An Investigation on *E. coli* Host Strain Influences and Strategies to Improve Supercoiled Plasmid DNA Production for Gene Therapy and Vaccination

A thesis submitted to University College London
for the degree of
DOCTOR OF PHILOSOPHY
in Biochemical Engineering

By Sin Yee YAU (MEng, AMIChemE)

2010

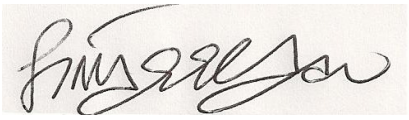
Advanced Centre of Biochemical Engineering
Department of Biochemical Engineering
University College London
Torrington Place, London WC1E 7JE, UK

Declaration

The author confirms that this thesis presents work which is the author's own except in the area listed below:

Chapter 3: The Microsoft Excel-based program which applies Principal Components Analysis on the data set was written and coded in-house by Dr Simon Edwards-Parton.

Signed:



Copyright

Acknowledgments

I would like to take this opportunity to thank everyone who has helped and inspired me throughout the past three years of my PhD studies. I would especially like to thank my supervisor Professor Eli Keshavarz-Moore and advisor Professor John Ward for their much appreciated advice, unwavering support and help.

Above all, I would like to acknowledge and thank my boyfriend Michael Rose, who has constantly provided me with his support, encouragement and ideas especially during the stressful times. He has also on many occasions willingly stayed late and overnight with me during my fermentation runs. I hope I have been just as supportive with his PhD!

Many thanks also to Gareth Mannall and Ian Buchanan for their technical support and help in the ACBE fermentation hall and MMP lab, and to Simon Edwards-Parton for his sharing his knowledge on PCA, the PCA program he wrote and coded and of course, teaching me how to use it.

Thanks to the lab members of the MMP lab and everyone in Vineyard office in the ACBE, particularly Michael Wong, Brenda Parker, Sally Hassan, Gaik Sui Kee, Carol Knevelman and Darren Nesbeth- I won't forget the times we spent sharing our joys, worries and gossip!

To my other adopted lab, Foster Court and its members, I thank you all for letting me come over and use the equipment! Steven Branston, Emma Stanley, Adam Stonier, Simyee Kong and Daniel Gibbons- thank you all for your help and advice.

Thanks to Dr David Summers and Chin Chin Chen from the Department of Genetics at Cambridge University for providing me with the Q-cells and for the induction tips during the week I was there.

Also many thanks to the Rose family for letting us stay with them during our writing up and job hunting period, which unbelievably coincided with the worst economic downturn in decades.

Finally, I would like my parents and family back in Northern Ireland for their support and understanding during the past 5 years.

I hope in the end, the hard work paid off and it was all worth it.

Sin Yee Yau
2010

Abbreviations

A	Absorbance Unit
AAV	Adeno-Associated Virus
AGE	Agarose gel electrophoresis
CAT	Chloramphenicol acetyltransferase
CCC	Covalently Closed Circles
DCW	Dry cell weight
DF	Diafiltration
DNA	Deoxyribonucleic acid
DOT	Dissolved Oxygen Tension
FDA	Food and Drug administration (USA)
gDNA	Genomic DNA
GFP	Green Fluorescent Protein
H-NS	Histone-like nucleoid structuring protein
OC-DNA	Open Circular DNA
OD600nm	Optical density at 600nm
PC	Principal Component
PCA	Principal Components Analysis
pDNA	Plasmid DNA
Q-cells	Quiescent cells
RFU	Relative Fluorescence Unit
RNA	Ribonucleic acid
RO	Reverse osmosis
scFv	Single-chain antibody fragment
SC	Supercoiled
SC-DNA	Supercoiled DNA
SDCAS	Semi defined medium with casamino acids
SEC	Size exclusion chromatography
ssDNA	Single stranded DNA
TB	Terrific Broth
TBE	Tris-Borate Electrophoresis buffer
TCA	Tricarboxylic Acid
UF	Ultrafiltration

An Investigation on Host Strain Influences and Strategies to Improve Supercoiled Plasmid DNA Production

Abstract

The growing demand for quick and effective methods of producing large amounts of plasmid DNA for human therapy and vaccination has increased the practical challenges associated with process optimisation to improve supercoiled plasmid DNA yields obtained through current methods. The supercoiled isoform of DNA is the preferred form for use in gene therapy and vaccination as this isoform is known to produce higher levels of *in vitro* and *in vivo* transgene expression than other forms of plasmid DNA.

This study was designed to investigate whether different strategies can be implemented early on in a process to improve supercoiled plasmid DNA yields obtained upstream, with the view to aid and/or ease further downstream stages. The main theme investigated is the influence of the host strain on supercoiled plasmid DNA production. Seventeen strains of *Escherichia coli* and three different plasmids were investigated at shake flask scale, before two strains were selected for scale up to 7L fermentation scale. The results obtained indicated that the host strain plasmid combination can heavily influence both the quantity and quality of plasmid DNA obtained and this behaviour cannot simply be determined by looking at the host strain genotype. Fermentation runs on the two strains selected for scale up (BL21 DE3 gWiz and HB101 gWiz) demonstrated that these two strains scale up very well, maintaining high specific pDNA yields (1.5mg/L/OD for BL21 DE3 gWiz) and high SC-DNA yields (98% for HB101 gWiz).

Temperature amplification studies using strains harbouring pUC18 have shown that although most strain-plasmid combinations yielded more plasmid at a higher temperature of 40℃, the extent of this increase is highly influenced by the host strain. Indeed in some cases, such as for the

strains ABLE K, W3110, W1485, a higher plasmid yield was obtained at 37℃. However, as the growth rates of these cultures were not measured, the extent of the accumulation of plasmid DNA due to the effects of the growth rate and/or temperature during the exponential phase of growth is unknown at this time. Similarities to what has been reported as temperature induced runaway plasmid replication have been observed in this study, although no experiments were conducted to confirm whether these observations were indeed as result of runaway replication as defined in the literature.

Potential alternative strategies investigated included implementing anaerobiosis to test if these conditions can improve supercoiled plasmid DNA production at fermentation scale, and whether a 'Quiescent cell expression system' (a state where chromosomal replication and expression is temporarily shut down but residual proteins remain metabolically active) can be implemented to improve plasmid DNA yields by redirecting resources away from biomass production. The results suggest that under the conditions set in this study, these strategies do not increase plasmid DNA production or the percentage of supercoiled plasmid obtained.

In conclusion, the results from this investigation have demonstrated that a highly effective and influential strategy for improving the quality and quantity of plasmid DNA obtained is the initial choice of the host strain-plasmid combination. Further improvements can then be obtained through the application of other reported fermentation strategies.

Contents Page

List of Tables **Page**

List of Figures **Page**

1. Introduction

1. Introduction

The growing demand for quick and effective methods of producing large amounts of plasmid DNA for human therapy and vaccination has increased the practical challenges associated with the optimisation of plasmid DNA yield obtained through current methods. The popularity of using plasmid DNA for human applications is primarily due to its well documented safety, the absence of specific immune responses to the plasmid and also due to its lack of genetic integration (Prather et al., 2003). To date there have been 1579 completed, ongoing or pending gene therapy trials globally since 1989, of which 18% have used naked/plasmid DNA (Wiley, 2009). The current demand is specifically for plasmid DNA in the supercoiled form for administration in vaccines and for human gene therapy.

1.1. Plasmid Gene Therapy

Recombinant plasmid DNA is used both as a raw material as in Adeno-Associated Virus (AAV) vector production as well as an active ingredient in the form of DNA vaccines. The naked DNA containing the gene of interest is delivered in the form of a plasmid and in some cases the plasmid maybe also formulated with a carrier to improve delivery. Plasmid therapy involves using plasmids vectors to deliver the gene of interest to the patient and this transferred gene should replace, repair or repress a defective gene in the patient. More than 4,000 human diseases are caused by a single faulty or missing gene (Gottschalk & Chan, 1998) and these include cystic fibrosis and sickle cell anaemia.

1.1.1. DNA Vaccines

The advantage of using a plasmid vector is that it contains no viral component and is unaffected by pre-existing immunogenicity. The simplicity of this method has led to the development of DNA vaccines. Like most vaccines, DNA vaccines induce antigen-specific antibody responses allowing the body to recognise and react quickly to a real infection. However the production of DNA vaccines do not require the cultivation of dangerous or infectious agents (like live viruses) and there is

no risk of an attenuated virus mutating back to its virulent form, hence making DNA vaccines a safer alternative to traditional vaccines. There is also potential for DNA vaccines to treat diseases which have proven difficult to be administered through an attenuated virus, such as HIV. Current issues concerning the use of plasmid therapy is that it is difficult to target specific tissues and the low level of expression in vivo due to the transient expression of the plasmids; however there have been a few reports on preliminary tests that have provided some positive results on a vaccine for Multiple Sclerosis (Stuve et al., 2007). A DNA vaccine has already been approved and licensed for animal use in the form of a West Nile Virus DNA vaccine for horses (CDC press release, 2005; Davis et al., 2001).

1.2. Plasmid DNA

1.2.1. General Properties

In general most plasmids are circular double-stranded DNA molecules that are separate from the chromosomal DNA. Bacterial plasmids are extrachromosomal and plasmids may appear in different forms, covalently closed circles (CCC), open circles (OC) or linear, and can also be multimeric as concatamers or catenates. Their size varies from 1 kilo bases up to 1.7 mega bases. There are anywhere from one copy, for large plasmids, to hundreds of copies of the same plasmid present in a single cell. Generally for recombinant DNA studies, most plasmid-based vectors are derivatives of ColE1 or pMB1.

1.2.2. Advantages of Plasmid DNA

Plasmids often contain genes or gene-cassettes that confer a selective advantage to the bacterium hosting them, such as antibiotic resistance, heavy metal ion resistance, virulence, nitrogen fixation and fertility. Plasmids may also contain genes that are essential for host cell growth and division. A plasmid contains one or more DNA sequences that serves as an origin of replication or ori (a starting point for DNA replication), enabling the plasmid DNA to be replicated autonomously.

Plasmids used in gene therapy and vaccines are used as vectors to transfer genes from one organism to another and typically contain a genetic marker conferring a phenotype that can be used as a selection agent. Most plasmids also contain a multiple cloning site which is a region on the plasmid containing several commonly used restriction sites. This region enables the easy insertion of DNA fragments with the correct ligated ends.

1.2.3. Plasmid Transfer

Horizontal plasmid transfer can occur in three ways, conjugation, transformation and transduction (Dale, 1998; Ingraham et al., 1983; Lodish, 1999). Conjugative plasmids contain *tra*-genes, which perform the conjugation process, allowing the sexual transfer of plasmids to another bacterium. These plasmids can then be sub-classed as repressed or derepressed. Most naturally occurring plasmids carry conjugation systems which are repressed, whereby the transfer functions switched off and are only transferred by the isolation of derepressed mutants (*drd*$^-$) or if the repression genes fail (Glass et al., 1982).

Non-conjugative plasmids are incapable of initiating conjugation and can only be transferred with the assistance of conjugative plasmids. Non-conjugative plasmids are either mobilisable or non-mobilisable. Mobilisable plasmids carry only a subset of the genes required for transfer and are not self-transmissible, but use the transfer apparatus of other plasmids to mediate their transfer. In nature plasmids are typically conjugative or mobilisable, however these natural features are removed or disabled when the plasmids are used for human therapy to prevent the possible uncontrollable spread of genetic engineered DNA.

1.2.4. Conjugation

Conjugation differs in gram-negative and in gram-positive bacteria. In gram-negative bacteria where a double membrane is present, plasmids confer a fertility function encoded by the F-plasmid and produce sex pili. These pili are either thin flexible, thick flexible or rigid and they aid the

conjugation process by bringing the plasmids together, acting like a bridge to transfer the genetic material.

1.2.5. Transformation

Transformation involves the genetic alteration of a cell due to the introduction, uptake and expression of foreign DNA or RNA from the environment such as those released from dead cells. Natural genetic transformation by bacterial cells involves the active uptake of extracellular chromosomal or plasmid DNA and the heritable incorporation of its genetic information (Lorenz and Wackernagel, 1994). Natural genetic transformation constitutes a major horizontal gene-transfer mechanism in prokaryotes (Vries and Wackernagel, 2002). Artificial competence involves passively making the cell permeable for the introduction of extracellular DNA into bacterial cells, for example, by using divalent cations like calcium chloride ($CaCl_2$) or electroporation techniques (Sambrook et al., 1989).

1.2.6. Transduction

Transduction can also be used to transfer plasmids using bacteriophages. Plasmids incorporated into the bacteriophage DNA packaging can be introduced to other bacterial cell upon infection.

1.3. Choice of Insert and Vector

When choosing an insert and vector for clinical applications, the safety and potency of the plasmid should be considered. The FDA requires that the entire plasmid is sequenced before clinical trials may commence (Smith and Klinman, 2001).

The behaviour of the plasmid in its bacterial host is also an important factor in the selection of the vector and host. The plasmid yield and stability and the ability for plasmid selection are areas in which manufacturers continuously want to optimise. The plasmid yield is determined by the origin of replication which determines the plasmid copy number. It is generally recommended that a pUC replication origin is used as pUC plasmids have high copy numbers (usually around 500-700) and

are also well characterised (Prazeres et al., 1999; Marquet et al., 1995; Sambrook et al., 1989).

1.3.1. Plasmid Selection

Selecting the plasmid is usually carried out in the form of utilising antibiotic resistance. This method ensures that only the cells containing the resistant plasmid are able to propagate further in the culture. For pSVβ and many other commonly used plasmids in the laboratory, the antibiotic used as a selection tool is ampicillin.

Plasmids with the amp^r gene are ampicillin resistant. Ampicillin binds to and inhibits a number of enzymes in the bacterial membrane that are involved in the synthesis of the cell wall. However ampicillin is unsuitable for full scale production (Marquet et al., 1995) as even in minute quantities, it can cause severe allergic reactions in some people. Such safety risk would limit its use and the applications of the plasmids which are ampicillin resistant, unless they also express resistance to another antibiotic.

The FDA recommends the use of aminoglycoside antibiotics such as kanamycin and neomycin as selection agents instead of ampicillin (FDA 1996a, 1996b). Kanamycin and neomycin are deoxystreptamine aminoglycosides that bind to ribosomal components and inhibits protein synthesis. Phosphorylation of these antibiotics is thought to interfere with their active transport into the cell. However kanamycin is not widely used in medicines due its unwanted side effects such as ototoxcity (Humes, 1984) and foetal damage.

1.3.2. Plasmid Size

Present clinical trials use plasmids generally less than 10kb in size although studies have shown that in the near future, larger plasmids will be required to accommodate newer concepts and applications as the field progresses into more complicated areas (Levy et al., 2000). It is hypothesised by Levy et al., (2000), that as plasmid size increases, the

general stress on the host cell would also increase which would lead to a decrease in the maximum yield. Previous studies on a few *E. coli* strains with plasmids of different sizes up to 8kb showed that the increase in plasmid size did not significantly alter the growth rate, but a decrease in the host growth rate was observed with a 21kb plasmid when the metabolic load on the host cell was far greater (Cheah et al., 1987, Warnes and Stephenson, 1986). These reports however do stress that the impact of increased plasmid size on the growth rate is strictly dependent on the host/vector system and not only on the plasmid size. This is further supported by studies on the impact of plasmid size on cellular oxygen demand in *E. coli* (Kay et al., 2003), where it has been shown that it is not the plasmid size that causes an increase in host cellular oxygen demand but it is rather due to the increased amount of protein produced by larger plasmids.

1.3.3. Copy Number & ColE1 Plasmid DNA Replication

The copy number of a plasmid is defined as the average number of plasmids per bacterial cell or per chromosome under normal growth conditions. Many copies of the same plasmid can exist in the same cell and the plasmid copy number varies generally with the size of the plasmid (large plasmids tend to have low copy numbers and small plasmids tend to have high copy numbers). Since most plasmid-based vectors are derivatives of ColE1, the ColE1 mode of plasmid replication and copy number regulation is one of the most well-known. Here the copy number is controlled by the rate of initiation of replication and this generally involves RNA primers (Dale, 1998).

1.3.3.1. Relaxed and Stringent Control

Controlled by the plasmid replicon, the copy number of a plasmid can fluctuate within a narrow range in response to changes in the growth conditions of the bacterial culture. At steady state, the copy number remains constant as the population of plasmid doubles at exactly the same rate as the population of host cells (Sambrook et al., 2001).

Plasmids that have RNA as a positive regulatory molecule controlling the frequency of initiation of plasmid DNA synthesis, such as pMB1/ColE1 replicon, generally have high copy numbers. These plasmids do not require plasmid-encoded proteins for replication; rather these plasmids replicate at random and rely on enzymes and proteins supplied by the host (Bazaral and Helinski, 1970). These are known as “relaxed” plasmids and will continue to replicate when protein synthesis is inhibited, for example, by chloramphenicol addition (Clewell and Helinski, 1972; Clewell, 1972) or amino acid starvation (Hoffman 1990).

“Stringent” plasmids such as pSC101 and the F plasmid require the ongoing synthesis of the RepA protein for replication, thus their copy numbers cannot be amplified by the inhibition of protein synthesis through amino acid starvation or chloramphenicol addition. Stringent plasmids tend to have low copy numbers and seem to be replicated during a short interval during the cell division cycle (Finkelstein and Helmstetter, 1977).

1.3.3.2. ColE1 Replication Control Mechanism

The ColE1 replication origin is a 600 base-pair region from which DNA replication proceeds (Tomizawa et al., 1974 & 1975). When a RNA preprimer, RNA II, is synthesised from a region close to the origin, it is cleaved by the host-encoded enzyme RNase H to release the 3′OH for elongation by the *E. coli* DNA polymerase I, thus allowing replication to occur. When the complementary inhibitor RNA I binds to RNA II and forms a ‘kissing complex’ (which is stabilised by the rop protein, aka the rom protein), this cleavage is prevented. In this case, no free replication primer is available for DNA replication and therefore replication does not occur. RNA I regulates the frequency of initiation of replication, and thus the plasmid copy number.

Another gene, the *rom/rop* gene synthesizes the rom/rop protein that facilitates the binding between the RNA I and RNA II molecules (affecting the rate of formation of the RNA I/RNA II complex). These mechanisms ensure that less RNA II is available to serve as an initiation signal for the

DNA polymerase, slowing DNA replication and thereby regulating the copy number. If this regulating mechanism is left uncontrolled, the replication process will continue and many copies of DNA would be made. The pUC series of plasmids are an example of this as they lack the *Rom/Rop* gene and have a single point mutation in the replication primer RNA II (Lin-Chao et al.,1992), resulting in high copy numbers per cell during replication (around 500-700 per cell).

Excessive uncharged tRNAs can result from amino acid starvation and it has been reported that the accumulation of uncharged tRNAs can also have an impact on plasmid replication control in relaxed plasmids. Uncharged tRNAs (tRNAs that are not charged with the specific amino acid) interact with the RNA I and/or the RNA II, preventing the 'kissing complex' from forming (Wrobel and Wegrzyn, 1998; Yavachev and Ivanov, 1988). This interaction occurs due to the tRNAs' sequence homology to three RNA-loop structures (Yavachev and Ivanov, 1988), present in RNA I and RNA II of the origin of replication which can interfere with the plasmid copy number control, leading to the rapid increase in plasmid copy number.

Wrobel and Wegrzyn, (1998), have reported that the extent of amplification is strongly dependent on the particular limiting amino acid and the specific origin of replication. Various kinds of tRNA have different homology to RNA I and/or RNA II and different kinds of uncharged tRNA molecules are predominant during starvation of particular amino acids; thus the interaction of tRNA with RNA I and/or RNA II may be stronger or weaker depending on the nature of the deprived amino acid. The authors have also found that for ColE1 plasmids, the greatest amplification was obtained by leucine starvation. Furthermore, the amplification rates for some replicons are higher when amino acid limitation is used compared to amino acid starvation (Wrobel and Wegrzyn, 1997).

1.4. Host Selection

Plasmids are most commonly produced in *E. coli*. Generally *E. coli* strains with high plasmid copy numbers and high plasmid retention levels are used, particularly if they are compatible with the latter downstream recovery and purification stages (Durland and Eastman, 1998). To date there has been no published data regarding the common consensus on choosing a particular *E. coli* strain based on its genotypic or phenotypic properties. However the strain of *E. coli* used to propagate a plasmid can have sizeable influence on the quality of the purified DNA, for example, DH1, DH5α and C600 yields high quality DNA (O'Mahoney et al., 2007) whereas HB101 contains large amounts of carbohydrates which can inhibit enzyme activities if not completely removed (Prazeres et al., 1999).

In a commercial environment, the host cell line selection is a critical part of any platform process development strategy; hence the decision regarding the choice of host strain will be carefully made based on characterisation, safety, source history and traceability, and intellectual property issues (Nail and Akers, 2002). With these criteria in mind and the time and resources needed to establish each master cell bank, it is likely that many commercial organisations will only have a few established and well-characterised standard cell lines available to choose from.

Benefits may be obtained through the evaluation of a number of these well characterised strains and even some less common ones to evaluate the properties of each and how well suited these strains are as hosts to plasmid vectors. It is now a requirement that the source and history of bacterial cells used, including confirmation of its stated phenotype and genotype and what procedures are used to generate the cells, are documented as part of a new drug application. (FDA- 'Points to Consider on Plasmid DNA Vaccines for Preventive Infectious Disease Indications', 1996).

Most strains used for cloning are derived from *E. coli* K-12 since the safety of K-12 strains have been the most thoroughly investigated and hence is

preferred by biosafety regulatory bodies, compared to others such as derivatives of *E. coli* B (Bachman, 1996). K-12 strains also possess the *EcoK* restriction-modification system (Raleigh et al., 1989), coded by *hsd* genes. Strains used in cloning are usually hsdR or hsdS so that there is no possibility of *EcoK* restriction sites in the cloned DNA being cleaved (Brown, 2000). Similarly, *mcrA* and *mcrBC* strains also lack functional McrA and McrB restriction systems which cleave DNA at short target sequences (Raleigh et al., 1988).

For supercoiled pDNA production, the mutations of interest include *endA1* and *recA*, which are thought to prevent plasmid degradation and nicking during lysis (Carnes, 2005; Schoenfeld et al., 1995) and to improve structural plasmid stability respectively. Another mutation of interest is the *gyrA96* mutation, conveying resistance to nalidixic acid. Strains that have the *gyrA96* mutation are reported to have reduced levels of supercoiled DNA as they increase the plasmid linking number (Ramirez and Villarejo, 1991).

1.5. Supercoiled Plasmid DNA

Plasmids exist in three main forms, linear, open circular (OC) and supercoiled (SC). In *E. coli*, DNA supercoiling is maintained by the opposite activities of at least two topoisomerases; topoisomerase I which relaxes negatively supercoiled DNA, and the ATP-dependant DNA gyrase, a type II topoisomerase that introduces negative supercoils (Dorman, 1991). DNA gyrase effectively converts the chemical energy released on hydrolysis of ATP into torsional energy in a negatively supercoiled DNA chain. As a result of DNA twisting, the Gibbs free energy content of supercoiled DNA is high. Twisting of DNA can occur in both directions, by rotation clockwise (negative twisting) or anticlockwise (positive twisting) around the axis of the helix. Different growth conditions may affect the level of negative supercoiling such as osmotic shock, temperature increase, anaerobiosis and nutrient starvation (Hsieh et al., 1991 a & b; Conter et al., 2003; Goldstein and Drlica, 1984; Cortassa and Aon, 1993; Jensen et al., 1995).

Both type I and type II topoisomerases cause regional relaxation or increased DNA supercoiling. Topoisomerase II also has the ability to catenate (interlink), decatenate, knot, unknot and relax circular double-stranded DNA. This ability to catenate can result in plasmid multimer formation. Multimeric plasmid molecules contain multiple copies of a functional gene and can in theory be more efficient in transfection experiments than the monomeric form. The structural stability of supercoiled dimers and trimers has been reported to be sufficient during batch cultivation in a bioreactor (Voss et al., 2003), although more research is needed to confirm if higher gene expression can be accomplished after transfection. However, the use of multimeric plasmids in a process may be a regulatory issue as different forms of plasmid molecules may have different potencies (Durland and Eastmann, 1998).

Plasmid topology is particularly important since the current perception is that the monomeric supercoiled form is the preferred form of DNA for use in gene therapy and vaccination (Prazeres et al., 1999). The supercoiled plasmid isoform is known to produce higher levels of *in vitro* and *in vivo* transgene expression when compared with open circular and linear isoforms (Voss et al., 2003; Schleef and Schmidt, 2004). FDA requirements specify that a minimum level of supercoiling in the final product should be stated. Therefore from a pharmaceutical point of view, the aim is to obtain high levels of SC-DNA and less of the contaminating genomic DNA, RNA and proteins without compromising the stability of the DNA product.

1.5.1. Linking Number and Linking Number Distribution

The linking number has been used to describe the degree of supercoiling for closed loops of DNA. The change in the linking number (ΔLk) depicts the number of times the two strands of DNA double helix are intertwined for circular DNA (Dale 1998). This change is given by the equation:

$$\Delta Lk = \Delta Wr + \Delta Tw \qquad [1.6.1.1.]$$

where ΔWr is the change in DNA writhe and ΔTw is the change in DNA twist. DNA writhe describes the path of the DNA duplex (supercoiling) in space and DNA twist is the number of base pairs per turn of the DNA helix. The helical twist of DNA depends of the temperature and on the salt concentration (Gould, 1998). As the shape of the DNA duplex is not fixed, the writhing number and the twist number can vary continuously, although their sum total will stay constant providing the DNA strands are not broken.

Supercoils can decrease the number of base pairs per helical turn and cause the DNA helix to form a helix of higher order (Green and Rao, 1998). Since the degree of supercoiling of different circular DNAs may differ due to different DNA lengths, the superhelix density, σ, is sometimes used to refer to the number of supercoils per helical turn of DNA, as this is independent of DNA length. This is defined by the equation:

$$\sigma = \Delta Lk / \text{DNA twist number} \qquad [1.6.1.2.]$$

B-DNA is the most common form of DNA *in vivo*. The DNA twist number for B-DNA in solution equals 10.5bp/turn (Crick, 1976), therefore:

$$\sigma = \Delta Lk / (PL/10.5) \qquad [1.6.1.3.]$$

where PL is the length of plasmid in base pairs.

Most supercoiled plasmid molecules isolated from prokaryotes have σ values between -0.05 and -0.07 (Vologodskii et al., 1992).

In a supercoiled plasmid DNA sample the distribution of linking numbers is quantitatively symmetrical and follows the Boltzmann distribution defined by the free energy of supercoiling (Keller, 1975; Pulleyblank, 1975). A minimal linking number distribution of a supercoiled plasmid DNA sample conveys homogeneity (Pruss, 1985).

1.5.2. Detection of Linking Number

In order to detect and estimate the plasmid linking number distribution, toposiomers first need to be separated. Supercoiled plasmids from *E. coli* generally have a large number of supercoils, making the task of separating these from each other difficult. However the addition of a DNA intercalator such as chloroquine or ethidium bromide can increase the DNA writhe. This effectively increases the size of the plasmid and slows the electrophoretic mobility of the DNA, allowing the separation of plasmids based on linking number. The difference in linking number can be calculated by counting the number of bands (topoisomers) from the top band down on an agarose gel. Nicked plasmids are unaffected this way by intercalating agents since the superhelical stress can be removed by the rotation of one DNA strand around the other. Nicked plasmids therefore have a linking number of zero, the maximum size and the lowest mobility on a gel (Xu and Bremmer, 1997).

1.6. Plasmid DNA Manufacturing

Similar to recombinant protein production, large scale plasmid production can generate a number of recovery and purification issues, especially during the downstream processing operations. During these stages the purity, potency, identity, efficacy and safety specifications are most important and the key aim is to eliminate cellular components of the host strain. For *E. coli* B, the average composition of biomass present by dry cell weight has been reported by Ingram et al., (1983). The amount of plasmid present can vary considerably depending on the copy number and high copy number plasmids can make up at least 1-3% of the biomass present (Varley et al., 1999; Levy et al., 2000).

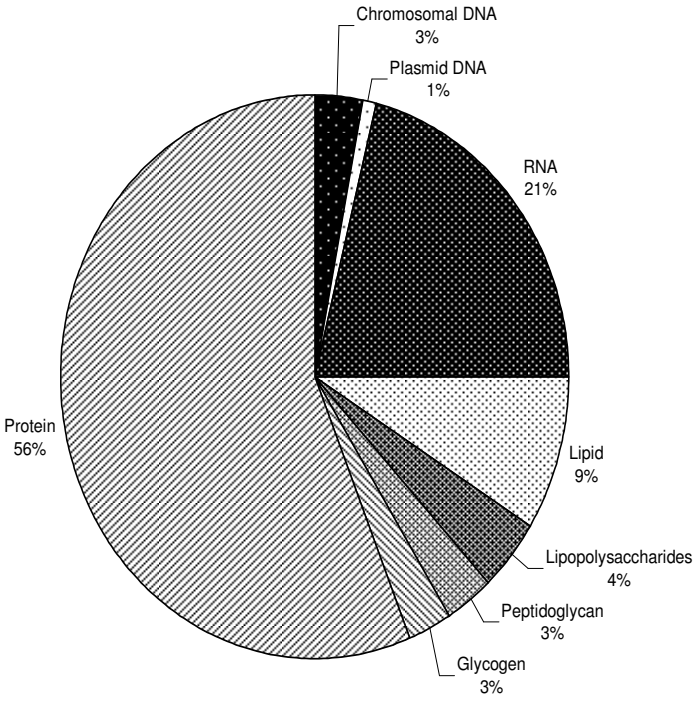


Figure 1.6. An average composition of biomass in *E. coli* B by % dry cell weight. Adapted from Ingrahm et al., 1983.

1.6.1. Cell Growth

Most pharmaceutical grade DNA use *E. coli* as host of which the latter, transformed with the plasmid DNA of choice, is grown in suitable conditions in a growth medium containing the selection agent (typically an antibiotic) before being harvested and lysed. The *E. coli* biomass then needs to be lysed to obtain the plasmid DNA. As plasmids are susceptible to shear (Levy et al., 1999), mechanical methods are typically not used and instead chemical lysis of the biomass is preferred, particularly alkaline lysis (Birnhoim & Doly, 1979).

1.6.2. Alkaline Lysis

Alkaline lysis involves using an alkaline solution usually containing SDS, sodium hydroxide (NaOH) and typically RNAse (Birnhoim & Doly, 1979). The SDS solubilises the phospholipid and protein components of the cell membrane causing cell lysis and also denatures proteins by forming a

SDS-polypeptide complex. The sodium hydroxide temporarily denatures the plasmid DNA but irreversibly denatures chromosomal DNA and host proteins lysed from the biomass. After becoming denatured, plasmid DNA immediately renatures properly in a neutralizing solution (typically potassium acetate which causes the SDS to precipitate); however a long incubation period in the NaOH solution may irreversibly denature the plasmid DNA and so is best avoided. The addition of the neutralising agent also results in the precipitation of protein and genomic DNA. RNAse or a similar enzyme degrades the host RNA to minimise RNA contamination. The denatured proteins and contaminants (together known as the "floc") trapped by the salt-detergent complexes of the alkaline solution are then removed. Other less common techniques are available for plasmid DNA recovery such as lysis by boiling (Sambrook et al., 1989).

1.6.3. Clarification

The plasmid DNA is separated from the lysate containing the floc by centrifugation, filtration or floculation. As mentioned, due to the susceptibility of the plasmids to shear, particularly larger plasmids, centrifugation is avoided for large scale production due to the high shear, semi-continuous mode of operation of industrial centrifuges (Theodossiou et al., 1998). In some cases, this harvesting and clarification step may be completely bypassed with the lysate being purified directly by chromatography. The latter option however is an expensive one considering the cost of chromatography columns and matrices and is not recommended for large volumes where fouling would be a serious problem.

1.6.4. Affinity Precipitation

Most large scale processes involve affinity precipitation using alcohols such as isopropanol or ethanol. Expensive chromatography stages can be avoided as the majority of contaminants can be removed by the addition of cationic detergents such as CTAB, polyethylene glycol, certain salts such as $LiCl_2$ and $CaCl_2$ or polyamines such as spermidine and polyethylene imine. These methods usually result in plasmid DNA yields

of greater than 90%, but are sensitive to temperature, pH, the method of washing of the precipitates and the degree of mixing (Prather et al., 2003).

1.6.5. Purification

Chromatography is the method of choice for large scale purification of supercoiled plasmid DNA (Ferreira et al., 2000; Prazeres et al., 1998). Chromatography can play a part both as a process step and as an analytical tool used for monitoring process development and quality control. However due to the large size of the plasmid molecule compared to the average pore size of common chromatographic supports (which tend to be designed for proteins), plasmid DNA cannot access pores in standard chromatographic media and can only bind to the outer surface of the particles. The result of this is that only 0.2-3g plasmid DNA bind per litre of resin, hence a significant amount (500-2000L) of chromatography media is needed (Prather et al., 2003).

Ion-exchange, reversed-phase, hydrophobic interaction, affinity and size exclusion have been used for the large scale separation of plasmid DNA (Durland and Eastman, 1998; Ferreira et al., 2000; Prather et al., 2003; Prazeres et al., 1998). Separation is achieved largely based on the size of the plasmid molecule, however the chemical properties (charge and hydrophobicity), the accessibility of the nucleotide bases to ligands and the topological constraints due to supercoiling should all be taken into consideration when choosing the chromatography approach and the order of each chromatography step if more than one is needed. In some anion exchangers, more compact SC forms which have a higher charge density elute later than the OC forms (Prazeres et al., 1998); therefore the presence of different plasmid forms and multimers may complicate purification and characterisation (Durland and Eastman, 1998).

1.6.6. Buffer Exchange

Buffer exchange and desalting can be performed by ultrafiltration (UF), diafiltration (DF), size exclusion chromatography (SEC) or alcohol-based precipitation of DNA. The issue associated with UF and DF is that at large

scale, a large membrane surface area and high fluid flow rates are required in order to minimise premature membrane fouling (Prather et al., 2003). The disadvantages of using SEC are its limited capacity and the dilution of plasmid (Prazeres and Ferreira, 2004). The use of alcohol precipitation of DNA will require safety measures in place such as the design of proof facilities or use of adequate protection masks (Marquet et al., 1995).

1.6.7. Final Product

The final plasmid DNA for gene therapy or DNA vaccine must be 0.22µm sterile filtered and must meet the minimum release specifications after purification, usually with >90% supercoiled form DNA, with an absorbance A260/A280 between 1.8-2.0. Contaminant specifications are made based on dose and regulatory authorities have specified a maximum concentration of contaminant protein, genomic DNA and endotoxin per mg, protein or dose. RNA should be undetectable by gel electrophoresis and there should be less than 0.01µg per dose of protein; less than 0.05µg per µg plasmid or 0.01µg per dose of genomic DNA and less than 0.1EU per µg plasmid endotoxin impurities present (Prazeres and Ferreira, 2004).

1.6.8. DNA recovery and purification using Mini-, Midi- and Maxi-prep kits

For laboratory-scale volumes of culture, Qiagen offer kits to recover DNA such as miniprep, midiprep and maxiprep kits all to suit different concentrations of DNA to be recovered. These kits often enable a faster and more straightforward way of recovering purified DNA based on the Birnhoim and Doly, (1979), alkaline lysis technique.

1.6.9. Storage of Plasmid DNA Samples

A recent study by Freitas et al., (2007), has reported that pDNA remains relatively stable compared to RNA if stored under appropriate conditions. In this report, *E. coli* cell pellets were able to be stored safely without compromising the pDNA stability for at least 3 weeks at 4°C and at least

12 weeks at -20℃ prior to processing. Alkaline lysates however, should be stored at -20℃ to avoid supercoiled pDNA degradation. Kong et al., (2008), reported that freeze-thawing cell biomass after recovery from various centrifugation methods resulted in supercoiled plasmid DNA degradation to OC form and this degradation is affected by the temperature and time at which the cells are stored. The authors found that the highest supercoiled yields were obtained when fresh *E. coli* cell biomass were harvested using a tubular bowl centrifuge and then resuspended in TE buffer for less than 2 hours at temperatures lower than 13℃.

1.7. Potential Alternative Methods for Increasing Quality and Quantity of Plasmid Product

1.7.1. Anaerobiosis

To date, several studies have reported on different culture conditions that affect the level of DNA supercoiling (superhelical density) and as example of this is anaerobic shock or anaerobiosis (Cortassa and Aon, 1993; Hsieh et al., 1991; Dorman et al., 1988) which is said to increase DNA supercoiling (plasmid and chromosomal DNA). Anaerobiosis has been reported to increase DNA supercoiling due to its affects in decreasing the activity of toposiomerase I (Cortassa & Aon 1993) or increasing the [ATP] / [ADP] ratio (Hsieh et al., 1991 a & b).

1.7.2. 'Quiescent Cell Expression System'

1.7.2.1. Segregational Instability and Plasmid Multimerisation

Multicopy plasmids are distributed randomly at cell division and plasmid-free cells rarely arise providing the copy number remains high (Summers and Sherratt, 1988; Summers 1989). However, the presence of plasmid multimers can confuse plasmid-encoded control circuits, leading to copy number depression (Summers, 1998). Plasmid multimerisation reduces the number of independently segregating units leading to segregational instability of high copy number plasmids (Summers and Sherratt, 1988).

ColE1 and other naturally occurring plasmids are maintained stably because their multimers are resolved efficiently to monomers by site-specific recombination at the *cer* site (a binding site for a recombinase that reverses dimer and higher oligomer formation). These plasmids also contain at least four host-encoded proteins essential for recombination at *cer* (XerC, XerD, ArgR and PepA) (Summers and Sherratt, 1988).

The presence of multimers also triggers the expression of a small 70nt RNA called Rcd. Since it has been observed that Rcd is only expressed in multimer-containing cells, Chant et al., (2007), and Rowe and Summers, (1999), proposed that Rcd is a component of a checkpoint which delays cell division until Xer-cer multimer resolution is complete, thus avoiding the production of plasmid-free daughter cells. When Rcd is overexpressed for an extended period in a cell, it causes a distinct cell cycle arrest phenotype to occur. This proposition forms the basis of recent studies on the 'Quiescent cell expression system', initially through Rcd induction and overexpression.

1.7.2.2. Definition of a 'Quiescent Cell Expression System'

A fundamental problem associated with the expression of recombinant DNA in fast growing *E. coli* cells is that most of the energy and nutrient resources are directed toward biomass production. The metabolic stress imposed by hosting a multicopy plasmid reduces the growth rate and viability of the host cell. However, it is important to note that it is not the plasmid carriage itself that imposes a large metabolic burden but the expression of plasmid-borne genes (Kay et al., 2003).

A promising alternative to the traditional high-biomass fermentations for plasmid DNA production is to employ a 'Quiescent cell expression system', where the cells go into a stationary but metabolically active phase, devoting a large proportion of their resources into plasmid-borne gene expression rather than for growth and other functions during quiescence. Since the cells cease dividing, the redirection of energy and resources from biomass production may lead to an improved yield of

recombinant product without the need for additional nutrients in the media (Rowe and Summers, 1999).

1.8. An Overview of Viral and Non-Viral Vaccines

This section will give an outline of the current vectors used for gene therapy and vaccination.

1.8.1. Viral Vaccines

Historically important human viral vaccines have been produced in animals such as the vaccines for rabies (1885), polio (1955) and foot-and-mouth disease (1962) and more successful commercial products such as vaccines for Inactivated Influenza (Flushfield®, Fluzone®), Hepatitis A (Vaqta®, Havrix®) and Rotavirus (RotaTeq®).

1.8.2. Live and Inactivated Viral Vaccines

The two general classes of viral vaccines are considered either as non pathogenic vaccines derived from live, attenuated viruses (such as Adenovirus, Polio and Measles) and the vaccines derived from inactivated viruses such as Hepatitis A, Rabies and Foot-and-Mouth.

Other types of vaccines are subunit vaccines which contain a part of the virus either natural or synthetic (e.g. Influenza A and B and HBsAg derived from plasma of carriers) or recombinant viral proteins which have been expressed in bacteria, yeast, mammalian cells and viruses (Ada, 1997; Arvin and Greenberg, 2006; Aunins et al., 2000; Eliss, 1999).

1.8.3. Viral Vectors

The genetic modification of genomes from different virus families (Adenoviridae, Retroviridae, Herpesviridae) led to the development of gene therapy vectors with a similar capacity to infect cells or tissues as that of wild type viruses. Although the viral vectors have been engineered to remove their pathogenicity, their use for gene therapy still carries some safety concerns. Despite considerable progress over the past decade in the generation of reduced immunogenic response gene delivery systems,

the remaining immunogenicity of many gene therapy viral vectors is still the major hurdle, preventing their frequent application in clinical trials (Ritter et al., 2002).

1.8.4. Non-Viral Vectors and Delivery

Non-viral methods of gene delivery have generated a lot of attention in recent years as their use in gene therapy eliminates the concerns over the risks of immunogenicity that lingers with the use of viral vectors. The most commonly used methods of non-viral DNA transfer utilise naked plasmid DNA, liposomes or molecular conjugates as the non viral delivery vehicle.

1.8.4.1. Naked Plasmid DNA

Direct gene transfer with naked plasmid DNA is the simplest approach to non viral delivery systems. Intramuscular injections of a naked plasmid DNA can result in an immune response to the encoded viral antigen (Wolff et al., 1990). However, intravascular injection of plasmid DNA generally results in limited gene expression, leading to the development of rapid methods for plasmid delivery such as the use of a gene gun and electroporation. These methods may allow DNA to directly penetrate cell membrane and bypass endosome/lysosome, thus avoiding enzymatic degradation.

1.8.4.1.1. Gene gun

The gene gun method of delivery utilises naked plasmid DNA that has been adsorbed onto gold or tungsten microparticles into the target cells. The device uses helium as an accelerant and fires the dense gold or tungsten particles coated with the marker gene into the desired tissue. This method may also allow the DNA to directly reach the nucleus (Li and Huang, 2000).

1.8.4.1.2. Electroporation

The electroporation method for gene delivery involves the use of electrical pulses to permealise the cell membrane and permit the transfer of the naked plasmid DNA (Rols et al., 1998).

1.8.4.2. Liposomes

Liposomes are lipid bilayers entrapping a fraction of aqueous fluid and are the most widely studied and used of the non-viral vectors (Robbins et al., 1998, Templeton et al 1999). Anionic plasmid DNA will spontaneously associate to the external surface of cationic liposomes and these liposomes will interact with the cell membrane (Felgner et al., 1994). However there are some disadvantages as studies have shown that cationic lipids can cause direct cell toxicity, generate inflammatory responses on their own and/or increase the immune response to the co-administered adenovirus (Al Fasbender et al., 1998).

1.8.4.3. Molecular Conjugates

Examples of molecular conjugates used for in vitro non-viral gene delivery include using polymeric delivery systems such as polyethylenimine (PEI) (Kichler et al., 2001) and Chitosans. Like liposomes, polymer-DNA complexes (polyplexes) work on the basis that the generation of positively charged complex owing to electrostatic interaction of these cationic polymers with anionic DNA. The cationic polyplex can interact with the negatively charged cell surface to improve DNA uptake.

The efficiency of gene transfer by these applications is still very low (Ritter el al., 2002) and most require the use serum based media. Serum is a predominant cost factor for large-scale transient gene expression and is subject to batch to batch variation that makes standardisation of production protocols difficult. Additionally, the FDA has recommended against the use of bovine derived materials from countries where BSE has been diagnosed (such as New Zealand), but where possible it is preferable to avoid all use. Recent studies have shown some promise in serum-free, large scale transfections of CHO cells using PEI as the delivery system (Derouazi et al., 2004).

1.8.4.4. RNA/DNA Chimera

This relatively novel gene therapy approach is based on targeted correction of diseased mutations by mismatch repair. This involves the delivery into diseased cells of a short length, chimeric RNA:DNA oligonucleotide (Li and Huang, 2000). This molecule is designed to align with a specific target genetic sequence and triggers the endogenous DNA repair systems to alter the one target nucleotide in the desired way. Since the RNA:DNA oligomer is relatively small in size, it is readily able to form complexes with a variety of delivery or targeted systems.

1.8.4.5. Bacteriophage-based DNA Vaccines

Whole phage particles can be used to deliver vaccines in the form of immunogenic peptides attached to modified phage coat proteins or as delivery vehicles for DNA vaccines (Clark and March, 2004; March et al., 2006; McGrath and van Sindren, 2007). The gene encoding the antigen of choice is inserted into the genome of a bacteriophage vector and once taken up, the phage coat is removed. The vaccine DNA component is then expressed and vaccine protein is made within the host, leading to significant immune responses within a few weeks. Bacterial virus particles, such as filamentous bacteriophage, have been found to exhibit an adjuvant-like effect when applied in vaccination strategies (Willis et al., 1993). Since phages are unable to replicate in eukaryotic cells and do not carry antibiotic resistance genes, these potential barriers to regulatory approval are mitigated (March et al., 2006).

However it is important to note that one key disadvantage of using phages is the potential risk of phage infection/contamination to any process that relies on bacterial fermentation, such as for enzyme or antibiotic production. Steps to control phage contamination include the development and implementation of practical approaches at the production plant and laboratory level, such as adopting good aseptic techniques and hygiene measures and adequate HEPA air filtration (Moineau, 2005).

1.9. Introduction to Commonly Used Analytical Techniques & Methods

1.9.1. Agarose Gel Electrophoresis (AGE)

A common method used to determine plasmid form distribution is agarose gel electrophoresis. Plasmid DNA samples are applied to an agarose gel stained with ethidium bromide. Ethidium bromide is an intercalating agent commonly used as a nucleic acid stain in agarose gels and can be added to the gel prior to electrophoresis or afterwards before exposure to UV light.

AGE works by applying an electric field through a gel matrix. DNA molecules being negatively charged are drawn towards the anode positioned at the base of the gel. In doing so the molecule passes through the gel, the microstructure of which resembles a sieve. The migration distance of the DNA molecules depend on their chain length, with smaller DNA molecules moving through the gel matrix more readily than larger molecules and such that molecules of different length migrate as distinct bands.

After the gel is run, the gel is then scanned with a UV camera which detects and highlights the bands of DNA which the ethidium bromide has bound to. Each band represents a form of the plasmid and these forms can be determined by the use of a marker of known molecular weight and size distribution, if the molecular weight of a particular plasmid form is known.

If a sample of uncut plasmid DNA is run, the molecular weight of supercoiled species can be determined by comparison to a set of supercoiled standards of known molecular weight. However, this will depend to a certain extent on the linking number distribution of the uncut plasmid DNA sample. A wider linking number distribution conveys a more heterogeneous plasmid DNA sample. An alternative method of

determining molecular weight is to cut or nick the plasmid DNA samples using restriction enzymes and compare these with linear standards.

Although molecular weight distribution is an important application of AGE, for plasmid DNA processing AGE is primarily used for estimation of the relative proportions of different plasmid DNA species (SC, OC, linear) and thus it is important to correct for binding differences of ethidium bromide to different species (Bauer and Vinograd, 1968). Typically one or two prominent bands are observed when a small amount of plasmid DNA is applied to a gel, the most common one representing the SC form and the other representing the OC form. In some cases additional bands are observed representing the linear form, dimers of the SC, OC or linear forms and in some cases genomic DNA (gDNA).

Most gel cameras come with software that allows for densitometric analysis of the band intensity. This tool allows for the quantification of plasmid DNA in a band based on its intensity which can be used to determine SC amounts in a particular sample. It is noted by Projan et al., 1983, that there is an error of up to 20% associated with this method and for %SC analysis, a correction factor of 1.36 must first be applied to the SC intensity to account for the preference of ethidium bromide binding to the OC form.

1.9.2. Chloroquine Agarose Gel Electrophoresis

Chloroquine agarose gel electrophoresis has traditionally been used to estimate the linking number distribution of a supercoiled DNA sample and allowing the determination of the linking number, ΔLk. The addition of an optimal concentration of chloroquine, which like ethidium bromide is an intercalating agent, results in the supercoiled topoisomers present in a plasmid DNA sample to be resolved. When chloroquine binds to a close circular DNA molecule the DNA twist is reduced, thereby increasing the DNA writhe, since ΔLk cannot change providing the DNA duplex remains intact. This relaxation of the supercoiled topoisomers reduces their electrophoretic mobility causing the topoisomers to become separated into

a group of bands representing the number of superhelical turns of the plasmid (Keller, 1975).

In order to measure Δ*Lk*, the correct chloroquine concentration must be chosen to maximise resolution. In low concentrations of chloroquine the plasmid binds less chloroquine, resulting in some or all the topoisomers being resolved, the loss of some negative supercoils and slower migration. At high concentrations, more chloroquine is bound to the plasmid allowing the unwinding of all negative supercoils and this results in relaxed topoisomers that migrate with nicked circles ($\Delta Lk = 0$). A lower chloroquine concentration will result in topoisomers that are negatively supercoiled and a higher chloroquine concentration should resolve the topoisomers as positive supercoils (Gould et al., 1998). Generally, the more negatively supercoiled a plasmid DNA sample is, the more chloroquine is required to resolve these topoisomers as positive supercoils.

To identify the absolute value of the linking number difference, standards were run in parallel to the samples and analysed by the method described by Keller, (1975). Plasmid DNA standards differing in their average number of superhelical turns can be prepared by treating the plasmid DNA with an excess of a relaxing enzyme in the presence of increasing concentrations of chloroquine. Increasing the concentration of chloroquine causes the plasmid DNA to become increasingly more negatively supercoiled. After gel electrophoresis, the average value of the topological winding number will differ in each sample, depending on the amount of DNA binded by chloroquine. A relaxed sample of plasmid DNA where the plasmid DNA has been completely relaxed using topoisomerase I, and another consisting of untreated plasmid DNA should also be included as standards.

1.9.3. A260nm Spectrophotometry

UV spectrophotometry at 260nm is a common method used to determine the DNA yield, but more so as a rough indicator of the purity of the plasmid DNA sample. A sample of the resuspended plasmid is diluted in 1xTE

(Tris-HCL, EDTA) buffer (typically 10 or 20 times dilution so that the absorbance fall within the linear range of 0.1 – 1.0 AU). The sample of diluted plasmid is transferred to a quartz cuvette and is then analysed in a spectrophotometer at A260nm. This absorbance reading gives a value of plasmid concentration as 1AU is equal to 50μgml^{-1} double-stranded DNA. The same sample will then be analysed at wavelength 280nm. The ratio of A260 / A280 reveals information about the plasmid purity. If the ratio >1.8 the plasmid preparation may be contaminated with RNA. If the ratio is <1.8 the plasmid preparation is contaminated with protein.

The reason this technique is better suited for measuring DNA purity rather than quantity is that high levels of contaminants such as RNA and proteins can interfere with the reading, suggesting that there are higher amounts of DNA present than the actual amount.

Recently, a company under the name Nanodrop Technologies (Wilmington, DE, USA) has produced a few unique analytical instruments that quantify nucleic acids, proteins and fluorescent dyes based on UV/Vis spectrophotometry. Only 2μL of undiluted sample is required, enabling easy handling of DNA especially when only relatively small quantities are available.

1.9.5. Fluorescence-based Nucleic Acid Dyes

Nucleic acid binding dyes are widely used for DNA quantification, visualisation and purification. The most frequently used fluorescent dyes are PicoGreen, Hoechst, Ethidium bromide, DAPI and the dimeric cyanine nucleic acid dyes available from Molecular Probes such as YOYO-1 and YO-PRO-1.

1.9.5.1. PicoGreen Analysis

PicoGreen analysis is a sensitive assay for detecting double-stranded DNA (dsDNA) in solution with the use of PicoGreen dsDNA quantitation reagent (Molecular Probes, Inc., Eugene, OR, USA). This reagent binds to nucleic acids and fluoresces intensely upon doing so. The amount of

dsDNA in a sample is determined by comparing the fluorescent intensity to a high and low range standard obtained during each experiment. Calf thymus DNA or Lambda DNA are common DNAs used to obtain the standard linear graphs.

The advantage of using the PicoGreen dsDNA quantitation reagent is that it binds very rapidly to dsDNA and has a much broader DNA concentration range than compared to other nucleic acid dyes such as Hoechst 33342 (Molecular Probes, Inc., Eugene, OR, USA) or Ethidium bromide. However the fluorescence intensity emitted when PicoGreen is bound to dsDNA can be affected in conditions of high ionic strengths such as in certain buffers or samples containing high levels of salts like NaCl (Singer et al., 1997; O'Mahoney et al., 2007).

1.9.5.2. Modified PicoGreen Analysis

This assay is developed from a method used in Rock et al., (2003), to measure the supercoiled DNA ratio by measuring the fluorescence intensities of the plasmid DNA samples at pH 8 and then at pH 12.4. This assay judges the integrity of DNA by looking at the increasing rates of strand separation of dsDNA in a strong alkali for molecules with increased single stranded nicks.

The assay uses the well established principle of DNA denaturation in alkali, where DNA denaturation increases with the number of nicks. Hence the SC DNA ratio of a sample is reflected by the fluorescence intensity at pH12.4 (where the alkali has denatured the dsDNA into the OC form) divided by the intensity at pH 8 (where the PicoGreen reagent binds to mostly SC DNA). PicoGreen is stable over a wide range of pH (pH 8 – pH 13), so the fluorescence signals should not be affected by the changing pH conditions. NaOH (sodium hydroxide) is the common alkali added to the DNA samples in TE buffer solution to denature the dsDNA and changes the pH from pH 8 to pH12.4. In Rock et al., (2003), the PicoGreen-DNA samples were left to incubate at room temperature, protected from light for 5 minutes after 0.1M NaOH was added. Much

longer periods of incubation with NaOH is not recommended as in the presence of alkali, denaturation begins nicking the dsDNA until eventually even the OC-DNA denatures to become ssDNA.

Rock et al., (2003), did not look into the changing ionic strengths of the samples with the addition of NaOH and the effects this change may have on the fluorescence signal. It is mentioned in Singer et al., (1997), that in conditions of high ionic strength, the fluorescence intensity emitted from PicoGreen binding to dsDNA is indeed affected. This is supported by the findings reported by O'Mahoney et al., (2007), where overestimation of plasmid titres was observed using PicoGreen due to salt impurities present in the pDNA samples.

Some observations have shown that the 5 minute incubation period after the addition of NaOH may not be enough to fully denature the SC-DNA into OC-DNA and the concentration used 0.1M NaOH seems slightly low compared to the 0.2M used in Birnhoim and Doly's (1979) alkaline lysis method. The alternative is to heat denaturate the SC-DNA into OC-DNA, however, an optimum time of heating needs to be established as too much denaturation can cause the dsDNA to form ssDNA permanently.

1.9.5.3. Ethidium Bromide, Hoescht, YOYO-1 and YOPRO-1 Dyes

Ethidium bromide is DNA-intercalating agent commonly used as a nucleic acid stain in agarose gels, but it was traditionally used as a stand alone nucleic acid dye. The use of ethidium bromide as a stand alone fluorescent dye has been mostly replaced by newer commercial dyes where the dynamic range is much higher and in some cases much less toxic e.g. SYBR Gold (Molecular Probes, Inc., Eugene, OR, USA).

Fluorescent cationic Hoechst dyes such as Hoechst 33342 and Hoechst 33258 (both from Molecular Probes, Inc., Eugene, OR, USA) are cell permeable nucleic acid stains that enable the sensitive detection of DNA. Fluorescence of these two dyes in particular is enhanced upon binding to the minor groves of dsDNA at stretches of at least three AT base pairs

(Molecular Probes supplier's manual). Hoechst dye binding to DNA is strongly AT selective and not as selective as PicoGreen reagent as the Hoechst dyes also bind to RNA and ssDNA. Samples containing divalent cations such as Ca2+ or Mg2+ are deleterious to Hoechst dyes and EDTA is often required to chelate them (Dolezel, 1991).

