

The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

Investigation into the Biotechnological Applications of *Pantoea agglomerans* for the Production of High value Terpenoids such as Taxadiene

James Williamson BSc(Hons)

Thesis submitted to the University of Nottingham for the degree of Doctorate of Philosophy

September 2016

ABSTRACT

Plant derived terpenoids represent a diverse class of chemicals with important roles such as fragrances, flavours and pharmaceuticals as well as high value medicinal compounds such as the antimalarial Artemisinin and the anti-cancer drug Paclitaxel (Taxol). High volume production of these products can be difficult due to reliance on plant production or economic difficulties in synthetic production. These high value terpenoids represent an interesting challenge to metabolic engineering and synthetic biology, not only the requirement to produce these chemicals at scale, but also because novel products can be developed via the intermediates of these pathways.

Research into methods of microbial production of Taxol and its derivatives have focused on the engineering of *Escherichia coli* and the precursor pathways leading to geranylgeranyl diphosphate (GGPP), since GGPP is the last non-dedicated precursor to Taxol. The work reported here has taken a different approach and has instead attempted to utilise natural diversity in search of a better host for production of these high value terpenoids, by selecting an organism that already produces a large amount of GGPP. As well as being a precursor to Taxol, GGPP can also be converted into carotenoid pigments such as lycopene, β-carotene and zeaxanthin. Therefore production of carotenoids can be used as a surrogate measure for predicting how well a species will perform in the production of heterologous isoprenoids. Genes for carotenoid biosynthesis from *Pantoea* species have been used in many heterologous expression experiments, however *Pantoea* have not been considered as a possible chassis for the production of isoprenoids. As a relatively close phylogenetic relative to *E.coli*, these Gram negative bacteria are easily transformable, genetically tractable and many tools which have been developed for *E. coli* function in *Pantoea*.

The work reported here addresses the developments of tools and techniques to assess the ability of *Pantoea agglomerans* to produce terpenoids including lycopene and taxadiene. The tools developed included: 1) a range of promoters, which were characterised in *E. coli* and *P. agglomerans* for their regulation of lycopene

biosynthesis, 2) methods for the editing of the genome of *P. agglomerans* and 3) a range of vectors for the expression of the first 3 dedicated steps in Taxol biosynthesis, taxadiene synthase (*txs*), taxadiene 5α hydroxylase and taxadiene-5α-ol-O-acetyl transferase (5α) and taxane 10β-hydroxylase (10β). These genes initially could not be cloned in an operon together due to apparent toxicity. However by designing measures to minimise the burden imposed it was possible to express all four enzymes, of the three step process in *P. agglomerans*.

This work discusses the difficulties associated with working with a less well characterised organism (such as *P. agglomerans*), how we overcame these issues and the unexpected phenotypes exhibited by this genus.

ACKNOWLEDGMENTS

I want to acknowledge my supervisors Dr Phil Hill and Dr Rupert Fray for all their support and insight into a fascinating subject, which crosses both of their interests.

I would also like to thank my colleagues for all their support throughout my studies. My fellow DTP students who have been Guinea pigs along with me, and all the members of the School of Food Sciences for their support, friendship and cake throughout my studies.

I am especially grateful to my friends and family for getting me through this, escpecially the ordeal that was writing this thesis. Finally a special thanks goes to my grandad, whom got to see me finish this work, but who will not be with me at graduation, you always encouraged me to be a scientist.

-Thank you

TABLE OF CONTENTS

Abstract i

Acknowledgments iii

Table of Contents iv

List of figures xii

List of tables xvii

Abbreviations xviii

Chapter 1 Literature Review 1

1.1 High value products from plants 1

1.2 The biological purpose of Taxol 1

1.3 Chemical structure of Taxol 3

1.4 Pharmacology and pharmaceutical use 4

1.5 Biosynthesis of Taxol 6

1.5.1 MEP and Mevalonate pathway 6

1.5.2 Taxadiene synthase (TXS) 7

1.5.3 Cytochrome p450s: taxadiene 5α hydroxylase and taxane 10β hydroxylase 8

1.6 Industrial production 9

1.6.1 Initial production 9

1.6.2 Full and Semi synthesis 10

1.6.3 Cell culture 10

1.7 Heterologous production methods 11

1.7.1 Focus of the field 11

1.7.2 Plants 12

1.7.3 Microbes 12

1.7.4 Other Examples of microbial production of pharmaceuticals 17

1.8 The Genus *Pantoea* 19

1.8.1 Taxonomy 19

1.8.2 Genetics and physiology 21

1.8.3 Environments 21

1.8.4 The Large *Pantoea* Plasmid (LPP-1) 22

1.8.5 Carotenoid Biosynthesis 23

1.8.6 Biotechnology applications 25

1.9 Project aims and rationale 26

Chapter 2 Materials and methods 28

2.1 Bacterial Strains 28

2.2 Plasmids 28

2.3 Chemical reagents 31

2.3.1 Water 31

2.4 Media and reagents 31

2.4.1 Luria Bertani media 31

2.4.2 2YT media 32

2.4.3 Congo Red agar 32

2.4.4Antibiotics 32

2.5 Buffers 32

2.5.1 5X Isothermal Reaction Buffer (5X IRB) 32

2.5.2 50X TAE Buffer 32

2.6 Methods 33

2.6.1 Culturing 33

2.6.2 Measurement of bacterial growth 33

2.6.3 Enumeration of viable cells 33

2.7 DNA 34

2.7.1 DNA synthesis 34

2.7.2 Genomic DNA extraction 34

2.7.3 Plasmid DNA Extraction 34

2.7.4 Polymerase chain reaction 34

2.7.5 Primers 35

2.7.6 Restriction digestion 37

2.7.7 Agarose gel electrophoresis 37

2.7.8 DNA purification 38

2.7.9 Ligation of DNA 38

2.7.10 One step isothermal DNA assembly reaction (Gibson Assembly) 39

2.8 Electrocompetent cells 39

2.9 Electroporation 40

2.10 Insertion of transgenes using Tn*7* 40

2.11 DNA sequencing 41

2.11.1 Sanger sequencing 41

2.11.1a Plasmid 41

2.11.1b PCR product 41

2.11.2 Genome sequencing 41

2.11.2a Genomic DNA extraction 41

2.10.2b Sequencing 41

2.11.2c Assembly 42

2.11.2d Annotation 42

2.12 Carotenoid measurements 42

2.12.1 Lycopene extraction and quantification 42

2.12.2 Zeaxanthin extraction 43

2.13 Taxane extraction 43

2.13.1 Taxadiene extraction and analysis 43

2.13.2 GC-MS 44

2.13 Dry cell weight 44

2.14 Hard colony material extraction 45

2.15 Microscopy 45

2.15.1 Gram staining 45

2.15.2 Congo red stain 45

2.15.3 CCD imaging 46

2.16 Statistical methods 46

Chapter 3 Characterisations of parts and procedures for the transformation of *P. agglomerans*, and the production of lycopene 47

3.1 Introduction 47

3.2 Plasmid selection and transformation 47

3.2.1 Antibiotic sensitivity 47

3.2.2 Transformation protocol 48

3.2.3 Plasmid origins of replication 48

3.3 Assessing promoter functionality in *P. agglomerans* 49

3.3.1 pBT-GFP 49

3.3.2 pCT-GFP 50

3.3.3 pSU-GFP 50

3.4 Dynamic tracking of promoter activity through use of the bioluminescent *lux* operon 53

3.4.1Replacment of the GFP cassette with the *lux* operon 53

3.4.2 Expression of the *lux* operon under the control of P_{BAD} in *P. agglomerans* and *E. coli* 56

3.4.3 Control of the expression of the *lux* operon under the control of P_{CRT} in *P. agglomerans* and *P. agglomerans* ΔLPP-1 60

3.5 Comparison of the control and production of lycopene in *E. coli* and *P. agglomerans* 63

3.5.1 Construction of the Lycopene biosynthetic operon 63

3.4.4 Transformation of the lycopene biosynthetic pathway into *P. agglomerans* ΔLPP-1 66

3.4.5 Testing production of lycopene in *E. coli* and *P. agglomerans* using P_{LAC} 66

3.4.6 Effect of expression level on lycopene production, using pBT-CrtEIB and pBT-CrtE-idi-IB 69

3.4.7 Induction and the introduction of *idi* 72

3. 5 Discussion 80

Chapter 4 Development of chromosomal integration and gene knockout systems for use in *P. agglomerans* 85

4.1 Transposon mediated chromosomal integration 85

4.2 Chromosomal integration of *txs* into *P. agglomerans* using pGRG36t-Lac-TXS 89

4.2.1 Integration of P_{LAC}-*txs* 89

4.2.2 pGRG36t-KanR 92

4.3 Deletion/ integration system pDEL 94

4.3.1 Construction 96

4.3.2 Effectiveness of counter selection 98

4.3.3 Knock out constructs 100

4.4 Discussion 103

Chapter 5 Taxane biosynthetic operon 107

5.1 Design of the synthetic taxane operon 107

5.2 pBMH-AB and pBMH-CDE transformation 109

5. 2.1 Full operon construction. 112

5.2.2 Deletion of 10β hydroxylase 112

5. 3 Separation of the taxane genes by PCR 116

5.3. 1 pSU-TXS, pBT-5α, pBT-10β and pBT-5α10β 116

5.4 Discussion 120

Chapter 6 Introduction of taxadiene synthase into *P. agglomerans* 121

6.1 Potential DNA methylation and restriction systems in *P. agglomerans* 121

6.1.1 Wild-type morphology colonies 122

6.2 Deletion of LPP-1 from *P. agglomerans* 127

6.2.1 Selection for the loss of LPP-1 from *P. agglomerans* using pLBT-TXS 127

6.2.2 Evidence that *P. agglomerans* has a restriction system 128

6.3 Induction of *txs* under the control of P_{BAD} 130

6.4 Hard colony morphology and propagation 136

6.4.1 Culture and re-streak 136

6.5 Transformation method 137

6.5.1 Addition of glycerol and glucose to the agar plates 139

6.6 Increasing Hard colony production 140

6. 6. 1 Propagation of Hard colonies 141

6.6.2 Mucoid colony restreak 142

6.6.3 Material collection 144

6.6.4 Solvent extraction 144

6.6.5 Congo red staining 145

6.6.6. Microscopy 147

6.7 Discussion 148

Chapter 7 Co-expression of taxadiene synthase with the p450 enzymes taxadiene 5α hydroxylase and taxane 10β hydroxylase 152

7.1 Transformation of pSU-TXS into *P. agglomerans* 152

7.1.1 Differences in zeaxanthin production between *P. agglomerans* transformed with pSU8 or pSU-TXS 153

7. 2 Characterisation of pBT-5α, pBT-10β, pBT-5α10β 154

7.2.2 Growth curves 155

7.3 Co transformation of *P. agglomerans* with pSU-TXS and pBT-5α, pBT-10β, pBT-5α10α 157

7.3.1 Competent cells 157

7.3.2 Co-transformation 157

7.4 *P. agglomerans* pSU- TXS pBT-10β 158

7. 5 pSU-TXS pLBT-5α and pBT-5α 161

7.5.1 pLBT-5α 161

7.6 The pLBT-TXS series 162

7.6.1 Creation of pLBT-TXS series 162

7.6.2 Transformation into *P. agglomerans* 166

7.7 Taxane production 170

7.7.1 Culturing 170

7.7.2 Taxadiene extraction and analysis 172

7.7.3 Stability 173

7. 8 Discussion 175

Chapter 8 Discussion 179

8.1 Tools that have been developed.....179

8.2 Improvements to stability.....181

References.....185

Appendix.....202

LIST OF FIGURES

Figure 1-1 Needles and aril of *Taxus baccata*.....2

Figure 1-2 Chemical structure of taxa-(4,5),(11,12)-diene (taxadiene).....3

Figure 1-3 Chemical structure of Taxol.....4

Figure 1-4 Biosynthetic pathway for Taxol.....9

Figure 1-5 Cladogram of the Genera within the Enterobacteriaceae.....20

Figure 1-6 Genome diagram of the Group I LPP-1 pan-genome.....23

Figure 1-7 Zeaxathin biosynthesis.....24

Figure 2-1 Promega 1Kb ladder.....38

Figure 3-1 Plasmid maps of pBT-GFP, pCT-GFP and pSU-GFP.....52

Figure 3-2 Plasmid maps of pBT-LUX and pCT-LUX.....55

Figure 3-3 relative light production from *E. coli* and *P. agglomerans* ΔLPP-1 transformed with pBT-LUX.....58

Figure 3-4 Area under the curve of the data from figure *3*.4.....59

Figure 3-5 assessment of the dynamics of P_{CRT} in wild-type *P. agglomerans* and *P. agglomerans* ΔLPP-1. 62

Figure 3-6 Cloning strategy for creating *crtEIB* and *crtE-idi-IB* operons. 65

Figure 3-7 comparison of *E. coli* transformed with pBT-crtEIB and pBT-crtE-idi-IB ..66

Figure 3-8Production of lycopene in *E. coli* and *P. agglomerans* LPP-1 transformed with pSU-crtE-idi-IB. 68

Figure 3-9 Production of lycopene from *E. coli* and *P. agglomerans* LPP-1 transformed with pBT-crtEIB with and without induction by the addition of arabinose. 71

Figure 3-10 Production of lycopene per litre after 24 hours in *E. coli* and *P. agglomerans* ΔLPP-1 transformed with either pBT-crtEIB or pBT-crtE-idi-IB under three levels of induction from the addition of arabinose.....74

Figure 3-11Production of lycopene from *E. coli* and *P. agglomerans* LPP-1 transformed with pBT-crtEIB or pBT-CrtE-idi-I under three levels of induction 75

Figure 3-12 Production of lycopene after 72 hours in *E. coli* and *P. agglomerans* ΔLPP-1 transformed with pCT-crtEIB or pCT-crtE-idi-IB.....77

Figure 3-13 Production of lycopene over 72 hours in *E. coli* and *P. agglomerans* ΔLPP-1 transformed with pCT-crtEIB or pCT-crtE-idi-IB. 78

Figure 3-14 Production of lycopene after 72 hours in *E. coli* and *P. agglomerans* ΔLPP-1 transformed with pCT-crtEIB or pCT-crtE-idi-IB. 79

Figure 4-1 Plasmid maps of pGRG36 and pGRG36T. 87

Figure 4-2 Plasmids maps of pSU-TXS and pGRG36T-lacTXS. 88

Figure 4-3 Colonies of *P. agglomerans* following attempted chromosomal integration of P_{LAC}-TXS. 91

Figure 4-4 Re-attempted chromosomal integration with and without the addition of ampicillin during the culture stage. 92

Figure 4-5 Plasmid map of pGRG36T-KanR.....94

Figure 4-6 Diagram of homologous recombination for the knockout of chromosomal genes with use of a counter selectable marker. 95

Figure 4-7 Plasmid maps of pDEL2 and pDEL-*crtE*-dsCrtZ 97

Figure 4-8 The effect of sucrose and increased levels of nickel to the toxicity of *P. agglomerans* harbouring a plasmid carrying t*et*R and *sacB* 99

Figure 4-9 *P. agglomerans* zeaxthin biosynthesis operon *crtEXYIBZ*.....101

Figure 4-10 Patched out colonies being screened for the deletion of *crt*XYIBZ, ...102

Figure 5-1 Organisation of the Taxane synthetic operon. 109

Figure 5-2. Map of pBMH-AB and pBMH-CDE from Biomatik 111

Figure 5-3. pBMH-CDE *Nco*I restriction digest. 115

Figure 5-4 pSU9 and pSU9-TXS. 118

Figure 5-5 Plasmid map of pBT-GFP pBT Taxane variants. 119

Figure 6-1 Colony morphology of *P. agglomerans* transformed with pBT-TXS.....124

Figure 6-2 pBT-DAM1.....125

Figure 6-4 Colony recovery in *P. agglomerans* transformed with plasmid DNA prepared from *E. coli* and *P. agglomerans* ΔLPP-1. 129

Figure 6-5 Effect of induction on stability of pBT-TXS 131

Figure 6-6 pBT-DAM2 132

Figure 6-7 Stability of pBT_TXS over 6 days. 134

Figure 6-8 6 day course of plasmid restrictions for induced and uninduced 134

Figure 6-9 pBT-DAM3 135

Figure 6-10 Re-streaked hard colony. 138

Figure 6-11 Effect of glycerol and glucose on wild-type *P. agglomerans* morphology. 140

Figure 6-12 Effect of glycerol and glucose on Hard colony production in *P. agglomerans* transformed with pBT-TXS. 141

Figure 6-13 Propagation of Hard colonies and their growth. 142

Figure 6-14 Hard and mucoid colony restreak 143

Figure 6-15 Hard colonies forming on Congo Red LB agar. 146

Figure 6-16 Microscopy of hard colony and hard colony debris 148

Figure 7-1 Production of zeaxanthin in wild-type *P. agglomerans* with and without the plasmids pSU8 or pSU-TXS 154

Figure 7-2 Growth curve of several strains of *P. agglomerans*. 156

Figure 7-3 Growth curve of co transformed *P. agglomerans* 159

Figure 7-4 Differing colony morphology of *P. agglomerans* 160

Figure 7-5 pLT-1 and pLBT-1. 164

Figure 7-6 pLBT-GFP map and the pLBT-TXS series 165

Figure 7-7 Restriction digest confirmation of the production of pLBT-TXS-5α, pLBT-TXS-10β and pLBT-TXS-5α10β 166

Figure 7-8 Restriction digest confirmation of the transformation of *P. agglomerans* ΔLPP-1 with pLBT-TXS-10β 167

Figure 7-9 Restriction digest confirmation of the transformation of *P. agglomerans* with pLBT-TXS-10β 168

Figure 7-10 Restriction digest confirmation of the transformation of *P. agglomerans* ΔLPP-1 with pLBT-TXS-5α, and pLBT-TXS-5α10β 169

Figure 7-11 Restricion digest confirmation of the transformation of *P. agglomerans* with pLBT-TXS-5α and pLBT-TXS-5α10β 169

Figure 7-12 OD600 and viable cell counts of cultures of *P. agglomerans* pLBT-TXS series. 171

Figure 7-13 Plasmid Stability of pLBT-TXS series in P. agglomerans over 6 days of culture. 175

LIST OF TABLES

Table 1.Heterlogous production of terpenoids 16

Table 2 Bacterial strains 28

Table 3 Plasmids.....30

Table 4 Oligonucleotides used for cloning insertional sequences and plasmids 36

Table 5 Oligonucleotides used for verification of DNA sequences.....37

Table 6 pBMH-CDE restriction digest.....114

Table 8 Percent change in light production under different levels of induction in *E. coli* and *P. agglomerans* ΔLPP-1 transformed with pBT-LUX.....59

Table 9 Influenece of incubation temperature on Hard colony formation 138

Table 10 . The addition of glucose to media and the influence on Hard colony production.....139

ABBREVIATIONS

10β	Taxane 10β-hydroxylase
5α	Taxadiene 5α- hydroxylase and taxadiene-5α-ol-O-acetyltransferase
*crt*B	Phytoene synthase
*crt*E	Geranylgeranyl diphosphate synthase
*crt*I	Phytoene desaturase
DMAPP	Dimethylallyl diphosphate
FPP	Farnesyl diphosphate
GFP	Green florescent protein
GGPP	Genranylgenranyl diphosphate
GPP	Geranyl diphosphate
idi	Isopentenyl diphosphate isomerase
IPP	Isopentenyl diphosphate
LPP-1	Large *Pantoea* Plasmid-1
PCF	Plant cell fermentation
RPM	Revolutions per minute
txs	Taxadiene synthase

Chapter 1 LITERATURE REVIEW

1.1 High value products from plants

For millennia humans have appreciated the pharmaceutical uses provided by plants and products derived from them. From the pain relief provided by the opioids derived from poppies and salicylic acid identified from willow bark to anti-malarials such as artemisinin from *Artemisia annua.* In the examples of opioids and artemisinin, the plants from which these drugs are derived can be farmed in order to produce industrially relevant quantities, however not all plant derived chemicals can be harvested in such a manner. Either the plants they are produced by are not amenable to agriculture, requiring conditions that are difficult to reproduce on scale, or they produce such small amounts that production in this manner is not economically viable. For these chemicals, alternative methods such as chemical synthesis from petrochemicals or chemical semi synthesis from plant intermediates are a route to industrial production. For compounds which are more complex in structure, or for which suitable quantities of feedstock can be secured, industrial biotechnology approaches can be used. Biosynthetic pathways can be identified and then be expressed in other hosts, more amenable to industrial production such as food crops or microbial systems such as yeast or *Escherichia coli*.

One natural product of note is Taxol (generic name paclitaxel), an anti-cancer chemotherapy agent derived from the bark of the yew tree *Taxus brevifolia*. It was discovered during the 1960's through work between the National Cancer Institute (NCI) and the U.S. Department of Agriculture (USDA) which screened over 100,000 plant extracts looking to identify naturally occurring anticancer compounds (Wani *et al.*, 1971)

1.2 The biological purpose of Taxol

Taxol occurs naturally in the bark of *T. brevifolia* and is thought to have evolved alongside a range of chemical defences of *T. brevifolia,* as cell cultures can be

stimulated to produce more Taxol with the addition of methyl jasmonate (a plant hormone associated with plant defence regulation) (S. Li *et al.*, 2012). Extracts of the plant have a bitter taste and are toxic to deter predation by animals. Ingestion of Taxus material can be lethal when consumed by humans, with several reported incidents of attempted and successful suicides by this method as well as accidental ingestion (Grobosch *et al.*, 2013). Consumption is also one of the most common causes of livestock and companion animal poisonings in Europe (Cortinovis & Caloni, 2013), which tend to occur due to incorrect disposal of tree clippings. All parts of the tree are poisonous except for the scarlet aril (berry), however the seeds within are toxic (Grobosch *et al.*, 2013).

In livestock the first symptom that individuals have consumed yew tree material is the death. Examination of gut contents usually reveals visible evidence of needles from the tree. In humans symptoms include cardiac failure and ventricular fibrillation (Soumagne *et al.*, 2011)



Figure 1-1 Needles and aril of *Taxus baccata* (www.en.wikipedia.org/wiki/Taxus)

Extracts from Taxus species have also been shown to have antimicrobial properties against some species of bacteria such as *Staphylococcus aureus*, *Pseudomonas aeruginosa* and *Salmonella typhi* (as well as fungi such as *Trichophyton longifusus, Microspoum canis, and Fusarium solani* (Nisar *et al.*, 2008).

The full contents of chemicals extracted from Taxus is extensive, but the main constituents which confer toxicity to animals are the taxine alkaloids which were identified in the 19th century. After the discovery of the anti-cancer properties of *Taxus* extract, it was found that Taxol was the active agent responsible for the cytotoxicity towards cancers cells (Wani *et al.*, 1971).

1.3 Chemical structure of Taxol

Taxol is a member of the taxoids, a group of alkaloid diterpenes characterised by their possession of a taxane skeleton *Fig.1.2* often referred to as taxadiene. The structure of Taxol, shown in *Fig.1.3,* is composed of: a taxane skeleton, 2 chain arms (carbon 2 and 13), several hydroxide groups (carbon 1 and 7) and an oxetane ring (carbon 4,5 and 20). The oxetane ring is a chemical structure very rarely found in nature outside of the taxanes.

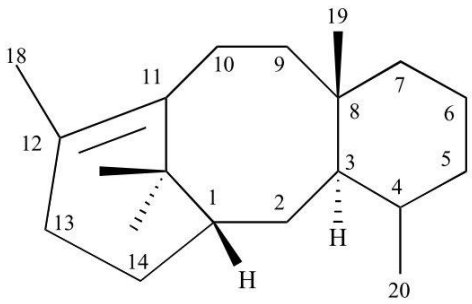


Figure 1-2 Chemical structure of taxa-(4,5),(11,12)-diene (taxadiene) The first dedicated intermediate of Taxol biosynthesis, and the core skeleton of the taxanes. (taxadiene) (Expósito *et al.*, 2009)

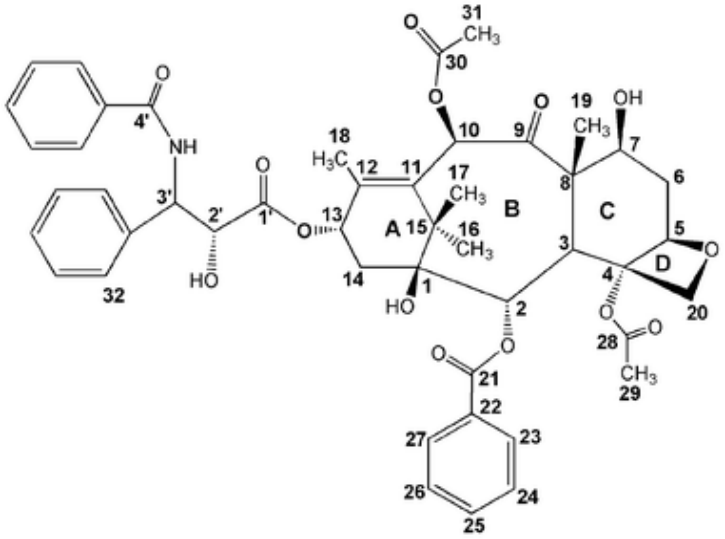


Figure 1-3 Chemical structure of Taxol (Heider, Harper, & Grant, 2007)

1.4 Pharmacology and pharmaceutical use

Taxol has been used to treat a range of cancers including breast (Wang *et al.*, 2009), lung (Kamath *et al.*, 2004) and neck cancer (Argiris *et al.,* 2008) among others.

Taxol acts to stabilise tubulin polymerisation, and so disrupts dynamic instability of microtubules. The structure of tubulin was solved by x-ray crystallography, and as a by-product of using Taxol to preserve the samples, it was also shown how Taxol binds to tubulin. This helped elucidate the way Taxol stabilises microtubules (Nogales *et al.*, 1998). Tubulin consists of a hetero dimer of α and β subunits that can bind guanine residues and their subunit counterpart, this binding traps GTP. This GTP is eventually hydrolysed to produce GDP, which shortens the dimer by 3-6%. When bound to GTP, the filament is more likely to grow; when GDP is bound the filament is more unstable. Taxol binds to a pocket in the β subunit. This binding stabilises the polymerisation of tubulin by countering the hydrolysis of GTP (Amos & Löwe, 1999).

Yeast β tubulin subunits are not susceptible to the stabilising effect of Taxol. Humans have several isotypes of β tubulin, βIII tubulin (found mainly in the brain) has been

shown to be resistant to Taxol in a similar manner to yeast β tubulin and therefore can play a role in treatment failing (Das *et al.,* 2012).

Compared to other tubulin stabilising agents such as epothilones, from *myxobacterium Sorangium cellulosum* (Mühlradt *et al.,* 1997), Taxol also promotes cell death via apoptosis. In eukaryotic cells such as human, Bcl-2 is a regulatory protein, which normally acts to suppress the apoptosis pathway. Paradoxically it is often down regulated in Taxol resistant cancer cells whereas, most chemotherapy resistant cancers over express anti-apoptosis factors. Ferlini *et al.* found that Bcl-2 can be activated by Nur77 to initiate apoptosis (Ferlini *et al.*, 2009). Nur77 can also bind to β tubulin and produces similar effects to Taxol. They also showed that the oxetane ring interacts hydrophobically with both tubulin and Bcl-2 by mimicking a proline residue found in Nur77. All these factors suggest that Taxol mimics Nur77 in order to modulate the apoptosis pathway and was derived by convergent evolution.

Taxol has also been found to increase chromosomal instability in tumour cells. This is initiated by the unequal separation of chromosomes during cell division leading to the erosion of the genome and cell death, adding to its cytotoxicity towards fast growing cancer cells (Lee *et al.*, 2016).

Tumour cell resistance to Taxol is a common occurrence. Resistance is normally mediated by the upregulation of ABC transporters which allow the cancer cells to export Taxol, and other chemotherapy agents, therefore reducing retention time and cytotoxicity. Several novel taxanes have been shown to have potential as agents to reverse tumour multi drug resistance (Ojima *et al.*, 2005) (Brooks *et al.*, 2004) (Brooks *et al.*, 2003) (Minderman *et al.*, 2004), for example Ortataxel which can be up 400-fold more active against multi drug resistant cells, this shows that the taxanes could still yield more drugs, with improved or novel pharmacological properties.

As Taxol and related products have been on the market for over 20 years now, several other uses for Taxol and other taxanes have been investigated. Applications range from treating degenerative conditions such as Alzheimer's and Parkinson's disease

by offsetting tau protein sequestration (a protein that stabilises microtubules within neurons) (Zhang *et al.*, 2005). Other applications focus on tackling fast replicating cells such as in psoriasis (Ehrlich *et al.*, 2004). Taxol is also being used to coat medical equipment, such as arterial stents, to prevent restenosis after coronary artery angioplasty (Loh & Waksman, 2012). With the rise of antibiotic resistance in many pathogenic bacteria, there has also been investigations into the potential use of precursors to and derivatives of Taxol for the treatment of *Mycobacterium tuberculosis* (Huang *et al.,* 2007)(Huang *et al.*, 2006). With the continued use of Taxol and the exploration into related chemicals with other properties the demand for these chemicals is increasing and improvements to production methods need to be investigated.

1.5 Biosynthesis of Taxol

The full biosynthetic pathway for Taxol is still being elucidated, the reasons for the difficulty in solving this pathway is in the complexity of the molecule in question and its unusual structure and chemistry, this is compounded by low yields of intermediate products from plant extracts and the range of other products present in plant extracts (Guerra-Bubb *et al.* 2013).

1.5.1 MEP and Mevalonate pathway

The early stages of biosynthesis are well understood, as Taxol biosynthesis stems from the diterpenoid branch of terpenoid biosynthesis which includes other products like the C40 carotenoids such as lycopene. Terpenoids are a broad class of chemical, which include a large proportion of natural products including: monoterpeoids, carotenoids and steroids, and they are the basis of many plant secondary metabolites such as fragrances and pigments. In bacteria they are also used to produce pigments, but also coenzymes essential for cellular function such as ubiquinone for respiration (Meganathan, 2001). Terpenoids are built up from two, five carbon (C5) subunits Dimethylallyl diphosphate (DMAPP) and Isopentenyl diphosphate (IPP). These

building blocks can be synthesised via two different pathways, the mevalonate pathway and the MEP pathway.

The mevalonate pathway is utilised by eukaryotes and some bacteria. Acetyl-CoA is used as a starting point to produce mevalonate which in turn is used to produce IPP, this is then interconverted to DMAPP via Isopentenyl pyrophosphate isomerase (IDI).

The MEP pathway (sometimes known as the DXS pathway), is used by some bacteria and within the plastid of plants. The MEP pathway uses glyceraldehyde 3-Phosphate and pyruvate as a starting material which is converted to 2-c-methyl-d-erythritol 4-phosphate from which IPP and DMAPP can be produced 1-deoxy-D-xylulose 5-phosphate reductase (DXR) and then interconverted by the isomerase IDI (Expósito *et al.*, 2009).

DMAPP and IPP are then reacted to produce geranyl diphosphate (GPP), with the further addition of 2 units of IPP producing farnesyl diphosphate (FPP) and then geranylgeranyl diphosphate (GGPP) by Geranylgeranyl-diphosphate synthase. GPP is the branching point for monoterpenoid biosynthesis of products such as limonene, menthol and camphor. FPP can be used in the production of steroids and GGPP is the branching point for the diterpenoids such as carotenoids (lycopene, zeaxanthin and astaxanthin) and Taxol.

Biosynthesis of isoprenoids and terpenoids in *Taxus spp.* uses both the mevalonate and MEP pathway. Cytosolic production of IPP and DMAPP occurs via the mevalonate pathway whilst the plastid uses the MEP pathway and is the site of taxadiene production (Kumaran, Kim, & Hur, 2010), which is the first dedicated step in Taxol biosynthesis which is carried out by taxadiene synthase (*txs*).

1.5.2 Taxadiene synthase (TXS)

For the biosynthesis of Taxol, GGPP is then cyclised by taxadiene synthase (TXS) to produce taxa-(4,5),(11,12)-diene (taxadiene). Alternatively 2 molecules of GGPP can be combined to produce phytoene, the first dedicated step in C40 carotenoid

biosynthesis. Taxadiene then leaves the plastid into the cytosol where it is decorated in a variety of hydroxyl groups.

Taxadiene synthase was identified by being purified from yew tree extracts and shown to cyclise GGPP into taxadiene (Hezari *et al.,* 1995), and through the use of cDNA the sequence of *txs* was isolated (Wildung & Croteau, 1996). Further studies showed that the N-terminal of TXS has an extra ~80 residues which are removed during processing, and are most probably a signal peptide, once TXS was crystallised without ~106 residues from the N-terminus (Köksal *et al.,* 2011).

1.5.3 Cytochrome p450s: taxadiene 5α hydroxylase and taxane 10β hydroxylase

Taxol biosynthesis is summarised in figure 1.4. Taxadiene is hydroxylated at the C5 position to produce taxa-4(20),11(12)-dien-5α-ol (Jennewein *et al.,* 2004) via taxadiene-5α-hydroxylase (5α) a cytochrome p450. 5α also mediates the migration of the double bond from the 4(5) to 4(20) (Croteau *et al.*, 2006) which plays a role in the formation of the oxetane ring. The 5α- hydroxyl group is then acetylated by taxadiene-5α-ol-O-acetyl transferase, forming taxa-4(20),11(12)-dien-5α-yl-acetate (Walker *et al.,* 2000). Following this the C10 position is hydroxylated by taxane 10β-hydroxylase (10β) another cytochrome p450 (Schoendorf *et al.,* 2001). After this point the exact mechanism for the oxetane ring formation is not fully understood (Kaspera *et al.,* 2010). Following the oxetane ring formation, further hydroxylations and the addition of the chain arms occur. In this study we are interested in attempting to produce precursors that occur before the formation of the oxetane ring.

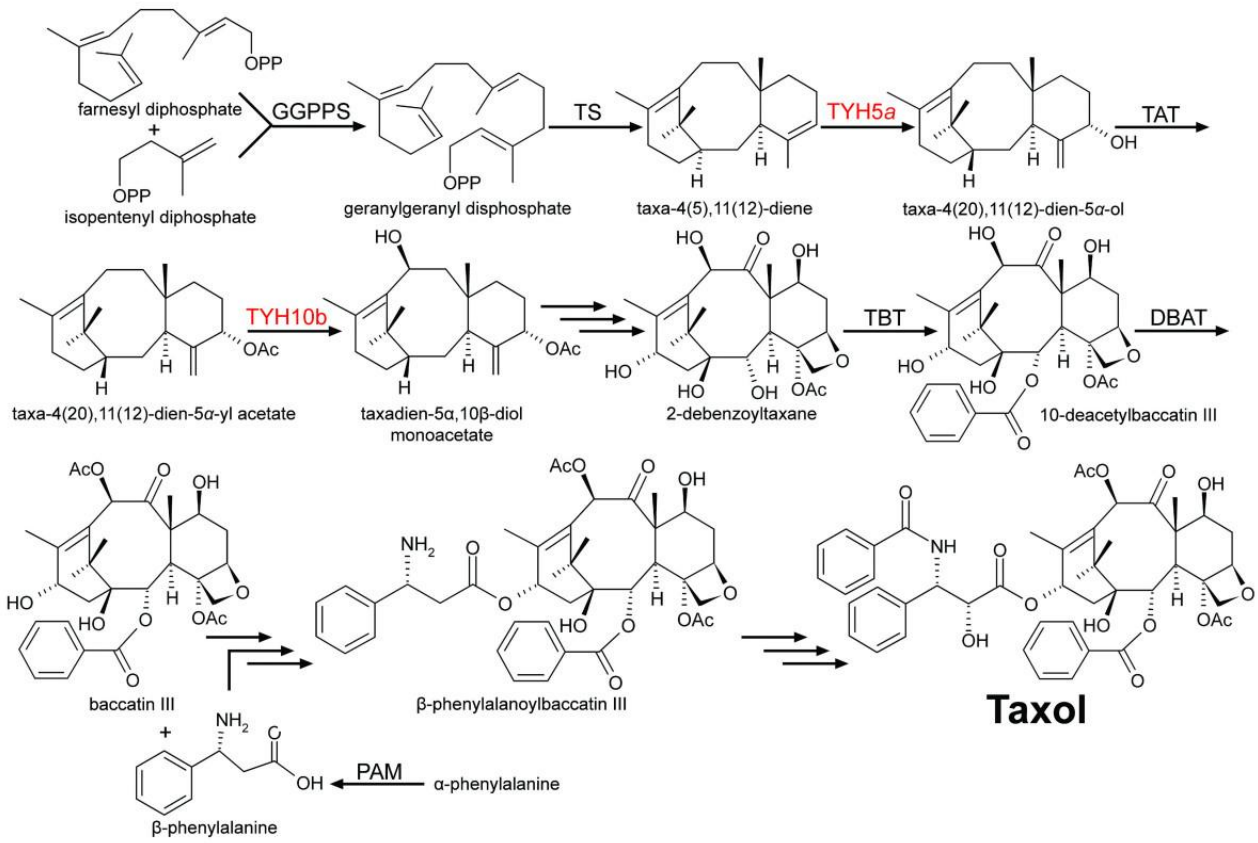


Figure 1-4 Biosynthetic pathway for Taxol, GGPPS: Geranylgeranyl diphosphate synthase; TS: taxadiene synthase (later referred to as TXS); TYH5a: taxadiene 5α- hydroxylase; TAT: taxa-4(20), 11(12)-dien-5a-ol-O-acetyltransferase (later referred to as 5α); TYH10b: taxane 10β-hydroxylase (later referred to as 10β); TBT: taxane 2a-O-benzoyltransferase; DBAT: 10-deacetyl baccatin III-O-acetyltransferase. PAM: phenyalanine aminomutase. Multiple arrows indicate several as yet undefined steps. Red text indicates a cytochrome P450 enzyme (Chemler *et al.,* 2006)

1.6 Industrial production

1.6.1 Initial production

Initial production of Taxol required the use of an entire 100 year old Pacific yew tree to produce 3kg of bark leading to 300mg of Taxol, which was the amount used per patient in initial experimental trials Clearly this production method was not sustainable or ethical due to the limited supply and environmental cost (Horwitz, 1994). The low yield per tree and the slow growth of yew trees caused an ethical

dilemma in balancing the environmental impact of harvesting yew trees and not being able to produce enough Taxol to treat all patients that require it. To mitigate this alternative methods of production were looked for.

1.6.2 Full and Semi synthesis

A fully chemically synthetic route was discovered in 1994 by Holton *et al.*, (1994), however this was not scalable or commercially viable. Instead semi synthesis methods were developed. It was shown that baccatin III and 10-deacetyl baccatin III, precursors to Taxol, could be extracted from the needles of the European yew tree and then chemically converted into Taxol (Yue *et al.,*1996) (Baloglu & Kingston, 1999). This route of semi-synthesis allowed for a greater amount of Taxol to be produced per tree and was more sustainable. However; this process still had limitations and was expensive due to the amount of solvents required for the difficult extraction of the precursors.

1.6.3 Cell culture

In 1990 Python Biotech (a spin out from Cornell University) was formed with the aim of commercialising the production of cell cultures from the yew tree. Plant Cell Fermentation (PCF) greatly reduced the need for solvents and increased yields. For this work Bristol-Myers Squibb (BMS) were awarded the Presidential Green Chemistry Challenge Award. This method has been run since 1993 at their plant in Germany with a capacity of 75,000L (Biotech, 2013). PCF produces higher amounts of Taxol and the precursors, baccatin III and 10-deacetyl baccatin III than previous methods, but like the other methods that use plant material, only small quantities of the earlier intermediates are extracted.

Several strains of endophytic fungi have been isolated from different Yew trees from around the world that appear to produce Taxol. While studies such as the culturing of the Iranian strain (SBU-16) produce such small quantities of Taxol (6.9µg/L)

(Mirjalili *et al.,* 2012) other groups have produced higher concentrations, for example 4786.9 µg/L (Kumaran *et al.*, 2010) These strains may become a source of Taxol in their own right, but more likely they will be used in a co-culturing technique with PCF. It has been shown in PCF that levels of Taxol increase when cells are exposed to jasmonic acid, a compound known to be involved in shifting metabolism in defence against microbes (Wilson & Roberts, 2012), It may be that these endophytes elicit a similar response.

1.7 Heterologous production methods

1.7.1 Focus of the field

Heterologous production of Taxol and its intermediates is a particularly desirable endeavour, firstly due to the cost of Taxol, but also because of the poorly understood intermediates and their potentially interesting chemical and biochemical properties that currently cannot be easily explored. The oxetane ring is a subject of particular interest, due to its bioactivity and its rarity in biological chemistry.

The terpenoids contain a large diversity of interesting natural products. For this reason, the field of metabolic engineering has been very interested and active in the aim of producing organisms that can produce these compounds. The areas of isoprenoid metabolic engineering was recently reviewed by Immethun *et al.* (2013), and Table 1 was adapted from this review and lists examples of strategies used to produce isoprenoids relevant to this thesis.

In the attempts to produce Taxol in a heterologous systems, most projects have focussed on the production of the first dedicated precursor, taxadiene. While taxadiene itself is not considered to have any pharmacological properties, it is an interesting compound for synthetic chemists. Due to the low yield of intermediates of Taxol from natural sources and from chemical synthesis it has been difficult to

produce and study novel taxanes. There have been a wide selection of organisms and strategies employed to produce it in a heterologous system.

1.7.2 Plants

Research into transgenic plants has investigated expressing taxadiene synthase in the whole plant and in specific tissues. Attempts have been made to use *Arabidopsis thaliana* and produced up to 600 ng/g of dry weight(Besumbes *et al.*, 2004), whilst initial attempts produced only 25 ng/g dry cell weight, growth was retarded and leaves were pale, indicating a reduction in carotenoids. Li *et al.*(2015) were able to engineer the *Artemisia annua L,* which naturally produces artemisinin, to produce 129.7 µg/g dry weight of taxadiene. Instead of using the whole plant Kovacs *et al.* (2007) produced taxadiene in tomato plants. They expressed TXS in the fruit, which naturally produces large amounts of carotenoids and depletion of which was unlikely to interfere with developmental signals. This study achieved a yield of 160 mg/kg of fruit, although they commented that the methods of extraction could have been improved and that due to the volatility of taxadiene some product was lost in processing.

1.7.3 Microbes

Saccharomyces cerevisiae, has been used to produce taxadiene however compared to some of the *E. coli* systems its yields are low. 8.7 mg/l (Engels *et al.,* 2008). In *E. coli* Ajikumar *et al.*, (2010) managed to produce taxadiene at a relatively high concentration for microbial chemical production (1 g/l) (Ajikumar *et al.* 2010). This was achieved in *E. coli* by modularising the MEP pathway and the terpenoid pathway of taxadiene (GGPPS and TXS), and then varying expression of the two parts for optimal production. The terpenoid module encoded for an *E. coli* codon optimised version of Taxadiene synthase from *Taxus canadensis.* After the initial high production, the addition of the 5α hydroxylation step caused a drop in production to around 58 mg/litre.

They showed that the native MEP pathway, which has been previously attempted to be optimised for greater production of isoprenoids with little success, was inhibited by an accumulation of indole during production of taxadiene. The mechanism by which this occurs is not clear, but it is suspected to be due to a synergistic effect on growth by taxadiene and indole.

Heterologous taxadiene production in B and K derived *E. coli* were analysed by Boghigian and colleagues (2012). This study also used a modular approach to determine the effect of promoter strength and strain background on production levels; most notably testing the MEP pathway under control of the inducible T7 promoter.

Boghigian *et al.* (2012) modified an isoprenoid computer simulation to include GGPPs and TS. From this 12 new target genes were identified for overexpression studies (8 from the isoprenoid pathway 4 from outside). From these 12 targets: overexpression of Isopentenyl diphosphate isomerase (IDI) (From within the isoprenoid pathway), Polyphosphate kinase (Ppk), Pyridine nucleotide transhydrogenase (SthA) and Phosphoribosylglycinamide formyltransferase (PurN) were found to significantly increase the production of taxadiene, when overexpressed and also produced synergistic effects.

Isoprenoid	Approach	Microbial production (fold improvement)	Microbe	Natural source	Chemical synthesis
Lycopene	Express Erwinia carotenoid biosynthesis gene cluster and idi from *Haematococcus pluvialis*	1.03 mg/g dcw (4.5-fold) (Kajiwara, *et al.* 1997)	*E. coli*	Tomatoes 0.15–0.25 mg/g (Rath *et al.,* 2007)	0.13 mg/g and 70% trans configurations (Olempska-Beer, 2006)
	Redesign the global regulatory system, the Ntr regulon	160 mg/l (from 0mg/l) (Farmer & Liao, 2000)	*E. coli*		
	Overexpress the catalytic domain of HMG and disrupt ERG9	7.8mg/g dcw (7-fold) (Shimada *et al.*, 1998)	*C. utilis*		
	Overexpress genes identified by the FSEOF strategy combined with gene knockouts	12.32 mg/g dcw (4-fold) (Choi *et al.,*2010)	*E. coli*		
	Overexpress and knockout genes selected from a metabolic landscape	16 mg/g dcw (4-fold) (Jin & Stephanopoulos, 2007)	*E. coli*		
	Use “global transcription machinery engineering” to improve phenotypes	7.7 mg/l (1.8-fold) (Alper & Stephanopoulos, 2007)	*E. coli*		
	Optimize DXP pathway with “multiplex automated genome engineering”	9 mg/g dcw (5-fold) (Wang *et al.*, 2009)	*E. coli*		
	Promoter engineering to optimise pathway	529.45 mg/L	*E. coli*	(Shen *et al.,* 2015)	

	balance. Mevalonate pathway of S. cerevisiae, was also optimised for E.coli and all pathways were chromosomally integrated.	20.25 mg per gram of dry cell weight.			
Zeaxanthin	Overexpress different combinations of idi from *Xanthophyllomyces dendrorhous*, *dxr* from *Sulfolobus acidocaldarius*, and native *dxs*	1.6 mg/g dcw (3.5-fold) (Albrecht, Misawa, & Sandmann, 1999)	*E. coli*	Tagetes erecta's red flowers 23% dry weight (Stankovic, 2004)	12% Overall yield of racemic mix (Khachik & Chang, 2009)
	Use the "ordered gene assembly in *Bacillus subtilis* (OGAB) method" to determine optimal gene order	820 µg/g dcw (4.4-fold) (Nishizaki, Tsuge, Itaya, Doi, & Yanagawa, 2007)	*E. coli*		
Taxadiene	Express genes for GGPPS, taxadiene synthase, three cytochrome P450 hydroxylases, and three acyl/aroyl CoA dependent transferases and build a five step taxoid pathway	1 mg/l (100-fold) (Dejong *et al.*, 2006)	*S. cerevisiae*	*Taxus brevifolia* 0.01–0.1% dry bark weight (Hezari *et al.*, 1995)	18–20% Overall yield (Mendoza, Ishihara, & Baran, 2012)
	Express genes for geranylgeranyl diphosphate synthase from Sulfolobus acidocaldarius and a codon-optimized taxadiene synthase from *Taxus chinensis*	8.7 mg/l (40-fold) (Engels *et al.*, 2008)	*S. cerevisiae*		
	Vary small pathway modules simultaneously to determine the optimally balanced	1 g/l (15,000-fold) (Ajikumar *et al.*, 2010)	*E. coli*		

	complete pathway ("multivariate modular pathway engineering")			

Table 1. Heterologous production of terpenoids Summary of strategies used for the production of a range of terpenoids pertinent to this project. Adapted from (Immethun *et al.*, 2013)

1.7.4 Other Examples of microbial production of pharmaceuticals

The commercial production of semi synthetic artemisinin is possibly currently the greatest achievement by the field of synthetic biology.

The aim of the project was to produce the antimalarial drug in a microbial system (initially *E. coli*), so that it could be produced cheaply through fermentation, rather than from extraction from crops of *Artemisia annua.* Artemisinin biosynthesis diverts from the terpenoid biosynthesis pathway at the point of FP, and so the project started off by trying to produce amorphadiene in *E. coli*, by introducing the amorphadiene synthase (ADS) from *Artemisia annua.* Alongside this they adjusted the expression of DXS, which had been identified as a rate limiting step this led to a modest increase in production (Reiling *et al.*, 2004). The addition of the mevalonate pathway from yeast and the use of an *E. coli* codon optimised ADS gene greatly increased the production of IPP and DMAPP, and this produced much higher yields (Martin, Pitera, Withers, Newman, & Keasling, 2003). Further improvements were made to this strain based upon systems biology data about the accumulation of potentially toxic intermediates, and improvements were made to the fermentation conditions, which led to a strain capable of producing 25 g/l of amorphadiene (Tsuruta *et al.*, 2009).

Having produced a high yield of the first dedicated precursor to artemisinin, the next step requires the hydroxylation of amorphadiene to artemisinic acid, by a cytochrome p450 (CVP71AV1). It was known that functional expression of this gene would be difficult in *E. coli*, and so two streams of work were set up. One group worked on producing artemisinic acid in *E. coli*, and the other in *S. cerevisiae*. Eventually this led to an *E. coli* strain which could produce 1 g/l (Chang *et al.,* 2007) and a *S. cerevisiae* strain that could produce 2.5 g/l (Lenihan *et al.,* 2008) of artemisinic acid. However the *E. coli* strain could only do so at 20°C, a temperature too low to be maintained economically at an industrial scale. Further improvements were made to the expression levels of the pathways involved in order to maximise production, however this also led to a rise in production of artemisinic aldehyde. To

remove this aldehyde dehydrogenase and alcohol dehydrogenases from *A. annua* were screened. The uses of these dehydrogenases and the removal of promoter regulators produced a strain capable of producing more than 25 g/l of artemisinic acid, which could then be extracted and chemically converted into artemisinin (Paddon *et al.*, 2013) .

This work highlighted the need to understand the biosynthetic pathway, the bottlenecks and the non-desirable product branch points. As well as understanding the advantages and disadvantages of using different organisms as the chassis for production.

1.7.5 Cytochrome p450s

The ability to functionally express p450 enzymes was essential to this work and it proved a defining feature of which organism was ultimately used for production. Cytochrome p450s are found in all domains of life, but to different degrees of rarity. As a family of enzymes they are found commonly found throughout the eukaryotes, however several species of bacteria (including *E. coli*) do not have them. P450s bind iron ions to heam, this iron ion is used (alongside O_2) to oxidise the substrate, the remaining oxygen is then reduced to H_2O. These enzymes also require the expression of their reductase, in order to function optimally. Prokaryotic and eukaryotic p450s vary in structure. Eukaryotic p450s normally have an N-terminus signal peptide, for membrane attachment, whereas prokaryotes express their p450's are soluble in the cytosol. This means that eukaryotic p450s expressed directly in bacteria, tend to be insoluble unless their tag are removed or other engineering is undertaken. One approach that has been taken is to conjoin the p450 to its cognate reductase (Leonard & Koffas 2007). However as shown in the work by Keasling *et al.* to produce artemisinic acid, over expression of the reductase can have an adverse effect of product yield (Paddon & Keasling 2014).

1.8 The Genus *Pantoea*

Following from the work of Ajikumar and colleauges (2010) and Kovacs and colleauges (2007) it was decided to attempt production of taxadiene using bacteria which already produce large quantities of GGPP for carotenoid biosynthesis, and therefore metabolism could be diverted for the production of taxadiene.

The possibility of using a species of *Pantoea* was investigated due to several characteristics. The Genus are renowned for their ability to produce carotenoids, this being a distinguishing feature of the genus. As members of the Enterobacteriaceae and being closely related to *E. coli,* they were considered to be amenable to similar lab culturing and genetic manipulation as *E. coli.* Also *Pantoea* has a history of being used as source for carotenogenic genes, such as *P. ananatis crtE* and *crtI* for golden rice and the heterologous production of lycopene in *E. coli* (Beyer *et al.,* 2002; Yoon *et al.*, 2007)

1.8.1 Taxonomy

Pantoea are a genus of the family Enterobacteriaceae, and were delineated from the *Erwinia* in the last 25 years. As a genus they have been difficult to identify biochemically, leading to successive taxonomic misplacement. *P. agglomerans* has had several other classifications including *Erwinia herbicola* and *Enterobacter agglomerans.* Therefore much of the relevant literature refers to *Erwinia herbicola*. Figure 1.5 shows the phylogeny of *Pantoea* in the family Enterobacteriaceae. The field was recently reviewed by Walterson & Stavrinides (2015), and so all of the areas of *Pantoea* research will not be mentioned here, however the points pertinent to the use of *Pantoea* as an organism for the bioproduction of Taxanes will be.

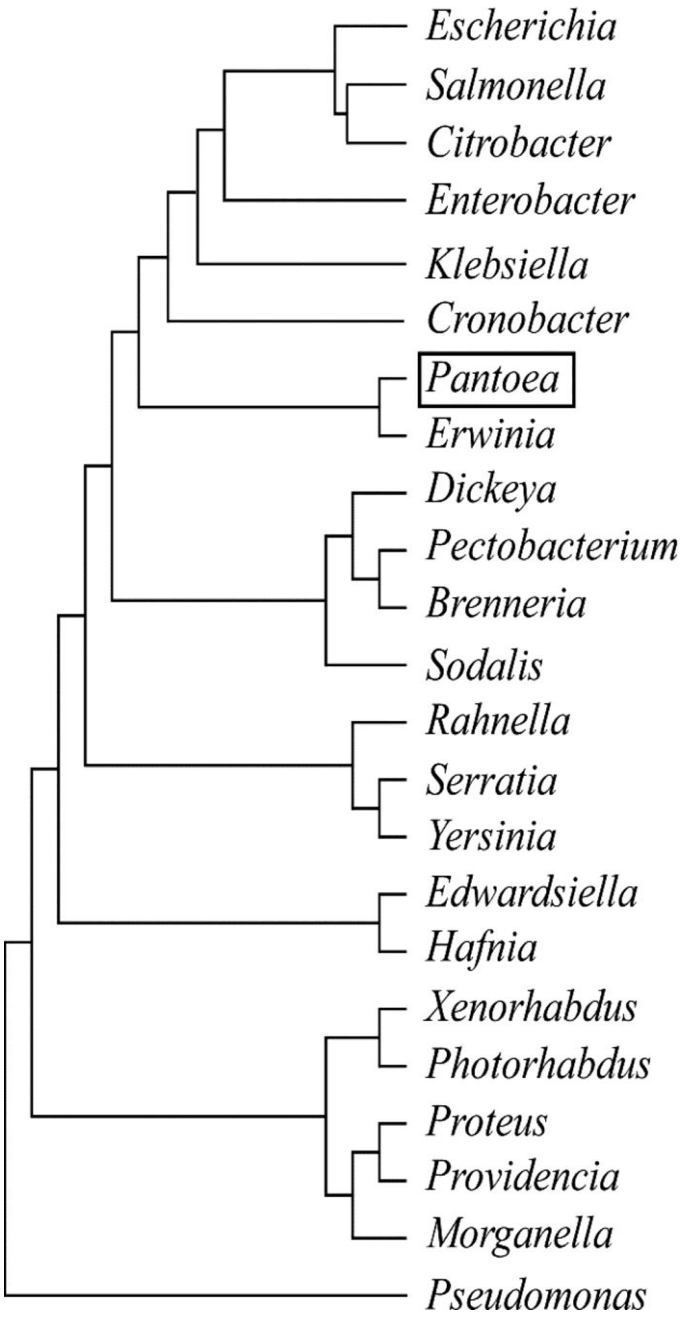


Figure 1-5 Cladogram of the Genera within the Enterobacteriaceae created using the Pathosystems Resource Integration Centre (Wattam *et al.*, 2014). Constructed using shared protein homologs derived from whole genomes of entire species groups within each genus in the. The relative position of *Pantoea* is indicated with a box. Pseudomonas (Pseudomonadaceae) was used as the root. Adapted from (Walterson & Stavrinides, 2015)

1.8.2 Genetics and physiology

Pantoea agglomerans are gram negative short rods, which produce round, regularly shaped yellow colonies when grown on solid media such as LB. *P. agglomerans* can be cultured at a wide range of temperatures (between -6°C to 42°C) and pH (4.5-8.6) (Costa *et al.,* 2002), with the optimum conditions for growth being 30°C at pH 7.0 (Kim *et al.*, 2012).

