Engineering Rhizobacteria as Synthetic Biology Chassis



Kyle C Grant
Wadham College
University of Oxford

A Thesis submitted in the Field of Synthetic Biology
for the Degree of Doctor of Philosophy

University of Oxford
Trinity 2018

For Lucinda, Colin, Iris May, Graeme, Gogo, Mum and Dad



Engineering Rhizobacteria as Synthetic Biology Chassis

Kyle C Grant, Wadham College, University of Oxford, Trinity 2018

A thesis submitted for the degree of Doctor of Philosophy

Abstract

In the next half-century the world will face an enormous food deficit due to a raising birth rate and longer life-spans of the global population. By 2050 there will be 9.5 billion people in the world who will require a 69% increase from current total food calories to prevent global starvation. Widespread hunger and starvation are linked directly to the productivity of crop growth and productivity in agriculture, especially in the third world. By 2030, Areas such as sub-Saharan Africa and Australia are predicted to have a 37% and 48% decrease in crop productivity respectively due to global warming. The first 'Green Revolution' increased the world's crop output by an order of magnitude using synthetic nitrogen fertilisers derived from the Haber-Bosch process. A process which last year yielded 180 million tonnes of ammonia, 82% of which was used for the manufacture of nitrogenous fertilisers. Global usage of synthetic nitrogenous fertilisers is increasingly unsustainable and a new approach to increasing crop productivity is required to face the ever-increasing food demands over the next

30 years. In this project we have contributed to the engineering of biofertilisers and plant growth promoting rhizobacteria that could alleviate the use of inorganic fertilisers and chemicals and serve as an economically viable alternative in developing countries. Synthetic biology tools and techniques are employed to engineer two model rhizobacteria. *Azorhizobium caulinodans* ORS571 and *Azospirillum brasilense* FP2, two nitrogen-fixing and known plant-associative rhizobacteria, are used as chassis. A suite of synthetic biology tools are characterised for the two rhizobacteria species as exemplars of the toolkit being used for the engineering of other significant rhizobacteria. Advanced logic-gate based genetic circuitry is developed to yield functional genetic circuits that can be engineered with characterised plant growth promotion (PGP) effectors to yield agriculturally-relevant phenotypes. A suite of these PGP mechanisms including auxin biosynthesis, phosphate solubilisation and antifungal compound (DAPG) biosynthesis, are built and characterised in both rhizobacterial strains

The novel implementation of a CRISPR interference technology (CRISPRi) is demonstrated in both chassis. Genomic targeting of dCas9 is implemented to control the ammonia assimilation pathway of *A. brasilense* FP2 with the effect of significant ammonia release into growth media. This work builds the foundations for the future engineering of both logical genetic circuitry and the CRISPRi system with an input from a unique associative crop plant signal. This will allow a rhizosphere-specific phenotypic switch of the rhizobacteria from a fast-growing competitive state to a plant beneficial ammonia secreting state.

Acknowledgments

"Gradatim Ferociter"

- *Blue Origin*

Science is hard. Synthetic Biology is harder. This thesis has been a pleasure to undertake with incredibly hard failures and elating successes in abundance. Many others and I involved with the work in this field have been making incremental yet solid steps towards our vision of a sustainable and fertiliser-free agricultural industry. It is without a doubt that the advent of such an age will bring with it not only food security for billions but also a stability in the World's ecosystems that has not been seen since the advent of the Green Revolution.

My supervisor Philip Poole has been an immense inspiration over the past four years. He has worked tirelessly to understand how the powerful forces of nature have yielded one of the most exquisite examples of symbiosis seen in the natural world. It is only with his wealth of knowledge, and passion for sharing that knowledge, that this work has been possible. Thank you, Phil, for being so supportive throughout this work and also for keeping me on track when times got tough, I'm incredibly grateful.

Our team leader, and behind the scenes heroine, Alison East has been steadfast in her support of my work and has given fantastic advice on experimental designs as well as rhizobacterial knowledge I could never have found elsewhere. Thank you, Alison, for putting up with my errant ways and for your support throughout this project.

This project would not have been possible without the support and direction of Dr Barney Geddes. Without your enthusiasm for this subject I may well never have studied rhizobacteria and would have missed out on something I have found incredibly inspiring. In many ways Barney has been the rock of much of our work in the lab and his attention to detail and scientific method are second to none. I wish him every success in the new chapter of his life in Canada.

Finally, and most importantly, I wish to acknowledge and thank my partner Lucinda O'Kane for her unwavering support of my DPhil over four long years. Despite ups and downs, I would never have been able to even contemplate studying for a DPhil without your love, support and patience. I owe you more than you know.

Abbreviations

Abbreviation	Full Description
2, 4-D	2,4-dichlorophenoxyacetic acid
3O6-HSL	3O6 homoserine lactone
ACC	1-aminocyclopropane-1-carboxylic acid
ADP	Adenosine diphosphate
ARA	Acetylene reduction assay
aTc	Anhydrous tetracycline
ATP	Adenosine triphosphate
CAD	Computer aided design
Cas	CRISPR associate proteins
CFU	Colony-forming units
CRISPR	Clustered regularly interspaced palindromic repeats
CRISPRi	CRISPR Interference Technology
crRNA	CRISPR non-coding RNA
DAPG	2,4-diaceylphloroglucanol
dCas9	Dead cas9
dCas9	Dead Cas9
DNA	Deoxyribonucleic acid

DPI	Days post-inoculation
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
EZ	Elongation zone
FRET	Förster resonance energy transfer
gDNA	Genomic DNA
gDNA	Genomic DNA
GOGAT	Glutamine oxoglutarate aminotransferase
GS	Glutamine Synthase
IAA	Indole-3-acetic acid
iGEM	International competition for genetically engineered machines
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IRLC	Inverted repeat lacking clade
LB	Luria-burtani
Lux	Luciferase (Luminescence)
MAPG	Monoaceylphloroglucanol
MOPS	3-(N-morpholino)propanesulfonic acid
MSX	L-methionine sulfoximine
MZ	Maturation zone
ncRNA	Non-coding RNA

OD	Optical density
ORF	Open reading frame
PAM	Protospacer adjacent motif
PGPR	Plant growth promoting rhizobacteria
Pi	Inorganic phosphate
PQQ	Pyrroloquinoline
PSI	Phosphate solubilisation index
REU	Relative expression units
SAM	S-adenosy-L-methionine
SDM	Site-directed mutagenesis
SEM	Standard error of the mean
sfGFP	Super folder green fluorescent protein
sgRNA	Single guide RNA
SNP	Single nucleotide polymorphism
ssDNA	Single-stranded DNA
ssrA	small specialised stable RNA
T7 RNAP	T7 RNA Polymerase
tracRNA	crRNA-transacting RNA
UMS	Universal minimal salts
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Table of Contents

ENGINEERING RHIZOBACTERIA AS SYNTHETIC BIOLOGY CHASSIS.....

ABSTRACT.....

ACKNOWLEDGMENTS.....

ABBREVIATIONS.....

LIST OF TABLES.....

LIST OF FIGURES.....

1 INTRODUCTION.....

1.1 The Importance of Nitrogen in Agricultural Systems 1

1.2 Natural Diazotrophic Bacterial Interactions with Crop Plants..... 2

1.3 Engineering Non-leguminous Diazotrophic Bacteria-Plant Interactions 4

1.3.1 Global Agriculture Overview..... 4

1.3.2 Rhizobacterial Plant Colonisation..... 8

1.3.3 Azorhizobium caulinodans ORS571..... 8

1.3.4 Azospirillum brasilense FP2 11

1.4 Synthetic Biology..... 14

1.4.1 Introduction to Synthetic Biology..... 14

1.4.2 Computational Logic in Bacteria..... 15

1.4.3 CRISPR Interference Technology (CRISPRi)..... 17

1.5 Genetic Engineering of Diazotrophic Rhizobacteria 20

1.5.1 Diazotrophic Rhizobacteria 20

1.5.2 Ammonia Secretion and Plant Growth Promotion..... 22

1.5.3 Plant Growth Promotion Phenotypes..... 25

1.5.4 Synthetic Phenotypic Switches..... 28

1.6 SPECIFIC PLANT SIGNALLING MOLECULE 30
1.6.1 Plant – Rhizosphere Chemical Signals 30
1.6.1.1 Rhizopine: A unique and specific plant signalling molecule 31
1.7 AIMS AND OBJECTIVES 32
2 GENERAL MATERIALS AND METHODS 33
2.1 STRAINS USED IN THIS STUDY 33
2.2 PLASMIDS USED IN THIS STUDY 34
2.3 PRIMERS USE IN THIS STUDY 43
2.4 MEDIA AND ANTIBIOTICS AND OTHER CHEMICALS 49
2.4.1 Media 49
2.4.2 Antibiotics and other Chemicals 50
2.4.3 Aerobic Growth 51
2.4.4 Microaerobic Diazotrophic Growth 52
2.5 MOLECULAR TECHNIQUES 54
2.5.1 DNA Isolation 54
2.5.2 DNA Amplification by PCR 55
2.5.3 DNA Gel Electrophoresis 56
2.5.4 DNA Fragmentation by Restriction Enzyme Digests 56
2.5.5 DNA Ligation 57
2.5.6 Recombination-based Cloning HD (BD) Cloning 57
2.5.7 Type IIS Restriction Enzyme Cloning 58
2.5.8 DNA Sequence Verification 61
2.5.9 Oligonucleotide Annealing 61
2.5.10 Overlap PCR and Site Directed Mutagenesis (SDM) 62
2.5.11 DNA Consensus Sequence Analysis 62
2.6 TRANSFORMATION TECHNIQUES 64
2.6.1 Thermal 64
2.6.2 Electroporation 65
2.6.3 Conjugation 66

2.7 ASSAYS 68
2.7.1 Growth of Rhizobacteria 68
2.7.2 Fluorescent Protein Excitation and Emission Determination 69
2.7.3 Fluorescent Reporter Gene Expression Analysis 69
2.7.4 Single Cell Gene Expression Analysis – Flow Cytometry 71
2.7.5 Nitrogen Fixation Assay – Acetylene Reduction Assay 72
2.7.6 Cellular Weight and Protein Content Determination 75
2.7.7 Lux Activity in Rhizobacteria 77
2.8 STATISTICAL ANALYSIS 78

3 RHIZOKIT: A SYNTHETIC BIOLOGY TOOLBOX FOR RHIZOBACTERIA 80

3.1 ABSTRACT 80
3.2 INTRODUCTION 82
3.3 RESULTS 88
3.3.1 Constitutive Promoter Design and Characterisation 88
3.3.2 Ribosome Binding Site Sequence and Analysis 94
3.3.3 Terminator design and characterisation 98
3.3.4 Chemically Inducible Promoters 103
3.3.5 Tunable Protein Degradation Analysis 110
3.3.6 Orthogonal Transcription System (T7 RNAP) 115
3.3.7 Fluorescent Reporter Library 119
3.4 DISCUSSION 127

4 GENETIC LOGIC GATE CIRCUITRY IN RHIZOBACTERIA 133

4.1 ABSTRACT 133
4.2 INTRODUCTION 134
4.2.1 NOT Gates 137
4.3 RESULTS 140
4.3.1 Orthogonal Transcriptional Repressor and Promoter Selection 140
4.3.1.1 Transcriptional regulator and promoter dynamics 146

4.3.1.2 Fitness cost analysis 152
4.3.2 Transcriptional repressor orthogonality 154
4.3.3 Inducible NOT gate in Rhizobacteria 156
4.3.3.1 NOT Gate Functional dynamics 162
4.3.3.2 Fitness analysis 164
4.4 DISCUSSION 166

5 ENGINEERING OF PLANT GROWTH PROMOTION PHENOTYPES 173

5.1 ABSTRACT 173
5.2 INTRODUCTION 174
5.2.1 Nitrogen Fixation 178
5.2.2 Phosphate solubilisation 180
5.2.3 Indole-3-Acetic Acid (Auxin) 182
5.2.4 Antifungal Protection 183
5.3 RESULTS 186
5.3.1 Engineering IAA Biosynthesis 186
5.3.1.1 Constitutive auxin biosynthesis 187
5.3.1.2 Inducible IAA Biosynthesis 195
5.4 COMPETITIVE *IN VIVO* PLANT GROWTH PROMOTION 202
5.4.1 Root hair length promotion 202
5.4.2 Root length promotion 208
5.4.3 Root diameter analysis 215
5.5 ENGINEERING DAPG BIOSYNTHESIS 218
5.5.1 Constitutive DAPG Biosynthesis 221
5.6 PHOSPHATE SOLUBILISATION 226
5.6.1 Constitutive Citric Acid Biosynthesis 228
5.7 DISCUSSION 234
5.7.1 IAA Biosynthesis 234
5.7.2 DAPG Biosynthesis 240
5.7.3 Phosphate Solubilisation 242

6 CRISPRI OF RHIZOBACTERIA 246

6.1 ABSTRACT 246

6.2 INTRODUCTION 247

6.3 RESULTS 254

6.3.1 dCas9 Toxicity Screen in Rhizobacteria 254

6.3.2 dCas9 Activity in vivo 262

6.3.3 Golden Gate CRISPRi Strategy 268

6.3.4 CRISPRi of endogenous rhizobacterial gene 273

6.3.4.1 Glutamine Synthetase Chemical Inhibition 275

6.3.4.2 GlnA promoter mapping and analysis 277

6.3.4.3 GlnA sgRNA design 284

6.3.4.4 *In vivo* CRISPRi of GlnA 287

6.4 DISCUSSION 297

6.4.1 dCas9 Function in rhizobacteria 297

6.4.2 CRISPRi of glnA in A. brasilense FP2 300

7 GENERAL CONCLUSIONS AND FUTURE DIRECTIONS 304

7.1 OVERVIEW 304

7.2 SUMMARY OF RESULTS AND FUTURE WORK 304

7.2.1 RhizoKit: Rhizobacterial DNA Toolkit 304

7.2.2 Computational logic in rhizobacteria 307

7.2.3 Plant Growth Promotion 308

7.2.4 CRISPRi in rhizobacteria 311

7.3 FINAL CONCLUSION 313

8 BIBLIOGRAPHY 316

9 APPENDICES 345

List of Tables

Table 2-1 – Bacterial strains used in this thesis 33

Table 2-2 – Plasmids used in this thesis 34

Table 2-3 – Primers used in this thesis 43

Table 2-4 - Working concentrations of antibiotics used in this work. Units in μg ml^{-1} 51

Table 2-5 -Excitation and emission filters used for the determination of fluorescent protein expression in *A. caulinodans* ORS571 and *A. brasilense* FP2. 71

Table 3-1 – Table of primers used to amplify and domesticate the IPTG inducible rPU promoter for rhizobacterial synthetic biology. 104

Table 3-2 - Table of primers used to amplify and domesticate the aTc inducible rPU promoter for rhizobacterial synthetic biology. 105

Table 3-3 - Table of primers used to amplify and domesticate the DAPG inducible rPU promoter for rhizobacterial synthetic biology. 105

Table 3-4 – ssrA degradation tag sequences utilised to tune protein degradation within rhizobacteria. 111

Table 3-5 - Fluorescent Protein Reporter Library 120

Table 4-1 – Truth tables, symbols and matrices for the common logic gates present in electronics and synthetic biology. Note that **A** & **B** are inputs and **X** is the output. 135

Table 5-1 - Table of primers used to amplify and domesticate the *phl*ABC operon for rhizobacterial synthetic biology. 220

Table 6-1 – S. pyogenes dCas9 variants, unique amino acid substitutions and PAM sequences. 251

List of Figures

Figure 1-1 - Nitrogen-fixing nodules formed on the root of the legume *Medicago trucatula*[10].. 3

Figure 1-2 – Top 5 global agricultural crops in 2017..... 5

Figure 1-3 - Endophytic colonisation of O. sativa by Rhizobium IRBG74 via both lateral root junctions and root hairs..... 7

Figure 1-4 – Infection of wheat root tissue with the rhizobacteria *Azorhizobium caulinodans* ORS571 with different application methods.. 10

Figure 1-5 - Colonisation of the lateral root emergence of wheat (*Triticum aestivum* var. CD104) by *Azospirillum brasilense* FP2:..... 12

Figure 1-6 – The CRISPRi system derived from the naturally occurring CRISPR-Cas9 gene editing system. 19

Figure 1-7 - Refactoring of the *K. oxytoca nif* cluster.. 21

Figure 1-8 - A schematic of the traits that support aerobic nitrogen fixation and their significance to transferring nitrogen (ammonium) to cereals. 23

Figure 1-9 - Model for the regulation of intracellular NH_4^+ by GlnK. 24

Figure 1-10 – Overview of the several plant growth promoting phenotypes of rhizobacteria and their mechanism with plant root metabolism. 27

Figure 1-11 - Models of both toggle switches and oscillators (that can exhibit switch-like behaviour)..... 29

Figure 2-1 - Overview of the modular cloning system used in Golden Gate assembly. 59

Figure 2-2 – Conversion of acetylene to ethylene by nitrogenase. Cofactors illustrated. 73

Figure 2-3 - Pellicle formed by *A. caulinodans* ORS571 in semisolid 0.175% (w/v) agarose UMS media supplemented with 10 mM succinate and 30 μM nicotinic acid. Image taken after 42 hours growth. 74

Figure 2-4 - Standard curves of protein content and dry weight for rhizobacteria. 76

Figure 3-1 – Consensus sequence for σ70 promoters derived after a literature review. 89

Figure 3-2 – Alignment of the chosen constitutive promoters used in this study. 89

Figure 3-3 – Example plasmid map of the constitutive promoter and RBS analysis plasmids (pOPS1167, pOPS1168, pOPS1169, pOPS1171, pOPS1172, pOPS1173 & pOPS1174). 91

Figure 3-4 – Constitutive promoter library expression strengths. 93

Figure 3-5 – Consensus sequence for bacterial and archaeal ribosome binding site. 94

Figure 3-6 – Nucleotide alignment of the 4 RBS sequence library. 95

Figure 3-7 -Ribosome binding site (RBS) strengths. 97

Figure 3-8 – Secondary structure of the terminator DT14. 98

Figure 3-9 -Alignment of the terminator library DNA sequences.. 99

Figure 3-10 – Example plasmid map of the pOGG156-161 family of plasmids. 100

Figure 3-11 – Terminator strength profiles across all three test strains. Blue = Rhizobium sp. IRBG74. 102

Figure 3-12 – Plasmid map of the aTc induction test plasmid. 106

Figure 3-13 – Chemically inducible promoter induction curves and dynamic ranges. 109

Figure 3-14 – Plasmid maps of the degradation tag reporter plasmid and DAPG inducible LacI repressor plasmid. 112

Figure 3-15 – Degradation rate profiles of the suite of ssrA degradation tags. 114

Figure 3-16 – T7 promoter library alignment. 115

Figure 3-17 – A - Plasmid map of the aTc inducible T7 RNA polymerase system. 117

Figure 3-18 – Orthogonal T7 promoter transcriptional strength profiles. 118

Figure 3-19 - Plasmid map of the constitutive mRuby3 expression plasmid. ..121

Figure 3-20 – Excitation and emission spectra for the fluorescent protein reporter library in *A. caulinodans* ORS571. 123

Figure 3-21 – Excitation and emission spectra for the fluorescent protein reporter library in *A. brasilense* FP2. 125

Figure 3-22 – Absolute fluorescence responses of the library of fluorescent protein.. 126

Figure 4-1 – Architecture of a biological NOT gate constructed from the transcriptional repressor TetR and its cognate promoter Ptet. 138

Figure 4-2 - Plasmid map of the Plac reporter plasmid (pOPS1011). The other cognate promoter reporter plasmids (pOPS0457 (Ptet), pOPS0589 (Pphl) and pOPS0590) are identical to this plasmid other than different 'P' modules. 143

Figure 4-3 - Plasmid map of the LacI constitutive expression plasmid (pOPS1177). The other transcriptional repressor expression plasmids (pOPS1175 (TetR), pOPS1176 (PhlF) & pOPS1178 (ClR) are identical to this plasmid other than different 'SC' modules. 145

Figure 4-4 – Repression profiles of all the cognate promoter and transcriptional repressor pairs in *E. coli*. 148

Figure 4-5 - Repression profiles of all the cognate promoter and transcriptional repressor pairs in *A. caulinodans* ORS571 149

Figure 4-6 - Repression profiles of all the cognate promoter and transcriptional repressor pairs in *A. brasilense* FP2 151

Figure 4-7 -Growth curves of the dual reporter and transcriptional repressor expression plasmids. 152

Figure 4-8 – Orthogonality matrices for the NOT gate transcriptional repressor and cognate promoter pairs in all three test strains. 155

Figure 4-9 – DNA sequence alignment of the four orthogonal promoters and transcriptional binding regions (operator). 156

Figure 4-10 – Architecture of the synthetic rhizopine inducible NOT gate circuit. 157

Figure 4-11 – Plasmid maps of the rhizopine inducible NOT gate level one cassettes. 159

Figure 4-12 -Plasmid map for the rhizopine inducible NOT gate circuit. The LacI and PhlF based NOT gates were identical in architecture to the above plasmid other than the replacement of the transcriptional repressor ORF (above labelled tetR) and the cognate promoter pair (above labelled TetR operator region). ... 161

Figure 4-13 – Rhizopine inducible NOT gate dynamic range in both rhizobacterial test strains. ... 163

Figure 4-14 -Growth and fitness analysis for both the induced and uninduced states of the rhizopine inducible NOT gates in the engineered *R. leguminosarum* 3841 rhizopine sensitive strain. ... 165

Figure 4-15 – Hypothestical AND gate genetic circuit architecture utilising the constituent parts of multiple NOT gates to yield sfGFP expression when both inputs (DAPG and Arabinose) are present. ... 168

Figure 5-1 – Representation of various rhizosphere classes and characteristics. ... 176

Figure 5-2 – Mechanism for biological nitrogen fixation in diazotrophic PGPR. ... 180

Figure 5-3. Phosphate mineralisation and solubilisation cycle in the rhizosphere. ... 181

Figure 5-4 - A putative model for the mechanism of action of DAPG. ... 184

Figure 5-5 - IAA biosynthesis pathway from *Pseudomonas savastanoi*. ... 186

Figure 5-6 – Plasmid maps of the two level one IAA biosynthesis pathway plasmids. ... 188

Figure 5-7 - Plasmid maps of the two constitutive IAA biosynthesis pathway circuits. ... 189

Figure 5-8 - IAA biosynthesis analysis from the constitutive IaaM and IaaH gene circuits in *E.* coli and two rhizobacterial strains after 48 hours diazotrophic growth. ... 193

Figure 5-9 - Plasmid maps of the two level one DAPG inducible IAA biosynthesis circuit plasmids. ... 196

Figure 5-10 – Plasmid map of the DAPG inducible IAA expression plasmid in the pOGG027 broad host range vector (pOPS1166). ... 197

Figure 5-11 - IAA biosynthesis analysis from the DAPG inducible *iaaM* and *iaaH* gene circuit in *E. coli* and two rhizobacterial strains after 48 hours diazotrophic growth. 199

Figure 5-12 – *H. vulgare* Golden Promise growing in competitive soil environment for root hair length analysis. 204

Figure 5-13 – Root tip zone of *H. vulgare* Golden Promise. 205

Figure 5-14 – Light microscopy analysis of the effect of a synthetic constitutive IAA biosynthesis circuit (pOPS0667) in *A. brasilense* FP2 on the root hair length of *H. vulgare* plants. 206

Figure 5-15 – Impact of the synthetic constitutive IAA biosynthesis circuit on the root hair length of *H. vulgare* in a competitive soil system. 207

Figure 5-16 - *H. vulgare* Golden Promise growing in competitive soil environment for root length analysis. 210

Figure 5-17 – Processing sequence of operations to analyse whole root systems of *H. vulare* in a competitive soil environment. 211

Figure 5-18 - Impact of the synthetic constitutive IAA biosynthesis circuit on the root length of *H. vulgare* in a competitive soil system. 213

Figure 5-19 - Impact of the synthetic constitutive IAA biosynthesis circuit on the root diameter of *H. vulgare* plants in a competitive soil system. 216

Figure 5-20 -The putative DAPG biosynthesis pathway from *P. fluorescens* F113. 218

Figure 5-21 - Plasmid maps of the three level one DAPG biosynthesis pathway plasmids and level two complete constitutive DAPG expression circuit. 222

Figure 5-22 - DAPG biosynthesis from the synthetic constitutive circuit (pOPS0910) in *E.* coli and *A. caulinodans* ORS571. 225

Figure 5-23 - Plasmid maps of the two level one novel citric acid biosynthesis pathway plasmids. 229

Figure 5-24 - Plasmid maps of the two novel constitutive citric acid biosynthesis pathway circuits. 230

Figure 5-25 – Citric acid biosynthesis and phosphate solubilisation analysis from the constitutive GltA and CitC gene circuits in *E.* coli and two rhizobacterial strains after 72 hours growth. 232

Figure 5-26 -Predicted IAA induced *H. vulgare* root hair elongation model. 238

Figure 5-27 – Factors affecting the biosynthesis of active DAPG in rhizobacteria. 241

Figure 5-28 – Phylogenetic tree based on the CitC gene DNA sequence from a selection of bacterial species. 243

Figure 6-1 – Traditional CRISPR-Cas9 DNA cleavage and alternative CRISPRi-dCas9 mechanisms of action *in vivo*. 248

Figure 6-2 sgRNA molecule architecture. 250

Figure 6-3 – Architecture of the dCas9 wt, EQR and VRER variants expression plasmid (pOPS1150 (wt), pOPS1151 (EQR) and pOPS1152 (VRER)). 254

Figure 6-4 – Plasmid maps of the three dCas9 expression plasmids for use in rhizobacteria species. 256

Figure 6-5 - Effect of dCas9 wt expression on the growth of *A. caulinodans* ORS571 in liquid media. 258

Figure 6-6 - Effect of dCas9 wt expression on the growth of *A. brasilense* FP2 in liquid media. 261

Figure 6-7 - Architecture of the dCas9 wt, EQR and VRER variants reporter test plasmids (pOPS1153 (wt), pOPS1154 (EQR) and pOPS1155 (VRER)). 263

Figure 6-8 - Plasmid maps of the dCas9 reporter plasmids for analysing dCas9 variants activity in rhizobacteria. 264

Figure 6-9 – dCas9 variant repression profiles in both *A. caulinodans* ORS571 and *A. brasilense* FP2. 267

Figure 6-10 – Plasmid map of the sgRNA expression plasmid (pOGG230) with Golden Gate compatible LacZ selection cassette for BsaI-mediated integration of promoter and 20 bp targeting portion of sgRNA molecule. 269

Figure 6-11 – Mechanism for the Golden Gate assembly of sgRNA molecules into the stable dCas9 variants expression plasmid. 271

Figure 6-12 - Architecture of the Golden Gate dCas9 wt, EQR and VRER and sgRNA expression plasmids with integrated sgRNA Golden Gate cassette ... 272

Figure 6-13 - CRISPRi repression of the GS-GOGAT cycle in *A. brasilense* FP2. 274

Figure 6-14 – MSX Inhibition of Glutamine Synthetase in *A. brasilense* FP2....277

Figure 6-15 – Genetic context of the *A. brasilense* FP2 GlnA gene.....278

Figure 6-16 - Plasmid maps for the analysis of GlnA expression in *A. brasilense* FP2. 280

Figure 6-17 – *A. brasilense* FP2 *glnB*/*glnA* promoter expression analysis under various growth conditions. 282

Figure 6-18 – sgRNA targets for the repression of GlnA expression by CRISPRi in *A. brasilense* FP2.....285

Figure 6-19 – DNA sequence alignment of the six sgRNA promoter and 20 bp target regions for CRISPRi of PglnA in *A. brasilense* FP2.....286

Figure 6-20 – CRISPRi qPCR quality control data for *glnA* gene repression in *A. brasilense* FP2. 290

Figure 6-21 – Chemical overview of the GS glutamyl transferase reaction used to detect GS activity in *A. brasilense* FP2. 291

Figure 6-22 – Equation to calculate activity of GS. 292

Figure 6-23 – Analysis of the dCas9 wt repression of the GlnA gene of *A. brasilense* FP2 in diazotrophic growth conditions. 294

1 Introduction

1.1 The Importance of Nitrogen in Agricultural Systems

In modern agriculture there are two main methods for providing the nitrogen requirements of crops, especially the cereals *Triticum aestivum* (common wheat), *Oryza sativa* (Asian rice) and *Zea mays*, which are the three main crop species farmed by mankind. Historically the primary method has been to rotate a nitrogen-fixing variety of 'green-manure' plants such as Hairy Vetch (*Vicia villosa*), Crimson Clover (*Trifolium incarnatum*) or alfalfa (*Medicago sativa*) with the cereal crop year on year[1]. This allows for several advantages with respect to cereal crop yield including promoting the nitrogen content of the soil, lowering the pH of basic soils and the aggregation of dry soils to prevent the 'dust-bowl' effect[2]. The second, and vastly more common approach is to use synthetic fertilisers derived from the energy-intensive Haber-Bosch process[3]. However, it is increasingly clear that this use of fertilisers is environmentally unsustainable due to detrimental impacts such as watercourse eutrophication and the production of vast quantities of potent greenhouse gasses[4]. Twenty four percent of global greenhouse emissions are as a result of fertiliser manufacturing and use and cause a massive impact on the rate of global warming, especially due to nitrous oxide release from fertilised land[5]. Economic inaccessibility of fertilisers to developing countries causes a massive

yield gap in farmers from different parts of the world. Only one farmer in twenty is able to buy fertilisers in sub-Saharan Africa, meaning crop yields are low due to the nutrient-poor soil[6]. Therefore, there has been an increasing movement in scientific enquiry into the potential use of biological fertilisers, especially the implementation of bacteria that fix atmospheric nitrogen naturally.

1.2 Natural Diazotrophic Bacterial Interactions with Crop Plants

Reactive nitrogen, usually in the form of ammonia (NH_3) ammonium (NH_4^+) or nitrates (NO_3^-), is required by the plant for the synthesis of amino acids and therefore proteins[7]. Although nitrogen is extremely abundant on the planet, it mainly exists in the inert and unreactive noble gas form (N_2). One major group of plants that are able to successfully access fixed nitrogen from the atmosphere are the legumes. These plants, including *Pisum sativum* (pea), *Phaseolus vulgaris* (common beans) and *Trifolium pratense* (red clover), do so via their symbiotic relationship with diazotrophic bacteria known as the rhizobia[8]. The mechanism by which rhizobia fix nitrogen for the plant is one of the most well studied symbioses in nature. The bacteria contain the enzyme nitrogenase that reduces N_2 gas into ammonia (NH_3) and then by protonation to ammonium (NH_4^+). The formation of ammonia from molecular hydrogen (H_2) and nitrogen (N_2) has a total negative enthalpy of reaction (ΔH^0 = - 45.2 kJ) and a very high activation energy (E_A = 230 – 420 kJ mol^{-1}, Modak 2002). Nitrogenase is an enzymatic catalyst that reduces

this energy barrier to the level that the reaction can take place at ambient environmental temperatures (10 – 40˚C).



Figure 1-1 - Nitrogen-fixing nodules formed on the root of the legume *Medicago trucatula*[10]. The pink colour of the nodules indicates the presence of leghaemoglobin, which keeps the oxygen level of the nodule low enough for the nitrogenase enzyme to function at its optimum level.

Rhizobia occupy root nodules on specific plants of the *Leguminosae* and *Rosid* families of plants (Figure 1-1). Root nodules are adapted to produce a low-oxygen flux environment for the optimum function of the oxygen-inhibited nitrogenase enzymes within the rhizobia[11]. This low-oxygen flux environment achieved by the formation of an oxygen-absorbing protein called leghaemoglobin that provides the respiring rhizobia with oxygen but at a low enough concentration as not to inhibit nitrogenase activity (Figure 1-1). Within nodules rhizobia alter their morphology and become bacteroids, which in IRLC (inverted repeat lacking clade) legumes are terminally differentiated[12]. At this stage in their life-cycle bacteroids fix atmospheric nitrogen to ammonia that is then transferred to the root nodule cytosol as ammonium ions. Here the ammonium is assimilated mainly to L-

asparagine which is then actively transported throughout the plant, especially the shoots, for biosynthesis[13,14]. In turn the host plant supplies the bacteroids with a dicarboxylate carbon source with which to oxidise and release energy. In temperate legumes such as peas the major compound supplied is L-malate.

1.3 Engineering Non-leguminous Diazotrophic Bacteria-Plant Interactions

1.3.1 Global Agriculture Overview

For thousands of years cereal crops have been the primary component of the human diet and have been central to shaping human civilisation from the fertile crescent to Southern America. The global agricultural landscape is dominated by the growth of cereals, specifically rice (*Oryza sativa*), wheat (*Triticum aestivum*) and maize (*Zea mays*; Figure 1-2). More than 50% of global caloric intake is derived directly from the above cereal grain crops making it the primary focus in the field of sustainable agriculture, especially biofertilisation.

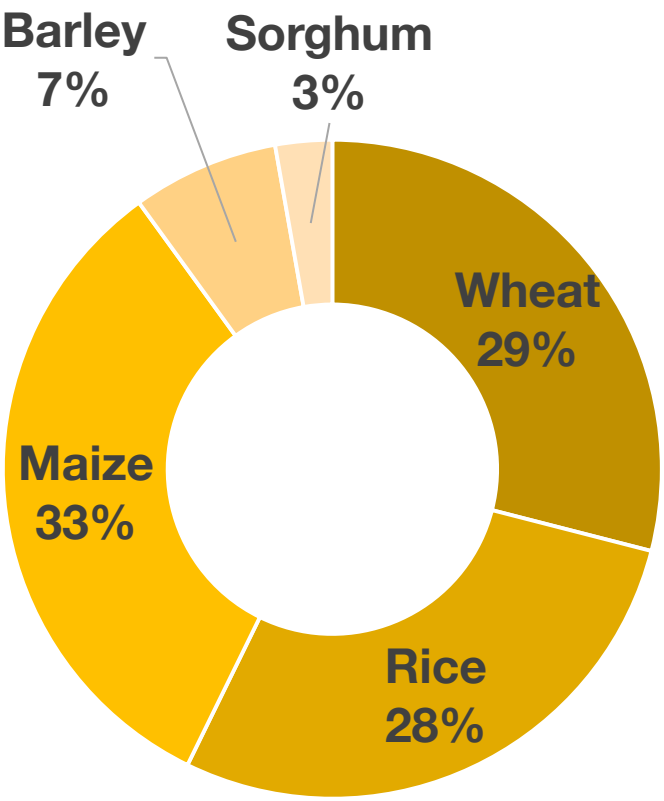


Figure 1-2 – Top 5 global agricultural crops in 2017. Data sourced from the Food and Agriculture Organisation of the United States of America.

For decades there has been interest in engineering a rhizobia-like symbiosis in cereal crops. Organisations such as the Bill and Melinda Gates Foundation and the Food and Agriculture Organisation of the United Nations (FAO) are funding efforts to close the gap between the demand for cereal crops and the yields produced by ecologically safe farming methods[15,16]. Approaches for the transfer of nitrogen fixation into cereals fall into one of three categories: The first being the synthetic formation of nodules on non-nodulating plant species. A recent discovery of the common symbiotic pathway in cereal crops has suggested that engineering the recognition of rhizobia by cereal crops might be possible[17]. There

is evidence that the nodule induction pathway in legumes and the initiation of mycorhizzal interactions in all plants is very similar[18]. However, the organogenesis of nodules on plant roots is a very complex multi-stage development that is not yet fully understood[19].

The second approach is to directly engineer a functional nitrogenase enzyme into the cells and organelles of the cereal plant. An approach that as of yet has shown little success as the enzyme is extremely oxygen sensitive and engineering an anoxic environment in plant tissues is particularly difficult[10,20]. However, some recent success was observed when a functionally active Fe-Iron component of nitrogenase from *Azotobacter vinelandii* was successfully expressed in yeast mitochondria[21].

The final, and most promising approach, is to utilise endophytic bacteria that already associate with cereals as a chassis to engineer nitrogen fixation into plants roots and within the rhizosphere[22,23]. Some success has been observed when in 2013 a group from the Instituto de Genética Ewald A. Favret, Argentina showed a successful transfer of an entire nitrogen-fixation gene cluster from *Pseudomonas stuzeri* to *Pseudomonas protogens* Pf-5[24]. This allowed *P. protogens* to grow micro-aerobically utilising N_2 as its only nitrogen source. A feat possible due to the high respiration rate and therefore rapid assimilation of oxygen out of the cellular cytoplasm. Even more remarkably, when Arabidopsis, Alfalfa (*Medicago sativa*), Tall Fescue (*Festuca arundinacea*) and Common Wheat

(*Triticum aestivum*) were inoculated with the modified nitrogen-fixing *P. protogens*, significant plant growth effects were observed in sterile systems[24]. These advances in transferring nitrogen fixation across species have been incremental and undefined. With the advent of the field of synthetic biology, techniques are available to manipulate nitrogen fixation in far greater detail and therefore allow its controlled transfer across species with far greater efficiency.

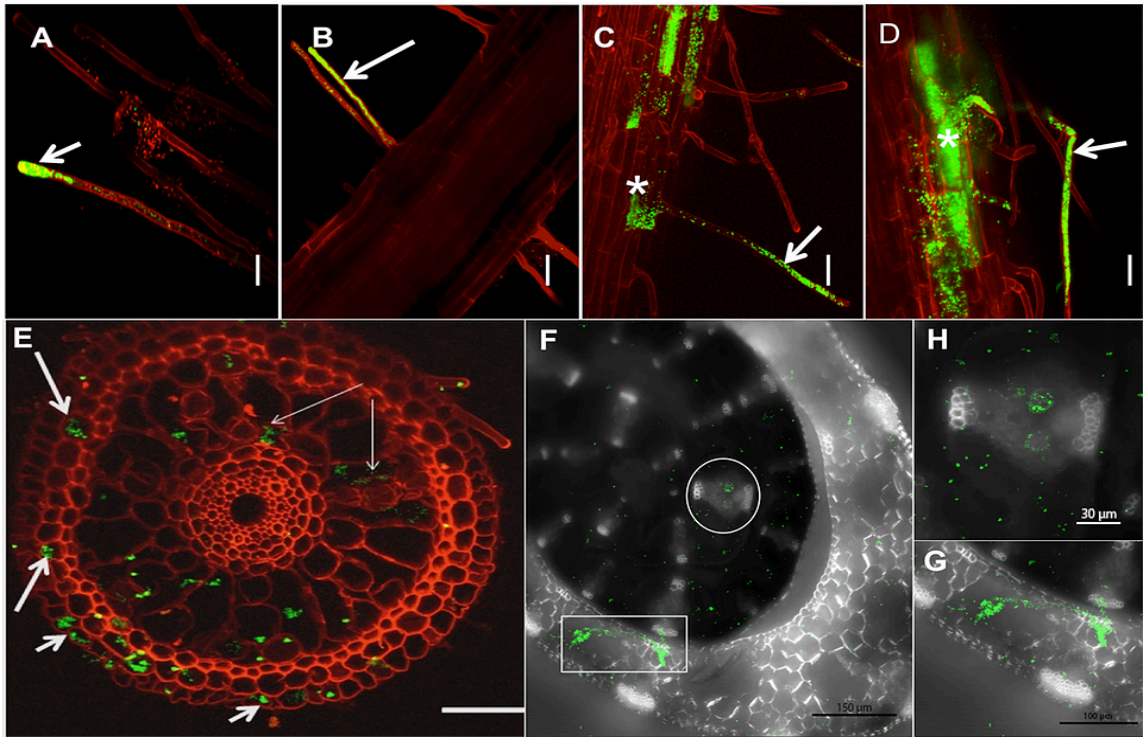


Figure 1-3 - Endophytic colonisation of O. sativa by Rhizobium IRBG74 via both lateral root junctions and root hairs. IRBG74 enters the root cortex via the root hair (A-D) as well as lateral root junctions. Arrows indicate the infection thread of rhizobia. The cells localise within the intercellular spaces, within the xylem and stems (E-F). Arrows indicate groups of rhizobia cells within the plant root cortex. Adapted from Shubhajit, 2014.

1.3.2 Rhizobacterial Plant Colonisation

Plant colonisation by endophytic bacteria is a trait that is governed by hundreds of genes in a complex host-symbiont interaction[25]. To engineer this trait into a novel bacterium would be extremely difficult. Instead it has been proposed to utilise a rhizobacteria that are known to already associate strongly with target crop plants. There are several known characterised rhizobacteria species capable of associating with the roots of cereals two of which will be discussed in more detail below[26,27]. For example, the *Sesbania rostrata* endophyte *Rhizobium* str. IRBG74 has been shown to form high-density colonisation on the root of *O. sativa*, as well as effectively colonising the plant root cortex and other vascular tissues[28] (Figure 1 – 3). In this study we chose to focus our engineering efforts on two rhizobacterial strains that are confirmed to associate with desired cereal crop plants and have the ability to fix atmospheric nitrogen.

1.3.3 *Azorhizobium caulinodans* ORS571

Azorhizobium caulinodans ORS571 represents a particularly important chassis for engineering as it has been shown to be an endophyte as well as a free-living nitrogen fixing rhizobacteria [29–31]. *A. caulinodans* ORS571 can fix nitrogen under microaerobic conditions (3-6% oxygen) in the laboratory and has also been confirmed to be an endophyte of *O. sativa* and *T. aestivum*[29,31] (Figure 1-4). These

two characteristics make *A. caulinodans* ORS571 an ideal target for engineering nitrogen fixation and ammonia secretion to target crop plants. Ammonia production by nitrogen fixation *in vivo* can be analysed effectively in liquid culture, an advantage that is almost unique within the rhizobacteria [32].

A. caulinodans ORS571 has been studied extensively in terms of its nitrogen regulatory network. Coupled with the full annotation of its genome the organism is a model in the field of nitrogen fixation and a target of many research groups attempting to engineer the fixation pathway [33].

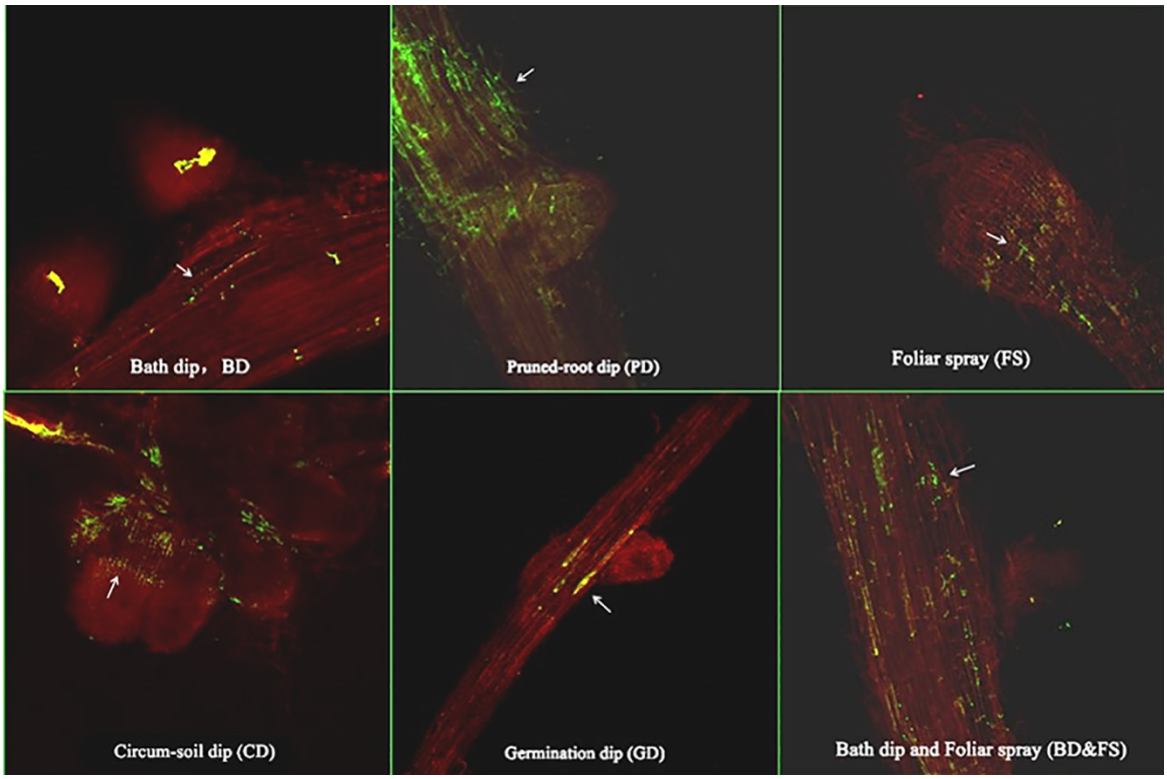


Figure 1-4 – Infection of wheat root tissue with the rhizobacteria *Azorhizobium caulinodans* ORS571 with different application methods. 2, 4-D was added as an inducer to all methods to form para-nodules. Arrows refer to the appearance and position of *A. caulinodans* ORS571. (A) BD, a few of the rhizobacteria existed within para-nodules. (B) PD, a certain amount of *A. caulinodans* ORS571 is found within para-nodules and root cortex. (C) CD, a high amount of rhizobacteria can be detected in para-nodules. (D) FS, low amount of rhizobacteria was observed in para-nodules. (E) GD, A minimal amount of *A. caulinodans* ORS571 was observed in para-nodules with germination dip application. (F) SD&FS, there were rhizobacteria within para-nodule primordia, but none observed inside. Adapted from Huawei *et al.,* 2017.

1.3.4 *Azospirillum brasilense* FP2

A second chassis used for synthetic biology engineering is the rhizobacterium *Azospirillum brasilense* FP2. In recent years this organism has come to the forefront in terms of its potential colonization and plant growth promotion potential. Several groups are working with this organism to understand its inherent plant growth promotion mechanisms and to engineer them in different capacities[34]. Several studies have shown that *A. brasilense* FP2 colonises the root systems of wheat and maize through the cracks that form in lateral root emergences (Figure 1-5).

A. brasilense FP2 is a free-living soil rhizobacteria that has the ability to affect the growth of numerous agricultural crops worldwide through the excretion of various hormones and the bacteria's ability to fix atmospheric nitrogen[35,36]. *A. brasilense* FP2 secretes multiple phytohormones, the most common being Indole Acetic Acid (IAA). *A. brasilense* FP2 produces IAA by utilizing the amino acid tryptophan as a precursor. IAA along with various other phytohormones alter the metabolism, growth profile and the morphology of the associated plant root system in favor of better mineral and water adsorption[37]. Just like the aforementioned *A. caulinodans* ORS57 (1.3.3 above), *A. brasilense* FP2 is capable of diazotrophic growth through the direct assimilation of ammonium and the activity of nitrogenase. Nitrogen fixation occurs under microaerobic conditions in liquid media as well as in the microaerobic regions of biofilms[38].

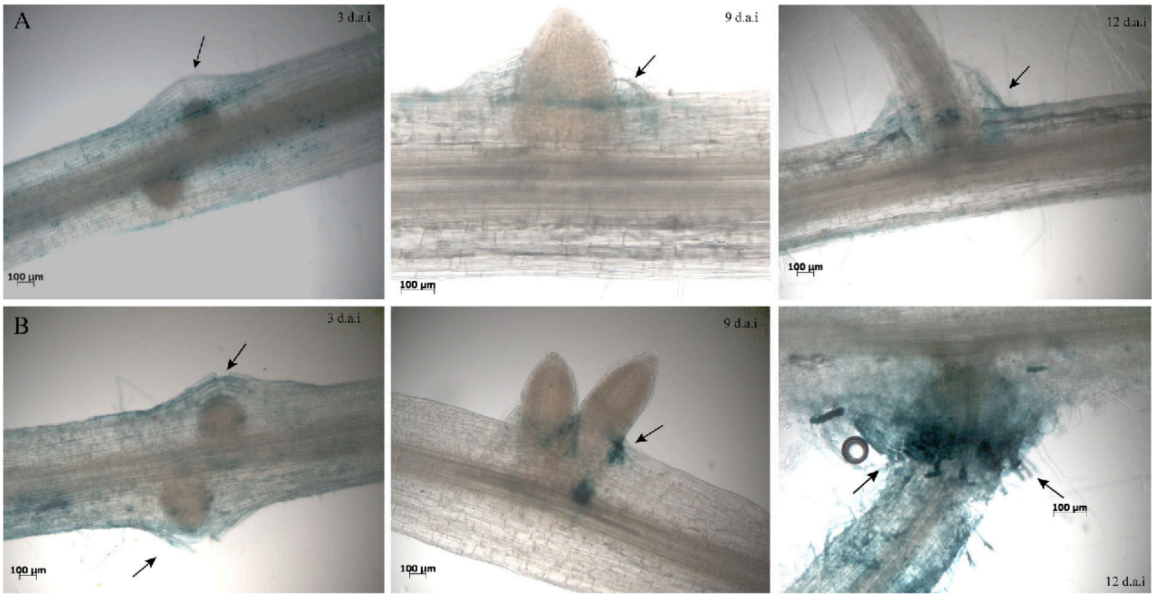


Figure 1-5 - Colonisation of the lateral root emergence of wheat (*Triticum aestivum* var. CD104) by *Azospirillum brasilense* FP2: **A** - 3, 9, and 12 days after inoculation (d.a.i). **B** - Wheat root inoculated with HM053 strain: 3, 9, and 12 d.a.i. Seedlings were inoculated with *Azospirillum brasilense* FP2 (10^7cells per plant) and the GUS activity (blue pigment) from a nifH-gusA reporter was analysed by light microscopy. Adapted from Santos *et al.,* 2017[39].

Many countries such as Brazil and Argentina use bacterial inoculants containing *A. brasilense* FP2 alone or in concert with other plant growth promoting bacteria[40]. *A. brasilense* FP2 belongs to the alpha subclass of Proteobacteria[41]. The genome consists of multiple small chromosomes instead of a single large circular chromosome, the DNA sequence of which has been drafted and is currently being validated by collaborators at the Universidad Nacional de Río Cuarto (UNRC), Córdoba, Argentina. In this study we used a draft genome mapped to the fully assembled *A. brasilense* Az39 isolate for gDNA sequence based work[42].

In the past decade much attention has been focused on a spontaneous mutant of *A. brasilense* FP2 (Sp7 ATCC 29145, SmR, NalR) named HM053, first characterised by Machado et al. 1991[43]. *A. brasilense* HM053 has been verified to excrete significant amounts of ammonium and remain in a nitrogen-fixing state in the presence of high concentrations of NH_4^+, a unique phenotype in diazotrophic rhizobacteria. For this reason *A. brasilense* HM053 has been nominated as a candidate biofertilizer to supply nitrogen to *Gramineae* plant family[44]. Machado et al. (1991) suggested that HM053's ability to excrete ammonium is related to its low glutamine synthetase (GS) activity, resulting in a deficient NH_4^+ assimilation flux and therefore the excretion of excess NH_4^+ produced after nitrogen fixation. A recent study showed that *A. brasilense* HM053 can promote the growth of *Setaria viridis* and was capable of providing 100% of the plants nitrogenous requirements in gnotobiotic growth conditions[45].

1.4 Synthetic Biology

1.4.1 Introduction to Synthetic Biology

In the past decade the nascent and multidisciplinary field of 'Synthetic Biology' has been rapidly growing. Synthetic biology has allowed for the more precise and practical engineering of organisms. Techniques such as CRISPR/Cas9 genome editing and rapid computational modelling have been rapidly adopted over multiple domains of life from *Arabidopsis thaliana* to *human* embryos[46–48]. In bacterial engineering, the current trend is moving away from model organisms such as *E. coli* into more unusual but potentially more applicable and marketable species such as the *Staphylococci* and *Streptomyces*[49].

Many new *in silico* tools have become available to control gene expression *in vivo*. Tools such as the 'Salislab RBS Predictor' allows researchers to upload gene to be expressed and the algorithm will predict the various synthetic ribosome binding sites that will give various levels of translation[50]. There are also new libraries of DNA 'parts' being updated regularly and characterised in multiple different host organisms. One such library is the iGEM biobrick repository that stocks and distributes various DNA sequences ranging from transcriptional regulators to biosynthesis clusters[51]. The standardisation of expression cassettes has been one area of intense research to reduce the metabolic context and species dependence of gene expression[52]. New DNA parts such as the RiboJ

insulator have been developed to reduce the impact of promoter sequence on translation due to secondary structure and mRNA folding[53,54]. Precise termination of gene expression is also extremely important to make sure any synthetic expression targets only the desired gene and not downstream elements and coding sequences. It is for this reason that powerful synthetic terminators have been designed and characterised using both *in vivo* DNA-fold based approaches and genome mining and testing from multiple organisms[55,56].

DNA synthesis costs have been plummeting rapidly over the past ten years from $20 per base pair down to $0.07[57]. New synthesis techniques such as the high-throughput silicon-based DNA writing platform from Twist Bioscience allows 9,600 synthesis reactions to take place concurrently, yielding strands of up to 3.5 kb[58]. The advent of high-throughput and accurate gene synthesis allows for the realisation of more complex genetic designs that are composed of multiple synthetic parts, such as the *nif* cluster and synthesis of methyl halides from heterologous biomass[59–61].

1.4.2 Computational Logic in Bacteria

The ability for cells to become processors has been successfully demonstrated on multiple occasions, from simple NOT-gate inductions to a fully automated system for designing advanced genetic[54,62]. This allows for the tight control and

tuning of different phenotypes under different environmental conditions and inputs, especially in unique niches such as those found in plant-bacteria symbioses. Circuits and sensors have begun to be transferred to plants directly allowing for example the expression of reporters in response to toxins[63] and this engineering approach will be essential for engineering interactions between the plant root and modified rhizosphere bacteria.

Taking principles of modularity and control systems from engineering, synthetic biologists are defining minimal subsets of genes and DNA parts with well-defined functions. The minimisation of parts such as promoters, ribosome binding sites and in some cases entire biosynthesis clusters yield far smaller DNA sequences that are cheaper to make and more readily transferrable into host chassis[64]. One of the most promising advances in the field has been the characterisation and implementation of systems to edit DNA sequences anywhere in an organism's genome, quickly and cheaply. The below chapter discusses this new technology and its applications in more detail.

1.4.3 CRISPR Interference Technology (CRISPRi)

Rhizobacteria have complex gene expression regulation networks like most living organisms. A significant portion of the information encoding the function and various phenotypes of a rhizobacterial cell are dictated by its transcriptome. Transcription is the first step in gene expression and is therefore a nexus of phenotypic regulation. To understand this fundamental cellular process requires modern experimental tools that are capable of a high-throughput interrogation of the cells transcriptional landscape. The recently characterised CRISPR interference technology (CRISPRi) enables highly specific control of both genomic and exogenous gene expression across a broad host range of chassis[65].

The CRISPRi system was developed after the discovery of the CRISPR-Cas RNA-mediated immune system found in bacteria and archaea[66]. Most early CRISPR-Cas research was performed in the bacteria *Streptococcus pyogenes* and is the source of many of the tools available commercially to this day. CRISPR-Cas is a system divided into five classes, of which the type II is the most ubiquitously utilised for genome-engineering applications[67]. The discovery of the discreet components of the type II CRISPR system and their mechanisms were essential for the development of a widely-applicable genome-engineering applications[68]. One of the most significant early applications was the demonstration that *Streptococcus thermophilus* could specifically cleave double-stranded DNA in a

system mediated by the Cas9 protein[69]. The later discovery of a short DNA sequence adjacent to the RNA-hybridising site was termed the protospacer-adjacent motif (PAM; Mojica et al. 2009). This mechanism enables the CRISPR-Cas system to discriminate self from non-self. Subsequent engineering of the transactivating CRISPR RNA (tracRNA) and CRISPR RNA (crRNA) molecules resulted in the fusion of the two essential RNA molecules into one single guide RNA molecule (sgRNA)[71]. By combining these applications and discoveries, genome editing from simple single nucleotide modifications to whole biosynthetic pathway integrations have been demonstrated[72,73].

In addition to utilising the type II CRISPR system for editing genomic sequences, the CRISPR-Cas9 has subsequently been modified to be used as a sequence-specific, non-mutagenic transcription regulation tool. The first application of the system was demonstrated by introducing two-point mutations into the *S. pyogenes* Cas9 in the two nuclease domains, HNH and RuvC1 (Figure 1-6B). The resulting dead Cas9 (dCas9) is a nuclease-deficient and therefore unable to cleave DNA[74]. It does however retain the full ability to bind specifically to DNA as guided by a sgRNA molecule. This allows for the dCas9 protein to be implemented as a recruiter for various effector proteins for gene regulation at the transcriptional level, yielding both repression and activation characteristics[75].

The CRISPRi system has been shown to function well in several model organisms such as *E. coli*[66], *Saccharomyces cerevisiae*[76] and mammalian cell lines[77] but is yet to be implemented in any rhizobacteria species.

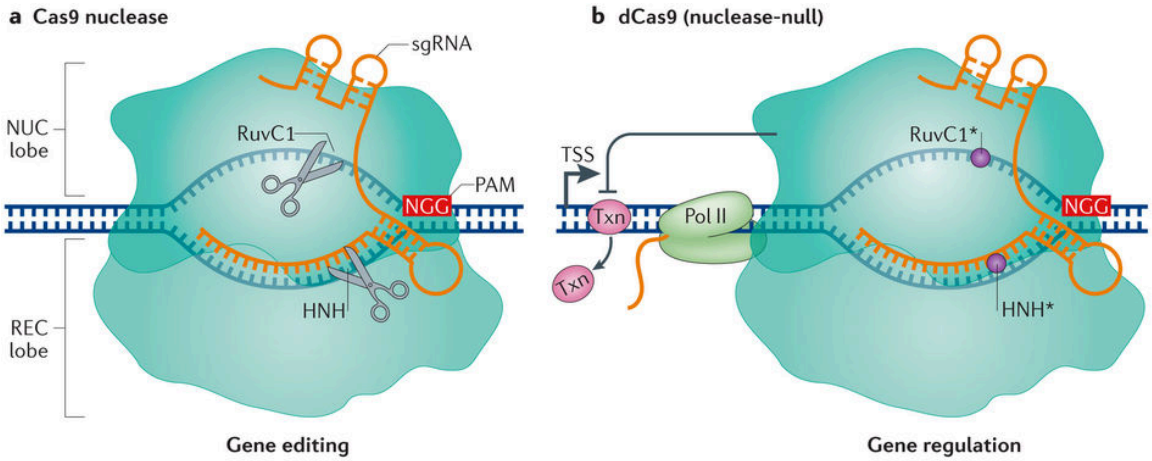


Figure 1-6 – The CRISPRi system derived from the naturally occurring CRISPR-Cas9 gene editing system. **A** - The *S. Pyogenes* Cas9 protein consists of both a nuclease (NUC) domain and a recognition (REC) domain. The Cas9 is targeted to specific 20 bp DNA sequences by the direct hybridisation of the sgRNA with the target DNA. This presence of a 5' PAM sequence in the DNA is essential to the hybridisation and in *S. pyogenes* consists of the sequence NGG. **B** – The engineered dCas9 protein contains two mutations, one in the RuvC1 (D10A) and HNH (H841A) domains. This inactivates the nuclease activity but retains the full ability of the dCas9 protein to bind to specific 20 bp sequences of DNA. Binding of the dCas9 protein downstream or on top of a transcription start site (TSS) can block transcription initiation or elongation by blocking RNA polymerase II interaction with the DNA. Figure adapted from Dominguez *et al.,* 2016[74]

1.5 Genetic Engineering of Diazotrophic Rhizobacteria

1.5.1 Diazotrophic Rhizobacteria

The mechanism and genetic regulatory network of nitrogen-fixing enzymes in rhizobia is complex. Previous studies have shown that when the expression of a single constituent gene of the *nif* cluster is sub-optimal, the efficiency of nitrogen fixation drops dramatically[22]. Many multimillion-pound research projects and international collaborative teams are working on modelling nitrogen fixation in rhizobia. The goal being to engineer novel minimalised and refactored systems for insertion across a broad range of both free-living and endophytic rhizobacteria[78]. Modelling the nitrogen-fixation metabolic pathways by flux analysis allows for the determination of pinch-points and targets for synthetic engineering[79]. One particularly successful study showed a way to refactor an entire nitrogen fixation cluster from *Klebsiella oxytoca* by completely re-shuffling essential gene codon usage and removing all regulatory (known and unknown) genes and operators from the sequence. The genes were then placed under the control of synthetic, well-characterised 'parts' including promoters, operators, ribosome binding sites and terminators[80] (Figure 1-7). By multiple iterations of optimisation, the efficiency of nitrogen fixation was slowly increased to 7.4% of the wild-type nitrogen fixation output, a remarkable result considering the complexity of the *nif* cluster.

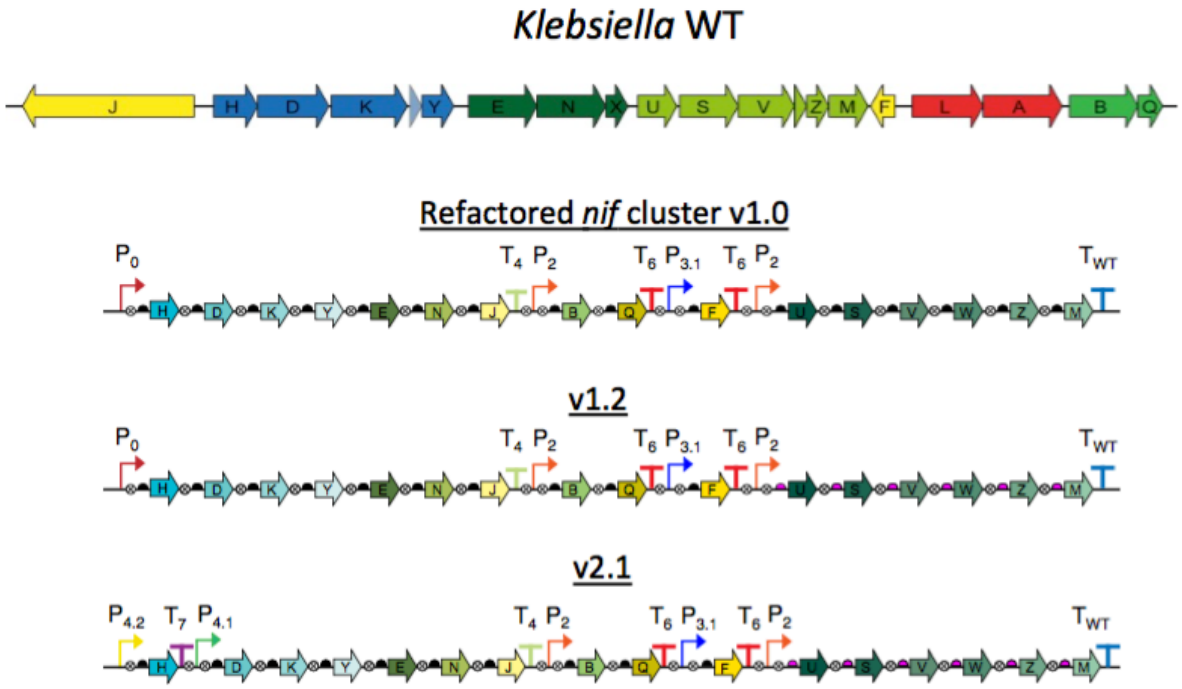


Figure 1-7 - Refactoring of the *K. oxytoca nif* cluster. Firstly, all the codon usage of the *nif* genes were randomly changed to remove any DNA sequence specific regulation. Next the genes were isolated as distinct units without the promoters, ribosome binding sites and terminators and were shuffled in their order. Next was a series of iterations to replace the promoters, ribosome binding sites and terminators with synthetic, well-characterised replacements. By doing so the activity of the nif-cluster was slowly restored towards wild-type levels (adapted from Temme *et al.,* 2012).

1.5.2 Ammonia Secretion and Plant Growth Promotion

Significant ammonium release by the rhizobia in the proximity or within plants is a crucial goal for the successful promotion of plant biosynthesis[81]. No matter how well an organism fixes nitrogen, if there is no ammonium secreted out of the bacterial cell, there will be no benefit to the target crop plant, except upon the death and breakdown on the microbe. Disrupted regulation of the core nitrogen fixation pathway has been shown to lead to ammonium excretion in *Azotobacter vinlandii,* where mutations in the key regulators *nifA* and *nifL* resulted in ammonium release[82,83]. Characterised mutations in the glutamine synthetase (*glnA*) gene within *Azospirillum brasilense* caused significant ammonium release that possibly increased the growth of *T. aestivum* under nitrogen-limiting conditions[84,85]. The deletion of two ammonium uptake transporters encoded by *amtB1/2* caused a significant increase in ammonium secretion in *P. stuzeri*[86] (Figure 1-8).

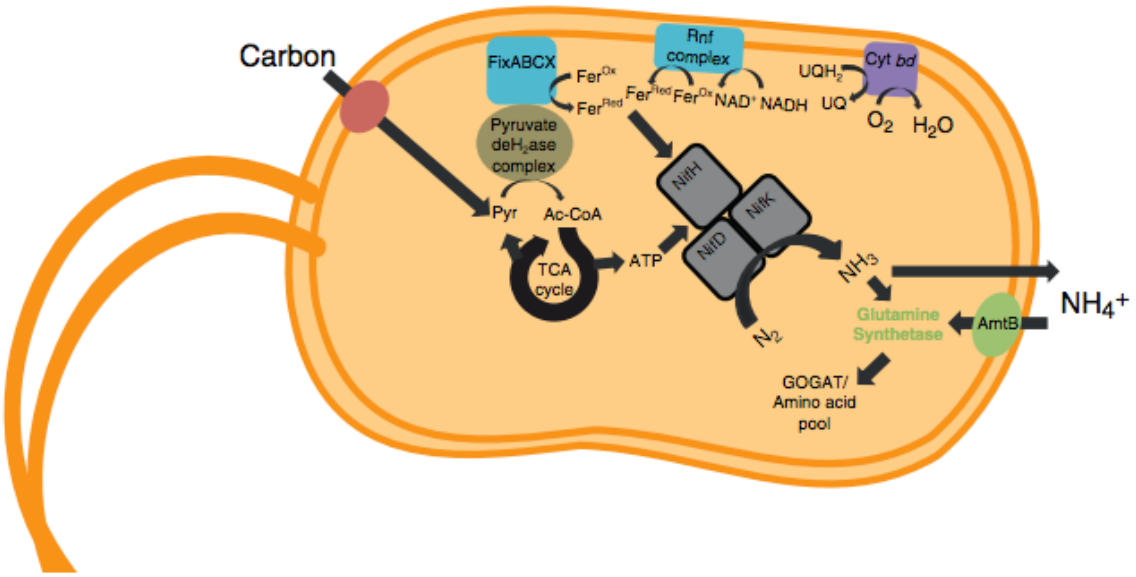


Figure 1-8 - A schematic of the traits that support aerobic nitrogen fixation and their significance to transferring nitrogen (ammonium) to cereals. Energy to supply ATP to the intensive nitrogenase reaction comes from the rhizosphere and could be an avenue for biasing the rhizosphere community. Electron transfer to the nitrogenase can be carried out by FixABCX or potentially the Rnf complex (seen in blue). The mechanism by which this occurs has not been fully characterised but could be mediated by the reduction of ferredoxin by a coupled exergonic reaction. Oxygen protection could be carried out by high respiration rates inferred by the use of alternative terminal oxidases (seen in purple). As well as adjusting the regulation of the nitrogenase reaction, the enzymes involved in ammonium uptake and assimilation (seen in green) represent ideal targets for genetic modification and engineering. Thus yielding more ammonium released to the plant rather than assimilated by the rhizobia. Adapted from Geddes *et al.,* 2015.

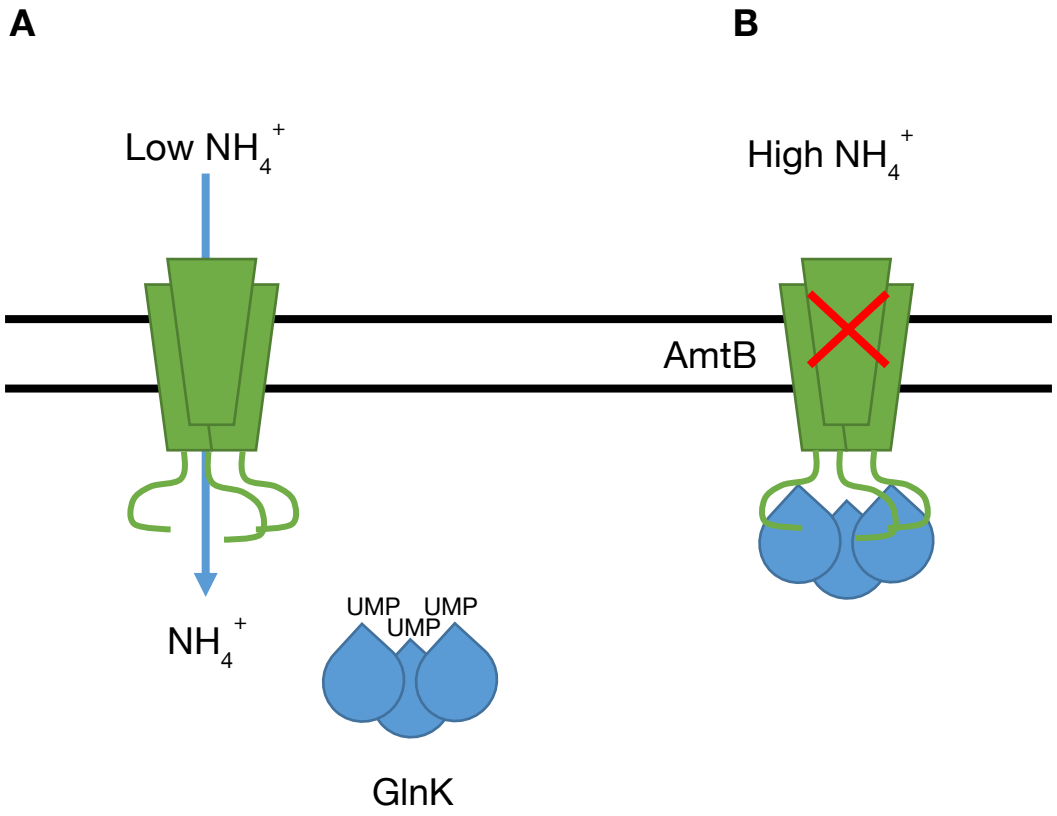


Figure 1-9 - Model for the regulation of intracellular NH_4^+ by GlnK. **A** - In conditions of nitrogen limitation, GlnK is uridylylated and does not associate strongly with the membrane or AmtB. AmtB is therefore active and will efficiently uptake ammonium from the surrounding rhizosphere/environment. **B** – When ammonium is in high enough concentration GlnK is rapidly de-uridylylated and the unmodified form of the protein associates tightly with AmtB on the inner membrane. This association causes the inhibition of AmtB and ammonium cannot be transported across the membrane into the cell until the concentration of intracellular NH_4^+ decreases.

AmtB is regulated by the nitrogen-regulatory protein GlnK, a member of the P-II protein family. When ammonia is in abundance within the cell GlnK becomes de-uridylylated and blocks the cytoplasmic pore exit of the AmtB pore-forming protein, thus preventing ammonia from entering the cell. By adjusting this

feedback loop it may be possible to reduce ammonia uptake, and therefore ammonia efflux by diffusion[87] (Figure 1-9). Ammonia and amino acid release have also been observed when the core enzymes of nitrogen fixation within *Rhizobium etli* were altered. By adding an orthogonal glutamate dehydrogenase (GDH) from *E. coli*, an alternative nitrogen assimilation reaction was available to the bacterium by forming L-glutamate directly from ammonia and a-ketoglutarate[88].

1.5.3 Plant Growth Promotion Phenotypes

The field of biological plant growth promotion focuses on phenotypes other than just ammonia secretion. Rhizobacteria have an exceptional ability to promote the growth of associated plants in mechanisms such as the production of siderophores, the solubilization of phosphate, the production of active antifungal compounds and the production of plant growth hormones[27,89–91]. After the aforementioned nitrogen, phosphorus is the most limiting nutrient for crop plant growth. Despite abundant phosphorus content in most agricultural soils, it is not available in a form accessible to plant uptake. Plants are only able to absorb soluble phosphorus in the form of mono- and di-basic phosphate[92]. Certain rhizobacteria species are able to solubilse phosphate by way of secreting organic acids such as citric and gluconic acid[93]. These acids released into the rhizosphere by the rhizobacteria and lower the pH of the environment, allowing the phosphate to solubilize and be taken up by associated plants[92] (Figure 10).

Phytohormones, such as auxins (IAA) control many stages in plant growth and development such as cell elongation, division and differentiation[11]. The microbial biosynthesis pathways and fundamental mechanisms of auxin action on plants is a highly researched field. Auxin (IAA) is one of the most important phytohormones produced by several strains of rhizobacteria such as *Micrococcus luteus* and has been commercialized as a plant growth promoting inoculant[94].

The production of antifungal compounds by rhizobacteria is an expanding field of study due to intense farming techniques reducing the barrier to entry for diseases such as Take-All caused by several *Phytophthora* species. 2,4-diacetylphloroglucinol (DAPG) is a well-known antifungal compound produced naturally by biocontrol fluorescent *Pseudomonas* species[95]. At low concentration (< 10 μM) DAPG has been shown to be a signal for systemic resistance in plants[96], stimulating root exudation[97] and enhancing root branching[98].

One of the key benefits in utilizing synthetic biology to engineer the above plant growth promoting phenotypes, is that their combination into a single chassis can be achieved (Figure 1-10). Therefore, a single engineered inoculant could provide a multi-faceted plant growth promotion effect rather than just one.

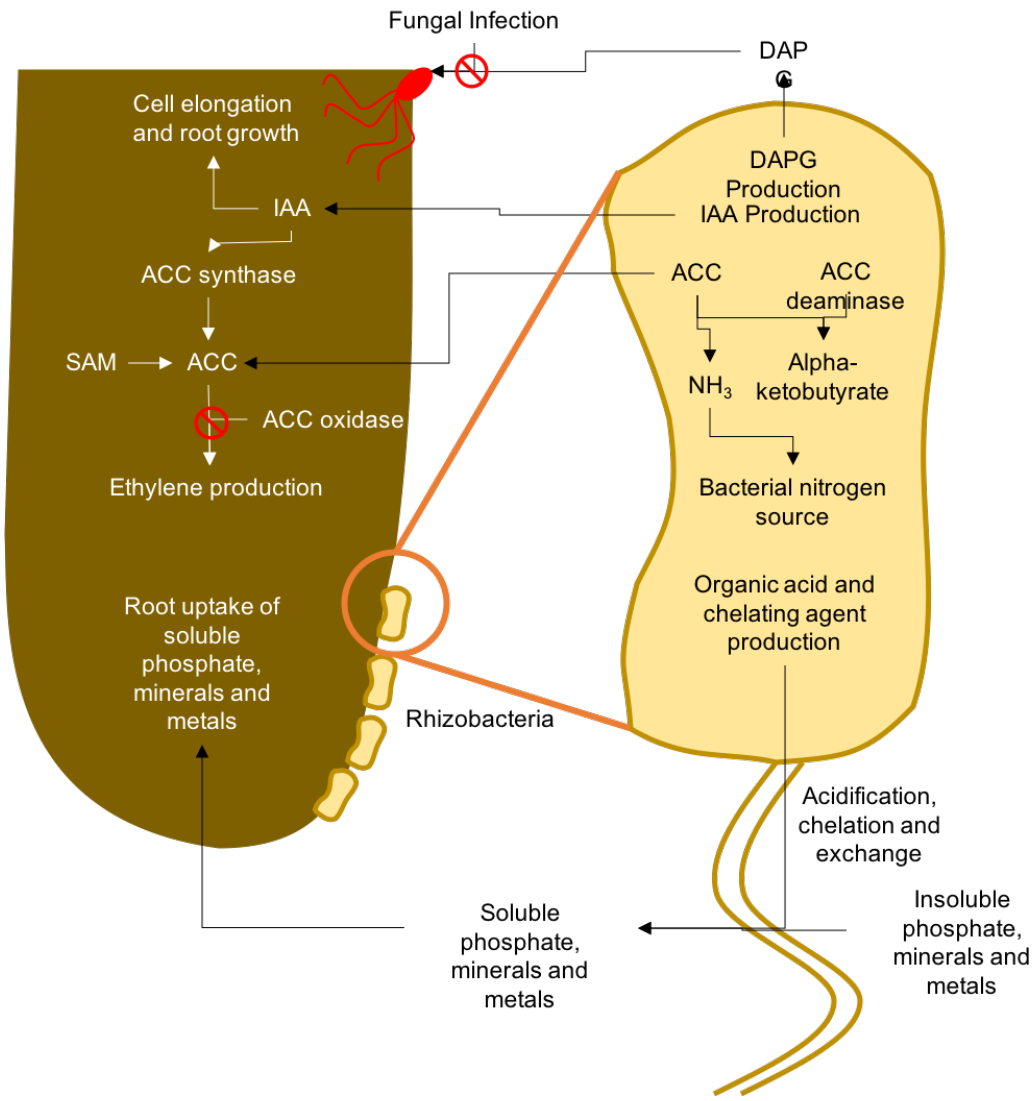


Figure 1-10 – Overview of the several plant growth promoting phenotypes of rhizobacteria and their mechanism with plant root metabolism. ACC = 1-aminocyclopropane-1-carboxylic acid. IAA = Indole-3-acetic acid. DAPG - 2,4-diacetylphloroglucanol. SAM - S-adenosy-L-methionine.

1.5.4 Synthetic Phenotypic Switches

Despite the clear advantage of having ammonia-secreting and other PGP rhizobia in association with target plants, all of the significant plant growth results were achieved in sterile growth environments. This was because when any of the rhizobacterial chassis were altered from their wild-type phenotypes, they become uncompetitive in complex communities (such as soil) and do not survive. To combat this, I propose that a genetic 'toggle switch' could be used to switch between wild-type and ammonia-secreting phenotypes. The switch would be triggered by key plant-specific signals (either natural or engineered). Hence, only when the rhizobia are in close proximity to a target plant, or indeed have colonised intra or extracellularly, will the ammonia-secretion phenotype switch occur. In the absence of a plant signal, the bacteria are able to utilise its fixed nitrogen for growth and this will facilitate the maintenance of a competitive population in the soil microbiome.

Genetic toggle switches have been successfully made in *E. coli* using transcriptional regulators engineered in such a way as to turn on or off depending on input molecules[99] (Figure 1-11). More complex versions of binary systems can cause both toggle switch and oscillatory behaviours in the expression of genes depending on various combinations of inducer molecules[100] (Figure 1-11). By using orthogonal transcriptional regulators and their cognate promoters, it is possible to make genetic switches function independently of the hosts genetic

machinery and regulatory apparatus[101]. Thus, the control of any functional gene can be regulated and expressed upon addition of a signal in the form of an inducer molecule, especially in the rhizosphere where plants exude many unique compounds such as flavonoids. Previous studies that combine synthetic biology techniques and protein engineering in rhizobacteria have successfully engineered novel chimeric transcription factors that integrate with the host bacteria's regulatory network[102]. This work is significant as it allows for the integration of synthetic regulatory networks with the endogenous networks of the host chassis.

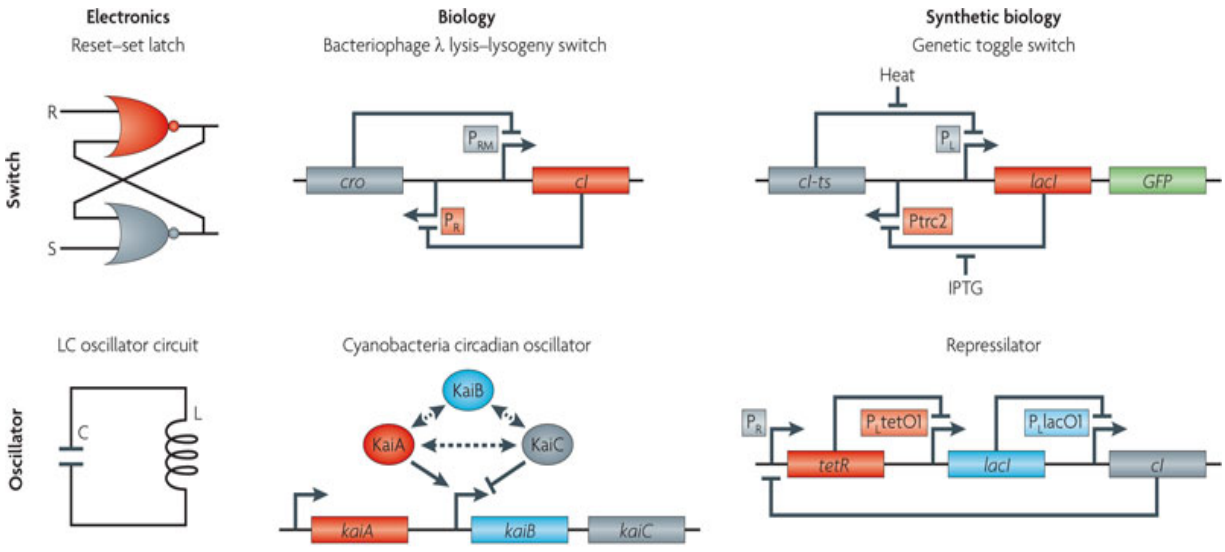


Figure 1-11 - Models of both toggle switches and oscillators (that can exhibit switch-like behaviour). The circuit diagram indicates the engineering representation of both a reset-set latch (toggle switch) and oscillator. Two examples from biology are the lambda phage lysis-lysogeny switch and the cyanobacteria circadian oscillator that controls the cells respiratory phenotype. Based on these biological circuits, representations of a functional toggle switch and oscillator driven by transcriptional repressors (repressilator) are given. Adapted from Khalil *et al.,* 2010.

1.6 Specific Plant Signalling Molecule

1.6.1 Plant – Rhizosphere Chemical Signals

Plants and their rhizosphere-specific microbiome communicate through the release and subsequent detection of primary (amino acids & carbohydrates) and secondary (alkaloids & terpenoids) metabolites. This two-way communication between plants and their rhizosphere microbiome plays an important role in shaping the structure of the rhizosphere and mediating the interactions between plants and their resident rhizobacteria[103,104]. Rhizobacteria directly impact the growth, nutrition and defence of associated plants[105,106].

By harnessing the potential chemical signalling pathways between rhizobacteria and plants, it is possible to control the expression of traits that would otherwise be detrimental to the rhizobacteria chassis (1.5.4 above). The ideal signal is orthogonal to other rhizosphere signals, minimising the cross-talk with existing integrations and signalling networks. By making the signal orthogonal, only engineered microbes can respond to the signal and not other naturally occurring rhizobacteria. Dr Barney Geddes of the Poole lab group has been developing an orthogonal plant signal that could be used as a unique plant to rhizobacteria signal. The chosen trans-kingdom signal is based on a group of compounds called rhizopines.

1.6.1.1 Rhizopine: A unique and specific plant signalling molecule

Rhizopine compounds (scyllo-inosamine (SIA) and 3-O-methyl-scyllo-inosamine (3-O-MSI)) are biosynthesised by some rhizobia in legume nodules during nitrogen fixing symbiosis[107]. Rhizopine biosynthesis is rare and found in only a small portion of two species of rhizobia: *Sinorhizobium meliloti* and *Rhizobium leguminosarum* bv. viciae[108]. Rhizopines are therefore orthogonal to the vast majority of plant rhizospheres, especially those of cereal-rhizobacteria models. The Rhizobia that are able to synthesise rhizopines are also able to utilise them as a sole carbon and nitrogen source and genetic loci responsible for both the anabolism and catabolism of rhizopines have been identified in *S. meliloti* L5-30[107,109,110]. There is no consensus as to the natural purpose of rhizopine biosynthesis, it has been hypothesised that they are produced to nourish sibling rhizobacteria in the rhizosphere[109,110]. If this is indeed true, rhizopine molecules may be readily exuded by engineered cereal plants into their associative rhizosphere.

1.7 Aims and Objectives

To date there have been no significant efforts to fully characterise rhizobacteria as synthetic biology chassis. Studies have used DNA parts from other unrelated strains to perform some novel applications. This work will characterise fully a suite of DNA parts for use across a broad variety of host rhizobacteria (Chapter 3). The two rhizobacteria *A. caulinodans* OR571 and *A. brasilense* FP2 will be used as exemplar rhizobacteria throughout this work. The DNA parts will then be used in tandem with transcriptional repressors to construct logical genetic circuitry in the two key rhizobacterial species (Chapter 4). By harnessing natural plant growth promotion phenotypes from other soil bacteria, the construction of novel PGP circuitry will be attempted (Chapter 5). The main aim of which is exhibiting measurable plant growth promotion of the target crop plant *Hordeum vulgare* (Barley) in a competitive soil environment. Finally, the design, construction and implementation of a CRISPRi system will be demonstrated in a rhizobacteria species for the first time (Chapter 6). The main aim of this chapter is to demonstrate endogenous gene regulation, without genome manipulation of any kind, and produce a plant growth promotion phenotype; ammonia excretion.

2 General Materials and Methods

N.B Specific materials and methods are provided in the subsequent chapters.

