoglucanase]-cbhA-cbhB* constructs spotted onto MUC plates show positive β-glucosidase activity compared to a *C. freundii* positive control **(B).**

3.8.3 Assembly of xylanase and mannanase encoding constructs

For the utilization and deconstruction of xylan, three constructs were assembled; *cex-xynD, bxyF-cex,* and *xynF-xynD*. These are illustrated in Figure 3.20. Each construct encodes two xylanases from *C. fimi*, with expression under the control of the IPTG inducible P_{lac} promoter. Each construct is suitable for expression in either *E. coli* or *C. freundii*. In addition co-cultures of strains each expressing one of the xylanase constructs can be performed as each construct is placed in the pSB1C3 vector encoding chloramphenicol resistance. Successful assembly of composite devices and host cell transformation was determined by PCR cloning and restriction digest analysis of plasmid DNA (see Appendix).

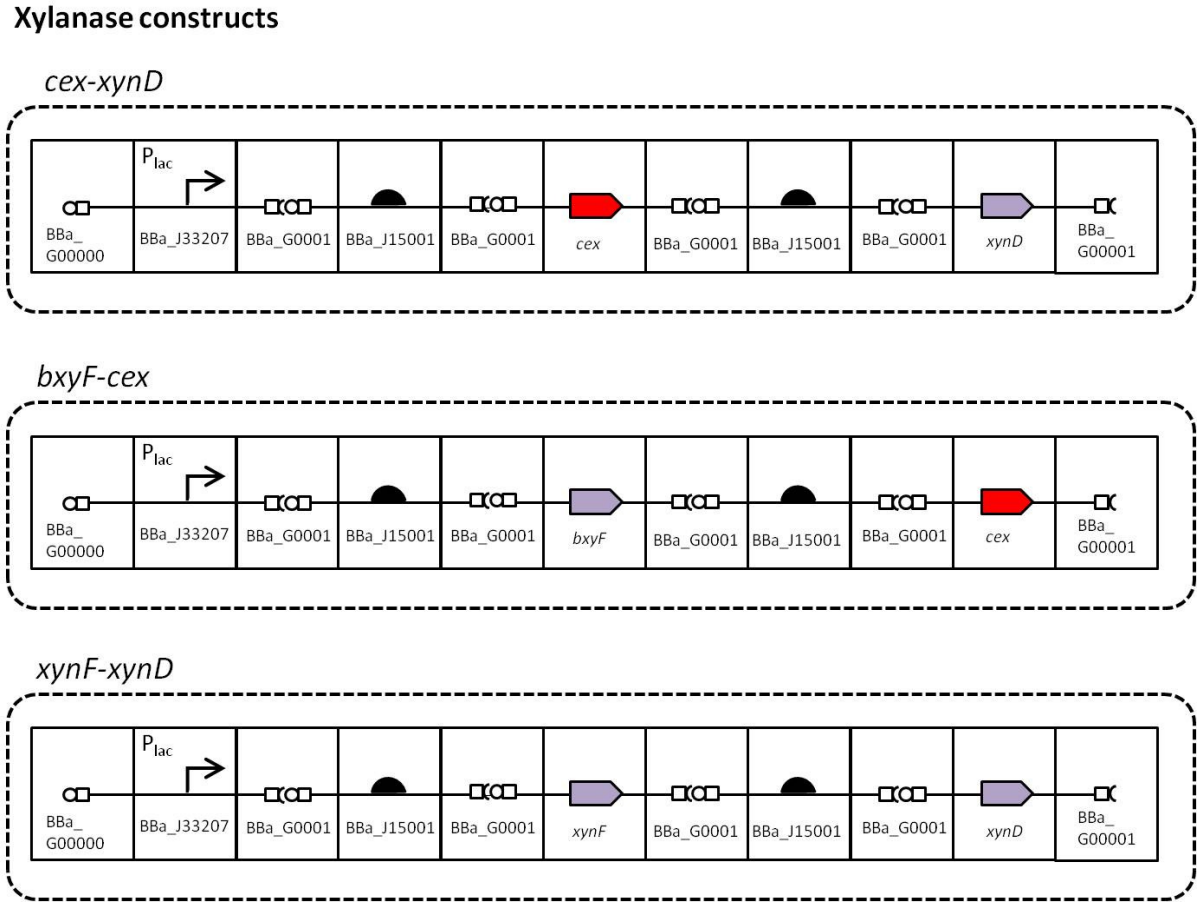

Figure 3.20: Xylanase encoding constructs assembled for expression in *Escherichia coli* and *Citrobacter freundii* described according to the Synthetic Biology Open Language v.1.0. Constructs described here include; *cex-xynD* encoding the endo-xylanases Cex and XynD, *bxyF-cex* encoding the exoxylanase BxyF and endoxylanase Cex, *xynF-xynD* encoding the xylanase XynF and endoxylanase XynD. Expression of each construct is under the control of P_{lac}. Parts are listed according to their BioBrick Foundation Parts Registry part number where applicable (www.partsregistry.org).

BioBrick parts encoding mannanases for the deconstruction and utilization of mannan were assembled onto a single construct; *man2A-man26A*. This is illustrated in Figure 3.21. Expression of Man2A and Man26A was put under the control of the P_{lac} promoter. According to the literature, complete hydrolysis of mannan is achievable in the presence of both Man2A and Man26A (Stoll, Stålbrand, and Warren 2001). The construct described is suitable for expression in either *E. coli* or *C. freundii* as both species are able to utilize mannose sugars liberated from mannan (Holt and Bergey 2000). Successful assembly of composite devices and host cell transformation was determined by PCR cloning and restriction digest analysis of plasmid DNA (see Appendix).

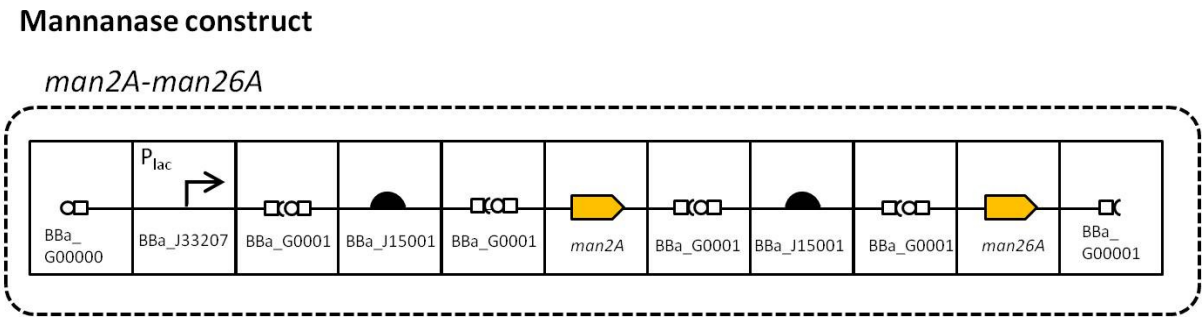

Figure 3.21: Mannanase encoding construct assembled for expression in *Escherichia coli* and *Citrobacter freundii* described according to the Synthetic Biology Open Language v.1.0. The construct described encodes the exo-mannosidase Man2A and the endo-mannanase Man26A under the control of the IPTG inducible P_{lac} promoter. Parts are listed according to their BioBrick Foundation Parts Registry part number where applicable (www.partsregistry.org).

3.9 Discussion

The intention of the characterization assays and results described within this chapter was to inform the design and assembly of multiple constructs with defined activities against the major polysaccharides present in plant biomass; cellulose, mannan and xylan. The resultant composite devices were purposely designed for expression within suitable expression hosts; in this case those hosts being *E. coli* and *Citrobacter freundii*. Hosts were chosen on the basis of the following criteria; (i) ability to utilize glucose, mannose and xylose liberated from plant biomass, (ii) sufficient secretion of extra-cellular proteins, (iii) ease of genetic manipulation and cultivation within a laboratory environment.

Though *E. coli* is not well suited to the expression of extracellular proteins, there is evidence within the literature for the non-specific leakage of *C. fimi* cellulases from the periplasm and into the culture medium (Guo *et al.* 1988). Having confirmed such observations, we decided to perform all cloning and assembly of BioBrick parts encoding *C. fimi* cellulases in *E. coli* with the intention of transferring expression to a more suitable host. As demonstrated within this chapter, the constructs developed in *E. coli* were repurposed for expression in *C. freundii*, since both bacteria share transcriptional and translational elements. *C. freundii* exhibiting a type II secretion system offers itself as a suitable host for expression of extracellular endo- and exoglucanases. Moreover, cellobiose liberated by the concerted action of endo- and exoglucanases can be internalised by *C. freundii* and hydrolysed to glucose by a native β-glucosidase. In addition, results presented within this chapter highlight an interesting aspect to the use of *C. freundii* in that the bacterium is able to tolerate saline media. Growth is almost entirely unaffected in the presence of 0.6 M sodium chloride, opening up avenues in the use of sea water in large scale bioprocessing projects.

One of the major goals in the field of Synthetic Biology is the development of standardized genetic parts and the assembly of these into devices with predictable properties. With this in mind, work presented here on the characterization and testing of parts active on cellulosic substrates highlights several of the challenges inherent to standardizing methodologies in measuring glycanase activity. Most notably, the β-glucosidase Nag3A is shown to give positive results when activity is measured on the substrate 4-methylumbelliferyl-β-1,4-D-glucopyranoside, though expression by a recombinant host is insufficient for growth on cellobiose. In addition, reaction conditions are shown to affect the observed activity of glycanases being assayed. This is exemplified in the case of the endo-xylanases Cex and XynD which show strong exo-xylosidase activity after an extended incubation with the substrate of 24 h. These results illustrate the need to standardize methods and experimental conditions when measuring glycanase activities within such metabolic engineering projects, and emphasize the importance of detailed parts characterization in order for the design of predictable genetic devices.

Chapter 4

Utilization of defined cellulosic and hemicellulosic substrates

Summary

To assay for the cellulolytic capacity encoded by the assembled constructs described in Chapter 3, expression host strains *E. coli* MG1655*, C. freundii* NCIMB11490 and SBS197 were transformed and assayed for growth on defined cellulosic and hemicellulosic substrates, including; mannan from *Saccharomyces cerevisiae*, beechwood xylan, carboxymethyl cellulose and avicel. To accommodate parallel screening and the reliable assay of cell growth on these substrates, cell density was measured as a function of protein content using the Coomassie protein assay reagent. This method was chosen as it is more suitable for high throughput studies compared to performing colony counts. Preliminary trial experiments using the reagent are performed on *C. fimi*, *E. coli* and *C. freundii* strains so as to troubleshoot potential problems, and in addition establish benchmark results for *C. fimi* grown on various carbon sources with which to compare results of our recombinant strains. *E. coli* transformed with the mannanase encoding construct was shown to utilize mannan as a carbon source in M9 minimal medium; however *C. freundii* NCIMB11490 exhibiting the same construct was not able to utilize mannan. Xylan utilization assays identified *E. coli* transformed with either xylanase construct was able to utilize the substrate, this was echoed for *C. freundii* NCIMB11490 strains except in the case of *xynF-xynD* which did not show appreciable growth. Of the recombinant cellulase expressing *E. coli* strains grown on CMC, those expressing CenA best utilized the substrate. Moreover expression of the cellobiohydrolases CbhA and CbhB contributed significantly towards utilization of the substrate. A synergistic cooperation between strains expressing endoglucanases CenB and CenD was observed on CMC. Assays performed for *C. freundii* cellulase strains showed that NCIMB11490 strains were not able to utilize CMC, whilst SBS197 strains were able. *C. freundii* SBS197 cellulase strains were shown to utilize crystalline avicel as a carbon source. It was noted from these results that the expression of the cellobiohydrolases CbhA and CbhB did not significantly contribute towards utilization of the substrate when expressed by *C. freundii*.

4.1 Introduction

The constructs described in Chapter 3 encoding enzymes for the hydrolysis of mannan, cellulose and xylan were introduced into the expression hosts *E. coli* MG1655, *Citrobacter freundii* NCIMB11490 and SBS197. Minpreps to isolate plasmid DNA encoding the recombinant cellulases from the relevant strains were performed. Subsequent restriction digest analysis to confirm the presence and size of those assemblies were conducted (see Appendix Figures A-C). Presented within this chapter are the results for growth assays of the various recombinant strains on defined substrates; namely mannan from *Saccharomyces cerevisiae*, beechwood xylan, carboxymethyl cellulose (CMC) and avicel.

4.2 Coomassie protein assay

To systematically and accurately assay for growth of multiple recombinant hosts and strains across a range of cellulosic substrates, a reproducible high throughput assay method was required. Notably, the measurement of cell density as a function of culture absorbance would not be a practical solution in this instance, as the insoluble nature of some of the cellulosic substrates results in cultures having a high particulate content, thus skewing absorbance measurements. With these considerations in mind, we chose to measure cell growth as a function of protein content within the culture by means of a Coomassie protein reagent assay.

The protein reagent assay is based on the binding of proteins to the coomassie dye which results in a spectral shift from a reddish/brown colour (maximum absorbance at 465 nm) to the blue form of the dye (maximum absorbance at 610 nm). The difference between the two forms of the dye is greatest at 595 nm, and is thus the optimal wavelength for measurement of the blue colour generated by the coomassie dye/protein complex. The coomassie dye is known to be unaffected by the presence of free amino acids, peptides and low molecular weight proteins. Importantly, the assay measures for total protein content and so the results of samples taken from cultures will include the presence of metabolically active cells as well as inactive dead cells (Bradford 1976).

4.2.1 Preliminary assays and benchmarks for *C. fimi* growth on various carbon sources

As a preliminary test to determine the efficacy of the Coomassie protein assay, M9 minimal media containing various cellulosic carbon sources were set up and inoculated with *Cellulomonas fimi* to measure cell growth. Carbon sources present in those cultures included

one of the following; avicel, CMC, cellulose paper, cellulose paper in 0.6 M NaCl media, or xylan. Cultures were incubated at 30°C on a rotary shaker at 200 rpm. 100ul samples were taken every 24 hours and assayed for protein content (see Chapter 2 for full methods). Absorbance at 595 nm was measured and the results are shown in Figure 4.1.

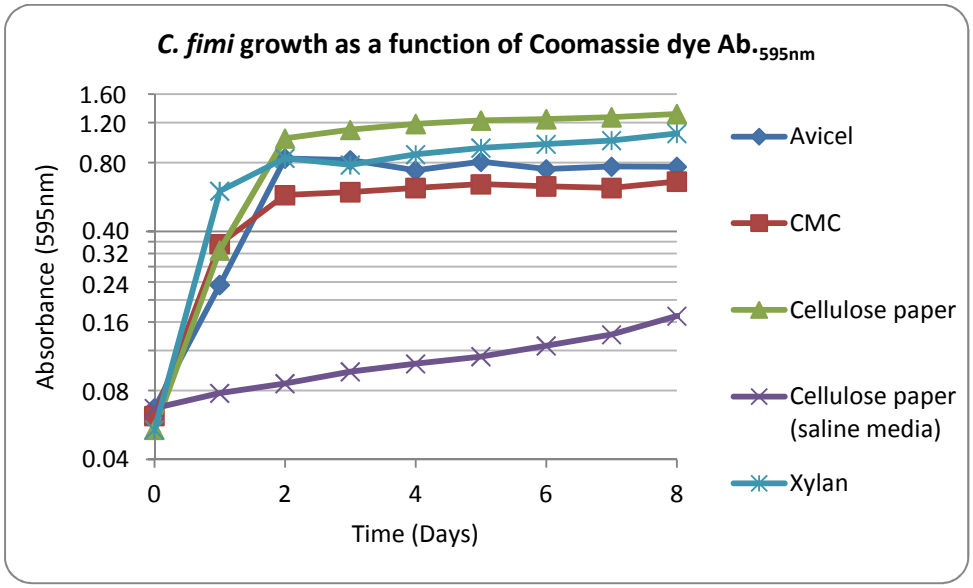

Figure 4.1: Growth of *Cellulomonas fimi* on various carbon sources measured as a function of Coomassie dye absorbance at 595 nm. To test for the efficacy of the Coomassie protein assay, cultures of *C. fimi* were grown in 5 ml M9 medium containg 0.34 g/l thiamine and a single carbon source; 50 mg avicel, 50 mg carboxymethyl cellulose, 125 mg cellulose paper, 125 mg cellulose paper in saline medium, or 50 mg beechwood xylan. 100 µl samples were taken every 24hrs over the course of 8 days and absorbance of the Coomassie dye measured at 595 nm to assay protein content. **Note:** Ab_{595} results above 1 are considered unreliable as samples were not diluted before readings were recorded.

Figure 4.1 shows that *C. fimi* is able to utilize cellulose paper, xylan, avicel and CMC. Notably, *C. fimi* is unable to tolerate 0.6 M NaCl medium as no appreciable growth is observed on cellulose paper in saline conditions. Highest growth is achieved on cellulose paper which is a mix of both amorphous and crystalline cellulose. Growth on xylan is also considerable since this substrate does not exhibit a crystalline form like that of avicel and as such is easier to degrade. Strikingly though growth on avicel is higher than that of CMC. This observation may be explained as CMC is a methylated form of cellulose, with every 1 in 10 glucopyranoside residues being methylated and so unusable by *C. fimi*.

The purpose of this preliminary assay was not only to establish a benchmark for the activities of *C. fimi* against these substrates for comparison to our recombinant hosts, but also as a means of troubleshooting potential problems within the assay protocol. One such challenge

was measuring data for cultures grown on avicel and xylan. As these substrates contain a high particulate content, absorbance readings were inconsistent and highly variable in some cases. In an attempt to address this issue, absorbance was measured at 530 nm and deducted from that measured at 595 nm so as to eliminate absorbance generated from turbidity.

4.2.2 Comparing Coomassie assay results across bacterial species

To determine how suitable the Coomassie assay is for measuring and comparing protein content across different bacterial species, cultures of *C. freundii* NCIMB11490, *E. coli* MG1655 and *C. fimi* were conducted in M9 medium with 1% w/v glucose and grown to stationary phase. Samples were taken and serial dilutions made in fresh media. The culture optical density at 600 nm was measured versus the absorbance of those same samples treated with the Coomassie reagent at 595 nm. Results are shown in Figure 4.2 below.

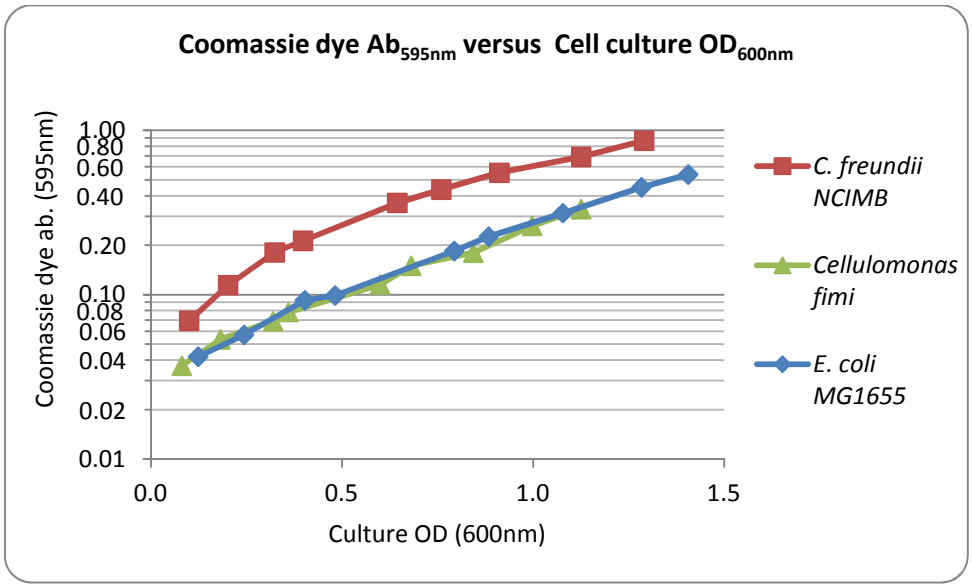

Figure 4.2: Coomassie dye absorbance at 595 nm measured versus cell culture optical density at 600 nm across the three species *C. freundii, E. coli* and *C. fimi*. Cultures of *C. freundii, E. coli* and *C. fimi* were grown to stationary phase in LB. Serial dilutions were made and their OD measured at 600 nm. Protein content was measured using the Coomassie protein assay and results recorded as absorbance at 595 nm. Plotted against one another results show the differential relationships of protein content versus cell density across the three species. **Note:** OD_{600} results above 1 are considered unreliable as samples were not diluted before readings were recorded.

Figure 4.2 shows that although two bacterial cultures exhibit the same OD_{600}, each may give significantly different results for the protein content measured by the Coomassie assay. *E. coli* MG1655 and *C. fimi* show a very similar trend when culture absorbance is measured versus Coomassie dye absorbance, indicating those cultures of both species share a similar

protein content at any given cell density. Conversely, *Citrobacter freundii* will show significantly higher protein content at any given cell density. Results described here could be interpreted in a number of ways. One such interpretation is that protein content of each species can vary greatly and so sensitivity of the Coomassie dye to any given protein mix can vary too, giving higher or lower values for protein assayed.

The relevance and importance of these findings to the present chapter is that it would be unfair to compare results for Coomassie dye absorbance across the different bacterial species, as the values generated do not necessarily correlate to cell density in a linear fashion. As such, throughout this chapter we will discuss results for Coomassie dye absorbance according to one bacterial species at a time and not compare across species.

4.3 Utilization of mannan

Expression hosts *E. coli* MG1655 and *C. freundii* NCIMB11490 transformed with *man2A-man26A* in pSB1K3 were assayed for their ability to utilize 0.5% w/v mannan from *Saccharomyces cerevisiae*. M9 medium containing 0.5% w/v mannan and 0.1% w/v yeast extact was inoculated with cell pellets from overnight cultures of the two strains resuspended in M9 minimal media so that the initial OD_{600} of the culture was 0.1 at time 0. Cultures were incubated at 37^{o}C on a rotary shaker at 200 rpm. Over the course of 11 days 100ul samples were taken at 24 hr intervals and protein content measured by the Coomassie protein assay. All cultures were performed in duplicate. Results are shown in Figure 4.3 below.

It can be seen that the difference in observed protein content between the *E. coli* MG1655 vector control and *man2A-man26A* strains is very similar over the first 4 days (Fig. 4.3A), after which, *man2A-man26A* shows a higher protein content that is steadily maintained over the remainder of the assay. Results for *C. freundii* NCIMB11490 *man2A-man26A* are however less encouraging as protein content is almost identical to that as the vector control strain, indicating that mannan is not being utilized as an additional carbon source (Fig. 4.3B). To confirm that *C. freundii* is able to utilize mannose, future growth assays should be conducted using this as the sole carbon source.

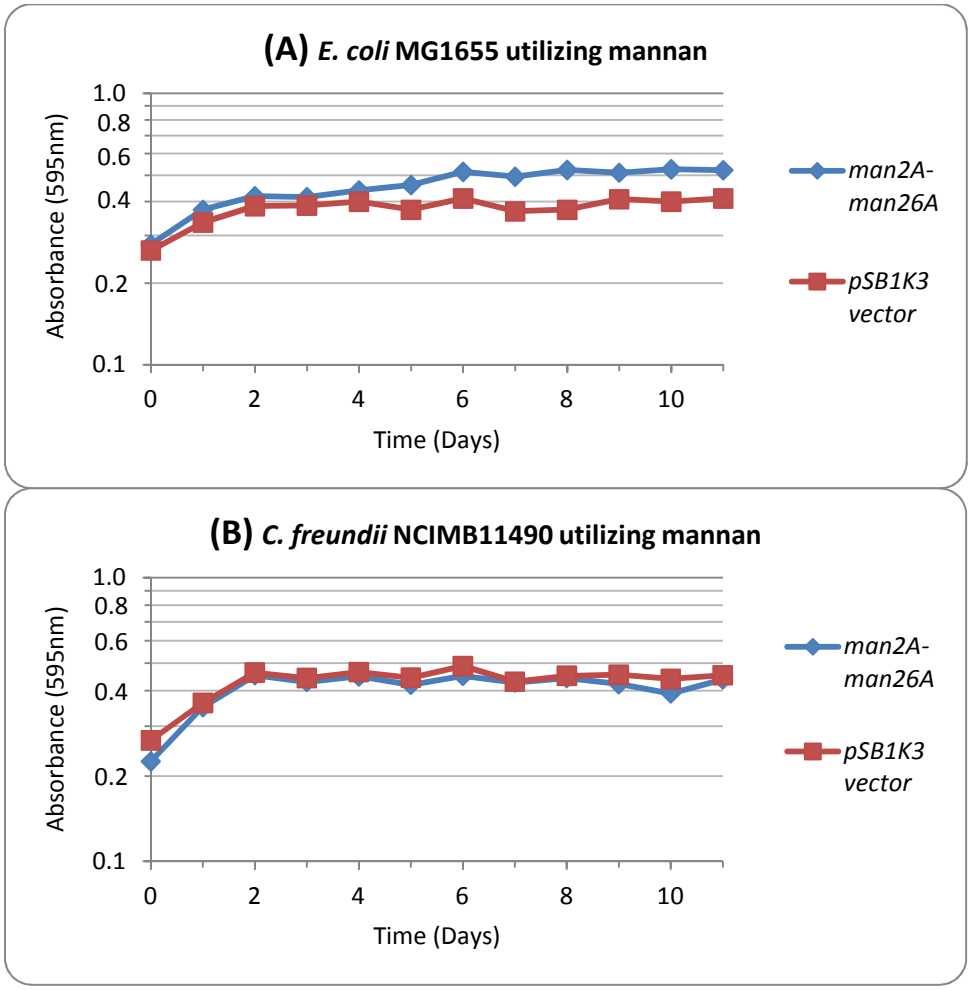

Figure 4.3: Coomassie dye absorbance measured for *E. coli* MG1655 and *C. freundii* NCIMB *man2A-man26A* strains grown in M9 media containing 0.5% mannan. Results show cell growth measured as a function of protein content of *E. coli* MG1655 *man2A-man26A* **(A)** and *C. freundii* NCIMB *man2A-man26A* **(B)** strains in M9 minimal media containing 0.5% manan, 0.34 g/l thiamine and 0.1% w/v yeast extract. Controls in each instance are the respective expression hosts transformed with the pSB1K3 vector.

