Expression and Characterization of Exoglucanases in *Saccharomyces cerevisiae*

Niël van Wyk

Dissertation presented for the degree of Doctor of Microbiology at Stellenbosch University

Supervisor: W.H. van Zyl
Co-supervisor: R. den Haan
March 2010

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the owner of the copyright thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature:……. Date: .....

SUMMARY

Currently a world-wide tendency exists to shift away from relying on fossil fuels as a primary energy source and to focus on sustainable, environmentally-friendly alternatives. Ethanol is one such alternative and shows potential to replace petroleum as a transport fuel. Plant biomass, deemed a renewable energy source, can be converted to ethanol. The process of conversion via biologically-mediated events is problematic mainly due to the recalcitrance of the chief components of plant biomass namely cellulose, hemicellulose and lignin towards enzymatic degradation. A concept of consolidated bioprocessing (CBP) aims to make the process of bioconversion of plant biomass to ethanol cost-effective. For such a bioconversion, a biocatalyst is needed which can depolymerize the complex carbohydrates i.e. the cellulose and hemicellulose to their respective monomers for concurrent fermentation to ethanol. *Saccharomyces cerevisiae* shows potential as a candidate CBP-biocatalyst due to its high ethanol productivity, general robustness and amenability to genetic manipulation. However, *S. cerevisiae* does not possess the ability to break down the abovementioned carbohydrates.

This study attempted to address certain aspects of yeast strain development for CBP. Genes encoding cellulases responsible for major crystalline cellulose hydrolysis i.e. exoglucanases were expressed in *S. cerevisiae* and the recombinant proteins were characterized. Further work involved exploring different ways of increasing the cellulolytic capability of recombinant *S. cerevisiae*.

Both the *cel9A* of *Thermobifida fusca* and *Npcel6A* of *Neocallimastix patriciarum* were functionally expressed in *S. cerevisiae*. Expression of *cel9A* enabled *S. cerevisiae* to grow on phosphoric acid swollen cellulose reaching an aerobic growth rate (μ_{MAX}) of 0.088 h^{-1}. This is the first report of *S. cerevisiae* growing on such a substrate while producing only one heterologous protein. An increase in the cellulolytic capability of recombinant *S. cerevisiae* was observed when *cel9A* was co-expressed with *Trcel6A*, *cel7A* and *cel7B* of *Trichoderma reesei*. These results proved that the synergy between cellulases can contribute towards increasing the cellulolytic capability of recombinant *S. cerevisiae*.

NpCel6A has the highest reported individual activity on a crystalline cellulose substrate. However, expression of *Npcel6A* by *S. cerevisiae* resulted in lower levels of exoglucanase activity on Avicel of 0.540±0.062 mU/gDCW compared to the recombinant *S. cerevisiae* strains that produces Cel6A of *T. reesei* (4.101±0.243 mU/gDCW). This observation could be ascribed to glycosylation of the catalytic domain of NpCel6A. The replacement of the carbohydrate-binding module (CBM) and asparagine-rich linker of NpCel6A with the CBM and serine/threonine-rich linker of TrCel6A resulted in a decrease in recombinant cellulolytic activity produced by *S. cerevisiae*. In contrast, when the CBM and linker of NpCel6A were appended to the N-terminus of the catalytic domain of TrCel6A, significantly higher levels of cellulase activity were observed when produced by *S. cerevisiae*. This observation was largely attributed to the difference in glycosylation of the linkers. These results showed the value of domain swapping for obtaining increased cellulase secretion by *S. cerevisiae*.

The native *S. cerevisiae* genes *PSE1* and *SOD1*, were individually overexpressed in the *S. cerevisiae* strain producing NpCel6A, Cel3A of *Saccharomycopsis fibuligera* and Cel7B of *Trichoderma reesei*. The *DDI1* gene of *S. cerevisiae* was also disrupted in the strain producing NpCel6A. In all cases, transformants were identified which displayed higher levels of cellulase activity compared to the original strain. This demonstrated the potential of *S. cerevisiae* to be considered as a "chassis"-strain that can, with the help of metabolic engineering, produce more recombinant cellulases.

The swelling factor protein called swollenin, a contributor in the disruption of the crystallinity of cellulose, was co-expressed with *cel9A* and *Npcel6A* individually in *S. cerevisiae.* Even though functionality of swollenin was confirmed, no noteworthy increase in the levels of cellulase activity was observed for recombinant strains.

The recombinant yeast strains generated during this study represent significant progress towards developing *S. cerevisiae* as a CBP organism.

OPSOMMING

Tans heers daar wêreldwyd ‘n tendens om weg te beweeg vanaf fossielbrandstowwe as ‘n primêre energiebron en om te fokus op volhoubare, omgewingsvriendelike alternatiewe. Etanol is een só ‘n alternatief en toon potensiaal om petroleum as ‘n vervoerbandstof te vervang. Plantbiomassa, wat as ‘n hernubare energiebron beskou word, kan na etanol omgeskakel word. Die proses van omskakeling via biologies-gefasiliteerde gebeurtenisse is problematies hoofsaaklik aangesien die hoofkomponente van plantbiomassa naamlik sellulose, hemisellulose en lignien weerstandig is teen ensimatiese afbraak. ‘n Konsep genaamd gekonsolideerde bioprosessering (CBP) poog om die proses van bio-omskakeling van plantbiomassa na etanol koste-effektief te maak. Vir só ‘n bio-omskakeling, word ‘n biokatalis benodig om die komplekse koolhidrate i.e. die sellulose en hemisellulose te kan depolimeriseer tot hul onderskeie monomere en tegelykertyd te fermenteer na etanol. *Saccharomyces cerevisiae* toon potensiaal as ‘n kandidaat CBP-biokatalis vanweë sy hoë etanol-produktiwiteit, algehele robuustheid en geskiktheid vir genetiese manipulering. *S. cerevisiae* besit egter nie die vermoë om bogenoemde koolhidraatpolimere af te breek nie.

Hierdie studie het gepoog om sekere aspekte van gisrasontwikkeling vir CBP te adresseer. Sellulase-koderende gene wat in staat is om kristallyne sellulose te hidroliseer naamlik eksoglukanases, is in *S. cerevisiae* uitgedruk en die rekombinante proteïene is gekarakteriseer. Verdere werk het behels die verkenning van verskillende maniere om die sellulolitiese vermoëns van rekombinante *S. cerevisiae* te verbeter.

Beide die *cel9A* van *Thermobifida fusca* en *Npcel6A* van *Neocallimastix patriciarum* is funksioneel uitgedruk in *S. cerevisiae*. Uitdrukking van *cel9A* het *S. cerevisiae* die vermoë gegee om op fosforsuur-geswelde sellulose te groei teen ’n groeitempo (μ_{MAX}) van 0.088 h^{-1}. Dit is die eerste melding van *S. cerevisiae* wat kon groei op so ‘n substraat deur net een heteroloë proteïen te produseer. Daar is ook ‘n toename in die sellulolitiese vermoëns van *S. cerevisiae* waargeneem toe *cel9A* saam met *Trcel6A*, *cel7A* en *cel7B* van *Trichoderma reesei* uitgedruk is wat bewys dat die sinergie tussen sellulases kan bydra tot ‘n toename in die sellulolitiese vermoëns van *S. cerevisiae*.

NpCel6A het die hoogste aangemelde indiwiduele aktiwiteit op ‘n kristallyne sellulose substraat. Rekombinante uitdrukking van *Npcel6A* in *S. cerevisiae* het egter laer eksoglukanase aktiwiteit op

Avicel getoon (0.540±0.062 mU/gDCW) in vergelyking met die rekombinante *S. cerevisiae* ras wat Cel6A van *T. reesei* produseer (4.101±0.243 mU/gDCW). Die waarneming kan aan die glikosilering van die katalitiese domein van NpCel6A toegeskryf word. Die vervanging van die koolhidraat-bindingsmodule (CBM) en asparagien-ryke koppeleenheid van NpCel6A met die CBM en serien/treonien-ryke koppeleenheid van TrCel6A het tot 'n afname in rekombinante sellulase aktiwiteit deur *S. cerevisiae* gelei. In teenstelling, toe NpCel6A se CBM en koppeleenheid voor die katalitiese domein van TrCel6A geplaas is, het dit tot 'n beduidende hoër sellulase aktiwiteit deur *S. cerevisiae* gelei. Dié waarneming is grootliks toegeskryf aan die verskil in glikosilering van die koppeleenhede. Hierdie resultate bewys die waarde wat domein-vervanging kan hê om 'n toename in sellulase aktiwiteit waar te neem in rekombinante *S. cerevisiae*.

Gene eie aan *S. cerevisiae, PSE1* en *SOD1,* is indiwidueel ooruitgedruk in die *S. cerevisiae* ras wat NpCel6A, Cel3A van *Saccharomycopsis fibuligera* en Cel7B van *T. reesei* produseer. Die *DDI1* geen van *S. cerevisiae* is ook ontwrig in die ras wat NpCel6A produseer. In alle gevalle is *S. cerevisiae* transformante geïdentifiseer wat hoër vlakke van sellulase aktiwiteit getoon het. Dit wys die potensiaal om *S. cerevisiae* as 'n "onderstel"-organisme te beskou waarvolgens dit, met behulp van metaboliese ingenieurswese, meer heteroloë sellulases kan produseer.

Die swelfaktorproteïen swollenin, wat bydra tot die ontwrigting van die kristalliniteit van sellulose, is saam met *cel9A* en *Npcel6A* onderskeidelik in *S. cerevisiae* uitgedruk. Die funksionaliteit van swollenin is bevestig, maar geen noemenswardige toename in sellulase aktiwiteit vir die rekombinante rasse is gevind nie.

Die rekombinante gisrasse wat tydens hierdie studie gegenereer is, dui op beduidenswaardige vordering in die ontwikkeling van *S. cerevisiae* as 'n CBP organisme.

To my parents, for all the sacrifices you made to further my education.

Acknowledgments

Just like no man and no gene, I believe that no thesis is an island and would like to extent my sincere gratitude toward the following:

- My parents who supported me (financially, emotionally) through all my years at university.
- Janean Jeppe who made life on campus infinitely easier for me.
- Nina van Wyk, it's now more than ten years.
- Aletta Eksteen and Angie Basson with whom I've walked a road from my second year on campus.
- Lisa du Plessis, Nicolette Fouché, Elinda Viljoen and Danie Diedericks who were my early-morning coffee buddies that became a bad, but essential habit.
- All my fellow colleagues in lab 335 that in some way or another helped me (some much more than others) during my stay. Let's see if I can remember everyone (in no particular order): Annie Chimpango, Dr Danie la Grange, Dr Jacques Smith, Dr Tania de Villiers, Gillian de Villiers, Dr Ronél van Rooyen, Jan-Nico Coetzee, Ike James, Lara Kotze, Annatjie Hugo, Arista Kloppers, Christa Sadie, Helba Bredell, Ilse Wepener, Astrid Sidle and Isa Marx (although not formally part of the lab).
- Everyone in the Microbiology Department.
- All the institutions that financially aided me during my study: the National Research Foundation, the Harry Crossley Foundation, the Ethel and Ernst Eriksen Trust and the University of Stellenbosch.
- Prof. Florian Bauer, Dr Amanda Swart and Prof. Marinda Bloom who granted me a Masters upgrade.
- My examiners Dr Heinrich Volschenk, Prof. Lisbeth Olsson and Prof. Edward Bayer for granting me the degree and for their valuable comments of the thesis.
- My apprentice Heinrich Kroukamp for making a dreadful 2009 a little bit pleasant in the lab.
- Dr Shaunita Rose for being the only one who took a scientific interest in my project. You were always at hand with extremely valuable tips, had an admirable work-ethic and never seemed too irritated (maybe I refused to see) when I came for help.
- Dr Riaan den Haan for being my co-supervisor for the 5 years I've been in lab 335 and especially for reading and re-reading the manuscript.
- Prof. Willem Heber van Zyl who acted as my supervisor while I was in lab 335. It is a well-established and well-equipped lab in which I could do my experiments, because of all the work that was done mostly by you way before I became part of the lab.
- The Lord Almighty for the strength and guidance.

PREFACE

The thesis is presented as a compilation of scientific manuscripts. The experimental chapters (Chapter 3, 4 and 5) are written in a style of a journal to which the manuscripts were submitted. The experimental chapters (Chapters 6 and 7) are currently not considered for publication.

Chapter 3 CO-EXPRESSION OF *THERMOBIFIDA FUSCA cel9A* WITH OTHER CELLULASES IN *SACCHAROMYCES CEREVISIAE*
Submitted for publication to Applied Microbiology and Biotechnology

Chapter 4 IMPROVEMENT OF THE HETEROLOGOUS PRODUCTION OF NpCel6A FROM *NEOCALLIMASTIX PATRICIARUM* IN *SACCHAROMYCES CEREVISIAE*
Published in Enzyme and Microbial Technology Vol 46: 378-383 (2010).

Chapter 5 OVEREXPRESSION OF NATIVE *SACCHAROMYCES CEREVISIAE* GENES FOR IMPROVEMENT OF HETEROLOGOUS PROTEIN PRODUCTION
In preparation for publication

Chapter 6 EFFECT OF THE DISRUPTION OF THE *DDI1* GENE ON THE HETEROLOGOUS PRODUCTION OF *NEOCALLIMASTIX PATRICIARUM* Cel6A BY *SACCHAROMYCES CEREVISIAE*

Chapter 7 THE ROLE OF SWOLLENIN ON THE CELLULOLYTIC EFFECT OF *SACCHAROMYCES CEREVISIAE* WHEN HETEROLOGOUSLY CO-EXPRESSED WITH CELLULASES.

Appendix A A METHOD OF CONVERTING CELLULOSICS TO FERMENTING CELLULOSE. A South African Provisional Patent (nr. 2007/06861) that has been filed and is an expansion to work done in chapter 3.

INDEX

LIST OF ABBREVIATIONS

3D	three-dimensional
5-HMF	5-hydroxymethyl furfural
ATP	adenosyl triphosphate
BC	bacterial cellulose
BGL	β-glucosidase
BMCC	bacterial microcrystalline cellulose
C5	pentoses
C6	hexoses
CAI	codon adaptation index
CAZy	carbohydrate-active enzymes
CBI	codon bias index
CBP	consolidated bioprocessing
CBH	cellobiohydrolase
CBM	carbohydrate-binding module
CD	catalytic domain
CMC	carboxymethylcellulose
DCW	dry cell weight
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DNS	dinitrosalicyclic acid
DP	degree of polymerization
E10,85,95,100	ethanol and percentage blend with petroleum
E.C.	enzyme class
EG	endoglucanase
ELISA	enzyme-linked immunosorbent assay
GH	glycoside hydrolase
GRAS	generally recognized as safe
HEC	hydroxyethylcellulose
HPAEC	high-performance anionic exchange chromatography
IUB-MB	International union of biochemistry and molecular biology
LiOAc	lithium acetate
*p*NPC	*p*-nitrophenyl β-D-glucopyranoside
*p*NPG	*p*-nitrophenyl β-D-cellobioside
*p*NPP	*p*-nitrophenylphosphate
OPEC	Organization of petroleum-exporting countries
ORF	open reading frame
PASC	phosphoric acid-swollen cellulose
PCR	polymerase chain reaction
mRNA	messenger-ribonucleic acid
ROS	reactive oxygen species
rpm	revolutions per minute
SC	synthetic complete
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SHF	separate hydrolysis and fermentation
spp	species
SSCF	simultaneous saccharification and co-fermentation
SSF	simultaneous saccharification and fermentation
TC	terminal complex
TNP	trinitrophenyl-
UDP	uridine diphospho-
UNFCCC	United Nations framework convention on climate change
USA	United States of America
YNB	yeast nitrogen base
YPD	yeast extract peptone dextrose

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

GENERAL INTRODUCTION AND PROJECT AIMS

1.1. GENERAL INTRODUCTION

There is great demand world-wide to increase the output from alternative energy resources and to minimize the world's dependency on fossil fuel-based energy resources. The overall negative impact on the environment, the finiteness of fossil fuels and energy security are all major driving forces for the search and development of alternatives. These alternative energy resources encompass every aspect of energy usage and include amongst many others solar, wind, nuclear and tidal energy resources. Currently, none of the abovementioned energy resources can be considered as feasible candidates for the replacement of fossil fuels in the existing transport industry. Ethanol, however, which was used in the first car engines, received a resurgence in appeal as a viable candidate to replace fossil fuels ever since the energy crises in the 1970s (Putsche & Sandor, 1996). The majority of ethanol is produced from starch and sugar-based resources albeit current practices would not be sustainable for future demands (Hammerschlag, 2006). By producing ethanol from the complex carbohydrates found in lignocellulose instead of starch and sucrose-based resources could address these shortcomings. Lignocellulose or plant biomass, which is largely composed of cellulose, hemicellulose and lignin, is much more difficult to break down than starch and sucrose and technologies for the conversion of lignocellulose to ethanol are largely in its developmental phase (Hahn-Hägerdal et al., 2006). One decidedly expensive step in the process is the addition of enzymes, most notably cellulases, to lignocellulose (Stephanopolous, 2007).

A process strategy called consolidated bioprocessing (CBP) addresses the high cost of enzyme-additions (Lynd et al., 2005). When realized, CBP can dramatically reduce the cost involved in the large-scale conversion of lignocellulose to ethanol. What CBP entails is to combine all the biologically-mediated transformations necessary for the conversions into a single step. CBP requires an organism, or biocatalyst, which is able to hydrolyze, or depolymerize, the cellulose and hemicellulose to the respective monomers and ferment the monomers to ethanol.

Naturally, such an organism has not yet been described although many organisms exist that, with the aid of genetic engineering, could potentially be utilized in CBP-applications.

Baker's yeast, *Saccharomyces cerevisiae,* is one such organism that holds tremendous promise as a candidate organism for CBP (Van Zyl et al., 2007). Amongst its favourable characteristics are its ability to vigorously ferment simple sugars (hexoses) to ethanol, its tolerance to high concentrations of ethanol and certain byproducts of lignocellulosic hydrolysates and its well-described genetic system. The major drawback of *S. cerevisiae* for consideration as a CBP-organism is that it cannot depolymerize either cellulose or hemicellulose. Thus, to address the shortcoming of *S. cerevisiae* many have attempted to introduce cellulase- and hemicellulase-coding genes in *S. cerevisiae* to enable the yeast to hydrolyze these substrates. With regard to the construction of a cellulose-hydrolyzing or cellulolytic *S. cerevisiae*, breakthroughs include the description of yeasts capable of growing of cellobiose and amorphous cellulose while concomitantly producing ethanol in both cases (Den Haan et al., 2007a; Van Rooyen et al., 2005). There is, however, major progress needed before *S. cerevisiae* can be considered as a feasible CBP-organism especially since it was calculated that at least a 100-fold increase in cellulase production is needed before substantial growth on a crystalline source of cellulose can be observed (Den Haan et al., 2007b).

1.2. AIMS OF PRESENT STUDY

The principal aim of the study is to engineer cellulolytic *S. cerevisiae* strains and to explore several ways of improving the cellulolytic capability of *S. cerevisiae*.

The objectives identified to satisfy these aims are:

- To heterologously express *cel9A* of *Thermobifida fusca* in *S. cerevisiae* and characterize the recombinant protein
- To investigate to what extent the synergy that exists between Cel9A and other cellulases could contribute toward increasing cellulolytic capability when co-expressed in *S. cerevisiae*

- ∠ To heterologously express *Npcel6A* of *Neocallimastix patriciarum* in *S. cerevisiae* and characterize the recombinant protein
- ∠ To investigate what effect replacing the carbohydrate binding module and linker region of NpCel6A with that of TrCel6A of *Trichoderma reesei* could have on its production in *S. cerevisiae*
- ∠ To overexpress and disrupt native genes in *S. cerevisiae* and determine its effect on heterologous protein production
- ∠ To heterologously express the swelling factor swollenin *swoI* of *Trichoderma reesei* in *S. cerevisiae* and characterize the recombinant protein
- ∠ To investigate the effect of swollenin when co-expressed with cellulases in *S. cerevisiae*

1.3. LITERATURE CITED

Den Haan R, Rose SH, Lynd LR, Van Zyl WH 2007a. Hydrolysis and fermentation of amorphous cellulose by recombinant *Saccharomyces cerevisiae*. Metab Eng 9: 87-94.

Den Haan R, Mcbride JE, La Grange DC, Lynd LR, Van Zyl WH 2007b. Functional expression of cellobiohydrolases in *Saccharomyces cerevisiae* towards one-step conversion of cellulose to ethanol. Enzyme Microb Technol 40: 1291-1299.

Hahn-Hägerdal B, Galbe M, Gorwa-Grauslund MF, Lidén G, Zacchi G 2006. Bio-ethanol – the fuel of tomorrow from the residues of today. Trends Biotechnol 24: 549-556.

Hammerschlag R 2006. Ethanol's energy return on investment: A survey of the literature 1990 – present. Environ Sci Technol 40: 1744-1750.

Lynd LR, Van Zyl WH, McBride JE, Laser M 2005. Consolidated bioprocessing of cellulosic biomass: an update. Curr Opin Biotechnol 16: 577-583.

Putsche V, Sandor D 1996. Strategic, economic, and environmental issues for transportation fuels. In: Handbook on bioethanol: production and utilization. (Wyman ed.), Taylor and Francis, Washington DC: pp 21-35.

Stephanopoulos G 2007. Challenges in engineering microbes for biofuels production. Science 315: 801-803.

Van Rooyen R, Hahn-Hägerdal B, La Grange DC, Van Zyl WH 2005. Construction of cellobiose-growing and fermenting *Saccharomyces cerevisiae* strains. J Biotechnol 120: 284-295.

Van Zyl WH, Lynd LR, Den Haan R, McBride JE 2007. Consolidated bioprocessing for ethanol production using *Saccharomyces cerevisiae*. Adv Biochem Eng Biot 108: 205-235.

CHAPTER 2

THE USE OF *SACCHAROMYCES CEREVISIAE* AS ORGANISM FOR CONSOLIDATED BIOPROCESSING OF CELLULOSE-CONTAINING MATERIALS

2.1. CELLULOSE AND CELLULASES

2.1.1. CELLULOSE

2.1.1.1. A SHORT HISTORY ON THE DISCOVERY OF CELLULOSE

The French chemist Anselme Payen described in 1838 a resistant fibrous solid that remained behind after a variety of plant tissues were treated with acids and ammonia along with a subsequent extraction with water, alcohol and ether (Payen, 1838). Payen also determined the molecular formula to be $C_6H_{10}O_5$ and observed it to be isomerically similar to starch. The term "cellulose" was first coined in 1839 in a report of the French Academy of Science on the work of Payen.

Prior to its discovery and designation, cellulose was used for thousands of years by humans as building materials, an energy source and clothing. Ever since ancient Egyptians used cellulose fibres as writing material known as papyrus, cellulose has played an integral part in shaping human culture.

2.1.1.2. THE ABUNDANCE OF CELLULOSE

Cellulose, the primary structural component of plant cell walls, constitutes more than half of the bound carbon in the earth (Brown, 2004). Billions of tonnes of cellulose are synthesized annually by plants through carbon dioxide fixation. Even though the exact composition of plant cell walls differs greatly among all the plant taxa (Table 2.1), cellulose usually comprises about 35 to 50 % of plant dry weight. The stage of development of the plant cell can also affect the amount of cellulose present. Cellulose constitutes about 20 to 40 % of cell wall dry weight in primary cell walls (Wood, 1992). In secondary cell walls, this percentage increases to about 40 to 60 %. Cellulose fibres usually occur in plant cell walls as an embedded component along with hemicellulose and lignin, together forming an insoluble matrix commonly known as

lignocellulose.

Cellulose is mainly of vegetable origin, but can also be produced by several microorganisms notably the bacterial genera *Acetobacter*, *Rhizobium*, *Agrobacterium* and *Sarcina* (Delmer, 1999) and also occurs in the stiff outer mantles of certain marine invertebrates known as tunicates (Watanabe & Tokuda, 2001). Cellulose, as in plants, forms part of the cell wall structure in several bacterial species (including all cyanobacteria) and the slime mold *Dictyostelium*.

TABLE 2.1: Percentage (%) of cellulose content of selected plant biomass materials (Hon, 1996; Sun & Chen, 2002; Howard et al., 2003).

Cellulose content			
0-35%	**35-55%**	**55-80%**	**80-100%**
Wheat straw	Softwood	Hemp	Sunn
Rice straw	Bagasse	Istle	Cotton
Corn Stalks	Corn Stover	Jute	
	Grasses	Henequen	
	Kenaf	Ramie	
	Hardwood	Flax[a]	
	Coir	Sisal	
	Switchgrass		

a: includes retted and unretted

Where it is possible for some bacteria to survive without cellulose synthesis, vascular plants can not (Saxena & Brown, 2005). Apart from resisting turgor pressure in the plant cell, it also has a definite role in maintaining the size, shape and division/differentiation potential of plant cells. Structures based on old walls from dead plant cells are pivotal functional units in the formation of xylem – a crucial component in the transportation of water and water-soluble minerals throughout the plant body.

2.1.1.3. THE STRUCTURE OF CELLULOSE

Cellulose is a linear homopolymer consisting of D-anhydroglucopyranose units linked by β-(1,4)-glycosidic bonds (Wood, 1991). As a result, each glucose residue is rotated 180° with respect to the preceding and following residue along the main axis of the glucan chain, as can be seen in Figure 2.1. ^{1}H-NMR spectroscopy has shown that the β-linked D-anhydro-glucopyranoses

take on the $^4C_1$ chair conformation, the lowest free energy conformation of the molecule (Krässig, 1993). This leads to the hydroxyl groups being positioned in an equatorial (ring-like) plane, while the hydrogen atoms are in an axial position. Each pyranose ring contains free hydroxyl groups at the C-2, C-3 and C-6 atoms.

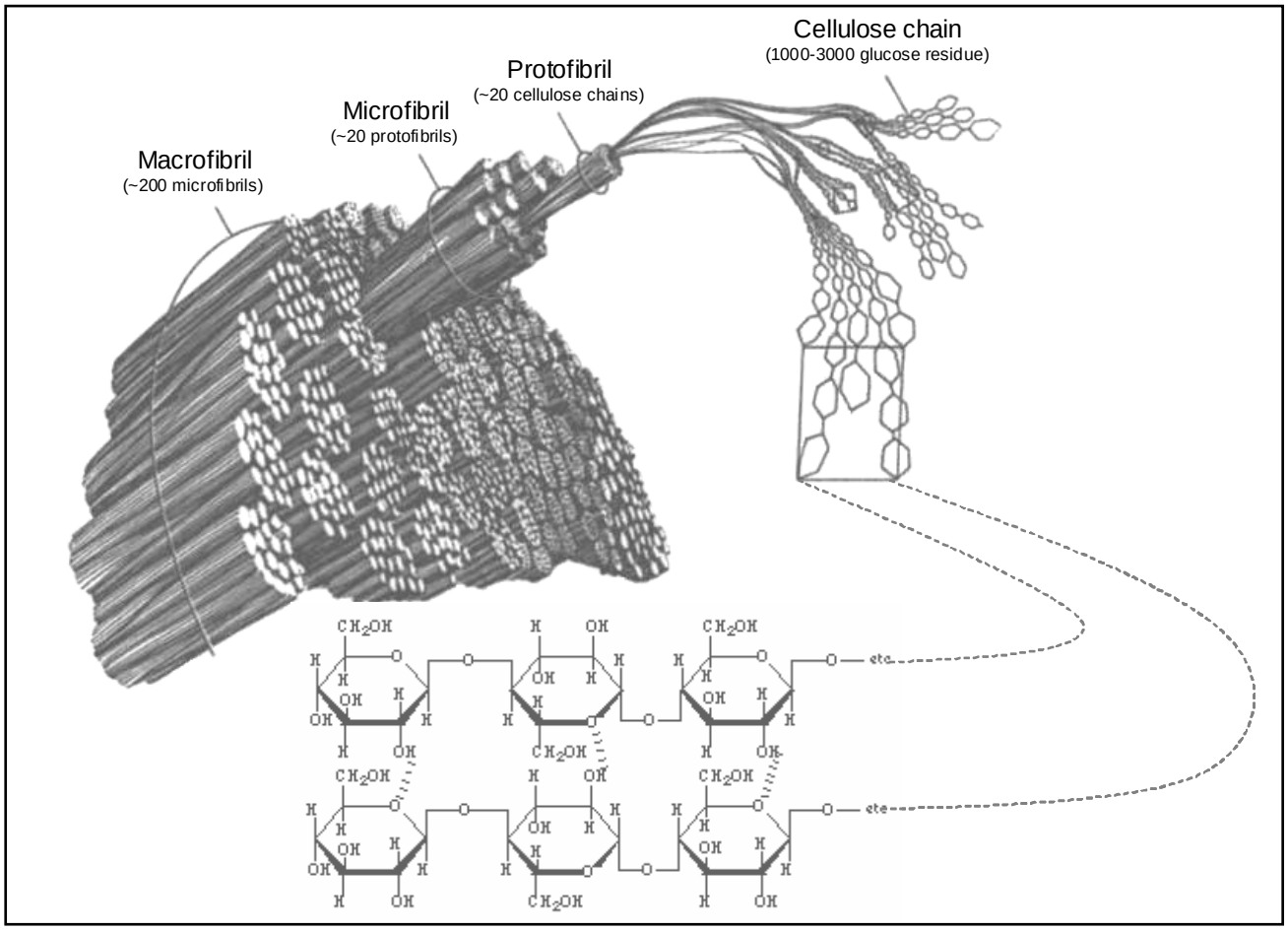


FIGURE 2.1: The primary and macromolecular structure of cellulose (Van Rooyen, 2002).

The physical properties of cellulose, which include the crystalline state, degree of crystallinity and molecular weight, are extremely variable and are dependant on the source from which the cellulose was obtained (Saxena & Brown, 2005). The arrangement of the individual glucan chains with respect to one another primarily determines the crystalline state of cellulose. It has been found that cellulose can occur in four different crystalline forms, so-called polymorphs, designated as cellulose I-IV (Eriksson et al., 1990). The most common crystalline form of cellulose in nature is cellulose I. The glucan chains of cellulose I are parallel to one another and

up to seventy neighbouring unipolar glucan chains interact with one another through hydrogen bonds and weak Van der Waals interactions and form microfibrils that are highly ordered and crystalline. Microfibrils occur as entities with irregular lateral width (Chanzy, 1990). In primary cell walls of certain plants the diameter of microfibrils are usually between 2-3 nm, but microfibrils can be ten times wider in the cell walls of some algae or tunicates. Tsekos et al. (1999) found that the microfibrillar width in the red alga *Erythrocladia subintegra* can get as wide as 68 nm.

These microfibrils combine to create fibres or macrofibrils. Due to the intermolecular and intrachain hydrogen bonds, these cellulose fibres are rendered insoluble in water. There are, however, regions within the structure of the cellulose fibre that are less ordered and are called amorphous regions. Apart from crystalline and amorphous regions within the fibre structure, it also may contain abnormalities like kinks of the microfibrils or voids such as micropores, large pits and capillaries (Blouin et al., 1970; Cowling, 1975; Fan et al., 1980). These anomalies within the crystal structure lend a measure of heterogeneity to the fibre. This means that cellulose fibres can be partially wetted when immersed in aqueous solutions. Some micropores and capillaries are even large enough to allow the penetration of large molecules like enzymes (Stone & Scallan, 1968).

High resolution ^{13}C solid-state NMR spectroscopy on cellulose I showed that it consists of two crystallographic phases or sub-allomorphs, I_α (triclinic) and I_β (monoclinic) (Atalla & VanderHart, 1984). I_α and I_β are different with respect to their crystal packing, molecular conformation and hydrogen bonding (Nishiyama et al., 2003). The I_α phase is metastable and through the process of annealing can be converted to I_β. Due to the differences in the two sub-allomorphs, one can deduce that the physical properties of fibres may vary due to the specific amount of each sub-allomorph. Sugiyama et al. (1991) found that cellulose from some bacteria and algae are mostly I_α, whereas cellulose from cotton, wood, tunicates and ramie are rich in I_β.

There are a few organisms that are able to produce cellulose II, such as mutants of the acetate-

producing bacteria *Acetobacter xylinum* (Shibazaki et al., 1998). The arrangements of the glucan chains in cellulose II are, unlike cellulose I, anti-parallel. Cellulose II is the most thermodynamically stable allomorph of cellulose due to an additional hydrogen bond per glucose residue. It contains two different anhydroglucoses (A and B) with different backbone structures (Kono & Numata, 2004). Cellulose III is formed by mercerizing cellulose with liquid ammonia below -30°C (Wada et al., 2004). It is similar to cellulose II, but the glucan chains are parallel. The last polymorph of cellulose that has been described, cellulose IV can be formed when cellulose III is treated in a glycerol-containing liquid at high temperature and pressure (Klemm et al., 2002).

Usually the size of a cellulose molecule can be given as the degree of polymerization (DP), which spans from 30 to 15 000 glucose moieties (Klemm et al., 2002). The DP in secondary cell walls of plants are usually 7 000 to 14 000 glucose moieties per molecule and can be less than 500 moieties in primary cell walls.

2.1.1.4. THE BIOGENESIS OF CELLULOSE IN PLANTS

Cellulose as found with most other large polysaccharides, has no demarcated size and in contrast with proteins and nucleic acids, there is no genetic predetermined template that steers its synthesis.

Cellulose biogenesis in vascular plants starts with the coordinated polymerization of the nucleotide sugar uridine diphospho-D-glucose (UDP-D-glucose) by enzymatic action known as chain initiation (Koyama et al., 1997). This is followed by the extrusion and simultaneous crystallization of the cellulose microfibrils. Making use of electron microdiffraction, tilting analysis and silver labelling of the reducing ends microfibril, it was revealed that cellulose biosynthesis takes place at the non-reducing end of the growing chain.

The main enzyme complexes involved in biogenesis are called terminal complexes (TCs) and are localized at the cell membranes (Scheible et al., 2001). The key enzyme in TCs is cellulose

syntase (Ces) and polymorphs of these enzymes are arranged in a rosette-like structure (Figure 2.2), however linear TCs are commonly found in algae (Mizuta & Brown, 1992). A single rosette subunit enzyme complex consists of three isoforms of Ces polypeptides (α_1, α_2 and β) encoded by the *CesA* gene family. These Ces polypeptides contain eight transmembrane helices suggesting that TCs also play a role in forming pore-like structures on the cell membrane through which the nascent chain is extruded into the cell wall.

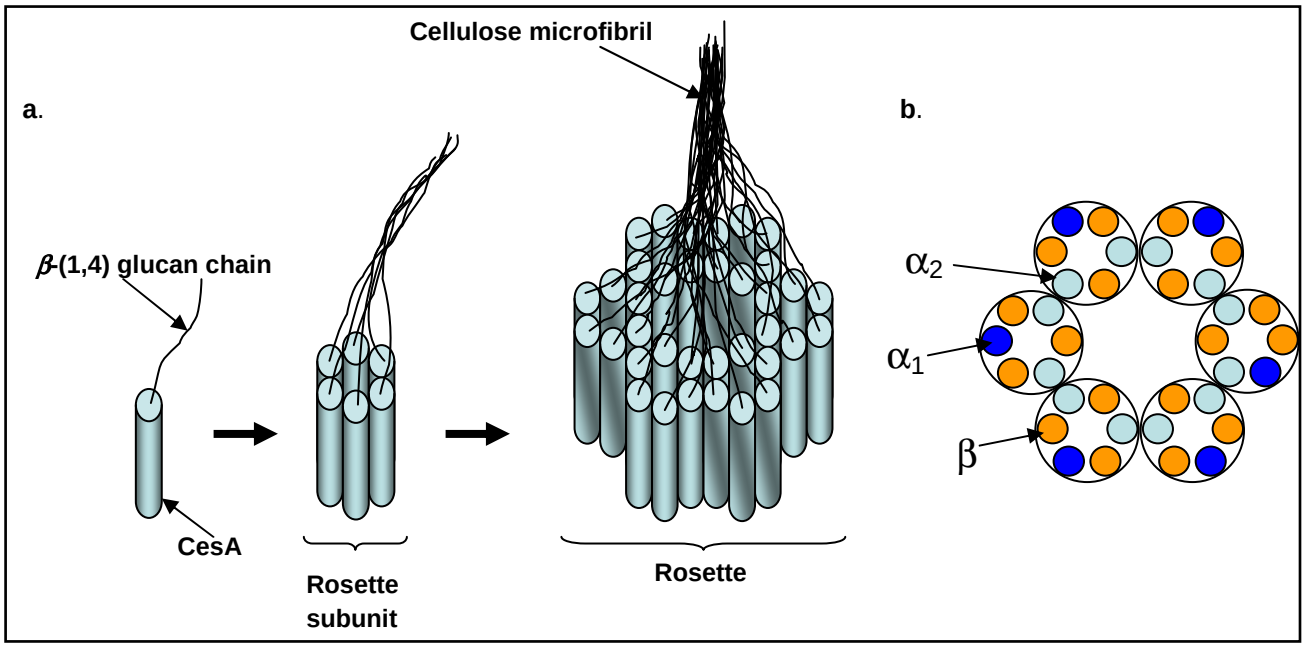


FIGURE 2.2: (a) A model for the structure of the rosette involved in cellulose biosynthesis. Six CesA polypeptides interact to form a rosette subunit. Each CesA polypeptide is shown to be involved in the synthesis of one β-(1,4)-glucan chain. Once the 36 chains emerge from the rosette, they combine to form an elementary cellulose microfibril. (b) Each rosette subunit consists of three different types CesA polypeptides (α_1, α_2 and β) and are assembled in such a way that α_2 from different subunits interact with each other (Scheible et al., 2001).

The first stage of in vivo crystalline cellulose biosyntesis, known as glucan sheet assembly, is catalyzed by the rosettes consisting of three different cellulose synthase dimers (Saxena & Brown, 2005). The CesA dimers utilize UDP-D-glucose and is responsible for the polymerization of individual β-(1,4)-glucan chains. The glucan chains formed from each rosette associate through Van der Waals interactions to produce a glucan chain sheet. Adjacent sheets assemble through interchain hydrogen-bonding to from the crystalline cellulose I microfibril while being simultaneously extruded from the cell through a pore-like structure.

Although no additional proteins are known to be directly involved in the crystallisation process, it has been proposed that the proteins involved in the organisation of the CesA dimers and the extrusion of the glucan sheets may contribute to this process (Arioli et al., 1998). Among the proteins identified that might have a potential role are annexin-like proteins which have UDP-D-glucose binding activity (Shin & Brown, 1999).

2.1.1.5. CELLULOSIC SUBSTRATES USED IN RESEARCH

There are a variety of cellulose substrates commercially available that are used in academic research and for industrial applications (Zhang et al., 2006). Table 2.2 shows the wide diversity in cellulose and cellulose-like substrates. The recalcitrance of these substrates towards hydrolysis is directly linked to their degree of similarity with naturally occurring cellulose. Broadly, these substrates can be distinguished with regards to their solubility in water.

A brief description of a selection of the cellulose substrates shown in Table 2.2 will follow with emphasis on cellulose substrates used in the current study. Carboxymethyl cellulose (CMC) is an ionic-substituted cellulose derivative that has both commercial and research application (Wood, 1988). It is prepared with the alkali-catalyzed reaction of insoluble crystalline cellulose with chloroacetic acid. CMC is water-soluble due to the polar carboxyl groups. There are two physical parameters that determine the quality of CMC used: (i) the degree of substitution i.e. the amount of carboxyl groups added onto the cellulose chain and (ii) the degree of polymerization. CMC is widely used in the food and non-food industries as a non-toxic viscosity modifier, food thickener, emulsion stabilizer and lubricant (Klein & Snodgrass, 1993). In the research field, CMC is commonly used to determine the hydrolytic activity of endoglucanases (Zhang et al., 2006). Endoglucanase activity can be measured in two ways using CMC: (i) the increase in reducing sugar formation or (ii) the reduction of its viscosity. Ionic strength, pH fluctuations and polyvalent cation concentrations can influence the viscosity of CMC, mainly due to its ionic nature (Wood & Bhat, 1988). Therefore, nonionic substituted cellulose substrates like hydroxyethyl cellulose (HEC) can be used as a replacement for viscosity assays. Both CMC and HEC can be mixed with certain dyes like Remazol Brilliant Blue R (Fülöp & Ponyi,

1997), Oztazin Brilliant Red H-3B (Biely et al., 1985) or Ruthenium Red (Rescigno et al., 1994) to form soluble dyed derivatives of CMC and HEC, respectively. These substrates can also be used to test endoglucanase activity as the colours released from these substrates can be quantitatively measured and are indicative of hydrolysis of the cellulose backbone (Zhang et al., 2006).

TABLE 2.2: Substrates that contain β-1,4-glucosidic bonds hydrolysed by cellulases. Adapted from Zhang et al., (2006).

Soluble	Insoluble
Short chain (low DP)	Crystalline cellulose
Cellodextrins	Cotton
Radio-labeled cellodextrins	Hydrocellulose (Avicel)
Cellodextrin derivatives	Bacterial cellulose
β-methylumbelliferyl-oligosaccharides	Whatman filter paper nr 1
p-nitrophenol-oligosaccharides	*Valonia* cellulose
Long chain cellulose derivatives	Amorphous cellulose
Carboxymethyl cellulose (CMC)	Phosphoric acid swollen cellulose (PASC)
Hydroxyethyl cellulose (HEC)	Alkali swollen regenerated amorphous cellulose
Dyed HEC, CMC	Dyed cellulose
	Fluorescent cellulose
	Chromogenic and fluorephoric derivatives
	Trinitrophenyl-carboxymethylcellulose (TNP-CMC)
	Fluram-cellulose
	Practical cellulose-containing derivatives
	α-cellulose
	Pretreated lignocellulosic biomass

p-Nitrophenyl glycosides such as *p*-nitrophenyl β-D-cellobioside (*p*NPC) and *p*-nitrophenyl β-D-glucopyranoside (*p*NPG) are chromogenic substrates derived from soluble cellodextrins (Zhang et al., 2006). The release of the *p*-nitrophenol moiety during a cellulase enzyme reaction causes a colour reaction which can be read spectrophotometrically at 405nm. These substrates are commonly used to study initial cellulase kinetics (Tuohy et al., 2002), reaction specificity (Zverlov et al., 2002), binding site thermodynamics (Barr & Holewinski, 2002) and the effect of inhibitors on cellulases (Tuohy et al., 2002).

Microcrystalline cellulose, also known as hydrocellulose or by its commercial name avicel is a popular insoluble crystalline substrate for determining the hydrolytic activity of some cellulases – especially exoglucanases (Zhang & Lynd 2004). It is prepared by hydrolyzing wood pulp with dilute hydrochloric acid to remove any non-crystalline cellulose followed by formation of colloidal dispersions using high shear fields and ending with spray drying of the washed pulp slurry (Fleming et al., 2001). Despite its preparation, microcrystalline cellulose still contains a significant fraction of amorphous cellulose, but can still be considered as a good substrate for exoglucanase activity assays due to its low DP and relatively low accessibility (Zhang et al., 2006). Even though generally considered pure, most avicel preparations will still contain hemicellulose or lignin components.

Bacterial cellulose (BC) is usually synthesized as microfibrils by the aerobic Gram-negative, acetic acid bacteria *Gluconacetobacter xylinus*, previously known as *Acetobacter xylinum* (Yamada et al., 1997). *G. xylinus* uses linear terminal enzyme complexes for the assembly of cellulose microfibrils (Yamanaka et al., 2000). BC is different from the cellulose derived from plants with regards to its high level of crystallinity and purity as it is free from lignin and other biogenic products (Keshk, 2002). It can be easily separated, has a high water-absorption and has improved mechanical strength.

Swollen, insoluble cellulose is usually prepared by the conversion of the crystalline fraction of cellulose to the amorphous with either mechanical or chemical methods (Zhang et al., 2006). Mechanically made amorphous cellulose is usually prepared by ball milling or severe blending (Wood, 1988). Alkali-swollen amorphous cellulose can be prepared by swelling cellulose powder in a high concentration of sodium hydroxide (e.g., 16% wt/wt) which results in the production of cellulose type II from type I (O'Sullivan, 1997). Phosphoric acid swollen cellulose (PASC) also known as Walseth cellulose is often made by swelling cellulose powder by adding 85% *o*-phosphoric acid (Walseth, 1952). It has been shown that the high concentrations of *o*-phosphoric acid result in some degree of conversion of type I cellulose to type II (Weimer et al., 1990). The characteristics of amorphous cellulose made by ball milling, sodium hydroxide

and *o*-phosphoric acid can vary significantly, depending on the origins of the crystalline cellulose powders, the reaction temperature and time and the reagent types and concentrations. As a result, it is almost impossible to compare hydrolysis rates of different kinds of cellulases on various types of amorphous cellulose from different batches of amorphous cellulose preparations (Zhang et al., 2006).

Other polymeric substrates commonly used to determine the activity of cellulases include barley-β-glucan, laminaran and lichenan. Barley-β-glucan is a linear homopolysaccharide made up of three or four glucose residues linked with β-1,4-glucosidic bonds separated with a single β-1,3-linkage (Bielecki & Galas, 1991). Laminarin, extracted from brown seaweeds, is a polysaccharide composed primarily of β-1,3-linked glucose residues with occasional intrastrand β-1,6-linked glucose linkages or branch points (McGrath & Wilson, 2006). Lichenan, extracted from Icelandic moss, is a linear, (1,3:1,4)-β-glucan with a structure similar to that of barley β-glucans (McCleary, 1988). The ratio of 1,4- to 1,3-β-linkages is approximately 2:1. The three abovementioned substrates, known as substrates with mixed linkages, are all decidedly more soluble in water than naturally occurring cellulose.

2.1.2. THE ENZYMATIC HYDROLYSIS OF CELLULOSE

2.1.2.1. INTRODUCTION

Glycosidases (*O*-glycoside hydrolases, EC 3.2.1.*x*) are an assorted collection of enzymatic proteins that have widespread application in the biochemical, medical and industrial fields (Henrissat et al., 1998). Glycosidases, in general, catalyse the hydrolysis of the glycosidic bonds in oligo- and polysaccharides. Cellulolytic enzymes, known commonly as cellulases or β-glucanases, represent the largest groups in the modern structural classification of glycoside hydrolases and are responsible for the hydrolysis of cellulose. A diverse group of organisms produce cellulases that work together to depolymerise cellulose (Rabinovich et al, 2002). The enzymes involved in cellulose hydrolysis have traditionally been classified in three major classes in accordance to their enzymatic activity: (i) endoglucanases or 1,4-β-D-glucan-4-glucanohydrolases (E.C 3.2.1.4); (ii) exoglucanases which are subdivided into cellodextrinases or

1,4-β-D-glucan-glucanohydrolases (E.C. 3.2.1.74) and cellobiohydrolases or 1,4-β-D-glucan cellobiohydrolases (E.C. 3.2.1.92); and (iii) β-glycosidase or β-glucoside glucohydrolases (E.C. 3.2.1.21).

2.1.2.2. THE CATALYTIC MECHANISM OF CELLULASES

The enzymatic hydrolysis of glycosidic bonds by cellulases proceeds via the general catalysis that involves two critical amino acid residues: a proton donor and a nucleophilic base (Sinnott, 1990). The hydrolysis occurs via two major mechanisms where the configuration of the anomeric carbon of the sugar ring is either retained or inverted and cellulases are often classified as being retaining or inverting enzymes (Davies & Henrissat, 1995). Figure 2.3 shows a schematic representation of the two mechanisms. The position of the proton donor in both retaining and inverting mechanisms are identical and is in close proximity to form a hydrogen bond with the glycosidic oxygen. In retaining enzymes, the nucleophilic base is also in close proximity of the sugar anomeric carbon. The distance between the two carboxylates i.e. the proton donor and the nucleophilic base is approximately 5.5 Å (McCarter & Withers, 1994). This nucleophilic base is, however, more distant in inverting enzymes, since a water molecule needs to be accommodated between the base and the sugar. Thus, the distance between the two carboxylates of inverting enzymes can be anything between 6.5-9.5 Å. Thus, the important difference between the retaining and inverting mechanisms is that retaining enzymes can catalyse both transglycosylation and hydrolysis due to the production of a covalent intermediate on the enzyme. No known inverting enzyme can catalyze transglycosylation (Blanchard & Withers, 2001).

2.1.2.3. ENDOGLUCANASES

Endoglucanases cleave intramolecular β-1,4-glucosidic linkages at random sites within amorphous regions of the cellulose chains (Teeri, 1997). During hydrolysis, endoglucanases generate cello-oligosaccharides of various lengths and, as a result, produce more reducing chain ends. The catalytic domains (CD) of all endoglucanases, whose three-dimensional structures have been determined, contain an open or exposed active site (Juy et al, 1992). This

feature gives endoglucanases a distinctive cleft-like, also known as groove-like, topology (Figure 2.4) and allows it to bind to the interior of long cellulose molecules. Activities of endoglucanases are often measured on a soluble high DP derivative of cellulose, like CMC (Zhang & Lynd, 2004). Endoglucanases decrease the specific viscosity of CMC at an appreciable rate with little hydrolysis taking place. Quantifying endoglucanase activities would thus involve measuring the reduction on viscosity of the cellulose substrate, but can also be assessed by measuring the increase in reducing chain ends generated. For a semi-quantitative approach endoglucanase activity can also be detected on agar plates by staining residual polysaccharides such as CMC with various dyes like Congo Red that can only absorb to long chains of polysaccharides (Fülöp & Ponyi, 1997).

