Increasing ethanol tolerance through the heterologous expression of the *Geobacillus thermoglucosiadius* heat-shock proteins (GroEL/GroES) in an *Escherichia* coli host

Wellington Charewa

A thesis submitted in partial fulfillment of the requirements for the degree of Magister Scientiae in the Department of Biotechnology, University of the Western Cape

January 11, 2012

Supervisor: Professor Don Cowan

Co- supervisor: Dr Mark Paul Taylor

Abstract

Due to economic and environmental concerns associated with use of fossil fuels, humanity is seeking alternative fuels. Ethanol is one of the alternative fuels produced commercially. Current ethanol production technologies using first generation ethanol processes is criticised for depleting the food supply and escalating food prices. Biomass is a target feedstock for use in bioethanol production and would resolve the criticism associated with the current bioethanol industry. Bacterial strains such as *Geobacillus thermoglucosidasius* NCIMB 11955 can be used to produce ethanol from biomass because they assimilate hexose and pentose sugars, a property that is lacking in first generation ethanol producing microbes (*Saccharomyces cerevisiae* and *Zymomonas mobilis*) (Riyanti and Rogers, 2009). Due to the low ethanol tolerance (4 % v/v (maximum)) of *G. thermoglucosidasius*, use of this species for bioethanol production is not economically viable. GroES and GroEL genes are involved in stress tolerance in bacteria: activation of these genes has been observed in stress induced bacteria (Rasouly and Ron, 2009).

In this study the ethanol tolerance of *G. thermoglucosidasius* NCIMB 11955 was characterised by culturing at 45 ºC and 55 ºC in the presence of ethanol. A greater ethanol tolerance was observed at the sub-optimal growth temperature of 45 ºC. *Escherichia coli* metabolic systems are well understood. Aiming to improve the ethanol tolerance *G. thermoglucosidasius* NCIMB 11955, the GroES and GroEL genes of the organism were cloned in an expression vector and expressed in *E.coli* before testing their ability to confer an increased tolerance to ethanol. Proteomic analysis of the recombinant *E. coli* strain showed that GroES was over-expressed while GroEL was not. After over expression of GroES, the optical density of cultures was periodically measured. Over-expression of the *G. thermoglucosidasius* NCIMB 11955 GroES gene improved the ethanol tolerance of *E. coli* Rosetta pLySs growing in 4% (v/v) ethanol.

Keywords

Geobacillus

Ethanol tolerance

Stress response

Bioethanol

Growth temperature

GroESL

Heat shock proteins

Declaration

I declare that “Increasing ethanol tolerance through the heterologous expression of the *Geobacillus thermoglucosiadius* heat-shock proteins (GroEL/GroES) in an *Escherichia coli* host” is my own work, that it has not been submitted for any degree or examination in any other university, and that all the sources l have used or quoted have been indicated and acknowledged by complete references.

Wellington Charewa 15/12 2013

Signed……

Acknowledgements

My initial thanks go to Professor Donald Cowan for securing funding for me, his guidance and support throughout my research. My thanks go to Professor Marla Tuffin, Dr. Mark. P. Taylor and Dr. Heide Goodman for advice and help throughout my research. Thanks go to Professor B. Ndimba for helping me with the proteomic analysis.

My thanks extend to the MCB laboratory at the University of Cape Town for their service of DNA sequencing. Thanks to the Inqaba Biotechnical Laboratories who gave me an internship position that had a great impact on my molecular techniques. My thanks extend to the TMO team who supplied us with all the information we required on the *Geobacillus thermoglucosidasius* NCIMB 11955 genome sequence.

On a personal note I would like to thank my Mum and Dad for the continued support they gave to me throughout my life. Thanks to my siblings Carlington, Grace and Emerald who are always there for me.

Thanks go to my close friend Asaniel Mavugara who always made time outside the laboratory great. Thanks go to my closest friend who has been in my life since 2006; I appreciate your time and attention that you have given me in the past five years.

Table of contents

List of figures

List of tables

List of frequently used symbols and abbreviations

Symbol/abbreviation	Expansion/meaning
× g	Centrifugal force
BLAST	Basic local alignment sequencing tool
bp	Base pair
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetra-acetic acid
g	Gram
ng	nanogram
kV	Kilo volt
L	Liter
M	Molar
mM	Millimolar
ºC	Degrees Celsius
OD	Optical density
TAE	Tris-acetate EDTA buffer
TE	Tris EDTA
v/v	Volume per volume
w/v	Weight per volume
µl	Microlitre
bp	Base pair

Chapter 1: Literature review

1.1 The current state of fuel usage

Due to environmental and sustainability concerns surrounding the use of fossil fuels, the use of alternative and renewable fuels has become a major scientific and social area of interest (Solomon *et al*., 2007). A number of renewable forms of energy are currently being developed to replace or supplement fossil fuels. These include wind and solar derived energy (electricity), chemical energy, biological gas (mainly hydrogen gas for transportation, heating and cooking) and biological liquid biofuels for transportation (Pernick and Wilder, 2007). The latter covers the current feasible liquid renewable alternatives for the automobile industry as a replacement for or a blending agent with gasoline.

Other petrol alternatives suffer from a number of disadvantages. The volumetric energy density of hydrogen gas is less than that of petrol (Savage, 2011). Hydrogen requires an infrastructure of pipe lines for transportation which, in Africa at least, is largely non-existent and expensive (Gardener, 2004). Electric powered cars are generally less efficient than those running on hydrogen, biofuels and fossil fuels (Gardner, 2004). Electricity has to be sourced from coal, solar or wind energy and at present it is a technologically challenging task to produce sufficient quantities to meet demand. Biofuels are possibly the most efficient and cost effective replacements to petroleum but are again less energy efficient in comparison to petrol. Bioethanol is 1.67 times less energy efficient than petrol (Samson, 1991). Amongst the renewable fuels

(principally ethanol, butanol, higher chain or branched chain alcohols and biodiesel) bioethanol is arguably the most efficient and has the added benefit of potentially addressing a number of environmental concerns that have arisen from oil consumption in the last 50 - 100 years (Solomon *et al*., 2007).

In comparison to other liquid fuels, ethanol possesses a number of innate characteristics that make it attractive as a liquid transportation fuel. It has a higher octane value than petrol (Chandel *et al*., 2007) and, being renewable, secures energy rights for a country. Bioethanol can be generated locally while oil is localised globally and is therefore controlled by holder states (Green, 2004). Bioethanol can also be integrated as a blended product into the current transport petroleum infrastructure with little modification. Most cars currently tolerate addition of 5 - 10% ethanol to the petrol supply of the internal combustion engine (Antoni *et al*., 2007; Solomon *et al*., 2007).

Currently industrial bioethanol production is based on mature technology involving the fermentation of maize and sugarcane-derived carbohydrates (so-called first generation ethanol production) (Butzen and Haefele, 2008). These processes are biocatalysed by industrial strains of *Saccharomyces cerevisiae* (Butzen and Haefele, 2008). Although this is a proven technology which is capable of producing ethanol at a commercial level, ethanol production from maize and sugarcane has been criticised for diverting carbohydrates from the human food chain and for driving up the price of maize (Figure 1.1). In the short term this has been less of a concern in western countries but is a hotly

debated topic in third world nations (Strydom, 2009). Global implementation of this technology is seen as a transitional mitigation of the problems concerning fuel security because of its ethical greyness.

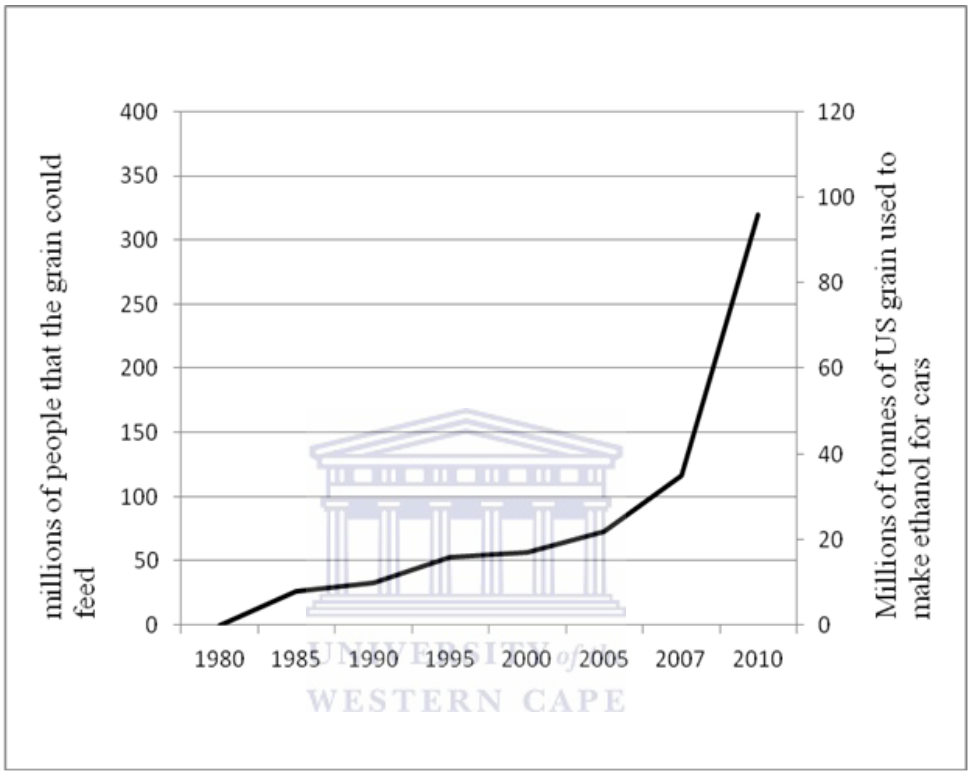


Figure 1.1: Comparison of the amount of US grain used to produce ethanol and the population size the grain could feed (Vidal, 2010 http://www.theguardian.com/environment/2010/jan/22/quarter-us-grain-biofuels-food)

Based on the model above, for every 90 million tonnes of maize used to produce ethanol, nearly 325 million people are deprived of food or could be fed from grain diverted for fuel use). This is approximately three times the population of Africa (ICPD Program of Action, 1998). The region is faced with a shortage in the supply of food, has insufficient arable land for food production and suffers from erratic environmental factors (such as rainfall)

which, in the absence of intensive farming systems and widespread fertiliser use, may impact the efficacy of first generation biofuel production (Chirwa, 2007).

North America currently uses 21 million barrels of oil per day, which equates to maize ethanol derived from 11.2 million tons of maize (Vidal, 2010). At its highest maize yield, 14.2 billion square metres of land would be required to provide North America with approximately 30 hours of ethanol at the current usage. Therefore first generation ethanol production is probably not feasible for any country, developed or otherwise, that wished to shift entirely from petroleum to biofuel products in the long term.

1.1.1 Second generation biofuel production

Critics of the use of food chain carbohydrate for ethanol production have motivated for the development of second generation bioethanol production (Antoni *et al*., 2007; Demain, 2009). Second generation processes utilise biomass (the lignocellulosic quotient of biological materials) as the raw source material from which to derive carbohydrates for biofuel production.

The benefits of this approach over first generation technologies are numerous. Firstly, the issues of food security are bypassed since lignocellulose is inedible. Secondly, the majority of biomass available is regarded as waste material as it is derived from agricultural waste feed stocks, wood and paper waste, municipal wastes and unfermented materials from current fuel production

processes (Galbe and Zacchi, 2002; Chandel *et al*., 2007). The use of these waste materials to produce ethanol would not require additional arable land. The biomass could even be produced on non-arable land if a dedicated drought resistant or tolerant crop was sourced. Unlike first generation biofuel processes, the potential for economic growth and job creation exists, as does the production of a valuable product with applications in transportation (Solomon *et al.*, 2007). However these technologies are mostly at the laboratory stage of development and few commercially viable processes exist at present (Demain, 2009).

British Petroleum (BP) is one of the fuel processing companies that supplies petrol and other transportation fuels on a large scale. The company plays a major role in the production of ethanol from sugar cane in Brazil (Butamax, December 2012). With the realised need to shift from corn/sugarcane ethanol production to biomass ethanol production, BP has been developing technologies that will reduce biomass ethanol production costs in partnership with the Verenium Corporation. The partnership, which started in 2008, is currently producing 1.4 million gallons of cellulosic ethanol in its demonstration facility (Butamax, December 2012). BP, which acquired the Verenium Corporation's biofuels business for $98.3 million, intends to build an industry leading cellulosic ethanol facility in the USA (Verenium Industry, July 2010).

Other companies such as TMO Renewables are also focussing on cellulosic ethanol production (Taylor *et al*., 2008). In 2008 TMO Renewables introduced a pilot cellulosic ethanol plant near Surrey, England which utilises a thermophilic bacterium (a *Geobacillus thermoglucosidasius* mutant) to produce high levels of ethanol and smaller quantities of by-products (such as lactate and formate) than was previously possible (Taylor *et al*., 2008). The process cuts the costs of lignocellulose pretreatment as the mutant *Geobacillus* strain can ferment complex sugars and has an optimum growth temperature of 55 to 60 ˚C. Due to the high temperature (60 to 70 ˚C) of the biocatalytic process, cooling costs during biomass processing to ethanol have been reduced (Taylor *et al*., 2008).

1.1.2 Limitations to the development of second generation processes

Biomass in its raw form is recalcitrant to biological degradation or fermentation (Antoni *et al*., 2007). Costly and laborious chemical and enzymatic pre-treatments are required in order to liberate fermentable sugars (Samson, 1991). Costing, based on the current best biomass treatment and fermentation technologies (acid treatment of biomass and delignification processes), showed that in 1991 a gallon of ethanol cost $4 to produce, compared to $0.63 for a gallon of petrol (Samson, 1991).

There are two main obstacles blocking the full realisation of second generation technologies. First, the added cost implications required to pre-treat biomass needs to be reduced either through the optimisation of chemical methods or the

reduction in cost of production by specific biological catalysts for deconstruction (Green and Mugica, 2005). Secondly, a versatile ethanologenic strain with a broad catabolic phenotype is required (Cripps *et al*., 2009). *Geobacillus* spp. are good candidates, as they have many of the required qualities.

Refinement of the pre-treatment and enzymatic catalysis steps and raw material modifications may reduce the cost of bioethanol at the pump (Solomon *et al*., 2007). This research and development, together with the coupling of ethanol production to other processes such as electricity generation, may further reduce the cost of ethanol production from biomass (Green and Mugica, 2005). Genetically modified lignocellulosic materials which would facilitate pre-treatment are being investigated (Weng *et al*., 2008). Lignocellulose contains lignin which is an inhibitor of the saccharification enzymes that are responsible for breaking down hemicellulose (Demain, 2009). Reducing the amount of lignin in plants used for bioethanol production will greatly reduce pre-treatment and ethanol production costs.

The choice of organism used in the catalysis of biomass to ethanol is critical for improving efficiency and lowering production costs. An organism that utilises both pentose and hexose sugars is the only feasible production strain for the efficient processing of biomass to ethanol. The focus taken by some groups has been to investigate thermophilic prokaryotes with known ethanol production

phenotypes and catabolic promiscuity. One such organism is *G. thermoglucosidasius* NCIMB 11955 (Blumer-Schuette *et al*., 2008).

G. thermoglucosidasius utilises both pentose and hexose sugars, making it an ideal candidate for biomass hydrolysate fermentation. This species tolerates up to 4 % v/v ethanol at an optimum growth temperature of 55 ºC. The drawbacks of using this strain for commercial ethanol production include the low ethanol tolerance, low ethanol yields (they are mixed acid fermenters) and growth inhibition both by their metabolic by-products (lactate, pyruvate, acetate and formate) and pre-treatment derived inhibitors, for example furfural (Blumer-Schuette *et al*., 2008).

1.2 Sources of fuel

1.2.1 Fossil fuels

Fossil fuels are defined as any natural occurring organic fuels that are formed in the earth's crust. These include petroleum, coal, and natural gas. The overriding limitation to their efficacy is that fossil fuels are finite and oil will not last into the 21st century, if the Hubbert peak theory is to be believed (Hubbert, 1956; Campbell and Laherrère, 1995).

A decline in oil production can cause an increase in oil demand and therefore an oil price increase (Galbe and Zacchi, 2002). As more countries are becoming industrialised, the demand for oil is increasing and petrol producers have cited the need for more oil reserves to sustain future demands (Galbe and

Zacchi, 2002). Some countries have responded to the crisis by developing methods for petrol production through coal liquefaction. For example, Sasol in South Africa produces one third of the country's fuel using coal liquefaction (Wakeford, 2008). Coal is more abundant than oil, but like oil, coal is a finite resource and therefore the long term sustainability of these technologies is questionable.

A long term replacement that is renewable and has a low environmental impact is the ultimate goal for meeting energy requirements (Strydom, 2009). Environmental and energy security concerns (as witnessed in the 1970s) have re-emerged and the need for alternative fuel has once again seen revalidation of the concept of biofuel (in particular ethanol) as a good alternative fuel (Sun and Cheng, 2001; Galbe and Zacchi, 2002; Chandel *et al*., 2007).

1.2.2 Bioethanol

Ethanol yields and the process economics/technicalities of bioethanol production are the driving variables determining the success of current efforts to produce ethanol (Mishima *et al*., 2008). Ethanol is a good alternative fuel and it can potentially assist in the reduction of greenhouse gas emissions by combusting to completion to produce water and carbon dioxide (Mishima *et al*., 2008).

Ethanol enhances engine performance by increasing its thermal efficiency, therefore reducing heat loss (Wyman, 1996). Based on its stoichiometry,

ethanol flame temperature is 1930˚C, lower than that of petrol (1977˚C) and diesel (2054˚C) (Wyman, 1996). Ethanol can replace the use of tetraethyl lead as an octane enhancer in gasohol and decrease smog emission due to its lower volatility (Demain, 2009). Besides these advantages, ethanol is less toxic to the environment as it does not contain sulphur (Demain, 2009). Sulphur burns in air to produce sulphur dioxide which reacts with water to form acid rain with detrimental effects to the environment (Jones, 1950). Bio-ethanol has also been used as an oxygenate to replace methyl-tert-butyl ether (MTBE) (Antoni *et al*., 2007; Demain, 2009). MTBE was used as a petrol oxygenate for years but its use was phased out due to its toxicity and pollution of large volumes of ground water (Demain, 2009). In addition, ethanol is twice as efficient as MTBE at oxygenating petroleum (Sun and Cheng, 2001).

There are drawbacks to using ethanol as an alternative fuel or supplement to gasoline. These include the fact that bioethanol production is currently dependent on food supplies (maize and sugar cane) and that raw materials and production costs are high resulting in ethanol costing more than petrol (Demain, 2009). Due to the industrial production of ethanol from food-derived carbohydrate, costs of staple foods (corn) have increased (Vidal, 2010). As a result, many families are unable to afford a basic meal and human rights groups have condemned bioethanol production as unethical (Demain, 2009). The condemnation of this first generation bioethanol production has shifted the focus of ethanol producers to using biomass as a raw material. Second generation bioethanol production may be a solution of two major problems

faced by southern Africa i.e. a lack of usable energy and food security (Strydom, 2009).

Southern Africa has no viable oil deposits and is always hit hard by increases in fuel prices (Wakeford, 2008). Creating a viable second generation ethanol production industry in southern Africa will result in a region that can sustain itself in terms of energy security and food supply. A biomass based ethanol industry is a possible solution to some of the economic problems facing southern Africa.

1.2.3 Biobutanol

Biobutanol is produced from sugar by the acetone-butanol-ethanol (ABE) fermentation process (Demain, 2009). The organism used for the industrial production of acetone from starch is *Clostridium acetobutylicum* (Antoni *et al*., 2007).

Butanol has a low octane number, a factor that limits its application as an oxygenate for gasoline. The octane number of a fuel is the measure of how resistant the fuel is to engine knocking and does not relate to the energy content of the fuel (Surisetty et al., 2011). Engine knocking can rapidly damage an engine (Surisetty et al., 2011). A fuel with a higher octane number is less likely to cause engine knocking, a property that motor vehicle manufacturers have taken advantage of by adjusting the ignition timing (Surisetty et al., 2011). Ethanol is a more desirable fuel because it has a higher motor octane number

(89 MON) than butanol (78 MON)

(http://www.eia.gov/oiaf/analysispaper/biodiesel/).

The development of biobutanol infrastructure is cheaper than that of bioethanol infrastructure. Butanol does not require engine modification until it reaches a concentration of 40 % v/v of the total fuel. By comparison ethanol requires engine modification if it is going to be used at concentrations beyond 15 % v/v of the total fuel volume (Demain, 2009).

High butanol production yields can be achieved by gas stripping of liquefied maize starch during fed batch fermentation; this ABE process has been shown to produce 56 g/l butanol, 24 g/l acetone and 1 g/l ethanol from 225 g/l sugar (Demain, 2009). Despite the advantages, the production of butanol and the genetic modification of the organisms involved are currently not at the stage whereby commercialisation can be realised. Although companies such as DuPont and BP are focusing on commercialisation of this fuel, butanol production methods require much improvement to bring it to the level of current ethanol production technologies.

DuPont and BP merged to form Butamax whose headquarters are in Wilmington, USA (http://www.butamax.com/Portals/0/pdf/butamax_advanced _biofuels_llc_fact_sheet.pdf). Butamax is responsible for the completion of the butanol research of DuPont and BP and the company has advanced research to

a point that the technology is at a demonstration phase. Butamax expects to launch the first commercial butanol production plant during the year 2013.

1.2.4 Biodiesel

Biodiesel refers to a non-petroleum based diesel fuel that consists of short chain alkyl (methyl or ethyl) esters made by transesterification of vegetable oil or animal fat (Demirbas, 2005). Biodiesel is produced through acid or base esterification (http://www.eia.gov/oiaf/analysispaper/biodiesel/). Oil feed stocks containing about 4 % free fatty acids by volume are put through an acid esterification process to maximise the output of biodiesel (http://www.eia.gov/oiaf/analysispaper/biodiesel/). The acid esterification reaction is catalysed by sulphuric acid; the acid is mixed with methanol prior to conversion of free fatty acids to produce biodiesel. In an acid/base biodiesel process, acid esterification of oils is done to free fatty acids that are used to maximise extraction of more fatty acids in the base catalysed reaction (Demirbas, 2005). The products of acid esterification and water free oils are mixed and the mixture undergoes transesterification, a reaction catalysed by potassium hydroxide containing methanol (Demirbas, 2005). Once the reaction is complete, the products (biodiesel and glycerine) are separately extracted before methanol is removed. Removing methanol last is necessary because methanol prevents the reverse reaction from proceeding (Van Gerpen, 2005).