1.8.3 Environments

The name *Pantoea* comes from the Greek 'Pantos', meaning 'every' as *Pantoea* species can be isolated from a diverse array of habitats with examples found in soils and water, as well as being associated with plants. Many species have also been linked to a variety of animal species, including insects and humans. Several plant isolated species are considered pathogens, whilst several species have been shown to be beneficial to the growth of plants (Chalupowicz, *et al.,* 2009).

Within the populations of *Pantoea* isolated from humans, there is debate as to their role in infection, as to and whether they are pathogenic or just versatile opportunists. Clinical isolates of *Pantoea* tend to be identified as *P. agglomerans* by biochemical testing such as API 20E. In clinical settings ranging from catheter related urinary tract infections to infected lacerations (Walterson & Stavrinides, 2015). Several clinical isolates were investigated on a genetic level to better identify their true taxonomy. This study was able to remove a large proportion of the clinical isolates from *P. agglomerans* and attribute them to the correct species of *Pantoea*, or in some cases other genera completely (Rezzonico *et al.,*2009). However, this method was not able to separate plant associated and human associated isolates indicating that clinical isolates can colonise plant hosts and that environmental isolates can colonise humans (Nadarasah & Stavrinides, 2014).

As well as many other animals, many insect species have been identified which have a symbiotic relationship. It has been shown that axenic (devoid of bacteria) populations of locust (*Schistocerca gregaria*) do not produce the pheromone

precursor guaiacol and also produced smaller amounts of phenol. It was then shown that certain gut biota were able to produce guaiacol from axenic faecal pellets, one of these being *Pantoea agglomerans* (Dillon & Vennard 1999). This could mean that *P. agglomerans* is capable of producing potentially useful amounts of specialised biochemicals.

1.8.4 The Large *Pantoea* Plasmid (LPP-1)

The large *Pantoea* Plasmid-1 (LPP-1) contributes greatly to the diversity seen within *Pantoea*. LPP-1 is an approximately 500 kilobase (Kb) plasmid with large variation between and within species of *Pantoea*. It hosts several carbohydrate, amino acid and organic acid utilisation genes, secretion systems and resistance genes to antimicrobial agents and arsenic. This makes it useful for the adaption to new niches (De Maayer *et al.*, 2012). Three groups of LPP-1 have been identified with *P. agglomerans* having type I. Of the 20 genomes analysed, 46 coding sequences were found to be conserved among all LPP-1, the zeaxanthin biosynthetic pathway (*crtEYIBZ*) was amongst these conserved genes, with the further modification of zeaxanthin to zeaxanthin diglucoside encoded by *crtX* absent from group III. Other core functions encoded on LPP-1 include the biosynthesis of thiamine. Thiamine is an essential cofactor, having this vital cellular function encoded on LPP-1 presumably helps to maintain selection of LPP-1 after cell division. The carotenoids biosynthetic pathway clearly also benefits *Pantoea*, most probably as a defence against oxidative stress and UV damage (Dussault *et al.,* 2008) (Mohammadi, *et al.,* 2012). LPP-1 clearly plays a crucial role in the ability of *Pantoea* to live in a variety of habitats. Figure 1.6 shows a genome diagram of the group I LPP-1 adapted from (De Maayer *et al.*, 2012).

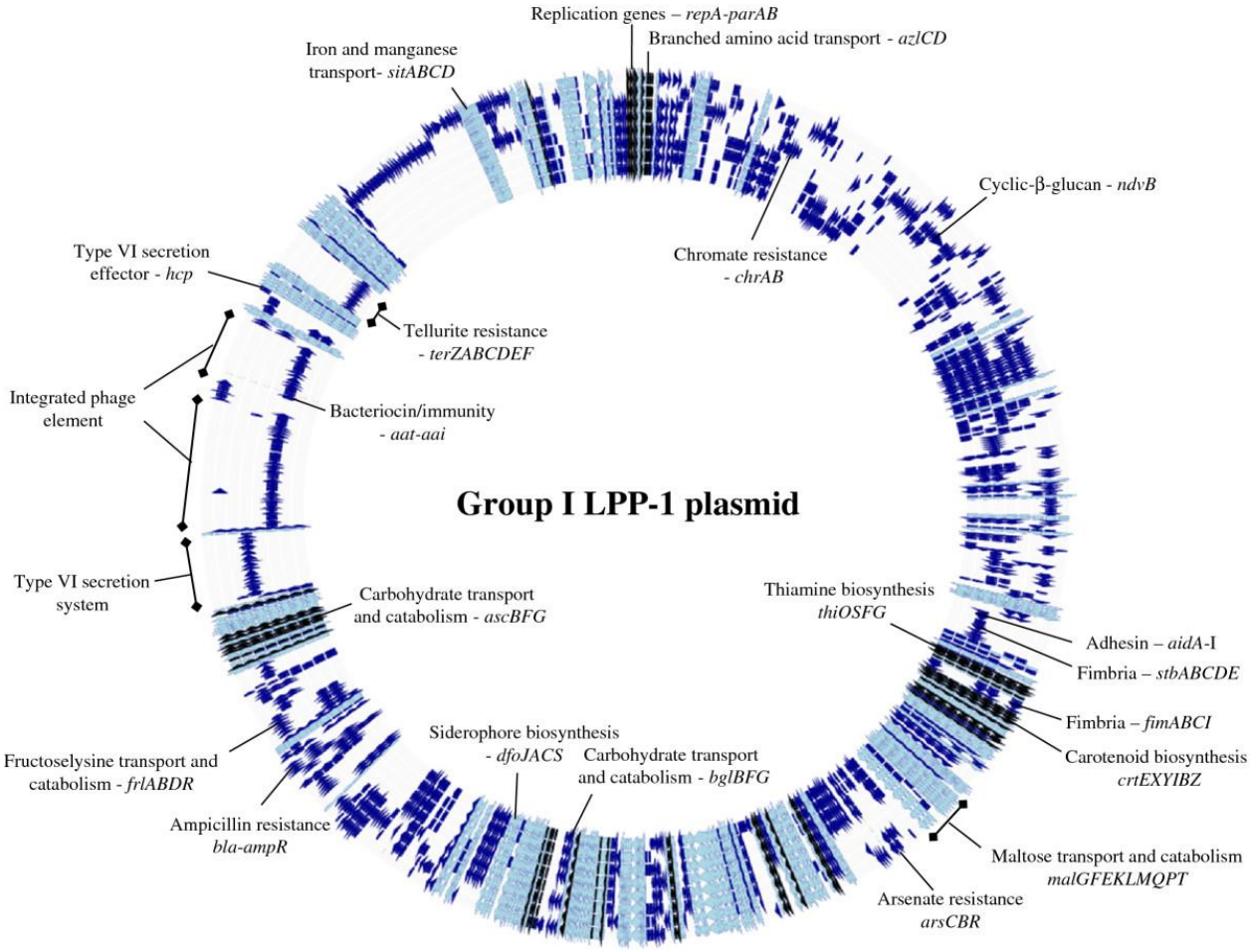


Figure 1-6 Genome diagram of the Group I LPP-1 pan-genome. Diagram showing the major phenotypes encoded on the Group I LPP-1 pan-genome. The circles represent the LPP-1 genomes of 8 *Pantoea* strains shown to belong to group I. CDS in the flexible portion of Group I pan-genome are coloured dark blue, while those conserved among all Group I LPP-1 genomes are coloured light blue. Black arrows indicate those CDS conserved among all 20 LPP-1 genomes screened by De Maayer *et al.* . Adapted from (De Maayer *et al.*, 2012).

1.8.5 Carotenoid Biosynthesis

As mentioned in the previous section *P. agglomerans* is able to synthesize zeaxanthin, a yellow tetraterpenoid via the *crtEXYIBZ* operon. Figure 1.7 summarises zeaxanthin biosynthesis which stems from the diterpenoid biosynthesis pathway. Two molecules of GGPP (produced from FPP by CrtE) are combined to produce phytoene, catalysed by phytoene synthase CrtB. Phytoene is then desaturated via CrtI to produce the red pigment lycopene. Lycopene is then cyclised at both termini

by lycopene cyclase, CrtY, producing β-carotene. CrtZ then facilitates the hydroxylation of β-carotene to produce zeaxanthin. As mentioned previously not all *Pantoea* have a copy of *crt*X, but in those that do CrtX glycosylates zeaxanthin (Misawa *et al.*, 1990). As well as the addition of *crtX* in some carotenoid operons there are also examples of *Pantoea* with *idi* (IPP isomerase) in the operon (*crtE-idi-crtXYIBZ*). These were found to produce fivefold more pigment when cloned into *E. coli* than operons without *id* (Sedkova & Tao, 2005).

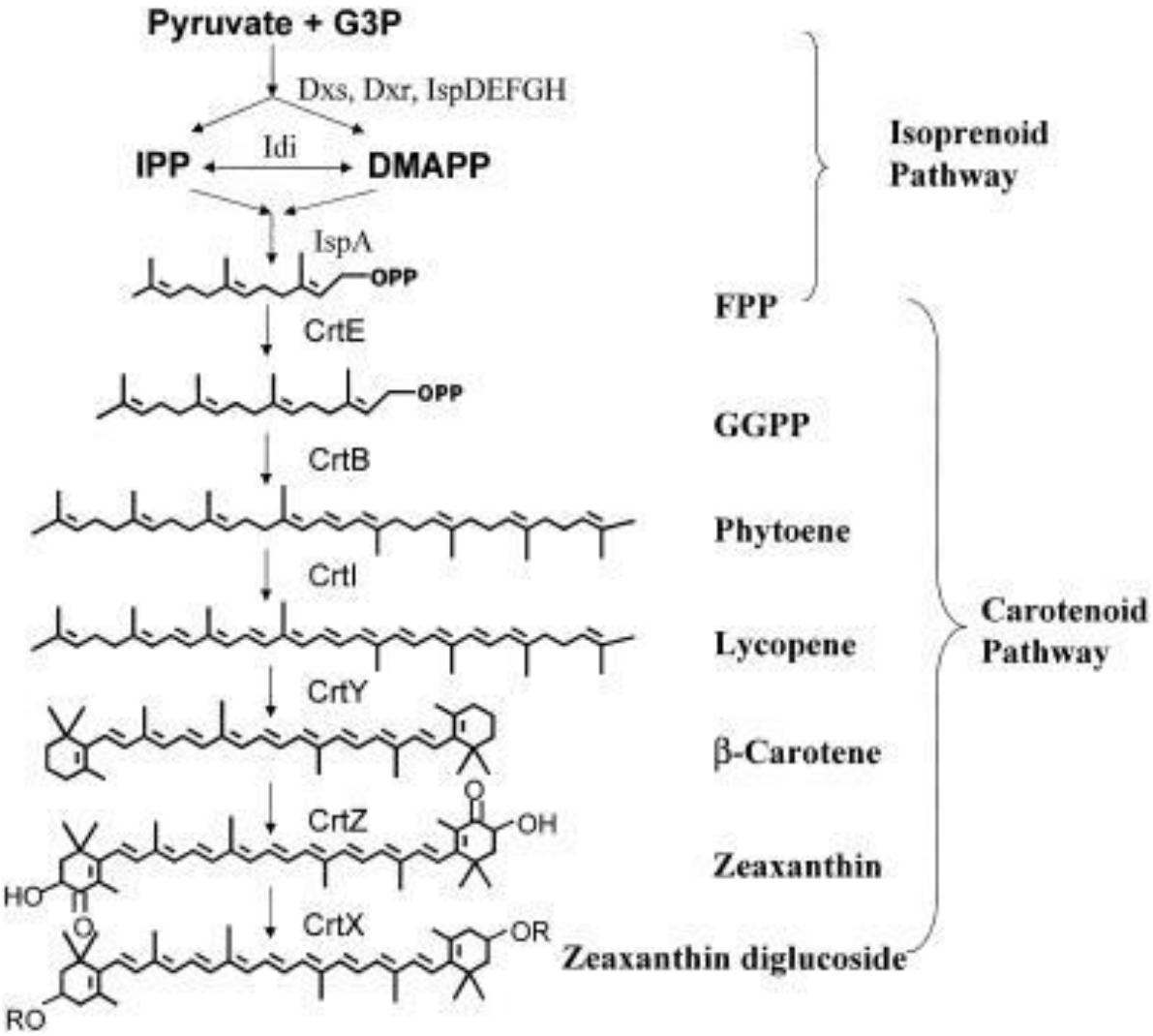


Figure 1-7 Zeaxanthin biosynthesis idi, isopentenyl pyrophosphate isomerase; IspA, farnesyl pyrophosphate synthase; CrtE, geranylgeranyl pyrophosphate synthase;

CrtB, phytoene synthase; CrtI, phytoene desaturase; CrtY, lycopene β-cyclase; CrtZ, 3,(3')-β-ionone hydroxylase; CrtX, zeaxanthin glucosyl transferase. Adapted from (Sedkova & Tao, 2005).

1.8.6 Biotechnology applications

Research involving *Pantoea* and its potential biotechnology applications covers a range of industries, with potential applications in food production, potential symbionts for plants and immunomodulation for helping overcoming cancer (Hara *et al.*, 2012; Nakata, *et al.,* 2011; Walterson & Stavrinides, 2015). A particularly novel use for *P. agglomerans* has been explored which exploit its association with the insect gut. *P. agglomerans* has a potential to be used as a vector for the delivery of anti-parasitic agents *in vivo*, which could be used to prevent the spread of diseases such as malaria. Wang & Ghosh (2012) identified the midgut of the anopheles mosquito as a site where the parasite goes through its most vulnerable developmental stages. A strain of *P. agglomerans* was produced containing the *E. coli* haemolysin secretion system giving it the ability to secrete a variety of anti-malaria effector molecules into the midgut. Bacterial numbers rapidly increase after a blood meal, and so would anti-malarials. This is a novel strategy which managed to reduce parasite development by up to 95% for both the human and mouse model parasite, blocking transmission (Wang & Ghosh, 2012).

Despite being a rather obscure organism, it is in fact being increasingly investigated for its use in biotechnology, with several companies having commercial products using *Pantoea* based technology. Several species of *Pantoea* have been shown to reduce the instances of certain diseases in fruit bearing trees such as apple and pear. Bloomtime™ produces several commercial sprays for the prevention of fire blight, which is caused by *Erwinia amylovora*. The species of *Pantoea* used has been shown to produce an antibiotic product effective against *E. amylovora*. *Pantoea* acts to exclude *E. amylovora* from the phytosphere thus not allowing it to infect the trees. This is preferable to the previous method of control, which was spray with antibiotics such as streptomycin (Walterson *et al.* 2014, Braun-Kiewnick *et al.* 2012).

Ajinimoto, one of Japans biggest mono sodium glutamate manufacturers has recently switched the bacteria it uses for fermentation for the production of L-glutamic acid from *Corynebacterium glutamicum* to *Pantoea ananatis* AJ13355. This strain of *P. ananatis* was found to be able to grow in low pH conditions and is resistant to saturating concentrations of L-glutamic acid. Whilst this strain has obvious application in the food and flavour industry, it has great potentially to be used for production of other chemicals using amino acids as a starting point of the synthesis (Hara *et al.*, 2012).

From the literature there are several potential advantages of using *P. agglomerans* for the production of taxadiene and other precursors to Taxol.

- Fast growing organism which can be grown well between 15°C to 40°C
- Already grown commercially, so it can be scaled up effectively.
- Genetically tractable, with a large array of tools, characterised in *E. coli*
- Potentially broader utilisation of carbon sources with potentially greater utilisation of more complex feedstocks than *E. coli.*
- Large natural diversity which can be used to screen for specific functions.
- Optimised pathways for the production of high value carotenoids such as zeaxanthin, which can be exploited to produces other products.

1.9 Project aims and rationale

This study aimed to investigate the possibilities of using *Pantoea* as a chassis organism for the production of high value terpenoids. The aim was to produce the already established intermediates of the Taxol biosynthesis pathway: Taxadiene, Taxadiene-5α-ol, Taxadiene-5α-yl acetate, Taxadiene-5α, 10β-diol monoacetate, and attempt to produce the novel taxane Taxadiene-10β-ol. *Pantoea* was selected because of its biotechnological potential as described previously in this chapter, with the main advantage being the high levels of zeaxanthin it naturally produces.

Pantoea would be transformed with genes from *Taxus cuspidata* to express Taxadiene synthase (TXS), 5α hydroxylase and taxadien-5alpha-ol O-acetyltransferase (5α acetyltransferase and 5α hydroxylase are referred to as one unit, 5α) and 10β Hydroxylase (10β). The proposed first, second and third step, respectively, of Taxol biosynthesis (Figure 1-4).

To achieve these aims several objectives were identified:

-Evaluate the potential of *P. agglomerans* as a host organism for terpenoid production. Whilst developing tools and materials for the transformation of *P. agglomerans* and the expression of terpenoid biosynthetic pathways.

-Use these tools to characterise the behaviour of promoters and reporter genes in *P. agglomerans*.

-Develop methods to edit the genome of *P. agglomerans*

- Clone the first 3 steps of Taxol biosynthesis, preferably in a modular operon. Then express TXS in *P. agglomerans* and then co-express TXS with the p450 hydroxylases to produce functionalised taxanes.

Chapter 2 MATERIALS AND METHODS

2.1 Bacterial Strains

Organism	Genotype	Source
E. coli DH5	F– Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *sup*E44 λ– *thi*-1 *gyr*A96 *rel*A1	(Taylor *et al.,* 1993)
E. coli Top10	F– *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara leu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	InvitrogenTM
P. agglomerans	Wild isolate	Dr Rupert Fray, University of Nottingham

Table 2 Bacterial strains

2.2 Plasmids

Name	Origin of replication	Resistance markers	Promoter	Insert	Chapter/ reference
pBMH-AB	ColE1	*kanR*	T7	*txs* and 3/4 10β hydroxylase	Chapter 5
pBMH-CDE	ColE1	*KanR*	T7	~1/4 of 10β starting at a *Sph*I site, then contains 5α hydroxylase and acetylase	Chapter 5

pSU9	p15	*cmR*	P_{LAC}	-	
pSU-TXS	p15	*cmR*	P_{LAC}	*txs* and a small portion of 10β hydroxylase (*Pst*I *Sma*I)	Chapter 6
pGRG36	pSC101	*bla* (*ampR*)	-	-	McKenzie & Craig (2006)
pGRG36T	pSC101	*tetR*	-	-	Chapter 4
pGRG36T- L-TS	pSC101	*tetR*	-	Lac *txs* (*Swa*I/*Pvu*II - *Xma*I)	Chapter 4
pBT-GFP	ColE1	*tetR*	P_{BAD}	*gfp*	Chapter 3
pBT-TXS	ColE1	*tetR*	P_{BAD}	*txs*	Chapter 5
pBT-5a	ColE1	*tetR*	P_{BAD}	5α hydroxylase and acetylase	Chapter 5
pBT-10B	ColE1	*tetR*	P_{BAD}	10β Hydroxylase	Chapter 5
pBT-5a10B	ColE1	*tetR*	P_{BAD}	5α hydroxylase, acetylase and 10β Hydroxylase	Chapter 5
pLT1	pSC101	*tetR*	-	-	Chapter 4
pDEL	pSC101	*tetR*	-	*sacB*	Chapter 4
pLBT	pSC101	*tetR*	P_{BAD}	P_{BAD} *gfp*	Chapter 7
pLBT-TXS	pSC101	*tetR*	P_{BAD}	*txs*	Chapter 7
pLBT-TXS-10B	pSC101	*tetR*	P_{BAD}	*txs* and 10Beta Hydroxylase	Chapter 7
pLBT-TXS-5a	pSC101	*tetR*	P_{BAD}	*txs* and 5alpha hydroxylase, acetylase and 10β Hydroxylase	Chapter 7
pLBT-TXS-5a-10B	pSC101	*tetR*	P_{BAD}	*txs*, 5α hydroxylase, acetylase	Chapter 7

				and 10β Hydroxylase	
pBT-CrtE	ColE1	*tetR*	P_{BAD}	*crtE*	Chapter 3
pBT-CrtIB	ColE1	*tetR*	P_{BAD}	*crtIB*	Chapter 3
pBT-CrtEIB	ColE1	*tetR*	P_{BAD}	*crtEIB*	Chapter 3
pBT-Idi	ColE1	*tetR*	P_{BAD}	*idi*	Chapter 3
pBT-CrtE-Idi	ColE1	*tetR*	P_{BAD}	*crtE-idi*	Chapter 3
pBT-CrtE-Idi-IB	ColE1	*tetR*	P_{BAD}	*crtE-idi-CrtIB*	Chapter 3
pSU-GFP-i	p15	*cmR*	P_{LAC}	*gfp* (*Not*I *Xba*I- *gfp*- *Spe*I *Xho*I)	Chapter 3
pSU-CrtEIB	p15	*cmR*	P_{LAC}	*crtEIB*	Chapter 3
pSU-CrtE-Idi-IB	p15	*cmR*	P_{LAC}	*crtE-idi-crtIB*	Chapter 3
pBT-dsCrtZ	ColE1	*tetR*	P_{BAD}	750bp downstream region of *crtZ*	Chapter 4
pBT-CrtE-dsCrtZ	ColE1	*tetR*	P_{BAD}	*crtE* and downstream region of *crtZ*	Chapter 4
pDEL-CrtE-dsCrtZ	pSC101	*tetR*	-	*crtE* and downstream region of *crtZ*	Chapter 4
pBT-LUX	ColE1	*tetR*	P_{BAD}	*luxABCDE*	Chapter 3
pCT-GFP	ColE1	*tetR*	P_{CRT}	P_{CRT} from *P. agglomerans*	Chapter 3
pCT-LUX	ColE1	*tetR*	P_{CRT}		Chapter 3
pCT-EIB	ColE1	*tetR*	P_{CRT}		Chapter 3
pCT-E-idi-IB	ColE1	*tetR*	P_{CRT}		Chapter 3
pGRG36T-KanR	pSC101	*tetR* + *kanR*	-	*kanR*	Chapter 4

Table 3 Plasmids

2.3 Chemical reagents

All general chemical reagents were analytical grade and purchased from Sigma (Missouri, USA), Fisher scientific (Massachusetts, USA) and VWR (Pennsylvania, USA) unless otherwise stated.

2.3.1 Water

Reverse osmosis water (RO water) was used for the making culturing media and buffer solutions unless stated otherwise.

For molecular biology techniques such as: PCR, restriction digest, ligation and Gibson assembly, water filtered using the Millipore system (Millipore, Massachusetts, USA) was used (Deionised, organics removed and UV treated)

2.4 Media and reagents

Culturing of Bacteria to a single colony was carried out on LB agar, supplemented with the appropriate antibiotics and other additives as appropriate.

2.4.1 Luria Bertani media

LB media was prepared by adding 10 g tryptone, 5 g yeast extract and 5 g of sodium chloride to 1 l of RO water, for LB agar 15 g/l of agar was added (1.5% w/v). LB media was sterilised by autoclaving at 120°C for 20 minutes.

For Agar plates, media was allowed to cool to below 55°C, appropriate supplements were then added, including antibiotics. Plates were then poured with approximately

20 ml of media per plate, in a laminar flow hood. Plates were then allowed to set and stored at 4°C before being used.

2.4.2 2YT media

2YT media was produced by using 16 g Tryptone, 10 g Yeast extract and 5 g NaCl per litre of media. It was then sterilised as described for Luria Bertani media.

2.4.3 Congo Red agar

LB media was prepared and supplemented with filter sterilised 0.2% glycerol and 0.08% (w/v) Congo Red dye (benzidinediazo-bis-1-naphthylamine-4-sulfonic acid) before pouring.

2.4.4 Antibiotics

Ampicillin was used at 100 µg/ml, dissolved in water and filter sterilised

Tetracycline was used at 10 µg/ml dissolved in 50% ethanol

Chloramphenicol was used at 34 µg/ml in 100% ethanol

2.5 Buffers

2.5.1 5X Isothermal Reaction Buffer (5X IRB)

25% PEG-8000, 500 nM Tris HCl pH 7.5, 50 mM $MgCl_2$, 50 mM DTT, 1 mM of each dNTP, 5 mM NAD made up with sterile ultrapure Millipore water.

2.5.2 50X TAE Buffer

50X TAE was made by combining 242 g Tris base, 57 ml of Glacial acetic acid 100 ml 0.5M EDTA in 1 l of RO water and then adjusted to pH 8.0 with the addition of concentrated sodium hydroxide.

2.6 Methods

2.6.1 Culturing

Bacterial cultures were set up using LB media (unless stated otherwise) with the addition of the appropriate antibiotics. *E. coli* was grown at 37°C 150 RPM, unless transformed with a pSC101ts replicon plasmid which needed to be maintained, then cultures were grown at 30°C. *P. agglomerans* was grown at 30°C 150 RPM.

2.6.2 Measurement of bacterial growth

Bacterial broth cultures were measured using a spectrophotometer to measure absorbance at OD600nm with a path length of 1 cm in a Pharmac biotech ultraspec 2000.

2.6.3 Enumeration of viable cells

To quantify total viable cell counts, samples were serial diluted to 10^{-6} and 100 µl was plated onto 3 replicate plates and spread. Plates were allowed to dry on the bench before being inverted and incubated. Plates were incubated at the appropriate temperature based on organism and plasmid (30°C for *P. agglomerans* and any organism transformed with a plasmid with pSC101 origin of replication, 37°C for *E. coli)*. After 18-24 hours of incubation colony forming units were counted. CFU/ml was then calculated using the equation bellow.

CFU/ mL = average number of colonies for a dilution x 100 x dilution factor

2.7 DNA

2.7.1 DNA synthesis

DNA to be synthesised was designed using DNA2.0's Gene designer software.

DNA was synthesised by Biomatik (Ontario, Canada).

2.7.2 Genomic DNA extraction

Genomic DNA was extracted using the GeneJET Genomic DNA Purification Kit (Thermo scientific, Loughborough, United Kingdom), following the manufacturer's instructions.

2.7.3 Plasmid DNA Extraction

Plasmid preparation was carried out using the Zyppy^{TM} Plasmid Miniprep Kit (Zymo research, California, USA) following the manufacturer's instruction.

2.7.4 Polymerase chain reaction

Oligonucleotide primers were synthesised by Eurofins (Ebersberg, Germany) (Table 4 and Table 5). Polymerase chain reaction (PCR) amplification were carried out using Phusion high-fidelity DNA polymerase (New England Biolabs, Massachusetts, United States) following the methods provided by the manufacturer. 20 µl or 50 µl reactions were carried out at a final concentration of 1X Phusion HF or GC Buffer (provided by manufacturer), dNTP 200 µM, forward primer 0.5 µM, reverse primer 0.5 µM, 1 unit of polymerase per 50 µl reaction. For PCR from genomic DNA 50-100 ng of DNA was used as a template, for PCR amplification from plasmid template 1 pg-10 ng of DNA was used.

Annealing temperature was calculated using New England Biolabs Tm calculator using the binding sequence of the primer. Temperature was cycled using the C1000 BioRad PCR block. Conditions were an initial 98°C for 30 seconds, then 25-35 cycles of: 98°C for 10s, calculated annealing temperature for 20s, 72°C for 15s per Kb to be amplified. This was then followed by a final extension at 72°C for 5 minutes. PCR products were separated using agarose gel electrophoresis.

2.7.5 Primers

Name	5' – 3' sequence	Restriction enzyme
pGRG tet gib f 3	CTAGAGTCGACCTGCAGGTTCTCATGTTTGACAGCTTATC ATCGATAAGC	*SbfI*
pGRG tet gib r 2	TCTGCGAGGCTGGCCGCCGCCGGCTTCCATTCA	
Ts_r_NcoI_ Xhoi	CGTTCTAAGCTCGAGCCATGGATTAATTTACACTTGGATT GGGTCG	*Xho*I *Nco*I
10B_SbfI_f	TGCCTCCCTGCAGGACATGTAGGAGGTTCTACAGATGAA TACG	*SbfI*
10B_XhoI_ r	CCGCGGCTCGAGCTAAGAGCGAGGGAAGACTTTAAGCG	*Xho*I
5A_10B_G A	ATATGTGGCGAAAGTCTGATCTAGAGGCAGGAGGTTCTA CAGATGAATACGTTCTC	
10b_GBT_x hoi_GA	GTTGCGTCGCGGTGCAT	
Ori31_f	TAGCAATTTAACTGTGATAAACTACCGCATTAAAGCTTAT CAACACTCAACCCTATCTCGGG	
Ori12_r	CCTTTACTCATAATTATTTCCTCCTCCTGCAGGGCGGCCGC AAAAAAACGGGTATGGAGAAACAG	
Gfp2_f	GCGGCCGCCCTGCAGGAGGAGGAAATAATTATGAGTAA AGGCG	
Gfp2_r	CTCGAGTTTGTATAGTTCATCCATGCCATGTGTAATCCC	
Tet23_f	TGGGATTACACATGGCATGGATGAACTATACAAACTCGA GTCAGGTCGAGGTGGCCCG	
Tet3_r	ATAAGCTTTAATGCGGTAGTTTATCACAG	
pSU9_INT_ R	AGGTGCACTCGAGTTTTTGCGGCCGCGTAATCATGGTCA TAGCTGTTTCC	*Xho*I *Not*I
pSU9_INT_ F	AAAAACTCGAGACTGGCCGTCGTTTTACAAC	*Xho*I
int_crtE_f	GGCATGGCGGCCGCATCTAGAAGGAGGAAAAAAAAATG ACGGTCTGTG	*Not*I *Xba*I

int_crtE_r	CGCATTCTCGAGAACTAGTGCTTTTAACTGACGGCAGCG AG	*Xho*I *Spe*I
CrtZ DS_F	GGCATGGCGGCCGCATCTAGAAGAGGGATTTGCCAGG	*Not*I *Xba*I
CrtZ DS_R	CGCATTCTCGAGAACTAGTTAATCCCACCGTCTGC	*Xho*I *Spe*I
CrtI_Int_F_ 2	GGCATGGCGGCCGCATCTAGACAACAGGCGATGCAATG	*Not*I *Xba*I
CrtB_Int_R	CGCATTCTCGAGAACTAGTTTAAACGGGGCGCTG	*Xho*I *Spe*I
IDI_F	GGCATGGCGGCCGCATCTAGAAATAATCTCAAGGAAACG CC	*Not*I *Xba*I
IDI_R	CGCATTCTCGAGAACTAGTTTTATGCGGTAACTTGCG	*Xho*I *Spe*I
pCRT_ClaI_ f	GCACCCATCGATGCTGGCAAGACACCCTAAACTC	*Cla*I

Table 4 Oligonucleotides used for cloning insertional sequences and plasmids

Name	5’ – 3’ sequence
M13f	GTTTTCCCAGTCACGAC
M13r	AACAGCTATGACCATG
pGBT_ins_f	CCACATTGATTATTTGCACG
pGBT_ins_r	GCATGGATTGTAGGCGC
TS us r	GATATCATCCACAATGGTACTGG
TS ds F	CGGAATCGCCAATGAGG
Tn*7*l usTS f	GGGAACTGGGTGTAGC
Tn*7*R dsTS r	GCAAATTGCCCGTCGTATT
TS middle f	CTATTCAAAGAGATTGAGTACGTGG
TS middle r	GCCAAGGGTACTCCACC
TS US r	TTTCGGAAACCACCTTACTCGT
Ts DS f	GGTTTACATCGACCCAATCCAA
5a Ac trans DS f	GTCCAAAATGAAGTCTTTCAAAATTGA
10b Hdxyl DS f	CCATTGCCAAACAACGGCTT
5a_OH_dsr	CCATTGCCAAACAACGGCTT
5a_AC_usf	ATGGAGAAAACGGACTTGCATG
TS_SYN_F	TCACACATGATGAGCAGTTCG
TS_SYN_R	CCATCGGGATGTCGTTACTC
ts_syn_int_f	CTCTTTAGGGCCTCAGACCTGG
ts_syn_int_r	GCTCATGTGCCAAGGATATTCAACG
5aOh_US_r	TCACCTCATTGAATTTAGCAACG
ts_DS_f	GGTTTACATCGACCCAATCCAA
5aAc_usf	ATGGAGAAAACGGACTTGCATG
5aAc_us_r	CAATTGAACTCAGCTGCAAAG
10B_us_r	CGTCAGAAAAAGGATGCTCTG

5Ac_Ds_f	TCCTCAAGCCGTCGAAAAATAAAC
pDEL_r	GCAAATTGCCCGTCGTATTAAAG

Table 5 Oligonucleotides used for verification of DNA sequences

2.7.6 Restriction digestion

Enzymes were purchased from New England Biolabs. Digests were carried out following the manufacturer's instructions. A typical restriction digest would contain between 0.1- 1 µg of DNA, 1-2 U restriction endonuclease and 1X buffer (supplied by manufacturer). Final volumes were between 5-30 µl with sterile water being used to make up the final volume. Digests were incubated for at least 2 hours at the appropriate temperature.

2.7.7 Agarose gel electrophoresis

Agarose gels were prepared by dissolving between 0.6-1.2% (w/v) agarose in 1X TAE buffer. After the solution cooled 5µl of ethidium bromide was added and the gel was poured and allowed to set. Ladder from Promega (Wisconsin, USA) was used, Figure 2.1 shows the DNA band pattern. Before loading on to the gel DNA was mixed with loading dye provided by Promega. The gel was submerged in 1X TAE and run at a voltage of 80V DC through the gel for approximately 1 hour, or until the loading dye markers had sufficiently separated.

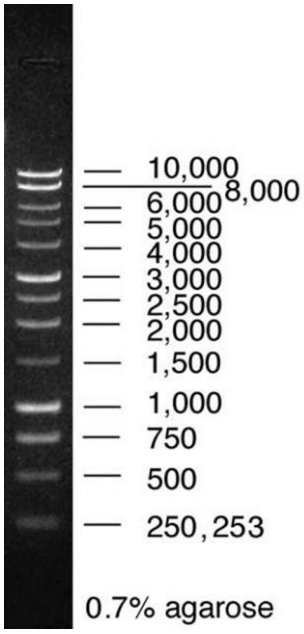


Figure 2-1 Promega 1Kb ladder

2.7.8 DNA purification

DNA was extracted from agarose gel following electrophoresis by first excising the band with a scalpel, avoiding touching any other DNA bands. This Agarose block was then dissolved and the DNA removed using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research)

2.7.9 Ligation of DNA

DNA ligation was performed using T4 DNA ligase (Invitrogen, California, USA). Purified DNA inserts were ligated into purified DNA vector at a ratio of 3:1 when ligating a single insert and 10:1 when using 2 DNA inserts. Ligations were carried out in 10µl volume with 1X T4 DNA ligase buffer with 1U T4 DNA ligase, with the remaining volume made up with ultrapure Millipore water.

2.7.10 One step isothermal DNA assembly reaction (Gibson Assembly)

Gibson assembly was designed to be a method of stitching DNA fragments together without the use of restriction enzymes and ligases. DNA fragments are designed with a 20-40 bp overlap of sequence between fragments. Through the use of a T5 exonuclease, the 5' end of the fragment is degraded, leaving a 3' overhang. These overhangs are then able to bind other fragments due to their complementarity, DNA polymerase then fills in the gaps left by the T5 exonuclease, and a ligase seals the final nicks. This method is a powerful tool in the establishment of synthetic biology approaches to the production of large DNA constructs, such as synthetic operons or chromosomes (Gibson *et al.*, 2009).

12X reactions were made by combining 36µl 5X IRB, 3.6µl T5 exonuclease (0.2 U/µl), 18 Taq DNA ligase (40 U/µl), 2.24µl Phusion DNA Polymerase (2 U/µl) and 120.16 µl ultrapure Millipore water. These were then mixed and dispensed into 12X 15 µl aliquots and stored at -20°C until use.

DNA assembly was carried out by mixing the individual fragments to be assembled in stoichiometrically equal amounts up to a total of between 0.2–1 pmols in 5 µl of water. This was then mixed with 15 µl of the reaction mix and incubated at 50°C for 1 hour then used to transform *E. coli*.

2.8 Electrocompetent cells

Electrocompetent *E. coli* and *P. agglomerans* cells were made using the method described by Sambrook and Russell (2001). Inoculum was grown overnight in 10ml of LB broth (supplemented with the appropriate antibiotics) at the appropriate temperature for the organism (37°C *E. coli* , 30°C, *P. agglomerans*) shaking at 150 RPM. This was then used to inoculate fresh media 1:100 1 l) which was then grown to an OD 600nm of between 0.6-0.8. The culture was then chilled on ice for 15-30 minutes before centrifugation at 4°C for 10 minutes. The supernatant was discarded

and the cells were washed in a full volume of cold sterile RO water. This process was repeated twice more each time halving the volume of water used. Cells were then resuspeneded in 1 ml of 10% glycerol and aliquots of 50 µl were dispensed into sterile Eppendorf tubes and immediately frozen using liquid nitrogen. These were then stored at -80°C until required.

2.9 Electroporation

Aliquots of electrocompetent cells (section 2.8) were thawed on ice. 1-2 µl of dialysed DNA was mixed into the cells. This mix was then transferred to a cold 0.2mm electroporation cuvette (Biorad, California, USA). The cuvette was then loaded into the electroporation chamber and a pulse of 2.5kV 925 µF and 300W was delivered using the Gene Pulser Xcell™ Electroporation Systems (BioRad). 1ml of LB + 0.2% glucose was added to the cuvette, then the culture was removed and transferred to a sterile flat bottom universal. Cultures were then incubated shaking for 1 hour at 30°C for *P. agglomerans* and any organism transformed with a plasmid with pSC101 origin of replication, 37°C for *E. coli*, before plating onto selective media and incubating overnight at the appropriate temperature.

2.10 Insertion of transgenes using Tn*7*

The pGRG36 Tn*7* delivery system devised by McKenzie & Craig (2006) was used as the basis for developing a method to use the transposon Tn*7* to insert genes into the chromosome downstream of *glm*S by transposition.

After the transgene had been cloned within the Tn arms the plasmid was transformed into the intended host at 30°C (*E. coli* or *P. agglomerans*). From there the procedure outlined by McKenzie & Craig (2006) and summarised in a document available at https://media.addgene.org/data/58/96/162baa3c-af64-11e0-90fe-003048dd6500.pdf was used.

2.11 DNA sequencing

2.11.1 Sanger sequencing

2.11.1a Plasmid

Plasmid DNA was sequenced by first extracting plasmid DNA using the Zyppy Plasmid Miniprep Kit (Zymo) following the manufacturer's instructions. 5 µl of approximately 100 ng/µl and 5 µl of 3.2 pmol/µl of primer per sequencing reaction was sent to Source Bioscience Nottingham.

2.11.1b PCR product

To sequence PCR product, first it was purified by agarose gel electrophoresis using the Zymo kit, following the manufacturer's instructions. Then 5 µl at a concentration of 1ng/µl per 100 bp of PCR product and 5µl of 3.2 pmol/µl of primer per reaction was sent to Source Bioscience, Nottingham, for Sanger sequencing.

Sequencing results were analysed using Vector NTI and ApE (A plasmid Editor)

2.11.2 Genome sequencing

2.11.2a Genomic DNA extraction

Genomic DNA was extracted using the Thermo scientific geneJET genomic DNA extraction kit. The quality of the extraction was assessed using a NanoDrop® ND-1000 Spectrophotometer (Thermo Scientific) to measure the 260/280nm absorbance ratio and ensure it was above 1.5.

2.10.2b Sequencing

Genomic DNA was sequenced by the University of Nottingham's Deep seq facility using the Illumina MiSeq 300PE machine. The samples were first prepped for use on the Illumina system and then sequenced so that there were approximately 50x reads per isolate.

2.11.2c Assembly

The sequencing data was trimmed using PRINSEQ (Schmieder and Edwards, 2011), quality checkedwith FastQC (Andrews, 2010) and then assembled using SPAdes (Nurk *et al*., 2013).

2.11.2d Annotation

Gene annotation was carried by using the online genome annotation server, RAST (Aziz *et al.*, 2008).

2.12 Carotenoid measurements

2.12.1 Lycopene extraction and quantification

Lycopene was extracted using the hexane methanol method described by Bóna-Lovász *et al.* 2013. Samples under 1.5 ml were centrifuged at 15,000 RPM in a bench top centrifuge for 2 minutes, samples above 1 ml were centrifuged at 4000 g in a benchtop centrifuge. Supernatant was discarded and the cell pellet then suspended in a 0.5 volume (compared to the original culture) of methanol then 1 volume of hexane was added and mixed using a bench top vortex at full power for 5 minutes. 0.5 volume of RO water was then added, shaken by hand, and then centrifuged in the same manner as previously stated. Absorption of the hexane phase at 502nm was then measured using Pharmac biotech ultrospec 2000. If the pellet was still visibly red (or light pink) all supernatant was removed and the process repeated.

To calculate Lycopene concentration, absorption data from all washes was combined. Concentration was then calculated using the Beer Lambert equation.

$A = \varepsilon \times l \times C$

A= absorption, ε= wavelength dependant molar absorption coefficient, l = path length, C = concentration.

To calculate mass the molar mass of lycopene molar absorption coefficient of 172,000 l mol^{-1} cm^{-1} at 502 nm was used. Molecular weight = 536.873 g mol^{-1}

2.12.2 Zeaxanthin extraction

Zeaxanthin was extracted using acetone. Samples under 1.5 ml were centrifuged at 15,000 RPM in a bench top centrifuge for 2 minutes, samples above 1 ml were centrifuged at 4000 g. Supernatant was discarded and the cell pellet then suspended in acetone (scaling factor was recorded), this was then mixed using a benchtop vortex for 5 minutes. Samples were then centrifuged again as previous. The supernatant was then measured using a spectrophotometer at 452 nm using quartz cuvettes.

The concentration of zeaxanthin was determined using the beer lambert equation,

Molar absorption coefficient =133,000 lmol^{-1} cm^{-1} at 452 nm

Molecular weight= 568.88 g mol^{-1}

2.13 Taxane extraction

2.13.1 Taxadiene extraction and analysis

Cultures were subjected to a 1:1 hexane extraction in a separating funnel. These were mixed for 10 minutes and then allowed to settle for 20 minutes. The water phase was tapped off, stopping once the emulsion reached the tap. The emulsion and the hexane layer were then tapped off. This was allowed to settle and the free flowing hexane was decanted off. The remaining emulsion was then frozen at -20°C. After being frozen any free flowing hexane was decanted off and the remaining emulsion was allowed to thaw at room temp. The water phase and emulsion were then subjected to a second 1:1 hexane extraction as before.

After the two hexane washes, the water phase was mixed with equal parts chloroform in a separating funnel and mixed for 10 minutes. The mixture was allowed

to settle then the tap was partially opened and chloroform slowly left the funnel without dragging emulsion through. Once the emulsion started to move towards the tap, the tap was closed and the collected chloroform was kept, the emulsion was then run out through the tap along with a small amount of the water phase. This was then frozen at -20°C. It was then removed from the freezer, agitated to break up the partially frozen emulsion and was then added back into the separating funnel and allowed to fully thaw out. Free flowing chloroform was then tapped off. The water phase and the emulsion were then washed again with chloroform following the same procedure.

2.13.2 GC-MS

Gas chromatography mass spectrometry (GC-MS) was carried out as in Kovacs *et al.* (2007). Samples were injected in splitless mode (injector 250°C) onto a 30m x 0.25 mm silica capillary column with a 1.0 lm film thickness. The initial oven temperature was 160°C, this was increased to 300°C at 10/min after a 1 min delay (using helium as the carrier gas).

Taxadiene prepared from the transgenic tomato fruit strains described by Kovacs *et al.* (2007) was used as a control.

2.14 Dry cell weight

Dry cell weight was calculated by centrifuging cell cultures at 4000 g in 50 ml falcon tubes. Supernatant was discarded and the cell pellet was then suspended in 5ml of water. This was then pipetted into pre weighed weighing boats and left in an incubator at 42°C until the mass remained the same between 2 measurements 24 hours apart, approximately a week.

2.15 Hard colony material extraction

Following the collection of hard colonies from agar plates, they were washed in a succession of solvents: water, Hexane, methanol and then chloroform. For each solvent the colonies were manually mixed and agitated with a pipette to attempt to resuspend them in the solvent, and then mixed by vortex for 3 minutes, they were then centrifuged at 15,000 RPM in benchtop centrifuge for 2 minutes, the supernatant was removed and kept for analysis and the next solvent was added to the remaining material.

2.16 Microscopy

2.16.1 Gram staining

Slides were produced from liquid culture and from agar plate bacterial colonies. In the case of liquid culture a few drops were spread on the slide, for solid media culture a small amount of material was spread over the slide in a few drops of sterile water. Slides were then allowed to air dry and fixed using a Bunsen flame. Slides were then immersed in crystal violet stain for 60 seconds, rinsed in water. The slides were flooded with Gram's iodine and left to stand for 30 seconds and then washed again. The slides were then decolourised by flooding them in 95% ethyl alcohol for 30 seconds, the slides were then washed in water. Finally the slides were flooded with 0.1% basic fuchsin solution for 30 seconds then washed in water and allowed to dry. Slides were examined under oil immersion.

2.16.2 Congo red stain

Hard material from hard colony solvent extraction or a fresh hard colony were first Placed onto a slide with a drop of water. The slide was then pressed against another slide and moved around to disrupt the material and spread it. Both slides were then flooded with 0.01% Congo Red stain, excess stain was allowed to run off and the

slides left to air dry. The slides were then flooded with acid alcohol for 20 seconds, excess poured off and the left to air dry. Slides were examined under oil immersion.

2.16.3 CCD imaging

Visualisation of bioluminescence was carried out using the Biospace system (charged coupled device imaging equipment). Agar plates with colonies to be screened were placed into the machine and the level of light production was measured for 2 minutes. This data was then visualised using Biospace visualisation software, which enabled identification of individual colonies that were bioluminescent.

2.17 Statistical methods

Graphpad prism software was used for statistical analysis. The T test was used to compare production of products from two populations (be it lycopene production or cell growth).

For all graphs (unless otherwise stated) the mean value of three replicates was plotted alongside error bars denoting +/- the standard deviation.

Chapter 3 CHARACTERISATIONS OF PARTS AND PROCEDURES FOR THE TRANSFORMATION OF *P. AGGLOMERANS*, AND THE PRODUCTION OF LYCOPENE

3.1 Introduction

To be able to transform *Pantoea* with plasmid DNA, three things were required: a procedure to introduce DNA into *Pantoea*, functional origins of replication and antibiotic selection markers to maintain the plasmids.

Having a range of origins of replication, with different compatibility groups allows the introduction and maintenance of more than one plasmid at once allowing for a modular approach to investigating how best to express a biosynthetic pathway. Having access to a range of origins of replication with different copy numbers also adds to the ability to fine tune the expression levels required to produce large quantities of products as demonstrated in (Ajikumar *et al.*, 2010).

This section deals with the establishment of a set of tools for the transformation of *P. agglomerans*, and then the characterisation of promoters and reporter genes.

3.2 Plasmid selection and transformation

3.2.1 Antibiotic sensitivity

The antibiotic susceptibility of *P. agglomerans* was initially tested as described in methods. Of the antibiotic cassettes that were available for use, *P. agglomerans* was resistant to ampicillin, but susceptible to kanamycin, tetracycline and chloramphenicol.

3.2.2 Transformation protocol

Transformation by electroporation was first attempted using pSU8; *P. agglomerans* cells were prepared as in (Sambrook & Russell, 2001). 1µl of pSU (~50ng/µl) plasmid prepared from *E. coli* top10 was used for the electroporation; the recovered transformants were then plated on LB chloramphenicol overnight at 30°C.

This method was successful and produced Chloramphenicol resistant colonies.

3.2.3 Plasmid origins of replication

In order to determine which plasmid origins of replication were functional in *Pantoea*, a range of plasmids with a range of origins of replication were transformed into *P. agglomerans.* Table 4 shows the plasmids used, their resistance markers and origin of replication. All of these origins of replication produced transformants. Transformed colonies were then grown overnight in culture and then plasmid prepped to extract the plasmid and confirm its integrity by restriction digest.

Plasmid name	Origin of replication	Antibiotic resistance marker	Copy number	Compatibility class	Maintained in *P. agglomerans*
pBT-GFP	ColE1	Tetracycline	~15-20	A	Yes
pSU	p15A	Chloramphenicol	~10	B	Yes
pLBT-GFP	pSC101	Tetracycline	~5	C	Yes

Table 3. Functional origins of replication in *P. agglomerans*

3.3 Assessing promoter functionality in *P. agglomerans*

Three promoters were identified as having interesting properties for use in *Pantoea*: the lactose inducible promoter P_{LAC}, the arabinose inducible promoter P_{BAD} and the native promoter of the carotenoid operon P_{CRT}. P_{LAC} was of interest because it has been well studied and characterised in *E. coli.* P_{BAD} has also been well characterised in *E.coli* but unlike P_{LAC} the induction from the addition of arabinose is 'tuneable', with higher levels of arabinose leading to a higher expression of protein (to a limit). The control of the carotenoid operon by P_{CRT} could help to elucidate why *Pantoea* is able to produce large amounts of zeaxanthin, controlling taxadiene synthase in a similar manner could help improve yield.

Initial testing of promoters for use in *Pantoea* was concerned with confirming their functionality in *Pantoea*.

The principle was to clone GFP downstream of the promoter under investigation, and then transform this plasmid into *Pantoea* and screen for levels of green fluorescence in the resulting transformants.

3.3.1 pBT-GFP

In order to put GFP under the control of the promoter P_{BAD}, pBT-GFP was created using Gibson assembly. Several components from different plasmids were used: the ColE1 origin of replication of pBAD18 was chosen as a foundation for pBT-GFP as it is compatible with pSU9. Using pBAD18 as a template also allowed for the P_{BAD} promoter to be cloned in the same reaction. The tetracycline resistance cassette *tetR* from pBR322 was chosen as the selection marker. GFP was cloned from pUC18-NotI. The nomenclature of pBT refers to the promoter ($P_{BAD:}$ B) and the resistance marker (*tetR*: T) that are present.

The three fragments were amplified by PCR using Primers Ori31_f, Ori12_r, with P_{BAD}18 as a template, Gfp2_f, Gfp2_r, with pUC18-NotI as a template and Tet23_f, Tet3_r with pBR322 as the template. The products were cleaned up by gel extraction

and then joined together using Gibson assembly. The reaction was then transformed into *E. coli* by electroporation. Transformants were selected on LB tetracycline and colonies were then patched on to LB tetracycline + 0.2% arabinose added. Colonies were screened under a 480nm light source and green colonies were selected and grown overnight for plasmid preparation. The plasmid was confirmed by restriction digest and by sequencing.

3.3.2 pCT-GFP

To ensure that a functional promoter had been cloned from *P. agglomerans* P_{CRT} was amplified from genomic DNA of *P. agglomerans* using primers pCRT_ClaI_f and pCRT_NotI_r to produce a 2Kb fragment of DNA upstream of the start of *crt*E (the first gene in the *crt* operon). This fragment was restricted with *Not*I and *Cla*I. pBT-GFP was also restricted with *Not*I and *Cla*I to remove P_{BAD}. These two fragments were then ligated and transformed into *E. coli.* Selection was on LB tetracycline at 37°C. Constructs were confirmed by restriction digest and Sanger sequencing, then named pCT-GFP (figure *3*.1). Colonies were visibly expressing GFP when observed with 480nm light. pCT was named using the same nomenclature as for pBT. The C refers to the promoter $p_{CRT.}$

3.3.3 pSU-GFP

To put GFP under the control of P_{LAC}, GFP was cloned into pSU9 to produce pSU-GFP. pSU9 was amplified by PCR using the primers pSU9_INT_R and pSU9_INT_F. This fragment was then restricted with *Not*I and *Xho*I, GFP was restricted from pBT-GFP using *Not*I and *Xho*I. These fragments were then ligated together to produce pSU-GFP. The ligation was transformed into *E. coli* and green fluorescent colonies (under 480nm light) were selected and plasmid construction was confirmed by restriction digest. pSU was not named using the previously used nomenclature, as it was a pre-existing plasmid created by

pSU-GFP was transformed into *P. agglomerans* by electroporation, colonies were then selected for on LB chloramphenicol. pBT-GFP and pCT-GFP were transformed into *P. agglomerans* by electroporation and then selected for on LB tetracycline. All strains were grown over night at 30°C. Colonies formed from all transformation reactions. For pSU-GFP and pCT-GFP green fluorescent colonies were observed under 480nm light. pBT-GFP colonies were streaked onto LB tetracycline agar plates with and without the addition of 0.2% arabinose. Colonies plated onto media containing arabinose produced visible amounts of GFP when observed under 480nm light source; the colonies streaked onto media without arabinose did not.

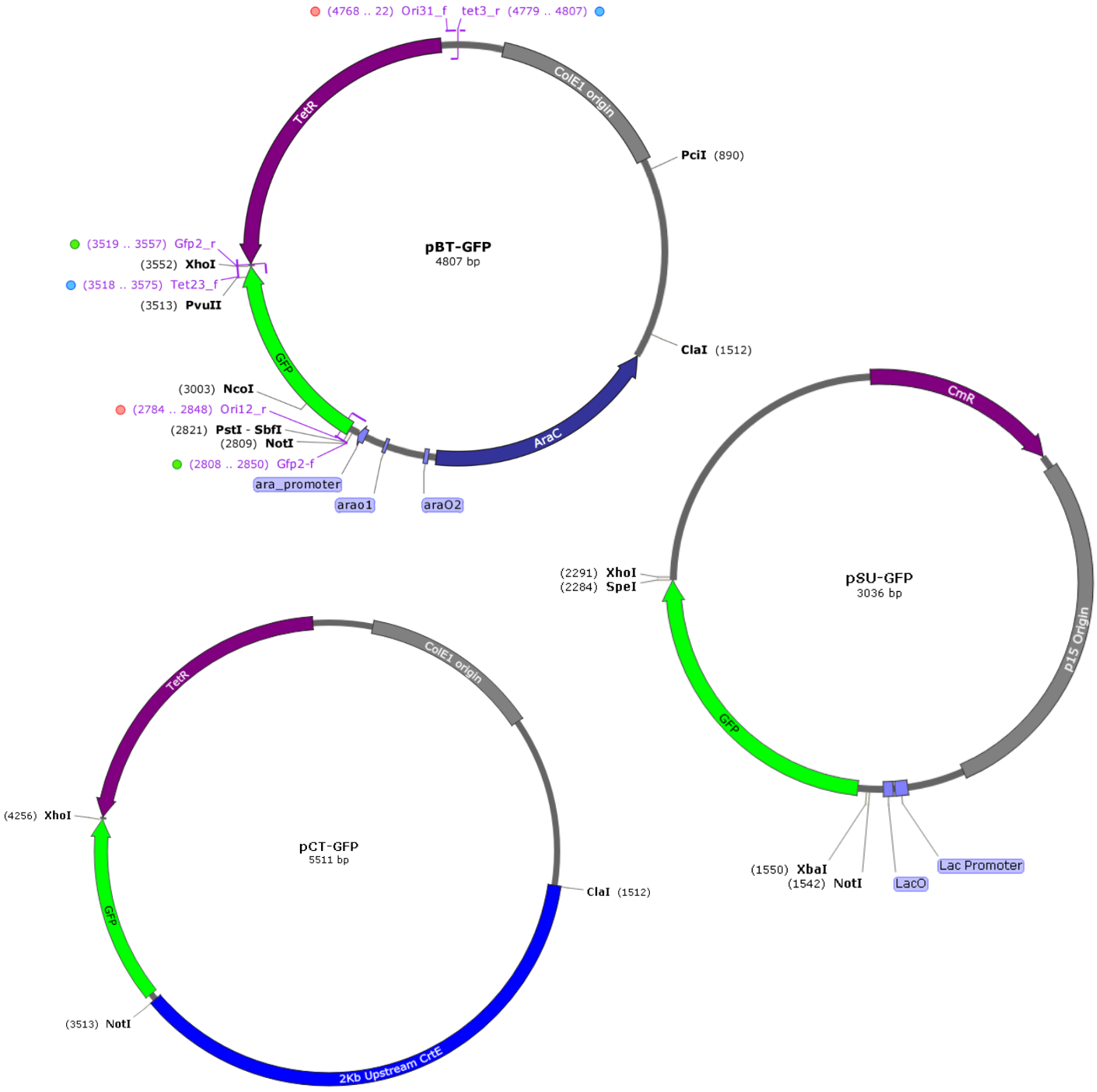


Figure 3-1 Plasmid maps of pBT-GFP, pCT-GFP and pSU-GFP

3.4 Dynamic tracking of promoter activity through use of the bioluminescent *lux* operon

To achieve a better understanding of the change in expression from promoters over time, GFP was replaced by the *lux* operon. The *lux* operon produces bioluminescence by synthesising aldehyde which is then oxidised by a luciferase to produce light. This is a metabolically active process, requiring ATP to fuel the production of aldehyde which is then reacted with $FMNH_2$ and 0_2 by luciferase to produce fatty acid, FMN, water and light. The operon can also be used as an indication of the fitness of cells. Expression of the *lux* operon could therefore be measured over a time course to give a more accurate reading of expression at that moment, compared to the expression of GFP which is a cumulative measure of protein production. The proteins of the *lux* operon also have a much shorter half-life than GFP, and so it is possible to detect when a promoter switches off.