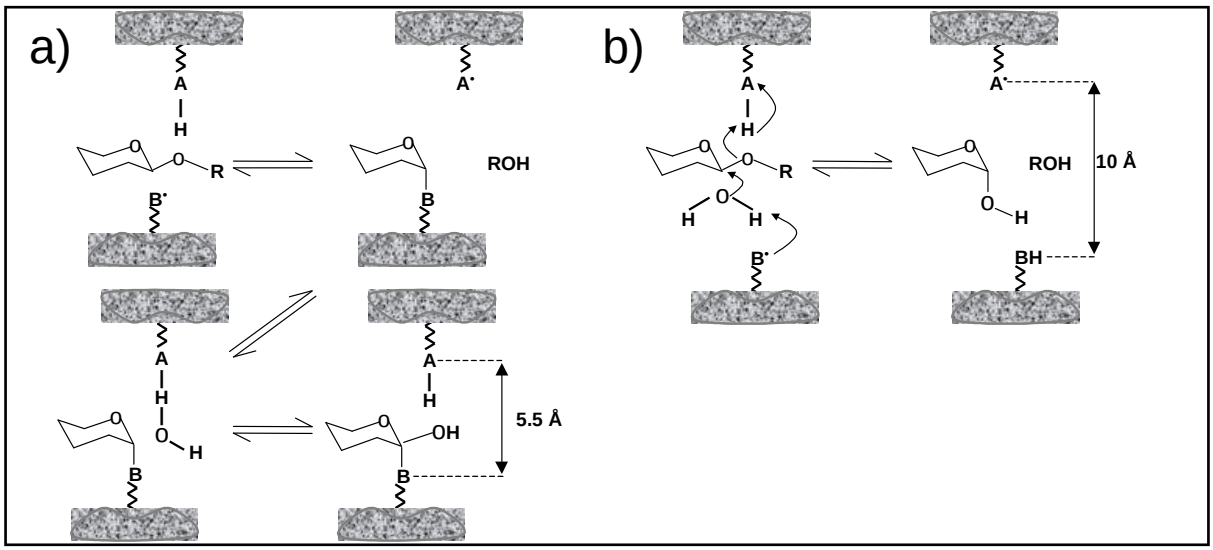


FIGURE 2.3: The retaining (a) and inverting (b) mechanism of enzymatic glycosidic bond hydrolysis (Davies & Henrissat, 1995). (a) In the retaining mechanism, the glycosidic oxygen is protonated by the acid catalyst (AH) or proton donor and nucleophilic assistance to the leaving group departure (ROH) is supplied by the base residue (B^-). This is the first step of this double displacement mechanism. A water molecule hydrolyses the resulting glycosyl enzyme and this second nucleophilic substitution at the anomeric carbon creates a product with a β-configuration matching the stereochemistry of the original substrate. (b) In the inverting mechanism, a single nucleophilic displacement leads to the inversion of the anomeric carbon. Protonation of the glycosidic oxygen and leaving group departure are accompanied by a simultaneous attack of a water molecule which is activated by the nucleophilic base (B^-). This mechanism yields a product with an α-configuration – thus of an opposite stereochemistry than the original substrate.

2.1.2.4. EXOGLUCANASES

Exoglucanases work in a processive manner on either the reducing or nonreducing ends of cellulose chains and liberate either glucose (glucanohydrolases) or cellobiose (cellobio-

hydrolase) as major products (Teeri, 1997). Occasionally, longer cello-oligosaccharides such as cellotriose and cellotetraose are also cleaved off (Divne et al., 1994; Sakon et al., 1997). Exoglucanases are also implicated to act on microcrystalline cellulose by presumably peeling off cellulose chains from the microcrystalline structure.

The CD of most exoglucanases are covered by long loops giving a tunnel-like topology (Figure 2.4) (Divne et al., 1998; Varrot et al., 1999). Three-dimensional studies have revealed that the loops can undergo large movements which leads to the opening and closing of the "roof" of the tunnel. Once the "roof" is in an open position, the polymeric substrate can become entrapped inside the tunnel and when threaded through either a cellobiose or glucose unit is hydrolysed at a time. Once hydrolysis is initiated at the reducing or non-reducing end of a cellulose chain, the enzyme moves along the length of the cellulose chain in a processive manner (Tomme et al., 1995).

Traditionally, cellulases with little activity on soluble CMC, but relatively high activity on avicel are classified as exoglucanases. Unlike endoglucanases and β-glucosidases, there are no cellulose substrate specific for measuring exoglucanase activity within the cellulase mixture (Sharrock, 1988). The activity of exoglucanases that attack the cellulose chain from the reducing end like *T. reesei* CBH I can be measured by 4-methylumbelliferyl-β-D-lactoside, *p*-nitrophenyl β-D-cellobioside or *p*-nitrophenyl β-D-lactoside, but these abovementioned substrates cannot be used for exoglucanases that attack from the non-reducing end e.g. *T. reesei* CBH II (van Tillbeurgh et al., 1982). Agar plate screening is also not efficient for measuring exoglucanases semi-quantitatively due to the blocking of the processive nature of exoglucanases by carboxymethyl substitutions of CMC (Demain et al., 2005).

2.1.2.5. β-D-GLUCOSIDASES

β-D-Glucosidases, sometimes called cellobiases, hydrolyze soluble cellobiose and other cellodextrins with a DP of up to six (Sternberg et al., 1977). There is a marked decrease in the rate of hydrolysis as the DP of the substrate increases (Zhang & Lynd, 2004).

The active site of most β-D-glucosidases as shown in Figure 2.4 are located in a pocket-like, sometimes called crater-like, region within the enzyme's topology (Jenkins et al., 1995). This type of topology is ideal for the recognition of non-reducing ends of a oligosaccharide chain. Most β-D-glucosidases hydrolyse substrates containing β-glycosidic linkages by cleaving off the terminal, non-reducing D-glucose residue. β-D-glucosidases are adapted to accommodate substrates with a large number of easy accessible, non-reducing chain ends at the surface. This excludes fibrous substrates such as native cellulose as it has almost no free chain ends.

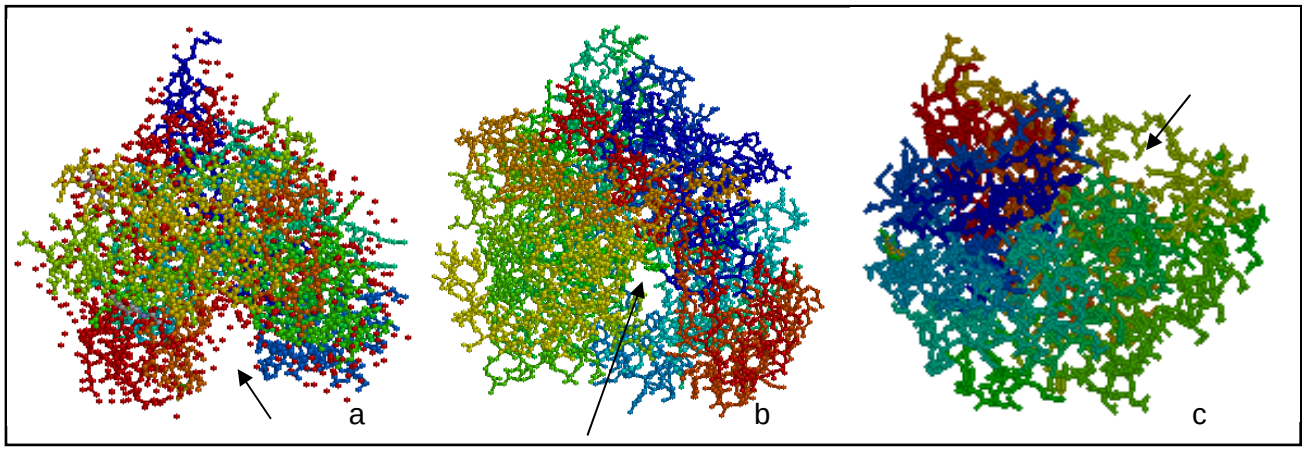


FIGURE 2.4: Three-dimensional structures generated by RasMol V2.7.3 (©Copyright Herbert J. Bernstein 1998-2005) of (a) the endoglucanase Cel6A of *Humicola insolens* Uniprot ref Q9C1S9, (b) the exoglucanase Cel48A of *Clostridium thermocellum* Uniprot ref P37698 and (c) the β-glucosidase and Cel1A of *Paenibacillus polymyxa* Uniprot ref P22037. The characteristic cleft-like topology of (a), the tunnel-shaped topology of (b) and the pocket-like topology of (c) are indicated with arrows.

Most β-D-glucosidases are also subjected to end product inhibition (Chen et al., 1992). However, glucose levels appear to adversely affect the activity of the β-D-glucosidases produced by *T. reesei* more so than other β-D-glucosidases (Decker et al., 2000).

Substrates used for measuring β-D-glucosidase activity are based on the release of coloured or fluorescent products like *p*-nitrophenyl β-D-1,4-glucopyranoside (Deshpande et al., 1983), but cellobiose can also be used since neither endo- or exoglucanases can hydrolyze it (Zhang & Lynd, 2004).

2.1.2.6. SYNERGY AMONG CELLULASES

A common feature among cellulases is their ability to synergize with each other, implying that the specific activity of a mixture of different cellulases is significantly higher than the sum of the individual cellulases in the mixture on the same substrate (Wood & McRae, 1979). Three major factors are known to affect the extent to which cellulases synergize (Jeoh et al., 2006): (i) the ratio and concentrations of the cellulase reaction mixture; (ii) the ease of access to binding sites for the cellulases in mixture and (iii) the physical and chemical heterogeneity of the cellulose substrate.

Four types of synergy between cellulases have been described: (i) endo-exo synergy which occurs between endoglucanases and exoglucanases; (ii) exo-exo synergy which occurs between two different exoglucanases that attack from different ends of the cellulose chain, either reducing or non-reducing; (iii) synergy between exoglucanases and β-glucosidases where β-glucosidases hydrolyse cellobiose and cello-oligosaccharides as end products of exoglucanases and thus removing feedback inhibition molecules and (iv) intramolecular synergy between the catalytic domains and the carbohydrate-binding modules (CBMs) (Lynd et al., 2002; Teeri, 1997; Din et al., 1994).

Synergy is usually most pronounced on crystalline cellulose, where the specific activity of some cellulase mixtures containing four to six different cellulases can be up to 15 times that of the sum of the individual activities of the cellulases in the mixture (Irwin et al., 1993). It appears that synergism only occurs when two cellulases attack different regions of the cellulose microfibril and each cellulase creates new sites for attachment and subsequent hydrolysis for the other enzymes in the mixture. The reason why most endoglucanases will show synergism with any exoglucanase, but not with each other, is thus due to the fact no new sites for attachments are created (Jeoh et al., 2006). Exoglucanases, however do show synergism with other exoglucanases, but only if one attacks from the reducing end and the other attacks from the non-reducing end of the cellulose chain. Processive endoglucanases show synergy with both endoglucanases and exoglucanases.

There is no support for the idea that synergism requires species-specific interactions between the synergising cellulases, since it is known that cellulases of organisms from different niches show similar synergy to those from the same organism (Wilson, 2008).

2.1.2.7. NON-HYDROLYTIC PROTEINS THAT CONTRIBUTE TOWARD CELLULOSE DEGRADATION

According to an early theory on the degradation of cellulose, Reese et al. (1950) proposed two components of activity that microorganisms utilize: (i) the C_x factor which comprises all the hydrolytic enzymes that convert cellulose to the end product of glucose and (ii) the non-enzymatic C_1 factor known as the swelling factor which, in general, makes the cellulose substrate more accessible for the hydrolytic enzymes. The C_1 components of microorganisms are not as clearly defined as the C_x factors, especially since they are not as easily quantified as hydrolytic enzymes.

Possibly the best described protein which has been classified as a swelling factor is swollenin (SWO1) of *T. reesei* (Saloheimo et al., 2002). Encoded by the *swoI* gene, swollenin is a protein that contains two domains: a N-terminal CBM similar to other CBMs found in *T. reesei* and a C-terminal plant expansin-like domain. Expansins are plant cell wall proteins that are known to disrupt the hydrogen bonds between cellulose chains and other polysaccharides, but without generating any significant reducing ends (Cosgrove et al., 2002). Through a series of experiments, SWO1 was shown to contribute toward the disintegration of the cellulose fibres in plant cell walls (Figure 2.5) without adding to the hydrolysis of the substrate (Saloheimo et al., 2002).

Along with C_1 factors, microorganisms produce many other non-hydrolytic enzymes that contribute towards the depolymerization of cellulose. These include cellobiose phosphorylases and cello-dextrin phosphorylases that perform the inorganic phosphate dependent phosphorolysis of β-glucosidic bonds (Reichenbecher et al., 1996). The end products of this type of enzyme catalysis are α-D-glucose-1-phosphate and equimolar amounts of either β-D-glucose for cellobiose phosphorylases or β-D-cellodextrin with reduced chain length for

cellodextrin phosphorylases. These enzymes are energetically advantageous over β-glucosidases due to the production of α-D-glucose-1-phosphate – an intermediate in the glycolysis pathway (Lou et al., 1996).

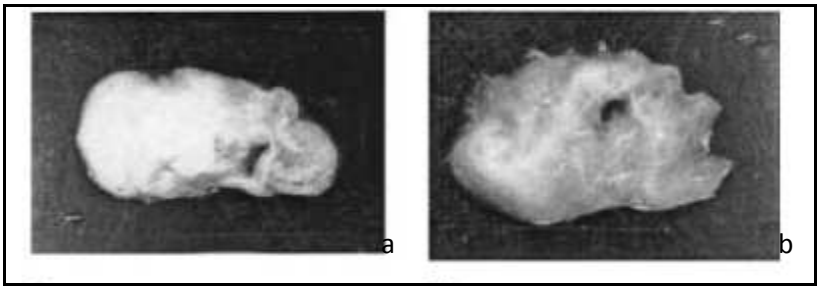


FIGURE 2.5: Structure of *Valonia* algae cells treated with buffer alone (a); and SWO1 in (b). Treatment with SWOI led to the decrease in the density of the cell wall (Saloheimo et al., 2002).

Cellulose dehydrogenases, also called cellobiose oxidases, can oxidize cellobiose to cellobiolactone under aerobic conditions (Henriksson et al., 2000). Cellobiolactone acts spontaneously with water to form cellobionic acid. Although the biological function of cellobiose dehydrogenases are not yet clear, its binding to microcrystalline cellulose and the enhancement of cellulose hydrolysis have been reported (Bao & Renganathan, 1992). It has been proposed that cellobiose dehydrogenases may generate hydroxyl radicals that could aid in the degradation of cellulose (Henriksson et al., 2000).

2.1.2.8. CARBOHYDRATE-BINDING MODULES

A carbohydrate-binding module (CBMs) is a contiguous amino acid sequence with a discreet folding pattern and has the function of binding carbohydrate-containing materials (Boraston et al., 1999). Most CBMs form part of carbohydrate active enzymes (Boraston et al., 2004) or form part of a cellulosomal scaffoldin (see Cellulase enzyme systems). Cases where CBMs are produced independently have also been reported (Moser et al., 2008). The sizes of CBMs range from small fungal binding domains of 36 amino acid residues to bacterial modules of over 200 residues (Lehtiö, 2002). They can be found either on the N-terminal or C-terminal ends in relation to the catalytic domain. The overall function of CBMs in carbohydrate active enzymes is to bring the catalytic domain in close proximity to the substrate thereby increasing the local substrate concentration for the enzyme active site (Boraston et al., 2004). CBMs are thought to

promote the hydrolysis in the catalytic domain by decreasing the dilution effect of the enzyme at the substrate surface by enhancing solubilization of individual glucan chains from the cellulose surface (Linder & Teeri, 1997). Glycoside hydrolases with no CBMs were shown to have a higher affinity toward crystalline cellulose when CBMs were appended (Ong et al., 1991). Alternatively, glycoside hydrolases with CBMs lose significant binding capacity to crystalline substrates when the CBM is removed (Bolam et al., 1998). It appears that the CBM has little to no effect on the activity of glycoside hydrolases towards soluble substrates (Lehtiö, 2001).

Certain CBMs also appear to have the ability to disrupt the crystalline nature of the polysaccharide structure – thus adding to the C_1 factors as discussed previously (Gao et al., 2001; Moser et al., 2008). In addition, this disruptive effect on the crystallinity improved the eventual degradative capacity of the catalytic modules.

2.1.2.9. CELLULASE ENZYME SYSTEMS

Cellulolytic microorganisms employ diverse strategies to degrade lignocellulosic materials, but can be broadly divided into two: (i) the non-complexed or free cellulase system (Warren, 1996) and (ii) the complexed or cellulosome system (Doi & Kosugi, 2004). A third strategy where cellulolytic bacteria are able to completely degrade cellulose without possessing any processive cellulases is currently under investigation (Wilson, 2008).

Most aerobic cellulolytic organisms utilize the free cellulase system which includes the soft-rot fungus *T. reesei* (Claeyssens et al., 1990), the white-rot fungus *Phanerochaete chrysosporium* (Broda et al., 1994) and the coryneform bacteria *Cellulomonas fimi* (Chaudhary et al., 1997) – all considered to be model organisms in cellulose degradation. These microorganisms produce several soluble cellulases along with associated polysaccharide depolymerases like xylanases that are secreted extracellularly. However, some of these enzymes, particularly β-glucosidases, are not completely secreted and will remain associated with the cell wall (Usami et al., 1990).

Cellulosomes are protuberances found on the cell wall of certain obligate anaerobic microorganisms when grown on cellulose-containing materials (Doi, 2008). Cellulosomes are stable enzyme complexes that are tightly bound to the cell wall, but flexible so as to also bind to microcrystalline cellulose. The cellulosome consist of two key components: (i) scaffolding proteins and (ii) several bound enzymes (Bayer et al., 1985). The scaffolding protein is a large non-enzymatic protein that contains several enzyme binding sites called cohesions and at least one carbohydrate binding module (CBM). It also helps with the attachment of the cellulosome to the microorganism through cell-surface-associated proteins (Doi & Kosugi, 2004). The bound cellulosomal enzymes are all involved in the depolymerization of the plant material substrate and contain so-called dockerin domains which interact with the scaffoldin's cohesion domains.

Advantages of producing and maintaining cellulosomes include keeping concerted enzyme activity in a close proximity to the bacterial cell which enables optimum synergy between the cellulases and other glycoside hydrolases (Schwarz, 2001). The cellulosome also lessens the distance over which cellulose hydrolysis products must diffuse, which allows for more effective uptake of these oligosaccharides by the host cell. Microorganisms that produce cellulosomes like most *Clostridia* spp. that utilizes cellulose (Doi, 2008), can also produce non-cellulosomal cellulases that are secreted by the organism similar to the non-complexed systems (Berger et al., 2007).

The analyses of genomic sequences of the aerobic bacterium *Cytophaga hutchinsonii* (Xie et al., 2007) and the obligate anaerobe *Fibrobacter succinogenes* (Jun et al., 2007) have shown that these organisms contain no genes that encode for processive cellulases. This is an unusual phenomenon since it is widely considered that all true cellulolytic microorganisms contain at least one processive cellulase. Both *C. hutchinsonii* and *F. succinogenes* contain genes that putatively only encode for endoglucanases – most of them without any carbohydrate binding modules (Wilson, 2008). Neither contains any scaffoldin gene or any genes that contain dockerin domains. Currently, it appears that these two organisms possess a yet unknown strategy to degrade cellulose.

2.1.2.10. CLASSIFICATION OF CELLULASES

Before the current nomenclature system, cellulases were given relatively random names as they were discovered and were given appropriate designations as prescribed by the IUB-MB enzyme nomenclature (Henrissat et al., 1998). This particular enzyme nomenclature is based on the enzyme's substrate specificity, but rarely on its molecular mechanism and never on its structural features. Initially, this nomenclature was sufficient since only a few enzymes had been characterized and comparisons between different glycoside hydrolase systems were not compromised. Later-on, in the mid-1990s, it became clear that a new classification system was required since enzyme systems derived from widely diverse organisms shared similarities that were not reflected from their enzymes' names (Warren, 1996).

In early 1998, Henrissat et al. proposed a new scheme for designating glycoside hydrolases which was mainly based on amino acid sequence similarities. It was found that there was a direct relationship between the amino acid sequence and the folding pattern of the enzyme. Advantages of such a classification system are threefold: (i) the structural features of enzymes are a better representation of the enzyme than substrate specificity since many enzymes exhibit specificity toward more than one type of substrate; (ii) the evolutionary relationships between these enzymes can be revealed; and (iii) the mechanistic behaviour of these enzymes can be easily obtained from their protein sequence. In this designation, enzymes were grouped into families. At the time of writing this literature review, there were 115 (as of 25/01/2010) different glycoside hydrolase families described (http://www.cazy.org/fam/acc_GH.html). Cellulases (endo- and exoglucanases) can be found in families 5, 6, 7, 8, 9, 12, 44, 45, 48 and 60 – a total of 10 families. Several other families (10, 26, 51, 74) also contain cellulase enzymes, but the majority of the enzymes that belong to these families are non-cellulolytic. Indirectly, a family also indicates the mechanism of the particular enzyme for e.g. all the enzymes in family 5 are retaining and all the enzymes in family 6 are inverting.

This designation prescribes that all cellulose enzymes have to start with Cel, with the gene name being *cel*, followed by the family number in which the particular enzyme is classified e.g.

the endoglucanase II (previously EGII) from *T. reesei* which is in glycoside hydrolase family 5 would be called Cel5A according to this scheme. Examples of how this designation system changed the naming of the cellulases of *Thermobifida fusca* are listed in Table 2.3. If more than one enzyme from the same organism belong to the same family, for instance in the case of *T. reesei* CBH1 and EG1 – which both belong to glycoside hydrolase family 7 – the enzyme that was described first will be called Cel7A followed by Cel7B. To distinguish between two similar enzymes from different organisms like Cel48A of *Clostridium thermocellum* and Cel48A of *T. fusca*, the first letters of both the genus and species name of the organism must be used (in this case Ct and Tf, respectively).

TABLE 2.3. Previous and current designation of all the endo- and exoglucanases produced by the actinomycete *Thermobifida fusca*.

Enzyme	Previous designation	Current designation
Endoglucanase	E1	TfCel9B
Endoglucanase	E2	TfCel6A
Exoglucanase	E3	TfCel6B
Processive endoglucanase	E4	TfCel9A
Endoglucanase	E5	TfCel5A
Exoglucanase	E6	TfCel48A
Endoglucanase	None	TfCel5B

This scheme is mostly concerned with the folding structure of the catalytic domain of the cellulase enzymes. Most cellulases are also modular and contain other ancillary domains, usually one carbohydrate-binding module CBM. CBMs have also been grouped into separate families of related amino acid sequences (Tomme et al., 1995).

Information on glycoside hydrolases, from putative to fully characterised, is maintained and regularly updated on the Carbohydrate-Active Enzyme or CAZy website (http://afmb.cnrs-mrs.fr/CAZY/) (Coutinho & Henrissat, 1999). Similarly, the carbohydrate binding modules (as discussed above) are also classified in 59 (as of 25/01/2010) different families based on amino acid sequence, binding specificity, and structure. More information and classification can be found in the Carbohydrate-Binding Module Family Server (http://afmb.cnrs-mrs.fr/pedro/CAZY/cbm.html).

2.1.2.11. CELLULASES USED IN THIS STUDY

2.1.2.11.1. CEL9A OF *THERMOBIFIDA FUSCA*

The moderately thermophilic actinomycete *Thermobifida fusca* (previously known as *Thermomonospora fusca*) is a well characterized cellulolytic organism commonly found in compost heaps and rotting hay (Wilson, 1988). *T. fusca* secretes at least seven different cellulose-degrading enzymes. These enzymes include four endoglucanases (Cel9B, Cel6A, Cel5A and Cel5B), two exoglucanases (Cel6B which attacks from the non-reducing end and Cel48A which attacks from the reducing end of the glucan chain) and one unusual enzyme known as a processive endoglucanase, Cel9A (previously known as E4).

From a structural basis, Cel9A is a modular enzyme consisting of four domains: a family 9 catalytic domain (CD) adjacent to a family IIIc carbohydrate-binding module (CBM), a fibronectin III-like (FN III) linker domain and a family II carbohydrate-binding module (Irwin et al., 1998). The CD is homologous to the CD of the family 9 enzyme Cel9W of the anaerobic bacteria *Clostridium thermocellum*. The family IIIc CBM is also homologous to a CBM of the scaffoldin protein of *C. thermocellum.* A study on several mutants of Cel9A determined the exact amino acid residues which are important for the catalytic mechanism of Cel9A (Zhou et al., 2004). Glu424 was identified as the active catalytic acid whereas both Asp55 and Asp58 act as a catalytic base to deprotonate the nucleophilic water. Tyr206 was found to take part in binding and catalysis and Tyr318 was found to be involved in the specific binding of crystalline cellulose substrates. Both tyrosine residues in the catalytic cleft were required for the processive action of Cel9A. The amino acid residues Asp55, Asp58 and Glu424 were found to be conserved among all family 9 glycoside hydrolases (Jung et al., 1993).

Cel9A was, at first, classified as an endoglucanase due to its low, but detectable activity on carboxymethyl cellulose (Irwin et al., 1993). However, further research showed that during cellulose degradation, it mainly produces soluble oligosaccharides (87%) from insoluble cellulose – which is a general trait of exoglucanases. Thus, Cel9A displayed both endoglucanase and exoglucanase activity. This finding prompted further investigation into the structure and

mechanism of Cel9A (Sakon et al., 1997). It directly led to the uncovering of the crystal structure of Cel9A-68 – a 68kDa fragment of the intact Cel9A enzyme where the family II cellulose binding domain was removed. This particular study provided evidence that revealed the mechanistic manner with which Cel9A can be both an endoglucanase and an exoglucanase. It also indicated that when Cel9A acts as an exoglucanase, the main product of liberation is cellotetraose. This study further presented the first evidence of a cellulase enzyme where the catalytic domain and the cellulose binding domain had direct interaction with each other during cellulose hydrolysis. The CD and the family IIIc CBM interact in two loop regions in such a way that the shallow cleft of the CD is aligned with the flat face of the CBM, enabling the binding of both domains along the glucan chain.

The current thinking of how Cel9A operates concerning the hydrolysis of cellulose starts with an initial cleavage being made along the glucan chain (Irwin et al., 1998). From then on Cel9A hydrolyzes the glucan chain in a processive manner, from the non-reducing end with the aid of the family IIIc CBM, which appears to be responsible for the loose binding of the enzyme to crystalline cellulose (Li et al., 2007).

Apart from the endo/exoglucanase activity of Cel9A, it also exhibits good synergism in cellulase mixtures (Irwin et al., 1993). Cel9A has been shown to synergize with Cel6B (an exoglucanase) and Cel5A (an endoglucanase) of *T. fusca* in binary mixtures with a synergistic effect i.e. the activity in binary mixture divided by the sum of individual activities of 1.9 and 1.7 respectively on Avicel. It also synergizes well in binary mixtures with both types of cellobiohydrolases from *T. reesei,* namely Cel6A and Cel7A. It has been proposed that Cel9A is able to exhibit synergy with classical endoglucanases due to its processive activity (i.e. act as an exoglucanase), while it is able to synergize with exoglucanases since it can bind and cleave internal sites within the glucan chain (i.e. act as an endoglucanase) (Irwin et al., 1998).

The DNA sequence of the *cel9A* gene was determined in 1993 (Jung et al., 1993) and was successfully cloned and expressed in *Escherichia coli* and *Streptomyces lividans.*

2.1.2.11.2. CEL6A OF *NEOCALLIMASTIX PATRICIARUM*

The anaerobic chytrid *Neocallimastix patriciarum* – commonly found in the rumen of sheep and cattle – is known for its diverse hydrolase activity towards plant polysaccharides, its capacity for generating fermentable sugars from cellulose with relative efficiency, and its capability to grow on cellulose as its only carbohydrate source (Williams & Orpin, 1987).

The *cel6A* (previously named *celA*) gene of *N. patriciarum* was isolated from a cDNA library (Denman et al., 1996). It has an open reading frame of 1,284 base pairs and encodes for a polypeptide of 428 amino acid residues. It has a characteristic secretion signal peptide consisting of 20 amino acid residues. Cel6A is a modular enzyme in that it contains both a catalytic domain (CD) and a family 1 carbohydrate-binding module (CBM1). The catalytic domain, located on the C-terminus of the mature enzyme, is homologous to Cel6A (or CBHII) of *T. reesei* with the level of similarity being 53% and the level of identity in the catalytic domain region, 37%. Like the CD, the CBM shows a high degree of homology with the CBM of the *T. reesei* Cel6A of (level of identity, 57%; level of similarity, 77%). An unusual asparagine-rich linker is found between the CBM and the CD of which 33 of the 44 residues are asparagine.

The specific activity of the immunoaffinity-purified Cel6A against a crystalline type of cellulose (Avicel) was determined to be 9.7 U/mg of protein (Denman et al., 1996). This is the highest specific activity for an individual cellulase against Avicel reported in literature (Lynd et al., 2002). Cel6A also exhibited activity against amorphous cellulose, carboxymethylcellulose and lichenan, but not against the substrates xylan and *p*-nitrophenyl cellobioside.

The *cel6A* gene has been cloned and expressed successfully in *E. coli* (Denman et al., 1996), *Clostridium beijerinckii* (Lopez-Contreras et al., 2001) and the Gram-positive rumen bacterium *Streptococcus bovis* (Ekinci et al., 2002). A bacterial dockerin module was fused onto the C-terminal end of *cel6A*, for subsequent incorporation into chimeric minicellulosomes in combination with various cellulosomal cellulases from *Clostridium cellulolyticum* (Mingardon et al., 2007). In a separate study, the CBM was successfully fused to the *Candida antarctica* lipase

B gene and expressed in *Pichia pastoris* (Rotticci-Mulder et al., 2001). This was done to purify the lipase B protein by immobilizing it on Avicel.

2.2. ETHANOL PRODUCTION FROM LIGNOCELLULOSIC BIOMASS

2.2.1. INTRODUCTION

Civilization is largely dependent on energy utilization (Wyman et al., 1996). Currently, the principle source of energy is derived from the combustion of fossil fuels. Fossil fuels give rise to electricity and provide millions of people a daily means of transportation. Ever since the dawn of the Industrial Revolution in the late 1800s, the use of fossil fuels as an energy source has created many problems, to such an extent that some of them have to be addressed immediately. An obvious way of addressing these problems is to look for renewable alternatives.

The combustion of fossil fuels and the search for alternatives have long been a contentions subject among policy-makers world-wide (Farrell et al., 2006). The general consensus for the implementation of viable alternatives is that it should provide what fossil fuels cannot. These include several socio-political and environmental aspects of which job creation, energy security and a favourable carbon footprint are among the most important.

There are a multitude of alternatives when it comes to the generation of electricity and many of them have already replaced a significant section previously held by the combustion of fossil fuels (Whittington, 2002). These include nuclear (Pearsson & Pena-Tores, 2000), wave or tidal (Gross et al., 2003), solar (Şen, 2005), wind (Joselin Herbert et al., 2007) hydro-electric (Bakis, 2007), geothermal (Fridleifsson et al., 2001) and the combustion of biologically-derived wastes (Montgomery, 2004). However, none of the abovementioned alternatives can currently be used in the transportation sector in any substantial manner – even though solar and hydrogen power holds promise for the future (Şen, 2005).

Even though ethanol was used as a transportation fuel in the earliest forms of cars including the Model T Ford, the current interest in its use as a viable alternative to fossil fuels has its roots in a major political incident in the early 1970s (Kerr, 1998). The oil embargo in 1973 enforced by most of the oil-producing countries on the USA and the majority of the developed world had a dramatic effect on the cost of living, reshaped the geopolitical landscape and sparked interest in the use of ethanol and many other alternative energies. In the mid-1970s, Brazil became the first country to launch a program, known as Proálcool, to increase the amount of transport fuel produced locally. This directly led to Brazil's independency from foreign-produced oil. This program was based on the ethanol production derived from molasses and sugar cane juice. Mainly due to the Proálcool initiative, Brazil is still one of the major players in ethanol production and provides an excellent model for other countries as they are moving toward similar enterprises (Tan et al., 2008).

2.2.2. ETHANOL AS TRANSPORTATION FUEL

Ethanol is the common name given to ethyl alcohol, but has several synonyms found in literature based on the composition, the feedstock from which it was derived and the end uses (Lyons, 2003). It is a clear, colourless, flammable oxygenated hydrocarbon with the chemical formula of C_2H_5OH. About 95 % of all ethanol output is from the fermentation and distillation of agricultural feedstocks, whereas the rest is of synthetic origin derived from crude oil or gas and coal.

Ethanol for use in vehicle engines has long been recognized as a high-quality transportation fuel despite being rather costly (Bailey, 1996). It outperforms petroleum on the key criteria of octane rating, energy efficiency and emission performance. Ethanol can be used in the low-level petroleum car fleet with spark-ignition engines in a E10 blend – meaning 10 % of the transportation fuel is ethanol – without any modification needed for the car engine. However, dedicated or flexible-fuel engines are needed for blends with higher ethanol content, like E85, E95 or even E100. Ethanol cannot blend with diesel fuel, but can fully replace it in compressed-ignition engines. Ethanol has a lower energy density than conventional petroleum and even

though it does not affect the engine performance, it does influence the volume and weight of the fuel tank (Sinor & Bailey, 1993). Difficulties also arise regarding the transportation and storage of ethanol and ethanol blends mainly due to the presence of water (Bailey, 1996). Ethanol is miscible in water which would result in the ethanol to separate into the water phase, disturbing the blend composition. Thus underground storage tanks and distribution pipes would have to be secured free of water.

2.2.3. ENVIRONMENTAL AND SOCIO-POLITICAL IMPACTS OF ETHANOL AS TRANSPORTATION FUEL

Many industrial countries, excluding the USA, signed the Kyoto Protocol in 1997 and are committed to reduce their carbon footprint by using less fossil fuels (UNFCCC, 1998). By including ethanol in their transportation fuel industry, it is believed that these countries will reach their intended targets successfully.

In general, ethanol and ethanol blends used in vehicles cause a cleaner combustion mainly due to ethanol's oxygen content (Morris, 2000). This results in a significant reduction in emissions that are seen as harmful to the environment. The most striking reduction is that of carbon monoxide in E10 blends of 20 %. Ethanol dilutes most other toxic elements in petroleum such as benzene, toluene and xylene. A concern is that with E85 blends, the emissions of five major air pollutants i.e. carbon monoxide, volatile organic compounds, particulate matter with aerodynamic diameter ≤ 10 µm, oxides of sulphur (SO_x) and oxides of nitrogen (NO_x) are higher than with petroleum (Brinkman et al., 2005). On balance, it appears that combustion of ethanol and ethanol blends are significantly less toxic than pure petroleum.

As shown in Figure 2.6, the combustion of any fossil fuel reintroduces carbon – mostly in the form of carbon dioxide – into the Earth's atmosphere that was trapped in sediments and removed from the carbon cycle millions of years ago (Kerr, 2007). Even though through the process of photosynthesis, plants and cyanobacteria can reincorporate the carbon back in the cycle, rapid deforestation and the increase in fossil fuel utilization resulted in a dramatic increase of carbon dioxide levels in the atmosphere. This is the leading cause of global warming

– a serious concern to modern-day humankind. Ethanol made from plant biomass can, in theory, leave a much more favourable carbon footprint as the plant biomass can be regenerated. This statement is, however, subject to agricultural practices and waste management that put renewability and sustainability as its highest priority. The petro-chemical industry is also responsible for considerable adverse effects on land and water resources with oil spills (Wyman, 1996). Spillage of ethanol is not considered a major problem as it is biodegradable.

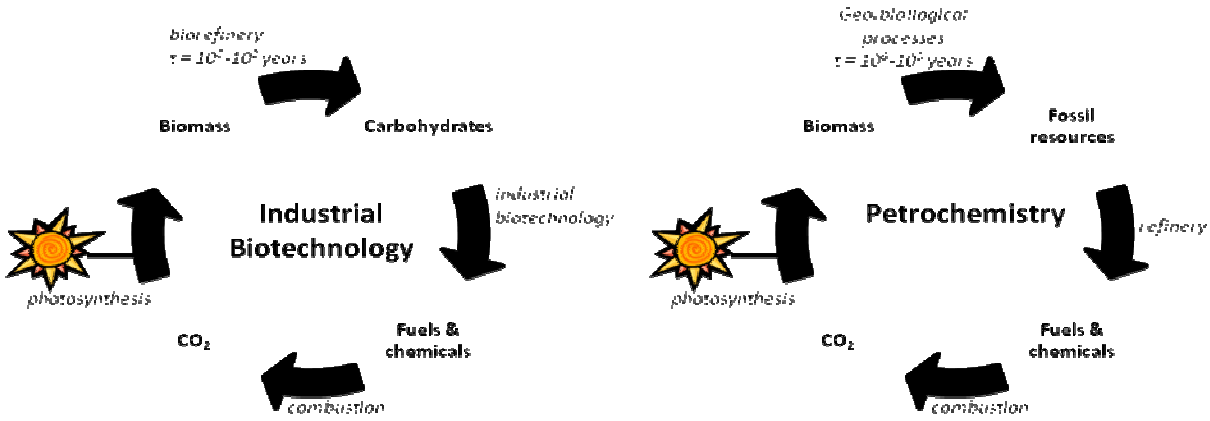


FIGURE 2.6. Simplified diagram of the carbon cycle between ethanol and petroleum. Note the time difference for biomass to be converted to carbohydrates and fossil resources for each process.

From an economic and political standpoint, many countries, most notably the USA, have already made significant strides to implement ethanol and biofuels as a whole to replace fossil fuels for energy security reasons (Kerr, 1998). It is estimated that about two-thirds of oil reserves are situated in the Middle East region which holds a majority in the Organization of Petroleum-Exporting Countries (OPEC). OPEC has a key influence on the determination of the prevailing price of oil per barrel. For example, in 2008 the world has seen the oil price fluctuate from above $140 to below $40 per barrel. Oil prices have often been set on reasons that go far beyond supply and demand such as political unrest and speculation. The authority that OPEC has on the expenditure of countries with a net import of oil can be diminished by the local production of biofuels and thus aid in political security world-wide.

As many of the current ethanol initiatives require vast amounts of agricultural land, it is seen as an opportunity for the social upliftment of rural communities world-wide (Mangoyana, 2009). Ethanol production will create new jobs of which many would have to be near the site of feedstock production.

2.2.4. THE PRODUCTION OF ETHANOL

Ethanol production for the biofuel industry is categorized in first- and second-generations based on the type of substrate used (Tan et al., 2008). Ethanol made from agriculturally important crops are known as first-generation ethanol of which sugar cane and maize (corn) are the most widely used. Sugary and starchy feedstocks provide a relatively easy upstream processing towards generating fermentable sugars with mature technologies existing for both types of feedstocks. However, these feedstocks are important staple foods and along with palm oil and soybeans, which are used for biodiesel production, led to the "food for fuel" debate. With the escalating demands of these agricultural feedstocks to cater for the production of biofuels around the world, a fear exists that it will result in the dramatic rise of the price of essential foods. Along with the pressure the biofuels industry puts on arable land, it is also heavily linked with the deforestation of tropical regions in South America, Africa and South-East Asia as more plantation areas are required for the growing demand.

Second-generation ethanol is derived from lignocellulosic biomass and can include waste or by-products generated from the agricultural, municipal, industrial, food and forestry sectors and dedicated energy crops like grasses and short rotation crops and is not in direct competition with food production (Gomez et al., 2008). Second-generation ethanol is sometimes called cellulosic ethanol since cellulose is the biggest component of lignocellulose. The large-scale production of second-generation ethanol can alleviate the stress put on arable land and would also result in a more positive carbon footprint as opposed to the production of first-generation ethanol (Hill et al., 2006). The cost in feedstock for second-generation ethanol is significantly lower compared to agricultural crops since the input cost for energy crops is much less.

The biggest difference between first- and second-generation ethanol is that the latter requires an additional step of extensive pretreatment using a combination of physical, chemical and enzymatic methods before enough fermentable sugars can be released for ethanol production. This extra processing step dramatically increases the cost involved in the process.

Acid treatment, steam explosion, liquid hot water treatment, ammonia fibre explosion and liming are some of the most widely used physical and chemical pretreatments - each with their own mode of action (Mosier et al., 2005). Due to lignocellulosic substrates' heterogeneity, different pretreatments are usually combined. Some of the major functions of physical and chemical pretreatment procedures include the separation of the lignin matrix from the carbohydrates, disruption of the crystalline nature of lignocellulosic material and increasing the exposure of the fibres for better enzyme hydrolysis (Prasad et al., 2007). Important prerequisites for pretreatment procedures are the prevention of carbohydrate degradation, avoidance of the formation of inhibitory compounds and cost-effectiveness. It is important to note that different pretreatment protocols are required to release optimal amounts of fermentable sugars from different feedstocks. It is therefore necessary to optimize every biomass to ethanol process with regard to its pretreatment to maximize the amount of product.

After the initial breakdown of the lignocellulose structure, enzymes are added that are able to depolymerize the complex carbohydrates into fermentable sugars (Mielenz, 2001). Depending on the specific lignocellulosic biomass, different enzyme combinations – or cocktails – are used and would include cellulases, hemicellulases, ligninases, amylases and pectinases. The enzyme cocktails are usually obtained from large-scale cultivations of filamentous fungi such as *Trichoderma* and *Aspergillus* spp., but bacterial and fungal enzyme mixtures proved to be similarly effective (Baker et al., 1995). Despite recent significant cost-reduction reported by enzyme companies Novozyme and Genencor (Stephanopolous, 2007), enzyme production, especially that of cellulases, remains a major cost consideration.

2.3. CONSOLIDATED BIOPROCESSING

2.3.1. INTRODUCTION

The steps used to process commercial-scale biomass, where enzymes and/or microbial hydrolysis are involved, are usually divided into four distinct biologically facilitated transformations: (i) the production of saccharolytic (polysaccharide-degrading) enzymes including cellulases and hemicellulases; (ii) the hydrolysis of carbohydrate components present in the pretreated biomass to easily fermentable sugars; (iii) the fermentation of six-carbon sugars (glucose, mannose and galactose) and (iv) the fermentation of five-carbon sugars (xylose, arabinose) (Lynd et al., 2005). Figure 2.7 shows process configurations where some of the abovementioned transformations have been combined. Figure 2.7 also shows the process configuration called consolidated bioprocessing (CBP), first coined by Lee Lynd (Lynd, 1996) as a variation to direct microbial hydrolysis (Wang et al., 1983) that would combine all four transformations into one step. During a CBP process, an organism or a consortium of organisms would thus be able to produce a sufficient amount of saccharolytic enzymes to hydrolyze the lignocellulosic substrate to soluble, short-chained oligosaccharides i.e. fermentable sugars and subsequently ferment the sugars to ethanol. What makes CBP different from the other process configurations is that it has no dedicated cellulase/hemicellulase production step. Using an ATP-centered model for anaerobic cellulose fermentation, Van Walsum and Lynd (1998) showed the feasibility of CBP from a kinetic and bioenergetic perspective when compared to the processing strategy known as simultaneous saccharification and fermentation (SSF).

As shown in Table 2.4, CBP can potentially decrease the cost and increase the efficacy of other process models especially in cases where dedicated cellulase production is necessary (Lynd et al., 2005). These include the avoidance of cost for capital, resources and dedicated facility space required for cellulase production. If higher hydrolysis rates of plant biomass can be realized, CBP can potentially also reduce the reactor volume needed for fermentation and could also decrease the capital cost involved. Higher hydrolysis rates might be achieved if slightly thermophilic organisms are used. The so-called enzyme-microbe synergy can be increased if

more complexed cellulase systems are applied and if organisms that can adhere to cellulose are used. Especially the enzyme-microbe synergy proves to be a major determining factor as it was shown that the anaerobic bacterium *C. thermocellum* can surpass the specific cellulase rates exhibited by the soft-rot fungus *T. reesei* by 20-fold and a substantial part of the dramatic increase was attributed to the cellulases' involvement in the cellulose-enzyme-microbe complexes (Lu et al., 2006).

Even though the concept of CBP can be applied to most processes where a cheap, renewable substrate needs to be converted to a commodity bioproduct, most emphasis concerning CBP has been put on the production of ethanol from plant biomass (Lynd, 1997).

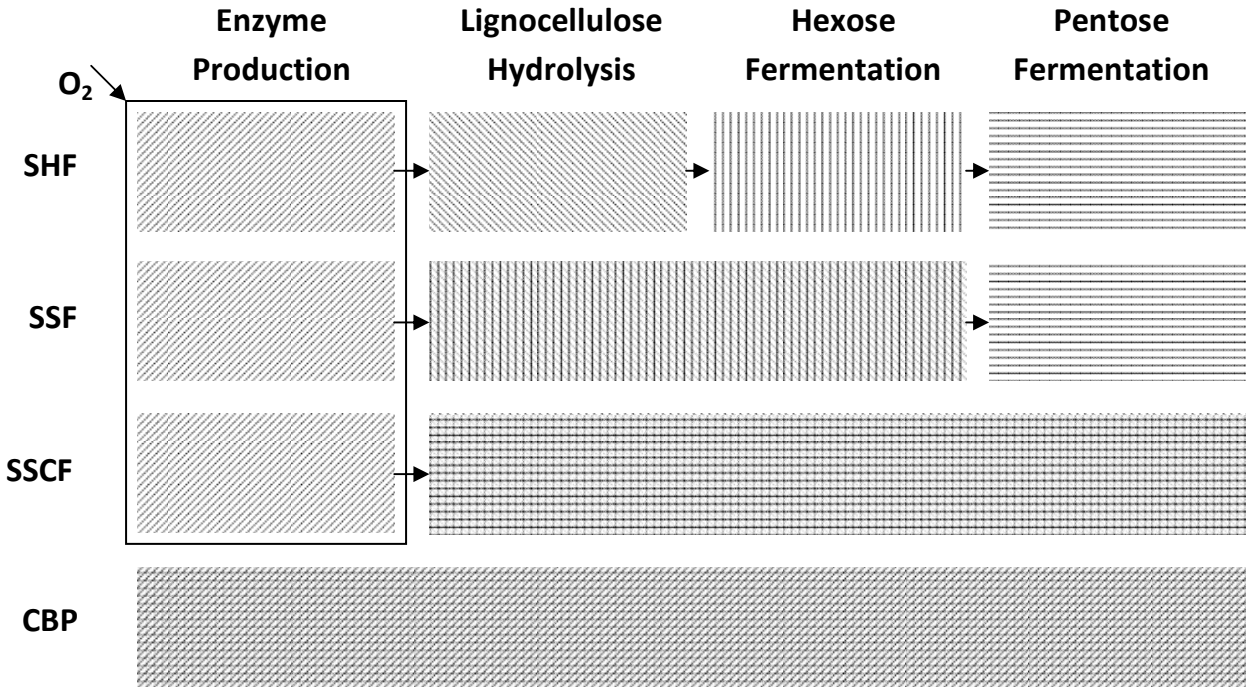


FIGURE 2.7. Different process strategies for the biologically facilitated conversion of pretreated lignocellulosic biomass (Lynd et al., 2002). Each rectangular box represents a separate bioreactor – not to scale. Separate hydrolysis and fermentation (SHF) has four distinct process steps and can have as many as four different biocatalysts. With simultaneous saccharification and fermentation (SSF) the hydrolysis and hexose fermentation steps were combined while in simultaneous saccharification and cofermentation (SSCF) the hydrolysis and the fermentation of both the hexoses and pentoses were combined. Consolidated bioprocessing (CBP) shows a process configuration where all abovementioned process steps are combined.

2.3.2. DEVELOPMENT OF A CBP ORGANISM

There is currently no organism that can sufficiently meet all the requirements of what a CBP-organism ideally should possess, but many have strong characteristics that make them potential candidates (Van Zyl et al., 2007). These candidate organisms can broadly be divided into two groups: (i) the native cellulolytic and (ii) the recombinant cellulolytic. Addressing the shortcomings of these organisms is the objective of many research groups around the world.

TABLE 2.4. Projections of the cost in US cent involved of ethanol production (in US$) by consolidated bioprocessing (CBP) versus simultaneous saccharification and co-fermentation (SSCF) with the inclusion of dedicated cellulase production. Adapted from Lynd et al., 2005.

	Cellulase Production (CP)	SSCF	Total (CP + SSCF)	CBP
Lost Yield	5.59	-	5.59	-
Utilities	1.63	2.27	3.90	1.71
Raw Materials	0.83	1.02	1.85	0.68
Capital and Related	1.80	5.69	7.49	1.84
			18.83	4.23

In the native cellulolytic strategy, one would begin with organisms that can naturally utilize cellulose and other biomass polysaccharides (Baskaran et al., 1995; Lynd et al., 2001). Many of these organisms are thermophilic like the Gram-positive *Thermoanaerobacterium* spp. and some *Clostridium* spp. Among the major challenges for the development of such organisms include channelling energy fluxes toward ethanol as end product. Many of these organisms are known to have a low ethanol yield during the fermentation of both hexoses and pentoses. With the aid of metabolic engineering of these organisms, it has been shown that a significant reduction of the amount of lactic acid and acetic acid can occur once the genes that encode for enzymes responsible for the catalysis of pyruvate or acetyl-CoA to these organic acids such as lactate dehydrogenase, acetate kinase and phosphotransacetylase are knocked out (Desai et al., 2004). However, the genetic tools needed to knock out certain genes in these types of organisms are still in a developmental stage with relatively few successes (Tyurin et al., 2004). The growth of these organisms is also known to be restricted by high concentrations of ethanol due to end product inhibition of glycolytic enzymes and extensive cell wall damage due to increased membrane fluidity (Ingram, 1990). A promising phenomenon as shown with *C. thermocellum* is that it developed an increased tolerance to ethanol during prolonged

exposures (Taillez et al., 1989).

There are also several filamentous fungi which possess strong cellulolytic capability which have also been implicated as possible CBP-candidates. Some of the species include *Fusarium oxysporum* (Ruiz et al., 2007), *Neurospora crassa* (Rao et al., 1985), *Mucor indicus* and *Rhizopus oryzae* (Abedinifar et al., 2009). The abovementioned fungi can even produce limited amounts of ethanol when subjected to anaerobic conditions. Genetic tools to engineer these fungi would also need to address its ethanol productivity by reducing its, amongst others, lactic acid production.