2.1 Strains Used in this Study

Table 2-1 – Bacterial strains used in this thesis

Strain	Description	Reference
DH5 alpha	F- deoR endA1 recA1 relA1 gyrA96 hsdR17(rkmk+) supE44 thi-1 - phoA Δ(lacZYA-argF)U169 Φ80lacZΔM15 λ	Bioline
S17-1	Conjugative donor strain; recA pro hsdR RP4-2-Tc::Mu-Km::Tn7 integrated into the chromosome	ATCC47055
ORS571	*Azorhizobium caulinodan*s wild-type strain	Dreyfus and Dommergues, 1981
FP2	*Azospirillum brasilense* FP2 wild-type strain	Tarrand *et al.,* 1978
3841	*Rhizobium leguminosarum* bv. viciae; Str^r derivative of strain 300	Johnston and Beringer, 1975
L530	*Sinorhizobium meliloti* L530	De Lajudie *et al*. 1994
IRBG74	*Rhizobium* (Agrobacterium) sp.	Tan *et al.,* 2001

2.2 Plasmids Used in this Study

Table 2-2 – Plasmids used in this thesis

Name	Description	Reference
pOGG021	pL1V-F1; Amp^r	Weber *et al.,* 2011
pOGG022	pL1M-ELB-1; Amp^r	Weber *et al.,* 2011
pOGG024	pL1V-Lv1-Gent-pBBR1-ELT3; Gent^r	Weber *et al.,* 2011
pOGG026	pL1V-Lv1-Neo-Rk2-Par-ELT4; Kan/Neo^r	Weber *et al.,* 2011
pOGG027	pL2V-Lv2-Neo-pBBR-ELT3; Kan/Neo^r	Weber *et al.,* 2011
pOGG031	pL0M-PU-Plac; Amp^r	Weber *et al.,* 2011
pOGG036	pL0M-S-sfGFP; Amp^r	Weber *et al.,* 2011
pOGG037	pL0M-SC-sfGFP; Amp^r	Weber *et al.,* 2011
pOGG038	pL0M-C2-sfGFP; Amp^r	Weber *et al.,* 2011
pOGG039	pL0M-T-T7; Amp^r	Weber *et al.,* 2011
pOGG054	pL1V-F2; Amp^r	Weber *et al.,* 2011
pOGG055	pL1V-F3; Amp^r	Weber *et al.,* 2011
pOGG056	pL1M-ELB-2; Amp^r	Weber *et al.,* 2011
pOGG057	pL1M-ELB-3; Amp^r	Weber *et al.,* 2011
pOGG058	pL1V-R1; Amp^r	Weber *et al.,* 2011
pOGG059	pL1V-R2; Amp^r	Weber *et al.,* 2011
pOGG060	pL1V-R3; Amp^r	Weber *et al.,* 2011
pOGG063	pL0M-SC-mCherry; Amp^r	Weber *et al.,* 2011

pOGG064	pL0V-SC1; Amp^r	Weber *et al.,* 2011
pOGG065	pL0V-C1; Amp^r	Weber *et al.,* 2011
pOGG066	pL0V-C2; Amp^r	Weber *et al.,* 2011
pOGG067	pL0V-U; Amp^r	Weber *et al.,* 2011
pOGG068	pL0V-PU; Amp^r	Weber *et al.,* 2011
pOGG069	pL0V-S; Amp^r	Weber *et al.,* 2011
pOGG070	pL0V-PUSCT; Amp^r	Weber *et al.,* 2011
pOGG071	pL0V-T; Amp^r	Weber *et al.,* 2011
pOGG072	pL0V-SC; Amp^r	Weber *et al.,* 2011
pOGG073	pL0V-P; Amp^r	Weber *et al.,* 2011
pOGG074	pL0V-C; Amp^r	Weber *et al.,* 2011
pOGG096	PL2V-Lv2-Neo-RK2-par-ELT4; Kan/Neo^r	Weber *et al.,* 2011
pOGG097	pL0M-PU-PmocR; Amp^r	Geddes, 2018 (Manuscript in preparation)
pOGG098	pL0M-PU-PmocB; Amp^r	Geddes, 2018 (Manuscript in preparation)
pOGG109	pL0M-SC-mocR in pUC57-Kan; Kan/Neo^r	Geddes, 2018 (Manuscript in preparation)
pOGG110	pL0M-SC-mocBintBC in pUC57-Kan; Kan/Neo^r	Geddes, 2018 (Manuscript in preparation)
pOGG111	pL0M-SC-AtiolG in pUC57-Kan; Kan/Neo^r	Geddes, 2018 (Manuscript in preparation)
pOGG112	pL0M-SC-AtmosB in pUC57-Kan; Kan/Neo^r	Geddes, 2018 (Manuscript in preparation)
pOGG113	pL0M-PU-empty (use as No Promoter); Amp^r	Mendoza-Suarez, 2018 (Manuscript in preparation)

pOGG120	pL0M-P-J23104; Amp^r	This work
pOGG121	pL0M-P-J23106; Amp^r	This work
pOGG122	pL0M-P-J23115; Amp^r	This work
pOGG123	pL0M-SsrA Fast-C2; Amp^r	This work
pOGG124	pL0M-SsrA Moderate-C2; Amp^r	This work
pOGG125	pL0M-SsrA Very fast-C2; Amp^r	This work
pOGG126	pL0M-SC-sYFP2; Amp^r	This work
pOGG127	pL0M-SC-eCFP; Amp^r	This work
pOGG128	pL0M-SC-mRuby3; Amp^r	This work
pOGG129	pL0M-SC-mTagBFP; Amp^r	This work
pOGG130	pL0M-SC-YFP; Amp^r	This work
pOGG131	pL0M-PU-Plux; Amp^r	This work
pOGG132	pL0M-PU-Pbad; Amp^r	This work
pOGG134	pL0M-P-Ptet; Amp^r	This work
pOGG135	pL0M-rPU-PtetA-tetR (tetracycline inducible); Amp^r	This work
pOGG136	pL0M-PU-Pphl; Amp^r	This work
pOGG137	pL0M-rPU-PphlA-phlF (DAPG inducible); Amp^r	This work
pOGG138	pL0M-PU-PnodA; Amp^r	This work
pOGG139	pL0M-rPU-PnodA-nodD (flavonoid inducible); Amp^r	This work
pOGG140	pL0M-PU-Plac; Amp^r	This work
pOGG142	pL0M-PU-PtauA; Amp^r	This work
pOGG143	pL0M-U-Rstd; Amp^r	This work

pOGG144	pL0M-U-R1; Amp^r	This work
pOGG145	pL0M-U-R13; Amp^r	This work
pOGG146	pL0M-U-R22; Amp^r	This work
pOGG147	pL0M-U-R31; Amp^r	This work
pOGG149	pL0M-SC1-sfGFP; Amp^r	This work
pOGG151	pL0M-SC-T7RNAP* in pUC57-Kan; Kan/Neo^r	This work
pOGG152	pL0M-P-PT7wt; Amp^r	This work
pOGG153	pL0M-P-PT71; Amp^r	This work
pOGG154	pL0M-P-PT72; Amp^r	This work
pOGG155	pL0M-P-PT73; Amp^r	This work
pOGG156	pL0M-T-DT15 short terminator; Amp^r	This work
pOGG157	pL0M-T-DT16 short terminator; Amp^r	This work
pOGG158	pL0M-T-DT25 short terminator; Amp^r	This work
pOGG159	pL0M-T-F1 short terminator; Amp^r	This work
pOGG160	pL0M-T-F2 terminator; Amp^r	This work
pOGG161	pL0M-T-F4 terminator; Amp^r	This work
pOGG162	pL0M-T-F6 short terminator; Amp^r	This work
pOGG163	pL0M-SC-araC; Amp^r	This work
pOGG164	pL0M-SC-clR; Amp^r	This work
pOGG165	pL0M-SC-lacI; Amp^r	This work
pOGG166	pL0M-SC-luxR; Amp^r	This work
pOGG167	pL0M-SC-phlF; Amp^r	This work

pOGG168	pL0M-SC-tauR; Amp^r	This work
pOGG169	pL0M-SC-tetR; Amp^r	This work
pOGG179	pL0M-P-J23100; Amp^r	Jorrin, 2018 (Manuscript in preparation)
pOGG180	pL0M-P-J23101; Amp^r	Jorrin, 2018 (Manuscript in preparation)
pOGG181	pL0M-P-J23111; Amp^r	Jorrin, 2018 (Manuscript in preparation)
pOGG187	pL1V-F4; Amp^r	Weber *et al.,* 2011
pOGG188	pL1M-ELE4; Spec^r	Weber *et al.,* 2011
pOGG189	pL1M-ELE3; Spec^r	Weber *et al.,* 2011
pOGG190	pL0M-P-Pcl; Amp^r	This work
pOGG191	pL2V-Lv2-Tet-RK2-par-ELT4; Tet^r	Weber *et al.,* 2011
pOGG192	pL0M-SC-IaaM; Amp^r	This work
pOGG193	pL0M-SC-IaaH; Amp^r	This work
pOGG194	pL0M-PU-PmocC; Amp^r	Geddes, 2018 (Manuscript in preparation)
pOGG195	pL0M-PU-PmocD; Amp^r	Geddes, 2018 (Manuscript in preparation)
pOGG196	pL0M-SC-mocRshort; Amp^r	Geddes, 2018 (Manuscript in preparation)
pOGG197	pL0M-SC-mocB; Amp^r	Geddes, 2018 (Manuscript in preparation)
pOGG199	pL0M-SC-phlABC; Amp^r	This work
pOGG200	pL0M-SC-phlD; Amp^r	This work
pOGG201	pL0M-SC-phlE; Amp^r	This work

pOGG208	pL1M-F1 J23106 RBSstd phlABC; Amp^r	This work
pOGG209	pL1M-F2 PmocB phlD DT15; Amp^r	This work
pOGG210	pL1M-F1 PmocB iaaH F6; Amp^r	This work
pOGG211	pL1M-R2 PmocD iaaM DT16; Amp^r	This work
pOGG212	pL1M-F1 PmocB phlABC F6; Amp^r	This work
pOGG213	pL1M-F3 PmocD phlE F2; Amp^r	This work
pOGG214	pL1M-F3 J23106 RBSstd phlE F2; Amp^r	This work
pOGG227	pL0M-PU-PglnA (FP2); Amp^r	This work
pOGG228	pL0M-PU-PglnB1&2 (FP2); Amp^r	This work
pOGG229	pL0M-PU-PglnB1&2 GlnB PglnA (FP2); Amp^r	This work
pOGG230	pL1V-dSpCas9 sgRNA expression plasmid with GG lacZ cassette; Gent^r	This work
pOGG231	pL1V-dSpCas9 Expression Plasmid with LacZ sgRNA cassette; Kan/Neo^r	This work
pOGG232	pL1V-dSpCas9 EQR Expression Plasmid with LacZ sgRNA cassette; Kan/Neo^r	This work
pOGG233	dSpCas9 VRER Expression Plasmid with LacZ sgRNA cassette; Kan/Neo^r	This work
pOGG234	PL1M-F1 PL0M-J23106 PL0M-R1-U - IaaM PL0M-SC - F2 terminator PL0M-T; Amp^r	This work
pOGG235	PL1M-R2 PL0M-J23115-P - PL0M-Rstd-U - IaaH PL0M-SC - DT15 short terminator PL0M-T; Amp^r	This work
pOGG236	PL1M-F1 IaaM PL0M-SC - F2 terminator PL0M-T - pOGG021 (reversed) into PphlF-rPU-PL0M; Amp^r	This work
pOGG237	PL1M-R2 Pphl-PU-PL0M - IaaH PL0M-SC - DT15 short terminator PL0M-T into pOGG059; Amp^r	This work
pOGG238	PL1M-F1 PL0M-J23106 PL0M-Rstd-U - GltA PL0V-SC - F1 short terminator PL0M-T	This work
pOGG239	PL1M-R2 PL0M-J23104-P - PL0M-R1-U - CitC PL0V-SC - DT16 Short terminator; PL0M-T; Amp^r	This work

pOGG240	PmocB LacI DT16S Rhizopine Inducible LacI F1 module; Amp^r	This work
pOGG241	PmocB PhlF DT16S Rhizopine Inducible PhlF F1 module; Amp^r	This work
pOGG242	PmocB TetR DT16S Rhizopine Inducible TetR F1 module; Amp^r	This work
pOGG243	Plac RBSstd sfGFP F1 module; Amp^r	This work
pOGG244	Pphl RBSstd sfGFP F1 module; Amp^r	This work
pOGG245	Ptet RBSstd sfGFP F1 module; Amp^r	This work
pOGG246	pL0M-SC-CitC; Amp^r	This work
pOGG247	pL0M-SC-GltA; Amp^r	This work
pAJM.847	PhlF Pphl RBSstd sfGFP pBBR1; Tet^r	Meyer *et al.,* 2018
pAN4137	Ptet RBSstd dSpCas9; Tet^r	Nielsen *et al.,* 2014
pAN4181	Ptet RBSstd dSpCas9VRER; Tet^r	Nielsen *et al.,* 2014
pAN4183	Ptet RBSstd dSpCas9EQR; Tet^r	Nielsen *et al.,* 2014
pAN4244	dSpCas9-2NToperator-J23101-YFP3snp-Pp-Sp-sgRNA-2NT; Spec^r	Nielsen *et al.,* 2014
pAN4245	dSpCas9VRER-2NToperator-J23101-YFP3snp-Pp-Sp-sgRNA-2NT; Spec^r	Nielsen *et al.,* 2014
pAN4249	dSpCas9EQR-2NToperator-J23101-YFP3snp-Pp-Sp-sgRNA-2NT; Spec^r	Nielsen *et al.,* 2014
pOPS0457	Ptet RBSstd sfGFP DT16S pBBR1; Gent^r	This work
pOPS0589	Pphl RBSstd sfGFP DT16S pBBR1; Gent^r	This work
pOPS0590	Plac RBSstd sfGFP DT16S pBBR1; Gent^r	This work
pOPS0666	Constitutive J23016 RBSstd IaaM & J23115 RBS1 IaaH expression plasmid; Gent^r	This work
pOPS0667	Constitutive J23016 RBSstd IaaM & J23115 RBS1 IaaH expression plasmid with *par* gene cluster; Kan/Neo^r	This work
pOPS0910	Constitutive DAPG biosynthesis plasmid with *par* gene cluster for stability; Kan/Neo^r	This work

pOPS1011	Plac RBSstd sfGFP DT16S pBBR1; Gent^r	This work
pOPS1150	aTc inducible dSpCas9wt cassette with *par* gene cluster for stability; Kan/Neo^r	This work
pOPS1151	aTc inducible dSpCas9EQR cassette with *par* gene cluster for stability; Kan/Neo^r	This work
pOPS1152	aTc inducible dSpCas9VRER cassette with *par* gene cluster for stability; Kan/Neo^r	This work
pOPS1153	dSpCas9wt YFP reporter plasmid. Pphl sgRNA expression cassette; Gent^r	This work
pOPS1154	dSpCas9EQR YFP reporter plasmid. Pphl sgRNA expression cassette; Gent^r	This work
pOPS1155	dSpCas9VRER YFP reporter plasmid. Pphl sgRNA expression cassette; Gent^r	This work
pOPS1156	dSpCas9wt YFP reporter plasmid. J23100 sgRNA expression cassette; Gent^r	This work
pOPS1157	dSpCas9EQR YFP reporter plasmid. J23100 sgRNA expression cassette; Gent^r	This work
pOPS1158	dSpCas9VRER YFP reporter plasmid. J23100 sgRNA expression cassette; Gent^r	This work
pOPS1159	FP2 PglnA NT1 sgRNA expression plasmid; Gent^r	This work
pOPS1160	FP2 PglnA NT2 sgRNA expression plasmid; Gent^r	This work
pOPS1161	FP2 PglnA NT3 sgRNA expression plasmid; Gent^r	This work
pOPS1162	FP2 PglnA T1 sgRNA expression plasmid; Gent^r	This work
pOPS1163	FP2 PglnA T2 sgRNA expression plasmid; Gent^r	This work
pOPS1164	FP2 PglnA T3 sgRNA expression plasmid; Gent^r	This work
pOPS1165	Constitutive J23016 RBSstd IaaM & J23115 RBS1 IaaH expression plasmid; Gent^r	This work
pOPS1166	DAPG inducible IaaM & IaaH plasmid; Gent^r	This work
pOPS1167	J23104 RBSstd sfGFP promoter analysis plasmid; Gent^r	This work
pOPS1168	J23106 RBSstd sfGFP promoter analysis plasmid; Gent^r	This work
pOPS1169	J23115 RBSstd sfGFP promoter analysis plasmid; Gent^r	This work
pOPS1170	J23104 RBSstd sfGFP RBS analysis plasmid; Gent^r	This work

pOPS1171	J23104 RBS1 sfGFP RBS analysis plasmid; Gent^r	This work
pOPS1172	J23104 RBS13 sfGFP RBS analysis plasmid; Gent^r	This work
pOPS1173	J23104 RBS22 sfGFP RBS analysis plasmid; Gent^r	This work
pOPS1174	J23104 RBS31 sfGFP RBS analysis plasmid; Gent^r	This work
pOPS1175	J23106 RBSstd TetR DT16S pBBR1; Gent^r	This work
pOPS1176	J23106 RBSstd PhlF DT16S pBBR1; Gent^r	This work
pOPS1177	J23106 RBSstd LacI DT16S pBBR1; Gent^r	This work
pOPS1178	J23106 RBSstd ClR DT16S pBBR1; Gent^r	This work
pOPS1179	Rhizopine inducible LacI with Plac RBSstd sfGFP cassette (NOT gate) plasmid with *par* gene cluster; Kan/Neo^r	This work
pOPS1180	Rhizopine inducible PhlF with Plac RBSstd sfGFP cassette (NOT gate) plasmid with *par* gene cluster; Kan/Neo^r	This work
pOPS1181	Rhizopine inducible TetR with Plac RBSstd sfGFP cassette (NOT gate) plasmid with *par* gene cluster; Kan/Neo^r	This work
pOPS1182	J23016 RBSstd GltA, J23104 RBS1 CitC; Gent^r	This work
pOPS1183	J23016 RBSstd GltA, J23104 RBS1 CitC with *par* gene cluster; Kan/Neo^r	This work
pOPS1184	FP2 PglnA sfGFP ssrA Fast DegTag promoter analysis plasmid; Gent^r	This work
pOPS1185	FP2 PglnB1&2 sfGFP ssrA Fast DegTag promoter analysis plasmid; Gent^r	This work
pOPS1186	FP2 PglnB1&2 GlnB PglnA sfGFP ssrA Fast DegTag promoter analysis plasmid; Gent^r	This work

2.3 Primers Use in this Study

Table 2-3 – Primers used in this thesis

Primer	Sequence	Description
M13 rev (-29)	AAGAACATCGATTTTCCATGGC AG	Sequencing/mapping primer; pJET1.2
M13 uni (-21)	GACTCACTATAGGGAGAGCGG C	Sequencing/mapping primer; pJET1.2
oxp0715	CAAACCGCCTCTCCCCG	pL0V screening FWD
oxp0716	AAAGTGCCACCTGACGTCT	pL0V screening REV
oxp0824	ATGGTCCTGTTCTGCTGCC	sfGFP FWD Out
oxp0825	TACCAGAGTCGGCCAAGGT	sfGFP REV Out
oxp1011	TTTTGAAGACAAGGAGTAGCGGATCC TACCTGAC	Pbad-PU-Primer FWD
oxp1012	TTTTGAAGACAACATTATGTATATCTC CTTCCCCGG	Pbad-PU-Primer REV
oxp1013	TTTTGAAGACAAGGAGGTTGACTTGT GAGCGGATAA	Plac-PU-Primer FWD
oxp1014	TTTTGAAGACAACATTGGTTAATTTCT CCTCTTTAATTCTAGATG	Plac-PU-Primer REV
oxp1015	TTTTGAAGACAAGGAGTACCTGTAGG ATCGTACAGGTTTAC	Plux-PU-Primer FWD
oxp1016	TTTTGAAGACAACATTGGTTAATTTCT CTCTTTAATTCTAGATTTATTCGA	Plux-PU-Primer REV
oxp1017	TTTTGAAGACAAGGAGTTGACATGAT ACGAAACGTACCGTAT	Pphl-PU-Primer FWD
oxp1018	TTTTGAAGACAACATTGGTTAATTTCT CCTCTTTAATTCTAGACT	Pphl-PU-Primer REV
oxp1019	TTTTGAAGACAAAATGCAATATGGAC AATTGGTTTCTTCTC	AraC-SC-Primer FWD
oxp1020	TTTTGAAGACAAAAGCTTATGACAACT TGACGGCTAC	AraC-SC-Primer REV
oxp1023	TTTTGAAGACAAAATGGCACGTACCCC GAGCC	PhlF-SC-Primer FWD

oxp1024	TTTTGAAGACAAAAGCTTAACGCTGTGTACCCGGAC	PhlF-SC-Primer REV
oxp1025	TTTTGAAGACAAAATGTCCAGATTAGATAAAAGTAAAGTGATTAAC	TetR-SC-FWD
oxp1026	TTTTGAAGACAAAAGCTTAGGACCCACTTTCACATTTAAG	TetR-SC-REV
oxp1029	TTTTGAAGACAAGGAGTTTATCAAAAAGAGTCGGTGGAATCC	Ptet PU FWD
oxp1030	TTTTGAAGACAACATTGGTTAATTTCTCCTCTTTAATTCTAGAGTGC	Ptet PU REV
oxp1031	TTTTGAAGACAAGGAGTTAGGACCCACTTTCACATTTAAGTTG	Ptet rPU FWD A
oxp1032	TGAAGACGTTTTCACCTCGAGAAAATTTATCAAAAAGA	Ptet rPU FWD B
oxp1033	TGAAGACGTGAAAACGAAAGGGCCTCG	Ptet rPU REV A
oxp1034	TTTTGAAGACAAGGAGTTAACGCTGTGTACCCGGAC	Pphl rPU FWD A
oxp1035	TGAAGACGTGTTCACCTCGAGAAAATTTATCAAAAAG	Pphl rPU FWD B
oxp1036	TGAAGACGTGAACACGAAAGGGCCTC	Pphl rPU REV B
oxp1037	TTTTGAAGACAACATTGGTTAATTTCTCCTCTTTAATTCTAGACT	Pphl-PU-Primer REV
oxp1064	TTTTGAAGACAAGGAGTGTTCACCTCGCCCAACTG	Ptau-PU-FWD
oxp1065	TTTTGAAGACAACATTGGTTAATTTCTCCTCTTTAATTTTTCTCCTCT	Ptau-PU-REV
oxp1066	TTTTGAAGACAAGGAGTTATGACAACTTGACGGCTACATCA	Pbad-rPU-FWD
oxp1067	TTTTGAAGACAAGGAGGGGTTGAATATCCATTCCATAGATG	Pnod-PU-FWD
oxp1068	TTTTGAAGACAACATTGGTTAATTTCTCCTCTTTAATTTTTCTCC	Pnod-PU- REV
oxp1286	TTTTGAAGACAAAATGCGTAAAGGCGAAGAGCTGTTC	sfGFP SC1 FWD
oxp1287	TTTTGAAGACAACACCTTTGTACAGTTCATCCATACCATGC	sfGFP SC1 REV
oxp1288	TTTTGAAGACAAAATGAGCGAACTGATCAAAGAGAACATG	mTagBFP SC FWD
oxp1289	TTTTGAAGACAAAAGCTTATTAATTCAGTTTATGACCCAGCTTGC	mTagBFP SC REV
oxp1300	TTTTGAAGACAAGGAGTCACTGCCCGCTTTCCAG	Plac-rPU-FWDA

	TTTTGAAGACAAACCAAAGACAGCTC	
oxp1301	ATGTTATATCCC	Plac rPU A REV
	TTTTGAAGACAATGGTATCGTCGTATC	
oxp1302	CCACTACC	Plac rPU B FWD
	TTTTGAAGACAACATAAAGACGGTAC	
oxp1303	GCGACTG	Plac rPU B REV
	TTTTGAAGACAATATGGGAGAAAATA	
oxp1304	ATACTGTTGATGGGT	Plac rPU C FWD
	TTTTGAAGACAAGGTAAAGACGAAAG	
oxp1305	GGCCTCG	Plac rPU C REV
	TTTTGAAGACAATACCTCGAGAAAATT	
oxp1306	TATCAAAAAGAGTG	Plac rPU D FWD
	TTTTGAAGACAACATTGGTTAATTTCT	
oxp1307	CCTCTTTAATTCTAGATGTGT	Plac rPU/PU D REV
	TTTTGAAGACAAGGAGTTACGACGTA	
oxp1308	CAACTTTCCATCTCA	Pnod rPU A FWD
	TTTTGAAGACAAGAAGCCTTCCGTAG	
oxp1309	AGCAATG	Pnod rPU A REV
	TTTTGAAGACAACTTCAGACCACGTCC	
oxp1310	GAACTTA	Pnod rPU B FWD
	TTTTGAAGACAACTGCTCTTCTTCGAC	
oxp1311	AAGATAATACTG	Pnod rPU B REV
	TTTTGAAGACAAGCAGTGCCATGAAG	
oxp1312	TCTGAGA	Pnod rPU C FWD
	TTTTGAAGACAACATTGGTTAATTTCT	
oxp1313	CCTCTTTAATTTTTCTCC	Pnod rPU C REV
	TTTTGAAGACAAAATGGCGGGCTCGA	
oxp1314	CCGTC	TauR SC FWDA
	TTTTGAAGACAAGTAAAAGACCCGGC	
oxp1315	CGCA	TauR SC REV A
	TTTTGAAGACAATTACATCGGCAGTTT	
oxp1316	CTCCAAGT	TauR SC FWDB
	TTTTGAAGACAAATCAAGACCTTCCGG	
oxp1317	CCC	TauR SC REV B
	TTTTGAAGACAATGATGCGGACTTGCT	
oxp1318	GATGAA	TauR SC FWDC
	TTTTGAAGACAAAAGCCTAGATATCG	
oxp1319	ATACCCCGCTCC	TauR SC REV C
	TTTTGAAGACAAAATGAAACCAGTAA	
oxp1320	CGTTATACGATGTCG	LacI SC FWD
	TTTTGAAGACAAAAGCTCACTGCCCGC	
oxp1321	TTTCCAG	LacI SC FWD
	TTTTGAAGACAAAATGGTTAGCAAGG	
oxp1330	GCGAAGAACTTTT	sYFP2 SC FWD

	TTTTGAAGACAAAAGCTTATTATTTAT	
oxp1331	ACAGCTCATCCATACCCAG	sYFP2 SC REV
	TTTTGAAGACAAAATGGTGAGCAAGG	
oxp1332	GCGAGGAG	eCFP SC FWD
	TTTTGAAGACAAAAGCTTATTACTTGT	
oxp1333	ACAGCTCGTCCATG	eCFP SC REV
	TTTTGAAGACAAAATGAGCACAAAAA	
oxp1336	AGAAACCATTAACACAAG	clR FWD
	TTTTGAAGACAAAAGCTTATTAGCCAA	clR Rev with double stop
oxp1337	ACGTCTCTTCAGGC	codon overhang
	TTTTGAAGACAAAATGTTTGGACCGG	IaaM (and IAA operon)
oxp1338	ATTTCCCGTTT	FWD
	TTTTGAAGACAAAGATTGCATGTAAC	
oxp1339	GCGTGCGA	IaaM A REV
	TTTTGAAGACAAATCTTCTAAAACTTTT	
oxp1340	GTGATGGTTGATCGG	IaaM B FWD
	TTTTGAAGACAAAAGCTTATTAGCCTT	IaaH (and IAA operon)
oxp1341	TTAACACTTTTTCGATCAAA	REV
	TTTTGAAGACAAAATGCGCGAAATGA	
oxp1342	TTACGCTTGAATC	IaaH FWD
	TTTTGAAGACAAAAGCTTATTAGTCCC	
oxp1343	CCAGCGCG	IaaM REV
	TTTTGAAGACAAAATGGTGAGCAAGG	mCherry SC1 FWD
oxp1352	GCGAGGAG	Primer
	TTTTGAAGACAACACCTTACTTGTACA	
oxp1353	GCTCGTCCATGC	mCherry SC1 REV Primer
		Golden Gate Level 0 Module FWD Sequencing
oxp1362	TTTGCTGGCCTTTTGCTCAC	Primer
	TTTTGAAGACAAAATGGGTTTCATGG	GS (*glnA*) *A. brasilense*
oxp1473	AGGACTTCATCAAGAAGAACCGCATC	FP2 SC FWD
	TTTTGAAGACAAAAGCTCAAACGTTCA	GS (*glnA*) *A. brasilense*
oxp1474	GCAGAAGATTCTC	FP2 SC REV
	CACTCTGTGGTCTCAGGAGGCTTTACT	
oxp1920	CCACCGTTGGCT	dSpCas9 VRER FWD
	CACTCTGTGGTCTCAAGCGAAAATATT	
oxp1921	TGCTCATGAGCCCGAAG	dSpCas9 VRER REV
	CACTCTGTGGTCTCAGGAGGGACCAA	dSpCas9VRER YFP
oxp1922	AACGAAAAAAGGCC	Reporter FWD pAN4245
	CACTCTGTGGTCTCAAGCGAAAAAAG	dSpCas9VRER YFP
oxp1923	CACCGACTCGGTG	Reporter REV pAN4245
	TTTTGAAGACAATCCCAGAGCGTTCAC	FWD primer for
oxp1951	CGACAAACA	amplifying cognate

		promoter and sfGFP reporter cassette
oxp1952	TTTTGAAGACAAGCAATCAACAGGAG TCCAAGAGCG	REV primer for amplifying cognate promoter and sfGFP reporter cassette
oxp2040	GGGGTTTTGAAGACAAAATGACCAAC ATGACCCAGGCTTC	*S. typhimurium citC* 2.0 FWD
oxp2041	CCCCTTTTGAAGACAAAAGCTTACACC ATCATGCTGAACACGATGC	*S. typhimurium citC* 2.0 REV
oxp2074	TGTGAAGACAATGCCGAATTC	*phlABC* FWD for operon
oxp2077	GCAGAAGACAATAGTAGCGAAGGA	*phlD* REV for operon
oxp2138	TTTTGAAGACAAGGAGCCGTGCGTGG ATGCCC	FP2 PglnA - PglnA PL0V-PU FWD
oxp2139	TTTTGAAGACAACATTATCTCTTTTTCC ACGCTTAAGAGGT	FP2 PglnA - PglnA PL0V-PU REV
oxp2146	GTTTTTCAGCAAGATGGATCCTAGGG ATAACAGGGTAATGTGCATTTCGTCG ACGGC	FP2 *glnA* LF FWD with *SceI* and *BamHI*
oxp2147	GAATTCTCCATCGGACTTCGGTCCA	FP2 *glnA* FP2 LF REV with *EcoRI*
oxp2148	TCCGATGGAGAATTCATCTCTTTTTCC ACGCTTAAGAGGT	FP2 *glnA* FP2 RF FWD with *EcoRI*
oxp2149	ATCTTCTAGAAAGATGGATCCGATGG ATGGGACTCGAACCC	FP2 *glnA* FP2 RF REV + *BamHI*
oxp2193	TTTTGAAGACAAGGAGCATCCGAACG GACAAGGC	PglnBA FP2 PU FWD
oxp2194	TTTTGAAGACAACATTGCGTCTCTCGT GTACGAGG	PglnBA FP2 PU REV
oxp2451	GGAGATCGCCAAGAAGATCA	FP2 *glyA* FWD qPCR
oxp2452	GCTCTTGGCGTAGGTCTTGA	FP2 *glyA* REV qPCR
oxp2453	TCACCGACGAAGAGCTGATG	FP2 *gyrB* FWD qPCR
oxp2454	CTCTCGATCTCGGTCTTGG	FP2 *gyrB* REV qPCR
oxp2455	TCCCGGACGAGCGGAC	FP2 *glnA* FWD for qPCR Set A
oxp2456	GCGGATAATCGGGGTCCTTG	FP2 *glnA* REV for qPCR Set A

oxp2457	AAGAACCGCATCACCGAAGT	FP2 GlnA FWD for qPCR set B
oxp2458	AATTCTCGTCCGAGGTGACC	FP2 GlnA REV for qPCR set B
oxp2459	GACCCCGATTATCCGCTGG	FP2 GlnA FWD for qPCR set C
oxp2460	TGGTTGAAGTTGATCTCGATCT	FP2 GlnA REV for qPCR set C
oxp2461	GAATCCGTACCTCGCCATCG	FP2 GlnA FWD for qPCR set D
oxp2462	CCAGGAATTTTTCGCCCAGC	FP2 GlnA REV for qPCR set D
oxp2581	GGAGGAGAAAGAAGGATCATGAATA AAGTAGGAATTGTGAGCT	DAPG BD Cloning primer for Rhizpine Induction A
oxp2582	GGGTATGGAGAAGGATCTCACTTATC CTCCAGGGTCA	DAPG BD Cloning primer for Rhizpine Induction B
oxp3110	CCCTGGGCTGGATTCTCACCAATAAAA AACGC	sgRNA cassette PasI transfer REV
oxp3111	CCCTGGGTCTAGGGCGGCGGATTTG	sgRNA cassette PasI transfer FWD
pJET1.2 for	TGTAAAACGACGGCCAGT	Sequencing primer for lacZ-containing vectors
pJET1.2 rev	CAGGAAACAGCTATGACC	Sequencing primer for lacZ-containing vectors
pOT forward far	GACCTTTTGAATGACCTTTA	Universial sequencing/mapping primer

2.4 Media and Antibiotics and other Chemicals

2.4.1 Media

E. coli strains were grown at 37°C in Luria Bertani broth (LB) (10g L^{-1} tryptone, 5 g L^{-1} yeast extract, 5 g L^{-1} NaCl). Cultures were shaken at 200 rpm unless otherwise stated. For growth on solid media, agar was added 1.4 (w/v) prior to autoclaving. *A. caulinodans* and *A. brasilense* were grown on tryptone yeast (TY) media (5 g L^{-1} tryptone, 3 g L^{-1} yeast extract, 6 mM $CaCl_2$) (Beringer, 1974). For growth on solid media, agar was added 1.75 (w/v) before autoclaving. When a minimal growth media was required, rhizobacteria were grown in Universal Minimal Salts (UMS). UMS contains 0.5 mM K_2HPO_4, 0.5 g L^{-1} $MgSO_4.7H_2O$, 0.2 g L^{-1} NaCl, 4.19 g L^{-1} MOPS, adjusted to pH 7.0 and autoclaved. Following autoclaving media was supplemented with 1 mL of iron stock solution (12 g L^{-1} $FeSO_4.7H_2O$), 1 mL of calcium stock solution (75 g L^{-1} $CaCl_2.2H_2O$) and 1 mL of a stock solution containing 1 g L^{-1} thiamine hydrochloride, 2 g L^{-1} D-Pantothenic acid calcium salt and 100 mg L^{-1} biotin. For *A. caulinodans* and *A. brasilese* growth in UMS was supplemented with filter-sterilised 20 mM succinic acid (succinate), 10 mM NH_4Cl and 1 µL of 1000x vitamin solution. One litre of vitamin stock contains 0.375 g L^{-1} EDTA-Na_2, 0.16 g L^{-1} $ZnSO_4.7H_2O$, 0.2 g L^{-1} , $NaMoO_4$, 0.25 g L^{-1} H_3BO_3, 0.2 g L^{-1} $MnSO_4.4H_2O$, 0.02 g L^{-1} $CuSO_4.5H_2O$, 1 g L^{-1} $CoCl_2.6H_2O$. *A. caulinodans* was grown at 37°C, with 200 rpm shaking for liquid media. *A.*

caulinodans is auxotrophic for nicotinate, so 30 μM nicotinic acid was added to the growth medium. *A. brasilense* was grown at 30°C, with 200 rpm shaking for liquid media. All bacterial strains were stocked in 15% (v/v) glycerol and snap-frozen in liquid nitrogen for storage at -80°C.

2.4.2 Antibiotics and other Chemicals

Appropriate antibiotics were added to media in concentrations as shown in Table 2-4. *A. caulinodans* ORS571 has chromosomal resistance to nitrofurantoin and streptomycin, the latter was used as a selection antibiotic. *A. brasilense* FP2 has chromosomal resistance to streptomycin and ampicillin, the latter was used as a selection antibiotic. In cases of *E. coli* transformation where beta-galactosidase blue-white screening could be used for screening colonies, X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) was added to LB agar at 40 μg mL^{-1}.

Table 2-4 - Working concentrations of antibiotics used in this work. Units in μg ml^{-1}.

	E. coli	*A. caulinodans* ORS571	*A. brasilense* FP2
Ampicillin	100	100	50
Chloramphenicol	10		
Gentamycin	10	20	20
Kanamycin	20	20	20
Neomycin	20	80	80
Nitrofurantoin		10	10
Spectinomycin	50	100	100
Streptinomycin	25	250	250
Tetracycline	10	5	5

2.4.3 Aerobic Growth

E. coli strains were grown at 37°C in 10 ml Luria Bertani broth (LB) (10 g L^{-1} tryptone, 5 g L^{-1} yeast extract, 5 g L^{-1} NaCl). Cultures were shaken at 200 rpm unless otherwise stated. For growth on solid media, 1.4% (w/v) agar was added at prior to autoclaving [112]. *A. caulinodans* and *A. brasilense* were grown on tryptone yeast (TY) media (5 g L^{-1} tryptone, 3 g L^{-1} yeast extract, 6 mM $CaCl_2$) (Beringer 1974). For growth on solid media, 1.75 % (w/v) agar was added before autoclaving. When a minimal growth media was required, bacteria were grown in

10 ml of Universal Minimal Salts (UMS). UMS contains 0.5 mM $KHPO_4$, 0.5 g L^{-1} $MgSO.7H_2O$, 0.2 g L^{-1} NaCl, 4.19 g L^{-1} MOPS, adjusted to pH 7.0 and autoclaved. Following autoclaving media was supplemented with 1 mL of iron stock solution (12 g L^{-1} $FeSO.7H_2O$), 1 mL of calcium stock solution (75 g L^{-1} $CaCl.2H_2O$) and 1 mL of a stock solution containing 1 g L^{-1} thiamine hydrochloride, 2 g L^{-1} D-Pantothenic acid calcium salt and 100 mg L-biotin. For and growth in UMS was supplemented with filter-sterilised 20 mM succinic acid ($(CH_2)_2(CO_2H)_2$), 20 mM NH_4Cl and 1 µL of 1000x vitamin solution. One litre of vitamin stock contains 0.375 g L^{-1} EDTA-Na_2, 0.16 g L^{-1} $ZnSO.7H_2O$, 0.2 g L^{-1} , $NaMoO_4$, 0.25 g L^{-1} H_3BO_3, 0.2 g L^{-1} $MnSO_4.4H_2O$, 0.02 g L^{-1} $CuSO_4.5H_2O$, 1 g L^{-1} $CoCl_2.6H_2O$. *A. caulinodans* and *A. brasilense* were grown at 37˚C and 30˚C respectively with 200 rpm shaking for liquid media. *A. caulinodans* is auxotrophic for nicotinate, so 30 µM nicotinic acid was added to the growth medium. All bacterial strains were stocked in 15% glycerol (v/v) and snap-frozen in liquid nitrogen for storage at -80°C.

2.4.4 Microaerobic Diazotrophic Growth

A. caulinodans ORS571 and *A. brasilense* FP2 were grown under microaerobic conditions in an atmospheric regulatory cabinet at 3% (v/v) oxygen to facilitate diazotrophic growth. Growth was performed in gas tight universal tubes with rubber gas sampling apertures (VWR, Pennsylvania, USA). Cultures of rhizobacteria were grown in 10 ml of semisolid UMS media with the addition of

0.175% agarose (w/v). The media was supplemented with 20 mM succinate, vitamins and relevant antibiotics (2.4.2 above). The semisolid UMS media was autoclaved prior to inoculation to both sterilize and dissolve the agarose before leaving to set overnight at room temperature. Cultures for inoculation were grown to an OD595 of 0.6 in 10 ml of UMS media supplemented with 20 mM succinate, 20 mM NH_3Cl and vitamins. The 10 ml universals were inoculated with $1X10^5$ cells by direct pipette injection into the center of the media.

Please note that as A. caulinodans ORS571 is auxotrophic for nicotinate, any UMS growth media was supplemented with 30 μM nicotinic acid. *E. coli* and *A. caulinodans* ORS571 was grown at 37°C with no shaking, *A. brasilense* FP2 was grown at 30°C with no shaking. Unless otherwise stated the cultures were grown in the dark until a visible pellicle had formed in the growth media (Figure 2-3).

2.5 Molecular Techniques

2.5.1 DNA Isolation

Genomic DNA was isolated from *A. brasilense* and *A. caulinodans* cultures using Purelink Genomic DNA Mini Kit (ThermoFisher Scientific), following the manufacturers protocol using $1x10^9$ cells per isolation. Plasmid DNA was extracted from E. coli using either the GeneJET Plasmid Minirep Kit (ThermoFisher Scientific) or QIAprep Spin Miniprep Kit (Qiagen). Both kits were used following the manufacturer protocols. When DNA purification was required following restriction digest or PCR either GeneJET PCR Purification kit (ThermoFisher Scientific) or QIAquick PCR Purification kit (Qiagen) was used, according to manufacturers protocols. DNA isolation from electrophoresis gels was carried out using GENEJet Gel Extraction Kit (ThermoFisher Scientific) or QIAquick Gel Extraction Kit (Qiagen) according to manufacturers protocols. Isopropanol (100% w/w) was used to improve efficiency of purification of small products.

2.5.2 DNA Amplification by PCR

PCR primers were designed using Geneious R9 (Biomatters Ltd). Primers for HD (formerly BD) cloning were designed using the ClonTech InFusion™ online tool. Primers were synthesised by Eurofins MWG Operon. Mapping/screening PCR reactions were carried out in 10 µL reactions using GoTaq® Green Master Mix (Promega). Amplification PCR reactions were carried out using in 25 µL or 50 µL reactions using Phusion® High-fidelity PCR Master Mix (ThermoFisher Scientific) or Q5 High-Fidelity DNA Polymerase (NEB). When difficult-to-amplify or high-GC regions of DNA were to be amplified a Phusion® reaction was set up using GC buffer with 5% DMSO included in the reaction. ThermoFishercycler conditions were set according to the manufacturer's instructions for polymerases, with Tm determined by Geneious R9 (Biomatters Ltd). PCR amplification for cloning or sequencing was carried out using genomic DNA at 5 pmol. Screening of E. coli transformants was carried out using colony PCR, where a single colony dip was transferred into the reaction using a sterile wooden toothpick or pipette tip. Where large-scale screening of rhizobacteria colonies was required colonies were transferred into 10 µL sterile H2O using a sterile pipette tip, and this was heated to 60°C for five minutes. The resulting suspension (1 µL) was used in the PCR reaction. When custom DNA sequences were required synthesis was carried out by GeneWiz (Takeley, UK), Integrated DNA Technologies (IDT, Leatherhead, UK) or by TWIST Bioscience (San Fransisco, USA).

2.5.3 DNA Gel Electrophoresis

GeneRuler® 1 kb or 100 bp (ThermoFisher Scientific) was used as a marker ladder for DNA electrophoresis. Electrophoresis was used to separate PCR products and restriction digests of DNA. Electrophoresis was carried out in 0.9% agarose (Sigma Aldrich) in TAE buffer (400 mM Tris acetate, 1 mM EDTA). Gels were run at 100 V for 30-40 minutes. Agarose percentage was raised to 1.2% for better resolution when isolating small DNA fragments, and running time was extended to 60 minutes at 90 mV accordingly. DNA was stained using Sybr® Safe (Invitrogen). Sybr® Safe was added to molten agar (1:10,000) before gels were cast. Stained agarose gels were visualized using the GelDoc EZ System (BioRad).

2.5.4 DNA Fragmentation by Restriction Enzyme Digests

Restriction enzyme digests were carried out on purified DNA using restriction endonucleases and their respective buffers (Roche, ThermoFisher Scientific or NEB) following manufacturer's instructions. Fragmented DNA was separated by gel electrophoresis as described above, and when required DNA was purified from the restriction reaction or agarose gels using the GeneJET PCR Purification kit (ThermoFisher Scientific) as per the manufacturer's instructions.

2.5.5 DNA Ligation

DNA ligation was carried out using T4 DNA Ligase and 10X T4 DNA Ligase Buffer (ThermoFisher Scientific). Reactions were carried out according the manufacturers protocol. Ligations into the pJET1.2/blunt vector were carried out at room temperature for 30 minutes. All other ligations were carried out at room temperature overnight.

2.5.6 Recombination-based Cloning HD (BD) Cloning

HD cloning (previously known as BD cloning) was carried out using InFusion® HD Cloning Kit (ClonTech). Primers were designed with a 15bp overhang corresponding to the ends of the destination vector or adjacent DNA fragment, designed using the online Primer Design Tool provided by ClonTech and tested in silico using Geneious R9 (Biomatters Ltd). HD cloning was carried out in a ThermoFishercycler according to the manufacturers protocols.

2.5.7 Type IIS Restriction Enzyme Cloning

Golden Gate cloning uses Type IIS restriction enzymes BsaI, BpiI, and DraIII to allow direct cloning of several fragments into a predefined order. Type IIS restriction sites cleave at a defined distance from their recognition site, leaving a specific 4 base pair overhang (Figure 2-1B). These unique overhangs can be designed as part of primer and vector design and so fragments can be efficiently cloned in a predefined and rational order. The generation of 4 base pair unique overhangs allows multiple DNA fragments to be assembled in a single reaction. A modular cloning (MoClo) system has subsequently been developed with a standardized set of -63- overhangs for parts used in the Golden Gate system. This relies on a level-based hierarchical system. Level 0 vectors carry the individual parts required for an assembly, such as promoters or coding sequences (Figure 2-1A). These parts are maintained inside replicating domesticated plasmids for long-term storage and efficient DNA recovery purposes. Level 0 modules can then be assembled into a full transcriptional unit, known as a Level 1 module. These Level 1 modules can be assembled into a vector and used for expression of genes in a compatible host bacterium. Alternatively, multiple level 1 modules can be combined together in a vector to form a Level 2 multigene construct (Figure 1A; Weber et al. 2011).

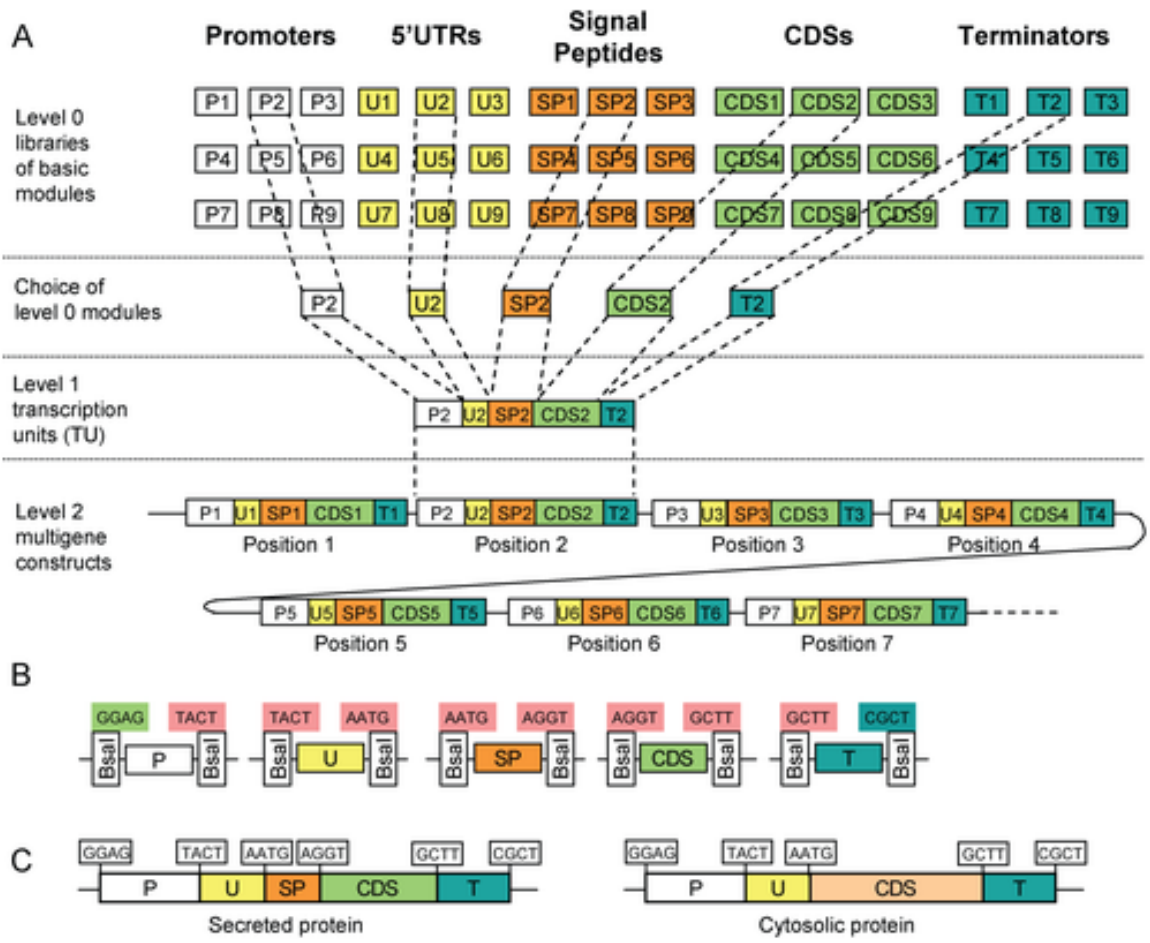


Figure 2-1 - Overview of the modular cloning system used in Golden Gate assembly. **A**: A suite of compatible level 0 modules are assembled to make a transcriptional unit 'TU' (level 1). These level 1 units can then be assembled into multigene units (level 2); **B**: Each level 0 module has a unique overhang allowing determined orientation and order in the TU **C**: Level 0 modules can be used in different combinations to design various forms of transcriptional unit, for example with and without degradation tags. Reproduced with permission [113].

A set of standardised Level 0 vectors are available for cloning parts of interest from various databases such as Addgene (Cambridge, MA) or ENSA (Norwich, UK). It is also possible to create new level 0 modules by PCR with the addition of specific overhangs. Once a Level 0 module has been constructed by PCR it can be subsequently be used in a single reaction to assemble a transcriptional unit.

Level 0 modules (‘parts’) are assembled into Level 0 vectors using Type IIS enzyme BpiI (ThermoFisher Scientific). Each Level 0 vector contains unique recognition sites and overhangs for Type IIS enzyme BsaI (ThermoFisher Scientific) either side of a LacZ selectable marker cassette. A Level 1 assembly reaction can be carried out with BsaI (ThermoFisher Scientific) in order to assemble the DNA parts in the desired order, from multiple level 0 plasmids. Golden Gate reactions were carried out using 1 µL of the desired Type IIS enzyme, 1 µL T4 DNA ligase, 1.5 µL DNA ligase buffer (ThermoFisher Scientific), 1.5 µL BSA (2 mg mL-1) and a total of 40 fmol of the required parts and vector backbone, made up to 15 µL with dH_2O. A series of vector backbones based on conventional cloning vectors as well as a pipeline to construct new ones has now been developed in the Poole lab (Geddes et al., in production). If DNA fragments of interest contain recognition sites for Type IIS enzymes, these sites must be 'domesticated'. Domestication involves changing the specific nucleotides of the type IIS enzyme recognitions site whilst conserving codons in any open reading frame. This may be carried out by overlap PCR or site directed mutagenesis (SDM) see 2.5.8 below. Primers were designed using Geneious R9 (Biomatters Ltd). In cases where there are multiple domestication reactions required in one sequence, DNA fragments may be synthesised commercially with codon optimisation as part of the design process.

2.5.8 DNA Sequence Verification

All DNA sequences once cloned into a stable vector were verified using the EuroFins TubeSeq service (Eurofins, Luxemborg). 10 uL of plasmid DNA of a concentration of > 100 ng μl^{-1} was combined with 10 pmol μl^{-1} of a sequencing primer of choice designed using the Geneious R8 primer design software (Geneious, Biomatters Ltd). Two sequencing reactions were performed on a region of DNA, one in the forward and one in the revrese direction to ensure double coverage of any DNA fragment. Verified sequences were aligned to predicted DNA sequences using the MUSCLE alignment algorithym 3.8.425 with 8 iterations[114].

2.5.9 Oligonucleotide Annealing

Complementary single stranded DNA (ssDNA) oligonucleotides were annealed together at a concentration of 4 μM in a buffer containing 60 mM Tris HCl (pH 7.5), 500 mM NaCl, 60 mM $MgCl_2$ and 10 mM DTT. The mixture was heated to 95˚C for five minutes and then a decreasing temperature gradient of 0.05˚C sec^{-1} was used to anneal the oligonucleotides until a temperature of 20˚C was reached. Annealed linear dsDNA sequences were then diluted to 0.15 μM for use in Golden Gate cloning reactions and with other cloning techniques.

2.5.10 Overlap PCR and Site Directed Mutagenesis (SDM)

SDM reactions were performed using the Q5 SDM Kit (New England Biolabs) following the manufacturer's instructions. Primers were designed using Geneious R9 (Biomatters Ltd). Overlap Extension PCR / SDM can be used to direct mutagenesis within DNA regions. This was used for domestication, or for targeted alterations or small single nucleotide deletions within DNA fragments used in conventional cloning. Overlap PCR primers were designed using complementary 15 bp primer overhangs. These overhangs can be altered to allow deletion or mutation. All overlap PCR primers were designed using Geneious R9 (Biomatters Ltd). In cases where there are multiple domestication reactions required in one sequence, DNA fragments may be synthesised commercially with codon optimisation as part of the design process.

2.5.11 DNA Consensus Sequence Analysis

DNA consensus sequence analysis was utilised for determining the constituent promoter sequences for the rhizobacteria *Azospirillum brasilense* FP2 and *Azorhizobium caulinodans* ORS571. The prediction of prokaryotic promoter elements and regulons (PePPER) software package was used to determine the

consensus sequence of the two rhizobacteria [115]. A FASTA file containing intergenic DNA sequences (lacking ORFs) was produced after analyzing both the *A. caulinodans* ORS571 and *A. brasilense* FP2 genomes with Glimmer3. This file was then analysed subsequently by the software package and a consensus sequence per nucleotide position was generated.

2.6 Transformation Techniques

2.6.1 Thermal

Chemically competent *Escherichia coli* DH5α cells were used for regular transformations and for HD cloning. These were either obtained commercially or generated in the laboratory. For the latter, DH5α was inoculated in 500 mL LB for 3 hours. Once the culture reached OD595 0.3-0.4 cells were incubated on ice for 10 minutes before being pelleted at 4000 rpm at 4°C for 10 minutes. The cell pellet was resuspended in 10 mL ice cold 0.1 M $CaCl_2$ 15% glycerol (v/v) and incubated on ice for 15- 30 minutes. Suspensions were pelleted at 4°C at 4000 rpm for 10 minutes and $CaCl_2$ solution decanted. Pellets were resuspended in 20 mL ice cold 0.1 M $CaCl_2$ 15% (v/v) glycerol and incubated at 4°C for 30 minutes. 50-200 µL aliquots were dispensed into chilled Eppendorf® tubes and snap frozen in liquid nitrogen. Competent cells were stored at -80°C. S17.1 competent cells containing the λpir gene for conjugal DNA transfer were generated using the same treatment. For Golden Gate cloning commercial Gold or Silver standard competent cells (Bioline) were used for high efficiency transformations as per the manufacturer's instructions.

For transformation of plasmids into *E. coli*, 10 µL of cells were thawed on ice and 1 µL of plasmid or ligation reaction (50-200 ng μL^{-1}) was added. In the case of

Golden Gate reactions 1 µL of reaction was added to 10 µL of competent cells. Cells were incubated on ice for 30 minutes before being heat-shocked at 42°C for 45 seconds. Cells were incubated on ice for a further two minutes. 400 µl of SOC medium (Melford) was added, and cells shaken at 37°C for 1 hour at 300 rpm. Cells were spread-plated on LB agar containing appropriate antibiotics and plates incubated overnight at 37°C.

2.6.2 Electroporation

Electrocompetent stocks of both *A. caulinodans* and *A. brasilense* were made in preparation for transfer of plasmid DNA by electric pulse application. Cultures of both rhizobacteria were grown overnight in TY media with 200 rpm shaking at 30˚C for *A. caulinodans* and 37˚C for *A. caulinodans.* Cells from the overnight culture were sub-cultured into 500 ml of fresh TY media and grown to an OD of 0.6 with the same conditions as for the overnight growth period. Cells were harvested by centrifugation at 3000 g for 15 minutes at 4˚C. The bacterial pellet from the 500 ml culture was washed twice with 500 ml of sterile distilled water and once with 500 ml of sterile 10% (v/v) glycerol. The final bacterial pellet was resuspended in 2 mL of sterile 10% glycerol (v/v) and fractioned into 100 µL aliquots. The aliquots were then shock-frozen in liquid nitrogen and stored at -80˚C.

Transformation of the electrocompetant rhizobacteria was performed by adding 50 ng of plasmid DNA to a 100 µL aliquot of thawed electrocompetant cells. The mix was left for 30 minutes on ice before transfer to a sterile electroporation cuvette (Cell Projects Ltd). A electric pulse was applied using a MicroPulser™ Electroporator (Bio-Rad Ltd) wit settings of 21 kV cm^{-1} with a time constant fixed at 5 ms. The cells were immediately resuspended in 1 ml of fresh TY and transferred to a fresh 1.5 ml Eppendorf® microfuge tube. The cultures were incubated for 3 hours at 30˚C for *A. brasilense* and 37˚C for *A. caulinodans* with 200 rpm shaking. 200 µL of the culture was plated in series dilution onto TY plates with appropriate antibiotics. Plates were incubated at 30˚C for *A. brasilense* and 37˚C for *A. caulinodans* for 2-3 days.

2.6.3 Conjugation

Triparental mating was used to conjugate plasmids from *E. coli* into the rhizobacteria *A. caulinodans* ORS571 and *A. brasilense* FP2. Helper plasmid pRK2013[116] was used in all triparental conjugations. For low efficiency conjugations biparental mating was utilised. Plasmids were transformed into competent S17.1 cells (see 2.6.1 above), and pRK2013 was not used. Three days

prior to conjugation the recipient rhizobacteria strain was grown on a TY slope containing relevant antibiotics at 30˚C for *A. brasilense* FP2 and 37˚C for *A. caulinodans* ORS571. One day prior to conjugation, the donor strain containing the plasmid and strain containing pRK2013 were inoculated in 10 mL LB containing appropriate antibiotics. Overnight cultures were sub-cultured 500 µL into 5 mL fresh LB containing appropriate antibiotics, and grown for 3-4 hours at 37°C shaking at 100 rpm until OD595 0.5. *E. coli* strains were pelleted at 4000 rpm for 10 minutes and gently washed with TY medium three times in order to remove traces of antibiotics. The rhizobacteria slope was washed with 3-5 mL of TY. Recipient rhizobacteria (400 µL) was mixed with 400 µL -66- donor E. coli and 200 µL pRK2013. This suspension was centrifuged at 6000 rpm for 5 minutes. The resultant pellet was resuspended in 30 µL TY medium and spotted on a TY plate, incubated at 30˚C for and 37˚C for overnight. A sterile loop was then used to streak bacteria from the TY plate spot onto a fresh TY plate containing appropriate antibiotics, and incubated for 2 days at 30˚C for *A. brasilense* FP2 and 37˚C for *A. caulinodans* ORS571. Nitrofurantoin (5 µg mL^{-1}) was used to select against *E. coli* when conjugating into *A. caulinodans* ORS571. Ampicillin (100 µg mL^{-1}) was used to select against *E. coli* when conjugating into *A. brasilense* FP2.

2.7 Assays

2.7.1 Growth of Rhizobacteria

Growth curves of *A. caulinodans* and *A. brasilense* were obtained using a CLARIOStar Omega Microplate Reader (BMG). and were grown on TY slopes plus appropriate antibiotics for 3 days at 37˚C and 30˚C respectively. Slopes were washed with 3-5 mL TY and diluted to OD595 0.01 ($1x10^7$ cfu). 400 µL cultures were grown per well in black PVC 24-well microcell plates (Vision Plate™ 24, 4titude) with transparent, gas permeable seals (Gas Permeable Moisture Barrier Seal, 4titude). Plates were shaken at 500 rpm orbital shaking at 30˚C for *A. brasilense* and 37˚C for *A. caulinodans*. OD595 was measured every 20 minutes for 24-36 hours. Data was processed using the MARS data analysis suite (BMG). Flocculation of *A. caulinodans* was common in the microplate growth assays but reliable OD595 measurements could be obtained by using orbital averaging of 2 mm for each microwell using the bottom optic of the plate reader.

2.7.2 Fluorescent Protein Excitation and Emission Determination

The excitation and emission spectra of a suite of fluorescent reporters was determined using a CLARIOStar Omega Microplate Reader (BMG). Fluorescent proteins were cloned by Golden Gate into the stable broad host-range expression vector pOGG024. The reporter proteins were expressed with constitutive sigma70 promoters of a medium strength (J23106) and a RBS of standard strength (RBSstd). Plasmids were transformed into recipient *A. caulinodans* and *A. brasilense* by electroporation, see 2.6.2 above. Cultures were grown until an OD595 of 0.6 overnight in TY media with appropriate antibiotics with 200 rpm shaking at 30˚C for *A. caulinodans* and 37˚C for *A. brasilense*. 1 ml of the test cultures was transferred to a black PVC 24-well microcell plates (Vision Plate™ 24, 4titude) with transparent, gas permeable seals (Gas Permeable Moisture Barrier Seal, 4titude). Spectral scans were performed for both excitation and emission between 340 - 740 nm in 5 nm bandwidth increments.

2.7.3 Fluorescent Reporter Gene Expression Analysis

Fluorescent protein expression was assayed concurrently with the high throughput growth assay as described in 2.7.1 above as to enable normalisation of fluorescence per cell post assay. Fluorescence determinations were made

using a CLARIOStar Omega Microplate Reader (BMG) with the MARS Data Suite for post-run analysis (BMG). Relative gene expression values were determined at an OD595 of 0.6 unless otherwise stated and normalised to an empty vector without the specific gene expression cassette. This value was then normalised to the optical density and therefore cell number of the culture, yielding a relative expression unit or REU.

Filter settings varied depending on the variant of fluorescent reporter protein used in each assay. Determination of mTagBFP fluorescence was made using a 399 nm excitation with 5 nm bandpass and emission of 454 nm with 5 nm bandpass filtration. Determination of eCFP fluorescence was made using an excitation of 433 nm with 5 nm bandpass and an emission of 475 nm with 5 nm bandpass filtration. Determination of sfGFP fluorescence was made using an excitation of 485 nm with 5 nm bandpass and an emission of 510 nm with 5 nm bandpass filtration. Determination of sYFP2 fluorescence was made using an excitation of 515 nm with 5 nm bandpass and an emission of 527 nm with 5 nm bandpass filtration. Determination of mCherry fluorescence was made using an excitation of 587 nm with 5 nm bandpass and an emission of 610 nm with 5 nm bandpass filtration. Determination of mRuby3 fluorescence was made using an excitation of 558 nm with 5 nm bandpass and an emission of 592 nm with 5 nm bandpass filtration. Table 2-5 indicates the excitation and emission maxima of the fluorescent proteins used in this study.

Fluorescent Protein Reporter	Excitation Wavelength (nm)	Emission Wavelength (nm)	Data Reference
mTagBFP2	399	454	[117]
eCFP	433	475	[118]
sfGFP	485	510	[119]
sYFP2	515	527	[120]
mCherry	587	610	[121]
mRuby3	558	592	[122]

Table 2-5 -Excitation and emission filters used for the determination of fluorescent protein expression in *A. caulinodans* ORS571 and *A. brasilense* FP2.

2.7.4 Single Cell Gene Expression Analysis – Flow Cytometry

Plasmids were transformed into *A. caulinodans* ORS571 and *A. brasilense* FP2 (2.6.2 above) and grown overnight in 1 mL TY with appropriate antibiotics at 37°C and 30°C respectively with 200 rpm shaking. Cell populations were diluted into 1X phosphate buffered saline (2:1000) and loaded for detection into a Attune NxT Flow Cytometer (ThermoFisher Scientific, CA, USA). sfGFP/sYFP2 detection (FL1, Excitation 488 nm, Emission 525/30 nm) and mCherry/mRuby detection (FL3, Excitation 561 nm, Emission 615/25 nm) for approximately 48 000 cells per sample, was carried out and analysed using FlowJo (v10.15) software (FlowJo, LLC, Oregon, USA). In order to analyse the library populations, a gating strategy

was used in which reference regions were quantified. To do this, Level 1 plasmids were prepared with either sfGFP or mCherry assembled with low (J23115), medium (J23106) and high (J23104) promoter strengths, the RBStd and DT16S terminator. The distribution of library populations was then estimated by overlapping with the known reference standards.

2.7.5 Nitrogen Fixation Assay – Acetylene Reduction Assay

An acetylene reduction assay was used to determine the activity of the nitrogenase enzyme inactive cultures of rhizobacteria. This assay takes advantage of the non-specific activity of the nitrogenase enzyme, which is able to break triple bonds in several chemical species[123]. Nitrogenase breaks the triple carbon-carbon bond in acetylene to result in ethylene. Figure 2-2 demonstrates the reaction carried out by nitrogenase on acetylene.

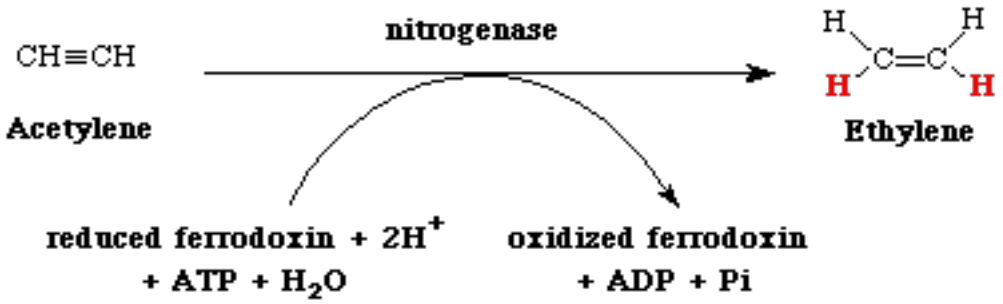


Figure 2-2 – Conversion of acetylene to ethylene by nitrogenase. Cofactors illustrated.

A. caulinodans ORS571 and *A. brasilense* FP2 were grown under nitrogen-fixing conditions, as described in 2.4.4 above. The typical 'bell-shaped' pellicle formed by *A. caulinodans* is illustrated in Figure 2-3. The sealed universal tubes were removed from the low-oxygen cabinet after 48 hours growth, maintaining the 3% oxygen internal atmosphere within the universal tubes. Gas impermeable silicon lid tops were used to allow injection into the universal tubes without introducing ambient atmospheric gasses. Air (0.67 mL) was removed from each universal tube and 0.56 mL acetylene was added to create a 2% (v/v) acetylene concentration within the headspace of each universal tube. Tubes were then incubated at 37°C for *A. caulinodans* and 30°C for *A. brasilense* for 8 hours with no shaking. Gas samples (1 mL) were removed using a syringe and needle from the headspace of the universal tubes and run through a Clarus® 480 (Perkin-Elmer) to determine conversion of acetylene to ethylene. Nitrogenase activity was expressed as µmol ethylene produced per mg protein per hour.



Figure 2-3 - Pellicle formed by *A. caulinodans* ORS571 in semisolid 0.175% (w/v) agarose UMS media supplemented with 10 mM succinate and 30 μM nicotinic acid. Image taken after 42 hours growth.

2.7.6 Cellular Weight and Protein Content Determination

For the quantification of biochemical values per mg protein, a bicinchoninic acid assay was utilised. Samples of OD 0.7 cultures of rhizobacteria (10-25 µL) were made up to 25 µL in sterile water. When required, samples were further diluted to obtain a concentration 0.2-10 mg mL^{-1} to fall within the linear range of the assay. NaOH (25 µL 2M) was added to each tube, and samples incubated at 100°C for 10 minutes. Cu_2SO_4 (20 µL of a 4% (w/w) stock solution) and 1 mL bicinchoninic acid ($(HO_2CC_9H_5N)_2$) was added to each sample. Samples were vortexed thoroughly and incubated at 60 °C for 30 minutes. Samples were then cooled, and absorbance measured at 562 nm using a Helios Spectrophotometer (Thomas Scientific). BSA 0-1 mg mL^{-1} was used to construct a standard curve (Figure 2-4).

Dry weight determinations were performed by growing 200 ml of rhizobacteria to an OD595 of 1.2. Graduated aliquots of $1x10^{10}$ cells up to $5x10^{10}$ cells were then harvested by centrifugation at 4500 g for 10 minutes at 4˚C. The supernatants were then disposed and the pellet dried on a nitrocellulose membrane at 80˚C for 2 days. The dry weight was then determined using a Scientific Weighing Scale (Kern, Balingen).

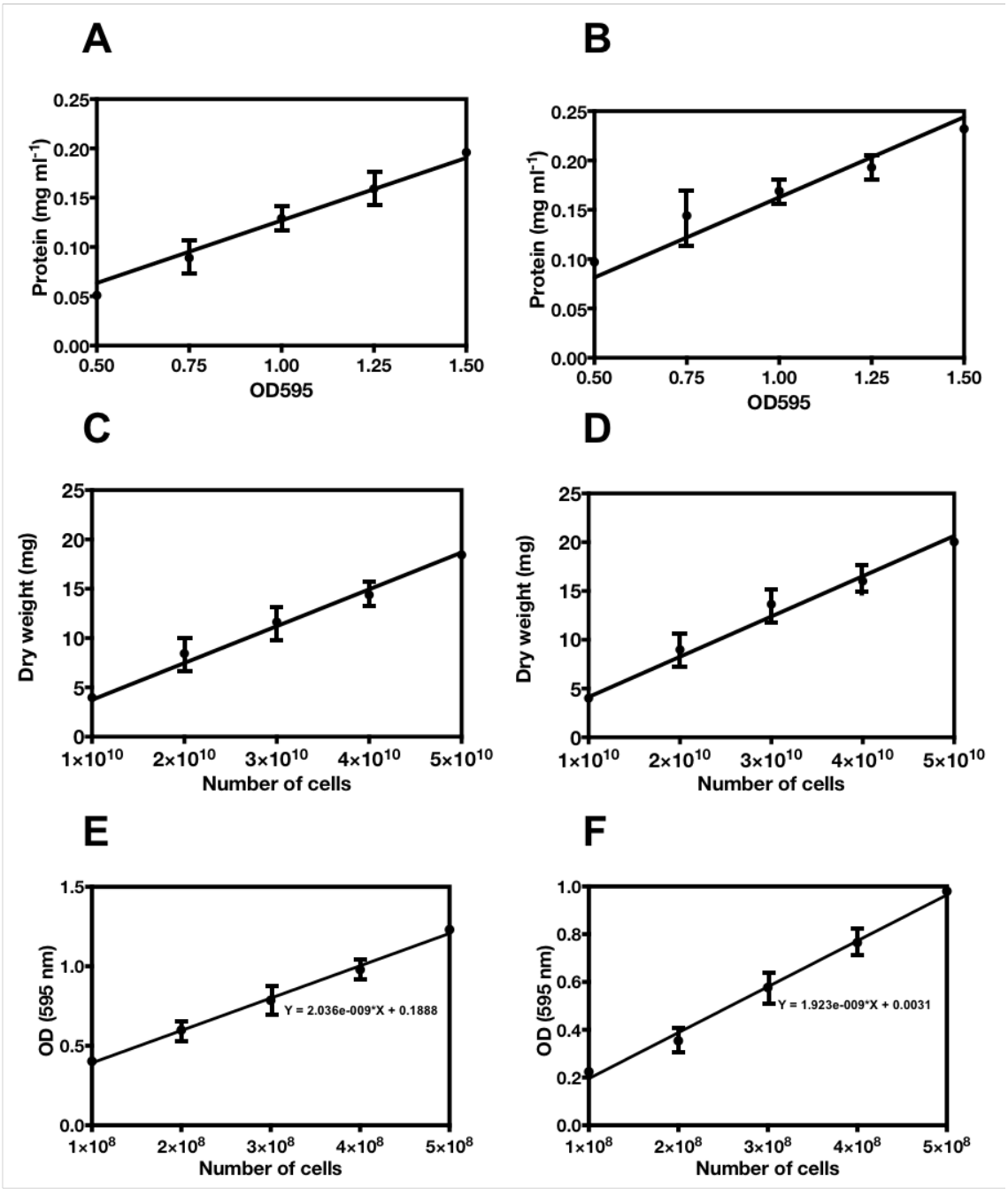


Figure 2-4 - Standard curves of protein content and dry weight for rhizobacteria. **A -** *A. caulinodans* ORS571 protein content. **B -** *A. brasilense* FP2 protein content. **C -** *A. caulinodans* ORS571 dry weight. **D -** *A. brasilense* FP2 dry weight. **E** - *A. caulinodans* ORS571 cell count per OD595. **F** - *A. brasilense* FP2 cell count per OD595. Error bars indicate the standard error of the mean (SEM).

2.7.7 Lux Activity in Rhizobacteria

A luciferase-based reporter system (lux) has been described before in the field[124]. This system allows *in vivo* measurement of promoter activity by the detection of photons emitted in result of the luciferase operon reaction. A promoter or reporter region of interest is cloned upstream of the *luxCDABE* operon from *Vibrio fischeri*. The bacterial luciferase reaction consumes oxygen and therefor the reaction may be influenced by local oxygen tension in various biological contexts such as that of a biofilm or nodule. The operon has been domesticated and made into a pL0M-SC module via SDM (see 2.5.8 above). Promoter regions of interest were either amplified by PCR from genomic DNA or selected from the library of pL0M-P/PU parts available. BsaI sites were added to the ends of PCR products if required, and the product was cloned via Golden Gate reaction ahead of the lux operon into the vector pOGG24. Correctly constructed lux expression plasmids were electroporated into *A. caulinodans* ORS571 and *A. brasilense* FP2 (2.6.2 above). Plasmid pOGG24 in the wild type background was used as an empty vector control. Lux activity was assayed on both solid media and liquid media by high throughput microwell analysis. Growth on agar plates was measured in a NightOwl II camera (Berthold Technologies). Plates were imaged with a 20 second exposure and the response normalised to the empty vector negative control.

The high throughput lux expression analysis was determined using a CLARIOStar Omega Microplate Reader (BMG). Cultures were grown until an OD595 of 0.6 overnight in TY media with appropriate antibiotics with 200 rpm shaking at 30˚C for *A. caulinodans* ORS571 and 37˚C for *A. brasilense*. $1x10^6$ cells of the test cultures was transferred to a black PVC 24-well microcell plates (Vision Plate™ 24, 4titude) with transparent, gas permeable seals (Gas Permeable Moisture Barrier Seal, 4titude). 500 µl of UMS media with 20 mM Succinic Acid and 10 mM NH_4Cl and appropriate antibiotics were added to each well. Cultures were grown for 48 hours with 200 rpm shaking at 30˚C for *A. caulinodans* ORS571 and 37˚C for *A. brasilense* FP2. Luminescence measurements were taken every 20 minutes with a gain function of 1000 (of a total 4000). A concurrent analysis script was used to measure both OD595 and luminescence in relative light units (rlu). Maximum rlu was determined for each strain and expressed as rlu $cell^{-1}$.

2.8 Statistical Analysis

For all the experimental data presented as a single point, at least 3 biological and 3 technical replicates were used to derive a mean data value unless otherwise stated. Error bars on figures indicate the standard error of the mean unless otherwise stated. To obtain 95% confidence intervals between data points and variables, a paired t-test was implemented and the level of margin significance (p-value) is presented in the text where relevant. Where two independent variables are being compared between samples a two-way ANOVA was used to test for

significant effects of the two independent variables on the dependent variable being quantified. The Prism Software Suite (Graphpad, USA) was used to generate all graphical figures and perform statistical analysis *in silico*.

3 RhizoKit: A Synthetic Biology Toolbox for Rhizobacteria

3.1 Abstract

Golden Gate cloning is an adaptable DNA cloning tool in synthetic biology for the assembly of plasmid constructs often used in combinatorial pathway optimization, with a number of assembly kits developed specifically for *Escherichia coli*[125] yeast[126] and plant-based expression[127]. However, its use for synthetic biology in commonly utilised and environmentally-significant bacterial systems such as the Rhizobacteria has unexpectedly been overlooked. Here, we produce RhizoKit; a simple modular toolkit of DNA parts for a variety of applications in rhizobacteria from protein synthesis and degradation to the hierarchical assembly of transcription units based on the MoClo assembly standard[113]. The kit features a library of constitutive promoters, T7 RNAP orthologous promoters, RBS strength variants, synthetic terminators, protein degradation tags and fluorescent reporter proteins.

We validated RhizoKit in two significant rhizobacteria, *Azorhizobium caulinodans* ORS571 and *Azospirillum brasilense* FP2 by assembling multi-gene expression circuits, characterised *in vivo* and *in vitro* DNA parts, and performed combinatorial

pathway assembly, using pooled libraries of fluorescent proteins as a proof-of-concept. In summary, RhizoKit provides a standardised and adaptable synthetic biology toolkit for a variety of applications in *Rhizobacteria* synthetic biology.

KEYWORDS: synthetic biology, Golden Gate, Rhizobacteria, logic gates, DNA parts, RhizoKit.

3.2 Introduction

One of the major aims of synthetic biology is the production of natural and commodity chemical products such as drug precursors, fine and specialized chemicals and biofuels using renewable feedstocks[128]. This is driven by the need to transition away from a current reliance on petroleum-derived chemical products and production processes. However more recently, other applications of synthetic biology have been successfully achieved such as bacterial photography[129], biological computation[130] and *in situ* mercury bioremediation[131]. The latter of which is just one application of synthetic biology to maintaining the environment. One of the most significant global environmental issues is the access to fixed or available forms of nitrogen which limits global crop productivity. Nitrogenous fertilizer production currently represents a significant expense for the efficient growth of various crops in the developed world[132].

A more sustainable and environmentally friendly solution to this problem is in the application of synthetic biology to engineer synthetic genetic circuits within diazotrophic crop-associative rhizobacteria to produce and release nitrogenous compounds such as ammonia (NH_3). The development of a rapid, efficient and hierarchical DNA assembly technique is essential to achieving this goal and a number of methods have been proposed[133]. One such method called Golden Gate cloning, is a high-efficiency, hierarchical, cloning technique used for a wide variety

of DNA assembly applications, from the construction of simple gene expression cassettes to diverse biochemical pathway libraries[113,134]

Golden Gate cloning primarily utilises the non-palindromic Type IIS restriction endonuclease BsaI, which cuts outside of its recognition sequence (GGTCTCN↓ NNNN), producing a four base pair overhang. This overhang, also known as a fusion site, can be designed to provide both directional and sequential DNA part assembly. BsaI sites are placed at both the 5′ and 3′ ends of a DNA part in inverse orientation, allowing the restriction site to be removed during the cloning reaction by T4 DNA ligase. DNA parts are held within 'base vectors' and can include constitutive and regulatable promoters, protein tags, ribosome-binding sites (RBS), open reading frames (ORF), and terminators. A one-pot reaction of base vectors and occasionally PCR products with Golden Gate overhangs are then digested and ligated into 'destination vectors' using BsaI, T4 DNA ligase and a reaction buffer containing Mg^{2+-} ions. Golden Gate assembly has been reported to assemble up to 15 DNA fragments with greater than 98% efficiency[135] and relies solely on the use of the BsaI enzyme.

In the past half-decade, significant modifications to the original Golden Gate cloning technique have yielded such methods as Golden Braid and MoClo standards[113,136]. These cloning methods utilise alternating antibiotic marker plasmids and dual type IIS endonucleases (BsaI and BpiI) to rapidly construct hierarchical multigene pathways up to 50 kb in size[137]. In summary, several

plasmid libraries have been constructed including MoClo and GoldenBraid for plants[127] *Saccharomyces cerevisiae*[126] and *Escherichia coli*[125]. For a detailed review of gene expression circuit assembly strategies utilizing Type IIS endonucleases please see Casini et al. 2015[138]. The MoClo assembly technique allows for flexible design and construction of multiple transcriptional units into circuits with each gene under the control of its own individual promoter, RBS and terminator. As mentioned several MoClo-based gene assembly kits have been published, however a kit for rhizobacteria, one of the most abundant and agriculturally relevant bacteria group, has surprisingly been overlooked. In this study I built a toolkit by modifying and creating new flexible Golden Gate DNA 'parts' specifically for use in rhizobacteria. The toolkit is a custom-designed modular library of standard biological parts for the hierarchical assembly and optimisation of synthetic genetic constructs, using the MoClo assembly standard[113].

Herein, I present the rhizobacteria synthetic biology toolkit, a user-friendly, characterized and customizable Golden Gate kit with parts that can be used for example in combinatorial pathway optimisation, logic gate circuit construction, protein expression and degradation control and CRISPR interference technology. Included is a custom-designed library of constitutive σ70 promoters, inducible expression based on the T7 RNA polymerase, a suite of chemically inducible promoters, a standardized RBS library and a range of synthetic terminators enhanced beyond previously published characteristics [139]. The rhizobacteria kit

also provides an orthogonal T7 RNAP and a suite of cognate T7 promoters for multigene expression control with discreet gene-specific stochiometric transcription regulation[140]. The characterisation of a library of fluorescent protein reporters that can be used to monitor multigene circuit expression profiles as well as *in vivo* and *in vitro* spatial colonisation analysis is achieved. The validation of custom-designed synthetic biology parts was performed using a fluorescent protein reporter to monitor real-time expression profiles. The lack of inducible promoter DNA parts in all of the MoClo kits to date was an omission this study has fixed with the construction and characterisation of a suite of integrated inducible promoters and their cognate repressor or activator transcription factors.

In synthetic biology, there is an ever-increasing demand for a repertoire of biological parts (promoters, RBS, terminators) characterized in the context of the chosen DNA assembly technique and expression strain. For example, DNA parts behave very differently when placed in a different context (known as “context dependency”), which directly challenges the engineering dogma surrounding synthetic biology. A number of previous studies have tried to address this problem through the design of neutral insulator sequences such as Ribozyme post transcriptional processing and orthogonal homology-based DNA sequence optimisation[141,142]. To understand such effects on the DNA parts built here, the validation of the library of biological parts in several distinct rhizobacteria species including *Azorhizobium caulinodans* ORS571, *Azorpirillum brasilense* FP2 and to some degree *Rhizobacterium* sp. IRBG74 is performed. Furthermore the

application of the DNA parts characterised in this chapter form the basis of the latter part of the thesis. To demonstrate the advanced utility of the toolkit, the assembly of several orthogonal NOT gates to verify advanced computational power with cellular gene expression circuits is achieved (Chapter 4). With a focus on the agricultural relevance of rhizobacteria, the construction of several plant growth promoting (PGPR) phenotypes based on DNA parts is completed and their efficacy *in vitro* and *in vivo* is analysed (Chapter 5). The plant growth promotion hormone indole-3-acetic acid (IAA), also known as auxin, has been demonstrated to increase plant biomass and enhance root growth and branching[94]. In this thesis we engineered an orthogonal IAA biosynthesis pathway into *A. caulinodans* ORS571 and *A. brasilense* FP2 and demonstrate both auxin biosynthesis in liquid culture but also significant morphological changes in associated *Hordeum vulgare* plants when grown with our modified rhizobacteria (5.4 below). The potent cereal crop devastating disease 'damping-off' was our primary focus in the development of an antifungal producing biosynthetic pathway. 2,4-diacetylphluroglucanol (DAPG) is a powerful antifungal agent that has been shown to prevent zoospore formation in the damping-off disease causing *Pythium sporangiiferum*[143]. In this work the successful adaptation of the DAPG biosynthetic pathway from *Pseudomonas fluorescens* with the Rhizokit parts was demonstrated, and the production of of active DAPG under culture conditions is analysed (5.5.1 below)

In previous literature the species *E. coli* is well known as one of the major mainstays in synthetic biology, providing a very well-characterised chassis for the

manipulation of genetic circuits and implementation of biosynthetic pathways. Using the toolkit of DNA parts characterised here, modular pathways and circuits can be efficiently assembled using a common technique (Golden Gate) for a broad range of significant rhizobacteria. In addition, other cloning methods such as PCR, SDM and Gibson assembly[144] can be utilised to construct or modify the DNA parts further. In this study the demonstration of a universal cloning system for several rhizobacterial chassis is completed and demonstration that an ORF or DNA part (e.g., promoter, RBS or fluorescent protein) can be easily transferred between multiple systems and exhibit comparable characteristics is verified. The aim in the construction of the rhizobacterial toolkit is to provide future research users with a flexible and fully characterised synthetic biology toolkit that can be customised for a wide range of agriculturally-relevant phenotypes and uses in rhizobacteria.

3.3 Results

3.3.1 Constitutive Promoter Design and Characterisation

The sequence of constitutive promoters in the synthetic biology workhorse *E. coli* have been studied and published in several previous reports[145,146]. These studies often concentrate on the σ70 promoters that are common in the expression of housekeeping genes such as ruvB and gyrB[147]. Sequence characterisation of rhizobacterial constitutive promoters has not been a focus of study yet some characterisation in relevant species such as *Azospirillum* has been accomplished[148]. From previous studies focusing on *E. coli* promoter sequences we were able to extract a library of different constitutive promoters with a linear scale of transcriptional strength. The promoter sequences in question have been previously published for their characteristics in other species several times but were first utilised by JC Anderson for an iGEM project[149]. We hypothesised that a characterisation of a selection of the Anderson promoters in the two rhizobacteria species would yield a broad-host range library of promoter sequences for use in constitutive gene expression cassettes.

From the sequence data we extracted the consensus sequence for a rhizobacterial-*E. coli* hybrid promoter (see Figure 3-1). The -10 box (also known as the Pribnow box) TATAAT and the -35 box

5′-WTWNNNNNNNAWWNNNTTC**TTGACA**TNTNNNNNNNWTNTNN**TATAAT**DNNTNNRTWW-3′

-35 box -10 box

Figure 3-1 – Consensus sequence for σ70 promoters derived after a literature review. D = 33.33% A, G and C. N = 25% A, T, C and G. R = 50% each A and G. W = 50% each A and T.

TTGACA are well known to be present in many prokaryotic promoters and are well conserved across *E. coli* and *A. brasilense* in our comparative analysis. In addition, the sequence TTC and TNT are often found immediately upstream and downstream the -35 box, respectively. The sequence TG is frequently found 1 bp immediately upstream the -10 box. Based on this data a library of 35 bp long sequences containing 24 conserved (A, G, C and T) with point mutations in the -35 and -10 regions were chosen.

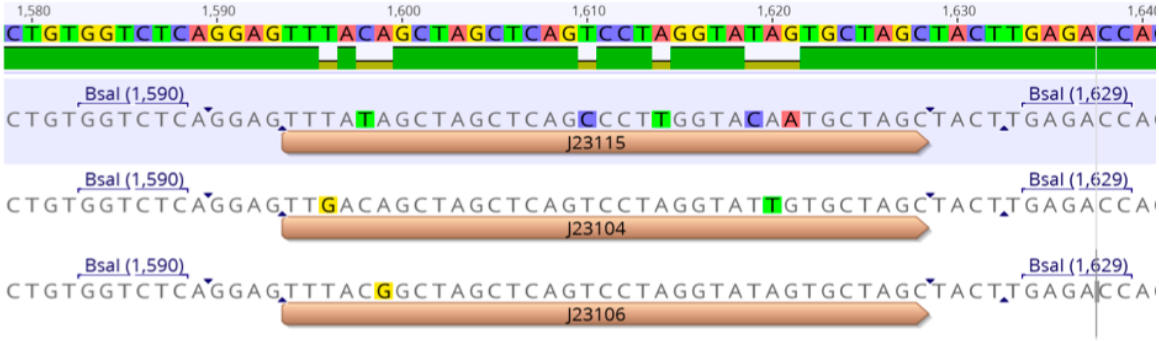


Figure 3-2 – Alignment of the chosen constitutive promoters used in this study. Note the Golden Gate enzyme recognition sites flanking the operational promoter ('P') sequence in orange. Alignment performed with MUSCLE 3.8.425 with 8 iterations [114].

Consensus promoter sequences were made into Golden Gate ‘P’ parts (Figure 3-2, as described in 2.5.7 above) and cloned with an RBSstd sequence ahead of a sfGFP fluorescent reporter gene and rrnB terminator into a stable broad host range plasmid with pBBR1 based replication system (Figure 3-3). These plasmids (pOPS1167, pOPS1168 & pOPS1169) was transformed into both Rhizobium sp. IRBG74, *A. caulinodans* ORS571 and *A. brasilense* FP2 and the relative strengths of several constituent promoters was analysed with the promoter Plac as a strong positive control (pOPS1011, see 2.7.3 above).

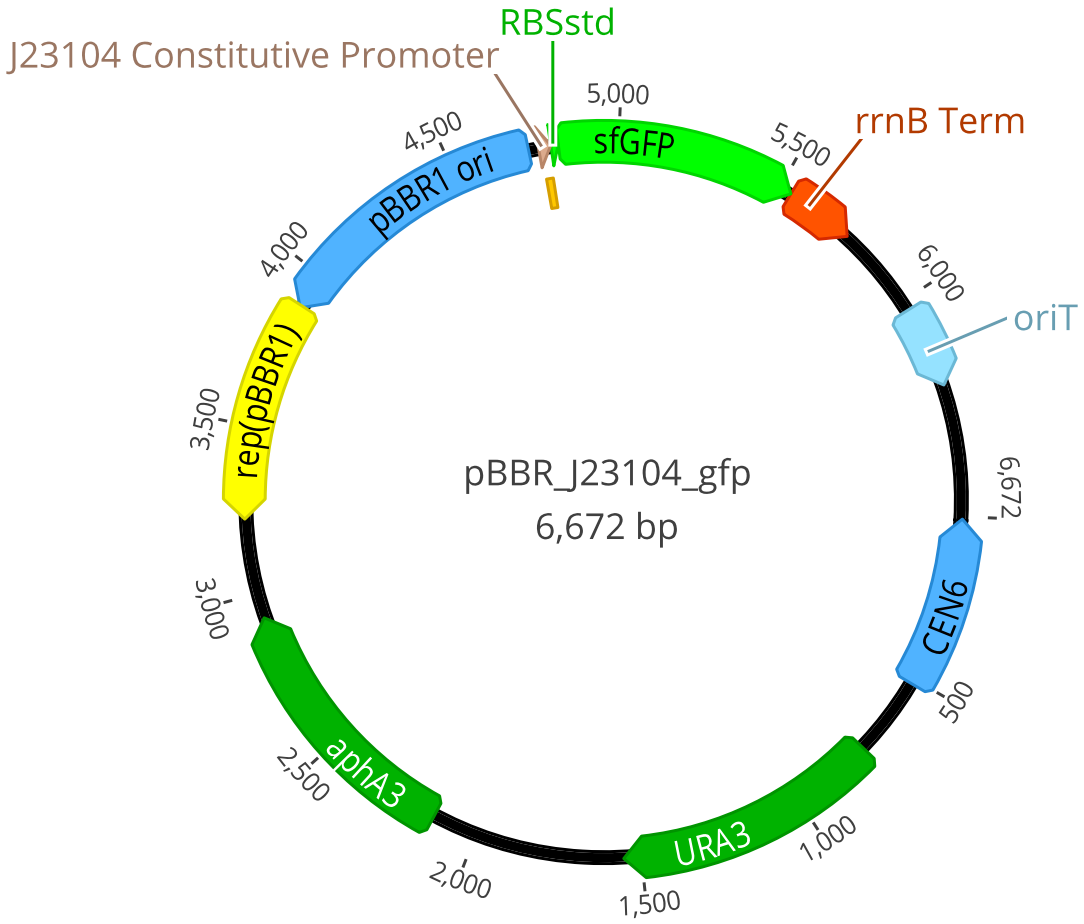


Figure 3-3 – Example plasmid map of the constitutive promoter and RBS analysis plasmids (pOPS1167, pOPS1168, pOPS1169, pOPS1171, pOPS1172, pOPS1173 & pOPS1174). The constitutive promoters and RBS sequences were altered to allow analysis of specific transcription and translation strengths. Note pBBR1 broad host replicon and *aphA3* gene than confers Kanamycin and Neomycin resistance for selection.