4.4 Utilization of beechwood xylan

To assay for the ability to utilize xylan of expression hosts transformed with constructs encoding xylanase activity, cultures were set up in M9 medium containing 1% w/v xylan, supplemented with 0.1% w/v yeast extract. Following the same methodologies as previously described, media was inoculated with a cell pellet from overnight strains spun down and resuspended in minimal media to achieve a final OD_{600} of 0.1 at time 0 of the growth assay. Cultures were incubated at 37^{o}C on a rotary shaker at 200 rpm. Over the course of 8 days 100 µl samples were taken at 24 hour intervals and protein content measured by the

Coomassie protein assay. All cultures were performed in duplicate. Results are shown in Figure 4.4.

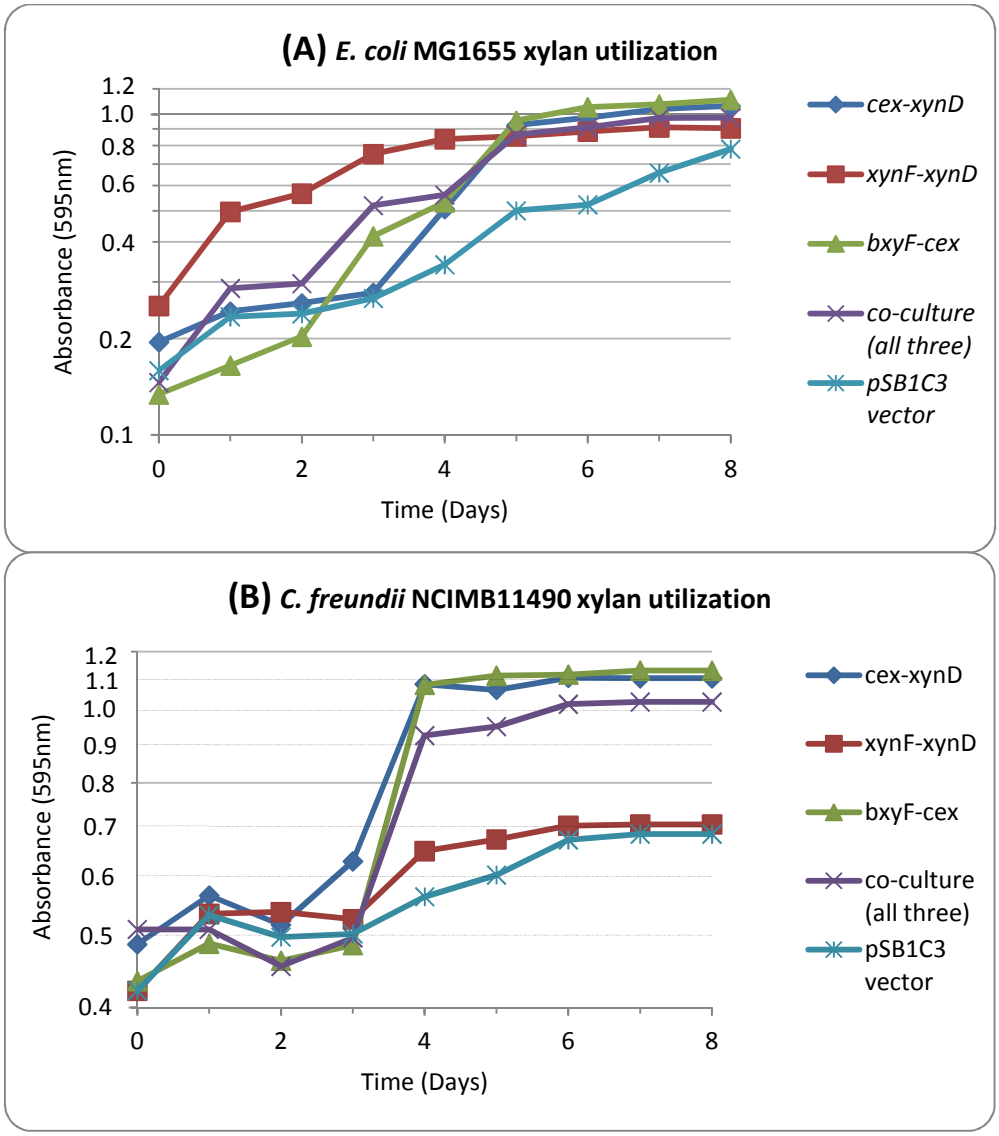

Figure 4.4: Coomassie dye absorbance measured for *E. coli* MG1655 and *C. freundii* NCIMB strains transformed with the xylanase constructs *cex-xynD, xynF-xynD* or *bxyF-cex* grown in M9 media containing 1% xylan. Results show cell growth measured as a function of protein content of *E. coli* MG1655 xylanase strains **(A)** and *C. freundii* NCIMB xylanase strains **(B)** grown in M9 minimal media containing 1% w/v xylan, 0.34 g/l thiamine and 0.1% w/v yeast extract. Controls in each instance are the respective expression hosts transformed with the pSB1C3 vector. **Note:** Ab_{595} results above 1 are considered unreliable as samples were not diluted before readings were recorded.

E. coli MG1655 *xynF-xynD* shows the highest rate in the increase of protein content over the course of the assay, though the highest total protein content is observed in cultures of *bxyF-cex* and *cex-xynD* strains (Fig. 4.4A). A co-culture of all three recombinant *E. coli* xylanase strains however did not show any appreciable difference in protein content compared to the

individual strains, suggesting that synergistic interaction or complementarities between all four xylanases is minimal if at all existent when assayed on beechwood xylan. Protein content of the vector control measured by the Coomassie dye reaches an absorbance of ~0.8 at the end of 8 days, which is comparable to that of the xylanase positive strains. It can argued that the addition of 0.1% w/v yeast extract as a starter carbon source is perhaps too much for the proper assay of xylan utilization and that the addition of less would yield more encouraging results compared to a negative control.

The change in protein content measured for *C. freundii* xylanase strains grown on xylan is by comparison more striking than that of *E. coli*, as a clear exponential phase is observed between days 3 and 4 (Fig. 4.4B). Consistent with that of *E. coli*, cultures of *cex-xynD* and *bxyF-cex* strains as well as the co-culture of all three strains show fairly similar rates of change in protein content. Cultures of *xynF-xynD* and the vector control show an almost identical change in protein content over the course of 8 days. Such results are unexpected as the same construct when expressed in *E. coli* shows clear xylanase activity as growth on xylan is observed.

4.5 Utilization of amorphous cellulose: Carboxymethyl cellulose

4.5.1 Growth assays of *C. freundii* SBS197 in 1% CMC M9 media with 0.01% yeast extract

Initial growth assays measuring OD_{600} of cultures grown in 1% w/v CMC M9 media were conducted with the addition of 0.01% w/v yeast extract. These assays performed measured growth of *C. freundii* SBS197 cellulase strains over the course of 8 days. All cultures were performed in duplicate. Results are shown in Figure 4.5. It can be noted that of the strains cultured, only *cenA-cex* reached an appreciably high cell density when grown on CMC; 0.865 at 8 days. *cenA-cex-cbhA-cbhB* constructs in high and low copy vectors also showed significant growth compared to the vector control, though not as significant as *cenA-cex* as a maximum OD_{600} of only 0.438 was met. It was inferred that the expression of the larger CbhA and CbhB exoglucanases with respective atomic masses of 89 and 115kDa compared to Cex with a mass of 54kDa, places a larger metabolic load upon the host and as such additional starter carbon is required for sufficient growth on CMC. This hypothesis is later confirmed within this chapter. Strains expressing CenB or CenD showed comparable growth to that of the control, a potential reason being that both endoglucanases are noted to have a

lower activity against CMC as evidenced within the literature (Tomme *et al.* 1996), and during characterization assays described within Chapter 3.

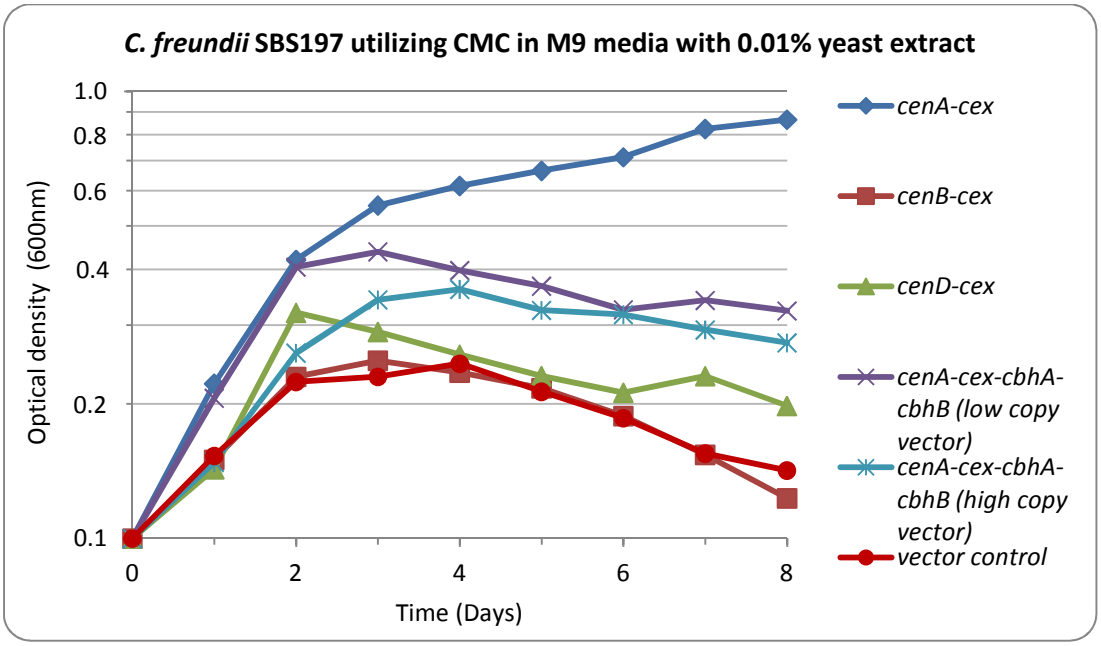

Figure 4.5: Growth of *C. freundii* SBS197 cellulase strains in 1% CMC M9 media containing 0.01% yeast extract. Growth of *C. freundii* SBS197 cellulase strains was measured as OD_{600} in M9 minimal media containing 1% w/v CMC, 0.34 g/l thiamine and 0.01% w/v yeast extract as carbon sources. Control strains were those transformed with the pSB1C3 vector.

4.5.2 Growth assays of *E. coli* MG1655 cellulase strains

Based on the above results, further assays of growth of *E. coli* and *C. freundii* cellulases strains on CMC were conducted in M9 media supplemented with 0.1% w/v yeast extract and 1% w/v CMC so as to encourage initial growth of strains expressing CbhA and CbhB. As described before, cultures were inoculated with cell pellets resuspended in minimal medium for a final OD_{600} of 0.1 at time 0. Growth on CMC was measured as a function of Coomassie dye absorbance over the course of 9 days. Cultures were performed in duplicate. Results are shown in Figure 4.6.

The most efficient strain to utilize CMC was that of *cfbglu-cenA-cex*, with Coomassie dye absorbance measuring 0.478 at day 8 (Fig. 4.6A). Those cultures with strains not expressing an exoglucanase did not show comparable protein content, though in the case of *cfbglu-cenA* and *cfbglu-cenB* Coomassie dye absorbance measured was ~1.7 and ~1.4 times greater than that of the vector control. Figure 4.6B shows results for various co-cultures of the *cfbglu-[endoglucanase]* strains. Strains *cfbglu-cenA* and *cfbglu-cenC* were used to prepare a single

co-culture with the intention of determining whether these two endoglucanases, which are known to be highly active on CMC (Tomme *et al.* 1996), could allow for enhanced utilization of CMC by the host strains. Likewise, co-cultures of *cfbglu-cenB* and *cfbgl-cenD* strains were intended to confirm if the lower activity of CenB and CenD against CMC would impair growth on the substrate. Results suggest that CenA and CenC in concert allow for enhanced growth on CMC as Coomassie dye absorbance measured was ~1.30 and ~1.26 times greater than that of the *cfbglu-cenB/cfbglu-cenD* co-culture and pSB1C3 vector control culture, respectively. However, the expression of an exoglucanase such as Cex seems to give much better utilization of CMC as a carbon source (Fig. 4.6A).

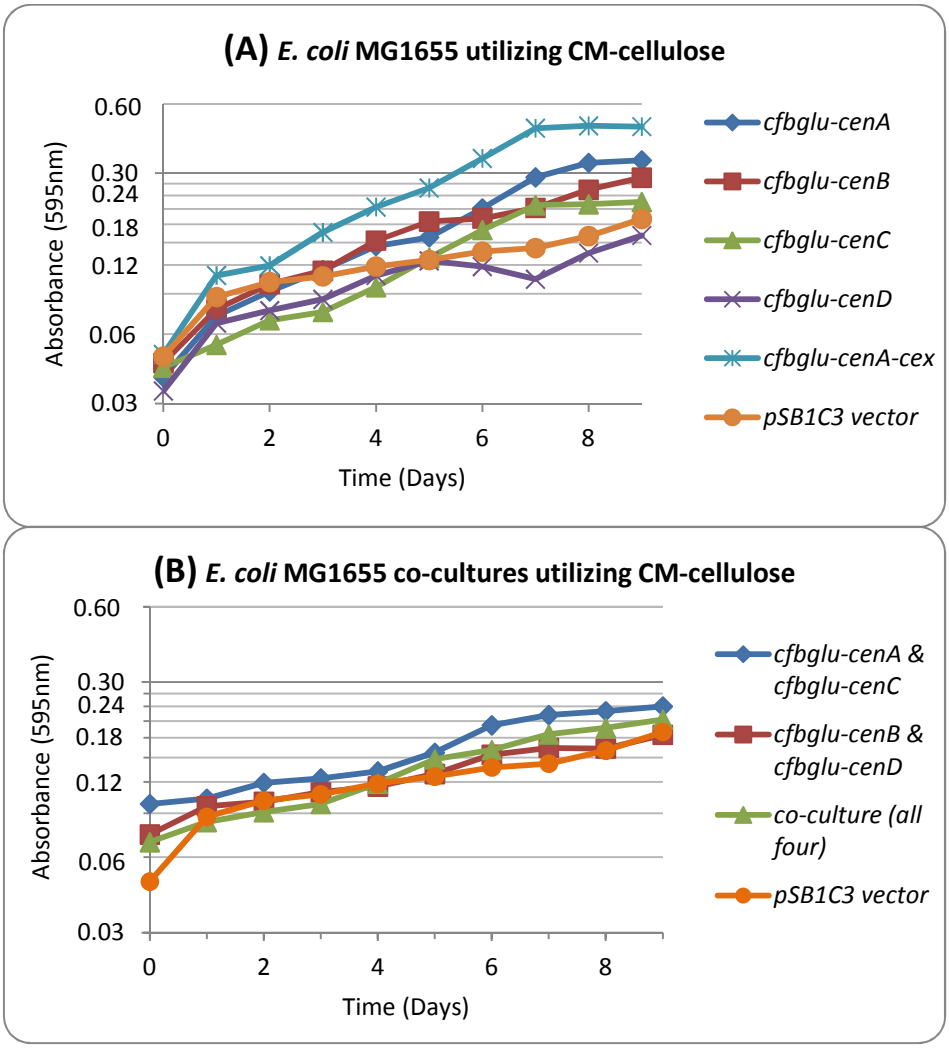

Figure 4.6: Coomassie dye absorbance measured for *E. coli* MG1655 transformed with the *cfbglu-[endoglucanase]* constructs grown in M9 media containing 1% w/v CMC. Results show cell growth measured as a function of protein content of *E. coli* MG1655 *cfbglu-[endoglucanase]* strains **(A)** and *cfbglu-[endoglucanase]* co-cultures **(B)** grown in M9 minimal medium containing 1% w/v CMC, 0.34 g/l thiamine and 0.1% yeast extract. Controls in each instance are *E. coli* MG1655 transformed with the pSB1C3 vector.

E. coli MG1655 *cfbglu-[endoglucanase]-cbhA-cbhB* strains were assayed for growth on CMC using the same methodologies as those previously described for the assay of *cfbglu-[endoglucanase]* strains. These cultures were performed in duplicate. Figure 4.7 shows results for the measurement of Coomassie dye absorbance of *cfbglu-[endoglucanase]-cbhA-cbhB* strains (Fig. 4.7A) and also co-cultures of those strains (Fig. 4.7B).

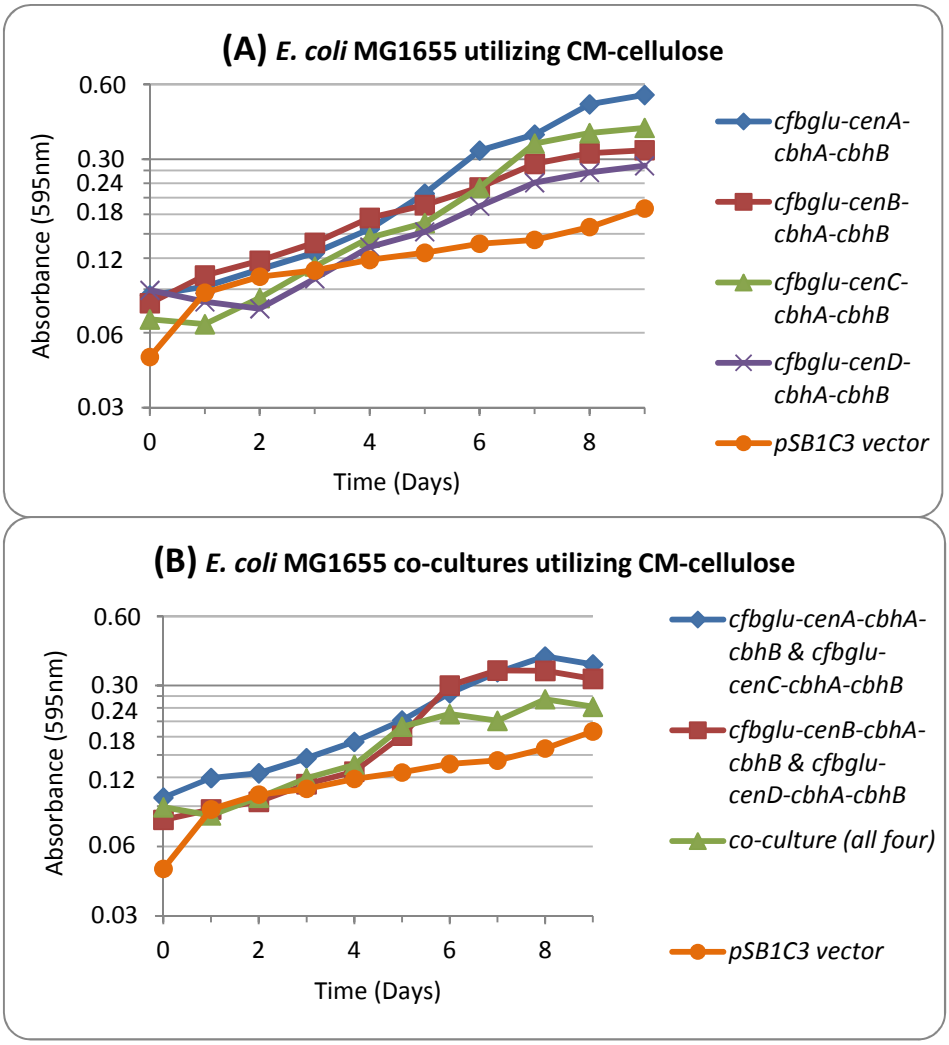

Figure 4.7: Coomassie dye absorbance measured for *E. coli* MG1655 transformed with the *cfbglu-[endoglucanase]-cbhA-cbhB* constructs grown in M9 media containing 1% CMC. Results show cell growth measured as a function of protein content of *E. coli* MG1655 *cfbglu-[endoglucanase]-cbhA-cbhB* strains **(A)** and *cfbglu-[endoglucanase]-cbhA-cbhB* co-cultures **(B)** grown in M9 minimal media containing 1% w/v CMC, 0.34 g/l thiamine and 0.1% yeast extract. Controls in each instance are *E. coli* MG1655 transformed with the pSB1C3 vector.

Results presented in Figure 4.7 suggest that the expression of CbhA and CbhB contributes significantly towards the utilization of CMC by the expression host. Strain *cfbglu-cenA-cbhA-cbhB* showed the highest Coomassie dye absorbance of 0.544 after 9 days, which was followed by that of *cfbglu-cenC-cbhA-cbhB* with 0.401 (Fig. 4.7A). These figures indicate approximately 2.9 and 2.1 times greater protein content than that of the vector control after 9 days incubation. Results for *cfbglu-cenB-cbhA-cbhB* and *cfbglu-cenD-cbhA-cbhB* strains show a higher Coomassie dye absorbance than the vector control, though noticeably less than the *cfbglu-cenA-cbhA-cbhB* or *cfbglu-cenC-cbhA-cbhB* strains. Again, this difference in utilization of CMC may be attributed to the difference between the endoglucanases in their binding preference and catalytic activity against the amorphous substrate (Tomme *et al.* 1996; Warren 1996; Brun *et al.* 2000).

Co-cultures of *cfbglu-cenA-cbhA-cbhB* and *cfbglu-cenC-cbhA-cbhB* strains indicate that CMC is being utilized; however a lower Coomassie dye absorbance measurement than that of a single culture of *cfbglu-cenA-cbhA-cbhB* is evident; 0.402 compared to 0.544 (Fig.4.7B). Perhaps competitive binding for sites on the substrate between CenA and CenC hinders enzymatic productivity, reducing host strain fitness. Interestingly, co-cultures of *cfbglu-cenB-cbhA-cbhB* and *cfbglu-cenD-cbhA-cbhB* show greater protein content than cultures of those individual strains, suggesting a synergistic or at least cooperative activity between CenB and CenD improves utilization of CMC by the host strain (Fig. 4.7B). A co-culture of all four strains shows the least protein content measured.

4.5.3 Growth assays of *C. freundii* cellulase strains

The *C. freundii* NCIMB11490 cellulase strains *[endoglucanase]-cex-cbhA-cbhB* were assayed for their ability to utilize 1% w/v CMC in M9 media containing 0.1% w/v yeast extract. As before, cell pellets of overnight cultures were resuspended in minimal media to give a final OD_{600} of 0.1 at time 0. Samples taken at 24 hour intervals were measured for total protein content using the Coomassie protein reagent dye and the absorbance of which measured at 595 nm. Results are presented in Figure 4.8 below.

