Prototyping and Development of Synthetic Metabolic Pathways for the Biosynthesis of Flavour & Fragrance Compounds

Jonathan Emmanuel Muller

Department of Life Sciences

Imperial College London

A thesis submitted for the degree Doctor of Philosophy

2022

Abstract

There is an urgent need to develop sustainable methods for producing renewable chemicals. Photosynthetic cyanobacteria could be used to produce sustainable chemicals from CO_2 and sunlight. However, bioproduction using cyanobacteria has yet to be commercialised due to the high cultivation costs, and low product titres and values. To develop this technology, biosynthesis of higher value compounds should be targeted first. Flavour and fragrance compounds are a promising target, due to their high-value and large global market. In this work, synthetic flavour and fragrance biosynthesis pathways were prototyped in *Escherichia coli* for a faster Design-Build-Test-Learn Cycle, with the intention to thereafter transfer them to cyanobacteria. A novel pathway was proposed for the biosynthesis of decalactones, which are of commercial interest. The first step in the pathway was the biosynthesis of decanoic acid by a C10 specific acyl-ACP thioesterase. Such thioesterases had not yet been reported, so one with a C10 specificity of 77.1% was engineered. The second step in the pathway was hydroxylation of the decanoic acid by BM3 to generate 4-hydroxy and 5-hydroxydecanoic acid, which spontaneously condensed to form γ-decalactone and δ-decalactone, respectively. Native BM3 did not hydroxylate decanoic acid at the required positions, so mutants of BM3 were prepared that showed altered regioselectivity. The complete biosynthetic pathway was then expressed in *E. coli* and biosynthesis of γ-decalactone (3.53 mg/L) and δ-decalactone (0.53 mg/L) was achieved. The biosynthesis of other flavour and fragrance compounds using synthetic pathways that required less enzyme engineering was also investigated, to reduce the amount of work required in *E. coli* prior to transfer to cyanobacteria. The biosynthesis of butyl acetate (153 mg/L), medium chain length methyl ketones (3.73mg/L) and 2-dodecenol (22 mg/L) was achieved by adapting previously developed synthetic pathways, while biosynthesis of ethyl butyrate (224 mg/L) was achieved using a novel pathway.

Statement of Originality

I hereby confirm that the work presented in this thesis is my own. All external sources of information referred to in this thesis have been properly acknowledged and cited.

-Jonathan Muller

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution 4.0 International Licence (CC BY 4.0).

Under this licence, you may copy and redistribute the material in any medium or format for both commercial and non-commercial purposes. You may also create and distribute modified versions of the work. This is on the condition that you credit the author.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Acknowledgements

Undertaking a PhD is a substantial challenge, and I have been fortunate to receive support from a wide range of individuals during this time.

First and foremost, I would like to thank my primary supervisor, Professor Patrik Jones. He has been an endless source of encouragement, guidance, support, and mentorship throughout my PhD. I owe a lot of my development professionally and as a scientist to him.

I would also like to thank my secondary supervisor, Professor Klaus Hellgardt, and my Progress Review Panel, Dr Karen Polizzi and Dr Rodrigo Ledesma Amaro for their advice and feedback during the course of my PhD. I would also like to thank my examiners, Professor Nigel Scrutton and Professor Geoff Baldwin for their feedback on this thesis and an enjoyable and stimulating discussion during the viva.

I also want to thank all the members of the Jones lab who I have been fortunate enough to work alongside. I would like to thank Dr Ian Yunus, Dr Pachara Sattayawat, Dr Daniel Alvarez Zabala and Dr Marine Valton, who all helped mentor, train and support me during this journey. I would also like to thank Elisabeth Bubna-Litic, Helen (HeEun) Yang, George Sabin, Xiangming Xu, Angel Xia and all other previous members who I have not had the chance to name, for creating a research group that has been welcoming, supportive and pleasant to work in even during a global pandemic.

I would also like to thank my colleagues from the neighbouring Jiminez lab, Dr Alice Banks and Dr Satya Prakash for also giving me advice and helping me troubleshoot problems with the molecular biology. I would also like to thank Eleonora Moratto for her friendship and support, and Lucas Planelles, for helping me troubleshoot problems with the Basic assembly.

I also want to thank the Customer Success team at Synthace for hosting me during my PIPs placement, and would like to thank in particular Victoria González Sabater Budilov, for her mentorship and support, and Lauren Fovargue, for the very helpful career advice. My time with you was a real highlight of the PhD.

I would also like to my family for their support, especially my parents. They have always believed in me and supported me, and without them I wouldn’t have made it onto this PhD program. I also want to thank them for how they have helped me figure out the next steps in my career now that my PhD is finishing.

I would also like to thank Shengxin Chen, for her patience, understanding and support during these years. Having you there has helped me persevere through the hardships. I also want to thank Jody He for her friendship and support, and always being someone I can talk to. I would also like to thank Joanna C. Chang, Mingke Pan, Andrea Basso, Dani Raza and Edita Pileckyte for their help and friendship during this time.

I also want to thank Andrew G. Walton CBE for his sage words of advice over the past year.

Finally, science is not cheap. I would like to thank the BBSRC, UKRI, and the British taxpayers who have made this work possible.

Table of Contents

Abstract……2

Statement of Originality……3

Copyright Declaration……4

Acknowledgements……5

Table of Contents……7

List of Figures……17

List of Tables……21

Chapter 1 Introduction……23

1.1 Motivation……23

1.1.1 The need for sustainability……23

1.1.2 Sustainable bioproduction using cyanobacteria……23

1.1.3 Flavours and Fragrances……25

1.1.4 Prototyping biosynthetic pathways in *Escherichia coli*……25

1.2 Aims……26

1.3 Implementation……26

1.4 Thesis organisation……27

Chapter 2 Literature Review……28

2.1 *Escherichia coli*……28

2.1.1 *E. coli* as a model organism for synthetic biology……28

2.1.2 Fatty acid biosynthesis and β-oxidation……29

2.2 Lactone Biosynthesis……32

2.2.1 Lactones……32

2.2.2 Previously reported biosynthesis of γ- and δ-decalactone……33

2.2.3 Proposed pathway for γ- and δ-decalactone biosynthesis……34

2.2.4 Acyl-ACP thioesterases……..34

2.2.5 CYP450 BM3……36

2.2.6 Glucose dehydrogenase……38

2.2.7 Conversion of hydroxyfatty acids to lactones……...39

2.3 Ester Biosynthesis……40

2.3.1 Esters……40

2.3.2 Previously reported biosynthesis of esters……...41

2.3.3 Proposed pathway for butyl acetate biosynthesis……41

2.3.4 C4 specific acyl-ACP thioesterase……42

2.3.5 Carboxylic acid reductase……...42

2.3.6 Alcohol dehydrogenase……..42

2.3.7 Alcohol-acyl transferase……..43

2.3.8 Proposed pathway for ethyl butyrate biosynthesis……..43

2.3.9 Acetyl-CoA synthase……...43

2.4 Methyl Ketone Biosynthesis……45

2.4.1 Methyl ketones……..45

2.4.2 Native biosynthesis in *Penicillium*……..45

2.4.3 Native biosynthesis in plants……..46

2.4.4 Previously reported biosynthesis of methyl ketones……47

2.4.5 The spontaneity of β-ketofatty acid decarboxylation……...48

2.4.6 Proposed pathway for methyl ketone biosynthesis……48

2.5 Green Leaf Volatile Biosynthesis……49

2.5.1 Green Leaf Volatiles……49

2.5.2 Native GLV biosynthesis pathway……49

2.5.3 Previously reported biosynthesis of GLVs……....50

2.5.4 Proposed pathway for biosynthesis of GLVs……50

2.5.5 DfsA……51

Chapter 3 Materials and Methods……52

3.1 *Escherichia coli*……52

3.1.1 *E. coli* strains……52

3.1.2 LB media and LB-Agar……52

3.1.3 M9 minimal media……52

3.1.4 Antibiotics used for cultivation of *E. coli*……53

3.1.5 Cryopreservation and revival of *E. coli* glycerol stocks……54

3.1.6 Preparation of calcium competent *E. coli*……54

3.1.7 Heat-shock transformation of *E. coli*……55

3.1.8 Preparation of electrocompetent *E. coli*……55

3.1.9 Electroporation of *E. coli*……56

3.1.10 *E. coli* bioproduction and feeding experiments……56

3.2 *Synechocystis* Sp. PCC 6803……57

3.2.1 *Synechocystis* strains……57

3.2.2 BG11-Co media and BG11-Co-Agar……57

3.2.2 Antibiotics used for cultivation of *Synechocystis*……58

3.2.3 Cryopreservation and revival of *Synechocystis* strains……59

3.2.4 Maintenance of revived *Synechocystis* strains……59

3.2.5 Triparental conjugation of *Synechocystis*……59

3.2.6 *Synechocystis* bioproduction experiments……60

3.3 Molecular Biology Methods……61

3.3.1 Extraction of plasmid DNA……61

3.3.2 Quantification of DNA……61

3.3.3 Primers and gBlocks……61

3.3.4 pJET cloning…….....62

3.3.5 Colony PCR…….....62

3.3.6 Agarose gel electrophoresis…….....63

3.3.7 Gel extraction…….....64

3.3.8 Acyl-ACP thioesterase truncation…….....64

3.3.9 Amplification of genes from the genome…….....65

3.3.10 Production of hybrid CYP450s…….....65

3.3.11 Site-directed mutagenesis…….....67

3.3.12 BASIC DNA assembly…….....69

3.3.13 DNA sequencing…….....71

3.3.14 Knock-out of *iscR* from the genome of *E. coli*…….....71

3.4 Product Analysis…….....74

3.4.1 Measurement of *E. coli* OD_{600}…….....74

3.4.2 Measurement of *Synechocystis* OD_{730}…….....74

3.4.3 Gas-Chromatography Mass-Spectrometry (GC-MS) …….....74

3.4.4 High-Performance Liquid Chromatography (HPLC) …….....74

3.4.5 Extraction of fatty alcohols in the presence of an isopropyl myristate overlay…74

3.4.6 Extraction of decanoic acid and decalactones from the culture media……75

3.4.7 Extraction of decanoic acid and decalactones in the presence of a decane overlay…….....75

3.4.8 Extraction of C4 esters in the presence of a hexadecane overlay……76

3.4.9 Extraction of methyl ketones in the presence of an isopropyl overlay……..76

3.4.10 Extraction and decarboxylation of β-ketofatty acids from the culture media…..76

3.4.11 Analysis of soluble products……77

3.4.12 Analysis of glucose consumption……77

3.5 Bioinformatics Methods……..78

3.5.1 Prediction of the chloroplast localisation peptide……78

3.5.3 Protein sequence alignments……78

3.5.3 Generation of predicted protein structures……78

3.5.4 Structure guided mutagenesis……78

3.5.5 Alignment of protein structures……78

3.5.6 Blast……78

3.6 Statistical Analysis……79

Chapter 4 Engineering a C10 Acyl-ACP Thioesterase……80

4.1 Aims……80

4.2 Results……81

4.2.1 Truncation of *Cuphea viscosissima* FatB1……81

4.2.2 Truncation of *Ulmus americana* FatB1……84

4.2.3 Screening CvFatB1 truncations in *Synechocystis* Sp. PCC 6803……87

4.2.4 Inducer optimisation of CvFatB1 truncation expression……89

4.2.4.1 Inducer optimisation of CvT3 expression……90

4.2.4.2 Inducer optimisation of CvT4 expression……91

4.2.4.3 Inducer optimisation of CvT5 expression……92

4.2.4.4 Inducer optimisation of CvT6 expression……93

4.2.4.5 Inducer optimisation of CvT7 expression……94

4.2.4.6 Inducer optimisation of CvT8 expression……95

4.2.4.7 Comparison of optimally expressed CvFatB1 truncations……96

4.2.5 Overexpression of ACP……98

4.2.5.1 ACP expression in combination with CvT3……99

4.2.5.2 ACP expression in combination with CvT4……100

4.2.5.3 ACP expression in combination with CvT5……101

4.2.5.4 ACP expression in combination with CvT6……102

4.2.5.5 ACP expression in combination with CvT7……103

4.2.5.6 ACP expression in combination with CvT8……104

4.2.6 Site-directed mutagenesis of CvT5……105

4.2.7 Screening CvT5 mutants……..106

4.2.7.1 Screening and inducer optimisation of CvT5 M90G……106

4.2.7.2 Screening and inducer optimisation of CvT5 W122G……..107

4.2.7.3 Screening and inducer optimisation of CvT5 W140G……..108

4.2.8 Optimising the mutation at the CvT5 W122 site……109

4.3 Discussion……...112

4.3.1 The effect of truncation upon FatB acyl-ACP thioesterase activity in *E. coli*…..112

4.3.2 The difference in thioesterase C10 specificity between *E. coli* and *Synechocystis*……...113

4.3.3 The effects of ACP expression upon alcohol biosynthesis……...114

4.3.4 Engineering CvT5 for increased C10 specificity in *E. coli*……115

4.4 Conclusion……..117

Chapter 5 Engineering BM3 for the ω-5 and ω-6 Hydroxylation of Decanoic Acid……118

5.1 Aims……118

5.2 Results……119

5.2.1 Decanoic acid feeding assay……119

5.2.2 Assaying methyl decanoate as an alternative substrate for feeding assays…121

5.2.3 Could sample acidification improve decalactone titres?.....122

5.2.4 Could an overlay be used for product extraction?.....123

5.2.5 Optimisation of the M9 media for BM3 expression……125

5.2.6 Localisation of decanoic acid, δ-decalactone and γ-decalactone……..127

5.2.7 Screening alternative CYP450s……...128

5.2.8 Could deletion of *iscR* improve the activity of the new CYP450s?.....129

5.2.9 Mutagenesis of BM3……..…...131

5.2.10 Mutagenesis of the BM3 F87 site……..…....133

5.2.11 Adding activity enhancing mutations to BM3 F87G……134

5.2.12 Extension of the decalactone biosynthesis pathway……..136

5.2.13 Could increasing oxygen supply improve BM3 activity?.....140

5.2.14 Could NADPH recycling improve BM3 activity?.....141

5.2.15 Media supplementation with caffeine……...…143

5.2.16 Knock-out of *iscR* in *E. coli* ΔfadD……145

5.2.17 Screening new BM3 activity enhancing mutations……..146

5.3 Discussion……..….....149

5.3.1 The regioselectivity and substrate specificity of BM3……149

5.3.2 The spontaneity of hydroxyfatty acid condensation……..……149

5.3.3 Product localisation……...….....150

5.3.4 Media optimisation…….……...….....150

5.3.5 Mutagenesis of BM3…….……..….....151

5.3.6 Decalactone biosynthesis pathway extension……..……152

5.3.7 Co-factor balance…….……...….....153

5.4 Conclusions……..….....154

Chapter 6 *De novo* Biosynthesis of δ-Decalactone and γ-Decalactone……155

6.1 Aims……..….....155

6.2 Results……..….....156

6.2.1 *De novo* biosynthesis of decalactones in *E. coli*……..……156

6.2.2 Was ferric ammonium citrate detrimental to thioesterase activity?.....158

6.2.3 Redesign of plasmid for expressing pathway enzymes and RBS optimisation of CvT5 W122F expression……..……..159

6.2.4 *De novo* biosynthesis using new BM3 mutants and screening the effect of a decane overlay……..……..161

6.2.5 Screening hybrid CYP450s……..……163

6.2.6 New operon order for EcACP and CvT5 W122F expression……..165

6.2.7 Media supplementation with Cas amino acids……...166

6.2.8 RBS optimisation of EcACP expression…….167

6.2.9 RBS optimisation of BM3 expression…….....168

6.2.10 Biosynthesis of decalactones in *Synechocystis* Sp. PCC 6803……..170

6.3 Discussion……..…….172

6.3.1 *De novo* biosynthesis of decalactones in *E. coli*…….……..172

6.3.2 Low decanoic acid titres…….……...173

6.3.3 Conversion of decanoic acid to decalactones…….……..175

6.3.4 Biosynthesis of decalactones in *Synechocystis*…….……176

6.4 Conclusions……..…….178

Chapter 7 Prototyping Biosynthetic Pathways for Alternative Flavour and Fragrance Compounds……..……..179

7.1 Aims……..……..…179

7.2 Results: Butyl Acetate Biosynthesis……..……..……..….180

7.2.1 Screening C4 specific acyl-ACP thioesterases…….……180

7.2.2 Screening AATs for butyl acetate biosynthesis…….……181

7.2.3 Screening a library of new AATs for butyl acetate biosynthesis……..182

7.2.4 The effect of host background and pACYC-PrhaBAD-GmR on butanol biosynthesis……..……..……..……..184

7.2.5 Screening alternative plasmid backbones for expressing AAT……186

7.2.6 AAT feeding assay……..……....……..……..187

7.2.7 Butyl acetate biosynthesis……..……....……..……..188

7.3 Results: Ethyl Butyrate Biosynthesis……..……....……..……..……..190

7.3.1 Proof-of-concept biosynthesis of ethyl butyrate……190

7.3.2 Screening new AAE7s……...191

7.3.3 Screening different host backgrounds for ethyl butyrate biosynthesis……193

7.3.4 Could LiAAT4 and PDC be used to improve ethyl butyrate titres?.....195

7.4 Results: Methyl Ketone Biosynthesis……..198

7.4.1 Could Group 2 ALTs be predicted from protein structure? ……198

7.4.2 Screening Group 2 ALTs……...200

7.4.3 Screening YneP and YtpA……...201

7.5 Results: Unsaturated Alcohol Biosynthesis……204

7.5.1 Screening DfsA……..205

7.5.2 Structure guided mutagenesis of DfsA……..205

7.5.3 Screening DfsA mutants……206

7.6 Discussion……..208

7.6.1 Biosynthesis of butyl acetate……...208

7.6.2 Biosynthesis of ethyl butyrate……..209

7.6.3 Potential for future biosynthesis of butyl butyrate……..210

7.6.4 Activity of the library of new AATs……..210

7.6.5 Biosynthesis of methyl ketones……..211

7.6.7 The application of DfsA for biosynthesis of unsaturated fatty acid derivatives..212

7.7 Conclusions……214

Chapter 8 Conclusions and Recommendations for Future Work……..215

8.1 Conclusions……215

8.2 Recommendations for future work……...218

8.2.1 Improving the C10 thioesterase……...218

8.2.2 Assaying the effect of ACP expression in *Synechocystis*……218

8.2.3 Improving BM3 activity……...218

8.2.4 Expanding the lactone biosynthesis pathway to other chain length products…219

8.2.5 Biosynthesis of unsaturated fatty acid derived products……..220

8.2.6 Porting synthetic metabolic pathways to cyanobacteria……...220

8.2.7 Development of cyanobacteria biotechnology……...221

8.2.8 Transition to a bio-based economy……..221

References……..223

Appendices……261

Appendix A: *Escherichia coli* strains used for bioproduction and feeding experiments in this work……...261

Appendix B: *Synechocystis* Sp. PCC 6803 strains used for bioproduction experiments in this work…….....271

Appendix C: Primers…….....272

Appendix D: Preparation of partially double-stranded linkers for BASIC assembly……276

Appendix E: Partially double stranded linker sequences…….277

Appendix F: List of expression vectors used in this work…….278

Appendix G: List of genes sequences used in this work…….293

List of Figures

Figure 2.1 The E. coli fatty acid biosynthesis pathway……..30

Figure 2.2 The β-oxidation pathway……31

Figure 2.3 γ- and δ-lactone structure……32

Figure 2.4 Biotransformation of ricinoleic acid to γ-decalactone……..33

Figure 2.5 Proposed pathway for biosynthesis of γ- and δ-decalactone……34

Figure 2.6 Regioselectivity of BM3……38

Figure 2.7 Acetate and ethyl ester structure……40

Figure 2.8 Proposed pathway for biosynthesis of butyl acetate……42

Figure 2.9 Proposed biosynthetic pathway for ethyl butyrate……43

Figure 2.10 Methyl ketone biosynthesis pathway in *Penicillium*……..45

Figure 2.11 Methyl ketone biosynthesis pathway in plants……47

Figure 2.12 Green leaf volatile biosynthesis pathway……50

Figure 2.13 Proposed cis-2-hexenal and cis-2-hexenol biosynthetic pathway……..51

Figure 4.1 Pathway and plasmid design for screening CvFatB1……..83

Figure 4.2 Screening CvFatB1 truncations……..84

Figure 4.3 Alignment of CvFatB1 and UaFatB1……..85

Figure 4.4 Screening UaFatB1 truncations……87

Figure 4.5 Plasmid design for expressing CvFatB1 truncations in *Synechocystis*……88

Figure 4.6 Screening CvFatB1 truncations in *Synechocystis*……...89

Figure 4.7 Plasmid design for inducer optimisation of CvFatB1 truncations……..90

Figure 4.8 Inducer optimisation of CvT3 expression……..91

Figure 4.9 Inducer optimisation of CvT4 expression……..92

Figure 4.10 Inducer optimisation of CvT5 expression……93

Figure 4.11 Inducer optimisation of CvT6 expression……94

Figure 4.12 Inducer optimisation of CvT7 expression……95

Figure 4.13 Inducer optimisation of CvT8 expression……96

Figure 4.14 Optimised expression of CvFatB1 truncations……...97

Figure 4.15 Plasmid design for ACP expression……99

Figure 4.16 ACP expression in combination with CvT3……..100

Figure 4.17 ACP expression in combination with CvT4……..101

Figure 4.18 ACP expression in combination with CvT5……..102

Figure 4.19 ACP expression in combination with CvT6……..103

Figure 4.20 ACP expression in combination with CvT7……..104

Figure 4.21 ACP expression in combination with CvT8……..105

Figure 4.22 CvT5 predicted structure……..……..106

Figure 4.23 Screening and inducer optimisation of CvT5 M90G……107

Figure 4.24 Screening and inducer optimisation of CvT5 W122G……108

Figure 4.25 Screening and inducer optimisation of CvT5 W140G……109

Figure 4.26 Screening and inducer optimisation of CvT5 W122 mutants……111

Figure 5.1 Decanoic acid feeding assay……..……120

Figure 5.2 Assaying methyl decanoate as an alternative substrate for feeding assays…….122

Figure 5.3 Assaying the effect of sample acidification……123

Figure 5.4 Assaying the effect of a decane overlay……125

Figure 5.5 M9 media optimisation……..……127

Figure 5.6 Partitioning of decanoic acid, δ-decalactone and γ-decalactone……128

Figure 5.7 Screening new CYP450s……..……..……129

Figure 5.8 The effect of ΔiscR……..……..……..……131

Figure 5.9 Screening BM3 mutants……..……..……133

Figure 5.10 Screening BM3 F87 mutants……..……..……134

Figure 5.11 Adding activity enhancing mutations to BM3 F87G……136

Figure 5.12 Extension of the decalactone biosynthesis pathway……138

Figure 5.13 Screening AlkK and PpAAT1……139

Figure 5.14 Effect of increasing oxygen supply……141

Figure 5.15 Recycling NADPH using GDH……142

Figure 5.16 Co-expression of BM3 and GDH……143

Figure 5.17 Supplementation of M9-Fe with caffeine……145

Figure 5.18 Screening knockout of ΔiscR in E. coli ΔfadD……146

Figure 5.19 Screening new BM3 activity enhancing mutations……148

Figure 6.1 Pathway and plasmid design for δ- and γ-decalactone biosynthesis……157

Figure 6.2 *De novo* biosynthesis of decalactones……158

Figure 6.3 Effect of ferric ammonium citrate and a second plasmid on decanoic acid titres...159

Figure 6.4 New plasmid designs for δ- and γ-decalactone biosynthesis……160

Figure 6.5 RBS Optimisation of CvT5 W122F expression……161

Figure 6.6 Biosynthesis of decalactones using new BM3 mutants with and without a decane overlay……163

Figure 6.7 Screening hybrid CYP450s……164

Figure 6.8 New plasmid design for expression of CvT5 W122F and EcACP……165

Figure 6.9 RBS Optimisation of CvT5 W122F expression with the new operon order……166

Figure 6.10 Supplementing M9 with cas amino acids……167

Figure 6.11 RBS optimisation of EcACP expression……168

Figure 6.12 RBS optimisation of BM3 expression……169

Figure 6.13 Plasmid design for expression of CvT5 and BM3 in Synechocystis……170

Figure 6.14 De novo biosynthesis of decalactones in Synechocystis……171

Figure 7.1 Pathway and plasmid design for butyl acetate biosynthesis……180

Figure 7.2 Screening C4 thioesterases……181

Figure 7.3 Screening AATs for butyl acetate biosynthesis……182

Figure 7.4 Screening new AATs for butyl acetate biosynthesis……184

Figure 7.5 Screening effect of host background and pACYC-PrhaBAD-GmR on butanol biosynthesis……185

Figure 7.6 Screening alternative plasmid backbones for expressing AAT……187

Figure 7.7 AAT feeding assay……188

Figure 7.8 Butyl acetate biosynthesis……189

Figure 7.9 Pathway and plasmid design for ethyl butyrate biosynthesis……190

Figure 7.10 Proof-of-concept biosynthesis of ethyl butyrate……191

Figure 7.11 Screening new AAE7s……193

Figure 7.12 Screening different host backgrounds for ethyl butyrate biosynthesis……195

Figure 7.13 Construct design for expressing AAT and PDC……195

Figure 7.14 Effect of LiAAT4 and PDC expression on ethyl butyrate titre……197

Figure 7.15 Pathway and plasmid design for methyl ketone biosynthesis……198

Figure 7.16 Phylogeny of predicted ALT structures……199

Figure 7.17 Screening Group 2 ALTs……201

Figure 7.18 Screening YneP and YtpA……203

Figure 7.19 Pathway and plasmid design for cis-2-alkenol biosynthesis……204

Figure 7.20 Screening DfsA……205

Figure 7.21 Structure guided mutagenesis of DfsA……206

Figure 7.22 Screening DfsA mutants……207

List of Tables

Table 2.1 Examples of products biosynthesised by engineered *E. coli*……..29

Table 3.1 Composition of M9 minimal media with 2% glucose……..……..53

Table 3.2 Preparation of antibiotics……..……..……..54

Table 3.3 Composition of BG11-Co media……..……..……..57

Table 3.4 Composition of BG11……..……..……..……..58

Table 3.5 Composition of Trace Minerals solution……..…..58

Table 3.6 pJET cloning reaction mix……..……..62

Table 3.7 Colony PCR reaction mix……..……..62

Table 3.8 PCR conditions for colony PCR……..……..……..63

Table 3.9 Agarose gel percentages and DNA ladders for PCR products of different lengths...63

Table 3.10 Q5 DNA polymerase reaction mix……..…..64

Table 3.11 Q5 DNA polymerase PCR conditions……..……..64

Table 3.12 Phusion DNA polymerase reaction mix……..……..66

Table 3.13 Phusion DNA polymerase PCR Conditions……..……..66

Table 3.14 Annealing temperatures and extension times for generating hybrid CYP450s….67

Table 3.15 Reaction mix for site-directed mutagenesis primer phosphorylation……67

Table 3.16 Reaction mix for site-directed mutagenesis……..68

Table 3.17 Thermocycler program for site-directed mutagenesis……68

Table 3.18 Second thermocycler program for Site directed mutagenesis……..…68

Table 3.19 Reaction mix for preparing double-stranded linkers……..69

Table 3.20 Reaction mix for the BASIC clip reaction……70

Table 3.21 Incubation conditions for the BASIC clip reaction……70

Table 3.22 Reaction mix for the BASIC assembly reaction……...71

Table 4.1 CvFatB1 Truncations……82

Table 4.2 UaFatB1 Truncations……85

Table 4.3 Optimum IPTG concentrations for CvFatB1 truncation expression……96

Table 6.1 RBSs used for gene expression optimisation……...160

Table 7.1 New AAT library……...183

Chapter 1

Introduction

1.1 Motivation

1.1.1 The need for sustainability

The global population is growing and becoming richer. As it does so demand for resources such as energy, food and commodities will rapidly increase. However, we live in a world of finite resources, and concerns about limited resources, along with the environmental unsustainability of our current way of life has led to an increased demand for the renewable production of fuels, chemicals and commodities[1]. Furthermore, climate change linked to the emission of greenhouse gases has led to concerns about the use of fossil fuels as an energy source[2]. Fossil fuels are not only used to generate energy, but are also used in petrochemical refinery processes to produce chemicals and everyday materials[3], further contributing to climate change.

So far, the world has already experienced 1.1°C of warming above pre-industrial levels. With this degree of warming, it is possible that we could already experience multiple climatic tipping points; collapse of the Greenland ice sheet, collapse of the West Antarctic ice sheet, breakdown of the Labrador-Irminger sea convection current, abrupt thaws of Boreal permafrost and die offs of low latitude coral reefs. Currently, under the terms of the 2015 Paris Agreement, signatories will aim to limit global warming to 1.5-<2°C. With this amount of warming the previously mentioned climatic tipping points will become likely, and further tipping points become possible. Even more concerningly, with current policies global warming is expected to exceed the Paris Agreement target and reach 2.6°C[4]. The warming we have experienced so far is already causing climate change, with extreme weather events becoming more likely. For example, the August 2022 heatwave in the United Kingdom where recorded temperatures exceeded 40°C for the first time is thought to have been made 10 times more likely by climate change[5].

1.1.2 Sustainable bioproduction using cyanobacteria

Bioproduction could serve as method of producing renewable chemicals, to help meet demand for commodities while reducing emissions[6]. Bioproduction also has some further advantages over traditional chemical synthesis. Chemical production has a high energy use, and often

uses hazardous or toxic chemicals as part of the manufacturing process, and may also produce toxic by-products that need removal. Chemical reactions also have lower regio- and stereo-specificity[7]. On the other hand, bioproduction uses less harmful chemical reagents, and enzyme catalysed reactions are highly regio- and stereo-specific[3].

Bioproduction could also be used to produce sustainable biofuels, which have advantages over other renewable energy sources as they are more compatible with existing infrastructure and are more energy dense[2]. However, a drawback to biofuels is that currently the main feedstock for biofuel production is agricultural products. As the biofuel market expands it poses a risk of helping drive rising food insecurity, by competing for land with food production, or by driving up the prices for agricultural products[8].

Cyanobacteria, which are photosynthetic microalgae, are a promising chassis for sustainable bioproduction as they could serve as “green catalysts” for the production of renewable chemicals from light and CO_2[9]. Cyanobacteria could be used to produce biofuels, as many species accumulate high lipid contents (>70%) in their cells, which is attractive for making biodiesel[2]. Furthermore, some cyanobacteria naturally produce attractive value-added products such as pigments, vitamins, and biopolymers. Cyanobacteria could also be applied to wastewater remediation[10]. Compared to higher plants, cyanobacteria grow faster and are easier to engineer, have a higher photosynthetic efficiency and do not waste biomass on roots or stems[9,11,12]. Engineered cyanobacteria strains have already been reported that biosynthesise fatty acyl-ACP derived products such as biofuels[9,13].

However, there are challenges to using cyanobacteria for sustainable bioproduction. There are two main cultivation methods. Firstly, there are open pond systems which are inexpensive. However, these require large land areas, are vulnerable to contamination, suffer from water loss and it is challenging to control the culture conditions. Secondly, there are closed photobioreactors, which avoid the drawbacks to open pond systems but are very expensive[2]. For bioproduction to be economically viable, the revenue from selling the products needs to cover the costs of production. For cyanobacteria to grow continuous lighting is needed, which requires energy and the associated costs, and growth is also limited by a slow rate of carbon fixation[12]. Meanwhile the unit values of biofuels are very low, meaning current yields are below economic feasibility[12,14]. Therefore, while engineered biofuel producing cyanobacteria strains have been reported in the literature, industrial production using cyanobacteria has not yet been achieved[9]. If commercial cyanobacteria bioproduction is to be realised, there is a need to target production of higher-value products first.

1.1.3 Flavours and Fragrances

Flavours and Fragrances are high value compounds, for example the first microbially-synthesised γ-Decalactone was worth €10,000 /kg when it entered the European market[15].There is a large global market for Flavours and Fragrances that was valued at $27.1 billion in 2016 and predicted to reach $33.5 billion in 2021. The market is also rapidly growing at a compound annual growth rate of 2.1%[7].

Flavours and Fragrances are typically made by extraction from plant biomass or by chemical synthesis[16,17]. Within the Flavour and Fragrance market there is a "Naturals" subsection, for which there is strong customer preference, leading to premium prices that are often higher by a factor of 10 or more[18-21]. Chemically synthesised Flavours and Fragrances are sometimes labelled as "nature-identical", but this is seen as unfriendly to consumers[21,22]. If the Flavour and Fragrance compound is produced by biotechnology, it can be marketed and sold as "Natural", according to EU and US legislation, so long as the fed substrate occurs in nature and the final product can be found in raw natural materials[18-21].

The premium Naturals subsection therefore presents a market opportunity for cyanobacteria, as the high prices could allow for bioproduction using cyanobacteria to be economically viable. Another advantage for cyanobacteria bioproduction over chemical synthesis in the Flavour and Fragrance market is the stereospecificity of enzymatic reactions, as particular enantiomers of a Flavour and Fragrance compound can have a stronger scent than their counterpart[23], making the ability to specifically synthesise a certain isomer desirable.

There are a wide range of Flavour and Fragrance compounds, with key Flavours and Fragrances in fruit aromas and artificial fragrances including alcohols, aldehydes, ketones, esters, lactones and terpenoids[19,24,25]. Aliphatic aldehydes, ketones and esters are more volatile than their counterpart alcohols, resulting in stronger aromas. Presence of specific compounds is positively associated with consumer preference for different products, for example presence of butyl acetate, ethyl butyrate, γ-Decalactone and 2-heptanone are positively correlated with customer preferences in strawberries[26].

1.1.4 Prototyping biosynthetic pathways in *Escherichia coli*

The slow growth of cyanobacteria poses a challenge to their engineering as it increases the time required for a Design-Build-Test-Learn (DBTL) cycle, meaning fewer iterations can be achieved within a given time period. This problem becomes especially acute when implementing synthetic biosynthetic pathways where multiple components require engineering. *E. coli* is fast growing and a well-established chassis for synthetic biology[27]. *E. coli* can therefore be used as a biological prototype to validate and optimise synthetic metabolic

pathways and their components. By using *E. coli*, the time taken for a DBTL cycle can be reduced, allowing for more iterations and optimisation. Once optimised, the pathway will be ported to cyanobacteria to validate its functionality in a photosynthetic host and enable renewable biosynthesis of the target compound from sunlight and CO_2. This approach has already been used by the Jones group, demonstrating that synthetic pathways that are functional in *E. coli* can be transferred to cyanobacteria[28].

1.2 Aims

1) To identify key Flavour and Fragrance compounds to target for bioproduction and design metabolic pathways for their biosynthesis

2) To validate key enzymes required for the biosynthetic pathway and if necessary, engineer them for improved activity and/or specificity

3) To validate and if necessary, optimise the biosynthetic pathways in *E. coli*

4) To port the biosynthetic pathways to the cyanobacteria *Synechocystis* Sp. PCC 6803 to validate activity in a photosynthetic host.

1.3 Implementation

A comprehensive literature review was performed to identify target Flavour and Fragrance compounds, their native biosynthetic pathways and if/how they had been previously biosynthesised in heterologous hosts. Genes were also identified that could be applied to the creation of novel metabolic pathways for the biosynthesis of target compounds, with the novel pathways having advantages over existing previously reported pathways such as requiring fewer genes, therefore reducing cellular burden, or by being more carbon efficient.

Individual genes for the proposed biosynthetic pathways were then validated in *E. coli* by expressing the gene and feeding the gene's substrate or incorporating the gene into a synthetic metabolic pathway where the other enzymes' function had already been validated. If the gene required engineering for altered substrate specificity or increased activity, this was achieved by incorporating previously reported mutations to the gene from the literature or by using the protein structure to design new mutations.

The genes were then incorporated into the proposed biosynthetic pathways and the pathway validated in *E. coli*. Optimisation was carried out by controlling gene expression levels by altering the amount of inducer used or by using RBSs with different strengths, or by investigating alternative operon designs. Finally, validated biosynthetic pathways were ported to cyanobacteria.

1.4 Thesis organisation

Chapter 2 presents a literature review on *E. coli* as a model organism for synthetic biology and metabolic engineering, decalactones, butyl acetate and ethyl butyrate, methyl ketones and green leaf volatiles, and their respective biosynthetic pathways, previously reported attempts at their biosynthesis, and the genes used in their proposed biosynthetic pathways in this study.

Chapter 3 presents all the materials and methods used in this work.

Chapter 4 presents the engineering of a C10 specific acyl-ACP thioesterase.

Chapter 5 presents the engineering of BM3 for the ω-5 and ω-6 hydroxylation of decanoic acid, and optimisation of culture conditions for expression of BM3.

Chapter 6 presents the validation of a synthetic metabolic pathway for decalactone biosynthesis and work to optimise the pathway in *E. coli*, along with porting the pathway to *Synechocystis* Sp. PCC 6803.

Chapter 7 presents work prototyping synthetic pathways the biosynthesis of butyl acetate, ethyl butyrate, methyl ketones and green leaf volatiles in *E. coli*

Chapter 8 concludes key results from this work and provides recommendations for future work.

Chapter 2

Literature Review

2.1 *Escherichia coli*

2.1.1 *E. coli* as a model organism for synthetic biology

E. coli is a Gram-negative facultative anaerobe[27]. There is a deep pool of knowledge about *E. coli*, with the whole genome having been sequenced for multiple strains[29,30] and annotated[31,32], and a central database, EcoCyc, developed that provides information on the metabolites, genes, metabolic pathways, and regulatory networks found within *E. coli*[33-37]. This is complemented by the Keio collection, a set of 3985 different gene knock-out strains of *E. coli*[38]. There are also well-established molecular biology techniques for the engineering of *E. coli*, such as electroporation[39], transformation of chemically competent cells by heat shock[40], and knock-out or chromosomal integration of genes using λ-Red recombinase[41,42]. Recently, CrispR has become an alternative tool to λ-Red recombinase, and can even allow for the editing of multiple genes at once[43]. Furthermore, *E. coli* is natively able to grow on a wide variety of carbon sources ranging from carbohydrates such as glucose and lignocellulosic biomass derived sugars, to glycerol and fatty acids. *E. coli* is also able to grow and metabolise substrates under both aerobic and anaerobic conditions[27,44].

This combination of knowledge about *E. coli*'s genetic makeup, established tools for genetic engineering, ease of culturing and a short-doubling time have led *E. coli* to become a workhorse for the development of synthetic biology and metabolic engineering[27,45]. Classic synthetic biology toy circuits such as the bi-stable toggle switch, where two mutually inhibiting promoters can be flipped between two stable states by transient induction[46], or the repressilator, where a network of three repressible promoters lead to the periodic transient expression of GFP[47], were first developed in *E. coli*. More complex circuits incorporating logic were also first reported in *E. coli*[48]. Beyond toy circuits, *E. coli* has been engineered to biosynthesise a truly diverse array of products, including aromatics, dicarboxylic acids, fatty acid derivatives, flavonoids, isoprenoids and terpenoids, lactams, organics acids, polyhydroxyalkanoates and short-chain alcohols[27,44,49-52]. It is beyond the scope of this review to provide a truly comprehensive list; however, some examples are provided in Table 2.1.

Product	Class	Titre	Culture Volume	Time	Reference
Cinnimaldehyde	Aromatic	75 mg/L	50 mL Shake Flask	48 hrs	53
Shikimate	Aromatic	27.41 g/L	5 L Bioreactor	48 hrs	54
Riboflavin	Vitamin	2.7 g/L	50 mL Shake Flask	60 hrs	55
Isoprene	Isoprenoid	320 mg/L	100 mL Shake Flask	26 hrs	56
τ-Caudinol	Terpenoid	133.5 mg/L	100 mL Shake Flask	40 hrs	57
Octanoic Acid	Aliphatic Acid	1 g/L	300 mL Bioreactor	120 hrs	58
Octyl Glucoside	Alkyl Glucoside	550 mg/L	25 mL Shake Flask	48 hrs	28
Propane	Alkane	32 mg/L	20 mL Serum Bottle	19 hrs	14
2-Pentanone	Methyl Ketone	240 mg/L	3 mL Vacutainer	72 hrs	59
L-Malate	Dicarboxylic Acid	21.65 g/L	5L Bioreactor	72 hrs	60

Table 2.1 Examples of products biosynthesised by engineered E. coli

2.1.2 Fatty acid biosynthesis and β-oxidation

All target Flavour and Fragrance compounds in this work will be biosynthesised using *E. coli*'s native fatty acid biosynthesis as a starting point. Prokaryotes such as *E. coli* have a Type II fatty acid biosynthesis system where the steps in the biosynthetic cycle are carried out by individual enzymes[61,62].

The first step in fatty acid biosynthesis is the carboxylation of acetyl-CoA to malonyl-CoA. This is catalysed by the multimeric protein AccABCD (acetyl-CoA carboxylase). The reaction proceeds by the ATP dependent carboxylation of biotin to form carboxybiotin, followed by transfer of carbonate from the carboxybiotin to acetyl-CoA, to form malonyl-CoA[63,64]. The second step is catalysed by FabD, (malonyl-CoA:ACP transacylase) which transfers malonate from CoA to ACP[65,66]. FabH (β-ketoacyl-ACP synthase III) then catalyses the condensation of malonyl-ACP with acetyl-CoA, forming β-ketohexanoyl-ACP[67].

From this point on, the fatty acid biosynthesis pathway becomes cyclical. The β-ketoacyl-ACP is reduced by FabG (β-ketoacyl-ACP reductase) to β-hydroxyacyl-ACP in an NADPH dependent reaction[68]. B-hydroxyacyl-ACP is then dehydrated by either FabA (β-hydroxyacyl-ACP dehydratase/isomerase) or FabZ ((β-hydroxyacyl-ACP dehydratase) to form 2-enoyl-ACP. FabA and FabZ have different substrate specificities, FabA is active on medium chain length saturated and unsaturated β-hydroxyacyl-ACP, and short and long chain length saturated β-hydroxyacyl-ACP. FabZ is active on short chain and long chain length saturated and unsaturated β-hydroxyacyl-ACP. Additionally, FabA can also reversibly isomerise 2-decenoyl-ACP to 3-decenoyl-ACP. This reaction is the branching point between saturated and unsaturated fatty acid biosynthesis, and FabA cannot perform this isomerisation reaction on other chain length 2-enoyl-ACPs[61,62,69,70]. Enoyl-ACP is then reduced to acyl-ACP by FabI (Enoyl-ACP reductase) in an NADH dependent reaction[71]. The final step in the cycle

is performed by either FabB (β-ketoacyl-ACP synthase I) or FabF, (β-ketoacyl-ACP synthase II) which catalyses the condensation of acyl-ACP and acetyl-CoA to reform a molecule of β-ketoacyl-ACP that is 2 carbon atoms longer than at the beginning of the cycle. FabB can additionally use 3-decenoyl-ACP as a substrate for condensation with acetyl-CoA, to biosynthesise unsaturated fatty acids[72,73].

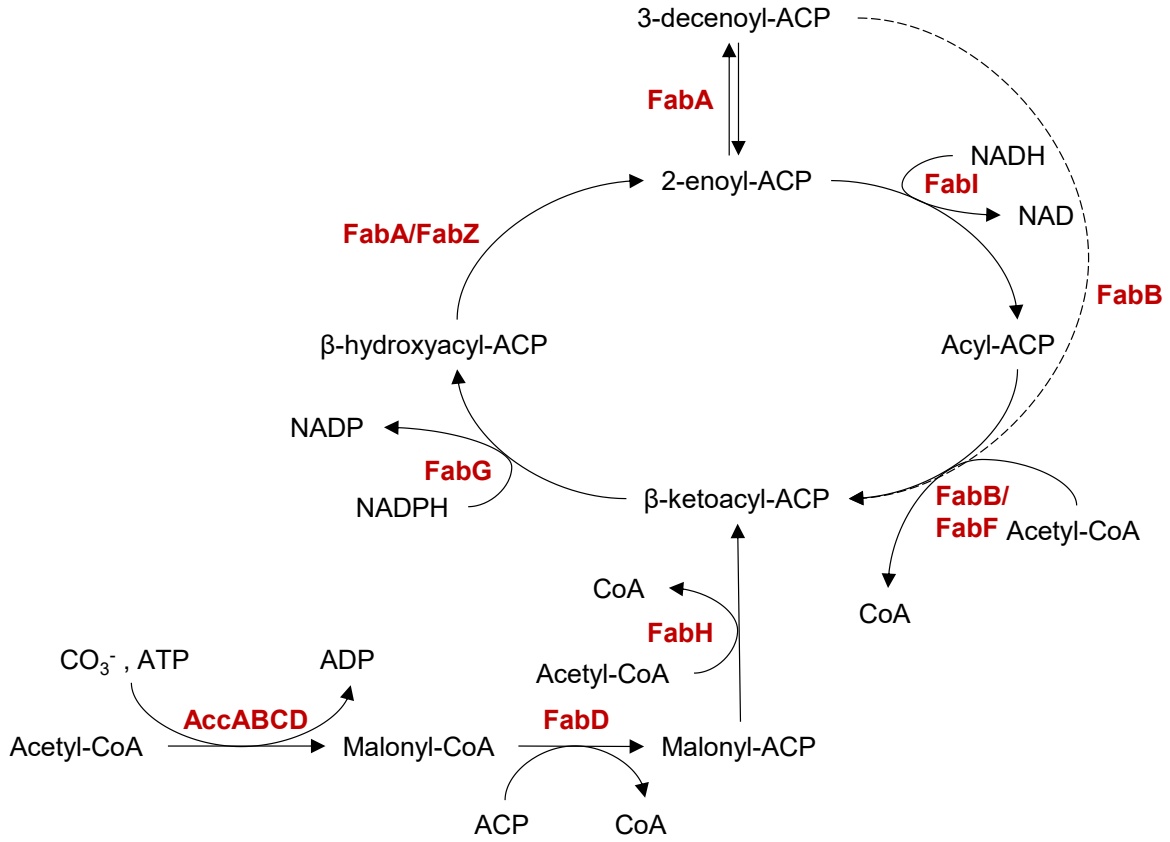


Figure 2.1 The E. coli fatty acid biosynthesis pathway *This figure is adapted from the literature[62]. Pathway enzymes are highlighted in red. AccABCD: Acetyl:CoA carboxylase, FabD: Malonyl-CoA:ACP transacylase, FabH: β-ketoacyl-ACP synthase III, FabG: β-ketoacyl-ACP reductase, FabA: β-hydroxyacyl-ACP dehydratase/isomerase, FabZ: β-hydroxyacyl-ACP dehydratase, FabI: Enoyl-ACP reductase, FabB: β-ketoacyl-ACP synthase I, FabZ: β-ketoacyl-ACP synthase II.*

Some target Flavour and Fragrance compounds in this work have previously been synthesised using the β-Oxidation pathway for the degradation of fatty acids. This pathway begins with the ATP dependent ligation of free fatty acids with CoA by FadD or FadK (Fatty Acid CoA ligases). FadD is active on C10 and longer fatty acid substrates, while FadK is anaerobically expressed and is active on C6 and C8 fatty acids[74,75]. From this point on β-Oxidation proceeds in a cyclical series of reactions akin to those of fatty acid biosynthesis. The first step in the cycle is FadE (acyl-CoA dehydrogenase) which catalyses the FAD dependent dehydrogenation of acyl-CoA to enoyl-CoA[76]. The next steps in the cycle are catalysed by FadB, which is a multifunctional enzyme. FadB first acts as an enoyl-CoA hydratase and catalyses the hydration of enoyl-CoA

to β-Hydroxyacyl-CoA. FadB then acts as a β-Hydroxyacyl-CoA dehydrogenase and catalyses the NAD dependent dehydrogenation of β-Hydroxyacyl-CoA to β-Ketoacyl-CoA. FadB also has cis-Δ3:trans-Δ2-enoyl-CoA isomerase and β-Hydroxyacyl-CoA epimerase activities for the degradation of unsaturated fatty acids[77]. The final step in the β-Oxidation cycle is a catalysed by FadA (β-Ketoacyl-CoA thiolase), which produces acetyl-CoA and a new molecule of acyl-CoA that is two carbons shorter[78].

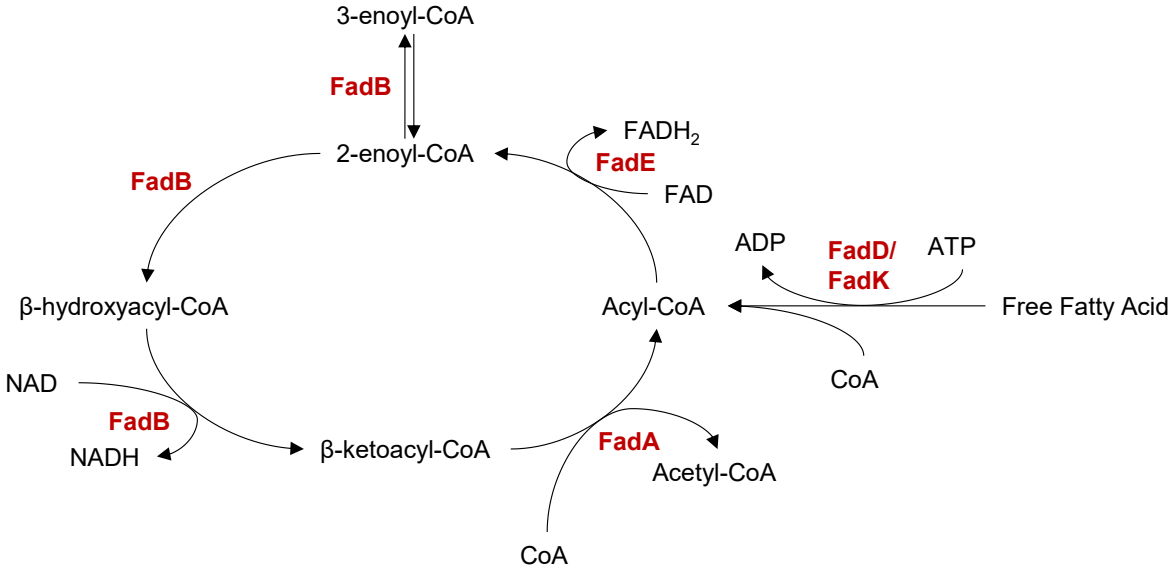


Figure 2.2 The β-oxidation pathway *Figure adapted from the literature[62]. Pathway enzymes are highlighted in red. FadD: Fatty Acid CoA Ligase, FadK: Fatty Acid CoA Ligase, FadE: Aycl-CoA Dehydrogenase, FadB: Enoyl-CoA Hydratase, β-Hydroxyacyl-CoA Dehydrogenase, cis-Δ3:trans-Δ2-enoyl-CoA Isomerase and β-Hydroxyacyl-CoA Epimerase, FadA: β-Ketoacyl-CoA Thiolase.*

2.2 Lactone Biosynthesis

2.2.1 Lactones

Lactones are cyclic esters formed by the spontaneous condensation of hydroxyfatty acids. Γ- and δ-lactones are respectively formed from the condensation of 4- and 5-hydroxyfatty acids and are the most stable kinds of lactones (Figure 2.3)[79,80]. Medium chain length lactones are important and potent aroma compounds and are found in the aromas of more than 120 foodstuffs, such as fruit, and are used in the Flavour and Fragrance industry[18,80-82]. The native biosynthetic pathways for γ- and δ-lactones are not well understood, and while it is known that 4- and 5-hydroxyfatty acids are precursors to γ- and δ-lactone biosynthesis, how the precursors themselves are biosynthesised is unknown.

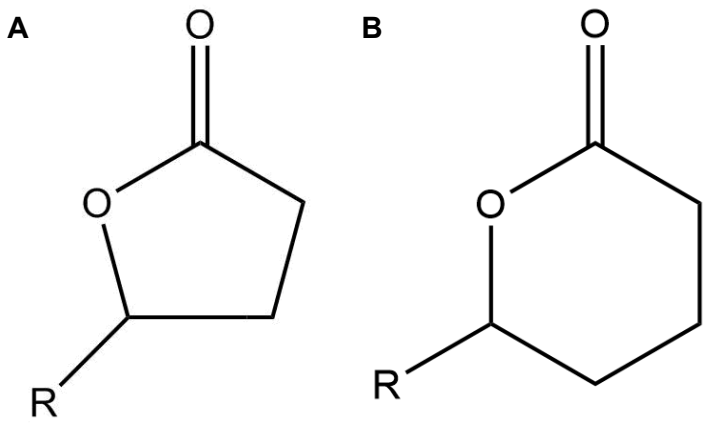


Figure 2.3 γ- and δ-lactone structure *A) γ-lactone structure B) δ-lactone structure.*

The most widely produced and commercially important lactone is γ-decalactone[83]. Γ-Decalactone is an important component of peach, apricot, and strawberry aromas, and has an odour usually described as having a powerful peach-like scent, and has also been reported to have an aroma of coconut[79,84-86]. Both R- and S-enantiomers are found in natures, with the S-enantiomer found in mangoes and the R-enantiomer found in peaches and other fruits[21]. There is a large global demand for γ-decalactone, with a global market size of 10,000 kg in 1997[87]. It is also high value, with the first microbially synthesised γ-decalactone worth €10,000 /kg, and after optimisation of production the price dropped to $300 /kg[15,84]. Unsaturation of the ring structure can also be used to modify the aroma, for example unsaturation at the Δ3 position results in an even stronger peach aroma. However, it is difficult to introduce the double bond at the correct position, and unsaturation at the Δ2 position results in a mushroom aroma. Due to this challenge unsaturated γ-decalactones are not produced industrially[79]. The closely related δ-decalactone has a similar aroma to γ-decalactone, that has been reported as being intensely fruity, peach-like and having a peach-coconut odour[17,86]. Like

γ-decalactone, both enantiomers are found in nature, with the S-enantiomer found in raspberries and the R-enantiomer found in peaches[23].

2.2.2 Previously reported biosynthesis of γ- and δ-decalactone

So far previously reported biosynthesis of decalactones has focused on γ-decalactone, with the majority of industrial scale production of γ-decalactone using *Yarrowia lipolytica* to biotransform ricinoleic acid (12-hydroxy-9-octadecenoic acid) to γ-decalactone[88-91]. Methyl ricinolate can also be used as an alternative substrate to ricinoleic acid[83,91]. The fed substrate enters the β-oxidation pathway[85], and after multiple rounds of the β-oxidation cycle 4-hydroxydecanoyl-CoA is produced. It is not clear how the γ-decalactone is then produced from the 4-hydroxydecanoyl-CoA (Figure 2.4).

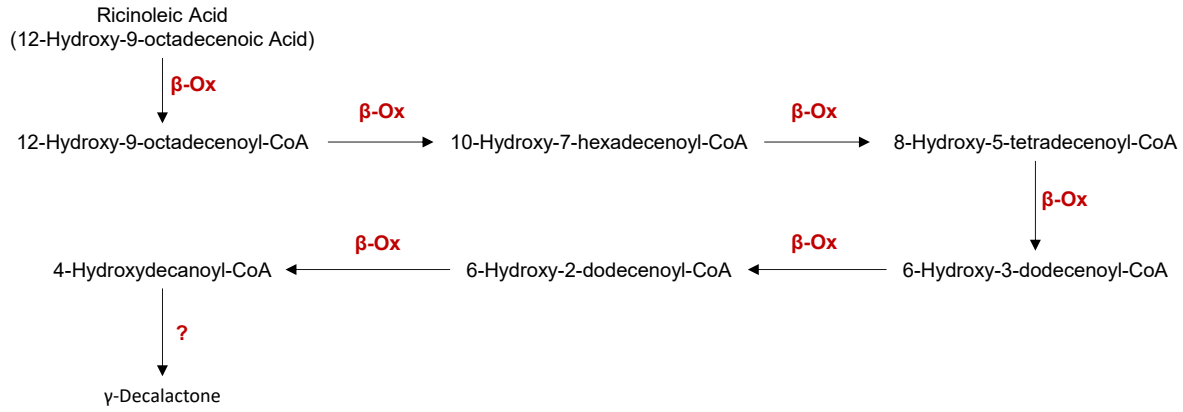


Figure 2.4 Biotransformation of ricinoleic acid to γ-decalactone *Pathway steps are highlighted in red. β-Ox: β-Oxidation, ?: Exact details of the pathway steps and catalysis between these two compounds is not known.*

This concept of feeding a long-chain fatty acid substrate into the β-oxidation cycle has since been expanded to other pure substrates such as oleic acid and 12-hydroxystearic acid, and mixed substrates such as grapeseed and peanut oil, and to other host species including bacteria, with a range of medium chain length γ- and δ-lactones being produced[81,82,92-95]. This concept was then further expanded to achieve *de novo* biosynthesis of lactones by engineering the native fatty acid biosynthesis of *Ashbya gossypii* to overproduce linoleic acid, which then entered the β-oxidation pathway, producing a range of γ-lactones[90].

Alternative methods of biosynthesising lactones have also been reported. Glycolic acid has been used as a substrate for the biosynthesis of 3-hydroxy-γ-butyrolactone[96,97], and CYP450s have been used to hydroxylate decanoic acid and dodecanoic acid at the C5 position, with the resulting hydroxyfatty acids condensing under acidic conditions to form δ-decalactone and δ-dodecalactone[98,99]. Production of δ-decalactone has also been reported following the incubation of baker's yeast with massoia lactone, an unsaturated derivative of

δ-decalactone[100]. Various medium chain length γ- and δ-lactones were also produced by incubating oxofatty acids with cell extracts from *E. coli* that had expressed a carbonyl reductase, forming 4- and 5-hydroxyfatty acids that were then lactonized using acidic and high heat conditions[101].

2.2.3 Proposed pathway for γ- and δ-decalactone biosynthesis

In this work, a novel metabolic pathway for the biosynthesis of γ- and δ-decalactone is proposed. The pathway is comprised of a C10 acyl-ACP thioesterase to produce free decanoic acid from endogenous fatty acid biosynthesis. A CYP450 is then used to hydroxylate the decanoic acid at the 4- or 5- position. The 4- or 5-hydroxydecanoic acids then spontaneously condense to respectively form the γ- and δ-decalactones. The pathway can be expanded upon by the expression of Glucose Dehydrogenase (GHD) for co-factor recycling, and should the condensation prove to not be spontaneous, by the expression of AlkK and PpAAT1 to ligate the 4- and 5-hydroxydecanoic acids to CoA and catalyse the formation of γ- and δ-decalactone from 4- and 5-hydroxydecanoyl-CoA (Figure 2.5).

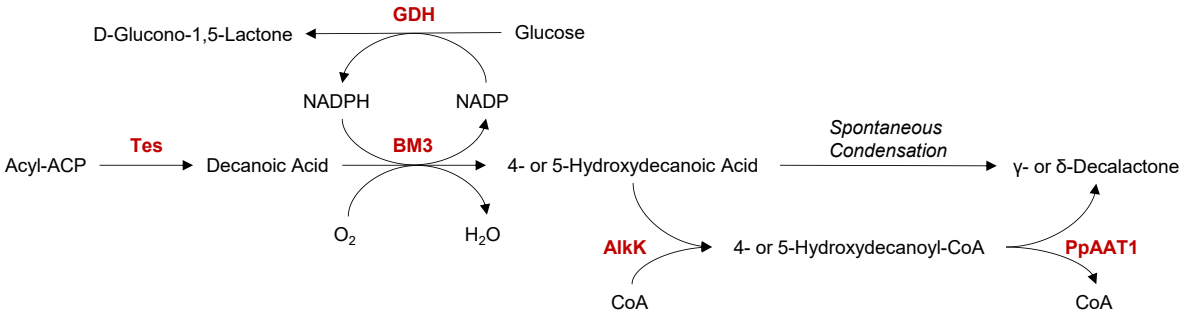


Figure 2.5 Proposed pathway for biosynthesis of γ- and δ-decalactone *Pathway enzymes are highlighted in red. Tes: C10 Acyl-ACP Thioesterase, BM3: CYP450 CYP102A1 (Priestia megaterium), GDH: Glucose Dehydrogenase, AlkK: Hydroxyfatty acid-CoA ligase, PpAAT1: Alcohol-Acyl Transferase (Prunus persica).*

2.2.4 Acyl-ACP thioesterases

The first step in the proposed biosynthetic pathway is catalysed by an acyl-ACP thioesterase. These enzymes terminate fatty acid biosynthesis by hydrolysing the thioester bond between ACP and the growing acyl chain[102]. Different thioesterases have different specificities for acyl-ACP of various chain lengths. In order to biosynthesise decalactones, decanoic acid will need to be first produced. To do so, a C10 specific thioesterase will be required.

Microbial acyl-ACP thioesterases tend to be active across a wide range of acyl-ACP chain lengths while plant acyl-ACP thioesterases tend to be more specific[103,104]. Plant thioesterases are classified into two main families, FatA and FatB. FatA acyl-ACP thioesterases are specific

for unsaturated C18 acyl-ACP, while FatB acyl-ACP thioesterases are specific for unsaturated acyl-ACP of chain lengths ranging from C8 to C18[104,105]. The plant genus *Cuphea* is known for the high proportion of medium-chain length fatty acids contained within their seeds, and while the seed's fatty acid distribution profile doesn't perfectly match thioesterase specificity, as the activities of other enzymes also contribute to the profile[105-108], FatB thioesterases from the genus *Cuphea* are therefore a promising source for a C10 specific acyl-ACP thioesterase.

While many FatB thioesterases have been reported[104-106,109-122], only seven FatB thioesterases with medium chain length specificity have been reported, and of these five originated in *Cuphea*[102,104,112,123-125]. Fatty acid biosynthesis proceeds in a cyclical manner, with each turn of the cycle elongating the growing acyl chain by two carbons. However, acyl-ACP thioesterase specificities proceed in units of four, with C4, C8 and C12 acyl-ACP thioesterases reported prior to this work commencing. It is not clear why this was the case and led to the need to engineer a C10 specific acyl-ACP thioesterase. During the course of this work, in 2022, the acyl-ACP thioesterase FatB3 from *Cuphea lanceolata* was reported to be C10 specific (70% specificity) when transgenically expressed in *E. coli*[125]. This is the only C10 acyl-ACP thioesterase to have been reported in the literature, however by the time of publication work on engineering our own C10 specific acyl-ACP thioesterase had already been completed.

FatB thioesterases have N-terminal chloroplast localisation peptides, with truncation needed required to form the mature polypeptide[105,109]. Upon heterologous expression in bacteria, this N-terminal chloroplast localisation peptides renders the thioesterase insoluble, causing a loss of activity. Truncation of the thioesterase is therefore required prior to transgenic expression[126]. While it is known that truncation is required, where the site of truncation should be is not. Truncating the same thioesterase at different locations leads to different activity and specificity profiles[126-128], and so optimisation of truncation should be considered a necessary step when transgenically expressing FatB thioesterases in bacteria.

There has been work on engineering FatB thioesterases to have altered activity and specificity. Through random mutagenesis, a highly active and C8 specific variant of *Cuphea palustris* FatB1 was obtained, which had a unique truncation and had two point mutations[126]. Introducing the same truncation and point mutations onto the thioesterases *Cuphea avigera var pulcherrima* FatB3 and *Cuphea lanceolata* FatB3 also led to increases in activity[28,125]. *Cuphea hookerina* FatB2 was engineered to be more C10 specific by incorporating a point mutation that modulated the interactions between the thioesterase and ACP. While this increased the proportion of decanoic acid from 26.6 to 60% of total fatty acid production, there was also a 4-fold decrease in specific activity on decanoyl-ACP and overall fatty acid production dropped from 34 mg/L to 22 mg/L[129]. Another approach has been to alter the

specificity by engineering the binding pocket, either by introducing residues from another FatB thioesterase with a different chain-length specificity or by mutating residues to enlarge the binding pocket, and this approach has been used several times to generate mutant FatB thioesterase with increased specificity for longer chain-length substrates[120,130,131].

Domain swapping has also been used as a tool to investigate and modulate FatB thioesterase activity and specificity. In early work, swapping the active site of UcFatB1 and GarmFatA1 led to hybrid thioesterases with altered substrate specificity[132], and in contrasting results it was reported that either the N-terminal domain[133] or C-terminal two-thirds[120] of the protein that were key to activity. Later, more systematic studies split FatB thioesterases into six domains, arranged N- to C-terminal, allowing for a more granular investigation of the effects of domain swapping. These studies found that swapping Domains 1-4 had an effect on activity and specificity, with Domains 2-4 having the greatest effect[134,135]. These results can serve to explain the discrepancies in early reports, because depending on where the earlier authors decided to divide the thioesterases into N- and C-terminal domains, Domains 2-4 may have been assigned to either the N- or C-terminal. It was also observed that Domain 5 had an effect on protein expression in *E. coli*, and that by introducing domain 5 from *Cuphea palustris* FatB2, which was highly expressed in *E. coli*, into *Cuphea palustris* FatB1, expression of *Cuphea palustris* FatB1 could be increased[135]. During these studies, a hybrid thioesterase with domains from *Cuphea viscosissima* FatB1 and FatB2 was generated which showed increased C10 specificity (43% of total fatty acid product compared to 26.6% for the wild-type *Cuphea viscosissima* FatB1), however this came at a cost of an ~10-fold decrease in activity[134].

FatB thioesterases have been widely used to produce free fatty acids in various heterologous hosts such as plants[136,137], eukaryotic algae[127,138-140], yeasts[141], cyanobacteria[142] and *E. coli*[58,121,127,143-145]. FatB thioesterases have also been incorporated into synthetic pathways for the biosynthesis of various fatty acid derived products, including alcohols[128,146-150], alkanes[13], alkyl glucosides[28], esters[28,127,151-153], hydroxyfatty acids[146,154] and methyl ketones[127,155,156].

2.2.5 CYP450 BM3

The second step in the proposed decalactone biosynthesis pathway is the hydroxylation of decanoic acid at the 4- and 5-positions, to form the appropriate 4- and 5-hydroxydecanoic acid precursors to γ-decalactone and δ-decalactone.

CYP450 BM3 or CYP102A1 (BM3) is a soluble catalytically self-sufficient mono-oxygenase from *Priestia megaterium* (formerly called *Bacillus megaterium*[157]). BM3 natively hydroxylates C12 to C20 fatty acids at the ω-1, ω-2 and ω-3 positions in a highly regioselective manner (98% R at the ω-1 and ω-2 positions, 72% R at the ω-3 position), using oxygen and NADPH as

co-factors[158-160]. Typically, microbial CYP450s interact with a ferrodoxin reductase to obtain electrons, however in BM3 a haem-binding CYP450 hydroxylase domain is fused to a CYP450 reductase domain containing eukaryotic-like flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN) binding domains. This fusion allows for efficient electron transfer between the reductase and haem domain and for BM3 to be catalytically self-sufficient. Due to this efficiency, BM3 has the highest reported catalytic activity for a P450 mono-oxygenase, and is 100-1000 times faster than eukaryotic systems[158,159]. It has also been found that caffeine can be used to further increase the activity of BM3[161]. BM3 can oxidise both saturated and unsaturated fatty acids, and oxidation of unsaturated fatty acids can also result in the production of epoxyfatty acids. Under excess oxygen conditions, further oxidation can occur producing dihydroxyfatty acids and ketofatty acids[162]. The mechanism by which BM3 binds the fatty acid substrate has also been determined. An R47/Y51 motif at the mouth of the substrate access channel binds the carboxyl group of the fatty acid, and then the alkyl chain enters the active site channel, where hydrophobic amino acid residues stabilise the alkyl chain in the correct orientation and in proximity to the haem group[163,164]. Natural variants of BM3 have also been reported with similar regioselectivities of hydroxylation but differing turnover numbers on dodecanoic acid[165].

Changing the regioselectivity of hydroxylation to the ω-5 and ω-6 positions would enable the production of 4- and 5-hydryoxydecanoic acid and therefore the decalactones (Figure 2.6). Regioselectivity altering mutations at S72Y, V78, F87 and I263 residues were reported, with hydroxylation up to the ω-9 position on dodecanoic acid achieved, however these mutations also led to a loss of activity[166]. Mutations that increase BM3's activity on fatty acids have also been reported in the literature, with different mutations increasing activity on C8-C16 fatty acids[164,167-169]. Mutants with altered substrate specificity have also been observed to also have altered regioselectivities[167-170], meaning combinatorial mutants with regioselectivity-altering and activity-enhancing mutations may not have the expected regioselectivities. Beyond fatty acids, BM3 has been engineered to accept a wide range of unnatural substrates, including alcohols[171], alkanes[127,167,172-177], aromatics[170,172,176-182], flavonoids[183,184], statins[185] and steroids[186,187].

BM3 has previously been expressed in *E. coli* for the bioconversion of long chain fatty acids to the corresponding ω-1 to ω-3 hydroxyfatty acids[188], or incorporated into a synthetic metabolic pathways in *E. coli* with long chain acyl-ACP thioesterases to biosynthesise long chain ω-1 to ω-3 hydroxyfatty acids[154,189].

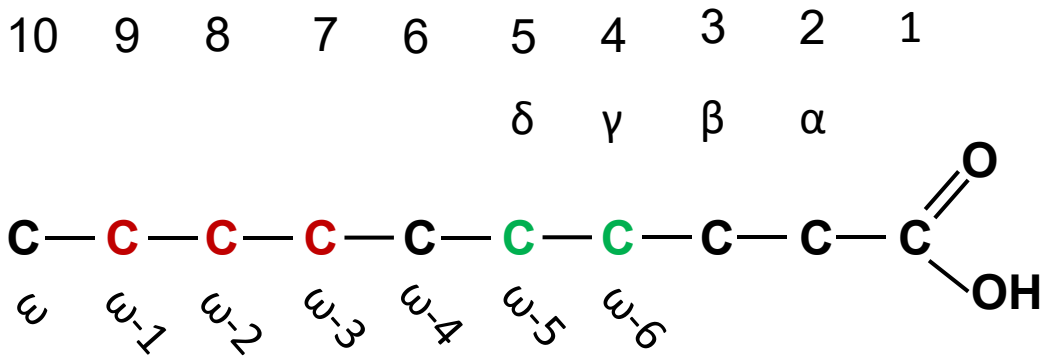


Figure 2.6 Regioselectivity of BM3 *A schematic diagram of decanoic acid is shown, with the carbons numbered, α to δ carbons and ω to ω-6 carbons labelled. BM3 natively hydroxylates the carbons in red, but in order to produce the hydroxyfatty acid precursors to γ- and δ-decalactone the carbons highlighted in green need to be hydroxylated.*

Various self-sufficient CYP450s have since been reported that hydroxylate fatty acids[190-202], however few have either been shown to have activity on decanoic acid, or to hydroxylate other fatty acids with the appropriate regioselectivities[98,99,203-206]. Of these, the most promising is *Tepidiphilus thermophilus* CYP116B46 (TtCYP450). In an *in vitro* assay using cell free extracts, TtCYP450 converted 83% of decanoic acid substrate to hydroxydecanoic acid, with 90% specificity for the ω-5 position[98], by far the highest reported specificity for this position. TtCYP450 differs from BM3 in that its reductase domain contains FMN and an Fe-S cluster[205]. There are crystal structures available[207,208], and simulations have determined how the substrate binding channel opens to accommodate the substrate[209], and how electrons are transferred between the haem and reductase domain[210], making TtCYP450 the best understood CYP450 of this type. Therefore, TtCYP450 represents a promising alternative to BM3 and is also worth investigation.

2.2.6 Glucose dehydrogenase

Regeneration of the NADPH cofactor could be used to increase the efficiency of the proposed pathway. It was reported that fusion of BM3 with formate dehydrogenase, which was used to increase NADPH supply, led to increased substrate conversion[211].

Glucose dehydrogenase (GDH) dehydrogenates glucose to D-glucono-1,5-lactone, reducing either NAD or NADP to NADH or NADPH. *P. megaterium* has four GDHs which accept both NAD and NADP as substrates, with GDH1 and GDH2 having the highest kcat/Km for NADP[212]. GDH from *P. megaterium* and *Bacillus subtilis* have been expressed in *E. coli* to regenerate NADPH during the bioconversion of ketoacids to hydroxyacids[213-217]. GDH from *P. megaterium* and *B. subtilis* have also been co-expressed or added *in vitro* when characterising

CYP450s[99,199,202,203,205,206], and immobilised BM3 and GDH have been used to hydroxylate dodecanoic acid[218]. Co-expression of GDH with BM3 or other CYP450s has also been reported to improve product conversion and CYP450 activity compared to expression of just the CYP450, and to increase intracellular NADPH concentration[219-223].

2.2.7 Conversion of hydroxyfatty acids to lactones

It has been suggested that 4- or 5-hydroxyfatty acids will spontaneously condense to form lactones[166], however acidic conditions have been used in the literature to catalyse the lactonization reaction[93,98,99,101,224]. Of particular note, it was found that at neutral pH in aqueous media, 4-hydroxypentanoate and γ-pentalactone were in equilibrium. Lowering the pH from 7 to 6 shifted the equilibrium position ~5 fold towards the lactone, and using a polymer to remove the lactone from the aqueous phase could further bias the equilibrium position towards the lactone[224]. This suggests that while some spontaneous condensation may occur, product extraction and purification procedures will require optimisation to favour the lactone over the hydroxyfatty acid.

Alternatively, enzymatic means could be used to synthesise the lactone from the hydroxyfatty acid. *Pseudomonas* are known to produce medium chain length polyhydroxyalkanoates, which are used for carbon and energy storage. One of the genes in the polyhydroxyalkanoate biosynthesis pathway, AlkK from *Pseudomonas oleovorans* (PoAlkK), was found to be a hydroxyfatty acid-CoA ligase. It was found to be able to ligate 3- and 4- hydroxybutanoic acid with CoA, albeit with an activity loss relative to butanoic acid, and be more active on medium chain length aliphatic substrates than short chain substrates. A homologue of PoAlkK, Pp0763 from *Pseudomonas putida* was identified and found to function as a 3-hydroxydecanoic acid-CoA ligase[225,226]. These genes could therefore be used to synthesise 4- and 5-hydroxydecanoyl-CoA from the respective fatty acids. PpAAT1 from *Prunus persica* is an alcohol-acyl transferase, and has been shown to catalyse the condensation of 4-hydroxydecanoyl-CoA to γ-decalactone when transgenically expressed in *E. coli*[227,228]. Together, AlkK and PpAAT1 could therefore serve as extension to proposed biosynthetic pathway should there be issues with the spontaneous or acid-catalysed condensation of 4- and 5-hydroxydecanoic acid to γ- and δ-decalactone.

2.3 Ester Biosynthesis

2.3.1 Esters

Esters are formed by the condensation of an alcohol and acyl-CoA, this reaction is typically catalysed by an alcohol-acyltransferase (AAT) for the formation of short and medium chain length esters, and by Wax Synthases (WS) for the formation of long chain length esters[151]. High molecular weight esters have uses as biofuels and lubricants[229], while lactate esters can be used as green solvents[230].

Low molecular weight esters (less than 10 carbons) are volatile, and are flavour and fragrance compounds that are important contributors to fruit and plant aromas[26,229,231,232]. There are two main classes of flavour and fragrance ester, acetate esters and ethyl esters (Figure 2.7) [233,234]. Esters are also key contributors to the aromas of fermented goods, such as dairy goods and alcoholic beverages[233,235-240]. During the fermentation process, by lactic acid bacteria or yeast, a range of volatile compounds are produced, of which esters are the most important[235,237]. Due to their effect on the beverage's aroma and flavour, the ester profile is key to the overall quality of the product[235,240,241]. Understanding and being able to modify the flavour profiles of fermented drinks is of industrial interest, as this will affect consumer reception of the product[233]. To this end there have been attempts to modify the ester production by yeast to create new flavour profiles, such as by co-culturing different yeast species during the fermentation process[242], changing the culture conditions[238], or using yeast that have been engineered to produce more esters[243].

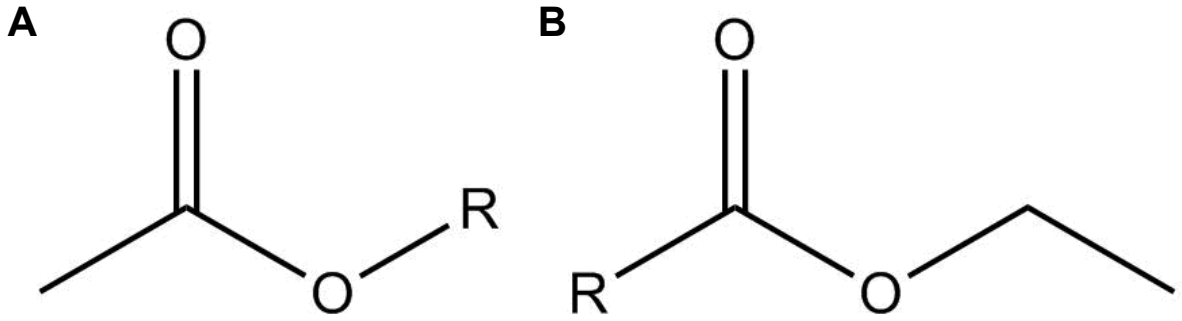


Figure 2.7 Acetate and ethyl ester structure *A) Acetate ester structure B) Ethyl ester structure.*

In this work, biosynthesis of butyl acetate and ethyl butyrate was targeted. Butyl acetate has a fruity, banana-like smell and has applications as a flavour compound and as an industrial solvent. There is a large global market, which from 1999-2009 ranged in volume from 89 million to 107 million kg at an average price of $1.06 /kg[241,244]. Ethyl butyrate has a fruity,

rum-like aroma, and is a key aroma compound in wines. It has a significant role in the flavour and fragrance industry, where it is used to add fruity tones to food and beverages[244,245].

2.3.2 Previously reported biosynthesis of esters

There has been great interest in the biosynthesis of esters, with many groups reporting the biosynthesis of a diverse array of esters using various host organisms and pathways[28,151-153,230,246-271]. There are commonalities however, such as production of an alcohol, production of acyl-CoA and synthesis of the final ester. For production of the alcohol, approaches including feeding the alcohol[230,250,259,261,269], using Pyruvate Decarboxylase (PDC) to produce ethanol[127,246,255,264], engineering 2-keto acid pathways to biosynthesise the alcohol[259,265,266], using the *Clostridium* butanol biosynthesis pathway[247,248,253,256,257,263,268,272] and expressing a thioesterase and carboxylic acid reductase[28] have all been used. For production of acyl-CoA, ligating a fatty acid to CoA using fadD[127,151,152,246,249,264], propionyl-CoA transferase (PCT)[230,255] and the *Clostridium* butanol biosynthesis pathway[127,248,256,257,263,268,272] have been used. Finally, to synthesise the ester, AAT[28,151,230,247,250,254,255,257,259,262,265,266,268,270-272], WS[152,246,249,264,267], Chloramphenicol Transferase (CAT)[260,261] and Eat1, a novel ethyl acetate synthesising enzyme[252], have all been used.

Prior to this work commencing, biosynthesis of both butyl acetate and ethyl butyrate had been reported in *E. coli* but not in cyanobacteria. This was mostly done by feeding various substrates[250,255,259], and the *de novo* biosynthesis of both butyl acetate and ethyl butyrate from glucose had only been reported once in *E. coli*, with titres obtained of 1.29 mg/L and 134 mg/L respectively[254]. There was therefore scope for improvement before attempting biosynthesis in cyanobacteria. In 2022, after work on biosynthesis of butyl acetate had been completed and while work on biosynthesis of ethyl butyrate was underway, two groups reported the *de novo* biosynthesis of butyl acetate and ethyl butyrate in *E. coli*, reaching titres of 22.8 g/L[253] and 450 mg/L[256], respectively.

2.3.3 Proposed pathway for butyl acetate biosynthesis

In this work a previously reported pathway for the biosynthesis of acetate esters from the Jones lab[28] will be adapted to biosynthesise butyl acetate. A C4 specific acyl-ACP thioesterase will be used to produce butyric acid from endogenous fatty acid biosynthesis. Sfp activates Carboxylic Acid Reductase (CAR) which reduces free fatty acids to aldehydes. Endogenous Alcohol Dehydrogenase (ADH) then reduces aldehydes to alcohols. AAT is then used to produce butyl acetate from butanol and acetyl-CoA. (Fig 2.8)

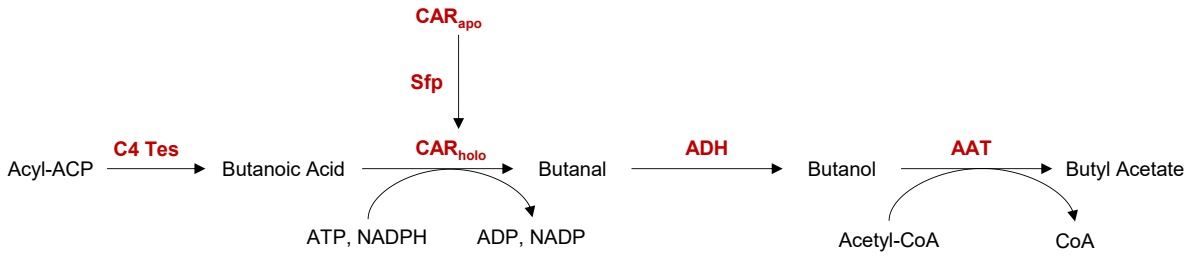


Figure 2.8 Proposed pathway for biosynthesis of butyl acetate *Pathway enzymes are highlighted in red. C4 Tes: C4 specific acyl-ACP thioesterase, Sfp: Phosphopantetheine transferase (Bacillus subtilis), CAR: Carboxylic acid reductase, ADH: Alcohol Dehydrogenase, AAT: Alcohol-acyltransferase.*

2.3.4 C4 specific acyl-ACP thioesterases

To produce butanoic acid a C4 specific acyl-ACP thioesterase will be used. In work characterising multiple thioesterases from plants and bacteria in *E. coli*, microbial thioesterases from *Bacteroides fragilis* (Tes4), *Bacteroides thetaiotamicron* (BtTes) and *Marvinbryantia formatexigens* (MfTes) led to the accumulation of butanoic acid within the cell (>10% of total fatty acid product)[104]. Tes4, BtTes and MfTes have since been used in work by other groups to biosynthesise butanoic acid, butanol, and propane[14,273-276]. These thioesterases can therefore be used to produce butanoic acid for the proposed butyl acetate biosynthesis pathway.

2.3.5 Carboxylic acid reductase

Carboxylic acid reductase (CAR) catalyses the second reaction in the butyl acetate biosynthesis pathway. CAR catalyses the reduction of carboxylic acids to aldehydes using NADPH and ATP co-factors. CARs have broad substrate specificity, and are able to reduce aliphatic, aromatic, heterocyclic, and di- and tricarboxylic acid substrates[277-282]. CAR requires phosphopantetheinylation to form the holoenzyme, which is catalysed by a phosphopantetheine transferase (PPTase)[283]. In the Jones lab, CAR from *Mycobacterium marinum* (MmCAR) and PPTase from *B. subtilis* (Sfp) have been used to reduce fatty acids to aldehydes as part of pathways for the biosynthesis of alkanes[13], alcohols[128,150], esters and glucosides[28]. MmCAR and Sfp have also been used by other groups to biosynthesise alcohols[1], and bioconvert fatty acids to aldehydes[127,284,285] in *E. coli*.