Biodiesel was made available to the public when diesel engines were introduced. Petroleum separation and processing to petrol is cheaper and more efficient than processing of oil (plant and animal fats) to biodiesel (Demirbas, 2005). This resulted in the phasing out of biodiesel as a transportation fuel. As a result of the limited oil supply in the 1970's (Cavallo, 2004), production of biodiesel increased noticeably in the 1990's (Demirbas, 2005). In 1999, the National Biodiesel Board of the USA recorded production of biodiesel at 1.9 million litres per day and 25.4 million litres per day in 2000 (http://www.eia.gov/oiaf/analysispaper/biodiesel/). Regardless of its advantages, biodiesel production is limited by a limited supply of raw material. Biomass for ethanol production is more abundant than biomass for biodiesel production (Demirbas, 2005).

1.2.5 Commercialisation and economics of bioethanol production

Brazil has been producing bioethanol from sugarcane since 1970 (Kamimura and Sauer, 2008). Brazil's fuel stations supply 100 % bioethanol and car manufacturers in Brazil have adjusted to the demand of ethanol powered engines. For example Ford and Honda have been supplying cars able to use 95 % ethanol as fuel since the 1980's (Blumer-Schuette *et al*., 2008). Due to high ethanol outputs facilitated by infrastructure development, Brazil was not severely impacted by the fuel price rise in 2008 (Chandel *et al*., 2007). Brazil leads the world in the use of ethanol to replace fossil fuels.

The ethanol industries of Brazil and the USA cannot ethically be reproduced in continents such as Africa due to the short supply of food in the continent. Consequently the idea of the production of ethanol from grain was criticised in southern Africa before the technology was put into practise (Strydom, 2009). Due to food supply concerns, producers of ethanol are trying to shift from using food carbohydrate (maize and sugar cane) to the use of biomass as the raw material (Demain, 2009). The capital cost of ethanol production from biomass is five times higher than that of ethanol production from maize (Van Gerpen, 2005). Based on the Energy Information Administration findings, the cost of biomass fermentation to ethanol was $2.6 per litre in 2008. Fermentation of biomass to ethanol still requires developments to bring it to the level of starch fermentation by yeast (Galbe and Zacchi, 2002; Chandel *et al*., 2007).

1.2.6 Industrial production of ethanol using *Saccharomyces cerevisiae*

S. cerevisiae is a facultative anaerobe that is used in the industrial production of ethanol. *S. cerevisiae* produces ethanol from the pyruvate formed by the Embden-Meyerhof pathway with the aid of pyruvate decarboxylase (*pdc*) and alcohol dehydrogenase (*adh*) enzymes (Li *et al.*, 2012) (Figure 1.2).

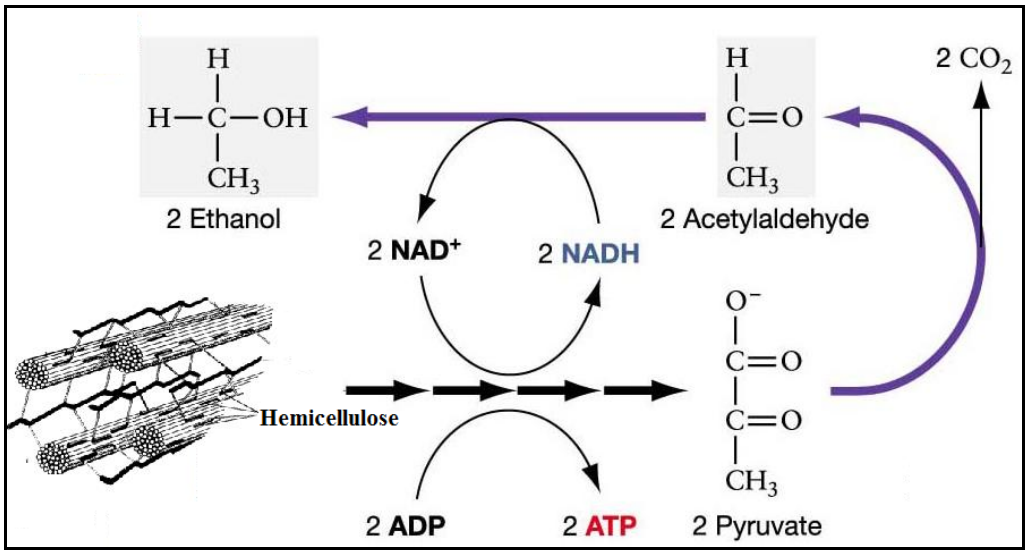


Figure 1.2: Ethanol production in *S. cerevisiae*. Adapted from Li *et al.*, 2012

Sugars (glucose) are converted to pyruvate through the Embden Meyerhof or Entner Doudoroff pathways (Gunasekaran and Chadra Raj, 1999). Pyruvate is then converted into acetaldehyde and carbon dioxide (Zhang *et al.*, 2007). Carbon dioxide is excreted from the cell and the acetaldehyde is further converted into ethanol through catalysis by alcohol dehydrogenase (Zhang et al., 2007*). S. cerevisiae* produces ethanol efficiently at 30 °C, pH 5 and at glucose concentrations of 25 g/l.

S. cerevisiae cannot ferment pentose sugars such as xylose (Li *et al.*, 2012). As a result, research is focussed on the constitutive expression of genes coding for enzymes involved in the fermentation of pentose sugars to ethanol. Strategies to improve *S. cerevisiae* for the fermentation of D-xylose have been implemented. These include the insertion of a bacterial D-xylose isomerase gene (Brat *et al.*, 2009) and the heterologous expression of pentose utilizing

genes (Walfridsson *et al.*, 1995). The expression of the xylose isomerase gene from *Clostridium phytofermentans* in *S. cerevisiae* enhanced the growth of *S. cerevesiae* in xylose containing media but did not enhance ethanol production (Brat *et al.*, 2009). The engineered strain converted xylose to xylitol which accumulated in the cells. Xylitol is an inhibitor of the xylose isomerase gene.

Further attempts made to engineer *S. cerevisiae* for xylose fermentation included the introduction of aldose reductase, xylitol dehydrogenase and xylulokinase genes from *Pichia stipitis* into the genome of *S. cerevisiae* (Walfridsson *et al*., 1995). Expression of aldose reductase alone did not produce the desired effects. The engineered strain did not grow in xylose containing media but did grow in a xylose-glucose mixture, producing xylitol at a high yield. However, the excess xylitol caused a redox imbalance within the cell (Meinander *et al.*, 1996). Expression of both aldose reductase and xylitol dehydrogenase allowed the engineered strain to grow on xylose containing media (Meinander *et al*., 1996). The genes for xylose fermentation have been cloned into *S. cerevisiae* from other organisms including *Pichia tannophilus* (Stevis *et al*., 1987). Although pentose utilisation by *S. cerevisiae* has been successfully achieved, optimisation of the metabolic system for simultaneous, exogenous sugar metabolism is still required (Young *et al*., 2010).

For efficient fermentation of lignocellulose hydrolysates, an ethanologenic bacterial strain such as *G. thermoglucosidasius* that co-ferments hexose and

pentose sugars is required. Developing a yeast strain that co-ferments hexose and pentose sugars efficiently has been an obstacle to the application of yeast in the commercial fermentation of biomass (Young *et al.*, 2010). The ethanol industry has diverted focus from developing yeast strains to ferment xylose to the improvement of bacteria that naturally ferment xylose. Organisms such as *G. thermoglucosidasius* are prime candidates.

1.3 Ethanologenic prokaryotes

1.3.1 *Escherichia coli*

E. coli is a Gram negative bacterium with an optimum growth temperature of 37°C (Yu *et al.*, 2000). For many years *E. coli* has been the workhorse for production of many economically viable chemicals due to its well understood genetic background, amenability to genetic modification and good growth properties with low nutrient requirements (Yu *et al.*, 2000).

E. coli produces very little ethanol (Ingram *et al.*, 1987). Due to the demand for renewable fuels and the dwindling supply of petroleum, researchers have attempted to develop strains such as *E. coli* for commercial ethanol production. The major factors that limit *E. coli* as a commercial strain for ethanol production are its poor ethanol tolerance and low ethanol yield (Ingram *et al.*, 1987; Yu *et al.*, 2010). The poor ethanol yield from *E. coli* is associated with the lack of *pdc* and *adh* genes in the strain, which instead produces succinate, acetate and lactate as major fermentation products (Ingram *et al.*, 1987). A number of strategies have been implemented to improve its ethanol yield. The

pdc and *adhB* genes from *Zymomonas mobilis* were integrated into *E. coli* (Wang *et al*., 2008). When integrating *pdc* and *adh*, an artificial operon of the two genes was developed and integrated into the pZY507 vector under the control of a *lac* promoter. Both genes were expressed under the control of *lacI*^{q}*-tac* (Wang *et al*., 2008). Expression of the *pdc* and *adh* genes in *E. coli* reduced acetate accumulation by shifting the carbon flow to ethanol production (Wang *et al*., 2008). Expression of the *pdc* and *adhB* genes from *Z. mobilis* in *E. coli* improved ethanol tolerance from undetectable amounts to 18 mmol/l (Ingram *et al*., 1987).

1.3.2 *Klebsiella* spp.

Klebsiella is a genus of Gram negative non-motile pathogenic bacteria that is associated with diseases such as pneumonia and urinary tract infections in hospitalised patients or alcoholics (Podshun and Ullmann, 1998). Species from this genus are oxidase negative and have a polysaccharide based capsule (Podshun and Ullmann, 1998). *Klebsiella* species ferment D-xylose to produce ethanol, butanediol and mixed acids. The concentration of ethanol produce by *Klebsiella* species is low compared to that produced by *Z. mobilis* and *S. cerevisiae* (Tolan and Finn, 1987). *Klebsiella* species produce a maximum of 7 g/l ethanol when cultured on pressed sugar beet pulp (Sutton and Peterson, 2001). Genetic engineering of this species has been attempted to improve ethanol yields. The *pdc* gene from *Z. mobilis* was cloned into *Klebsiella* to

enhance ethanol production. Ethanol production increased to 25.1 g/l on xylose and reduced the production of mixed acids (formate, acetate, lactate and butanediol) (Tolan and Finn, 1987). The engineered *Klebsiella* sp. expressing the *Z. mobilis pdc* gene produced 9.1 g/l more ethanol compared to an *E. coli* mutant expressing the same gene (Tolan and Finn, 1987). *Klebsiella* are pathogenic and this factor limits their commercial application as ethanol producers.

1.3.3 *Clostridium* spp.

Clostridium species are Gram positive bacteria belonging to the Firmicutes. This genus is characterised by pathogenic and non-pathogenic species. Pathogenic *Clostridium* species cause food poisoning (*C. perfringens*), tetanus (*C. tetani*) and post natal death (*C. sordellii*) (Bruggemann and Gottschalk, 2009**).** There are some species in this genus that are useful in industrial processes. *C. thermocellum* ferments biomass hydrolysates to ethanol (Bruggemann and Gottschalk, 2009**)** and *C. acetobutylicum* is used to produce acetone and biobutanol from starch (Bruggemann and Gottschalk, 2009**).**

Clostridium species ferment both hexose and pentose sugars into acetone, butanol and ethanol through the ABE process (Lovitt *et al*., 1984). *Clostridium* species can generally tolerate a maximum ethanol concentration of 1.6 % v/v; at 2 % v/v ethanol, fermentation of sugars is totally inhibited (Lovitt *et al*., 1984). In the presence of mixed sugars, glucose is the preferred substrate (Ezeji *et al*., 2008).

1.3.4 *Zymomonas mobilis*

The genus *Zymomonas* is sub-divided into *Z. mobilis and Z. anaerobia* (Gunasekaran and Chandra Raj, 1999). *Z. mobilis* is a Gram negative facultative anaerobic bacterium with a number of industrially desirable characteristics (Yang *et al.*, 2010). *Z. mobilis* has a high specific productivity, a high ethanol yield and a unique anaerobic Entner Doudoroff pathway (Yang *et al*., 2010). This species can tolerate up to 13 % ethanol v/v within a pH range of 3.5-7.5 (Yang *et al*., 2010). Due to the activity of the Entner Doudoroff pathway, *Z. mobilis* produces ethanol with a concurrent low diversion of carbon to biomass formation (Gunasekaran and Chandra Raj, 1999). Comparing it to ethanologenic yeasts, Z. *mobilis* has advantages for the industrial production of ethanol: it does not require the controlled addition of oxygen during fermentation and it is amenable to genetic manipulation.

Regardless of its advantages, there are limitations in using *Z. mobilis* for biomass fermentation. For example, it can only utilise glucose, fructose and sucrose as substrates, although co-culturing *Z. mobilis* with *C. saccharolyticum* has been proposed to circumvent this (Table 1.1). *Z. mobilis* is also inhibited by its by-products (formate, lactate, and acetate) and more severely by biomass pre-treatment products (hydroxymethylfufural and furfural) (Gunasekaran and Chandra Raj, 1999).

Table 1.1: Ethanol yields from various glucose-xylose sugar mixtures by a coculture fermentation consisting of *Z. mobilis* and *C. saccharolyticum* (*pdc* negative) (Murray and Asther, 1984)

Glucose-xylose sugar mixture (g/l)	Ethanol yield (mmoles/l)	Efficiency (%)
20 glucose + 10.8 xylose	322	98.5
40 glucose + 21.6 xylose	620	92.0
60 glucose + 32.4 xylose	903	89.3
80 glucose + 43.2 xylose	907	67.3

Attempts have been made to introduce pathways that are involved in pentose sugar fermentation into *Z. mobilis* with a goal of improving ethanol production from biomass (Gunasekaran and Chandra Raj, 1999). Metabolic engineering of *Z. mobilis* to produce a xylose fermenting strain was based on supplementing the wild type genome with genes responsible for pentose uptake and catabolism (Zhang *et al*., 1997).

To introduce the pentose metabolic pathway from *E. coli*, genes coding for xylose isomerase, L-ribulokinase, xylulokinase, L-arabinose isomerase, L-ribulose-5-phosphate, 4-epimerase, transaldolase and transketolase were integrated into the *Z. mobilis* genome. A multicopy plasmid was used to introduce these pentose pathway genes into *Z. mobilis* (Zhang *et al*., 1997). This multicopy plasmid mutant is a stable pentose and hexose sugar fermenting strain and is able to ferment xylose, glucose and arabinose produced by biomass hydrolysis (Zhang *et al*., 1997).

1.3.5 *Geobacillus* spp.

The genus *Geobacillus* consists of thermophilic microorganisms capable of fermenting sugars into alcohol (ethanol) and organic acids (Popova *et al*., 2002; Blumer-Schuette *et al*., 2008). These bacteria utilise pentose and hexose sugars as carbon sources. Members of the genus *Geobacillus* were first classified as phenotypically and phylogenetically coherent *Bacilli* (group 5) before their reclassification to the genus *Geobacillus* where *Geobacillus stearothermophilus* was the type strain (Nazina *et al*., 2001).

Geobacillus species are isolated from a variety of high temperature environments: hot springs, compost and artificial high temperature biotopes (Maugeri *et al*., 2002). Members of the genus *Geobacillus* are used in a number of industrial processes. They are used in the field of biotechnology processing as a source of thermostable enzymes including pullanases, proteases and amylases and for the production of exopolysaccharides (Moriello *et al*., 2003).

Geobacillus species also produce bacteriocins. Bacteriocins have lytic activity against bacteria, including those that are medically important such as *Salmonella typhimurium* (Novotny and Perry, 1992). These bacteriocins reduce the risk of contamination in bioreactors where these organisms are used as biocatalysts.

The growth temperature range of members of the genus varies from 45 to 75 °C, with an optimum growth temperature of 55 ˚C – 65 ˚C and a pH range of 6

- 8 (Thompson *et al*., 2008). *Geobacillus* species are potential ethanologens for biomass fermentation due to the flexibility of their cultivation conditions (they are facultative anaerobes and utilise pentose and hexose sugars) (Payton, 1984). Higher culturing temperatures of up to 60 °C can be used when working with *Geobacillus* cultures (Riyanti and Rogers, 2009).

Production of ethanol at high temperatures offers many advantages. These include obviating the need for process cooling and increasing the solubility of carbohydrates. In addition there is no risk of contamination by mesophiles (Zhang *et al*., 2007). Fermentation at high temperatures reduces the solubility of gases making it easy to maintain a near-anaerobic environment. The use of thermophiles such as *Geobacillus* spp. for biomass fermentation is limited by high by-product formation, poor ethanol tolerance and growth inhibition from the by-products furfural (a pre-treatment product) and mixed acids (fermentation products) (Galbe and Zacchi, 2002).

1.4 Use of *Geobacillus* spp. for ethanol production

Members of the genus *Geobacillus* are ideal biocatalysts to use in second generation bioethanol production because of their ability to utilise both hexose and pentose sugars. *G. thermoglucosidasius* produces ethanol concentrations of 0.2 g/l during the exponential growth phase (Riyanti and Rogers, 2009). This limits its application in its current form as a commercial strain for bioethanol production.

Thermophilic ethanologens such as members of the genus *Geobacillus* lack *pdc* and *adh* genes (Zhang et al., 2007; Riyanti and Rogers, 2008). To achieve a higher ethanol yield and simultaneously solve the mixed acid production problem in *G. thermoglucosidasius* 11955, genetic manipulation was employed to redirect mixed acid (acetate, pyruvate, formate *etc*.) pathways towards ethanol production (Cripps *et al*., 2009). Pyruvate dehydrogenases were over-expressed in *G. thermoglucosidasus* NCIMB 11955 and both the *ldh* and *pflB* genes were knocked out (Cripps *et al*, 2009). After genetic manipulation, a mutant strain (*G. thermoglucosidasius* TM242) was created. *G. thermoglucosidasius* TM242 produces no lactate and formate and less acetate than the wild type strain (Cripps *et al*., 2009). Accumulation of pyruvic acid in the mutant strain was also lower than that in the wild type strain (Cripps *et al*., 2009).

Although these *Geobacillus* species utilise a wide range of sugars and their development to improve ethanol yield was successful, they have a low ethanol tolerance. *G. thermogucosidasius* M10EXG has a higher ethanol tolerance than other *Geobacillus* species with *G. thermoglucosidasius* M10EXG and *G. thermoglucosidasius* NCIMB 11955 have maximum ethanol tolerances of 10 and 4 % v/v respectively (Galbe and Zacchi, 2002; Maugeri *et al*., 2002).

Should these species be engineered to tolerate ethanol concentrations to approximate that of *S. cerevisiae*, they would be highly efficient for biomass fermentation. *Geobacillus thermoglucosidasius* M10EXG utilises xylose and

glucose simultaneously to produce ethanol (Riyanti and Rogers, 2009). When glucose is abundant, xylose metabolism is suppressed through glucose repression of the xylose uptake mechanism (Riyanti and Rogers, 2009).

1.5 Ethanol tolerance and the general stress response

Bacteria encounter a variety of stressful conditions during growth (Sardessai and Bhosle, 2002). Heat and solvent stress stimulates the induction of stress response mechanisms (Laksanalamai and Robb, 2003). Stress response mechanisms of bacteria include oxidative, solvent and temperature stress response mechanisms (Amartey *et al*., 1998). The site of action for solvent (ethanol) stress is the cell membrane (Sardessai and Bhosle, 2002) which plays a vital role in the survival of bacteria by controlling movement of compounds in and out of the cell.

The toxicity of a solvent depends on its log P value (Sardessai and Bhosle, 2002). The log P value of a compound is the ratio of the compound's organic to aqueous phase concentrations (Hirakawa *et al*, 2005). Solvents with lower log P values are more toxic as they have a higher efficiency of partitioning into the cell membrane. The toxicity of a solvent is dependent on the amount which accumulates in the cell membrane rather than on its chemical composition (Sardessai and Bhosle, 2002). Solvent tolerance levels differ between organisms, a factor determined by physiological properties such as the amount of unsaturated fatty acids in the membrane lipid bilayer (Taylor *et al*., 2008). The extent to which the lipid bilayer of an organism tolerates solvent stress is

dependent on its composition (Michel *et al*., 1985). Accumulation of alcohol in the cell membrane causes the cell to leak by disrupting its normal structure (Sardessai and Bhosle, 2002; Taylor *et al*., 2008; Ding *et al*., 2009).

When partitioning into the lipid bilayer, ethanol resides at the water/lipid interface. This phenomenon is dependent on the chemical composition of the polar groups at the water/lipid interface (Nizza and Gawrisch, 2009). Lipid bilayers with phosphatidylglycerol, phosphatidylserine and sphingomyelin at their lipid/water interface are less susceptible to ethanol partitioning. These lipids bilayers have higher partition coefficients compared to those with phosphatidylethanolamine bilayers (Nizza and Gawrisch, 2009). The functional implications of ethanol binding to proteins are poorly understood. A number of research groups are focussing on elucidating the effect of ethanol on the functionality of biopolymers (Westerman *et al*., 1988).

Physiological characteristics of organisms that tolerate high ethanol concentrations have been studied. The ethanol tolerance of *S. cerevisiae* is associated with up regulation of zinc finger proteins (MacPherson *et al*., 2006) and prolyl hydroxylase (PHD) finger 1 proteins (Betz *et al*., 2004) which are involved in the expression of stress related genes. Changes in expression of heat shock protein regulator genes due to ethanol stress suggests heat shock proteins play a role in the ethanol tolerance of bacteria (MacPherson *et al*., 2006).

A number of metabolic pathways are involved in the stress response mechanism. Increased production of heat shock proteins is one of the mechanisms used to combat solvent and heat stresses (Park *et al*., 2001). As the concentration of solvent increases, a point is reached when stress response mechanisms cannot withstand the enzyme denaturing effect of the stress stimuli and bacterial growth is hampered (Sardessai and Bhosle, 2002; Taylor *et al*., 2008). Some heat shock proteins are up regulated due to temperature changes but not by high ethanol concentrations while for others the opposite is true (Dubaquié *et al*., 1997). Much stress tolerance research focuses on temperature stress while solvent (ethanol) stress is less studied (Michel *et al*., 1985).