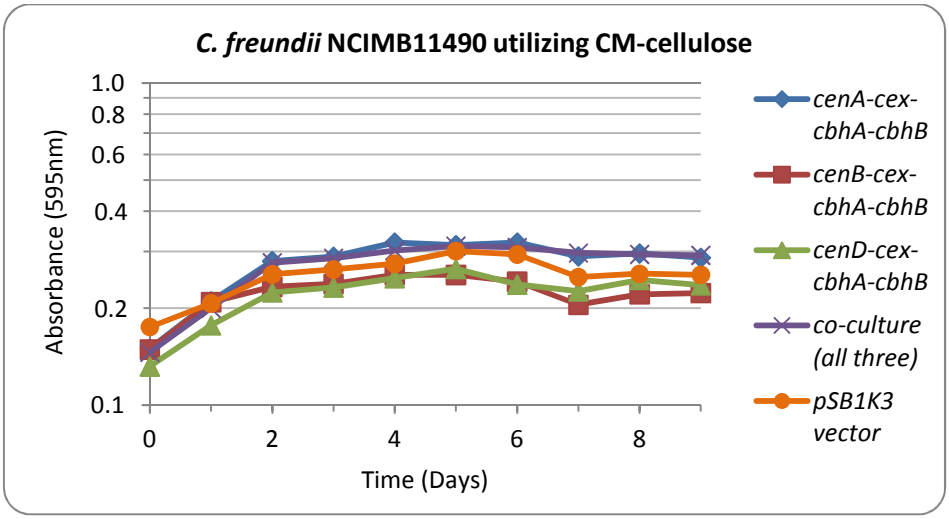

Figure 4.8: Coomassie dye absorbance measured for *C. freundii* NCIMB11490 transformed with the *[endoglucanase]-cex-cbhA-cbhB* constructs grown in M9 media containing 1% CMC. Results show cell growth measured as a function of protein content of *C. freundii* NCIMB *[endoglucanase]-cex-cbhA-cbhB* strains grown in M9 minimal media containing 1% w/v CMC, 0.34 g/l thiamine and 0.1% w/v yeast extract. Control cultures are those of *C. freundii* NCIMB transformed with the pSB1K3 vector.

Assays for measuring growth on CMC of the *C. freundii* NCIMB11490 cellulase strains yielded unexpected results, in that protein content was almost identical if not less than the vector control indicating that the host was unable to utilize CMC (Fig. 4.8). These results are surprising as *E. coli* transformed with the full endoglucanase and exoglucanase complement *cfbglu-[endoglucanase]-cbhA-cbhB* was able to utilize CMC (Fig. 4.7A). Moreover, Chapter 5 of this thesis shows *C. freundii* NCIMB11490 *cenA-cex-cbhA-cbhB* to effectively degrade and utilize cellulose paper. Perhaps differences in the substrate characteristics of CMC are responsible for these observations in *C. freundii* NCIMB11490. Perhaps methylated-glucose inhbits glucose uptake by *C. freundii*, or the low-viscosity of 1% w/v CMC media prevents proper aeration of the culture.

Results for that of *C. freundii* SBS197 are however more encouraging (Figure 4.9). *C. freundii* SBS197 transformed with the constructs *[endoglucanase]-cex* and *cenA-cex-cbhA-cbhB* were assayed for their ability to utilize 1% w/v CMC in M9 media containing 0.1% w/v yeast extract. Figure 4.9A shows that Coomassie dye absorbance measured for *cenA-cex* reached a maximum of 0.75 compared to 0.34 for the vector control, a 2.2 fold increase. This was the highest reading measured, followed by that of the co-culture of all three *[endoglucanase]-cex* strains with an absorbance of 0.665. *C. freundii* SBS197 transformed with *cenA-cex-cbhA-cbhB* in both high and low copy vectors showed broadly similar trends

in protein content measured over the 9 day incubation period, with a protein content of approximately 1.85 times that of the vector control on the 9th day (Fig. 4.9B).

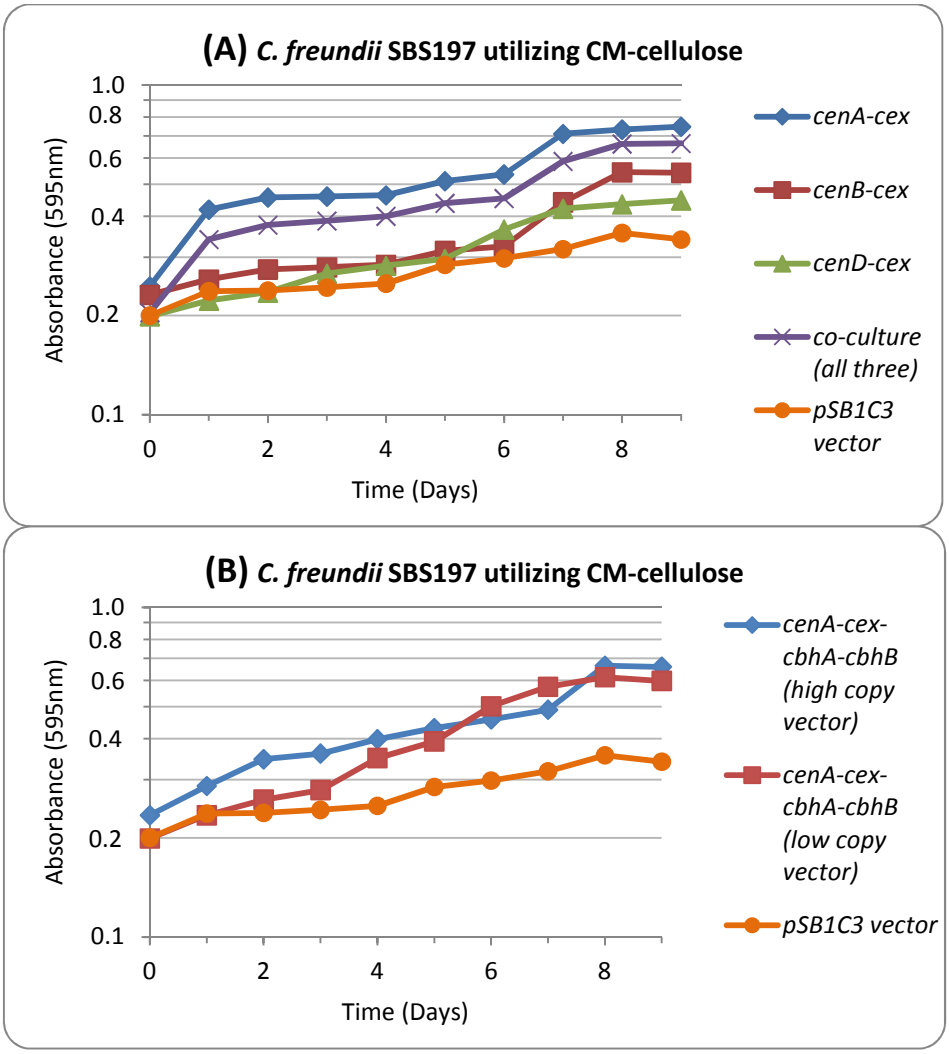

Figure 4.9: Coomassie dye absorbance measured for *C. freundii* SBS197 transformed with the *[endoglucanase]-cex* and *cenA-cex-cbhA-cbhB* constructs grown in M9 media containing 1% CMC. Results show cell growth measured as a function of protein content of *C. freundii* SBS197 *[endoglucanase]-cex* strains **(A)** and *cenA-cex-cbhA-cbhB* strains **(B)** grown in M9 minimal media containing 1% w/v CMC, 0.34 g/l thiamine and 0.1% yeast extract. Controls are those of *C. freundii* SBS197 transformed with the pSB1C3 vector.

4.6 Utilization of crystalline cellulose: Avicel

The use of the Coomassie protein reagent for measuring protein content of cultures with avicel as the carbon source was unsuccessful. Although preliminary trials using *C. fimi* on avicel gave good positive results, the assay method did not translate well to those recombinant hosts of *E. coli* and *Citrobacter freundii*. Most likely due to *C. fimi* being able to

efficiently utilize crystalline cellulose, the signal-to-noise ratio was high enough to effectively cancel any false positive absorbance generated by turbidity from avicel particulates. However in the case of recombinant *E. coli* and *C. freundii*, the level of false positive absorbance generated from avicel particulates prevented effective measurement of protein content. As such, no conclusions could be drawn from the inconsistent data generated and was excluded from the thesis. Instead, the measurement of cell growth on avicel was determined by calculating colony forming units per millilitre of culture. It was the intention to avoid such assay methods as these are not suitable for scaling up to a high-throughput approach, but as it seems, they are necessary in the case of assaying growth on substrates such as avicel.

4.6.1 Colony forming units of *C. freundii* SBS197 in 2% avicel M9 media with 0.01% yeast extract

M9 medium containing 2% w/v avicel and 0.01% w/v yeast extract was inoculated with *C. freundii* SBS197 transformed with one of the following cellulase encoding constructs; *[endoglucanase]-cex* or *[endoglucanase]-cex-cbhA-cbhB.* Cultures of 5 ml volume in 1 oz glass vials were incubated at 37^{o}C on a rotary shaker at 200 rpm over the course of 4 days. 100 µl samples were taken at 24hr intervals and serial dilutions made in distilled water. Dilutions were spread on LB agar plates and incubated at 37^{o}C overnight. Colonies were counted and colony forming units per ml calculated. Cultures were performed in duplicate. Results are shown in Figure 4.10 below.

C. freundii SBS197 cellulase expressing strains showed appreciable growth in avicel media (Fig. 4.10). Notably colony forming units per ml of strain *cenB-cex* reached a maximum of 4.25x10^{8} at day 2. Strain *cenD-cex* showed 3.6x10^{8} CFU per ml at day 2, whereas *cenA-cex* did not give results higher than that of the vector control. An important point to note is the amount of starter carbon added to the media to support initial growth, as this was only 0.01% w/v yeast extract. As shown before in the case of assaying growth on CMC (Fig. 4.5), the addition of 0.01% w/v yeast extract did not support growth of strains expressing both CbhA and CbhB cellobiohydrolases, and as such it can be argued that a higher content of starter carbon is needed to support inital growth. Results presented in Figure 4.10 seem to support this hypothesis, which is further tested in the next section.

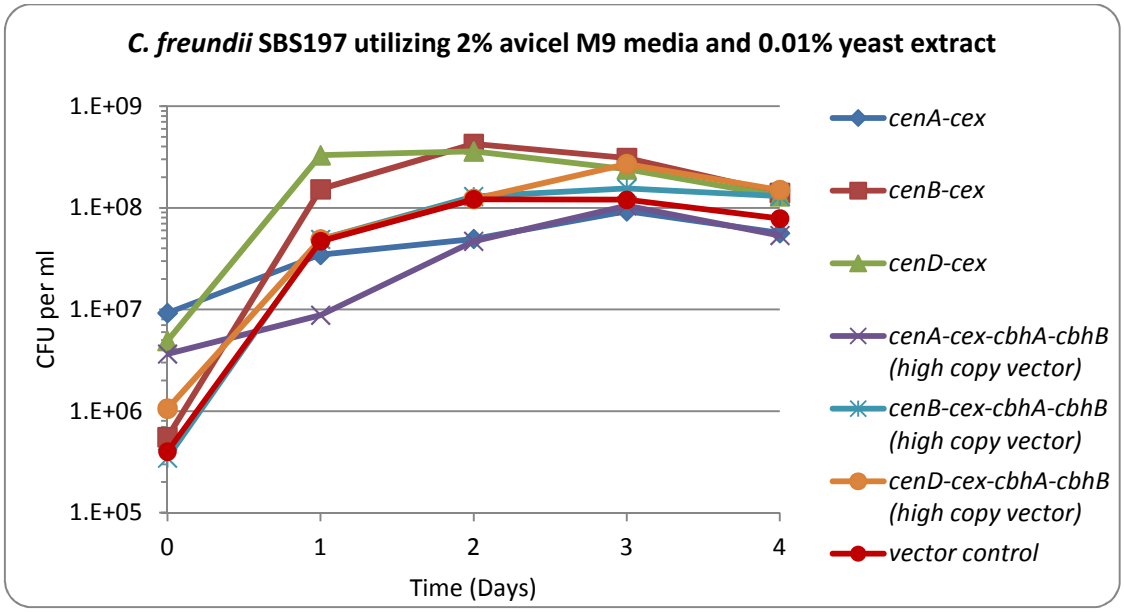

Figure 4.10: Growth of *C. freundii* SBS197 cellulase strains in 2% avicel M9 media with 0.01% yeast extract measured as colony forming units per millilitre. Results show cell growth measured as colony forming units per ml of *C. freundii* SBS197 *[endoglucanase]-cex* strains and *[endoglucanase]-cex-cbhA-cbhB* strains grown in M9 minimal media containing 2% w/v avicel, 0.34 g/l thiamine and 0.01% w/v yeast extract. Control cultures are those of *C. freundii* SBS197 transformed with the pSB1C3 vector.

4.6.2 Colony forming units of *C. freundii* SBS197 in 1% avicel M9 media with 0.1% yeast extract

C. freundii SBS197 cellulase strains grown in M9 media with 1% w/v avicel and 0.1% w/v yeast extract looked to have utilized the substrate more effectively than those grown in 2% w/v avicel media with 0.01% w/v yeast extract. Figure 4.11A shows colony forming units per ml for *[endoglucanase]-cex* strains. Consistent with previous results *cenB-cex* showed the highest CFU per ml with $1x10^9$. This was followed by that of *cenA-cex* and *cenD-cex* with $7.9x10^8$ and $8x10^8$ CFU per ml, respectively. Data for *cenA-cex-cbhA-cbhB* expression in low and high copy vectors as well as a co-culture of the *[endoglucanase]-cex-cbhA-cbhB* strains is shown in Figure 4.11B. Comparing data for *cenA-cex* and *cenA-cex-cbhA-cbhB* strains shows that both achieve approximately $8x10^8$ CFU per ml at days 5 and 4, respectively. Similarly, co-cultures of the *[endoglucanase]-cex* strains compared to co-cultures of the *[endoglucanase]-cex-cbhA-cbhB* strains also show comparable results as these reach approximately $8x10^8$ CFU per ml at days 5 and 4, respectively. Such findings suggest that in these conditions, the expression of the cellobiohydrolases do not significantly affect the

ability of the host to utilize avicel compared to that of a host expressing solely Cex as an exoglucanase.

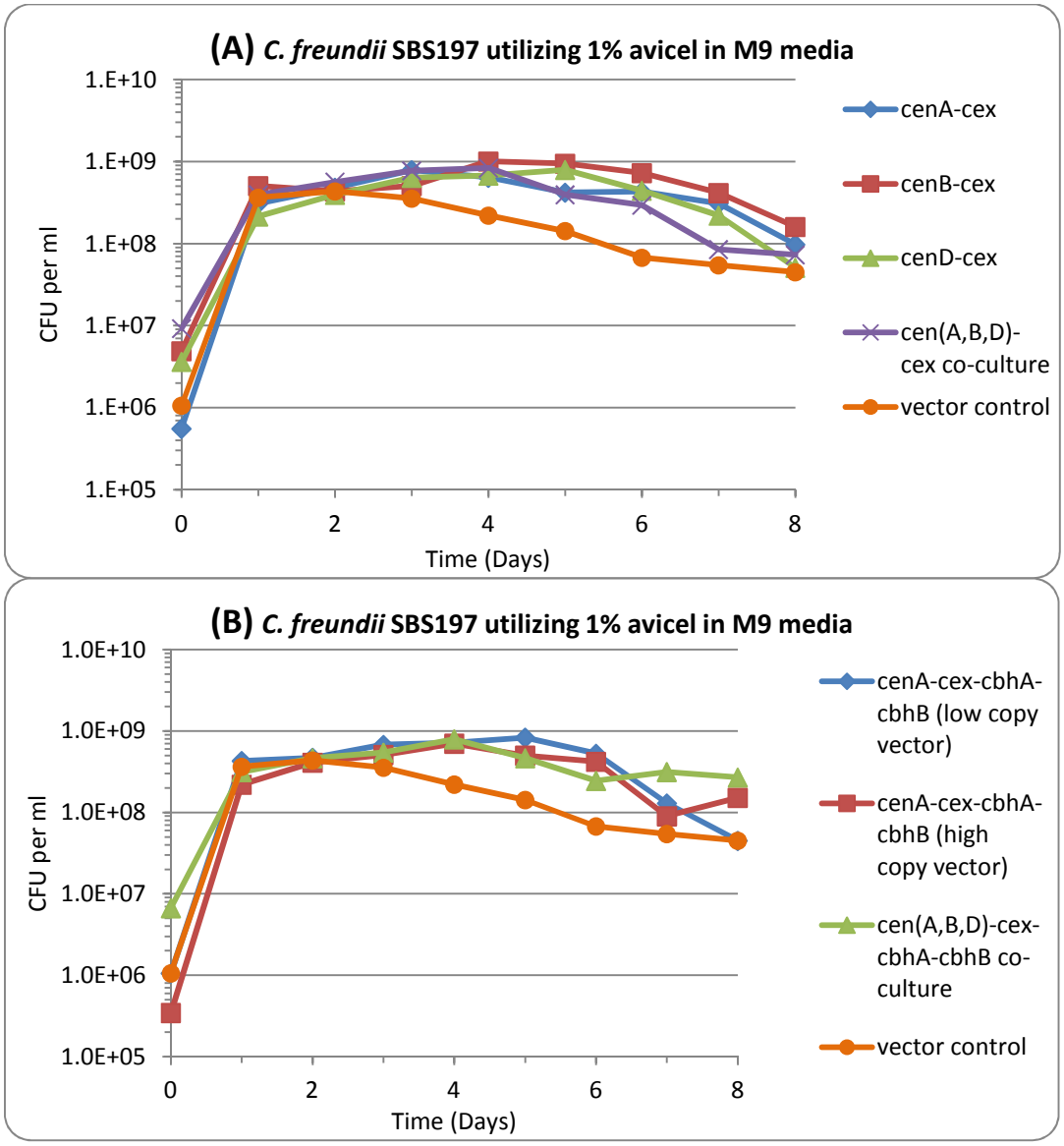

Figure 4.11: Growth of *C. freundii* SBS197 cellulase strains in 1% avicel M9 media with 0.1% w/v yeast extract measured as colony forming units per millilitre. Results show cell growth measured as colony forming units per ml of *C. freundii* SBS197 *[endoglucanase]-cex* strains **(A)** and *[endoglucanase]-cex-cbhA-cbhB* strains **(B)** grown in M9 minimal media containing 1% w/v avicel, 0.34 g/l thiamine and 0.1% w/v yeast extract. Controls are those of *C. freundii* SBS197 transformed with the pSB1C3 vector.

Data presented here shows the *C. freundii* SBS197 cellulase strains to be more effective at utilizing avicel in M9 media containing 1% w/v avicel and 0.1% w/v yeast extract compared to 2% w/v avicel and 0.01% w/v yeast extract. It is possible that the reduction in avicel content from 2% to 1% avicel may improve initial loading of cellulases onto the substrate and

enhance the concerted action of endo- and exo-glucanases, improving the rate of hydrolysis of avicel and subsequent viability of the host cell. Moreover increasing the concentration of yeast extract from 0.01% to 0.1% may be necessary for the expression of the larger constructs, so as to support the initial growth until sufficient substrate is degraded to support further growth. A final point to consider regarding media composition within these assays is the fact that a population decline is evident within each strain assayed. This decline is most evident after days 4 and 5 where the colony forming units per ml drop in some cases by a factor of 2. It has been discussed within our lab that this may be the result of *C. freundii* producing excess acid, a characteristic which is observed when the bacterium utilizes glucose but interestingly not cellobiose (data not shown). As such, the addition of suitable buffering agents may be required to remedy this.

4.7 Discussion

Introduced within this chapter is the Coomassie protein assay which is used to determine cell growth as a function of protein content for recombinant strains utilizing cellulose, mannan and xylan substrates. Preliminary assays using the reagent on *Cellulomonas fimi* cultures grown on those cellulosic substrates highlighted several considerations within the protocol; (i) the assay is indiscriminate of metabolically active cells and dead cells, (ii) substrates with high particulate content can skew absorbance readings, (iii) and the protein content measured across different bacterial species can vary at a given cell optical suggesting results cannot be fairly compared across different expression hosts.

Further to this, assays of recombinant *E. coli* and *Citrobacter freundii* cellulase strains were conducted so as to optimize media compositions for growth on cellulosic substrates. Results showed that in some instances growth of the negative vector control is too great, or that conversely more yeast extract is beneficial for expression of the larger cellulase encoding constructs. As such, the amount of yeast extract supplemented in the growth media would ideally need to be defined according to the needs of the expression host. These observations suggest the use of a more defined carbon source such as glycerol is required so that the initial carbon present within the media can be calculated and from there determine the optimum starter carbon necessary to support initial growth. Although the vitamins and amino acids provided for by the addition of yeast extract are likely to aid in supporting the inital growth of the culture and as such should not be disregarded.

Such considerations are important for the proper assay of our recombinant *E. coli* and *C. freundii* strains for the utilization of lignocellulosic substrates. In the case of recombinant *E. coli* and *C. freundii* strains transformed with *man2A-man26A*, it was shown that although both expression hosts are able to utilize mannose (Holt and Bergey 2000), only *E. coli* expressing *man2A-man26A* is able to utilize mannan. A possible reason for this is that perhaps a reduced level of yeast extract is needed to properly differentiate growth between the *C. freundii* vector control and test strains, thus a defined starter carbon source such as glycerol is optimal.

Within this chapter, recombinant strains of *E. coli* MG1655 and *C. freundii* NCIMB11490 transformed with the xylanase encoding constructs were shown to grow in M9 media supplemented with 1% w/v beechwood xylan. The highest levels of protein content measured by the Coomassie protein reagent in both expression hosts were that of the strains *cex-xynD* and *bxyF-cex*. Interestingly, although *E. coli* strains expressing *xynF-xynD* were able to utilize xylan, *C. freundii xynF-xynD* was not. A possible reason for this is that the XynF and XynD xylanases are not expressed in an active form by *C. freundii*, and further parts characterization in *C. freundii* as an expression host is required.

In assaying for growth of *E. coli* and *C. freundii* cellulase strains on CMC it was noted that in order for the initial growth of strains expressing CbhA and CbhB, 0.01% w/v yeast extract as a carbon source was insufficient. As such, assays were instead conducted in M9 media supplemented with 0.1% w/v yeast extract. Assays for *E. coli* MG1655 cellulases strains showed that for appreciable growth on CMC the expression of at least one exoglucanase is required. In accordance with the literature which describes CenA and CenC to be the most active on CMC (Tomme *et al.* 1996), cultures of our strains expressing either CenA or CenC grown on CMC showed the highest protein content. Interestingly though, co-cultures of strains expressing CenB and CenD showed comparable growth to co-cultures of CenA and CenC expressing strains, suggesting a synergistic or cooperative activity between CenB and CenD on CMC.

Results for assays conducted for the growth of *C. freundii* NCIMB11490 cellulase strains on CMC were shown to be comparable to that of the vector control, indicating that CMC was not effectively utilized as a carbon source. Results for *C. freundii* SBS197 cellulase strains were however more encouraging with *cenA-cex* showing 2.2 times greater absorbance generated by the Coomassie dye compared to the vector control. Strains transformed with the

cenA-cex-cbhA-cbhB construct in either a high or low copy vector showed similar growth on CMC suggesting that the copy number is not a limiting factor in cellulase expression and utilization of CMC.

Attempting to measure utilization of avicel by our recombinant cellulase strains proved to be unsuccessful when using the Coomassie protein reagent. Although successful when measuring that of *C. fimi* avicel utilization, data for *E. coli* and *C. freundii* strains was highly variable and inconclusive. It was deduced that the comparably less efficient cellulase systems designed for expression in our expression hosts did not support the generation of high protein content in those cultures. Therefore the absorbance of Coomassie dye measured at 595 nm was comparatively lower and so more likely affected by turbidity generated by avicel particulates, leading to highly variable absorbance measurements. As such, cell growth in cultures with avicel as a carbon source was measured by calculating colony forming units per millilitre, rather that the Coomassie protein reagent.

In assaying *C. freundii* SBS197 cellulase strains for their ability to utilize avicel, it was noted that 0.01% yeast extract was insufficient to support initial growth and so medium containing 0.1% yeast extract was used instead. Of those strains assayed, *cenB-cex* showed the highest growth with a maximum of 1.01×10^9 CFU per ml met at day 4. By comparison *cenA-cex* strains showed slightly less growth with 7.9×10^8 CFU per ml at day 3, results which are consistent with the literature describing CenB as an endoglucanase with preferential activity against crystalline cellulose and that of CenA with a preference against amorphous cellulose. Moreover it was noted that the expression of the cellobiohydrolases in concert with CenA did not significantly improve growth on avicel compared to that of CenA expressed in concert with the single exoglucanase Cex, as CFU per ml measured was comparable. These observations in combination with those assays conducted in CMC media suggest that the cellobiohydrolases are best suited to processing amorphous cellulose, and perhaps in order to process crystalline cellulose require the action of an additional agent capable of inducing amorphogenesis.