2.3.6 Alcohol dehydrogenase

Alcohol Dehydrogenase (ADH) catalyses the reversible reduction of aldehydes to alcohols. ADH overexpression has been used in the literature to catalyse this reaction[273], however *E. coli* has 13 known endogenous ADHs[286], and in previous work by the Jones the activity of

endogenous ADHs was sufficient in both *E. coli* and cyanobacteria to reduce fatty aldehydes to alcohols[28,128,150].

2.3.7 Alcohol-acyl transferase

The final step in the butyl acetate biosynthesis pathway is catalysed by an alcohol-acyl transferase (AAT). AATs catalyse the condensation of an alcohol and an acyl-CoA to form an ester and free CoA[287,288]. AATs can accept a wide range of substrates, including aliphatic, aromatic, cyclic, terpenoid and thioether alcohols[287,289,290], and many AATs have been characterised[127,227,288,289,291-314]. AATs are used in the production of esters in fruits and alcoholic beverages, determining the final aroma of the product[295,315-318]. AATs generally have wide substrate specificities but different preferred substrates, and ester production in fruit is determined by both the availability of precursor compounds and AAT specificity[289,305,309,319,320]. For example, VpAAT1 from papaya is capable of synthesising esters not found in papaya[289]. AAT overexpression has been used to alter the flavour profiles of dairy products[321] and alcoholic drinks[316,322], and to biosynthesise esters as previously discussed in Section 2.3.2

2.3.8 Proposed pathway for ethyl butyrate biosynthesis

In this work, a novel metabolic pathway for the biosynthesis of ethyl butyrate is proposed. A C4 specific acyl-ACP thioesterase will be used to produce butanoic acid from endogenous fatty acid biosynthesis. Acetyl-CoA Synthase (ACS) will be used to ligate butanoic acid with CoA. AAT will then catalyse the condensation of ethanol and butanoyl-CoA to form ethyl butyrate (Fig 2.9).

C4 thioesterases and AAT have been previously discussed in Sections 2.3.4 and 2.3.7.

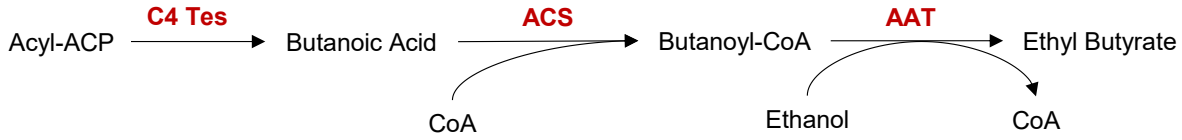


Figure 2.9 Proposed biosynthetic pathway for ethyl butyrate *Pathway enzymes are highlighted in red. C4 Tes: C4-specific acyl-ACP thioesterase, ACS: Acetyl-CoA synthase, AAT: alcohol-acyltransferase.*

2.3.9 Acetyl-CoA synthase

To synthesise ethyl butyrate, butanoyl-CoA needs to be synthesised first. To do so, butanoic acid will be ligated with free CoA. *E. coli* has two fatty acid-CoA ligases, FadD and FadK, of which FadD has been previously used to ligate fatty acids to CoA as part of pathways for the biosynthesis of esters, as discussed in Section 2.3.2. However, both FadD and FadK lack the appropriate substrate specificity for the biosynthesis of butanoyl-CoA, and are instead active on longer chain length substrates[75]. Instead, an Acetyl-CoA Synthase (ACS) can be used.

These typically ligate acetate with CoA, however a C4 specific ACS, AAE7 from *Arabidopsis thaliana* has been reported[323], as well as a C4 specific mutant of *A. thaliana* ACS[324], and ACSs from *Metalloshpaera sedula* that can accept butanoic acid as a substrate[325]. In unpublished work by the Jones group, homologues of *A. thaliana* AAE7 were identified and found to ligate butanoic acid and CoA together. These genes can therefore be used to synthesise butanoyl-CoA.

2.4 Methyl Ketone Biosynthesis

2.4.1 Methyl ketones

Medium chain length methyl ketones are potent flavour and fragrance compounds, that are important constituents of dairy products and plant aromas[16,22,26,326-328]. Long chain length methyl ketones help protect plants from pests and can be used as biofuels[22].

2.4.2 Native biosynthesis in *Penicillium*

During fermentation of blue cheese by *Penicillium* methyl ketones are produced, however the exact pathway is still unknown[327]. It is known that during the ripening of blue cheeses there is an accumulation of free fatty acids due to the activity of lipase, and that the addition of lipase during the fermentation process results in increased production of methyl ketones[329,330]. When *Penicillium roqueforti* was incubated with 2-^{14}C labelled octanoic acid, 1-^{14}C labelled 2-heptanone was produced, demonstrating that fatty acids are precursors in the biosynthesis of methyl ketones. Production of ^{14}C labelled CO_2 was also observed, and from this it was deduced that the octanoic acid entered the β-oxidation pathway, and that some of the octanoic acid was not converted into 2-heptanone, but was instead degraded to CO_2. One of the intermediaries of the β-oxidation cycle is β-ketoacyl-CoA, and it was suspected that a thiolase acted upon the β-ketoacyl-CoA to produce free β-ketofatty acids, which were then decarboxylated to methyl ketones by a decarboxylase[331]. When *Penicillium camemberti* was incubated with lauric and oleic acid, the major product was 2-nonanone, with minor amounts of 2-heptanone and 2-undecanone, and trace amounts of 2-pentanone and 2-tridecanone produced[330]. This result suggested that the fed acids underwent multiple rounds of the β-oxidation cycle, and that the enzymes responsible for synthesising methyl ketones from intermediates of the β-Oxidation cycle had medium chain length specificity. This pathway is summarised in Figure 2.10.

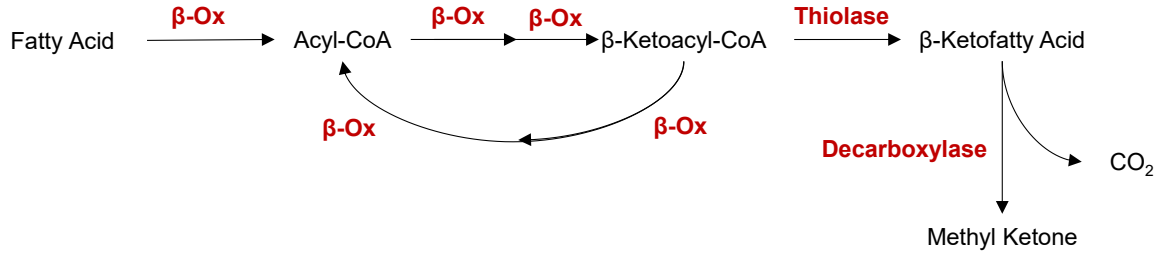


Figure 2.10 Methyl ketone biosynthesis pathway in Penicillium *Pathway enzymes are highlighted in red. B-Ox: β-oxidation, Thiolase: Proposed β-Ketoacyl-CoA thioesterase, Decarboxylase: Proposed β-Ketofatty Acid Decarboxylase.*

2.4.3 Native biosynthesis in plants

Tomato (*Solanum haborchaites f glabratum*) trichomes were found to have high levels of methyl ketones. Fatty acid biosynthetic enzymes were found to be more highly expressed than fatty acid degradation enzymes, suggesting a pathway based on fatty acid biosynthesis, unlike in *Penicillium*. The most expressed unknown sequence was investigated and found to catalyse *in vitro* the production of methyl ketones from β-ketoacyl ACP, and was named *S. habrochaites f glabratum* methyl ketone synthase 1 (ShMKS1)[332]. Follow up work identified a second unknown gene, named ShMKS2, and determined that ShMKS2 was a C14 specific β-ketoacyl ACP thioesterase and ShMKS1 was a β-ketofatty acid decarboxylase[333]. The pathway is summarised in Figure 2.11. Homologues of ShMKS2 were subsequently identified in *A. thaliana*[333], and identified in and characterised from soybean (*Glycine max*)[334] and aubergine (*Solanum melongena*)[335]. The characterised homologues were also C14 specific β-ketoacyl ACP thioesterases.

A family of four thioesterases that were related to ShMKS2 was Identified in *A. thaliana* and named Acyl Lipid Thioesterases (ALT)[336]. Further work found that ALTs were widely distributed in plants, and a small library of ALTs were characterised in *E. coli*. This work found that ALTs could be classified into four distinct groups based on activity. Group 1 ALTs were C14 specific β-ketoacyl ACP thioesterases, like ShMKS2. Group 2 ALTs had mixed activities, and were C6-C10 specific acyl-ACP and β-ketoacyl ACP thioesterases. Group 3 ALTs were C12-C14 specific acyl-ACP thioesterases, and Group 4 ALTs were C16 specific acyl-ACP thioesterases. Which group an ALT belonged to could not be predicted using the protein sequence, however alignment of predicted protein structures found that Group 2 ALTs occupied a distinct branch within the phylogeny[337].

Homologues of ShMKS1 and ShMKS2, YtpA and YneP, were identified in *Bacillus nematocida*, a pathogen of nematodes. Overexpression of YneP in *E. coli* led to the biosynthesis of C5-C11 methyl ketones, while overexpression of YtpA in *E. coli* did not lead to biosynthesis of methyl ketones. A homologue of YneP, YneP' was identified in *Bacillus subtilis* and overexpression of YneP' in *E. coli* also led to the biosynthesis of C5-C11 methyl ketones[338]. From these results the authors concluded that YneP and YneP' functioned as methyl ketone synthases, while YtpA did not contribute to methyl ketone production. However, YtpA is homologous to ShMKS1, a β-ketofatty acid decarboxylase, and so it is unlikely that expression of only YtpA, without a β-ketofatty acid substrate, whether fed or produced by enzymatically, would lead to biosynthesis of methyl ketones, and so these results should not be used to conclude that YtpA does not contribute to methyl ketone production.

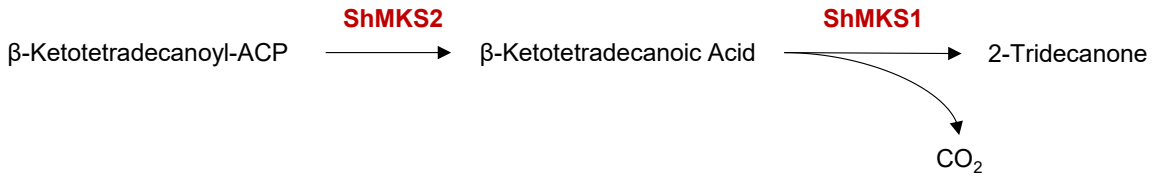


Figure 2.11 Methyl ketone biosynthesis pathway in plants *Pathway enzymes are highlighted in red. ShMKS2: Solanum habrochaites f glabratum Methyl Ketone Synthase 1, a C14 specific β-ketoacyl ACP thioesterase, ShMKS1: Solanum habrochiates f glabratum Methyl Ketone Synthase 2, a β-ketoacid decarboxylase.*

2.4.4 Previously reported biosynthesis of methyl ketones

Biosynthesis of methyl ketones using both the *Penicillium* β-oxidation and plant methyl ketone synthase pathways has been reported.

Using the β-oxidation pathway, 2-undecanone, 2-tridecanone and 2-pentanone were biosynthesised in *E. coli*. To do so a thioesterase was used to biosynthesise free fatty acids. These free fatty acids were then fed into a modified β-oxidation pathway. Native FadE was replaced with an acyl-CoA oxidase from *Micrococcus luteus* and FadA was knocked out to prevent degradation to shorter chain length fatty acids. Finally, FadM, a promiscuous acyl-CoA thioesterase with a wide substrate specificity[339], was expressed to produce β-ketofatty acids from β-ketoacyl-CoA. Spontaneous decarboxylation of the β-ketofatty acids to methyl ketones was then observed[155]. In follow up work methyl ketone titres were improved by optimising flux through the pathway[340] and co-factor balance[341]. Substituting the thioesterase for a C8 specific thioesterase and screening for FadM homologues with higher activity was then used to synthesise medium chain length methyl ketones[156]. This approach was then used to engineer a range of host organisms to biosynthesise methyl ketones[342-346]. A modification of the pathway was also reported, where PhaG from *Pseudomonas koreensis* was used as a transacylase, converting β-hydroxyacyl-ACP to β-hydroxydacyl-CoA, an intermediate in β-oxidation, removing the need for a thioesterase[347].

The plant pathway has also been used to biosynthesise methyl ketones in *E. coli*. Expression of only ShMKS2 in *E. coli* led to biosynthesis of 2-tridecanone, suggesting some spontaneous decarboxylation of β-ketotetradecanoic acid to 2-tridecanone occurred. When ShMKS2 was co-expressed with ShMKS1 there was then a 4-fold increase in 2-tridecanone titre[333]. However, another group reported that *E. coli* expressing ShMKS1 and ShMKS2 actually biosynthesised less 2-tridecanone than *E. coli* expressing only ShMKS2[155]. A third group also reported the biosynthesis of 2-tridecanone in *E. coli* after co-expression of ShMKS1 and ShMKS2, but did not investigate whether co-expression improved 2-tridecanone titres relative to expression of only ShMKS2[348]. C11-C15 methyl ketones were also biosynthesised in *Saccharomyces*

cerevisiae by integrating ShMKS2 into an engineered fungal fatty acid synthase, and integration of ShMKS1 and ShMKS2 led to an increase in titres[349].

Alternative biosynthetic pathways beyond the native pathways have also been reported. CoA dependent pathways were used to biosynthesise 2-butanone and 2-pentanone. Free β-ketofatty acids were synthesised, and then *Clostridium acetobutylicum* acetoacetate decarboxylase was used to catalyse the decarboxylation of the β-ketofatty acids to methyl ketones[59,350]. An engineered Polyketide Synthase was also used to biosynthesise 2-butanone and 2-pentanone in *E. coli*[351] and *Streptomyces albus*[352]. 2-butanone was also produced in *E. coli* from butan-2,3-diol using a glycerol dehydratase[353].

2.4.5 The spontaneity of β-ketofatty acid decarboxylation

While it has been reported that decarboxylation is spontaneous, and biosynthesis of methyl ketones without a decarboxylase has been observed[155,156,338,340-345,347,351,352], there have also been reports showing that co-expression of a decarboxylase[59,348,350], or that heat and acidic conditions can be used to improve methyl ketone production[326,333-337]. Due to these contradictory reports, investigation will be needed to determine how spontaneous the decarboxylation reaction is, and whether co-expression of a decarboxylase improves methyl ketone titres.

2.4.6 Proposed pathway for methyl ketone biosynthesis

In this work, the plant methyl ketone synthase pathway will be used to biosynthesise methyl ketones. If decarboxylation is spontaneous, then only one heterologous gene will be needed for methyl ketone biosynthesis, reducing burden upon the cell. Group 2 ALTs and YneP show promise for the biosynthesis of medium chain length methyl ketones, however due to the somewhat bizarre units used when characterising ALTs (nmol /unit OD , ODs not provided)[337] and the fact that when characterising YneP methyl ketone production was not quantified, only a GC-MS chromatogram was shown[338], it is not clear how active these genes are.

2.5 Green Leaf Volatile Biosynthesis

2.5.1 Green Leaf Volatiles

Green leaf volatiles (GLVs) are a set of C6 saturated and unsaturated alcohols, aldehydes and esters, and are called so because they produce the “green odour” of crushed leaves[354-356]. GLVs are used as flavours and fragrances in the food industry for their green odour, which is used to reconstitute the fresh aroma of fruits and vegetables following processing[22,357]. In plants GLVs are produced after wounding, herbivory, or pathogen attack[355,358,359]. GLVs have antibacterial[360-362], antifungal[363-366] and nematocidal[367,368] activity. Unsaturated GLVs are more potent inhibitors of bacterial growth than their saturated counterparts[360]. They can also serve to attract predators, for example after herbivory by cabbage white butterfly larvae, GLV production increased attracting pathogenic wasps[366]. Due to these activities GLVs are also used in the food industry as food preservatives[369].

2.5.2 Native GLV biosynthesis pathway

It was initially observed that linolenic acid could be oxidised by Ginko and Tea leaf extracts to *cis*-3-hexenal, which then isomerised to *trans*-2-hexenal[370,371]. After further research the whole pathway was elucidated. Linolenic is oxidised by a lipoxygenase (LOX). LOXs can have C9 or C13 specificity, in the biosynthesis of GLVs a C13 LOX is used to oxidise linolenic (9,12,15-octadecatrienoic) acid to 13-hydroperoxy-9,11,15-octadecatrienoic acid[355,359,372]. Hydroperoxide lyase (HPL) then cleaves the 13-hydroperoxy-9,11,15-octadecatrienoic acid to produce *cis*-3-hexenal and 12-oxo-9-dodecenoic acid[355,359,373]. The *cis*-3-hexenal can then be isomerised to *trans*-2-hexenal by a 3Z:2E-enal isomerase[374-376]. *Trans*-2-hexenal can be further converted by alkenal hydrogenase (ALH) to hexanal[377]. The aldehydes can then be reduced to the corresponding alcohols by ADH[378]. *Cis*-3-hexenal can also be converted to *cis*-3-hexenyl acetate by an AAT[379]. The pathway is summarised in Figure 2.12.

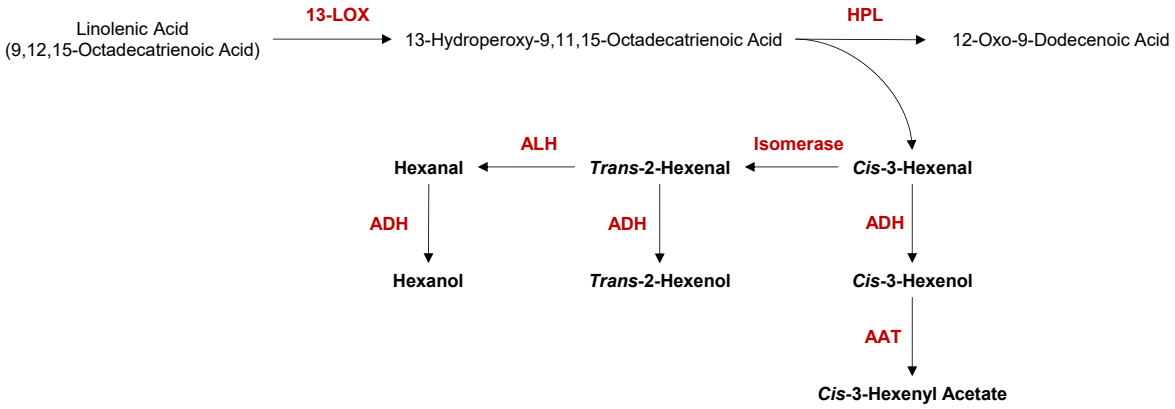


Figure 2.12 Green leaf volatile biosynthesis pathway *GLVs are highlighted by bold text. Pathway enzymes are shown in red. 13-LOX: C13 specific lipoxygenase, HPL: Hydroperoxide lyase, Isomerase: 3Z:2E-enal isomerase, ALH: Alkenal hydrogenase, ADH: Alcohol dehydrogenase, AAT: Alcohol-acyltransferase.*

2.5.3 Previously reported biosynthesis of GLVs

GLVs have been produced in *E. coli* by overexpressing LOX and feeding linolenic acid to make *cis*-3-hexenal or feeding arachidonic acid to make hexanal[380]. Alternatively, GLVs have been produced in *E. coli* by expressing HPL and feeding 13-hydroperoxides of linoleic and linolenic acid to make hexanal, *cis*-3-hexenal and *cis*-3-hexenol[381,382]. LOX and HPL have also been expressed together in *S. cerevisiae* and *Yarrowia lipolytica* to bioconvert linoleic and linolenic acid to hexanal, and *cis*-3-hexenol and *trans*-2-hexenal[383,384].

2.5.4 Proposed pathway for biosynthesis of GLVs

In this work, the following pathway for the biosynthesis of *cis*-2-hexenal and *cis*-2-hexenol is proposed (Fig 2.13). A C6 specific mutant of DfsA will be used to produce *cis*-2-hexenoic acid from endogenous fatty acid biosynthesis. CAR will then reduce the *cis*-2-hexenoic to *cis*-2-hexenal, and endogenous ADH will reduce *cis*-2-hexenal to *cis*-2-hexenol. While the *trans* isomers are produced in native GLV biosynthesis, a study found that both *cis* and *trans* isomers had similar odour profiles, with leafy-grassy green, fresh, sweet and fruity notes. There was also no difference in threshold values in parts per million for odour perception, and only slight differences in threshold values for taste perception[385]. Therefore, *cis*-2-hexenal and *cis*-2-hexenol could be used as GLVs in flavours and fragrances.

CAR and ADH have been previously discussed in Sections 2.3.5 and 2.3.6.

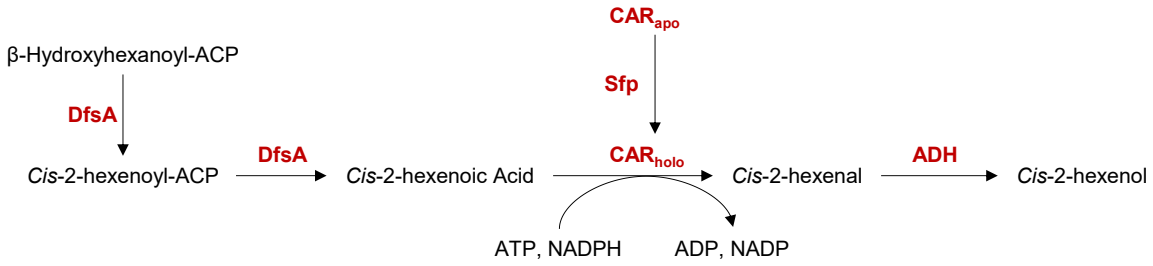


Figure 2.13 Proposed cis-2-hexenal and cis-2-hexenol biosynthetic pathway *Pathway enzymes are highlighted in red. DfsA: Bifunctional β-Hydroxyacyl-ACP dehydratase and cis-2-enoyl-ACP thioesterase (Burkholderia cenocepacia), Sfp: Phosphopantetheine transferase (Bacillus subtilis), CAR: Carboxylic acid reductase.*

2.5.5 DfsA

Diffusible signalling factors (DSFs) are a family of quorum signalling molecules, comprised of straight chain fatty acids that are either methylated, unsaturated or both. The first identified DSF was 11-methyl-2-dodecenoic acid, named DSF, in *Xanthomonas campestris pv campestris*, and the gene *rpfF* was found to be associated with DSF biosynthesis[386,387]. Another DSF, *cis*-2-dodecenoic acid, named BDSF, was identified in *Burkholderia cenocepacia*, and the ORF *Bcam0581*, a homology of *rpfF*, was found to be essential for production of BDSF[388]. BDSF was then identified in *Xanthomonas oryzae pv oryzae*[389], and in other strains within the *B. cenocepacia* complex which contained homologues of *Bcam0581*[390]. *Cis*-2-decenoic acid was also identified in *Pseudomonas aeruginosa*[391] and *Streptococcus mutans*[392].

In *P. aeruginosa* a homologue of *rpfF* was identified, *PA0745/dspI*, and reported to be responsible for the biosynthesis of *cis*-2-decenoic acid[393]. However, another group reported that a different gene, *PA4980*, produced *cis*-2-decenoic acid, based on protein-protein interactions[394], leaving it unclear which gene is responsible for *cis*-2-decenoic acid biosynthesis. On the other hand, *Bcam0581* was characterised, and found to be a bifunctional enzyme, with combined β-hydroxydodecanoyl-ACP dehydratase and *cis*-2-dodecenoyl-ACP thioesterase activity. *Bcam0581* was named DfsA, and it was found that heterologous expression of DfsA in *E. coli* led to the biosynthesis of *cis*-2-dodecenoic acid[395]. In later work, the crystal structure for DfsA was resolved, with a dodecanoic acid ligand within the substrate binding pocket[396].

Due to its better characterisation and available crystal structure, DfsA represents a better target for engineering, and if its chain length specificity could be shifted to being C6 specific, it could be used to biosynthesise *cis*-2-hexenal and *cis*-2-hexenol.

Chapter 3

Materials and Methods

3.1 *Escherichia coli*

3.1.1 *E. coli* strains

E. coli DH5α (ThermoFisher Scientific) were used for molecular cloning of all plasmids produced during this work. *E. coli* BW25113 Wild-Type, BW25113 derivative strains JW1794-1 (ΔfadD) and JW5020-1 (ΔfadE) from the Keio collection[38] (with integrated kanamycin resistance cassettes removed during prior work from the Jones lab), BW25113 ΔiscR (this study) and BW25113 ΔyciA (made during prior work from the Jones lab) were used for bioproduction and feeding experiments. *E. coli* BL21 Wild-Type and BL21 ΔiscR[397] were used for some feeding experiments.

3.1.2 LB media and LB-Agar

LB Media was prepared by dissolving 2 g of LB powder (Sigma-Aldrich) per 100 mL of MilliQ ultrapure water (Merck). To prepare LB-Agar, 2 g of LB powder and 1.5 g Bacto Agar (Appleton Woods) were dissolved per 100 mL MilliQ. LB and LB-Agar were sterilised by autoclaving at 121°C and 15 psi for 20 mins. After autoclaving, LB and LB-Agar were stored at room temperature. To make petri dishes, the LB-Agar was microwaved until fully melted, then allowed to cool until the bottle could be comfortably held without a glove. If required, antibiotics were added at this point, and 10-15 mL LB-Agar was poured into a 90 mM petri dish (VWR). The LB-Agar was then allowed to set at room temperature under sterile conditions.

3.1.3 M9 minimal media

M9 minimal media with 2% glucose was prepared under sterile conditions from stock solutions as set out in Table 3.1. Antibiotics were added to the M9 as necessary.

Component	Stock Conc.	Final Conc.	Volume	Source
M9 Salts	5X	1X	200 mL	Sigma-Aldrich
Glucose	20%	2%	100 mL	Sigma-Aldrich
$MgSO_4$	1 M	2 mM	2 mL	Sigma-Aldrich
$CaCl_2$	1 M	0.1 mM	100 µL	Sigma-Aldrich
MilliQ	-	-	Up to 1L	Merck

Table 3.1 Composition of M9 minimal media with 2% glucose

The stock solutions were individually prepared by dissolving the components in MilliQ and then autoclaved at 121°C and 15 psi for 20 mins. Autoclaved MilliQ, 5X M9 and 20% Glucose stock solutions were stored at 4°C, while 1 M $MgSO_4$ and 1 M $CaCl_2$ were stored at room temperature.

To make M9-Fe, the M9 was prepared as described above. A stock solution of 6 g/L Ferric Ammonium Citrate (Sigma-Aldrich) was prepared by dissolving the Ferric Ammonium Citrate in MilliQ and filter-sterilised using a sterile 0.2 µm syringe filter. The stock solution was then stored in the dark at 4°C. To make M9-Fe, 1 mL of stock solution was added, giving a final concentration of 6 mg/L Ferric Ammonium Citrate.

To supplement M9 with 5-aminoleuvinic acid a 1 M stock solution was prepared by dissolving 5-aminoleuvinic acid (Sigma-Aldrich) in MilliQ, and filter-sterilised using a sterile 0.2 µm syringe filter. The stock solution was then stored in the dark at 4°C. 5-aminoleuvinic acid stock was then added to M9 to give a final concentration of 0.2 mM.

To supplement M9 with caffeine, a 50 mM stock solution was prepared by dissolving caffeine in MilliQ and sterilised by autoclaving at 121°C and 15 psi for 20 mins. Caffeine stock was then added to M9 to give different caffeine concentrations.

To supplement M9 with cas amino acids, a 2% stock solution was prepared by dissolving casein hydrolysate (Millipore) in MilliQ. The stock solution was then sterilised by autoclaving at 121°C and 15 psi for 20 mins at stored at room temperature. Cas amino acid stock solution was then added to M9 to give a final concentration of 0.05%.

3.1.4 Antibiotics used for cultivation of *E. coli*

Stock solutions of antibiotics were prepared by dissolving antibiotic powder in either MilliQ or ethanol, according to Table 3.2. Antibiotics dissolved in MilliQ were then filter-sterilised using a sterile 0.2 µm syringe filter. Antibiotics were stored at -20°C and thawed at room temperature prior to use.

Antibiotic	Stock Conc.	Working Conc.	Solvent	Source
Carbenicillin	100 mg/mL	100 mg/L	MilliQ	Sigma-Aldrich
Kanamycin	50 mg/mL	50 mg/L	MilliQ	Sigma-Aldrich
Spectinomycin	50 mg/mL	50 mg/L	MilliQ	Sigma-Aldrich
Gentamycin	50 mg/mL	10 mg/L	Ethanol	Sigma-Aldrich
Erythromycin	20 mg/mL	200 mg/L	Ethanol	Sigma-Aldrich
Chloramphenicol	37 mg/mL	37 mg/L	Ethanol	Sigma-Aldrich

Table 3.2 Preparation of antibiotics

3.1.5 Cryopreservation and revival of *E. coli* glycerol stocks

E. coli strains were cryopreserved in a 16% (v/v) glycerol solution. An 80% glycerol solution was prepared by mixing 20 mL glycerol (Sigma-Aldrich) with 5 mL MilliQ and autoclaved at 121°C and 15 psi for 20 mins. To prepare glycerol stocks, 600 µL of overnight culture in LB media and 150 µL 80% (v/v) sterile glycerol were added to a 2 mL CryoTube (VWR). The CryoTube was then stored in a -80°C freezer.

To revive a strain, the glycerol stock was removed from the freezer and kept on ice. An inoculating loop was used to re-streak the stock onto an LB-Agar plate with appropriate antibiotics. The plate was then incubated overnight at 37°C.

Alternatively, an inoculation loop was used to inoculate the stock into 5 mL LB with appropriate antibiotics in a 50 mL falcon tube. The LB culture was then incubated overnight at 37°C, 180 rpm.

3.1.6 Preparation of calcium competent *E. coli*

Calcium competent cells were prepared using the following protocol, adapted from a previously reported protocol[398].

5 mL LB with appropriate antibiotics in a 50 mL falcon tube was inoculated with the desired *E. coli* strain and incubated overnight at 37°C, 180 rpm. 1 mL of overnight cell culture was used to inoculate 100 mL LB with appropriate antibiotics in a 250 mL Erlenmeyer flask. The 100 mL cell culture was incubated at 37°C, 180 rpm for 2 hrs, then removed and placed on ice for 30 mins. While the cell culture was on ice, a centrifuge was cooled to 4°C. The 100 mL cell culture was then divided into two 50 mL falcon tubes, and centrifuged at 4000, 4°C rpm for 5 mins. The supernatants were discarded, and the pellets resuspended in 5 mL ice-cold sterile 100 mM $MgCl_2$. The cell cultures were then centrifuged at 4000, 4°C rpm for 5 mins. The supernatants were discarded, and the pellets resuspended in 10 mL ice-cold sterile

100 mM $CaCl_2$. The cell cultures were then centrifuged at 4000, 4°C rpm for 5 mins. The supernatants were discarded, and the pellets resuspended in 5 mL ice-cold sterile 85 mM $CaCl_2$, 15% glycerol solution. The cell cultures were then centrifuged at 4000 rpm, 4°C for 5 mins. The supernatants were discarded, and the pellets resuspended in 2.5 mL ice-cold sterile 85 mM $CaCl_2$, 15% glycerol solution. The cell cultures were then aliquoted into sterile Eppendorf tubes and immediately stored at -80°C.

3.1.7 Heat-shock transformation of *E. coli*

Prior to transformation, a 50 µL aliquot of calcium competent cells of the appropriate *E. coli* strain were thawed on ice for 20-45 mins. If the *E. coli* was being transformed with a single plasmid, 1.5 µL of ~150 ng/µL plasmid was added to the competent cells. If the *E. coli* was being co-transformed with two plasmids, 2 µL of ~150 ng/µL of each plasmid was added to the competent cells. The cells were incubated on ice for 20 mins. During this time, a Thermoshaker (Grant Bio) was pre-heated to 42°C. After incubation on ice, the cells were incubated at 42°C for 45 s, then immediately placed back on ice for 2 mins. 1 mL of LB without antibiotics was then added, and the cells incubated at 37°C, 180 rpm for 1 hr. The cells were pelleted by centrifugation at 8000 rpm for 1 min. 950 µL supernatant was removed and discarded, and the cells resuspended in the remaining 100 µL. The cells were then transferred to an LB-Agar plate supplemented with appropriate antibiotics and spread using an L-shaped spreader (VWR).

3.1.8 Preparation of electrocompetent *E. coli*

Electrocompetent *E. coli* cells were prepared using a protocol provided by a previous member of the Jones lab.

5 mL LB with appropriate antibiotics in a 50 mL falcon tube was inoculated with the desired *E. coli* strain and incubated overnight at 37°C, 180 rpm. 50 µL of overnight culture was used to inoculate 50 mL LB with antibiotics in a 250 mL Erlenmeyer flask. The 50 mL culture was grown to an OD_{600} of ~0.6, then transferred to a 50 mL falcon and cooled on ice for 10 mins. The culture was then centrifuged at 4000rpm, 4°C for 10 mins. The supernatant was discarded, and the pellet resuspended in 50 mL ice-cold sterile MilliQ. The culture was then centrifuged at 4000rpm, 4°C for 10 mins. The supernatant was discarded, and the pellet resuspended in 25 mL ice-cold sterile MilliQ. The culture was then centrifuged at 4000rpm, 4°C for 10 mins. The supernatant was discarded, and the pellet resuspended in 25 mL ice-cold sterile 10% glycerol. The culture was then centrifuged at 4000rpm, 4°C for 10 mins. The supernatant was then discarded, and the pellet resuspended in 500 µL ice-cold sterile 10% glycerol. The culture was then aliquoted into Eppendorf tubes and immediately stored at -80°C.

3.1.9 Electroporation of *E. coli*

Electroporation of electrocompetent *E. coli* was performed according to a protocol provided by a previous member of the Jones lab. In this work, this protocol was only used to transform *E. coli* with PCR products for making gene knockouts.

25 µL of electrocompetent cells of the appropriate *E. coli* strain were thawed on ice for 20-45 mins. 1-2 µL of PCR product (100 ng total DNA) was added to the competent cells, and the cells kept on ice for 5 mins. The cells were carefully transferred to the gap within an electroporation cuvette (SLS), making sure no air bubbles were introduced. The cuvette was gently tapped to settle the cells within the bottom of the gap and remove any bubbles. The metal sides of the cuvette were then wiped with Kimwipe (Sigma-Aldrich) to remove any moisture. The cuvette was then placed within an electroporator and pulsed with the following settings: 1.8kV, 200Ω and 25µF. 500 µL SOC media without antibiotics was added to the cuvette to resuspend the electroporated cells, and then transferred to an Eppendorf. The cells were then incubated at 30°C, 150 rpm for 2 hrs. 50 µL and 450 µL of the SOC were then plated on LB-Agar plates supplemented with 30% of the normal working concentration.

3.1.10 *E. coli* bioproduction and feeding experiments

If a new *E. coli* strain was being assayed, the appropriate host strain was transformed with the necessary plasmids and plated onto LB-Agar supplemented with the appropriate antibiotics. Otherwise, the strain was revived from the glycerol stock by re-streaking onto an LB-Agar plate supplemented with the appropriate antibiotics. The next day, three colonies were picked per plate and inoculated into 5 mL LB preculture with appropriate antibiotics in a 50 mL falcon tube. The precultures were incubated overnight at 37°C, 180 rpm. The next day, the precultures were pelleted by centrifugation at 4000 rpm for 5 mins. The supernatant was discarded, and the cells washed by resuspension in 2.5 mL M9. The precultures were centrifuged again 4000 rpm for 5 mins. The supernatant was discarded, and the pellets resuspended in 0.5 mL M9. The precultures were then used to inoculate 25 mL M9 with appropriate antibiotics in a 100 mL Erlenmeyer flask at an OD_{600} of 0.1. If an overlay was used, at this stage a 10% v/v overlay was added to M9. The cultures were then incubated at 37°C, 150 rpm for 4 hrs, and then induced. If a feeding experiment was to be performed, the fed substrate was added at this time. The cultures were then incubated at 30°C, 150 rpm for 48 hrs, at which point samples were taken.

3.2 *Synechocystis* Sp. PCC 6803

3.2.1 *Synechocystis* strains

Two modified strains of *Synechocystis* were used in this study, which had been produced by a previous member of the Jones lab[13,28]. For biosynthesis of fatty alcohols, *Synechocystis* Sp. PCC 6803 Δaas with Sfp-CAR integrated into the genome was used. For biosynthesis of decalactones, *Synechocystis* Sp. PCC 6803 Δaas was used.

3.2.2 BG11-Co media and BG11-Co-Agar

In this work, a cobalt inducible promoter was used to express heterologous genes within *Synechocystis*, so BG11-Co (modified BG11 media[399] with cobalt removed) was used for the cultivation of *Synechocystis*. BG11-Co was prepared as set out in Table 3.3 and sterilised by autoclaving at 121°C, 15 psi for 20 mins. After autoclaving, BG11-Co was stored at room temperature.

Component	Stock Conc.	Final Conc.	Volume	Source
BG11	100X	1X	10 mL	-
Ferric Ammonium Citrate	1000X	1X	1 mL	-
Na_2CO_3	1000X	1X	1 mL	-
K_2HPO_4	1000X	1X	1 mL	-
Pentahydrate Na-Thiosulfate	-	-	4.7 G	Sigma-Aldrich
MilliQ	-	-	Up to 1L	Merck

Table 3.3 Composition of BG11-Co media

1000X Ferric Ammonium Citrate was prepared by dissolving 0.6 g Ferric Ammonium Citrate (Sigma-Aldrich) in 100 mL MilliQ. 1000X Na_2CO_3 was prepared by dissolving 2 g Na_2CO_3 (Sigma-Aldrich) in 100 mL MilliQ. 1000X K_2HPO_4 was prepared by dissolving 3.05 g K_2HPO_4 (Sigma-Aldrich) in 100 mL MilliQ. 100X BG11 was prepared according to Table 3.4.

Trace minerals were prepared according to Table 3.5. The individual components were dissolved in MilliQ.

Component	Amount	Source
$NaNO_3$	149.6 g	Sigma-Aldrich
$MgSO_4(H_2O)_7$	7.49 g	Sigma-Aldrich
$CaCl_2(H_2O)_2$	3.6 g	Sigma-Aldrich
Citric Acid	0.6 g	Sigma-Aldrich
0.25 M NaEDTA pH 8.0	1.12 mL	Sigma-Aldrich
Trace Minerals	100 mL	-
MilliQ	Up to 1L	-

Table 3.4 Composition of BG11

Component	Amount	Source
H_3BO_3	2.86 g	Sigma-Aldrich
$MnCl_2(H_2O)_4$	1.816 g	Sigma-Aldrich
$ZnSO_4(H_2O)_7$	0.22 g	Sigma-Aldrich
$Na_2MoO_4(H_2O)_2$	0.39 g	Sigma-Aldrich
$CuSO_4(H_2O)_5$	0.079 g	Sigma-Aldrich
MilliQ	1 L	-

Table 3.5 Composition of Trace Minerals solution

To prepare BG11-Co-Agar, 7.5g of Bacto Agar (Appleton Woods) was dissolved in 500 mL BG11-Co. BG11-Co-Agar was sterilised by autoclaving at 121°C, 15 psi for 20 mins, and then stored at room temperature.

To make petri dishes, the BG11-Co-Agar was microwaved until fully melted, then allowed to cool until the bottle could be comfortably held without a glove. If required, antibiotics were added at this point, and 5-10 mL BG11-Agar was poured into a 55 mM petri dish (Merck). The BG11-Co-Agar was then allowed to set at room temperature under sterile conditions.

3.2.2 Antibiotics used for cultivation of *Synechocystis*

A stock solution of erythromycin was prepared and used as described in Section 3.1.4, with the exception that the final working concentration was 20 µg/mL.

3.2.3 Cryopreservation and revival of *Synechocystis* strains

Synechocystis strains were cryopreserved in a 16% (v/v) glycerol solution. An 80% glycerol solution was prepared by mixing 20 mL glycerol (Sigma-Aldrich) with 5 mL MilliQ, and autoclaved at 121°C and 15 psi for 20 mins. To prepare glycerol stocks, 1.2 mL of cell culture in BG11-Co media and 300 µL 80% (v/v) sterile glycerol were added to a 2 mL CryoTube (VWR). The CryoTube was then stored in a -80°C freezer.

To revive a strain, the glycerol stock removed from the -80°C freezer and kept on ice. An inoculating loop was used to transfer the strain onto a BG11-Co-Agar plate. The plate was incubated for 2 days at 30°C in Sanyo MIR-253 light incubator (Panasonic Biomedical). The plate was then transferred to an Algaetron AG230 (Photon Systems Instruments) and incubated at 30°C with continuous warm-white LED illumination at 50 µmol photons/m^2 and a 1% CO_2 atmosphere (v/v). After approximately a week, growth would be sufficient to transfer the strains to liquid media.

5 mL BG11-Co with appropriate antibiotics was added to a 6-well plate well (VWR). An inoculation needle (VWR) was used to gently scrape the surface of the BG11-Co-Agar, removing the cyanobacteria. The needle was then gently rubbed against the edge of the well, transferring the cyanobacteria to the liquid media.

3.2.4 Maintenance of revived *Synechocystis* strains

Revived *Synechocystis* strains were maintained in liquid culture in 5 mL BG11-Co with appropriate antibiotics within a 6-well plate. Cultures were grown in an Algaetron AG230 and incubated at 30°C with continuous warm-white LED illumination at 50 µmol photons/m^2 and a 1% CO_2 atmosphere (v/v). Once a week, 1 mL cell culture was used to inoculate 4 mL of fresh BG11-Co with appropriate antibiotics.

3.2.5 Triparental conjugation of *Synechocystis*

In this work triparental conjugation was used to transform *Synechocystis* with plasmids[400]. Two *E. coli* strains were used in this process; HB101, carrying helper plasmid pRL623[401] with a chloramphenicol resistance marker, and DM82, carrying conjugal plasmid pRL443 with a carbenicillin resistance marker.

Chemically competent HB101 were transformed by heat-shock with the plasmid to be transformed into *Synechocystis*. After transformation, a colony was picked and inoculated into 5 mL LB with appropriate antibiotics in a 50 mL falcon tube and incubated overnight at 37°C, 180 rpm. Meanwhile, DM82 was revived into liquid media as described in Section 3.1.5 and incubated overnight at 37°C, 180 rpm.

The next day, 2 mL of overnight culture for each *E. coli* strain was centrifuged at 8000 rpm for 1 min. The supernatant was discarded, and the pellet resuspended in 1 mL fresh LB without antibiotics. This process was repeated three times to completely remove the antibiotics. Meanwhile, 1 mL of *Synechocystis* Sp. PCC 6803 from a 6-well plate culture was centrifuged at 8000 rpm for 1 min. The supernatant was discarded, and the pellet resuspended in 1 mL fresh LB without antibiotics. This process was repeated three times to completely remove the antibiotics.

After the washing steps had been completed, 100 µL of HB101, 100 µL of DM82 and 100 µL of *Synechocystis* were combined in a 1.5 mL Eppendorf and incubated for 2 hrs in an Algaetron AG230 at 30°C with continuous warm-white LED illumination at 50 µmol photons/m^2 and a 1% CO_2 atmosphere (v/v). After incubation, the mixture was transferred to a BG11-Co-Agar plate without antibiotics, and gently spread with an L-shape spreader and left to dry. The plates were then incubated for 2 days in a in Sanyo MIR-253 light incubator at 30°C.

After 2 days, 1 mL BG11-Co was added to the plate. The cells were gently scraped with an L-shaped spreader, and the BG11-Co transferred to a new BG11-Co-Agar plate with appropriate antibiotics and left to dry. Once dry, the plates were incubated in an Algaetron AG230 at 30°C with continuous warm-white LED illumination at 50 µmol photons/m^2 and a 1% CO_2 atmosphere (v/v). After 10-14 days, colonies could be observed, and after a further 7 days, the colonies had grown to form a lawn, which could be used to inoculate liquid media as described in Section 3.2.3

3.2.6 *Synechocystis* bioproduction experiments

Synechocystis 6-well plate cultures were used to inoculate 25 mL BG11-Co with appropriate antibiotics in a 100 mL Erlenmeyer flask at a starting OD_{730} of 0.2. Cultures were then incubated in an Algaetron AG230 at 30°C, 180 rpm with continuous warm-white LED illumination at 50 µmol photons/m^2 and a 1% CO_2 atmosphere (v/v). After 2 days, the cultures were induced with cobalt (5 µM $CO(NO_3)_2(H_2O)_6$. Samples were then taken on Day 5 and Day 10.

3.3 Molecular Biology Methods

3.3.1 Extraction of plasmid DNA

Plasmid DNA was extracted from *E. coli* using a Qiagen Miniprep kit (Qiagen). A 5 mL LB culture with appropriate antibiotics in a 50 mL falcon tube, cultured overnight at 37°C, 180 rpm, was used for the extraction. The extraction was performed according to the manufacturer's protocol.

3.3.2 Quantification of DNA

DNA concentration and purity were measured using a Tecan Infinite200 Pro Nanoquant plate. Wells on the Nanoquant plate were cleaned by pipetting 1.5 µL of MilliQ onto them and then wiping the well with Kimwipe. Depending on the solvent that the DNA was dissolved in, 1.5 µL of EB buffer or MilliQ were then used to blank the wells on the Tecan Infinite200 Pro. The blank was then removed by wiping the well with Kimwipe, and 1.5 µL of sample pipetted onto the blanked well. The Tecan Infinite200Pro then measured the absorbance of the sample at 260 and 280 nm. The absorbance at 260 nm was used to quantify the DNA concentration. The ratio between the absorbances at 260 and 280 nm was calculated, and a range of 1.8-2.1 was considered acceptably pure.

3.3.3 Primers and gBlocks

All primers used in this study were synthesised by Integrated DNA Technologies (IDT). Primers were synthesised using IDT's DNA oligos service, and were ordered in tubes at a 25 nmole scale, using their Standard Desalting Purification.

All gBlocks used in this study were synthesised by IDT. DNA sequences were first codon-optimised for expression in *Synechocystis* Sp. PCC 6803 using IDT's Codon Optimization Tool, using the protein amino acid sequence as a template. Codon-optimised gBlocks were checked to ensure the absence of BsaI recognition sites, and if any were present either a new codon-optimised DNA sequence was generated, or the sequence manually edited to remove the BsaI recognition site without causing a mutation. Basic Prefix (5'-TCTGGTGGGTCTCTGTCC-3') was added to the DNA sequence before the start codon. Basic Suffix (5'- GGCTCGGGAGACCTATCG-3') was added to the DNA sequence after the stop codon. The gBlocks were then synthesised using IDT's gBlocks Gene Fragments service and were ordered in tubes. BM3 was synthesised using IDT's gene synthesis tool and was provided already cloned into the pUCIDT backbone. BM3 was also codon optimised for expression in *Synechocystis*.

3.3.4 pJET cloning

The blunt-end ligation vector pJET was used for storage of gBlocks and other DNA parts for BASIC assembly. CloneJET PCR cloning kit was used (ThermoFisher Scientific). The reaction mixture was adapted from the manufacturer's protocol as set out in Table 3.6. The reaction mix was incubated at 20°C for 25 mins, before the entire reaction mix was used to transform chemically competent *E. coli* DH5α by heat shock.

Component	Volume	Source
2X Reaction Buffer	2 µL	Thermo Fisher Scientific
pJET Blunt Vector	0.2 µL	Thermo Fisher Scientific
gBlock or PCR Product	1-3 µL	-
T4 DNA Ligase	0.5 µL	Promega

Table 3.6 pJET cloning reaction mix

3.3.5 Colony PCR

Colony PCR was used to screen colonies, after transformation with the pJET cloning reaction mix, for successful ligation of the insert into pJET. Colony PCR was performed using DreamTaq Green PCR Master Mix (2X) (Thermo Fisher Scientific). The reaction mix per reaction was prepared in PCR tubes, as set out in Table 3.7.

Component	Concentration	Volume
DreamTaq Green Mastermix	2X	12.5 µL
Fw Primer	10 µM	0.5 µL
Rv Primer	10 µM	0.5 µL
MilliQ	-	11.5 µL

Table 3.7 Colony PCR reaction mix

To screen a colony a pipette tip was used to pick half the colony. The tip was then placed inside the reaction mix and washed by pipetting up and down. After the desired colonies had been picked, the PCR tubes were transferred to a Thermocycler. The PCR conditions in Table 3.8 were then run on the Thermocycler.

Step	Temperature	Time	No of Cycles
Initial Denaturation	95°C	3 mins	1
Denaturation	95°C	30 s	35
Annealing	*	30 s	
Extension	72°C	1 min†	
Final Extension	72°C	10 mins	1
Hold	4°C	∞	1

Table 3.8 PCR conditions for colony PCR ** Dependent on primers. Annealing temperature was calculated using the online NEB Tm Calculator (New England Biolabs) †For PCR products up to 2kb. For longer products extension time was increased by 1*

3.3.6 Agarose gel electrophoresis

Gel electrophoresis was used to check that PCR products were of the appropriate length. Agarose gel was added to TAE buffer and heated in a microwave until melted. Depending on the expected length of the PCR product, different percentage agarose gels were used (Table 3.9). Once melted, the gel was left to cool to room temperature. Once cooled, SYBR Safe DNA Gel Stain (Invitrogen) was added to the gel at a ratio of 10 µL SYBR Safe per 100 mL gel. The gel was then poured into a casting basin and a comb placed to form wells. The gel was then allowed to set. Once set, the comb was removed, and the gel moved to an electrophoresis tank (Bio-Rad laboratories) containing 1X TAE buffer. 6X Loading Dye (New England Biolabs) was added to the PCR product before it was loaded into the well. If DreamTaq Green PCR 2X Master Mix was used, this step was not needed. A DNA ladder was also loaded into one of the wells. The ladder used depended on the expected length of the PCR product (Table 3.9). The gel was then run at 100V for 25-30 mins. Once finished, the gel was placed on a blue-light transilluminator (Clare Chemical Research). The DNA ladder was then used to verify that the PCR product was of the correct length.

PCR Product Length	Agarose Gel %	DNA Ladder	DNA Ladder Supplier
<0.3 kb	4	O'RangeRuler 20bp	Thermo Fisher Scientific
0.3-3 kb	1	1 kb Plus	New England Biolabs
>3 kb	0.5	1 kb Plus	New England Biolabs

Table 3.9 Agarose gel percentages and DNA ladders for PCR products of different lengths

3.3.7 Gel extraction

DNA was purified from agarose gels using the QIAquick Gel Extraction Kit (Qiagen). The purification was performed according to the manufacturer's protocol.

3.3.8 Acyl-ACP thioesterase truncation

The thioesterases CvFatB1 and UaFatB1 were truncated by a PCR reaction. The pJETs containing the wild-type thioesterases were used as a template for the PCR reaction. The variable forward primers bound at the site where the truncation was to be introduced and amplified the thioesterase from this point, while adding the BASIC prefix and if necessary, a start codon. Meanwhile the reverse primer was kept constant and amplified the thioesterase while adding the BASIC suffix. The result was a N-terminally truncated thioesterase with BASIC prefix and suffix. The amplification was performed using Q5 DNA polymerase (New England Biolabs), the reaction mixture was prepared as set out in Table 3.10, and the PCR reaction run as set out in Table 3.11. Primers for thioesterase truncation are provided in Appendix C.

Component	Concentration	Volume
Q5 Reaction Buffer	5X	5 µL
dNTPs	10 mM	0.5 µL
Fw Primer	10 µM	1.25 µL
Rv Primer	10 µM	1.25 µL
Template DNA	100 ng/µL	1 µL
Q5 DNA Polymerase	2 U/µL	0.25 µL
MilliQ	-	Up to 25 µL

Table 3.10 Q5 DNA polymerase reaction mix

Step	Temperature	Time	No of Cycles
Initial Denaturation	98°C	30 s	1
Denaturation	98°C	10 s	35
Annealing	*	30 s	
Extension	72°C	30 s/kb	
Final Extension	72°C	2 mins	1
Hold	4°C	∞	1

Table 3.11 Q5 DNA polymerase PCR conditions **Annealing temperature was calculated using the online NEB Tm Calculator (New England Biolabs)*

3.3.9 Amplification of genes from the genome

In this work, *acpP* was amplified from the genome of *E. coli*, *acp* from the genome of *Synechocystis* and *Pp0763* from the genome *Pseudomonas putida*. The genes were amplified from the genome using primers designed to do so and add the BASIC prefix and suffix. Primers for amplifying genes from the genome are provided in Appendix C.

For amplification of *E. coli acpP*, *E. coli* BW25113 Wild-Type was re-streaked onto LB-Agar. Reaction mix for Q5 DNA polymerase was prepared as described in Table 3.10, with the exception that there was no template DNA. A colony was picked using a pipette tip, and the tip washed in the reaction mix. The PCR reaction then proceeded as described in Table 3.11.

For amplification of *Synechocystis acp*, a plate containing a lawn of *Synechocystis* Sp. PCC 6803 Wild-Type was kindly provided by Pachara Sattayawat from the Jones group. The lawn was scraped into 25 µL MilliQ in a PCR tube, and the cells were subjected to 3 rounds of freeze-thaw lysis by heating the PCR tube at 95°C for 3 mins followed by freezing at -80°C for 3 mins. Reaction mix for Q5 DNA polymerase was prepared as described in Table 3.10, with the exception that instead of template DNA, 5 µL of the lysate was used as a template. The PCR reaction then proceeded as described in Table 3.11.

For amplification of *Pseudomonas putida Pp0763*, Dr Alice Banks from the Jiminez group at Imperial College London kindly provided a plate of re-streaked *Pseudomonas putida* KT2440. *Pp0763* was then amplified as described for *E. coli acpP*.

3.3.10 Production of hybrid CYP450s

Two hybrid CYP450s were made. The first consisted of the haem domain of *Rhodocyclales bacterium* CYP450 (RbCYP450) fused to the reductase domain of *Priestia megaterium* CYP102A1 BM3 (BM3). The second consisted of the haem domain of *Tepidiphilus thermophilus* CYP450 (TtCYP450) fused to the reductase domain of BM3.

Hybrid CYP450s were generated using overlap-extension PCR with Phusion DNA polymerase (New England Biolabs). The reaction mix was prepared as described in Table 3.12, and PCR conditions are described in Table 3.13. The different PCR reactions had different annealing temperatures and extension times, which are listed in Table 3.14. Primer sequences for making hybrid CYP450s are provided in Appendix C.

Firstly, the haem domains of RbCYP450 and TtCYP450 were amplified from their respective pJETs to include the BASIC prefix and an 8 bp overlap with the reductase domain of BM3. Meanwhile, the reductase domain of BM3 was amplified from pJET to include the BASIC suffix and an 8bp overlap with either the haem domain of RbCYP450 or TtCYP450. This gave a total

16 bp overlap between the two domains. This length was chosen as it provided the longest overlap possible while keeping the annealing temperatures between the two domains below 72°C, according to the NEB Tm calculator. The PCR reaction mix was then run on an agarose gel, and the appropriate bands excised and purified.

To assemble the hybrid CYP450s, a second set of PCR reaction mixes were prepared as in Table 3.12, with the exceptions that for the template DNA, 5 ng of each domain was used, and primers were not added. This reaction was run as in Table 3.13, except only 8 cycles of the PCR reaction were run. This reaction was performed to hybridise the two domains together. After this reaction was complete, the primers were added, and the PCR reaction run according to Table 3.13 to amplify the hybrid CYP450s. The PCR reaction mix was then run on an agarose gel, and the appropriate bands excised, purified, and cloned into pJET.

Component	Concentration	Volume
Phusion HF Buffer	5X	4 µL
dNTPs	10 mM	0.4 µL
Fw Primer	10 µM	1 µL
Rv Primer	10 µM	1 µL
Template DNA	100 ng/µL	1 µL
Phusion DNA Polymerase	2 U/µL	0.2 µL
MilliQ	-	Up to 20 µL

Table 3.12 Phusion DNA polymerase reaction mix

Step	Temperature	Time	No of Cycles
Initial Denaturation	98°C	30 s	1
Denaturation	98°C	10 s	35
Annealing	See Table 3.14	30 s	
Extension	72°C	See Table 3.14	
Final Extension	72°C	10 mins	1
Hold	4°C	∞	1

Table 3.13 Phusion DNA polymerase PCR Conditions

Reaction	Annealing Temperature	Extension Time
RbCYP450 Haem Domain Amplification	67°C	60 s
Rb x BM3 Reductase Domain Amplification	69°C	60 s
TtCYP450 Haem Domain Amplification	70°C	60 s
Tt x BM3 Reductase Domain Amplification	69°C	60 s
RbCYP450 Haem Domain x BM3 Reductase Domain Overlap	61°C	60 s
TtCYP450 Haem Domain x BM3 Reductase Domain Overlap	71°C	90 s
RbCYP450 x BM3 Final Amplification	69°C	90 s
TtCYP450 x BM3 Final Amplification	69°C	90 s

Table 3.14 Annealing temperatures and extension times for generating hybrid CYP450s

3.3.11 Site-directed mutagenesis

Primers were designed that were homologous 25 bp upstream and downstream of the codon that was to be mutated, and to code for the desired new codon. Primers were phosphorylated by incubating the reaction mix in Table 3.15 at 37°C for 30 mins. Primer sequences for site-directed mutagenesis are provided in Appendix C.

Component	Volume	Source
T4 Ligase Buffer	5 µL	Promega
100 µM Primer	5 µL	-
T4 Polynucleotide Kinase	1 µL	Promega
MilliQ	39 µL	-

Table 3.15 Reaction mix for site-directed mutagenesis primer phosphorylation

Once the primer was phosphorylated, the reaction mix in Table 3.16 was prepared in a PCR tube and incubated in a thermocycler running the program in Table 3.17. pJET containing the Wild-Type gene was used as template DNA.

After the program had run, 1 µL Dpn1 (New England Biolabs) was added to the reaction mix, and the reaction mix was incubated overnight at 37°C. The next day, 0.5 µL of Q5 DNA polymerase was added to the reaction mix and the program in Table 3.18 was run.

Component	Volume	Source
T4 Ligase Buffer	5	Promega
Taq DNA Ligase Buffer	2.5	New England Biolabs
10 mM dNTPs	4	New England Biolbas
Phosphorylated Primer	1	-
Template DNA	1.5	-
Q5 DNA Polymersae	0.5	New England Biolabs
Taq DNA Ligase	0.5	New England Biolabs
MilliQ	39 µL	-

Table 3.16 Reaction mix for site-directed mutagenesis

Step	Temperature	Time	No of Cycles
Preheating	68°C	5 mins	1
Initial Denaturation	98°C	3 mins	1
Denaturation	98°C	30s	25
Annealing	55°C	30s	
Extension	72°C	5 mins	
Final Extension	72°C	10 mins	1
Hold	4°C	∞	1

Table 3.17 Thermocycler program for site-directed mutagenesis

Step	Temperature	Time	No of Cycles
Initial Denaturation	98°C	3 mins	1
Denaturation	98°C	30s	2
Annealing	55°C	30s	
Extension	72°C	5 mins	
Final Extension	72°C	10 mins	1
Hold	4°C	∞	1

Table 3.18 Second thermocycler program for Site directed mutagenesis

After this program had finished, 5-10 µL of reaction mix was used to transform *E. coli* DH5α. On average, 1 in 4 colonies after transformation would contain a plasmid with the mutated gene. Successful mutagenesis was confirmed by DNA sequencing.

3.3.12 BASIC DNA assembly

Biopart Assembly Standard for Idempotent Cloning (BASIC) [402] was used to assemble all DNA constructs used in this work. To do so GFP was excised from an expression vector and replaced with the gene(s) of interest. RBS linkers were used to anneal the gene(s) of interest downstream of the expression vector's promoter sequence. The RBS linkers used incorporated different 5'-Untranslated Regions (UTR), and when assembling multiple genes in an operon, a different 5'-UTR had to be used for each gene, otherwise any 5'-UTR could be used during assembly. In this work, most genes were expressed using RBS3. After the final gene in the operon, a neutral linker was used to anneal the operon upstream of a terminator encoded on the expression vector. In this work neutral linkers L1, L2 and L3 were used depending on availability of the linker stocks, and any neutral linker could be used during assembly.

DNA sequences for plasmid backbones used in this work are provided in Appendix F. DNA sequences for genes used in this work are provided in Appendix G.

Partially double-stranded linkers for BASIC assembly were purchased from Biolegio, alternatively some double-stranded linkers were prepared from single-stranded linker and adapter oligos, synthesised by IDT. To prepare the double-stranded linkers, the reaction mix in Table 3.19 was prepared in a 1.5 mL Eppendorf tube. The reaction mix was incubated at 95°C in a heatblock for 2 hrs, then allowed to cool to room temperature before being stored at -20°C. DNA sequences for adapter and linker oligonucleotides for preparing double-stranded linkers are provided in Appendix D. DNA sequences for double stranded linkers are provided in Appendix E.

Component	**Volume**
Annealing Buffer (10 mM TRIS-HCl, 100 mM $NaCl_2$, 10 mM $MgCl_2$, pH 7.9)	5 µL
Linker Oligo (100 µM)	5 µL
Adapter Oligo (100 µM)	1 µL

Table 3.19 Reaction mix for preparing double-stranded linkers

In the first clip reaction, GFP is cut out of the cloning backbone by BsaI. Partially double-stranded linkers are ligated by T4 DNA ligase to the resulting 4 bp overhangs, creating a new 21 bp overhang. In parallel, the genes of interest are cut out of the pJET storage plasmid and have double-stranded linkers ligated to them. The reaction mix for the clip reaction is described in Table 3.20. A separate reaction mix for each individual part in the assembly was prepared in parallel and incubated as described in Table 3.21.

Component	Volume	Source
T4 DNA Ligase Buffer	3 µL	Promega
Prefix Linker	5 µL	-
Suffix Linker	5 µL	-
Biopart	4 µL	-
Bsa1-HfV2	1 µL	New England Biolabs
T4 DNA Ligase	0.5 µL	Promega
MilliQ	11.5 µL	-

Table 3.20 Reaction mix for the BASIC clip reaction

Temperature	Time
37°C	2 hrs
65°C	20 mins
20°C	20 mins
4°C	∞

Table 3.21 Incubation conditions for the BASIC clip reaction

After the Clip reaction had finished, DNA was purified from the reaction mix using Ampliclean Magnetic Beads (Nimagen), and a neodymium magnet, according to the manufacturer's protocol. For the final step, the purified DNA parts were eluted in 30 µL of MilliQ.

The purified parts were then assembled in a PCR tube according to the reaction mix in Table 3.22. The reaction mix was then incubated at 50°C for 45 mins. After incubation, the entire reaction mix was transformed into *E. coli* DH5α. Transformant cells were plated onto LB-Agar supplemented with appropriate antibiotics for the backbone and incubated overnight. The next day the plate was placed on a blue-light transilluminator (Clare Chemical Research) to check for the absence of GFP. Non-GFP colonies were picked, incubated overnight, miniprepped and sent for sequencing for confirmation of successful assembly.

Component	Volume	Source
Cutsmart Buffer	1 µL	Promega
Part (per part)	1 µL	-
MilliQ	Up to 10 µL	-

Table 3.22 Reaction mix for the BASIC assembly reaction

3.3.13 DNA sequencing

DNA sequencing was used to screen for mutants during site directed mutagenesis, to check for successful generation pJET cloning and to check for successful BASIC assembly. NGS Sanger Sequencing by SourceBioscience was used for DNA sequencing. Sequencing samples and primers were prepared according to their specifications, and samples were dropped off at a collection point at Imperial College London. Sequencing results were analysed using Benchling. Primer sequences used for sequencing are provided in Appendix C.

3.3.14 Knock-out of *iscR* from the genome of *E. coli*

iscR was knocked-out of the genome of *E. coli* BW25113 by adapting previously reported protocols[41,42]. Three plasmids were used, pKD46 encoding λ-Red recombinase under control of an L-arabinose inducible promoter, pCP20 encoding FLP-ase and pKD3 encoding a chloramphenicol resistance cassette flanked by FRT sites. Both pKD46 and pCP20 had temperature sensitive origins of replication, and needed to be maintained at 30°C. All three plasmids had a carbenicillin resistance marker.

Firstly, primers were designed to amplify the chloramphenicol resistance cassette flanked by FRT sites from pKD3. The primers had 25 bp homology upstream and downstream of the FRT sites and had 50 bp overhangs that were homologous to the regions immediately upstream and downstream of the *iscR* locus. Amplification of the resistance cassette was done using Q5 DNA polymerase as described in Section 3.3.8, except that only 10 ng of template DNA was used. The annealing temperature used was for that of the 25 bp overlaps between the primer and template, not for the entire primer sequence. The PCR reaction was repeated four times. Primer sequences used for amplification of pKD3 are provided in Appendix C. Afterwards, 0.5 µL Dpn1 was added to the PCR tubes, and the reaction mix was incubated at 37°C for 1 hr. The reaction mix was then run on an agarose gel as described in Section 3.3.6, and all four bands purified from the gel, as described in Section 3.3.7, and pooled together. The gel purification product was then stored at -20°C.

pKD46 was transformed by heat-shock into chemically competent *E. coli* ΔfadD, as described in Section 3.1.7, except that 30°C, instead of 37°C, was used for recovery of cells after heat-shock and for overnight incubation of the plates. A colony was picked and inoculated into 5 mL LB with carbenicillin in a 50 mL falcon tube and grown overnight at 30°C, 150 rpm. This overnight culture was then used to make electrocompetent cells, as described in Section 3.1.8 with the exception that during the growth of the 50 mL culture after inoculation with the overnight culture, when the 50 mL culture reached an OD_{600} of 0.1 it was induced with 15 mM filter-sterilised L-arabinose. Electrocompetent *E. coli* ΔfadD + pKD46 were transformed with the gel purification product by electroporation as described in Section 3.1.9. LB-Agar plates with 30% the normal working concentration of chloramphenicol were grown overnight at 37°C. Individual colonies were then restreaked to a fresh LB-Agar plate with 30% the normal working concentration of chloramphenicol and grown overnight at 37°C. Colonies were then screened by colony PCR for successful integration of the chloramphenicol resistance cassette into the genome and removal of *iscR*. Primer sequences used for colony PCR during knock-out of *iscR* are provided in Appendix C.

To remove the chloramphenicol resistance cassette, calcium competent *E. coli* ΔfadD ΔiscR:Frt-ChlR-Frt were prepared and transformed with pCP20 as described in Sections Section 3.1.6 and 3.1.7, except that 30°C, instead of 37°C, was used for recovery of cells after heat-shock and for overnight incubation of the plates. A colony was picked in 5 mL LB with no antibiotics in a 50 mL falcon tube and grown at 30°C, 150 rpm for 2 hrs. The 5 mL culture was then grown at 42°C for 6 hrs. 20 µL, 200 µL and 500 µL were plated onto LB-Agar with no antibiotics and grown overnight at 37°C. The next day, two LB-Agar plates were prepared, one with 30% the normal working concentration of chloramphenicol and one without. A matching grid was drawn on each plate. Colonies were picked from the original plate using a pipette tip. The pipette was then used to inoculate the two new plates by first touching the pipette tip to the LB-Agar-Chl plate followed by touching the pipette tip to the LB-Agar plate. Using the grid, the colony was inoculated to the same location on both plates. The two plates were then incubated overnight at 37°C. The next day colony growth was observed on the two plates. If a colony grew at the same location on both plates, this was interpreted as failure to remove the chloramphenicol resistance cassette. If a colony grew on only the LB-Agar-Chl plate this was interpreted as being inconclusive. If a colony grew on only the LB-Agar plate this was interpreted as potential removal of the chloramphenicol resistance cassette. These colonies were then screened by colony PCR, as described to confirm removal of the chloramphenicol resistance cassette. Positive colonies were then used to inoculate 5 mL LB with and without chloramphenicol in a 50 mL flacon tube and incubated overnight at 37°C, 180 rpm. If growth

was only observed in the overnight culture without chloramphenicol, the culture was then used to make a glycerol stock for cryopreservation.

3.4 Product Analysis

3.4.1 Measurement of *E. coli* OD_{600}

20 µL of cell culture was diluted in 180 µL MilliQ in a 96-well plate (SLS). A blank was prepared by adding 200 µL MilliQ to a separate well. Absorbance at 600 nm was measured using a Tecan Infinite M200 Pro Spectrophotometer.

3.4.2 Measurement of *Synechocystis* OD_{730}

100 µL of cell culture was diluted in 900 µL BG11-Co in a cuvette (VWR). Absorbance at 730 nm was measured using a Tecan Infinite M200 Pro Spectrophotometer that had been blanked using a cuvette containing 1 mL BG11-Co.

3.4.3 Gas-Chromatography Mass-Spectrometry (GC-MS)

GC-MS was used for the detection and quantification of volatile products. Samples were analysed using an Agilent 7890B GC with an Agilent DB-WAXetr column and an Agilent 7693 autosampler, connected to an Agilent 5977B Mass Selective Detector. Samples were injected in Pulsed-Split mode in a 3:1 ratio with a flow rate of 1.5 mL/min. The inlet temperature was 250°C and inlet pressure was 10.3 psi. The carrier gas helium at a flow rate of 1.5 mL/min. The autosampler needler was washed with acetone and hexane before injecting samples. Acetone samples were run before samples to clean the column from any residue from previous experiments.

3.4.4 High-Performance Liquid Chromatography (HPLC)

HPLC was used for the detection and quantification of soluble products. Samples were analysed using an Agilent 1200 Series HPLC equipped with Aminex HPX-87H column (Bio-Rad Laboratories). Product detection and quantification was performed using a Refractive Index Detector (RID). The liquid phase was 5 mM H_2SO_4 at a flow rate of 0.6 mL/min. The column was heated to 60°C.

3.4.5 Extraction of fatty alcohols in the presence of an isopropyl myristate overlay

A 10% v/v isopropyl myristate overlay was used to extract fatty alcohols from the culture media during bioproduction experiments. After the cultures had grown for 48 hrs the overlay developed a frothy appearance due to emulsion of the overlay and culture media. 1 mL of this emulsion was transferred to an Eppendorf tube and centrifuged at 13000 rpm for 1 min. After centrifugation 100 µL of the top layer (the overlay) was transferred to a 2 mL glass screw-cap

vial (Agilent) with a 250 µL insert (Agilent). Samples were then capped with a blue screw cap (Agilent).

Samples were run on the GC-MS with the following program. The oven temperature was initially set at 100°C and held for 30 s. The temperature was then ramped at a rate of 30°C /min to a final temperature of 250°C, where it was held for 1 min.

3.4.6 Extraction of decanoic acid and decalactones from the culture media

Cell extraction with chloroform-methanol was used to extract decanoic acid and decalactones from the culture media during feeding and bioproduction experiments. After 48 hrs cultivation 1 mL cell culture was transferred to a capped glass tube. If the cell and aqueous phase were to be separately examined 1 mL cell culture was transferred to a 1.5 mL Eppendorf and centrifuged at 13000 rpm for 10 mins. The supernatant was then transferred to a capped glass tube. The pellet was resuspended in 1 mL fresh media and transferred to a capped glass tube. 2 mL of 2:1 chloroform methanol was added to the glass tube. If the sample was to be acidified 150 µL of 6 M HCl was also added to the glass tube, if not this step was skipped. The glass tube was vortexed and then centrifuged at 3500 rpm for 3 mins. After centrifugation, the bottom phase (chloroform) was transferred to a new capped glass tube using a glass Pasteur pipette. Samples were then dried under a nitrogen gas flow. Once dry, 100 µL of chloroform was added to the glass tube, and the glass tube was vortexed again. The chloroform was then transferred using a glass Pasteur pipette to a 2 mL glass screw-cap vial with a 250 µL insert. Samples were then capped with a blue screw cap.

Samples were run on the GC-MS with the following program. The oven temperature was initially set at 160°C and held for 30 s. The temperature was then ramped at a rate of 50°C /min to 200°C, where it was held for 2 mins. The temperature was then ramped again at 50°C /min to a final temperature of 250°C, where it was held for 1 min.

3.4.7 Extraction of decanoic acid and decalactones in the presence of a decane overlay

A 10% v/v decane overlay was used to extract decanoic acid and decalactones from the culture media during some feeding and bioproduction experiments. After the cultures had grown for 48 hrs the overlay developed a frothy appearance due to emulsion of the overlay and culture media. 1 mL of this emulsion was transferred to an Eppendorf tube and centrifuged at 13000 rpm for 1 min. After centrifugation 100 µL of the top layer (the overlay) was transferred to a 2 mL glass screw-cap vial. with a 250 µL insert. Samples were then capped with a blue screw cap.

Samples were run on the GC-MS with the following program. The oven temperature was initially set at 100°C and held for 30 s. The temperature was then ramped at a rate of 30°C /min to a final temperature of 250°C, where it was held for 2.5 mins.

3.4.8 Extraction of C4 esters in the presence of a hexadecane overlay

A 10% v/v hexadecane overlay was used to extract butyl acetate and ethyl butyrate from the culture media during bioproduction and feeding experiments. After the cultures had grown for 48 hrs the overlay developed a frothy appearance due to emulsion of the overlay and culture media. 1 mL of this emulsion was transferred to an Eppendorf tube and centrifuged at 13000 rpm for 1 min. After centrifugation 100 µL of the top layer (the overlay) was transferred to a 2 mL glass screw-cap vial with a 250 µL insert. Samples were then capped with a blue screw cap.

Samples were run on the GC-MS with the following program. The oven temperature was initially set at 100°C and held for 30 s. The temperature was then ramped at a rate of 30°C /min to a final temperature of 250°C, where it was held for 1 min.

3.4.9 Extraction of methyl ketones in the presence of an isopropyl myristate overlay

A 10% v/v isopropyl myristate overlay was used to extract methyl ketones from the culture media during bioproduction experiments. After the cultures had grown for 48 hrs the overlay developed a frothy appearance due to emulsion of the overlay and culture media. 1 mL of this emulsion was transferred to an Eppendorf tube and centrifuged at 13000 rpm for 1 min. After centrifugation 100 µL of the top layer (the overlay) was transferred to a 2 mL glass screw-cap vial with a 250 µL insert. Samples were then capped with a blue screw cap.

Samples were run on the GC-MS with the following program. The oven temperature was initially set at 100°C and held for 30 s. The temperature was then ramped at a rate of 30°C /min to a final temperature of 250°C, where it was held for 1 min.

3.4.10 Extraction and decarboxylation of β-ketofatty acids from the culture media

A previously reported protocol[337] was used to extract β-ketofatty acids from the culture media and decarboxylate them to methyl ketones by heat and acid treatment. 500 µL cell culture was transferred to a 2 mL Eppendorf. 500 µL of 2M H_2SO_4 was added, and the Eppendorf incubated at 75°C for 30 mins, then at 30°C for 30 mins using a heatblock. After incubation 500 µL of hexane was added to the Eppendorf, and the Eppendorf was vortexed then centrifuged at 5000 rpm for 10 mins. After centrifugation 100 µL of the top layer (hexane) was transferred to a 2 mL glass screw-cap vial with a 250 µL insert. Samples were then capped with a blue screw cap.

Samples were run on the GC-MS with the following program. The oven temperature was initially set at 100°C and held for 30 s. The temperature was then ramped at a rate of 30°C /min to a final temperature of 250°C, where it was held for 1 min.

3.4.11 Analysis of soluble products

In this work HPLC was used to quantify butanol, butyric acid, 2-pentanone, 2-petanol and 2-butenol. After bioproduction and feeding experiments 1 mL of cell culture was transferred to a 1.5 mL Eppendorf and centrifuged at 13000 rpm for 10 mins. 500 µL of supernatant was transferred to 2 mL glass screw-cap vial and the sample was capped. Samples were then run on the HPLC. For ethanol and butyric acid samples were run for 25 mins, for butanol, 2-butenol, 2-pentanone and 2-pentanol samples were run for 40 mins.

3.4.12 Analysis of glucose consumption

After bioproduction and feeding experiments 1 mL of cell culture was transferred to a 1.5 mL Eppendorf and centrifuged at 13000 rpm for 10 mins. 500 µL of supernatant was transferred to 2 mL glass screw-cap vial and the sample was capped. Samples were then run on the HPLC for 12 mins.

3.5 Bioinformatics Methods

3.5.1 Prediction of the chloroplast localisation peptide

The chloroplast localisation peptide was predicted using the ChloroP server (http://www.cbs.dtu.dk/services/ChloroP/)[403]. At the time of writing this server has now been discontinued.

3.5.2 Protein sequence alignments

Protein sequence alignments were either performed using Benchling (Clustal Omega) or by using EMBL-EBI Multisequence Alignment Tool (Clustal Omega)[404].

3.5.3 Generation of predicted protein structures

Predicted protein structures were generated using the Phyre2 server (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index)[405]. The amino acid sequence was inputted, and normal modelling mode was used. To generate multiple predicted protein structures at once Batch Processing mode was used.

3.5.4 Structure guided mutagenesis

Protein structures were visualised using PyMol[406]. The Mutagenesis Wizard feature was used to investigate potential mutations that would cause steric hindrance at the appropriate location on the bound ligand.

3.5.5 Alignment of protein structures

Alignment of protein structures was performed using the mtm-align server (https://yanglab.nankai.edu.cn/mTM-align/)[407]. At the time of writing, this server appears to be no longer be available.

3.5.6 Blast

Blast (Basic Local Alignment Search Tool) was performed using the online web server from NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi)[408]

3.6 Statistical Analysis

A Welch T-test was performed using R Studio. A p-value of <0.05 was taken to be significant, and all p-values are displayed to 3 significant figures. Data presented is the average of 3 biological replicates, and error bars of ± 1 Standard Deviation are shown.

Chapter 4

Engineering a C10 Acyl-ACP Thioesterase

4.1 Aims

γ-Decalactone and δ-decalactone are commercially important high impact and value flavour and fragrance compounds[15,79,84,87]. The native pathway for their biosynthesis is unknown[87], and prior work in the literature has focused on the biotransformation of long chain fatty acids to medium chain length γ- and δ-lactones[81,83,85,88-95].

In this work a novel metabolic pathway for the biosynthesis of γ- and δ-decalactone has been proposed. In order to biosynthesise γ- and δ-decalactone, decanoic acid first needs to be produced. To do so, a C10 specific acyl-ACP thioesterase will be used.

No C10 specific thioesterase, defined here as >66.6% C10 specificity had yet been reported in the literature, therefore engineering of an existing thioesterase to improve the C10 specificity will be required. The FatB family of acyl-ACP thioesterases are a promising starting point for engineering a C10 thioesterase. While it is known that FatB acyl-ACP thioesterases require N-terminal truncation to remove a Chloroplast Localisation Peptide for optimal expression in bacteria[105,109,126], and that different truncations of the same thioesterase can result in different activities and specificities[126,128], the effect of truncation upon thioesterase activity and specificity has not yet been systematically studied. It was therefore not clear where the optimal site for truncation of a FatB thioesterase was located.

The aims of this Chapter are therefore to investigate the effect of truncation upon FatB acyl-ACP thioesterase activity and specificity, in order to identify the most optimal truncation site. Once this has been achieved, and if necessary, the truncated thioesterase will then be engineered to be C10 specific.

4.2 Results

4.2.1 Truncation of *Cuphea viscosissima* FatB1

The acyl-ACP thioesterase *Cuphea viscosissima* FatB1 (CvFatB1) had the highest reported C10 specificity (25.7%) in the literature, as measured in moles of decanoic acid as a percentage of total moles of fatty acids produced in *Escherichia coli* overexpressing CvFatB1[104]. It was therefore chosen as a candidate for engineering a C10 specific thioesterase. The first step in engineering CvFatB1 was to investigate the effect of N-terminal truncation.

A set of systematic truncations was designed and produced. The truncations were designed using three approaches. Firstly, CvFatB1 was truncated to excise the chloroplast localisation peptide predicted by the ChloroP server. Secondly, CvFatB1 was truncated according to previously reported truncations of FatB thioesterases in the literature[105,109,123,126,129,130,134,135]. Finally, to cover for the possibility that the optimal truncation site was located elsewhere, CvFatB1 was systematically truncated every 10 amino acids, extending from the site of greatest truncation reported in the literature. In total this process generated 21 unique truncations of CvFatB1, some of which had already been produced in prior work by the Jones lab (Table 4.1).

To screen the CvFatB1 truncations, a previously reported alcohol biosynthesis pathway[409], that had also been used by the Jones lab[28,150], was used (Fig 4.1A). All heterologous genes in the pathway were expressed on a single plasmid (Fig 4.1B). By screening for alcohol biosynthesis, rather than acid biosynthesis, product analysis was simpler and faster, allowing for a larger set of truncations to be screened.