3.4.1Replacment of the GFP cassette with the *lux* operon

To produce plasmids with the lux operon under the control of the different promoters being investigated, pBT-GFP, pCT-GFP and pSU-GFP had their GFP cassette replaced with the *lux* operon. *lux*ABCDE was amplified from pUC18-*Not*IDual using primers Lux_INT_F and Lux_INT_R, it was then restricted with *Not*I and *Xho*I. pBT-GFP was also restricted with *Not*I and *Xho*I to remove GFP. The back bone of pBT-GFP and the amplified *lux* fragment were then ligated together and transformed into *E. coli* and selected for on LB tetracycline agar at 37°C overnight. Colonies were patched onto LB tetracycline 0.2% arabinose agar and grown overnight at 37°C. These were then screened by photon detection using a cooled CCD system (Biospace labs). Colonies that were actively producing light were picked and grown in LB culture overnight; the plasmid was then extracted and profiled by restriction digest to confirm the construct was correct. This plasmid was then named pBT-LUX.

pSU-GFP and pCT-GFP and pBT-LUX were then restricted with *Not*I and *Xho*I. pCT-GFP and pSU-GFP were run on an electrophoresis gel in order to separate the backbone of the plasmid from GFP. The backbones were then extracted. pBT-LUX was separated by electrophoresis to yield the *lux*ABCDE fragment. pSU-GFP backbone and the *lux* fragment were then ligated, transformed and selected for on LB chloramphenicol agar at 37° overnight. The same procedure was carried out for the pCT-GFP backbone, but was selected for on LB tetracycline agar. Colonies were then screened for light producing colonies using a cooled CCD system (Biospace labs). The pCT backbone reaction produced a large number of colonies that produced light; these were then confirmed by restriction digest and named pCT-LUX.

The pSU backbone reaction, which would have produced a *lux* operon driven by P_{lac}, produced fewer colonies only some of which were light after several days of incubation. These colonies struggled to grow when restreaked onto LB chloramphenicol plates or in LB chloramphenicol culture with the addition of 0.2% glucose. For this reason this construct was not used going forward, but are discussed at the end of this chapter.

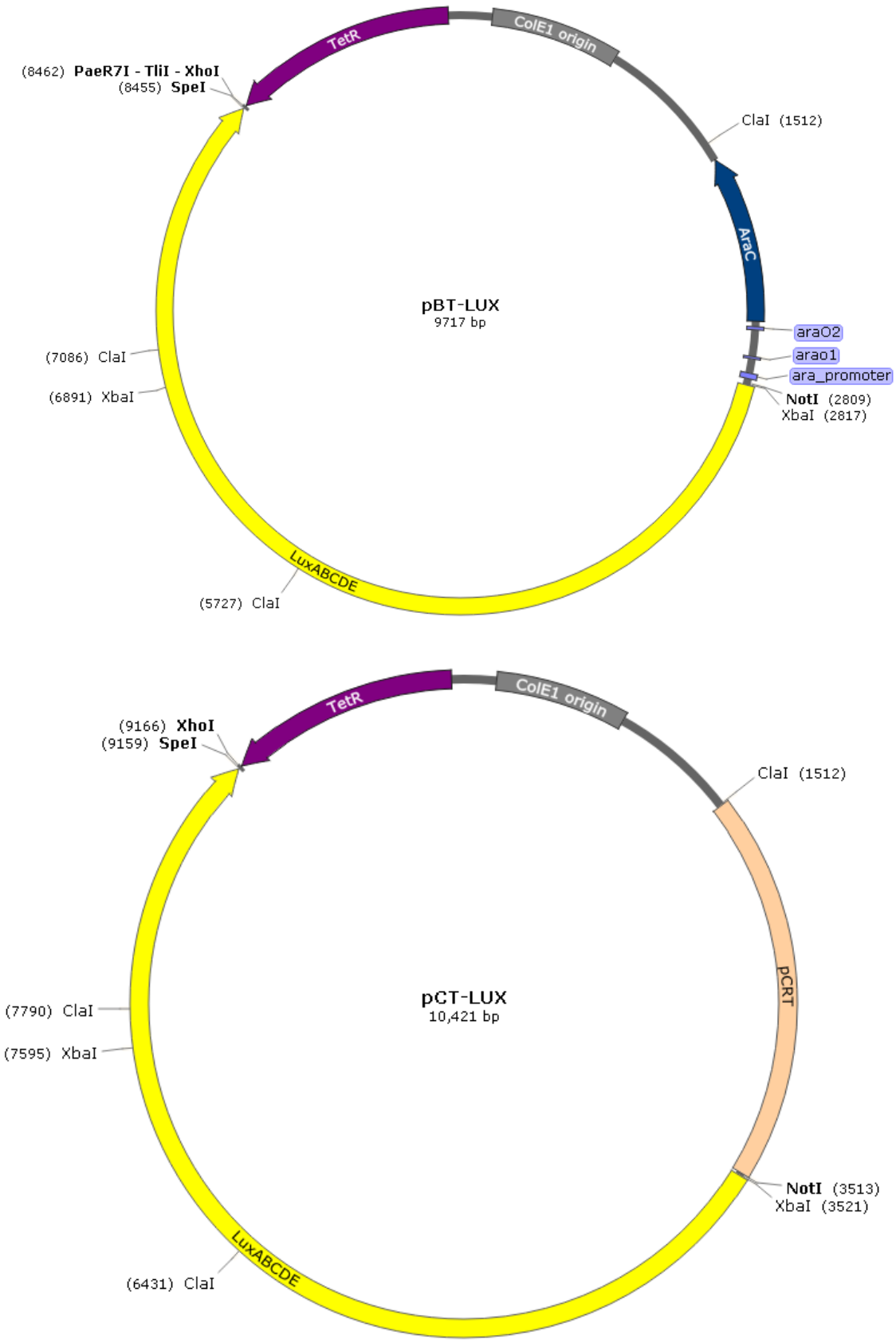


Figure 3-2 Plasmid maps of pBT-LUX and pCT-LUX

3.4.2 Expression of the *lux* operon under the control of P_{BAD} in *P. agglomerans* and *E. coli*

pBT-LUX was transformed into *P. agglomerans* ΔLPP-1. *P. agglomerans* ΔLPP-1 (described in chapter 6) is a mutant strain deficient in carotenoid biosynthesis; this was selected due to concerns that the production of zeaxanthin potentially could absorb light produced by the *lux* operon (490nm). Colonies were selected for on LB tetracycline and incubated at 30°C overnight. Colonies were then patched onto LB tetracycline with arabinose and incubated at 30°C overnight; these were then screened using a cooled CCD system (Biospace labs) for bioluminescence. One bioluminescent colony was selected and grown in LB tetracycline culture overnight at 30°C, *E. coli* harbouring pBT-LUX was grown in LB tetracycline culture at 37°C overnight as well. These two overnight cultures were then diluted 1:100 into fresh media and incubated at the same temperature used for the previous overnight incubation until an OD600 of 0.6 was achieved (around 4-5 hours).

A 96 well plate was then prepared with varying concentrations of glucose or arabinose to repress or induce expression of the *lux* operon from P_{BAD}. The concentrations used were 0.1% glucose (as repression) 0% glucose 0% arabinose, 0.001% arabinose, 0.01% arabinose and 0.1% arabinose. These were set up in triplicate for both species (6 wells in total), 200µl of bacterial culture was then added to each well. The plate was then incubated at 30°C in a Genios Pro plate reader (Tecan) for 20 hours. Measurements of OD600 and relative light units were taken every 15 minutes for the first hour and then every half an hour for the rest of the time course, results are shown in figure *3*.3.

It was found that both *E. coli* and *P. agglomerans* can express *lux* under the control of P_{BAD}, and this expression can be tuned. Compared to *E. coli, P. agglomerans* does not repress P_{BAD} as strongly under 0.1% glucose, but it is repressed compared to the baseline of no additional sugars. Differentiation between the expression levels in *P. agglomerans* was greater than in *E. coli* most notably for the addition of 0.001% and 0.01% arabinose. To compare the overall expression levels over the time course of

the experiment, the area under the curve (AUC) was calculated and is shown in Figure *3*.4.

Comparing the percentage increase from the baseline (no added glucose or arabinose) serves to highlight how *P. agglomerans* is able to produce a more graduated response to increasing levels of inducer. The most notable difference the *P. agglomerans* system produces is at the 0.001% arabinose induction level, which is able to provide 135% increase from the baseline compared to 1705% increase in *E. coli* (Table 4).

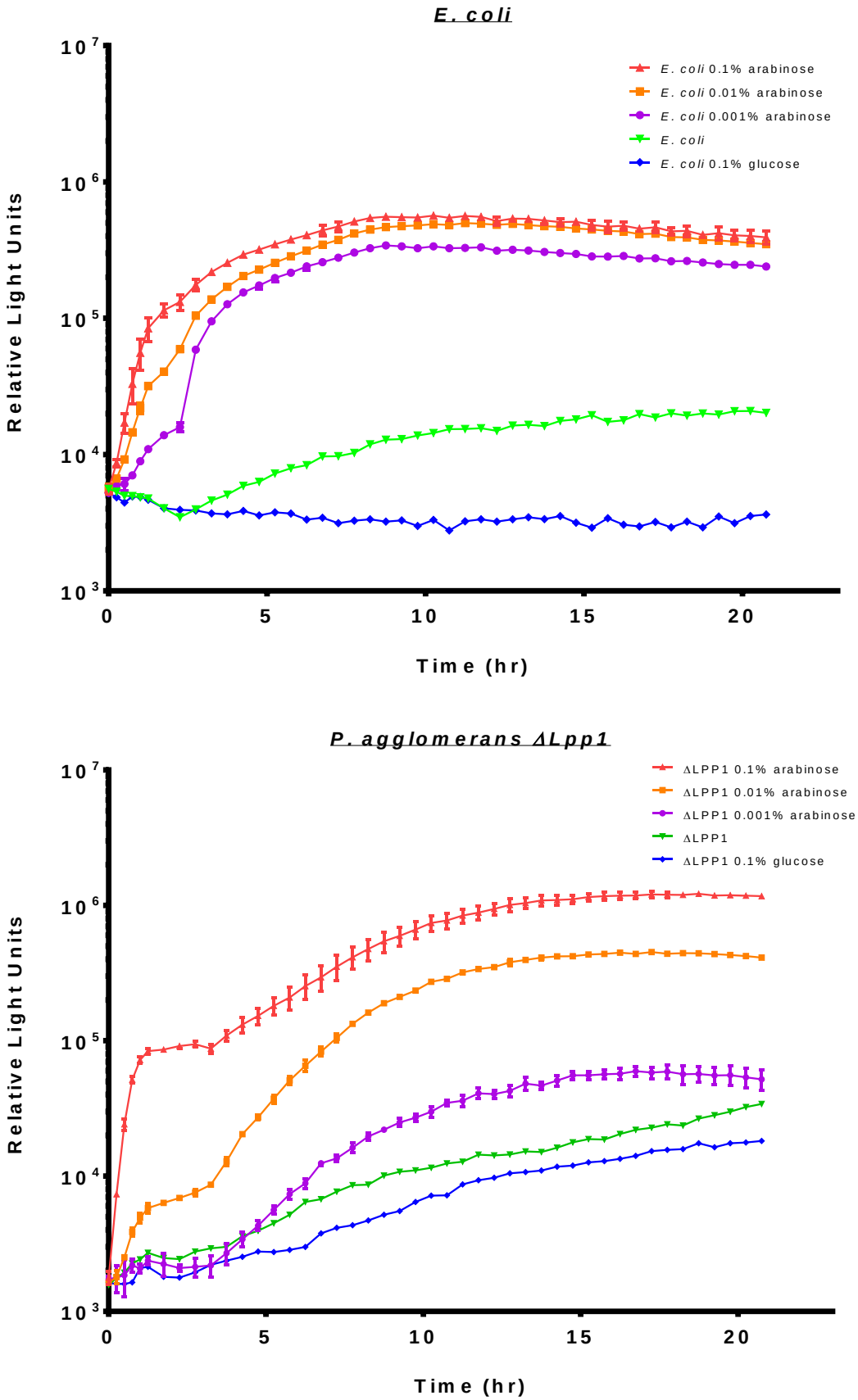


Figure 3-3 relative light production from *E. coli* and *P. agglomerans* ΔLPP-1 transformed with pBT-LUX

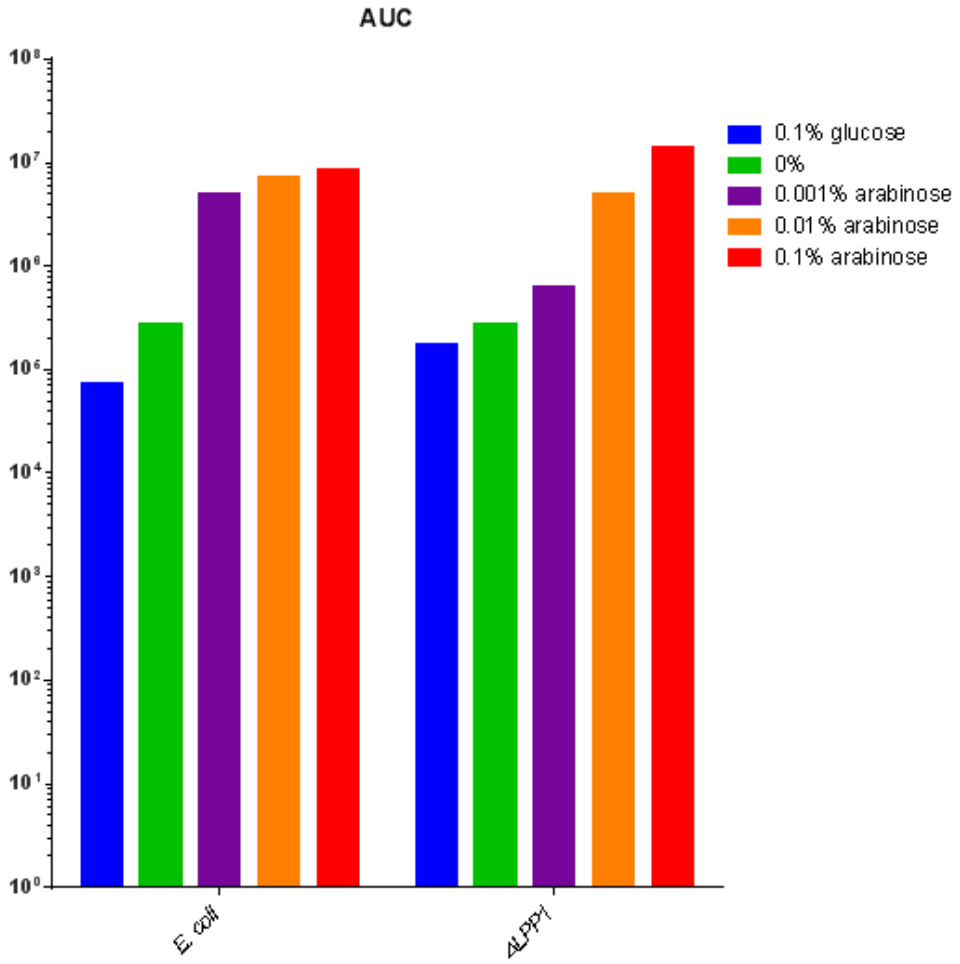


Figure 3-4 Area under the curve of the data from figure *3*.4. Area under the curve was calculated to express the total relative light unit production over the course of the whole experiment.

percent change from baseline*					
	0.1% glucose	0%*	0.001% arabinose	0.01% arabinose	0.1% arabinose
E. coli	-127	0	1705	2552	3024
ΔLPP1	-163	0	135	1766	5135

Table 6 Percent change in light production under different levels of induction in *E. coli* and *P. agglomerans* ΔLPP-1 transformed with pBT-LUX

Ultimately from a tuneable system it is desirable to be able to predictably achieve a range of expression levels in order to find the perfect level of expression for the task at hand. From the range of induction levels used in this experiment*, P. agglomerans* with P_{BAD} has been shown to provide a greater range of expression levels, compared with the same system in *E. coli* which shows finer differentiation over a narrower range of expression levels. Further work needs to be carried out to determine if there is a threshold level of arabinose required for activity in this *E. coli* system, or if the lower levels of expression seen in the *P. agglomerans* system, could be replicated by yet lower concentrations of arabinose. For *Pantoea* the next step would be to see if the intermediate expression levels can be achieved by changing the concentration of arabinose in the system, and exploring the upper limit of induction.

3.4.3 Control of the expression of the *lux* operon under the control of P_{CRT} in *P. agglomerans* and *P. agglomerans* ΔLPP-1

pCT-LUX was transformed into *P. agglomerans* ΔLPP-1 and wild-type *P. agglomerans*. Colonies were selected on LB tetracycline and incubated at 30°C overnight. Colonies were then screened using a cooled CCD system (Biospace labs) for bioluminescence. For each strain one bioluminescent colony was selected and grown in LB tetracycline culture 30°C. These two overnight cultures were then diluted 1:1000 into fresh media and incubated at 30° 150 RPM for 4 hours.

The cultures were then loaded onto a 96 well plate (4x 200µl of bacterial culture). The plate was then incubated at 30°C in a Genios Pro plate reader (Tecan) for 24 hours. Measurements of OD600 and relative light units were taken every 30 minutes, results are shown in figure *3*.5.

The growth of both *P. agglomerans* and *P. agglomerans* ΔLPP-1 strains were similar with and without the addition of glucose as shown by the OD600 data. When looking at light production, wild-type *P. agglomerans* initially produces more light compared to *P. agglomerans* ΔLPP-1, but around the ten hour point *P. agglomerans* ΔLPP-1 is

at the same level. The addition of glucose does however increase the level of light production at later time points, in the wild-type, but not in *P. agglomerans* ΔLPP-1. Plotting RLU/OD600 against OD600 shows that *P. agglomerans* ΔLPP-1has an increase in the amount of light produced by cell once the cell culture reaches near to OD600 0.045, this could signify that P_{CRT} is controlled by quorum sensing mechanism only in *P. agglomerans* ΔLPP-1.

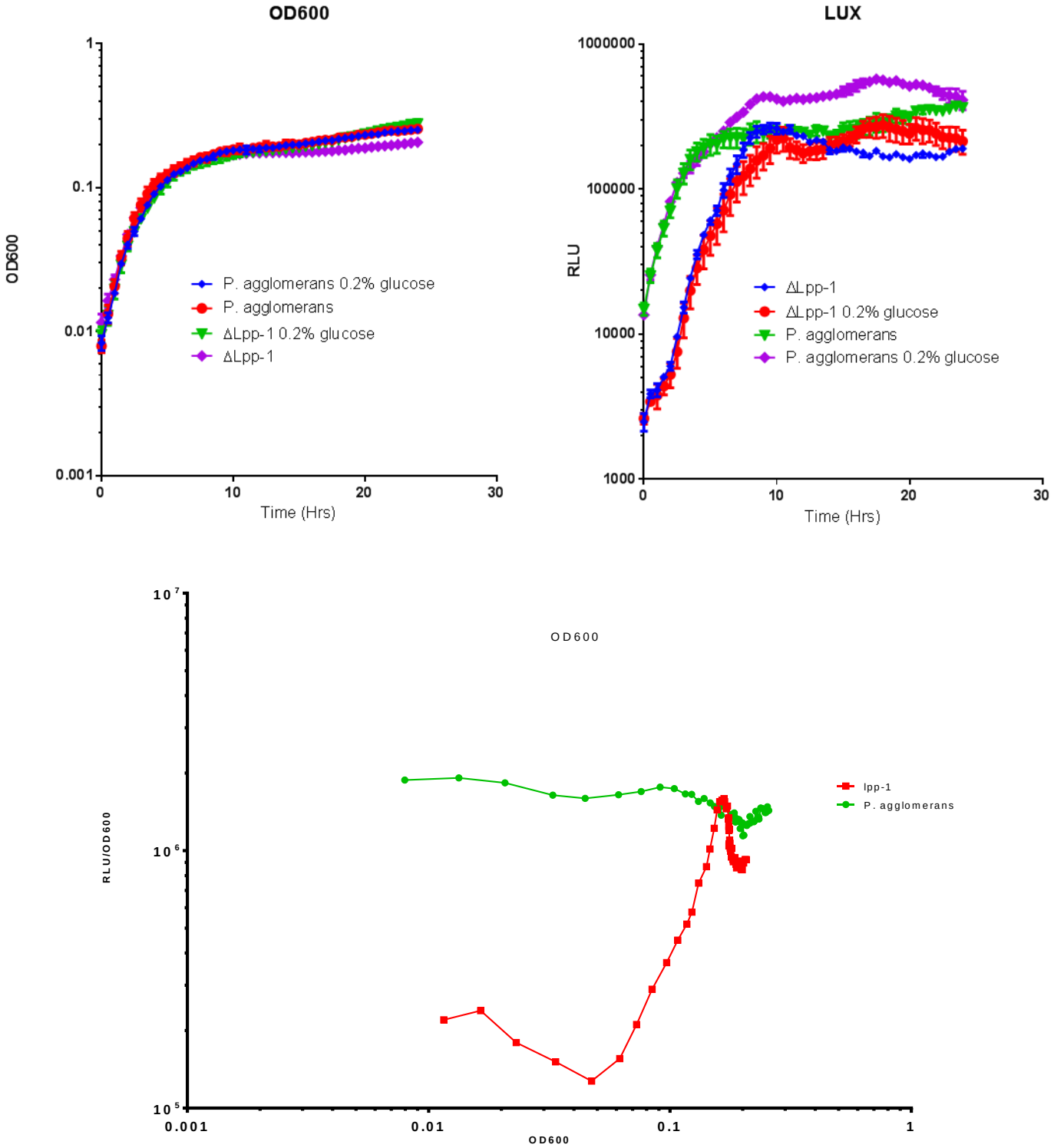


Figure 3-5 assessment of the dynamics of P_{CRT} in wild-type *P. agglomerans* and *P. agglomerans* ΔLPP-1. OD600, measuring cell growth B. shows light production over time, C. amount of light produced relative to the cell density, plotted against cell density

3.5 Comparison of the control and production of lycopene in *E. coli* and *P. agglomerans*

3.5.1 Construction of the Lycopene biosynthetic operon

After developing these tools for expression of genes and pathways in *Pantoea* the next step was to test their ability to produce terpenoids and to compare the capacity for terpenoid production of *E. coli and P. agglomerans*. To investigate how useful the previously characterised promoters could be for the production of terpenoids the lycopene biosynthetic operon was put under the control of P_{BAD}, P_{CRT} and P_{LAC}. Lycopene was chosen as it has a relatively short pathway (*crt*EIB), with no coloured intermediates, unlike the production of zeaxanthin. Lycopene is also easily extracted from bacterial cultures with hexane, and its concentration can be measured using a bench top spectrophotometer. These factors allow for fast screening of samples, thus enabling more samples to be taken and processed. The addition of Isopentenyl pyrophosphate *isomerase* (ipp isomerase) gene, *idi,* to the pathway was also investigated following the work by Sedkova & Tao, (2005), mentioned in chapter 1.

In order to construct the lycopene pathway the *crtE*, *crtI*, *crtB* and *idi* genes were cloned from *P. agglomerans* genomic DNA: *crtE* was amplified using primers int_crtE_f and int_crtE_r, *idi* was amplified using primers IDI_F and IDI_R, *crtI* and *crtB* were amplified together using primers int_CrtI_F_2 and int_crtB_R. These fragments were then restricted with *Not*I and *Xho*I and ligated into pBT-GFP using the same sites (in separate reactions). These reactions were transformed into *E. coli* and selected on LB tetracycline at 37°C overnight. Colonies were then screened by restriction digest; the final construct was sequenced and named pBT-CrtE. pBT-idi and pBT-CrtIB respectively.

These constructs were flanked by the compatible sites *Xba*I (upstream) and *Spe*I (downstream), this feature allowed for combinations of genes to be constructed more quickly. pBT-CrtE-idi and pBT-*crt*EIB were constructed by restricting pBT-CrtE with *Spe*I and *Xho*I, pBT- idi and pBT-CrtIB were restricted with *Xba*I and *Xho*I to extract *idi* and *crt*IB, these were then ligated into the cut backbone of pBT-CrtE

(separately). These reactions were transformed into *E. coli* and selected for on LB tetracycline at 37°C overnight. Colonies were then screened and confirmed by restriction digest.

pBT-CrtE-idi was then restricted by *Spe*I and *Xho*I, pBT-CrtIB was restricted with *Xba*I and *Xho*I, *crt*IB was then ligated into the cut pBT-CrtE-idi backbone, transformed into *E. coli* and selected for on LB tetracycline at 37°C. Colonies were screened and confirmed using restriction digest, this construct was named pBT-CrtE-idi-IB.

pSU-CrtE-idi-IB and pCT- CrtEIB/ CrtE-idi-IB

The GFP cassette of pSU-GFP and pCT-GFP were replaced with the lycopene synthesis constructs of pBT-CrtEIB and pBT-CrtE-idi-IB via the *Not*I and *Xho*I sites. pCT-CrtEIB and pCT-CrtE-idi-IB were transformed into *E. coli* and selected for on LB tetracycline at 37°C. pSU-CrtEIB and pSU-CrtE-idi-IB were transformed into *E. coli* and selected for on LB chloramphenicol at 37°C. Plasmid maps are shown in figure *3*.6 and *3*.1.

Colonies from pSU- CrtEIB and pCT-CrtEIB were unobservable after 24 hours. An additional 48 hours led to small colonies. To encourage the growth of these colonies, 2µl of 20% glucose solution was pipetted on to these colonies and they were then incubated overnight at 37°C. Colonies were observably producing lycopene, they were then grown in LB, with appropriate antibiotics and 0.2% glucose, plasmids were then prepped and the constructs were confirmed by restriction digest.

Colonies from pSU- CrtE-idi-IB and pCT-CrtE-idi-IB, were observable after 24 hours and were producing observable amounts of lycopene. Constructs were confirmed by restriction digest. Figure *3*.8 shows the difference in pCT-Crt-EIB and pCT-CrtE-idi-IB after 3 days growth at 37°C.



Figure 3-6 Cloning strategy for creating *crtEIB* and *crtE-idi-IB* operons. Selection of fragments cloned between the *Not*I and *Xho*I sites of pBT-GFP to produce: i. pBT-crtE, ii. pBT-idi, iii. pBT-crtE-idi, iv. pBT-crtIB, v. pBT-crtEIB, pSU-crtEIB and pCT-crtEIB and vi. pBT-crtE-idi-IB., pSU-crtE-idi-IB, and pCT-crtE-idi-IB. Genes not to scale. Plasmid maps for pBT-GFP, pCT-GFP and pSU-GFP shown in figure 3.1.

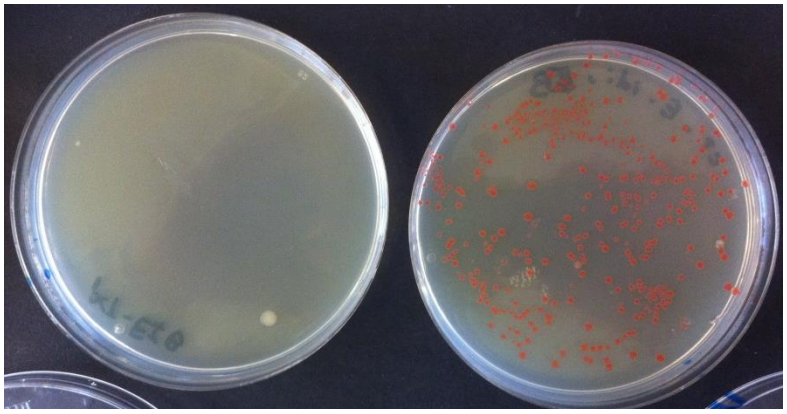


Figure 3-7 comparison of *E. coli* transformed with pBT-crtEIB and pBT-crtE-idi-IB Transformants with pBT-crtEIB (left), were very small colonies with a pinkish hue. However transformants with pBT-crtE-idi-IB (right) grew well and produced visible amounts of lycopene.

3.4.4 Transformation of the lycopene biosynthetic pathway into *P. agglomerans* ΔLPP-1

P. agglomerans ΔLPP-1 is an unpigmented strain of *P. agglomerans*, devoid of the carotenoid biosynthesis operon construction of this strain is described in chapter 6. This strain was used so that lycopene would not be converted to zeaxanthin by the native pathway. pSU-CrtEIB and pSU-CrtE-idi-IB were transformed into *P. agglomerans* ΔLPP-1 by electroporation and selected for on LB chloramphenicol with 0.2% glucose added to the media at 30°C overnight. pCT-CrtEIB, pCT-CrtE-idi-IB, pBT-CrtEIB and pBT-CrtE-idi-IB were transformed by electroporation and selected for on LB tetracycline media with 0.2% added glucose at 30°C overnight. Colonies were obtained for all six transformations and were confirmed by observation of lycopene production.

3.4.5 Testing production of lycopene in *E. coli* and *P. agglomerans* using P_{LAC}

As *E. coli* pSU-crtEIB exhibited a slow growth phenotype, measurements of the production of lycopene were not taken. However *E. coli* and *P. agglomerans* ΔLPP-1

harbouring pSU-CrtE-idi-IB were assessed for their ability to produce lycopene in culture.

Cultures of *E. coli* and *P. agglomerans* ΔLPP-1 harbouring pSU-CrtE-idi-IB were grown overnight at 37°C and 30°C, respectively, in LB Chloramphenicol with both 0.2% glycerol and 0.2% glucose. 2ml of each of these cultures was then added to 200 ml of fresh media (same as the overnight). This was grown to OD600 of 0.6 at 37°C for the *E. coli* culture and 30°C for *P. agglomerans*. The cultures were then split three ways and incubated at 30°C at 150rpm. After 24 and 48 hours 2ml samples were taken. Lycopene was extracted using the hexane methanol method (2.12.1 Lycopene extraction and quantification). Absorbance at 502nm was measured and the concentration of Lycopene was calculated using the Beer Lambert equation with a molar extinction coefficient at 502 nm in hexane of $1.72 \times 105\ M^{-1} \cdot cm^{-1}$ (Bóna-Lovász *et al.*, 2013).

P. agglomerans produced 24% more lycopene per litre after 24 hours and 21% more after 48 hours (4.2 mg/l and 4.7 mg/l) compared to *E. coli* (3.2 mg/l after 24 hours and 3.8 mg/l after 48 hours) (Figure *3*.9). A two tailed t-test shows that both of these time points show significant difference in production (p=0.0006 at 24 hours and p=0.0085 at 48 hours, with a threshold of $P< 0.05$). After 48 hours the dry cell weights per litre of culture were very similar between *E. coli* and *P. agglomerans* (2.4 g and 2.5 g respectively and not statistically significant with a two tailed t- test P= 0.2906). When lycopene production is normalised with the dry cell weight, *P. agglomerans* ΔLPP-1 does produce more lycopene per gram of dry cell weight (1.9 mg g^{-1} compared to 1.6 mg g^{-1} in *E. coli*) but this was not statistically significant (two tailed t-test, P= 0.0701).

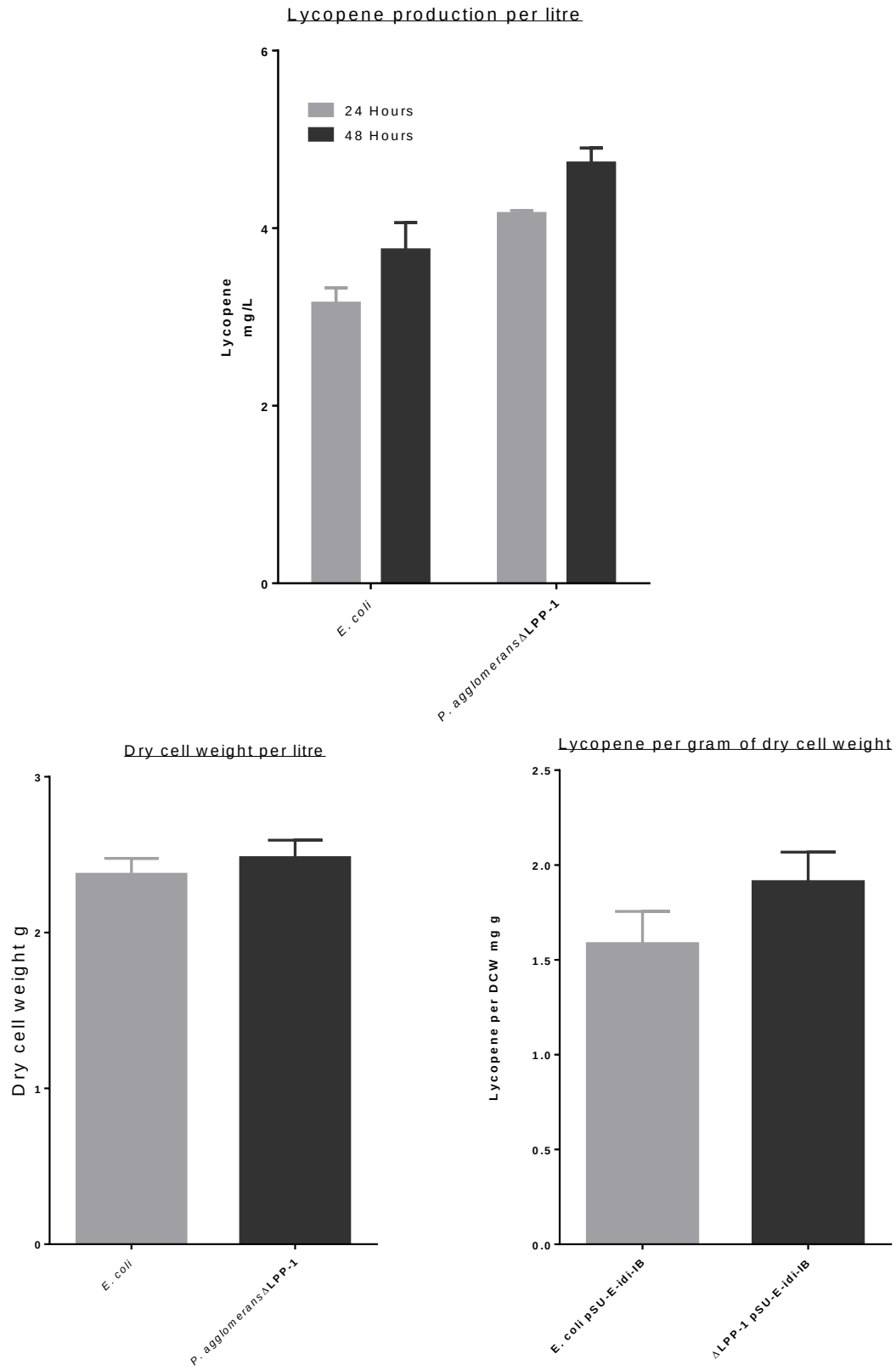


Figure 3-8Production of lycopene in *E. coli* and *P. agglomerans* LPP-1 transformed with pSU-crtE-idi-IB. A. lycopene production after 24 and 48 hours, B. dry cell weight after 48 hours, C. lycopene per gram of dry cell weight after 48 hours.

3.4.6 Effect of expression level on lycopene production, using pBT-CrtEIB and pBT-CrtE-idi-IB

Following the establishment that both organisms can produce lycopene and that the production can be quantified it was decided to analyse how different induction levels can affect production of lycopene. For this the P_{BAD} constructs, pBT-CrtEIB and pBT-CrtE-idi-IB, were used.

LB culture was replaced with 2YT media in this experiment as this is a richer media anticipated to promote the production of lycopene.

Cultures of *E. coli* and *P. agglomerans* harbouring pBT-CrtEIB were grown overnight in LB tetracycline + 0.2% glycerol at 37°C and 30°C respectively. For both organisms 1:100 dilutions were made into 6X 50ml 2YT tetracycline +0.2% glycerol and grown until OD600 0.8 at 37°C for *E. coli* and 30°C for *P. agglomerans*. Arabinose (0.2%) was added to 3 cultures of each organism. All cultures were then incubated at 30°C at 150rpm. After 22 and 46 hours lycopene was extracted from 5ml of culture and absorption was measured as previously described. After 46 hours the dry cell weight of 25ml of culture was determined.

Figure *3*.10 shows the production of lycopene, induced and un-induced, was higher in *P. agglomerans* (146% increase from *E. coli* to *P. agglomerans* ΔLPP-1 after 46 hours un-induced) however, the addition of 0.2% arabinose had a negative effect on the production of lycopene. For *E. coli,* production dropped 64% from 3.0mg/l to 1.1mg/l and for *P. agglomerans* ΔLPP-1 74% from 7.4mg/l to 1.9mg/l. The addition of arabinose caused a significant reduction in dry cell weight in both *E. coli* and *P. agglomerans* ΔLPP-1*. E. coli weight* dropped 60% and *P. agglomerans* ΔLPP-1 was reduced by 59%. However when lycopene per dry cell weight is compared, *E. coli* suffers an insignificant reduction (two tailed t-test, p=0.5208) compared to the significant 37% decrease seen in *P. agglomerans* ΔLPP-1 (two tailed t-test, p= 0.0010).

The reduction in mass accumulation indicates that the bioburden imposed by the induction of the lycopene biosynthetic operon was taxing on the population and reduced growth. This highlights the importance of balancing cell growth and maintenance with production of secondary metabolites. This possibly helps to explain why *E. coli* struggled to grow whilst transformed with pSU-CrtEIB, as P_{LAC} is known for being 'leaky' with a higher basal level of transcription than other promoters. This could be further exacerbated by trace amounts of lactose in LB media carried over from the production of tryptone from milk.

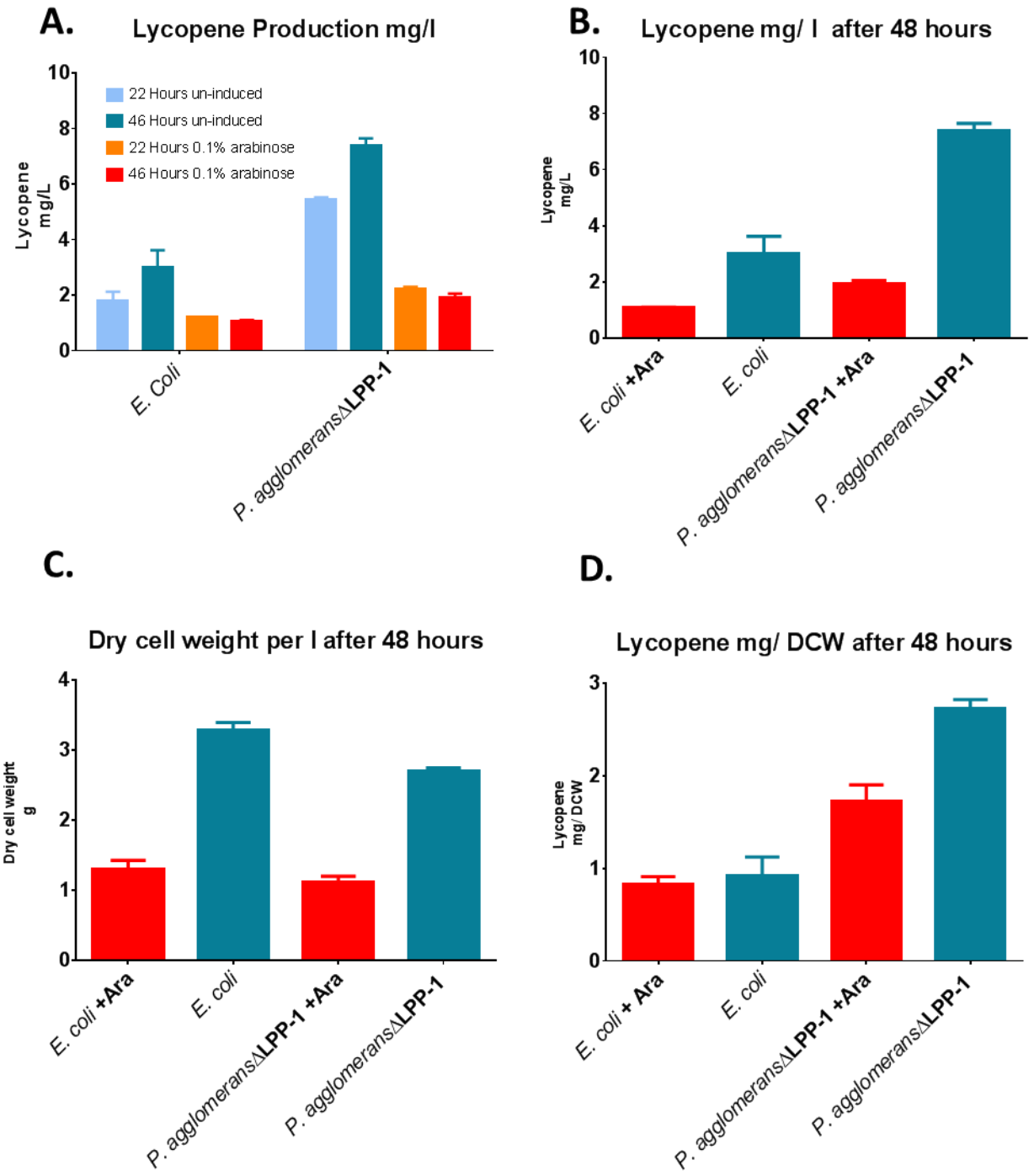


Figure 3-9 Production of lycopene from *E. coli* and *P. agglomerans* LPP-1 transformed with pBT-crtEIB with and without induction by the addition of arabinose. A. shows lycopene production per litre after 24 and 48 hours with and without induction, B. lycopene per litre after 46 hours with and without induction, C. dry cell weight with and without induction, D. lycopene per dry cell weight with and without induction.

3.4.7 Induction and the introduction of *idi*

Following the previous two experiments, it was apparent that the production of lycopene was linked to the ability to balance growth of cells with the carbon flux down the lycopene biosynthetic pathway, and that the level of induction and the inclusion of *idi* into the *crtEIB* operon both played a role in control the level of carbon flux diverted to lycopene metabolism. To explore the interaction of these factors in the production of lycopene in both *E. coli* and *P. agglomerans* ΔLPP-1; Lycopene concentration was measured after 20 hours, under three different induction levels of arabinose, with or without *idi* included into the crtEIB operon, under the control of the promoter P_{BAD}.

E. coli pBT-CrtEIB*, E. coli* pBT-CrtE-idi-IB, *P. agglomerans* ΔLPP-1 pBT-CrtEIB and *P. agglomerans* pBT-CrtE-idi-IB were cultured in 100ml of 2YT media containing tetracycline and 0.2% glycerol. Cultures were started by 1:100 dilution from overnight cultures in the same media and conditions. *E. coli* cultures were cultured at 37°C at 150 rpm and *P. agglomerans* ΔLPP-1 cultures were grown at 30°C at 150 rpm for the duration of the experiment. At OD600 of 0.8 cultures were split into 9x 50ml Falcon tubes, 10ml into each. Arabinose was then added so as to achieve either 0%, 0.02% or 0.2% (w/v), these were then put back in to shaking incubators set at the temperatures they had previously been at. After 20 hours 5ml of culture was spun down and lycopene was extracted using the hexane, methanol method. The other 5ml was spun down and the dry weight calculated. Results are shown in figure *3*.11

For *E. coli* the inclusion of *idi* clearly increases lycopene production per litre. The impact of the addition of *idi* is most obvious in the populations that were induced with arabinose. In the non *idi* population the addition of 0.02% arabinose increased lycopene production by 39.6% and 0.2% arabinose had negligible effect, whereas in the *idi* populations the 0.02% arabinose increased lycopene production by 103.7%

compared to no induction and 0.2% arabinose caused a 120.1% increase. When the *idi* and non *idi* groups were compared directly based on their arabinose concentration there was an 82.2% increase in the 0% arabinose group, 165.9% increase for the 0.02% arabinose population and 325.4% for the 0.2% arabinose cultures in the *idi* group. This shows that the addition of *idi* to the lycopene biosynthetic pathway enables a greater yield of lycopene, but also overcomes underlying issues with induction of the pathway. This is more easily seen when comparing the dry cell weight of the cultures. Without *idi* the cell mass drops with the addition of arabinose, however this drop does not occur in the *idi* group.

In *P. agglomerans* without *idi*, the addition of 0.02% arabinose increased production of lycopene per litre by 137.8% compared to the uninduced culture, however increasing the concentration of arabinose to 0.2% did not increase production of lycopene further. The addition of *idi* to the operon causes a 144.9% increase in lycopene production in the uninduced cultures, but this does not exceed the production of the non *idi* strain with induction. Induction with arabinose of the *idi* cultures show no further increase at 0.02% arabinose and a 62.3% decrease with the addition of 0.2% arabinose.

In the non *idi* plasmid the addition of increasing concentrations of arabinose leads to a drop in the cell mass generated. For the *idi* plasmid the average mass is higher than the non *idi* uninduced and is not impacted as strongly by the addition of 0.2% arabinose compared to the non *idi* equivalent.

When the production of lycopene was normalised by taking into account the dry cell weight, the addition of *idi* to the *crt*EIB operon appears to have a negative effect on production, despite increasing yield per litre. This is caused by a larger increase in dry cell weight than in lycopene production.

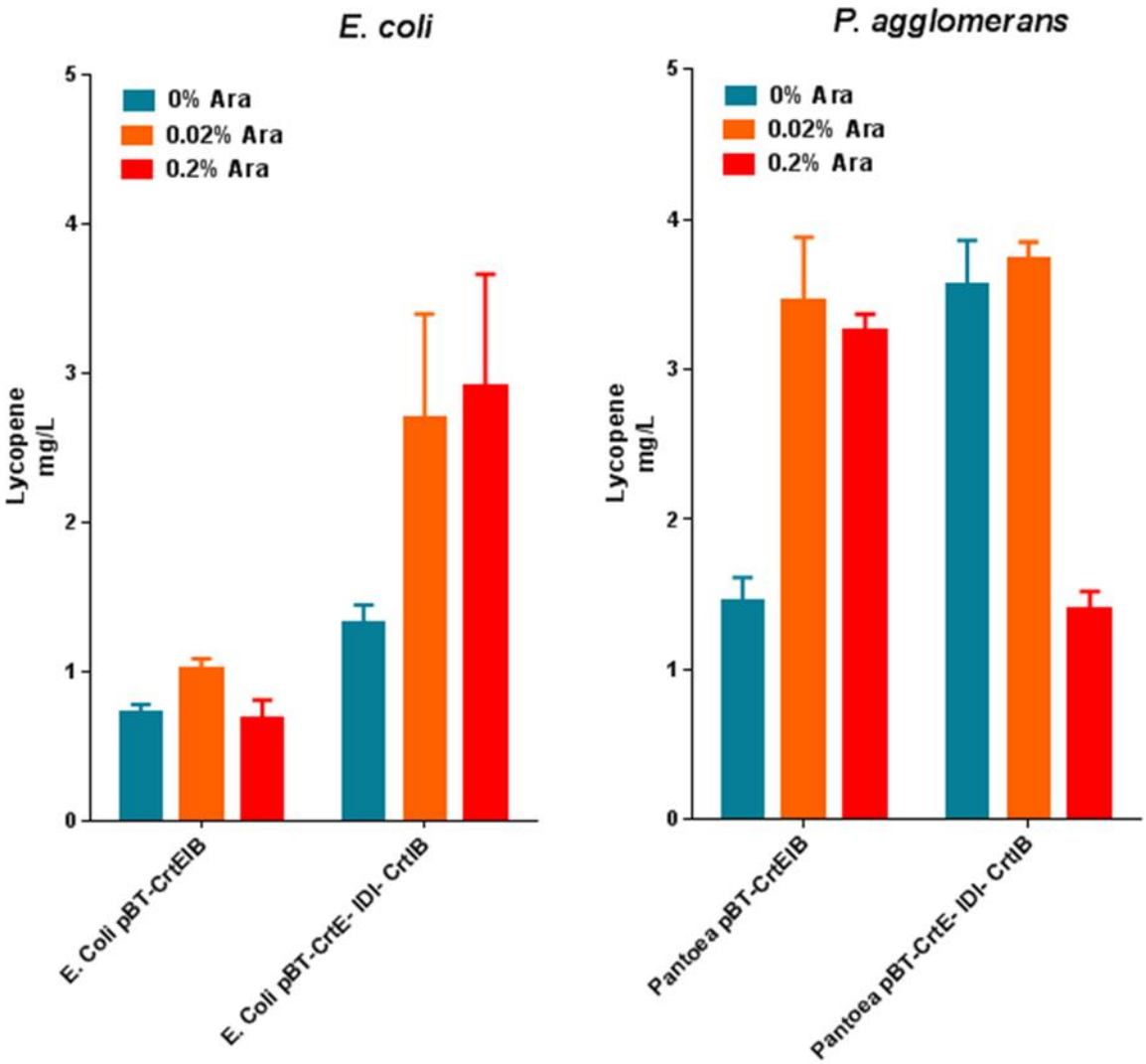


Figure 3-10 Production of lycopene per litre after 24 hours in *E. coli* and *P. agglomerans* ΔLPP-1 transformed with either pBT-crtEIB or pBT-crtE-idi-IB under three levels of induction from the addition of arabinose.

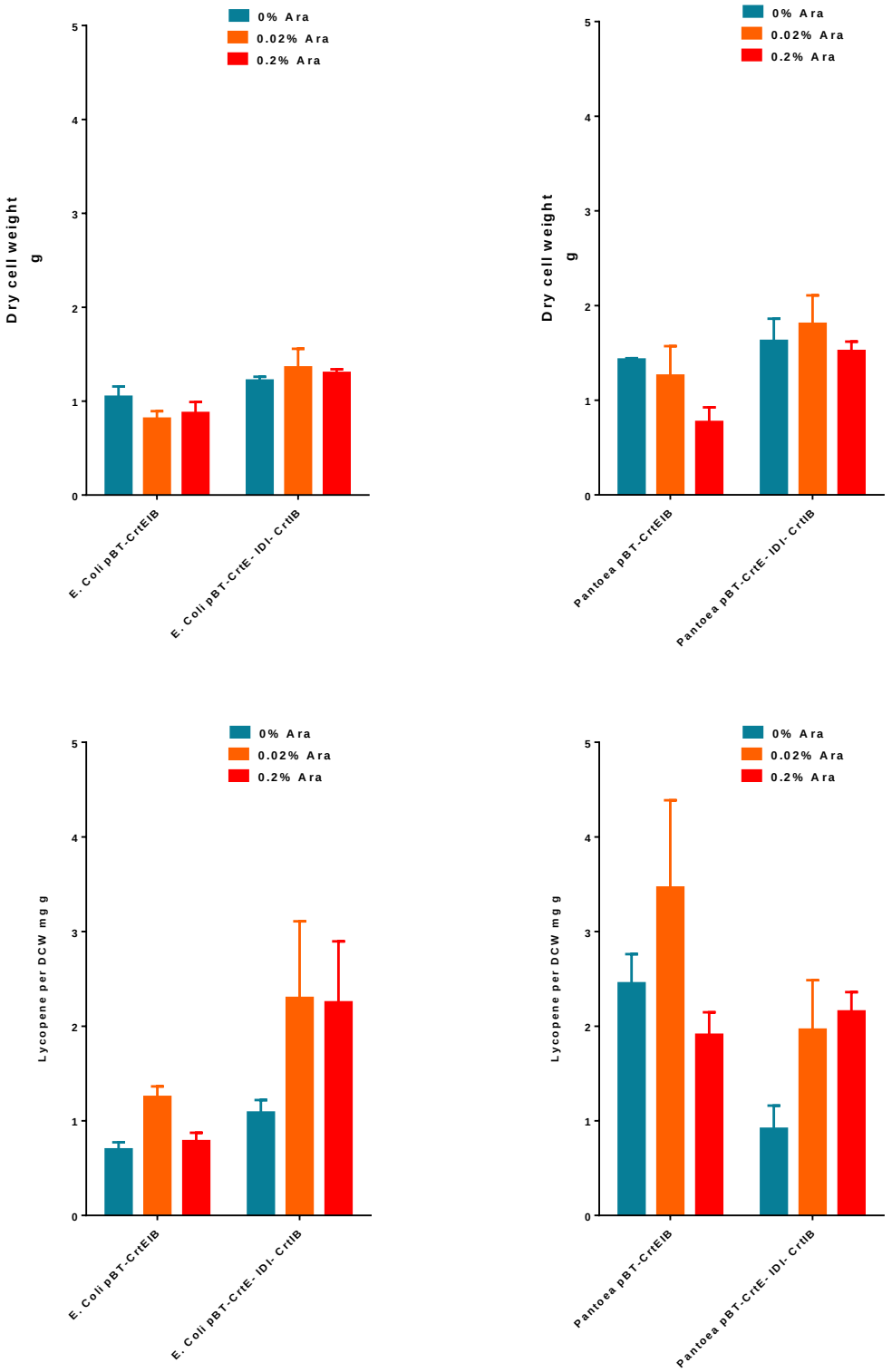


Figure 3-11Production of lycopene from *E. coli* and *P. agglomerans* LPP-1 transformed with pBT-crtEIB or pBT-CrtE-idi-I under three levels of induction A.*E. coli* dry cell weight B. *P. agglomerans* ΔLPP-1 dry cell weight, C. *E. coli* lycopene per gram of dry cell weigh, D. *P. agglomerans* ΔLPP-1 lycopene per gram of dry cell weight.

3.4.8 Lycopene production under control of P_{CRT}

The need for carotenoids, to protect against UV damage and other stress, and the constraints in producing them have selected P_{CRT} and its regulation to balance growth and the production of carotenoids. Following the observations of the behaviour of pCT-LUX in both *P. agglomerans* wild-type *and P. agglomerans* ΔLPP-1 it was decided to look at how *E. coli* and *P. agglomerans* ΔLPP-1 express *crt*EIB and *crt*E-idi-IB under the control of P_{CRT}. Wild type *P. agglomerans* could not be used as it would have converted any lycopene into zeaxanthin, Efforts were made to edit the LPP-1 of wild *type P. agglomerans* in order to knockout *crtXYZ* to make a strain which only produced lycopene. However these efforts were not successful (chapter 5).

The influence of P_{CRT} on the production of lycopene was tested using *E. coli* and *P. agglomerans* ΔLPP-1 transformed with either pCT-CrtEIB and pCT-CrtE-idi-IB. Cultures were grown in 2YT 0.2% glycerol with tetracycline added, *E. coli* was grown overnight at 37°C 150RPM and *P. agglomerans* ΔLPP-1 was grown at 30°C 150RPM. 1:100 dilutions were then made into 3X 100ml of fresh media for each strain in 250ml conical flasks. *E. coli* cultures were then incubated at 37°C 150RPM until an OD600 of 0.6 and then transferred to 30°C 150RPM alongside *P. agglomerans* ΔLPP-1. Lycopene was measured after 22, 48 and 72 hours, by using the hexane, methanol method on 2ml of culture. After 48 hours of culturing, 400µl of 50% glycerol and 50µl of 10mg/ml tetracycline were added to the cultures to maintain growth and selection pressure. After 72 hours the dry weight of 85ml of culture was measured.

The final colour of the cultures is shown in Figure *3*.13 and lycopene production is shown in Figure *3*.14 and *3*.15. In *E. coli,* the introduction of *idi* raised the production of lycopene after 72 hours by 545% from 4.0mg/l to 26.1 mg/l. In *P. agglomerans* ΔLPP-1 the increase from the addition of *idi* to the *crt*EIB operon was more modest at 125% (13.8mg/l to 31.0 mg/l). The addition of *idi* appears to increase the rate of production of lycopene in *P. agglomerans* ΔLPP-1 to such an extent that *P. agglomerans* ΔLPP-1 pCT-*crt*E-idi-IB has a rate of production that is approximately double its counterpart without *idi* on the operon, 24 hours in advance.