Regarding the media composition of cultures assayed within those experiments containing avicel, several amendments could be made to improve experimental accuracy. Firstly, it was observed that growth in M9 containing 1% w/v avicel was higher than that in M9 containing 2% w/v avicel. A phenomenon that could be attributed to an enhanced loading of cellulases onto the substrate, as the substrate surface area within the media is effectively reduced,

resulting in improved complementary activity between those cellulases acting in concert. Therefore reducing the avicel content within the media could improve productivity of the recombinant hosts. A second amendment would be the addition of suitable buffering agents to the media to improve buffering capacity. It has been observed within the lab that *C. freundii* SBS197 produces high levels of acid during metabolism of glucose, which could be the reason for the apparent population decline observed around days 4 and 5. Making such changes to carbon content and buffering capactiy in the culturing medium could improve the viability of our recombinant hosts and so facilitate the further study of those cellulase encoding constructs already assembled.

Chapter 5

Deconstruction and utilization of cellulose paper

Summary

To complement those growth assays on defined substrates in Chapter 4, further assays for growth were conducted using cellulose paper as a carbon source. This substrate was chosen as the deconstruction of the material can easily be observed, so as to highlight any differences in the disruptive capacity of the various expression hosts on the substrate. Of the recombinant *E. coli* cellulase strains assayed, a co-culture of two strains expressing either endoglucanase CenA or CenC, in addition to the two cellobiohydrolases CbhA and CbhB showed the highest protein content as measured by the Coomassie assay. A co-culture of two strains expressing either endoglucanase CenB or CenD in addition to the cellobiohydrolases showed appreciable synergistic cooperation compared to cultures of the individual strains. This observation is consistent with results presented in Chapter 4 where a co-culture of CenB and CenD expressing strains were shown to utilize CMC in a synergistic manner. Growth assay results for *C. freundii* NCIMB11490 cellulase strains showed that each strain was capable of utilizing cellulose paper with the highest efficiency being those expressing the endoglucanase CenA. Similar results were obtained for *C. freundii* SBS197 cellulase strains; though a higher effective utilization of the substrate was observed for NCIMB11490 suggesting it is a more suitable expression host of the two. Images taken of cultures of recombinant expression hosts utilizing cellulose paper illustrate the deformation of the substrate over the course of a nine day culturing period. Those strains expressing the endoglucanase CenA had shown the most efficient utilization of cellulose paper and in accordance also showed the highest destruction of the substrate. Though notably some exceptions to this were observed where substrate utilization and deconstruction were not synonymous. Within this chapter, cellulose paper utilization and deconstruction by the *C. freundii* cellulase strains is shown to be unaffected when conducted in saline media. Both NCIMB11490 and SBS197 cellulase strains showed identical growth on the substrate in media containing 0.6 M sodium chloride as compared to the respective cultures performed in media lacking sodium chloride. Wildtype *Cellulomonas fimi* growth on cellulose paper is shown to be completely inhibited in saline media. These results highlight the novelty in developing *C. freundii* as a bioprocessing host capable of culturing in sea water, as a means of reducing the demand on fresh water resources and improving utilization of marine biomass.

5.1 Introduction

Recombinant *E. coli* and *C. freundii* cellulase strains assayed for utilization of amorphous and crystalline substrates within Chapter 4 were assayed for the ability to utilize pure cellulose paper as a carbon source. This substrate was chosen for further assay experiments as its deconstruction can be verified without the need for additional experimental techniques and also allows for conclusions to be drawn regarding the destructive potential of the different cellulase strains on cellulosic biomass. Results for the utilization and deconstruction of cellulose paper are discussed within this chapter.

5.2 Utilization of cellulose paper as a carbon source

The composition and preparation of M9 media for culturing of cellulase strains utilizing cellulose paper was as described in Chapter 4. M9 medium supplemented with 0.1% w/v yeast extract was inoculated with cell pellets of overnight cultures of those cellulose strains to be tested. Cell pellets were resuspended in M9 medium to give a final OD_{600} of 0.1 within 5 ml 1 oz glass vials. As a source of carbon, four individual cellulose paper squares measuring 0.5 cm^2 with a combined mass of 125 mg were added to each culture. Cultures were incubated at 37^oC on a rotary shaker at 200 rpm over a 9 day period. All cultures were performed in duplicate.

5.2.1 Coomassie protein assays for *E. coli* MG1655 cellulase strains utilizing cellulose paper

E. coli MG1655 cellulase strains *cfbglu-[endoglucanase]-cbhA-cbhB* and *cfbglu-cenA-cex* were used to inoculate M9 medium containing 125 mg cellulose paper. Co-cultures of *cfbglu-[endoglucanase]-cbhA-cbhB* strains were also prepared and assayed for growth on cellulose paper. Results are presented in Figure 5.1.

Of those individual strains assayed, *cfbglu-cenA-cbhA-cbhB* showed the highest protein content as measured by the absorbance of Coomassie dye (Fig. 5.1A). An absorbance of 0.395 was recorded at day 9 for that strain, compared to 0.295 for that of the vector control. Notably, the remaining cellulase strains showed a final protein content that was very similar to that of the vector control at day 9. However each strain also showed considerably higher protein content between days 5 and 8 compared to the vector control. This is best exemplified by strains *cfbglu-cenA-cex* and *cfbglu-cenC-cbhA-cbhB* which show respectively a 1.97 and 1.99 fold greater absorbance in Coomassie dye measured at day 6 compared to the vector

control. The fact that a greater rate of protein generation is measured in the cellulase strains compared to the vector indicates that cellulose paper is being utilized, though the culture may not be stable as the rate of protein generation levels off until the final content at day 9 is comparable to that of the control.

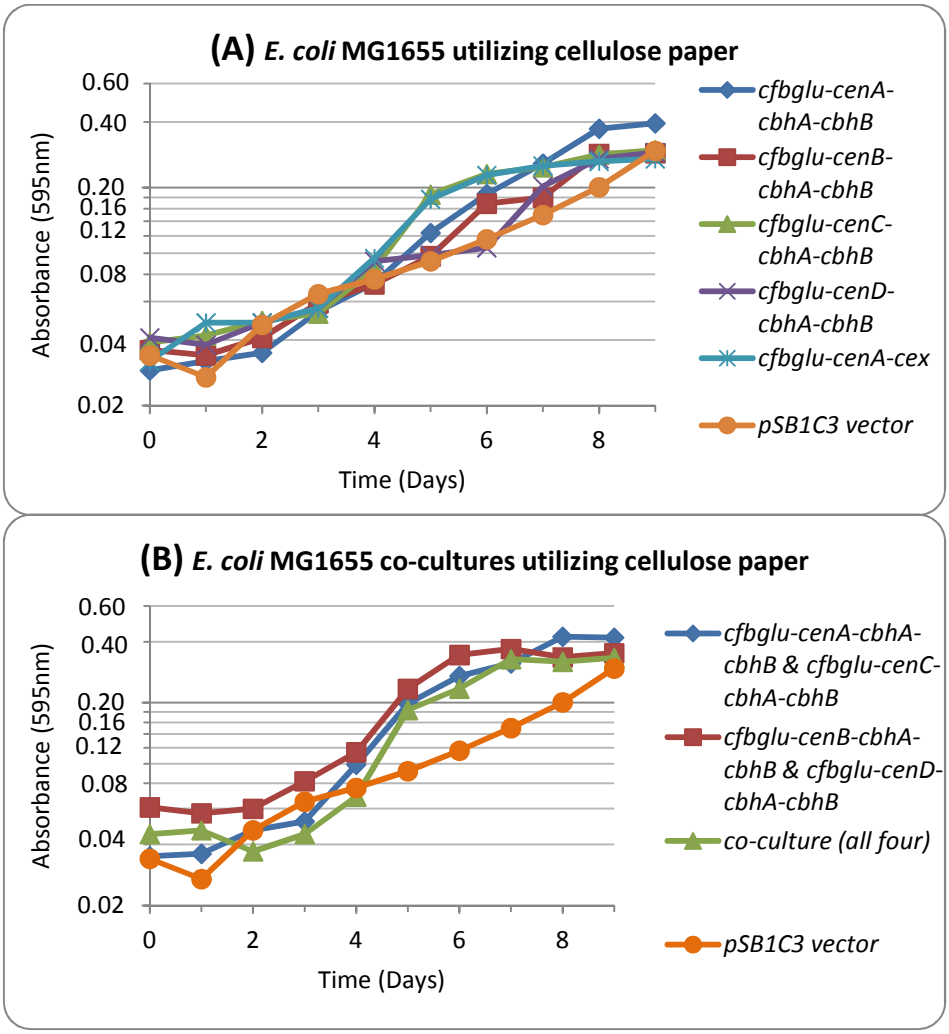

Figure 5.1: Coomassie dye absorbance measured for cultures of *E. coli* MG1655 cellulase strains in 5 ml M9 medium containing 125 mg cellulose paper. Cellulase strains *cfbglu-[endoglucanase]-cbhA-cbhB* and *cfbglu-cenA-cex* were assayed for the ability to utilize 125 mg cellulose paper in 5 ml M9 minimal medium containing 0.34 g/l thiamine and 0.1% yeast extract, by measurement of protein content using the Coomassie protein reagent **(A).** Co-cultures of those strains were then assayed for the utilization of cellulose paper so as to study the existence of any synergistic or cooperative effects between endoglucanases for the enhanced utilization of the substrate **(B).**

Co-cultures of the *E. coli* cellulase strains yielded interesting results in that protein content measured was markedly greater than that of the individual strains (Fig. 5.1B). Notably, a co-culture of strains expressing CenB and CenD showed a maximal absorbance measured for Coomassie dye reagent of 0.368, compared to 0.288 and 0.290 for the individual strains expressing CenB or CenD, respectively. The co-culture of CenB and CenD expressing strains showed a maximal absorbance 2.95 times greater than the vector control at day 6, whereas the individual strains did not show an absorbance more than 1.46 times that of the vector. Similarly, a co-culture of strains expressing CenA and CenC reached a maximal absorbance for Coomassie dye of 0.423, compared to 0.395 and 0.297 for individual strains expressing CenA or CenC, respectively. These results indicate that there may be cooperative or synergistic actions at play between the endoglucanases during the utilization of cellulose paper. Interestingly, a co-culture of all four endoglucanase strains does not measure as high a protein content compared to the previous co-cultures. This may due to competitive binding between the endoglucanases for sites on the substrate, limiting overall enzymatic productivity.

5.2.2 Coomassie protein assays for *C. freundii* cellulase strains utilizing cellulose paper

C. freundii NCIMB11490 cellulase strains *[endoglucanase]-cex-cbhA-cbhB* were assayed for the ability to utilize cellulose paper. A co-culture of all three strains was also assayed. Results are shown in Figure 5.2. Cellulase strain *cenA-cex-cbhA-cbhB* showed the highest maximal protein content measured with an absorbance of 0.456 at day 9. This was followed by the co-culture of all three strains with an absorbance of 0.321, compared to 0.151 for the pSB1K3 vector control. These results are encouraging as each strain showed at least 1.9 times greater absorbance measured at the end of the 9 day incubation compared to the vector control. Moreover each strain was shown to be capable of utilizing cellulose paper, whereas in Chapter 4 the same strains were unable to utilize CMC. No synergistic effects were noted within the co-culture as total protein content measured was not appreciably greater than that of the individual strains.

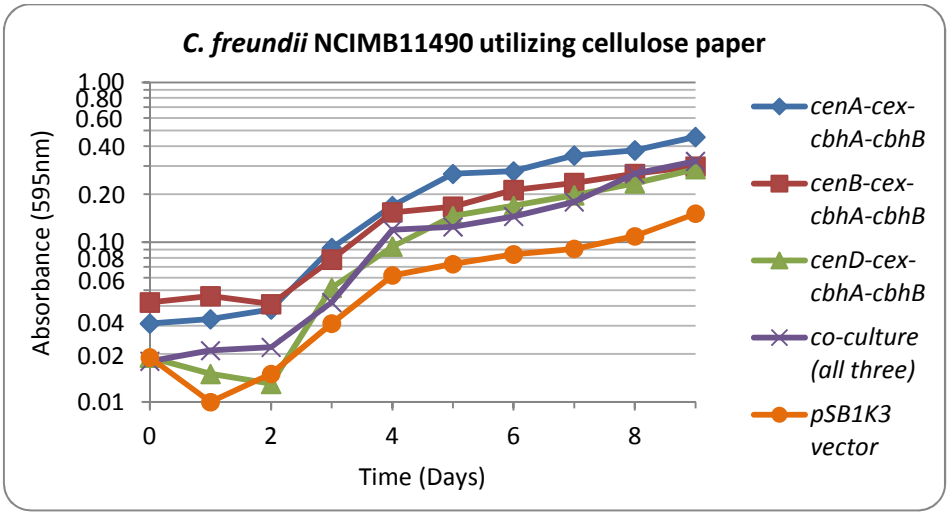

Figure 5.2: Coomassie dye absorbance measured for cultures of *C. freundii* NCIMB11490 cellulase strains in 5 ml M9 media containing 125 mg cellulose paper. Cellulase strains *[endoglucanase]-cex-cbhA-cbhB* are assayed for the ability to utilize 125 mg cellulose paper in 5 ml M9 minimal medium containing 0.34 g/l thiamine and 0.1% yeast extract, by measurement of protein content using the Coomassie protein reagent**.** A co-culture of those strains was also assayed so as to study the existence of any synergistic or cooperative effects between endoglucanases for the enhanced utilization of the substrate.

C. freundii SBS197 cellulase strains were assayed for utilization of cellulose paper. SBS197 transformed with *[endoglucanase]-cex* or *cenA-cex-cbhA-cbhB* constructs were used to inoculate M9 minimal medium supplemented with 0.1% w/v yeast extract and 125 mg cellulose paper. In accordance with previous methods, cultures were assayed for protein content using the Coomassie protein assay at regular 24 h intervals over the course of 9 days. Cultures were conducted in duplicate and results are presented in Figure 5.3.

Highest protein content observed was that of *cenA-cex* strains with an absorbance measured for Coomassie dye of 0.611 at day 9, compared to 0.364 for the vector control (Fig. 5.3A). Strains *cenB-cex* and *cenD-cex* also showed greater protein content compared to the vector control on cellulose paper, as did the co-culture of all three strains. Strains transformed with *cenA-cex-cbhA-cbhB* constructs in either low or high copy vectors showed similar rates of change in protein content over the 9 day incubation period, with respective maximal absorbances of 0.493 and 0.550 recorded at day 9 (Fig. 5.3B). Results show that *C. freundii* SBS197 cellulase strains are able to utilize cellulose; however the *cenA-cex-cbhA-cbhB* strain shows only 1.51 times greater protein content compared to the control whereas *C. freundii* NCIMB11490 transformed with the same construct shows 3.02 times greater protein content

than that of the respective vector control. Ultimately, it may be interpreted from this data that *C. freundii* NCIMB11490 is a more effective host chassis for cellulase expression and cellulose paper utilization when comparing recombinant strains against the respective vector controls.

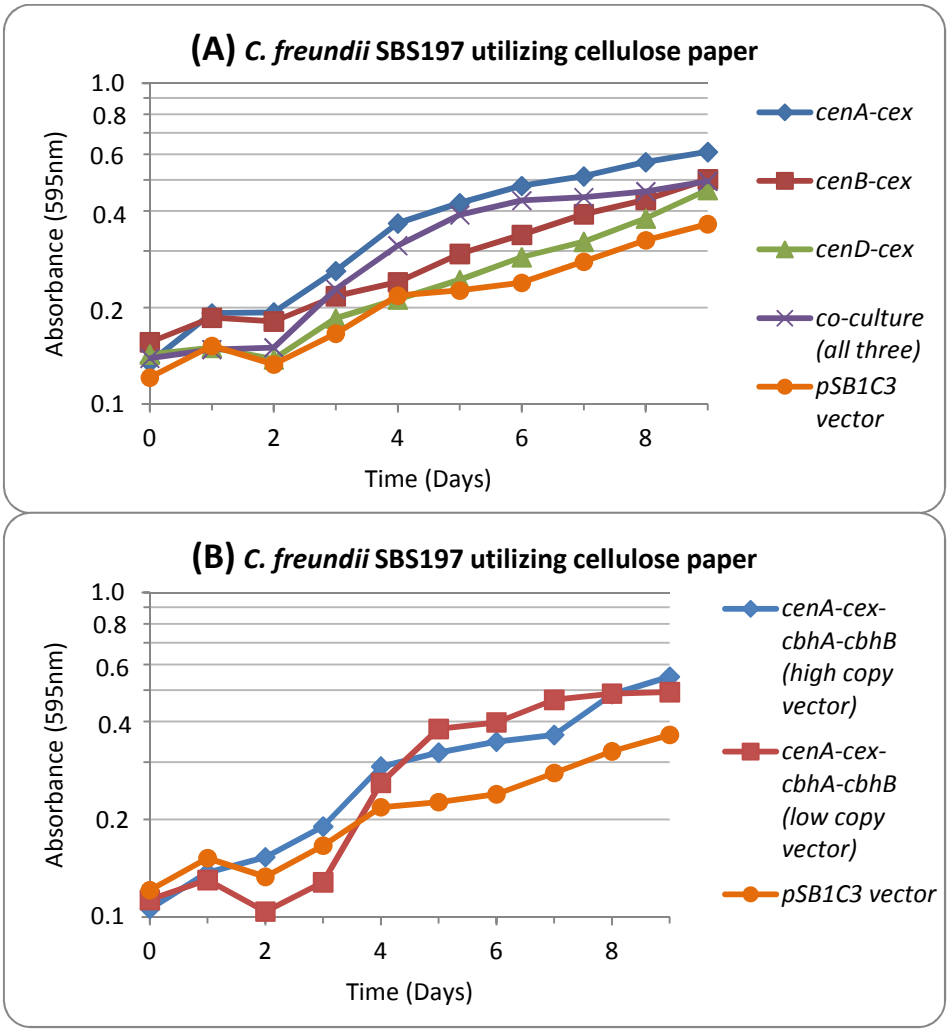

Figure 5.3: Coomassie dye absorbance measured for cultures of *C. freundii* SBS197 cellulase strains in 5 ml M9 medium containing 125 mg cellulose paper. Cellulase strains *[endoglucanase]-cex* were assayed for the ability to utilize 125 mg cellulose paper in 5 ml M9 minimal medium containing 0.34 g/l thiamine and 0.1% yeast extract, by measurement of protein content using the Coomassie protein reagent**.** A co-culture of those strains was also included **(A).** Cellulase strain *cenA-cex-cbhA-cbhB* in both high and low copy vectors was assayed for the ability to utilize cellulose paper compared to the pSB1C3 vector control **(B).**

5.2.3 Colony forming units for *C. freundii* SBS197 cellulase strains utilizing cellulose paper

To complement assays for determining cell growth on cellulose paper by use of the Coomassie assay reagent, cultures were also measured for colony forming units per millilitre

(CFU per ml). Medium preparation and composition was the same as that of the cultures assayed using the Coomassie protein reagent. Cultures were performed in duplicate, and results are presented in Figure 5.4 below.

C. freundii SBS197 cellulase strain *cenA-cex* showed the highest CFU per ml with $1x10^9$ measured at day 3, which was followed by a co-culture of all three strains *cen(A,B,D)-cex* with $7x10^8$ CFU per ml measured at day 4 (Fig. 5.4A). As noted with previous colony counts measured for *C. freundii* SBS197 strains utilizing avicel in Chapter 4, a population decline is observed. Results for strains transformed with the *cenA-cex-cbhA-cbhB* constructs in high and low copy vectors show a comparable colony forming count with maxima of $9.1x10^8$ and $9.9x10^8$ CFU per ml, respectively.

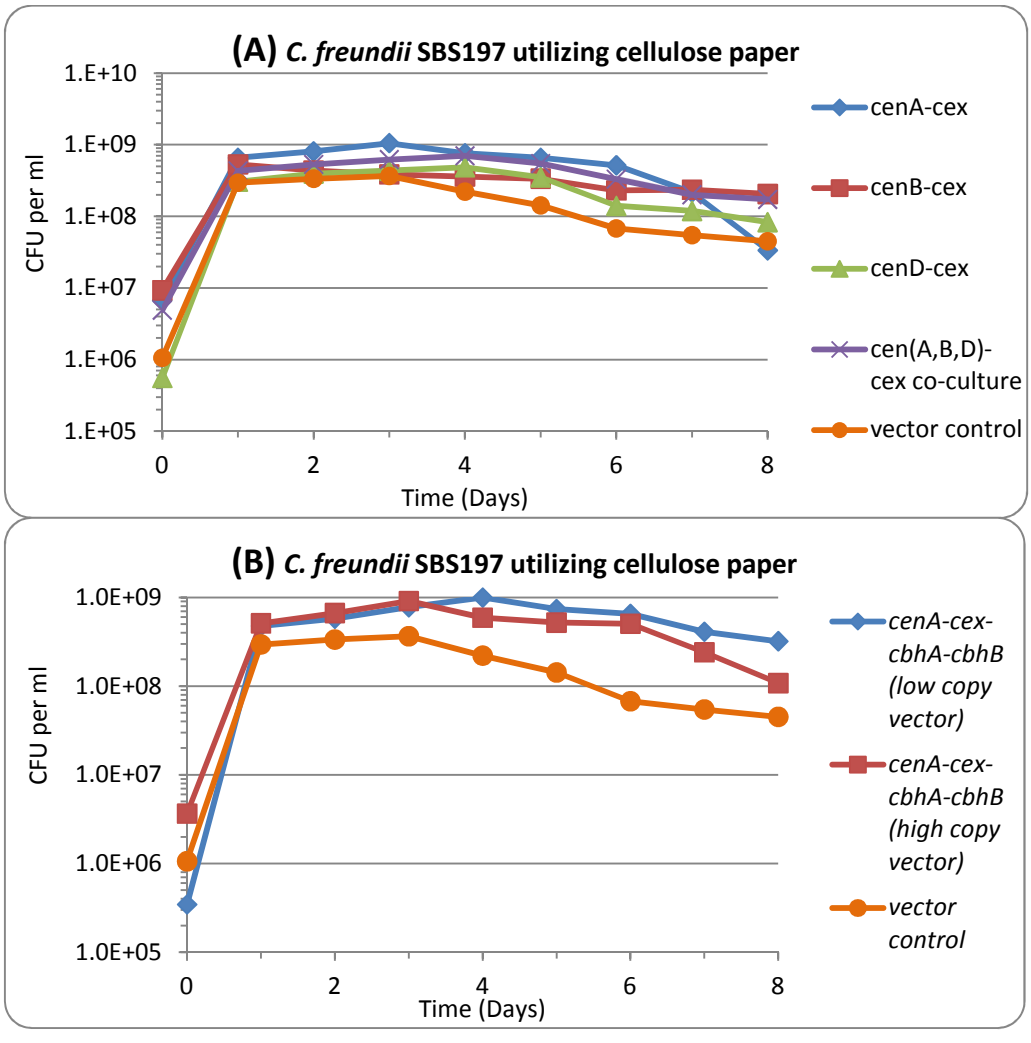

Figure 5.4: Growth of *C. freundii* SBS197 cellulase strains on cellulose paper measured as colony forming units per millilitre. Cultures of *C. freundii* SBS197 cellulase strains were assayed for growth in 5 ml volumes of M9 minimal medium supplemented with 0.1% w/v yeast extract, 0.34 g/l thiamine and 125 mg cellulose paper over 8 days. Cell growth was measured as a function of colony forming units per millilitre.

5.3 Deconstruction of cellulose paper

As noted previously, the choice of cellulose paper as an additional cellulosic substrate for growth assays was intended to facilitate observations made in the deconstruction of the substrate over the course of the assay. Within this section, images of cultures of recombinant expression hosts grown on cellulose paper are presented and discussed. In addition, results for our recombinant cellulase strains are compared against a wildtype *C. fimi* benchmark. All cultures were conducted in M9 medium supplemented with 0.1% w/v yeast extract and 125 mg cellulose paper. *E .coli* and *C. freundii* strains were cultured at 37^{o}C to achieve optimum cell growth, whilst wildtype *C. fimi* cultures are conducted at the species' optimum growth temperature of 30^{o}C.