The results of the constitutive promoter strength analysis indicates comparable characteristics of the four chosen promoters across all three test strains (Figure 3-4). The strength of the J23104 promoter was not significantly different to that of the Plac promoter in all of the test strains. The σ70 Anderson promoters yielded

the predicted linear ‘strong, medium and weak’ transcriptional profiles across all three strains, although to various relative profiles in each strain. There was some variation observed in the absolute level of sfGFP expressed for each strain at the sampling point. The transcriptional strength profiles of the promoters in the strains *Rhizobium* sp. IRBG74 and *A. caulinodans* ORS571 are very similar with the relative transcriptional expression strengths not being significantly different from each other when a paired t-test was performed ($P>0.05$).

In the strain *A. brasilense* FP2 the difference transcriptional strength of the medium (J23106) and weak (J23115) promoters is not as pronounced as in the other two test strains. There is only a 22% difference in the strengths of the J23016 and J23115 promoters compared to a 51% difference in the strengths of the strong J23104 and J23106 promoters.

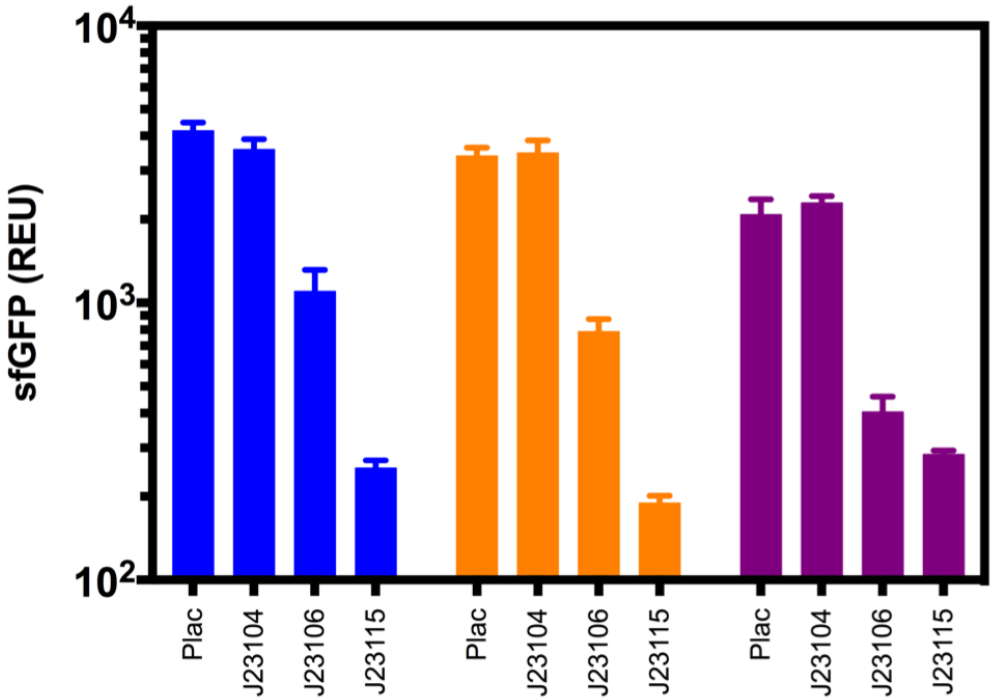


Figure 3-4 – Constitutive promoter library expression strengths. Blue = Rhizobium sp. IRBG74. Orange = *Azorhizobium caulinodans* ORS571. Purple = *Azospirillum brasilense* FP2. Error bars indicate the standard error of the mean (SEM). N = 3 biological and 3 technical replicates.

3.3.2 Ribosome Binding Site Sequence and Analysis

The sequences for ribosome binding sites (RBS) with different translational strengths were chosen from a full library of RBS sequences previously characterised for gram negative bacteria [50].

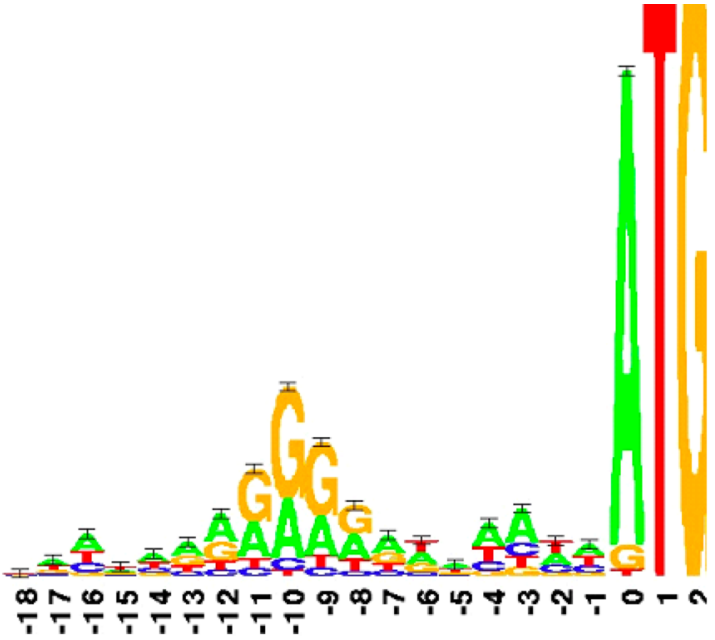


Figure 3-5 – Consensus sequence for bacterial and archaeal ribosome binding site. The Shine-Dalgarno sequence is clearly illustrated as the conserved AGGGG region -12 to -8 nucleotides ahead of the ATG start codon.

The consensus sequence in gram negative bacteria is known as the Shine-Dalgarno sequence [150]. The sequence recruits ribosomes to the mRNA transcript and aligns the first methionine tRNA anticodon with the first coding codon (AUG).

The consensus sequence for prokaryotic RBS sequences is illustrated in Figure 3-5.

The relative translational strengths of the selected RBSs were predicted to yield a logarithmic variation of strength across the library chosen. An alignment of the RBS sequences shows a suite of SNPs in both the conserved Shine-Dalgarno sequence but also the regions upstream and downstream have variation that changes the affinity of the DNA sequence for ribosomes. Interestingly the RBS sequence R1 has a 5bp region (AAATT) ahead of the ATG start codon which appears to allow for high levels of ribosome recruitment, possibly due to increased secondary structure of the mRNA transcript (Figure 3-6).

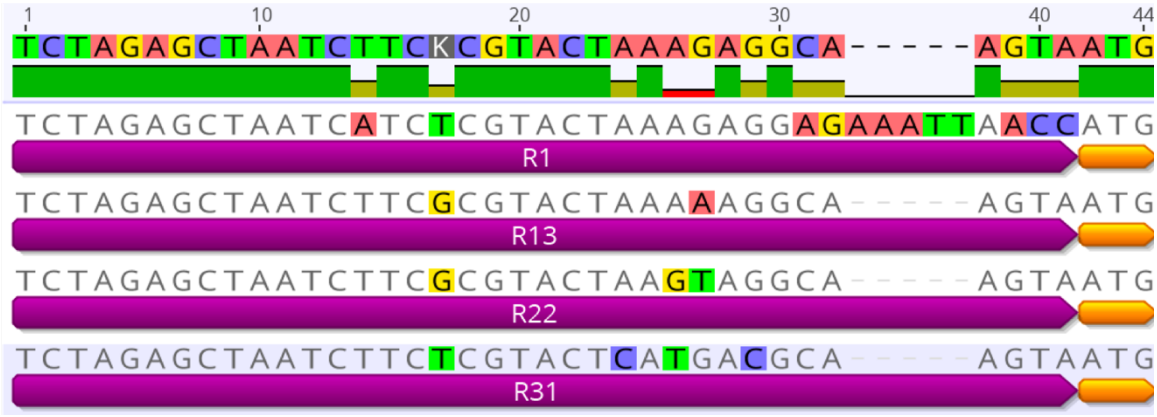


Figure 3-6 – Nucleotide alignment of the 4 RBS sequence library. Note the ATG start codon on the right hand side of the alignment where translation initiation occurs. Alignment performed with MUSCLE with 8 iterations [114].

Ribosome binding site sequences were made into Golden Gate 'U' parts (as described in 2.5.7 above) and cloned with the J23104 promoter sequence ahead of a sfGFP fluorescent reporter gene and rrnB terminator into a stable broad host range plasmid with the pBBR1 based replication system (Figure 3-3). These plasmids (pOPS1167, pOPS1171-74) were transformed into both *Rhizobium* sp. IRBG74, *A. caulinodans* ORS571 and *A. brasilense* FP2 and the relative translational strengths of 4 different ribosome binding sites were analysed with the RBSstd as a strong positive control (see section 2.7.3 above). The RBSstd sequence was derived as the consensus RBS from the bacteriophage T7, shown to have strong translational strength across many gram negative bacteria[151].

The observed translational profile of the RBS library showed a clear logarithmic range in strengths ranging from the strongest RBSstd to the weakest RBS31 sequence. The order of the RBS sequences strengths maintained across all three strains tested. There was however some variation in the relative strengths of the RBSs within each strain. For example the RBS13 and RBS22 sequences in *Azorhizobium caulinodans* ORS571 yielded very similar translational strengths unlike the differences observed in *Rhizobium* sp. IRBG74. The translational profiles of the RBS sequences in both *Azorhizobium caulinodans* ORS571 and *Azospirillum brasilense* FP2 were very similar with the relative strength patterns closely resembling each other more than that of *Rhizobium* sp. IRBG74.

There is a marked 20-fold difference in the RBS strengths between RBS1 and RBS 13, this was greater than the expected 15-fold difference predicted by the Salislab RBS predictor [50]. The full range of translational strengths is 140-fold across the library of RBS sequences in Rhizobium sp. IRBG74, 120-fold in *Azorhizobium caulinodans* ORS571 and 110-fold in *Azospirillum brasilense* FP2. There were significant differences in translational strengths of the RBS sequences within all three test strains (P>0.05).

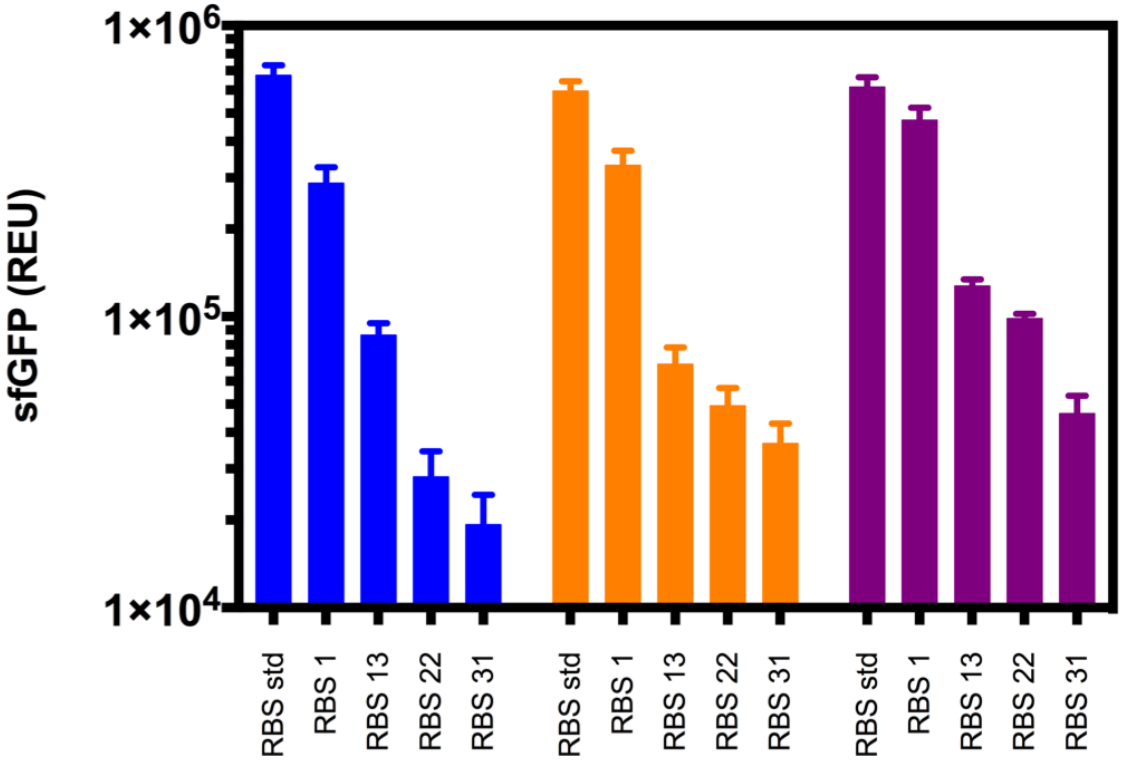


Figure 3-7 -Ribosome binding site (RBS) strengths. Blue = Rhizobium sp. IRBG74. Orange = *Azorhizobium caulinodans* ORS571. Purple = *Azospirillum brasilense* FP2. Error bars indicate the standard error of the mean (SEM). N = 3 biological and 3 technical replicates.

3.3.3 Terminator design and characterisation

The terminator library used in this work was created from a selection of terminators available in the laboratory already and from our collaborators at MIT (Voigt Lab). By predicting the secondary structure of the DNA in the terminator region, it is possible to predict the strengths of terminators depending on the level of stem-loop formation[152]. Terminators with larger and more complex stem-loop structures (caused by DNA sequence homology) are more likely to halt transcription by causing steric blocking of the DNA polymerase from migrating down the DNA polymer[153].

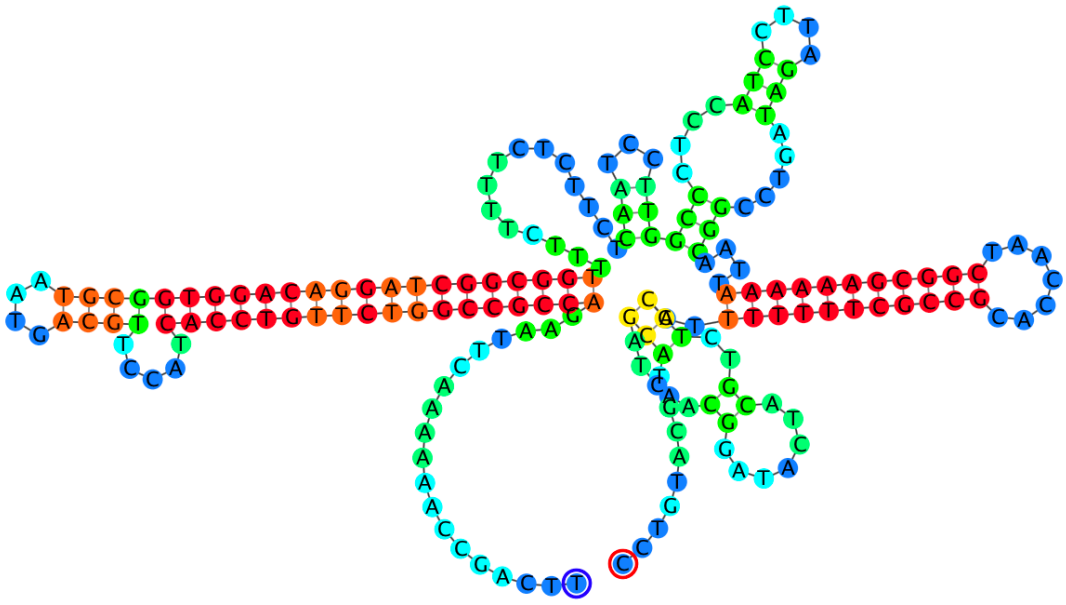


Figure 3-8 – Secondary structure of the terminator DT14. Structure determined using DNAfold plugin on Geneious 11.2.1. The scale from red to blue indicates increased stem-loop formation.

The suite of terminators were designed to have little to no homology between them as being over around 100bp in length, homologous recombination can occur between identical terminators if used in the same synthetic construct (Figure 3-9).

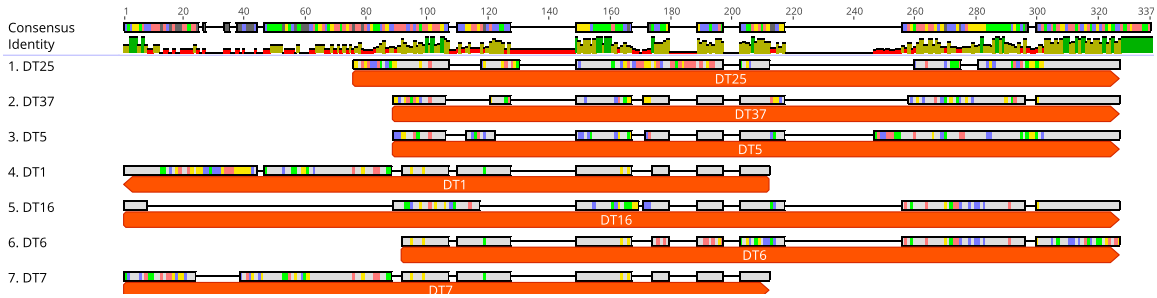


Figure 3-9 -Alignment of the terminator library DNA sequences. Alignment performed with MUSCLE with 8 iterations [114]. Note the high level of divergence in sequence between the terminator sequences.

Terminators were constructed as 'T' Golden Gate parts by oligonucleotide synthesis and annealing (see 2.5.8 above). The terminator sequences were then cloned between an sfGFP and mCherry fluorescent reporter operon in a stable broad host range plasmid. The sfGFP and mCherry operon is expressed from a J23104 constitutive promoter and both the sfGFP and mCherry genes are translated from RBSstd sequences (Figure 3-10). The family of terminator test plasmids (pOGG156-161) were transformed into *Rhizobium* sp. IRBG74, *A. caulinodans* ORS571 and *A. brasilense* FP2 and then the expression levels of both sfGFP and mCherry were measured in late log phase (2.7.3 above). Please note that in later parts of this work the 'S' denomination behind the promoter indicates

a shortened version with no change in termination performance. These parts were constructed in later work to increase Golden Gate cloning efficiencies of complex genetic circuitry.

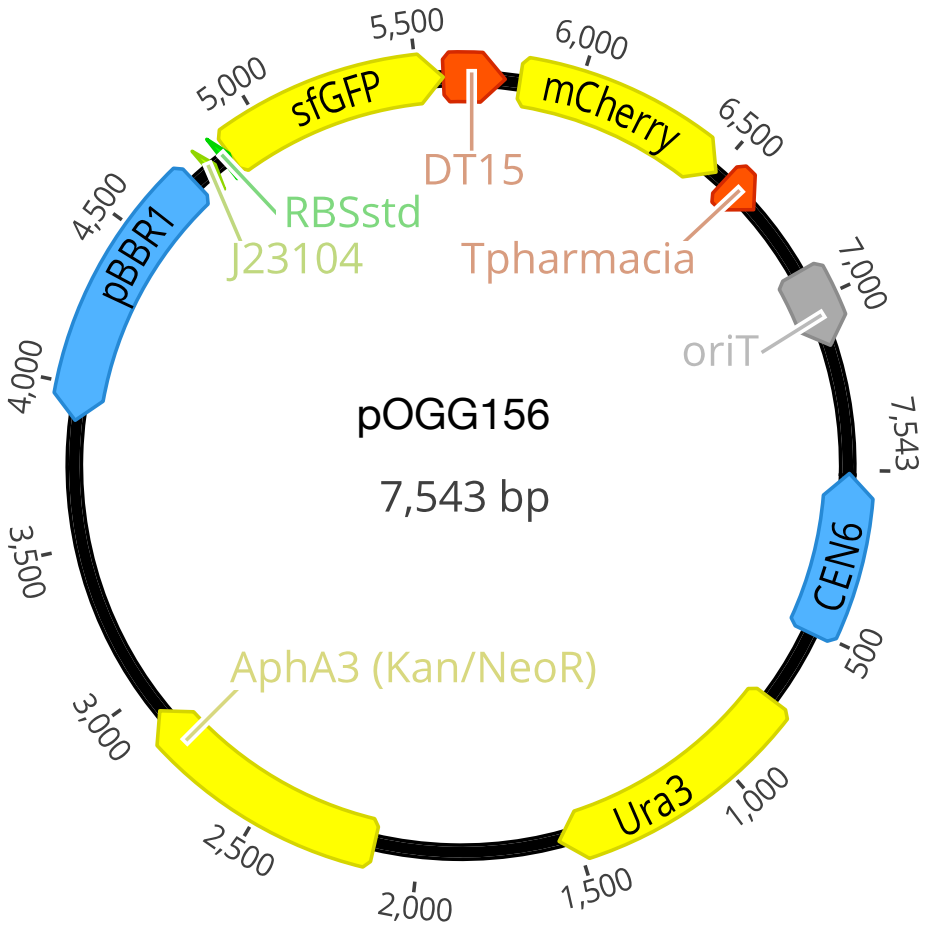


Figure 3-10 – Example plasmid map of the pOGG156-161 family of plasmids. The terminator sequences (in the above example DT15) were altered between the plasmids to allow analysis of specific termination strengths of each terminator in turn. Note pBBR1 broad host replicon and *aphA3* gene than confers Kanamycin and Neomycin resistance for selection.

The termination strength (Ts) unit is derived from the ratio of mCherry to sfGFP. For example, a Ts of 100 means the fluorescence signal from mCherry was 100 times weaker than that of sfGFP after normalisation to the fluorescent reporters quantum yield and extinction coefficients (see 2.7.2 above). The commonly used terminators T7, Tpharmacia and TrrnBT1 were added as positive controls and also to benchmark the termination strength of our modified synthetic terminators in rhizobacteria.

The results from the termination strength assay in the three rhizobacteria strains indicated that the all of the new synthetic terminator library functioned better than the incumbent synthetic biology terminators T7, Tpharmacia and TrrnBT1 (Figure 3-11). Across all three strains there was a range of termination strengths with all terminators exhibiting unique termination strengths. There is a 90-fold range between the weakest and the strongest of the new synthetic terminators across all the test strains with the weakest being TF4 with a Ts of 11 and the strongest was T15 with a Ts of 112. There appeared to be no correlation between the degree of secondary structure (the free energy of the ensemble) and the terminator strength in the library of terminators we tested.

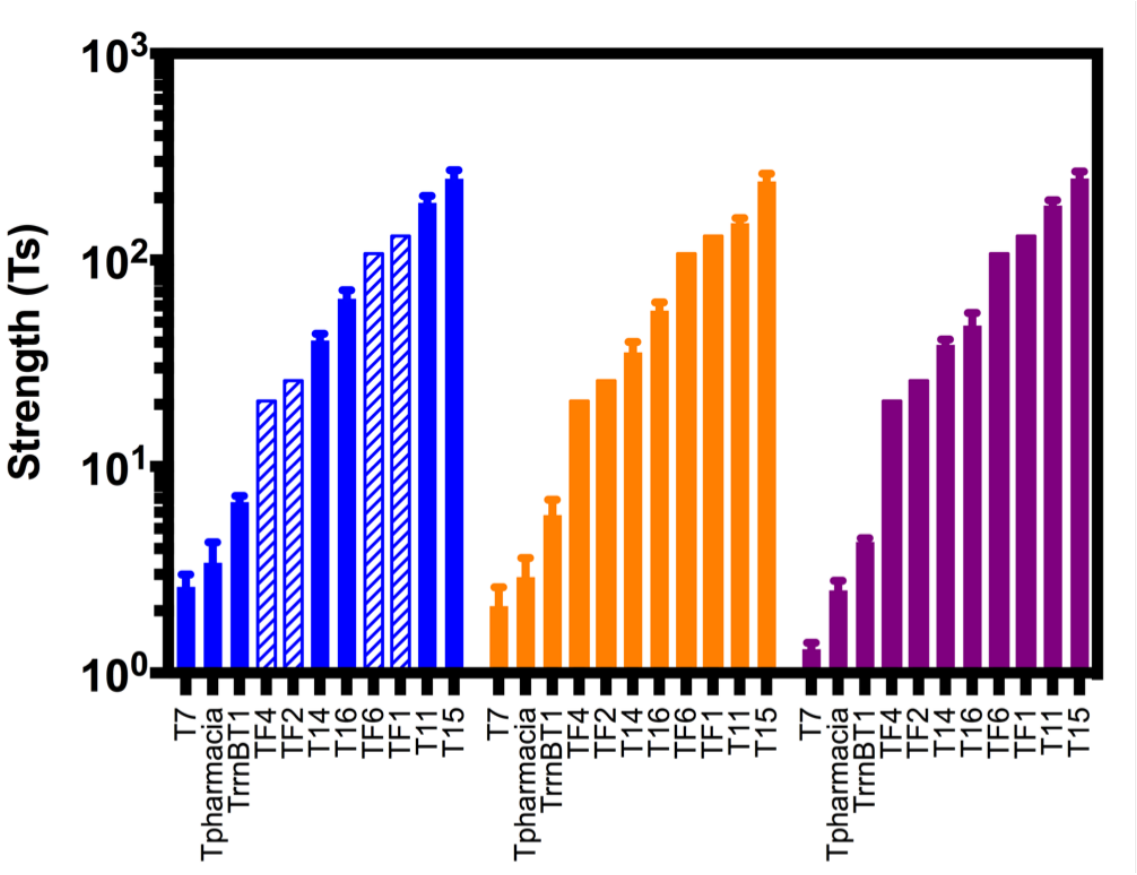


Figure 3-11 – Terminator strength profiles across all three test strains. Blue = Rhizobium sp. IRBG74. Orange = *Azorhizobium caulinodans* ORS571. Purple = *Azospirillum brasilense* FP2. Error bars indicate the standard error of the mean (SEM). N = 3 biological and 3 technical replicates. Hashed bars indicate where only a single biological replicate was tested due to experimental constraints.

3.3.4 Chemically Inducible Promoters

A suite of chemically inducible promoters were constructed as Golden Gate 'rPU' modules, the 'r' indicating a regulated promoter ('P') and an associated RBS ('U') module. These DNA parts consisted of a constitutively expressed transcriptional regulator (activator or repressor) as well as the cognate promoter for that specific promoter. A library of rPU modules were constructed using the same method as other 'PU' Golden Gate parts, therefore mimicking the four base pair overhangs of the 5' end of the promoter part and the 3' end of the RBS part. Each segment of DNA required for the part was PCR amplified from sequenced plasmid or genomic DNA by by PCR (2.5.2 above). An arabinose inducible rPU was constructed by amplifying the AraC and associated Pbad promoter from the iGEM plasmid K808000 using primers oxp1066 and oxp1012. At the same time the Pbad promoter was amplified without the AraC regulator to make a 'PU' module using primers oxp1011 and oxp1012. The AraC transcriptional regulator was amplified as a 'SC' module using the primers oxp1019 and oxp1020. A homoserine lactone (3O6-HSL) inducible promoter was constructed as a 'rPU' module by amplifying the LuxR and cognate Plux promoter from the iGEM plasmid K575025 using primers oxp1067 and oxp1016. Concurrent to the construction of the rPU module a 'SC' module of the LuxR protein was amplified using primers oxp1021 and oxp1022 and a Plux 'PU' part was amplified by primers oxp1015 and oxp1016. An Isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible promoter was constructed as a 'rPU' module from the iGEM plasmid

K1085014 using a suite of primers designed to domesticate the amplicon prior to being cloned into the pL0V-PU plasmid (Table 3-1).

Table 3-1 – Table of primers used to amplify and domesticate the IPTG inducible rPU promoter for rhizobacterial synthetic biology.

Primer number	**Primer name**
oxp1300	Plac-rPU -A-Fwd
oxp1301	Plac-rPU -A-Rev
oxp1302	Plac-rPU-B-Fwd
oxp1303	Plac-rPU -B-Rev
oxp1304	Plac-rPU-C-Fwd
oxp1305	Plac-rPU -C-Rev
oxp1306	Plac-rPU-D-Fwd
oxp1307	Plac-rPU -D-Rev

At the same time a Plac 'PU' part was constructed from the same template plasmid (K1085014) using primers oxp1013 and oxp1014. A LacI 'SC' module was amplified using primers oxp1320 and oxp1321. An anhydrous tetracycline (aTc) inducible promoter 'rPU' part was constructed from the iGEM plasmid K1132043 by amplification and domestication using a suite of primers (Table 3-2). A TetR 'SC' part was constructed from the same template plasmid (K1132043) by amplification using primers oxp1025 and oxp1026. A Ptet 'PU' part was also constructed using primers oxp1029 and oxp1030. A 2,4-diaceylphloroglucanol (DAPG) inducible 'rPU' part was constructed by PCR amplification from

Pseudomonas fluorescens CHA0 genomic DNA using a suite of primers (Table 3-3).

Table 3-2 - Table of primers used to amplify and domesticate the aTc inducible rPU promoter for rhizobacterial synthetic biology.

Primer Number	**Primer Name**
oxp1030	Ptet-rPU -A-Fwd
oxp1031	Ptet-rPU -A-Rev
oxp1032	Ptet-rPU -B-Fwd
oxp1033	Ptet-rPU -B-Rev

A PhlF 'SC' module was amplified from the same genomic DNA template using primers oxp1023 and oxp1024 and a Pphl 'PU' part was amplified using primers oxp1017 and oxp1018.

Table 3-3 - Table of primers used to amplify and domesticate the DAPG inducible rPU promoter for rhizobacterial synthetic biology.

Primer Number	**Primer Name**
oxp1034	Pphl-rPU-A-Fwd
oxp1035	Pphl-rPU-A-Rev
oxp1955	Pphl-rPU-B-Fwd
oxp1036	Pphl-rPU-B-Rev

To test the induction profiles of the suite of 'rPU' parts, each was cloned ahead of a sfGFP reporter gene and terminated with a DT15S terminator by Golden Gate

cloning (2.5.7 above). The resulting plasmids were then cloned into both *A. caulinodans* ORS571 and *A. brasilense* FP2 by electroporation (2.6.2 above).

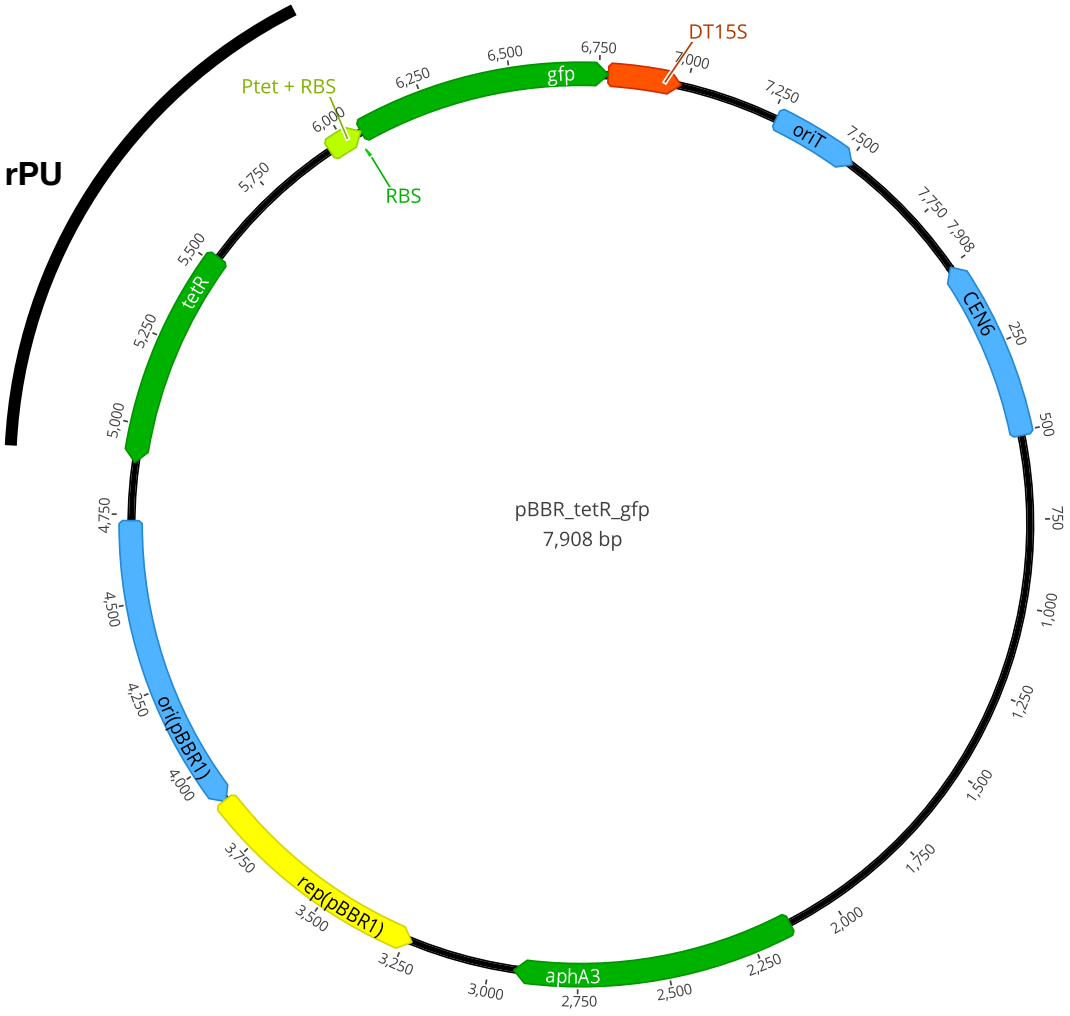


Figure 3-12 – Plasmid map of the aTc induction test plasmid. Note the TetR and Ptet 'rPU' module and sfGFP reporter gene. The other induction test plasmids are identical to this plasmid other than different rPU modules. The aphA3 gene confers resistance to gentamycin for selection and the pBBR1 replication origin functions across a broad spectrum of rhizobacteria.

The test strains were then tested for their induction profiles by a high throughput microwell growth and induction (2.7.3 above). The chemical inducer compounds were bought from Sigma Aldrich (St Louis, USA) and then dissolved in either sterile dH_2O or 100% ethanol (v/v) as per the manufacturer's instructions. Inductions were performed in a range of concentrations: 0 – 10^3 µM for the arabinose inducible 'rPU'; 0 – 10^2 nM for the 3O6-HSL inducible 'rPU'; 0 – 10^6 µM for the IPTG inducible 'rPU'; 0 – 1 µM for the aTc inducible 'rPU' and 0 – 100 mM for the DAPG inducible 'rPU' promoter. Induction occurred when cultures reached an OD595 of 0.5, mid log phase for both rhizobacteria.

The results for the chemically inducible promoters in both *A. caulinodans* ORS571 and *A. brasilense* FP2 showed a variety of dynamic ranges and induction sensitivities. In both strains the arabinose induction did not show a significant increase of sfGFP expression at any concentration of arabinose. There was a slight linear increase of sfGFP induction in *A. brasilense* FP2 but it was not significant across the induction range ($p>0.05$). The 3O6 HSL induction in both strains did not yield any significant increase in the expression of sfGFP at any of the concentrations 3O6 HSL added. The IPTG induction analysis in *A. caulinodans* ORS571 showed a dynamic range of induction of 5x across the range of IPTG added. The addition of 1mM of IPTG yielded the most significant ($p<0.05$) induction of sfGFP, any concentration increase used for induction above this concentration did not yield a significant increase in sfGFP expression. The IPTG induction *A. brasilense* FP2 yielded a comparable dynamic range to that of *A.*

caulinodans ORS571, with a 4.8x increase in sfGFP expression across the range of IPTG concentrations utilised and an optimum induction concentration of 1 mM IPTG. However, the absolute response of sfGFP (REU) in *A. brasilense* FP2 was 6.5x lower than that of *A. caulinodans* ORS571. The aTc induction in *A. caulinodans* ORS571 showed a 18.9x dynamic range of induction across the aTc concentrations utilised with an optimum induction concentration of 100 μM aTc. In *A. brasilense* FP2 the aTc induction dynamic range was observed to be 17.5x with an optimum induction observed with the addition of 100 μM aTc. The dynamic ranges in both *A. caulinodans* ORS571 and *A. brasilense* FP2 were not significantly different ($p>0.05$) but there was a difference ($p<0.05$) between the absolute sfGFP expression (REU) with *A. brasilense* FP2 being 1.8x higher than *A. caulinodans* ORS571. The DAPG induction in *A. caulinodans* ORS571 showed a 15.5x dynamic range of induction across the DAPG concentrations utilised with an optimum induction concentration of 1 mM DAPG. In *A. brasilense* FP2 the DAPG induction dynamic range was observed to be 16.2x with an optimum induction observed with the addition of 1 mM DAPG. The dynamic ranges in both *A. caulinodans* ORS571 and *A. brasilense* FP2 were similar but there was a difference ($p<0.05$) between the absolute sfGFP expression (REU) with *A. brasilense* FP2 being 5.2x higher than *A. caulinodans* ORS571. In summary the aTc and DAPG inducible promoter profiles show the highest dynamic ranges across the suite of inducible promoters tested.

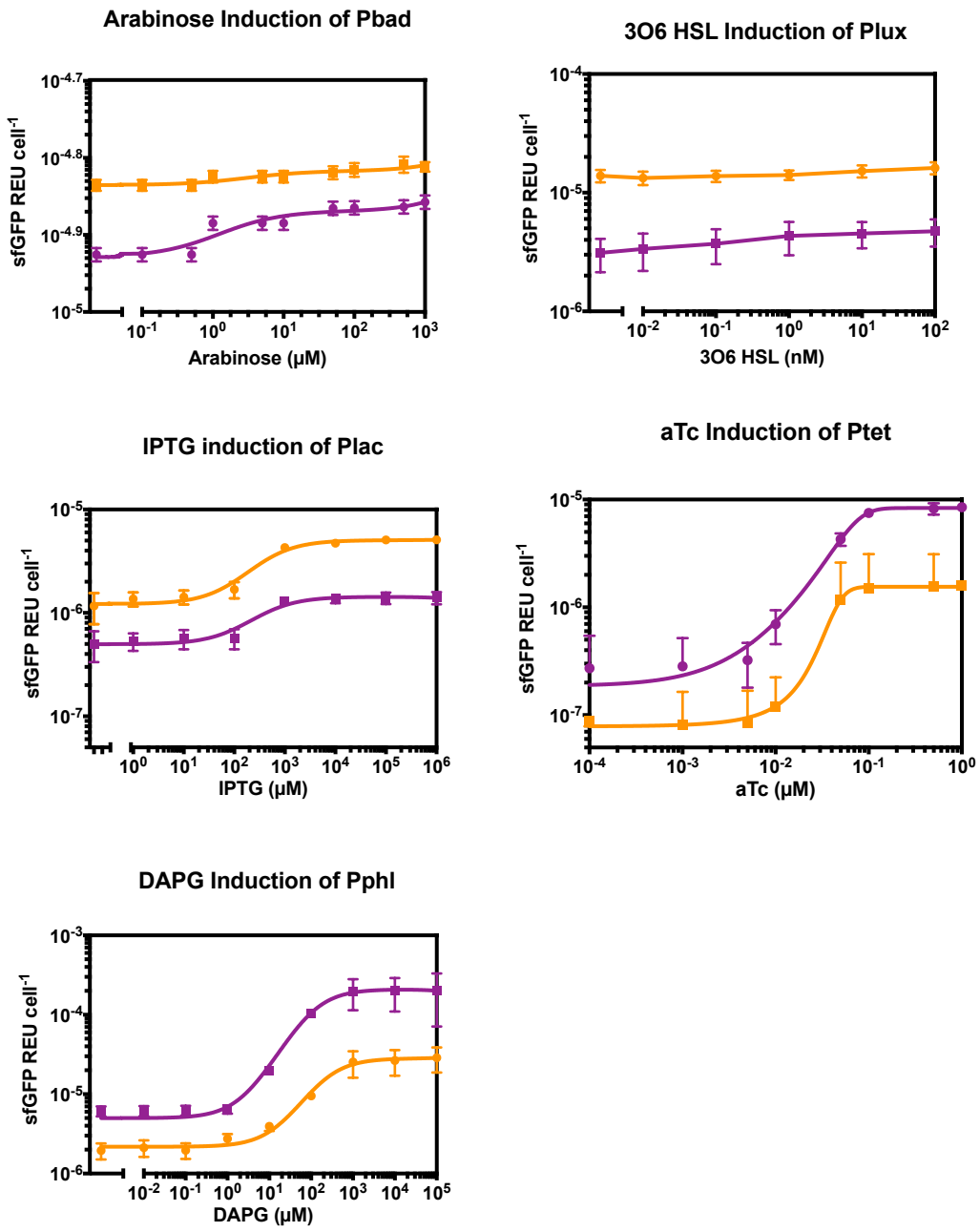


Figure 3-13 – Chemically inducible promoter induction curves and dynamic ranges. Orange = *Azorhizobium caulinodans* ORS571. Purple = *Azospirillum brasilense* FP2. Error bars indicate the standard error of the mean (SEM). N = 3 biological and 3 technical replicates.

3.3.5 Tunable Protein Degradation Analysis

The protein degradation mechanism adopted in this work for the tunable degradation of expressed proteins is based on the endogenous ssrA system found in many species of bacteria. Degradation of ssrA proteins is a central quality control mechanism that prevents truncated and misfolded, non-functional proteins from populating the cell. In *E. coli,* the model synthetic bacterial species, the ATP-dependent ClpXP and ClpAP proteases are thought to be responsible for the proteolysis with two adaptor proteins ClpS and SspB modulating the process. Degradation tags are short peptide sequences that are added to the -COOH terminus of a protein to mark it for degradation. In doing do the half-life of the protein within a bacterial cell is significantly decreased. When a truncated mRNA, and therefore associated polypeptide develops within the cell, the ribosome stalls and becomes stuck as there is no stop codon present allowing the ribosome to terminate peptide synthesis and therefore dissociate. A small specialised stable RNA A (ssrA) molecule rescues the ribosome by adding an eleven codon ssrA degradation tag followed by a stop codon. This allows the ribosome to dissociate and therefore continue to function whilst the truncated polypeptide is degraded by the ClpX – ClpA system. The consensus sequence of the eleven codon ssrA tag has been identified as 'AANDENYALAA' [154]. However, by varying the terminal 3 amino acids, the specificity of the ssrA tag to the ClpX protein can be altered and therefore the half-life of the polypeptide can be tuned to different levels. Three

degradation tags were designed with varying terminal end sequences that I predict will give different degradation rate characteristics (Table 3-4).

Table 3-4 – ssrA degradation tag sequences utilised to tune protein degradation within rhizobacteria.

Degradation Tag Name	**Degradation tag amino acid sequence**
ssrA moderate	AANDENYADAS
ssrA fast	AANDENYNYADAS
ssrA very fast	AANDENYALAA

Both *A. caulinodans* ORS571 and *A. brasilense* FP2 contain homologues to the ClpX, ClpA and ClpS genes when analysed with the genomic DNA search algorithm Basic Local Alignment Search Tool (BLAST; Altschul et al. 1990). Therefore we hypothesise that the ssrA degradation tags we have designed, given previous analysis in *E. coli,* will function in both *A. caulinodans* ORS571 and *A. brasilense* FP2.

The ssrA tags were synthesised and codon optimised for rhizobacteria using the IDT gBlock codon optimisation tool and DNA synthesis platform (IDT, Chicago, USA). The degradation tags were designed as 'C2' -COOH terminal tags that can be cloned in frame onto the end of any coding sequence that had been designed with no 'stop' codon ('SC1' part). These parts were then stably cloning into the pL0V-C2 plasmid.

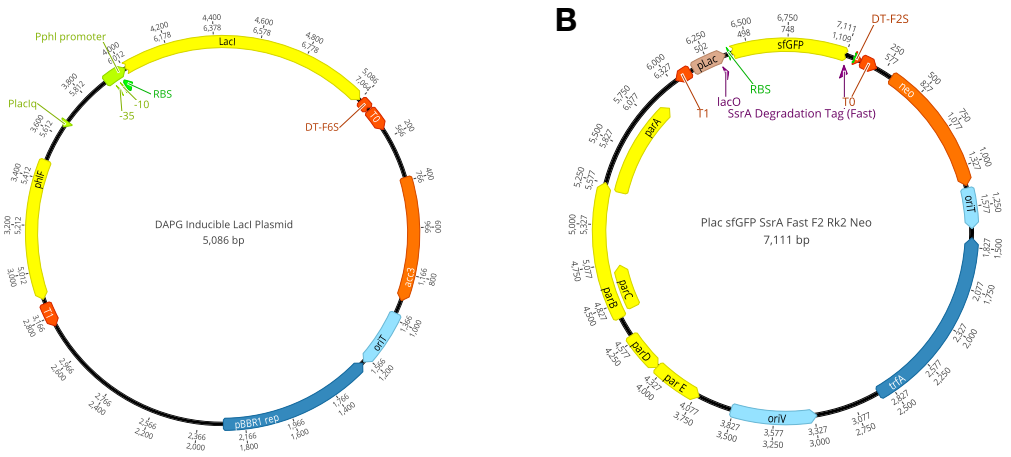


Figure 3-14 – Plasmid maps of the degradation tag reporter plasmid and DAPG inducible LacI repressor plasmid. **A** – DAPG inducible LacI plasmid map. Note the PphI rPU module containing the PhlF transcriptional repressor gene and cognate PphI promoter driving LacI expression. pBBR1 based origin of replication and acc3 gentamycin selection gene. **B** – Degradation tag reporter plasmid containing an sfGFP reporter gene with chosen degradation tag driven by a regulatable Plac promoter. Note that the ssrA Fast degradation tag in this plasmid is swapped out for the moderate or very fast degradation tags for testing their characteristics *in vivo*. RK2 based origin of replication and kanamycin/neomycin selection genes. The *parABCDE* operon allows for the increased maintenance of the plasmid in culture.

To analyse the degradation characteristics of the ssrA tags, reporter plasmids with tagged sfGFP reporter genes were created by Golden Gate cloning (2.5.7 above). An sfGFP reporter is expressed from a Plac regulatable promoter and terminated by a DT-25S terminator (Figure 3-14A). The backbone for this assembly, pOGG26 contains the *par* operon allowing for increased maintenance of the plasmid in the culture population. A control plasmid lacing a degradation tag was constructed using the same method. The second plasmid in the analysis pipeline, a DAPG inducible LacI expression plasmid, allows for the repression of

sfGFP expression and therefore degradation analysis of the reporter to be possible. To express LacI, a PphI 'rPU' part drives expression of a LacI repressor gene, terminated by a DT-F6S terminator (Figure 3-14B). The cassette was assembled into the stable backbone pOGG24 allowing for maintenance of the plasmid in the same strain as the reporter plasmid. Both plasmids were transformed into *A. caulinodans* ORS571 and *A. brasilense* FP2 by two rounds of conjugation (2.6.3 above), the reporter plasmids were transformed into the chassis first to check for sfGFP expression from the Plac promoter. The subsequent conjugation allowed for the repression of the sfGFP on the reporter plasmid from the LacI on the repressor plasmid to be confirmed.

To characterise the degradation profiles of the tags, the dual-plasmid system in each strain was analysed by high-throughput microwell growth and fluorescence (2.7.1 above and 2.7.2 above). The strains were grown until mid-log phase of an OD595 = 0.6. At which time 1mM of DAPG was added to each well to induce the LacI repressor expression. After 30 minutes the sfGFP response begins to decrease and the degradation rate of sfGFP per well (ΔsfGFP REU min^{-1}) was calculated. Negative controls into which no DAPG were added allowed for the ssrA tag dependant degradation to be confirmed.

The results from the analysis indicated that the three ssrA degradation tags did confer an increase in degradation of the tagged sfGFP reporter protein. In *A. caulinodans* ORS571 the suite of degradation tags followed the hypothesised

characteristics with the 'very fast', 'fast' then 'moderate' tags conferring degradation rates of 0.265, 0.195 and 0.150 ΔsfGFP REU min^{-1} respectively. This pattern of degradation was also observed in *A. brasilense* FP2 with degradation rates of 0.510, 0.375 and 0.185 observed for the 'very fast', 'fast' then 'moderate' tags respectively. The degradation of the sfGFP reporter with no tag was negligible and not significant (p.0.05).

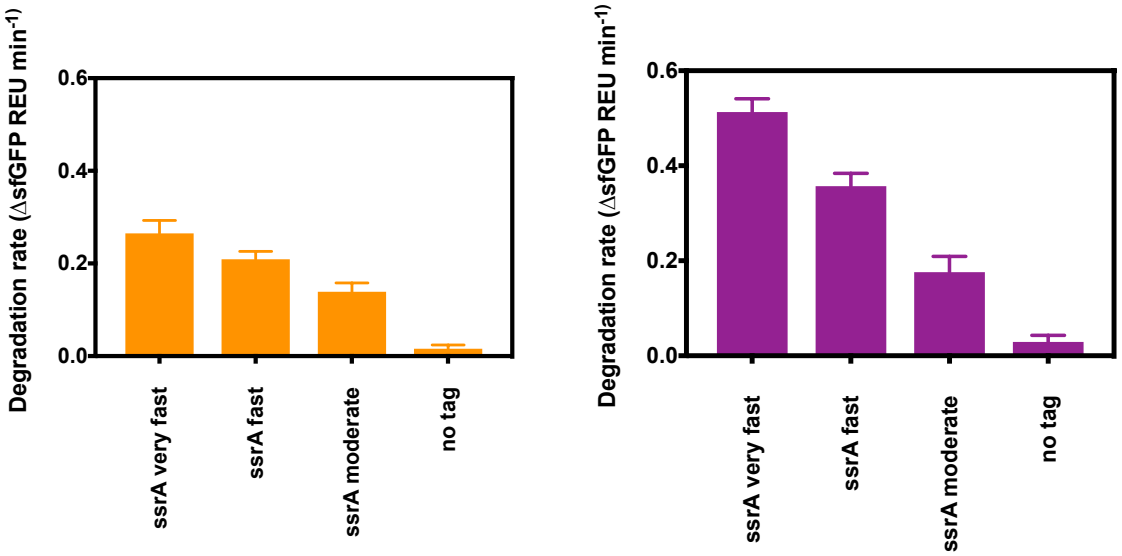


Figure 3-15 – Degradation rate profiles of the suite of ssrA degradation tags. Orange = *Azorhizobium caulinodans* ORS571. Purple = *Azospirillum brasilense* FP2. Error bars indicate the standard error of the mean (SEM). N = 3 biological and 3 technical replicates.

3.3.6 Orthogonal Transcription System (T7 RNAP)

The orthogonal polymerase system adopted for RhizoKit was based on the T7 RNA polymerase (T7 RNAP) from the T7 bacteriophage[156]. T7 RNAP recognises a heterologous promoter sequence to that of the endogenous bacterial RNA polymerase. The consensus sequence of the T7 promoter is TAATACGACTCACTATAG, with transcription initiating at the final 'G' base (Sousa & Mukherjee 2003; Figure 3-16).

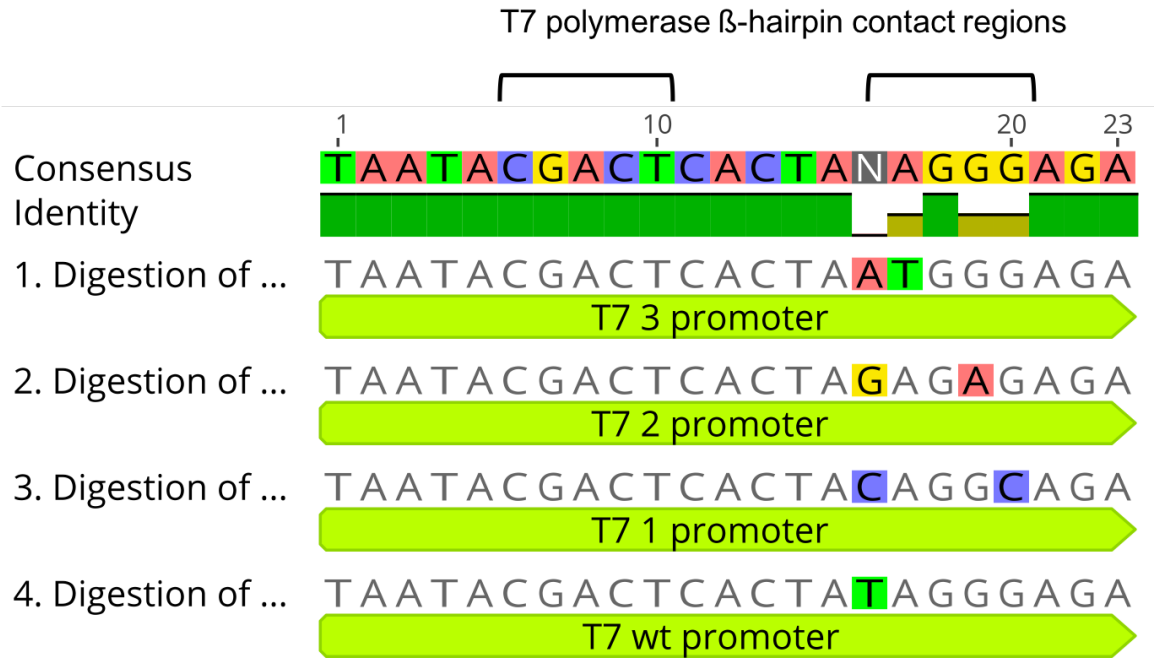


Figure 3-16 – T7 promoter library alignment. Alignment performed with MUSCLE with 8 iterations [114]. Note the SNPs present in the second contact region, proximal to the TSS.

For the T7RNAP system to be tuneable as an effective synthetic biology tool, cognate T7 promoter strengths as well as induction profiles of the T7RNAP are required. To generate a library of different strength T7 promoters, I designed 3 promoters with differing SNPs in the proximal ß-hairpin contact region of the promoter (Figure 3-16). I predicted that SNPs in that region would reduce the specificity of the T7RNAP to bind to the promoter and therefore the level of transcription from the promoter would be lower. The T7 promoter parts were made as 'P' modules by oligonucleotide synthesis and then subsequent annealing (2.5.8 above).

To test both the T7RNAP activity in rhizobacteria, as well as the strength of the 4 T7 promoters, an aTc inducible T7RNAP expression plasmid and T7 reporter plasmid were constructed. The aTc inducible T7RNAP plasmid was assembled by Golden Gate with a Ptet 'rPU' driving a T7RNAP gene synthesised by IDT using rhizobacterial codon optimisation (7.5.7 above; Karsten Temme et al. 2012). The expression cassette was terminated by a DT-F1S terminator and assembled into an RK2 origin-based plasmid for stable rhizobacterial maintenance (Figure 3-17A). The T7RNAP reporter plasmids contain the reporter gene sfGFP driven by each of the four T7 promoters (PT7wt-PT73) and terminated by the DT16S terminator (Figure 3-17B). The cassette was assembled in the pOGG24 vector with a compatible origin of replication (pBBR1) to that of the T7RNAP expression plasmid.

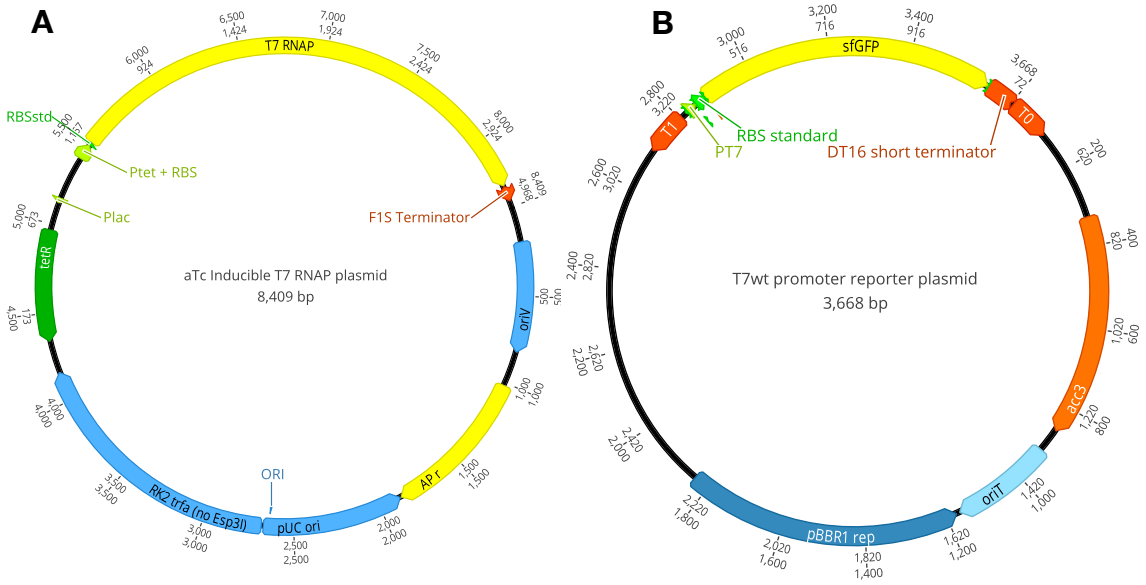


Figure 3-17 – A - Plasmid map of the aTc inducible T7 RNA polymerase system. Note the RK2 broad host replicon and ampicillin resistance gene for selection (Apr). B – Plasmid map of the T7wt promoter reporter plasmid. The other T7 reporter plasmids look identical other than different promoter sequences in the 'P' position. Note the pBBR1 replicon that is compatible in rhizobacteria with the RK2 plasmid origin on the T7RNAP expression plasmid.

Both the T7RNAP expression plasmid and the T7 reporter plasmids were transformed into *A. caulinodans* ORS571 and *A. brasilense* FP2 by two rounds of conjugation (2.6.3 above), the reporter plasmids were transformed into the strain first to check for any sfGFP expression from the PT7 promoters. The subsequent conjugation transformed the strains to contain both a T7 reporter and T7RNAP expression plasmid.

To characterise the T7RNAP expression profile and strength of the T7 promoters, the dual-plasmid system in each chassis was analysed by high-throughput microwell growth and fluorescence (2.7.1 above and 2.7.2 above). The strains were grown until mid-log phase of an OD595 = 0.6. At which time 1mM of aTc was added to each well to induce T7RNAP expression. After 240 minutes the sfGFP response was measured and calculated after normalisation to an empty pOGG024 negative control vector. Negative controls into which no aTc were added allowed for the basal T7RNAP expression levels to be analysed.

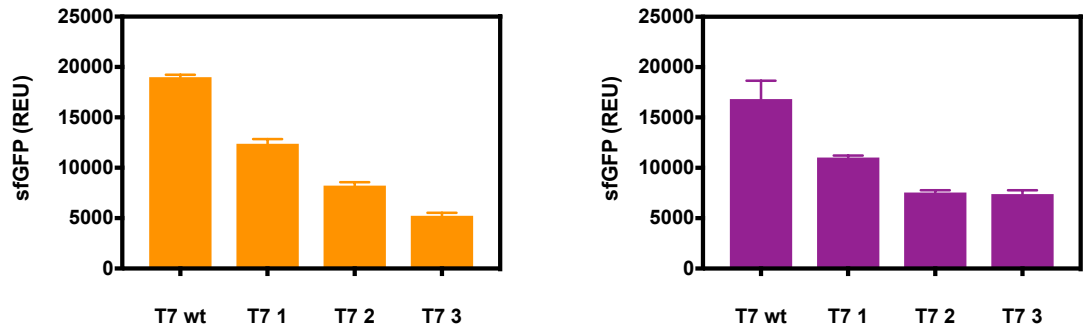


Figure 3-18 – Orthogonal T7 promoter transcriptional strength profiles. Orange = *Azorhizobium caulinodans* ORS571. Purple = *Azospirillum brasilense* FP2. Error bars indicate the standard error of the mean (SEM). N = 3 biological and 3 technical replicates.

The results from the T7 promoter strength characterisation indicated that the T7 wild type promoter (T7 wt) was the strongest out of the library of promoters in both *A. caulinodans* ORS571 and *A. brasilense* FP2. The strengths of the promoters then decreased across the library of promoters. In *A. caulinodans* ORS571 the strengths of the promoter decreased in a stepwise pattern with the T7 wt being 4x more powerful than the T7 3 promoter. The difference between each of the promoters was roughly 6000 REU. The pattern of promoter strengths in *A. brasilense* FP2 was not linear, the T7 wt and T7 1 promoters had a 5500 REU difference in strength between them, but the T7 2 and T7 3 promoters had no significant difference in transcriptional strength ($p>0.05$).

3.3.7 Fluorescent Reporter Library

For complex synthetic biology circuitry as well as *in vivo* biological imaging, multiple fluorescent protein reporters with discreet excitation and emission wavelengths are required. Combinations of fluorescent proteins can also be combined for Förster resonance energy transfer experiments (FRET), for determining protein localisation and interaction dynamics [158]. Several fluorescent protein reporters were chosen including mTabBFP2, eCFP, sfGFP, YFP, sYFP2, mCherry and mRuby3 (Table 3-5). To generate the mTagBFP2 'SC' part the primers oxp1288 and oxp1289 were used to amplify the part using PCR from the iGEM plasmid K592100 (2.5.2 above).

Table 3-5 - Fluorescent Protein Reporter Library

Fluorescent protein reporter	**Reference**
mTagBFP2	[159]
eCFP	[120]
sfGFP	[119]
YFP	[121]
sYFP2	[120]
mCherry	[121]
mRuby3	[122]

To generate the eCFP ‘SC’ part the primers oxp1332 and oxp1333 were used to amplify the part using PCR from the iGEM plasmid E0020 (2.5.2 above). The sfGFP ‘SC’ part was already part of the ENSA Golden Gate database (pOGG037) and we were able to source the plasmid directly from them (John Innes Centre, Norwich, UK). To generate the YFP ‘SC’ part the primers oxp1027 and oxp1028 were used to amplify the part using PCR from the plasmid pAN4248 kindly donated by Dr Alec Nielsen of the Voight laboratory, MIT (2.5.2 above). To generate the sYFP2 ‘SC’ part the primers oxp1331 and oxp1331 were used to amplify the part using PCR from the iGEM plasmid K864100 (2.5.2 above). The mCherry ‘SC’ part was already part of the ENSA Golden Gate database (pOGG037) and we were able to source the plasmid directly from them (John Innes Centre, Norwich, UK). The mRuby3 DNA sequence published by Bajar et al. indicated many BsaI and BpiI sites within the open reading frame of the mRuby3

gene. We therefore codon optimized and excluded the restriction enzymes recognition sequences from a synthesized DNA 'part' from IDT (IDT, Leatherhead, UK; 2.5.2 above).

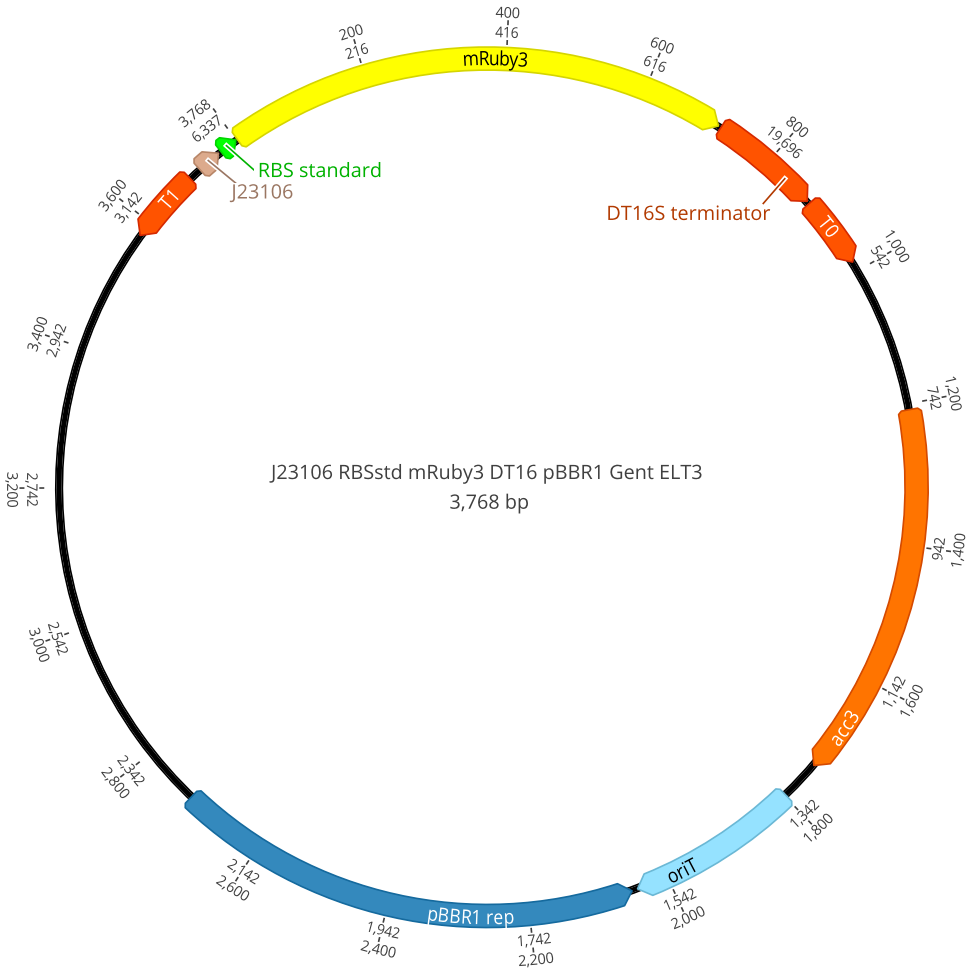


Figure 3-19 - Plasmid map of the constitutive mRuby3 expression plasmid. Note the medium strength J23106 promoter and standard RBS sequence. The other expression test plasmids are identical to this plasmid other than different fluorescent protein SC modules. The acc3 gene confers resistance to gentamycin for selection and the pBBR1 replication origin functions across a broad spectrum of rhizobacteria.

To test the fluorescence response of the reporters as well as determine their excitation and emission spectra in rhizobacteria, a suite of constitutive expression plasmids were constructed (Figure 3-19). The fluorescent reporter proteins 'SC' parts were cloned by Golden Gate into the pOGG024 base vector with the J23106 constitutive 'P' and a standard RBS 'U' (2.5.7 above). The transcriptional unit was terminated by a DT 16S terminator. These plasmids (pOP0468-472) were transformed into both *A. caulinodans* ORS571 and *A. brasilense* FP2 by electroporation (2.6.2 above). Cultures of the test strains were grown to an OD595 of 0.6 before 400 µL of the cultures were loaded in triplicate into a 24-well microwell plate for high throughput fluorescence screening (2.4.3 above). To analyse both the absolute fluorescence response and the excitation and emission spectra, a bandpass from 350 to 700 nm in 2 nm increments was performed with a CLARIOStar Omega Microplate Reader (BMG). Absolute fluorescence measurements were calculated after normalisation to an empty pOGG024 vector, excitation and emission spectra were normalised to 100 as the maxima point for both excitation and emission responses.

The results from the excitation and emission spectra in both *A. caulinodans* ORS571 and *A. brasilense* FP2 indicated that all the fluorescent reporter proteins function in both chassis and reproduce expected fluorescence maxima (Figure 3-20 & Figure 3-21).

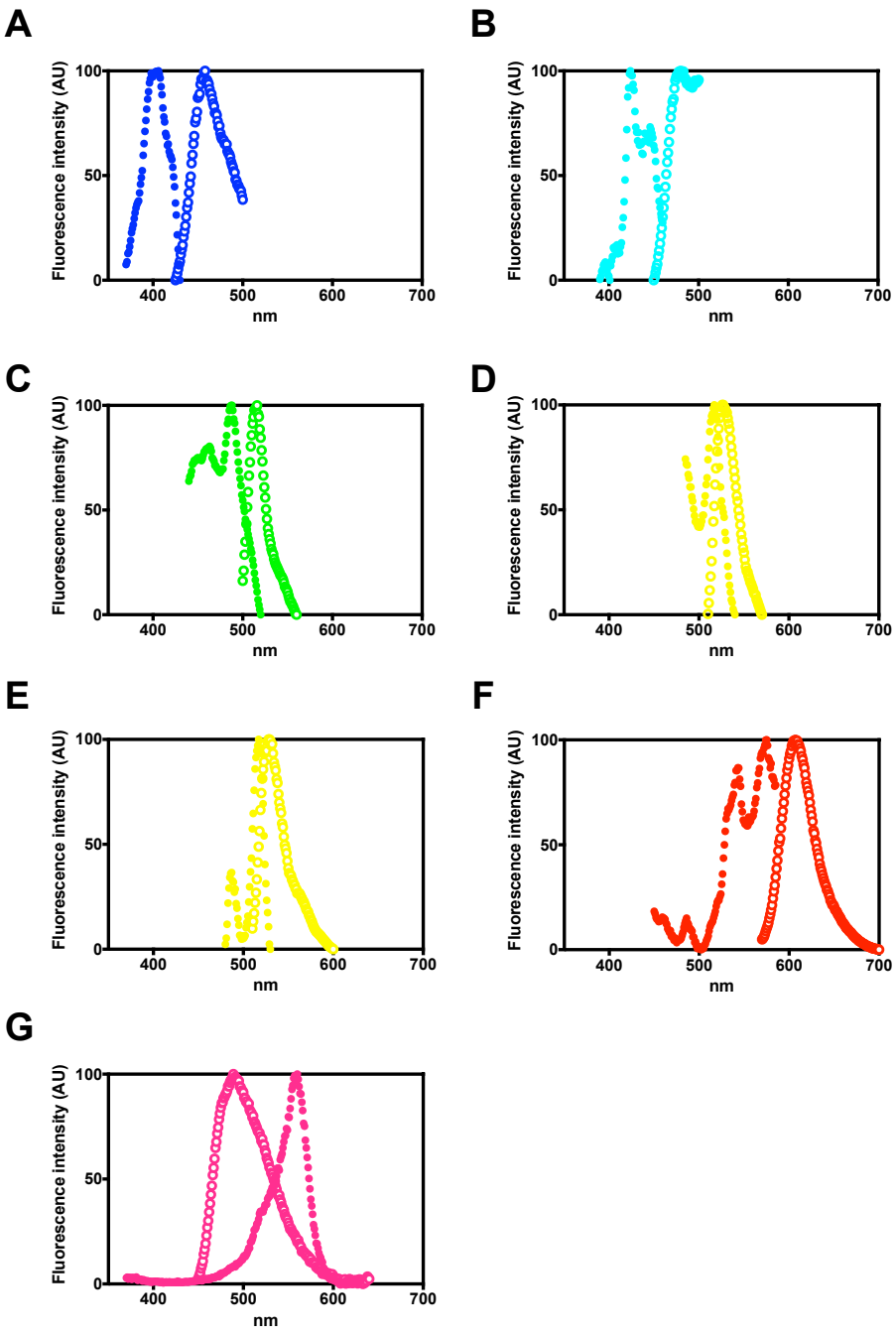


Figure 3-20 – Excitation and emission spectra for the fluorescent protein reporter library in *A. caulinodans* ORS571. Solid icons indicate the excitation spectra and hollow icons indicate the emission spectra. **A** – mTagBFP2. **B** – eCFP. **C** – sfGFP. **D** – YFP. **E** – sYFP2. **F** – mCherry. **G** – mRuby. N = 3 biological and 3 technical replicates.

The fluorescence profiles in both *A. caulinodans* ORS571 and *A. brasilense* FP2 were very similar with only minor differences in the secondary peaks and drop-off response after maxima. There was no significant difference in the excitation and emission maxima across the chassis ($p>0.05$).

Maximum mTagBFP2 fluorescence was observed at 399 nm excitation at an emission of 454 nm. Maximum eCFP fluorescence was observed at an excitation of 433 nm and an emission of 475 nm. Maximum sfGFP fluorescence was observed at an excitation of 485 nm with an emission of 510 nm. Maximum YFP and sYFP2 fluorescence was observed with an excitation of 515 nm and an emission of 527 nm. Maximum mCherry fluorescence was observed at an excitation of 587 nm and an emission of 610 nm. Finally maximum mRuby3 fluorescence was observed at excitation of 558 nm with 5 nm bandpass and an emission of 592 nm with 5 nm bandpass filtration (Figure 3-20 & Figure 3-21).

The absolute fluorescence response of the reporter library indicated that sYFP2 had the highest response with an emission maxima of 225000 AU in both chassis, this was three times higher than the next highest response of mRuby3 of 75000 AU. eCFP had a maximum emission response of 55400 AU; mCherry of 43550 AU; mTagBFP2 of 15000 AU; sfGFP of 14330 AU and YFP of 13100 AU (Figure 3-22).

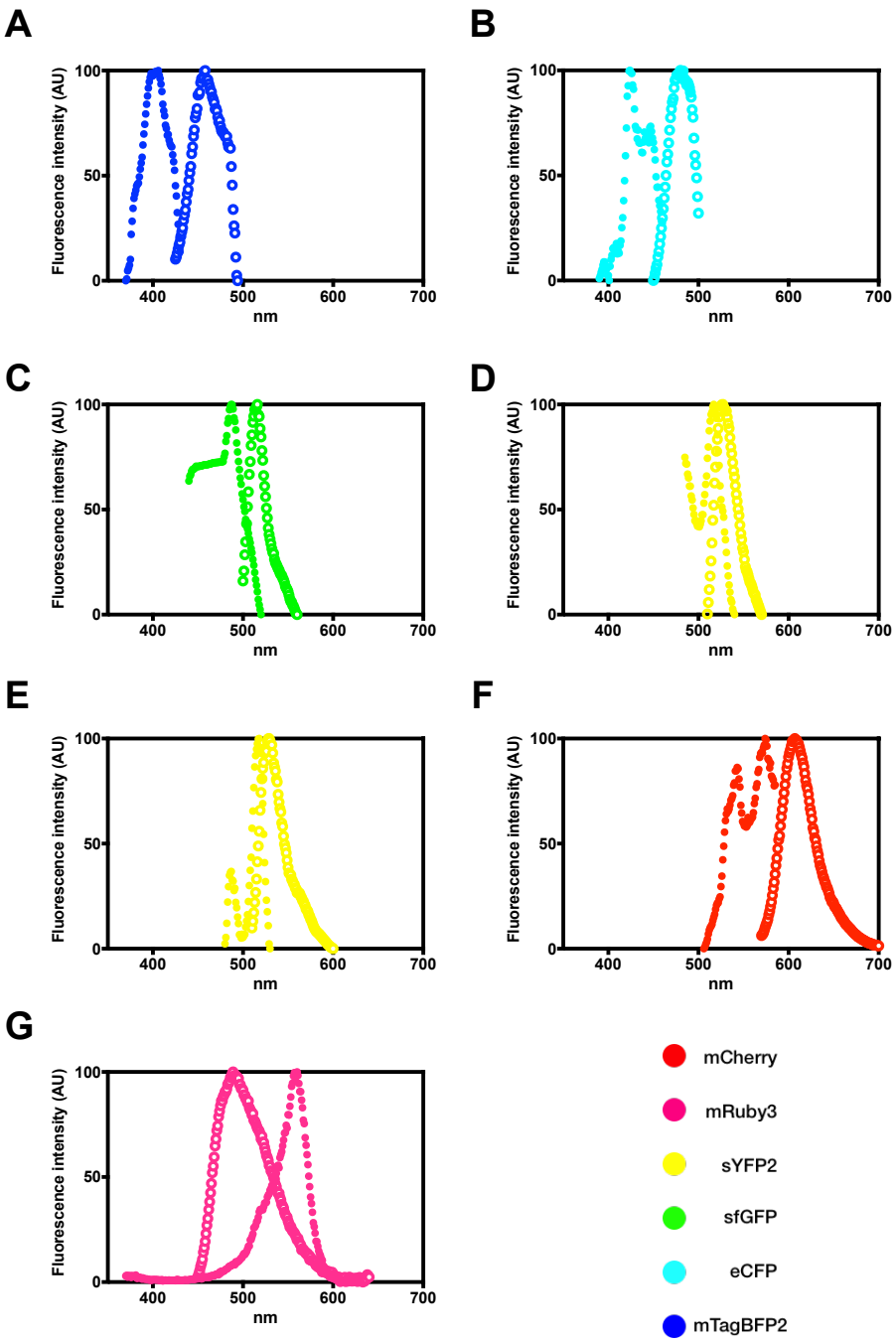


Figure 3-21 – Excitation and emission spectra for the fluorescent protein reporter library in *A. brasilense* FP2. Solid icons indicate the excitation spectra and hollow icons indicate the emission spectra. **A** – mTagBFP2. **B** – eCFP. **C** – sfGFP. **D** – YFP. **E** – sYFP2. **F** – mCherry. **G** – mRuby3. N = 3 biological and 3 technical replicates.

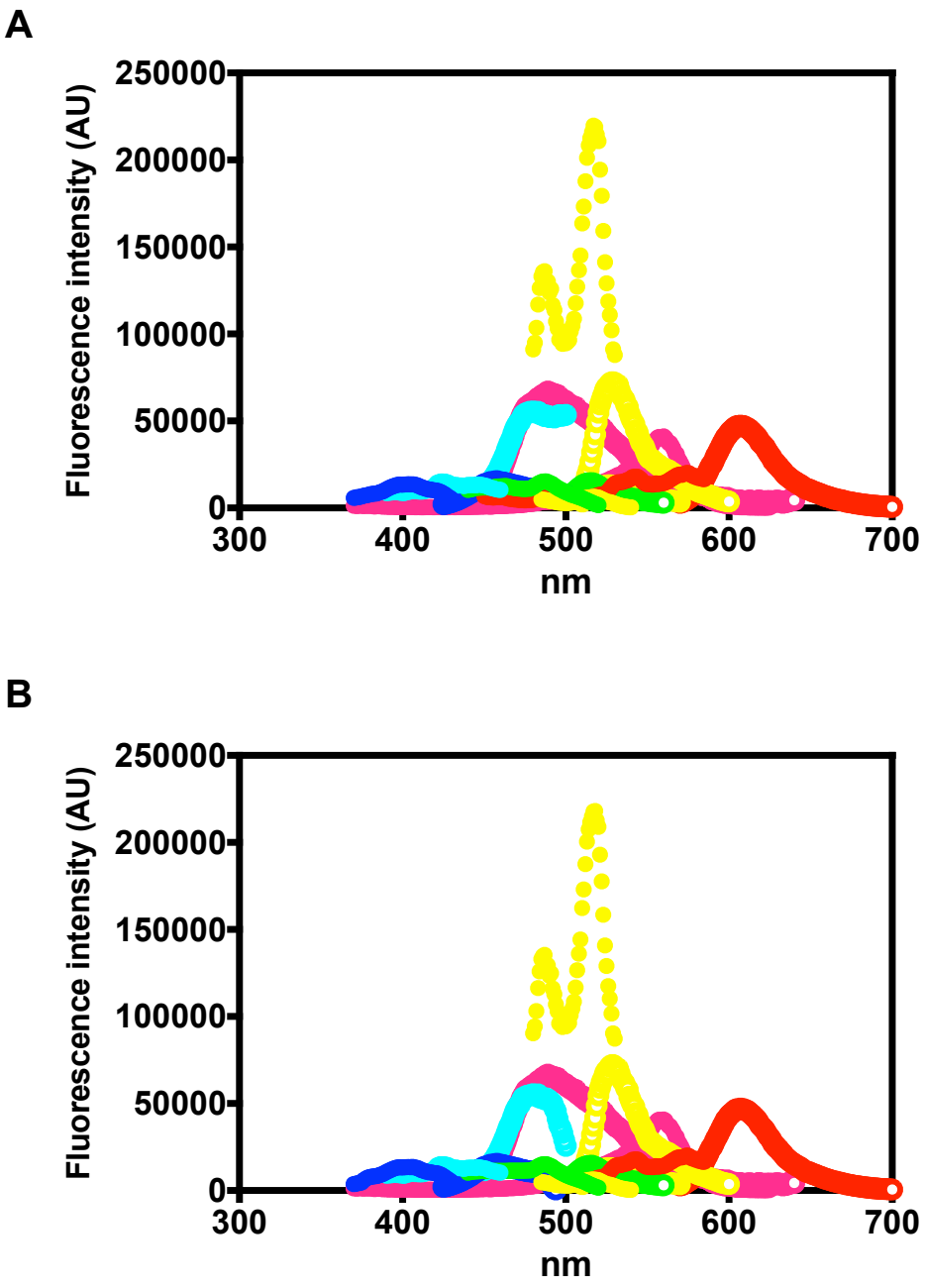


Figure 3-22 – Absolute fluorescence responses of the library of fluorescent protein. **A** – *A. caulinodans* ORS571. **B** – *A. brasilense* FP2. N = 3 biological and 3 technical replicates.

3.4 Discussion

In this chapter I have built and characterised a suite of DNA 'parts' for rhizobacteria that can be combined in different combinations to yield many of the core synthetic biology tools available to other model organisms. I found that many parts designed for other organisms such at *E. coli* function well in both rhizobacteria chassis as well as new parts that have been designed in this study.

The first suite of parts, the constitutive σ70 promoters used to tune transcription in rhizobacterial gene circuits functioned well in all three test chassis. The strength of the J23104 promoter was not significantly different to that of the widely-used Plac promoter indicating the affinity of the σ70 binding to the -10 and -35 regions of the promoter sequence are of similar strength[56]. The variation in promoter strengths across the chassis could be due to the different transcriptional profiles and recognition of the transcription initiation sequences, it may also be interpreted that the folding time of sfGFP in each host chassis is different and that the absolute copy number of the plasmid-borne TEU varies across host chassis[160]. Therefore we cannot directly compare the strengths of each promoter across the different chassis, only that their relative strengths remain intact across the chosen strains, for example J23104 is always the strongest of the constitutive promoters and J23115 the weakest, this is in support of the initial promoter characterisation in *E. coli* [125]. The constitutive promoter suite has predictable characteristics across

both rhizobacterial strains and therefore is a strong tool for TEU design and optimisation of synthetic biology circuits in rhizobacteria.

The second set of parts, the RBSs used to tune the translation of chosen genes in a genetic circuit, also functioned very well in all the test chassis. The variability in the RBS strengths across the library were of a greater magnitude than expected, likely due to the SNP locations being scattered across the entire ribosome footprint and therefore significantly changing the affinity of the ribosome to the DNA molecule. The significantly greater change in RBS strength between the RBS1 and RBS13 sequences is due to the addition of the AAATT sequence immediately following the RBS footprint on the RBS1 DNA sequence. This additional sequence for the RBS1 was expected to enhance the DNA secondary structure to allow translation initiation to occur more readily *in vivo* as demonstrated in previous studies[161]. The RBS13, 22 and 31 sequences did not have this accessory sequence and with the addition of multiple SNPs in the RBS sequence, have decreased translational efficiencies. The RBS suite has predictable translational characteristics in both of the rhizobacteria strains and as such any synthetic TEUs constructed with the parts should have comparable characteristics across a broad rhizobacterial host range.

The generation of a range or terminators with highly heterologous DNA sequences was successful. Although the range of terminator strengths was unexpected in our experiment, the library of strengths is not without merit. Previous studies in

the field by the Che-Lin group at the National Tsing Hua University of Taiwan proposed that terminators could be used as transcriptional filters in complex gene expression operons[162]. Our terminators could be used in the future to control the stoichiometric transcription of genes in a complex operon, when expressed from a common single upstream promoter.

The suite of chemically inducible promoters did not perform as we expected. In traditional inducible gene expression experiments for the past several decades, IPTG has been the most widely-used chemical inducer and has performed well in the *E. coli* chassis. In our experiments the IPTG induction in rhizobacteria was extremely poor and cannot be used as an effective inducer molecule. This is likely due to several reason, the first being that the IPTG molecule is not being actively up taken by the cell by the endogenous lactose permease[163]. Alternative reasons could be that the LacI repressor gene to which IPTG binds and therefore activates, has not folded correctly in the rhizobacteria chassis[52,164]. Both the 3O6 HSL and arabinose inducible promoters did not function in the rhizobacteria chassis. The 3O6 HSL inducible system is derived from the quorum regulatory system of the marine bacterium *Vibrio fischeri* and is a complex and highly specific system[165,166]. There are several reasons why this system does not function well ranging from the inability of the 3O6 HSL to cross the cell membrane to its active degradation by endogenous *aiiA* genes within the cell cytoplasm[167,168]. The arabinose inducible promoter system may well be unfunctional due to similar reasons. Usually the AraC gene transcriptional activator dimerizes, in high arabinose concentrations,

and binds two adjacent operator sequences allowing the Plac promoter to be exposed and therefore for transcription to initiate[169]. It is likely that in the two rhizobacterial strains that this dimerization does not occur and the polymerase binding region remains occluded. Both the 3O6HSL and arabinose inducible systems should not be used further for rhizobacterial engineering in the future without further research and optimization. The aTc and DAPG inducible promoter systems functioned very well in both bacterial strains with comparable induction dynamic profiles. Both systems have been used widely across both bacterial and mammalian chassis, aTc specifically is actively transported into all bacterial species by the ubiquitous OmpF and OmpC porin proteins[170]. DAPG has also been shown previously to be a portable induction system in not only a wide range of bacteria but also in mammalian cell lines indicating its robust functionality in many biological environments[171]. Both aTc and DAPG are used as the inducer molecules of preference moving forwards with this study and should be the two primary modes of chemically inducible gene circuit induction in rhizobacteria.

The protein degradation tag suite functioned well across both rhizobacterial chassis with comparable characteristics in both. As the system is based on the endogenous *ssrA* system found in all bacteria, we were confident that the ClpXP and ClpAP degradation proteins would function in both *A. caulinodans* ORS571 and *A. brasilense* FP2[172–174]. The absolute degradation rate of the sfGFP reporter gene cannot be directly compared across both strains as the sfGFP maturation rates in both strains are not constant[172]. However the profile of degradation rates

across the suite of tags is constant and therefore protein degradation can be accurately tuned in synthetic circuits upon normalization to the untagged target proteins stability *in vitro*.

The fluorescent protein library generated and characterised functioned better than initially predicted. Several studies have shown that the spectral profiles of fluorescent proteins shift between host chassis and are therefore not compatible with precise imaging, FRET, laser-based cell sorting or reporter expression analysis techniques[175]. However I was not expecting the chosen fluorescent proteins to alter significantly in their excitation and emission spectra due to their already well established broad host range consistency[117–119,176]. Due to the complex and robust beta-barrel structure of the fluorophores selected, the excitation and emission spectra of the reporter proteins are highly conserved and were shown to be almost exactly the same across both *A. caulinodans* ORS571 and *A. brasilense* FP2[120]. Therefore the library of reporter proteins can be utilised in future synthetic biology circuits with confidence that their characteristics will be consistent in multiple environments. Combinations of the reporter proteins characterised here, such as mTagBFP2, sYFP2 and mRuby3, can be combined to yield multi-output synthetic biology circuits in rhizobacteria. These circuits can be characterised without the risk of spectral overlap and increased noise in the system.

The characterisation of a suite of synthetic biology parts was successful and allows for the robust base from which to build more complex synthetic biology circuits and functionality in rhizobacteria.