CvFatB1 Truncation	Truncation Site	Rationale	Origin	Reference
CvT1	A58	Removal of CLP predicted by ChloroP	Prior work	403
CvT2	L89	Truncation from the literature	Prior work	105, 109
CvT3	M94	Truncation from the literature	This study	126, 135
CvT4	E107	Truncation from the literature	This study	134
CvT5	M112	Truncation from the literature	Prior work	126
CvT6	R119	Truncation from the literature	This study	123
CvT7	K130	Truncation from the literature	This study	129, 130
CvT8	R140	Systematic truncation	This study	-
CvT9	I150	Systematic truncation	This study	-
CvT10	T160	Systematic truncation	This study	-
CvT11	I170	Systematic truncation	This study	-
CvT12	N180	Systematic truncation	This study	-
CvT13	M189	Systematic truncation	This study	-
CvT14	K200	Systematic truncation	This study	-
CvT15	T210	Systematic truncation	This study	-
CvT16	W220	Systematic truncation	This study	-
CvT17	M229	Systematic truncation	This study	-
CvT18	T240	Systematic truncation	This study	-
CvT19	V250	Systematic truncation	This study	-
CvT20	R260	Systematic truncation	This study	-
CvT21	Q270	Systematic truncation	This study	-

Table 4.1 CvFatB1 Truncations *Truncations were designed using the ChloroP server, the literature or by systematic truncation of the protein. Some truncations were generated in prior work by the Jones lab.*

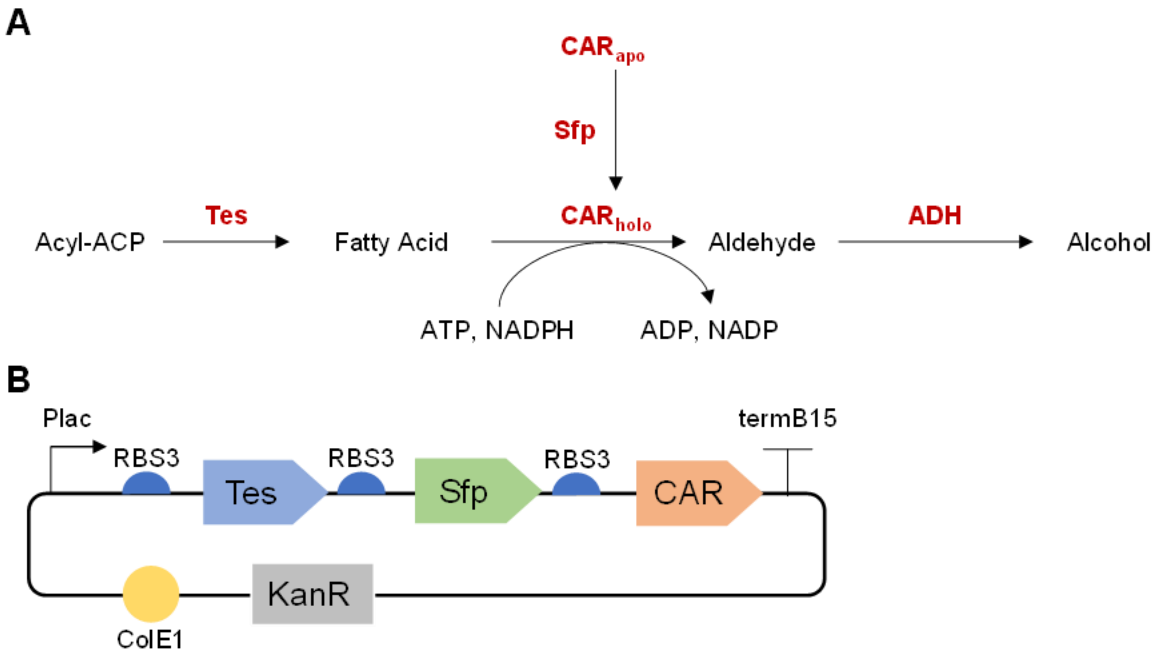


Figure 4.1 Pathway and plasmid design for screening CvFatB1 truncations *A) Synthetic metabolic pathway for alcohol biosynthesis. Pathway enzymes are highlighted in red. Tes: CvFatB1 truncation, Sfp: Phosphopantetheine transferase (Bacillus subtilis), CAR: Carboxylic acid reductase (Mycobacterium marinum), ADH: Native alcohol dehydrogenases B) Plasmid design for expressing the alcohol biosynthesis pathway. A pET plasmid with ColE1 origin of replication and kanamycin resistance cassette was used to express pathway enzymes. Plac and RBS3 were used to express a CvFatB1 truncation, Sfp and CAR in a single operon.*

No alcohol biosynthesis was observed when expressing the untruncated CvFatB1 wild-type (WT). The highest alcohol titres were observed when truncations CvT4 and CvT5 were expressed. The majority of alcohols produced after expression of CvT4 and CvT5 were octanol (127.1 mg/L and 157.7 mg/L, respectively) and decanol (85.5 mg/L and 90.8 mg/L, respectively). From CvT8 onwards, there was a loss of activity with only low titres of alcohols produced, suggesting that these truncations were detrimental to thioesterase activity (Fig 4.2A).

Every truncation except CvT6 and CvT7 showed improved growth relative to the WT (OD_{600} = 1.10). Both CvT6 (OD_{600} = 0.54, p = 0.0314) and CvT7 (OD_{600} = 0.57, p = 0.00646) showed significant reductions in OD_{600} after 48 hrs cultivation relative to the WT (Fig 4.2C). When alcohol yield was investigated, truncations CvT4, CvT5, CvT6 and CvT7 were the most productive, with respective octanol yields of 18.7, 23.6, 17.1 and 21.4 mg/g glucose, and respective decanol yields of 12.6, 13.5, 4.2 and 7.8 mg/g glucose (Fig 4.2B).

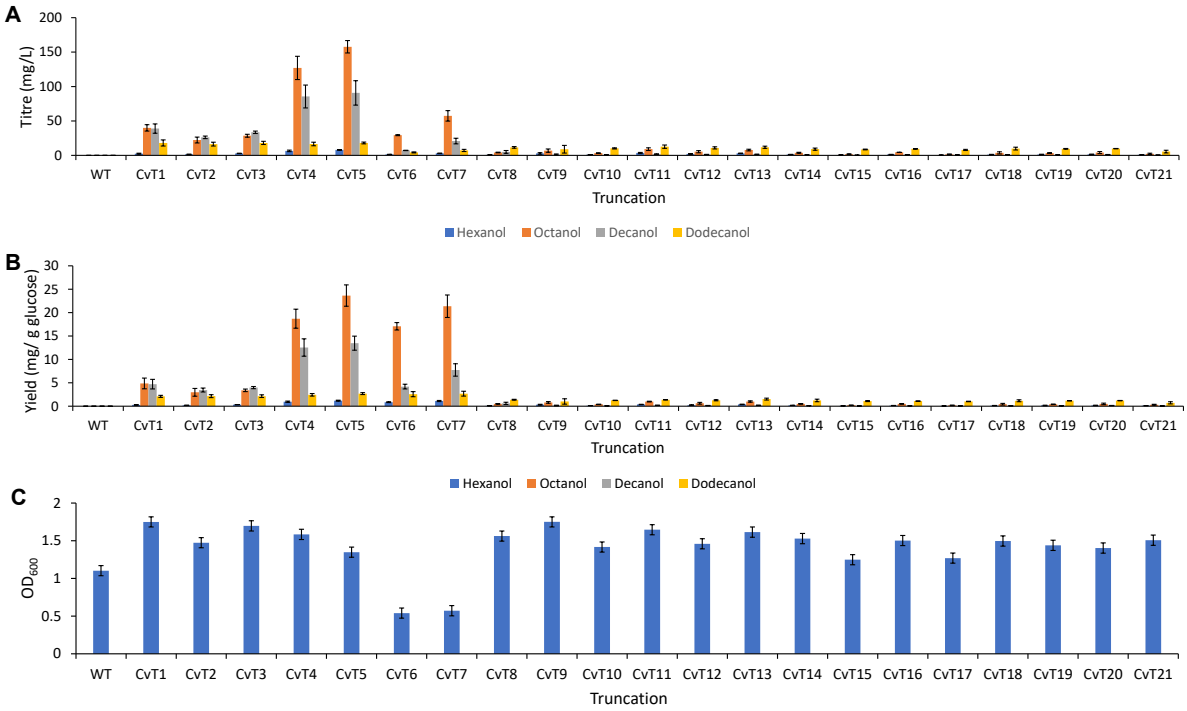


Figure 4.2 Screening CvFatB1 truncations *E. coli ΔfadD were transformed with a plasmid for alcohol biosynthesis with different CvFatB1 truncations A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with 0.5 mM IPTG. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.2 Truncation of *Ulmus americana* FatB1

Ulmus americana FatB1 (UaFatB1) was reported to have similar C10 specificity to CvFatB1 (23.4% vs 25.7%) while producing 10-fold higher total fatty acid titres (1098 nmol/mL vs 79.2 nmol/mL) when expressed in *E. coli*[104]. UaFatB1 was therefore also selected as a candidate for engineering a C10 thioesterase. To assay the effects of truncation upon UaFatB1, the amino acid sequences of CvFatB1 and UaFatB1 were aligned, and truncations of UaFatB1 were designed that incorporated the same truncations as on CvFatB1 (Fig 4.3). As an activity loss from CvT8 onwards had previously been observed, a smaller set of 9 truncations of UaFatB1 was produced (Table 4.2).

```
CvFatB1 MVAAAATSAFFPVPAPGTSPKPGKSGNWPSSLSPTFKPKSIPNGGFQVKANASAHPKANG 60
UaFatB1 ------------------------------------MGSGALQVKASSQAPPKLNG 20

CvFatB1 SAVNLKS-GSLNTQEDTSSSPPPRAFLNQLPDWSMLLTAITTVFVAAEKQWTMLDRKSKR 119
UaFatB1 SNVGLVKSSQIVKKGDDTTSPPARTFINQLPDWSMLLAAITTLFLAAEKQWMMLDWKPKR 80

CvFatB1 PDMLVDSVGLKSIVRDGLVSRHSFSIRSYEIGADRTASIETLMNHLQETTINHCKSLGLH 179
UaFatB1 PDMLVDPFGLGRFVQDGLVFRNNFSIRSYEIGADRTASIETLMNHLQETALNHVKSVGLL 140

CvFatB1 NDGFGRTPGMCKNDLIWVLTKMQIMVNRYPTWGDTVEINTWFSQSGKIGMASDWLISDCN 239
UaFatB1 EDGLGSTREMSLRNLIWVVTKMQVAVDRYPTWGDEVQVSSWATAIGKNGMRREWIVTDFR 200

CvFatB1 TGEILIRATSVWAMMNQKTRRFSRLPYEVRQELTPHFVDSPHV--IEDNDQKLRKFDVKT 297
UaFatB1 TGETLLRATSVWVMMNKLTRRISKIPEEVWHEIGPSFIDAPPLPTVEDDGRKLTRFDESS 260

CvFatB1 GDSIRKGLTPRWNDLDVNQHVSNVKYIGWILESMPIEVLETQELCSLTVEYRRECGMDSV 357
UaFatB1 ADFIRSGLTPRWSDLDINQHVNNVKYIGWLLESAPPEIHESHEIASLTLEYRRECGRDSV 320

CvFatB1 LESVTAVDPSENGGRS--QYKHLLRLEDGTDIVKSRTEWRPKNAGTNG-AISTSTAKTSN 414
UaFatB1 LNSATKVSDSSQLGKSAVECNHLVRLQNGGEIVKGRTVWRPKRPLYNDGAVVDVPAKTS- 379

CvFatB1 GNSVS 419
UaFatB1 ----- 379
```

Figure 4.3 Alignment of CvFatB1 and UaFatB1 *Amino acid sequences of CvFatB1 and UaFatB1 were aligned. Sites where CvFatB1 and UaFatB1 were truncated are highlighted in red. Alignment was performed using EMBL-EBI.*

UaFatB1 Truncation	Truncation Site
UaT1	L17
UaT2	L49
UaT3	M54
UaT4	E67
UaT5	M72
UaT6	K79
UaT7	G90
UaT8	R100
UaT9	I110

Table 4.2 UaFatB1 Truncations *Truncations of UaFatB1 were designed that incorporated the same truncations used in CvFatB1.*

UaFatB1 truncations were screened using the same alcohol biosynthesis pathway and plasmid design as used in Section 4.2.1 to screen the CvFatB1 truncations.

Alcohol titres were lower than those observed for CvFatB1 truncations, however most of the truncations were observed to produce more decanol than octanol. It was also observed that unlike the CvFatB1 WT, the untruncated UaFatB1 wild-type (WT) was active, and its expression led to an octanol titre of 17.0 mg/L and the highest observed decanol titre of 29.9 mg/L. The most active truncation was UaT1, with respective octanol and decanol titres of 22.8 and 28.5 mg/L (Fig 4.4A).

UaT1 produced significantly higher yield of octanol (10.1 mg/g glucose, $p = 0.0442$) than the WT, while the truncation UaT4 produced higher yields of octanol (11.2 mg/g glucose) than the WT, albeit not significantly higher ($p = 0.0849$). Both UaT1 and UaT4 had higher decanol yields than the WT, but not significantly so (12.7 mg/g glucose, $p = 0.0741$ and 11.1 mg/g glucose, $p = 0.185$, respectively) (Fig 4.4B).

UaT1 and UaT4 also showed reduced OD_{600} relative to the UaFatB1 WT, suggesting their expression had a negative effect on cell growth.

As higher decanol than octanol titres were observed, C10 specificity, measured as the percentage of moles of total fatty alcohols that were decanol, was calculated. UaFatB1 WT had the highest C10 specificity of 51.3% (Fig 4.4D).

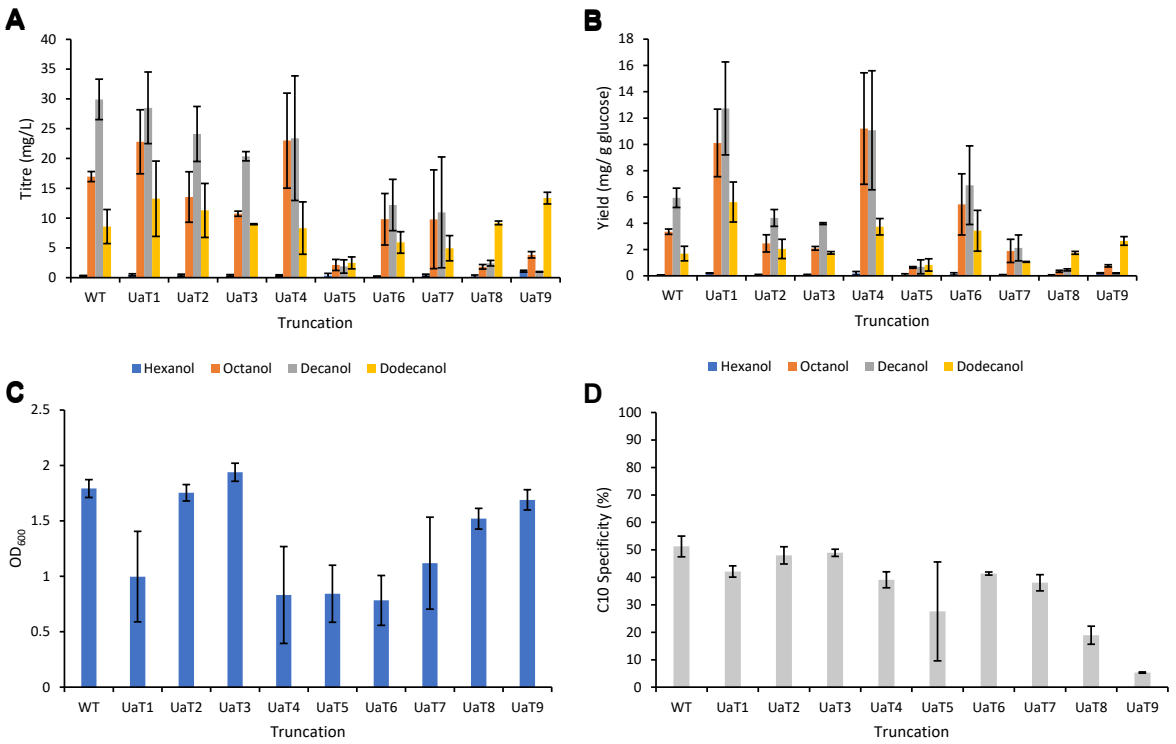


Figure 4.4 Screening UaFatB1 truncations *E. coli ΔfadD were transformed with a plasmid for alcohol biosynthesis with different UaFatB1 truncations A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. Cells were induced with 0.5 mM IPTG. The overlay was sampled and analysed by GC-MS. D) C10 specificity, calculated as the percentage of total moles of fatty alcohol that were decanol. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.3 Screening CvFatB1 truncations in *Synechocystis* Sp. PCC 6803

As truncations of CvFatB1 had higher product titres than truncations of UaFatB1, it was decided to further investigate only the CvFatB1 truncations. Truncations CvT3-T8 were chosen for further investigation.

In future work, the C10 thioesterase will be incorporated into a synthetic pathway for the biosynthesis of δ-decalactone and γ-decalactone, and it will be of interest to express this pathway in cyanobacteria. The truncations CvT3-T8 were therefore screened in the cyanobacteria *Synechocystis* Sp. PCC 6803. An engineered strain from the Jones lab, *Syenchocystis* Δaas with Sfp and CAR integrated into the genome[28] was used to screen the thioesterases. As Sfp and CAR were integrated into the genome, only the truncation needed to be expressed using a plasmid (Fig 4.5).

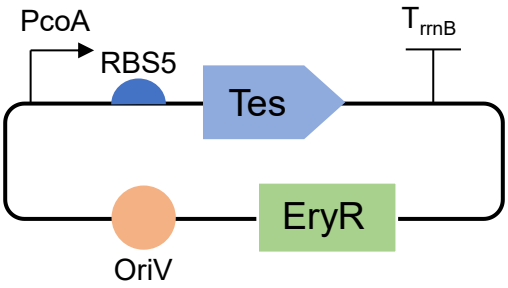


Figure 4.5 Plasmid design for expressing CvFatB1 truncations in Synechocystis *A pRSF1010 plasmid with erythromycin resistance cassette was used to express a truncation of CvFatB1 (Tes). The truncation was expressed using the PcoA promoter and RBS5.*

Expression of truncations CvT4-6 led to the highest alcohol titres in *Synechocystis*. Unlike in *E. coli* more decanol than octanol was produced. For example, following expression of CvT5 higher decanol (112.4 mg/L) than octanol (24.9 mg/L) titres were observed after 10 days cultivation. Decanol titres after expression of CvT5 were notably higher than the decanol titres obtained when CvT6, which was the second most active thioesterase, was expressed (90.5 mg/L), however due to the large variation within the triplicas this difference was not significant ($p = 0.214$) (Fig 4.6A).

Synechocystis expressing CvT5 and CvT6 had the least growth after 10 days (OD_{730} = 2.10 and OD_{730} = 1.82, respectively), with no significant difference between the two strains ($p = 0.249$) (Fig 4.6B).

C10 specificity was calculated as before, with CvT4 showing the highest C10 specificity (71.9%) which was significantly higher than the C10 specificities of CvT5 (63.1%, $p = 0.0135$) and CvT6 (59.4%, $p = 0.0376$). There was no significant difference in C10 specificity between CvT5 and CvT6 ($p = 0.276$) (Fig 4.6C).

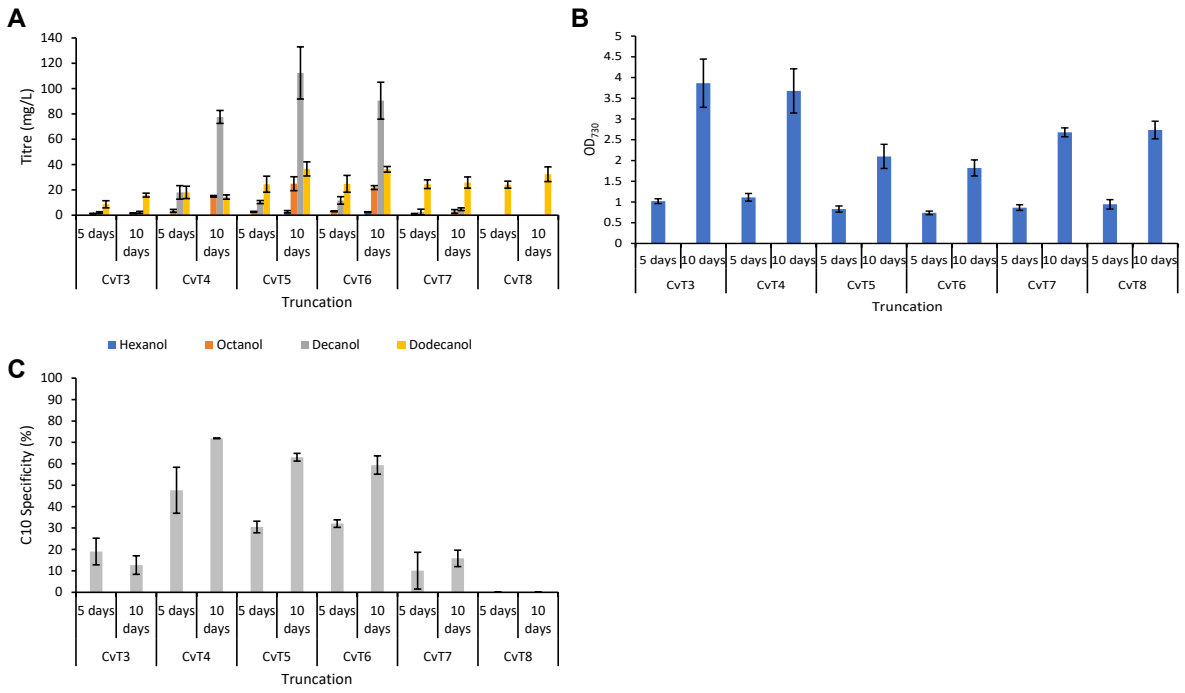


Figure 4.6 Screening CvFatB1 truncations in Synechocystis *Synechocystis Δaas with Sfp-CAR integrated into the genome were transformed with a plasmid expressing a CvFatB1 truncation. A) Alcohol titres B) OD_{730} after 5 and 10 days cultivation in BG11-Co media with a 30% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with 5 μM Cobalt. C) C10 specificity, calculated as the percentage of total moles of fatty alcohol that were decanol. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.4 Inducer optimisation of CvFatB1 truncation expression

It was previously observed in Section 4.2.1 that alcohol titres and growth were lower in *E. coli* expressing CvT6 and CvT7 compared to those expressing CvT4 and CvT5. A possible explanation was that overexpression of these truncations had toxicity effects. Therefore, if expression of CvT6 and CvT7 was optimised, growth and therefore alcohol titres could be improved. To optimise expression of the truncations, the inducer concentration was varied.

Previously, a truncation, Sfp and CAR were expressed in a single operon under Plac, meaning varying IPTG concentration would also affect expression of Sfp and CAR, which could also affect alcohol titres. To control for this, the operon was split across two plasmids, with Plac used to express a truncation, and the rhamnose inducible PrhaBAD promoter used to express Sfp and CAR (Fig 4.7).

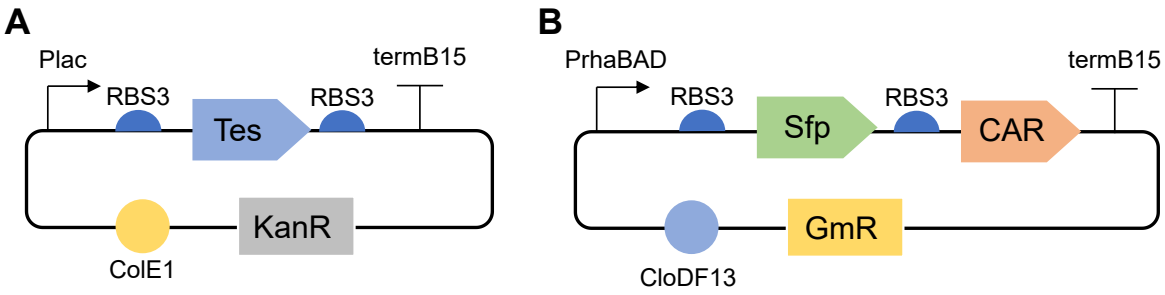


Figure 4.7 Plasmid design for inducer optimisation of CvFatB1 truncations *A) A pET plasmid with kanamycin resistance cassette was used to express a truncation of CvFatB1 (Tes). The truncation was expressed using the Plac promoter and RBS3 B) A pCDF plasmid with gentamycin resistance cassette was used. Sfp and CAR were expressed in a single operon using PrhaBAD and RBS3.*

4.2.4.1 Inducer Optimisation of CvT3 expression

Changing the concentration of IPTG used to induce expression of CvT3 had no effect on dodecanol titres. As IPTG concentration was reduced from 500 µM to 250 µM there was small increase in decanol titres from 28.1 mg/L to 32.9 mg/L, however this was not significant ($p = 0.304$). As IPTG concentration was further reduced, octanol and decanol titres fell. Alcohol production was still observed even when no IPTG was added, indicating that there was some leaky expression from the promoter (Fig 4.8A). The same trend was also observed for the alcohol yields (Fig 4.8B). Conversely, as IPTG concentration was reduced, OD_{600} increased, indicating that expression of CvFatB3 imposed a growth burden upon the cell (Fig 4.8C).

As induction with 250 µM IPTG led to the highest, albeit not significantly so, alcohol titres, this was taken as the optimum IPTG concentration for expression of CvT3.

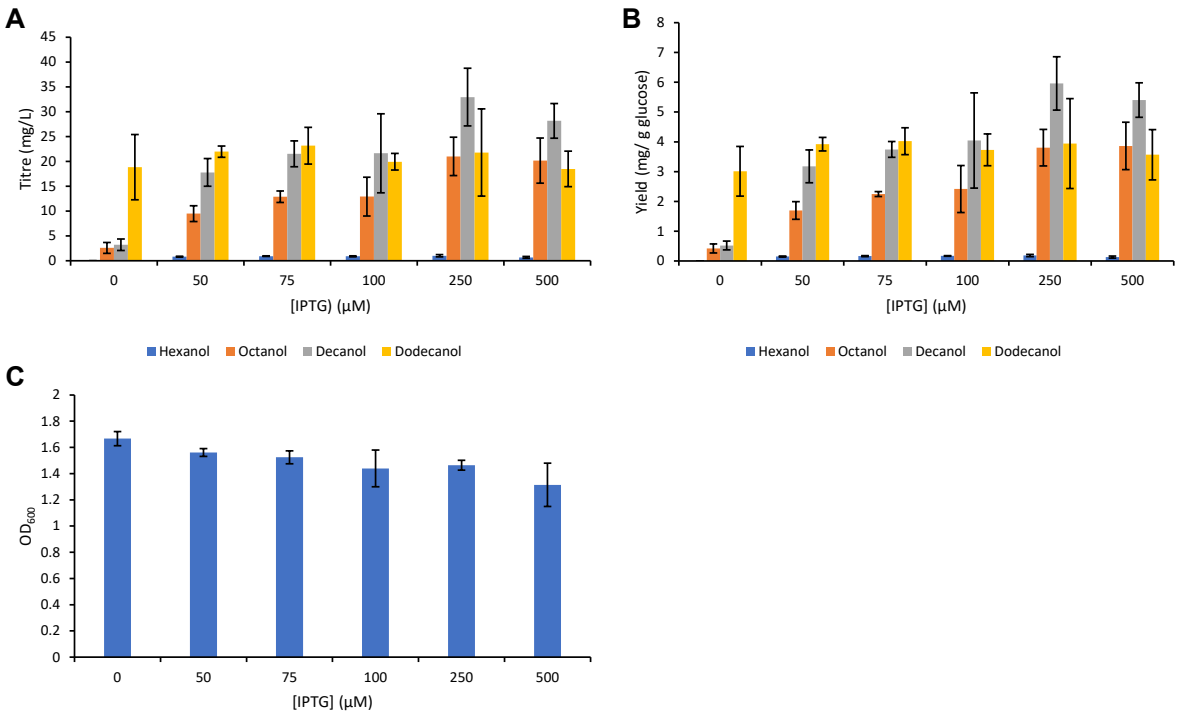


Figure 4.8 Inducer optimisation of CvT3 expression *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with a variable amount of IPTG and 1% rhamnose. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.4.2 Inducer Optimisation of CvT4 expression

As the concentration of IPTG used to induce CvT4 decreased to 250 µM, octanol titres initially increased from 81.4 to 88.4 mg/L, but this was not significant (p = 0.1125). As the concentration of IPTG further decreased, octanol titres trended downwards albeit not consistently. Meanwhile, as IPTG concentration decreased decanol titres increased and peaked at induction with 75 µM IPTG, when 63.7 mg/L of decanol was produced. However, this was not a significant increase compared to decanol titres at induction with 500 µM IPTG (47.9 mg/L, p = 0.117) (Fig 4.9A). The highest decanol yields were also seen at induction with 75 µM IPTG, whereas the highest octanol yields were observed at induction with 500 µM IPTG (Fig 4.9B). Changing the concentration of IPTG used for induction had no effect on OD_{600} (Fig 4.9C).

An IPTG concentration of 75 µM was taken as the optimum for CvT4 expression, as this led to the highest decanol titre and yield.

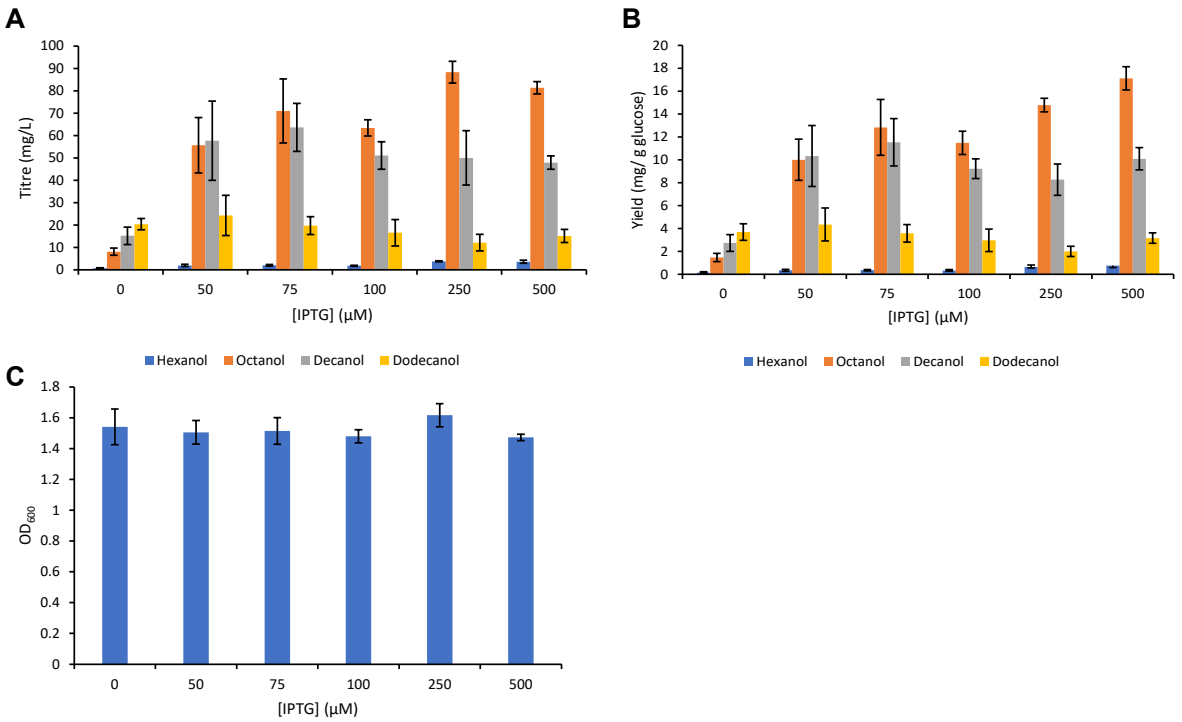


Figure 4.9 Inducer optimisation of CvT4 expression *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with a variable amount of IPTG and 1% rhamnose. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.4.3 Inducer Optimisation of CvT5 expression

As IPTG concentration used for induction of CvT5 decreased, alcohol titres increased, peaking at 75 μM IPTG. Higher titres of octanol (118.0 mg/L) and decanol (81.5 mg/L) were produced following induction with 75 μM IPTG, compared to induction with 500 μM IPTG, however this increase was only significant for decanol ($p = 0.00747$) (Fig 4.10A). On the other hand, as IPTG concentration used for induction decreased, octanol yields also decreased. Decanol yields increased as IPTG concentration decreased, with very similar yields at 50 and 75 μM IPTG (13.2 and 13.1 mg/g glucose) (Fig 4.10B). As IPTG concentration decreased, OD_{600} initially increased before plateauing (Fig 4.10C).

As induction with 75 μM IPTG led to the highest octanol and decanol titres, and the second-highest decanol yield, 75 μM IPTG was selected as the optimum for CvT5 expression.

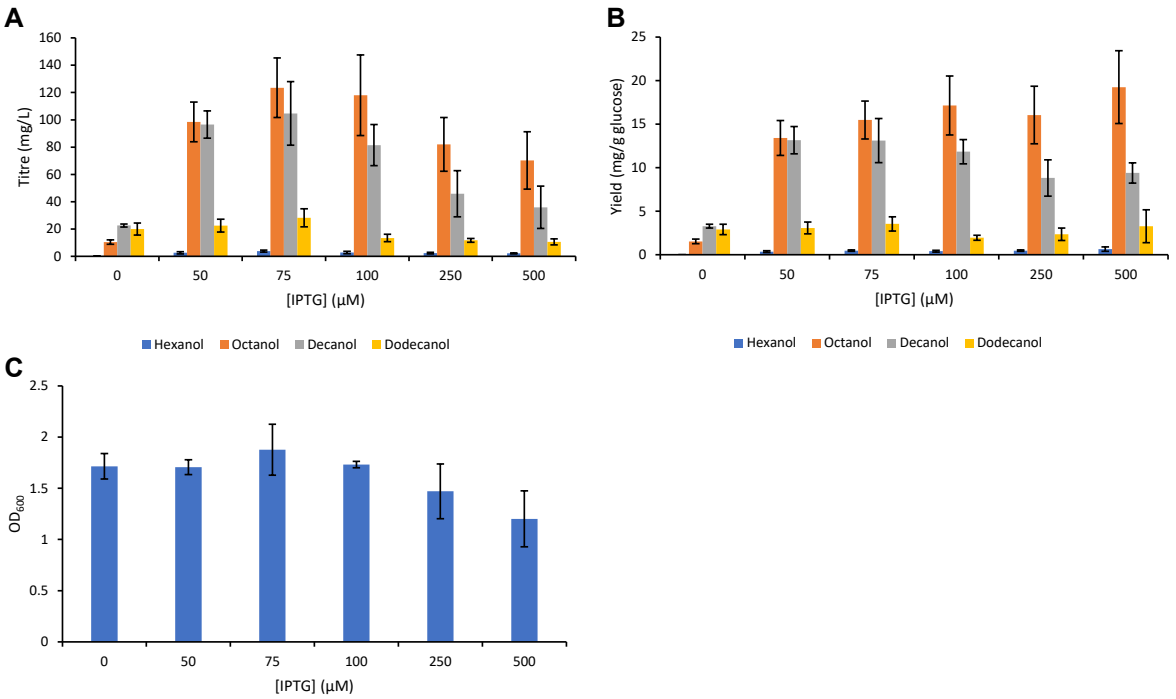


Figure 4.10 Inducer optimisation of CvT5 expression *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with a variable amount of IPTG and 1% rhamnose. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.4.4 Inducer Optimisation of CvT6 expression

As IPTG concentration used for induction of CvT6 decreased, alcohol titres increased, peaking at 50 μM IPTG. Significantly higher titres of octanol (99.3 mg/L, p = 0.000529) and decanol (62.0 mg/L, p = 0.0620) were produced, compared to induction with 500 μM IPTG (Fig 4.11A). Alcohol yields could not be calculated for induction with 250 and 500 μM IPTG, because higher glucose concentrations (>20 g/L) were measured in the cultures after 48 hrs growth than were present in the initial media. Once yields could be calculated, it was observed that as IPTG concentration used for induction decreased, alcohol yields decreased (Fig 4.11B). As IPTG concentration decreased, OD_{600} increased (Fig 4.11C).

As induction with 50 μM IPTG led to the highest alcohol titres, this was chosen as the optimum concentration for CvT6 expression.

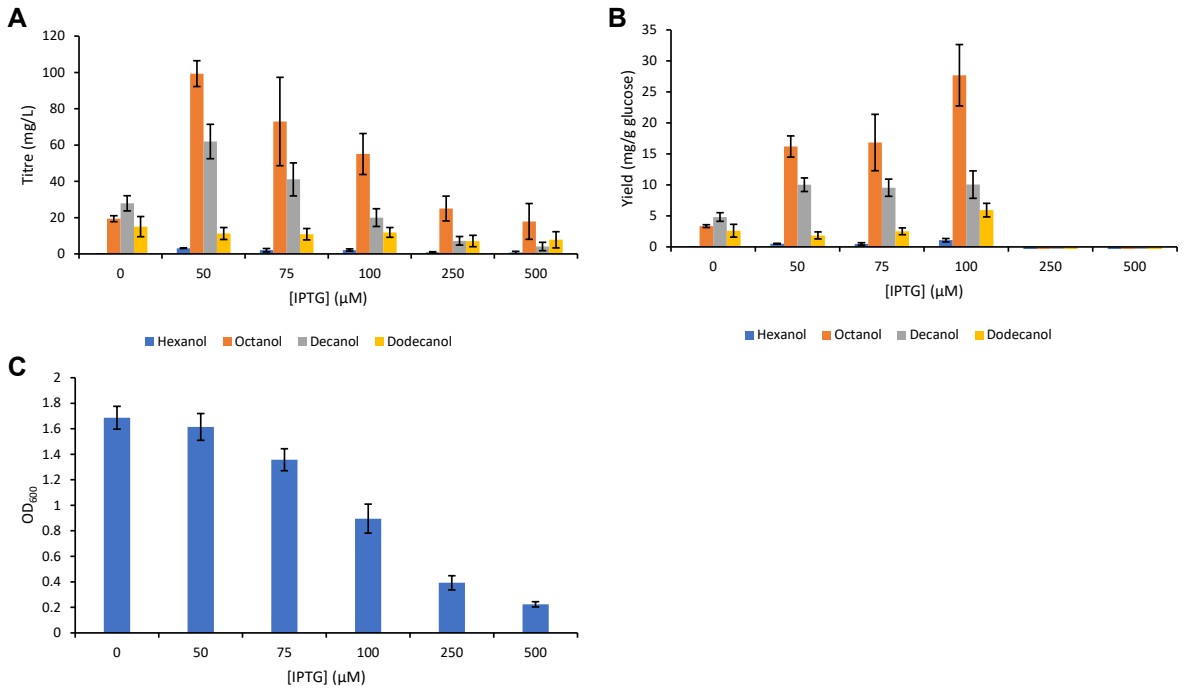


Figure 4.11 Inducer optimisation of CvT6 expression *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. Cells were induced with a variable amount of IPTG and 1% rhamnose. The overlay was sampled and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.4.5 Inducer Optimisation of CvT7 expression

As IPTG concentration used for induction of CvT6 decreased, octanol titres increased, peaking at 50 µM IPTG. Significantly higher octanol titres (86.8 mg/L, p = 0.00749) were obtained, compared to induction with 500 µM IPTG. The highest decanol titre (44.6 mg/L) was observed when no IPTG was used to induce the culture (Fig 4.12A). The highest octanol yields were observed following induction with 50 µM IPTG (47.8 mg/g glucose) and with 100 µM IPTG (62.6 mg/g glucose), with no significant difference between the two (p = 0.258). Altering IPTG concentration did not appear to have an effect upon the decanol yields (Fig 4.12A). Altering IPTG concentration did not have a significant effect upon OD_{600}, although increased OD_{600} was observed when no IPTG was used to induce expression of CvT7.

50 µM IPTG was chosen as the optimum concentration for induction as this led to the highest octanol titre and yield.

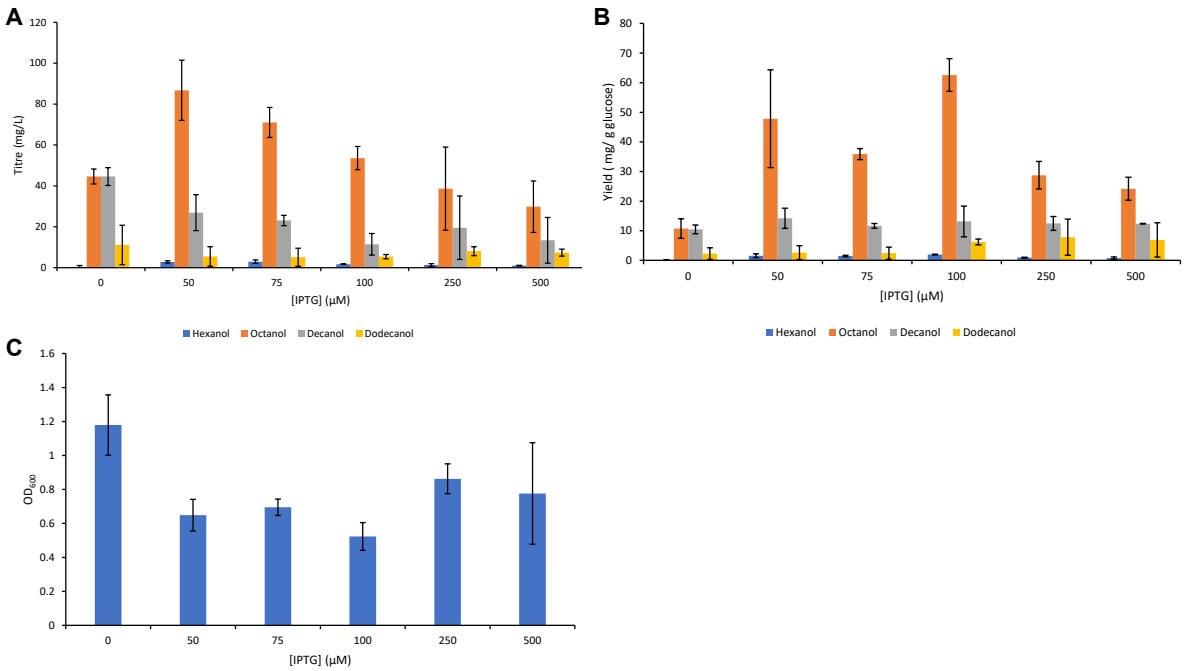


Figure 4.12 Inducer optimisation of CvT7 expression *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with a variable amount of IPTG and 1% rhamnose. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.4.6 Inducer Optimisation of CvT8 expression

As IPTG concentration used to induce CvT8 varied, no significant change in alcohol titres was observed (Fig 4.13A). The same was observed for alcohol yields (Fig 4.13B) and OD_{600} (Fig 4.13C). As there seemed to be no optimal concentration for CvT8 expression, 500 μM IPTG was chosen to induce CvT8 in future experiments.

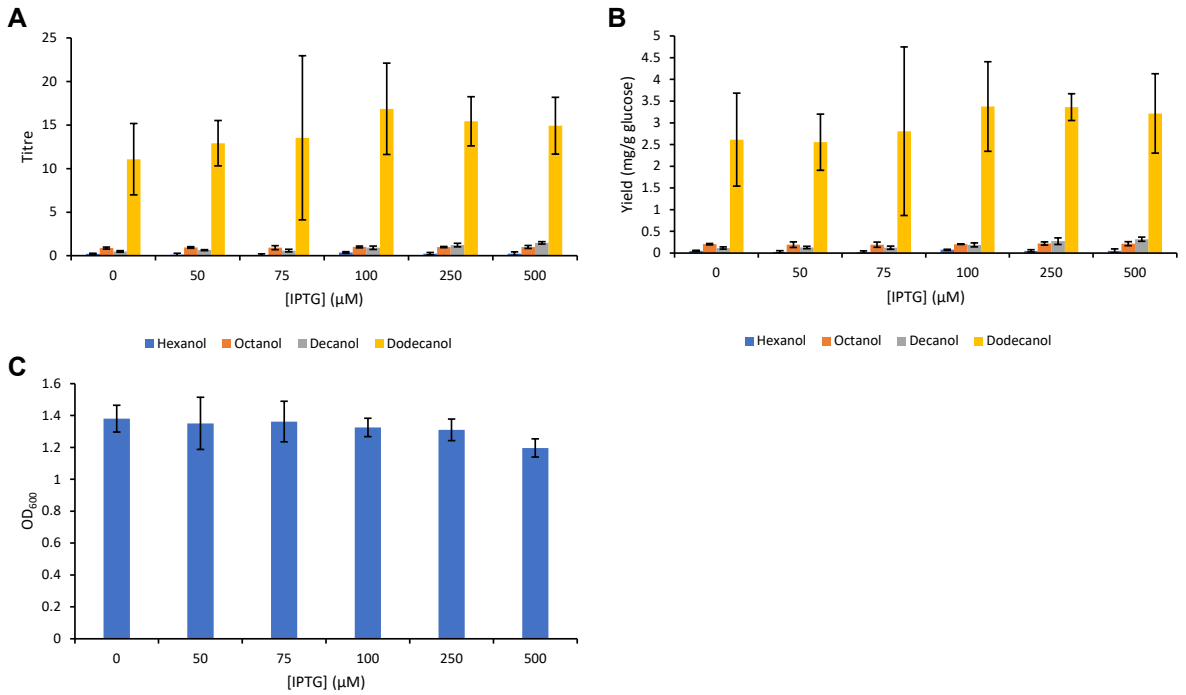


Figure 4.13 Inducer optimisation of CvT8 expression *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with a variable amount of IPTG and 1% rhamnose. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.4.7 Comparison of optimally expressed CvFatB1 truncations

After the optimum IPTG concentrations for each truncation had been determined (summarised in Table 4.3), the CvT3-8 were again screened, with each truncation expressed using its individual optimum IPTG concentration. This allowed the direct comparison of activities and specificities between the truncations.

CvFatB1 Truncation	Optimum [IPTG] (μM)
CvT4	250
CvT4	75
CvT5	75
CvT6	50
CvT7	50
CvT8	500

Table 4.3 Optimum IPTG concentrations for CvFatB1 truncation expression

Expression of CvT5 led to the highest alcohol titres, with octanol and decanol titres of 101.5 mg/L and 81.8 mg/L. However, this was not significantly higher than the octanol titres when CvT6 was expressed (85.3 g/L. p = 0.220), or the decanol titres when CvT4 was expressed (76.7 mg/L, p = 0.566). Overall, alcohol titres increased with truncation before peaking at CvT5, and then decreasing with further truncation, suggesting that the optimal site for truncation is around CvT5 (Fig 4.14A).

Truncations CvT3 and CvT8 had the lowest alcohol yields, and there were no significant differences in alcohol yields between truncations CvT4, CvT5, CvT6 and CvT7 (Fig 4.14B).

Even with inducer optimisation, CvT7 showed reduced growth relative to the other truncations, which showed similar growth to each other (Fig 4.14C).

As the thioesterase became more truncated, C10 specificity decreased, with CvT3 being the most C10 specific (44.9%), and CvT8 being the least specific (14.5%).

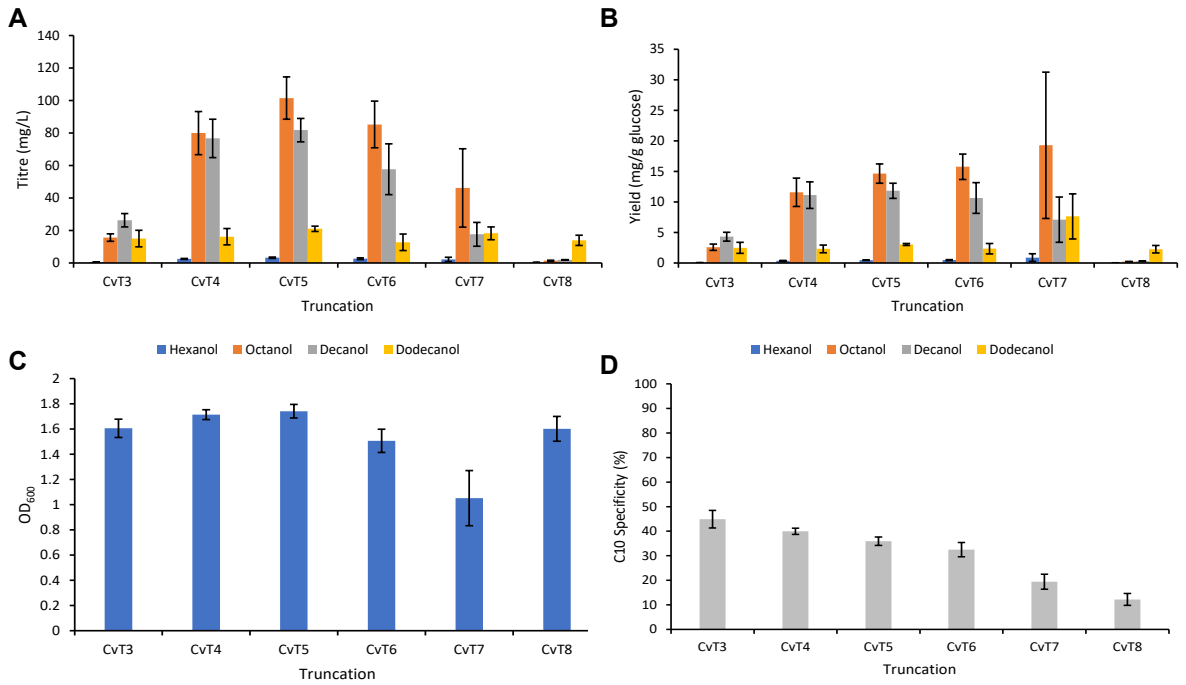


Figure 4.14 Optimised expression of CvFatB1 truncations *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis, expressing different truncations of CvFatB1 A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with a variable amount of IPTG and 1% rhamnose. D) C10 specificity, calculated as the percentage of total moles of fatty alcohol that were decanol. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.5 Overexpression of ACP

It was observed that the truncations had different activities and specificities, and that the same truncation had a different activity and specificity when expressed in *E. coli* and *Synechocystis*. It was theorised that these differences were due to interactions between the truncations and the ACPs, with different truncations interacting differently with the same ACP, and the ACPs from *E. coli* and *Synechocystis* Sp. PCC 6803 interacting differently to each other with the same truncation.

It had been previously reported that engineering a heterologous thioesterase to optimise its interaction with *E. coli* ACP led to a three-fold increase in fatty acid titres[410], and it was also reported that a mutation that decreased thioesterase interaction with ACP increased C10 specificity. It was theorised that this decreased interaction allowed for more decanoyl-ACP production, whereas before fatty acid biosynthesis would have been terminated at octanoyl-ACP by the thioesterase, therefore increasing C10 specificity[129]. These observations supported the theory that differing interactions between a thioesterase and an ACP influenced product biosynthesis.

ACP is an essential co-factor during fatty acid biosynthesis, to which the acyl chain is attached by a thioester bond during the cyclical fatty acid elongation reactions[411]. Bacteria only have one ACP gene while plants have multiple isoforms[412]. ACP overexpression has been observed to lead to altered fatty acid composition in plants[413,414] and *E. coli*[412,415], and thioesterases have been shown to have different specificities for different ACPs[133,416]. ACP overexpression could be therefore used in combination with a truncation to specifically produce decanoic acid.

It has been reported that knock-out of ACP from the genome both can[417,418] and cannot[415] be complemented by expressing ACP on a plasmid. Due to these conflicting reports, it was decided not to attempt knock-out of ACP from the genome of *E. coli* when overexpressing a heterologous ACP.

There is no ACP sequence available for *Cuphea viscosissima*, however the closely related *Cuphea lanceolata* is known to have two ACP isoforms[419,420], and of these ACP2 led to a higher fatty acid synthesis than ACP1 *in vitro* and increased the proportion of medium chain length products[421]. It was therefore decided to screen the overexpression of *Synechocystis* Sp. PCC 6803 ACP (6803ACP) and *Cuphea lanceolata* ACP2 (ClACP2) in combination with the different CvFatB1 truncations to investigate if this resulted in increased C10 specificity. Overexpression of *E. coli* ACP was also investigated as a control. To overexpress the

thioesterases, the plasmid used to express Sfp and CAR (Fig 4.7B) was modified to include ACP in the synthetic operon (Fig 4.15).

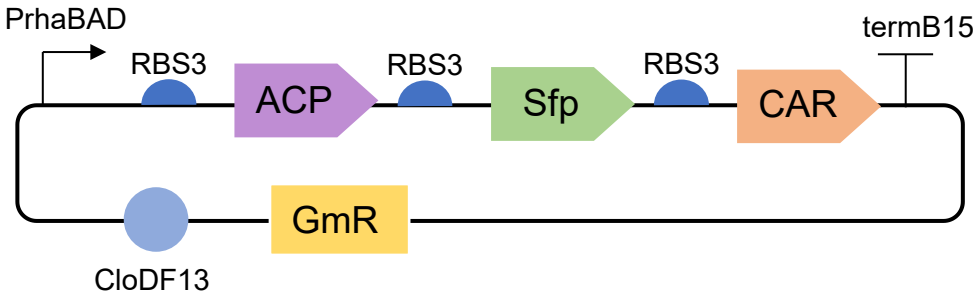


Figure 4.15 Plasmid design for ACP expression *A pCDF plasmid with gentamycin resistance cassette was used. ACP, Sfp and CAR were expressed in a single operon using PrhaBAD and RBS3.*

4.2.5.1 ACP expression in combination with CvT3

Additional overexpression of EcACP and expression of 6803ACP led to an increase in alcohol titres compared to expression of only CvT3, Sfp and CAR. The greatest improvement was seen when expressing 6803ACP, with significantly higher octanol (19.2 mg/L, $p = 0.0185$) and higher decanol (27.7 mg/L) titres relative to no ACP overexpression. On the other hand, expression of ClACP2 led to no significant change in alcohol titres (Fig 4.16A). A similar pattern was seen for alcohol yields (Fig 4.16B). Overexpression of ACP led to small but not significant decreases in OD_{600} (Fig 4.16C). There were small and not significant increases in C10 specificity when overexpressing an ACP (Fig 4.16D).

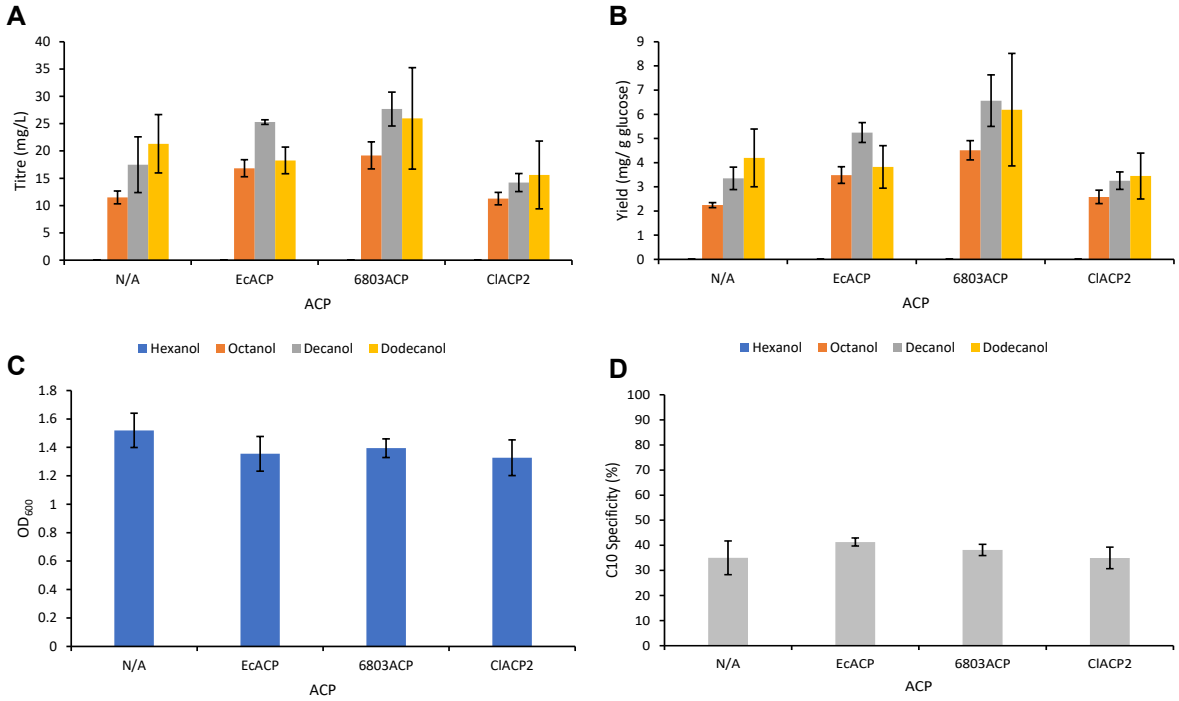


Figure 4.16 ACP expression in combination with CvT3 *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis, expressing different ACPs. N/A: No ACP A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with 250 μM IPTG and 1% rhamnose. D) C10 specificity, calculated as the percentage of total moles of fatty alcohol that were decanol. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.5.2 ACP expression in combination with CvT4

Additional expression of every ACP led to an increase in alcohol titres when combined with CvT4. The largest increases were seen with overexpression of EcACP, with significant increases in octanol (91.7 mg/L, p = 0.0175) and decanol (57.1 mg/L, p =0.00354) titres (Fig 4.17A). A similar pattern was seen with alcohol yield, however, increases in yield were smaller and not significant (Fig 4.17B). Again, additional expression of an ACP led to small but not significant decreases in OD_{600} (Fig 4.17C). There were also no significant changes in C10 specificity after expressing an ACP (Fig 4.17D).

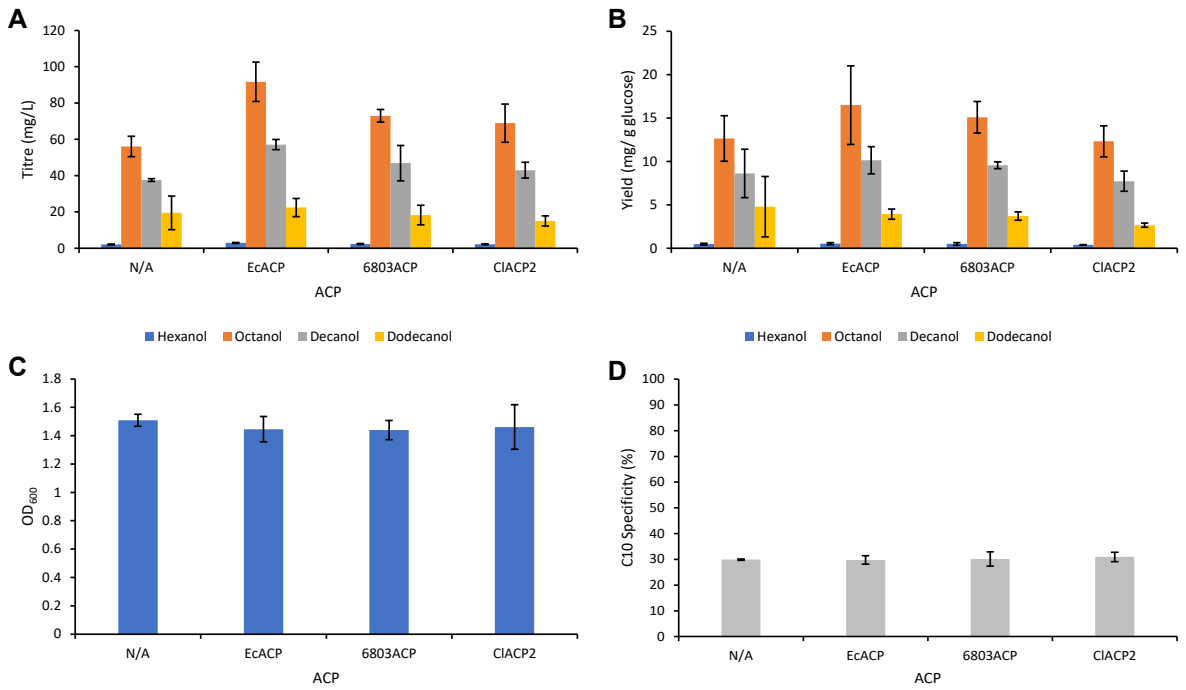


Figure 4.17 ACP expression in combination with CvT4 *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis, expressing different ACPs. N/A: No ACP A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with 75 μM IPTG and 1% rhamnose. D) C10 specificity, calculated as the percentage of total moles of fatty alcohol that were decanol. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.5.3 ACP expression in combination with CvT5

Additional overexpression of EcACP or expression 6803ACP in combination with CvT5 led to an increase in alcohol titres. Overexpression of EcACP led to the highest octanol (133.3 mg/L) and decanol (66.6 mg/L) titres, however only the increase in octanol titre was significant ($p = 0.0227$). Overexpression of ClACP2 had no effect on titres (Fig 4.18A). A similar pattern was observed for alcohol yields (Fig 4.18B). Again, addition of ACP overexpression led to small but not significant decreases in OD_{600} (Fig 4.18C). There were also no significant changes in C10 specificity after overexpressing an ACP (Fig 4.18D).

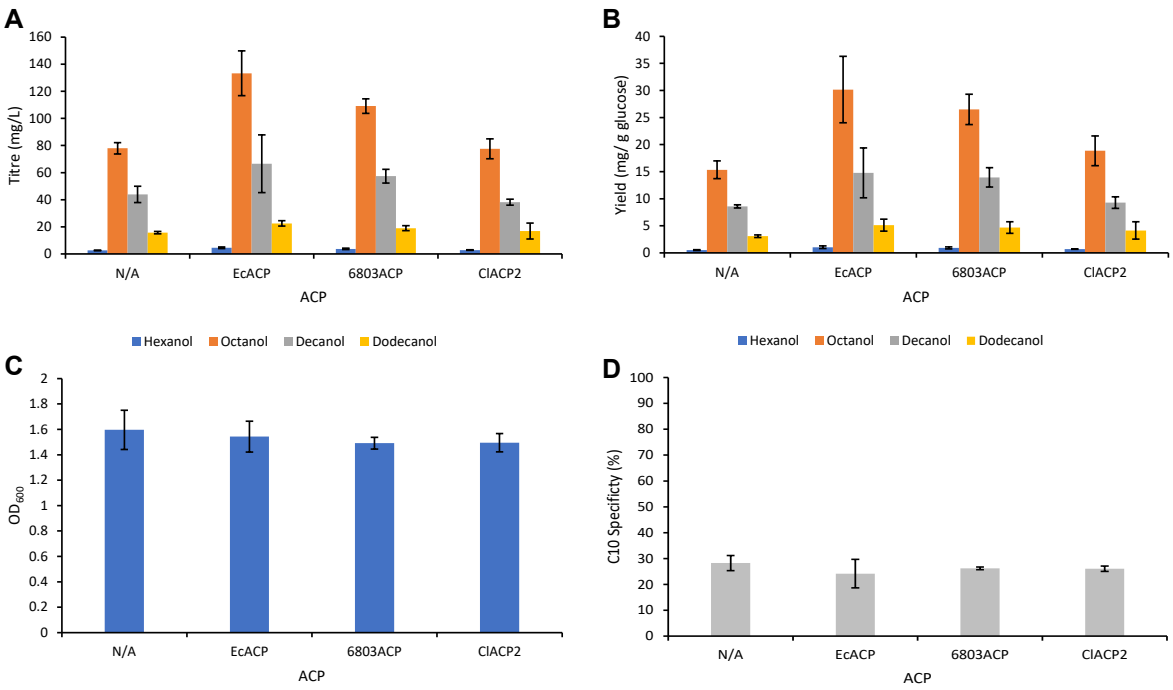


Figure 4.18 ACP expression in combination with CvT5 *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis, expressing different ACPs. N/A: No ACP A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with 75 μM IPTG and 1% rhamnose. D) C10 specificity, calculated as the percentage of total moles of fatty alcohol that were decanol. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.5.4 ACP expression in combination with CvT6

Additional overexpression of EcACP in combination with CvT6 led to a significant ~2 fold increase in octanol (197.3 mg/L, p = 0.0281) and decanol (150.1 mg/L, p = 0.000767) titres. On the other hand, expression of 6803ACP had no effect on titres, and expression of ClACP2 led to decreases in octanol (63.9 mg/L) and decanol (45.7 mg/L) titres (Fig 4.19A). A similar pattern was observed for alcohol yields (Fig 4.19B). Addition of ACP expression led to no significant changes in OD_{600} (Fig 4.19C). There were also no significant changes in C10 specificity after overexpressing an ACP (Fig 4.19D).

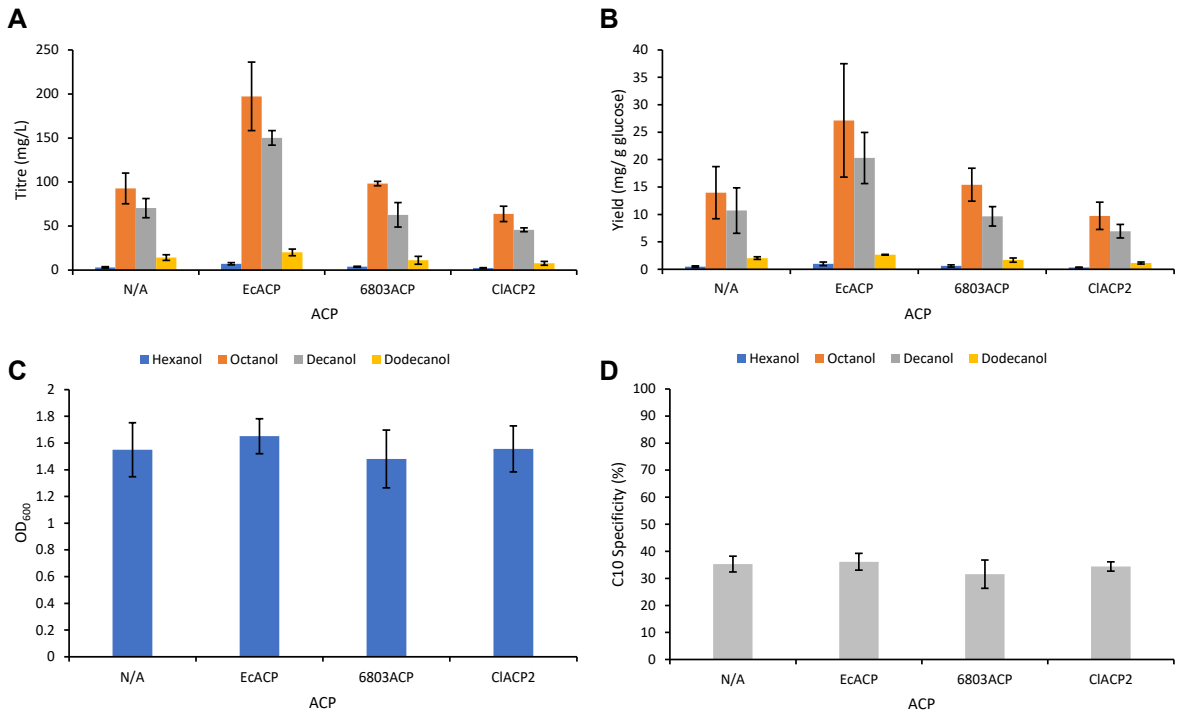


Figure 4.19 ACP expression in combination with CvT6 *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis, expressing different ACPs. N/A: No ACP A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with 50 μM IPTG and 1% rhamnose. D) C10 specificity, calculated as the percentage of total moles of fatty alcohol that were decanol. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.5.5 ACP expression in combination with CvT7 expression

Additional overexpression EcACP or expression of 6803ACP in combination with CvT7 led to increases in octanol titre, with the highest titre observed when overexpressing EcACP (98.2 mg/L). However, this was not a significant increase ($p = 0.0934$). No significant changes in decanol titre were observed. Expression of ClACP2 led to a decrease in octanol titres (83.3 mg/L) (Fig 4.20A). A similar outcome was also observed for alcohol yields (Fig 4.19B). ACP expression led to no significant changes in OD_{600} (Fig 4.20C). Overexpression of EcACP and expression 6803ACP led to small decreases in C10 specificity (17.8% and 17.5%, respectively) (Fig 4.20D).

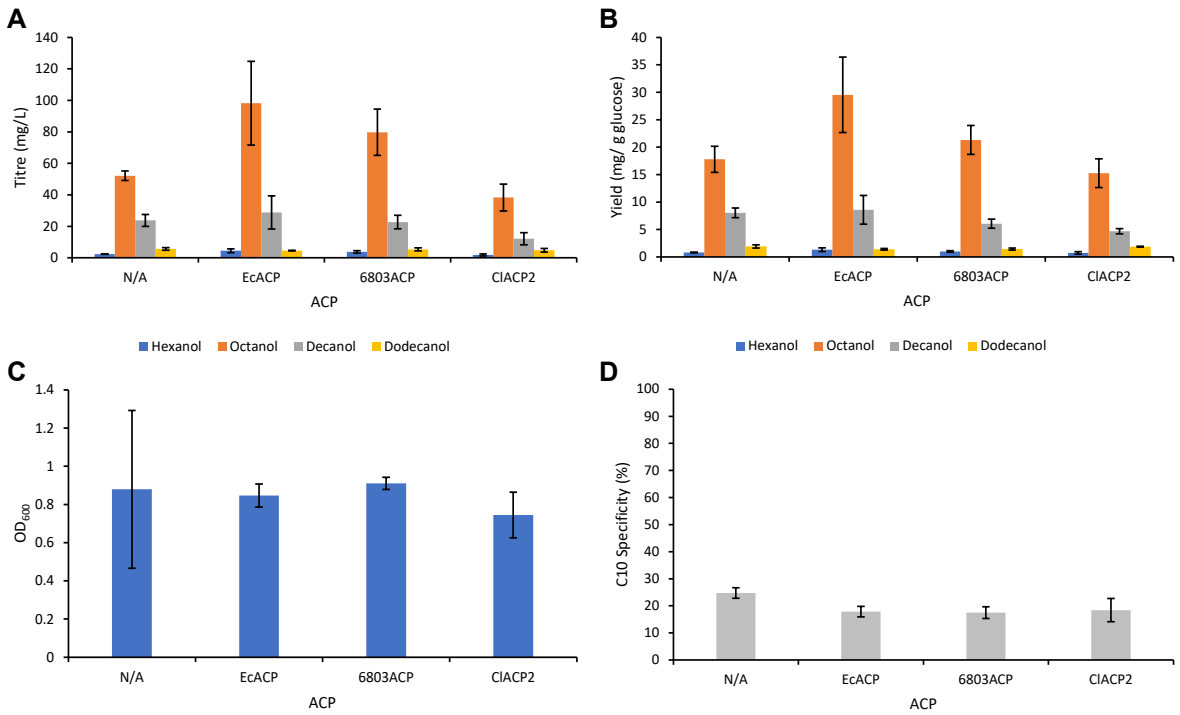


Figure 4.20 ACP expression in combination with CvT7 *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis, expressing different ACPs. N/A: No ACP A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with 50 μM IPTG and 1% rhamnose. D) C10 specificity, calculated as the percentage of total moles of fatty alcohol that were decanol. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.5.6 ACP expression in combination with CvT8 expression

Additional expression of every ACP in combination with CvT8 expression led to a decrease in dodecanol titres. When EcACP and 6803ACP were expressed, small increases in decanol titres were observed (Fig 4.21A). A similar pattern was observed for alcohol yields (Fig 4.21B). ACP expression again led to small decreases in OD_{600} (Fig 4.21C). Significant increases in C10 specificity were observed when overexpressing EcACP (17.0%, p = 0.0262) and 6803ACP (20.2%, p = 5.44e-5) (Fig 4.21D).

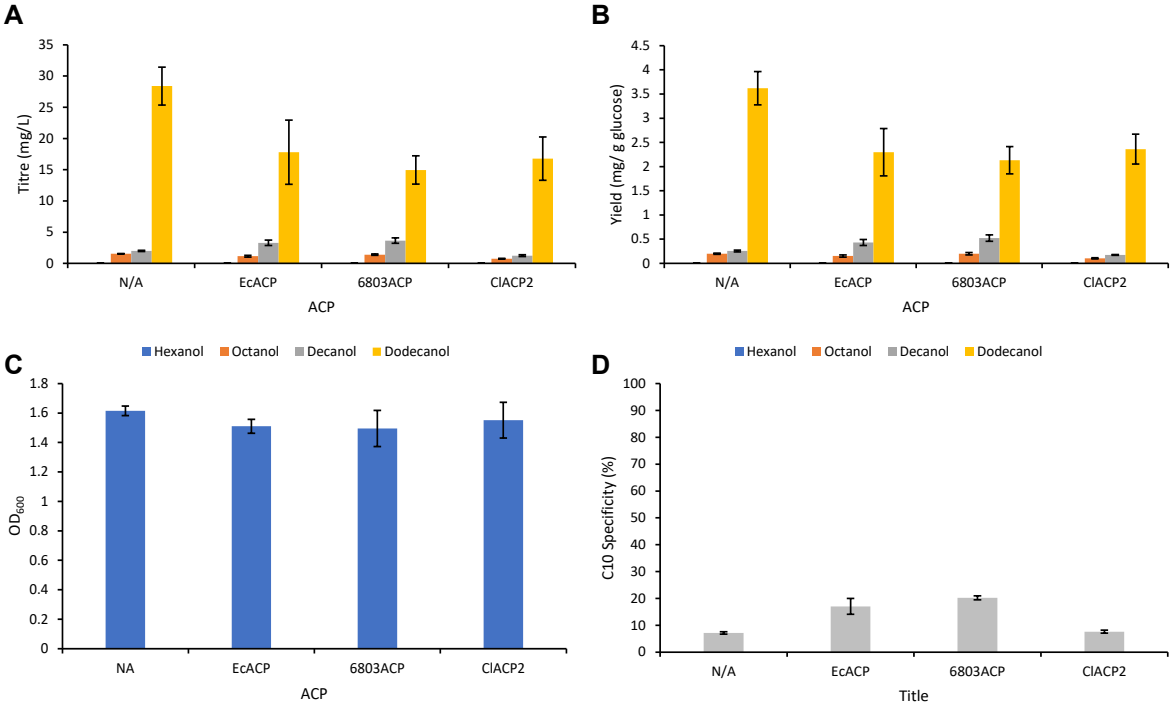


Figure 4.21 ACP expression in combination with CvT8 *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis, expressing different ACPs. N/A: No ACP A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with 500 μM IPTG and 1% rhamnose. D) C10 specificity, calculated as the percentage of total moles of fatty alcohol that were decanol. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.6 Site-directed mutagenesis of CvT5

It was decided to engineer the truncation CvT5 to be more C10 specific, as expression of CvT5 led to the highest alcohol titres in Section 4.2.4.7. A model of CvT5's structure was generated using the Phyre2 server[405] and visualised using PyMol. The catalytic triad was identified, and upon examination of the protein structure, an internal cavity whose mouth was flanked by the catalytic triad was observed. This cavity was taken to be the substrate binding pocket. Upon inspection of the pocket, three large hydrophobic residues, M90, W122 and W140 were identified that lined the bottom of the pocket (Fig 4.22). These were mutated to glycine, with the aim of enlarging the substrate binding pocket, therefore shifting CvT5's specificity from being mixed C8 and C10 specific to being more C10 specific.

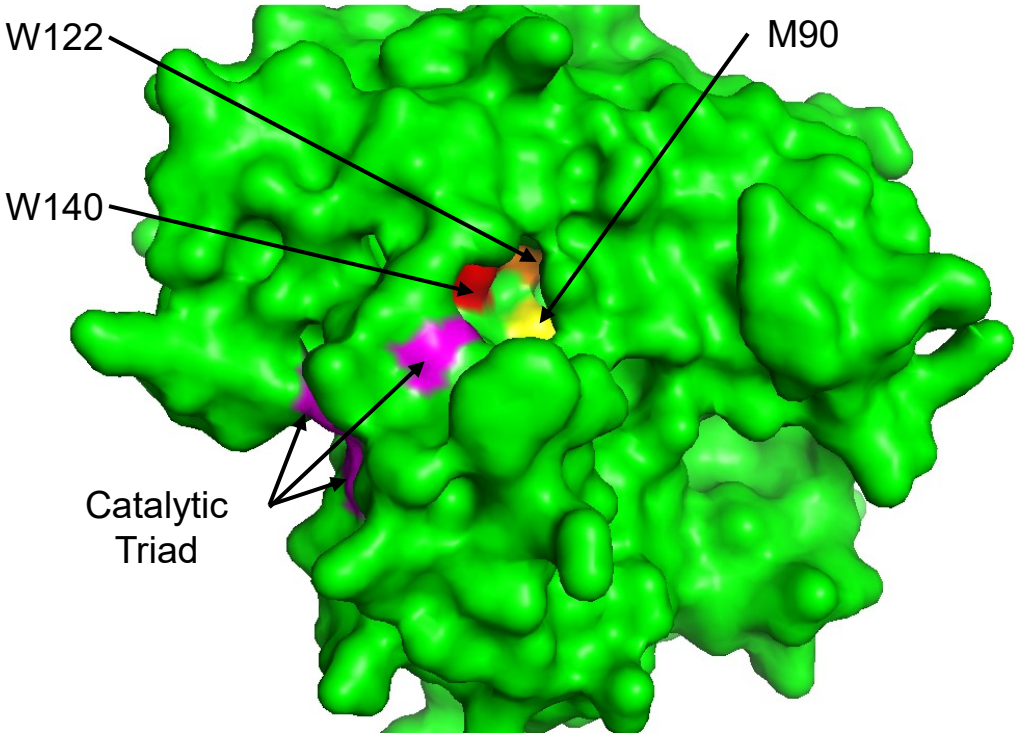


Figure 4.22 CvT5 predicted structure *The protein structure of CvT5 was predicted using Phyre2 and visualised using PyMol. The catalytic triad was identified (purple). Three large hydrophobic amino acid residues lining the bottom of the substrate binding pocket were identified; M90 (yellow), W122 (orange) and W140 (red).*

4.2.7 Screening CvT5 mutants

CvT5 mutants were screened by incorporating the mutants into the alcohol biosynthesis pathway, in combination with EcACP overexpression. It has previously been reported that engineering thioesterases to alter the specificity results in an activity loss[129,130], and so inducer optimisation was performed for each mutant in case the optimum IPTG concentration had changed.

4.2.7.1 Screening and inducer optimisation of CvT5 M90G

The M90G mutation led a severe activity loss, with significantly lower alcohol titres (Fig 4.23A) and yields (Fig 4.23B) observed. CvT5 M90G also showed slightly reduced growth (Fig 4.24C). IPTG concentration used for induction appeared to have an effect on specificity, with higher IPTG concentrations having higher C10 specificities. The highest C10 specificity was observed when induced with 500 µM IPTG (45.4%), which was a significant increase ($p = 0.0326$) (Fig 4.23D).

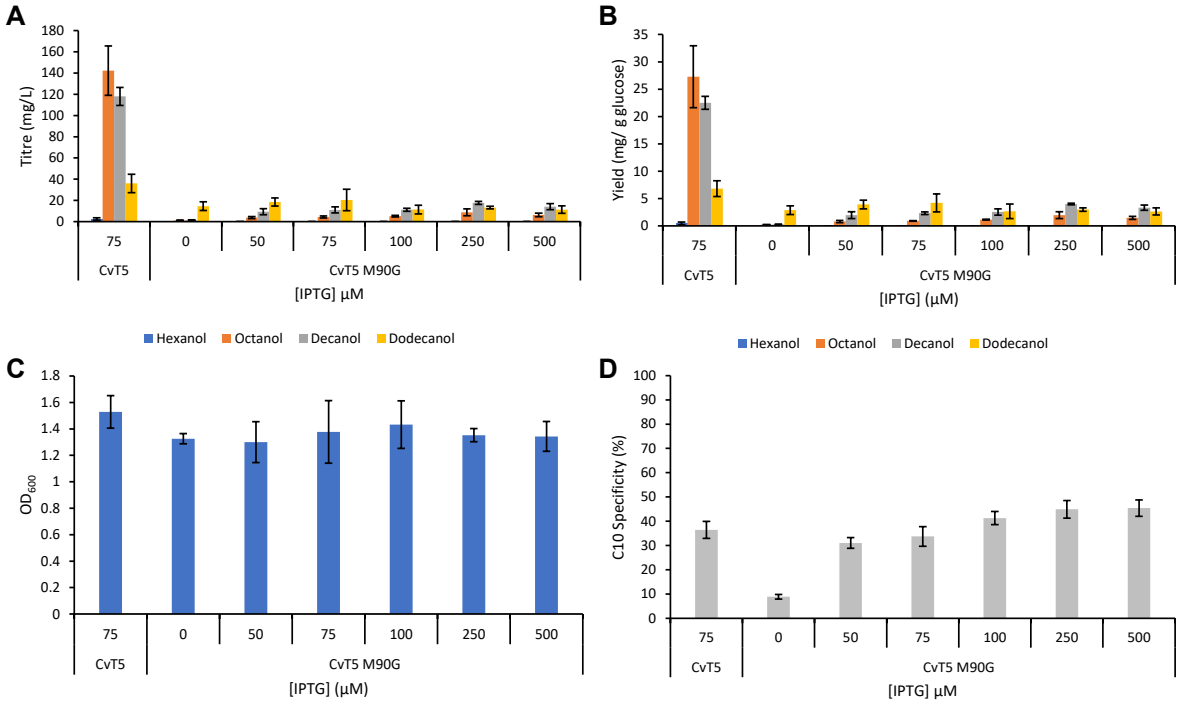


Figure 4.23 Screening and inducer optimisation of CvT5 M90G *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis, expressing either CvT5 or CvT5 M90G and EcACP, Sfp and CAR. A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with a variable amount of IPTG and 1% rhamnose. D) C10 specificity, calculated as the percentage of total moles of fatty alcohol that were decanol. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.7.2 Screening and inducer optimisation of CvT5 W122G

The highest alcohol titres were observed when CvT5 W122G was induced with 250 µM IPTG. Relative to CvT5 there were significant decreases in octanol titre (9.8 mg/L, p = 0.0103), and decanol titres were also reduced (66.4 mg/L). Moreover, no hexanol production was observed, and there was also a decrease in dodecanol titres (Fig 4.24A). A similar pattern was also observed for the yield (Fig 4.24B). There were no significant differences in OD_{600} between CvT5 and CvT5 W122G at any IPTG concentration (Fig 4.24C). CvT5 W122G had significantly increased C10 specificity compared to CvT5, with the highest specificity observed when induced with either 250 µM IPTG (63.7%) or 500 µM IPTG (64.5%), with no significant difference between the two (p = 0.8356) (Fig 4.24D).

250 µM IPTG was selected as the optimum concentration for induction of CvT5 W122G, as this resulted in the highest alcohol titres and C10 specificity.

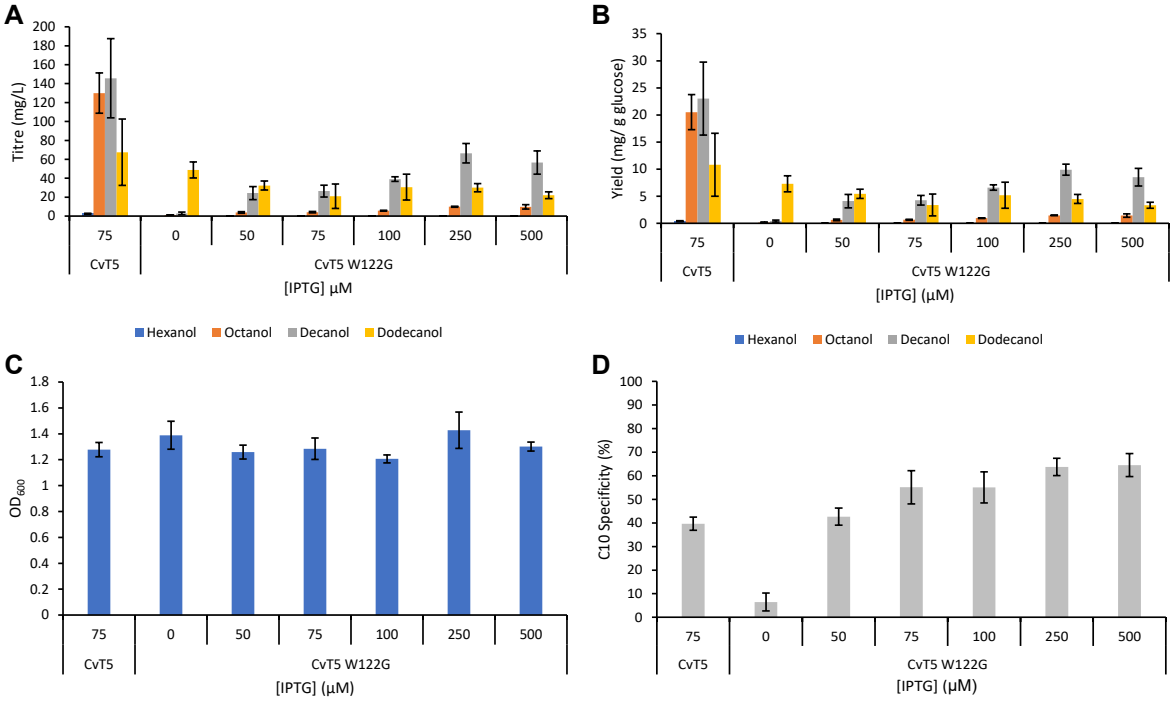


Figure 4.24 Screening and inducer optimisation of CvT5 W122G *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis, expressing either CvT5 or CvT5 W122G and EcACP, Sfp and CAR. A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with a variable amount of IPTG and 1% rhamnose. D) C10 specificity, calculated as the percentage of total moles of fatty alcohol that were decanol. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.7.3 Screening and inducer optimisation of CvT5 W140G

The W140G mutation led to a severe activity loss, with significantly lower alcohol titres (Fig 4.25A) and yields (Fig 4.25B) observed. CvT5 W140G also showed slightly increased growth (Fig 4.24C), and reduced C10 specificity (Fig 4.25D).