1.5.1 Heat-shock proteins

Heat shock proteins are a class of functionally related proteins, whose expression is up regulated when an organism is exposed to stress. Increased expression of heat shock proteins is transcriptionally regulated by heat shock factors (Rasouly and Ron, 2009). Heat shock proteins are found in every organism, from bacteria to humans (Laksanalamai and Robb, 2003). Heat shock protein nomenclature is based on their molecular weight. For example, Hsp60 and Hsp10 are 60 kDa and 10 kDa in size respectively (Kim *et al*., 1996). Heat shock proteins with molecular weights between 15 kDa and 42 kDa are known as small heat shock proteins (Laksanalamai and Robb, 2003).

Heat shock proteins are intracellular chaperones for other proteins: they play a role in protein folding and assist the formation of functional protein conformation (Kim *et al*., 1998). Some heat shock proteins are expressed under ambient conditions. These heat shock proteins are part of the house keeping proteins and they differ from organism to organism and, in some species, from organ to organ (Rasouly and Ron, 2009).

Chaperones are proteins that assist the formation of non-covalent interactions between proteins or within proteins that result in their folding or unfolding and assembly or disassembly, as well as in protein transportation (Chen *et al*., 2006). Chaperones do not take part in their substrate's biological functions (Segal and Ron, 2006). Chaperonins are a subclass of chaperones directly involved in the folding process of non-native polypeptides into their native functional state.

The chaperone group is sub-divided into two groups (Luo *et al*., 2009). Group 1 chaperonins are found in bacteria and organelles of endosymbiosis (Luo *et al*., 2009). In *E. coli*, GroEL and GroES are referred to as chaperone 60 and chaperone 10 respectively. The group 1 chaperonins found in prokaryotes are characterised by α-crystalline conserved sequences (Luo *et al*., 2009). Group 2 chaperonins are found in the eukaryotic cytosol and in archaea (Fenton and Horwich, 2003).

The pathways through which unfolded proteins are processed are shown in figure 1.3. Once a protein is produced by the ribosome, it is taken up by the DnaK/DnaJ/GrpE complex (Luo *et al*., 2009). The unfolded protein is pre-processed into a conformation that is identified by the GroES/GroEL complex (GroESL) which further processes the protein to its functional conformation (Ranford *et al*., 2000). The two complexes (GroES/GroEL and DnaK/DnaJ/GrpE) work hand in hand but can also refold proteins independently (Ziemienowicz *et al*., 1993). The GroESL complex is more efficient than the DnaK/DnaJ/GrpE complex (Ranford *et al*., 2000).

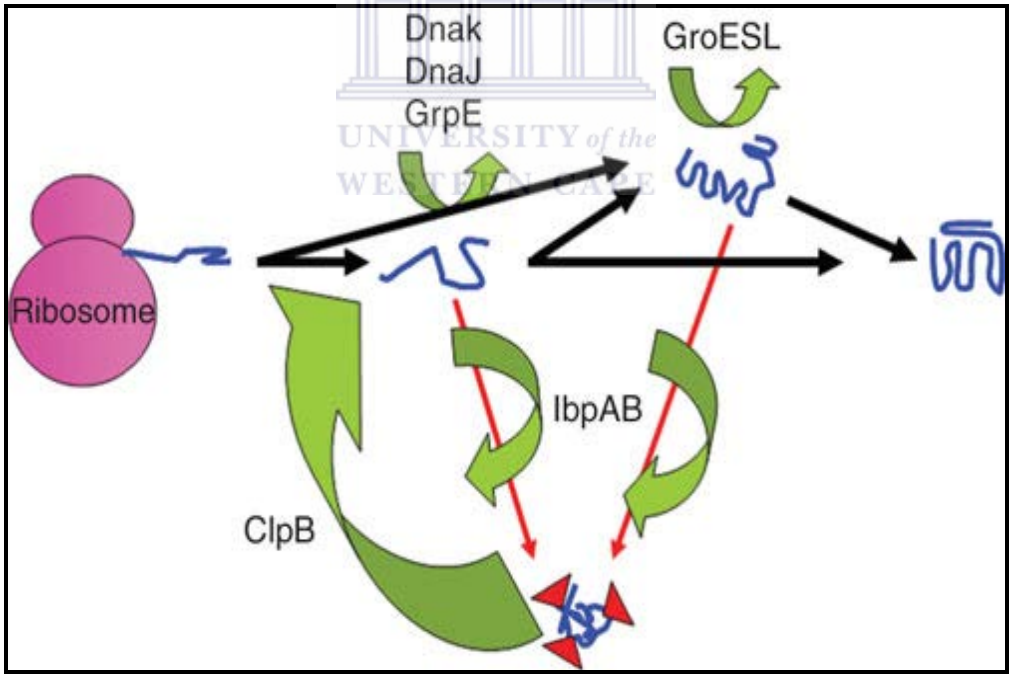


Figure 1.3: The pathways through which unfolded or denatured proteins are processed (adapted from Luo *et al*., 2009)

IbpAB and ClpB are small heat shock proteins that bind to unfolded or denatured proteins to prevent the proteins from aggregating (Figure 1.3) (Ziemienowicz *et al*., 1993). Small heat shock proteins bind to denatured proteins and give unfolded proteins a conformational structure that is recognisable by the GroESL and DnaK/DnaJ/GrpE complexes (Ziemienowicz *et al*., 1993).

1.5.2 Role of GroES and GroEL in protein folding

The GroES and GroEL proteins have been extensively studied since the observation of their involvement in stress tolerance. The mechanism and structure of the GroESL complex of *E. coli* was elucidated (Gupta, 1995) (Figure 1.4).

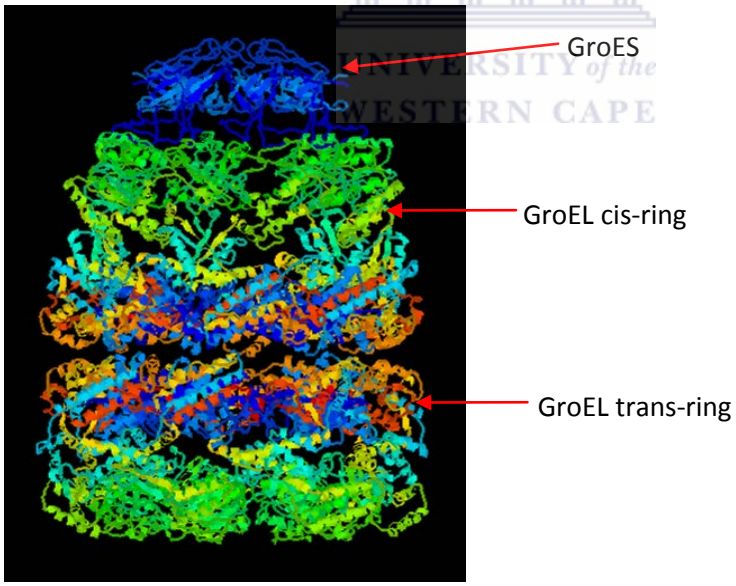


Figure 1.4: The GroESL complex of *E. coli* (Gupta, 1995)

The GroESL complex is composed of a 7 member GroES ring that caps the GroEL cis-ring. The GroEL complex has a membrane that separates the cis and

trans-ring. The other end of the complex (the trans side) is closed by a membrane. The rings communicate through allosteric structural changes (Dubaquié *et al*., 1997).

GroEL is composed of 14 subunits of 58 000 relative molecular weight each. The subunits attach to form a 2 stacked heptameric ring enclosing a central cavity (the substrate binding site) (Xu *et al*., 1997). Each subunit of GroEL is composed of three domains: an apical domain, an equatorial domain and the intermediate domain (Ranford *et al*., 2000). The domains are grouped into a cis-ring and a trans-ring. The cis-ring is composed of the apical domain while the trans-ring is composed of the intermediate and equatorial domains.

Prior to protein binding, the intermediate domain swings towards the equatorial domain, causing the central channel to pivot 25 degrees around Pro 137 and Gly 410 (Xu *et al*., 1997). Such a conformational change closes the binding site at the top part of the equatorial domain, simultaneously creating a number of interactions within the subunits and with neighbouring subunits. These interactions sterically hinder dissociation of ADP from the cis-ring and link the binding of GroES to the cis-ring ATP hydrolysis site. GroES is a single heptameric ring that forms an asymmetrical 1:1 complex with the ATP-bound GroEL cylinder. ATP binds at the equatorial domain (the cis-ring) of GroEL subunits. Once ATP is bound to the intermediate domain, it effects a conformational change that results in the alteration of the hydrophobic and

hydrophilic binding sites. This enlarges the central cavity and increases the efficacy of the complex for protein folding (Xu *et al*., 1997; Lou *et al*., 2009).

GroEL has ATPase activity equivalent to 5U/minute/monomer **(**Ranford *et al*., 2000). GroES attaches to the GroEL at the apical domain on the cis-ring, forming a dome like cover. GroES can only bind to ATP-activated GroEL. Through protein crystallography the conformational changes that occur as the GroESL/ATP complex prepares for protein folding was elucidated (Xu *et al*., 1997). On its own, GroEL interacts with unfolded proteins and prevents them from aggregating into an irreversible state (Truscott *et al*., 1994). This requires ATP. GroEL without a core-chaperone (ATP or GroES) increases the rate of protein aggregation (Luo *et al*., 2009).

1.5.3 Regulation of heat shock genes

As discussed in section 1.5.1, heat shock proteins play a major role in bacterial stress tolerance. Their expression is vital for cellular survival under normal or stressful conditions. Like all genes, the heat shock protein genes are regulated by promoters and other repression factors. In most bacteria, sigma factor 32 and sigma 70 (vegetative sigma factor) regulate the expression of heat shock genes (Burdon, 1986).

1.5.3.1 Sigma factor 32

An understanding of the mechanisms and factors that play a role in the regulation of heat shock protein expression is necessary before attempts can be

made to over-express them. In *E. coli*, the heat shock gene operons have heat shock promoters that are specific to sigma factor 32 transcription factors (Segal and Ron, 2006).

Sigma factor 32 is an activator which as a result of protease (HflB) activity, has a short half-life of approximately 4 min (Segal and Ron, 2006). When a cell is exposed to temperatures at which proteins denature, the accumulation of denatured proteins stimulates a series of events that lead to the stabilisation of sigma factor 32 thereby up-regulating the heat shock proteins. In bacteria such as *E.coli*, *Pseudomonas aeruginosa* and *Vibrio cholera* only sigma factor 32 is detected as a regulatory factor for heat shock operons (Rasouly and Ron, 2009).

Expression of heat shock proteins is dependent on the expression of sigma factor 32. An increase in production and stabilisation of this heat shock factor is therefore necessary for stress tolerance in bacteria such as *E. coli* (Rasouly and Ron, 2009). HflB protease is not the only protease that plays a role in regulating levels of sigma factor 32. HsiVU protease also degrades sigma factor 32 through an ATP dependent pathway (Yura *et al.*, 1993). Under ambient conditions, the synthesis of sigma factor 32 is repressed post transcriptionally (Straus *et al*., 1987).

The sigma factor 32 mechanism of heat shock regulation does not apply to all bacteria. Some bacteria use sigma factor 32 together with an inverted repeat, some use sigma factor 70 alone, and some use sigma factor 70 together with an

inverted repeat (IR) while some bacteria use more than one sigma factor (Straus *et al*., 1987).

1.5.3.2 Regulation of heat shock protein expression in *Geobacillus* species

Expression of heat shock genes in *Geobacillus* species is regulated by sigma factors (Segal and Ron, 2006). Unlike in *E. coli* where sigma factor 32 is the major player in heat shock protein expression, *Geobacillus* spp. heat shock protein expression is regulated by a number of sigma factors. Proteomic analysis of the soluble sub-proteome of *G. thermoleovorans* T80 revealed that sigma factors 37, w, 70 and 43 were produced in response to nutrient limitation and oxidative stress (Graham *et al*., 2006).

Expression of sigma factors in response to stress suggests that *G. thermoleovorans* T80 prepares itself for the production of proteins that would allow the cells to survive the imposed stress (Hecker *et al*., 1996). Sigma factors 43 and 70 belong to the group sigma factor A, which transcribes genes under the control of HrcA-CIRCE in Gram positive bacteria (Hecker *et al*., 1996). CIRCE is known as the Controlling Inverted Repeat of Chaperonin Expression (Reischl *et al*., 2002). HrcA is a negative control transcriptional repressor to CIRCE in group 1 heat shock operons (Hecker *et al*., 1996). Under optimum growth temperatures for bacteria, HrcA is bound to CIRCE, thereby repressing the expression of heat shock proteins. Under stress conditions HrcA dissociates from its operators thereby allowing the induction of the heat shock operons and expression of heat shock proteins (Reischl *et al*., 2002). The

activity of the HrcA system is modulated by the GroE chaperon system, where the GroE system maintains HrcA in a conformation that is able to bind CIRCE (Reischl *et al*., 2002).

When proteins are denatured in the cell, the GroE system and not HrcA bind to the proteins because the GroE system has higher affinity for denatured proteins than HrcA (Reischl *et al*., 2002). Since there is no GroE system to maintain a functional conformation of HrcA, HrcA cannot bind to CIRCE and repression of heat shock proteins ceases (Reischl *et al*., 2002). GroESL and DnaK operons are regulated by the sigma A promoter (sigma factor 43 and 70) and a conserved CIRCE (Reischl *et al*., 2002).

1.6 Aims and objectives of this study

Geobacillus thermoglucosidasius TM242 cannot tolerate ethanol concentrations beyond 4 % v/v (Maugeri *et al*., 2002). In previous proteomics studies, it was observed that heat shock protein 60 is up-regulated in *G. thermoglucosidasius* as ethanol concentrations increase (Charewa, 2008). Studies in other thermophiles reveal that heat shock proteins are up-regulated due to temperature and solvent stress (Michel *et al*., 1985). Following these observations, it was hypothesised that over-expression of heat shock proteins may result in an increased ethanol tolerance in *G. thermoglucosidasius* NCIMB 11955.

The small heat shock protein from *Pyrococcus furiosus* when over-expressed in *E. coli* BL21 was able to maintain the activity of *Taq* DNA polymerase, DNA restriction endonuclease *Hindlll* and lysozyme at elevated temperatures (Chen *et al*., 2006). Crystals of the GroEL protein of *Thermococcus litoralis* were produced when a recombinant *E. coli* strain harbouring the cloned GroEL gene was grown at temperatures exceeding 80 °C. It is evident that heat shock genes from thermophiles can be expressed in *E. coli* (Osipiuk *et al.*, 2000) and can enhance the protection of proteins from denaturation due to stress (Chen *et al*., 2006).

In this study the focus was to further understand ethanol tolerance in *G. thermoglucosidasius*. The effect of over-expressing *G. thermoglucosidasius* NCIMB 11955 heat shock proteins (GroES and GroEL) on ethanol tolerance will be tested by expressing these proteins in *Escherichia coli*. The major objectives of this research are listed:

I. **Further characterisation of the relationship between ethanol tolerance and growth temperature in *G. thermoglucosidasius* NCIMB 11955.** The ethanol tolerance of *G. thermoglucosidasius* NCIMB 11955 at growth temperature of 45 and 55 °C in complex media will be studied.

II. **Amplification and expression of the GroELS operon of *G. thermoglucosidasius* NCIMB 11955 in *E. coli* Rosetta pLysS.** The GroELS operon of *G. thermoglucosidaius* NCIMB 11955 has been implicated in acquired ethanol tolerance in this strain (Charewa, 2008). In

order to characterise this operon *in vitro* and *in vivo*, expression constructs within heterologous hosts are required.

III. ***In vivo* characterisation of imparted stress tolerance in expression hosts**. Expression strains will be subjected to high exogenous ethanol concentrations in order to determine if the GroELS complex could enhance strain stress tolerance and complement the general stress response.

Chapter 2: Materials and methods

2.1Chemicals and reagents

Chemicals were obtained from the following suppliers: Sigma Aldrich, Merck, Kimix Chemicals and Laboratory Suppliers.

Endonuclease enzymes, polymerase enzymes, DNA and protein markers were purchased from Fermentas Life Science Ltd. Primers for polymerase chain reactions were synthesised and supplied by Inqaba Biotech.

2.2 Buffers and solutions

Distilled water was used to prepare buffers and solutions. The pH of buffers was adjusted using 5 M NaOH or 1 M HCl. A Crison pH-meter (Basic 20+) was used for pH determination. Solutions that required autoclaving were autoclaved at 121˚C for 20 min.

10 mM Tris, 1 mM EDTA buffer pH 8 (TE)

1 ml of 1 M Tris (pH 8) and 0.2 ml of 0.5 M EDTA were added to a 250 ml autoclaved bottle containing 98.8 ml of sterile water. The mixture was shaken and aliquoted into 10 ml aliquots. The 1 x TE was stored at room temperature.

Sodium phosphate buffer 1 M (pH 8)

14.2 g of Na_2HPO_4 was dissolved in 10 ml of sterile water. Simultaneously and separately, 1.56 g of NaH_2PO_4 was dissolved in 10 ml of sterile water. A volume

of 9.3 ml of the Na_2PO_4 stock was mixed with 6.8 ml of the NaH_2PO_4 stock. The mixture was adjusted to pH 8 and filter sterilised before use.

2.3 Media

Media components used in this study were supplied by Oxoid and Biolabs. Autoclaving of media and components was performed at 121 °C for 20 minutes.

2 x TY media (Taylor *et al.*, 2008)

2x TY media was used for culturing overnight cultures of *G. thermoglucosidasius*.

Constituent	g/L
Tryptone	16
Yeast extract	10
NaCl	5

The constituents were added to distilled water and autoclaved at 120 °C for 15 minutes. Ampicillin, Xgal and IPTG were added to cooled (50 °C) medium where necessary.

Luria Bertani broth and agar (LB)

LB broth was used as a cellular propagation medium. LB agar was used for colony culturing of *E. coli* cells.

Constituent	g/L
Tryptone	10
Yeast extract	5
NaCl	10

LB agar was prepared from LB broth by adding 13 g/L of bacteriological agar. The constituents were added to distilled water and autoclaved at 120 °C for 15 minutes. Ampicillin, Xgal and IPTG were added to cooled (50 °C) medium where necessary.

SOC medium

SOC media was used for reviving newly transformed *E.coli* cells.

Constituent	g/L
Tryptone	20
Yeast extract	5
NaCl	0.5
KCl (250Mm)	10 mL

Constituents were added to distilled water and the medium was autoclaved and cooled to 50 °C prior to the aseptic addition of 5 ml of filter sterilised 2 M $MgCl_2$ and 20 ml of 1 M glucose.

2.4 Bacterial strains and plasmids

The following strains (Table 2.1) and plasmids (Table 2.2) were used in this study.

Table 2.1: **Bacterial strains used in this study**

Bacterial strains	**Supplier**
E. coli Gene Hog	Invitrogen (USA)
E. coli Rosetta2 (DE3) pLysS	Novagen (USA)
Geobacillus thermoglucosidasius NCIMB 11955	TMO Renewables (Surrey, UK)
Geobacillus thermoglucosidasius M10EXG	Department of Biochemistry, Ohio State University, USA

Table 2.2: Plasmids used in this study

Plasmid	**Characteristics**	**Supplier**
pET21a (+)	Expression vector containing an ampicillin resistance gene and a T7 *lac* promoter	Novagen
pGem-T-Easy	Cloning vector containing an ampicillin resistance gene.	Fermentas Life Sciences

2.5Effect of temperature on the growth and ethanol tolerance of *G. thermoglucosidasius* NCIMB 11955

2.5.1 Effect of temperature on the growth

G. thermoglucosidasius NCIMB 11955 cells were plated on LB agar plates and incubated at 55 °C overnight. A single colony was inoculated into 12 ml of SOC broth and incubated at 55 °C overnight with shaking. Three flasks containing 50ml 2 x TY broth were pre-warmed at 55 °C and 3 others were pre-warmed at

45 °C. 200 µl of the overnight culture was inoculated into each flask. The flasks were incubated at 55 °C and 45 °C respectively with shaking. The OD_{600} of the cultures were measured periodically.

2.5.2 Effect of temperature on ethanol tolerance

Geobacillus thermoglucosidasius NCIMB 11955 cells were streaked on LB agar plates and incubated at 55 °C overnight. A single colony was inoculated into 12 ml of SOC broth and incubated at 55 °C overnight with shaking. In duplicate, 200 µl of the overnight culture was added to 50 ml of pre-warmed 2x TY broth containing ethanol concentrations of 0, 2, 4, 6, 5 and 8 % v/v. One of each duplicate culture was incubated at 55 °C and the other at 45 °C with shaking.

2.6 General recombinant DNA procedures

2.6.1 DNA quantification

For routine quantification, DNA concentrations were determined using a Nanodrop ND-1000 instrument. DNA was re-suspended in double distilled water overnight at 4 °C. For more accurate quantification, DNA concentrations were measured using the Qubit™ DNA assay kit according to the recommended procedures.

2.6.2 Gel extraction and DNA purification

DNA fragments were briefly visualized under UV illumination at a peak wavelength of 302 nm and excised from agarose gels using a sterile scalpel blade. A GFX PCR DNA gel band purification kit (GE Healthcare Life Sciences) was

used to purify the DNA from the gel slices according to the manufacturer's instructions. The DNA was eluted in 10 mM Tris-buffered double distilled water at pH 8.0.

2.6.3 Plasmid DNA purification

The Macherey-Nagel Nucleospin Extract ll PCR Clean Up and Gel Extraction kit was used to purify DNA from solutions and agarose gels according to the manufacturer's instructions.

2.6.4 Agarose gel electrophoresis

Analysis of DNA using agarose gel electrophoresis was performed according to the method of Sambrook and Russell (2001) with minor adaptations. TAE agarose gels containing 1 % agarose w/v were cast and electrophoresis was performed at 100 v in 0.5 X TAE buffer. A concentration of 0.5 µg/ml of ethidium bromide was added to the agarose gels to allow visualisation of DNA under UV transillumination. DNA fragments were sized based on their migration on gels in comparison to the migration of fragments of a *Pst* I digested lambda DNA marker. An Alpha Imager was used to visualise agarose gels.