In the recombinant cellulolytic strategy, one would begin with organisms with an inherent high ethanol yield and productivity from either hexoses, pentoses or both. Along with *S. cerevisiae* other potential candidates include the Gram-negative γ-proteobacterium *Zymomonas mobilis* and certain *Pichia* species (Yu & Zhang, 2004). These organisms, in general, lack the capability to efficiently depolymerise the polysaccharides found in plant biomass and thus cannot utilize cellulose or hemicellulose as sole carbohydrate source. Significant efforts have been made towards the construction of strains of these types of organisms by means of genetic engineering to produce enzymes that are able to depolymerise the cellulose and hemicellulose components in plant biomass (Van Zyl et al., 2007) and will be discussed later-on. One of the biggest challenges for these types of organisms is to increase the titer of heterologously-produced enzyme(s) used to depolymerise the polysaccharides in plant biomass.

The enzymatic hydrolysis of plant biomass has been approached in literature mainly from the perspective of an enzymatically-oriented intellectual paradigm, where enzymes are characterised almost independent of the organisms from which it was isolated (Lynd et al., 2002). The CBP processing strategy requires that the hydrolysis of plant biomass – especially that of cellulose – be perceived in terms of a microbial paradigm. The microbial paradigm perspective lends itself to place emphasis on different fundamental issues, organisms and cellulase systems as an integrative entirety.

2.3.3. FEASIBILITY OF *S. CEREVISIAE* AS A CANDIDATE FOR CBP

2.3.3.1. FERMENTATION ROBUSTNESS

S. cerevisiae possesses both essential and desirable traits which makes it a prime candidate for future CBP applications for a variety of reasons as shown in Table 2.5 (Van Zyl et al., 2007). Table 2.5 also shows the traits still needed for *S. cerevisiae* to be a competitive CBP organism.

TABLE 2.5. Current feasibility status of *S. cerevisiae* for CBP application (Van Zyl et al., 2007)

Essential traits
Ability to ferment most hexoses found in plant biomass natively
High ethanol yield and productivity (Crabtree-positive)
High ethanol tolerance
High tolerance toward inhibitors found in pretreated plant biomass
General robustness for industrial application
High level of heterologous gene expression
High level of secreted heterologous proteins
Desirable traits
Generally Recognized as Safe (GRAS) status
Recyclable
Simultaneous fermentation of sugars
Minimum nutrient supplementation
Amenable to DNA manipulation and transformation
Lacking traits
Cannot ferment any pentoses found in plant biomass natively
Possesses no depolymerases to degrade the complex polysaccharides of lignocellulose

S. cerevisiae has been exploited in the food and beverage industries for many generations and holds the advantage over competitor CBP organisms since it has been selected for increased fermentative capability for centuries by brewers and breadmakers alike. *S. cerevisiae* is a Crabtree-positive organism, implying that it can ferment glucose at high concentrations of more than 20-40 mM to ethanol under aerobic conditions (Verstrepen et al., 2003). This is mainly due to the yeast's overriding glucose suppression circuit that inhibits the Krebs cycle and oxidative phosphorylation of glucose. *S. cerevisiae* employs the Emden-Meyerhof-Parnas glycolytic pathway (Figure 2.8) to dissimilate glucose. Under anaerobic conditions, an ethanol production rate of 30 mmol.g biomass^{-1}h^{-1} can be attained (Bakker et al., 2000). Ethanol concentrations exceeding 4% v/v are toxic to the majority of microorganisms especially bacteria (Casey & Ingledew, 1986). *S. cerevisiae* has one of the highest ethanol tolerances of microorganisms and can exceed well beyond 10% v/v. Glucose is transported across the cell membrane by means of facilitated diffusion with an elaborate hexose transporter system

containing up to 32 members of the hexose transporter family (Boles & Hollenberg, 1997). When it comes to other sugar monomers found in plant biomass and its utilization by *S. cerevisiae*, three key criteria have to be fulfilled: (i) the presence of functional transporters; (ii) the presence of enzymes that are able to direct the metabolism of the particular sugar monomer to the main glycolytic pathway and (iii) the maintenance of a closed redox balance (Van Maris et al., 2006). *S. cerevisiae* is known to ferment the glucose isomers mannose and fructose comparatively efficiently, as they are also transported via the hexose transporter system. After phosphorylation by hexokinase, only isomerization has to take place for both of these sugars to be incorporated into the glycolytic pathway. The hexose galactose can also be metabolized by *S. cerevisiae* (Melcher, 1997) albeit differently from fructose and mannose. Galactose is also taken up by a hexose transporter and is incorporated into the glycolytic pathway via the Leloir pathway. Unlike, fructose and mannose, galactose metabolism is tightly regulated and it is not metabolized in the presence of glucose (Johnston et al., 1994).

During physico-chemical pretreatment of lignocellulosic biomass, a variety of compounds are released that are inhibitory to growth of microorganisms (Klinke et al., 2004). These compounds can be classified in three major groups: (i) weak acids, (ii) furan derivatives and (iii) phenols (Palmqvist et al. 1999). Even though the number and identity of these compounds can vary with the particular plant biomass feedstock being used and the pretreatment conditions, acetic acid, furfural which is derived from pentoses and 5-hydroxymethylfurfural (5-HMF) which is derived from hexoses are the most prevalent (Taherzedeh et al., 1997; Ingram et al., 1999). As opposed to most bacteria, *S. cerevisiae* is known to have a general hardiness towards many of these compounds (Fischer et al., 2008). With regards to furfural and 5-HMF, *S. cerevisiae* is capable of reducing these compounds to the less toxic furfuryl and 5-hydroxymethyl furfuryl alcohols respectively with relative efficacy (Taherzedeh et al., 2000a; Taherzedeh et al., 2000b). Overexpression of the *ADH6* gene in *S. cerevisiae*, which is known to be a 5-HMF-reducing enzyme, has improved conversion of 5-HMF to its furfuryl alcohol derivative (Petersson et al., 2006). Overexpression of *ZWF1*, which encodes a glucose-6-phosphate dehydrogenase and

forms part of the pentose phosphate pathway, resulted in higher furfural tolerance (Gorsich et al., 2006). This might be due to an increase in the amount of reducing equivalents produced.

2.3.3.2. GENETIC AMENABILITY

S. cerevisiae is a well established eukaryotic model organism to gain more information on a myriad of fundamental biological processes (Nevoigt, 2008). It was the first eukaryote whose complete genome was mapped (Goffeau et al., 1996). Currently, several internet databases exist that contain an enormous amount of information on *S. cerevisiae* genes, open reading frames, and gene products such as the *Saccharomyces* Genome Database (http://www.yeastgenome.org/) and the Comprehensive Yeast Genome Database (http://mips.gsf.de/genre/proj/yeast/). There are also databases that provide access to results from genome-wide microarray studies (http://transcriptome.ens.fr/ymgv/index.php) and networks of protein interactions (http://www.thebiogrid.org/). Another noteworthy example of useful information on *S. cerevisiae* is the *Saccharomyces* Genome Deletion Project (Winzeler et al., 1999) which has developed a collection of knockout strains covering 96% of the yeast genome.

The first genetic transformation of *S. cerevisiae* was recorded in 1978 (Hinnen et al., 1978). Since then, significant progress have been made in transforming both laboratory and industrial strains of yeast (Akada, 2002). Heterologous DNA can be delivered via episomal, integrative or centrosomal plasmids. Ploidy and availability for selectable markers are essential considerations for successful transformations. Laboratory yeasts are usually haploid and can contain several auxotrophic markers. Industrial strains – the strains most likely to be used in large-scale ethanol production – are diploid or polyploid and lack most of the auxotrophic markers found in laboratory strains. Even though the auxotrophic markers for *URA3*, *TRP1* and *ECM31* have been used in industrial strains (Kitamoto et al., 1990; Shimoi et al., 2000), transformation of industrial strains relies heavily on the use of heterologous drug-resistant markers, yeast drug-resistant markers and yeast overexpression markers conferring drug resistance (Akada, 2002). Although most of these types of markers have been tested on

episomal plasmids, integration plasmids will most likely be implemented for industrial strains due to the loss-of-plasmid that occurs when no selective pressure is maintained. Popular genes or sequences used for targeted integration are rRNA genes (Fujiti et al., 1990) and δ sequences (Sakai et al., 1990) due to the multiple areas in the genome where copies of these genes or sequences occur. Most of the shortcomings of *S. cerevisiae* as a candidate CBP organism in which genetic manipulation might be necessary, can be dealt with comparative ease.

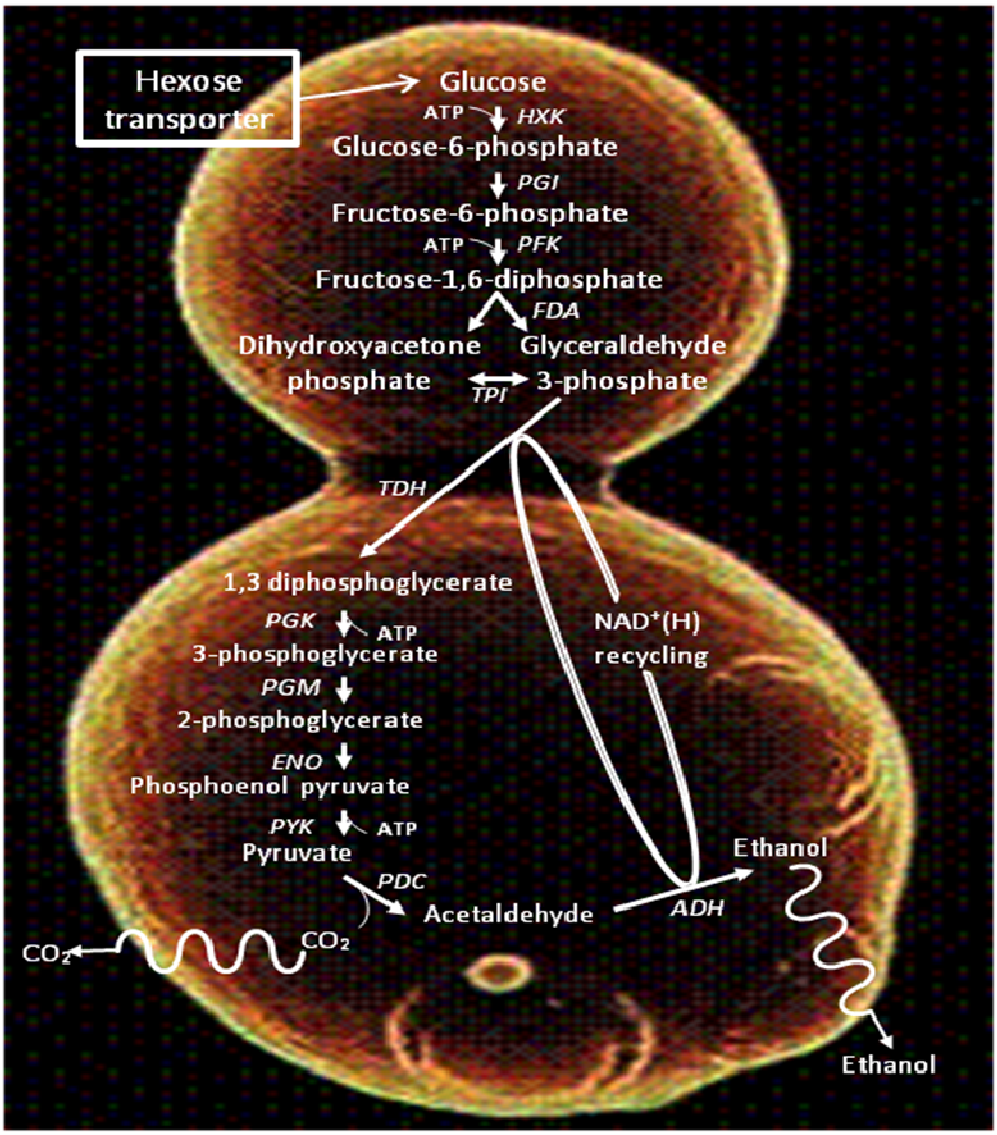


FIGURE 2.8. The Embden-Meyerhof-Parnas pathway employed by *S. cerevisiae* to dissimilate glucose and produce ethanol (Russell, 2003). Under anaerobic conditions, or aerobic conditions where sugar levels are high, the NADH formed by glyceraldehyde-3-phosphate dehydrogenase (TDH1) is reoxidised via the combined activity of pyruvate decarboxylase (PDC1) and alcohol dehydrogenase (ADH1) – which ensures redox balancing and the production of ethanol (Van Maris et al., 2006).

2.3.4. STRIDES MADE IN THE CONSTRUCTION OF *S. CEREVISIAE* AS A CANDIDATE CBP ORGANISM

Despite all the advantages of using *S. cerevisiae* as a CBP organism, it has several shortcomings especially with regards to its narrow substrate range (Van Zyl et al., 2007). As shown in Table 2.5, *S. cerevisiae* cannot natively ferment the pentoses commonly found in lignocellulosic biomass i.e. arabinose and xylose. It also cannot ferment L-rhamnose or galacturonic acid, a hexose sugar and a hexose sugar acid commonly found in certain lignocellulosic hydrolysates with high pectin content (Van Maris et al., 2006). From a cellulose fermentation perspective, the most important shortcoming is the lack of any native cellulase activity.

Several advances have been made in the construction of a cellulolytic *S. cerevisiae* (Van Zyl et al., 2007). The groundwork relies on observing how the most effective native cellulolytic organims degrade cellulose and attempt to mimic that in a recombinant *S. cerevisiae*. As was mentioned earlier in this review, cellulolytic microorganisms employ three different strategies to degrade cellulose. Of the three, most emphasis has been put on mimicking the non-complexed strategy employed by aerobic fungi and bacteria, although the successful heterologous expression of the scaffoldin protein that forms part of the cellulosome was reported recently (Lilly et al., 2009). Thus, for full enzymatic hydrolysis of cellulose in *S. cerevisiae*, successful heterologous expression of a β-D-glucosidase, endoglucanase and exoglucanase is needed. In a study where only a β-D-glucosidase (Cel3A) of *Saccharomycopsis fibuligera* was expressed in *S. cerevisiae*, ethanol productivity on cellobiose was comparable to that obtained on glucose showing success in the final step in cellulose hydrolysis (Van Rooyen et al., 2005). In another study where the same β-D-glucosidase was co-expressed with an endoglucanase (Cel7B) of *T. reesei*, anaerobic growth and ethanol production on PASC was reported (Den Haan et al., 2007a). Where all three types of enzymes have been co-expressed, higher ethanol yields were reported on PASC although much higher cell densities were used (Fujita et al., 2004). No one has as of yet reported in the literature any substantial growth or ethanol production on a crystalline form of cellulose like Avicel without the external addition of enzymes. Since exoglucanases are the only type of enzyme with any substantial activity on crystalline celluloses, the expression thereof would be of great interest. Table 2.6 lists some of

the exoglucanases that have been successfully expressed in *S. cerevisiae* – most of them individually. In general, exoglucanases have a much lower specific activity compared to endoglucanases and β-D-glucosidases mainly due to the crystalline substrate. In addition, most of the exoglucanases expressed in *S. cerevisiae* showed a relatively low titer (Den Haan et al., 2007b). It has been reported that hyperglycosylation of these secreted proteins by the yeast might decrease their specific activity (Reinikainen et al., 1992).

Exoglucanase expression has been identified as a limiting factor for yeast as a candidate CBP organism (Lynd et al., 2005) with the amount of exoglucanase expressed by *S. cerevisiae* required to be up to 10 % of the total cell protein to enable growth on a crystalline cellulose (Den Haan et al., 2007b). It is thus essential to investigate ways of improving the titer of exoglucanases produced and thereby improve the overall cellulolytic capability of the recombinant *S. cerevisiae*.

TABLE 2.6. Exoglucanases that have been heterologously produced by *S. cerevisiae*.

Host organism (Fungal/Bacterial)	Glycoside hydrolase family	Titer produced by *S. cerevisiae* (mg/L)	Reference*
Penicillium chrysogenum (F)	7	ND	Hou et al. (2007)
Thermoascus aurantiacus (F)	7	0.1	Hong et al. (2003)
Aspergillus aculeatus (F)	7	7	Takada et al. (1998)
Aspergillus niger (F)	7	ND	Den Haan et al. (2007b)
Phaenerochaete chrysosporium (F)	7	ND	Den Haan et al. (2007b)
Trichoderma reesei (F)	7	ND	Den Haan et al. (2007b)
Penicillium janthinellum (F)	7	ND	Koch et al. (1993)
Melanocarpus albomyces (F)	7	ND	Voutilainen et al. (2007)
Cellulomonas fimi (B)	6	2.5	Curry et al. (1988)
Agaricus bisporus (F)	6	ND	Chow et al. (1994)
Trichoderma reesei (F)	6	10	Zurbriggen et al. (1990)

*Note that some of these exoglucanases have been expressed in *S. cerevisiae* on several occasions while only one reference is given.
ND: Not determined

2.3.5. ADVANCES IN THE METABOLIC ENGINEERING OF *S. CEREVISIAE*

Traditionally, strain improvements of *S. cerevisiae* were conducted via random mutagenesis or cross-breeding (Ostergaard et al., 2000). As recombinant DNA techniques became more advanced, it enabled researchers to manipulate strains at a quicker rate and in a more specific manner.

Metabolic engineering of an organism is defined as the directed improvement of a specific phenotype of the organism which can be obtained with a combination of theoretical modelling, data from biochemical analyses and genetic engineering (Bailey, 1991; Ostergaard et al., 2000). This strategy can be divided into two components (i) analysis of cell properties to identify likely targets for genetic engineering followed by the (ii) genetic engineering of cells where after cells can be assayed to observe if the desired phenotype was achieved.

Since *S. cerevisiae* is widely considered as a suitable host for biofuel production, metabolic engineering could add great value in improving its performance capacity e.g. its ethanol production, tolerance towards inhibitors and substrate utilization (Lee et al., 2008). Table 2.7 shows a few examples in literature from the past 17 years where native *S. cerevisiae* genes were either disrupted or overexpressed to alter its metabolism. Examples of note are the overexpression of *PSE1* (Chow et al., 1992) and the disruption of *DDI1* (Lustgarten & Gerst, 2000) – two studies which unintentionally led to significant improvements in the secretion of native proteins of *S. cerevisiae*. One of the drawbacks of *S. cerevisiae* as a heterologous protein producer is its relatively low level of secretion capacity compared to other yeast like *Pichia pastoris*, *Hansenula polymorpha*, *Kluyveromyces* spp. and *Yarrowia lipolytica* (Müller et al., 1998; Buckholz & Gleeson, 1991). It is thus important to investigate improvements in the secretion capacity of *S. cerevisiae* by means of metabolic engineering as many of the heterologous gene products required for CBP should preferably be secreted in large quantities.

Thus, besides the heterologous production of hydrolytic enzymes like cellulases and xylanases in *S. cerevisiae* it is also vital to explore how the host's metabolism can be streamlined to make it a better ethanol producer from lignocellulosic biomass.

In this study we have attempted to further develop *S. cerevisiae* as a CBP organism for the conversion of lignocellulosic biomass to fuel ethanol by examining the production of novel cellulases and enhancing cellulase production by this host organism.

TABLE 2.7. Examples of native genes genetically engineered to alter phenotype of *S. cerevisiae*.

Name of gene	Difference in phenotype compared to parental strain	Reference
PSE1^{O}	Improved secretion of native proteins	Chow et al., 1992
MIG1$^{\Delta}$, *MIG2*$^{\Delta}$	Increase in specific growth rate on glucose	Klein et al., 1999
DDI1$^{\Delta}$	Increased secretion of native proteins	Lustgarten & Gerst, 1999
GDH1$^{\Delta}$, *GLT1*^{O}, *GLN1*^{O}, *GDH2*^{O}	Increase in ethanol productivity	Nissen et al., 2000
MTT1^{O}	Increased maltotriose uptake	Dietvorst et al., 2005
GRE3$^{\Delta}$, *TAL1*^{O}, *TKL1*^{O}, *RPE1*^{O}, *XKS1*^{O}, *RPI1*^{O}	Improved anaerobic xylose utilization	Kuyper et al., 2005
ADH6^{O}	Increased HMF reduction capability	Peterson et al., 2006
TPI1$^{\Delta}$, *ADH1*$^{\Delta}$, *GPD1*^{O}	Increased glycerol production	Cordier et al., 2007
TKL1$^{\Delta}$, *TKL2*$^{\Delta}$, *XKS2*$^{\Delta}$, *DOG1*^{O}	Increased xylitol production from xylose	Toivari et al., 2007
DAK1$^{\Delta}$, *DAK2*$^{\Delta}$	Improved dihydroxyacetone production	Nguyen & Nevoigt, 2009

O - Overexpression
Δ - Disruption

2.4. LITERATURE CITED

Abedinifar S, Karimi K, Khanahmadi M, Taherzadeh MJ 2009. Ethanol production by *Mucor indicus* and *Rhizopus oryzae* from rice straw by separate hydrolysis and fermentation. Biomass Bioenerg 33: 828-833.

Akada R 2002. Genetically modified industrial yeast ready for application. J Biosci Bioeng 94: 536-544.

Atalla RH, Vanderhart DL 1984. Native cellulose: a composite of two distinct crystalline forms. Science 223: 283-285.

Arioli T, Peng L, Betzner AS, Burn J, Wittke W, Herth W 1998. Molecular analysis of cellulose biosynthesis in *Arabidopsis*. Science 279: 717-720.

Bailey BK 1996. Performance of ethanol as a transportation fuel. In: Handbook on bioethanol: production and utilization (Wyman CE ed.) Taylor & Francis, Washington DC, USA. pp 37-58.

Bailey JE 1991. Toward a science of metabolic engineering. Science 252: 1668-1675.

Baker JO, Adney WS, Thomas SR, Nieves RA, Chou YC, Vinzant TB, Tucker MP, Laymon RA, Himmel ME 1995. Synergism between purified bacterial and fungal cellulases. In: Enzymatic degradation of insoluble carbohydrates ACS symosium series 618. (Himmel ME, Saddler JN eds.) Americal Chemical Society, Washington DC, USA. pp 113-141.

Bakis R 2007. Electricity production opportunities from multipurpose dams (case study). Renew Energ 32: 1723-1738.

Bakker BM, Bro C, Kötter P, Luttik MAH, van Dijken JP, Pronk JT 2000. The mitochondrial alcohol dehydrogenase Adh3p is involved in a redox shuttle in *Saccharomyces cerevisiae*. J Bacteriol 182: 4730-4737.

Bao W, Renganathan V 1992. Cellobiose oxidase of *Phanerochaete chrysosporium* enhances crystalline cellulose degradation by cellulases. FEBS Lett 302: 77-80.

Barr BK, Holewinski RJ 2002. 4-Methyl-7-thioumbelliferyl-beta-D-cellobioside: a fluorescent, nonhydrolyzable substrate analogue for cellulases. Biochemistry 41: 4447-4452.

Baskaran S, Ahn H-J, Lynd LR 1995. Investigation of the ethanol tolerance of *Clostridium thermosaccharolyticum* in continuous culture. Biotechnol Prog 11: 276-281.

Bayer EA, Setter E, Lamed R 1985. Organization and distribution of the cellulosome in *Clostridium thermocellum*. J Bacteriol 163: 552-559.

Berger E, Zhang D, Zverlov VV, Schwarz WH 2007. Two noncellulosomal cellulases of *Clostridium thermocellum*, Cel9I and Cel48Y, hydrolyse crystalline cellulose synergistically. FEMS Microbiol Lett 268: 194-201.

Bielecki S, Galas E 1991. Microbial β-glucanase different from cellulases. Crit Rev Biotechnol 10: 275-305.

Biely P, Mislovicová D, Toman R 1985. Soluble chromogenic substrates for the assay of endo-1,4-beta-xylanases and endo-1,4-beta-glucanases. Anal Biochem 144: 142-146.

Blanchard JE, Withers SG 2001. Rapid screening of the aglycone specificity of glycosidases: applications to enzymatic synthesis of oligosaccharides. Chem Biol 8: 627-633.

Blouin FA,. Martin LF, Rowland SP 1970. Gel permeation properties of cellulose III-Measurement of pore structure of unmodified and of mercerized cottons in fibrous form. Textile Res J 40: 809-813.

Bolam DN, Ciruela A, Mcqueen-Mason S, Simpson P, Williamson MP, Rixon JE, Boraston AB, Hazlewood GP, Gilbert, HJ 1998. *Pseudomonas* cellulose-binding domains mediate their effects by increasing enzyme substrate proximity. Biochem J 331: 775-781.

Boles E, Hollenberg CP 1997. The molecular genetics of hexose transport in yeast. FEMS Microbiol Rev 21: 85-111.

Boraston AB, McLean BW, Kormos JM, Alam M, Gilkes NR, Haynes CA, Tomme P, Kilburn DG, Warren RAJ 1999. Carbohydrate-binding modules: diversity of structure and function. In: Recent Advances in Carbohydrate Bioengineering (Gilbert HJ, Davies GJ, Henrissat B, Svensson B eds.) Royal Society of Chemistry, Cambridge, UK. pp 202-211.

Boraston AB, Bolam DN, Gilbert HJ, Davies GJ 2004. Carboydrate-binding modules: fine-tuning polysaccharide recognition. Biochem J 382: 769-781.

Brinkman N, Wang M, Weber T, Darlington T 2005. Well-to-wheels analysis of advanced fuel/vehicle systems: a North American study of energy use, greenhouse gas emissions and criteria pollutant emissions. Argonne Natl Lab, Argonne, Illinois, USA. pp 62-68.

Broda P, Birch P, Brooks P, Copa-Patino JL, Sinnott ML, Tempelaars C, Wang Q, Wyatt A, Sims P 1994. *Phanerochaete chrysosporium* and its natural substrate. FEMS Microbiol Rev 13: 189-196.

Brown RM 2004. Cellulose structure and biosythesis. What is in store for the 21st century? J Polym Sci Pol Chem 42: 487-495.

Buckholz RG, Gleeson MAG 1991. Yeast systems for the commercial production of heterologous proteins. *Bio/Technology* 9: 1067-1072.

Casey GP, Ingledew WM 1986. Ethanol tolerance in yeasts. Crit Rev Microbiol 13: 219-280.

Chanzy H 1990. Aspects of cellulose structure. In: Cellulose source and exploitation. Industrial utilization, biotechnology and physico-chemical properties, (Kennedy JF, Phillips GO, Williams PA eds.) Ellis Horwood, New York, USA. pp 3-12.

Chaudhary P, Kumar NN, Deobagkar DN 1997. The glucanases of *Cellulomonas*. Biotechnol Adv 15: 315-331.

Chen HZ, Hayn M, Esterbauer H 1992. Purification and characterization of two extracellular β-glucosidases from *Trichoderma reesei*. Biochim Biophys Acta 1121: 54-60.

Chow CM, Yague E, Raguz S, Wood DA, Thurston CF 1994. The Cel3 gene of *Agaricus bisporus* codes for a modular cellulase and is transcriptionally regulated by the carbon source. Appl Environ Microbiol 60: 2779-2785.

Chow TY-K, Ash JJ, Dignard D, Thomas DY 1992. Screening and identification of a gene, *PSE-1,* that affects protein secretion in *Saccharomyces cerevisiae.* J Cell Sci 101: 709-719.

Claeyssens M, Van Tilbeurgh, Kamerling JP, Berg J, Vrsanska M, Biely P 1990. Studies of the cellulolytic system of the filamentous fungus *Trichoderma reesei* QM 9414. Substrate specificity and transfer activity of endoglucanase I. Biochem J 270: 251-256.

Cordier H, Mendes F, Vasconcelos I, Francois JM 2007. A metabolic and genomic study of engineered *Saccharomyces cerevisiae* strains for high glycerol production. Metab Eng 9: 364-378.

Cosgrove DJ, Li LC, Cho H, Hoffmann-Benning S, Moore RC, Blecker D 2002. The growing world of expansins. Plant Cell Physiol 43: 1436-1444.

Coutinho PM, Henrissat B 1999. The modular structure of cellulases and other carbohydrate-active enzymes: an integrated database approach. In: Genetics, biochemistry and ecology of cellulose degradation (Ohmiya K, Hayashi K, Sakka K, Kobayashi Y, Karita S, Kimura T eds.) Uni Publishers Co, Tokyo, Japan. pp 15-23.

Cowling EB 1975. Physical and chemical constraints in the hydrolysis of cellulose and lignocellulosic materials. Biotechnol. Bioeng. Symp. 5: 163-181.

Curry C, Gilkes N, O'Neill, G, Miller RC, Skipper N 1988. Expression and secretion of a *Cellulomonas fimi* exoglucanase in *Saccharomyces cerevisiae*. Appl Environ Microbiol 54: 476-484.

Davies GJ, Henrissat B 1995. Structures and mechanisms of glycosyl hydrolases. Structure 3: 853-859.

Decker CH, Visser J, Schreier P 2000. β-Glucosidases from five black *Aspergillus* species: study of their physico-chemical and biocatalytic properties. J Agric Food Chem 48: 4929-4936.

Delmer DP 1999. Cellulose biosynthesis: Exciting times for a difficult field of study. Annu Rev Plant Physiol Plant Mol Biol 50: 245-276.

Demain AL, Newcomb M, Wu JHD 2005. Cellulase, clostridia and ethanol. Microbiol Mol Biol Rev 69: 124-154.

Den Haan R, Rose SH, Lynd LR, Van Zyl WH 2007a. Hydrolysis and fermentation of amorphous cellulose by recombinant *Saccharomyces cerevisiae*. Metab Eng 9: 87-94.

Den Haan R, Mcbride JE, La Grange DC, Lynd LR, Van Zyl WH 2007b. Functional expression of cellobiohydro-lases in *Saccharomyces cerevisiae* towards one-step conversion of cellulose to ethanol. Enzyme Microb Technol 40: 1291-1299.

Denman S, Xue GP, Patel B 1996. Characterization of a *Neocallimastix patriciarum* cellulase cDNA (celA) homologous to *Trichoderma reesei* cellobiohydrolase II. Appl Environ Microbiol 62: 1889-1896.

Desai SG, Guerinot ML, Lynd LR 2004. Cloning of L-lactate dehydrogenase and elimination of lactic acid production via gene knockout in *Thermoanaerobacterium saccharolyticum* JW/SL-YS485 Appl Microbiol Biotechnol 65: 600-605.

Deshpande V, Raman HS Rao M 1983. Simultaneous saccharification and fermentation of cellulose to ethanol using *Penicillium funiculosum* cellulase and free or immobilized *Saccharomyces uvarum* cells. Biotechnol Bioeng 25: 1679-1684.

Dietvorst J, Londesborough J, Steensma HY 2005. Maltotriose utilization in lager yeast strains: *MTT1* encodes a maltotriose transporter. Yeast 22: 775-788.

Din N, Damude HG, Gilkes NR, Miller RC, Warren RAJ, Kilburn DG 1994. C1-Cx revisited: intramolecular synergism in a cellulase. Proc Natl Acad Sci USA 91: 11383-11387.

Divne, C, Ståhlberg J, Reinikainen T, Ruohonen L, Pettersson G, Knowles JKC, Teeri TT, Jones TA 1994. The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. Science 265: 524-528.

Divne C, Ståhlberg J, Teeri TT, Jones TA 1998. High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei*. J Mol Biol 275: 309-325.

Doi RH, Kosugi A 2004. Cellulosomes: Plant cell wall degrading enzyme complexes. Nature Reviews 2: 541-551.

Doi RH 2008. Cellulases of mesophilic microorganisms cellulosome and noncellulosome producers. Ann NY Acad Sci 1125: 267-279.

Ekinci MS, Martin JC, Flint HL 2001. Expression of a cellulase gene, *celA*, from the rumen fungus *Neocallimastix patriciarum* in *Streptococcus bovis* by means of promoter fusions. Biotechnol Lett 24: 735-741.

Eriksson KE, Blanchette RA, Ander P 1990. Biodegradation of cellulose. In: Microbial and enzymatic degradation of wood and wood components (Eriksson KE ed.) Springer-Verlag, Berlin, Germany. pp 89-180.

Fan LT, Lee Y, Beardmore DH 1980. Major chemical and physical features of cellulose materials as substrates for enzymatic hydrolysis. Adv Biochem Eng 14: 101-117.

Farrell AE, Plevin RJ, Turner BT, Jones AD, O'Hare M, Kammen DM 2006. Ethanol can contribute to energy and environmental goals. Science 311: 506-508.

Fischer CR, Klein-Marcuschamer D, Stephanopoulos G 2008. Selection and optimization of microbial hosts for biofuels production. Metab Eng 10: 295-304.

Fleming K, Gray DG, Matthews S 2001. Cellulose crystallites. Chemistry 7: 1831-1835.

Fridleifsson IB 2001. Geothermal energy for the benefit of the people. Renew Sust Energ Rev 5: 299-312.

Fujita, Y., J. Ito, M. Ueda, H. Fukuda, and A. Kondo. 2004. Synergistic saccharification, and direct fermentation to ethanol, of amorphous cellulose by use of an engineered yeast strain codisplaying three types of cellulolytic enzyme. Appl Environ Microbiol 70: 1207-1212.

Fujiti T, Kondo K, Shimizu F, Sone H, Tanaka J, Inoue T. 1990. Application of a ribosomal DNA integration vector in the construction of a brewer's yeast having alpha-acetolactate decarboxylase activity. Appl Environ Microbiol 56: 997-1003.

Fülöp L, Ponyi T, 1997. Rapid screening for endo-β-1,4-glucanase and endo-β-1,4-mannanase activities and specific measurement using soluble dye-labelled substrates. J Microbiol Methods 29: 15-21.

Gao PJ, Chen GJ, Wang TH, Zhang YS, Liu J. 2001. Non-hydrolytic disruption of crystalline structure of cellulose by cellulose binding domain and linker sequence of cellobiohydrolase I from *Penicillium janthinellum*. Shengwu Huaxue Yu Shengwu Wuli Xuebao 33: 13-18.

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG 1996. Life with 6000 genes. Science 274: 563-567.

Gomez LD, Steele-King CG, McQueen-Mason SJ 2008. Sustainable liquid biofuels from biomass: the writing's on the walls. New Phytol. 178: 473-485.

Gorsich SW, Dien BS, Nichols NN, Slininger PJ, Liu ZL, Skory CD. 2006. Tolerance to furfural-induced stress is associated with pentose phosphate pathway genes *ZWF1*, *GND1*, *RPE1*, and *TKL1* in *Saccharomyces cerevisiae*. Appl Microbiol Biotechnol 71: 339-349.

Gross R, Leach M, Bauen A 2003. Progress in renewable energy. Environ Int 29: 105-122.

Henriksson G, Johansson G, Pettersson G 2000. A critical review of cellobiose dehydrogenases. J Biotechnol 78: 93-113.

Henrissat B, Teeri TT, Warren RAJ 1998. A scheme for designating enzymes that hydrolyse the polysaccharides in the cell walls of plants. FEBS Lett. 425: 352-354.

Hinnen A, Hicks JB, Fink GR 1978. Transformation of yeast. Proc Natl Acad Sci USA 75: 1929-1933.

Hill J, Nelson E, Tilman D, Polasky S, Tiffany D 2006. Environmental, economic and energetic cost and benefits of biodiesel and ethanol biofuels. PNAS 103: 11206-11210.

Hon DN 1996. Functional biopolymers: a new dimensional creativity in lignocellulosic chemistry. In: Chemical modification of lignocellulosic materials (Hon DN ed.) Marcel Dekker, New York, USA. pp 1-10.

Hong J, Tamaki H, Yamamoto K, Kumagai H 2003. Cloning of a gene encoding thermostable cellobiohydrolase from *Thermoascus aurantiacus* and its expression in yeast. Appl Microbiol Biotechnol 63: 42-50.

Hou Y, Wang T, Long T, Zhu H 2007. Cloning, sequencing and expression analysis of the first cellulase gene encoding cellobiohydrolase 1 from a cold-adaptive *Penicillium chrysogenum* FS010. Acta Bioch Bioph Sin 39: 101-107.

Howard RL, Abotsi E, Jansen van Rensburg EL, Howard S 2003. Lignocellulose biotechnology: issues of bioconversion and enzyme production. African Journal of Biotechnology 2: 602-619.

Ingram LO 1990. Ethanol tolerance in bacteria. Crit Rev Biotechnol 9: 305-319.

Ingram LO, Aldrich HC, Borges AC, Causey TB, Martinez A, Morales F, Saleh A, Underwood SA, Yomano LP, York SW, Zaldivar J, Zhou S 1999. Enteric bacterial catalyst for fuel ethanol production. Biotechnol Prog 15: 855-866.

Irwin D, Shin DH, Zhang S, Barr BK, Sakon J, Karplus PA, Wilson DB 1998. Roles of the catalytic domain and two cellulose binding domains of *Thermomonospora fusca* E4 in cellulose hydrolysis. J Bacteriol 180: 1709-1714.

Irwin DC, Spezio M, Walker LP, Wilson DB 1993. Activity studies of eight purified cellulases: specificity, synergism, and binding domain effects. Biotechnol Bioeng 42: 1002-1013.

Jenkins J, Leggio LL, Harris G Pickersgill R 1995. β-Glucosidase, β-galactosidase, family A cellulases, family F xylanases and two barley glycanases for a superfamily with 8-fold/architecture and with two conserved glutamates near the carboxy-terminal ends of strands four and seven. FEBS Lett 362: 281-285.

Jeoh T, Wilson DB, Walker LP 2006. Effect of cellulase mole fraction and cellulose recalcitrance on synergism in cellulose hydrolysis and binding. Biotechnol Prog 22: 270-277.

Johnston M, Flick JS, Pexton T 1994. Multiple mechanisms provide rapid and stringent glucose repression of *GAL* gene expression in *Saccharomyces cerevisiae*. Mol Cell Biol 14: 3834-3841.

Joselin Herbert GM, Iniyan S, Sreevalsan E, Rajapandian S 2007. A review of wind energy technologies. Renew Sust Energ Rev 11: 1117-1145.

Jun HS, Qi M, Gong J, Egbosimba EE, Forsberg CW 2007. Outer membrane proteins of *Fibrobacter succinogenes* with potential roles in adhesion to cellulose and in cellulose digestion. J Bacteriol 189: 6806-6815.

Jung ED, Lao G, Irwin D, Barr BK, Benjamin A, Wilson DB 1993. DNA sequences and expression in *Streptomyces lividans* of an exoglucanase gene and an endoglucanase gene from *Thermomonospora fusca*. Appl Environ Microbiol. 59: 3032-3043.

Juy M, Amit AG, Alzari PM, Poljak RJ, Béguin P, Aubert J, Claeyssens M 1992. Three-dimensional structure of a thermostable bacterial endoglucanase. Nature 357: 89-91.

Kerr RA 1998. OPEC's second coming. Science 281: 1129.

Kerr RA 2007. Global warming is changing the world. Science 316: 188-190.

Keshk S 2002. *Gluconacetobacter xylinus*: A new resource for cellulose. Egypt J Biotechnol 11: 305-310.

Kitamoto K, Oda K, Gomi K, Takahashi K 1990. Construction of uracil and tryptophan auxotrophic mutants from *sake* yeasts by disruption of *URA3* and *TRP1* genes. Agric Biol Chem 54: 2979-2987.

Klein CJL, Rasmussen JJ, Rønnow B, Olsson L, Nielsen J 1999. Investigation of the impact of *MIG1* and *MIG2* on the physiology of *Saccharomyces cerevisiae*. J Biotechnol 68: 197-212.

Klein GL, Snodgrass WR 1993. Cellulose. In: Encyclopedia of food Science, food technology and nutrition. (Macrae R, Robinson RK, Saddler MJ eds.) Academic Press, London, UK. pp 758-767.

Klemm D, Schmauder H-P, Heinze T 2002. Cellulose. In: Biopolymers: biology, chemistry, biotechnology, application, polysaccharide II. (Vandamme E, De Baets S, Steinbüchel A eds.), Wiley-VCH, Weinheim, Germany. pp 275-319.

Klinke HB, Thomsen AB, Ahring BK 2004. Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. Appl Microbiol Biotechnol 66: 10-26.

Koch A, Weigel CTO, Schulz G 1993. Cloning, sequencing, and heterologous expression of a cellulase encoding cDNA (*cbh1*) from *Penicillium janthinellum.* Gene 124: 57-65.

Kono H, Numata Y 2004. Two-dimensional spin-exchange solid-state NMR study of the crystal structure of cellulose II. Polymer 45: 4541-4547.

Koyama M, Helbert W, Imai T, Sugiyama J, Henrissat B 1997. Parallel-up structure evidences the molecular directionality during biosynthesis of bacterial cellulose. Proc Nat Acad Sci USA 94: 9091-9095.

Krässig HA 1993. Cellulose: Structure accessibility and reactivity. In: Polymer monographs. (Krässig HA ed.) Gordon and Breach Science Publishers, Switzerland. pp 167-323.

Kuyper M, Hartog MMP, Toirkens MJ, Almering MJH, Winkler AA, van Dijken JP, Pronk JT 2005. Metabolic engineering of a xylose-isomerase-expressing *Saccharomyces cerevisiae* strain for rapid anaerobic xylose fermentation. FEMS Yeast Res 5: 399-409.

Lee SK, Chou H, Ham TS, Lee TS, Keasling JD 2008. Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. Curr Opin Biotechnol 19: 556-563.

Lehtiö J 2002. Functional studies and engineering of family 1 carbohydrate binding modules. Ph. D. thesis.

Li Y, Irwin DC, Wilson DB 2007. Processivity, substrate binding, and mechanism of cellulose hydrolysis by *Thermobifida fusca* Cel9A. Appl Environ Microbiol 73: 3165-3172.

Lilly M, Fierobe H-P, Van Zyl WH, Volschenk H 2009. Heterologous expression of a *Clostridium* mini-cellulosome in *Saccharomyces cerevisiae*. FEMS Yeast Res 9: 1236-1249.

Linder M, Teeri TT 1997. The roles and function of cellulose binding domains. J Bacteriol 57: 15-28.

López-Contreras AM, Smidt H, Van der Oost J, Claassen PA, Mooibroek H, de Vos WM 2001. *Clostridium beijerinckii* cells expressing *Neocallimastix patriciarum* glycoside hydrolases show enhanced lichenan utilization and solvent production. Appl Environ Microbiol 67: 5127-5133.

Lou J, Dawson KA, Strobel HJ 1996. Role of phosphorolytic cleavage in cellobiose and cellodextrin metabolism by the ruminal bacterium *Prevotella ruminicola*. Appl Environ Microbiol 62: 1770-1773.

Lu YP, Zhang Y-HP, Lynd LR 2006. Enzyme–microbe synergy during cellulose hydrolysis by *Clostridium thermocellum*. Proc Natl Acad Sci USA 103: 16165-16169.

Lustgarten V, Gerst JE 1999. Yeast *VSM1* encodes a v-SNARE binding protein that may act as a negative regulator of constitutive exocytosis. Mol Cell Biol 19: 4480-4494.

Lynd LR 1996. Overview and evaluation of fuel ethanol production from cellulosic biomass: technology, economics, the environment, and policy. Annu Rev Energy Environ 21: 403-465.

Lynd LR 1997. Cellulose Ethanol: Technology in relation to environmental goals and policy formulation. Peer-reviewed chapter, In: Transportation, energy, and environment: how far can technology take us? (DeCicco J, DeLucchi M eds.) American council for an energy-efficient economy press, Washington DC, USA. pp 109-134.

Lynd LR, Baskaran S, Casten S, 2001. Salt accumulation resulting from base added for pH control, and not ethanol, limits growth of *Thermoanaerobacterium thermosaccharolyticum* HG-8 at elevated feed xylose concentrations in continuous culture. Biotechnol Prog 17: 118-125.

Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS 2002. Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev 2002 66: 506-577.

Lynd LR, Van Zyl WH, McBride JE, Laser M 2005. Consolidated bioprocessing of cellulosic biomass: an update. Curr Opin Biotechnol 16: 577-583.

Lyons TP 2003. Ethanol around the world: rapid growth in policies, technology and production In: The alcoholic textbook fourth edition. (Jacques KA, Lyons TP, Kendall DR eds). Nottingham University Press, Nottingham, UK. pp 85-121.

Mangoyana RB 2009. Bioenergy for sustainable development: an African context. Phys Chem Earth 34: 59-64

McCarter J, Withers SG 1994. Mechanisms of enzymatic glycoside hydrolysis. Curr Opin Struc Biol 4: 885-892.

McCleary BV 1988. Solubility properties of barley β-glucan. In: Alternative end uses of barley. (Sparrow DHB, Lance RCM, Henry RJ eds.) Glen Osmond, Australia. pp 117-120.

McGrath CE, Wilson DB 2006. Characterization of a *Thermobifida fusca* β-1,3-glucanase (Lam81A) with a potential role in plant biomass degradation. Biochemistry 45: 14094-14100.

Melcher K 1997. Galactose metabolism in *Saccharomyces cerevisiae*: a paradigm for eukaryotic gene regulation. In: Yeast sugar metabolism. Technomic publishing Inc, Pennsylvania. pp 235-269.

Mielenz JR 2001. Ethanol production form biomass: technology and commercialization status. Curr Opin Microbiol 4: 324-329.

Mingardon F, Chanal A, Lopez-Contreras AM, Dray C, Bayer EA, Fierobe H-P 2007. Incorporation of fungal cellulases in bacterial minicellulosomes yields viable, synergistically acting cellulolytic complexes. Appl Environ Microbiol 73: 3822-3832.

Mizuta S, Brown RM 1992. High resolution analysis of the formation of cellulose synthesizing complexes in *Vaucheria hamata*. *Protoplasma* 166: 187-199.

Montgomery R 2004. Development of biobased products. Bioresource Technol 91: 1-29.

Morris D, "The MTBE crisis and the future of renewable fuels," Testimony before the U.S. Senate Agriculture, Nutrition and Forestry Committee, April 11, 2000.

Moser F, Irwin D, Chen S, Wilson DB 2008. Regulation and characterization of *Thermobifida fusca* carbohydrate-binding module proteins E7 and E8. Biotechnol Bioeng 100: 1066-1077.

Mosier N, Wyman C, Dale B, Elander R, Lee YY, Holtzapple M, Ladisch M 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. Bioresour Technol 96: 673-686.

Müller S, Sandal T, Kamp-Hansen P, Dalbøge H 1998. Comparison of expression systems in the yeasts *Saccharomyces cerevisiae*, *Hansenula polymorpha*, *Klyveromyces lactis*, *Schizosaccharomyces pombe* and *Yarrowia lipolytica*. Cloning of two novel promoters from *Yarrowia lipolytica* Yeast 14: 1267-1283.

Nevoigt E 2008. Progress in metabolic engineering of *Saccharomyces cerevisiae.* Microbiol Mol Biol Rev 72: 379-412.

Nguyen HTT, Nevoigt E 2009. Engineering of *Saccharomyces cerevisiae* for the production of dihydroxyacetone (DHA) from sugars: A proof of concept. Metab Eng 11: 335-346.

Nishiyama Y, Sugiyama J, Chanzy H, Langan P 2003. Crystal structure and hydrogen bonding system in cellulose I from synchrotron X-ray and neutron fiber diffraction. J Am Chem Soc 125: 14300-14306.

Nissen TL, Kielland-Brandt MC, Nielsen J, Villadsen J 2000. Optimization of ethanol production in *Saccharomyces cerevisiae* by metabolic engineering of the ammonium assimilation. Metab Eng 2: 69-77.

Ong E, Gilkes NR, Miller RC,Warren RAJ, Kilburn DG 1991. Enzyme immobilization using a cellulose-binding domain: properties of a β-glucosidase fusion protein. Enzyme Microb Technol 13: 59-65.

Ostergaard S, Olsson L, Nielsen J 2000. Metabolic engineering of *Saccharomyces cerevisiae.* Microbiol Mol Biol Rev 64: 34-50.

O'Sullivan AC 1997. Cellulose: the structure slowly unravels. Cellulose 4: 173-207.

Palmqvist E, Grage H, Meniander NQ, Hahn-Hägerdahl B 1999. Main and interaction effects of acetic acid, furfural and para-hydrobezoic acid on growth and ethanol productivity of yeast. Biotechnol Bioeng 63: 46-55.

Payen M 1838. Report to the societé d'encouragement, on two prizes, - one for perfecting the fabrication of dextrine, and its application to the useful arts; - the other for the extraction of sugar from dextrine. J Franklin Inst 2: 264-265.

Pearson P, Pena-Tores H 2000. Carbon abatement and new investment in liberalised electricity markets: a nuclear revival in the UK? Energ Policy 28: 115-35.

Petersson, A, Almeida JR, Modig T, Karhumaa K, Hahn-Hagerdal B, Gorwa-Grauslund, MF, Liden G 2006. A 5-hydroxymethyl furfural reducing enzyme encoded by the *Saccharomyces cerevisiae ADH6* gene conveys HMF tolerance. Yeast 23: 455-464.

Prasad S, Singh A Joshi HC 2007. Ethanol as an alternative fuel from agricultural, industrial and urban residues. Resour Conserv and Recy 50: 1-39.

Rabinovich ML, Melnik MS, Boloboba AV 2002. Microbial cellulases (review). Appl Biochem Microbiol 38: 305–321.

Rao M, Mishra C, Keskar S, Srinivasan MC 1985. Production of ethanol from wood and agricultural residues by *Neurospora crassa.* Enzyme Microb Technol 7: 625-628.

Reese ET, Sui RGH, Levinson HS 1950. The biological degradation of soluble cellulose derivates and its relationship to the mechanism of cellulose hydrolysis. J Bacteriol 59: 485-497.

Reichenbecher M, Lottspeich F, Bronnenmeier K 1996 Purification and properties of a cellobiose phosphorylase (CepA) and a cellodextrin phosphorylase (CepB) from the cellulolytic thermophile *Clostridium stercorarium.* Eur J Biochem 247: 262-267.

Reinikainen T, Ruohonen L, Nevanen T, Laaksonen L, Kraulis P, Jones TA, Knowles JKC, Teeri TT 1992 Investigation of the function of mutated cellulose-binding domains of *Trichoderma reesei* cellobiohydrolase I. Proteins 14: 475-482.