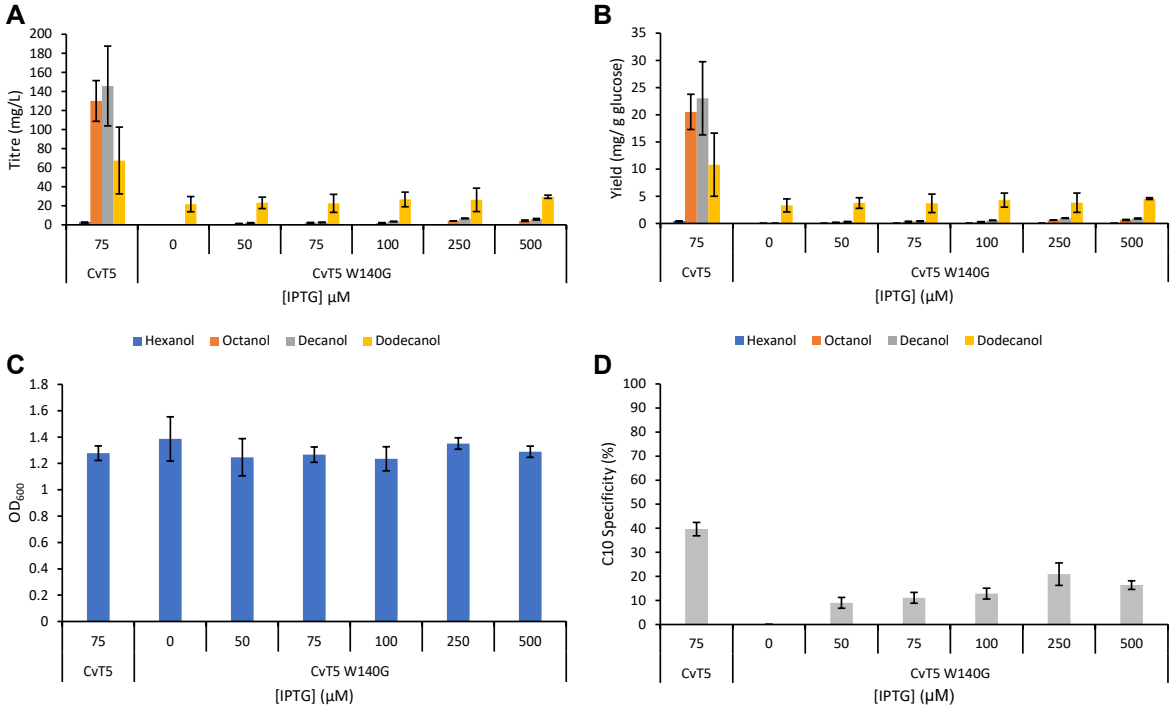


Figure 4.25 Screening and inducer optimisation of CvT5 W140G *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis, expressing either CvT5 or CvT5 W140G and EcACP, Sfp and CAR. A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with a variable amount of IPTG and 1% rhamnose. D) C10 specificity, calculated as the percentage of total moles of fatty alcohol that were decanol. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.2.8 Optimising the mutation at the CvT5 W122 site

CvT5 W122G showed an increase in C10 specificity relative to CvT5, however this came at a cost to activity. It was theorised that mutagenesis of W122 to a differently sized hydrophobic amino acid residue than glycine could produce a mutant of CvT5 that had increased C10 specificity without such a large activity loss. The following mutants; W122V, W122M and W122F were generated, and screened and inducer optimised as before.

The CvT5 W122V and CvT5 W122M showed similar activity profiles to CvT5 W122G, with reduced alcohol titres, and with decanol being the most produced alcohol. For both CvT5 W122V and CvT5 W122M the highest decanol titres were observed when induced with 500 µM IPTG. Expression of CvT5 W122F led to more decanol than octanol being produced. CvT5 W122F had low activity when no IPTG was used to induce expression, and other IPTG

concentration had little effect on overall alcohol titres. The highest decanol titres (222.5 mg/L) were observed when CvT5 W122F was induced with 50 µM IPTG, and the highest octanol titres (200.3 mg/L) were observed when CvT5 W122F was induced with 500 µM IPTG (Fig 4.26A).

A similar pattern was observed for the yields (Fig 4.26B). The CvT5 W122V, CvT5 W122M and CvT5 W122F mutants all showed similar growth to CvT5 and CvT5 W122G at the differing IPTG concentrations (Fig 4.26C).

The highest C10 specificity (77.1%) was observed when CvT5 W122G was induced with 250 µM IPTG. There was no significant difference in C10 specificity between the CvT5 W122G, CvT5 W122V and CvT5 W122F mutants. The CvT5 W122F mutant showed an intermediate C10 specificity between CvT5 (36.2%) and the CvT5 W122G, CvT5 W122V and CvT5 W122M mutants. Apart from when no IPTG was used to induce its expression, CvT5 W122F was ~45% C10 specific (Fig 4.26D).

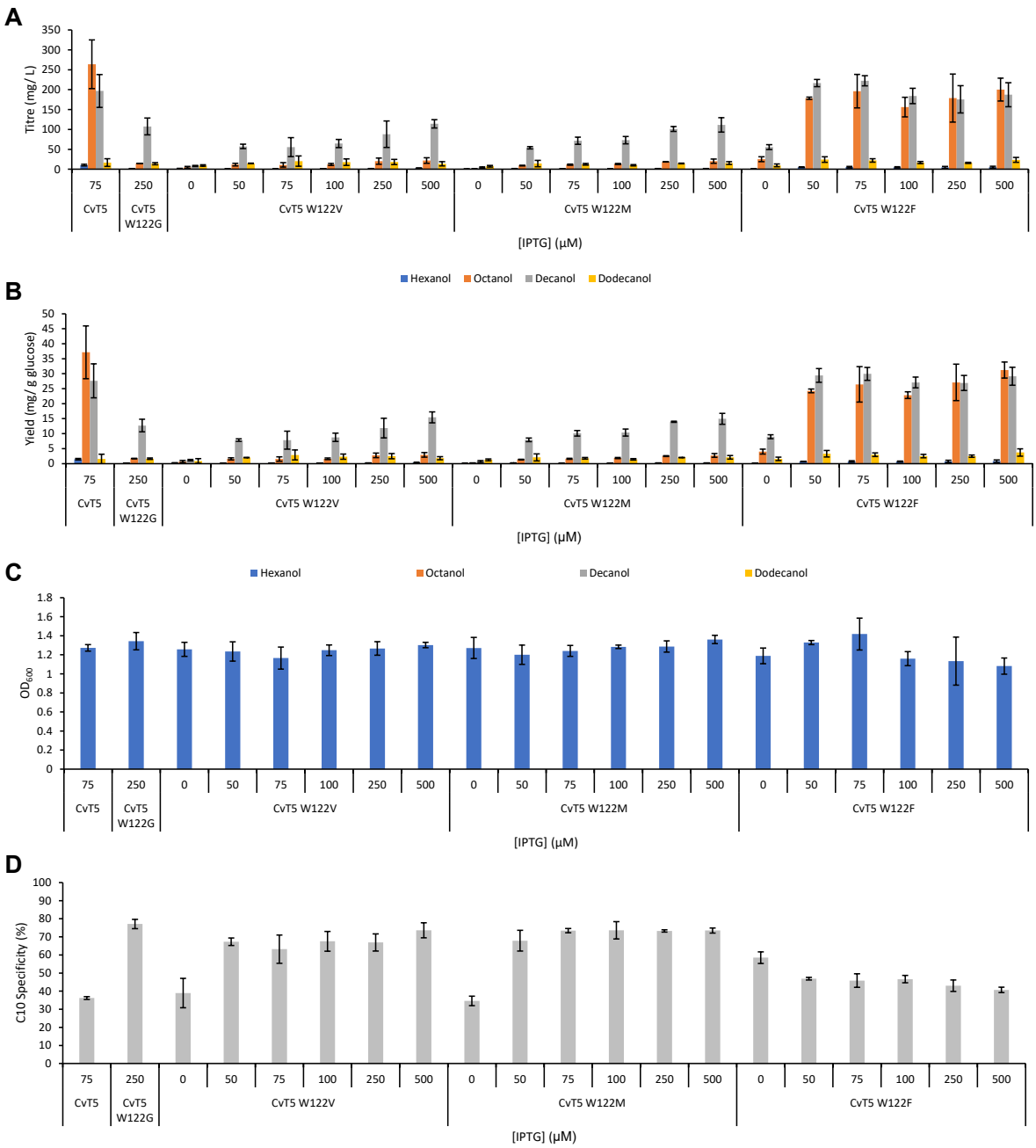


Figure 4.26 Screening and inducer optimisation of CvT5 W122 mutants *E. coli ΔfadD were co-transformed with two plasmids for alcohol biosynthesis, expressing a CvT5 or a CvT5 W122 mutant and EcACP, Sfp and CAR. A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media with 2% glucose and a 10% v/v isopropyl myristate overlay. The overlay was sampled and analysed by GC-MS. Cells were induced with a variable amount of IPTG and 1% rhamnose. D) C10 specificity, calculated as the percentage of total moles of fatty alcohol that were decanol. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

4.3 Discussion

4.3.1 The effect of truncation upon FatB acyl-ACP thioesterase activity in *E. coli*

In this work a much larger set of truncations was investigated during the truncation assay of CvFatB1 than had previously been examined in the literature when truncating other thioesterases.

From the truncation CvT8 onwards, there was a clear loss of activity with low alcohol titres observed, of which dodecanol followed by octanol were the major products. *E. coli* has an endogenous acyl-ACP thioesterase, TesA, which has a wide chain length substrate specificity, and has low C8 and C12 activity[422]. TesA may have therefore been responsible for production of low amounts of medium chain length free fatty acids, which were then converted to the corresponding alcohols by CAR and ADH, resulting in the biosynthesis of some alcohols even when the CvFatB1 truncations were not active. Truncation beyond the site of truncation used in CvT7 was likely detrimental to the thioesterase's activity as both the N-terminal chloroplast localisation peptide and part of the thioesterase proper were removed. Truncation of a FatB thioesterase that resulted in an activity loss had not previously been reported, and knowledge of this cut-off point will be useful for future work when designing and screening FatB thioesterase truncations.

Truncations CvT4-7 had the highest alcohol yields out of all the assayed truncations and, following inducer optimisation, higher alcohol titres than truncations CvT3 and CvT8. This suggests that the optimal site for truncation of CvFatB1 is somewhere within the region between the truncations used to generate CvT4 and CvT7, likely around the truncation site of CvT5, which had the highest alcohol titres. Truncations CvT4-7 are all located within the N-terminal region of CvFatB1 Domain 1 from a thioesterase domain swapping study, while CvT8 is located in the C-terminal region of Domain 1[134]. This suggests that the 10 residues in Domain 1 between CvT7 and CvT8 are key to thioesterase activity. To my knowledge, the optimal truncation site of a FatB thioesterase has not previously been so precisely located, and it was not known that the residues between CvT7 and CvT8 played such a key, albeit still to be determined role, in thioesterase activity.

When the thioesterase UaFatB1 was truncated in the same manner, truncations UaT1-4 had the highest alcohol titres. This suggests that the optimal site for truncation is not universal among FatB thioesterases, with each thioesterase having its own optimal truncation site. This means that when expressing a new thioesterase, screening a range of truncations to identify the optimum truncation site would a useful step. Truncations UaT8 and UaT9 also had very low alcohol titres, supporting the theory that truncation beyond this point is detrimental to

thioesterase activity. The highest decanol titres were observed when expressing the wild-type thioesterase sequence. The wild-type thioesterase was not expected to be active, due to the presence of the Chloroplast Localisation Peptide, which would render the thioesterase insoluble and therefore inactive upon expression in bacteria[126]. However, the "wild-type" sequence was already partially truncated, explaining why activity could be observed.

The activity and specificity profiles of CvFatB1 and UaFatB1 did not match what had previously been reported, with CvFatB1 being relatively more active and UaFatB1 being relatively more C10 specific than previously reported[104]. This highlights the importance of verifying the reported activity of genes from the literature before incorporating them into synthetic metabolic pathways.

4.3.2 The difference in thioesterase C10 specificity between *E. coli* and *Synechocystis*

When the truncations CVT3-8 were expressed in *Synechocystis* a significant difference in specificity was observed, with more decanol than octanol being produced. C10 specificity reached 71.9% when CvT4 was expressed. This was the highest C10 specificity that had been observed for any acyl-ACP thioesterase and exceeded the target of 66.6% C10 specificity set at the start of this Chapter.

This increased C10 specificity when the thioesterase was heterologously expressed in *Synechocystis* matched previous observations from the Jones lab where a thioesterase, *Cuphea hookerina* FatB2, that had mixed C8 and C10 specificity upon expression in *Synechocystis*[128] was then found in unpublished work to be C8 specific when expressed in *E. coli*.

One potential explanation was that ACP and thioesterase interactions were modulating the specificity. It had previously been reported that expression of a heterologous ACP led to altered fatty acid composition in plants and *E. coli*[127,412,414,415], supporting the theory that *Synechocystis* ACP could be responsible for this altered C10 specificity. This was investigated, and the results are discussed in the following section.

An alternative explanation could be that there are differences in fatty acid biosynthetic cycle between *Synechocystis* and *E. coli* that leads to a relatively larger amount of decanoyl-ACP being present in *Synechocystis*. This increased substrate availability would therefore lead to more decanoic acid and therefore decanol being biosynthesised, leading to higher apparent C10 specificities.

Regardless of the cause of the higher observed C10 specificity in *Synechocystis*, this ability to specifically produce decanoic acid is useful when trying to biosynthesise decanoic acid derived products such as γ- and δ-decalactone. This could give *Synechocystis* an advantage

over other hosts such as *E. coli* for the biosynthesis of decanoic acid derived products, creating a competitive niche within which *Synechocystis* can be developed as a chassis for sustainable bioproduction.

4.3.3 The effects of ACP expression upon alcohol biosynthesis

Heterologous ACP expression was investigated as a method to tailor thioesterase specificity. Expression of three ACPs was investigated. Firstly, expression of 6803ACP was investigated to examine if this would alter the thioesterases' C10 specificity in *E. coli* to resemble that in *Synechocystis.* Secondly, expression of ClACP2 was investigated to determine if it would have an effect on fatty acid and therefore alcohol biosynthesis. Finally, overexpression of the *E. coli* ACP (EcACP) was investigated as a control.

Expression of ClACP2 either had no effect on alcohol titres, or even in some cases even led to decreased alcohol titres. It has previously been reported that plant ACPs are less active than bacterial ACPs when used in fatty acid synthesis reactions *in vitro*[423]. ClACP2 therefore likely had low activity relative to the endogenous *E. coli* ACP, and thus did not have a significant effect on fatty acid biosynthesis. The decreased alcohol titres were sometimes observed, this could either be due to increased burden upon the cells from the expression of an additional heterologous protein, or ClACP2 may have acted as a competitive inhibitor, competing with endogenous acyl-ACP to bind to the thioesterase, and therefore reducing titres.

Expression of 6803ACP sometimes led to increases in alcohol titres and sometimes had no effect. The observed increases in titres suggest that 6803ACP was able to interact with the *E. coli* fatty acid biosynthesis enzymes, and that ACP availability may be a limiting factor for fatty acid biosynthesis. It has also been previously observed that expressing a cyanobacterial ACP in *E. coli* leads to increased fatty acid biosynthesis[415], matching this result. Expression of 6803ACP did not lead to any changes in C10 specificity. This could be that the theory that the different interactions between 6803ACP and the truncations modulated the specificity to be more C10 specific was incorrect, or that not enough of the 6803ACP was incorporated into the fatty acid biosynthesis cycle to have an effect. Either way, the results demonstrated that overexpression of 6803ACP on a plasmid could not be used to modulate thioesterase specificity.

Overexpression of EcACP consistently led to increased alcohol titres. This suggests that availability of ACP is a limiting factor in fatty acid biosynthesis, and that overexpression of EcACP can be applied to boost the production of fatty acid derived products in future works. To my knowledge, it has not previously been reported that overexpression of EcACP can be used to improve product biosynthesis in *E. coli*.

In future work it would be of interest to examine the effects of ACP overexpression in *Synechocystis*, as low product titres are a hinderance to the development of *Synechocystis* as a host for renewable bioproduction[12]. If overexpression of ACP could be used to increase titres of fatty acid derived products in *Synechocystis*, this would be highly desirable. Constructs for the expression of thioesterase truncations in combination with the different ACPs were prepared, and triparental conjugation to transform *Synechocystis* with these constructs was attempted. Despite repeated attempts at triparental conjugation, colony growth after transformation could not be observed for all the required strains, and so this experiment could not be performed. The triparental conjugation should therefore be attempted again in future work.

4.3.4 Engineering CvT5 for increased C10 specificity in *E. coli*

Structure guided mutagenesis of CvT5 led to the following three mutants being designed and tested, CvT5 M90G, CvT5 W122G and CvT5 W140G. Of these, CvT5 M90G and CvT5 W140G had activity losses, suggesting mutation at these sites was detrimental to the thioesterase activity. M90 and W140 line the substrate binding pocket, and their mutation to glycine could have destabilised the protein structure, leading to the observed activity loss.

On the other hand, CvT5 W122G had increased C10 specificity, which reached as high as 77.1% in Section 4.2.8, but at a cost to overall activity, with lower product titres. This C10 specificity exceeded the target of 66.6% set at the start of this chapter and was more specific than the then most C10 specific engineered thioesterase reported in the literature (~60%)[129]. After this work had been completed, an engineered version of *Cuphea lanceolata* FatB3 (ClFatB3) that incorporated a truncation and activity enhancing mutations was reported to have 70% C10 specificity, exceeding the target set at the beginning of this chapter. While the C10 specificity was lower than that of CvT5 W122G, this was achieved without the need for specificity-altering mutations[125]. The application of the engineering techniques used on CvT5 to the engineered ClFatB3 could therefore result in even more specific C10 acyl-ACP thioesterases in future work.

The CvT5 residues M90 and W122 are located in the C-terminus of CvFatB1 Domain 2, while W140 is located in the N-terminus of CvFatB1 Domain 3. The fact that mutagenesis at these sites had such large impacts on activity and specificity matched a prior report that Domains 1-4 are important in controlling thioesterase activity, although the fact that the W122G mutation resulted in a C10 specific thioesterase didn't match the reported conclusion that Domains 3 and 4 are the most important in controlling thioesterase activity[134].

Attempts at optimising the mutation at the W122 site to find a balance between increasing C10 specificity without compromising overall activity had limited success. The CvT5 W122V and CvT5 W122M mutants showed very similar outcomes to the CvT5 W122G mutant. On the other hand, the CvT5 W122F mutant showed increased C10 specificity (46.9%) without an activity loss. In future work it would be of interest to express CvT5 W122F in cyanobacteria. CvT5 had the highest alcohol titres when expressed in *Synechocystis*, however it's specificity of 63.1% was just short of the target. CvT5 W122F is therefore likely to exceed the target when expressed in *Synechocystis*, without compromising on alcohol titres.

4.4 Conclusions

In this Chapter, truncation of two thioesterases, CvFatB1 and UaFatB1 was assayed in *E. coli* and in *Synechocystis*. Truncation of CvFatB1 was systematically investigated to a greater extent than had previously been reported in the literature, and it was found that the optimal site for truncation was located within the N-terminus of CvFatB1 Domain 1. It was also found that truncation beyond a certain point abolished thioesterase activity, which had not previously been reported. The truncation CvT4 was found to be a C10 specific (71.9%) thioesterase when expressed in *Synechocystis*. At the time this was the highest C10 specificity reported for any acyl-ACP thioesterase.

ACP overexpression was investigated in *E. coli* as a tool to modulate thioesterase specificity. While this did not result in altered substrate specificity, it was found that overexpression of EcACP led to increased product titres, which had not been previously reported in the literature.

The truncation CvT5 was engineered for improved C10 specificity using the predicted protein structure, and one of the designed mutants, CvT5 W122G, had a high C10 specificity (77.1%). CvT5 W122G is the most C10 specific acyl-ACP thioesterase that has been reported.

Chapter 5

Engineering BM3 for the ω-5 and ω-6 Hydroxylation of Decanoic Acid

5.1 Aims

γ-Decalactone and δ-decalactone are commercially important high impact and value flavour and fragrance compounds[15,79,84,87]. The native pathway for their biosynthesis is unknown[87], and prior work in the literature has focused on the biotransformation of long chain fatty acids to medium chain length γ- and δ-lactones[81,83,85,88-95].

In this work a novel metabolic pathway for the biosynthesis of γ- and δ-decalactone has been proposed. In order to biosynthesise γ- and δ-decalactone the respective hydroxyfatty acid precursors, 4- and 5-hydroxydecanoic acid, first need to be produced. To produce 4- and 5-hydroxydecanoic acid, hydroxylation of decanoic acid is required at the ω-6 and ω-5 positions, respectively.

To do so, Cytochrome P450 BM3 will be used. BM3 does not natively hydroxylate decanoic acid[158], or natively hydroxylate fatty acids at these positions[166]. It is also unclear whether the 4- or 5-hydroxydecanoic acids will spontaneously condense to form the γ- and δ-decalactone, or whether sample treatment or a pathway extension is required.

The aims of this chapter were therefore to engineer BM3 to hydroxylate decanoic acid at the 4- and 5-positions and investigate the spontaneity of condensation of the hydroxyfatty acids to the corresponding decalactones.

5.2 Results

5.2.1 Decanoic acid feeding assay

BM3 natively hydroxylates C12-C20 fatty acids[158], however low activity has been reported on decanoic acid[164]. To confirm whether BM3 could hydroxylate a decanoic acid substrate, BM3's ability to consume a fed decanoic acid substrate was investigated. It was expected that if BM3 was active on decanoic acid, 7- ,8- and 9-hydroxydecanoic acid would be produced, however chemical standards were not available. Therefore, it was decided that consumption of the fed decanoic acid would be used as a proxy for BM3 activity. While not expected to be biosynthesised, standards for γ- and δ-decalactone were subjected to GC-MS analysis. BM3 was expressed in *Escherichia coli* using a plasmid (Fig 5.1E) and the *E. coli* cultures were fed with 2.5 mM decanoic acid. Cultures that expressed GFP instead of BM3 were used as a control, and also fed with 2.5 mM decanoic acid. As a second control, called Spike, cultures expressing GFP were fed with 2.5 mM decanoic acid, γ-decalactone and δ-decalactone. These controls allowed the product analysis technique to be tested, to ensure that the compounds could be extracted and to investigate if there was any product loss during the experiment. Along with the BM3 wild-type, two mutants generated in previous work by the Jones lab were also investigated. These were BM3 F87A[166], which had been reported to have altered regioselectivity, and BM3 L188Q[169], which had been reported to have increased activity upon decanoic acid.

It was observed that not all of the fed decanoic acid could be recovered by cell extraction. In the GFP strain 399.6 mg/L (2.32 mM) decanoic acid was measured, and in the Spike strain 330.4 mg/L (1.92 mM) decanoic acid, 277.6 mg/L (1.63 mM) δ-decalactone and 366.3 mg/L (2.15 mM) γ-decalactone was measured (Fig 5.1A). These data suggests that there was either product loss during the cultivation period, or that the product analysis technique was not able to extract all the decanoic acid and decalactones. After 48 hrs cultivation, a ring of white fatty material was observed on the flasks, suggesting some of the decanoic acid had precipitated out of solution (Fig 5.1D). It was also observed that the Spike strain (OD_{600} = 0.59) showed significantly reduced growth relative to the GFP strain (OD_{600} = 1.25, p = 0.0123), indicating that the decalactones were toxic at these concentrations (Fig 5.1C).

Expression of the BM3 wild type (BM3 WT) led to significantly lower decanoic acid titres (297.1 mg/L, p = 0.00334) compared to expression of GFP. Expression of the BM3 L188Q mutant led to a further significant decrease in decanoic acid titres (143.2 mg/L, p = 0.00868). This indicated that BM3 was able to consume the fed decanoic acid, and that the L188Q mutation did increase BM3's activity on decanoic acid (Fig 5.1A).

Feeding decanoic acid to BM3 was not expected to result in the biosynthesis of decalactones, however low titres of δ-decalactone (1.45 mg/L) and γ-decalactone (1.64 mg/L) were observed upon expression of the BM3 WT. Expression of BM3 F87A led to significantly higher δ-decalactone tires (2.55 mg/L, p = 0.0359) and an insignificant decrease in γ-decalactone titres (0.71 mg/L, p = 0.122), indicating that the F87A mutation did have an effect on hydroxylation regioselectivity (Fig 5.1B). Decanoic acid titres were significantly higher compared to the wild type (352.1 mg/L, p = 0.0243), indicating the F87A mutation led to decreased activity, as had been previously reported[166].

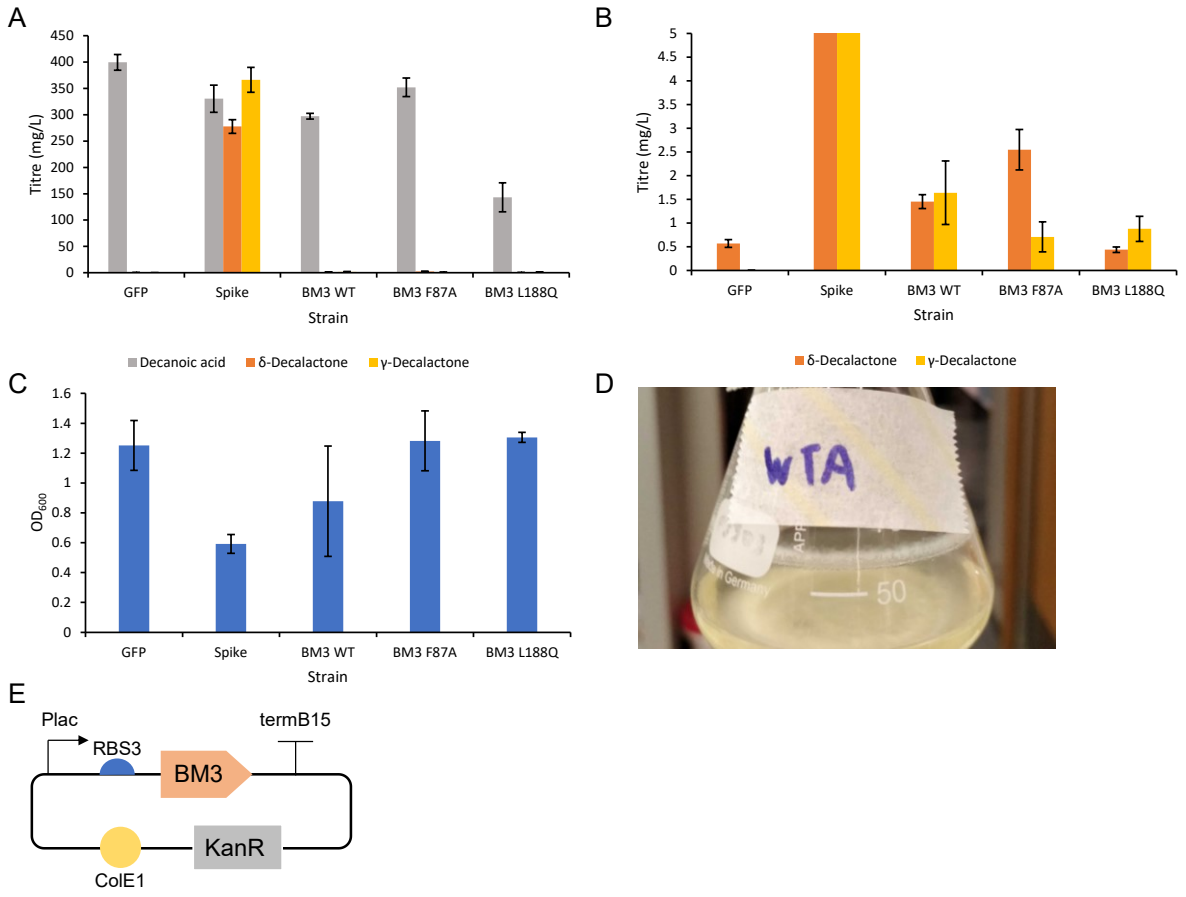


Figure 5.1 Decanoic acid feeding assay *E. coli ΔfadD were transformed with a plasmid expressing either GFP, BM3 or a BM3 mutant. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) δ-decalactone and γ-decalactone titres C) OD_{600} after cultivation in M9 media for 48 hrs at 30°C, 150rpm. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM decanoic acid. The Spike strain was also fed with 2.5 mM δ-decalactone and 2.5 mM γ-decalactone. Products were extracted from the media using chloroform-methanol and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation. D) A ring of fatty material was observed on the flasks after 48 hrs cultivation. E) Plasmid design for expressing BM3. A pET plasmid with ColE1 origin of replication and kanamycin resistance cassette was used. Plac and RBS were used to express GFP, BM3 WT or a BM3 mutant.*

5.2.2 Assaying methyl decanoate as an alternative substrate for feeding assays

The decanoic acid precipitating out of solution presented a challenge for assaying BM3, as it reduced the amount of substrate available to the BM3. As an alternative substrate methyl decanoate was investigated. While methyl decanoate has been reported to be insoluble in water[424], methyl decanoate is a liquid at room temperature, unlike the solid decanoic acid,d which may lead to improved mixing between the substrate and the media. It was therefore decided to investigate the effect of feeding methyl decanoate upon decalactone titres.

E. coli expressing either GFP, the BM3 F87A mutant or a new BM3 V78A F87A L188Q mutant, incorporating regioselectivity altering[166] and activity enhancing mutations[169], were assayed in a feeding experiment in the same manner as before. The *E. coli* strains were fed with either 2.5 mM decanoic acid or 2.5 mM methyl decanoate. After feeding with methyl decanoate, no methyl decanoate was observed after product extraction with chloroform-methanol, instead decanoic acid was observed. Lower decanoic acid titres were observed after feeding methyl decanoate then after feeding decanoic acid (Fig 5.2A).

Feeding methyl decanoate instead of decanoic acid led to a 2-4 fold increase in δ-decalactone and γ-decalactone titres. BM3 V78A F87A L188Q fed with methyl decanoate had the highest titres of both δ-decalactone (3.96 mg/L) and γ-decalactone (6.52 mg/L), however neither of these were significantly higher than BM3 F87A titres ($p = 0.268$ and $p = 0.119$) (Fig 5.2B).

There was no significant difference in OD_{600} between *E. coli* expressing BM3 mutants when fed with decanoic acid or methyl decanoate, however some reduced growth when feeding methyl decanoate was observed in the GFP expressing strain (Fig 5.2C).

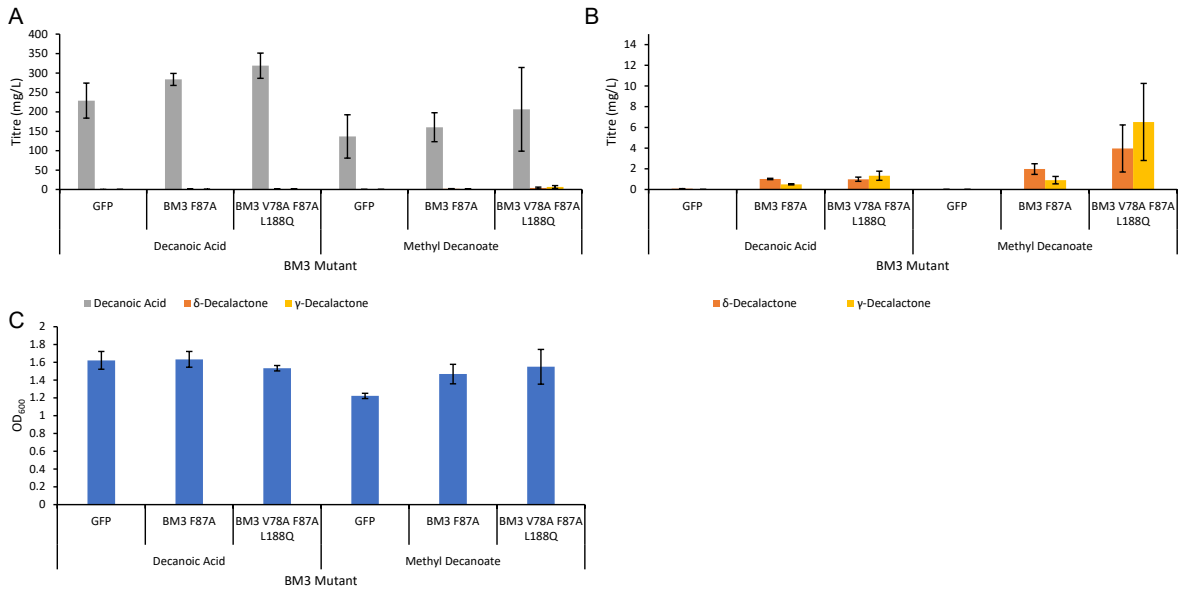


Figure 5.2 Assaying methyl decanoate as an alternative substrate for feeding assays *E. coli ΔfadD were transformed with a plasmid expressing either GFP or a BM3 mutant. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) δ-decalactone and γ-decalactone titres C) OD_{600} after cultivation in M9 media for 48 hrs at 30°C, 150rpm. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM decanoic acid or 2.5 mM methyl decanoate. Products were extracted from the media using chloroform-methanol and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

5.2.3 Could sample acidification improve decalactone titres?

It had previously been reported that condensation of hydroxyfatty acids to lactones could be catalysed by acidic conditions[98,99,127,224]. In parallel with the experiment in Section 5.2.2, cell cultures were sampled twice, and one of the samples was extracted with chloroform-methanol as before, and one sample acidified during the extraction process with chloroform-methanol by the addition of sulfuric acid, to catalyse the condensation of hydroxydecanoic acids to decalactones. This process was repeated for strains fed with both decanoic acid and methyl decanoate, but only the data for strains fed with methyl decanoate is presented here.

Lower decanoic acid titres were observed when samples were treated with sulfuric acid during the extraction process. However, none of these decreases in titre were significant (Fig 5.3A). On the other hand, increased δ-decalactone and γ-decalactone titres were observed in strains expressing BM3 mutants when the samples were treated with sulfuric acid, although again these changes were not significant. Sample acidification led to ~1.5 fold higher δ-decalactone (6.73 mg/L) and γ-decalactone (10.7 mg/L) titres when expressing BM3 V78A F87A L188Q (Fig 5.3B).

The fact that sample acidification led to increased decalactone titres, albeit not significantly, indicated that while there was spontaneous condensation of the 4- and 5-hydroxydecanoic acid to δ- and γ-decalactone, not all of the hydroxyfatty acids spontaneously condensed.

While individually feeding methyl decanoate instead of decanoic acid and sample acidification did not lead to significant increases in titre, taken together these changes had a large effect on decalactone titres, and so were used going forward.

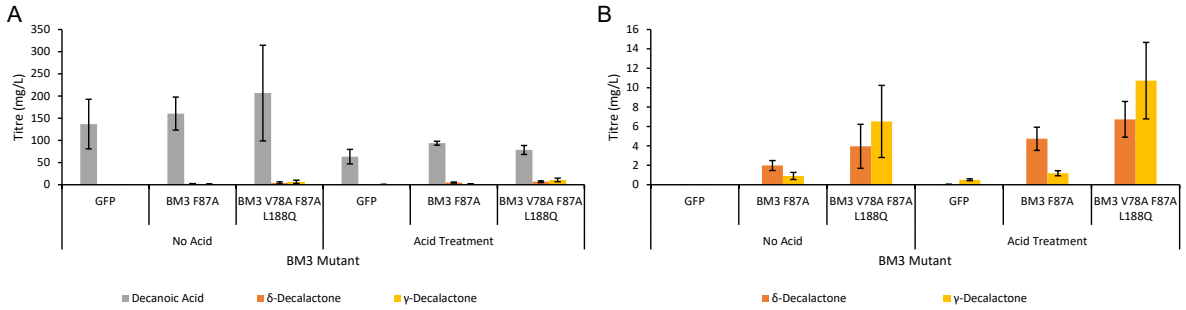


Figure 5.3 Assaying the effect of sample acidification *E. coli ΔfadD were transformed with a plasmid expressing either GFP or a BM3 mutant. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) δ-decalactone and γ-decalactone titres after cultivation in M9 media for 48 hrs at 30°C, 150rpm. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM methyl decanoate. Products were extracted from the media using chloroform-methanol (No Acid), or chloroform-methanol and sulfuric acid (Acid Treatment) and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

5.2.4 Could an overlay be used for product extraction?

An overlay could allow for extraction of the decalactones from the culture media, allowing for simpler and faster product analysis. Furthermore, an overlay could help alleviate toxicity effects[150] from the decalactones at higher concentrations by physically partitioning the *E. coli* and the decalactones.

A feeding experiment was performed, where the different *E. coli* strains were cultured with and without a 10% decane overlay. For both sets of strains, extraction with chloroform-methanol and sulfuric acid was performed upon the culture media to determine decanoic acid and decalactone titres in this phase, and for the strains with an overlay, the overlay was sampled to determine the decanoic acid and decalactone titres in this phase.

It was observed that the decane overlay extracted decanoic acid from the culture media, leading to decreased decanoic acid titres in the media (Fig 5.4A). It was observed that upon addition of a decane overlay, decalactones could be observed in both the culture media and overlay phases, with higher decalactone titres in the overlay phase. However, decalactone titres in both phases were very low. For example, after expression of BM3 V78A F87A L188Q,

γ-decalactone titres of 0.12 mg/L and 0.43 mg/L were respectively observed in the culture media and overlay phases. This gave a total γ-decalactone titre of 0.55 mg/L, which was significantly lower (p= 0.0159) than γ-decalactone titres in the culture media phase in the absence of an overlay (12.5 mg/L) (Fig 5.4B). Presence of an overlay had no clear effect on OD_{600}, with GFP and BM3 F87A expressing strains showing slightly reduced growth, and BM3 V78A F87A L188Q expressing strains showing slightly increased growth.

These data suggested the presence of an overlay was detrimental to bioconversion of decanoic acid to decalactones, due to the overlay extracting decanoic acid out of the media, reducing the amount of substrate available. An overlay was therefore not used going forward.

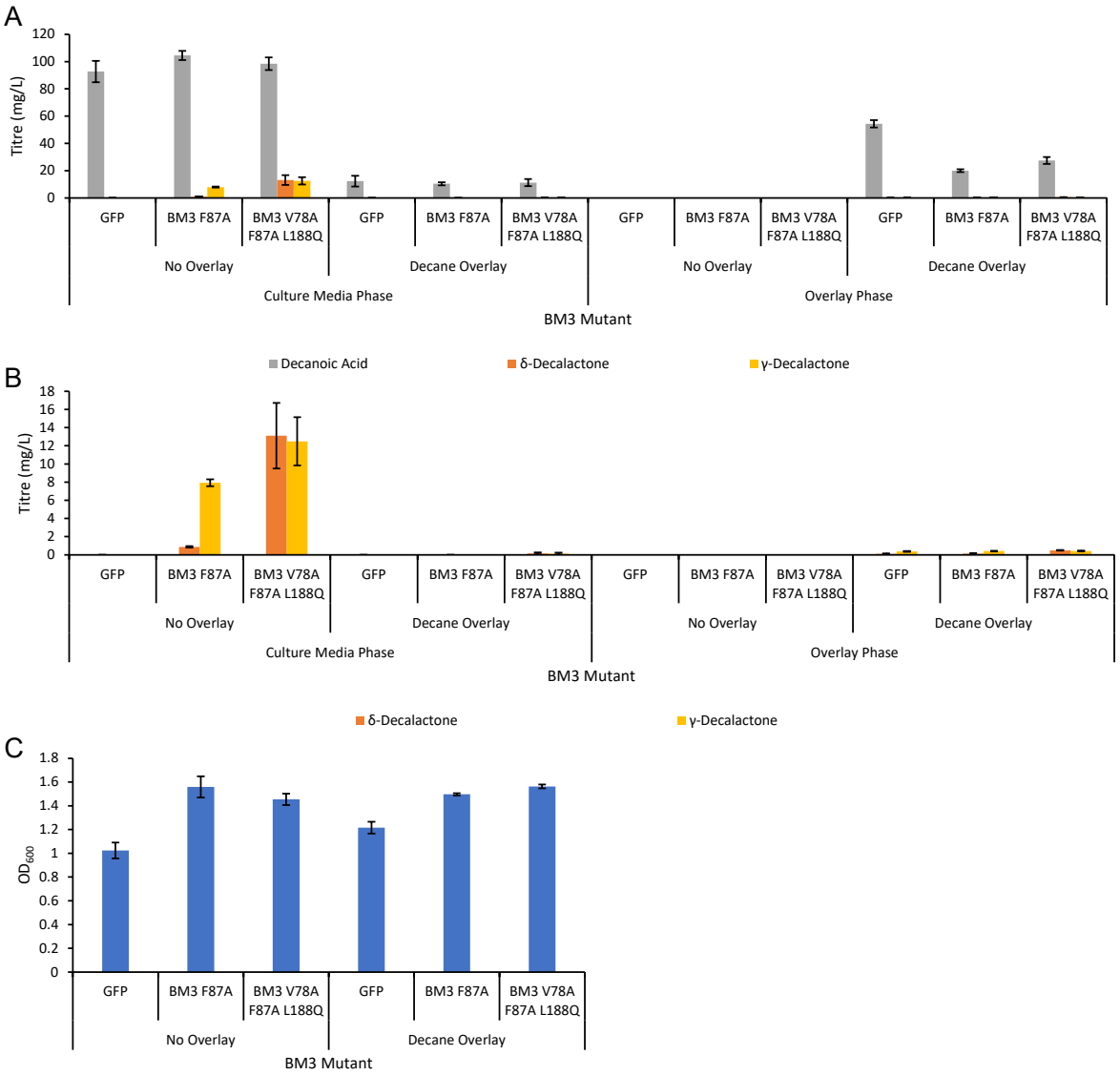


Figure 5.4 Assaying the effect of a decane overlay *E. coli ΔfadD were transformed with a plasmid expressing either GFP or a BM3 mutant. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) δ-decalactone and γ-decalactone titres C) OD_{600} after cultivation in M9 media for 48 hrs at 30°C, 150rpm. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM methyl decanoate. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. If an overlay was present, the overlay was sampled and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

5.2.5 Optimisation of the M9 media for BM3 expression

In the literature some groups have supplemented their media with either 5-aminoleuvinic acid (ALA) or ALA and iron chloride (either $FeCl_2$ or $FeCl_3$) when expressing CYP450s[202,205,206,425]. Iron is a co-factor for the haem prosthetic group, and ALA is a precursor in the haem biosynthesis pathway. Therefore, supplementing the M9 with either iron or ALA could be used

to improve BM3 activity. In the literature, the media was always supplemented, and it was not investigated whether supplementation did have an effect. Therefore, it was investigated whether supplementation did have an effect by comparing M9 with M9 supplemented with either iron, ALA or both.

When attempting to supplement the M9 with the same iron concentration used in the literature (0.05% $FeCl_3$)[206] the iron precipitated out of solution. To solve this problem, M9 was supplemented with iron by adding 6 mg/L ferric ammonium citrate. This is the same concentration of iron as in BG11, and the citrate acts as a chelating agent, allowing the iron to dissolve in the media.

To assay the effect of media supplementation, a feeding experiment was performed using *E. coli* expressing the BM3 V78A F87A L188Q mutant, cultured in the different medias.

When *E. coli* were cultured in supplemented M9 decanoic titres decreased, with significantly lower titres (48.2 mg/L, p = 0.0257) observed when *E. coli* were cultured in M9 supplemented with iron and ALA (Fig 5.5A). On the other hand, increased decalactone titres were only observed when *E. coli* were cultured in M9 supplemented with ferric (Fe3+) ammonium citrate. There was a significant (p = 0.00587) increase in δ-decalactone titres to 18.1 mg/L, and no significant change in γ-decalactone titres (Fig 5.5B). Culturing the *E. coli* in M9 supplemented with ALA led to higher OD_{600}, while supplementation of the M9 with iron only had no significant effect on OD_{600}.

Based on these results, M9 supplemented with iron (M9-Fe) was used going forward.

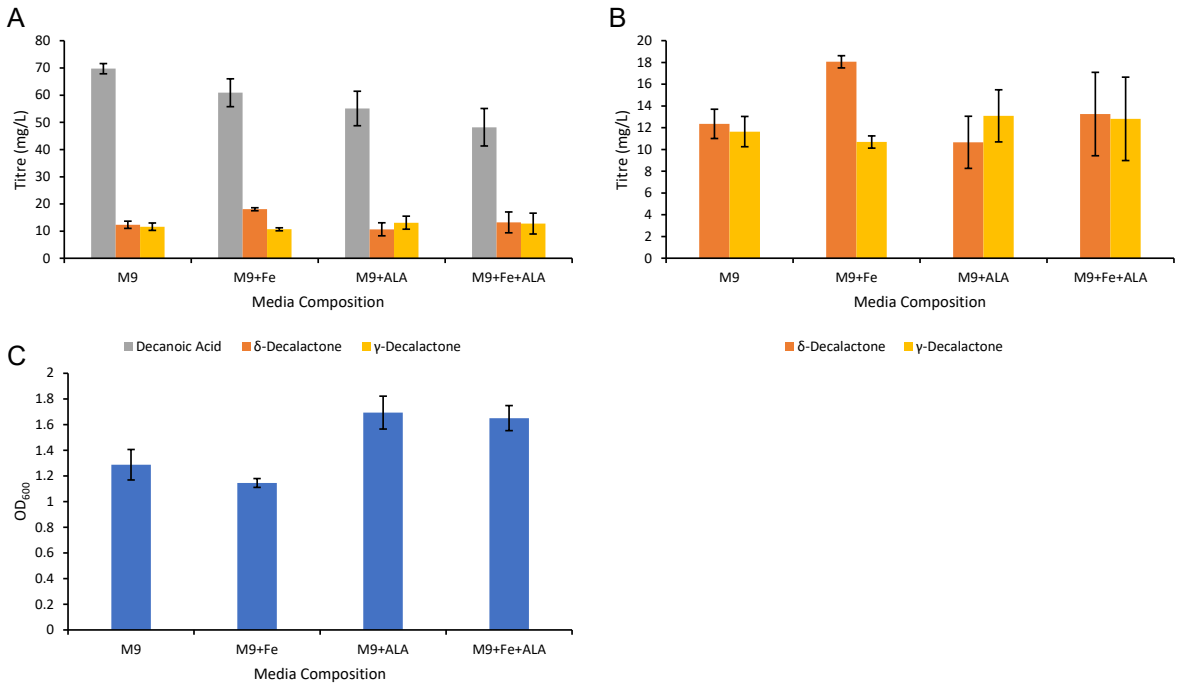


Figure 5.5 M9 media optimisation *E. coli ΔfadD were transformed with a plasmid expressing BM3 V78A F87A L188Q. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) δ-decalactone and γ-decalactone titres C) OD_{600} after cultivation in M9 media, or M9 supplemented with either 6 mg/L ferric ammonium citrate (Fe), 0.2 mM 5-aminoleuvincic acid (ALA) or Fe and ALA for 48 hrs at 30°C, 150rpm. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM methyl decanoate. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

5.2.6 Localisation of decanoic acid, δ-decalactone and γ-decalactone

In parallel with the work in Section 5.2.5, the partitioning of the different products was investigated. During sampling, the strain grown in non-supplemented M9 was sampled twice. One sample (culture media phase) was extracted with chloroform-methanol and sulfuric acid as normal, while the second sample was centrifuged to pellet the cells. The supernatant was removed (aqueous phase) and extracted with chloroform-methanol and sulfuric acid, and the pellet was resuspended in fresh M9 (cell phase) and then extracted with chloroform-methanol and sulfuric acid.

Decanoic acid was found in both cell and aqueous phases, with more decanoic acid localised in the aqueous phase. These data showed that decanoic acid was able to enter the cell, and therefore be available as a substrate for BM3, however the majority remained outside the cell. The total decanoic acid titres for the cell (40.0 mg/L) and aqueous phase (61.9 mg/L) exceeded the decanoic acid titre of the culture media phase (69.7 mg/L) (Fig 5.6A).

Almost all the δ-decalactone and γ-decalactone partitioned to the aqueous phase, with only trace amounts observed in the cell phase. There were no significant differences in δ-decalactone and γ-decalactone titre between the culture media and aqueous phases (Fig 5.6B).

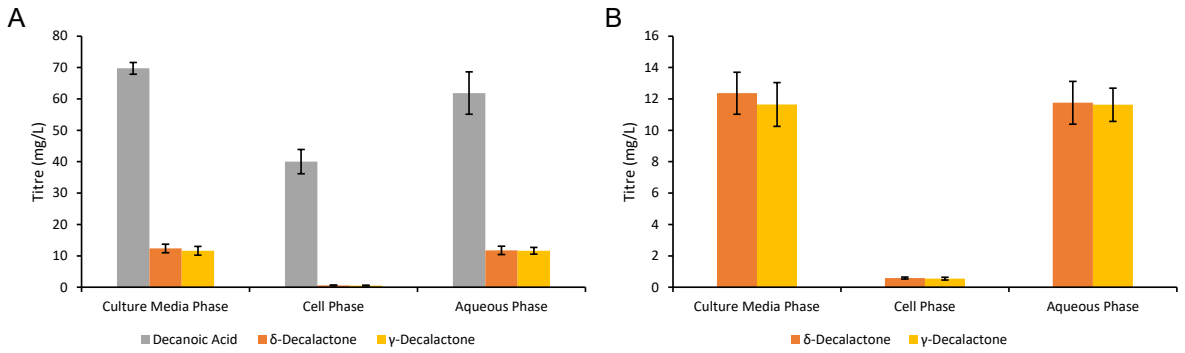


Figure 5.6 Partitioning of decanoic acid, δ-decalactone and γ-decalactone *E. coli ΔfadD were transformed with a plasmid expressing BM3 V78A F87A L188Q. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) δ-decalactone and γ-decalactone titres after cultivation in M9 media for 48 hrs at 30°C, 150rpm. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM methyl decanoate. Products were extracted from the media (Culture Media Phase), the supernatant after the culture was centrifuged (aqueous phase) or the pellet after centrifugation resuspended in fresh M9 (cell phase) using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

5.2.7 Screening alternative CYP450s

A CYP450 from *Tepidiphilus thermophilus* (TtCYP450) was reported to hydroxylate decanoic acid with 90% ω-5 regioselectivity[205]. TtCYP450 could therefore be used to convert decanoic acid to 5-hydroxydecanoic acid and therefore δ-decalactone. A BlastP search using TtCYP450 as a template found that an uncharacterised CYP450 from *Rhodocyclales bacterium* (RbCYP450) had 92.8% sequence similarity to TtCYP450. It was hypothesised that due to the high sequence similarity RbCYP450 would have similar activity to TtCYP450. Therefore, both RbCYP450 and TtCYP450 were expressed in *E. coli* and assayed using a feeding experiment, with *E. coli* expressing GFP, BM3 WT and BM3 V78A F87A L188Q used as controls.

Upon expression of either RbCYP450 or TtCYP450, no significant changes in decanoic acid titres were observed compared to the GFP control ($p = 0.640$ and $p = 0.748$) (Fig 5.7A). No γ-decalactone production was observed and only trace amounts of δ-decalactone were observed (Fig 5.7B). No significant changes in OD_{600} were observed either (Fig 5.7C).

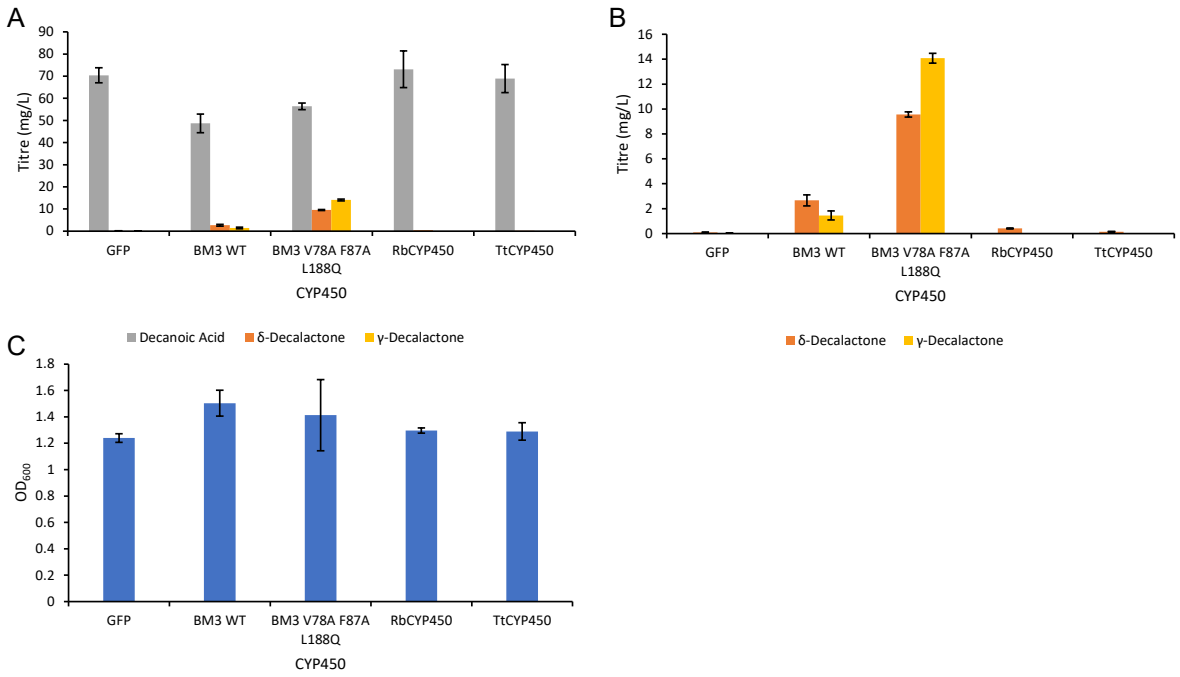


Figure 5.7 Screening new CYP450s *E. coli ΔfadD were transformed with a plasmid expressing either GFP or a CYP450. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) δ-decalactone and γ-decalactone titres C) OD_{600} after cultivation in M9-Fe media for 48 hrs at 30°C, 150rpm. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM methyl decanoate. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

5.2.8 Could deletion of *iscR* improve the activity of the new CYP450s?

Unlike BM3, RbCYP450 and TtCYP450 have reductase domains that contain iron-sulphur clusters[205]. An *E. coli* BL21 ΔiscR strain was available in the lab. *iscR* is a transcriptional repressor of the isc locus which is involved in the assembly of iron-sulphur clusters, and deletion of *iscR* in this strain had been found to increase the activity of a hydrogenase that contained iron-sulphur clusters[397]. Therefore, deletion of *iscR* could be used to improve the activity of the new CYP450s.

The new CYP450s, and controls of GFP, BM3 WT and BM3 V78A F87A L188Q were expressed in both BL21 wild-type (BL21 WT) and BL21 ΔiscR and fed methyl decanoate. As the BL21 ΔiscR strain had a kanamycin resistance cassette, a slight modification of the backbone in Fig 5.1E was used, where the kanamycin resistance cassette was replaced with a gentamycin resistance cassette.

Upon deletion of *iscR*, decanoic acid titres in *E. coli* expressing RbCYP450 and TtCYP450 increased compared to the BL21 WT strains, and were higher than the GFP expressing strain (Fig 5.8A). Deletion of *iscR* also had no effect on δ-decalactone titres, with only trace amounts

produced in both BL21 WT and BL21 ΔiscR when expressing RbCYP450 or TtCYP450. Production of γ-decalactone could not be observed in either host background (Fig 5.8B). Deletion of *iscR* had no impact on growth (Fig 5.8C).

Unexpectedly, deletion of *iscR* approximately doubled δ-declactone and γ-decalactone titres when BM3 V78A F87A L188Q was expressed (Fig 5.9B), however these changes were not significant ($p = 0.215$ and $p = 0.139$). After reviewing the literature, no clear explanation for why deletion of *iscR* could lead to an increase in BM3 activity could be proposed.

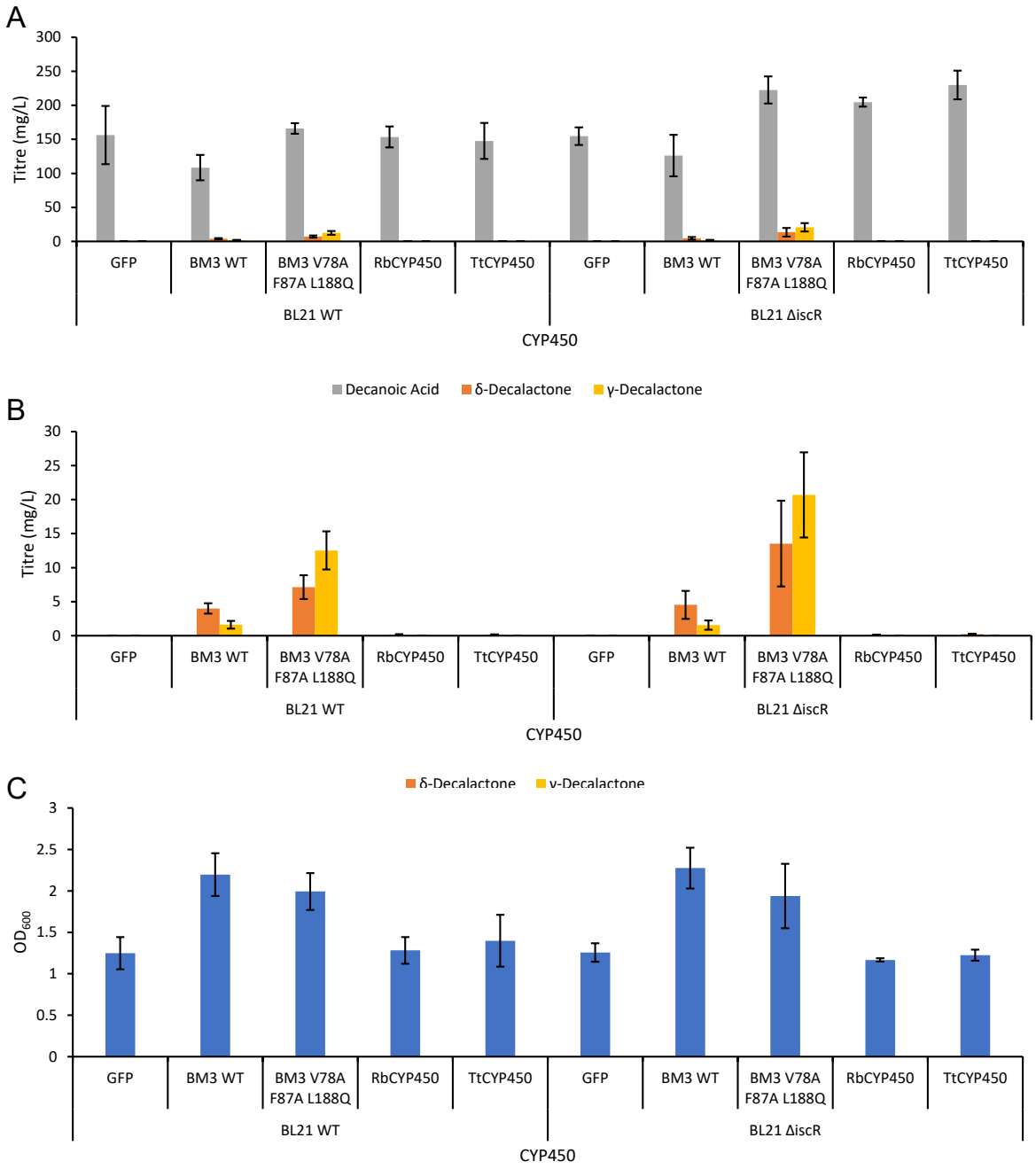


Figure 5.8 The effect of ΔiscR *E. coli BL21 WT and BL21 ΔiscR were transformed with a plasmid expressing either GFP or a CYP450. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) δ-decalactone and γ-decalactone titres C) OD_{600} after cultivation in M9-Fe media for 48 hrs at 30°C, 150rpm. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM methyl decanoate. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

5.2.9 Mutagenesis of BM3

It was observed that the BM3 V78A F87A L188Q mutant consistently made roughly equal amounts of δ-decalactone and γ-decalactone, implying the mutant had roughly equal ω-5 and

ω-6 regioselectivity. However, it had been reported that the BM3 V78A F87A double mutant had much greater ω-5 regioselectivity than ω-6[166]. Moreover, in this work BM3 V78A F87A had not been compared to BM3 V78A F87A L188Q to confirm the L188Q mutation was having the expected activity enhancing effects. Therefore, mutagenesis of BM3 was systematically investigated. Two mutants with altered regioselectivity, V78A and V78A F87A were assayed individually, in combinatorial mutants with the activity enhancing mutation L188Q[169] and in combinatorial mutants with L188Q and a second activity enhancing mutation, I401P[177].

All BM3 mutants had higher decanoic acid titres higher then the BM3 WT, but lower than the GFP control. BM3 mutants with both V78A and F87A had higher decanoic titres than BM3 mutants with V78A only, although the titres were not significantly higher (Fig 5.9A).

As previously observed, the BM3 F87A mutant produced significantly more δ-decalactone (8.17 mg/L, $p = 0.00233$) than the BM3 WT. Addition of L188Q increased δ-decalactone titre to 11.8 mg/L, and addition of L188Q and I401P further increased the titre to 15.2 mg/L, significantly higher than BM3 F87A titres ($p = 0.000997$) (Fig 5.9B)

The BM3 V78A F87A mutant produced significantly more δ-decalactone (8.96 mg/L, $p = 0.0193$) and γ-decalactone (13.2 mg/L, $p = 0.0160$) than the BM3 WT. Addition of L188Q increased both δ-decalactone and γ-decalactone titres to 10.8 mg/L and 16.2 mg/L, but this increase was not significant. Further addition of I401P led to no γ-decalactone production being observed, and observation of only trace amounts of δ-decalactone (Fig 5.9B).

All BM3 mutants showed similar growth, with no visible trends in OD_{600} (Fig 5.9C).

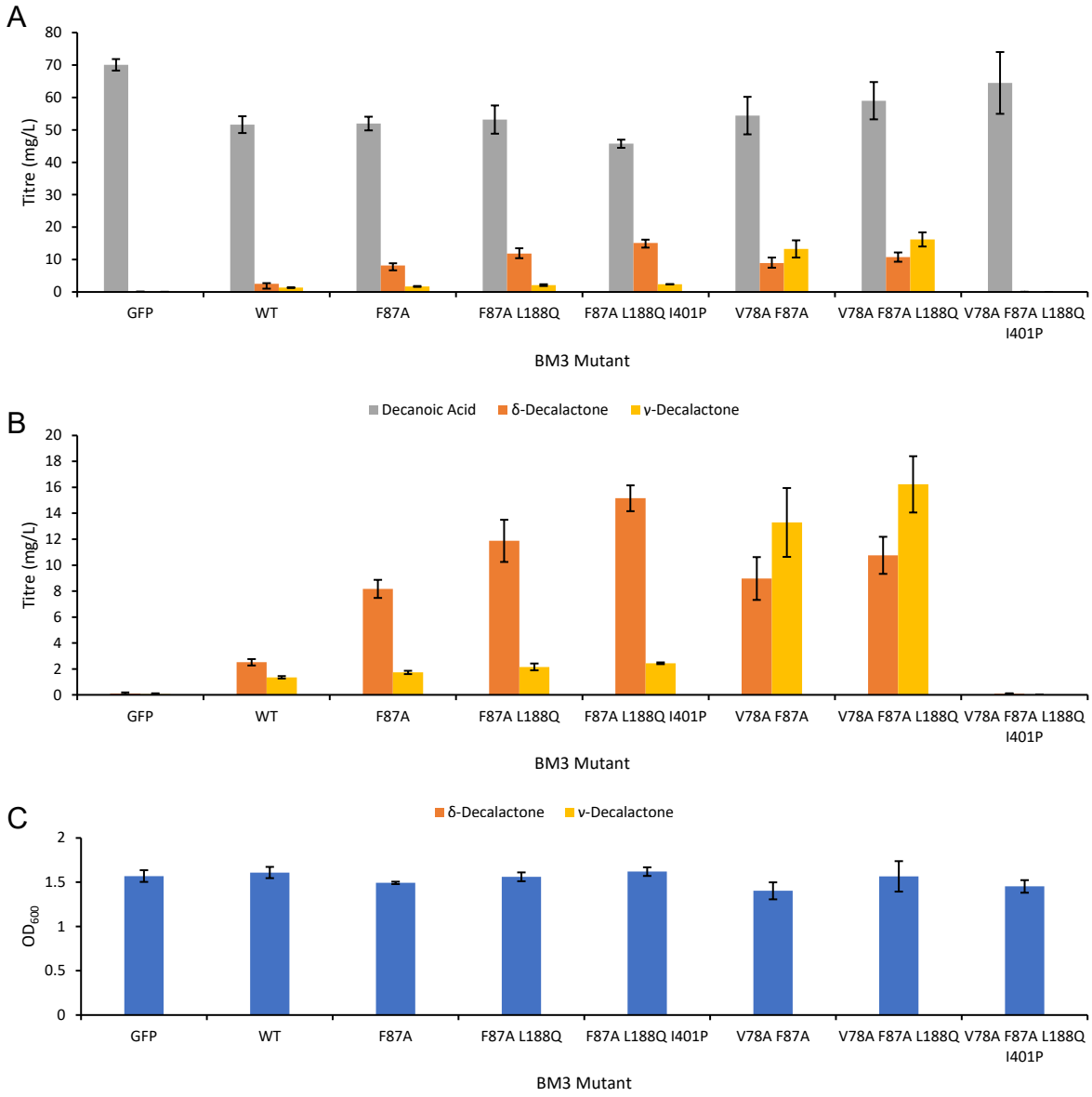


Figure 5.9 Screening BM3 mutants *E. coli ΔfadD were transformed with a plasmid expressing either GFP, BM3 WT or a BM3 mutant. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) δ-decalactone and γ-decalactone titres C) OD_{600} after cultivation in M9-Fe media for 48 hrs at 30°C, 150rpm. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM methyl decanoate. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

5.2.10 Mutagenesis of the BM3 F87 site

Different point mutations of the BM3 F87 residue had been reported to have different regioselectivities from F87A. F87G and F87V mutants from the literature[166], as well as a new

F87I mutation were screened by a feeding assay. GFP, BM3 WT, BM3 F88A and BM3 V78A F87A were used as controls.

Decanoic acid titres were similar in *E. coli* expressing BM3 F87G and F87A, while F87V and F87I had reduced decanoic acid titres relative to F87A (Fig 5.10A). The BM3 F87G mutant had significantly higher δ-decalactone titres (17.0 mg/L, p = 0.0160) and γ-decalactone titres (4.63 mg/L, p = 0.00302) than the BM3 F87A mutant. δ-decalactone and γ-decalactone titres in the BM3 V78A F87A mutant were lower than previously observed, making comparison of BM3 F87G and BM3 V78A F87A difficult. The F87V and F87I mutants had slightly higher δ-decalactone titres than the BM3 WT, but titres were lower than those of the original F87A mutant (Fig 5.11B). None of the mutants had a significant effect on OD_{600} (Fig 5.11C).

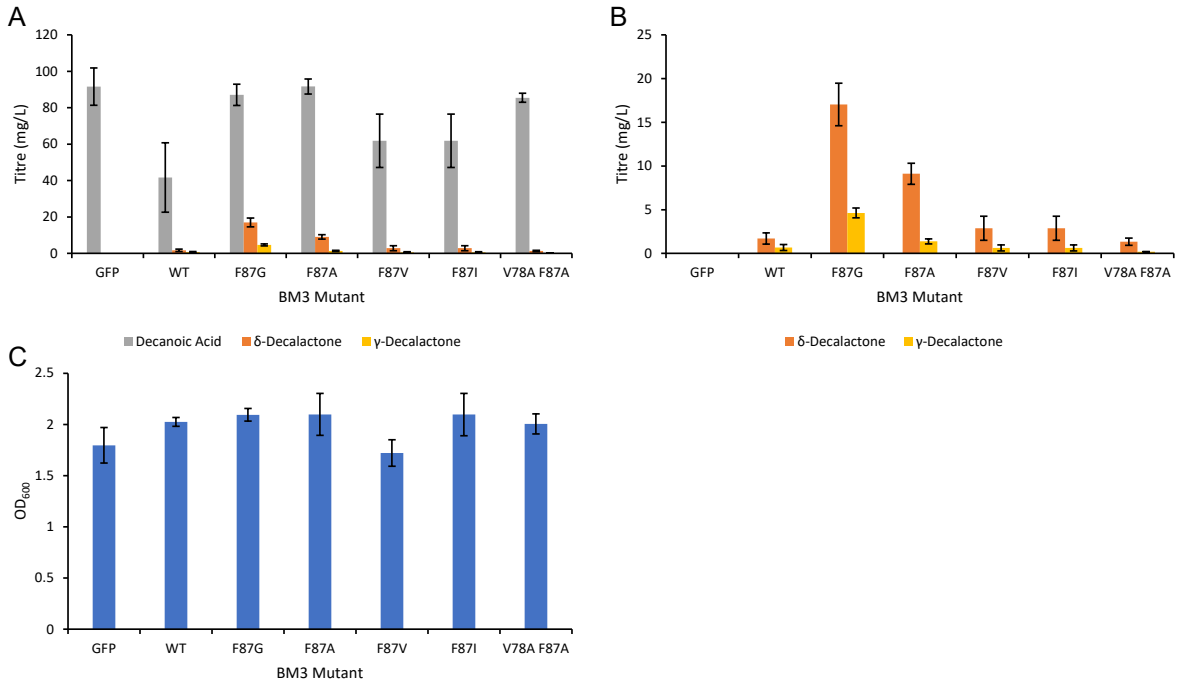


Figure 5.10 Screening BM3 F87 mutants *E. coli ΔfadD were transformed with a plasmid expressing either GFP, BM3 WT or a BM3 mutant. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) δ-decalactone and γ-decalactone titres C) OD_{600} after cultivation in M9-Fe media for 48 hrs at 30°C, 150rpm. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM methyl decanoate. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

5.2.11 Adding activity enhancing mutations to BM3 F87G

The BM3 F87G mutant was then engineered to be more active by combining F87G with the activity enhancing mutations L188Q and I401P. Combinatorial mutants were generated and screened using a feeding experiment. *E. coli* expressing GFP, BM3 WT, BM3 V78A F87A L188Q and BM3 F87A L188Q I401P were used as controls.

It was noticed that titres of decanoic acid, δ-decalactone and γ-decalactone were higher than normally observed in the controls. Despite this, inferences could be made about the relative activities of the new mutants.

Combining activity enhancing mutations with F87G unexpectedly led to higher decanoic titres, although the increases in titre were not significant (Fig 5.11A). Only the BM3 F87G L188Q mutant showed increased δ-decalactone titres, which reached 59.7 mg/L. This was an increase relative to both BM3 F87G (25.7 mg/L) and BM3 F87A L188Q I401P (32.8 mg/L), however due to large variation observed for all the mutants these increases were not significant. The I401P mutation appeared detrimental to activity, with both the BM3 F87G I401P and BM3 F87G L188Q I401P mutants producing less δ-decalactone than BM3 F87G (Fig 5.11B). Observed OD_{600} in the controls was also higher than previously observed and adding the activity enhancing mutations to BM3 F87G did not have significant effects on growth (Fig 5.11C).

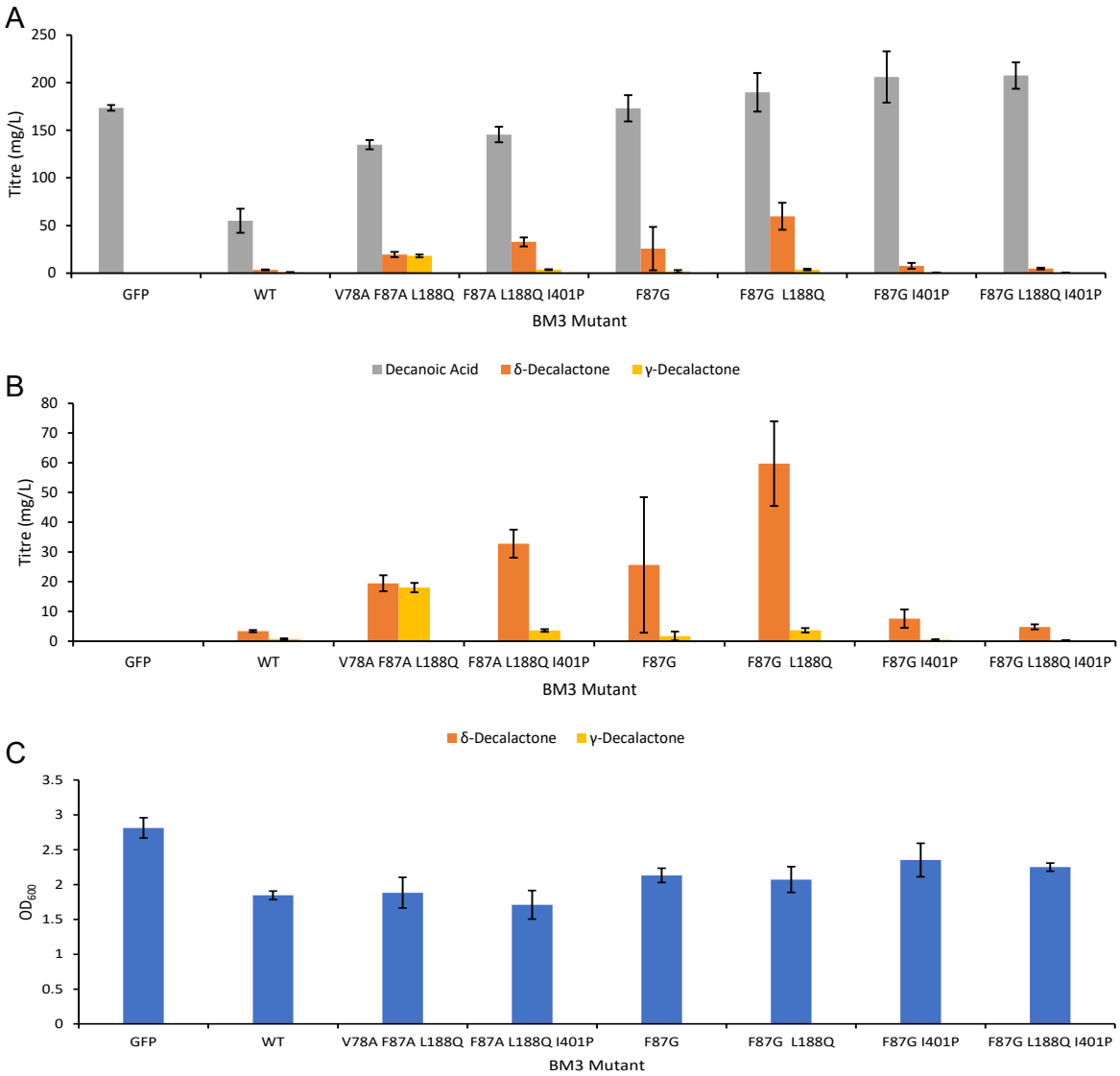


Figure 5.11 Adding activity enhancing mutations to BM3 F87G *E. coli ΔfadD were transformed with a plasmid expressing either GFP, BM3 WT or a BM3 mutant. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) δ-decalactone and γ-decalactone titres C) OD_{600} after cultivation in M9-Fe media for 48 hrs at 30°C, 150rpm. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM methyl decanoate. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

5.2.12 Extension of the decalactone biosynthesis pathway

This work was performed in parallel to Sections 5.2.10 and 5.2.11.

Previous results in Section 5.2.3 had demonstrated that not all of the 4- and 5-hydroxydecanoic acid spontaneously condensed to γ-decalactone and δ-decalactone, and that sample acidification could help catalyse the condensation reaction. A hydroxyfatty acid

CoA ligase had been reported in the literature, AlkK from *Pseudomonas oleovorans* (PoAlkK)[225], as well as a homologue with 3-hydroxydecanoic acid CoA ligase activity from *Pseudomonas putida* (Pp0763)[226]. An alcohol acyltranserase, PpAAT1 that catalyses the condensation of 4-hydroxydecanoyl CoA to γ-decalactone had also been reported in peaches (*Prunus persica*) [227]. Together these genes could be used to extend the proposed decalactone biosynthesis pathway and enzymatically condense 4- and 5-hydroxydecanoic acid to γ- and δ-decalactone, removing the requirement for acid treatment of the samples.

To test these genes, BM3 V78A F87A L188Q, AlkK and PpAAT1 were incorporated into a synthetic pathway for bioconversion of decanoic acid to decalactones (Fig 5.12A). BM3 V78A F87A L188Q was expressed on one plasmid (Fig 5.12B), and a second plasmid was used to express either AlkK (Fig 5.12C) or AlkK and PpAAT1 (Fig 5.12D). Alternatively, as a control GFP was expressed on the second plasmid using the design in Fig 5.12C. This would allow the ligation of 4- and 5-hydroxydecanoic acids with CoA to be observed, and then the subsequent conversion of the 4- and 5-hydroxydecanoyl-CoA to γ- and δ-decalactone.

A significant challenge to this plan was the lack of availability of standards. Standards for 4-hydroxydecanoic acid, 4-hydroxydecanoyl CoA and 5-hydroxydecanoyl-CoA were not available, and only a small amount (100 mg) of Sodium 5-hydroydecanoate was available, which was not sufficient for preparing GC-MS standards. To overcome this, a workaround was devised. When samples were taken, the culture media was sampled twice, and once extracted with chloroform-methanol and once with chloroform-methanol and sulfuric acid. The extraction process with only chloroform-methanol would give the titres of decalactones that had spontaneously condensed. The extraction process with chloroform-methanol and sulfuric acid would give the combined titres of hydroxydecanoic acids and decalactones, and so the two values could be used to estimate the hydroxydecanoic acid titres.

The expectation was that expression of AlkK would lead to a drop in hydroxydecanoic acid titres, as they were ligated to CoA, and that this would lead to a drop in decalactone titres without acidification due to the amount of precursor being reduced. Expression of AlkK and PpAAT1 would lead to a drop in hydroxydecanoic acid titres as before, but then lead to an increase in decalactone titres without acidification, as the hydroxydecanoyl-CoA was condensed to decalactones.

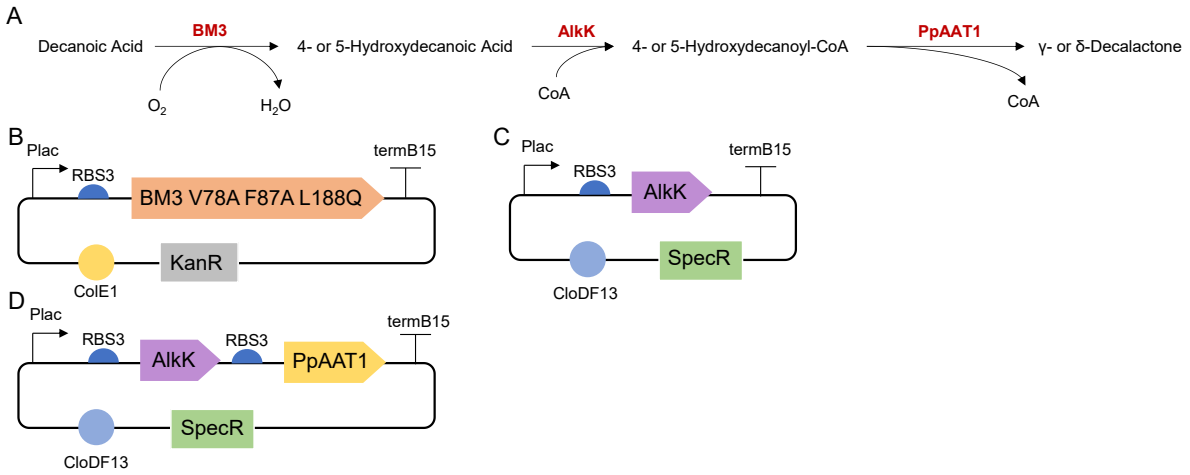


Figure 5.12 Extension of the decalactone biosynthesis pathway *A) Pathway for conversion of decanoic acid to decalactones. Pathway enzymes are highlighted in red. BM3: CYP450 CYP102A1 (Priestia megaterium), AlkK: Hydroxyfatty acid CoA ligase, PpAAT1: Prunus persica alcohol acyltransferase 1 B) Plasmid design for expressing BM3 V78A F87A L188Q. A pET plasmid with ColE1 origin of replication and kanamycin resistance cassette was used. Plac and RBS3 were used to express BM3 V78A F87A L188Q. C) Plasmid design for expressing AlkK. A pCDF plasmid with CloDF13 origin of replication and spectinomycin resistance cassette was used. Plac and RBS3 were used to express AlkK. As a control GFP was expressed instead. D) Plasmid design for expressing AlkK and PpAAT1. A pCDF plasmid with CloDF13 origin of replication and spectinomycin resistance cassette was used. Plac and RBS3 were used to express AlkK and PpAAT1 in a single operon.*

E. coli strains co-expressing BM3 V78A F87A L188Q and either GFP, AlkK or AlkK and PpAAT1 were assayed using a feeding experiment.

Expression of PoAlkK had no significant effect on decanoic acid titres, however expression of Pp0763 significantly ($p = 0.0130$) reduced decanoic acid titres from 79.3 mg/L to 44.9 mg/L (Fig 5.13A). Expression of PoAlkK or Pp0763 had no significant effect on estimated 5-hydroxydecanoic acid titres. Expression of PoAlK slightly decreased estimated 4-hydroxydecanoic acid titres, while expression of Pp0763 resulted in 4-hydroxydecanoic acid titres that were estimated to be negative. Co-expression of PoAlkK and Pp0763 also led to negative estimated 4-hydroxydecanoic acid titres (Fig 5.13B).

Expression of PoAlkK and PpAAT1 led to a small but significant ($p = 0.0262$) increase in γ-decalactone titres from 13.5 mg/L to 15.4 mg/L, but had no significant impact on δ-decalactone titres. Expression of Pp0763 and PpAAT1 had no significant effects on either δ-decalactone or γ-decalactone titres (Fig 5.13B).