4 Genetic Logic Gate Circuitry in Rhizobacteria

4.1 Abstract

Synthetic genetic logic circuits perform robust computational operations based on freely-diffusing transcriptional regulators within a cell. Logic gate functions can be utilised to allow complex control of desired phenotypes such as nitrogen fixation and phytohormone production in rhizobacterial chassis. In this chapter we design, construct and analyse four different transcriptional regulator based NOT gates. Their orthogonality and the minimization of cross-talk between the gates is presented and confirmed in the rhizobacteria *Azorhizobium caulinodans* ORS571 and *Azospirillum brasilense* FP2.

To demonstrate the adaptability of the NOT gates in a relevant synthetic symbiosis, a unique plant signaling input molecule, rhizopine, was implemented. Rhizopine- dependent induction was integrated into the NOT gate circuitry to control logical functions of three out of the four NOT gates in the rhizobia species *Sinorhizobium meliloti* L530 as well as a rhizopine sensitive engineered *Rhizobium leguminosarum* 3841 chassis.

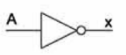
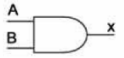
4.2 Introduction

Synthetic biology is by definition an approach to constructing biological systems and circuits with engineering-level precision and robustness. This rationale gave rise to the conceptualisation of standard 'parts' (see 2.5.7 above), hierarchical composition, logical functions and the Computer-Aided Design (CAD) of biological systems and circuits [54,177]. As discussed a vast number of promoters, RBSs and proteins have been standardised and characterised in various significant chassis (3.2 above) as well as scaffolds [178] and noncoding RNAs (ncRNAs; Win & Smolke 2008). Additionally, complex synthetic devices that mimic electronic circuits such as oscillators, toggle switches and riboregulators have been constructed and function successfully in a wide range of chassis [100,180,181]. In the past half-decade these synthetic parts, devices and synthetic circuits have been integrated further either together to form more complex networks, or with the endogenous cellular networks causing network-scale manipulation [182–184].

One of the more important aspects of synthetic biology is the focus on the mechanism of biological logic gate circuitry, increased computational speed and potential biotechnological and agricultural applications. In the field of electronics, a logic gate is a physical device with exhibits a Boolean function based on a set of discreet inputs and outputs of 0 (no input) and 1 (input). The gate executes a consistent logical function on one or more inputs that then produces a single

output of its own (Table 8; Jaeger 1997). For example a NOT gate will output a 1 when the input is 0, and will output a 0 when the input is 1 (Table 4-1)

Table 4-1 – Truth tables, symbols and matrices for the common logic gates present in electronics and synthetic biology. Note that **A** & **B** are inputs and **X** is the output.

Name	NOT	AND	NAND	OR	NOR	XOR	XNOR
Alg. Expr.	$\overline{A}$	AB	$\overline{AB}$	$A+B$	$\overline{A+B}$	$A \oplus B$	$\overline{A \oplus B}$
Symbol							

Truth Table																					
	A	X	B	A	X	B	A	X	B	A	X	B	A	X	B	A	X	B	A	X
	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1
	1	0	0	1	0	0	1	1	0	1	1	0	1	0	0	1	1	0	1	0
			1	0	0	1	0	1	1	0	1	1	0	0	1	0	1	1	0	0
			1	1	1	1	1	0	1	1	1	1	1	0	1	1	0	1	1	1

Initially logic gates were implemented using transistors and diodes that acted as electronic switches, but gates have also been constructed from electromagnetic relays, fluidic systems, optical systems and even mechanical elements. When logic gates are combined in hierarchical and modular systems, devices such as multiplexers, registers, arithmetic logic units and computer memory are created[186]. Logic gates are used for the rapid and robust storage of digital data by construction of several gates in 'Flip-Flop' circuits, forming the central architecture of digital electronics systems in computers and communications systems.

Taking inspiration from electronics rationale and engineering, synthetic biologists have successfully engineered a number of biological logic gates comprising of coding DNA, promoters, transcriptional regulators, RNA polymerase ncRNA, DNA binding elements and small signaling molecules[139,187,188]. Genetic elements such as promoters interact with regulatory proteins such as transcription factors to switch any chosen gene ON of OFF. Inputs can range from chemical inducer molecules to RNS/protein concentrations and even environmental factors such as light intensity, temperature, pH and oxygen tension[189,190]. The long-term goal of synthetic biology is complete biocomputation; to completely reprogram the logical decision making gene networks in a given chassis in order to perform novel functions in areas such as fine chemical manufacture, data storage and agricultural yield enhancement[191]. Biocomuputation *in vivo* is a revolutionary step in biotechnology and will allow for paradigm shifts in the field once the level of reproducibility of genetic circuitry rivals that of electronic computation[192]. In the past decade a number of studies have successfully constructed and characterised biological logic gates in chassis including mammalian, bacterial and fungal cells[187,193,194]. Two groundbreaking studies in recent years were the generation of a numerical counter and the construction of an advanced programming language implemented in *E. coli* that also had automated circuit architecture generation built into the platform[54,195].

4.2.1 NOT Gates

The NOT gate is a Boolean gate that always produces an inverted version of the input as its output. It is the simplest and most useful biological logic gate as it allows for the construction of all the other gates when layered on complex hierarchies. In one example of how a biological NOT gate is constructed, the enzyme RNA polymerase binds to a promoter sequence therefore allowing transcription of a chosen gene. The inducible expression of a transcriptional repressor that interferes with that process causes a NOT gate function. In the example given in Figure 4-1 the input IPTG causes the induced expression of the TetR transcriptional regulator. The helix-turn-helix DNA binding domains of TetR recognise the 15 bp sequence of the TetA operator sequence within the promoter sequence of the Ptet promoter. This causes a widening of the major groove and kinking of the DNA preventing RNA polymerase from binding thus preventing transcription initiation of the sfGFP reporter gene. In summary in the example given here, the addition of IPTG (input of 1) stops the expression of the sfGFP reporter gene (output of 0).

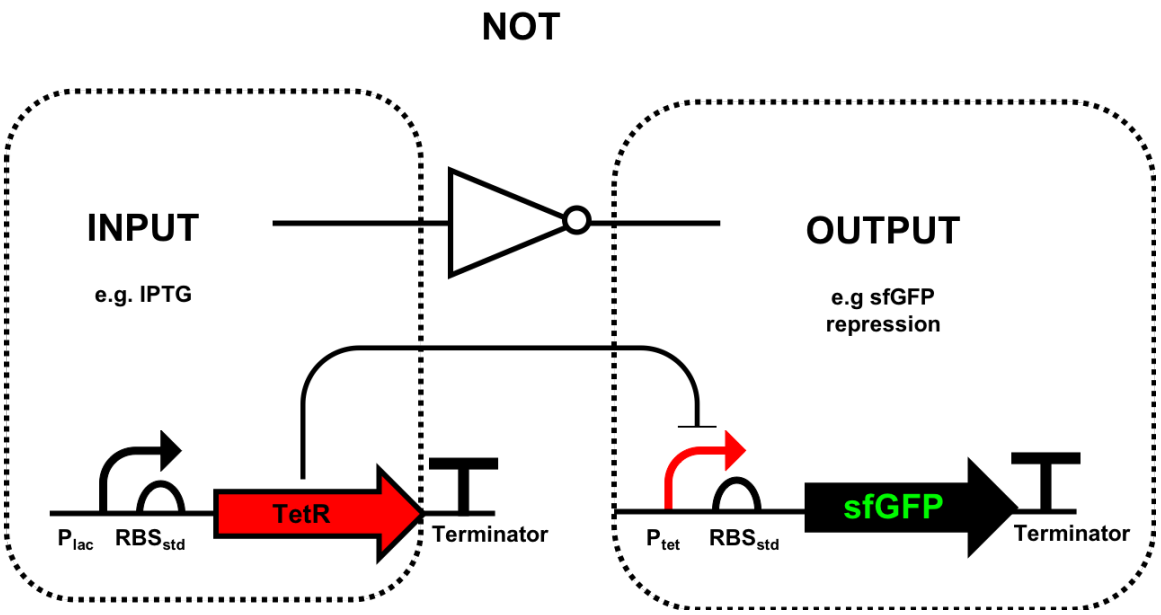


Figure 4-1 – Architecture of a biological NOT gate constructed from the transcriptional repressor TetR and its cognate promoter Ptet. Note that the input in this example is IPTG and the output is the expression of the sfGFP reporter.

Recently, a part mining exploration has been performed across 82,000 members of the EMBL-EBI prokaryotic transcriptional regulator database yielding a library of 73 TetR family repressors[187]. From this library 16 were identified that gave strong repression profiles of >200-fold with little to no cross-talk between the regulators observed. Each repressor and its cognate promoter was converted into a biological NOT gate, the combination of which yielded more than 10 different logic gates including NAND and AND gates.

Logic gate functions could be powerful tools in the engineering of rhizobacteria. As described above the nitrogen fixation regulatory networks in diazotrophic

rhizobacteria are incredibly complex and offer a significant challenge in re-engineering (1.5.1 above). Over 16 core *nif* genes are expressed under the tight regulatory control of numerous regulators, therefore allowing the highly energy-dependent nitrogen fixation process to occur[23]. We hypothesis that through the rational refactoring of the global regulators of nitrogen fixation (GlnA and NifA) could allow for the synthetic control of the process and subsequent ammonia release into the rhizosphere. Additionally phenotypes such as the expression of *nif* genes only under low oxygen tension and in the proximity of a specific plant signal (1.6.1.1 above) would allow efficient crop-specific plant growth promotion effectors to be constructed. In this section we design, construct and validate several logic gate components in the two agriculturally-relevant rhizobacteria *Azorhizobium caulinodans* ORS571 and *Azospirillum brasilense* FP2. The orthogonality of the logic gates, their dynamic range of repression and the effect on the cells viability are all analysed and presented.

4.3 Results

4.3.1 Orthogonal Transcriptional Repressor and Promoter Selection

The selection of transcriptional repressors for use in rhizobacteria was influenced by studies in the synthetic biology mainstay chassis *E. coli.* As with the previous section, many tools designed for prokaryotic function are based in *E. coli* and hence represent the best base from which to develop rhizobacterial logic gates (3.2 above). Once of the most implemented and well understood transcriptional repressors is the LacI repressor from the *lac* operon. Utilised for several decades in biotechnology as the 'go-to' transcriptional repressor, this was the first transcriptional repressor selected. When present *in vivo,* and without the presence of lactose or IPTG, the LacI repressor protein forms a tetramer that binds to two out of three corresponding operators adjacent to the Plac promoter. This causes a conformational change of the DNA into a loop structure, obscuring the RNA polymerase binding sites and therefore repressing transcription of genes downstream of the Plac promoter[164,196].

The second transcriptional repressor selected was the CIR repressor of bacteriophage lambda (∅λ), used in several of the premier groundbreaking synthetic biology studies [100,180]. CIR is a transcriptional repressor that allows ∅λ to establish and maintain latency in a host bacterial cell. The CIR protein is a

homooctamer that binds to two OL operators and two OR operators together, therefore forming the intervening DNA into a loop. This conformational change of the DNA prevents RNA polymerase from binding to and therefore initiating transcription from the Pcl promoter [197].

The final two transcriptional regulators selected were from the TetR family of transcriptional repressors, well characterised as some of the most effective regulators in prokaryotes[130,187]. The third transcriptional regulator selected was TetR itself, an important regulator of bacterial resistance to tetracycline antibiotic compounds. In nature TetR is responsible for repressing the TetA tetracycline membrane efflux pump that exports tetracycline and other related compounds from the cell. TetR is a homodimer that has helix-turn-helix DNA binding domains able to bind to 15 base pair palindromic sequences of the TetA operator[198]. This process cases widening of the major grooves as well as kinking of the DNA, preventing transcription initiation from occurring from the Ptet promoter[199]. The final transcriptional repressor selected was the PhlF repressor from the 2,4-diacetylphloroglucinol biosynthesis operon of *Pseudomonas fluorescens* Pf-5. In a mechanism similar to that of the TetR repressor, the PhlF protein contains a helix-turn-helix DNA binding domain that recognises the phlA operator within the -10 box of the Pphl promoter[200]. Upon binding to the phlA operator the DNA molecule is folded and transcription from the sigma 70 Pphl promoter is prevented[201].

Golden Gate cloning was selected as the cloning pipeline for the rapid assembly and flexibility it provides (2.5.7 above). Each of the transcriptional repressors was amplified and domesticated as a Golden Gate 'SC' part from either genomic DNA (gDNA) of a target organism or from plasmid DNA by PCR (2.5.2 above). The LacI SC part was constructed by amplifying the LacI coding region from the plasmid pAN4271 (kindly donated by Alec Nielsen of the Voigt Lab, MIT) using primers oxp1320 and oxp1321. The amplicon was then immortalized into the Golden Gate vector poGG072 by level 0 Golden Gate reaction to form plasmid pL0M-SC-lacI (2.5.7 above). The cIR SC part was constructed by amplifying the cIR coding region from the iGEM biobrick plasmid C0051 using primers oxp1336 and oxp1337. The amplicon was then cloned into the Golden Gate vector pGGG072 by level 0 Golden Gate reaction to form plasmid pL0M-SC-CIR (2.5.7 above). The TetR SC part was constructed by amplifying the TetR coding region from the iGEM biobrick plasmid C0040 using primers oxp1025 and oxp1026. The amplicon was then cloned into the Golden Gate vector pOGG072 by level 0 Golden Gate reaction to form plasmid pL0M-SC-TetR (2.5.7 above). The final transcriptional repressor, PhlF was constructed by amplifying the PhlF coding region from gDNA of the strain *P fluorescens* Pf-5, (donated by the MaClean group, University of Oxford) using primers oxp1023 and oxp1024. The amplicon was then cloned into the Golden Gate vector pOGG072 by level 0 Golden Gate reaction to form plasmid pL0M-SC-PhlF (2.5.7 above). The cognate promoters were constructed as PL0M-PU modules by oligonucleotide annealing (2.5.8 above). Sequences were derived from the same template DNA as the transcriptional regulators above.

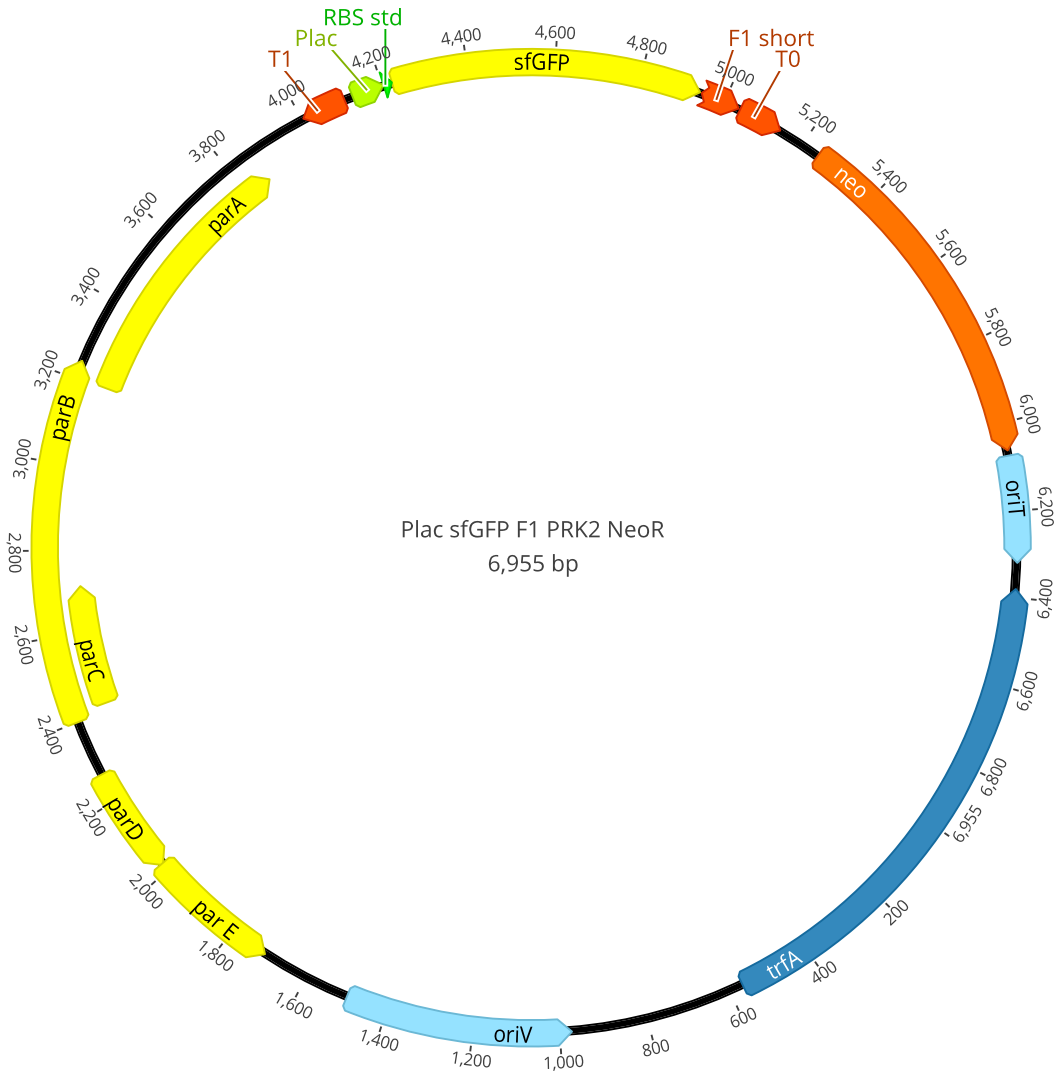


Figure 4-2 - Plasmid map of the Plac reporter plasmid (pOPS1011). The other cognate promoter reporter plasmids (pOPS0457 (Ptet), pOPS0589 (Pphl) and pOPS0590) are identical to this plasmid other than different ‘P’ modules. The neo gene confers resistance to kanamycin and neomycin for selection and the RK2 replication origin (TrfA replication initiator protein and oriV) functions across a broad spectrum of rhizobacteria.

To test the repression profiles of each of the transcriptional repressors in the two rhizobacterial chassis (*A.caulinodans* ORS571 & *A. brasilense* FP2) a dual plasmid system was employed. Construction of the genetic circuitry was performed using characterised parts from the RhizoKit toolbox (see chapter 3 above).

The first plasmid was termed a 'reporter' plasmid and contained a sfGFP reporter protein expressed under each of the four repressible cognate promoters selected; Plac, Pcl Ptet and Pphl (Figure 4-2). Each of these four plasmids (pOPS1011, pOPS0457, pOPSpOPS0589 & pOPS0590) were constructed by level one Golden Gate reaction, the promoter P part was added to the reaction along with an RBSstd U part (pL0M-U- RBsstd), sfGFP SC part (pL0M-SC-sfGFP) and F1S terminator part (pL0M-T-F1S) to insulate the transcriptional unit. The backbone for all the reporter plasmids was pOGG026, an RK2 origin of replication plasmid (4-7 copies per cell) with kanamycin/neomycin selection and the *par* gene operon for enhanced maintenance of the plasmid in culture (2.5.7 above).

The second set of plasmids in the repression analysis was the constitutive transcriptional regulator expression plasmid, assembled by Golden Gate level 1 reaction from a suite of parts (2.5.7 above). Each of the four transcriptional repressors was expressed from a J23106 constitutive promoter (PL0M-P-J23016) and RBSstd (PL0M-U-RBSstd) with a DT16S terminator (PL0M-T-16S). The backbone of the expression plasmids was pOGG024, a pBBR1 replication of origin (50-70 copies per cell) plasmid with gentamycin selection.

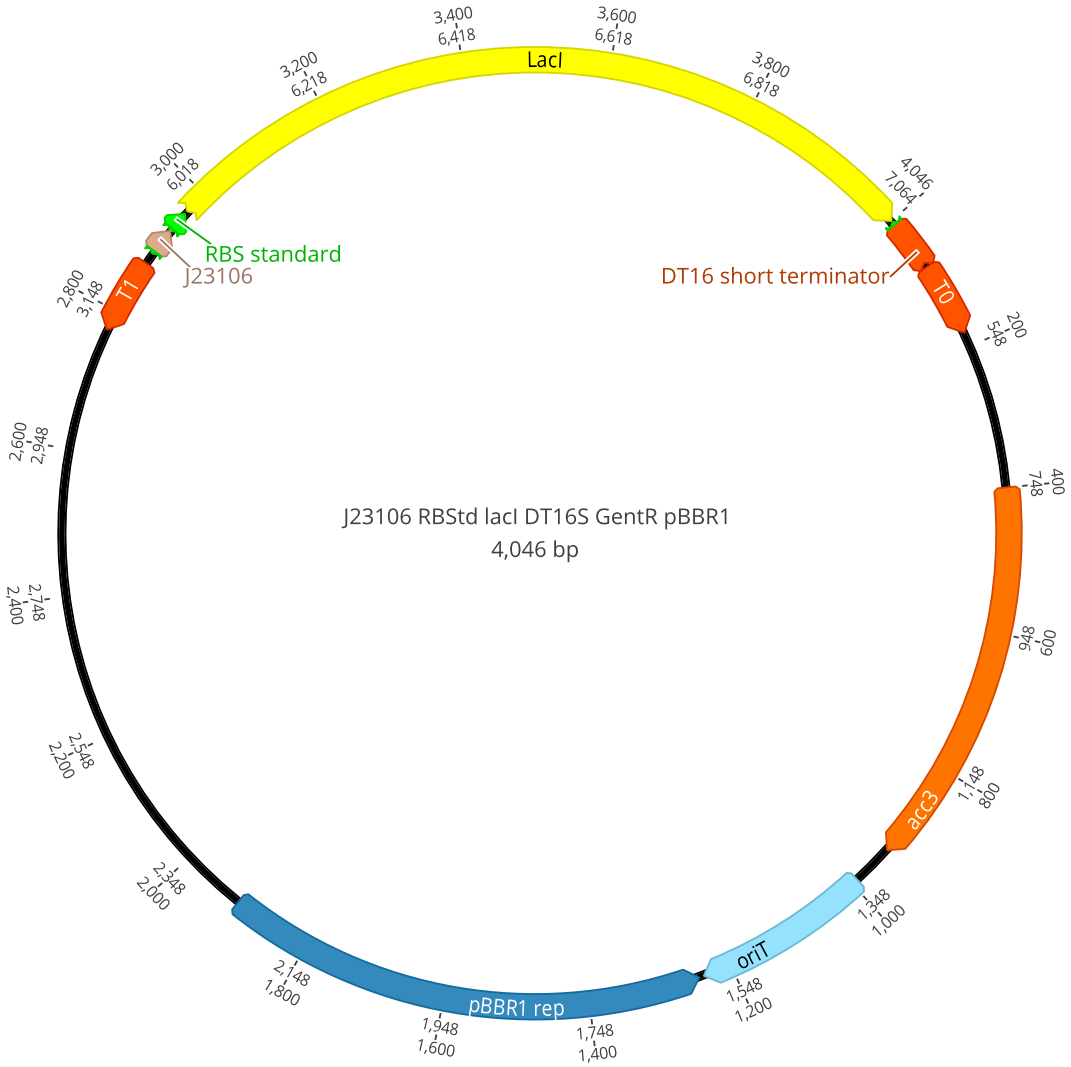


Figure 4-3 - Plasmid map of the LacI constitutive expression plasmid (pOPS1177). The other transcriptional repressor expression plasmids (pOPS1175 (TetR), pOPS1176 (PhlF) & pOPS1178 (CIR) are identical to this plasmid other than different 'SC' modules. The acc3 gene confers resistance to gentamycin for selection and the pBBR1 replication origin functions across a broad spectrum of rhizobacteria.

Each of the reporter and expression plasmids were introduced into both *A. caulinodans* ORS571 and *A. brasilense* FP2 by sequential electroporation of the reporter plasmid (2.6.2 above) and subsequent conjugation of the transcriptional repressor expression plasmid (2.6.3 above). The reporter plasmids were transformed first into the two chassis so that the unrepressed characteristics of the cognate promoters could be analysed. Each of the four repressor expression plasmids were then transformed into each of the reporter plasmid-containing chassis so that orthogonality and repression dynamic analysis could be performed for each of the 16 combinations of transcriptional repressor and cognate promoters.

4.3.1.1 Transcriptional regulator and promoter dynamics

To analyse the unrepressed dynamics of the cognate repressible promoters, the two rhizobacteria strains containing solely the reporter plasmid were grown overnight and then sub-cultured (1/50, v/v) in fresh UMS media with 20 mM succinate and 10 mM NH_4Cl with appropriate antibiotics (2.4.3 above) and grown to an OD595 of 0.6. High-throughput gene expression analysis was then performed with three biological and three technical replicates for each rhizobacteria and cognate reporter plasmid combination. Temporal promoter expression analysis was performed in a microwell plate reader by observing both the OD595 and sfGFP fluorescence (2.7.3 above) every 20 minutes while growing the rhizobacteria under aerobic conditions with an initial cell density of 10^5 cells

per 500 mL microwell (2.7.3 above). An *E. coli* control with the four reporter plasmids was also grown and tested as a positive control in LB media with appropriate antibiotics (2.4.1 above and 2.4.2 above). Fluorescence measurements for the purpose of analysing repression dynamics were taken at an OD595 of 0.5 for *A. caulinodans* ORS571 and an OD595 of 0.8 for *A. brasilense* FP2, and the *E. coli* control when cells were at late log phase.

To determine the repressed characteristics of the cognate promoters, the strains containing both the reporter and transcriptional repressor expression plasmids were grown using the same above methods plus the extra antibiotic gentamycin (2.4.2 above). These dual plasmid repressed reporter strains were assayed concurrently with the primary unrepressed strains using the same high throughput microwell plate method (2.7.1 above and 2.7.3 above). By comparing the relative expression units (REU) of both the repressed and unrepressed systems at the same phase of growth, the repression dynamics normalised to the number of cells in the culture was possible.

The results of the repression dynamics analysis in the control chassis *E. coli* indicated that there was a high variation in the transcriptional strengths of the unrepressed cognate promoters. Ptet and Plac were the two strongest promoters with no significant difference in sfGFP expression at the time point selected ($p>0.05$). The Pcl promoter was observed to be 6X weaker than the Ptet and Plac

promoters with the Pphl promoter the weakest out of those analysed at 8X the strength of the Ptet and Plac promoters.

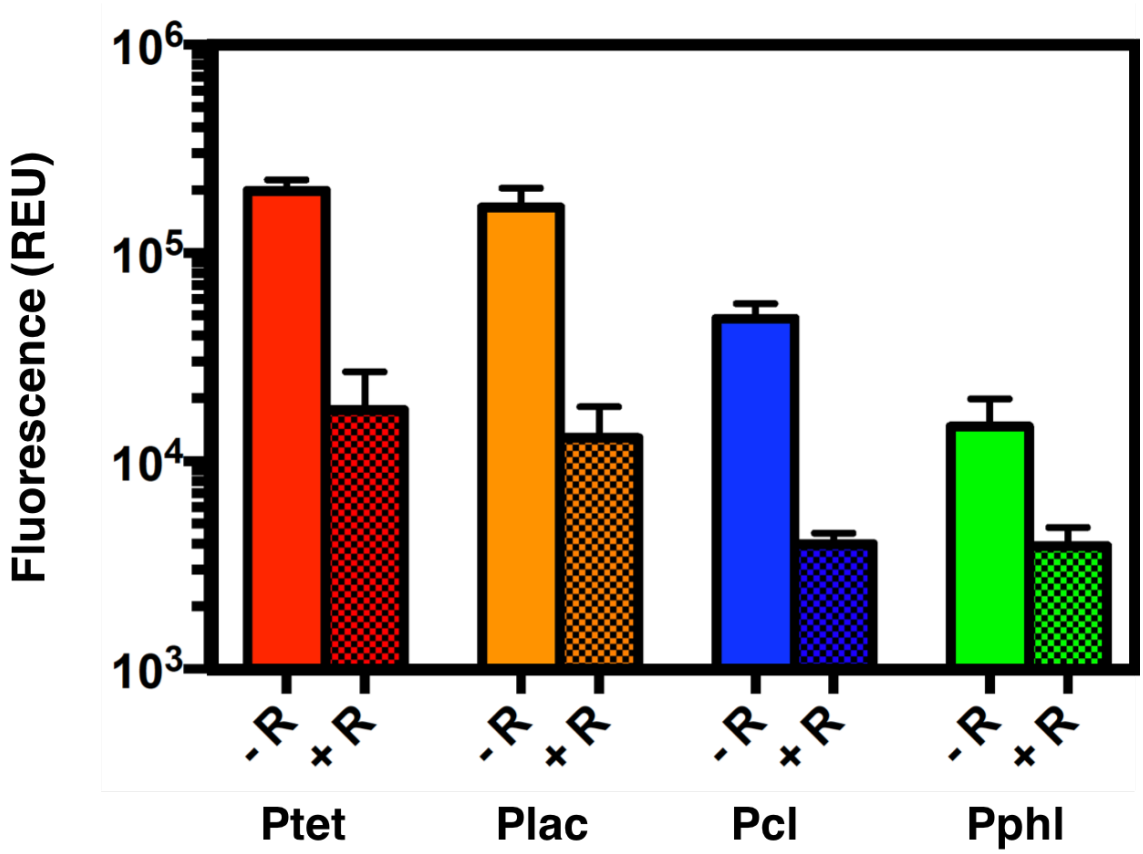


Figure 4-4 – Repression profiles of all the cognate promoter and transcriptional repressor pairs in *E. coli*. -R denotes the test bacteria containing only the reporter plasmid. + R denotes the bacteria contains both the reporter plasmid and transcriptional repressor plasmid. Error bars indicate the standard error or the mean. N = 3 biological and 3 technical replicates.

The repression profiles of the promoter-transcriptional repressor pairs indicated that all four had dynamic ranges of approximately one order of magnitude (ten times less expression of sfGFP under repression). The Ptet and Plac promoters

both had dynamic ranges of 10.1X, the Pcl promoter of 10.9X and the Pphl promoter of 8.3X.

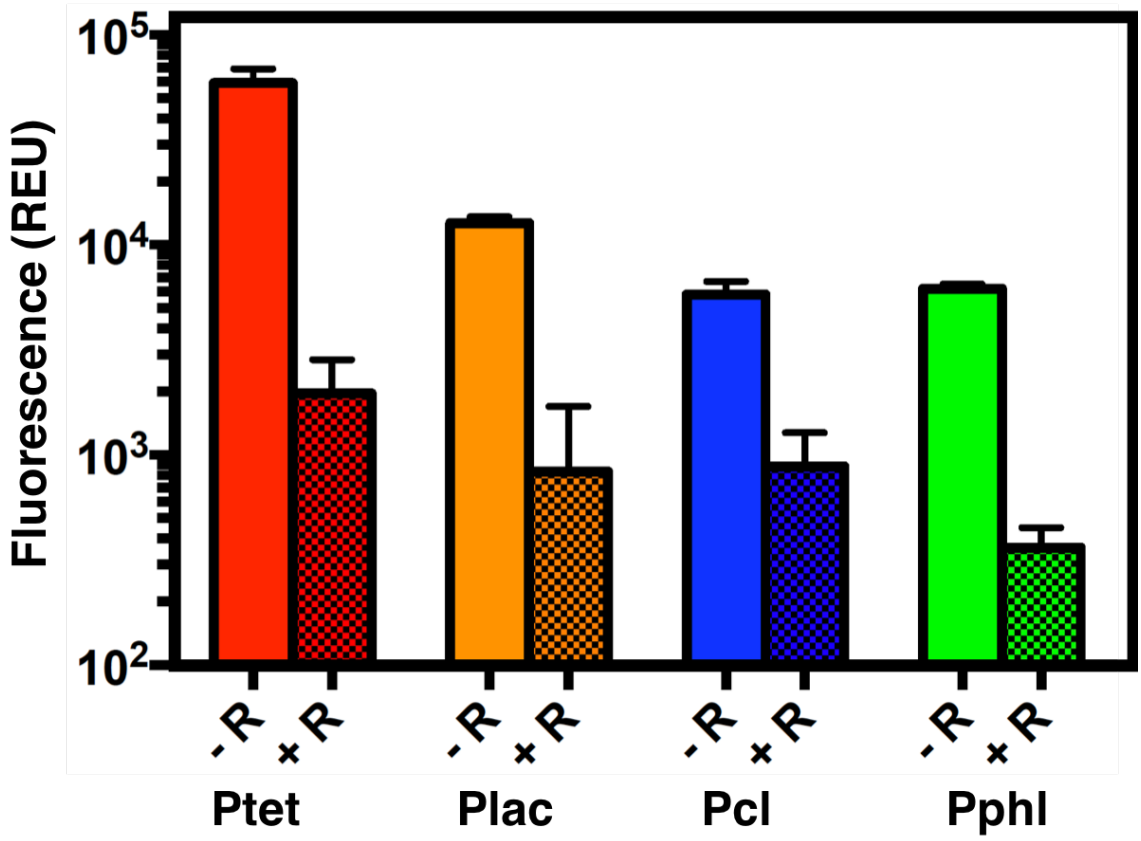


Figure 4-5 - Repression profiles of all the cognate promoter and transcriptional repressor pairs in *A. caulinodans* ORS571 -R denotes the test bacteria containing only the reporter plasmid. + R denotes the bacteria contains both the reporter plasmid and transcriptional repressor plasmid. Error bars indicate the standard error or the mean. N = 3 biological and 3 technical replicates.

The results of the promoter transcriptional analysis in the rhizobacteria chassis *A. caulinodans* ORS571 indicated that there was again variation in the strengths of the unrepressed cognate promoters (Figure 4-5). Ptet was the strongest promoter analysed at 4.5X the strength of the Plac promoter. The Pcl and Pphl promoters

were both 11.2X weaker than the Ptet promoter and were not significantly different in strength to each other ($p>0.05$).

The repression profiles of the promoter-transcriptional repressor pairs in *A. caulinodans* ORS571 indicated that there was far greater variability in dynamic range than in the *E. coli* chassis (Figure 4-5). The Ptet promoter had a dynamic repression range of 15.3X, the Plac promoter a range of 13.9X, the Pcl promoter a range of 7.5X and finally the Pphl a range of 12.9X.

The unrepressed promoter transcriptional strengths in the *A. brasilense* FP2 chassis followed the same relative strength profile as in the *A. caulinodans* FP2 chassis (Figure 4-6). Ptet was the strongest promoter analysed with the Plac promoter only 0.5X weaker. The Pcl and Pphl promoters were not significantly different in unrepressed strength ($p>0.05$) and were observed to be 9.2X weaker than the Ptet promoter.

The results of the repression dynamics analysis in *A. brasilense* FP2 indicated that there was less of a dynamic range across all four transcriptional repressor-promoter pairs than in *E. coli* and *A. brasilense* FP2 (Figure 4-6). The Ptet promoter had a dynamic repression range of 10.5X, the Plac promoter a smaller range of 9.3X and the Pcl and Pphl promoters a range of only 5.1X each. Therefore in the *A. brasilense* FP2 chassis the Pcl and Pphl promoters functioned identically with

regards to their transcriptional strengths and repression characteristics under the experimental conditions implemented.

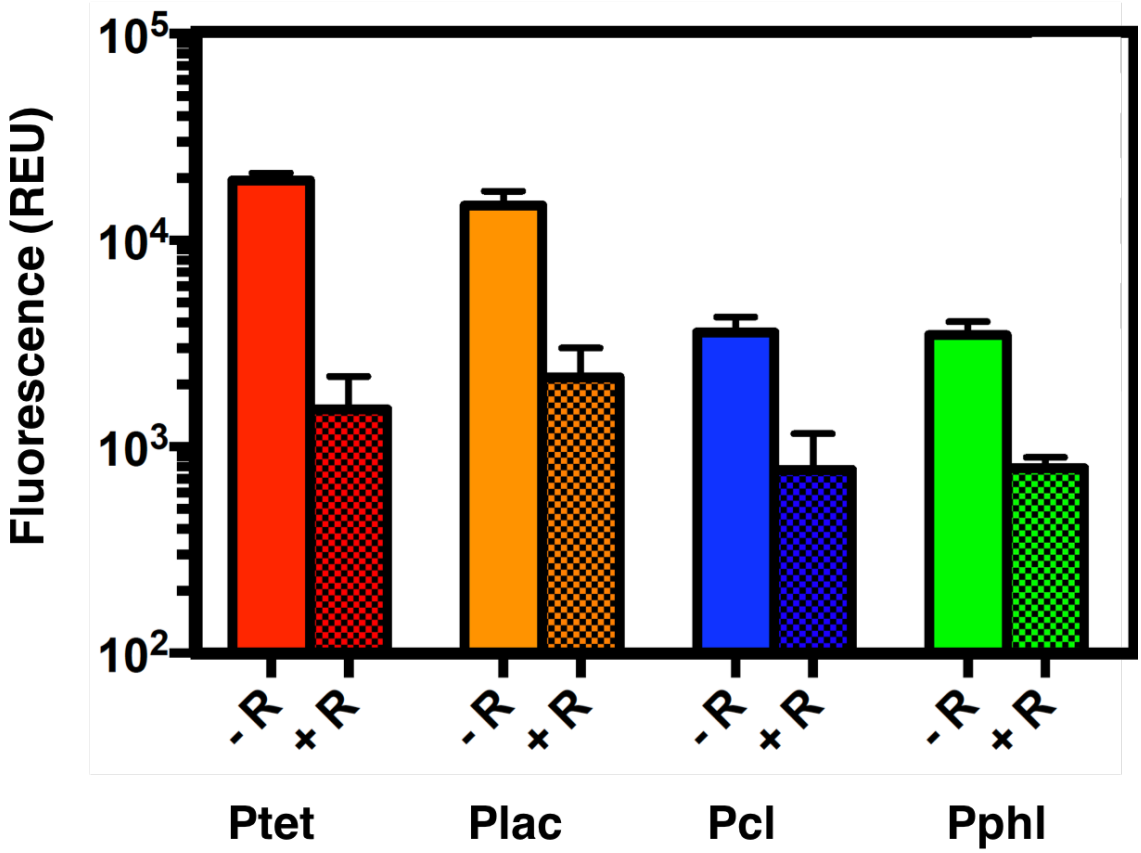


Figure 4-6 - Repression profiles of all the cognate promoter and transcriptional repressor pairs in *A. brasilense* FP2 -R denotes the test rhizobacteria containing only the reporter plasmid. + R denotes the rhizobacteria contains both the reporter plasmid and transcriptional repressor plasmid. Error bars indicate the standard error or the mean. N = 3 biological and 3 technical replicates.

4.3.1.2 Fitness cost analysis

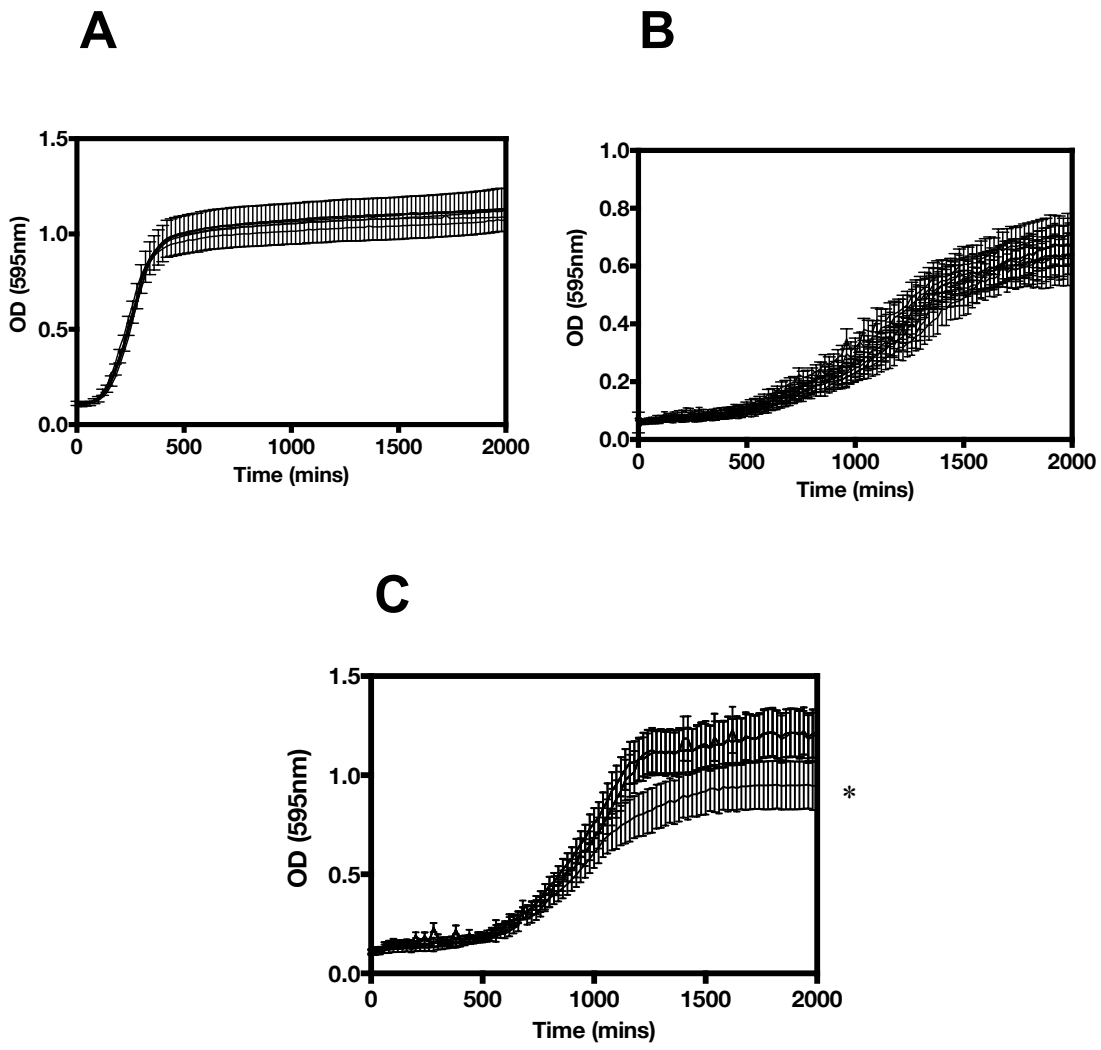


Figure 4-7 -Growth curves of the dual reporter and transcriptional repressor expression plasmids. **A** – *E. coli,* **B** – *A. caulinodans* ORS571 and **C** – *A. brasilense* FP2. N = 3 biological and 3 technical replicates. Error bars indicate the standard error of the mean. ***** PphI reporter plasmid + PhlF transcriptional repressor plasmid in *A. brasilense* FP2. All other reporter and transcriptional repressor expression plasmids are visualised and overlap on the graph above.

Concurrently with the transcriptional and dynamic range analysis (4.3.1.1 above), the growth rates of the chassis with both the single reporter plasmid and dual reporter and expression plasmids was analysed by OD595 measurements (2.7.1 above). A control of a wild-type for each chassis was also included. This was performed to analyse the effect of the expression of both the reporter sfGFP gene from the reporter plasmid and more importantly the cognate transcriptional repressor from the expression plasmid.

The results from the fitness analysis indicated that in *E. coli* there was no significant difference ($p>0.05$) across any of the transcriptional repressors and cognate promoter pairs. There was also no significant difference ($p>0.05$) between any of the transcriptional plasmid pairs and the wild-type *E. coli* bacteria (Figure 4-7 A). The results from the fitness analysis in *A. caulinodans* ORS571 mirrored that of the *E. coli* with no significant difference ($p>0.05$) between the reporter plasmid and expression plasmid systems across any of the pairs assayed. There was also no significant difference ($p>0.05$) between the single and dual plasmid circuits compared to the wild-type *A. caulinodans* ORS571 strain (Figure 4-7 B). The fitness analysis in the *A. brasilense* FP2 chassis indicated that all but one of the combinations of reporter and expression plasmids was not significantly different ($p>0.05$) from one another or the wild type (Figure 4-7 C). However, the combination of the Pphl reporter plasmid the PhlF constitutive expression plasmid indicated a significant ($p<0.05$) reduction in the culture density of *A. brasilense* FP2 after 20 hours compared to the other combinations

of transcriptional promoter and repressor assayed. The maximum density of *A. brasilense* FP2 containing the Pphl reporter plasmid the PhlF constitutive expression plasmids was 22% less than that of the other combinations assayed (Figure 4-7 C).

4.3.2 Transcriptional repressor orthogonality

For the purpose of developing complex logic gates beyond the foundational NOT gate, orthogonality of the four NOT gates is essential. Cross-talk between transcriptional repressor and cognate promoters adds noise to the system that leads to analog rather than digital genetic circuit outputs. To measure all possible cross-reactions, we assayed the activity of each transcriptional repressor expression plasmid against the other four cognate promoter reporter plasmids (4.3.1 above), yielding 16 combinations of promoter and repressor in a matrix. To achieve this the promoter reporter and transcriptional repressor plasmids were co-transformed in all combinations by sequential conjugation (2.6.3 above). The resulting strains were assayed by high-throughput fluorescent reporter gene expression analysis in a microwell plate (2.7.3 above). Three biological and three technical replicates were grown for each of the 16 combinations. The dynamic range of repression for each combination of promoter and repressor was calculated as the ratio between the sfGFP expression per cell for the promoter reporter alone versus the promoter and repressor plasmids being present. The

ratio was then normalised to a maximum of 100 for visualisation of the orthogonality values as a colour gradient. This data was used to construct an orthogonality matrix that indicates the specificity and therefore orthogonality of each reporter and repressor combination.

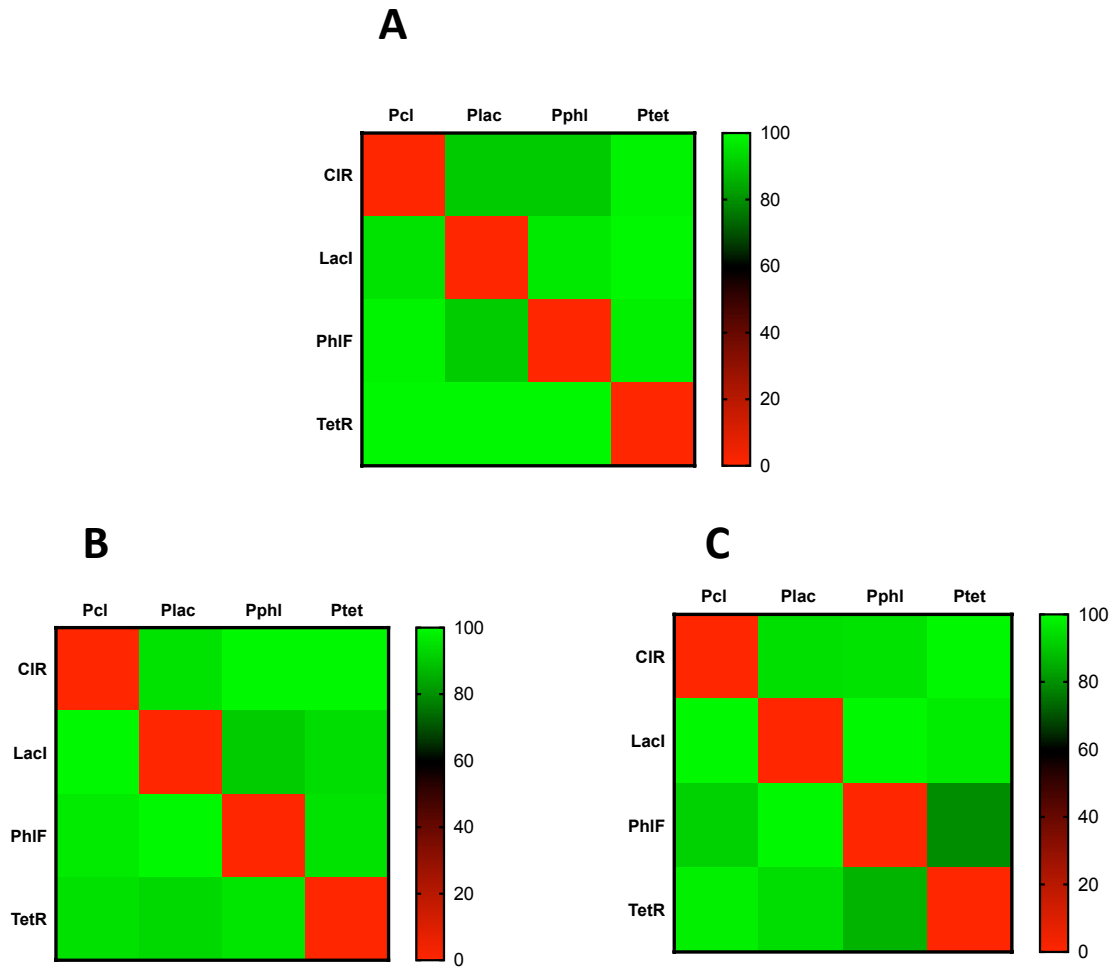


Figure 4-8 – Orthogonality matrices for the NOT gate transcriptional repressor and cognate promoter pairs in all three test strains. **A –** *E. coli.* **B –** *A. caulinodans* ORS571. **C –** *A. brasilense* FP2.

The normalised delta (between 0-100) between the repressed and unrepressed states of the gates were calculated by dividing the mean unrepressed fluorescence with that of the repressed fluorescence. The data represents the mean of three biological and three technical replicates.

The transcriptional repressors assayed were highly orthogonal, all four had negligible cross-reactions ($p<0.05$). Amongst this set of orthogonal transcriptional repressors, the sequence diversity of the DNA-binding regions is highly divergent (Figure 4-9).

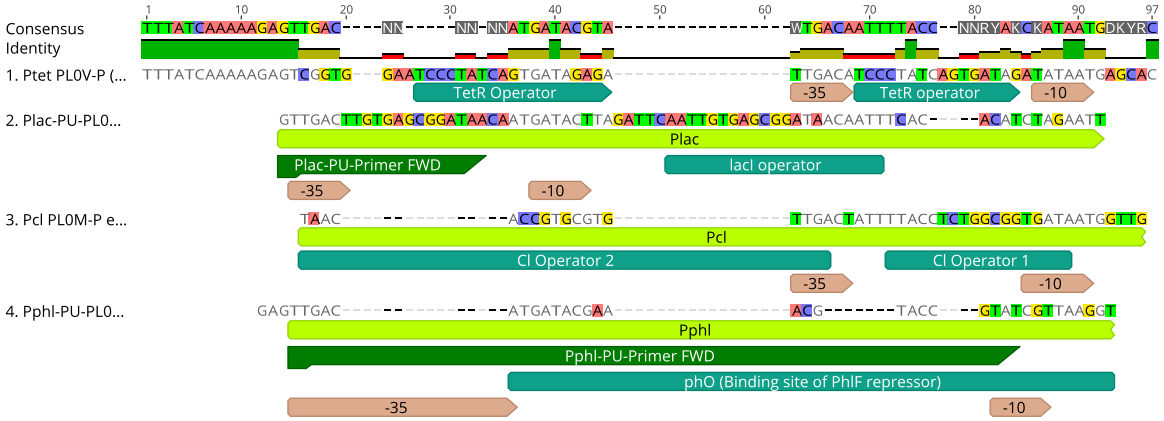


Figure 4-9 – DNA sequence alignment of the four orthogonal promoters and transcriptional binding regions (operator). Alignment was performed using the MUSCLE 3.8.425 alignment algorithm with 8 iterations. Note -35 and -10 regions annotated and lack of coherent sequence alignment across the four DNA sequences.

4.3.3 Inducible NOT gate in Rhizobacteria

To fully utilise the power of logical computation in rhizobacteria, a variable input must be utilised rather than a constitutively active gate. Therefore, we tested the

NOT gate functions in response to the unique plant signaling molecule rhizopine (1.6.1.1 above). At the time of writing there was no circuit able to process the input of rhizopine in *A. caulinodans* ORS571 or *A. brasilense* FP2. However, a rhizopine inducible system has been characterised in both the naturally rhizopine sensitive *Sinorhizobium meliloti* L530 chassis as well as an engineered *Rhizobium leguminosarum* 3841 chassis (Geddes et al., manuscript in preparation). Therefore, we utilised these two chassis for testing the function of three of the NOT gates. The CIR based NOT gate was not included in the analysis pipeline due to cloning issues (4.4 below).

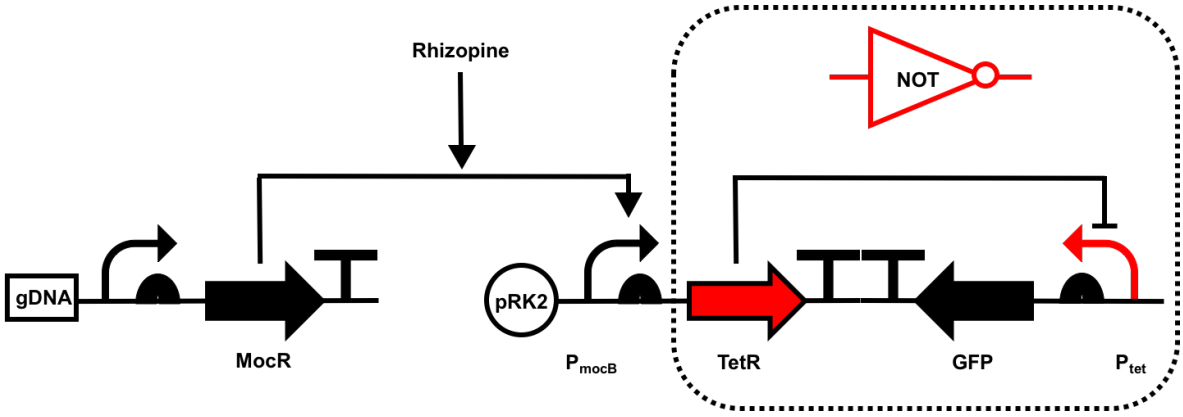


Figure 4-10 – Architecture of the synthetic rhizopine inducible NOT gate circuit. The three transcriptional repressor-based circuits (TetR, LacI and PhlF) were identical to the above circuit save for the replacement of the repressor ORF (above labelled TetR) and cognate promoter (above labelled Ptet). gDNA – genomic DNA integrated MocR transcriptional activator cassette. pRK2 – Plasmid-based NOT gate circuit.

The architecture for the rhizopine inducible NOT gates was based on a two-part system. The first part of the architecture was the constitutive MocR transcriptional activator cassette. In *S. meliloti* L530 the endogenous genomic MocR gene is constitutively expressed from its own PmocR promoter and therefore did not require engineering (Figure 4-10). In the *R. leguminosarum* 3841 the constitutive MocR transcriptional activator from *S. meliloti* L530 was expressed from the native PmocB promoter. The accessory genes IntB and IntC were included in the same integrated construct to enhance the active transport of rhizopine molecules across the cell membrane. The second part of the NOT gate circuit was the transcriptional repressor and cognate promoter pair with a sfGFP reporter (Figure 4-10). The transcriptional repressor is expressed from a rhizopine-inducible PmocB promoter, whilst the sfGFP reporter is expressed from the repressible cognate promoter corresponding to the transcriptional repressor.

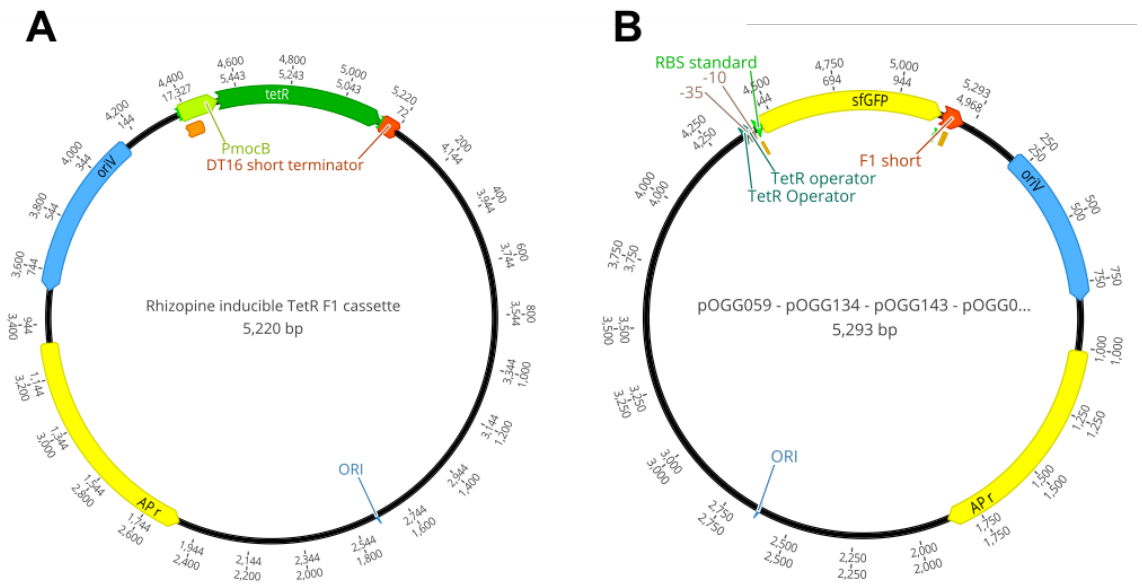


Figure 4-11 – Plasmid maps of the rhizopine inducible NOT gate level one cassettes. **A** – Map of the rhizopine inducible transcriptional regulator (tetR in the above image and lacI and phlF in the other two variants). **B** – Map of the repressible sfGFP reporter plasmid expressed from the cognate promoter (Ptet above but Plac and Pphl in the other two variants.

The rhizopine inducible NOT gate plasmids were constructed by Golden Gate cloning in the pOGG096 backbone for enhanced stability in culture (2.5.7 above). The first level one Golden Gate construct consisted of the rhizopine inducible transcriptional repressors in a pL1V-F1 (pOGG021) backbone (Figure 4-11A). The pL0M-PU-PmocB promoter (pOGG098), pL0M-T-DT16S terminator (pOGG157) and relevant transcriptional repressor pL0M-SC (pOGG169 for TetR, pOGG165 for LacI and pOGG167 for PhlF) were assembled into a level one forward facing module for subsequent level two Golden Gate cloning reaction (2.5.7 above). The second level one Golden Gate plasmid constructed was the sfGFP reporter

expressed from each of the cognate promoters in the pL1V-R2 (pOGG059) plasmid (Figure 4-11B). Each of the cognate promoters (pOGG134 for Ptet, pOGG140 for Plac and pOGG136 for Pphl) was added to the pL0M-SC-sfGFP reporter ORF (pOGG037) and pL0M-T-F1S terminator (pOGG159) in a level one Golden Gate reaction. The level one plasmids with both the transcriptional repressor cassettes (pOGG240-242) and sfGFP reporter cassettes (pOGG243045) were then integrated into the pOGG096 plasmid by a level two Golden Gate cloning reaction (Figure 4-12; 2.5.7 above). The resulting plasmids constituted the final rhizopine inducible NOT gates pOPS1179 (LacI), pOPS1180 (PhlF) and pOPS1181 (TetR).

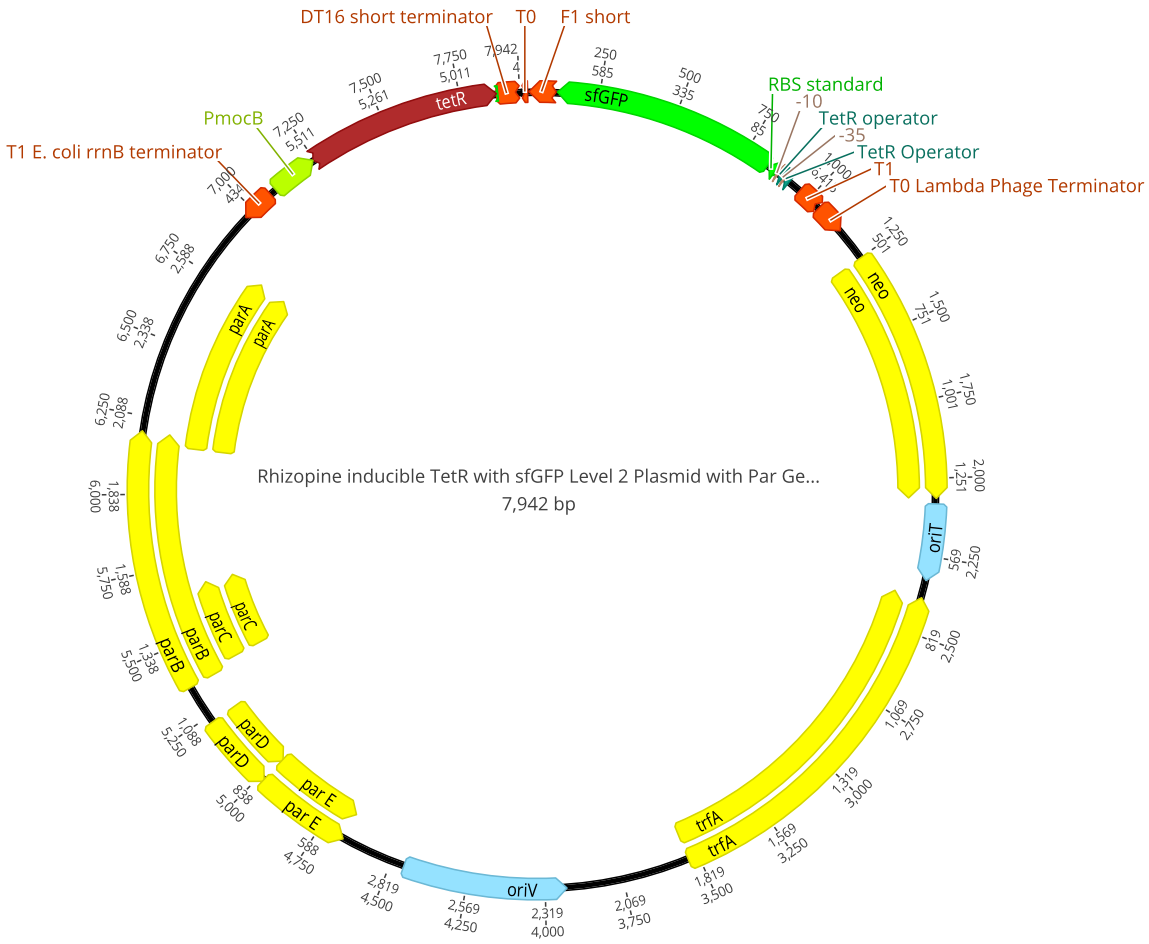


Figure 4-12 -Plasmid map for the rhizopine inducible NOT gate circuit. The LacI and PhlF based NOT gates were identical in architecture to the above plasmid other than the replacement of the transcriptional repressor ORF (above labelled tetR) and the cognate promoter pair (above labelled TetR operator region). Please note the accessory *par* gene operon for enhanced stability in bacterial culture, the *trfA* – oriT RK2 origin of replication regions and *neo* gene conferring kanamycin/neomycin selection in culture.

The rhizopine inducible NOT gate plasmids (pOPS1179-1181) were then transformed into both *S. meliloti* L530 as well as the engineered *R. leguminosarum* 3841 strain by conjugation (2.6.3 above). To test the NOT gate function, a high-throughput microwell growth (2.7.1 above) and fluorescent reporter analysis (2.7.3

above) pipeline was utilised. Both *S. meliloti* L530 and the engineered *R. leguminosarum* 3841 strains were grown at 28˚C with 300 rpm shaking with an initial cell density of $1x10^5$ cells per 500 µl microwell. Each of the NOT gate containing strains consisted of three biological and three technical replicates per data point. Rhizopine induction was carried out at the beginning of the assay with the addition of 10 µM rhizopine (scyllo-inosamine) at t = 0.

4.3.3.1 NOT Gate Functional dynamics

The results from the rhizopine inducible NOT gate function in both *S. meliloti* L530 and the engineered *R. leguminosarum* 3841 strain indicated comparable functionality in both. However, the dynamic ranges of all the NOT gates were far lower than those observed in the dual plasmid NOT gate architecture in 4.3.1.1 above (Figure 4-13).

In *S. meliloti* L530 strain the dynamic range of the TetR NOT gate was the greatest at 4.6X, the LacI NOT gate was 4.3X and the PhlF NOT gate was the lowest 3.2X (Figure 4-13A). There was a significant difference in the uninduced state between all of the gates ($p<0.05$) but not the induced states of the gates ($p>0.05$).

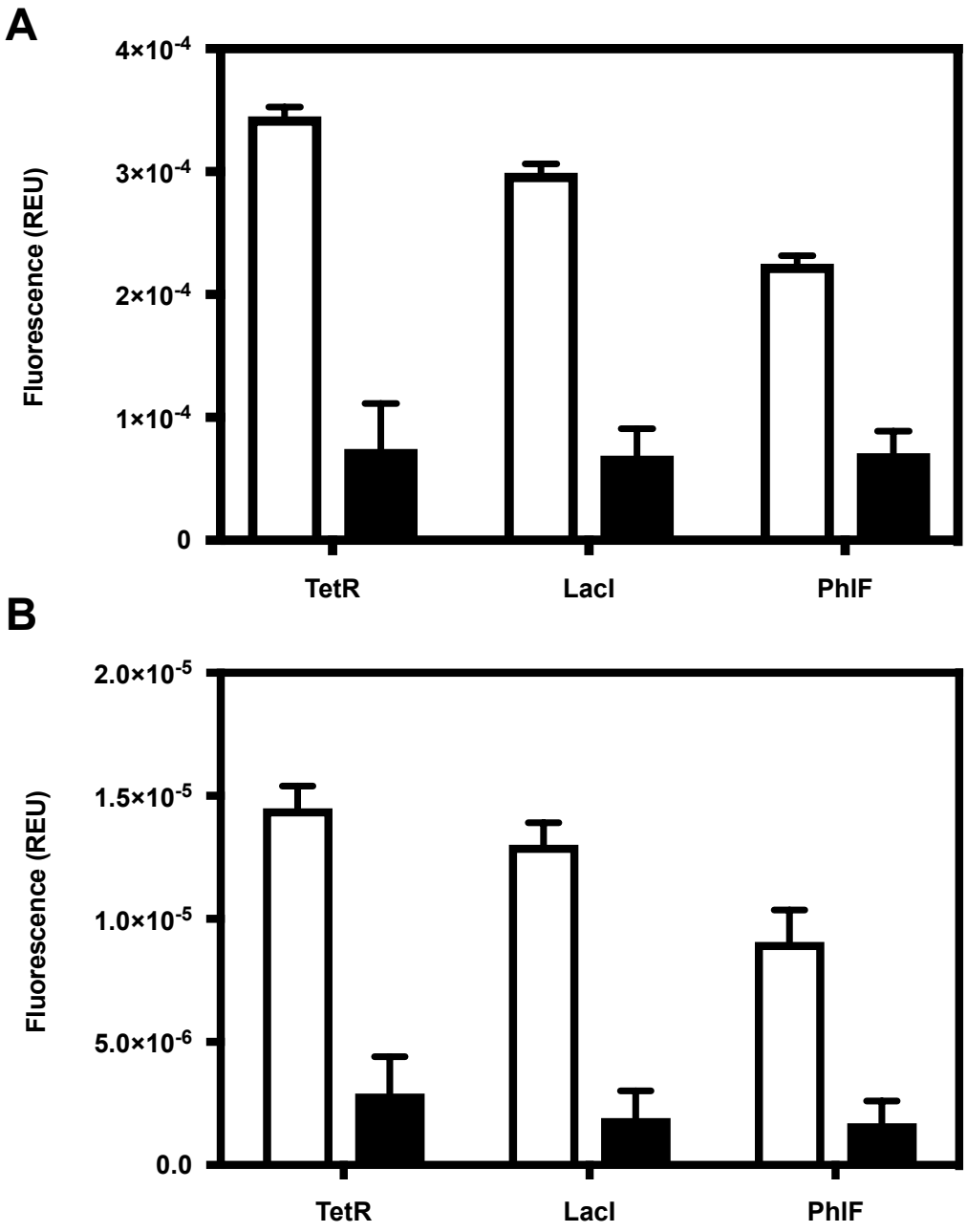


Figure 4-13 – Rhizopine inducible NOT gate dynamic range in both rhizobacterial test strains. **A** – *S. meliloti* L530. **B** – *R. leguminosarum* engineered with rhizopine sensitivity (Geddes et al., 2018 manuscript in preparation. White bars indicate uninduced state (input of 0) while black bars indicate the induced state (input of 1) with 10 μM rhizopine. Errors bars indicate the SEM. N = 3 biological and 3 technical replicates.

In the engineered *R. leguminosarum* 3841 rhizopine sensitive strain the dynamic range of the TetR NOT gate was 4.7X, the LacI NOT gate was the highest at 6.5X and the PhlF NOT gate was the lowest at 4.5X (Figure 4-13B). There was a significant difference in the uninduced state between all of the gates ($p<0.05$) but similarly to the *S. meilioti* L530 strain there was no significant difference in the induced states of the gates ($p>0.05$).

4.3.3.2 Fitness analysis

There was no significant difference ($p>0.05$) observed in either the growth rates or final culture density across all three NOT gates and induced states in *S. meliloti* L530 (data not shown). However, in the engineered *R. leguminosarum* 3841 rhizopine sensitive strain a significant difference was observed ($p<0.05$) in the generation time between the cluster of induced NOT gates and uninduced NOT gates (Figure 4-14). The cluster of induced NOT gates had a shared generation time of 4.33 hours whereas the cluster of uninduced NOT gates had a shorter generation time of 3.81 hours at a t= 30 hours. The final culture density observed at stationary phase for all of the test chassis was consistent at 1.9×10^9 cells ml^{-1}.

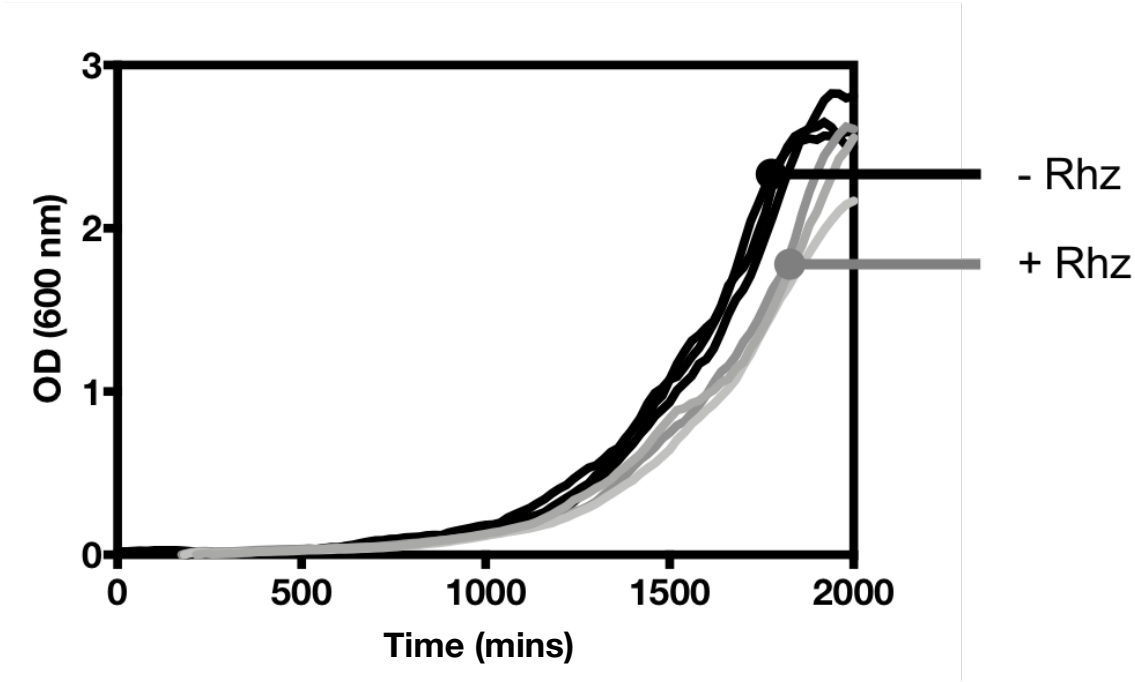


Figure 4-14 -Growth and fitness analysis for both the induced and uninduced states of the rhizopine inducible NOT gates in the engineered *R. leguminosarum* 3841 rhizopine sensitive strain. **– Rhz** indicates the growth of the strains with an input of 0, no rhizopine added. **+ Rhz** indicates the growth of the strains with an input of 1, 10 μM rhizopine added at t=0.

4.4 Discussion

The ability to finely control and manipulate gene expression and therefore relevant phenotypes is one of the most important aspects of synthetic biology and genetic engineering as a whole[202]. Implementing the structures and functions that underpin computation in rhizobacteria has the potential to influence many agriculturally-relevant issues such as eutrophication and yield optimization[203–205]. However until recently the field of genetic computation has been very limited in scope and often lacked any level of sophistication employing single transcriptional regulators and therefore only very small case-specific circuits could be constructed[191]. In this study we have characterised four transcriptional regulators that were made into NOT gates (TetR, LacI, PhlF and CIR) that all function in both the model rhizobacterial strains, *A. caulinodans* ORS571 and *A. brasilense* FP2 (4.3.1 above). By demonstrating the comparable function of the transcriptional repressors in both strains, portable circuits that can be implemented in other rhizobacteria chassis with predictable functions can now be constructed (4.3.1.1 above). The most promising of the four transcriptional repressors was that of the TetR repressor. With the largest dynamic range and no visible effect on cell viability, the TetR protein is our repressor of preference for NOT gate construction (4.3.1.1). Several previous studies have shown the TetR protein and its homologues to be the most promising of the known repressor families and as such the mining of more orthogonal regulators from the same

family is an area of intense interest in the field[187,198]. The dynamic range of the LacI-based NOT gate was comparable between *E. coli* and *A. brasilense* FP2 (Figure 4-4 and Figure 4-6). However the gate did not function as well in *A. caulinodans* ORS571 (Figure 4-5). We hypothesise that this is due to the presence of an endogenous *lacI* gene in the *A. caulinodans* ORS571 genome that has a 26.8% identity to that of the gene utilised in our circuit construction. Therefore it is possible that some cross-talk exists between the endogenous LacI gene and the Plac promoter of the NOT gate even when the synthetic LacI repressor is not present in the system.

Future work on tuning the dynamic range of the NOT gates using degradation tags on the transcriptional repressor could minimise the time taken for bifurcation of the gate and therefore create truly digital circuits[206]. Conversely, adding noise to the system in the form of feedback loops and nonspecific repressor operator sequences could convert the gates into a dynamic analogue output for applications such as circadian rhythm simulator[180].

Furthermore, we have measured the cross-reactivity of the transcriptional repressors and show there is minimal cross-talk allowing multi-regulator circuits to be constructed (4.3.2 above). A small amount of cross talk (>10%) was observed between the PhlF repressor and the Ptet promoter in the *A. brasilense* FP2 chassis (Figure 4-8C). This is possibly due to the fact that the TetR and PhlF repressors are both members of the TetR protein family and share a similar tertiary

structure[201]. Therefore some non-specific binding of the PhlF repressor to the Ptet operator sites exists, causing slight repression of the downstream sfGFP expression. Further analysis into the footprint of the PhlF DNA binding regions is required to establish the extent of the cross-reactivity. However the low level of cross talk is small enough not to cause major perturbations to genetic circuits constructed with both PhlF and Ptet promoters.

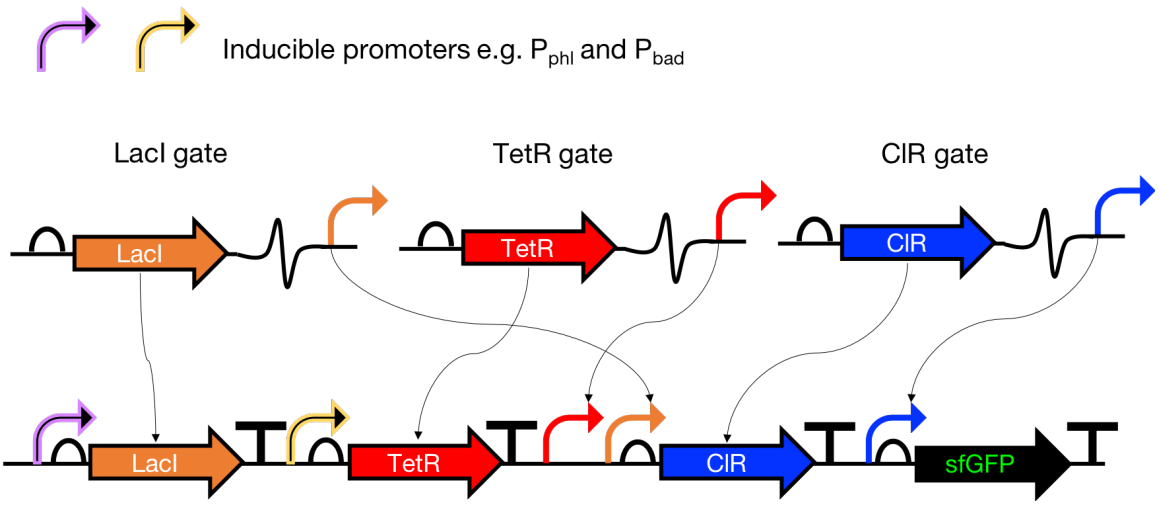


Figure 4-15 – Hypothetical AND gate genetic circuit architecture utilising the constituent parts of multiple NOT gates to yield sfGFP expression when both inputs (DAPG and Arabinose) are present. Note the genetic parts required for Pphl and Pbad activation (PhlF and AraC) are not shown in this diagram.

Previous studies have shown that NOT gates can be combined in sequence to yield more complex logical functions such as NOR and AND gates[187,193]. With the NOT gates characterised in this work a AND gate could hypothetically be constructed (Figure 4-15). Future work is required to characterise more inducible

promoter systems that are orthogonal to the repressors used in this study and then tune the system at both the transcriptional and translational levels for maximum response.

We were able to demonstrate the function of three out of the four NOT gates in response to a plant specific signal, rhizopine (4.3.3 above). The CIR repressor based NOT gate failed to clone by Golden Gate reaction (2.5.7 above) or traditional overlap PCR (2.5.10 above) with the rhizopine inducible PmocB promoter part. Due to time constraints in this project, the cloning reaction was not pursued and will require further work to successfully construct the rhizopine inducible CIR NOT gate circuit.

The two rhizobia strains utilised were *S. meliloti* L530 and an engineered rhizopine-sensitive *R. leguminosarum* 3841 strain. In both strains the three NOT gates based on the TetR, LacI and PhlF repressors functioned well with varying dynamic ranges. As expected the TetR based gate had the highest dynamic range in *S. meliloti* L530 (Figure 4-13A), however the LacI based gate had the highest range in the engineered *R. leguminosarum* 3841 strain (Figure 4-13B). This was unexpected and is possibly due to the LacI repression being tighter in than the TetR repressor In *R. leguminosarum* 3841. This may be due to a greater half-life of the protein, an optimum tertiary structure confirmation or even the secondary structure of the local region of the operator DNA sequence in the chassis being more favourable for binding for the LacI repressor[207]. The reduction in the

engineered *R. leguminosarum* 3841 rhizopine sensitive chassis generation time (Figure 4-14) when the NOT gates were induced with rhizopine is likely due to the increased burden on the cells energetic and translational machinery (ribosomes, ATP and cofactors)[208]. To decrease the metabolic burden on the cell, future work could tune the NOT gate circuitry to increase transcriptional initiation (with stronger promoters) and decrease translational speed (with weaker RBSs) yielding a more energy efficient expression genetic expression cassette[209].

I was unable to perform the rhizopine induction in either of the two model rhizobacterial strains *A. caulinodans* ORS571 or *A. brasilense* FP2. At the time of writing, the rhizopine responsive genetic circuits characterised in the lab by Geddes *et al.,* 2018 (manuscript in preparation) have been transferred to *A. caulinodans* ORS571 and show tentative functionality. No such success has yet been observed in *A. brasilense* FP2. I believe the rhizopine active transport accessory genes *intBC* are not functioning optimally in the *A. caulinodans* ORS571 strain and therefore refactoring of the synthetic circuits are required to enhance the rhizopine uptake and sensitivity[210]. It is also possible that the MocR rhizopine regulator does not exhibit optimal function due to a variation in the proteins tertiary structure or codon usage bias preventing mature protein formation[211].

The four transcriptional repressors were selected because of their specificity, stability *in vivo* and published ability to operate within synthetic circuits[197,198,201,212].

Other biochemical interactions could be used to expand the suite of orthogonal gates for future use in rhizobacteria. To be compatible with our computational architecture, the inputs and outputs of each of the logic gates must be promoters, therefore allowing the gates to be layered in a discreet hierarchy with transcriptional regulation as the core mechanism of the circuit. The transcriptional repression may be achieved with other classes of proteins that bind to DNA such as zinc finger proteins (ZFPs), transcription activator-like effector nucleases (TALENs) or even the sgRNA of a CRISPRi effector[71,213–215]. In future studies large computational circuits could comprise of mixtures of these transcriptional effectors whereby expanding the number of available gates beyond the limit of a single family of repressors[216].

The set of four orthogonal NOT gates presented here is sufficient to construct nontrivial genetic circuits with powerful logic relevant to agricultural biotechnology. For example the construction of a multi-input AND gate circuit may be constructed in a rhizobacteria. Only when a plant specific signal molecule such as rhizopine, as well as a robust concentration of the amino acid tryptophan will the chassis initiate auxin biosynthesis. If either of the inputs is 0 the circuit will not produce an output therefore preventing off-target and tryptophan depletion phenotypes from occurring. Other applications based on the four logic gates presented here include the construction of biological timers to control temporal gene expression, memory storage and simple control theory algorithms[184,195,216]. With the field of computation in bacteria progressing rapidly, there are now tools

available to rapidly automate the design of advanced computational circuitry[202]. However, the application of these computational circuits to alleviate real-world agricultural as well as wider biotechnological issues remains elusive. I believe that the characterisation of specific agriculturally relevant phenotypes in model rhizobacterial strains should be of the highest priority. By characterising natural plant growth promotion phenotypes from rhizobacteria and controlling them via computational circuits in novel hosts, it is possible to create crop-specific and highly regulated plant growth promotion in agricultural systems.

5 Engineering of Plant Growth Promotion Phenotypes

5.1 Abstract

In this chapter the design, construction and implementation of a suite of plant growth promotion mechanisms encoded in genetic circuits was successfully achieved. Constitutive Indole-3-acetic acid (auxin) biosynthesis was achieved and so was orthogonal chemical signal (DAPG) induction in both *Azorhizobium caulinodans* ORS571 and *Azospirillum brasilense* FP2. The PGP effect of the engineered rhizobacterial strain in a competitive soil environment showed similar significant plant root (1.5X) and root hair (2.5X) elongation of *Hordeum vulgare* (Barley) plants. The roots of those plants with significant increases in length showed a negative correlation in their root diameter. The active antifungal compound 2,4-diaceylphloroglucanol (DAPG) was successfully expressed *in vitro* from a constitutive biosynthesis circuit active in both *E. coli* ($2.8x10^{-7}$ μmol $cell^{-1}$ DAPG) and *A. caulinodans* ORS571 ($9.7x10^{-7}$ μmol $cell^{-1}$ DAPG). A phosphate solubilisation circuit comprising of citric acid biosynthesis and efflux modules was constructed. The circuit produced tentative phosphate solubilisation *in vitro* of insoluble di-calcium phosphate in *E. coli* and *A. brasilense* FP2 with a maximum phosphate solubilisation index of 2.33. The results of this study lay robust foundations for the future engineering of agriculturally-relevant PGPR.

5.2 Introduction

The soil environment is a complex and dense system with a multitude of bacteria, fungi, protozoa and algal cells. Of all of the microorganisms the bacteria are by far the most common constituting over 95% of the biomass in a typical soil system. Studies have shown that soil systems contain anywhere between 10^4 and 10^9 cells per gram of soil and of those species, only around 1% have been cultured sucesfully[217]. In environmentally stressed soils such as those found in intensive agricultural systems or in arid climates, the bacterial cell population may be as low as 10^4 cells per gram of soil[218]. Many factors influence the number and population breadth of the bacteria found in the soil including the temperature, moisture content, salinity and the number and types of plants present[219]. The distribution of bacteria in soil systems is not homogenous, the concentration of bacteria found around the root systems of plants is usually significantly higher than the bulk soil[220]. This is due to the increased concentration of organic acids, sugars, amino acids and other nutrients that exude from plant roots[221]. Previous studies have shown that off the total fixed carbon of a plant system, a third may be exuded through the root system in to the soil[222]. The region around the root systems of plants that contains many plant-associative bacteria is known as the 'rhizosphere' and bacteria that live in this niche are classified as 'rhizobacteria'. Those species that promote plant growth (PGP) are classified as 'plant growth promoting rhizobacteria' (PGPR)[92]. The interaction between rhizobacteria and

plants may be beneficial, detrimental or neutral from the point of view of the plant[223], however one species may have very different effects on the same plant in different environmental situations. For example, a species that provides a source of nitrogen to a plant may be beneficial when the plant is growing in a low nitrogen environment, but will be neutral to a plant growing in a high nitrogen system such as one with excess synthetic fertiliser application. Previous studies have also shown that one particular species can have entirely disparate effects on plant growth and development. For example an IAA producing mutant of the soil bacterium *Pseudomonas fluorescens* BSP53a stimulated the development of *Ribes nigrum* (blackcurrent) cutting roots whilst inhibiting root development in *Prunus avium* (wild cherry) cuttings [224].

PGPR include free-living species that may also form symbiotic relationships with plants, for example the *Rhizobia* and *Frankia* spp. and bacterial endophytes such as the *Azospirillum* and *Azorhizobium* spp. which both colonise portions of a plants tissue by various means[225–227]. The mechanisms by which the PGPR promote plant growth are wide and varied however the majority of species either facilitate resource acquisition, modulate plant growth hormone concentration or directly/indirectly offset the inhibitory effects of pathogenic fungal species[228]. Previous studies in the field of PGPR over the past 50 years focused on the physiological, molecular and biochemical perspectives of specific symbioses[8,227]. These studies therefore became the consensus approach to studying PGPR mechanisms and their effects in nature. However, conversely to the *Rhizobia* spp.,

the interactions between PGPR species and associative target plants are less well characterised. The most broadly studied mechanism for PGPR is that of the supply of nutrients that the soil environment is unable to adequately provide a target plant. Nutrients such as fixed nitrogen, iron, phosphorus and other trace elements. Many agricultural soils are suboptimal in one or more of these key nutrients and also consequently have degraded rhizospheres with limited bacterial populations.

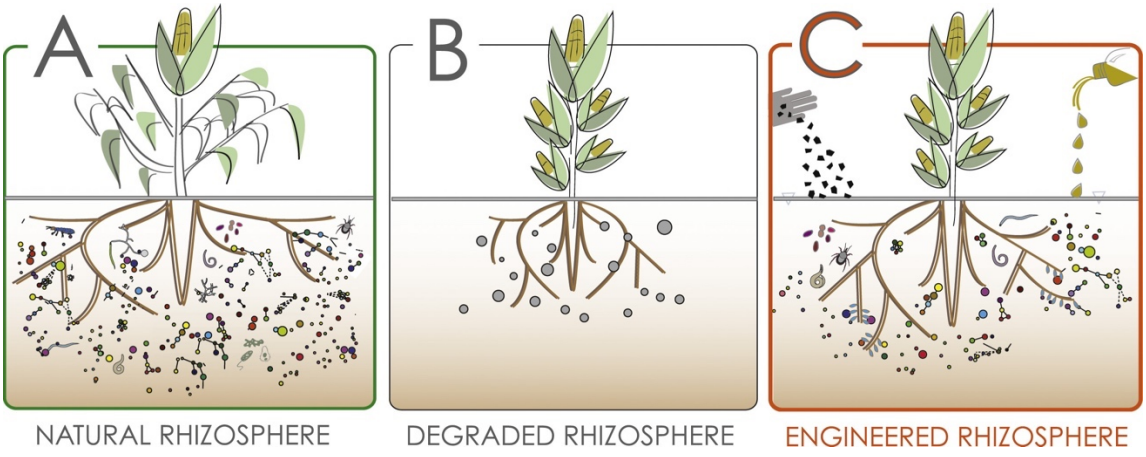


Figure 5-1 – Representation of various rhizosphere classes and characteristics. **A** – A natural rhizosphere with a robust and dense bacterial population. **B** – A degraded rhizosphere often found in intense agricultural systems with excess fertiliser application. **C** – An engineered rhizosphere, the addition of inoculants and PGPR specific adjuncts may rejuvenate degraded rhizospheres. Adapted from Wallenstein, 2017[229].