Rescigno A, Rinaldi AC, Curreli N, Olianas A, Sanjust E 1994. A dyed substrate for the assay of endo-1,4-beta-glucanases. J Biochem Biophys Methods 28: 123-129.

Rotticci-Mulder JC, Gustavsson M, Holmquist M, Hult K, Martinelle M 2001. Expression in *Pichia pastoris* of *Candida antarctica* lipase B and lipase B fused to a cellulose-binding domain Protein Expres Purif 21**:** 386-392.

Ruiz E, Romero I, Moya M, Sanchez S, Bravo V, Castro E 2007. Sugar fermentation by *Fusarium oxysporum* to produce ethanol. World J Microbiol Biotechnol 23: 259-267.

Russell I 2003 Understanding yeast fundamentals. In: The alcoholic textbook fourth edition. (Jacques KA, Lyons TP, Kendall DR eds). Nottingham University Press, Nottingham, UK. pp 85-121.

Sakai A, Shimizu Y, Hishinuma F 1990. Integration of heterologous genes into the chromosome of *Saccharomyces cerevisiae* using a delta sequence of yeast retrotransposon Ty. Appl Microbiol Biotechnol 33: 302-306.

Sakon J, Irwin D, Wilson DB, Karplus PA 1997. Structure and mechanism of endo/exocellulase E4 from *Thermomonospora fusca*. Nat Struct Biol 4: 810-818.

Saloheimo M, Paloheimo M, Hakola S, Pere J, Swanson B, Nyyssönen E, Bhatia A, Ward M, Penttila M 2002. Swollenin, a *Trichoderma reesei* protein with sequence similarity to the plant expansins, exhibits disruption activity on cellulosic materials. Eur J Biochem 269: 4202-4211.

Saxena IM, Brown RM 2005. Cellulose biosynthesis: current views and evolving concepts. Ann Bot (Lond) 96: 9-21.

Scheible W, Eshed R, Richmond T, Delmer D, Somerville C 2001. Modifications of cellulose synthase confer resistance to isoxaben and thiazolidinone herbicides in *Arabidopsis Ixr1* mutants. Proc Natl Acad Sci USA 98: 10079-10084.

Schwarz WH 2001. The cellulosome and cellulose degradation by anaerobic bacteria. Appl Microbiol Biotechnol 56: 634-649.

Şen Z 2005. Solar energy in progress and future research trends. Prog Energ Combust Sci 30: 367-416.

Shimoi H, Okuda M, Ito K 2000. Molecular cloning and application of a gene complementing pantothenic acid auxotrophy of sake yeast Kyokai no.7. J Biosci Bioeng 90: 643-647.

Sharrock KR 1988. Cellulase assay methods: a review. J Biochem Biophys Methods 17: 81-105.

Shibazaki H, Saito M, Kuga K, Okano T 1998. Native cellulose II production by *Acetobacter xylinum* under physical constraints. Cellulose 5: 165-173.

Shin H, Brown Jr RM 1999. GTPase Activity and biochemical characterization of a recombinant cotton fiber annexin. Plant Physiol 119: 925-934.

Sinnott ML 1990. Catalytic mechanism of enzymatic glycosyl transfer. Chem Rev 90 :1171-1202.

Sinor J, Bailey BK 1993. Current and potential future performance of ethanol fuels. SAE paper #930376, Society of automotive engineers. Warrendale, Pennsylvania.

Stephanopoulos G 2007. Challenges in engineering microbes for biofuels production. Science 315: 801-803.

Sternberg D, Vijayakumar P, Reese ET 1977. β-Glucosidase: microbial production and effect on enzymatic hydrolysis of cellulose. Can J Microbiol 23: 139-147.

Stone JE, Scallan AM 1968. A structural model for the cell wall of water-swollen wood pulp fibres based on their accessibility to macromolecules. Cellulose Chem Technol 3: 343-358.

Sugiyama J, Vuong R, Chanzy H 1991. Electron diffraction study on the two crystalline phases occurring in native cellulose from an algal cell wall. Macromolecules 24: 4168-4175.

Sun Y, Chen J 2002. Hydrolysis of lignocellulosic materials for ethanol production: a review. Bioresource Technol 83 :1-11.

Taherzadeh MJ, Eklund R, Gustafsson L, Niklasson C, Lidén G 1997. Characterization and fermentation of dilute-acid hydrolyzates from wood. Ind Eng Chem Res 36: 4659-4665.

Taherzadeh, MJ, Gustafsson L, Niklasson C, Liden G 2000a. Inhibition effects of furfural on aerobic batch cultivation of *Saccharomyces cerevisiae* growing on ethanol and/or acetic acid. J Biosci Bioeng 90: 374-380.

Taherzadeh, MJ, Gustafsson L, Niklasson C, Liden G 2000b. Physiological effects of 5-hydroxymethylfurfural on *Saccharomyces cerevisiae*. Appl Microbiol Biotechnol 53: 701-708.

Taillez P, Girard SH, Longin R, Beguin P, Millet J 1989. Cellulose fermentation by an asporogenous mutant and an ethanol-tolerant mutant of *Clostridium thermocellum*. Appl Environ Microbiol 55: 203-206.

Takada G, Kawaguchi T, Sumitani J, Arai M 1998. Expression of *Aspergillus aculeatus* No. F-50 cellobiohydrolase I (*cbh1*) and β-glucosidase 1 (*bgl1*) genes by *Saccharomyces cerevisiae*. Biosci Biotechnol Biochem 62: 1615-1618.

Tan KT, Lee KT, Mohamed AR 2008. Role of energy policy in renewable energy accomplishment: The case of second-generation bioethanol. Energ Policy 36: 3360-3365.

Toivari MH, Ruohonen L, Miasnikov AN, Richard P, Penttila M 2007. Metabolic engineering of *Saccharomyces cerevisiae* for conversion of D-glucose to xylitol and other five-carbon sugars and sugar alcohols. Appl Environ Microbiol 73: 5471-5476.

Teeri TT 1997. Crystalline cellulose degradation: new insight into the function of cellobiohydrolases. Trends Biotechnol 15: 160-167.

Tomme P, Warren RAJ, Gilkes NR 1995. Cellulose hydrolysis by bacteria and fungi. Adv Microb Physiol 37: 1-81.

Tsekos I, Orologas N, Herth W 1999. Cellulose microfibril assembly and orientation in some bangiophyte red algae: relationship between synthesizing terminal complexes and microfibril structure, shape, and dimensions. Phycologia 38: 217-224.

Tuohy MG, Walsh DJ, Murray PG, Claeyssens M, Cuffe MM, Savage AV, Coughlan MP 2002. Kinetic parameters and mode of action of the cellobiohydrolases produced by *Talaromyces emersonii*. Biochim Biophys Acta 1596: 366-380.

Tyurin M, Desai SG, Lynd LR 2004. Electrotransformation of *Clostridium thermocellum*. Appl Environ Microbiol 70: 883-890.

UNFCCC, 1998. Kyoto protocol to the united nations framework convention on climate change. Available at: http://unfccc.int/resource/docs/convkp/kpeng.pdf. Website operational on 29 November 2009.

Usami S, Kirimura K, Imura M, Morikawa S 1990. Cellular localization of the constitutive β-glucosidase in *Trichoderma viride*. J Ferment Bioeng 70: 185-187.

Van Maris AJ, Abbott DA, Bellissimi E, Jan den Brink J, Kuyper M, Luttik MA, Wisselink HW, Scheffers WA, Van Dijken JP, Pronk JT. 2006. Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: current status. Antonie van Leeuwenhoek 90: 391-418.

Van Rooyen R, Hahn-Hägerdal B, La Grange DC, Van Zyl WH 2005. Construction of cellobiose-growing and fermenting *Saccharomyces cerevisiae* strains. J Biotechnol 120: 284-295.

Van Rooyen R 2002. Cloning of a novel *Bacillus pumilus* cellobiose utilising system: Functional expression in *Escherichia coli.* M. Sc. thesis.

Van Tilbeurgh H, Claeyssens M, de Bruyne CK 1982. The use of 4-methylumbelliferyl and other chromophoric glycosides in the study of cellulolytic enzymes. FEBS Lett 149: 152-156.

Van Walsum P, Lynd LR 1998. Allocation of ATP to synthesis of cells and hydrolytic enzymes in celluloytic fermentative microorganisms: Bioenergetics, kinetics, and bioprocessing. Biotech Bioeng 58: 316-320.

Van Zyl WH, Lynd LR, Den Haan R, McBride JE 2007. Consolidated bioprocessing for ethanol production using *Saccharomyces cerevisiae*. Adv Biochem Eng Biot 108: 205-235.

Varrot A, Schülein M, Davies GJ 1999. Structural changes of the active site tunnel of *Humicola insolens* cellobiohydrolase, Cel6A, upon oligosaccharide binding. Biochemistry. 38: 8884-8891.

Verstrepen KJ, Derdelinckx G, Dufour JP, Winderickx J, Thevelein JM, Pretorius IS, Delvaux FR 2003. Flavor-active esters: adding fruitiness to beer. J Biosci Bioeng 96: 110-118.

Voutilainen SP, Boer H, Linder MB, Puranen T, Rouvinen J, Vehmaanperä J, Koivula A 2007. Heterologous expression of *Melanocarpus albomyces* cellobiohydrolase Cel7B, and random mutagenesis to improve its thermostability. Enzyme Microb Technol 41: 234-243.

Wada M, Chanzy H, Nishiyama Y, Langan P 2004. Cellulose IIII crystal structure and hydrogen bonding by synchrotron X-ray and neutron fiber diffraction. Macromolecules 37: 8548-8555.

Walseth CS. 1952. Occurrence of cellulases in enzyme preparations from microorganisms. TAPPI 35: 228-233.

Wang DIC, Avgerinos GC, Biocic I, Wang SD, Fang HY 1983. Ethanol from cellulosic biomass. Philos Trans R Soc Lond Ser B 300: 323-333.

Warren RAJ 1996. Microbial hydrolysis of polysaccharides. Annu Rev Microbiol 50: 183-212.

Watanabe H, Tokuda G 2001. Animal cellulases. Cell Mol Life Sci 58: 1167-1178.

Weimer PJ, Lopez-Guisa JM, French AD 1990. Effect of cellulose fine structure on kinetics of its digestion by mixed ruminal microorganisms in vitro. Appl Environ Microbiol 56: 2419-2421.

Whittington HW 2002. Electricity generation: options for reduction in carbon emissions. Phil Trans R Soc Lond A 360: 1653-1668.

Williams AG, Orpin CG 1987. Polysaccharide-degrading enzymes formed by three species of anaerobic rumen fungi grown on a range of carbohydrate substrates. Can J Microbiol 33: 418-426.

Wilson DB 1988. Cellulases of *Thermomonospora fusca*. Method Enzymol 160: 314-323.

Wilson DB 2008. Three microbial strategies for plant cell wall degradation. Ann NY Acad Sc 1125: 289-297.

Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connelly C, Davis K, Dietrich F, Dow SW, El Bakkoury M, Foury F, Friend SH, Gentalen E, Giaever G, Hegemann JH, Jones T, Laub M, Liao H, Liebundguth N, Lockhart DJ, Lucau-Danila A, Lussier M, M'Rabet N, Menard P, Mittmann M, Pai C, Rebischung C, Revuelta JL, Riles L, Roberts CJ, Ross-MacDonald P, Scherens B, Snyder M, Sookhai-Mahadeo S, Storms RK, Veronneau S, Voet M, Volckaert, G, Ward TR, Wysocki R, Yen GS, Yu K, Zimmermann K, Philippsen P, Johnston M, Davis RW 1999. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. Science 285: 901-906.

Wood TM 1988. Preparation of crystalline, amorphous, and dyed cellulase substrates. Method Enzymol 160: 19-25.

Wood TM 1991. Fungal cellulases. In: Biosynthesis and biodegradation of cellulose. (Haigler CH, Weimer PJ eds.) Marcel Dekker, New York, USA. pp 491-533.

Wood TM 1992. Fungal cellulases. Biochem Soc Trans 20: 46-53.

Wood TM, Bhat KM 1988. Methods for measuring cellulase activities. Methods Enzymol 160: 87-117.

Wood TM, McRae ST 1979. Synergism between enzymes involved in the solubilization of the native cellulose. Adv Chem Ser. 18: 181-210.

Wyman CE 1996. Ethanol production form lignocellulosic biomass: overview. In: Handbook on bioethanol. (Wyman CE ed.) Taylor & Francis, Washington DC, USA. pp 1-15.

Xie G, Bruce DC, Challacombe JF, Chertkov O, Detter, JC, Gilna P, Han CS, Lucas S, Misra M, Myers GL, Richardson P, Tapia R, Thayer N, Thompson LS, Brettin TS, Henrissat B, Wilson DB, McBride MJ 2007. Genome sequence of the cellulolytic gliding bacterium *Cytophaga hutchinsonii*. Appl Environ Microbiol 73: 3536-3546.

Yamada Y, Hoshino K, Ishikawa T 1997. The phylogeny of acetic acid bacteria based on the partial sequences of 16S ribosomal RNA : The elevation of the subgenus *Gluconoacetobacter* to the generic level. Biosci Biotechnol Biochem 61: 1244-1251.

Yamanaka S, Ishihara M, Sugiyama J 2000. Structural modification of bacterial cellulose. Cellulose 7: 213-225.

Yu ZS, Zhang HX 2003. Pretreatments of cellulose pyrolysate for ethanol production by *Saccharomyces cerevisiae*, *Pichia* sp. YZ-1 and *Zymomonas mobilis*. Biomass Bioenerg 24: 257-262.

Zhang Y-HP, Lynd LR 2004. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. Biotechnol Bioeng 88: 797-824.

Zhang Y-HP, Himmel ME, Mielenz JR 2006. Outlook for cellulase improvement: Screening and selection strategies. Biotechnol Adv 24: 452-481.

Zhou W, Irwin DC, Escovar-Kousen J, Wilson DB 2004. Kinetic studies of *Thermobifida fusca* Cel9A active site mutant enzymes. Biochemistry 43: 9655-9663.

Zurbriggen B, Bailey MJ, Penttilä ME, Poutanen K, Linko M 1990. Pilot scale production of a heterologous *Trichoderma reesei* cellulase by *Saccharomyces cerevisiae.* J Biotechnol 13: 267-78.

Zverlov VV, Velikodvorskaya GA, Schwarz WH 2002. A newly described cellulosomal cellobiohydrolase, CelO, from *Clostridium thermocellum*: investigation of the exo-mode of hydrolysis, and binding capacity to crystalline cellulose. Microbiology 148: 247-255.

CHAPTER 3

HETEROLOGOUS CO-PRODUCTION OF *THERMOBIFIDA FUSCA CEL9A* WITH OTHER CELLULASES IN *SACCHAROMYCES CEREVISIAE*

N. van Wyk, R. den Haan, W.H. van Zyl

Department of Microbiology, University of Stellenbosch, Matieland 7602, South Africa.

3.1. ABSTRACT: The processive endoglucanase Cel9A of the moderately thermophilic actinomycete *Thermobifida fusca* was functionally produced in *Saccharomyces cerevisiae.* Recombinant Cel9A displayed activity on both soluble (carboxymethylcellulose) and insoluble (Avicel) cellulose substrates confirming its dual role as an endoglucanase and an exoglucanase. HPAEC analyses of soluble sugars released from Avicel revealed a cellobiose/glucose ratio of 2.5 ± 0.1. Growth by the recombinant strain on amorphous cellulose was possible due to the sufficient amount of glucose cleaved from the cellulose chain. This is the first confirmed report of *S. cerevisiae* growing on a cellulosic substrate as sole carbohydrate source while only expressing one recombinant gene. To improve the cellulolytic capability of *S. cerevisiae* and to investigate the level of synergy among cellulases produced by a recombinant host, the *cel9A* gene was co-expressed with four cellulase-coding genes of *Trichoderma reesei*: two endoglucanases *cel5A* (*egII*) and *cel7B* (*egI*), and two cellobiohydrolases *cel6A* (*cbh2*) and *cel7A* (*cbh1*). Synergy, especially between the Cel9A and the two cellobiohydrolases, resulted in a higher cellulolytic capability of the recombinant host.

3.2. INTRODUCTION

Lignocellulosic biomass has been identified as an almost inexhaustible and renewable raw material for the production of biofuels, in particular ethanol (Lin & Tanaka, 2006). A major obstacle for further progress in this field is the apparent lack of low-cost technology to address the recalcitrance of lignocellulose (Lynd et al., 2002). A promising approach is that of consolidated bioprocessing (CBP), where the production of cellulolytic enzymes, the enzymatic hydrolysis of the biomass and the fermentation of the resulting sugars are all combined in a

single bioprocessing step (Lynd et al., 2005). CBP has the potential to significantly reduce the costs involved in converting lignocellulosic biomass to fermentable sugars which can in turn be fermented to ethanol. For CBP to be realized, an organism must readily hydrolyze cellulose and convert the products of hydrolysis to a desired product – in this case ethanol – on an industrial scale. *Saccharomyces cerevisiae* is a promising CBP candidate as it produces ethanol at high concentrations, has GRAS status, can be genetically manipulated with relative ease and is generally robust in large-scale applications (Van Zyl et al., 2007). Major drawbacks of *S. cerevisiae* include the complete lack of a cellulose-degrading enzyme system and its relatively low secretion capacity. Attempts to construct *S. cerevisiae* strains capable of cellulose hydrolysis are based on mimicking natural cellulolytic organisms by expressing and secreting a multitude of synergizing cellulases. Strides made in the construction of such *S. cerevisiae* strains include the expression of a β-D-glucosidase (*cel3A* – previously known as *BGL1*) of *Saccharomycopsis fibuligera* in *S. cerevisiae* where the growth rate and ethanol productivity of the recombinant strain on cellobiose were comparable to that obtained on glucose (Van Rooyen et al., 2005). The same β-D-glucosidase was co-expressed with endoglucanase I (*cel7B*) of *Trichoderma reesei* and anaerobic growth with concomitant ethanol production on phosphoric acid-swollen cellulose (PASC) was observed (Den Haan et al., 2007a). The heterologous expression of exoglucanases, such as cellobiohydrolases, is problematic since in most cases low titers are achieved (Den Haan et al., 2007b). Exoglucanases are particularly important since they are capable of disrupting the crystalline structure of cellulose – the most prevalent state of cellulose found in lignocellulosic biomass.

The cellulolytic system of the moderately thermophilic actinomycete *Thermobifida fusca* is well-described and serves as a model for aerobic bacteria (Wilson, 2004). The seven known cellulases it produces include a processive endoglucanase (Cel9A) which acts as both an endo and an exoglucanase. Cel9A also shows significant levels of synergy in binary reactions with both endo and exoglucanases (Irwin et al., 1993).

In this study, the processive endoglucanase Cel9A of *T. fusca* was heterologously produced in *S. cerevisiae*. In an attempt to evaluate to what extent synergy between cellulases could contribute towards increasing the cellulolytic capacity of *S. cerevisiae*, the *cel9A* gene was further co-expressed with two endoglucanases *cel5A* (*egII*) and *cel7B* (*egI*), and two cellobiohydrolases *cel6A* (*cbh2*) and *cel7A* (*cbh1*) – all from the soft-rot fungus *T. reesei*.

3.3. MATERIALS AND METHODS

3.3.1. MEDIA AND CULTURE CONDITIONS

The chemicals and media components used in all experiments were of laboratory grade standard. Phosphoric acid swollen cellulose (PASC) was prepared as previously described (Zhang et al., 2006) using Avicel PH-101 (Fluka, Buchs, Switzerland) as crystalline source. The *Escherichia coli* strain DH5α was used for plasmid transformation and propagation. *E. coli* cells were cultivated in Luria Bertani medium (5 g/L yeast extract, 10 g/L NaCl, 10 g/L tryptone) supplemented with ampicillin (100 mg/L). *S. cerevisiae* Y294 transformants were selected and maintained on synthetic complete medium without uracil (SC^{-URA}) or without uracil and leucine ($SC^{-URA-LEU}$) plates (1.7 g/L yeast nitrogen base without amino acids and ammonium sulphate [Difco laboratories, Detroit, MI, USA], 5 g/L $(NH_4)_2SO_4$, 20 g/L glucose, 20 g/L agar and supplemented with amino acids as required). Autoselective *S. cerevisiae* strains were cultured in YPD medium (10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose). The recombinant strain Y294[CEL9A] and the reference recombinant strain Y294[REF] were cultured on YP-PASC medium (10 g/L yeast extract, 20 g/L peptone, ~10 g/L PASC). Yeast strains were routinely cultured in 250-mL erlenmeyer flasks containing 50 mL medium at 30°C, on a rotary shaker at 100 rpm. For aerobic growth experiments yeast strains were grown in baffled 250-mL Erlenmeyer flasks containing 50 mL growth medium. Precultures of the strains were grown on YPD medium. For growth on liquid YP-PASC medium at least three cultures of the strains tested were inoculated simultaneously. Samples were periodically taken and yeast cells in the media were counted in duplicate on a haemocytometer.

3.3.2. NUCLEIC ACID MANIPULATIONS AND TRANSFORMATIONS

Standard procedures for DNA manipulation were followed as outlined by Sambrook & Russel (2001). All restriction endonuclease enzymes were purchased from Fermentas (Vilnius, Lithuania). For polymerase chain reactions, the ExTaq polymerase (TakaraBio, Otsu, Japan) was used as recommended by the manufacturer with a Perkin Elmer GeneAmp® PCR System 2400 (The Perkin-Elmer Corporation, Norwalk, CT, USA). Details of all the primers used in this study are given in Table 3.1. All plasmids and strains used in this study are summarized in Table 3.2. Primers Tfuscel90-left and Tfuscel90-right were used to retrieve the *cel9A* gene of *T. fusca* from the plasmid D1020 (donated by David Wilson, Cornell University). A subsequent PCR product was cloned into the TA-cloning vector pGEM (Promega, Madison, WI, USA) – generating pGEM-Cel9A. The pGEM-Cel9A construct was digested with the restriction endonucleases *Nru*I and *Xho*I to obtain a 2518-bp fragment. This fragment was inserted in frame with the *xyn11B* secretion signal of *T. reesei* on the *E. coli/S. cerevisiae* shuttle vector yXYNSEC and named yXS+Cel9A. The complete $PGK1_P$-XYNSEC-*cel9A*-$PGK1_T$ gene cassette was amplified with PGK1H-left and PGK1H-right primers as a 3566-bp fragment and inserted into the TA-cloning vector pTZ57R/T (Fermentas) – generating pTZ-Cel9AHindIII. The pTZ-Cel9AHindIII construct was digested with *Hin*dIII and the $PGK1_P$-XYNSEC-*cel9A*-$PGK1_T$ cassette was inserted in the *Hin*dIII-digested plasmids pAZ40 and pegII. The final plasmid constructs were named pCel9A+EgI and pCel9A+EgII respectively. The $PGK1_P$-XYNSEC-*cel9A*-$PGK1_T$ gene cassette was also amplified with PGK1S-left and PGK1S-right primers with the resulting fragment inserted into the TA-cloning vector pTZ57R/T – generating pTZ-Cel9ASalI. The pTZ-Cel9ASalI construct was digested with *Sal*I and the $PGK1_P$-XYNSEC-*cel9A*-$PGK1_T$ cassette was inserted in the *Sal*I-digested plasmid pDLG77 resulting in the plasmid pCel9A+CBH1. The gene cassette $ENO1_P$-*cel6A*-$ENO1_T$ on the plasmid pAZ21 was amplified using the primers ENO1-left and ENO1-right resulting in a DNA fragment of the size 2542 bp. The fragment was inserted into pTZ57R/T and renamed pTZ-CBH2. The pTZ-CBH2 was partially digested with *Hin*dIII and was inserted into yXS+Cel9A resulting in plasmids pCel9A+CBH2.

S. cerevisiae Y294 strains were transformed using the LiOAc/DMSO protocol described by Hill et al. (1991). For the construction of autoselective strains the native *FUR1* gene, (Kern et al., 1991) was disrupted using plasmid pDF1 (La Grange et al., 1996) that was linearized with *Nco*I and *Ava*III. Double cross-over transformants were confirmed by PCR using the FUR1-L and FUR1-R primers.

TABLE 3.1: Primers used in this study. Relevant restriction sites are shown in bold.

Primer name	Sequence (5'-3')	Relevant restriction sites	Genbank accession number
Tfuscel90-left	AAGCT**TCGCGA**GAACCGGCGTTCAACTAC	*Nru*I	L20093
Tfuscel90-right	AGGC**CTCGAG**GTTGCCGTGTTAGGCG	*Xho*I	
ENO1-left	**GGATCC**ACTAGTCTTCTAGGCGGGTTATC	*Eco*RI	X99228
ENO1-right	**AAGCTT**GCGGCCGCAAAGAGGTTTAGACATTGG	*Hin*dIII	
PGK1H-left	ACTG**AAGCTT**GGATCCTTAAAGATGCCG	*Hin*dIII	NC001135
PGK1H-right	ACTG**AAGCTT**GGCCAAGCTTTAACGAAC	*Hin*dIII	
PGK1S-left	ACTG**GTCGAC**GGATCCTTAAAGATGCCG	*Sal*I	
PGK1S-right	ACTG**GTCGAC**GGCCAAGCTTTAACGAAC	*Sal*I	
FUR1-L	TCCGTCTGGCATATCCTA		M36485
FUR1-R	TTGGCTAGAGGACATGTA		

3.3.3. CELLULASE ACTIVITY ASSAYS

Yeast cells cultured in double-strength SC media for four days were spun down at 2,000 rpm for three min. The supernatants were used to determine endoglucanase and exoglucanase activity. The dinitrosalicyclic acid (DNS) method was used to determine endoglucanase activity, based on the protocol described by Miller (1960) with carboxymethylcellulose (CMC) serving as cellulose substrate. Four-hundred-and-fifty microliters of 1% CMC in 0.05 M acetate buffer pH 5.0 was placed into test tubes and incubated in a waterbath at 50°C for five min. Fifty microliters of appropriately diluted yeast supernatants were added, vortexed and incubated for 15 min after which 750 µL of the DNS solution (1% 3,5-dinitrosalicyclic acid [Sigma, St. Louis, MO, USA], 20% potassium sodium tartrate, 1% NaOH, 0.2% phenol, 0.05% Na_2SO_3) was added and vortexed. Subsequently, samples were boiled for 15 min. Cooled-down samples were read spectrophotometrically at 540 nm.

TABLE 3.2: Plasmids and strains used in this study.

Plasmids propagated in *E. coli* DH5α	Abbreviated name	Relevant genotype	Source/Reference
Escherichia coli DH5α		*supE44 ΔlacU169 (ϕ80lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	Sambrook 2001
pGEM-Cel9A		*bla cel9A* (secretion signal removed)	This work
pTZ-Cel9AHindIII		*bla* $PGK1_P$-*XYNSEC-cel9A-* $PGK1_T$	This work
pTZ-Cel9ASalI		*bla* $PGK1_P$-*XYNSEC-cel9A-* $PGK1_T$	This work
pTZ-Cbh2		*bla* $ENO1_P$-*XYNSEC-cel6A-* $ENO1_T$	This work
yXYNSEC		*bla URA3* $PGK1_P$-*XYNSEC-PGK1*$_T$	Van Rooyen 2005
pAZ40		*bla URA3* $ENO1_P$-*cel7B-* $ENO1_T$	Den Haan 2007a
pLEGII		*bla URA3* $ENO1_P$-*cel5A-* $ENO1_T$	Du Plessis 2009
pDLG77		*bla URA3* $ADH2_P$-*cel7A-* $ADH2_T$	Den Haan 2007b
pAZ21		*bla* $ENO1_P$-*XYNSEC-cel6A-* $ENO1_T$	Den Haan 2007b
yXS+Cel9A		*bla URA3* $PGK1_P$-*XYNSEC-cel9A-* $PGK1_T$	This work
pCel9A+EgI		*bla URA3* $PGK1_P$-*XYNSEC-cel9A-PGK1*$_T$ $ENO1_P$-*cel7B-* $ENO1_T$	This work
pCel9A+EgII		*bla URA3* $PGK1_P$-*XYNSEC-cel9A-PGK1*$_T$ $ENO1_P$-*cel5A-* $ENO1_T$	This work
pCel9A+CBH1		*bla URA3* $PGK1_P$-*XYNSEC-cel9A-PGK1*$_T$ $ADH2_P$-*cel7A-* $ADH2_T$	This work
pCel9A+CBH2		*bla URA3* $PGK1_P$-*XYNSEC-cel9A-PGK1*$_T$ $ENO1_P$-*XYNSEC-cel6A-* $ENO1_T$	This work
pDF1		*bla fur1::LEU2*	La Grange 1996
Yeast strains constructed			
Saccharomyces cerevisiae Y294		α *leu2-3,*112 *ura3-52 his3 trp1-289*	ATCC 201160
(*fur1::LEU2* yXYNSEC)	Y294[REF]	*bla ura3/URA3* $PGK1_P$-*XYNSEC-* $PGK1_T$	Van Rooyen 2005
(*fur1::LEU2* pAZ40)	Y294[EGI]	*bla ura3/URA3* $ENO1_P$-*cel7B-* $ENO1_T$	Den Haan 2007a
(*fur1::LEU2* pLEGII)	Y294[EGII]	*bla ura3/URA3* $ENO1_P$-*cel5A-* $ENO1_T$	Du Plessis 2009
(*fur1::LEU2* pDLG77)	Y294[CBH1]	*bla ura3/URA3* $ADH2_P$-*cel7A-* $ADH2_T$	Den Haan 2007b
(*fur1:LEU2* pAZ21)	Y294[CBH2]	*bla ura3/URA3* $ENO1_P$-*XYNSEC-cel6A-* $ENO1_T$	Den Haan 2007b
(*fur1:LEU2* yXS+Cel9A)	Y294[CEL9A]	*bla ura/URA3* $PGK1_P$-*XYNSEC-cel9A-* $PGK1_T$	This work
(*fur1:LEU2* pCel9A+EgI)	Y294[CEL9A+EGI]	*bla ura3/URA3* $PGK1_P$-*XYNSEC-cel9A-PGK1*$_T$ $ENO1_P$-*cel7B-* $ENO1_T$	This work
(*fur1:LEU2* pCel9A+EgII)	Y294[CEL9A+EGII]	*bla ura3/URA3* $PGK1_P$-*XYNSEC-cel9A-PGK1*$_T$ $ENO1_P$-*cel5A-* $ENO1_T$	This work
(*fur1:LEU2* pCel9A+CBH1)	Y294[CEL9A+CBH1]	*bla ura3/URA3* $PGK1_P$-*XYNSEC-cel9A-PGK1*$_T$ $ADH2_P$-*cel7B-* $ADH2_T$	This work
(*fur1:LEU2* pCel9A+CBH2)	Y294[CEL9A+CBH2]	*bla ura3/URA3* $PGK1_P$-*XYNSEC-cel9A-PGK1*$_T$ $ENO1_P$-*XYNSEC-cel6A-* $ENO1_T$	This work

To determine exoglucanase activity, 300 µL of the yeast culture supernatant was added to deep-well microtiter plates with each well containing a 300 µL solution with 2% Avicel PH-105, 0.1 M acetate buffer pH 5.0, 0.04% NaN_3 and 0.3 µL Novozyme-188 (Sigma). The deep-well plate was sealed and incubated on a microtiter plate shaker and shaken at ~1000 rpm for 24 h

at 50°C. One hundred microliter samples of the enzyme-substrate mixture were taken at the 0 h and 24 h time intervals to determine background sugars that were present and the amount of sugars released respectively, using a modified DNS method. The samples were transferred to a 96-well PCR plate using a multi-channel pipette and were centrifuged at 2000 rpm for two min. Fifty microliters of the supernatant were subsequently pipetted along with 100 µL DNS solution into a clean 96-well PCR plate, sealed, heated at 99°C for five min and cooled at 4°C for one min. Absorbance values at 540 nm were read on a X-mark™ microtitre plate reader (Biorad, Hercules, CA, USA).

Glucose was used for the standard curve (0.125 mg/mL-2 mg/mL) from which the Units/mL of each sample was calculated. A unit of enzyme activity is equal to 1 µmol of monomeric sugar released per min. All values were normalized with the dry cell weight in mg/mL of the relevant yeast culture.

3.3.4. HIGH-PERFORMANCE ANION EXCHANGE CHROMATOGRAPHY

Concentrated enzyme solutions were made by adding 130 mL of the supernatant of Y294[CEL9A] grown for four days in YPD to 1% PASC, 2.5% Avicel and 2.5% BMCC, respectively. The enzyme-substrate mixtures were stirred overnight at 4°C. Subsequently, the enzyme-substrate mixtures were filtered through a glass fiber filter (Whatman, Kent, UK) and the enzyme-bound cellulose that remained on the filter was resuspended in 0.05 M citrate buffer pH 5.0. The cellulose-bound enzyme solutions were incubated at 37 °C on an orbital shaker and samples taken at 0, 2, 5 and 24 h. These samples were analysed for glucose and cellobiose content by high-performance anion exchange chromatography with a Dionex DX 500 chromatography system coupled to a pulsed amperometric detector (PAD, Dionex ED 40) using a CarboPac PA-10 column (Dionex, Sunnyvale, CA, USA). The column was initially activated with 20 mM NaOH and a flow-rate of 1 mL/min was maintained. After 25 min the NaOH concentration was increased to 225 mM for 6 min and then decreased to 20 mM. External standards (glucose and cellobiose) were used for quantification.

3.3.5. SDS-PAGE AND WESTERN BLOT ANALYSIS

Supernatants of yeast cultures grown for four days in YPD were boiled in denaturing buffer (200 mM Tris–HCl pH 6.8, 8% (w/v) SDS, 0.4% bromophenol blue, 40% (v/v) glycerol, and 400 mM DTT) for 4 min and subjected to SDS-PAGE (Laemmli, 1970). Before SDS-PAGE, some protein samples were de-*N*-glycosylated with the endoglycosidase PNGase F (New England Biolabs Ipswich, MA, USA) (Plummer & Tarentino, 1991). A 7.5% acrylamide gel was used and electrophoresis was carried out at 120 V for one hour in Tris-glycine buffer (25 mM Tris, 250 mM glycine, 0.1% SDS). After electrophoresis, proteins were transferred onto a nitrocellulose membrane using voltage of 4 V for 1.5 h with a Semiphor Mini Trans-Blot Cell (Hoefer Scientific, Holliston, MA, USA). Blocking of the membrane was carried out at room temperature overnight with gentle shaking in Western blocking buffer (WBB) containing TBST (50 mM Tris base, 150 mM NaCl, pH 7.5 with 0.1% (v/v) Tween-20) and 5% skim milk. Primary antibody raised in rabbit (a kind gift from David Wilson, Cornell University) was diluted 1:1000 in WBB and added to the membrane for two hours and incubated at room temperature with gentle shaking. The membrane was then washed with TBST and WBB for 15 min. Subsequently, the secondary antibody, anti-rabbit immunoglobulin G conjugated with alkaline phosphatase (Sigma), diluted 1:2000 in WBB was added to the membrane and incubated for two hours at room temperature. The blot was then washed with TBST and bands were visualized by pipetting 750 µL of CDP-Star (Roche, Mannheim, Germany) over the membrane followed by signal development with autoradiography.

3.3.6. INDIRECT ENZYME-LINKED IMMUNOSORBENT ASSAY (ELISA)

Flat-bottomed 96-well microtiter plates (Nalge Nunc, Rochester, NY, USA) were coated with 50 µL of yeast supernatant and incubated overnight at room temperature. All steps thereafter were also conducted at room temperature. Coated plates were washed three times with deionized water after which wells were filled with blocking buffer (BB) containing borate buffered saline (0.017 M $Na_2B_4O_7 \cdot 10H_2O$ and 0.12 M NaCl pH 8.5) with 0.05% Tween-20, 1 mM EDTA, 0.25% bovine serum albumin and 0.02% NaN_3) and incubated for 30 min. Fifty microliters of primary antibody diluted 1:1000 in BB were pipetted into wells and incubated for

two hours after which plates were again washed with deionized water. Fifty microliters of secondary antibody diluted 1:2000 in BB were then added and incubated for two hours. Plates were washed after which 75 µL of a freshly-made *p*NPP substrate solution (3 mM *p*-nitrophenylphosphate (Sigma), 0.05 M Na_2CO_3, 0.05 mM $MgCl_2$) were added in each well. The colour-development reaction was stopped after approximately 30 min by adding 15 µL of 0.5 M NaOH and values were spectrophotometrically measured with a microtiter plate reader at a wavelength of 405 nm. Purified Cel9A protein (a kind gift from David Wilson, Cornell University) was used to prepare a standard curve (5.7 µg/mL - 0.71 µg/mL).

3.4. RESULTS

3.4.1. CLONING AND EXPRESSION OF *T. FUSCA* AND *T. REESEI* GENES IN *S. CEREVISIAE*

The *T. fusca cel9A* gene, with its secretion signal coding region replaced with that of the xylanase II gene (*xyn11B*) of *Trichoderma reesei,* was placed under the transcriptional regulation of the constitutive *S. cerevisiae* promoter and terminator of the *PGK1* gene. The additional cellulase-encoding genes of *T. reesei* were placed under the transcriptional control of either the constitutive promoter of *ENO1* (*cel5A*, *cel6A* and *cel7B*) or the derepressive promoter of *ADH2* (*cel7A*). Confirmation of the insertion of heterologous cellulase genes in yeast transformants was done with PCR analysis on total yeast DNA isolates using primers that annealed to the individual cellulase genes (results not shown).

3.4.2. GROWTH OF Y294[CEL9A] ON PASC

The recombinant *S. cerevisiae* strain Y294[CEL9A] exhibited cellulase activity on both soluble amorphous (Figure 3.1 and Table 3.3) and insoluble crystalline (Table 3.3) cellulose substrates. HPAEC analyses of Cel9A-treated insoluble cellulosic substrates showed that both cellobiose and glucose were cleaved from the cellulose chain (Figure 3.2). The amount of cellobiose units cleaved per glucose molecule was calculated to be 2.5 ± 0.1 and was relatively similar for all three insoluble cellulose substrates. It was also calculated that roughly 14.3 µg/ml glucose was released per hour by the recombinantly produced Cel9A. Y294[CEL9A] was shown to grow on YP-PASC agar plates and in liquid media as shown in Figure 3.3 whereas the reference strain

Y294[REF] showed no significant growth in the allotted time on the same substrate. This indicated that the level of glucose cleaved off by Cel9A was sufficient for growth on insoluble amorphous cellulose as sole carbohydrate source.

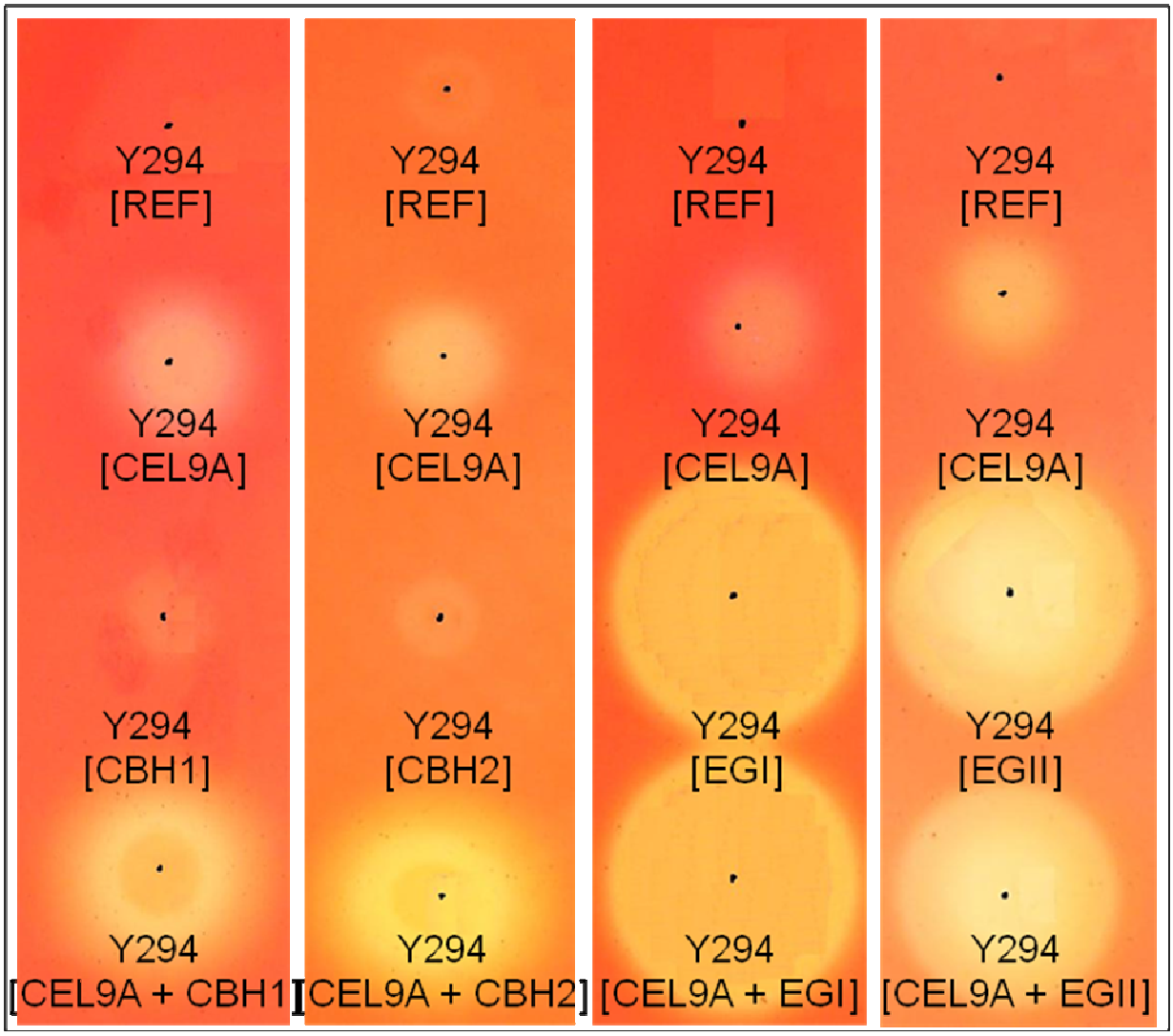


FIGURE 3.1: Carboxymethyl-cellulose plate assays on *S. cerevisiae* Y294 expressing individual cellulases and the combinations of Cel9A with other cellulases. Yeast colonies were grown for two days at 30 °C on 1 % CMC-containing plates and washed off. Plates were stained with 0.1 % Congo Red and destained with 1 M NaCl.

3.4.3. CELLULASE ACTIVITIES OF CO-EXPRESSION STRAINS

Table 3.3 shows the cellulase activities on amorphous (CMC) and crystalline cellulose (Avicel) produced by the recombinant yeast strains. The CMC data showed no significant increase in cellulase activity when cellulases were co-expressed in yeast. With Avicel as substrate, the strains where Cel9A was co-expressed with endoglucanases of *T. reesei* (*cel5A* and *cel7B*) had

significantly higher activities than with the cellulases expressed individually. Strain Y294[CEL9A+EGI] had a 1.3-fold increase in cellulase activity than the sum of the strains expressing the cellulases individually. The same was not observed with strain Y294[CEL9A+EGII] as the sum of strains Y294[CEL9A] and Y294[EGII] expressing the cellulases individually was nearly the same. Strain Y294[CEL9A+CBHII] also gave a 1.3-fold increase in comparison to the sum of the strains expressing the cellulases individually. The biggest increase was observed with strain Y294[CEL9A+CBHI] where a 1.9-fold increase in cellulase activity was found in comparison to the cellulase activities of the individual strains. This is consistent with data found by Irwin et al. (1993) which found higher levels of synergy between purified Cel7A and Cel9A than with any other cellulase in binary reactions.

TABLE 3.3: Cellulase activity profiles on CMC and Avicel for the cellulases heterologously expressed and co-expressed in the *S. cerevisiae* strain Y294 and Cel9A concentration produced by recombinant strains. Activity was determined as Units/mg dry cell weight for CMC and Units/g dry cell weight for Avicel. Plus/minus signs indicate the standard variation from the mean of three repeats.

Yeast strain[a]	Activity on CMC Units/mg DCW[b]	Activity on Avicel Units/g DCW[b]	Titer of Cel9A produced (µg/ml)[c]
Y294[CEL9A]	1.888± 0.088	0.660 ± 0.037	1.59
Y294[EGI]	6.491± 0.467	1.324 ± 0.061	
Y294[CEL9A+EGI]	6.427± 0.415	2.635 ± 0.061	1.32
Y294[EGII]	4.983± 0.497	1.151 ± 0.159	
Y294[CEL9A+EGII]	5.415± 0.292	1.704 ± 0.094	1.35
Y294[CBHI]	1.512± 0.310	1.071 ± 0.106	
Y294[CEL9A+CBHI]	2.211± 0.140	3.327 ± 0.253	1.27
Y294[CBHII]	1.426± 0.591	4.101 ± 0.243	
Y294[CEL9A+CBHII]	1.839± 0.055	6.302 ± 0.468	1.53

[a] The value obtained for the reference strain Y294[REF] was subtracted from all values above.
[b] Activity was determined as Units/mg dry cell weight for CMC and Units/g dry cell weight for Avicel.
[c] Concentrations represent averages of supernatants of three cultures with standard deviations of less than 10%

3.4.4. GLYCOSYLATION AND RECOMBINANT CEL9A PRODUCTION BY *S. CEREVISIAE*

Results obtained from Western blot analyses suggested that the recombinant Cel9A protein was heavily glycosylated (Figure 3.4). The glycosylation pattern spans an area between approximately 90 kDa to over 200 kDa. Analysis of the translated protein sequence revealed six

potential *N*-glycosylation sites (Asn-Xaa-Ser/Thr) of which two on the C-terminal CBM have a likeliness to be glycosylated of more than 50% using NetNGlyc 1.0 Software (http://www.cbs.dtu.dk/services/NetNGlyc/). An indirect ELISA method was used to quantify the amount of Cel9A enzyme produced by Y294[CEL9A] and the co-expression strains. ELISA values obtained of cultures are shown in Table 3.3. Y294[CEL9A] was shown to produce Cel9A at a level of 1.59 µg/mL while the strains where Cel9A were co-expressed with other cellulases led to a titer of slightly less than the strain where Cel9A was produced individually.

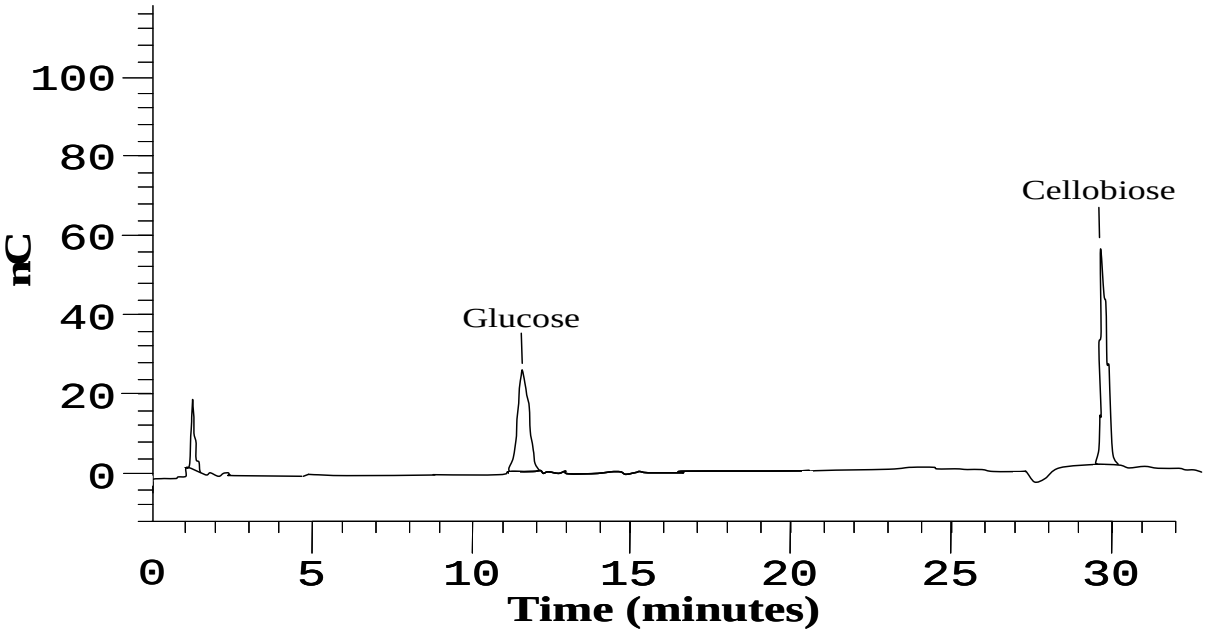


FIGURE 3.2: HPAEC chromatogram of products released from an insoluble cellulose substrate, in this case BMCC, by Cel9A expressed in *S. cerevisiae* Y294[CEL9A] incubated for 2h at 50 °C.

3.5. DISCUSSION

Cel9A of the moderately thermophilic actinomycete *Thermobifida fusca* (previously known as *Thermomonospora fusca*) is known as a processive endoglucanase that exhibits both endoglucanase and exoglucanase activity (Irwin et al., 1993). Of all the cellulases characterized from *T. fusca*, Cel9A has the highest activity towards crystalline cellulose, especially BMCC (Li et al., 2007). Another notable feature of Cel9A is the synergy it has with both endoglucanases and exoglucanases (Irwin et al., 1993). This is the first account where the *cel9A* was heterologously produced in an eukaryotic host and its dual role as an endoglucanase and an exoglucanase was retained in the recombinant host. The successful expression and secretion of the actinomycete

cellulase-encoding gene in *S. cerevisiae* is somewhat surprising as the gene sequence of *cel9A* has a GC content of 68% – much higher than that of the recombinant host's 38 %. Adding to that, the calculated codon adaptation index for the gene calculated according to Carbone (2003) and Sharp & Li (1987) methods were rather low (0.0929 and 0.0727, respectively).