Expression of only Pp0763 reduced growth, otherwise no significant differences in OD_{600} were observed (Fig 5.13C).

While a potential increase in γ-decalactone titres was observed, due to a lack of understanding of what was happening in the system due to a lack of standards, and difficulties in making reliable estimates of what might be occurring, highlighted by repeated negative titres, this idea was not taken further.

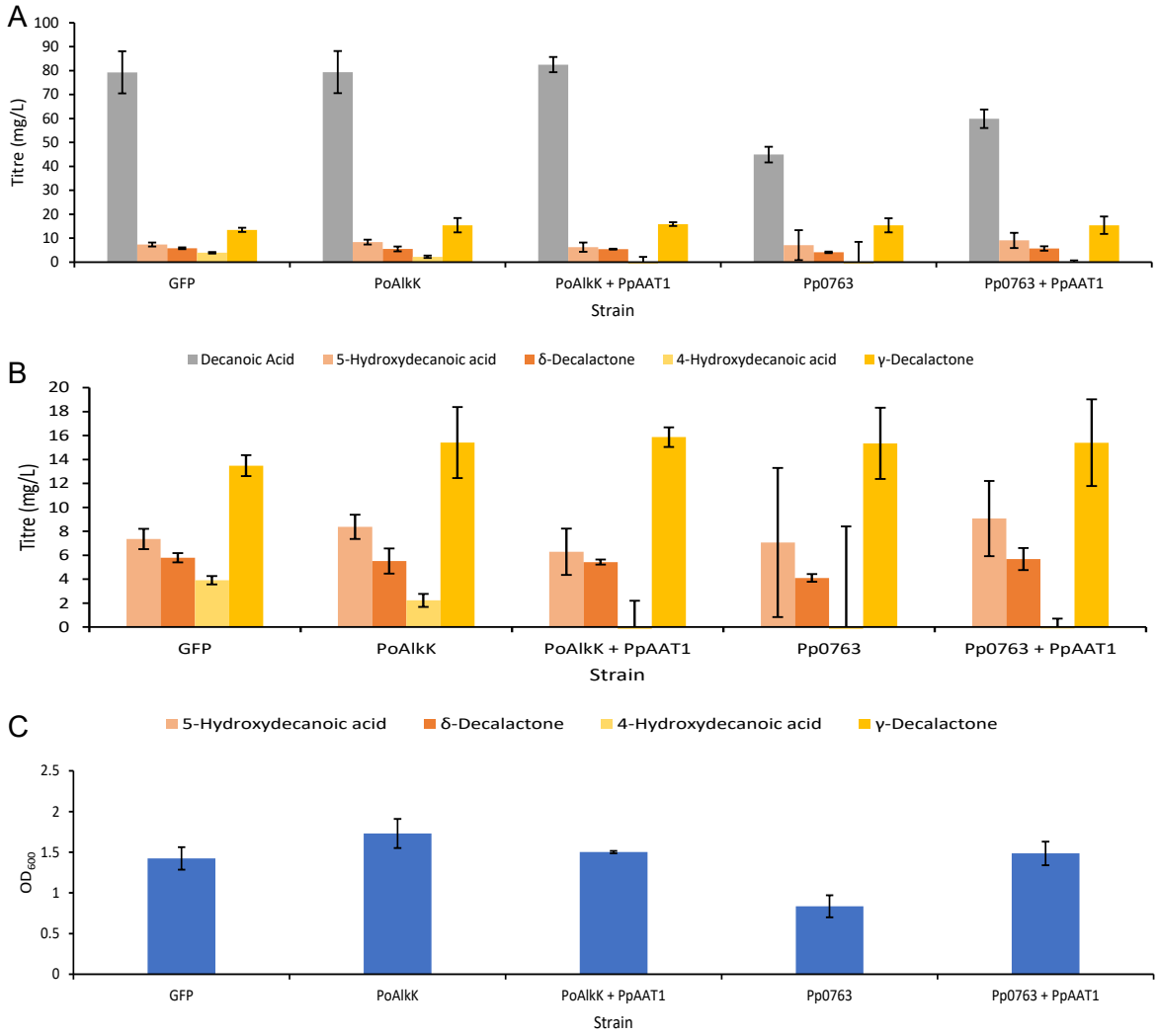


Figure 5.13 Screening AlkK and PpAAT1 *E. coli ΔfadD were co-transformed with two plasmids, one expressing BM3 V78A F87A L188Q, and one expressing either GFP, an AlkK or an AlkK and PpAAT1 A) Decanoic acid, 5-hydroxydecanoic acid, δ-decalactone, 4-hydroxydecanoic and γ-decalactone titres B) 5-hydroxydecanoic acid, δ-decalactone, 4-hydroxydecanoic acid and γ-decalactone titres C) OD_{600} after cultivation in M9-Fe media for 48 hrs at 30°C, 150rpm. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM methyl decanoate. Products were extracted from the media using chloroform-methanol or chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

5.2.13 Could increasing oxygen supply improve BM3 activity?

This work was performed in parallel to Sections 5.2.10 and 5.2.11.

Oxygen is a cofactor for the reaction catalysed by BM3[159], so increasing oxygen supply was investigated as a method of improving BM3 activity. To increase oxygen supply baffled shake flasks were used. Only 250 mL baffled shake flasks were available in the lab, so the *E. coli* were cultured in 50 mL M9 in 250 mL shake flasks with and without baffling. To investigate the effects of baffling, a feeding experiment using *E. coli* expressing either GFP or BM3 V78A F87A L188Q was performed.

Higher than normal titres of decanoic acid, δ-decalactone and γ-decalactone were observed in the strains expressing GFP and BM3 V78A F87A L188Q cultured in non-baffled shake flasks. When *E. coli* expressing BM3 V78A F87A L188Q were cultured in baffled shake flasks slightly increased decanoic acid titres (Fig 5.14A) and δ-decalactone and γ-decalactone titres (Fig 5.14B) were observed compared to non-baffled shake flasks. However, none of these changes were significant. *E. coli* expressing GFP showed significantly improved growth when cultured in baffled shake flasks, otherwise no changes were observed (Fig 5.14C).

When analytical standards are run on the GC-MS a calibration curve of the form y = mx is generated and used to calculate product titres in mg/L. Investigation of the calibration curve found the m values to be approximately half of the m values of prior calibration curves. This would have resulted in higher calculated titres, explaining why titres in the non-baffled flasks were much higher than previously observed. As a precaution, the standards were remade for future work. Despite this issue with the standards, from these results it could be concluded that baffling did not have a significant effect on decalactone titres, and so the experiment was not repeated.

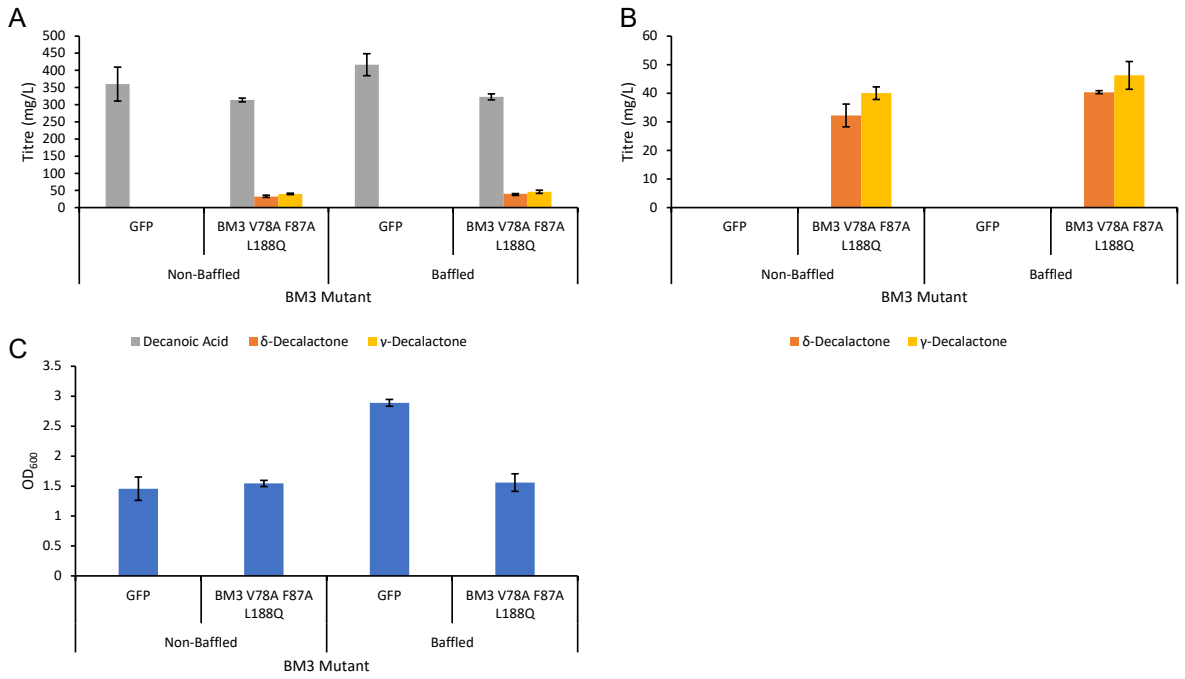


Figure 5.14 Effect of increasing oxygen supply *E. coli ΔfadD were transformed with a plasmid expressing either GFP or BM3 V78A F87A L188Q. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) δ-decalactone and γ-decalactone titres C) OD_{600} after cultivation in M9-Fe media for 48 hrs at 30°C, 150rpm. Cells were cultured in either a 250 mL shake flask without baffling or a baffled 250 mL shake flask. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM methyl decanoate. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

5.2.14 Could NADPH recycling improve BM3 activity?

The reaction catalysed by BM3 also uses NADPH as a cofactor[194]. Glucose dehydrogenase (GDH) oxidises glucose to D-glucono-1,5-lactone, reducing either NAD or NADP in the process. GDHs have been used to recycle NADPH when characterising CYP450s[99,199,202,203,205,206] and co-expression of GDH with BM3 has been observed to increase product yields[220,221,223]. It was therefore decided to investigate co-expression of BM3 and GDH to see if this would improve BM3 activity and therefore lactone titres.

The modified pathway for bioconversion of decanoic acid to decalactones is shown in Fig 5.15A. As NADPH was oxidised to NADP during the hydroxylation of decanoic acid, GDH would recycle the NADP to NADPH, regenerating the co-factor. Either only a BM3 variant was expressed as control (Fig 5.15B), or a BM3 variant was co-expressed with a GDH (Fig 5.15C).

For the BM3 variants, either BM3 WT, BM3 V78A F87A L188Q or BM3 F87G L188Q were expressed. For the GDH, either *Priestia megaterium* GDH1 (BmGDGH1) or *P. megaterium*

GDH2 (BmGDH2) were expressed, as these had been reported to have the highest specificities for NADP amongst *P. megaterium*'s native GDHs[212]. Also available in the lab was a glyceraldehyde-3-phosphate dehydrogenase from *Clostridium acetobutylicum* (GapN). In unpublished work it had been shown to regenerate NADPH, improving activity of a transhydrogenase. GapN was therefore also investigated and expressed like the GDHs.

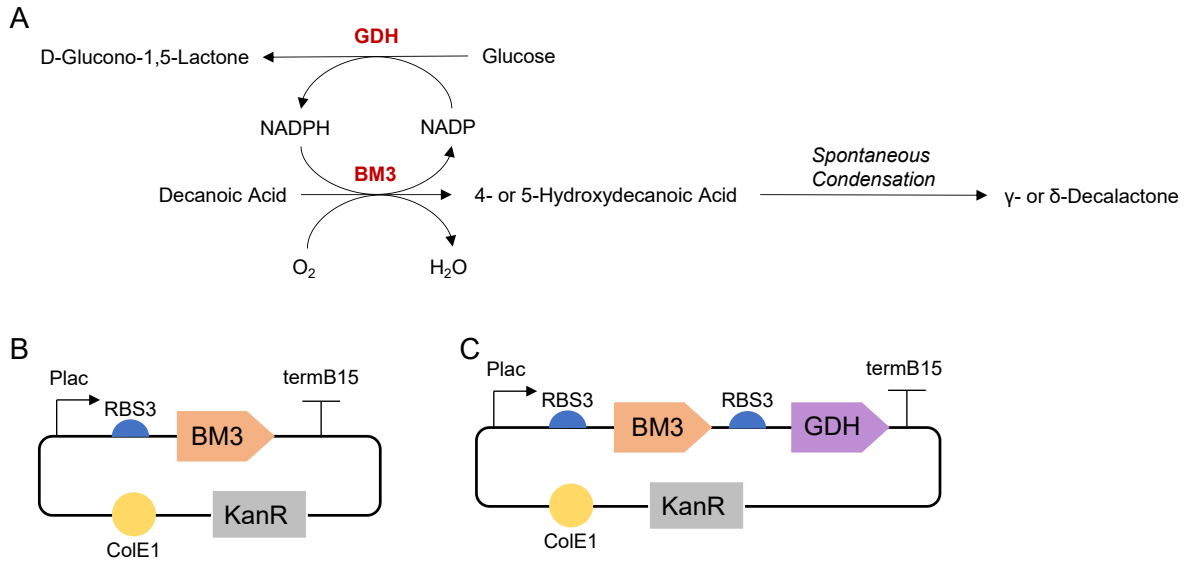


Figure 5.15 Recycling NADPH using GDH *A) Pathway for conversion of decanoic acid to decalactones. Pathway enzymes are highlighted in red. BM3: CYP450 CYP102A1 (Priestia megaterium), GDH: Glucose dehydrogenase B) Plasmid design for expressing BM3 variants. A pET plasmid with ColE1 origin of replication and kanamycin resistance cassette was used. Plac and RBS3 were used to express BM3 variants. C) Plasmid design for expressing BM3 variants and GDH. A pET plasmid with ColE1 origin of replication and kanamycin resistance cassette was used. Plac and RBS3 were used to express BM3 variants and GDH in a single operon.*

When BM3 WT was co-expressed with GapN or BmGDH3, no significant change in decanoic acid titres was observed. When BM3 WT was co-expressed with BmGDH2 a decrease in decanoic acid titre was observed. On the other hand, when either BM3 V78A F87A L188Q or BM3 F87G L188Q was co-expressed with GapN or a GDH, decanoic acid titres increased slightly, but not significantly (Fig 5.16A).

δ-decalactone and γ-decalactone titres remained low when BM3 WT was co-expressed with GapN or a GDH. Significant increases in δ-decalactone were observed when BM3 V78A F87A L188Q was co-expressed with GapN (18.7 mg/L, $p = 0.0467$) or BmGDH2 (19.0 mg/L, $p = 0.0392$). Co-expression of BM3 V78A F87A L188Q with GapN or a GDH did not lead to significant changes in γ-decalactone titres. On the other hand, co-expression of BM3 F87G L188Q led to mostly insignificant decreases in δ-decalactone and γ-decalactone

titres (Fig 5.16B). There were no apparent trends in effects on growth when co-expressing a BM3 variant and a GDH (Fig 5.16C).

As significant increases in both δ-decalactone and γ-decalactone were not observed, this idea was not investigated further.

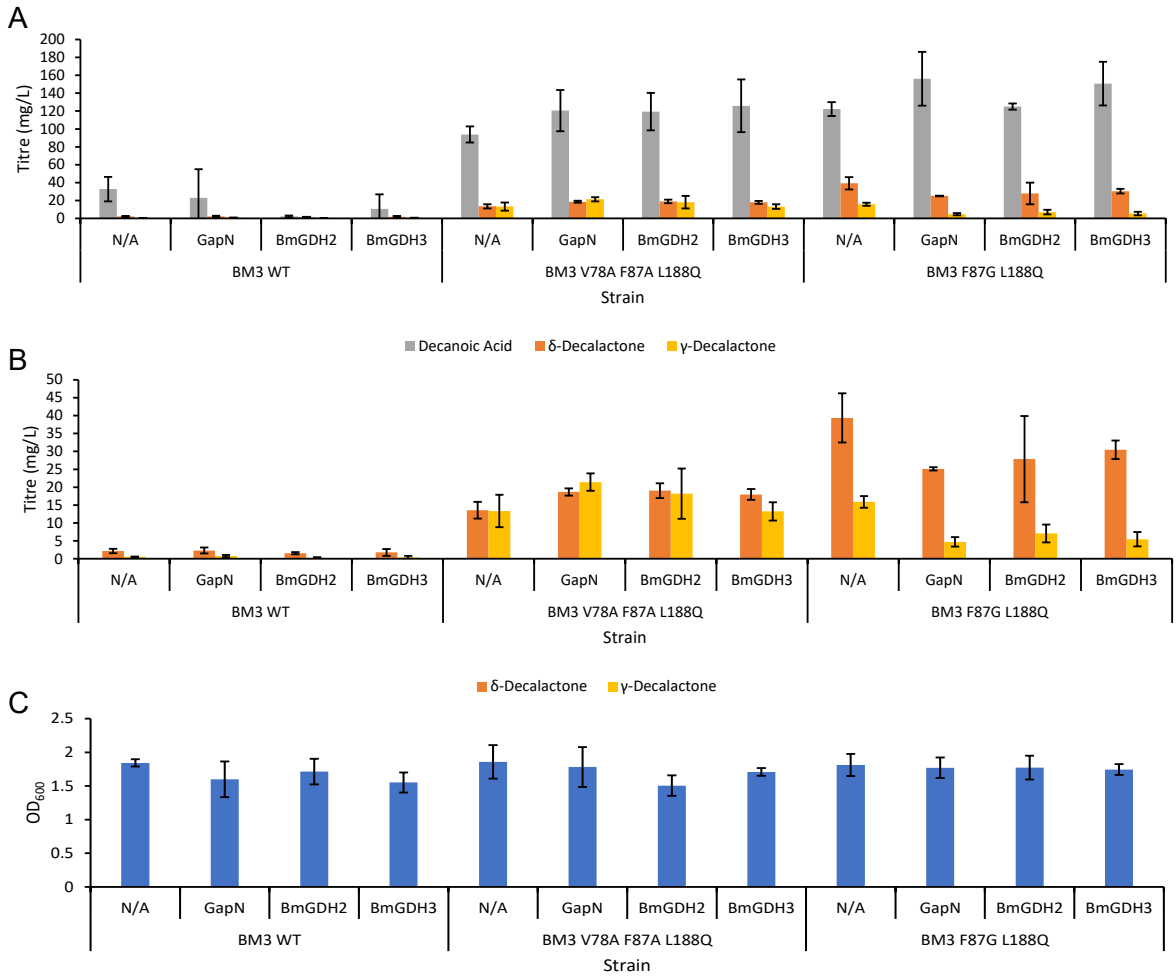


Figure 5.16 Co-expression of BM3 and GDH *E. coli ΔfadD were transformed with a plasmid expressing either a BM3 variant or a BM3 variant and a GDH. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) δ-decalactone and γ-decalactone titres C) OD_{600} after cultivation in M9-Fe media for 48 hrs at 30°C, 150rpm. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM methyl decanoate. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

5.2.15 Media supplementation with caffeine

It had been reported that addition of caffeine to the culture media increased BM3 activity[161]. Supplementation of the M9-Fe media with caffeine was therefore investigated to see if BM3

activity and therefore decalactone titres could be increased. *E. coli* expressing BM3 V78A F87A L188Q were cultured in M9-Fe supplemented with different concentrations of caffeine and fed methyl decanoate.

As the concentration of caffeine in the media increased, decanoic acid titres trended downwards (Fig 5.17A). However, as the caffeine concentration increased, there was no significant change in either δ-decalactone or γ-decalactone titres (Fig 5.17B). As caffeine concentration increased, OD_{600} trended downwards, indicating that there were some toxicity effects (Fig 5.17C).

Based on these results it was concluded that caffeine supplementation had no effect on decalactone titres, and so M9-Fe without caffeine was used going forwards.

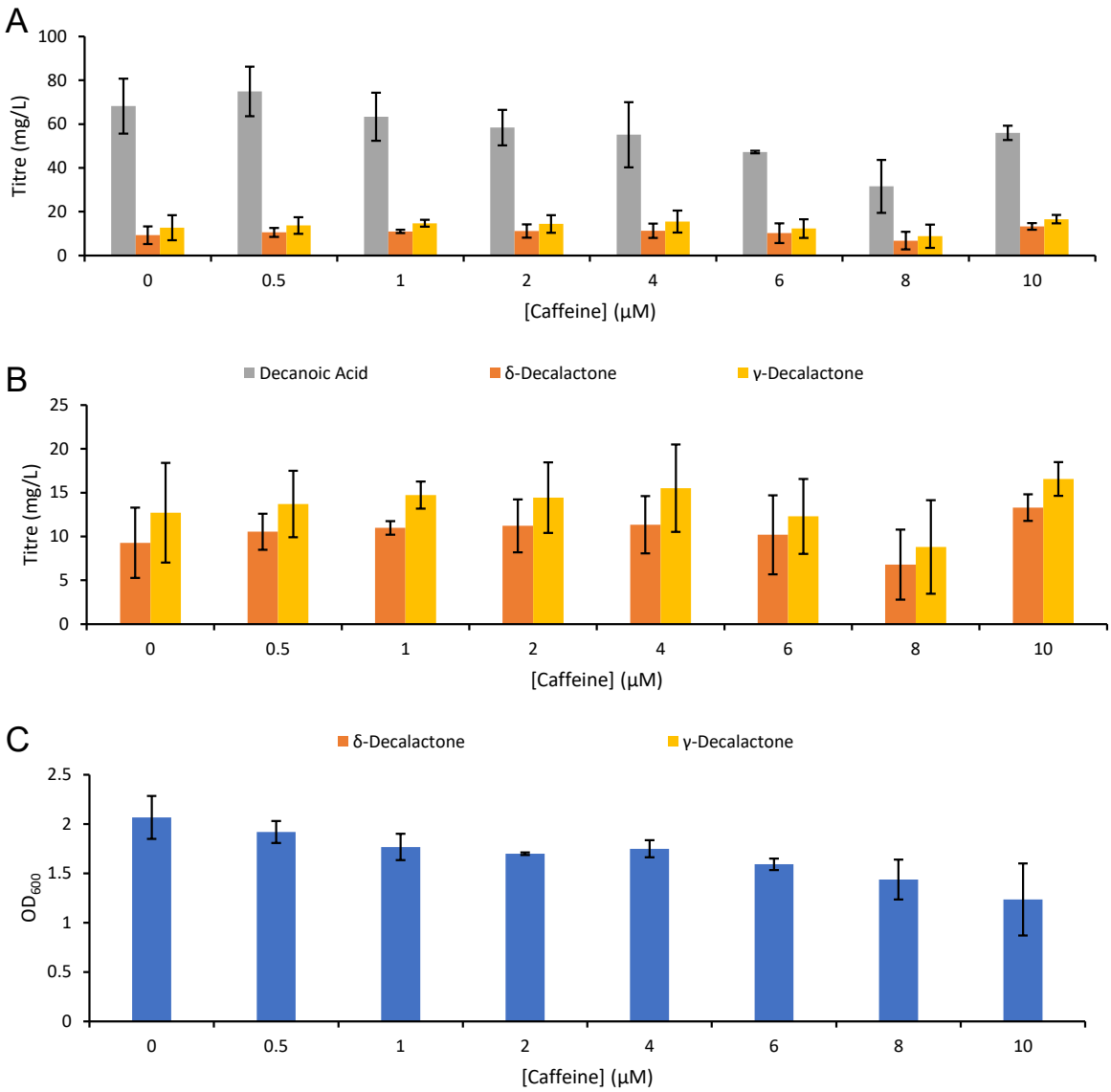


Figure 5.17 Supplementation of M9-Fe with caffeine *E. coli ΔfadD were transformed with a plasmid expressing BM3 V78A F87A L188Q. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) δ-decalactone and γ-decalactone titres C) OD_{600} after cultivation in M9-Fe media supplemented with varying concentrations of caffeine for 48 hrs at 30°C, 150rpm. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM methyl decanoate. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

5.2.16 Knock-out of *iscR* in *E. coli* ΔfadD

It had been observed in Section 5.2.8 that deletion of *iscR* in a BL21 background resulted in increased decalactone titres. It was therefore investigated whether ΔiscR could be combined with ΔfadD to improve decalactone titres. In parallel to work in Sections 5.2.9-15, iscR was knocked out of the *E. coli* ΔfadD strain that had been used for feeding experiments, to make a new ΔfadD ΔiscR strain. The effect of this knockout was then investigated by performing a

feeding experiment with expression of GFP, BM3 WT, BM3 V78A F87A L188Q and BM3 F87G L188Q in both the ΔfadD and ΔfadD ΔiscR backgrounds.

No significant differences in decanoic acid titres (Fig 5.18A), δ-decalactone titres, γ-decalactone titres (Fig 5.18B) or OD_{600} (Fig 5.18C) could be observed. From this it was concluded that ΔiscR did not have an effect on BM3 activity, and the ΔfadD strain was used going forward.

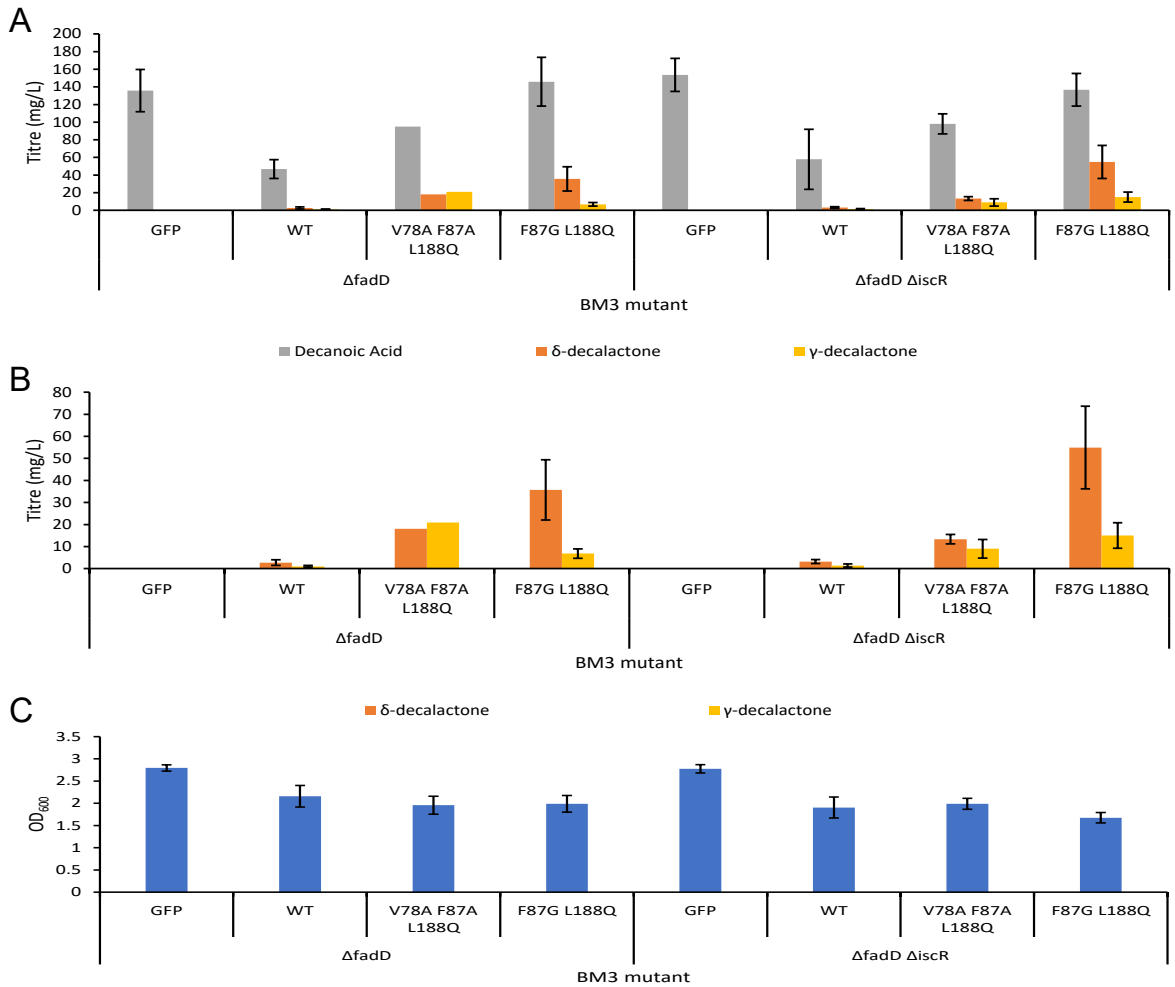


Figure 5.18 Screening knockout of ΔiscR in E. coli ΔfadD *E. coli ΔfadD and ΔfadD ΔiscR were transformed with a plasmid expressing either GFP, BM3 WT or a BM3 mutant. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) δ-decalactone and γ-decalactone titres C) OD_{600} after cultivation in M9-Fe media for 48 hrs at 30°C, 150rpm. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM methyl decanoate. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

5.2.17 Screening new BM3 activity enhancing mutations

Further mutagenesis of BM3 was investigated. Combinatorial mutants of BM3 V78A F87A L188Q and BM3 F87G L188Q with new activity enhancing mutations from the literature, D182S and H266T[426], and A328V[168] were generated and screened by a feeding assay.

Incorporating the D182S, H266T and A328V point mutations had no effect on decanoic acid titres (Fig 5.19A). When D182S, H226T or A328V were combined with BM3 F87G L188Q titres of both δ-decalactone and γ-decalactone fell significantly. On the other hand, combining D182S and H266T with BM3 V78A F87A L188Q had no significant effect on δ-decalactone and γ-decalactone titres. Combination of A328V with BM3 V78A F87A L188Q led to significantly higher γ-decalactone titres (30.5 mg/L, $p = 0.0327$) and an insignificant decrease in δ-decalactone titres (5.53 mg/L, $p = 0.0806$) (Fig 5.19B). Addition of D182S, H266T and A328V led to decreases in OD_{600} (Fig 5.19C).

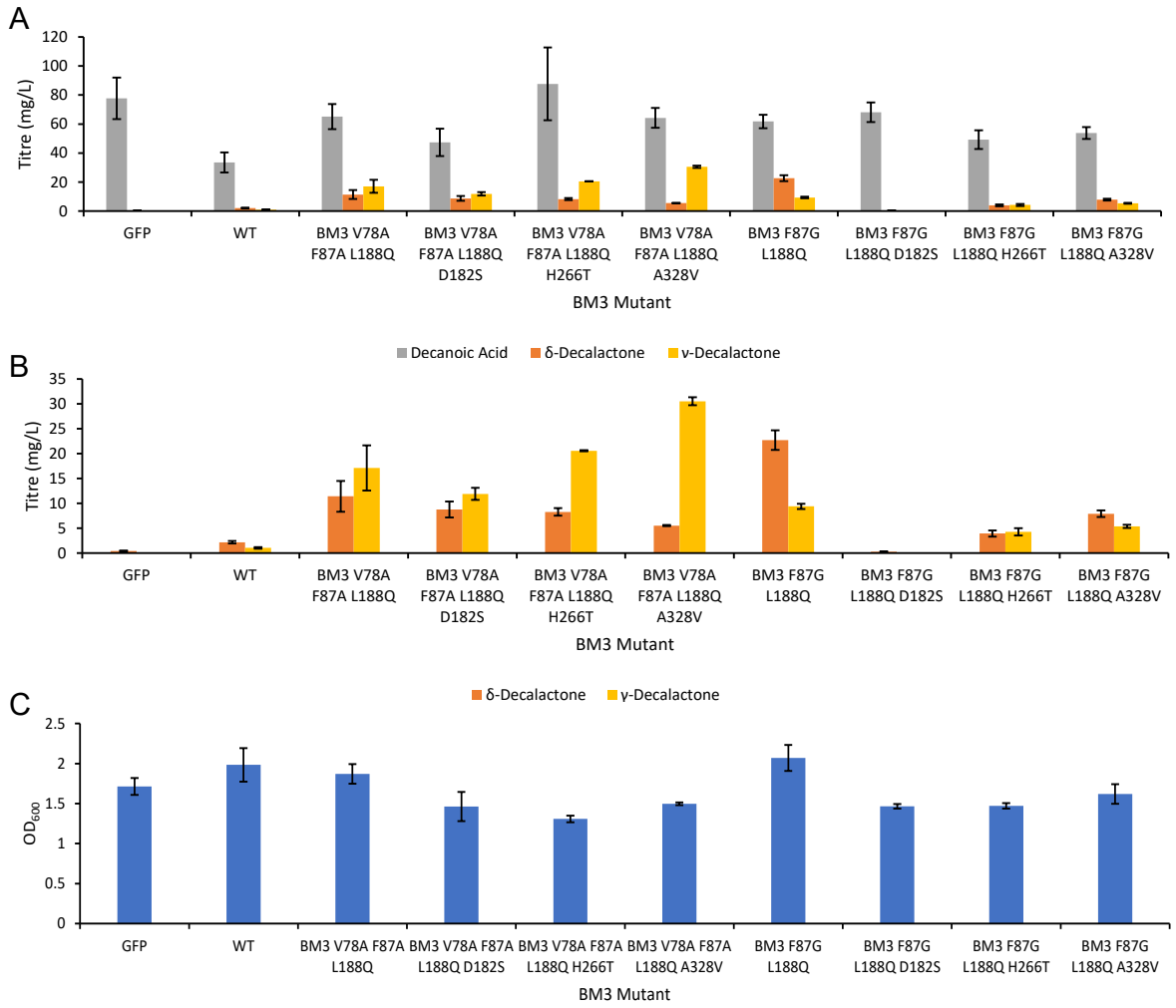


Figure 5.19 Screening new BM3 activity enhancing mutations *E. coli ΔfadD were transformed with a plasmid expressing either GFP, BM3 WT or a BM3 mutant. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) δ-decalactone and γ-decalactone titres C) OD_{600} after cultivation in M9-Fe media for 48 hrs at 30°C, 150rpm. Cells were induced with 0.5 mM IPTG and fed with 2.5 mM methyl decanoate. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

5.3 Discussion

5.3.1 The regioselectivity and substrate specificity of BM3

It was observed that feeding decanoic to the BM3 WT resulted in low amounts of δ-decalactone and γ-decalactone being produced. In previous work characterising the regioselectivity of BM3 and BM3 mutants when hydroxylating dodecanoic acid, hydroxylation beyond the ω-4 position was not observed[166]. However, to produce δ-decalactone and γ-decalactone, ω-5 and ω-6 hydroxylation is necessary, meaning that BM3 has a much wider native regioselectivity on decanoic acid. This is the first time that hydroxylation at these positions by non-engineered BM3 has been reported.

While BM3 was able to convert some of the decanoic acid, the majority was not hydroxylated. Hydroxylation of decanoic acid by BM3 has previously been reported, but BM3 is more active on longer chain length substrates[164,169]. Adding the activity enhancing mutation L188Q did increase decanoic acid consumption, but regioselectivity altering mutations were observed to and have been reported to decrease BM3 activity. While in this work combining multiple activity enhancing mutations was used to try and overcome this challenge, targeting production of dodecalactones may be an alternative solution. While this would also require the use of regioselectivity altering mutations, BM3 would have a higher native activity on the dodecanoic acid substrate, potentially leading to higher product titres. This idea formed the basis of an undergraduate research project under my supervision, where bioconversion of dodecanoic acid to and biosynthesis of dodecalactones was achieved.

5.3.2 The spontaneity of hydroxyfatty acid condensation

δ-Decalactone and γ-decalactone production was observed after feeding decanoic acid to BM3 or BM3 mutants, indicating that the 4- and 5-hydroxydecanoic acids did spontaneously condense to form the corresponding lactones. However, treating the samples with acid during extraction with chloroform-methanol was found to nearly double decalactone titres. These results suggested that not all the hydroxydecanoic acids spontaneously condensed to decalactones. The fact that acidification nearly doubled titres implies that the hydroxydecanoic acids and decalactones are in an equilibrium, and that acidification helps bias the equilibrium towards the decalactones, matching previously reported results in the literature for 4-hydroxypentanoic acid and γ-pentalactone[224]. However, it is not clear due to a lack of chemical standards if the acid treatment resulted in all the hydroxydecanoic acids condensing to decalactones, or if uncondensed hydroxydecanoic acids still remained. Only acidification using 150 µL 6 M H_2SO_4 was investigated, and in future work it should be investigated whether acidification with even stronger acidic conditions can further improve decalactone titres.

5.3.3 Product localisation

A key question with the feeding assays is whether the substrate can enter the cell and therefore be available to the BM3. When partitioning of the products was investigated, decanoic acid was found in both the cell and aqueous phase, with more decanoic acid found in the aqueous phase. This suggests that while some decanoic acid was able to enter the cell, the majority remained outside and had not been available to BM3, which would have limited decalactone titres. To overcome this challenge, a solution could be to produce decanoic acid within the cell using a C10 acyl-ACP thioesterase, where it would immediately be available to the BM3. A potential issue with this solution is that it is not clear whether efflux of the decanoic acid would occur faster than conversion of the decanoic acid by BM3. To overcome this, knock-out of fatty acid exporters could be considered, however it is not known how the decalactones are exported from the cell and so this could have a detrimental effect on decalactone export. Alternatively, *in vitro* conversion of decanoic acid by BM3 could be used to characterise BM3 mutants. Ultimately though, the mutants will need to be expressed and convert decanoic acid *in vivo*, and so biosynthesising the decanoic acid with a thioesterase would be a better next step.

A key question for commercial bioproduction is the localisation of the final product, in this case the decalactones. If the product remains within the cells, then processing will be required for its extraction, raising production costs. However, if the product is naturally excreted purification becomes simpler. Localisation experiments found that the decalactones were partitioned within the aqueous phase. This means that after hydroxylation of the decanoic acid and condensation, the decalactones were natively exported from the cell. Furthermore, in experiments with a decane overlay, decalactones were observed in the overlay, meaning that they had been extracted and purified out of the media. Titres were low, but this was probably due to the overlay also extracting the decanoic acid substrate. Overall, this native export of decalactones from the cell and the ability of an overlay to extract the decalactones from the aqueous phase is very promising when considering commercial bioproduction of decalactones, as it means fewer downstream processing steps are required.

5.3.4 Media optimisation

While media supplementation with iron and ALA had been reported in the literature before[202,205,206,425], it had not been systematically investigated whether this supplementation actually had an effect. This investigation was performed in this work, and it was found that only supplementation with iron led to increased activity.

M9 does not normally contain iron, which is a co-factor for the haem prosthetic group. BM3 activity without iron supplementation was observed, a potential explanation is that the cells already contained some iron from being pre-cultured in LB, and this was enough for some BM3 activity, but iron availability was a limiting factor. Upon supplementation, this limiting factor was removed, leading to increased activity. When trying to replicate the literature, the solubility of iron chloride proved to be an issue, and when iron chloride was added to M9 at the reported concentrations, it precipitated out of solution. The use of a chelating agent, in ferric ammonium citrate, helped overcome this problem, at the cost of not perfectly replicating the conditions used in the literature. In future work, it would be of interest to examine supplementing M9 with different ferric ammonium citrate concentrations, to identify the optimum for BM3 activity.

ALA is a precursor in the haem biosynthesis pathway. Supplementation with ALA did not have an effect on BM3 activity, suggesting that native haem biosynthesis was not a limiting factor. This negative result highlights the need to test assumptions, as the groups who reported its use did not report investigating whether ALA supplementation had an effect.

Caffeine had been reported to be an activator of BM3[161], so supplementation of M9 with caffeine was investigated to see if this could improve activity. However, no increases in activity were observed, and toxicity effects were observed. This contrasts previously reported results, however that work was done *in vitro*, and so toxicity would not have been a problem. Furthermore, in this work it was not examined if caffeine entered the cell or not, and so may not have been able to activate BM3. Even if this was the reason for the observed results, the outcome would still be that supplementing M9 with caffeine doesn't improve BM3 activity and so this idea was not pursued further.

5.3.5 Mutagenesis of BM3

It was observed that the regioselectivity altering mutations had different apparent regioselectivities than had been previously reported[166]. These mutations were characterised using dodecanoic acid, and so may have different regioselectivities on different chain length fatty acids, such as decanoic acid. These mutations resulted in activity losses, as had been reported in the literature.

Due to these activity losses, and the need to improve BM3's activity on decanoic acid, different activity enhancing mutations from the literature were also screened[168,169,177,426]. Three, L188Q, A328V and I401P were found to increase the activity of BM3 regioselectivity mutants. While L188Q could be combined with the different regioselectivity mutants, combining A328V and I401P with regioselectivity mutants sometimes increased decalactone titres and sometimes

decreased them. This suggests that combining certain mutations may have been destabilising to the protein, resulting in a loss of activity. Alternatively, mutations that alter the fatty acid substrate specificity have also been reported to affect the regioselectivity[167-170], and could have resulted in hydroxylation of decanoic acid at the incorrect positions. This was observed when A328V was combined with V78A F87A L188Q, where a shift from similar titres of δ-decalactone and γ-decalactone to increased specificity for the γ-decalactone was observed. Either way, the combinatorial mutants were not suitable for the bioconversion of decanoic acid to decalactones.

Through the mutagenesis assays, two combinatorial mutants were obtained, BM3 F87G L188Q and BM3 V78A F87A L188Q A328V that respectively specifically produced δ-decalactone and γ-decalactone. These mutants would allow tailoring of the proposed decalactone biosynthesis pathway to produce either the δ- or γ-decalactone.

5.3.6 Decalactone biosynthesis pathway extension

An extension to the pathway to allow for the enzymatic conversion of 4- and 5-hydroxydecanoic acid to γ- and δ-decalactone was investigated. Expression of Pp0763 led to a drop in decanoic acid titres. While Pp0763 was reported to be a 3-hydroxydecanoyl CoA ligase, PoAlkK, a homologue, was reported to be more active on non-hydroxylated substrates. Pp0763 could have therefore ligated the decanoic acid with CoA, leading to the observed drop in decanoic acid titres.

Due to the lack of analytical standards, 4- and 5-hydroxydecanoic acid titres had to be estimated, with estimated titres used to infer the activities of PoAlkK and Pp0763. This had limited success, with negative estimated titres. However, negative estimated titres were far more common for 4-hydroxydecanoic acid than 5-hydroxydecanoic acid, suggesting that PoAlkK and Pp0763 may have been more specific for 4-hydroxydecanoic acid, and ligated it with CoA. Furthermore, the only significant increase in decalactone titre observed was an increase in γ-decalactone titre when PoAlkK and PpAAT1 were co-expressed. This could have been because 4-hydroxydecanoyl-CoA was produced and then converted to the ester by PpAAT1, however it does not explain why expression of Pp0763 led to a decrease in estimated 4-hydroxydecanoic acid titres but additional expression of PpAAT1 did not significantly increase titres in this instance.

Since this experiment was performed, a new BM3 mutant, V78A F87A L188Q A328V was generated that produced increased γ-decalactone and therefore 4-hydroxydecanoic acid titres. It may be interesting to repeat this assay with this new mutant, to see if the greater production of 4-hydroxydecanoic acid helps clarify what is going on.

Availability of free CoA may also have been a limiting factor when the pathway extension was screened, and engineering *E. coli* to produce more free CoA could also be considered if this pathway is investigated in future work.

Ultimately, the lack of standards is a major hinderance to investigating this idea, and so it was not investigated further. Even if the pathway extension was functional, it would increase the number of heterologous genes required, increasing burden on the cell, which may have its own negative impacts on decalactone titre.

5.3.7 Co-factor balance

Improved co-factor supply was investigated, by using baffled shake flasks to improve culture oxygenation, and by using BmGDH1, BmGDH2[212] and GapN to increase NADPH supply. No changes to decalactone titre were observed, suggesting that co-factor supply was not an issue in these conditions.

However, during *de novo* biosynthesis of decalactones from glucose this may not be the case, particularly for NADPH supply. NADPH is a co-factor for FabG[68], which catalyses one of the steps in the fatty acid biosynthesis cycle. This means that when decalactones are being synthesised from glucose, there will be greater demand for NADPH, and NADPH may become a limiting factor. NADPH recycling using GDH could therefore be reinvestigated in future work once *de novo* biosynthesis of decalactones has been achieved.

5.4 Conclusions

In this chapter optimised media for the expression of BM3 was developed. This was the first time that optimisation of M9 media for BM3 expression had been systematically investigated.

Product partitioning was investigated, and decalactones were found to localise in the aqueous phase, from which they could be extracted using an overlay. To my knowledge this is the first time a solvent overlay has been shown to extract lactones from the culture media.

The spontaneity of hydroxydecanoic acid condensation to decalactones was investigated, and while some spontaneous condensation was observed, it was found that treatment with acid could be used to catalyse the condensation. While acidic conditions had been reported to catalyse the condensation of hydroxyfatty acids to lactones[224], this was the first-time acidification had been shown to improve the condensation of 4- and 5-hydroxydecanoic acid to γ- and δ-decalactone.

An extension to the pathway that would allow for enzymatic conversion of hydroxyfatty acids to decalactones was investigated, but due to a lack of analytical standards results were inconclusive.

Co-factor supply was investigated and found to not be a limiting factor for BM3 activity.

Through multiple rounds of mutagenesis combinatorial BM3 mutants that were able to specifically convert decanoic acid to δ- and γ-decalactone were obtained. This was the first time that BM3 has been engineered to convert fatty acids to lactones, and the first time that a CYP450 has been used to convert a fatty acid to a γ-lactone. A promising alternative CYP450 to BM3 from the literature was also investigated[98], along with a homologue, however they were found to have very low activity. The combinatorial BM3 mutants developed in this work therefore also represent an improvement in the state-of-the-art in the conversion of decanoic acid to δ-decalactone.

Chapter 6

De novo Biosynthesis of δ-Decalactone and γ-Decalactone

6.1 Aims

γ-Decalactone and δ-decalactone are commercially important high impact and value flavour and fragrance compounds[15,79,84,87]. The native pathway for their biosynthesis is unknown[87], and prior work in the literature has focused on the biotransformation of long chain fatty acids to medium chain length γ- and δ-lactones[81,83,85,88-95].

In this work a novel metabolic pathway for the biosynthesis of γ- and δ-decalactone has been proposed. The pathway consists of two engineered enzymes. Firstly, a C10 specific acyl-ACP thioesterase produces free decanoic acid from endogenous acyl-ACP. In Chapter 4 engineering was performed to generate C10 specific acyl-ACP thioesterases for the pathway. Secondly, an engineered BM3 mutant hydroxylates the decanoic acid at the 4- and 5-positions, and the 4- and 5-hydroxydecanoic acid are then condensed to respectively produce the γ-decalactone and δ-decalactone. In Chapter 5 engineering of BM3 to carry out these hydroxylation reactions was performed.

The aim of this chapter is to combine these enzymes to construct a novel synthetic pathway for the *de novo* biosynthesis of decalactones from glucose, without requiring the feeding of an additional substrate. The pathway will first be expressed and optimised in *Escherichia coli* to confirm its functionality.

Bioproduction using cyanobacteria as a host could allow for the renewable production of chemicals from CO_2 and sunlight[11], however this is currently commercially inviable due to high costs[2] and low product values and titres[150]. Decalactones are high value products[15,84], and so could help commercialise this technology. Therefore, after the pathway is validated in *E. coli* it will be ported to the cyanobacterium *Synechocystis* Sp. PCC 6803, to confirm the pathway is also functional in this heterologous host.

Some work in this chapter was done in parallel with work in chapter 5.

6.2 Results

6.2.1 *De novo* biosynthesis of decalactones in *E. coli*

The synthetic pathway for δ- and γ-decalactone biosynthesis was expressed (Fig 6.1A). The mutant thioesterase CvT5 W122F from Chapter 4 (Section 4.2.8) was used to biosynthesise decanoic acid. While CvT5 W122F was the not the most C10 specific thioesterase engineered in Chapter 4, its expression led to the highest decanol titres. BM3 has very low activity on octanoic acid[164], meaning off-target activity of BM3 on any octanoic acid that was also produced would not be an issue. Therefore, it was decided that using a thioesterase that produced more decanoic acid over a thioesterase that specifically produced decanoic acid would lead to the highest decalactone titres. CvT5 W122F was expressed on plasmid using the Plac promoter (Fig 6.1B), and a second plasmid was used to co-express a BM3 mutant and EcACP using the PrhaBAD promoter (Fig 6.1C). This allowed the expression of CvT5 W122F to be induced using 75 μM IPTG, which had been found to be the optimal concentration in Section 4.2.8, without affecting expression of the BM3 mutant or EcACP. CvT5 W122F was co-expressed in *E. coli* with either the BM3 wild-type (BM3 WT), BM3 F87A L188Q I401P or BM3 V78A F87A L188Q, as these were the best mutants that had been engineered so far in Chapter 5 (Section 5.2.9). Finally, EcACP was also co-expressed as this had been show in Chapter 4 (Section 4.2.5) to increase product titres. As a control instead of EcACP and a BM3 mutant, GFP was co-expressed with CvT5 W122F.

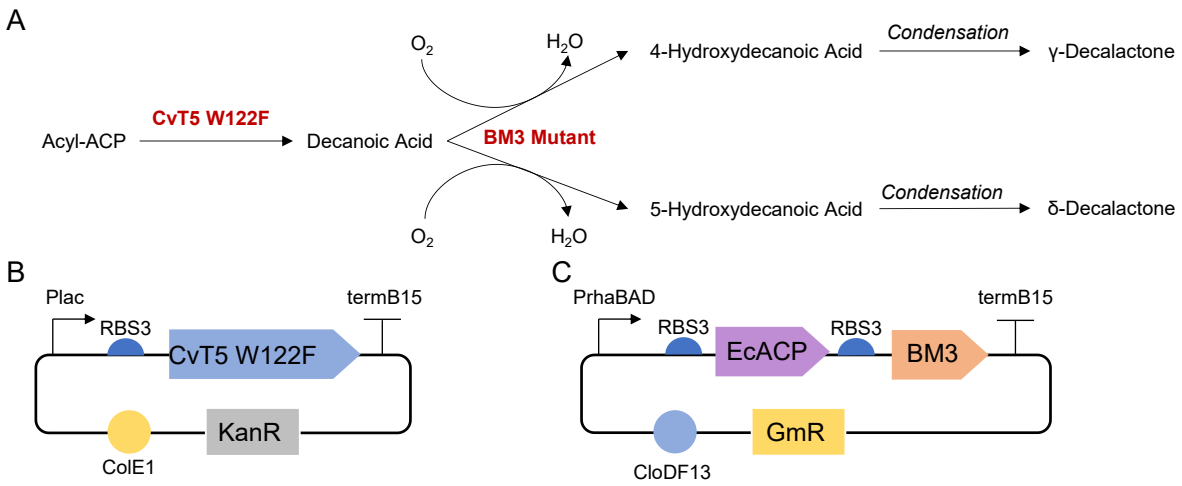


Figure 6.1 Pathway and plasmid design for δ- and γ-decalactone biosynthesis *A) Synthetic metabolic pathway for decalactone biosynthesis. Pathway enzymes are highlighted in red. CvT5 W122F: Cuphea viscosissima FatB1 acyl-ACP thioesterase truncated at M112 and with W122F mutation. BM3 Mutant: CYP450 CYP102A1 (BM3) (Priestia megaterium) with combinatorial mutations detailed in Chapter 5. B) Plasmid design for biosynthesis of decanoic acid. A pET plasmid with ColE1 origin of replication and kanamycin resistance cassette was used Plac and RBS3 were used to express a CvT5 W122F. C) Plasmid design for hydroxylation of decanoic acid at the 4- and 5-positions. A pCDF plasmid with CloDF13 origin of replication and spectinomycin resistance cassette was used. Plac and RBS3 were used to express EcACP and BM3 variants in a single operon.*

Titres of decanoic acid were lower than expected, at only 22.7 mg/L in the control expressing CvT5 W122F and GFP. Based on the results of Section 4.2.8, decanoic acid titres of around 200 mg/L had been predicted. Due to large variation within the control, there was no significant differences in decanoic acid titres between strains expressing GFP or BM3 mutants (Fig 7.2A).

Very low titres of both δ-decalactone (0.46 mg/L) and γ-decalactone (0.45 mg/L) were observed when the BM3 WT was expressed. Expression of BM3 F87A L88Q led to no γ-decalactone biosynthesis, and a decrease in δ-decalactone titres (0.17 mg/L). Expression of BM3 V78A F87A L188Q led to a small but significant ($p = 0.0284$) increase in γ-decalactone titres (0.72 mg/L) (Fig 7.2B). The BM3 mutants' relative activities and specificities did not match those observed in Section 5.2.9.

Reduced growth was observed when the different BM3 variants were expressed instead of GFP (Fig 7.2C).

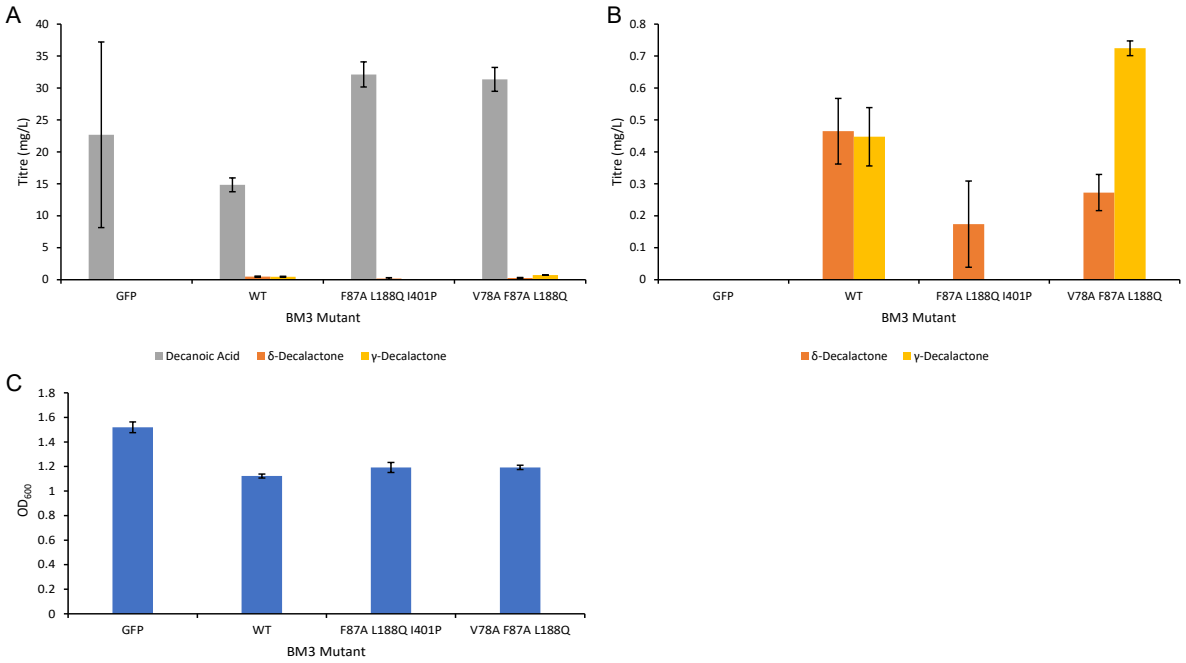


Figure 6.2 De novo biosynthesis of decalactones *E. coli ΔfadD were transformed with a plasmid expressing CvT5 W122F and a second plasmid expressing either GFP or EcACP and a BM3 mutant. A) Decanoic acid, δ-decalactone and γ-decalactone titres. B) δ-decalactone and γ-decalactone titres. C) OD_{600} after cultivation in M9-Fe media for 48 hrs at 30°C. Cells were induced with 75 μM IPTG and 0.1% rhamnose. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

6.2.2 Was ferric ammonium citrate detrimental to thioesterase activity?

When the thioesterases were screened in Chapter 4 M9 media was used. However, in Section 6.2.1 M9-Fe media was used to culture the *E. coli*, as this media had been shown to improve BM3 activity in Chapter 5 (Section 5.2.5). A possible explanation for the low decanoic acid titres observed in the control strain in Section 6.2.1 was that the ferric ammonium citrate was detrimental to thioesterase activity. Another explanation was that co-transformation with the second plasmid used to express GFP or EcACP and a BM3 variant was having a negative impact on titres, which had been observed elsewhere in this work (Section 7.2.4). This was not considered likely, as this backbone had been used to express EcACP, Sfp and CAR in Chapter 4 without any problems, but was still investigated in order to rule it out. *E. coli* were either transformed with only the plasmid for expression of CvT5 W122 from before, or co-transformed with that plasmid and the second plasmid from Section 6.2.1, expressing GFP. The two different strains where then cultured in either M9 or M9-Fe media.

The lowest decanoic acid titres (71.95 mg/L) were observed when *E. coli* were transformed with only the plasmid for expression of CvT5 W122F and cultured in M9. Higher decanoic acid

titres were observed for the other different combinations of culture conditions and plasmids used for transformation, with no significant differences between the combinations. Decanoic acid titres (114.0 mg/L) in *E. coli* co-transformed to express CvT5 W122F and GFP and cultured in M9-Fe were much higher than observed in the control in Section 6.2.1, which used the same *E. coli* strain and culture conditions, suggesting there may have been an issue when this experiment was performed. These data suggested there was no negative effect on thioesterase activity from ferric ammonium citrate or the second plasmid.

However, titres were still lower than expected based on results in Chapter 4. No significant differences in growth were observed between the different combinations (Fig 6.3B).

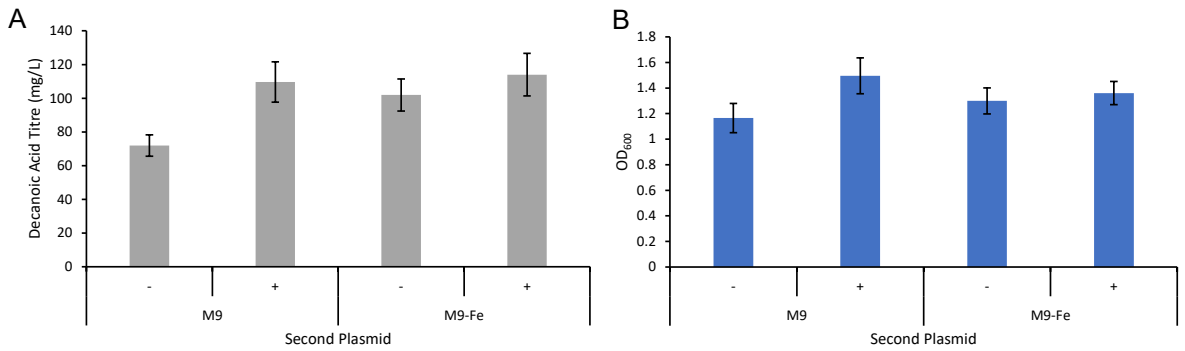


Figure 6.3 Effect of ferric ammonium citrate and a second plasmid on decanoic acid titres *E. coli ΔfadD were transformed with a plasmid expressing CvT5 W122F, or a plasmid expressing CvT5 W122F and a second plasmid expressing GFP. A) Decanoic acid titres. B) OD_{600} after cultivation in M9 or M9-Fe media for 48 hrs at 30°C. Cells were induced with 75 μM IPTG and 0.1% rhamnose. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

6.2.3 Redesign of plasmids for expressing pathway enzymes and RBS optimisation of CvT5 W122F expression

The plasmids were redesigned so that all genes were expressed using the Plac promoter, meaning only a single inducer would be required. EcACP and CvT5 W122F were expressed on one plasmid (Fig 6.4A), and BM3 mutants on a second (Fig 6.4B). In Chapter 4, inducer optimisation of CvT5 W122F had been shown to improve product titres. As inducer optimisation was no longer possible without also affecting the expression of EcACP and BM3, CvT5 W122F expression was optimised by varying RBS strength, which has previously been shown to modulate protein expression and product titres[427]. The Biolegio linker kit included RBSs with different relative expression levels, and a set of these RBSs was chosen for screening (Table 6.1).

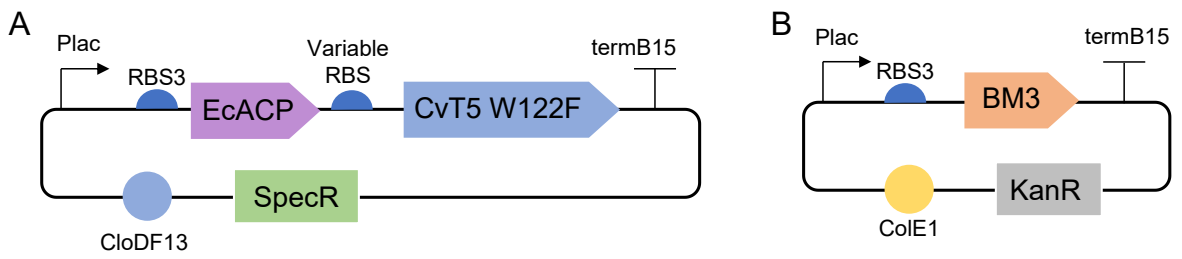


Figure 6.4 New plasmid designs for δ- and γ-decalactone biosynthesis *A) Plasmid design for biosynthesis of decanoic acid. A pCDF plasmid with CloDF13 origin of replication and spectinomycin resistance cassette was used. Plac, RBS3 and a variable RBS were used to express EcACP and CvT5 W122F in a single operon. B) Plasmid design for hydroxylation of decanoic acid at the 4- and 5-positions. A pET plasmid with ColE1 origin of replication and kanamycin resistance cassette was used. Plac and RBS3 were used to express a BM3 variant.*

RBS	Relative Expression (%)
3	100
A07	69
2	56
A03	46
A08	35
A01	23
A09	14
1	5

Table 6.1 RBSs used for gene expression optimisation *Expression of the gene controlled by the RBS is given relative to expression with RBS3. Data provided by Biolegio in their Basic linker kit.*

E. coli were co-transformed with plasmids expressing EcACP and CvT5 W122F using different RBSs, and a plasmid expressing GFP instead of BM3. The highest decanoic acid titres (33 mg/L) were observed when CvT5 W122F was expressed using RBSA07 (Fig 6.5A). This was still lower than the expected titre of ~ 200 mg/L. No significant changes in OD_{600} were observed when CvT5 W122F was expressed with different RBSs (Fig 6.5B).

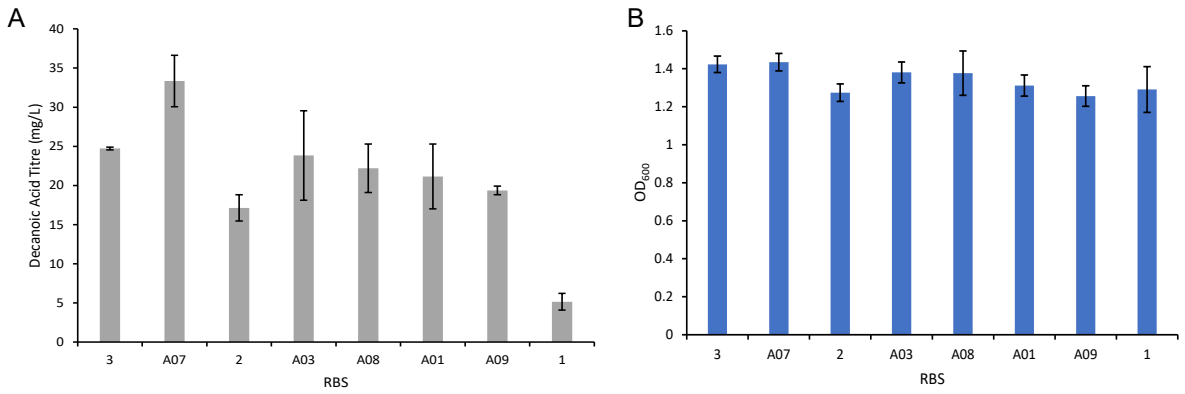


Figure 6.5 RBS Optimisation of CvT5 W122F expression *E. coli ΔfadD were transformed with a plasmid expressing EcACP and CvT5 W122F expressed using different RBSs and a second plasmid expressing GFP. A) Decanoic acid titres. B) OD_{600} after cultivation in M9-Fe media for 48 hrs at 30°C. Cells were induced with 0.5 mM IPTG. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

6.2.4 *De novo* biosynthesis using new BM3 mutants and screening the effect of a decane overlay

At this stage, new BM3 mutants had been generated in Chapter 5 (Section 5.2.17), which in feeding experiments produced higher titres of both δ-decalactone and γ-decalactone. *De novo* biosynthesis of the decalactones was therefore attempted again using these new mutants and the plasmid designs in Section 6.2.3. In parallel, the effect of a decane overlay during *de novo* biosynthesis was investigated. *E. coli* were co-transformed with a plasmid expressing RBS3-EcACP-RBSA07-CvT5 W122F operon, and a second plasmid expressing either the BM3 WT, BM3 V78A F87A L188Q, BM3 F87G L188Q, BM3 V78A F87A L188Q A328V or as a control GFP.

In the absence of an overlay, decanoic acid titres of 31.8 mg/L were observed in the culture media phase when GFP was expressed on the second plasmid, consistent with the results in Section 6.2.3. When a decane overlay was present, decanoic acid titres in the culture media phase dropped to 2.71 mg/L. In the absence of an overlay, decreased decanoic acid titres in the culture media phase were observed when a BM3 mutant was expressed on the second plasmid. When BM3 WT was expressed on the second plasmid, only trace amounts of decanoic acid were observed (0.27 mg/L). In the presence of an overlay decanoic acid titres in the culture media phase sharply decreased (Fig 6.6A), like when GFP was expressed on the second plasmid.

No decalactones were observed in the culture media phase when either GFP or the BM3 Wild-Type were expressed on the second plasmid. In the absence of an overlay, the highest

γ-decalactone titres (3.53 mg/L) were observed when BM3 V78A F87A L188Q A328V was expressed on the second plasmid. δ-Decalactone titres were lower, reaching only 0.53 mg/L when BM3 F87G L188Q was expressed on the second plasmid. The BM3 V78A F87A L188Q and F87G L188Q appeared to have altered regioselectivities compared to the feeding experiments, shifting from respectively producing roughly equal amounts of δ- and γ-decalactone, and producing more δ-decalactone to producing more γ-decalactone and roughly equal amounts of δ- and γ-decalactone. In the presence of an overlay, there was also a sharp decrease in decalactone titres in the culture media phase. (Fig 6.6B).

In the presence of an overlay, decanoic acid was observed in the overlay. When GFP was expressed on the second plasmid, decanoic acid titres in the overlay were 34.0 mg/L, comparable to decanoic acid titres in the culture media phase in the absence of an overlay. Decanoic acid was also observed in the overlay when a BM3 mutant was expressed on the second plasmid, and titres were higher than in the culture media phase in the absence of an overlay (Fig 6.6C). Due to an issue with washing the ALS injector needle after decalactone standards were run on the GC, it could not be determined if any decalactones were present in the overlay.

Expression of the BM3 WT, V78A F87A L188Q and F87G L188Q led to improved growth relative to the GFP control. Expression of the V78A F87A L188Q A238V mutant had no significant effect on growth. There was no significant difference in growth when an overlay was used (Fig 6.6D).

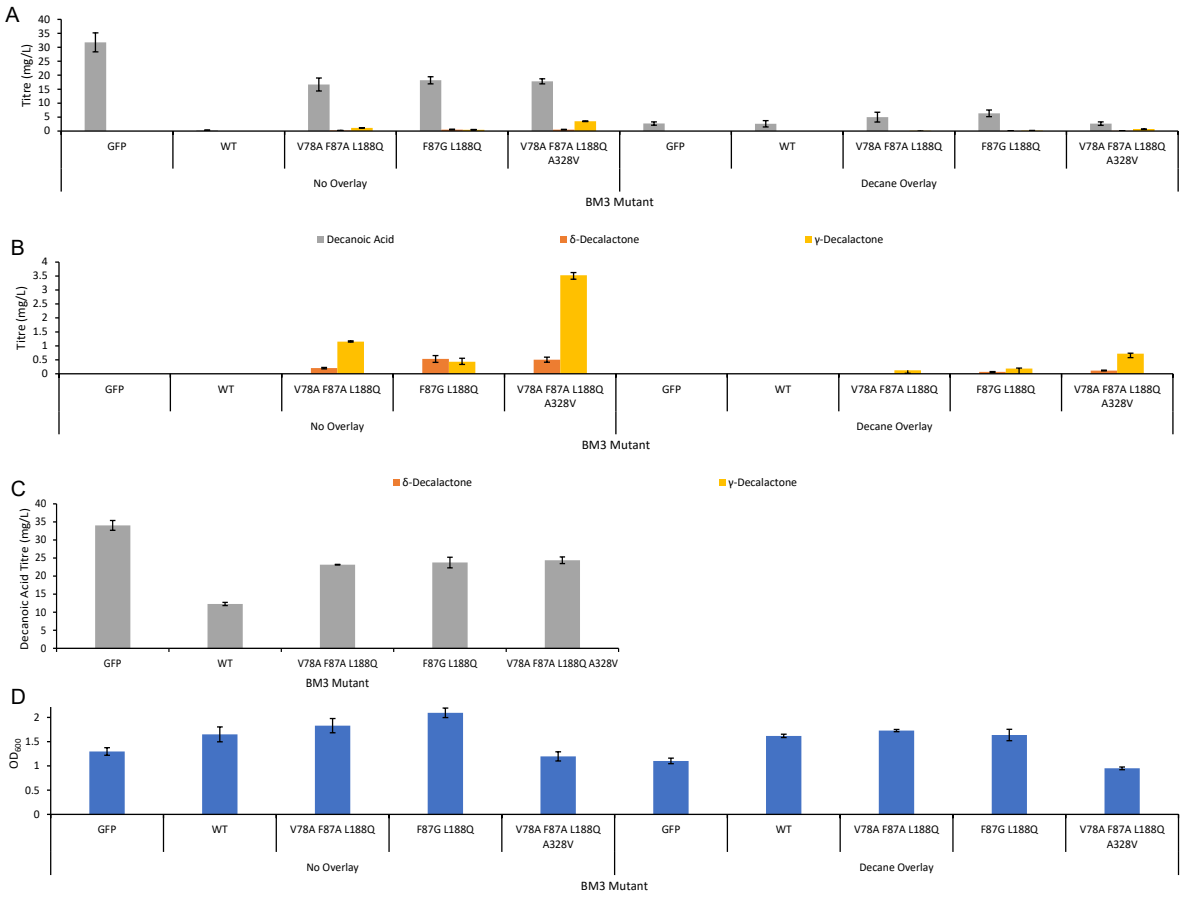


Figure 6.6 Biosynthesis of decalactones using new BM3 mutants with and without a decane overlay *E. coli ΔfadD were transformed with a plasmid expressing EcACP and CvT5 W122F and a second plasmid expressing GFP or a BM3 mutant. A) Decanoic acid, δ-decalactone and γ-decalactone titres in the culture media phase. B) δ-decalactone and γ-decalactone titres in the culture media phase. C) Decanoic acid titres in the overlay phase. D) OD_{600} after cultivation in M9-Fe media with and without a 10% decane overlay for 48 hrs at 30°C. Cells were induced with 0.5 mM IPTG. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. The decane overlay was sampled directly and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

6.2.5 Screening hybrid CYP450s

Two new CYP450s, RbCYP450 and TtCYP450 had been investigated in Chapter 5, on the basis of TtCYP450's reported specificity for hydroxylation of decanoic acid at the ω-5 position[98]. However, no activity could be observed (Sections 5.2.7 and 5.2.8). The hydroxylation of decanoic acid is catalysed by the haem domain, and it was theorised that fusing the haem domains of the new CYP450s to the reductase domain of BM3 could generate active hybrid CYP450s that would perform the desired hydroxylation reaction. Two hybrid CYP450s were

generated, RbxBM3 (haem domain of RbCYP450 fused to reductase domain of BM3) and TtxBM3 (haem domain of TtCYP450 fused to reductase domain of BM3). The hybrid CYP450s were then screened by incorporation into the decalactone biosynthesis pathway and expressed using the same constructs as in Section 6.2.4.

Decanoic acid titres for *E. coli* expressing GFP, BM3 WT and BM3 mutants were similar to those observed in Section 6.2.4. There was no significant difference in decanoic acid titres between *E. coli* expressing GFP (32.3 mg/L) and RbxBM3 (25.9 mg/l, p = 0.0921) or TtxBM3 (29.5 mg/L, p = 0.507) (Fig 6.7A).

δ-decalactone and γ-decalactone titres for *E. coli* expressing BM3 mutants were consistent with those observed in Section 6.2.4. Low titres of δ- and γ-decalactone were observed when the BM3 Wild-Type was expressed. No decalactones could be observed when RbxBM3 or TtxBM3 were expressed (Fig 6.7B).

OD_{600} in *E. coli* expressing GFP, BM3 WT and BM3 mutants was also consistent with those observed in Section 6.2.4. No significant differences in OD_{600} were observed between *E. coli* expressing GFP and RbxBM3 or TtxBM3 (Fig 6.7C).

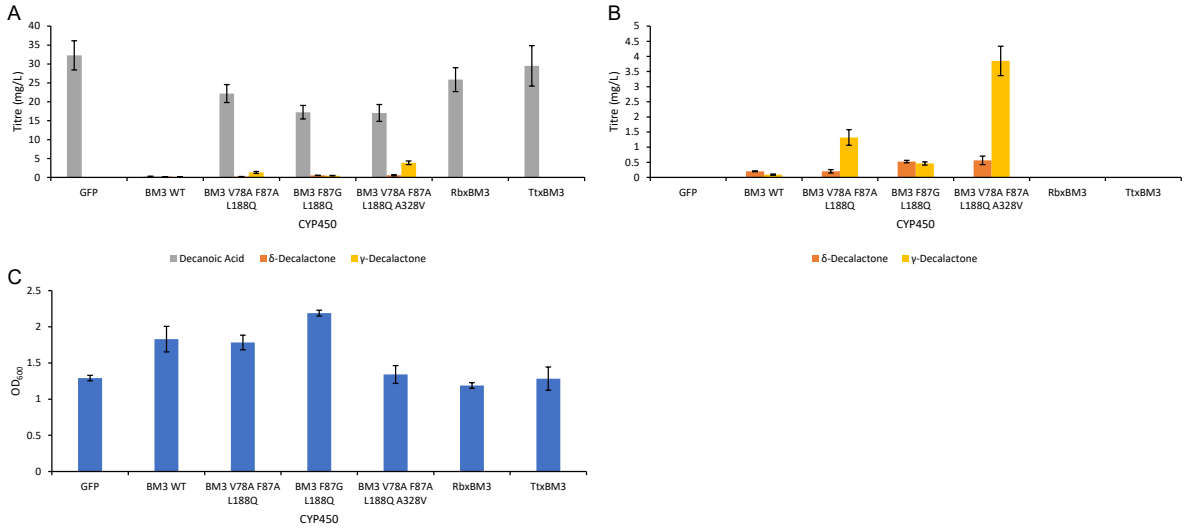


Figure 6.7 Screening hybrid CYP450s *E. coli ΔfadD were transformed with a plasmid expressing EcACP and CvT5 W122F and a second plasmid expressing GFP or a CYP450 A) Decanoic acid, δ-decalactone and γ-decalactone titres. B) δ-decalactone and γ-decalactone titres. C) OD_{600} after cultivation in M9-Fe media for 48 hrs at 30°C. Cells were induced with 0.5 mM IPTG. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

6.2.6 New operon order for EcACP and CvT5 W122F expression

In prior work by the Jones lab, it was observed that gene order within an operon had an effect on product titres[153]. It was therefore decided to investigate this, by flipping the gene order so that CvT5 W122F was expressed before EcACP (Fig 6.8). RBS optimisation of CvT5 W122F expression was also repeated in combination with the new gene order.

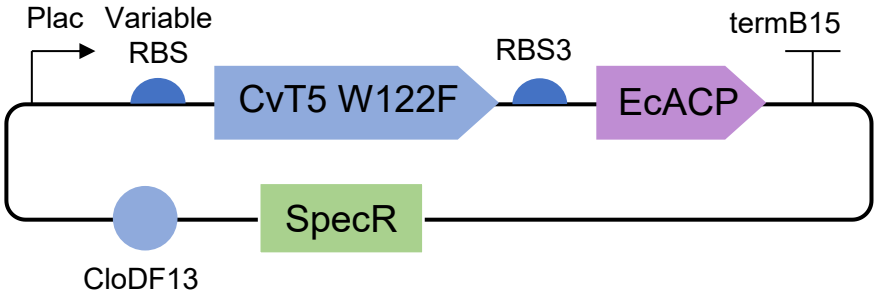


Figure 6.8 New plasmid design for expression of CvT5 W122F and EcACP *A pCDF plasmid with CloDF13 origin of replication and spectinomycin resistance cassette was used. Plac, a variable RBS and RBS3 were used to express CvT5 W122F and EcACP in a single operon.*

E. coli were co-transformed with different plasmids to express CvT5 W122F and EcACP with the different RBSs, and a second plasmid expressing GFP. As a control, *E. coli* was co-transformed with the plasmid expressing the old RBS3-EcACP-RBSA07-CvT5 W122F operon, and a second plasmid expressing GFP.

The highest decanoic acid titres (44.7 mg/L) were observed when CvT5 W122F was expressed using RBS3 and the new operon order. However, this was not significantly ($p = 0.632$) different to decanoic acid titres (41.7 mg/L) when CvT5 W122F was expressed using the old operon order and RBSA07 (Fig 6.9A). No significant differences in OD_{600} were observed when CvT5 W122F was expressed with different RBSs (Fig 6.9B).

The old operon RBS3-EcACP-RBSA07-CvT5 W122F was therefore used going forward.

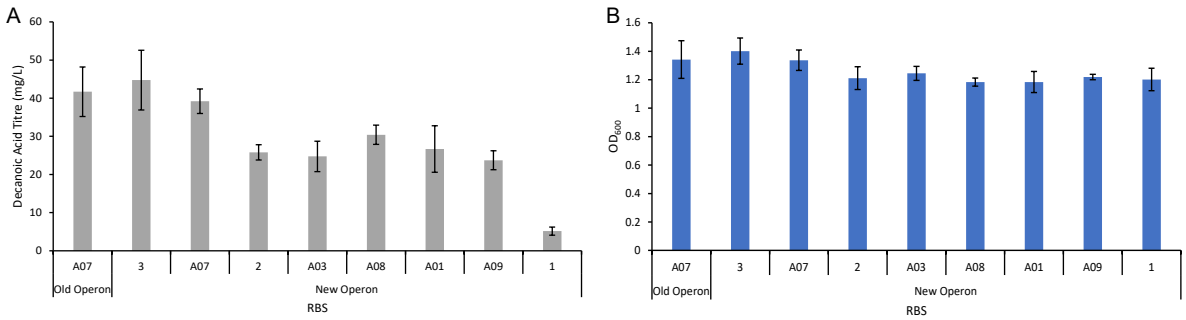


Figure 6.9 RBS Optimisation of CvT5 W122F expression with the new operon order *E. coli ΔfadD were transformed with a plasmid expressing CvT5 W122F expressed using different RBSs and EcACP and a second plasmid expressing GFP. A) Decanoic acid titres. B) OD_{600} after cultivation in M9-Fe media for 48 hrs at 30°C. Cells were induced with 0.5 mM IPTG. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

6.2.7 Media supplementation with Cas amino acids

In previous work by the Jones lab, it had been observed that supplementing M9 media with cas amino acids improved *E. coli* growth and product titres (unpublished data). *E. coli* expressing EcACP-RBSA07-CvT5 W122F and different BM3 mutants were therefore cultured in M9-Fe with and without supplementation with 0.05% cas amino acids.

Increased decanoic acid titres were observed when the M9 was supplemented with cas amino acids (Fig 6.10A). However, the effects of supplementation with cas amino acids upon decalactone titres were less clear. When the BM3 WT was expressed, supplementation with cas amino acids led to small increases in both δ- and γ-decalactone titre. When BM3 V78A F87A L188Q and F87G L188Q were expressed, no significant changes in decalactone titres were observed. When BM3 V78A F87A L188Q A328V was expressed, a significant ($p = 0.0152$) decrease in γ-decalactone titres was observed (Fig 6.10B). Slightly higher growth was observed when the M9 was supplemented with cas amino acids (Fig 6.10C).

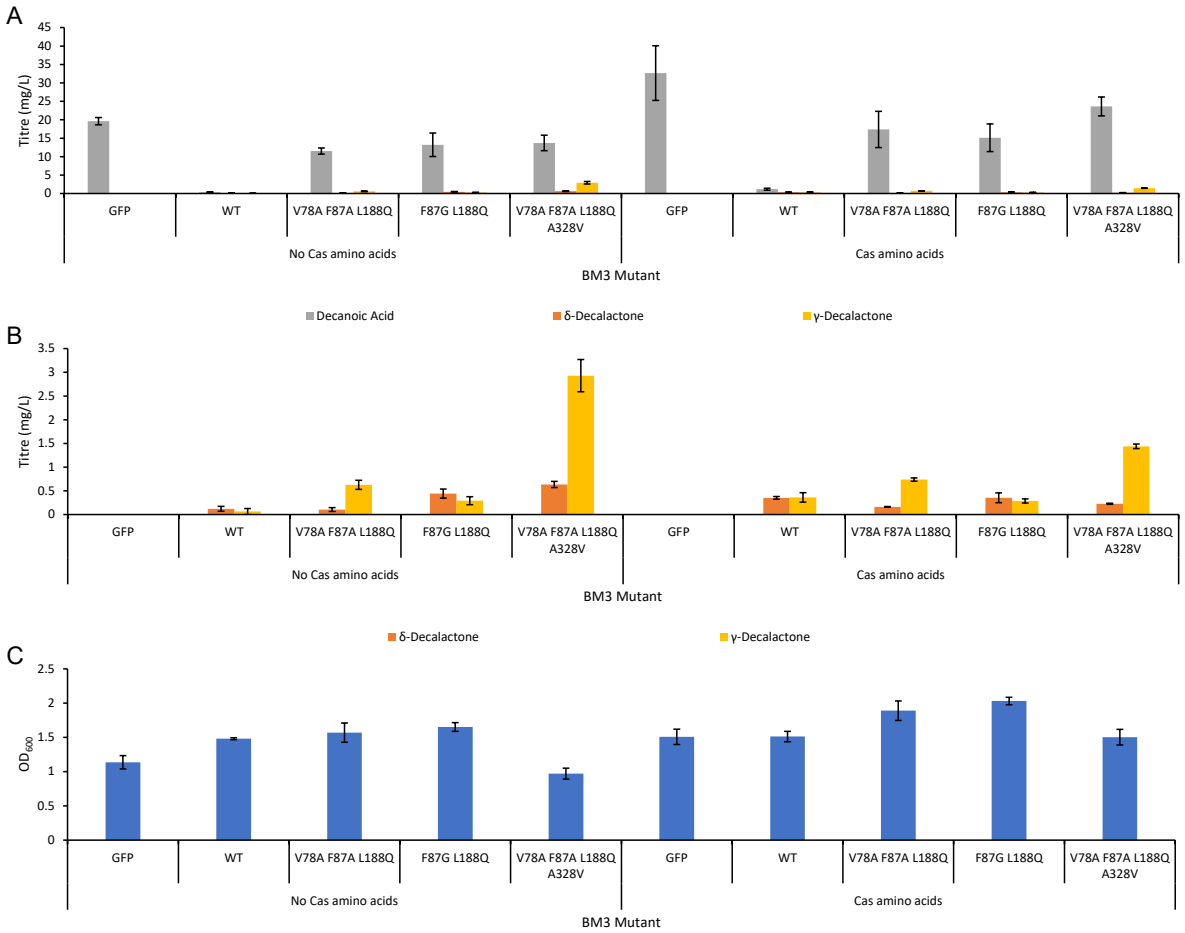


Figure 6.10 Supplementing M9 with cas amino acids *E. coli ΔfadD were transformed with a plasmid expressing EcACP-RBSA07-CvT5 W122F and a second plasmid expressing GFP or a BM3 variant. A) Decanoic acid, δ-decalactone and γ-decalactone titres. B) δ-decalactone and γ-decalactone titres. C) OD_{600} after cultivation in M9-Fe media with and without 0.05% cas amino acids for 48 hrs at 30°C. Cells were induced with 0.5 mM IPTG. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

6.2.8 RBS optimisation of EcACP expression

It was investigated whether RBS optimisation of EcACP expression could be used to improve decanoic acid titres. The operon EcACP-RBSA07-CvT5 W122F was used, and the RBS used to express EcACP varied. *E. coli* were co-transformed with different plasmids for expressing EcACP with the different RBSs and CvT5 W122F, and a second plasmid expressing GFP.