E. coli pCT-CrtE-idi-IB and *P. agglomerans* ΔLPP-1 pCT-CrtE-idi-IB are not significantly different in their lycopene production in the first 48 hours (t-test 22 hours p= 0.068, 48 hours p= 0.399). After 72 hours however, *P. agglomerans* ΔLPP-1 pCT-CrtE-idi-IB produces 19% more than *E. coli* pCT-CrtE-idi-IB. This result is statistically significant (t-test p= 0.007). This might be attributed to the addition of glycerol and tetracycline after 48 hours.

The dry cell weight analysis after 72 hours shows that the inclusion of *idi* increased the cell weight of *P. agglomerans* ΔLPP-1 cultures whilst reducing the dry cell weight of *E. coli.* Once lycopene production was normalised using dry cell weight, *P. agglomerans* ΔLPP-1 pCT-crtE-idi-IB produces 22% more lycopene than *E. coli* pCT-crtE-idi-IB (12.5mg l^{-1} and 10.28mg l^{-1} respectively), this difference was shown to be statistically significant with a two tailed t-test giving a P value = 0.022.

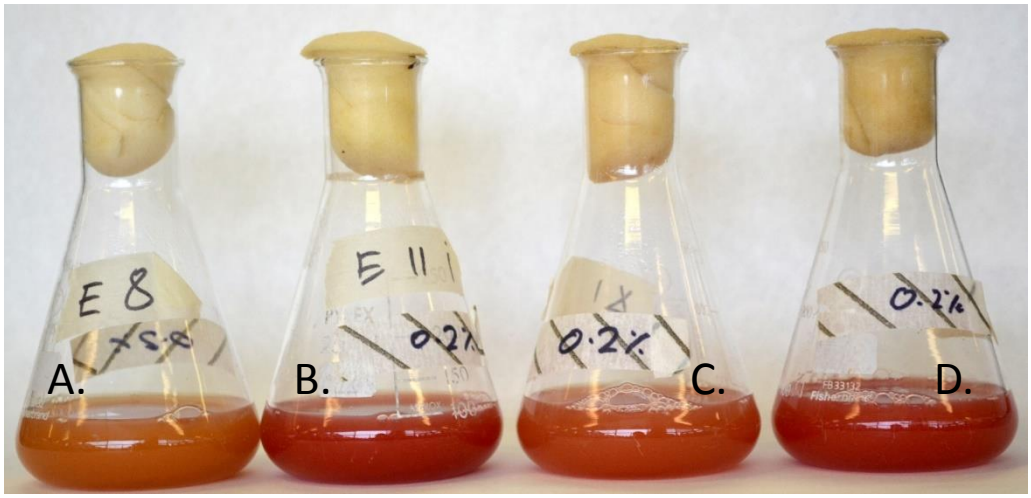


Figure 3-12 Production of lycopene after 72 hours in *E. coli* and *P. agglomerans* ΔLPP-1 transformed with pCT-crtEIB or pCT-crtE-idi-IBA. *E. coli* pCT-crtEIB, B. *E. coli* pCT-crtE-idi-IB, C. *P. agglomerans* ΔLPP-1 pCT-crtEIB, D. *P. agglomerans* pCT-crtE-idi-IB

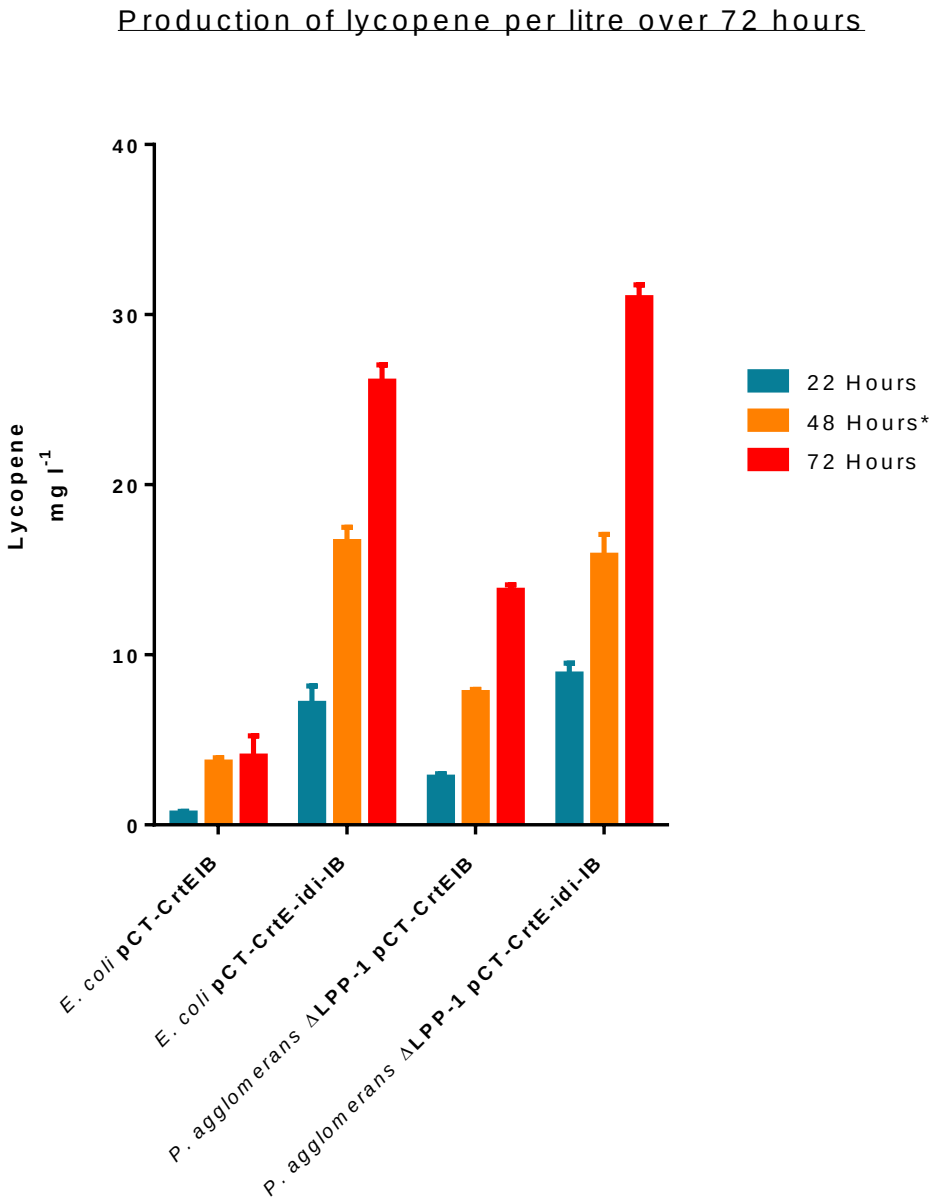


Figure 3-13 Production of lycopene over 72 hours in *E. coli* and *P. agglomerans* ΔLPP-1 transformed with pCT-crtEIB or pCT-crtE-idi-IB.

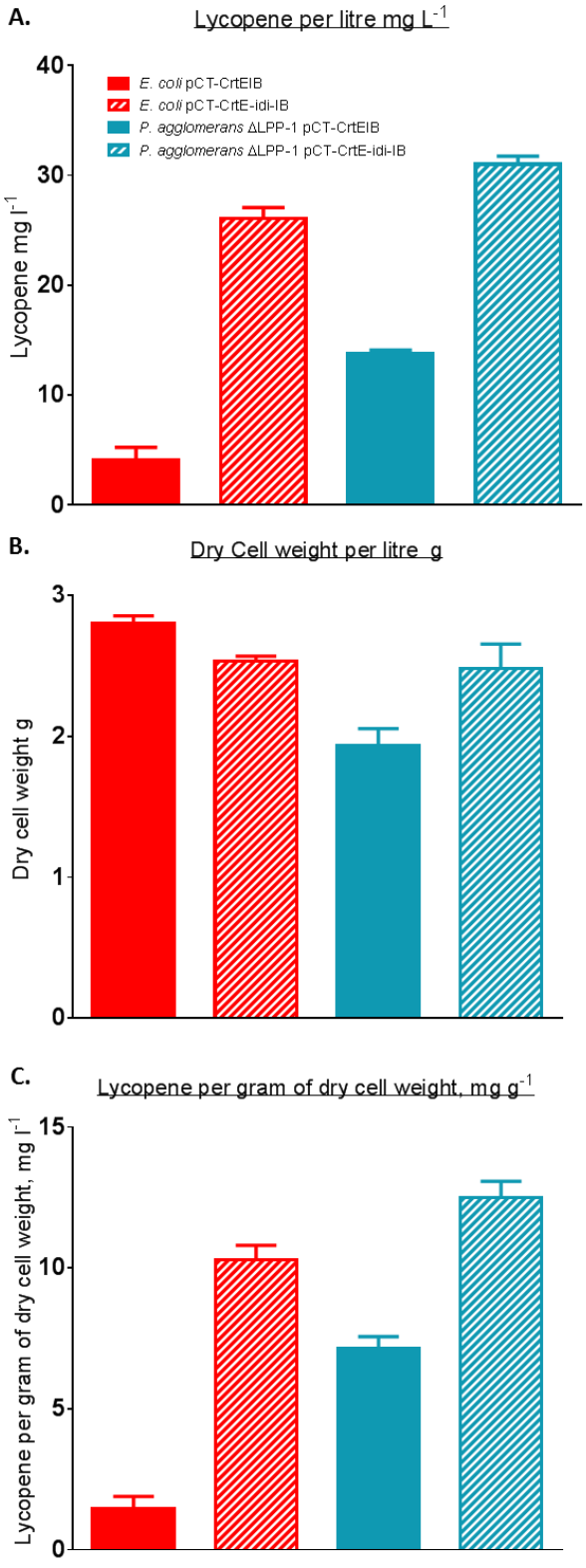


Figure 3-14 Production of lycopene after 72 hours in *E. coli* and *P. agglomerans* ΔLPP-1 transformed with pCT-crtEIB or pCT-crtE-idi-IB. A. lycopene per litre, B. dry cell weight, C. lycopene per dry cell weight.

3. 5 Discussion

In this chapter several plasmids have been produced with a range of reporter systems for the characterisation of promoters in *P. agglomerans,* with the aim to improve diterpenoid production.

Having data on the characterisation of several promoters in *P. agglomerans* was extremely useful moving forward with the project and to design expression systems. The screening process which moves from testing initial functionality of a promoter, then to its dynamics in culture using the lux operon and then finally characterising its effect of the production of lycopene, provides a useful screening process for promoters. It also offers the possibility to be scaled up in order to test hundreds of promoters. Potential promoters from *P. agglomerans*, other *Pantoea* and from other organisms could be cloned with standardised restriction sites at the end (Much like *Not*I and *Xho*I are used in this study), these could then be cloned upstream of GFP for an initial viability screen. The plasmid modularity seen in this chapter could then be used to connect viable promoters to *lux* and GFP and then measure their output using the plate reader. This data could be visualised and the promoters organised based on their behaviours and characteristics. Finally a range of promoters with varying characteristics could be screened for their ability to produce high yields of lycopene, helping to build up a better understanding of the best strategy for high yields of diterpenoid production.

Most of these steps can now be automated (with advances in robotics and computer science) allowing for large number of promoters to be screened. This automation also allows for viable novel synthetic promoters to be discovered more quickly, a service which several companies are now offering, these companies screen libraries of synthetic promoters for activity asked for by the client. Promoters can be screened for a wide range of activities including switching on or off in response to specific chemicals or conditions, having a certain periodicity or just having the optimal constitutive induction level.

From the work presented in this chapter the most promising promoters are P_{BAD} and P_{CRT}. From the range of induction levels used in the P_{BAD} *lux* experiment*, P. agglomerans* with P_{BAD} has been shown to provide a greater range of expression levels, compared with the same system in *E. coli,* which shows finer differentiation of light output over the same range of expression levels. Further work needs to be carried out to determine if there is a threshold level of arabinose required for activity in this *E.coli* system, or if the lower levels of expression seen in the *P. agglomerans* system, could be replicated by yet lower concentrations of arabinose. For *Pantoea,* the next step would be to see if the intermediate expression levels can be achieved by changing the concentration of arabinose in the system, and exploring the upper limit of induction. The ability to tune P_{BAD} seen in the *lux* experiments, offers the possibility of finding 'sweet spot' for induction of products without having to screen large libraries of promoters

The control of P_{CRT} still needs to be elucidated, LPP-1 appears to have some role in the regulation of P_{CRT} and appears to regulate P_{CRT} to make it express constitutively in the wild-type background, and in a quorum sensing like manner in the ΔLPP-1 deletion strain. LPP-1 may potentially regulate P_{CRT} in such a way that a lower threshold of population density is required for induction, but this would require further investigation. (As P_{CRT} was isolated a lot later in the project and so not mentioned in the proceeding chapters. However a plasmid combining P_{CRT} with *crt*E and *txs* could potentially be very interesting to see if applying the control imposed on to the biosynthesis of one diterpenoid can help improve the yield of another.)

The methods used to measure lycopene production in this chapter, while not necessarily being as accurate as other methods such a HPLC, were easy to carry out on a large number of samples. Also extracting lycopene instead of zeaxanthin had the added benefit of using disposable plastic cuvettes instead of the glass ones required when undertaking an acetone extraction of zeaxanthin, this sped up measurements considerably.

Overexpression of *crt*EIB can clearly retard the growth of *E. coli* and *P. agglomerans* as shown in section *3*.4 5 and *3*.4 6. In section *3*.4.5, *3*.4.7 and *3*.4.8 it was shown that the addition of *idi* helps to alleviate these growth issues and in the example of the pSU-crtEIB were the addition of *idi* allows colonies to be grown. The potential reason for this is the fact that *idi* helps to balance the production of IPP and DMAPP. At first resolving this imbalance would appear to just increase the flux of carbon towards lycopene, thus exacerbating the imbalance between growth and production of secondary metabolites, unless other products of the terpenoid biosynthesis pathway are required for growth, and their production is hampered by the drain imposed by CrtEIB. This drain can be explained by the fact that FPP, as well as being a precursor to GGPP is a precursor to respiratory quinones which are essential for cell growth. I hypothesise that without increased expression of *idi,* in synchronicity with expression of *crtEIB* can lead to a depletion of FPP and therefore reduce quinone production, and therefore slow growth. Inclusion of *idi* in operons including *txs* should be considered in future projects, as well as attempting to use a copy of *idi* from other organisms should be considered, so as to avoid any potential homologous recombination due to multiple copies of *idi* being present at different locus in the genome.

P_{BAD} and P_{CRT} both show potential to be useful promoters for the production of high yields of diterpenoids in *P. agglomerans*. The experiments they were tested in were however not comparable due to scale and time period. Future experiments could directly compare these systems in *P. agglomerans,* and also use a greater range of concentrations of arabinose for induction of P_{BAD} and potentially looking at allowing the culture to reach a higher OD600 before induction.

In the three experiments in which *E. coli* and *P. agglomerans* were grown in parallel conditions, *P. agglomerans* ΔLPP-1 was able to produce more lycopene per litre than *E. coli* (21% more using pSU-crtE-idi-IB, 145% more using pBT-crtEIB and 19% more with pCT-crtE-idi-IB). The maximum recorded concentration of lycopene per litre was 31.9mg l^{-1} (12.5mg g^{-1} dcw) was achieved by using *P. agglomerans* ΔLPP-1 pCT-crtE-

idi-IB after 72 hours, cultured at 30°C 150 RPM in 2YT tetracycline +0.2% glycerol (plus the addition of 0.2% glycerol w/v after 48 hours), chapter *3*.4.8.

The production of lycopene shown here is not so high when compared to recent developments in production of lycopene in *E. coli* which have now reached production levels of 529.45 mg/L and 20.25 mg per gram of dry cell weight in the fed-batch culture. This was achieved by integrating the pathway into the chromosome to improve stability, the addition of the mevalonate pathway from *S. cerevisiae* and promoter engineering (Shen *et al.*, 2015). However the literature prior to this ranges from 7.7mg l^{-1} to 160mg l^{-1} (Alper & Stephanopoulos, 2007) (Farmer & Liao, 2000) and a yield of 7.8mg g^{-1} dcw to 16mg g^{-1} dcw (Shimada *et al.*, 1998)(Jin & Stephanopoulos, 2007), showing that the work presented in this thesis is comparable with those earlier achievements.

Looking at an experiment with a similar design to that undertaken here, Yoon *et al.*, (2007) compared the use of *P. agglomerans* and *P. ananatis crt*EIB in *E. coli* . They were able to produce a strain of *E. coli* capable of producing 27mg l^{-1} of lycopene after 48 hours in 2YT +0.5% glycerol, they then improved this to 60mg l^{-1} by the addition of the 'bottom half' of the mevalonate pathway, which converts mevalonate into DMAPP and IPP, and supplementing the culture with mevalonate. This gives a bench mark to compare the efforts to produce lycopene in *P. agglomerans,* and suggests future investigations into the potential addition of the mevalonate pathway to *P. agglomerans.*

Despite the differences in lycopene production between *E. coli* and *P. agglomerans* not being high when *idi* is included in the operon, *P. agglomerans* still shows potential as a host for diterpenoid production. Currently these experiments have been set up using experience of *E. coli* preferences for diterpenoid production, but being able to easily measure lycopene production makes it a useful tool for identifying culturing conditions favourable to *P. agglomerans.* These can then be used to inform the conditions used for the production of taxadiene, which is much more difficult to

measure. This should help to reduce the test and redesign cycle time and improve yields of taxadiene.

An interesting avenue of investigation will be to look at the ability of *P. agglomerans* to utilise carbon sources and especially larger quantities of glycerol. The addition of glycerol in the last 24 hours of the pCT-crtE-idi-IB experiment, the addition of glycerol increased production of lycopene more in *P. agglomerans* than in *E. coli.*

Chapter 4 DEVELOPMENT OF CHROMOSOMAL INTEGRATION AND GENE KNOCKOUT SYSTEMS FOR USE IN *P. AGGLOMERANS*

Manipulation of an organism's chromosome is vital metabolic engineering. Deletion of competitive pathways can improve the yield of products, and removal of certain recombinases and insertional elements can improve strain stability, for example the clean gene strain produced by Scarab Genomics (www.scarabgenomics.com). Chromosomal integration of pathways can also improve their stability and remove the need for added selection agents (such as antibiotics) to fermentation, therefore reducing the cost of production. This section deals with the attempted chromosomal integration of *txs* using Tn*7* transposition and the establishment of a suicide plasmid-based homologous recombination gene knockout system.

4.1 Transposon mediated chromosomal integration

An integration system that could be used in *Pantoea* was the first of these tools to be investigated. pGRG36 was built by McKenzie & Craig, (2006) and uses the Tn*7* transposon machinery combined with the temperature sensitive origin of replication pSC101ts to integrate DNA downstream of *glm*S, after using a recognition site within this highly conserved gene. The genes to be integrated are cloned between the left and right Tn*7* transposon ends in a MCS with unique sites *Avr*II, *Not*I, *Pac*I and *Xho*I. The plasmid is then transformed into the bacteria of interest (or mated in via the OriT) and selected for with ampicillin. Single colonies are then grown overnight at 30°C then plated without ampicillin selection at 37°C to cure the cells of the temperature sensitive plasmid. PCR can then be used to screen for integration.

4.1.1 Replacement of the antibiotic resistance cassette of pGRG36, from ampicillin to tetracycline resistance

As the strains of *P. agglomerans being* examined is ampicillin resistant the resistance cassette of pGRG36 needed to be changed, pGRG36 and the subsequent tetracycline resistant pGRG36T are shown in Figure *4*.1. pGRG36 was obtained from the Craig lab, via add gene (McKenzie & Craig, 2006). pGRG36 was converted into pGRG36T by amplifying *tetR* from pBR322 using primers tetgib_f3 and tetgib_r2. This was then restricted with *Sbf*I leaving one end blunt (KOD polymerase was used in PCR amplification to create blunt ended product). pGRG36 was restricted with *Sbf*I and *Nae*I (blunt ended cutter) to remove ampicillin resistance. The *tetR* fragment was then ligated into this site. *E. coli* were then transformed and transformants were selected for on LB tetracycline at 30°C overnight. The number of colonies recovered was low (less than 10) these were grown overnight in broth and the plasmids were then extracted and screened by restriction digest.

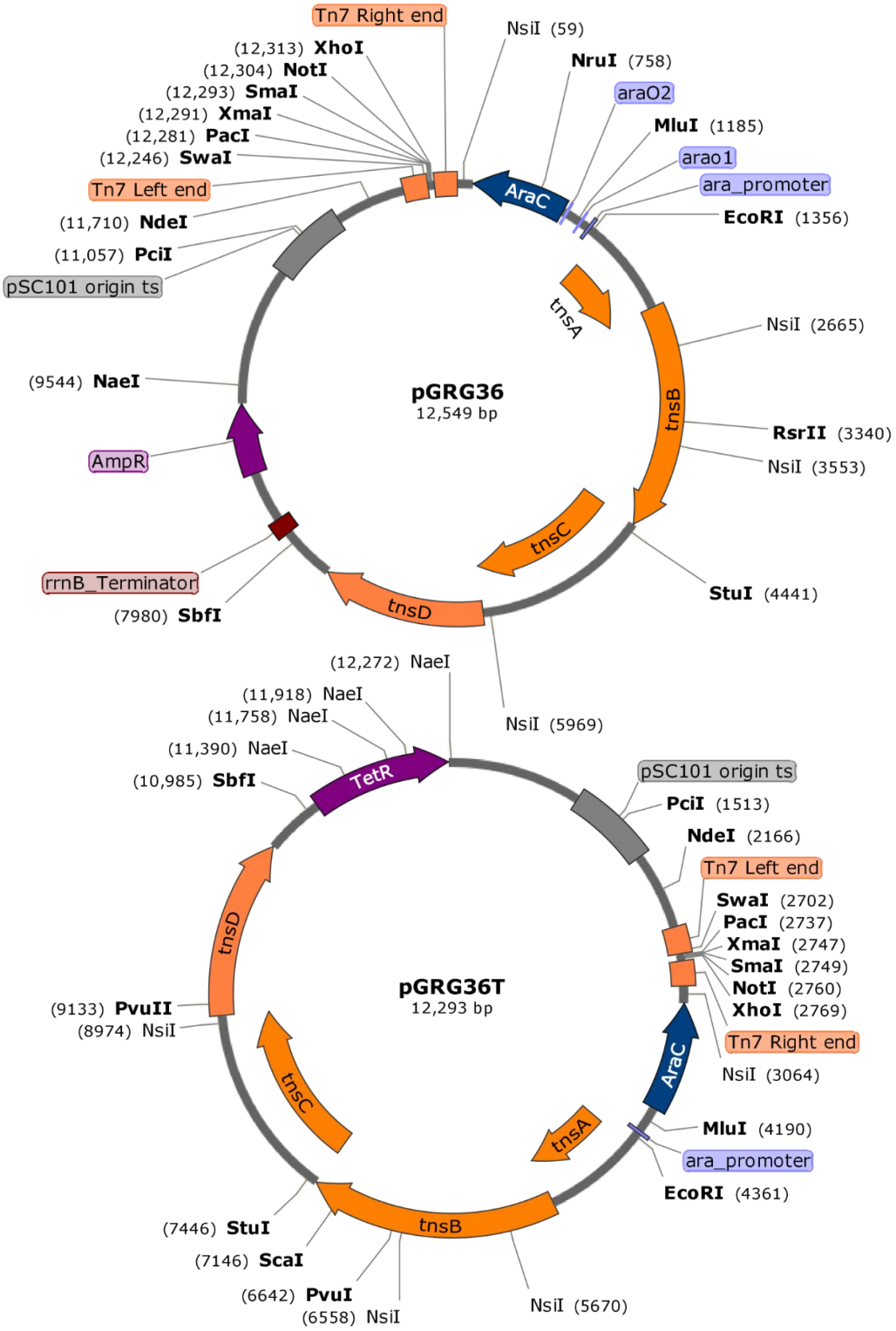


Figure 4-1 Plasmid maps of pGRG36 and pGRG36T. pGRG36T was created by replacing the ampR cassette of pGRG36 with a *tetR* cassette. Genes to be integrated into the hosts chromosome are cloned between *Pac*I and *Xho*I sites

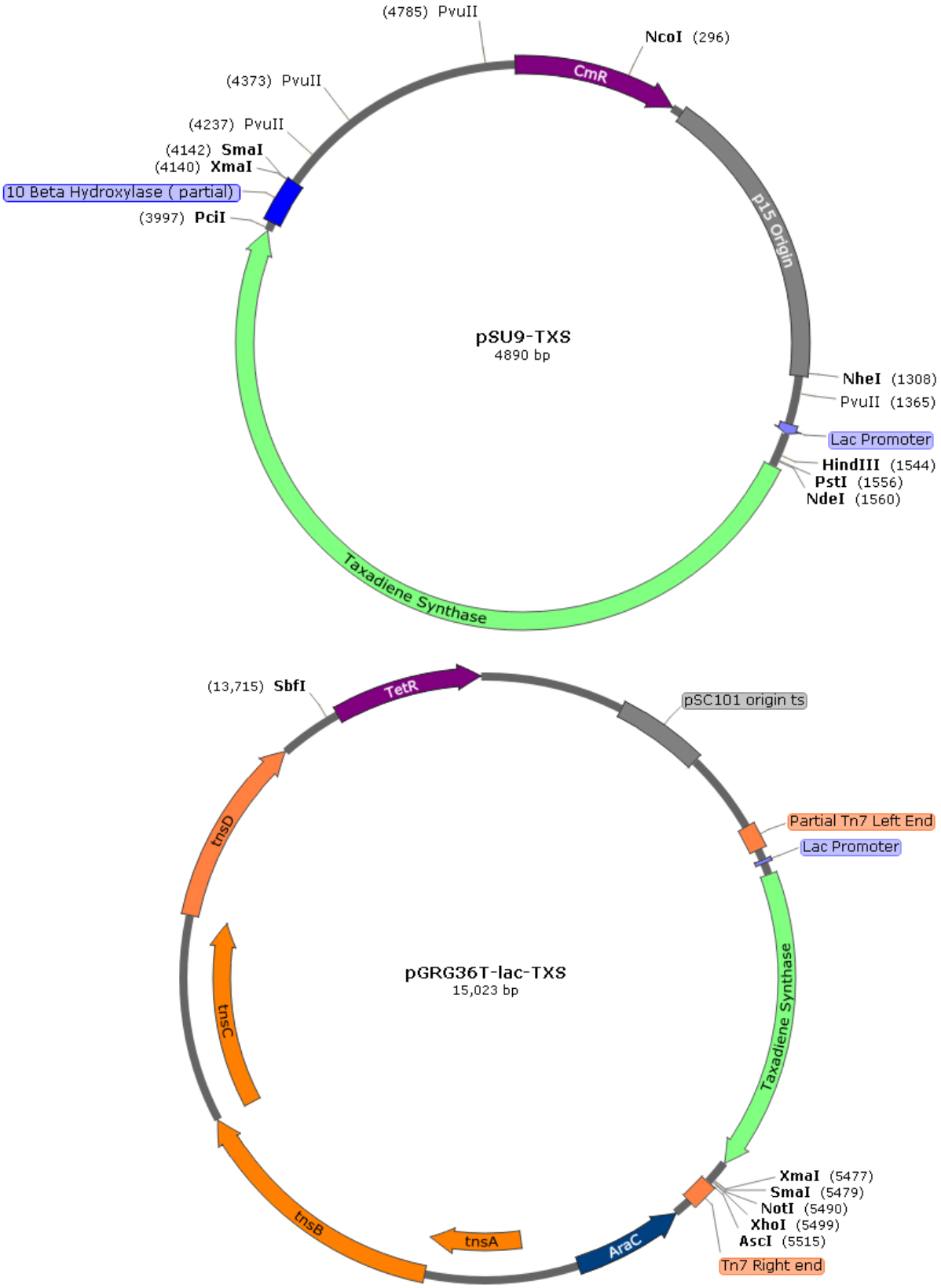


Figure 4-2 Plasmids maps of pSU-TXS and pGRG36T-lacTXS. TXS was restricted from pSU-TXS and cloned in between *Swa*I and *Xma*I

4.2 Chromosomal integration of *txs* into *P. agglomerans* using pGRG36t-Lac-TXS

Having converted pGRG36 to be usable in *P. agglomerans,* the next step was to attempt to integrate Taxadiene synthase (*txs*) to the chromosome. To ensure that *txs* was under the direct control of a promoter pSU-TXS (described in chapter 5) was restricted with *PvuII* and *XmaI* to produce a P_{LAC} *txs* fragment, this was then ligated between the *SwaI* and *XmaI* site of pGRG36T. This construct was then confirmed by restriction mapping.

4.2.1 Integration of P_{LAC}-*txs*

pGRG36T-Lac-*txs* was then transformed into WT *P. agglomerans* and the integration procedure was followed. After the curing of pGRG36t-Lac-*txs*, colonies were patched on to LB agar and LB tetracycline 10µg/ml. Colonies that grew on tetracycline were assumed to still harbour pGRG36T-Lac-*txs* and were disregarded. Colonies were then screened by PCR for *txs*, but none were found.

Whilst screening for insertions, several colony morphologies were noticed on the original serial dilution plates: White, yellow, pale yellow and white and yellow sectored colonies Figure *4*.3.

It was hypothesised from the difference in colour and the pattern in which they formed that the white patches could be cells that had lost the carotenoid biosynthesis pathway operon c*rtEXYIBZ*, which is held on the ~500Kb plasmid LPP-1. As it was possible to select for the presence of LPP-1 by adding ampicillin to the media, colonies and differentially pigmented sections of colonies were picked on to LB, LB Amp 200µg/ml and LB tet 10µg/ml Figure*4*.3C&D.

No colonies grew on tetracycline, showing that pGRG36T-Lac-TXS had been cured successfully. Completely white colonies and patches taken from white areas of sectored colonies grew on LB but not on Ampicillin plates. This indicated that white

colonies had lost the native ampicillin resistance; this combined with the loss in pigmentation confirmed the loss of LPP-1. The reasoning behind this loss was that the deletion of LPP-1 deleted *CrtE* (geranylgeranyl diphosphate synthase) and therefore denied *txs* any substrate for the synthesis of taxadiene, potentially pointing to a toxic effect of the production of taxadiene.

Yellow colonies and yellow area of sectored colonies appeared to go paler on Ampicillin than on plain LB. Some yellow colonies appear to be negatively affected by Amp, but this could be due to colonies being made of a mixed population of LPP-1 positive and negative colonies. Alternatively it could be due to the high dose of Ampicillin used (200µg/ml opposed to the usual 50µg/ml) having an adverse effect on growth, even in the presence of resistance genes.

In order to try to prevent the loss of LPP-1 during the integration process, ampicillin (50µg/ml) was added during the overnight growth stage. Two cultures were set up, with ampicillin 100µg/ml and without, these were grown overnight at 30°C and then serially diluted, plated on LB and grown for 48 hours at 42°C. Single colonies were then patched in replicate onto LB and LB amp 200µg/ml (Figure *4*.4). The addition of ampicillin clearly reduced the loss off LPP-1 during culturing as none of the patched colonies appeared to have any deleterious phenotype from being patched onto LB amp 200µg/ml compared to half of those that were grown overnight without ampicillin.

Colonies were screened for integrants using primers Tn*7*l usTS f and Tn*7*R dsTS r, but no positive results were found. It was unclear whether this was due to a lack of integrants or if the primers, which were optimised for use in *E. coli,* were not priming correctly. At this point genome sequence wasn't available so new *glm*S primers could not be designed.

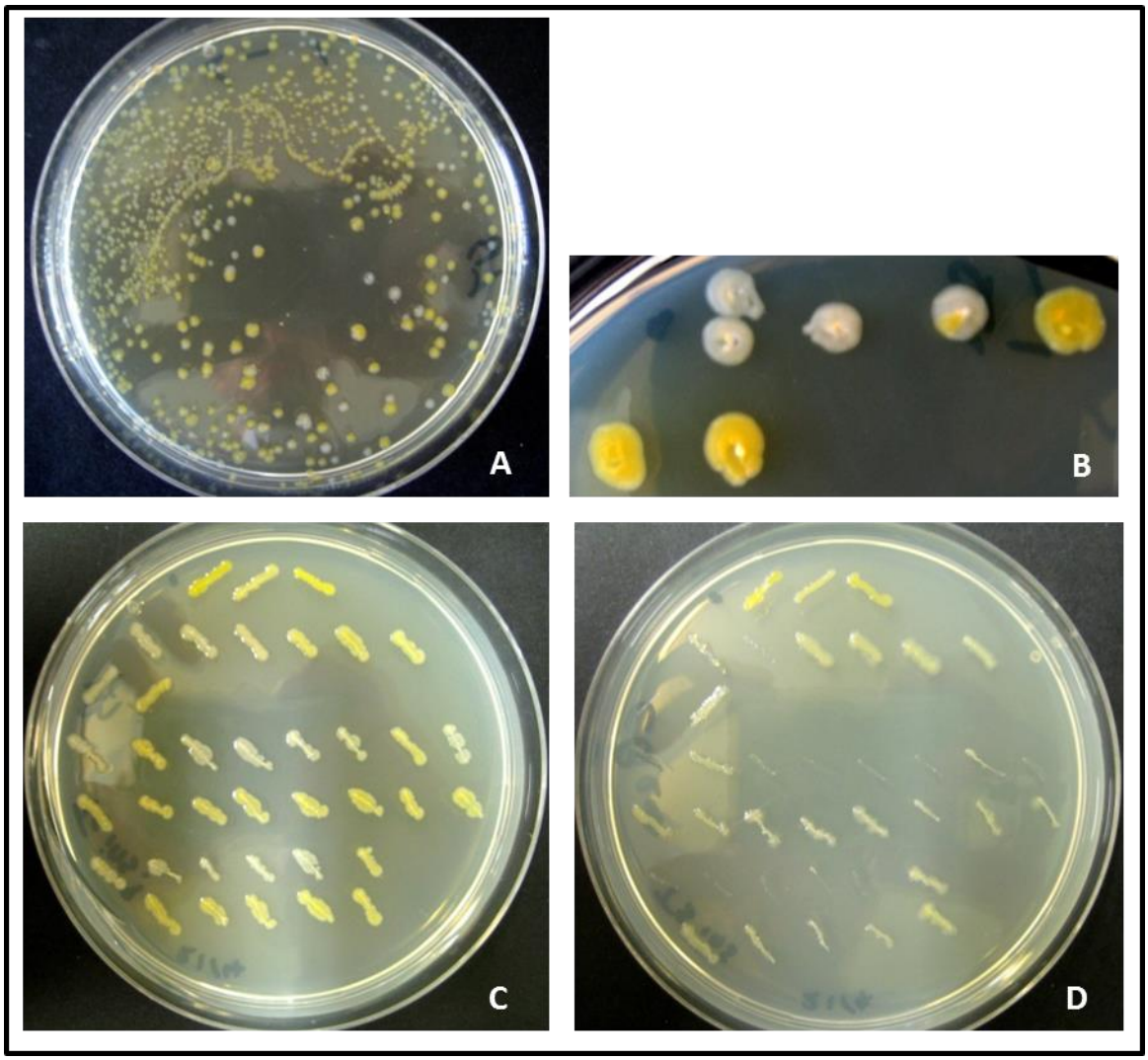


Figure 4-3 Colonies of *P. agglomerans* following attempted chromosomal integration of P_{LAC}-TXS. A. colonies after 48 hours at 42°C to block replication of pSC101ts. B. the range of colony morphologies seen, White, sectored and yellow. C colonies patched onto LB media and incubated overnight at 30°C, D. Same colonies patched as in C. but onto LB ampicillin.

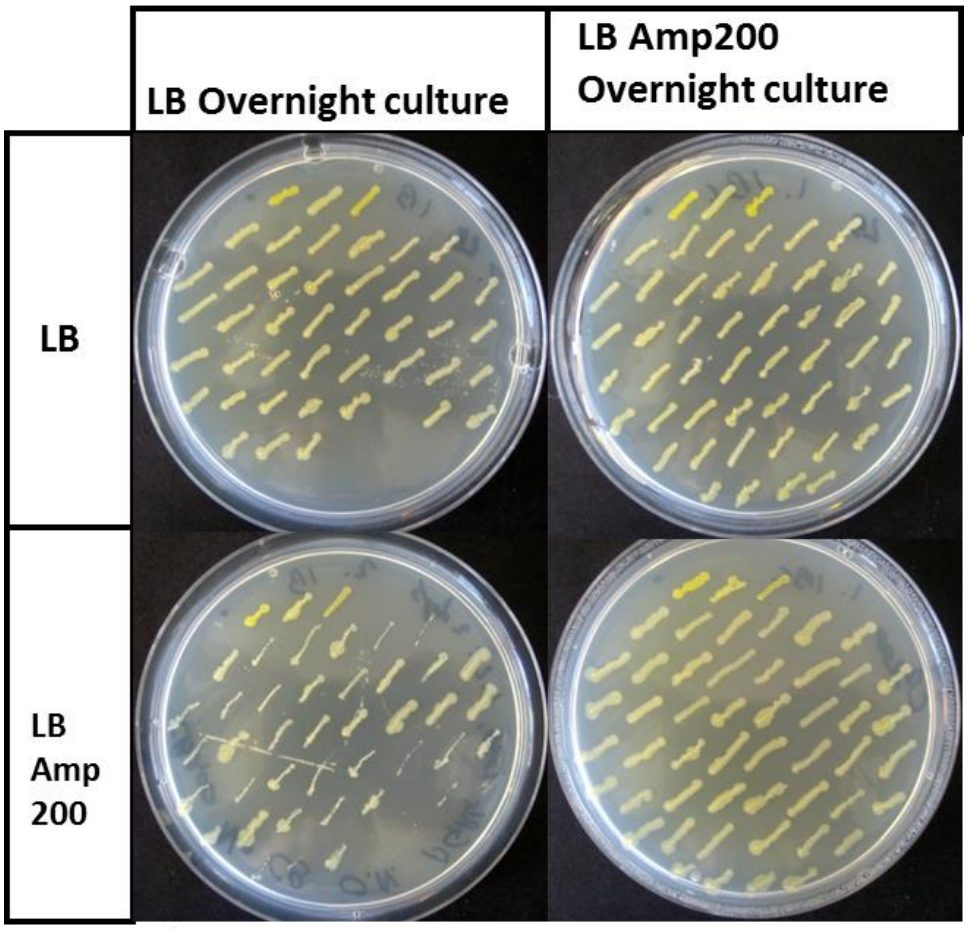


Figure 4-4 Re-attempted chromosomal integration with and without the addition of ampicillin during the culture stage.

4.2.2 pGRG36t-KanR

To confirm that the pGRG36T system could work with *Pantoea*, integration of the kanamycin resistance cassette *KanR* was attempted, as this allowed direct selection and was not expected to have a deleterious effect on the cell. *KanR* was amplified from pDONR221 using primers KanR_XhoI_F and KanR_ANA_R and was ligated into pGRG36T between *Xho*I and *AscI*. Transformants were selected on LB kanamycin 50µg/ml and confirmed by restriction digest.

The plasmid was then extracted and transformed into *P. agglomerans* ΔLPP-1 by electroporation. Cells were plated on LB containing tetracycline and kanamycin and incubated at 30°C for two nights.

The method given by McKenzie & Craig, (2006) was then followed. This involved growing a culture over night at 30°C, then plating it out on LB containing only kanamycin and incubating at 37°C to block replication of pSC101 and therefore only allowing *KanR* integrants to grow.

These colonies grew slower than previous transformations, and after two days two distinct populations were observable, larger and smaller colonies. These were patched onto LB tetracycline and LB kanamycin plates and incubated at 30°C over night. Large colonies were resistant to both kanamycin and tetracycline, whilst the small colonies were only resistant to kanamycin, indicating that the colonies had integrated *Kan*R and been cured of pGRG36t-KanR. Genomic DNA extraction and PCR confirmation still needs to be carried out.

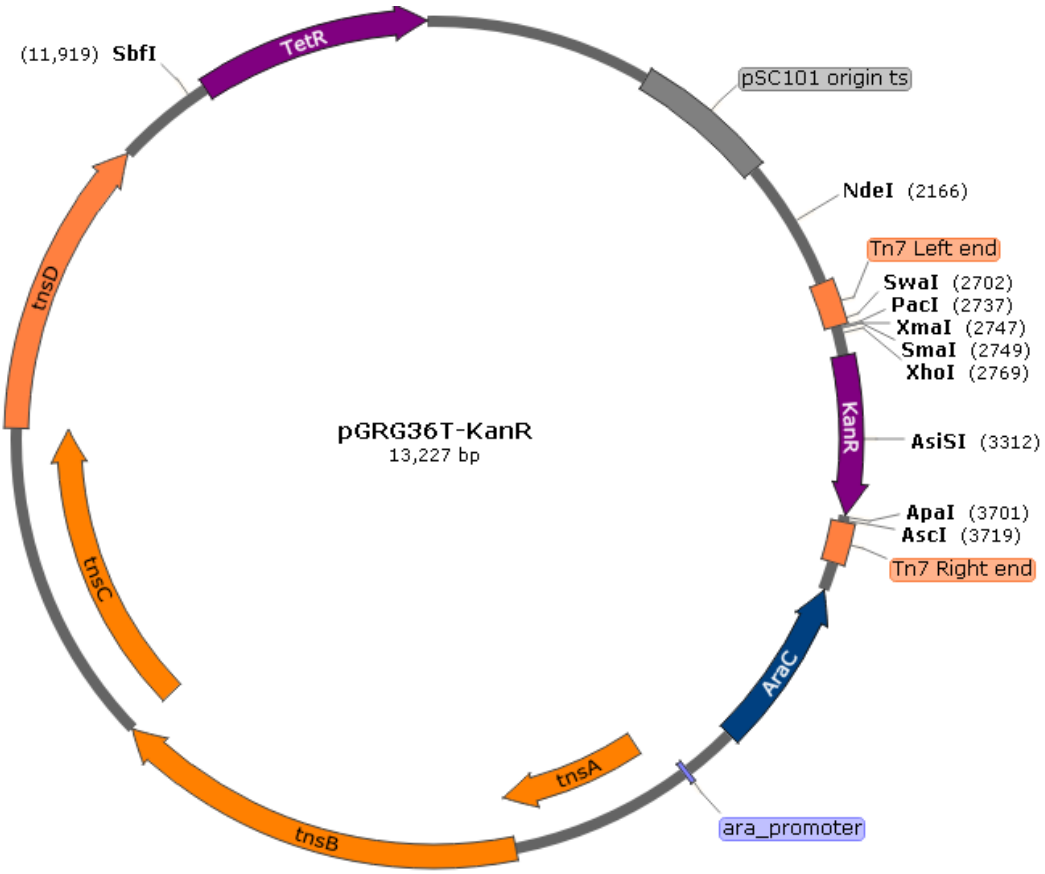


Figure 4-5 Plasmid map of pGRG36T-KanR

4.3 Deletion/ integration system pDEL

For development of a gene knockout protocol for *Pantoea,* the well-established methods of homologous recombination were used as a starting point. These systems uses homologous regions to integrate a deletion construct into the chromosome via positive selection (pop in), and then negative selection is used to force and select for recombination and loss of genetic material (pop out) (figure *4*.6).

The pSC101ts origin of replication was used for the construction of this knockout system, as the temperature sensitive origin of replication allowed for selection of

cells which have integrated the construct. Cells would be grown above 37°C to block plasmid replication and in the presence of tetracycline. Only cells with chromosomally integrated *tetR*, would be able to replicate.

Gene knockout was mediated by production of deletion constructs, these were designed by amplifying between 500bp-1Kb upstream and downstream of a deletion site, these were then ligated together.

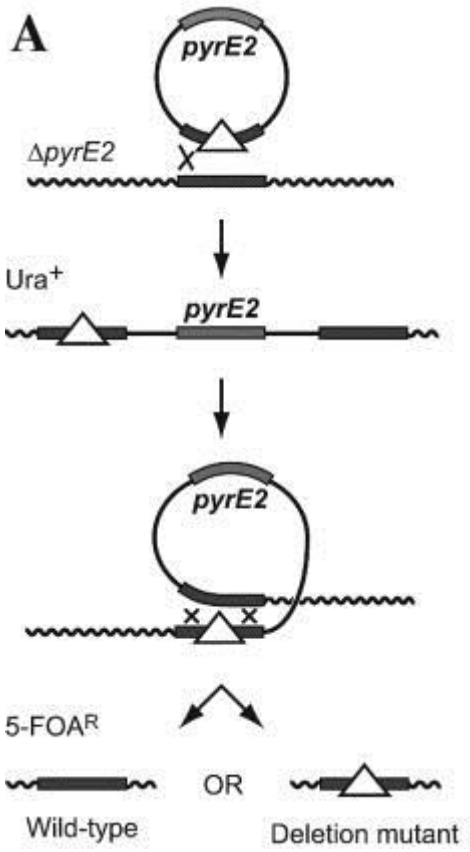


Figure 4-6 Diagram of homologous recombination for the knockout of chromosomal genes with use of a counter selectable marker. ΔpyrE2 strain is auxotrophic for uracil, and so the integration of the plasmid can be selected for by the removal of uracil from the media. The addition of 5-FOA to the media can then be used as a counter selection for the loss of pyrE2 (after a suitable time without selection pressure for its maintenance), this leads to a 50/50 loss or retention of the original construct. Adapted from Allers *et al.*,(2004)

It was decided to use two markers *sacB* and *tetR*. *sacB* is a negative selection marker that encodes for levansucrase which converts sucrose into the toxic compound levan. *tetR* is a dual selection marker; it confers resistance to tetracycline, but also increases sensitivity to nickel ions (Li *et al.,*2013). Initial integration of the plasmid into the chromosome via a single cross is selected by the addition of tetracycline at 42°C (a temperature at which pSC101ts replication is blocked). A second crossover event is then selected for with the addition of sucrose and nickel to the media. The resulting colonies then need to be screened to ensure they are knockout mutants.

4.3.1 Construction

pLT-1 was used as a starting point for the construction (Description of its construction can be found in chapter 7). *sacB* was restricted from pDM4 using *NdeI* and *StuI* (blunt) and was ligated into *NdeI* and *SwaI* (a blunt ended restriction enzyme) of pLT-1, this was then transformed into *E. coli* and colonies were screened by restriction digest the construct was named pDEL-2 (figure *4*.7).

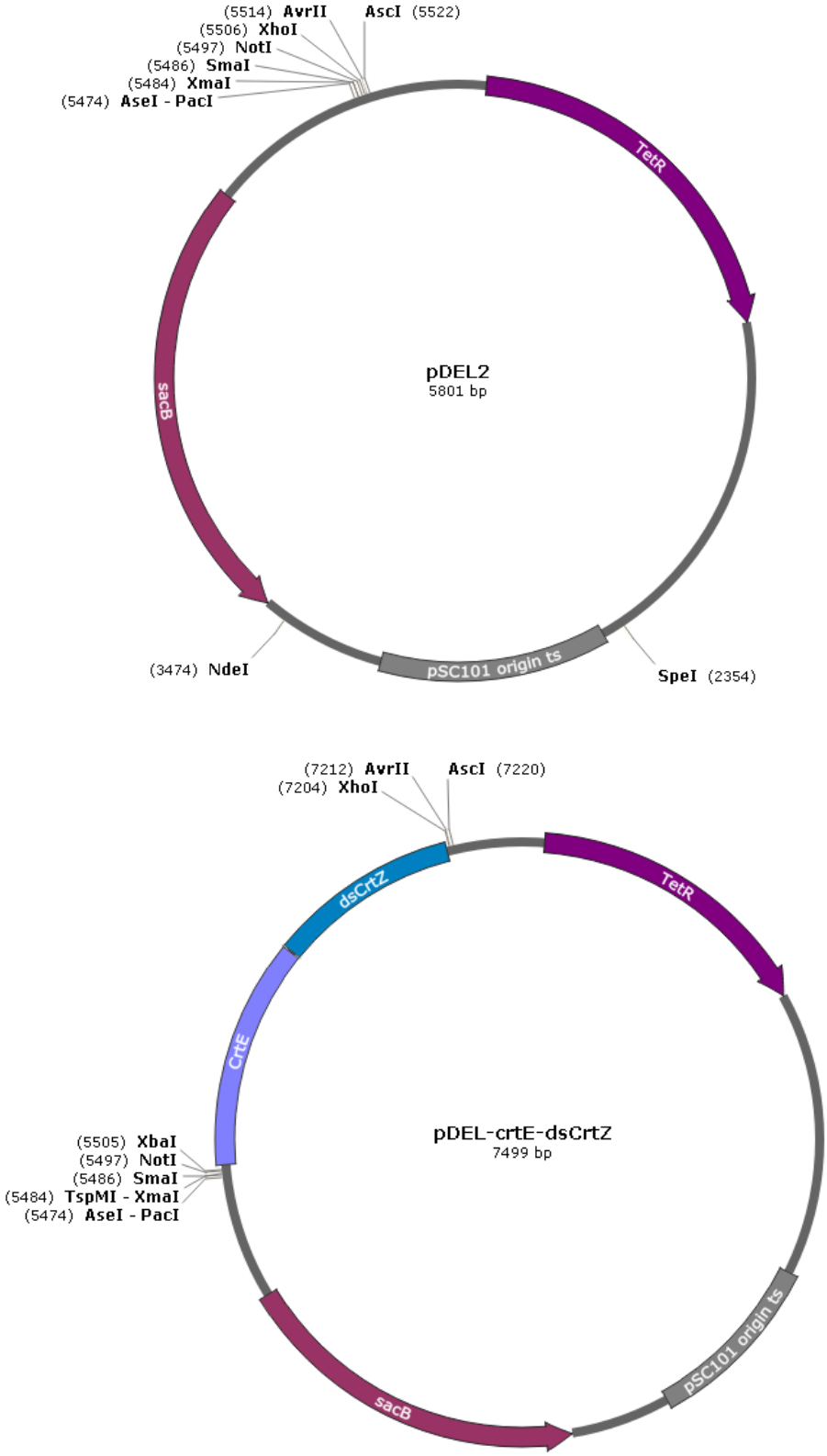


Figure 4-7 Plasmid maps of pDEL2 and pDEL-*crtE*-dsCrtZ

4.3.2 Effectiveness of counter selection

sacB and *tetR* needed to be tested to confirm that the addition of Sucrose (for *sacB*) or Nickel (*tetR*) caused a negative phenotype, and would therefore act as a negative selection pressure. Cultures of *P. agglomerans* or *P. agglomerans* harbouring pDEL-2 were grown over night in LB media at 30°C, the cells containing pDEL2 had tetracycline 10 µg/ml added. Cultures were then diluted 1 in 100 in 200 µl of LB containing: 2 mM, 1 mM, 0.5 mM, 0.1 mM or 0 mM nickel chloride solution ($NiCl_2$), with and without the addition of 6% sucrose, Cells containing pDEL-2 also had 10 µg/ml of tetracycline added to maintain selection of pDEL-2 these were then pipetted into a 96 well plate and grown for 20 hours at 30°C in a Genios Pro plate reader. Figure *4*.8 shows the growth of these cultures over 20 hours.

The presence of *sacB* and the addition of sucrose alone did not produce as strong as a negative selection pressure as was expected. In fact, the addition of sucrose improved the growth in both the plasmid free strain and the strain harbouring pDEL-2. The addition of $NiCl_2$ also causes a difference of growth in favour of the plasmid free strain, but not by very much. However, the combination of both negative selection pressures led to a reduction in the growth of the pDEL-2 harbouring strain, compared to the plasmid free strain, when the $NiCl_2$ concentration was raised to 2 mM. This combination seemed to produce a synergistic effect to retard growth in the plasmid harbouring strain.

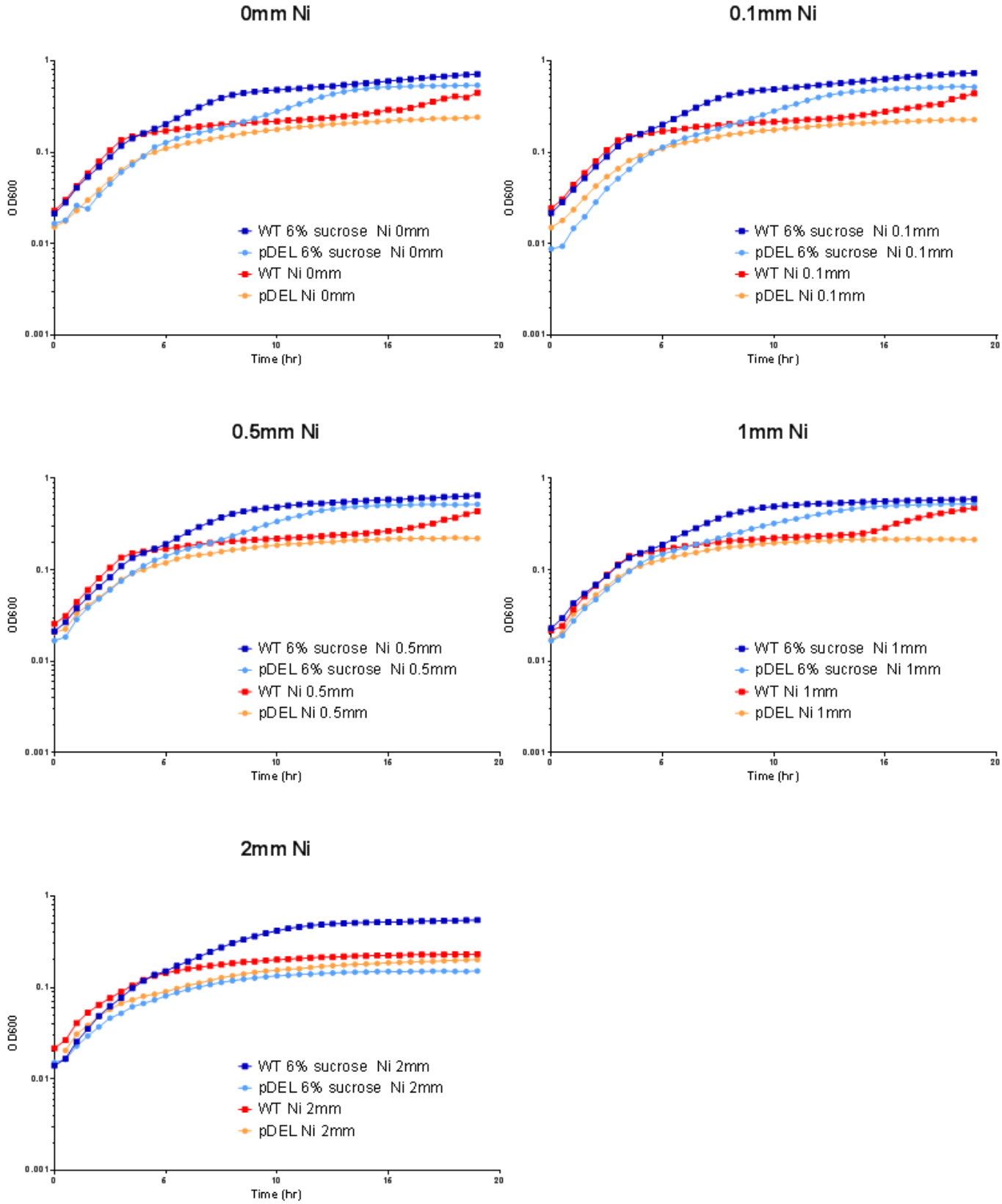


Figure 4-8 The effect of sucrose and increased levels of nickel to the toxicity of *P. agglomerans* harbouring a plasmid carrying te*t*R and *sacB*

4.3.3 Knock out constructs

The phytoene synthase *crt*B was chosen as the target for the development of this knockout protocol; not only is it the first step of the carotenoid biosynthesis pathway, which is competitive with the production of Taxol, its deletion would lead to loss of colony colour. This would make screening for deletions easier and at the same time produce a strain that would be useful for future work. Figure *4*.9 shows the structure of the carotenoid biosynthesis operon in *P. agglomerans*. As *crtE* was required for the production of taxadiene and was and is located at the start of the operon, it was decided to use it as one end of the knockout construct (908bp). The other half was a 753bp sequence, located 677bp downstream of *crtEXYIBZ*.