YOYO-1 and YO-PRO-1 are among the many dimeric cyanine nucleic acid dyes available from Molecular Probes. They are both ultrasensitive fluorescent dyes with a strong binding affinity for nucleic acids. Similar to ethidium bromide, the fluorescent yield fromYOYO-1 and YO-PRO-1 are dependent on the DNA concentration and not on G+C or A+T content. The benefits of using cyanine dyes as opposed to traditional nucleic acid stains such as ethidium bromide or propidium iodide is that there is a higher affinity for nucleic acid binding with a lower inherent fluorescence in the absence of nucleic acids with the cyanine dyes.

However like PicoGreen, YOYO-1 also shows significant sensitivity to RNA and to high ionic strengths and YO-PRO-1 was found to give poor results when used with whole cell DNA analysis (Marie et al., 1996).

1.10. Principal Components Analysis

Plasmid-based gene expression systems are currently being widely investigated due to their potential for use in human therapy and vaccination. Currently such studies range from research at the cellular level to developing fermentation, recovery and purification strategies for large scale plasmid production. These tasks are aided by recent influxes of highly advanced and automated systems that have been introduced to increase the amount of data that can be captured and reinforce the analytical process. However for most of these tasks and systems the main drawback is the time-consuming data analysis required to comprehend the substantial amounts of data captured. Established statistical techniques such as Design of Experiments (DOE) and Principal Components Analysis are already widely in use and can aid analysis by

selecting the most interesting experimental variables or to extract useful data.

Principal Components Analysis (PCA) is a classical statistical method used in the analysis of large amounts of data by transforming a number of possibly correlated variables into a smaller number of uncorrelated variables called principal components (Wold et al., 1987; Wold and Sjöström, 1998). This is done by establishing relationships between the original variables, using a linear transform to reduce the dimensionality of the data matrix.

PCA relies upon an eigenvector decomposition of the covariance or correlation matrix of the variables. Data for processing by PCA are normally represented as an *m x n* matrix of *m* samples for each of the *n* variables. The goal of PCA is to reduce the dimensionality of the data matrix by finding *r* new variables, where *r* is less than *n*. Provided the columns of the *m x n* data matrix, X, have been mean centred (adjusted to have a zero mean by subtracting the original mean of each column) then the covariance matrix of X is defined as:

$$Cov(X) = \frac{X^T X}{m-1} \qquad [1]$$

If the columns of X have been auto scaled (adjusted to zero mean and unit variance by dividing each column by its standard deviation), then equation [1] gives the correlation matrix of X.

PCA decomposes the data matrix X into a series of paired column vectors, scores, t_i, and loadings, p_i, referred to as the principal components. The first principal component captures the most variance of the data, with the subsequent (i.e. second, third and so forth) principal components capturing the remaining variance. Generally, since the majority of the variance in the data matrix can be expressed by using only the first few

principal components, the remaining variance is represented in a residual error matrix, E, that accounts for background noise:

$$X = t_1 p_1^T + t_2 p_2^T + ... + t_k p_k^T + E \quad [2]$$

k must be less than or equal to the smaller dimension of X (the minimum of *m* and *n*).

The scores, t_i, contain information on how the samples relate to each other, whereas the loadings, p_i, contain information on how the variables relate to each other.

In the PCA decomposition, the loadings, p_i, are eigenvectors (λ_i) of the covariance matrix such that for each p_i:

$$Cov(X)\, p_i = \lambda_i p_i \quad [3]$$

The scores, t_i, form an orthogonal set ($t_i^T t_j = 0$ for $i \neq j$) whereas the loadings, p_i, are orthonormal ($p_i^T p_j = 0$ for $i \neq j$; $p_i^T p_j = 1$ for $i = j$). The scores, t_i, is the linear combination of the original X variables defined by p_i:

$$X p_i = t_i \quad [4]$$

The PCA scores and loading data is shown graphically and can be plotted in two or three dimensional graphs.

Existing reports on the use of PCA in scientific fields include using PCA to summarise microarray experiments that measure gene expression and clustering (Raychaudhuri et al., 2000, Yeung and Ruzzo, 2001); and the use of PCA in conjunction with chromatography to aid product analysis (Chandwani et al., 1997; Edwards-Parton et al., 2005 & 2008; Malmquist and Danielsson, 1994). These reports apply PCA on a large, genomic-

scale or are used to model scale up predictions of chromatographic separation.

The advantage of using PCA on data from host strain selection experiments for plasmid DNA production is that certain relationships and trends can be highlighted that may not have been initially apparent through traditionally used techniques and graphical presentations. PCA also has the potential to aid and ease data analysis when the data increases to an unmanageable amount to manually analyse efficiently.

1.11. Research Objectives

The main objective of this study is to investigate a rational approach for the processing of plasmids from host strain selection to large scale fermentation. In particular this project will investigate initial host strain selection and process conditions and the impact of these choices on supercoiled plasmid DNA production.

The strain of *E. coli* used to propagate a plasmid can have sizeable influence on the quantity and quality of the purified plasmid DNA. Consequently it would be significantly beneficial if a systematic approach is conducted to obtain the data of various *E. coli* strains at shake flask scale such as growth profiles, plasmid copy number, supercoiled DNA amount and plasmid purity, and to then see whether there are any similarities in the genotype of better performing strains. From these results, two of the better performing host-plasmid combinations will be taken forward for bioreactor studies to help determine the effects of scale up on the strain performance.

In line with investigating different strategies that may improve supercoiled plasmid DNA production, a part of this research will include implementing anaerobiosis, to test if this approach can improve plasmid DNA supercoiling at fermentation scale under the conditions set. If so, this technique may be beneficial in improving upstream supercoiled pDNA yields prior to cell harvest and recovery.

The final part of this research aims to investigate whether a 'Quiescent cell system' (a state where chromosomal replication and expression is temporarily shut down but residual proteins remain metabolically active, involving indole induction of *hns* mutated *E. coli* cells) works for plasmid production as well as it has been reported for plasmid-borne protein expression. The main objective is to establish whether or not the yield of plasmid DNA increases with a 'Quiescent cell system', or if plasmid production is inhibited with the onset of quiescence. If the induction is

deemed beneficial then optimum conditions such as inducer concentration, induction time and scalability will need to be established.

2. Materials and Methods

2. Materials and Methods

2.1. Materials

2.1.1. Plasmids

The plasmid vectors pSVβ, gWiz and pQR150 were chosen for this project as they were previously used in similar studies within the Department of Biochemical Engineering, UCL in recent years (Cooke et al., 2004; Kay et al., 2003; Kong et al., 2006; O'Kennedy et al., 2000 and 2003). Characteristics of these three plasmids can also be found in Chapter 3, Table 3.2.3. Plasmids pGlo and pUC18 were used in for GFP expression in Chapter 6 and temperature amplification studies in Chapter 4 respectively.

2.1.1.1 pSVβ

pSV-β-Galactosidase plasmid vector (referred to in this case as just pSVβ) (Promega) is a 6.9kb pUC-based plasmid encoding the β-gal gene and is resistant to ampicillin. pSVβ is a modification of pRSV-β-Gal with SV40 and pUC18 sequences substituted for RSV and pBR322 sequences. pSVβ was obtained from Professor John Ward's culture collection (purchased from Promega) and was originally transformed into DH5α and stored in 20% glycerol at -80°C. This plasmid contains the pMB1 origin of replication.

2.1.1.2. gWiz GFP

gWiz GFP (Aldevron, Fargo, ND, USA) is one of a series of plasmids that have been engineered by Gene Therapy Systems (San Diego, CA) to produce high levels of transient gene expression in a wide variety of mammalian cells and tissues as well as high yield of plasmid production in *E. coli*. This gWiz vector contains a proprietary modified promoter followed by Intron A from the human cytomegalovirus (CMV) immediate-early (IE) gene. This 5.8kb plasmid of pUC origin is resistant to Kanamycin. gWiz was originally transformed into DH1 and stored as cell

paste (Terrific Broth) at -80°C. This plasmid contains the pMB1 origin of replication.

2.1.1.3. pQR150

pQR150 is a 20kb plasmid derived from pBGS18, itself a derivative of pBS8 but without the duplicate *Acc*I site (Spratt et al., 1986). pQR150 was constructed by Professor John Ward and is described in Jackson et al., (1995) and Sheridan et al., (1997). The benzoate *meta*-cleavage pathway of *Pseudomonas putida* encoding the catechol 2,3-dioxygenase activity was cloned into pBGS18 to yield pQR150. This plasmid is resistant to kanamycin. This plasmid was originally transformed into JM107 and stored in 20% glycerol at -80°C. This plasmid contains the pMB1 origin of replication.

2.1.1.4. pGlo

pGlo (Bio-Rad, Hercules, CA, USA) is a 5.4kb pUC-derived plasmid that expresses GFP in *E. coli* upon arabinose induction (Mosher, 2002). This plasmid contains the pMB1 origin of replication. The promoter sequence controlling GFP expression usually regulates the transcription from the *ara* operon (P_{BAD}). No (or very little) transcription occurs from the *ara* operon unless the activator protein AraC first binds to a specific activator sequence (*araI*) just upstream of P_{BAD}. This only occurs in the presence of the sugar arabinose. The *bla* gene coding for beta-lactamase confers ampicillin resistance. This plasmid was purchased from Bio-Rad and stored at -80°C.

2.1.1.5. pUC18 and pUC19

pUC18 and pUC19 vectors are small, high copy number, *E. coli* plasmids, 2686 bp in length. pUC18 and pUC19 are identical except that they contain multiple cloning sites arranged in opposite orientations. These plasmids contain the pMB1 origin of replication. The high copy number of pUC plasmids is a result of the lack of the *rop* gene and a single point mutation in *rep* of pMB1. The *bla* gene coding for beta-lactamase confers

ampicillin resistance. pUC18 and pUC19 were purchased from Stratagene and stored at -80℃.

The maps of the above plasmids can be found in Appendices section A.1.

2.1.2. *E. coli* Strains Used in this Study

Different strains of *E. coli* to be experimented on initially before the best performing strains are taken further to fermentation scale. Since there are thousands of potential *E. coli* strains to select from, these following strains were chosen based on their commercial availability, common usage in laboratories or to establish their potential as efficient hosts. The latter criterion is aimed mainly at older strains used in the early era of microbiology that have since been overlooked. Most of the chosen strains are derivatives of *E. coli* K-12 since the safety of K-12 strains have been the most thoroughly investigated and hence is preferred by biosafety regulatory bodies, compared to others such as derivatives of *E. coli* B (Bachman, 1996). By commercially available, it is implying that a company owns, developed and or patented the strain. Strains that are sold by some companies may not necessarily be commercial strains if they are "old" strains and are not proprietary; some strains such as DH5α and DH10β were originally introduced as non-commercial strains but were later sold to a company and are now owned by Invitrogen.

The strains used in Chapter 3 are listed in Tables 3.2.1. and 3.2.2. In Chapter 4, another wild type *E. coli* K-12 strain, W1485, ATCC 12435 (λ-, F+), (Jensen et al., 1993) was used. This strain is the parent of W3110 and MG1655.

The strains were either purchased commercially, or kindly obtained from Professor John Ward (UCL Research Department of Structural and Molecular Biology, University College London, London UK). The strains obtained from Professor Ward were originally stored in 20% glycerol at -80℃, either from mixing broth cultures (LB or Nutrient Broth) 50:50 with 40% glycerol (in water) solution, or by flooding an agar plate with sterile

20% glycerol (in water) solution and suspending the colonies. Commercial strains were purchased from Invitrogen, Stratagene or New England Biolabs in the form of competent cells and stored at -80°C.

2.1.2.1. 'Quiescent' Cell *E. coli* Strain

Hereby known as Quiescent cells or Q-cells, a *hns* mutated *E. coli* K-12 strain W3110 *hns* 205::Tn*10* was kindly provided by Dr David Summers from the Department of Genetics at Cambridge University. The *hns*-205::Tn*10* deriviative of W3110 can be constructed by P1 transduction from strain GM230, selecting for mucoid, tetracycline resistant colonies. These *hns* mutant cells were then transformed singly with gWiz, and pGlo using the method described in section 2.2.1.

2.1.3. Culture Media

All media and stock solutions were prepared in reverse osmosis (RO) water.

Table 2.1.3.1. Nutrient Broth CM0001 powder (Oxoid) formula:

Lab-Lemco (beef extract) powder	1g/L
Yeast Extract	2g/L
Peptone	5g/L
Sodium Chloride	5g/L

13g of the nutrient broth powder was dissolved in 1L of RO water and mixed thoroughly before being autoclaved at 121℃ for 15 minutes. The culture pH was around pH 7.4 ± 0.2.

Table 2.1.3.2. SDCAS medium ingredient list

Component	**Amount per litre**
D-glucose	10.0g
Magnesium Sulphate Heptahydrate	1.2g
Ammonium Sulphate	4.0g
Potassium Dihydrogen Orthoposphate	13.3g
Casamino Acids	10.0g
Citric acid	1.7g
EDTA sodium salt	8.4mg
Cobalt chloride hexahydrate	2.5mg
Manganese sulphate	15.0mg
Copper Sulphate Dihydrate	1.5mg
Boric Acid	3.0mg
Sodium Molybdate Dihydrate	2.5mg
Zinc Chloride	13.0mg
Iron III citrate	100.0mg
Thiamine Hydrochloride	4.5mg

2.1.3.2.1. SDCAS medium preparation

SDCAS medium was prepared as follows in order to prevent precipitation or the breakdown of amino acids when steam sterilised in the presence of high glucose concentration (Maillard reaction).

1. Trace Elements, per litre: EDTA (8.4g), Cobalt chloride (2.5g), manganese sulphate (15g), copper sulphate (1.5g), boric acid (3g), sodium molybdate (2.5g), zinc chloride (13g) were prepared and autoclaved.
2. Stock solutions of Iron III citrate (4g/L), D-glucose (200g/L) and Magnesium Sulphate (48g/L) and Casamino acids (200g/L) were prepared and autoclaved separately.
3. Thiamine Hydrochloride (4.5g/L) was prepared and filter-sterilised and stored in the fridge.
4. Base medium, per litre: Ammonium Sulphate (4g), Potassium Dihydrogen orthophosphate (13.3g) and citric acid (1.7g) was prepared and adjusted to pH 7.0 with the addition of 4M NaOH and before being autoclaved or steam sterilised in situ in the fermenter.

The other components were sterilised separately and then added to the shake flask or fermenter containing the base medium before inoculation, to obtain the final concentrations of each component as listed above in the ingredient list.

Table 2.1.3.3. LB broth powder (Sigma-Aldrich) formula:

Yeast Extract	5g/L
Tryptone (Pancreatic Digest of Casein)	10g/L
Sodium Chloride	5g/L

20g of the nutrient broth powder was dissolved in 1L of RO water and mixed thoroughly before being autoclaved at 121℃ for 20 minutes. The culture pH was around pH 7.4 ± 0.2.

Table 2.1.3.4. Terrific Broth (TB) (Sigma-Aldrich) formula:

Yeast Extract	24g/L
Tryptone (Pancreatic Digest of Casein)	12g/L
K_2HPO_4	9.4g/L
KH_2PO_4	2.2g/L

48.2g of TB powder was dissolved in 1L of distilled water. 8mL of glycerol (Sigma-Aldrich) was added to this mixture before it was autoclaved for 20 minutes at 121 ℃.

2.1.4. Agar Plates

Table 2.1.4.1. Nutrient Agar CM0003 powder (Oxoid) formula:

Lab-Lemco (beef extract) powder	1g/L
Yeast Extract	2g/L
Peptone	5g/L
Sodium Chloride	5g/L
Agar	15g/L

28g of the nutrient agar powder was dissolved in 1L of RO water and mixed thoroughly before being autoclaved at 121℃ for 20 minutes. The culture pH was around pH 7.4 ± 0.2.

2.1.5. Antibiotic Stock Solutions

Ampicillin (Sigma-Aldrich) stock solutions were prepared and filter-sterilised at a stock concentration of 50mg/mL in RO water and this amount was diluted to a working concentration of 100µg/mL for cultures. Kanamycin (Sigma-Aldrich) stock solution was prepared in the same way except the stock solution concentration for Kanamycin is 10mg/mL and the working concentration for cultures is 50µg/mL. Stock solutions of both

antibiotics were transferred to single-use eppendorf tubes and stored at -20℃.

2.1.6. Buffers

The following buffers were used throughout the experiments conducted:

1. TE Buffer (10x concentrated) (1.0 M Tris-HCl, pH approx. 8.0, containing 0.1 M EDTA) (Sigma-Aldrich Company Ltd., Gillingham, Dorset, UK). Working solutions of 1xTE buffer were used.
2. TAE buffer (10x) (0.4 M TRIS acetate, pH 8.3±0.2 (25℃); 0.01 M EDTA.Na_2) (Sigma-Aldrich). Working concentrations of 1xTAE were used for gel electrophoresis.
3. TBE buffer (5x or 10x) (0.445 M TRIS borate, pH 8.4±0.1 (25℃); 0.01 M EDTA.Na_2) (Sigma-Aldrich). Working concentrations of 0.5xTBE were used for gel electrophoresis.
4. Gel loading buffer (Sigma-Aldrich) consists of Bromophenol blue (0.05% w/v), sucrose (40% w/v), EDTA (0.1 M, pH 8.0), and SDS (0.5% w/v).

2.1.7. Qiagen Buffers for Qiaprep Mini-, Midi- and Maxi- Kits

1. Buffer P1 resuspension buffer added with 100µg/mL RNase A and Lyse Blue solution (50mM Tris-HCl pH 8.0, 10mM EDTA.Na_2). Stored at 4℃.
2. Buffer P2 contains 1% (w/v) sodium dodecyl sulphate (SDS) and 200mM sodium hydroxide. Stored at room temperature.
3. Buffer N3 contains 25-50% guanidine hydrochloride and 10-25% acetic acid. Stored at room temperature.
4. Buffer PB contains 25-50% guanidine hydrochloride and 25-50% isopropanol. Stored at room temperature.
5. Buffer PE is a wash buffer added with 70% ethanol. Stored at room temperature.
6. Buffer EB is an elution buffer (10mM Tris-HCL, pH 8.5). Stored at room temperature.
7. Buffer P3 is a neutralisation buffer containing 3M potassium acetate, pH 5.5. Stored at room temperature.

8. Buffer QBT is an equilibration buffer containing 750mM NaCl; 50mM MOPS, pH 7.0; 15% isopropanol (v/v); 0.15% Triton X-100 (v/v). Stored at room temperature.
9. Buffer QC is a wash buffer containing 1M NaCl; 50mM MOPS, pH 7.0; 15% isopropanol (v/v). Stored at room temperature.
10. Buffer QF is an elution buffer containing 1.25M NaCl; 50mM Tris-Cl, pH 8.5; 15% isopropanol (v/v). Stored at room temperature.
11. Buffer TE in this case is an elution buffer consisting of 10mM Tris-Cl, pH 8.0; 1Mm EDTA. Stored at room temperature.

2.1.8. Restriction Enzymes Used For Nicking/Cutting dsDNA to Open Circular (OC) or Linear Forms

1. *Hind III* (New England Biolabs, Hitchin, UK)
2. *EcoRI* (New England Biolabs)
3. *NtBstNBI* (New England Biolabs)
4. *BamHI* (New England Biolabs)

2.1.9. Agarose Gel Electrophoresis Molecular Weight Markers

1. λBstE II (New England Biolabs)
2. λHind III (New England Biolabs)
3. DNA ladder, supercoiled (Sigma-Aldrich)
4. Hyperladder-1 (Bioline, London, UK)

Markers 1 and 2 need to be heated at 60℃ for 3 minutes prior to use/placing on ice to dissociate the 12-base sticky ends of the cohesive ends. Without heating λHind III, the 23kb band and the 4.3kb band anneal by H-bonding between the complementary bases to form a 27kb band. Without heating λBstE II, the 8.4kb band and the 5.6kb band anneal to form a 14kb band. This results in a decrease in intensity of the 8.4 and the 5.7kb bands plus the appearance of a 14kb band.

Although both markers were heated following the manufacturer's instructions, on a few occasions for λBstE II this did not work and a 14kb band was observed (see Figure 2.1.9.1.).

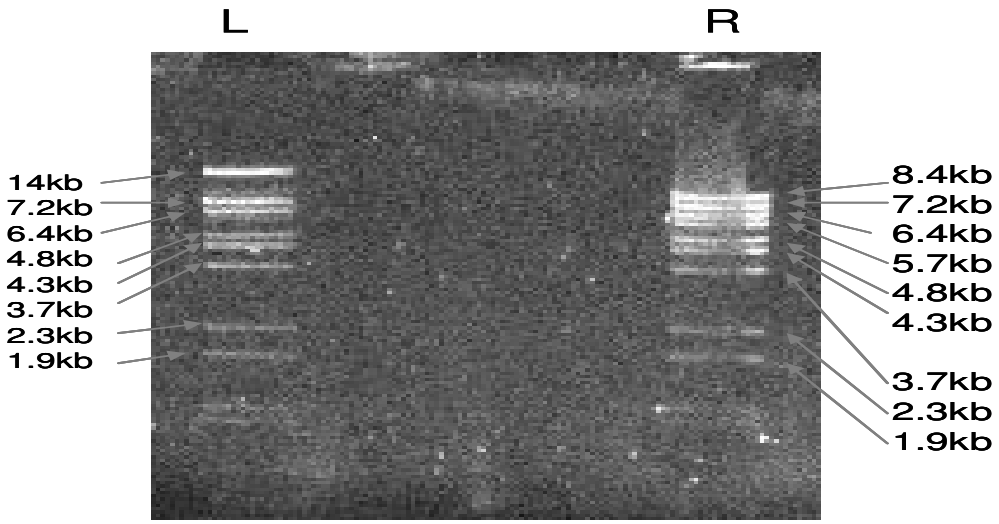


Figure 2.1.9.1. λBstE II marker with correct weightings (R) and the version with annealed 8.4kb and 5.7kb bands due to insufficient heating prior to use (L).

2.2. Methods

2.2.1. Transformation

Non-competent strains obtained from Professor John Ward were subsequently grown up into working vials of competent cells in 75mM $CaCl_2$ 20%(v/v) glycerol, before they were used for transformations; with plasmids pSVβ, gWiz, pQR150 and pUC18 using the standard calcium chloride technique (Sambrook et al., 1989). The remaining competent cells were frozen at -80℃ and the newly transformed cells were transferred to 20% (v/v) glycerol solution and stored at -20℃. The same transformation procedure was carried out on commercially purchased competent strains.

2.2.2. Plate Cultures

The glycerol transformed cells were thawed and then streaked onto nutrient agar plates (Nutrient Agar CM0003, Oxoid, Basingstoke, UK) containing the appropriate antibiotic in concentrations that were sufficient to select against the host (100μg/mL ampicillin or 50μg/mL kanamycin).

2.2.3. 10mL Cultures

10mL cultures containing media and the appropriate antibiotic were grown in 50mL Falcon tubes for 16.5 hours at 37℃, 250rpm. These cultures were inoculated with single picked colony from a fresh grown plate culture.

2.2.4. 100mL Cultures

100mL working volume shake flask cultures containing media and the appropriate antibiotic were grown in 500mL shake flask for 8 hours at 37℃, 250rpm. These cultures were inoculated with a seed 10mL falcon tube culture that had been left to grow overnight at 37℃, 250rpm, prior to the early morning shake flask inoculation. These seed cultures were prepared as in 2.2.3.

2.2.5. Fermentation Cultures

5L working volume fermentation cultures were grown in a 7L fermenter (LH Series, Bioprocess Engineering Services) containing 4.5L SDCAS medium. A 500mL SDCAS seed culture grown for 16 hours at 37℃, 250rpm in a baffled 2L shake flask was used to inoculate the fermentation media. This seed culture was inoculated with a 10mL SDCAS pre-seed culture grown for 8 hours at 37℃, 250rpm and was inoculated by a single colony from a freshly grown plate. Initial fermentation conditions were set at 37℃, 700rpm agitation, 1vvm air inlet, at pH 7.0. The temperature and pH were maintained automatically by the PID controls, with the pH was adjusted automatically with additions of 2M phosphoric acid or 5M ammonium hydroxide. Antifoam control was also automatically controlled and when necessary, neat polypropylene glycol (PPG) was added. The agitation was initially set at 700rpm, with cascade control linked to the DOT settings so that once the DOT drops below 30%, the stirrer speed will automatically increase to maintain a DOT level above 30%.

2L LB media fermentations were conducted in a 2L bench-top fermenter (LH Series 210, Bioprocess Engineering Services) with a 1.5L working volume. 1.4L of LB media was autoclaved in the 2L fermenter before kanamycin (50µg/mL) was added to the media shortly before inoculation

with 100mL seed inoculum. The cultures were maintained at 37ºC, at an agitation rate of 600rpm. The airflow was maintained at 1.5L/min and the pH was controlled at pH 7.0 ±0.2 (10% HCL or 2M NaOH addition as required). Antifoam was added manually when needed.

The exit gas data for all fermentations was monitored by an online mass spectrometer (Prima 600, VG-gas Analysis, Winsford, Cheshire, UK) connected to the fermenter controls. Online data was logged by Propack data logging and acquisition software (Acquisition Systems, Guildford, Surrey), Bioview Software (Adaptive Biosystems, Watford, UK) or BioXpert data logging and acquisition software.

OUR and CER (both in mmol L^{-1} h^{-1}) were calculated using the following equations:

$$Gas_{in} \cdot \%N_{2in} = Gas_{out} \cdot \%N_{2out} \quad [1]$$

$$Gas_{in} = Q \cdot \frac{P}{RT} \quad [2]$$

$$OUR = Gas_{in} \cdot \%O_{2in} - Gas_{out} \cdot \%O_{2out} \quad [3]$$

$$CER = Gas_{out} \cdot \%CO_{2out} - Gas_{in} \cdot \%CO_{2in} \quad [4]$$

where

Gas_{in}	=	Air flow into the fermenter (mmol L^{-1} h^{-1})
Gas_{out}	=	Air flow out of the fermenter (mmol L^{-1} h^{-1})
$\%N_{2in}$	=	Inlet concentration of nitrogen (%mol)
$\%N_{2out}$	=	Outlet concentration of nitrogen (%mol)
Q	=	Air flow rate (L h^{-1})
P	=	Pressure (atm)
R	=	Ideal gas constant
T	=	Temperature (ºK)
$\%O_{2in}$	=	Inlet concentration of oxygen (%mol)
$\%O_{2out}$	=	Outlet concentration of oxygen (%mol)
$\%CO_{2out}$	=	Inlet concentration of carbon dioxide (%mol)
$\%CO_{2in}$	=	Outlet concentration of carbon dioxide (%mol)

2.2.6. Anaerobic Transition

For the 7L batch fermentation cultures where anaerobiosis is employed, the fermentation conditions were initially set out with the same conditions as stated in 2.2.5. with the same seed culture volume. After 4 hours of normal growth (during mid-exponential phase) the air-inlet feed was replaced with pure nitrogen (1vvm) and the DOT cascade control was turned off. The %DOT values dropped almost immediately to 0% after the switch to nitrogen. The mass spectrometer readings of the outlet air were checked after this transition to ensure that only nitrogen is being fed into the fermenter.

Culture samples taken from the fermenter were kept anaerobic until after the cells were lysed. This was done by continually sparging nitrogen into the sample tubes during handling before carefully sealing the tubes.

2.2.7. 'Quiescent Cell' Studies

2.2.7.1. 'Quiescent Cell' Indole Preparation and Induction

Indole (Sigma-Aldrich) was dissolved in absolute ethanol (Sigma-Aldrich) and this solution is kept at 4℃ only for a week before being discarded to minimise instability of the indole solution.

W3110 *hns* 205:: Tn10 cells containing one of the plasmids (gWiz, or pGlo) were induced into quiescence by adding indole solution of the desired concentration to the culture at a specific time of growth. The addition of indole solution must be carefully conducted since indole is relative volatile and must remain in solution. The tip of the pipette should be submerged beneath the media surface during the ejection of the indole and the shake flask or pipette should be swirled to mix the indole thoroughly during the addition. A cloudy white precipitate should be observed during the ejection of the indole into the culture which will disappear after a few swirls. The culture will have an aromatic smell after the addition of indole.

2.2.7.2. 'Quiescent Cell' Culture Conditions

Cultures were grown on LB broth (Sigma-Aldrich) at either 400mL or 100mL working volume scale (in 2L or 500mL shake flasks respectively), inoculated with overnight cultures grown at 37℃, 250rpm.

A 2L W3110 *hns* 205:: Tn10 gWiz fermentation (1.5L working volume) run was conducted as described in Chapter 2 section (2.2.5.). 3mM indole was used to induce quiescence after 5 hours of growth.

When control cultures were carried out to represent non-induction (0mM indole), the addition of equal amounts of absolute ethanol was used to replace indole. Bacterial GFP expression from the plasmid pGlo was induced by adding 0.25% (w/v) L-(+)-arabinose (Sigma-Aldrich) to cultures (Mosher, 2002).

2.2.8. Optical Density (OD_{600}) Measurements

Samples were taken hourly during the growth profile analysis of the shake flask cultures and these samples were divided up to enable various measurements to be made. One of these measurements was the optical density (OD) of the cultures at a wavelength of 600nm using a spectrophotometer (Kontron Uvikon Spectrophotometer) or the Amersham 1100 Spectrophotometer (after the Kontron spectrophotometer broke down and was replaced).

For pSVβ cultures in Chapter 3, the maximum specific growth rates (μmax) from 10mL cultures were calculated using measurements taken from Tecan Safire^2 plate reader at 600nm. 200μL of sample were aliquoted into microwell plates and measured hourly for 6 hours to calculate the growth rates and μmax.

2.2.9. Dry and Wet Cell Weights

For dry cell weight determination, duplicate 10mL cultures were spun at 3500g in a GS6R centrifuge (Beckman, High Wycombe, Bucks.) for 10 minutes at 4℃ in pre-heated and pre-weighed plastic Falcon tubes. These

tubes were pre-heated to evaporate any residual moisture in the plastic that would otherwise affect the weight readings after drying. The tubes were left to cool down to room temperature before the cultures were aliquoted into them. The cell pellets were then washed in sterile RO water and then dried for 48 hours in a 100℃ oven to obtain a constant weight reading.

For wet cell weight determination, 1.5mL samples of culture broth were transferred to pre-weighed, sterile 2.2mL microtubes. These were spun at 13000g in a bench top centrifuge (Heraeus Sepatech Biofuge 13) for 5 minutes. The supernatant was then carefully decanted off before the tubes were re-weighed. The weights of the pellets were determined by the difference between the initial weight of the tubes and the final tube-pellet weight.

2.2.10. Plasmid Extraction DNA and Purification

Plasmid DNA samples were extracted and purified using Qiagen Mini-, Midi or Maxi- prep kits (Qiagen Ltd. West Sussex, UK) according to the protocols provided with these kits. A list of buffers and solutions and their compositions used can be found in section 2.1.7.

2.2.10.1. Qiagen Miniprep Procedure

1mL of culture sample is spun down at 13,000rpm for 3 minutes. The cell pellet is then resuspended in 250µL Buffer P1. 250µL Buffer P2 is then added and mixed thoroughly by inverted the eppendorf tube 4 times. When the LyseBlue reagent in the P2 buffer turns homogenously blue, 350µL of Buffer N3 is added and the contents of the tube are then mixed immediately and thoroughly by inverted 4-6 times. The mixture is then centrifuged for 10 minutes at 13,000rpm using a benchtop centrifuge (Heraeus Sepatech Biofuge 13). The supernatant is then transferred by decanting into a Qiaprep spin column and the contents are transferred through by centrifuging for 1 minute at 13,000rpm, discarding the flow-through after the spin (the plasmid is retained by binding to the column membrane). 0.5mL Buffer PB is then added to the spin column to wash

the column by centrifugation for 1 minute at 13,000rpm, again the flow-through is discarded. 0.75mL of Buffer PE is added to the column and the column is washed by centrifugation for 1 minute. After this flow-though is discarded, the column is centrifuged again for another minute to removed any residual wash buffer. The column is then placed in a fresh eppendorf tube and 50µL of EB buffer is added to the column. This is left to stand for 1 minute before centrifuging for another 1 minute to elute the DNA which is collected in the clean eppendorf tube. For high cell density cultures, i.e. samples from fermentations, lysate from one of the triplicate samples taken was diluted 1:1 with sterile RO water before it was transferred to the spin columns. This procedure was performed to check whether the miniprep columns had become saturated and to ensure that the amount of plasmid DNA retained was accurate.

2.2.10.2. Qiagen Hispeed Midi- and Maxi- Prep Procedure

Qiagen Hispeed Midi- and Maxi- prep kits follow the same procedure except a higher volume of buffers is needed during the Maxi-prep since there is more culture volume initially. The Maxi-prep buffer volume is bracketed. A known volume of frozen culture is thawed and then spun down in a Falcon tube using a GS6R centrifuge (Beckman, High Wycombe, Bucks.) for 10 minutes at 4℃. The cell pellet is then resuspended in 6mL (10mL) Buffer P1 before 6mL (10mL) Buffer P2 is added. The mixture is then inverted several times before being left to stand at room temperature for 5 minutes. 6mL (10mL) of chilled Buffer P3 is added to the lysate and mixed by inverting 4-6 times. The lysate is then immediately transferred to a QIAfilter cartridge barrel and left to incubate at room temperature for 10 minutes. The cap from the QIAfilter outlet nozzle was removed and the contents of the QIAfilter cartridge were expelled into a previously equilibrated tip (equilibrated by applying 4mL or 10mL Buffer QBT and allowing the column to empty by gravity flow) using a plunger. The cleared lysate was then left to enter and clear the resin by gravity flow. 20mL (60mL) Buffer QC was added to the HiSpeed tip and allowed to flow through by gravity flow. The DNA was eluted with 5mL (15mL) Buffer QF into a 50mL Falcon tube and the DNA was precipitated

by adding 3.5mL (10.5mL) room temperature isopropanol. This mixture was left to stand for 5 minutes at room temperature. The eluate-isopropanol mixture was then transferred into a 20mL (30mL) syringe and the mixture was expelled through a QIAprecipitator Module at the end of the syringe using a plunger. This flow-through mixture is discarded. The QIAprecipitator is then removed before the plunger is pulled out of the syringe to allow 2mL 70% ethanol to be added to the syringe. After re-applying the QIAprecipitator, the DNA is then ethanol-washed by reinserting the plunger to push the ethanol through the QIAprecipitator. This last step is repeated again but instead of adding ethanol, nothing is added to allow the membrane to dry by pressing air through it. The outlet nozzle of the QIAprecipitator is dried with a clean tissue to prevent ethanol carryover. Transfer the QIAprecipitor to the end of a new 5mL syringe and hold this syringe over a 1.5mL or 2.2mL fresh eppendorf tube. Add 1mL of TE buffer to the new syringe and insert the plunger to elute the DNA from the QIAprecipitator into the eppendorf tube. Remove the QIAprecipitator from the 5mL syringe, pull out the plunger, reattach the QIAprecipitator and transfer the eluate from the eppendorf back into the syringe. Elute for the second time to ensure the maximum amount of DNA in the QIAprecipitator is solubilised and recovered.

2.2.10.3. Qiagen Miniprep kits vs. Isopropanol Precipation Method

For larger plasmids like the 20kb pQR150, slight nicking of SC plasmids to the OC form has been observed possibly due to the stress of the DNA passing through the Qiagen spin column matrix structure. This observation is noted when Qiagen miniprepped samples of pQR150 were compared with isopropanol-precipitated DNA samples of the same plasmid from the same culture. The isopropanol precipitation technique applied was the same as the Qiagen buffer stages up to just before the sample is added to the spin column, where instead the isopropanol was added to the sample and then ice chilled before an ethanol wash was conducted to remove the excess salts. However this isopropanol precipitation method yielded a less pure sample of DNA and the high ionic strength of the samples led to problems with the PicoGreen analysis that is sensitive to high ionic

conditions. Slight nicking of the SC pQR150 DNA using Qiagen Hi-speed Midi- and Maxi-kits was not observed but this is likely due to the fact that the Midi- and Maxi-kits used involved syringe filters rather than a spin column.

2.2.11. Plasmid DNA Yield Determination

The concentration of pDNA from a purified sample was determined using a Nanodrop spectrophotometer ND-1000 (NanoDrop, Willington, DE, USA) which measures the nucleic acid concentration by an absorbance of 260nm. For nucleic acid quantification, a form of the Beer-Lambert equation is used to correlate the calculated absorbance with concentration:

$$c = \frac{A \cdot e}{b} \qquad [1]$$

Where

c is the nucleic acid concentration in ng μL^{-1}

A is the absorbance in AU

e is the wavelength-dependent extinction coefficient in ng.cm μL^{-1}

b is the path length in cm

The generally accepted extinction coefficients for nucleic acids are double-stranded DNA: 50; single-stranded DNA: 33 and RNA: 40.

The total DNA yield was then determined by multiplying the DNA concentration by the amount of eluant per 1 mL sample.

2.2.12. Purity

The purity of a sample was determined by the A260nm/A280nm ratio using the Nanodrop device and to a lesser extent through agarose gel electrophoresis. The A260nm/A280nm ratio can tell if a sample is pure or contaminated by RNA or protein. A pure sample of DNA should have an

A260nm/A280nm ratio of 1.8. If the ratio is less than 1.8, then the sample is contaminated with protein and if the ratio is greater than 1.8, then the contaminant is RNA.

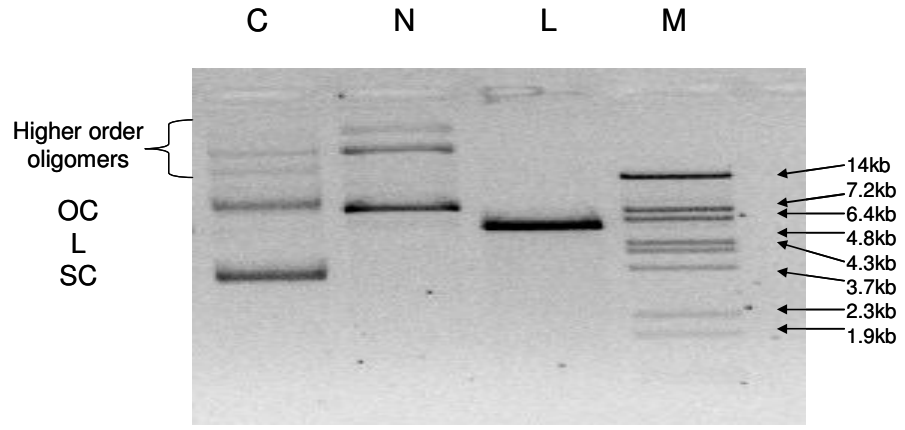
2.2.13. % Cells Containing Plasmid

The level of plasmid loss at one time point was measured by conducting serial dilutions of cultures after a specific period of growth on selective agar plates (with antibiotic) and non-selective (no antibiotic addition) and then comparing the number of colonies on each plate. A 'stable' strain, i.e. one that retains the plasmid, would provide approximately the same number of colonies on both the selective and non-selective plates.

2.2.14. Nicking and Linearising dsDNA

The band(s) on a gel that corresponds to the SC form was determined by nicking dsDNA to OC form or linear form using restriction enzymes and then running these nicked and linear samples on a separate agarose gel before comparing these forms to a control unrestricted sample. For nicking to OC form, 2µL *NtBstNBI* was added to 8µL dsDNA with 8µL EB buffer and 2µL 10x Buffer RB or NEB buffer 3 and this mixture was incubated at 55℃ for 1 hour. For linear DNA samples, the procedure was the same except 2µL of *HindIII* or 2µL *BamHI* was added instead of *NtBstNBI* (for BamHI, 2µL NEB buffer 4 replaced the other buffers) and the mixture was incubated at 37℃ for 1-2 hours.

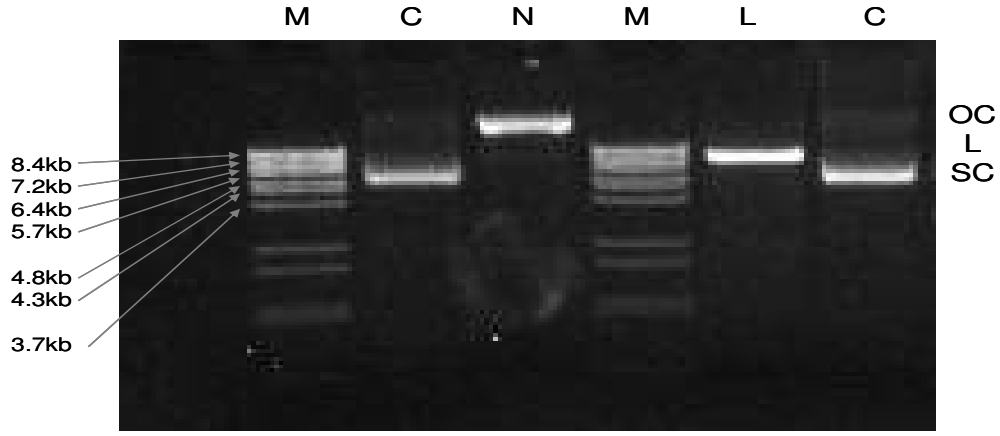
Figure 2.2.14.1. gWiz isoforms



C: Purified gWiz (5.8kb) sample
N: Nicked gWiz sample (OC-DNA)
L: Linear gWiz sample
Marker: λBstEII digest*
(*in this example, the 8.4kb band and the 5.7kb band annealled to form a 14kb band due to insufficient heating of the marker prior to use)

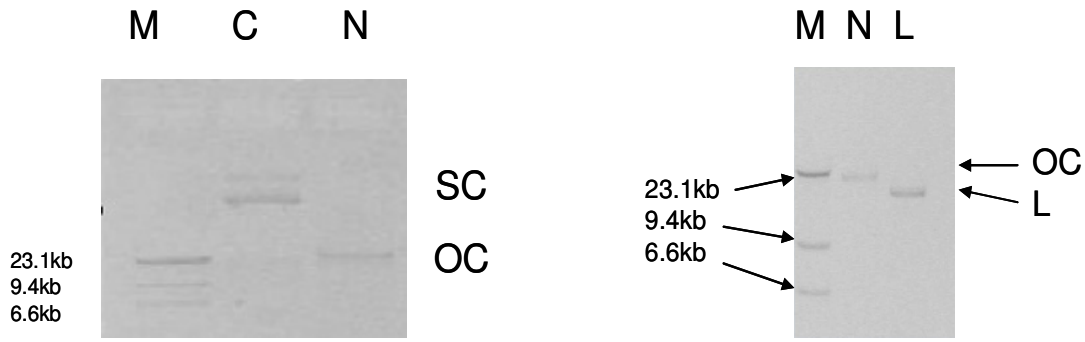
The higher order oligomers cannot be resolved any further by enzymatic nicking therefore the isoform(s) of these dimers cannot be determined. These bands were therefore omitted when estimating the total pDNA intensity.

Figure 2.2.14.2. pSVβ isoforms



C: Purified pSVβ (6.9kb) sample
N: Nicked pSVβ sample (OC-DNA)
L: Linear pSVβ sample
Marker: λBstEII digest

Figure 2.2.14.3. pQR150 isoforms



M: λHind III digest
C: Purified pQR150 (20kb) samples
N: Nicked sample (OC-DNA)
L: Linear sample

2.2.15. Copy Number Determination

10mL cultures were grown overnight for 18 hours in Nutrient Broth before samples were taken to determine the optical density (OD_{600nm}), DNA concentration and number of cells per mL. The number of cells per mL was determined when 1mL of sample was serially diluted in sterile RO water, before 0.05mL was aliquoted onto nutrient agar plates containing the appropriate antibiotic. The cells were then spread evenly onto the surface of the agar using a glass spreader that had been ethanol treated, flame sterilised and cooled before each spreading. The plates were then left to incubate for 24 hours at 37°C. After 24 hours only the plates that had between 30-350 colonies were counted. The number of bacterial cells per mL was then calculated as in [1]:

$$Number\ of\ cells\ per\ mL = \frac{Number\ of\ colonies}{Dilution\ factor \times volume\ of\ sample\ plated\ (mL)} \quad [1]$$

The size of each plasmid in kb was used to calculate the plasmid mass in µg using Avogadro's number (6.02×10^{23}) and then the number of plasmid molecules present in 1 µg mL^{-1} as shown in [2].

$$Number of\ plasmids\ per\ mL =$$

$$\frac{6.02 \times 10^{23}\ plasmids}{1 mole} \times \frac{1 mole}{plasmid\ mass\ (\mu g)} \times \frac{amount\ of\ plasmid\ DNA\ (\mu g)}{1 mL\ of\ solution} \qquad [2]$$

The copy number was then calculated by dividing the number of plasmid molecules per mL by the number of cells per mL. This method of determining plasmid copy number was adapted from Smith and Biochka, 1998.

2.2.16. *gyrA96* Reversion Test

Control (non-selective) nutrient agar plates and selective nutrient agar plates containing 30µg/mL nalidixic acid (Sigma-Aldrich) were prepared. Strains that have the *gyrA96* mutation should have resistance to nalidixic acid and therefore grow on both plates, whereas strains without the *gyrA96* mutation should only grow on the control plates. If a *gyrA96* strain does not grow on the selective plate it could suggest that the mutation has reverted through many generations.

2.2.17. Glucose Concentration Determination

The concentration of D-glucose was measured using a Nova 400 Bioanalyser (Nova Biomedical Corporation). Culture samples were spun down and the supernatant was collected in a fresh tube and filtered using a 0.22µm syringe filter before being analysed by the bioanalyser.

2.2.18. Acetate Concentration Determination

The concentration of acetate produced during the *E. coli* fermentation runs was determined using HPLC analysis. Culture samples were spun down and the supernatant was collected in a fresh tube and filtered using a

0.22µm syringe filter, before being analysed by HPLC Aminex HPx-78H column (Bio-Rad, Hercules, CA, USA) in both the Ultraviolet (UVD-1700, Dionex UK Ltd, Leeds UK) and Refractive Index, RI, (RI-101, Dionex UK Ltd) spectra. The sample readings were compared to an acetate standard calibration curve at concentrations of 1.25g/L, 2.5g/L, 5g/L, 10g/L and 20g/L. For each sample, the injection volume was 20µL (Dionex AS1-100 Automated sample injector, Dionex UK Ltd) and the run time was 25 minutes at a flow rate of 0.6mL/min.

2.2.19. Agarose Gel Electrophoresis

Agarose gels were run to determine the topology of the plasmid and the respective SC, OC and linear band intensities. 0.8% (w/v) agarose gels pre-stained with 5µL of ethidium bromide (500mg/mL) were run in either 1xTAE (Tris–acetate EDTA) or 0.5xTBE (Tris-borate EDTA) buffer (Sigma-Aldrich) at 80V (TAE) or 100V (TBE) for approximately 2 hours. A UV gel cam (Biorad Geldoc 2000, Biorad, Hercules, CA, USA) was used to highlight the bands and the software (Quantity One, Biorad) was used to as the densitometry tool, analysing the band contour (area) and intensity. A correction factor of 1.36 was applied to all the supercoiled bands as mentioned in section 1.9.1. in Chapter 1.

2.2.20. Chloroquine Agarose Gel Electrophoresis

The initial method used in this investigation for chloroquine agarose gel preparation is to dissolve the agarose in 1xTBE running buffer (0.8% agarose gel) and then wait for the gel to cool slightly before adding 5µg/mL chloroquine (dissolved in 1xTBE running buffer). The gels were left to set and were covered with aluminium foil to protect the light-sensitive chloroquine. The same concentration of chloroquine was added to the running buffer in the gel tank. Gels were run at 15V overnight at room temperature. The gels were then washed with fresh RO water twice for 1 hour on a gentle rocking platform. After washing out the chloroquine, the gels were stained with 1µg/mL ethidium bromide (in 1xTBE buffer) for 2 hours with gentle shaking. The depicted method is described as “initial” since these were the initial conditions set. Further changes were

implemented when multiple attempts at resolving the toposiomers of gWiz using this technique failed. Individual changes tested included varying the chloroquine concentration (various concentrations between 1µg/mL to 100µg/mL were tested); changing the agarose concentration from 0.8% to 1% and 1.2%; changing the running buffer to 0.5x TBE or 1xTAE buffer; changing the gel running time up to 23 hours; changing the washing solution from fresh RO water to fresh running buffer (same type and concentration as gel running buffer) and to RO water with 1mM $MgSO_4$. Unsuccessful attempts continued even when different plasmids were tested (pSVβ, pUC18 and pUC19) and when the pDNA samples were treated with Topoisomerase I from wheat germ (Sigma-Aldrich) with increasing concentrations of a helix overwinding agent, $CaCl_2$ (based on method described by Xu and Bremer, 1997).

2.2.21. Microscope Slide Preparation

Microscope analysis was conducted using a Leica DMRA2 microscope (Leica Microsystems UK Ltd.) and Leica QWin imaging software. 5µL of culture broth from 10mL overnight cultures were aliquoted onto a microscope slide and the spread evenly across the surface of the slide using a glass cover-slip. This cover-slip was then discarded and the cell samples were left to air dry until the liquid evaporated. The samples were then heat fixed when the slides were passed over hot bunsen flame. 3mg/mL of methylene blue was then pipetted onto each slide and left to stain for 1 minute before the methylene blue was rinsed out with RO water. A fresh cover-slip was then placed over the slide and fixed in place around the edges using clear nail varnish. These prepared slides were then viewed at 1000x magnification using immersion oil. Negative controls were set in place by viewing slides containing no cells. The cells were removed by spinning down the samples at 13,000rpm for 3 minutes. 5µL of the supernatant was then aliquoted onto the slides and the above heat fixing and staining procedure was carried out.

2.2.22. Cell Counting Using A Haemocytometer

A Bright-Line Haemocytometer (Sigma-Aldrich) was used with a standard bench top microscope (Olympus CH series, Olympus UK Ltd, Watford, UK). *E. coli* cell culture samples were diluted in 1xPBS (100-500x dilution, depending on cell density) and mixed thoroughly to minimise cell clumping. The haemocytometer was cleaned with ethanol and wiped dry before use. With the cover-slip in place over the haemocytometer chamber, a small amount of cell suspension was transferred into both chambers of the haemocytometer using a pipette, allowing each chamber to fill by capillary action. Care was taken to ensure that each chamber was not under- or overfilled. The chambers were then viewed separately under a microscope at 400x magnification. For each chamber, the cells were counted in the 1mm centre square and four 1mm corner squares, counting only the cells on the top and left middle line. The aim is to view between 20-50 cells per square so the dilution factor of the cell samples in PBS was adjusted each time in order to achieve this. After counting both chambers, the chambers were wiped clean with ethanol and dried with a lens tissue before repeating the procedure again with a second sample. With the cover-slip in place, each square of the haemocytometer represents a total volume of 0.1mm^3 and since 1cm^3 is approximately equivalent to 1mL, the subsequent cell concentration per mL can be determined from counting 10 squares [1]:

$$Cells\ per\ mL = Average\ count\ per\ square \times dilution\ factor \times 10^4 \quad [1]$$

The total number of cells can then be determined using [2]:

$$Total\ number\ of\ cells = cells\ per\ mL \times original\ volume\ of\ culture \quad [2]$$

2.2.23. GFP Extraction and Analysis

1mL culture samples were spun down and resuspended in 1xPBS buffer before the washed cell pellets were frozen at -20℃ for 24-48 hours. Thawed samples were then treated with 300µl Bugbuster® HT Protein Extraction Reagent (Bio-Rad, Hercules, CA, USA). Treated samples were left to incubate with gentle shaking at room temperature before they were spun down at 13000rpm for 20 minutes at 4℃ using an Eppendorf CF 5415 R centrifuge. The supernatant from each sample were collected in fresh eppendorf tubes and the samples were used immediately for fluorescence analysis or were frozen at -20℃.

The fluorescence intensity of GFP from the supernatant samples was measured in triplicate in 96 well plates using a Tecan Safire^2 plate reader. 100µL of 40x diluted samples (diluted in 1xPBS) were aliquoted into the wells and measured at excitation and emission wavelengths of 395nm and 509nm respectively.

2.3. Error Analysis

The distribution of the values measured from a repeated experiment is specified as either the standard deviation or as the standard error of the mean.

2.3.1. Standard deviation of a sample (non-biased) equation:

$$\sigma_x = \sqrt{\frac{\sum (x - \overline{x})^2}{(N-1)}} \qquad [1]$$

2.3.2. Standard error of the mean equation:

$$\sigma_{\overline{x}} = \frac{\sigma_x}{\sqrt{N}} \qquad [2]$$

Where

σ_x = standard deviation of the sample

x = each value in the data set

$\bar{x}$ = mean of all values in the data set

N = number of values in the data set

$\sigma_{\bar{x}}$ = standard error of the mean

2.3.3. Linest

To determine the error associated with specific growth rates, the Linest function in Excel (Microsoft) was implemented to calculate the standard error in the data by using the "least squares" method to calculate an array that best fits the data. Linest is a linear regression method and is only appropriate to fit linear data, therefore the natural log was applied to the data prior to regression.

The output of the Linest function (least squares results) is displayed in multiple cells (2x5) and include the values for the slope, constant and the standard deviation of both:

2.3.3. Table showing the output parameters of the Linest function

	A	B
1	Slope	Constant value
2	Standard deviation of the slope	Standard deviation of the constant value
3	Coefficient determination, r^2	Residual deviation, s(y)
4	Fisher statistic, F	Degrees of freedom (n-2)
5	Regression SS	Residual SS

SS= sum of squares
n= number of data points

2.4. Principal Components Analysis (PCA)

The Microsoft Excel-based program which applies PCA on the data set was written and coded in-house by Dr Simon Edwards-Parton and the resulting data is plotted onto 2D or 3D graphs using MATLAB version R 2006a. The Excel-based PCA program facilitates the input of the culture

data and automatically auto scales the data prior to analysis, i.e. the data is mean centred and scaled to unit variance. Variance scaling, such as auto scaling was chosen over standard mean centring as auto scaling as it focuses upon the deviations from the variable mean and investigates the variance from the mean data rather than the data as a whole, allowing the input of data of differing scales and units and improves the level of resolution. Mean centering would be the appropriate choice in systems where all the variables have the same units, since variables showing significant variation are more important and contain more information than those with small variation in such a system.

The PCA program analyses the data by establishing relationships between the original variables using a linear transform that decompose the data matrix into several 'directions' known as principal components. The key notion is that the smaller data set obtained is easier to manage and analyse, without compromising the original variation present in the data set. Please refer to Chapter 1 (section 1.10.) for a more in depth description of PCA principles.

The PCA program takes the original data with its multiple-dimensionality and translates these into fewer dimensions, for example, if you have 8 factors (8 dimensions), PCA can translate the variation present in these 8 dimensions and reduce this down to 3 dimensions (or 3 prinicipal components). This translation equation is known as the loadings data and this data also provides a weighting for each factor and contains information on how the factors relate to each other. In the context of plasmid DNA production, the principal components can be expressed in terms of the original data e.g. PC1= aOD + bμmax + cpDNA Yield, with a, b and c representing the weighting of each factor. The scores data contain information on how the samples relate to each other and provide the individual coordinates of each point (in this case, each strain-plasmid combination) for each principal component for example, when PC1 is plotted against PC2, the scores for PC1 form the co-ordinates on the x-

axis and the scores for PC2 form the co-ordinates for the y-axis for each strain-plasmid combination.

The principal components are ordered by variance, so the first principal component, PC1, will have captured the most data variation. This is important as it allows one to differentiate between the principal components that are useful and those that have mainly only captured experimental noise.

3. Host Strain Influences on Supercoiled Plasmid DNA Production in *E. coli*

3.1. Introduction

The growing demand for quick and effective methods of producing large amounts of plasmid DNA for human therapy and vaccination has increased the practical challenges associated with the optimisation of supercoiled (SC) plasmid DNA (herein after called SC-DNA) yield obtained through current methods. Although much effort has been directed towards downstream processing, there is still much room for improvement at the earlier stages of production namely the cell host/plasmid selection where the choice of the *E. coli* host strain used to propagate the product containing-plasmid may have sizeable influence on the quality and quantity of the purified DNA. This would influence both production as well as the downstream operations.

Generally *E. coli* strains with high plasmid copy numbers and high plasmid retention levels (segregational stability) are used as hosts for plasmid production, particularly if they are compatible with the latter downstream recovery and purification stages and have minimum potential for genetic alterations (Durland and Eastman, 1998).