2.7 Genomic DNA extraction

Genomic DNA extraction was carried out according to the method described by Sambrook and Russelll (Sambrook and Russell, 2001).

A single colony of *G. thermoglucosidasius* NCIMB 11955 was inoculated into 5 ml volumes of pre-warmed (pre-incubated at 55 °C for at least 30 min) 2x TY medium. The culture was incubated at 55 °C overnight. From the overnight culture, a 0.5 ml aliquot was inoculated into two separate 10 ml volumes of pre-warmed 2x TY medium. The cultures were incubated overnight at 55 °C. The overnight culture was centrifuged (Eppendorf 5810 R centrifuge) at 10 000 x g at 4 °C for 10min to harvest cells. The cell pellet was re-suspended in 0.5 ml of phosphate buffer saline (PBS) lysis buffer. The cell suspension was frozen on ice for 15 to 20 min, thawed and sonicated with a Bandelin Sonopulse 2070 sonicator. Sonication was done three times for 10 s at a power setting of 30 W. A final concentration of 0.1 mg/ml proteinase K was added to the lysate suspension and the lysate was incubated for 60 min at 37 °C. An equal volume of 25:24:1 phenol/chloroform/isoamyl alcohol was added to the suspension. The mixture was centrifuged for 5 min at 16 000 x g. The aqueous layer was transferred to a clean 1.5ml micro-centrifuge tube. DNA was re-extracted with an equal volume of 25:24:1 phenol/chloroform/isoamyl alcohol. The aqueous layer was transferred to a clean tube after every phenol/chloroform/isoamyl alcohol extraction. DNA was further extracted by adding equal volumes of 24:1 chloroform /isoamyl alcohol. To every 400 µl of aqueous phase, 15 µl of 5 M NaCl was added. The Eppendorf

tube was filled with ice cold 99.6 % ethanol and incubated overnight at -20 °C. The solution was centrifuged for 15 min at 16 000 x g at 4 °C. The pellet was washed with 500 µl of 70 % ethanol and air dried for 15 min. The pellet was re-suspended in 50 µl of sterile water and stored at -20 °C.

2.8 Amplification of the GroES and GroEL genes

The GroESL operon genes were amplified independently. Primer sets employed in this study are listed in Table 2.3. When designing primers, the Integrated DNA Technologies tool Oligonucleotide Analyser was used to determine useful sequences. Restriction sites were incorporated into the primer sequences for directional cloning (Table 2.3).

The highlighted region in primer W1 represents an incorporated *Nde* I restriction site while the highlighted region in primer W3 is an *Nde* I restriction site. An *EcoR* 1 restriction site is highlighted in the W2 reverse primer. Highlighted on reverse primer W4 is the *Not* 1 restriction site.

Table 2.3: Primer sequences and PCR conditions for amplification of genes in the GroESL operon

Genes amplified	Primers used	PCR conditions	Reference
GroES	W1: Forward primer 5'CGGGGTCATATGC-TAATAAGAACGGCG3'	95°C for 30s, 30x (95°C for 30s, 55°C for 30s, 72°C for 40s), 72°C for 10 min	This study
	W2: Reverse primer 5'AGGAGGGCTCATT-AACCAATCACAGCC3'		
GroEL	W3: Forward primer 5'-CGGCGCGGCGCA TATGGCAAAAGAAAT-3'	95°C for 30s, 30x (95°C for 30s, 55°C for 30s, 72°C for 40s), 72°C for 10 min	This study
	W5: Reverse primer 5'-TAGCGAGCTCTTAC ATCATTCCGCCCATC-3'		
M13 Primers	Forward primer 5'-GTTTTCCCAGTC ACGAC-3'	95°C for 30s, 30x (95°C for 30s, 55°C for 30s, 72°C for 40s), 72°C for 10 min.	Life Technologies
	Reverse primer 5'-AGCGGATAACAATT TCACACAGGA- 3'		

PCR amplification was performed in 0.2 ml thin walled tubes using an Eppendorf master-cycler gradient thermo-cycler. A standard 20 µl PCR reaction contained 10µl of sterile distilled water, 0.5µM of forward primer, 0.5 µM of reverse primer (Table 2.3), 0.2 mM of each dNTP (dATP, dTTP, dCTP, dGTP), 1U of Taq DNA polymerase, 1ng/µl of bovine serum albumin, 2 µl of 1x Taq buffer and 4 ng of

plasmid DNA or 25 ng of genomic DNA. The annealing temperature used for GroES and GroEL amplifications was 55 °C.

2.9 Cloning of PCR products

A ligation reaction of 10 µl final volume was used throughout. Reactions consisted of a 3:1 ratio of insert to vector, 1 µl of T4 DNA ligase enzyme, 5µl of 2 X T4 DNA ligase buffer. The volume was made up to 10 µl with the addition of sterile water. For the 3:1 insert to vector addition 20 ng of plasmid DNA and 60 ng of each gene fragment were added to the ligation reagents. The ligation mixture was incubated overnight at 4 °C. The GroES and GroEL fragments were independently cloned into the pGem-T-Easy vector (Fementas Life Sciences) before they were excised and cloned into the pET21 (a) + vector. The GroES fragment was cloned into the pET21 a (+) multiple cloning site using restriction sites *Nde* I and *EcoR* I. The clone was designated pET21 (a) +/ GroES. GroEL was subsequently cloned downstream of the GroES fragment at the *Not* 1 restriction site situated in the multiple cloning site of the vector. The pET21 (a) +/ GroES construct was treated with shrimp alkaline phosphatase prior to ligation with the GroEL gene fragment.

2.10 Preparation of *Escherichia coli* competent cells

The Sambrook and Russell (2001) method for preparation of competent cells was used. Glassware was pre-washed with 70 % ethanol and autoclaved before use. An overnight culture was prepared by inoculating a single colony of E. coli Rosetta pLySs into 5ml of LB broth and incubated overnight at 37 °C. From the

overnight culture, 1ml was inoculated into 50ml of LB broth and incubated at 37 °C until the culture reached an OD_{600} of 0.4. The culture was cooled on ice for 20 min. Cells were harvested by centrifugation (Eppendorf 5810R fixed rotor centrifuge) at 3000 for 10min. The supernatant was decanted and the cells were re- suspended in 10ml of ice cold 0.1 M $CaCl_2$ and incubated on ice for 20 min. The cells were harvested by centrifugation as described above and suspended in 2 ml of 0.1 M $CaCl_2$ containing 15 % v/v glycerol. Cells were aliquoted into 50 µl volumes and stored at -80 °C.

2.11 Transformation of competent cells

Chemically competent cells stored at -80 °C were thawed on ice for 5min prior to the addition of 5 µl of a pre-cooled ligation mixture. The mixture was incubated on ice for a further 20min. The cells were heat shocked at 42 °C for 1 min and were returned to ice for 2 min. A volume of 400 µl pre-warmed SOC broth (Section 2.3.4) was added. The solution was incubated at 37 °C for 60min with shaking. A volume of 50 µl of the cells was plated onto LB agar plates containing ampicillin (100 µg/ml), 0.1mM IPTG and 40 µg/ml X-Gal. The plates were incubated overnight at 37 °C.

2.12 Screening for GroES and GroEL habouring clones

Positive transformants (white colonies) were inoculated into 5 ml of LB broth containing 100 µg/ml of ampicillin. The cultures were incubated overnight at 37 °C prior to plasmid extraction. Restriction digestion (section 2.11.1), PCR amplification (Section 2.8) and DNA sequencing were used to screen for positive

clones after each cloning step. Plasmid DNA purification was done as described in section 2.6.2.

2.12.1 Restriction enzyme digestion

Restriction enzyme digests were done in 0.6 ml microcentrifuge tubes. A final volume of 10 µl was used. The restriction enzymes *EcoR* l and *Nde* l were used to confirm the presence of the GroES insert after each cloning attempt. Restriction enzyme *Not* 1 was used to confirm the presence of the GroEL gene fragment. The pET21/GroESL construct was digested with restriction enzyme *EcoR* 1 to predict the arrangement of the GroES and GroEL genes in the pET21 (a)+ vector. Fragment sizes of 1600, 1000, 600 and 400 bps were expected after digestion with *EcoR* I if the GroEL fragment lay in the same orientation as the GroES fragment. Reactions were set up so that 1U of each restriction enzyme was used to digest 1 µg of plasmid DNA. Reactions were incubated at 37 °C for 2-24 hours. The restriction products were analysed by gel electrophoresis on 1 % agarose gels (section 2.5.4).

2.12.2 Sequencing

Sequencing of cloned insert DNA was performed by the DNA Sequencing Group, Department of Molecular and Cell Biology (MCB) at the University of Cape Town (UCT) using M13 forward and reverse primers (Table 2.3). The Bioedit and DNAMAN programmes were used for sequence analyses and annotation of the sequences. The edited and annotated sequences were submitted to the National

Centre for Bioinformatics (NCBI) server (http://www.ncbi.nlm.nih.gov/blast/) for Blast analysis.

2.13 Protein analysis

2.13.1 Over-expression of the GroES and GroEL genes

A single *E. coli* Rosetta colony carrying the pET21/GroESL construct was inoculated and incubated overnight at 37 °C in 5 ml LB broth containing 100 mg/ml ampicillin. Simultaneously, the negative control (*E. coli* Rosetta pLysS carrying the pET21 a (+) vector without insert) was inoculated and incubated as above. A volume of 1ml of each overnight culture was inoculated separately into 50 ml LB broth containing 100 mg/ml ampicillin. The cultures were incubated at 37 °C and growth was monitored by measuring optical density (OD) using a spectrophotometer at 600 nm (Biomate 3, Thermo Electron Corporation). At an OD_{600} of 0.4 the cultures were induced for expression of the GroESL genes by the addition of 500 µl of 0.1 M IPTG. The induced cultures were incubated at 37 °C for 6 hours. Cells were harvested by centrifugation in an Eppendorf 5810 R centrifuge at 10 000 for 5 min. Cell pellets were re-suspended in 350 µl of PBS buffer. The cell suspensions were sonicated using a Bandelin Sonopulse 2070 sonicator at 30 W; sonication was performed three times for 30 s. The lysed cells were transferred to a 1.5 ml microcentrifuge tube before centrifugation at 16 000 for 15 min to pellet cell debris. The supernatant (intracellular protein solution) was transferred to a clean 1.5 ml micro tube. The concentration of intracellular protein was determined using the Bradford assay (section 2.13.2). The cell debris pellet containing insoluble protein was re-suspended in 20 µl of protein loading

dye. Proteins in the supernatant and insoluble fractions were separated and visualized by SDS PAGE. All protein samples were analysed on 12 % SDS polyacrylamide gels (section 2.14.3).

2.13.2 Determination of protein concentration

The Bradford assay (Bradford, 1976) was used to determine protein concentration. A bovine serum albumin stock solution of 2.4 mg/ml was used for preparing the standards of 10, 20, 40 and 50 mg/ml.

2.13.3 SDS polyacrylamide gel electrophoresis (PAGE)

Separation of proteins by SDS PAGE was performed using a Mighty Small vertical slab unit (Hoefer SE 280).Gels (1.5 mm thick) containing 12% acrylamide resolving and 4% acrylamide stacking gels were prepared using the Hoefer SE 245 dual gel caster. Protein samples (40 ng) were added to an equal volume of 2 X SDS PAGE loading buffer (Sambrook and Russell, 2001) prior to incubation at 90 °C for 5 min. An unstained protein marker (6 µl) was used for size determination. Electrophoresis was performed at a constant voltage of 70 V in 1 X running buffer (Sambrook and Russell, 2001). Gels were stained overnight with Coomassie Blue and destained four times using PAGE destaining solution (Sambrook and Russell, 2001).

2.14 Characterisation of the tolerance to ethanol of *E. coli* Rosetta pLySs/ pET21 (a) +/ GroESL

A single colony of *E. coli* pLySs/ pET21 (a) + / GroESL was inoculated into 5ml of pre warmed LB medium. The culture was incubated at 37 °C overnight. 4 ml of the overnight culture was inoculated into 630 ml 2TY broth and the culture was grown at 55°C with shaking. When the culture reached an OD_{600} of 0.8, the culture was divided into 12 x 250 ml flasks (50 ml per flask). All Flask where incubated at 37 °C with shaking. 6 duplicates where made from the 10 flask. Ethanol concentrations of 0, 2, 3, 4, 5 and 6 % v/v were added to the flasks, in a situation where there where 2x of each ethanol concentration containing flask. *E. coli* Rosetta pLySs was inoculated into one set of flasks containing 0, 2, 3, 4, 5 and 6 % ethanol v/v. *E. coli* Rosetta pLySs/ pET21 (a) +/ GroESL was inoculated into the other set of flask containing 0, 2, 3, 4, 5 and 6 % ethanol v/v. The cultures were incubated at 37 °C with shaking for a further 16 hrs. The OD_{600} of the cultures was measured periodically.

Chapter 3: Construction of the GroESL recombinant plasmid

3.1 Introduction

In a previous proteomic study, it was observed that heat shock proteins of *G. thermoglucosidasius* NCIMB 11955 were up-regulated as the ethanol concentration of the growth medium was increased (Figure 3.1) (Charewa, 2008). Protein spot P2 (Figure 3.1) was identified as the *G. thermoglucosidasius* NCIMB 11955 GroEL protein. Based on the above proteomic study and published literature (Kim *et al*., 1996), it was hypothesised that up-regulation of the GroESL operon of *G. thermoglucosidasius* NCIMB 11955 could improve the tolerance of the bacterium to ethanol. The aim of this section was to clone the GroESL operon of *G. thermoglucosidasius* NCIMB 11955 into the pET21 (a) + vector.

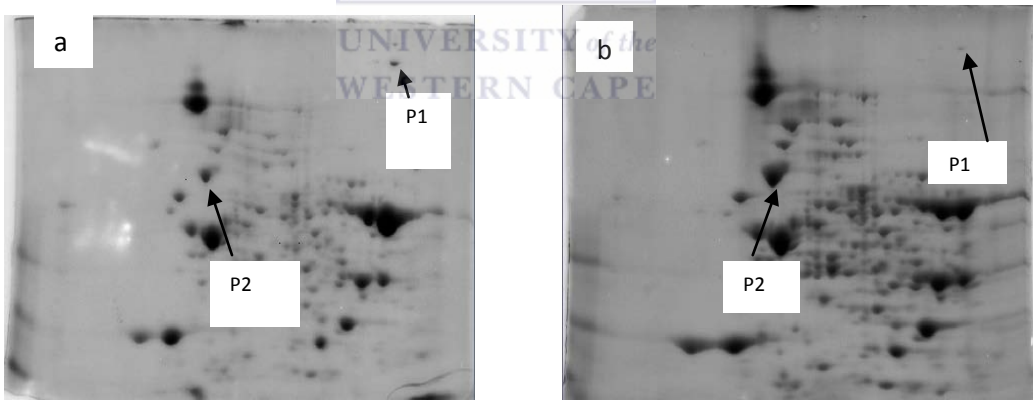


Figure 3.1: 2-D gels of the soluble proteins extracted from cultures of *G. thermoglucosidasius* NCIMB 11955 stressed with 0 and 5% ethanol, respectively. (a) 0% ethanol condition and (b) 5% ethanol conditions. Protein spot P1 was down and protein spot P2 was up-regulated as ethanol concentrations in the growth medium were increased. The gels were reproduced in triplicates. Similar patterns for each condition were noted obtained (Charewa, 2008)

Gene regulation factors control the expression of heat shock proteins (Segal and Ron, 2006). CIRCE secondary structure hairpin loops are regulatory elements of

the GroESL operon of bacteria such as *E. coli* and *G. thermoglucosidasius* (Section 1.6.3). These secondary structures are located downstream of the GroES gene and upstream of the GroEL gene within the 72 base pair sequence linking the two genes in *G. thermoglucosidasius* (Segal and Ron, 2006) (Figure 3.2). Hairpin loops affect PCR amplification - the effect of hairpin loops depends on loop size and the number of complementary bases (Singh *et al.*, 2000). Hairpin loops with less than four complementary bases have no effect on PCR amplification, but anything longer will hinder PCR amplification. The hairpin loop in the *G. thermoglucosidasius* GroESL operon is nine base pairs long and thus long enough to affect PCR amplification (Singh *et al.*, 2000). The size of the hairpin loop was a concern when attempting to replicate the GroESL operon via PCR amplification. Amplification of the entire GroESL operon was unsuccessful (data not shown) and the two genes (GroES and GroEL) were amplified independently.

GTGCTAATAAGAACGGCGAAAATTATGTTAAGGAGGTTGTTTTCCGTGATAAAGCCATTAGGTGATCGCGT

TGTCATTGAAATCGTTGAAACGGAAGAAAAAACTGCAAGCGGTATCGTATTGCCAGATACTGCAAAAGAAA

AACCGCAAGAAGGCAAAGTTGTTGCCGTTGGAAAAGGACGCGTACTTGACAACGGTCAACGCGTAGCTCCA

GAAGTGGAAGTTGGCGATCGCATTATCTTCTCGAAATATGCGGGTACAGAAGTGAAATATGACGGCAAAGA

ATACTTAATTTTGCGTGAAAGCGA**TATTTTGGCTGTGATTGGTTAA**TATATAGCGTTGATAACATAGATGTG

CAAAAAAATACTTAACGATT**TCATTTTACAAGGAGGTAACGGGGT**ATGGCAAAAGAAATTAAATTCAGC

GAAGAAGCTCGTCGTGCGATGCTGCGCGGTGTTGACAAACTAGCGATGCAGTAAAAGTAACGTTAGGTCCAA

AAGGCCGTAACGTTGTATTAGAGAAAAAATTCGGTTCTCCATTAATTACAAACGACGGTGTTACGATCGCGA

AAGAAATCGAATTAGAAGACCCATTTGAAAACATGGGTGCGAAGCTTGTTGCTGAAGTTGCAAGCAAAACA

AACGATGTTGCTGGGGACGGTACAACAACAGCGACAGTTTTAGCTCAAGCGATGATCCGTGAAGGCTTAAAG

AACGTAACAGCTGGCGCAAACCCAATGGGAATCCGCAAAGGTATTGAAAAAGCGGTTGCTGTAGCGGTAGA

AGAATTAAAAGCAATCTCCAAACCAATCCAAGGAAAAGAATCGATCGCGCAAGTTGCGGCTATTTCTGCGGC

TGACGAAGAAGTTGGCCAATTAATTGCAGAAGCAATGGAACGCGTCGGCAACGACGGTGTTATCACATTAG

AAGAATCAAAAGGTTTCACAACAGAATTAGATGTTGTGGAAGGTATGCAATTTGACCGCGGTTATGCGTCTC

CATACATGATCACAGATACAGAAAAAATGGAAGCAGTGCTTGAAAATCCATATATCTTAATCACTGACAAAA

AAATCTCGAACATTCAAGACATCTTGCCTATCTTAGAACAAGTTGTTCAACAAGGCAAACCATTGTTAATCAT

CGCGGAAGACGTCGAAGGCGAAGCGCTTGCAACATTAGTTGTTAACAAACTTCGCGGCACGTTCACTGCGGT

AGCGGTTAAAGCGCCTGGCTTCGGTGATCGCCGTAAAGCAATGTTGGAAGACATCGCAATCTTAACTGGCGG

TGAAGTCATCTCCGAAGAATTAGGACGCGAATTAAAATCAACAACAATTGCATCACTTGGCCGCGCTTCGAA

AGTTGTTGTAACGAAAGAAAATACAACAATCGTTGAAGGCGCTGGCGATTCTGAACGCATTAAAGCTCGCAT

CAACCAAATCCGCGCTCAATTAGAAGAAACTACTTCTGAATTCGACCGCGAAAAATTACAAGAACGTTTGGC

AAAACTTGCTGGCGGCGTAGCGGTCATCAAAGTTGGTGCAGCGACAGAAACAGAATTGAAAGAACGCAAAT

TGCGCATTGAAGACGCGCTCAACTCTACTCGTGCGGCTGTCGAAGAAGGTATCGTAGCCGGCGGTGGTACGG

CATTAATGAACGTATATAACAAAGTTGCTGCGATCGAAGCAGAAGGCGACGAAGCAACTGGTGTGAAAATC

GTTCTTCGCGCAATCGAAGAGCCAGTTCGCCAAATCGCGCAAAACGCTGGTTTGGAAGGCTCTGTCATTGTT

GAACGCTTAAAATCCGAAAAACCTGGCATCGGCTTCAACGCTGCTACTGGCGAATGGGTAAACATGATCGAA

GCTGGTATTGTTGACCCAACGAAAGTAACTCGCTCCGCTCTGCAAAACGCAGCTTCTGTTGCCGCTATGTTCT

TAACAACAGAAGCAGTTGTCGCTGACAAACCAGAAGAAAACAAAGGCGGCAATAGCGGAAT**GCCTGACAT**

GGGCGGAATGATGTAA

Figure 3.2: The GroESL operon of the *G. thermoglucosidasius* NCIMB 11955 genome (Li and Wong, 1992). The DNA sequence highlighted in blue is the GroES sequence which is followed by the 72 bp spacer (highlighted in green). The sequence highlighted in yellow following the 72 bp spacer is the GroEL sequence. The regions of the operon used to design the primers W1, W2, W3 and W5 are in bold and are located at the ends of each gene

The GroES and GroEL genes of *G. thermoglucosidasius* NCIMB 11955 were cloned independently into the pET21 (a) + vector multiple cloning site simulating their arrangement in the *G. thermoglucosidasius* NCIMB 11955 GroESL operon.

3.2 Results

3.2.1 PCR amplification of the GroES and GroEL genes from genomic DNA

The GroES and GroEL genes were amplified from the *G. thermoglucosidasius* 11955 genome using primers W1 and W2 (GroES) and W3 and W5 (GroEL) as described in Section 2.7. A volume of 8 µl of the PCR products was electrophoresed on a 1% agarose gel. The GroES (Figure 3.3) and GroEL (data not shown) fragments were gel excised and independently cloned into the pGem-T-Easy vector (Figures 3.4 and 3.5).