The DNA fragments containing *crtE* and the downstream region of *crtEXYIBZ* (referred to as dscrtZ) were amplified from *P. agglomerans* genomic DNA using primers crtZ DS_F and crtZ DS_R. These primers incorporated the Restriction sites *Not*I, *Xba*I upstream of the fragment and *Spe*I, *Xho*I downstream. The use of *Xba*I and *Spe*I was inspired by the synthetic biology format of Biobricks (Brown, 2007; Knight, 2007). *Xba*I and S*pe*I are compatible sites, and can therefore be ligated together to form a scar site which cannot be restricted by either enzyme. This was done in order to make this system modular. Constructs could be made with the same sites on either end, and quickly be flanked by the homology arms, and then be integrated into the genome.

The *crt*E fragment was restricted with *Not*I and *Spe*I, whilst the dscrtZ fragment was restricted with *Xba*I and *Xho*I, these were then ligated into pBT-GFP which had been restricted with *Not*I and *Xho*I (removing GFP). The ligation was transformed and colonies were screened by restriction digest and the knockout construct was Sanger sequenced using primers pGBT_F and pGBT_R. This plasmid was named pBT-CrtE-dsCrtZ. The CrtE-dsCrtZ construct was then cut from pBT-CrtE-dsCrtZ with *Not*I and *Xho*I and ligated into pDEL-2 using the same sites. Transformants were screened by restriction digest and named pDEL-CrtE-dsCrtZ.

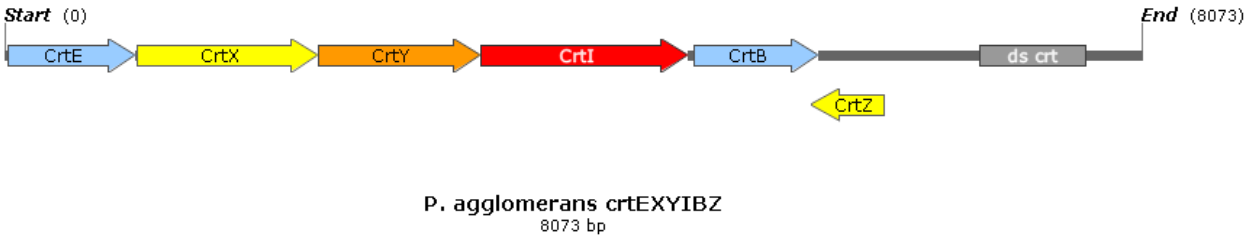


Figure 4-9 *P. agglomerans* zeaxanthin biosynthesis operon *crtEXYIBZ*

4.3.4 Attempted deletion of *crt*XYIBZ

pDEL-CrtE-dsCrtZ was transformed into *P. agglomerans* and recovered on LB agar containing tetracycline and ampicillin at 30°C overnight. Transformation was confirmed by restriction digest of plasmid DNA isolated from transformant colonies. A single colony was then inoculated into LB containing tetracycline and ampicillin and grown overnight at 30°C. 100µl of this culture was then added to 10ml of LB containing tetracycline and ampicillin (1:100 dilution) and grown for 8 hours at 37°C. This culture was then serial diluted onto LB agar containing ampicillin and tetracycline and grown over night at 37°C. Colonies were then patched onto LB amp 100µg/ml at 37°C overnight; these were then patched onto LB agar with ampicillin, $NiCl_2$ 2 mM and 6% sucrose at 37°C.

Every day for a week after this initial patching, patching was repeated from the previous day's plate, all plates were then incubated at 37°C. Throughout the week pale colonies or patches were streaked onto LB ampicillin agar at 30°C to isolate colourless individuals. Figure 4-10 shows the resulting patches after 5 days of incubation. Patches had irregular borders compared to wild-type *P. agglomerans* plated on LB plates, and showed variation in pigmentation. White areas were streaked onto LB ampicillin agar plates and grown at 30°C overnight. Smaller pale sections (Figure 4.10 A) grew to produce patches that appeared to be wild-type, yellow with a regular border. Larger areas of white growth did not grow when

patched on to LB ampicillin agar, indicating that these had lost LPP-1 through heat stress (Chatterjee & Gibbins, 1971). This loss was made possible by the degradation of ampicillin in the media, both enzymatically and by heat, therefore removing the selection pressure on LPP-1. Following several attempts there were unfortunately no unpigmented, ampicillin resistant colonies isolated through this process.

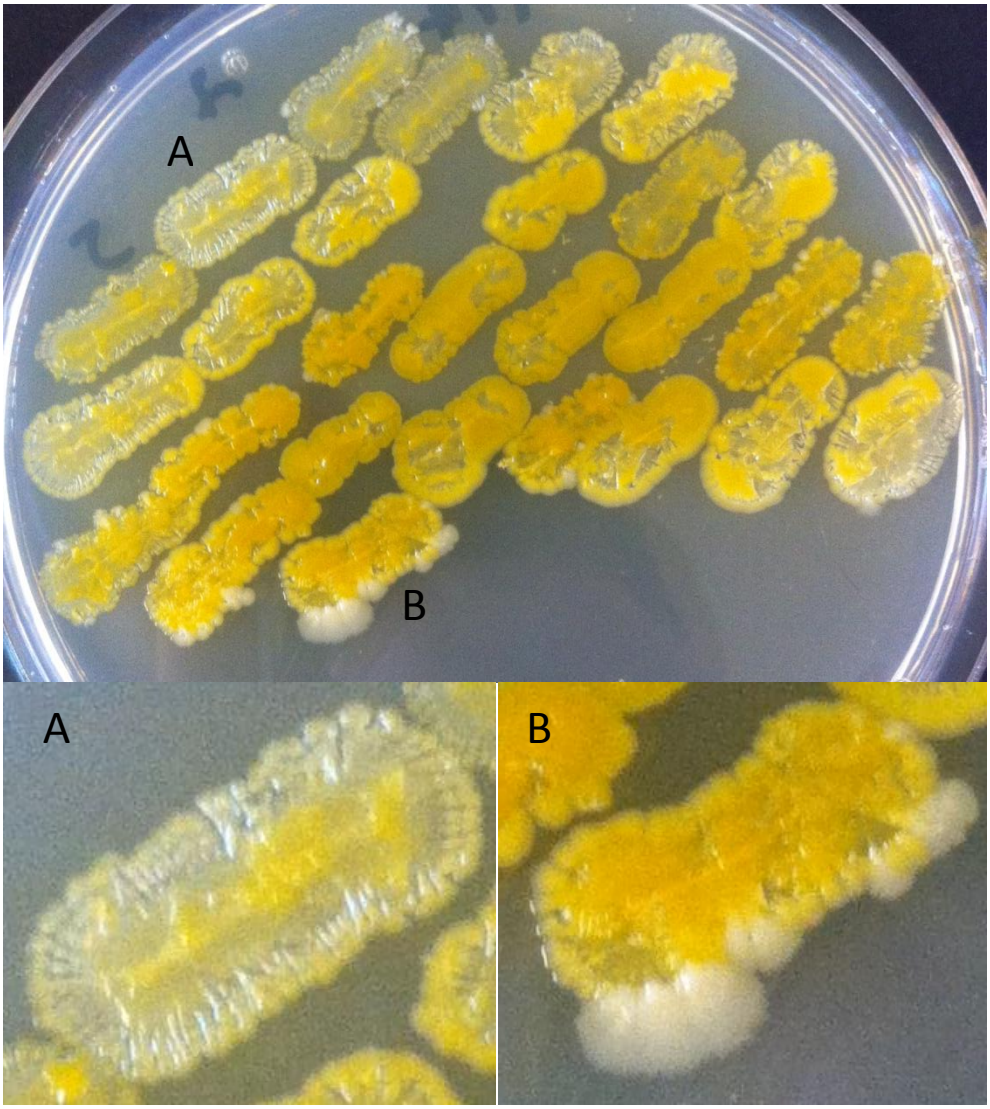


Figure 4-10 Patched out colonies being screened for the deletion of *crt*XYIBZ, Bacteria developed a rough appearance and showed some signs of differences in pigmentation (seen clearly in patch A), however these did not persist when re-patched out. Patch B shows clear signs of unpigmented bacteria, however when re-patched onto LB ampicillin there was no growth, indicating that these unpigmented cells had lost the LPP-1 plasmid.

4.4 Discussion

Both the integration and gene deletion methods require a greater selective pressure in order to improve their efficiency. For the pGRG36T Tn*7* integration the use of the *kan*R antibiotic resistance marker appears to have improved integration, however there are still several colonies forming that have not integrated and are surviving due to a plasmid copy of KanR as the pSC101ts plasmid is not completely blocked from replicating at 37°C. Also the size difference observed between integrants and those still harbouring pGRG36t-KanR occurs because KanR is more effective at in the higher copy number state provided by being hosted on pSC101ts, compared to being on the chromosome. This inefficiency could make the future integration of taxane genes difficult, if as it appears in this section, there is a selection pressure against integrating them. However, being able to screen for true integrants by patching onto LB tetracycline in some way overcomes these issues. An alternative way to improve this would be to add a reporter gene to the backbone of pGRG36T, for example the red fluorescent protein (RFP), this would make colonies that have not been cured of pSC101 visible and easily screened against. Following the proof of concept that pGRG36T can be used to integrate genes into the chromosome of *P. agglomerans* the next step was to re-attempt to integrate Taxadiene synthase alongside KanR.

Subsequent to carrying out the attempted integration of *txs*, it was discovered that the *Swa*I site should not have been used to clone into. This cut into the Tn*7* left arm and reduced it by 45bp, (19.4%) and will have had an effect on the efficiency of integration. However taking this into account, it was still considered that the effects seen on the loss of LPP-1 indicate that the presence of TXS would select against integration of this construct into the chromosome by Tn*7*.

For the gene deletion method using the pDEL system, there are several possibilities as to why deletion of *crt*XYIBZ was not achieved. As shown during the integration of *KanR* using pGRG36T, despite the pSC101 origin of replication being temperature sensitive and being grown at a prohibitive temperature, colonies still had *KanR*

showing that there is some limited replication of pSC101ts at 37°C. This limited replication could have diluted the initial selection pressure for individuals that had integrated pDEL-CrtE-dsCrtZ into the Carotenoid operon. Secondly there is a possibility that the loss of carotenoid biosynthesis has an associated cost, especially when one of the counter selection agents $NiCl_2$ is indirectly linked to increasing oxidative stress (Macomber & Hausinger, 2011).

A more effective counter selection system is required. Higher levels of $NiCl_2$ could be attempted, but preliminary work with a culture of wild-type *P. agglomerans* with 5mM $NiCl_2$ failed to grow, and so there is not much scope for further investigation in this area. There are several other counter selection systems which could be investigated such as *pyr*E, *phe*S or *thy*A (Reyrat *et al.,*1998). As previously mentioned the pSC101ts plasmids propensity to not be completely cured when grown at 37°C also compounds any issues borne out of the selection system not being 100% effective. An alternative to pSC101 would be to move to a λpir based plasmid system like R6K origin, this origin can only replicate in strains containing λ*pir* (Skrzypek, *et al.,* 1993). Additionally the knockout construct could flank a selection marker such as *KanR.* This could then be used to force selection towards integration and then to retaining the knockout construct rather than the original gene during the subsequent homologous recombination step.

In this attempted gene knockout, the fact that the genes were plasmid based may have complicated the procedure further. As *crtEXYIBZ* is on LPP-1 this leads to a higher level of ploidy than if the knockout target was chromosomal. Future knockout methods should focus on attempting the deletion of a chromosomal gene to verify the technique and then apply it to editing LPP-1.

Alternative methods of gene deletion may involve the use of Lambda red recombination. This method uses a deficient prophage that expresses a number of

proteins that facilitate the integration of linear dsDNA via homologous recombination. Gam, inhibits the host nucleases, which protects linear dsDNA from cell mediated degradation. Exo, degrades linear dsDNA from the 5' end, leaving 3' overhangs. Beta, binds to ssDNA to mediate pairing. This method has been used extensively in *E. coli* to produce libraries of knockout mutants (Datsenko & Wanner 2000) (Yu *et al.* 2000). This system has also been adapted for the use in *P. ananatis*, however it required screening for mutants which did not find expression of the system toxic (Katashkina *et al.* 2009).

During the time spent on this project, the technology of CRISPR has developed from a lab curiosity to a functional tool with a wide range of applications in biology, and is accessible for use in a variety of organisms (Choi & Lee, 2016). Also adding to its success is its accessibility and low price point compared to other similar techniques like zinc nuclease fingers. Many plasmid based systems have been produced for use in bacteria, and several have been made available by plasmid sharing services such as addgene (www.addgene.com). CRISPR/Cas9 system introduces double strand breaks (DBS) at specific sites, based on a RNA guide sequence. This essentially allows the user to programme a novel restriction enzyme. The length of the recognition site is also long enough that it is possible that in an entire genome, the site only appears once. This ability produces an extremely powerful selection pressure, as unrepaired DBS are lethal. The CRISPR system can be set up to restrict in a gene of interest, this can then not be repaired unless the original recognition site is destroyed. DNA can be transformed in along with the CRISPR plasmids which complement the restriction site, through homologous recombination this can then be incorporated and therefore remove the restriction site. Without this removal of the restriction site the cell will die. This is a basic explanation of the uses of CRISPR for gene deletion, but further uses are reviewed by (Choi & Lee, 2016).

This technique is extremely powerful because of the strong selection pressure it exerts, but also because of the versatility it demonstrates to be applied in different

situations. It can be used for gene integration, gene deletion as well as transient gene silencing by using Cas9 with mutated nuclease domain, these bind to the recognition site but do not restrict (Larson *et al.*, 2013). This technique known as CRISPRi, can be used to interfere with transcription of a gene or genes allowing for the investigation to genes which are essential, and therefore the deletion of which are lethal. Alternatively this method could be used to switch off biosynthetic pathways in order to divert resources down an alternative pathway.

There are however a few pitfalls with using CRISPR. There are currently no reports in the literature of CRISPR being used in *Pantoea*, however it has been used extensively in *E. coli.* From the literature the chances of it not working in *Pantoea* are slim, and initial issues can be worked through. The most likely problems would come from ensuring that there is no cross reactivity with any cut site selected or non-specific restriction. Overall CRISPR would be the next tool for investigation to continue to edit the genome of *P. agglomerans*.

Chapter 5 TAXANE BIOSYNTHETIC OPERON

5.1 Design of the synthetic taxane operon

After the initial project design, it was decided to have a synthetic taxane operon commercially synthesised was the best starting point. Having the operon synthesis *de novo* offered several advantages: Firstly it could be codon optimised for *P. agglomerans*; *Taxus* has a different preferred codon usage to *P. agglomerans*, using several codons which are rare for *P. agglomerans*. Having the operon synthesised meant that these could be replaced. Secondly restriction enzyme sites could be removed from the sequence by choosing alternative codons. Thirdly it was much easier to make the operon modular, the order of genes could be decided and sites put in for easy rearrangement.

Figure *5*.1 shows the original plan for the operon; it was designed to be flanked by 2 *Not*I sites so that it could be restricted out and ligated straight into several plasmids used in previous studies from our group. If directionality of the insert was required *Sbf*I *Pst*I and *Xho*I could be used to cut the insert to ensure ligation in a specific orientation. Sites were also added for the quick deletion of genes from the operon. Flanking each gene were a pair of compatible sites which could be cut with two separate enzymes, giving compatible overhangs that could be ligated to produce a scar site which cannot be restricted with either of the original enzymes. These pairs were: for *txs Nde*I and *Ase*I, for 10β *Pci*I and *Nco*I, 5α *Nhe*I, *Spe*I (in between the hydroxylase and acetyltransferase) and *Xba*I.

Protein sequences from *Taxus cuspidata* (the Japanese yew) were used as a template, taxadiene synthase had the first 60 amino acids removed, in line with Ajikumar *et al.*, (2010), to remove the signal peptide, the other three genes did not have any sequence removed as the signal peptide detection software signal P (Emanuelsson *et al.,* 2007) did not detect any. DNA 2.0 software was used to design the operon; each protein sequence was put through the codon optimisation software

using the *Pantoea agglomerans* codon preference table from www.kazusa.or.jp/codon. These were then assembled in order and had a ribosome binding site (RBS) added upstream. These RBS were based on sites used in previous studies expressing pathways in *E. coli,* the Shine Dalgano sequence (AGGAGG) was used followed by 8 nucleotides which were varied for each gene, to reduce homology. The restriction enzyme sites listed in Table 1 were then removed by changing codons whilst still attempting to maintain the most favourable codons.

Finally the operon was flanked by the Gateway™ cloning sites attL1 and attR2, this was to allow for recombination in to gateway vectors alongside a variety of promoters and reporter constructs, however due to production issues (see below) these functions were never used.

Biomatik were commissioned to produce the construct, but experienced difficulty in cloning the full operon. They were able to produce the operon in two halves, pBMH-AB and pBMH-CDE. pBMH-AB contained attL1, *txs*, and the first 880bp of 10β and pBMH-CDE contained the last 580bp of 10β, both 5α genes and attR2.

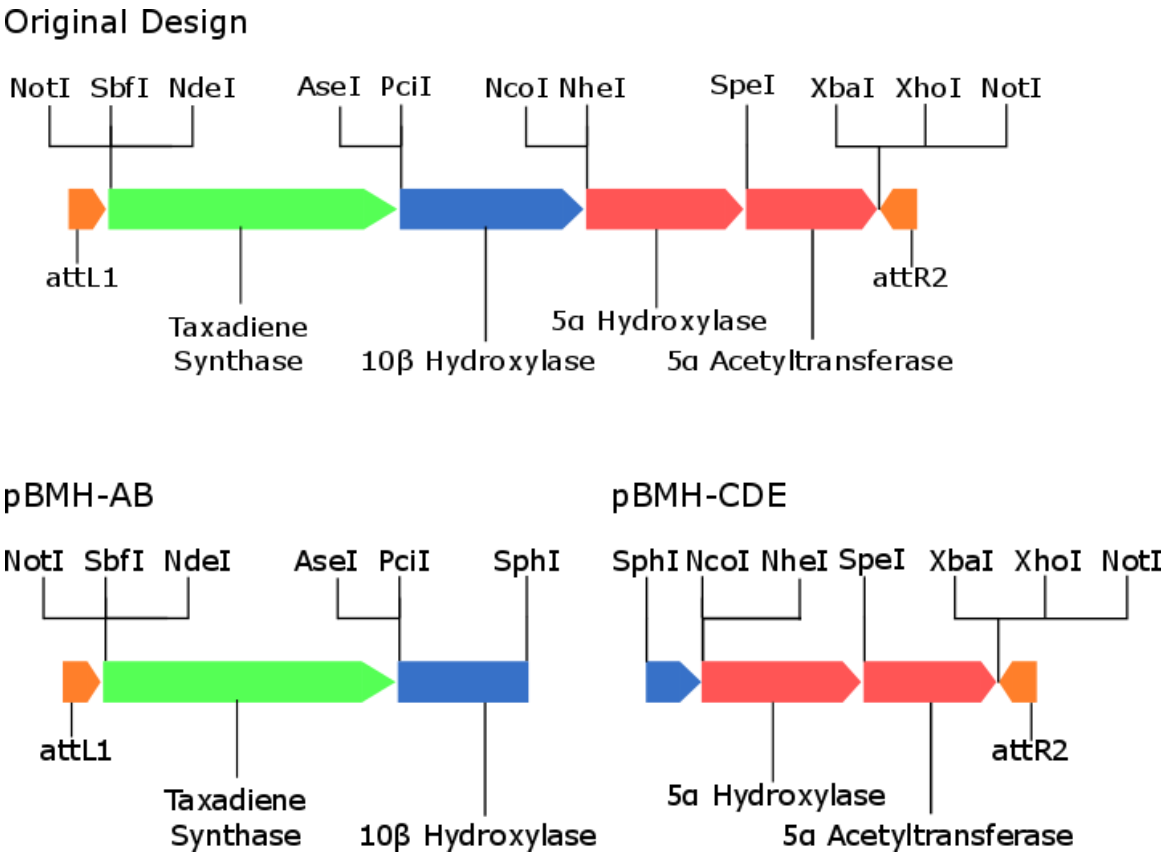


Figure 5-1 Organisation of the Taxane synthetic operon. Original design of the Taxane operon consisting of Taxadiene synthase (*txs*), Taxane 10β Hydroxylase (10β), taxadiene 5α hydroxylase and taxadien-5alpha-ol O-acetyltransferase (5α acetyltransferase)(5α) flanked by the Gateway cloning att sites attL1 and attR2. pBMH-AB and pBMH-CDE show the two halves made by Biomatik. Attempts to ligate the two halves of the operon at *Sph*I site produced no correct clones.

5.2 pBMH-AB and pBMH-CDE transformation

Upon arrival of the plasmids they were transformed into *E.coli* Top10 via electroporation and transformants were selected for on kanamycin. Single colonies were then grown in LB broth at 37°C and the plasmid DNA was extracted. These were then confirmed by restriction digest to check the plasmid size, however as will be

discussed further in section *5*.2.3 this was carried out incorrectly and pBMH-CDE was in fact larger than expected, a mistake made by myself and the manufacturers.

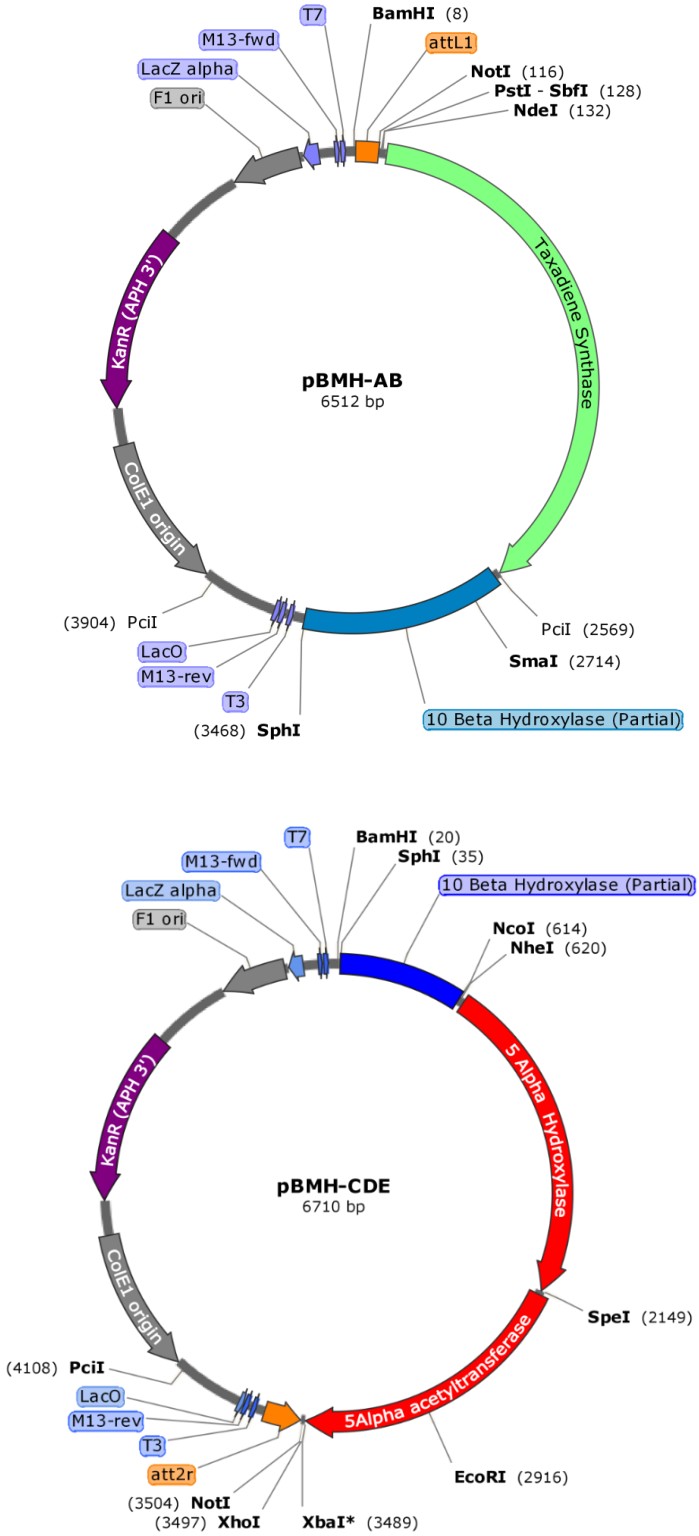


Figure 5-2. Map of pBMH-AB and pBMH-CDE from Biomatik. pBMH is a high copy number plasmid with a kanamycin resistance marker used for cloning synthetic constructs produced by Biomatik. pBMH-AB contains the first half of the synthetic taxane operon, up to the SphI site in 10β hydroxylase. pBMH-CDE contains the operon downstream of the SphI site.

5. 2.1 Full operon construction.

Figure *5*.2 shows the plasmid maps of pBMH-AB and pBMH-CDE. Attempts were made to ligate the whole operon together as Biomatik had attempted to. Ligation of the whole operon was via the *SphI* sites. pBMH-AB and pBMH-CDE were restricted with *BamH*I and *SphI*, and purified by gel extraction. The *txs* fragment of pBMH-AB was ligated into pBMH-CDE and then transformed into *E. coli* via electroporation. No colonies were recovered.

5.2.2 Deletion of 10β hydroxylase

As 10β was the site of the break in the operon it was considered a likely candidate for the cloning issues. This was supported by the fact that *txs* and 5α had previously been co expressed in *E. coli* by Ajikumar and colleagues (2010). In order to test this, it was decided to attempt to clone the *txs* and 5α fragments without the 10β hydroxylase. Cloning was complicated by the fact that the *Pci*I site required to delete 10β was also present near the origin of replication. pBMH-AB was restricted with *Pci*I and pBMH-CDE was restricted with *Pci*I and *Nco*I. Both were run on a gel and the backbone of pBMH-AB and the 5α hydroxylase fragment from pBMH-CDE were isolated and ligated together.

The only construct detected from this ligation were self-re-ligated pBMH-AB, showing that the *Pci*I sites was in fact outside of the origin of replication, and that this strategy was flawed and a different approach was required.

5.2.3 Unexpected band patterns from *BamH*I and *Nco*I restriction digest

Instead of trying to insert the 5α portion of the taxane operon into pBMH-AB, the opposite was attempted. pBMH-AB was to be cut with *BamH*I and *Pci*I (and *Nsi*I to remove the backbone), while pBMH-CDE was restricted with *BamH*I and *Nco*I. These were then going to be ligated; however there was a discrepancy between the data and the map when restricting pBMH-CDE. As shown in Figure *5-3*.3 restriction with

*BamH*I and *Nco*I produces four bands: 4-5Kb, >3Kb, 500-750bp and <250bp. opposed to the expected 6.1Kb and ~600bp bands.

To investigate whether this was an enzyme based issue, due to star activity or an error in reaction set up, several reactions were set up in parallel. Figure *5-3*.3 and Table 7 pBMH-CDE restriction digestion show the restrictions used and the resulting band pattern.

Restriction Enzyme Used	Expected bands	Observed bands	Total
Uncut			
*Nco*I + *BamH*I	6116bp, 594bp	4Kb-5Kb, 3Kb-4Kb, 500bp-750bp + >250bp	7.75-10
*BamH*I	6710bp	8Kb-10Kb	8-10
*Nco*I	6710bp	~5Kb , ~4Kb + >250bp	9.25
*Nco*I + *Not*I	3821bp, 2889bp	4Kb, 3Kb, 1-1.5Kb + >250bp	8.25-8.75
*Nco*I-*HF* +*BamH*I	6116bp, 594bp	4Kb-5Kb, 3Kb-4Kb, 500bp-750bp + >250bp	7.75-10

Table 7 pBMH-CDE restriction digest Expect and actual bands derived from several digests of pBMH-CDE.

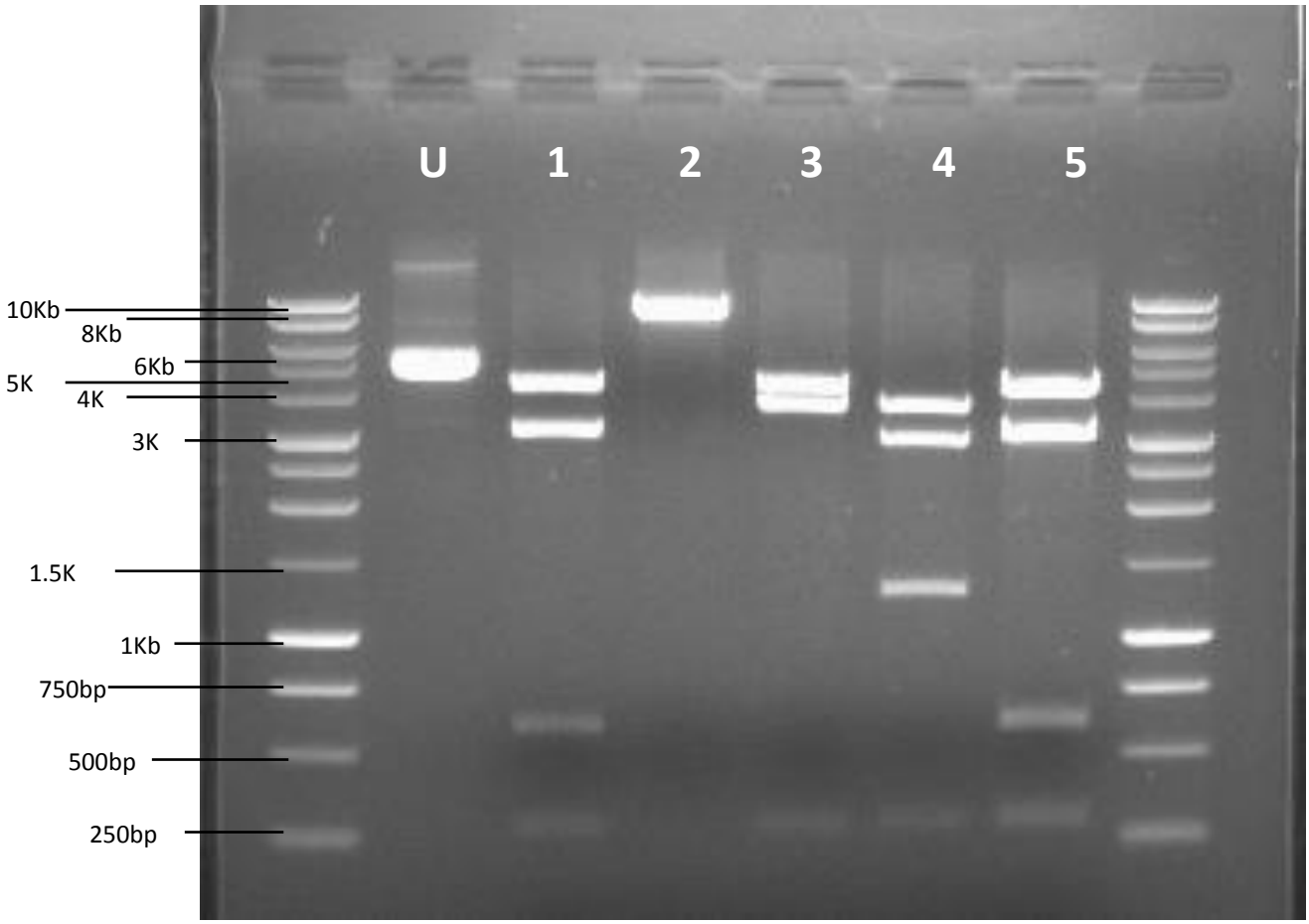


Figure 5-3. pBMH-CDE *Nco*I restriction digest. pBMH-CDE was restricted with a combination of enzymes following the discovery of an unusual band pattern following digestion with *Nco*I. pBMH-CDE was cut with: 1. *BamHI* and *Nco*I, 2. *BamHI*, 3.*Nco*I, 4. *Not*I and *Nco*I, 5. *BamHI* and *Nco*I-HF and U. uncut.

From these results the discrepancy was clearly not due to an error in the handling of *Nco*I, as both the old stock and the new high fidelity version cut to produce the same pattern. Restriction with *BamH*I alone produces a single band, showing that there are not multiple plasmids of varying sizes. The pBMH-CDE plasmid appears to be between 8Kb and 10Kb, compared to the designed 6.7Kb. The combined restriction digest with *Nco*I and *BamH*I and *Nco*I and *Not*I show that *Not*I and *BamH*I are on separate bands of the *Nco*I digest, with *Not*I being on the larger band. *BamH*I and *Not*I are not present in duplicate.

5. 3 Separation of the taxane genes by PCR

Following the realisation that pBMH-CDE was larger than expected a new strategy for building the operon was devised. As the operon was designed to be modular it seemed logical to take all three separate parts of the operon (*txs*, 10β, 5α) into separate plasmids, which could then be re-combined, by co-transforming into *P. agglomerans*.

5.3. 1 pSU-TXS, pBT-5α, pBT-10β and pBT-5α10β

For co-transformation two plasmids were required with compatible origins of replication and different selection markers. Work carried out in chapter 7.3.2 shows that pSU9 and pBT-GFP could both be maintained in *P. agglomerans*. pSU9 carries the chloramphenicol resistant gene CmR (*CatR*) and uses the p15A origin of replication. *txs* was ligated into pSU9 using the *Pst*I and *Sma*I sites (Figure *5-4*). This construct was then transformed into *E. coli* by electroporation, recovered for 1 hour at 37°C and then plated onto LB chloramphenicol and incubated at 37°C overnight. Colonies were screened by restriction digest and the sequence confirmed by Sanger sequencing.

The other taxane genes 5α and 10β, as well as *txs* were then ligated separately and in combination in to a second plasmid pBT-GFP (construction described in chapter 4), which was custom built for this project.

pBT-TXS was constructed by PCR amplification of Taxadiene synthase from pBMH-AB with Primers M13-f and TS_R_*Nco*I_*Xho*I, this fragment was then restricted with *Not* and *Xho*I and ligated into the same sites in pBT-GFP. pBT-5a was constructed by amplifying 5α hydroxylase and acetyltransferase from pBMH-CDE with primers M13-f and M13-r, this product was then digested with *Nco*I and *Xho*I and ligated into pBT-GFP using the same sites. pBT-10B was constructed by first restricting pBMH-AB and pBMH-CDE with *Sph*I and then ligating them together, this reaction was then used as

a template for 10β hydroxylase PCR amplification using primers 10B_f and 10B_R. The 10β hydroxylase fragment was then cloned into pBT-GFP using the *Sbf*I and *Xho*I sites. pBT-5a10B was constructed by amplifying 10β hydroxylase from pBT-10B using primers 5a_10B_GA and 10B_GBT_*Xho*I_GA. These produced a product with 28bp up stream and 40bp of downstream homology to pBT-5a relative to the *Xho*I restriction site. This fragment was then inserted into pBT-5a via Gibson assembly after being restricted with *Xho*I. These constructs were transformed into *E. coli* by electroporation and recovered for 1 hour at 37°C 150 RPM, they were then plated onto LB tetracycline and incubated at 37°C overnight. Colonies were then screened by restriction digest to confirm the correct construct, following this their sequences was confirmed by Sanger sequencing, figure *5*.5 shows the plasmid maps.

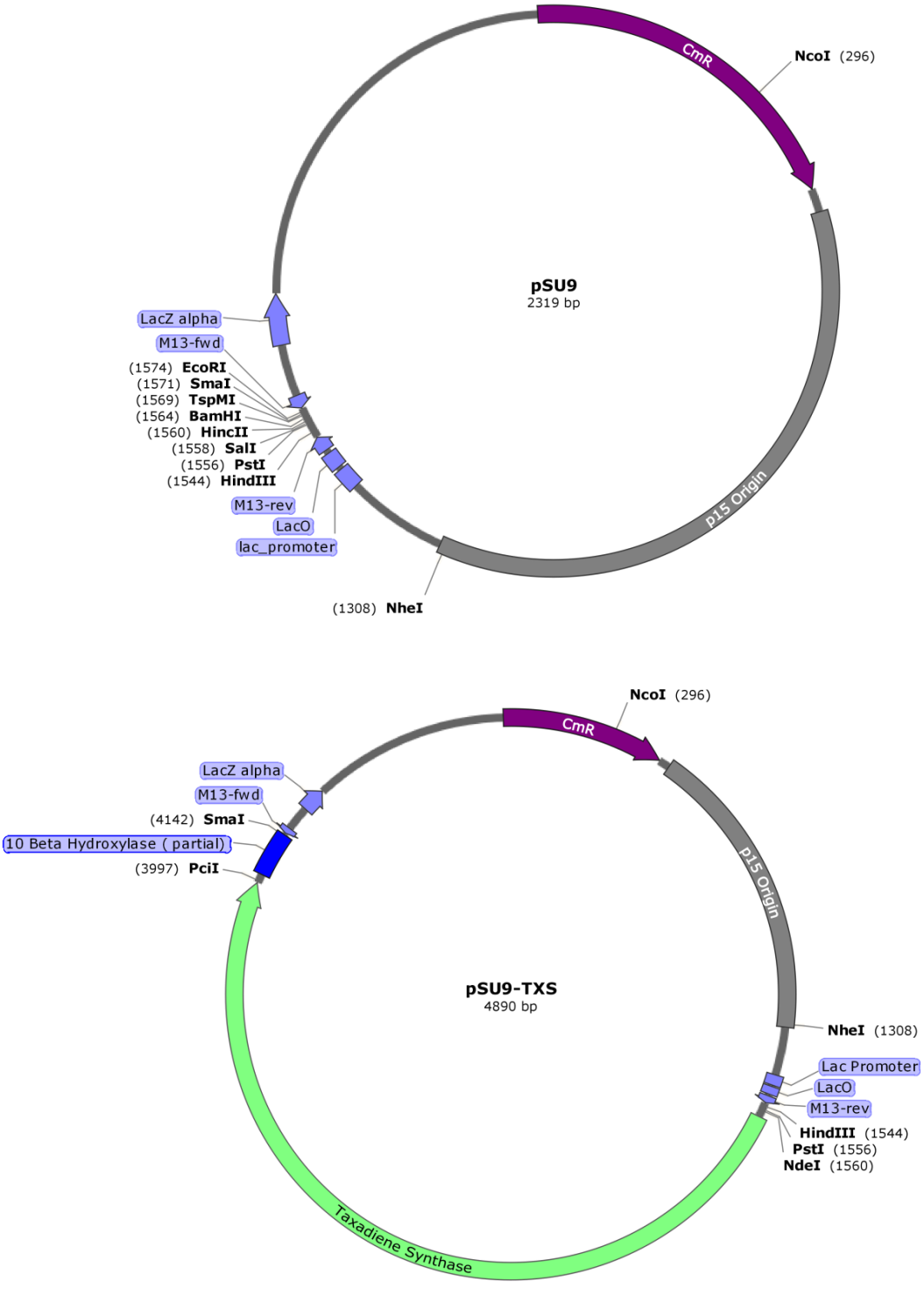


Figure 5-4 pSU9 and pSU9-TXS. pSU9 contains the p15A origin of replication with the *cat* chloramphenicol resistance marker with a MCS downstream of P_{LAC}. pSU9-TXS was created by restriction digest of pBMH with *Pst*I and *Sma*I to liberate the *txs* fragment and then ligate it into pSU9 previously restriction digested with *Pst*I and *Sma*I.

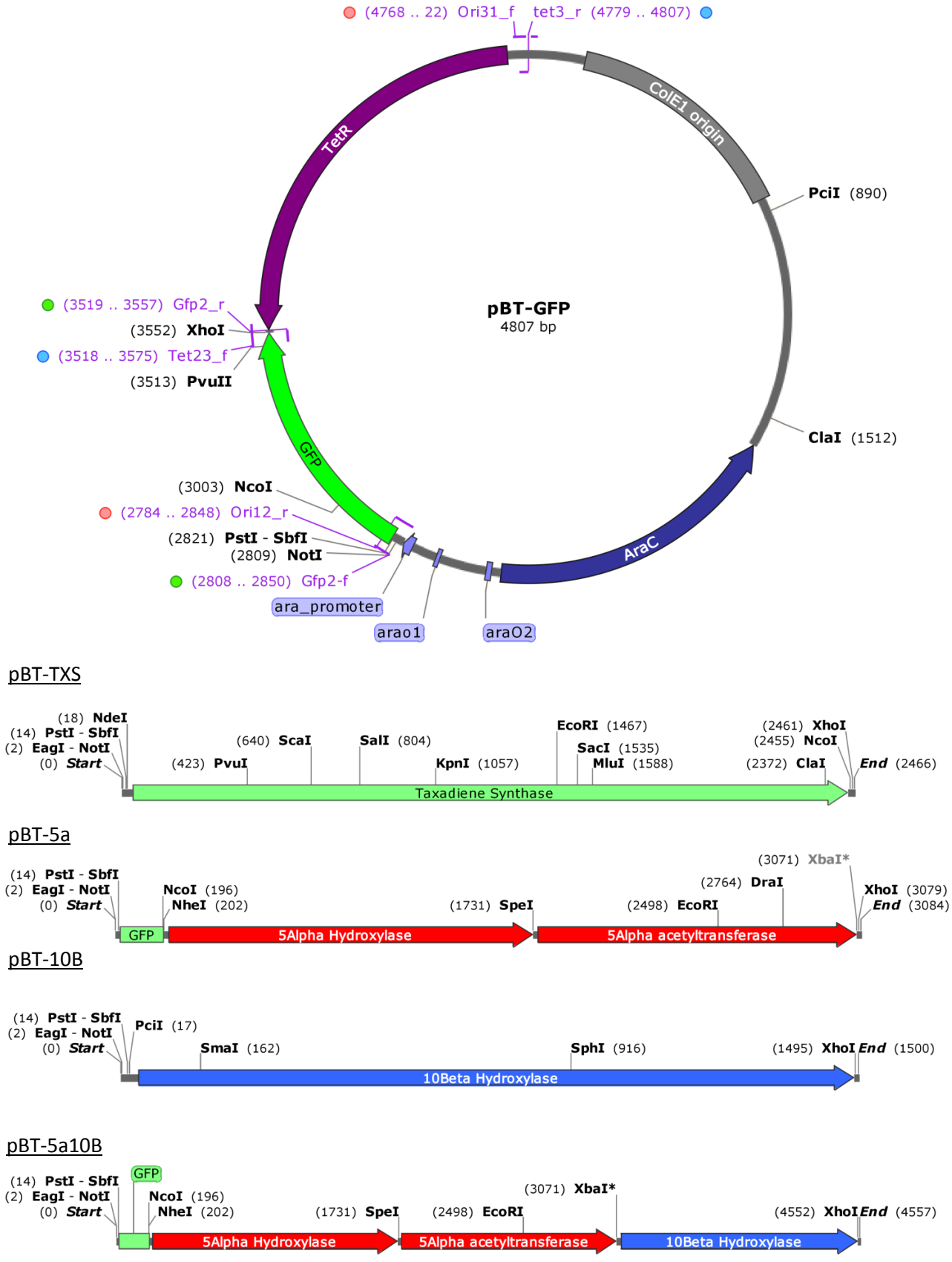


Figure 5-5 Plasmid map of pBT-GFP pBT Taxane variants. pBT-GFP was produced by Gibson assembly, the *Not*I and *Xho*I cloning sites were then used to replace GFP with the individual units of the synthetic taxane operon, TXS, 5α, 10β and the combination of 5α10β.

5.4 Discussion

This chapter covered the initial struggles encountered when working with the taxane genes and the methods used to try to tackle these issues. The ability to remove each of the taxane gene units and clone them into a separate plasmid shows the use of modularity in designing synthetic operons as it allows for an efficient disassembly and reassembly of constructs. This chapter was also formative for the strategy used to take the project forward. Instead of focusing on trying to get the full operon into *P. agglomerans*, the focus was redirected to expressing *txs* in *P. agglomerans* and trying and devise a better strategy for the co-transforming *txs* and 10β or 5α.

Chapter 6 INTRODUCTION OF TAXADIENE SYNTHASE INTO *P. AGGLOMERANS*

After developing a set of tools for use in *Pantoea*, especially *txs* expression vectors the next step was to transform *P. agglomerans* with these *txs* containing plasmids and express it in order to produce taxadiene. pBT-TXS was the first *txs* plasmid to be transformed into *P. agglomerans* because it is a high copy number plasmid with the ability to tune to expression level of *txs* via P_{BAD} and the addition of arabinose. Faithful transformation with pBT-TXS was not achieved as readily as expected, this chapter details the issues faced and the steps taken to mitigate them.

6.1 Potential DNA methylation and restriction systems in *P. agglomerans*

pBT-TXS extracted from *E. coli* was transformed into *P. agglomerans* by electroporation, cells were allowed to recover at 30°C for one hour and then transformants were selected for on LB ampicillin and tetracycline plates overnight at 30°C. This transformation provided fewer colonies than transformation of *P. agglomerans* with pBT-GFP. There were also a range of colony morphologies. Initially after 1 day of incubation pale colonies that looked much like wild-type were observed, some of which were larger than the others. After 2-3 days of incubation these colonies developed a deeper shade of yellow, and the difference in colony size was more pronounced. There was also the appearance of a new morphology, small raised nodules with a crinkled texture, which in some cases appeared to dig into the agar.

6.1.1 Wild-type morphology colonies

Initially the big and small phenotype colonies were analysed for the fidelity of their transformation with pBT-TXS, as this difference in phenotypes could be caused by a deletion mutation in *txs*. To start, one big and one small wild-type colony were restreaked on to fresh media and incubated at 30°C overnight. Big colonies produced similar colonies after overnight growth whilst the small colonies produced small colonies and a few that developed into slightly larger and more yellow colonies like those seen in figure 6.1 C.

A single colony from the big colony plate and a single small colony from the small colony plate were picked and grown in LB ampicillin and tetracycline culture overnight at 30°C 150 RPM, plasmid DNA was then extracted and Sanger sequenced using the primer GBT_R. GBT_R was designed to sequence the 3' end of sequences inserted between the *Not*I and *Xho*I sites of the pBT series of plasmids, in this case the *txs* gene. GBT_R sequence is based on the terminal region of *tetR*, ensuring that it could prime and sequence potential deletion mutants, as tetracycline resistance is selected for by tetracycline in the media. The sequencing from the small colony confirmed that *txs* was present, however sequencing from the large colony showed a 2263bp deletion in the middle of *txs*, which has then been repaired to produce an 189bp remnant of *txs*, removing almost all of the original sequence. This plasmid was named pBT-DAM1 (Damaged) and is shown in Figure 6-2 pBT-DAM1 alongside a description of the differences between pBT-DAM1 and pBT-TXS.

This same deletion was isolated from three separate transformants (3 biological replicates). Due to the reproducibility of this deletion it raised the possibility that *P. agglomerans* has a methylation and restriction system. Some hypothesis of how this system might work were then made from the work already carried out, in order to create a strategy to avoid restriction of *txs* and transform *P. agglomerans* with pBT-TXS faithfully.

The restriction system is not 100% effective at digesting all DNA that comes in to the cell as pBT-GFP and other plasmids had been transformed in without issue. The alternative theory here is that previous plasmids have been linearised, and then repaired faithfully. In either case DNA can be transformed into *P. agglomerans*.

To delete a section of *txs* at least 2 restrictions are required and then repaired.

There is potentially a negative selection pressure against the maintenance of *txs*, which would help to select for deletion mutants to flourish.

It possibly is not a type II restriction enzyme as the cut sites were not palindromes or share homology, however type II s restriction sites are not palindromes and cut outside of the recognition site, but no sequences were identified in the vicinity of the

cut sites. This leads to the possibility of it being a type III restriction enzyme, with a recognition site much further from the cut site.

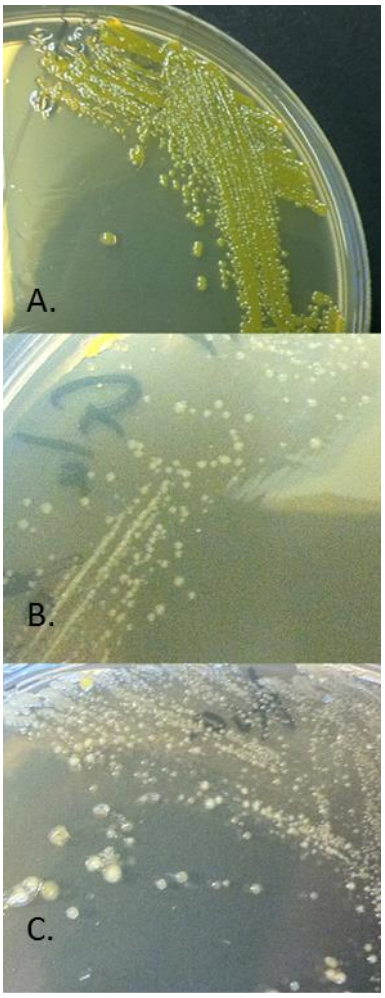


Figure 6-1 Colony morphology of *P. agglomerans* transformed with pBT-TXS A. large colony produces similar colonies when restreaked, B. Small colonies also propagate small colonies, however as shown in C. small colonies can also yield big colonies.

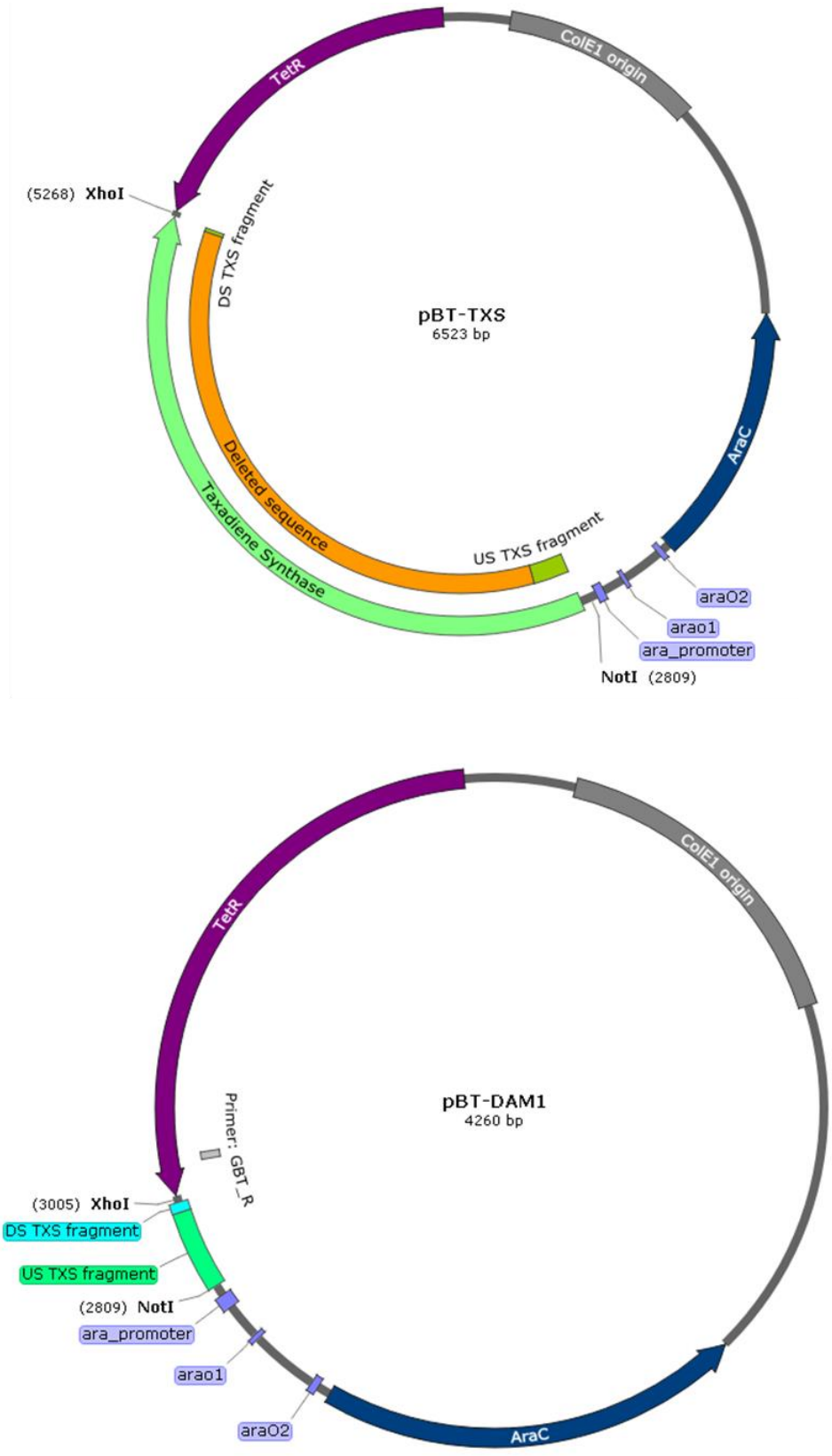


Figure 6-2 pBT-DAM1 Produced by the deletion of the majority of TXS

As DNA can be transformed into *P. agglomerans*, and that the issue of deletion of blocks of DNA from plasmids had not been observed prior to this experiment, it suggested that there was a selection pressure to remove *txs*. Much like in the attempted integration of *txs* into the chromosome and the loss of LPP-1 (chapter 4), this event led to the loss of a potentially functional taxadiene biosynthetic pathway.

Methylation of pBT-TXS, at the correct motif, would block restriction by the nuclease. One method for producing DNA with the correct methylation pattern would involve cloning the methylases from *P. agglomerans* and integrating them into the chromosome of a cloning strain of *E. coli.* Plasmids could then be transformed into this strain, to methylate them, DNA would then be extracted and transformed into *P. agglomerans*. This method is used for transformations of several organisms including *staphylococcus aureus* and *Haloferax volcanii*. *S. aureus* is sensitive to cytosine methylation, this can be avoided by transforming plasmid into *E. coli* DC10B which is deficient in cytosine methyltransferase (Monk *et al.,*2012). *H. volcanii* has a restriction system which recognises adenine-methylated GATC sites and so plasmid DNA needs to be transformed into a *dam*$^{-}$ strain of *E. coli* deficient in GATC methylation before transformation into *H. volcanii* (Holmes *et al.,*1991). Without having access to the completed genome sequence for this strain of *P. agglomerans*, it was not possible to take this approach. Instead it was decided to remove the selection pressure for the production of non-functional *txs*. By knocking out *crtE P. agglomerans* cannot make GGPP and therefore TXS would have no substrate for the production of taxadiene. The transformed plasmid would then be able to be methylated and replicated. After extraction, this could then be used for transformation of the wild type strain.

6.2 Deletion of LPP-1 from *P. agglomerans*

The suspected loss of LPP-1 during the attempted integration of *txs* into the chromosome of *P. agglomerans*, suggested that loss of *crtE* removed the selection pressure imposed by having a functional taxadiene biosynthetic pathway. Removal of LPP-1 was therefore suggested as a time efficient way of performing a knockout of *crtE* without having a full genome sequence. To do this a method was devised that used the observation that that the presence of *txs* can select for the loss of LPP-1. *txs* would be introduced into *P. agglomerans* on the temperature sensitive origin of replication pSC101. Then it would be selected for, whist removing antibiotic selection on LPP-1 leading to the maintenance of *txs* and the loss of LPP-1.