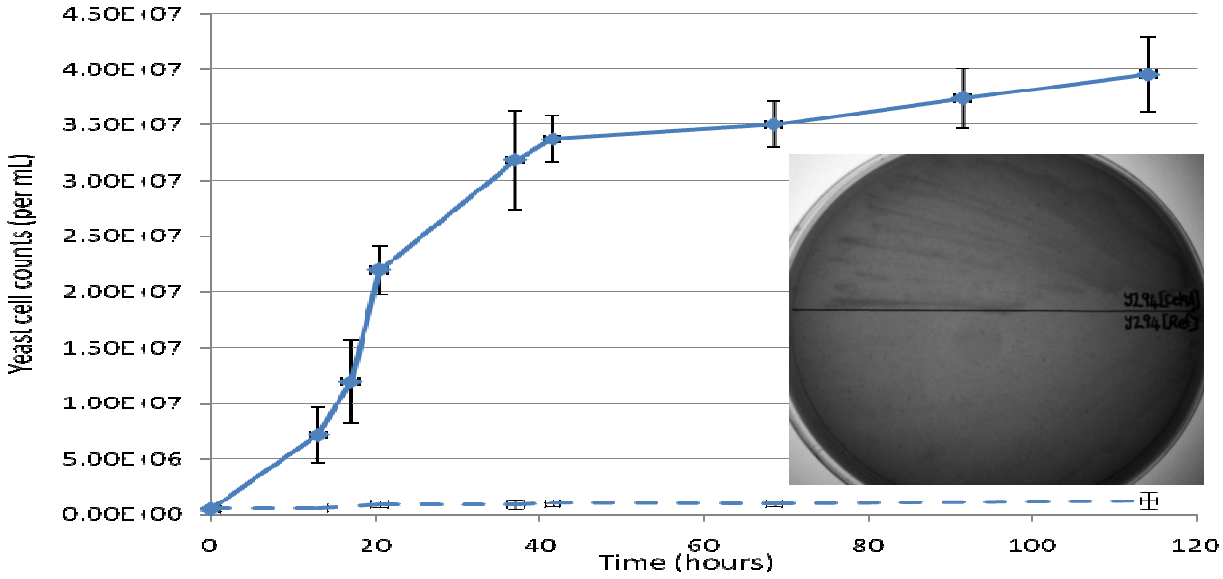


FIGURE 3.3: Aerobic growth curve of Y294[CEL9A] (solid line) on 1 % phosphoric acid swollen cellulose supplemented with 2 % peptone and 1 % yeast extract. Insert is a petri dish containing YP-PASC agar showing growth by Y294[CEL9A] on the top part of the plate. In both the liquid and solid media the reference strain Y294[REF] (dashed line) showed no significant growth during the allotted time frame.

The fact that significant amounts of glucose were cleaved from the cellulose chain by Cel9A is consistent with previous thin-layer chromatography data that showed significant amounts of glucose present after cellulose hydrolysis (Zhou et al., 2004). The cellobiose/glucose ratio is extremely low for an exo-acting enzyme as most exoglucanases produce nearly no glucose from the cellulose chain. In both the plate assays and liquid assays Cel9A did not show higher cellulase activity than other cellulases heterologously produced by *S. cerevisiae*, but the fact that relatively high titers of Cel9A is produced and that significant amounts of glucose were cleaved off by Cel9A enabled the yeast to grow on an insoluble cellulose substrate. To our knowledge, this is the first account where significant growth was found by a *S. cerevisiae* strain on an insoluble cellulose substrate while expressing only one heterologous gene. This is also the first account where yeast growth on cellulose was observed without the need of heterologous β-glucosidase expression.

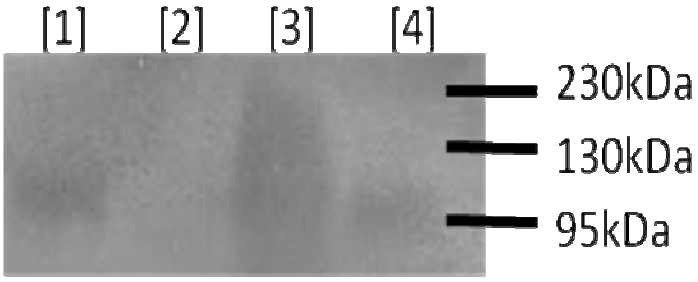


FIGURE 3.4: Western blot analysis of protein samples using Cel9A primary antibody. [1] Purified Cel9A used as positive control; [2] supernatant of Y294[REF]; [3] untreated Y294[CEL9A] supernatant [4] supernatant of Y294[CEL9A] treated with PNGaseF.

De-*N*-glycosylation of recombinant Cel9A samples removed the poly-disperse protein smear and yielded a protein species with a size similar to that of the native protein, suggesting that only *N*-glycosylation contributed toward the hyperglycosylation. However, recombinant Cel9A was still functional despite the level of hyperglycosylation observed.

The fact that an increase in the cellulase activities on Avicel was always observed when cellulase genes from *T. reesei* was co-expressed with Cel9A, implied that the increase in episomal plasmid size and the expected reduction on plasmid copy number did not affect the overall recombinant cellulolytic capacity negatively. The ELISA results showed that the levels of Cel9A produced when co-expressed with other cellulases were not dramatically less than when produced individually. Saturation with the expression of two cellulases in *S. cerevisiae* was thus not achieved and the expression of more cellulases could further improve the cellulolytic capability of *S. cerevisiae*. The apparent increase in halo formation on the plate assay of the Cel9A co-expressed with the two cellobiohydrolases (Figure 3.1) could be explained in terms of the much longer duration of incubation of the plates whereas the quantitative liquid assays were conducted for only 15 minutes. Improved growth of these recombinant *S. cerevisiae* strains co-expressing cellulase genes on cellulose substrates will not be expected mainly due to the small amounts of glucose produced by the *T. reesei* enzymes from cellulose in comparison to Cel9A.

The study showed that Cel9A of *T. fusca* should be strongly considered in *S. cerevisiae* strain development for CBP due to its relatively high level of production by *S. cerevisiae* and the synergy it demonstrated with other recombinant cellulases. The recombinant yeast strains generated in this study thus signify major progress toward the realization of CBP of lignocellulosic materials by *S. cerevisiae*.

3.6. LITERATURE CITED

Carbone A, Zinovyev A, Képès F 2003. Codon adaptation index as a measure of dominating codon bias. Bioinformatics 19: 2005-2015.

Den Haan R, Rose SH, Lynd LR, Van Zyl WH 2007a. Hydrolysis and fermentation of amorphous cellulose by recombinant *Saccharomyces cerevisiae*. Metab Eng 9: 87-94.

Den Haan R, Mcbride JE, La Grange DC, Lynd LR, Van Zyl WH 2007b. Functional expression of cellobiohydrolases in *Saccharomyces cerevisiae* towards one-step conversion of cellulose to ethanol. Enzyme Microb Technol 40: 1291-1299.

Du Plessis L, Rose SH, Van Zyl WH 2009. Exploring improved endoglucanase expression in *Saccharomyces cerevisiae* strains. Appl Microbiol Biotechnol DOI 10.1007/s00253-009-2403-z.

Hill J, Donald KA, Griffiths DE 1991. DMSO-enhanced whole cell yeast transformation. Nucleic Acids Res 19: 5791.

Irwin DC, Spezio M, Walker LP, Wilson DB 1993. Activity studies of eight purified cellulases: Specificity, synergism and binding domain effects. Biotechnol Bioeng 42:1002-1013.

Kern L, De Montigny J, Jund R, Lacroute F 1990. The *FUR1* gene of *Saccharomyces cerevisiae*: cloning, structure and expression of wild-type and mutant alleles. Gene 88: 149-157.

Laemmli UK 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.

La Grange DC, Pretorius IS, Van Zyl WH 1996. Expression of a *Trichoderma reesei* beta-xylanase gene (XYN2) in *Saccharomyces cerevisiae*. Appl Environ Microb 62:1036-1044.

Li Y, Irwin DC, Wilson DB 2007. Processivity, substrate binding, and mechanism of cellulose hydrolysis by *Thermobifida fusca* Cel9A. Appl Environ Microbiol 73: 3165-3172.

Lin Y, Tanaka S 2006. Ethanol fermentation from biomass resources: current state and prospects. Appl Microbiol Biotechnol 69: 627-642.

Lynd LR, Van Zyl WH, McBride JE, Laser M 2005. Consolidated bioprocessing of cellulosic biomass: an update. Curr Opin Biotech 16: 577-583.

Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS 2002. Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev 66: 506-577

Miller GL, Blum R, Glennon WE, Burton AL 1960. Measurement of carboxymethyl cellulase activity. Anal Biochem 2: 127-132.

Plummer Jr TH, Tarentino AL, 1991. Purification of the oligosaccharide-cleaving enzymes of *Flavobacterium meningosepticum*. Glycobiology 1: 257-63.

Sambrook J, Russel DB. 2001. Molecular cloning: A laboratory manual. Cold spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sharp PM, Li WH 1987. The codon adaptation index - a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res 15: 1281-1295.

Wilson DB 2004. Studies of *Thermobifida fusca* plant cell wall degrading enzymes. Chem Rec 4: 72-82.

Van Rooyen R, Hahn-Hägerdal B, La Grange DC, Van Zyl WH 2005. Construction of cellobiose-growing and fermenting *Saccharomyces cerevisiae* strains. J Biotechnol 120: 284-295.

Van Zyl WH, Lynd LR, Den Haan R, McBride JE 2007. Consolidated bioprocessing for ethanol production using *Saccharomyces cerevisiae*. Adv Biochem Eng Biot 108: 205-235.

Zhang YH, Cui J, Lynd LR, Kuang LS 2006. A transition from cellulose swelling to cellulose dissolution by *o*-phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure. Biomacromolecules 7: 644-648.

Zhou W, Irwin DC, Escovar-Kousen J, Wilson DB 2004. Kinetic studies of *Thermobifida fusca* Cel9A active site mutant enzymes. Biochemistry 43: 9655-9663.

CHAPTER 4

HETEROLOGOUS PRODUCTION OF NpCEL6A FROM *NEOCALLIMASTIX PATRICIARUM* IN *SACCHAROMYCES CEREVISIAE*

N. van Wyk, R. den Haan, W.H. van Zyl

Department of Microbiology, University of Stellenbosch, Matieland 7602, South Africa.

4.1. ABSTRACT: The cellobiohydrolase-encoding gene from the anaerobic chytrid *Neocallimastix patriciarum* (*Npcel6A*) was heterologously produced in *Saccharomyces cerevisiae*. Recombinant cellulase activity was detected on soluble cellulose substrates however little activity was found on Avicel. The unusual Asn-rich linker region along with the carbohydrate binding module (CBM) of NpCel6A was replaced with a Ser/Thr/Pro-rich linker and CBM from the TrCel6A of *Trichoderma reesei*. The yeast producing the hybrid cellulase had significantly lower activity on filter paper than the original NpCel6A. When the Asn-rich linker was fused to the catalytic domain of TrCel6A, significantly higher cellulase activity was found on filter paper than the original TrCel6A.

4.2. INTRODUCTION

Glycoside hydrolases are a group of enzymes that hydrolyze the glycosidic bonds in carbohydrates (Coutinho & Henrissat, 1999). Cellulases represent a large subgroup of glycoside hydrolases and are specific for cleaving the β-(1,4)-bonds of cellulose. Complete hydrolysis of crystalline cellulose found in lignocellulosic biomass requires a synergistic mix of different cellulases that include exoglucanases, endoglucanase, and β-glucosidase activities.

Cellulases are crucial for the economic production of cellulosic ethanol. Even though the cost involved in cellulase production has been dramatically decreased, large-scale cellulase production is still considered to be a key economic barrier for the production of ethanol from lignocellulosics (Stephanopolous, 2007). For the cost reduction of ethanol production from lignocellulosics, a promising strategy consists of engineering noncellulolytic ethanol-producers,

like *Saccharomyces cerevisiae*, to produce cellulases or to engineer cellulase producers to become ethanologenic (Lynd et al., 2005). This strategy is known as consolidated bioprocessing (CBP) and combines all the biologically-facilitated processes needed for lignocellulosics to be converted to ethanol.

When heterologous proteins are produced in a recombinant host, unpredictable co- and post-translational events occur which may alter the structure and function of the mature protein (Nevalainen et al., 2005). An important post-translational event is protein glycosylation which has been shown to influence the function, structural framework, and stability of proteins (Imperiali et al., 1999). Cellobiohydrolases, like *Trichoderma reesei* Cel6A and Cel7A, have been shown to be adversely impacted by glycosylation when heterologously produced by a recombinant host (Den Haan et al., 2007; Górka-Nieć et al., 2007; Jeoh et al., 2008)

Strides made in the construction of *S. cerevisiae* strains for the realization of CBP include the report of *S. cerevisiae* co-expressing the endoglucanase *cel7B* of *T. reesei* with the β-glucosidase (*cel3A*) of *Saccharomycopsis fibuligera* which showed growth on amorphous cellulose along with concomitant ethanol production (Den Haan et al., 2007). The heterologous expression of exoglucanases are pivotal for *S. cerevisiae* to grow on the cellulose found in lignocellulosic materials as they are known to disrupt the crystalline nature of cellulose. The exoglucanases expressed in *S. cerevisiae* have thus far shown disappointingly low titers (Mertz et al., 2005). Several aspects that ranges from hyperglycosylation to codon-usage have been given as possible reasons for the low titers obtained.

The glycoside hydrolase family 6 (GH6) comprises a relatively diverse group of cellulases including both endoglucanases and cellobiohydrolases derived from a wide range of species that includes both eukaryotes and prokaryotes (Mertz et al., 2005). Phylogenetic analyses have grouped all the GH6 cellulases into 8 subfamilies. The cellobiohydrolase Cel6A of *Neocallimastix patriciarum* (NpCel6A) is in subfamily 3 that not only comprises of chytridiomycotal cellulases, but also includes endoglucanases. It is also clear from the

phylogenetic data that the members of subfamily 3 have C- and N-terminal loops slightly shorter than those found in members of subfamilies 4 to 8 – all representing cellobiohydrolases, but C-terminal loops are longer than those members of subfamily 1 and 2 – all representing endoglucanases. A noteworthy attribute of NpCel6A is that it has the highest individual specific activity on Avicel of any cellulase reported thus far (Denman et al., 1996). Another interesting feature of NpCel6A is its asparagine-rich linker domain that connects the carbohydrate binding module to the catalytic domain. It was recently shown that exchanging linkers and CBMs of exocellulases greatly impacted on their catalytic activities (Voutilainen et al., 2008).

In this study, the gene encoding NpCel6A was heterologously produced in *S. cerevisiae*. Attempts were made to improve the overall cellulolytic capability of the yeast producing NpCel6A by means of replacing its linker region.

4.3. MATERIALS AND METHODS

4.3.1. MEDIA AND CULTURE CONDITIONS

All chemicals and media components used in this study were of laboratory grade standard. *Escherichia coli* DH5α was used for initial plasmid construction and propagation. The laboratory strains *S. cerevisiae* Y294 and oversecretor NI-C-D4 were used for all yeast experiments. *E. coli* cells were cultivated in LB medium (5 g/L yeast extract, 10 g/L NaCl, 10 g/L tryptone) supplemented with ampicillin (100 mg/L). The genotypes of all the strains constructed and used in this study are shown in Table 4.1. *S. cerevisiae* Y294 and NI-C-D4 transformants were selected and maintained on synthetic complete (SC) medium plates (1.7 g/L yeast nitrogen base without amino acids and ammonium sulphate [Difco laboratories, Detroit, MI, USA], 5 g/L $(NH_4)_2SO_4$, 20 g/L glucose, 20 g/L agar and supplemented with amino acids as required) with the pH adjusted to 6.0. Autoselective *S. cerevisiae* Y294 strains were cultured in YPD medium (10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose. For activity assays, *S. cerevisiae* cultures were either cultivated in 25 mL YPD or double-strength SC which was buffered containing 20 g/L

succinic acid with pH adjusted to 6.0 with NaOH without uracil in 125 mL erlenmeyer flasks at 30°C on a rotary shaker at 100 rpm.

4.3.2. PLASMID CONSTRUCTION

Standard procedures for DNA manipulation were followed as prescribed by Sambrook & Russel (2001). All restriction DNA endonuclease enzymes were purchased from Fermentas (Vilnius, Lithuania). For polymerase chain reactions, the High Fidelity PCR Enzyme Mix (Fermentas) was used as recommended by the manufacturer with a Perkin Elmer GeneAmp® PCR System 2400 (The Perkin-Elmer Corporation, Norwalk, CT, USA). Details of the primers used in this study are given in Table 4.2.

TABLE 4.1: Microbial strains used in this study.

Plasmids propagated in *E. coli* DH5α	**Abbreviated name**	**Relevant genotype**	**Source/Reference**
Escherichia coli DH5α		*supE44 ΔlacU169 (ϕ80lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	Sambrook & Russel, 2001
pUC-57-NpCel6A		*bla NpCel6A* (native secretion signal replaced)	This work
pJC1		*bla* $PGK1_P$-$PGK1_T$	Crous et al., 1995
yXYNSEC		*bla* $PGK1_P$-*XYNSEC-* $PGK1_T$	Van Rooyen et al., 2005
pNpCel6A		*bla* $PGK1_P$-*XYNSEC-Npcel6A-* $PGK1_T$	This work
pAZ21		*bla* $ENO1_P$-*XYNSEC-Trcel6A-* $ENO1_T$	Den Haan et al., 2007b
pJET-TrNCBM-NpCel6A		*bla XYNSEC-CBM+linkerTrCel6A-catNpCel6A*	This work
pJET-NpNCBM-TrCel6A		*bla XYNSEC-CBM+linkerNpCel6A-catTrCel6A*	This work
pNCBM-NpCel6A		*bla* $PGK1_P$-*XYNSEC-CBM+linkerTrCel6A-catNpCel6A-* $PGK1_T$	This work
pNCBM-TrCel6A		*bla* $PGK1_P$-*XYNSEC-CBM+linkerNpCel6A-catTrCel6A-* $PGK1_T$	This work
pDF1		*bla fur1::LEU2*	La Grange et al., 1996
Yeast strains constructed			
Saccharomyces cerevisiae Y294		α *leu2, ura3, his3, trp1*	ATCC 201160
fur1::LEU2 yXYNSEC	Y294[REF]	*bla ura3/URA3* $PGK1_P$-*XYNSEC-* $PGK1_T$	Van Rooyen et al., 2005
fur1::LEU2 pNpCel6A	Y294[NpCEL6A]	*bla ura3/URA3* $PGK1_P$-*XYNSEC-Npcel6A-* $PGK1_T$	This work
fur1::LEU2 pAZ21	Y294[TrCEL6A]	*bla ura3/URA3* $ENO1_P$-*XYNSEC-Trcel6A-* $ENO1_T$	Den Haan et al., 2007b
fur1::LEU2 pTrNCBM-NpCel6A	Y294[TrNCBM-NpCEL6A]	*bla ura3/URA3* $PGK1_P$-*XYNSEC-CBM+linkerTrCel6A-catNpCel6A-*$PGK1_T$	This work
fur1::LEU2 pNpNCBM-TrCel6A	Y294[NpNCBM-TrCEL6A]	*bla ura3/URA3* $PGK1_P$-*XYNSEC-CBM+linkerNpCel6A-catTrCel6A-* $PGK1_T$	This work
Saccharomyces cerevisiae NI-C-D4		α *trp1, ura3, pep4*	Wang et al., 2001
pNpCel6A	NI-C-D4 [NpCEL6A]	*bla ura3/URA3* $PGK1_P$-*XYNSEC-Npcel6A-* $PGK1_T$	This work

The gene sequence of *NpCel6A* was codon-optimised using the Codon Adaptation Index (CAI) calculator for expression in *S. cerevisiae* (Carbone et al., 2003) with a CAI value of 0.932. The native secretion signal was replaced with the secretion signal from *Xyn11B* of *T. reesei*. The complete gene sequence including *Eco*RI and *Xho*I restriction endonuclease sites were sent for de novo gene synthesis to GenScript Corporation (Piscataway, NJ, USA). The pUC57 vector containing the synthetic 1227-bp *NpCel6A* gene was digested with *Eco*RI and *Xho*I and ligated into the *E. coli*/*S. cerevisiae* shuttle vector yXYNSEC containing the phosphoglycerate kinase 1 (*PGK1*) promoter and terminator regions with the resulting plasmid designated pNpCel6A.

TABLE 4.2: Primers used in this study. Sites for restriction enzymes are in boldface.

Primer name	Oligonucleotide sequence (5'-3')	Relevant restriction sites	Genbank accession nr
ENO1-L	**GGATCC**ACTAGTCTTCTAGGCGGGTTATC	*Eco*RI	X99228
ENO1-R	**AAGCTT**GCGGCCGCAAAGAGGTTTAGACATTGG	*Hin*dIII	
FUR1-L	TCCGTCTGGCATATCCTA		M36485
FUR1-R	TTGGCTAGAGGACATGTA		
Npcel6A-L	**GAATTC**CAACATGGTTTCCT	*Eco*RI	ACC49315
Npcel6A-R	**CTCGAG**TTAGAAGGATGGTC	*Xho*I	
Trcel6A-L	**GAATTC**ATAATGGTCTCCTTC	*Eco*RI	CAA49757
Trcel6A-R	CAGT**CTCGAG**TTACAAGAAAGATGGGTTAGC	*Xho*I	
TrCBMlinkerNpcat	GGCCATGACGATGAATGAGAAGGCCAAGATGGTTC		
NpCBMlinkerTrcat	CCGACGAAAGGGTTACCACCTCCGTTGTTGTTGTTGTTG		

For the domain swapping experiments, the overlap primer TrCBMlinkerNpcat was used along with Trcel6A-L to amplify the N-terminal secretion signal, carbohydrate-binding module (CBM) and linker region of the *Trcel6A* gene of *Trichoderma reesei*. The resulting PCR-product was used along with Npcel6A-R to fuse the abovementioned N-terminal region to the catalytic domain of *cel6A* of *N. patriciarum*. The resulting PCR fragment was ligated into the pJET1.2/blunt vector and designated pJET-TrCBMlinkerNpcat. The plasmid was digested with *Eco*RI and *Xho*I and the 1365-bp fragment was ligated into a *Eco*RI-and-*Xho*I-digested pJC1 called pTrCBMlinkerNpcat which was transformed into yeast. The overlap primer NpCBMlinkerTrcat was used with Npcel6A-L to amplify the N-terminal secretion signal, carbohydrate-binding module (CBM) and linker region of *Npcel6A* gene of *N. patriciarum*. The resulting PCR-product was used along with Trcel6A-R to fuse the abovementioned N-terminal region to the catalytic domain of *cel6A* of *T. reesei*. The resulting PCR fragment was ligated into the pJET1.2/blunt vector and designated pJET-NpCBMlinkerTrcat.

The plasmid was digested with *Eco*RI and *Xho*I and the 1432-bp fragment was ligated into a *Eco*RI-and *Xho*I-digested pJC1 called pNpCBMlinkerTrcat which was transformed into yeast. Diagrammatic representation of hybrid constructs is shown in Figure 4.1.

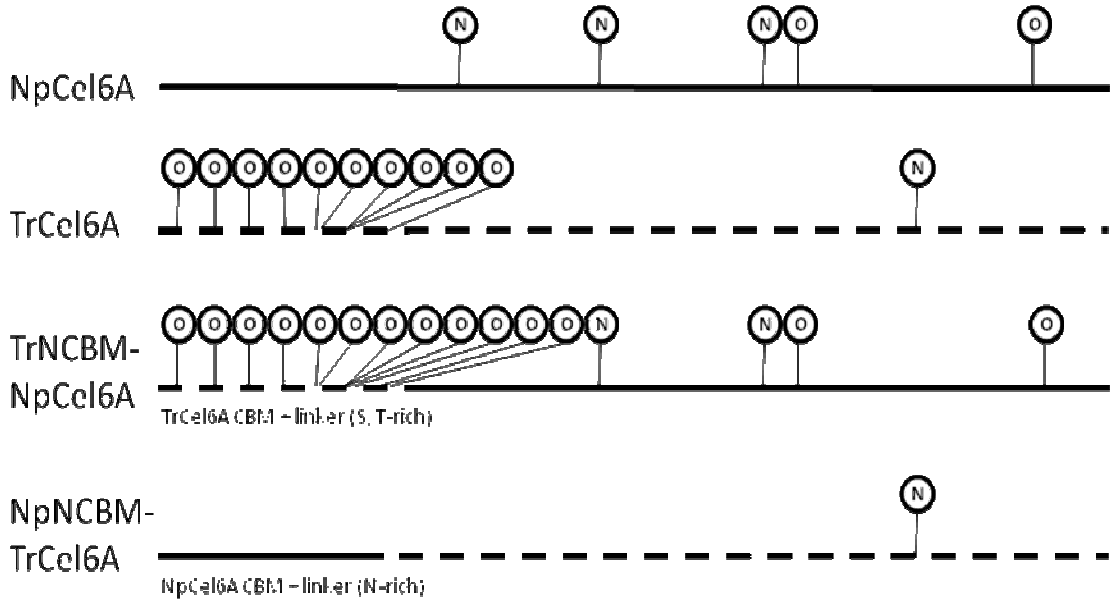


FIGURE 4.1: Schematic representation of hybrid gene constructs generated in this study. Potential *N*- and *O*-glycosylation sites are shown on each sequence with a likeliness of being glycosylated of over 60%.

Verification of relevant sequences was conducted after sequence determination by the dideoxy chain termination method, using the ABI PRISM^{TM} 3100 Genetic Analyzer. Analyses of sequences were mainly done using DNAMAN (version 4.1) software from Lynnon BioSoft and the BLAST program at the National Center for Biotechnology Information (www.ncbi.nih.gov/BLAST/).

4.3.3. YEAST TRANSFORMATIONS

S. cerevisiae Y294 and NI-C-D4 strains were transformed using the LiOAc/DMSO protocol described (Hill et al., 1991). For the construction of autoselective strains, the native *FUR1* gene was disrupted using pDF1 (La Grange et al., 1996) that was linearized with *Nco*I and *Ava*III. Double cross-over transformants were verified with FUR1-L and FUR1-R primers.

4.3.4. HIGH-PERFORMANCE ANIONIC EXCHANGE CHROMATOGRAPHY

Concentrated enzyme solutions were made by adding 130 mL of the supernatant of Y294[REF], Y294[NpCEL6A] or Y294[TrCEL6A] grown for four days in YPD to 2.5% Avicel. The enzyme-substrate mixtures were stirred overnight at 4°C. Subsequently, the enzyme-substrate mixture was filtered through a glass fiber filter (Whatman, Kent, UK) and the cellulose-bound enzyme that remained on the filter was resuspended in six mL of 0.05 M citrate buffer pH 5.0. Samples were then taken of the concentrated enzyme-substrate solutions that were incubated at 37°C on an orbital shaker at time period 0, 2, 5 and 24 h. These samples were analysed for glucose and cellobiose content by high-performance anion exchange chromatography (HPAEC) with a Dionex DX 500 chromatography system coupled to a pulsed amperometric detector (PAD, Dionex ED 40) using a CarboPac PA-10 column (Dionex, Sunnyvale, CA, USA). The column was initially activated with 20 mM NaOH. A flow-rate was maintained at 1 mL/min. After 25 min the NaOH concentration was increased to 225 mM for 6 min and then decreased to 20 mM. External standards of glucose (Saarchem, Krugersdorp, South Africa), cellobiose (Sigma, St. Louis, MO, USA) and cellotriose (Fluka) were used for quantification.

4.3.5. SDS-PAGE AND ZYMOGRAM ANALYSIS

Yeast cultures grown for four days in YPD were centrifuged, filtered through a glass fiber filter and concentrated approximately ten-fold on the Diaflo Ultrafilter (Amicon Inc, Beverly, MA, USA) before being analyzed by SDS-PAGE. Boiled samples were run under non-reducing conditions on a 10% polyacrylamide gel. After electrophoresis at 120 V, gels were silver-stained according to O'Connell & Stults (1997). For zymogram analyses, boiled and unboiled samples were run under non-reducing conditions on a 10% polyacrylamide gel containing 0.3% lichenan. After electrophoresis, gels were washed 4 x 30 min with 0.05 M citrate buffer pH 5.0 containing 0.05% DTT at room temperature, whereafter gels were incubated at 40°C in 0.05 M citrate buffer pH 5.0. Gels were subsequently flooded with 0.1% Congo Red and destained with 1 M NaCl.

4.3.6. PLATE AND LIQUID CELLULASE ACTIVITY ASSAYS

Yeast strains were spotted on SC media with YNB and $(NH_4)_2SO_4$ with 0.1% lichenan from Icelandic moss (Sigma) and incubated at 30°C for three days. Plates were then washed to remove the yeast colonies and flooded with 0.1% Congo Red for 30 min followed by destaining with 1 M NaCl. For quantitative endoglucanase liquid assays, a scaled-down version of the DNS method was used to determine the amount of reducing sugars generated according to Miller et al. (1960): 50 µL of a dilution of yeast supernatant were added to 450 µL of a 0.1% lichenan solution in 0.05 M citrate buffer pH 5.0, vortexed and incubated in a waterbath at 40°C for five min. Subsequently 750 µL of a DNS (1% 3,5-dinitrosalicyclic acid, 20% potassium sodium tartrate, 1% NaOH, 0.2% phenol, 0.05% Na_2SO_3) solution was added, vortexed, boiled for 15 min and allowed to cool down before optical densities were measured at 540 nm.

To determine exoglucanase activity, 300 µL of the yeast culture supernatant was added to deep-well microtiter plates with each well containing a 300 µL solution with 2% Avicel PH-105, 0.1 M acetate buffer pH 5.0, 0.04% NaN_3 and 0.3 µL Novozyme-188 (Sigma). The deep-well plate was sealed and incubated on a microtiter plate shaker and shaken at ~1000 rpm for 24 h at 50°C. One hundred microliter samples of the enzyme-substrate mixture were taken at the 0 h and 24 h time intervals to determine background sugars that were present and the amount of sugars released respectively, using a modified DNS method (Miller et al., 1960). The samples were transferred to a 96-well PCR plate using a multi-channel pipette and were centrifuged at 2000 rpm for two min. Fifty microliters of the supernatant were subsequently pipetted along with 100 µL DNS solution into a clean 96-well PCR plate, sealed, heated at 99°C for five min and cooled at 4°C for one min. Absorbance values at 540 nm were read on a X-mark™ microtitre plate reader (Biorad, Hercules, CA, USA).

Filter paper activity assays were conducted according to procedures recommended by the National Renewable Energy Laboratory (Denver, CO, USA) (Adney & Baker, 1996). One by six cm Whatman nr 1 filter paper strips were immersed in one mL 0.05 M citrate buffer pH 4.8 in test tubes and placed in a water bath at 50°C for five minutes. Five-hundred microlitres of the

yeast culture supernatants, appropriately diluted in citrate buffer, were added and incubated at 50°C for 60 min. Three mL of DNS solutions were added, vortexed and boiled for five minutes, allowed to cool down and spectrophotometrically measured at 540 nm. Glucose was used to draw a standard curve from which the amount of Units per milliliter (U/mL) of each sample was calculated. A Unit is defined as the amount of enzyme required to release one µmol of reducing sugar released per minute. All values obtained were normalized with the dry cell weight (in mg/mL) of the yeast cultures from which the samples were taken (Den Haan et al., 1996).

4.4. RESULTS

4.4.1. CLONING AND EXPRESSION OF *N. PATRICIARUM*, *T. REESEI* AND HYBRID GENES IN *S. CEREVISIAE*

The synthetic, codon-optimized *N. patriciarum cel6A* gene sequence (*Npcel6A*), with its secretion signal replaced with that of the xylanase II gene (*xyn11B*) of *T. reesei,* was expressed under the transcriptional control of the constitutive *S. cerevisiae* promoter of the phosphoglycerate kinase 1 (*PGK1*) gene. The *cel6A* gene of *T. reesei* (*Trcel6A*) with its native secretion signal also replaced with that of *xyn11B* was expressed under the transcriptional control of the constitutive promoter of Enolase 1 (*ENO1*). The hybrid genes – TrCBM-NpCel6A and NpCBM-TrCel6A – generated in the study were expressed under the *PGK1* promoter. Confirmation of the insertion of heterologous cellulase genes in yeast transformants were done with PCR analysis on total yeast DNA isolates using primers that annealed to the cellulase genes (results not shown).

4.4.2. SDS-PAGE AND ZYMOGRAM ANALYSIS

A polydisperse smear at the approximate size of NpCel6A (~47 kDa) was detected in the supernatant of Y294[NpCEL6A] cultivated in YPD, but not in that of lane Y294[REF] (Figure 4.2). The smear indicates possible glycosylation of the protein. Due to the interferance of a yeast protein of 50-57 kDa, the extent of the smear is not clear. The zymogram with lichenan as

substrate (Figure 4.2) showed unboiled NpCel6A with a small clearing zone corresponding to the molecular weight of around 47 kDa.

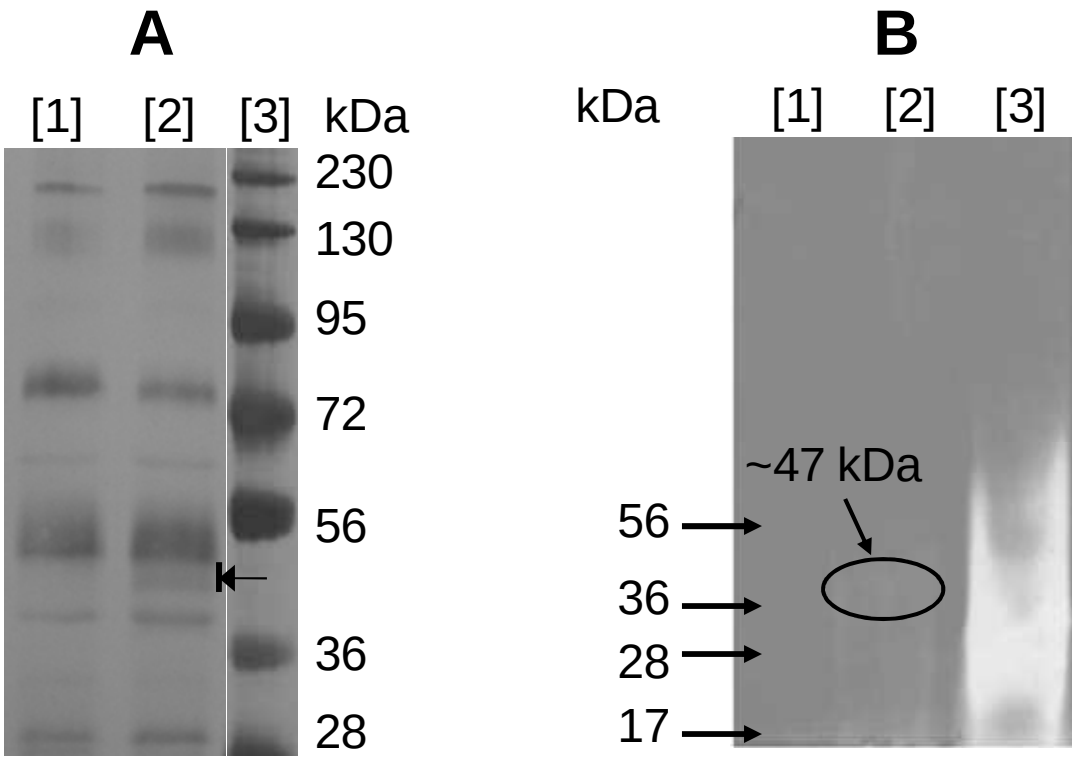


FIGURE 4.2: A: Silver-stained SDS-PAGE of Y294[NpCel6A] supernatant compared to Y294[REF]. Lane 1 contains 20 µL of supernatant of the reference strain Y294[REF]. Lane 2 contains 20 µL of supernatant of Y294[NpCel6A]. Lane 3 contains the Page Ruler^{TM} Prestained protein Ladder Plus (Fermentas). The arrow potentially indicates a polydispersed NpCel6A protein due to glycosylation. B: Zymogram of (boiled – lane 2 – and unboiled – lane 3) samples of the supernatant of Y294[NpCel6A]. Unboiled supernatant of Y294[REF] in in lane 1.

4.4.3. HPAEC

HPAEC (Figure 4.3) revealed that the yeast produced recombinant NpCel6A released glucose, cellobiose and cellotriose from Avicel at a ratio of 1:5.6:1.3. In contrast, yeast produced TrCel6A released glucose, cellobiose and cellotriose from Avicel at a ratio of 1:24.4:1. A chromatogram of Avicel treated with negative control supernatant showed no detectable peaks.

4.4.4. ACTIVITY ASSAYS OF RECOMBINANT YEAST STRAINS

Y294[NpCEL6A] showed clearing halos on agar plates containing lichenan – a β-linked 1-3,1-4 polysaccharide (Figure 4.4) confirming the earlier report of Npcel6A having activity on this

substrate (Denman et al., 1996). NpCel6A also showed a clearing halo on Ostazin-linked-hydroxyethylcellulose which also corresponded to previous data of having activity on another substituted cellulose substrate i.e. carboxymethylcellulose (Denman et al., 1996). Y294[TrCEL6A] did not show activity on either soluble substrates and thus liquid assays on this substrate were not conducted. Despite the proof of enzyme activity on Avicel in the HPAEC experiments, little activity was detected for Y294[NpCEL6A] in liquid assays (Table 4.3) compared to Y294[TrCEL6A]. Activity assays on filter paper, which are often used to determine total cellulolytic activity and not just exoglucanase activity, were chosen for comparisons. Y294[NpCEL6A] showed significantly more activity on filter paper compared to TrCel6A (Table 4.3). With regards to the normalized activity of NpCel6A on lichenan and Avicel, the activity obtained on lichenan was over 100 000-fold more in relation to what was obtained on Avicel. This did not correlate to the specific activities on NpCel6A on the respective substrates where the ratio was only 35-fold (Denman et al., 1996). NpCel6A produced by the NI-C-D4 strain, deficient in *N*-glycosylation (NI-C-D4[NpCEL6A]), showed less activity on both soluble and insoluble cellulose substrates albeit not dramatic.

4.4.5. DOMAIN SWAPPING

To investigate what effect the unusual Asn-rich linker could have on the overall heterologous production of NpCel6A, it was, along with the family 1 CBM of NpCel6A, replaced with the family 1 CBM and linker region of TrCel6A by means of overlap-PCR. The strain generated - Y294[TrNCBM-NpCel6A] - showed a 60% decrease in filter paper-hydrolyzing activity than Y294[NpCel6A] (Table 4.3). The original protein NpCel6A and the hybrid protein TrNCBM-NpCel6A have three and two likely *N*-glycosylation sites (Asn-Xaa-Ser/Thr) respectively (Figure 4.1) according to NetNGlc 1.0 software, but differ with regard to their *O*-glycosylation. According to NetOGlc 1.0 software, TrNCBM-NpCel6A have 14 potential *O*-glycosylation sites on its Ser or Thr residues with a likeliness of over 60% whereas NpCel6A have only two.

The unusual Asn-rich linker and CBM of NpCel6A was then fused to the TrCel6A catalytic domain with the resulting yeast strain designated Y294[NpNCBM-TrCEL6A]. An approximate

two-fold increase in filter paper activity was observed with Y294[NpNCBM-TrCEL6A] in comparison to Y294[TrCEL6A] (Table 4.3). Both the hybrid protein NpNCBM-TrCel6A and TrCel6A contain one potential *N*-glycosylation site (Figure 4.1), but TrCel6A contains ten potential *O*-glycosylation sites with a likeliness of higher than 60% whereas NpNCBM-TrCel6A contains none.

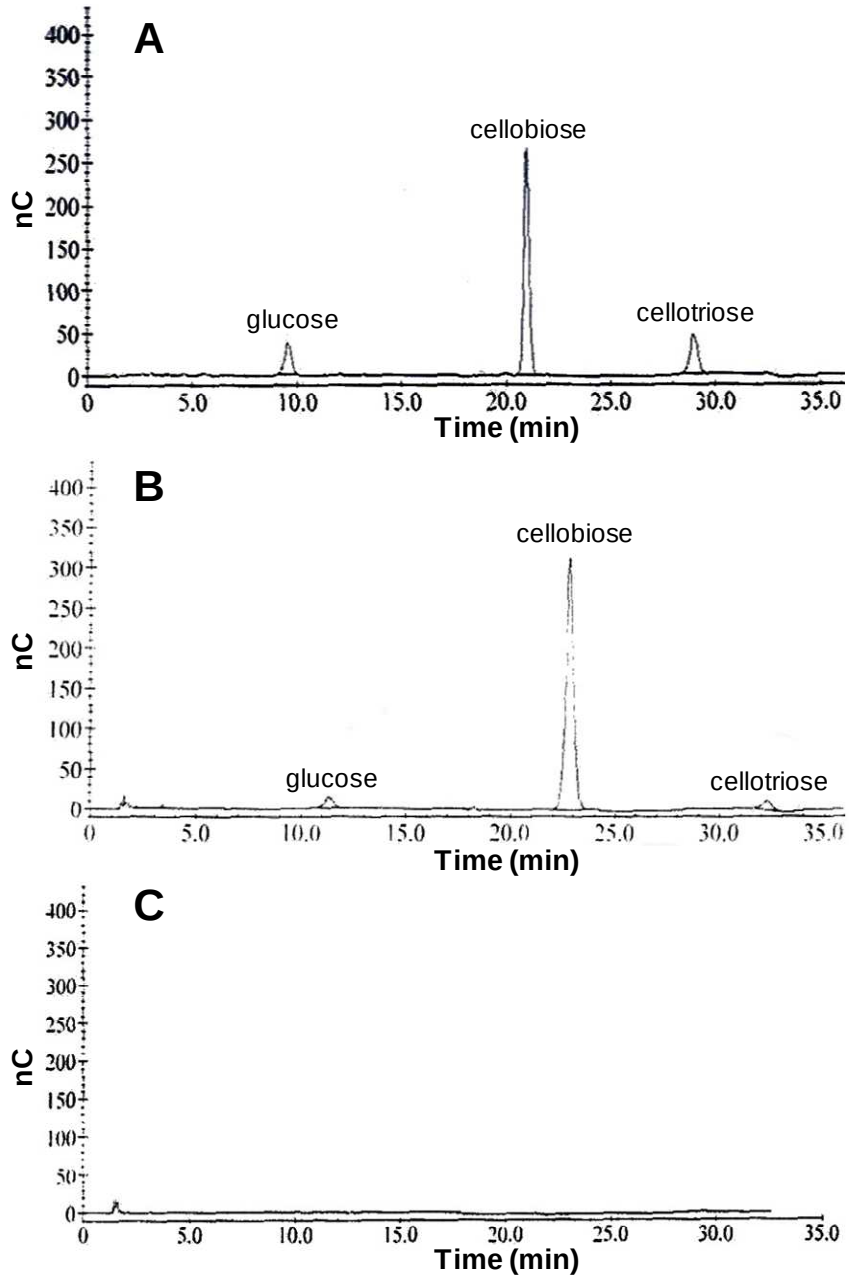


FIGURE 4.3: High anionic exchange chromatogram of the soluble sugars released from Avicel when treated with the supernatant of (A) Y294[NpCel6A], (B) Y294[TrCel6A] and (C) Y294[REF] after two hours.

4.5. DISCUSSION

NpCel6A was previously claimed to possess the highest individual specific activity on Avicel (Denman et al., 1996). Although some endoglucanases show activity on Avicel, it is generally considered as a substrate for exoglucanases. The fact that NpCel6A showed pronounced activity on substituted amorphous cellulose substrates is particularly unusual as the processive nature of most exoglucanases are blocked by substitutions on the cellulose chain due to its tunnel-forming loops. Thus even though formally classified as a cellobiohydrolase, NpCel6A displayed definite endoglucanase-like behaviour. The apparent overglycosylation of the recombinantly produced protein dramatically compromised its processive capability hence the barely detectable activity on crystalline cellulose. Due to the little difference in cellulase activities between the heterologous production of NpCel6A by Y294 and the *N*-glycosylation deficient strain NI-C-D4, it is deduced that *O*-glycosylation in the catalytic domain is responsible for the compromised processive capability. It is presumed that *O*-glycosylation as shown with the polydisperse smear on the protein gel interfered with the tunnel-froming loops that prevent non-reducing ends of the cellulose chain to enter. The glycosylation did not seem to affect the endoglucanase activity of the catalytic domain.

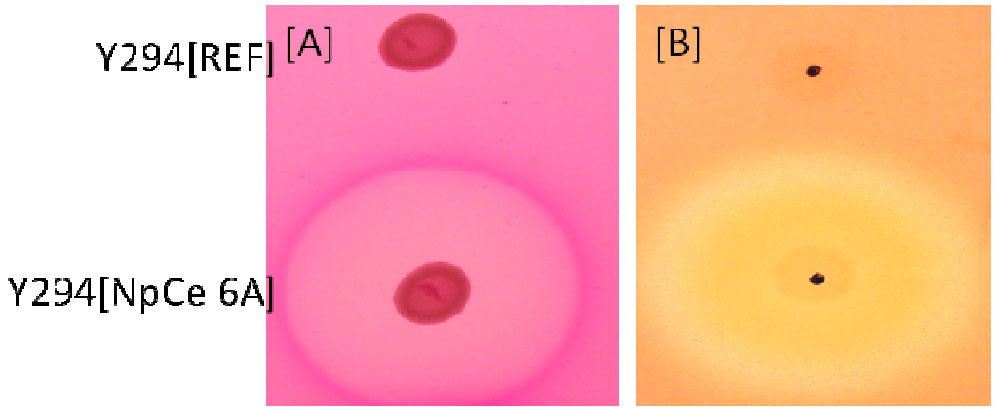


FIGURE 4.4: Plate assays on soluble cellulose substrates. The two recombinant yeast strains were spotted on SC media containing 1% Ostazin-linked hydroxyethylcellulose [A] and 0.1% lichenan [B] respectively and allowed to grow for two days at 30°C.

An unusual feature of NpCel6A is the large percentage of asparagine (Asn) residues found in the linker region between the catalytic domain and the carbohydrate-binding module (CBM). An Asn-rich linker is however not unique for NpCel6A and is common among glycoside hydrolases

of rumen organisms. Even though TrCel6A and NpCel6A share a high level of identity in their catalytic domains and CBMs, their respective linker regions are completely different: TrCel6A have a Ser/Thr/Pro-rich linker common among many linker regions of glycoside hydrolases found in aerobic fungi. It was initially hypothesized that the heterologous NpCel6A production by *S. cerevisiae* might be impeded by the incorporation of a series of Asn-residues since highly transcribed genes in *S. cerevisiae* are often depleted in Asn-codons (Jansen & Gerstein, 2000). This was, however, not proven true since replacement of the Asn-rich linker resulted in a decrease in the cellulase activity of the recombinant yeast strain. Furthermore, when the Asn-rich linker was appended in front of the catalytic domain of TrCel6A, a marked increase in cellulase activity was observed. A possible explanation for the observations could be that the difference in *O*-glycosylation of the Ser/Thr-rich linker of TrCel6A impedes the overall production of the heterologous proteins although the effect of the different CBMs cannot be excluded. When *Trcel6A* was expressed in *pmt* mutants of *S. cerevisiae* it was shown that *O*-glycosylation is not necessary for heterologous production of the protein (Górka-Nieć et al., 2007).

TABLE 4.3: Activity assays of recombinant yeast supernatants when grown on double-strength SC on various cellulose substrates*. Plus/minus signs indicate the standard variation from the mean of three repeats.

	Lichenan (Units/mg DCW)	Filter paper (mUnits/mg DCW)	Avicel (mUnits/mg DCW)	Lichenan: Avicel activity ratio
Y294[NpCEL6A]	99.622±8.901	25.120±2.244	0.540±0.062	184 485
NI-C-D4[NpCEL6A]	77.346±6.828	ND	0.353±0.071	ND
Y294[TrCEL6A]	ND	16.532±0.466	4.101±0.243	ND
Y294[TrNCBM-NpCEL6A]	48.516±4.022	15.329±1.570	ND	ND
Y294[NpNCBM-TrCEL6A]	ND	38.444±3.471	ND	ND
Specific activity (Unit/mg protein)†	343.5		9.7	35.4

*Values obtained for reference strain were subtracted.

†Denman et al., 1996

ND not determined

In this study, we attempted to construct a recombinant *S. cerevisiae* strain with significant Avicel-hydrolyzing ability by heterologously producing a cellobiohydrolase namely *N. patriciarum cel6A* (*NpCel6A*). The recombinant yeast displayed disappointingly low Avicel-

hydrolyzing ability, compared to a strain that heterologously produced *cel6A* of *T. reesei* (*TrCel6A*). This was mainly attributed to glycosylation of the catalytic domain. By switching domains including the linker regions of TrCel6A and NpCel6A, it was shown that their respective production by *S. cerevisiae* could be dramatically altered with a preference for the Asn-rich linker.

4.6. LITERATURE CITED

Adney B, Baker J 1996. Measurement of cellulase activities. In: Labaratories analytical procedures: National Renewable Energy Laboratories. pp 1-8.

Carbone A, Zinovyev A, Képès F 2003. Codon adaptation index as a measure of dominating codon bias. Bioinformatics 19: 2005-2015.

Coutinho PM, Henrissat B 1999. The modular structure of cellulases and other carbohydrate-active enzymes: an integrated database approach. In: Genetics, biochemistry and ecology of cellulose degradation (K. Ohmiya K, Hayashi K, Sakka K, Kobayashi Y, Karita S, Kimura T eds.) Uni Publishers Co, Tokyo. Japan. pp 15-23.

Crous JM, Pretorius IS, Van Zyl WH 1995. Cloning and expression of an *Aspergillus kawachii* endo-1,4-β-xylanase gene in *Saccharomyces cerevisiae*. Curr Genet 28: 467-473.

Den Haan R, Rose SH, Lynd LR, Van Zyl WH 2007a. Hydrolysis and fermentation of amorphous cellulose by recombinant *Saccharomyces cerevisiae*. Metab Eng 9: 87-94.