Many reports discuss fermentation strategies and factors that can enhance process performance such as the effects of growth medium, feeding mode and employment of high cell-density cultures (Lee, 1996; O'Kennedy et al., 2000 & 2003; O'Mahony et al., 2007; Prather et al., 2003). Whilst additional research on these well documented factors and conditions may further aid process optimisation, there are some less explored areas that merit investigation such as the influences of the host strain. To date there has been limited published reports regarding the common consensus on choosing a particular *E. coli* strain based on its genotypic or phenotypic properties (Luli and Strohl, 1990; Taylor et al., 1993) for the purpose of preparing recombinant SC-DNA. These reports were relatively brief or were not directly aimed at SC-DNA production. Consequently it would be beneficial if a systematic approach were to be conducted to derive and collate the data of various *E. coli* strains and use

this data as a starting point for designing the most suitable host for a specific plasmid-based product.

It has been 85 years since *E. coli* K-12 was first cultivated (Bachmann, 1996) and the accumulation of selected and unknown mutations through individual pathways of descent in several different laboratories since this time has made it very difficult to establish if there are any links between these mutations to the overall growth characteristics of these strains. This problem is not helped by the fact that gaps in laboratory records mean that the history of some strains cannot be properly established (Bachmann, 1972 & 1996). In addition to this, newer, commercial strains are constantly being developed with claims that these strains are better suited for molecular biology than the older strains. It is now a requirement that the source of bacterial cells used, including its phenotype and genotype and what procedures are used to generate the cells are documented as part of a new drug application (FDA,1996a).

It is probable that in the near future, larger plasmids will be required to accommodate larger genes, control regions or multiple immunological proteins (Levy et al., 2000). It is hypothesised that as the plasmid size increases, the general stress on the host cell would also increase which would lead to a decrease in the maximum yield. Previous studies on a few *E. coli* strains with plasmids of different sizes up to 8kb showed that the increase in plasmid size did not significantly alter the growth rate, but a decrease in the host growth rate was observed with a 21kb plasmid when the metabolic load on the host cell was far greater (Cheah et al., 1987; Warnes and Stephenson, 1986). These reports however do stress that the impact of increased plasmid size on the growth rate is strictly dependent on the host/vector system and not only on the plasmid size. This is supported by studies on the impact of plasmid size on cellular oxygen demand in *E. coli* (Kay et al., 2003), where it has been shown that it is not the plasmid size that causes an increase in host cellular oxygen demand but it is rather due to the proteins expressed by the plasmids.

3.2. Aim and Approach

In this chapter we report the effects of host strain selection on plasmid stability, purity and the ratio of supercoiled DNA obtained at shake flask scale, using 3 plasmids of different sizes (gWiz, 5.8kb; pSVβ, 6.9kb and pQR150, 20kb).

Whilst each strain may exhibit different trends under optimised conditions, by keeping the growth conditions constant for each strain-plasmid combination, we are only looking at one variable which is the host strain. A list of the 17 strains used in this study and their sources can be found in Tables 3.2.1. and 3.2.2. More information on the storage conditions of the strains can also be found in Chapter 2 section 2.1.2. A description of the mutations listed in the genotypes can be found in the Appendices section A.2. The strains were selected based on common lab usage in an attempt to uncover what might make some strains better at plasmid production than others and were not only selected with industrial applications in mind.

10mL cultures containing Nutrient broth and the appropriate antibiotic were grown in 50mL Falcon tubes for 16.5 hours at 37°C, 250rpm.

100mL working volume shake flask cultures containing Nutrient Broth and the appropriate antibiotic were grown in 500mL shake flasks for 8 hours at 37°C, 250rpm.

Table 3.2.1. A list of commercial *E. coli* strains and sources used.

Strain	Genotype	Source	Derivation	Reference
ABLE K	*lac (lacZ⁻) [Kan^r, mcrA⁻, mcrCB⁻, mcrF⁻, mrr⁻, hsdR(r_K^- m_K^-)] [F'proAB, lacI^q,ZΔ M15, Tn10(Tet^r)]*	Purchased from Stratagene in 2006	*E. coli* C	Stratagene
BL21 DE3	*F⁻ dcm ompT hsdS($r_B^-m_B^-$) galΔ (DE3)*	Purchased from Stratagene in 2006	*E. coli* B	Studier and Moffat, 1986
DH5α	*F⁻ endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(r_K^- m_K^+), λ–*	JMW culture collection (UCL), obtained from Invitrogen	*E. coli* K-12; derived from DH1	Invitrogen; Hanahan, 1985; Grant et al., 1990.
DH10β	*F⁻ endA1 recA1 galU galK deoR nupG rpsL ΔlacX74 Φ80lacZΔM15 araD139 Δ(ara,leu)7697 mcrA Δ(mrr-hsdRMS-mcrBC) λ⁻*	JMW culture collection (UCL), obtained from Invitrogen	*E. coli* K-12; MC1061 derivative	Invitrogen; Grant et al., 1990
SURE 2	*e14⁻(mcrA⁻), Δ (mcrCB-hsdSMR-mrr)171, endA1, supE44, thi-1, gyrA96, relA1, lac, recB, recJ, sbcC, umuC::Tn5(Kan^r), uvrC, [F' proAB, lacI^{q}ΔM15, Tn10(Tet^r), Amy, Cam^r]*	Purchased from Stratagene in 2006	*E. coli* ER1451 derivative	Stratagene
TURBO	*[F' proA$^+$B$^+$ lacI^q ΔlacZM15/fhuA2 Δ(lac-proAB) glnV zgb-210::Tn10 (Tet^R)] endA1 thi-1 Δ(hsdS-mcrB)5*	Purchased from New England Biolabs in 2006	*E. coli* K-12	New England Biolabs
XL1 Blue	*endA1 gyrA96 (nal^R) thi-1 recA1 relA1 lac glnV44 [F'::Tn10 proAB$^+$ lacIq Δ(lacZ)M15] hsdR17(r_K^- m_K^+)*	JMW culture collection (UCL), obtained from Stratagene	*E. coli* K-12	Stratagene; Bullock et al., 1987
XL10 Gold	*endA1 glnV44 recA1 thi-1 gyrA96 relA1 lac Hte Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 tet^R [F'proAB lacI^{q}ZΔM15 Tn10(Tet^R Amy Cm^R)]*	Purchased from Stratagene in 2006	*E. coli* K-12	Stratagene

JMW= Professor John Ward

Table 3.2.2. A list of non-commercial *E. coli* strains and sources used.

Strain	Genotype	Source	Derivation	Reference
C600	*(F⁻ tonA21 thi-1 thr-1 leuB6 lacY1 glnV44 rfbC1 fhuA1 λ⁻)*	JMW culture collection from the University of Bristol, 1976-1980	*E. coli* K-12	Appleyard, 1954; Hanahan, 1983
DH1	*endA1 recA1 gyrA96 thi-1 glnV44 relA1 hsdR17 (r_K^- m_K^+) λ^-*	JMW culture collection (UCL), purchased from the NCTC #50236	*E. coli* K-12; parent of DH5α	Low, 1968; Bachman, 1996
HB101	*F⁻ mcrB mrr hsdS20 (r_B^- m_B^-) recA13 leuB6 ara-14 proA2 lacY1 galK2 xyl-5 mtl-1 rpsL20 (Sm^R) glnV44 λ^-*	JMW culture collection (UCL), purchased from the ATCC	*E. coli* K-12; hsd genes derived from *E. coli* B	Boyer and Roulland-Dussoix, 1969
J53	*proA met*	JMW culture collection from the University of Bristol, 1976-1980	*E. coli* K-12*;* J5 derivative	Clowes and Hayes, 1968
JM107	*supE endA1 glnV44 thi-1 relA1 gyrA96 Δ(lac-proAB) [F' traD36 $proAB^+$ $lacI^q$ lacZΔM15] hsdR17(R_K^- m_K^+) λ^-*	JMW culture collection from the University of Bristol, 1976-1980	*E. coli* K-12: DH1 derivative	Yanisch-Perron et al., 1985
MG1655	*F⁻ λ^- ilvG- rfb-50 rph-1*	JMW culture collection (UCL), purchased from the ATCC	*E. coli* K-12 wild type	Blattner *et al.,* 1997; Hayashi et al., 2006; Riley et al., 2006
TG1	*supE thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5 ($r_K^-m_K^-$) [F' traD36, proAB, $lacI^q$ ΔM15]*	JMW culture collection (UCL) obtained from the Eastman Institute, UCL	*E. coli* K-12	Guyer et al., 1981
TOP10	*F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 deoR nupG recA1 araD139 Δ(ara-leu)7697 galU galK rpsL(Str^R) endA1 λ^-*	JMW culture collection (UCL), obtained from Invitrogen	*E. coli* K-12; genotype very similar to DH10β	Grant *at el.,* 1990; Casdaban and Cohen, 1980
W3110	*F⁻ λ^- rph-1 INV (rrnD, rrnE)*	JMW culture collection (UCL), purchased from the ATCC	*E. coli* K-12 wild type	Hayashi et al., 2006; Riley et al., 2006

JMW= Professor John Ward; NCTC= National Collection of Type Cultures; ATCC= American Type Culture Collection

Table 3.2.3. A list of plasmids and their reported characteristics used (plasmid maps can be found Appendices section A.1.)

Plasmid	Origin of replication	Characteristics
pSVβ	pUC-based containing the pMB1 origin of replication. A single point mutation within the replication primer RNA II leads to an increase in copy number	pSV-β-Galactosidase plasmid vector encoding the β-gal gene. Designed as a positive control vector for monitoring transfection efficiencies of mammalian cells. The SV40 early promoter and enhancer initiate transcription of the bacterial lacZ gene. pSVβ is 6.9kb in size and is resistant to ampicillin. Colonies of *E. coli* containing pSVβ will appear blue when plated on media containing X-gal.
gWiz GFP	pUC-based containing the pMB1 origin of replication. The presence of the point mutation in RNA II has not been confirmed	The gWiz series of vectors have been designed to produce high levels of transient gene expression in a wide variety of mammalian cells and tissues as well as high yield of plasmid production in *E. coli*. The plasmid contains a proprietary modified promoter followed by Intron A from the human cytomegalovirus (CMV) immediate-early (IE) gene. gWiz GFP is 5.8kb in size and is resistant to kanamycin.
pQR150	pMB1 origin of replication	Designed to encode the entire *meta*-cleavage pathway, genes for the *meta*-cleavage pathway, *xyl*XYZLTEGFJQKIH, were cloned into pBGS18. pQR150 is 20kb in size and is resistant to kanamycin.

3.3. Results

The study was initiated to determine the effect different host strains of *E. coli* have on the yields of plasmid DNA and the quality of the plasmid DNA, as determined by the supercoiled proportion of the plasmid. There has been data on these two aspects published for single strains and also anecdotal evidence for differences between some *E. coli* strains. Since there are thousands of potential *E. coli* strains to select from, the 17 strains used here (Table 3.2.1. and Table 3.2.2.) were chosen based on their commercial availability, common usage in laboratories or to establish their potential as efficient hosts. The latter criterion is aimed mainly at "older" strains (C600, HB101, MG1655, J53, W3110, JM107, TG1) used in the early era of microbiology that have since been overlooked. MG1655, J53 and W3110 represent some of the early *E. coli* K-12 strains that have been exposed to the fewest cycles of mutations. MG1655 and J53 are two mutational steps from the original *E. coli* K-12 whilst W3110 has undergone 3 rounds of mutation (Chart 8 in Bachmann, 1996). Most of the chosen strains are derivatives of *E. coli* K-12 since the safety of K-12 strains have been the most thoroughly investigated and hence is preferred by biosafety regulatory bodies, compared to others such as derivatives of *E. coli* B (Bachman, 1996). Durland and Eastman (1998) also cite that current trends appear to favour well-characterised K-12 strains such as DH5α and DH10β (Invitrogen).

Figure 3.3.1. shows the growth profiles of the fastest and slowest growing *E. coli* strains for gWiz. As expected the results in Figure 3.3.1. and Table 3.3.1. show that commercial strains do not necessarily have a faster or slower growth rate than non-commercial strains.

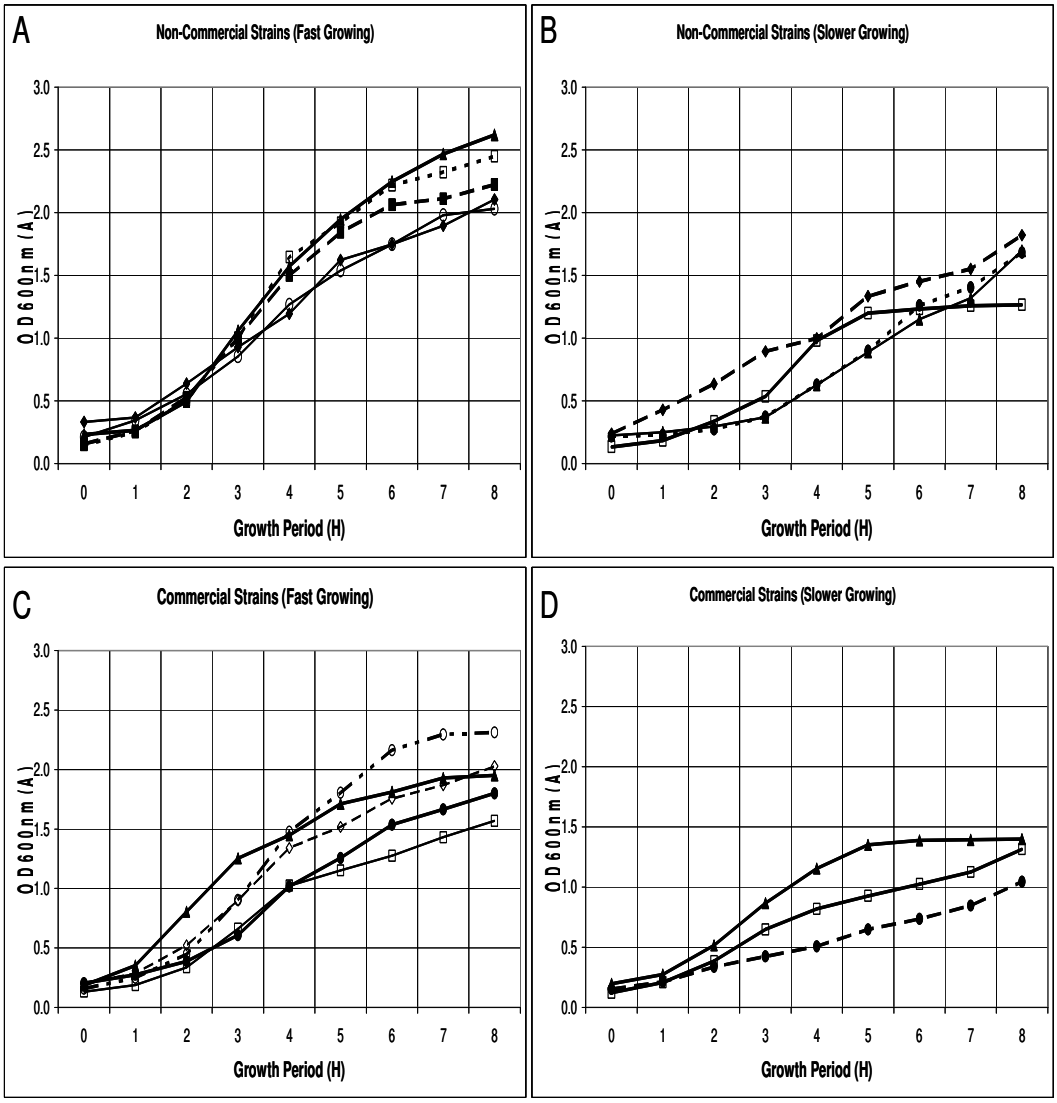


Figure 3.3.1. Average growth profiles of fast and slow growing *E. coli* strains harbouring the plasmid gWiz, 5.8kb, grown on Nutrient Broth at 37℃, 250rpm for 16.5 hours (100mL working volume scale). Values averaged from 2-3 replicates. A & B shows the growth profiles of the fastest and slowest growing non-commercial *E. coli* strains and C & D the fastest and slowest growing commercial *E. coli* strains investigated.

Presenting the performance of the host strains with each plasmid, Figures 3.3.2. (B) and 3.3.3. (B) also highlight (circled) the best performing strains in terms of high plasmid DNA yield and high level of supercoiling for the plasmid cultures gWiz and pSVβ respectively. Generally the same strains performed well with the two smaller plasmids (gWiz and pSVβ) and these are BL21 DE3, HB101, DH5α and TG1. However there were some exceptions as some strains yielded more DNA for one small plasmid and not for the other, such as MG1655, J53, DH10β, W3110 and XL1 BLUE. The differences in substantial stretches of DNA due to the presence of different genes in gWiz and pSVβ such as the GFP and β-galactosidase genes may be a possible cause for this. Another probable reason is the metabolic burden on the host cell caused by the expression of antibiotic resistance at an unreasonably high level. Plasmid-host interactions can also affect some metabolic pathways in the host cell for ColE1-like plasmids (Wang et al., 2006).

Some fast growing strains were not affected by high growth rates and yielded large amounts of gWiz plasmid, such as DH5α, MG1655, BL21 DE3 and TURBO. A boundary can be drawn around the qualitative trend in Figure 3.3.2. (D) (gWiz cultures) to show this with these 4 strains of the trend lying outside of this boundary. A similar boundary can be drawn for cultures harbouring pSVβ (a steeper line) as shown in Figure 3.3.3. (D), with DH5α and BL21 DE3 again outside the boundary. However our investigation also highlighted some strains that seemed to be more susceptible to the effects of high growth rate for the plasmids investigated such as DH1 and TOP10. DH1 (μ_{max} = 0.89h^{-1} for gWiz and 1.05h^{-1} for pQR150 cultures) and TOP10 (μ_{max} = 0.74h^{-1} for gWiz and 0.76h^{-1} for pQR150 cultures) yielded relatively low DNA concentrations and no supercoiling for gWiz cultures (Figure 3.3.2., Table 3.3.1.).

The results from the large plasmid pQR150 cultures (Figure 3.3.4, Table 3.3.3.) show that these cultures grew significantly faster than the smaller plasmid cultures, but yielded less plasmid DNA. Since pQR150 generally has a lower copy number (Table 3.3.4.), the expression of constitute

genes would be lower and this may explain why the plasmid DNA yields are lower. As similarly observed with the two smaller plasmids (Figures 3.3.2. (C), 3.3.3. (C) and Tables 3.3.1. and 3.3.2.), fast growth did not seem to influence the levels of supercoiling. Repeated experiments have however noted inconsistencies in the %SC-DNA obtained from certain *E. coli* strains harbouring pQR150 and we are not sure the exact reason for this. Possible reasons could be nicking of the plasmid during purification, stress caused by an increased metabolic burden, or plasmid instability (segregational and structural) and these factors are further explained in the discussions section.

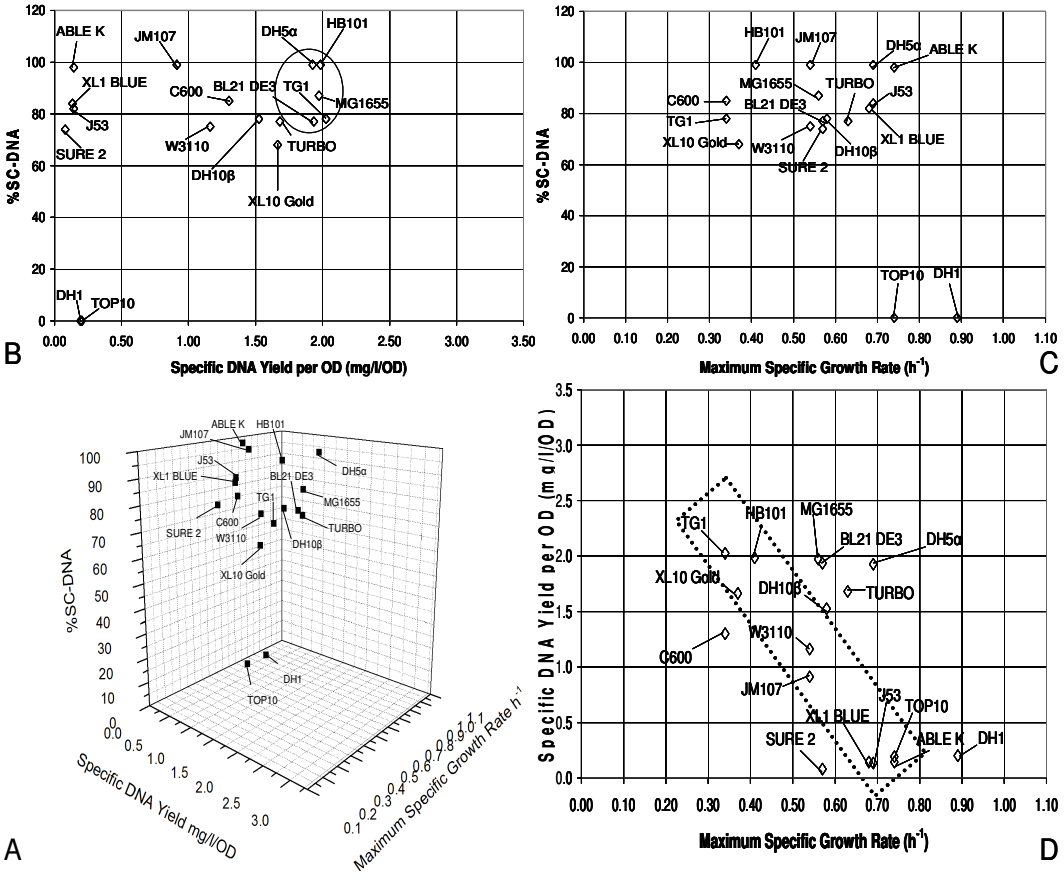


Figure 3.3.2. Scatter plot and specific yield, maximum growth rate and %SC-DNA for strains harbouring the plasmid gWiz. (A) is a 3-D scatter plot showing the relationships between the specific DNA yield per OD, maximum specific growth rate and %SC-DNA for all 17 strains transformed with the plasmid gWiz (5.8kb). These relationships can be more clearly observed in the 2-D scatter plots (B) showing the %SC-DNA vs. specific DNA yield per OD; (C) showing the %SC-DNA vs. the maximum specific growth rate and (D) showing the specific DNA yield per OD vs. the maximum specific growth rate. The circled area in (B) highlights the strains that provided consistently high yields of plasmid DNA. ------ Boundary area around the trend. All the data sets were averages obtained from a minimum of 3 replicates, 10mL scale cultures grown on Nutrient Broth at 37℃, 250rpm. Maximum specific growth rates were calculated from 100mL cultures also grown on Nutrient Broth at 37℃, 250rpm.

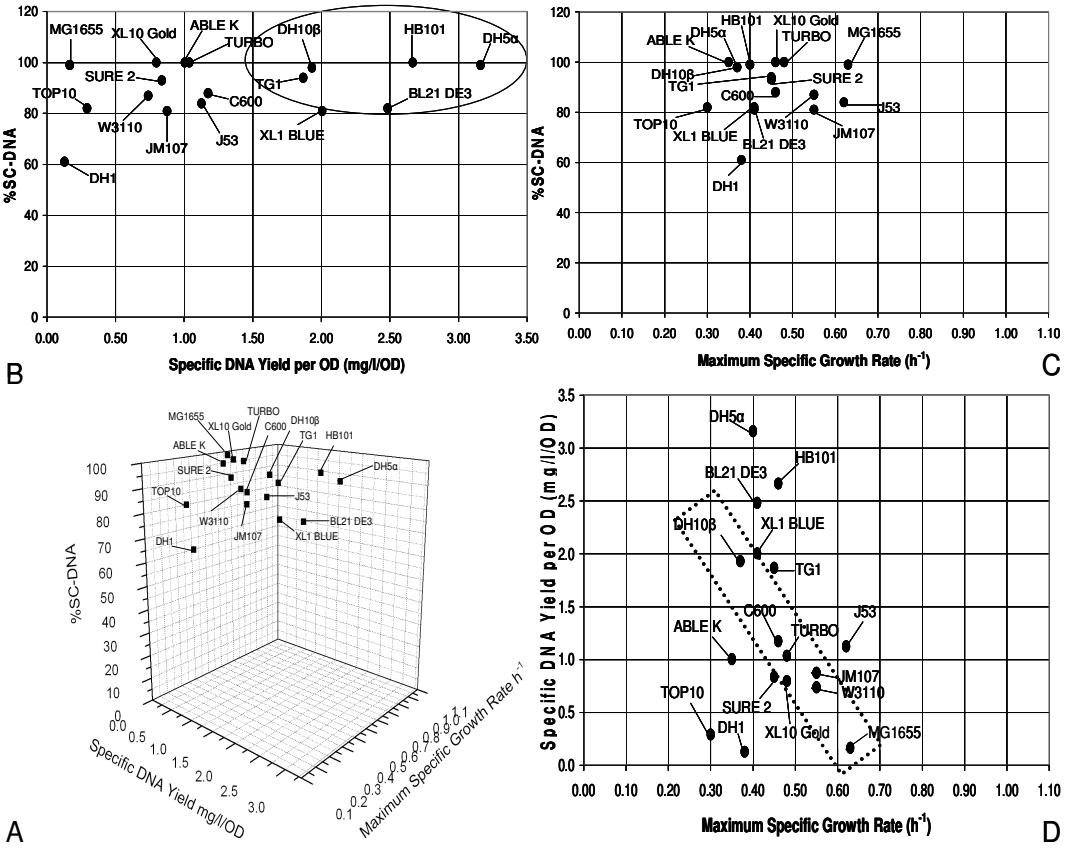


Figure 3.3.3. Scatter plot and specific DNA yield per OD, maximum growth rate and %SC-DNA for pSVβ. (A) is a 3-D scatter plot showing the relationships between the specific DNA yield per OD, maximum specific growth rate and %SC-DNA for all 17 strains transformed with the plasmid pSVβ (6.9kb). These relationships can be more clearly observed in the 2-D scatter plots (B) showing the %SC-DNA vs. specific DNA yield per OD; (C) showing the %SC-DNA vs. the maximum specific growth rate and (D) showing the specific DNA yield per OD vs. the maximum specific growth rate. The circled area in (B) highlights the strains that provided consistently high yields of plasmid DNA. ------ Boundary area around the trend. Data sets were averages obtained from a minimum of 3 replicates, 10mL scale cultures, except for the μmax values which were measured in microwell plates from one experiment and not replicated.

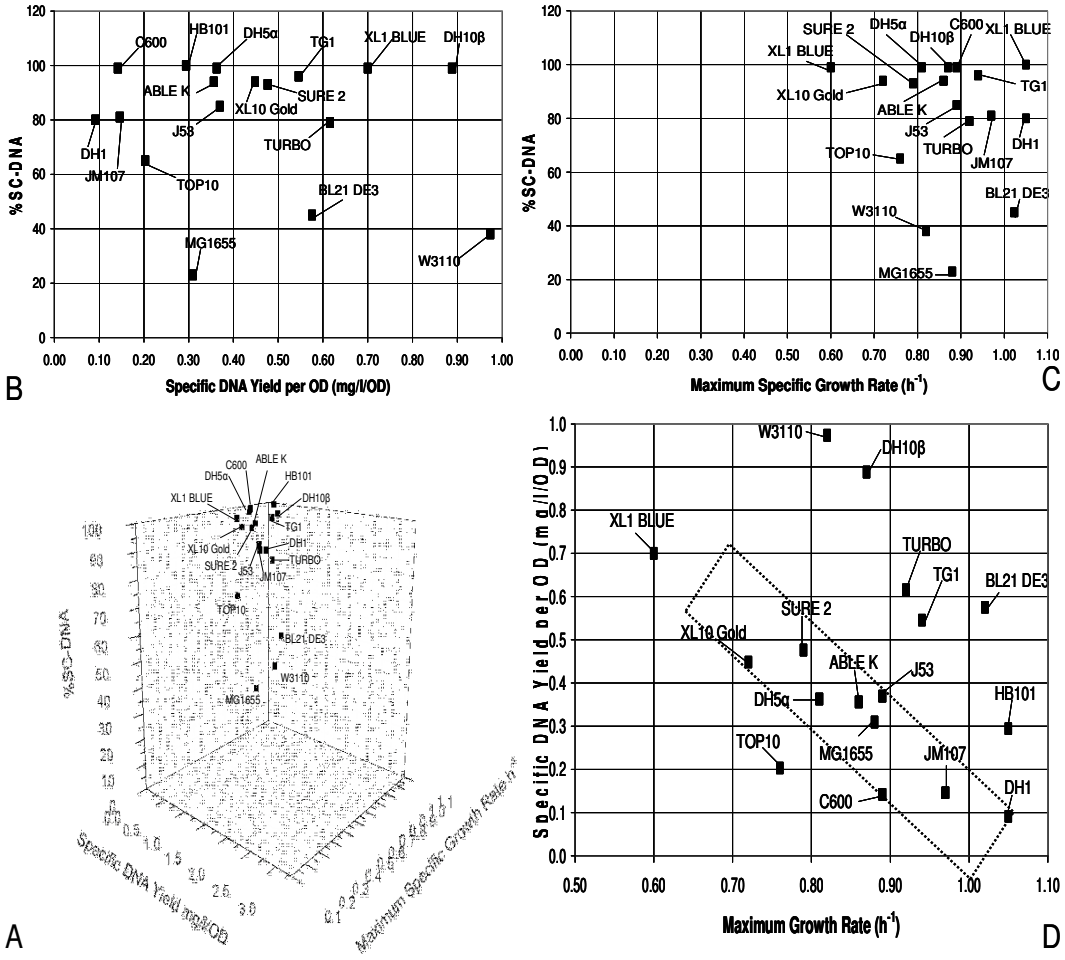


Figure 3.3.4. Scatter plot and specific yield, maximum growth rate and %SC-DNA for pQR150. (A) is a 3-D scatter plot showing the relationships between the specific DNA yield per OD, maximum specific growth rate and %SC-DNA for all 17 strains transformed with the plasmid pQR150 (20kb). These relationships can be more clearly observed in the 2-D scatter plots (B) showing the %SC-DNA vs. specific DNA yield; (C) showing the %SC-DNA vs. the maximum specific growth rate and (D) showing the specific DNA yield per OD vs. the maximum specific growth rate. ------ Boundary area around the trend. All the data sets were averages obtained from a minimum of 3 replicates, 10mL scale cultures except for the μmax values which were measured from 100mL shake flask experiments and not replicated. The scales for (B) and (D) differ to that in Figures 3.3.2. (B & D) and 3.3.3 (B & D).

Table 3.3.1. Comparison of biomass, total plasmid DNA and SC-DNA yields obtained from each *E. coli* strain harbouring the plasmid gWiz, 5.8kb, grown on Nutrient Broth at 37℃, 250rpm.

gWiz plasmid *E. coli* Strain	Volumetric pDNA yield (mg/L) from 10mL scale	Volumetric pDNA yield (mg/L) from 100mL scale	OD_{600} (10mL scale after 16.5 hours growth)	OD_{600} (100mL scale after 8 hours growth)	μ_{max} (h^{-1}) ± LINEST calculated standard error of μ	% SC-DNA ± 20% error	DCW (g/L) from 10ml scale culture	Specific pDNA yield (mg/L/OD) from 10mL scale culture	Specific pDNA yield (mg/L/OD) from 100mL scale culture
BL21 DE3	4.97 ± 0.11	5.90 ± 1.27	2.6 ± 0.16	2.0 ± 0.11	0.57 ± 0.05	77	1.70	1.93	2.91
MG1655	4.74 ± 0.32	1.95 ± 0.66	2.4 ± 0.05	1.8 ± 0.22	0.56 ± 0.03	87	1.64	1.97	1.11
HB101	4.61 ± 0.73	2.72 ± 0.48	2.3 ± 0.12	2.0 ± 0.06	0.41 ± 0.03	99	1.90	1.98	1.34
TG1	4.11 ± 0.54	0.91 ± 0.23	2.0 ± 0.03	1.7 ± 0.01	0.34 ± 0.02	78	2.19	2.02	0.54
TURBO	3.15 ± 0.61	1.13 ± 0.01	1.9 ± 0.09	1.6 ± 0.13	0.63 ± 0.05	77	1.78	1.68	0.72
DH5α	3.03 ± 0.21	1.26 ± 0.22	1.6 ± 0.31	1.8 ± 0.22	0.69 ± 0.05	99	2.07	1.93	0.70
DH10β	2.40 ± 0.09	0.75 ± 0.06	1.6 ± 0.10	1.4 ± 0.18	0.58 ± 0.04	78	1.83	1.53	0.53
XL10	2.33 ± 0.65	1.14 ± 0.06	1.4 ± 0.18	1.0 ± 0.15	0.37 ± 0.02	68	1.84	1.66	1.09
C600	2.29 ± 0.72	1.95 ± 0.08	1.8 ± 0.33	1.8 ± 0.17	0.34 ± 0.02	85	1.79	1.30	1.07
W3110	1.94 ± 0.55	0.75 ± 0.30	1.7 ± 0.32	2.1 ± 0.16	0.54 ± 0.04	75	1.56	1.16	0.35
JM107	1.31 ± 0.07	0.37 ± 0.04	1.4 ± 0.08	1.2 ± 0.06	0.54 ± 0.05	99	1.67	0.91	0.31
DH1	0.39 ± 0.11	0.43 ± 0.04	1.9 ± 0.09	2.6 ± 0.07	0.89 ± 0.07	0*	1.52	0.20	0.16
TOP10	0.36 ± 0.07	0.57 ± 0.21	1.8 ± 0.09	2.0 ± 0.21	0.74 ± 0.07	0*	1.22	0.19	0.41
ABLE K	0.33 ± 0.00	0.27 ± 0.07	2.3 ± 0.11	2.0 ± 0.05	0.74 ± 0.07	98	1.90	0.14	0.14
XL1 BLUE	0.26 ± 0.04	0.46 ± 0.14	1.8 ± 0.08	2.3 ± 0.11	0.68 ± 0.06	82	2.07	0.14	0.20
J53	0.25 ± 0.01	0.88 ± 0.02	1.8 ± 0.09	2.5 ± 0.09	0.69 ± 0.06	84	1.29	0.13	0.36
SURE 2	0.17 ± 0.03	0.12 ± 0.01	2.1 ± 0.22	1.3 ± 0.11	0.57 ± 0.05	74	1.47	0.08	0.09

*No SC-DNA bands were detected even after 5 replicates. Average values shown: 10mL cultures replicated 3-6 times, 100mL cultures replicated 2-3 times. DCW experiment was not replicated.

Table 3.3.2. Comparison of biomass, total plasmid DNA and SC-DNA yields obtained from each *E. coli* strain harbouring the plasmid pSVβ (6.9kb), grown on Nutrient Broth at 37℃, 250rpm.

pSVβ plasmid *E. coli* Strain	Volumetric pDNA yield (mg/L) from 10mL scale culture (16.5 hour culture)	OD_{600} from 10mL scale culture (after 16.5 hours)	μ_{max} (h^{-1}) ± LINEST calculated standard error of μ	%SC-DNA ± 20% error	DCW (g/L)	Specific pDNA yield (mg/L/OD) from 10mL scale culture
HB101	6.75 ± 0.94	2.5 ± 0.32	0.46 ± 0.05	100	1.69	2.67
DH5α	6.00 ± 0.77	1.9 ± 0.03	0.40 ± 0.04	99	1.58	3.16
BL21 DE3	5.22 ± 0.04	2.1 ± 0.07	0.41 ± 0.05	82	1.80	2.48
TG1	4.17 ± 0.24	2.2 ± 0.17	0.45 ± 0.04	94	1.75	1.87
XL1 BLUE	3.69 ± 0.47	1.8 ± 0.17	0.41 ± 0.03	81	2.18	2.01
DH10β	2.98 ± 0.21	1.5 ± 0.11	0.37 ± 0.03	98	2.68	1.93
C600	2.63 ± 0.37	2.2 ± 0.38	0.46 ± 0.05	88	1.27	1.17
ABLE K	2.20 ± 0.08	2.2 ± 0.04	0.35 ± 0.04	100	1.41	1.00
TURBO	1.98 ± 0.13	1.9 ± 0.08	0.48 ± 0.05	100	1.24	1.04
J53	1.79 ± 0.52	1.6 ± 0.22	0.62 ± 0.07	84	1.08	1.13
JM107	1.51 ± 0.26	1.7 ± 0.19	0.55 ± 0.07	81	1.34	0.88
XL10	1.30 ± 0.22	1.6 ± 0.07	0.48 ± 0.04	100	1.54	0.80
W3110	1.24 ± 0.49	1.7 ± 0.15	0.55 ± 0.05	87	1.81	0.74
SURE 2	0.90 ± 0.13	1.1 ± 0.05	0.45 ± 0.05	93	1.55	0.84
TOP10	0.62 ± 0.13	2.1 ± 0.17	0.30 ± 0.02	82	1.35	0.29
MG1655	0.29 ± 0.04	1.7 ± 0.20	0.63 ± 0.08	99	1.21	0.17
DH1	0.24 ± 0.01	1.9 ± 0.16	0.38 ± 0.05	61	1.55	0.13

Average values shown: 10mL cultures replicated 3-6 times. DCW and μmax experiments were not replicated.
μmax values measured using microwell plates (200μL of sample were aliquoted into microwell plates and measured hourly at 600nm).

Table 3.3.3. Comparison of biomass, total plasmid DNA and SC-DNA yields obtained from each *E. coli* strain harbouring the plasmid pQR150 (20kb), grown on Nutrient Broth at 37℃, 250rpm.

pQR150 plasmid *E. coli* Strain	Volumetric pDNA yield (mg/L) from 10mL scale	Volumetric pDNA yield (mg/L) from 100mL scale	OD_{600} (10mL scale after 16.5 hours growth)	OD_{600} (100mL scale after 8 hours growth)	μ_{max} (h^{-1}) ± LINEST calculated standard error of μ	% SC-DNA ± 20% error	DCW (g/L) from 10mL scale culture	Specific pDNA yield per OD (mg/L/OD) from 10mL scale culture	Specific pDNA yield per OD (mg/L/OD) from 100mL scale culture
W3110	1.49 ± 0.34	2.70	1.5 ± 0.10	2.8	0.82 ± 0.07	38	2.11	0.97	0.95
TG1	1.32 ± 0.16	1.16	2.4 ± 0.11	3.9	0.94 ± 0.08	96	2.46	0.55	0.30
DH10β	1.32 ± 0.16	1.54	1.5 ± 0.08	2.2	0.87 ± 0.07	99	1.00	0.89	0.71
BL21 DE3	1.30 ± 0.18	1.44	2.3 ± 0.20	3.0	1.02 ± 0.09	45	1.79	0.58	0.48
XL1 BLUE	1.19 ± 0.11	1.16	1.7 ± 0.21	2.2	0.60 ± 0.03	99	3.74	0.70	0.53
TURBO	0.99 ± 0.25	1.34	1.6 ± 0.19	3.2	0.92 ± 0.08	79	0.83	0.62	0.42
SURE 2	0.92 ± 0.23	1.59	1.9 ± 0.17	2.2	0.79 ± 0.06	93	1.01	0.48	0.73
ABLE K	0.86 ± 0.35	0.55	2.4 ± 0.09	2.9	0.86 ± 0.08	94	1.01	0.36	0.19
XL10	0.76 ± 0.13	0.50	1.8 ± 0.08	2.5	0.72 ± 0.05	94	1.66	0.45	0.20
HB101	0.67 ± 0.12	0.44	2.3 ± 0.05	3.5	1.05 ± 0.09	100	1.05	0.29	0.13
MG1655	0.64 ± 0.15	0.68	2.1 ± 0.20	3.1	0.88 ± 0.08	23	1.43	0.31	0.22
J53	0.61 ± 0.06	1.09	1.7 ± 0.18	3.5	0.89 ± 0.08	85	1.07	0.37	0.31
DH5α	0.52 ± 0.05	0.66	1.4 ± 0.05	2.8	0.81 ± 0.07	99	1.51	0.36	0.23
TOP10	0.40 ± 0.13	0.71	2.0 ± 0.18	2.4	0.76 ± 0.04	65	1.23	0.20	0.29
C600	0.33 ± 0.06	0.32	2.3 ± 0.09	3.0	0.89 ± 0.07	99	1.77	0.14	0.11
JM107	0.33 ± 0.04	0.14	2.2 ± 0.11	2.2	0.97 ± 0.06	81	1.20	0.15	0.06
DH1	0.16 ± 0.04	0.09	1.7 ± 0.24	2.1	1.05 ± 0.07	80	1.12	0.09	0.80

Average values shown: 10mL cultures replicated 3-6 times. 100mL cultures, DCW and μmax experiments were not replicated.

3.4. Discussion

3.4.1. Host Genotype and Selective Mutations

Certain selective mutations were introduced into several laboratory *E. coli* strains to improve their performance, such as *endA1* and *recA*, which are thought to prevent plasmid degradation and nicking during lysis (Carnes, 2005; Schoenfeld et al., 1995) and to improve plasmid stability respectively. However this study indicates that the effect of these selective mutations is very strain/plasmid dependent.

Of the strains used here (Tables 3.2.1 and 3.2.2.), eight do not have the *endA* mutation but the data from Figures 3.3.2, 3.3.3. and 3.3.4. show that several non-*endA* strains give high proportions of SC plasmid DNA. For gWiz (Figure 3.3.2.), ABLE K, HB101, C600 and MG1655 all give good % SC plasmid preparations and are not *endA* strains. Even for the larger plasmid pQR150 two strains that have a wild type endonuclease I (HB101 and C600) produce 100% and 99% respectively, SC-DNA whilst the *endA* strains TURBO, TOP10, JM107 and DH1 give 81-65%. These findings are supported by Taylor et al., 1993, which reported that even *endA* strains can result in plasmid degradation and plasmid nicking.

The *recA* mutation leaves about 10% of cells anucleate at each cell division and can generate up to 50% dead cells (Capaldo et al., 1974), yet the majority of the *recA* strains used here gave fast growth rates and only DH10β and XL10 Gold had slow growth out of the 7 *recA* strains in Tables 3.2.1. and 3.2.2.

Another mutation of interest is the *gyrA96* mutation. This mutation conveys resistance to nalidixic acid and strains that have this gene have reduced levels of supercoiled DNA (Cashel and Rudd, 1987; Ramirez and Villarejo, 1991). However DH5α, DH1, JM107, SURE 2, XL1 BLUE and XL10 Gold all have the *gyrA96* mutation, but the supercoiling of the three plasmid in these strains differ greatly. There are some combinations that manage to

achieve 99% SC-DNA whilst some combination of plasmid and host have zero. DH1 seems to fare the worst.

gyrA96 mutation reversion tests conducted on all 17 strains showed that all of the strains that have the *gyrA96* mutation still show resistance to nalidixic acid, implying that the mutation is still present. Control tests were put in place and strains without the *gyrA96* mutation grew only on the non-selective control plates, with the exception of TOP10 and DH10β. Many reports that mention TOP10 and DH10β refer to Grant et al. 1990, however upon closer examination of this paper, it is apparent that TOP10 is never mentioned. Since the genotypes of TOP10 and DH10β are very similar (Tables 3.2.1. and 3.2.2.), it is very likely that TOP10 is a derivative of DH10β and both may have the *gyrA96* mutation carried over from a common parent. This example highlights the difficultly in tracing the origins of some strains and the added complexities of correlating gene mutations since so many have silent mutations. Although TOP10 and DH10β are similar genotypically, it is apparent through this investigation that they behave very differently, with DH10β performing much better in terms of pDNA yield, specific DNA yield per OD and level of supercoiling for all three plasmid cultures.

3.4.2. Commercial vs. Non-Commercial Strains

Commercial *E. coli* strains are widely used in laboratories and published reports on the high quality performances of some of these commercial strains have justified their use. However the data compiled in this investigation indicate that commercial strains do not always necessitate good productivity and it is important to make note of this as some commercial strains are designed to suit specific uses (Table 3.4.2.1.) and these claimed phenotypes do not necessarily improve the yield or quality of DNA produced. In this investigation some of these commercial strains perform poorly compared to the older, less mutated strains. Some manufacturers of plasmid-based therapies or treatments also prefer not to use commercial strains in their research and development since there are licensing issues associated with the use of strains owned by a third party.

It is also useful to note that strains like DH5α and DH10β were originally constructed non-commercially as reported in Grant et al., 1990; however in this study they are listed as commercial strains since they were sold or licensed to a company soon after their publication and are both currently owned by Invitrogen.

The commercial strain TURBO is stated to have the highest growth rate on agar plates (8 hours after transformation, New England Biolabs) which we have also observed, however its actual growth in both Nutrient Broth and LB media was surprisingly moderate and it did not grow as fast as some other strains, whether in its transformed or non-transformed state. The design of these commercial strains for specific applications may explain their relatively poor expression levels in our investigation. A possible negative correlation present in some commercial strains like SURE 2, ABLE K and XL10 Gold is that these strains already have inserted genes that code for antibiotic resistance. Depending on the nature of the resistance mechanism, the presence of these genes can increase the metabolic burden on the host strain particularly if the host cell chromosome is also harbouring a plasmid with the same antibiotic resistance gene, resulting in the expression of both set of genes when the corresponding antibiotic is added. Many antibiotic genes are from large, low copy number resistance plasmids such as R1 and normally there would be no more than 3-4 copies of these antibiotic genes per cell when on the original R-plasmids (Nordstrom et al., 1984). At this level the effect of the resistance genes should be minor, however in the case of SURE 2 and XL10 Gold, there are two different resistance genes on the F plasmid, and as a result this could increase the negative effect the expression of these genes may have.

Table 3.4.2.1. Table listing a few of the commercial strains investigated and their quoted purposes.

Strain	Quoted Purpose	Source
ABLE K	Reduces plasmid copy number of ColE1-derived plasmids 10-fold; helps in cloning toxic gene products	Stratagene
BL21 DE3	Contains the T7 RNA polymerase gene under control of the *lacUV5* promoter. The polymerase gene is integrated into the bacterial chromosome from l(DE3). Use for non-toxic protein expression.	Stratagene
DH5α	Ideal for routine sub cloning and high quality plasmid preparation.	Invitrogen
DH10β	Suitable for cloning DNA that contains methylcytosine and methyladenine.	Invitrogen
SURE 2	This strain is good for sub cloning genes with Z-DNA secondary structures and unstable clones.	Stratagene
TOP10	High efficiency for cloning. Genotypically very similar to DH10β.	Invitrogen
XL10 Gold	The high transformation efficiency (hte) phenotype is ideal for large plasmids.	Stratagene

3.4.3. Wild Type K-12 Strains

Research on the use of wild type K-12 strains such as MG1655 and W3110 is becoming more popular since their genomes have been completely sequenced and these strains are relatively free from undetected base changes in genetic DNA known as cryptic mutations (Lamb, 1975) that are present in newer strains. MG1655 and W3110, performed variably when compared to their newer, commercial counterparts. With the large plasmid pQR150 (20kb), W3110 produced the most pDNA and specific DNA yield per OD from a 100mL culture compared to the other 17 strains, albeit only 38% of this DNA was in the supercoiled form. Maximum specific growth rates for both MG1655 and W3110 transformed with each of the three plasmids were all very similar.

MG1655 (Blattner et al., 1997) is a very robust strain and is most similar to the original K-12 strain. Hayashi et al., (2006), has summarised that there are only 282 bases different between the MG1655 genome and that of W3110. When grown on Nutrient Broth, W3110 generally performed better than MG1655 when propagating 2 of the 3 plasmids tested, however with the gWiz plasmid, MG1655 performed significantly better than W3110. Since these two strains are so similar genotypically and in terms of growth, the differences in plasmid DNA propagation are hard to reconcile. It is possible that minor differences in one strain and not the other may affect the propagation of some plasmids.

3.4.4. Variation in Supercoiling of the Large Plasmid, pQR150

Our investigation has highlighted some inconsistencies in the level of supercoiling of the large plasmid pQR150 in certain *E. coli* strains even after many replicates and although we do not fully understand why this is, we have some probable suggestions.

One initial theory is that the spin columns used in the miniprep kits were nicking the SC-DNA into the OC form since pQR150 is a larger plasmid. This theory was based on observations during previous unpublished experiments using larger plasmids (>13kb), however in our study when

this potential problem was mitigated by precipitating pQR150 with isopropanol after the clearing spin and not using a spin column, low SC-DNA levels were still observed for some strains (Table 3.3.3.). Larger plasmids maybe more prone to nicking hence it is plausible that there maybe a positive correlation between the size of plasmid and the amount of SC-DNA obtained. However the results obtained from this investigation do not show a reduced level of SC-DNA from the pQR150 cultures for some strains and thus do not support this nicking theory.

Kay et al., (2003) reported that it is the large amount of proteins produced by some plasmids and not the increased plasmid size that creates an increased metabolic load leading to increased oxygen demand, and this is supported by our findings which show that the growth rates for the pQR150 cultures are not affected by the increased plasmid size. The 11 genes on pQR150 need to be induced with IPTG to see the increased metabolic load (Kay et al., 2003).

Segregational instability of plasmids, whereby multicopy plasmids are lost from cultures during cell division due to uneven plasmid partition into daughter cells (Summers, 1998; Summers and Sherratt, 1984), may be a problem with a larger plasmid as reported by Warnes and Stephenson 1986. The performance of XL1 BLUE transformed with a plasmid with different sized DNA inserts ranging from 2.96kb to 11.96kb, as reported by Smith and Bidochka, (1998), concludes that the rate of plasmid loss is positively correlated with the size of the plasmid. Plasmid retention tests on the strains harbouring pQR150 did not show more plasmid loss than what was measured from the smaller sized plasmids gWiz and pSVβ (Table 3.4.5.1.) so this issue remains inconclusive. Other possible factors contributing to this problem may include the structural stability of this plasmid. Cooke et al., 2004, showed that triplex DNA sequences will form structures that make a plasmid more prone to some physical treatments. pQR150 is not known to contain any of these structures in the TOL plasmid parts that have been sequenced (Greated et al., 2002).

3.4.5. Correlation of Copy Number and Growth Rate

The data from this investigation suggest that the correlation between the copy number and the growth rate is dependant on the host/plasmid system. In some strains pQR150 has a lower copy number than pSVβ and gWiz (both high copy pUC-based plasmids) but there are some notable exceptions (Table 3.4.5.1) where some strains give higher copy numbers for pQR150 than for pSVβ e.g. DH1, J53 and DH10β or gWiz e.g. JM107. Reports discussing the effects of copy number on the growth rate in continuous cultures have resulted in opposing findings, either with a positive correlation where high dilution rates (up to a dilution rate of 1 h^{-1}) resulted in increasing copy numbers of pBR322 containing the origin of replication from pMB1 (Reinikainen and Virkajarvi, 1989) or a negative correlation where high dilution rates (above 0.6 h^{-1}) correlated with lower copy numbers (Seo and Bailey 1986). Kim and Ryu, (1991), have also reported a negative correlation between high growth rates and low copy number whilst employing a different plasmid in batch cultures.

3.4.6. Runaway Plasmid Replication?

This study shows that the copy number is a function of both the host cell and the plasmid. The results obtained (Table 3.4.5.1.) show that most of the strains have a copy number in the 100's range (around 100 to 800) which is within the range of pUC-based plasmids such as gWiz and pSVβ with a pBM1 origin of replication. However there are a few strains (HB101, MG1655, TG1 and W3110) that provided very high copy numbers, much higher than this "normal" range for pUC-based plasmids (i.e. a copy number over 800). Cultures propagating the larger plasmid pQR150 (although also containing the pBM1 origin of replication, pQR150 is a pBGS18 derivative), were not affected by this phenomenon which seemed to be affecting a few of the gWiz and pSVβ cultures. The very high copy numbers produced in these strains suggests the loss of plasmid replication control, resulting in a large accumulation of plasmid DNA within the host cell. This characteristic that has been linked to the phenomenon known as 'runaway plasmid replication', where plasmids are hugely overproduced

during replication and this is usually observed when conditional runaway replication vectors are used.

It has been reported by Remaut et al., (1983); Togna et al., (1993) and Uhlin et al., (1978) & (1979), that runaway replication severely increases the metabolic burden on the cell that can result in a decreased growth rate and productivity. The growth rate and productivity of the affected strains in this investigation do not seem to be affected and it is also important to note the significant positive implications of this occurrence in bioprocessing. If this occurrence can be controlled, high plasmid titres may be potentially obtained from relatively low cell density cultures, easing fermentation and the latter downstream processing stages.

In this study the plasmids gWiz and pSVβ have not been altered to induce runaway replication but the high copy numbers obtained suggest something similar to runaway replication has occurred. When comparing the region known to be mutated in pUC plasmids that leads to high copy number (Lin-Chao et al., 1992), the DNA sequence for gWiz (Aldevron and Gelantis Inc.) shows that the mutation is not present in gWiz, but pSVβ does have this mutation. Whether the data for gWiz represents actual sequenced plasmid data or whether it has been compiled from the published sequence data for the various segments that make up gWiz is not known. There were early DNA sequences of pUC in the databases that were derived from the known parts of pBR322 and only later when pUC was actually sequenced was it shown to contain the G to A mutation (Lin-Chao et al., 1992).

Furthermore, these observations only occur in the same few strains for both plasmids which seemingly suggest that runaway replication is not only plasmid dependant but also host strain dependant. Uhlin et al., (1978), and Wong et al., (1982), have reported that the copy numbers of these mutated vectors increase when the temperature is raised to 40℃ since the Lambda repressor in these vectors are temperature sensitive. Initial tests on HB101, MG1655, TG1 and W3110 gWiz and pSVβ cultures

have shown that some of these strains provide higher specific plasmid DNA yields at 40℃ than at 37℃ or 30℃ (see Chapter 4).

It cannot be confirmed if runaway replication occurred as the results obtained can only suggest that something similar to this phenomenon was observed. Experiments required to confirm runaway replication were not conducted.

Table 3.4.5.1. Copy number range and percentage of plasmid containing cells of the transformed *E. coli* strains investigated.

Plasmid	**gWiz™ (5.8kb)**			**pSVβ (6.9kb)**			**pQR150 (20kb)**		
E. coli Strain	Copy Number		% cells containing plasmid	Copy Number		% cells containing plasmid	Copy Number		% cells containing plasmid
	Minimum	Maximum		Minimum	Maximum		Minimum	Maximum	
ABLE K	129	144	93	231	248	90	151	241	88
BL21 DE3	263	447	85	455	491	63	167	484	97
C600	315	740	94	837	891	84	81	104	98
DH5α	406	588	92	125	774	97	184	219	94
DH1	210	330	98	147	195	97	308	554	96
DH10β	610	789	80	184	243	94	237	476	93
HB101	988*	2708*	73	856	2448*	80	121	193	95
J53	335	338	95	87	125	78	240	326	93
JM107	94	192	92	475	668	84	116	318	90
MG1655	1840*	2206*	90	108	154	91	261	745	97
SURE 2	194	228	95	117	678	60	178	310	93
TG1	1425*	1715*	82	648	1872*	87	356	509	91
TOP10	458	619	90	152	226	83	181	186	91
TURBO	94	496	90	124	203	79	186	199	92
W3110	123	478	63	508	1967*	79	385	565	90
XL1 BLUE	153	274	93	217	755	92	207	336	91
XL10 Gold	554	736	89	118	322	96	296	344	94

The copy number and % cells containing plasmid values were obtained from 10mL cultures of *E. coli* propagating the corresponding plasmid, grown on Nutrient broth at 37℃ for 18 hours. The copy number ranges and average stability percentages were obtained from triplicate experiments. Values marked with * show extremely high copy numbers (too high to define accurately).

3.4.7. Principal Components Analysis

Principal Components Analysis was conducted using the results obtained from this chapter. Eight factors including four factors that were derived factors (12 factors in total), were inserted into the PCA program

Table 3.4.7. Factors from 10mL cultures of all 51 host strain-plasmid combinations (17 strains, 3 plasmids) inserted into the PCA program.