The highest decanoic acid titres (14.0 mg/L) were observed when EcACP was expressed using RBS3 (Fig 6.11A). No significant differences in OD_{600} were observed when EcACP was expressed using different RBSs (Fig 6.11B).

These decanoic acid titres were significantly lower than had previously been observed, even when RBS3 was used to express EcACP. These apparent low titres are likely due to an issue with the GC-MS, as integrated peak areas for both the samples and standards were abnormally high, especially for the standards. Nevertheless, from these results it could be concluded that the optimal RBS for expressing EcACP was RBS3, which was the RBS already being used.

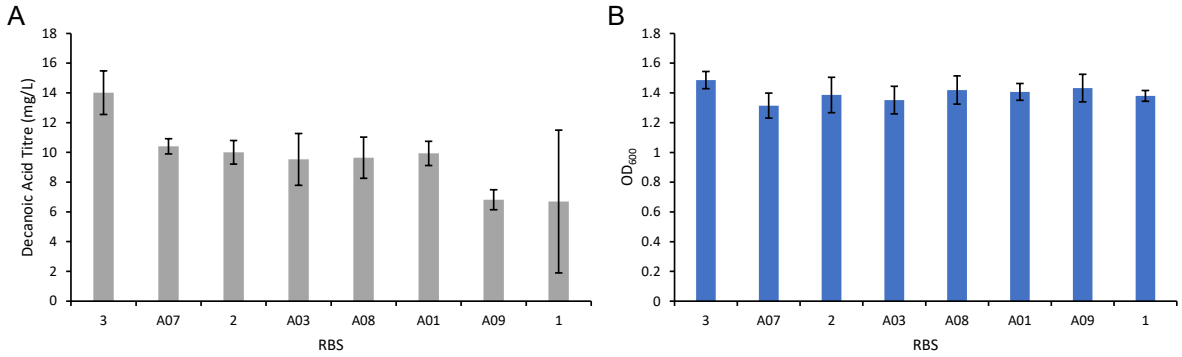


Figure 6.11 RBS optimisation of EcACP expression *E. coli ΔfadD were transformed with a plasmid expressing EcACP expressed using different RBSs and CvT5 W122F and a second plasmid expressing GFP. A) Decanoic acid titres. B) OD_{600} after cultivation in M9-Fe media for 48 hrs at 30°C. Cells were induced with 0.5 mM IPTG. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

6.2.9 RBS optimisation of BM3 expression

It was investigated whether RBS optimisation of BM3 expression could be used to improve decalactone titres. *E. coli* were co-transformed with a plasmid expressing RBS3-EcACP-RBSA07-CvT5 W122F and a second plasmid expressing either BM3 F87G L188Q or BM3 V78A F87A L188Q A328V expressed using different RBSs.

As the RBS used to express a BM3 mutant weakened, decanoic acid titres increased. Decanoic acid titres observed when RBS3 was used to express BM3 F87G L188Q (39.2 mg/L) and BM3 V78A F87A L188Q A328V (96.1 mg/L) were higher than previously observed in Sections 6.2.4, 6.2.5 and 6.2.7, when the same constructs were used to express those BM3 mutants (Fig 6.12A).

As the strength of the RBS used to express BM3 F87G L188Q decreased, there was initially no significant effect on decalactone titres. As RBS strength further weakened, first δ-decalactone and then γ-decalactone biosynthesis could no longer be observed. As the strength of the RBS used to express BM3 V78A F87A L188Q A328V weakened there was an immediate and significant decrease in δ-decalactone and γ-decalactone titres. As RBS

strength further weakened biosynthesis of δ-decalactone could no longer be observed. Low titres of γ-decalactone could be observed for every RBS used (Fig 6.12B).

Varying RBS strength had no significant effect on OD_{600} for either BM3 mutant (Fig 6.12C).

While the results when the BM3 mutants were expressed using RBS3 were inconsistent with previous results, it could be observed that using a weaker RBS to express a BM3 mutant led to higher decanoic acid titres and lower decalactone titres. These data therefore indicated that the optimal RBS for expression of BM3 mutants was RBS3, which was already being used.

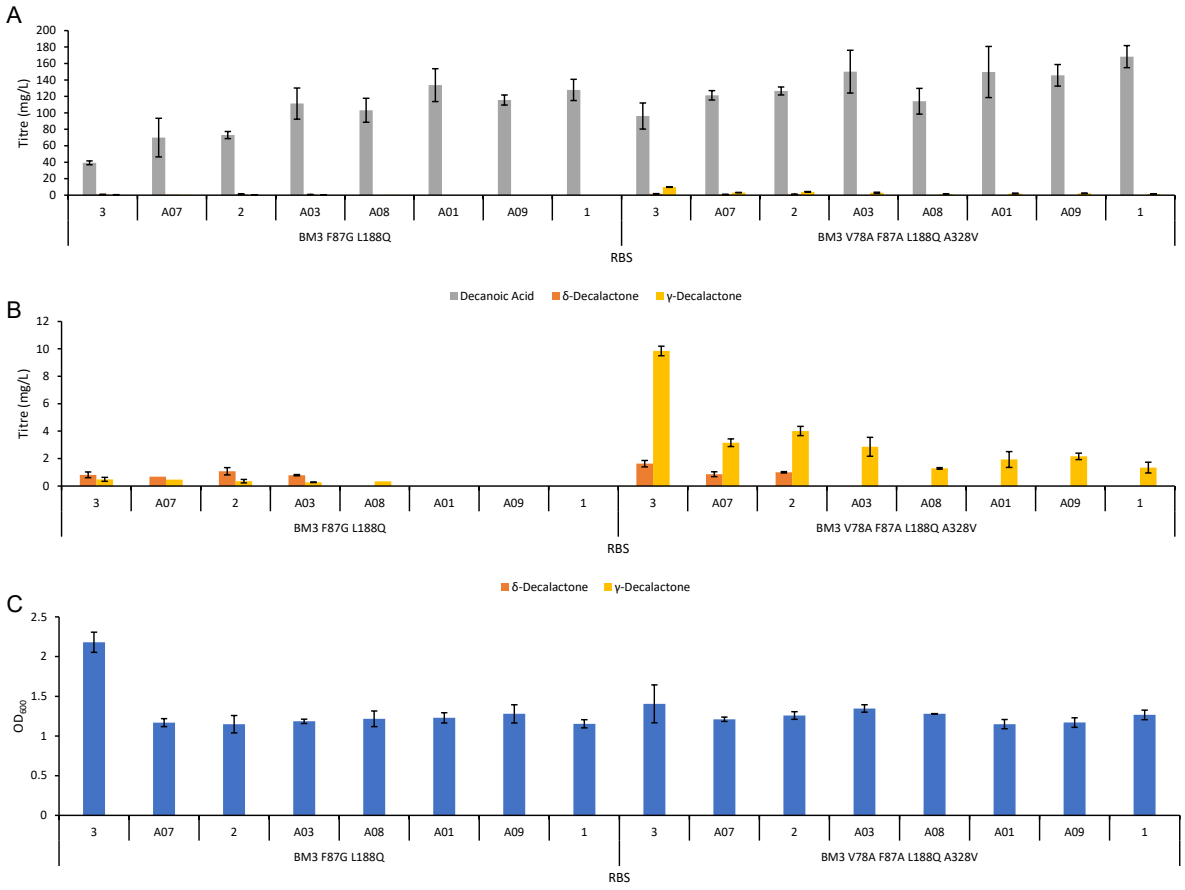


Figure 6.12 RBS optimisation of BM3 expression *E. coli ΔfadD were transformed with a plasmid expressing RBS3-EcACP-RBSA07-CvT5 W122F and a second plasmid expressing either BM3 F87G L188Q or BM3 V78A F87A L188Q A328V, expressed using different RBS strengths. A) Decanoic acid, δ-decalactone and γ-decalactone titres. B) δ-decalactone and γ-decalactone titres. C)OD_{600} after cultivation in M9-Fe media for 48 hrs at 30°C. Cells were induced with 0.5 mM IPTG. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

While attempts to improve decanoic acid and decalactone titres in *E. coli* in Sections 6.2.6-9 had not been successful, *de novo* biosynthesis of decalactones had been achieved in *E. coli*

in Section 6.2.4. It was therefore decided to use the remaining time in the project to attempt *de novo* biosynthesis of decalactones in *Synechocystis*.

6.2.10 Biosynthesis of decalactones in *Synechocystis* Sp. PCC 6803

The decalactone biosynthesis pathway was assayed in *Synechocystis*. A Δaas strain was used to prevent recycling of free fatty acids. The thioesterase CvT5 was used, as its expression led to the highest decanol titres in Chapter 4 (Section 4.2.3). CvT5 was either expressed individually on a plasmid as a control (Fig 6.13A), or co-expressed with a BM3 variant (Fig 6.13B). Expression of the BM3 Wild-Type, BM3 F87G L188Q and BM3 V78A F87A L188Q A328V was investigated. During triparental conjugation, growth of the transformant *Synechocystis* strains was poor. After nearly a month of incubation of the BG11-Co-Agar plates in an Algaetron, only small colonies were observed instead of the expected lawn of cyanobacteria. Nevertheless, the cyanobacteria were successfully transferred to liquid media and precultures generated to attempt biosynthesis of decalactones.

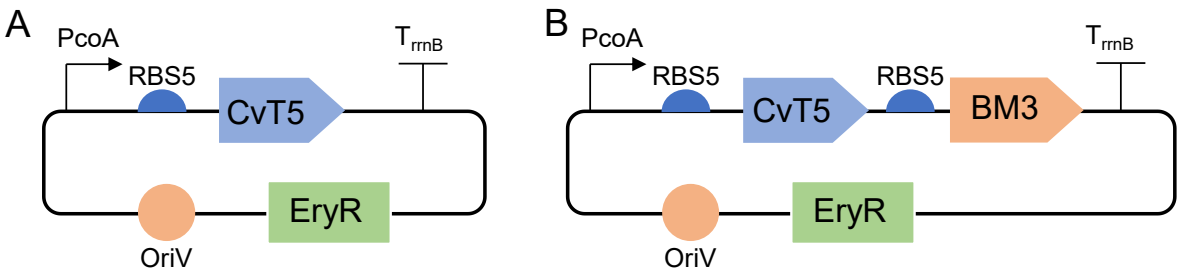


Figure 6.13 Plasmid design for expression of CvT5 and BM3 in Synechocystis *A pRSF1010 plasmid with OriV origin of replication and erythromycin resistance cassette was used. PcoA and RBS5 were used to express A) CvT5 or B) CvT5 and BM3 in a single operon.*

No product formation was observed after 5 days cultivation. After 10 days low titres of decanoic acid were observed, and no decalactone biosynthesis was observed. When only CvT5 was expressed decanoic acid titres reached 0.71 mg/L. This was much lower than previously observed decanol titres (112.4 mg/L) when CvT5 was previously characterised in *Synechocystis* (Section 4.2.3). Co-expression of a BM3 variant with CvT5 did not have significant effects on decanoic acid titre (Fig 6.14A).

After 10 days cultivation there was no significant difference in OD_{730} between the different strains (Fig 6.14B). Observed OD_{730}s were much lower (OD_{730} = 0.4-1.2) than those observed in Section 4.2.3 when CvT5 was characterised in *Synechocystis*. (OD_{730} = 2.1).

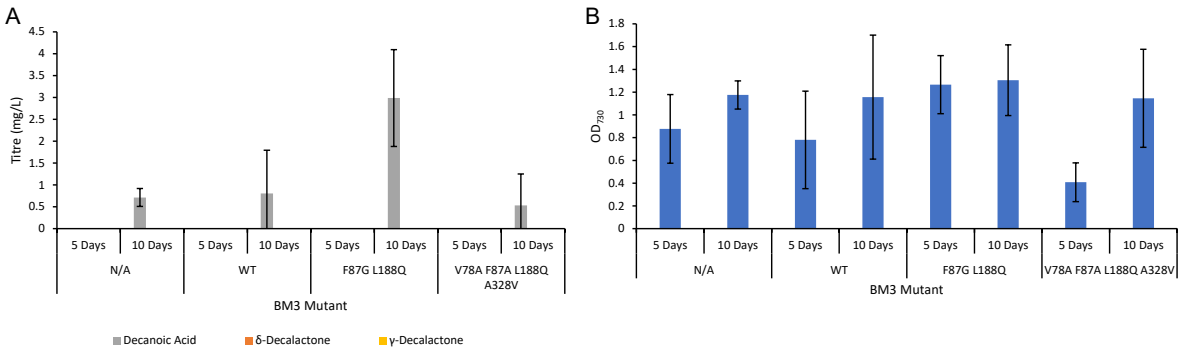


Figure 6.14 De novo biosynthesis of decalactones in Synechocystis *Synechocystis Δaas was transformed with a plasmid expressing either CvT5 or co-expressing CvT5 and a BM3 variant. A) Decanoic acid, δ-decalactone and γ-decalactone titres B) OD_{730} after cultivation in BG11-Co media for 10 Days at 30°C. Cells were induced with 5 μM Cobalt. Products were extracted from the media using chloroform-methanol and sulfuric acid and analysed by GC-MS. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

6.3 Discussion

6.3.1 *De novo* biosynthesis of decalactones in *E. coli*

In this chapter *de novo* biosynthesis of both δ-decalactone and γ-decalactone was achieved in *E. coli*, with titres in Section 6.2.4 reaching 0.53 mg/L and 3.53 mg/L. While low, this result demonstrates that the novel metabolic pathway designed in this work is functional. To my knowledge, *de novo* biosynthesis of lactones has not been reported in *E. coli* and cyanobacteria.

De novo biosynthesis of γ-lactones has been reported in *Ashbya gossypii*, and total lactone titres were also low, reaching ~3 mg/L. In that instance, a range of γ-lactones of different chain lengths were biosynthesised, whereas in this work γ- and δ-lactones of the same chain length were biosynthesised. To biosynthesis the γ-lactones in that report, oleic acid was first biosynthesised and then fed into the β-oxidation cycle to produce shorter acyl-CoA of the appropriate chain length. It was then unclear how the shortened fatty acids were hydroxylated and condensed to lactones[90].

The γ-decalactones titres achieved in this work were therefore higher than those previously reported, although it was not a big improvement. More importantly, the system in this work is better understood – biosynthesis and consumption of the decanoic acid could be observed, concurrent to the biosynthesis of decalactones. Furthermore, by using different BM3 mutants the output of the pathway could be tailored to production of a mix of δ- and γ-decalactones, or to the more specific production of γ-decalactone. In future work, by using different thioesterases and different BM3 mutants the pathway could be tailored to the specific biosynthesis of different chain length δ- and γ-decalactones. In fact, an undergraduate student under my supervision was successfully able to biosynthesise δ- and γ-dodecalactones using a C12 thioesterase available in the Jones lab and BM3 mutants developed during this work and from the literature (Jones lab, unpublished data). While titres achieved in this work were low, two clear limiting factors could be identified: lower than expected decanoic acid biosynthesis, and incomplete consumption of decanoic acid by the BM3 mutants. These will be further discussed below, but if they could be overcome there is clear scope to improve decalactone titres.

Overall, the pathway developed in this work is functional and shows promise for the biosynthesis of different range δ- and γ-decalactones. Further work will be required to improve product titres.

6.3.2 Low decanoic acid titres

Decanoic acid titres in Sections 6.2.1 and 6.2.2 were lower than expected. From the results in Chapter 4 (Section 4.2.8), decanoic acid titres of ~200 mg/L had been predicted. In Section 6.2.2 decanoic acid titres of ~100 mg/L were achieved, and it was concluded that there were still issues with decanoic acid biosynthesis. However, in these experiment CvT5 W122F was expressed using the same plasmid as in Chapter 4, and in the controls GFP was expressed on the second plasmid. When CvT5 W122F was characterised in Chapter 4, it was co-expressed with EcACP, Sfp and CAR. Co-expression of EcACP with an acyl-ACP thioesterase had been observed in Chapter 4 to increase product titres by ~1.5-2 fold. In retrospect, it is not surprising that decanoic acid titres in Section 6.2.2 were lower than expected, as the expectation included co-expression of EcACP, which did not occur.

The experiment in Section 6.2.2 should be repeated, with the following changes. Instead of co-expressing CvT5 W122F and GFP, CvT5 W122F should be co-expressed with EcACP. This should account for any negative impacts on titre due to lack of EcACP expression. Along with EcACP, Sfp and CAR were expressed when CvT5 W122F was characterised in Section 4.2.8. This could have led to increased flux through the alcohol biosynthesis pathway, resulting in higher alcohol titres. Furthermore, an isopropyl myristate overlay was used when CvT5 W122F was characterised, which could have further increased flux through the pathway by extracting the alcohols from the culture media. To account for these factors, *E. coli* co-expressing CvT5 W122F, EcACP, Sfp and CAR, as in Section 4.2.8, cultured with and without an overlay should be used as controls. This would determine whether these factors lead to higher product titres, and therefore unrealistic expectations for decanoic acid biosynthesis. Biosynthesis of decanoic acid without these factors could then be determined from the *E. coli* strain co-expressing CvT5 W122F and EcACP.

Decanoic acid titres in Section 6.2.1 were much lower than in Section 6.2.2. Upon reviewing the peak areas for the standards used in Section 6.2.2, it was noticed that the integrated peak areas for the standards were smaller than normal, which would have led to higher calculated decanoic acid titres, causing inconsistencies between the two experiments. Once the experiment in Section 6.2.2 has been repeated as discussed above, the experiment in Section 6.2.1 should also be repeated, using the *E. coli* strain co-expressing CvT5 W122F and EcACP as the new control. These repeats would clarify how much decanoic acid is biosynthesised, and how much is then converted to the δ-decalactone and γ-decalactone using these construct designs. Additionally, the more active BM3 mutants that were unavailable at the time could also be expressed instead of the less active mutants used when the experiment was originally performed.

The plasmids used to express the pathway were then redesigned so that CvT5 W122F and EcACP were expressed using a single plasmid, and that both plasmids would be induced using Plac. Despite efforts to optimise the EcACP-CvT5 W122F operon, including varying RBS strength and swapping the gene order, decanoic acid titres remained low, at about ~40 mg/L, lower than in Section 6.2.2. These results suggest that co-expression of EcACP and CvT5 W122F in an operon was not as effective as expressing the two genes separately.

RBS optimisation of EcACP expression within the EcACP-RBSA07-CvT5 W122F operon was also investigated in Section 6.2.8. Decanoic acid titres were lower than had previously been observed, even when the original RBS, RBS3, was used to express EcACP, due to abnormal integrated peak areas on the GC-MS for both the samples and the standards. Despite this challenge, from the data it could be concluded that RBS3 was the optimal RBS for expressing EcACP. A repeat of this experiment without these analytical issues should be performed to lend more certainty to this conclusion.

A challenge with RBS optimisation is that varying the strength of an upstream RBS can also affect the expression of downstream genes. It has previously been reported that increasing the strength of an upstream RBS led to accumulation of both proteins in an operon, with abundance of the downstream protein correlated to abundance of the upstream protein[427]. This means that using RBS strength to optimise expression of EcACP and CvT5 W122F may not have been an optimal approach. When the RBS used to express CvT5 W122F was varied in Section 6.2.3, the upstream RBS3 used to express EcACP would have also led to expression of CvT5 W122F, minimising the effect of varying the RBS in front of CvT5 W122F. Conversely, when the upstream RBS was varied, as in Sections 6.2.6 and 6.2.8, this would have affected expression of both genes, when the aim was to modulate expression of only gene. To investigate this potential effect, abundance of EcACP and CvT5 W122F when expressed with different RBSs could be quantified. To identify the optimal RBSs for expressing the entire operon, a systematic investigation of each possible pair of RBSs should be performed.

In this work, different construct designs were investigated, however the two designs were not directly compared. Once Section 6.2.2 has been repeated, the *E. coli* co-expressing CvT5 W122F and EcACP on two different plasmids should be screened alongside *E. coli* expressing the RBS3-EcACP-RBSA07-CvT5 W122F operon, to determine which construct design is optimal for decanoic acid biosynthesis.

6.3.3 Conversion of decanoic acid to decalactones

A limiting factor in the biosynthesis of decalactones was the conversion of decanoic acid to decalactones by BM3. From the feeding experiments in Chapter 5, it was not clear how much of the fed substrate was being consumed, as it was not known exactly how much entered the cell and was available to BM3. However, in these *de novo* biosynthesis experiments, the decanoic acid was produced in the cell where it would be immediately available to the BM3. In Section 6.2.4, biosynthesis of decanoic acid was observed in the GFP control (31.8 mg/L), however upon expression of the BM3 WT almost none was observed (0.27 mg/L). These data shows that BM3 WT was capable of consuming the decanoic acid that was produced. However, as had been reported in the literature[166], and observed in Chapter 4, mutating BM3 for altered regioselectivity resulted in reduced activity. Decanoic acid could be observed when combinatorial BM3 mutants incorporating regioselectivity altering and activity enhancing mutations were expressed, indicating that even though the mutants had activity enhancing mutations, there was still an overall activity loss compared to the BM3 WT.

When BM3 V78A F87A L188Q A328V was expressed, decanoic acid titres of 17.8 mg/L were observed. From this it can be estimated that approximately only 44.0% (14.0 mg/L) of the available decanoic acid substrate was converted by the BM3 mutant to the various hydroxydecanoic acids. Decanoic acid titres may also have been lower due to increased burden from expressing the BM3 mutant. To control for this, an inactive BM3 mutant could be expressed instead of GFP as a control. Nevertheless, the fact that decanoic acid could be observed indicates there was incomplete consumption of the substrate.

Furthermore, δ-decalactone and γ-decalactone titres were 0.51 mg/L and 3.53 mg/L, respectively, indicating that only 3.6% and 25.2% of the decanoic acid was hydroxylated at the 5- and 4- positions, assuming all of the 4- and 5-hydroxydecanoic acid was condensed to γ- and δ-decalactone during the cell extraction and sample acidification process, and that no product loss occurred. This means that the majority of hydroxylation performed by the BM3 mutant was off-target.

RBS optimisation of BM3 expression was investigated in Section 6.2.9. Even when the BM3 mutants were expressed with RBS3, as in Section 6.2.4, titres of decanoic acid, δ-decalactone and γ-decalactone were substantially higher than had previously been observed. Again, abnormally high integrated peak areas were observed for both the standards and the samples, leading to higher apparent product titres. From the data, it could be concluded that the RBS originally used to express the BM3 mutants, RBS3, was optimal from BM3 expression, however repeating this experiment to overcome the analytical issues would provide more clarity.

There is therefore a need for further work on engineering BM3 for increased activity and improved regioselectivity. As discussed in Chapter 5, GDH overexpression could be attempted again in parallel with *de novo* biosynthesis, as NADPH may be a limiting factor in this context but not in the feeding experiments. A challenge with further engineering is that few mutations that alter the regioselectivity of hydroxylation on fatty acids have been reported[166], and the most promising were already explored in this work. Furthermore, further altering the regioselectivity may result in further activity losses.

Finding a different CYP450 to use instead of BM3 will also be challenging, as the most promising CYP450 for δ-decalactone biosynthesis from the literature, TtCYP450, which was reported to specifically hydroxylated decanoic acid at the ω-5 position[98], was found to have very low activity in Chapter 5 (Section 5.2.7). Attempts to optimise the host background for TtCYP450 expression (Section 5.2.8), or to generate active hybrids of TtCYP450 and BM3 were unsuccessful (Section 6.2.5). When TtCYP450 was initially characterised, it was initially expressed in *E. coli* BL21 DE3 using a pET28a backbone before being extracted and characterised *in vitro*[98,205]. In Section 5.2.8, TtCYP450 was also expressed in BL21, using a pET backbone, so this is not likely to have been an issue. TtCYP450 was characterised alongside a panel of other new CYP450s, and BM3 was used by the authors as a control. They reported that TtCYP450 converted 83% of the decanoic acid substrate, while BM3 converted >99% of the substrate. These data indicate that BM3 is more active than TtCYP450, but as BM3 converted all the substrate, it cannot be determined by how much BM3 is more active. An explanation for the low δ-decalactone titres observed when expressing TtCYP450 could therefore be that there was not an issue expressing TtCYP450 *in vivo*, but that it has very low activity compared to BM3. Fusing the haem domain of TtCYP450 with the reductase domain of BM3 did not lead to improved activity, suggesting that the haem domain of TtCYP450 is the cause of TtCYP450's low activity. TtCYP450 has also been reported to hydroxylate aromatic substrates *in vitro*[205], and so may have poor substrate specificity for aliphatic acids.

Recently a high-throughput colourimetric assay for the biosynthesis of aldehydes and ketones was reported, and the assay could also be used to detect γ-butyrolactone[428]. This assay is therefore likely to be able to detect other δ- and γ-lactones. This assay could be paired with random mutagenesis of BM3 to screen a large library of mutants, with the goal of identifying highly-active BM3 mutants that can convert fatty acids to the corresponding δ- and γ-lactones.

6.3.4 Biosynthesis of decalactones in *Synechocystis*

Biosynthesis of decalactones was not observed when the pathway was expressed in *Synechocystis*. However, this is likely not due to the pathway not being functional in

Synechocystis. Instead, the most likely cause is the very low biosynthesis of decanoic acid. From the experiments in *E. coli*, it can be seen that only a minority of decanoic acid is converted to decalactones, and so when decanoic acid titres are so low it is not surprising that no decalactones were observed. While low decanoic acid titres could be due to an issue with the thioesterase, this is unlikely as the same plasmid was used to biosynthesise octanol and decanol in *Synechocystis* Δaas with Sfp-CAR integrated into the genome in Chapter 4 (Section 4.2.3), with observed decanol titres of 112.4 mg/L, as opposed to the decanoic acid titres of 0.71 mg/L observed in Section 6.2.10. Again, the expression of Sfp-CAR and an isopropyl myristate overlay in Chapter 4 could have increased flux, leading to higher product titres, however in prior work by the Jones lab free fatty acid titres and alcohol titres were similar when either only a thioesterase or a thioesterase, Sfp and CAR were expressed in *Synechocystis* Δaas[150]. On the other hand, in prior work also by the Jones lab, alcohol titres were higher than fatty acid titres when a thioesterase was expressed in *Syenchocystis* Δaas with Sfp-CAR integrated into the genome then when the thioesterase was expressed in *Syenchocystis* Δaas without Sfp-CAR[128]. The extent of this effect depended on the thioesterase used, however there was never a case where alcohol biosynthesis, but not fatty acid biosynthesis was observed. This suggests that the low titres were not caused by low activity of the thioesterase.

The most likely explanation for the low titres is the poor growth of the transformant strains. After 10 days cultivation, observed OD_{730}s in Section 6.2.10 were low, ranging from 0.4-1.2, whereas in Section 4.2.3 an OD_{730} of 2.1 was achieved. In prior work by the Jones group higher OD_{730}s of up to 6 have also been observed[13,153].

During triparental conjugation, colonies took longer than normal to develop, and during the decalactone biosynthesis experiment much lower OD_{730}s were observed than in the alcohol biosynthesis experiment in Chapter 4. While toxicity from the decanoic acid or decalactones could be a cause of this poor growth, at such low concentrations this is unlikely. Triparental conjugation has successfully been used to generate transformant *Synechocystis* strains both in this work and in prior work by the Jones group, however the growth problems in this case started from triparental conjugation. This suggests that an issue may have occurred during the process of triparental conjugation, such as contamination. To check for contamination, the transformant *Synechocystis* could be streaked onto LB-Agar, and if growth is observed this would indicate that contamination of the *Synechocystis* cultures had occurred. The transformation of the *Synechocystis* Δaas should be also repeated, using freshly revived *Synechocystis* Δaas, and *E. coli*, and the biosynthesis experiment repeated with the newly generated strains.

6.4 Conclusions

In this chapter a novel synthetic pathway for the biosynthesis of δ-decalactone and γ-decalactone was expressed in *E. coli* and found to be functional. Biosynthesis of both δ-decalactone and γ-decalactone was achieved, although titres were low. To my knowledge this was the first time *de novo* biosynthesis of lactones has been achieved in *E. coli*. This was also the first time that *de novo* biosynthesis of δ-decalactone had been achieved in an engineered microbial host.

Decalactone titres were limited by lower-than-expected titres of decanoic acid, and poor conversion of the decanoic acid to BM3. Future work will be needed to address these issues, and due to analytical issues, some experiments in this chapter should be repeated.

Overall, the pathway shows promise for the tailored biosynthesis of both δ- and γ-lactones of different chain lengths.

Chapter 7

Prototyping Biosynthetic Pathways for Alternative Flavour and Fragrance Compounds

7.1 Aims

In this study the development of synthetic pathways for the biosynthesis of flavour and fragrance compounds was targeted, to drive the development of cyanobacteria biotechnology. Initial prototyping and development of the pathways was initially performed in *E. coli* for a faster design-build-test-learn cycle.

In Chapters 4-6 a synthetic pathway for the biosynthesis of γ-decalactone and δ-decalactone was developed, however intensive enzyme engineering was required, increasing the amount of work that needed to be performed in *E. coli*.

Biosynthesis of other flavour and fragrance compounds was therefore investigated, where enzyme engineering would not be required, or an existing synthetic pathway could be adapted for their biosynthesis. This would reduce the amount of work that needed to be performed in *E. coli*, enabling a more rapid transfer of the pathway to cyanobacteria.

Butyl acetate, ethyl butyrate, methyl ketones and green leaf volatiles were identified as target flavour and fragrance compounds for bioproduction. The aim of this chapter is therefore to prototype pathways for the biosynthesis of these compounds.

7.2 Results: Butyl Acetate Biosynthesis

Bioproduction of butyl acetate, which has uses as both a flavour and fragrance compound and as an industrial solvent[241,244], was targeted due to its large commercial market, which from 1999-2009 ranged from 89 to 107 million kg[241].

Modification of a previously reported acetate ester biosynthesis pathway[28] by substituting the previously used C8 specific acyl-ACP thioesterase for a C4 specific acyl-ACP thioesterase could allow for the biosynthesis of butyl acetate (Fig 7.1A). The pathway is expressed using a two-plasmid system, the first plasmid expresses the genes used for alcohol biosynthesis (Fig 7.1B), and the second a gene for conversion of the alcohol to an acetate ester (Fig 7.1C).

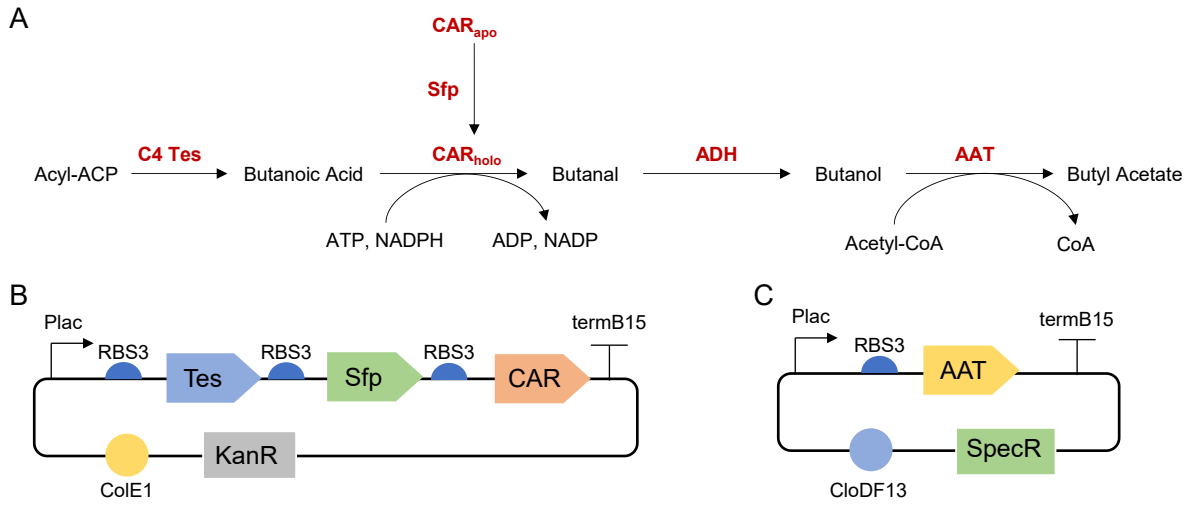


Figure 7.1 Pathway and plasmid design for butyl acetate biosynthesis *A) Synthetic metabolic pathway for alcohol biosynthesis. Pathway enzymes are highlighted in red. C4 Tes: C4 specific acyl-ACP thioesterase Sfp: Phosphopantetheine transferase (Bacillus subtilis), CAR: Carboxylic acid reductase (Mycobacterium marinum), ADH: Native alcohol dehydrogenases, AAT: Alcohol-acyl transferase B) Plasmid design for biosynthesis of butanol. A pET plasmid with ColE1 origin of replication and kanamycin resistance cassette was used. Plac and RBS3 were used to express a C4 Tes, Sfp and CAR in a single operon. C) Plasmid design for conversion of butanol to butyl acetate. A pCDF plasmid with CloDF13 origin of replication and spectinomycin resistance cassette was used. Plac and RBS3 were used to express the AAT.*

7.2.1 Screening C4 specific acyl-ACP thioesterases

Three thioesterases from the literature with C4 activity were identified, from *Bacteroides fragilis* (Tes4), *Bacteroides thetaiotamicron* (BtTEs) and *Marvinbryantia formatexigens* (MfTes)[104,276]. Biosynthesis of alcohols using a thioesterase, Sfp and CAR was well established in the Jones lab, so the candidate thioesterases were screened by co-expression

with Sfp and CAR using a plasmid as in Fig 7.1B. Biosynthesis of butanol and longer chain length alcohols was investigated to assay thioesterase activity and specificity.

Very high titres of butanol were observed when expressing either Tes4 or BtTes. Butanol titres were higher when expressing Tes4 (1562 mg/L) than when expressing BtTes (1165 mg/L), albeit not significantly so ($p = 0.190$). Very low titres of longer chain length alcohols were observed (Fig 7.2A). Similar results were also observed for the alcohol yields (Fig 7.2B). There was no significant difference in growth between *E. coli* expressing Tes4 or BtTes, however expression of MfTes had a negative impact on growth (Fig 7.2C).

Based on these results, Tes4 was selected for work going forward due to its higher butanol titres.

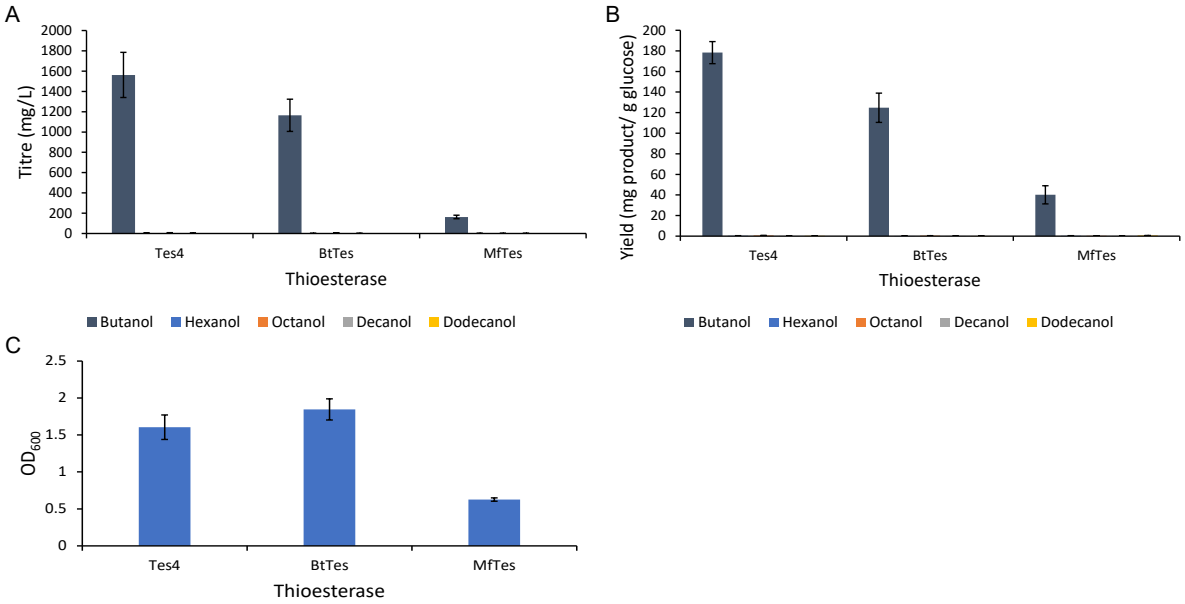


Figure 7.2 Screening C4 thioesterases *E. coli ΔfadD were transformed with a plasmid for alcohol biosynthesis expressing different C4 thioesterases, Sfp and CAR A) Alcohol titres B) Alcohol yields C) OD_{600} after 48 hrs cultivation in M9 media and a 10% v/v isopropyl myristate overlay for 48 hrs at 30°C. The overlay was sampled and analysed by GC-MS. Cells were induced with 0.5 mM IPTG. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

7.2.2 Screening AATs for butyl acetate biosynthesis

The final step in the butyl acetate biosynthesis pathway is conversion of butanol to butyl acetate, catalysed by an alcohol-acyltransferase (AAT). Some AATs were available in the Jones lab; AtF1 from yeast[288], Chloramphenicol transferase (CAT)[307], SAAT from cultivated strawberry[127], VAAT from wild strawberry[291] and VpAAT1 from papaya[289]. These AATs were screened for butyl acetate biosynthesis by co-transforming *E. coli* with a plasmid for butanol biosynthesis using Tes4, and a second plasmid expressing an AAT as in Fig 7.1C, which had

already been produced in prior work by the Jones lab. As a control, a second plasmid expressing GFP instead was used.

Biosynthesis of butyl acetate was observed in strains expressing Atf1, CAT, SAAT and VAAT. The highest titres of butyl acetate were observed when expressing Atf1 (571.4 mg/L). Expression of Atf1 led to a decrease in butanol titres compared to the GFP control, from 999.4 mg/L to 343 mg/L, however due to the large variation in titres within the triplica this was not significant ($p = 0.0573$) (Fig 7.3A). When titres were converted from mg/L to mM a decrease in total product titres was observed when expressing Atf1 or SAAT, indicating that product loss had occurred (Fig 7.3B). The product yields gave similar results to the product titres (Fig 7.3C). No significant changes in OD_{600} were observed when expressing an AAT instead of GFP (Fig 7.3D).

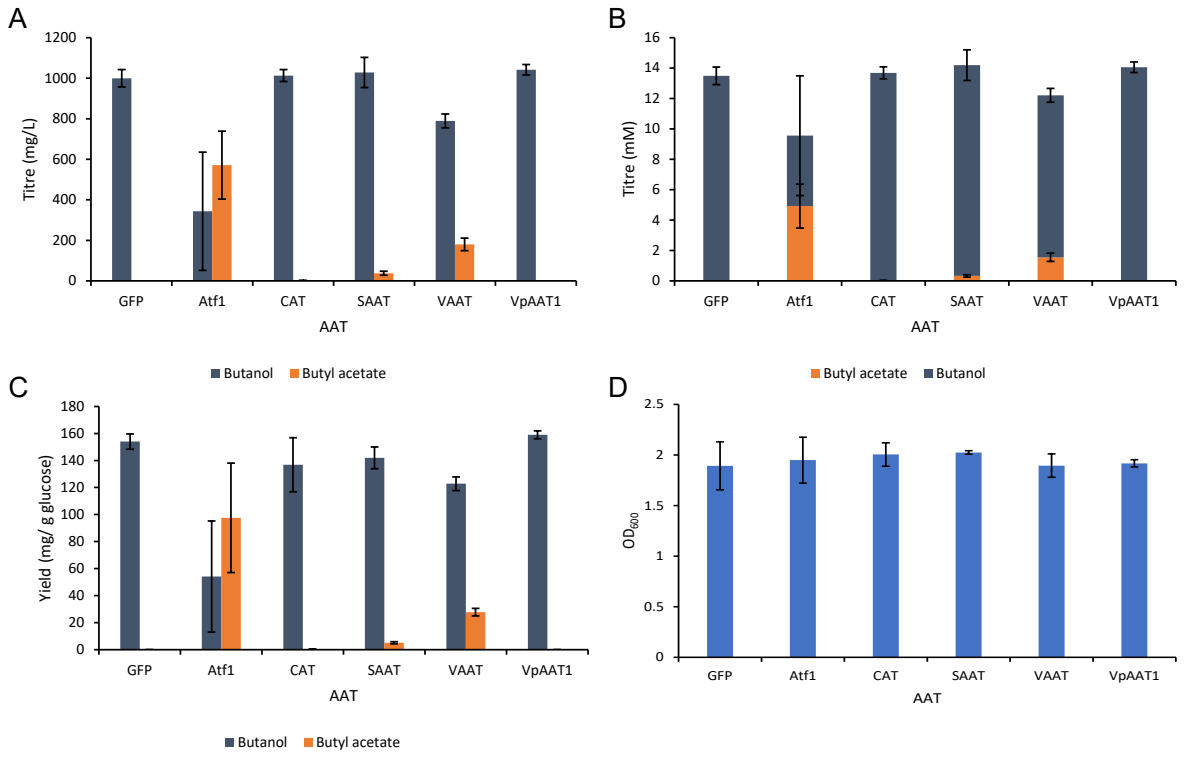


Figure 7.3 Screening AATs for butyl acetate biosynthesis *E. coli ΔfadD were transformed with a plasmid for butanol biosynthesis expressing Tes4, Sfp and CAR, and a second plasmid expressing either GFP or an AAT A) Titres in mg/L B) Titres in mM C) Alcohol yields D) OD_{600} after 48 hrs cultivation in M9 media and a 10% v/v hexadecane overlay for 48 hrs at 30°C. The overlay was sampled and analysed by GC-MS. Cells were induced with 0.5 mM IPTG. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

7.2.3 Screening a library of new AATs for butyl acetate biosynthesis

An undergraduate working on a bioinformatics project in the Jones lab had identified a library of new uncharacterised AATs (unpublished work, Table 7.1). It was therefore decided to screen these AATs for butyl acetate biosynthesis as well. The new AATs were screened in

the same manner as before, with the exception that the plasmid used to express AATs and GFP control was changed (Fig 7.4C). This new plasmid had used the PrhaBAD promoter rather than Plac, so that future optimisation of butanol biosynthesis and esterification could be performed individually.

Name	Species	Enzyme Annotation
AtHPFT	*Arabidopsis thaliana*	Omega-hydroxypalmitate O-feruloyl Transferase
CcVS	*Citrus clementina*	Vinorine Synthase
CsVS	*Citrus sinensis*	Vinorine Synthase
LaAAT1	*Lavandula angustifolia*	Putative Alcohol acyltransferase 1
LaAAT2	*Lavandula angustifolia*	Putative Alcohol acyltransferase 2
LaAT1	*Lavandula angustifolia*	Rosmarinic Acid Synthase LaAT1
LiAAT4	*Lavandula x intermedia*	AAT-4 Alcohol Acetyltransferase
NtBBT	*Nicotania tabacum*	Benzyl Alcohol O-benzoyltransferase
NtSHCT	*Nicotania tabacum*	Shikimate O-hydroxycinnamoyltransferase
RsVS	*Rauvolfia serpentina*	Vinorine Synthase
SsMMT	*Salvia splendens*	Malonyl-coenzyme:anthocyanin 5-O-glucoside-6-O-malonyltransferase

Table 7.1 New AAT library

Butanol titres were far lower than in previous experiments. Butanol titres in the GFP control were only 150 mg/L. Large variation within the triplicas were also observed. Butyl acetate titres when expressing Atf1, SAAT and VAAT were also lower than previously observed, and no butyl acetate biosynthesis could be observed when expressing CAT. Amongst the new AATs, butyl acetate biosynthesis could only be observed when expressing LiAAT4. Butyl acetate titres when expressing LiAAT4 (157.3 mg/L) were significantly ($p = 0.0273$) higher than when expressing Atf1 (72.4 mg/L) (Fig 7.34). OD_{600}s were also lower than previously observed, and expression of the new AATs did not have significant effects on growth (Fig 7.4B). When the experiment was repeated similar results were obtained (data not shown).

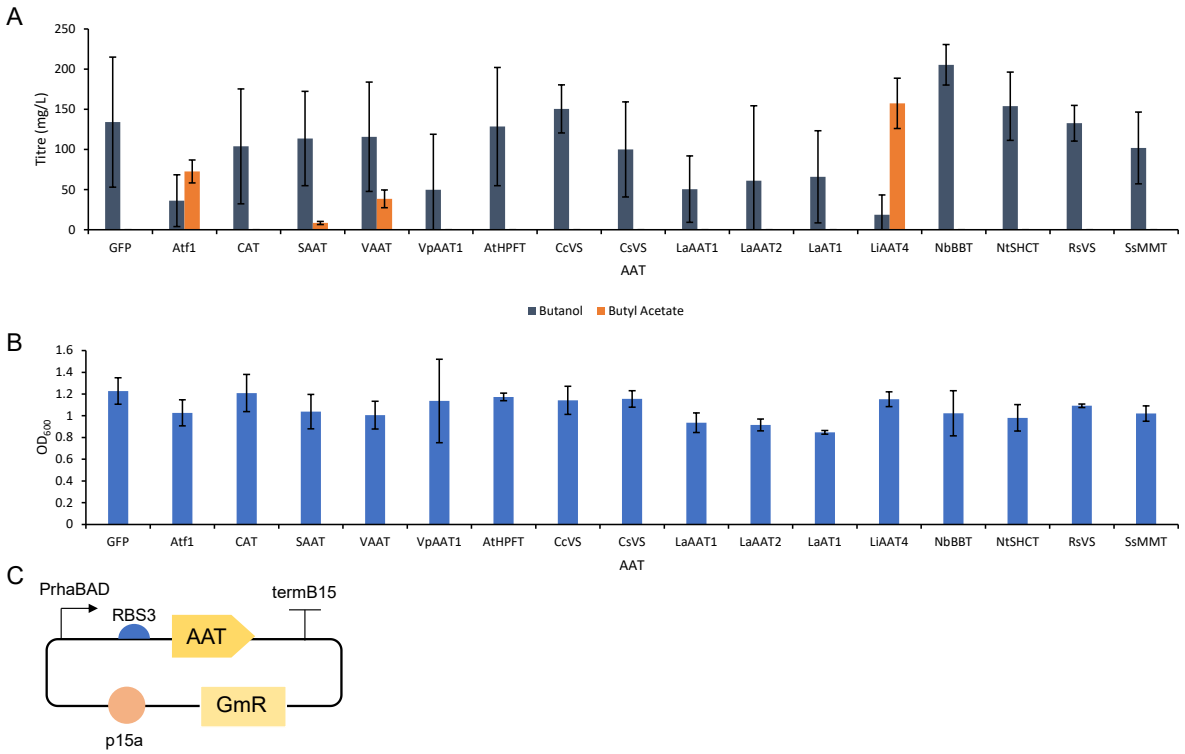


Figure 7.4 Screening new AATs for butyl acetate biosynthesis *E. coli ΔfadD were transformed with a plasmid for butanol biosynthesis expressing Tes4, Sfp and CAR, and a second plasmid expressing either GFP or an AAT A) Titres B) OD_{600} after 48 hrs cultivation in M9 media and a 10% v/v hexadecane overlay for 48 hrs at 30°C. The overlay was sampled and analysed by GC-MS. Cells were induced with 0.5 mM IPTG and 0.1% rhamnose. Data are the average of 3 biological replicates and error bars are of 1 standard deviation C) Plasmid design for conversion of butanol to butyl acetate. A pACYC plasmid with p15a origin of replication and gentamycin resistance cassette was used. PrhaBAD and RBS3 were used to express the AAT.*

7.2.4 The effect of host background and pACYC-PrhaBAD-GmR on butanol biosynthesis

In the previous experiments an *E. coli* ΔfadD background had been used for butanol and butyl acetate biosynthesis. It was decided to investigate the effect of different *E. coli* host backgrounds on butanol biosynthesis. *E. coli* BL21, BW25113, ΔfadD and ΔfadE were tested. It was also suspected that the new plasmid used to express AATs, pACYC-PrhaBAD-GmR, was having a negative effect on butanol biosynthesis, so the different *E. coli* backgrounds were either transformed with a plasmid for butanol biosynthesis and induced with IPTG, or co-transformed with the plasmid for butanol biosynthesis and pACYC-PrhaBAD-GFP-GmR and induced with IPTG and rhamnose. Biosynthesis of butanol was then assayed.

Butanol titres of 701.1 mg/L were observed in the ΔfadD strain transformed with a single plasmid. This was lower than observed in Section 7.2.1 and 7.2.2, but higher than in Section 7.2.3. The highest butanol titres were observed in the BW25113 (892 mg/L) and ΔfadE (926.4 mg/L) backgrounds, with no significant difference between the two ($p = 0.787$). Addition of the second plasmid led to significant decreases in butanol titre in the BW25133, ΔfadD and ΔfadE backgrounds. On the other hand, addition of the second plasmid actually improved butanol titres in the BL21 host background (Fig 7.5A). Similar results were observed for the butanol yields (Fig 7.5B). Decreased OD_{600} was observed in the BW25133, ΔfadD and ΔfadE strains following co-transformation with the second plasmid, while in the BL21 strain no significant changes in OD_{600} were observed (Fig 5.7C).

Based on these results, the ΔfadE background was used in the following experiments.

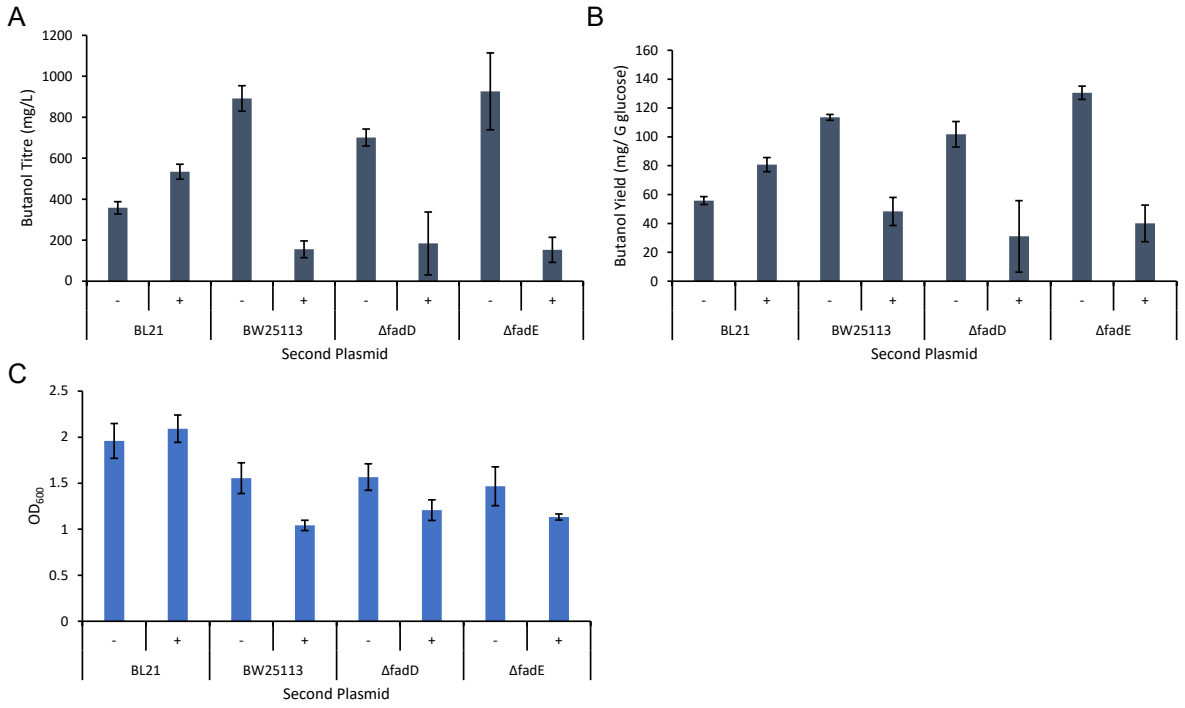


Figure 7.5 Screening effect of host background and pACYC-PrhaBAD-GmR on butanol biosynthesis *E. coli BL21, BW25113, ΔfadD and ΔfadE were transformed with either plasmid for butanol biosynthesis expressing Tes4, Sfp and CAR (-), or the butanol biosynthesis plasmid and pACYC-PrhaBAD-GmR expressing GFP (+) A) Titres B) OD_{600} after 48 hrs cultivation in M9 media and a 10% v/v hexadecane overlay for 48 hrs at 30°C. The overlay was sampled and analysed by GC-MS. Cells were induced with 0.5 mM IPTG and if co-transformed with the second plasmid 0.1% rhamnose. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

7.2.5 Screening alternative plasmid backbones for expressing AAT

As pACYC-PrhaBAD-GmR had been observed to have a negative impact on butanol titres, different backbones that could be used to express AAT were screened to see if they affected butanol biosynthesis. *E. coli* ΔfadE were co-transformed with the butanol biosynthesis plasmid and different plasmid backbones that could be used to express an AAT. If both plasmids had the Plac promoter, the *E. coli* were induced with only IPTG, if the second plasmid had the PrhaBAD promoter, the *E. coli* were induced with both IPTG and rhamnose.

Only the pACYC-PrhaBAD-GmR backbone was found to have a negative impact on butanol titres. Co-expression of the pACYC-Plac-GmR backbone resulted in increased butanol titres. Apart from these two backbones, the backbone used did not have an impact on butanol titres (Fig 7.6A). Similar results were observed for the butanol yield (Fig 7.6B). Co-expression of pACYC-PrhaBAD-EryR and pACYC-PrhaBAD-GmR led to decreased OD_{600}, while co-expression of pACYC-Plac-GmR led to improved growth (Fig 7.6C).

The pCDF-Plac-SpecR backbone was chosen for future work as it had no significant impact on titre, yield or growth, and had already been used in Section 7.2.2, meaning fewer new constructs needed to be assembled.

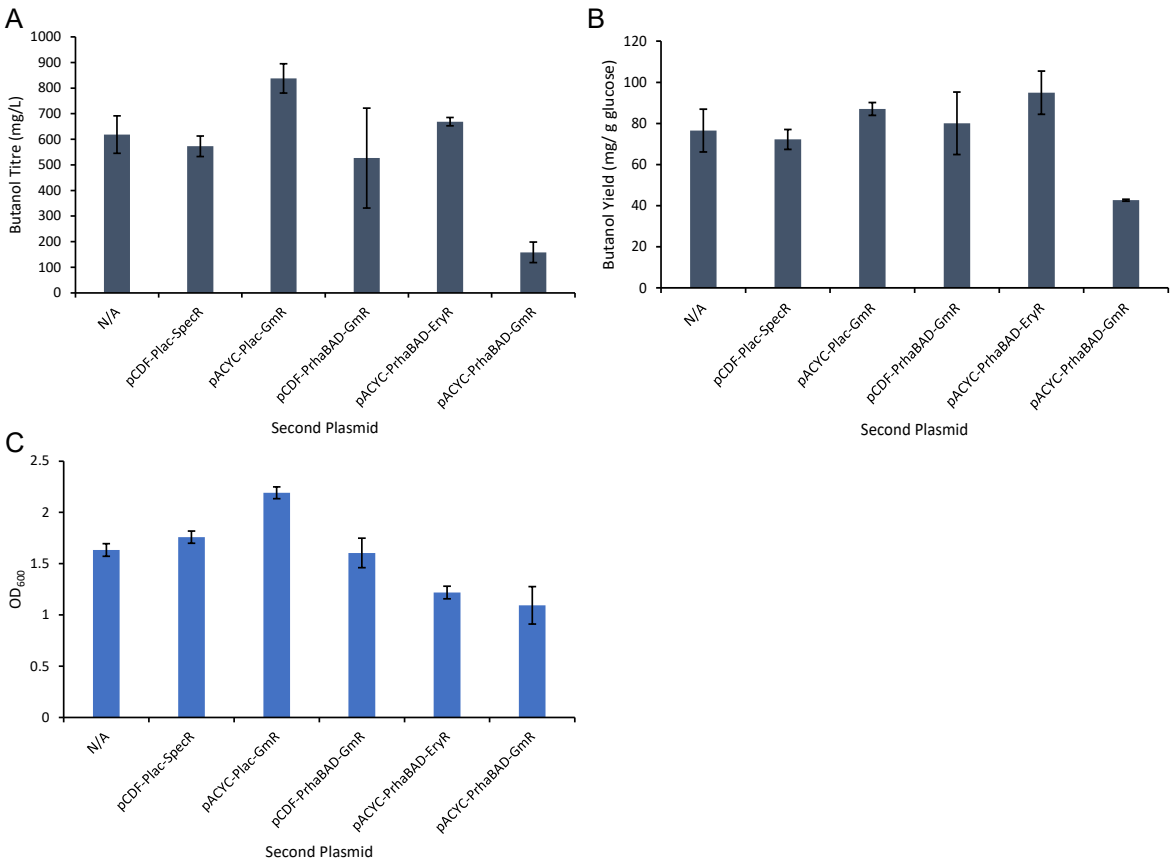


Figure 7.6 Screening alternative plasmid backbones for expressing AAT *E. coli ΔfadE were transformed with either plasmid for butanol biosynthesis expressing Tes4, Sfp and CAR (N/A), or the butanol biosynthesis plasmid and different plasmid backbones expressing GFP A) Titres B) OD_{600} after 48 hrs cultivation in M9 media and a 10% v/v hexadecane overlay for 48 hrs at 30°C. Cells were induced with 0.5 mM IPTG and if co-transformed with the second plasmid with PrhaBAD 0.1% rhamnose. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

7.2.6 AAT feeding assay

In parallel to the work in Section 7.2.5, the new AATs were screened again, by using a feeding assay. *E. coli* expressing the AATs using the constructs made for Section 7.2.3 were fed with 1g/L butanol and bioconversion of the butanol to butyl acetate was investigated.

Bioconversion of butanol to butyl acetate was observed in strains expressing Atf1, CAT, SAAT, VAAT and LiAAT4 confirming previous results. In this instance higher butyl acetate titres were observed when Atf1 was expressed (768.9 mg/L) then when LiAAT4 was expressed (467.9 mg/L), however the difference was not significant ($p = 0.105$). While 1 g/L of butanol was fed, 800 mg/L was observed in the GFP control, indicating that there was loss of the fed substrate. When Atf1 was expressed no butanol could be observed and when LiAAT4 was

expressed very low amounts of butanol (3.90 mg/L) were observed (Fig 7.7A). No significant differences in OD_{600} were observed when expressing the different AATs (Fig 7.7B).

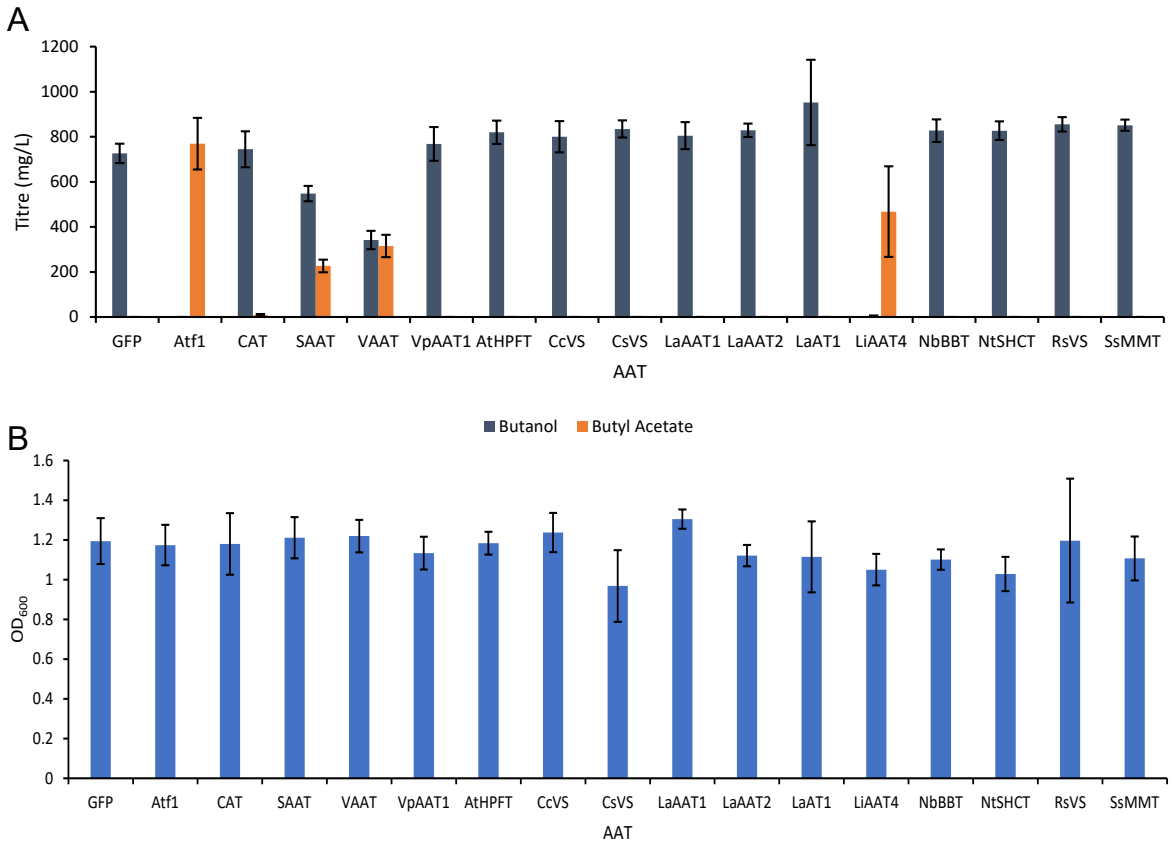


Figure 7.7 AAT feeding assay *E. coli ΔfadE were transformed with a plasmid expressing GFP or an AAT A) Titres B) OD_{600} after 48 hrs cultivation in M9 media and a 10% v/v hexadecane overlay for 48 hrs at 30°C. The overlay was sampled and analysed by GC-MS. Cells were induced with 0.5 mM IPTG and fed with 1 g/L butanol. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

7.2.7 Butyl acetate biosynthesis

Butyl acetate biosynthesis was repeated as described in Section 7.2.2, except that only AATs that had been shown to bioconvert butanol to butyl acetate in the feeding experiment, and a ΔfadE background was used.

The highest butyl acetate titres were observed when Atf1 was expressed (153 mg/L). This was lower than previously observed in Section 7.2.2. The next highest butyl acetate titres were observed when LiAAT4 was expressed (83.1 mg/L), which was significantly lower than Atf1 ($p = 0.0124$) (Fig 7.8A). When the titres were converted to mM, decreases in overall titres could be observed when Atf1 and LiAAT4 were expressed, indicating that product loss had

occurred (Fig 7.8B). Similar results to product titres were observed for product yields (Fig 5.8C). Again, expression of AAT had no significant effect on growth (Fig 5.8D).

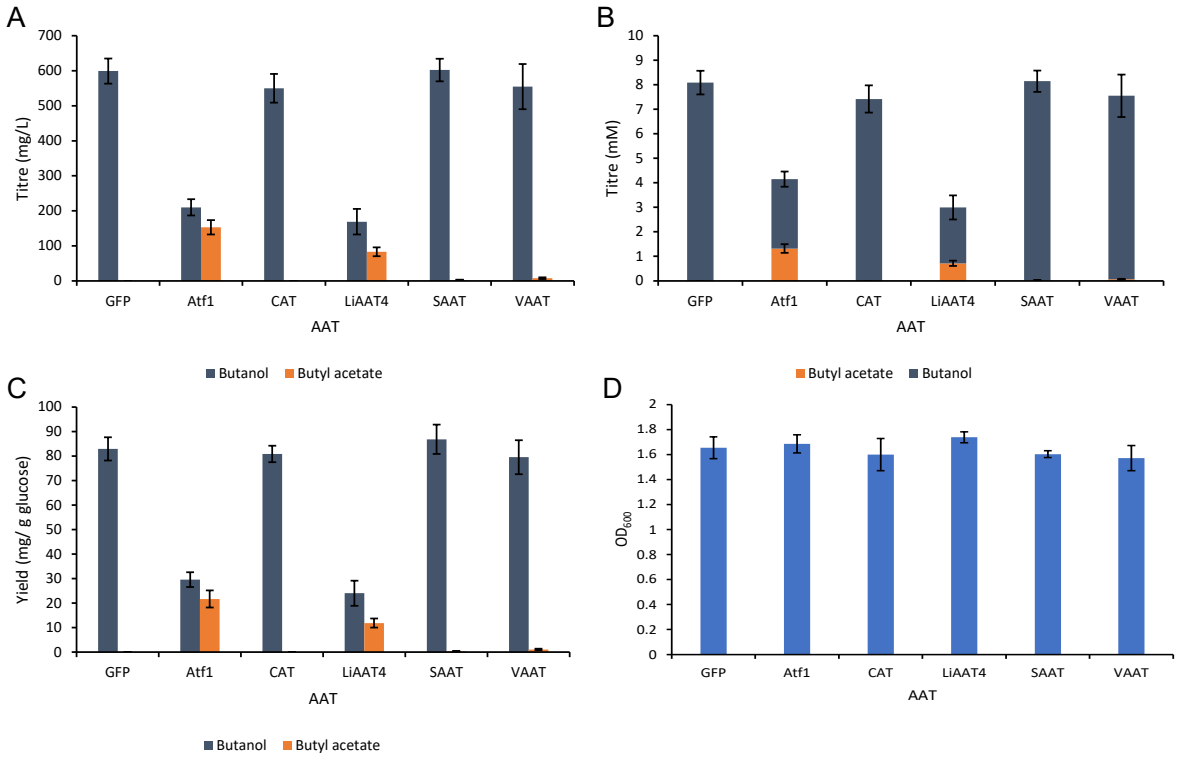


Figure 7.8 Butyl acetate biosynthesis *E. coli ΔfadE were transformed with a plasmid for butanol biosynthesis expressing Tes4, Sfp and CAR, and a second plasmid expressing either GFP or an AAT A) Titres in mg/L B) Titres in mM C) Alcohol yields D) OD_{600} after 48 hrs cultivation in M9 media and a 10% v/v hexadecane overlay for 48 hrs at 30°C. The overlay was sampled and analysed by GC-MS. Cells were induced with 0.5 mM IPTG. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

7.3 Results: Ethyl Butyrate Biosynthesis

Ethyl butyrate is an important compound in the flavour and fragrance industry and is used to fruity tones to foods and beverages[245]. Bioproduction of ethyl butyrate was therefore also targeted.

A novel pathway for the biosynthesis of ethyl butyrate was designed (Fig 7.9A). The pathway was expressed using a two-plasmid system, the first plasmid expresses the genes used for butanoyl-CoA biosynthesis (Fig 7.9B), and the second a gene for conversion of the butanoyl-CoA to ethyl butyrate (Fig 7.9C).

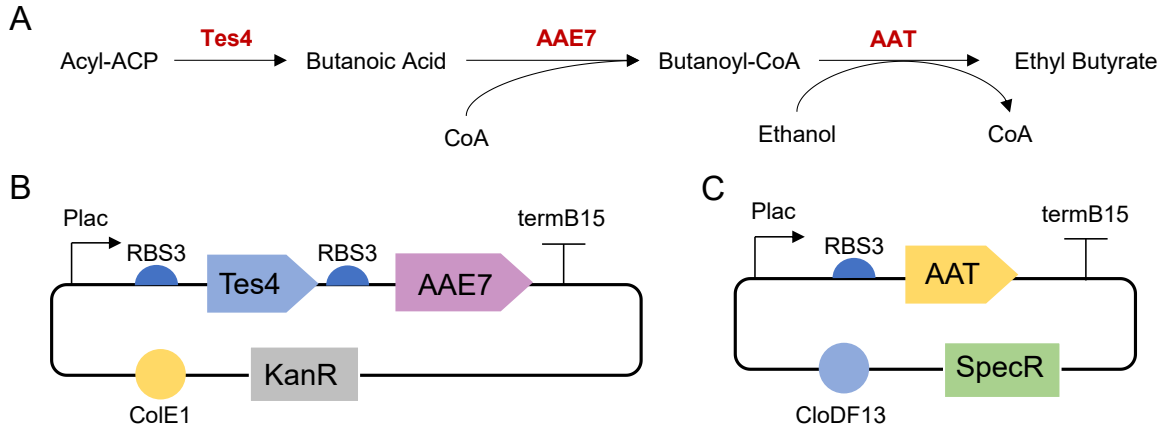


Figure 7.9 Pathway and plasmid design for ethyl butyrate biosynthesis *A) Synthetic metabolic pathway for alcohol biosynthesis. Pathway enzymes are highlighted in red. Tes4: C4 specific acyl-ACP thioesterase (Bacteroides fragilis), AAE7: Butanoic acid CoA ligase (Arabidopsis thaliana) AAT: Alcohol-acyl transferase B) Plasmid design for biosynthesis of butanoyl-CoA. A pET plasmid with ColE1 origin of replication and kanamycin resistance cassette was use. Plac and RBS3 were used to express Tes4 and an AAE7 in a single operon. C) Plasmid design for conversion of butanoyl-CoA to ethyl butyrate. A pCDF plasmid with CloDF13 origin of replication and spectinomycin resistance cassette was used. Plac and RBS3 were used to express an AAT.*

7.3.1 Proof-of-concept biosynthesis of ethyl butyrate

Arabidopsis thaliana AAE7 (AtAAE7) had been reported to specifically ligate butanoic acid with CoA[323]. Meanwhile, in Section 7.2.1 Tes4 had been found to be a highly active and specific C4 thioesterase. Therefore co-expression of these enzymes could be used to biosynthesise butanoyl-CoA. As a proof-of-concept experiment, Tes4 and AtAAE7 were co-expressed with GFP or different AATs from the initial smaller library used in Section 7.2.2.

Ethyl butyrate biosynthesis was observed in strains expressing Atf1, SAAT and VAAT, with the highest ethyl butyrate titres observed when VAAT was used (14.5 mg/L) (Fig 7.10A).

Similar results were observed for the ethyl butyrate yield (Fig 7.10B), and no significant differences in OD_{600} were observed when expressing the different AATs (Fig 7.10C).

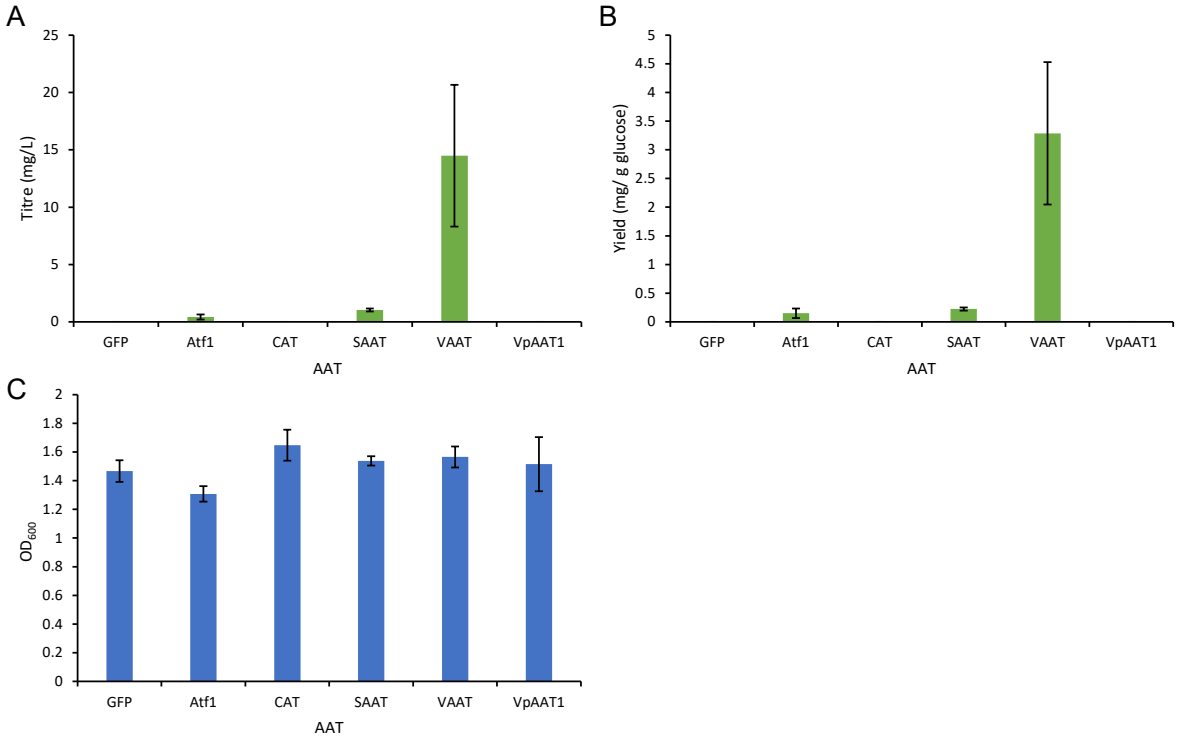


Figure 7.10 Proof-of-concept biosynthesis of ethyl butyrate *E. coli ΔfadE were transformed with a plasmid for butanoyl-CoA biosynthesis expressing Tes4 and AtAAE7, and a second plasmid expressing either GFP or an AAT A) Titres B) Yields C) OD_{600} after 48 hrs cultivation in M9 media and a 10% v/v hexadecane overlay for 48 hrs at 30°C. The overlay was sampled and analysed by GC-MS. Cells were induced with 0.5 mM IPTG. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

7.3.2 Screening new AAE7s

A student under my supervision generated and investigated a library of AtAAE7 homologues. In feeding experiments with butanoic acid, two of the new AAE7s, from *Gossypium arboreum* (GaAAE7) and from *Theobroma cacao* (TcAAE7) showed increased consumption of butanoic acid (Jones lab, unpublished work).

Unfortunately, production of butanoyl-CoA could not be quantified. This is because the analytical technique to measure butanoyl-CoA required a new HPLC column and new HPLC method, which had a higher operating pressure. A new cell extraction technique was also required to extract the butanoyl-CoA for analysis. These changes triggered a series of cascading faults, including multiple leaks, a broken purge valve, a blocked needle seat, autosampler arm faults in parallel to unrelated failure of the CPU cooling fan and burn out of the DAD bulb due to old age of the system. These problems required a lot of repair work,

including three separate disassemblies and reassemblies of the autosampler unit to install new parts. This meant during the student's project the HPLC was often out of order, and unfortunately never managed to perform a complete run of their butanoyl-CoA samples and standards. Once the HPLC was finally fixed, it was considered too risky to attempt running butanoyl-CoA samples again, and so this was not investigated in this work.

As a control, *E. coli* was transformed with a plasmid expressing only Tes4 (Fig 7.11F), and a second plasmid expressing GFP as in Section 7.3.1. This strain would only produce butanoic acid. *E. coli* were then transformed with a plasmid expressing Tes4 and AAE7 as in Section 7.3.1, and a second plasmid expressing GFP. If butanoic acid titres decreased in this strain, this would be an indication that the AAE7 was active and ligating the butanoic acid to CoA. Finally, *E. coli* were transformed with the Tes4 and AEE7 plasmid, and a plasmid expressing VAAT. This strain should then be able to biosynthesise ethyl butyrate. This design would allow ligation of butanoic acid to CoA to be indirectly observed, and the corresponding effects of the new AAE7s on ethyl butyrate titre.

Co-expression of an AAE7 with Tes4 led to a drop in butanoic acid titres. Expression of AtAAE7 led to the smallest drop in titres, while expression of GaAAE7 led to the lowest butanoic acid titres (Fig 7.11A). This suggested the new AAE7s were more active than AtAAE7. When Tes4, an AAE7 and VAAT were co-expressed biosynthesis of ethyl butyrate was observed. Although GaAAE7 appeared to consume the most butanoic acid, the highest ethyl butyrate titres were observed when expressing TcAAE7 (46.1 mg/L). This was a significant ($p = 0.00961$) 15-fold increase in ethyl butyrate titres compared to when AtAAE7 was expressed (2.92 mg/L). Ethyl butyrate titres were lower than previously observed in Section 7.3.1, potentially due to the different BW25133 host strain used in this experiment (Fig 7.11B). When product titres were converted to mM product loss could be observed when expressing Tes4, an AEE7 and VAAT (Fig 7.11C). Similar results to the titres in mg/L could be observed for product yields (Fig 7.11D). There was no significant difference in OD_{600} between the strains (Fig 7.11E).

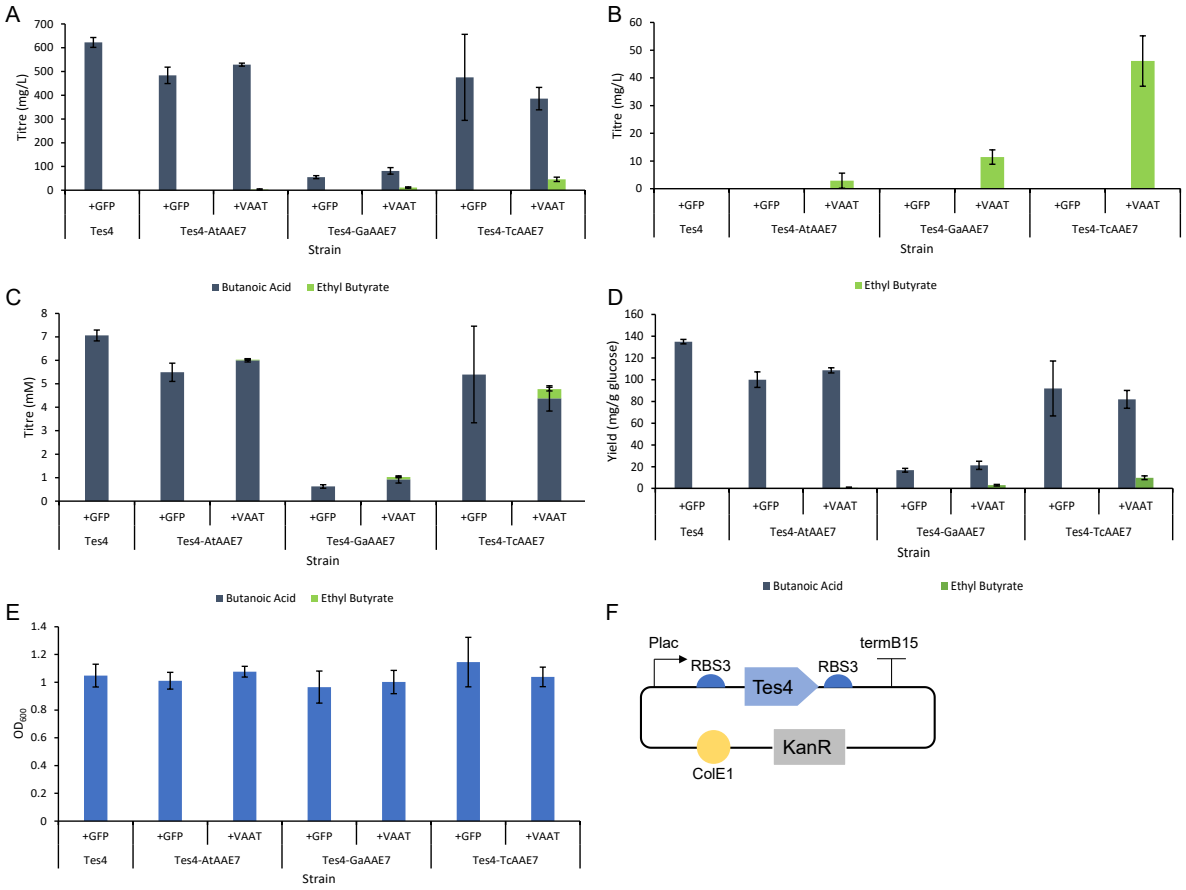


Figure 7.11 Screening new AAE7s *E. coli BW25113 were transformed with a plasmid expressing Tes4 or a plasmid for butanoyl-CoA biosynthesis expressing Tes4 and an AAE7, and a second plasmid expressing either GFP or VAAT A) Butanoic acid and ethyl butyrate titres in mg/L B) Ethyl butyrate titres in mg/L C) Butanoic acid and ethyl butyrate titres in mM D) Yields E) OD_{600} after 48 hrs cultivation in M9 media and a 10% v/v hexadecane overlay for 48 hrs at 30°C. The overlay was sampled and analysed by GC-MS. Cells were induced with 0.5 mM IPTG. Data are the average of 3 biological replicates and error bars are of 1 standard deviation F) Plasmid design for expressing only Tes4. A pET plasmid with ColE1 origin of replication and kanamycin resistance cassette. Plac and RBS3 were used to express Tes4.*

7.3.3 Screening different host backgrounds for ethyl butyrate biosynthesis

Knock-out of FadE has been reported to increase ester titres[246], as this prevents degradation of acyl-CoA. The student had also generated a ΔyciA strain and found removing this thiolase improved butanoyl-CoA titres (Jones lab, unpublished work). These host backgrounds were therefore screened to see if they improved ethyl butyrate titres. *E. coli* BW25113, ΔfadE and ΔyciA were transformed with Tes4, the different AAE7s and VAAT. While in Section 7.3.2 expression of TcAEE7 resulted in the highest ethyl butyrate titres, and it was not expected that the different AAE7s would have differing activities in the different host strains, all three AAE7s were expressed in each host strain in case there were combinatorial effects.

Ethanol is also required to biosynthesise ethyl butyrate, however in previous experiments ethanol titres had not been measured. Ethanol standards were therefore to improve understanding of the system.

Again, lower butanoic titres were observed in BW25113 when expressing GaAAE7 and TcAAE7. However, this effect was not as pronounced when the AAE7s were expressed in ΔyciA, and in ΔfadE higher butanoic acid titres were actually observed when expressing GaAAE7 (Fig 7.12A).