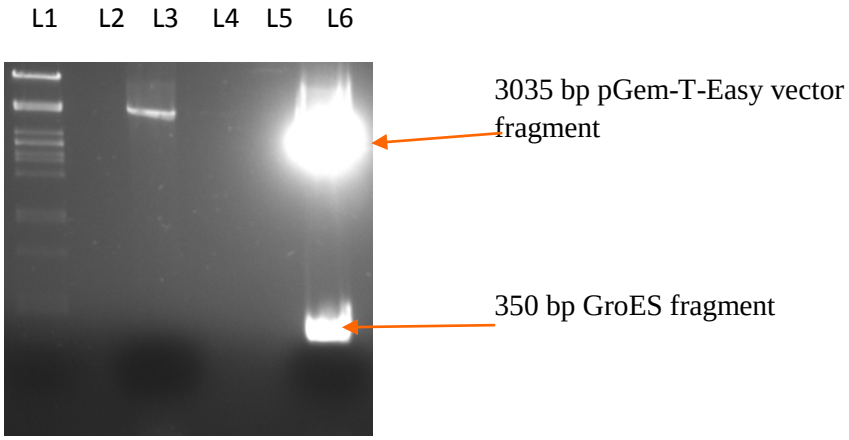


Figure 3.3: Agarose gel electrophoresis of restriction enzyme digests to assess the presence of GroES in the p-Gem- T- easy vector. Lane 1: λ DNA digested with *Pst* I. Lane 3: p-Gem-T-Easy digested with *Nde* 1/ *EcoR* 1. Lane 6: p-Gem-T-easy/ GroES digested with *Nde* 1/ *EcoR* 1

3.2.2 Cloning of the GroES and GroEL genes into the p-Gem-T-easy vector

The GroES (350 bp) and GroEL (1680 bp) fragments were gel excised from the gels and sequentially cloned into the p-Gem-T-easy vector.

Digestion of the p-Gem-T-easy/ GroEL construct using restriction enzymes *Nde* 1 and *EcoR* 1 released two bands corresponding to the p-Gem-T-easy backbone (3045 bp) and a fragment corresponding in size to GroEL (1680 bp) (Figure 3.4).

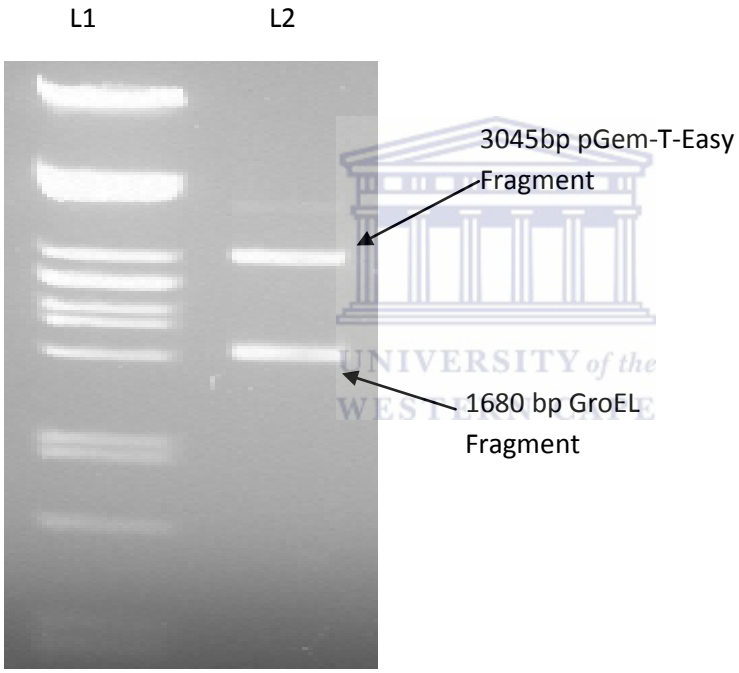


Figure 3.4: Agarose gel electrophoresis of the restriction enzyme digests to assess the presence of GroEL in the p-Gem- T- easy/ GroEL construct. Lane 1: λ DNA digested with *Pst* I. Lane 2: Restriction digestion of the pGem- T- easy/ GroEL construct digested with *Not* 1

3.2.3 Construction and analysis of the pET21 (a) +/ GroESL plasmid

The 350 bp GroES fragment (section 3.2.2) was cloned into the pET 21 (a) + vector using the restriction sites *Nde* 1 and *EcoR* 1. The resultant plasmid was designated pET21 (a) +/ GroES. pET21 (a) +/ GroES was subsequently linearized using the restriction enzyme *Not* 1 and was treated with shrimp alkaline phosphatase to prevent recircularisation. This was ligated with the 1680 bp GroEL fragment (section 3.2.2) resulting in the pET21 (a) +/ GroESL plasmid. Following both cloning procedures the plasmids were transformed into *E. coli* Rosetta pLysS and plated onto LB Agar containing ampicillin (100 µg/ml), 0.1mM IPTG and 40 µg/ml X-Gal.

Using the blue/ white clone selection positive clones were selected and the presence of the GroES and GroEL inserts were verified by restriction digestion, PCR amplification and by sequence analysis of the pET21 (a) +/ GroESL recombinant plasmid (Figures 3.5 - 3.9).

Restriction digestion using enzymes *Not* 1, *Nde* 1 and *EcoR* 1 confirmed the presence and direction of the GroES and GroEL fragments in the pET21 (a) +/ GroESL recombinant plasmid. A 1680 bp band corresponding to GroEL fragment and a 5720 bp band corresponding to the pET21 (a) +/ GroES were released when pET21 (a) +/ GroESL was digested with Not 1 enzyme (Figure 3.5).

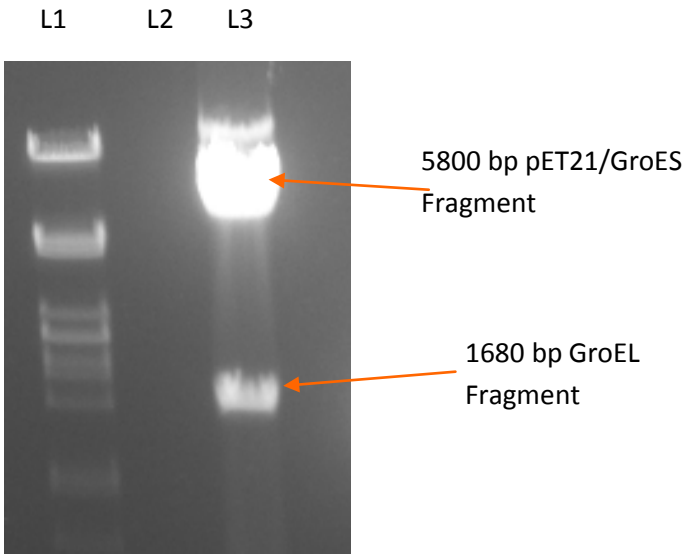


Figure 3.5: Restriction digest to assess the presence of GroEL in the pET21 (a) +/ GroESL construct. *Not* 1 was used to digest the construct. Lane 1: λ DNA digested with *Pst* I. Lane 3: pET21 (a) +/ GroESL construct digested with *Not* 1

Agarose gel electrophoresis of the pET21 (a) +/ GroESL recombinant plasmid digested with *Nde* 1 and *EcoR* 1 showed four bands of approximately 350, 570, 1000 and 5400 bp (Figure 3.6).

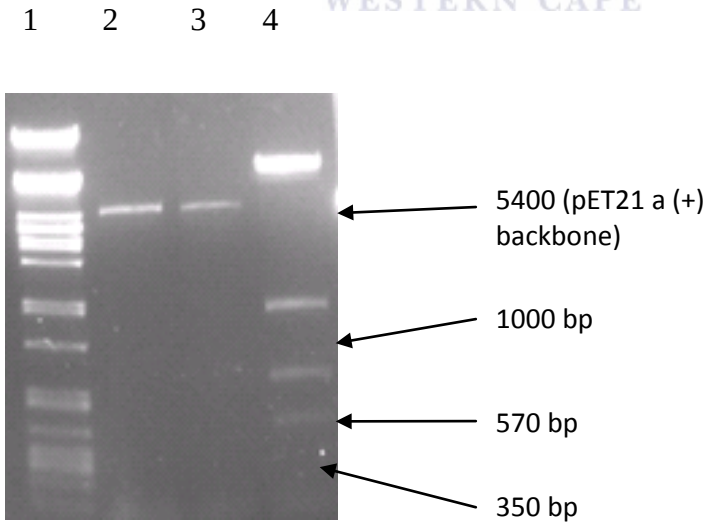


Figure 3.6: Agarose gel electrophoresis of a restriction digest of the pET21 (a) + / GroESL construct using *Nde* 1 and *EcoR* 1 enzymes. Lane 1: λ DNA digested with *Pst* I. Lane 2: p-Gem-T-easy vector digested with *EcoR* 1. Lane 3: p-Gem-T-easy vector digested with *Nde* 1. Lane 4: pET21 (a) + / GroESL construct digested with *EcoR* 1 and *Nde* 1 enzymes.

PCR amplification of the pET21 (a)+/ GroESL recombinant plasmid with the forward primer W1 and reverse primer W5 resulted in a 2000 bp band as expected (Figure 3.7). Both analyses (restriction digestion and PCR amplification) confirmed the presence and orientation of the GroES and the GroEL fragments in the pET21 (a) +/ GroESL recombinant plasmid.

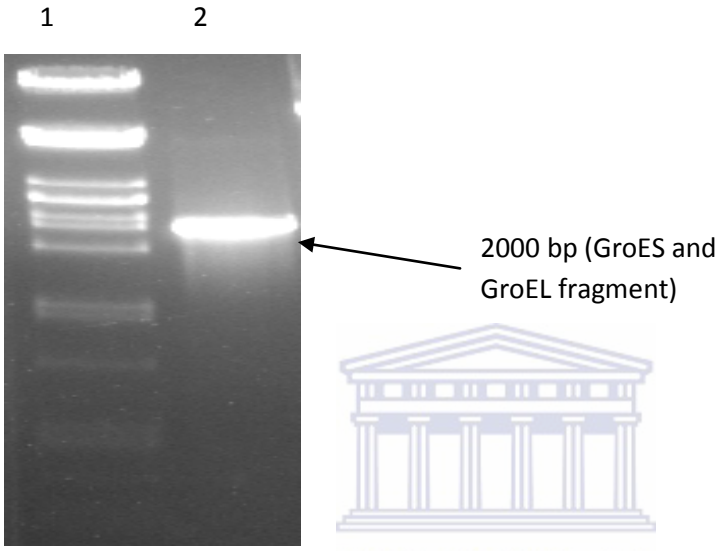


Figure 3.7: Agarose gel electrophoresis of PCR amplification products of the pET21 (a)+/ GroESL construct using primers W1 and W5, to verify the presence and direction of the GroES and GroEL fragments in the construct. Lane 1: λ DNA digested with *Pst* I. Lane 2: 2000 bp PCR amplicon.

A non-proof reading *Taq* polymerase was used in the initial amplification of the GroES and GroEL genes. DNA sequencing of the pET21 (a) +/ GroESL recombinant plasmid confirmed that the GroES and GroEL gene sequences were not altered during PCR amplification. DNA sequencing was also done to confirm whether the GroES start codon was in-frame with the T7 promoter of the pET21 (a) + vector, both genes were in the correct orientation and that the vector was intact.

3.2.4 Sequence analysis

Pairwise alignment was done using the BioEdit sequence analysis software. The sequence of the pET21 (a) +/ GroESL plasmid was aligned to the sequence of the pET21 (a) + vector published by Novagen and to the sequences of the GroES and GroEL genes provided by TMO Renewables. The cloned genes were identical in sequence to the database sequences. Figure 3.8 shows the alignment of the cloned GroES gene to the GroES sequenced provided but TMO Renewables. Pairwise alignment of the resultant sequence to the pET 21 (a) + sequence confirmed that the cloned GroES and GroEL genes were in-frame and adjacent to the T7 promoter in the pET21 (a) + vector fragment. The cloned GroES and GroEL genes and gene expression elements such as the pET21 (a) + T7 promoter and the ribosomal binding site sequences of the pET21 vector were correctly oriented.

```
....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
     65         75         85         95        105        115
WC Construct ACATATGCTA ATAAGAACGG CGAAAATTAT GTTAAGGAGG TTGTTTTCCG TGATAAAGCC
GroES        ----GTGCTA ATAAGAACGG CGAAAATTAT GTTAAGGAGG TTGTTTTCCG TGATAAAGCC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
    125        135        145        155        165        175
WC Construct ATTAGGTGAT CGCGTTGTCA TTGAAATCGT TGAAACGGAA GAAAAAACTG CAAGCGGTAT
GroES        ATTAGGTGAT CGCGTTGTCA TTGAAATCGT TGAAACGGAA GAAAAAACTG CAAGCGGTAT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
    185        195        205        215        225        235
WC Construct CGTATTGCCA GATACTGCAA AAGAAAAACC GCAAGAAGGC AAAGTTGTTG CCGTTGGAAA
GroES        CGTATTGCCA GATACTGCAA AAGAAAAACC GCAAGAAGGC AAAGTTGTTG CCGTTGGAAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
    365        375        385        395        405        415
WC Construct GCGTGAAAGC GATATTTTGG CTGTGATTGG TTAATGAGCC CTCCTAATCA CTAGTGAATT
GroES        GCGTGAAAGC GATATTTTGG CTGTGATTGG TTAA------ ---------- ----------
```

Figure 3.8: Pair wise alignment of pET21 (a) +/ GroESL (WC) to the GroES sequence provided by TMO Renewables. The *Nde* l restriction site is highlighted in red.

Based on the alignment in Figure 3.8, no DNA replication errors occurred during PCR amplification and cloning of the GroES gene. The GroEL sequence was observed on analysis of the T7 promoter sequencing results: GroEL starts at position 473 base pairs (data not shown). Pairwise alignment verified that the GroES and GroEL genes were cloned in-frame with the pET21 vector expression system and that there were no errors during PCR amplification.

A restriction map of the pET21 (a) +/ GroESL construct (Figure 3.9) was constructed from the DNA sequence using DNAMAN and BioEdit software. The restriction fragments obtained from the restriction digest of the recombinant pET21 plasmid (Figure 3.5 and 3.6) correlated with the fragments expected from restriction analysis of the pET21/ GroESL construct. Figure 3.9 shows the arrangement of the GroES and GroEL genes in the recombinant plasmid. The *G. thermoglucosidasius* NCIMB 11955 GroESL operon was successfully simulated in the pET21 vector, although the simulation of the 72 base pair sequence interlinking the two genes in *G. thermoglucosidasius* was not successful.

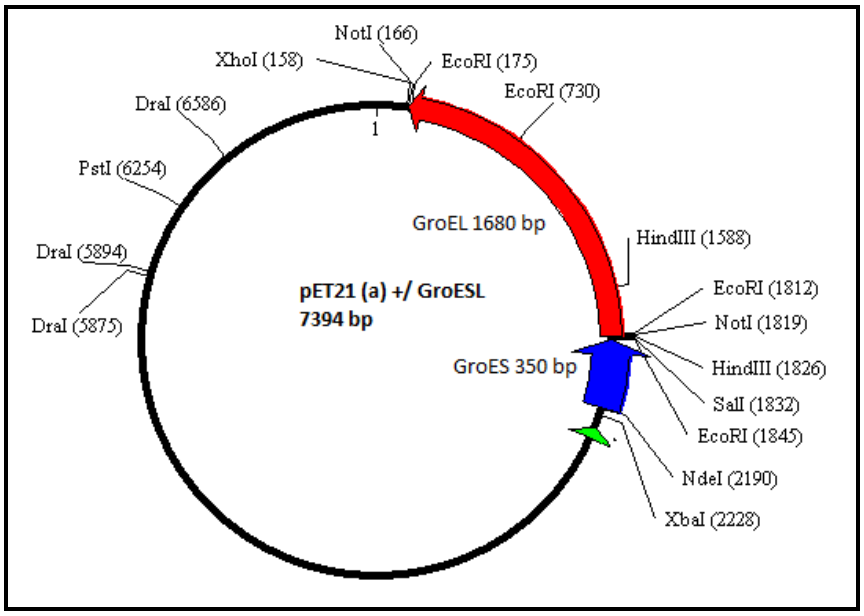


Figure 3.9: Restriction map 0f the pET21 (a) +/ GroESL construct. The green arrow represents the T7 promoter upstream of GroES. The GroES gene is represented by a blue arrow and the GroEL gene is represented by a red arrow.

3.3 Discussion

The GroES and GroEL genes of *G. thermoglucosidasius* NCIMB 11955 were successfully cloned into the pET21 (a) + vector. A number of challenges were faced during the cloning procedures.

3.3.1 PCR amplification of the GroESL operon from genomic DNA

An initial attempt to amplify the entire GroESL operon of *G. thermoglucosidasius* NCIMB 11955 was not successful. The hairpin loop structure situated within the 72 base pair sequence that separates the GroES from the GroEL gene in *G. thermoglucosidasius* NCIMB 11955 (Segal and Ron, 2006) may have affected the attempts at amplification. A similar 9 bp inverted repeat structure in the *G. stearothermophilus* GroESL operon is located at the same position as the hairpin

loop in *G. thermoglucosidasius* NCIMB 11955 (Schuön and Schumann, 1993). Hairpin loops characterized by more than three complementary base pairs greatly affect PCR amplification (Schuön and Schumann, 1993). In some cases, it is possible to adjust PCR parameters so as to amplify across hairpin loops. These adjustments include altering the annealing temperatures and the magnesium chloride concentrations used during amplifications. All the attempts to amplify the complete GroESL operon were unsuccessful.

An alternative strategy using primers designed to independently amplify the GroES and GroEL genes of *G. thermoglucosidasius* NCIMB 11955 was successful. The amplification of both genes was achieved using a 55 ºC annealing temperature. The 72 base pair sequence interlinking the GroES and GroEL genes in *G. thermoglucosidasius* NCIMB 11955 was not amplified, suggesting that the ribosomal binding site of the GroEL gene was not amplified (Figure 3.9).

3.3.2 Cloning and gene analysis of clones

The GroES and GroEL genes of *G. thermoglucosidasius* NCIMB 11955 were first cloned into the pGem- T- easy vector to increase the efficiency of digestion by restriction enzymes, since the efficiency of restriction enzymes is determined by the number of nucleotide base pairs present on either side of the restriction site (Berg *et al.*, 2002).

Restriction map analysis of the genes and plasmids used was performed using NEBcutter V2.0 and appropriate restriction sites were identified. The restriction

enzymes *Nde* 1 and *EcoR* 1 were used to excise the GroES gene from the p-Gem-T- easy/ GroES construct and to clone the gene into the pET21 (a) + vector. This resulted in directional cloning of GroES into the multiple cloning site of the pET21 vector (Figure 3.9). Furthermore the GroES gene was inserted in-frame with the T7 promoter of the vector. Thereafter pET21 (a) +/ GroES construct was digested with the restriction enzyme *Not* 1 and was treated with shrimp alkaline phosphatase to prevent the plasmid from circularization (Sambrook and Russell, 2001). Restriction analysis (Figure 3.6) and PCR amplification using primers W1 and W5 were used to verify the presence of the desired insert in the recombinant clone.

The recombinant clone was named pET21 (a) +/ GroESL (WC construct) and sequenced. Alignment of sequence of the cloned GroESL construct with the GroES and GroEL sequences of the *G. thermoglucosidasius* NCIMB 11955 provided by the TMO Renewables confirmed that the cloned sequences were identical to those from the parent strain (Figure 3.8). The alignment verified that there were no DNA replication errors made during the initial amplification of the genes from genomic DNA even though a non- proof DNA polymerase enzyme (Lab Taq) was used during amplification. The recombinant plasmid map of pET21 (a) +/ GroESL was constructed using DNAMAN and indicated that the GroES and GroEL genes were in-frame with the T7 promoter of the vector (Figure 3.9). The recombinant plasmid was transformed into the expressional host *E. coli* Rosetta pLysS and used to determine whether the cloned GroESL complex was able to enhance the ethanol stress tolerance of the strain.

Chapter 4: Chapter 4: Effect of growth temperature on bacterial growth and ethanol tolerance

4.1 Introduction

4.1.1 Effect of growth temperature on bacterial growth and ethanol tolerance

In order to determine growth conditions which improve the ethanol tolerance of *G. thermoglucosidasius* NCIMB 11955 it was necessary to understand the effects of temperature on the growth of the organism. The effect of growth temperature on *G. thermoglucosidasius* NCIMB 11955 was determined.

An increase in growth temperature typically results in two major changes in bacteria. The fluidity of the cell membrane increases and the expression of heat shock proteins is induced (Banat *et al*, 1998). The increased fluidity of the cell membrane leads to the cell membrane becoming more susceptible to ethanol toxicity (Georgieva *et al.*, 2007). The change in the fluidity of the cell membrane is known as homeoviscous adaptation - a process whereby microorganisms alter the proportion of saturated and unsaturated fatty acids in their lipid membrane (Hazel, 1995).

Analysis of the growth of cultures of *G. thermoglucosidasius* in media containing ethanol at varying concentrations and at various growth temperatures will provide insight into the effects of growth temperature on the ethanol tolerance of the organisms and will determine the optimum conditions for improved ethanol production. These observations are necessary when developing strains for the industrial production of ethanol.

4.1.2 *In vivo* characterisation of the effect of GroES and GroEL over-expression

Heat shock proteins play a major role in protein folding and in the stabilisation of proteins against denaturation. Denaturation of proteins is caused by a number of factors including temperature, pH and solvent stresses (Geogieva *et al*., 2007). The expression of heat shock proteins is a mechanism of stress tolerance in organisms (Laksanalamai, 2003; Silveira *et al*., 2004). Generally, the higher the intensity of the stress, the more heat shock protein expression is required (Silveira *et al.*, 2004). Therefore for bacteria to tolerate high ethanol concentrations, expression of ethanol tolerance specific heat shock proteins is required in order to refold denatured proteins and to prevent the formation of protein aggregates. Bacteria control the levels of heat shock proteins they produce; they do however have a limited turnover capacity which is determined by their protein production mechanisms (Silveira *et al.*, 2004). Stress (in this case an increase in the ambient ethanol concentration) stimulates the induction of heat shock production efficiency to a certain maximum rate (Georgieva *et al.*, 2007). Once the maximum heat shock protein production rate is achieved, the cell cannot further increase heat shock protein production, resulting in insufficient heat shock proteins for reviving the cellular enzymes. The result is cell death (Georgieva *et al.*, 2007).