Factor	Derived factors
Volumetric plasmid DNA yield (mg/L)	-
Optical density (OD_{600nm})	-
Percentage of supercoiled DNA, %SC-DNA	-
Maximum specific growth rate, μ_{max}, (h^{-1})	-
Percentage of cells containing plasmid after 18 hours (%)	-
Dry cell weight, DCW, (g/L)	-
Maximum copy number value	-
Minimum copy number value	-
Specific pDNA yield per OD (mg/L/OD)	Derived from the pDNA yield and OD values
pDNA yield per DCW (mgDNA/gDCW)	Derived from the pDNA yield and DCW values
Specific productivity per OD per hour (mg/L/OD/h)	Derived from the pDNA yield, OD values and the culture time (16.5 hours)
Specific productivity per DCW per hour (mgDNA/gDCW/h)	Derived from the pDNA yield and DCW values and the culture time (16.5 hours)

In Figure 3.4.7.1. the eigenvalues (representing variation within the data captured by each principal component) show that 57% of the data variation is captured by the first principal component, PC1, and that 82%

of data variation is captured by the first, second (PC2) and third principal (PC3) components combined. This level of variation captured is sufficient to confirm and support the selection of these three components for further analysis and to disregard the fourth and fifth components, which have most likely captured noise inherent within the experimental data.

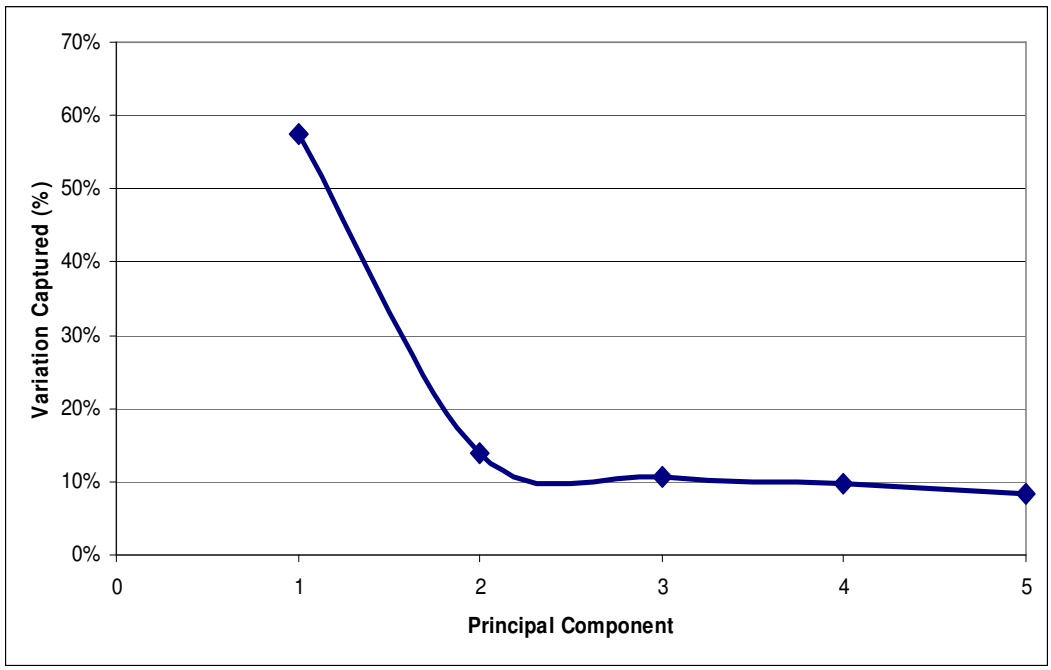


Figure 3.4.7.1. Principal Component Analysis (PCA) plot showing the % variation captured by five principal components. The majority of the data variation (57%) is captured by the first principal component (PC1) and 82% variation is captured in total by PC1, PC2 and PC3. This amount of variation captured is sufficient to safely select these three components for further analysis.

3.4.7.1. PCA Results

The graphical representations of the data obtained using PCA confirm initial trends observed in Figures 3.3.1.-3.3.4. in fewer graphs. When the scores data from PC1 is plotted against PC2 and PC2 plotted against PC3 (Figure 3.4.7.2.) the main trends in the form of data clusters are observed. In the PC1 vs. PC2 plot (Figure 3.4.7.2. top), the secondary cluster on the right shows a very positive correlation (as indicated by its top right hand side position on the chart) with PC1 and PC2. The plot showing PC2 vs. PC3 (Figure 3.4.7.2. bottom) also show two main clusters, however these

clusters are less spread out compared to the PC1 vs. PC2 plot, which can perhaps be explained by the lower percentage of data variation captured by the second and third PCs (14% and 11% data variation captured respectively).

The loadings data signify the weighting of each factor analysed and together with the standardised data, the positioning of certain host-plasmid combinations in the scores plots can be established. The loadings data of PC1 vs. PC2 (Figure 3.4.7.3. top), show that the variables which contribute most significantly to PC1 are plasmid DNA yield (volumetric) and the data derived from this factor (specific pDNA yield per OD, specific pDNA yield per DCW and the specific productivities per OD per hour or per DCW per hour) and the maximum specific growth rate, μ_{max}. μ_{max} has a negative correlation with PC1 whereas the volumetric pDNA yield, specific pDNA yields and productivities all have positive correlations, indicating that these factors are anti-correlated with μ_{max}. This makes sense as the growth rate can have a negative correlation with the copy number (and subsequently the plasmid yield obtained) as reported by Seo and Bailey, 1986 and Kim and Ryu, 1991.

No clear clusters are apparent in the PC1 vs. PC2 loadings plot (Figure 3.4.7.3. top), suggesting that none of the factors are very similar to each other and have differing correlations with either PC1 and/or PC2. Although no strong correlations with each other are shown, one or two factors are positioned closer to some than others, for example, the minimum and maximum copy numbers; and the volumetric pDNA yield and its derivatives as expected. The specific productivity values were derived from the specific pDNA yields by dividing the latter by the culture age at the time of sampling and as they were all mean centred and scaled by the same factor (16.5 hours), there is no effect on the relative position of each value. This would explain why the loading values are exactly the same for the productivity as for the specific plasmid yields.

Clusters are apparent in the PC2 vs. PC3 loadings plot (Figure 3.4.7.3. bottom), where %SC, pDNA yield, specific plasmid yields (per OD and per DCW) and specific productivities (per OD per hour and per DCW per hour) are shown to be clustering together on the left hand side of the plot, suggesting that these factors show similar negative correlation with PC2 and/or positive correlation with PC3.

3.4.7.2. High plasmid yield combinations

Based on the loadings and standardised data, the host-plasmid combinations found in the PC1 vs. PC2 scores plot (Figure 3.4.7.2. top, cluster B) are positioned where they are due to the high plasmid DNA yields and high copy numbers obtained from these combinations. Even within this secondary cluster, we can see that some combinations stand out more than others, namely MG1655 gWiz, HB101 pSVβ, HB101 gWiz, TG1 gWiz and TG1 pSVβ, all of which provided high copy numbers. However, upon initial examination of the known genotypes of these three host strains, there does not appear to be any obvious genotype markers that can be correlated with high specific pDNA yield.

Whilst PCA finds factors that capture the most variance in the predictor variables, it does not seek to find a single factor that best correlates predictor variables with predicted variables. Occasionally there may be factors in PCA models which do not contribute positively to the predictive ability of the model and this happens because the principal components are chosen without consideration of how they relate to the predicted variable. Another technique called Partial Least Squares (PLS) regression (Wold and Sjöström, 1998) is able to find factors that both capture variance and achieve correlation. Factors not contributing positively in PLS models happen much less frequently because the latent variables are chosen depending on how correlated the scores are to the predicted variable. PLS analysis was not conducted on the data during this investigation.

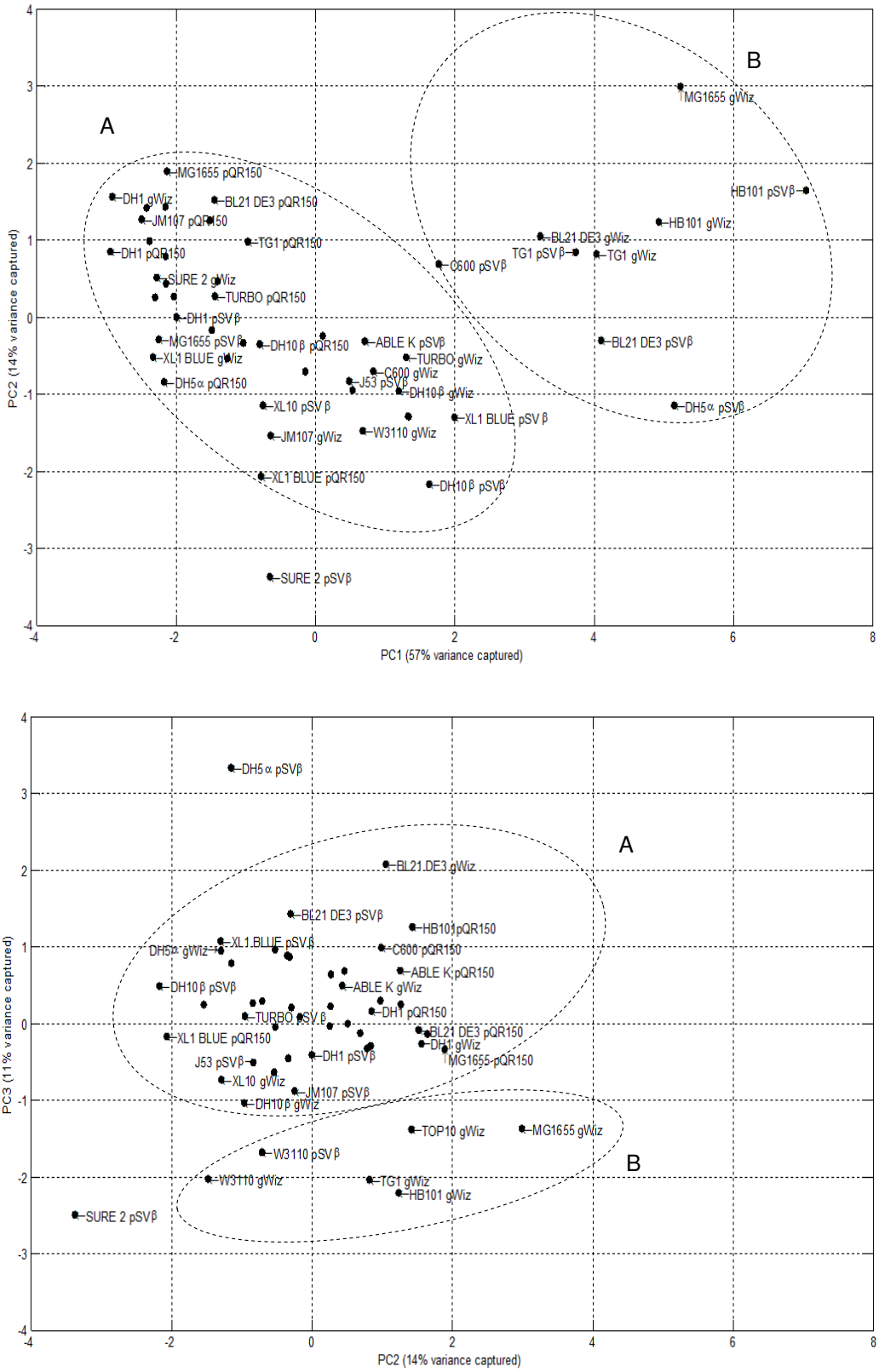


Figure 3.4.7.2. PCA scores plots showing PC1 vs.PC2 (top) and PC2 vs. PC3 (bottom). The main cluster circled (A) is where most of the strain-plasmid combinations are found, representing the general trend. The secondary cluster (B) is also circled.

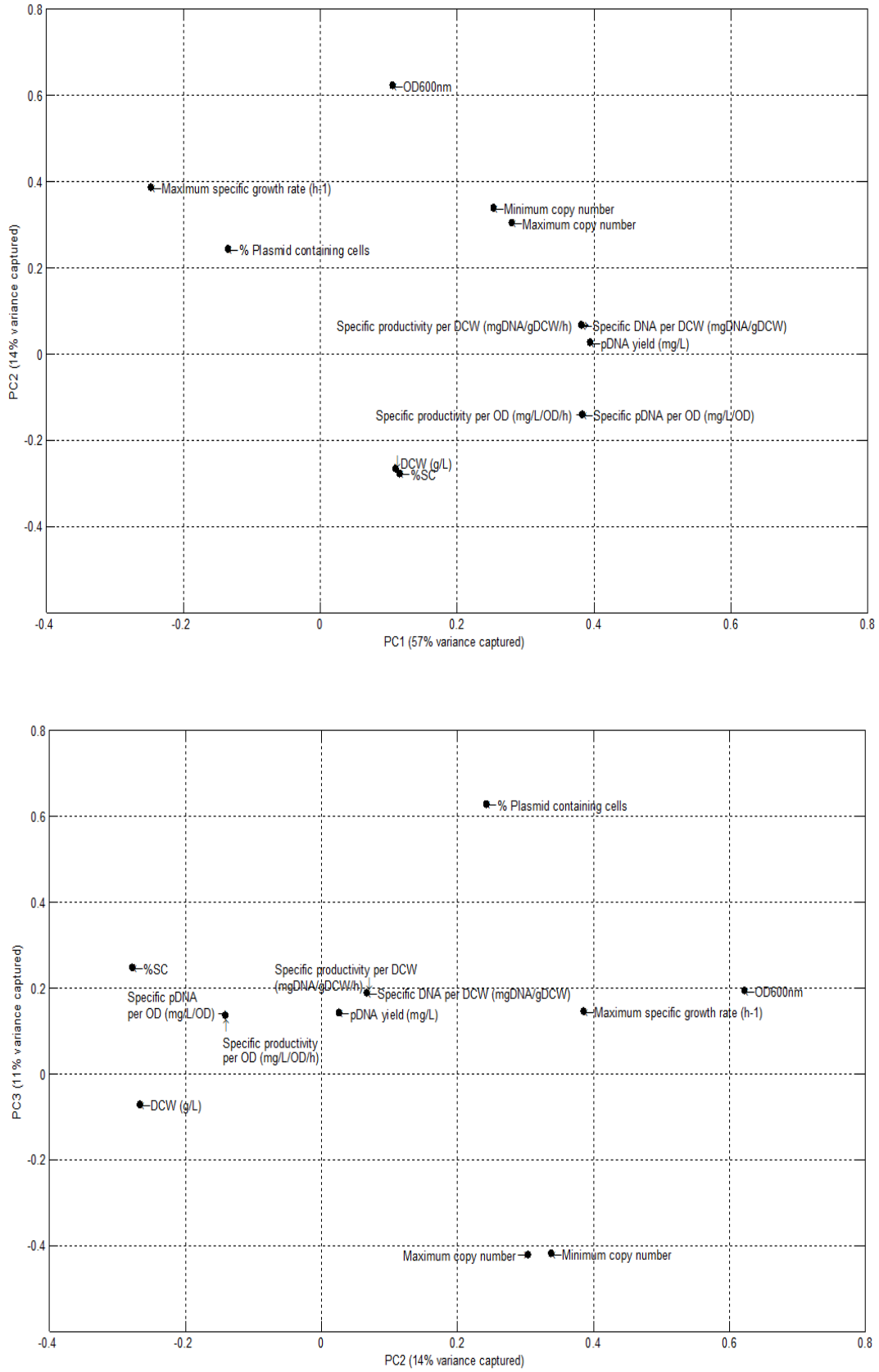


Figure 3.4.7.3. PCA plot of the loadings data from PC1 vs.PC2 (top) and PC2 vs. PC3 (bottom).

The results obtained show that when all of the selected factors were included, specific strain-plasmid were mostly affected by the plasmid pDNA yield, copy number values (maximum and minimum), OD and plasmid stability. These were the most influential factors when analysing the variation captured by the first 2 principal components, i.e. the most important two components. These highlighted influential factors and the %SC-DNA are considered to be good attributes when selecting a strain-plasmid combination for plasmid DNA production. In theory, one would expect a high plasmid DNA yield from a culture of high OD, copy number and plasmid stability, therefore the strain-plasmid combination that is most positively affected by these factors should be considered as a potential candidate for selection.

The presence of strain-plasmid combinations that do not follow the general trend (outliers) can highlight the likelihood of any experiment error in a data set or simply points of interest. In this study the few strains that stand out i.e. TOP10 gWiz, SURE 2 pSVβ and W3110 gWiz, all provided low SC-DNA percentages, low OD and low plasmid stability values. In terms of productivity, it would seem that these would be the candidates to avoid. Due to the complexities of host-plasmid interactions that are not yet fully understood, there may not be a simple reason as to why these plasmid-strain combinations performed poorly.

3.5. Conclusion

A systematic approach has been conducted to compare the performance of 17 commonly used *E. coli* strains. The compiled data strongly indicate that the choice of the host strain has a significant influence on plasmid DNA yield and quality. The accumulation of selected and unknown mutations during the last 85 years has made it very difficult if not impossible to define all the mutations in these strains and the effects of such mutations. It will be of great interest to see whether the same strains that produce high yields also do that on scaling up to bioreactors and whether changing the growth medium and mode of feeding (fed-batch vs. batch) operation can exert additional gains.

Intriguing observations regarding the possible occurrence of a type of runaway plasmid replication when pUC-based plasmids are used in certain strains may provide a completely new outlook on the concept of plasmid production. Future studies may confirm that there is a possibility that runaway replication or something similar to it, is not only plasmid dependant but also host strain dependant.

The results have also shown that PCA can be a very useful analysis and selection tool when applied to a relatively small and simple data set. Applying PCA in an investigation of this kind can quickly enable the selection of the best performing candidates or highlight interesting points early on in an investigation that may merit further analysis. Traditional techniques require analysis through direct visual inspection, usually involving many, sometimes complicated graphs and the advantage of PCA is that it can reduce the entire data set to just one or two important plots, with less bias on the data. Had PCA been applied earlier on, the same trends and correlations may have been obtained quicker, more easily and reduced the complexity of the data (and hence the number of graphs and tables) presented.

Importantly, not only is PCA specifically designed for analysis of very large data sets where manual data analysis will be too time consuming and

inefficient, the application of PCA to strain selection is novel. The loadings data provides useful information on the variables that most influence the performance (in this case the specific pDNA yield and percentage supercoiling) of a strain-plasmid combination.

The nature of the F-plasmid present in 7 of the 17 host strains should also be noted. The use of a conjugation-proficient F episome has long been discouraged by NIH guidelines (Yanisch-Perron et al., 1985) since there are certain phages (such as M13) that use the F pilus as an attachment point on *E. coli*. Although some consideration should be given to the use of strains containing the F-plasmid in order to mitigate the potential risk of phage infection, it is important to stress that only a few phages use the F pilus as an attachment point whereas the majority of phages use other parts of the cell surface such as TonB for T phages (Garen and Puck 1951), maltose uptake protein for lambda (Hazelbauer, 1975; Randall-Hazelbauer and Schwartz, 1973) and lipopolysaccharides for T5 phage (Heller and Braun, 1979).

4. Interaction between Host Strain Selection and Temperature Amplification on Plasmid DNA Production

4.1. Introduction

In Chapter 3, a few strains (HB101, MG1655, TG1 and W3110) harbouring pSVβ or gWiz, were observed to provide very high copy numbers, much higher than the reported copy number range for pUC-based plasmids (100-800). It was suggested in Chapter 3 that these strains displayed similar characteristics to what has been described as runaway plasmid replication, where the plasmid is hugely overproduced during replication (Remaut et al., 1983; Togna et al., 1993 and Uhlin et al., 1978 & 1979).

It has been reported that the copy numbers of some pMB1-derived vectors increase 30- to 40- fold at 42°C when compared to growth at 30°C, due to the presence of a temperature sensitive copy number mutant, *cop*^{ts}, (Wong et al., 1982). This phenomenon normally tends to be apparent in relaxed replicons with higher copy numbers such as the pUC based vectors (Molin et al., 1989; Reinikainen et al., 1989; Wong et al., 1982) caused by alterations in RNA I that affect the regulation of RNA II (Lin-Chao et al., 1992).

Throughout this investigation, references to the term “runaway replication” are used only to describe the observed increase in copy number to very large amounts beyond the 800 copies per cell. It may be caused by a type of unrestricted and uncontrolled replicon control mechanism sensitive to temperature based amplification, whereby when induced, normally through a temperature shift, excessive plasmid DNA production over and beyond the average copy number under “normal” conditions (in this chapter, “normal” is defined as growth at 37°C on Nutrient Broth and the copy number criteria for a pUC-based plasmid is a copy number of 800) is observed. The increase in plasmid copy number seen by virtue of the mutation in the pUC series that lessens the interaction of RNA I and RNA II at 40-42^{o}C (Lin-Chao et al., 1992), could be described as a stage on the way to runaway replication, since the complete removal of RNA I control would lead to uncontrolled runaway replication from the RNA II dependent

only on the RNA II promoter and the conversion of the RNA II primer to DNA.

It has been noted that runaway replication has also been reported to occur through other techniques such as decoupling protein synthesis and DNA replication through chloramphenicol amplification (Clewell, 1972; Frenkel and Bremer, 1986; Wegrzyn 1995) or amino acid starvation through the accumulation of uncharged tRNAs that interact with RNA I or RNA II formation (Hofmann et al., 1990; Wrobel and Wegrzyn, 1998; Yavachev and Ivanov, 1988; O'Kennedy et al., 2003).

Initial tests conducted on HB101, MG1655, TG1 and W3110 gWiz and pSVβ cultures have shown that some of these high copy number host-plasmid combinations also provide higher specific plasmid DNA yields at 40℃ than at 37℃ or 30℃, suggesting that the large increase in plasmid yield observed is not only plasmid dependant but also host strain dependant. Thus, following on from Chapter 3, this chapter will investigate temperature induced amplification of *E. coli* strains harbouring a pUC-based plasmid, to examine the combined effects of the host strain and temperature amplification on plasmid DNA production.

4.2. Aim and Approach

This study will look into the effects of temperature amplification by measuring the specific plasmid yield per OD of different host strain-plasmid combinations harbouring (individually) pUC18, pSVβ and gWiz when grown on Nutrient Broth at 30℃, 37℃ and 40℃.

Nutrient Broth was chosen to remove any potential effects from tRNA induced copy number increases seen in defined media where uncharged tRNAs occur in partial amino acid starvation (Hofmann et al., 1990; Wrobel and Wegrzyn, 1998; Yavachev and Ivanov, 1988; O'Kennedy et al., 2003).

4.2.1. Criteria for Temperature Induced Amplification

Vectors which exhibit runaway replication have been reported to replicate in a normal fashion at temperatures of up to 34°C but plasmid replication becomes uncontrolled at temperatures above 39°C (Sambrook et al., 2001). The criteria set out in this study to measure temperature induced amplification (similar to what some authors have described as runaway replication), is to observe a substantial increase in copy number (or specific plasmid yield per OD) in cultures grown at 40°C than in cultures grown at 30°C under the same conditions (except temperature).

It has been reported that some vectors increase 30- to 40- fold at 42°C when compared to growth at 30°C (Wong et al., 1982), however this rate of increase was observed under different conditions to the conditions set in this chapter (Wong et al., 1982, used *E. coli* strain DG75 harbouring a ColE1-derived plasmid, pEW2762, grown on Vogel-Bonner medium containing 0.2% glucose and 100µg/mL ampicillin at 30°C which was shifted to 42°C when the culture reached an OD_{650} of 0.3; the culture scale was not mentioned by the authors of this report).

pUC plasmids were originally derived from pBR322 (pMB1 origin), as a result of a single point mutation within the replication primer RNA II. This C to T mutation on the origin of replication (when compared with pBR322) alters the secondary structure of the RNA II primer, decreasing the interaction of RNA II with RNA I, leading to an increase in copy number that is temperature dependant, i.e. the copy number is increased at 42°C but is normal at 30°C (Lin-Chao et al., 1992).

In this chapter pUC18 was transformed into the 17 strains listed in Table 3.1. in Chapter 3, including another K-12 wild type strain W1485 (ATCC 12435) (λ^-, F^+) (Jensen, 1993), the parent strain of MG1655 and W3110. A few gWiz and pSVβ harbouring strains (W3110, MG1655, TG1, HB101, SURE 2 and J53) from Chapter 3 were also investigated to confirm earlier observations. 10 mL cultures were grown on the same broth as in Chapter 3 (Nutrient Broth CM0001), in 50mL BD Falcon tubes at 30°C, 37°C or

40℃, with shaking at 250rpm for 16 hours (incubated vertically). These temperatures were checked to be correct using independent thermometers placed inside the shakers.

4.2.2. Issues with Enumerating Bacterial Cells

A temperature amplification strategy was implemented to increase the plasmid copy number as reported by Uhlin et al., (1978) and Wong et al., (1982). Although the specific yield of pDNA per cell can be used as an approximate indication of plasmid copy number, to compare with previous literature, the plasmid copy number per cell should also be measured to correct for increases in cell numbers. However during this study it became clear that this would not be straightforward partly due to the difficulty in accurately enumerating bacterial cells.

The method used to enumerate bacterial cells in Chapter 3 (serial diluting before plating onto agar plates), although used quite often to count bacterial cells, results in only viable cell counts and this may not reflect the total number of cells in the culture. If the cells in some of the cultures were entering or have entered the death phase, then they may not all form colonies, resulting in a lower number of cells counted on the plates than the actual total number of cells. Consequently this would lead to higher copy number per cell determination if these dead or dying cells contained pDNA. Lee et al., (1994), reported that with *E. coli* cells producing poly(3-hydroxybutyric acid) (PHB), the number of colony forming units (CFUs) is found to be significantly lower compared to the true total cell count as determined by a haemocytometer. The report suggests that as the cells accumulate PHB with time, once they contain a critical amount of PHB (not determined in that study), the cells did not form colonies.

Although plasmid DNA is not deemed to be as metabolically demanding or toxic to the cell as the polymer PHB, high yields of plasmid DNA may have had some effect on the cells viability during stationary phase (when the cultures were sampled in Chapter 3) and their ability to form colonies. Initial attempts at testing this theory were unsuccessful as the

haemocytometer method used to enumerate bacterial cells resulted in variable readings, often with lower cell counts than the plate cultures of the same sample. Since haemocytometers were designed for use with much larger blood cells and the chambers are made to "fit" the width of a blood cell, inaccurate measurements were obtained when using a haemocytometer to count *E. coli* cells under a microscope.

Other techniques reported to have been used to enumerate bacterial cells include using Coulter counters, flow cytometry and specialised flow cytometry systems such as the FACS machine. Some of these techniques are aided by the use of fluorescent labels and probes which are used to measure culture viability.

Electric counters, such as the Coulter counter, have traditionally been used to count bacterial cells (Carrillo et al., 1985; Kubitschek and Friske, 1986; Harvey and Marr, 1966; Jaffe et al., 1985). This technique determines the particle size by measuring the change in resistance as a particle moves through a small channel filled with electrolyte. The problem with this method for bacterial cell enumeration is that these electrical counters cannot screen samples showing microbial heterogeneity.

Advances in standard flow cytometry based techniques, commonly used for mammalian cell screening, have improved this technique for bacterial counting applications (Shapiro, 2000; Steen, 2000). Flow cytometry counts and examines cells in a stream of fluid based on the physical and/or chemical characteristics of single cells flowing through the detection apparatus. The combination of scattered and fluorescent light is used to analyse the physical and chemical structure of each particle.

Fluorescent-Activated Cell Sorting (FACS) (Becton Dickinson and Company, Franklin Lakes, NJ, USA) is a specialised type of flow cytometer which can sort through a sample containing a heterogeneous mixture of cells. This system has been reportedly used effectively with high sensitivity for bacterial cell sorting (Francisco et al., 1993; Fu et al.,

1999). The FACS system uses a vibrating mechanism to cause the stream of cells break into individual droplets. The system is adjusted so that a droplet will most likely contain only one cell. The fluorescence characteristics of the droplets are measured before a particular charge is placed on each droplet. The charged droplets then fall through an electrostatic deflection component that sorts the droplets into containers based upon their charge.

Another method for bacterial enumeration is the use of the nucleic acid stain Acridine Orange (3,6-bis[dimethylamino]acridinium chloride) in fluorescent direct counts (Kepner and Pratt, 1994). Acridine Orange (AO) is a nucleic acid selective metachromatic stain commonly used in flow cytometry for cell cycle determination and can be used to measure DNA-RNA content (Darzynkiewicz, 1990; Myc et al., 1992). AO interacts with DNA and RNA by intercalation or electrostatic attraction respectively. When excited by blue light, AO fluoresces green (515 to 575nm) when intercalated (as a monomer) in double-stranded nucleic acids and red due to dye-dye interaction when bound to single-stranded nucleic acids. RNA electrostatically bound AO fluoresces red (600 to 640nm). AO does not differentiate between live and dead cells. Fixed or preserved cells are stained with AO, followed by filtration and direct counting of bacterial cells on the filter surfaces.

Unfortunately due to the lack of access to these devices, the discussed alternative techniques for bacterial enumeration and/or sorting could not be tested during the course of this investigation. As a result, this investigation will focus on temperature induced amplification of specific plasmid DNA yield per OD as an indicator of copy number rather than calculating the copy number directly.

4.3. Results and Discussions

4.3.1. pUC18 Cultures

The general trend of the results obtained in Figure 4.3.1.1. show that most of the strains provided higher specific yields of pUC18 pDNA per OD at 40℃ than at 30℃ or 37℃, especially the strains that resulted in very high copy numbers for the plasmids gWiz and pSVβ (HB101, TG1 and MG1655) which provided between 4- to 6- fold copy number amplification between 30℃ and 40℃. Other strains that resulted in high copy number amplification between 30℃ and 40℃ are C600, TURBO, JM107, and DH10β, with an 8-, 6-, 5- and 5- fold copy number amplification respectively. There were however, a few exceptions to this trend such as the strains W3110, W1485 and ABLE K. These three strains provided higher specific pDNA yields per OD at 37℃, with ABLE K providing the highest specific pDNA yields per OD of pUC18 out of all 18 strains at 37℃.

When ABLE K pUC18 was scaled up to 100mL working volume shake flask culture, the results show that more plasmid was obtained at an even lower temperature (30℃). The ABLE K strain is based on a C strain of *E. coli* crossed with a DNA polymerase I deficient strain of a K-12 strain and is designed to enhance the probability of retrieving clones that are toxic to *E. coli* through the reduction of the plasmid copy number of ColE1-based replicons (Stratagene).

It is unclear if the ABLE K *pol* mutation has resulted in a temperature sensitive strain, however a study carried out by Christie et al., (1995), have reported that *E. coli* C strains have a different nucleotide sequence of the *rpoD* gene (encoding the primary σ^{70} subunit of RNA polymerase essential for regulating transcription initiation) to K-12 strains. Mutations in the *rpoD* gene can lead to a temperature sensitive growth phenotype (sensitive to temperatures above 40-43.5℃) (Grossman et al., 1985) and may explain why the level of cloned gene products from ABLE K is higher at lower temperatures.

The wild type strains of *E. coli* K-12, MG1655, W3110 and W1485, performed variably with the latter two strains providing more specific pDNA per OD at 37℃. However it has to be noted that the W3110 pUC18 cultures also provided the largest variation in data from independent triplicate cultures, suggesting that this host-strain plasmid combination is quite sensitive to environmental conditions. W1485 is the parent strain of W3110 and MG1655 and as there are very few mutational differences in the genotypes between these strains, it is difficult to pinpoint why these strains performed differently. This observation again emphasises the need to assess each host strain-plasmid combination before generalising on the outcome of mutations.

In contrast, it is also important to note that the specific productivity and specific yields of many relaxed plasmids is inversely proportional to the growth rate (Seo and Bailey, 1985). When the growth rate of the host cell decreases, the plasmid population increases which can result in high copy numbers, which may or may not respond to temperature induced amplification. The growth rates of the 10mL cultures were not measured in this chapter, so it cannot be determined if the observed increases in copy number are growth rate independent.

The plasmid yields and ODs in this chapter were all measured after 16 hours of growth and considering the volume of the Nutrient Broth cultures (10mL), it is reasonable to assume that these cultures were in the stationary phase of growth during sampling. During the stationary phase, the growth rate slows as a result of nutrient depletion and plasmid copy numbers have been shown to increase during this phase (Fitzwater et al., 1988).

Taking these reports into consideration, there is a possibility that besides the increase in temperature, the copy numbers measured in this chapter may have been affected (in conjunction with or not) by the increased growth rate during the exponential phase or the decrease growth phase

during the stationary phase. *E. coli* cultures grown at lower temperatures tend to have lower growth rates, thus higher copy numbers would be expected from the 30℃ cultures than at 37℃ or 40℃, however, this was not observed at 10mL scale. As the growth rates of the 10mL cultures were not measured in this chapter, it cannot be determined whether the growth rates at 40℃ were lower than at 37℃.

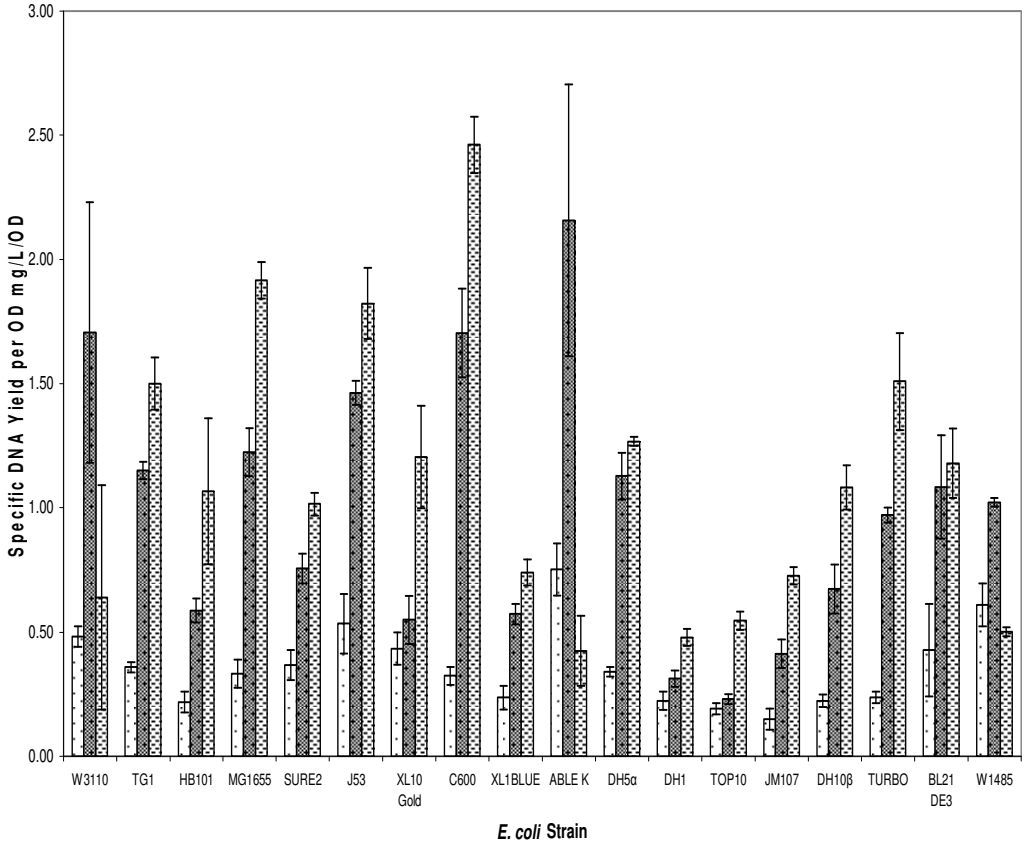


Figure 4.3.1.1. Specific plasmid DNA Yield per OD from 10mL cultures of *E. coli* strains propagating pUC18 grown on Nutrient Broth CM0001 at 30℃, 37℃ & 40℃ for 16 hours at 250rpm (incubated vertically). The error bars show the standard error from independent triplicate readings for each strain-plasmid combination (duplicate readings for W1485). Strains W3110 and ABLE K provided the largest variation in data. With the exception of ABLE K, W3110 and W1485, all of these pUC18-harbouring strains performed better at 40℃ in terms of specific plasmid DNA yield per OD.

4.3.1.2. 100mL W3110 pUC18 cultures

Preliminary results from W3110 pUC18 cultures grown at 100mL scale, do however confirm that an increase in growth rate has a negative effect on the specific pDNA yield, but this was only observed for the culture grown at 30°C and not at 37°C or at 40°C. Figure 4.3.1.2. (A) shows the specific pDNA yields per OD and growth rate profiles of three 100mL working volume shake flask cultures of W3110 that were grown on Nutrient Broth at 30°C, 37°C and 40°C at 250rpm for 8 hours. The results show that the growth rates did not differ significantly between the cultures grown at 37°C and at 40°C, even though there was generally more specific pDNA per OD obtained from the culture grown at 40°C. However in comparison to these two cultures, the growth rate profile for the culture grown at 30°C was generally lower, except between 6 to 8 hours of growth which coincides with the period when the specific pDNA yields per OD obtained from the culture grown at 30°C were at their lowest.

In a separate set of experiments (Figure 4.3.1.2. (B)), three more 100mL shake flask cultures of W3110 pUC18 were grown on Nutrient Broth at 30°C, 37°C and 40°C at 250rpm; however these cultures were not sampled until the period between 11 to 16 hours of growth. This sampling period took place during the late stationary phase of growth as indicated by the levelling of the ODs (and thus growth rates) measured. Similar to the 10mL cultures of W3110 pUC18, the specific pDNA yields per OD obtained from these 100mL cultures were also higher from the cultures that were grown at 30°C and 37°C, which is more likely to be a result of the cultures being in the stationary phase of growth.

However it should be noted that these two sets of 100mL W3110 pUC18 cultures have not been corroborated by replicates, thus appropriate caution should be taken when considering this data. The Linest function in Excel was used to calculate the standard error in the data by using the "least squares" method to calculate an array that best fits the data.

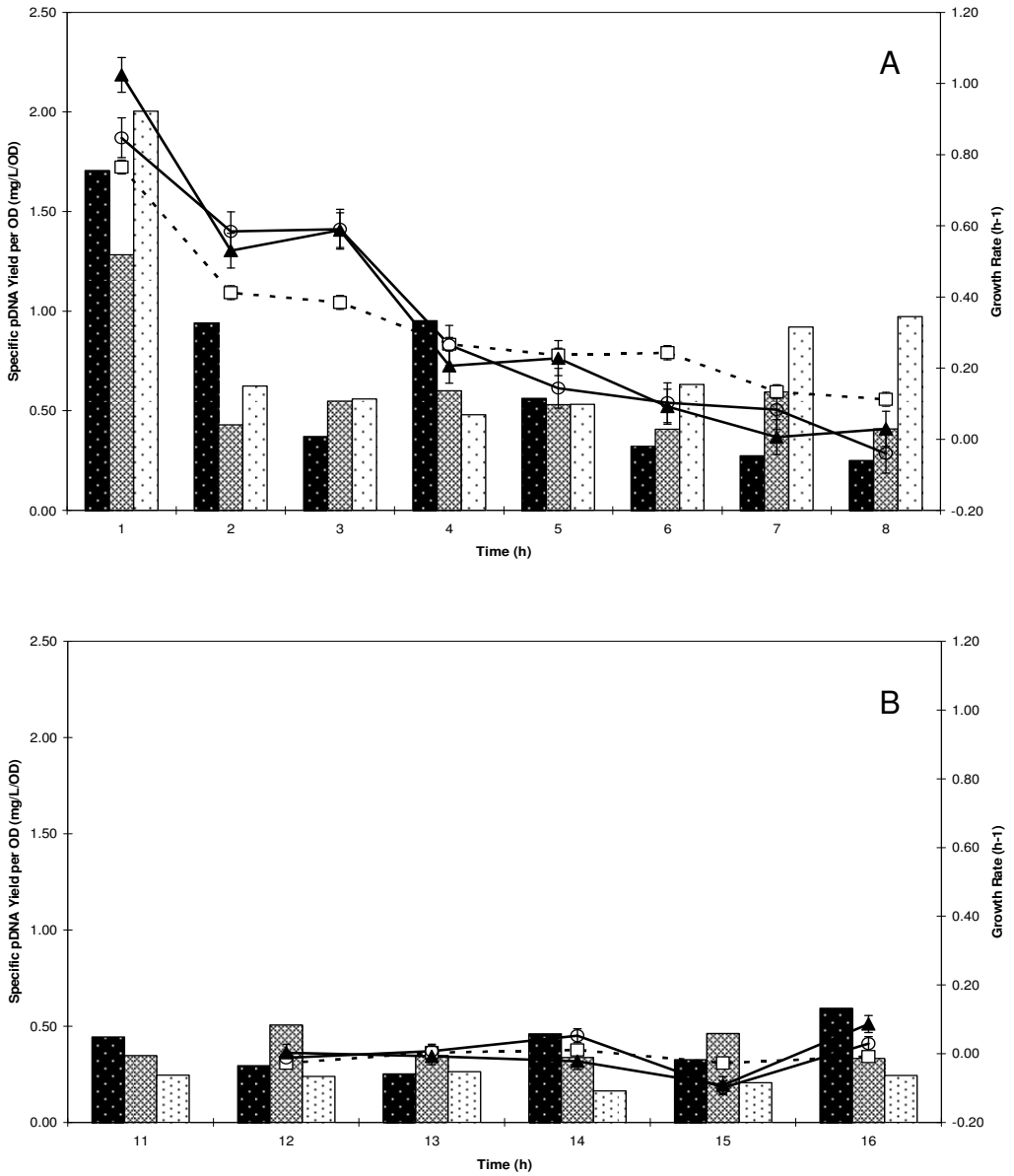


Figure 4.3.1.2. Specific plasmid DNA yield per OD and growth rate profiles from 100mL working volume cultures of W3110 pUC18 grown on Nutrient Broth at 30℃, 37℃ & 40℃ for 8 hours (A) or and in a separate experiment for 16 hours (B), at 250rpm in 500mL shake flasks. Error bars indicate the standard error calculated using the Linest function in Excel. Abbreviation: 'dC' = degrees Celcius.

4.3.2. pSVβ and gWiz Cultures

As mentioned in Chapter 3, when a few *E. coli* gWiz and pSVβ strains were cultured at 30℃, 37℃, and 40℃, the strain-plasmid combinations that had very high copy numbers also tended to produce high specific yield of pDNA per OD at 40℃. The strain-plasmid combinations of HB101 gWiz and HB101 pSVβ and J53 gWiz achieved almost 12-, 7- and 6- fold amplification between 30℃ and 40℃ (Figure 4.3.2.1.). However, J53 pSVβ, SURE 2 pSVβ and SURE 2 gWiz did not achieve temperature induced amplification at 40℃ when compared to cultures grown at 30℃.

Again since no growth rates were measured for these 10mL cultures there is perhaps some correlation between growth rate and specific pDNA yield however, this cannot be determined by these experiments. Nonetheless, 100mL cultures of these strain-plasmid combinations conducted in Chapter 3, showed that some strain-plasmid combinations did not appear to be negatively affected by fast growth rates and still achieved high specific pDNA yields (e.g. MG1655 gWiz).

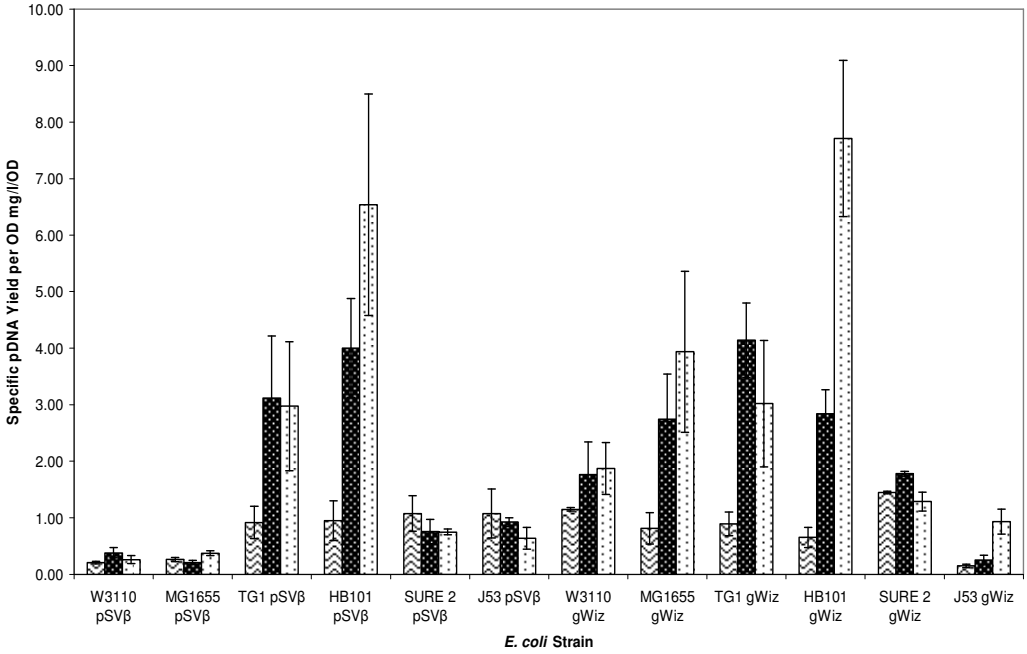


Figure 4.3.2.1. Specific pDNA yields per OD obtained from selected *E. coli* strains harbouring plasmids gWiz and pSVβ. Results are from 10mL cultures grown on Nutrient Broth CM0001 at 30℃, 37℃ & 40℃ for 16 hours at 250rpm. The error bars show the standard error from independent triplicate cultures for each strain-plasmid combination.

Growth at 30 degrees Celcius Growth at 37 degrees Celcius Growth at 40 degrees Celcius

4.3.3. Differences in morphology of *E. coli* strains

Since initial attempts at enumerating the cells using a haemocytometer were unsuccessful, accessible alternative methods were sought to attempt to see if there were any differences in the host strain-plasmid combinations that resulted in very high copy numbers. There have been reports (Lee at al., 1994; Rinas et al., 1993) that cell stress from the overexpression of some recombinant products can result in changes in the physical appearance of the host cells. Lee et al., (1994), observed the elongation of the *E. coli* cells during the accumulation of PHB in the cytoplasm as granules, distorting the normal cell shape as observed under a microscope. Another study has also reported cell elongation of *E. coli* when overexpressing a recombinant protein product (Rinas et al., 1993). With these reports in mind, this section will investigate whether there are differences in the physical appearance of the some strain-plasmid combinations as a result of their very high copy number.

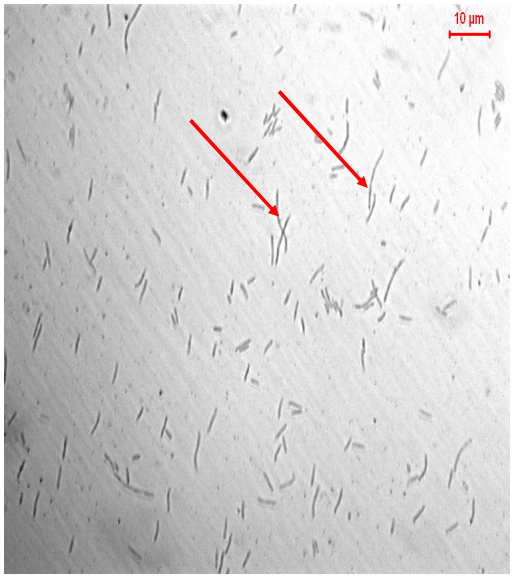
Attempts were made to view a few of the high copy number strain-plasmid combinations under a microscope and compare these with cells containing no plasmid or one with a lower copy number. These cultures were all grown at 37℃ (Table 4.3.3.1.). Microscopic analysis on samples of the cultures grown at 30℃ or 40℃ were not conducted, thus this section will focus on the possibility of cell elongation due to the high plasmid copy number of certain strain-plasmid combinations at 37℃ and not on cell elongation due to temperature induced plasmid amplification (the latter of which has already been reported in literature). Rinas et al., (1993) have reported observing abnormal *E. coli* cell elongation in response to a temperature shift (from 30℃ to 42℃) for cells bearing a temperature-sensitive runaway replicon that is not encoding a recombinant protein product. Wong et al., (1982), have also observed cell elongation of a runaway vector upon temperature shift from 30℃ to 42℃ that was not observed in cells at 30℃ or in control (non-temperature-sensitive) samples at 42℃.

Table 4.3.3.1. A list of *E. coli* strains or strain-plasmid combinations prepared for microscopic analysis at 1000x magnification

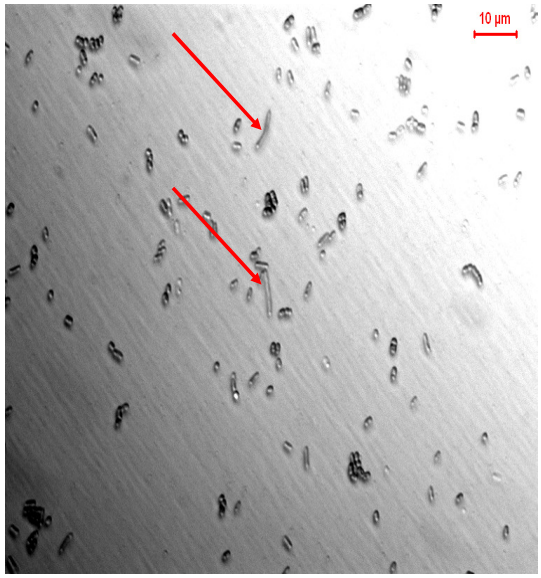
***E. coli* Strain/ Strain-plasmid combination**	**Copy number (as determined in Chapter 3)**	**Growth conditions**	**Approximate % elongated cells from microscope image***
HB101 gWiz	High (>1000)	10mL cultures grown vertically in 50mL BD Falcon tubes on Nutrient Broth at 37°C, 250rpm for 18 hours	24%
MG1655 gWiz	High (>1000)		3%
TG1 gWiz	High (>1000)		3%
ABLE K gWiz	Low (≈100)		0%
HB101 (no plasmid)	N/A		2%
MG1655 (no plasmid)	N/A		0%
TG1 (no plasmid)	N/A		0%
ABLE K (no plasmid)	N/A		0%
Control (spun down, cell-free supernatant and methylene blue sample)	N/A		N/A

*Cells that appear to be longer than twice the length of 'normal' cells (i.e. non-transformed cells) were counted as 'elongated' and this number was divided by the total number of cells to obtain the % of elongated cells per image. Please note that these are approximate values based on the field of view shown in Figures 4.3.3.1. to 4.3.3.4.

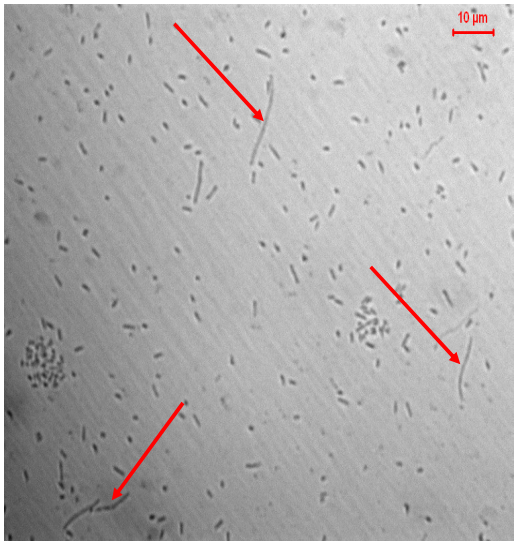
When viewed under the microscope, heterogeneity was observed for the strain HB101 as some cells were longer than others (both when transformed with gWiz and also in its plasmid-free state) (Figure 4.3.3.1.). Elongated cells were observed for MG1655 gWiz (Figure 4.3.3.2.), but not in its non-transformed, plasmid-free state suggesting that the accumulation of a large concentration of plasmid DNA maybe causing the cells to elongate. TG1 gWiz cells were also elongated and generally wider than TG1 plasmid-free cells (Figure 4.3.3.3.). These three strain-plasmid combinations were found to have copy numbers >1000 in Chapter 3 using the serial dilution and plating method. The lower copy number commercial strain ABLE K looked the same in its transformed and non-transformed, plasmid-free state (Figure 4.3.3.4.).



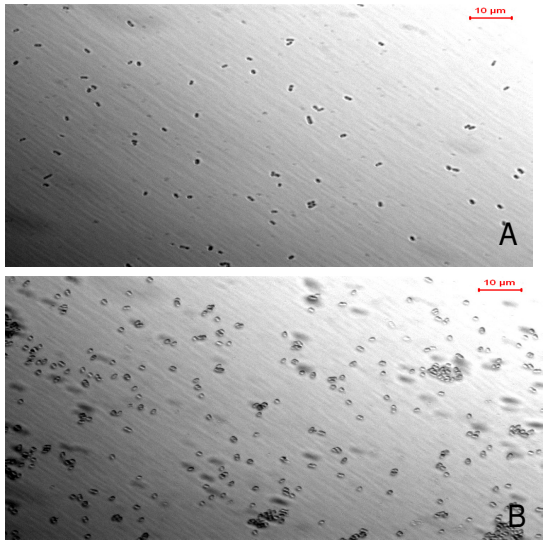
HB101 gWiz



HB101 (non-transformed)

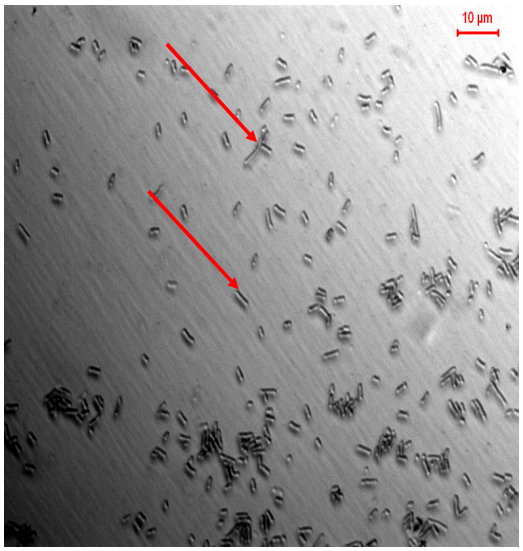


MG1655 gWiz

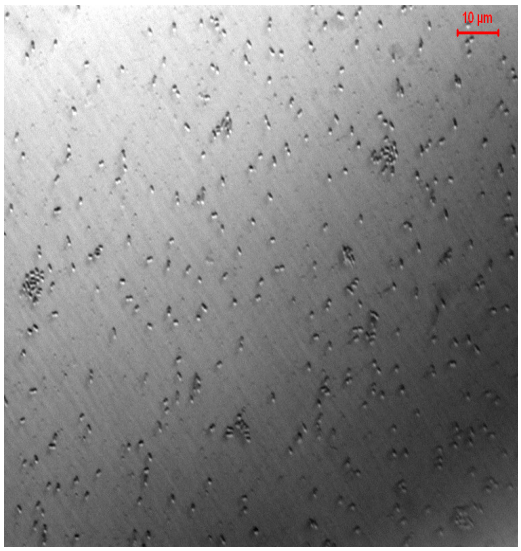


(A & B) MG1655 (non-transformed)

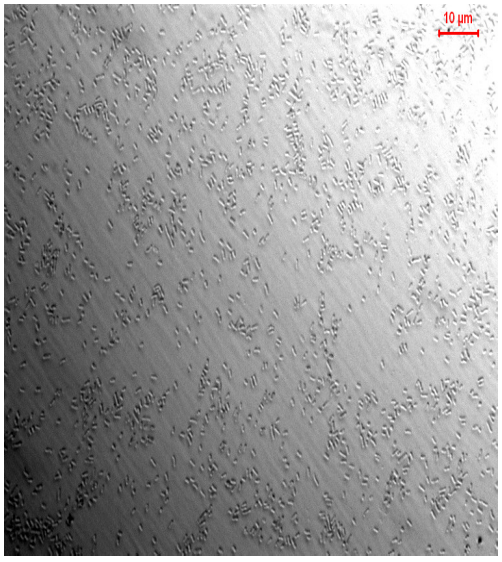
Figures 4.3.3.1. & 4.3.3.2. Microscope images of HB101 (Figure 4.3.3.1.) and MG1655 (Figure 4.3.3.2.) cells in their transformed, plasmid containing and non-transformed, plasmid-free state. Viewed under 1000x magnification. The cells were grown in 10mL of Nutrient Broth in BD Falcon tubes for 18 hours at 37℃, 250rpm (incubated vertically). Arrows indicate what appear to be elongated cells present in both cultures of HB101 plasmid-containing and plasmid-free cells, but only in cultures of MG1655 plasmid-containing cells.