5.3.1 Extent of cellulose paper deconstruction after a 9 day culturing period

Pictures of recombinant *E. coli* MG1655, *C. freundii* NCIMB11490 and SBS197 cellulase strains at the end of a 9 day culturing period are presented in Figures 5.6, 5.7 and 5.8 respectively. Cultures were conducted in 1 oz glass vials containing 5 ml of M9 medium supplemented with 0.1% w/v yeast extract and four 0.5 cm^2 cellulose paper squares. Flasks were laid horizontally on a light-box and images taken with a digital camera.

Figure 5.5 illustrates the extent of cellulose paper deconstruction by *E. coli* MG1655 cellulase strains *cfbglu-[endoglucanase]-cbhA-cbhB* and *cfbglu-cenA-cex*. It can be observed that the strain *cfbglu-cenA-cbhA-cbhB* shows the greatest extent of deconstruction at the end of 9 days as evidenced by the presence of small fibrous particles of cellulose paper (Fig. 5.5ii). By comparison the vector control shows no rounding of the cellulose paper squares or the presence of any particulate matter (Fig. 5.5i). These results are consistent with results observed for the Coomassie protein assay which shows *cfbglu-cenA-cbhA-cbhB* to have the highest protein content at the end of the 9 days (Fig. 5.1A), indicating that the strain effectively deconstructs and subsequently utilizes cellulose paper as a source of carbon. Strains expressing CenB, CenC or CenD show comparatively less deconstruction of the substrate (Fig. 5.5iii-v), which again is consistent with results from the Coomassie assay showing these strains to have a lower total protein content compared to that of *cfbglu-cenA-cbhA-cbhB* (Fig. 5.1A). The strain *cfbglu-cenA-cex* shows comparatively less deconstruction of the substrate, and significantly less than *cfbglu-cenA-cbhA-cbhB*, which again is consistent with Coomassie results showing less protein generated by *cfbglu-cenA-cex* (Fig. 5.1A). Although both strains exhibit a full cellulase complement, differences are seen in the

utilization and deconstruction of cellulose paper. This may be due to the fact that overexpression of Cex in *E. coli* has been known to cause a toxic effect upon the host, reducing cell viability (Fu *et al.* 2005).

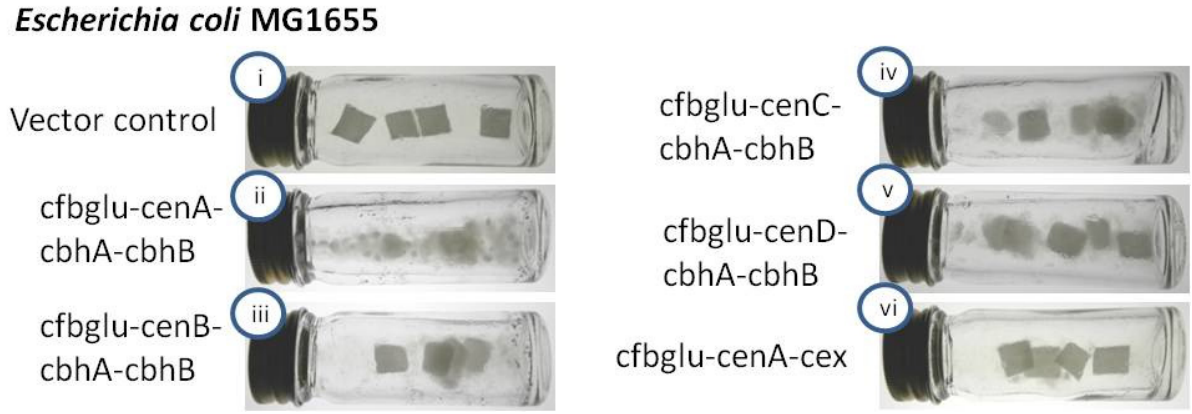

Figure 5.5: Extent of cellulose paper deconstruction by *E. coli* MG1655 cellulase strains after 9 days culturing in M9 media. Cellulase strains *cfbglu-[endoglucanase]-cbhA-cbhB* were used to inoculate 5 ml M9 media supplemented with 0.1% w/v yeast extract, 0.34 g/l thiamine and 125 mg cellulose paper. A culture of the strain *cfbglu-cenA-cex* was also included. Cultures were incubated at 37^{o}C on a rotary shaker at 200 rpm over 9 days. A negative control strain is that transformed with the pSB1C3 vector.

Figure 5.6 illustrating deconstruction of cellulose paper by *C. freundii* NCIMB11490 cellulase strains shows similar results to those presented for *E. coli* MG1655 cellulase strains. Specifically, the strain *cenA-cex-cbhA-cbhB* shows the greatest extent of substrate deconstruction (Fig. 5.6ii). This is followed by that of a co-culture of the three strains expressing CenA, CenB and CenD endoglucanases (Fig. 5.6v). Again, results presented here are consistent with those of the Coomassie protein assay where the highest protein content observed was that of *cfbglu-cenA-cbhA-cbhB* followed by the co-culture of all three strains (Fig 5.2). It is noteworthy that although the co-culture of the three strains showed almost identical protein content to that of CenB and CenD expressing strains at the end of 9 days (Fig. 5.2), the level of substrate deconstruction observed here is greatest for the co-culture (Fig 5.6v). This observation suggests that effective substrate deconstruction and efficient substrate utilization are not synonymous in every instance.

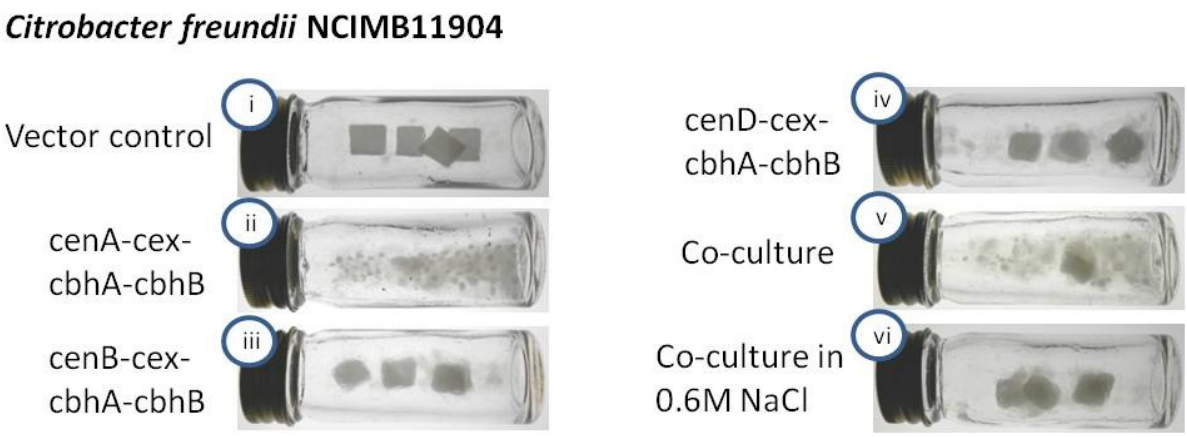

Figure 5.6: Extent of cellulose paper deconstruction by *C. freundii* NCIMB cellulase strains after 9 days culturing in M9 media. Cellulase strains *[endoglucanase]-cex-cbhA-cbhB* were used to inoculate 5 ml M9 media supplemented with 0.1% w/v yeast extract, 0.34 g/l thiamine and 125 mg cellulose paper. A co-culture of all three strains is also included. Cultures were incubated at 37^{o}C on a rotary shaker at 200 rpm over 9 days. A negative control strain is that transformed with the pSB1K3 vector.

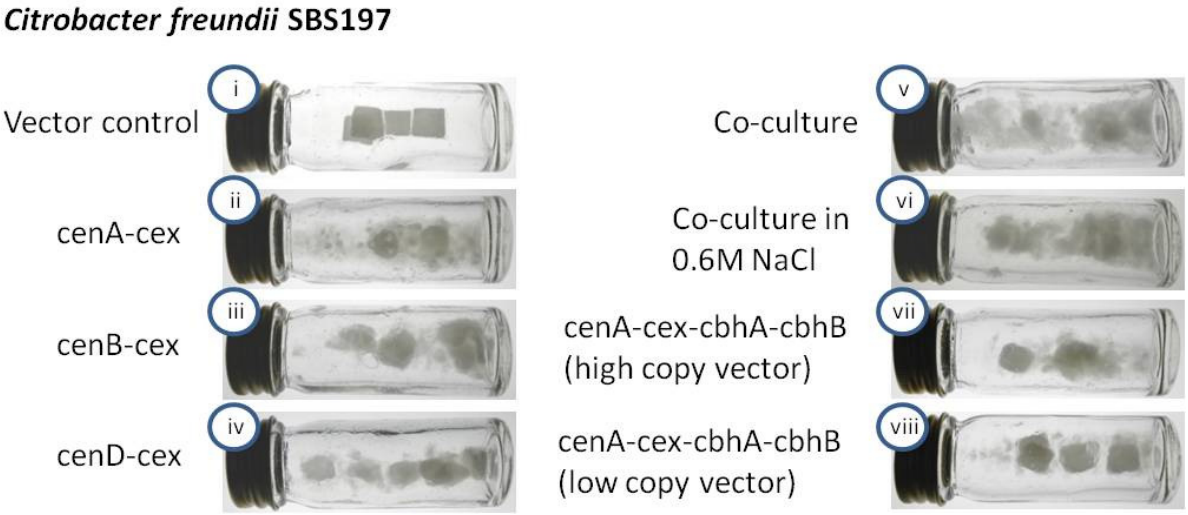

Figure 5.7: Extent of cellulose paper deconstruction by *C. freundii* SBS197 cellulase strains after 9 days culturing in M9 media. Cellulase strains *[endoglucanase]-cex* were used to inoculate 5 ml M9 medium supplemented with 0.1% w/v yeast extract, 0.34 g/l thiamine and 125 mg cellulose paper. A co-culture of all three strains is included, as well as strains expressing *cenA-cex-cbhA-cbhB* in both high and low copy vectors. Cultures were incubated at 37^{o}C on a rotary shaker at 200 rpm over 9 days. A negative control strain is that transformed with the pSB1C3 vector.

The destructive potential of *C. freundii* SBS197 cellulase strains on cellulose paper is presented in Figure 5.7. Of those strains assayed, *cenA-cex* showed the greatest level of substrate deconstruction as observed by the amount of small fibrous particles (Fig. 5.7ii). A co-culture of all three *[endoglucanase]-cex* strains also showed a high extent of cellulose paper deconstruction (Fig. 5.7v). Once again, these results are consistent with those observed within the Coomassie protein assay, as *cenA-cex* showed the highest maximal protein content, followed by a *[endoglucanase]-cex* co-culture (Fig. 5.3). Deconstruction of cellulose paper by *cenA-cex-cbhA-cbhB* strains does not appear to be comparable to that of *cenA-cex* (Fig. 5.7vii, viii); even though those strains were shown to reach a similar maximum cell growth as measured by calculating CFU per ml of those cultures (Fig. 5.4). These results seem to indicate again that deconstruction of cellulose paper is not synonymous with utilization of the substrate.

5.3.2 Time course of cellulose paper deconstruction by *C. freundii* SBS197 *cen(A,B,D)-cex* co-cultures

Figure 5.8 below shows a time course of the observed deconstruction of cellulose paper by a co-culture of *C. freundii* SBS197 *cen(A,B,D)-cex* cellulase strains. Respective positive and negative control cultures shown include wildtype *C. fimi* and the *C. freundii* SBS197 pSB1C3 vector control strain. Observed deconstruction of the substrate was recorded over a 9 day incubation period. Cultures of *C. freundii* SBS197 were conducted at 37^{o}C and that of *C. fimi* at 30^{o}C.

Results identify the stark differences between recombinant *C. freundii* and wildtype *C. fimi* in the ability to degrade cellulose paper. Whilst *C. fimi* shows considerable deformation of the cellulose paper squares at day 2, a similar extent of deformation is only noted for *C. freundii* at day 8. *C. fimi* shows almost complete solubilisation of the paper at day 7, after which no further deconstruction is observed. As noted from previous assays such as the Coomassie protein assay, there is an observably longer lag phase for the generation of protein within recombinant *C. freundii* cultures grown on cellulosic substrates. This is echoed in Figure 5.8 as a significant lag in substrate deformation is noted until days 5 or 6, after which the paper squares lose structure and the media becomes cloudy with particulate matter. This extended lag phase may potentially be due to the inefficient expression of recombinant *C. fimi* genes which exhibit a high GC content. As such, effective cellulolytic activity for the deconstruction and utilization of the substrate is hindered by low expression levels of the cellulases.

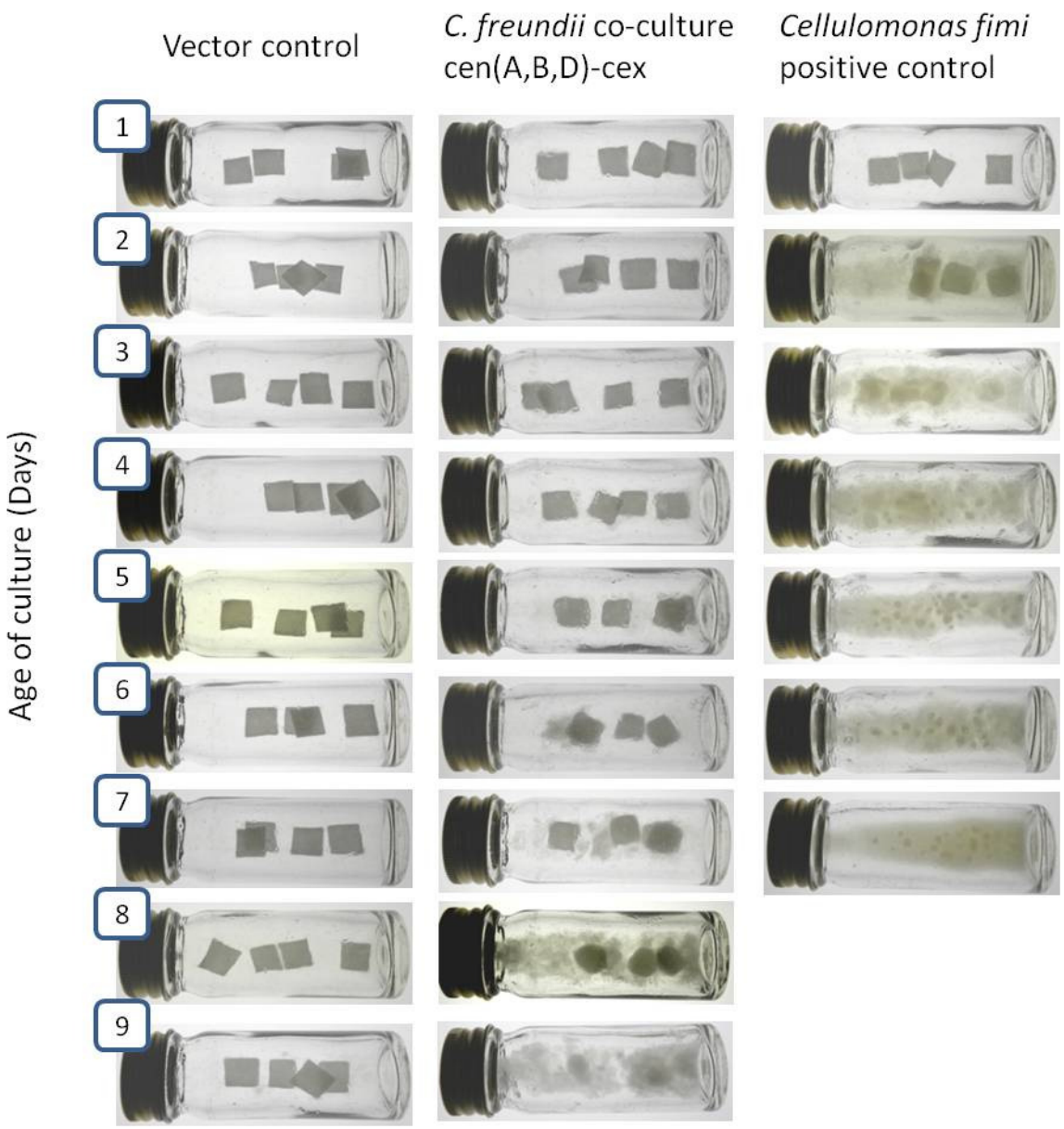

Figure 5.8: Time course of cellulose paper deconstruction by a co-culture of *Citrobacter freundii* cellulase strains compared to wildtype *Cellulomonas fimi*. Deconstructioon of cellulose paper by a co-culture of *C. freundii S*BS197 strains *cenA-cex, cenB-cex* and *cenD-cex* in M9 media containing 0.1% w/v yeast extract, 0.34 g/l thiamine and 125 mg cellulose paper was observed over the course of 9 days. Results are compared to a *C. freundii S*BS197 pSB1C3 vector control culture and a wildtype *C. fimi* culture. *C. freundii* cultures were conducted at 37^{o}C and *C. fimi* at 30^{o}C.

5.4 Utilization and deconstruction of cellulose paper in saline media

Characterization assays of *C. freundii* strains conducted in Chapter 2 identified the ability of the expression host to grow in M9 minimal medium containing 1% w/v glucose and 0.6 M sodium chloride. To further study these observations, *C. freundii* cellulase strains were

assayed for the ability to deconstruct and utilize cellulose paper in saline media. The results of this are presented within this section.

5.4.1 Coomassie protein assay of cultures grown on cellulose paper in saline media

C. freundii cellulase strains were used to inoculate 5 ml M9 minimal medium supplemented with 0.1% w/v yeast extract, 0.6 M NaCl and 125 mg cellulose paper. Cultures were conducted in 1 oz glass vials and incubated at 37^{o}C on a rotary shaker at 200 rpm over a 9 day period. Assays were conducted in duplicate. Results for protein content measured using the Coomassie protein assay reagent are shown in Figure 5.9.

The *C. freundii* SBS197 co-culture of *cen(A,B,D)-cex* strains showed a protein content 2.15 times greater than that of the respective vector control at day 9 (Fig. 5.9A). As for *C. freundii* NCIMB11490, a co-culture of *cen(A,B,D)-cex-cbhA-cbhB* strains showed a 2.05 fold greater absorbance measured at day 9 compared to the vector control.

It should be noted that results measured for each strain are not compared directly, as indicated within Chapter 4 the Coomassie assay will generate variable data across different bacterial species. In comparing each respective co-culture to the appropriate vector control strain, it can be seen that the relative utilization of cellulose paper in saline media by each expression host is comparable. Moreover, the expression of the two cellobiohydrolases CbhA and CbhB do not significantly contribute towards substrate utilization compared to that of Cex.

An interesting observation made in comparing results for growth in saline media against growth in media lacking 0.6 M NaCl, is that almost an identical measurement of protein content is observed in both conditions, for both strains (Fig. 5.9B). Specifically, the *C. freundii* NCIMB11490 co-culture of cellulase strains in M9 media showed an absorbance for Coomassie dye of 0.321, and that of the same co-culture in M9 saline media was 0.350. The *C. freundii* SBS197 co-culture of cellulase strains in M9 media showed an absorbance for Coomassie dye of 0.495, and that of the same co-culture in M9 saline media was 0.521. Results suggest that utilization of the cellulose paper substrate is unaffected in the presence of 0.6 M NaCl.

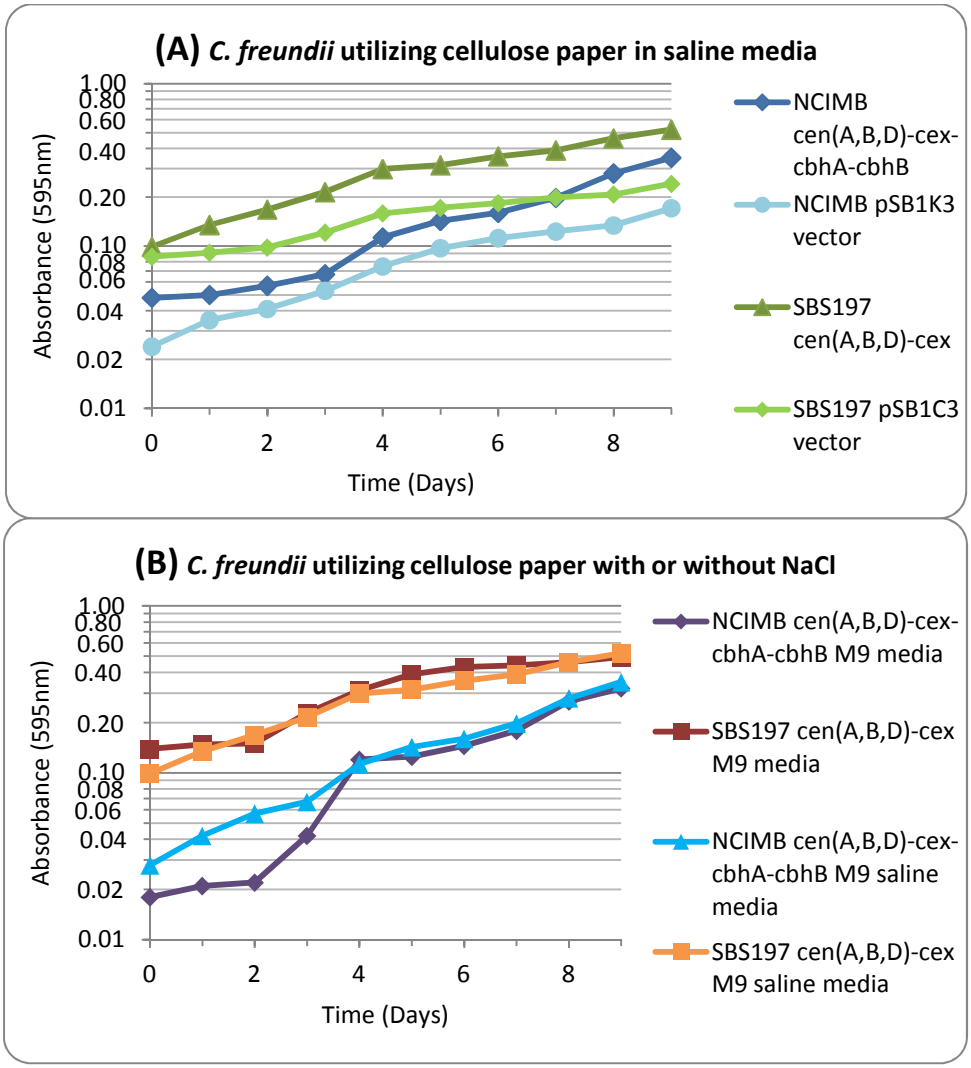

Figure 5.9: Growth of *C. freundii* cellulase strains in co-cultures on cellulose paper in saline media measured as a function of Coomassie dye absorbance at 595 nm. A *C. freundii* NCIMB co-culture of cellulase strains *[endoglucanase]-cex-cbhA-cbhB* and a co-culture of *C. freundii* SBS197 cellulase strains *[endoglucanase]-cex* are assayed for the ability to utilize 125 mg cellulose paper in 5 ml M9 media containing 0.6 M NaCl, 0.34 g/l thiamine and 0.1% yeast extact **(A).** Co-cultures of *C. freundii* cellulase strains grown in saline conditions are compared to identical co-cultures grown in non-saline media **(B).** Total protein content is measured using the Coomassie protein reagent with absorbance recorded at 595 nm**.**

As a comparison to assays conducted for recombinant *C. freundii* strains, cultures for the assay of *C. fimi* growth in M9 media supplemented with 0.1% w/v yeast extract, 125 mg cellulose paper and 0.6 M NaCl were conducted in parallel. Figure 5.10 shows protein content measured by absorbance of the Coomassie protein reagent for *C. fimi* cultures in the presence or absence of 0.6 M NaCl. It can be noted that *C. fimi* cultured in saline media generates a 2.6 fold increase in protein content from day 0 to 9, whilst *C. fimi* cultured in the absence of 0.6 M NaCl shows a 24 fold increase in protein generated.