Den Haan R, Mcbride JE, La Grange DC, Lynd LR, Van Zyl WH 2007b. Functional expression of cellobiohydrolases in *Saccharomyces cerevisiae* towards one-step conversion of cellulose to ethanol. Enzyme Microb Technol 40: 1291-1299.

Denman S, Xue GP, Patel B 1996. Characterization of a *Neocallimastix patriciarum* cellulase cDNA (celA) homologous to *Trichoderma reesei* cellobiohydrolase II. Appl Environ Microbiol 62: 1889-1896.

Górka-Nieć W, Bańkowska R, Palamarczyk G, Krotkiewski H, Kruszewska JS 2007. Protein glycosylation in *pmt* mutants of *Saccharomyces cerevisiae*. Influence of heterologously expressed cellobiohydrolase II of *Trichoderma reesei* and elevated levels of GDP-mannose and *cis* prenyltransferase activity. BBA Gen Subjects 1770: 774-780.

Hill J, Donald KA, Griffiths DE 1991. DMSO-enhanced whole cell yeast transformation. Nucleic Acids Res 19: 5791.

Imperiali B, O'Connor SE 1999. Effect of N-linked glycosylation on glycopeptide and glycoprotein structure. Curr Opin Chem Biol 3: 643-649.

Jansen R, Gerstein M 2000. Analysis of the yeast transcriptome with structural and functional categories: characterizing highly expressed proteins. Nucleic Acids Res 28: 1481-1488.

Jeoh T, Michener W, Himmel ME, Decker SR, Adney WS 2008. Implications of cellobiohydrolase glycosylation for use in biomass conversion. Biotechnology for Biofuels 1: 10-22.

La Grange DC, Pretorius IS, Van Zyl WH 1996. Expression of a *Trichoderma reesei* beta-xylanase gene (XYN2) in *Saccharomyces cerevisiae*. Appl Environ Microbiol 62: 1036-1044.

Lynd LR, Van Zyl WH, McBride JE, Laser M 2005. Consolidated bioprocessing of cellulosic biomass: an update. Curr Opin Biotechnol 16: 577-583.

Mertz B, Kuczenski RS, Larsen RT, Hill AD, Reilly PJ 2005. Phylogenetic analysis of family 6 glycoside hydrolases. Biopolymers 79: 197-206.

Miller GL, Blum R, Glennon WE, Burton AL 1960. Measurement of carboxymethyl cellulase activity. Anal Biochem 2: 127-132.

Nevalainen KMH, Te'o VSJ, Bergquist PL 2005. Heterologous protein expression in filamentous fungi. Trends Biotechnol 23: 468-474.

O'Connell K, Stults, J 1997. Identification of mouse liver proteins on two-dimensional electrophoresis gels by matrix-assisted laser desorption/ionization mass spectrometry of in situ enzymatic digests. Electrophoresis 18: 349-359.

Sambrook J, Russel DB 2001. Molecular cloning: A laboratory manual. Cold spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Stephanopoulos G 2007. Challenges in engineering microbes for biofuels production. Science 315: 801-803.

Van Rooyen R, Hahn-Hägerdal B, La Grange DC, Van Zyl WH 2005. Construction of cellobiose-growing and fermenting *Saccharomyces cerevisiae* strains. J Biotechnol 120: 284-295.

Voutilainen SP, Puranen T, Siika-aho M, Lappalainen A, Alapuranen M, Kallio J, Hooman S, Viikari L. Vehmaanperä J, Koivula A 2008. Cloning, expression, and characterization of novel thermostable family 7 cellobiohydrolases. Biotechnol and Bioeng 101: 515-528.

Wang B-D, Chen DC, Kuo TT 2001. Characterization of a *Saccharomyces cerevisiae* mutant with oversecretion phenotype Appl Microbiol Biotechnol 55: 712-720.

CHAPTER 5

OVEREXPRESSION OF NATIVE *SACCHAROMYCES CEREVISIAE* GENES FOR IMPROVEMENT OF HETEROLOGOUS PROTEIN PRODUCTION

Niël van Wyk, Heinrich Kroukamp[1], Riaan den Haan, Willem H. van Zyl

Department of Microbiology, University of Stellenbosch, Matieland 7602, South Africa.

5.1. ABSTRACT: Two native genes *PSE1* and *SOD1* were individually overexpressed in *Saccharomyces cerevisiae* by placing each gene under the transcriptional control of the constitutive *PGK1* promoter. Their effect on heterologous protein production of three cellulases – *cel6A* of *Neocallimastix patriciarum*, *cel3A* of *Saccharomycopsis fibuligera* and *cel7B* of *Trichoderma reesei* was investigated. In both cases, transformants were obtained that showed significantly higher heterologous protein activity that ranged between a 17% to 152% increase compared to the parental strains. When both *PSE1* and *SOD1* were overexpressed in the yeast that produces cel3A, a dramatic 420% increase in β-glucosidase activity was observed.

5.2. INTRODUCTION

Saccharomyces cerevisiae is seen as a platform organism that, with the aid of metabolic engineering and synthetic biology techniques, can be manipulated so that industrially important processes and pathways can be optimized (Lee et al., 2008). Successes in metabolic engineering of *S. cerevisiae* resulted in improved ethanol production (Bro et al., 2006) and substrate utilization (Roca et al., 2003). Protein secretion is often an important factor for heterologous protein production strategies. Several native proteins of *S. cerevisiae* have been identified that, when either disrupted or overexpressed, resulted in improved protein secretion (Chow et al., 1992; Lustgarten & Gerst, 1999).

[1] Mr H. Kroukamp made *PSE1* and *SOD1* overexpressing strains of Y294[EG1] and Y294[SFI] and conducted corresponding enzyme assays on yeast transformants as part of his HonsB.Sc. laboratory project.

Chow et al. (1992) reported enhanced secretion when the Protein secretion enhancer 1 protein (Pse1) of *S. cerevisiae* was overexpressed. The native gene cassette that encoded *PSE1* was placed on an episomal vector and resulted in higher levels of secreted killer toxin, α-factor and acid phosphatase. Pse1 is a member of the karyopherin family and has been implicated in the export of mRNA out of the nucleus to the cytosol (Seedorf & Silver, 1997).

Sod1 is a Cu-Zn superoxide dismutase which is involved in the detoxification of reactive oxygen species (ROS) in the cytosol and is also implicated in copper ion buffering (Culotta et al., 1995). Although the native *SOD1* has been overexpressed previously in *S. cerevisiae* (Harris et al., 2003), the effect it had on the phenotype of heterologous protein production was not investigated. The *SOD1* gene was recently overexpressed in *Kluyveromyces lactis* which resulted in improved secreted production of heterologous glucoamylase (Raimondi et al., 2008). It was postulated that the burden experienced by the yeast cell with the increased load of heterologous proteins resulted in the release of more ROS. The effect of an increased level of ROS could be relieved by the presence of more Sod1 protein in the cytosol.

In this study, the native *S. cerevisiae* genes *PSE1* and *SOD1* were each placed under the transcriptional control of the constitutive *PGK1* promoter and terminator and gene cassettes were integrated into the recombinant laboratory yeast strains Y294[NpCEL6A], Y294[SFI] and Y294[EGI] – which heterologously expressed *cel6A* of *Neocallimastix patriciarum*, *cel3A* of *Saccharomycopsis fibuligera* and *cel7B* of *Trichoderma reesei*, respectively (Van Wyk, 2010; Den Haan et al., 2007; Van Rooyen et al., 2005). Both *PSE1* and *SOD1* were also simultaneously overexpressed in Y294[SFI].

5.3. MATERIALS AND METHODS

5.3.1. RECOMBINANT YEAST STRAIN CONSTRUCTIONS

For the construction of *PSE1* and *SOD1* overexpressing strains, the open reading frames of the *PSE1* and *SOD1* genes of *Saccharomyces cerevisiae* Y294 were amplified using the primer sets PSE1-L/R and SOD1-L/R, respectively (Table 5.1). A 3270-bp PCR fragment for *PSE1* and a 489-

bp PCR fragment for *SOD1* were digested with *Asc*I and *Pac*I and ligated into the yeast expression vector pBKD1 – to yield pBKD1-PSE1 and pBKD1-SOD1. The primer set hph-L/R was used to amplify the hygromycin resistance gene (*hph*) from the plasmid pMU1037 (Mascoma Corporation, USA). The *Mun*I/*Spe*I digested *hph* gene cassette was ligated into the *Eco*RI/*Spe*I digested pBKD1_SOD1 and replaced the geneticin marker to generate pBHD_SOD1. The integration plasmids were linearized with *Xho*I after which transformation of Y294[NpCel6A], Y294[SFI] and Y294[EGI] were done under 200 µg/ml geneticin selection according to the LioAC/DMSO protocol (Hill et al., 1991). The transformant, Y294[SFI]PSE1 was also transformed with linearized pBHD_SOD1 under 100µg/ml hygromycin B selection. The total genomic DNA of each yeast transformant was isolated and successful integration of either the *SOD1* of *PSE1* overexpression cassette was confirmed with PCR analyses with PGK1term-R and SOD1-L or PSE1-L primers. Yeast transformants thus possessed the native copy of either *PSE1* or *SOD1* plus one or more integrated copies of the gene under constitutive transcriptional regulation. All yeast strains used and generated in this study are shown in Table 2.

TABLE 5.1. Primers used in this study. Restriction sites are shown in boldface.

Primer name	Oligonucleotide sequence (5'-3')	Restriction sites	Genbank accession nr
PSE1-L	GACT**TTAATTAA**ATGTCTGCTTTACCGGAAG	*Asc*I	CAA89141
PSE1-R	GCTA**GGCGCGCC**TTATGCAAACCATTTATG	*Pac*I	
SOD1-L	GACT**TTAATTAA**ATGGTTCAAGCAGTCGCAG	*Asc*I	CAA89634
SOD1-R	GCTA**GGCGCGCC**TTAGTTGGTTAGACCAATGACACC	*Pac*I	
PGK1term-R	CATAGAAATATCGAATGGGAA		NC001135
hph-L	CCTCATTCT**ACTAGT**GGATC	*SpeI*	AAA92252
hph-R	GATC**CAATTG**CAACCCTTAATATAACTTCG	*MunI*	

5.3.2. ASSAYS

All yeast transformants obtained were initially screened in 5 ml YPD-containing test tubes for improved heterologous protein production by determining enzyme activity of the yeast supernatant in liquid assays with lichenan, carboxymethylcellulose (CMC) and *p*-nitrophenyl-β-D-glucopyranoside (*p*NPG) as substrates for Y294[NpCEL6A], Y294[EGI] and Y294[SFI], respectively, using modified protocols as described previously (Van Wyk et al., 2010; Den Haan et al., 2007; Van Rooyen et al., 2005). All the volumetric values were normalized with the dry cell weight of the corresponding yeast cultures in mg/ml (Meinander et al., 1996). Of the up to

67 transformants screened, the ones with the highest normalized activity were assayed in triplicate and compared to the parental strain.

TABLE 5.2. The recombinant yeast strains used in this study.

Yeast Strains	Abbreviated name	Relevant genotype	Reference
Parental yeast strains:			
S. cerevisiae Y294:		α *leu2*-3,112 *ura3*-52 *his3 trp1*-289	ATCC 201160
(*fur1*::*LEU2* YEp352)	Y294[REF]	*bla ura3/URA3 PGK1$_P$-XYNSEC-PGK1$_T$*	Van Rooyen et al., (2005)
(*fur1*::*LEU2* pNpCel6A)	Y294[NpCEL6A]	*bla ura3/URA3 PGK1$_P$-XYNSEC-NpCel6A-PGK1$_T$*	Van Wyk et al., 2010)
(*fur1*::*LEU2* pAZ40)	Y294[EG1]	*bla ura3/URA3 ENO1$_P$-cel7B-ENO1$_T$*	Den Haan et al., (2007a)
(*fur1*::*LEU2* ySFI)	Y294[BGL1]	*bla ura3/URA3 PGK1$_P$-XYNSEC-cel3A-PGK1$_T$*	Van Rooyen et al., (2005)
Overexpression yeast strains:			
S. cerevisiae Y294 (*fur1::LEU2* pNpCel6A):			
Y294_*SOD1* overexpressed	Y294[NpCEL6A]SOD1	*bla ura3/URA3 PGK1$_P$-NpCel6A-PGK1$_T$ kanMX PGK1$_P$-SOD1-PGK1$_T$*	This work
Y294_*PSE1* overexpressed	Y294[NpCEL6A]PSE1	*bla ura3/URA3 PGK1$_P$-NpCel6A-PGK1$_T$ kanMX PGK1$_P$-SOD1-PGK1$_T$*	This work
S. cerevisiae Y294 (*fur1::LEU2* pAZ40):			
Y294_*SOD1* overexpressed	Y294[EG1]SOD1	*bla ura3/URA3 ENO1$_P$-cel7B-ENO1$_T$ kanMX PGK1$_P$-SOD1-PGK1$_T$*	This work
Y294_*PSE1* overexpressed	Y294[EG1]PSE1	*bla ura3/URA3 ENO1$_P$-cel7B-ENO1$_T$ kanMX PGK1$_P$-PSE1-PGK1$_T$*	This work
S. cerevisiae Y294 (*fur1::LEU2* ySFI):			
Y294_*PSA1* overexpressed	Y294[BGL1]PSA1	*bla ura3/URA3 PGK1$_P$-XYNSEC-cel3A-PGK1$_T$ kanMX PGK1$_P$-PSA1-PGK1$_T$*	This work
Y294_*SOD1* overexpressed	Y294[BGL1]SOD1	*bla ura3/URA3 PGK1$_P$-XYNSEC-cel3A-PGK1$_T$ kanMX PGK1$_P$-SOD1-PGK1$_T$*	This work
Y294_*PSE1*, Y294_*SOD1* overexpressed	Y294[BGL1-PSE1]SOD1	*bla ura3/URA3 PGK1$_P$-XYNSEC-cel3A-PGK1$_T$ kanMX PGK1$_P$-PSE1-PGK1$_T$ PGK1$_T$ hphMX PGK1$_P$-SOD1-PGK1$_T$*	This work

5.3.3. SDS-PAGE

Supernatants of yeast cultures grown for two days in YPD and subjected to denaturing SDS-PAGE according to the method described by Laemmli (1970). The volume loaded in each well was determined by normalizing with the optical density (OD_{600}) of the yeast culture. The supernatant proteins were visualized on the gel with silver-staining (O'Connell and Stults, 1997).

5.4. RESULTS AND DISCUSSION

As shown in Figure 5.1, transformants were obtained for all recombinant strains that showed improved secreted enzymatic activity when *PSE1* or *SOD1* were overexpressed. We assumed that increased secreted activity levels in transformants indicated enhanced levels of protein secretion as the specific activity of the recombinant enzymes should have remain unaltered. For *PSE1* overexpressing strains, the best transformants showed improvements of 99%, 152% and 17% compared to the respective Y294[NpCEL6A], Y294[SFI] and Y294[EGI] parental strains. In a previous study where *PSE1* was overexpressed, a four-fold increase in total protein secretion was obtained (Chow et al., 1992). With the *SOD1* overexpressing strains, the best transformants gave improvements of 56%, 47%, 37% compared to the parental strains. This is thus the first study that showed improved protein secretion by *S. cerevisiae* when *SOD1* was overexpressed. When the native *SOD1* was overexpressed in *Kluyveromyces lactis* an approximate four-fold increase in heterologous protein secretion of human serum albumin was observed (Raimondi et al., 2008). When *SOD1* and *PSE1* was overexpressed simultaneously a remarkable 420% increase in *β*-glucosidase activity was observed compared to the parental strain Y294[SFI]. This observation was substantiated with SDS-PAGE analysis which clearly shows a general increase in the amount of secreted proteins (Figure 5.2).

During initial screening, yeast transformants showed a wide range of enzyme activity of which some were significantly lower than the corresponding parental strain (results not shown). This was a clear example of clonal variation since the difference in copy number and location of integration of the expression cassette could greatly affect the phenotype of the yeast. Apart from the noticeable changes in enzyme activity, in the case where *PSE1* and *SOD1* were simultaneously overexpressed in the Y294[SFI] a dramatic decline in the cell density of the yeast was observed. This could imply a relative metabolic burden experienced by the yeast even though a significant increase in the volumetric enzyme activity was also observed.

For the cost-efficient production of ethanol from lignocellulosic biomass it is of great interest to construct *S. cerevisiae* strains capable of hydrolyzing cellulose without a large input of

additional cellulases (Van Zyl et al., 2007). Although many studies have shown recombinant cellulase production by *S. cerevisiae*, growth on a crystalline form of cellulose still requires at least a 100 fold increase in cellulase – particularly exoglucanases – production (Den Haan et al., 2007b). Here, we have shown that the recombinant production of the three enzymes needed for complete cellulose hydrolysis i.e. an exoglucanase (NpCel6A), an endoglucanase (Cel7B) and a β-glucosidase (Cel3A) could be significantly increased with the aid of metabolic engineering.

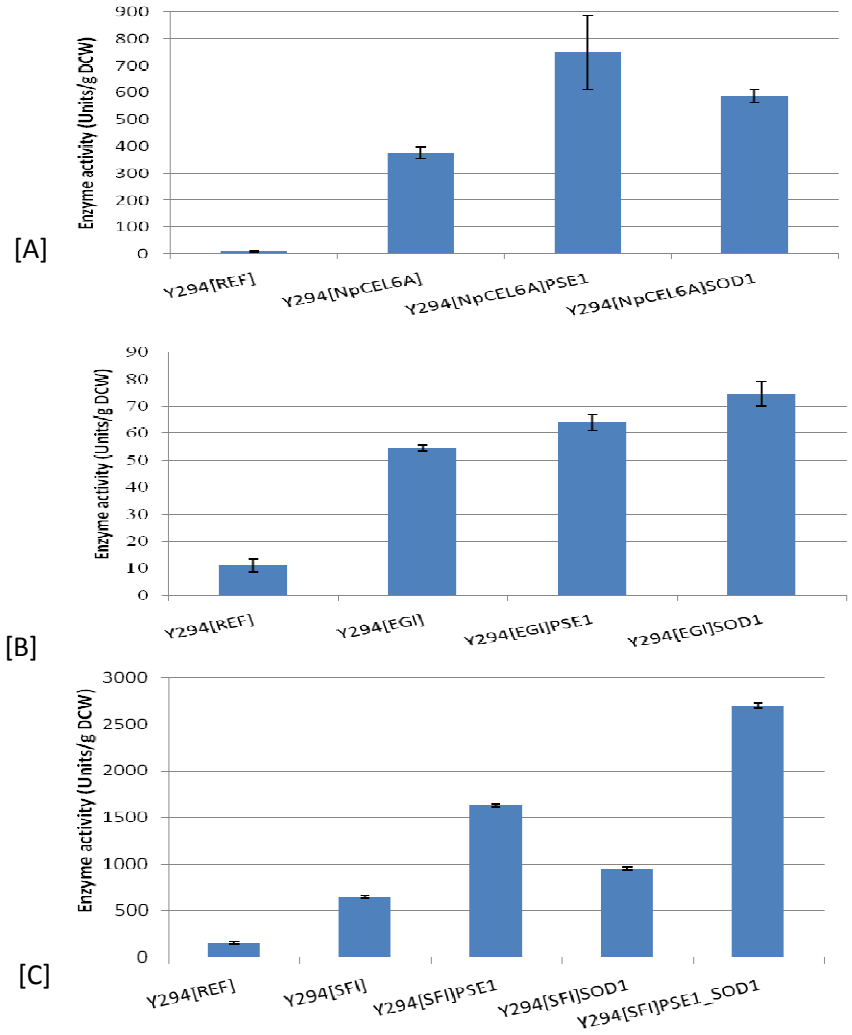


FIGURE 5.1: Enzyme activity profiles of *PSE1* and *SOD1* overexpression strains in comparison to parental strains. [A] shows Y294[NpCEL6A] with improved lichenase activity when *PSE1* and *SOD1* were individually overexpressed. Similarly results for Y294[EGI] on CMC and Y294[SFI] on *p*NPG are shown in [B] and [C] respectively. [C] also shows the simultaneous overexpression of *PSE1* and *SOD1*. A unit of enzyme per mL was defined as the amount needed to release one µmol of reducing sugar or equivalent per minute. Obtained values were normalized with the dry cell weight (DCW) (in g/mL) of the yeast. Values were obtained from independently conducted enzyme assays in triplicate. Error bars indicate standard deviation from the mean value.

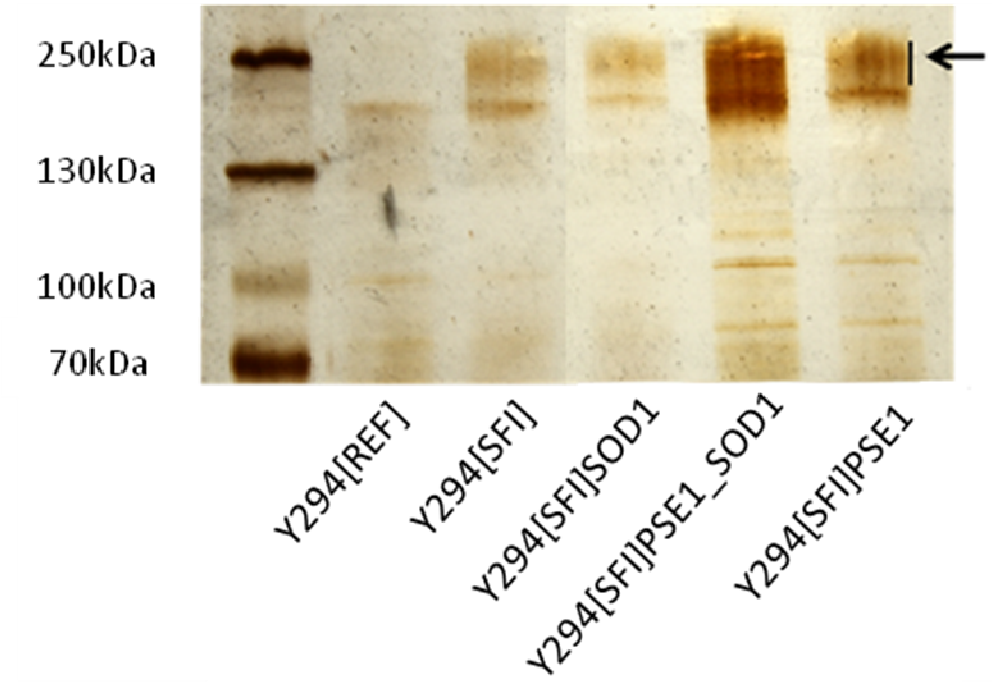


FIGURE 5.2: Photo of the silver-stained polyacrylamide gel with separated supernatant proteins and a molecular weight marker. Arrow indicates putative position of Cel3A.

This study shows the tremendous potential for metabolic engineering of *S. cerevisiae* to attain increased heterologous protein production. By overexpressing two different genes individually or simultaneously, it was shown that significant improvements in heterologous protein production could be achieved.

5.5. LITERATURE CITED

Bro C, Regenberg B, Forster J, Nielsen J 2006. In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. Metab Eng 8: 102-111.

Chow TY-K, Ash JJ, Dignard D, Thomas Y 1992. Screening and identification of a gene, *PSE-1,* that affects protein secretion in *Saccharomyces cerevisiae.* J Cell Sci 101: 709-719.

Culotta VC, Joh HD, Lin SJ, Slekar KH, Strain J 1995. A physiological role for *Saccharomyces cerevisiae* copper/zinc superoxide dismutase in copper buffering. J Biol Chem 270: 29991-29997.

Den Haan R, Rose SH, Lynd LR, Van Zyl WH 2007a. Hydrolysis and fermentation of amorphous cellulose by recombinant *Saccharomyces cerevisiae*. Metab Eng 9: 87-94.

Den Haan R, Mcbride JE, La Grange DC, Lynd LR, Van Zyl WH 2007b. Functional expression of cellobiohydrolases in *Saccharomyces cerevisiae* towards one-step conversion of cellulose to ethanol. Enzyme Microb Technol 40: 1291-1299.

Harris N, Bachler M, Costa V, Mollapour M, Moradas-Ferreira P, Piper PW 2005. Overexpressed Sod1p acts either to reduce or to increase the lifespans and stress resistance of yeast, depending on whether it is Cu^{2+}-deficient or an active Cu, Zn superoxide dismutase. Aging Cell 4: 41-52.

Hill J, Donald KA, Griffiths DE 1991. DMSO-enhanced whole cell yeast transformation. Nucleic Acids Res 19: 5791.

Laemmli, UK 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.

Lee SK, Chou H, Ham TS, Lee TS, Keasling JD 2008. Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. Curr Opin Biotechnol 19: 556-563.

Lustgarten V, Gerst JE 1999. Yeast *VSM1* encodes a v-SNARE binding protein that may act as a negative regulator of constitutive exocytosis. Mol Cell Biol 19: 4480-4494.

Meinander N, Zacchi G, Hahn-Hägerdal, B 1996. A heterologous reductase affects the redox balance of recombinant *Saccharomyces cerevisiae*. Microbiology 142: 165-172.

O'Connell K, Stults, J 1997. Identification of mouse liver proteins on two-dimensional electrophoresis gels by matrix-assisted laser desorption/ionization mass spectrometry of in situ enzymatic digests. Electrophoresis 18: 349-359.

Raimondi S, Zanni E, Talora C, Rossi M, Palleschi C, Uccelletti D 2008. *SOD1*, a new *Kluyveromyces lactis* helper gene for heterologous protein secretion. Appl Environ Microbiol 74: 7130-7137.

Roca C, Nielsen J, Olsson L 2003. Metabolic engineering of ammonium assimilation in xylose-fermenting *Saccharomyces cerevisiae* improves ethanol production. Appl Environ Microbiol 69: 4732-4736.

Seedorf M, Silver PA 1997. Importin/karyopherin protein family members required for mRNA export from the nucleus. PNAS 94: 8590-8595.

Van Rooyen R, Hahn-Hägerdal B, La Grange DC, Van Zyl WH 2005. Construction of cellobiose-growing and fermenting *Saccharomyces cerevisiae* strains. J Biotechnol 120: 284-295.

Van Wyk N, Den Haan R, Van Zyl WH 2010. Heterologous production of NpCel6A from *Neocallimastix particiarum.* Accepted for publication in Enzyme and Microbial Technology. *doi:10.1016/j.enzmictec.2009.11.005*

Van Zyl WH, Lynd LR, Den Haan R, McBride JE 2007. Consolidated bioprocessing for bioethanol production using *Saccharomyces cerevisiae*. Adv Biochem Eng Biotechnol 108: 205-235.

CHAPTER 6

EFFECT OF THE DISRUPTION OF THE *DDI1* GENE ON THE HETEROLOGOUS PRODUCTION OF *NEOCALLIMASTIX PATRICIARUM* Cel6A BY *SACCHAROMYCES CEREVISIAE*

N. van Wyk, R. den Haan, W.H. van Zyl

Department of Microbiology, University of Stellenbosch, Matieland 7602, South Africa.

6.1. ABSTRACT: The DNA damage inducible (*DDI1*) gene of *Saccharomyces cerevisiae* has been implicated as a possible negative regulator of constitutive exocytosis. In this study, the *DDI1* gene was disrupted with a *TRP1* gene cassette. The disruption plasmid was integrated into *S. cerevisiae* Y294[NpCEL6A] which heterologously produced the cellulase NpCel6A of *Neocallimastix patriciarum*. A 1.9-fold increase was observed in normalized lichenase activity of the strain where the *DDI1* gene was disrupted compared to the parental strain. However, a significant decrease in the amount of biomass produced was also observed.

6.2. INTRODUCTION

Despite the ease-of-use and the GRAS (generally recognised as safe) status of *S. cerevisiae*, most heterologous protein production strategies employ other yeast genera such as *Pichia* and *Kluyveromyces* as hosts (Gellissen & Hollenberg, 1997). This is due to one of the major drawbacks of *S. cerevisiae,* namely its relatively limited secretion capacity compared to the abovementioned yeast genera (Buckholz & Gleeson, 1991). For the cost-effective application of *S. cerevisiae* as an ethanol producer from lignocellulosic biomass, heterologous cellulases and hemicellulases need to be secreted at high titers to enable it to grow on and concomitantly produce ethanol from lignocellulose (Van Zyl et al., 2007). It is thus of great interest to improve the secretion capacity of *S. cerevisiae*.

In general terms, metabolic engineering is the directed enhancement of a specific phenotype of an organism which can be obtained with the aid of theoretical modeling, biochemical data and genetic engineering (Ostergaard et al., 2000). Due to the relative ease that *S. cerevisiae* can be

genetically manipulated and the vast amount of information available on the metabolic processes of *S. cerevisiae*, metabolic engineering can provide dramatic improvements with regard to the secretion capacity of *S. cerevisiae*.

Several studies have shown that the secretion of mostly native proteins can be increased by either disrupting or overexpressing certain genes (Ruohonen et al., 1997; Chow et al., 1992). One such study showed the 20 % increase of overall protein secretion when the DNA damage inducible gene *DDI1* (then called *VSM1*) was disrupted (Lustgarten & Gerst, 1998). Ddi1 is a soluble *N*-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) protein regulator. Ddi1 binds to the *N*-terminal autoinhibitory domain of Sso1 which is a target SNARE protein and also to Snc1 and Snc2 which are vesicular SNARE-proteins (Marash & Gerst, 2003; Lustgarten & Gerst, 1998). Both these v-SNARE and t-SNARE proteins are essential for membrane fusion between transport vesicles and the plasma membrane. It inhibits SNARE-complex formation or vesicle-plasma membrane fusion and is thus considered a negative regulator of constitutive exocytosis. Ddi1 has three domains: an *N*-terminal ubiquitin-like (UBL) domain, a C-terminal ubiquitin-associated (UBA) domain and a retroviral aspartyl-protease (RVP) domain (Gabriely et al., 2008). Thus, in addition to its role in excytosis, Ddi1 has also been linked to various other cellular functions like its role in S-phase checkpoint control (Clarke et al., 2001).

In this study, the *DDI1* gene of *S. cerevisiae* was disrupted. The effect the disruption has on the secreted production of the heterologous protein *Npcel6A* of *Neocallimastix patriciarum* was investigated.

6.3. MATERIALS AND METHODS

6.3.1. MEDIA AND CULTIVATING CONDITIONS

Escherichia coli DH5α was used for all plasmid propagations. *E. coli* cells were cultivated at 37°C in Luria Bertani broth and agar (5 g/L yeast extract, 10 g/L NaCl, 10 g/L tryptone) supplemented with ampicillin (100 mg/L). The haploid lab yeast strain *S. cerevisiae* Y294 was used for all

heterologous protein production experiments. Yeast transformants were selected on synthetic complete agar (20 g/L glucose, 1.7 g/L yeast nitrogen base, 5 g/L $(NH_4)_2SO_4$ with appropriate amino acids minus tryptophan. For activity assays, precultured yeast cells were cultivated at 30°C in 25 mL YPD (20 g/L glucose, 20 g/L peptone, 10 g/L yeast extract) for three days in 250 mL erlenmeyer flasks on a rotary shaker at 100 rpm.

6.3.2. VECTOR CONSTRUCTION AND YEAST TRANSFORMATIONS

Standard procedures for all DNA manipulation work were performed as prescribed by Sambrook & Russel (2001). Polymerase chain reaction experiments were conducted with the High-Fidelity PCR enzyme (Fermentas, Vilnius, Lithuania). All restriction enzymes were purchased from Fermentas. The 1287-bp open reading frame (ORF) of the *DDI1* gene of *S. cerevisiae* was amplified from the strain Y294 using the DDI1-L/R primer set with the resulting fragment ligated into the pJET2.1 blunt-end vector (Fermentas) – designated pJET-DDI1. The intact *TRP1* cassette was amplified from the prototrophic *S. cerevisiae* D5A strain using primer set TRP1L/R and ligated into pJET2.1. The resulting plasmid was digested with *Pvu*II and *Smi*I and ligated into the *Mlu*NI-digested pJET-DDI1 and designated pJET-DDI1ΔTRP1. The pJET-DDI1ΔTRP1 disruption plasmid was linearized with *Aha*III and *Bcu*I and transformed into yeast using the LioAc/DMSO method (Hill et al., 1991). Total DNA of yeast transformants was isolated and polymerase chain reaction analyses (PCR) was conducted with DDI1-L/R primers to determine whether the *DDI1* gene was disrupted. The yeast strains used in this study are shown in Table 6.1 and details on the primers used in this study to amplify the *TRP1* gene cassette and *DDI1* ORF and are shown in Table 6.2.

TABLE 6.1: Strains used and generated in this study.

	Relevant genotype on episomal plasmid	Chromosomal genotype	Reference
***S. cerevisiae* D5A**		*TRP1*	ATCC 200062
***S. cerevisiae* Y294**		*leu2-3*,112 *ura3-52 his3 trp1-289*	ATCC 201160
Y294[REF]	*bla URA3 $PGK1_P$-XYNSEC- $PGK1_T$*	*fur1*::*LEU2*	Van Rooyen et al., 2005
Y294[NpCEL6A]	*bla URA3 $PGK1_P$-XYNSEC-Npcel6A- $PGK1_T$*	*fur1*::*LEU2*	Van Wyk et al., 2010
Y294[NpCEL6A]Δddi1	*bla URA3 $PGK1_P$-XYNSEC-Npcel6A- $PGK1_T$*	*fur1*::*LEU2, ddi1::TRP1*	This study

TABLE 6.2: Primers used in this study.

Primer name	Sequence 5'-3'	Genbank accession nr
DDI1-L	ATGGATTTAACAATTTCAAACG	ACFL01000111
DDI1-R	TCATTGGAAAAGGAGGGATG	
TRP1-L	TACCATGGTCCATCTCTATTCTGAAAACG	AAFW02000145
TRP1-R	CGATCGATTAAACGGATCTCGCATTG	

6.3.3. ENZYME ACTIVITY ASSAYS

After yeast cultures were grown for three days, cells were spun down and the supernatants were used for enzyme activity assays using lichenan (Sigma, St. Louis, MO, USA) as substrate. Reducing sugars released from 0.1 % lichenan in 0.05M citrate buffer pH 5.0 after five minutes of incubation with yeast supernatant at 40°C were detected using the dinitrosalicyclic method (Miller et al., 1960). A standard curve was set using glucose from which each sample's enzyme activity was calculated in Units/mL (µmol of reducing sugar released per minute per mL). Lichenase values were normalized with the dry cell weight (mg/mL) of the yeast cultures (Meinander et al., 1996).

6.4. RESULTS

6.4.1. CONSTRUCTION OF A *DDI1* DISRUPTION STRAIN

PCR analyses of yeast transformants revealed disruption of the *DDI1* as shown in Figure 6.1.

6.4.2. ENZYME ACTIVITY ASSAYS

NpCel6A produced by *S. cerevisiae* possesses good activity on lichenan – a linear polymer of β-D-glucopyranose residues of (1,3),(1,4)-linkages – as demonstrated previously (Van Wyk et al., 2010; Denman et al., 1996). Table 6.3 shows the lichenase activity obtained for the yeast strain Y294[NpCEL6A]Δddi1 which contained a disrupted *DDI1* gene compared to the parental strain Y294[NpCEL6A]. The volumetric enzyme activity between the two strains were similar, but the dry cell weight of Y294[NpCEL6A]Δddi1 was less than half of that of the parental strain. This caused the normalized lichenase activity of Y294[NpCEL6A]Δddi1 to be more than 1.9-fold higher than that obtained for the parental strain.

6.5. DISCUSSION

The observation in this study that disruption of the *DDI1* gene led to a dramatic decrease in the dry cell weight of the yeast culture is inconsistent with previous data which showed no significant difference in the vegetative growth compared to the parental strain (Lustgarten & Gerst, 1998). It is not clear why a difference was observed. The laboratory yeast strain Y294 was generated through random mutagenesis and is routinely used in our research group as a relatively good protein secretor. Apart from its auxotrophic markers, little is known about its genotype. Since Ddi1 is known to function in central metabolic processes apart from exocytosis, it might have a different effect on different laboratory yeast strains or in different media. Also, the *N*-terminal UBL domain might still be functional as disruption with the *TRP1* gene cassette occurred later-on in the sequence. The UBL domain is responsible for Ddi1 to bind to the proteasome (Kaplun et al., 2005).

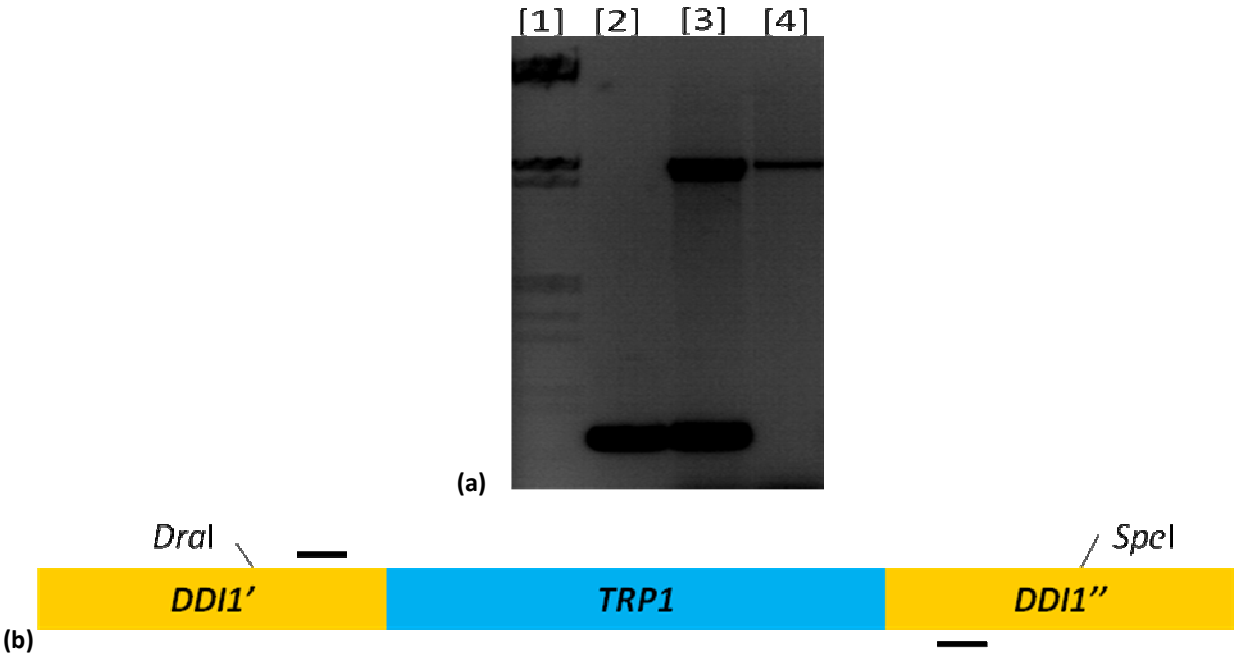


FIGURE 6.1: (a) Agarose gel electrophoresis of PCR amplicons amplified from total DNA isolated from yeast transformants using DDI1L/R primer set. Lane 1 is lambda DNA digested with Hin*dIII* and *Eco*RI. Lane 2 contains yeast with intact *DDI1* gene. Lane 3 contains yeast transformant where a single crossover integration took place. Lane 4 contains double cross-over integration i.e. gene replacement. (b) The linearized construction of *DDI1* disruption plasmid with an intact *TRP1* gene cassette. The black lines indicate the annealing regions of primers used to demonstrate *DDI1* gene disruption.

TABLE 6.3: Enzymatic activities of supernatant of yeast transformants. Plus/minus signs indicate the standard variation from the mean of three repeats.

	Volumetric activity (Units/mL)	Dry cell weight of yeast cells (mg/mL)	Normalized lichenase activity (Units/mg DCW)
Y294[NpCEL6A]	26.985 ± 0.878	3.321 ± 0.006	8.129 ± 0.349
Y294[NpCEL6A]Δddi1	22.970 ± 2.425	1.480 ± 0.008	15.522 ± 1.287
Y294[REF]	2.850 ± 2.308	3.356 ± 0.005	0.850 ± 0.690

Despite the lower yield in cells, roughly the same amount of the heterologous cellulase was produced in the strain where *DDI1* gene was disrupted compared to the parental strain. This is a promising observation as less cell mass might ease post-fermentation clean-up procedures. By combining several efforts of metabolic engineering, a *S. cerevisiae* strain might be generated that could produce desirable amounts of heterologous hydrolytic enzymes and make it an even more attractive host for ethanol production from lignocellulose.

6.5. LITERATURE CITED

Buckholz RG, Gleeson MAG 1991. Yeast systems for the commercial production of heterologous proteins. Bio/Technology 9: 1067-1072.

Clarke DJ, Mondesert G, Segal M, Bertolaet BL, Jensen S, Wolff M, Henze M, Reed SI. 2001. Dosage suppressors of *pds1* implicate ubiquitin-associated domains in checkpoint control. Mol Cell Biol 21: 1997-2007.

Gabriely G, Kama R, Gelin-Licht R, Gerst JE 2008. Different domains of the UBL-UBA ubiquitin receptor, Ddi1/Vsm1, are involved in its multiple cellular roles. Mol Cell Biol 19: 3625-3637.

Gellissen G, Hollenberg CP 1997. Application of yeasts in gene expression studies: a comparison of *Saccharomyces cerevisiae, Hansenula polymorpha* and *Kluyveromyces lactis* - a review. Gene 190: 87-97.

Hill J, Donald KA, Griffiths DE 1991. DMSO-enhanced whole cell yeast transformation. Nucleic Acids Res 19: 5791.

Kaplun L, Tzirkin R, Bakhrat A, Shabek N, Ivantsiv Y, Raveh D 2005. The DNA damage-inducible UbL-UbA protein Ddi1 participates in Mec1-mediated degradation of Ho endonuclease. Mol Cell Biol 25: 5355-5362.

Lustgarten V, Gerst JE 1999. Yeast *VSM1* encodes a v-SNARE binding protein that may act as a negative regulator of constitutive exocytosis. Mol Cell Biol 19: 4480-4494.

Marash M, Gerst JE 2003. Phosphorylation of the autoinhibitory domain of the Sso t-SNAREs promotes binding of the Vsm1 SNARE regulator in yeast. Mol Biol Cell 14: 3114-3125.

Meinander N, Zacchi G, Hahn-Hägerdal B 1996. A heterologous reductase affects the redox balance of recombinant *Saccharomyces cerevisiae*. Microbiology 142: 165-172.

Miller GL, Blum R, Glennon WE, Burton AL 1960. Measurement of carboxymethyl cellulase activity. Anal Biochem 2: 127-132.

Ruohonen L, Toikkanen J, Tieaho V, Outola M, Soderlund H, Keranen S 1997. Enhancement of protein secretion in *Saccharomyces cerevisiae* by overproduction of Sso protein, a late-acting component of the secretory machinery Yeast 13: 337-351.

Sambrook J, Russel DB 2001. Molecular cloning: A laboratory manual. Cold spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Van Rooyen R, Hahn-Hägerdal B, La Grange DC, Van Zyl WH 2005. Construction of cellobiose-growing and fermenting *Saccharomyces cerevisiae* strains. J Biotechnol 120: 284-295.

Van Wyk N, Den Haan R, Van Zyl WH 2010. Heterologous production of NpCel6A from *Neocallimastix particiarum. Accepted for publication in Enzyme and Microbial Technology*

Van Zyl WH, Lynd LR, Den Haan R, McBride JE 2007. Consolidated bioprocessing for ethanol production using *Saccharomyces cerevisiae*. Adv Biochem Eng Biot 108: 205-235.

CHAPTER 7

THE ROLE OF SWOLLENIN ON THE CELLULOLYTIC EFFECT OF *SACCHAROMYCES CEREVISIAE* WHEN HETEROLOGOUSLY CO-EXPRESSED WITH CELLULASE-CODING GENES.

N. van Wyk, S.H. Rose[1], R. den Haan, W.H. van Zyl

Department of Microbiology, University of Stellenbosch, Matieland 7602, South Africa.

7.1. ABSTRACT: Swollenin is an expansin-like protein of fungal origin which causes the weakening of filter paper and makes cellulose substrates more accessible to cellulases. In this study, the impact of swollenin on the cellulolytic effect of recombinant *Saccharomyces cerevisiae*, was investigated. A synthetically-designed swollenin-encoding gene (*swoI*) of *Trichoderma reesei* was expressed in *Saccharomyces cerevisiae* under the transcriptional control of the *ENO1* promoter. No dramatic improvement in Avicel hydrolysis was observed when swollenin was co-expressed with two cellulase-coding genes (*cel9A* of *Thermobifida fusca* and *cel6A* of *Neocallimastix patriciarum*). Filter paper treated with supernatants of recombinant yeast cultures showed that treatment with swollenin led to a statistically significant decrease in the amount of force needed to break the filter paper. However, the force needed to break the filter paper was still considerably more than what was needed with filter paper treated with an ionic liquid [BMiM]$MeSO_4$.

7.2. INTRODUCTION

Cellulose is the main polymer of the cell walls of plant cells and is the most abundant polysaccharide found on earth. The chemical composition of cellulose is relatively simple as it consists of D-glucose residues linked by β-(1,4)-glycosidic bonds which can form linear polymeric chains of over 10 000 glucose residues (Klemm et al., 2002). The cellulose found in plant cell walls is mostly of a highly crystalline nature where individual chains are linked to each other, interspersed by amorphous regions. Cellulose has a complex morphology, mainly due to its intermolecular bonding pattern (Hon, 1994).

[1] Dr SH Rose designed the SwoI gene and constructed the pBBH-SwoI plasmid.

Cellulose is widely regarded as an important renewable energy resource and holds potential as a substrate for the production of biofuels, in particular ethanol (Farrell et al., 2006). A major impediment for the economical conversion of cellulose to ethanol remains the lack of low-cost technologies that can address the recalcitrance of cellulose towards enzymatic hydrolysis (Stephanopolous, 2007). A promising strategy for the cost-reduction of the conversion of cellulose to ethanol is known as consolidated bioprocessing (CBP) which combines all the biological-facilitated transformations needed for the conversion into one step (Lynd et al., 2005). For CBP to be realized, an organism is required that is able to depolymerize the cellulose substrate and ferment the resulting D-glucose monomers to ethanol at a feasible rate. *Saccharomyces cerevisiae* shows potential to satisfy the requirements needed for CBP as it is a prolific ethanol producer and could be made cellulolytic with the aid of recombinant DNA techniques (Van Zyl et al., 2007). Significant strides have been made in constructing a cellulolytic yeast (Den Haan et al., 2007a; Van Rooyen et al., 2005) although not enough cellulases are being produced to support growth on crystalline cellulose as carbohydrate source (Den Haan et al., 2007b).

Early theories on the biological degradation of cellulose involved a two-component system called C_x and C_1 (Reese et al., 1950). C_x would include all the hydrolytic enzymes or cellulases and C_1, also called the swelling factor, would include all the non-hydrolytic enzymes that improve the accessibility of cellulose for the C_x components. When cultivated in cellulose-containing media, *Trichoderma* spp. produce swollenin (SwoI) – a protein capable of disrupting the crystallinity of cellulose fibres without generating any significant amount of reducing sugars and can thus be classified as a C_1 component (Saloheimo et al., 2002). The SwoI protein contains an N-terminal family 1 carbohydrate binding module (CBM1) and a C-terminal plant expansin-like domain. In general, plant expansins stimulate the rapid extension of cell walls by weakening the non-covalent bonds that aid in sustaining the integrity of the cell walls (Cosgrove et al., 2002).

Other studies have shown that overexpression of the swollenin gene led to improved colonization of *Trichoderma asperellum* on cucumber roots (Brotman et al., 2008). However, when the CBM region was disrupted, no improved colonization was found implying the importance of this moiety. Swollenin was also fused to the feruloyl esterase A of *Aspergillus niger* which led to increased release of ferulic acid from wheat bran (Levasseur et al., 2006).

In an attempt to investigate the degree to which swollenin could contribute toward the overall cellulolytic effect of *S. cerevisiae*, the *swoI* gene of *T. reesei* was co-expressed with two cellulases – the *cel9A* gene of *Thermobifida fusca* and the *cel6A* gene of *Neocallimastix patriciarum*.

7.3. MATERIALS AND METHODS

7.3.1. MEDIA AND CULTURING CONDITIONS

All the chemicals and media components used in this study were of laboratory grade standard. The *Escherichia coli* strain DH5α was used for plasmid transformation and propagation. *E. coli* cells were cultivated in Luria Bertani medium (5 g/L yeast extract, 10 g/L NaCl, 10 g/L tryptone) supplemented with ampicillin (100 mg/L). *S. cerevisiae* Y294 transformants were selected and maintained on synthetic complete medium without uracil (SC^{-URA}) or without uracil and leucine ($SC^{-URA-LEU}$) plates (1.7 g/L yeast nitrogen base without amino acids and ammonium sulphate [Difco laboratories, Detroit, MI, USA], 5 g/L $(NH_4)_2SO_4$, 20 g/L glucose, 20 g/L agar and supplemented with amino acids as required). Autoselective *S. cerevisiae* strains were cultured in YPD medium (10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose) or buffered SC media containing 20 g/L succinic acid with pH adjusted to 6.0 with NaOH. Yeast strains were routinely cultured in 250-mL erlenmeyer flasks containing 50 mL medium at 30°C, on a rotary shaker at 100 rpm.