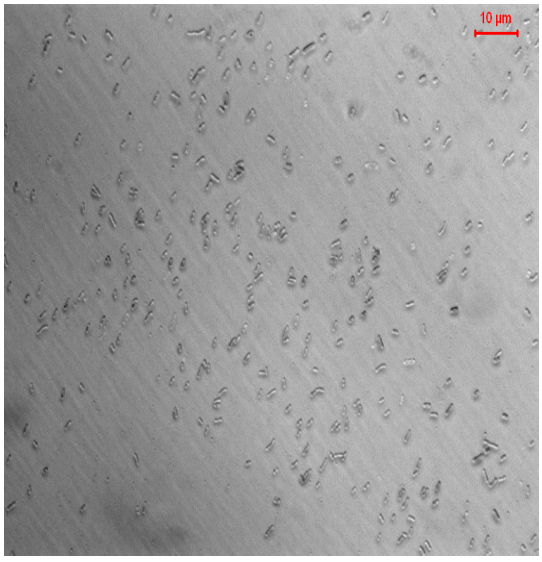
TG1 gWiz



TG1 (non-transformed)



ABLE K gWiz



ABLE K (non-transformed)

Figures 4.3.3.3. & 4.3.3.4. TG1 (Figure 4.3.3.3.) and ABLE K (Figure 4.3.3.4.) cells in their transformed, plasmid-containing and non-transformed, plasmid-free state. Viewed under 1000x magnification. The cells were grown in 10mL of Nutrient Broth for 18 hours at 37℃, 250rpm in BD Falcon tubes (incubated vertically). Arrows indicate what appears to be TG1 plasmid containing cells wider in diameter than plasmid-free cells. ABLE K cells did not appear different in its plasmid-containing or plasmid-free state.

It has been reported that *E. coli* cell sizes can decrease during the exponential phase of growth (Steen, 1990). Cell size changes can also result in response to changes in the growth rate (Trueba and Woldringh, 1980). Generally, cell sizes decrease at lower growth rates and under nutrient limitations to increase the surface to volume ratio in order to improve nutrient flux through the cell surface and non-growth-dependent energy consumption (for maintenance). In batch cultures various physiological states during the different growth phases result in sub-populations exhibiting different cell sizes.

The number of cells determined by measuring the optical density, OD, is based on the assumption that the size distribution of the cells (light scattering) remains constant, since the OD is the product of cell number and cell size as measured by light scattering. However this assumption does not hold true as cells from one phase of growth are not in the same state as cells from another phase and as discussed, different growth rates can affect the size of the cells, resulting in under- or overestimation of the cell number. To mitigate this potential problem of the changing OD to dry cell weight correlation, all the cultures were kept growing until they were in the stationary phase to achieve homogeneity and it is at this stage of growth when samples from all cultures were taken. Samples were immediately heat fixed onto microscope slides before being stained with methylene blue and viewed under the microscope.

Other than the study carried out by Lee at al., (1994), changes in the morphology of *E. coli* can be due to other factors and these have been well documented by other authors as shown in Table 4.3.3.2.

Table 4.3.3.2. Changes in cell morphology and the reported cause of the change

Change in *E. coli* cell morphology	**Reported cause of change in cell morphology**	**Literature reference**
Cell elongation	Accumulation of PHB in the cytoplasm as granules	Lee et al., (1994)
Cell elongation	Response to a temperature shift (from 30℃ to 42℃) for cells bearing a temperature-sensitive runaway replicon	Rinas et al., (1993); Wong et al., (1982)
Cell elongation	Overexpression of GrpE protein, affecting the molecular chaperone DnaK	Sugimoto et al., (2008)
Cell elongation and abnormal cell division	Inclusion bodies	Bowden et al., (1991); Hart et al., (1990); Lee et al., (2008); Williams et al., (1982)
Spherical morphology	Effects of cell envelope synthesis on mutants of *E. coli* K-12	Tamaki et al., (1980)
Inhibited cell elongation	Effects of β-lactam antibiotics	Spratt (1975)

4.4. Conclusion

It would appear that the choice of *E. coli* host strain and not just the type of plasmid has a sizeable influence on whether plasmid amplification due to increased temperature is successful. Indeed many of the host plasmid combinations exhibited a temperature induced plasmid amplification response between 30℃ to 40℃. However, there were a few strains that did not follow this trend such as W3110 and ABLE K. These results have again emphasised that the known genotype alone cannot determine how certain strains will behave, reiterating the need to assess each host strain-plasmid combination before generalising on the outcome of mutations.

As the growth rates of the 10mL cultures were not measured, it is impossible to decouple the effect of growth rate on the specific yield and the effects of temperature amplification. The extent of the accumulation of plasmid DNA due to the effects of the growth rate and/or temperature during the exponential phase of growth is unknown at this time. Preliminary results from 100mL W3110 pUC18 cultures do show that the growth rate profile for the culture grown at 30℃ was generally lower than those obtained from the 37℃ and 40℃ cultures, except between 6 to 8 hours of growth, when the growth rate for the 30℃ culture was higher. This coincides with the period when the specific pDNA yields per OD obtained from the culture grown at 30℃ were at its lowest. However, replicates are needed to confirm this initial observation and above all, the growth rates need to be measured for other strain-plasmid combinations that achieved very high copy numbers.

Microscope images of plasmid-containing and plasmid-free cells show that some host strain-plasmid combinations appear to exhibit morphological response (cell elongation and increase in diameter) which may or may not be associated with higher plasmid copy number. Regardless of the true cause of these morphological changes, these images do point out that each *E. coli* strain can be different in size and morphology; therefore it is highly recommended that assumptions made based on biomass (OD or

DCW), are made with caution and should be correlated to a specific number of bacterial cells between different *E. coli* strains.

Similarities to what has been reported as temperature induced runaway plasmid replication have been observed in this study, although no experiments were conducted to confirm whether these observations were indeed a result of runaway replication as defined in the literature. To do so, the characteristics of runaway replication need to be clearly defined as do the characteristics of "normal" replication. Additionally, an accurate method of copy number determination per cell is needed, for example, the use of Acridine Orange staining where fixed or preserved cells are stained with AO, followed by filtration and direct counting of bacterial cells on the filter surfaces (Kepner and Pratt, 1994).

5. The Effect of Anaerobiosis on Supercoiled Plasmid DNA Production

5. Strategies to Improve pDNA Yields and DNA Supercoiling

5.1. Introduction

Chapter 3 investigated the growth and productivity of a number of plasmid-strain combinations grown on complex Nutrient Broth (Oxoid CM0001) at 10mL and 100mL scale. From this initial investigation, it was observed that a few strain-plasmid combinations performed noticeably better than their counterparts at providing high yields of plasmid DNA and SC-DNA, for example, when harbouring the plasmid gWiz, the strains MG1655, HB101, TG1, BL21 DE3 and DH5α produced the most supercoiled plasmid DNA yield. The data obtained in Chapter 3 has clearly shown that the choice of host strain has a significant impact on the quality and quantity of plasmid DNA obtained and for any process, the early screening of host-plasmid performance before further large scale optimisation or verification takes place is highly recommended. Following on from this the next stage of this investigation is to look at how two of these host strain-plasmid combinations perform at a larger scale, such as a 7L fermenter, with a 5L working volume.

In Chapter 3, the cultures were grown on a complex medium, Nutrient Broth CM0001 (recipe mentioned in Chapter 2, Table 2.1.3.1.). Nutrient Broth (NB) is not as rich in nutrients as other well-known complex media, such as Terrific Broth (TB) or LB broth, and this may explain why initial 7L fermentation runs using Nutrient Broth resulted in poorer *E. coli* growth in terms of cell density than what has been observed when using TB or LB media. Complex media such as LB, NB and TB are unsuitable for commercial manufacturing due to the animal-derived components present, however complex media recipes are often used as an easy route to high cell density batch cultures. The limiting factors associated with the use of complex media are: the cost of non-animal derived raw products; oxygen transfer limitations upon scale up; cooling limitations due to the metabolic heat generated from high cell density cultures; batch to batch variations in nutrient composition; foaming caused by proteins present in the complex

media and higher biological oxygen demand (BOD) in terms of waste treatment (Junker 2007; Lee, 1996; Losen et al., 2004; Shojaosadati et al., 2008).

Defined media have become more popular for the production of biological products at an industrial scale since the nutrient compositions are known and can be controlled (Lee, 1996; van Hoek et al., 2000) especially when optimised to achieve production levels equal or greater than those achieved with complex media (Zhang and Greasham, 1999); however it is sometimes necessary to boost product formation and this can be done by choosing a semi-defined media. Studies have reported that some semi-defined media such as SDCAS and SDSOY were found to support higher cell densities and yield higher plasmid DNA and SC-DNA levels than complex media such as LB (O'Kennedy et al., 2000).

Since SDCAS medium (based on Hellmuth et al., (1994), with added casamino acids) has been successfully implemented for fermentation cultures in previous studies (O'Kennedy et al., 2003; Kay et al., 2003), this is the medium of choice of any further scale up work reported here. However, it is important to note that as SDCAS medium contains materials of animal origin (casamino acids), SDSOY would be the more acceptable choice of medium for industrial plasmid DNA manufacturing.

To date, several studies have reported on different conditions that affect the level of DNA supercoiling such as osmotic stress or salt shock (Conter, 2003; Hsieh et al., 1991), nutrient upshift or downshift (Balke and Gralla, 1987; Jensen et al., 1995), decreasing temperature (Goldstein and Drlica, 1984) and the presence of certain mutations such as *gyrB, topA10* or *ΔtopA* (Drlica, 1992; Sternglanz et al., 1981; Pruss et al., 1982). Another example is anaerobic shock or anaerbiosis (Cortassa and Aon, 1993; Hsieh et al., 1991; Dorman et al., 1988) which is said to increase DNA supercoiling (plasmid and chromosomal DNA).

Anaerobiosis has been reported to increase DNA supercoiling due to its effects in decreasing the activity of topoisomerase I (Cortassa & Aon, 1993) or increasing the [ATP] / [ADP] ratio (Hsieh et al., 1991 a&b), however these studies have only been conducted at shake flask scale, using rudimentary methods to implement and maintain anaerobic conditions. It would therefore be interesting to see if an increase in % pDNA supercoiling is observed under anaerobic conditions at fermentation scale. If successful, this application could have significant implications on plasmid DNA production as it would be an easy way to increase plasmid DNA supercoiling.

5.2. Aim and Approach

In this chapter, the impact of anaerobiosis on supercoiled plasmid DNA production was investigated at 7L scale (5L working volume). Firstly, the performance (yield of plasmid DNA and SC-DNA obtained) of 2 strains harbouring gWiz when grown on antibiotic containing (selective) SDCAS medium was evaluated at 5L working volume. Based on the findings in Chapter 3, the strains of *E. coli* harbouring gWiz selected were HB101 and BL21 DE3. HB101 gWiz was selected since it provided high percentage of plasmid DNA supercoiling (99%) and thus it would be interesting to see if this is maintained at 5L scale. BL21 DE3 gWiz was selected as it provided high yields of volumetric and specific pDNA (5.9mg/L and 2.9mg/L/OD respectively at 100mL scale) but low plasmid DNA supercoiling (77%). BL21 DE3 gWiz would be the strain-plasmid combination selected to undergo anaerbiosis, since any improvement in % plasmid supercoiling as a result of anaerobiosis, would be more evident from a culture that provides a lower level of plasmid supercoiling under normal aerobic conditions.

As osmotic stress or salt shock with NaCl has been reported to increase supercoiling (Conter et al., 2003 and Hsieh et al., 1991a), to ensure that anaerobiosis is the cause of any increase in supercoiling and not the effect of increased Na^+ ions, ammonium hydroxide was used as the base to adjust the fermentation pH instead of sodium hydroxide.

5.3. Results and Discussion

5.3.1. 7L Batch Fermentations Results

The results obtained from the 7L (5L working volume) fermentations of HB101 gWiz and BL21 DE3 gWiz show that HB101 gWiz (Figures 5.3.1.1., 5.3.1.2. and Table 5.3.1.1.), provided 12mg/L of pDNA (volumetric yield) and 0.70mg/L/OD of specific pDNA per OD. BL21 DE3 gWiz (Figures 5.3.1.3., 5.3.1.4. and Table 5.3.1.1.) provided even more pDNA and specific pDNA per OD (23mg/L and 1.49mg/L/OD respectively).

HB101 gWiz was able to maintain a high percentage of DNA supercoiling at 5L scale (98%), however, the percentage of supercoiled DNA obtained from BL21 DE3 gWiz fermentation was low (63%). Faint multimers and higher order oligomers are also observed in the plasmid DNA gel scans (Figures 5.3.2.5. and 5.3.2.6.). The presence of multimers and higher order oligomers may complicate purification and characterisation since more compact SC forms have a higher charge density and therefore elute later than the OC forms (Prazeres et al., 1998). Some multimers may change plasmid stability since the presence of plasmid multimers can confuse plasmid-encoded control circuits, leading to copy number depression (Summers, 1998). Plasmid multimerisation reduces the number of independently segregating units leading to segregational instability of high copy number plasmids (Summers and Sherratt, 1988). The use of multimeric plasmids in a process may also be a regulatory issue as different forms of plasmid molecules may have different potencies (Durland and Eastmann, 1998).

HB101 has been reported to produce large amounts of carbohydrate (Carnes et al., 2006; Prazeres et al., 1999); however, the carbohydrate content of HB101 was not measured in this study to confirm or disprove this. For future reference, taking an absorbance measurement of the lysate at 320nm can provide an indication of the polysaccharide content.

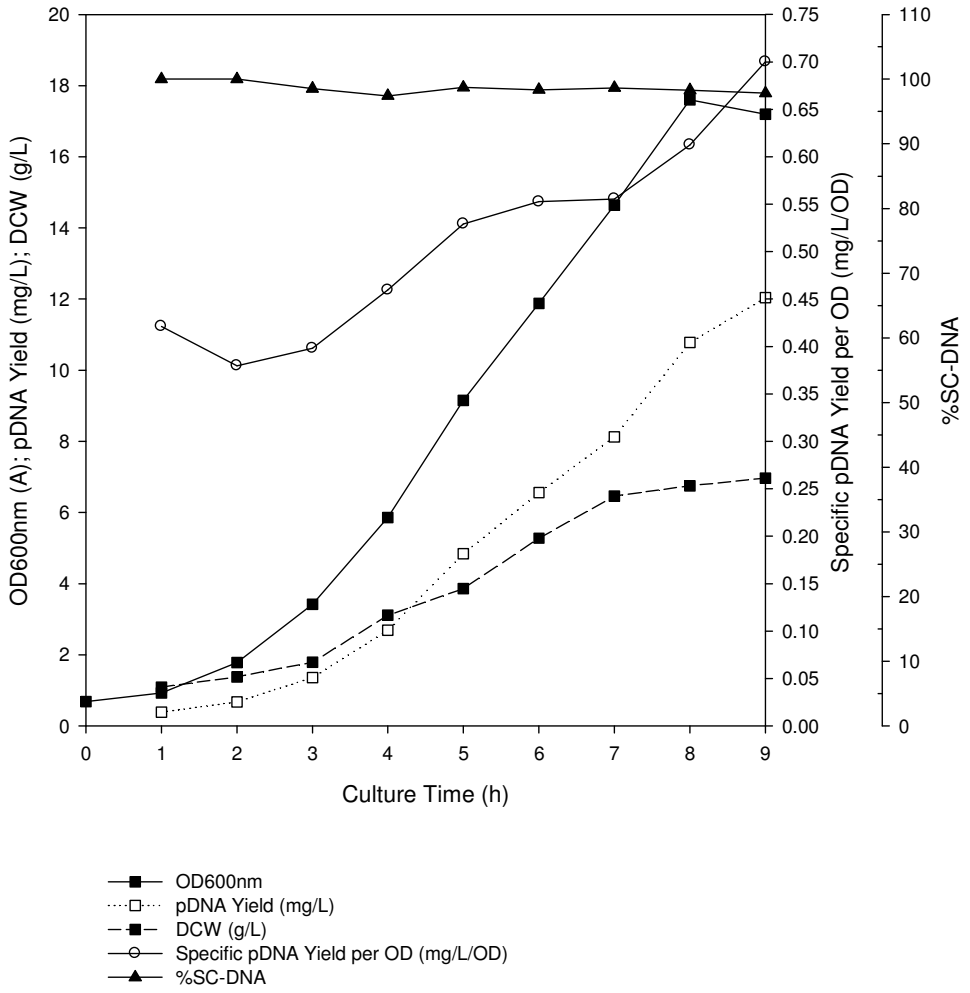


Figure 5.3.1.1. HB101 gWiz (7L SDCAS batch fermentation, 5L working volume grown at 37°C), biomass and plasmid profiles.

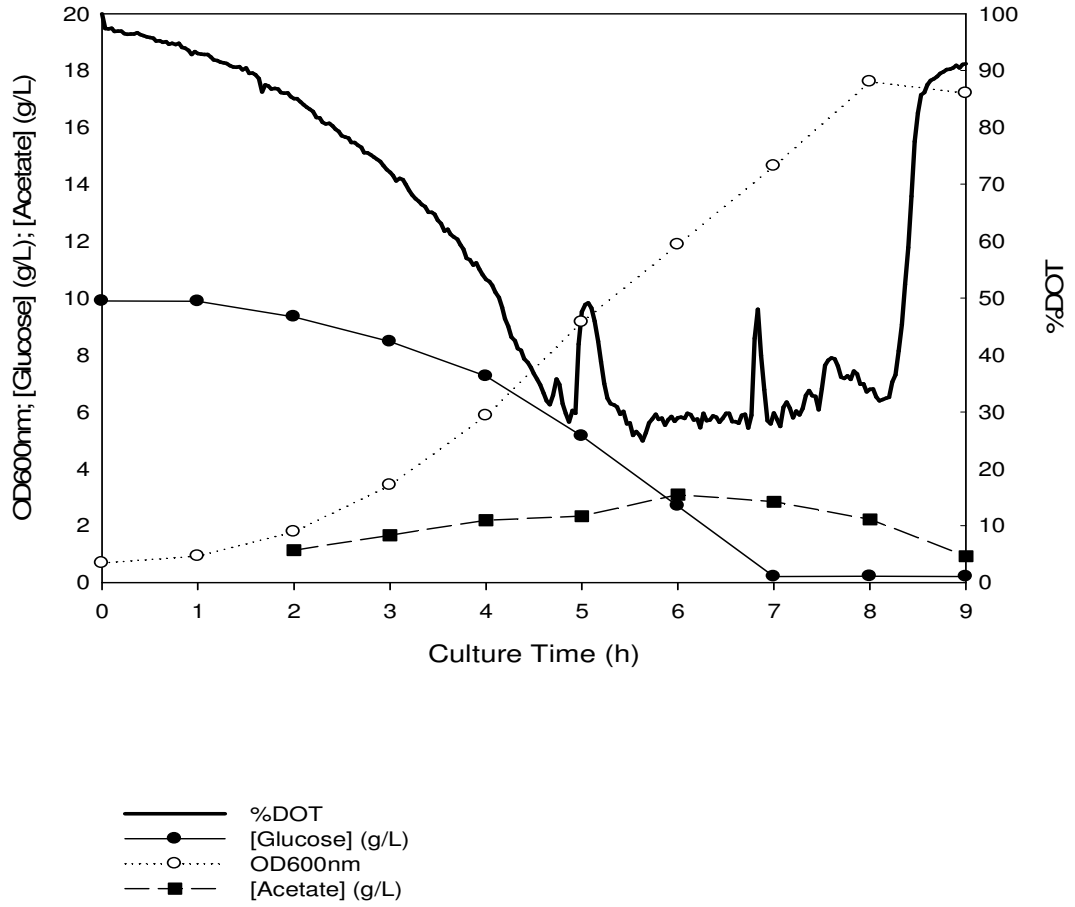


Figure 5.3.1.2. HB101 gWiz (7L SDCAS batch fermentation grown at 37℃, 5L working volume), dissolved oxygen tension, biomass, [glucose] and [acetate] profiles. Although DOT cascade control was implemented to maintain the DOT at 30% by automatically adjusting the agitation speed, fluctuations were observed after 4.5 hours of growth which were due to overcompensation of the agitation speed.

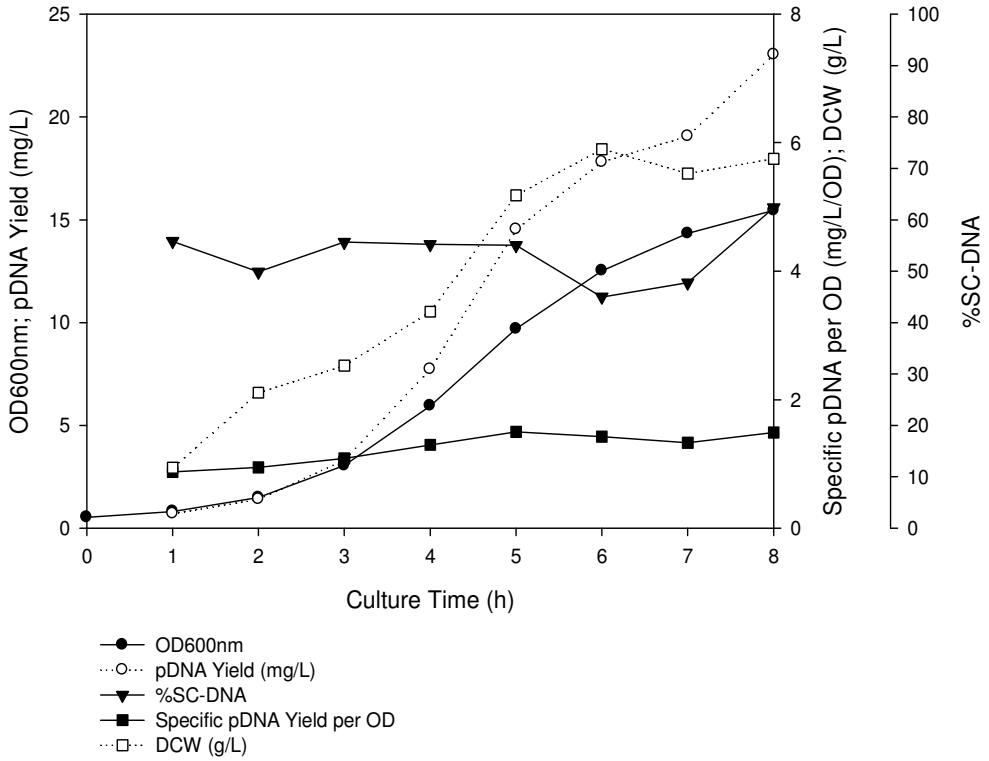


Figure 5.3.1.3. *E. coli* BL21 DE3 gWiz biomass and plasmid profiles from a 7L (5L working) batch SDCAS fermentation grown at 37℃.

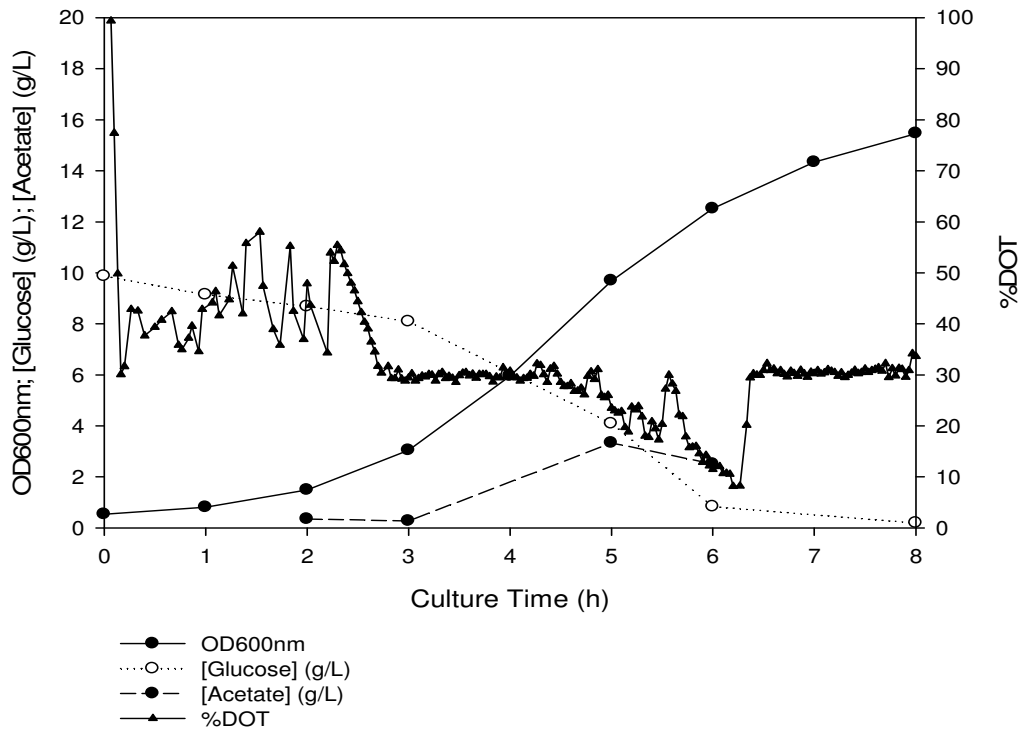


Figure 5.3.1.4. Biomass, [glucose] and [acetate] and dissolved oxygen tension profiles from an *E. coli* BL21 DE3 gWiz 7L SDCAS batch fermentation grown at 37℃. DOT cascade control was implemented to maintain the DOT at 30% by automatically adjusting the agitation speed, however, the %DOT briefly dropped to 8% after 6 hours when the agitation reached its maximum speed. The fluctuating %DOT readings observed between 0 and 2 hours of growth suggest a problem with the DOT probe or probe calibration.

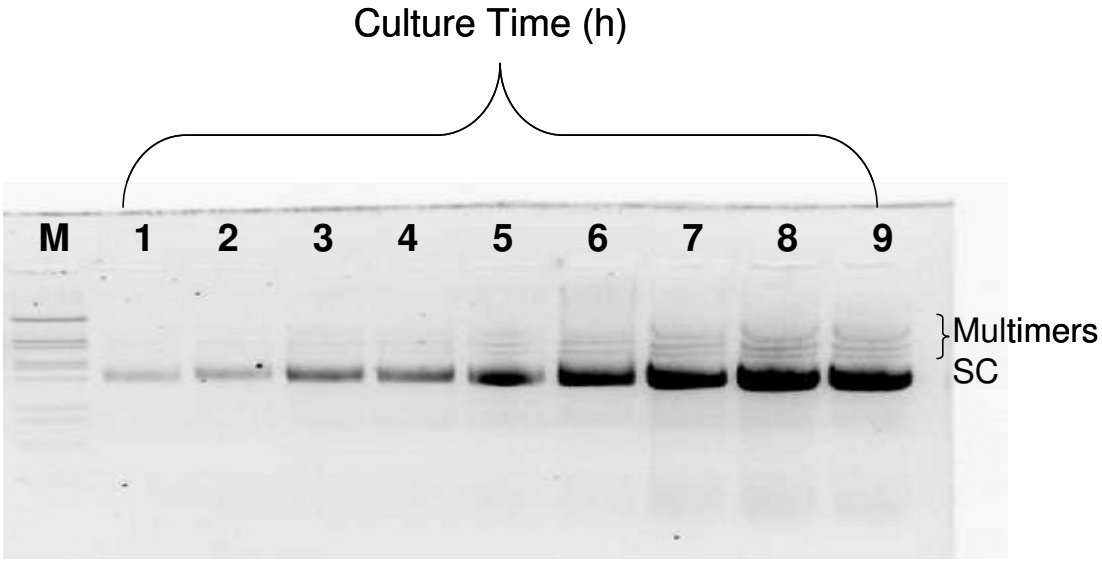


Figure 5.3.1.5. Gel scan of miniprepped HB101 gWiz samples from a 7L fermentation run (SDCAS medium, 37℃). M: λBstE II Marker. The same volume of cells was lysed at each interval during the miniprep process hence why the plasmid amount is steadily increasing with the increasing OD of the fermentation culture.

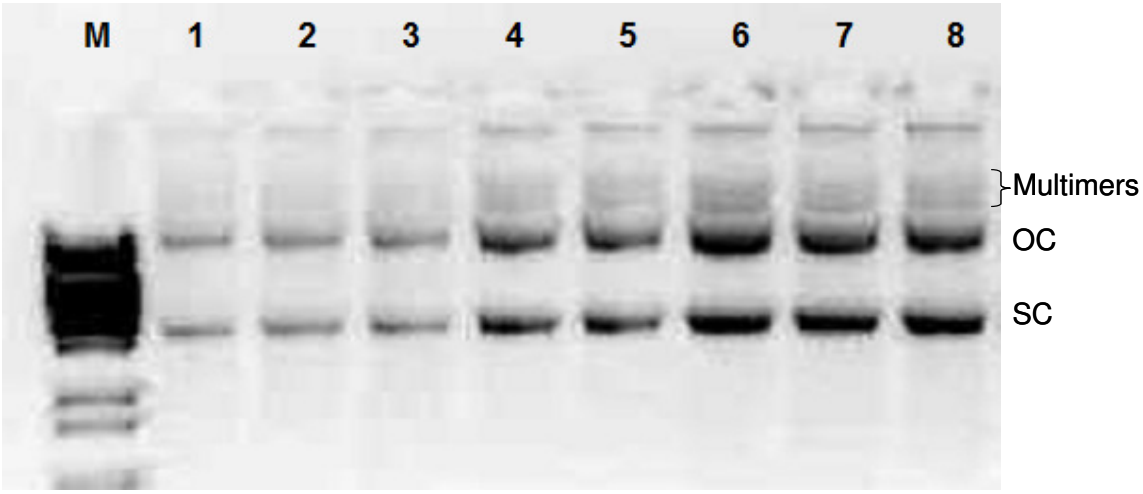


Figure 5.3.1.6. Gel scan of miniprepped BL21 DE3 gWiz samples from a 7L fermentation run (SDCAS medium, 37℃). M: λBstE II Marker.

Table 5.3.1.1. Comparison of growth and plasmid DNA yields of HB101 gWiz and BL21 DE3 gWiz at 5L working volume scale, grown on selective SDCAS medium at 37°C

Strain plasmid	Culture scale (working volume)	μmax (h^{-1})[a]	Mid exponential phase				Stationary phase			
			OD_{600nm}	Volumetric pDNA yield (mg/L)	Specific pDNA yield (mg/L/OD)	%SC-DNA ± 20%[c]	OD_{600nm}	Volumetric pDNA yield (mg/L)	Specific pDNA yield (mg/L/OD)	%SC-DNA ± 20%[c]
HB101 gWiz	5L	0.65 ± 0.03	9.2	4.8 ± 0.3[b]	0.5	99	17.2	12.0 ± 0.9[b]	0.7	98
BL21 DE3 gWiz	5L	0.69 ± 0.04	9.7	14.5 ± 1.5[b]	1.5	55	15.5	23.0 ± 1.1[b]	1.5	63

[a] Linest calculated standard deviation; [b] Standard error of the mean from a minimum of duplicate readings; [c] Error associated with densitometry method (Projan et al., 1983).

The accumulation of acetate in a culture can inhibit cell growth and reduce recombinant protein production. Acetate and other acidic by-products are most predominately formed in *E. coli* under aerobic conditions in the presence of excessive glucose what is known as the bacterial Crabtree effect (Rinas et al., 1989). HPLC acetate measurements of samples taken between 2-9 hours of growth indicates (Figure 5.3.1.2.) that the HB101 gWiz fermentation produced a maximum of 3.1g/L acetate, with the BL21 DE3 gWiz fermentation producing a similar amount, 3.3g/L acetate, although some BL21 DE3 gWiz samples were lost prior to the HPLC analysis therefore a hourly profile could not be obtained (Figure 5.3.1.4.). Once the glucose was consumed, both cultures reutilised acetate as observed by a decrease in the concentration of acetate obtained at the end of both fermentations. Acetate production and glucose consumption rates (Table 5.3.1.2.) show that the BL21 DE3 gWiz culture produces acetate at a higher rate than the HB101 gWiz culture (1.25 h^{-1} and 0.27 h^{-1} respectively).

Luli and Strohl, (1990), have reported that HB101 is generally a low acetate producer (final acetate concentration of 0.88 g/L; production rate of 0.57 h^{-1} from a batch 10L fermentation grown on SD-8 medium with 20g/L glucose at 37°C). The final acetate concentration obtained from this study was similar to Luli and Strohl's, (1990), at 0.93g/L. However in this study the maximum acetate concentration reached 3.1g/L which is quite high relative to the initial glucose concentration in the medium (10g/L). The maximum acetate concentration produced by HB101 was not mentioned in Luli and Strohl's, (1990) study and it is not known if glucose was used up at any point during the HB101 batch fermentation conducted by the authors.

Although this study has found that the BL21 DE3 culture has a high acetate production rate (1.25 h^{-1}) and produced a maximum of 3.3g/L of acetate, this *E. coli* B strain has been reported to produce a much lower concentration of acetate than some *E. coli* K-12 strains like JM109 (Shiloach et al, 1996; van de Walle and Shiloach, 1998). This has been

reported to be due to a mutation in one of the two enzymes catalyzing acetate production from acetyl CoA, with acetyl-P as the intermediate (Kleman and Strohl, 1994). Shiloach et al., (1996), have also reported that the amount of acetate produced by BL21 DE3 was low (2g/L) from an initial glucose concentration of 40g/L, although during exponential growth on high glucose concentration, this strain produces acetate at a specific rate equal to or higher than JM109 (which produced up to 14g/L acetate from 40g/L glucose). The authors also propose that this observation is also due to a control mechanism present in *E. coli* B derivatives (Kleman and Strohl, 1994; Shiloach et al., 1996). This proposition suggests that when the acetate concentration is above 1g/L, a control mechanism (suggested by Van de Walle and Shiloach, 1998, as the glyoxyate shunt pathway) activates at a certain point along the growth, changing the metabolic activity and reducing the glucose uptake rate. This in turn enhances acetate consumption and lowers acetate accumulation. Phue & Shiloach, (2004) have also suggested that, transcription of key metabolic genes (*aceA* and *aceB* and no transcription of the *iclR*) is the cause for the activation of glyoxylate pathway in BL21 DE3 that is inactive in JM109.

In general, acetate formation is dependent on the medium used and the specific growth rate, often occurring when the specific growth rate exceeds 0.2 or 0.35 h^{-1} in complex or defined media respectively (Doelle et al., 1974; Lee, 1996); however the critical specific growth rate that leads to acetate formation can vary among strains (Lee, 1996). The maximum specific growth rates of the HB101 gWiz and BL21 DE3 gWiz fermentations were around 0.7h^{-1} when grown on SDCAS medium which may explain why high concentrations of acetate was produced. A high acetate concentration reduces growth rate and biomass yield and the impact of this is more detrimental in a fed-batch system due to the extended culture period. Strategies can be implemented to reduce the detrimental effects of acetate such as controlled feeding of the carbon source to control the specific growth rate or by removing the acetate through electrodialysis (Kleman et al., 1991; Wong et al., 2010).

Table 5.3.1.2. Maximum glucose consumption and acetate production rates from 7L (5L working volume) SDCAS fermentations grown at 37°C

Strain plasmid	Final biomass yield on glucose (g DCW / g glucose consumed)	Glucose consumption		Acetate production	
		Consumption rate (h^{-1})[a]	Final concentration (g/L)	Production rate (h^{-1})[b]	Final concentration (g/L)
HB101 gWiz	0.71	1.60	0.21	0.27	0.93
BL21 DE3 gWiz	0.59	1.02	0.19	1.25	2.51[c]

[a, b] To allow for comparison with Luli and Strohl's, (1990) data, these rates were determined using the same slope calculation mentioned by Luli and Strohl, (1990)

[a] Calculated as Ln [(acetate concentration at T_2 – acetate concentration at T_1) / (T_2-T_1)], which yields h^{-1}

[b] Calculated as Ln [(glucose concentration at T_2 – glucose concentration at T_1) / (T_2-T_1)], which yields h^{-1}

[c] Last sample taken 2 hours prior to end of fermentation at 6 hours of growth

5.3.2. 7L BL21 DE3 gWiz Anaerobic Fermentations (SDCAS Medium (10g/L D-glucose and 10g/L casamino acids)

A preliminary 7L aerobic-anaerobic fermentation culture was conducted in the same way as two aerobic control runs (BL21 DE3 gWiz grown on selective SDCAS medium at 37℃, 5L working volume) except 100% nitrogen was sparged into the culture in place of the air supply when anaerobiosis was implemented. Similar to what was conducted by Cortassa & Aon, (1993) and Hsieh et al., (1991), the anaerobic transition occurred during the exponential phase of growth; in this case after 4 hours of growth during the mid exponential phase as determined by the aerobic fermentations conducted prior to the anaerobic fermentations and lasting for another 3 hours (remaining period of exponential phase) before the air supply was returned to the culture for the remainder of the fermentation.

Results from the aerobic (control) and aerobic-anaerobic fermentations are summarised in Table 5.3.2.1. During the anaerobic period, biomass production and growth was inhibited (Figure 5.3.2.1.), however the specific pDNA yields obtained during this period were similar to those obtained from the control cultures during the same time period (Figure 5.3.2.2). As shown in Figure 5.2.2.3., the amount of supercoiling from the anaerobic culture seemed to increase during the anaerobic period; however it is difficult to compare these values with the values obtained from the aerobic cultures due to the high level of variability, which may suggest issues with reproducibility. It is important to note that although the standard errors of each sample show the variation in the data, there is also a ±20% error associated with the densitometry determination method as reported by Projan et al., (1983).

Hsieh et al., (1991), reports that the initial onset of anaerobiosis resulted in a decrease in pDNA and chromosomal DNA supercoiling corresponding to a decrease in the [ATP]/[ADP] ratio; however after a period of 20 minutes, the level of supercoiling (measured by the change in linking number) elevated to exceed aerobic levels. Cortassa and Aon, (1993), reported that

in their investigation, DNA supercoiling (also measured by the change in linking number) increased after the onset of anaerobiosis but this increase was not maintained after the glucose level was depleted.

DNA supercoiling is controlled by the level of DNA gyrase and Topoisomerase I enzymes present and as gyrase is ATP-dependant, the intracellular [ATP]:[ADP] ratio affects how DNA gyrase functions. However since this ratio was kept constant in Cortassa and Aon's (1993) investigation, the observed increase in supercoiling was presumed by the authors to have been caused by decreasing Topoisomerase I levels. The authors then linked the amount of available glucose as the factor affecting the level of supercoiling. When cells are depleted their carbon and energy source, the resulting drop in energy state is accompanied by an increase in the linking number, corresponding to a decrease in the level of supercoiling (Jensen et al., 1995).

Although the level of % plasmid supercoiling appeared to have increased in the plasmid samples from the anaerobic culture during the 3 hour anaerobic period, this increase was not sustained and it is possible that glucose depletion after 5 hours of growth maybe the cause. To investigate if glucose depletion has affected the supercoiling levels in the anaerobic cultures, further fermentations were carried out that used SDCAS medium with 20g/L D-glucose and 20g/L casamino acids to ensure that glucose was available during the anaerobic period.

Table 5.3.2.1. Biomass and plasmid measurements from aerobic (control) and anaerobic (between 4-7 hours) 5L working volume SDCAS (10g/L D-glucose) fermentations of BL21 DE3 gWiz grown at 37°C

	Control 1 & 2	Anaerobic* culture A	Control 1 & 2	Anaerobic* culture A	Control 1 & 2	Anaerobic* culture A	Control 1 & 2		Anaerobic* culture A
Time (h)	**Mean OD_{600}** (**μmax:** 0.69 ± 0.04[a] h^{-1}) [a]	**Mean OD_{600}** (**μmax:** 0.69 ± 0.05 [a] h^{-1}) [a]	**Mean vol. pDNA yield (mg/L)**[b]	**Mean vol. pDNA yield (mg/L)**[d]	**Mean specific pDNA yield (mg/L/OD)**[b]	**Mean specific pDNA yield (mg/L/OD)**	**%SC-DNA 1**[e]	**%SC-DNA 2**[e]	**%SC-DNA C**[e]
0.0	0.6 ± 0.1	0.7	NS	NS	NS	NS	NS	NS	NS
1.0	1.0 ± 0.2	1.3	0.7[c]	1.0	0.9[c]	0.8	56 ± 5	NA	58
2.0	1.8 ± 0.3	2.3	1.4[c]	1.4	1.0[c]	0.6	50 ± 1	NA	56
3.0	3.6 ± 0.6	5.0	3.3[c]	1.4	1.0[c]	0.3	56 ± 3	NA	55
4.0	6.5 ± 0.6	9.2	10.5 ± 2.7	2.9 ± 0.3	1.6 ± 0.3	0.3	55 ± 4	55 ± 2	54 ± 6
5.0	10.2 ± 0.5	9.6	13.5 ± 1.0	11.6 ± 0.6	1.3 ± 0.2	1.2	55 ± 4	54 ± 8	57 ±10
6.0	12.7 ± 0.2	9.6	18.0 ± 0.2	13.3 ± 0.7	1.4 ± 0.0	1.4	50 ± 3	67 ± 4	61 ± 7
7.0	14.4 ± 0.1	9.7	16.1 ± 3.0	13.0 ± 1.0	1.1 ± 0.2	1.3	48 ± 3	49 ± 6	61 ± 7
8.0	14.8 ± 0.7	11.6	18.6 ± 4.4	14.1 ± 1.4	1.3 ± 0.2	1.2	62 ± 2	65 ± 1	60 ± 2
9.0	14.9 ± 1.3	12.7	13.7 ± 6.9	16.0 ± 1.3	0.9 ± 0.4	1.3	NA	72 ± 4	67 ± 7

NS= not sampled; NA= not available (sample lost). [a] Linest calculated standard deviation; [b] standard error of the mean from the two fermentation cultures; [c] from Control 1 only, as samples from Control 2 were lost; [d] standard error of the mean from at least duplicate samples; [e] standard error of the mean from duplicate samples, the densitometry method also has an associated error of ± 20%.*Samples at 4 hours of growth were taken just prior to anaerobic transition. Where no error is stated, only one sample was measured at that time point.

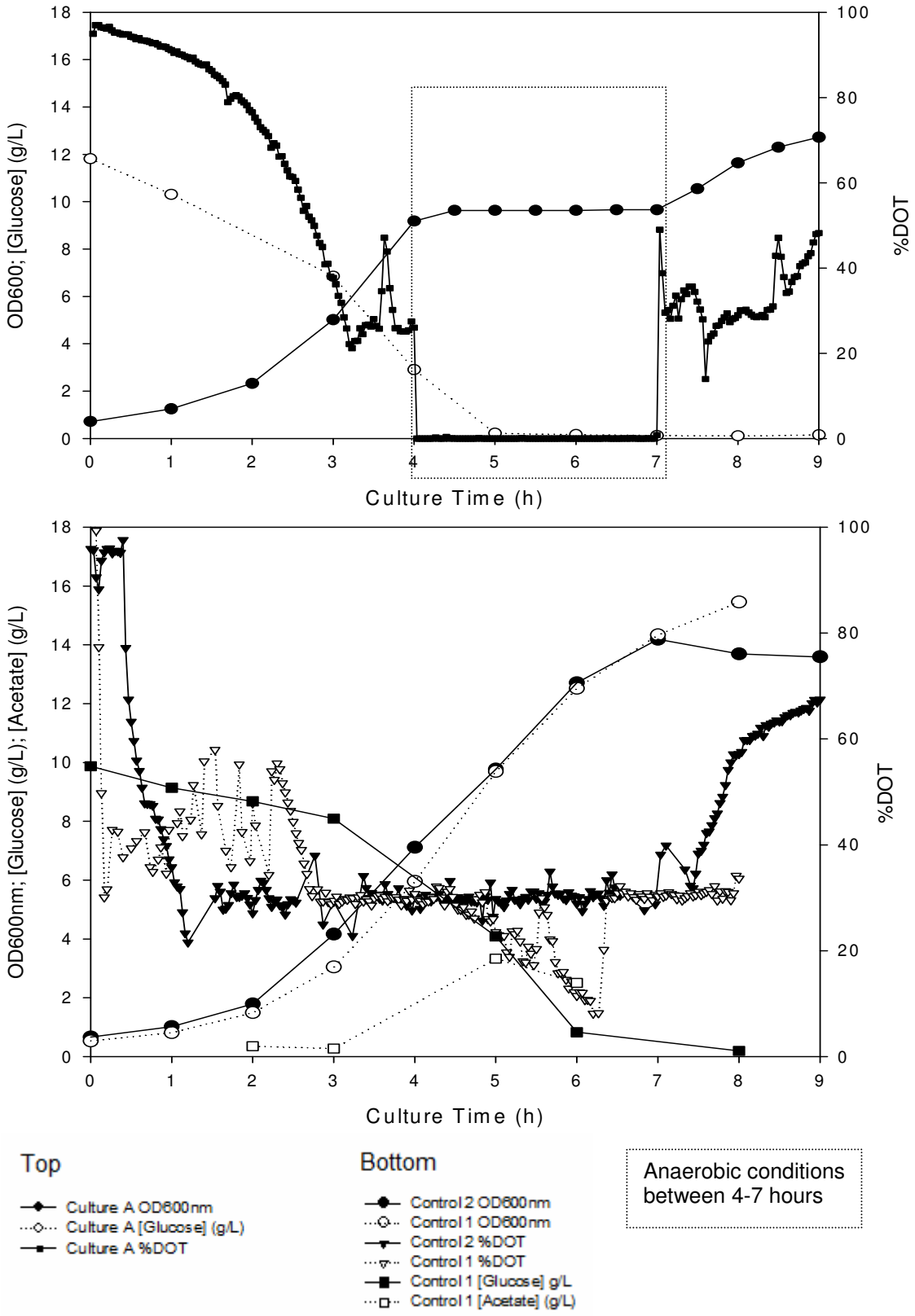


Figure 5.3.2.1. (Top): OD and D-glucose concentration profiles of a 7L batch fermentation culture (A) of BL21 DE3 gWiz grown at 37℃ that was switched to anaerobic conditions between 4 and 7 hours of growth. (Bottom): OD profiles from two control 7L fermentations (completely aerobic throughout) of BL21 DE3 gWiz at 37℃. The D-glucose concentration profile is only available for one of the control fermentations.

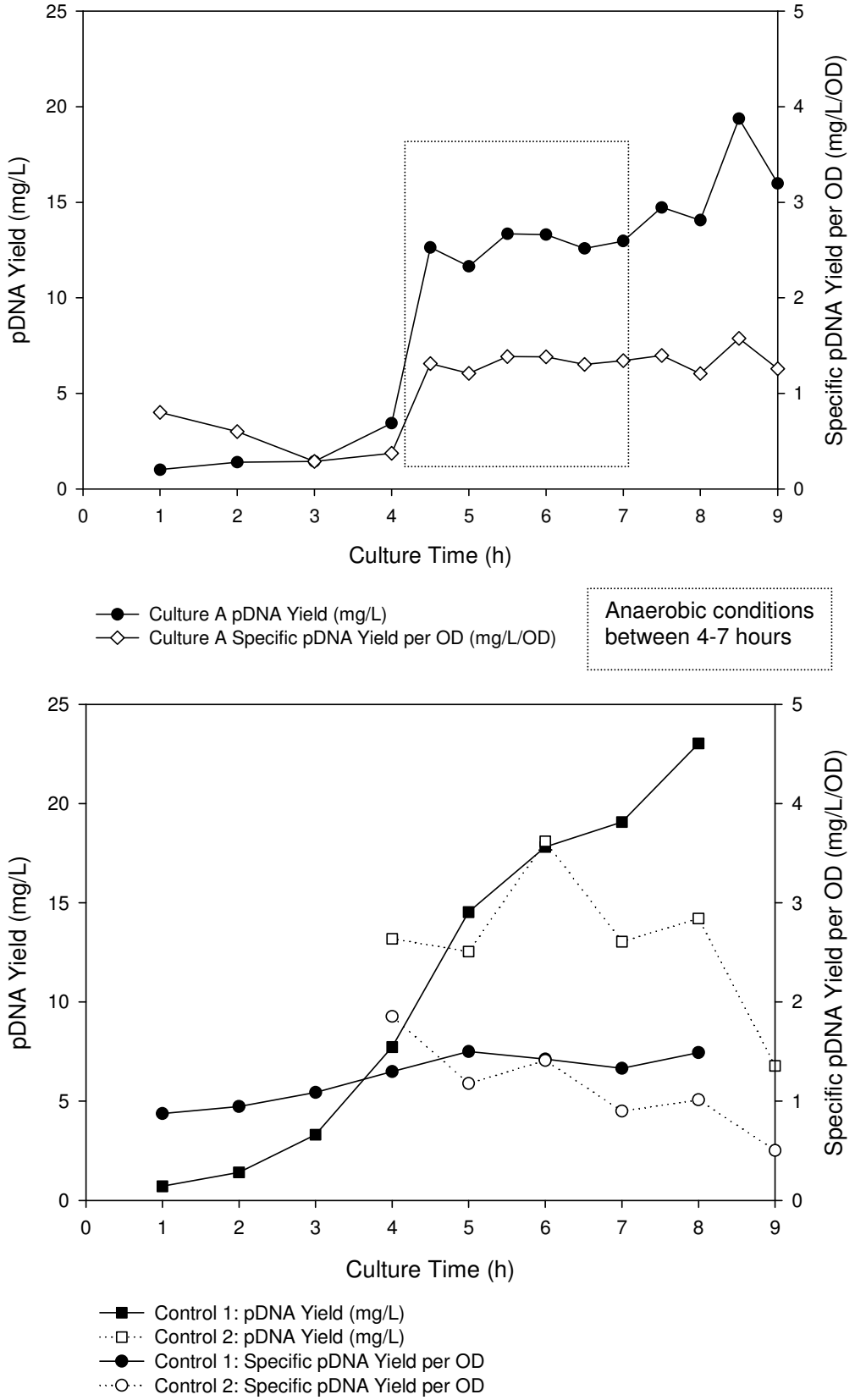


Figure 5.3.2.2. (Top): Plasmid profiles from aerobic-anaerobic (anaerobic between 4 and 7 hours growth) and completely aerobic 7L batch fermentation cultures (control) (bottom) of *E. coli* BL21 DE3 gWiz grown at 37℃. Samples taken between 1-3 hours from control culture 2 were lost.

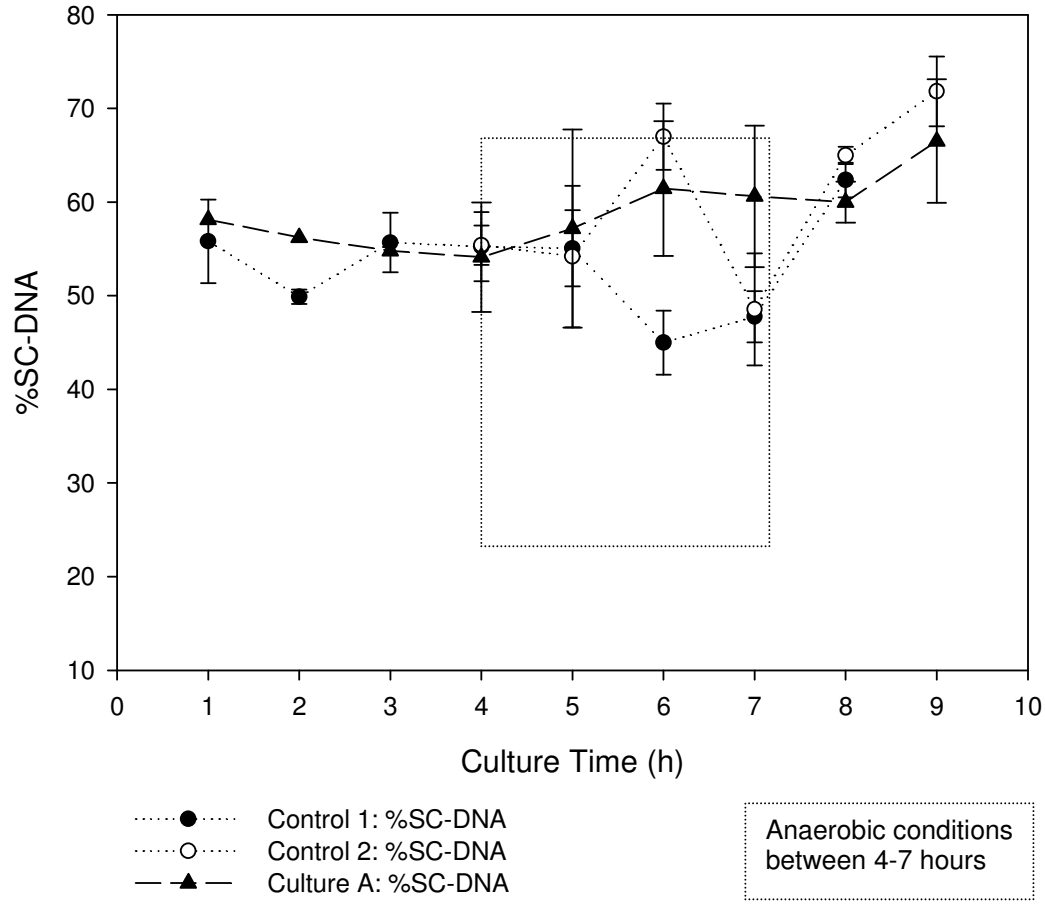


Figure 5.3.2.3. %SC-DNA profiles of aerobic-anaerobic (anaerobic between 4 and 7 hours growth) culture (A) and completely aerobic 7L batch SDCAS fermentation cultures (control) of *E. coli* BL21 DE3 gWiz grown at 37℃. Error bars reflect the standard error of the mean from a minimum of duplicate samples (except in Culture A between 1-3 hours when only one sample was measured at each time point). Samples taken between 1-3 hours from control culture 2 were lost. Samples at 4 hours of growth were taken just prior to anaerobic transition.

5.3.3. 7L BL21 DE3 gWiz Anaerobic Fermentations (with 20g/L D-glucose) fermentations

Fermentation runs using 20g/L D-glucose and 20g/L casamino acids SDCAS medium were conducted under complete aerobic and anaerobic (between 4-7 hours) conditions. Glucose levels did not become depleted until after the anaerobic period ended when the air supply was returned to the fermenter (Figure 5.3.3.1.).

The maximum concentration of acetate obtained from the control and anaerobic cultures (out of the samples measured) were 3.1g/L and 4.7g/L respectively (Table 5.3.3.1.). This slight increase in the maximum acetate produced from the anaerobic culture is a good indicator of anaerobic growth on glucose since facultative anaerobic bacteria such as *E. coli* carries out mixed-acid fermentation in the absence of oxygen, producing substrates that facilitate acetate production (Smith and Neidhardt, 1983). A comparison the maximum acetate production rates (albeit calculated based only on the acetate samples available) show similar rates between the aerobic and anaerobic cultures (0.27 h^{-1} and 0.21 h^{-1} respectively); however it is important to note that the maximum acetate production rate of the anaerobic culture occurred during the anaerobic period, when the maximum rate from the control aerobic culture was only 0.15 h^{-1} during the same period.

Although slow growth was observed during anaerobic period (the maximum growth rate was 0.14 h^{-1} compared to a μmax of 0.55h^{-1} from the control culture during the same time period), this slower growth rate did not result in higher pDNA yields (volumetric and specific) obtained from the anaerobic culture when compared to the control culture (Figure 5.3.3.2.). Generally plasmid copy numbers have been shown to increase as a result of reduced growth rates (Fitzwater et al., 1988); however since anaerobiosis is less efficient than aerobiosis, less metabolic energy is available for biosynthesis (Neidhardt et al., 1990) and may explain why a lower volumetric pDNA yield was obtained during anaerobic growth.

Although the %DOT of the control fermentation dropped below 20% after 5.5 hours of growth (down to 2% after 6.5 hours) when the stirrer reached its maximum speed, a comparison of the % SC-DNA levels obtained from both the control and anaerobic cultures between 4 and 6 hours of growth showed that the level of % pDNA supercoiling did not increase during anaerobiosis (Figure 5.3.3.3.).