The need for thermophilic bacteria which tolerate higher ambient ethanol concentrations has encouraged researchers to investigate ways in which heat shock protein expression can be employed in improving the ethanol tolerance of

industrially important microorganisms. Based on observations cited in previous studies (Joon Cha and Je Yoo, 1995; Banat *et al*, 1998) over-expression of heat shock proteins improves ethanol tolerance. There are two ways in which over-expression of heat shock proteins in bacteria can be achieved; by integration of heat shock coding genes into plasmid DNA or by integration of these genes into the bacterial chromosome (Joon Cha and Je Yoo, 1995).

When over-expressing genes in bacteria, cloning of the genes into vectors such as the pET21 (a) + vector is favoured over the integration of the genes into the bacterial chromosome (Joon Cha and Je Yoo, 1995). Over-expression of plasmid borne genes results in higher protein turnover than expression of chromosomally integrated genes. However plasmid borne genes are less stable than those integrated into the chromosome (Joon Cha and Je Yoo, 1995).

In this study the GroES and GroEL genes of *G. thermoglucosidasius* NCIMB 11955 were over-expressed in *E. coli* by integrating both genes into the pET21 (a) + vector multiple cloning site.

E. coli is a workhorse for the expression of a variety of compounds due to its well-researched genetic background. Although *E. coli* is efficient at expressing genes from other organisms, it is important to understand *E. coli* codon usage in relation to the composition of the gene being studied. Codon usage, which is greatly affected by the G+C composition of a gene, is the difference in frequency of synonymous codons in coding DNA (Rocha, 2004). The codon usage of *E. coli*

has been analysed in relation to the expression of a number of thermophilic genes. Thermophilic heat shock genes are expressed in *E. coli*; the GroES and GroEL genes from *Thermoanaerobacter brockii* have been expressed and their gene products purified from *E. coli* (Truscott *et al.,* 1994). The alcohol dehydrogenase gene from *G. thermoglucosidasius* M10EXG was expressed in *E. coli* without experiencing codon usage problems (Jeon *et al.*, 2008). This hinted that codon usage would not be problematic when expressing the GroES and GroEL genes from *G. thermoglucosidasius* NCIMB 11955 in *E. coli* Rosetta pLysS.

4.2 Results

4.2.1 Effect of temperature on the growth of *G. thermoglucosidasius* NCIMB 11955

G. thermoglucosidasius NCIMB 11955 was independently cultured at 45°C and 55°C in 2x TY medium (Section 2.5). The experiment was done in triplicate. The optical density at 600 nm (OD_{600}) of each culture was plotted against time (Figure 4.1). The lag phase of both cultures was similar, although the culture growing at 45 °C had an hour longer lag time. The cultures growing at 55 °C and 45 °C entered the stationary phase of growth after 11 and 12 hours respectively and attained maximum OD_{600}'s of 1.9 and 2.3 respectively (Figure 4.1).

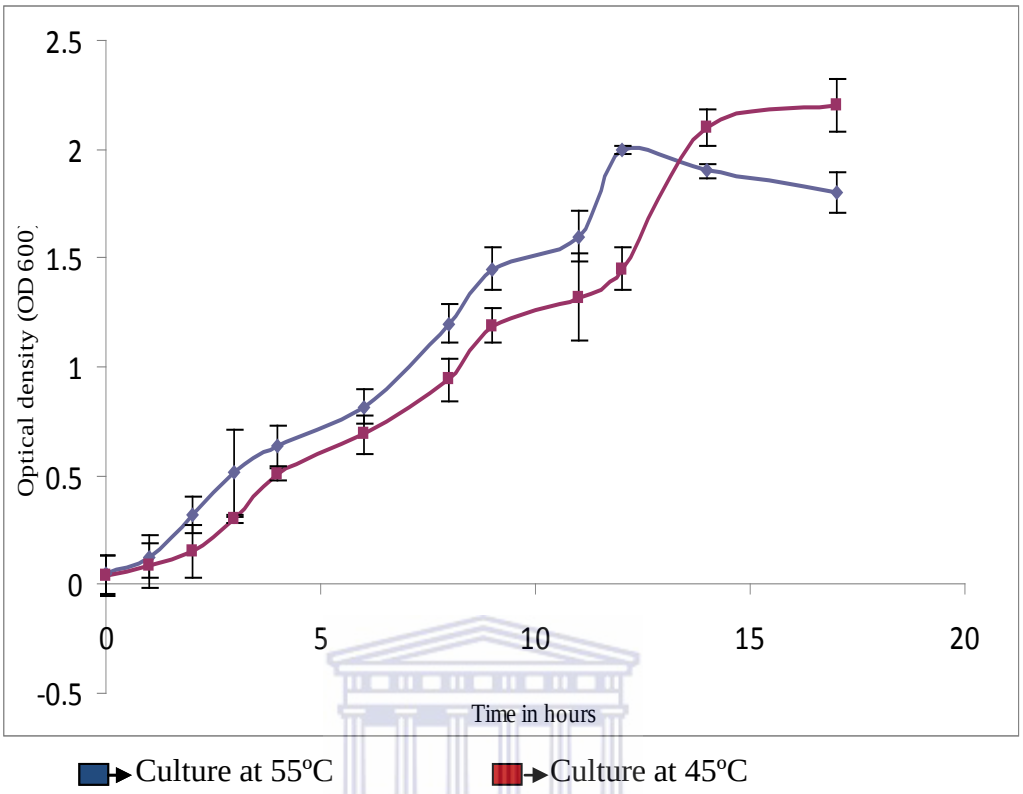


Figure 4.1: The effect of temperature on the growth of *G. thermoglucosidasius* NCIMB 11955 in 2x TY medium. Experiments were performed in triplicate. Error bars indicate the measurement error at each point.

4.2.2 Effect of temperature on the ethanol tolerance of *G.thermoglucosidasius* NCIMB 11955

G. thermoglucosidasius strain NCIMB 11955 was cultured at temperatures of 45°C and 55°C in 2x TY medium supplemented with varying concentrations of ethanol as described in section 2.5.1. The optical density of cultures after 16 hours incubation was determined (Figure 4.2). Cultures with final optical density readings ≥ 0.3 were regarded as viable cultures.

G. thermoglucosidasius NCIMB 11955 cultures grown in the presence of ethanol consistently had a higher cell density when growing at 45 ºC than when growing at 55 ºC (Figure 4.2). At 55 ºC, the growth of cultures of *G. thermoglucosidasius* NCIMB 11955 was inhibited by ethanol concentrations above 4 % v/v. At 45 ºC, cultures of *G. thermoglucosidasius* NCIMB 11955 grew in a maximum ethanol concentration of 6 % v/v (Figure 4.2). The largest difference in final cell densities between the cultures grown at 45 ˚C and 55 ˚C was observed at 4 % v/v ethanol.

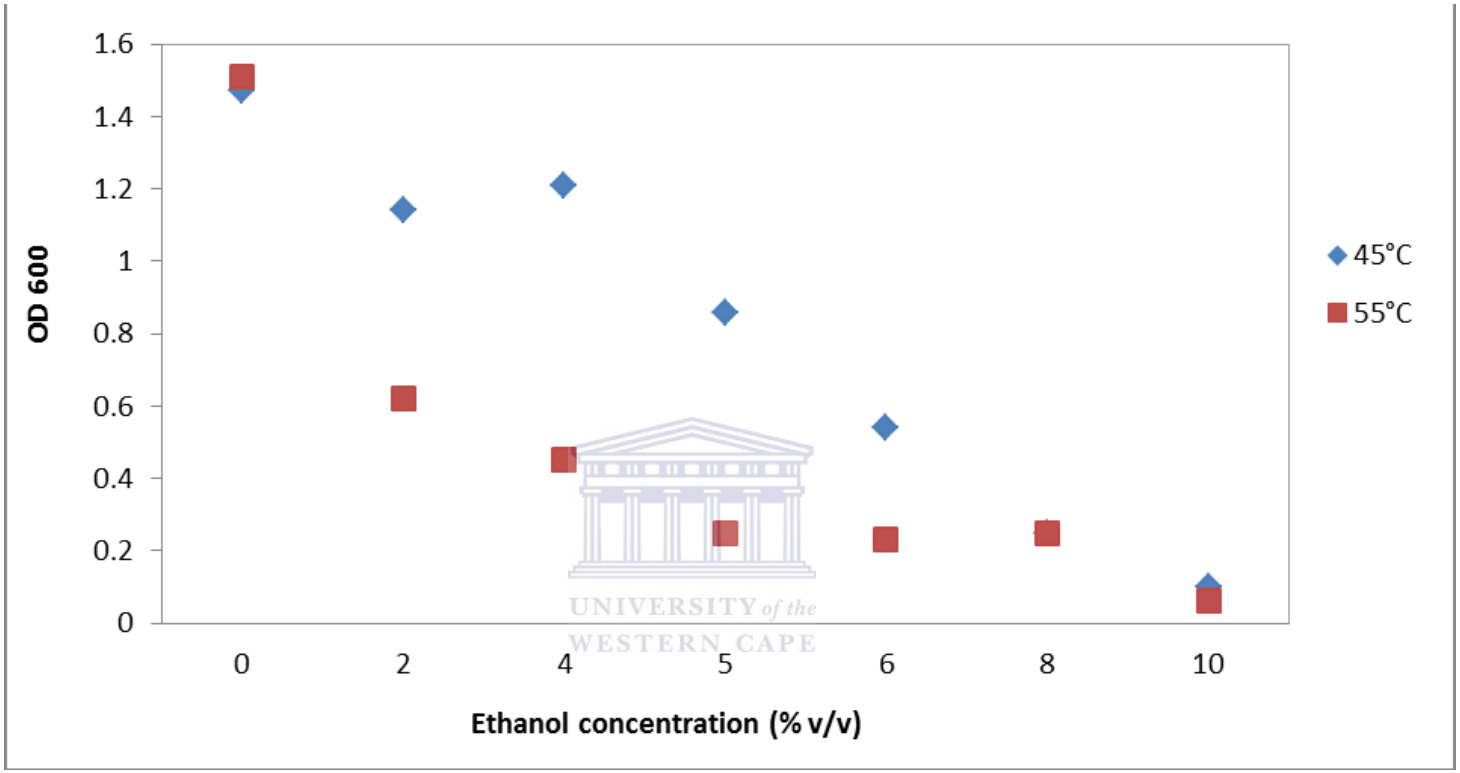


Figure 4.2: The effect of growth temperature on the ethanol tolerance of *G. thermoglucosidasius* NCIMB 11955. Cultures were grown for 16 hours on 2 x TY medium supplemented with ethanol before the final optical density (OD_{600}) was recorded. Cultures with a final OD_{600} of ≥ 0.3 were regarded as viable.

4.2.3 *In vivo* characterisation of GroES and GroEL over-expression

Expression of a 10kDa protein coded for by the recombinant plasmid pET21 (a) +/ GroESL was observed in *E. coli* Rosetta pLySs cells induced with 0.7 mM IPTG (Figure 4.3). The 60 kDa GroEL heat shock protein was not over expressed (Figure 4.3). The 10 kDa protein band was excised and subjected to tryptic digestion and mass spectrometry. It was positively identified as the GroES heat shock protein and had a protein score of 106 when compared to the GroES heat shock protein from *G. stearothermophilus.*

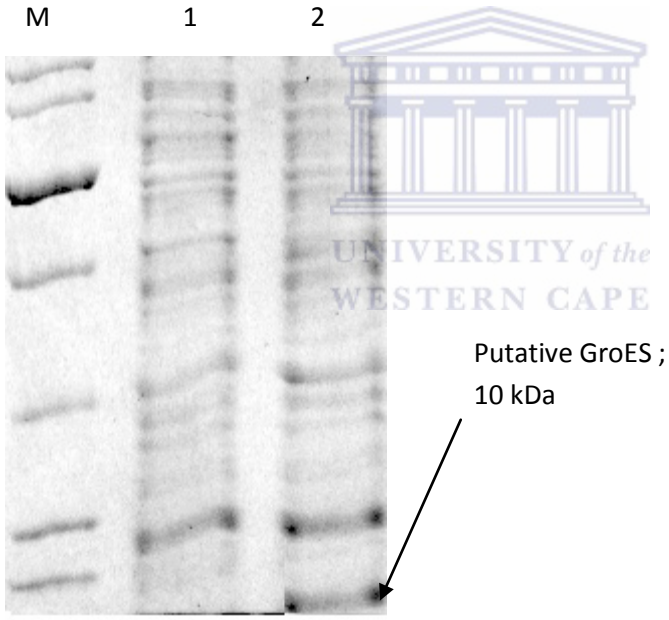
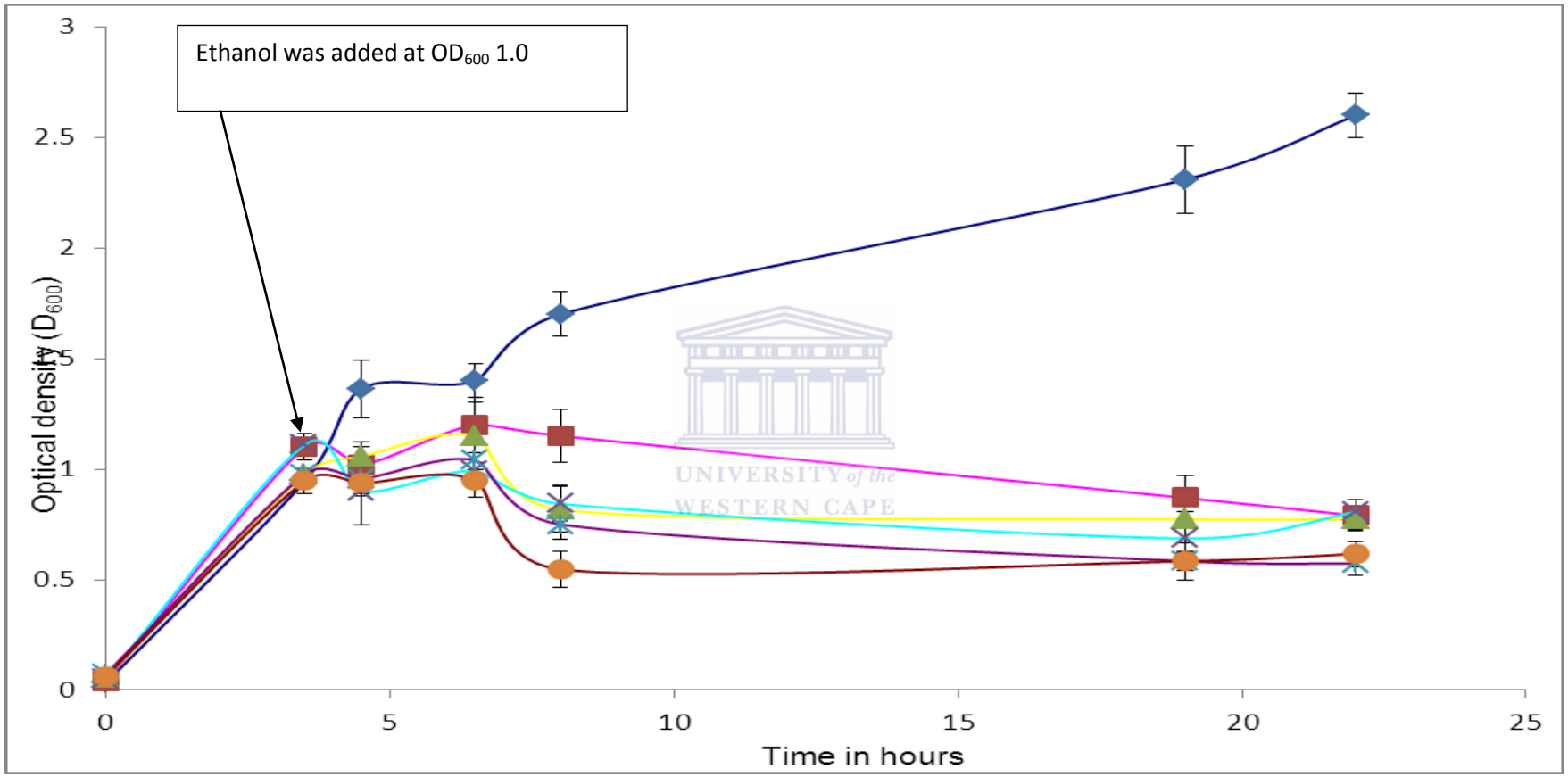


Figure 4.3: Polyacrylamide gel electrophoresis of the intracellular proteome of *E. coli* Rosetta pLySs and *E. coli* Rosetta pLySs harbouring the pET21 (a) +/ GroESL construct. Lane 1: Protein marker (Fermentas) Lane 2: intracellular proteome of wild-type *E. coli* Rosetta pLySs. Lane 3: intracellular proteome of *E. coli* Rosetta pLySs harbouring the pET21 (a) +/ GroESL recombinant plasmid.

To characterize the ethanol tolerance capacity of the recombinant strain, the viability was compared to that of the wild type strain when challenged by increasing concentrations of ethanol. Optical density measurements were taken to follow the growth of the *E. coli* Rosetta pLySs strain and the *E. coli* Rosetta pLySs harboring the pET21 (a) +/ GroESL construct and their survival following an ethanol challenge. Survival curves of the cultures are shown in figures 4.4, 4.5 and 4.6. Figure 4.4 follows cultures of *E. coli* Rosetta pLySs which were grown to an OD_{600} of 1.0 before being challenged by concentrations of 0, 4, 6, 8, 9 and 10 % (v/v). Cultures exposed to 4% and higher ethanol concentrations were inhibited while the cell density of the non-stressed culture increased over the entire 22 hour observation period. In figure 4.5 the growth/survival of the *E. coli* Rosetta pLySs culture transformed with pET21 (a) +/ GroESL recombinant plasmid is shown. The cultures were exposed to concentrations of 0, 4, 6, 8, 10 % ethanol (v/v) when the OD_{600} of the starter culture reached 1.00. The optical densities of cultures of *E. coli* Rosetta pLysSs and *E. coli* Rosetta pLySs pET21 (a) +/ GroESL growing in medium containing no added ethanol mirrored each other (the standard error bars overlapped) (Figure 4.6). The OD_{600} of cultures harbouring the recombinant plasmid exposed to 4% ethanol continued to increase for a period of an hour following the addition of ethanol (an OD_{600} increase of 0.4 was noted) before plateauing while that of the culture of *E.coli* Rosetta pLysSs plateaued on addition of the ethanol stress (Figures 4.5 and 4.6). This suggests that the differences in the optical densities of the cultures exposed to 4% ethanol were due to the improved ethanol tolerance of the recombinant strain.



Key: 0% ethanol 4% ethanol 6% ethanol 8% ethanol 9% ethanol 10% ethanol

Figure 4.4: Growth of unchallenged and ethanol challenged *E. coli* Rosetta pLySs in 2 x TY. Ethanol (4, 6, 8, 10% v/v) was added when cultures reached an OD_{600} of 1.0. The survival of the cultures was monitored following the addition of ethanol.

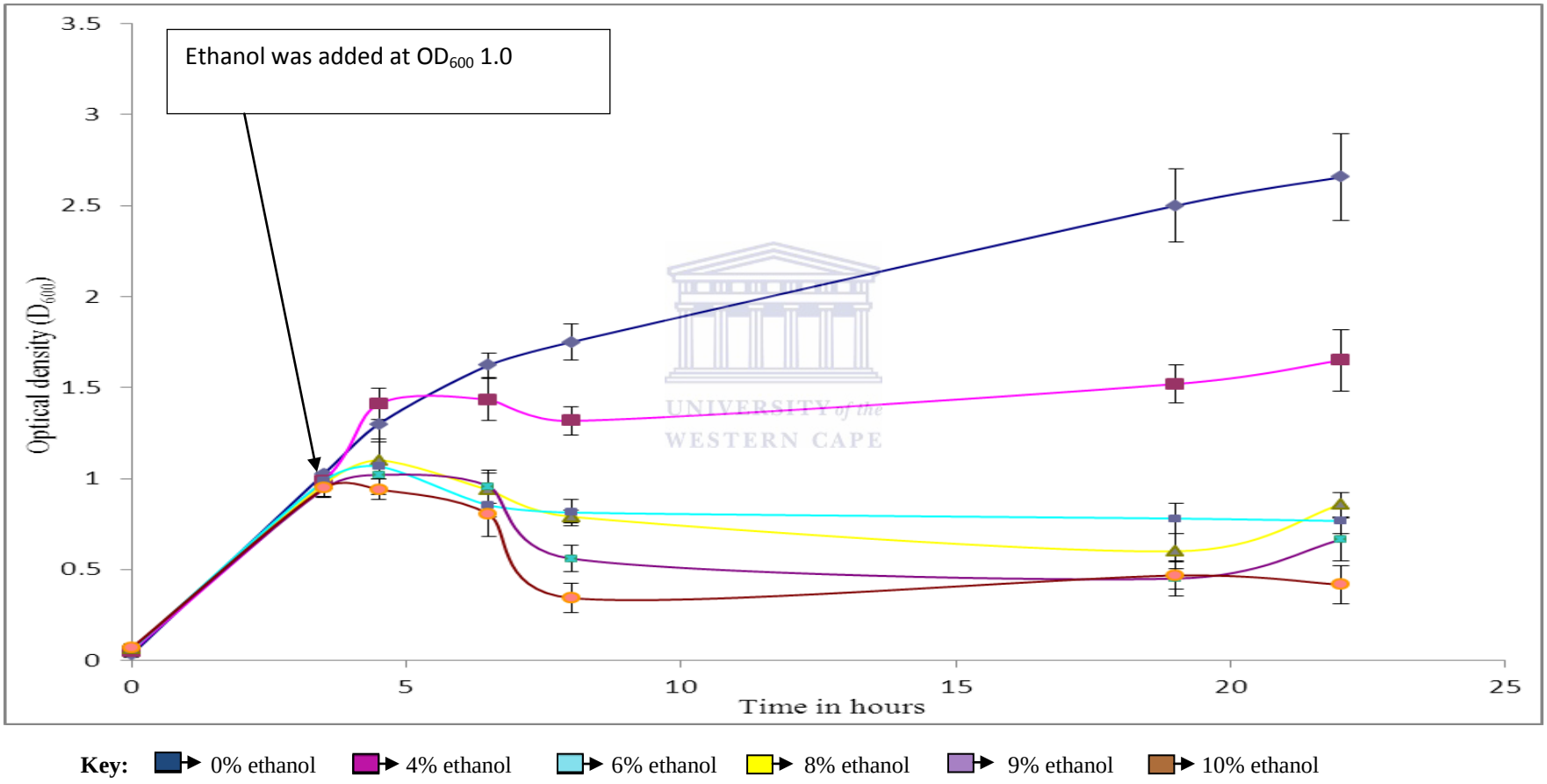


Figure 4.5: Growth of unchallenged and ethanol challenged *E. coli* Rosetta pLySs/ pET21 (a) +/ GroESL in 2 x TY. Ethanol (4, 6, 8, 10 v/v) was added when cultures reached an OD_{600} of 1.0. The survival of the cultures was monitored following the addition of ethanol.