7.3.2. VECTOR CONSTRUCTION

Standard procedures for DNA manipulation were performed according to protocols described by Sambrook & Russel (2001). All restriction DNA endonuclease enzymes were purchased from

Fermentas (Vilnius, Lithuania). For polymerase chain reactions, the High Fidelity PCR Enzyme Mix (Fermentas) was used as recommended by the manufacturer with a Perkin Elmer GeneAmp® PCR System 2400 (The Perkin-Elmer Corporation, Norwalk, CT, USA). Details of all the primers used in this study are given in Table 7.1. All plasmids used in this study are shown in Table 7.2. The s*woI* gene sequence including its native secretion signal and sites for *Eco*RI and *Xho*I endonuclease digestion was codon-optimized (Carbone et al., 2003) and a gene sequence with a codon adaptation index (CAI) value of 0.914 was sent for de novo synthesis to GenScript Corporation (Piscataway, NJ, USA). The resulting pUC-57 vector containing the 1494-bp *swoI* gene was digested with *Eco*RI and *Xho*I and ligated into the corresponding sites of the *E. coli*/*S. cerevisiae* shuttle vector pBBH containing the Enolase 1 (*ENO*1) promoter and terminator sequences – designated pBBH-SwoI. The primer set ENO1-L and ENO1-R was used to amplify the *ENO*1$_{P/T}$-*swoI* gene cassette from pBBH-SwoI and was ligated into the pTZ57R/T vector. The plasmid was digested with *Hin*dIII obtaining a 2580-fragment which was ligated into a *Hin*dIII-digested pNpCel6A and yXS+Cel9A. The resulting plasmids were designated pNpCel6A+SwoI and pCel9A+SwoI, respectively.

TABLE 7.1: Primers used in this study. Relevant restriction enzymes in boldface.

Primer name	Sequence (5'-3')	Relevant restriction sites	Genbank accession number
SWOI-L	**GAATTC**ATGGCTGGTAAGTTG	*Eco*RI	CAB92328
SWOI-R	**CTCGAG**TTAGTTTTGAGAGAATTG	*Xho*I	
Tfuscel90-L	AAGCT**TCGCGA**GAACCGGCGTTCAACTAC	*Nru*I	L20093
Tfuscel90-R	AGGC**CTCGAG**GTTGCCGTGTTAGGCG	*Xho*I	
ENO1-L	**GGATCC**ACTAGTCTTCTAGGCGGGTTATC	*Eco*RI	X99228
ENO1-R	**AAGCTT**GCGGCCGCAAAGAGGTTTAGACATTGG	*Eco*RI	
PGK1-L	ACTG**AAGCTT**GGATCCTTAAAGATGCCG	*Hin*dIII	NC001135
PGK1-R	ACTG**AAGCTT**GGCCAAGCTTTAACGAAC	*Hin*dIII	
FUR1-L	TCCGTCTGGCATATCCTA		M36485
FUR1-R	TTGGCTAGAGGACATGTA		
Npcel6A-L	**GAATTC**CAACATGGTTTCCT	*Eco*RI	ACC49315
Npcel6A-R	**CTCGAG**TTAGAAGGATGGTC	*Xho*I	

7.3.3. YEAST TRANSFORMATION

S. cerevisiae Y294 strains were transformed with plasmids using the LiOAc/DMSO protocol as described by Hill et al. (1991). Autoselective strains were generated by disruption of the *FUR1*

gene with pDF1 linearized with *Ava*III and *Nco*I (La Grange et al., 1996). All yeast strains constructed and used in this study are shown in Table 7.2.

7.3.4. AVICELASE ASSAY

To determine Avicelase activity, 300 µL of the yeast culture supernatant grown in buffered, double-strength SC media for four days at 30°C was added to deep-well microtiter plates with each well containing a 300 µL solution with 2% Avicel PH-105 (Fluka, Buchs, Switzerland), 0.1 M acetate buffer pH 5.0, 0.04% NaN_3 and 0.3 µL Novozyme-188 (Sigma, St. Louis, MO, USA). The deep-well plate was sealed and incubated on a microtiter plate shaker and shaken at ~1000 rpm for 24 h at 40°C. One hundred microliter samples of the enzyme-substrate mixtures were taken at the 0 h and 24 h time intervals to determine background sugars that were present and the amount of sugars released respectively, using a modified DNS method (Miller et al., 1960). The samples were transferred to a 96-well PCR plate using a multi-channel pipette and were centrifuged at 2000 rpm for two min. Fifty microliters of the supernatant were subsequently pipetted along with 100 µL DNS solution (1% 3,5-dinitrosalicyclic acid, 20% potassium sodium tartrate, 1% NaOH, 0.2% phenol, 0.05% Na_2SO_3) into a clean 96-well PCR plate, sealed, heated at 99°C for five min and cooled at 4°C for one min. Absorbance values were determined at 540 nm on an X-mark™ microtiter plate reader (Biorad, Hercules, CA, USA). Glucose was used to construct of a standard curve (0.125 mg/mL- 2 mg/mL) from which the Units/mL of each sample was calculated. A unit was defined as the amount of enzyme that released 1 µmol of monomeric sugar from the substrate per min. All values were normalized with the dry cell weight in mg/mL of the relevant yeast culture. Dry cell weight (DCW) of each sample was determined as previously described (Den Haan et al. 2007b).

7.3.5. MEASUREMENT OF FILTER PAPER MECHANICAL STRENGTH

Whatman nr 1 filter paper strips (1×6 cm) (Whatman, Kent, UK) were treated with supernatants of recombinant yeasts cultured for three days in YPD. Filter papers were also treated with 1-butyl-3-methyl-imidazolium-methylsulfate ([BMiM]$MeSO_4$) (Fluka), an ionic liquid with known cellulose solvent properties (Van Rantwijk et al., 2003). Fifty filter paper strips were used for

each treatment. Twenty microliters of yeast supernatant were spotted in the middle of the strip and incubated overnight at 37 °C. The force (in Newton) needed to break treated filter paper strips was measured on an Instron Universal Strength Testing Machine (Instron Ltd, High Wycombe, UK). Clamps used in the experiment were spaced 4 cm apart. A load cell of 50 kg was used. The peak load measured for breakage of every filter paper was recorded and used for further calculations.

TABLE 7.2: Plasmids and yeast strains used in this study.

Plasmids propagated	Abbreviated name	Relevant genotype	Reference
E. coli DH5 α		*supE44 ΔlacU169 (ϕ80lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	Sambrook & Russel, 2001
yXYNSEC		*bla* $PGK1_P$-XYNSEC-$PGK1_T$	Van Rooyen et al., 2005
pBBH		*bla* $ENO1_{P/T}$	Rose et al., 2009
pUC-57-SwoI		*bla swoI*	This work
pBBH-SwoI		*bla* $ENO1_P$-*swoI*-$ENO1_T$	This work
yXS+Cel9A		*bla* $PGK1_P$-XYNSEC-*cel9A*-$PGK1_T$	Van Wyk et al., 2010a
pCel9A+SwoI		*bla* $PGK1_P$-XYNSEC- $PGK1_T$ $ENO1_P$-*swoI*- $ENO1_T$	This work
pNpCel6A		*bla* $PGK1_P$-XYNSEC-*Npcel6A*-$PGK1_T$	Van Wyk et al., 2010b
pNpCel6A+SwoI		*bla* $PGK1_P$-XYNSEC-*Npcel6A*-$PGK1_T$ $ENO1_P$-*swoI*-$ENO1_T$	This work
pDF1		*bla fur1::LEU2*	La Grange et al., 1996
Yeast strains			
S. cerevisiae Y294		*leu2, ura3, his3, trp1*	ATCC 201160
fur1::LEU2 yXYNSEC	Y294[REF]	*bla ura3/URA3* $PGK1_P$-XYNSEC-$PGK1_T$	Van Rooyen et al., 2005
fur1::LEU2 pSwoI	Y294[SWOI]	*bla ura3/URA3* $ENO1_P$-*swoI*- $ENO1_T$	This work
fur1::LEU2 yXS+Cel9A	Y294[CEL9A]	*bla ura3/URA3* $PGK1_P$-XYNSEC-*cel9A*-$PGK1_T$	Van Wyk et al., 2010a
fur1::LEU2 pCel9A+SwoI	Y294[CEL9A+SWOI]	*bla ura3/URA3* $PGK1_P$-XYNSEC-*cel9A*-$PGK1_T$ $ENO1_P$-*swoI*-$ENO1_T$	This work
fur1::LEU2 pNpCel6A	Y294[NpCEL6A]	*bla ura3/URA3* $PGK1_P$-XYNSEC-*Npcel6A*- $PGK1_T$	Van Wyk et al., 2010b
fur1::LEU2 pNpCel6A+SwoI	Y294[NpCEL6A+SWOI]	*bla ura3/URA3* $PGK1_P$-XYNSEC-$PGK1_T$ $ENO1_P$-*swoI*-$ENO1_T$	This work

7.4. RESULTS

7.4.1. CLONING AND EXPRESSION OF SWOLLENIN IN *S. CEREVISIAE*

The synthetic *swoI* gene of *T. reesei* was placed under the transcriptional control of the constitutive *S. cerevisiae ENO1* promoter. For the co-expression of swollenin with cellulase-encoding genes, the processive endoglucanase *T. fusca cel9A* gene sequence, with its secretion

signal replaced with the secretion signal of the xylanase II gene (*xyn11B*) of *T. reesei* and the synthetic cellobiohydrolase *cel6A* of *N. patriciarum* were placed under the transcriptional regulation of the constitutive *S. cerevisiae PGK1* promoter (Van Wyk et al., 2010a; Van Wyk et al., 2010b). The cellulase gene cassettes were placed on the same plasmid as the swollenin gene cassette. Confirmation of the insertion of heterologous genes in yeast transformants was done with PCR analyses on total yeast DNA isolates using primers (Table 7.1) that annealed to the specific heterologous genes (results not shown).

7.4.2. CELLULASE ACTIVITY OF RECOMBINANT *S. CEREVISIAE* STRAINS

A relatively small halo could be observed around Y294[SWOI] in the plate assay with lichenan as substrate implying a small amount of hydrolytic activity or structural changes to the substrate (Figure 7.1). Comparatively large halos were seen around Y294[CEL9A], Y294[CEL9A+SWOI], Y294[NpCEL6A] and Y294[NpCEL6A+SWOI]. Figure 7.2 shows the activity experiments with Avicel as substrate. In this case, no hydrolytic activity was detected with Y294[SWOI]. The activity observed with Y294[NpCEL6A] and Y294[NpCEL6A+SWOI] was similar. Slightly higher levels of cellulase activity was obtained with Y294[CEL9A+SWOI] compared to Y294[CEL9A].

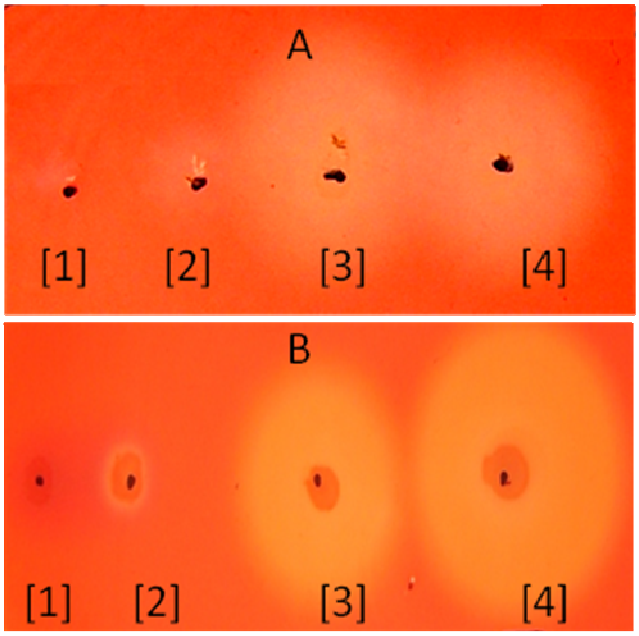


FIGURE 7.1: Plate assay of recombinant yeast strains grown for three days at 30 °C on a SC-plate containing 0.1 % lichenan. **A** [1] Y294[REF], [2] Y294[SWOI], Y294[CEL9A], Y294[CEL9A+SWOI]; **B** [1] Y294[REF], [2] Y294[SWOI], [3] Y294[NpCEL6A], [4] Y294[NpCEL6A+SWOI].

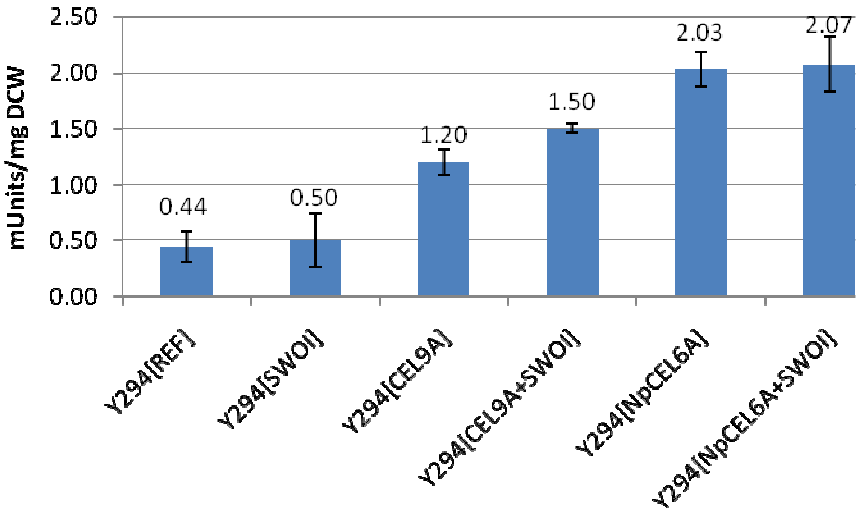


FIGURE 7.2: Avicelase activity assays of recombinant yeast cultures.

7.4.3. FILTER PAPER TENSILE STRENGTH

The level of the weakening of filter paper strips by swollenin was determined by measuring the force needed to break the filter paper. The datasets generated were subjected to Lilliefors analyses along with histograms to ascertain whether each dataset was normally distributed. With the exception of the filter paper treated with ionic liquid, treatments showed normal distribution. The one-tailed Student's t-tests were performed on all generated datasets (except ionic liquid) to determine whether there is any difference in the distribution between them. To establish the reproducibility of the experiment, the supernatants of two different cultures of Y294[REF] and Y294[SWOI] were used on different batches of filter paper on different days. As shown in Table 7.3, the H_0 hypothesis, which assumes the distributions of the two datasets compared are the same, could in both cases not be rejected. Table 7.4.1 and 7.4.2 show the Student's t-test comparisons with all the relevant treatments. Compared to the reference treatment (Y294[REF]), all treatments excluding Y294[CEL9A] and Y294[CEL9A+SWOI] led to a significant decrease in the force needed to break the filter paper. As shown in Table 7.4.1 ten microliters of each of Y294[CEL9A] and Y294[SWOI] led to a significant reduction in force needed to break the filter paper when compared to all the other compared treatments. Ten microliters of each of Y294[NpCEL6A] and Y294[SWOI] also led to a significant reduction in the force needed to break the filter paper except for when compared to Y294[NpCEL6A] (Table

7.4.2). With both individual cellulase treatments (Y294[CEL9A] and Y294[NpCEL6A]) similar distributions were found as with individual swollenin treatment (Y294[SWOI]). No decrease in reduction of tensile strength was observed in the co-expression strains (Y294[CEL9A+SWOI] and Y294[NpCEL6A+SWOI]) compared to the strains where the cellulases and swollenin were expressed individually. Treatment with the ionic liquid [BMiM]$MeSO_4$, compared to the individual treatment of swollenin (the mean of the two separate treatments), as shown in Figure 7.3 led to a four-fold reduction in force needed to break the filter paper.

TABLE 7.3: Evaluation of effectiveness of filter paper breakage experiments. Error bars indicate the standard variation from the mean of three repeats.

	Y294[REF] (1)	Y294[REF] (2)	Y294[SWOI] (1)	Y294[SWOI] (2)
Mean	15.60655	15.19674	14.34578	14.4073312
Variance	10.81732	5.2181	8.722791	3.087897994
Observations	50	50	50	50
t Stat	0.723656		-0.12664	
t Critical one-tail	2.369977		2.373868	
P(T<=t) one-tail	0.235609		0.449771	
H_0 hypothesis	Not Reject		Not Reject	

7.5. DISCUSSION

Swollenin is an expansin-like protein of fungal origin which has shown the ability to weaken insoluble cellulose substrates by supposedly disrupting the hydrogen bonds between different cellulose microfibrils without possessing any hydrolytic activity (Saloheimo et al., 2002). Swollenin is produced by *Trichoderma* spp. and presumably other cellulolytic fungi when grown on cellulose. Therefore, it would be of interest for the construction of a cellulolytic *S. cerevisiae* to evaluate swollenin's effect when heterologously produced by the yeast.

The small halo around Y294[SWOI] which point to some hydrolytic activity is consistent with previous data that the expansin-like domain shows statistically significant homology to endoglucanases of the glycoside hydrolase family 45 (Saloheimo et al., 2002). Thus, even though by definition expansin-like proteins possess no hydrolytic activity, swollenin exhibited slight β-glucanase activity or merely altered the structure of lichenan on plate assays. The relatively large halos visible around Y294[CEL9A], Y294[CEL9A+SWOI], Y294[NpCEL6A] and

Y294[NpCEL6A+SWOI] reconfirmed previous findings of Cel9A and NpCel6A having activity on cellulose substrates (Irwin et al., 1998; Denman et al., 1996). As plate assays are only valuable for qualitative proof of activity, no deductions can be made with regard to the observation of a larger halo around Y294[NpCEL6A+SWOI] in comparison to Y294[NpCEL6A]. Avicel assays revealed no dramatic improvement in the co-expression strains with only a slight increase in reducing sugar released when Cel9A was co-expressed with SwoI. In previous studies, to determine the effect of swollenin on cellulose substrates, the swollenin protein was either purified or the supernatant containing the swollenin was concentrated (Levasseur et al., 2006; Saloheimo et al., 2002) whereas in this study the effect of the crude swollenin protein was evaluated as it was produced by the recombinant yeast in the supernatant. For CBP applications, a yeast that requires down-stream concentration of components for cellulose degradation, would be highly undesirable.

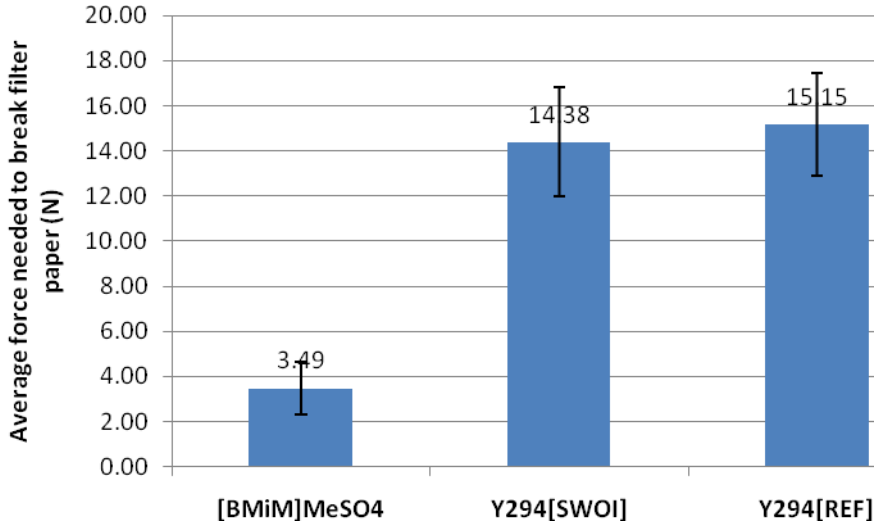


FIGURE 7.3: Comparison between treatments of filter paper with the ionic liquid and supernatant of yeast producing swollenin. Error bars indicate the standard variation from the mean of three repeats.

The experimental set-up of measuring the tensile strength of filter paper was similarly performed previously (Saloheimo et al., 2002) with the main difference being that the supernatant used in this study was not concentrated. High levels of variation of all the treatments were observed, possibly due to the heterogeneity of the filter paper and the relatively low concentrations of the swollenin and cellulases. Swollenin did show an improved effect on filter paper breakage, but the yeast supernatant treatments which led to the most

reduction in force needed to break the filter paper were combinations of supernatants of yeast which individually produced a cellulase and swollenin. This could suggests that, when co-expressed, less of the cellulase and swollenin were produced. This might be attributed to the increase in plasmid size of the co-expression plasmids, which results in lower plasmid copy number per cell or a congestion in the yeast secretory apparatus leading to a decrease in the amount of enzymes produced.

TABLE 7.4.1: Summary of Student's t-test comparisons of datasets of different treatments of filter paper.

		Y294[REF]	Y294[CEL9A]	Y294[SWOI]	Y294[CEL9A+ SWOI]	10µl of Y294 [CEL9A] + 10µl of Y294[SWOI]
Mean of force needed to break filter paper (in N)		15.19674	14.54479	14.40733	14.49131	12.84777
Variance		5.2181	4.7745	3.0879	17.5386	8.1901
t Stat	**10µL Y294 [CEL9]+ 10 µL Y294 [SWOI]**	4.536047	3.332666	3.283762	2.291163	
t Critical (α=0.05)		1.661404	1.661585	1.663884	1.662557	
P(T<=t) one-tail		8.54E-06	0.000620	0.000757	0.012185	
H_0 hypothesis		Reject	Reject	Reject	Reject	
t Stat	**Y294 [CEL9A+ SWOI]**	1.936827	0.080053	-0.130749		
t Critical (α=0.05)		1.661585	2.377802	1.668271		
P(T<=t) one-tail		0.027917	0.468206	0.448186		
H_0 hypothesis		Not reject	Not Reject	Not Reject		
t Stat	**Y294 [SWOI]**	1.936828	0.346634			
t Critical (α=0.05)		1.661585	1.661229			
P(T<=t) one-tail		0.027917	0.364820			
H_0 hypothesis		Reject	Not Reject			
t Stat	**Y294 [CEL9A]**	1.45835				
t Critical (α=0.05)		2.365002				
P(T<=t) one-tail		0.073971				
H_0 hypothesis		Not reject				

Recently, the use of ionic liquids has become an attractive option for the pretreatment of lignocellulosic biomass since some of these liquids can form hydrogen bonds with cellulose at high temperatures which result in the dissolution of cellulose crystal structure (De Costa Sousa et al., 2009). Even though there are a number of drawbacks with the use of ionic liquids, including the cost and the inhibitory effect on enzymatic activity, with optimized strategies, ionic liquids could be applied in large-scale pretreatment of all kinds of lignocellulose. With simple tensile strength experiments, it was shown in this study that an ionic liquid could have a dramatically improved effect on disrupting the crystallinity of the cellulose in comparison to swollenin which implies that much more swollenin is needed for a similar effect.

TABLE 7.4.2: Summary of Student's t-test comparisons of datasets of different treatments of filter paper.

		Y294[REF]	Y294 [NpCEL6A]	Y294[SWOI]	Y294 [NpCEL6A+ SWOI]	10µl of Y294 [NpCEL6A] + 10µl of Y294[SWOI]
Mean of force needed to break filter paper (in N)		15.60655	13.147812	14.345781	14.090207	13.011591
Variance		10.81732	2.508513	8.722791	9.167826	3.479861
t Stat	**10µL Y294 [NpCEL6A]+ 10 µL Y294 [SWOI]**	4.852784	0.393617	2.700695	2.144600	
t Critical (α=0.05)		1.664624	1.661052	1.663420	1.663650	
P(T<=t) one-tail		3.06E-06	0.347373	0.004193	0.017472	
H_0 hypothesis		Reject	Not Reject	Reject	Reject	
t Stat	**Y294 [NpCEL6A + SWOI]**	2.398446	-1.950135	0.427256		
t Critical (α=0.05)		0.009189	1.665707	1.660551		
P(T<=t) one-tail		1.660715	0.027474	0.335065		
H_0 hypothesis		Reject	Not Reject	Not Reject		
t Stat	**Y294 [SWOI]**	2.016780	-2.527641			
t Critical (α=0.05)		1.660714	1.665425			
P(T<=t) one-tail		0.023241	0.006794			
H_0 hypothesis		Reject	Not Reject			
t Stat	**Y294 [NpCEL6A]**	4.762673				
t Critical (α=0.05)		1.666600				
P(T<=t) one-tail		4.90E-06				
H_0 hypothesis		Reject				

Even though it can be assumed that any improvements in cellulose degradation described in this study could be ascribed to the effect of the expansin-like domain of swollenin, it is still unclear to what extent the family 1 CBM has contributed to cellulose degradation. A recent finding showed that hydrolase-less CBMs are produced by *T. fusca* that have an increased effect on cellulose hydrolysis when cellulases are at a low concentration and at long incubation times (Moser et al., 2008). It has been shown that the removal of the CBM from the expansin-like domain resulted in a decreased effect of swollenin (Brotman et al., 2008), suggesting that it is more likely that both domains contributed to any improved cellulolytic effects.

Large-scale enzyme cocktails prepared by biotechnology companies like Genencor and Novozyme often employ the enzymes produced by lignocellulosic aerobic fungi like *Trichoderma* spp. and *Aspergillus* spp. These enzyme cocktails which include swelling factor proteins form an integral part in conversion strategies of lignocellulose to ethanol. Even in many CBP-related strategies, these enzyme additions will most likely still be incorporated. It is not clear to what degree swollenins or any other swelling factor proteins contribute to the hydrolysis of cellulose in these cocktails, but the amount produced by *S. cerevisiae* and the

negligible effect it had when co-expressed with a cellulase does not render it useful to *S. cerevisiae* as a candidate CBP-organism.

7.6. LITERATURE CITED

Brotman Y, Briff E, Viterbo A, Chet I 2008. Role of swollenin, an expansin-like protein from *Trichoderma*, in plant root colonization. Plant Physiol 147: 779-789.

Carbone A, Zinovyev A, Képès F 2003. Codon adaptation index as a measure of dominating codon bias. Bioinformatics 19: 2005-2015.

Cosgrove DJ, Li LC, Cho H, Hoffmann-Benning S, Moore RC, Blecker D 2002. The growing world of expansins. Plant Cell Physiol 43: 1436-1444.

Da Costa Sousa L, Chundawat SPS, Balan V, Dale BE 2009. 'Cradle-to-grave' assessment of existing lignocellulose pretreatment technologies. Curr Opin Biotechnol 20:1-9.

Den Haan R, Rose SH, Lynd LR, Van Zyl WH 2007a. Hydrolysis and fermentation of amorphous cellulose by recombinant *Saccharomyces cerevisiae*. Metab Eng 9: 87-94.

Den Haan R, Mcbride JE, La Grange DC, Lynd LR, Van Zyl WH 2007b. Functional expression of cellobiohydrolases in *Saccharomyces cerevisiae* towards one-step conversion of cellulose to ethanol. Enzyme Microb Technol 40: 1291-1299.

Denman S, Xue GP, Patel B 1996. Characterization of a *Neocallimastix patriciarum* cellulase cDNA (celA) homologous to *Trichoderma reesei* cellobiohydrolase II. Appl Environ Microbiol 62: 1889-1896.

Farrell AE, Plevin RJ, Turner BT, Jones AD, O'Hare M, Kammen DM 2006. Ethanol can contribute to energy and environmental goals. Science 311: 506-508.

Hill J, Donald KA, Griffiths DE 1991. DMSO-enhanced whole cell yeast transformation. Nucleic Acids Res 19: 5791.

Hon DNS 1994. Cellulose: a random walk along its historical path. Cellulose 1: 1-25.

Irwin D, Shin DH, Zhang S, Barr BK, Sakon J, Karplus PA, Wilson DB 1998. Roles of the catalytic domain and two cellulose binding domains of *Thermomonospora fusca* E4 in cellulose hydrolysis. J Bacteriol 180: 1709-1714.

Klemm D, Schmauder H-P, Heinze T 2002. Cellulose. In: Biopolymers: Biology, Chemistry, Biotechnology, Application, Polysaccharide II. (Vandamme E, De Baets S, Steinbüchel A eds.), Wiley-VCH, Weinheim, Germany. pp 275-319.

La Grange DC, Pretorius IS, Van Zyl WH 1996. Expression of a *Trichoderma reesei* beta-xylanase gene (XYN2) in *Saccharomyces cerevisiae*. Appl Environ Microbiol 62: 1036-1044.

Levasseur A, Saloheimo M, Navarro D, Andberg M, Monot F, Nakari-Setälä T, Asther M, Record E 2006. Production of a chimeric enzyme tool associating the *Trichoderma reesei* swollenin with the *Aspergillus niger* feruloyl esterase A for release of ferulic acid. Appl Microbiol Biotechnol 73: 872-880.

Lynd LR, Van Zyl WH, McBride JE, Laser M 2005. Consolidated bioprocessing of cellulosic biomass: an update. Curr Opin Biotechnol 16: 577-583.

Miller GL, Blum R, Glennon WE, Burton AL 1960. Measurement of carboxymethyl cellulase activity. Anal Biochem 2: 127-132.

Moser F, Irwin D, Chen S, Wilson DB 2008. Regulation and characterization of *Thermobifida fusca* carbohydrate-binding module proteins E7 and E8. Biotechnol Bioeng 100: 1066-1077.

Reese ET, Sui RGH, Levinson HS 1950. The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. J Bacteriol. 59: 485-497.

Rose SH, Fouché N, Van Zyl WH 2009. Comparing the expression of *Saccharomyces cerevisiae* and *Kluyveromyces lactis* by evaluating their ability to recombinantly produce the *Aspergillus aculeatus* Man1 enzyme. Submitted for publication.

Saloheimo M, Paloheimo M, Hakola S, Pere J, Swanson B, Nyyssonen E, Bhatia A, Ward M, Penttila M 2002. Swollenin, a *Trichoderma reesei* protein with sequence similarity to the plant expansins, exhibits disruption activity on cellulosic materials. Eur J Biochem 269: 4202-4211.

Sambrook J, Russel DB 2001. Molecular cloning: A laboratory manual. Cold spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Stephanopoulos G 2007. Challenges in engineering microbes for biofuels production. Science 315: 801-803.

Van Rantwijk F, Lau RM, Sheldon RA 2003. Biocatalytic transformations in ionic liquids. Trends Biotechnol 21: 131-138.

Van Rooyen R, Hahn-Hägerdal B, La Grange DC, Van Zyl WH 2005. Construction of cellobiose-growing and fermenting *Saccharomyces cerevisiae* strains. J Biotechnol 120: 284-295.

Van Wyk N, Den Haan R, Van Zyl WH 2010a. Co-expression of *Thermobifida fusca cel9A* with other cellulases in *Saccharomyces cerevisiae. Submitted to Applied and Environmental Microbiology.*

Van Wyk N, Den Haan R, Van Zyl WH 2010b. Heterologous production of NpCel6A from *Neocallimastix particiarum. Accepted for publication in Enzyme and Microbial Technology.*

Van Zyl WH, Lynd LR, Den Haan R, McBride JE 2007. Consolidated bioprocessing for ethanol production using *Saccharomyces cerevisiae*. Adv Biochem Eng Biot 108: 205-235.

CHAPTER 8

GENERAL CONCLUSIONS AND COMMENTARY

Projected models have shown that consolidated bioprocessing (CBP) could dramatically reduce the cost involved in the large-scale production of ethanol from lignocellulose (Lynd et al., 1996). This would ease the inevitable transition for transport sectors world-wide to shift from being petroleum-based to biofuels-based – in particular ethanol. A biocatalyst is needed that is able to depolymerize the complex carbohydrates found in lignocellulose (cellulose and hemicellulose) in an efficient manner and convert the resulting monomers (C6 and C5 sugars) to ethanol via fermentation. A main obstacle for this process is addressing the recalcitrance of lignocellulose – especially the cellulose component. The only depolymerizing enzymes that have a dramatic effect on the disruption of crystalline cellulose are the β-(1,4)-exoglucanases such as cellobiohydrolases. It is thus important that the biocatalyst in question should be able to produce substantial amounts of β-(1,4)-exoglucanases. As *S. cerevisiae* shows tremendous potential as a candidate CBP-biocatalyst, it is of great interest to investigate the heterologous production of exoglucanases in this organism.

The main aim of this study was to construct recombinant *S. cerevisiae* strains with exoglucanase-coding genes and explore different ways to improve on the cellulolytic capability of *S. cerevisiae*.

From the data presented in this study, the following can be concluded:

- The *cel9A* of *Thermobifida fusca* could be successfully expressed in *S. cerevisiae.*
- *T. fusca* Cel9A, despite being hyperglycosylated by *S. cerevisiae,* was still functional.
- Growth on phosphoric acid swollen cellulose as sole carbohydrate source was shown by *S. cerevisiae* that only express *cel9A.*
- That synergy between *T. fusca cel9A* and other cellulases resulted in an increase in Avicel hydrolysis when co-expressed in *S. cerevisiae.*
- Co-expression with other cellulases did not decrease Cel9A production in *S. cerevisiae.*
- The *Npcel6A* of *Neocallimastix patriciarum* was successfully expressed in *S. cerevisiae.*

- ∠ Recombinant NpCel6A showed pronounced activity on a soluble cellulose source, but little on a crystalline cellulose source.
- ∠ It was presumed that glycosylation of its catalytic domain diminished NpCel6A activity on insoluble cellulose.
- ∠ NpCel6A with its carbohydrate-binding module (CBM) and Asn-rich linker replaced with the CBM and Ser/Thr-rich linker of TrCel6A of *Trichoderma reesei* was shown to yield significantly decreased cellulase activity when expressed in *S. cerevisiae.*
- ∠ TrCel6A with its carbohydrate-binding module (CBM) and Ser/Thr-rich linker replaced with the CBM and Asn-rich linker of NpCel6A of *N. patriciarum* was shown to have significantly increased cellulase activity when expressed in *S. cerevisiae.*
- ∠ Overexpression of the native Protein secretion enhancer 1 gene (*PSE1*) led to an increase in heterologous production of NpCel6A, Cel3A of *Saccharomycopsis fibuligera* and Cel7B of *Trichoderma reesei* in *S. cerevisiae.*
- ∠ Overexpression of the native Superoxide dismutase 1 gene (*SOD1*) led to an increase in heterologous production of NpCel6A, Cel3A of *Saccharomycopsis fibuligera* and Cel7B of *Trichoderma reesei* in *S. cerevisiae.*
- ∠ Disruption of the *DDI1* gene led to an increase in normalized heterologous production of NpCel6A in *S. cerevisiae*.
- ∠ The swelling factor protein swollenin (*swoI*) of *T. reesei* could be functionally expressed in *S. cerevisiae*.
- ∠ Co-expression of *swoI* with cellulases could not be shown to improve the cellulolytic capability of *S. cerevisiae.*

6.1. COMMENTARY

In addition to the abovementioned conclusions, comments on certain aspects regarding the study will follow.

CODON OPTIMIZATION

It is generally regarded that codon usage bias is an important bottleneck for heterologous protein production – particularly for bacterial hosts (Lithwick & Margalit, 2003). Thus *cel9A* of *Thermobifida fusca* with its unfavourable codon usage bias for expression in yeast should in theory be a difficult protein to be produced by *S. cerevisiae*, but relatively high titers of the protein were obtained. It shows that the emphasis put on codon usage bias for heterologous protein production in yeast may not be as high as for bacteria.

Ed Rybicki of the Institute of Infectious Disease and Molecular Medicine, University of Cape Town, South Africa (personal communication, 2008) demonstrated that genes coding virus proteins with a higher GC content yielded higher titers when expressed in tobacco plants than a gene coding the same protein, but with a lower GC content. Since *T. fusca cel9A* also has a high GC content (68%), it might be an interesting future endeavour to compare expression levels of genes that code for the same protein, but with completely different GC contents in *S. cerevisiae*. It has been shown previously that "preferred" codons in yeast genes that are highly expressed have a GC content comparable to the mean content for the whole genome (Sharp & Lloyd, 1993) implying that GC content of a gene does not play a large role in whether or not a gene is highly expressed. But, in contrast, work has been done that shows that there is not a good correlation between predicted and actual expression data while only relying on characteristics of coding DNA sequences (Friberg et al., 2004).

A question arises of whether one should express a codon-optimized, which also implies synthetically-made, version of the gene as oppose to the native version – especially for *S. cerevisiae*. If the opportunity arises to synthetically design a gene, and codon usage is the only criteria, codon-optimization might be the prudent choice since one would eliminate any possibility that unfavourable codon usage could have on eventual heterologous expression of the gene.

However, it is debatable what is considered codon-optimized since there are various models that one can follow. Two of the most common are the codon bias index (CBI) (Bennetzen & Hall, 1982) and the codon adaptation index (CAI) (Sharp & Li, 1987). In short, CBI considers the codon usage of the whole genome of an organism whereas the CAI considers only the codon usage of the most expressed genes. Besides codon indices, another important factor is that of the codon context of a gene, as it considers the primary structure of mRNA of a transcribed ORF and to what degree certain mRNA codons next to each other could lead to translational inaccuracies (Moura et al., 2005). One must also contemplate whether one should use the term "codon-optimized" if it is not known if the expression of such a gene in a recombinant host is by any means better than that of the native gene. Perhaps the term "codon-altered" could be a better alternative.

The first genes to be heterologously expressed in the late 1970s were in fact synthetically designed (Goeddel et al., 1979). Due to the financial implications involved in synthesizing such genes, most research teams opted for expressing native genes in recombinant hosts. However, in recent times technologies became available that dramatically lowered the cost for the synthesis of such genes. Thus, if the primary aim of a project is to express and characterize a gene in a recombinant host and the gene is not freely available, it would be a time-saving and most likely money-saving exercise to synthesize the gene. Obtaining a gene from the native organism, constructing gene libraries or undergoing material transfer agreements with a research group that possesses the given gene might in many cases lead to a drawn-out process that may not yield the desired result. It was for the abovementioned reasons that the synthesis of a codon-altered gene was chosen for the study of the expression of the *Neocallimastix patriciarum cel6A* gene in *S. cerevisiae*.

GLYCOSYLATION OF HETEROLOGOUS PROTEINS

Arguably, the most important post-translational modification of secreted proteins in *S. cerevisiae* is glycosylation. It is well-known that *S. cerevisiae* tends to hyperglycosylate *N*-glycosylation sites on proteins destined for secretion. It is easy to predict the likeliness of a

given protein to be glycosylated as there are useful software available on the Expasy search engine (NetNGlc 1.0 and NetOGlc 1.0). What is not possible is to predict what effect glycosylation could have on a protein – especially if the three-dimensional (3D) structure of the protein is not available. One could assume from a 3D structure that if the possible glycosylation should occur in an area of the protein where functional domains are situated, that the protein's functionality could be severely reduced. If the possible glycosylation should occur in regions distant from functional domains, a negligible effect could be expected. In this study, both the Cel9A and NpCel6A were shown to be glycosylated by *S. cerevisiae*. Cel9A was shown to be functional and glycosylation on this protein was not predicted in its catalytic domain. NpCel6A was shown to have reduced processive capability and glycosylation was predicted to occur in its catalytic domain. It is thus extremely important to identify sites of glycosylation before expression studies in *S. cerevisiae* should commence. Caution should thus be taken if glycosylation is predicted within the catalytic domain of the protein. Removal of the possible glycosylation sites could have an adverse effect on the protein function, as the residues involved; asparagine, serine and threonine are all polar and might have a non-specific role in the catalytic domain. It would thus be a more feasible option if other functionally similar genes with no glycosylation sites within their catalytic domain could be used instead.

Although it appears that glycosylation of heterologously-produced enzymes has either a negative of negligible effect, many studies have shown the opposite. Glycosylation was shown to increase the stability of enzymes (Imperiali & Rickert, 1995), minimize proteolytic cleavage (Qin et al., 2008), improve proteosolubility (Murakami-Sekimata et al., 2009) and increase pH range (Lee et al., 2003). It has even been shown that the addition of either N-terminal or C-terminal N-glycosylation sites on heterologous proteins resulted in improved secretion (Sagt et al., 2000). Thus, considering the glycosylation of a given protein in *S. cerevisiae* is imperative before embarking on heterologous expression studies as it would most likely have a dramatic impact on the eventual outcome.

SYNERGY AMONG HETEROLOGOUSLY-PRODUCED CELLULASES

Synergy among cellulases in an important, if not essential, component of crystalline cellulose hydrolysis. It has been shown in numerous studies that the ratio between the cellulases is central for maximal hydrolysis (Kim et al., 1998; Irwin et al., 1993). The cellulase that is in most cases required in the highest proportion for efficient hydrolysis of crystalline cellulose is the exoglucanase or cellobiohydrolase. Thus, for synergy to play an important role in the recombinant cellulolytic *S. cerevisiae*, exoglucanases need to be produced at a higher level than endoglucanases and β-glucosidases. A strategy that can be employed, is to integrate the expression cassettes containing the endoglucanase and β-glucosidase and to place the exoglucanase(s)-containing cassettes on an episomal plasmid which would yield a higher copy number. It is important that the yeast cell benefit from the heterologous production of the given enzyme. For instance, if large levels of β-glucosidase enzymes are produced, but no substrate is available i.e. the products of endo-exoglucanase action, then the cell spent energy on making an enzyme that it would not directly benefit from. In most heterologous expression studies, this would not be a problem since the fate of the producer is not of concern, but for the development of a CBP organism, the producer serves as a biocatalyst as it converts the end product of cellulose hydrolysis to ethanol – the final end product.

To determine what each cellulase contributes toward cellulose hydrolysis when produced by recombinant *S. cerevisiae* is problematic due to synergy and with the fact that each type of cellulase has some individual activity on any kind of cellulose substrate. The amount of enzyme produced can be determined with an ELISA-type experiment. It is important to determine the amount of each cellulase enzyme produced by recombinant *S. cerevisiae* as it would reveal which are not produced in enough quantities when co-expressed. Raising antibodies for each cellulase would be cumbersome, but a strategy which could eliminate the amount of antibodies needed to be raised, is to target the CBM of the enzyme since most – especially eukaryotic – cellulases have CBMs that are homologous to one another.

At the moment, to ensure production of cellulases, all the cellulases-coding genes are placed under similar constitutive regulation since inducible promoters of *S. cerevisiae* (the *GAL* family) cannot be used on a large scale due to the cost of inducers. To draw the maximum effect of synergy by regulating the ratio of the different cellulases being produced by the recombinant host is a challenging problem to overcome since *S. cerevisiae* lacks the native capability to regulate cellulose hydrolysis.

PREDICTION OF SUCCESSFUL EXPRESSION OF GENES IN A RECOMBINANT HOST

There are many heterologous genes that for some unknown reason perform extremely well in *S. cerevisiae* whereas functionally similar genes do not yield nearly the same results. A clear example is that of β-glucosidase of *Saccharomycopsis fibuligera* (Van Rooyen et al., 2005). In this particular study, several other β-glucosidase-coding genes were also expressed in *S. cerevisiae*, but did not show nearly the same level of enzyme production. Currently, no accurate model seems to exist that can predict how well a heterologous gene will be transcribed, translated, modified and secreted in *S. cerevisiae*. It is unfortunate that "negative results" of expression studies where extremely low or no yield of the desired product were found, are rarely published as it might give a better understanding of what characteristics are preferred (or rather not preferred), by *S. cerevisiae*.

There does not appear to be any good correlation between the specific activities of a particular enzyme and the normalized activities obtained when expressed in a recombinant host: enzymes with lower specific activities might perform better when expressed recombinantly and the reasons are not always clear. Thus, it seems that the most feasible option for the construction of a cellulolytic *S. cerevisiae* is to work on the trial-and-error principle: express a wide variety of cellulase-coding genes, assay transformants and select the best – without necessarily understanding why certain genes work better.

IMPROVEMENT ON THE YEAST CHASSIS

The heterologous expression of a multitude of enzymes is needed for complete hydrolysis of lignocellulose. It is highly unlikely that even when all required hydrolase-coding genes (which include the hemicellulases) are inserted in recombinant *S. cerevisiae* that it would be sufficient for a feasible hydrolysis rate for the conversion of lignocellulose to ethanol. It is thus important to consider the yeast strain that was constructed as a chassis strain and investigate several ways of improving on the chassis.

These improvements – known as metabolic engineering – could include the overexpression and deletion or disruption of native genes – as was done in this study. With regards to overexpression, yeasts and all organisms have an innate way of buffering any unexpected increases of a gene product. It is thus important to test whether there is in fact a difference in phenotype when overexpression occurs.

Regarding integration of heterologous genes, an increase in copy number alone is not the only factor to be considered to ensure an increased effect in heterologous protein production. The loci of the inserted gene cassettes are just as important since certain regions in the genome are far better transcibed than others. Thus, determining the copy number by means of Southern hybridization or real-time PCR should be accompanied with data that would show where the integration of the overexpression gene cassettes took place.

In this study noteworthy increases of up to two-fold compared to the parental strain in cellulase activities were detected when just one gene was overexpressed. This holds tremendous promise for future studies where whole pathways could be altered. Even though altering the metabolism to increase heterologous protein production could contribute greatly towards improving yeast strain development for consolidated bioprocessing, one must not lose sight of the primary goal of producing ethanol. Thus, one should steer away from altering the yeast's metabolism if it would lead to a dramatic reduction of the yeast's capability to produce ethanol or leads to a decrease in ethanol tolerance.

Yeast strain development for consolidated bioprocessing is in an exciting stage as there is still a multitude of genes and gene products that need to be considered that might add value to the already growing chassis. As there are around 6000 genes in *S. cerevisiae* (Goffeau et al., 1996), with many gene products still unassigned, the scope of improving the yeast chassis remains vast.

6.2. LITERATURE CITED

Bennetzen JL, Hall BD 1982. Codon selection in yeast. J Biol Chem 257: 3026-3031.

Friberg M, Von Rohr P, Gonnet G 2004. Limitations of codon adaptation index and other coding DNA-based features for prediction of protein expression in *Saccharomyces cerevisiae.* Yeast 21: 1083-1093.

Goeddel DV, Kleid DG, Bolivar F, Heyneker HL, Yansura DG, Crea R, Hirose T, Kraszewski A, Itakura K, Riggs AD 1979. Expression in *Escherichia coli* of chemically synthesized genes for human insulin. Proc Natl Acad Sci USA 76: 106-110.

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG 1996. Life with 6000 genes. Science 274: 563-567.

Imperiali B, Rickert KWC 1995. Conformational implications of asparagine-linked glycosylation. Proc Natl Acad Sci USA 92: 97-101.

Irwin DC, Spezio M, Walker LP, Wilson DB 1993. Activity studies of eight purified cellulases: specificity, synergism, and binding domain effects. Biotechnol Bioeng 42: 1002-1013.

Kim E, Irwin DC, Walker LP, Wilson DB 1998. Factorial optimization of a six-cellulase mixture. Biotechnol Bioeng 58: 494-501.

Lee J, Park JS, Moon JY, Kim KY, Moon HM 2003. The influence of glycosylation on secretion, stability, and immunogenicity of recombinant HBV pre-S antigen synthesized in *Saccharomyces cerevisiae*. Biochem Biophys Res Commun 30: 427-432.

Lithwick G, Margalit H 2003. Hierarchy of sequence-dependent features associated with prokaryotic translation. Genome Res 13: 2665-2673.

Lynd LR, Elander RT, Wyman CE 1996. Likely features and costs of mature biomass ethanol technology. Appl Biochem Biotechnol 57/58: 741-761.

Moura G, Pinheiro M, Silva R, Miranda I, Afreixo V, Dias G, Freitas A, Oliveira JL, Santos MAS 2005. Comparative context analysis of codon pairs on an ORFeome scale. Genome Biol 6: R28(14).

Murakami-Sekimata A, Sato K, Sato K, Takashima A, Nakano A 2009. *O*-Mannosylation is required for the solubilization of heterologously expressed human β-amyloid precursor protein in *Saccharomyces cerevisiae*. Genes to Cells 14: 205-215.

Qin Y, Wei X, Liu X, Wang T, Qu Y 2008. Purification and characterization of recombinant endoglucanase of *Trichoderma reesei* expressed in *Saccharomyces cerevisiae* with higher glycosylation and stability. Protein Expres Purif 58: 162-167.

Sagt CMJ, Kleizen B, Verdwaal R, De Jong MDM, Muller WH, Smits A, Visser C, Boonstra J, Verkleij, AJ, Verrips CT 2000. Introduction of an *N*-glycosylation site increases secretion of heterologous proteins in yeasts. Appl Environ Microbiol 66: 4940-4944.

Sharp PM, Lloyd. 1993. Regional base composition variation along yeast chromosome III: evolution of chromosome primary structure. Nucleic Acids Res 21: 179-183.

Sharp PM, Li WH 1987. The codon adaptation index - a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res 15: 1281-1295.

Van Rooyen R, Hahn-Hägerdal B, La Grange DC, Van Zyl WH 2005. Construction of cellobiose-growing and fermenting *Saccharomyces cerevisiae* strains. J Biotechnol 120: 284-295.

APPENDIX A

TITLE OF INVENTION:

Method for fermenting cellulosics

FIELD OF INVENTION:

The present invention relates to a method for fermenting cellulosics.

More particularly, the invention relates to a method for fermenting amorphous cellulosics using a yeast strain by introducing a recombinant gene into the yeast's genome and improving the cellulosics-degrading capability by introducing more recombinant genes into the yeast's genome.

BACKGROUND OF INVENTION:

Policy-makers world-wide are moving towards managing energy reserves more efficiently [1]. Increased emphasis is being put on making renewable energy supply more sustainable and to reduce the amount of carbon emissions in the atmosphere. This would imply that the use of petroleum as main energy source should be reduced and a serious shift to alternatives should be encouraged. These alternative energy sources include solar power [2], wind power [3] wave power [4] and energy derived from biologically mediated events like methane, hydrogen, biodiesel and ethanol [5]. Ethanol, or bioethanol, has received a lot of attention recently and has been identified as a promising replacement candidate for petroleum as a transport fuel [6]. Ethanol is largely made via the fermentation of monomeric sugars and usually Baker's yeast, *Saccharomyces cerevisiae,* is employed in the large-scale production of ethanol. There are many reasons why *S. cerevisiae* is such a popular choice [7]: its behaviour in large-scale bioreactors is well-studied; it is able to convert certain simple sugars to ethanol at high rates and efficiencies. Furthermore, it is not susceptible to certain inhibitory compounds and high ethanol concentrations and enjoys GRAS (generally recognised as safe) status. One limitation of

S. *cerevisiae* is its narrow range of substrate utilisation. For instance, it is unable to utilise and convert the complex polysaccharides found in lignocellulose to ethanol.