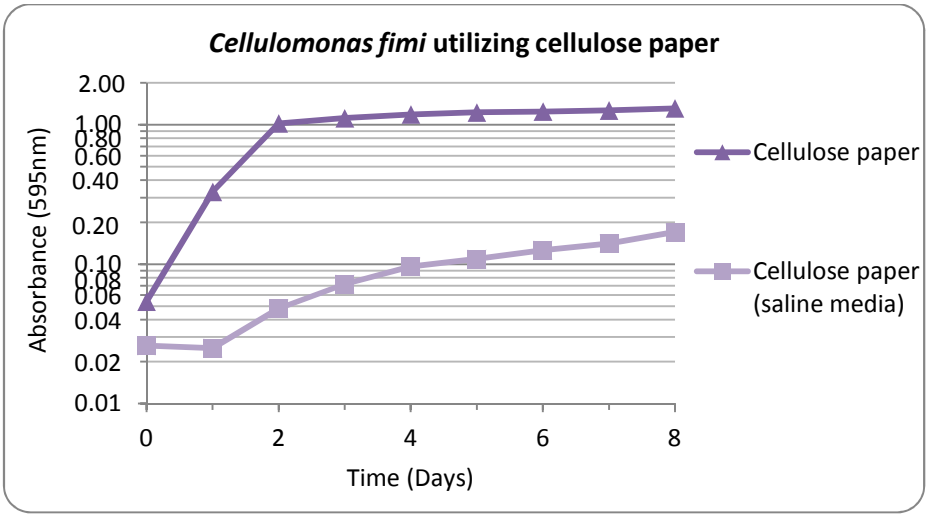

Figure 5.10: Growth of *Cellulomonas fimi* on cellulose paper in saline media measured as a function of Coomassie dye absorbance at 595 nm. Wildtype *Cellulomonas fimi* is assayed for the ability to utilize cellulose paper in M9 media containing 0.6 M NaCl, 0.34 g/l thiamine and 0.1% yeast extract. A second culture set up in M9 media without 0.6 M NaCl is used as a positive control. **Note:** Ab_{595} results above 1 are considered unreliable as samples were not diluted before readings were recorded.

5.4.2 Deconstruction of cellulose paper by *C. freundii* SBS197 co-cultures in saline media

Presented in Figure 5.11 is the time course of a *C. freundii* SBS197 co-culture of strains *cen(A,B,D)-cex* in M9 saline media containing 125 mg cellulose paper. An identical co-culture in M9 media lacking 0.6 M NaCl is presented for comparison as a control. The figure also presents wildtype *Cellulomonas fimi* cultures in M9 media with or without 0.6 M NaCl.

The rate of protein generation measured using the Coomassie protein assay reagent is almost identical for co-cultures grown in media with or without 0.6 M NaCl. These findings are echoed within the time course shown above in Figure 5.11, as both co-cultures show similar degrees of substrate deconstruction over the course of 9 days. By comparison, wildtype *Cellulomonas fimi* shows considerable degradation of the cellulose paper squares to the point of almost complete solubilisation within 7 days in non-saline media. However in saline conditions there is no observable deformation of the substrate and only a slight clouding of the media by fibrous particulate material is observed at days 6 and 7. These results highlight the ability of *C. freundii* to tolerate saline conditions and utilize cellulose paper unhindered in media with a high salt content.

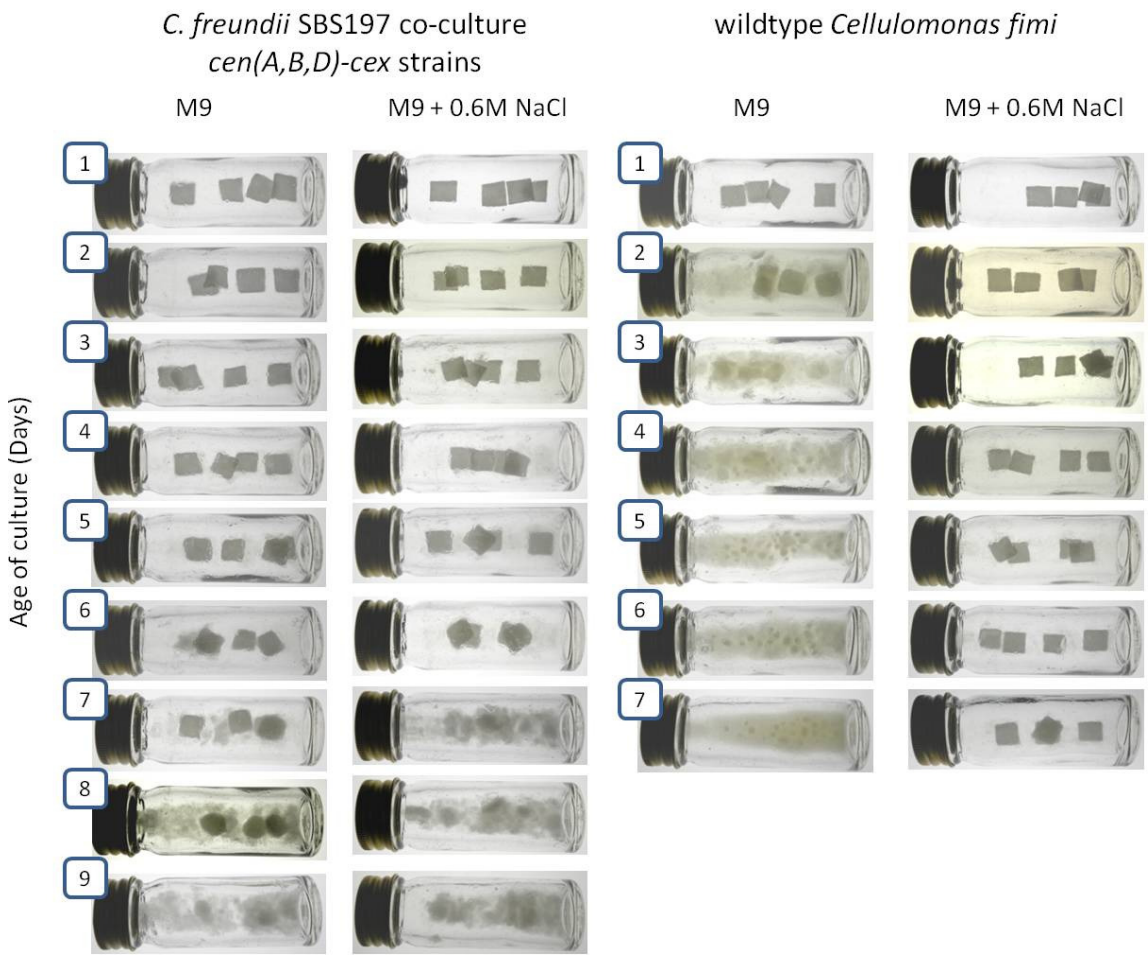

Figure 5.11: Time course of cellulose paper deconstruction by a co-culture of *C. freundii* SBS197 *[endoglucanase]-cex* strains in saline media, compared to wildtype *Cellulomonas fimi*.
Deconstruction of cellulose paper by a co-culture of three *C. freundii* SBS197 strains *cen(A,B,D)-cex* in M9 media containing 0.1% w/v yeast extract, 0.34 g/l thiamine and 125 mg cellulose paper was observed over the course of 9 days. Results are compared to an identical control co-culture of *C. freundii* SBS197 *cen(A,B,D)-cex* strains in M9 lacking 0.6 M NaCl. Wildtype *C. fimi* cultures in M9 media in the absence or presence of 0.6 M NaCl are also shown. *C. freundii* cultures were conducted at 37°C and *C. fimi* at 30°C.

5.5 Discussion

Within this chapter, the utilization and deconstruction of cellulose paper by our recombinant strains compared to wildtype *C. fimi* was documented. Of the individual *E. coli* MG1655 cellulase strains, cultures of *cfbglu-cenA-cbhA-cbhB* showed the greatest total protein content as measured by the Coomassie protein assay. Of the co-cultures assayed, those expressing CenA and CenC grew best on the substrate. Interestingly, individual strains expressing CenB and CenD did not show Coomassie dye absorbance more than 1.46 times that of the vector control, whereas when combined in a co-culture showed a maximal Coomassie dye absorbance 2.95 fold greater than the vector control. These results strongly suggest a

synergistic action between CenB and CenD endoglucanases on cellulose paper, allowing for greater utilization of the substrate for growth. However co-cultures of all four endoglucanase expressing strains did not show an appreciable synergistic activity on cellulase paper, perhaps due to competitive binding between the enzymes limiting enzymatic efficiency, or perhaps lower overall production of those productive enzymes. *E. coli* MG1655strains *cfbglu-cenA-cex* and *cfbglu-cenA-cbhA-cbhB* both exhibit a full cellulase complement, though notably *cfbglu-cenA-cbhA-cbhB* is shown to generate higher protein content. This may be due to the overexpression of Cex in *cfbglu-cenA-cex* reducing cell viability as previous studies identify a toxicity effect to be associated with Cex expression in *E. coli* (Fu *et al.* 2005; Fu *et al.* 2006).

Results for *C. freundii* NCIMB11490 cellulase strains showed that the *cenA-cex-cbhA-cbhB* strain generated the highest protein content as measured by the Coomassie assay. A co-culture of all three strains did not show any appreciable synergistic activities. Comparing *C. freundii* NCIMB11490 and SBS197 strains transformed with the construct *cenA-cex-cbhA-cbhB* it can be noted that SBS197 showed only a 1.51 fold greater absorbance than the vector control, whereas NCIMB11490 expressing the same construct showed a 3.01 fold increase in absorbance compared to the respective vector control. These results suggest NCIMB11490 is more suited than SBS197 for the expression of the larger cellulase constructs. In the case of SBS197, the cellulase strain generating the highest protein content measured by the Coomassie assay was that of *cenA-cex* in cultures with cellulose paper. Colony counts for *C. freundii* SBS197 cellulase strains on cellulose paper confirmed these results in that *cenA-cex* showed the highest CFU per ml compared to the other cellulase strains. Colony counts also identified the population decline observed amongst recombinant SBS197 strains grown in M9 media. As discussed within Chapter 4, this decline may be due to an accumulation of acid produced by the host and a lack of sufficient buffering resulting in a drop in pH.

The fact that all three expression hosts when expressing CenA showed the greatest generation of protein content compared to the other endoglucanase strains suggests that CenA expression contributes the most towards utilization of cellulose paper. This finding may be supported with the observation that CenA plays a role in the non-hydrolytic disruption of cellulose paper as described in Chapter 3, and so contributes to overall deconstruction and utilization of the substrate.

With regard to deconstruction of cellulose paper, the substrate offers itself as a good candidate for observing the physical deformation process within cultures of our recombinant hosts. Images taken of cellulase strains cultured in M9 medium containing cellulose paper showed the extent of substrate deconstruction over the course of 9 days. For the most part, it was observed across each expression host that cultures showing the greatest extent of substrate deconstruction were also those showing the highest generation of protein content, namely those expressing CenA. However, there are examples of *C. freundii* SBS197 cellulase strains where deconstruction of the substrate and subsequent utilization are not synonymous. This is exemplified by the strains *cenA-cex* and *cenA-cex-cbhA-cbhB*; both strains exhibited comparable results for the Coomassie assay and colony counts, indicating similar cell densities and utilization of the substrate is achieved. Yet, deconstruction of cellulose paper was markedly greater in the strain expressing Cex than that expressing the two cellobiohydrolases.

At first these observations may come across as a paradox, as deconstruction of the substrate to liberate fermentable sugars is required for growth and protein synthesis, therefore deconstruction and utilization should be proportional. However, the differences in cellulase expression profiles of each strain may lead to differences in the manner in which the substrate is deconstructed. This has been previously reviewed in the literature as Cex is known to bind irreversibly to crystalline cellulose and processively move along the substrate in a unilateral fashion (Jervis, Haynes, and Kilburn 1997; Mclean *et al.* 2000), whereas both CbhA and CbhB reversibly bind and process from opposing ends of the glucan chain (Gilkes *et al.* 1997). Differences such as these may contribute towards variations seen in the extent of substrate deconstruction and utilization. Alternatively*, cenA-cex* strains could exhibit higher CenA activity compared to *cenA-cex-cbhA-cbhB* strains, leading to greater substrate deconstruction by the non-hydrolytic disruption of CenA.

A time course showing the extent of cellulose paper deconstruction by a co-culture of *C. freundii* SBS197 cellulase strains highlights the inefficiencies of our recombinant host compared to that of wildtype *C. fimi*. Extensive deformation of the substrate by *C. fimi* is observed at day 2 of culturing, whereas a similar result for that of recombinant *C. freundii* is only met at day 8. Moreover, this illustrated lag in cellulolytic activity is suggested to be partly due to low level expression of a very limited set of recombinant cellulases by *C. freundii*. Moreover, the fact that genes cloned from *C. fimi* exhibit a high GC content of 74%

also impedes sufficient expression of those cellulases by the host chassis for effective cellulolytic activity.

In accordance with characterization assays conducted in Chapter 3 for growth of *C. freundii* in saline media, results presented here showed that the utilization and deconstruction of cellulose paper by recombinant *C. freundii* is unhindered in high salt conditions. That of *C. fimi* however showed complete inhibition of cellulolytic activity in saline media. As such, we introduce here the use of *C. freundii* as a suitable expression host for possible future bioprocessing projects conducted in sea water, which may be a more environmentally sustainable strategy as the use of fresh water is limited. Moreover, *C. freundii* offers itself as a feasible expression host for the bioprocessing of alginate derived from brown macroalgae or seaweed, which has previously been shown in *E. coli* (Wargacki *et al.* 2012).

Chapter 6

Conclusions and future prospects

6.1 Premise of the present research

In an effort to reduce carbon dioxide emissions, diversify energy resources and improve future energy security, many industrialised economies have shown a resurgence of interest in the development of renewable fuels and chemicals. The development and commercialisation of these biorenewable products is subject to the engineering of microbial hosts capable of utilizing a sustainable and abundant feedstock such as lignocellulosic plant biomass. However, due to its complex structure and recalcitrant nature, utilization of plant biomass is far from being economically viable and considerable development of microbial strains capable of utilizing the feedstock is necessary. In consideration of this, research presented here is focused on overcoming the biological barriers to lignocellulose hydrolysis through the development and characterization of modular genetic devices encoding defined activities against the major polysaccharides of lignocellulose.

6.2 Summary of contributions

Utilizing the well characterized celluloytic system of *Cellulomonas fimi*, we constructed a library of fourteen standardized gene parts, each encoding a defined activity against three of the major polysaccharides present in lignocellulose, namely; cellulose, mannan or xylan. Parts were designed and assembled according to the original BioBrick assembly method described by Knight (2003). The power of employing this approach and methodology is that multiple parts can be combined in any number or order for the design of tailor made genetic constructs. Moreover, the development of such a library conforming to the BioBrick standard allows for parts to be submitted to a centralized Parts Registry (http://partsregistry.org/), facilitating the distribution to and contribution by peers within the field of study. The co-ordinated development of a centralized standard library of active parts is of great value in the study of lignocellulose hydrolysis, as multiple enzymatic activities are required for the complete hydrolysis of the polysaccharides which make up lignocellulose.

Based on previous characterization studies reported within the literature, as well as characterization assays performed within this thesis, results informed the rational design and

assembly of composite devices made up of multiple lignocellulase encoding BioBrick parts. Nineteen devices in total were assembled with defined activities against mannan, xylan or cellulose. These constructs were designed with the intention of transformation within the two expression hosts *Escherichia coli* and *Citrobacter freundii*, for the purpose of assaying for growth of the various recombinant strains on defined lignocellulosic substrates, namely; mannan from *Saccharomyces cerivisiae*, beechwood xylan, carboxymethyl cellulose, avicel and cellulose paper. Since the constructs are assembled as BioBricks, these may be repurposed and studied further for activities against other cellulosic substrates not included within the present work.

Of the expression hosts transformed with the construct *man2A-man26A* which encodes hydrolysis of mannan, only *E. coli* showed utilization of the substrate in growth assays, whilst *C. freundii* did not. As both hosts are able to assimilate and utilize mannose (Holt and Bergey 2000), the reason for a difference in utilization of mannan may be attributed to differences in expression of the recombinant mannanases. As for *E. coli* and *C. freundii* expressing constructs for xylanase hydrolysis, both hosts were shown to utilize 1% w/v beechwood xylan as a sole carbon source, except in the case of the *C. freundii* strain *xynF-xynD*. This was surprising as the *E. coli* strain *xynF-xynD* was able to utilize xylan; such a difference again may be due to differences in the expression capacity of each host. Further study is required to resolve the discrepancies observed between the two expression hosts.

Assays for the growth of *E. coli* and *C. freundii* cellulase strains on amorphous carboxymethyl cellulose showed that those strains expressing the endoglucanase CenA were best suited to utilization of the substrate. This is consistent with the literature describing CenA as preferentially acting against amorphous forms of cellulose (Tomme *et al.* 1996). In addition it was shown that expression of the cellobiohydrolases CbhA and CbhB contributed significantly towards utilization of CMC, which is consistent with published studies (Warren 1996). Co-cultures of strains expressing the endoglucanase CenB and CenD indicated a synergistic cooperation between the cellulases on CMC as a marked increase in growth was observed as compared to individual cultures of the strains. This relationship was also shown to be present on cellulose paper as a substrate. These findings are as yet unreported within the literature and contribute towards the study of synergism within the endoglucanases of *C. fimi*, which is currently lacking within the published literature. *C. freundii* strains assayed for growth on crystalline avicel cellulose by measurement of colony forming units showed that strains expressing the endoglucanase CenB achieved the highest colony forming units per ml

as compared to CenA or CenD expressing strains. These results complement those presented within the literature describing CenB as an endoglucanase with preferential activity against crystalline cellulose (Tomme *et al.* 1996).

To further assess the enzymatic activities of those cellulase encoding constructs, transformed expression hosts were assayed for growth on cellulose paper as a sole source of carbon. Of the *E. coli* cellulase strains assayed, a co-culture of CenA and CenC expressing strains showed the highest growth and utilization of the substrate. Similarly, *C. freundii* strains expressing CenA exhibited the highest growth on cellulose paper. Enhanced utilization of cellulose paper by CenA expressing strains may be a result of the endoglucanase's role in the non-hydrolytic disruption and deformation of the substrate. Evidence within the literature shows CenA to disrupt the structure of cotton and Ramie fibres as observed by scanning electron microscopy (Din *et al.* 1994). Presented within this thesis, it was shown that the same disruptive action initiated by CenA also occurs on cellulose paper, previously undocumented within the literature. These results support the suggestion that non-hydrolytic disruption increases substrate surface area and accessibility of the cellulases, enhancing utilization of the substrate (Din *et al.* 1994). Moreover, results within this thesis documenting observable deconstruction of cellulose paper over the culturing period confirmed an enhanced deformation of the substrate by CenA, as *E. coli* and *C. freundii* strains expressing CenA showed the highest degree of deconstruction compared to other endoglucanase expressing strains.

A novelty presented within the characterization assays of *C. freundii*, which to our knowledge has not been reported within the literature, is the ability for *C. freundii* to grow in media with an equivalent salinity to that of sea water. Deconstruction and utilization of cellulose paper by *C. freundii* cellulase strains was shown to be unhindered in the presence of 0.6 M sodium chloride, whereas by comparison that of *C. fimi* was entirely inhibited in saline media. These results highlight the potential in developing *C. freundii* as a bioprocessing host capable of culturing in sea water, as a means of reducing the demand on fresh water resources.

6.3 Development of *Citrobacter freundii* as an expression host

We have shown that the assembly of composite devices with defined activities against a cellulosic substrate is achievable from individually characterized parts, though admittedly, more sophisticated and quantifiable characterization assays can facilitate the design process

and improve predictability of functioning devices. These characterization assays would need to be performed with specific considerations for the desired expression host. Consequently, the development of *C. freundii* as an expression host would require the comprehensive assay of our library of lignocellulase encoding parts within the host chassis.

Other considerations for *C. freundii* as a host chassis for recombinant cellulase expression include determining the optimal starter carbon source for cellulose growth assays. Yeast extract may offer itself as a good source of carbon for industrial processes as it can be derived from by-products from the brewing industry. However for the purposes of scientific study it is not an optimal choice as within this thesis a better defined starter carbon source was needed to better quantify the required carbon for supporting initial growth of our various recombinant strains. *C. freundii* is able to utilize glycerol as a sole carbon source, therefore the use of this starter carbon source may aid in the characterization of our strains during growth assays. Moreover glycerol will not affect regulation of the native *C. freundii* β-glucosidase. It should be noted that the use of yeast extract provides additional vitamins and amino acids to support protein synthesis, and so future characterization assays of the host chassis utilizing either glycerol or yeast extract should be compared.

Noted within this work was the population decline of *C. freundii* SBS197 cellulase strains. Amongst members of the lab using this expression host, it was suggested that excess acid formed during metabolism of glucose resulted in a drop in pH, consequently reducing cell viability. The pH of future growth assays should be measured to investigate this, and determine whether additional buffering agents such as MOPS are needed for optimal growth.

To improve the stability of our assembled genetic constructs, which in some instances exceed 12 kilo-bases and also improve transformation efficiencies of the expression host, recombination of our constructs onto the genome of the host could be attempted. Methods for the scarless addition or deletion of recombinant DNA could afford more sustainable expression of cellulases and consequently improved host cell viability (Sun *et al.* 2008; Wang *et al.* 2006). Protocols for achieving this have been reported with the use of linear PCR products exhibiting complementary sequences to recombination sites on the host genome (Datsenko and Wanner 2000). Such a method is ideal for our standardized constructs which can be easily modified to exhibit a range of flanking complementary sequences using standard BioBrick assembly methods. However to perform this method in *C. freundii*, the

sequencing of its genome would be necessary, in order to prospect for suitable recombination sites.

An interesting characteristic of *C. freundii*, previously unmentioned within this work is the ability of the host to process propanediol (Pang, Warren, and Pickersgill 2011). Within a subcellular microcompartment, metabolic processes for the utilization of propanediol are kept in isolation from other host cell activities. With regards to developing the host for biomass bioprocessing projects, the presence of such microcompartments can facilitate the metabolism of secondary products such as long chain alcohols or hydrocarbons which are toxic to the host. Isolating the synthesis of these products could improve host cell viability and improve the economics of a consolidated bioprocessing organism.

As *C. freundii* exhibits a type-II secretion system and offers itself as a suitable host for the recombinant expression of extracellular cellulase proteins, additional modifications at the sequence level can be made to parts within our library to leverage the secretion system. Signal sequences with enhanced activity in coordinating protein secretion can be added to cellulase encoding constructs. Such examples of highly effective signal sequences within the literature have been shown to exhibit cross-host operability and offer exciting routes to improving productivity of the bioprocessing host (Tan, Ho, and Ding 2002).

6.4 Defined expression of enzyme cocktails for cellulose hydrolysis

As presented within this thesis, the attempt to introduce a full cellulase complement from *C. fimi* into that of the expression host *C. freundii* yielded positive results. Notably, though, the efficiency of the recombinant hosts in utilizing cellulose paled in comparison to that of wildtype *C. fimi* from which all cellulases were cloned, and this was exemplified in results documenting the time course of cellulose paper deconstruction. In addition, protein content measured for the most productive *E. coli* strains was a co-culture of strains *cfbglu-cen(A,C)-cbhA-cbhB* giving an absorbance of 0.419, compared to *C. freundii* SBS197 *cenA-cex* of 0.611 and wildtype *C. fimi* of 1.312, all utilizing cellulose paper. Attempting to redesign natural biological systems in an efficient manner is a common challenge reported within the literature surrounding metabolic engineering projects.

Chan *et al.* (2005) in an attempt to modify the genome of bacteriophage T7 for more predictable manipulation and study, observed that a refactored surrogate was active in generating cell lysis on bacterial lawns; however the efficiency of which was only 22% as

compared to the wildtype. More recently, work by Temme *et al.* (2012) showed success in rebuilding the gene cluster of *Klebsiella oxytoca* encoding fixation of atmospheric nitrogen, from well characterized parts in a bottom up approach. Though successful, the refactored gene cluster was shown to perform at only about 7% efficiency as compared to the wildtype.

These examples along with the present study highlight the need in synthetic biology projects for the proper characterization and measurement of biological components, as well a more comprehensive understanding of transcriptional control within natural systems. As such, the development of recombinant hosts for cellulase expression requires the comprehensive characterization and assay of those encoding parts.

An opportunity for improving the functionality of our assembled constructs is that of precisely controlling expression of the exoglucanase Cex and the endoglucanase CenB, both of which were shown to have a toxic effect upon the host cell. As described within the literature, expression of Cex under a weaker promoter improved host cell viability, a strategy that could be extended to that of CenB expression (Fu *et al.* 2006).

Moreover, controlling the relative expression levels of the multiple cellulases encoded within our devices can improve hydrolysis of cellulose. As discussed within Chapter 1, cellulolytic species tend to exhibit an optimized expression profile of cellobiohydrolases type I and II and a single exoglucanase in a molar ratio of 70:30:1 (Meyer, Rosgaard, and Sørensen 2009). To achieve such a defined expression profile, the need for well documented and characterized transcriptional and translational control elements is required. The development of a library of such parts is currently in underway and is being over seen by members of the International Open Facility Advancing Biotechnology (BIOFAB) (unpublished, 2012). Initial reports have announced as a high as >90% predictability in gene expression, results which are within a suitable window for achieving the defined molar ratio of 70:30:1 for optimized cellulase expression and cellulose hydrolysis.