6.2.1 Selection for the loss of LPP-1 from *P. agglomerans* using pLBT-TXS

txs was inserted in to pLBT (described in chapter 7) under the control of P_{BAD} on the pSC101 origin of replication with tetracycline resistance. pBT-TXS and pLBT were restricted with *Not*I and *Xho*I. *txs* from pBT-TXS and the linearised back bone of pLBT were recovered from their respective reactions by gel electrophoresis. *txs* was then ligated into pLBT. The reaction was then transformed into *E. coli,* recovered at 30°C for 1 hour. Transformants were selected for on LB tetracycline agar at 30°C overnight. Colonies were screened by restriction digest. The finished construct was named pLBT-TXS.

pLBT-TXS was transformed into *P. agglomerans* via electroporation, transformants were selected for on LB tetracycline agar at 30°C and then confirmed by restriction digest. A single colony was diluted in 1ml of LB, serial diluted and then plated on LB tetracycline agar (without ampicillin). Plates were incubated at 30°C for several days and colonies were observed daily. After 2-3 days colonies were around 1cm in diameter and zeaxanthin pigmentation was clear. Several colonies showed a white

ring around the central yellow section of the colony. These were picked and restreaked onto fresh plates and incubated at 30°C overnight. These colonies grew and remained unpigmented. Colonies were screened for pLBT-TXS using plasmid purification and restriction digest. The presence of pLBT was confirmed. This strain was then patched onto LB ampicillin and incubated at 30°C, but did not grow. The loss of carotenoid production and the sensitivity to ampicillin suggested the loss of LPP-1. The strain was then streaked on to LB agar and incubated at 37°C to cure it of pLBT-TXS. Colonies that grew were patched on to LB and LB tetracycline and incubated at 30°C. One colony from the LB plate that did not grow on LB tetracycline was selected and designated *P. agglomerans* ΔLPP-1

6.2.2 Evidence that *P. agglomerans* has a restriction system

To investigate if the ΔLPP-1 strain of *P. agglomerans* could be used to improve transformation efficiency into *P. agglomerans* wild-type, *P. agglomerans* ΔLPP-1 was transformed with pBT-GFP and pBT-TXS, separately, and selected for on LB tetracycline at 30°C overnight. Colonies were screened by plasmid purification followed by restriction digest. Plasmids gave the expected restriction digest pattern and so no mutants were detected (10 colonies screened). Plasmid DNA was then extracted from these strains and from the equivalent strains of *E. coli.*

Wild-type *P. agglomerans* was transformed in 4 reactions with 100ng of pBT-GFP and pBT-TXS derived from *E. coli* and *P. agglomerans* ΔLPP-1 DNA. Transformations were then recovered at 30°C and then plated on LB tetracycline and incubated overnight at 30°C. Colony counts were performed the next day. Wild-type *P. agglomerans* transformed with pBT-GFP produced 30X more colonies when transformed with plasmid DNA extracted from *P. agglomerans* ΔLPP-1. For the transformation with pBT-TXS. The efficiency was increased by 120X between DNA extracted *from E. coli* versus *P. agglomerans* ΔLPP-1.

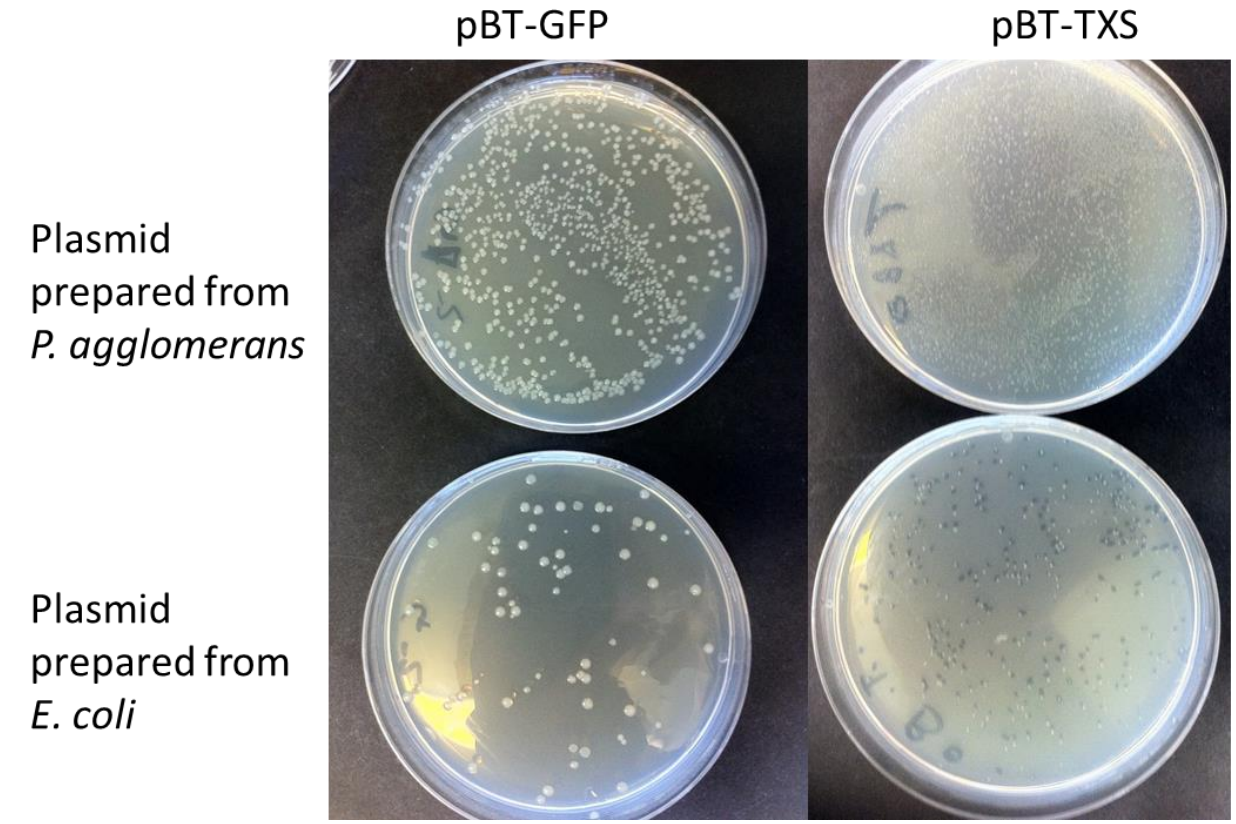


Figure 6-3 Colony recovery in *P. agglomerans* transformed with plasmid DNA prepared from *E. coli* and *P. agglomerans* ΔLPP-1.

Colonies of *P. agglomerans* transformed with pBT-TXS extracted from *P. agglomerans* ΔLPP-1 were screened by plasmid purification and restriction digest, all produced bands of the expected size, suggesting that the methylation was successful at protecting DNA. However this was not an extensive screen to prove this restriction was no longer happening. Big colonies were not observed, and transformation efficiency had been improved. These results are consistent with the presence of a restriction modification system in *P. agglomerans,* and enabled the improvement of the transformation process for *P. agglomerans.*

6.3 Induction of *txs* under the control of P_{BAD}

After isolating *P. agglomerans* faithfully transformed with pBT-TXS and with a fully functional zeaxanthin biosynthetic pathway, the next step was to culture it and attempt to produce taxadiene.

An overnight culture of *P. agglomerans* pBT-TXS was set up in 10 ml of LB tetracycline and ampicillin culture media at 30°C 150RPM. 2x 1:10 dilutions were then set up in 50ml of fresh LB tetracycline and ampicillin culture media, these were incubated at 30°C 150RPM until OD600 of 0.6 was achieved at which point one of the cultures was induced with arabinose to bring the final concentration to 0.1% arabinose. Both cultures were incubated at 30°C 150 RPM for 2 days. Plasmid DNA was then extracted from both cultures. 300 ng of this DNA alongside controls of the original pBT-TXS extracted from *E. coli* were then digested with *Xho*I (single restriction site) and *Sal*I-HF (two restriction sites) in separate reactions. These restriction digests were then separated by electrophoresis alongside uncut plasmid DNA. Figure 6-4 Effect of induction on stability of pBT-TXS shows the final stained gel. The un-induced culture has maintained the pBT-TXS plasmid with a band around the 6 Kbp point for the single cut and 2 bands one around 4 Kbp and the other close to 2.5 Kbp, matching the *E. coli* controls. The induced samples however present a different pattern. All of the bands present in the *E. coli* and un-induced samples are present, but in both the single cut and double cut restriction digest reactions an identical band can be seen which is between 3 and 4Kb in size. This band was extracted, and self-ligated overnight, it was then transformed into *E. coli* and selected for on LB tetracycline. Transformants were obtained and 3 of these were cultured and plasmid DNA extracted, examined by restriction digest and sent for Sanger sequencing using the primer pGBT_R. Figure 6.6 shows the plasmid map derived from the sequencing of the self-ligated fragment (here after known as pBT-DAM2, figure) alongside the original pBT-TXS. pBT-DAM2 was formed by a 3.6Kb deletion occurring from the terminal end of *txs* (and *tetR*) to the terminus of *ara*C. This deletion encompasses the

entirety of *txs*, the promoter P_{BAD} and removes any functionality from the transcriptional regulator *araC*.

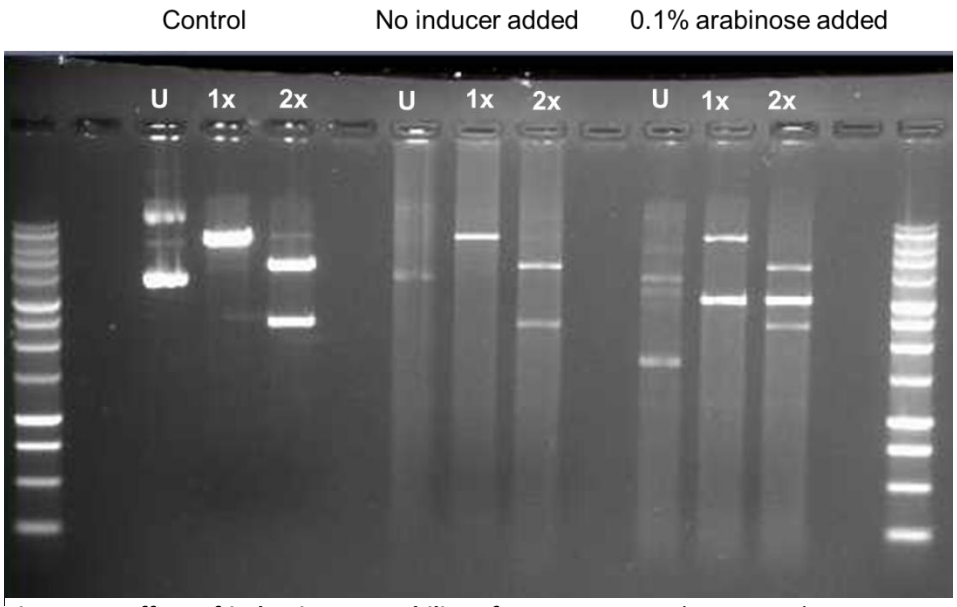


Figure 6-4 Effect of induction on stability of pBT-TXS Control was *E. coli* pBT-TXS, No inducer and 0.1% arabinose cultures were grown from *P. agglomerans* transformed with pBT-TXSU, after culturing for 2 days. U. uncut DNA, 1x single restriction with *Xho*I, 2x double restriction with *Sal*I-HF

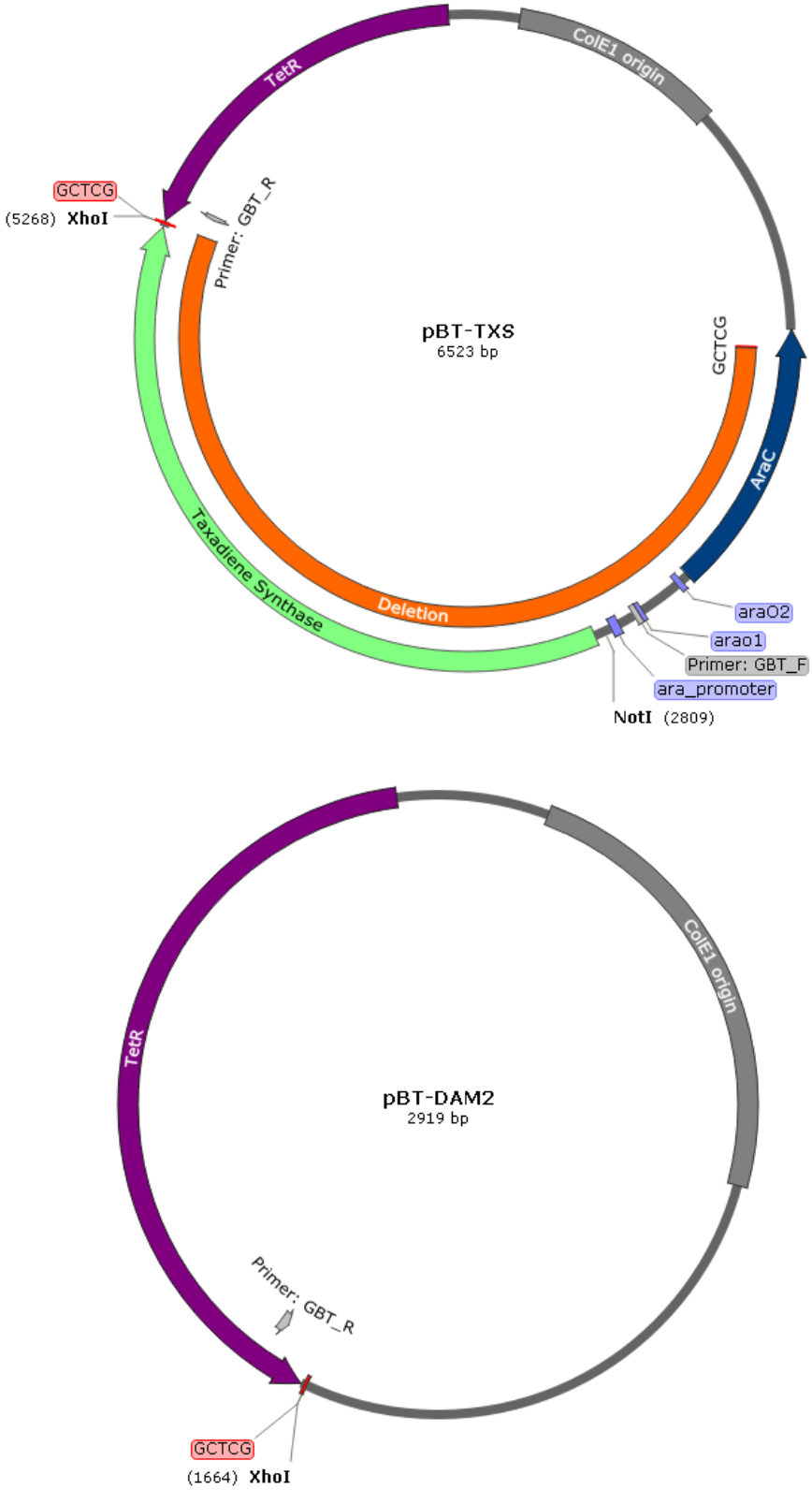


Figure 6-5 pBT-DAM2 Deletion in *txs* and *araC* leads to a reduced pBT with the minimum functions required, tetracycline resistance and an origin of replication.

To see if this result was reproducible six, 10ml cultures were set up in the same manner as previous experiment (3 induced 3 un-induced). After the cultures were induced (or not induced) they were incubated for 24 hours at 30°C 150RPM, after this point 1ml of culture was frozen for plasmid extraction at the end of the experiment and a 1:100 dilution into fresh media was also made from each culture, and this process was repeated over 6 days.

After 6 days plasmid DNA was extracted from all cultures and was restricted with *Sal*I-HF. pBT-TXS extracted from *E. coli* was also restricted in the same way and run alongside as a control, the results can be seen in Figure 6-6 Stability of pBT_TXS over 6 days. All cultures had undergone a mutation in pBT-TXS. The next step was to track down the point at which this mutation occurred. All frozen samples of culture X (induced) and A (un-induced) were treated to extract plasmid DNA and then restricted with *Xho*I. pBT-TXS extracted from *E. coli* and *P. agglomerans* was also used as a control and restricted with *Xho*I. These were then visualised using gel electrophoresis, Figure 6-7 6 day course of plasmid restrictions for induced and uninduced. The plasmid DNA used for the *P. agglomerans* control was the wrong sample and should be ignored, it appears that sample X2 was used for this digest. The figure clearly shows that within the first 24 hours both cultures had had a deletion within pBT-TXS, probably even before the first time point of the experiment. However the addition of arabinose appears to have added to the selection pressure for the loss of *txs* as only a faint band is visible after 24 hours for sample A. DNA from day 6 of both sets A and X were sent for Sanger sequencing using primer pGBT_R. These revealed the same sequence, suggesting that the mutation occurred during the first overnight culture, from which all 6 cultures were set up. This deletion was larger than that of pBT-DAM1 or pBT-DAM2, going from 6.5Kb to 2.5Kb. It was named pBT-DAM3 and a map is shown in Figure 6.9.

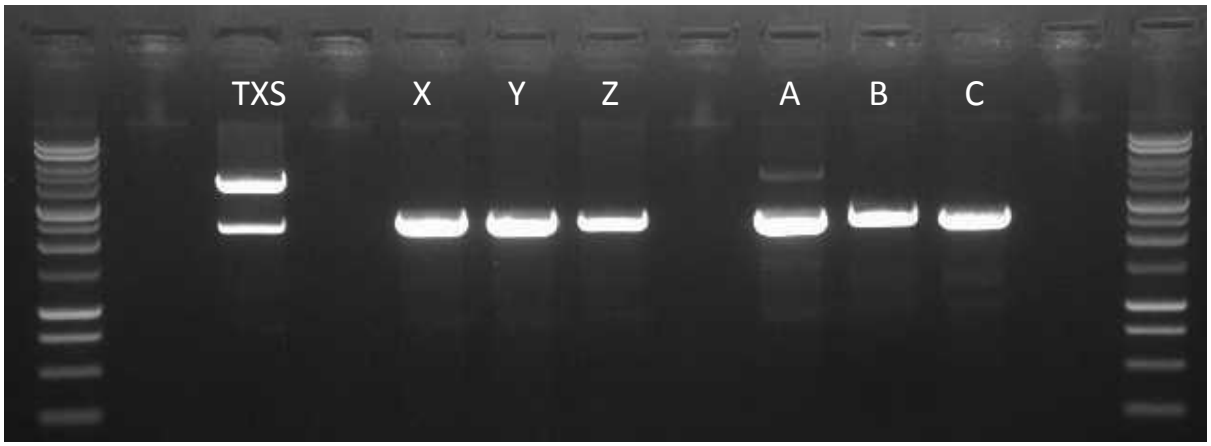


Figure 6-6 Stability of pBT_TXS over 6 days. 3 uninduced cultures (X,Y,Z) and 3 cultures induced with arabinose had their plasmid DNA extracted after 6 days and restricted with *Sa*lI-HF (double cut in pBT-TXS). *txs* control was extracted from *E. coli*.

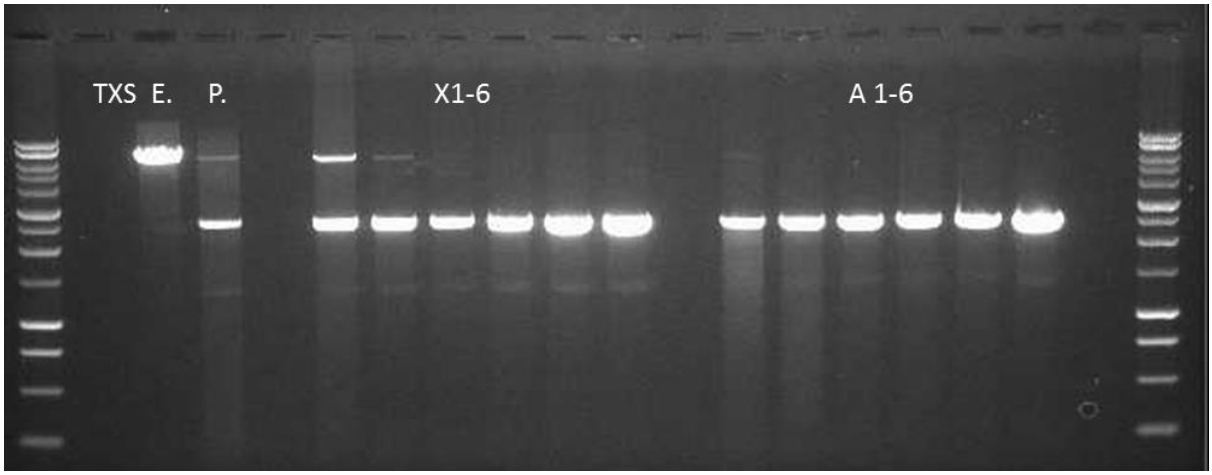


Figure 6-7 6 day course of plasmid restrictions for induced and uninduced Culture X (Uninduced) and A (induced with arabinose) had plasmid extractions carried out on every day of the culturing experiment. Samples were restricted with *Xho*I. E. *txs* Control was from *E. coli*. Sample P was mistakenly loaded

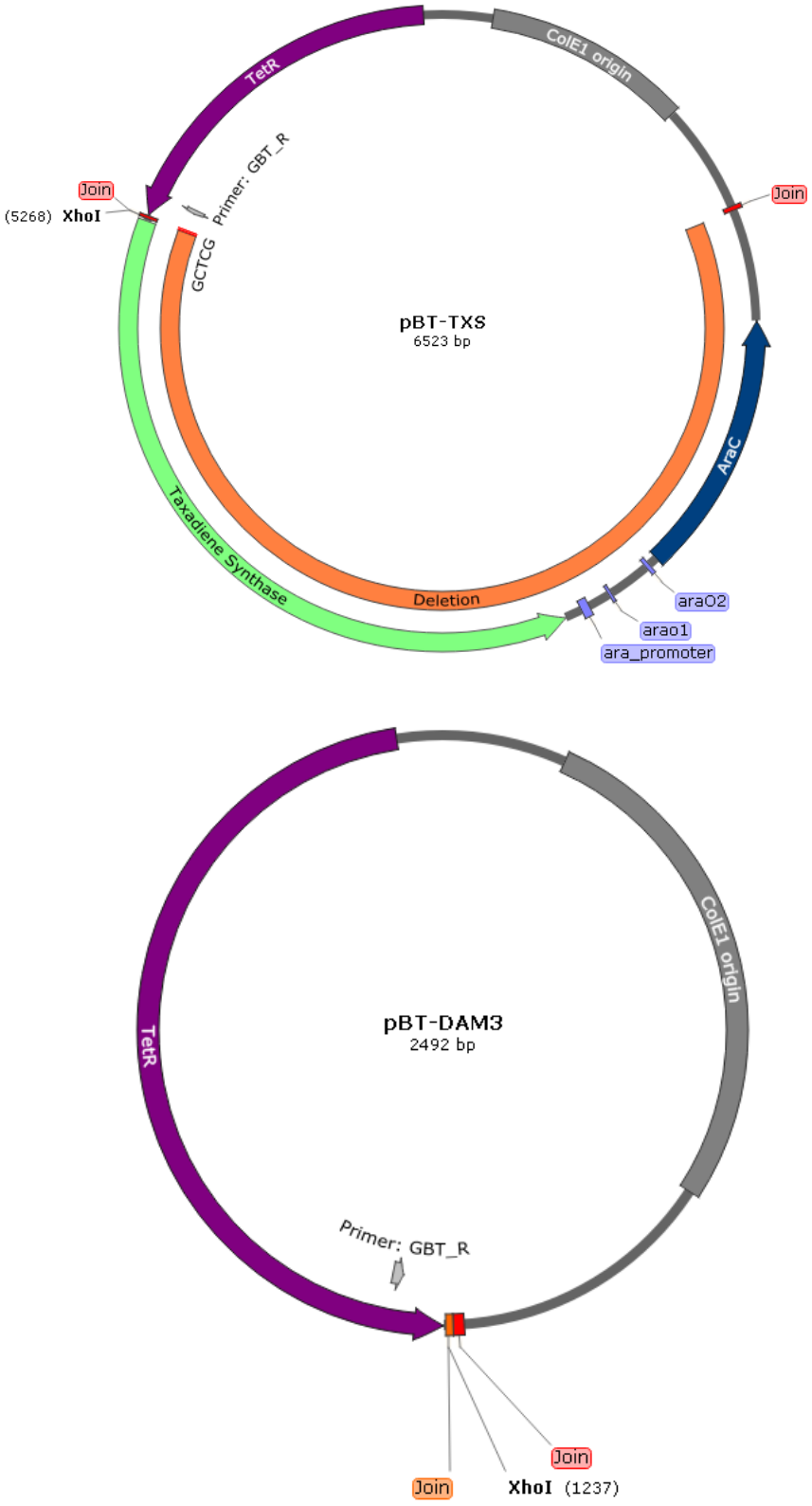


Figure 6-8 pBT-DAM3

6.4 Hard colony morphology and propagation

As previously presented in section 6. 1, during the initial transformation of *P. agglomerans* with pBT-*txs*, there were three observed phenotypes. As well as the different sized colonies there were also the crinkled hard colonies (from now on referred to as hard colonies) which had not previously been seen whilst working with *P. agglomerans.*

After further optimisation of the transformation of *P. agglomerans* with pBT-*txs* via the presumed methylation of pBT-*txs* using *P. agglomerans* ΔLPP-1, the transformation plates were still not homogenous in colony morphology, with hard colonies occurring at a high prevalence.

Colony PCR was carried out on these hard colonies using primers GBT_F and GBT_R which confirmed the presence of a 2.5Kb band which is equal to that of an intact TXS.

6.4.1 Culture and re-streak

To investigate the cause of this phenotype, it was first decided to devise methods for the production of enough material for further analysis.

Collecting material from one of these colonies was difficult, unlike a normal colony, which is pliable, these did not leave visible residue on the metal loop when probed. In order to remove any visible amount of material the easiest method was to remove the whole colony by applying pressure to the side or base, often this would lead to the removal of the whole colony and some of the agar from underneath.

Whole colonies were inoculated into 5ml of LB tetracycline and ampicillin media at 30°C 150 RPM overnight. Cultures were then observed. It was unclear whether the original colony was still aggregated as the media was cloudy. Plate based methods were then considered a better option.

Whole colonies were picked and then streaked on LB tetracycline and ampicillin agar. The colony was encouraged to remain within the confines of the loop whilst it was dragged along the plate in a cross hatch pattern. Plates were then incubated over night at 30°C.

Colonies were not observed on the following day, and were placed into cold storage at between 4-5°C. The procedure was repeated, but the same results were seen, the hard colonies appeared to not be amenable to propagation on solid media in this manner. However after 1 week, the plates were recovered from cold storage and found to have grown hard colonies (Figure 6-9 Re-streaked hard colony.. Colonies developed as ridges and as smaller individual hard colonies, however it is not certain if these can be considered to have been formed by a single cell.

6.5 Transformation method

Before the discovery that hard colonies could be propagated by restreaking on LB tetracycline and ampicillin agar, attempts were made to optimise the transformation procedure for the production of hard colony material to analyse what these colonies were made from. Recovery temperature and media content were both factors that were investigated.

To investigate the effect of the temperature of overnight incubation, 3 transformation reactions were performed, to transform wild-type *P. agglomerans* with pBT-TXS extracted from *E. coli* (this would have been better if carried out using plasmid DNA extracted from *P. agglomerans* ΔLPP-1, and was in future experiments). These were then recovered at 30°C 150 RPM for 1 hour then plated on to LB tetracycline and ampicillin agar. For each transformation 9 plates were spread with 100 µl of culture, 6 plates with undiluted transformation mix at neat concentration and 3 at a diluted of 10^{-1}. These were then incubated in sets of three (2X 10^{0} and 1x

10^{-1}) at either 25°C, 30°C or 37°C for 3 days. Hard colonies were counted every 24 hours. Results are show in Table 8 Influence of incubation temperature on Hard colony formation.

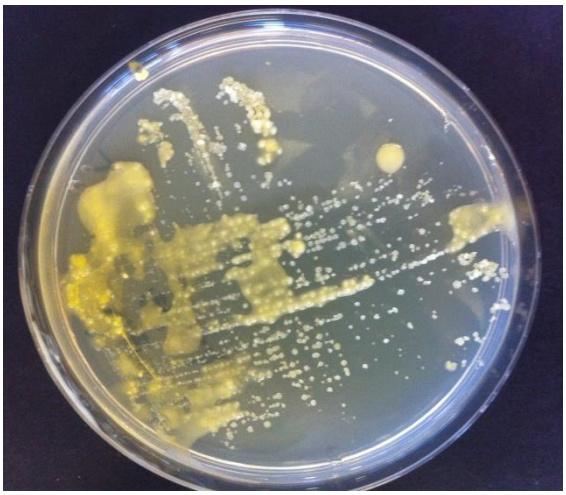


Figure 6-9 Re-streaked hard colony. Hard colonies developed several days after initial incubation, and form at temperatures as low as 4°C

		Number of hard colonies a different incubation temperatures		
		Incubation temperature		
		25°C	30°C	37°C
Transformation	A	0	8	135
	B	5	88	3580
	C	9	7	842

Table 8 Influence of incubation temperature on Hard colony formation

From this work it was apparent that overnight incubation at 37°C promoted the production of hard colonies, however these colonies were small and difficult to pick compared to those grown at lower temperatures. This could be due to competition for resources as the 37°C plates were much more densely populated, and so recovery at 30°C was favoured for future production.

6.5.1 Addition of glycerol and glucose to the agar plates

As this phenotype was observed after the transformation with pBT-TXS, the possibility was considered that this phenotype was being caused by the low level induction of *txs* by P_{BAD}. In order to try and reduce this it was proposed to add 0.2% glycerol and 0.2% glucose to the media (Guzman *et al.,* 1995).

Transformation was performed on *P. agglomerans* with pBT-TXS extracted from *P. agglomerans* ΔLPP-1 pBT-TXS. The transformation was recovered at 30°C at 150 RPM for 1 hour and then 100 µl of the recovery media was then plated out. 4x 100 µl at 10^0 dilution and 4x 100µl at 10^{-1} dilution, on either LB tetracycline and ampicillin + 0.2% glycerol or LB tetracycline and ampicillin + 0.2% glycerol +0.2% glucose. These were then split and incubated at either 30°C or 37°C over 2 nights. Results are show in table 9.

	Number of hard colonies a different incubation temperatures	
	Incubation Temperature	
	30°C	**37°C**
0.2% glycerol +0.2% glucose	6	152
0.2% glycerol No glucose	Too many to count	500

Table 9 . The addition of glucose to media and the influence on Hard colony production. Averages of plates that had colonies

The addition of glucose appears to have hampered the production of hard colonies. However hard colonies grown in the presence of glucose tended to be larger. Plates with glucose added also produced colonies with a mucoid phenotype. This made the extraction of hard colonies more difficult, as the phenotype spread far from the centre of the colony and would often make contact with other colonies. This phenotype occurred on glycerol plates, but at a lower frequency and with reduced spread.

P. agglomerans transformed with pBT-GFP was streaked onto LB tetracycline and ampicillin + 0.2% glycerol and LB tetracycline and ampicillin + 0.2% glycerol +0.2% glucose plates and incubated over night at 30°C as a control to ascertain how wild-type *P. agglomerans* would interact with the different media. As shown in Figure 6-10 both plates produced growth, the plate with 0.2% glucose had a defined mucoid phenotype, although compared to LB media with no additional glycerol or glucose, the plate with only 0.2% glycerol also had a mucoidal quality. Hard colony morphology was not observed under these conditions.

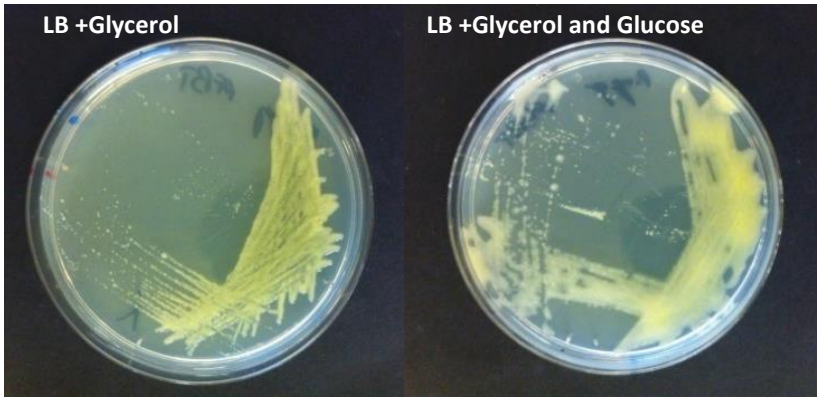


Figure 6-10 Effect of glycerol and glucose on wild-type *P. agglomerans* morphology. Addition of glucose and glycerol made P. agglomerans become more mucoid.

6.6 Increasing Hard colony production

To make material for further study, transformation was performed on *P. agglomerans* with pBT-TXS extracted from *P. agglomerans* ΔLPP-1 pBT-TXS. The transformation was recovered at 30°C at 150 RPM for 1 hour and then diluted 1:10 and then 100 µl was plated out onto 10 plates each of LB tetracycline and ampicillin, LB tetracycline and ampicillin + 0.2% glycerol or LB tetracycline and ampicillin + 0.2% glycerol +0.2% glucose. These were incubated at 30°C for 3 days. Some of the resulting plates can be seen in figure 6. 12, which shows the difference appearance of colonies when plated on LB, LB +0.2% glycerol or LB +0.2% glycerol and 0.2% glucose.

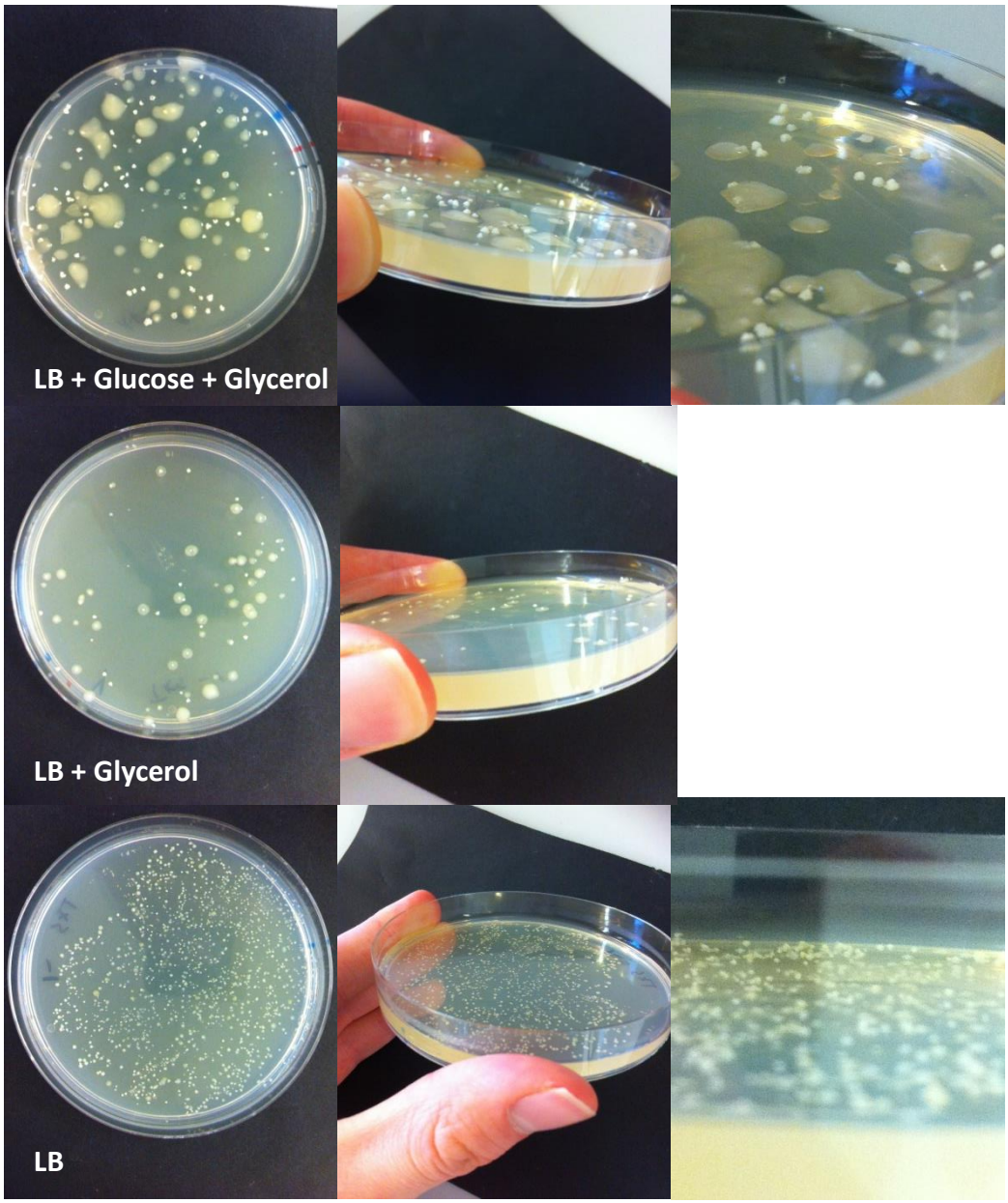


Figure 6-11 Effect of glycerol and glucose on Hard colony production in *P. agglomerans* transformed with pBT-TXS.

6. 6. 1 Propagation of Hard colonies

Hard colonies from this trial were picked and patched onto fresh LB tetracycline and ampicillin +0.2% glycerol agar plates. And grown at 30°C. Figure 6. 13 shows the

development of these plates over 4 days of incubation. After 4 days of incubation all colonies have produced zeaxanthin, showing that they have a functional *crtE* and can produce GGPP, so taxadiene could potentially be produced in these colonies, if *txs* is still functional. The hard colonies had grown, suggesting that this method could be used to propagate colonies, by growing them, splitting them into smaller colonies and placing them onto fresh media. Several colonies had become mucoid after 4 days of incubation.

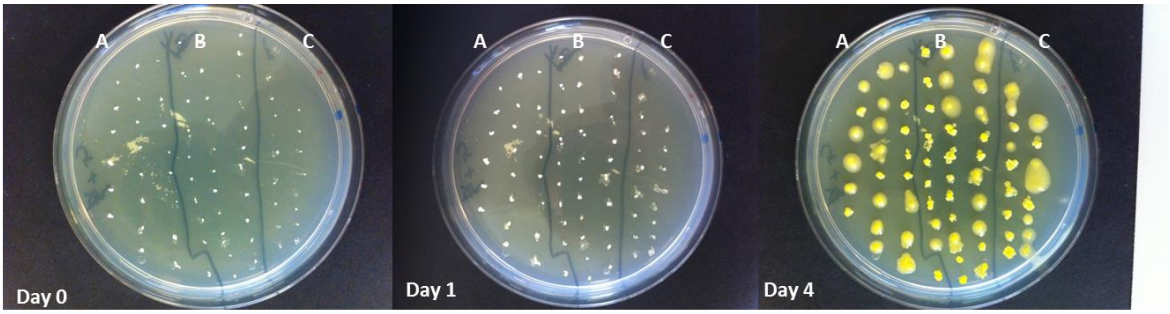


Figure 6-12 Propagation of Hard colonies and their growth. Hard colonies were picked from plated out transformation of *P. agglomerans* with pBT-TXS (Figure 6.12), they were then deposited on to LB with ampicillin, tetracycline and 0.2% glycerol.

6.6.2 Mucoid colony restreak

The change of hard colonies to mucoid phenotype observed when colonies were moved onto new media raised the question of how these two phenotypes were linked. Hard colonies and mucoid material were spread on to LB tetracycline and ampicillin +0.2% glycerol plates, incubated at 30°C overnight and then moved to 25°C for a further 3 days of incubation. During this second stage of incubation, plates were checked every 24 hours. What was observed was that hard colonies propagated hard colonies, some of which became mucoid. Mucoid material spread on a plate led to generation of more mucoid phenotype material, but hard colonies could be seen developing within the mucoid mass. This indicated that both phenotypes could inter convert and that potentially the phenotype was not due to a genotypic difference between the two phenotypes.

To ensure that this was not just the result of contamination and carry over of mucoid material with hard colonies (and vice versa). In triplicate hard colonies were streaked onto fresh LB tetracycline and ampicillin +0.2% glycerol media, grown at 30°C for 24 hours and then moved to 25°C. From this plate a single hard colony was picked and the process repeated. This happened for 8 passages, yet the mucoid phenotype was still present among the resulting restreak.

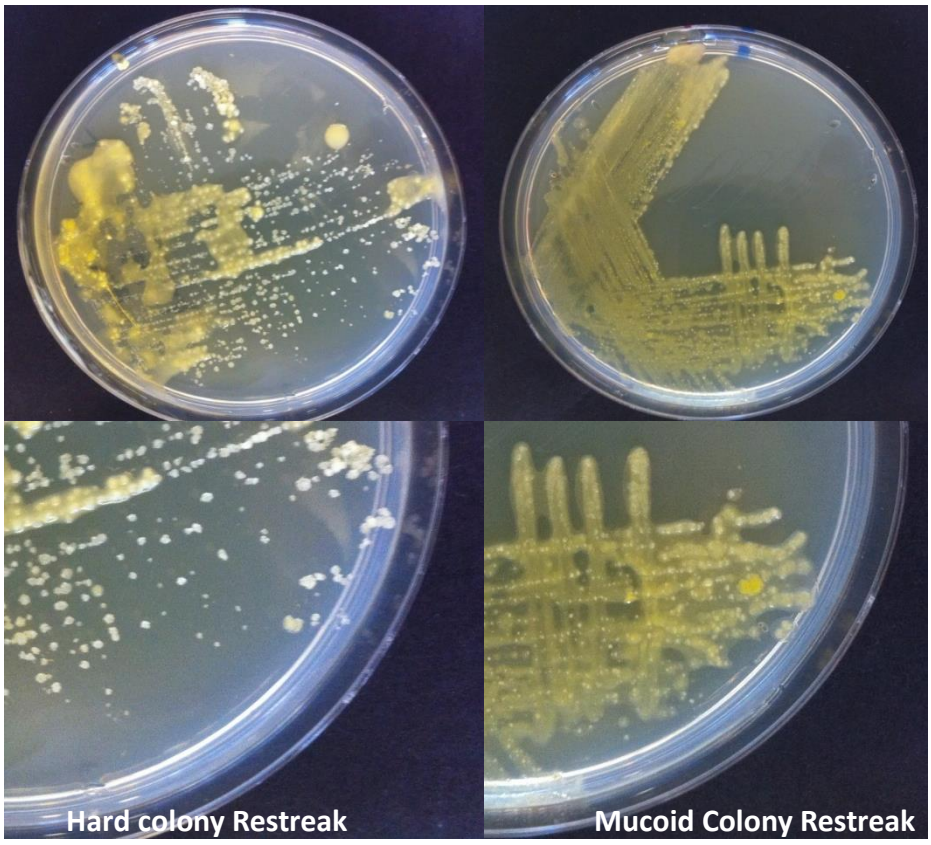


Figure 6-13 Hard and mucoid colony restreak Hard colonies can be restreaked on LB with ampicillin, tetracycline and 0.2% glycerol to propagate hard colonies, and produce some mucoid phenotype. Mucoid colonies when restreaked lead to the production of more colonies with a similar morphology, however hard colonies are clearly visible.

6.6.3 Material collection

After managing to reproduce the hard colony phenotype on a larger scale, the next step was to collect enough material to analyse the composition of these structures.

Colonies were collected using two methods. For larger colonies, these were lanced with a sterile wooden toothpick, then placed into a small volume of sterile distilled water (to help remove them from the pick). For small colonies, plates were flooded with sterile distilled water, a plastic spreader was then dragged over the plate to dislodge hard colonies. This had the undesired effect of suspending mucoid colonies as well and so the toothpick method was preferable. The plates were then drained into a funnel containing a mirror cloth filter. More sterile distilled water was then used to rinse the plates again, then the contents of the mirror cloth were washed with 50ml (an excess) of sterile distilled water. The flow though was discarded the hard colonies were allowed to dry and then transferred to a dry falcon tube.

6.6.4 Solvent extraction

As the hard colonies initially had a white, crystalline appearance it was theorised that this exterior could be made from taxadiene due to sharing similar properties. To investigate this hard colonies were washed in hexane (one of the solvents taxadiene is most soluble in). 1 ml of hexane was added to dried hard colonies in an Eppendorf tube and mixed by vortexing for 5 minutes. Hard colonies did not appear to dissolve, break apart or in any way reduce in size. The hexane was removed (and kept for further analysis). This process was then repeated with methanol instead of hexane, this also yielded no change in the hard colonies. Finally the process was repeated with chloroform as the solvent. This caused a pellicle to form, presumably from cell debris, yet there was still a clear shell left of the hard colony. After this, the previous solvent washes of hexane and methanol were repeated, as it was speculated that the chloroform may have removed any cells, connective proteins and lipids from the hard

colony scaffold. However these second washes yielded no further degradation of the hard colony.

The remaining hard colony material was air dried and along with the solvent extracts sent to our collaborators in the School of Chemistry for further analysis.

Solvent extracts were screened for taxadiene by GCMS, but none was detected.

The remaining hard colony material was subjected to several more organic solvents, none had any effect. Eventually the only method tested which caused the material to dissolve was heating to 100°C in 1 M HCl.

6.6.5 Congo red staining

From the discovery that the hard colony matrix could be dissolved by boiling in 1 M HCl, and that it wasn't soluble in a range of organic solvents, including chloroform, it was suggested that this material was cellulose or another carbohydrate polymer. This hard colony morphology is similar in several ways to the Red-Dry-Rough phenotype (RDAR) which is often seen in the genus Salmonella as well as some *E. coli* strains (Anwar *et al.,* 2014; Milanov *et al.,* 2015). RDAR is associated with the production of curli and production of cellulose (and other exopolysaccharides) for the formation of biofilms, the regulation of which is controlled by the secondary messenger signal cyclic-di-guanylic acid (c-di-GMP) (Römling, 2005). Congo red stain can be used to stain biofilms and Congo red agar is often used to screen for bacteria capable of biofilm formation.

To produce a Congo Red agar designed to promote the growth of hard colonies, previous recipes were adapted with the media conditions already optimised for the production of hard colonies. LB agar was supplemented with 0.08% Congo Red (Kaiser *et al.*, 2013) and 0.2% glycerol and the appropriate antibiotics, this was then called LBCR agar.

P. agglomerans was transformed with pBT-TXS as previously described, but plated on to LBCR tetracycline and ampicillin. Plates were then incubated at 30°C for 24 hours and then transferred to 25°C. From then they were monitored every 24 hours. Hard colonies formed readily on this media, as well as a few mucoid colonies, similar to what happened on media without the addition of 0.08% Congo red. Hard colonies that grew on LBCR were stained orange/red indicating that they had taken up the stain. Some colonies appeared to revert to mucoid phenotype, whilst still maintaining some characteristics of a hard colony. The raised structure would sink and become softer, easier to remove from the agar and the mucoid portion of the colony would develop. Figure 6.15.

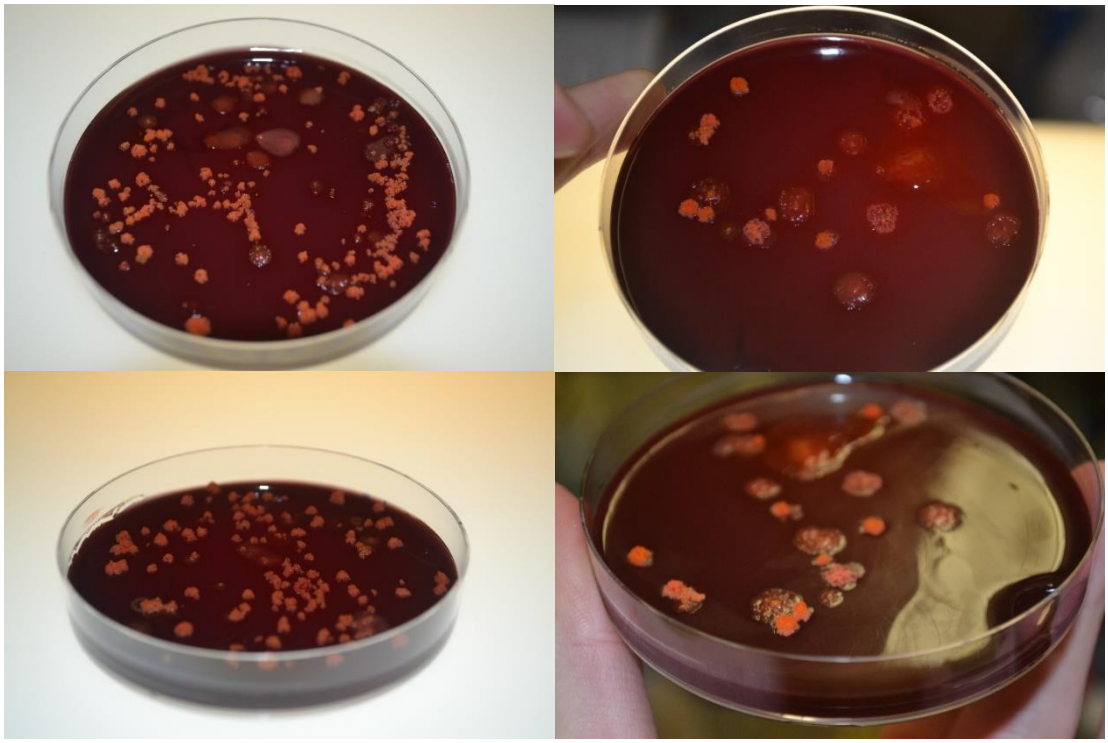


Figure 6-14 Hard colonies forming on Congo Red LB agar. *P. agglomerans* was transformed with pBT-TXS and recovered on Congo Red agar. The colonies were allowed to grow over several days to produce larger constructs to observer their development. The colonies took up the Congo Red stain as they were formed, the colonies protruded from the plate, embedded into the agar and were difficult to remove.

6.6.6. Microscopy

Congo red stain was used to try and investigate the structure of hard colonies. Several staining approaches were tried using a combination of Gram staining techniques and Congo red staining.

Hard colonies or hard colony debris (following chloroform extraction) were prepared for staining by first being placed on to a slide with a drop of sterile distilled water, a second slide was then placed on top of the colony and pressure was applied, the colony was spread out over an area of approximately 2cm in diameter. Both slides were then flooded with 0.1% Congo red stain, and the excess was allowed to run off, and the slide was left to air dry. Acid ethanol was then used to wash the slides of any unbound stain and to reduce the pH and change Congo red stain from red to blue. Figure 6.16 shows hard colony debris and hard colonies prepared in this manner. The slides were observed under oil immersion 100x magnification. The structures observed were larger than individual bacterial cells (although cells are not clearly visible in the images). The structures were irregularly shaped and the three dimensional aspect of them was clear to see by moving the stage up and down. The structures stained strongly with Congo Red, which although not specific for cellulose, does help to support the hypothesis that this structure is constructed from cellulose, among other materials.

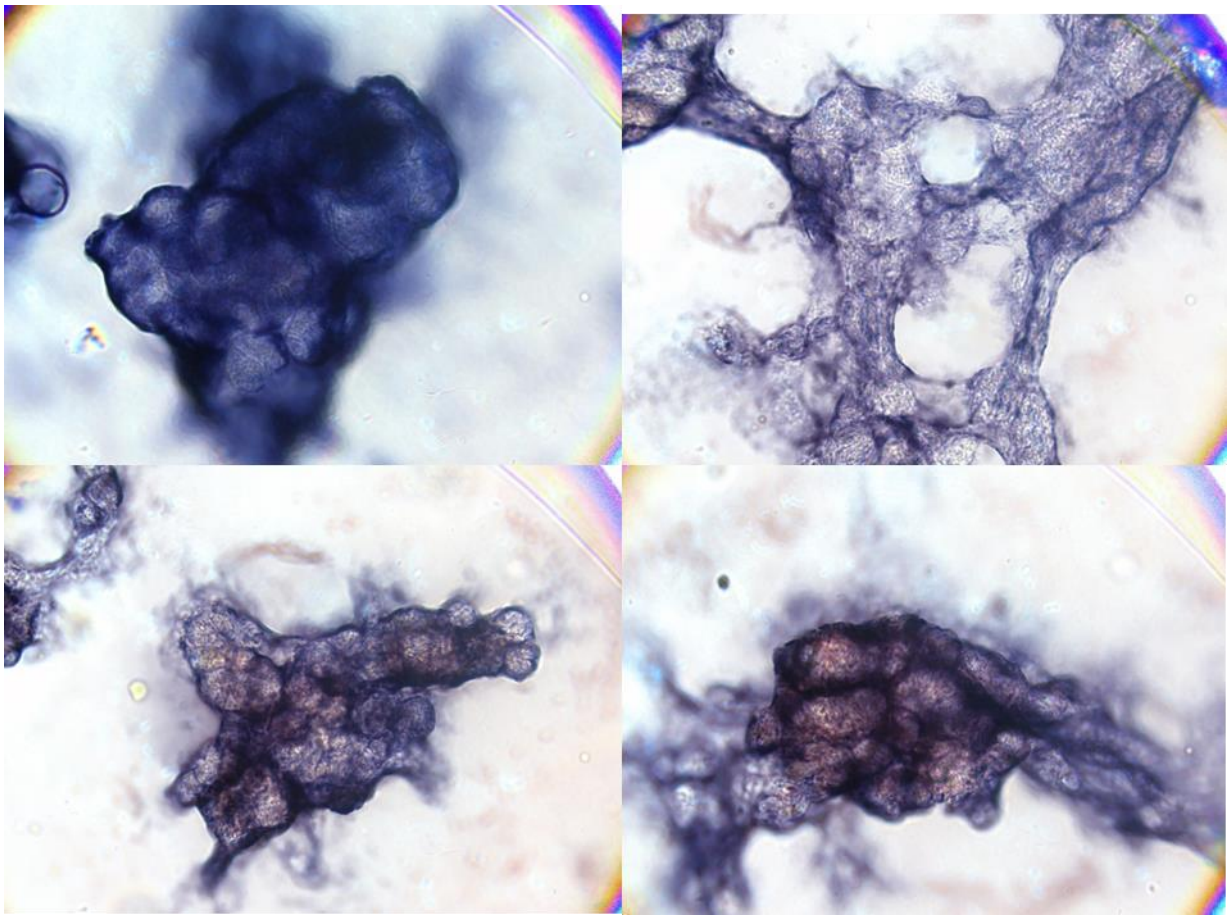


Figure 6-15 Microscopy of hard colony and hard colony debris Congo red followed by acidic ethanol was used as a stain. Slides were observed under oil emersion, 100X magnification.

6.7 Discussion

The initial transformation of pBT-TXS into *P. agglomerans* showed that *P. agglomerans* had a restriction system not previously noticed during the rest of the project. This implies that this system does not completely preclude plasmid DNA derived from *E. coli.* This may be because *E. coli* has some of the same methylation pattern. To test this DNA from strains of *E. coli* with different methyltransferases could be used to compare the transformation efficiency into *P. agglomerans*. Alternatively the use of single molecule real time (SMRT) DNA sequencing of the *P. agglomerans* genome could be used. This method is able to detect modifications to nucleotides and could help to identify the exact methylation behaviour of this strain

of *P. agglomerans* (Davis, Chao, & Waldor, 2013). The transformation of pBT-TXS appears to have highlighted the existence of this restriction system as expression of *txs* appears to be a selective disadvantage. This meant that individuals which restricted *txs* and then repaired the pBT backbone were at a competitive advantage compared to P. agglomerans which had faithfully been transformed with pBT-TXS. The mutants then proliferated quicker during the recovery step of transformation and produced a different phenotype to faithful transformants, making them a target of study.

The development of *P. agglomerans* ΔLPP-1 has improved the transformation of *P. agglomerans* greatly, especially for genes that form part of the Taxol biosynthetic pathway. As it is a Δ*crtE* background it removes the selective pressures for the loss of *txs* and the following genes, and presumably allows the plasmids with these genes to become methylated. This however needs to be confirmed. Despite not being a true restriction negative, methylation modification positive strain (it still has the restriction systems) it is a useful first step on producing one. From this point on thee are two choices for the production of a methylation strain, either the methylases are identified from the genome sequence and integrated into the chromosome of a cloning strain of *E. coli* , or the restriction systems of *P. agglomerans* ΔLPP-1 are knocked out. Short term neither of these are required. However if genes need to be transformed into *P. agglomerans* with a selective disadvantage that is not rescued by deletion of *crtE*, then a true methylase strain would need to be created.

pBT-DAM2 and pBT-DAM3 both occurred during culturing experiments. This would suggest that it is during rapid growth that these mutations are most likely to occur or be selected for, leading to their proliferation within the population. As previously noted, expression of *txs* appears to have a negative effect on the fitness of *P. agglomerans*. The only observation from the literature of taxane toxicity; other than the studies in to the use of taxanes as antibiotics against *Mycobacteria tuberculosis* (Huang *et al.*, 2006), is during production of taxadiene and derivatives in *E. coli* by

Ajikumar *et al.*, (2010). In this study the yield dropped significantly with the addition of the 5α hydroxylation step. They saw the yield fall from 1 g l^{-1} of taxadiene to around 50 μg l^{-1} taxadiene-5α-ol.