Ethyl butyrate titres in BW25113 were not consistent with the previous experiment in Section 7.3.2, with GaAAE7 and TcAAE7 expressing strains producing much less ethyl butyrate. However, the AtAAE7 expressing strain produced 2.93 mg/L of ethyl butyrate, which was consistent with the previous result.

Moving from BW25113 to ΔfadE and ΔyciA gave mixed results. When TcAAE7 was expressed in the ΔfadE and ΔyciA strains, ethyl butyrate titres were lower than when TcAAE7 was expressed in the BW251133 strain. On the other hand, ethyl butyrate titres were higher when AtAEE7 and GaAAE7 were expressed in either the ΔfadE or the ΔyciA strains compared to when AtAEE7 and GaAAE7 were expressed in the BW25113 strain. The highest ethyl butyrate titres were observed when expressing AtAAE7 in the ΔyciA strain (14.2 mg/L).

There was no significant difference in ethyl butyrate titres following expression of AtAAE7 or GaAAE7 in the ΔyciA strain, and no significant different in ethyl butyrate titres following expression of AtAAE7 in the ΔfadE or the ΔyciA strains. Ethyl butyrate titres when AtAAE7 were expressed in the ΔfadE strain were 11.9 mg/L, broadly in line with previous results in Section 7.3.1 (Fig 7.12B).

Ethanol titres were much lower than butanoic acid titres (Fig 7.12A), and sometimes no ethanol production could be observed (Fig 7.12B). There was no discernible pattern between ethanol titres and ethyl butyrate titres, but as ethanol titres were much lower than butanoic acid titres these results suggested ethanol supply could be a limiting factor in ethyl butyrate biosynthesis.

The products yields gave similar results to product titres (Fig 7.12C) and there were no discernible trends in OD_{600} when expressing the different genes in different host backgrounds (Fig 7.12D).

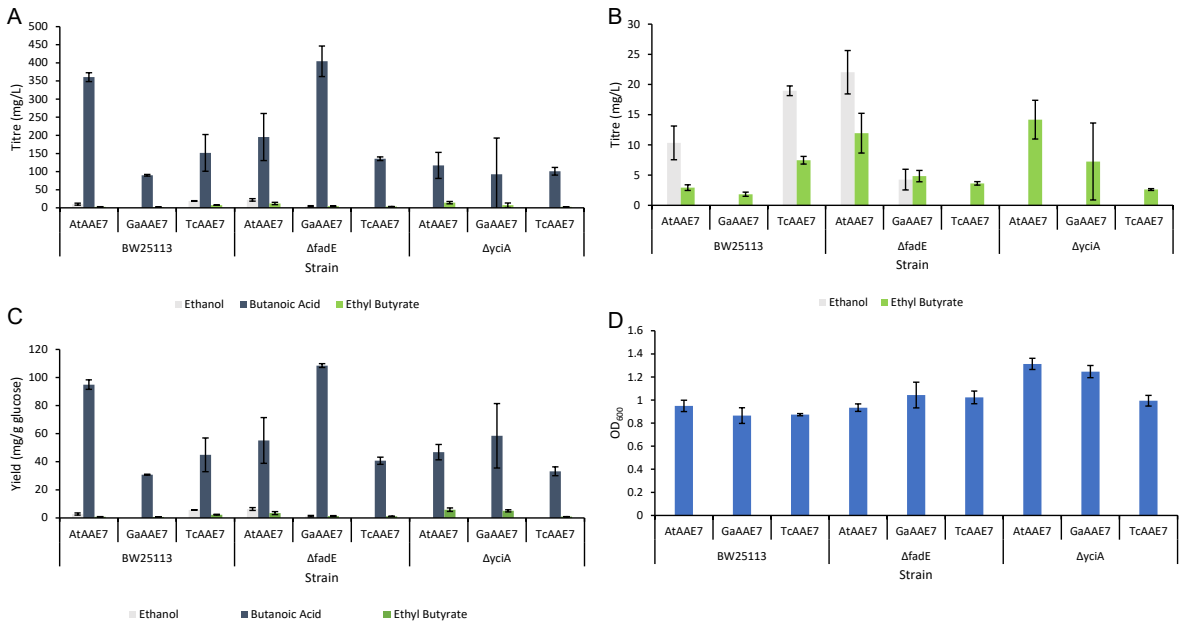


Figure 7.12 Screening different host backgrounds for ethyl butyrate biosynthesis *E. coli BW25113, ΔfadE and ΔyciA were transformed with a plasmid expressing Tes4 and an AAE7, and a second plasmid expressing VAAT A) Butanoic acid and ethyl butyrate titres B) Ethanol and ethyl butyrate titres C) Yields D) OD_{600} after 48 hrs cultivation in M9 media and a 10% v/v hexadecane overlay for 48 hrs at 30°C. The overlay was sampled and analysed by GC-MS. Cells were induced with 0.5 mM IPTG. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

7.3.4 Could LiAAT4 and PDC be used to improve ethyl butyrate titres?

LiAAT4 was the only active AAT from the new library. It was therefore decided to only investigate whether LiAAT4 could be used instead of VAAT to biosynthesise ethyl butyrate and improve titres. The results in Section 7.3.3 suggested that ethanol may have been a limiting factor in ethyl butyrate biosynthesis, so it was investigated whether expression of PDC^{429} could be used to increase ethanol and therefore ethyl butyrate biosynthesis. To do so, the construct used for expressing AAT was modified (Fig 7.13).

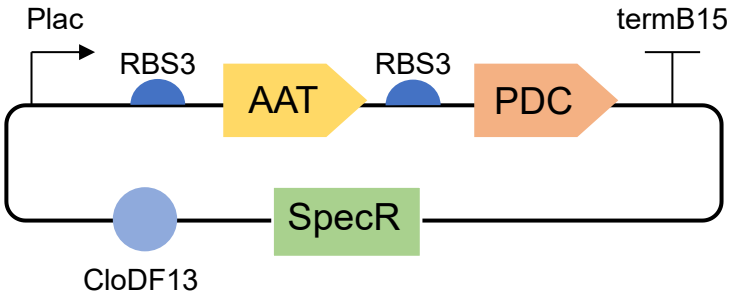


Figure 7.13 Construct design for expressing AAT and PDC *A pCDF plasmid with CloDF13 origin of replication and spectinomycin resistance cassette was used. Plac and RBS3 were used to express the AAT and PDC (Zymonas mobilis) in a single operon.*

E. coli were co-transformed with a plasmid expressing Tes4 and an AAE7, and with a second plasmid expressing either VAAT or LiAAT4 either alone or with PDC.

Upon expression of PDC very high ethanol titres were observed, reaching up to 3.47 g/L. (Fig 7.14A). When VAAT and PDC were co-expressed on the second plasmid a significant decrease in butanoic titres was observed, with zero or near-zero titres observed. At the same time, there was a significant and large rise in ethyl butyrate titres. The highest titres were observed when Tes4, GaAAE7, VAAT and PDC were co-expressed, reaching 224.0 mg/L, a more than 90-fold improvement compared to the strain without PDC (2.42 mg/L). On the other hand, no ethyl butyrate production was observed when LiAAT4 was expressed, suggesting it cannot accept butanoyl-CoA as a substrate (Fig 7.14B). Similar results were observed for product yields (Fig 7.14C). Expression of TcAAE7, PDC and AAT had a negative impact on growth, otherwise no significant effects on growth were observed (Fig 7.14D).

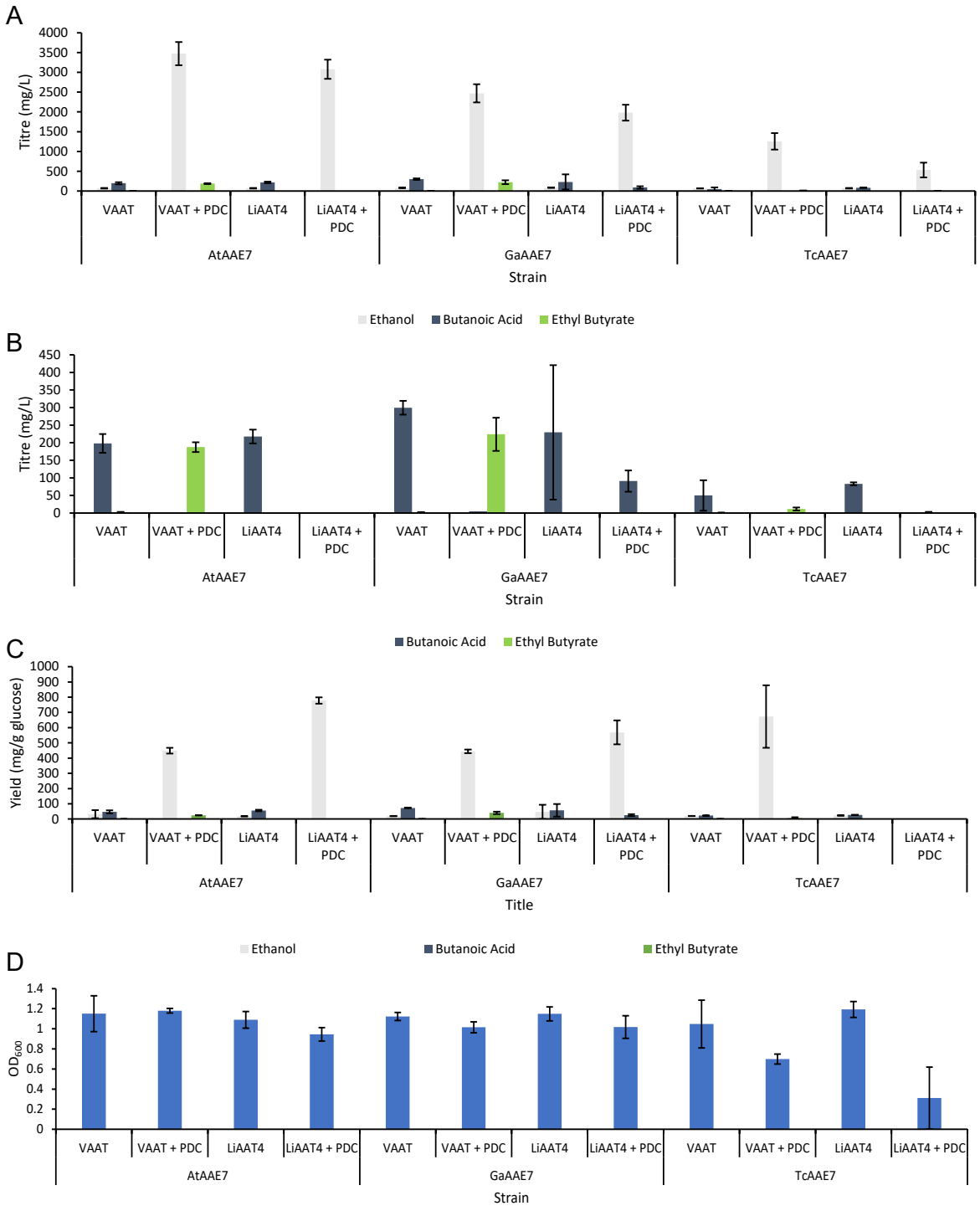


Figure 7.14 Effect of LiAAT4 and PDC expression on ethyl butyrate titre *E. coli ΔyciA were transformed with a plasmid expressing Tes4 and an AAE7, and a second plasmid expressing VAAT or LiAAT4 by themselves or with PDC A) Ethanol, butanoic acid and ethyl butyrate titres B) Butanoic acid and ethyl butyrate titres C) Yields D) OD_{600} after 48 hrs cultivation in M9 media and a 10% v/v hexadecane overlay for 48 hrs at 30°C. The overlay was sampled and analysed by GC-MS. Cells were induced with 0.5 mM IPTG. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

7.4 Results: Methyl Ketone Biosynthesis

Medium chain length methyl ketones are flavour and fragrance compounds that contribute to fruit and blue cheese aromas[327,328]. Previously, biosynthesis of medium chain length ketones has been reported using a modified β-oxidation pathway[156,347]. Medium chain length methyl ketone synthases (MKS) have been reported[336-338], which would allow for the biosynthesis of methyl ketones while using fewer heterologous genes than in the modified β-oxidation pathway. The biosynthesis of methyl ketones using a medium chain length methyl ketone synthase (MKS) followed by spontaneous decarboxylation of the β-ketofatty acids (Fig 7.15A) was therefore investigated. MKSs were expressed using a plasmid (Fig 7.15B).

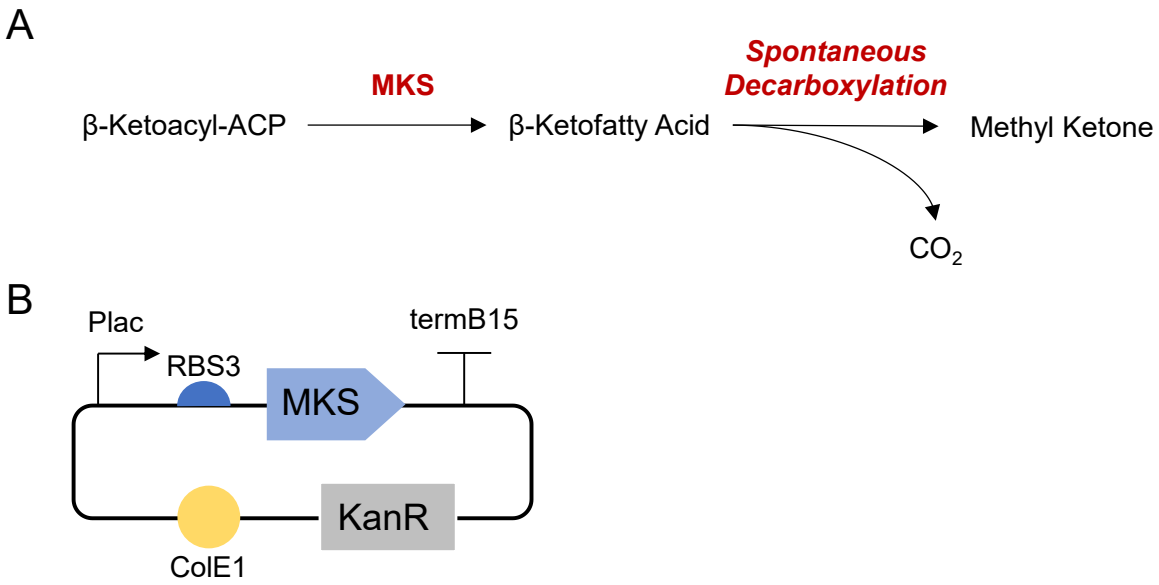


Figure 7.15 Pathway and plasmid design for methyl ketone biosynthesis *A) Synthetic metabolic pathway for methyl ketone biosynthesis. Pathway reactions are highlighted in red. MKS: Methyl ketone synthase B) Plasmid design for biosynthesis of β-ketofatty acids. A pET plasmid with ColE1 origin of replication and kanamycin resistance cassette was used. Plac and RBS3 were used to express the MKS.*

7.4.1 Could Group 2 ALTs be predicted from protein structure?

A new family of plant thioesterases, called Acyl Lipid Thioesterases (ALTs) was reported[336]. In follow up work it was found this family could be classified into four groups based on activity. Group 1 ALTs had long chain MKS activity, Group 2 ALTs had mixed activity, acting as both medium chain length acyl-ACP thioesterases and MKSs, and Group 3 and Group 4 had long chain acyl-ACP thioesterase activity. Therefore, only Group 2 ALTs were of interest. It was reported that the group an ALT belonged to could not be predicted from the ALT's sequence,

however when a protein structure based phylogeny was generated Group 2 ALTs formed a separate branch[337].

This could therefore allow for the identification of new Group 2 ALTs by bioinformatics, allowing a library to be generated and screened for new Group 2 ALTs that functioned as active and specific MKSs. To test this idea, it was attempted to recreate the reported phylogeny using the protein sequences of the characterised ALTs.

In that report[337], protein structures were generated using SWISS-MODEL and then aligned and a phylogeny generated using STAMP (Structural Alignment of Multiple Proteins) and Visual Molecular Dynamics (VMD). In this work, protein structures were instead generated using Phyre2's batch input feature[405], to allow the higher throughput generation of multiple protein structures at once. When attempting to use STAMP and VMD, error messages occurred that could not be resolved. The mtm-align server was used instead[407]. In the generated phylogeny, the Group 2 ALTs did not form a distinct branch (Fig 7.16). SWISS-MODEL[430] was then used to individually generate the predicted protein structures as in the report, but the Group 2 ALTs again did not form a distinct branch.

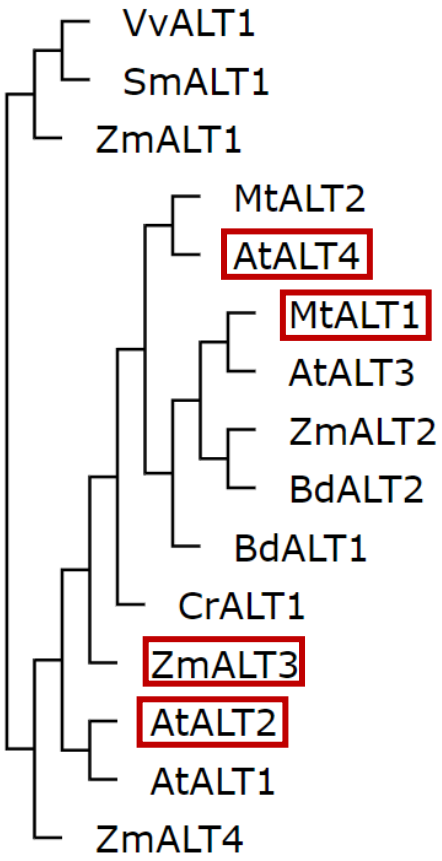


Figure 7.16 Phylogeny of predicted ALT structures *Group 2 ALTs are marked with red squares*

7.4.2 Screening Group 2 ALTs

Due to the units used (µg or nmol /unit OD) when the ALTs were characterised[336,337], from the reported data it was unclear how active the ALTs were and what product titres could be achieved upon heterologous expression of the ALTs in *E. coli*. Additionally, it was not clear from the literature whether decarboxylation of the β-ketofatty acids to methyl ketones was spontaneous or not. Furthermore, it had not been reported in the literature whether reduction of the methyl ketones to secondary alcohols occurs, which would reduce product titres.

To investigate these questions, two Group 2 ALTs, *Arabidopsis thaliana* ALT4 (AtALT2) and *Medicago trunculata* ALT1 (MtALT1) were expressed in *E. coli*. GFP was expressed as a control.

When AtALT2 was expressed, low titres of 2-heptanone (3.73 mg/L) and 2-nonanone (3.45 mg/L) were observed. In contrast when MtALT2 was expressed higher albeit still low titres of 2-pentanone were observed (6.62 mg/L). Unexpectedly 2-pentanone was also observed when GFP was expressed. No 2-heptanol or 2-nonanone was observed, however some 2-pentanol was observed (Fig 7.17A).

To examine the spontaneity of decarboxylation, the remaining β-ketofatty acids were extracted chemically and decarboxylated to methyl ketones using hot and acidic conditions. Only when expressing AtALT2 were 2-heptanone (0.61 mg/L) and 2-nonanone (0.95 mg/L) observed, and these titres were lower than those of spontaneously decarboxylated methyl ketones (Fig 7.17B).

Both AtALT2 and MtALT1 had been reported to also have medium chain length acyl-ACP thioesterase activity, however no significant differences in medium chain length fatty acid titres were observed upon their expression (Fig 7.17C).

Expression of AtALT2 led to significantly ($p = 6.53e-6$) reduced growth, while expression of MtALT2 had no effect (Fig 7.17D).

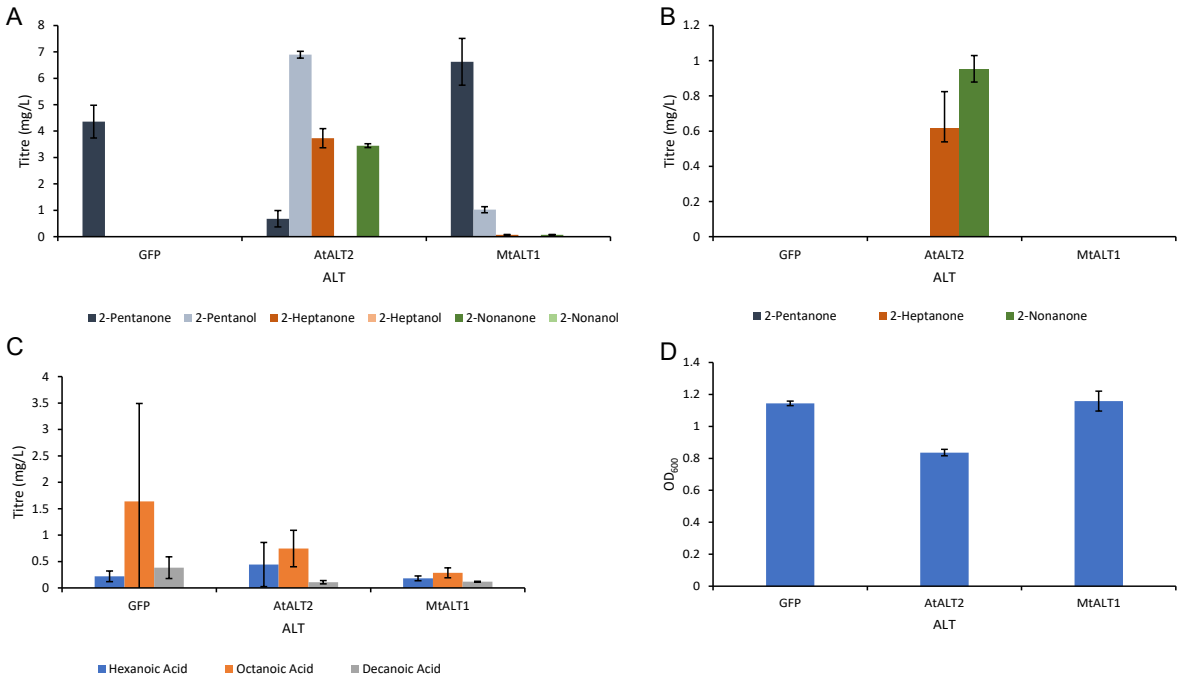


Figure 7.17 Screening Group 2 ALTs *E. coli ΔfadD were transformed with a plasmid expressing a MKS or GFP A) Methyl ketone and secondary alcohol titres B) Methyl ketone titres after extraction and chemical decarboxylation of β-ketofatty acids C) Fatty acid titres D) OD_{600} after 48 hrs cultivation in M9 media and a 10% v/v isopropyl myristate overlay for 48 hrs at 30°C. The overlay was sampled and analysed by GC-MS. β-Ketofatty acids were decarboxylated to methyl ketones using heat and acid treatment and extracted using hexane. Cells were induced with 0.5 mM IPTG. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

7.4.3 Screening YneP and YtpA

A potential medium chain length MKS, YneP, was reported in *Bacillus nematocida*[338]. Heterologous expression of YneP in *E. coli* led to biosynthesis of medium chain length methyl ketones, however titres were not quantified. Therefore, it was not clear what kind of methyl ketone titres could be expected in *E. coli* following expression of YneP.

In the same paper, a homologue of a β-ketotetradecanoic acid decarboxylase, YtpA, was identified, however the authors concluded YtpA was not active, as expression of only YtpA did not lead to methyl ketone biosynthesis[338]. I theorised that this was an incorrect conclusion. Due to its homology with a β-ketotetradecanoic acid decarboxylase, it was likely that YtpA was also a β-ketofatty acid decarboxylase. Expression of only a β-ketofatty acid decarboxylase will not lead to biosynthesis of methyl ketones, as the β-ketofatty acid precursors will be absent. The reported experiment was therefore unsuitable for testing YtpA's activity, and it was therefore incorrect of the authors to conclude YtpA is inactive, for such an experiment cannot

answer this question. I also hypothesised that if YtpA was active, YtpA would have medium chain length substrate specificity. This is because the *B. nematodica* YneP had been shown to be a medium chain length MKS, which produces medium chain length β-ketofatty acids, and it was therefore likely that the *B. nematodica* β-ketofatty acid decarboxylase had a corresponding substrate specificity.

It was therefore investigated whether expression of YneP would improve methyl ketone titres compared to expression of an ALT, and whether co-expression of a MKS with YtpA would improve methyl ketone titres.

E. coli were transformed with a plasmid expressing either GFP or a MKS as before, or a plasmid for co-expression of a MKS and YtpA (Fig 7.18C)

When only YneP was expressed biosynthesis of only 2-nonanone was observed (0.74 mg/L). When YneP was co-expressed with YtpA biosynthesis of 2-heptanone was observed, and there was a small but significant ($p = 0.00288$) increase in 2-nonanone titres (1.10 mg/L). When only AtALT2 was expressed biosynthesis of both 2-heptanone (2.79 mg/L) and 2-nonanone (2.82 mg/L) was again observed, and co-expression of YtpA again led to small but significant increases in 2-heptanone (3.18 mg/L, $p = 0.0249$) and 2-nonanone (3.73 mg/L, $p = 0.00306$) titres. Despite this, overall titres remained low (Fig 7.18A). Extraction and chemical decarboxylation of β-ketofatty acids was attempted, however no methyl ketones could be observed (data not shown). Expression of AtALT2 and YneP had negative effects on growth, however co-expression of YtpA with AtALT2 appeared to somewhat alleviate this effect (Fig 7.18B).

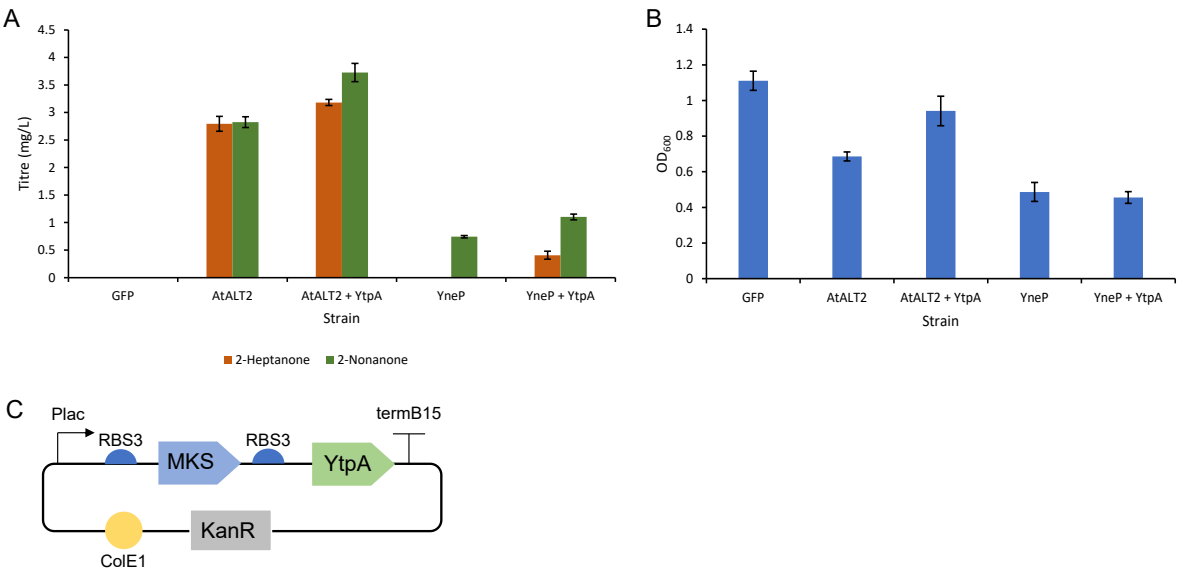


Figure 7.18 Screening YneP and YtpA *E. coli BW25113 were transformed with a plasmid expressing a GFP, a MKS or a MKS and YtpA A) Methyl ketone titres B) OD_{600} after 48 hrs cultivation in M9 media and a 10% v/v isopropyl myristate overlay for 48 hrs at 30°C. The overlay was sampled and analysed by GC-MS. Cells were induced with 0.5 mM IPTG. Data are the average of 3 biological replicates and error bars are of 1 standard deviation. C) Plasmid design for co-expression of MKS and YtpA. A pET plasmid with ColE1 origin of replication and kanamycin resistance cassette was used. Plac and RBS3 were used to express the MKS and YtpA in a single operon.*

7.5 Results: Unsaturated Alcohol Biosynthesis

Green Leaf Volatiles (GLVs) are a set of saturated and unsaturated C6 alcohols, aldehydes and esters, and are so-called because of their green odour[354-356]. Green leaf volatiles are important compounds in the food industry, where they are used to improve the fresh green odour of fruits and vegetables post-processing, and as food preservatives[22,357,369]. Biosynthesis of GLVs was therefore targeted.

A synthetic metabolic pathway was designed for the biosynthesis of *cis* 2-alkenols by adapting a previously reported pathway for the biosynthesis of unsaturated alcohols[409] (Fig 7.19A). The pathway was expressed on a single plasmid (Fig 7.19B). A key gene in this pathway was DfsA, a bifunctional β-hydroxydodecanoyl-ACP dehydrogenase and 2-dodecenoyl-ACP thioesterase. Engineering DfsA for altered substrate specificity would allow for the biosynthesis of different chain length *cis* 2-alkenols, including *cis* 2-hexenol, a stereoisomer of the GLV *trans* 2-hexenol.

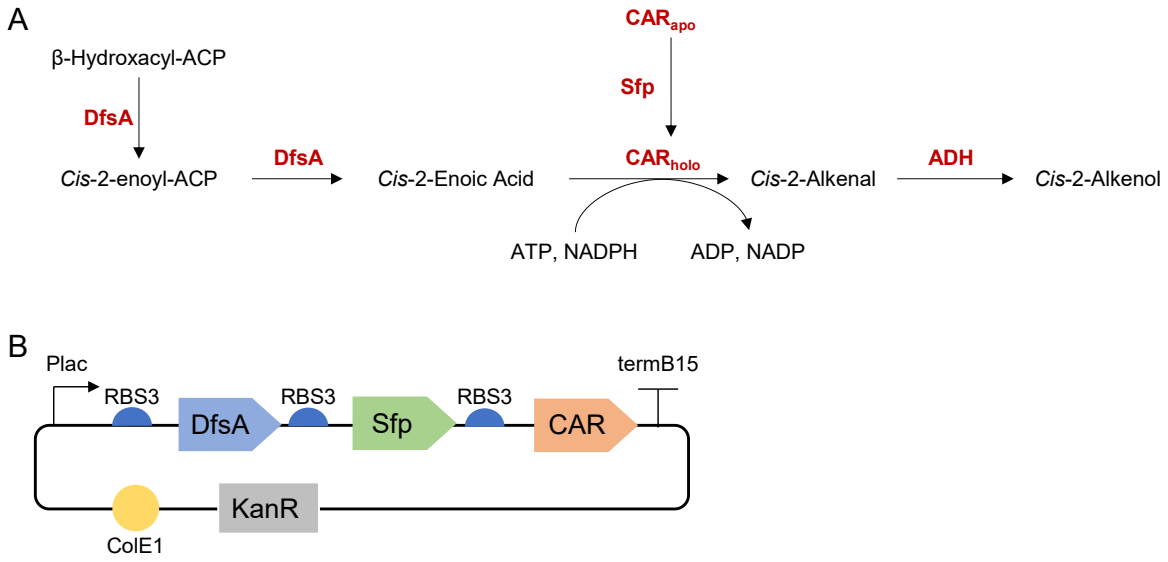


Figure 7.19 Pathway and plasmid design for cis-2-alkenol biosynthesis *A) Synthetic metabolic pathway for cis-2-alkenol biosynthesis. Pathway enzymes are highlighted in red. DfsA: Bifunctional β-hydroxydodecanoyl-ACP dehydrogenase and enoyl-ACP thioesterase (Burkholderia cenocepacia), Sfp: Phosphopantetheinyl transferase (Bacillus subtilis), CAR: Carboxylic acid reductase (Mycobacterium marinum), ADH: Endogenous alcohol dehydrogenase. B) Plasmid design for biosynthesis of cis-2-alkenols. A pET plasmid with ColE1 origin of replication and kanamycin resistance cassette were used. Plac and RBS3 were used to express DfsA, Sfp and CAR in a single operon.*

7.5.1 Screening DfsA

Expression of DfsA in *E. coli* had been reported to lead to the biosynthesis of *cis*-2-dodecenoic acid. The *cis*-2-dodecenoic acid was identified by a colourimetric assay and then confirmed by GC-MS, but titres were not quantified. Therefore, it remained unclear how active DfsA was[395]. To clarify this and confirm that transgenic expression of DfsA does lead to biosynthesis of *cis*-2-dodecnoic acid, DfsA was incorporated into the *cis*-2-alkenol biosynthesis pathway and screened for biosynthesis of 2-dodecenol.

Upon expression of DfsA, Sfp and CAR biosynthesis of 2-dodecenol could be observed, with titres reaching 22 mg/L (Fig 7.20A). 2-dodecenol has an aroma of coriander, and a strong scent of coriander was also observed when sampling the cultures. No significant difference in OD_{600} was observed between *E. coli* expressing GFP only or DfsA, Sfp and CAR (Fig 7.20B).

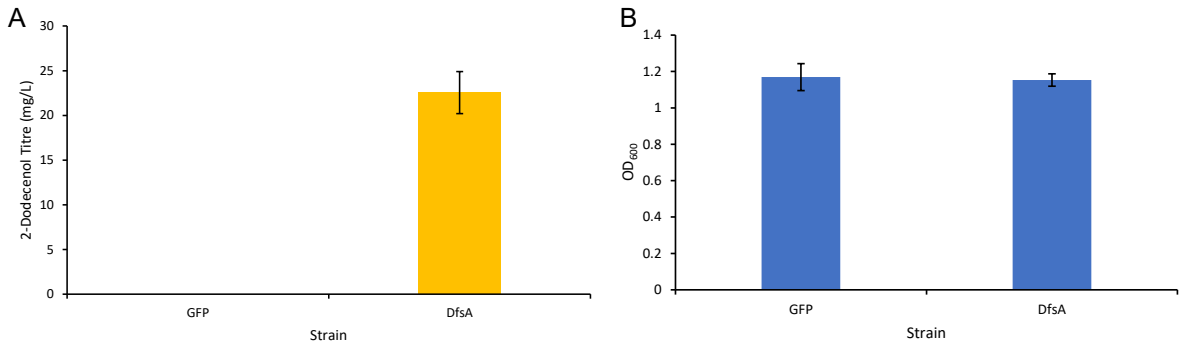


Figure 7.20 Screening DfsA *E. coli ΔfadD were transformed with a plasmid expressing a GFP, or DfsA, Sfp and CAR A) 2-dodecenol titres B) OD_{600} after 48 hrs cultivation in M9 media and a 10% v/v isopropyl myristate overlay for 48 hrs at 30°C. The overlay was sampled and analysed by GC-MS. Cells were induced with 0.5 mM IPTG. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

7.5.2 Structure guided mutagenesis of DfsA

A crystal structure of DfsA with a dodecanoic acid ligand docked in the substrate biding pocket was available on PDB (5FUS) [396]. This structure was visualised using PyMol, and residues lining the substrate binding pocket were examined. The Mutagenesis Wizard was used to mutate these residues, and potential rotamers of the mutations were examined. The rotamers were examined to see if they caused steric hinderance of the dodecanoic acid ligand after the C6 position. This would reduce the favourability of binding longer than C6 chain length substrates, shifting DfsA's specificity to be more C6 specific. Three mutations were designed, F44W (Fig 7.21A), F44Y (Fig 7.21B) and G135C (Fig 7.21C). The mutations were then

generated by site directed mutagenesis, however during the process of generating G135C, a downstream point mutation was also introduced, which resulted in a premature stop codon.

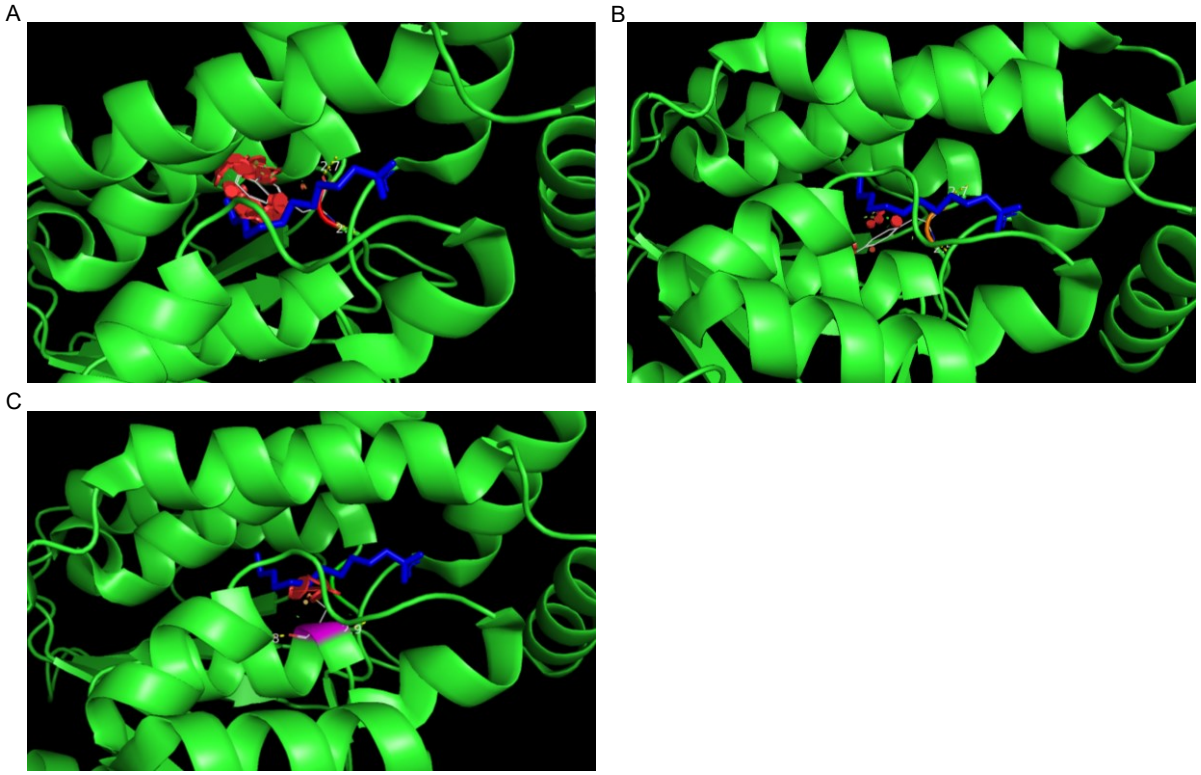


Figure 7.21 Structure guided mutagenesis of DfsA *DfsA (green) crystal structure with dodecanoic acid ligand (blue) from PDB (5FUS) was visualised using PyMol. Three mutations A) F44W (red) B) F44Y (orange) C) G135C (magenta) were designed that caused steric hinderance (red circles) of the dodecanoic acid ligand after the C6 position.*

7.5.3 Screening DfsA mutants

The DfsA mutants F44W and F44Y were screened by incorporation into the *cis*-2-alkenol biosynthesis pathway. Chemical standards were run for 2-dodecenol, 2-octenol, 2-hexenol and 2-butenol. Unfortunately, standards for 2-decenol were not available.

2-butenol biosynthesis was observed in all strains, even when only GFP was expressed, with no significant difference in 2-butenol titres between the different strains. Both DfsA mutants showed reduced 2-dodecenol titres compared to the wild type (Fig 7.22A).

Trace amounts of 2-octenol (0.07 mg/L) were observed when only GFP was expressed (Fig 7.22B), and expression of DfsA wild type, Sfp and CAR led to a small but significant ($p = 0.0500$) increase in titres (0.19 mg/L). The DfsA mutants had reduced 2-octenol titres relative to the wild type, however the differences were not significant. No biosynthesis of 2-hexenol was observed (Fig 7.22B).

When the DfsA wild type, Sfp and CAR were expressed there was a small decrease in OD_{600} relative to expression of only GFP. When a DfsA mutant was expressed instead of the wild type there was no significant effect on growth (Fig 7.22C).

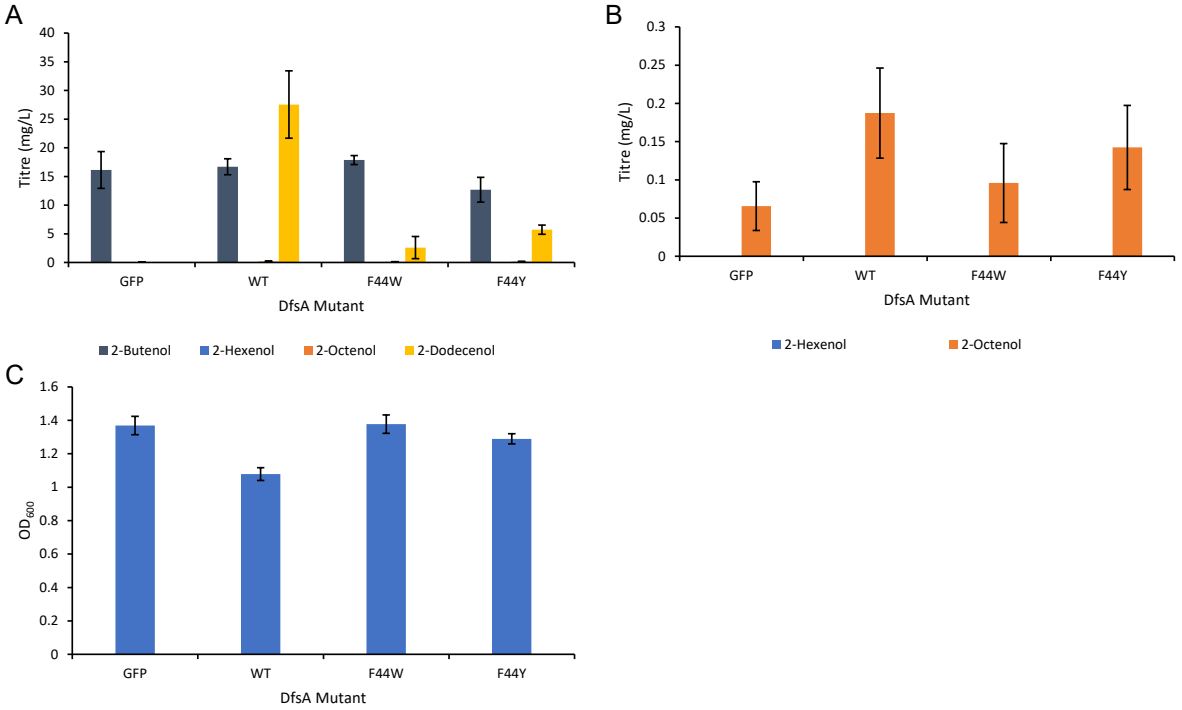


Figure 7.22 Screening DfsA mutants *E. coli ΔfadD were transformed with a plasmid expressing a GFP, or a DfsA variant, Sfp and CAR A) 2-butenol, 2-hexenol, 2-octenol and 2-dodecenol titres B) 2-hexenol and 2-octenol titres C) OD_{600} after 48 hrs cultivation in M9 media and a 10% v/v isopropyl myristate overlay for 48 hrs at 30°C. The overlay was sampled and analysed by GC-MS. Cells were induced with 0.5 mM IPTG. Data are the average of 3 biological replicates and error bars are of 1 standard deviation.*

7.6 Discussion

7.6.1 Biosynthesis of butyl acetate

In initial experiments in this work butyl acetate titres of 571.4 mg/L were observed. Butanol titres were also high, reaching 999.4 mg/L in the GFP control. However, in subsequent experiments butanol titres never reached this level, and butanol titres remained at 600 mg/L or lower despite using the same experimental conditions throughout this work. In subsequent experiments butyl acetate titres were also lower, only reaching 153 mg/L. Due to these discrepancies in titre the experiments in Sections 7.2.4 and 7.2.7 were repeated, however these repeats confirmed the reported results (data not shown). Therefore, it is likely that the initial results are unreliable. In particular, initial butanol titres in Section 7.2.1 exceeded 1 g/L, titres that were significantly higher than had been observed for any other product biosynthesised under similar conditions in the Jones lab. On the other hand, when the same plasmids used in Section 7.2.2 were used by another member of the Jones lab to transform a formate dependent strain of *E. coli*, butyl acetate titres of ~1.5 g/L were reached (Jones lab, unpublished data).

Prior to this work the highest reported titres of butyl acetate biosynthesised *de novo* from glucose were only 1.29 mg/L[254]. The butyl acetate titres in Section 7.2.7 therefore represent an over 100-fold increase in titre. However, since the experiments in this work had been completed, two new papers in 2022 reported higher *de novo* butyl acetate titres which reached 0.64 g/L[256] and 1.5 g/L prior to scale-up in a fed-batch bioreactor at which point titres of 22.8 g/L were acheived[253].

However, in both these reports TB media was used, rather than M9 as in this work. To account for these differences, product yields can be analysed. For one of the reports yield data was not provided[256], however for the other where titres of 1.5 g/L were reported, a yield of 115 mg/ g glucose could be calculated[253]. This was also higher than the butyl acetate yields achieved in this work (21.6 mg/ g glucose).

While these titres are higher than the titres achieved in this work, there is scope to improve the titres achieved in this work. The most active AAT, based on butyl acetate titres was Atf1, yet not all of the butanol was converted to butyl acetate. Moreover, product loss was observed, likely due to volatility of the butyl acetate, with total molar titres of butanol and butyl acetate when expressing Atf1 lower than butanol titres when GFP was expressed. Based on the butanol titres in the GFP strain, a theoretical maximum of ~900 mg/L butyl acetate could be achieved under these culture conditions if complete conversion of butanol to butyl acetate could be achieved, and product loss averted. Incomplete conversion of butanol may be due to

a limited supply of acetyl-CoA, so engineering *E. coli* for increased acetyl-CoA supply and recycling of the free CoA produced by the esterification reaction should be investigated. However, higher titres (768.9 mg/L) of butyl acetate in parallel with complete consumption of butanol were observed during the feeding experiment, suggesting acetyl-CoA availability was a limiting factor under these conditions. Acetyl-CoA is also required for fatty acid and therefore butanol biosynthesis, and so the biosynthesis of butanol could have competed with the esterification reaction for acetyl-CoA. Butyl acetate titres could therefore be improved by balancing expression of the pathway enzymes to ensure availability of acetyl-CoA for both butanol biosynthesis and the esterification reaction. Optimisation of intracellular NAD/NADH concentration could also be investigated to improve butyl acetate titres, as optimisation of the NAD/NADH ratio was reported to improve butyl acetate titres from 0.39 g/L to 1.5 g/L[253]. Furthermore, in the two new papers a rich TB media was used to culture the *E. coli*, rather than the minimal M9 media used in this work. It has been observed in the Jones lab that switching from M9 to TB media improves product titres (unpublished data), and so this approach could also be used to improve titres.

7.6.2 Biosynthesis of ethyl butyrate

Over the course of this work ethyl butyrate titres were improved from 14.5 mg/L to 224.0 mg/L, a 15-fold improvement. A new pathway was developed using AAE7 to ligate butanoic acid to CoA. New AAE7s were screened for butanoyl-CoA and therefore ethyl butyrate biosynthesis, however there was a lot of inconsistency within the results from one experiment to the next, with the best AAE7 being unclear as a result. In the final experiment in Section 7.3.4 overexpression of PDC was examined, as previously observed ethanol titres were low, suggesting ethanol was a limiting factor. When PDC was overexpressed with VAAT, high ethanol titres were observed and ethyl butyrate titres improved by around 90-fold, while butanoic titres dropped to zero or near-zero. This suggests that the high concentrations of ethanol improved carbon flux through the pathway, leading to nearly all the butanoic acid being ligated to CoA by the AAE7, and that under these conditions AAE7 activity was not a limiting factor. As butanoyl-CoA titres could not be measured, it is not clear if VAAT was a limiting factor. As a control of *E. coli* expressing only the C4 thioesterase Tes4 was not used in that experiment, it cannot be determined if product loss occurred. However, in Section 7.3.2 such a control was used, and butanoic acid titres of ~600 mg/L were observed, suggesting a theoretical maximum of ~900 mg/L ethyl butyrate under these conditions. Nevertheless, the experiment should be repeated with this control to determine if and how much product loss occurred.

Prior to this work, the highest reported titre of ethyl butyrate biosynthesised *de novo* was 134 mg/L[254]. The titres obtained here therefore represented a much smaller 1.5-fold increase in titre. Since this work commenced, a new paper reported an ethyl butyrate titre of 0.45 g/L biosynthesised *de novo* from glucose[256], exceeding the titres reported in this work. There is again scope to improve the titres reported in this work - the theoretical maximum titre from this pathway is also ~900 mg/L. Furthermore, switching from M9 to TB media could again be used to improve ethyl butyrate titres.

7.6.3 Potential for future biosynthesis of butyl butyrate

The two precursor compounds to butyl acetate are butanol and butanoyl-CoA, both of which were biosynthesised during the production of butyl acetate and ethyl butyrate. Three AATs were also found to be able to catalyse the production of both butyl acetate and ethyl butyrate, Atf1, SAAT and VAAT. Of these VAAT shows the most promise, as only very low titres of ethyl butyrate were observed when Atf1 or SAAT were expressed. Therefore, biosynthesis of butyl butyrate is theoretically possible with the genes represented in this work and could be an avenue for future work.

7.6.4 Activity of the library of new AATs

In this work a library of potential AATs that had been identified by a student in Jones lab was screened for butyl acetate production. However, out of this library only three of the genes, LaAAT1, LaAAT2 and LiAAT4 were annotated as alcohol-acyl transferases, and only LiAAT4 was found to be active. In follow up work by another student in the Jones lab, screening the AAT library for production of acetate esters of octanol, menthol, and eugenol, again only LiAAT4 was found to be active (unpublished data). LiAAT4 was specifically labelled as an alcohol acetyltransferase, and could not produce ethyl butyrate, suggesting LiAAT4 specifically only accepts acetyl-CoA as a substrate. Given that the majority of the library did not have activity on any of the screened substrates, in future work care should be taken to review the enzyme annotations when designing a gene library for screening.

7.6.5 Biosynthesis of methyl ketones

In this work biosynthesis of methyl ketones using the methyl ketone synthase pathway was investigated. Medium chain length MKSs had been reported in the literature, however due to the reported units (µg or nmol /unit OD)[336-338] it was not clear what kind of titres would be achieved following expression of the MKSs in *E. coli*. Upon expression of the MKSs, very low titres (>5 mg/L) of methyl ketones were observed. Negative impacts on growth were also observed when expressing some of the MKSs. This could be due to toxicity effects from the β-ketofatty acids and methyl ketones, or due to burden imposed by overexpression of the MKS.

Toxicity effects are unlikely, due to the low product titres and the presence of an overlay that extracted the methyl ketones out of the media, keeping them away from the cells. To investigate whether toxicity effects were a factor, *E. coli* could be cultured in media supplemented with different concentrations of the methyl ketones to investigate if methyl ketones are toxic, and if so at what concentrations. To investigate whether cellular burden caused by overexpression of the MKS was an issue, inducer optimisation of expression of the MKSs could be performed.

Another outstanding question was whether the decarboxylation of β-ketofatty acids to methyl ketones was spontaneous or not. Methyl ketones were detected in the overlay, indicating that spontaneous decarboxylation does occur. Lower methyl ketone titres were observed after β-ketofatty acids were extracted and chemically decarboxylated to methyl ketones, suggesting the majority of β-ketofatty acids do spontaneously decarboxylate to methyl ketones. From the literature review a potential medium chain length β-ketofatty acid decarboxylase, YtpA, was identified. Co-expression of YtpA with a MKS led to small but significant improvements in methyl ketone titre, suggesting YtpA is active, in contradiction to previous reports[338]. However, after extraction and chemical decarboxylation of β-ketofatty acids, no methyl ketones could be detected in any of the strains, meaning increases in methyl ketone titre could not be tied to a corresponding decrease in β-ketofatty acids. Despite this, from the observed results, and YtpA's homology to β-ketotetradecanoic acid decarboxylase, it can be concluded that YtpA is a medium chain length β-ketofatty acid decarboxylase.

Methyl ketone titres in this work reached only 3.18 mg/L for 2-heptanone and 3.73 mg/L for 2-nonanone. This is far lower than the highest reported titres for 2-heptanone and 2-nonanone in the literature, which in shake flasks respectively reached ~0.2 g/L and ~1.2 g/L, using the alternative β-oxidation pathway for methyl ketone biosynthesis[431]. While using the Methyl Ketone Synthase pathway has theoretical advantages, such as reduced cellular burden due to needing fewer heterologous genes, these much lower titres are a far more significant drawback to using the Methyl Ketone Synthase pathway.

Furthermore, while screening for newer more active MKSs could help alleviate these low titres, work to replicate the predicted protein structure phylogeny of ALTs was unsuccessful. This prevents identification of new Group 2 ALTs by bioinformatics, reducing the chances of a generated library containing a new more active MKSs.

Due to these challenges of not been able to identify new ALTs and much lower titres than in the literature, it is not recommended to continue pursuing this idea in future work.

7.6.7 The Application of DfsA for biosynthesis of unsaturated fatty acid derivatives

In this work DfsA was incorporated into a 2-alkenol biosynthesis pathway and used to biosynthesise 2-dodecenol. This is the first time to my knowledge that *de novo* microbial biosynthesis of 2-dodecenol has been reported, with only bioconversion of 2-dodecenal to 2-dodecenol in *Penicillium chrysogenum* previously reported[432]. 2-dodecenol is also an attractive flavour and fragrance compound, with a potent aroma of coriander. This aroma could be smelled when sampling the cultures after 48 hrs cultivation, and even emanating from LB-Agar plates with transformant *E. coli* containing the uninduced operon of DfsA, Sfp and CAR. At no other point in this work could an aroma be smelled from just the LB-Agar plate. Due to its potency, biosynthesis of 2-dodecenol could be of industrial interest in future work.

To my knowledge, the only *de novo* microbial biosynthesis of any similar unsaturated fatty acid derived product using the native fatty acid biosynthesis pathway is the biosynthesis of 2-butenoic acid, achieved by inhibiting FabI to cause accumulation of enoyl-ACP, at the cost of blocking the fatty acid biosynthetic cycle[275]. This is not required when biosynthesising 2-dodecenoic acid with DfsA. Biosynthesis of 2-enoic acids has also been reported using an engineered reversal of the β-oxidation pathway[433]. However, this approach required the overexpression of five genes to biosynthesise a 2-enoic acid, as opposed to just one using DfsA, increasing cellular burden. DfsA and its homologues therefore represent a promising route for the biosynthesis of 2-enoic acid derived compounds.

In this work it was attempted to engineer DfsA to have C6 activity, in order to biosynthesise *cis*-2-hexenol, a stereoisomer of the GLV *trans*-2-hexenol. Unfortunately, only two of the designed mutants were successfully generated. These mutants had an activity loss relative to the wild-type and no changes in specificity were detected. A small increase in 2-octenol titres was observed when the wild-type, Sfp and CAR were expressed, relative to a GFP control. This suggests that DfsA has some, even if very low, substrate specificity for β-hydroxyoctanoyl-ACP and 2-octenoyl-ACP, which had not been previously reported. Standards for 2-decenol were not available so DfsA's activity on this chain length could not be determined. Still, this indication that DfsA does accept a range of substrate chain lengths, along with the crystal structure, meaning another attempt at engineering DfsA for C6 specificity should be considered for future work. The G135C mutant that was designed but not successfully generated in this work should also be screened.

2-butenol biosynthesis was also observed, even when only GFP was expressed. This may have been due to YdiL, an endogenous butenyl-CoA thiolase of *E. coli*[433].

Potential C10 homologues of DfsA have been reported in *Pseudomonas aeruginosa*, although it is not entirely clear which gene is responsible for 2-decenoic acid biosynthesis, with two candidates[393,394]. Given that DfsA is functional in *E. coli*, it is likely that these homologues are also functional. A C10 homologue of DfsA could be used as an alternative to the C10 thioesterase in the decalactone biosynthesis pathway developed in this work to biosynthesise unsaturated derivatives of δ-decalactone and γ-decalactone. Of these, the Δ2 unsaturated δ-decalactone, known as massoia lactone, is a highly desirable target, due to having an even more potent aroma than δ-decalactone[434].

Therefore, DfsA and its homologues represent a promising avenue for future work, as they would allow for the biosynthesis of potent and desirable flavour and fragrance compounds.

7.7 Conclusions

In this chapter an existing pathway for the biosynthesis of acetate esters was tailored to biosynthesise butyl acetate, and new AAT capable of producing butyl acetate was identified. Titres achieved exceeded those in the literature at the start of the work, but were subsequently exceeded in later reports.

A novel pathway for the biosynthesis of ethyl butyrate was developed and optimised, and titres achieved exceeded those in the literature at the start of the work, but were subsequently exceeded in later reports.

Methyl ketone biosynthesis using the methyl ketone synthase pathway was investigated. Medium chain length methyl ketones from the literature were screened but titres were low. The spontaneity of decarboxylation of β-ketofatty acids was investigated and shown to be mostly spontaneous. A medium chain length β-ketofatty acid decarboxylase was also identified.

DfsA was incorporated into a pathway for 2-alkenol biosynthesis, and 2-dodecenol was biosynthesised. Subsequent engineering did not result in C6 specificity, however the wild type was shown to have some activity on different chain length substrates. DfsA and its homologues are promising avenue for future work biosynthesising unsaturated fatty acids and their derivatives, including desirable flavour and fragrance compounds.

Chapter 8

Conclusions and Recommendations for Future Work

8.1 Conclusions

Cyanobacteria biotechnology is a promising tool for the renewable production of chemicals from CO_2 and light. Currently this technology is limited by the high costs of culturing cyanobacteria, and low titres and values of the products that have been produced in cyanobacteria so far. As a result, bioproduction using cyanobacteria is commercially inviable. To help commercialise this technology, production of high-value compounds should be targeted first. Flavours and fragrances are an appealing target due to their high value and large global market. *Escherichia coli* was chosen as a host within which synthetic metabolic pathways for the biosynthesis of flavour and fragrance compounds could be prototyped and developed, as *E. coli* is well established as a host organism for synthetic biology and is easy to culture, allowing for an accelerated design-build-test-learn cycle.

γ-Decalactone and δ-decalactone were selected as target compounds as they are commercially important, having both a high value and strong aroma. As the native biosynthesis pathway is unknown, a novel synthetic pathway for their biosynthesis was proposed. This comprised of a C10 specific acyl-ACP thioesterase that produced free decanoic acid from endogenous fatty acid biosynthesis, and the CYP450 BM3, which would hydroxylate the decanoic acid at the 4- and 5-positions. The 4-hydroxydecanoic acid and 5-hydroxydecanoic would then spontaneously condense to respectively form the γ-decalactone and δ-decalactone.

In Chapter 4 an acyl-ACP thioesterase, *Cuphea viscosissima* FatB1 (CvFatB1) was engineered to be C10 specific, as no C10-specific thioesterase had yet been reported in the literature. FatB thioesterases, such as CvFatB1 require truncation to remove an N-terminal Chloroplast Localisation Peptide prior to heterologous expression in bacteria. Truncation of CvFatB1 was systematically investigated, and the optimum site for truncation located. It was also found that truncation beyond a certain point was detrimental to thioesterase activity, which had not previously been reported. It was also found that expressing the CvFatB1 truncations in the cyanobacteria *Synechocystis* Sp. PCC 6803 resulted in higher C10 specificity than in *E. coli*. Heterologous ACP expression was then investigated as a tool to modulate thioesterase specificity in *E. coli*. While altered substrate specificity was not achieved, it was discovered

that overexpression of *E. coli* ACP resulted in increased product titres, which had not previously been reported. A model of the truncation CvT5's structure was then generated, and structure guided mutagenesis was performed. The mutant CvT5 W122G showed increased C10 specificity of up to 77.1%. CvT5 W122G is therefore the most specific C10 thioesterase to have been reported.

In Chapter 5 the CYP450 BM3 was engineered to hydroxylate decanoic acid at the 4- and 5-positions, to form the 4-hydroxydecanoic acid and 5-hydroxydecanoic acid precursors to γ-decalactone and δ-decalactone. It was unclear whether the hydroxydecanoic acids would spontaneously condense to the corresponding lactones, so this was investigated. It was found that some spontaneous condensation did occur, but that condensation could be improved using acidic conditions during the product extraction process. This was the first time that acidic conditions had been shown to improve the condensation of 4- and 5-hydroxydecanoic acid to γ- and δ-decalactone. BM3 does not natively hydroxylate decanoic acid at these positions, and so multiple rounds of engineering were performed to generate novel combinatorial mutants of BM3 that incorporated previously reported activity enhancing and regioselectivity altering mutations. Two mutants, BM3 F87G L188Q and BM3 V78A F87A L188Q A328V were obtained that respectively specifically converted decanoic acid to δ-decalactone and γ-decalactone. This was the first time that BM3 had been engineered to produce lactones from fatty acids. This was also the first that a CYP450 had been used to produce a γ-lactone from a fatty acid. An alternative CYP450 from the literature that been reported to hydroxylate decanoic acid at the 5- position was investigated and found to have low activity. BM3 F87G L188Q therefore represented an improvement in the start of the art in the conversion of decanoic acid to δ-decalactone.

In Chapter 6 expression of the complete pathway and *de novo* biosynthesis of the decalactones was attempted. In *E. coli*, following expression of engineered enzymes from Chapter 4 and 5, titres of 0.53 mg/L δ-decalactone and 3.53 mg/L γ-decalactone were achieved. This was the first time *de novo* biosynthesis of lactones had been reported in *E. coli*, and the first time *de novo* biosynthesis of δ-decalactones had been reported in a microbial host. However, the pathway was limited by lower-than-expected titres of decanoic acid following expression of the thioesterase, and poor conversion of the decanoic acid to the lactones. Moreover, there were repeated analytical issues leading to large variance between experiments. Future work will be required to optimise both the biosynthesis of decanoic acid, and subsequent conversion to decalactones. With the remaining time in the project, biosynthesis of decalactones was attempted in *Synechocystis*. However, the transformant strains grew poorly, and only low titres of decanoic acid and no decalactones could be observed.

In Chapter 7 pathways for the biosynthesis of other commercially attractive flavour and fragrance compounds were investigated. Ester biosynthesis was studied, with an existing pathway adapted for the biosynthesis of butyl acetate, and a novel pathway designed for the biosynthesis of ethyl butyrate. *De novo* biosynthesis of both these compounds was achieved and following optimisation titres of butyl acetate were over 100-fold higher than the highest *de novo* titres of butyl acetate previously reported at the start of this work. Following optimisation ethyl butyrate titres were 1.5-fold higher than the highest *de novo* titres of ethyl butyrate previously reported at the start of this work. However, in the course of this work new reports of *de novo* biosynthesis of both butyl acetate and ethyl butyrate were published, where greater titres of both butyl acetate and ethyl butyrate were achieved than in this work. Biosynthesis of methyl ketones was also investigated, using medium chain length Methyl Ketone Synthases. However, these were found to have very low activity, and methyl ketone titres achieved in this work were far lower than those previously reported in the literature using an alternative pathway. Finally, biosynthesis of the green leaf volatiles 2-hexenol was attempted by engineering DfsA, a bifunctional β-hydroxydodecanoyl-ACP dehydrogenase and dodecanoyl-ACP thioesterase, and co-expressing the engineered DfsA with Sfp and CAR to biosynthesis unsaturated alcohols. While biosynthesis of 2-hexenol was not achieved, biosynthesis of 2-dodecenol was. DfsA therefore showed potential for the future biosynthesis of 2-enoic acid derived products.

8.2 Recommendations for future work

8.2.1 Improving the C10 thioesterase

In Chapter 4, the thioesterase CvT4 was found to have 71.9% C10 specificity when expressed in *Synechocystis*, and the thioesterase CvT5 W122G was found to have 77.1% specificity when expressed in *E. coli*.

While these specificities exceeded the target of 66.6% set at the start of Chapter 4, these still fall short of the highest reported specificity for a FatB thioesterase, with *Cuphea palustris* FatB1 (CpFatB1) reported to have 97.5% C8 specificity[104]. Meanwhile a C10 specificity of 70% has since been reported in *Cuphea lanceolata* FatB3 (ClFatB3) that had been truncated and had activity-enhancing mutations[125]. This was higher than the C10 specificity of CvT5 (36.3%). Therefore, there is clear scope to engineer a thioesterase with an even higher specificity than those achieved in this work. The engineering techniques used in this work to generate the CvT5 mutants should be applied to ClFatB3 to obtain improved C10 thioesterases.

While CvT5 W122G had high C10 specificity, there was also an activity loss leading to decanol titres that were lower overall compared to CvT5. Activity enhancing mutations of CpFatB1 have been reported[126], and engineering other *Cuphea* FatB thioesterases to incorporate these mutations has also led to increased activity[28,125]. Adding these mutations to CvT5 W122G to improve its activity should therefore be investigated in future work.

8.2.2 Assaying the effect of ACP expression in *Synechocystis*

In Chapter 4 overexpression of *E. coli* ACP was found to substantially increase product titres in *E. coli*. One of the limiting factors for the development of bioproduction using cyanobacteria is the low product titres. Overexpression of endogenous ACP or expression of a heterologous ACP has not been reported in cyanobacteria, however if ACP expression could be used in *Synechocystis* as in *E. coli* to improve titres of fatty acid derived products, this could assist in the development of this technology. Plasmid constructs have already been prepared for the co-expression of CvT4, CvT5 and CvT6 with EcACP, 6803ACP and ClACP2 in *Synechocystis*. Generation and screening of the appropriate strains should therefore be performed to investigate this.

8.2.3 Improving BM3 activity

In Chapter 5 and 6, incomplete consumption of decanoic acid by BM3 was observed. Moreover, only a minority of the consumed decanoic acid was converted to decalactones, suggesting there was either off-target regioselectivity or incomplete condensation of the 4-hydroxydecanoic acid and 5-hydroxydecanoic acid to γ-decalactone and δ-decalactone, or

the loss of the product through volatilisation. There is therefore a need to further improve BM3's activity.

One of the challenges in investigating condensation of the hydroxyfatty acids to lactones was the unavailability of appropriate chemical standards. Acidification of the samples during the product extraction process was found to increase decalactone titre, as this catalysed the condensation reaction. The effect of acidification upon condensation should be further investigated, with different amounts of acid used to catalyse the condensation reaction. If decalactone titres further increase before plateauing as the amount of acid increases, this would indicate that complete condensation of the hydroxyfatty acids to decalactones has occurred.

A challenge in engineering BM3 was that the required regioselectivity altering mutations also caused a loss of activity[166]. While activity enhancing mutations from the literature[168,169,177,426] were screened, these were not sufficient, and in the *de novo* biosynthesis experiments the combinatorial BM3 mutants showed reduced consumption of decanoic acid compared to the wild-type. In future work there is therefore a need to continue engineering these BM3 mutants for improved activity. Recently a high-throughput colourimetric assay that can detect γ-butyrolactone has been reported[428]. Pairing this high-throughput assay with random mutagenesis of BM3 should be investigated to identify novel activity enhancing mutations. Moreover, as this assay will detect lactones, it will allow for the screening of mutants that are both active and have the appropriate regioselectivity for the biosynthesis of lactones.

In Chapter 5, supplementation of the M9 media with iron, 5-aminoleuvinic acid (ALA) and caffeine was investigated, based on prior reports in the literature. Of these, only iron supplementation was found to improve BM3 activity. However, for iron and ALA supplementation, only a single concentration point was investigated. This media optimisation should be investigated in more depth, by examining a range of different concentrations of iron and ALA.

8.2.4 Expanding the lactone biosynthesis pathway to other chain length products

Adapting the synthetic lactone biosynthesis pathway to the production of octalactones and dodecalactones should also be investigated in future work. One of the challenges with biosynthesising decalactones was the low activity of the BM3 mutants. However, wild-type BM3 is more than 10-fold more active on dodecanoic acid than decanoic acid[164]. Therefore, while the necessary regioselectivity altering mutations will still cause a loss of activity, this will not be as great an issue as the baseline activity is higher. C12 specific acyl-ACP thioesterases are also available[102]. An undergraduate student under my supervision has already started this

work and was able to biosynthesise γ-dodecalactone and δ-dodecalactone in *E. coli* using the adapted pathway.

Biosynthesis of octalactones should also be investigated in future work. While BM3 natively has 30-fold lower activity on octanoic acid then decanoic acid[164], BM3 can natively hydroxylate octanoic acid at the 5- position, producing the 5-hydroxyoctanoic acid precursor to δ-octalactone. This means that regioselectivity altering mutations will not be required, avoiding the corresponding activity loss. C8 specific acyl-ACP thioesterase are also available.

8.2.5 Biosynthesis of unsaturated fatty acid derived products

In Chapter 7, engineering of DfsA in order to biosynthesise 2-hexenol was unsuccessful. However, one of the designed mutants, DfsA G135C could not be obtained in time and so was not screened. Firstly, this mutant should be generated and screened. While the wild-type DfsA could be used to biosynthesise 2-dodecenol, titres of only 22 mg/L were achieved. Homologues of DfsA have been reported[386,390], and these should be investigated for improved activity relative to DfsA. Should a more active homologue be identified, engineering of the homologue for C6 specificity can then be investigated.

Beyond 2-hexenol, biosynthesis of other unsaturated fatty acid derived products should be investigated. Putative C10 specific homologues of DfsA have been reported[393,394], and should be investigated to confirm their activity. If active, co-expression of these C10 homologues with combinatorial BM3 mutants from this work could be used to biosynthesise massoia lactone, a Δ2 derivative of δ-decalactone. Massoia lactone is of commercial interest as it has a much higher impact aroma than δ-decalactone, with an odour threshold value that is over thirty times lower[434].

8.2.6 Porting synthetic metabolic pathways to cyanobacteria

In this work, transfer of the synthetic metabolic pathway to *Synechocystis* was unsuccessful, likely due to poor growth of these strains. The transfer should be repeated, and the C12 adaptation of the pathway transferred to *Synechocystis*. This will confirm that the pathway is functional in this heterologous host. Furthermore, the unsaturated alcohol biosynthesis pathway using DfsA should also be transferred to *Synechocystis* to verify that DfsA is functional in this heterologous host. Not only is it important that the pathways prototyped in *E. coli* are functional in *Synechocystis,* but it is also likely that further development of the pathways within *Synechocystis*, to account for differences between the two hosts, will be required to achieve commercially viable titres.

8.2.7 Development of cyanobacteria biotechnology

In this work, the focus was on the prototyping and development of synthetic metabolic pathways for the biosynthesis of flavour and fragrance compounds. The rationale for targeting production of these compounds was that their high price would enable the commercialisation of cyanobacteria biotechnology, and that in the future production of cheaper compounds, such as biofuels could be targeted. While further development of these pathways is still needed, along with development of pathways for other industrially relevant chemicals, this is not the only challenge that needs to be overcome to develop this technology.

There are many challenges when engineering cyanobacteria, of which the most significant is their slow growth rate, which in this work led to the pathway prototyping and development being performed in *E. coli* for a faster DBTL cycle. This slow growth rate is not only a hinderance to technological development, but will also be a hinderance during biomanufacturing using cyanobacteria, as it will increase the time required to grow the cultures to the necessary ODs for bioproduction. Recently, new fast-growing strains of cyanobacteria with doubling times of around 2 hrs have been reported[435,436]. The development of these strains as chassis for metabolic engineering and synthetic biology should be investigated to overcome this challenge of slow growth.

Another challenge in the engineering of cyanobacteria is their genetic instability[437], which leads to loss of expression of heterologous pathways over long term cultivation. This must be investigated, to understand the cause of this instability, and to ensure that heterologous pathway expression is maintained during long-term continuous cultivation. If this is not resolved, then cyanobacteria will be limited to batch cultivation, which given their current slow growth rate, is likely to be inefficient and limit productivity.

Beyond the engineering of cyanobacteria, further development of the bioreactors used for their cultivation will be required. Currently growing cyanobacteria in a bioreactor is energy intensive, and as most energy is currently produced using fossil fuels, will therefore lead to GHG emissions, reducing any potential benefits of using cyanobacteria biotechnology to remove CO_2. To mitigate this, development of more energy-efficient bioreactors should be explored, along with powering bioproduction facilities with renewable energy to avoid incidental GHG emissions.

8.2.8 Transition to a bio-based economy

The development of this cyanobacteria bioproduction biotechnology needs to be performed concurrently with a transition to a bio-based economy[438]. Under such an economy, microbial production, green chemistry, and advances in chemical and physical processing are combined

to produce cost-effective biorefineries. The entire value chain would be bio-based and allow for the renewable production of goods. Beyond sustainable chemicals, land should be used more efficiently, freeing up space to expanded protected areas for the conservation of biodiversity. Furthermore, to avoid the creation of new economically left-behind areas, these new biorefineries should be distributed to provide economic benefits and opportunities across a wide-range of local areas, rather than being concentrated in a few locations. To achieve such a transition strong governmental support will be required, such as by subsidising development of key technologies or creating favourable conditions for private investment. Governments should also oppose further development of the fossil fuel industry, such as by withholding permission for drilling new oil and gas wells.

References

1 Jiang, W., Qiao, J. B., Bentley, G. J., Liu, D. & Zhang, F. Modular pathway engineering for the microbial production of branched-chain fatty alcohols. *Biotechnology Biofuels* **10**, 244 (2017). https://doi.org:10.1186/s13068-017-0936-4

2 Liao, J. C., Mi, L., Pontrelli, S. & Luo, S. Fuelling the future: microbial engineering for the production of sustainable biofuels. *Nature Reviews Microbiology* **14**, 288-304 (2016). https://doi.org:10.1038/nrmicro.2016.32

3 Lee, S. Y. *et al.* A comprehensive metabolic map for production of bio-based chemicals. *Nature Catalysis* **2**, 18-33 (2019). https://doi.org:10.1038/s41929-018-0212-4

4 Armstrong McKay, D. I. *et al.* Exceeding 1.5 degrees C global warming could trigger multiple climate tipping points. *Science* **377**, eabn7950 (2022). https://doi.org:10.1126/science.abn7950

5 Attribution, W. W. *Without human-caused climate change temperatures of 40°C in the UK would have been extremely unlikely*, 2022).

6 Lennen, R. M. & Pfleger, B. F. Microbial production of fatty acid-derived fuels and chemicals. *Current Opinions in Biotechnology* **24**, 1044-1053 (2013). https://doi.org:10.1016/j.copbio.2013.02.028

7 SÁ, A. G. A., Meneses, A. C. d., Araújo, P. H. H. d. & Oliveira, D. d. A review on enzymatic synthesis of aromatic esters used as flavor ingredients for food, cosmetics and pharmaceuticals industries. *Trends in Food Science & Technology* **69**, 95-105 (2017). https://doi.org:10.1016/j.tifs.2017.09.004

8 Koizumi, T. Biofuels and food security. *Renewable and Sustainable Energy Reviews* **52**, 829-841 (2015).

9 Knoot, C. J., Ungerer, J., Wangikar, P. P. & Pakrasi, H. B. Cyanobacteria: Promising biocatalysts for sustainable chemical production. *Journal of Biological Chemistry* **293**, 5044-5052 (2018). https://doi.org:10.1074/jbc.R117.815886

10 Lau, N. S., Matsui, M. & Abdullah, A. A. Cyanobacteria: Photoautotrophic Microbial Factories for the Sustainable Synthesis of Industrial Products. *Biomed Research International* **2015**, 754934 (2015). https://doi.org:10.1155/2015/754934

11 Branco Dos Santos, F., Du, W. & Hellingwerf, K. J. *Synechocystis*: Not Just a Plug-Bug for CO2, but a Green *E. coli*. *Frontiers in Bioengineering and Biotechnology* **2**, 36 (2014). https://doi.org:10.3389/fbioe.2014.00036

12 Kanno, M., Carroll, A. L. & Atsumi, S. Global metabolic rewiring for improved CO2 fixation and chemical production in cyanobacteria. *Nature Communications* **8**, 14724 (2017). https://doi.org:10.1038/ncomms14724

13 Yunus, I. S. *et al.* Synthetic metabolic pathways for photobiological conversion of CO2 into hydrocarbon fuel. *Metabolic Engineering* **49**, 201-211 (2018). https://doi.org:10.1016/j.ymben.2018.08.008

14 Kallio, P., Pasztor, A., Thiel, K., Akhtar, M. K. & Jones, P. R. An engineered pathway for the biosynthesis of renewable propane. *Nature Communications* **5**, 4731 (2014). https://doi.org:10.1038/ncomms5731

15 Berger, R. G. Biotechnology of flavours--the next generation. *Biotechnology Letters* **31**, 1651-1659 (2009). https://doi.org:10.1007/s10529-009-0083-5

16 Carroll, A. L., Desai, S. H. & Atsumi, S. Microbial production of scent and flavor compounds. *Current Opinions in Biotechnology* **37**, 8-15 (2016). https://doi.org:10.1016/j.copbio.2015.09.003

17 Serra, S., Fuganti, C. & Brenna, E. Biocatalytic preparation of natural flavours and fragrances. *Trends in Biotechnology* **23**, 193-198 (2005). https://doi.org:10.1016/j.tibtech.2005.02.003

18 Krings, U. & Berger, R. G. Biotechnological production of flavours and fragrances. *Applied Microbiology and Biotechnology* **49**, 1-8 (1998).

19 Panten, J. & Surburg, H. in *Ullmann's Encyclopedia of Industrial Chemistry* (2015).

20 Schempp, F. M., Drummond, L., Buchhaupt, M. & Schrader, J. Microbial Cell Factories for the Production of Terpenoid Flavor and Fragrance Compounds. *Journal of Agricultural and Food Chemistry* **66**, 2247-2258 (2018). https://doi.org:10.1021/acs.jafc.7b00473

21 Vandamme, E. J. & Soetaert, W. Bioflavours and fragrances via fermentation and biocatalysis. *Journal of Chemical Technology & Biotechnology* **77**, 1323-1332 (2002). https://doi.org:10.1002/jctb.722

22 Schwab, W., Davidovich-Rikanati, R. & Lewinsohn, E. Biosynthesis of plant-derived flavor compounds. *Plant Journal* **54**, 712-732 (2008). https://doi.org:10.1111/j.1365-313X.2008.03446.x

23 Byrne, B. & Rowe, D. in *Perfumer & Flavorist* Vol. 34 36-46 (2009).

24 El Hadi, M. A., Zhang, F. J., Wu, F. F., Zhou, C. H. & Tao, J. Advances in fruit aroma volatile research. *Molecules* **18**, 8200-8229 (2013). https://doi.org:10.3390/molecules18078200

25 Song, J. & Forney, C. F. Flavour volatile production and regulation in fruit. *Canadian Journal of Plant Science*, 537-550 (2008).

26 Klee, H. J. & Tieman, D. M. The genetics of fruit flavour preferences. *Nature Reviews Genetics* **19**, 347-356 (2018). https://doi.org:10.1038/s41576-018-0002-5

27 Pontrelli, S. *et al. Escherichia coli* as a host for metabolic engineering. *Metabolic Engineering* **50**, 16-46 (2018). https://doi.org:10.1016/j.ymben.2018.04.008

28 Sattayawat, P., Yunus, I. S. & Jones, P. R. Bioderivatization as a concept for renewable production of chemicals that are toxic or poorly soluble in the liquid phase. *Proceedings of the National Academy of Sciences* **117**, 1404-1413 (2020). https://doi.org:10.1073/pnas.1914069117

29 Blattner, F. R. *et al.* The Complete Genome Sequence of *Escherichia coli* K-12. *Science* **277**, 1453-1462 (1997).

30 Hayashi, K. *et al.* Highly accurate genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110. *Molecular Systems Biology* **2**, 2006 0007 (2006). https://doi.org:10.1038/msb4100049

31 Karp, P. D. *et al.* Multidimensional annotation of the *Escherichia coli* K-12 genome. *Nucleic Acids Research* **35**, 7577-7590 (2007). https://doi.org:10.1093/nar/gkm740

32 Riley, M. *et al. Escherichia coli* K-12: a cooperatively developed annotation snapshot--2005. *Nucleic Acids Research* **34**, 1-9 (2006). https://doi.org:10.1093/nar/gkj405

33 Karp, P. D. *et al.* The EcoCyc Database. *Nucleic Acids Research* **30**, 56-58 (2002).

34 Keseler, I. M. *et al.* EcoCyc: a comprehensive database resource for *Escherichia coli*. *Nucleic Acids Research* **33**, D334-337 (2005). https://doi.org:10.1093/nar/gki108

35 Keseler, I. M. *et al.* EcoCyc: a comprehensive database of *Escherichia coli* biology. *Nucleic Acids Research* **39**, D583-590 (2011). https://doi.org:10.1093/nar/gkq1143

36 Keseler, I. M. *et al.* EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Research* **41**, D605-612 (2013). https://doi.org:10.1093/nar/gks1027

37 Keseler, I. M. *et al.* The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic Acids Research* **45**, D543-D550 (2017). https://doi.org:10.1093/nar/gkw1003

38 Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology* **2**, 2006 0008 (2006). https://doi.org:10.1038/msb4100050

39 Dower, W. J., Miller, J. F. & Ragsdale, C. W. High efficiency transformation of *E. coli* by high volatage electroporation. *Nucleic Acids Research* **16** (1988).

40 Cohen, S. N., Chang, A. C. Y. & Hsu, L. Nonchromosomal Antibiotic Resistance in Bacteria: Genetic Transformation of *Escherichia coli* by R-Factor DNA. *Proceedings of the National Academy of Sciences* **69**, 2110-2114 (1972).

41 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences* **97**, 6640-6645 (2000).

42 Kuhlman, T. E. & Cox, E. C. Site-specific chromosomal integration of large synthetic constructs. *Nucleic Acids Research* **38**, e92 (2010). https://doi.org:10.1093/nar/gkp1193

43 Jiang, Y. *et al.* Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Applied and Environmental Microbiology* **81**, 2506-2514 (2015). https://doi.org:10.1128/AEM.04023-14

44 Clomburg, J. M. & Gonzalez, R. Biofuel production in *Escherichia coli*: the role of metabolic engineering and synthetic biology. *Applied Microbiology and Biotechnology* **86**, 419-434 (2010). https://doi.org:10.1007/s00253-010-2446-1

45 Cameron, D. E., Bashor, C. J. & Collins, J. J. A brief history of synthetic biology. *Nature Reviews Microbiology* **12**, 381-390 (2014). https://doi.org:10.1038/nrmicro3239

46 Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339-342 (2000).