6.5 Protein engineering for modifying glycanase activities

Methods in protein engineering offer potential for the design of glycanases with activities and functionalities unobserved in nature. For example, the substitution of the CenA linker connecting both catalytic and cellulose binding domains with that of the the human antibody IgA1 linker region, was shown to be unaffected by the native *C. fimi* serine protease and was instead cleaved by an IgA protease from *Neisseria gonorrhoea* (Miller *et al.* 1992). The

controlled disassociation the catalytic and cellulose binding domains within the modular cellulases allows for the control of each domain's respective roles in cellulose hydrolysis of soluble oligosaccharides, or disruption of crystalline cellulose by binding to the substrate.

The modularity of the cellulases can also be exploited for the engineering of recombinant hosts with improved binding capacities to the cellulose surface. Francisco *et al.* (1993) showed that expression of a the Cex cellulose binding domain as a fusion protein to an *E. coli* outer membrane protein allowed for the host to bind cellulose. Such a technique offers significant promise in replicating the efficiencies observed in complexed cellulase systems which tether the celluloytic host to the substrate surface, improving hydrolysis of the substrate and metabolic fitness of the host.

The shuffling of catalytic and substrate binding modules within and between the glycanases offers a route to manipulating their activities. Gilkes *et al.* (1991) demonstrated that catalytic domains from different families can be associated with the same type of CBD. Moreover it was shown that a given organism could possess cellulases from several families but only one type of CBD. Furthermore, it was demonstrated that the linker domain shows some sequence identity between enzymes from the same organism but that there is little, if any, identity between linkers from different organisms. These observations show that a degree of shuffling of modules between both species and protein encoding genes is a common occurrence. As such, the high variability between the cellulases arose due to duplication events occurring whereby sequences encoding distinct modules are shuffled in and amongst other repeating gene sequences, and propagated through horizontal gene transfer. The same process could be practised in the lab for the generation of an extensive library of novel cellulase encoding parts. The high-throughput generation and expression of this library could be achieved through the use of a synthetic integron which leverages the ability of transposases to initiate recombination events and shuffling of gene sequences (Bikard *et al.* 2010). Screening of strains expressing variations of the library can be performed on a defined substrate of interest, with the strains exhibiting enhanced activity being selected for by assaying for growth on that substrate.

6.6 Tailor made cellulolytic communities for a defined feedstock

As mentioned before, problems exist in characterizing individual lignocellulases on defined substrates as experimental conditions can greatly affect observed results. Consequently, building composite devices from these parts does not necessarily yield assemblies with truely

defined activities against a substrate as complexities in synergistic activities or competitive inhibition of individual enzyme components is not entirely predictable. Therefore rather than attempt to build large gene assemblies encoding all the enzymatic functions of a wildtype cellulolytic species for transformation in a single recombinant host, the construction of smaller assemblies for distribution across a number of recombinant hosts may be performed.

The advantage of such a distributed approach to the design of an engineered cellulase system is that the labour intensive steps in constructing large gene assemblies can be significantly reduced. Moreover, the assemblies constructed can be repurposed across multiple assays, each composed of a different co-culture or a different cellulosic substrate. This approach may reduce emergent behaviour and improve robustness of the system. Ultimately, the approach allows for the study of an engineered cellulolytic community and can offer insight into hidden synergies and complementarities between the various strains, as evidenced within this thesis. Four different assemblies were transformed into *E. coli*, each encoding a full cellulase complement but with each exhibiting a different endoglucanase. Co-cultures of the resultant strains yielded insights into synergistic activities between CenB and CenD expressing strains when grown on CMC and cellulose paper.

The concept of distributed synthetic gene networks within the field of synthetic biology has been shown to be applicable to the study of many biological phenomena, such as distributed logic operations carried out across multiple individual cells (Li and You 2011). The applications of this are widespread and notably valuable in the design of predictable engineered biological systems.

With specific regards for the development of bioprocessing projects, this concept of a distributed network was shown by Bayer *et al.* (2009) who effectively developed a stable culture of the wildtype cellulolytic bacterium *Actinotalea fermentans* and an engineered *Saccharomyces cerevisiae* yeast strain to synthesise methyl halides from biomass. *A. fermentans* was shown to utilize a number of unprocessed lignocellulose substrates as a sole source of carbon, including; switchgrass, corn stover, sugar bagasse and poplar. Fermentation of biomass by *A. fermentans* yielded acetate and ethanol which could be respired by *S. cerevisiae* and utilized as a carbon source for synthesis of methyl halides, commodity chemicals which can be polymerized by synthetic chemistry processes to produce long chain hydrocarbons.

6.7 Foundation of a bioeconomy

Published studies presented here, as well as results from this thesis, highlight the advances made in predictably engineering biology for societal benefit. Specifically we have shown the potential for developing a bioprocessing host capable of utilizing cellulose as a sole carbon source. In the case of *C. freundii*, we also showed that the utilization of cellulose was uninhibited in saline media, allowing for the potential use of sea water for culturing and so reducing the use of fresh water resources. Further work in developing the host for utilizing raw biomass such as agriculture waste streams is still needed, though the use of parts based biology in designing strains for the utilization of defined lignocellulose polysaccharides including mannan and xylan shows promise towards this goal.

A scenario in which waste materials are utilized for the generation of high value chemicals and fuels could be the foundation of a biologically driven economy, or bioeconomy (Carlson 2007). Given the rapidly increasing global human population and by consequence the dwindling of traditional energy resources, innovative and more sustainable technologies are needed for development. The studies and concepts presented here can contribute toward that effect, though should be considered with regards to a larger picture; whereby waste streams within society are extensively repurposed and recycled across all aspects of industry, whilst in addition natural resources are carefully managed and invested in for future generations.

Appendix

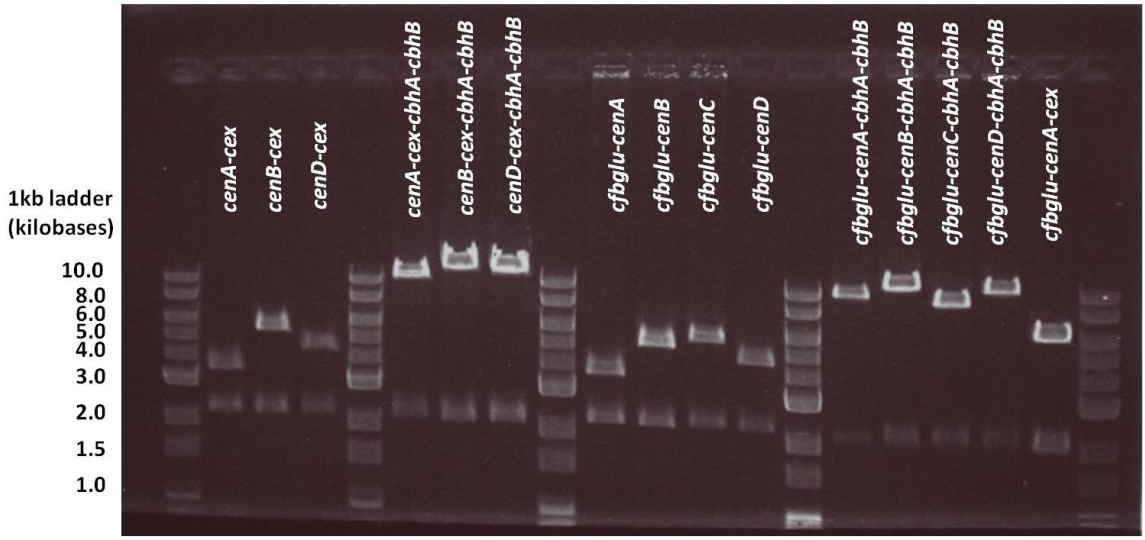

Fig. A: Restriction digest analysis of plasmid DNA from recombinant cellulase expressing strains. Miniprep plasmid DNA isolated from *Citrobacter freundii* NCIMB11490, *Citrobacter freundii* SBS197 and *Escherichia coli* MG1655 recombinant cellulase expressing strains was digested with the endonucleases EcoRI and PstI to confirm the presence of plasmids housing the cellulase encoding constructs. Constructs *[endoglucanase]-cex* and *[endoglucanase]-cex-cbhA-cbhB* are housed in the plasmid pSB1C3. Constructs *cfbglu-[endoglucanase]* are housed in the plasmid pSB1K3. Constructs *cfbglu-[endoglucanase]-cbhA-cbhB* are housed in the plasmid pSB1C3.

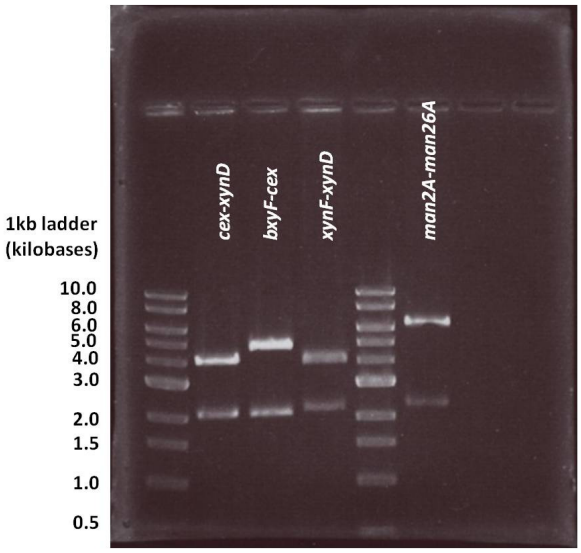

Fig. B: Restriction digest analysis of plasmid DNA from recombinant xylanase and mannanase expressing *Citrobacter freundii* NCIMB11490 strains. Miniprep plasmid DNA isolated from *Citrobacter freundii* NCIMB11490 recombinant xylanase and mannanase expressing strains was digested with the endonucleases EcoRI and PstI to confirm the presence of plasmids housing the cellulase encoding constructs. Xylanase constructs *cex-xynD, bxyF-cex* and *xynF-xynD* were housed in the plasmid pSB1C3. The mannanase construct *man2A-man26A* was housed in the plasmid pSB1K3.

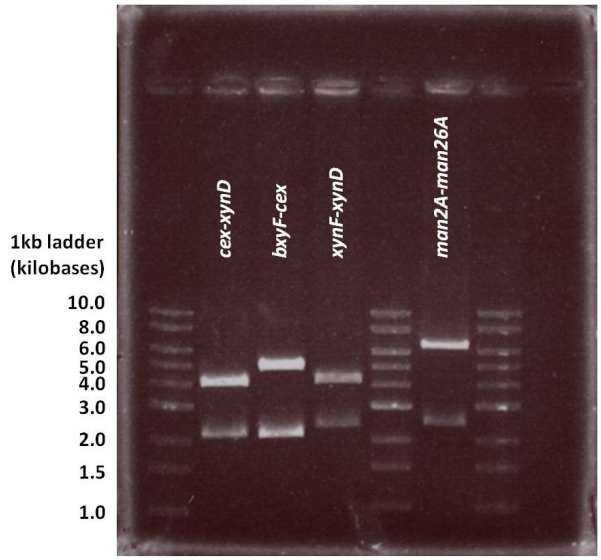

Fig. C: Restriction digest analysis of plasmid DNA from recombinant xylanase and mannanase expressing *Escherichia coli* MG1655 strains. Miniprep plasmid DNA isolated from *Escherichia coli* MG1655 recombinant xylanase and mannanase expressing strains was digested with the endonucleases EcoRI and PstI to confirm the presence of plasmids housing the cellulase encoding constructs. Xylanase constructs *cex-xynD, bxyF-cex* and *xynF-xynD* were housed in the plasmid pSB1C3. The mannanase construct *man2A-man26A* was housed in the plasmid pSB1K3.

References

Arantes, V., and J. N. Saddler. 2010. "Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis." *Biotechnology for Biofuels* 3: 4. doi:10.1186/1754-6834-3-4.

Baker, J O, C I Ehrman, W S Adney, S R Thomas, and M E Himmel. 1998. "Hydrolysis of cellulose using ternary mixtures of purified cellulases." *Applied Biochemistry and Biotechnology* 70-72: 395-403.

Bayer, T. S., D. M. Widmaier, K. Temme, E. A. Mirsky, D. V. Santi, and C. A. Voigt. 2009. "Synthesis of Methyl Halides from Biomass Using Engineered Microbes." *Journal of the American Chemical Society* 131 (18): 6508-6515.

Biely, P., and V. Puchart. 2006. "Recent progress in the assays of xylanolytic enzymes." *Journal of the Science of Food and Agriculture* 86: 1636-1647. doi:10.1002/jsfa.

Bikard, D., S. Julié-Galau, G. Cambray, and D. Mazel. 2010. "The synthetic integron: an in vivo genetic shuffling device." *Nucleic acids research* 38 (15): 153. doi:10.1093/nar/gkq511.

Black, G. W,, G. P. Hazlewood, S. J. Millward-Sadler, J. I. Laurie, and H. J. Gilbert. 1995. "A modular xylanase containing a novel non-catalytic xylan-specific binding domain." *The Biochemical Journal* 307 (1): 191-5.

Boisset, C., C. Pétrequin, H. Chanzy, B. Henrissat, and M. Schülein. 2001. "Optimized mixtures of recombinant *Humicola insolens* cellulases for the biodegradation of crystalline cellulose." *Biotechnology and Bioengineering* 72 (3): 339-345.

Bokinsky, G., P. P. Peralta-Yahya, A. George, B. M. Holmes, E. J Steen, J. Dietrich, T. Soon Lee. 2011. "Synthesis of three advanced biofuels from ionic liquid-pretreated switchgrass using engineered *Escherichia coli*." *Proceedings of the National Academy of Sciences of the United States of America* 108 (50): 19949-19954. doi:10.1073/pnas.1106958108.

Bolam, D. N., H. Xie, P. White, P. J. Simpson, S. M. Hancock, M. P. Williamson, and H. J. Gilbert. 2001. "Evidence for synergy between family 2b carbohydrate binding modules in *Cellulomonas fimi* xylanase 11A." *Biochemistry* 40 (8): 2468-2477.

Boudeti, A., C. Lapierre, and J. Pettenati. 1995. "Biochemistry and molecular biology of lignification." *New Phytologist* 129 (2): 203-236.

Bradford, M. M. 1976. "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." *Analytical Biochemistry* 72 (7): 248-54.

Brun, E., P. E. Johnson, A. L. Creagh, P. Tomme, P. Webster, C. A. Haynes, and L. P. McIntosh. 2000. "Structure and binding specificity of the second N-terminal cellulose-binding domain from *Cellulomonas fimi* endoglucanase C." *Biochemistry* 39 (10): 2445-2458.

Cantarel, B. L., P. M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, and B. Henrissat. 2009. "The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics." *Nucleic Acids Research* 37 (Databse issue) (January): D233-8. doi:10.1093/nar/gkn663.

Carlson, R. 2007. “Laying the foundations for a bio-economy.” *Systems and Synthetic Biology* 1 (3): 109-17. doi:10.1007/s11693-007-9010-z.

Carlson, R. 2009. “The changing economics of DNA synthesis.” *Nature Biotechnology* 27 (12): 1091-1094. doi:10.1038/nbt1209-1091.

Cavaco–Paulo, A., J. Morgado, J. Andreaus, and D. Kilburn. 1999. “Interactions of cotton with CBD peptides.” *Enzyme and Microbial Technology* 25 (8-9): 639-643. doi:10.1016/S0141-0229(99)00101-5.

Chan, L. Y., S. Kosuri, and D. Endy. 2005. “Refactoring bacteriophage T7.” *Molecular Systems Biology* 2005 (18): 1-10. doi:10.1038/msb4100025.

Cheng, A., and T. K. Lu. 2012. “Synthetic Biology: An Emerging Engineering Discipline.” *Annu. Rev. Biomed. Eng. 2012.* 14: 155-178. doi:10.1146/annurev-bioeng-071811.

Chung, C. T., S. L. Niemela, and R. H. Miller. 1989. “One-step preparation of competent E*scherichia coli:* transformation and storage of bacterial cells in the same solution.” *Proceedings of the National Academy of Sciences of the United States of America* 86 (7): 2172-5.

Clarke, J., K.. Davidson, J. Gilbert, C. Fontes, and G. P. Hazlewood. 1996. “A modular xylanase from mesophilic *Cellulomonas fimi* contains the same cellulose-binding and thermostabilizing domains as xylanases from thermophilic bacteria.” *FEMS Microbiology Letters* 139 (1): 27-35. doi:10.1016/0378-1097(96)00101-2.

Cosgrove, D. J. 2000. “Loosening of plant cell walls by expansins.” *Nature* 407 (21): 321-6. doi:10.1038/35030000.

Coughlan, M. 1985. “The properties of fungal and bacterial cellulases with comment on their production and applications.” *Biotechnol Genet Eng Rev* 3: 39-109.

Coughlan, M., and F. Mayer. 1992. The cellulose-decomposing bacteria and their enzyme systems. In *The Prokaryotes: a Handbook on theBbiology of Bacteria*, 460-516. 2nd ed. New York, NY.: Springer-Verlag.

Coutinho, J .B., N. R. Gilkes, R. A. Warren, D. G. Kilburn, and R. C. Miller. 1992. “The binding of *Cellulomonas fimi* endoglucanase C (CenC) to cellulose and Sephadex is mediated by the N-terminal repeats.” *Molecular Microbiology* 6 (9): 1243-1252.

Coutinho, J. B., B. Moser, D. G. Kilburn, R. A. Warren, and R. C. Miller. 1991. “Nucleotide sequence of the endoglucanase C gene (cenC) of *Cellulomonas fimi*, its high-level expression in *Escherichia coli*, and characterization of its products.” *Molecular Microbiology* 5 (5): 1221-1233.

Damude, H. G., V. Ferro, S. G. Withers, and R. A. Warren. 1996. “Substrate specificity of endoglucanase A from *Cellulomonas fimi*: fundamental differences between endoglucanases and exoglucanases from family 6.” *Biochem. J.* 315: 467-472.

Datsenko, K. A., and B. L. Wanner. 2000. “One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products.” *Proceedings of the National Academy of Sciences of the United States of America* 97 (12): 6640-6645. doi:10.1073/pnas.120163297.

Din, N., H. G. Damude, N. R. Gilkes, R. C. Miller, R. A. Warren, and D. G. Kilburn. 1994. "C1-Cx revisited: intramolecular synergism in a cellulase." *Proceedings of the National Academy of Sciences of the United States of America* 91 (24): 11383-7.

Din, N., N. R. Gilkes, B.Tekant, R. C. Miller, R. A. Warren, and D. G. Kilburn. 1991. "Non–Hydrolytic Disruption of Cellulose Fibres by the Binding Domain of a Bacterial Cellulase." *Nat. Biotechnol* 9: 1096-9.

Diouf, J., and J. Sheeran. 2010. *The State of Food Insecurity in the World: Addressing food insecurity in protracted crises*. Food and Agricultural Organization of the United Nations. Rome.

Dudley, B. 2011. *BP Statistical Review of World Energy June 2011*. British Petroleum plc, London.

Endy, D. 2005. "Foundations for Engineering Biology." *Nature* 438 (7067): 449-453.

Esteghlalian, S., N. R. Gilkes, D. G. Kilburn, R. A. Warren, and J. N. Saddler. 2001. "Do cellulose binding domains increase substrate accessibility?" *Applied Biochemistry And Biotechnology* 91: 575-592.

Eveleigh, D. E. 1987. "Cellulase: a perspective." *Philosophical Transactions of the Royal Society of London. Series A, Mathmetical and Physical Sciences* 321: 435-447.

Francisco, J. A., C. Stathopoulos, R. A. Warren, D. G. Kilburn, and G. Georgiou. 1993. "Specific Adhesion and Hydrolysis of Cellulose by intact *Escherichia coli* Expressing Surface Anchored Cellulase or Cellulose Binding Domains." *Nature Biotechnol.* 11: 491.

French, C. E. 2009. "Synthetic biology and biomass conversion: a match made in heaven?" *Journal of the Royal Society, Interface* 6 Suppl 4 (June): S547-58. doi:10.1098/rsif.2008.0527.focus.

Fu, Z. B., K. L. Ng, C. C. Lam, K. C. Leung, W. H. Yip, and W. K. R. Wong. 2006. "A two-stage refinement approach for the enhancement of excretory production of an exoglucanase from *Escherichia coli*." *Protein expression and purification* 48 (2): 205-14. doi:10.1016/j.pep.2006.01.013.

Fu, Z. B., K. L. Ng, T. L. Lam, and W. K. R. Wong. 2005. "Cell death caused by hyper-expression of a secretory exoglucanase in *Escherichia coli*." *Protein Expression and Purification* 42(1): 67-77. doi:10.1016/j.pep.2005.03.029.

Gilbert, H. J. 2010. "The biochemistry and structural biology of plant cell wall deconstruction." *Plant Physiology* 153(2): 444-455. doi:10.1104/pp.110.156646.

Gilbert, H. J,, H. Stålbrand, H. Brumer, and H. Sta. 2008. "How the walls come crumbling down: recent structural biochemistry of plant polysaccharide degradation." *Current Opinion in Plant Biology* 11 (3): 338-348. doi:10.1016/j.pbi.2008.03.004.

Gilbert, J., and P. Hazlewood. 1993. "Bacterial cellulases and xylanases." *Journal of General Microbiology* 139: 187-194.

Gilkes, N., E. Kwan, D. G. Kilburn, and C. Miller. 1997. "Attack of carboxymethylcellulose at opposite ends by two cellobiohydrolases from *Cellulomonas fimi*." *Journal of Biotechnology* 57: 83-90.

Gilkes, N. R., R. Antony, J. Warren, R. C. Miller, D. G. Kilburn, N. R. Gilkes, and R. A. Warren. 1988. "Precise Excision of the Cellulose Binding Domains from two *Cellulomonas fimi* Cellulases by a Homologous Protease and the Effect on Catalysis." *The Journal of biological Chemistry* 263(21) : 10401-10407.

Gilkes, N. R., B. Henrissat, D. G. Kilburn, R C. Miller, and R. A. Warren. 1991. "Domains in microbial beta-1, 4-glycanases, sequence conservation, function, and enzyme families." *Microbiological Reviews* 55 (2): 303-315.

Greenberg, N. M., R. A. Warren, G. Kilburn, and C. Miller. 1987. "Regulation, Initiation, and Termination of the *cenA* and *cex* Transcripts of *Cellulomonas fimi*." *Journal of Bacteriology* 169 (2): 646-653.

Greenberg, N., R. Warren, R. C. Miller, and D. G. Kilburn. 1987. "Regulation and initiation of cenB transcripts of *Cellulomonas fimi*." *Journal of Bacteriology* 169 (10): 4674-4677.

Guo, Z., A. Ong, N. R. Gilkes, G. Kilburn, R. A. Warren, and R. C. Miller. 1988. "Leakage of *Cellulomonas fimi* cellulases from *Escherichia coli*." *FEMS Microbiology Letters* 49 (2): 279-283. doi:10.1016/0378-1097(88)90474-0.

Hansen, C. K. 1992. "Fibronectin type III-like sequences and a new domain type in prokaryotic depolymerases with insoluble substrates." *FEBS Letters* 305 (2): 91-6.

Haruta, S., Z. Cui, Z. Huang, M. Li, M. Ishii, and Y. Igarashi. 2002. "Construction of a stable microbial community with high cellulose-degradation ability." *AppliedMmicrobiology and Biotechnology* 59 (4-5): 529-34. doi:10.1007/s00253-002-1026-4.

Hekmat, O., C. Florizone, Y. W. Kim, L. D. Eltis, R A. Warren, and S. G. Withers. 2007. "Specificity fingerprinting of retaining beta-1,4-glycanases in the *Cellulomonas fimi* secretome using two fluorescent mechanism-based probes." *Chembiochem* 8 (17): 2125-2132. doi:10.1002/cbic.200700481.

Hekmat, O., Shouming He, R A. Warren, and S. G. Withers. 2008. "A Mechanism-Based ICAT Strategy for Comparing Relative Expression and Activity Levels of Glycosidases in Biological Systems." *Journal of Proteome Research* 7: 3282-3292.

Hekmat, O., Y. W. Kim, S. J. Williams, Shouming He, and S. G. Withers. 2005. "Active-site peptide 'fingerprinting' of glycosidases in complex mixtures by mass spectrometry. Discovery of a novel retaining beta-1,4-glycanase in *Cellulomonas fimi*." *The Journal of Biological Chemistry* 280 (42): 35126-35135. doi:10.1074/jbc.M508434200.

Henrissat, B. 1994. "Cellulases and their interaction with cellulose." *Cellulose* 1: 169-196.