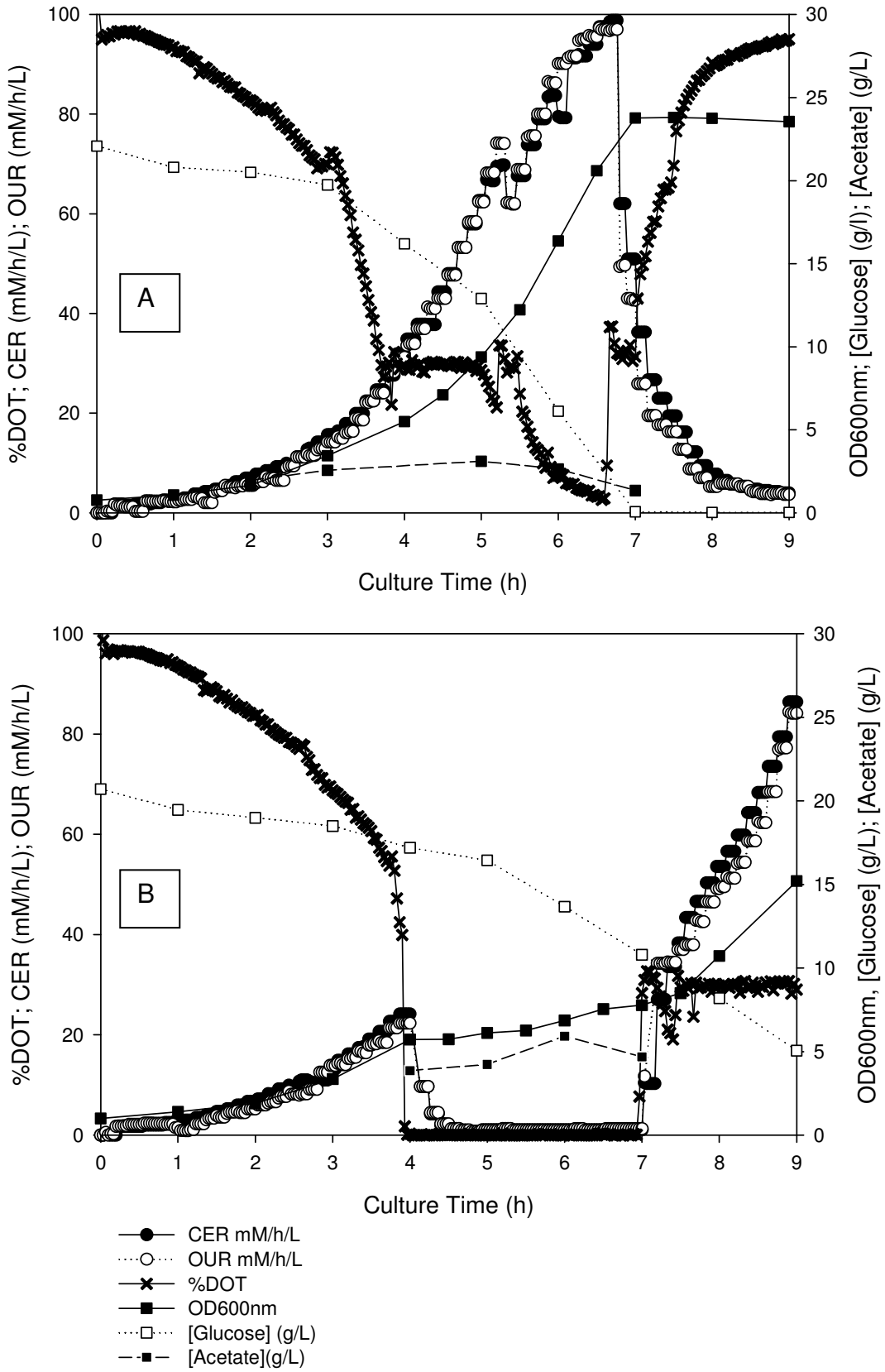


Figure 5.3.3.1. A & B: %DOT, OUR, CER, optical density, glucose and acetate profiles from (A) aerobic (control) and (B) anaerobic (4-7hours) 7L batch fermentation cultures of BL21 DE3 gWiz grown on SDCAS medium consisting of 20g/L D-glucose and 20g/L casamino acids grown at 37℃.

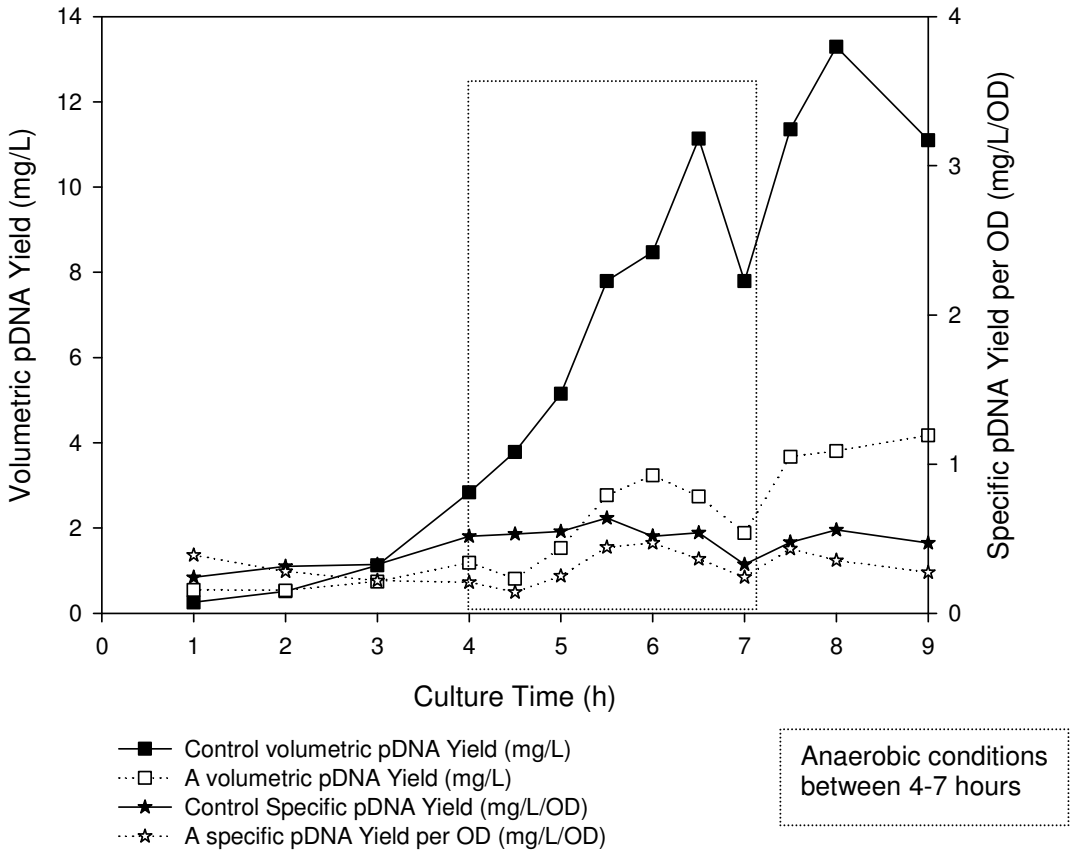


Figure 5.3.3.2. Plasmid profiles from anaerobic (A) and aerobic (control) 7L batch fermentation cultures of BL21 DE3 gWiz grown on SDCAS medium consisting of 20g/L D-glucose and 20g/L casamino acids at 37℃. Samples at 4 hours of growth were taken immediately prior to anaerobic transition.

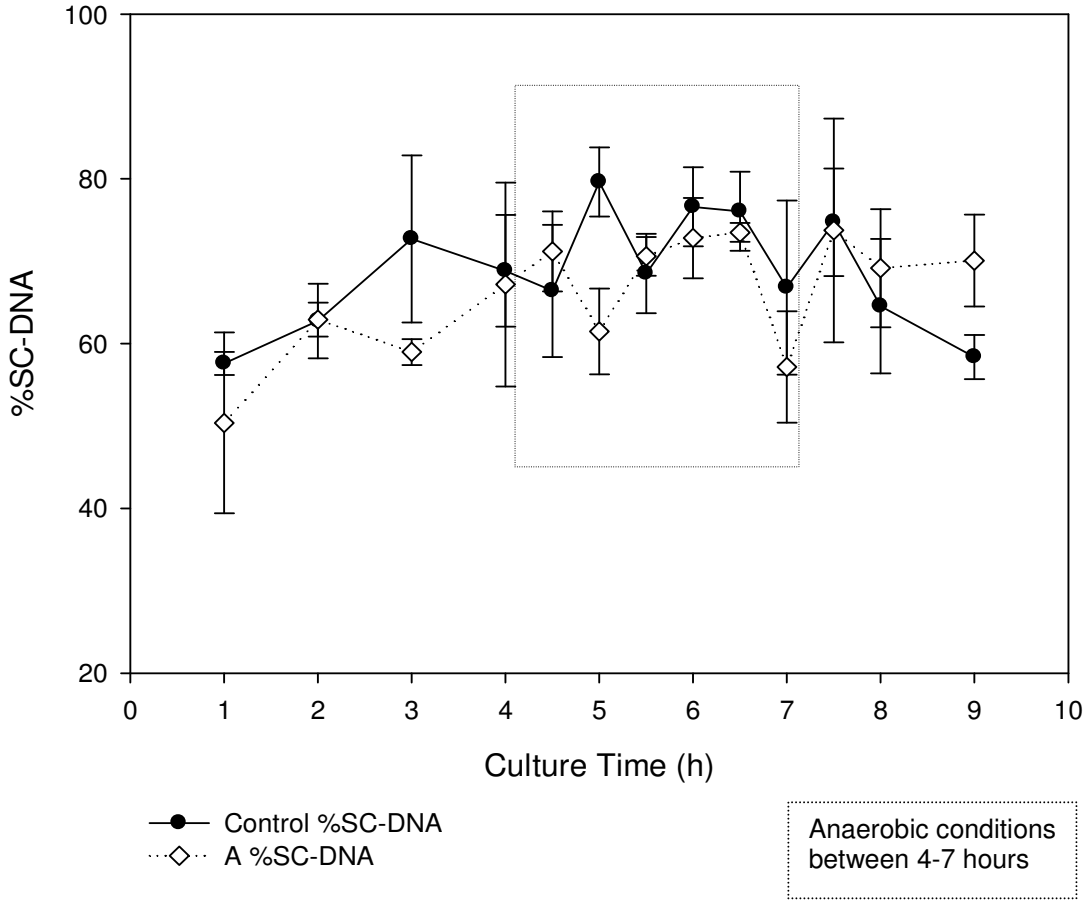


Figure 5.3.3.3. %SC-DNA profiles from anaerobic (A) (anaerobic period highlighted in box) and aerobic (control) 7L batch fermentation runs of BL21 DE3 gWiz grown on SDCAS medium consisting of 20g/L D-glucose and 20g/L casamino acids at 37°C. Error bars reflect the standard error of the mean from a minimum of duplicate samples. Samples at 4 hours of growth were taken immediately prior to anaerobic transition.

Table 5.3.3.1. Biomass and plasmid measurements from aerobic (control) and anaerobic (between 4-7 hours) 5L working volume SDCAS (20g/L D-glucose) fermentations of BL21 DE3 gWiz grown at 37°C

	Control	Anaerobic*	Control		Anaerobic*		Control	Anaerobic*	Control	Anaerobic*	Control	Anaerobic*
T (h)	OD_{600} (µmax: 0.60 h^{-1}; between 4-7 hours: 0.55 h^{-1}± 0.03 [a])	OD_{600} (µmax: 0.55h^{-1}; between 4-7 hours: 0.14 h^{-1}± 0.02 [a])	Maximum glucose utilisation (G) \ Acetate production (A) rate (h^{-1})				Volumetric pDNA yield (mg/L)[b]	Volumetric pDNA yield (mg/L)[b]	Specific pDNA yield (mg/L/OD)	Specific pDNA yield (mg/L/OD)	%SC-DNA [c]	%SC-DNA [c]
			G	A	G	A						
0.0	0.8	1.0	4.62	0.27	0.48	0.21	NS	NS	NS	NS	NS	NS
1.0	1.0	1.4					0.3 ± 0.1	0.6 ± 0.1	0.24	0.39	58 ± 1 [d]	50 ± 11
2.0	1.7	1.9					0.5 ± 0.1	0.5 ± 0.2	0.32	0.28	63 ± 5	63 ± 2 [d]
3.0	3.4	3.4					1.1 ± 0.3	0.8 ± 0.1	0.33	0.22	73 ± 10	59 ± 2 [d]
4.0	5.5	5.7					2.8 ± 0.2	1.2 ± 0.1	0.52	0.21	69 ± 7	67 ± 12 [d]
5.0	9.4	6.1					5.2 ± 1.4	1.5 ± 0.4	0.55	0.25	80 ± 4	61 ± 5
6.0	16.4	6.9					8.5 ± 1.6	3.2 ± 0.4	0.52	0.47	77 ± 5	73 ± 5
7.0	23.8	7.8					7.8 ± 1.1	1.9 ± 0.1	0.33	0.24	69 ± 11	57 ± 7
8.0	23.8	10.7					13.3 ± 1.2	3.8 ± 0.2	0.56	0.36	65 ± 8 [d]	69 ± 7
9.0	23.6	15.2					11.1 ± 0.7	4.2 ± 0.4	0.47	0.27	58 ± 3 [d]	70 ± 6

NS= Not sampled; [a] Linest calculated standard deviation; [b] standard error of the mean from triplicate samples; [c] standard error of the mean from triplicate samples ([d] from duplicate samples), please also note that the densitometry method has an associated error of ± 20%.*Samples at 4 hours of growth were taken just prior to anaerobic transition. No error range is available for the OD as only one measurement was taken at each time point.

5.3.4. Anaerobic SDCAS 1L Shake Flask Cultures

To test if the results were affected by the scale of the fermentation, independent triplicate 1L shake flask SDCAS cultures (200mL working volume) of BL21 DE3 gWiz were set up and 100% nitrogen from a nitrogen cylinder was pumped into the sealed flasks via sterile tubing and an air filter for the anaerobic transition during the mid-exponential phase (between 3 to 6 hours of growth). These cultures were incubated in a shaking water bath at 37°C, 100 strokes per minute, pH 7.0.

Similar to the fermentations cultures, results from these shake flask cultures did not show an increase in the percentage of pDNA supercoiling during anaerobiosis (Figure 5.3.4.1.), suggesting that anaerobiosis does not increase % pDNA supercoiling for our system under the different conditions investigated.

However, it is interesting to note that unlike the fermentation cultures, volumetric and specific plasmid DNA yields obtained from the anaerobic shake flask cultures were higher during the anaerobic period than from the control (aerobic) culture during the same time period. As mentioned previously, this increase in plasmid production is likely a result of the slower growth rate of the anaerobic culture ($0.11h^{-1}$) compared to the aerobic culture ($0.22h^{-1}$) during the anaerobic period (3-6 hours of growth); however the reasons why plasmid production increased under anaerobic shake flask conditions and not under fermentation conditions are unclear. Less efficient control of the growth conditions (such as dissolved nitrogen or oxygen, cooling or heat transfer and pH) of the shake flask cultures compared to the fermentation cultures may have affected plasmid production to some extent. Acidic conditions and higher temperatures (40-42°C) have been reported to increase plasmid production of relaxed plasmids (Reinikainen et al., 1989; Wong et al., 1982); and since these cannot be consistently controlled as effectively under shake flask conditions, any variation from the initial conditions together with a slower growth rate may have resulted in increased plasmid yields obtained during anaerobiosis.

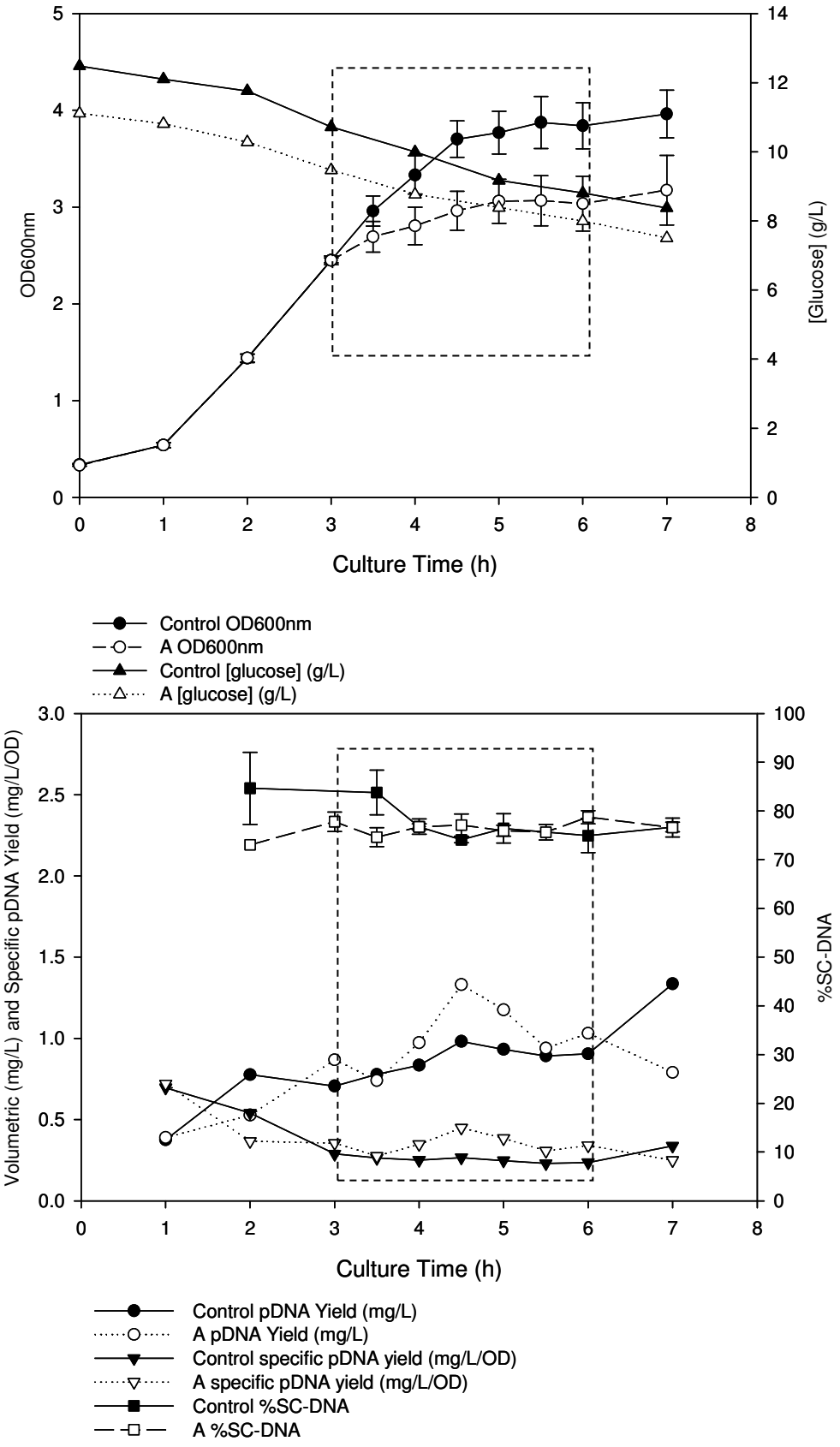


Figure 5.3.4.1. (top) OD and glucose concentration profiles (bottom) plasmid profiles averaged from independent triplicate 1L (200mL working volume) SDCAS shake flask cultures of BL21 DE3 gWiz grown at 37℃ aerobically (control) and anaerobically (A) during the period indicated by the dashed box.

5.4. Conclusion

Results from 7L (5L working volume) fermentation runs of HB101 gWiz and BL21 DE3 gWiz show that HB101 gWiz was able to maintain a high percentage of DNA supercoiling (98%) throughout the course of the fermentation and provide 12mg/L volumetric and 0.70mg/L/OD specific pDNA yields; whereas BL21 DE3 gWiz provided high volumetric and specific yields (23mg/L and 1.49mg/L/OD) albeit the percentage of supercoiled DNA obtained from BL21 DE3 gWiz fermentation was low (63%).

The main finding from this chapter is that under the conditions investigated, the results obtained show that anaerobic shock or anaerobiosis does not increase the percentage of supercoiled plasmid DNA. It is important to note that in previous studies where anaerobiosis has been reported to increase plasmid supercoiling (Cortassa and Aon, 1993; Hsieh et al., 1991), the level of supercoiling measured focused specifically on the linking number distribution and not the percentage of plasmid supercoiling as investigated in this study. The %SC-DNA value indicates the percentage of plasmid DNA in the total population that is supercoiled, whereas the supercoil linking number (number of superhelical turns) indicates how tightly supercoiled the plasmid is out of the supercoiled plasmid DNA in the population. Since the level of DNA superhelicity varies with changes in cellular energy, the linking number can change more rapidly in response to altered cellular and environmental conditions than the level of % plasmid supercoiling. This study attempted to investigate whether the altered environmental condition of anaerobiosis effecting changes in topoisomerase I and DNA gyrase activity would shift the linking number distribution, enough to convert relaxed plasmids to the supercoiled form, increasing the proportion of supercoiled DNA in the population. This remains undetermined.

Measuring the linking number distribution of the plasmid samples during aerobic and anaerobic conditions would have been invaluable to this study to confirm whether anaerobiosis can increase pDNA supercoiling as

determined by the superhelical density (linking number); however in this study multiple attempts at running chloroquine gels to determine the linking number of the plasmid samples were unsuccessful. Agarose gels containing chloroquine (but no ethidium bromide) were ran at a low voltage for up to 23 hours before they were washed for an hour and then stained with ethidium bromide. Many factors were investigated such as various chloroquine concentrations and electrophoresis conditions. Treating the plasmid samples with Topoisomerase I from wheat germ as reported by Xu and Bremer, (1997), also failed to resolve the plasmid. Since different plasmids have different sensitivity to these factors, many published reports mention the use of different concentrations of chloroquine, running and washing buffers and running times to resolve a variation of plasmid types and sizes; thus the optimum chloroquine gel conditions for gWiz obtained from BL21 DE3 remains to be established before linking number determination can be carried out.

6. A ‘Quiescent Cell System’ for Plasmid DNA Production

6. A 'Quiescent Cell System'

6.1. Introduction

A fundamental problem associated with the expression of recombinant DNA in fast growing *E. coli* cells is that most of the energy and nutrient resources are directed toward biomass production. The metabolic stress imposed by hosting a multicopy plasmid reduces the growth rate and viability of the host cell. A promising alternative to the traditional high-biomass fermentations for plasmid DNA production is to employ a 'Quiescent cell system', a state where chromosomal replication and expression is temporarily shut down but residual proteins remain metabolically active, using indole induction of *hns* mutated *E. coli* cells (Chant et al., 2007). During quiescence the cells cease dividing and go into a stationary but metabolically active phase, devoting a large proportion of their resources into plasmid-borne gene expression rather than for growth and other functions. Since the cells are no longer growing to produce biomass, this redirection of energy and resources may provide an improved yield of recombinant product without the need for additional nutrients in the media. This advantage would ease downstream processing and purification stages (Rowe & Summers, 1999).

The *E. coli* histone-like nucleoid structuring (H-NS) protein is a nucleoid-associated protein involved in gene regulation and DNA compaction. It has been reported that when the *E. coli* H-NS protein is overexpressed, growth inhibition, filamentation and nucleoid compaction is observed (Spurio et al., 1997). As mentioned previously in Chapter 1 (Section 1.7.2.1.), Rowe and Summers, (1999), have reported that a mutation in *hns* is necessary to generate an Rcd-induced quiescent state in broth culture. *hns* mutants have also been reported to be better suited to high cell density fermentations since they cope better with higher osmolarities and lower pHs than wild type cells (Mukherjee et al., 2004).

Previous studies on Rcd quiescent cell induction, involving the overexpression of a 70nt RNA called Rcd have provided stable

quiescence for at least 24 hours (Mukherjee et al., 2004, Rowe & Summers, 1999). During quiescence, the cells were able to synthesize the plasmid-encoded chloramphenicol acetyltransferase (CAT) protein for at least 10 hours after Rcd-induction, although there was a decline in the rate of synthesis over time (Rowe & Summers, 1999). This decline was not observed for single-chain antibody (scFv) expression during quiescence (Mukherjee et al., 2004), in fed-batch 5L fermentation studies, where protein expression was sustained. Both these studies utilised a Rcd-induction method that involves a two-step heating preparation from 30℃ to 42℃ to induce Rcd expression on a two-plasmid system, consisting of plasmid B8 with *rcd* under control of the λP_R promoter and p*cIts*857, encoding the *cIts*857 temperature-sensitive repressor. The shift in the temperature induces Rcd expression from the λP_R promoter since the *cIts*857 protein does not function as a transcriptional repressor at 42℃. However, a recent study has shown that an aromatic molecule indole is required for Rcd-mediated stabilisation of multicopy plasmids by acting as an intracellular signalling molecule (Chant et al., 2007). Chant et al., 2007 proposed that the upregulation of indole synthesis by Rcd is responsible for delaying cell division and it was deduced that indole can also be used to induce quiescence in place of the two-step Rcd induction.

Since indole induction involves only one step, the addition of indole to a culture containing *hns* mutated *E. coli* cells would be much easier to implement than Rcd induction. Studies carried out by Summers' team on indole-induction have mainly been conducted on cultures induced at an OD_{600nm} of 0.2 although a few cultures have been successfully induced at higher ODs of 15, eventually reaching ODs of around 50 after induction. This increase in OD after induction is perhaps better explained by the elongation of existing cells as previously observed rather than biomass production, since the cells were observed to be in a quiescent state. According to Rowe and Summers, (1999), the majority of cells remain undivided and of uniform size, but elongated to two to four cell lengths with correctly portioned and condensed nucleoids.

The addition of 2mM-3mM indole to culture medium has been found to result in optimum indole-induced quiescence for *hns* mutated W3110 cells (personal communication with Summers' team; Summers and Chant, 2009). Chant et al., (2007), have reported that an exogenous indole concentration ranging between 3 to 6mM results in effectively suppressing the growth of *E. coli* BW25113 containing the Rcd expression plasmid pCm(ss)-19. These studies suggest that the concentration of indole is important in maintaining a quiescent state and this concentration may differ for different strains or conditions.

Chant et al., (2007), also reported that the cell growth phase affects the level of intracellular indole as Rcd levels in late exponential/early stationary phase cells are higher. Increased Rcd levels were reported to increase the binding affinity of the enzyme tryptophanase to tryptophan which in turn stimulates indole production. Since it is also presumed that an excessive indole concentration can cause cell lysis, further research is needed to find the optimum indole concentration, and induction time for plasmid DNA production.

Summers and Chant, (2009), have briefly reported that replication of some plasmids (CAT and scFv) continue during quiescence, so should this study confirm that a 'Quiescent cell system' works for plasmid production, the industrial application of such a system is significant. It may result in the development of a system where cells are grown to a high cell density, induced to stop growing, but continue to produce plasmid DNA. Since the cells are no longer dividing and the host chromosome synthesis is down-regulated; more energy and cellular resources should in theory be available for plasmid production.

6.2. Aim and approach

The aim of this chapter is to firstly attempt to achieve a 'Quiescent cell system' that continues to produce plasmid DNA (proof of concept) and in doing so, find the optimum conditions for inducing and maintaining quiescence for the strain and plasmids used. This chapter will focus on indole induction using *E. coli* strain W3110 *hns*-205::Tn10, which was transformed singly with the plasmids gWiz (5.8kb) (Aldevron) for plasmid DNA analysis and with pGlo (5.4kb) (Biorad) for GFP expression studies.

6.3. Results

6.3.1. Inducing Quiescence Using Indole for Plasmid Production

This initial part of the investigation was performed to test whether quiescence can be achieved and maintained during plasmid (gWiz) production. Several experiments on 100mL shake flask cultures (inoculated with the same overnight culture) all showed that the optimum concentration of indole for inducing quiescence is between 3-4mM (Figure 6.3.1.1.). 3mM of indole maintained quiescence more effectively than 4mM as observed by the slight decreasing OD in the 4mM culture. This decrease in OD may be explained by the fact that the uptake of excessive indole can cause cell lysis, especially when the cells are producing indole *in vivo* as well.

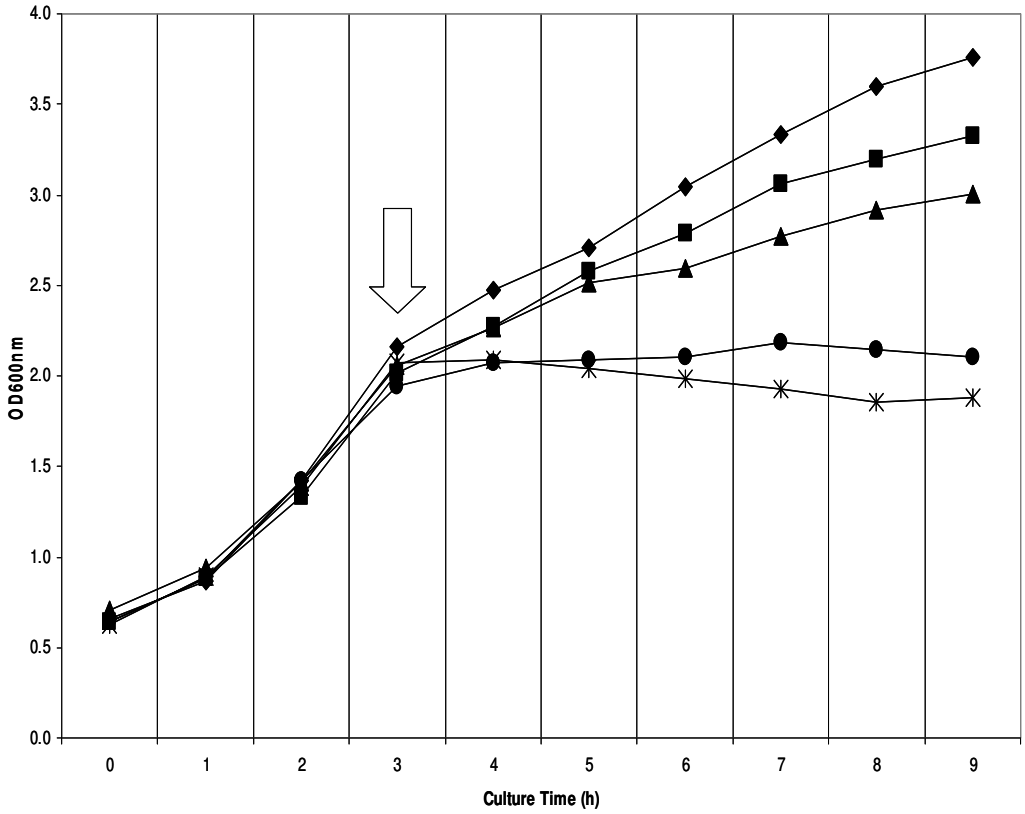


Figure 6.3.1.1. Average OD_{600nm} readings from W3110 Hns 205:: Tn10 gWiz 100mL shake flask cultures grown at 37℃, 250rpm (averaged from duplicate experiments). For each set of experiments, cultures were inoculated with 10mL of the same 100mL overnight culture and then induced with 0mM (absolute ethanol), 1mM, 2mM, 3mM or 4mM indole after 3 hours of growth. The arrow indicates the time of induction or absolute ethanol addition. The sample at 3 hours of growth was taken immediately prior to induction, hence only samples at 4-9 hours of growth are post-induction samples.

Although the results shown in Figure 6.3.1.1. suggest that quiescence is maintained more effectively in the 3mM induced culture in terms of inhibiting biomass production, DNA analysis showed no significant increase in specific plasmid DNA yields (Figure 6.3.1.2.1. & Figure 6.3.1.2.2. and Table 6.3.1.).

Induction with up to 2mM indole had no observed effect on the specific pDNA yield per OD or %SC, whereas induction with concentrations of indole above 2mM resulted in decreased yields of specific pDNA per OD (when compared to the 0mM control) and no significant change to %SC. The slight decline in specific pDNA yields per OD as the indole concentration increases is observed again in Figure 6.3.1.2.2. F, a dose response graph illustrating the effect on indole concentration on specific pDNA yields at 6 and 8 hours of growth (generally across all cultures the specific pDNA yields are at their peak at either 6 or 8 hours growth).

The onset of quiescence is reported to delay cell division until multimer resolution is complete (Chant et al., 2007; Rowe and Summers, 1999), however higher order oligomers of gWiz are still present in the gel scans (Figure 6.3.1.3.) suggesting that multimer resolution did not take place. This can perhaps be explained by the fact that the gWiz plasmid does not have the *cer* site for the *cer*-acting recombinases to act upon; hence these higher order oligomers are not resolved.

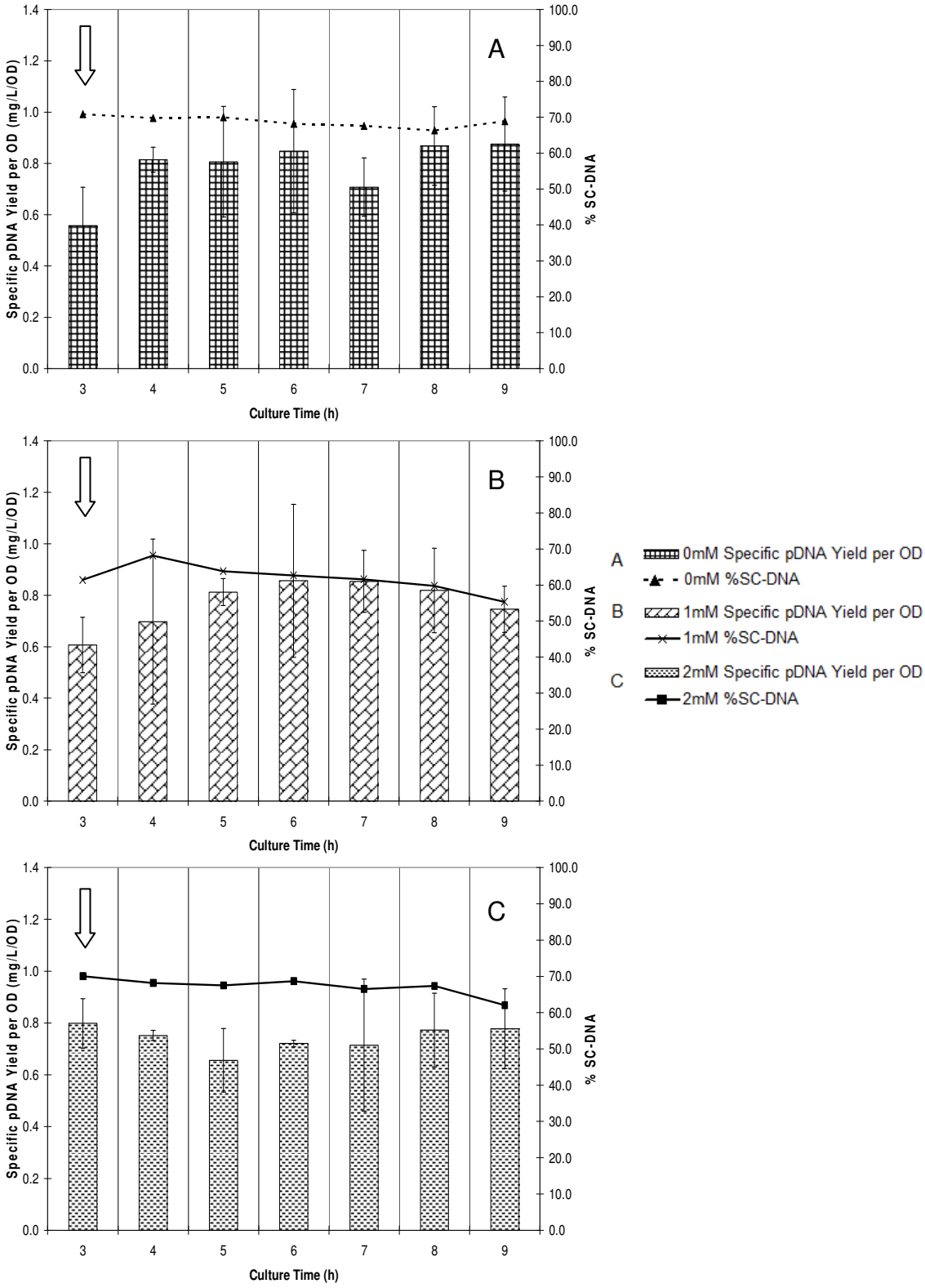


Figure 6.3.1.2.1. Specific pDNA yields and %SC-DNA obtained from W3110 Hns 205::Tn10 gWiz 100mL shake flask cultures grown at 37℃, 250rpm (averaged from duplicate experiments). For each set of experiments, cultures were inoculated with 10mL of the same 100mL overnight culture and then induced with 0mM (absolute ethanol) (A), 1mM (B) or 2mM (C) indole. The arrow indicates the time of indole induction or absolute ethanol addition. The sample at 3 hours of growth was taken immediately prior to induction. Error bars indicate standard error of the mean of the specific pDNA yield from the duplicate cultures.

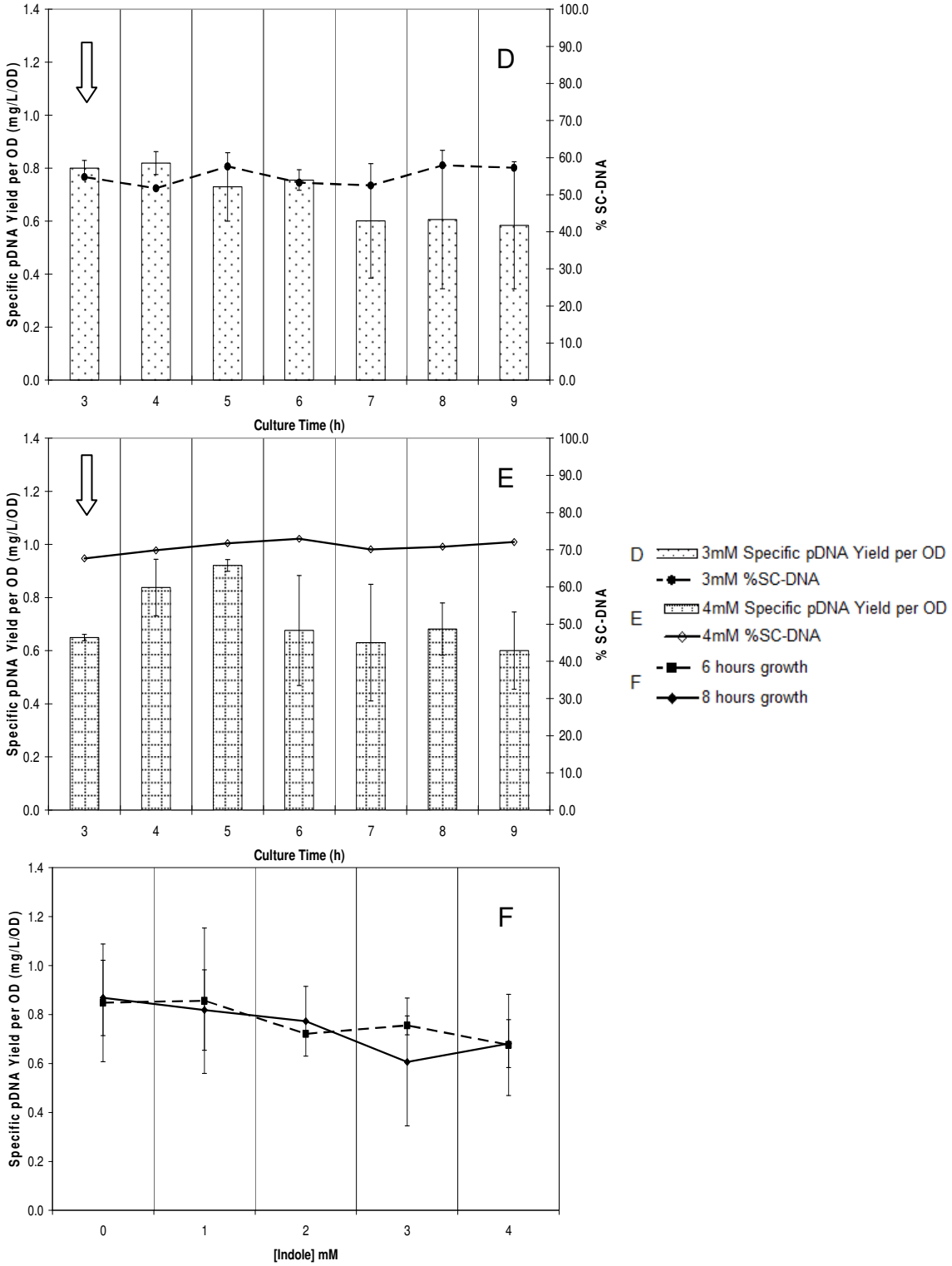


Figure 6.3.1.2.2. Specific pDNA yields and %SC-DNA obtained from W3110 Hns 205::Tn10 gWiz 100mL shake flask cultures grown at 37℃, 250rpm (averaged from duplicate experiments). Cultures were inoculated with 10mL of the same 100mL overnight culture and then induced with 3mM (D) or 4mM (E) indole. The arrow indicates the time of indole induction. The sample at 3 hours of growth was taken immediately prior to induction. (F) Dose response curves showing the effect of indole concentration on the specific pDNA yield obtained at 6 and 8 hours of growth from these cultures. Error bars indicate standard error of the mean of the specific pDNA yield from the duplicate cultures.

Table 6.3.1. Specific pDNA yields (mg/L/OD) and %SC-DNA obtained from Q-cell gWiz cultures pre- and post-induction between 3 and 9 hours of growth

[Indole] (mM)	3 hours of growth (pre-induction)		4 hours of growth (1 hour post-induction)		6 hours of growth (3 hours post-induction)		9 hours of growth (6 hours post-induction)	
	Specific pDNA yield (mg/L/OD)[b]	%SC-DNA ±20% error[c]	Specific pDNA yield (mg/L/OD)[b]	%SC-DNA ±20% error[c]	Specific pDNA yield (mg/L/OD)[b]	%SC-DNA ±20% error[c]	Specific pDNA yield (mg/L/OD)[b]	%SC-DNA ±20% error[c]
0[a]	0.56 ± 0.15	71	0.81 ± 0.05	70	0.85 ± 0.24	68	0.87 ± 0.18	69
1	0.61 ± 0.11	61	0.70 ± 0.32	68	0.86 ± 0.30	63	0.75 ± 0.09	55
2	0.80 ± 0.10	70	0.75 ± 0.02	68	0.72 ± 0.01	69	0.78 ± 0.15	62
3	0.80 ± 0.03	55	0.82 ± 0.04	52	0.76 ± 0.04	53	0.58 ± 0.24	57
4	0.65 ± 0.01	68	0.84 ± 0.11	70	0.68 ± 0.21	73	0.60 ± 0.15	72

[a] Absolute ethanol
[b] Standard error of the mean from duplicate cultures
[c] Error associated with densitometry method (Projan et al., 1983)

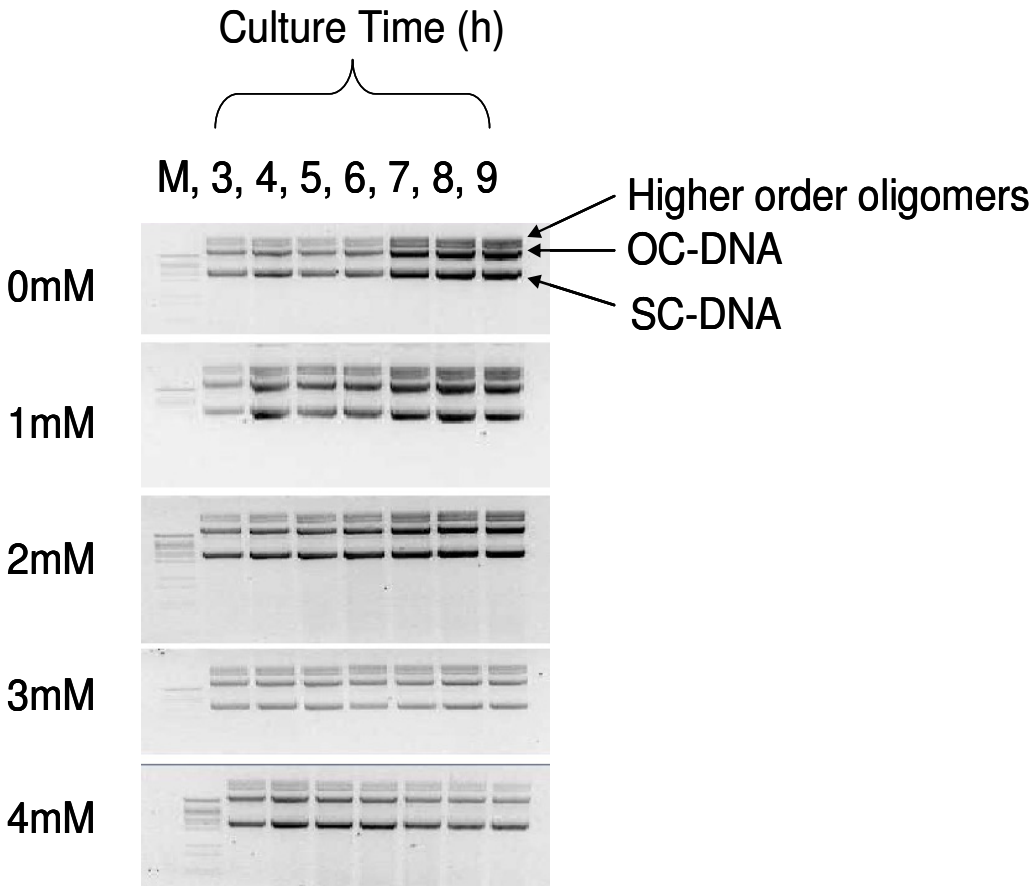


Figure 6.3.1.3. Gel scans of pDNA (gWiz) samples from each 100mL W3110 Hns 205::Tn10 gWiz culture, grown on LB broth at 37℃, 250rpm. The cultures were induced with different concentrations of indole (or absolute ethanol in the 0mM culture) after 3 hours of growth (sample 3 is a pre-induction sample). 15µL of miniprepped pDNA, 5µL of DNA loading buffer and 10µL of λBstEII digest marker (M) were loaded onto 1xTBE 0.8% agarose gels, pre-stained with ethidium bromide and ran at 85V for 2 hours.

6.3.2. Effect of Cell Density

Chant et al., (2007) reported that the level of intracellular indole produced by the cells is dependent on the phase of growth, with more indole produced during the late exponential/early stationary phase of growth i.e. at a higher cell density. Indole is removed from the cell by xenobiotic exporters, but in low density cultures, the exported indole will be insufficient to significantly increase the extracellular indole concentration. To investigate if these changes would have any effect on plasmid production, different volumes of the same overnight culture were used to inoculate 100mL shake flask cultures and these cultures were then induced after 3 hours of growth with 3mM indole.

The results indicate that the density of the culture does have an effect on how quickly the cells can be induced into a quiescent state. Figure 6.3.2.1. shows that quiescence is achieved almost immediately at higher cell densities (cultures induced with a larger volume of inoculum), but requires an additional two hours to reach quiescence in lower cell density cultures. Chant et al., 2007, reported that late exponential/early stationary growth phase cultures resulted in higher intracellular indole production. As the concentration of intracellular indole was not measured in these experiments, it can only be assumed that Chant et al., (2007)'s observations are also valid for these cultures. Subsequently, these findings can only suggest that the quiescent state is reached much faster at a higher cell density due to increased intracellular indole concentration.

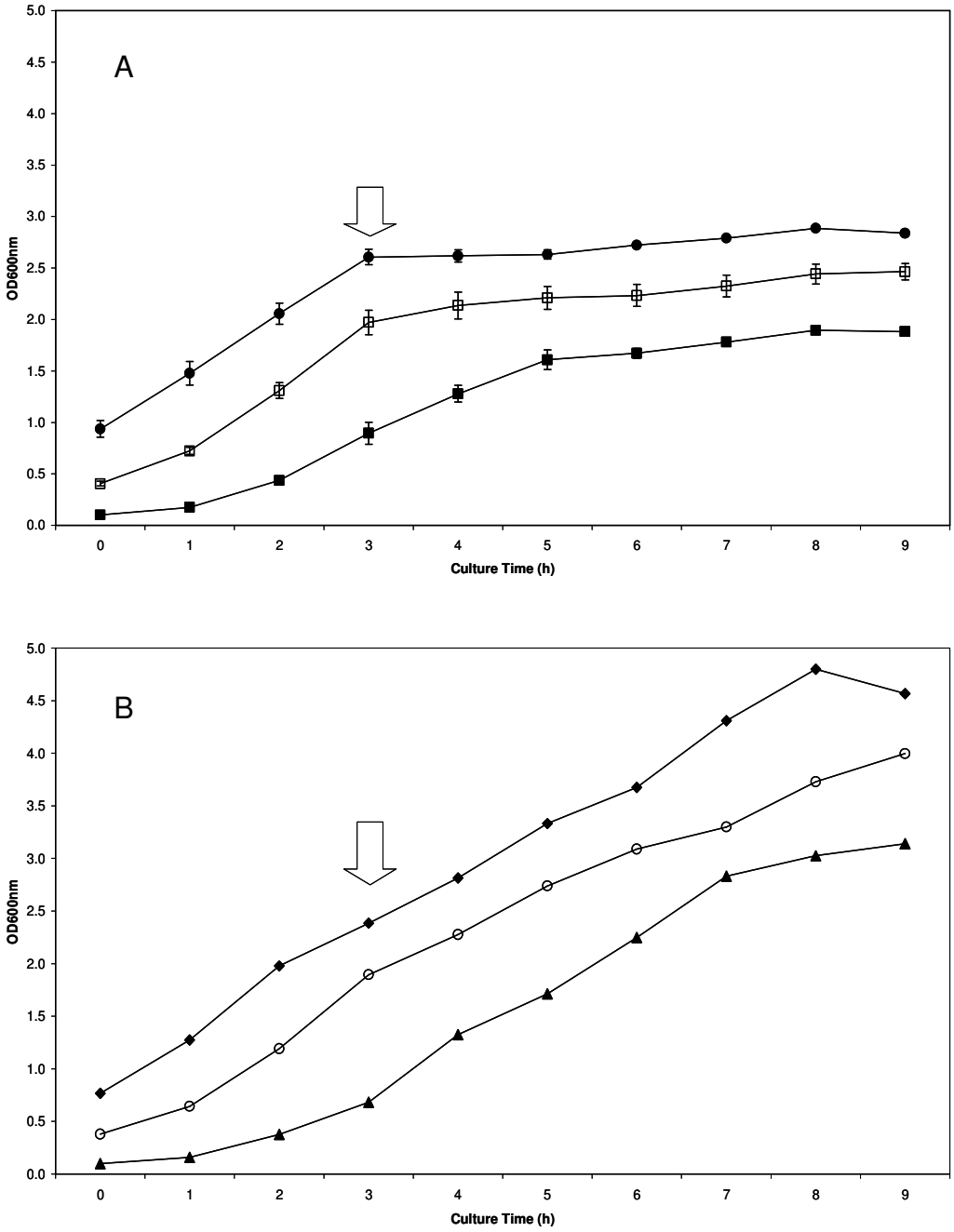


Figure 6.3.2.1. Average OD_{600nm} measurements from duplicate 100mL cultures induced with 3mM indole (A) and 0mM control cultures (B) grown at 37℃, 250rpm. For each set of experiments, the cultures were inoculated with 2mL, 10mL or 20mL of inoculum from the same overnight culture. Arrow indicates the time of induction or absolute ethanol addition. Error bars indicate the standard error of the mean from the duplicate cultures.

The results suggest that immediate quiescence can only be achieved through exogenous indole addition if there is already sufficient cellular indole, either through intracellular production during late exponential/early stationary phase; or in a high cell density culture. The effect of the latter condition is possibly due to the effects of indole acting as signalling molecule in high cell density cultures where the cells are closely packed together. In theory if indole is present in sufficient levels *in vivo*, then perhaps it is possible for a signal to be transferred to neighbouring cells via close contact. This theory is plausible since has been reported that indole acts as a stationary phase extracellular signal for many bacteria (Hirakawa et al., 2005). Although indole is a signalling molecule which shares some modes of action with quorum sensing systems, it itself is not a quorum sensing molecule (Di Martino et al., 2003).

Indole is removed from the cell by xenobiotic exportation (Chant et al., 2007 and Hirakawa et al., 2005), a type of multidrug resistance mechanism (MDR) developed by an organism to resist certain xenobiotic (considered foreign or toxic to the organism) drugs. MDR transporters can be split into two main groups according to the mode of energy coupling to drug export across the cytoplasmic and other membranes. One of these is the ATP binding cassette (ABC) group, which utilises the free energy of ATP hydrolysis to expel the drug from the cell against its concentration gradient. *E. coli* contains five putative ABC-type MDR transporters and these are likely to be involved in the transportation of specific subsets of molecules, such as metabolism-derived compounds (Lubelski et al., 2007). The other main MDR transporter group consist of proton motive force (PMF) dependent transporters, secondary transporters that utilise the PMF (H^+) or sodium (Na^+) motive force for drug expulsion. In *E. coli* seven different proton-dependent MDR pumps have been identified (Sulavik et al., 2001), one of which is AcrEF. AcrEF has been reported by Kawamura-Sato et al., 1999, to excrete indole.

Other than acting as a signalling molecule, indole has also been shown to induce the expression of a number of MDR pumps (Hirakawa et al., 2005).

Subsequently, whether in conjunction of indole acting as a signalling molecule or not, it is plausible that quiescence is achieved sooner in high cell density cultures because there is more total intracellular indole available to induce MDR pump expression. This elevated MDR expression would increase the extracellular indole concentration significantly, enough to help induce quiescence quickly when indole is added to the medium. Although indole is released when tryptophanase cleaves tryptophan, Summers and Chant (2009), have reported that supplementing the growth medium with 0.5mM tryptophan does not make a difference to the response of the cultures to indole.

6.3.3. Plasmid DNA Production During Quiescence

6.3.3.1. 2L Shake Flask Cultures

Plasmid DNA analysis of samples from 2L shake flask cultures (400mL working volume) confirms the observations made with the 500mL cultures (100mL working volume) in that indole induction does not appear to offer significant additional benefit over non-induced cultures in terms of plasmid DNA production. The specific plasmid DNA yields obtained at 400mL induced culture were not as high as those achieved from the non-induced culture of the same scale (Figure 6.3.3.1.) suggesting that instead, induction may be inhibiting plasmid production.

6.3.3.2. 2L Fermentation Culture

At 2L fermentation scale, both an increase and then a drop in pDNA yields were observed during quiescence (Figure 6.3.3.2.). Quiescence was achieved and maintained for a minimum of 6 hours (no samples were taken between 7 and 18 hours post-induction, hence it is unknown whether the cells remained in quiescence during this period). DNA analysis conducted show that the cells continued producing pDNA for another 4 hours post-induction, after which the yield of pDNA began to drop.

Since the 2L shake flask cultures conducted previously already show that indole induction does not appear to offer any additional benefits in terms of plasmid production when compared to non-induced cultures, a comparable non-induced fermentation culture was not carried out. Instead, the aim of conducting the 2L fermentation was to confirm if plasmid production continues post-induction. The results shown in Figure 6.3.3.2. do arguably confirm and show in more detail (when compared to the shake flask cultures) that quiescence does not stop plasmid production immediately as it does with biomass production. Whether this brief continuation of plasmid production is a result of indole induction remains inconclusive.

Overall, the results obtained from all the cultures analysed, suggest that an achievable ‘Quiescent cell system’ does not seem to increase plasmid DNA production as previously reported for plasmid-encoded protein expression (Rowe and Summers 1999, Mukherjee et al., 2004).