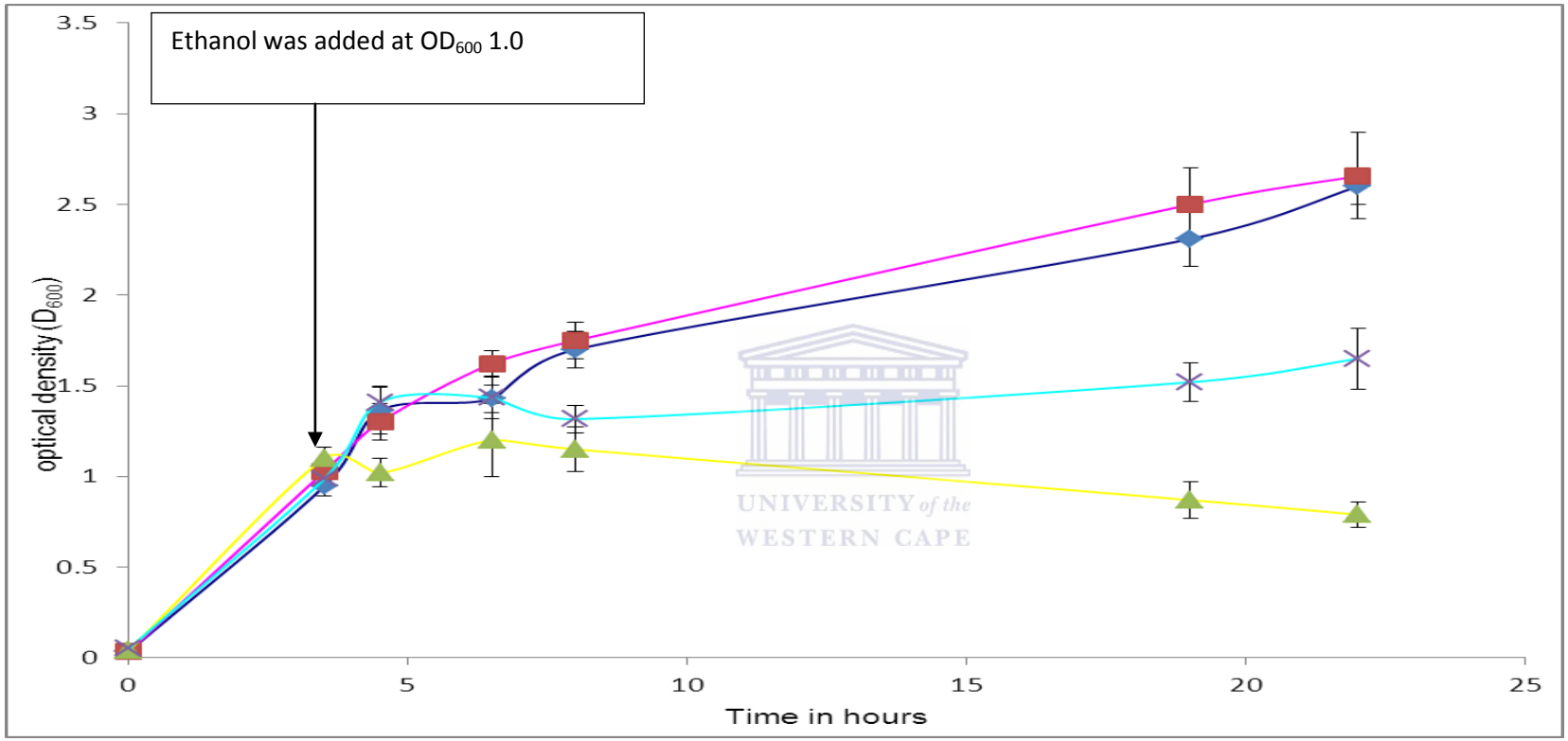


Figure 4.6: Effect of the addition of ethanol (4 % v/v) on the growth of *E. coli* Rosetta pLySs and *E. coli* Rosetta pLySs/ pET21 (a) +/ GroESL. The survival of the cultures was monitored following the addition of ethanol.

Using OD_{600} measurements the growth rates of stressed and unstressed cultures were estimated using the simple growth rate equation: growth rate (μ) = 2.303 ($\log OD_2 - \log OD_1$) / ($t_2 - t_1$) (Zwietering *et al*., 1990) (Table 4.1). t_1 was the time point 3h after ethanol was added to the stressed cultures (i.e. 3h after starter cultures had reached an OD_{600} of 1.0). t_2 was the time point 15h after the addition of ethanol to stressed cultures.

The growth rate of unstressed (no ethanol added) wild type and recombinant cultures was 0.20 and 0.202 respectively (Figure 4.6 and Table 4.1). Negative growth rates were recorded for all cultures (recombinant and wild type) stressed by the addition of ethanol bar for the recombinant culture stressed by the addition of 4% ethanol (Figures 4.6 and Table 4.1). This culture, harbouring the pET21 (a) + / GroESL construct, recorded a growth rate of 0.102 while the host strain experienced cell death (a negative growth rate of -0.014) under the same conditions.

Table 4.1: Growth rates of *E.coli* Rosetta pLysSs *and E.coli* Rosetta pLysSs/pET 21 (a) +/GroESL cultures grown at 37°C in 2 x TY medium with and without an ethanol challenge. Cultures were grown to an OD_{600} of 1.0 prior to the addition of 0, 4, 6, 8, 9 and 10% (v/v) ethanol. Growth rates were calculated over a 12 hour period starting 3h after the addition of ethanol.

Percentage of ethanol added	***E.coli* Rosetta pLysSs**	***E.coli* Rosetta pLysSs/pET 21 (a) +/GroESL**
0	0.200	0.202
4	-0.014	0.102
6	-0.04	-0.10

The final turbidity of ethanol challenged and unchallenged cultures after 18h exposure to ethanol (in the case of unchallenged cultures this was the 18h period after reaching an OD600 reading of 1.0) was determined. A decrease in the final turbidities of cultures with increasing ethanol concentrations was recorded . The final turbidity of cultures of *E. coli* Rosetta pLySs/ pET21 (a) +/ GroESL was higher than that of the *E. coli* Rosetta pLySs cultures exposed to the same conditions. The biggest difference in final culture turbidity was observed when challenged with 4 % ethanol.

4.3 Discussion

4.3.1 Effect of temperature on the growth of *G. thermoglucosidasius* NCIMB 11955

The optimum growth temperature of *Geobacillus thermoglucosidasius* NCIMB 11955 is 55 °C (Cripps *et al*., 2008). A growing bacterial culture typically exhibits four growth phases: a lag phase, an exponential growth phase, a stationary phase and an exponential death phase (Zwietering, 1990). The lag phase in cultures of *G. thermoglucosidasius* NCIMB 11955 in 2x TY medium at 45 °C and 55 °C was 1.5 hours (Figure 4.1). The culture growing at 55 ºC exhibited a higher growth rate than the culture growing at 45 ºC (Figure 4.1). The culture at 45 ºC achieved higher cell density (OD_{600} 2.2) than the culture at 55 ºC (OD_{600} 1.9). These observations do not agree with the Arrhenius law (Ratkowsky *et al.*, 1982) which describes the temperature dependence of the specific reaction rate constants in a chemical reaction. Based on the Arrhenius law, the growth rate of cultures at 45 and 55 ºC was predicted to be marginally different. However because bacterial

growth is a complex reaction involving a variety of enzymes and substrates (Ratkowsky *et al*., 1982) the results observed in this study i.e. the similar growth rates of cultures at 45 and 55 ºC were not nullified as the Arrhenius law does not adequately describe the effect of temperature on the bacterial growth rate (Ratkowsky *et al.*, 1982).

4.3.2 Effect of growth temperature on the ethanol tolerance of *Geobacillus* NCIMB 11955

After culturing *G. thermoglucosidasius* NCIMB11955 at 45 ˚C and 55 ˚C in 2x TY medium containing increasing concentrations of ethanol (0 - 10 % v/v) it was apparent that growth temperature influenced the ethanol tolerance of the cultures. The cell turbidity readings of *G. thermoglucosidasius* NCIMB 11955 cultures after an hour growth in 2x TY medium at 45 ºC and 55 ºC differed by OD 0.3. The cell turbidities of this strain at 45 ºC and 55 ºC in 2% and 4% v/v ethanol differed by 0.64 and 0.76, respectively . A higher cell density in the presence of ethanol implies better ethanol tolerance (Georgieva *et al.*, 2007). These findings suggest that *G. thermoglucosidasius* NCIMB 11955 has a higher tolerance to ethanol when growing at 45 ºC than at 55 ˚C.

Similar observations were made when *B. stearothermophilus* LLD 15 was grown in liquid media containing increasing concentrations of ethanol (Amartey, 1991). In this strain the effect of ethanol on the growth rate of cultures was slight, but the growth rate was consistently greater at 70 °C than at 60 °C (Amartey, 1991). The decrease in the ethanol tolerance of *B. stearothermophilus* LLD 15 at increased

temperatures was linked to high concentrations of cardiolipin protein in the cell membranes of the organism (Amartey, 1991). The cardiolipin content of the cell membranes of *B. stearothermophilus* LLD 15 increased with increased growth temperatures under anaerobic conditions (Mosley *et al.*, 1976).

The observations in figures 4.2 that cultures of *G. thermoglucosidasius* NCIMB 11955 were more tolerant to ethanol when growing at 45 °C than at 55 °C may be linked to the phenomenon of 'bacterial suicide'. The bacterial suicide hypothesis suggests that when subjected to a stress a rapidly growing bacterial culture is more likely to suffer growth arrest than a slow growing culture (Aldsworth *et al.*, 1999). The hypothesis suggests that when cells undergo growth arrest their metabolism does not stop. Due to the growth arrest and the continued metabolism an imbalance between catabolism and anabolism is experienced. This imbalance results in free radical formation (Aldsworth *et al.*, 1999). The activity of the free radicals which accumulate is lethal to the cell i.e. the lethality is not due to the stress (Aldsworth *et al.*, 1999). Accumulation of free radicals reduces the fitness of the cell due to an imbalance between the endogenous superoxide stress and the antioxidant defences of the cell (Strohmeier *et al*., 1998). If the imbalance favours the accumulation of endogenous superoxide, DNA and cellular enzymes are damaged resulting in cell death (Walker, 1996).

Regardless of homeoviscous adaptation and bacterial suicide through stress, changes in growth temperature induce the production of heat shock proteins in bacterial cells (Cooper and Ho, 1983). Heat shock proteins play a major role in the

stress tolerance of bacteria (Silveira *et al.*, 2004). GroEL, Group E small heat shock protein and GroES are examples of group 1 heat shock proteins present in bacteria such as *E. coli* (Luo *et al*., 2009). It is evident that these heat shock proteins are over expressed in stressed cells. To determine whether the GroES and GroEL gene products of *G. thermoglucosidasius* NCIMB 11955 were able to influence the tolerance of *E. coli* Rosetta pLySs to ethanol the genes were cloned and expressed in *E. coli* Rosetta pLySs.

4.3.3 *In-vivo* characterisation of GroES and GroEL over-expression

Analysis of the proteome of *E. coli* Rosetta pLySs harbouring the pET21 (a) + / GroESL construct to that of the wild type strain was used to indicate whether GroES and GroEL were over-expressed (Figure 4.3). No obvious protein band corresponding to the GroEL gene product was observed on a SDS PAGE gel while expression of a 10Kda band corresponding in size to the GroES gene product was observed (Figure 4.3). Analysis of the pET21 (a) +/ GroESL insert sequence indicated that the construct lacked a Shine-Dalgarno sequence upstream of GroEL. A Shine-Dalgarno sequence is a ribosomal binding site located eight base pairs upstream of the start codon for a gene (Staples and Hindley, 1971; Shine and Dalgarno, 1975). This binding site facilitates a complementary base pairing between DNA and RNA polymerase and signals the start of a new gene (Stamples and Hindley, 1971). Without a Shine-Dalgarno sequence, RNA polymerase is unable to recognise the GroEL start codon downstream of the GroES stop codon and the GroEL gene would not be transcribed.

Optical density was used to monitor culture growth. Bacterial cells in a culture scatter light, reducing the amount of light that reaches the photoelectric cell, and the reduction in light is measured and recorded as optical density (Widdel, 2010). Spectrophotometry does not differentiate dead cells from growing cells, a factor that need to be addressed when using optical density to monitor growth (Lin *et al*., 2010). To overcome errors associated with the conversion of optical density to cell density, a standard curve of optical density versus colony counts should be created (Lin *et al*., 2010). From the standard curve, a proportionality factor can be calculated (Lin *et al*., 2010). The proportionality factor is then used to calculate the cell number of a culture at the time when optical density was measured. Due to time constraints colony counts were omitted in this study. In this study the growth rate over a particular period during culture growth was calculated from the optical density using the equation growth rate (μ) = $2.303(\log OD_2 - \log OD_1) / (t_2 - t_1)$ (Table 4.1).

Growth rates were used to compare ethanol tolerance of cultures of *E. coli* Rosetta pLySs to that of the strain harbouring the pET21 (a) +/ GroESL construct. A higher growth rate in the presence of ethanol would suggest better ethanol tolerance of that strain. Ethanol-stressed *E. coli* Rosetta pLySs/ pET21 (a) +/ GroESL cultures had a greater tolerance to the presence of 4% ethanol than the corresponding *E. coli* Rosetta pLySs cultures (Figure 4.6). This may be due to the presence of increased levels of the GroES heat shock protein in the cultures. Based on these findings, it is suggested that over-expression of the GroES heat

shock protein in cultures prior to the addition of ethanol improves the ethanol tolerance of the strain (Dugas, 1996; Chen *et al.*, 2006; Georgieva *et al.*, 2007).

Increasing the cyclic rate of the protein folding mechanism of the GroESL complex by increasing the amount of the GroES protein present in a cell increases the ability of bacteria to tolerate higher ethanol concentrations (Dugas, 1996; Chen *et al.*, 2006; Georgieva *et al.*, 2007). The GroESL mechanism of protein folding (section 1.4) in which the GroES protein binds to the GroEL protein improves the efficiency of protein folding (Chen *et al.*, 2006). ATP and substrate proteins bind to the hydrophobic cavity of GroEL, inducing a conformational change that attracts binding of GroES to the other end of the GroEL cavity. Binding of the GroES protein together with ATP hydrolysis induces a conformational change in GroEL, pushing the substrate protein into the hydrophilic cavity of GroEL (Chen *et al.*, 2006). In the hydrophilic cavity, unfolded proteins bury their hydrophobic groups, thereby refolding. Binding of unfolded protein to the hydrophobic cavity of GroEL sends an allosteric signal to the hydrophilic cavity. This results in the release of GroES and folded protein, and the cycle begins again (Dugas, 1996; Chen *et al.*, 2006).

The formation of the GroESL/substrate complex is the rate determining step of the protein folding reaction (Gupta *et al.*, 2006). Based on the co-enzyme principle, having the GroES population dominating the GroEL population suggests that there are always GroES molecules available to activate GroEL for protein

refolding. Therefore, the rate of protein folding/refolding is completed more quickly. Although expression of the GroEL protein from the pET21 (a) +/ GroESL construct was not detected, the presence of overexpressed GroES protein in *E. coli* Rosetta pLySs/ pET21 (a) +/ GroESL increased the ability of the bacteria to tolerate the presence of 4% ethanol in the growth medium.

An extensive amount of work has been done on the co-over expression of GroES in *E. coli* together with other foreign proteins (Gupta *et al.,* 2006). The observations in these studies are similar to observations made in this study: over expression of GroES improves stress tolerance. To compare the effect of expression of the GroES gene on tobacco mosaic virus coat protein (TMV CP) stability and solubility, a recombinant plasmid harbouring the gene coding for TMV CP and the GroES gene was expressed in *E. coli* while a plasmid carrying only the TMV CP gene was simultaneously expressed in *E. coli* (Hwang *et al.*, 1998). A 3 to 5 fold increase in the amount of soluble TMV CP at 30°C was observed in the strain harbouring the GroES gene.

In this study the *G. thermoglucosidasius* NCIMB11955 GroESL operon was successfully cloned into the pET 21 (a) + expression vector but the 72bp sequence interlinking the genes in the *G. thermoglucosidasius* genome was not cloned. Expression of a 10kDa protein coded for by the recombinant plasmid pET21 (a) + / GroESL was observed in *E. coli* Rosetta pLySs cells induced with IPTG (Figure 4.3). This protein was identified as the GroES heat shock protein (section 4.1.3).

The 60kDa GroEL heat shock protein was not overexpressed, possibly due to the absence of structures within the interlinking sequence which would allow for the binding of RNA polymerase. To fully characterise the effect of overexpression of the GroES and GroEL genes on the ethanol tolerance of strains it is essential to develop constructs containing the full complement of regulatory sequences necessary for the expression of the genes. Overexpression of the GroES gene together with the GroEL gene would potentially influence protein stability and solubility in host cells and result in an improved tolerance to stress.

Chapter 5: Conclusion and future perspectives

5.1 Conclusion

The objectives of this project were partially achieved. The first objective was to determine whether growth temperature affects the tolerance to ethanol of *G. thermoglucosidasius* NCIMB 11955. Changes in the growth temperature clearly affected the tolerance of the strain to ethanol: when the growth temperature was lowered to 45ºC (the optimum growth temperature of the strain is 55ºC cultures tolerated the presence of higher ethanol concentrations.

The second objective was to clone the *G. thermoglucosidasius* GroES and GroEL genes into the pET21 (a) + expression vector for expression in *E. coli* Rosetta pLySs. The GroES and GroEL genes were successfully cloned into the multiple cloning site of the expression vector. This was confirmed by restriction digestion, DNA sequencing and PCR amplification.

The final objective was to express the pET21/GroESL construct in *E. coli* Rosetta pLySs and to assess the effect of expression of the cloned genes on the growth, stability and ethanol tolerance of the host strain. Transformation of the recombinant plasmid into *E. coli* was successful and the resultant clone was named *E. coli* Rosetta pLySs WC. Proteomic analysis of *E. coli* Rosetta pLySs WC proteome on one dimensional SDS-PAGE gels showed that GroES was over-expressed while GroEL was not. GroEL was probably not over-expressed because the gene lacked a Shine-Dalgarno sequence upstream of its start codon. However, over-expression of GroES alone improved the ethanol tolerance of the *E. coli*

host. The difference in ethanol tolerance of cells of the host and recombinant strains was observed in media containing 4% v/v ethanol: the recombinant strain showed growth in 4% v/v ethanol while growth of the host strain was inhibited.

5.2 Future perspectives

It is critical for this study that a Shine-Dalgarno sequence be inserted upstream of the GroEL gene so that both genes are over-expressed. Over-expression of GroEL is expected to result in a further increase in the ethanol tolerance of the strain. The GroEL protein is involved in protein folding. Chromosomal integration of the GroESL operon into the *G. thermoglucosidasius* chromosome is an alternative strategy to overexpress the heat shock proteins. Chromosomal integration of the GroESL operon into the *G. thermoglucosidasius* genome would increase the copies of the operon in the chromosome and may result in overexpression of the GroESL genes.

Further characterisation of the *G. thermoglucosidasius* ethanol response proteins would result in a better understanding of ethanol tolerance-related problems in this species. Once a database of ethanol stress response proteins is available, a recombinant clone that caters for all heat shock proteins - which are up-regulated as ethanol concentration increases – could be designed.