DNA is constantly being damaged and repaired within cells, some damage such as mismatched bases are relatively easily repaired and have little mutagenic effect as there is a direct template to compare to. Double strand breaks are potentially much more mutagenic and are repaired by homologous recombination, and can often lead to the inversion, duplication or deletion of genes. Homologous recombination is widely regarded as a major force in the production of diversity from which evolution then selects the fittest individuals (Tham *et al.,* 2016). Stability of pathways is a major research focus within the metabolic engineering community. Sleight & Sauro, (2013) investigated the stability of 24 genetic circuits expressing the florescent proteins YFP, CFP and RFP. They found the most robust circuits had no repeated units and produced lower expression levels. They also found that circuits with expression 10% lower than the maximum had significantly higher evolutionary stability. Taking this into account helps to design more stable expression vectors for *txs*. To improve stability, in the following chapter, lower copy number plasmids were used to reduce the probability of mutations occurring in the taxane biosynthetic pathway.

From the deletion to *txs* and the hard colony phenotype produced by *P. agglomerans* transformed with pBT-TXS, there is a clear indication that expression of *txs* causes *P. agglomerans* stress. From the investigation of the hard colony phenotype, it appears to be similar to biofilm formation such as the RDAR phenotype of *salmonella spp.,* which has been identified as a mechanism for environmental resistance (Milanov *et al.*, 2015), or that of the wrinkly spreader phenotype seen in *Pseudomonas fluorescens* (Spiers *et al.,* 2002). Biofilms often form to protect a community from environmental factors such as desiccation and antibiotics but also from host defences such as antibodies or organic peroxides (Pamp *et al.,* 2007; Sauer, 2003). *P. agglomerans* is a plant epiphyte (Walterson & Stavrinides, 2015), and so it is possible

in its evolutionary past it has had to adapt to surviving on or inside plants. Taxadiene is the first dedicated precursor to Taxol and other taxanes, providing *Taxus* with many defensive chemicals. However the pathways for biosynthesis of Taxol are long and so it is inconceivable that they evolved without some of the intermediate products having a selective advantage for *Taxus spp.* If taxadiene is in some way bioactive, it could be hypothesised to modulate the formation of cellulose and biofilm production in *P. agglomerans*, allowing the bacteria to resist host defence mechanisms. From analysis of the genome of *P. agglomerans* it was found that an organic peroxide resistance operon is upstream of the cellulose synthase operon. This could indicate a connection between the regulations of both of these pathways, however in *E. coli* it has been shown that cellulose synthase is allosterically regulated by the bacterial secondary messenger cyclic di-GMP (Morgan *et al.,* 2014), meaning that the increase in cellulose production could be a by-product of an increase in cyclic di-GMP caused by many other cell regulatory pathways. Further investigation to the increase in biofilm formation would need to be carried out in order to compare this phenotype to other similar phenotypes in the Enterobacteriaceae. This work may require transcriptomic analysis, aiming at looking what changes in transcription the production of taxadiene causes. Alternatively the methods used by other groups that have studied biofilm producing phenotypes could be used such as creating a Tn5 mutant library and screening for individuals that could not produce the hard colony phenotype, much like in the work on the wrinkly spreader phenotype (Spiers *et al.*, 2002). Work could also look at what survival benefit the phenotype provides.

This chapter highlights the need for stable and sustainable expression of biosynthetic pathways, especially as for the most part heterologous expressed pathways have little benefit to the host's fitness and are therefore under constant selection pressure for their reduction in efficiency. The following chapter will deal with the methods used to attempt to stabilise this pathway

Chapter 7 Co-expression of taxadiene synthase with the p450 enzymes taxadiene 5a hydroxylase and taxane 10b hydroxylase

Following the first attempts to reconstruct the synthetic taxane operon and the subsequent separation of the individual modules, producing pSU-TXS, pBT5α, pBT10β and pBT-5α10β, these individual parts needed to be characterised for their effect on *P. agglomerans* , with the aim of identifying how to mitigate the perceived toxicity of the pathway.

7.1 Transformation of pSU-TXS into *P. agglomerans*

P. agglomerans was transformed with pSU-TXS by electroporation, and transformants selected on LB chloramphenicol at 30°C. DNA from colonies were screened by restriction digest to identify correctly transformed colonies, all colonies screened produced the correct band pattern. One sample was then Sanger sequenced using primers TS middle_f, TS middle_r, M13f and M13r. These were shown to have the correct sequence.

Transformation of *P. agglomerans* with pSU-TXS did not produce any abnormal colonies that were noticed in the transformation of *P. agglomerans* with pBT-TXS.

P. agglomerans was also transformed with pSU8 as a control. This was performed in the same manner as previously stated. When colonies were grown overnight to produce enough cells for plasmid extraction, it was noticed that the colour of the culture was much more yellow than the wild-type, the culture also had a distinctive sweet fruity smell.

Visual inspection of cultures of wild-type *P. agglomerans* , *P. agglomerans* transformed with pSU8 and *P. agglomerans* transformed with pSU-TXS indicated that all three cultures appeared to be slightly different shades of yellow, with *P. agglomerans* transformed with pSU-TXS being the palest and *P. agglomerans* transformed with pSU being the brightest. The zeaxanthin content of these strains were investigated to confirm if this observation was due to zeaxanthin accumulation.

7.1.1 Differences in zeaxanthin production between *P. agglomerans* transformed with pSU8 or pSU-TXS

Overnight cultures were set up in 10 ml of LB for *P. agglomerans,* and in LB chloramphenicol for *P. agglomerans* transformed with pSU8 or pSU-TXS. These were incubated at 30°C at 150 RPM. The next morning 1:100 dilutions were made into fresh media (same as previously stated) in triplicate. These were then incubated for 24 hours at 30°C 150RPM. After 24 hours zeaxanthin was extracted and measured from 5ml of culture (method 2.12.2). The dry cell weight was then measured from 90 ml of culture. Figure 7.1 shows the production of zeaxanthin by cell weight. The dry cell weight of *P. agglomerans* transformed with pSU-TXS was not significantly different from wild-type *P. agglomerans* (Two tailed t-test P value = 0.7432) or *P. agglomerans transformed* with pSU8 (two tailed t-test P value = 0. 1019), but the difference in weight between wild-type *P. agglomerans* and *P. agglomerans* transformed with pSU8 was significantly different (two tailed t-test P value = 0.0016). With the zeaxanthin measurements normalised to account for any difference in mass, it is clearly shown that the introduction of *txs* into *P. agglomerans* reduces the levels of zeaxanthin, whilst the introduction of pSU8 increases production of zeaxanthin by some unknown mechanism.

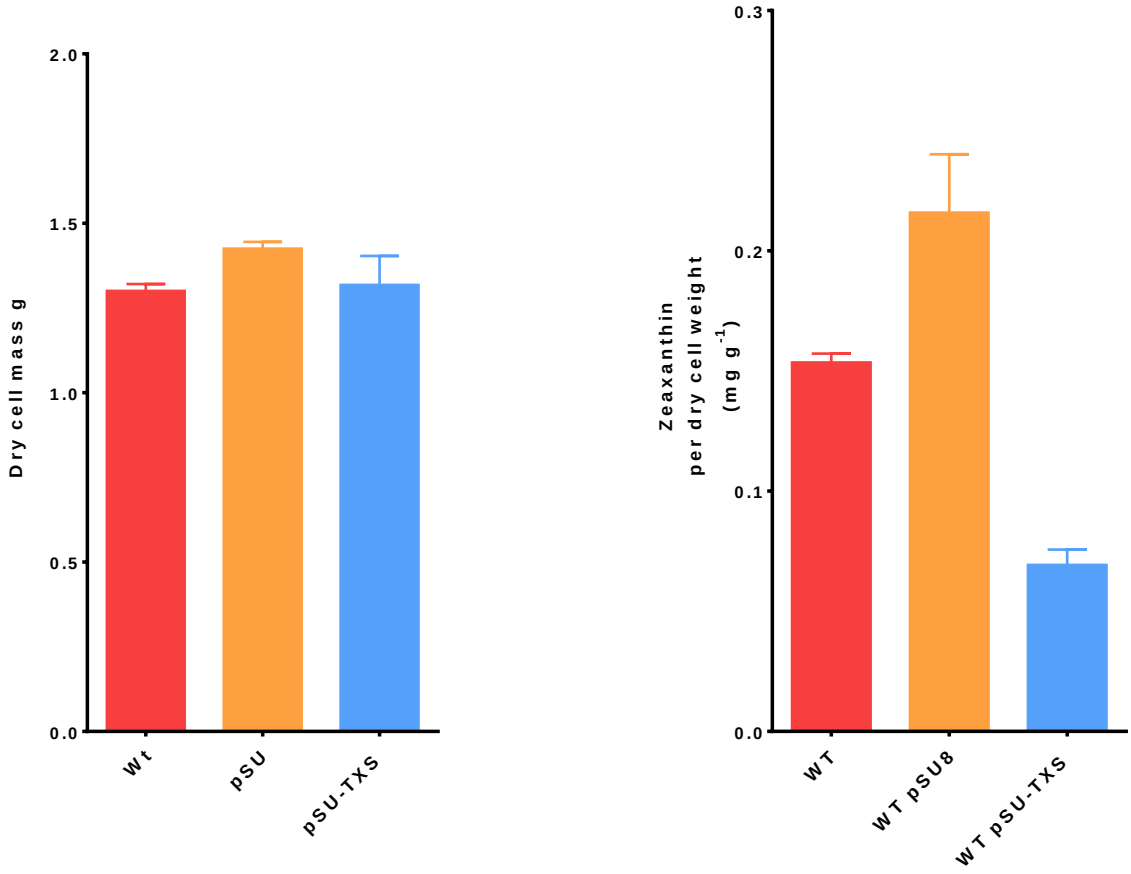


Figure 7-1 Production of zeaxanthin in wild-type *P. agglomerans* with and without the plasmids pSU8 or pSU-TXS A. Dry cell weight, B. Zeaxanthin production by gram of dry cell weight.

7. 2 Characterisation of pBT-5α, pBT-10β, pBT-5α10β

With *txs* transformed into *P. agglomerans* and producing an observable phenotype, the next step was to attempt to co-express the p450 enzymes, taxadiene 5α hydroxylase (5α) and taxane 10β hydroxylase (10β) alongside pSU-TXS. The first step was to characterise each of the individual plasmids for their effect on the growth of *P. agglomerans*.

7.2.1 Transformation into *P. agglomerans*

P. agglomerans was transformed by electroporation with pBT-5α, pBT-10β, and pBT-5α10β in 3 separate reactions. They were then recovered for 1 hours at 30°C 150 RPM then plated onto LB tetracycline and incubated at 30°C overnight. Colonies grew for all transformation reactions and the colonies looked like wild-type *P. agglomerans* colonies. Colonies were screened by restriction digest.

7.2.2 Growth curves

Cultures were set up from single colonies of: *P. agglomerans* transformed with either pBT-5α, pBT-10β, pBT-5α10β or pBT-GFP in LB tetracycline media, *P. agglomerans* transformed with pSU9 or pSU-TXS in LB chloramphenicol, and wild-type *P. agglomerans* in LB. These were grown overnight at 30°C shaking at 150RPM. These cultures were then diluted 1:100 into the fresh media with appropriate antibiotics and 200 µl of culture was pipetted into each of three wells each in a 96 well plate. The plate was then incubated in a Genios Pro plate reader (Tecan) for 24 hours and the OD600 was measured every 30 minutes, the results are shown in figure 7.2.

P. agglomerans transformed with pSU8 or pSU-TXS grows to the same cell density as wild-type *P. agglomerans*. Transformation with pBT-GFP lowered the maximum cell density achieved compared to wild-type *P. agglomerans*. The replacement of GFP for 5α did not lower the cell density however transformation with pBT-10β reduces the growth rate, but not the overall density of the culture. Transformation with pBT-5α10β leads to a reduced growth rate and lower final cell density.

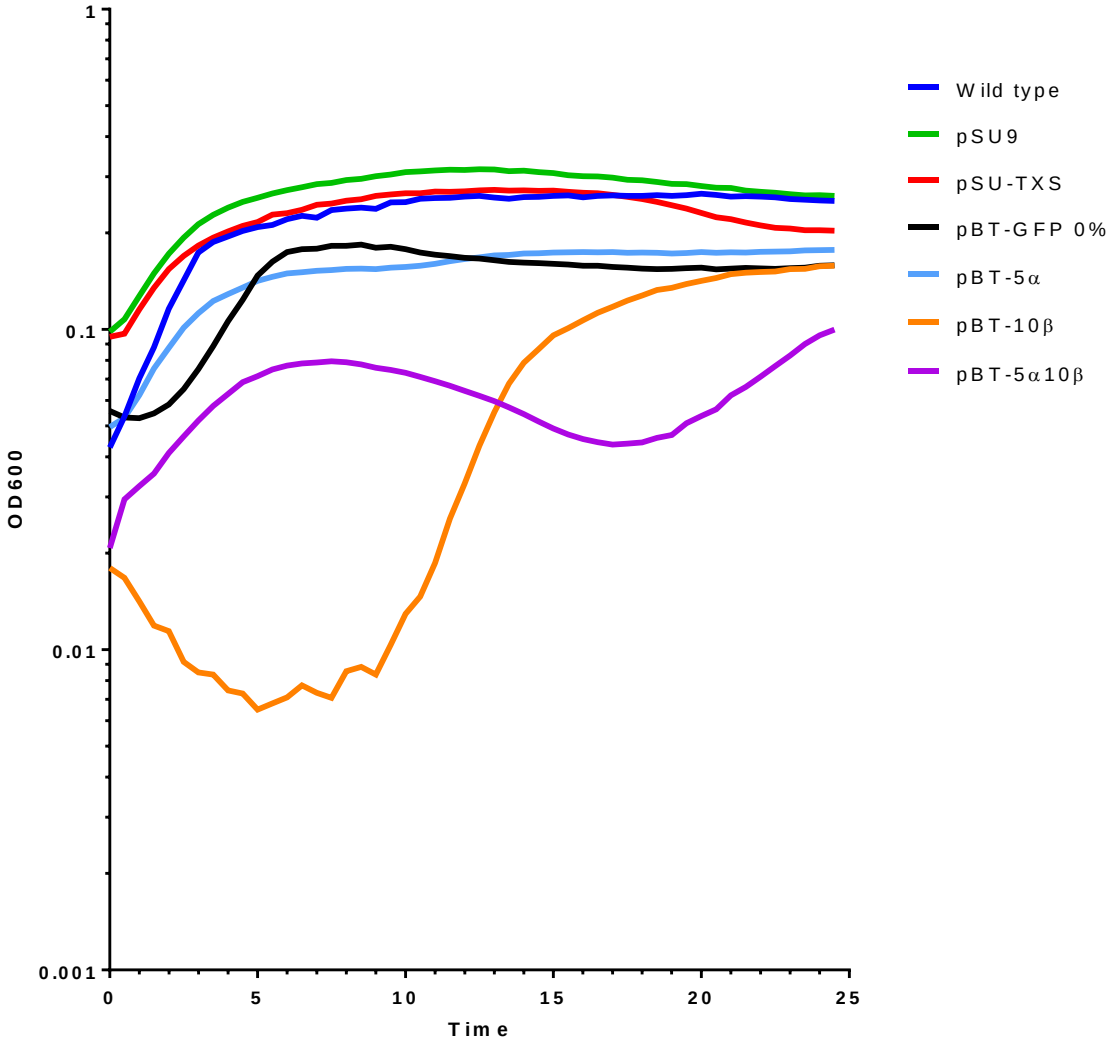


Figure 7-2 Growth curve of several strains of *P. agglomerans*. Strains used: Wild-type *P. agglomerans*, *P. agglomerans* transformed with either pSU9, pSU-TXS, pBT-GFP, pBT-5α, pBT-10β or pBT-5α10β.

7.3 Co transformation of *P. agglomerans* with pSU-TXS and pBT-5α, pBT-10β, pBT-5α10α

7.3.1 Competent cells

In order to co-transform *P. agglomerans* with pSU-TXS and another compatible plasmid (such as pBT-5α, pBT-10β, pBT-5α10β or pBT-GFP), *P. agglomerans* already transformed with pSU-TXS was prepared for transformation by using the same method used in chapter 3. When the culture was between OD600 of 0.6 and 0.8 the cells were pelleted and washed. The cell pellet was noticeably more viscous than previous attempts to make competent cells from wild-type *P. agglomerans.*

The cells were tested for their competency by transforming with pBT-GFP. Transformants were recovered for 1 hour at 30°C and shaking at 150RPM and then incubated overnight on LB ampicillin, tetracycline + chloramphenicol agar at 30°C. Colonies grew and were screened for the presence of both plasmids by plasmid extraction and restriction mapping.

7.3.2 Co-transformation

Following the successful co transformation of *P. agglomerans* with both pSU-TXS and pBT-GFP, *P. agglomerans* pSU-TXS was then transformed with pBT-5α, pBT-10β and pBT-5α10β in three separate transformations. Plasmid DNA used was extracted from *E. coli* and transformants were recovered at 30°C for 1 hour before being plated on LB ampicillin, tetracycline chloramphenicol at 30°C overnight. No colonies were observable after 24 hours, but after 48 a few (< 10) were observable on the plates of *P. agglomerans* pSU-TXS transformed with pBT-10β.

These results supported the suggestion (see section 5.1) that the synthetic operon (in its entirety) was in some way toxic, however this result suggested that in fact it

was the combination of *txs* and 5α and that *txs* and 10β could be transformed with each other into *P. agglomerans.*

7.4 *P. agglomerans* pSU- TXS pBT-10β

The colony morphology of *P. agglomerans* pSU-TXS, pBT-10β was distinct from wild-type *P. agglomerans.* Figure 7.4 shows the difference between wild-type *P. agglomerans* (a), *P. agglomerans* ΔLPP-1 (b) and *P. agglomerans* pSU-TXS pBT-10β. Colonies of *P. agglomerans* pSU-TXS pBT-10β have rough irregular edges and raised sections. The colonies started off pale, but eventually turned more yellow over the course of a few days.

These colonies were verified as *P. agglomerans* pSU-TXS pBT-10β by restriction digest of plasmid DNA.

Growth curves were carried out for the following strains: *P. agglomerans* co transformed with pSU-TXS and pBT-GFP, *P. agglomerans* co transformed with pSU-TXS and pBT-10β, *P. agglomerans* transformed with pBT-GFP, *P. agglomerans* transformed with pSU-TXS. The growth curve was set up using the same method as in chapter 3.

Co transformation of *P. agglomerans* with pSU-TXS and either pBT-GFP or pBT-10β reduces the growth rate and overall cell density of the culture, with *P. agglomerans* pSU-TXS pBT-10β growing slower than *P. agglomerans* co transformed with pSU-TXS and pBT-GFP. This indicated that co transformation with pSU-TXS and any of the pBT series has a noticeable cost to *P. agglomerans*.

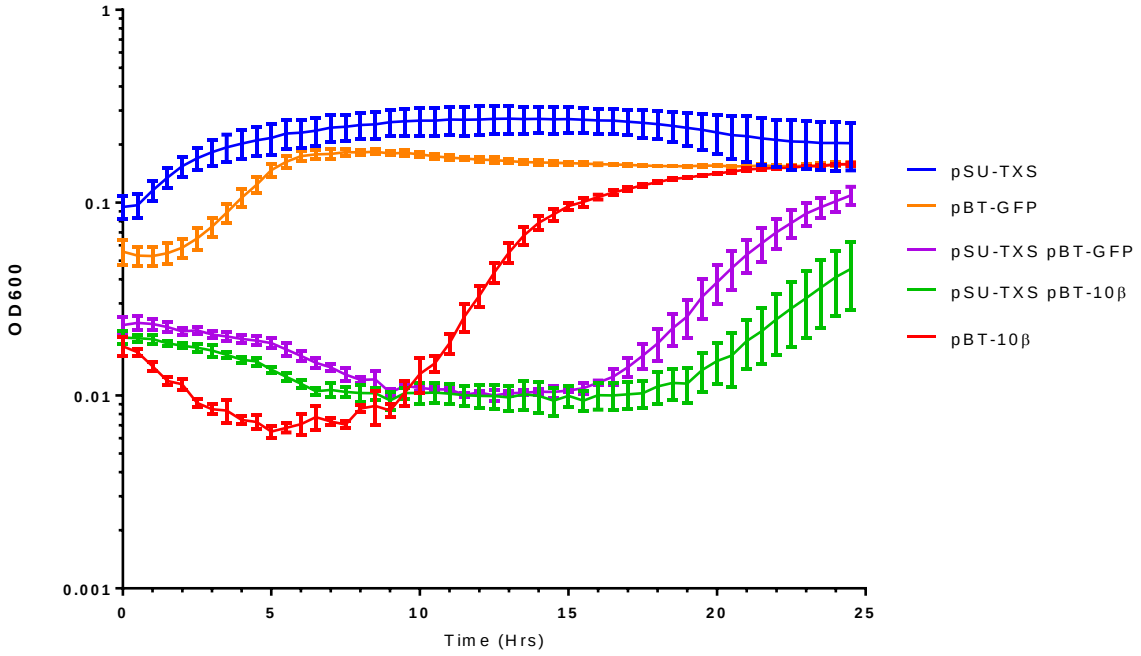


Figure 7-3 Growth curve of co transformed *P. agglomerans* P. agglomerans transformed with pSU-TXS, pBT-10β, or co transformed with pSU-TXS and pBT-GFP or pSU-TXS and pBT-10β

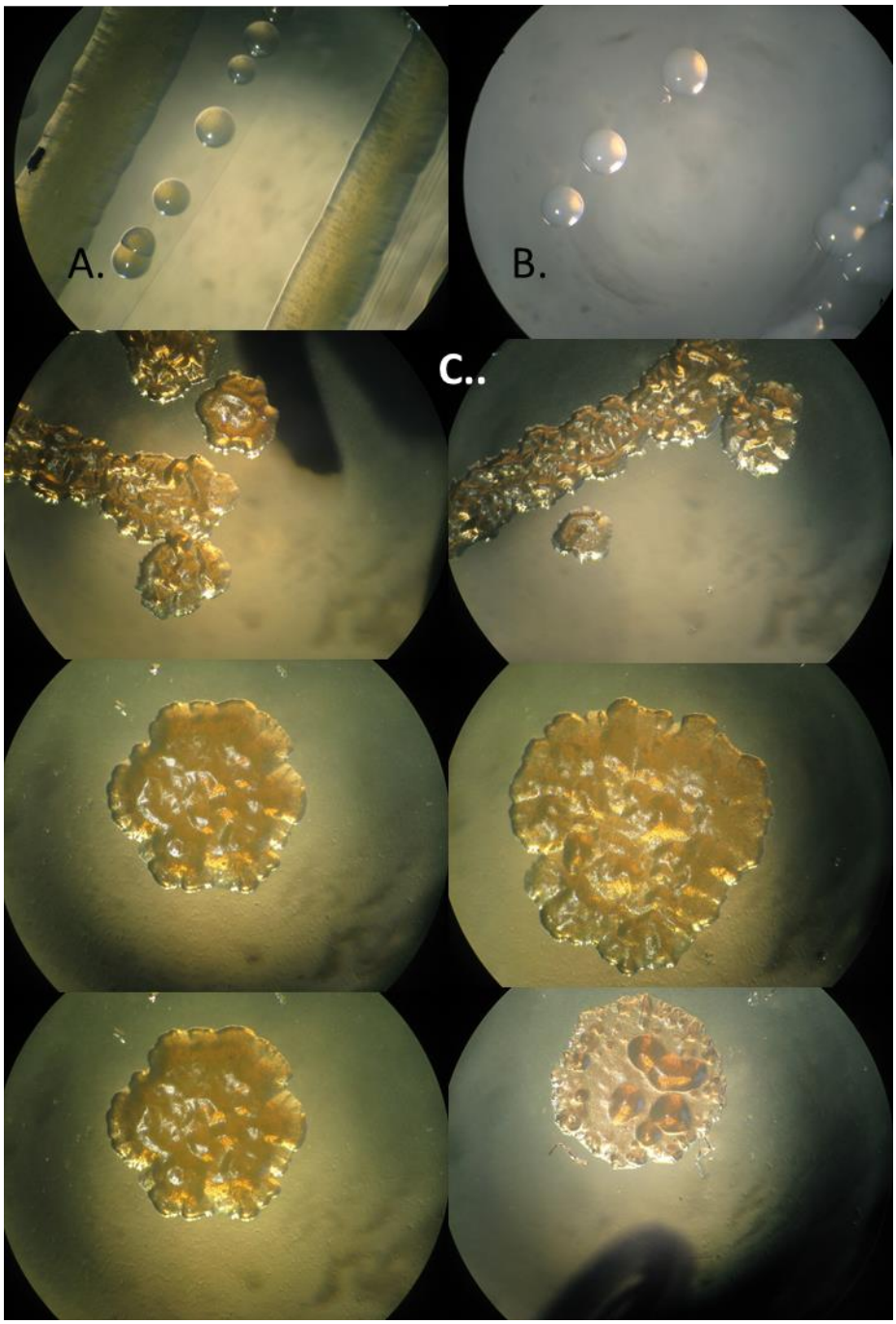


Figure 7-4 Differing colony morphology of *P. agglomerans*, A. *P. agglomerans* B. *P. agglomerans* ΔLPP-1 and C. *P. agglomerans* co-transformed with pSU-TXS and pBT-10β

7. 5 pSU-TXS pLBT-5α and pBT-5α

7.5.1 pLBT-5α

In an attempt to reduce the burden on *P. agglomerans* caused by the co-transformation with *txs* and 5α, it was decided to lower the copy number of the plasmid 5α was expressed on. To do this pLBT-5α was produced by restricting pLBT-GFP and pBT-5α with *Not*I and *Xho*I. The 5α fragment was then ligated into the pLBT backbone. The ligation reaction was transformed into *E. coli* by electroporation.

P. agglomerans ΔLPP-1 was then transformed with pBT-5α and pLBT-5α (in two separate reactions) in order to methylate the plasmid DNA.

Plasmid DNA was extracted from *P. agglomerans* ΔLPP-1 and then used to transform *P. agglomerans* cells containing pSU-TXS by electroporation. Transformations were selected on LB tetracycline, chloramphenicol and ampicillin +0.2% glycerol +0.2% glucose and incubated at 30°C for 48 hours.

The pLBT-5α transformation produced colonies looked like wild-type *P. agglomerans.* Transformation with pBT-5α yielded colonies, which were not expected following the previous results. The success of the transformation can be put down to the improvements made to the transformation of *P. agglomerans* such as; the additional step of using plasmid DNA extracted from *P. agglomerans* ΔLPP-1, and the addition of 0.2% glucose and 0.2% glycerol to LB media to attempt to repress expression from P_{BAD}, (and P_{LAC} to some extent).

The colonies produced by co-transformation with pSU-TXS and pBT-5α had similar morphology to those of pBT-TXS, there were hard irregular colonies which were difficult to remove from the agar plate.

7.6 The pLBT-TXS series

7.6.1 Creation of pLBT-TXS series

As pBT uses a high copy number origin of replication (ColE1, between ~15-20 copies) a proposed solution to the instability problem was to reduce the copy number of the plasmid.

A lower copy number plasmid can have increased stability because having fewer copies of a gene lead to fewer instances of damage and therefore fewer instances of incorrect DNA repair. A lower copy number also makes repression tighter as there is a lower level of basal expression therefore reduces the selection pressure to lose function in these genes.

The temperature sensitive origin of replication pSC101 was chosen as the backbone for this new series of plasmids as it was compatible with the p15 origin of replication of pSU9 as well as ColE1, it has a copy number of ~5 and could be maintained by *Pantoea*. The temperature sensitivity of this origin of replication was also not a major concern when considering it as it replicates faithfully at temperatures around 30°C, similar to *Pantoea's* preferred temperature. pSC101 was already being used in *Pantoea* to integrate genes in to the chromosome via Tn*7* by using pGRG36T (Figure 4-1), this was used as a starting point for the construction of a new, low copy, expression vector.

First the Tn*7* transposase was removed from pGRG36T by restriction with *SbfI* and *NsiI*, the 4.4 Kb backbone was extracted and self-ligated (compatible sites) and was renamed pLT-1 (Figure 7-5). The P_{BAD} promoter and *ara*C repressor of pBT-GFP was ligated into pLT-1 via *NdeI* and *Not*I and was named pLBT-1.

txs and GFP were then inserted into pLBT-1 by *Not*I and *Xho*I digest and ligation to produce pLBT-GFP and pLBT-TXS (Figure 7-6). pBT-5α and pBT-5α10β were restricted with *Nco*I and *Xho*I, pBT-10β was restricted with PciI and *Xho*I, and these fragments

were then ligated into pLBT-TXS in the *Nco*I and *Xho*I sites to produce pLBT-TXS-5α, pLBT-TXS-10β and pLBT-TXS-5α-10β. These were confirmed by restriction mapping and sequencing, Figure.

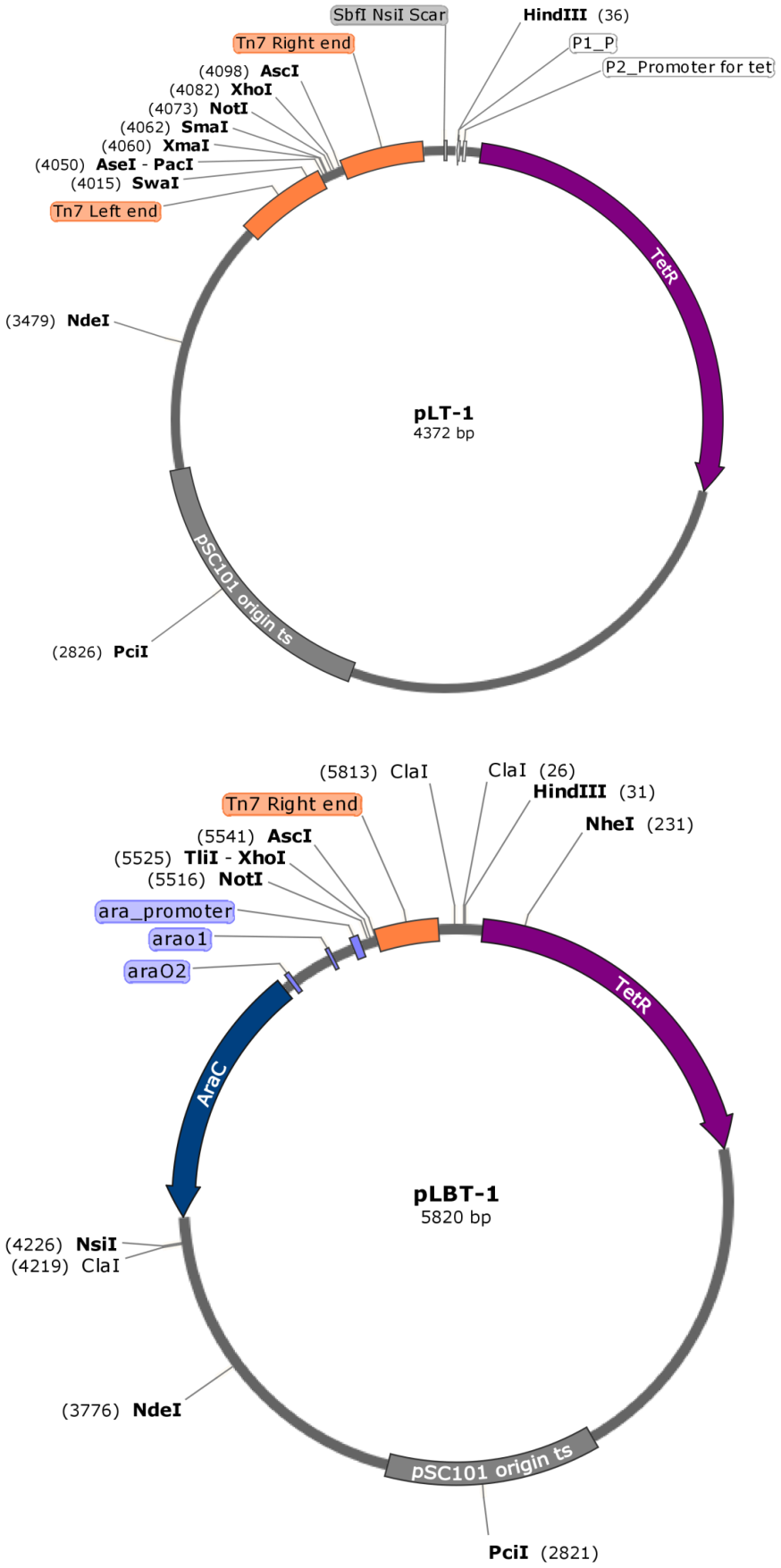


Figure 7-5 pLT-1 and pLBT-1. pLT-1 was constructed by restricting pGRG36T with *SbfI* and *NsiI* and then self-ligating the 4.3Kb fragment. pLBT-1 was produced by restricting both pBT-GFP and pLT with *Not*I and *NdeI* to ligating P_{BAD} into pLT-1.

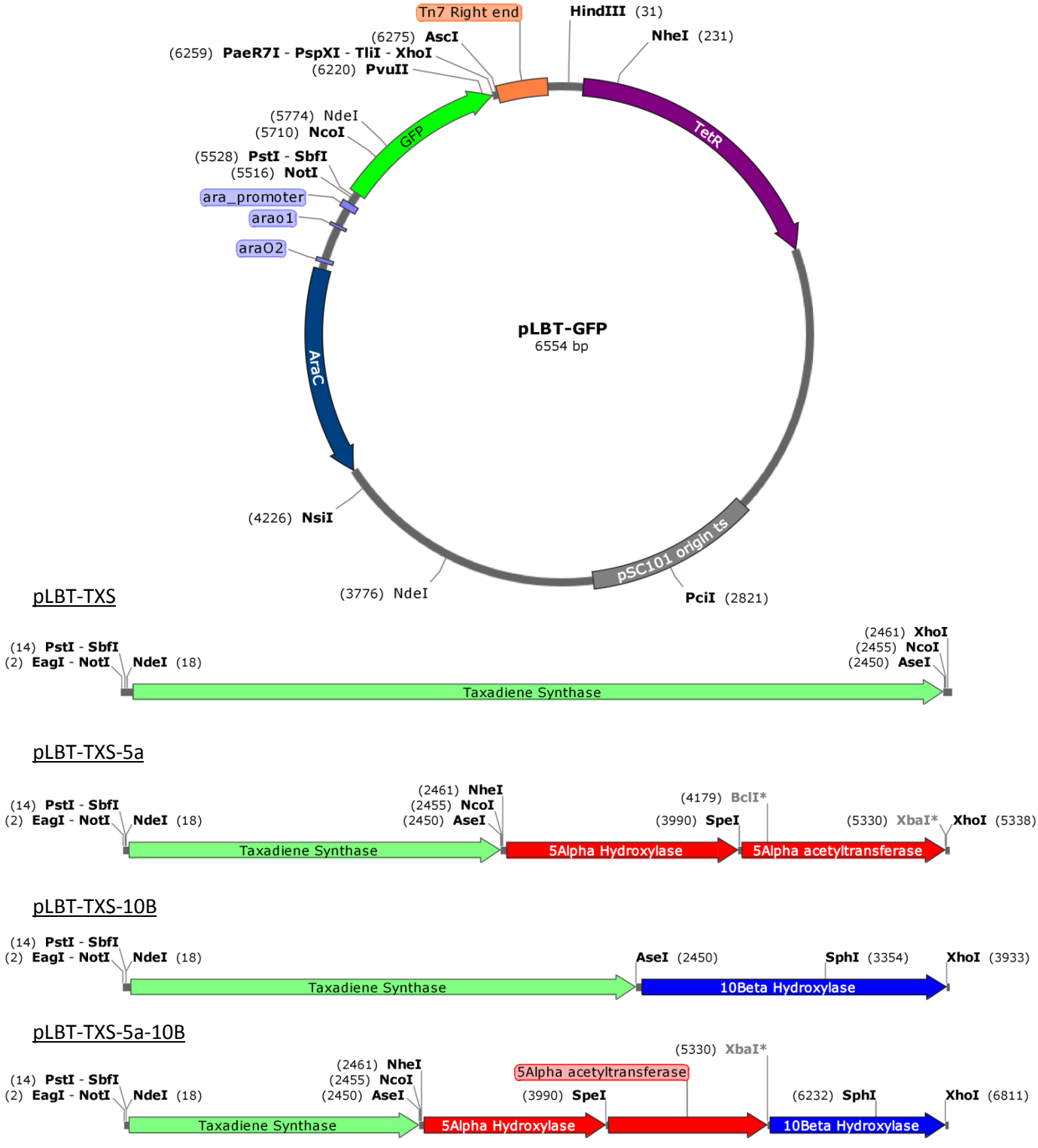


Figure 7-6 pLBT-GFP map and the pLBT-TXS series

7.6.2 Transformation into *P. agglomerans*

Following the construction of the pLBT-TXS series (pLBT-TXS, pLBT-TXS-5α, pLBT-TXS-10β and pLBT-TXS-5α10β) of plasmids in *E. coli,* these were transformed into *P. agglomerans* ΔLPP-1, in order to methylate the DNA to protect it from restriction when transformed into wild-type *P. agglomerans*. Colonies were then picked, restreaked, plasmid DNA prepared and screened by restriction digest. Figure 7. 8 and 7.10 show the confirmation by restriction digest.

DNA extracted from these *P. agglomerans* ΔLpp-1 pLBT-TXS series strains were then used to transform wild-type *P. agglomerans* in the same manner as described for transformation into *P. agglomerans* ΔLPP-1 but with the addition of ampicillin to the LB plates (LB tetracycline and ampicillin +0.2% glycerol +0.2% glucose). These were then picked, restreaked and confirmed by restriction digest, Figure 7. 9 and 7.11.

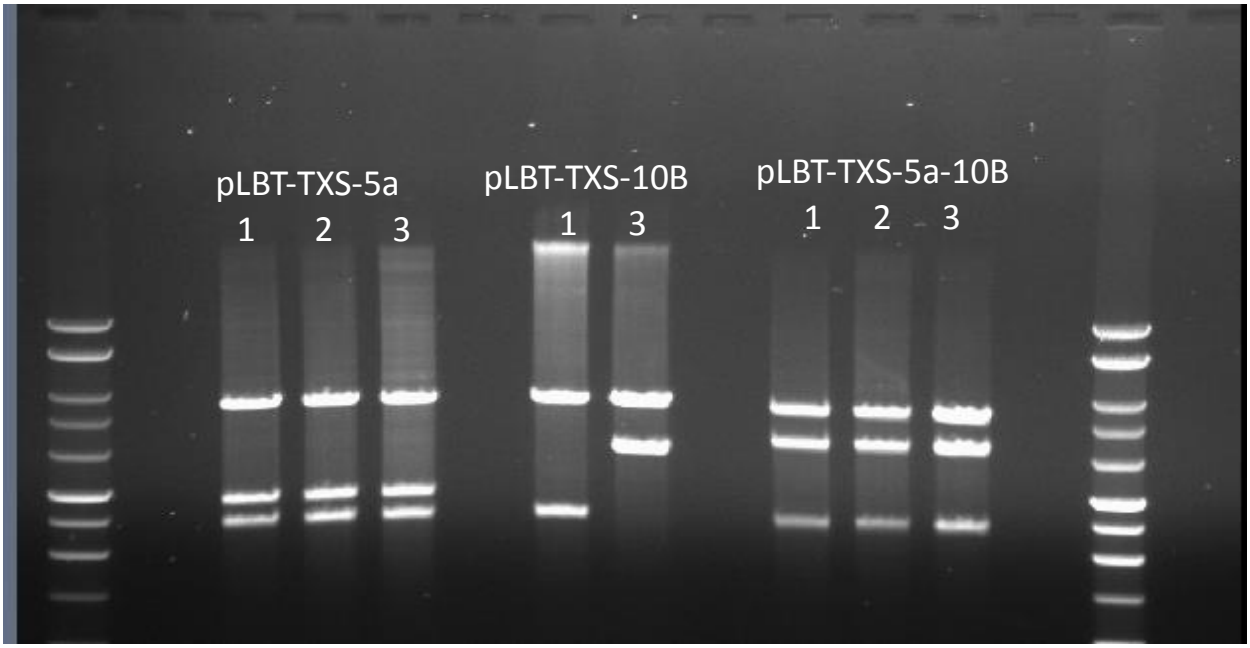


Figure 7-7 Restriction digest confirmation of the production of pLBT-TXS-5α, pLBT-TXS-10β and pLBT-TXS-5α10β pLBT-TXS-5α restricted with *Not*I, *Nco*I and *Xho*I, Expected bands of 5.8Kb, 2.8Kb and 2.4Kb were observed DNA from lane 1 used for further work. pLBT-TX-10β was digested with *Not*I and *Xho*I, expected bands of 5.8Kb and 3.9Kb were observed, DNA from lane 3 used for further work. pLBT-TXS-5α10β was digested with *Not*I *Nco*I and *Xho*I, expected bands of 5.8Kb, 4.3Kb and 2.4Kb were observed, DNA from lane 1 was used for further work

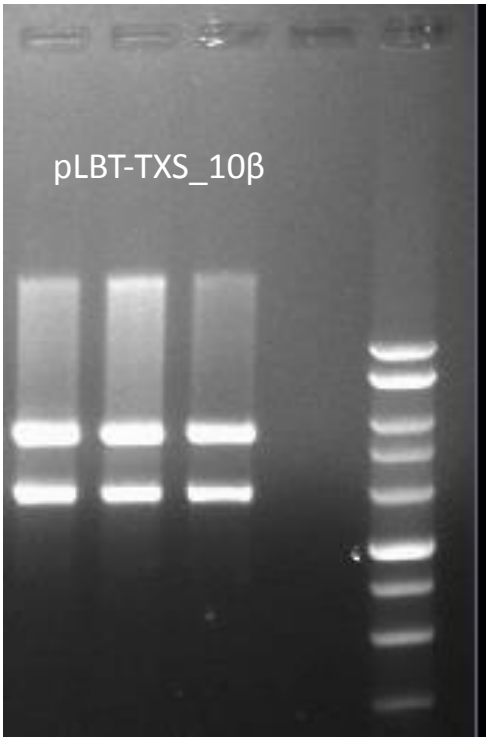


Figure 7-8 Restriction digest confirmation of the transformation of *P. agglomerans* ΔLPP-1 with pLBT-TXS-10β Restriction digest was carried out using *Not*I and *Xho*I, 5.8Kb and 3.9Kb bands were expected and observed.

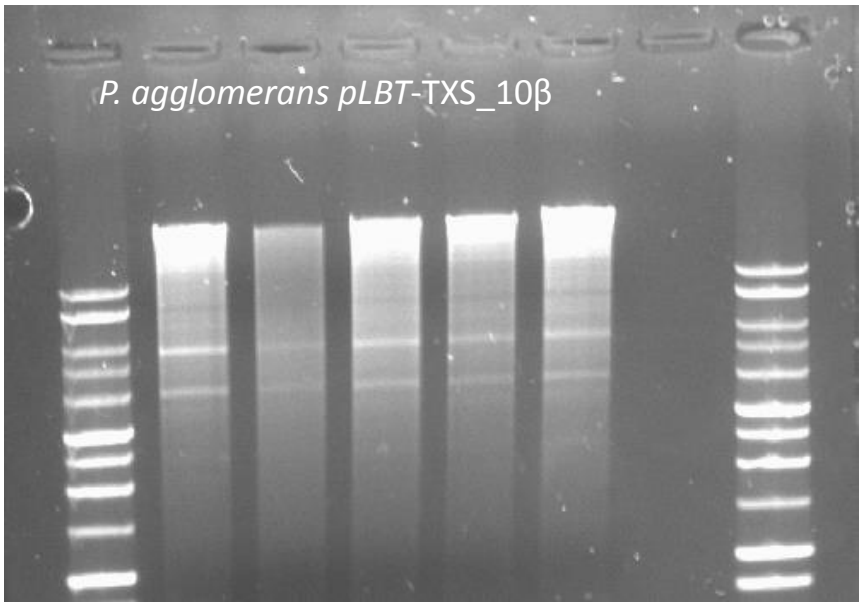


Figure 7-9 Restriction digest confirmation of the transformation of *P. agglomerans* with pLBT-TXS-10β Restriction digest was carried out using *Not*I and *Xho*I, 5.8Kb and 3.9Kb bands were expected and observed.

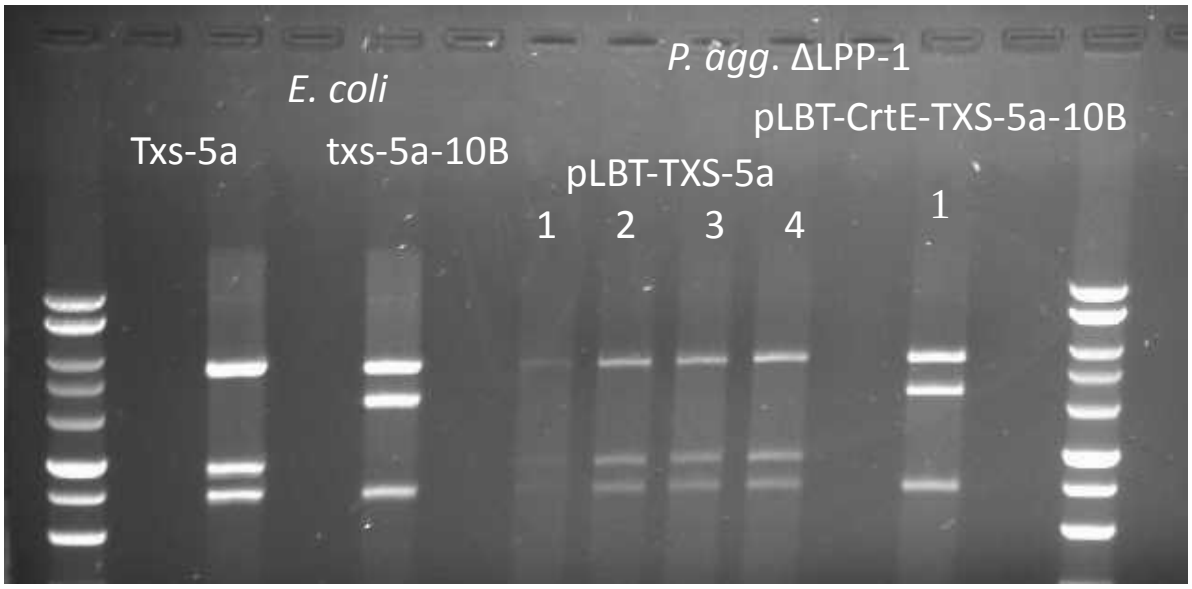


Figure 7-10 Restriction digest confirmation of the transformation of *P. agglomerans* ΔLPP-1 with pLBT-TXS-5α, and pLBT-TXS-5α10βpLBT-TXS-5α restricted with *Not*I, *Nco*I and *Xho*I, Expected bands of 5.8Kb, 2.8Kb and 2.4Kb. pLBT-TXS-5α10β was digested with *Not*I *Nco*I and *Xho*I, expected bands of 5.8Kb, 4.3Kb and 2.4Kb were observed.

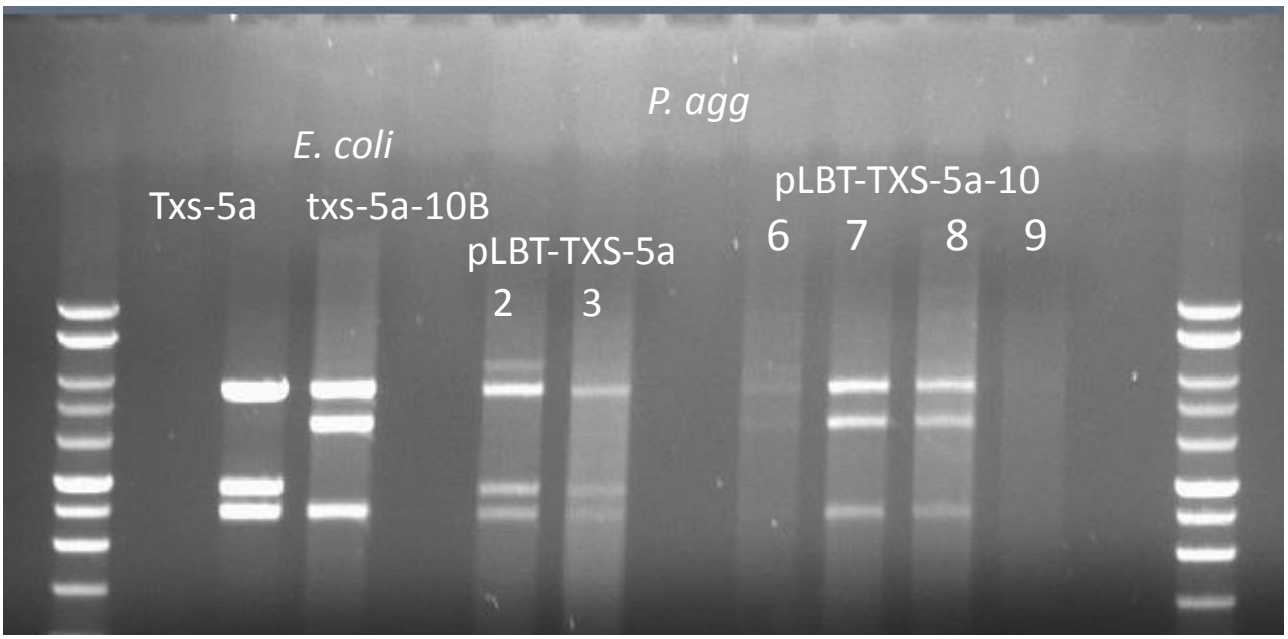


Figure 7-11 Restricion digest confirmation of the transformation of *P. agglomerans* with pLBT-TXS-5α and pLBT-TXS-5α10β pLBT-TXS-5α restricted with *Not*I, *Nco*I and *Xho*I, Expected bands of 5.8Kb, 2.8Kb and 2.4Kb were observed pLBT-TXS-5α10β was digested with *Not*I *Nco*I and *Xho*I, expected bands of 5.8Kb, 4.3Kb and 2.4Kb were observed. *E. coli* plasmid DNA used as controls

7.7 Taxane production

With the successful transformation of *P. agglomerans* with the pLBT-TXS series, therefore giving a range of possible taxanes that could be produced from *P. agglomerans,* the next step was to culture these strains and extract the taxanes they produced.

7.7.1 Culturing

For each of the 4 strains, (*P. agglomerans* transformed with either: pLBT-TXS, pLBT-TXS-5α, pLBT-TXS-10β or pLBT-TXS-5α10β), culture was set up from a single colony in 20 ml of LB tetracycline ampicillin +0.2% glycerol +0.2% glucose, these were then incubated at 30°C overnight. The next morning the cultures were used to inoculate 250ml of LB tetracycline ampicillin +0.2% glycerol +0.2% glucose, these were incubated for 24 hours then the cultures were centrifuged and the pellet re-suspended into 500ml of 2YT ampicillin tetracycline +1% glycerol and were incubated at 22°C. Every 24 hours 50ml of culture was sampled and 50 ml of 2YT +1% glycerol +50 µl tetracycline 10mg l^{-1} 500µl ampicillin 100 mg l^{-1} was added to the culture.

From the 50ml removed from culture several measurements and samples were taken:

- 6ml of culture was frozen so that plasmid DNA could be extracted at a later date.

-OD600 measurements were taken

-Colony forming units were counted by plating out 100µl of samples diluted down to 10^{-6}

Figure 7.12 shows the OD600 measurements and the colony counts.

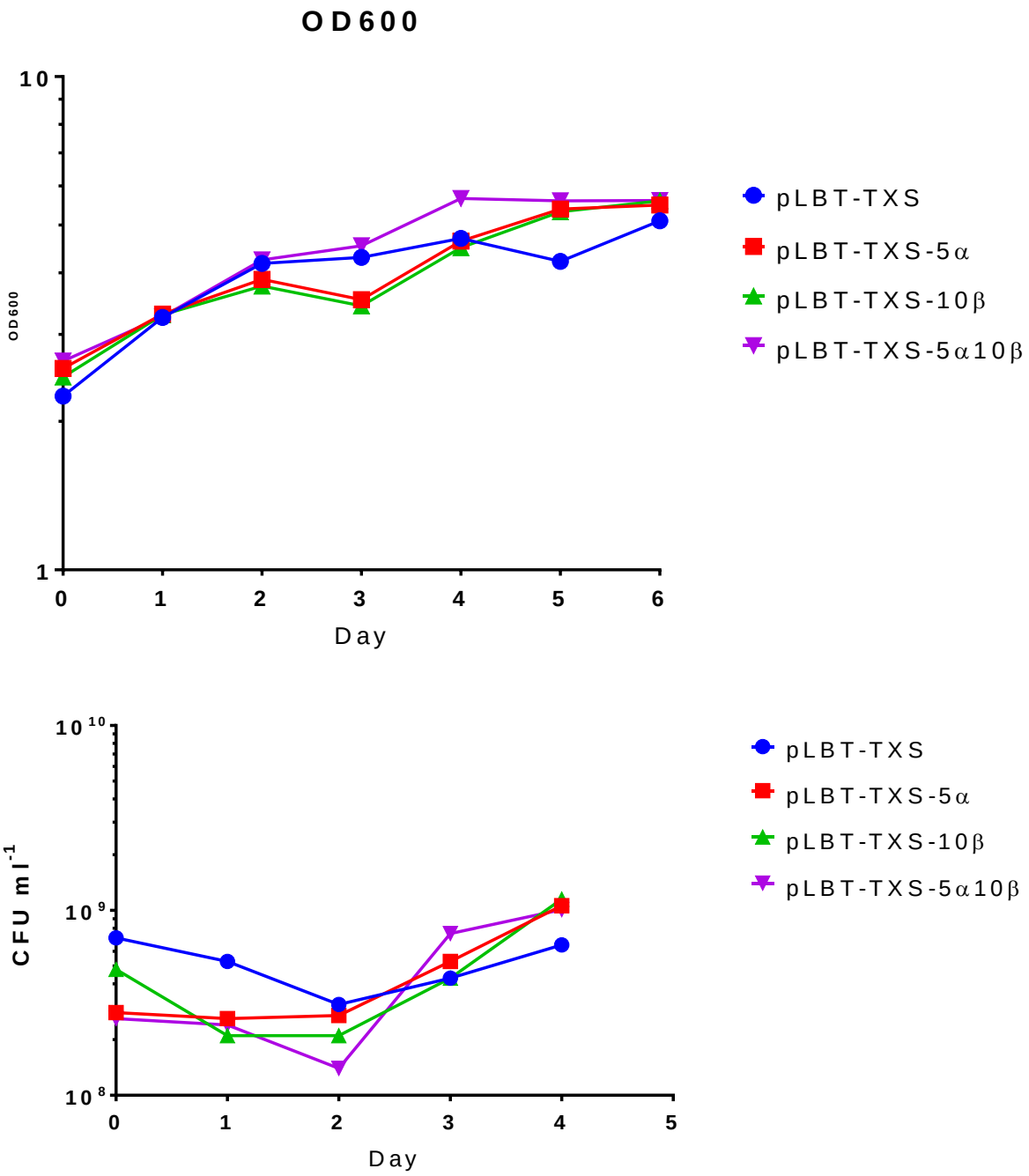


Figure 7-12 OD600 and viable cell counts of cultures of *P. agglomerans* pLBT-TXS series.

A wild-type *P. agglomerans* culture was also set up in 250ml of 2YT 1%glycerol and incubated for 5 days at 22°C, but the measurements previously described were not taken. This culture was designed to be a control for the GCMS analysis.

7.7.2 Taxadiene extraction and analysis

On the 6th day all cultures were taken out of the 22°C shaking incubator, decanted into clean Duran bottles and placed at 4°C. Cultures were then subjected to a 1:1 hexane extraction, 500ml of culture and 500ml of hexane were combined in a 1l separating funnel. These were mixed for 10 minutes and then allowed to settle for 20 minutes. During the mixing process a stable emulsion formed at the interface, which did not readily break down with agitation or over time. As the hexane was the top phase, the mixture was allowed to settle as much as possible and then the water phase was tapped off, stopping once the emulsion reached the tap. The emulsion and the hexane layer were then tapped off. This was allowed to settle and the free flowing hexane was decanted off. The remaining emulsion was then frozen at -20°C. After being frozen any free flowing hexane was decanted off and the remaining emulsion was allowed to thaw at room temp. The water phase and emulsion were then subjected to a second 1:1 hexane extraction as before.