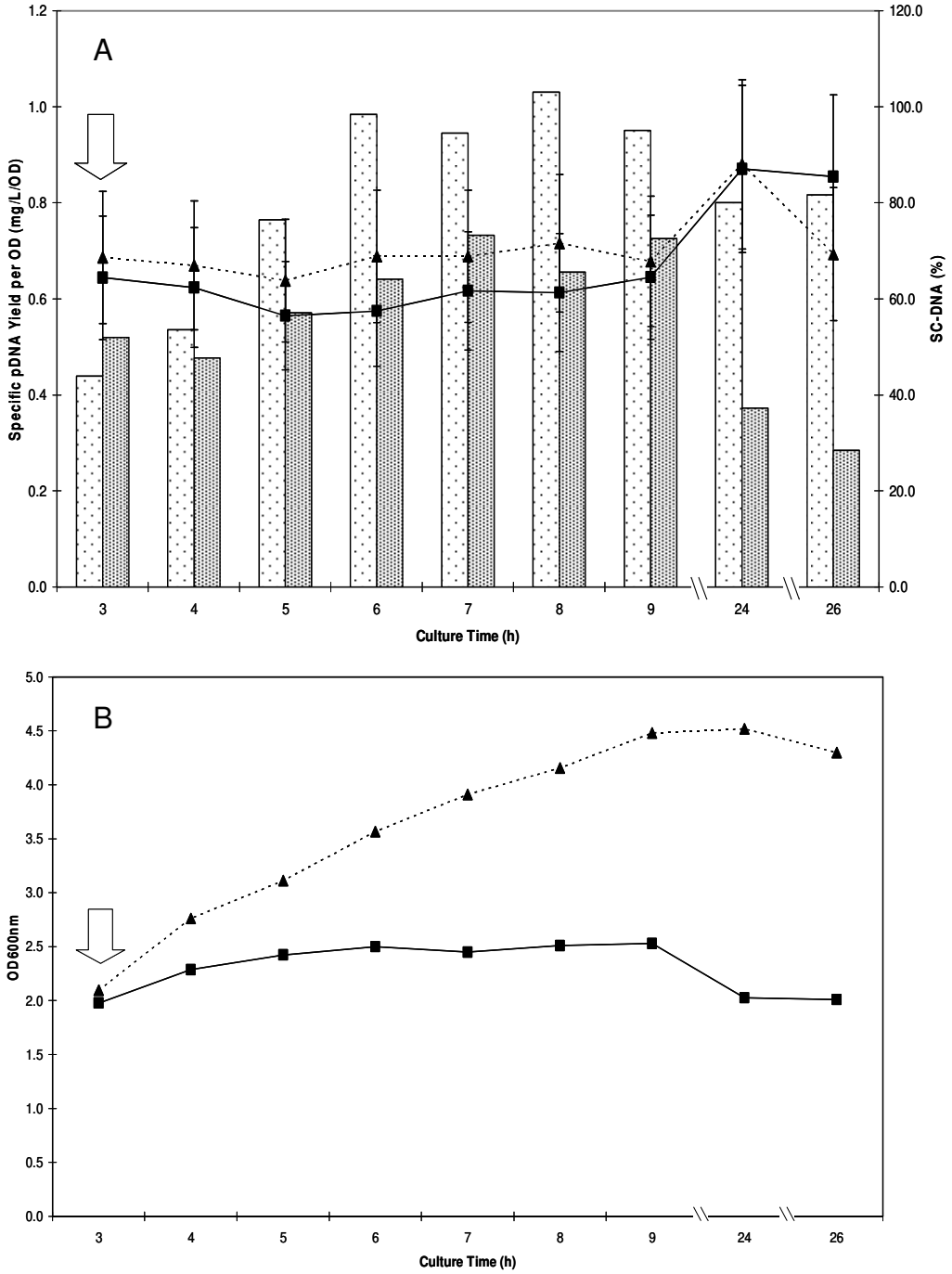


Figure 6.3.3.1. (A): Specific pDNA yield per OD and %SC-DNA obtained from 400mL working volume W3110 Hns 205::Tn10 gWiz quiescent and non-quiescent (control) cultures grown at 37℃, 250rpm in 2L shake flasks containing LB broth. Arrow indicates time of 3mM indole induction in quiescent cultures (sample at 3 hours taken pre-induction). (B) Growth profiles (optical density at 600nm) of quiescent and non-quiescent cultures.

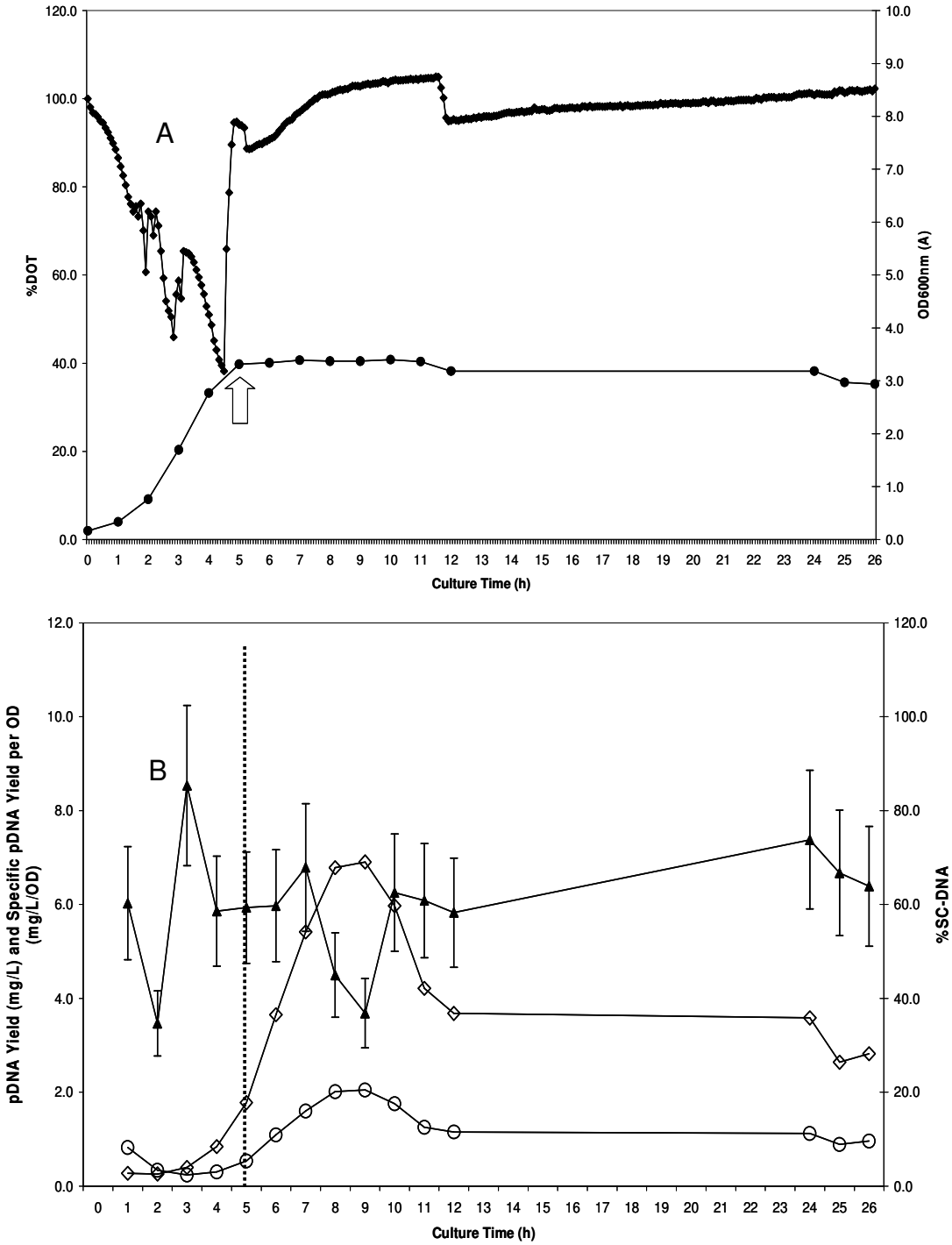


Figure 6.3.3.2. A & B: show the DOT, biomass (A) and plasmid DNA profiles (B) obtained from a 2L fermentation Q-cell gWiz culture (1.5L working volume) grown on LB broth at 37℃, 600rpm. The culture was induced with 3mM indole after 5 hours of growth (sample at 5 hours is pre-induction). Error bars indicate ±20% error associated with densitometry method (Projan et al., 1983). Induction time is indicated by arrow (A) or dashed line (B) (sample at 5 hours taken pre-induction).

6.4. Protein Expression During Quiescence

It is possible that the down-regulation of chromosomal genes, together with other side effects of quiescence may have affected the cells' ability to redirect available resources freed up from cell non-division into plasmid replication. The cellular machinery required for protein synthesis and folding is not affected in this way and this has been confirmed in 2L shake flask cultures (400mL working volume) using the plasmid pGlo (5.4kb) that expresses GFP in *E. coli* upon arabinose induction. In these cultures the specific activity per OD (specific activity proportional to the fluorescence signal (RFU) divided by the OD at time of cell lysis) in induced cultures were higher after 3 hours post-indole induction than in non-induced cultures indicating that plasmid-borne protein expression continues and actually increases during quiescence (Figure 6.4.1.).

Although bacterial GFP expression seems to continue and increase in quiescent cells in this study, it has been reported by Rowe and Summers, (1999), that not all plasmid-encoded proteins are continually expressed during quiescence. In their study using the proteins CAT and cI*ts*857, the latter protein did not show the same high-level of accumulation in quiescent cells as CAT. Instability and degradation at the temperature required for Rcd expression (42°C) or perhaps differences between the efficiency of transcription and translation of these two proteins may have been the cause of this.

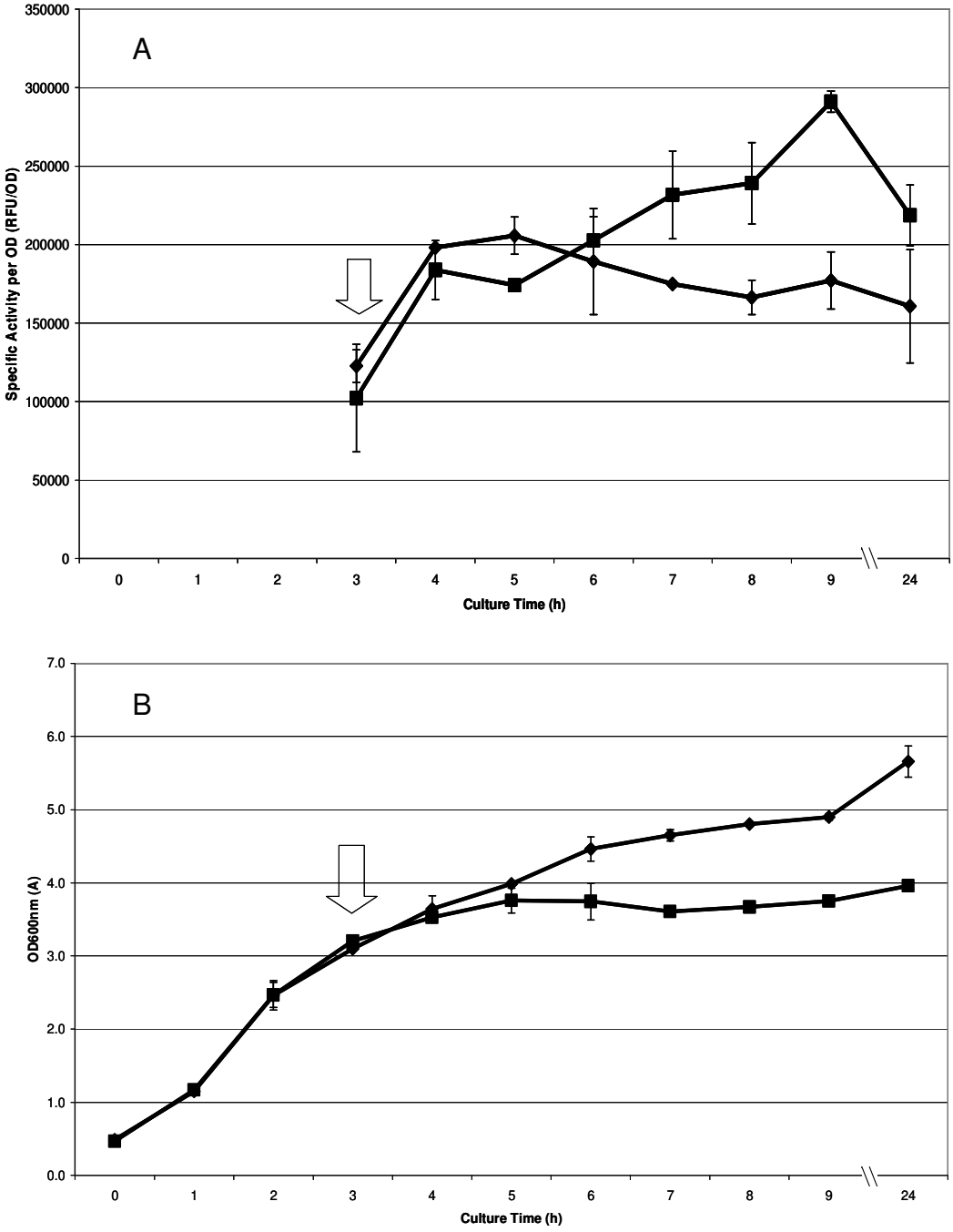


Figure 6.4.1. A & B: Specific activity per OD (RFU/OD) (A) and OD profile (B) of W3110 Hns 205::Tn10 pGlo samples treated with Bugbuster® HT Protein Extraction Reagent (Novagen). Triplicate readings were obtained from two sets of 400mL working volume cultures (2x non-quiescent (control), 2x quiescent) grown on LB broth at 37°C, 250rpm. Error bars indicate the standard error of the mean. Fluorescence intensity (RFU) was measured at excitation and emission wavelengths of 395nm and 509nm respectively. 0.25% (w/v) L-(+)-arabinose was added to cultures 1 hour after inoculation (not shown in figure) to induce bacterial GFP expression. Quiescence was induced 3 hours after inoculation as indicated by the arrow in the figure (sample shown at 3 hours is pre-induction sample). —◆— Control —■— Quiescent

6.5. Conclusion

The aim of this study was to investigate whether a 'Quiescent cell system' can increase plasmid DNA production in *hns*-mutated *E. coli* cells. It was demonstrated that although plasmid DNA continues to increase initially during quiescence, this is not maintained. The pDNA yield obtained is lower from quiescent cultures than from non-quiescent cultures, suggesting that prolonged quiescence has a detrimental impact on pDNA production. Normally, the presence of plasmid multimers may depress the copy number of each plasmid, causing instability and thereby increasing the number of plasmid-free cells (Summers, 1998). As the cells are not dividing in a 'Quiescent cell system', segregational instability should not be an issue and therefore it would be an unlikely source of instability.

The concept of inhibiting chromosomal DNA replication in a 'Quiescent cell system' is similar to the effect of chloramphenicol amplification, whereby the addition of chloramphenicol at the mid-log phase of growth inhibits protein synthesis allowing replication of relaxed plasmid DNA long after chromosomal replication has ceased (Clewell, 1972; Clewell and Helinski, 1972; Reinikainen et al., 1989).

Chloramphenicol inhibits protein synthesis by binding to the 50S ribosomal subunit where it blocks the peptidyltransferase reaction during translation (Foster, 1983; Gale et al., 1981; Garrett et al., 2000). This is turn prevents the initiation of chromosomal DNA replication, however, active (on-going) chromosomal replication is completed (Seo, 1986). Plasmid DNA initiation does not appear to be coupled with the stage of replication of the host chromosome since pDNA replication continues to a level determined by growth conditions long after chromosomal DNA replication has ceased, yielding an increase in plasmid content relative to the chromosomal DNA (Clewell, 1972; Clewell and Helinski 1972). Plasmid DNA replication continues in cells treated with chloramphenicol because relaxed plasmids do not require any plasmid-encoded proteins for replication *in vivo* (Crosa et al., 1975; Hershfield et al., 1974). However it is assumed that the

replication process is limited by biosynthetic precursors, replication enzymes and initiation factors, and metabolic energy (Clewell, 1972).

In a 'Quiescent cell system' chromosomal expression is shut down, presumably due to the condensation of chromosomal DNA as a result of indole induction into a quiescent state; however, unlike the effects of chloramphenicol amplification, *de novo* protein synthesis reportedly continues in quiescent cells (Rowe and Summers, 1999). While chloramphenicol addition has been proven to amplify plasmid DNA replication relative to chromosomal DNA, our investigation has found no significant amplification in plasmid DNA content when employing a 'Quiescent cell system'. Instead, this study has found that the equivalent yield of plasmid produced in a quiescent culture is lower than in non-quiescent cultures. Table 6.5.1. compares the methodologies of the 'Quiescent cell system' with chloramphenicol plasmid DNA amplification.

Table 6.5.1. A comparison of the chloramphenicol plasmid DNA amplification and 'Quiescent cell system' methodologies

	Chloramphenicol plasmid DNA amplification	**Indole induced 'Quiescent cell system'**
Concept	Chloramphenicol inhibits protein synthesis by binding to the 50S ribosomal subunit where it blocks the peptidyltransferase reaction during translation. Initiation of chromosomal DNA replication is prevented but active chromosomal DNA replication is completed. Plasmid DNA replication continues even after chromosomal replication has ceased. Plasmid DNA yield is amplified relative to chromosomal DNA.	Growth arrest is achieved after *E. coli* cells containing the *hns205* mutation are induced with indole. Nucleoid condensation leads to the shut down of chromosomal expression and the preferential expression of plasmid genes. Cells remain metabolically active and *de novo* transcription and translation continues. The copy numbers of some plasmids are also reportedly amplified during quiescence.
Limitations	Plasmid DNA replication continues until resources needed such as biosynthetic precursors, replication enzymes and initiation factors, and metabolic energy are used up. The final plasmid content obtained depends on the growth medium used.	Cells have been maintained in a stable quiescent state for at least 24 hours; however the accumulation of indole in a culture (through addition and intracellular production) can cause cell lysis. Not all plasmid-borne proteins are expressed during quiescence. This study has shown that the plasmid DNA yield is not amplified during quiescence.
Literature sources	Clewell, (1972); Clewell and Helinski, (1972); Crosa et al., (1975); Hershfield et al., (1974); Reinikainen et al., (1989); Seo, (1986).	Rowe and Summers (1999); Mukherjee et al., (2004); Chant et al., (2007) and Summers and Chant, (2009).

In addition to replication enzymes and initiation factors, the regulation of some chromosomal genes, such as carbon transport genes, are required to maintain plasmid replication. As plasmid DNA production has been observed to continue for a short while after the onset of quiescence, there is perhaps a delayed response in the down-regulation of these chromosomal genes, resulting in a delayed effect on plasmid DNA production. Wang et al., (2006), conducted DNA microarray experiments to analyse transcription profiles in BL21, BL21 containing one of two plasmids derived from ColE1, pOri1 and pOri2. The authors found that a number of carbon transport genes were up-regulated significantly in the plasmid-containing cells, suggesting that these groups of genes were helpful for maintaining plasmid DNA in *E. coli*. The authors also concluded that the *rpiA* gene (encoding ribose-5-phosphate isomerase A) is one of the important limiting factors for ColE1-type plasmid DNA replication.

Metabolic pathways such as glycolysis, the tricarboxylic acid (TCA) cycle and the pentose phosphate pathway are involved in the generation of NADPH for biosynthesis, recruiting essential metabolites for nucleic acids and amino acids. These are important for plasmid DNA replication, especially during the replication of high copy number plasmids where an extra synthesis of nucleotides is needed. A dependence of plasmid DNA synthesis on RNA synthesis has also been reported at a time when protein synthesis is no longer required (Lark, 1972). It is possible that if any of these pathways or genes on the host chromosome are affected or down-regulated during quiescence, then plasmid replication can also be affected.

Alternatively, if plasmid DNA production during quiescence is placing substantial metabolic demands on the host cells, then potential changes in fitness may lead to plasmid DNA degradation and cell lysis. This may explain why specific pDNA yields obtained from quiescent cultures in this study appear to drop slightly with time after the cells have been induced into a quiescent state. Utilising a low copy number plasmid may reduce the level of stress on the host cells during the growth phase (pre-induction

period) of the culture; however, it remains to be seen if a significant increase in copy number during quiescence (when compared to a non-quiescent culture) is observed when using a low copy number plasmid.

Although this study has shown that a 'Quiescent cell system' does not appear to offer any significant advantages over a non-quiescent system in terms of plasmid DNA production; it has confirmed that a 'Quiescent cell system' can be used to increase the expression of GFP. This is in agreement with studies that have also shown an increase in the expression of some plasmid-encoded proteins such as CAT and scFv, although it has been reported that this is likely to be protein specific (Rowe and Summers, 1999, and Mukherjee et al., 2004).

7. Conclusion and Future Work

7. Thesis Conclusion and Future Work

This project has examined three main areas of optimised plasmid DNA production; investigating the influences of the *E. coli* host strain on three different plasmids, studying the implementation of anaerobiosis to improve supercoiled plasmid DNA yields, and examining whether a 'Quiescent cell system' for plasmid production is viable. Over the course of the investigation, other factors of interest emerged such as the combined effects of the host strain-plasmid system and increased temperature on plasmid yields.

The discussion and conclusions of the results have been covered in depth within each of the results chapters. This chapter summarises the findings of the thesis and discusses potential future work.

7.1. Host Strain Influences on Supercoiled Plasmid DNA Production:

It was demonstrated in Chapter 3, at 10mL and 100mL scales, and verified at 5L scale in Chapter 5, that the host strain-plasmid combination has a significant impact on plasmid DNA yield and quality. Seventeen *E. coli* strains and three plasmids (gWiz, 5.8kb; pSVβ, 6.9kb and pQR150, 20kb) were investigated. *E. coli* strains BL21 DE3, HB101, DH5α and TG1 were found to produce large yields of volumetric and specific plasmid DNA when harbouring either of the two smaller plasmids (gWiz and pSVβ). Interestingly some strains yielded more DNA for one of these plasmids but not for the other and some strains that look genotypically very similar (e.g. DH10β and TOP10) behaved differently with the same plasmid. The effects of certain specific or known mutations that were introduced into several laboratory *E. coli* strains to improve their performance, such as *endA1* and *recA* (Carnes, 2005; Schoenfeld et al., 1995), were also found to be strain/plasmid dependent. The results demonstrated that it is not possible predict the productivity and quality of plasmid DNA produced in that strain by known genotypic information alone. The accumulation of specific and unknown mutations during the last 88 years has made it very difficult if not impossible to define all the mutations in these strains and the

effects of such mutations; therefore the early screening of host-plasmid performance before further large scale optimisation or verification takes place is highly recommended, as it can potentially reduce manufacturing costs by improving the quantity and quality of the plasmid product obtained.

However, it should be noted that this recommendation may not be feasible industrially. Commercial products for human use require that all cell lines and strains involved in the product manufacturing process to comply with the criteria and guidelines set out by regulatory bodies such as the FDA and EMEA. In such cases, detailed information on each cell line in a GMP cell bank such as safety and characterisation data, ownership/IP, source traceability and genealogy is required as part of the overall product and process validation program. To ensure each strain or cell line in a Master cell bank meets all of the necessary criteria, the validation process is expensive and time consuming (Nail and Akers, 2002); therefore it is likely that many commercial organisations may have only a few cell lines in their working cell bank. In light of this, screening a large selection of hosts would not be feasible since only a few prospective host cells would be available for screening.

The scientific reasons behind the findings of Chapter 3 may offer an opportunity for future work certainly in the area of cell engineered strains. Strain and vector engineering strategies have been developed to provide improve pDNA sequence stability, yields and safety, and to address upstream and/or downstream issues (Bower and Prather, 2009; Choi et al., 2003; Han et al., 2003; Hodgson and Williams, 2006; Ow et al., 2009 and Posfai et al., 2006). These reports have proposed strategic modifications to *E. coli* strains including: single gene modifications to improve specific pDNA yields (Ow et al., 2009) or plasmid segregational stability (Zhao et al., 2007); and genome reduction to enable stable propagation of previously unstable recombinant genes (Posfai et al., 2006).

An interesting step would be to look into inserting/deleting selective mutations into older and less mutated strains such as MG1655 or W3110 and see whether these mutations or deletions improve strain performance when compared to a defined strain. Similarly, another example would be to introduce mutations to HB101, which has been shown in this study to be able to maintain high yields of supercoiled plasmid (gWiz), to lower its carbohydrate content and in turn mitigating any issues that have been suggested to occur during the downstream purification stages.

7.2. Interaction between Host Strain Selection and Temperature Amplification on Plasmid DNA Production

Temperature amplification studies conducted using strains harbouring pUC18 have shown that although most strain-plasmid combinations yielded more plasmid at a higher temperature of 42°C, the extent of this increase is highly influenced by the host strain. Indeed in some cases, such as for the strains ABLE K, W3110, W1485, a higher plasmid yield was obtained at 37°C (by at least 2- fold compared to plasmid yields obtained at 40°C).

One important issue that still needs to be addressed is whether the observed accumulation of plasmid DNA at 40°C compared to at 30°C and 37°C is growth rate independent. As the growth rates of the 10mL cultures were not measured, the extent of the accumulation of plasmid DNA due to the effects of the growth rate and/or temperature during the exponential phase of growth is unknown at this time. Further investigations that involve measuring i) the growth rate profiles, whether at a larger scale or by measuring samples in microwell plates, ii) the specific pDNA yield and iii) the productivity of each strain-plasmid combination at these three temperatures (30°C, 37°C and 40°C) are needed to determine this.

Microscopic images of plasmid-containing and plasmid-free cells show that some host strain-plasmid combinations (MG1655 gWiz, HB101 gWiz and TG1 gWiz) appear to exhibit morphological response (cell elongation and increase in diameter) which may or may not be associated with higher

plasmid copy number. Apart from the possible association with higher plasmid copy number, morphological changes of *E. coli* have been reported to be caused by other factors including the effect of inclusion bodies, excessive accumulation of the polymer PHB and a response to shifts in temperature (Bowden et al., 1991; Hart et al., 1990; Lee et al., 2008; Williams et al., 1982; Lee et al., 1994; Rinas et al., 1993; Wong et al., 1982).

Similarities to what has been reported as temperature induced runaway plasmid replication have been observed in this study, although no experiments were conducted to confirm whether these observations were indeed as result of runaway replication as defined in the literature. To do so, the characteristics of runaway replication need to be clearly defined as do the characteristics of "normal" replication.

The increase in plasmid copy number seen by virtue of the mutation in the pUC series that lessens the interaction of RNA I and RNA II at 40-42°C (Lin-Chao et al., 1992) could be described as a stage on the way to runaway replication and what one may be seeing in the host-plasmid combinations that give very large copy number increases either between temperatures or between strains, is some host mediated lessening of the control of RNA I over RNA II. What the mechanism of this host mediated enhancement of copy number is remains to be seen.

7.3. 7L Fermentations

7.3.1. Scale up (to 7L) of selected host-plasmid systems

Results from 7L (5L working volume) fermentation runs of HB101 gWiz and BL21 DE3 gWiz grown on SDCAS medium at 37°C, show that HB101 gWiz was able to maintain a high percentage of DNA supercoiling (98%) throughout the course of the fermentation and provide 12mg/L volumetric and 0.70mg/L/OD specific pDNA yields; whereas BL21 DE3 gWiz provided high volumetric and specific yields (23mg/L and 1.49mg/L/OD), but a low percentage of supercoiled DNA (63%). Plasmid supercoiling has

been shown to be significantly affected during the downstream recovery and purification processing stages (Ciccolini et al., 2002; Ferreira et al., 2000; Kong et al., 2006; Levy et al., 2000, Prazeres et al., 1998), therefore the ability to maximise supercoiled plasmid DNA production during the upstream stages is important to the final yield of supercoiled plasmid DNA obtained.

Carnes et al., (2005), and Prazeres et al., (1999), have suggested that *E. coli* HB101 and other high carbohydrate producers should be avoided as host strains at large scale because the high carbohydrate content can foul downstream processing operations. However to date, no such downstream studies using this strain have been discussed in the open literature, so it is not known if this has actually been proven or if it has just been anecdotally reported. The carbohydrate content of HB101 was not measured during this study so these reports cannot be confirmed or disproved.

7.3.2. Effects of Anaerobiosis on Supercoiled Plasmid DNA Production

It has been reported that anaerobiosis increases plasmid supercoiling through changes in the linking number distribution (Hsieh et al., 1991a; Cortassa and Aon, 1993); however it has not been confirmed if anaerobiosis effecting changes in topoisomerase I and DNA gyrase activity, would shift the linking number distribution enough to convert relaxed DNA into supercoiled plasmid DNA, thus increasing the proportion of supercoiled DNA in a population. The main finding from this chapter is that under the conditions investigated, the results obtained show that anaerobic shock or anaerobiosis does not increase the percentage of supercoiled plasmid DNA in the host plasmid system studied (BL21 DE3 gWiz).

Any increases in plasmid supercoiling through changes in linking number distribution (Hsieh et al., 1991a; Cortassa and Aon, 1993) could not be confirmed in this study as multiple attempts at running chloroquine gels to

resolve the plasmid DNA samples into topoisomers were unsuccessful and the optimum chloroquine gel conditions for gWiz obtained from BL21 DE3 remains to be established before linking number determination can be carried out. Further work investigating the whether there is a correlation between the linking number distribution and the percentage of plasmid supercoiling in a population is currently being conducted by the Department of Biochemical Engineering, UCL.

Further work could also consider other strategies such as osmotic stress or salt shock with NaCl (Conter et al., 2003 and Hsieh et al., 1991a) and implementing anaerobiosis during the stationary phase of growth. The stationary phase of growth is when DNA supercoiling levels are reportedly decreasing and are undergoing various metabolic changes (Goldstein and Drlica, 1984; O'Kennedy et al., 2003; Reyes-Dominguez et al., 2002). If anaerobiosis does indeed increase plasmid supercoiling levels through changes in the linking number distribution, then implementing it during the stationary phase prior to harvesting could potentially increase the yield of highly supercoiled plasmid DNA. From a processing view, hindering growth during exponential phase through anaerobiosis will only lead to a lower volumetric pDNA yield, whereas an aerobic growth before an anaerobic transition during late exponential/stationary phase will ensure a sufficient volumetric yield of pDNA is obtained before harvesting.

7.4. A 'Quiescent Cell System' For Plasmid DNA Production

A 'Quiescent cell system' (a state where chromosomal replication and expression is temporarily shut down but residual proteins remain metabolically active) (Rowe and Summers, 1999), was investigated to examine whether the reported benefits of such a system for protein expression can also benefit and improve plasmid production. It was demonstrated that although plasmid DNA continues to increase initially during quiescence, this is not maintained. The pDNA yield obtained is lower from quiescent cultures than from non-quiescent cultures, suggesting that prolonged quiescence has a detrimental impact on pDNA production. .

As discussed in Chapter 6, the concept of inhibiting chromosomal DNA replication in a 'Quiescent cell system' is similar to the effect of chloramphenicol amplification (Clewell, 1972; Clewell and Helinski, 1972; Reinikainen et al., 1989). Since chloramphenicol plasmid amplification was first reported in the 1970s, advances in plasmid research and design have resulted in improved plasmid characteristics such as increases in plasmid copy number. Except in cases where low copy number plasmids are specifically required, the copy numbers of some plasmids are now so high (e.g. pUC plasmids) that there is no longer the need for chloramphenicol amplification in order to achieve high yields of plasmid DNA. However, chloramphenicol is still added in cases where the goal is to reduce the bulk of bacterial cells harvested from large-scale preparations, specifically large amounts of viscous cell lysate (Sambrook and Russell, 2001). In such case, an equivalent yield of plasmid DNA can be obtained from smaller number cells that have been exposed to chloramphenicol as from larger number of cells that have not; easing and reducing the time and cost of downstream processing. The obvious drawback of this application is that the use of chloramphenicol for pharmaceutical products is discouraged by regulatory bodies to minimise the spread of antibiotic resistance traits to environmental microbes (FDA Points to Consider on Plasmid DNA Vaccines for Preventive Infectious Disease Indications, 1996a). The use of antibiotics is also not a viable option for gene therapy treatments that involve the direct introduction of plasmid DNA into patients (Cranenburgh et al., 2001).

As the cells in a 'Quiescent cell system' stop producing biomass but remain metabolically active, a 'Quiescent cell system' harbouring a plasmid with a non-antibiotic selection marker, can also potentially offer similar biomass-reducing benefits for pharmaceutical products without the drawbacks associated with chloramphenicol addition. However, as the findings of this study suggest, a 'Quiescent cell system' would not be beneficial for plasmid DNA production as such, but for the expression of some plasmid-borne proteins.

7.5. Overall conclusion

In conclusion, the results from this investigation have demonstrated that a highly effective and influential strategy for improving the quality and quantity of plasmid DNA obtained is the initial choice of the host strain-plasmid combination. Further improvements can then be obtained through the application of other strategies such as employing optimised fermentation strategies, high cell density cultures, optimised growth media and amino acid starvation to name a few (Hofmann et al., 1990; Kay et al., 2003; Lee, 1996; O'Kennedy et al., 2000 & 2003; Prazeres et al., 1999; Williams et al., 2009; Wrobel, B. and Wegrzyn, G. 1998).

8. References

8. References

Ada, G. 1997. Overview of vaccines. *Molecular Biotechnology* **8**: 123-134.

Al Fasbender, Lee, J.H., Walters, R.W., Moninger, T.O., Zabner, J., and Welsh, M.J. 1998. Incorporation of Adenovirus in Calcium Phosphate Precipitates Enhances Gene Transfer to Airway Epithelia In Vitro and In Vivo. *Journal of Clinical Investigation.* **102**: 184-193.

Appleyard, R.K. 1954. Segregation of New Lysogenic Types During Growth of a Doubly Lysogenic Strain Derived From *E. coli* K12. *Genetics* **39**: 440.

Arvin, A.M. and Greenberg, H.B. 2006. New viral vaccines. *Virology* **344**: 240-249.

Aunins, J. G., Lee. A. L., Volkin, D. B. 2000. The Biomedical Engineering Handbook: Second Edition. CRC Press LLC.

Bachmann, B.J. 1972. Pedigrees of Some Mutant Strains of *Escherichia Coli* K-12. *Bacteriological Reviews* **36**: 525-557.

Bachmann, B.J. 1996. Derivations and Genotypes of Some Mutant Derivatives of Escherichia coli K-12. *Escherichia coli* and *Salmonella*: cellular and molecular biology. 2nd ed. (Neidhardt, F.C., Curtiss, III R., Ingraham, J.L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., Umbarger, H. E., Eds), pp. 2460-2488. Washington, D.C.: ASM Press.

Balding, C., Blaby, I., Summers, D. 2006. A Mutational Analysis of the ColE1-encoded cell cycle regulator Rcd confirms its role in plasmid stability. *Plasmid* **56:** 68-73.

Balke, V.L. and Gralla, J.D. 1987. Changes in the Linking Number of Supercoiled DNA Accompany Growth Transitions in *Escherichia coli*. *Journal of Biotechnology*. **169**: 4499-4506.

Bauer, W. and Vinograd, J. 1968. The Interaction of Closed Circular DNA with Intercalative Dyes: The Superhelix Density of SV40 DNA in the Presence and Absence of Dye. *Journal* of *Molecular Biology **33**: 141-172.*

Bazaral, M. and Helinski, D. R. 1970. Replication Bacterial Plasmid and an Episome in *Escherichia coli*. *Biochemistry* **9**: 399-406.

Bellin, C.A. and Rao, P.S.C. 1993. Impact of Bacterial Biomass on Contaminant Sorption and Transport in a Subsurface Soil. *Applied and Environmental Microbiology.* **59**: 1813-1820.

Birnhoim, H.C. & Doly, J. 1979. A Rapid Alkaline Extraction Procedure for Screening Recombinant Plasmid DNA. *Nucleic Acid Research* **7**: 1512-1523.

Blattner F.R., Plunkett III, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., Shao, Y. 1997. The Complete Genome Sequence of *Escherichia coli* K-12. *Science*. **277**: 1453-1462.

Boyer, H.W. and Roulland-Dussoix, D. 1969. A Complementation Analysis of the Restriction and Modification of DNA in *Escherichia coli. Journal Molecular Biol*ogy **41**: 459.

Bowden, G.A., Paredes, A.M., Georgiou, G. 1991. Structure and Morphology of Protein Inclusion Bodies in *Escherichia coli*. *Nature Biotechnology.* **9**: 725-730.

Brown, T.A. Essential Molecular Biology. A Practical Approach. Volume 1, Second Edition. Oxford University Press. 2000.

Bullock, W.O., Fernandez, J.M., Short, J.M. 1987. XL1-Blue: A high efficiency plasmid transforming recA Escherichia coli strain with beta-galactosidase selection. *Biotechniques* **5**:376-378.

Carrillo, M., Estrada, E., Hazen, T.C. 1985. Survival and Enumeration of the Feral Indicators *Bifidobacterium adolescentis* and *Escherichia coli* in a Tropical Rain Forest Watershed. *Applied and Environmental Microbiology.* **50**:468-476.

Carnes, A. 2005. Fermentation Design for the Manufacture of Therapeutic Plasmid DNA . *BioProcess International* October 2005: 2-7.

Carnes, A., Hodgson, C., Williams, J. 2004. Optimization of *E. coli* Fermentation for Plasmid DNA Production. *Molecular Therapy* **9**: S310, 818.

Carnes, A.E., Hodgson, C., Williams, J.A. 2006. Inducible *Escherichia coli* fermentation for increased plasmid DNA production. *Biotechnology and Applied Biochemistry* **45:** 155-166.

Capaldo, F.N., Ramsey, G., Barbour, S.D. 1974. Analysis of the Growth of Recombination-Deficient Strains of *Escherichia coli* K-12. *Journal of Bacteriology* **118**:242-249.

Casdaban, M.J. and Cohen, S.N. 1980. Analysis of Gene Control Signals by DNA Fusion and Cloning in *E. coli. Journal Molecular Biol*ogy **138**:179-207.

Cashel, M. & Rudd, K. E. 1987. The Stringent Response. In *Escherichia coli* and Salmonella typhimurium. Cellular and Molecular Biology, Vol. 2

(Neidhart, F. C. Ingranham, J. L., Low, K. B., Magasanik, B., Schaechter, M. & Umbarger, H. E., Eds), pp. 1410-38. Washington, D.C.: ASM Press.

Centers for Disease Control and Prevention (CDC) press release dated July 18 2005. Last accessed July 2009. http://www.cdc.gov/od/oc/media/ pressrel/r050718.htm.

Cheah, G.E., Weigand, W.A., Stark, B.C. 1987. Effects of Recombinant Plasmid Size on Cellular Processes in *E.coli*. *Plasmid.* **18**:127-134.

Chandwani, R.M., Pate, M.E., Savery, J., Thornhill, N.F., Horstmann, B., Hoare, M., Titchner-Hooker, N.J. 1997. The Use of Principal Component Analysis as a Database Mining Tool for the Exploratory Diagnosis of Chromatographic Processes. *Bioprocess Engineering.* **17**:229-234.

Chant, E.L. and Summers, D.K. 2007. Indole Signalling Contributes to the Stable Maintenance of *E. coli* Multicopy Plasmids. *Molecular Microbiology.* **63**: 35-43.

Choi, J.H., Lee, S.J., Lee, S.J., Lee, S.Y. 2003. Enhanced Production of Insulin-Like Growth Factor I Fusion Protein in *Escherichia coli* by Coexpression of the Down-Regulated Genes Identified by Transcriptome Profiling. *Applied and Environmental Microbiology*. **69:** 4737-4742.

Christie, G.E. and Cale, S.B. 1995. The Sigma-70 Subunit From *Escherichia coli* C Differs From That of *Escherichia coli* K12. *Gene.* **162**: 161-162.

Ciccolini, L.A.S., Ayazi-Shamlou, P., Titchener-Hooker, N. 2002. A Mass Balance Study to Assess the Extent of Contaminant Removal Achieved in the Operations for the Primary Recovery of Plasmid DNA from *Escherichia coli* Cells. *Biotechnology and Bioengineering*. **77**: 796- 805.

Clewell, D.B. and Helinski, D.R. 1972. Effect of Growth Conditions on the Formation of the Relaxation Complex of Supercoiled ColE1 Deoxyribonuceic Acid and Protein in *Escherichia coli*. Journal *of Bacteriology* **110**: 1135-1146.

Clewell, D.B. 1972. Nature of ColE1, Plasmid Replication in *Escherichia coli*, in the Presence of Chloramphenicol. *Journal of Bacteriology* **110**: 667-676.

Clowes & Hayes. Experiments in Microbial Genetics. 1968. *Blackwell Scientfic Publications, Oxford.*

Cooke, G.D. Cranenburg, R.M., Hanak, J.A.J., Dunnill, P., Thatcher, D.R., Ward, J.M. 2001. Purification of Essentially RNA Free Plasmid DNA Using a Modified *Escherichia coli* Host Strain Expressing Ribonuclease A. *Journal of Biotechnology* **85**: 297- 304.

Cooke, J.R., McKie, E.A., Ward, J.M., Keshavarz-Moore, E. 2004. Impact of Intrinsic DNA Structure on Processing of Plasmids For Gene Therapy and DNA Vaccines. *Journal of Biotechnology.* **114**: 239-254.

Cortassa, S. and Aon, M.A. 1993. Altered Topoisomerase Activities May Be Involved in the Regulation of DNA Supercoiling In Aerobic-Anaerobic Transitions in *Escherichia coli*. *Molecular and Cellular Biochemistry*. **126**: 115-124.

Conter, A. 2003. Plasmid DNA Supercoiling and Survival in Long-Term Cultures of *Escherichia coli*: Role of NaCl. *Journal of Bacteriology*. **185**: 5324-5327.

Cranenburgh, R.M., Hanak, J.A.J., Williams, S.G., Sherratt, D.J. 2001. *E. coli* Strains That Allow Antibiotic-Free Plasmid Selection and Maintenance by Repressor Titration. *Nucleic Acids Research* **29**: No. 5 e26.

Cranenburgh, R.M., Lewis, K.S., Hanak, J.A.J. 2004. Effect of Plasmid Copy Number and LacOperater Sequence on Antibiotic-Free Plasmid Selection by Operator-Repressor Titration in *E. coli*. *Journal Molecular Microbiology and Biotechnology* **7**: 197-203.

Crosa, J. H., Luttropp, L.K., Falkow, S. 1975. Nature of R-Factor Replication in the Presence of Chloramphenicol. *PNAS USA* **72**: 654-658.

Crick, F.H.C. 1976. Linking Numbers and Nucleosomes. *Proceedings of the National Academy of Science USA.* **73**: 2639-2643.

Curling, J., and Smith, L. 2003. Chromatography for Plasmid DNA Manufacture. *Genetic Engineering News* 23, 11, 42.

Dale, J.W. Molecular Genetics of Bacteria (Third Edition). John Wiley & Sons. 1998.

Darzynkiewicz, Z. 1990. Differential Staining of DNA and RNA in Intact Cells and Isolated Cell Nuclei with Acridine Orange. *Methods in Cell Biology* **33**:285-98.

Davis, B.S., Chang, G., Cropp, B., Roehrig, J.T., Martin, D.A., Mitchell, C.J., Bowen, R., Bunning, M. 2001. West Nile Virus Recombinant DNA Vaccine Protects Mouse and Horse from Virus Challenge and Expresses *In Vitro* a Noninfectious Recombinant Antigen That Can Be Used in Enzyme-Linked Immunosorbent Assays. *Journal of Virology* **75**: 4040-4047.

Derouazi, M., Girard, P., Van Tilborgh, F., Iglesias, K., Muller, N., Bertschinger, M., Wurm, F. 2004. Serum-Free Large-Scale Transient Transfection of CHO Cells. Biotechnology and Bioengineering. **87**(4):537-45.

de Vries, J. and Wackernagel, W. 2002. Integration of Foreign DNA During Natural Transformation of Acinetobacter sp. by Homology-Facilitated Illegitimate Recombination. *PNAS* **99**: 2094-2099.

Di Martino, P., Fursy, R., Bret, L., Sundararaju, B., Phillips, R.S. 2003. Indole Can Act As An Extracellular Signal To Regulate Biofilm Formation of *Escherichia coli* and Other Indole-Producing Bacteria. *Canadian Journal of Microbiology.* **49**: 443-449.

Doelle, H. W., Hollywood, N. W., Westwood, A. W. 1974. Effect of Glucose Concentration on a Number of Enzymes Involved in the Aerobic and Anaerobic Utilization of Glucose in Turbidostat-Cultures of *Escherichia coli*. *Microbios* **9**: 221-232.

Dolezel, J. 1991. Flow cytomeric analysis of nuclear DNA content in higher plants. *Phytochem. Anal.* **2**:143-154.

Dorman, C. J. 1991. DNA Supercoiling and Environmental Regulation of Gene Expression in Pathogenic Bacteria. *Infection and Immunity* **59**: 745-749.

Dorman, C.J., Barr, G.C., Bhriain, N.N., Higgins, C.F. 1988. DNA Supercoiling and the Anaerobic and Growth Phase Regulation of *tonB* Gene Expression. *Journal of Bacteriology*. **170**: 2816-2826.

Drlica, K. 1992. Control of Bacterial DNA Supercoiling. *Molecular Microbiology.* **6**: 425-433.

Durland, R.H. and Eastman, E.M. 1998. Manufacturing and Quality Control of Plasmid-Based Gene Expression Systems. *Advanced Drug Delivery Reviews.* **30**:33-48.

Edwards-Parton, S., Thornhill, N.F., Bracewell, D.G., Liddell, J.M. and Titchener-Hooker, N.J. 2008. Principal Component Score Modelling for the

Rapid Description of Chromatographic Separations. *Biotechnology Progress* **24**: 202-208.

Edwards-Parton, S., Titchner-Hooker, N., Thornhill, N., Bracewell, D., Lidell, J. 2005. Use of Multivariate Statistical Analysis in the Modelling of Chromatographic Processes. IChemE 7th World Congress of Chemical Engineering, Glasgow July 10-14, 2005.

Ellis, R.W. 1999. New technologies for making vaccines. *Vaccine* **17**: 1596-1604.

FDA Points to Consider on Plasmid DNA Vaccines for Preventive Infectious Disease Indications. **1996a.**

FDA Addendum to the Points to Consider in Human Somatic Cell and Gene Therapy. 1991. *Human Gene Therapy.* **1996b.**

Felgner, J.H., Kumar, R., Sridhar, C.N., Wheeler, C.J., Tsai, Y.J., Border, R., Ramsey, P., Martin, M. 1994. Enhanced Gene Delivery and Mechanism Studies with a Novel Series of Cationic Lipid Formulations. *Journal of Biological Chemistry*. **269**: 2550-2661.

Ferreira, G.N.M, Monteiro, G.A. Prazeres, D.M.F., Cabral, J.M.S. 2000. Downstream Processing of Plasmid DNA for Gene Therapy and DNA Vaccine Applications. *TIBTECH*. September. **18**: 380-388.

Finkelstein, M. and Helmstetter, C.E. 1977. Cell cycle analysis of F'*lac* replication in *Escherichia coli* B/r. *Journal of Bacteriology* **132**: 884-895.

Fitzwater, T., Zhang, X-Y., Elble, R. and Polisky, B. 1988. Conditional High Copy Number ColE1 Mutants: Resistance to RNA1 Inhibition *in vivo* and *in vitro*. *The EMBO Journal* **7**: 3289-3297.

Foster, T.J. 1983. Plasmid-Determined Resistance to Antimicrobial Drugs and Toxic Metal Ions in Bacteria. *Microbiological Reviews* **47**: 361-409.

Francisco, J.A., Campbell, R., Iverson, B.L., Georgious, G. 1993. Production and Fluorescence-Activated Cell Sorting of *Escherichia coli* Expressing A Functional Antibody Fragment on the External Surface. *Proceedings of the National Academy of Science USA.* **90**: 10444-10448.

Freitas, S.S., Azzoni, A.R., Santos, J.A.L., Monteiro, G.A., Prazeres, D.M.F. 2007. On the stability of Plasmid DNA Vectors During Cell Culture and Purification. *Molecular Biotechnology.* **36**: 151–158.

Frenkel, L. and Bremer, H. 1986. Increased Amplification of Plasmids pBR322 and pBR327 by Low Concentrations of Chloramphenicol. DNA 5: 539-544.

Fu, A.Y., Spence, C., Scherer, A., Arnold, F.H., Quake, S.R. 1999. A Microfabricated Fluorescence-Activated Cell Sorter. *Nature Biotechnology*. **17**: 1109-1111.

Gale, E. F., E. Cundliffe, P. E. Reynolds, M. H. Richmond, and M. J. Waring. 1981. The Molecular Basis of Antibiotic Action. Wiley Interscience, London.

Garen, A. and Puck, T.T. 1951. The First Two Steps of the Invasion of Host Cells by Bacterial Viruses. II. *Journal of Experimental Medicine* **94**: 177-189.

Garrett, R., S. Douthwaite, A. Liljas, A. Matheson, P. Moore, and H. Noller. 2000. The Ribosome: Structure, Function, Antibiotics, and Cellular Interactions. ASM Press, Washington, D.C.

Glass, R.E. 1982. Gene Function: *E. coli* and its Heritable Elements. Illustrated Edition. University of California Press.

Goldstein, E. and Drlica, K. 1984. Regulation of Bacterial DNA Supercoiling: Plasmid Linking Numbers Vary With Growth Temperature. *Proceedings of the National Academy of Science USA.* **81**: 4046-4050.

Gould, H. Chromatin. The Practical Approach (ed. Rickwood, D., Hames, B.D.) Oxford University Press. 1998.

Gottschalk, U. and Chan, S. 1998. Somatic Gene Therapy Present Situation and Future Perspective. *Arzneimitel-Forschung/Drug Research.* **49**:1111-1120.

Grant, S.G.N., Jessee, J., Bloom, F.R., Hanahan, D. 1990. Differential Plasmid Rescue From Transgenic Mouse DNAs Into *E. coli* Methylation-Restriction Mutants. *Proceedings of the National Academy of Science USA.* **87**: 4645-4649

Greated, A., Lambertsen, L., Willimans, P.A., Thomas, C.T. 2002. Complete sequence of the IncP-9 TOL plasmid pWW0 from *Pseudomonas putida. Environmental Microbiology*. **4**: 856-871.

Green, J.J., Rao, V.B. Recombinant DNA Principles and Methodologies. CRC Press. 1998.

Grossman, A.D., Zhou, Y-N., Gross, C., Heilig, J., Christie, G.E., Calendar, R. 1985. Mutations in the *rpoh (htpR)* Gene of *Escherichia coli* K12 Phenotypically Suppress a Temperature-Sensitive Mutant Defective in the σ^{70} Sununit of RNA Polymerase. *Journal of Bacteriology.* **161**: 939-943.

Guyer, M.S., Reed, R.R., Steitz, J.A., Low, K.B. 1981. Identification of a Sex-Factor-Affinity Site in *E. coli* as Gamma Delta. *Cold Spring Harbor Symposia on Quantitative Biology.* **45**: 135-140.

Han, M., Jeong, K.J., Yoo, J.S., Lee, S.Y. 2003. Engineering *Escherichia coli* for Increased Productivity of Serine-Rich Proteins Based on Proteome Profiling. (2003) *Applied Environmental Microbiology.* **69:** 5772-5781.

Hanahan, D. 1983. Studies on transformation of Escherichia coli with plasmids. *Journal of Molecular Biology.* **166(4)**:577-80.

Hanahan, D. DNA Cloning: A Practical Approach (Glover, D.M., ed.) 1985. *IRL Press, McLean, Virginia.* **1**:109.

Hayashi, K., Morooka, N., Yamamoto, Y., Fujita, K., Isono, K., Choi, S., Ohtsubo, E., Baba, T., Wanner, B.L., Mori, H., Horiuchi, T. 2006. Highly Accurate Genome Sequence of *Escherichia coli* K-12 Strains MG1655 and W3110. *Molecular Systems Biology* **2**: article number 2006.0007

Hazelbauer, G.L. 1975. Role of the Receptor for Bacteriophage Lambda in the Functioning of the Maltose Chemoreceptor of *Escherichia coli*. *Journal of Bacteriology* **124**: 119-126.

Heller, K. and Braun, V. 1979. Accelerated Adsorption of Bacteriophage T5 to *Escherichia coli* F, Resulting from Reversible Tail Fiber-Lipopolysaccharide Binding. *Journal of Bacteriology* **139**: 32-38.

Hellmuth, K., Korz, D.J., Sanders, E.A., Deckwer, W. D. 1994. Effect of Growth Rate on Stability and Gene Expression of Recombinant Plasmids During Continuous and High Cell Density Cultivation of *Escherichia coli* TG1. *Journal of Biotechnology* **32**: 289-298.

Hirakawa, H., Inazumi, Y., Masaki, T., Hirata, T., Yamaguchi, A. 2005. Indole Induces the Expression of Multidrug Exporter Genes in *Escherichia coli*. *Molecular Microbiology.* **55**: 1113-1126.

Hodgson, C.P. and Williams, J.A. Improved Strains of *E. coli* for Plasmid DNA Production. *WIPO International Patent Application, 9 March 2006.*

Horn, N.A., Meek, J.A., Budahazi, M., Marquet, M. 1995. Cancer Gene Therapy Using Plasmid DNA: Purification of DNA for Human Clinical Trials. *Human Gene Therapy.* **6**: 565-573.

Hsieh, L.S., Burger, R.M., Drlica, K. 1991. Bacterial DNA Supercoiling and [ATP]/[ADP]. Changes Associated with a Transition to Anaerobic Growth. *Journal of Molecular Biology*. **219**: 443-450. (a)

Hsieh, L.S., Rouviere-Yaniv, J., Drlica, K. 1991. Bacterial DNA Supercoiling and [ATP]/[ADP]. Changes Associated with Salt Shock. *Journal of Bacteriology*. **173**: 3914-3917. (b)

Humes, L.E. 1984. Noise-Induced Hearing Loss As Influenced By Other Agents And By Some Physical Characteristics Of The Individual. *Journal of Acoustical Society of America.* **76**: 1318-1329

Ingrahm, J.L., Maaloe, O., Neidhardt, F.C. Growth of the Bacterial Cell. Sunderland, Mass. Sinauer Associates, 1983.

Jackson, G.A., Shrestha, S., Ward, J.M. 1995. Metabolic Engineering of the Toluene Degradation Pathway. *Journal Cell Biochemistry*. S21A C1A 48.

Jaffe, A., Ogura, T., Hiraga, S. 1985. Effects of CCD Function of the F Plasmid on Bacterial Growth. *Journal of Bacteriology.* **163**: 841-849.

Jensen, K.F. 1993. The *Escherichia coli* K-12 "Wild Types" W3110 and MG1655 Have an *rph* Frameshift Mutation That Leads to Pyrimidine Starvation Due to Low *pyrE* Expression Levels. *Journal of Bacteriology.* **175**: 3401-3407.

Jensen, P.R., Loman, L., Petra, B., van der Weijden, C., Westerhoff, H.V. 1995. Energy Buffering of DNA Structure Fails when *Escherichia coli* Runs Out of Substrate. *Journal of Bacteriology*. **177**: 3420-3426.

Junker, B. 2007. Foam and Its Mitigation in Fermentation Systems. *Biotechnology Progress* **23**: 767-784.

Kawamura-Sato, K., Shibayama, K., Horii, T., Iimuma, Y., Arakawa, Y., Ohta, M. 1999. Role of Multiple Efflux Pumps in *Escherichia coli* in Indole Expulsion. *FEMS Microbiology Letters* **179**: 345-352.

Kay, A., O'Kennedy, R., Ward, J., Keshavarz-Moore, E. 2003. Impact of Plasmid Size on Cellular Oxygen Demand in *E. coli. Biotechnology and Applied Biochemistry.* **38**:1-7.

Kepner, R.L, Pratt, J.R. 1994. Use of Fluorochromes for Direct Enumeration of Total Bacteria in Environmental Samples: Past and *Present*. Microbiological Reviews **58**: 603-615.

Kichler, A., Leborgne, C., Coeytaux, E. and Danos, O. 2001. Polyethylenimine -mediated gene delivery: a mechanistic study. *Journal of Gene Medicine.* **3**:135-144.