Henrissat, B., T. Teeri, and R. A. Warren. 1998. "A scheme for designating enzymes that hydrolyse the polysaccharides in the cell walls of plants." *FEBS Letters* 425: 352-354.

Higashide, W., Y. Li, Y. Yang, and J. C. Liao. 2011. "Metabolic engineering of *Clostridium cellulolyticum* for production of isobutanol from cellulose." *Applied and Environmental Microbiology* 77 (8): 2727-2733. doi:10.1128/AEM.02454-10.

Van der Hoeven, M. 2011. *CO2 emissions from fuel combustion: Highlights*. International Energy Agency, Paris, France.

Holt, J. G., and D. H. Bergey. 2000. *Bergey's Manual of Systematic Bacteriology*. Ed. Wilian R. Hensyl. *Genome*. 9th ed. Vol. 2. Philadelphia: Lippincott Williams and Wilkins. doi:10.1007/0-387-30743-5.

Hua, S., Meinke A., Tomme P., Damude H. G., Kwan E., Kilburn D. G., Miller R. C., Warren R A., and Gilkes N. R. 1996. *Cellulomonas fimi* Cellobiohydrolases. In *ACS Symposium Series, Vol. 618*, 618:174-196. Washington, DC.: American Chemical Society, May 5. doi:doi:10.1021/bk-1995-0618.ch012.

Irwin, D. C., M. Spezio, L. P. Walker, and D. B. Wilson. 1993. "Activity studies of eight purified cellulases: Specificity, synergism, and binding domain effects." *Biotechnology and Bioengineering* 42 (8): 1002-13. doi:10.1002/bit.260420811.

Jervis, E. J., C. A. Haynes, and D. G. Kilburn. 1997. "Surface diffusion of cellulases and their isolated binding domains on cellulose." *The Journal of Biological Chemistry* 272 (38): 24016-24023.

Jervis, E., L. Creagh, E. Ongt, D. G. Kilburn, and C. A. Haynes. 1996. "Binding of the cellulose-binding domain of exoglucanase Cex from *Cellulomonas fimi* to insoluble microcrystalline cellulose is entropically driven." *Proceedings of the National Academy of Sciences* 93 (22): 12229-12234. doi:10.1073/pnas.93.22.12229.

Johnson, P. E., P. Tomme, M. D. Joshi, and L. P. McIntosh. 1996. "Interaction of soluble cellooligosaccharides with the N-terminal cellulose-binding domain of *Cellulomonas fimi* CenC NMR and ultraviolet absorption spectroscopy." *Biochemistry* 35 (44): 13895-13906. doi:10.1021/bi961186a.

Kataeva, I. A., R. D. Seidel, A. Shah, L. T. West, X. L. Li, and L. G. Ljungdahl. 2002. "The Fibronectin Type 3-Like Repeat from the *Clostridium thermocellum* Cellobiohydrolase CbhA Promotes Hydrolysis of Cellulose by Modifying its Surface." *Applied and Environmental Microbiology* 68 (9): 4292-4300. doi:10.1128/AEM.68.9.4292.

Kato, S., S. Haruta, Z. J. Cui, and M. Ishii. 2005. "Stable Coexistence of Five Bacterial Strains as a Cellulose-Degrading Community." *Applied and Environmental Microbiology* 71 (11): 7099-7106. doi:10.1128/AEM.71.11.7099.

Kenyon, W. J., S. W. Esch, and C. S. Buller. 2005. "The curdlan-type exopolysaccharide produced by *Cellulomonas flavigena* KU forms part of an extracellular glycocalyx involved in cellulose degradation." *Antonie van Leeuwenhoek* 87 (2) : 143-8. doi:10.1007/s10482-004-2346-4.

Kerff, F., A. Amoroso, R. Herman, E. Sauvage, S. Petrella, P. Filée, P. Charlier. 2008. "Crystal structure and activity of *Bacillus subtilis* YoaJ (EXLX1), a bacterial expansin that promotes root colonization." *Proceedings of the National Academy of Sciences* 105 (44): 16876-81. doi:10.1073/pnas.0809382105.

Kim, H.K. K, and M.Y. Y Pack. 1989. "Cloning and expression of *Cellulomonas fimi* β-glucosidase genes in *Escherichia coli*." *Enzyme and Microbial Technology* 11 (5) (May): 313-316. http://www.sciencedirect.com/science/article/pii/0141022989900483.

Kleman-Leyer, K. M., N. R. Gilkes, R. C. Miller, T. K. Kirk. 1994. "Changes in the molecular-size distribution of insoluble celluloses by the action of recombinant *Cellulomonas fimi* cellulases." *The Biochemical Journal* 302: 463-469.

Knight, Tom. 2003. *Idempotent Vector Design for Standard Assembly of Biobricks*. http://dspace.mit.edu/handle/1721.1/21168.

Koivula, A., G. Carrard, H Söderlund, and P Béguin. 2000. "Cellulose-binding domains promote hydrolysis of different sites on crystalline cellulose." *Proceedings of the National Academy of Sciences of the United States of America* 97 (19): 10342-7. doi:10.1073/pnas.160216697.

Kumar, R., S. Singh, O. V., and Singh. 2008. "Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives." *Journal of Industrial Microbiology & Biotechnology* 35 (5): 377-391. doi:10.1007/s10295-008-0327-8.

Lamed, R., J. Naimark, E. Morgenstern, and E. A. Bayer. 1987. "Specialized cell surface structures in cellulolytic bacteria." *Journal of Bacteriology* 169 (8): 3792-800.

Lane, I. 2012. Brazil sets up $38 billion ethanol subsidy program to stimulate expansion. *Biofuels Digest*. http://www.biofuelsdigest.com/bdigest/2012/02/27/brazil-sets-up-38-billion-ethanol-subsidy-program-to-stimulate-expansion/.

Lane, J. 2012. Biofuels, materials raise $1.374B in equity in past 12 months. *BioFuels Digest*. http://www.biofuelsdigest.com/bdigest/2012/02/15/biofuels-materials-raise-1-374b-in-equity-in-past-12-months/.

Langsford, M. L., N. R. Gilkes, B. Singh, B. Moser, R. C. Miller, R. A. Warren, and D. G. Kilburn. 1987. "Glycosylation of bacterial cellulases prevents proteolytic cleavage between functional domains." *FEBS Lett* 225: 163-167.

Laurie, J. I., J. H. Clarke, A. Ciruela, C. Faulds, G. Williamson, H. J. Gilbert, J. Rixon, S. J. Millward-Sadler, and G. P. Hazlewood. 1997. "The NodB Domain of a multidomain xylanase from *Cellulomonas fimi* deacetylates acetylxylan." *FEMS Microbiology Letters* 148: 261-264.

Levy, I., Z. Shani, and O. Shoseyov. 2002. "Modification of polysaccharides and plant cell wall by endo-1,4-β-glucanase and cellulose-binding domains." *Biomolecular Engineering* 19: 17-30.

Li, Bochong, and Lingchong You. 2011. "Division of logic labour." *Nature* 469 (4): 171-172. doi:10.1029/2008GC002281.

Linder, M., and T. T. Teeri. 1997. "The roles and function of cellulose-binding domains." *Journal of Biotechnology* 57: 15-28. doi:10.1016/S0168-1656(97)00087-4.

Little, E., P. Bork, and R. F. Doolittle. 1994. "Tracing the Spread of Fibronectin Type III Domains in Bacterial Glycohydrolases." *Journal of Molecular Evolution* 39: 631-643.

Lucas, S., J. Han, A. Lapidus, J. F. Cheng, L. Goodwin, S. Pitluck, L. Peters. 2012. "Complete sequence of *Cellulomonas fimi* ATCC 484." *(Submitted)*. http://www.ncbi.nlm.nih.gov/genome/?term=cellulomonas fimi.

Lynd, L. R., P. J. Weimer, W. H. Van Zyl, and I. S. Pretorius. 2002. "Microbial Cellulose Utilization: Fundamentals and Biotechnology." *Microbiology and molecular biology reviews* 66 (3): 506-577. doi:10.1128/MMBR.66.3.506.

Mansfield, S. D., and R. Meder. 2003. "Cellulose hydrolysis – the role of monocomponent cellulases in crystalline cellulose degradation." *Cellulose* 10: 159-169.

Marris, E. 2006. “Drink the best and drive the rest: Brazil’s sugar-cane ethanol industry is the world's best and able to get better.” *Nature* 444: 670-672.

Martin, V., D. J. Pitera, S.T. Withers, J. D. Newman, and J. D. Keasling. 2003. “Engineering a mevalonate pathway in Escherichia coli for production of terpenoids.” *Nature Biotechnology* 21 (7): 796-802. doi:10.1038/nbt833.

Mayer, C., D. J. Vocadlo, M. Mah, K. Rupitz, D. Stoll, R. A. Warren, and S. G. Withers. 2006. “Characterization of a beta-N-acetylhexosaminidase and a beta-N-acetylglucosaminidase/beta-glucosidase from *Cellulomonas fimi*.” *FEBS Journal* 273 (13): 2929-2941. doi:10.1111/j.1742-4658.2006.05308.x.

Mclean, B. W., M. R. Bray, A. B. Boraston, N. R. Gilkes, C. A. Haynes, and D. G. Kilburn. 2000. “Analysis of binding of the family 2a carbohydrate-binding module from *Cellulomonas fimi* xylanase 10A to cellulose: specificity and identification of functionally important amino acid residues.” *Protein Engineering* 13 (11): 801-809.

Meinke, A., N. R. Gilkes, D. G. Kilburn, R. C. Miller, R. A. Warren. 1993. “Cellulose-binding polypeptides from *Cellulomonas fimi*: endoglucanase D (CenD), a family A beta-1,4-glucanase.” *Journal ofBbacteriology* 175 (7): 1910-1918.

Meinke, A., N. R. Gilkes, E. Kwan, D. G. Kilburn, R. A. Warren, and R. C. Miller. 1994. “Cellobiohydrolase A (CbhA) from the cellulolytic bacterium *Cellulomonas fimi* is a beta-1,4-exocellobiohydrolase analogous to *Trichoderma reesei* CBH II.” *Molecular Microbiology* 12 (3): 413-422.

Meinke, A., R. Gilkes, J. Warren, D. G. Kilburn, and R.C. Miller. 1991. “Multiple Domains in Endoglucanase B (CenB) from *Cellulomonas fimi*: Functions and Relatedness to Domains in Other Polypeptides.” *Journal of Bacteriology* 173 (22): 7126-7135.

Meinke, A., M. Schmuck, N. R. Gilkes, D. G. Kilburn, R. C. Miller, and R. A. Warren. 1992. “The tertiary structure of endo-β-1,4-glucanase B (CenB), a multidomain cellulase from the bacterium *Cellulomonas fimi*.” *Glycobiology* 2 (4): 321-326. doi:10.1093/glycob/2.4.321.

Meyer, A. S., L. Rosgaard, and H. R. Sørensen. 2009. “The minimal enzyme cocktail concept for biomass processing.” *Journal of Cereal Science* 50 (3): 337-344. doi:10.1016/j.jcs.2009.01.010.

Miller, P. B., H. Shen, N. R. Gilkes, D. G. Kilburn, R. C. Miller, G. Plaut, and R. A. Warren. 1992. “Endoglucanase A from *Cellulomonas fimi* in which the hinge sequence of human IgA1 is substituted for the linker connecting its two domains is hydrolyzed by IgA proteases from *Neisseria gonorrhoeae*.” *FEMS microbiology letters* 71 (2): 199-203.

Millward-Sadler, S. J., D. M. Poole, B. Henrissat, G. P. Hazlewood, J. H. Clarke, and H. J. Gilbert. 1994. “Evidence for a general role for high-affinity non-catalytic cellulose binding domains in microbial plant cell wall hydrolases.” *Molecular Microbiology* 11 (2): 375-382. doi:10.1111/j.1365-2958.1994.tb00317.x.

Moser, B., N. R. Gilkes, D. G. Kilburn, R. A. Warren, and R. C. Miller. 1989. “Purification and characterization of endoglucanase C of *Cellulomonas fimi*, cloning of the gene, and analysis of in vivo transcripts of the gene.” *Applied and Environmental Microbiology* 55 (10): 2480-2487.

Norton, J., A. Aaronson, K. Menzie, C. Skelly, J. Love, and S. Shagam. 2012. *World Agricultural Supply and Demand Estimates*. *World*. United States Department of Agriculture, Arlington, Virginia.

Notenboom, V., C. Birsan, R. A. Warren, S. G. Withers, and D. R. Rose. 1998. "Exploring the cellulose/xylan specificity of the beta-1,4-glycanase cex from *Cellulomonas fimi* through crystallography and mutation." *Biochemistry* 37 (14): 4751-8. doi:10.1021/bi9729211.

Le Nours, J., L. Anderson, D. Stoll, H. Stålbrand, and L. L. Leggio. 2005. "The structure and characterization of a modular endo-beta-1,4-mannanase from *Cellulomonas fimi*." *Biochemistry* 44 (38): 12700-12708. doi:10.1021/bi050779v.

Ong, E., G. Kilburn, R. C. Miller, and R. A. Warren. 1994. "*Streptomyces lividans* Glycosylates the Linker Region of a 3-1,4-Glycanase from *Cellulomonas fimi*." *Journal of Bacteriology* 176 (4): 999-1008. doi:10.1111/j.1751-7915.2010.00165.x.

O'Neill, G., S. H. Goh, R. A. Warren, D. G. Kilburn, and R. C. Miller. 1986. "Structure of the gene encoding the exoglucanase of *Cellulomonas fimi*." *Gene* 44 (2–3): 325-330.

Pang, A., M. J. Warren, and R. W. Pickersgill. 2011. "Structure of PduT, a trimeric bacterial microcompartment protein with a 4Fe-4S cluster-binding site." *Acta Crystallographica. Biological Crystallography* 67 (2): 91-6. doi:10.1107/S0907444910050201.

Patrinos, A. N., and R. Staffin. 2005. *Genomics:GTL Roadmap, DOE/SC-0090*. United States Department of Energy, Oakridge National Laboratory, Tennessee.

Perlack, R.D., and B. J. Stokes. 2011. *US Billion-ton Update: Biomass supply for a Bioenergy and Bioproducts Industry*. United States Department of Energy, Oakridge National Laboratory, Tennessee.

Poon, D. K. Y., S. G. Withers, and L. P. McIntosh. 2007. "Direct demonstration of the flexibility of the glycosylated proline-threonine linker in the *Cellulomonas fimi* Xylanase Cex through NMR spectroscopic analysis." *The Journal of biological chemistry* 282 (3): 2091-2100. doi:10.1074/jbc.M609670200.

Qiu, J. 2009. "China's climate target: Is it achievable?" *Nature News* 462: 550-1.

Quiroz-Castañeda, R. E., C. Martínez-Anaya, L. I. Cuervo-Soto, L. Segovia, and J. L. Folch-Mallol. 2011. "Loosenin, a novel protein with cellulose-disrupting activity from *Bjerkandera adusta*." *Microbial Cell Factories* 10 (8): 1-9. doi:10.1186/1475-2859-10-8.

Rabinovich, M. L., M. S. Melnick, and A. V. Bolobova. 2002. "The Structure and Mechanism of Action of Cellulolytic Enzymes." *Biochemistry* 67 (8): 850-871.

Reese, E. T., R. G. H. Siu, and H. S. Levinson. 1950. "The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis." *Journal of Bacteriology* 59 (4): 485-497.

Saloheimo, M., M. Paloheimo, S. Hakola, J. Pere, B. Swanson, E. Nyyssnen, A. Bhatia, M. Ward, and M. Penttil. 2002. "Swollenin, a *Trichoderma reesei* protein with sequence similarity to the plant expansins, exhibits disruption activity on cellulosic materials." *European Journal of Biochemistry* 269 (17): 4202-4211. doi:10.1046/j.1432-1033.2002.03095.x.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*. 4th ed. New York: Cold spring harbour laboratory press.

Sandercock, L. E., A. Meinke, R. A. J. Warren, N. R. Gilkes, and D. G. Kilburn. 1996. "Degradation of cellulases in cultures of *Cellulomonas fimi*." *FEMS Microbiology Letters* 143: 7-12.

Scheller, H. V., and P. Ulvskov. 2010. "Hemicelluloses." *Annual Review of Plant Biology* 61: 263-289. doi:10.1146/annurev-arplant-042809-112315.

Schirmer, A., M. A. Rude, X. Li, E. Popova, and S. B. del Cardayre. 2010. "Microbial biosynthesis of alkanes." *Science* 329 (5991): 559-562. doi:10.1126/science.1187936.

Shen, H., N. R. Gilkes, D. G. Kilburn, R. C. Miller, R. A. Warren. 1995. "Cellobiohydrolase B, a second exo-cellobiohydrolase from the cellulolytic bacterium *Cellulomonas fimi*." *The Biochemical Journal* 311: 67-74.

Shen, H., M. Schmuck, I. Pilz, N. R. Gilkes, D. G. Kilburn, R. C. Miller, R. A. Warren. 1991. "Deletion of the linker connecting the catalytic and cellulose-binding domains of endoglucanase A (CenA) of *Cellulomonas fimi* alters its conformation and catalytic activity." *The Journal of Biological Chemistry* 266 (17): 11335-11340.

Shoseyov, O., Z. Shani, and I. Levy. 2006. "Carbohydrate binding modules: biochemical properties and novel applications." *Microbiology and Molecular Biology Reviews* 70 (2): 283-295. doi:10.1128/MMBR.00028-05.

Sissine, F. 2007. *Energy Independence and Security Act of 2007*. United States Department of Energy, Washington, DC.

Smith, K. A., J. O. Sullivan, K. R. Vincent, J. L. Barden, P. D. Martin, C. Michael, L. Mellish, D. R. Kearney, and B. T. Murphy. 2011. *US Energy Information Administration: International Energy Outlook 2011*. United States Energy Information Administration, Washington, DC.

Somerville, C., S. Bauer, G. Brininstool, M. Facette, T. Hamann, J. Milne, E. Osborne. 2004. "Toward a systems approach to understanding plant cell walls." *Science* 306: 2206-11. doi:10.1126/science.1102765.

Sommer, M. O., G. M. Church, and G. Dantas. 2010. "A functional metagenomic approach for expanding the synthetic biology toolbox for biomass conversion." *Molecular Systems Biology* 6 (360). doi:10.1038/msb.2010.16.

Steen, E. J., Y. Kang, G. Bokinsky, Z. Hu, A. Schirmer, A. McClure, S. B. Del Cardayre, and J. D. Keasling. 2010. "Microbial production of fatty-acid-derived fuels and chemicals from plant biomass." *Nature* 463 (7280): 559-562. doi:10.1038/nature08721.

Stoll, D., A B., H. Stålbrand, B. W. McLean, D. G. Kilburn, and R. A. Warren. 2000. "Mannanase Man26A from *Cellulomonas fimi* has a mannan-binding module." *FEMS Microbiology Letters* 183 (2): 265-269.

Stoll, D., H. Stålbrand, and R. A. Warren. 2001. Hydrolysis of Mannans by Cellulomonas fimi. In *ACS Symposium Series,* 769:91-98. Washington, DC.: American Chemical Society. doi:10.1021/bk-2001-0769.ch005.

Stoll, D., H. Stålbrand, and R. A. J. Warren. 1999. "Mannan-Degrading Enzymes from *Cellulomonas fimi*." *Applied and Environmental Microbiology* 65(6): 2598-2605.

Stålbrand, H., S. D. Mansfield, J. N. Saddler, D. G. Kilburn, R. A. Warren, and N. R. Gilkes. 1998. "Analysis of Molecular Size Distributions of Cellulose Molecules during Hydrolysis of Cellulose by Recombinant *Cellulomonas fimi* β-1,4-Glucanases." *Applied and Environmental Microbiology* 64 (7): 2374-2379.

Sun, W., S. Wang, and R. Curtiss. 2008. "Highly efficient method for introducing successive multiple scarless gene deletions and markerless gene insertions into the *Yersinia pestis* chromosome." *Applied and Environmental Microbiology* 74 (13): 4241-4245. doi:10.1128/AEM.00940-08.

Tabor, S. 2001. "Expression using the T7 RNA polymerase/promoter system." *Current protocols in molecular biology* 16 (2). doi:10.1002/0471142727.mb1602s11.

Tan, N. S., B. Ho, and J. Ling Ding. 2002. "Engineering a novel secretion signal for cross-host recombinant protein expression." *Protein Engineering* 15 (4): 337-345.

Teather, R. M., and P. J. Wood. 1982. "Use of Congo Red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen." *Applied and Environmental Microbiology* 43 (4): 777-80.

Temme, K., D. Zhao, and C. A. Voigt. 2012. "Refactoring the nitrogen fixation gene cluster from *Klebsiella oxytoca*." *Proceedings of the National Academy of Sciences of the United States of America* 109 (18): 7085-7090. doi:10.1073/pnas.1120788109.

Tomme, P., A. Boraston, B. McLean, J. Kormos, A. L. Creagh, K. Sturch, N. R. Gilkes, C. A. Haynes, R. A. Warren, and D. G. Kilburn. 1998. "Characterization and affinity applications of cellulose-binding domains." *Journal of Chromatography* 715 (1): 283-96.

Tomme, P., E. Kwan, N. R. Gilkes, D. G. Kilburn, and R. A. Warren. 1996. "Characterization of CenC, an enzyme from *Cellulomonas fimi* with both endo- and exoglucanase activities." *Journal of Bacteriology* 178 (14): 4216-4223.

Turmes, C. 2009. *EU renewables directive 2009/28/EC*. The Parliament and the Council of the European Union, Luxembourgh.

Wakarchuk, W. W., D. G. Kilburn, R. C. Miller, and R. A Warren. 1984. "The Preliminary Characterization of the beta-Glucosidases of *Cellulomonas fimi*." *Journal of General Microbiology* 130: 1385-1389.

Wang, J., M. Sarov, J. Rientjes, J. Fu, H. Hollak, H. Kranz, W. Xie, F. Stewart, and Y. Zhang. 2006. "An improved recombineering approach by adding RecA to lambda Red recombination." *Molecular Biotechnology* 32 (1): 43-53.

Wargacki, J., E. Leonard, M. N. Win, D. D. Regitsky, C. N. S. Santos, P. B. Kim, S. R. Cooper. 2012. "An Engineered Microbial Platform for Direct Biofuel Production from Brown Macroalgae." *Science* 335 (6066): 308-313. doi:10.1126/science.1214547.

Warren, R. A. J. 1996. "Microbial hydrolysis of polysaccharides." *Annual Review of Microbiology* 50: 183-212. doi:10.1146/annurev.micro.50.1.183.

Warren, R. A. J., C. F. Beck, N. R. Gilkes, D. G. Kilburn, M. L. Langsford, R. C. Miller, G. P. O'Neill, M. Scheufens, and W. K. R. Wong. 1986. "Sequence conservation and region shuffling in an endoglucanase and an exoglucanase from *Cellulomonas fimi*." *Proteins: Structure, Function, and Bioinformatics* 1 (4): 335-341. doi:10.1002/prot.340010407.

Westfall, P. J., and T. S. Gardner. 2011. "Industrial fermentation of renewable diesel fuels." *Current Opinion in Biotechnology* 22 (3) : 344-50. doi:10.1016/j.copbio.2011.04.023.

White, A., S. G. Withers, N. R. Gilkes, and D. R. Rose. 1994. "Crystal structure of the catalytic domain of the beta-1,4-glycanase cex from *Cellulomonas fimi*." *Biochemistry* 33: 12546-12552.

Wilson, D. 2008. "Three Microbial Strategies for Plant Cell Wall Degradation." *Ann. N.Y. Acad. Sci.* 1125 (March): 289-297. doi:10.1196/annals.1419.026.

Wong, W. K., B. Gerhard, Z. Min Guo, D. G. Kilburn, R. Anthony, J. Warren, and R. C. Miller. 1986. "Characterization and structure of an endoglucanase gene cenA of *Cellulomonas fimi*." *Gene* 44 (2–3): 315-324.

Zechel, D. L., S. P. Reid, D. Stoll, O. Nashiru, R A. J. Warren, and S. G. Withers. 2003. "Mechanism, mutagenesis, and chemical rescue of a beta-mannosidase from *Cellulomonas fimi*." *Biochemistry* 42 (23): 7195-7204. doi:10.1021/bi034329j.

Zhang, Y. P., and L. R. Lynd. 2004. "Toward an Aggregated Understanding of Enzymatic Hydrolysis of Cellulose: Noncomplexed Cellulase Systems." *Biotechnology and Bioengineering* 88 (7): 797-824. doi:10.1002/bit.20282.