To date farmers have become dependent on synthetic and expensive forms of both nitrogenous and phosphorus-based fertilisers which counter-intuitively degrade the natural rhizospheres of plants, causing a decrease in the bacterial

populations (Figure 5-1A). Synthetic fungicides are also applied in large quantities, often decimating natural mycorrhizal populations and degrading the rhizosphere further. In addition to the fertilisers and pesticides being costly, the production of the compounds and their over-application cause significant human and environmental hazards such as eutrophication and massive greenhouse gas emissions[223]. It is therefore of critical importance that natural alternatives to fertilisers and fungicides be developed and applied in agricultural systems. A systems engineering approach to rhizosphere development has been used successfully to reconstitute the natural PGPR populations in degraded systems[230]. The advent of synthetic biology allows us to harness the PGPR phenotypes from many disparate organisms and combine them in strategic host chassis to target and tune PGP traits in target crop plants. In this chapter the selection, construction and characterisation of a suite of PGP phenotypes in confirmed plant-associative rhizobacteria is presented with a view to enhance on PGP systems found in nature.

5.2.1 Nitrogen Fixation

In addition to the *Rhizobia* spp. there are many other diazotrophic (nitrogen fixing) bacteria that inhabit the soil and specifically the rhizosphere. *Azospirillum* spp. for example are able to fix nitrogen as their sole nitrogen source and have been shown in previous studies to release fixed nitrogen on to associated plants[36]. The general consensus in the field is that free-living diazotrophs only supply minute amounts of their total fixed nitrogen to plants and do not supply enough to satisfy the nitrogen requirement of a healthy plant[35]. In section 6.2 of this thesis the discussion of the engineering of ammonia assimilation pathways in rhizobacteria will be presented. The field of metabolic engineering has allowed tools to tune and modify rhizobacterial cells into releasing more fixed nitrogen than they would naturally. Nitrogenase (*nif*) genes that are essential for the fixation of nitrogen are complex in number and mechanism. They include the structural genes *nifHDK*, genes required for the FeMo co-factor biosynthesis, electron transport chain and a suite of regulatory genes including the global activator *nifA* to control the process[231] (Figure 5-2).

The *nif* genes are typically found in a large cluster of operons around 22 kB in length and composing of 20 different proteins[232]. Because of the complexity and multi-layered feedback loops of the system, genetic engineering approaches to improve nitrogen fixation have been only few in number. Many different research

groups and projects are currently attempting to re-engineer the nitrogen fixation pathway either into plants themselves[21,233], into other eukaryotic systems[234] or into novel plant-associated rhizobacterial strains with thorough refactoring of the endogenous genetic circuitry[80,235]. Since the process of nitrogen fixation is so energy-intensive, requiring a large amount of energy in the form of ATP, many studies have attempted to direct the flux of carbon metabolism in a strain towards that of oxidative phosphorylation (yielding ATP) rather than glycogen synthesis. In one such study a strain of *Rhizobium tropici* was engineered with a glycogen synthase deletion. In co-culture with the host *Phaseolus vulgaris* (common bean) plant, a significant increase in plant dry weight was observed[236]. Although a success in controlled growth environments, the engineered *R. tropici* strain did not proliferate in a competitive rhizosphere environment, indicating the need for both temporal and spatial control of phenotypes such as a glycogen synthesis cut-off.

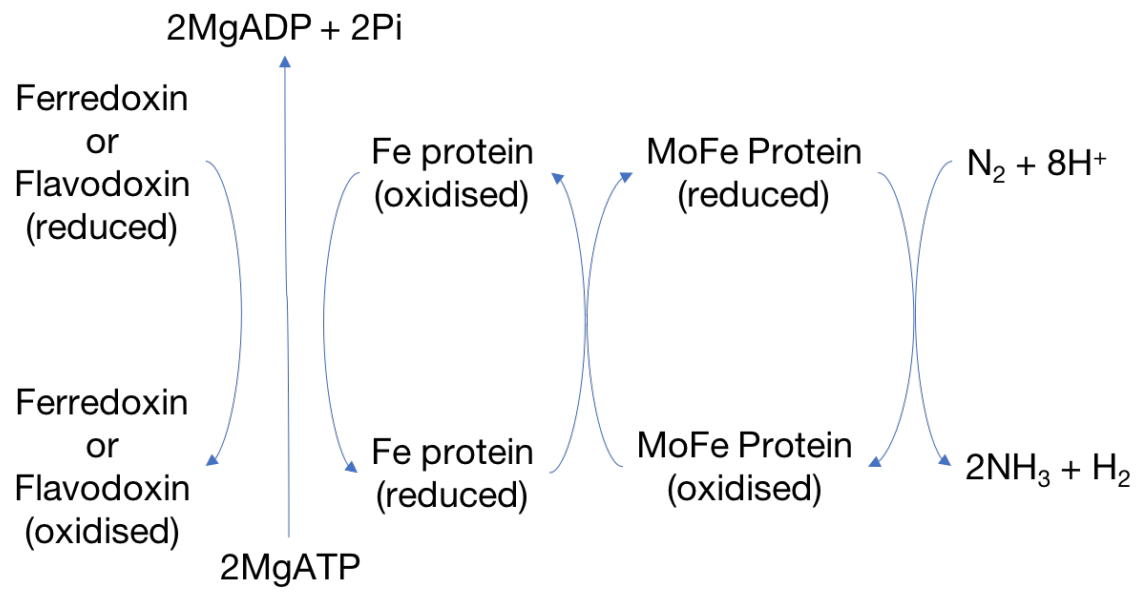


Figure 5-2 – Mechanism for biological nitrogen fixation in diazotrophic PGPR. ADP – Adenosine diphosphate. ATP – Adenosine triphosphate. Pi – inorganic phosphate.

5.2.2 Phosphate solubilisation

Phosphorus concentrations in agricultural soil environments are usually relatively high (between 350 and 1,150 mg $kg^{-1}$of soil)[237]. However, most of this phosphorus is insoluble and is therefore not able to cross the cell wall and membrane of plant root cells, providing no benefit to biosynthesis. Insoluble phosphate is often present in the form of an inorganic mineral such as dicalcium phosphate ($Ca_3(PO_4)_2$) or an organic compound such as inositol phosphate[238]. When additional inorganic phosphate is added to agricultural systems as fertiliser, it is

rapidly immobilised into an insoluble form and therefore sequestered away from the plant biosynthesis process. Phosphorus is essential to plant biosynthesis and the inability of plants to obtain sufficient amounts of the element causes reductions in yields and nutrient content[238]. Therefore the solubilisation of phosphorus in the rhizosphere of crop plants especially is an important trait from PGPR species as well as a target for future synthetic biology engineering[239,240]. The main mechanism by which inorganic phosphate is solubilized is by the action of organic acids such as gluconic and citric acid, both of which are synthesized and secreted by several soil bacteria including *Pseudonomas* spp[241]. Conversely the mineralisation of phosphate occurs through a wide variety of phosphatases in many bacterial strains, including those which perform beneficial solubilisation[242] (Figure 5-3).

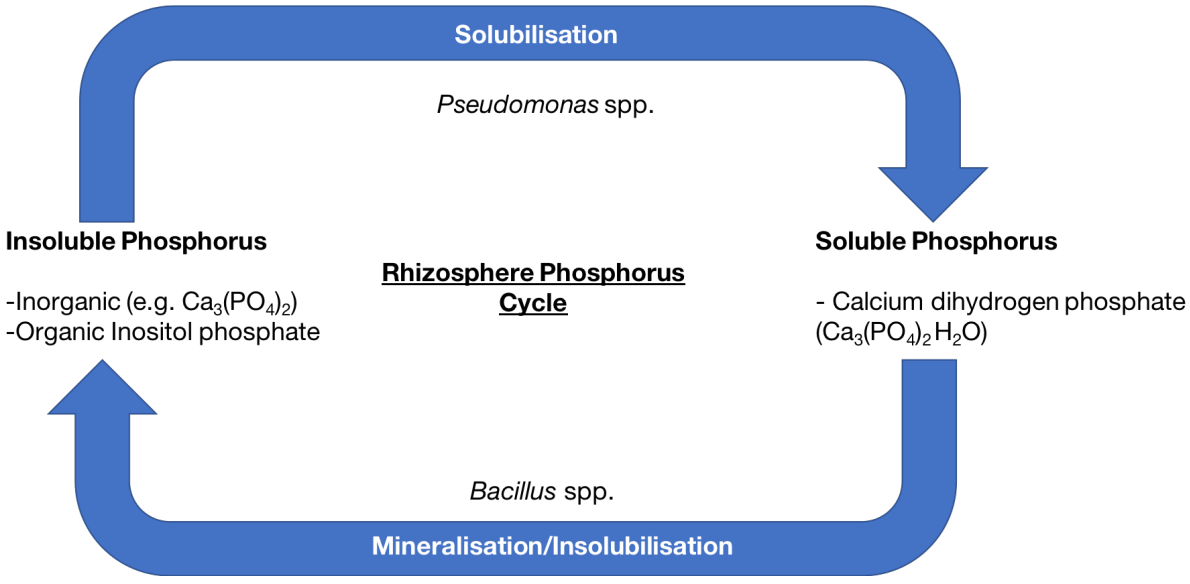


Figure 5-3. Phosphate mineralisation and solubilisation cycle in the rhizosphere.

Several studies have attempted to enhance the phosphate solubilisation capabilities of rhizobacteria and have had only limited success as the strains viability often decreases significantly with enhanced organic acid biosynthesis[243]. Often only when phosphate solubilising rhizobacteria are co-inoculated with other PGPR strains do any significant measurable PGP effects occur[244]. Synthetic biology approaches to reengineering phosphate solubilisation may maximise organic acid production with a lower effect on the strains fitness.

5.2.3 Indole-3-Acetic Acid (Auxin)

Throughout the study of the rhizosphere and plant development, several natural occurring auxins have been discovered and characterised, indole-3-acetic acid (IAA/Auxin) is widely regarded as the most well understood and powerful of the natural phytohormones[245,246]. Several rhizobacterial species have been shown to constitutively produce small amounts of IAA when in the rhizosphere[93]. The strain *Pseudomonas putida* GR12-2 was tested for its IAA biosynthesis and effect on *Brassica napus* (Oilseed rape) seedlings. The study showed that deletion of the cells IAA biosynthesis pathway caused a 50% reduction in the root length of the cocultured *B. napus* seedlings when compared to a coculture with the wild-type *P. putida* GR12-2 strain[247].

IAA has a wide variety of plant development effects from cell division, differentiation and elongation to seed and tuber genesis, lateral root proliferation

and photo/gravitropism[248,249]. The concentration of IAA in a system has a dramatic variation of effects depending on the plant species present (5.2 above). Indeed IAA has significantly different effects depending on the plant tissue. For example plant roots require five orders of magnitude less IAA for supporting plant growth than the shoot tissue. The endogenous concentration of IAA in a plant is often a balance of the biosynthesis systems of the plant itself as well as the accumulation of IAA from the rhizosphere. With this in mind the level of IAA synthesised by a plant, when coupled with that of the IAA from the rhizosphere must be carefully balanced to achieve an optimal concentration of IAA for plant growth promotion[250]. Novel synthetic biology circuitry can be utilised to effectively regulate IAA biosynthesis in engineered PGPR species and target crop systems thereby allowing optimal IAA concentrations for the plant species and environmental conditions present.

5.2.4 Antifungal Protection

Intense farming techniques in the past two decades have necessitated the application of synthetic agrochemicals ranging from fertilisers (see 5.2 above) to pesticides and antifungal compounds. These have many negative side effects from rhizosphere mycorrhizal degradation to multi-antifungal resistant pathogens[251,252]. Due to the large yield losses due to fungal pathogen diseases, the use of PGPR with antifungal properties are of significant interest in agriculture as an alternative to expensive synthetic compounds[253]. Several species of *Bacillus*

and *Pseudomonas* spp. are commercially available as biocontrol cultures that excrete hydrolytic enzymes that actively degrade the cell walls of fungi[254]. The *Pseudomonas* spp. especially excrete a wide variety of antimicrobial and antifungal compounds such as 2,4-diacetylphloroglucinol (2,4-DAPG or DAPG), pyoluteorin and even hydrogen cyanide[255,256].

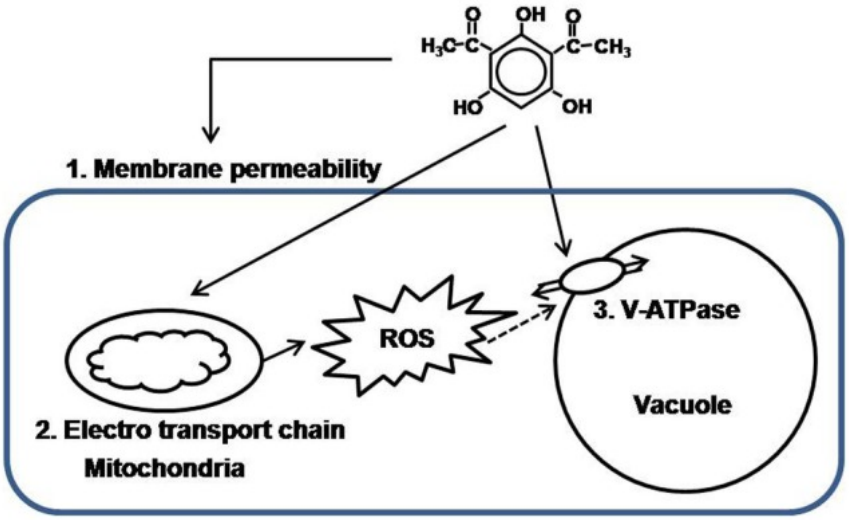


Figure 5-4 - A putative model for the mechanism of action of DAPG. DAPG likely causes membrane permeability to increase, disruption of the mitochondrial electron transport chain and inhibition of V-ATPase activity. Adapted from Kwak *et al.,* 2013[257].

DAPG has been shown in several studies to be a potent antifungal agent inhibiting the growth of several global pathogenic fungi such as *Gaeumannomy cesgraminis*, *Penicillium digitatum* and *Penicillium italicum*[258]. The direct mechanism by which DAPG inhibits fungal growth is not fully understood however a putative mode of action has been proposed in the literature (Figure 5-4). In several past studies the yeast *Saccharomyces cerevisiae* was utilised as a model

to analyse the effect of DAPG on fungal cell function. One study indicated that DAPG interfered with the mitochondrial function[259] of the cell whereas another indicated that DAPG perturbs the cell membrane increasing permeability leading to a reactive oxygen burst[257,260]. The study of the antifungal mechanisms of DAPG promise to be an area of significant research in the field of plant growth promotion in the future, especially as an effector that can be controlled using novel synthetic biology circuitry.

5.3 Results

5.3.1 Engineering IAA Biosynthesis

Although IAA is a plant hormone and is biosynthesised in small amounts by plants, many bacteria have the genetic capacity to perform IAA biosynthesis (see 5.2.3 above). There are several different biochemical pathways that have been elucidated[261] yet all use the substrate tryptophan as their precursor compound (Figure 5-5).

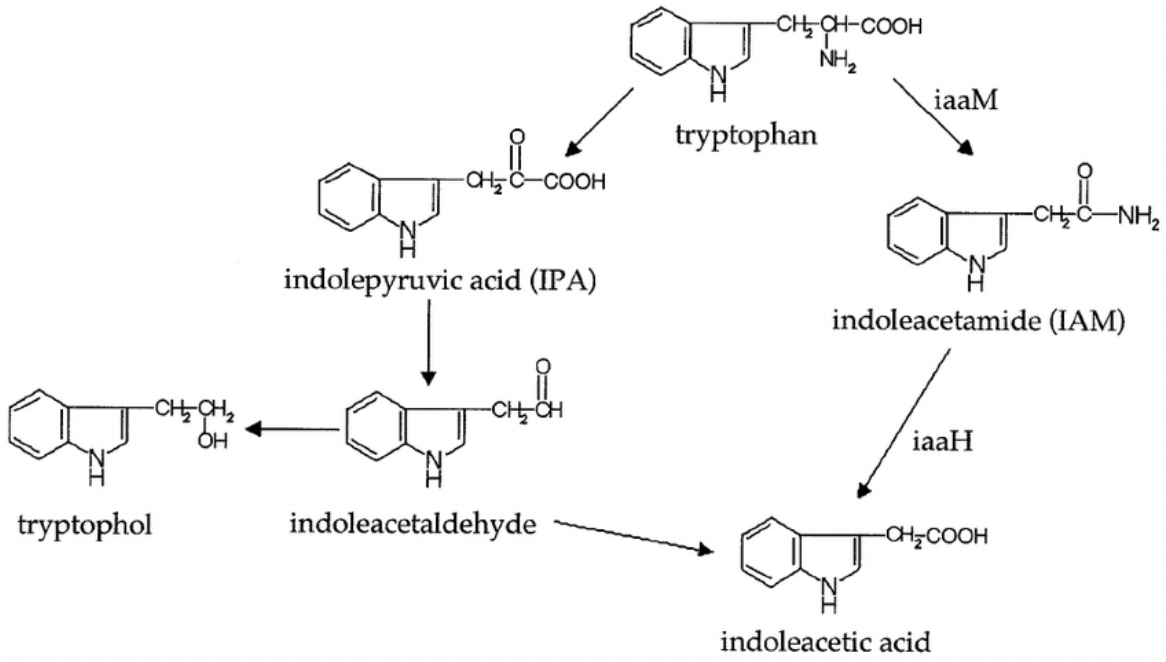


Figure 5-5 - IAA biosynthesis pathway from *Pseudomonas savastanoi*. Adapted from Cohen *et al.,* 2002[262].

We decided in this study to utilise the two gene tryptophan monooxygenase (IaaM) and indole-3-acetamide hydrolase (IaaH) system as it is a simple two gene pathway and has previously been demonstrated to function in *E. coli*[263] (Figure 5-5). As a basis for the work we utilised the IaaM and IaaH gene sequences from *Pseudomonas savastanoi*, a species known to produce and excrete large concentrations of IAA into host plants upon pathogenic interaction[264]. For the ease of modular synthetic circuit construction and cloning, the Type IIS Golden Gate cloning pipeline was utilised for the construction of all IAA biosynthesis circuits (2.5.7 above).

The *iaaM* and *iaaH* genes were made into pL0M-SC DNA parts by PCR amplification from the iGEM plasmid K515100 using primers oxp1339 and oxp1340 for *iaaM,* and oxp1338 and oxp1341 for *iaaH* (2.5.2 above). These parts were cloned by level 0 Golden Gate reaction into pOGG072 to create plasmids pL0M-SC-IaaM (pOGG192) and pL0M-SC-IaaH (pOGG193) and were sequence verified before use in downstream cloning reactions (2.5.8 above).

5.3.1.1 Constitutive auxin biosynthesis

To test the function of the IAA biosynthesis pathway, the constitutive expression of *iaaM* and *iaaH* in the synthetic biology mainstay *E. coli* and the two rhizobacterial strains *A. caulinodans* ORS571 and *A. brasilense* FP2 was performed. To construct the two-gene expression cassette, DNA parts from the

RhizoKit toolbox were implemented (see chapter 3 above). Initially two expression cassettes were constructed for the *iaaM* and *iaaH* genes in level one Golden Gate holding vectors before being combined into a single stable broad host range plasmid by a level 2 reaction. qPCR studies have shown that the stoichiometry of the IaaM and IaaH mRNA transcripts have a ratio of approximately 10:1 IaaM to IaaH[264,265]. Therefore we tuned our expression circuit to be of the same ratio to preserve maximum pathway flux *in vivo*.

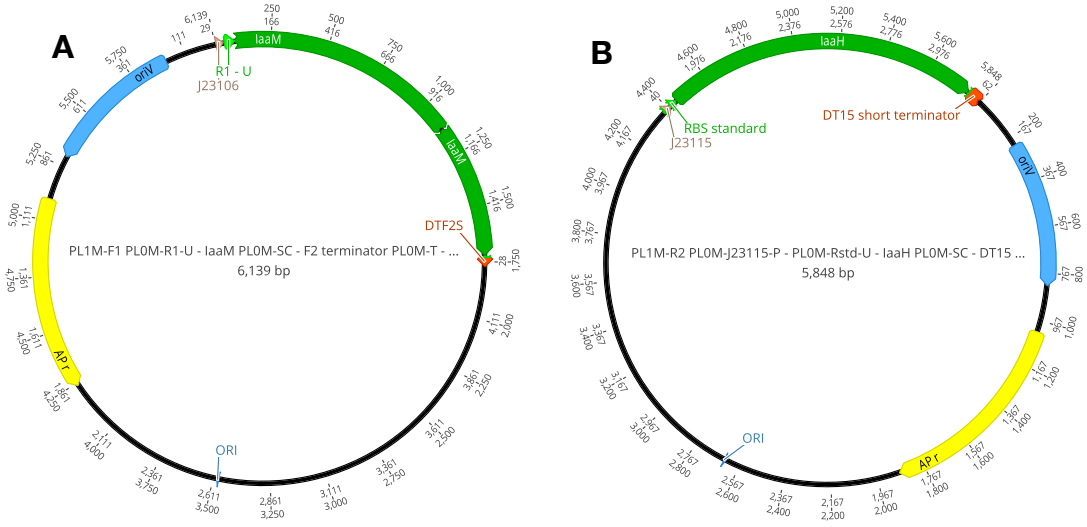


Figure 5-6 – Plasmid maps of the two level one IAA biosynthesis pathway plasmids. **A** – IaaM expression cassette in pL1V-F1 (pOGG234). **B** - IaaH expression cassette in pL1V-R2 (pOGG235). Note the oriV *E. coli* origin of replication and ampicillin selection genes.

The *iaaM* gene was cloned into the pL1V-F1 (pOGG021) vector with a medium strength pL0M-P-J23106 promoter (pOGG121), a pL0M-U-RBS1 (pOGG144) and

terminated by the pL0M-T-F2S terminator (pOGG160, Figure 5-6A). The *iaaH* gene was cloned into the pL1V-R2 (pOGG059) vector with a low strength pL0M-P-J23115 promoter (pOGG122), a pL0M-U-RBSstd (pOGG143) and terminated by the pL0M-T-DT15S terminator (pOGG156, Figure 5-6B). Both the level one plasmids were then purified (2.5.1 above) before being cloned in a level two Golden Gate reaction (2.5.7 above) into both the pOGG027 and pOGG096 pL2V vectors producing pOPS0666 and pOPS0667 respectively. We chose to clone into both level two vectors to analyse which of plasmids yielded enhanced biosynthesis profiles.

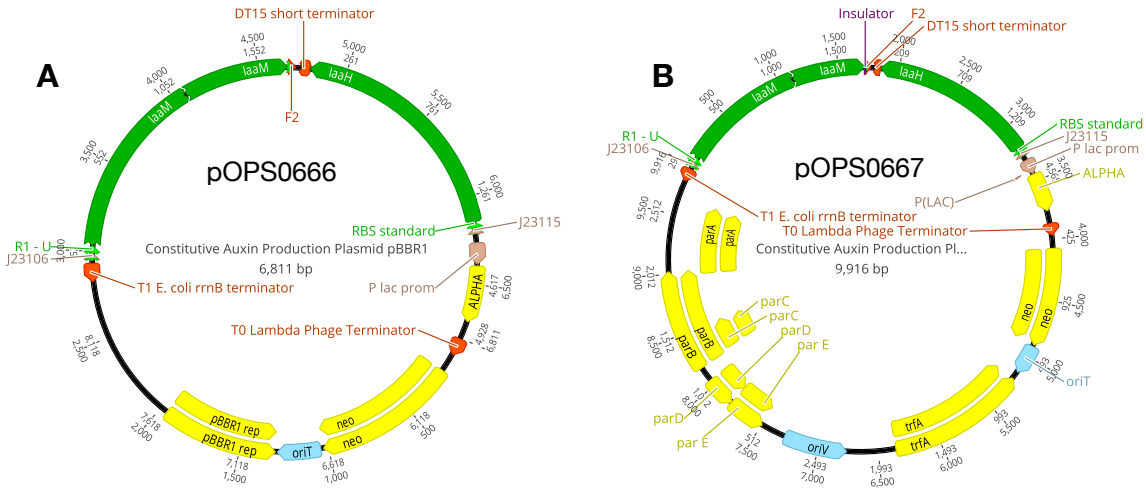


Figure 5-7 - Plasmid maps of the two constitutive IAA biosynthesis pathway circuits. **A** – Constitutive IAA circuit based in the pOGG027 broad host range plasmid (pOPS0666). Note the pBBR1 replication of origin and kanamycin/neomycin selection. **B** - Constitutive IAA circuit based in the pOGG096 broad host range plasmid (pOPS0667). Note the RK2 replication of origin, kanamycin/neomycin selection and *par* operon for enhanced maintenance of the plasmid in culture.

The expression cassettes for each of the IAA genes were orientated so that their transcriptional directions converged, preventing read through despite the terminator validation (3.3.3 above). Both pOPS0666 and pOPS0667 were transformed into *E. coli* by heat shock (2.6.1 above). The plasmids were transformed into the two rhizobacterial strains *A. caulinodans* ORS571 and *A. brasilense* FP2 by electroporation (2.6.2 above).

To analyse the production of IAA by the two rhizobacteria strains as well as the *E. coli* control, a high throughput colorimetric assay based on the Salkowski assay[266] was developed. To compare the concentrations of IAA produced by the two engineered rhizobacterial strains, the wild type *A. caulinodans* ORS571 and *A. brasilense* FP2 were used to determine basal endogenous IAA production and therefore normalise the data. Empty vector controls were also added to verify circuit-dependent IAA production. Where possible all cultures and samples were kept in the dark due to the photo-instability of IAA[267]. All the test strains were grown overnight in either LB (*E. coli*) or TY (rhizobacteria) media to a culture OD595 of 0.6 (2.7.1 above). The *E. coli* strains were then sub-cultured into 10 mL semisolid (0.175% (w/v) agarose) LB media plus relevant antibiotics in gas-sealed universal tubes (2.4.2 above and 2.4.4 above). The two rhizobacterial strains were sub-cultured into nitrogen-free 10 mL UMS + 20mM Succinate (*A. brasilense* FP2) and UMS + 20 mM Succinate + 30 μM nicotinic acid (*A. caulinodans* ORS571) at an initial cell density of 10^5 cells (2.4.4 above). The cultures were grown for 48 hours in the dark until a robust pellicle was visible for both rhizobacterial strains

(Figure 2-3), no pellicle was observed in the *E.coli* strain due to its ability to colonise semisolid media irrespective of oxygen concentration. After 48 hours the samples were vortexed to homogenise the cultures, and then centrifuged at 6000g for 15 minutes at 4˚C. The agarose layer was removed from above the cell pellet by cut pipette tips. The cell pellet was then weighed both wet and after drying at 80˚C to determine cell number and protein content for IAA production normalisation. The supernatant was filtered using vacuum filtration with a 0.22 µm cellulose filter (Advantec MFS, Dublin, Ireland) to remove any cellular debris.

To make the Salkowski reagent, 0.811 g of anhydrous $FeCl_3$ was dissolved in 10 ml of dH_2O to obtain a 0.5M solution. The final reagent is made by dissolving 1 mL of the 0.5 M $FeCl_3$ solution in 49 mL of a 35% (w/v) $HClO_4$. The Salkowski reagent was stored at room temperature in the absence of sunlight. For high throughput IAA quantification a microwell absorbance assay was used. Supernatant sample (250 µl) was transferred to a single 24-well microwell plate, to this 500 µL of the Salkowski reagent was added. The plate was then shaken for 15 minutes at 500 rpm orbital shaking before the OD530nm was measured and normalised to a blank control. A series dilution from 0 to 100 µmol of IAA (Sigma-Aldrich) was used to create a standard curve from which the amount of IAA produced from the samples was calculated (Supplementary Figure 1) The IAA amounts were then normalised to the cell number from the analysis of the pellet performed subsequent to the culture centrifugation (see above).

The results from the constitutive IAA biosynthesis circuits in *E. coli*, *A. caulinodans* ORS571 and *A. brasilense* FP2 indicated successful IAA biosynthesis in all three strains. The wild-type strain produced a negligible amount of IAA, as did both the empty pOGG027 and pOGG096 vector carrying strains. The amount of IAA produced from the pOGG096-based constitutive expression circuit (pOPS0667) was significantly higher ($p<0.05$) at $2.8x10^{-13}$ mol $cell^{-1}$ than the pOGG027 based circuit (pOPS0666), which produced $1.1x10^{-13}$ mol $cell^{-1}$ (Figure 5-8A).

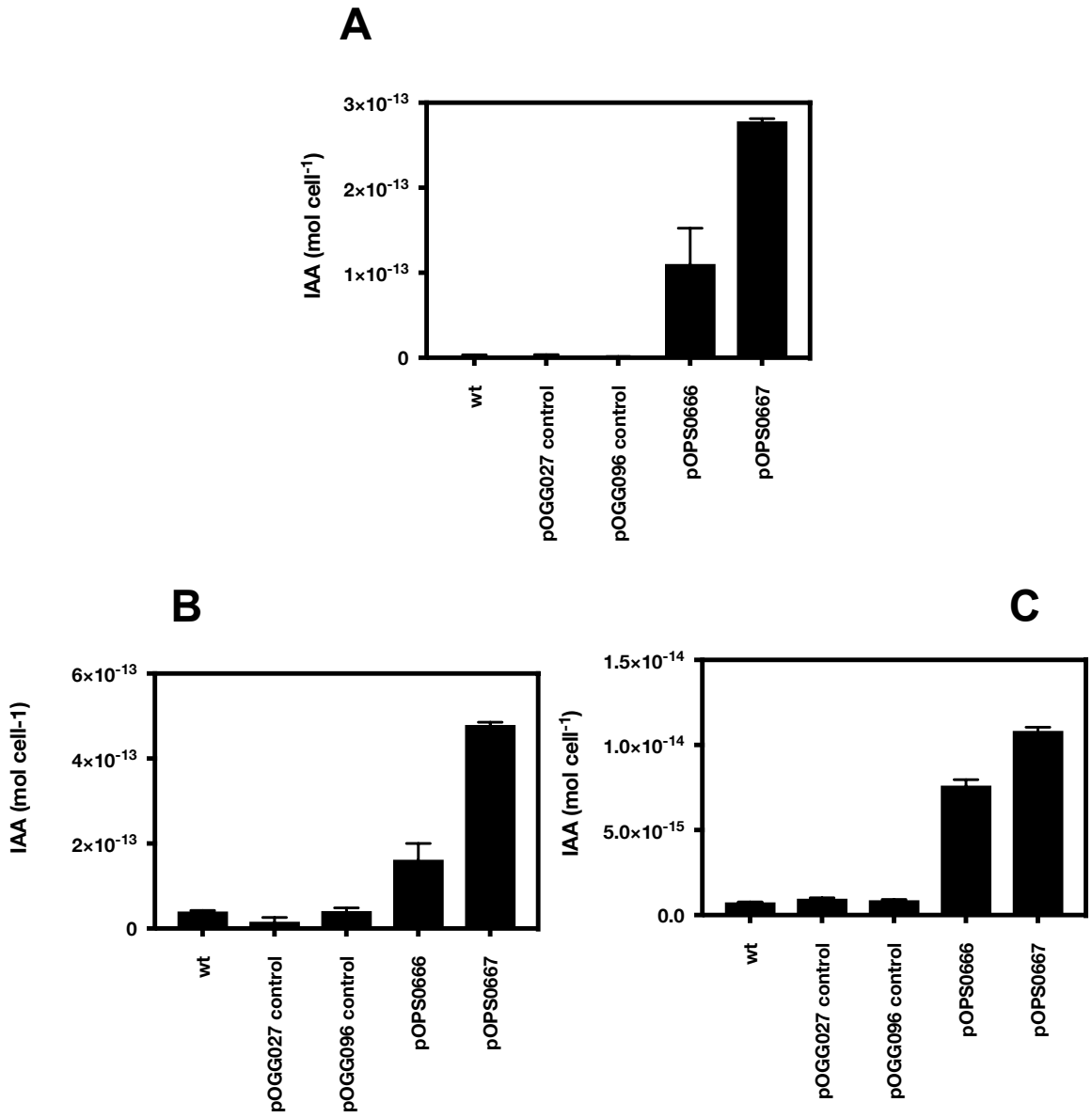


Figure 5-8 - IAA biosynthesis analysis from the constitutive IaaM and IaaH gene circuits in *E.* coli and two rhizobacterial strains after 48 hours diazotrophic growth. ***A*** – *E. coli.* **B –** *A. caulinodans* ORS571. **C** – *A. brasilense* FP2. wt – Wild-type. Error bars indicate the standard error of the mean. N = 3 biological and 3 technical replicates.

There was a high degree of variation from the biological replicates of the pOGG027 based IAA expression circuit (pOPS0666) with 10X the SEM compared to that of the pOGG096 based circuit (pOPS0667).

Out of the three strains tested IAA biosynthesis was highest in *A. caulinodans* ORS571. The wild-type control produced a nontrivial amount of IAA at $4.1x10^{-14}$ mol $cell^{-1}$ (Figure 5-8B). Both of the empty vector controls showed IAA amounts that were not significantly different ($p>0.05$) to that of the wild type control. The pOGG027-based constitutive IAA expression circuit (pOPS0666) produced $1.6x10^{-13}$ mol $cell^{-1}$ of IAA and was significantly lower ($p<0.05$) than the pOGG096-based circuit (pOPS0667), which produced $4.8x10^{-13}$ mol $cell^{-1}$ (Figure 5-8B). Similarly to the *E. coli* control the variation between biological replicates of the pOPS0666 circuit was much higher than that of the pOPS0667 circuit.

The *A. brasilense* FP2 strain produced the least IAA out of the three species tested, over 10X less than *A. caulinodans* ORS571 and *E. coli* (Figure 5-8C). The wild type control produced $7.5x10^{-16}$ mol $cell^{-1}$ of IAA which was not significantly different to that of the pOGG096 empty vector control ($p>0.05$) but was significantly lower than the pOGG027 empty vector control which produced 9.7x10 mol $cell^{-1}$ (Figure 5-8C). In correlation with the results observed in the other two chassis, the pOPS0666 circuit produced less than the pOPS0667 circuit. The pOPS0666 circuit produced $7.6x10^{-15}$ mol $cell^{-1}$ of IAA compared to $1.1x10^{-14}$ mol $cell^{-1}$ produced from the pOPS0667 circuit.

5.3.1.2 Inducible IAA Biosynthesis

For an IAA biosynthesis circuit to be tuned actively *in vitro*, an inducible version of the circuit is required. Several previous studies have shown that inducible biosynthesis circuits in engineered chassis both reduce the energetic burden of the process, as well as yielding higher amounts of the product post-induction[268]. For this reason a DAPG inducible IAA biosynthesis circuit was constructed for expression in both the *A. caulinodans* ORS571 and *A. brasilense* FP2 strains. The circuit was constructed using characterised parts from the RhizoKit toolbox (chapter 3 above) and cloned using the Golden Gate pipeline for rapid and modular assembly (2.5.7 above). Two expression cassettes were constructed for the *iaaM* and *iaaH* genes in level one Golden Gate holding vectors before being combined into a single stable broad host range plasmid by a level 2 reaction.

The *iaaM* gene (pOGG192) was cloned into the pL1V-F1 (pOGG021) vector with a pL0M-rPU-Pphl DAPG inducible promoter (pOGG137) and terminated by a pL0M-T-F2S terminator (pOGG160, Figure 5-9A). The *iaaH* gene (pOGG193) was cloned into the pL1V-R2 (pOGG059) vector with the pL0M-PU-Pphl promoter (pOGG136) and terminated by the pL0M-T-DT15S terminator (pOGG156, Figure 5-9B).

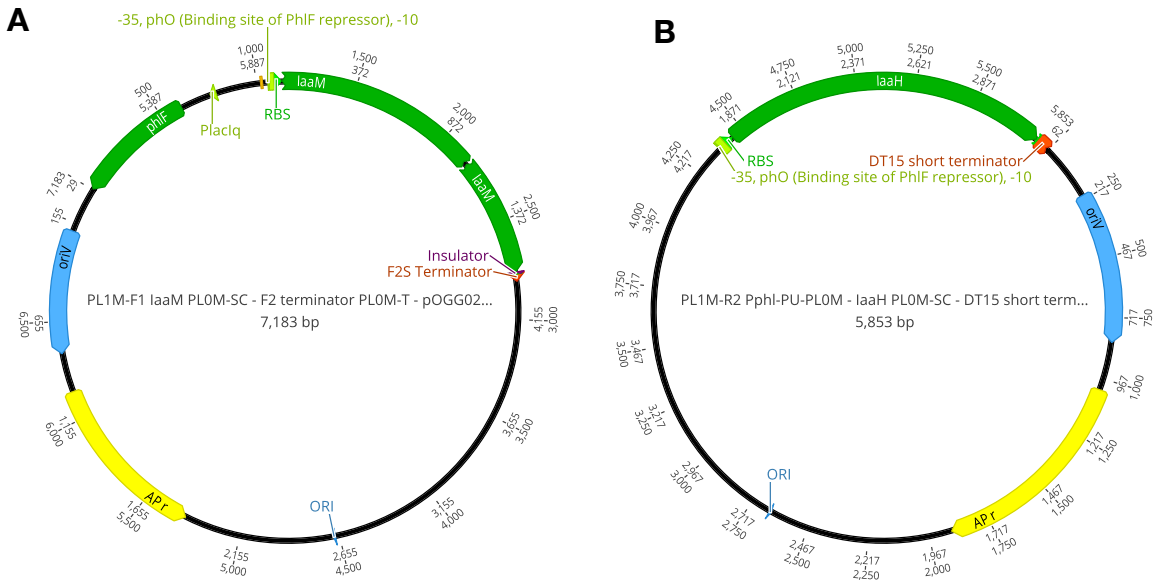


Figure 5-9 - Plasmid maps of the two level one DAPG inducible IAA biosynthesis circuit plasmids. **A** – IaaM expression cassette in pL1V-F1 (pOPS236). **B** - IaaH expression cassette in pL1V-R2 (pOPS237). Note the oriV *E. coli* origin of replication and ampicillin selection genes.

Both the level one plasmids were then purified (2.5.1 above) before being cloned in a level two Golden Gate reaction (2.5.7 above) into the pOGG027 pL2V vector along with the pL1M-ELB-2 module (pOGG56) for downstream integration of other accessory genes (Figure 5-10). We attempted to clone the DAPG inducible circuit into pOGG096 for enhanced maintenance in culture but were unsuccessful. The expression cassettes for each of the IAA genes were orientated so that their transcriptional directions converged, preventing read through despite the terminator validation (3.3.3 above).

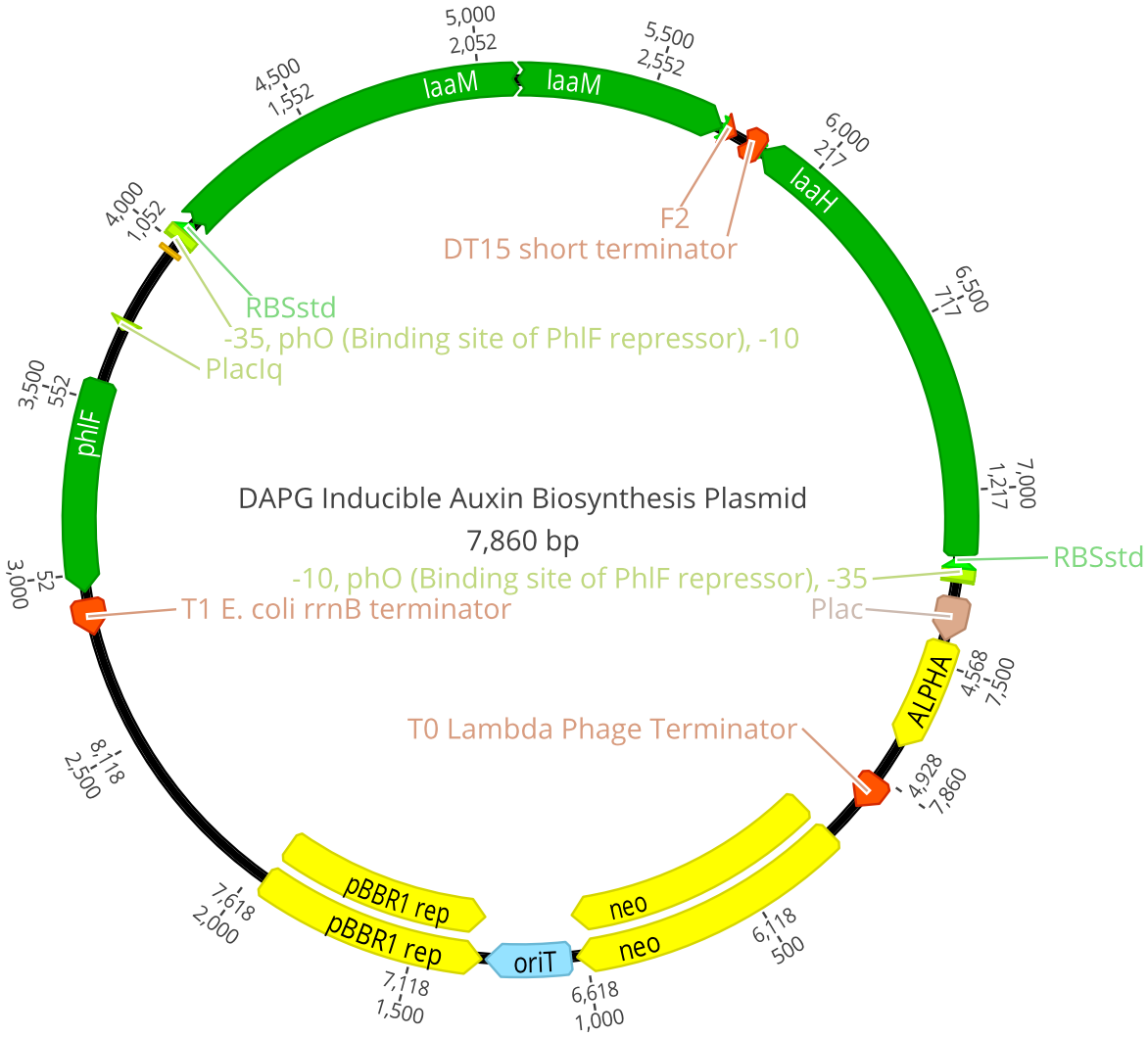


Figure 5-10 – Plasmid map of the DAPG inducible IAA expression plasmid in the pOGG027 broad host range vector (pOPS1166). Note the pBBR1 based origin of replication and kanamycin/neomycin selection gene. The ALPHA ELB3 fragment allows for downstream cassette integration into this vector by BsaI Golden Gate reaction (2.5.7 above).

The DAPG Inducible IAA circuit (pOPS1166) was transformed into *E. coli* by heat shock (2.6.1 above). The plasmids were transformed into the two rhizobacterial strains *A. caulinodans* ORS571 and *A. brasilense* FP2 by electroporation (2.6.2 above). To analyse the production of IAA by the two rhizobacteria as well as the *E. coli* control, a high throughput colorimetric assay was used identical to that used to analyse the constitutive IAA biosynthesis circuits (see 5.3.1.1 above). The wild-type chassis and empty pOGG027 vector controls were analysed in parallel to the circuit-carrying chassis to perform post-assay normalisation to basal IAA production. To induce the DAPG biosynthesis cassette in each test strain, 1mM DAPG was added to the 10 mL of semisolid media prior to inoculation. DAPG induction mid growth profile could not be attempted due to disruption of the rhizobacterial pellicle structure upon DAPG addition and centrifugation. Previous studies have shown that the pellicle structure formed when diazotrophic rhizobacteria are grown in semisolid media is essential for nitrogen fixation to occur and therefore cell viability[269].

The results from all three chassis indicated successful DAPG induction of the IAA biosynthesis circuit. The results from the *E. coli* control chassis indicated the amount of IAA was produced from the wild type strain was below detection levels for the modified Salkowski assay (Figure 5-11A).. The empty pOGG027 vector control indicated a similar result with no detectable IAA produced under the experimental conditions used here. There was a detectable amount of IAA produced (8.5×10^{-15} mol cell^{-1}) from the uninduced biosynthesis circuit that was

significantly greater ($p<0.05$) than that of both the wild-type and empty vector controls (Figure 5-11A).

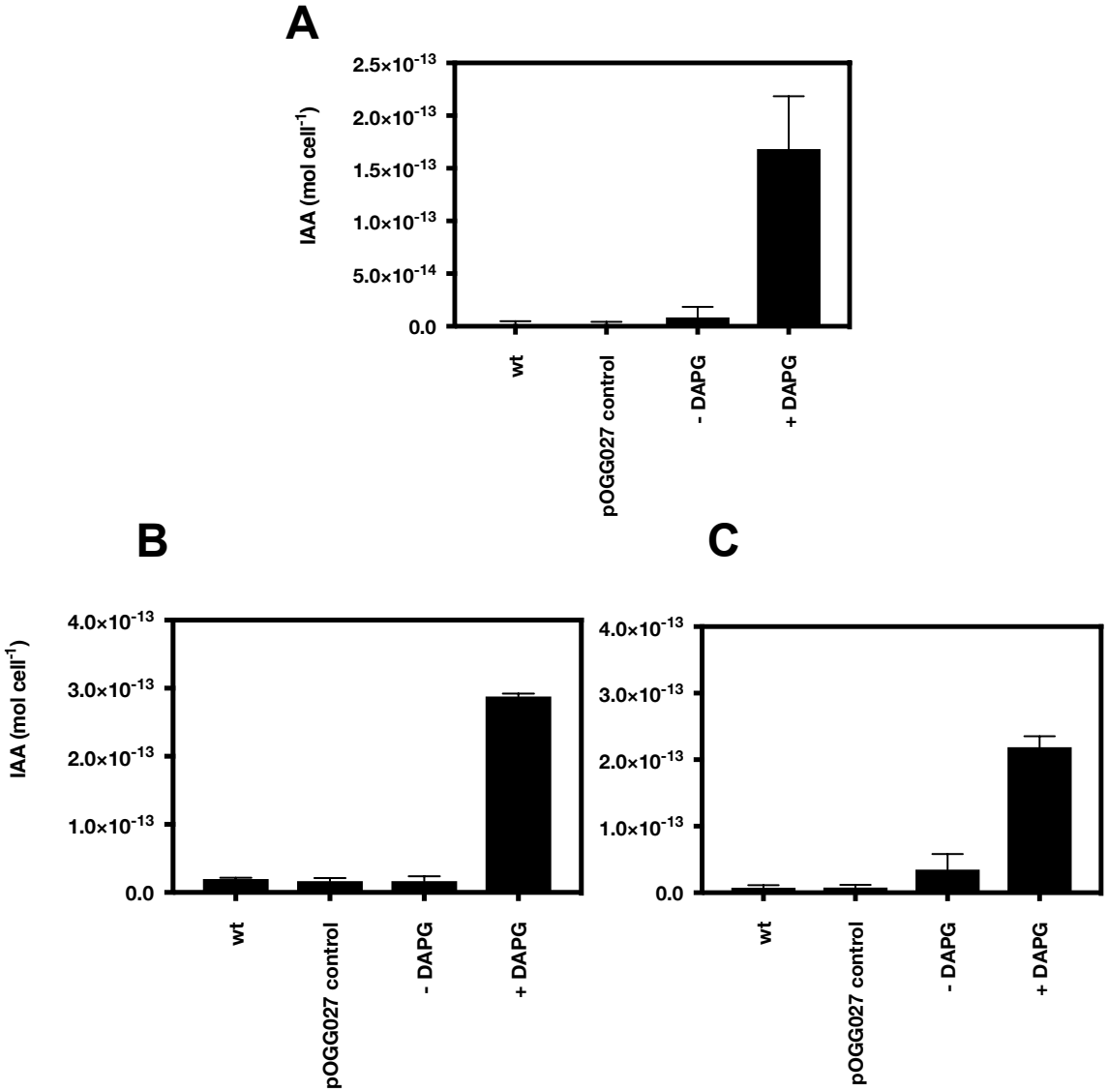


Figure 5-11 - IAA biosynthesis analysis from the DAPG inducible *iaaM* and *iaaH* gene circuit in *E. coli* and two rhizobacterial strains after 48 hours diazotrophic growth. ***A*** – *E. coli*. **B –** *A. caulinodans* ORS571. **C** – *A. brasilense* FP2. wt – Wild-type. Error bars indicate the standard error of the mean. N = 3 biological and 3 technical replicates.

The induction of the IAA biosynthesis circuit with 1 mM DAPG produced $1.7x10^{-13}$ mol $cell^{-1}$ of IAA, a significant 19.8X increase compared to that of the uninduced circuit (Figure 5-11A).

The largest amount of IAA produced of all three chassis was that of the rhizobacteria *A. caulinodans* ORS571. There was a nontrivial amount of IAA produced ($4.1x10^{-14}$ mol $cell^{-1}$) by the wild-type strain in the same order of magnitude as that observed in the constitutive IAA biosynthesis circuit (Figure 5-8B). The pOGG027 empty vector control and the uninduced circuit did not produce a significantly different amount ($p>0.05$) of IAA than that of the wild type control. Induction of the IAA biosynthesis circuit with 1 mM DAPG caused $2.91x10^{-13}$ mol $cell^{-1}$ of IAA to be produced, a 17.2X increase compared to the uninduced state of the circuit (Figure 5-11B).

IAA biosynthesis in the *A. brasilense* FP2 chassis was observed to be less than but within the same order of magnitude as that of the *A. caulinodans* ORS571 chassis (Figure 5-11C). There was a small amount of IAA produced ($7.5x10^{-15}$ mol $cell^{-1}$) by the wild type chassis which was not significantly different to that of the pOGG027 empty vector control ($p>0.05$). The uninduced IAA circuit produced significantly more ($p<0.05$) IAA than both of the control conditions at $3.5x10^{-14}$ mol $cell^{-1}$ (Figure 5-11C). When induced with 1 mM DAPG the biosynthesis circuit produced $2.2x10^{-13}$ mol $cell^{-1}$ IAA which was 6.5X more than when in the

uninduced state. Therefore the dynamic range of the DAPG inducible IAA biosynthesis circuit was 2.9X greater in the *A. caulinodans* ORS571 chassis than the *A. brasilense* FP2 chassis. However neither chassis had dynamic ranges greater than *E. coli* despite less IAA being produced in that chassis than in *A. caulinodans* ORS571.

5.4 Competitive *In vivo* Plant Growth Promotion

5.4.1 Root hair length promotion

An increase In root hair length is one of the primary indicators of early PGP in plant systems. Many studies have shown that optimum IAA concentrations in the rhizosphere cause significant root hair length elongation[270,271]. To analyse the effects of the engineered IAA biosynthesis circuits, a competitive *in vivo* growth experiment was performed. Barley (*Hordeum vulgare*) is one of the top 5 global cereal crops (Figure 1-2) and was the focus of our PGP attempts in this study. The Golden Promise variety of *H. vulgare* was selected as the target plant due to its importance in agricultural systems in the United Kingdom and abroad,[272] as well as its use as a plant transformation system in Bill and Melinda Gates-funded projects to engineer nodulation in cereals. The growth analysis was performed in a competitive rhizosphere environment with soil sampled from the University of Oxford Wytham Woods Research station (51°46'19.5"N 1°19'57.4"W). Covering vegetation was removed and soil collected from a depth of 20-40 cm. The soil is a well-graded and largely plastic clay-silt (BS1377: 1977) shown to contain P_{3-} 122.9 mg kg^{-1}, K_{+} 483.6 mg kg−1 and Mg_{2+} 304.9 mg $kg^{-1}$with a 22% natural moisture content, 16.78% organic content, 58% liquid limit and pH 7.22[273,274]. Soil samples were not sterilised, thus preserving the native microorganism

community, and were stored at room temperature in sealed vessels until required. Soil was sieved with a 2 mm pore steel soil sieve (PASCO, California) before being transferred to 245 cm^2 Nunc Square BioAssay plates (ThermoFischer Scientific, Loughborough, UK) with a 1cm^2 in the top and bottom to allow for shoot growth and drainage respectively. Soil (40g) was added to the top of the plate and then leveled with a ruler. The soil was maintained at 70% capacity by watering with 40 mL of sterile dH_2O every 4 days. No pesticide or other chemical treatment was applied during the duration of the experiment. *H. vulgare* Golden Promise seeds were surface sterilized using 70% (v/v) ethanol and 2% bleach (v/v) before being germinated on water agar plates for two days in the dark at room temperature.

The constitutive IAA biosynthesis circuit in the pOGG096 broad host range plasmid (pOPS0667) was selected as the primary circuit for *in vivo* PGP due to the enhanced IAA production *in vitro* as shown in 5.3.1.1 above. Both the *A. caulinodans* ORS571 and *A. brasilense* FP2 strains were tested for their *in vivo* PGP in a competitive soil environment. Wild-type and 15 µM IAA were used as negative and positive controls respectively in parallel to the IAA biosynthesis circuit-expressing strains. Cultures were grown at either 37˚C (*A. caulinodans* ORS571) or 30˚C (*A. brasilense* FP2) overnight in TY media with the appropriate antibiotics and 200 rpm shaking (2.4.2 above). On the day of inoculation the overnight cultures were sub cultured into fresh TY media and grown in the same growth conditions until an OD595 of 0.6. Cultures were then washed three times with sterile UMS before being diluted to the required inoculation density.



Figure 5-12 – *H. vulgare* Golden Promise growing in competitive soil environment for root hair length analysis. Note clear BioAssay plate and 45˚ growth incline allow for visualisation and microscopic analysis of the entire plant root system.

Seeds were inoculated prior to planting by incubating in 10 mL of UMS with 10^5 cells of the test culture or 10 mL of 15 µM IAA as a positive control. This concentration of IAA was added to the water of the positive control plants at each watering point of the experiment. Inoculation was performed at room temperature

for 30 minutes at 80 rpm on an orbital shaker. Seeds were then planted in the BioAssay plates 1 cm deep and 1 cm from the top edge. Plants were grown in a LGC-960 C growth chamber (Schneider Scientific, Paris) in 16 h light and 8 h dark conditions at 23˚C. The BioAssay plates were arranged randomly within the growth chamber at a 45˚ angle to encourage root growth along the bottom face of the plate (Figure 5-12).

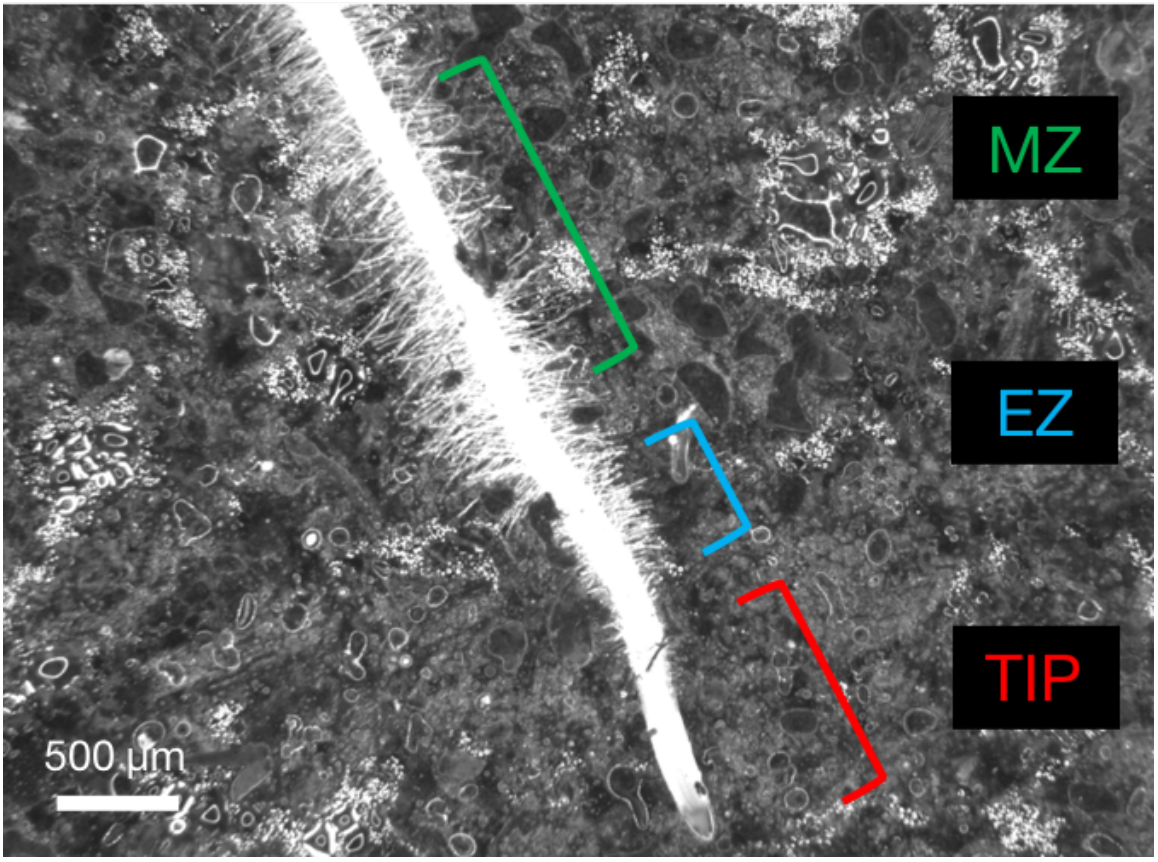


Figure 5-13 – Root tip zone of *H. vulgare* Golden Promise. MZ – Maturation zone. EZ – Elongation zone. Note high contrast post-image processing using Fiji (ImageJ) software[275] for more reliable measurements of root hair length.

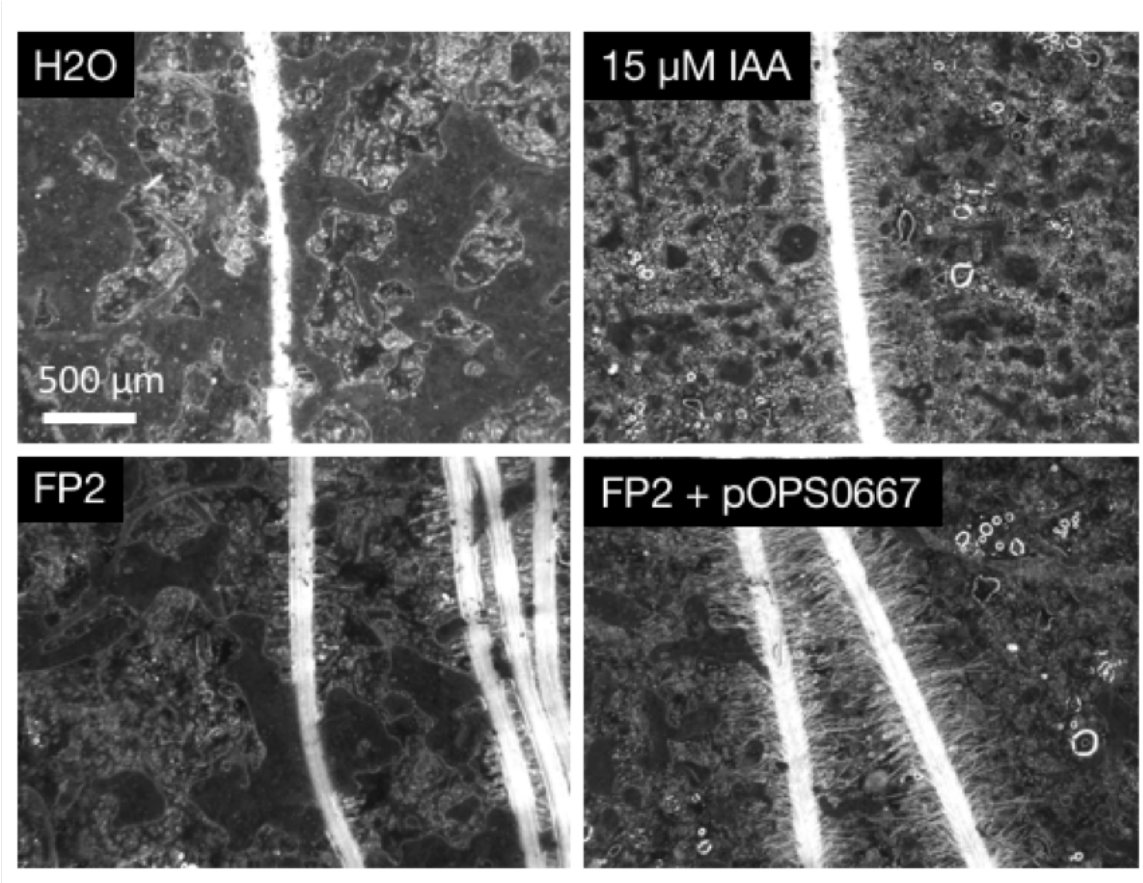


Figure 5-14 – Light microscopy analysis of the effect of a synthetic constitutive IAA biosynthesis circuit (pOPS0667) in *A. brasilense* FP2 on the root hair length of *H. vulgare* plants.

Light microscopical analysis of the root hair length was performed four days post inoculation (DPI) using a Leica stereomicroscope LZ M125 (Leica Scientific, Wetzlar) and Leica LAS X software suite. Images taken for root hair determination were taken exclusively in the maturation zone (MZ) behind the elongation zone (EZ) of the growing root tip (Figure 5-13). This was done due to the natural and wide variation of root hair lengths both below in the EZ and in the upper regions of the root system where lateral root junctions occur. Images of the root hairs in

the MZ were processed in the Fiji image processing suite[275] to enhance the contrast to 75% maximum before root hair length measurement was performed. The length of 10 root hairs on three root MZs was performed for three plants per condition, therefore producing a mean length from 3 biological (independent plants) and 30 technical replicates.

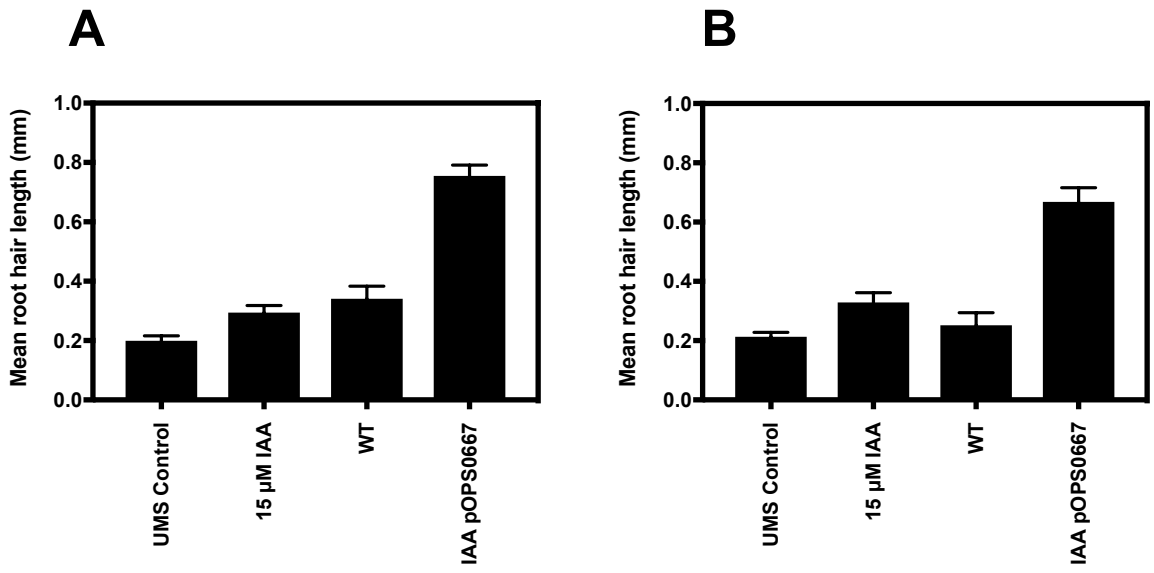


Figure 5-15 – Impact of the synthetic constitutive IAA biosynthesis circuit on the root hair length of *H. vulgare* in a competitive soil system. **A** – *A. caulinodans* ORS571. **B** – *A. brasilense* FP2. Error bars indicate the standard error of the mean, n = 3 biological and 30 technical replicates.

The synthetic constitutive IAA expression circuit (pOPS0667) in both rhizobacterial chassis promoted the length of *H. vulgare* root hairs in the competitive soil environment (Figure 5-14). The root hair length of plants inoculated with the wild type *A. caulinodans* ORS571 chassis (340 µm) were significantly greater ($p<0.05$) than the UMS negative control (199 µm). However

there was no significant difference ($p>0.05$) between the mean root hair lengths of the wild type and 15 µM positive control treatments (294 µm; Figure 5-15A). The root hairs of *H. vulgare* inoculated with the constitutive IAA expression circuit in *A. caulinodans* ORS571 were 2.2X longer than when inoculated with the wild type chassis and 3.8X longer compared to the negative UMS control.

The effect on the root hair length of the constitutive IAA expression circuit in *A. brasilense* FP2 was less pronounced than that of *A. caulinodans* ORS571 (Figure 5-15B). However the root hairs of *H. vulgare* inoculated with pOPS0667 (670 µm) were significantly longer ($p<0.05$) than those of any of the other treatments; 2.6X longer than the wild type treatment (250 µm), 2.1X longer than the 15 µM IAA positive control (330 µm) and 3.2X longer than the UMS negative control treatment (210 µm). There was no significant difference observed ($p>0.05$) between the mean root length of the plants inoculated with the wild type *A. brasilense* FP2 and the UMS negative control. However, the root hairs of the plants inoculated with the 15 µM IAA positive control were significantly longer ($p<0.05$) than both the negative control and wild type strains.

5.4.2 Root length promotion

Another quantitative metric often associated with IAA-induced PGP is that of net root length. Many studies have shown that upon the addition of IAA to an optimum

concentration, the root length of target plants are increased significantly[276]. However, suboptimal or excessive IAA concentrations can often significantly decrease root length[277]. To analyse the effect of the synthetic constitutive IAA biosynthesis circuit, *H. vulgare* was again selected as the target plant of choice. All the experimental preparation and conditions was identical to that of the root hair length assay (see 5.4.1 above) with the following modifications: The growth vessels (rhizoboxes) were custom designed for the analysis of net root length rather than root hair length and as such were much larger (30x60cm). Two sheets of clear PVC were placed 0.5cm using square PVC border spacers and clamped together using foldback clips (Figure 5-16).

Instead of a single *H. vulgare* seed being planted in each plate, three were planted equidistant to one another 1cm from the top of the rhizobox (Figure 5-16). To facilitate drainage of the much larger rhizoboxes, 2 cm of fine grade perlite (GardenersDream, Glasgow) was added to the bottom of the vessel ahead of being filled with the sieved Wytham soil.

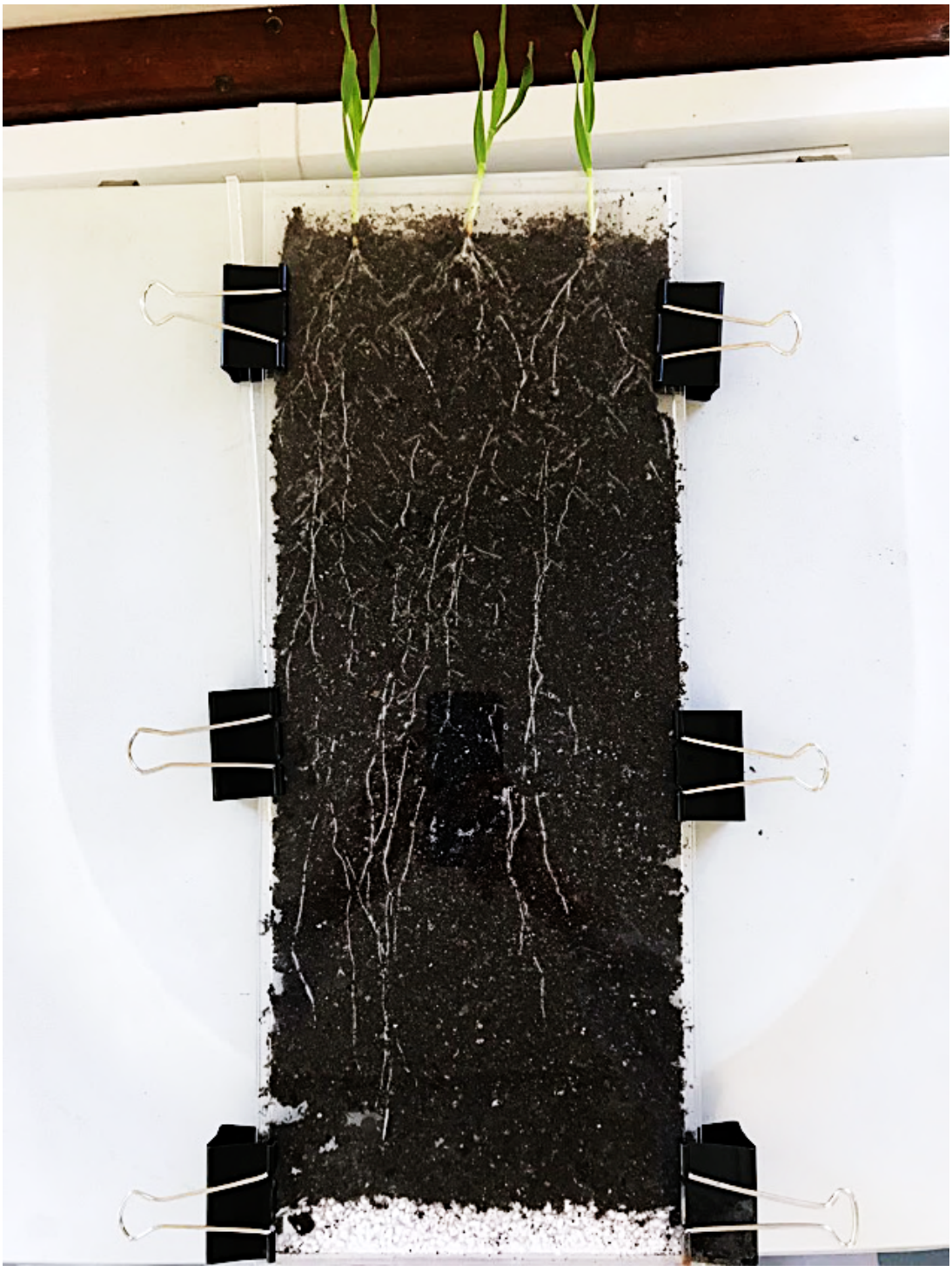


Figure 5-16 - *H. vulgare* Golden Promise growing in competitive soil environment for root length analysis. Note clear rhizobox PVC plates and 45˚ growth incline allow for visualisation and analysis of the entire plant root system.

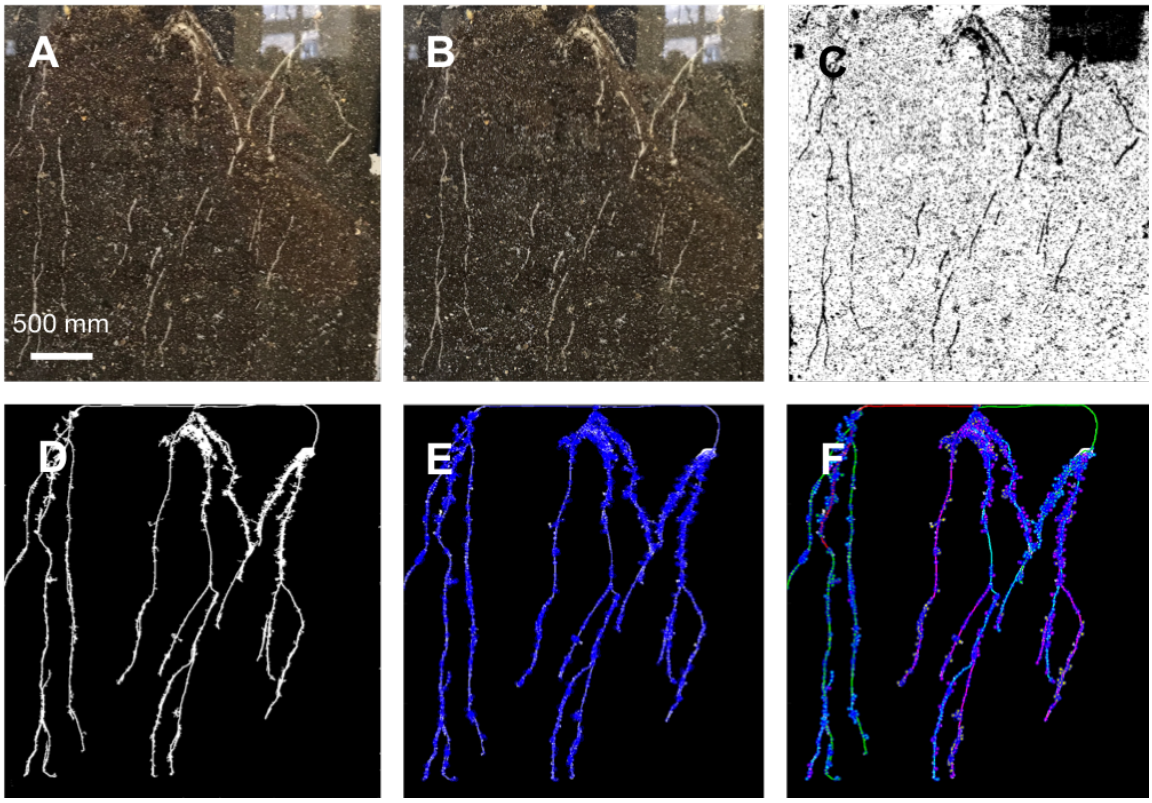


Figure 5-17 – Processing sequence of operations to analyse whole root systems of *H. vulare* in a competitive soil environment. **A** – Raw image taken using a Canon Eos 400D camera. **B** – Image sharpened by a factor of 4X in the Fiji software package[275]. **C** – Image colour table converted to BW. **D** – Image converted to binary bitmap and edge detection and gap-fill functions performed. **E** – Image files transferred to the Root System Analyser package[278] and initial root trace performed. **F** – Full root system analysis with individual root mapping and diameter analysis performed.

Images of the root systems were taken using a Canon EOS 400D DSLR camera (Canon, Reigate) at 4, 8 and 12 DPI (Figure 5-17).. The images were first processed in the Fiji image analysis platform[275] to perform edge detection of the root systems, gap-filling and then finally conversion into a binary bitmap image[279].

The images were then processed in the Root System Analyser software package for individual root tracing, branching structure and number as well as root diameter analysis[278]. Only the three main tap roots of the *H. vulgare* system were quantified and then represented in the root length analysis, all lateral roots were excluded for consistency. Each *H. vulgare* seed produces three tap roots representing the three technical replicates, and each test condition was performed in an individual rhizobox with three seeds representing three biological replicates.

The results from the PGP effect of the synthetic constitutive IAA expression circuit indicated that both rhizobacteria strains caused a significant increase ($p<0.05$) in root length 8 DPI (Figure 5-18). Before this time there was no significant difference ($p>0.05$) in the root lengths of any of the treatments. After 12 DPI the root systems of the *H. vulgare* plants hit the bottom of the rhizobox and the image analysis pipeline was unable to determine root length accurately.

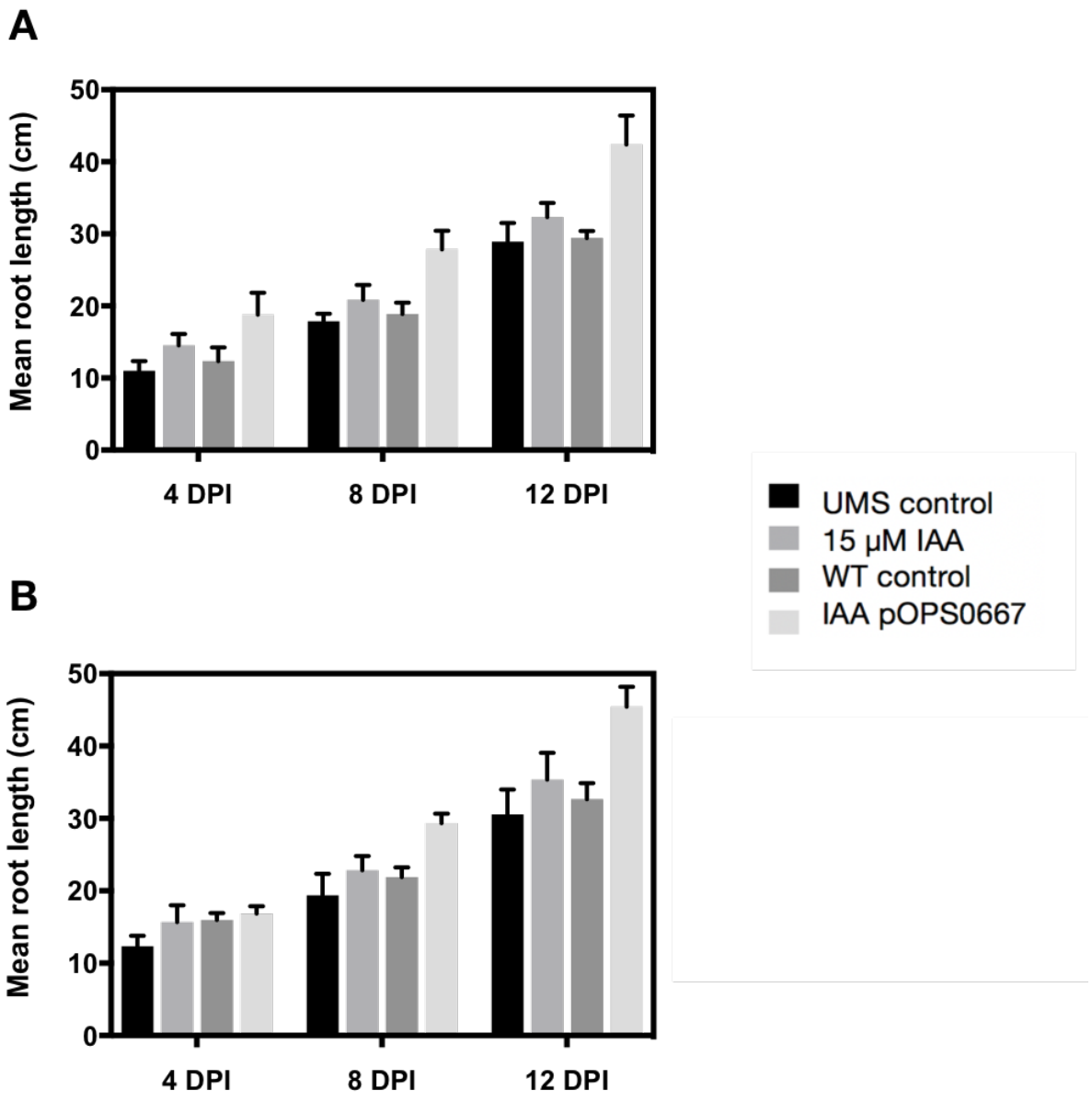


Figure 5-18 - Impact of the synthetic constitutive IAA biosynthesis circuit on the root length of *H. vulgare* in a competitive soil system. **A** – *A. caulinodans* ORS571. **B** – *A. brasilense* FP2. Error bars indicate the standard error of the mean, n = 3 biological and 3 technical replicates.

At 8DPI the root length of the *H. vulgare* plants inoculated with the A*. caulinodans* ORS571 strain containing the IAA biosynthesis circuit (pOPS0667) produced 52% greater root length (29 cm) than those inoculated with the wild type chassis (19 cm). The roots were also significantly longer ($p<0.05$) than those of the UMS negative control (19 cm) and the 15 μM IAA control (22 cm; Figure 5-18A). At 12 DPI the difference in root lengths between the treatments was more pronounced with the IAA circuit inducing 43% longer roots (45 cm) than the wild type control (31 cm), 35% longer than the 15 μM IAA positive control (34 cm) and 53% longer than the UMS negative control (31 cm; Figure 5-18A).

The impact of the IAA biosynthesis circuit (pOPS0667) in *A. brasilense* FP2 was comparable to that of *A. caulinodans* ORS571. At 8DPI the root length of the *H. vulgare* plants inoculated with the IAA biosynthesis circuit produced 31% greater root length (29 cm) than those inoculated with the wild type chassis (22 cm). The roots were also significantly longer ($p<0.05$) than those of the UMS negative control (19 cm) and the 15 μM IAA control (24 cm; Figure 5-18B). At 12 DPI the difference in root lengths between the treatments was again more pronounced with the IAA circuit inducing 42% longer roots (47 cm) than the wild type control (33 cm), 34% longer than the 15 μM IAA positive control (35 cm) and 51% longer than the UMS negative control (31 cm; Figure 5-18B).

5.4.3 Root diameter analysis

Concurrent to the analysis of the effect of the synthetic IAA biosynthesis circuit (pOPS0667) on *H. vulgare* root length, the average root diameter was determined using the Root System Analyser suite[278] (5.4.2 above). The plant roots that were used to determine the mean root diameter were identical to those used in the root length analysis to maintain consistency between the assays.

At 12 DPI the root systems were at their most developed and therefore this time point was selected for the analysis and presentation of mean root diameters. The root diameter for the *H. vulgare* plants inoculated with the synthetic IAA biosynthesis circuit was significantly less ($p<0.05$) than all other treatments (Figure 5-19). This was a pattern observed in both the *A. caulinodans* ORS571 and *A. brasilense* FP2 strains. The *H. vulgare* roots inoculated with the *A. caulinodans* ORS571 strain with the IAA expression circuit showed a 31% decrease in the root diameter (340 µm) compared to the plants inoculated with the wild type strain (580 µm; Figure 5-19A). There was no significant difference ($p>0.05$) in the plant root diameters observed when inoculated with the wild type strain compared to the UMS negative control. There was however a significant 12% decrease in the root diameters of the plants treated with the 15 µM IAA control compared to the wild-type inoculation (Figure 5-19A).

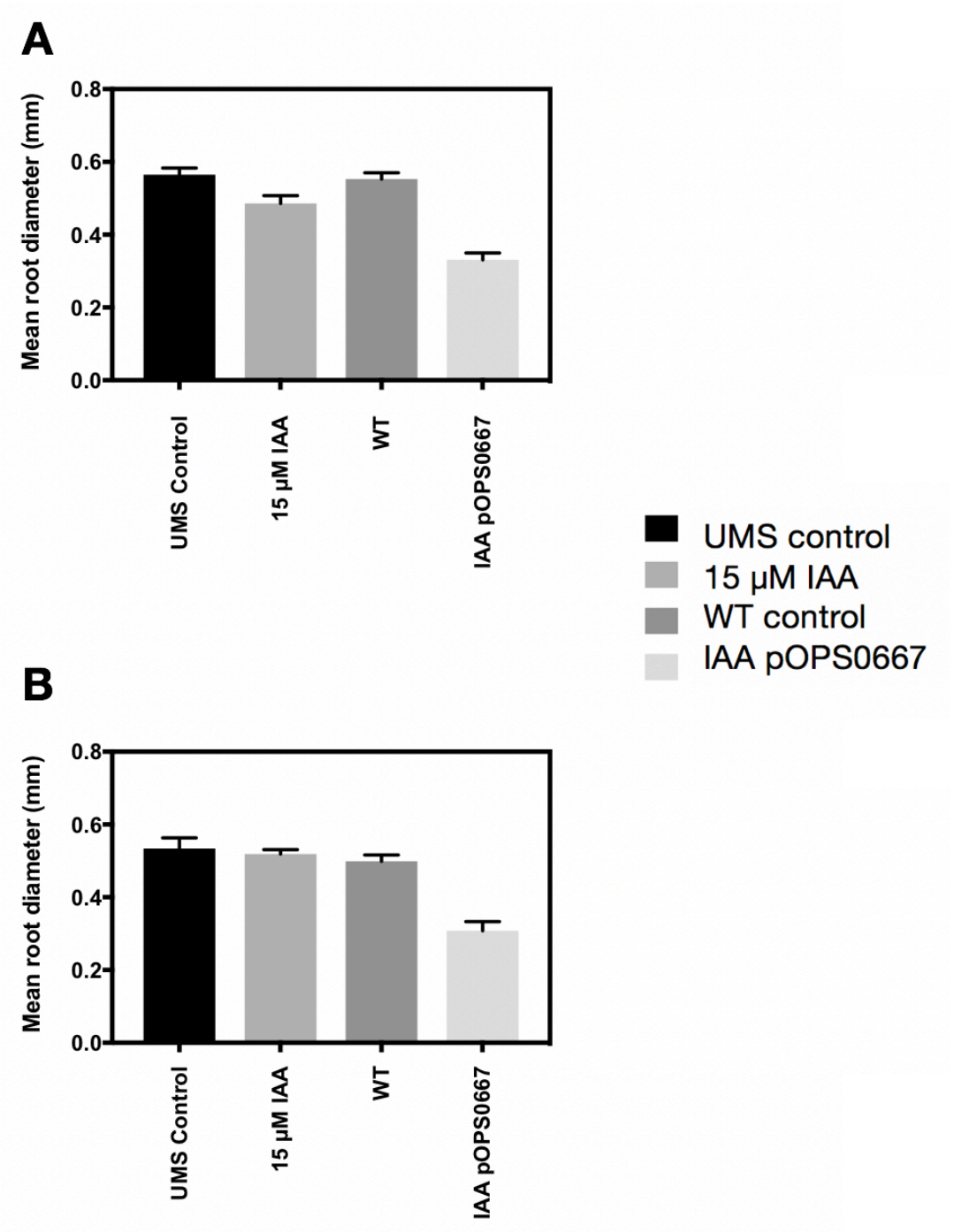


Figure 5-19 - Impact of the synthetic constitutive IAA biosynthesis circuit on the root diameter of *H. vulgare* plants in a competitive soil system. **A** – *A. caulinodans* ORS571. **B** – *A. brasilense* FP2. Error bars indicate the standard error of the mean, n = 3 biological and 3 technical replicates.

The *H. vulgare* plants inoculated with the *A. brasilense* FP2 chassis also showed a significant ($p<0.05$) 28% decrease in root diameter when inoculated with the constitutive IAA expression circuit (310 μm) compared to the wild type (480 μm; Figure 5-19B). There was no significant difference observed for the root diameters ($p>0.05$) between any of the control treatments in the experimental conditions used in this study.

5.5 Engineering DAPG Biosynthesis

The pathway we chose to engineer DAPG biosynthesis is based upon that of the *phl* pathway from the soil bacterium *Pseudomonas fluorescens* F113, a strain well known for its antifungal properties[280]. There are six open reading frames in the *phl* region, four of which comprise of the core operon *phlACBD*. The PhlD protein is a polyketide synthase which has shown homology to chalcone and stilbene synthases found in some plants[176]. The other two gene of the pathway, *phlE* and *phlF* encode a putative efflux and regulatory repressor protein respectively. The PhlF repressor has been used extensively in this work (see 3.3.4 aboveand 4.3.1 above).

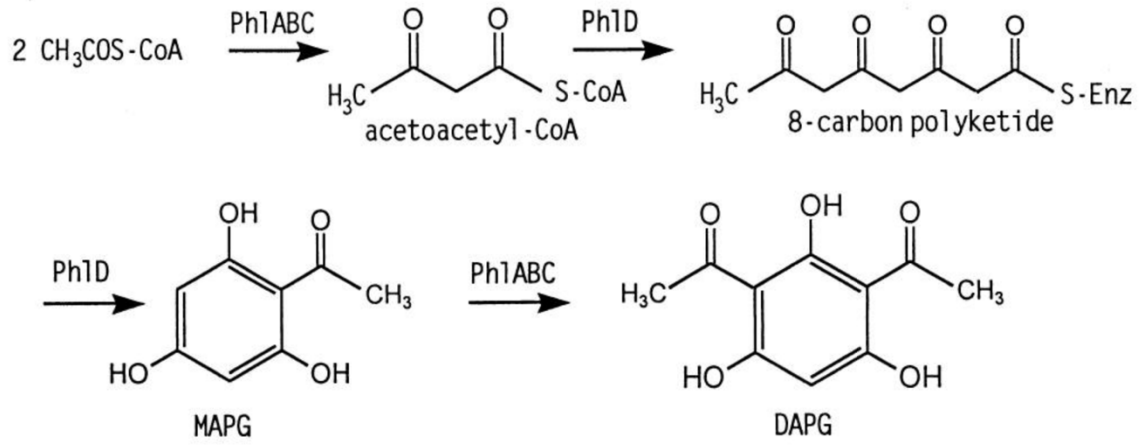


Figure 5-20 -The putative DAPG biosynthesis pathway from *P. fluorescens* F113. Note that the *phl* genes perform multiple steps in the pathway. Enz – enzyme. Adapted from Bangera and Thomashow, 1999.