References

- Aldsworth, T. G., Sharman, R. L. and Dodd, C. E. R. (1999) Bacterial suicide through stress. *Cell Molecular Biology and Life Science* 56: 378-383
- Amartey, S. A., Leak. D. J. and Hartley, B. S. (1991) Effects of temperature and medium composition on the ethanol tolerance of *Bacillus stearothermophilus* LLD 15. *Biotechnology Letters* 13: 627-632
- Amartey, S. A., Leung, P. C. J., Baghaei-Yazdi, N., Leak, D. J. and Hartley, B. S. (1998) Fermentation of a wheat straw acid hydrolysate by *Bacillus stearothermophilus* T-13 in continuous culture with partial cell recycle. *Process Biochemistry* 34: 289-294
- Antoni, D., Zverlov, V. V. And Schwarz, W.H.. (2007) Biofuels from microbes. *Applied Microbiology and Biotechnology* 77: 23-35
- Banat, I., Nigam, P., Singh, D., Marchant, R. and Mc Hale, A. (1998) Review: ethanol production at elevated temperatures and alcohol concentrations: Part I – Yeasts in general. *World Journal of Microbiology and Biotechnology* 14: 809-12
- Berg, J.M., Tymoczko, J.L. and Stryer, L. (2002) Biochemistry. 5th edition. New York: W H Freeman
- Betz, C., Schlenstedt, G. and Bailer, S. M. (2004) Asr1p, a novel yeast ring/PHD finger protein, signals alcohol stress to the nucleus. *Journal of Biological Chemistry* 279: 28174-28181

- Blumer-Schuette, S. E., Kataeva, I., Westpheling, J., Adams, M. W. and Kelly, R. M. (2008). Extremely thermophilic micro-organisms for biomass conversion. *Current Opinion in Biotechnology*. 19: 210-217
- Bradford, M.M. (1976) Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* 72: 248–254
- Brat, D., Boles, E. and Wiedemann, B. (2009) Functional expression of a bacterial xylose isomerase in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology* 75: 2304-2311
- Bruggemann, H. and Gottschalk, G. (editors). (2009) *Clostridia: molecular biology in the post-genomic era*. Caister Academic Press, U.K.
- Burdon, H. R. (1986) Heat shock and the heat shock proteins. *Biochemistry Journal* 240: 313-324.
- Butzen, S. and Haefele, D. (2008) Dry-grind ethanol production from corn. *Enzyme Research* 18: 1-10
- Campbell, C.J. and Laherrère, J. H. (1995) The end of cheap oil. *Scientific American* March: 80-86
- Cavallo, A. J. (2004) Hubbert's petroleum production model: an evaluation and implications for world oil production forecasts. *Natural Resources Research* 13: 1-16
- Chandel, A. K., Chan, E. S., Rudravaram, R., Lakshmi, M., Venkateswar, N., Rao, L. and Ravindra, P. (2007) Economics and environmental impact of bioethanology production technologies. *Biotechnology and Molecular Biology Reviews* 2: 14-32

- Charewa, W. (2008) Honours mini thesis: University of the Western Cape
- Chen, H., Chu, Z., Yang, S. and Y. (2006) Over-expression and characterization of the recombinant small heat shock protein from *Pyrococcus furiosus*. *Biotechnology Letters* 28: 1089-1094
- Chirwa, E. W. (2007) Sources of technical efficiency among smallholder maize farmers in Malawi. *African Economic Research Consortium,* Research paper 172
- Cooper, P. and Ho, T. D. (1983) Heat shock proteins in maize. *Plant Physiology*. 71: 215-222
- Cripps, R. E., Eley, K., Leak, D. J., Rudd, B., Taylor, M., Todd, M., Martin, S., Boakes, S. and Atkinson, T. (2009) Metabolic engineering of *Geobacillus thermoglucosidasius* for high yield ethanol production. *Metabolic Engineering* 11: 398-408
- Demain, A. L. (2009) Bio-solutions to the energy problem. *Journal of Microbiology and Biotechnology* 10: 521-528
- Demirbas, A. (2005) Biodiesel production from vegetable oils via catalytic and non-catalytic supercritical methanol transesterification methods. *Journal for Progress in Energy and Combustion Science* 31: 486-487.
- Ding, J., Huang, X., Zhang, L., Zhao, N., Yang, D. and Zhang, K. (2009) Tolerance and stress response to ethanol in *Saccharomyces cerevisiae*. *Applied Microbiology and Biotechnology* 85: 253-263.
- Dubaquié, Y., Looser, R. and Rospert, S. (1997) Significance of chaperonin 10-mediated inhibition of ATP hydrolysis by chaperonin 60. *Proceeding of the National Academy of Sciences* 90: 9011-9016

- Dugas, H. (1996) A chemical approach to enzyme action. *Biochemistry and Molecular Biology Education* 25: 173-188
- Ezeji, T., Qureshi, N. and Blaschek, H. P. (2007) Butanol production from agricultural residues: impact of degradation products on *Clostridium beijerinckii* growth and butanol fermentation. *Biotechnology and Bioengineering*. 97: 1460-1469
- Fenton, W. A. and Horwich, A. L. (2003) Chaperonin-mediated protein folding: fate of substrate polypeptide. *Quarterly Reviews of Biophysics* 36: 229–56
- Galbe, M. and Zacchi, G. (2002) A review of the production of ethanol from softwood. *Applied Microbiology and Biotechnology* 59: 618-628
- Gardener, M. (2004). Is "hydrogen highway" the answer? http://www.utsandiego.com/uniontrib/20041122/news_1n22hydrogen.html
- Georgieva, T. I., Mikkelsen. M. J. And Ahring, B. K. (2007) High ethanol tolerance of the thermophilic anaerobic ethanol producer *Thermoanaerobacter* BG1L1. *Central European Journal of Biology* 2: 364-377
- Graham, R. L. J., Pollock, E. C., Ternan, N. G. and McMullan, G. (2006) Top down proteomic analysis of the soluble sub-proteome of the obligate thermophile *Geobacillus thermoleovorans* T80: insights into its cellular process. *Journal of Proteome Research* 5: 822-828
- Green, N. (2004) Growing energy: how biofuels can help to end America's oil dependence, Washington D. C: *Natural Resources Defense Council Report*: 15-30

- Green, N. And Mugica, Y. (2005) Bringing biofuels to the pump: an aggressive plan for ending America's oil dependence, Washington D. C.: *Natural Resources Defence Council Report* 40-45
- Gunasekaran, P. and Chandra Raj, K. (1999) Ethanol fermentation technology- *Zymomonas mobilis*. *Current Science (India)* 77: 56-68
- Gupta, R. S. (1995) Evolution of the chaperonin families (Hsp60, Hsp10 and Tcp-1) of proteins and the origin of eukaryotic cells. *Journal of Molecular Microbiology* 15: 1–11
- Gupta, P., Aggarwal, N., Batra, P., Mishra, S. and Chaudhuri, T.K. (2006) Co-expression of chaperonin GroEL/GroES enhances in vivo folding of yeast mitochondrial aconitase and alters the growth characteristics of *Escherichia coli*. *International Journal of Biochemistry and Cell Biology* 38:1975-1985
- Hazel, J. R. (1995) Thermal adaptation in biological membranes: is homeoviscous adaptation the explanation? *Annual Review of Physiology* 57: 19-42
- Hecker, M., Schumann, W. and Völker, U. (1996) Heat shock and general stress response in *Bacillus subtilis*. *Molecular Microbiology* 19: 417-428
- Hirakawa, H., Kamiya, N., Kawarabayashi, Y. and Nagamune, T. (2005) Log P effect of organic solvents on a thermophilic alcohol dehydrogenase. *Biochemica et Biophysica Acta* 1748: 94 - 99
- Hubbert, M. K. (1956) Nuclear energy and the fossil fuels. *Drilling and production practice* (American Petroleum Institute publication 95) http://www.hubbertpeak.com/hubbert/1956/1956.pdf

- Hwang, D. J., Turner, N. E. and Wilson, T. M. (1998) Chaperone protein GrpE and the GroEL GroES complex promote the correct folding of tobacco mosaic virus coat protein for ribonucleocapsid assembly *in vivo*. *Archives of Virology* 143:2203-2214

- Ingram, L, O., Conway, T., Clark, D, P., Sewell, G, W. and Preston, J, F. (1987). Genetic engineering of ethanol production in *E.coli*. *Applied and Environmental Microbiology* 53: 2420-2425

- Jeon, Y. J., Young, J. C. N., Riyanti, E. I., Neilan, B. A., Rogers, P. L. and Svenson, C. J. (2008) Heterologous expression of alcohol dehydrogenase gene from *Geobacillus thermoglucosidasius* M10EXG. *Journal of Biotechnology* 135:127-133

- Jones, E. M. (1950) Chamber process manufacturing of sulphuric acid. *Industrial and Engineering Chemistry*. 42: 2208-2210

- Kamimura, A. and Sauer, L. I. (2008) The effect of flex fuel in the Brazilian light road transportation. *Energy Policy* 36: 1574-1576.

- Kim, R., Kim, K. K., Yokota, H. and Kim, S. (1998). Small heat shock protein of *Methanococcus jannaschii*, a hyperthermophile. *Proceedings of the National Academy of Sciences* 95: 9129-9133

- Laksanalamai, P., Jiemjit, A., Bu, Z., Maeder, D. L. and Robb, F. T. (2002) Multi-subunit assembly of the *Pyrococcus furiosus* small heat shock protein is essential for cellular protection at high temperature. *Extremophile* 7: 79-83

- Laksanalamai, P. and Robb, F. T. (2003) Small heat shock proteins from extremophiles. *Extremophiles* 8: 1-11
- Li, M. and Wong, S. L. (1992) Cloning and characterisation of the GroESL operon from *Bacillus subtilis*. *Journal of Bacteriology* 174: 3981-3992
- Li, Y., Tian, C., Tian, H., Zhang, J., He, X., Ping, W. and Lei, H. (2012) Improvement of bacterial cellulose production by manipulating the metabolic pathways in which ethanol and sodium citrate is involved. *Applied Microbiology and Biotechnology* 96: 1479 – 1487
- Lin, H.L., Lin, C.C., Lin, Y.J., Lin, H.C. and Shih, C.M. (2010) Revisiting with a relative-density calibration approach the determination of growth rates of microorganisms by use of optical density data from liquid cultures. *Applied and Environmental Microbiology* 76: 1683-1685
- Lovitt, R. W., Longin, R. and Zeikus, J. G. (1984) Ethanol production by thermophilic bacteria: physiological comparison of solvent effects on parent and solvent-tolerant strains of *Clostridium thermohydrosulfuricum*. *Applied and Environmental Microbiology* 48: 171-177
- Luo, H., Laksanalamai, P. and Robb, F. T. (2009) An exceptionally stable Group ll chaperonin from the hyperthermophile *Pyrococcus furiosus*. *Archives of Biochemistry and Biophysics*: 486: 12-18
- MacPherson, S., Larochelle, M. and Turcotte, B. (2006) A fungal family of transcriptional regulators: the zinc cluster proteins. *Microbiology and Molecular Biology*. 70: 583-604

- Maugeri, T. L., Gugliandolo, C., Caccamo D. And Stackbrandt, E. (2002) A polyhasic taxonomic study of thermophilic bacilli from shallow marine vents. *Journal of Applied Microbiology* 24: 572-587
- Meinander, N., Zacchi, G. and Hahn-Hägerdal, B. (1996). A heterologous reductase affects the redox balance of recombinant *Saccharomyces cerevisiae*. *Journal of Microbiology* 142. 165-172.
- Michel, G. P. F., Azoulay, T. and Starka, J. (1985) Ethanol effect on the membrane protein pattern of *Z. mobilis*. *Annales de l'Institut Pasteur Microbiology* 136: 173-179
- Mishima, D., Kunuki, M., Sei, K., Soda, S., Ike, M. and Fujita, M. (2008) Ethanol production from candidate crops. *Bioresource Technology* 99: 2495-2500
- Monod, J. (1949) The growth of bacterial cultures. *Annual Reviews of Microbiology* 3: 371-394
- Moriello, S. V., Lama, L., Poli, A., Gugliandolo, C., Maugeri, T. L., Gambcorta, A. and Nicolaus, B. (2003) Production of exopolysaccharides from a thermophilic microorganism isolated from a marine hot spring in flegrean areas. *Journal of Industrial Microbiology and Biotechnology* 30: 95-101
- Mosley, G.A., Card, G. L. and Koostra, W. L. (1976) Effect of calcium and anaerobiosis on the thermostability of *Bacillus stearothermophilus*. *Canadian Journal of Microbiology* 22: 468-474

- Murray, W. D., and Asther, M. (1984) Ethanol fermentation by hexose and pentose wood sugars produced by hydrogen-fluoride solvolysis of aspen chips. *Biotechnology Letters* 6: 323-326
- Nazina, T. N., Tourova, T. P. Poltaraus, A. B., Novikova, E. V., Grigoryan, A. A., Ivanova, A. E., Lysenko, A. M., Petrunyaka, V. V., Osipov, G. A., Belyaew, S. S. and Ivanov, M. V. (2001) Taxonomic study of aerobic thermophilic bacilli: descriptions of *Geobacillus subterraneus* gen. nov., sp. nov. and *Geobacillus uzenensis* sp. nov. from petroleum reservoirs and transfer of *Bacillus stearothermophilus*, *Bacillus thermocatenulatus*, *Bacillus thermoleovorans*, *Bacillus kaustophilus*, *Bacillus thermoglucosidasius* and *Bacillus thermodenitrificans* to *Geobacillus* as the new combinations *G. stearothermophilus*, *G. thermocatenulatus*, *G. thermoleovorans*, *G. kaustophilus*, *G. thermoglucosidasius* and *G. thermodenitrificans*. *International Journal of Systematic and Evolutionary Microbiology* 51: 433-446.
- Nizza, D. T. and Gawrisch, K. (2009) A layer model of ethanol partitioning into lipid membranes. *General Physiology and Biophysics* 28: 140-145
- Novotny, J. F. and Perry, J. J. (1992) Characterization of bacteriocins from two strains of *Bacillus thermoleovorans*, a thermophilic hydrocarbon-utilizing species. *Applied and Environmental Microbiology* 58: 2393-2396
- Osipiuk, J., Sriram, M., Mai, X., Adams, W. W. M. and Joachimiak, A. (2000) Cloning, expression and crystallization of GroEL proteins from *Thermococcus litoralis*. *Acta Biochimica Polonica* 47: 209-214

- Park, J.M., Werner,J., Kim, J.M., Lis, J.T. and Kim, Y-G. (2001) Mediator, not holoenzyme, is directly recruited to the heat shock promoter by HSF upon heat shock. *Molecular Cell* 8: 9–19
- Payton, M. A. (1984) Production of ethanol by thermophilic bacteria. *Trends in Biotechnology* 2: 153-158
- Pernick, R. and Wilder, C. (2007) *The clean tech revolution: The next big growth and investment opportunity*. HarperCollins
- Podshun, R. and Ullmann, U. (1998) *Klebsiells* sp. as nosocomial pathogens: epidemiology, taxonomy, typing methods and pathogenic factors. *Clinical Microbiology* 11: 589-603
- Popova, N. A., Nikolaev Yu, A., Tourova, T. P., Lysenko, A. M., Osipov, G. A., Verkhovtseva, V. N. and Panikov, N. S. (2002) *Geobacillus uralicus*, a new species of thermophilic bacteria. *Microbiology* 71: 335-341
- Ranford, J. C., Coates, A. R. and Henderson, B. (2000) Chaperonins are cell-signalling proteins: the unfolding biology of molecular chaperones. *Expert Reviews in Molecular Medicine* 2: 1–17
- Rasouly, A. and Ron, E. Z. (2009) Interplay between the heat shock response and translation in *Escherichia coli*. *Research in Microbiology* 160: 288-296
- Ratkowsky, D.A., Olley, J., McMeekin, T.A. and Ball, D.A. (1982) Relationship between temperature and growth rate of bacterial cultures. *Journal of Bacteriology* 149*:*1-5

- Reischl, S., Thomas, W. and Schumann, W. (2002) Isolation and analysis of mutant alleles of the *Bacillus subtilis* HrcA repressor with reduced dependency on GroE function. *The Journal of Biological Chemistry* 277: 32659-32667
- Riyanti, E. I. and Rogers, P. L. (2009) Kinetic evaluation of an ethanol-tolerant thermophile *Geobacillus thermoglucosidasius* M10EXG for ethanol production. *Indonesian Journal of Agriculture Science* 10: 34-41
- Rocha, E.P. (2004) Codon usage bias from tRNA's point of view: redundancy, specialization, and efficient decoding for translation optimization. *Genome Research* 14:2279-2286
- Sambrook, J. F. and Russell, D. W. (2001). Molecular cloning: a laboratory manual. 3rd edition. Cold Spring Harbor Laboratory Press
- Samson, R. (1991) Switchgrass: A living solar battery for the prairies. Ecological agriculture projects. *Human Ecology* 35: 653–668
- Sardessai, Y. and Bhosle, S. (2002) Tolerance of bacteria to organic solvents. *Research in Microbiology* 153: 263-268
- Savage, N. (2011). Fuel options: the ideal biofuel. *Nature*. 474: 9–11
- Schuön, U. and Schumann, W. (1993) Molecular cloning, sequencing, and transcriptional analysis of the GroESL operon from *Bacillus sterothermophilus. Journal of Bacteriology* 175: 2465-2469
- Segal, G. and Ron, E. Z. (2006) Regulation of heat shock response in bacteria. *Annals of the New York Academy of Sciences* 851: 1-6
- Shine, J. and Dalgarno, L. (1975) Determinant of cistron specificity in bacterial ribosomes. Nature. 254: 34–38

- Silveira, M. G., Baumgärtner, M., Rombouts, F. M. and Abee, T. (2004) Effect of adaptation to ethanol on cytoplasmic and membrane protein of *Oenococcus oeni*. *Applied and Environmental Microbiology* 70:2748-2755
- Singh, V. K., Govindarajan, R., Naik, S. and Kumar, A. (2000) The effect of hairpin loop structure on PCR amplification efficiency. *Molecular Biology*. 3: 67-69
- Solomon, B. D., Barnes, J. R. and Halvorsen, K. E. (2007) Grain and cellulosic ethanol: history, economics and energy policy. *Biomass and Bioenergy* 31:436-425
- Staples, D. H. and Hindley, J. (1971) Ribosomal binding site of RNA polymerase cistron. *Nature New Biology* 234: 211- 212
- Straus, D. B., Walter, W. A. and Gross, C. A. (1987) The heat shock response of *E. coli* is regulated by changes in the concentration of sigma 32. *Nature* 329: 348-351
- Strohmeier, H., Gort, A. and Imlay, J. A. (1998) Balance between endogenous superoxide stress and antioxidant defences. *Journal of Bacteriology* 180: 1402-1410
- Strydom, B. D. (2009). Master's thesis: The economic impact of maize based ethanol production in South African animal feed industry. University of the Free State, Bloemfontein
- Sun, Y. and Cheng, J. (2001) Hydrolysis of lignocellulosic materials for ethanol production: *Bioresources Technology* 83: 1-11

- Surisetty, V.R., Dalai, A.K. and Kozinski, J. (2011) Alcohols as alternative fuels: An overview. *Applied Catalysis A: General* 404: 1–11
- Sutton, M. D. and Peterson, J. B. (2001) Fermentation of sugar beet pulp for ethanol production using bioengineered *Klebsiella oxytoca* strain P2. *Journal of Sugar Beet Research* 38: 19-35
- Taylor, M., Tuffin, M., Burton, S., Eley, K. And Cowan, D. (2008) Microbial response to solvent and alcohol stress. *Biotechnology Journal* 3: 1388-1397
- Thompson, H. A., Studholme, J. D., Green, M. E. and Leak, D. J. (2008) Heterologous expression of pyruvate decarboxylase in *Geobacillus thermoglucosidaius*. *Biotechnology Letters* 30: 1359-1365
- Tolan, J.S. and Finn, R. K. (1987) Fermentation of d-xylose to ethanol by genetically modified *Klebsiella planticola. Applied and Environmental Microbiology* 53: 2039–2044
- Truscott, K. N., Høj, P. B. and Scopes, R. K. (1994) Purification and characterization of chaperonin 60 and chaperonin 10 from the anaerobic thermophile *Thermoanaerobacter brockii*. *European Journal of Biochemistry* 222: 277-284
- Van Gerpen, J. (2005). Biodiesel processing and production. *Fuel Processing Technology.* 86: 1097 – 1107
- Vidal, J. (2010) http://www.theguardian.com/environment/2010/jan/22/quarter-us-grain-biofuels-food

- Wakeford, J. (2008) Implications of global oil depletion for transport planning in South Africa. Proceedings of the 27th Southern African Transport Conference, Pretoria
- Walfridsson, M., Hallborn, J., Penttilā, M., Kerānen, S. and Hahn-Hāgerdal, B. (1995) Xylose metabolizing *Saccharomyces cerevisiae* strains over-expressing the TKL1 and TAL1 genes encoding the pentose phosphate pathway enzymes transketolase and transaldolase. *Applied and Environmental Microbiology* 61: 4184-4190
- Walker, G. C. (1996) The SOS response of *E. coli*. ASM Press, Washington
- Wang, M., Zhao, J., Yang, Z. and Du, Z. (2007). Electrochemical insights into the ethanol tolerance of *Saccharomyces cerevisiae*. *Bioelectrochemistry* 71: 107-112.
- Wang, Z., Yuquan Xu, W., Li, S., Lu, W., Ping, S., Zhang, W. and Lin, M. (2008) An ethanol-tolerant recombinant *Escherichia coli* expressing *Zymomonas mobilis pdc* and *adhB* genes for enhanced ethanol production from xylose. *Biotechnology Letters* 30: 657-663
- Weng, J., Li, X., Bonawitz, N. D. and Chapple, C. (2008) Emerging strategies of lignin engineering and degradation for cellulosic biofuel production. *Current Opinion in Biotechnology* 19: 166-172
- Westerman, P. W., Pope, J. M., Phonphok, N., Doane, J. W., Dubro, D. W. (1988) The interaction of n-alkanols with lipid bilayer membranes: a 2H NMR study. *Biochimica et Biophysica Acta 939*: 64-84

- Wyman, C. E. (1996). Ethanol production from lignocellulosic biomass. In: Wyman, C. E. (ed.) Handbook on bioethanol: production and utilization. Taylor & Francis, Bristol
- Xu, Z., Horwich, A. L. and Sigler, P. B. (1997) The crystal structure of the Asymmetric GroEL-GroES-(ADP) - chaperonin complex. *Nature* 388: 741-750
- Yang, X., Xu, L. and Yan, M. (2010) Modelling metabolic response to changes of enzyme amount in yeast glycolysis. *African Journal of Biotechnology* 9: 6894-6901
- Young, E., Lee, S., Alper, H. (2010). Optimizing pentose utilization in yeast: the need for novel tools and approaches. *Biotechnology for Biofuels*. 3 : 1-12
- Yu, D., Ellis, H. M., Lee, E. C., Jenkins, N. A., Copeland, N. G. and Court, D. L. (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proceedings of the National Academy of Sciences* 97: 5978–5983
- Yura, T., Nagai, H. and Mori, H. (1993) Regulation of the heat shock response in bacteria. *Annual Review of Microbiology* 47: 321-350
- Zhang, B., Yang, Q., Zhang, X. and Song, J. (2007) Mechanism and research progress of ethanol fermentation related genes *pdc* and *adhll* involved in fuel ethanol production. *KMITL Science and Technology Journal* 7:171-178

- Zhang, H., Lin, L., Zeng, C., Shen, P. and Huang, Y. (2007) Cloning and characterization of a haloarchaeal heat shock protein 70 functionally expressed in *Escherichia coli*. *FEMS Microbiology Letters* 275: 168-174
- Ziemienowicz, A., Zylicz, M., Floth, C. and Hubscher, U. (1993) Calf thymus Hsc 70 protein protects and reactivates prokaryotic and eukaryotic enzymes. *Journal of Biological Chemistry* 270: 15479-15484
- Zwietering, M. H., Jongenburger I., Rombouts F. M. and van 'T Riet, K. (1990) Modeling of the bacterial growth curve. *Applied and Environmental Microbiology* 56: 1875–1881