Kim, J.Y. and Ryu, D.D.Y. 1991. The Effects of Plasmid Content, Transcription Efficiency, and Translation Efficiency on the Productivity of a Cloned Gene Protein in *Escherichia coli*. *Biotechnology and Bioengineering.* **38**:1271-1279.

Kleman, G.L. and Strohl, W.R. 1994. Acetate Metabolism by *Escherichia coli* in High Cell Density Fermentation. *Applied Environmental Microbiology.* **60**: 3952-3958.

Kong, S., Rock, C.F., Booth, A., Willoughby, N., O'Kennedy, R. Relton, J., Ward, J.M., Hoare, M., Susana Levy, M.S. 2008. Large-scale Plasmid

DNA Processing: Evidence That Cell Harvesting and Storage Methods Affect Yield of Supercoiled Plasmid DNA. *Biotechnology Applied Biochemistry* **51**: 43-51.

Kong, S., Titchener-Hooker, N., Levy, M.S. 2006. Plasmid DNA Processing for Gene Therapy and Vaccination: Studies on the Membrane Sterilisation Filtration Step. *Journal of Membrane Science*. **280**: 824-831.

Kubitschek, H.E., Friske, J.A. 1986. Determination of Bacterial Cell Volume with the Coulter Counter. *Journal of Bacteriology.* **168**: 1466-1467.

Lahijani, R., Hulley, G., Soriano, G., Horn, N.A., Marquet, M. 1996. High-Yield Production of pBR322-Derived Plasmids Intended for Human Gene Therapy by Employing a Temperature-Controllable Point Mutation. *Human Gene Therapy.* **7**: 1971-1980.

Lamb, B.C. 1975. Cryptic mutations: Their Predicted Biochemical Basis, Frequencies and Effects on Gene Conversion. *Molecular and General Genetics MGG* **137**: 305-314.

Lark, K.G., 1972. Evidence for the Direct Involvement of RNA in the Initiation of DNA Replication in *Escherichia coli* 15T$^-$. *Journal of Molecular Biology* **64**: 47-60.

Lee, K.K., Jang, C.S., Yoon, J.Y., Kim, S.Y., Kim, T.H., Ryu, K.H., Kim, W. 2008. Abnormal Cell Division Caused by Inclusion Bodies in *E. coli* Increases Resistance Against External Stress. *Microbiological Research.* **163**: 394-402.

Lee, S.Y. 1996. High Cell-Density Culture of Escherichia coli. *Tibtech* **14**: 98-105.

Lee, S.Y., Yim, K.S., Chang, H.N., Chang, Y.K. 1994. Construction of plasmids, Estimation of Plasmid Stability, and Use of Stable Plasmids for the Production of Poly(3-hydroxybutyric acid) by Recombinant *Escherichia coli. Journal of Biotechnology.* **32**: 203-211.

Levy, M.S., Collins, I.J., Ward J.M., Titchener-Hooker, N., Ayazi Shamlou, P., Dunnill, P. 1999. Effect of Shear on Plasmid DNA in Solution. *Bioprocess Engineering.* **20**: 7-13.

Levy, M.S., Loftian, P., O'Kennedy, R., Lo-Yim, M., Shamlou, P.A. 2000. Quantitation of Supercoiled Circular Content in Plasmid DNA Solutions Using a Fluorescence-Based Method. *Nucleic Acids Research.* **28**: 257.

Levy. M.S., O'Kennedy, R., Ayazi-Shamlou, P., Dunnill, P. 2000. Biochemical Engineering Approaches to the Challenges of Producing Pure Plasmid DNA. *TIBTECH.* July. **18**: 296-305.

Li, S. and Huang, L. 2000. Nonviral Gene Therapy: Promises and Challenges. *Gene Therapy*. **7**: 31-34.

Lin-Chao, S., Chen, W-T, Wong, T-T. 1992. High copy number of the pUC plasmid results from a Rom/Rop-supressible point mutation in RNA II. Molecular Microbiology. **6**:3385-3393.

Listner, K., Bentley, L, Okonkowski, J., Kistler, C., Wnek, R., Caparoni, A., Junker, B., Robinson, D., Salmon, P., Chartrain, M. 2006. Development of a Highly Productive and Scalable Plasmid DNA Production Platform. *Biotechnology Progress*. **22**: 1335-1345.

Lodish, H., Berk, A., Zipursky, S.A., Matsudaira, P., Baltimore, D. and Darnell, J. *Molecular Cell Biology* (4th ed.) New York: W. H. Freeman. 1999.

Lorenz, M.G. and Wackernagel, W. 1994. Bacterial Gene Transfer by Natural Genetic Transformation in the Environment. *Microbiological Reviews* **58**: 563-602.

Losen, M., Frolich,B., Pohl, M., Buchs. 2004. Effect of Oxygen Limitation and Medium Composition on *Escherichia coli* Fermentation in Shake-Flask Cultures. *Biotechnology Progress* **20***:* 1062-1068.

Low, B. 1968. Formation of Merodiploids In Matings With A Class of Rec-Recipient Strains of *Escherichia Coli* K-12. *Proceedings of the National. Academy of Science USA* **60**: 16-167.

Lubelski, J., Konings, W. N., Driessen, A. J.M. 2007. Distribution and Physiology of ABC-Type Transporters Contributing to Multidrug Resistance in Bacteria. *Microbiology and Molecular Biology Reviews* **71**: 463-476.

Luli, G.W. and Strohl, W.R. 1990. Comparison of Growth, Acetate Production, and Acetate Inhibition of *Escherichia coli* Strains in Batch and Fed-Batch Fermentations. *Applied Environmental Microbiology* **56**: 1004-1011.

McGrath, S., and van Sinderen, D. 2007. Bacteriophage: Genetics and Molecular Biology. Caister Academic Press, First edition.

Macphee, D.G., Radford, A.J., Reanney, D.C. 1988. Preparation of Plasmid DNA and Products Thereof. US Patent # 4755464.

Malmquist, G. and Daneilsson, R. 1994. Alignment of Chromatographic Profiles for Principal Component Analysis: a Prerequisite for Finger Printing Methods. *Journal of Chromatography A* **687**: 71-88.

March, J.B., Clark, J.R., Jepson, C.D. 2004. Genetic Immunisation Against Hepatitis B Using Whole Bacteriophage Lambda Particles. *Vaccine* **22**: 1666-1671.

March, J.B., Jepson, C.D., Clark, J.R., Totsika, M., Calcutt, M.J. 2006. Phage Library Screening for the Rapid Identification and *In Vivo* Testing of Candidate Genes for a DNA Vaccine against *Mycoplasma mycoides* subsp. *mycoides* Small Colony Biotype. *Infection and Immunity* **74**: 167-174.

Marquet, M. Horn, N.A., Meek, J. 1995. Process Development for the Manufacture of Plasmid DNA Vectors for Use in Gene Therapy. *BioPharm.* September, 26-37.

Marie, D., Vaulot, D., Partensky, F., 1996. Application of the Novel Nucleic Acid Dyes YOYO-1, YO-PRO-1 and PicoGreen for Flow Cytomeric Analysis of Marine Prokaryotes. *Applied Environmental Microbiology* **62**: 1649-1655.

Moineauand, S. and Lévesque, C. 2004. Bacteriophages, Biology and Applications**.** Edited by Elizabeth Kutter and Alexander Sulakvelidze**.** CRC Press, Pp. 285-296.

Molin, S., Light, J.A., Larsen, J.E.L. 1989. Plasmids with Conditional Uncontrolled Replication Behaviour. US Patent #4806471.

Monteiro, A.C.M., Ferreira, G.P., Ferreira L.C.S., Costa, S.O.P. 2003. Environmental and genetic factors affecting mutability to aminoglycoside antibiotics among *E. coli* K12 strains. *Genetics and Molecular Biology* **26**: 2, 221-227.

Morgan, A.R., Lee, J.S., Pulleyblank, D.E., Murray, N.L., Evans, D.H. 1979. Ethidium Fluorescence Assays. Part 1. Physicochemical Studies. *Nucleic Acids Res.* **7**:547-569.

Mosher, R.H. 2002. Using pGlo to Demonstrate the Effects of Catabolite Repression on Gene Expression in Escherichia coli. *Bioscene*, **21**: 17-23.

Mukherjee, K.J., Rowe, D.C.D., Watkins, N.A., Summers, D.K. 2004. Studies of Single-Chain Antibody Expression in Quiescent *Escherichia coli*. *Applied Environmental Microbiology*. **70:** 3005-3012.

Myc, A., Pizzolo, J.G., Dygulski, K., Myron, R. 1992. Increase in Acridine Orange (AO) Fluorescence Intensity of Monocytes Cultured in Plastic Tissue Culture Plates as Measured by Flow Cytometry. *Cytometry* **13**:103-107.

Nail, S.L., Akers, M.J., 2002. Development and manufacture of protein pharmaceuticals. Kluwer Academic/Plenum Publishers. New York.

Neidhardt, F.C., Ingraham, J.L., Schaechter, M., 1990. Biosynthesis and Fueling, In Physiology of the Bacterial Cell. A Molecular Approach. Massachusetts: Sinauer Associates, USA.

Noites, I.S., O'Kennedy, R.D., Levy, M.S., Abidi, N., Keshavarz-Moore, E. 1999. Rapid Quantification and Monitoring of Plasmid DNA Using an Ultrasensitive DNA-Binding Dye. *Biotechnology and Bioengineering.* **66**: 195-201.

Nordstrom, K., Molin, S., Light, J. 1984. Control of Replication of Bacterial Plasmids: Genetics, Molecular Biology, and Physiology of the Plasmid Rl System. *Plasmid* **12**: 71-90.

O'Kennedy, R.D., Baldwin, C., Keshavarz-Moore, E. 2000. Effects of Growth Medium Selection on Plasmid DNA Production and Initial Processing Steps. *Journal of Biotechnology.* **76**: 175-183.

O'Kennedy, R.D., Ward, J.M., Keshavarz-Moore, E. 2003. Effects of Fermentation Strategy on the Characteristics of Plasmid DNA Production. *Biotechnology and Applied Biochemistry.* **37**: 83-90.

O'Mahoney, K., Freitag, R., Hilbrig, F., Müller, P., Schumacher, I. 2007. Strategies for High Titre Plasmid DNA Production in *Escherichia coli* DH5α. *Process Biochemistry.* **42**: 1039-1049.

Ow, D.S., Yap, M.G., Og, S.K. 2009. Enhancement of Plasmid DNA Yields During Fed-Batch Culture of a fruR-knockout Escherichia coli Strain. Biotechnology and Applied Biochemistry 52: 53-59.

Phue, J. and Shiloach, J. 2004. Transcription Levels of Key Metabolic Genes Are The Cause For Different Glucose Utilization Pathways in *E. coli* B (BL21 DE3) and *E. coli* K (JM109). *Journal of Biotechnology.* **109**: 21-30

Posfai, G., Plunkett, G. III., Feher, T., Frisch, D., Keil, G.M., Umenhoffer, K., Kolisnychenko, V., Stahl, B., Sharma, S.S., de Arruda, M., Burland, V., Harcum, S.W., Blattner, F.R. 2006. Emergent Properties of Reduced-Genome *Escherichia coli*. *Science.* **312**: 1044-1046.

Prather, K.J., Sagar, S., Murphy, J., Chartrain, M. 2003. Industrial Scale Production of Plasmid DNA for Vaccine and Gene Therapy: Plasmid Design, Production and Purification. *Enzyme and Microbial Technology.* **33**: 865-883.

Prazeres, D.M.F. and Ferreira, G.N.M. 2004. Design of Flowsheets for the Recovery and Purification of Plasmids for Gene Therapy and DNA Vaccination. *Chemical Engineering and Processing* **43**: 615–630.

Prazeres, DMF., Ferreira G.N.M., Monteiro, G.A., Cooney, C.L., Cabral, J.M.S. 1999. Large-Scale Production of Pharmaceutical-Grade Plasmid DNA for Gene Therapy: Problems and Bottlenecks. *Trends in Biotechnology*. **17**: 169-174

Projan, S.J., Carleton, S., Novick, R.P. 1983. Determination of Plasmid Copy Number by Fluorescence Densitometry. *Plasmid* **9**:182-190.

Pruss, G.J. 1985. DNA Topoisomerase I Mutants Increased Heterogeneity in Linking Number and other Replicon-dependent Changes in DNA Supercoiling. *Journal of Molecular Biology* **185**: 51-43.

Pruss, G.J., Manes, G.H., Drlica, K. *Escherichia coli* DNA Topoisomerase I Mutants: Increased Supercoiling is Corrected by Mutations Near Gyrase Genes. *Cell.* 1982. **31**: 35-42.

Pulleyblank, D.E., Shure, M., Tang, D., Vinograd, J., Vosberg, H.P. 1975. Action of Nicking-Closing Enzyme on Supercoiled and Nonsupercoiled Closed Circular DNA: Formation of a Boltzmann Distribution of Topological Isomers. *PNAS USA* **72**: 4280-4284.

Qiagen. Plasmid Purification Handbook, 2nd Edition. December 2006.

Raleigh, E.A., Lech, K., Brent, R. Current Protocols in Molecualr Biology (ed. Ausubel, F.M, Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K.) John Wiley, New York. 1989.

Ramirez, R.M., Villarejo, M. 1991. Osmotic signal transduction to proU is independent of DNA supercoiling in *Escherichia coli*. *Journal of Bacteriology.* **173**:879-885.

Randall-Hazelbauer, L. and Schwartz, M. 1973. Isolation of the Bacteriophage Lambda Receptor from *Escherichia coli*. *Journal of Bacteriology* **116**: 1436-1446.

Raychaudhuri, S., Stuart, J.M., Altman, R.B. 2000. Principal Components Analysis to Summarize Microarray Experiments: Application to Sporulation Time Series. *Pacific Symposium on Biocomputing* **5**:452-463.

Reinikainen, P., Korpela, K., Nissinen, V., Olkku, J., Soderlund, H. Markkanen, P. 1988. *E. coli* Plasmid Production in Fermenter. *Biotechnology and. Bioengineering.* **33**:386-393.

Remaut, E., Tsao, H., Fiers, W. 1983. Improved Plasmid Vectors with a Thermoinducible Expression and Temperature-Regulated Runaway Replication. *Gene.* **22**:103-113.

Reyes-Dominguez, Y., Contreras-Ferrat, G., Ramirez-Santos, J., Membrillo-Hernandez, J., Gomez-Eichelmann, C. Plasmid DNA Supercoiling and Gyrase Activity in *Escherichia coli* Wild-Type and *rpoS* Stationary-Phase Cells. *Journal of Bacteriology.* 2002. **185**: 1097-1100.

Riley, M., Abe, T., Arnaud, M.B., Berlyn, M.B., Blattner, F.R., Chaudhuri, R.R., Glasner, J.D., Mori, H., Horiuchi, T., Keseler, I.M., Kosuge, T., Perna, N.T., Plunkett III, G., Rudd, K.E., Serres, M.H, Thomas, G.H., Thomson, N.R., Wishart, D.S., Wanner, B.L. 2006. *Escherichia coli* K-12: a Cooperatively Developed Annotation Snapshot-2005. *Nucleic Acids Research*. **34**: 1-9.

Rinas, U., Kracke-Helm, H, Schugerl, K. 1989. Glucose as a Substrate in Recombinant Strain Fermentation Technology. *Applied Microbiology and Biotechnology*. **31**:163-167.

Ringer, M. 2008. What is Principal Component Analysis? *Nature Biotechnology.* **26**: 3.

Ritter T., Lehmann M., Volk H.D. 2002. Improvements in gene therapy: averting the immune response to adenoviral vectors. *BioDrugs.* **16**: 3-10.

Robbins, P.D. and Ghivizzani, S.C. 1998. Viral Vectors for Gene Therapy. *Pharmacology and Therapeutics.* **80**: 35-47.

Rock, C., Shamlou, P.A., Levy, M.S. 2003. An Automated Microplate-Based Method For Monitoring DNA Strand Breaks In Plasmids and Bacterial Artificial Chromosomes. *Nucleic Acids Research.* **31**: e65.

Rols, M.-P., Delteil, C., Golzio, M., Dumond, P., Cros, S., Teissie, J. 1999. *In vivo* electrically mediated protein and Gene Transfer in Murine Melanoma. *Nature Biotechnology*. **16**: 168-171.

Rowe, D.C.D. and Summers, D.K. 1999. The Quiescent-Cell Expression System for Protein Synthesis in Escherichia coli. *Applied Environmental Microbiology.* **65:** 2710-2715.

Ryan, W. and Parulekar, S.J. 1991. Recombinant Protein Synthesis and Plasmid Instability in Continuous Cultures of *Escherichia coli* JM103 Harbouring a High Copy Number Plasmid. *Biotechnology and Bioengineering.* **37**: 415-429.

Sambrook, J., Fritsch, E.F., Maniatis, T. Molecular Cloning- A Laboratory Manual. *Cold Spring Harbor Laboratory Press. 2nd Edition.* 1989. New York.

Sambrook, J., Russell, D.W. Molecular Cloning- A Laboratory Manual. Volume 1. *Cold Spring Harbor Laboratory Press.* 2001. New York.

Schleef, M. and Schmidt, T. 2004. Animal-free Production of CCC-Supercoiled Plasmids for Research and Clinical Applications. *Journal of Gene Medicine*. **6**: S45–S53.

Schmidt, T., Friehs, K., Schleef, M., Voss, C., Flaschel, E. 1999. Quantitative Analysis of Plasmid Forms by Agarose and Capillary Gel Electrophoresis. *Analytical Biochemistry.* **274**:235-240.

Schoenfeld, T., Mendez, J., Storts, D.R., Portman, E., Patterson, B., Frederiksen, J., Smith, C. 1995. Effects of Bacterial Strains Carrying the *endA1* Genotype on DNA Quality Isolated with Wizard™ Plasmid Purification Systems. *Promega Notes Magazine*. **53**:12.

Seo, J-H. 1986. Investigations of Plasmid-Host Cell Interactions in Recombinant *Escherichia coli* Populations. PhD Thesis, California Institute of Technology.

Seo, J. and Bailey, J.E. 1985. Effects of recombinant plasmid content on growth properties and cloned gene product formation in *Escherichia coli. Biotechnology and Bioengineering* **27**: 1668-1674.

Seo J. and Bailey J. 1986. Continuous cultivation of recombinant *Escherichia coli*: Existence of an Optimum Dilution Rate for Maximum Plasmid and Gene Product Concentration. *Biotechnology and Bioengineering* **28**:1590–1594.

Sheridan, R. Jackson, G.A., Regan, L., Ward, J. Dunnill, P. 1998. Rational Engineering of the TOL *meta*-Cleavage Pathway. *Biotechnology and Bioengineering* **58**: 240-249.

Sherratt, D.J., Williams, S.G., Hanak, J.A.J. *Plasmid Stabilisation*.1999. US Patent 5,972,708.

Shiloach, J., Kaufman, J., Guillard, A. S., Fass, R. 1996. Effect of Glucose Supply Strategy on Acetate Accumulation, Growth and Recombinant Protein Production by *Escherichia coli* BL21 (λDE3) and *Escherichia coli* JM109. *Biotechnology and Bioengineering.* **49**:421-428.

Shojaosadati, S.A., Kolaei, S.M.V., Babaeipour, V., Farnoud, A.M. 2008. Recent advances in high cell density cultivation for production of recombinant protein. *Iranian Journal of Biotechnology* **6**: 63-84.

Singer, V.L., Jones, L.J., Yue, S.T., Haugland, R.P. 1997. Characterization of PicoGreen Regeant and Development of a Fluorescence-Based Solution Assay for Double-Stranded DNA Quantitation. *Analytical Biochemistry*. **249**:228-238.

Smith, H.A. and Klinman, D.M. 2001. The Regulation of DNA Vaccines. *Current Opinion in Biotechnology.* **12**:299-303.

Smith, M.A. and Bidochka, J. 1998. Bacterial Fitness and Plasmid Loss: The Importance of Culture Conditions and Plasmid Size. *Canadian Journal of Microbiology*. **44**: 351-355.

Smith, M.W. and Neidhardt, F.C. 1983. Proteins Induced by Anaerobiosis in *Escherichia coli*. *Journal of Bacteriology* **154**: 336-343.

Spratt. B.G. 1975. Distinct Penicillin-Binding Proteins Involved in the Division, Elongation and Shape of *Escherichia coli* K-12. *Proceedings of the National Academy of Science USA.* **72**: 2999-3003.

Spratt, B.G., Hedge, P.J., te Heesen, S., Edelman, A., Broome-Smith, J.K. 1986. Kanamycin-resistant vectors that are analogues of plasmids pUC8, pUC9, pEMBL8 and pEMBL9. *Gene* **41**:337-342.

Spurio, R., Falconi, M., Brandi, A., L.Pon, C., O.Gualerzi, C. 1997. The Oligomeric Structure of Nucleoid Protein H-NS Is Necessary For Recognition of Intrinsically Curved DNA and For DNA Bending. *EMBO Journal.* **16**:1795-1805.

St. John, A.C., Goldberg, A.L. 1980. Effects of Starvation for Potassium and Other Inorganic Ions on Protein Degradation and Ribonucleic Acid Synthesis in *Escherichia coli*. *Journal of Bacteriology* **143**: 1223-1233.

Steen, H.B. 2000. Flow Cytometry of Bacteria: Glimpses from the Past With A View To The Future. *Journal of Microbiological Methods*. **42**: 65-74.

Steen H.B. 1990. Flow Cytometry Studies of Microorganisms. In Melamed, M.R., Lindmo, T., Mendelsohn, M.L. (Eds.), Flow Cytometry and Sorting. 2nd Edition. Wiley-Liss, New York, 605-622.

Sternglanz, R., DiNardo, S., Voelkel, K.A., Nishimura, Y., Hirota, Y., Becherer, K., Zumstein, L., Wang, J.C. 1981. Mutations in the Gene Coding for *Escherichia coli* DNA Topoisomerase I Affect Transcription and Transposition. *Proceedings of the National Academy of Science USA.* **78**: 2747-2751.

Studier, F.W. and Moffat, B.A. 1986. Use of Bacteriophage T7 RNA Polymerase to Direct Selective High Level Expression of Cloned Genes. *Journal of Molecular Biology.* **189**: 113-130.

Stuve, O., Eagar, T.N., Frohman, E.M., Cravens, P.D. 2007. DNA Plasmid Vaccination for Multiple Sclerosis. *Archives of Neurology.* **64** (10): 1385-1386.

Sugimoto, S., Saruwatari, K., Higashi, C., Sonomoto, K. 2008. The Proper Ratio of GrpE to DnaK is Important for Protein Quality Control by the DnaK-DnaJ-GrpE Chaperone System and for Cell Division. *Microbiology.* **154**: 1876-1885.

Sulavik, M., Houseweart, C., Cramer, C., Jiwani, N., Murgolo, N., Greene, J., DiDomenico, B., Shaw, K.J., Miller, G.H., Hare, R., Shimer, G. 2001. Antibiotic Susceptibility Profiles of *Escherichia coli* strains Lacking Multidrug Efflux Pump Genes. *Antimicrobial Agents and Chemotherapy.* **45**:1126-1136.

Summers, D.K. 1989. Derivatives of ColE1 cer Show Altered Topological Specificity in Site-Specific Recombination. *The EMBO Journal.* **8**: 309-315.

Summers, D. 1998. Timing, Self-Control and A Sense of Direction Are The Secrets Of Multicopy Plasmid Stability. *Molecular Microbiology.* **29**: 1137-1145.

Summers, D.K., and Chant, E., 2009. Chemical Induction in Quiescence in Bacteria. Patent Application, USPTO #20090004700

Summers, D.K. and Sherratt, D.J. 1984. Multimerisation of High Copy Number Plasmids Causes Instability: ColE1 Encodes A Determinant Essential For Plasmid Monomerisation And Stability. *Cell.* **36**: 1097-1103.

Summers, D.K. and Sherratt, D.J. 1988. Resolution of ColE1 Dimers Requires a DNA Sequence Implicated in the Three-Dimensional Organization of the *cer* Site. *The EMBO Journal.* **7**: 851-858.

Tamaki, S., Matsuzawa, H., Matsuhashi, M. 1980. Cluster of *mrdA* and *mrdB* Genes Responsible for the Rod Shape and Mecillinam Sensitivity of *Esherichia coli*. *Journal of Bacteriology.* **141**: 52-57.

Taylor, R.C., Walker, D.C., McInnes, R.R. 1993. *E. coli* Host Strains Significantly Affect the Quality of Small Scale Plasmid DNA Preparations Used For Sequencing. *Nucleic Acids Research***. 21**: 1677-1678.

Templeton, N.S. and Lasic, D.D. 1999. New Directions in Liposome Gene Delivery. *Molecular Biotechnology.* **11**: 175-180.

Theodossiou, I., Collins, I.J., Ward, J.M., Thomas, O.R.T., Dunnill, P. 1988. The Processing of a Plasmid-Based Gene from *E. coli* Promary Recovery by Filtration. *Bioprocess Engineering*. **16**: 175-183.

Togna, A.P., Shuler, M.L., Wilson, D.B. 1993. Effects of Plasmid Copy Number and Runway Plasmid Replication on Overproduction and Excretion of β-Lactamase from *Escherichia coli*. *Biotechnology Progress*. **9**: 31-39.

Trueba, F.J., Woldringh, C.L. 1980. Changes in Cell Diameter During the Division Cycle of *Escherichia coli*. *Journal of Bacteriology* **142**: 869-878.

Uhlin, B.E., Nordström, K. 1978. A Runaway-Replication Mutant of Plasmid R1drd-19: Temperature-Dependant Loss of Copy Number Control. *Molecular Genetics and Genomics.* **164**: 167-179

Uhlin, B.E., Molin, S., Gustafsson, P., Nordström, K. 1979. Plasmids with Temperature-Dependant Copy Numbers for Amplification of Cloned Genes and Their Products. *Gene.* **6**: 91-106.

Urthaler, J., Buchinger, W., Necina, R. 2005. Improved Downstream Process for the Production of Plasmid DNA for Gene Therapy. *Acta Biochemica Polonica*. **52(3)**: 703-711.

van de Walle, M. and Shiloach, J. 1998. Proposed Mechanism of Acetate Accumulation in Two Recombinant *Escherichia coli* Strains During High Density Fermentation. *Biotechnology and Bioengineering.* **57**: 71-78

van Hoek, P., de Hulster, E., van Dijken, J.P., Pronk, J.T. 2000. Fermentative Capacity in High-Cell-Density Fed-batch Cultures of Baker's Yeast. *Biotechnology and Bioengineering* **68**: 517-523.

Varley, D.L., Hitchcock, A.G., Weiss, A.M.E, Horler, W.A., Cowell, R., Peddie, L., Sharpe, G.S., Thatcher, D.R., Hanak, J.A.J. 1999. Production of Plasmid DNA for Human Gene Therapy Using Modified Alkaline Lysis and Expanded Bed Anion Exchange Chromatography. *Bioseparations*. **8**: 209-217.

Vologodskii, A.V., Levene, S.D., Klenin, K.V., Frank-Kamenetskii, M., Cozzarelli, N.R. 1992. Conformational and Thermodynamic Properties of Supercoiled DNA. *Journal Molecular Biology* **227**: 1224-1243.

Voss, C, Schmidt, T., Schleef, M., Friehs, K., Flaschel, E. 2003. Production of Supercoiled Multimeric Plasmid DNA for Biopharmaceutical Application. *Journal of Biotechnology.* **105**: 205-213.

Warnes, A. and Stephenson, J.R. 1986. The Insertion of Large Pieces of Foreign Genetic Material Reduces Stability of Bacterial Plasmids. *Plasmid*. **16**: 116-123.

Wan, N.C. and Goodrick, J.C. Method for Increasing Plasmid Yield. United States Patent #5, 1996, 487,986.

Wang, Z., Xiang, L., Shao, J., Wegrzyn, A., Wegrzyn, G. 2006. Effects of the Presence of ColE1 plasmid DNA in *Escherichia coli* on the Host Cell Metabolism. *Microbial Cell Factories.* **5**: 34.

Weikert, C., Sauer, U., Bailey, J.E. 1997. An *Escherichia coli* Host Strain Useful for Efficient Overproduction of Secreted Protein. *Biotechnology and Bioengineering.* **59**: 386- 391.

Wiley. http://www.wiley.co.uk/genmed/clinical/ Last accessed July 2009.

Williams, D.C., Van Frank, R.M., Muth, W.L., Burnett, J.P. 1982. Cytoplasmic Inclusion Bodies in *Escherichia coli* Producing Biosynthetic Human Insulin Proteins. *American Association for the Advancement of Science.* **215**: 687-689.

Williams, J.A., Luke, J. Langtry, S., Anderson, S., Hodgson, C.P., Carnes, A. 2009. Generic Plasmid DNA Production Platform Incorporating Low Metabolic Burden Seed-Stock and Fed-Batch Fermentation Processes. *Biotechnology and Bioengineering*. **103**: 1129-1142.

Willis, A.E., Perham, R.N., Wraith, D. 1993. Immunological Properties of Foreign Peptides in Multiple Display on a Filamentous Bacteriophage. *Gene* **128**: 79-83.

Wold, S., Esbensen, K. and Geladi, P. 1987. Principal Component Analysis. *Chemometrics and Intelligent Laboratory Systems* **2**: 37-52.

Wold, S. and Sjostrom, M. 1998. Chemometrics, Present and Future Success. *Chemometrics and Intelligent Laboratory Systems.* **44**: 3-14.

Wolff, J.A., Malone, R.W., Williams, P., Chong, W., Acsadi, G., Jani, A., Felgner, P.L. 1990. Direct Gene Transfer into Mouse Muscle *in vivo*. *Science.* **247**: 1465-1468.

Wong, E.M., Muesing, M.A., Polisky, B. 1982. Temperature-Sensitive Copy Number Mutants of ColE1 Are Located in an Untranslated Region of the Plasmid Genome. *Proceedings of the National Academy of Science U.S.A.* **79**: 3570-3574.

Wong, M., Woodley, J.M., Lye, G.J. 2010. Application of Bipolar Electrodialysis to *E. coli* Fermentation for Simultaneous Acetate Removal and pH Control. *Biotechnology Letters* **32**: 1053-1057.

Wrobel, B. and Wegrzyn, G. 1997. Replication and Amplification of λ Plasmids in *Escherichia coli* During Amino Acid Starvation and Limitation. *FEMS Microbiology Letters* **153**: 151-157.

Wrobel, B. and Wegrzyn, G. 1998. Replication Regulation of ColE1-Like Plasmids in Amino Acid-Starved *Escherichia coli*. *Plasmid.* **39**: 48-62.

Xu, Y., Bremer, H. 1997. Winding of the DNA Helix by Divalent Metal Ions. *Nucleic Acids Research.* **25**: 4067-4071.

Yanisch-Perron, C. Vieira, J. and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors. *Gene.* **33**: 103-119.

Yavachev, L. and Ivanov, I. 1988. What Does the Homology Between *E. coli* tRNAs and RNAs Controlling ColE1 Plasmid Replication Mean? *Journal of Theoretical Biology.* **131**: 235-241.

Yeung, K.Y. and Ruzzo, W. L. 2001. Principal Component Analysis for Clustering Gene Expression Data. *Bioinformatics.* **17**: 763-774.

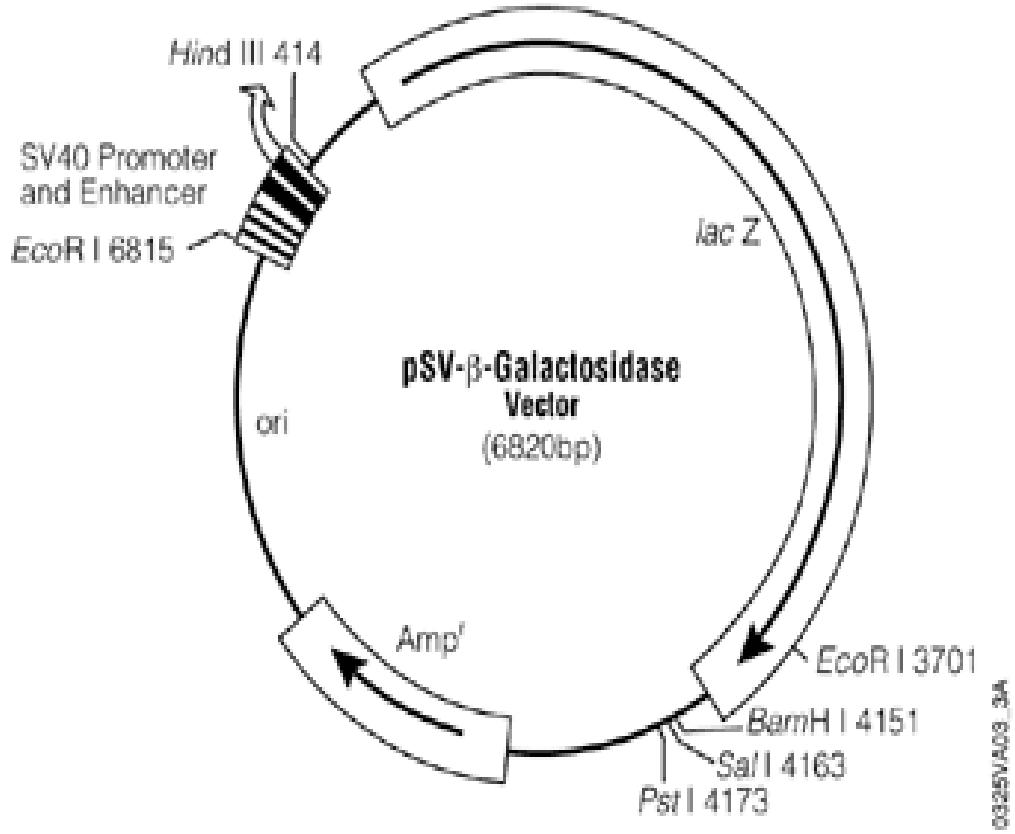
Zhao, J.B., Wei, D.Z., Tong, W.Y. 2007. Identification of *Escherichia coli* Host Cell for High Plasmid Stability and Improved Production of Antihuman Ovarian Carcinoma and Antihuman CD3 Single-Chain Bispecific Antibody. *Journal Applied Microbiology and Biotechnology* **76**: 795-800.

Zhang, J. and Greasham, R. 1999. Chemically Defined Media for Commercial Fermentations. *Applied Microbiology and Biotechnology* **51**: 407-421.

Appendices

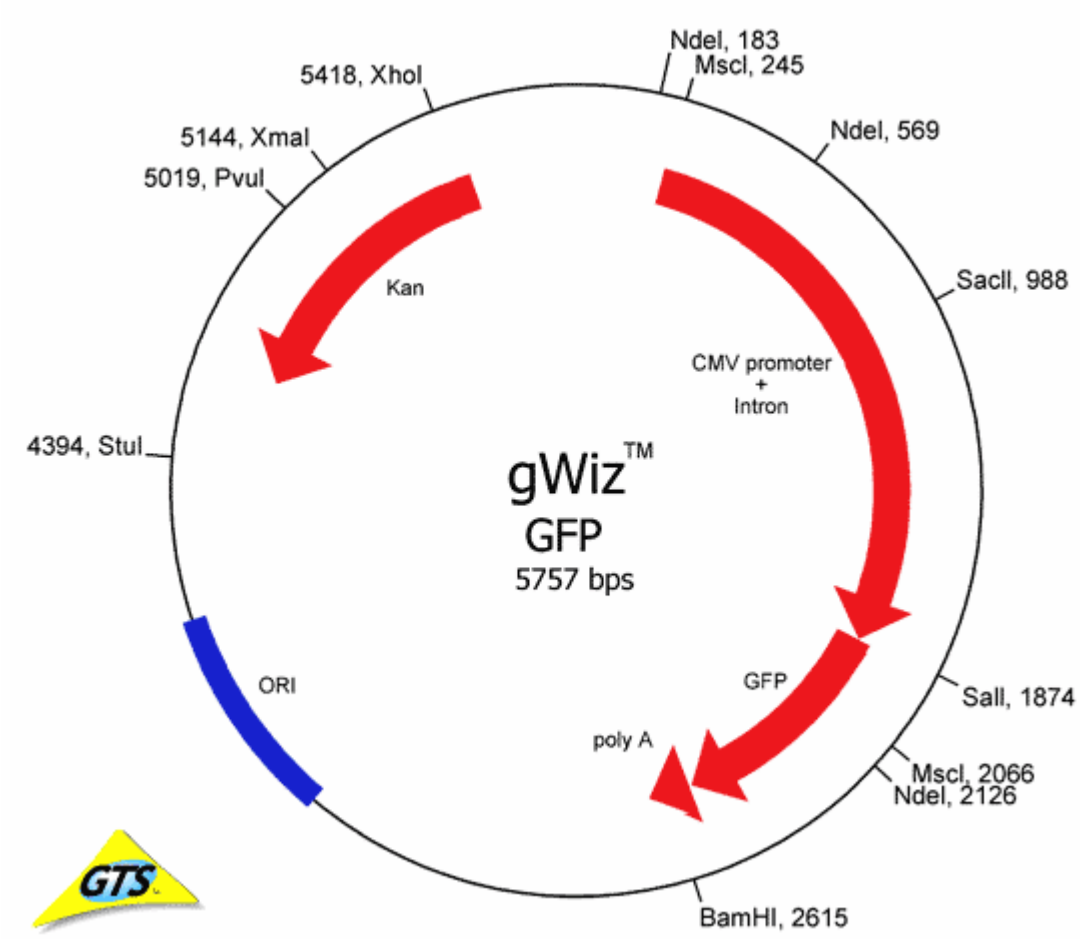
Appendix A.1.

A.1.1. pSVβ Plasmid Map



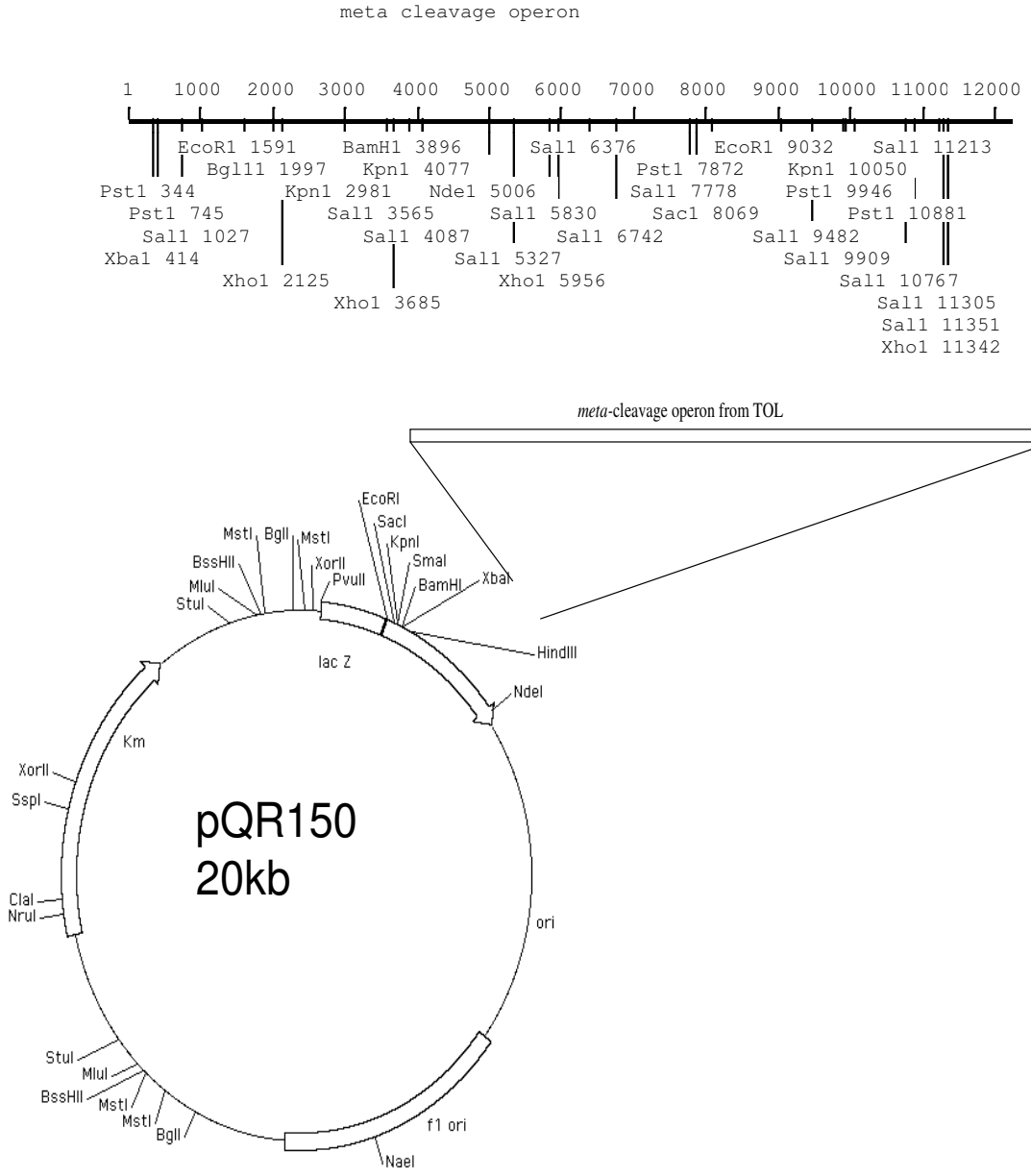
Source: Promega Corp., Madison WI, USA.

A.1.2. gWiz Plasmid Map



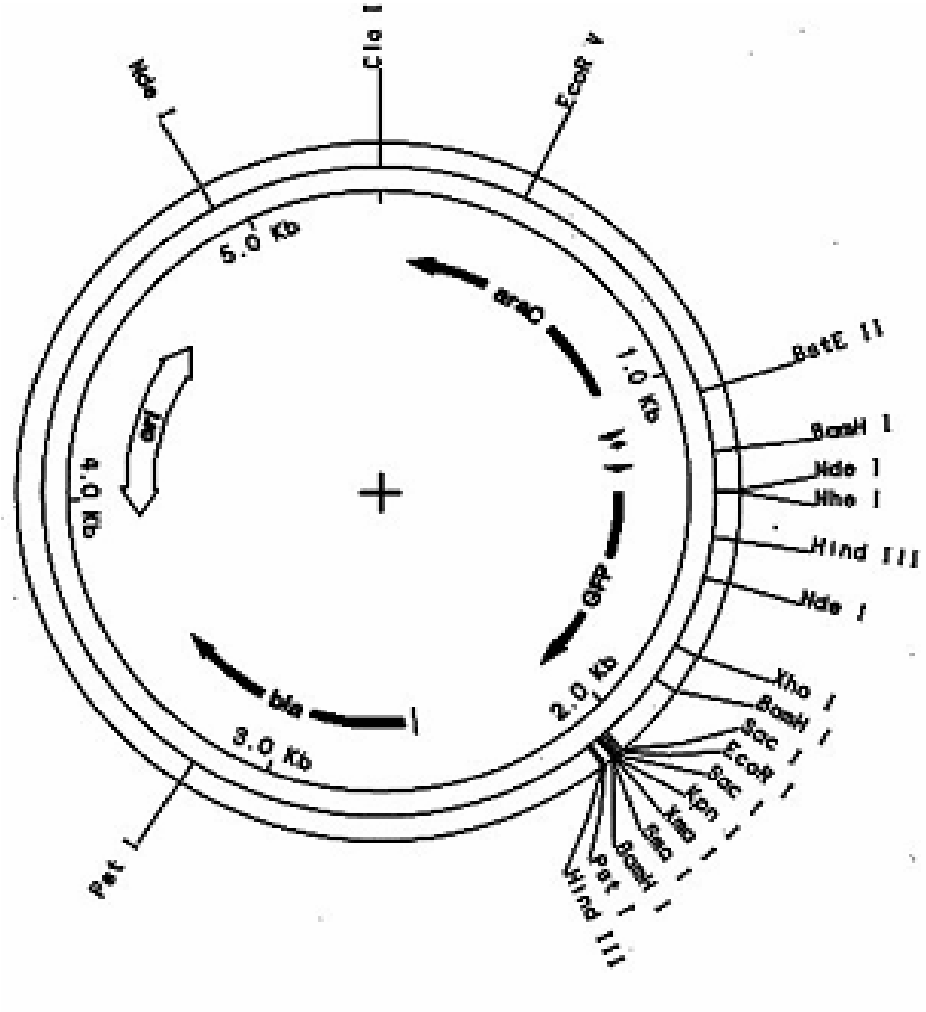
Source: www.Gelantis.com

A.1.3. pQR150 Plasmid Map



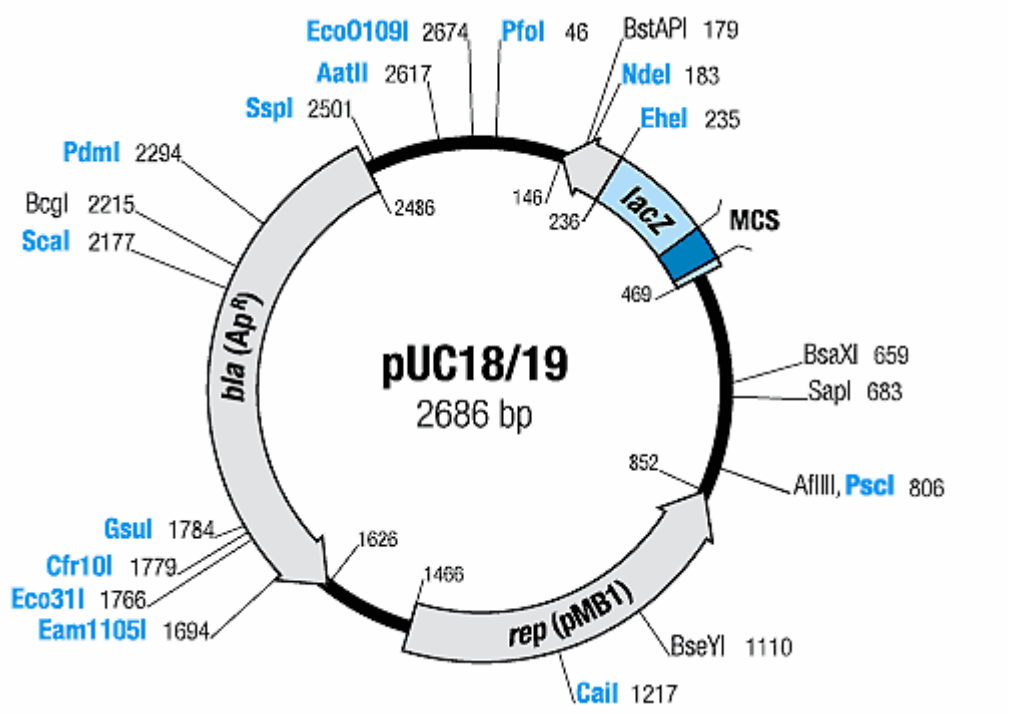
Map courtesy of Professor John Ward, UCL Research Department of Structural and Molecular Biology, University College London.

A.1.4. pGlo Plasmid Map



pGlo 5.4kb. Source: BIO-RAD

A.1.5. pUC18/19 Plasmid Map



Source: www.fermentas.com

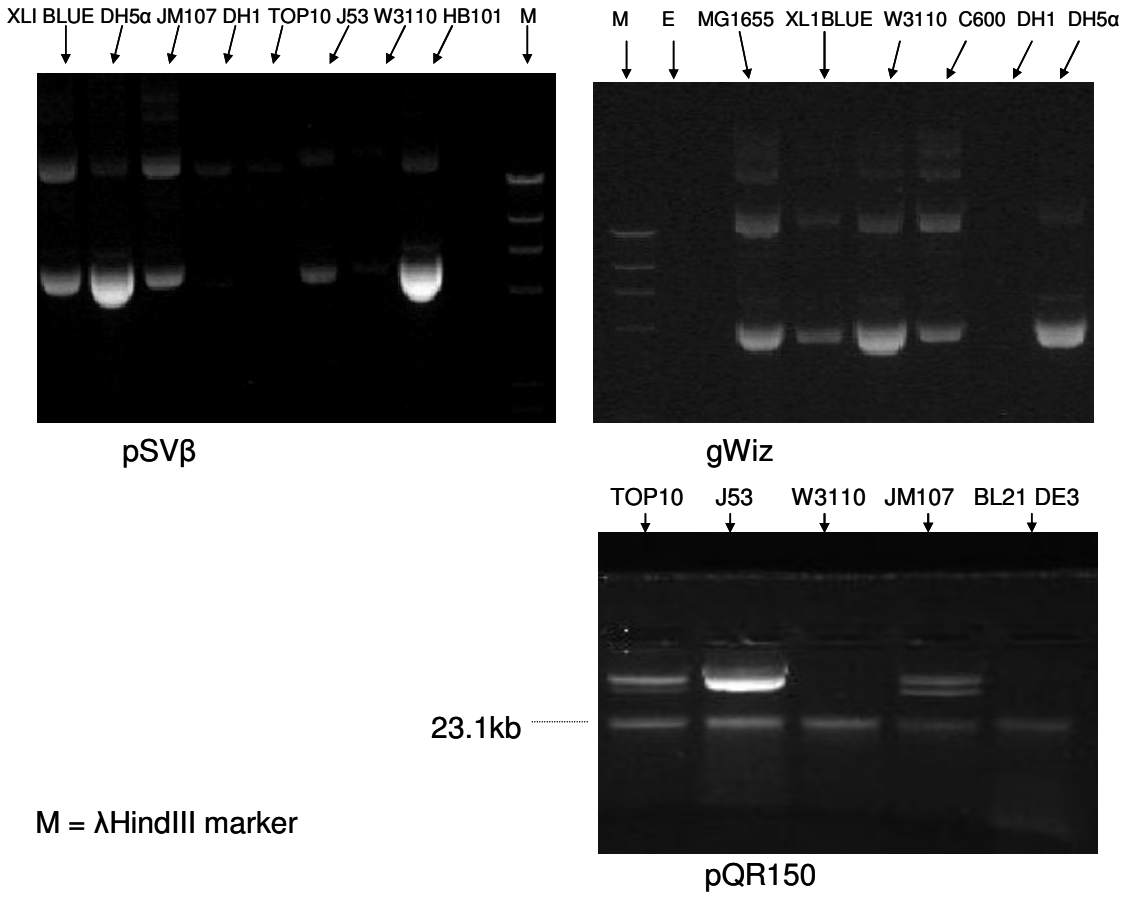
A.2. *E. coli* Mutations and Brief Descriptions

Mutation	Description
amy	Amylase producing
ara-14	Mutation in arabinose metabolism. Blocks arabinose catabolism
araD	Mutation in L-ribulose-phosphate 4-epimerase blocks arabinose catabolism. Part of an inducible operon araBAD repressed by L-arabinose
Cm^R	Encodes resistance to chloramphenicol
Δ ( )	Chromosomal deletion of genes between the listed genes ( )
deoR	Regulatory gene mutation allowing constitutive expression of genes for deoxyribose synthesis. Allows efficient propagation of large plasmids.
EndA, A1	Mutation in endonuclease I increases quality of plasmids and prevents plasmid nicking during lysis. These strains are useful for generating ss-DNA of M13, since the endonuclease digests dsDNA, but not ssDNA.
F^-	Host lacks the F' plasmid
F^+	Host carries the F plasmid. The cell is able to mate with F^- plasmids through conjugation.
F' [ ]	Carries an F plasmid that has host chromosomal genes from a previous recombination. Allows mating with F^- through conjugation. Chromosomal genes carried in the F plasmid are listed in brackets [ ].
galK	Galactokinase mutation blocks catabolism of galactose
galU	Glucose-1-phosphate uridylyltransferase mutation blocks ability to use galactose
gln V	Suppression of amber (UAG) stop codons by insertion of glutamine; required for some phage growth
gyrA96	DNA gyrase mutant produces resistance to nalidixic acid and can reduce plasmid supercoiling
hsdR, hsdS	DNA that does not contain methylation of certain sequences is recognised as foreign by EcoKI or EcoBI and restricted. These enzymes recognise different sequences and are encoded by different alleles of *hsdRMS*. *hsdR* mutations abolish restriction but not protective methylation (r^-m_+), while *hsdS* mutations abolish both (r^-m^-). DNA made in the latter will be restricted when introduced into a wild-type strain.
ilvG	Codes for acetohydroxy acid synthase (AHAS)
INV( )	Chromosomal inversion between locations indicated ( ) e.g. between *rrnD and rrnE*
Kan^R	Encodes kanamycin resistance
$lacI^q$	High levels of lac repressor protein produced, inhibits transcription from lac promoter. This repression can be overcome by binding to IPTG.
lacY1	Blocks use of lactose via lactose permease mutant
lacZ	ß-D-galactosidase gene mutation resulting in colourless colonies in the presence of X-gal.
lacZ Δ M15	A partial deletion of the NH_2-terminal region of ß-D-galactosidase to permit α-complementation with certain vectors that encode this region (e.g. pUC, M13). Δ U169,

	X111, & X74 all delete the entire lac operon. X111 also deletes proAB thus praline is required unless F'lac proAB is present. Often present on ϕ-80 or F'.
λ^-	Nonlysogenic i.e. phage will not incorporate into the chromosome.
λ(DE3)	λ phage containing a T7 RNA polymerase gene is integrated into host chromosome.
leuB	Requires leucine for growth on minimal media via a ß-isopropyl malate dehydrogenase mutation.
$m_{B/K}^{+/-}$	The (B/K) defines the strain lineage. The +/- indicates whether or not the strain contains the modification (methylation) system.
mcrA	Blocks restriction of DNA methylated at G(m)CGC.
mcrB	Blocks restriction of DNA methylated AG(mC)T.
met	Requires methionine for growth on minimal media via cystathionine ɣ-synthetase mutant.
mrr	Blocks adenine methylation; prevents cleavage of C(^{m})AG and G(^{m}A)C.
nal^R	Encodes resistance to nalidixic acid
nupG	Mutation for the transport of nucleosides
ompT	Activity of outer membrane protease (protease VII) is abolished.
Φ80	Cell carries Φ80 prophage. Some E. coli strains carry defective lac M15.
ProA, B	Requires proline for growth on minimal media
recA, A1, A13	Involved in DNA repair and recombination. Mutations reduce homologous recombination of vector with host DNA giving more stable inserts; however has been reported to result in slower growth. Cells are UV sensitive.
recJ	Exonuclease involved in recombination; alternate to recA pathway. Prevents plasmid-plasmid recombination. With recB, confers recA- phenotype. With sbcC, reduces Z-DNA rearrangements.
relA	RNA is synthesized in absence of protein synthesis (relaxed phenotype).
$r_{B/K}^{+/-}$	The (B/K) defines the strain lineage. The +/- indicates whether or not the strain contains the restriction system.
rfb-50	Is an IS5 mutation resulting in the absence of O-antigen synthesis.
rph-1	Mutation that results in frameshift over last 15 codons and has a polar effect on *pyrE* leading to suboptimal pyrimidine levels on minimal medium.
rpsL	Encoding streptomycin resistance via S12 mutant in 30S ribosome.
sbcC	Usually found with recB recC sbcB. However, strains carrying sbcC alone are recombination-proficient and stably propagate inverted repeats both in prophage l and in plasmids.
Sm^R	Encodes resistance to streptomycin
Str^R	Encodes resistance to streptomycin
supE,F	tRNA glutamine-inserting amber (UAG) (supE) or tyrosine (supF). supF is needed for growth of some phage vectors. Now called glnV or tyrT respectively.
Tet^R	Encodes tetracycline resistance

thi-1	Requires thiamine for growth on minimal media.
thr	Requires threonine for growth on minimal media.
Tn10	Encoding tetracycline resistance via a transposon.
traD36	Transfer factor mutation. Prevents transfer of F' plasmid.
umuC	Mutation in SOS gene that presents defects in some of the repair pathways induced by UV.
uvrC	Component of UV repair pathway. Reduces rearrangements of inverted repeats.
xyl-5	Blocks catabolism of xylose.

A.3. Agarose gel scans



A.2. Examples of agarose gels showing the different forms of pDNA from different *E. coli* strains. E= empty lane.