Previous studies expressing the *phlACBD* genes in *E. coli* were able to characterise the novel polyketide biosynthesis pathway[176]. PhlD was identified as the enzyme responsible for the production of monoacetylphloroglucinol (MAPG) the precursor to DAPG. PhlA, PhlB and PhlC are all necessary for the conversion of MAPG to DAPG and may also play a role in the synthesis of MAPG (Figure 5-20).

For the ease of modular synthetic circuit construction and cloning, the Type IIS Golden Gate cloning pipeline was utilised for the construction of all IAA biosynthesis circuits (2.5.7 above). The cloning of the following plasmids was performed in partnership with Dr Barney Geddes of the Department of Plant Sciences, University of Oxford. All assays were performed by the author. The *phlABC* portion of the pathway required extensive domestication to remove BsaI and BpiI restriction sites that would otherwise interfere with the downstream Golden Gate cloning pathway (2.5.7 above). The primers used to perform the domestication are listed in Table 5-1. Once the operon had been domesticated, the entire region was re-amplified using primers oxp2074 and oxp2075 and cloned into the pL0M-SC plasmid (pOGG072) to make the plasmid pL0M-SC-phlABC (pOGG199). The genes *phlD* and *phlE* were not able to be domesticated by conventional methods and was therefore synthesized with codon optimization and type IIS restriction enzyme site removal (2.5.2 above).

Table 5-1 - Table of primers used to amplify and domesticate the *phl*ABC operon for rhizobacterial synthetic biology.

Primer Number	**Primer Name**
oxp1137	*phlABC* Forward 1
oxp1138	*phlABC* Forward 2
oxp1139	*phlABC* Forward 3
oxp1140	*phlABC* Forward 4
oxp1141	*phlABC* Reverse 1
oxp1142	*phlABC* Reverse 2
oxp1143	*phlABC* Reverse 3
oxp1144	*phlABC* Reverse 4

The domesticated *phlD* gBlock was amplified with primers oxp2077 and oxp2078 and then cloned into the pL0M-SC plasmid (pOGG072) to make the plasmid pL0M-SC-phlD (pOGG200). The gene *phlE* was was amplified with primers oxp1151 and oxp1152 and then cloned into the pL0M-SC plasmid (pOGG072) to make the plasmid pL0M-SC-phlE (pOGG201).

5.5.1 Constitutive DAPG Biosynthesis

To analyse the DAPG biosynthesis ability of the rhizobacterial strains using the *P. fluorescens* F113 *phl* operon, a constitutive DAPG expression circuit was designed and constructed using the level one and level two Golden Gate cloning technique (2.5.7 above). Each of the parts of the operon were expressed from a medium strength J23104 promoter (3.3.1 above) and standard RBS sequence (3.3.2 above) and terminated with heterologous terminators (3.3.3 above). Firstly three level one plasmids were constructed with transcriptional units of the DAPG biosynthesis circuit in each. The first plasmid contained the three initial genes of the circuit *phlABC*. The *phl* operon pL0M-SC-phlABC (pOGG199) was cloned into the level one holding vector pL1M-F1 (pOGG021) with the medium strength promoter pL0M-P-J23106 (pOGG0121), pL0M-U-RBSstd (pOGG143) and no terminator to allow read through of the final operon (Figure 5-21A). The second plasmid (pOGG215), the *phlD* transcriptional unit was constructed in the pL1M-F2 plasmid with the pL0M-SC-phlD open reading frame (pOGG200), a medium strength promoter pL0M-P-J23106 (pOGG0121), pL0M-U-RBSstd (pOGG143) and pL0M-T-DT15S terminator (pOGG156; Figure 5-21B). The final level one plasmid (pOGG214) contained the pL0M-SC-*phlE* gene encoding the putative MFP permease expressed from a medium strength promoter pL0M-P-J23106 (pOGG0121) with the pL0M-U-RBSstd RBS (pOGG143) and pL0M-T-F2 terminator (pOGG160; Figure 5-21C).

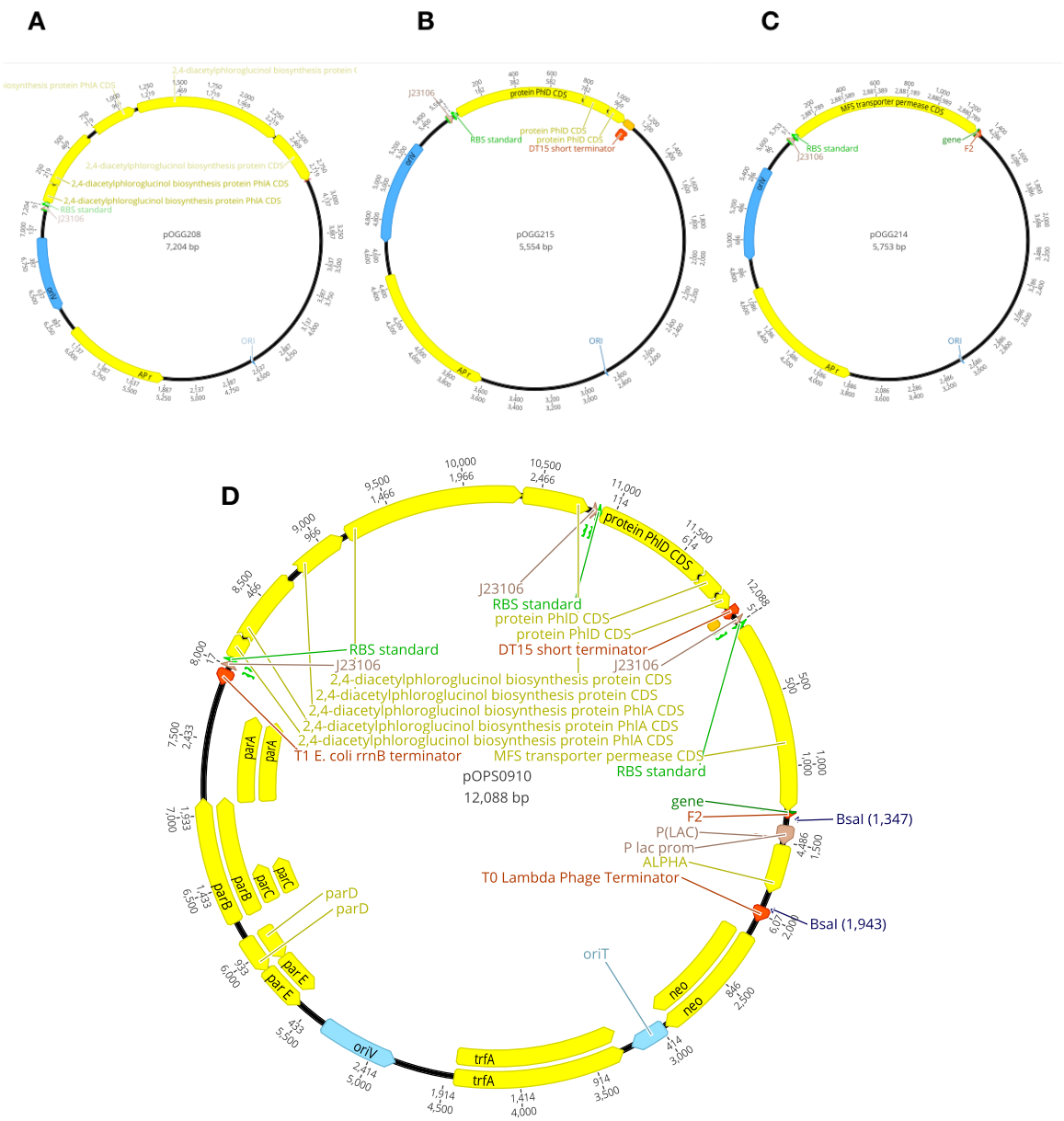


Figure 5-21 - Plasmid maps of the three level one DAPG biosynthesis pathway plasmids and level two complete constitutive DAPG expression circuit. **A** – pOGG208. PhlABC expression cassette in pL1V-F1. **B** - pOGG215. PhlD expression cassette in pL1V-F2. **C** – pOGG214. PhlE expression cassette in pL1V-F3. **D** – Final constitutive DAPG biosynthesis circuit pOPS0910. Note: The oriV *E. coli* origin of replication and ampicillin selection gene in the level 1 plasmids. The RK2 origin of replication and kanamycin/neomycin selection gene and *par* operon in the final level pOPS0910 circuit.

The three level one plasmids were combined in a level two Golden Gate reaction (2.5.7 above) in the pL2V-Lv2-Neo-RK2-Par-ELT4 (pOGG096) plasmid to make the full constitutive DAPG circuit pOPS0910 (Figure 5-21D).

The constitutive DAPG expression circuit (pOPS0910) was transformed into *E. coli* by heat shock (2.6.1 above) and into the rhizobacterial chassis *A. caulinodans* ORS571 by electroporation (2.6.2 above). Wild-type negative controls were included to normalise the DAPG biosensor response to the synthetic biosynthesis circuit. Due to transformation issues the circuit was not tested in *A. brasilense* FP2 despite 3 attempts to do so. To analyse the production of DAPG in both *E. coli* and *A. caulinodans* ORS571, a high-throughput biosensor assay was used. A fully characterised DAPG biosensor (pAJM.847) was donated to us by Dr Minhyung Ryu of the Voigt Lab, MIT[281]. The pAJM.847 biosensor is a pBBR1 based plasmid with a sfGFP reporter plasmid expressed from a Pphl promoter. The transcriptional regulator PhlF is expressed from a medium strength constitutive promoter, therefore repressing sfGFP expression until the addition of DAPG. The strains containing the synthetic DAPG biosynthesis circuits were grown to an OD595 of 0.6 in 10 ml of TY media (*A. caulinodans* ORS571) at 37˚C with 200 rpm shaking and the relevant antibiotics (2.7.1 above). Cultures were then centrifuged at 4500g for 10 minutes at 4˚C to pellet the cells and debris before the supernatant was filtered through a 0.22 µm filter. The cell number in the pellet was calculated by both OD595 and dry weight measurements. 100 µL of the filtered supernatant was then added to 900 µL of fresh TY media and

inoculated with 10^6 cells of the *E. coli* DH5a DAPG biosensor strain. These cultures were then transferred to a 24-well microwell plate before high-throughput growth and fluorescence response analysis as described in section 2.7.3 above. A DAPG standard curve was made in parallel to the growth and fluorescence measurements to calculate the amount of DAPG in the spent media (Supplementary Figure 2). The absolute DAPG produced in μmol $cell^{-1}$ was then calculated using the quantitative measurements of the original culture cell number. Five biological and three technical replicates were used to calculate the mean DAPG produced by each test culture.

The analysis of the DAPG biosynthesis circuit in *E. coli and A. caulinodans* ORS571 indicated DAPG was produced in both (Figure 5-22). There was no significant ($p>0.05$) DAPG produced or observed in any of the control conditions except the biosensor alone condition whereby some fluorescence response was observed above the minimum limit of the assay. The DAPG biosynthesis circuit (pOPS0910) in *E. coli* produced a maximum $2.8x10^{-7}$ μmol $cell^{-1}$ of DAPG. In comparison in *A. caulinodans* ORS571, 3.5X more DAPG at $9.7x10^{-7}$ μmol $cell^{-1}$ was produced.

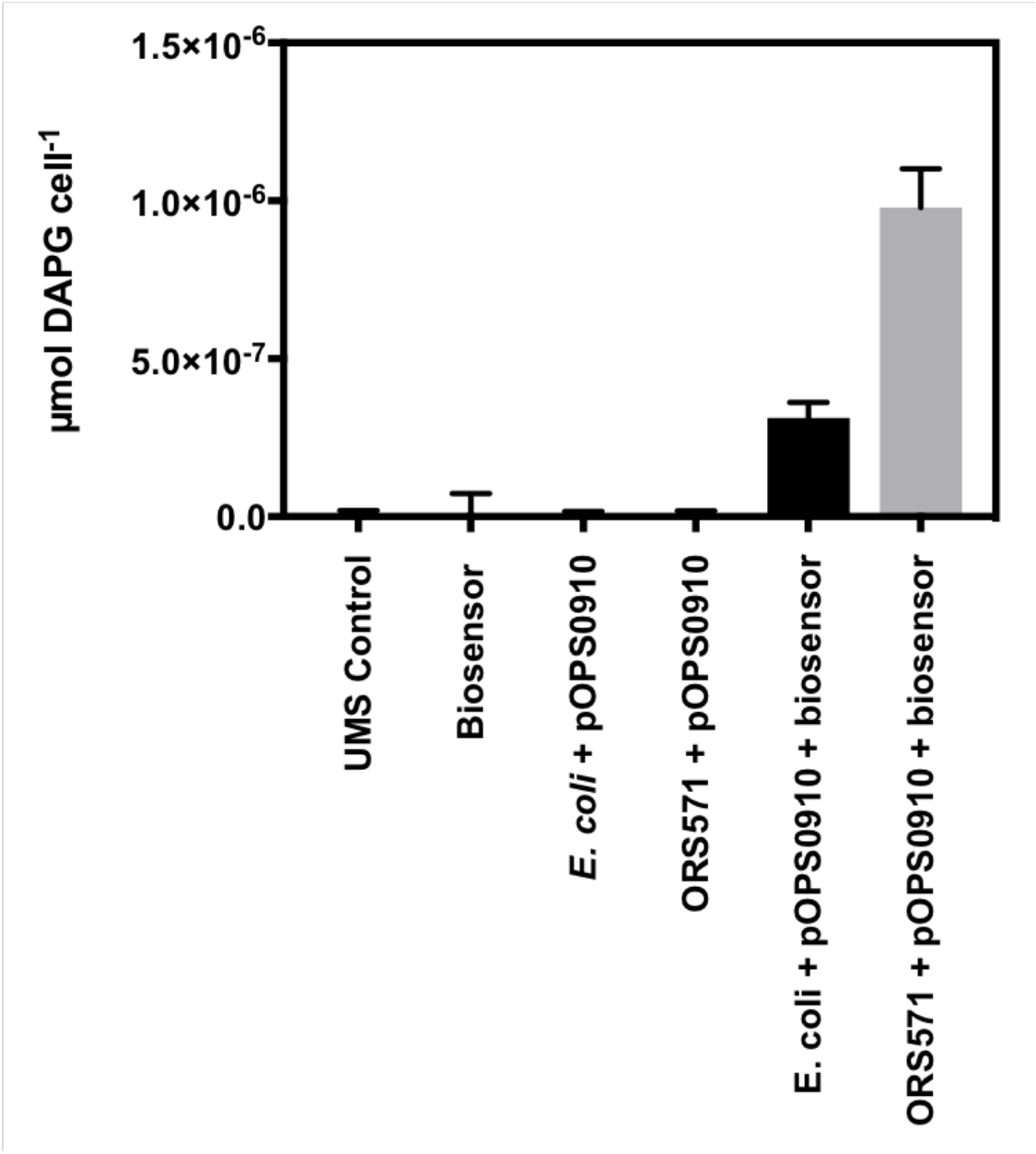


Figure 5-22 - DAPG biosynthesis from the synthetic constitutive circuit (pOPS0910) in *E.* coli and *A. caulinodans* ORS571. Biosensor = pAJM.847. Error bars indicate the standard error of the mean. N = 5 biological and 3 technical replicates.

5.6 Phosphate solubilisation

Phosphorus is one of the major nutrients limiting plant growth (see 5.2.2 above) and is therefore of significant interest in the field of engineering PGPR with synthetic biological techniques. Several studies have characterised and confirmed the phosphate solubilising ability of wild type bacteria such as the *Pseuodomonas* spp.[89]. It is well known that the main mechanism by which natural PGPR solubilise phosphate is by the secretion of organic acids[241] (see 5.2.2 above). The secretion of gluconic acid from the oxidation of cellular glucose is the primary mechanism in nearly all Gram-negative bacteria[282]. Studies have confirmed that gluconic acid biosynthesis is performed by the periplasmic glucose dehydrogenase involving pyrroloquinoline (PQQ) as the essential cofactor. Expression of the *pqq* gene cluster along with a transport protein in the non-phosphate-solubilising chassis *Herbaspirillum seropedicae* Z67 yielded a robust phosphate solubilization phenotype[283]. The efficacy of phosphate solubilization of PGPR is dependent on the concentration of aldose-based sugars present in the root exudates of associated plants. Therefore many phosphate solubilising PGPR are limited in their ability to produce gluconic acid and yield significant amounts of soluble phosphate in the rhizosphere. One solution to this inefficient efficacy is the production of stronger organic acids that solubilise phosphate more readily. Studies have revealed that the weakest organic acids are the acetic, succinic and lactic acids followed by gluconic, oxalic and tartaric acid.

The strongest and most powerful of the organic acids is citric acid[284]. We therefore predict that an engineered PGPR secreting citric acid will be significantly more effective at phosphate solubilisation than natural strains. Previous studies have determined that the overexpression of the citrate synthase gene *gltA* is from *E. coli* can cause enhanced citric acid biosynthesis within the PGP species *Pseudomonas fluorescens* ATCC13525[285]. However, the study found that the endogenous H^+-dependant citrate transporter was ineffective at exporting the high concentration of citric acid into the media. Therefore we believe that the use of a orthogonal transporter from the species *Salmonella typhimurium,* with confirmed high efflux rates of citric acid[286], may allow for more efficient export of citric acid in the rhizosphere.

The gene responsible for the citric acid efflux *citC* and the citric acid synthase gene *gltA* were made into pL0M-SC Golden Gate parts for efficient downstream cloning by Golden Gate reactions (2.5.7 above). The *citC* gene sequence was amplified from *S. typhimurium* DT2 gDNA (kindly donated by Professor Stephen Baker, University of Oxford) using primers oxp2040 and oxp2041 (2.5.2 above) and then cloned in to the plasmid pL0V-SC (pOGG072) in a level 0 Golden Gate reaction to create plasmid pL0M-SC-CitC (pOGG246). The *gltA* gene from *E. coli* DH5a was heavily populated with BsaI and BpiI restriction enzyme sites and was therefore synthesized with codon optimization and restriction enzyme site removal by IDT (2.5.10 above). The *gltA* gene sequence was flanked by BpiI sites

compatible with the pL0M-SC plasmid and therefore was directly cloned by level 0 Golden Gate reaction to create pL0M-SC-GltA (pOGG247).

5.6.1 Constitutive Citric Acid Biosynthesis

To test the function of the novel citric acid biosynthesis pathway, the constitutive expression of *gltA* and *citC* in the synthetic biology mainstay *E. coli* and the two rhizobacterial strains *A. caulinodans* ORS571 and *A. brasilense* FP2 was required. To construct the two-gene expression cassette, DNA parts from the RhizoKit toolbox were implemented (see chapter 3 above). Initially two expression cassettes were constructed for the *gltA* and *citC* genes in level one Golden Gate reactions before being combined into a single stable broad host range plasmid by a level 2 reaction. We predict that the flux of citric acid outside of the cell would be the limiting factor in the system and therefore expressed the *citC* gene more highly than the *gltA* gene.

The *gltA* gene was cloned into the pL1V-F1 (pOGG021) vector with a medium strength pL0M-P-J23106 promoter (pOGG121), a pL0M-U-RBSstd (pOGG143) and terminated by the pL0M-T-F1S terminator (pOGG159; Figure 5-23A). The *CitC* gene was cloned into the pL1V-R2 (pOGG059) vector with a high strength pL0M-P-J23104 promoter (pOGG120), a pL0M-U-RBSstd (pOGG143) and terminated by the pL0M-T-DT16S terminator (pOGG157; Figure 5-23B).

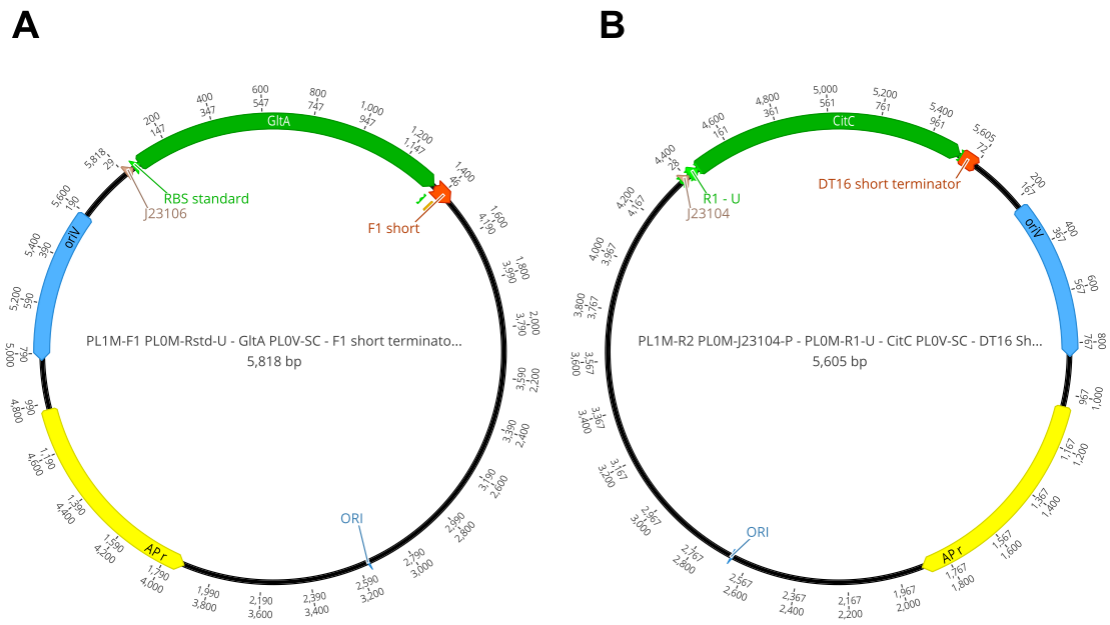


Figure 5-23 - Plasmid maps of the two level one novel citric acid biosynthesis pathway plasmids. **A** – GltA expression cassette in pL1V-F1 (pOGG238). **B** - CitC expression cassette in pL1V-R2 (pOGG239). Note the oriV *E. coli* origin of replication and ampicillin selection genes.

Both the level one plasmids were then purified (2.5.1 above) before being cloned in a level two Golden Gate reaction (2.5.7 above) into both the pOGG027 and pOGG096 pL2V vectors to create pOPS1182 and pOPS1183 (Figure 5-24). We chose to clone into both level two vectors to analyse which plasmids yielded enhanced citric acid biosynthesis profiles. The expression cassettes for each of the *gltA* and *citC* genes were orientated so that their transcriptional directions converged, preventing read through despite the terminator strength validation (3.3.3 above).

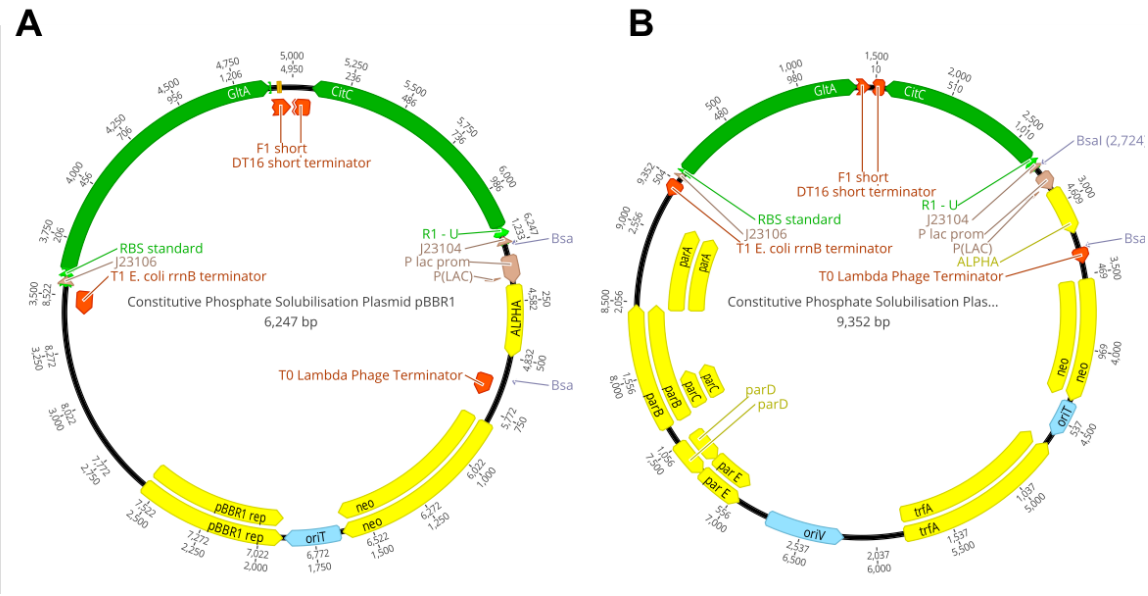


Figure 5-24 - Plasmid maps of the two novel constitutive citric acid biosynthesis pathway circuits. **A** – Constitutive citric acid biosynthesis circuit based in the pOGG027 broad host range plasmid (pOPS1182). Note the pBBR1 replication of origin and kanamycin/neomycin selection. **B** - Constitutive citric acid biosynthesis circuit based in the pOGG096 broad host range plasmid (pOPS1183). Note the RK2 replication of origin, kanamycin/neomycin selection and *par* operon for enhanced maintenance of the plasmid in culture.

Both pOPS1182 and pOPS1183 were transformed into *E. coli* by heat shock (2.6.1 above). The plasmids were transformed into the two rhizobacterial strains *A. caulinodans* ORS571 and *A. brasilense* FP2 by electroporation (2.6.2 above).

To analyse the phosphate solubilisation efficacy of the synthetic citric acid biosynthesis circuit, a di-calcium phosphate assay was developed based on that of the method used by Nautiyal, 1999 in screening for natural phosphate solubilising bacteria[287]. The media used to assay phosphate solubilising ability of

the engineered rhizobacteria was based on Pikovskayas agar (with modifications) and consists of: 3 g L^{-1} yeast extract, 6 g L^{-1} tryptone, 100 mM glucose, 5 g L^{-1} $Ca(H_2PO_4)_2$ and 15 5 g L^{-1} agar. The ingredients were all combined ahead of autoclaving and finally plate pouring and cooling at room temperature to set. Each of the test strains containing the pOPS1182 and pOPS1183 plasmids as well as wild-type controls were grown to an OD595 0.6 at either 37˚C (*E. coli* and *A. caulinodans* ORS571) or 30 ˚C (*A. brasilense* FP2) in TY media with the relevant antibiotics (2.4.3 above). The cultures were then washed three times with 10 mL of sterile UMS media at 4500g. The pellets were finally resuspended in 100 µL of sterile UMS media ahead of inoculation. Three microliters of $1x10^6$ cells or UMS was then aseptically spotted in triplicate on the prepared agar plates and dried completely in a sterile laminar flow hood. Plates were then incubated at either 37˚C (*E. coli* and *A. caulinodans* ORS571) or 30 ˚C (*A. brasilense* FP2) for 72 hours. The phosphate solubilisation index (PSI) was determined by using the below formula[288]:

$$Phosphate\ solubilisation\ Index\ (PSI) = \frac{colony\ diameter + halo\ zone\ diameter}{colony\ diameter}$$

Each presented value in the results below is a mean value generated from three biological and three technical replicates.

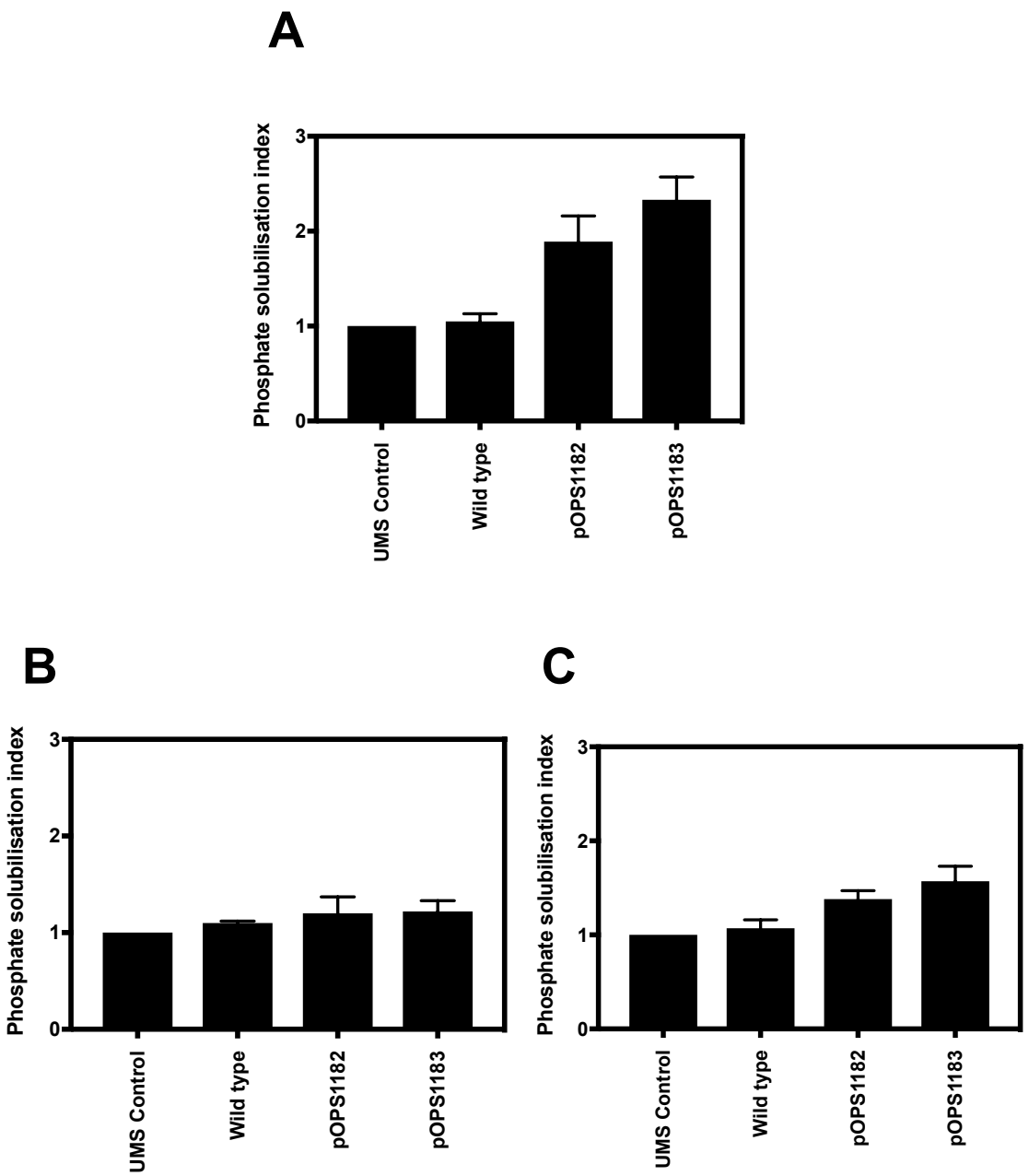


Figure 5-25 – Citric acid biosynthesis and phosphate solubilisation analysis from the constitutive GltA and CitC gene circuits in *E.* coli and two rhizobacterial strains after 72 hours growth. ***A*** – *E. coli.* **B –** *A. caulinodans* ORS571. **C** – *A. brasilense* FP2. Error bars indicate the standard error of the mean. N = 3 biological and 3 technical replicates.

The results from the phosphate solubilisation assay indicated that the synthetic citric acid circuit functioned well and yielded significant phosphate solubilisation ($p<0.05$) in both *E. coli* and to a lesser extent *A. brasilense* FP2 (Figure 5-25). We observed no significant difference ($p>0.05$) in phosphate solubilization for the circuit in *A. caulinodans* ORS571 compared to the wild type strain or UMS control (Figure 5-25B). In *E. coli* the circuit pOPS1182 caused an 85% increase in the PSI (1.89) and the circuit pOPS1183, with the additional *par* stability operon, yielded a significantly greater ($p<0.05$) 128% increase in the observed PSI (2.33) compared to the wild-type strain (Figure 5-25A). There was no significant difference ($p>0.05$) observed in the PSI between the wild type and UMS control. The PSIs observed for the circuits in the *A. brasilense* FP2 strain indicated phosphate solubilisation but to a much lower efficacy than that of the *E. coli* strain (Figure 5-25C). The pOPS1182 circuit caused a significant ($p<0.05$) 31% increase in the PSI (1.38) compared to the wild type. The pOPS1183 circuit caused significantly more ($p<0.05$) phosphate solubilisation with a PSI 50% more (1.57) than the wild type and 19% more than the pOPS1182 circuit. We therefore conclude that the circuits functioned most effectively in the *E. coli* strain with some phosphate solubilisation observed in *A. brasilense* FP2. There was no observable phosphate solubilisation in *A. caulinodans* ORS571. Due to time constraints the work on *in vivo* testing of the synthetic phosphate solubilising circuit was not completed. Future directions for this research is discussed in section 7 of this work.

5.7 Discussion

The goal of this chapter was to select, build and characterise a suite of the most fundamental PGP mechanisms: IAA (auxin) biosynthesis, DAPG antifungal biosynthesis and phosphate solubilisation. Our strategy to control nitrogen fixation and ammonia efflux is the basis of the next chapter. We were successful in the construction and analysis of all three PGP mechanisms but observed the most striking results for the biosynthesis of IAA.

5.7.1 IAA Biosynthesis

We were able to build and characterise an IAA biosynthesis circuit using the *iaaH* and *iaaM* genes from *P. savastanoi*. Both circuits (pOPS0666 and pOPS0667) with stoichiometrically optimised constitutive gene expressions were able to produce a maximum of 4.8×10^{-13} mol $cell^{-1}$ IAA in the rhizobacterial strain *A. caulinodans* ORS571. This amount of IAA excreted is less than observed in previous studies expressing the genes responsible for the indole-3-pyruvic acid pathway in an *E. coli* DH5a strain[289]. However in that study the strain was grown in rich media with a greater concentration of the substrate L-tryptophan. There was significantly more IAA excreted in strains with the pOPS0667 circuit, likely due to the *par* gene cluster providing enhanced stability of the circuit in the liquid culture. Further analysis of the presence of the plasmid and circuit within using qPCR during the

culture growth profile is required to test this hypothesis. The magnitude more IAA produced in the *A. caulinodans* ORS571 strain compared to that of the same circuits in *A. brasilense* FP2 was not expected (Figure 5-8). We predicted that the *A. brasilense* FP2 strain would produce significantly more IAA than the *A. caulinodans* ORS571 strain due to the endogenous IAA genes already present and characterised[42]. Previous studies have also shown that the expression of a single *ipdC* gene in *A. brasilense* FP2 augmented the strains endogenous IAA biosynthesis circuitry and produced significant amounts of IAA and even confirmed PGP effects in common wheat[37]. We hypothesise that the available tryptophan pool in *A. caulinodans* ORS571 is greater than that of the *A. brasilense* FP2, or that possibly the IaaH and IaaM proteins function more effectively in the intracellular environment of *A. caulinodans* ORS571. To test these hypotheses, reversed-phase high-performance liquid chromatography (HPLC) could be implemented to test for the tryptophan concentrations in the wild type *A. caulinodans* ORS571 and the *A. brasilense* FP2 strains. There may also be a variation in the feedback loops of the IAA and tryptophan biosynthesis pathways that diverts IAA biosynthesis more readily in *A. brasilense* FP2. This was observed in mutants of *A. brasilense* FP2 defective in the negative feedback repression of anthralinate synthetase that continued to produce large concentrations of tryptophan even with high intracellular IAA concentrations[290]. The characterisation of the constitutive promoters and RBS sequences in both strains (3.3.1 aboveand 3.3.2 above) proves that the expression of both *iaaH* and *iaaM* is likely to be

equivalent and therefore chassis-specific transcription and translation effects are not the cause of the difference in IAA biosynthesis[53,291].

The DAPG inducible IAA circuit (pOPS1166) functioned similarly in *A. caulinodans* ORS571 and *A. brasilense* FP2 (Figure 5-11). This was unexpected after the results from the constitutive IAA biosynthesis circuit analysis showed a magnitude difference in IAA biosynthesis (5.3.1.1 above). The DAPG inducible rPU part (pOGG137) utilised in the construction of the pOPS1166 circuit functions comparably in both strains with similar dynamic ranges (3.3.4 above). Therefore the induction of the circuit at the beginning of the experiment rather than in mid log phase, as was the case in the rPU characterisation, must create a context where the absolute amount of IAA measured at the time point selected was similar. The temporal IAA biosynthesis characteristics for both strains up to that point must vary greatly as has been seen in many studies using t=0 induction conditions[292–294]. To enhance the IAA biosynthesis in either the constitutive (pOPS0666 & pOPS0667) or DAPG inducible (pOPS1166) circuits for the purpose of IAA purification in large quantities, tryptophan addition to the growth media would remove the substrate as the limiting factor in the biosynthesis pathway.

The results from the PGP effects of the constitutive auxin biosynthesis circuit (pOPS0667) on *H. vulgare* Golden Promise plant roots are significant in respects of both root hair length, root length and root diameter. Both *A. caulinodans* ORS571 and *A. brasilense* FP2 strains with the pOPS0667 circuit produced robust

increases in both root hair length (Figure 5-15) and on the length of the three main tap roots (Figure 5-19). This result proves that the engineered strains not only compete with the endogenous rhizosphere population, but also excrete a significant amount of IAA that stimulates root hair development in the growing plant. This result is significant as other studies have been able to show enhanced IAA production in engineered chassis but at the detriment of the competitiveness of the strain to the point where they do not survive in competitive environments[224,290]. The increase in root hair length reveals that not only is the pOPS0667 circuit producing a measurable effect on plant development, but that the amount of IAA being absorbed by the root cells is in the optimum range required to promote and not repress root hair development (Figure 5-26)[295,296] It is likely that tuning the circuit to produce more IAA will both reduce the fitness of the host cell by reducing flux of protein biosynthesis from tryptophan, as well as cause cell hair growth repression. Therefore we believe this circuit does not require any more major transcriptional or translational tuning.

The results of the root length analysis (5.4.2 above) correlated with those of the root hair length (5.4.1 above). The increase in root length was only observable several days post inoculation due to the compounding promotion effects of the auxin over time. Studies have shown that plant root length promotion is an effect observed later than that of root hair length promotion[297].

.

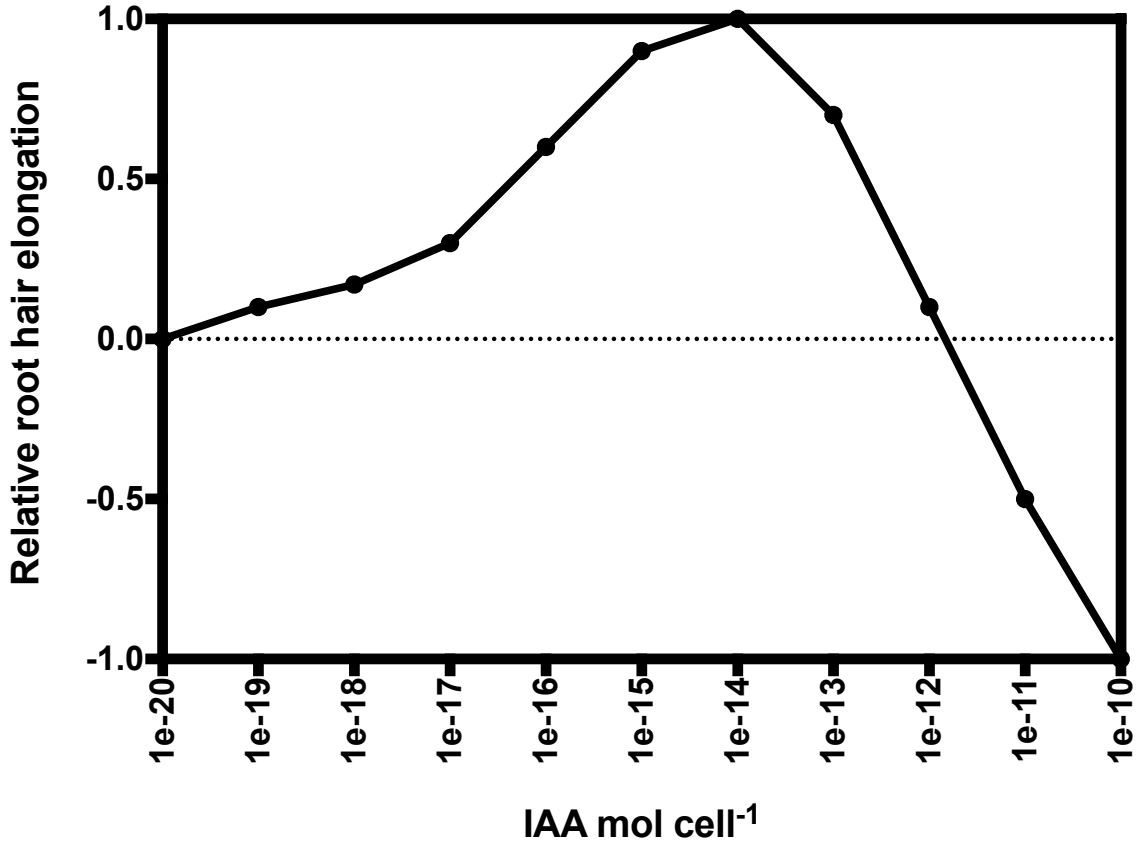


Figure 5-26 -Predicted IAA induced *H. vulgare* root hair elongation model.

Similarly to the rationale behind the root hair length elongation, to provide a growth promotion effect the concentration and therefore amount of IAA present in the rhizosphere must be at an optimum level[298]. It is therefore reasonable to conclude that the population of engineered *A. caulinodans* ORS571 and *A. brasilense* FP2 in the *H. vulgare* growth experiment not only persist for several DPI, but also continue to produce IAA to an optimum concentration in the rhizosphere. Future work to confirm this hypothesis may include rhizosphere

population screening using techniques such as the absolute quantitation of strain abundance using synthetic chimeric DNA spikes and metagenomic analysis[274]. Fluorescently tagging the engineered strains and visualizing their location *in vivo* would allow for the colonisation profile of the strains to be analysed. It is possible that the IAA biosynthesis may cause the rhizobacteria to colonise different areas of the plant root system.

An increase in root length elongation is often correlated with a decrease in root diameter in plant root systems altered with exogenous hormones[271,299]. This inverse relationship was confirmed in our experiments (Figure 5-19). The shift in biosynthetic output of the root elongation zone towards longer cells causes faster root growth (and therefore length) at the cost of the cells in the maturation zone of a developing root being thinner[300]. This phenotype may be either beneficial or detrimental to a *H. vulgare* plant in an agricultural context. If the soil environment has a high moisture content and there are high mechanical forces to the plant, then the longer and thinner root systems promoted by the pOPS0667 circuit in the rhizobacteria chassis would be detrimental to the plant. Conversely if the soil moisture content is low, and the mechanical forces on the plant are also low the promoted phenotype would likely be beneficial to the plant. This is because thinner roots are weaker at absorbing mechanical stress but deeper root systems often allow plants access to the moist subsoil layers below the topsoil[301]. In the future this hypothesis could be tested with the use of mechanical stress methods such as the pullout test[302] and soil moisture profiling apparatus[303].

5.7.2 DAPG Biosynthesis

The results from the constitutive DAPG biosynthesis circuit in *A caulinodans* ORS571 were significant with $9.7x10^{-7}$ µmol $cell^{-1}$ DAPG produced, equating to a maximum concentration of 203 µg ml^{-1} DAPG in the original culture media. This amount of DAPG in the media is 10X greater than the required inhibitory concentration required to prevent zoosporogenesis and reduce spore motility in the pathogenic fungus *Plasmopara viticola*, and a damping-off pathogen, *Aphanomyces cochlioides*[304]*.* Therefore we predict that experiments in the future that test the antifungal ability of the circuit will provide fungal protection as long as the colonization density of the rhizobacteria is around the 10^{4-5} cfu g^{-1} plant root tissue[11,305,306]. It is also unknown whether the DAPG biosynthesis circuit (pOPS0910) causes a significant detrimental effect to the growth of the cell. Previous studies have indicated that overexpression of DAPG within a host cell causes significant decreases in the host cells fitness[307]. Due to time constraints we were unable to perform growth curve analyses on the pOPS0910circuit in the *A. caulinodans* ORS571 strain and whether generation time of the cells is affected. This would be the first experiment to perform preceding *in vivo* DAPG and antifungal protection experiments in a competitive soil environment. The significant difference in the biosynthesis of DAPG between the *E. coli* and *A. caulinodans* ORS571 strains is likely due to differences in the malonyl-CoA substrate availability within the cells as the substrate plays a key role in fatty acid and polyketide biosynthesis[308]. There may also be differences in the efficiencies

of the *phl* gene products in either host strain due to variations in codon usage, tertiary structure, zinc/iron concentrations, *acS* and *gacA* gene expression and the negative feedback of fusaric acid, salicylate and pyoluteorin on DAPG biosynthesis[309] (Figure 5-27).

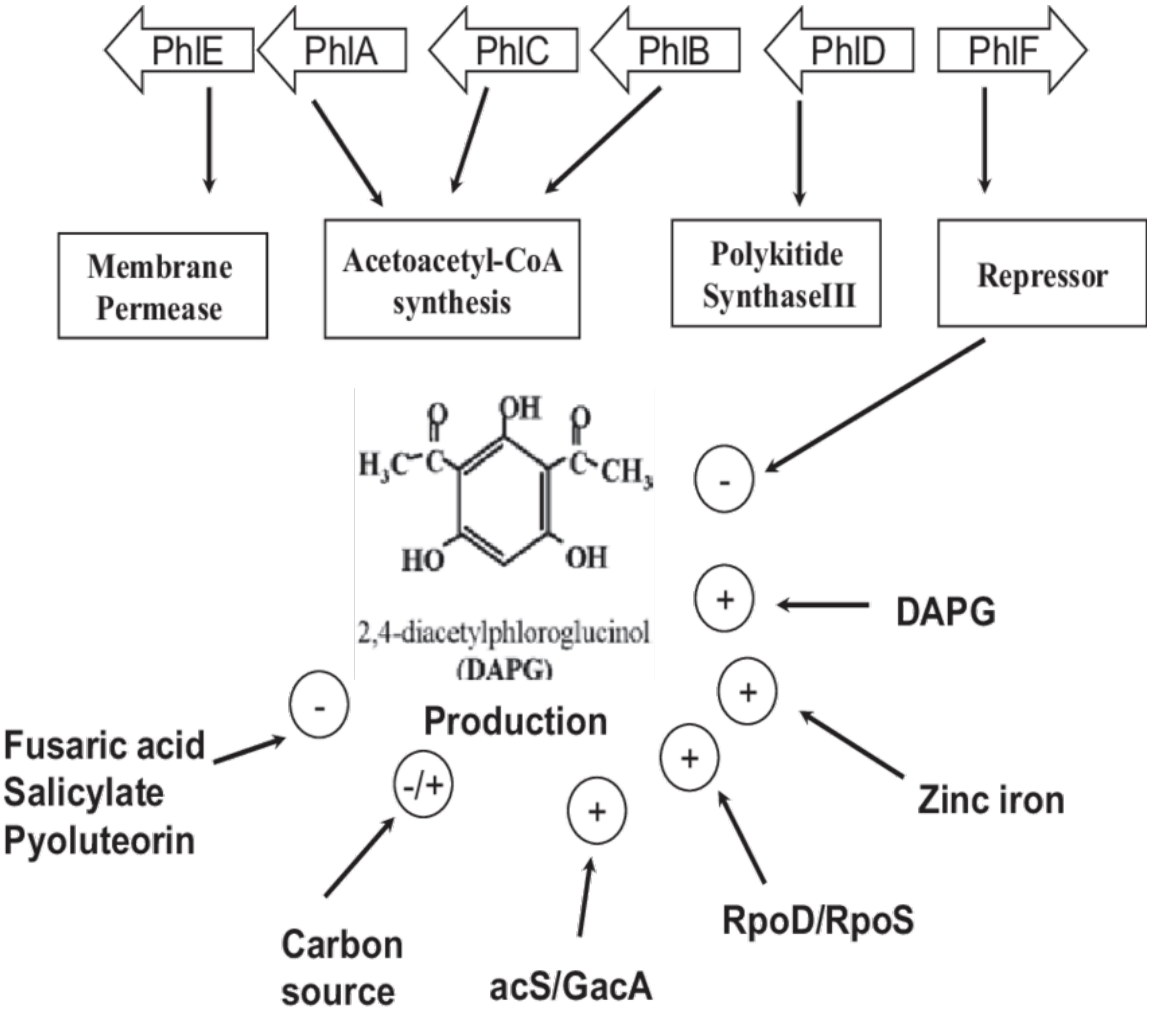


Figure 5-27 – Factors affecting the biosynthesis of active DAPG in rhizobacteria. Taken from Srinivasamurthy *et al*., 2014[309]. Note + and – symbols denote positive and negative feedback on DAPG biosynthesis.

Future gene transcription analysis (qRT-PCR) and enzyme activity analysis can be used to check the gene expression and then subsequent efficiencies of the *phl* enzymes in each rhizobacterial context. Changing the carbon source of the media may also give useful insight into which is the optimum for *in vitro* DAPG expression and purification. A source of biologically synthesized DAPG is useful for the orthogonal induction of many synthetic circuits.

5.7.3 Phosphate Solubilisation

The phosphate solubilisation circuits pOPS1182 and pOPS1183 did not function as well as expected in either of the two rhizobacterial chassis. The circuit did however perform well in the *E. coli* strain. It is likely that the GltA (citrate synthetase) is optimally folded in the *E. coli* strain due to the conserved codon usage and endogenous optimized control of the TCA cycle[310]. The CitC citrate transporter gene is also likely to be more active in the *E. coli* strain than the two rhizobacterial strains due to the gene coming from *Salmonella typhimurium* a closely-related species with similar codon usage to *E. coli* (Figure 5-28).

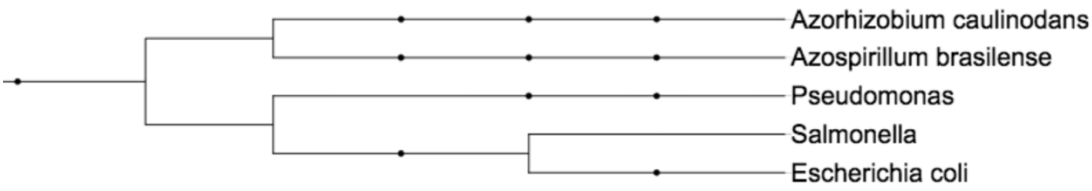


Figure 5-28 – Phylogenetic tree based on the CitC gene DNA sequence from a selection of bacterial species.

The citric acid biosynthesis circuits did not function in the *A. brasilense* FP2 strain and it is likely that refactoring of the circuit, qPCR gene expression and enzyme activity analyses are required to enhance the phosphate solubilisation capacity of the strain.

Previous studies have indicated that the citric acid biosynthesis was entirely dependent on large concentrations of glucose (~100 mM) within the media[243]. The 100 mM glucose added in our modified Pikovskayas agar is not able to be metabolized by *A. caulinodans* ORS571 or *A. brasilense* FP2 and therefore the citric acid biosynthesis of those strains is almost certainly limited by the carbon source availability. Conversely the *E. coli* strain is able to metabolise the glucose carbon-source and is therefore the main limiting factor to citric acid biosynthesis is removed. Future experimentation with the two citric acid biosynthesis circuits in both rhizobacterial strains on a range of aldose and ketose-based carbon

sources may optimise the circuits further. In the rhizosphere the prevalence of such a high concentration of carbon in the form of glucose (or another aldose sugar) is unlikely[222,311,312]. Therefore we predict the efficacy of our phosphate solubilisation circuit to be low and only effective in soil systems with high insoluble phosphate and low phosphate requiring target plant species.

In this chapter the successful design, construction and validation of three PGP mechanisms was achieved in two rhizobacterial strains. Future work will allow the combination of these mechanisms together to further enhance agricultural PGP. By moving the control of these systems under the control of a crop specific unique signaling molecule (rhizopine; 1.6.1.1 above) the PGP traits would be linked to those of the target plant rhizosphere and not any plants present in the system. This inducible system would also allow the engineered rhizobacteria to maintain in a wild-type phenotype state until such time as a plant signal is received, benefiting the system in so much as the strain would be more competitive in the soil environment and would persist longer. The rhizopine inducible PGP work is ongoing and at the time of writing, rhizopine-inducible IAA production by *A. caulinodans* ORS571 by the rhizopine inducible IAA biosynthesis circuit (pOPS1047) *in vivo* was observed (data not shown). A rhizopine inducible DAPG biosynthesis circuit (pOPS0909) was successfully constructed but at the time of writing has yet to be tested in either rhizobacterial chassis.

6 CRISPRi of Rhizobacteria

6.1 Abstract

In this chapter the successful design, construction and implementation of a CRISPR interference (CRISPRi) mechanism is demonstrated in two rhizobacteria species. The dCas9 wt protein as well as two engineered variants (EQR and VRER) with altered PAM sequence-targets are analysed for their on-target repression in both *Azorhizobium caulinodans* ORS571 and *Azospirillum brasilense* FP2. Successful repression of a sYFP2 reporter protein was successfully demonstrated in both strains with between 32-45-fold repression observed with high levels of sgRNA expression. Low level sgRNA expression caused significantly less target gene repression revealing the tunability of the system for tiered gene repression profiles. Construction and verification of three dCas9 plasmids (wt, EQR and VRER) with Golden Gate compatible sgRNA replacement cassettes for rapid sgRNA expression were constructed and verified for use in rhizobacteria. As a proof of concept and function, the CRISPRi system utilizing the dCas9 wt protein was implemented to control the expression of a core ammonia assimilation gene in the rhizobacteria *A. brasilense* FP2. Repressing *glnA* (glutamine synthetase) expression by the CRISPRi system produced a robust ammonia excretion phenotype with a maximum 1.7×10^{-10} mol $cell^{-1}$ ammonia release observed. qPCR and GS activity assays confirm the phenotype to be correlated to the *glnA* transcription level within the cell.

6.2 Introduction

Approximately 40% of all bacteria and over 90% of archaea contain an endogenous defense mechanism against foreign DNA elements such as bacteriophage, plasmids and viruses[313]. This mechanism is known as a Clustered Regularly Interspersed Short Palindromic Repeats (CRISPR) system[314]. The CRISPR system depends on the activity of CRISPR associated-(Cas) proteins and CRISPR non-coding RNA (crRNA) elements to guide the nuclease to target DNA regions[315]. In type II CRISPR systems a single mature crRNA-transacting RNA (tracrRNA) molecule interacts with a single endonuclease protein Cas9 to bind to a specific 20 bp target DNA sequence, therefore inducing a double stranded break. Previous studies have shown that the two RNA molecules required for an active Cas9 enzyme can be fused together into a single guide RNA (sgRNA) molecule and still enable sequence specific DNA targetting[316](Figure 6-1A and Figure 6-2). This development allows for the efficient programming of DNA targeting of the Cas9 endonuclease and the ability to multiplex reactions, targeting multiple locations in the genome simultaneously[317]. To date CRISPR systems have been used to engineer organisms across all kingdoms of life[47].

More recently, the CRISPR system has been modified to perform other novel functions such as DNA sequence localisation[318], self-cloning[319] and most importantly gene expression regulation by CRISPR interference (CRISPRi)

technology[66]. CRISPRi repurposes the native CRISR system from *Streptococcus pyogenes* for robust transcriptional regulation[71].

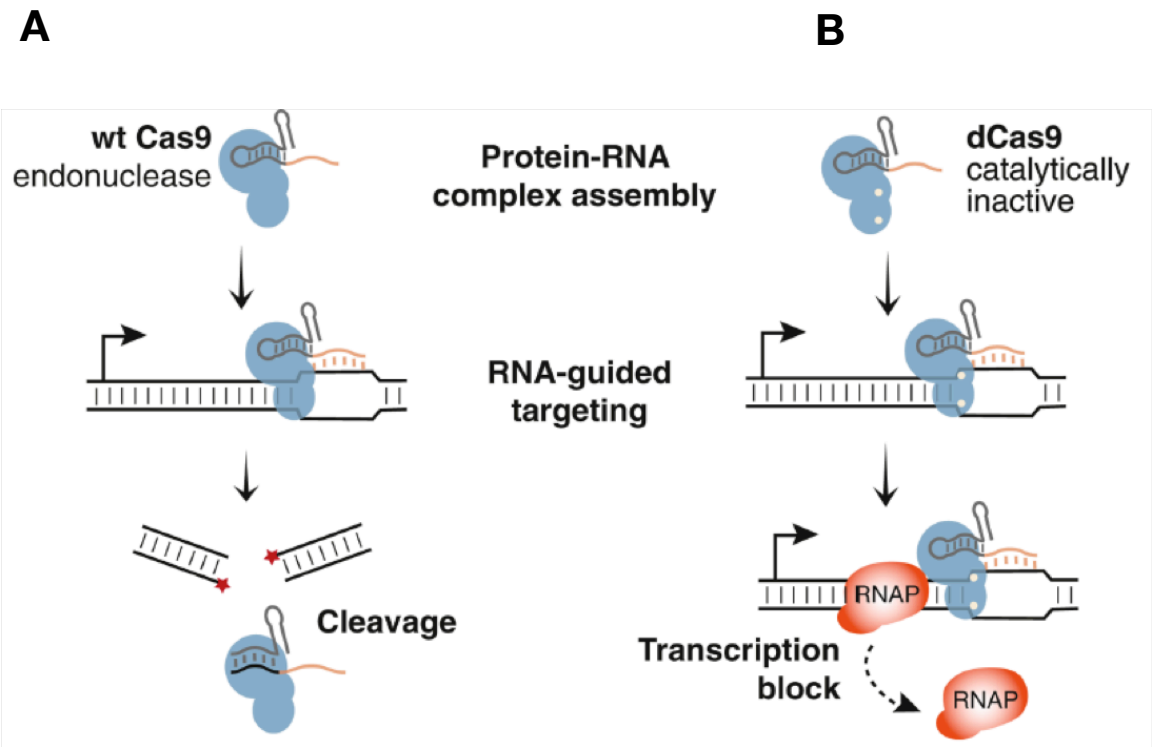


Figure 6-1 – Traditional CRISPR-Cas9 DNA cleavage and alternative CRISPRi-dCas9 mechanisms of action *in vivo*. **A** – Wild type Cas9-mediated cleavage of dsDNA. **B** – Engineered dCas9-mediated transcriptional blockage. Figure adapted from Qi *et al.,* 2013[66].

In the CRISPRi system a modified and catalytically dead version of the Cas9 enzyme (dCas9) is combined with sgRNA molecules to bind the protein to a 20 bp target DNA sequence[320](Figure 6-1B). The dCas9, once bound to the DNA, causes steric interference to the transcriptional process either through the prevention of DNA polymerase binding to the DNA, or preventing nascent RNA synthesis and prematurely terminating transcription[321]. This in effect causes gene

repression much like that of a traditional transcriptional regulator such as TetR or LacI (4.3.1 above) but with the added advantage of customizable DNA binding sites rather than conserved operator sequences. The repression profiles of CRISPRi systems are tunable by altering the specificity of the sgRNA to its target region[322], as well as using synthetic biology DNA parts such as inducible promoters (3.3.4 above) and characterised RBSs (3.3.2 above) to titrate the CRISPRi gene repression and even reverse its effects *in vivo*[323,324].

The widely conserved design of sgRNA molecules is that of a 102 bp long noncoding RNA comprising of a 20 bp target pairing region, a 42 bp dCas9/Cas9 binding region ('handle') and a 40 bp transcription terminator (Figure 6-2). The 20 bp target pairing region of the sgRNA can be readily altered allowing for a flexible gene targeting, the single requisite is that the corresponding 20 bp DNA sequence proceeds a NGG protospacer adjacent motif (PAM) site[70].

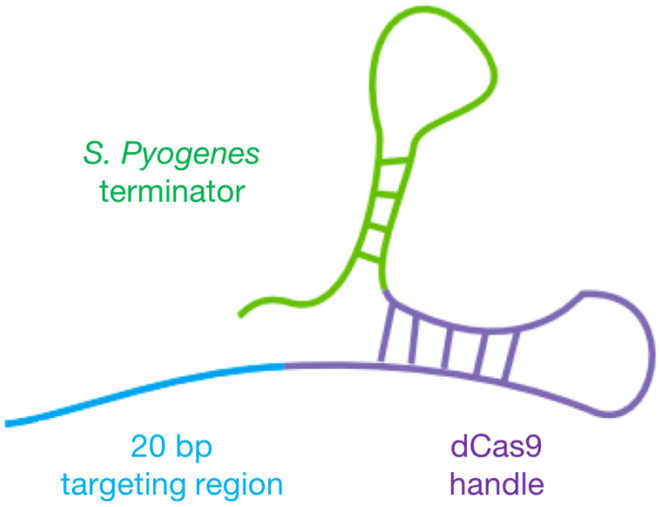


Figure 6-2 sgRNA molecule architecture. Note base pairing to form secondary stem-loop motifs.

The wild type (wt) dCas9 protein from *S. pyogenes* has been used repeatedly for gene repression in many studies[74,320,325,326], however the NGG PAM sequence constrains the number of 20 bp DNA sequences the wt dCas9 protein can effectively target. As a result the number of genes in a organisms genome that can reliably be repressed is reduced significantly. However, previous studies have managed to expand on the single NGG PAM sequence and engineer two variants of dCas9, EQR and VRER, that are able to bind to NGAG and NGCG PAM sequences respectively[327] (Table 6-1). The two variants were constructed by random mutagenesis of the dCas9 PAM interaction domain and then subsequent selection by directed evolution. Both the EQR and VRER variants have genome wide off-target specificities similar to that of the dCas9 wt protein as analysed by

GUIDE-seq analysis[328]. The wt, EQR and VRER dCas9 variants have been used widely in both genome editing and gene repression studies from zebrafish[329] to human cell lines[330].

Table 6-1 – S. pyogenes dCas9 variants, unique amino acid substitutions and PAM sequences.

dCas9 Variant	Amino Acid Substitutions	PAM sequence
Wild-type *S. pyogenes*	-	NGG
S. pyogenes EQR	D1135**E**/R1335**Q**/T1337**R**	NGAG
S. pyogenes VRER	D1135**V**/G1218**R**/R1335**E**/T1337**R**	NGCG

The modulation and engineering of a target organisms gene expression is a key goal of synthetic biology. Applications from genome mapping to metabolic flux and pathway engineering would benefit from the synthetic control of target gene expression/repression. The engineering of endogenous biosynthetic pathways and the expression of heterologous pathways usually require extensive balancing of gene expression stoichiometries to achieve maximum yields, product titer and efficiencies[331,332]. Traditional cloning methods to achieve these results typically take years to finish. The primary objective of metabolic pathway engineering is to produce the maximum amount of a target product whilst minimizing flux out of the pathway and toxic imbalances to the host cell metabolism. For the past 50 years this process has been accomplished by the knockout of endogenous genes responsible for the production of by-products that compete for the target

products precursor compounds[333]. However, traditional knock out approaches such as the phage λ Red recombinase consecutive deletion method are irreversible, time consuming and laborious[334]. The λ Red recombinase method of gene deletion also leaves DNA scarring sequences that can cause polar effects on surrounding gene expression. The CRISPRi system allows for the regulation of endogenous gene expression without the need for genome editing of any kind. There are however draw-backs, one being the effect on the host cell of the expression of the dCas9 protein, as well as the sgRNA. Several studies have shown that expression of dCas9 proteins without co-expression of a sgRNA may cause toxicity[130].

To date the use of CRISPR systems in plant growth promoting rhizobacteria (PGPR) species has been all but nonexistent with only a single study using CRISPR-Cas9 to analyse the interaction of two *Bacillus* spp. with *Brassica rapa* and *Lolium perenne* plants[335]. There have been no studies performed on rhizobacteria using the CRISPRi system to date, although some groups such as the Noirot Lab group at the Argonne National Laboratory have plans to start using CRISPRi to analyse the plant growth promoting *Pseudomonas fluorescens* biofilm formation phenotype.

In this chapter we successfully present the first evidence of the CRISPRi system functioning in two rhizobacteria, *Azorhizobium caulinodans* ORS571 and *Azospirillum brasilense* FP2. The repression of a reporter gene (sYFP2) was

achieved using three aTc inducible dCas9 variants on a separate plasmid to the reporter. The effect of dCas9 variants expression within the host cell is also analysed by real time growth assays. Successful tuning of the sgRNA expression level allowed enhanced reporter gene repression therefore confirming the flexibility of our CRISPRi system. A pipeline for the rapid cloning of sgRNA expression cassettes by Golden Gate is established and their integration into a single vector is achieved, paving the way for future multiplexing capabilities.

After characterisation of the CRISPRi system in the two rhizobacterial chassis, a suite of sgRNA molecules for the repression of a specific gene of interest (*glnA*) are designed and the efficacy of the CRISPRi system at repressing *glnA* transcription under diazotrophic growth conditions is analysed. The effect of the CRISPRi repression of GlnA expression is presented and a significant plant growth promoting phenotype is engineered: ammonia secretion.

6.3 Results

6.3.1 dCas9 Toxicity Screen in Rhizobacteria

To analyse the expression of dCas9 in the two rhizobacteria strains *A. caulinodans* ORS571 and *A. brasilense* FP2, an anhydrous tetracycline (aTc) inducible expression circuit was constructed. Using DNA parts characterised as part of the RhizoKit toolbox (3 above) and the Golden Gate Type IIS restriction enzyme cloning method (2.5.7 above) rapid construction of stable, broad host range dCas9 expression circuits were achieved (Figure 6-3).

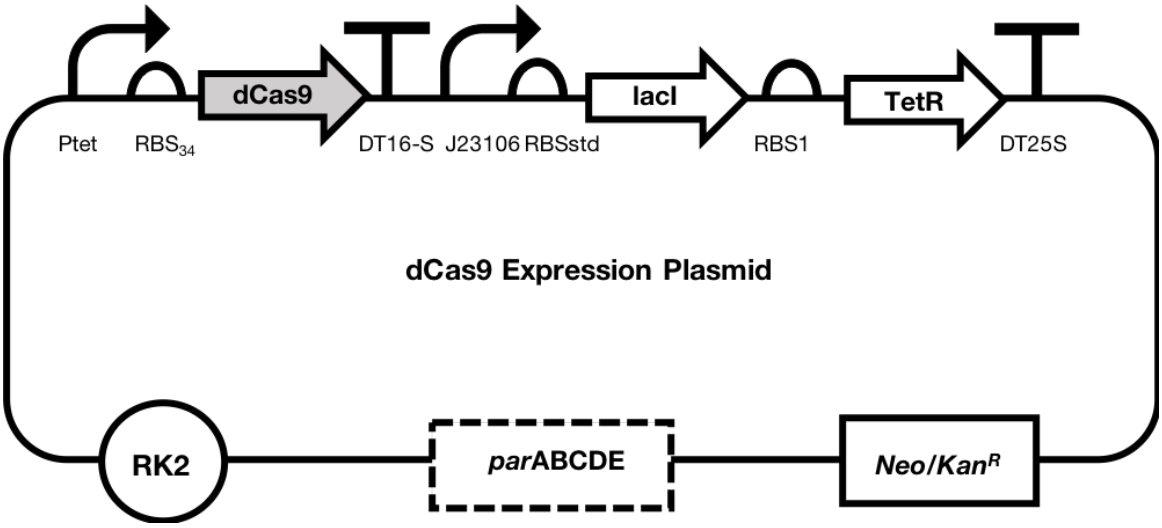


Figure 6-3 – Architecture of the dCas9 wt, EQR and VRER variants expression plasmid (pOPS1150 (wt), pOPS1151 (EQR) and pOPS1152 (VRER)). dCas9 – Open reading frame of dCas9 variants. RK2 – origin of replication. *parABCDE* - plasmid maintenance gene cluster. *Neo/Kan*^{R} – Neomycin and kanamycin selection gene. LacI and TetR – Transcriptional repressors.

Three variants of the dCas9 protein, the wild type, EQR and VRER engineered versions were kindly donated to us by Alec Nielsen of the Voigt lab, MIT. Unfortunately, the ORFs contained many BsaI and BpiI sites and were too large to codon optimise and synthesise using traditional synthesis platforms. Therefore, the dCas9 protein variants were cloned into the plasmid pOGG096 by recombinase-based BD reaction (2.5.6 above). Each of the aTc inducible dCas9 variants expression cassettes were amplified by PCR reaction using the primers oxp1920 and oxp1921 and the plasmids pAN4137, pAN4183 and pAN4181 as templates. The amplicons were purified (2.5.2 above) and then ligated into previously BsaI-digested pOGG096 vectors (2.5.5 above) to make plasmids pOPS1150-52 (Figure 6-4).

All three dCas9 variant expression circuits were transformed into *E. coli* by heat shock (2.6.1 above) for stocking at -80˚C. The plasmids were also mobilised into the two rhizobacterial chassis *A. caulinodans* ORS571 and *A. brasilense* FP2 by conjugation (2.6.3 above). To analyse the toxicity of the dCas9 protein expression in each of the rhizobacteria, a high throughput growth assay was used (2.7.1 above). Only the dCas9 wild type variant was used for toxicity analysis due to time constraints. The empty pOGG096 vector in each of the strains was used as a negative control to observe dCas9 specific toxicity. Strains were grown to an OD595 of 0.6 before being sub-cultured into 500 µL of fresh UMS supplemented with vitamins, 20 mM Succinate, 20 mM NH_3Cl, 30 µM nicotinic acid (for *A.*

caulinodans ORS571 only) and relevant antibiotics (2.4.2 above) at an initial density of $1x10^5$ cells per microwell.

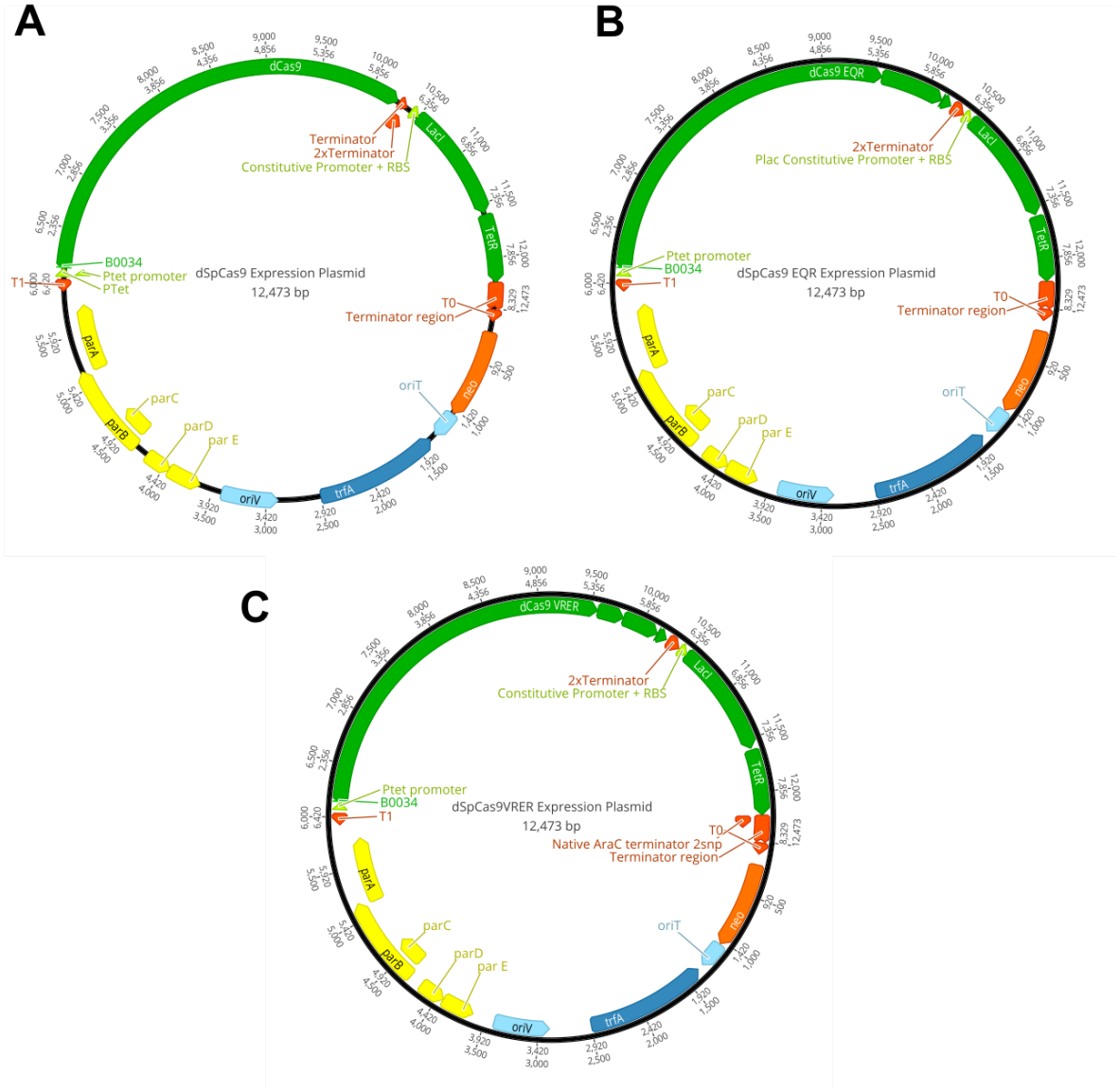


Figure 6-4 – Plasmid maps of the three dCas9 expression plasmids for use in rhizobacteria species. **A** – dCas9 wt variant pOPS1150. **B** – dCas9 EQR variant pOPS1151. **C** – dCas9 VRER variant pOPS1152. Note the RK2 origin of replication and kanamycin/neomycin selection gene and *par* operon for enhanced plasmid maintenance in culture.

To induce the dCas9 variant expression, either a normal (low) concentration (1 μM aTc) or high concentration (10 μM aTc) was added to each microwell at the start-point of the assay. An equal volume of UMS was added to the non-induced cultures. The cultures were grown at 37℃ for *A. caulinodans* ORS571 and 30℃ for *A. brasilense* FP2 with 200 rpm shaking and OD595 measurements taken every 20 minutes. Each strain and condition value presented is the mean product of three biological and three technical replicates.

The results from the dCas9 toxicity analysis in *A. caulinodans* ORS571 indicated that dCas9 wild type protein expression is detrimental to the growth cycle of the host rhizobacteria when no sgRNA is expressed (Figure 6-5A). There was a significant ($p<0.05$) increase in the lag time between the low induction level (12.5 h) and the high induction level (21.6h) compared to the uninduced condition (7.5 h). However, once the cultures were in log phase there was no significant difference ($p>0.05$) in the generation times (8.9 h), and no significant difference ($p>0.05$) in the final OD595 reached in stationary phase (0.750 abs) for the cultures induced with 1 μM aTc. The cultures induced with the higher 10 μM aTc concentration had a significantly lower ($p<0.05$) final OD595 than the normal 1 μM aTc and uninduced conditions indicating over-expression causes toxicity to the strain.

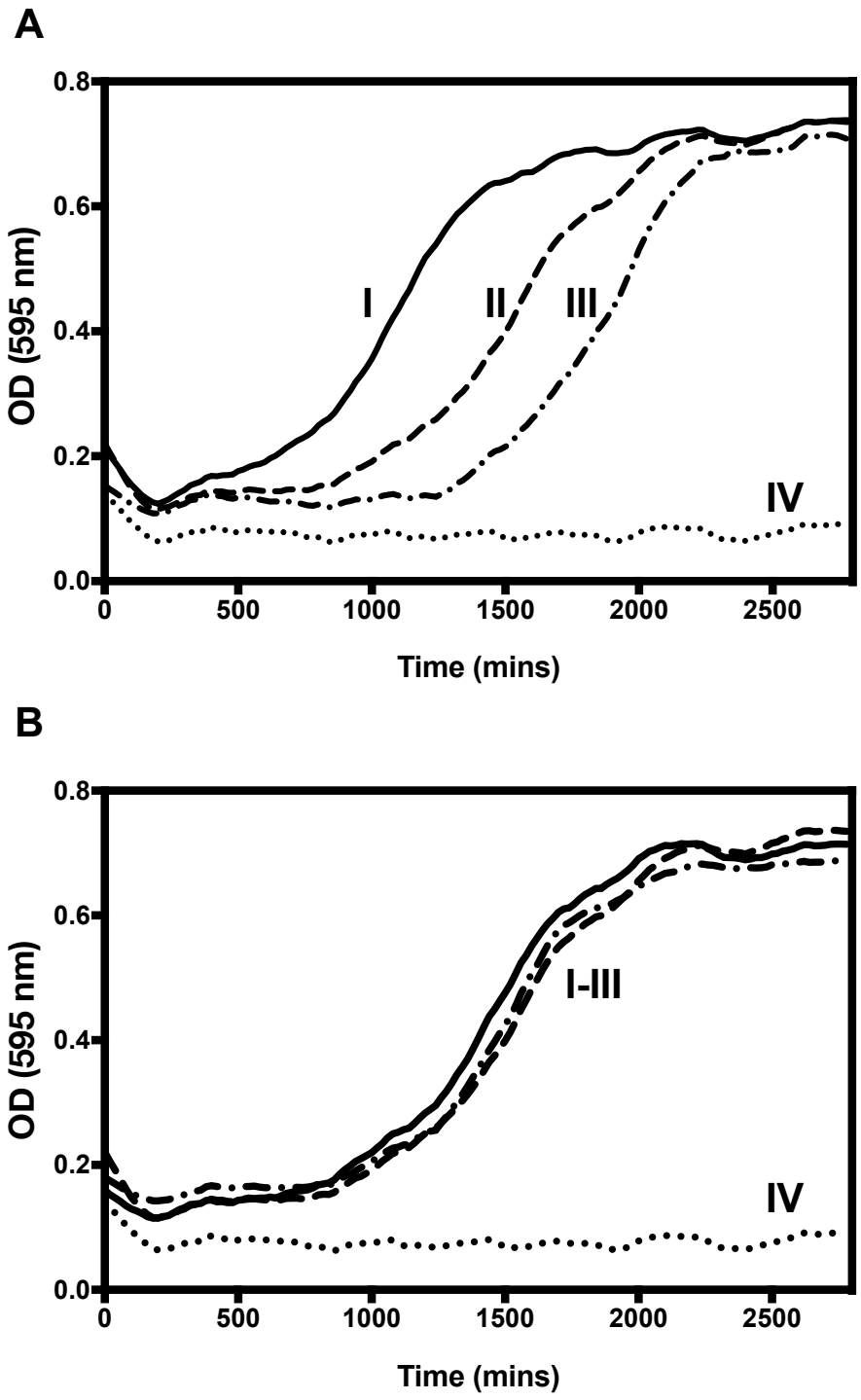


Figure 6-5 - Effect of dCas9 wt expression on the growth of *A. caulinodans* ORS571 in liquid media. **A** –pOPS1150 (aTc inducible dCas9 wt). **B** - pOGG096 (empty vector control). **I** – 0 µM aTc induction. **II** – 1 µM aTc induction. **III** – 10 µM aTc induction. **IV** – blank media control.

The negative control (pOGG096 vector) showed no significant difference ($p>0.05$) in the lag phase time (10.2 h), generation time (8.8 h), or final OD595 reached in stationary phase (0.765 abs) at either of the low or high aTc induction levels (Figure 6-5B). Therefore there was also no significant difference ($p>0.05$) in the generation times or the final culture OD595 observed between the dCas9 expression circuit strain and the pOGG096 negative control strain (Figure 6-5).

Expression of the dCas9 wild type protein in *A. brasilense* FP2 indicated less of a toxicity effect when induced at a low (normal) 1 μM aTc concentration, and a more noticeable drop in cell growth when induced at a high 10 μM aTc concentration (Figure 6-6). The results indicate a distinct biphasic growth profile for all the strains tested with two distinct generation time stages, the initial phase slower than the late log phase. In direct contrast to the *A. caulinodans* ORS571 results, the induced *A. brasilense* FP2 dCas9 expression strains showed no significant difference ($p>0.05$) in the length of the lag phase of growth (8.3 h). There was no significant difference observed ($p>0.05$) in the generation times of the uninduced (0 μM aTc) and low-level induction (1 μM aTc) dCas9 expression conditions at either the slower phase (4.8 h) or faster late log phase (2.6 h). However, there was a significant difference ($p<0.05$) observed in the early log phase generation time (7.8 h) between the high-level (10 μM aTc) induction and the lower and uninduced conditions (4.8 h; Figure 6-6A).

At the final time point observed in the growth assay performed there was a significant difference observed ($p<0.05$) in the OD595 between the uninduced/low induction and high level induction conditions (Figure 6-6A). The final OD595 reached for the uninduced/low level induction condition was 0.590 abs where the high-level induction condition was 0.485 abs. The OD595 of all the strains was still increasing at the end point of the assay and so the final maximum absorbance cannot be deduced under the experimental conditions chosen.

The negative control (pOGG096 vector) showed no significant difference ($p>0.05$) in the lag phase time (8.4 h), early log phase generation time (4.6 h), late log phase generation time (2.8 h), or final OD595 reached in stationary phase (0.585 abs) at either of the low or high aTc induction or uninduced conditions (Figure 6-6B). Therefore, there was also no significant difference ($p>0.05$) in the early and late log phase generation times or the final culture OD595 observed between the dCas9 expression circuit strain induced at low-level and the pOGG096 negative control strain (Figure 6-6).

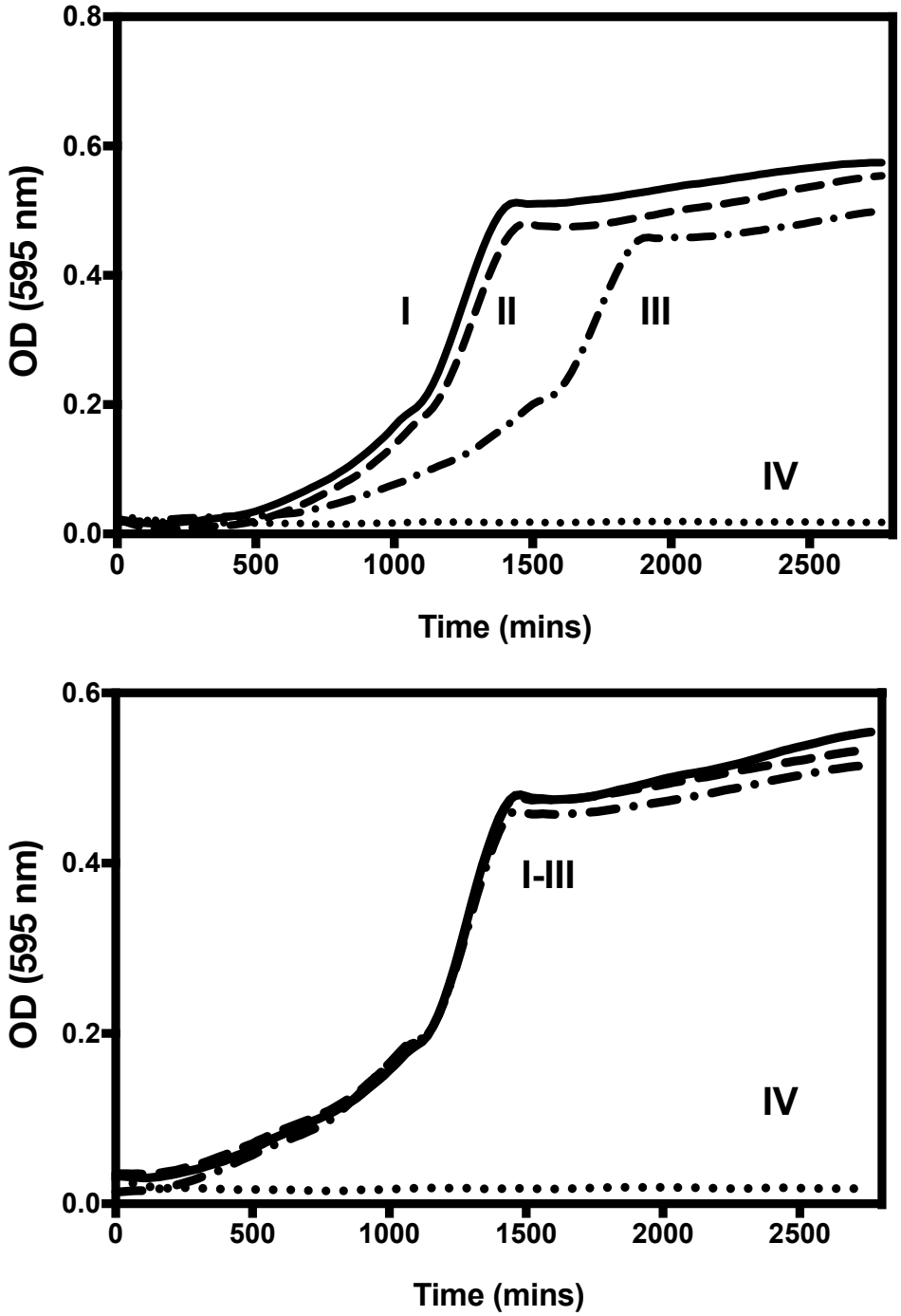


Figure 6-6 - Effect of dCas9 wt expression on the growth of *A. brasilense* FP2 in liquid media. **A** –pOPS1150 (aTc inducible dCas9 wt). **B** - pOGG096 (empty vector). **I** – 0 µM aTc induction. **II** – 1 µM aTc induction. **III** – 10 µM aTc induction. **IV** – blank media control.

6.3.2 dCas9 Activity *in vivo*

To analyse the efficacy of the dCas9 variants transcriptional interference *in vivo* a suite of 'reporter' plasmids (pOPS1153-1158) were constructed (Figure 6-7). The pBBR1 based replication of origin means the plasmid is compatible with the RK2 origin of the dCas9 expression plasmid (Figure 6-3). The reporter gene chosen was sYFP2 due to its robust and bright response observed in both *A. caulinodans* ORS571 and *A. brasilense* FP2 (3.3.7 above). The reporter was expressed from a J23101 promoter, the strongest of the characterised constitutive sigma70 Anderson promoters[182]. This second plasmid also contains the constitutive sgRNA expression cassette required for an active dCas9-sgRNA complex to form. Two promoter variants of each of the reporter plasmids were constructed driving the sgRNA transcription. The first weak sgRNA expression group (pOPS1153-55) were expressed from the Pphl promoter, previously characterised in the RhizoKit DNA toolbox as a relatively weak promoter in both rhizobacteria (Figure 3-18). In the second group of reporter plasmids (pOPS1156-58) the sgRNAs were expressed from the strong J23104 constitutive promoter (Figure 3-4). Therefore the final two-plasmid system consisted of an aTc inducible plasmid (pOPS1150-52; 6.3.1 above), as well as one of the reporter plasmids (pOPS1153-55 for weak sgRNA expression or pOPS1156-58 for strong sgRNA expression).