Hexane is a good choice of solvent for dissolving taxadiene (Ajikumar *et al.*, 2010; Kovacs *et al.*, 2007), but as the solubility of the other possible taxanes in hexane was unknown a 1:1 chloroform extraction was also undertaken, as chloroform is good at dissolving oils, fats, waxes and alkaloids due to its property of being a slightly polar organic solvent.

After the two hexane washes, the water phases was mixed with equal parts chloroform in a separating funnel and mixed for 10 minutes, this also produced an emulsion which did not readily breakdown. In this mixture chloroform formed the bottom phase and so the mixture was allowed to settle then the tap was partially opened and chloroform slowly left the funnel without dragging emulsion through. Once the emulsion started to move towards the tap, the tap was closed and the collected chloroform was kept, the emulsion was then run out through the tap along with a small amount of the water phase. This was then frozen at -20°C. It was then removed from the freezer, agitated to break up the partially frozen emulsion, and

was then added back into the separating funnel and allowed to fully thaw out. Free flowing chloroform was then tapped off. The water phase and the emulsion were then washed again with chloroform following the same procedure.

Extracts were then dried down using a rotary evaporator to reduce the volume of solvent to between 10-50ml, these extracts were sent for analysis by GCMS. However the results were inconclusive as the control taxadiene was not detected either. Further work will be conducted to confirm the presence or absence of taxadiene.

7.7.3 Stability

As with the culturing of *P. agglomerans* transformed with pBT-TXS, the cultures of *P. agglomerans* transformed with the pLBT-TXS series were screened by restriction digest following culturing. Plasmid DNA was extracted from the cultures frozen on day 1 (after the 2YT +1% glycerol ampicillin tetracycline media was inoculated) and from the last day (day 6). These were then screened by restriction digest to assess the integrity of the plasmids during the course of culturing. The gel was not as clear as it could have been, but this was in part due to using the low copy number plasmid origin of replication pSC101 and the low yield of plasmid it provides. This led to an overloading of material into the column meaning extra genomic DNA was present in the gel, more so than is normally seen in *P. agglomerans* plasmid DNA gels.

pLBT-TXS was restricted with *Pst*I and *Xho*I. After 1 day of culturing the bands of pLBT-TXS are clearly distinguished in the restriction digest, but it became apparent that *P. agglomerans* transformed with pLBT-crtE-TXS had been used instead of pLBT-TXS (6797bp, 2451bp, instead of 5819bp, 2451bp). pLBT-crtE-TXS is the same as pLBT-TXS except for the insertion of *crtE* downstream of P_{BAD} and upstream of *txs*. After 6 days the 6797bp band was faintly present, however the 2451bp band is hard to see due to the amount of DNA present on the gel, the 2451 band is the *txs* band, so it is unknown whether *txs* is intact or not (Figure 7-13).

pLBT-TXS-5α was restricted with *Xho*I and *Nco*I. On both days both bands are visible, however they are rather pale.

pLBT-TXS-10β was restricted with *Xho*I and *Pst*I. The *E. coli* sample shows an extra band due to using too much DNA in the restriction digest and the top band is a single restriction leaving linearised product. The restriction map looks correct for both day 1 and 6, however slightly paler on day 6.

pLBT-TXS-5α-10β was restricted with *Xho*I and *Nco*I, results from day 1 are obscured by excessive DNA however bands are visible after 6 days of culture.

Additionally the levels of tetracycline in the media were not maintained during the course of culturing, like ampicillin was, because the mechanism of resistance to tetracycline used (*tetA*), does not degrade the tetracycline unlike the mechanism of resistance to ampicillin. However the natural half-life of tetracycline in culture was not taken into consideration, in future culturing the addition of tetracycline should be increased to maintain selection. The effect of this drop in tetracycline is clearest in pLBT-TXS-10β on day 6 the bands are still present, showing that the plasmid has not degraded, but it is a lot fainter.

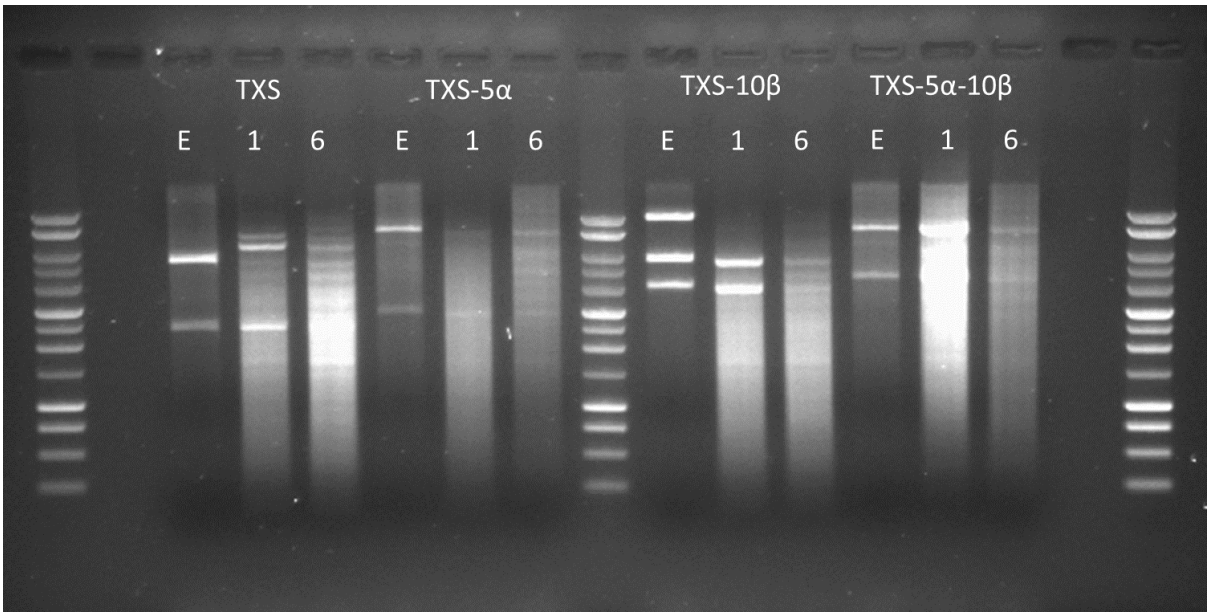


Figure 7-13 Plasmid Stability of pLBT-TXS series in P. agglomerans over 6 days of culture. E. *E. coli* plasmid DNA control 1. Day 1, 6. Day 6... TXS, E. pLBT-TXS restricted with *Pst*I and *Xho*I, 5819bp, 2451bp expected bands observed. 1 and 6. *P. agglomerans* transformed with pLBT-crtE-TXS, 797bp, 2451bp bands expected and observed. TXS-5α, pLBT-TXS-5α restricted with *Nco*I and *Xho*I, expected bands were 8.2Kb and 2.9Kb. TXS-10β. pLBT-TXS-10β digested with *Pst*I and *Xho*I expected bands were 5.8Kb and 3.9Kb. TXS-5α-10β. pLBT-TXS-5α10β restricted with *Nco*I and *Xho*I, 8.2Kb 4.4Kb bands expected.

7. 8 Discussion

The transformation of *P. agglomerans* with pSU-TXS clearly has some effect on the utilisation of GGPP as seen by the reduction in zeaxanthin production per gram of dry cell weight when compared to both the wild-type and the pSU9 control strain. Furubayashi *et al.*, (2014) have shown that a reduction in carotenoids produced by a culture can be related to the functionality of *txs*, by screening a range of mutant *txs* in lycopene producing *E. coli*. In *P. agglomerans* the inability to detect the presence of taxadiene by GCMS leaves a hole in the argument that *txs* is diverting GGPP into taxadiene rather than zeaxanthin. An additional control could be to create a non-functional *txs* by creating a point mutation in the active site. Using a greater volume

of culture could also help in detection of taxadiene, in this experiment there was a relatively small culture used compared to the later experiments with the pLBT-TXS series of plasmids.

Additionally it is not know where in the culture taxadiene is most likely to accumulate. Due to its lipophilic nature, the cell membrane has a high probability of being a location with a high concentration of taxadiene. It may be that if taxadiene does not pass into the aqueous section of the culture that solvent extraction first required a centrifugation step.

Reviewing the work of Ajikumar *et al.*, (2010) they found that 1:1 hexane culture extraction produced detectable levels of taxadiene, however they also showed that taxadiene can be volatile and leave the culture via air stripping. Addition of a layer of oil may be necessary for future work.

The phenotypes observed on *P. agglomerans* pBT-TXS and *P. agglomerans* pSU-TXS pBT-10β plates, strongly suggest that there is production of a chemical, or a change within the cells physiology which is causing a change in the behaviour of the bacteria. Whilst pBT-TXS appears to be unstable in *P. agglomerans* the combination of *txs* and 10β, be it in the form of pSU-TXS and pBT-10B or as pLBT-TXS-10B, appears to be stable. This either means that the addition of 10β has no effect, or it in fact helps to reduce the presumed toxicity of *txs* expression, although there is no direct evidence to support this. A potential experiment to test this would be to take 10β and express it on pSU9 and see if the co transformation of *P. agglomerans* with pSU-10β and pBT-TXS improves the stability of pBT-TXS. Either by transforming *P. agglomerans* pSU-10β cells with pBT-TXS extracted from *E. coli,* and comparing the length of time that *txs* can be detected in culture for compared to *P. agglomerans* pSU9 co-transformed pBT-TXS.

The initial inability to co-transform *txs* and 5α and then the observation of similar phenotype to that seen in *P. agglomerans* transformed with pBT-TXS also points to bioactivity of the products of these genes.

The issue of co-transformation of *P. agglomerans* with *txs* and 5α was overcome in two different manners, one with the improvement of the transformation procedure, therefore increasing efficiency after implementing the additional step of transforming into the non-counter selective background of *P. agglomerans* ΔLPP-1 in order to methylate the plasmid DNA, and the other the addition of glycerol and glucose to the agar in order to catabolically repress expression.

Despite the initial difficulties with cloning the completed synthetic operon, it was eventually possible to clone *txs*, 5α and 10β into a single operon. pLBT provided two important features in order to allow this, firstly the P_{BAD} inducible promoter, when repressed with glucose, gives a very low level of expression, secondly the low copy number origin of replication pSC101 ensured that any leak in transcription is reduced, compared to higher copy number plasmids. .

As a future experiment to attempt to improve the stability of the taxane pathways in *P. agglomerans*, the pLBT-TXS series have all had *kanR* added to the end of each set of genes, these were intended to be ligated into pGRG36T and then integrated into the chromosome of *P. agglomerans.*

Looking in the literature, the expression of taxane 10β-hydroxylase (10β) to produce taxadiene-5α-acetoxy-10β-ol has only been carried out by expressing 10β in yeast and supplementing it with taxadien-5α–yl-acetate (Schoendorf *et al.*, 2001), making this current work unique in attempting both hydroxylation steps simultaneously. This was carried out to confirm that the enzyme scoendorf *et al.* had identified and cloned from *Taxus spp.* was in fact the taxane 10β hydroxylase.

Future work will need to investigate appropriate p450 reductases for use within *P. agglomerans*, so as to efficiently cycle the hydroxylation reaction. This may involve

screening a selection of different plant derived sequences, and then redesign and codon optimise the gene to improve expression in *P. agglomerans*.

GCMS analysis of the pLBT-TXS series of cultures has still not detected taxadiene or any other taxanes, this comes as a disappointment, but with the final work of the project being the transformation *P. agglomerans* with the full original synthetic taxane operon, it brings the project back to the original aims, and opens up the possibility to improve the culturing of *P. agglomerans* and re attempt the search for taxadiene and novel taxanes.

Chapter 8 DISCUSSION

Pantoea is a diverse genus of bacteria with great potential for use in a wide range of biotechnological applications including: aiding agriculture, amino acid production and the therapeutics industry. This thesis has explored the potential for using *Pantoea* for the production of taxadiene and hydroxylated derivatives of it.

8.1 Tools that have been developed

Throughout this project developments have been made in the tool set available for use in *P. agglomerans* and attempts made to try and modularise them so that they can easily be adapted further. For example using the same MCS from pGRG36T when designing pBT-GFP so that any genes assembled in pBT-GFP could be lifted and ligated straight into pGRG36T. This design choice helped to save time and streamlined decision making as to which restriction sites to use.

Following on from chapter 7 and the production of the pLBT-TXS series, future work would be to evaluate this expression system with the *lux* and *crtEIB* reporter systems produced in chapter 3 by producing pLBT-LUX, pLBT-crtEIB and pLBT-crtE-idi-IB to assess the outputs under different levels of arabinose induction. This would help inform the culturing of *P. agglomerans* transformed with the pLBT-TXS series to produce taxanes.

It would also be interesting to co-transform *P. agglomerans* with the pLBT-TXS series and pCT-EIB (after replacing the *tetR* cassette with chloramphenicol resistance) to investigate the effect on the levels of lycopene produced.

Once taxadiene can be detected in these systems it would be useful to be able to correlate a drop in lycopene with the amount of taxadiene produced. This would allow for more culture conditions to be screened for their ability to produce taxadiene. This method has already been used to screen randomly mutated *txs*, firstly

by removing non-functioning mutants and then to select for the mutant with the biggest drop in lycopene (Furubayashi *et al.*, 2014), however this group did not correlate drops in lycopene with taxadiene production.

Genome editing was not as successful an endeavour as producing the reporter systems, but the pGRG36T system might now work better with the greater selection pressure granted by the addition of a *kanR* cassette to the end of the operon being integrated into the chromosome. Compared to other genome editing processes, the Tn*7* method can be a quick protocol. Once the cargo DNA has been ligated into pGRG36T, the protocol can be applied to several different strains and species simultaneously, unlike homologous recombination based methods. This means that multiple species of *Pantoea* could be screened in parallel for their ability to produce taxadiene in this manner.

In terms of producing deletion mutants there are several options moving forward, re attempt the current method attempted, using a different counter selection method, alternatively the Lambda Red system could be investigated as mentioned earlier. However, given the versatility of CRISPR as a tool, it could be worth investing time into transferring a system already used in *E. coli* for use in *Pantoea*. Theoretically the work would be minimal, as it should work identically in *P. agglomerans*, however it would be breaking new ground as there are no reported uses of CRISPR in the literature (however industrial users of *Pantoea*, such as Ajinomoto and Bloomtime, may be using it), but it functions in *E. coli*. As well as gene editing CRISPR has great potential for use as a genetic control in the form of CRISPRi. CRISPRi uses a modified Cas9 with a non-functional nuclease, this Cas9 binds to the PAM encoded by the guide RNA. These can be put under the control of an inducible promoter so that induction of the promoter switches off the targeted gene. This can be used to disrupt competitive biosynthetic pathways, which is an especially useful tool for competitive pathways which are vital to cell growth. Cultures can be grown to a high cell density,

and then the pathway essential for growth can be switched off and the desired biosynthetic pathway can be switched on.

An interesting tool in development for use in bacteria for production of potentially toxic products, or products with toxic intermediates is the use of microcompartments. Microcomparments are proteinaceous organelles composed of self-assembling subunits. These organelles house enzymes involved in a specific pathway with toxic intermediates, for example the catabolism of propanediol to propionic acid. The use of microcompartments in synthetic biology has been reviewed by Frank *et al.* (2013). If intermediates of Taxol biosynthesis are found to be toxic to *P. agglomerans*, or other bacterial systems used, then housing the biosynthesis in these organelles could dramatically increase the ability to generate taxanes. The group run by Danielle Tullman-Ercek (Northwestern University, Illinois, USA), are prominent in the field, and we are now in contact to try and express functional microcompartments in *P. agglomerans*, with an aim to eventually try and encapsulate some or all of the taxane pathway in order to improve production.

8.2 Improvements to stability

At the beginning of this project, the taxane operon seemed to be unstable and/ or toxic to both *E. coli* and *P. agglomerans*, however through the course of this work methods have been found to more stably express *txs*, 5α and 10β and allowed the transformation of *P. agglomerans* with a plasmid expressing all three taxane gene units. With the improved stability of the strain, it is now hoped that with improvements to the culturing and taxane extraction methods, that significant quantities of taxadiene and the following intermediates of taxadiene can be produced.

Work on stability will still need to be carried out, and will most likely be a constant problem moving forward unless methods to mitigate the bioburden and negative selection pressure associated with the production of secondary metabolites can be

found. Potential ways to mitigate stress imposed on cells would be to introduce a transport pump for the removal of taxadiene and its derivatives. As well as improving the removal of taxanes from the cell transport pumps could help improve the extraction of taxanes from cultures. ABC transporters from chemotherapy resistant cancer cells and endophytic fungus may provide a good starting point for investigation, but so might a broader search of transport pumps from fungi and bacteria associated with plants especially *Taxus spp.*

Once taxadiene, and some of the derivatives from it, have been produced, these can be used to try and elucidate if they are toxic to bacteria and the mechanism by which they are. This insight would help develop strategies to counter the potential toxicity and improve yields.

From this work, it appears that the potential toxicity of taxadiene production can be linked to an increase in biofilm producing behaviour in *P. agglomerans*, but no direct link can be found in the literature. However studies have been carried out on the toxicity of Taxol to pathogens such as *M. tuberculosis*, from these studies it was found that Taxol was able to inhibit *Fts*Z, one of the main cell division proteins in bacteria, despite *Fts*Z and tubulin only having 10% sequence similarity. Work since has looked at producing analogues which are still inhibitory to M. tuberculosis but less cytotoxic to humans (Huang *et al.*, 2006; Huang, *et al.,* 2007; Ojima *et al.,*2014). The development of analogues that still exhibit the inhibition of *Fts*Z provides the possibility that precursors to Taxol may hold similar properties. Inhibition of cell division is often associated with stress and survival responses (Boberek, *et al.,* 2012), stress responses are known to be linked to regulation of cyclic-di-GMP, which in turn regulates the formation of cellulose (by modulation of cellulose synthase) and therefore upregulates biofilm formation (Hengge, 2009). Although there is little experimental evidence for the link between taxadiene and FtsZ inhibition, this hypothesis opens up an intriguing line of investigation for future work.

The hard colony morphology is an interesting phenotype, and with some further work could help to understand the role *Pantoea* can play in infections of plants and humans, as well as the mechanics of biofilm formation and dispersal. As this phenotype is easily propagated on media, there are possibilities of using it to screen agents that could prevent the formation of biofilms or cause them to disassemble.

8.3 E. coli vs *P. agglomerans*

From this work it is hard to say conclusively either way if *P. agglomerans* is a better host organism for the biosynthesis of Taxol intermediates than *E. coli.* However this work has helped to show that *P. agglomerans* can be used for biotechnological applications, and if production of high value terpenoids such as Taxol turn out to not be viable in *P. agglomerans* (due to advances in *E. coli*), then *Pantoea* as a species may still have a niche to be exploited in industrial biotechnology.

The biosynthesis of high value terpenoids such as Taxol, economically allows the use of high grade starting materials; feedstock with few impurities or inhibitors of growth. This is an uncommon situation in an industrial setting. Bacterial production of products involve a large initial investment, and therefore the profit margins need to be significant to warrant it. Once running, feedstocks are incredibly important in dictating the profitability of a product. Through over a hundred years of lab culturing, *E. coli* has potentially lost some of its 'genetic vigour', usually maintained by constantly being challenged by the environment. This may lead it to be more fastidious on cheaper feedstocks than wild isolates.

With the diversity of niches that *Pantoea* can occupy, strain selection could be dictated by the conditions required for fermentation of a product. Ajinimoto™ have taken this approach when selecting their L-glutamic acid producing strain (Hara *et al.*, 2012). The genetic diversity of *Pantoea*, combined with its genetic similarity to *E. coli,* could make it an incredibly useful bio-production strain moving forward.

Products which *Pantoea* may be useful for producing are lower value terpenoids such a limonene, or following the work of Ajinomoto, amino acid biosynthesis could be used as a platform for production of an array of organic chemicals. An example of this is the work of McKenna & Nielsen, (2011), who managed to create a synthetic pathway for the production of styrene by expressing Phenylalanine ammonia lyase from *A. thaliana* and Ferulate decarboxylase from *S. cerevisiae* in *E. coli* . This pathway converted L-phenylalanine to trans-cinamic acid and then to styrene.

Work in this direction would involve investigating the range of feedstocks *Pantoea* can utilise, and prospecting for strains with unusual properties of benefit for biotransformation of feedstocks to high value products.

Future work on *Pantoea* should aim to further develop this bacteria for industrial use, as well as looking to the diversity of the genus for and additional advantages that can be gained.

REFERENCES

Ajikumar, P. K., Xiao, W.-H., Tyo, K. E. J., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon. T, H., Pfeifer, B. & Stephanopoulos, G. (2010). Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science (New York, N.Y.)*, *330*(6000), 70–4.

Albrecht, M., Misawa, N., & Sandmann, G. (1999). Metabolic engineering of the terpenoid biosynthetic pathway of *Escherichia coli* for production of the carotenoids beta-carotene and zeaxanthin. *Biotechnology Letters*, *21*, 791–795.

Allers, T., Ngo, H. P., Mevarech, M., & Lloyd, R. G. (2004). Development of Additional Selectable Markers for the Halophilic Archaeon *Haloferax volcanii* Based on the *leuB* and *trpA* Genes. *Applied and Environmental Microbiology*, *70*(2), 943–953.

Alper, H., & Stephanopoulos, G. (2007). Global transcription machinery engineering: A new approach for improving cellular phenotype. *Metabolic Engineering*, *9*(3), 258-267.

Amos, L., & Löwe, J. (1999). How Taxol® stabilises microtubule structure. *Chemistry & Biology*, *6*(3), 65–69.

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. *Http://Www.Bioinformatics.Babraham.Ac.Uk/Projects/Fastqc/:*.

Anwar, N., Rouf, S. F., Römling, U., & Rhen, M. (2014). Modulation of biofilm-formation in *Salmonella enterica* serovar Typhimurium by the periplasmic DsbA/DsbB oxidoreductase system requires the GGDEF-EAL domain protein STM3615. *PloS One*, *9*(8), e106095.

Argiris, A., Karamouzis, M. V, Raben, D., & Ferris, R. L. (2008). Head and neck cancer.

Lancet (London, England), *371*(9625), 1695–709.

Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, Andrei, L., Overbeek, R. A., McNeil, Leslie K. Paarmann, D., Paczian, T., Parrello, B., Pusch, G. D., Reich, C., Stevens, R. Vassieva, O., Vonstein, V., Wilke, A. & Zagnitko, O. (2008). The RAST Server: rapid annotations using subsystems technology. BMC Genomics, 9, 75

Baloglu, E., & Kingston, D. G. I. (1999). A new semisynthesis of paclitaxel from Baccatin III. *Journal of Natural Products*, *62*(7), 1068–1071.

Besumbes, Ó., Sauret-Güeto, S., Phillips, M. A., Imperial, S., Rodríguez-Concepción, M., & Boronat, A. (2004). Metabolic engineering of isoprenoid biosynthesis in *Arabidopsis* for the production of taxadiene, the first committed precursor of Taxol. *Biotechnology and Bioengineering*, *88*(2), 168–175.

Beyer, P., Al-Babili, S., & Ye, X. (2002). Golden rice: Introducing the β-carotene biosynthesis pathway into rice endosperm by genetic engineering to defeat vitamin A deficiency. *The Journal of Nutrition*, 506–510.

Boberek, J. M., Goh, S., Stach, J., & Good, L. (2012). Filamentous temperature-sensitive mutant Z (FtsZ) protein as an antibacterial target. *Antimicrobial Drug Discovery: Emerging Strategies*, (September 2016), 135–146.

Boghigian, B. a, Armando, J., Salas, D., & Pfeifer, B. a. (2012). Computational identification of gene over-expression targets for metabolic engineering of taxadiene production. *Applied Microbiology and Biotechnology*, *93*(5), 2063–73.

Boghigian, B. a, Salas, D., Ajikumar, P. K., Stephanopoulos, G., & Pfeifer, B. a. (2012). Analysis of heterologous taxadiene production in K- and B-derived *Escherichia coli*. *Applied Microbiology and Biotechnology*, *93*(4), 1651–61.

Bóna-Lovász, J., Bóna, A., Ederer, M., Sawodny, O., & Ghosh, R. (2013). A rapid

method for the extraction and analysis of carotenoids and other hydrophobic substances suitable for systems biology studies with photosynthetic bacteria. *Metabolites*, *3*(4), 912–30.

Braun-Kiewnick, A., Lehmann, A., Rezzonico, F., Wend, C., Smits, T. H. M., & Duffy, B. (2012). Development of species-, strain- and antibiotic biosynthesis-specific quantitative PCR assays for *Pantoea agglomerans* as tools for biocontrol monitoring. *Journal of Microbiological Methods*, *90*(3), 315–20.

Brooks, T. A., Kennedy, D. R., Gruol, D. J., Ojima, I., Baer, M. R., & Bernacki, R. J. (2004). Structure-activity Analysis of Taxane-based Broad-spectrum Multidrug Resistance Modulators. In *Anticancer Research* (Vol. 24, pp. 409–415).

Brooks, T. A., Minderman, H., O'Loughlin, K. L., Pera, P., Ojima, I., Baer, M. R., Bernacki, R. & Brooks, T. (2003). Taxane-based reversal agents modulate drug resistance mediated by P-glycoprotein, multidrug resistance protein, and breast cancer resistance protein. *Mol Cancer Ther*, *2*(11), 1195–205.

Brown, J. (2007). The iGEM competition: building with biology. *IET Synthetic Biology*, *1*(1–2), 3–6.

Chalupowicz, L., Barash, I., Panijel, M., Sessa, G., & Manulis-Sasson, S. (2009). Regulatory interactions between quorum-sensing, auxin, cytokinin, and the Hrp regulon in relation to gall formation and epiphytic fitness of P*antoea agglomerans* pv. *gypsophilae*. *Molecular Plant-Microbe Interactions : MPMI*, *22*(7), 849–56.

Chang, M. C. Y., Eachus, R. a, Trieu, W., Ro, D.-K., & Keasling, J. D. (2007). Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s. *Nature Chemical Biology*, *3*(5), 274–7.

Chatterjee, A., & Gibbins, L. (1971). Induction of nonpigmented variants of *Erwinia*

herbicola by incubation at supraoptimal temperatures. *Journal of Bacteriology*, *105*(I), 107–112.

Chemler, J. a, Yan, Y., & Koffas, M. a G. (2006). Biosynthesis of isoprenoids, polyunsaturated fatty acids and flavonoids in *Saccharomyces cerevisiae*. *Microbial Cell Factories*, *5*, 20.

Choi, H. S., Lee, S. Y., Kim, T. Y., & Woo, H. M. (2010). In silico identification of gene amplification targets for improvement of lycopene production. *Applied and Environmental Microbiology*, *76*(10), 3097–3105.

Choi, K. R., & Lee, S. Y. (2016). CRISPR technologies for bacterial systems: Current achievements and future directions. *Biotechnology Advances*.

Cortinovis, C., & Caloni, F. (2013). Epidemiology of intoxication of domestic animals by plants in Europe. *Veterinary Journal*. 197(2), 163-168.

Costa, E., Usall, J., Teixidó, N., Delgado, J., & Viñas, I. (2002). Water activity, temperature, and pH effects on growth of the biocontrol agent *Pantoea agglomerans* CPA-2. *Canadian Journal of Microbiology*, *48*(12), 1082–1088.

Croteau, R., Ketchum, R., & Long, R. (2006). Taxol biosynthesis and molecular genetics. *Phytochemistry Reviews*, *5*(1), 75–97. l

Das, L., Bhattacharya, B., & Basu, G. (2012). Rationalization of paclitaxel insensitivity of yeast β-tubulin and human βIII-tubulin isotype using principal component analysis. *BMC Research Notes*, *5*(1), 395.

Davis, B. M., Chao, M. C., & Waldor, M. K. (2013). Entering the era of bacterial epigenomics with single molecule real time DNA sequencing. *Current Opinion in Microbiology* 16(2), 192-198.

De Maayer, P., Chan, W.-Y., Blom, J., Venter, S. N., Duffy, B., Smits, T. H. M., &

Coutinho, T. A. (2012). The large universal *Pantoea* plasmid LPP-1 plays a major role in biological and ecological diversification. *BMC Genomics*, *13*(1), 625.

Dejong, J. M., Liu, Y., Bollon, A. P., Long, R. M., Jennewein, S., Williams, D., & Croteau, R. B. (2006). Genetic engineering of Taxol biosynthetic genes in *Saccharomyces cerevisiae. Biotechnology and Bioengineering*, *93*(2), 212–24.

Dussault, D., Caillet, S., Le Tien, C., & Lacroix, M. (2008). Carotenoids' influence on radiotolerance of *Pantoea agglomerans*, a plant pathogen. *Letters in Applied Microbiology*, *47*(3), 208–13.

Ehrlich, A., Booher, S., Becerra, Y., Borris, D. L., Figg, W. D., Turner, M. L., & Blauvelt, A. (2004). Micellar paclitaxel improves severe psoriasis in a prospective phase II pilot study. *Journal of the American Academy of Dermatology*, *50*(4), 533–40.

Emanuelsson, O., Brunak, S., von Heijne, G., & Nielsen, H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools. *Nature Protocols*, *2*(4), 953–971.

Engels, B., Dahm, P., & Jennewein, S. (2008). Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production. *Metabolic Engineering*, *10*(3–4), 201–6.

Expósito, O., Bonfill, M., Moyano, E., Onrubia, M., Mirjalili, M. H., Cusidó, R. M., & Palazón, J. (2009). Biotechnological production of Taxol and related taxoids: current state and prospects. *Anti-Cancer Agents in Medicinal Chemistry*, *9*(1), 109–21.

Farmer, W. R., & Liao, J. C. (2000). Improving lycopene production in *Escherichia coli* by engineering metabolic control. *Nature Biotechnology*, *18*(5), 533–537.

Ferlini, C., Cicchillitti, L., Raspaglio, G., Bartollino, S., Cimitan, S., Bertucci, C., Mozzetti, S., Gallo, D., Persico, M., Fattorusso, C., Campiani, G & Scambia, G.

(2009). Paclitaxel directly binds to Bcl-2 and functionally mimics activity of Nur77. *Cancer Research*, *69*(17), 6906–14.

Frank, S., Lawrence, A. D., Prentice, M. B., & Warren, M. J. (2013). Bacterial microcompartments moving into a synthetic biological world. *Journal of Biotechnology*, *163*(2), 273–279.

Furubayashi, M., Ikezumi, M., Kajiwara, J., Iwasaki, M., Fujii, A., Li, L., Saito, K. & Umeno, D. (2014). A high-throughput colorimetric screening assay for terpene synthase activity based on substrate consumption. *PloS One*, *9*(3), e93317.

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. a, Smith, H. O. & Hutchinson, C. A. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), 343–5.

Grobosch, T., Schwarze, B., Felgenhauer, N., Riesselmann, B., Roscher, S. & Binscheck, T. (2013). Eight cases of fatal and non-fatal poisoning with T*axus baccata. Forensic Science International*, *227*(1–3), 118–126.

Guzman, L. M., Belin, D., Carson, M. J., & Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose pBAD promoter. *Journal of Bacteriology*, *177*(14), 4121–30.

Hara, Y., Kadotani, N., Izui, H., Katashkina, J. I., Kuvaeva, T. M., Andreeva, I. G., Golubeva, L., Malko, D. B., Makeev, V. J., Mashko, S. V. & Kozlov, Y. I. (2012). The complete genome sequence of *Pantoea ananatis* AJ13355, an organism with great biotechnological potential. *Applied Microbiology and Biotechnology*, *93*(1), 331–41.

Heider, E., Harper, J., & Grant, D. (2007). Structural characterization of an anhydrous polymorph of paclitaxel by solid-state NMR. *Physical Chemistry Chemical Physics, 9*(46), 6083–6097.

Hengge, R. (2009). Principles of c-di-GMP signalling in bacteria. *Nature Reviews. Microbiology*, *7*(4), 263–273.

Hezari, M., Lewis, N. G., & Croteau, R. (1995). Purification and characterization of taxa-4(5),11(12)-diene synthase from Pacific yew (*Taxus brevifolia*) that catalyzes the first committed step of Taxol biosynthesis. *Archives of Biochemistry and Biophysics*. 322(2), 437-444

Holmes, M. L., Nuttall, S. D., & Dyall-Smith, M. L. (1991). Construction and use of halobacterial shuttle vectors and further studies on *Haloferax* DNA gyrase. *Journal of Bacteriology*, *173*(12), 3807–3813.

Holton, R. A., Somoza, C., Kim, H., Liang, F., Biediger, R. J., Boatman, P. D., Shindo, M., Smith, C. C., Kim, S., Nadizadeh, H., Suzuki, Y., Tao, C., Vu, P., Tang, S., Zhang, P., Murthi, K. K., Gentile, L. N.& Liu, J. H. (1994). First Total Synthesis of Taxol. 1. Functionalization of the B Ring. *American Chemical Society*, *116*(4), 1597–1598.

Horwitz, S. (1994). How to make Taxol from scratch. *Nature*, *367*, 593–594.

Huang, Q., Kirikae, F., Kirikae, T., Pepe, A., Amin, A., Respicio, L., Slayden, R. A., Tonge, P. J. & Ojima, I. (2006). Targeting FtsZ for antituberculosis drug discovery: Noncytotoxic taxanes as novel antituberculosis agents. *Journal of Medicinal Chemistry*, *49*(2), 463–466.

Huang, Q., Tonge, P., Slayden, R., Kirikae, T., & Ojima, I. (2007). FtsZ: A Novel Target for Tuberculosis Drug Discovery. *Current Topics in Medicinal Chemistry*, *7*(5), 527–543.

Immethun, C. M., Hoynes-O'Connor, A. G., Balassy, A., & Moon, T. S. (2013). Microbial production of isoprenoids enabled by synthetic biology. *Frontiers in Microbiology*, *4*(April), 75.

Jennewein, S., Long, R. M., Williams, R. M., & Croteau, R. (2004). Cythochrome p450

Taxadiene 5A-Hydroxylase, a Mechanistically Unusual Monooxygenase Catalyzing the First Oxygenation Step of Taxol Biosynthesis, *11*, 379–387.

Jennifer Guerra-Bubb, Rodney Croteau, & R. M. W. (2013). The Early Stages of Taxol Biosynthesis: An Interim Report on the Synthesis and Identification of Early Pathway Metabolites, *18*(9), 1199–1216.

Jin, Y. S., & Stephanopoulos, G. (2007). Multi-dimensional gene target search for improving lycopene biosynthesis in *Escherichia coli*. *Metabolic Engineering*, *9*(4), 337–347.

Kaiser, T. D. L., Pereira, E. M., dos Santos, K. R. N., Maciel, E. L. N., Schuenck, R. P., & Nunes, A. P. F. (2013). Modification of the Congo red agar method to detect biofilm production by *Staphylococcus epidermidis*. *Diagnostic Microbiology and Infectious Disease*, *75*(3), 235–239.

Kajiwara, S., Fraser, P. D., Kondo, K., & Misawa, N. (1997). Expression of an exogenous isopentenyl diphosphate isomerase gene enhances isoprenoid biosynthesis in *Escherichia coli*. *Biochem. J.*, *324*(2), 421–426.

Kamath, K., Braguer, D., Horwitz, S. B., Wilson, L., Briand, C., & Jordan, M. A. (2004). Synergistic Suppression of Microtubule Dynamics by Discodermolide and Paclitaxel in Non-Small Cell Lung Carcinoma Cells, (805), 4957–4964.

Kaspera, R., Cape, J., & Faraldos, J. (2010). Synthesis and in vitro evaluation of Taxol oxetane ring D precursors. *Tetrahedron Letters*, *51*(15), 2017–2019.

Khachik, F., & Chang, A. N. (2009). Total synthesis of (3R,3'R,6'R)-lutein and its stereoisomers. *Journal of Organic Chemistry*, *74*(10), 3875–3885.

Kim, I.-Y., Pusey, P. L., Zhao, Y., Korban, S. S., Choi, H., & Kim, K. K. (2012). Controlled release of *Pantoea agglomerans* E325 for biocontrol of fire blight disease of apple. *Journal of Controlled Release : Official Journal of the Controlled Release*

Society, *161*(1), 109–15.

Knight, T. (2007). Draft Standard for Biobrick Biological Parts. *BioBricks Foundation Request for Comments*.

Köksal, M., Jin, Y., Coates, R. M., Croteau, R., & Christianson, D. W. (2011). Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. *Nature*, *469*(7328), 116–20.

Kovacs, K., Zhang, L., Linforth, R. S. T., Whittaker, B., Hayes, C. J., & Fray, R. G. (2007). Redirection of carotenoid metabolism for the efficient production of taxadiene [taxa-4(5),11(12)-diene] in transgenic tomato fruit. *Transgenic Research*, *16*(1), 121–6.

Kumaran, R. S., Kim, H. J., & Hur, B.-K. (2010). Taxol-producing [corrected] fungal endophyte, Pestalotiopsis species isolated from *Taxus cuspidata*. *Journal of Bioscience and Bioengineering*, *110*(5), 541–6.

Larson, M. H., Gilbert, L. a, Wang, X., Lim, W. a, Weissman, J. S., & Qi, L. S. (2013). CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature Protocols*, *8*(11), 2180–96.

Lee, H. S., Lee, N. C. O., Kouprina, N., Kim, J. H., Kagansky, A., Bates, S., … Larionov, V. (2016). Effects of anticancer drugs on chromosome instability and new clinical implications for tumor-suppressing therapies. *Cancer Research*, *76*(4), 902–911.

Lenihan, J. R., Tsuruta, H., Diola, D., Renninger, N. S., & Regentin, R. (2008). Developing an industrial artemisinic acid fermentation process to support the cost-effective production of antimalarial artemisinin-based combination therapies. *Biotechnology Progress*, *24*(5), 1026–1032.

Li, M., Jiang, F., Yu, X., & Miao, Z. (2015). Engineering isoprenoid biosynthesis in *Artemisia annua* L. for the production of taxadiene: A key intermediate of Taxol.

BioMed Research International, *2015*.

Li, S., Zhang, P., Zhang, M., Fu, C., Zhao, C., Dong, Y., Guo, A. & Yu, L. (2012). Transcriptional profile of *Taxus chinensis* cells in response to methyl jasmonate. *BMC Genomics*, *13*(1), 295.

Li, X. T., Thomason, L. C., Sawitzke, J. A., Costantino, N., & Court, D. L. (2013). Positive and negative selection using the *tetA-sacB* cassette: Recombineering and P1 transduction in *Escherichia coli*. *Nucleic Acids Research*, *41*(22), 1–8.

Ling, L. L., Schneider, T., Peoples, A. J., Spoering, A. L., Engels, I., Conlon, B. P., Mueller, A., Schäberle, T. F., Hughes, D. E., Epstein, S., Jones, M., Lazarides, L. Steadman, V. A., Cohen, D. R., Felix, C. R., Fetterman, K. A., Millett, W. P., Nitti, A. G., Zullo, A. M., Chen, C. & Lewis, K. (2015). A new antibiotic kills pathogens without detectable resistance. *Nature*. 517

Loh, J. P., & Waksman, R. (2012). Paclitaxel drug-coated balloons: a review of current status and emerging applications in native coronary artery de novo lesions. *JACC. Cardiovascular Interventions*, *5*(10), 1001–12.

Macomber, L., & Hausinger, R. P. (2011). Mechanisms of nickel toxicity in microorganisms. *Metallomics : Integrated Biometal Science*, *3*(11), 1153–62.

Martin, V. J. J., Pitera, D. J., Withers, S. T., Newman, J. D., & Keasling, J. D. (2003). Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nature Biotechnology*, *21*(7), 796–802.

McKenna, R., & Nielsen, D. R. (2011). Styrene biosynthesis from glucose by engineered E. coli. *Metabolic Engineering*, *13*(5), 544–554.

McKenzie, G. J., & Craig, N. L. (2006). Fast, easy and efficient: site-specific insertion of transgenes into enterobacterial chromosomes using Tn*7* without need for selection of the insertion event. *BMC Microbiology*, *6*, 39.

Meganathan, R. (2001). Ubiquinone biosynthesis in microorganisms. *FEMS Microbiology Letters*, 203(2), 131-139.

Mendoza, A., Ishihara, Y., & Baran, P. S. (2012). Scalable enantioselective total synthesis of taxanes. *Nature Chemistry*, *4*(1), 21–5.

Milanov, D., Prunic, B., Velhner, M., Pajic, M., & Cabarkapa, I. (2015). RDAR morphotype: A resting stage of some *Enteroba cteriaceae*. *Food and Feed Research*, *42*(1), 43–50.

Minderman, H., Brooks, T. A., O'Loughlin, K. L., Ojima, I., Bernacki, R. J., & Baer, M. R. (2004). Broad-spectrum modulation of ATP-binding cassette transport proteins by the taxane derivatives ortataxel (IDN-5109, BAY 59-8862) and tRA96023. *Cancer Chemotherapy and Pharmacology*, *53*(5), 363–369.

Mirjalili, M. H., Farzaneh, M., Bonfill, M., Rezadoost, H., & Ghassempour, A. (2012). Isolation and characterization of *Stemphylium sedicola* SBU-16 as a new endophytic Taxol-producing fungus from *Taxus baccata* grown in Iran. *FEMS Microbiology Letters*, *328*(2), 122–9.

Misawa, N., Nakagawa, M., Kobayashi, K., Yamano, S., Izawa, Y., Nakamura, K., & Harashima, K. (1990). Elucidation of the *Erwinia uredovora* carotenoid biosynthetic pathway by functional analysis of gene products expressed in *Escherichia coli*. *Journal of Bacteriology*, *172*(12), 6704–12.

Mohammadi, M., Burbank, L., & Roper, M. C. (2012). Biological role of pigment production for the bacterial phytopathogen *Pantoea stewartii* subsp. *stewartii*. *Applied and Environmental Microbiology*, *78*(19), 6859–65.

Monk, I. R., Shah, I. M., Xu, M., Tan, M. W., & Foster, T. J. (2012). Transforming the untransformable: Application of direct transformation to manipulate genetically *Staphylococcus aureus* and *Staphylococcus epidermidis*. *mBio*, *3*(2).

Morgan, J. L. W., McNamara, J. T., & Zimmer, J. (2014). Mechanism of activation of bacterial cellulose synthase by cyclic di-GMP. *Nature Structural & Molecular Biology*, *21*(5), 489–96.

Mühlradt, P. F., Sasse, F., Activity, T. E., & Mãhlradt, P. F. (1997). Epothilone B Stabilizes Microtubuli of Macrophages Like Taxol without Showing Taxol-like Endotoxin Activity Advances in Brief Epothilone B Stabilizes Microtubuli of Macrophages Like Taxol without Showing, 3344–3346.

Nadarasah, G., & Stavrinides, J. (2014). Quantitative evaluation of the host-colonizing capabilities of the enteric bacterium *Pantoea* using plant and insect hosts. *Microbiology (United Kingdom)*, *160*(PART 3), 602–615.

Nakata, K., Inagawa, H., & Soma, G.-I. (2011). Lipopolysaccharide IP-PA1 from *Pantoea agglomerans* prevents suppression of macrophage function in stress-induced diseases. *Anticancer Research*, *31*(7), 2437–40.

Nisar, M., Khan, I., Ahmad, B., Ali, I., Ahmad, W., & Choudhary, M. I. (2008). Antifungal and antibacterial activities of *Taxus wallichiana* Zucc. *Journal of Enzyme Inhibition and Medicinal Chemistry*, *23*(2), 256–260.

Nishizaki, T., Tsuge, K., Itaya, M., Doi, N., & Yanagawa, H. (2007). Metabolic engineering of carotenoid biosynthesis in *Escherichia coli* by ordered gene assembly in *Bacillus subtilis*. *Applied and Environmental Microbiology*, *73*(4), 1355–1361.

Nogales, E., Wolf, S., & Downing, K. (1998). Structure of the αβ tubulin dimer by electron crystallography. *Nature*, *391*(January), 199–204.

Nurk, S., Bankevich, A., Antipov, D., Gurevich, A.,Korobeynikov, A., Lapidus, A., Prjibelsky, A., Pyshkin, A., Sirotkin, A., Sirotkin, Y., Stepanauskas, R., Mclean, J.,Lasken, R., Clingenpeel, S. R., Woyke, T., Tesler, G.,Alekseyev, M. A. & Pevzner,

P. A. (2013). Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads. *Research in Computation Molecular Biology,* 7821**,** 158-170.

Ojima, I., Borella, C. P., Wu, X., Bounaud, P.-Y., Oderda, C. F., Sturm, M., Miller, M. L., Chakravarty, S., Chen, J. Huang, Q. Pera, P., Brooks, T. Baer, M. R. & Bernacki, R. J. (2005). Design, synthesis and structure-activity relationships of novel taxane-based multidrug resistance reversal agents. *Journal of Medicinal Chemistry*, *48*(6), 2218–28.

Ojima, I., Kumar, K., Awasthi, D., & Vineberg, J. G. (2014). Drug discovery targeting cell division proteins, microtubules and FtsZ. *Bioorganic & Medicinal Chemistry*, *22*(18), 5060–77.

Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., … Newman, J. D. (2013). High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature*, *496*(7446), 528–32.

Pamp, S. J., Gjermansen, M., & Tolker-Nielsen, T. (2007). The Biofilm Matrix: A Sticky Framework. The Biofilm Mode of Life: *Mechanisms and Adaptations*.

Phyton Biotech. (2013). http://www.phytonbiotech.com/about_history.htm.

Rath S, Olempska-Beer Z, K. P. (2007). Lycopene extract from tomato is a lycopene-rich extract prepared from the ripe fruits of tomato (CTA). *Fao*, (9), 1–9.

Reiling, K. K., Yoshikuni, Y., Martin, V. J. J., Newman, J., Bohlmann, J., & Keasling, J. D. (2004). Mono and diterpene production in *Escherichia coli*. *Biotechnology and Bioengineering*, *87*(2), 200–212.

Reyrat, J., Pelicic, V., & Gicquel, B. (1998). MINIREVIEW Counterselectable Markers : Untapped Tools for Bacterial Genetics and Pathogenesis, *66*(9), 4011–4017.

Rezzonico, F., Smits, T. H. M., Montesinos, E., Frey, J. E., & Duffy, B. (2009). Genotypic

comparison of *Pantoea agglomerans* plant and clinical strains. *BMC Microbiology*, *9*, 204.

Rod J. Dillon, Chris T. Vennard, A. K. C. (1999). Exploitation of gut bacteria in the locust. *Nature*, *216*(1), 1. T

Römling, U. (2005). Characterization of the rdar morphotype, a multicellular behaviour in Enterobacteriaceae. *Cellular and Molecular Life Sciences*, 62(11), 1234-1246.

Sambrook, J., & Russell, D. W. (2001). Molecular Cloning - Sambrook & Russel - Vol. 1, 2, 3.pdf. *Human Mutation*.

Sauer, K. (2003). The genomics and proteomics of biofilm formation. *Genome Biology*, *4*(6), 219.

Schoendorf, A., Rithner, C. D., Williams, R. M., & Croteau, R. B. (2001). Molecular cloning of a cytochrome P450 taxane 10 beta-hydroxylase cDNA from *Taxus* and functional expression in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(4), 1501–6.

Schmieder, R. and Edwards, R. (2011). Quality Control and Preprocessing of Metagenomic Datasets. *Bioinformatics* 27 (6): 863–64

Sedkova, N., & Tao, L. (2005). Diversity of carotenoid synthesis gene clusters from environmental Enterobacteriaceae strains. *Applied and Environmental Microbiology*, *71*(12), 8141–8146.

Shen, H.-J., Hu, J.-J., Li, X.-R., & Liu, J.-Z. (2015). Engineering of *Escherichia coli* for lycopene production through promoter engineering. *Current Pharmaceutical Biotechnology*, *16*, 1094–1103.

Shimada, H., Kondo, K., Fraser, P. D., Miura, Y., Saito, T., & Misawa, N. (1998).

Increased carotenoid production by the food yeast *Candida utilis* through metabolic engineering of the isoprenoid pathway. *Applied and Environmental Microbiology*, *64*(7), 2676–2680.

Skrzypek, E., Haddix, P. L., Plano, G. V., & Straley, S. C. (1993). New Suicide Vector for Gene Replacement in *Yersiniae* and Other Gram-Negative Bacteria. Plasmid (Vol. 29).

Sleight, S. C., & Sauro, H. M. (2013). Visualization of Evolutionary Stability Dynamics and Competitive Fitness of *Escherichia coli* Engineered with Randomized Multigene Circuits, 20(9), 519-528.

Soumagne, N., Chauvet, S., Chatellier, D., Robert, R., Charrière, J. M., & Menu, P. (2011). Treatment of yew leaf intoxication with extracorporeal circulation. *American Journal of Emergency Medicine*, 29(3), 354

Spiers, A. J., Kahn, S. G., Bohannon, J., Travisano, M., & Rainey, P. B. (2002). Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. I. Genetic and phenotypic bases of wrinkly spreader fitness. *Genetics*, *161*(1), 33–46.

Stankovic, I. (2004). Zeaxanthin - Manufacturing. *Assessment*, *1*(7), 1–7.

Taylor, R. G., Walker, D. C., & Mclnnes, R. R. (1993). *E.coli* host strains, *21*(7), 1677–1678.

Tham, K. C., Kanaar, R., & Lebbink, J. H. G. (2016). Mismatch repair and homeologous recombination. *DNA Repair*, 38, 75-83.

Tsuruta, H., Paddon, C. J., Eng, D., Lenihan, J. R., Horning, T., Anthony, L. C., … Newman, J. D. (2009). High-level production of amorpha-4, 11-diene, a precursor of the antimalarial agent artemisinin, in *Escherichia coli*. *PLoS ONE*, *4*(2).

Walker, K., Schoendorf, A. & Croteau, R. (2000). Molecular cloning of a taxa-4(20),11(12)-dien-5alpha-ol-O-acetyl transferase cDNA from *Taxus* and functional expression in *Escherichia coli*. *Archives of Biochemistry and Biophysics*, *374*(2), 371–80.

Walterson, A. M., Smith, D. D. N., & Stavrinides, J. (2014). Identification of a *Pantoea* biosynthetic cluster that directs the synthesis of an antimicrobial natural product. *PloS One*, *9*(5), e96208.

Walterson, A. M. & Stavrinides, J. (2015). *Pantoea:* insights into a highly versatile and diverse genus within the Enterobacteriaceae. *FEMS Microbiology Reviews*, (August 2014), fuv027.

Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., Xu, G., Forest, C. R., & Church, G. M. (2009). Programming cells by multiplex genome engineering and accelerated evolution. *Nature*, *460*(7257), 894–8.

Wang, H., Liu, B., Zhang, C., Peng, G., Liu, M., Li, D., … Zhou, J. (2009). Parkin regulates paclitaxel sensitivity in breast cancer via a microtubule-dependent mechanism. *Journal of Pathology*, *218*(1), 76–85.

Wang, S., & Ghosh, A. (2012). Fighting malaria with engineered symbiotic bacteria from vector mosquitoes. *Proceedings of the …*, *109*(31), 12734–12739.

Wani, M., Taylor, H., & Wall, M. (1971). Plant antitumor agents. VI. Isolation and structure of Taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *Journal of the American Chemical Society*, *93*(9), 2325–2327.

Wattam, A. R., Abraham, D., Dalay, O., Disz, T. L., Driscoll, T., Gabbard, J. L., … Sobral, B. W. (2014). PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Research*, *42*(D1).

Wildung, M. R., & Croteau, R. (1996). A cDNA clone for taxadiene synthase, the

diterpene cyclase that catalyzes the committed step of Taxol biosynthesis. *Journal of Biological Chemistry*, *271*(16), 9201–9204.

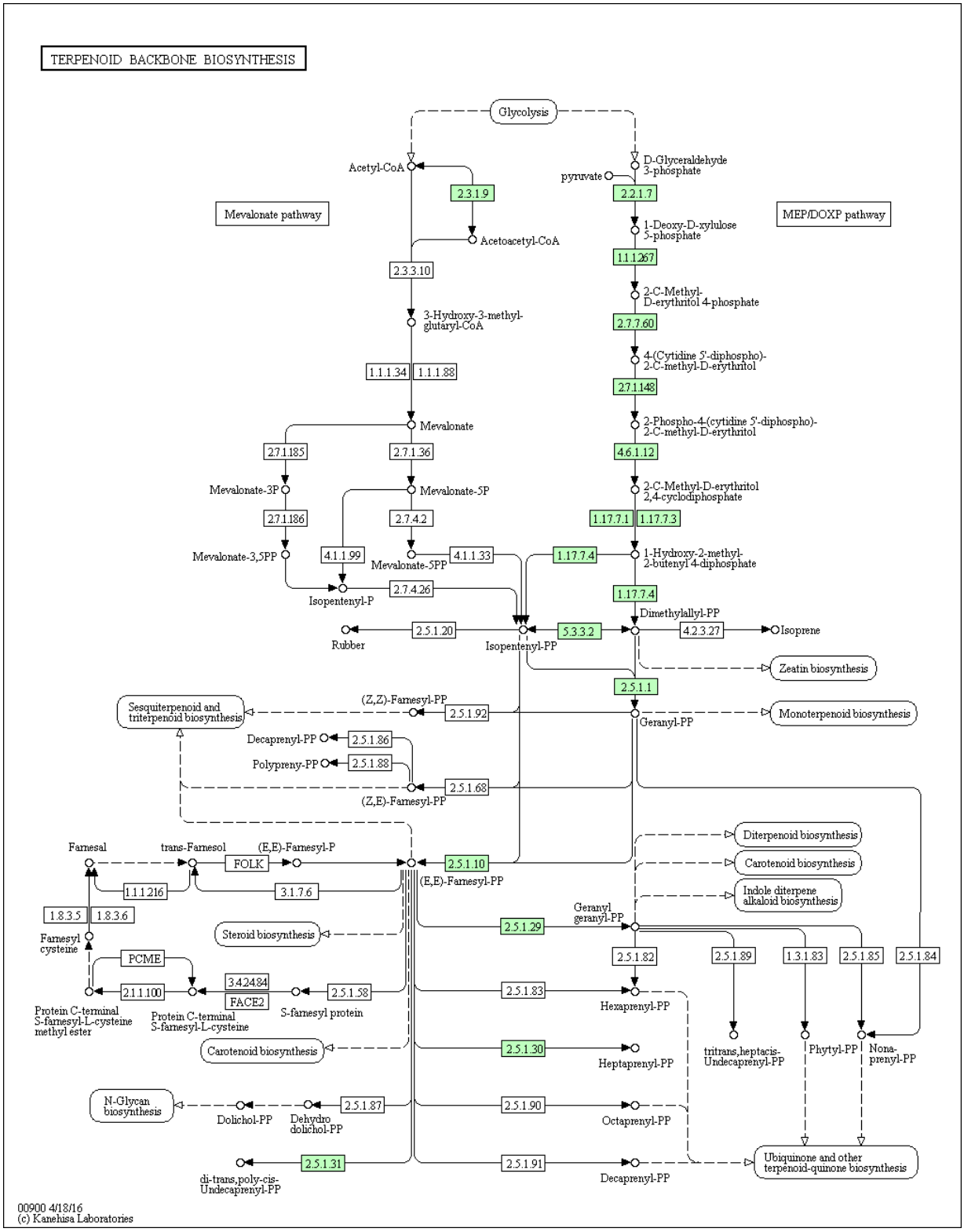
Wilson, S., & Roberts, S. (2012). Recent advances towards development and commercialization of plant cell culture processes for the synthesis of biomolecules. *Plant Biotechnology Journal*, *10*(3), 249–268.

Yoon, S. H., Kim, J. E., Lee, S. H., Park, H. M., Choi, M. S., Kim, J. Y., … Kim, S. W. (2007). Engineering the lycopene synthetic pathway in *E. coli* by comparison of the carotenoid genes of *Pantoea agglomerans* and *Pantoea ananatis*. *Applied Microbiology and Biotechnology*, *74*(1), 131–139.

Yue, Q., Fang, Q. C., & Liang, X. T. (1996). Semisynthesis of Taxol. *Yaoxue Xuebao*, *31*(12), 911–917.

Zhang, B., Maiti, A., Shively, S., Lakhani, F., McDonald-Jones, G., Bruce, J., … Trojanowski, J. Q. (2005). Microtubule-binding drugs offset tau sequestration by stabilizing microtubules and reversing fast axonal transport deficits in a tauopathy model. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(1), 227–31.

APPENDIX



http://www.genome.jp/kegg-bin/show_pathway?bsu00900