Lignocellulose is the collective name for the structural component of all plant biomass [8]. The composition of lignocellulose differs throughout plant taxa, but in general consists of 35-50 % cellulose, 20-35 % hemicellulose and 15-20 % lignin. It could potentially be beneficial to develop recombinant yeast that is able to produce the enzymes needed for the degradation of lignocellulose to monomeric sugars. This would allow the yeast to grow on the resultant monomeric sugars and ferment it to ethanol. This would make significant strides towards the development of an organism that can conduct consolidated bioprocessing (CBP) of lignocellulose to ethanol. CBP is the one-step conversion of a renewable substrate like lignocellulose to a desired product (in this case ethanol) and has the promise of reducing the costs involved in the large-scale conversion of lignocellulose to ethanol [9].

This project will concentrate on the degradation of the major component of lignocellulose i.e. cellulose by recombinant *S. cerevisiae.* Cellulose is a homopolymer composed of D-glucose units linked to each other via β-glycosidic bonds [10]. Due to intra- and interchain hydrogen bonding and Van der Waal interactions, cellulose is water-insoluble and extremely recalcitrant toward enzymatic degradation. There are several cellulolytic organisms that produce enzymes that can degrade cellulose and these enzymes are collectively called cellulases. Cellulases are traditionally split into three types: (1) exoglucanases, including cellodextrinases (1,4-β-D-glucan glucanohydrolases; EC 3.2.1.74) and cellobiohydrolases (1,4-β-D-glucan cellobiohydrolases; EC 3.2.1.91), (2) endoglucanases (1,4-β-D-glucan 4-glucanohydrolases; EC 3.2.1.4) and (3) β-glucosidases (β-glucoside glucohydrolases; EC 3.2.1.21) [11]. Exoglucanases are capable of attaching to the end of a cellulose chain and, in a processive manner, cleave off small soluble oligosaccharides. The majority of exoglucanases are cellobiohydrolases and almost exclusively cleave off cellobiose from the cellulose chain. Exoglucanases are of particular interest since they disrupt the crystalline nature of cellulose which is the most predominant state of cellulose found in nature with only a small proportion being amorphous. Exoglucanases work in concert with endoglucanases and β-glucosidases to degrade cellulose. Endoglucanases, which only

degrade the amorphous regions of cellulose, would generate new sites for attachment for the exoglucanases; exoglucanases disrupt the crystallinity of cellulose and generate more amorphous regions that endoglucanases could attack. β-Glucosidases cleave cellobiose, the major end product of exoglucanase-endoglucanase attack, to two glucose units and thus diminish the effect of end product inhibition of cellulases generated by high cellobiose concentrations [12]. Several unique enzymes have also been identified that are involved in cellulose degradation. Cel9A of the thermophilic actinomycete *Thermobifida fusca* is one of them in that it displays both endoglucanase and exoglucanase activity and is known in literature as a processive endoglucanase [14]. Both the specific endoglucanase and exoglucanase activity of Cel9A is not as high as other cellulases, but this enzyme has shown significant amount of synergy with both endoglucanases and exoglucanases.

Several attempts have been made to express exoglucanases in *S. cerevisiae*. In particular, Den Haan and colleagues [13] have expressed and secreted four cellobiohydrolases of fungal origin in *S. cerevisiae*. They have found disappointingly low levels of secretion of these enzymes and have calculated that an increase of at least 18 fold on the current levels of cellobiohydrolases is needed to enable growth of *S. cerevisiae* on a crystalline cellulose substrate. An increase could be obtained in several ways: codon-optimisation of the exoglucanase genes for more efficient expression; manipulation of the yeast in such a way that its secretion pathway could handle a higher through-put of enzymes to be secreted; and attempting the expression of other exoglucanases with possible increased specific activities.

It is an object of this invention to illustrate a novel method for the direct fermentation of amorphous cellulosics to ethanol by using a recombinant *S. cerevisiae* strain that expresses a processive endoglucanase.

This method is unlike that described by Fujita *et al*., [14], where three cellulases, *Trichoderma reesei* endoglucanase II and cellobiohydrolase II and *Aspergillus aculeatus* β-glucosidase I, were tethered on the yeast cell surface to generate a whole-cell biocatalyst for amorphous cellulose

conversion. The limiting feature of their approach was that cells had to be pre-grown to a high cell density before amorphous cellulose conversion to ethanol could occur with biomass loadings of 15 g/L.

It also unlike the method of Den Haan et al. [15], where two recombinant genes, the endo-1,4-β-endoglucanase I from *Trichoderma reesei* and β-glucosidase I from *Saccharomycopsis fibuligera,* where simultaneously expressed in *S. cerevisiae*. The recombinant yeast co-expressing the two cellulases was able to multiply on amorphous cellulose and concomitantly produce ethanol to a maximum of 1 g/L.

SUMMARY OF INVENTION

According to a first embodiment of the invention, there is provided a method of modifying yeast so as to enable growth on insoluble amorphous cellulosics, the method including the steps of:

transforming the yeast with a DNA sequence encoding the processive endoglucanase CEL9A of *Thermobifida fusca*; and

causing the yeast to express and secrete the mature *Thermobifida fusca* processive endoglucanase CEL9A heterologously.

The preferred yeasts as host cells may belong to the genera *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces*, and *Yarrowia*. Particularly preferred yeast species as host cells include *S. cerevisiae, S. bulderi, S. barnetti, S. exiguus, S. uvarum, S. diastaticus, S. carlsbergensis, K. lactis, K. marxianus,* and *K. fragilis*.

According to a second embodiment of the invention, there is provided an expression cassette containing a constitutive yeast promoter and a DNA sequence encoding a mature processive endoglucanse CEL9A.

The promoter sequence may be a *S. cerevisiae* phosphoglycerate kinase 1 (*PGK1*) gene promoter, the *S. cerevisiae* enolase 1 (*ENO1*) gene promoter or the *S. cerevisiae* alcohol dehydrogenase 2 (*ADH2*) gene promoter.

The native secretion signal of the *Thermobifida fusca Cel9A* gene may be replaced with the *Trichoderma reesei* β-xylanase 2 secretion signal to ensure efficient secretion of the mature enzyme from yeast strains.

The secretion signal fused to the gene sequence of *Thermobifida fusca Cel9A* may be the secretion signal of *K. marxianus* inulinase 1 (*INU1*), *S. cerevisiae* invertase 1 (*SUC2*), *S. cerevisiae* acid phosphatase (*PHO5*), *S. cerevisiae* mating type factor alpha (MFα) and *S. carlbergensis* alpha-galactosidase (*MEL1*).

The gene sequence may be from the glycosyl hydrolase family 9 of actinomycetes, including the genera *Streptomyces* and *Thermobifida*.

The gene sequence may be codon-optimised for efficient production in yeast, more particularly for *S. cerevisiae*.

According to a third embodiment of the invention there is provided a vector including the expression cassette.

The vector may be transformed into yeast, such as *Saccharomyces* strains, and maintained episomally. The vector may be a plasmid, and in particular, a multicopy, episomal plasmid such as yXS+Cel9A.

According to a fourth embodiment of the invention, there is provided a yeast strain transformed with the expression vector described above.

The yeast strain may be *S. cerevisiae* Y294 into which the expression vector has been chromosomally integrated.

According to a fifth embodiment of the invention, there is provided a yeast strain capable of hydrolyzing amorphous cellulosics

The transformed yeast strain may have the capacity to cleave off sufficient glucose from the amorphous cellulose chain to enable growth on amorphous cellulosics as sole carbohydrate source when compared to an untransformed yeast cell that does not grow significantly on the same medium.

According to a sixth embodiment of the invention, there is provided yeast strains with improved cellulolytic capability than of abovementioned strain

The strain may contain and produce the processive endoglucanase Cel9A along with any other cellulase; whether it is an endoglucanse, exoglucanase or β-glucosidase. More, particularly Cel9A may be co-expressed in a recombinant host along with *T. reesei* (teleomorph: *Hypocrea jadrii*) Cel5A, Cel6A, Cel7A or Cel7B.

DETAILED DESCRIPTION OF AN EXAMPLE OF AN EXPERIMENT THAT INCLUDES GRAPHS AND TABLES

GRAPH 1: Schematic representation of the plasmids used and constructed in this experiment. *ENO1*$_P$, *ENO1*$_T$, *PGK1*$_P$, *PGK1*$_T$, represents the *S. cerevisiae* enolase I and phosphoglycerate kinase I promoter and terminator DNA sequences respectively. *TrXYN2SEC* represents the 23-amino acid secretion signal of the xylanase II gene of *T. reesei.* TrCel7A represents cellobiohydrolase I (CBH1) of *Trichoderma reesei*. TrCel6A represents cellobiohydrolase II (CBH2) of *Trichoderma reesei*. TrCel7B represents endoglucanase I (EGI) of *Trichoderma reesei*. TrCel5A represents endoglucanase II (EGII) of *Trichoderma reesei*. TfCel9A represents the

processive endoglucanse of *Thermobifida fusca*. Relevant restriction endonucleases are also shown. All plasmids constructed are *E. coli/S. cerevisiae* shuttle vectors and also contains (1) a β-lactamase (*bla*) gene which acts as a selectable marker in *E. coli*; (2) an *ori* region which is responsible for replication of the plasmid in *E. coli*; (3) a 2µ sequence which is responsible for episomal replication of the plasmid in *S. cerevisiae* (4) and the *S. cerevisiae* orotidine-5'-phosphate decarboxylase gene (*URA3*) which is used as selectable marker in *S. cerevisiae*.

GRAPH 2: Typical HPAEC chromatogram of products released from an insoluble cellulose substrate, in this case BMCC, by Cel9A expressed in *S. cerevisiae.* Y294[Cel9A]. Cultures were grown for five days and the Cel9A enzymes secreted in the supernatant were absorbed to 2.5% BMCC. After overnight stirring at 4°C the cellulose-bound Cel9A was incubated at 37°C to allow for enzymatic hydrolysis and samples were taken at regular intervals and analysed for glucose and cellobiose levels. Results for other cellulosic substrates (PASC and Avicel) were similar.

GRAPH 3: *Saccharomyces cerevisiae* Y294[CEL9A] growing on yeast peptone containing 1% phosphoric acid swollen cellulose for four days at 30°C. The reference strain Y294[REF] showed no significant growth on the same media.

GRAPH 4: Aerobic growth curve of Y294[Cel9A] on 1% phosphoric acid swollen cellulose supplemented with 2% peptone and 1% yeast extract.

GRAPH 5: (a) Carboxymethyl-cellulose plate assays on *S. cerevisiae* Y294 expressing individual cellulases and the combinations of Cel9A with other cellulases. Yeast colonies were grown for 2 days at 30°C on CMC-containing plates and washed off. Plates were stained with 0.1% Congo Red and destained with 1 M NaCl. (b) Ostazin Brilliant Red-dyed hydroxyethyl-cellulose plate assays on *S. cerevisiae* Y294 expressing individual cellulases and the combinations of Cel9A with other cellulases. Yeast colonies were grown for 3 days at 30°C before photograph was taken.

GRAPH 6: HPLC chromatogram of a typical sample of anaerobic growth of Y294[CEL9A]. External standards for ethanol and cellobiose were used and peaks for these standards are shown.

TABLE 1: Microbial strains and plasmids used in this experiment.

TABLE 2: PCR primers used for gene isolation and plasmid construction in this experiment.

TABLE 3: Specific cellulase activity profile of Cel9A expressed in the *S. cerevisiae* strain Y294. Activity was determined as Units/mg cell dry weight. A unit is equal to 1 μmol of monomeric sugar released per minute per milliliter

Media and culture conditions

The chemicals and media components used in experiments were of laboratory grade standard. Phosphoric acid swollen cellulose (PASC) was prepared as previously described [16] using Avicel PH-101 (Fluka) as crystalline source. The *Escherichia coli* strain DH5α was used for plasmid transformation and propagation. *E. coli* cells were cultivated in Luria Broth medium (5 g/L yeast extract, 10 g/L NaCl, 10 g/L tryptone) supplemented with ampicillin (100 mg/L). *S. cerevisiae* Y294 transformants were selected and maintained on SC^{-URA} or $SC^{-URA-LEU}$ medium plates (1.7 g/L yeast nitrogen base without amino acids and ammonium sulphate [Difco laboratories, Detroit, MI, USA], 5 g/L $(NH_4)_2SO_4$, 20 g/L glucose, 20 g/L agar and supplemented with amino acids as required). Autoselective *S. cerevisiae* strains were cultured in YPD medium (10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose). The recombinant strain Y294[CEL9A] and the reference recombinant strain Y294[REF] were cultured on YP-PASC medium (10 g/L yeast extract, 20 g/L peptone, 10 g/ L PASC). Yeast strains were routinely cultured in 250-mL Erlenmeyer flasks containing 50 mL medium at 30°C, on a rotary shaker at 100 rpm. For aerobic growth experiments yeast strains were grown in baffled 250-mL Erlenmeyer flasks containing 50 mL growth medium. For anaerobic fermentation yeast strains were grown in rubber plugged 120-mL glass serum bottles containing 120 mL YP-PASC medium supplemented with 0.01 g/L ergosterol and 0.42 g/L Tween 80 [17]. Precultures of the strains were grown on YPD medium. For growth on liquid YP-PASC medium three cultures the strains tested were inoculated

simultaneously. Samples were periodically taken and yeast cells in the media were counted in duplicate on a haemocytometer.

Microbial strains and plasmids

The genotypes and sources of the yeast and bacterial strains and the plasmids that were constructed and used in this study, are summarized in Table 1. A diagrammatic representation of the plasmids that were transformed into yeast is given in Graph 1.

Nucleic acid manipulations and transformation

Standard procedures for DNA manipulation were followed as outlined by Sambrook et al. [18]. All restriction endonuclease enzymes were purchased from Fermentas. For polymerase chain reactions, the TAKARA ExTaq polymerase was used as recommended by the manufacturer with a Perkin Elmer GeneAmp® PCR System 2400 (The Perkin-Elmer Corporation, 30 Norwalk, CT, USA). Details of all the primers used in this experiment are given in Table 2. Primers TFuscel90-left and Tfuscel90-right were used to retrieve the *cel9A* gene of *T. fusca* from the plasmid D1020 (donated by David Wilson, Cornell University). A subsequent PCR product was cloned into the TA-cloning vector pGEM (Promega) – generating pGEM-Cel9A. The pGEM-Cel9A construct was digested with the restriction endonucleases *Nru*I and *Xho*I (Fermentas) to obtain a 2518-bp fragment. This fragment was inserted in frame with a *xyn2* secretion signal of *Trichoderma reesei* on the *E. coli/S. cerevisiae* shuttle vector yXYNSEC and named yXS+Cel9A. The complete $PGK1_P$-XYNSEC-*Cel9A*-$PGK1_T$ gene cassette was amplified with PGK1-left and PGK1-right primers as a 3566-bp fragment and inserted into the TA-cloning vector pTZ57R/T (Fermentas) – generating pTZ-Cel9A. The pTZ-Cel9A construct was digested with *Hin*dIII (Fermentas) and the $PGK1_P$-XYNSEC-*Cel9A*-$PGK1_T$ cassette was inserted in the plasmids pAZ41 [15] and pegII. The final plasmids constructs were named pCel9A+egI and pCel9A+egII respectively. The gene cassettes $ENO1_P$-*Cel6A*-$ENO1_T$ and $ENO1_P$-*Cel7B*-$ENO1_T$ on the plasmids pAZ21 and pAZ22 [13] were amplified using the primers ENO1-left and ENO1-right resulting in fragments of the sizes 2542 bp and 2498 bp respectively. The fragments were inserted into pTZ57R/T and renamed pTZ-CBH2 and pTZ-CBH1. Both pTZ-CBH2 and pTZ-CBH1 were partially digested with *Hin*dIII and

the respected fragments were inserted into yXS+Cel9A resulting in plasmids pCel9A+CBH2 and pCel9A+CBH1.

Activity assays

Cells were cultured in concentrated YPD media for four days after which cells were spun down at 2,000 rpm for 3 min. One hundred and thirty milliliters of the supernatant were filtered through a glass fiber filter (Whatman) to remove any remaining yeast cells. To absorb the cellulases present in the supernatant, cellulose (10 g/L PASC, 25 g/L BMCC and 25g/L Avicel) were added to a final concentration of 0.014%, 0.042% and 0.042% respectively along with 0.02% sodium azide. The cellulases were allowed to bind to the cellulose by incubating the solution overnight at 4 °C with continuous stirring. Cellulose with the attached exocellulases were captured on a glass fiber filter and washed with 20 mL of buffer (0.05 M NaOAc + 0.015 M $CaCl_2$, pH 5.5 for Cel9A and 0.05 M citrate buffer for Cel6A). Filters were put in capped tubes with 6 mL of respective buffers along with 0.02 % sodium azide. The capped tubes and its content were vortexed and a 1 mL sample was removed and used as a zero hour blank. The remainder was incubated at 55°C in a hybridization oven and shaked at a rate of 70 strokes per minute. One milliliter of the reaction volume were taken after two hours and centrifuged at 13,000 rpm for 10 min. Samples were then assayed for total sugar content by the phenol-sulfuric assay at 490 nm [19]. Cellobiose (Sigma) was used to make a linear standard curve.

To determine the endoglucanase activity of the recombinant Cel9A, the dinitrosalicyclic acid (DNS) method [20] for the detection of the amount of reducing sugars produced from carboxymethylcellulose (CMC) was used. One milliliter samples of cultures grown for four days were centrifuged for one minute at 13 K. Supernatants were appropriately diluted of which 50 µL were added to 450 µL of 1% CMC (Sigma) in a test tube, vortexed and incubated for 5 minutes at 55 °C. Seven hundred and fifty microliters of DNS were added and the samples were boiled at for 15 minutes. Activity was measured at 540 nm with glucose being used to construct a standard linear curve.

HPLC analyses

Samples were taken of concentrated enzyme solutions that were incubated at 37°C on an orbital shaker with a cellulose substrate (PASC, BMCC and Avicel) at time period 0, 2, 5 and 24h. These samples were analysed for glucose and cellobiose content by high-performance anion exchange chromatography with a Dionex DX 500 chromatography system coupled to a pulsed amperometric detector (PAD, Dionex ED 40) using a CarboPac PA-10 column (Dionex, Sunnyvale, CA, USA). Initially, the column was activated with 20mM NaOH. A flow-rate was maintained at 1 ml/min. After 25 minutes the NaOH concentration was increased to 225mM for 6 minutes when the NaOH was again decreased to 20mM. External standards (glucose and cellobiose) were used for any quantification.

During the anaerobic growth curve experiments samples were taken every twenty-four hours and the cellobiose and ethanol concentrations were determined with a Waters 717 injector (Milford, MA, USA) and Agilent 1100 pump (Palo Alto, CA, USA). The compounds were separated on an Aminex HPX-87H column (Bio-Rad, Richmond, CA), at a column temperature of 45°C with 5mM H_2SO_4 as mobile phase at a flow rate of 0.6 mL/min and subsequently detected with a Waters 410 refractive index detector.

RESULTS & DISCUSSION

Cel9A of the moderately thermophilic actinomycete *Thermobifida fusca* (previously known as *Thermomonospora fusca*) enzyme is unique among cellulases in that it exhibits both endoglucanase and exoglucanase activity [21]. Of all the cellulases characterised from *T. fusca*, Cel9A has the highest activity towards crystalline cellulose, especially BMCC [22]. Another distinguishing feature of Cel9A is the significant amount of synergy it has with both endoglucanases and exoglucanases [21].

In this study we have shown that the *Cel9A* gene sequence with its secretion signal replaced with that of the xylanase II gene (*xyn2*) of *Trichoderma reesei* (called XYNSEC) can be successfully produced by *S. cerevisiae* under the transcriptional control of the

phosphoglyceraldehyde kinase 1 (*PGK1*) promoter and terminator sequences. A major contribution to its successful expression could be assigned to its relatively high codon adaptation index for heterologous expression in *S. cerevisiae*, calculated as being 0.480, unusually high for bacterial genes expressed in *S. cerevisiae*. Cellulase activity was shown on both amorphous (CMC) and crystalline (Avicel) cellulose substrates. Its dual role as an endoglucanase and exoglucanases is thus maintained in the recombinant host. The specific activity on all the cellulosic substrates tested is, however, not as high as other cellulases expressed in *S. cerevisiae* (Table 3).

HPAEC on a Cel9A and cellulose (BMCC, Avicel and PASC) mixture incubated at 37°C revealed that both glucose and cellobiose are released from the cellulose chain (presumably from the non-reducing end) at detectable levels (Graph 2). The ratio of cellobiose to glucose cleaved off from the cellulose chain was calculated as being 2.367 ± 0.194 and was similar for all three substrates tested. This is an unusually low cellobiose/glucose ratio for a cellulase since most cellobiohydrolases would exclusively cleave off cellobiose from the cellulose chain.

The *S. cerevisiae* strain expressing *Cel9A* - Y294[CEL9A] - was shown to grow on YP-PASC agar plates (Graph 2). This indicates that the level of glucose being cleaved off by Cel9A was sufficient for growth on amorphous cellulose as sole carbohydrate source. A maximum specific growth rate (μ_{MAX}) of 0.088 h^{-1} was calculated during aerobic growth experiments. To our knowledge, this is the first account where significant growth was found by a *S. cerevisiae* strain on an insoluble cellulose substrate while expressing only one heterologous gene. This is also the first account where yeast growth on cellulose was observed without the need of heterologous β-glucosidase expression. In comparison, Van Dijken et al. [7] have shown that a prototropic, diploid yeast strain, CBS8066, with glucose as carbohydrate source reached a μ_{MAX} of 0.45. The Y294[CEL9A] strain constructed in this experiment performs exceptionally well if one considers the recalcitrant nature of the carbohydrate source in addition to the difference in rheology of the two growth substrates. During anaerobic growth, Y294[CEL9A] was able to produce ethanol reaching levels of more than 700mg/L (Graph 5). The cellobiose peak is a clear

indication of the presence and activity of Cel9A since no cellobiose is present in the media before inoculation and cellobiose cannot be utilised by the recombinant strain.

To improve the cellulolytic capability of Y294[CEL9A] and to investigate the extent that synergy can contribute to the improvement, the *cel9A* gene was co-expressed along with four other cellulases – all from the soft-rot fungus *Trichoderma reesei*. These cellulases: two endoglucanases *cel5A* (*egII*) and *cel7B* (*egI*), and two cellobiohydrolases *cel6A* (*cbh2*) and *cel7A* (*cbh1)* have already been individually expressed in *S. cerevisiae* [13, 15]. The *T. reesei* cellulase genes were all inserted under the transcriptional control of the Enolase 1 (*ENO1)* gene promoter and terminator sequences. Plate assays on CMC and Ostazin-linked hydroxyethylcellulose revealed that the strains where *cel9A* was co-expressed with the two cellobiohydrolases - Y294[CEL9A+CBH1] and Y294[CEL9A+CBH2] shown significant zones when compared to the strains only expressing an individual cellulase (Graph 6). This indicates that co-expression of two cellulases along with the synergy known to occur between Cel9A and other cellulases can increase the cellulolytic capability of *S. cerevisiae*.

In this invention, we were able to show that we can construct a yeast strain capable of growing on cellulose and concomitantly produce ethanol anaerobically. What makes this invention unique, is that the recombinant yeast used in the experiment is only heterologously expressing one gene – that of *cel9A* of *T. fusca*. The strain used, along with the strain co-expressing cellulase genes can contribute significantly toward the realisation of consolidated bioprocessing.

Literature list

[1] Schiermeier Q (2007) Europe moves to secure its future energy supplies. Nature **448**, 234-235.
[2] Şen Z (2005) Solar energy in progress and future research trends. Prog Energ Combust **30**, 367–416.

[3] Joselin Herbert GM, Iniyan S, Sreevalsan E, Rajapandian S (2007) A review of wind energy technologies. Renew Sust Energ Rev **11**, 1117–1145.

[4] Gross R, Leach M, Bauen A (2003) Progress in renewable energy. Environ Int **29**, 105– 122.

[5] Montgomery R (2004) Development of biobased products. Bioresource Technol **91**, 1–29.

[6] Lin Y, Tanaka S (2006) Ethanol fermentation from biomass resources: current state and prospects. Appl Microbiol Biotechnol **69**, 627–642.

[7] Van Dijken JP, Bauer J, Brambilla L, Duboc P, Francois JM, Gancedo C, Giuseppin MLF, Heijnen JJ, Hoare M, Lange HC, Madden EA, Niederberger P, Nielsen J, Parrou JL, Petit T, Porro D, Reuss M, Van Riel N, Rizzi M, Steensma HY, Verrips CT, Vindeløv J, Pronk JT (2000) An interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains. Enzyme Microb Tech **26**, 706–714.

[8] Sun Y, Cheng J (2002) Hydrolysis of lignocellulosic materials for ethanol production: a review. Bioresource Technol **83**, 1-11.

[9] Lynd LR, Van Zyl WH, McBride JE, Laser M (2005) Consolidated bioprocessing of cellulosic biomass: an update. Curr Opin Biotech **16**, 577–583.

[10] Eichhorn SJ, Baillie CA, Zafeiropoulous N, Mwaikambo LY, Ansell MP, Dufresne A, Entwistle KM, Herrera-Franco PJ, Escamilla GC, Groom L, Hughes M, Hill C, Rials TG, Wild PM (2001) Current international research into cellulosic fibres and composites. J Mater Sci **36,** 2107–2131.

[11] Rabinovich ML, Melnik MS, Bolobova AV (2002) Microbial cellulases (Review) Appl Biochem Micro **38**, 355-373.

[12] Hilde L, Johansson G (2004) Recent developments on cellulases and carbohydrate-binding modules with cellulose affinity. Biotechnol Lett **26**, 1683–1693.

[13] Den Haan R, Mcbride JE, La Grange DC, Lynd LR, Van Zyl WH (2007) Functional expression of cellobiohydrolases in *Saccharomyces cerevisiae* towards one-step conversion of cellulose to ethanol. Enzyme Microb Tech **40,** 1291-1299.

[14] Fujita Y, Ito J, Ueda M, Fukuda H, Kondo A (2003) Synergistic saccharification, and direct fermentation to ethanol, of amorphous cellulose by use of an engineered yeast strain codisplaying three types of cellulolytic enzyme. Appl Environ Microbiol **70**, 1207-1212.

[15] Den Haan R, Rose SH, Lynd LR, Van Zyl WH (2007) Hydrolysis and fermentation of amorphous cellulose by recombinant *Saccharomyces cerevisiae*. Metab Eng **9**, 87–94.

[16] Zhang YH, Cui J, Lynd LR, Kuang LS (2006) A transition from cellulose swelling to cellulose dissolution by *o*-phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure. Biomacromolecules **7**, 644–648.

[17] Yu S, Jeppsson H, Hahn-Hägerdal, B (1995) Xylulose fermentation by *Saccharomyces cerevisiae* and xylose-fermenting strains. Appl Microbiol Biotechnol **44**, 314–320.

[18] Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

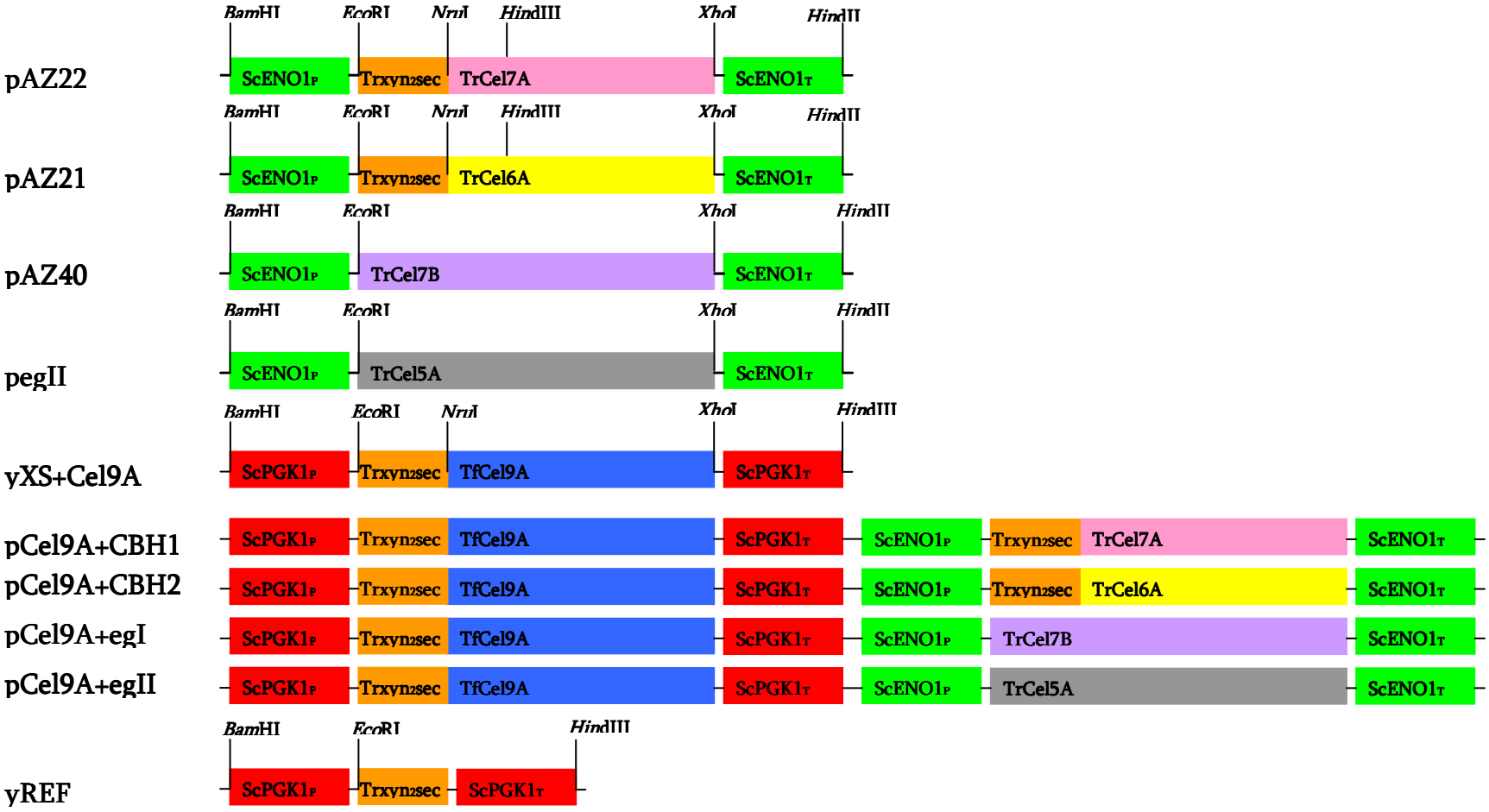
[19] Dubois M, Gilles KA, Hamilton JK, Rebers PA, Smith F (1956) Colorimetric method for determination of sugars and related substances. Anal Chem **28**: 350–356.

[20] Miller GL, Blum R, Glennon WE, Burton AL (1960) Measurement of carboxymethyl cellulase activity. Anal Biochem 2:127–132.

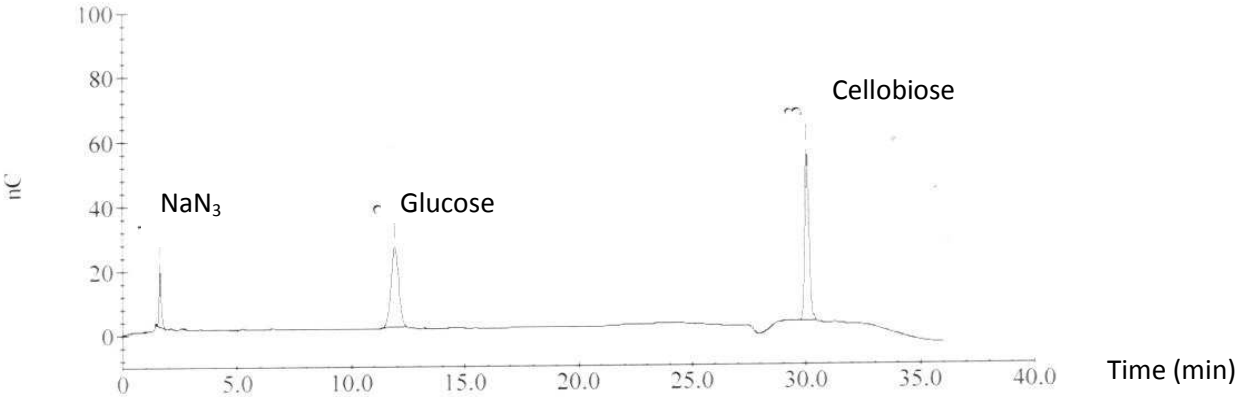
[21] Irwin DC, Spezio M, Walker LP, Wilson DB (1993) Activity studies of eight purified cellulases: Specificity, synergism and binding domain effects. Biotechnol Bioeng **42**, 1002-1013.

[22] Li Y, Irwin DC, Wilson DB (2007) Processivity, Substrate Binding, and Mechanism of Cellulose Hydrolysis by *Thermobifida fusca* Cel9A. Appl Environ Microbiol **73** 3165–3172.

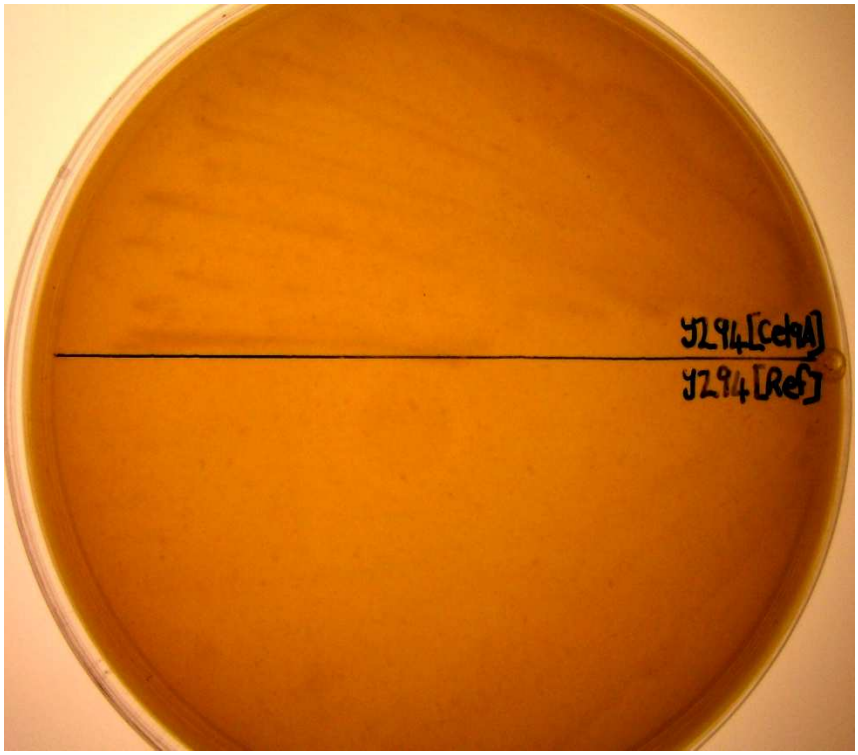
GRAPH 1


pAZ22
BamHI
EcoRI
NruI
HindIII
XhoI
HindIII
ScENO1P
Trxyn2sec
TrCel7A
ScENO1T
pAZ21
BamHI
EcoRI
NruI
HindIII
XhoI
HindIII
ScENO1P
Trxyn2sec
TrCel6A
ScENO1T
pAZ40
BamHI
EcoRI
XhoI
HindIII
ScENO1P
TrCel7B
ScENO1T
pegII
BamHI
EcoRI
XhoI
HindIII
ScENO1P
TrCel5A
ScENO1T
yXS+Cel9A
BamHI
EcoRI
NruI
XhoI
HindIII
ScPGK1P
Trxyn2sec
TfCel9A
ScPGK1T
pCel9A+CBH1
ScPGK1P
Trxyn2sec
TfCel9A
ScPGK1T
ScENO1P
Trxyn2sec
TrCel7A
ScENO1T
pCel9A+CBH2
ScPGK1P
Trxyn2sec
TfCel9A
ScPGK1T
ScENO1P
Trxyn2sec
TrCel6A
ScENO1T
pCel9A+egI
ScPGK1P
Trxyn2sec
TfCel9A
ScPGK1T
ScENO1P
TrCel7B
ScENO1T
pCel9A+egII
ScPGK1P
Trxyn2sec
TfCel9A
ScPGK1T
ScENO1P
TrCel5A
ScENO1T
yREF
BamHI
EcoRI
HindIII
ScPGK1P
Trxyn2sec
ScPGK1T

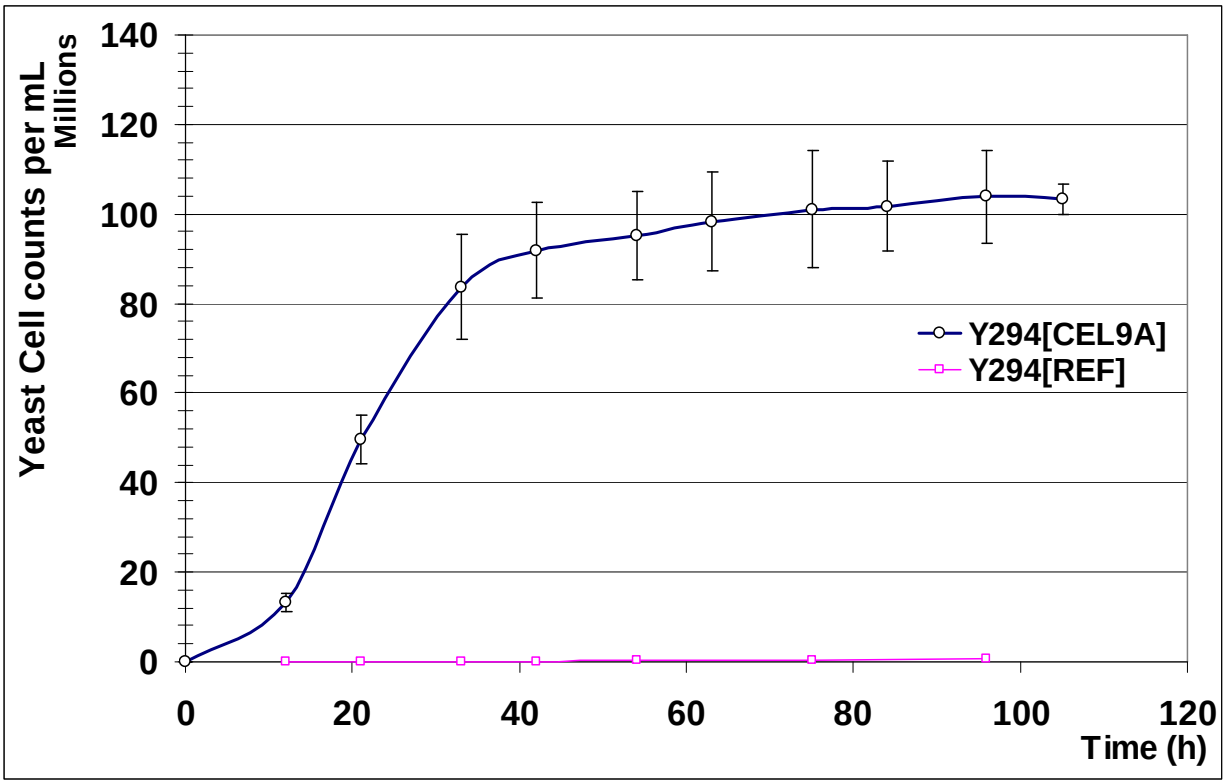
GRAPH 2



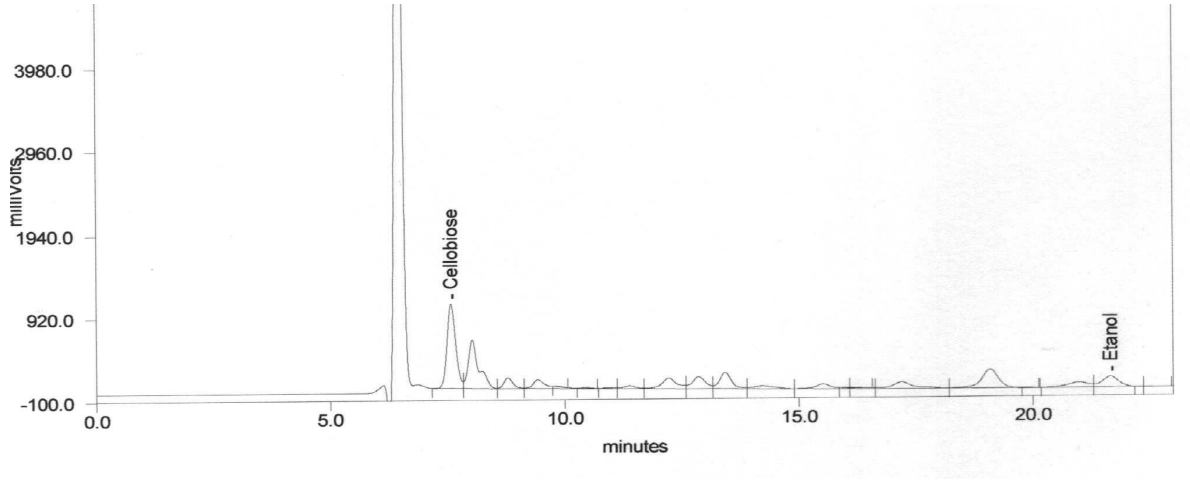
GRAPH 3



GRAPH 4



GRAPH 5



GRAPH 6

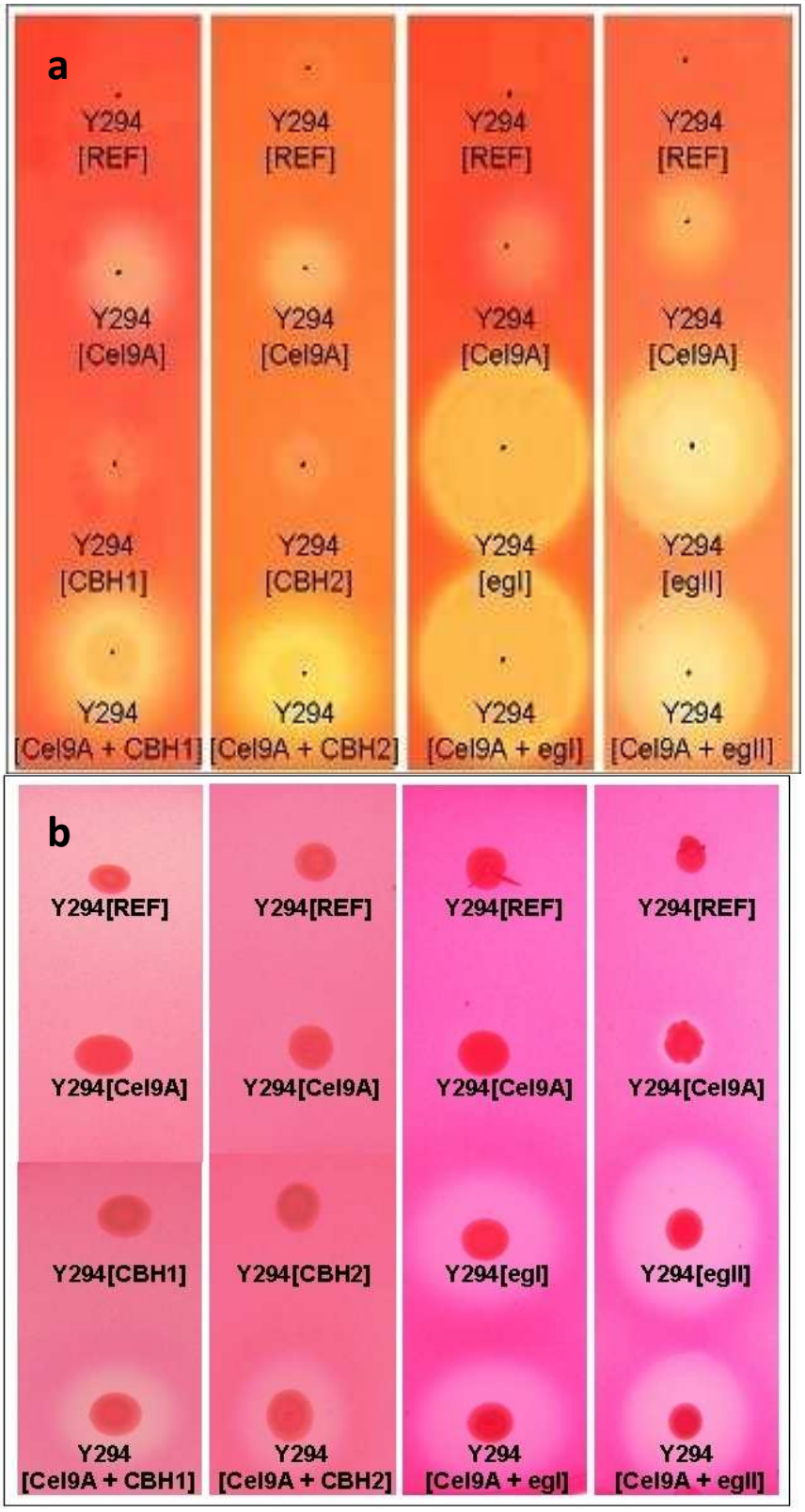

a
Y294 [REF]
Y294 [Cel9A]
Y294 [CBH1]
Y294 [Cel9A + CBH1]
Y294 [REF]
Y294 [Cel9A]
Y294 [CBH2]
Y294 [Cel9A + CBH2]
Y294 [REF]
Y294 [Cel9A]
Y294 [egl]
Y294 [Cel9A + egl]
Y294 [REF]
Y294 [Cel9A]
Y294 [egll]
Y294 [Cel9A + egll]
b
Y294[REF]
Y294[Cel9A]
Y294[CBH1]
Y294 [Cel9A + CBH1]
Y294[REF]
Y294[Cel9A]
Y294[CBH2]
Y294 [Cel9A + CBH2]
Y294[REF]
Y294[Cel9A]
Y294[egl]
Y294 [Cel9A + egl]
Y294[REF]
Y294[Cel9A]
Y294[egll]
Y294 [Cel9A + egll]

TABLE 1

Plasmids propagated in Escherichia coli DH5α	**Abbreviated name**	**Relevant genotype**	**Source/Reference**
pGEM-Cel9A		*cel9A* (secretion signal removed)	This work
pTZ-Cel9A		*PGK1P-XYNSEC-cel9A-PGK1T*	This work
pTZ-cbh1		*ENO1P-XYNSEC-cbh1-ENO1T*	This work
pTZ-cbh2		*ENO1P-XYNSEC-cbh2-ENO1T*	This work
Yeast strains constructed			
Saccharomyces cerevisiae Y294		α *leu2-3,*112 *ura3-52 his3 trp1-289*	ATCC 201160
(*fur1::LEU2* yXYNSEC)	Y294[REF]	*bla ura3/URA3 PGK1P-XYNSEC--PGK1T*	13
(fur1::LEU2 pAZ40)	Y294[egI]	*bla ura3/URA3 ENO1P-EGI-ENO1T*	19
(fur1::LEU2 pegII)	Y294[egII]	*bla ura3/URA3 ENO1P-EGII-ENO1T*	This work
(fur1::LEU2 pAZ22)	Y294[CBH1]	*bla ura3/URA3 ENO1P-XYNSEC-CBH1-ENO1T*	13
(*fur1:LEU2* pAZ21)	Y294[CBH2]	*bla ura3/URA3 ENO1P-XYNSEC-CBH2-ENO1T*	13
(*fur1:LEU2* yXS+Cel9A)	Y294[Cel9A]	*bla ura/URA3 PGK1P-XYNSEC-CEL9A-PGK1T*	This work
(*fur1:LEU2* pCel9A+egI)	Y294[Cel9A+egI]	*bla ura3/URA3 PGK1P-XYNSEC-CEL9A-PGK1T ENO1P-EGI-ENO1T*	This work
(*fur1:LEU2* pCel9A+egII)	Y294[Cel9A+egII]	*bla ura3/URA3 PGK1P-XYNSEC-CEL9A-PGK1T ENO1P-EGII-ENO1T*	This work
(*fur1:LEU2* pCel9A+CBH1)	Y294[Cel9A+CBH1]	*bla ura3/URA3 PGK1P-XYNSEC-CEL9A-PGK1T ENO1P-XYNSEC-CBH1-ENO1T*	This work
(*fur1:LEU2* pCel9A+CBH2)	Y294[Cel9A+CBH2]	*bla ura3/URA3 PGK1P-XYNSEC-cel9A-PGK1T ENO1P-XYNSEC-CBH2-ENO1T*	This work

TABLE 2

Primer name	**Sequence (5'-3')**	**Relevant restriction sites**
Tfuscel90-left	AAGCT**TCGCGA**GAACCGGCGTTCAACTAC	*Nru*I
Tfuscel90-right	AGGC**CTCGAG**GTTGCCGTGTTAGGCG	*Xho*I
ENO1-left	GGATCCACTAGTCTTCTAGGCGGGTTATC	
ENO1-right	**AAGCTT**GCGGCCGCAAAGAGGTTTAGACATTGG	*Hin*dIII
PGK1beginprom	ACTG**AAGCTT**GGATCCTTAAAGATGCCG	*Hin*dIII
PGK1endterm	ACTG**AAGCTT**GGCCAAGCTTTAACGAAC	*Hin*dIII
FUR1-L	TCCGTCTGGCATATCCTA	
FUR1-R	TTGGCTAGAGGACATGTA	

TABLE 3

Strain	**CMC**	**PASC**	**Avicel**	**BMCC**
Y294[Cel9A]	0.170 ± 0.018	1.241 ± 0.199	0.687 ± 0.102	0.399 ± 0.054
Y294[CBH2]	0.002 ± 0.001	3.570 ± 0.485	2.196 ± 0.320	1.336 ± 0.164
Y294[egI]	0.517 ± 0.021	Not determined	Not determined	Not determined
Y294[REF]	0.002 ± 0.001	0.265 ± 0.171	0.034 ± 0.017	0.047 ± 0.042