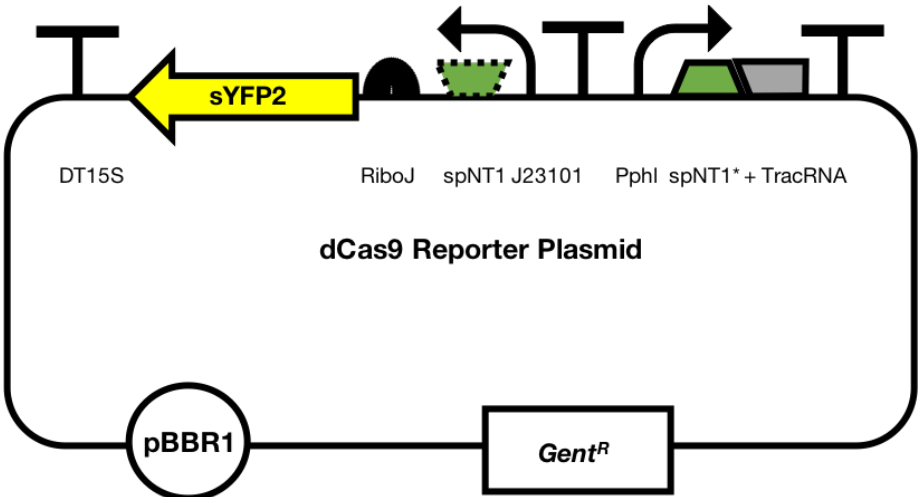


Figure 6-7 - Architecture of the dCas9 wt, EQR and VRER variants reporter test plasmids (pOPS1153 (wt), pOPS1154 (EQR) and pOPS1155 (VRER)). sYFP2 – Open reading frame of sYFP2 reporter gene. spNT1 – 20 bp target of sgRNA. spNT1* + TracRNA – spNT1 targetting sgRNA. RiboJ – translational insulator module. pBBR1 – origin of replication. *Gent*^{R} – gentamycin selection gene.

To construct the reporter plasmids for use in rhizobacteria, the primers oxp1922 and oxp1923 were used to PCR amplify the sYFP2 and sgRNA expression cassettes from the template plasmids pAN4244/pAN4244S (Wild type), pAN4249/pAN4249S (EQR) and pAN4245/pAN4245S (VRER) plasmids again kindly donated by Alec Nielsen. Plasmids denoted with an 'S' indicate the stronger sgRNA expression cassette.

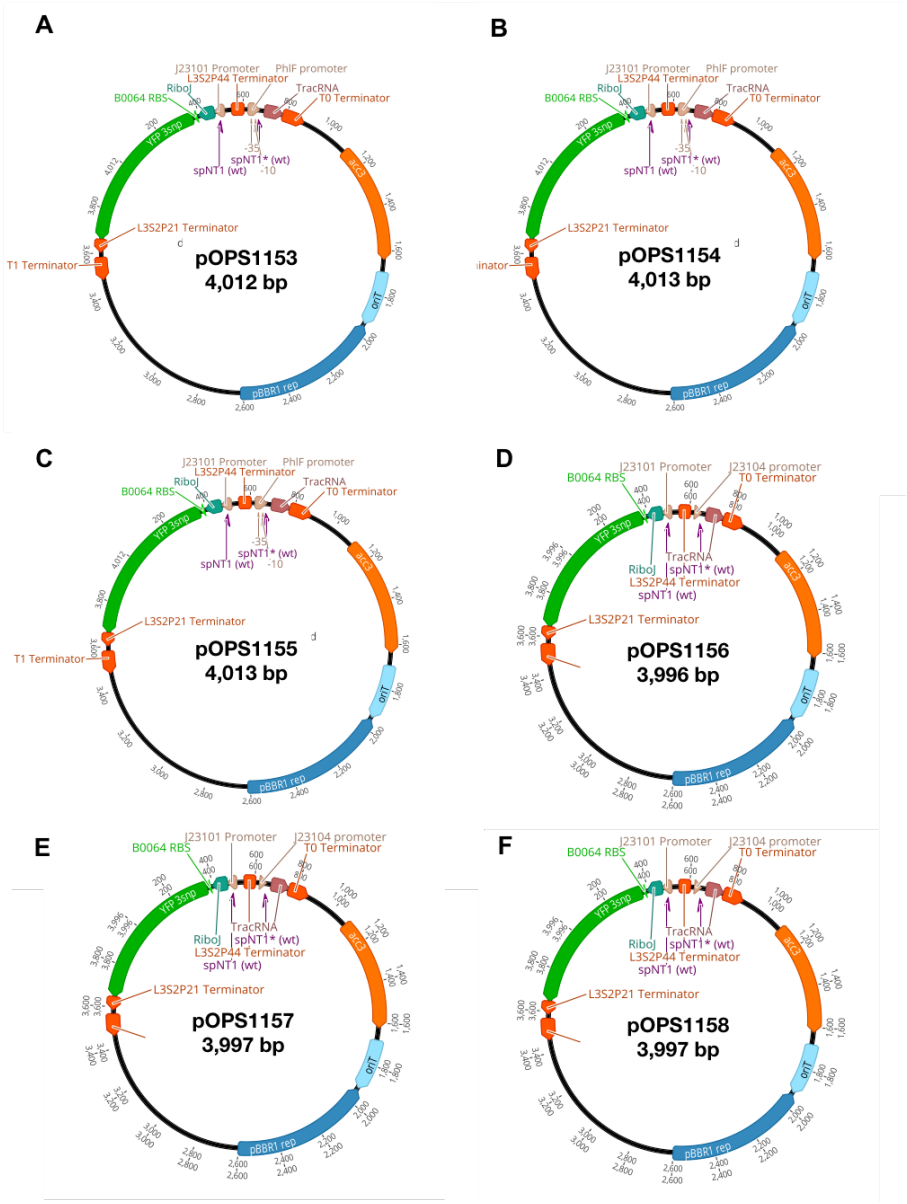


Figure 6-8 - Plasmid maps of the dCas9 reporter plasmids for analysing dCas9 variants activity in rhizobacteria. **A** – dCas9 wt PAM (NGG) variant with weak sgRNA expression pOPS1153. **B** – dCas9 EQR PAM (NGAG) variant with weak sgRNA expression pOPS1154. **C** – dCas9 VRER PAM (NGCG) variant with weak sgRNA expression pOPS1155. **D** – dCas9 wt PAM (NGG) variant with strong sgRNA expression pOPS1156. **E** – dCas9 EQR PAM (NGAG) variant with strong sgRNA expression pOPS1157. **F** – dCas9 VRER PAM (NGCG) variant with strong sgRNA expression pOPS1158. Note the pBBR1 origin of replication and gentamycin selection gene.

The amplicons were purified (2.5.2 above) and then cloned by recombination-based BD reaction (2.5.6 above) into BsaI-digested pOGG024 vectors (2.5.5 above) to make plasmids pOPS1153-58 (Figure 6-8).

All the dCas9 variant reporter plasmids were transformed into *E. coli* by heat shock (2.6.1 above) for stocking at -80˚C. The reporter plasmids were also mobilised into the two rhizobacterial strains *A. caulinodans* ORS571 and *A. brasilense* FP2 by conjugation (2.6.3 above). The conjugation was performed into the strains already harboring the corresponding dCas9 variant expression plasmid, yielding the final two plasmid full test strains. To analyse the transcriptional interference/repression of the dCas9 proteins in each of the rhizobacteria, a high throughput growth (2.7.1 above) and fluorescent reporter assay (2.7.3 above) was used. Strains were grown to an OD595 of 0.6 before being sub-cultured into 500 µL of fresh UMS supplemented with vitamins, 20 mM Succinate, 20 mM NH_3Cl, 30 µM nicotinic acid (for *A. caulinodans* ORS571 only) and relevant antibiotics (2.4.2 above) at an initial density of $1x10^5$ cells per microwell.

To induce the dCas9 variant expression, and therefore initiate sYFP2 repression, 1 µM aTc was added to each microwell at the start-point of the assay. An equal volume of UMS was added to the non-induced cultures. The cultures were grown at 37˚C for *A. caulinodans* ORS571 and 30˚C for *A. brasilense* FP2 with 200 rpm shaking. OD595 and sYFP2 fluorescence were analysed throughout the growth

of the cultures however the repression analysis based valued on cultures in mid log phase at an OD595 of 0.5 Each strain and condition value presented is the mean product of three biological and three technical replicates.

The repression profiles for all of the dCas9 variants with a low level of sgRNA expression showed very low levels of sYFP2 repression in both *A. brasilense* FP2 and *A. caulinodans* ORS571 (Figure 6-9A). The maximum fold-repression was observed at 3.95X in the *A. caulinodans* ORS571 strain with the dCas9 wt and EQR variants. There was significantly lower ($p<0.05$) repression observed in the same strain with the VRER variant at 1.9X repression. The dCas9 repression profiles observed in the *A. brasilense* FP2 strain were significantly lower ($p<0.05$) than those observed in *A. caulinodans* ORS571. There was no significant difference between any of the variants repression with all three causing 1.6X repression of the sYFP2 gene.

When the sgRNA molecule was transcribed to a high level, there was a robust and significant increase in the fold-repression of all the dCas9 variants in both rhizobacteria strains (Figure 6-9B). In *A. caulinodans* ORS571 the greatest repression was observed with the wt dCas9 variant causing 40.1X repression, the EQR 39.5X and finally the VRER variant causing significantly less ($p<0.05$) than the other two variants at 35.7X repression.

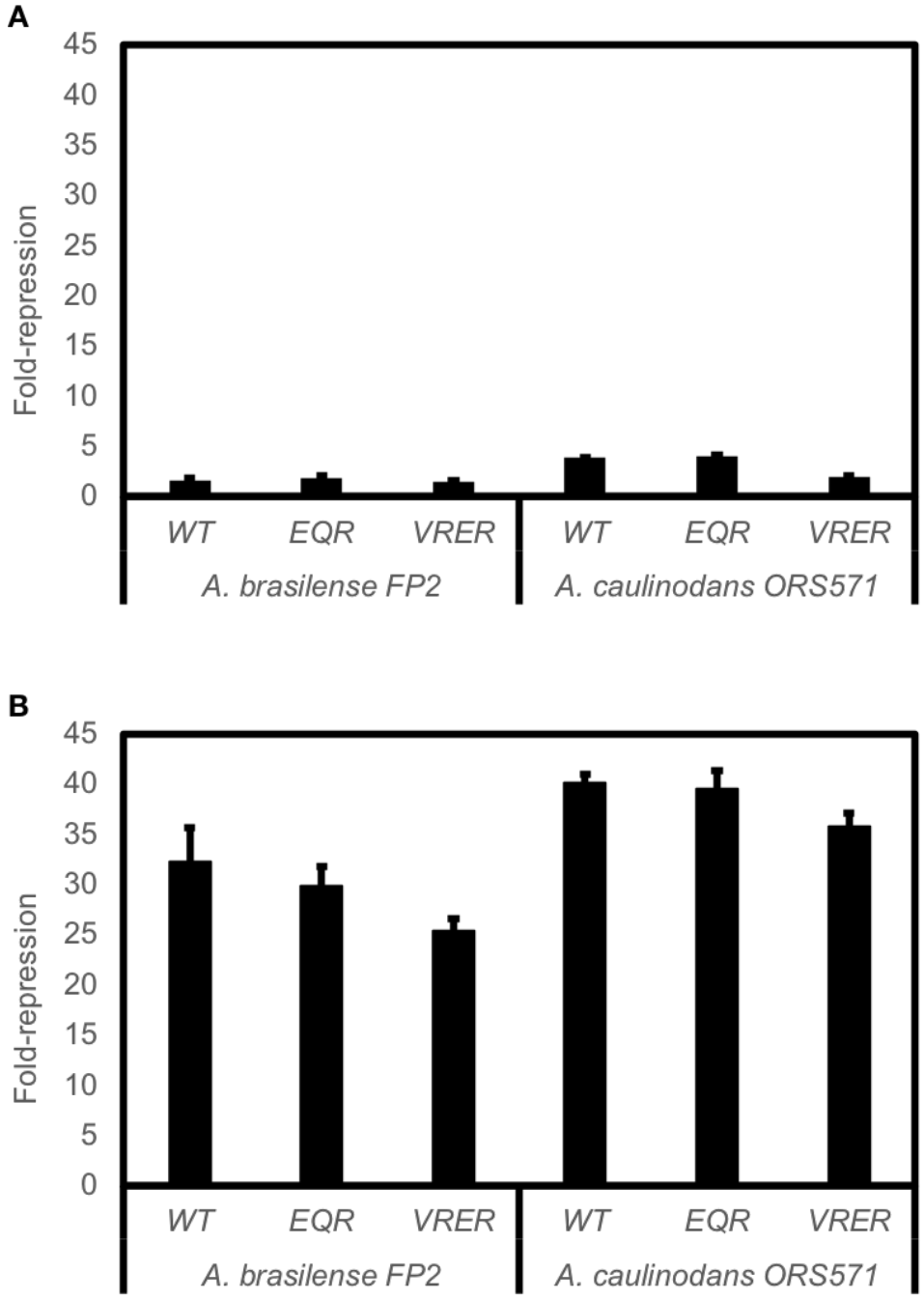


Figure 6-9 – dCas9 variant repression profiles in both *A. caulinodans* ORS571 and *A. brasilense* FP2. **A** – Reporter gene repression with weak sgRNA expression (Pphl promoter). **B** – Reporter gene repression with strong sgRNA expression (J23104 promoter). N = three biological and three technical replicates. Error bars indicate the standard error of the mean.

The repression profiles in *A. brasilense* FP2 were again significantly lower ($p<0.05$) than those observed in *A. caulinodans* ORS571 (Figure 6-9B). There was a clear step-wise decrease in the fold repression across the three variants with the WT dCas9 causing the greatest at 32.3X, the EQR variant next at 29.8X and finally the VRER variant causing 25.4X sYFP2 repression. In conclusion there is a significant correlation between the sgRNA transcription level and the dCas9 variants repression with more sgRNA causing more target gene repression.

6.3.3 Golden Gate CRISPRi Strategy

In order to be able to rapidly target genes for CRISPRi, a Golden Gate compatible sgRNA expression plasmid was constructed. A BD reaction was performed using amplicons of the sgRNA expression plasmid previously characterised above and a LacZ selection cassette for X-gal selection. By using the pOPS1153 sgRNA expression plasmid as a template, the primers oxp2136 and oxp2137 were used to PCR amplify the entire plasmid minus the promoter and 20 bp targeting portion of the sgRNA expression cassette. The second fragment, a LacZ selection cassette flanked by BsaI sites, was PCR amplified from the template plasmid pOGG024 using primers oxp2134 and oxp2135.The oxp2136 and 2137 primers have 15bp BD overhands compatible with the 5' and 3' ends of the LacZ cassette PCR amplicon allowing a BD recombinase reaction to be performed (2.5.6 above). The subsequent plasmid allows for the integration of a promoter and 20 bp targeting portion of the sgRNA into the TracRNA portion of the molecule with no

DNA scar formed (Figure 6-10). Sense and antisense oligonucleotides of approximately 100bp (including BsaI overhangs) can be annealed together to form the Golden Gate-compatible dsDNA required for integration into the pOGG230 plasmid by a level 1 Golden Gate cloning reaction (2.5.7 above).

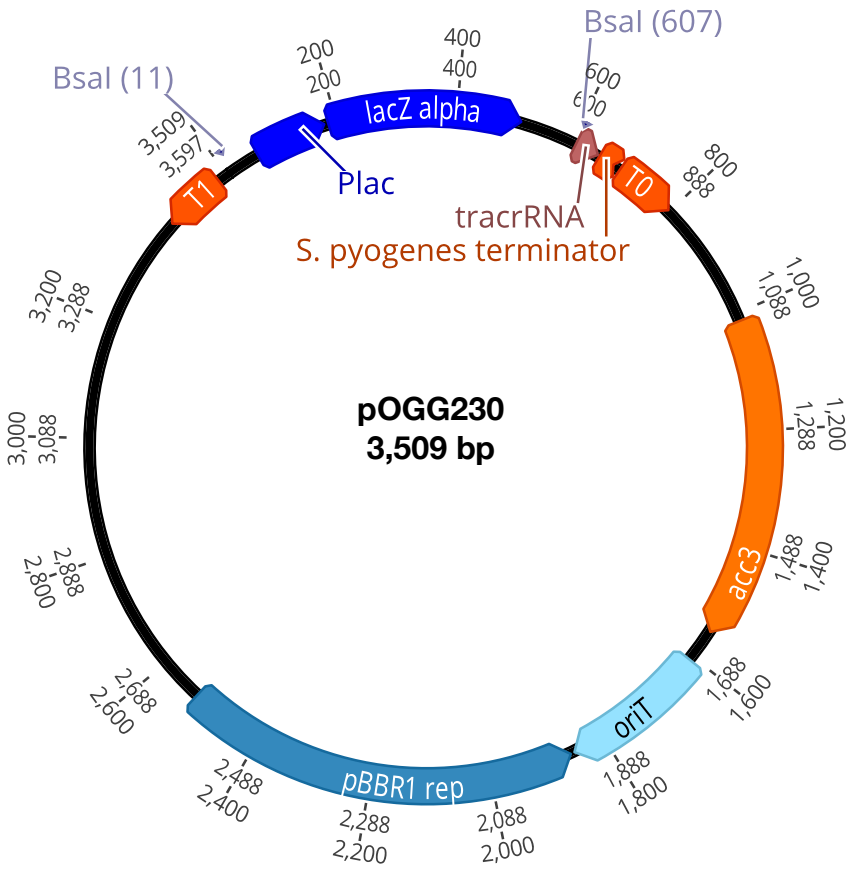


Figure 6-10 – Plasmid map of the sgRNA expression plasmid (pOGG230) with Golden Gate compatible LacZ selection cassette for BsaI-mediated integration of promoter and 20 bp targeting portion of sgRNA molecule. pBBR1 based origin of replication, *acc3* gentamycin selection gene and oriT origin of transfer for rhizobacterial conjugation.

To facilitate rapid sgRNA integration and expression with a dCas9 protein, without a double transformation, a single plasmid with a Golden Gate-compatible sgRNA integration cassette with LacZ selection marker was constructed (Figure 6-11). Three versions of these plasmids were created with each of the dCas9 variants to allow for genome-wide targeting with the ability to use all three PAM sites available (NGG, NGAG and NGCG). To integrate a Golden Gate compatible LacZ and TracrRNA cassette into the already constructed pOPS1150-52 plasmids, a PasI restriction enzyme digest and subsequent BD reaction were performed (2.5.4 aboveand 2.5.5 above). Firstly the three dCas9 variant expression plasmids were linearised by PasI restriction digest overnight and then purified. The entire LacZ and TracRNA cassette was PCR amplified using plasmid pOGG230 as a template using primers oxp3110 and oxp3111. These primers have 15bp BD overhangs compatible with the 5' and 3' ends of the PasI linearised dCas9 variants expression plasmids. Finally, the linearised dCas9 expression plasmids and LacZ tracRNA cassettes were combined by a BD recombinase reaction (2.5.6 above). The resulting plasmids pOGG231-233 are single aTc inducible dCas9 wt (pOGG231), EQR (pOGG232) and VRER (pOGG233) expression plasmids with a LacZ sgRNA Golden Gate cloning-compatible cassette (Figure 6-11). Due to time constraints in the cloning of these plasmids, their function has yet to be verified in ether rhizobacterial chassis. However DNA sequencing has confirmed its identity. Ongoing experiments in the Poole lab group will validate this plasmid for future use. All CRISPRi experiments performed in the following section consisted of dual dCas9 and sgRNA expression plasmids.

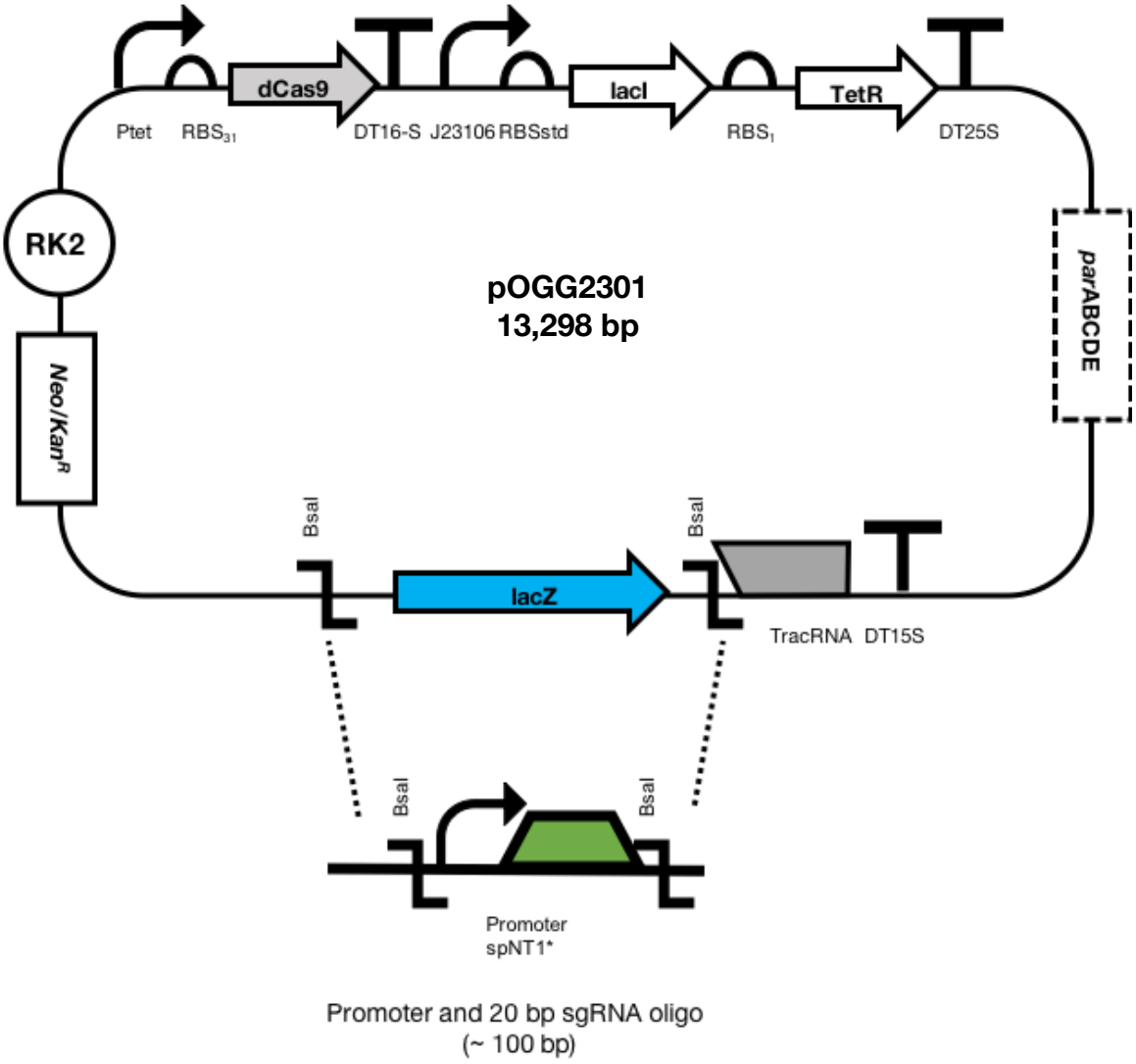


Figure 6-11 – Mechanism for the Golden Gate assembly of sgRNA molecules into the stable dCas9 variants expression plasmid. Example illustrated here is one of three dCas9 variant plasmids (pOGG231-233). A dsDNA molecule with promoter and 20bp target sequence is integrated into the plasmid by level one Golden Gate reaction. LacZ selection allows the selection of positive (white) colonies when plated on X-Gal.

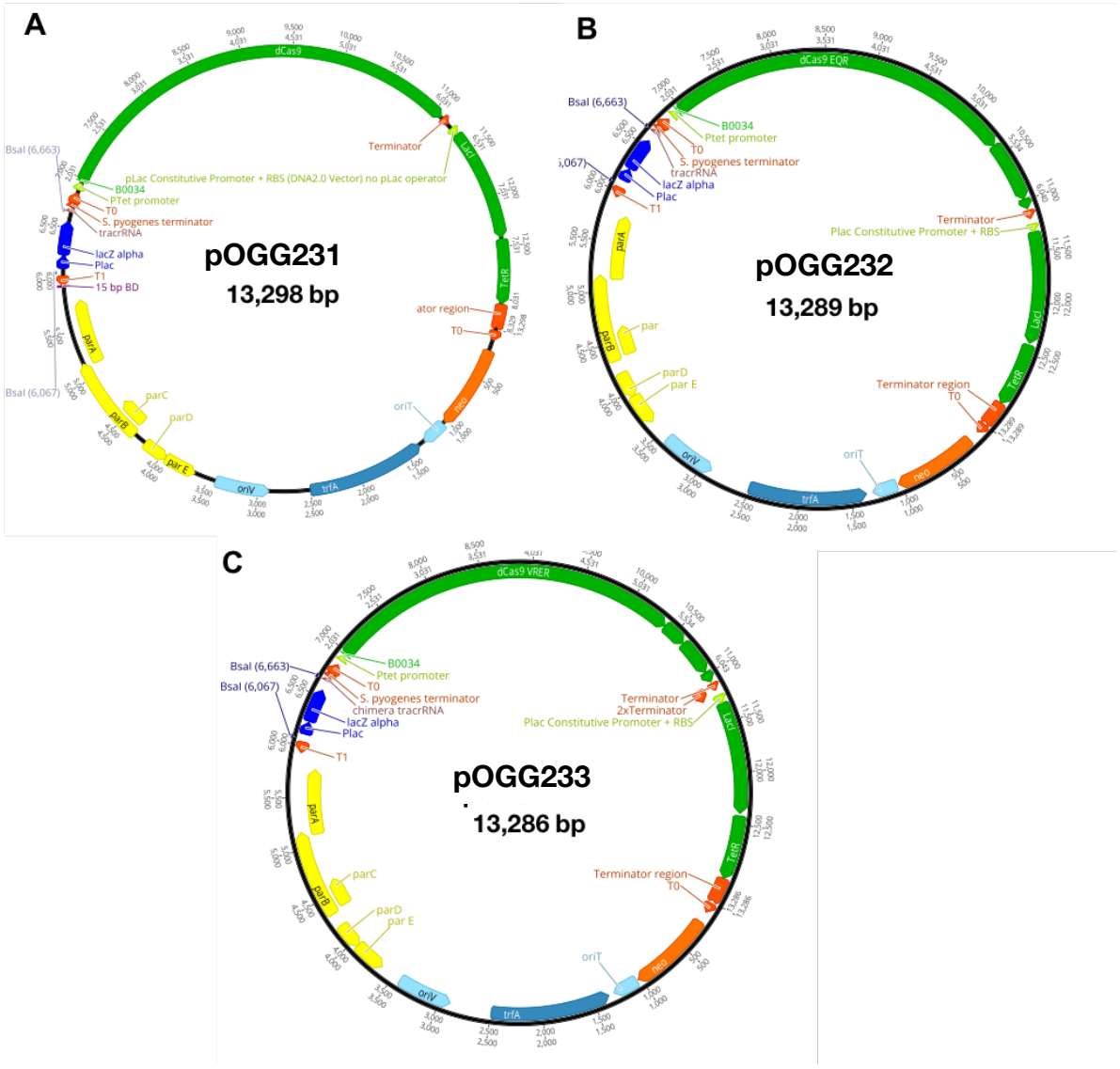


Figure 6-12 - Architecture of the Golden Gate dCas9 wt, EQR and VRER and sgRNA expression plasmids with integrated sgRNA Golden Gate cassette . **A** – dCas9 wt variant pOGG231 (wt NGG PAM). **B** – dCas9 EQR variant pOGG232 (NGAG PAM). **C** – dCas9 VRER variant pOGG234 (NGCG PAM). dCas9 – Open reading frame of dCas9 variants. RK2 – origin of replication. *parABCDE* - plasmid maintenance gene cluster. *Neo*/*Kan*^{R} – Neomycin and kanamycin selection gene. LacI and TetR – Transcriptional repressors. LacZ – Beta galactosidase selection marker gene.

6.3.4 CRISPRi of endogenous rhizobacterial gene

With the dCas9 CRISPRi system functioning in both *A. caulinodans* ORS571 and *A. brasilense* FP2, an endogenous gene repression experiment was performed. Due to time constraints only the rhizobacteria *A. brasilense* FP2 was used in the below study. The chosen target gene for repression was that of Glutamine Synthetase (*glnA*) and is a gene of significant interest in the field of plant growth promoting rhizobacteria (PGPR). Several studies have focused on the GlnA protein in *A. brasilense* FP2 as it is a key metabolic bottleneck in ammonia assimilation[336]. GS works in tandem with the protein glutamine oxoglutarate aminotransferase (GOGAT) in a cycle generating cellular glutamate and glutamine, the two major precursors for purine and pyrimidine as well as other amino acid biosynthesis pathways (Figure 6-13). Under diazotrophic growth conditions, atmospheric nitrogen is fixed into ammonium ions within the *A. brasilense* FP2 cytoplasm, the ammonium is then quickly assimilated into glutamine via the GS-GOGAT cycle.

To date no study has been able to knock out the *glnA* gene from the strain without a toxic phenotype and full loss of cell viability. This is probably due to the lack of glutamine and glutamate biosynthesis occurring even with the supplementation of extracellular glutamine/glutamate in growth media. We attempted several pK19

and pJQ200SK mutagenesis experiments targeting the *glnA* gene with no successful knock outs produced.

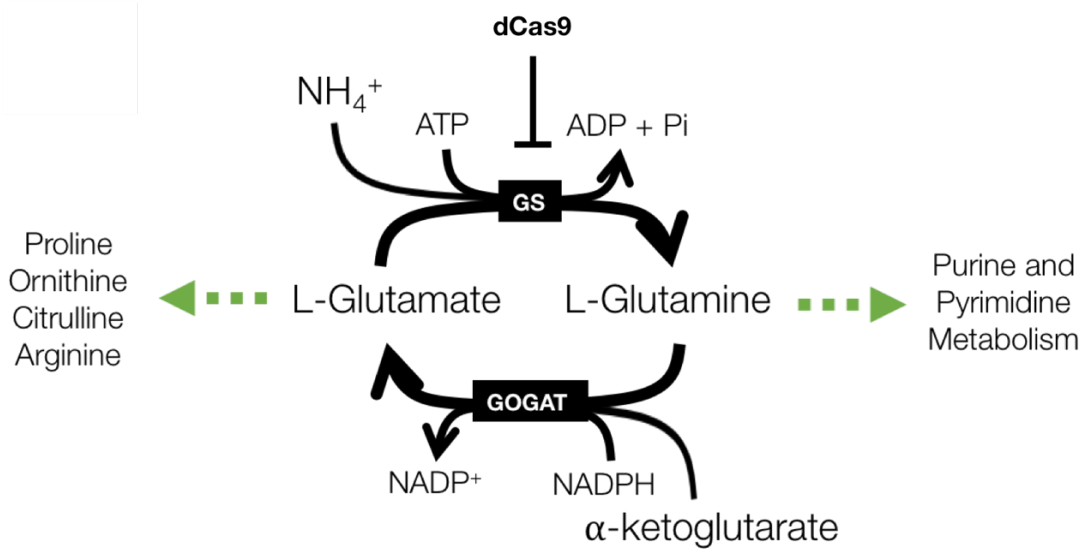


Figure 6-13 - CRISPRi repression of the GS-GOGAT cycle in *A. brasilense* FP2. GS – Glutamine Synthetase. GOGAT - Glutamine oxoglutarate aminotransferase. dCas9 – Represents hypothetical CRISPRi of GS gene (*glnA*) transcription.

Various studies have been able to partially mutagenise the *glnA* gene within the genome of the strain causing noticeable effects such as constitutive nitrogen fixation, and most importantly, ammonia secretion from the cell[43,85]. This phenotype is significant as the secretion of ammonia from mutagenise strains into a nitrogen starved rhizosphere of a plant has been shown to increase plant biomass[44,337]. Biophysical analysis has found that point mutations within the mutated *glnA* isoforms interfere with ammonia and glutamate binding in the primary and secondary active sites of the mature protein[85]. A previous study also showed that by replacing the endogenous genomic *glnA* promoter with an IPTG

inducible Plac promoter in the related rhizobacterium *Azotobacter vinaldii,* the strain was able to secrete ammonia into the growth medium upon low levels of *glnA* expression[338]. The *glnA* gene is therefore a fascinating and significant target for CRISPRi in the plant growth promoting rhizobacteria *A. brasilense* FP2.

6.3.4.1 Glutamine Synthetase Chemical Inhibition

To test that the GS protein activity in the cell is indeed responsible for ammonia assimilation, a chemical inhibition assay was used. The compound L-methionine sulfoximine (MSX) has been widely used as a glutamate analogue and inhibitor of GS activity in many bacterial species[339,340]. It is well known that MSX directly binds to GS at the glutamate binding pocket, competing directly against glutamate within the cell[341]. To test the effect of MSX on the GS activity in *A. brasilense* FP2, a diazotrophic growth and ammonium secretion assay was designed.

A. brasilense FP2 was grown at 30˚C with 200 rpm shaking in TY media to a culture OD595 of 0.6 (2.7.1 above). The initial cultures were then sub-cultured into 10 ml of nitrogen-free semisolid (0.175% agarose (w/v)) UMS media supplemented with 20mM succinate with an initial cell density of 10^5 cells (2.4.4 above). A range of MSX concentrations (0-75 µM) was added at the start point of the growth assay. The cultures were grown for 48 hours at 30˚C with no shaking in the until a robust pellicle was visible (Figure 2-3). After 48 hours the samples were vortexed to homogenise the cultures, an OD595 reading taken, and then

centrifuged at 6000g for 15 minutes at 4˚C. The cell pellet was then weighed both wet and after drying at 80˚C to determine cell number and protein content for ammonia production normalisation. The supernatant was filtered by vacuum filtration with a 0.22 µm cellulose filter (Advantec MFS, Dublin, Ireland) to remove any cellular debris. The ammonia concentration within the supernatant was then determined using an Orion™ High Performance ammonia electrode following the manufacturer's instructions (ThermoFisher Scientific). To calculate the ammonia concentration in the culture media samples from the mV electrode reading, a standard curve (10^{-1}- 10^{-6} M) was created in the same media as the culture growth (Supplementary Figure 3).

The results from the MSX assay revealed that an increasing concentration of MSX was correspondingly toxic to the *A. brasilense* FP2 cultures and after concentration of 50 µM MSX there was no growth observed (Figure 6-14). Between 0 and 2 µM MSX addition the OD595 of the culture decreased proportionally with a rise in the ammonia concentration in the media. The maximum ammonia secretion of $1.28X10^{-3}$ µmol $cell^{-1}$ was observed upon the addition of 2 µM MSX and was significantly greater ($p<0.05$) than at all other MSX concentrations analysed. However, at this concentration the OD595 of the culture was 4.3X lower than with no MSX addition. Upon the addition of 1 µM the OD595 of the cultures was half of that of the cultures with no MSX added, yet there was a robust $0.64X10^{-3}$ µmol $cell^{-1}$ of NH_3 excreted into the growth media. There was

no significant amount ($p>0.05$) of ammonia present in the media when no MSX was added to the culture (Figure 6-14).

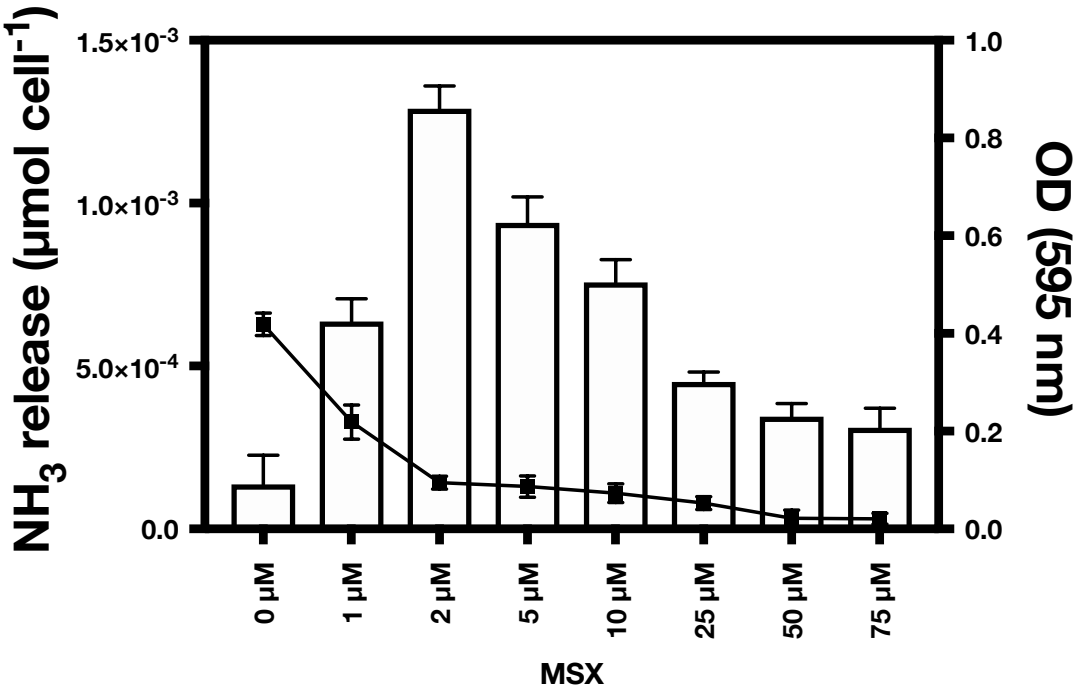


Figure 6-14 – MSX Inhibition of Glutamine Synthetase in *A. brasilense* FP2. Bar – NH_3 release into the media. Line – OD595 of the semisolid media culture. N = 3 biological and 5 technical replicates. Error bars indicate the standard error of the mean.

6.3.4.2 GlnA promoter mapping and analysis

To ensure the accuracy and on-target repression of the dCas9 CRISPRi system in *A. brasilense* FP2, promoter analysis to identify the active promoter from which *glnA* is expressed is required. The *glnA* genomic region in *A. brasilense* FP2 is

unusual in so much as the gene can be expressed from three different promoters (Figure 6-15A). Previous studies have been able to ascertain that the PglnB1 and PglnB2 promoters are sigma70 and sigma54 promoters, and therefore respectively constitutive and regulated by the RpoN low nitrogen regulator[342] (Figure 6-15B). The function and regulatory profile of the PglnA promoter remains undefined (Figure 6-15C).

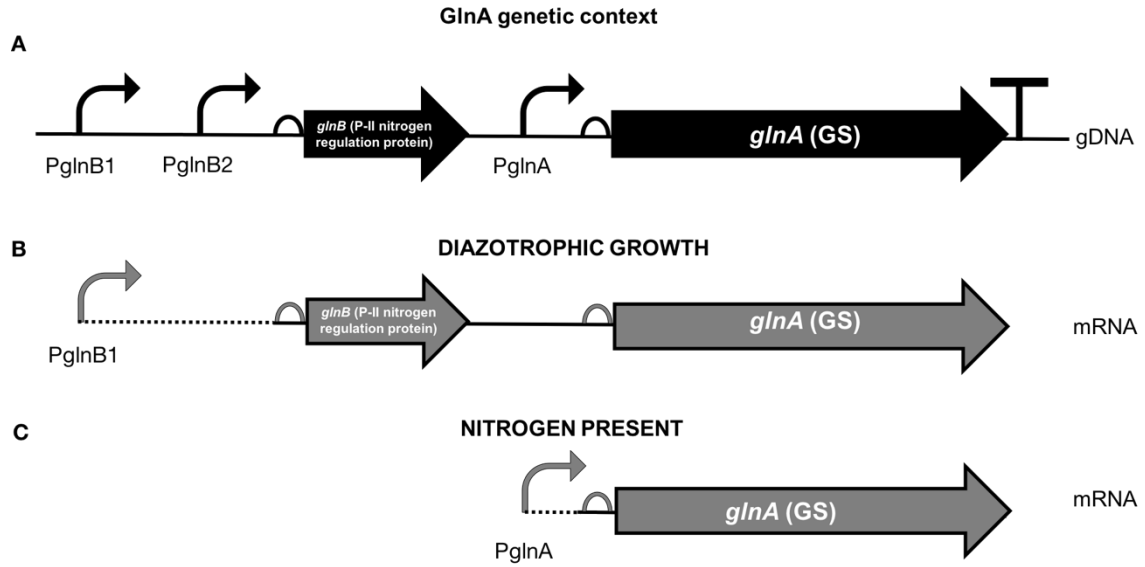


Figure 6-15 – Genetic context of the *A. brasilense* FP2 GlnA gene. **A** – Genomic DNA architecture for the GlnB and GlnA gene region. **B** – mRNA generated when under nitrogen limiting/diazotrophic conditions. **C** - mRNA generated when under nitrogen excess conditions.

To analyse the expression profiles of the *glnA* gene, a fluorescent reporter and flow cytometry analysis was performed. Three reporter plasmids were constructed, each with an sfGFP reporter protein transcribed from various

portions of the upstream *glnA* gene. Firstly Golden Gate promoter and RBS DNA parts (PL0M-PU) were constructed for the three upstream promoter regions (2.5.7 above). The first region, PglnA was PCR amplified from *A. brasilense* FP2 gDNA isolated using a PureLink™ Genomic DNA Isolation Kit (ThermoFisher Scientific) using primers oxp2138 and oxp2139 (2.5.2 above). The resulting amplicon was cloned into the pL0V-PU (pOGG068) vector by level 0 Golden Gate reation (2.5.7 above) to create the pL0M-PU-PglnA DNA part (pOGG227). The second upstream region, PglnB1&2 was again PCR amplified from the gDNA template using primers oxp2191 and oxp2192. The amplicon was then cloned into the vector pL0M-PU to form the pL0M-PU-PglnB1&2 plasmid (pOGG228). The final net upstream region including the *glnB* P-II nitrogen regulatory gene was amplified from gDNA using primers oxp2193 and oxp2194 and then cloned into the pL0V-PU plasmid to make the pL0M-PU-PglnB1&2GlnBPglnA plasmid (pOGG229).

To construct the three reporter plasmids (Figure 6-16), level one Golden Gate cloning was used to combine the three above pL0M-PU modules (pOGG227-29) with a sfGFP reporter SC1 module (pOGG037), a SsrA Fast C2 degradation tag (pOGG123) and finally a DT15S terminator (pOGG156). The broad host range plasmid backbone pOGG24 with pBBR1 and Gent^R selection was used as the vector into which the transcriptional cassettes were cloned (2.5.7 above). A SsrA fast degradation tag was included at the -COOH terminus of the sfGFP reporter protein to enhance the temporal resolution of expression. Without the degradation tag the sfGFP protein has a half-life too long for the resolution of hour by hour

expression[119]. The final three reporter plasmids (pOPS1184-6) were transformed into *A. brasilense* FP2 by electroporation (2.6.2 above).

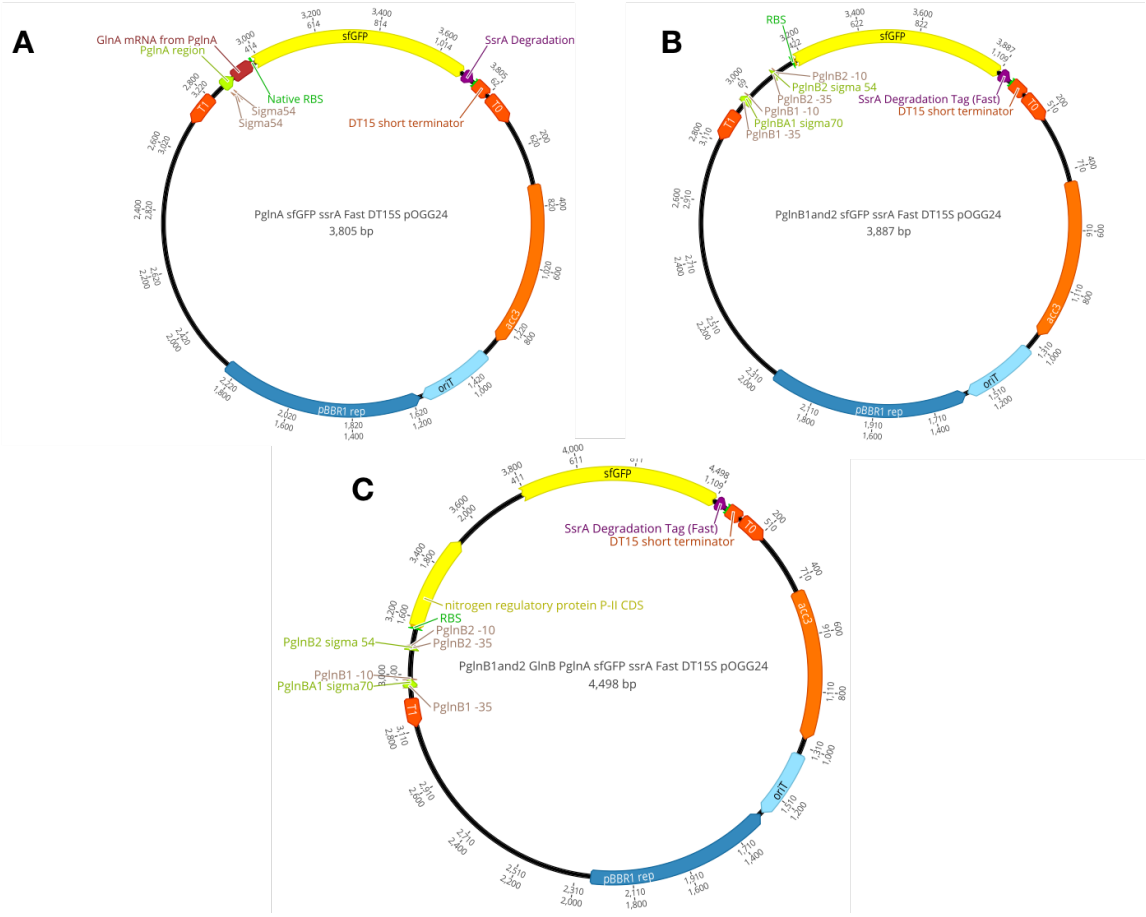


Figure 6-16 - Plasmid maps for the analysis of GlnA expression in *A. brasilense* FP2. **A** – PglnA promoter analysis plasmid pOPS1184. **B** – PglnB1&2 promoter analysis plasmid pOPS1185. **C** – PglnB1&2, GlnB, PglnA analysis plasmid pOPS1186. Note pBBR1 based origin of replication and acc3 gentamycin selection gene. Specific promoter motifs are annotated, e.g. Sigma 54 binding region in the PglnA promoter region.

To analyse the transcription profile of the three upstream fragments of the *glnA* gene a high throughput growth and fluorescent reporter assay was performed (2.7.1 and 2.7.3 above). Cultures of the three test strains were grown at 30˚C to

an OD595 of 0.6 in UMS media supplemented with 20 mM succinate and 20 mM NH_3Cl. Four test culture conditions were then used to assay the transcriptional profile of the upstream promoter regions of the *glnA* gene. This was done as to reveal the nitrogen regulatory properties of the promoters. Each of the four culture conditions was based on semisolid UMS (0.175% (w/v) agarose) with 20 mM succinate and vitamins and then supplemented with one of the following: 20 mM NH_3Cl, 20 mM L-glutamate, 20 mM L-glutamine or no additional nitrogen. Within a gas-impermeable universal tube, 10 ml of the test medium was added and inoculated with $1X10^5$ cells of each of the test reporter strains. The cultures were grown in parallel at 30˚C with no shaking. Every 4 hours a universal for each strain and culture condition combination was vortexed, an OD595 measurement taken. To analyse the population-wide response of the expression cassettes and present single-cell resolution transcriptional data, flow cytometry was utilised (2.7.4 above). This was only performed for this experiment and not others due to it being a small yet successful pilot study. The gating strategy we implemented involved the removal of doublet data points by gating in the FSC/FSC-area linear region, removal of dead cells by gating in the FSC/SSC plot and finally the sub-setting the sfGFP positive cells[343]. The final REU value for each data point is the product of $>1X10^6$ events and both three technical and biological replicates (2.7.3 above). The time point selected for the representation of the expression strength of each promoter fragment was 44h when the cultures were in late log phase.

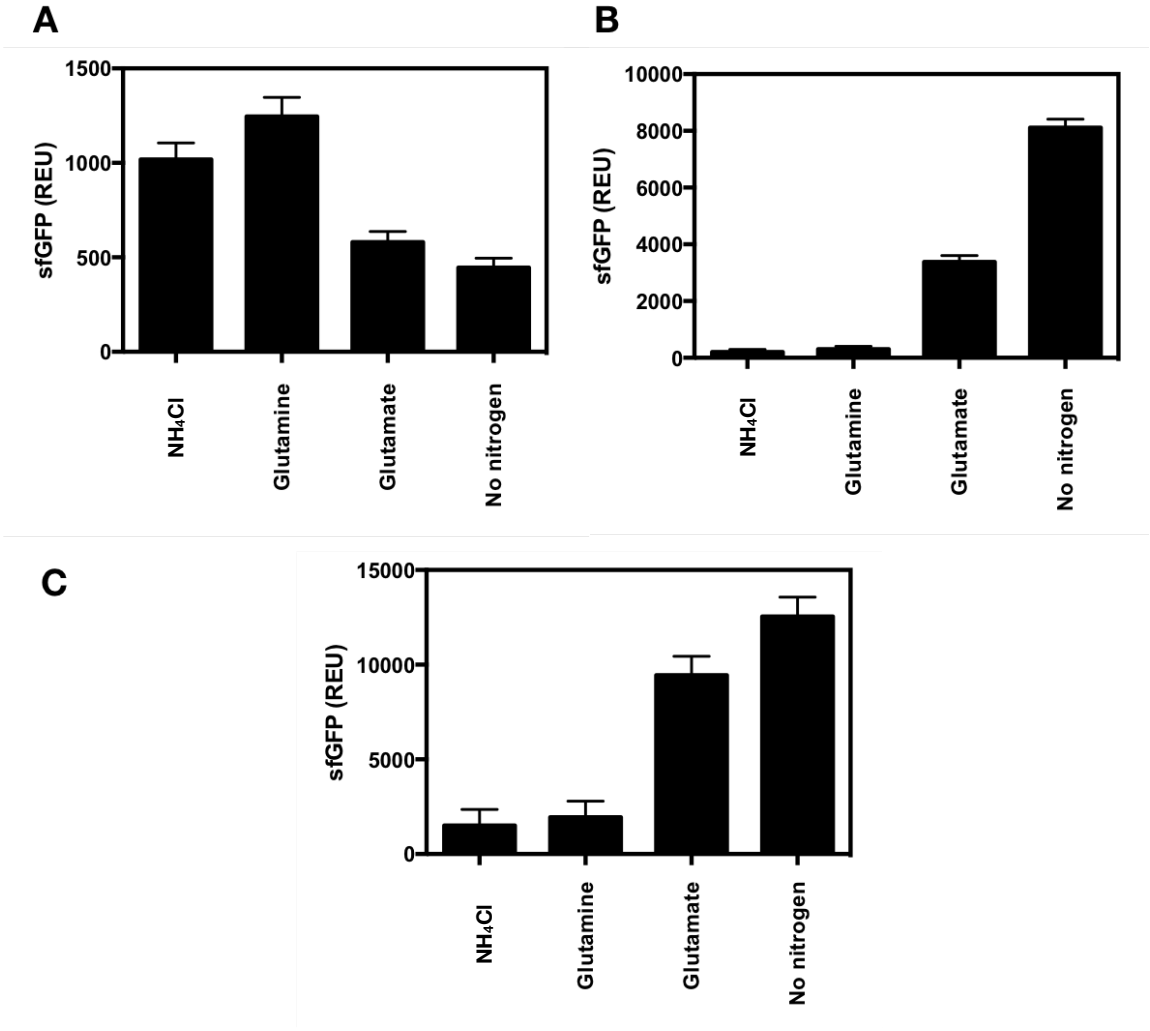


Figure 6-17 – *A. brasilense* FP2 *glnB*/*glnA* promoter expression analysis under various growth conditions. **A** – PglnA promoter. **B** – PglnB1&2 promoters. **C** – Combined PglnB1&2 – *glnB* – PglnA genetic region. N = three biological and three technical replicates. Error bars indicate the standard error of the mean.

The analysis of the upstream fragments of the *glnA* genomic region reveal varying transcription profiles in the nitrogen availability culture conditions used (Figure 6-17). The PglnA promotor region caused significantly greater ($p<0.05$) transcription under nitrogen excess conditions than nitrogen limiting conditions

(Figure 6-17A). The greatest transcription of the sfGFP reporter (1282 REU) was observed in the culture containing glutamine, which was significantly greater ($p<0.05$) than any of the other conditions. The second highest transcription was observed for ammonia (1057 REU) with both glutamate (575 REU) and no nitrogen (diazotrophic growth at 466 REU) revealing the lowest transcription.

The second region analysed for its transcriptional strength was that of the PglnB1&2 region (Figure 6-17B). The positive sfGFP response observed from the region was a magnitude greater than that of the PglnA promoter region. The highest transcription was observed for the cells grown under diazotrophic conditions (8681 REU). The second highest transcription was observed in the cells grown with glutamate as the carbon source, however the expression observed was 42% lower (3669 REU) than observed under diazotrophic growth conditions. Finally there was no significant ($p>0.05$) transcription observed from the PglnB1&2 promoter region with either ammonia of glutamine as the nitrogen source.

The final region analysed for transcriptional activity was the entire upstream region of the *glnA* gene, including the *glnB* gene and all three promoters; PglnB1, PglnB2 and PglnA (Figure 6-17C). The positive transcription observed from this region was 50% greater than that observed for the PglnB1&2 region (Figure 6-17B). The profile of transcriptional strengths resembled that of the PglnB1&2 region with the greatest transcription observed under diazotrophic growth conditions (13205 REU), followed by glutamate (9288 REU) and finally glutamine and ammonia (2124

REU). There was no significant difference observed ($p<0.05$) in transcription in either the glutamine or ammonia supplemented growth conditions.

In summary the PglnA promoter region produced the greatest transcription under nitrogen excess conditions. The PglnB1&2 region caused 10X higher transcription and conversely caused the greatest transcription under nitrogen-limitation conditions. Finally the full upstream region of the *glnA* gene caused the greatest transcription out of all the fragments analysed, with the transcription greatest under nitrogen limiting conditions.

6.3.4.3 GlnA sgRNA design

Subsequent to the *glnA* gene promoter analysis (6.3.4.2 above) the design of a suite of sgRNAs targeting the PglnA promoter region was performed. Using the sgRNA computational design tool embedded in the Geneious R11 software (BioMatters Ltd), the successful design of six sgRNAs was accomplished. The script, designed originally by Doench *et al.,* 2014[344] allows for the design of highly specific and active sgRNAs. Using a full draft *A. brasilense* FP2 genome as an input library, the off-target effects of sgRNA non-specific binding are minimised. Of the six sgRNAs designed, 3 bind to the non-template DNA strand (NT) and three bind to the template (T) strand (Figure 6-18). This was done to analyse the effects of strand specificity for the activity of the dCas9 transcriptional silencing. Many previous studies have shown stochiometric interference of the RNA

polymerase by dCas9 only occurs when the sgRNA targets the NT DNA strand[65,66,324]. The sgRNA molecules were designed to target 20 bp regions after a 3' NGG PAM sequence and are therefore only compatible with the dCas9 wt variant.

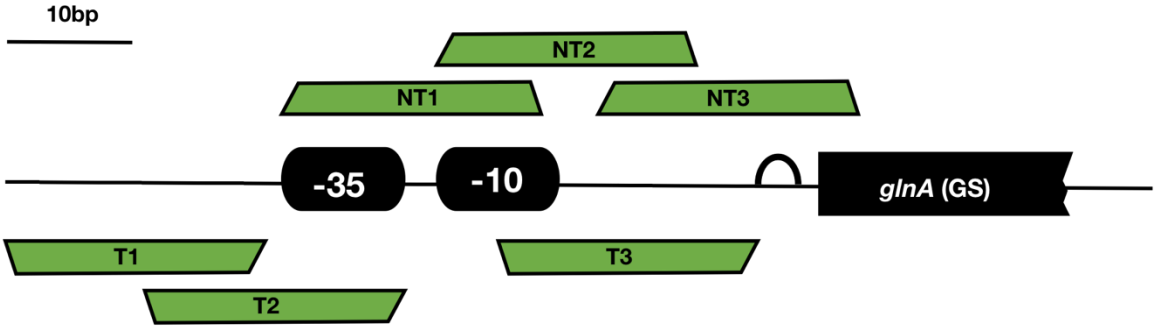


Figure 6-18 – sgRNA targets for the repression of GlnA expression by CRISPRi in *A. brasilense* FP2. NT – Non-template strand-binding sgRNA sites. T - Template strand-binding sgRNA sites.

The sgRNAs were designed to be expressed from the pOGG230 plasmid and as such the sgRNA molecules were constructed as promoter-target oligonucleotides (6.3.3 above). Each of the 20 bp sgRNA target modules were preceded by the DNA sequence of the J23119 strong sigma70 constitutive promoter[56] and the entire oligo was flanked with BsaI sites. Both a sense and antisense oligonucleotide were synthesised and then annealed to form linear dsDNA modules (2.5.9 above). The dsDNA promoter and sgRNA modules were each cloned into the pOGG230 sgRNA expression plasmid by level 1 BsaI Golden Gate reaction (2.5.7 above) creating plasmids pOPS1159-1164.

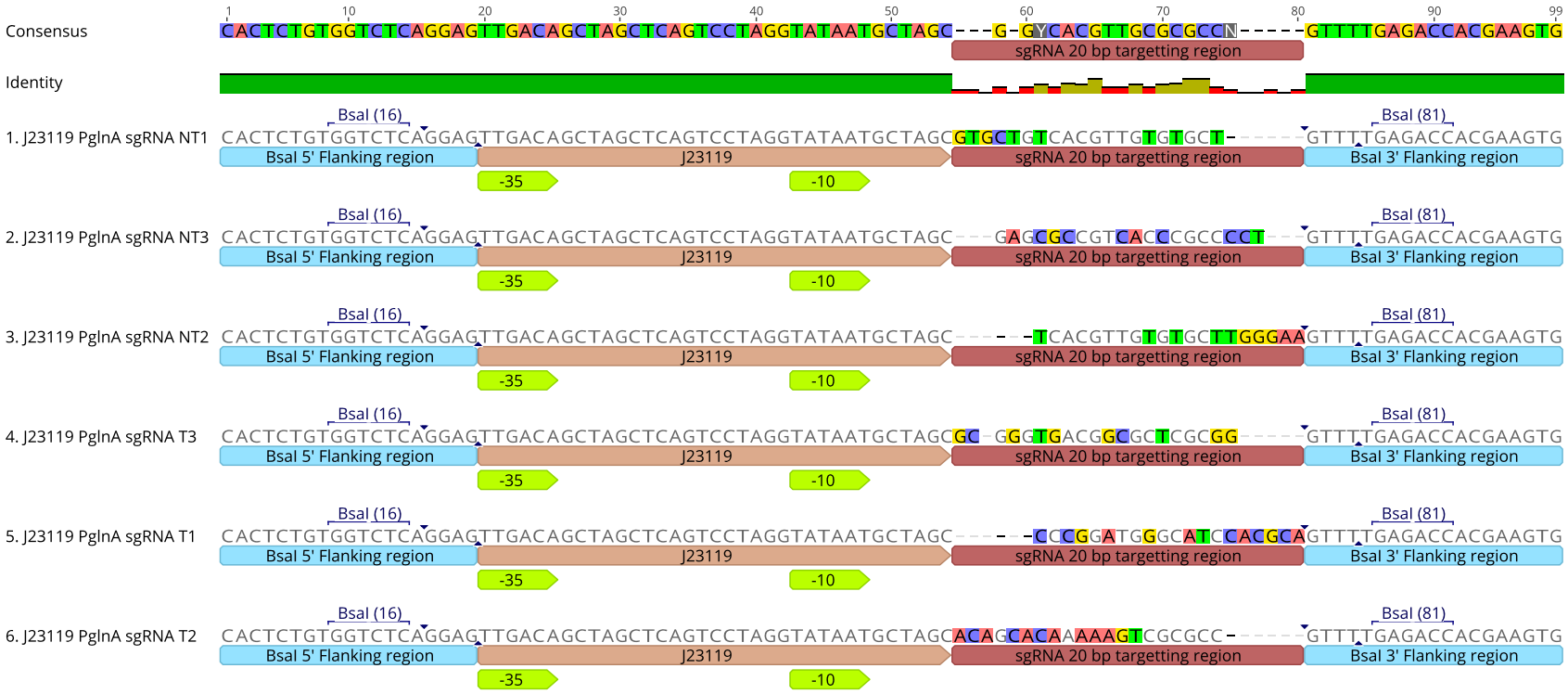


Figure 6-19 – DNA sequence alignment of the six sgRNA promoter and 20 bp target regions for CRISPRi of PglnA in *A. brasilense* FP2. J23119 – Strong constitutive promoter.

6.3.4.4 *In vivo* CRISPRi of GlnA

To analyse the effect of CRISPRi on the *glnA* gene in *A. brasilense* FP2, a dual dCas9 expression and sgRNA expression system was used. Firstly the pOPS1150 plasmid with aTc inducible dCas9 wt expression cassette was conjugated into *A. brasilense* FP2 by conjugation (2.6.3 above). An electrocompetent stock of this strain as well as the other two dCas9 expression variants (pOPS1151 (EQR) and pOPS1152 (VRER)) were made and stored at -80˚C for rapid sgRNA plasmid addition. Each of the six sgRNA expression plasmids targeting PglnA was separately transformed into the strain with the pOPS1150 plasmid by electroporation (2.6.2 above).

To test the dCas9 CRISPRi on the GS activity in *A. brasilense* FP2, a diazotrophic growth, ammonium secretion and qPCR composite assay was designed. Each of the six strains with dual plasmid CRISPRi circuits were grown at 30˚C with 200 rpm shaking in UMS media supplemented with 20 mM succinate and 20 mM ammonia to a culture density of OD595 of 0.6 (2.7.1 above). The initial cultures were then sub-cultured into gas impermeable universal tubes with 10 ml of nitrogen-free semisolid (0.175% agarose (w/v)) UMS media supplemented with 20mM succinate. The initial was 10^5 cells per 10 ml universal tube (2.4.4 above). Induction of the dCas9 wt expression cassette was performed by the addition of 1 µM aTc to the universal tubes at the start point of the assay.

The cultures were grown for 48 hours at 30˚C with no shaking the until a robust pellicle was visible (Figure 2-3). After 48 hours the samples were vortexed to homogenise the cultures, an OD595 reading taken, and then centrifuged at 6000g for 15 minutes at 4˚C. A fraction of the cell pellet was then weighed both wet and after drying at 80˚C to determine cell number and protein content for IAA production normalisation. The supernatant was filtered by vacuum filtration with a 0.22 µm cellulose filter (Advantec MFS, Dublin, Ireland) to remove any cellular debris. The ammonia concentration within the supernatant was then determined using the same method as described in section 6.3.4.1 above.

For the subsequent qPCR analysis and GS activity assays, five universal tubes were combined in triplicate to yield enough cellular material for robust measurements, therefore providing 5 biological and 3 technical replicates for the assays.

For the qPCR analysis of *glnA* mRNA quantification the total RNA of the pooled cellular material was extracted using a TRIzol MAX Bacterial Isolation Kit (Invitrogen, USA). The total RNA quality and composition was then verified on a 2100 Bioanalyser (Agilent, USA) prior to cDNA synthesis. cDNA was synthesised from random hexamer primer reactions using a SuperScript III First Strand Synthesis Kit (Invitrogen, USA). cDNA was then quantified using a NanoDrop ND-100 spectrophotometer (ThermoFisher, USA) prior to qPCR analysis. qRT-PCR reactions were performed on a QuantStudio 3 thermocycler (ThermoFisher, USA).

Each reaction comprised of 1X Applied Biosystems™ PowerUp™ SYBR™ Green Master Mix (ThermoFisher, USA), 0.5 µL of the forward primer (oxp2457), 0.5 µL of the reverse primer (oxp2458) and template cDNA transcribed from 10 ng of the total RNA. Each sample was measured in triplicate with negative (cDNA absent and primers absent) controls. For normalisation of the qPCR data, two reference housekeeping genes were used that have been previously characterised and represent consistent expression profiles across culture conditions and phases of the growth profile[345]. The two reference genes chosen, *gyrA* and *glyA* were amplified with primer pairs oxp2454/oxp2455 and oxp2451/oxp2452 respectively. The *glnA*, *gyrA* and *glyA* gene amplicons were each 200 bp long with a primer Tm of 60°C. PCR analysis on *A. brasilense* FP2 confirmed primer specificity and amplicon size verification (Figure 6-20 A).

Expression analysis of the *glnA* and reference genes was converted from the threshold cycle value (Ct) to a fg RNA / µg Total RNA by normalising with a standard curve of PCR amplified *glnA* DNA sequence. The amplification purity, efficiency and correlation coefficients were calculated using the QuantStudio Design and Analysis Software v1.4.3. The melt peak analysis (Figure 6-20B) verified a single consistent peak for all three primer sets used indicating amplification of a single product, primer dimer formation was not observed using the primer sets selected.

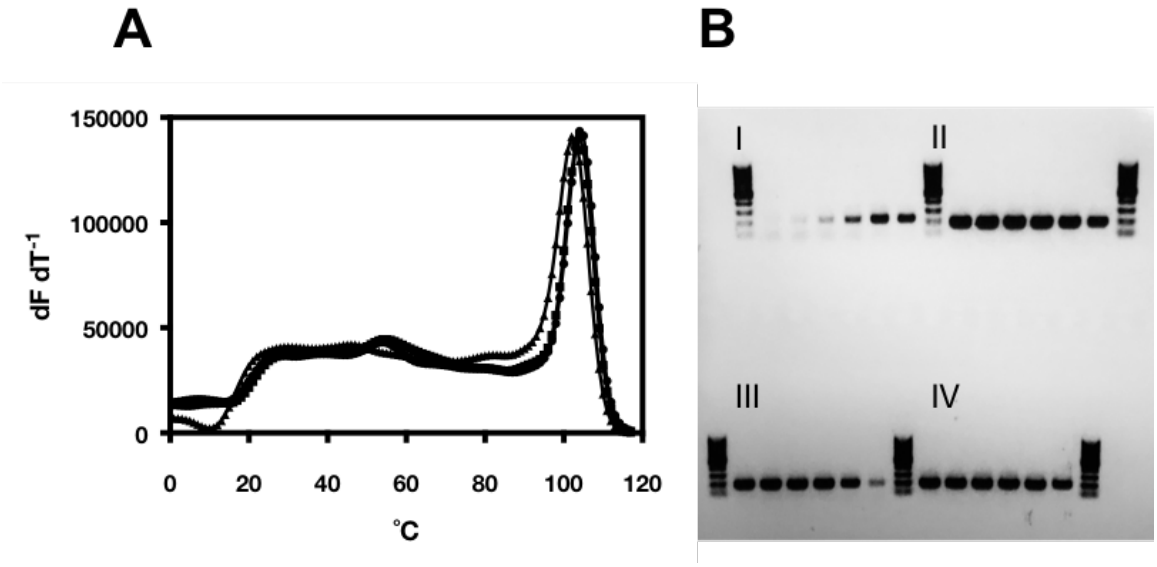


Figure 6-20 – CRISPRi qPCR quality control data for *glnA* gene repression in *A. brasilense* FP2. **A** - Melt curves for target and reference control genes. **B** – Thermal gradient (57.5-70˚C) agarose gel (0.9% w/v) showing amplification of a single product of expected 200 bp for each the r*ecA* positive control (I), *glnA* (II), *gyrA* (III) and *glyA* (IV) genes.

To assay GS activity *in vivo* a Glutamine Synthetase Microplate Assay Kit (Cohesion Biosciences, USA) was used. The assay relies on the γ-glutamyl transferase reaction, catalyzed by GS. γ-glutamylhydroxamate, synthesised from glutamine and hydroxylamine, reacts with $FeCl_3$ to form a complex with specific absorbance at 540 nm, therefore directly revealing GS activity (Figure 6-21). For the subsequent process all samples and reagents were kept on ice or at 4˚C. Each pooled bacterial pellet sample from five universal tubes was centrifuged at 4000g at 4˚C for 10 minutes. Assay buffer (containing ATP, Na_3AsO_4 and Mn^{2+} was added at 1 ml per $5X10^6$ cells and the mixture resuspended by vortex. The cells were

disrupted by sonication using a Q700 Sonicator (QSonica, USA) with a 20% power level for 3s. The mixture was allowed to cool for 10s in an ice bath before another 3s sonication, this process was repeated 30 times. The cell lysis solution was then fractionated by centrifugation at 8000g at 4˚C for 15 minutes to separate the cellular debris from the bacterial contents. For the final assay mixture, 70 µl of the supernatant of each sample was transferred to a microfuge tube, to which 160 µL of reaction buffer, and 70 µl of the substrate (L-glutamine and Hydroxylamine) was added. The mix was then incubated at 37˚C for 30 minutes before 100 µl of 90 mM $FeCl_3$ was added. The microfuge tubes were then mixed, centrifuged at 8000g at 4˚C for 10 minutes and then 200 µl of the supernatant was transferred to a 24-well microwell plate. Absorbance measurements were made at 560 nm after blanking to sample-free assay mix.

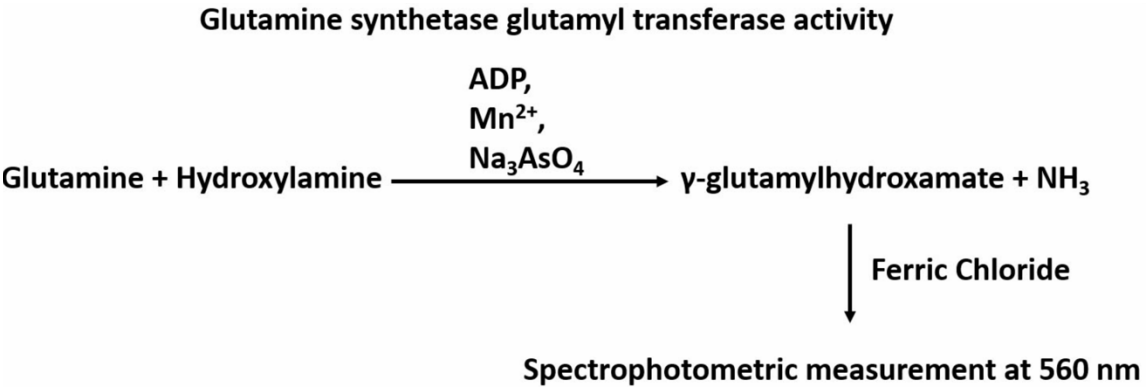


Figure 6-21 – Chemical overview of the GS glutamyl transferase reaction used to detect GS activity in *A. brasilense* FP2.

To quantify the amount of GS activity in the sample the following equation was used:

$$GS\ (U\ g^{-1}) = \left(OD_{sample} - OD_{control}\right) \times V_{total} \div \left(V_{sample} \times \frac{W}{V_{assay}}\right) \div 0.01 \div T$$

Figure 6-22 – Equation to calculate activity of GS. Where: **W** - the weight of sample, g; $\mathbf{V_{Total:}}$ - the total volume of the enzymatic reaction, 0.4 ml; $\mathbf{V_{Sample}}$ - the volume of sample, 0.07 ml; $\mathbf{V_{Assay}}$ - the volume of Assay buffer, 1 ml; **T** - the reaction time, 30 minutes. Each data point presented is the product of five biological and three technical replicates.

The results from the ammonia secretion assay of the six GS CRISPRi strains revealed significantly more ($p<0.05$) ammonia release for those sgRNA molecules that targeted the non-template strand as oppose to those that targeted the template strand when dCas9 expression was induced (Figure 6-23A). The greatest ammonia release ($1.7x10^{-10}$ mol $cell^{-1}$) was observed in the *A. brasilense* FP2 strain with the dCas9 with NT1 sgRNA. The ammonia quantity released was 7.1X greater when induced than uninduced. Both the NT2 and NT3 sgRNA CRISPRi strains produced significantly less ($p<0.05$) ammonia secretion than the NT1 sgRNA with $1.3x10^{-10}$ mol $cell^{-1}$ observed between them. There was no significant difference ($p>0.05$) between the three template strand targeting sgRNAs with $3.7x10^{-10}$ mol $cell^{-1}$ ammonia release observed. Both of the control strains, sgRNA negative and dCas9 negative showed no significant ammonia release ($p>0.05$) under the experimental conditions used. The unexpressed

conditions for the CRISPRi strains revealed that the NT1 and NT2 sgRNA strains released significantly more ammonia than all of the other sgRNAs ($2.4x10^{-11}$ mol $cell^{-1}$). The other sgRNA strains released $1.4x10^{-11}$ mol $cell^{-1}$ ammonia, only just above the detection limit of the assay.

The data from the OD595 measurements of the CRISPRi strains revealed that an increase in ammonia release corresponded with a decrease in the OD595 of the culture at the time of measurement (Figure 6-23B). The greatest decrease in OD595 was observed for the NT1 sgRNA strain at 0.67, 42% less than when in the uninduced dCas9 absent state. Both the NT2 and NT3 sgRNA CRISPRi strains had an OD595 of 0.95 when dCas9 was induced, 22% less than when in the uninduced state. There was no significant difference observed ($p>0.05$) between the OD595 of the template strand targeting sgRNAs ether when the dCas9 was induced (OD595 = 1.06) or uninduced (OD595 = 1.29). Both of the negative controls indicated OD595 readings that were not significantly different ($p>0.05$) to those of any of the sgRNAs apart from NT1. The NT1 sgRNA strain when uninduced had an OD595 (1.15) that was significantly lower ($p<0.05$) than the OD595 of all the other strains when in the uninduced state. There was a significant decrease in the OD595 of both negative controls when induced with 1 μM aTc. The dCas9 expression only control indicated a 22% decrease in OD595 when induced whereas the sgRNA expression only control indicated only a 7% decrease.

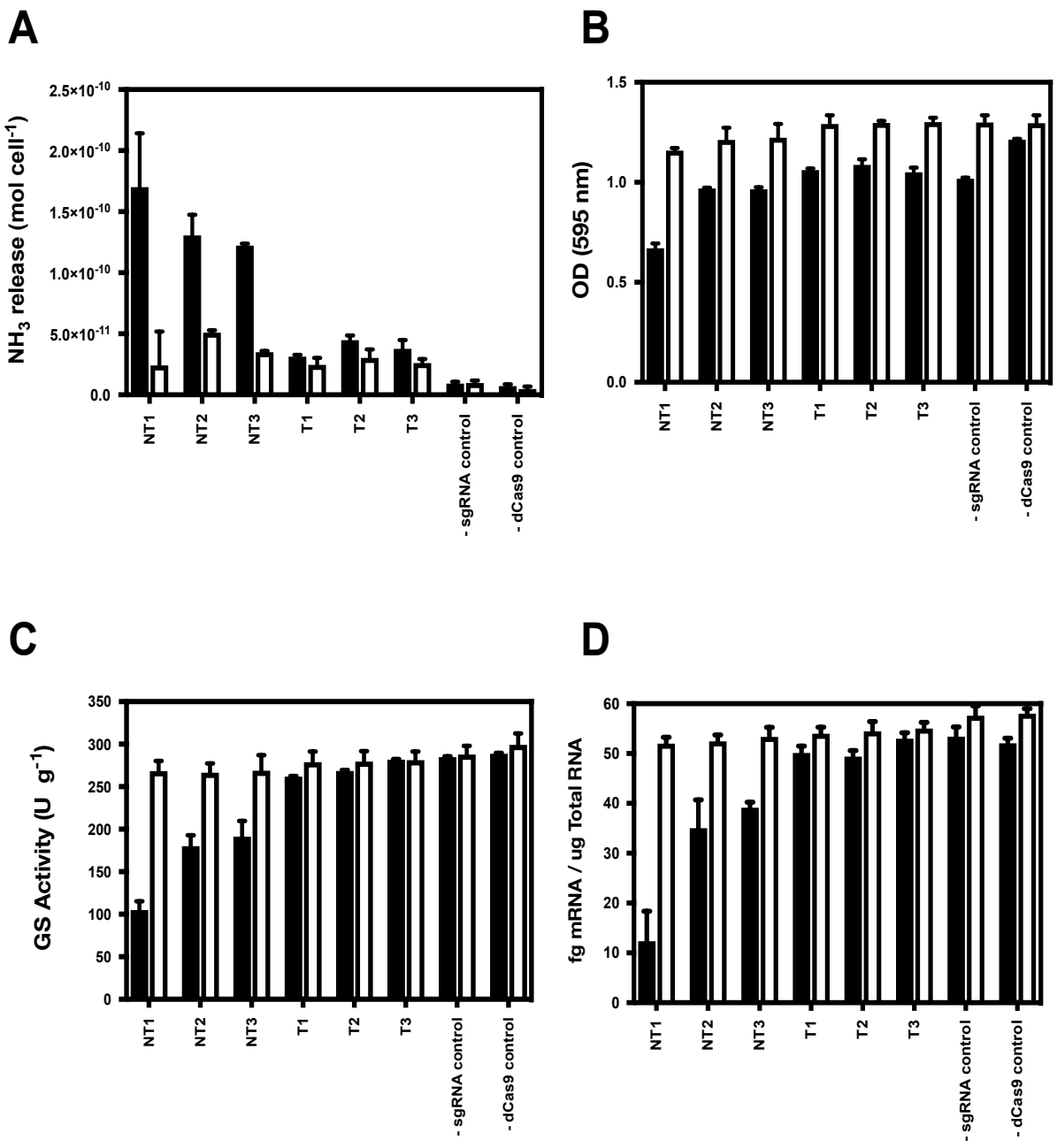


Figure 6-23 – Analysis of the dCas9 wt repression of the GlnA gene of *A. brasilense* FP2 in diazotrophic growth conditions. **A** – Ammonia (NH_3) release. **B** – OD595. **C** – Glutamine synthetase (GS) activity. **D** – qPCR on the GlnA (GS) transcript. N = three biological and three technical replicates. White bars indicate uninduced state (input of 0) while black bars indicate the induced state (input of 1) with 1 μM aTc. Error bars indicate the standard error of the mean.

Analysis of the GS activity of the CRISPRi *A. brasilense* FP2 strains revealed a significant decrease ($p<0.05$) in activity of those strains where the non-template strand was targeted (Figure 6-23C). In the dCas9 induced state the NT1 sgRNA strain had the least GS activity at 105 U g^{-1}, 62% less than when in the dCas9 uninduced state. The NT2 and NT3 sgRNA strains were not significantly different ($p>0.05$) in their GS activity of 185 U g^{-1} in the dCas9 induced state, 32% less than when in the uninduced state. There was no significant difference observed ($p>0.05$) between the dCas9 induced and uninduced states of all of the template strand targeting sgRNA strains, of the two negative control strains. The activity of 281 U g^{-1} of GS was consistent across all uninduced states of the sgRNA strains as well as the two negative control strains.

qPCR analysis of the *glnA* transcript in the CRISPRi *A. brasilense* FP2 strains revealed a negative correlation between the increase in ammonia release and the decrease in transcription of the *glnA* gene (Figure 6-23C). In reflection of the GS activity data, the induced dCas9 NT1 sgRNA strain had the least *glnA* transcription with a 77% decrease in expression compared to the uninduced state. Both the NT2 and NT3 sgRNA strains had significantly less ($p>0.05$) GS expression in the dCas9 induced state (37 sg mRNA / µg total RNA) compared to the uninduced state (53 sg mRNA / µg total RNA). There was no significant difference ($p>0.05$) observed between the three template strand targeting sgRNA strains when in each state. However there was a significant 11% decrease in *glnA* expression in the dCas9 induced state as opposed to the uninduced state.

Both of the negative controls had significantly greater ($p<0.05$) expression of the *glnA* gene when dCas9 was not expressed, compared to all other strains in the uninduced state. However the sgRNA expression-only negative control did show 10% less *glnA* expression in the induced (+1 μM aTc) as opposed to the uninduced state.

In summary there is a clear and statistically significant correlation between: an increase in ammonia release from the *A. brasilense* FP2 cells (Figure 6-23A), a decrease in the cultures OD595 (Figure 6-23B), a decrease in the cells GS activity (Figure 6-23C) and finally a decrease in the *glnA* expression within the cell. This pattern is only visible in the dCas9 induced state and therefore each of the characteristics is dependent on the expression of the dCas9 wt protein.

6.4 Discussion

6.4.1 dCas9 Function in rhizobacteria

The architecture of the CRISPRi system that we characterise in this work is relatively simple and builds off of the DNA parts characterised in the RhizoKit toolbox (3 above). The platform is a functional DNA editing-free system for targeted gene regulation. Our CRISPRi system does not require any endogenous host cellular machinery to function, instead relying solely on a dCas9 protein variant and corresponding sgRNA to be expressed[74,130]. This system can therefore be rapidly implemented in rhizobacteria to regulate target genes with only a small level of cloning required. The flexibility of the Golden Gate compatible single CRISPRi plasmids (6.3.3 above) allows for the rapid assembly of fully functional and unique gene targeting dCas9 variants in a broad host range of species. The three dCas9 variant and sgRNA Golden Gate plasmids are also compatible with sgRNA multiplexing whereby multiple sgRNAs can be integrated and expressed together[76,317,326]. Previous studies have shown that target gene silencing can be enhanced to the level of negligible gene transcription by multiplexing multiple dCas9 proteins to the 5' end of the target genes ORF[326,346].

The analysis of the dCas9 expression toxicity to the two rhizobacterial strains indicates that the impact was greatest in the *A. caulinodans* ORS571 strain

compared to *A. brasilense* FP2 (Figure 6-5 & Figure 6-6). However the higher induction concentration of 1 μM aTc caused severe toxicity to both strains. This is due to the over-expression of the dCas9 protein causing a significant burden to the cells biosynthetic capacity, preventing the cells growth rate being optimal[323]. This finding is in line with other dCas9 expression studies that have shown dCas9 to cause toxicity effects when overexpressed in strains with no sgRNA present[75]. The induction at the normal (optimum) 1 μM aTc concentration did cause a significant negative impact to the growth of the cells, but it is likely that future induction timing of the dCas9 expression cassette could alleviate the burden on the cells metabolic capacity[292]. aTc does not cause toxicity to cells due to its poor binding affinity to the ribosomal 30S subunit. This coupled with its over 30X greater binding affinity to the TetR protein makes the molecule a robust inducer of choice[347]. The lack of any significant decrease in cell growth for the induced negative control plasmid strains indicate that the effect on the cells growth is due to the dCas9 protein expression and not that of the plasmid maintenance or transcriptional regulator expression.

Our results from the activity of the dCas9 variants repression of the reporter sYFP2 protein were very promising (6.3.2 above). The fold-repression of the target gene with a low level sgRNA expression cassette shows a repression profile that is applicable for sensitive gene regulation but not for robust knock-down of target genes. However upon high levels of expression of the sgRNA, the repression profiles for all three variants in both rhizobacterial strains increased dramatically.

This is due to the stability of sgRNA molecules within the cellular environment being very low with a half-life of only 15 minutes when unbound to a dCas9 protein compared to >2 hours when bound[348]. Therefore expressing the sgRNA molecules to a higher concentration within the cell increases the probability of a full sgRNA-dCas9 complex forming and therefore DNA targeting and CRISPRi to take place[349]. The expression of sgRNAs within our CRISPRi system is therefore one key node in tuning the repression of a target gene. Gene expression interrogation in rhizobacteria to date has consisted almost wholly of mutagenesis or knock-outs, a binary and inflexible approach to gene regulation. By tuning the CRISPRi repression of a gene to a predictable level, the genes function and sensitivity of any biosynthesis or regulatory circuits can be illucidated[182,324]. In the work presented here the specificity of the sgRNA to the target DNA sequence of the sYFP2 promoter was 100%. We believe that causing mismatches to the sgRNA targeting region will further allow for tuning of the CRISPRi repression of target genes, a hypothesis supported in the work by Zhang *et al.,* 2017[350]. With the ability to induce the dCas9 protein to various levels upon the addition of differing concentrations of inducer, coupled with the sgRNA expression and specificity variations, we believe the CRISPRi system presented here will be a significant tool for future rhizobacterial engineering and wider synthetic biology as a whole.

6.4.2 CRISPRi of *glnA* in *A. brasilense* FP2

A. brasilense FP2 was chosen as the rhizobacterial chassis for CRISPRi for two reasons. Firstly the organism has been proven repeatedly to colonise target cereal plants such as Barley (*Hordeum vulgare*), Wheat (*Triticum aestivum*) and Rice (*Oryze sativa*) with robust cell counts per g plant root material observed[40,44,351,352]. Secondly, the glutamine synthetase *glnA* gene in the organism has been revealed in separate studies to be a key node in nitrogen assimilation and disruption of the genes function has produced the desirable ammonia excretion phenotype. As a first point of interrogation we successfully demonstrated that the inhibition of GS activity within the *A. brasilense* FP2 strain did indeed cause a robust and significant ammonia release phenotype (6.3.4.1 above). The maximum quantity of ammonia released observed were within the same order of magnitude as those observed when a similar MSX inhibition assay was performed on the plant-associative rhizobacteria *Azotobacter vinlandii* [353]. This result therefore confirmed the ammonia phenotype in the strain was replicable. We hypothesised that a CRISPRi approach to *glnA* repression would allow for an optimum balance between ammonia release and cell viability to be achieved.

To optimally target the CRSIPRi system to repress the *glnA* gene, a thorough promoter analysis was performed. The successful analysis of the endogenous *glnA* upstream promoter region of *A. brasilense* FP2 revealed a complex and highly nitrogen-dependent expression profile (6.3.4.2 above). Previous studies

have revealed that there are three upstream promoters of the *glnA* gene, a finding our results supported[342]. However, we were able to demonstrate that the expression strengths from the promoters under the various nitrogen-source conditions varied greatly. We hypothesise that the weak constitutive PglnA promoter maintains a constant basal expression level of the *glnA* gene even when glutamine concentrations within the cell are high. This then allows for the rapid induction of the positive feedback loop to activate nitrogenase when nitrogen is limited within the cell[85]. The PglnB2 promoter is another weak constitutive promoter that maintains a basal concentration of active GlnB within the cell, required when nitrogenase activation occurs. The PglnB1 promoter is a strong sigma54-regulated promoter that we have shown causes high levels of *glnB* and *glnA* expression under nitrogen-limiting conditions. The high *glnB* expression causes rapid nitrogenase activation, therefore causing more ammonia to be synthesised within the cell. The high concentration of GlnA within the cell therefore causes rapid ammonia assimilation into glutamine and subsequently glutamate (via GOGAT), both key precursors in amino acid, purine and pyrimidine biosynthesis[336]. Our results from the *glnA* promoter analysis supported all the above consensus mechanisms within *A. brasilense* FP2 and therefore informed the rational design of six sgRNA's to target the PglnA upstream promoter of *glnA.*

The results from the integrated diazotrophic repression of *glnA* by our dCas9 wt CRISPRi system revealed a robust correlation between *glnA* repression and ammonia release *in vitro* (6.3.4.4). This result was further verified by the negative

correlation between ammonia release and the *glnA* mRNA transcript levels within the cell and the GS activity of the cell. There was a clear non-template strand bias for successful stochiometric inhibition of the RNA polymerase binding to the PglnA promoter, and therefore *glnA* repression. This result is in line with the consensus of the field whereby dCas9 has a strong non-template strand bias for target gene repression[130,323]. The results of the qPCR analysis of the CRISPRi repressed strains revealed that the mRNA transcript of the *glnA* gene was of a lower concentration in the strains with an induced CRISPRi system. Verification of the mRNA transcript concentration being decreased is a key result confirming that the dCas9 system is repressing the transcription of the *glnA* gene rather than having some other effect that causes the ammonia excretion phenotype. Further RNAseq analysis of the strains with the active CRISPRi system would reveal valuable data as to the specificity of the sgRNA-dCas9 targeting and whether any other off-target effects were taking place[354]. With our data we cannot dismiss the possibility of other off-site dCas9 binding events on the genome. The correlation between an increase in ammonia excretion and decreases in *glnA* gene transcription and cellular GS activity proves that the GlnA protein activity levels are directly responsible for the ammonia excretion phenotype.

We hypothesise that with further tuning of the CRISPRi circuit, an optimum balance between the ammonia release of the cell and the viability of the cell can be found. Furthermore a fully characterised and orthogonal plant-produced signal such as rhizopine could be used to directly link the CRISPRi system and the

ammonia excretion phenotype to a host plant (1.6.1.1 above). This circuit would have the added benefit of leaving the engineered rhizobacterial strain with wild-type competitiveness levels within the rhizosphere, only shifting to an ammonia excreting and compromised state when associated with the target plat root system. Future efforts to enhance on the CRISPRi system created in this study may also look into the benefits of integrating the dCas9 protein and expression control modules into the host rhizobacterial genome. This would have the added advantage of reducing the cellular burden of multicopy plasmid maintenance on the cell, and studies have also shown greater dynamic responses for such circuits when integrated[355,356].

7 General Conclusions and Future Directions

7.1 Overview

This section of the thesis will provide a brief summary of the results (7.2 below). The wider implications of the findings are discussed and future avenues of enquiry and experimentation are also suggested. This section of the thesis ends with concluding remarks and conclusions (7.3 below).

7.2 Summary of Results and Future Work

7.2.1 RhizoKit: Rhizobacterial DNA Toolkit

The aim of this thesis was to characterise rhizobacteria as synthetic biology chassis and demonstrate the applications of various circuits to solve significant agricultural issues. In the first chapter the design, construction and characterisation of a full library of DNA 'parts' was successful and allowed all subsequent work to be performed. By characterising the behavior of each of the DNA parts in the two rhizobacterial species *Azorhizobium caulinodans* ORS571 and *Azospirillum brasilense* FP2, the rational design of synthetic circuits was achievable. Most of the DNA parts constructed in this study mirrored their characteristics in both strains and to the synthetic biology mainstay chassis

Escherichia coli with only minor variances observed[357]. While the number of parts in each category is limited, this robust study outlines pipelines for the rapid characterisation of future parts and their addition to the library. The first priority for future work is the expansion of the suite of constitutive promoters. This can be done in one of two ways. Firstly the other 49 members of the constitutive sigma70 Anderson promoter library could be generated as Golden Gate pL0M-P parts and their transcriptional strength in both rhizobacteria analysed[56]. Secondly, a consensus sequence for the constitutive promoters of housekeeping genes such as *recA* or *gyrB* can be determined[358]. Degenerate oligonucleotide libraries of every variation of this consensus sequence could be synthesised and then combined in a Golden Gate shuffling experiment to produce a library of potential rhizobacterial reporter plasmids[359]. These could then be analysed by the high throughput fluorescence screen of by FACS techniques to establish a range of different strength and sequence heterologous promoters.

The expansion of the fluorescence reporter suite would increase the number of reporters that can be used concurrently in any synthetic circuits. In this work we characterised seven reporters that can be divided into two compatibility groups (with separated excitation and emission spectra). The inclusion of a near infra-red protein such as TDsmURFP[360] and near UV protein such as Sirius[361] could expand the number of compatible reporter proteins to five in one strain.

The chemically inducible promoters analysed in this study were in the majority very poor. Only two (Ptet and Pphl) demonstrated dynamic ranges that are high enough for the use in constructing synthetic biology circuits. For future multi-input circuitry this library of inducible promoters must be expanded. One obvious choice for implementation in the rhizobacteria used in this thesis is that of the taurine inducible system regulated by the TauA activator. Previous studies have shown this system functioning well in rhizobia, therefore it is a strong candidate to use in rhizobacteria moving forwards[362]. The directed evolution pipeline used to construct the Marionette™ *Escherichia coli* strain[281] could be adapted and implemented in the two rhizobacterial species studied here to create libraries of regulatory proteins and output promoters with >100-fold dynamic ranges of induction.

With the RhizoKit toolbox many novel applications and genetic circuits can be constructed in the future. The construction of a chemotaxis circuit based on combining parts of the *che gene cluster* would enable rhizobacteria to migrate towards a plant specific signal. The characterisation of multiple fluorescent reporters could be used to interrogate the local and temporal changes of rhizobacterial gene whilst colonizing a target plant in real-time.

7.2.2 Computational logic in rhizobacteria

The characterisation of four orthogonal NOT gates in both rhizobacterial species was successfully demonstrated. The transcriptional repressors LacI, TetR, PhlF and ClR were combined with their cognate promoter pairs and Golden Gate cloning strategy to rapidly construct and demonstrate inducible NOT gate function and minimization of cross-talk between the gates. Although small in number, this core set of gates can be combined in the future to create more complex multi-input gates such as the AND and NAND gates[187,207]. By integrating AND and NAND gates with endogenous environmentally inducible inputs, such as oxygen concentration and unique plant signal molecules, the nitrogenase activity and ammonia assimilation activity of the cell can be regulated (via *glnA* and *nifA* outputs).

Further work to integrate the logic gates constructed here with the host cells own endogenous regulatory pathways can lead to novel and case-dependent synthetic circuitry and applications. Chimeric synthetic transcriptional repressors can be created due to the modular multi-domain nature of their structures. Work by Schumacher *et al.,* was able to demonstrate the combination of a sigma54 RNA polymerase activator input and output domains with the NtrC master nitrogen regulator to yield novel transcriptional activators integrated into the host cells metabolic regulatory network[102]. A similar approach could be used with the *A. caulinodans* ORS571 and *A. brasilense* FP2 strains used in this work to

integrate our characterised logic gate circuitry with the endogenous metabolic regulatory networks. Techniques such as RNAseq could then be employed to reveal the effect of interfacing synthetic logic gate circuitry on the global transcriptional profile of the cell.

7.2.3 Plant Growth Promotion

The engineering and demonstration of three PGP traits; auxin (IAA) biosynthesis, phosphate solubilisation and the antifungal DAPG biosynthesis was successful in both *A. caulinodans* ORS571 and *A. brasilense* FP2. By utilizing the two-gene pathway (IaaH & IaaM) from *Pseudomonas savastanoi,* the biosynthesis of significant amounts of IAA was observed under diazotrophic growth conditions. With the construction of a DAPG inducible IAA biosynthesis circuit, future experimentation can build on the ammonia excretion observed under the experimental conditions used here and improve on them. Induction timing, DAPG concentration and tuning of the PphI promoter and native RBS sequences could all increase DAPG synthesis *in vivo*[182].

Two plant growth promotion effects were observed in the rhizobox *Hordeum vulgare* plant growth experiments when inoculated with the constitutive auxin biosynthesis circuit pOPS0667 in both rhizobacterial strains. The fact that the same positive PGP effects, root hair elongation and increased net root length, is observed in the experiments with both strains, is valuable data proving the IAA

biosynthesis circuit may also function in other rhizobacterial species. Future studies using the pOPS0667 circuit in other rhizobacterial species that associate closely with other target crop plants may promote plant growth in those species[219,335]. With the proven PGP effects of IAA biosynthesis in associative rhizobacterial species confirmed, future work on including the system with rhizopine can be addressed. The construction of a rhizopine inducible IAA biosynthesis plasmid has been completed and ongoing work in the Poole Group at the University of Oxford will analyse its efficacy in the PGP of transgenic rhizopine producing *H. vulgare* lines. The possibility of using the DAPG inducible IAA biosynthesis circuit for PGP of target plants can also be a focus of further work if the DAPG biosynthesis cluster characterised in this work can be successfully transferred into target crop plants.

The phosphate solubilisation mechanism constructed and engineered in this work did prove to provide significant phosphate solubilisation efficacy on media plates when expressed in *Escherichia coli*. However, the circuit did not perform well in either of the two rhizobacterial species used in this work. As discussed in the chapter above, this may be due to many reasons including misfolding of the CitC and GltA genes, a lack of available metabolic capacity or translational issues due to rare codon usage in the two genes[363]. Future work to debug this circuit could focus on refactoring the codon usage or potentially look to scaffold the CitC and GltA proteins together near the cell membrane, causing enhanced flux and

potential metabolite channeling to enhance citric acid biosynthesis and efflux from the cell[364].

The final PGP trait engineered in this work was that of the DAPG biosynthesis circuit. The constitutive expression of DAPG was successfully demonstrated in both *E. coli* and *A. caulinodans* ORS571 in rich culture growth conditions. The *phl* operon clearly functions well in both strains but the robust DAPG production in the rhizobacterial strain allows for the strong basis for ongoing work with the circuit. Future studies could focus on confirming the DAPG produced from the biosynthesis circuit has active antifungal activity by performing growth inhibition assays on target fungal species such as *Penicillium digitatum* and *P. italicum*[258,280,304]. Once activity is confirmed, fungal protection assays with target crop plants such as *T. aestivum* can be performed to assay the engineered rhizobacterial efficacy in combating fungal pathogen infections. Directed evolution could be harnessed to enhance the DAPG antifungal pathway by putting competition pressure on strains growing in co-culture with target fungal species. This pressure would therefore allow for directed evolution towards enhanced DAPG and therefore antifungal activity *in vivo.*

7.2.4 CRISPRi in rhizobacteria

In this work the demonstration of the first synthetic CRISPRi system in rhizobacteria has been achieved. By utilizing the dCas9 wt, EQR and VRER variants of the Cas9 protein, we were able to demonstrate clear on-target gene repression in a dual plasmid system. There were clear differences in the dynamic ranges of repression between the dCas9 variants that were in keeping with previous studies utilizing the dCas9 system[330]. Possible reasons could include non-optimal DNA association of the REC domain of the dCas9 protein, as well as sgRNA association to the dCas9 variants differing in the cellular environment. The toxicity observed in both strains with dCas9 induction is common in studies where no sgRNA expression is present. Without a sgRNA molecule present within the cell, the dCas9 exhibits significantly more off-target binding to DNA sequences with homology to the sgRNA handle, therefore impacting endogenous gene expression to the detriment of the cell[139]. Future work using an inducible dCas9 system such as that implemented here could engineer tighter induction circuits whereby the dCas9 protein is not expressed until the pool of sgRNA molecules within the cell is at an optimum level[324]. An interesting alternative would be to build a pseudo sgRNA that does not target a rhizobial gene but is synthesised when the true sgRNA is not made. This would stabilise dCas9 and prevent toxicity. Upon induction by an appropriate signal the true guide sgRNA would be made and the pseudo sgRNA repressed.

The demonstration of the CRISPRi system regulating the *A. brasilense* FP2 glutamine synthetase gene (glnA) *in vivo* was successful and additionally produced a desired ammonia excretion phenotype. The release of fixed nitrogen by a diazotrophic rhizobacteria strain has only been achieved to date either by chemical inhibition of the enzymes involved in ammonia assimilation[365], or by extensive and time consuming genomic integration and gene expression regulation[366]. The ability to repress target genes in rhizobacteria, whereby causing PGP phenotypes such as ammonia release, is a novel achievement in the field and one that future work will develop and improve on. The CRISPRi system characterised in this work is compatible with multiplexing strategies[326] and therefore future work could focus on the optimisation of gene repression by multiple dCas9 binding events in the 5' and upstream regions of the *glnA* gene.

The potential for the dCas9 CRSIPRi system to regulate natural metabolic systems is vast. Synthetic sgRNAs can be rapidly synthesised, cloned and target both natural and synthetic DNA sequences. The integration of the entire CRISPRi genetic circuitry into a compact genomic construct would allow for allow for rapid CRISPRi experimentation and the paradigm of genomic mutagenesis and knock out techniques to be abolished. The CRISPRi circuitry characterised here can not only be used to regulate the host rhizobacteria's gene expression profiles, but also that of synthetic circuitry and constructs added subsequently by plasmid transformation or integration. Such a system would allow for a single-chemistry

system that controls a magnitude of gene expression profiles, all with one protein (dCas9) and a suite of sgRNAs under tightly regulated expression.

Other CRSIPRi systems such as those based on the Cpf1 system from the *Acidaminococcus* and *Lachnospiraceae* spp. can be harnessed to expand the PAM sequence targeting to more T-rich regions. The recent development of the adenine base editing (ABE) technique allows for the alteration of single base pairs without requiring double stranded DNA breaks. This system could be harnessed in the future to create reversible SNPs in target DNA sequences by the oscillatory control of the dCas9-adenosine deaminase and sgRNA machinery[367]

7.3 Final Conclusion

In this thesis a robust characterisation of the use of rhizobacteria as synthetic biology chassis has been achieved. By combining characterised DNA parts and transcriptional repressors, logical circuitry has been demonstrated in both *A. caulinodans* ORS571 and *A. brasilense* FP2 for the first time. The logic circuitry was able to demonstrate a unique plant signal inducible system in the strain *Sinorhizobium meliloti* by integrating the NOT gates with rhizopine activation machinery. Confirmed PGP traits such as auxin (IAA) biosynthesis, phosphate solubilisation and the antifungal compound DAPG production were demonstrated

in both rhizobacteria in novel synthetic circuits. Furthermore, the IAA biosynthesis circuit was used to demonstrate a robust PGP phenotype of Barley plant roots (*Hordeum vulgare*) in a competitive rhizosphere environment. This result paves the way for a rhizopine-induced IAA biosynthesis circuit that, while constructed, has yet to be verified under experimental conditions with transgenic rhizopine-producing *H. vulgare* plants. Demonstration of a fully customizable and modular CRISPRi system in both *A. caulinodans* ORS571 and *A. brasilense* FP2 we successful. This will enable future interrogation of any target gene within a rhizobacterial genome via transcriptional repression. The dCas9 wt CRISPRi system was implemented to target the endogenous glutamine synthetase gene (*glnA*) of *A. brasilense* FP2, thus yielding the desired ammonia excretion phenotype. Future tuning of the CRISPRi circuit and integration with the rhizopine signal from transgenic crop plants will allow for targeted ammonia release and thus PGP in competitive rhizosphere systems.

8 Bibliography

1. López-Bellido, L., López-Garrido, F. J., Fuentes, M., Castillo, J. E. & Fernández, E. J. Influence of tillage, crop rotation and nitrogen fertilization on soil organic matter and nitrogen under rain-fed Mediterranean conditions. *Soil Tillage Res.* **43,** 277–293 (1997).

2. Gan, Y. *et al.* Diversifying crop rotations with pulses enhances system productivity. *Sci. Rep.* **5,** 14625 (2015).

3. Smil, V. Improving Efficiency and Reducing Waste in Our Food System. *Environ. Sci.* **1,** 17–26 (2004).

4. Rockström, J. *et al.* A safe operating space for humanity. *Nature* **461,** 472–475 (2009).

5. Shcherbak, I., Millar, N. & Robertson, G. P. Global metaanalysis of the nonlinear response of soil nitrous oxide (N_2O) emissions to fertilizer nitrogen. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 9199–204 (2014).

6. Tilman, D., Balzer, C., Hill, J. & Befort, B. Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci.* **108,** 20260–20264 (2011).

7. Xu, G. H., Fan, X. & Miller, A. J. Plant Nitrogen Assimilation and Use Efficiency. *Annu. Rev. Plant Biol.* **63,** 153–182 (2012).

8. Dixon, R. & Kahn, D. Genetic regulation of biological nitrogen fixation. *Nat. Rev. Microbiol.* **2,** 621–631 (2004).

9. Modak, J. M. Haber Process for Ammonia Synthesis Jayant M Modak. *Resonance* **7,** 69–77 (2002).

10. Beatty, P. H. & Good, A. G. Plant science. Future prospects for cereals that fix nitrogen. *Science* **333,** 416–7 (2011).

11. Vessey, J. K. Plant growth promoting rhizobacteria as biofertilizers. *Plant Soil* **255,** 571–586 (2003).

12. Terpolilli, J. J., Hood, G. A. & Poole, P. S. *What determines the efficiency of N(2)-fixing Rhizobium-legume symbioses? Advances in microbial*

physiology **60,** (Elsevier Ltd., 2012).

13. Werner, G. D. A., Cornwell, W. K., Sprent, J. I., Kattge, J. & Kiers, E. T. A single evolutionary innovation drives the deep evolution of symbiotic N2-fixation in angiosperms. *Nat. Commun.* **5,** 4087 (2014).

14. Vitousek, P. M., Menge, D. N. L., Reed, S. C. & Cleveland, C. C. Biological nitrogen fixation: rates, patterns and ecological controls in terrestrial ecosystems. *Philos. Trans. R. Soc. B Biol. Sci.* **368,** (2013).

15. Birch, H. Gates pours cash into agriculture. *Nat. Biotechnol.* **27,** 1064–1064 (2009).

16. Montañez, A. Overview and Case studies on Biological Nitrogen Fixation: Perspectives and Limitations. *FAO* (2000).

17. Genre, A. & Russo, G. Does a Common Pathway Transduce Symbiotic Signals in Plant–Microbe Interactions? *Front. Plant Sci.* **7,** 96 (2016).

18. Oldroyd, G. E. D. Speak, friend and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nat. Rev. Microbiol.* **11,** 252–263 (2013).

19. Madsen, L. H. *et al.* The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. *Nat. Commun.* **1,** 1–12 (2010).

20. Hoffman, B. M., Lukoyanov, D., Yang, Z.-Y., Dean, D. R. & Seefeldt, L. C. Mechanism of Nitrogen Fixation by Nitrogenase: The Next Stage. *Chem. Rev.* **114,** 4041–4062 (2014).

21. López-Torrejón, G. *et al.* Expression of a functional oxygen-labile nitrogenase component in the mitochondrial matrix of aerobically grown yeast. *Nat. Commun.* **7,** 11426 (2016).

22. Geddes, B. A. *et al.* Use of plant colonizing bacteria as chassis for transfer of N 2 -fixation to cereals. *Curr. Opin. Biotechnol.* **32,** 216–222 (2015).

23. Rogers, C. & Oldroyd, G. E. D. Synthetic biology approaches to engineering the nitrogen symbiosis in cereals. *J. Exp. Bot.* **65,** 1939–1946 (2014).

24. Setten, L. *et al.* Engineering Pseudomonas protegens Pf-5 for Nitrogen Fixation and its Application to Improve Plant Growth under Nitrogen-Deficient Conditions. *PLoS One* **8,** (2013).

25. Iniguez, A. L. *et al.* Regulation of Enteric Endophytic Bacterial Colonization by Plant Defenses. *Mol. Plant-Microbe Interact.* **2**, 169-78 (2005).

26. Cummings, S. P. *et al.* Nodulation of Sesbania species by *Rhizobium* (Agrobacterium) strain IRBG74 and other rhizobia. *Environ. Microbiol.* **10**, 2510-25 (2009).

27. Vacheron, J. *et al.* Plant growth-promoting rhizobacteria and root system functioning. *Front. Plant Sci.* **4,** 356 (2013).

28. Mitra, S. Interaction of *Rhizobium* Sp. Strain IRBG74 with a Legume (*Sesbania Cannabina*) and a Cereal (*Oryza Sativa*). *Theses Diss.* (2014).

29. Gopalaswamy, G., Kannaiyan, S., O'Callaghan, K. J., Davey, M. R. & Cocking, E. C. The xylem of rice (*Oryza sativa*) is colonized by *Azorhizobium caulinodans*. *Proc. R. Soc. B Biol. Sci.* **268**, 103-107 (2000).

30. Cocking, E. C. & Cocking, E. C. Endophytic colonization of plant roots by nitrogen-fixing bacteria. *Plant Soil* **233,** 169–175 (2003).

31. Sabry, S. R. *et al.* Endophytic establishment of *Azorhizobium caulinodans* in wheat. *Proc. R. Soc. B Biol. Sci.* **264,** 341–346 (1997).

32. Michel-Reydellet, N. & Kaminski, P. A. *Azorhizobium caulinodans* PII and GlnK proteins control nitrogen fixation and ammonia assimilation. *J. Bacteriol.* **181,** 2655–8 (1999).

33. Lee, K.-B. *et al.* The genome of the versatile nitrogen fixer *Azorhizobium caulinodans* ORS571. *BMC Genomics* **9,** 271 (2008).

34. Michiels, K., Vanderleyden, J. & Van Gool, A. *Azospirillum* — plant root associations: A review. *Biol. Fertil. Soils* **8,** 356–368 (1989).

35. Rothballer, M., Schmid, M. & Hartmann, A. In situ localization and PGPR-effect of *Azospirillum brasilense* strains colonizing roots of different wheat varieties. *Symbiosis* **34,** 261–279 (2003).

36. Bashan, Y., Holguin, G. & de-Bashan, L. E. *Azospirillum* -plant relationships: physiological, molecular, agricultural, and environmental advances (1997-2003). *Can. J. Microbiol.* **50,** 521–577 (2004).

37. Baudoin, E. *et al.* Effects of *Azospirillum brasilense* with genetically modified auxin biosynthesis gene ipdC upon the diversity of the indigenous microbiota of the wheat rhizosphere. *Res. Microbiol.* **161,** 219–226 (2010).

38. Cassán, F. D., Okon, Y., Creus, C. M. & Cassa□n, F. D. *Handbook for azospirillum : technical issues and protocols*. (2015).

39. Santos, K. F. D. N. *et al.* Wheat colonization by an *Azospirillum brasilense* ammonium-excreting strain reveals upregulation of nitrogenase and

superior plant growth promotion. *Plant Soil* **415,** 245–255 (2017).

40. Morandini, P., Vignola, R. & Vittorioso, P. in *Plant Nutrition — Physiology and Applications* 127–131 (Springer Netherlands, 1990).

41. Steenhoudt, O. *et al. Azospirillum*, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects. *FEMS Microbiol. Rev.* **24,** 487–506 (2000).

42. Rivera, D. *et al.* Complete Genome Sequence of the Model Rhizosphere Strain *Azospirillum brasilense* Az39, Successfully Applied in Agriculture. *Genome Announc.* **2,** 683-714 (2014).

43. Machado, H. B., Funayama, S., Rigo, L. U. & Pedrosa, F. O. Excretion of ammonium by *Azospirillum brasilense* mutants resistant to ethylenediamine. *Can. J. Microbiol.* **37,** 549–553 (1991).

44. Santos, A. R. S. *et al.* Labeled *Azospirillum brasilense* wild type and excretion-ammonium strains in association with barley roots. *Plant Physiol. Biochem.* **118,** 422–426 (2017).

45. Pankievicz, V. C. S. *et al.* Robust biological nitrogen fixation in a model grass-bacterial association. *Plant J.* **81,** 907–19 (2015).

46. Chopra, P. & Kamma, A. Engineering life through Synthetic Biology. *In Silico Biol.* **6,** 401–10 (2006).

47. Bortesi, L. & Fischer, R. The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnol. Adv.* **33,** 41–52 (2015).

48. Kang, X., He, W., Huang, Y., Yu, Q. & Chen, Y. Introducing precise genetic modifications into human 3PN embryos by CRISPR / Cas-mediated genome editing. *J. Assist. Reprod. Genet.* **33**, 581-588 (2016).

49. Smanski, M. J. *et al.* Synthetic biology to access and expand nature's chemical diversity. **14**, 135-149 *Nature Reviews Microbiology* (2016).

50. Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. **27**, 946-950 *Nat. Biotechnol.* (2009).

51. Guan, Z., Schmidt, M., Pei, L., Wei, W. & Ma, K. Biosafety Considerations of Synthetic Biology in the International Genetically Engineered Machine (iGEM) Competition. *Bioscience* **63,** 25–34 (2013).

52. Kosuri, S. *et al.* Composability of regulatory sequences controlling transcription and translation in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S.*

A. **110,** 14024–9 (2013).

53. Lou, C., Stanton, B., Chen, Y. J., Munsky, B. & Voigt, C. A. Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat. Biotechnol.* **30**, 1137-42 (2012).

54. Nielsen, A. a K. *et al.* Genetic circuit design automation. *Science* **352**, (2016).

55. Chen, Y. J. *et al.* Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat. Methods* **10**, 659-64 (2013).

56. Davis, J. H., Rubin, A. J. & Sauer, R. T. Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res.* **39**, 1131-41 (2011).

57. Hughes, R. A. & Ellington, A. D. Synthetic DNA Synthesis and Assembly: Putting the Synthetic in Synthetic Biology. *Cold Spring Harb. Perspect. Biol.* **9** (2017).

58. Baker, B. Synthetic Biology and the Marketplace. *Bioscience* **67,** 877–883 (2017).

59. Bayer, T. S. *et al.* Synthesis of methyl halides from biomass using engineered microbes. *J. Am. Chem. Soc.* **131,** 6508–6515 (2009).

60. Czar, M. J., Anderson, J. C., Bader, J. S. & Peccoud, J. Gene synthesis demystified. *Trends Biotechnol.* **27,** 63–72 (2009).

61. Kosuri, S. & Church, G. M. Large-scale de novo DNA synthesis: technologies and applications. *Nat. Methods* **11,** 499–507 (2014).

62. Elowitz, M. B. & Leibier, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).

63. Schaumberg, K. A. *et al.* Quantitative characterization of genetic parts and circuits for plant synthetic biology. *Nat. Methods* **13**, 94-100 (2015).

64. Endy, D. Foundations for engineering biology. *Nature* **438**, 449–453 (2005).

65. Hawkins, J. S., Wong, S., Peters, J. M., Almeida, R. & Qi, L. S. Targeted Transcriptional Repression in Bacteria Using CRISPR Interference (CRISPRi). *Methods Mol. Biol.* **1311,** 349–62 (2015).

66. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152,** 1173–1183

(2013).

67. Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science (80-.).* **346,** 1258096–1258096 (2014).

68. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. a. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* **31,** 233–239 (2013).

69. Sapranauskas, R. *et al.* The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* **39,** 9275–9282 (2011).

70. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155,** 733–740 (2009).

71. Larson, M. H. *et al.* CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat. Protoc.* **8,** 2180–2196 (2013).

72. Pankowicz, F. P. *et al.* Reprogramming metabolic pathways in vivo with CRISPR/Cas9 genome editing to treat hereditary tyrosinaemia. *Nat. Commun.* **7,** 12642 (2016).

73. Shevidi, S., Uchida, A., Schudrowitz, N., Wessel, G. M. & Yajima, M. Single nucleotide editing without DNA cleavage using CRISPR/Cas9-deaminase in the sea urchin embryo. *Dev. Dyn.* **246,** 1036–1046 (2017).

74. Dominguez, A. A., Lim, W. A. & Qi, L. S. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat. Rev. Mol. Cell Biol.* **17,** 5–15 (2016).

75. Bikard, D. *et al.* Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* **41,** 7429–7437 (2013).

76. Ferreira, R., Skrekas, C., Nielsen, J. & David, F. Multiplexed CRISPR/Cas9 Genome Editing and Gene Regulation Using Csy4 in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **7,** 10–15 (2018).

77. Pham, H., Kearns, N. A. & Maehr, R. in 43–57 (Humana Press, New York, NY, 2016).

78. Inomura, K., Bragg, J. & Follows, M. J. A quantitative analysis of the direct and indirect costs of nitrogen fixation: a model based on *Azotobacter vinelandii*. *ISME J.* **11,** 166–175 (2017).

79. Resendis-Antonio, O., Reed, J. L., Encarnación, S., Collado-Vides, J. & Palsson, B. Ø. Metabolic Reconstruction and Modeling of Nitrogen Fixation in *Rhizobium etli*. *PLoS Comput. Biol.* **3,** e192 (2007).

80. Temme, K., Zhao, D. & Voigt, C. A. Refactoring the nitrogen fixation gene cluster from *Klebsiella oxytoca*. *Proc. Natl. Acad. Sci.* **109**, 7085-90 (2012).

81. Triplett, E. W. Diazotrophic endophytes: Progress and prospects for nitrogen fixation in monocots. in *Plant and Soi* **186**, 29-38 (1996).

82. Bali, A., Blanco, G., Hill, S. & Kennedy, C. Excretion of ammonium by a nifL mutant of *Azotobacter vinelandii* fixing nitrogen. *Appl. Environ. Microbiol.* **58**, 1711-18 (1992).

83. Brewin, B., Woodley, P. & Drummond, M. The basis of ammonium release in nifL mutants of *Azotobacter vinelandii*. *J. Bacteriol.* **181** 7356-7362 (1999).

84. Van Dommelen, A., Croonenborghs, A., Spaepen, S. & Vanderleyden, J. Wheat growth promotion through inoculation with an ammonium-excreting mutant of *Azospirillum brasilense*. *Biol. Fertil. Soils* **45,** 549-533 (2009).

85. Van Dommelen, A., Keijers, V., Wollebrants, A. & Vanderleyden, J. Phenotypic changes resulting from distinct point mutations in the *Azospirillum brasilense glnA* gene, encoding glutamine synthetase. *Appl. Environ. Microbiol.* **69,** 5699–5701 (2003).

86. Zhang, T. *et al.* Involvement of the ammonium transporter AmtB in nitrogenase regulation and ammonium excretion in *Pseudomonas stutzeri* A1501. *Res. Microbiol.* **163,** 332-9 (2012).

87. Conroy, M. J. *et al.* The crystal structure of the *Escherichia coli* AmtB-GlnK complex reveals how GlnK regulates the ammonia channel. *Proc. Natl. Acad. Sci.* **104,** 1213-1218 (2007).

88. Mendoza, A., Leija, A., Martinez-Romero, E., Hernandez, G. & Mora, J. The enhancement of ammonium assimilation in *Rhizobium etli* prevents nodulation of Phaseolus vulgaris. *Mol Plant Microbe Interact* **8,** 584-92 (1995).

89. Oteino, N. *et al.* Plant growth promotion induced by phosphate solubilizing endophytic Pseudomonas isolates. *Front. Microbiol.* **6,** 1–9 (2015).

90. Glick, B. R. *et al.* Promotion of Plant Growth by Bacterial ACC Deaminase. *CRC. Crit. Rev. Plant Sci.* **26,** 227–242 (2007).

91. Goswami, D., Thakker, J. N. & Dhandhukia, P. C. Portraying mechanics of plant growth promoting rhizobacteria (PGPR): A review. *Cogent Food Agric.* **2** (2016).

92. Jha, C. K., Patel, D., Rajendran, N. & Saraf, M. Combinatorial assessment on dominance and informative diversity of PGPR from rhizosphere of *Jatropha curcas. Basic Microbiol.* **50,** 211–217 (2010).

93. Goswami, D., Thakker, J. N. & Dhandhukia, P. C. Portraying mechanics of plant growth promoting rhizobacteria (PGPR): A review. *Cogent Food Agric.* **2** (2016).

94. Spaepen, S., Vanderleyden, J. & Remans, R. Indole-3-acetic acid in microbial and microorganism-plant signaling. *FEMS Microbiol. Rev.* **31,** 425–448 (2007).

95. Couillerot, O., Prigent-Combaret, C., Caballero-Mellado, J. & Moënne-Loccoz, Y. *Pseudomonas fluorescens* and closely-related fluorescent pseudomonads as biocontrol agents of soil-borne phytopathogens. *Lett. Appl. Microbiol.* **48,** 505–512 (2009).

96. Iavicoli, A., Boutet, E., Buchala, A. & Métraux, J.-P. Induced Systemic Resistance in *Arabidopsis thaliana* in Response to Root Inoculation with *Pseudomonas fluorescens* CHA0. *Mol. Plant-Microbe Interact.* **16,** 851–858 (2003).

97. Phillips, D. A., Fox, T. C., King, M. D., Bhuvaneswari, T. V & Teuber, L. R. Microbial Products Trigger Amino Acid Exudation from Plant Roots. *PLANT Physiol.* **136,** 2887–2894 (2004).

98. Brazelton, J. N., Pfeufer, E. E., Sweat, T. A., Gardener, B. B. M. & Coenen, C. 2,4-Diacetylphloroglucinol Alters Plant Root Development. *Mol. Plant-Microbe Interact.* **21,** 1349–1358 (2008).

99. Khalil, A. S. & Collins, J. J. Synthetic biology: Applications come of age. *Nature Reviews Genetics* **11,** 367-79 (2010).

100. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403,** 335–338 (2000).

101. Morey, K. J. *et al.* Crosstalk between endogenous and synthetic components - synthetic signaling meets endogenous components. *Biotechnology Journal* **7,** 846-55 (2012).

102. Schumacher, J., Wang, B., Bonatto, A. & Buck, M. Synthetic transcription factors allow regulon wide control and shifting the nitrogen/carbon balance

in bacteria. *N. Biotechnol.* **31,** (2014).

103. Venturi, V. & Keel, C. Signaling in the Rhizosphere. **21,** 187-98 *Trends in Plant Science* (2016).

104. Lareen, A., Burton, F. & Schäfer, P. Plant root-microbe communication in shaping root microbiomes. *Plant Molecular Biology* **90,** 575-87 (2016).

105. Berendsen, R. L., Pieterse, C. M. J. & Bakker, P. A. H. M. The rhizosphere microbiome and plant health. *Trends in Plant Science* **8,** 478-486 (2012).

106. Mendes, R. *et al.* Deciphering the rhizosphere microbiome for disease-suppressive bacteria. *Science* **322,** 1097-100 (2011).

107. Murphy, P. J., Wexler, W., Grzemski, W., Rao, J. P. & Gordon, D. Rhizopines-Their role in symbiosis and competition. *Soil Biol. Biochem.* **4-5,** 525-529 (1995).

108. Wexler, M., Gordon, D. & Murphy, P. J. The distribution of inositol rhizopine genes in Rhizobium populations. *Soil Biol. Biochem.* **4-5,** 531-537 (1995).

109. Murphy, P. J. *et al.* Genes for the catabolism and synthesis of an opine-like compound in *Rhizobium meliloti* are closely linked and on the Sym plasmid. *Proc. Natl. Acad. Sci. U. S. A.* **84,** 493-7 (1987).

110. Bahar, M. *et al.* A model for the catabolism of rhizopine in *Rhizobium leguminosarum* involves a ferredoxin oxygenase complex and the inositol degradative pathway. *Mol. Plant. Microbe. Interact.* **11,** 1057-68 (1998).

111. Beringer, J. E. R Factor Transfer in Bacterial distribution in the rhizosphere of wild barley under contrasting microclimates. *Microbiology* **84,** 188–198 (1974).

112. Atlas, R. M. *Handbook of microbiological media*. (Boca Raton, Fla., 2010).

113. Weber, E., Engler, C., Gruetzner, R., Werner, S. & Marillonnet, S. A Modular Cloning System for Standardized Assembly of Multigene Constructs. *PLoS One* **6,** e16765 (2011).

114. Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5,** 113 (2004).

115. de Jong, A., Pietersma, H., Cordes, M., Kuipers, O. P. & Kok, J. PePPER: a webserver for prediction of prokaryote promoter elements and regulons. *BMC Genomics* **13,** 299 (2012).

116. Ditta, G., Stanfield, S., Corbin, D. & Helinski, D. R. Broad host range DNA

cloning system for gram-negative bacteria: construction of a gene bank of Rhizobium meliloti. *Proc. Natl. Acad. Sci. U. S. A.* **77,** 7347–51 (1980).

117. Subach, O. M., Cranfill, P. J., Davidson, M. W. & Verkhusha, V. V. An Enhanced Monomeric Blue Fluorescent Protein with the High Chemical Stability of the Chromophore. **6,** e28674 (2011).

118. Rizzo, M. A., Springer, G. H., Granada, B. & Piston, D. W. An improved cyan fluorescent protein variant useful for FRET. *Nat. Biotechnol.* **22,** 445–449 (2004).

119. Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T. C. & Waldo, G. S. Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* **24,** 79–88 (2006).

120. Gert-Jan Kremers, Joachim Goedhart, Erik B. van Munster, and & Theodorus W. J. Gadella, J. Cyan and Yellow Super Fluorescent Proteins with Improved Brightness, Protein Folding, and FRET Förster Radius. **45,** 6570-80 (2006).

121. Shaner, N. C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat. Biotechnol.* **22,** 1567–1572 (2004).

122. Bajar, B. T. *et al.* Improving brightness and photostability of green and red fluorescent proteins for live cell imaging and FRET reporting. *Sci. Rep.* **6,** 20889 (2016).

123. Hardy, R. W., Holsten, R. D., Jackson, E. K. & Burns, R. C. The acetylene-ethylene assay for N_2 fixation: laboratory and field evaluation. *Plant Physiol.* **43,** 1185–207 (1968).

124. Frederix, M. *et al.* Mutation of praR in Rhizobium leguminosarum enhances root biofilms, improving nodulation competitiveness by increased expression of attachment proteins. *Mol. Microbiol.* **93,** 464–78 (2014).

125. Moore, S. J. *et al.* EcoFlex: A Multifunctional MoClo Kit for E. coli Synthetic Biology. *ACS Synth. Biol.* **5,** 1059–1069 (2016).

126. Tsai, C. S., Kwak, S., Turner, T. L. & Jin, Y. S. Yeast synthetic biology toolbox and applications for biofuel production. *FEMS Yeast Research* **15** (2015).

127. Engler, C. *et al.* A Golden Gate Modular Cloning Toolbox for Plants. *ACS Synth. Biol.* **3,** 839–843 (2014).

128. Keasling, J. D. Synthetic biology for synthetic chemistry. *ACS Chemical Biology* **3,** 64–76 (2008).

129. Fernandez-Rodriguez, J., Moser, F., Song, M. & Voigt, C. A. Engineering RGB color vision into *Escherichia coli*. *Nat. Chem. Biol.* **13,** 706–708 (2017).

130. Nielsen, A. A. K. & Voigt, C. A. Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. *Mol. Syst. Biol.* **10,** 763 (2014).

131. Tay, P. K. R., Nguyen, P. Q. & Joshi, N. S. A Synthetic Circuit for Mercury Bioremediation Using Self-Assembling Functional Amyloids. *ACS Synth. Biol.* **6,** 1841–1850 (2017).

132. Mus, F. *et al.* Symbiotic nitrogen fixation and the challenges to its extension to non-legumes. *Applied and Environmental Microbiology* **82,** 3698–3710 (2016).

133. Yamazaki, K., de Mora, K. & Saitoh, K. BioBrick-based 'Quick Gene Assembly' in vitro. *Synth. Biol.* **2,** (2017).

134. Engler, C., Gruetzner, R., Kandzia, R. & Marillonnet, S. Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes. *PLoS One* **4,** e5553 (2009).

135. Weber, E., Gruetzner, R., Werner, S., Engler, C. & Marillonnet, S. Assembly of Designer TAL Effectors by Golden Gate Cloning. *PLoS One* **6,** e19722 (2011).

136. Sarrion-Perdigones, A. *et al.* GoldenBraid: An Iterative Cloning System for Standardized Assembly of Reusable Genetic Modules. *PLoS One* **6,** e21622 (2011).

137. Werner, S., Engler, C., Weber, E., Gruetzner, R. & Marillonnet, S. Fast track assembly of multigene constructs using Golden Gate cloning and the MoClo system. *Bioengineered* **3,** 38–43 (2012).

138. Casini, A., Storch, M., Baldwin, G. S. & Ellis, T. Bricks and blueprints: methods and standards for DNA assembly. *Nat. Rev. Mol. Cell Biol.* **16,** 568–576 (2015).

139. Nielsen, A. A. & Voigt, C. a. Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. *Mol. Syst. Biol.* **10,** 763 (2014).

140. Temme, K., Hill, R., Segall-Shapiro, T. H., Moser, F. & Voigt, C. A. Modular control of multiple pathways using engineered orthogonal T7 polymerases. *Nucleic Acids Res.* **40,** 8773–81 (2012).

141. Lou, C., Stanton, B., Chen, Y.-J., Munsky, B. & Voigt, C. A. Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat. Biotechnol.* **30,** 1137–42 (2012).

142. Casini, A. *et al.* R2oDNA Designer: Computational Design of Biologically Neutral Synthetic DNA Sequences. *ACS Synth. Biol.* **3,** 525–528 (2014).

143. De Souza, J. *et al.* Effect of 2,4-diacetylphloroglucinol on pythium: cellular responses and variation in sensitivity among propagules and species. *Phytopathology* **93,** 966–975 (2003).

144. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 343–345 (2009).

145. Harley, C. B. & Reynolds, R. P. Analysis of *E. coli* promoter sequences. *Nucleic Acids Res.* **15,** 2343–61 (1987).

146. Goldstein, M. A. & Doi, R. H. Prokaryotic promoters in biotechnology. *Biotechnol. Annu. Rev.* **1,** 105–28 (1995).

147. Wösten, M. M. Eubacterial sigma-factors. *FEMS Microbiol. Rev.* **22,** 127–50 (1998).

148. Holguin, G., Patten, C. L. & Glick, B. R. Genetics and molecular biology of Azospirillum. *Biol. Fertil. Soils* **29,** 10–23 (1999).

149. Kelly, J. R. *et al.* Measuring the activity of BioBrick promoters using an in vivo reference standard. *J. Biol. Eng.* **3,** 4 (2009).

150. Shine, J. & Dalgarno, L. Determinant of cistron specificity in bacterial ribosomes. *Nature* **254,** 34–8 (1975).

151. Rangwala, S. H., Fuchs, R. L., Drahos, D. J. & Olins, P. O. Broad Host-Range Vector for Efficient Expression of Foreign Genes in Gram-Negative Bacteria. *Nat. Biotechnol.* **9,** 477–479 (1991).

152. Afzal, M., Shahid, A. A., Shehzadi, A., Nadeem, S. & Husnain, T. RDNAnalyzer: A tool for DNA secondary structure prediction and sequence analysis. *Bioinformation* **8,** 687–90 (2012).

153. Von Hippel, P. H. An integrated model of the transcription complex in elongation, termination, and editing. *Science* **281,** 660-5 (1998).

154. Keiler, K. C., Waller, P. R. & Sauer, R. T. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271,** 990–3 (1996).

155. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215,** 403–410 (1990).

156. Studier, F. W. & Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189,** 113–30 (1986).

157. Sousa, R. & Mukherjee, S. T7 RNA polymerase. *Prog. Nucleic Acid Res. Mol. Biol.* **73,** 1–41 (2003).

158. Jares-Erijman, E. A. & Jovin, T. M. FRET imaging. *Nat. Biotechnol.* **21,** 1387–1395 (2003).

159. Subach, O. M., Cranfill, P. J., Davidson, M. W. & Verkhusha, V. V. An Enhanced Monomeric Blue Fluorescent Protein with the High Chemical Stability of the Chromophore. *PLoS One* **6,** e28674 (2011).

160. Hebisch, E., Knebel, J., Landsberg, J., Frey, E. & Leisner, M. High Variation of Fluorescence Protein Maturation Times in Closely Related *Escherichia coli* Strains. *PLoS One* **8,** e75991 (2013).

161. Kushwaha, M. & Salis, H. M. A portable expression resource for engineering cross-species genetic circuits and pathways. *Nat. Commun.* **6,** 7832 (2015).

162. Lin, M.-T. *et al.* Novel Utilization of Terminators in the Design of Biologically Adjustable Synthetic Filters. *ACS Synth. Biol.* **5,** 365–374 (2016).

163. Tran, H., Oliveira, S. M. D., Goncalves, N. & Ribeiro, A. S. Kinetics of the cellular intake of a gene expression inducer at high concentrations. *Mol. BioSyst* **11,** 2579 (2015).

164. Fulcrand, G. *et al.* DNA supercoiling, a critical signal regulating the basal expression of the lac operon in *Escherichia coli*. *Sci. Rep.* **6,** 19243 (2016).

165. Ryan, P. R., Dessaux, Y., Thomashow, L. S. & Weller, D. M. Rhizosphere engineering and management for sustainable agriculture. *Plant and Soil* **321,** 363-383 (2009).

166. Ren, G., Ma, A., Liu, W., Zhuang, X. & Zhuang, G. Bacterial signals N-acyl homoserine lactones induce the changes of morphology and ethanol tolerance in *Saccharomyces cerevisiae*. *AMB Express* **6,** 117 (2016).

167. Houdt, R., Aertsen, A., Moons, P., Vanoirbeek, K. & Michiels, C. W. *N* -acyl-l-homoserine lactone signal interception by *Escherichia coli*. *FEMS Microbiol. Lett.* **256,** 83–89 (2006).

168. Lee, S. J. *et al.* Genes encoding the N-acyl homoserine lactone-degrading

enzyme are widespread in many subspecies of *Bacillus thuringiensis*. *Appl. Environ. Microbiol.* **68,** 3919–24 (2002).

169. Kosibat, B. E. & Chleif, R. Arabinose-inducible Promoter from *Escherichia coli* Its Cloning from Chromosomal DNA ,. *J. Mol. Biol.* **156,** 53–66 (1982).

170. Chopra, I. & Roberts, M. Tetracycline Antibiotics: Mode of Action, Applications, Molecular Biology, and Epidemiology of Bacterial Resistance. *Microbiol. Mol. Biol. Rev.* **65,** 232-260 (2001).

171. Stanton, B. C. *et al.* Systematic transfer of prokaryotic sensors and circuits to mammalian cells. *ACS Synth. Biol.* **3,** 880–891 (2014).

172. Cameron, D. E. & Collins, J. J. Tunable protein degradation in bacteria. *Nat. Biotechnol.* **32,** 1–8 (2014).

173. Farrell, C. M., Grossman, A. D. & Sauer, R. T. Cytoplasmic degradation of ssrA-tagged proteins. *Mol. Microbiol.* **57,** 1750–1761 (2005).

174. Gur, E. & Sauer, R. T. Evolution of the ssrA degradation tag in Mycoplasma: Specificity switch to a different protease. *Proc. Natl. Acad. Sci.* **105,** 16113–16118 (2008).

175. Barbier, M. & Damron, F. H. Rainbow Vectors for Broad-Range Bacterial Fluorescence Labeling. *PLoS One* **11,** e0146827 (2016).

176. Bangera, M. G. & Thomashow, L. S. Identification and characterization of a gene cluster for synthesis of the polyketide antibiotic 2,4-diacetylphloroglucinol from *Pseudomonas fluorescens*. *J. Bacteriol.* **181,** 3155-63 (1999).

177. Endy, D. Foundations for engineering biology. *Nature* **438,** 449–453 (2005).

178. Dueber, J. E. *et al.* Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* **27,** 753–759 (2009).

179. Win, M. N. & Smolke, C. D. Higher-order cellular information processing with synthetic RNA devices. *Science* **322,** 456–60 (2008).

180. Collins, J. J., Gardner, T. S. & Cantor, C. R. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403,** 339–342 (2000).

181. Na, D. *et al.* Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. *Nat. Biotechnol.* **31,** 170–174 (2013).

182. Arpino, J. A. J. *et al.* Tuning the dials of Synthetic Biology. *Microbiology* **159,** 1236–53 (2013).

183. Sowa, S. W., Baldea, M. & Contreras, L. M. Optimizing Metabolite Production Using Periodic Oscillations. *PLoS Comput. Biol.* **10,** e1003658 (2014).

184. Prindle, A. *et al.* Rapid and tunable post-translational coupling of genetic circuits. *Nature* **508,** 387–391 (2014).

185. Jaeger. in *Microelectronic circuit design* 226–233 (McGraw-Hill, 1997).

186. Tinder, R. in *Engineering digital design. Revised second edition* 317–31 (2000).

187. Stanton, B. C. *et al.* Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nat. Chem. Biol.* **10,** 99–105 (2014).

188. Liu, Y. *et al.* Synthesizing AND gate genetic circuits based on CRISPR-Cas9 for identification of bladder cancer cells. *Nat. Commun.* **5,** 5393 (2014).

189. Liu, T., Chen, J. Y., Zheng, Z., Wang, T. H. & Chen, G. Q. Construction of highly efficient *E. coli* expression systems containing low oxygen induced promoter and partition region. *Appl. Microbiol. Biotechnol.* **68,** 346–354 (2005).

190. Polstein, L. R. & Gersbach, C. A. A light-inducible {CRISPR-Cas9} system for control of endogenous gene activation. *Nat Chem Biol* **11,** 198–200 (2015).

191. Miyamoto, T., Razavi, S., DeRose, R. & Inoue, T. Synthesizing Biomolecule-Based Boolean Logic Gates. *ACS Synth. Biol.* **2,** 72–82 (2013).

192. Brophy, J. A. N. & Voigt, C. A. Principles of genetic circuit design. *Nat. Methods* **11,** 508–520 (2014).

193. Marchisio, M. A. In silico design and in vivo implementation of yeast gene Boolean gates. *J. Biol. Eng.* **8,** 6 (2014).

194. Kramer, B. P., Fischer, C. & Fussenegger, M. BioLogic gates enable logical transcription control in mammalian cells. *Biotechnol. Bioeng.* **87,** 478–484 (2004).

195. Friedland, A. E. *et al.* Synthetic Gene Networks That Count. *Science (80-.).* **324,** 1199–1202 (2009).

196. Becker, N. A., Peters, J. P. & Maher, L. J. Mechanism of promoter repression by Lac repressor-DNA loops. *Nucleic Acids Res.* **41,** 156–166 (2013).

197. Pedersen, M., Ligowska, M. & Hammer, K. Characterization of the CI repressor protein encoded by the temperate lactococcal phage TP901-1. *J. Bacteriol.* **192,** 2102–10 (2010).

198. Ramos, J. L. *et al.* The TetR Family of Transcriptional Repressors. *Microbiol. Mol. Biol. Rev.* **69,** 326-356 (2005).

199. Saenger, W., Hinrichs, W., Orth, P., Schnappinger, D. & Hillen, W. Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. *Nat. Struct. Biol.* **7,** 215–219 (2000).

200. Sheehan, M. M. *et al.* Regulation of production of the antifungal metabolite 2,4-diacetylphloroglucinol in *Pseudomonas fluorescens* F113: genetic analysis of phlF as a transcriptional repressor. *Microbiology* **146,** 537–546 (2000).

201. Abbas, A. *et al.* Characterization of interactions between the transcriptional repressor PhlF and its binding site at the phlA promoter in *Pseudomonas fluorescens* F113. *J. Bacteriol.* **184,** 3008–16 (2002).

202. Nielsen, A. a K. *et al.* Genetic circuit design automation. *Science* **352,** aac7341 (2016).

203. Mueller, N. D. *et al.* Closing yield gaps through nutrient and water management. *Nature* **490,** 254–257 (2012).

204. Lassaletta, L., Billen, G., Grizzetti, B., Anglade, J. & Garnier, J. 50 year trends in nitrogen use efficiency of world cropping systems: the relationship between yield and nitrogen input to cropland. *Environ. Res. Lett.* **9,** 105011 (2014).

205. Tilman, D., Balzer, C., Hill, J. & Befort, B. L. Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 20260-20264 (2011).

206. Ewen Cameron, D. & Collins, J. J. Tunable protein degradation in bacteria. **32,** 1276-1281 *Nat. Biotechnol.* (2014).

207. Reynolds, N., O'Shaughnessy, A. & Hendrich, B. Transcriptional repressors: multifaceted regulators of gene expression. **140,** 505-512 *Development* (2013).

208. de Jong, H., Geiselmann, J. & Ropers, D. Resource Reallocation in Bacteria by Reengineering the Gene Expression Machinery. *Trends in Microbiology* **25,** 480-493 (2017).

209. Frumkin, I. *et al.* Gene Architectures that Minimize Cost of Gene Expression. *Mol. Cell* **65,** 142-53 (2017).

210. Smanski, M. J. *et al.* Functional optimization of gene clusters by combinatorial design and assembly. *Nat. Biotechnol.* **32,** 1241–1249 (2014).

211. Sharp, P. M., Bailes, E., Grocock, R. J., Peden, J. F. & Sockett, R. E. Variation in the strength of selected codon usage bias among bacteria. *Nucleic Acids Res.* **23,** 1141-53 (2005).

212. Atkinson, M. R., Savageau, M. A., Myers, J. T. & Ninfa, A. J. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell* **113,** 597–607 (2003).

213. Holtz, W. J. Engineering scalable combinational logic in *Escherichia coli* using zinc finger proteins. *Electrical Engineering* (2011).

214. Gaj, T., Gersbach, C. A. & Barbas, C. F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology* **31,** 397-405 (2013).

215. Zhang, Y. *et al.* Transcription Activator-Like Effector Nucleases Enable Efficient Plant Genome Engineering. *Plant Physiol.* **161,** 20-7 (2013).

216. Itzkovitz, S., Tlusty, T. & Alon, U. Coding limits on the number of transcription factors. *BMC Genomics* **7,** 239 2006).

217. Schoenborn, L., Yates, P. S., Grinton, B. E., Hugenholtz, P. & Janssen, P. H. Liquid serial dilution is inferior to solid media for isolation of cultures representative of the phylum-level diversity of soil bacteria. *Appl. Environ. Microbiol.* **70,** 4363-6 (2004).

218. Timmusk, S. *et al.* Bacterial distribution in the rhizosphere of wild barley under contrasting microclimates. *PLoS One* (2011).

219. Glick, B. R. Plant Growth-Promoting Bacteria: Mechanisms and Applications. *Scientifica (Cairo).* **2012,** 1–15 (2012).

220. Badri, D. V. & Vivanco, J. M. Regulation and function of root exudates. *Plant, Cell Environ.* (2009).

221. Badri, D. V., Weir, T. L., van der Lelie, D. & Vivanco, J. M. Rhizosphere chemical dialogues: plant-microbe interactions. *Current Opinion in Biotechnology* (2009).

222. Haichar, F. el Z., Santaella, C., Heulin, T. & Achouak, W. Root exudates mediated interactions belowground. *Soil Biology and Biochemistry* (2014).

223. Udvardi, M., Brodie, E. L., Riley, W., Kaeppler, S. & Lynch, J. Impacts of Agricultural Nitrogen on the Environment and Strategies to Reduce these Impacts. *Procedia Environ. Sci.* **29,** 303 (2015).

224. Dubeikovsky, A. N., Mordukhova, E. A., Kochetkov, V. V., Polikarpova, F. Y. & Boronin, A. M. Growth promotion of blackcurrant softwood cuttings by recombinant strain *Pseudomonas fluorescens* BSP53a synthesizing an increased amount of indole-3-acetic acid. *Soil Biol. Biochem.* (1993).

225. Liu, H. *et al.* The infection and impact of Azorhizobium caulinodans ORS571 on wheat (*Triticum aestivum*). *PLoS One* **12,** e0187947 (2017).

226. Yegorenkova, I. V., Konnova, S. A., Sachuk, V. N. & Ignatov, V. V. *Azospirillum brasilense* colonisation of wheat roots and the role of lectin–carbohydrate interactions in bacterial adsorption and root-hair deformation. *Plant Soil* **231,** 275–282 (2001).

227. Hirsch, A. M., Lum, M. R. & Downie, J. A. What makes the rhizobia-legume symbiosis so special? *Plant Physiol.* **127,** 1484–92 (2001).

228. Glick, B. R. The enhancement of plant growth by free-living bacteria. *Can. J. Microbiol.* (1995).

229. Wallenstein, M. D. Managing and manipulating the rhizosphere microbiome for plant health: A systems approach. *Rhizosphere* (2017).

230. Hill, K. *et al.* Root Systems Biology: Integrative Modeling across Scales, from Gene Regulatory Networks to the Rhizosphere. *PLANT Physiol.* (2013).

231. Poza-Carrión, C., Jiménez-Vicente, E., Navarro-Rodríguez, M., Echavarri-Erasun, C. & Rubio, L. M. Kinetics of nif gene expression in a nitrogen-fixing bacterium. *J. Bacteriol.* (2014).

232. Dixon, R. & Kahn, D. Genetic regulation of biological nitrogen fixation. *Nature Reviews Microbiology* (2004).

233. Curatti, L. & Rubio, L. M. Challenges to develop nitrogen-fixing cereals by direct nif-gene transfer. *Plant Sci.* **225,** 130–137 (2014).

234. Burén, S. *et al.* Formation of Nitrogenase NifDK Tetramers in the Mitochondria of *Saccharomyces cerevisiae*. *ACS Synth. Biol.* (2017).

235. Setten, L. *et al.* Engineering *Pseudomonas protegens* Pf-5 for Nitrogen Fixation and its Application to Improve Plant Growth under Nitrogen-Deficient Conditions. *PLoS One* (2013).

236. Marroquí, S. *et al.* Enhanced symbiotic performance by *Rhizobium tropici*

glycogen synthase mutants. *J. Bacteriol.* **183,** 854–864 (2001).

237. Campbell, E. R., Warsko, K., Davidson, A. M. & Campbell, W. H. Determination of phosphate in soil extracts in the field: A green chemistry enzymatic method. *MethodsX* (2015).

238. Khan, M. S., Zaidi, A. & Wani, P. A. Role of phosphate-solubilizing microorganisms in sustainable agriculture — A review. *Agron. Sustain. Dev.* (2007).

239. Richardson, A. E. Prospects for using soil microorganisms to improve the acquisition of phosphorus by plants. *Aust. J. Plant Physiol.* (2001).

240. Rodríguez, H. & Fraga, R. Phosphate solubilizing bacteria and their role in plant growth promotion. *Biotechnol. Adv.* (1999).

241. Rodriguez, H., Gonzalez, T., Goire, I. & Bashan, Y. Gluconic acid production and phosphate solubilization by the plant growth-promoting bacterium *Azospirillum* spp. *Naturwissenschaften* (2004).

242. Yang, P. X. *et al.* Phosphate Solubilizing Ability and Phylogenetic Diversity of Bacteria from P-Rich Soils Around Dianchi Lake Drainage Area of China. *Pedosphere* (2012).

243. Adhikary, H., Sanghavi, P. B., Macwan, S. R., Archana, G. & Naresh Kumar, G. Artificial citrate operon confers mineral phosphate solubilization ability to diverse fluorescent pseudomonads. *PLoS One* **9,** (2014).

244. Rojas, A., Holguin, G., Glick, B. R. & Bashan, Y. Synergism between *Phyllobacterium* sp. (N2-fixer) and *Bacillus licheniformis* (P-solubilizer), both from a semiarid mangrove rhizosphere. *FEMS Microbiol. Ecol.* (2001).

245. Patten, C. L. & Glick, B. R. Bacterial biosynthesis of indole-3-acetic acid. *Can. J. Microbiol.* (1996).

246. Spaepen, S., Vanderleyden, J. & Remans, R. Indole-3-acetic acid in microbial and microorganism-plant signaling. *FEMS Microbiology Reviews* (2007).

247. Xie, H., Pasternak, J. J. & Glick, B. R. Isolation and characterization of mutants of the plant growth promoting rhizobacterium *Pseudomonas putida* GR12-2 that overproduce indoleacetic-acid. *Curr. Microbiol.* (1996).

248. Spaepen, S. & Vanderleyden, J. Auxin and plant-microbe interactions. *Cold Spring Harb. Perspect. Biol.* (2011).

249. Tsavkelova, E. A., Klimova, S. Y., Cherdyntseva, T. A. & Netrusov, A. I.

Microbial producers of plant growth stimulators and their practical use: A review. *Appl. Biochem. Microbiol.* (2006).

250. Pilet, P. E. & Saugy, M. Effect on Root Growth of Endogenous and Applied IAA and ABA: A Critical Reexamination. *Plant Physiol.* (1987).

251. Whipps, J. M. Microbial interactions and biocontrol in the rhizosphere. *J. Exp. Bot.* (2001).

252. Wiederhold, N. Antifungal resistance: current trends and future strategies to combat. *Infect. Drug Resist.* (2017).

253. Compant, S., Duffy, B., Nowak, J., Clément, C. & Barka, E. A. Use of plant growth-promoting bacteria for biocontrol of plant diseases: Principles, mechanisms of action, and future prospects. *Applied and Environmental Microbiology* (2005).

254. Syed Ab Rahman, S. F., Singh, E., Pieterse, C. M. J. & Schenk, P. M. Emerging microbial biocontrol strategies for plant pathogens. *Plant Science* (2018).

255. Ramette, A., Frapolli, M., Défago, G. & Moënne-Loccoz, Y. Phylogeny of HCN synthase-encoding hcnBC genes in biocontrol fluorescent pseudomonads and its relationship with host plant species and HCN synthesis ability. *Mol. Plant. Microbe. Interact.* (2003).

256. Raaijmakers, J. M., Vlami, M. & de Souza, J. T. Antibiotic production by bacterial biocontrol agents. *Antonie Van Leeuwenhoek Int. J. Gen. Mol. Microbiol.* (2002).

257. Kwak, Y. S. & Weller, D. M. Take-all of wheat and natural disease suppression: A review. *Plant Pathology Journal* (2013).

258. Gong, L. *et al.* Novel synthesized 2, 4-DAPG analogues: Antifungal activity, mechanism and toxicology. *Sci. Rep.* (2016).

259. Gleeson, O., O'Gara, F. & Morrissey, J. P. The Pseudomonas fluorescens secondary metabolite 2,4 diacetylphloroglucinol impairs mitochondrial function in *Saccharomyces cerevisiae*. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* (2010).

260. Kwak, Y. S. *et al. Saccharomyces cerevisiae* genome-wide mutant screen for sensitivity to 2,4-diacetylphloroglucinol, an antibiotic produced by *Pseudomonas fluorescens*. *Appl. Environ. Microbiol.* (2011).

261. Ahemad, M. & Kibret, M. Mechanisms and applications of plant growth promoting rhizobacteria: Current perspective. *Journal of King Saud*

University - Science (2014).

262. Cohen, B. a, Amsellem, Z., Maor, R., Sharon, A. & Gressel, J. Transgenically enhanced expression of indole-3-acetic Acid confers hypervirulence to plant pathogens. *Phytopathology* (2002).

263. Zhao, Y. Auxin biosynthesis: A simple two-step pathway converts tryptophan to indole-3-Acetic acid in plants. in *Molecular Plant* (2012).

264. Fett, W. F., Osman, S. F. & Dunn, M. F. Auxin production by plant-pathogenic pseudomonads and xanthomonads. *Appl. Environ. Microbiol.* (1987).

265. Yin, C., Park, J.-J., Gang, D. R. & Hulbert, S. H. Characterization of a Tryptophan 2-Monooxygenase Gene from *Puccinia graminis* f. sp. *tritici* Involved in Auxin Biosynthesis and Rust Pathogenicity. *Mol. Plant-Microbe Interact.* (2014).

266. Gordon, S. A. & Weber, R. P. Colorimetric Estimation of Inodoleacetic Acid. *Plant Physiol.* (1951).

267. Yamakawa, T. *et al.* Stability of indole-3-acetic acid to autoclaving, aeration and light illumination. *Agricultural and Biological Chemistry* (1979).

268. Geisel, N. Constitutive versus responsive gene expression strategies for growth in changing environments. *PLoS One* (2011).

269. Wang, D., Xu, A., Elmerich, C. & Ma, L. Z. Biofilm formation enables free-living nitrogen-fixing rhizobacteria to fix nitrogen under aerobic conditions. *ISME J.* **11,** 1602–1613 (2017).

270. Pitts, R. J., Cernac, A. & Estelle, M. Auxin and ethylene promote root hair elongation in Arabidopsis. *Plant J.* (1998).

271. Okamoto, T. *et al.* Genetic Dissection of Hormonal Responses in the Roots of Arabidopsis Grown under Continuous Mechanical Impedance. *PLANT Physiol.* (2008).

272. Thomas, W. T. B., Powell, W. & Wood, W. The chromosomal location of the dwarfing gene present in the spring barley variety golden promise. *Heredity (Edinb).* (1984).

273. Fenn, K., Malhi, Y., Morecroft, M., Lloyd, C. & Thomas, M. Comprehensive description of the carbon cycle of an ancient temperate broadleaved woodland. *Biogeosciences Discuss.* (2010).

274. Tkacz, A., Hortala, M. & Poole, P. S. Absolute quantitation of microbiota

abundance in environmental samples. *Microbiome* (2018).

275. Schindelin, J. *et al.* Fiji: An open-source platform for biological-image analysis. *Nature Methods* (2012).

276. Woodward, A. W. & Bartel, B. Auxin: Regulation, action, and interaction. *Annals of Botany* (2005).

277. Rahman, A. *et al.* Auxin, actin and growth of the *Arabidopsis thaliana* primary root. *Plant J.* (2007).

278. Leitner, D., Felderer, B., Vontobel, P. & Schnepf, A. Recovering Root System Traits Using Image Analysis Exemplified by Two-Dimensional Neutron Radiography Images of Lupine. *PLANT Physiol.* (2014).

279. Flavel, R. J., Guppy, C. N., Rabbi, S. M. R. & Young, I. M. An image processing and analysis tool for identifying and analysing complex plant root systems in 3D soil using non-destructive analysis: Root1. *PLoS One* (2017).

280. Michelsen, C. F. & Stougaard, P. A novel antifungal *Pseudomonas fluorescens* isolated from potato soils in Greenland. *Curr. Microbiol.* (2011).

281. Meyer, A. J., Segall-Shapiro, T. H. & Voigt, C. A. Marionette: *E. coli* containing 12 highly-optimized small molecule sensors. *bioRxiv* (2018).

282. Goldstein, A. H. Recent progress in understanding the molecular-genetics and biochemistry of calcium-phosphate solubilization by Gram-negative bacteria. *Biol. Agric. Hortic.* (1995).

283. Wagh, J., Shah, S., Bhandari, P., Archana, G. & Kumar, G. N. Heterologous expression of pyrroloquinoline quinone (pqq) gene cluster confers mineral phosphate solubilization ability to *Herbaspirillum seropedicae* Z67. *Appl. Microbiol. Biotechnol.* **98,** 5117–5129 (2014).

284. Archana, G., Buch, A. & Kumar, G. N. *Pivotal role of organic acid secretion by rhizobacteria in plant growth promotion*. *Microorganisms in Sustainable Agriculture and Biotechnology* (2013).

285. Buch, A. D., Archana, G. & Kumar, G. N. Enhanced citric acid biosynthesis in *Pseudomonas fluorescens* ATCC 13525 by overexpression of the Escherichia coli citrate synthase gene. *Microbiology* (2009).

286. Sobczak, I. & Lolkema, J. S. The 2-Hydroxycarboxylate Transporter Family: Physiology, Structure, and Mechanism. *Microbiol. Mol. Biol. Rev.* (2005).

287. Nautiyal, C. S. An efficient microbiological growth medium for screening

phosphate solubilizing microorganisms. *FEMS Microbiol. Lett.* (1999).

288. Singh, H. & Reddy, M. S. Effect of inoculation with phosphate solubilizing fungus on growth and nutrient uptake of wheat and maize plants fertilized with rock phosphate in alkaline soils. *Eur. J. Soil Biol.* (2011).

289. Romasi, E. F. & Lee, J. Development of indole-3-acetic acid-producing *Escherichia coli* by functional expression of IpdC, AspC, and Iad1. *J. Microbiol. Biotechnol.* (2013).

290. Hartmann, A., Singh, M. & Klingmüller, W. Isolation and characterization of *Azospirillum* mutants excreting high amounts of indoleacetic acid. *Can. J. Microbiol.* (1983).

291. Stanton, B. C. *et al.* Systematic Transfer of Prokaryotic Sensors and Circuits to Mammalian Cells. *ACS Synth. Biol* **3,** 45 (2014).

292. Blommel, P. G., Becker, K. J., Duvnjak, P. & Fox, B. G. Enhanced bacterial protein expression during auto-induction obtained by alteration of lac repressor dosage and medium composition. *Biotechnol. Prog.* (2007).

293. Sivashanmugam, A. *et al.* Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Sci* (2009).

294. Kelley, K. D., Olive, L. Q., Hadziselimovic, A. & Sanders, C. R. Look and see if it is time to induce protein expression in *Escherichia coli* cultures. *Biochemistry* (2010).

295. Niu, Y. *et al.* Auxin modulates the enhanced development of root hairs in *Arabidopsis thaliana* (L.) Heynh. under elevated CO 2. *Plant, Cell Environ.* (2011).

296. Zhang, S. *et al.* Multiple phytohormones promote root hair elongation by regulating a similar set of genes in the root epidermis in *Arabidopsis*. *J. Exp. Bot.* (2016).

297. Vanneste, S. & Friml, J. Auxin: A Trigger for Change in Plant Development. *Cell* (2009).

298. Eliasson, L., Bertell, G. & Bolander, E. Inhibitory action of auxin on root elongation not mediated by ethylene. *Plant Physiol.* (1989).

299. Wu, Q., Pagès, L. & Wu, J. Relationships between root diameter, root length and root branching along lateral roots in adult, field-grown maize. *Ann. Bot.* (2016).

300. Pacheco-Villalobos, D. *et al.* The Effects of High Steady State Auxin Levels

on Root Cell Elongation in Brachypodium. *Plant Cell* (2016).

301. Pierret, A. *et al.* Understanding deep roots and their functions in ecosystems: An advocacy for more unconventional research. *Annals of Botany* (2016).

302. Ji, X. dong, Cong, X., Dai, X. qing, Zhang, A. & Chen, L. hua. Studying the mechanical properties of the soil-root interface using the pullout test method. *J. Mt. Sci.* (2018).

303. Jackson, T., Mansfield, K., Saafi, M., Colman, T. & Romine, P. Measuring soil temperature and moisture using wireless MEMS sensors. *Measurement* (2008).

304. Islam, M. T. & von Tiedemann, A. 2,4-Diacetylphloroglucinol suppresses zoosporogenesis and impairs motility of *Peronosporomycete* zoospores. *World J. Microbiol. Biotechnol.* (2011).

305. Shishido, M., Breuil, C. & Chanway, C. P. Endophytic colonization of spruce by plant growth-promoting rhizobacteria. *FEMS Microbiol. Ecol.* (1999).

306. Mendis, H. C. *et al.* Strain-specific quantification of root colonization by plant growth promoting rhizobacteria *Bacillus firmus* I-1582 and *Bacillus amyloliquefaciens* QST713 in non-sterile soil and field conditidoi:10.1371/journal.pone.0193119

307. Yan, X. *et al.* Transcriptional regulator PhlH modulates 2,4-diacetylphloroglucinol biosynthesis in response to the biosynthetic intermediate and end product. *Appl. Environ. Microbiol.* (2017).

308. Foster, D. W. Malonyl-CoA: The regulator of fatty acid synthesis and oxidation. *Journal of Clinical Investigation* (2012).

309. Srinivasamurthy, R., Pratibha Singh, J. & Rai, A. K. *Biological control of bacterial wilt disease-causing pathogens: A sustainable approach for increasing crop production*. *Microbial Diversity and Biotechnology in Food Security* (2014).

310. Quandt, E. M. *et al.* Fine-tuning citrate synthase flux potentiates and refines metabolic innovation in the lenski evolution experiment. *Elife* (2015).

311. Derrien, D., Marol, C. & Balesdent, J. The dynamics of neutral sugars in the rhizosphere of wheat. An approach by 13C pulse-labelling and GC/C/IRMS. *Plant Soil* (2005).

312. Narula, N., Kothe, E. & Behl, R. K. Role of root exudates in plant-microbe interactions. *J. Appl. Bot. Food Qual.* **82,** 122–130 (2009).

313. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science.* (2007).

314. Wiedenheft, B., Sternberg, S. H. & Doudna, J. A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* (2012).

315. Brouns, S. J. J. *et al.* Small Crispr Rnas Guide Antiviral Defense in Prokaryotes. *Cancer Epidemiol. Biomarkers Prev.* (1993).

316. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* (2012).

317. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* (2013).

318. Roberts, B. *et al.* Systematic gene tagging using CRISPR/Cas9 in human stem cells to illuminate cell organization. *Mol. Biol. Cell* (2017).

319. Arbab, M., Srinivasan, S., Hashimoto, T., Geijsen, N. & Sherwood, R. I. Cloning-free CRISPR. *Stem Cell Reports* (2015).

320. Ji, W. *et al.* Specific Gene Repression by CRISPRi System Transferred through Bacterial Conjugation. *ACS. Synth. Biol*. **3,** 929-931 (2014).

321. La Russa, M. F. & Qi, L. S. The New State of the Art: Cas9 for Gene Activation and Repression. *Mol. Cell. Biol.* **35,** 3800-9 (2015).

322. Zheng, T. *et al.* Profiling single-guide RNA specificity reveals a mismatch sensitive core sequence. *Sci. Rep.* **7,** 40638 (2017).

323. Jang, S., Jang, S. & Jung, G. Y. Toward tunable dynamic repression using CRISPRi. *Biotechnology Journal* **13,** e1800052 (2018).

324. Fontana, J., Dong, C., Ham, J. Y., Zalatan, J. G. & Carothers, J. M. Regulated Expression of sgRNAs Tunes CRISPRi in E. coli. *Biotechnology Journal* **13,** e1800069 (2018).

325. Lowder, L. G., Malzahn, A. & Qi, Y. in *Methods in molecular biology (Clifton, N.J.)* (2018).

326. Zhao, Y. *et al.* CRISPR/dCas9-Mediated Multiplex Gene Repression in Streptomyces. *Biotechnology Journal* **13,** e1800121 (2018).

327. Kleinstiver, B. P. *et al.* Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523,** 481-485 (2015).

328. Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target

cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33,** 187-197 (2015).

329. Liu, J. *et al.* CRISPR/Cas9 in zebrafish: an efficient combination for human genetic diseases modeling. *Human Genetics* **136,** 1-12 (2017).

330. Cebrian-Serrano, A. & Davies, B. CRISPR-Cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. *Mamm. Genome* **28,** 247-61 (2017).

331. Alonso-Gutierrez, J. *et al.* Metabolic engineering of Escherichia coli for limonene and perillyl alcohol production. *Metab. Eng.* **19,** 33-41 (2013).

332. Yuan, Y., Du, J. & Zhao, H. Customized optimization of metabolic pathways by combinatorial transcriptional engineering. *Methods Mol. Biol.* **40,** e142 (2013).

333. Lee, S. Y. & Kim, H. U. Systems strategies for developing industrial microbial strains. *Nature Biotechnology* **33,** 1061-72 (2015).

334. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 6640-45 (2000).

335. Yi, Y., Li, Z., Song, C. & Kuipers, O. P. Exploring plant-microbe interactions of the rhizobacteria *Bacillus subtilis* and *Bacillus mycoides* by use of the CRISPR-Cas9 system. *Environ. Microbiol.* (2018).

336. Bozouklian, H., Fogher, C. & Elmerich, C. Cloning and characterization of the glnA gene of *Azospirillum brasilense* Sp7. *Ann. l'Institut Pasteur / Microbiol.* **137,** 3–18 (1986).

337. Van Dommelen, A., Croonenborghs, A., Spaepen, S. & Vanderleyden, J. Wheat growth promotion through inoculation with an ammonium-excreting mutant of *Azospirillum brasilense*. *Biol. Fertil. Soils* **45,** 549–553 (2009).

338. Ambrosio, R., Cesar, J., Ortiz-Marquez, F. & Curatti, L. Metabolic engineering of a diazotrophic bacterium improves ammonium release and biofertilization of plants and microalgae. *Metab. Eng.* **40,** 59–68 (2017).

339. Bruce Rowe, W., Ronzio, R. A. & Meister, A. Inhibition of Glutamine Synthetase by Methionine Sulfoximine. Studies on Methionine Sulfoximine Phosphate. *Biochemistry* **8,** 2674-80 (1969).

340. Ahmad, I. & Hellebust, J. A. Effects of methionine sulfoximine on growth and nitrogen assimilation of the marine microalga *Chlorella autotrophica*. *Mar. Biol.* **86,** 85-91 (1985).

341. Eisenberg, D., Gill, H. S., Pfluegl, G. M. U. & Rotstein, S. H. Structure–function relationships of glutamine synthetases. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1477,** 122-45 (2000).

342. de Zamaroczy, M., Paquelin, A. & Elmerich, C. Functional organization of the glnB-glnA cluster of *Azospirillum brasilense*. *J. Bacteriol.* **175,** 2507–15 (1993).

343. Tombolini, R., Unge, A., Davey, M. E., De Bruijn, F. J. & Jansson, J. K. Flow cytometric and microscopic analysis of GFP-tagged *Pseudomonas fluorescens* bacteria. *FEMS Microbiol. Ecol.* **22,** 17-28 (1997).

344. Doench, J. G. *et al.* Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.* **32,** 1262-1267 (2014).

345. McMillan, M. & Pereg, L. Evaluation of Reference Genes for Gene Expression Analysis Using Quantitative RT-PCR in *Azospirillum brasilense*. *PLoS One* **9,** e98162 (2014).

346. Vigouroux, A., Oldewurtel, E., Cui, L., Teeffelen, S. van & Bikard, D. Engineered CRISPR-Cas9 system enables noiseless, fine-tuned and multiplexed repression of bacterial genes. *bioRxiv* (2017).

347. Gossen, M. & Bujard, H. Anhydrotetracycline, a novel effector for tetracycline controlled gene expression systems in eukaryotic cells. *Nucleic Acids Res.* **21,** 4411-12 (1993).

348. Ma, H. *et al.* CRISPR-Cas9 nuclear dynamics and target recognition in living cells. *J. Cell Biol.* **214,** 529 (2016).

349. Yang, Y. jie *et al.* Increasing on-target cleavage efficiency for CRISPR/Cas9-induced large fragment deletion in *Myxococcus xanthus*. *Microb. Cell Fact.* **16,** 142 (2017).

350. Zhang, L., Rube, H. T., Bussemaker, H. J. & Pufall, M. A. The effect of sequence mismatches on binding affinity and endonuclease activity are decoupled throughout the Cas9 binding site. *BioRxiv* (2017).

351. Yegorenkova, I. V., Konnova, S. A., Sachuk, V. N. & Ignatov, V. V. *Azospirillum brasilense* colonisation of wheat roots and the role of lectin-carbohydrate interactions in bacterial adsorption and root-hair deformation. *Plant Soil* **231,** 275–282 (2001).

352. Bacilio-Jiménez, M. *et al.* Endophytic bacteria in rice seeds inhibit early colonization of roots by *Azospirillum brasilense*. *Soil Biol. Biochem.* **33,** 167–172 (2001).

353. Colnaghi, R., Green, A., He, L., Rudnick, P. & Kennedy, C. Strategies for increased ammonium production in free-living or plant associated nitrogen fixing bacteria. *Plant Soil* **194,** 145–154 (1997).

354. Jaitin, D. A. *et al.* Dissecting Immune Circuits by Linking CRISPR-Pooled Screens with Single-Cell RNA-Seq. *Cell* **167,** 1883-96 (2016).

355. Hao, N., Shearwin, K. E. & Dodd, I. B. Programmable DNA looping using engineered bivalent dCas9 complexes. *Nat. Commun.* **8,** 1628 (2017).

356. Ma, D., Peng, S. & Xie, Z. Integration and exchange of split dCas9 domains for transcriptional controls in mammalian cells. **3,** 13056 *Nat. Commun.* (2016).

357. Lai, H. E., Moore, S., Polizzi, K. & Freemont, P. EcoFlex: A multifunctional moclo kit for E. coli synthetic biology. *Methods Mol. Biol.* **1772,** 429–444 (2018).

358. Rhodius, V. A. & Mutalik, V. K. Predicting strength and function for promoters of the *Escherichia coli* alternative sigma factor, sigmaE. *Proc. Natl. Acad. Sci. U. S. A.* **16,** 2854-9 (2010).

359. Gilman, J. & Love, J. Synthetic promoter design for new microbial chassis. *Biochem. Soc. Trans.* **44,** 731-7 (2016).

360. Rodriguez, E. A. *et al.* A far-red fluorescent protein evolved from a cyanobacterial phycobiliprotein. *Nat. Methods* **13,** 763-9 (2016).

361. Tomosugi, W. *et al.* An ultramarine fluorescent protein with increased photostability and pH insensitivity. *Nat. Methods* **6,** 351-33 (2009).

362. Tett, A. J., Rudder, S. J., Bourdès, A., Karunakaran, R. & Poole, P. S. Regulatable vectors for environmental gene expression in Alphaproteobacteria. *Appl. Environ. Microbiol.* **78,** 7137–7140 (2012).

363. Bradley, R. W., Buck, M. & Wang, B. Tools and Principles for Microbial Gene Circuit Engineering. *J. Mol. Biol.* **428,** 862-888 (2016).

364. Pröschel, M., Detsch, R., Boccaccini, A. R. & Sonnewald, U. Engineering of Metabolic Pathways by Artificial Enzyme Channels. *Front. Bioeng. Biotechnol.* **3,** 168 (2015).

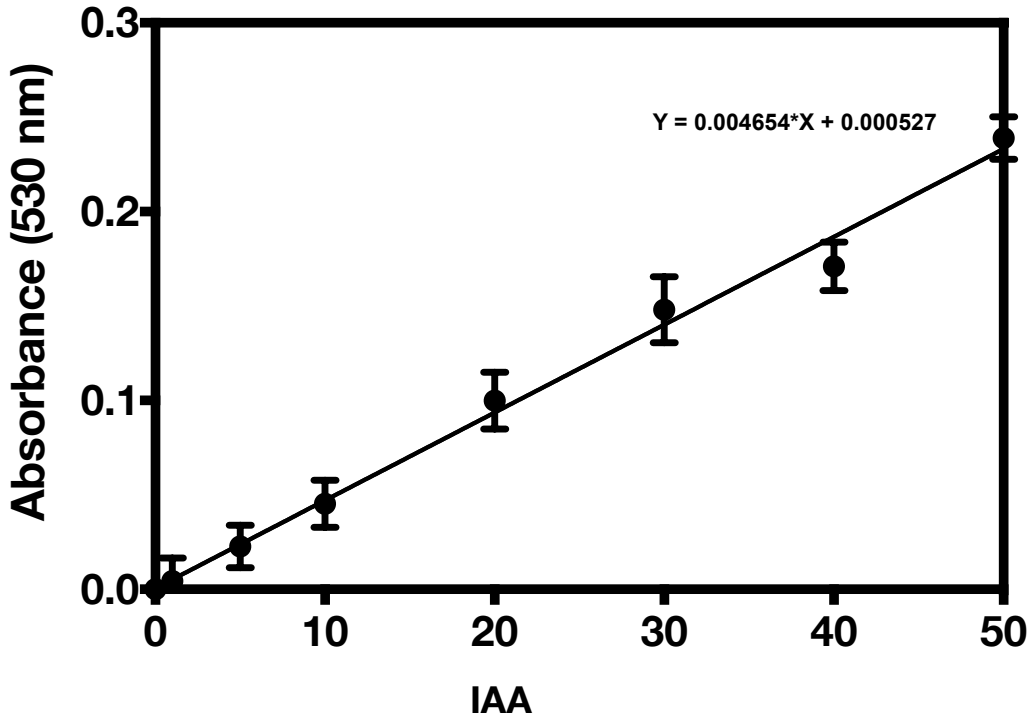
365. Vitorino, J. C. *et al.* Potential roles for the *glnB* and *ntrYX* genes in Azospirillum brasilense. *FEMS Microbiol. Lett.* **201,** 199–204 (2001).

366. Ambrosio, R., Ortiz-Marquez, J. C. F. & Curatti, L. Metabolic engineering of a diazotrophic bacterium improves ammonium release and biofertilization

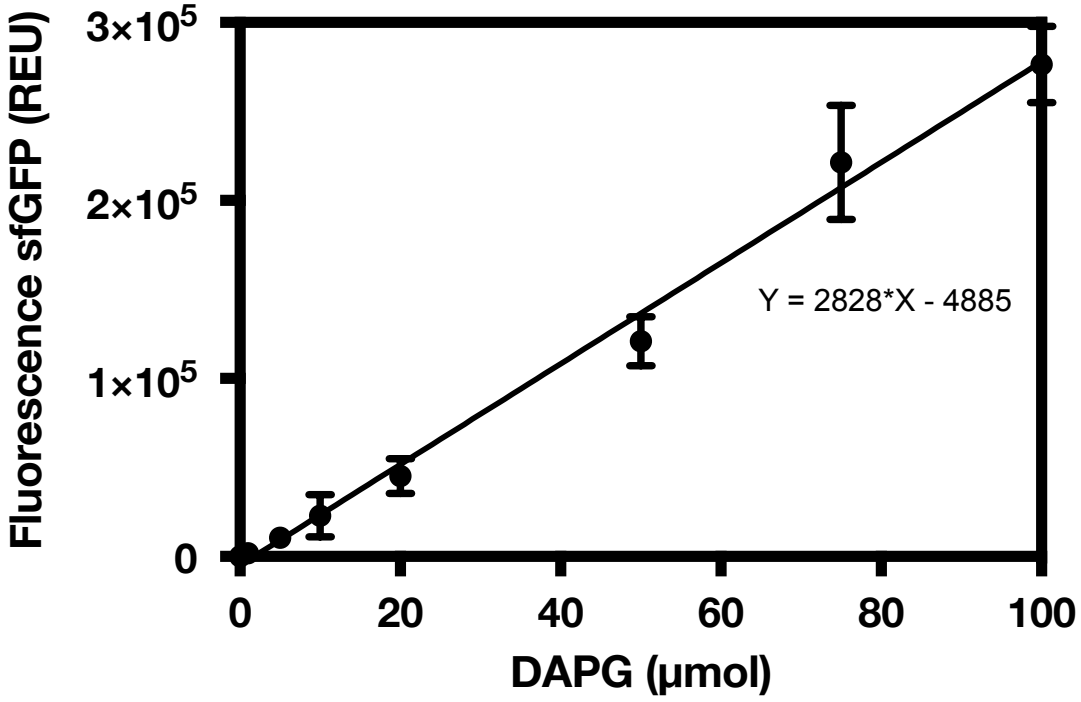
of plants and microalgae. *Metab. Eng.* **40,** 59-68 (2017).

367. Gaudelli, N. M. *et al.* Programmable base editing of A • T to G • C in genomic DNA without DNA cleavage. *Nat. Publ. Gr.* **551,** 464-471 (2017).

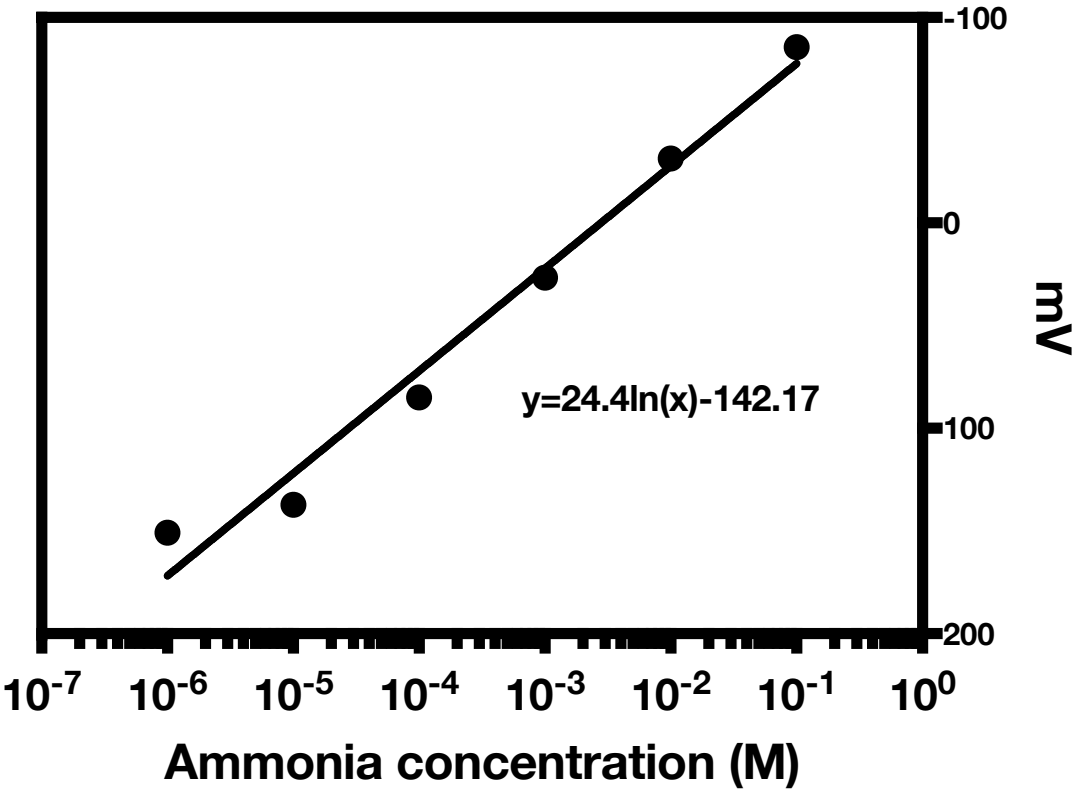
9 Appendices



Supplementary Figure 1 - IAA Standard curve for the quantification of IAA in liquid UMS media at room temperature and pressure. Error bars indicate the standard error of the mean. N = 3 independent technical replicates.



Supplementary Figure 2 - DAPG standard curve for the quantification of DAPG in liquid media at room temperature and pressure. Error bars indicate the standard error of the mean. N = 3 independent technical replicates.



Supplementary Figure 3 – Ammonia (NH_3) standard curve for the quantification of NH_3 in semisolid UMS media at room temperature and pressure. N = 3 independent technical replicates.

Due to the large size of the DNA sequences, plasmid maps and primer tables. The data will be made available at https://rhizosphere.org/lab-page/open-access-data/