47 Elowtiz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335-338 (2000).

48 Anderson, J. C., Voigt, C. A. & Arkin, A. P. Environmental signal integration by a modular AND gate. *Molecular Systems Biology* **3**, 133 (2007). https://doi.org:10.1038/msb4100173

49 Becker, J. & Wittmann, C. Systems metabolic engineering of *Escherichia coli* for the heterologous production of high value molecules-a veteran at new shores. *Current Opinion in Biotechnology* **42**, 178-188 (2016). https://doi.org:10.1016/j.copbio.2016.05.004

50 Matsumoto, T., Tanaka, T. & Kondo, A. Engineering metabolic pathways in *Escherichia coli* for constructing a "microbial chassis" for biochemical production. *Bioresource Technology* **245**, 1362-1368 (2017). https://doi.org:10.1016/j.biortech.2017.05.008

51 Ward, V. C. A., Chatzivasileiou, A. O. & Stephanopoulos, G. Metabolic engineering of *Escherichia coli* for the production of isoprenoids. *FEMS Microbiology Letters* **365** (2018). https://doi.org:10.1093/femsle/fny079

52 Yu, C., Cao, Y., Zou, H. & Xian, M. Metabolic engineering of *Escherichia coli* for biotechnological production of high-value organic acids and alcohols. *Applied Microbiology and Biotechnology* **89**, 573-583 (2011). https://doi.org:10.1007/s00253-010-2970-z

53 Bang, H. B., Lee, Y. H., Kim, S. C., Sung, C. K. & Jeong, K. J. Metabolic engineering of *Escherichia coli* for the production of cinnamaldehyde. *Microbial Cell Factories* **15**, 16 (2016). https://doi.org:10.1186/s12934-016-0415-9

54 Liu, X., Lin, J., Hu, H., Zhou, B. & Zhu, B. Site-specific integration and constitutive expression of key genes into *Escherichia coli* chromosome increases shikimic acid yields. *Enzyme and Microbial Technology* **82**, 96-104 (2016). https://doi.org:10.1016/j.enzmictec.2015.08.018

55 Lin, Z. *et al.* Metabolic engineering of *Escherichia coli* for the production of riboflavin. *Microbial Cell Factories* **13** (2014).

56 Zurbriggen, A., Kirst, H. & Melis, A. Isoprene Production Via the Mevalonic Acid Pathway in *Escherichia coli* (Bacteria). *BioEnergy Research* **5**, 814-828 (2012). https://doi.org:10.1007/s12155-012-9192-4

57 Sun, Y., Wu, S., Fu, X., Lai, C. & Guo, D. De novo biosynthesis of τ-cadinol in engineered *Escherichia coli*. *Bioresources and Bioprocessing* **9** (2022). https://doi.org:10.1186/s40643-022-00521-7

58 Tan, Z. *et al.* Engineering of *E. coli* inherent fatty acid biosynthesis capacity to increase octanoic acid production. *Biotechnology and Biofuels* **11**, 87 (2018). https://doi.org:10.1186/s13068-018-1078-z

59 Lan, E. I., Dekishima, Y., Chuang, D. S. & Liao, J. C. Metabolic engineering of 2-pentanone synthesis in *Escherichia coli*. *AIChE Journal* **59**, 3167-3175 (2013). https://doi.org:10.1002/aic.14086

60 Dong, X. *et al.* Metabolic engineering of *Escherichia coli* W3110 to produce L-malate. *Biotechnology and Bioengineering* **114**, 656-664 (2017). https://doi.org:10.1002/bit.26190

61 Chan, D. I. & Vogel, H. J. Current understanding of fatty acid biosynthesis and the acyl carrier protein. *Biochemical Journal* **430**, 1-19 (2010). https://doi.org:10.1042/BJ20100462

62 JanSen, H. J. & Steinbuchel, A. Fatty acid synthesis in *Escherichia coli* and its applications towards the production of fatty acid based biofuels. *Biotechnology for Biofuels* **7** (2014).

63 Choi-Rhee, E. & Cronan, J. E. The biotin carboxylase-biotin carboxyl carrier protein complex of *Escherichia coli* acetyl-CoA carboxylase. *Journal of Biological Chemistry* **278**, 30806-30812 (2003). https://doi.org:10.1074/jbc.M302507200

64 Cronan, J. E. & Waldrop, G. L. Multi-subunit acetyl-CoA carboxylases. *Progress in Lipid Research* **41**, 407-435 (2002).

65 Joshi, V. C. Mechanisms of Malonyl-Coenzyme A-Acyl-Carrier Protein Transacylase. *Proceedings of the Biochemical Society* **128**, 43-44 (1971).

66 Verwoert, I. I. G. S., Verbree, E. C., Van der Linden, K. H., Nijkamp, H. J. J. & Stuitje, A. R. Cloning, Nucleotide Sequence, and Expression of the *Escherichia coli fabD* Gene, Encoding Malonyl Coenzyme A-Acyl Carrier Protein Transacylase. *Journal of Bacteriology* **174**, 2851-2857 (1992).

67 Tsay, J. T., Oh, W., Larson, T. J., Jackowski, S. & Rock, C. O. Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (*fabH*) from *Escherichia col*i K-12. *Journal of Biological Chemistry* **267**, 6807-6814 (1992). https://doi.org:10.1016/s0021-9258(19)50498-7

68 Alberts, A. W., Majerus, P. W., Talamo, B. & Vagelos, P. R. Acyl-Carrier Protein. Intermediary Reactions of Fatty Acid Synthesis. *Biochemistry* **3**, 1563-1571 (1964).

69 Heath, R. J. & Rock, C. O. Roles of the FabA and FabZ beta-hydroxyacyl-acyl carrier protein dehydratases in *Escherichia coli* fatty acid biosynthesis. *Journal of Biological Chemistry* **271**, 27795-27801 (1996). https://doi.org:10.1074/jbc.271.44.27795

70 Kass, L. R. & Bloch, K. On The Enzymatic Synthesis of Unsaturated Fatty Acids in *Escherichia coli*. *Biochemistry* **58**, 1168-1173 (1967).

71 Heath, R. J. & Rock, C. O. Enoyl-Acyl Carrier Protein Reductase (*fabI*) Plays a Determinant Role in Completing Cycles of Fatty Acid Elongation in *Escherichia coli*. *Journal of Biological Chemistry* **270**, 26538-26542 (1995).

72 Edwards, P., Sabo Nelsen, J., Metz, J. G. & Dehesh, K. Cloning of the *fabF* gene in an expression vector and in vitro characterization of recombinant *fabF* and *fabB* encoded enzymes from *Escherichia coli*. *FEBS Letters* **402**, 62-66 (1997). https://doi.org:10.1016/s0014-5793(96)01437-8

73 Feng, Y. & Cronan, J. E. *Escherichia coli* unsaturated fatty acid synthesis: complex transcription of the *fabA* gene and in vivo identification of the essential reaction catalyzed by FabB. *Journal of Biological Chemistry* **284**, 29526-29535 (2009). https://doi.org:10.1074/jbc.M109.023440

74 Black, P. N., DiRusso, C. C., Metzger, A. K. & Heimert, T. L. Cloning, sequencing, and expression of the *fadD* gene of *Escherichia coli* encoding acyl coenzyme A synthetase. *Journal of Biological Chemistry* **267**, 25513-25520 (1992). https://doi.org:10.1016/s0021-9258(19)74070-8

75 Morgan-Kiss, R. M. & Cronan, J. E. The *Escherichia coli fadK* (*ydiD*) gene encodes an anerobically regulated short chain acyl-CoA synthetase. *Journal of Biological Chemistry* **279**, 37324-37333 (2004). https://doi.org:10.1074/jbc.M405233200

76 Campbell, J. W. & Cronan, J. E., Jr. The enigmatic *Escherichia coli fadE* gene is *yafH*. *Journal of Bacteriology* **184**, 3759-3764 (2002). https://doi.org:10.1128/JB.184.13.3759-3764.2002

77 Yang, S., Li, J., He, X., Cosloy, S. D. & Schulz, H. Evidence that the *fadB* Gene of the *fadAB* Operon of *Escherichia coli* Encodes 3-Hydroxyacyl-Coenzyme A (CoA) Epimerase, delta3-cis-delta2-*trans*-Enoyl-CoA Isomerase, and Enoyl-CoA Hydratase in Addition to 3-Hydroxyacyl-CoA Dehydrogenase. *Journal of Bacteriology* **170**, 2543-2548 (1988).

78 Spratt, S. K., Black, P. N., Ragozzino, M. M. & Nunn, W. Cloning, Mapping, and Expression of Genes Involved in the Fatty Acid-Degradative Multienzyme Complex of *Escherichia coli*. *Journal of Bacteriology* **158**, 535-542 (1984).

79 Gawdzik, B., Kamizela, A. & Szyskowska, A. in *CHEMIK* Vol. 69 342-349 (2015).

80 Sanchez-Sevilla, J. F., Cruz-Rus, E., Valpuesta, V., Botella, M. A. & Amaya, I. Deciphering gamma-decalactone biosynthesis in strawberry fruit using a combination of genetic mapping, RNA-Seq and eQTL analyses. *BMC Genomics* **15** (2014).

81 Marella, E. R. *et al.* A single-host fermentation process for the production of flavor lactones from non-hydroxylated fatty acids. *Metabolic Engineering* **61**, 427-436 (2020). https://doi.org:10.1016/j.ymben.2019.08.009

82 Xiao, Z., Ma, L., Zhao, J. Y., Zhong, H. & Sun, Y. 5-Hydroxy-gamma-decalactone production by *Bacillus* sp. 1s-1 and its complete genome sequence. *Journal of Biotechnology* **286**, 68-70 (2018). https://doi.org:10.1016/j.jbiotec.2018.07.033

83 Garcia, E. E., Belin, J. M. & Wache, Y. Use of a Doehlert factorial design to investigate the effects of pH and aeration on the accumulation of lactones by *Yarrowia lipolytica*. *Journal of Applied Microbiology* **103**, 1508-1515 (2007). https://doi.org:10.1111/j.1365-2672.2007.03379.x

84 Schrader, J., Etschmann, M. M. W., Sell, D., Hilmer, J. M. & Rabenhorst, J. Applied biocatalysis for the synthesis of natural flavour compounds-current industrial processes and future prospects. *Biotechnology Letters* **26**, 463-472 (2004).

85 Wache, Y. *et al.* Role of beta-oxidation enzymes in gamma-decalactone production by the yeast *Yarrowia lipolytica*. *Applied and Environmental Microbiology* **67**, 5700-5704 (2001). https://doi.org:10.1128/AEM.67.12.5700-5704.2001

86 Welsh, F. W., Murray, W. D., Williams, R. E. & Katz, I. Microbiological and Enzymatic Production of Flavor and Fragrance Chemicals. *Critical Reviews in Biotechnology* **9**, 105-169 (1989). https://doi.org:10.3109/07388558909040617

87 Sanchez, G. *et al.* An integrative "omics" approach identifies new candidate genes to impact aroma volatiles in peach fruit. *BMC Genomics* **14** (2013).

88 Kothari, S. D., Vadgama, R. N., Bhat, K. H., Lali, A. M. & Odaneth, A. A. Process optimization for production and purification of γ-decalactone from ricinoleic acid using *Yarrowia lipolytica* NCIM 3590. *Biocatalysis and Agricultural Biotechnology* **39** (2022). https://doi.org:10.1016/j.bcab.2022.102285

89 Małajowicz, J., Nowak, D., Fabiszewska, A. & Iuliano, A. Comparison of gamma-decalactone biosynthesis by yeast *Yarrowia lipolytica* MTLY40-2p and W29 in batch-cultures. *Biotechnology & Biotechnological Equipment* **34**, 330-340 (2020). https://doi.org:10.1080/13102818.2020.1749528

90 Silva, R. *et al.* Metabolic engineering of *Ashbya gossypii* for deciphering the de novo biosynthesis of gamma-lactones. *Microbial Cell Factories* **18**, 62 (2019). https://doi.org:10.1186/s12934-019-1113-1

91 Wache, Y. *et al.* Involvement of Acyl Coenzyme A Oxidase Isozymes in Biotransoformation of Methyl Ricinoleate into gamma-decalactone by *Yarrowia lipolytica*. *Applied and Environmental Microbiology* **66**, 1233-1236 (2000).

92 Boratynski, F., Szczepanska, E., De Simeis, D., Serra, S. & Brenna, E. Bacterial Biotransformation of Oleic Acid: New Findings on the Formation of gamma-Dodecalactone and 10-Ketostearic Acid in the Culture of *Micrococcus luteus*. *Molecules* **25** (2020). https://doi.org:10.3390/molecules25133024

93 Dufosse, L. *et al.* Production of gamma-decalactone and 4-Hydroxy-Decanoic Acid in the Genus *Sporidiobolus*. *Journal of Fermentation and Bioengineering* **86**, 169-173 (1998).

94 Rong, S. *et al.* Biotransformation of 12-hydroxystearic acid to gamma-decalactone: Comparison of two separation systems. *Journal of Microbiological Methods* **178**, 106041 (2020). https://doi.org:10.1016/j.mimet.2020.106041

95 Zia, H. *et al.* Biotechnological formation of dairy flavor inducing delta-lactones from vegetable oil. *Food Chemistry X* **13**, 100220 (2022). https://doi.org:10.1016/j.fochx.2022.100220

96 Dhamankar, H., Tarasova, Y., Martin, C. H. & Prather, K. L. Engineering *E. coli* for the biosynthesis of 3-hydroxy-gamma-butyrolactone (3HBL) and 3,4-dihydroxybutyric acid (3,4-DHBA) as value-added chemicals from glucose as a sole carbon source. *Metabolic Engineering* **25**, 72-81 (2014). https://doi.org:10.1016/j.ymben.2014.06.004

97 Martin, C. H. *et al.* A platform pathway for production of 3-hydroxyacids provides a biosynthetic route to 3-hydroxy-gamma-butyrolactone. *Nature Communications* **4**, 1414 (2013). https://doi.org:10.1038/ncomms2418

98 Manning, J. *et al.* Regio- and Enantio-selective Chemo-enzymatic C-H-Lactonization of Decanoic Acid to (S)-delta-Decalactone. *Angewandte Chemie International Edition* **58**, 5668-5671 (2019). https://doi.org:10.1002/anie.201901242

99 Maseme, M. J., Pennec, A., van Marwijk, J., Opperman, D. J. & Smit, M. S. CYP505E3: A Novel Self-Sufficient omega-7 In-Chain Hydroxylase. *Angewandte Chemie International Edition* **59**, 10359-10362 (2020). https://doi.org:10.1002/anie.202001055

100 Fronza, G. *et al.* Stereochemical aspects of flavour biogeneration through baker's yeast mediated reduction of carbonyl-activated double bonds. *Pure and Applied Chemistry* **68**, 2065-2071 (1996).

101 Wang, T., Zhang, X. Y., Zheng, Y. C. & Bai, Y. P. Stereoselective synthesis of chiral delta-lactones via an engineered carbonyl reductase. *Chemical Communications* **57**, 10584-10587 (2021). https://doi.org:10.1039/d1cc04542c

102 Voelker, T. A. *et al.* Fatty Acid Biosynthesis Redirected to Medium Chains in Transgenic Oilseed Plants. *Science* **257**, 72-74 (1992).

103 Bonner, W. M. & Bloch, K. Purification and Properties of Fatty Acyl Thioesterase I from *Escherichia coli*. *Journal of Biological Chemistry* **247**, 3123-3133 (1972).

104 Jing, F. *et al.* Phylogenetic and experimental characterization of an acyl-ACP thioesterase family reveals significant diversity in enzymatic specificity and activity. *BMC Biochemistry* **12**, 44 (2011). https://doi.org:10.1186/1471-2091-12-44

105 Jones, A., Davies, M. & Voelker, T. A. Palmitoyl-Acyl Carrier Protein (ACP) Thioesterase and the Evolutionary Origin of Plant Acyl-ACP Thioesterases. *The Plant Cell* **7**, 359-371 (1995).

106 Dormann, P., Spener, F. & Ohlrogge, J. B. Characterization of two acyl-acyl carrier protein thioesterase from developing *Cuphea* seeds specific for medium-chain- and oleoyl-acyl carrier protein. *Planta* **189**, 429-432 (1993).

107 Graham, S. A., Hirsinger, F. & Robbelen, G. Fatty acids of Cuphea (Lythraceae) seed lipids and their systematic significance. *American Journal of Botany* **68**, 908-917 (1981).

108 Iskandarov, U. *et al.* A Specialized Diacylglycerol Acyltransferase Contributes to the Extreme Medium-Chain Fatty Acid Content of *Cuphea* Seed Oil. *Plant Physiology* **174**, 97-109 (2017). https://doi.org:10.1104/pp.16.01894

109 Dormann, P., Voelker, T. A. & Ohlrogge, J. B. Cloning and Expression in *Escherichia coli* of a Novel Thioesteral from *Arabidopsis thaliana* Specific for Long-Chain Acyl-Acyl Carrier Proteins. *Archives of Biochemistry and Biophysics* **316**, 612-618 (1995).

110 Jha, J. K. *et al.* Cloning and functional expression of an acyl-ACP thioesterase FatB type from *Diploknema (Madhuca) butyracea* seeds in *Escherichia coli*. *Plant Physiology and Biochemistry* **44**, 645-655 (2006).

111 Jha, S. S. *et al.* Cloning and characterization of cDNAs encoding for long-chain saturated acyl-ACP thioesterases from the developing seeds of *Brassica juncea*. *Plant Physiology and Biochemistry* **48**, 476-480 (2010). https://doi.org:10.1016/j.plaphy.2010.02.006

112 Kim, H. J. *et al.* Toward production of jet fuel functionality in oilseeds: identification of FatB acyl-acyl carrier protein thioesterases and evaluation of combinatorial expression

strategies in *Camelina* seeds. *Journal of Experimental Botany* **66**, 4251-4265 (2015). https://doi.org:10.1093/jxb/erv225

113 Martins-Noguerol, R. *et al.* Characterization of the acyl-ACP thioesterases from *Koelreuteria paniculata* reveals a new type of FatB thioesterase. *Heliyon* **6**, e05237 (2020). https://doi.org:10.1016/j.heliyon.2020.e05237

114 Moreno-Perez, A. J., Sanchez-Garcia, A., Salas, J. J., Garces, R. & Martinez-Force, E. Acyl-ACP thioesterases from macadamia (*Macadamia tetraphylla*) nuts: cloning, characterization and their impact on oil composition. *Plant Physiology and Biochemistry* **49**, 82-87 (2011). https://doi.org:10.1016/j.plaphy.2010.10.002

115 Rodriguez-Rodriguez, M. F., Salas, J. J., Garces, R. & Martinez-Force, E. Acyl-ACP thioesterases from *Camelina sativa*: cloning, enzymatic characterization and implication in seed oil fatty acid composition. *Phytochemistry* **107**, 7-15 (2014). https://doi.org:10.1016/j.phytochem.2014.08.014

116 Sanchez-Garcia, A. *et al.* Acyl-ACP thioesterases from castor (*Ricinus communis* L.): an enzymatic system appropriate for high rates of oil synthesis and accumulation. *Phytochemistry* **71**, 860-869 (2010). https://doi.org:10.1016/j.phytochem.2010.03.015

117 Wu, P. Z. *et al.* Cloning and functional characterization of an acyl-acyl carrier protein thioesterase (JcFATB1) from *Jatropha curcas*. *Tree Physiology* **29**, 1299-1305 (2009). https://doi.org:10.1093/treephys/tpp054

118 Xia, W. *et al.* Identification and Validation of Candidate Genes Involved in Fatty Acid Content in Oil Palm by Genome-Wide Association Analysis. *Frontiers in Plant Science* **10**, 1263 (2019). https://doi.org:10.3389/fpls.2019.01263

119 Yeon, J., Park, J.-S., Lee, S. H., Lee, K.-R. & Yi, H. Overexpression of *Cuphea viscosissima* CvFatB4 enhances 16:0 fatty acid accumulation in Arabidopsis. *Journal of Plant Biotechnology* **46**, 282-290 (2019). https://doi.org:10.5010/jpb.2019.46.4.282

120 Yuan, L., Voelker, T. A. & Hawkins, D. J. Modification of the substrate specificity of an acyl-acyl carrier protein thioesterase by protein engineering. *Procedings of the National Academy of Science* **92**, 10639-10643 (1995).

121 Zhang, X., Li, M., Agrawal, A. & San, K. Y. Efficient free fatty acid production in *Escherichia coli* using plant acyl-ACP thioesterases. *Metabolic Engineering* **13**, 713-722 (2011). https://doi.org:10.1016/j.ymben.2011.09.007

122 Filichkin, S. A., Slabaugh, M. B. & Knapp, S. J. New FATB thioesterases from a high-laurate *Cuphea* species: Functional and complementation analyses. *European Journal of Lipid Science and Technology* **108**, 979-990 (2006). https://doi.org:10.1002/ejlt.200600158

123 Dehesh, K., Jones, A., Knutzon, D. S. & Voelker, T. A. Production of high levels of 8:0 and 10:0 fatty acids in transgenic canola by overexpression of *ChFatB2*, a thioesterase cDNA from *Cuphea hookeriana*. *The Plant Journal* **9**, 167-172 (1996).

124 Voelker, T. A., Jones, A., Crammer, A. M., Davies, H. M. & Knutzon, D. S. Broad-Range and Binary-Range Acyl-Acyl-Carrier-Protein Thioesterases Suggest an Alternative Mechanism for Medium-Chain Production in Seeds. *Plant Physiology* **114**, 669-677 (1997).

125 Banerjee, D., Jindra, M. A., Linot, A. J., Pfleger, B. F. & Maranas, C. D. EnZymClass: Substrate specificity prediction tool of plant acyl-ACP thioesterases based on ensemble learning. *Current Research in Biotechnology* **4**, 1-9 (2022). https://doi.org:10.1016/j.crbiot.2021.12.002

126 Hernandez Lozada, N. J. *et al.* Highly Active C8-Acyl-ACP Thioesterase Variant Isolated by a Synthetic Selection Strategy. *ACS Synthetic Biology* **7**, 2205-2215 (2018). https://doi.org:10.1021/acssynbio.8b00215

127 Aharoni, A. *et al.* Identification of the *SAAT* Gene Involved in Strawberry Flavor Biogenesis by Use of DNA Microarrays. *The Plant Cell* **12**, 647-661 (2000).

128 Yunus, I. S. *et al.* Improved Bioproduction of 1-Octanol Using Engineered *Synechocystis* sp. PCC 6803. *ACS Synthetic Biology* (2021). https://doi.org:10.1021/acssynbio.1c00029

129 Feng, Y. *et al.* Tuning of acyl-ACP thioesterase activity directed for tailored fatty acid synthesis. *Applied Microbiology and Biotechnology* **102**, 3173-3182 (2018). https://doi.org:10.1007/s00253-018-8770-6

130 Feng, Y. *et al.* Structural Insight into Acyl-ACP Thioesterase toward Substrate Specificity Design. *ACS Chemical Biology* **12**, 2830-2836 (2017). https://doi.org:10.1021/acschembio.7b00641

131 Wang, Q. *et al.* Manipulating fatty-acid profile at unit chain-length resolution in the model industrial oleaginous microalgae *Nannochloropsis*. *Metabolic Engineering* **66**, 157-166 (2021). https://doi.org:10.1016/j.ymben.2021.03.015

132 Facciotti, M. T. & Yuan, L. Molecular dissection of the plant acyl-acyl carrier protein thioesterases. *European Journal of Lipid Science and Technology* **100**, 167-172 (1998).

133 Salas, J. J. & Ohlrogge, J. Characterization of substrate specificity of plant FatA and FatB acyl-ACP thioesterases. *Archives of Biochemistry and Biophysics* **403**, 25-34 (2002).

134 Jing, F., Zhao, L., Yandeau-Nelson, M. D. & Nikolau, B. J. Two distinct domains contribute to the substrate acyl chain length selectivity of plant acyl-ACP thioesterase. *Nature Communications* **9**, 860 (2018). https://doi.org:10.1038/s41467-018-03310-z

135 Ziesack, M. *et al.* Chimeric Fatty Acyl-Acyl Carrier Protein Thioesterases Provide a Mechanistic Insight into Enzyme Specificity and Expression. *Applied and Environmental Microbiology* **84** (2018).

136 Facciotti, M. T., Bertain, P. B. & Yuan, L. Improved stearate phenotype in transgenic canola expressing a modified acyl-acyl carrier protein thioesterase. *Nature Biotechnology* **17**, 593-597 (1999).

137 Liu, Y. *et al.* Heterologous Expression of *Jatropha curcas* Fatty Acyl-ACP Thioesterase A (JcFATA) and B (JcFATB) Affects Fatty Acid Accumulation and Promotes Plant Growth and Development in *Arabidopsis*. *International Journal of Molecular Sciences* **23** (2022). https://doi.org:10.3390/ijms23084209

138 Lin, H. & Lee, Y. K. Genetic engineering of medium-chain-length fatty acid synthesis in *Dunaliella tertiolecta* for improved biodiesel production. *Journal of Applied Phycology* **29**, 2811-2819 (2017). https://doi.org:10.1007/s10811-017-1210-7

139 Radakovits, R., Eduafo, P. M. & Posewitz, M. C. Genetic engineering of fatty acid chain length in *Phaeodactylum tricornutum*. *Metabolic Engineering* **13**, 89-95 (2011). https://doi.org:10.1016/j.ymben.2010.10.003

140 Südfeld, C. *et al.* Accumulation of medium chain fatty acids in *Nannochloropsis oceanica* by heterologous expression of *Cuphea palustris* thioesterase FatB1. *Algal Research* **64** (2022). https://doi.org:10.1016/j.algal.2022.102665

141 Stefan, A., Hochkoeppler, A., Ugolini, L., Lazzeri, L. & Conte, E. The expression of the Cuphea palustris thioesterase CpFatB2 in *Yarrowia lipolytica* triggers oleic acid accumulation. *Biotechnology Progress* **32**, 26-35 (2016). https://doi.org:10.1002/btpr.2189

142 Liu, X., Sheng, J. & Curtiss, R., 3rd. Fatty acid production in genetically modified cyanobacteria. *Procedings of the National Academy of Science* **108**, 6899-6904 (2011). https://doi.org:10.1073/pnas.1103014108

143 Li, M., Zhang, X., Agrawal, A. & San, K. Y. Effect of acetate formation pathway and long chain fatty acid CoA-ligase on the free fatty acid production in *E. coli* exppressing acyl-ACP thioesterase from *Ricinus communis*. *Metabolic Engineering* **14**, 380-387 (2012).

144 Lu, X., Vora, H. & Khosla, C. Overproduction of free fatty acids in *E. coli*: implications for biodiesel production. *Metabolic Engineering* **10**, 333-339 (2008). https://doi.org:10.1016/j.ymben.2008.08.006

145 Torella, J. P. *et al.* Tailored fatty acid synthesis via dynamic control of fatty acid elongation. *Procedings of the National Academy of Sciences* **110**, 11290-11295 (2013). https://doi.org:10.1073/pnas.1307129110

146 Cao, Y. X. *et al.* Biosynthesis of odd-chain fatty alcohols in *Escherichia coli*. *Metabolic Engineering* **29**, 113-123 (2015). https://doi.org:10.1016/j.ymben.2015.03.005

147 Hernandez Lozada, N. J., Simmons, T. R., Xu, K., Jindra, M. A. & Pfleger, B. F. Production of 1-octanol in *Escherichia coli* by a high flux thioesterase route. *Metabolic Engineering* **61**, 352-359 (2020). https://doi.org:10.1016/j.ymben.2020.07.004

148 Jin, Z. *et al.* Engineering *Saccharomyces cerevisiae* to produce odd chain-length fatty alcohols. *Biotechnology and Bioengineering* **113**, 842-851 (2016). https://doi.org:10.1002/bit.25856

149 Youngquist, J. T. *et al.* Production of medium chain length fatty alcohols from glucose in *Escherichia coli*. *Metabolic Engineering* **20**, 177-186 (2013).

150 Yunus, I. S. & Jones, P. R. Photosynthesis-dependent biosynthesis of medium chain-length fatty acids and alcohols. *Metabolic Engineering* **49**, 59-68 (2018). https://doi.org:10.1016/j.ymben.2018.07.015

151 Chacon, M. G., Kendrick, E. G. & Leak, D. J. Engineering *Escherichia coli* for the production of butyl octanoate from endogenous octanoyl-CoA. *PeerJ* **7**, e6971 (2019). https://doi.org:10.7717/peerj.6971

152 Rahman, Z. *et al.* Enhanced Production of Fatty Acid Ethyl Ester with Engineered fabHDG Operon in *Escherichia coli*. *Microorganisms* **7** (2019). https://doi.org:10.3390/microorganisms7110552

153 Yunus, I. S., Palma, A., Trudeau, D. L., Tawfik, D. S. & Jones, P. R. Methanol-free biosynthesis of fatty acid methyl ester (FAME) in *Synechocystis* sp. PCC 6803. *Metabolic Engineering* **57**, 217-227 (2020). https://doi.org:10.1016/j.ymben.2019.12.001

154 Xiao, K. *et al.* Metabolic Engineering for Enhanced Medium Chain Omega Hydroxy Fatty Acid Production in *Escherichia coli*. *Frontiers in Microbiology* **9**, 139 (2018). https://doi.org:10.3389/fmicb.2018.00139

155 Goh, E. B., Baidoo, E. E., Keasling, J. D. & Beller, H. R. Engineering of bacterial methyl ketone synthesis for biofuels. *Applied Environmental Microbiology* **78**, 70-80 (2012). https://doi.org:10.1128/AEM.06785-11

156 Yan, Q. *et al.* Metabolic engineering of beta-oxidation to leverage thioesterases for production of 2-heptanone, 2-nonanone and 2-undecanone. *Metabolic Engineering* **61**, 335-343 (2020). https://doi.org:10.1016/j.ymben.2020.05.008

157 Gupta, R. S., Patel, S., Saini, N. & Chen, S. Robust demarcation of 17 distinct *Bacillus* species clades, proposed as novel *Bacillaceae* genera, by phylogenomics and comparative genomic analyses: description of *Robertmurraya kyonggiensis* sp. nov. and proposal for an emended genus *Bacillus* limiting it only to the members of the Subtilis and Cereus clades of species. *International Journal of Systematic and*

Evolutionary Microbiology **70**, 5753-5798 (2020). https://doi.org:10.1099/ijsem.0.004475

158 Boddupalli, S. S., Estabrook, R. W. & Peterson, J. A. Fatty Acid Monooxygenation by Cytochrome P-450 BM-3. *The Journal of Biological Chemistry* **265**, 4233-4239 (1990).

159 Munro, A. W. *et al.* P450 BM3: the very model of a modern flavocytochrome. *Trends in Biochemical Sciences* **27** (2002).

160 Truan, G., Komandla, M. R., Falck, J. R. & Peterson, J. A. P450BM-3: Absolute Configuration of the Primary Metabolies of Palmitic Acid. *Archives of Biochemistry and Biophysics* **366**, 192-198 (1999).

161 van Vugt-Lussenburg, B. M., Damsten, M. C., Maasdijk, D. M., Vermeulen, N. P. & Commandeur, J. N. Heterotropic and homotropic cooperativity by a drug-metabolising mutant of cytochrome P450 BM3. *Biochemical and Biophysical Research Communications* **346**, 810-818 (2006). https://doi.org:10.1016/j.bbrc.2006.05.179

162 Schneider, S., Wubboltds, M. G., Oesterhelt, G., Sanglard, D. & Witholt, B. Controlled Regioselectivity of Fatty Acid Oxidation by Whole Cells Producing Cytochrome P450 BM-3 Monooxygenase under Varied Dissolved Oxygen Concentrations. *Biotechnology and Bioengineering* **64**, 333-341 (1999).

163 Noble, M. A. *et al.* Roles of key active-site residues in flavocytochrome P450 BM3. *Biochemical Journal* **339**, 371-379 (1999).

164 Ost, T. W. B. *et al.* Rational re-design of the substrate binding site of flavocytochrome P450 BM3. *FEBS Letters*, 173-177 (2000).

165 Kang, J. Y. *et al.* Characterization of diverse natural variants of CYP102A1 found within a species of *Bacillus megaterium*. *AMB Express* **1**, 1 (2011). https://doi.org:10.1186/2191-0855-1-1

166 Dietrich, M., Do, T. A., Schmid, R. D., Pleiss, J. & Urlacher, V. B. Altering the regioselectivity of the subterminal fatty acid hydroxylase P450 BM-3 towards gamma- and delta-positions. *Journal of Biotechnology* **139**, 115-117 (2009). https://doi.org:10.1016/j.jbiotec.2008.10.002

167 Glieder, A., Farinas, E. T. & Arnold, F. H. Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase. *Nature Biotechnology* **20**, 1135-1139 (2002). https://doi.org:10.1038/nbt744

168 Haines, D. C. *et al.* A single active-site mutation of P450BM-3 dramatically enhances substrate binding and rate of product formation. *Biochemistry* **50**, 8333-8341 (2011). https://doi.org:10.1021/bi201099j

169 Jang, H.-H. *et al.* Role of Leu188 in the Fatty Acid Hydroxylase Activity of CYP102A1 from *Bacillus megaterium*. *Journal of Molecular Catalysis B: Enzymatic*, 35-42 (2016).

170 Lentz, O. *et al.* Modification of the fatty acid specificity of cytochrome P450 BM-3 from *Bacillus megaterium* by directed evolution: a validated assay. *Journal of Molecular Catalysis B: Enzymatic*, 123-133 (2001).

171 Zheng, Y. *et al.* High-specificity synthesis of novel monomers by remodeled alcohol hydroxylase. *BMC Biotechnology* **16**, 61 (2016). https://doi.org:10.1186/s12896-016-0291-8

172 Appel, D., Lutz-Wahl, S., Fischer, P., Schwaneberg, U. & Schmid, R. D. A P450 BM-3 mutant hydroxylates alkanes, cycloalkanes, arenes and heteroarenes. *Journal of Biotechnology* **88**, 161-171 (2001).

173 Chen, M. M., Snow, C. D., Vizcarra, C. L., Mayo, S. L. & Arnold, F. H. Comparison of random mutagenesis and semi-rational designed libraries for improved cytochrome P450 BM3-catalyzed hydroxylation of small alkanes. *Protein Engineering, Design and Selection* **25**, 171-178 (2012). https://doi.org:10.1093/protein/gzs004

174 Meinhold, P., Peters, M. W., Chen, M. M., Takahashi, K. & Arnold, F. H. Direct conversion of ethane to ethanol by engineered cytochrome P450 BM3. *Chembiochem* **6**, 1765-1768 (2005). https://doi.org:10.1002/cbic.200500261

175 Peters, M. W., Meinhold, P., Glieder, A. & Arnold, F. H. Regio- and Enantioselective Alkane Hydroxylation with Engineered Cytochromes P450 BM-3. *Journal of the American Chemical Society* **125**, 13442-13450 (2003).

176 Whitehouse, C. J. *et al.* Evolved CYP102A1 (P450BM3) variants oxidise a range of non-natural substrates and offer new selectivity options. *ChemComm*, 966-968 (2008). https://doi.org:10.1039/b718124h

177 Whitehouse, C. J. *et al.* A highly active single-mutation variant of P450BM3 (CYP102A1). *Chembiochem* **10**, 1654-1656 (2009). https://doi.org:10.1002/cbic.200900279

178 Carmichael, A. B. & Wong, L.-L. Protein engineering of *Bacillus megaterium* CYP102 The oxidation of polycyclic aromatic hydrocarbons. *European Journal of Biochemistry* **268**, 3117-3125 (2001).

179 Munday, S. D., Dezvarei, S. & Bell, S. G. Increasing the Activity and Efficiency of Stereoselective Oxidations by using Decoy Molecules in Combination with Rate-Enhancing Variants of P450Bm3. *ChemCatChem* **8**, 2789-2796 (2016). https://doi.org:10.1002/cctc.201600551

180 Neufeld, K., Marienhagen, J., Schwaneberg, U. & Pietruszka, J. Benzylic hydroxylation of aromatic compounds by P450 BM3. *Green Chemistry* **15** (2013). https://doi.org:10.1039/c3gc40838h

181 Urlacher, V. B., Makhsumkhanov, A. & Schmid, R. D. Biotransformation of beta-ionone by engineered cytochrome P450 BM-3. *Applied Microbiology and Biotechnology* **70**, 53-59 (2006). https://doi.org:10.1007/s00253-005-0028-4

182 Li, Y. & Wong, L. L. Multi-Functional Oxidase Activity of CYP102A1 (P450BM3) in the Oxidation of Quinolines and Tetrahydroquinolines. *Angewandte Chemie International Edition* **58**, 9551-9555 (2019). https://doi.org:10.1002/anie.201904157

183 Chu, L. L. *et al.* Hydroxylation of diverse flavonoids by CYP450 BM3 variants: biosynthesis of eriodictyol from naringenin in whole cells and its biological activities. *Microbial Cell Factories* **15**, 135 (2016). https://doi.org:10.1186/s12934-016-0533-4

184 Hong, L. L. & Kong, J. Q. Altering the Regioselectivity of Cytochrome P450 BM3 Variant M13 toward Genistein through Protein Engineering and Variation of Reaction Conditions. *ACS Omega* **5**, 32059-32066 (2020). https://doi.org:10.1021/acsomega.0c05088

185 Kim, K. H. *et al.* Generation of human chiral metabolites of simvastatin and lovastatin by bacterial CYP102A1 mutants. *Drug Metabolism and Disposition* **39**, 140-150 (2011). https://doi.org:10.1124/dmd.110.036392

186 Liu, X. & Kong, J. Q. Steroids hydroxylation catalyzed by the monooxygenase mutant 139-3 from *Bacillus megaterium* BM3. *Acta Pharmaceuteca Sinica B* **7**, 510-516 (2017). https://doi.org:10.1016/j.apsb.2017.04.006

187 Vottero, E. *et al.* Role of residue 87 in substrate selectivity and regioselectivity of drug-metabolizing cytochrome P450 CYP102A1 M11. *Journal of Biological Inorganic Chemistry* **16**, 899-912 (2011). https://doi.org:10.1007/s00775-011-0789-4

188 Schneider, S., Wubboltds, M. G., Sanglard, D. & Witholt, B. Biocatalyst Engineering by Assembly of Fatty Acid Transport and Oxidation Activities for In Vivo Application of Cytochrome P-450 BM-3 Monoxygenase. *Applied and Environmental Microbiology* **64**, 3784-3790 (1998).

189 Cao, Y. *et al.* Metabolic engineering of *Escherichia coli* for the production of hydroxy fatty acids from glucose. *BMC Biotechnology*, 16-26 (2016).

190 Bhattacharya, S. *et al.* A comprehensive characterization of novel CYP-BM3 homolog (CYP-BA) from *Bacillus aryabhattai*. *Enzyme and Microbial Technology* **148**, 109806 (2021). https://doi.org:10.1016/j.enzmictec.2021.109806

191 Choi, K. Y. *et al.* Cloning, expression and characterization of CYP102D1, a self-sufficient P450 monooxygenase from *Streptomyces avermitilis*. *The FEBS Journal* **279**, 1650-1662 (2012). https://doi.org:10.1111/j.1742-4658.2011.08462.x

192 Chowdhary, P. K., Alemseghed, M. & Haines, D. C. Cloning, expression and characterization of a fast self-sufficient P450: CYP102A5 from *Bacillus cereus*.

Archives of Biochemistry and Biophysics **468**, 32-43 (2007). https://doi.org:10.1016/j.abb.2007.09.010

193 Dietrich, M. *et al.* Cloning, expression and characterisation of CYP102A7, a self-sufficient P450 monooxygenase from *Bacillus licheniformis*. *Applied and Environmental Microbiology* **79**, 931-940 (2008). https://doi.org:10.1007/s00253-008-1500-8

194 Gustafsson, M. C. U. *et al.* Expression, Purification, and Characterization of *Bacillus subtilis* Cytochromes P450 CYP102A2 and CYP102A: Flavocytochrome Homologues of P450 BM3 from *Bacillus megaterium*. *Biochemistry* **43**, 5474-5487 (2004).

195 Hammerer, L., Winkler, C. K. & Kroutil, W. Regioselective Biocatalytic Hydroxylation of Fatty Acids by Cytochrome P450s. *Catalysis Letters* **148**, 787-812 (2017). https://doi.org:10.1007/s10562-017-2273-4

196 Kitazume, T., Takaya, N., Nakayama, N. & Shoun, H. *Fusarium oxysporum* fatty-acid subterminal hydroxylase (CYP505) is a membrane-bound eukaryotic counterpart of Bacillus megaterium cytochrome P450BM3. *Journal of Biological Chemistry* **275**, 39734-39740 (2000). https://doi.org:10.1074/jbc.M005617200

197 Kitazume, T. *et al.* Kinetic analysis of hydroxylation of saturated fatty acids by recombinant P450foxy produced by an Escherichia coli expression system. *European Journal of Biochemistry* **269**, 2075-2082 (2002). https://doi.org:10.1046/j.1432-1033.2002.02855.x

198 Munday, S. D., Maddigan, N. K., Young, R. J. & Bell, S. G. Characterisation of two self-sufficient CYP102 family monooxygenases from *Ktedonobacter racemifer* DSM44963 which have new fatty acid alcohol product profiles. *Biochimica et Biophysica Acta* **1860**, 1149-1162 (2016). https://doi.org:10.1016/j.bbagen.2016.01.023

199 Park, H. G. *et al.* CYP52A23 from *Candida albicans* and its Substrate Preference for Fatty Acid Hydroxylation. *Archives of Biochemistry and Biophysics* **671**, 27-34 (2019). https://doi.org:10.1016/j.abb.2019.06.002

200 Torres, S., Fjetland, C. R. & Lammers, P. J. Alkane-induced expression, substrate binding profile, and immunolocalization of a cytochrome P450 encoded on the *nifD* excision element of *Anabaena* 7120. *BMC Microbiology* **5**, 16 (2005). https://doi.org:10.1186/1471-2180-5-16

201 Zong, L. *et al.* Optimization and Engineering of a Self-Sufficient CYP102 Enzyme from *Bacillus amyloliquefaciens* towards Synthesis of In-Chain Hydroxy Fatty Acids. *Catalysts* **11** (2021). https://doi.org:10.3390/catal11060665

202 Yin, Y. C. *et al.* Unusually broad substrate profile of self-sufficient cytochrome P450 monooxygenase CYP116B4 from *Labrenzia aggregata*. *Chembiochem* **15**, 2443-2449 (2014). https://doi.org:10.1002/cbic.201402309

203 Khatri, Y. *et al.* A natural heme-signature variant of CYP267A1 from *Sorangium cellulosum* So ce56 executes diverse omega-hydroxylation. *The FEBS Journal* **282**, 74-88 (2015). https://doi.org:10.1111/febs.13104

204 Roberts, G. A., Grogan, G., Greter, A., Flitsch, S. L. & Turner, N. J. Identification of a new class of cytochrome P450 from a *Rhodococcus* sp. *Journal of Bacteriology* **184**, 3898-3908 (2002). https://doi.org:10.1128/JB.184.14.3898-3908.2002

205 Tavanti, M., Porter, J. L., Sabatini, S., Turner, N. J. & Flitsch, S. L. Panel of New Thermostable CYP116B Self-Sufficient Cytochrome P450 Monooxygenases that Catalyze C−H Activation with a Diverse Substrate Scope. *ChemCatChem* **10**, 1042-1051 (2018). https://doi.org:10.1002/cctc.201701510

206 Porter, J. L. *et al.* Cloning, expression and characterisation of P450-Hal1 (CYP116B62) from *Halomonas* sp. NCIMB 172: A self-sufficient P450 with high expression and diverse substrate scope. *Enzyme and Microbial Technology* **113**, 1-8 (2018). https://doi.org:10.1016/j.enzmictec.2018.02.005

207 Tavanti, M. *et al.* The crystal structure of P450-TT heme-domain provides the first structural insights into the versatile class VII P450s. *Biochemical and Biophysical Research Communications* **501**, 846-850 (2018). https://doi.org:10.1016/j.bbrc.2018.05.014

208 Zhang, L. *et al.* Structural insight into the electron transfer pathway of a self-sufficient P450 monooxygenase. *Nature Communications* **11**, 2676 (2020). https://doi.org:10.1038/s41467-020-16500-5

209 Faponle, A. S., Roy, A., Adelegan, A. A. & Gauld, J. W. Molecular Dynamics Simulations of a Cytochrome P450 from *Tepidiphilus thermophilus* (P450-TT) Reveal How Its Substrate-Binding Channel Opens. *Molecules* **26** (2021). https://doi.org:10.3390/molecules26123614

210 Wang, Z., Shaik, S. & Wang, B. Conformational Motion of Ferredoxin Enables Efficient Electron Transfer to Heme in the Full-Length P450TT. *Journal of the American Chemical Society* **143**, 1005-1016 (2021). https://doi.org:10.1021/jacs.0c11279

211 Kokorin, A. *et al.* Genetic fusion of P450 BM3 and formate dehydrogenase towards self-sufficient biocatalysts with enhanced activity. *Scientific Reports* **11**, 21706 (2021). https://doi.org:10.1038/s41598-021-00957-5

212 Nagao, T. *et al.* Cloning, Nucleotide Sequences, and Enzymatic Properties of Glucose Dehydrogenase Isozymes from *Bacillus megaterium* IAM1030. *Journal of Bacteriology* **174** (1992).

213 Kataoka, M. *et al. Escherichia coli* Transformant Expressing the Glucose Dehydrogenase Gene from *Bacillus megaterium* as a Cofactor Regenerator in a Chiral Alcohol Production System. *Bioscience, Biotechnology, and Biochemistry* **61**, 167-169 (1998).

214 Kataoka, M. *et al.* Stereoselective reduction of ethyl 4-chloro-3-oxobutanoate by *Escherichia coli* transformant cells coexpressing the aldehyde reductase and glucose dehydrogenase genes. *Applied Microbiology and Biotechnology* **51**, 486-490 (1999).

215 Park, H. J., Jung, J., Choi, H., Uhm, K. N. & Kim, H. K. Enantioselective bioconversion using *Escherichia coli* cells expressing *Saccharomyces cerevisiae* reductase and *Bacillus subtilis* glucose dehydrogenase. *Journal of Microbiology and Biotechnology* **20**, 1300-1306 (2010). https://doi.org:10.4014/jmb.1003.03025

216 Xu, Z., Jing, K., Liu, Y. & Cen, P. High-level expression of recombinant glucose dehydrogenase and its application in NADPH regeneration. *Journal of Industrial Microbiology and Biotechnology* **34**, 83-90 (2007). https://doi.org:10.1007/s10295-006-0168-2

217 Yun, H., Choi, H. L., Fadnavis, N. W. & Kim, B. G. Stereospecific Synthesis of (R)-2-Hydroxy Carboxylic Acids Using Recombinant *E. coli* BL21 Overexpressing YiaE from *Escherichia coli* K12 and Glucose Dehydrogenase from *Bacillus subtilis*. *Biotechnology Progress* **21**, 366-371 (2005).

218 Buergler, M. B., Dennig, A. & Nidetzky, B. Process intensification for cytochrome P450 BM3-catalyzed oxy-functionalization of dodecanoic acid. *Biotechnology and Bioengineering* **117**, 2377-2388 (2020). https://doi.org:10.1002/bit.27372

219 Pham, S. Q., Gao, P. & Li, Z. Engineering of recombinant *E. coli* cells co-expressing P450pyrTM monooxygenase and glucose dehydrogenase for highly regio- and stereoselective hydroxylation of alicycles with cofactor recycling. *Biotechnology and Bioengineering* **110**, 363-373 (2013). https://doi.org:10.1002/bit.24632

220 Schewe, H., Holtmann, D. & Schrader, J. P450(BM-3)-catalyzed whole-cell biotransformation of alpha-pinene with recombinant *Escherichia coli* in an aqueous-organic two-phase system. *Applied Microbiology and Biotechnology* **83**, 849-857 (2009). https://doi.org:10.1007/s00253-009-1917-8

221 Wang, F. *et al.* One-pot biocatalytic route from cycloalkanes to alpha,omega-dicarboxylic acids by designed *Escherichia coli* consortia. *Nature Communications* **11**, 5035 (2020). https://doi.org:10.1038/s41467-020-18833-7

222 Zhang, J. D., Li, A. T., Yu, H. L., Imanaka, T. & Xu, J. H. Synthesis of optically pure S-sulfoxide by *Escherichia coli* transformant cells coexpressing the P450 monooxygenase and glucose dehydrogenase genes. *Journal of Industrial*

Microbiology and Biotechnology **38**, 633-641 (2011). https://doi.org:10.1007/s10295-010-0809-3

223 Lu, Y. & Mei, L. Co-expression of P450 BM3 and glucose dehydrogenase by recombinant *Escherichia coli* and its application in an NADPH-dependent indigo production system. *Journal of Industrial Microbiology and Biotechnology* **34**, 247-253 (2007). https://doi.org:10.1007/s10295-006-0193-1

224 Dafoe, J. T. & Daugulis, A. J. Production of 4-valerolactone by an equilibrium-limited transformation in a partitioning bioreactor: impact of absorptive polymer properties. *Bioprocess and Biosystems Engineering* **37**, 533-542 (2014). https://doi.org:10.1007/s00449-013-1020-5

225 Satoh, Y., Murakami, F., Tajima, K. & Munekata, M. Enzymatic synthesis of poly(3-hydroxybutyrate-co-4-hydroxybutyrate) with CoA recycling using polyhydroxyalkanoate synthase and acyl-CoA synthetase. *Journal of Bioscience and Bioengineering* **99**, 508-511 (2005). https://doi.org:10.1263/jbb.99.508

226 Wang, Q., Tappel, R. C., Zhu, C. & Nomura, C. T. Development of a new strategy for production of medium-chain-length polyhydroxyalkanoates by recombinant *Escherichia coli* via inexpensive non-fatty acid feedstocks. *Applied and Environmental Microbiology* **78**, 519-527 (2012). https://doi.org:10.1128/AEM.07020-11

227 Peng, B., Yu, M., Zhang, B., Xu, J. & Ma, R. Differences in PpAAT1 Activity in High- and Low-Aroma Peach Varieties Affect gamma-Decalactone Production. *Plant Physiology* **182**, 2065-2080 (2020). https://doi.org:10.1104/pp.19.00964

228 Song, Z. Z. *et al.* Site-directed mutagenesis identified the key active site residues of alcohol acyltransferase PpAAT1 responsible for aroma biosynthesis in peach fruits. *Horticulture Research* **8**, 32 (2021). https://doi.org:10.1038/s41438-021-00461-x

229 Menendez-Bravo, S., Comba, S., Gramajo, H. & Arabolaza, A. Metabolic engineering of microorganisms for the production of structurally diverse esters. *Applied Microbiology and Biotechnology* **101**, 3043-3053 (2017). https://doi.org:10.1007/s00253-017-8179-7

230 Lee, J. W. & Trinh, C. T. Microbial biosynthesis of lactate esters. *Biotechnology and Biofuels* **12**, 226 (2019). https://doi.org:10.1186/s13068-019-1563-z

231 Bazzali, O., Tomi, F., Casanova, J. & Bighelli, A. Occurrence of C8-C10 esters in Mediterranean *Myrtus communis* L. leaf essential oil. *Flavour and Fragrance Journal* **27**, 335-340 (2012). https://doi.org:10.1002/ffj.3102

232 Xiao, Z. *et al.* Olfactory impact of esters on rose essential oil floral alcohol aroma expression in model solution. *Food Research International* **116**, 211-222 (2019). https://doi.org:10.1016/j.foodres.2018.08.015

233 Cordente, A. G., Curtin, C. D., Varela, C. & Pretorius, I. S. Flavour-active wine yeasts. *Applied Microbiology and Biotechnology* **96**, 601-618 (2012). https://doi.org:10.1007/s00253-012-4370-z

234 Saerens, S. M., Delvaux, F. R., Verstrepen, K. J. & Thevelein, J. M. Production and biological function of volatile esters in *Saccharomyces cerevisiae*. *Microbiology and Biotechnology* **3**, 165-177 (2010). https://doi.org:10.1111/j.1751-7915.2009.00106.x

235 Costello, P. J., Siebert, T. E., Solomon, M. R. & Bartowsky, E. J. Synthesis of fruity ethyl esters by acyl coenzyme A: alcohol acyltransferase and reverse esterase activities in *Oenococcus oeni* and *Lactobacillus plantarum*. *Journal of Applied Microbiology* **114**, 797-806 (2013). https://doi.org:10.1111/jam.12098

236 Liu, S. Q., Holland, R. & Crow, V. L. Esters and their biosynthesis in fermented dairy products: a review. *International Dairy Journal* **14**, 923-945 (2004). https://doi.org:10.1016/j.idairyj.2004.02.010

237 Sumby, K. M., Grbin, P. R. & Jiranek, V. Microbial modulation of aromatic esters in wine: Current knowledge and future prospects. *Food Chemistry* **121**, 1-16 (2010). https://doi.org:10.1016/j.foodchem.2009.12.004

238 Verstrepen, K. J. *et al.* Flavor-Active Esters: Adding Fruitiness to Beer. *Journal of Bioscience and Bioengineering* **96**, 110-118 (2003).

239 Cristiani, G. & Monnet, V. Food micro-organisms and aromatic ester synthesis. *Sciences des Aliments* **21**, 211-230 (2001). https://doi.org:10.3166/sda.21.211-230

240 Olaniran, A. O., Hiralal, L., Mokoena, M. P. & Pillay, B. Flavour-active volatile compounds in beer: production, regulation and control. *Journal of the Institute of Brewing* **123**, 13-23 (2017). https://doi.org:10.1002/jib.389

241 Park, Y. C., Shaffer, C. E. & Bennett, G. N. Microbial formation of esters. *Applied Microbiology and Biotechnology* **85**, 13-25 (2009). https://doi.org:10.1007/s00253-009-2170-x

242 Renault, P., Coulon, J., de Revel, G., Barbe, J. C. & Bely, M. Increase of fruity aroma during mixed *T. delbrueckii* / *S. cerevisiae* wine fermentation is linked to specific esters enhancement. *International Journal of Food Microbiology* **207**, 40-48 (2015). https://doi.org:10.1016/j.ijfoodmicro.2015.04.037

243 Zhang, C.-Y. *et al.* Increased esters and decreased higher alcohols production by engineered brewer's yeast strains. *European Food Research and Technology* **236**, 1009-1014 (2013). https://doi.org:10.1007/s00217-013-1966-1

244 Lee, J. W. & Trinh, C. T. Towards renewable flavors, fragrances, and beyond. *Current Opinions in Biotechnology* **61**, 168-180 (2020). https://doi.org:10.1016/j.copbio.2019.12.017

245 Xu, Y., Minhazul, K. A. H. M. & Li, X. The occurrence, enzymatic production, and application of ethyl butanoate, an important flavor constituent. *Flavour and Fragrance Journal* **35**, 601-615 (2020). https://doi.org:10.1002/ffj.3613

246 Duan, Y., Zhu, Z., Cai, K., Tan, X. & Lu, X. De novo biosynthesis of biodiesel by *Escherichia coli* in optimized fed-batch cultivation. *PLoS One* **6**, e20265 (2011). https://doi.org:10.1371/journal.pone.0020265

247 Fang, D. *et al.* Metabolic and Process Engineering of *Clostridium beijerinckii* for Butyl Acetate Production in One Step. *Journal of Agricultural and Food Chemistry* **68**, 9475-9487 (2020). https://doi.org:10.1021/acs.jafc.0c00050

248 Feng, J. *et al.* Renewable fatty acid ester production in *Clostridium*. *Nature Communications* **12**, 4368 (2021). https://doi.org:10.1038/s41467-021-24038-3

249 Guo, D., Zhu, J., Deng, Z. & Liu, T. Metabolic engineering of *Escherichia coli* for production of fatty acid short-chain esters through combination of the fatty acid and 2-keto acid pathways. *Metabolic Engineering* **22**, 69-75 (2014).

250 Horton, C. E. & Bennett, G. N. Ester production in *E. coli* and *C. acetobutylicum*. *Enzyme and Microbial Technology* **38**, 937-943 (2006). https://doi.org:10.1016/j.enzmictec.2005.08.025

251 Kalscheuer, R., Stolting, T. & Steinbuchel, A. Microdiesel: *Escherichia coli* engineered for fuel production. *Microbiology (Reading)* **152**, 2529-2536 (2006). https://doi.org:10.1099/mic.0.29028-0

252 Kruis, A. J. *et al.* Ethyl acetate production by the elusive alcohol acetyltransferase from yeast. *Metabolic Engineering* **41**, 92-101 (2017). https://doi.org:10.1016/j.ymben.2017.03.004

253 Ku, J. T., Chen, A. Y. & Lan, E. I. Metabolic engineering of *Escherichia coli* for efficient biosynthesis of butyl acetate. *Microbial Cell Factories* **21**, 28 (2022). https://doi.org:10.1186/s12934-022-01755-y

254 Layton, D. S. & Trinh, C. T. Engineering modular ester fermentative pathways in *Escherichia coli*. *Metabolic Engineering* **26**, 77-88 (2014). https://doi.org:10.1016/j.ymben.2014.09.006

255 Layton, D. S. & Trinh, C. T. Expanding the modular ester fermentative pathways for combinatorial biosynthesis of esters from volatile organic acids. *Biotechnology and Bioengineering* **113**, 1764-1776 (2016). https://doi.org:10.1002/bit.25947

256 Lee, J. W. & Trinh, C. T. Controlling selectivity of modular microbial biosynthesis of butyryl-CoA-derived designer esters. *Metabolic Engineering* **69**, 262-274 (2022). https://doi.org:10.1016/j.ymben.2021.12.001

257 Noh, H. J., Woo, J. E., Lee, S. Y. & Jang, Y. S. Metabolic engineering of *Clostridium acetobutylicum* for the production of butyl butyrate. *Applied Microbiology and Biotechnology* **102**, 8319-8327 (2018). https://doi.org:10.1007/s00253-018-9267-z

258 Ren, J. Y. *et al.* Enhanced Production of Ethyl Lactate in *Saccharomyces cerevisiae* by Genetic Modification. *Journal of Agricultural and Food Chemistry* **68**, 13863-13870 (2020). https://doi.org:10.1021/acs.jafc.0c03967

259 Rodriguez, G. M., Tashiro, Y. & Atsumi, S. Expanding ester biosynthesis in *Escherichia coli*. *Nature Chemical Biology* **10**, 259-265 (2014). https://doi.org:10.1038/nchembio.1476

260 Seo, H., Lee, J. W., Garcia, S. & Trinh, C. T. Single mutation at a highly conserved region of chloramphenicol acetyltransferase enables isobutyl acetate production directly from cellulose by *Clostridium thermocellum* at elevated temperatures. *Biotechnology and Biofuels* **12**, 245 (2019). https://doi.org:10.1186/s13068-019-1583-8

261 Seo, H., Lee, J. W., Giannone, R. J., Dunlap, N. J. & Trinh, C. T. Engineering promiscuity of chloramphenicol acetyltransferase for microbial designer ester biosynthesis. *Metabolic Engineering* **66**, 179-190 (2021). https://doi.org:10.1016/j.ymben.2021.04.005

262 Shi, W. *et al.* Enhancement of C6–C10 fatty acid ethyl esters production in *Saccharomyces cerevisiae* CA by metabolic engineering. *Lwt* **145** (2021). https://doi.org:10.1016/j.lwt.2021.111496

263 Sinumvayo, J. P., Zhao, C., Liu, G., Li, Y. & Zhang, Y. One-pot production of butyl butyrate from glucose using a cognate “diamond-shaped” *E. coli* consortium. *Bioresources and Bioprocessing* **8** (2021). https://doi.org:10.1186/s40643-021-00372-8

264 Steen, E. J. *et al.* Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **463**, 559-562 (2010). https://doi.org:10.1038/nature08721

265 Tai, Y. S., Xiong, M. & Zhang, K. Engineered biosynthesis of medium-chain esters in *Escherichia coli*. *Metabolic Engineering*, 20-28 (2015).

266 Tashiro, Y., Desai, S. H. & Atsumi, S. Two-dimensional isobutyl acetate production pathways to improve carbon yield. *Nature Communications* **6**, 7488 (2015). https://doi.org:10.1038/ncomms8488

267 Teo, W. S., Ling, H., Yu, A. Q. & Chang, M. W. Metabolic engineering of *Saccharomyces cerevisiae* for production of fatty acid short- and branched-chain alkyl esters biodiesel. *Biotechnology and Biofuels* **8**, 177 (2015). https://doi.org:10.1186/s13068-015-0361-5

268 Zhang, G. *et al.* Metabolic engineering of *Saccharomyces cerevisiae* for the biosynthesis of ethyl crotonate. *LWT* **168** (2022). https://doi.org:10.1016/j.lwt.2022.113908

269 Zhang, Y., Peng, J., Zhao, H. & Shi, S. Engineering oleaginous yeast *Rhodotorula toruloides* for overproductionn of fatty acid ethyl esters. *Biotechnology and Biofuels* **14**, 115-126 (2021).

270 Horton, C. E., Huang, K.-X., Bennett, G. N. & Rudolph, F. B. Heterologous expression of the *Saccharomyces cerevisiae* alcohol acetyltransferase genes in *Clostridium acetobutylicum* and *Escherichia coli* for the production of isoamyl acetate. *Journal of Industrial Microbiology and Biotechnology* **30**, 427-432 (2003). https://doi.org:10.1007/s10295-003-0070-0

271 Vadali, R. V., Horton, C. E., Rudolph, F. B., Bennett, G. N. & San, K. Y. Production of isoamyl acetate in ackA-pta and/or ldh mutants of *Escherichia coli* with overexpression of yeast ATF2. *Applied Microbiology and Biotechnology* **63**, 698-704 (2004). https://doi.org:10.1007/s00253-003-1452-y

272 Ma, Y. *et al.* Biosynthetic Pathway for Ethyl Butyrate Production in *Saccharomyces cerevisiae*. *Journal of Agricultural and Food Chemistry* **68**, 4252-4260 (2020). https://doi.org:10.1021/acs.jafc.0c00750

273 Jawed, K., Abdelaal, A. S., Koffas, M. A. G. & Yazdani, S. S. Improved Butanol Production Using FASII Pathway in E. coli. *ACS Synthetic Biology* **9**, 2390-2398 (2020). https://doi.org:10.1021/acssynbio.0c00154

274 Jawed, K. *et al.* Engineered Production of Short Chain Fatty Acid in *Escherichia coli* Using Fatty Acid Synthesis Pathway. *PLoS One* **11**, e0160035 (2016). https://doi.org:10.1371/journal.pone.0160035

275 Liu, X. *et al.* Biosynthesis of butenoic acid through fatty acid biosynthesis pathway in *Escherichia coli*. *Applied Microbiology and Biotechnology* **99**, 1795-1804 (2015). https://doi.org:10.1007/s00253-014-6233-2

276 Pásztor, A., Kallio, P., Malatinszky, D., Akhtar, M. K. & Jones, P. R. A synthetic O2 - tolerant butanol pathway exploiting native fatty acid biosynthesis in *Escherichia coli*. *Biotechnology and Bioengineering* **112**, 120-128 (2015). https://doi.org:10.1002/bit.25324

277 Akhtar, M. K., Turner, N. J. & Jones, P. R. Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities. *Procedings of the National Academy of Science* **110**, 87-92 (2013). https://doi.org:10.1073/pnas.1216516110

278 Butler, N. & Kunjapur, A. M. Carboxylic acid reductases in metabolic engineering. *Journal of Biotechnology* **307**, 1-14 (2020). https://doi.org:10.1016/j.jbiotec.2019.10.002

279 Finnigan, W. *et al.* Characterization of Carboxylic Acid Reductases as Enzymes in the Toolbox for Synthetic Chemistry. *ChemCatChem* **9**, 1005-1017 (2017). https://doi.org:10.1002/cctc.201601249

280 Moura, M. *et al.* Characterizing and predicting carboxylic acid reductase activity for diversifying bioaldehyde production. *Biotechnology and Bioengineering* **113**, 944-952 (2016). https://doi.org:10.1002/bit.25860

281 Winkler, M. Carboxylic acid reductase enzymes (CARs). *Current Opinion in Chemical Biology* **43**, 23-29 (2018). https://doi.org:10.1016/j.cbpa.2017.10.006

282 Cha, T.-Y. *et al.* Biosynthesis of C12 Fatty Alcohols by Whole Cell Biotransformation of C12 Derivatives Using *Escherichia coli* Two-cell Systems Expressing CAR and ADH. *Biotechnology and Bioprocess Engineering* **26**, 392-401 (2021). https://doi.org:10.1007/s12257-020-0239-7

283 Venkitasubramanian, P., Daniels, L. & Rosazza, J. P. Reduction of carboxylic acids by *Nocardia* aldehyde oxidoreductase requires a phosphopantetheinylated enzyme. *Journal of Biological Chemistry* **282**, 478-485 (2007). https://doi.org:10.1074/jbc.M607980200

284 Horvat, M. & Winkler, M. In VivoReduction of Medium‐ to Long‐Chain Fatty Acids by Carboxylic Acid Reductase (CAR) Enzymes: Limitations and Solutions. *ChemCatChem* **12**, 5076-5090 (2020). https://doi.org:10.1002/cctc.202000895

285 Maurer, S., Schewe, H., Schrader, J. & Buchhaupt, M. Investigation of fatty aldehyde and alcohol synthesis from fatty acids by alphaDox- or CAR-expressing *Escherichia coli*. *Journal of Biotechnology* **305**, 11-17 (2019). https://doi.org:10.1016/j.jbiotec.2019.08.011

286 Rodriguez, G. M. & Atsumi, S. Toward aldehyde and alkane production by removing aldehyde reductase activity in *Escherichia coli*. *Metabolic Engineering* **25**, 227-237 (2014). https://doi.org:10.1016/j.ymben.2014.07.012

287 Croteau, R. & Hooper, C. L. Metabolis of Monoterpenes Acetylation of (-)-Menthol By A Soluble Enzyme Preparation From Peppermint (*Mentha piperita*) Leaves. *Plant Physiology* **61**, 737-742 (1978).

288 Yoshioka, K. & Hashimoto, N. Ester Formation by Alcohol Acetyltransferase from Brewers' Yeast. *Agricultural and Biological Chemistry* **45**, 2183-2190 (1980). https://doi.org:10.1080/00021369.1981.10864861

289 Balbontin, C. *et al.* VpAAT1, a gene encoding an alcohol acyltransferase, is involved in ester biosynthesis during ripening of mountain papaya fruit. *Journal of Agricultural and Food Chemistry* **58**, 5114-5121 (2010). https://doi.org:10.1021/jf904296c

290 Lucchetta, L. *et al.* Biochemical and Catalytic Properties of Three Recombinant Alcohol Acyltransferases of Melon, Sulfur-Containing Ester Formation, Regulatory Role of CoA-SH in Activity, and Sequence Elements Conferring Substrate Preference. *Journal of Agricultural and Food Chemistry* **55**, 5213-5220 (2007).

291 Beekwilder, J. *et al.* Functional characterization of enzymes forming volatile esters from strawberry and banana. *Plant Physiology* **135**, 1865-1878 (2004). https://doi.org:10.1104/pp.104.042580

292 Cumplido-Laso, G. *et al.* The fruit ripening-related gene *FaAAT2* encodes an acyl transferase involved in strawberry aroma biogenesis. *Journal of Experimental Botany* **63**, 4275-4290 (2012). https://doi.org:10.1093/jxb/ers120

293 D'Auria, J. C., Chen, F. & Pichersky, E. Characterization of an acyltransferase capable of synthesizing benzylbenzoate and other volatile esters in flowers and damaged leaves of *Clarkia breweri*. *Plant Physiology* **130**, 466-476 (2002). https://doi.org:10.1104/pp.006460

294 Dhar, N. *et al.* Characterization of a sweet basil acyltransferase involved in eugenol biosynthesis. *Journal of Experimental Botany* **71**, 3638-3652 (2020). https://doi.org:10.1093/jxb/eraa142

295 El-Sharkawy, I. *et al.* Functional characterization of a melon alcohol acyl-transferase gene family involved in the biosynthesis of ester volatiles. Identification of the crucial role of a threonine residue for enzyme activity. *Plant Molecular Biology* **59**, 345-362 (2005). https://doi.org:10.1007/s11103-005-8884-y

296 Gonzalez, M. *et al.* Aroma development during ripening of *Fragaria chiloensis* fruit and participation of an alcohol acyltransferase (*FcAAT1*) gene. *Journal of Agricultural and Food Chemistry* **57**, 9123-9132 (2009). https://doi.org:10.1021/jf901693j

297 Goulet, C. *et al.* Divergence in the Enzymatic Activities of a Tomato and *Solanum penellii* Alcohol Acyltransferase Impacts Fruit Volatile Ester Composition. *Molecular Plant* **8**, 153-162 (2015).

298 Gunther, C. S., Chervin, C., Marsh, K. B., Newcomb, R. D. & Souleyre, E. J. F. Characterisation of two alcohol acyltransferases from kiwifruit (*Actinidia* spp.) reveals distinct substrate preferences. *Phytochemistry* **72**, 700-710 (2011).

299 Guterman, I. *et al.* Rose scent: genomics approach to discovering novel floral fragrance-related genes. *Plant Cell* **14**, 2325-2338 (2002). https://doi.org:10.1105/tpc.005207

300 Kamiyoshihara, Y. *et al.* Functional divergence of principal alcohol o-acyltransferase for biosynthesis of volatile acetate esters among tomato wild species *(Solanum Sect. Lycopersicon*). *Plant Science* **300**, 110612 (2020). https://doi.org:10.1016/j.plantsci.2020.110612

301 Li, D. *et al.* Molecular cloning and expression of a gene encodinng alcohol acyltransferase (*MdAAT2*) from apple (cv. Golden Delicious). *Phytochemistry* **67**, 658-667 (2006).

302 Maoz, I. *et al.* Concealed ester formation and amino acid metabolism to volatile compounds in table grape (*Vitis vinifera L.*) berries. *Plant Science* **274**, 223-230 (2018). https://doi.org:10.1016/j.plantsci.2018.05.020

303 Mitra, M. *et al.* Characterization of an alcohol acetyltransferase GcAAT responsible for the production of antifungal volatile esters in endophytic *Geotrichum candidum* PF005. *Microbiological Research* **260**, 127021 (2022). https://doi.org:10.1016/j.micres.2022.127021

304 Perez, A. G., Sanz, C. & Olias, J. Partial Purification and Some Properties of Alcohol Acyltransferase from Strawberry Fruits. *Journal of Agricultural and Food Chemistry* **41**, 1462-1466 (1993).

305 Qian, X. *et al.* Alcohol acyltransferase gene and ester precursors differentiate composition of volatile esters in three interspecific hybrids of *Vitis labrusca* x *V. Vinifera* during berry development period. *Food Chemistry* **295**, 234-246 (2019). https://doi.org:10.1016/j.foodchem.2019.05.104

306 Qiang, T. Y. *et al.* Identification, Molecular Cloning, and Functional Characterization of a Coniferyl Alcohol Acyltransferase Involved in the Biosynthesis of Dibenzocyclooctadiene Lignans in *Schisandra chinensis*. *Frontiers in Plant Science* **13**, 881342 (2022). https://doi.org:10.3389/fpls.2022.881342

307 Rottig, A. & Steinbuchel, A. Acyltransferases in bacteria. *Microbiology and Molecular Biology Reviews* **77**, 277-321 (2013). https://doi.org:10.1128/MMBR.00010-13

308 Shalit, M. *et al.* Volatile ester formation in roses. Identification of an acetyl-coenzyme A. Geraniol/Citronellol acetyltransferase in developing rose petals. *Plant Physiology* **131**, 1868-1876 (2003). https://doi.org:10.1104/pp.102.018572

309 Souleyre, E. J., Greenwood, D. R., Friel, E. N., Karunairetnam, S. & Newcomb, R. D. An alcohol acyl transferase from apple (cv. Royal Gala), MpAAT1, produces esters involved in apple fruit flavor. *FEBS Journal* **272**, 3132-3144 (2005). https://doi.org:10.1111/j.1742-4658.2005.04732.x

310 Wang, J. & De Luca, V. The biosynthesis and regulation of biosynthesis of Concord grape fruit esters, including 'foxy' methylanthranilate. *Plant Journal* **44**, 606-619 (2005). https://doi.org:10.1111/j.1365-313X.2005.02552.x

311 Yahyaoui, F. E. *et al.* Molecular and biochemical characteristics of a gene encoding an alcohol acyl-transferase involved in the generation of aroma volatile esters during melon ripening. *European Journal of Biochemistry* **269**, 2359-2366 (2002). https://doi.org:10.1046/j.1432-1033.2002.02892.x

312 Yan, X. *et al.* Functional characterization and catalytic activity improvement of BAHD acyltransferase from *Celastrus angulatus Maxim*. *Planta* **252**, 6 (2020). https://doi.org:10.1007/s00425-020-03413-2

313 Zhang, T. *et al.* Genome-wide identification, characterization, expression and enzyme activity analysis of coniferyl alcohol acetyltransferase genes involved in eugenol biosynthesis in *Prunus mume*. *PLoS One* **14**, e0223974 (2019). https://doi.org:10.1371/journal.pone.0223974

314 Zhou, W., Kong, W., Yang, C., Feng, R. & Xi, W. Alcohol Acyltransferase Is Involved in the Biosynthesis of C6 Esters in Apricot (*Prunus armeniaca* L.) Fruit. *Frontiers in Plant Science* **12**, 763139 (2021). https://doi.org:10.3389/fpls.2021.763139

315 Fuji, T. *et al.* Molecular Cloning, Sequence Analysis and Expression of the Yeast Alcohol Acetyltransferase Gene. *Applied and Environmental Microbiology* **60**, 2786-2792 (1994).

316 Lilly, M. *et al.* The effect of increased yeast alcohol acetyltransferase and esterase activity on the flavour profiles of wine and distillates. *Yeast* **23**, 641-659 (2006). https://doi.org:10.1002/yea.1382

317 Mason, A. B. & Dufour, J.-P. Alcohol acetyltransferases and the significance of ester synthesis in yeast. *Yeast* **16**, 1287-1298 (2000).

318 Shalit, M. *et al.* Acetyl-CoA: Alcohol Acetyltransferase Activity and Aroma Formation in Ripening Melon Fruits. *Journal of Agricultural and Food Chemistry* **49**, 794-799 (2001).

319 Echeverria, G., Graell, J., Lopez, M. L. & Lara, I. Volatile production, quality and aroma-related enzyme activities during maturation of 'Fuji' apples. *Postharvest Biology and Technology* **31**, 217-227 (2003).

320 Holland, D. *et al.* Development and Varietal Differences in Volatile Ester Formation and Acetyl-CoA: Alcohol Acetyl Transferase Activities in Apple (*Malus domestica* Borkh.) Fruit. *Journal of Agricultural and Food Chemistry* **53**, 7198-7203 (2005).

321 Hernandez, I., Molenaar, D., Beekwilder, J., Bouwmeester, H. & van Hylckama Vlieg, J. E. Expression of plant flavor genes in *Lactococcus lactis*. *Applied and Environmental Microbiology* **73**, 1544-1552 (2007). https://doi.org:10.1128/AEM.01870-06

322 Fukuda, K. *et al.* Balance of Activities of Alcohol Acetyltransferase and Esterase in *Saccharomyces cerevisiae* is Important for Production of Isoamyl Acetate. *Applied and Environmental Microbiology* **64**, 4076-4078 (1998).

323 Shockey, J. M., Fulda, M. S. & Browse, J. *Arabidopsis* contains a large superfamily of acyl-activating enzymes. Phylogenetic and biochemical analysis reveals a new class of acyl-coenzyme a synthetases. *Plant Physiology* **132**, 1065-1076 (2003). https://doi.org:10.1104/pp.103.020552

324 Sofeo, N. *et al.* Altering the Substrate Specificity of Acetyl-CoA Synthetase by Rational Mutagenesis of the Carboxylate Binding Pocket. *ACS Synthetic Biology* **8**, 1325-1336 (2019). https://doi.org:10.1021/acssynbio.9b00008

325 Hawkins, A. S., Han, Y., Bennett, R. K., Adams, M. W. & Kelly, R. M. Role of 4-hydroxybutyrate-CoA synthetase in the CO2 fixation cycle in thermoacidophilic archaea. *Journal of Biological Chemistry* **288**, 4012-4022 (2013). https://doi.org:10.1074/jbc.M112.413195

326 Buttery, R. G., Seifert, R. M., Guadagni, D. G. & Ling, L. C. Characterization of Some Volatile Constituents of Bell Peppers. *Journal of Agricultural and Food Research* **17**, 1322-1327 (1969).

327 Coton, E., Coton, M., Hymery, N., Mounier, J. & Jany, J.-L. *Penicillium roqueforti*: an overview of its genetics, physiology, metabolism and biotechnological applications. *Fungal Biology Reviews* **34**, 59-73 (2020). https://doi.org:10.1016/j.fbr.2020.03.001

328 Forney, F. W. & Markovetz, A. J. The biology of methyl ketones. *Journal of Lipid Research* **12**, 383-395 (1971). https://doi.org:10.1016/s0022-2275(20)39487-6

329 Kinsella, J. E. & Hwang, D. Biosynthesis of Flavors by *Penicillium roqueforti*. *Biotechnology and Bioengineering* **18**, 927-938 (1976).

330 Okumura, J. & Kinsella, J. E. Methyl Ketone Formation by *Penicillium camemberti* in Model Systems. *Journal of Diary Science* (1985).

331 Lawrence, R. C. & Hawke, J. C. The Oxidation of Fatty Acids by Mycelium of *Penicillium roqueforti*. *Journal of General Microbiology* **51**, 289-302 (1968).

332 Fridman, E. *et al.* Metabolic, genomic, and biochemical analyses of glandular trichomes from the wild tomato species *Lycopersicon hirsutum* identify a key enzyme in the biosynthesis of methylketones. *The Plant Cell* **17**, 1252-1267 (2005). https://doi.org:10.1105/tpc.104.029736

333 Yu, G. *et al.* Enzymatic functions of wild tomato methylketone synthases 1 and 2. *Plant Physiology* **154**, 67-77 (2010). https://doi.org:10.1104/pp.110.157073

334 Tran, H. T. D., Le, N. T., Khuat, V. L. U. & Nguyen, T. T. H. Identification and Functional Characterization of a Soybean (*Glycine max*) Thioesterase that Acts on Intermediates of Fatty Acid Biosynthesis. *Plants (Basel)* **8** (2019). https://doi.org:10.3390/plants8100397

335 Khuat, V. L. U. *et al.* Characterization of *Solanum melongena* Thioesterases Related to Tomato Methylketone Synthase 2. *Genes (Basel)* **10** (2019). https://doi.org:10.3390/genes10070549

336 Pulsifer, I. P. *et al.* Acyl-lipid thioesterase1-4 from *Arabidopsis thaliana* form a novel family of fatty acyl-acyl carrier protein thioesterases with divergent expression patterns and substrate specificities. *Plant Molecular Biology* **84**, 549-563 (2014). https://doi.org:10.1007/s11103-013-0151-z

337 Kalinger, R. S., Pulsifer, I. P. & Rowland, O. Elucidating the substrate specificities of acyl-lipid thioesterases from diverse plant taxa. *Plant Physiology and Biochemistry* **127**, 104-118 (2018). https://doi.org:10.1016/j.plaphy.2018.03.013

338 Zhu, M. *et al.* Biosynthesis of the Nematode Attractant 2-Heptanone and Its Co-evolution Between the Pathogenic Bacterium *Bacillus nematocida* and Non-pathogenic Bacterium *Bacillus subtilis*. *Frontiers in Microbiology* **10**, 1489 (2019). https://doi.org:10.3389/fmicb.2019.01489

339 Nie, L., Ren, Y. & Schulz, H. Identification and Characterization of *Escherichia coli* Thioesterase III That Functions in Fatty Acid Beta-Oxidation. *Biochemistry* **47**, 7744-7751 (2008).

340 Goh, E. B. *et al.* Substantial improvements in methyl ketone production in E. coli and insights on the pathway from in vitro studies. *Metabolic Engineering* **26**, 67-76 (2014). https://doi.org:10.1016/j.ymben.2014.09.003

341 Goh, E. B., Chen, Y., Petzold, C. J., Keasling, J. D. & Beller, H. R. Improving methyl ketone production in *Escherichia coli* by heterologous expression of NADH-dependent FabG. *Biotechnology and Bioengineering* **115**, 1161-1172 (2018). https://doi.org:10.1002/bit.26558

342 Dong, J. *et al.* Methyl ketone production by *Pseudomonas putida* is enhanced by plant-derived amino acids. *Biotechnology and Bioengineering* **116**, 1909-1922 (2019). https://doi.org:10.1002/bit.26995

343 Hanko, E. K. R. *et al.* Engineering beta-oxidation in *Yarrowia lipolytica* for methyl ketone production. *Metabolic Engineering* **48**, 52-62 (2018). https://doi.org:10.1016/j.ymben.2018.05.018

344 Muller, J. *et al.* Engineering of *Ralstonia eutropha* H16 for autotrophic and heterotrophic production of methyl ketones. *Applied Environmental Microbiology* **79**, 4433-4439 (2013). https://doi.org:10.1128/AEM.00973-13

345 Nies, S. C. *et al.* High titer methyl ketone production with tailored *Pseudomonas taiwanensis* VLB120. *Metabolic Engineering* **62**, 84-94 (2020). https://doi.org:10.1016/j.ymben.2020.08.003

346 Zhang, G. *et al.* Dual beta-oxidation pathway and transcription factor engineering for methyl ketones production in *Saccharomyces cerevisiae*. *Metabolic Engineering* **73**, 225-234 (2022). https://doi.org:10.1016/j.ymben.2022.08.004

347 Yan, Q. *et al.* Metabolic engineering strategies to produce medium-chain oleochemicals via acyl-ACP:CoA transacylase activity. *Nature Communications* **13**, 1619 (2022). https://doi.org:10.1038/s41467-022-29218-3

348 Park, J. *et al.* Synthesis of methyl ketones by metabolically engineered *Escherichia coli*. *Journal of Industrial Microbiology and Biotechnology* **39**, 1703-1712 (2012). https://doi.org:10.1007/s10295-012-1178-x

349 Zhu, Z. *et al.* Expanding the product portfolio of fungal type I fatty acid synthases. *Nature Chemical Biology* **13**, 360-362 (2017). https://doi.org:10.1038/nchembio.2301

350 Srirangan, K. *et al.* Engineering *Escherichia coli* for Microbial Production of Butanone. *Applied Environmental Microbiology* **82**, 2574-2584 (2016). https://doi.org:10.1128/AEM.03964-15

351 Yuzawa, S. *et al.* Comprehensive in Vitro Analysis of Acyltransferase Domain Exchanges in Modular Polyketide Synthases and Its Application for Short-Chain Ketone Production. *ACS Synthetic Biology* **6**, 139-147 (2017). https://doi.org:10.1021/acssynbio.6b00176

352 Yuzawa, S. *et al.* Short-chain ketone production by engineered polyketide synthases in *Streptomyces albus*. *Nature Communications* **9**, 4569 (2018). https://doi.org:10.1038/s41467-018-07040-0

353 Yoneda, H., Tantillo, D. J. & Atsumi, S. Biological production of 2-butanone in *Escherichia coli*. *ChemSusChem* **7**, 92-95 (2014). https://doi.org:10.1002/cssc.201300853

354 Hatanaka, A. The Biogeneration of Green Odour by Green Leaves. *Phytochemistry* **34**, 1201-1218 (1993).

355 Matsui, K. Green leaf volatiles: hydroperoxide lyase pathway of oxylipin metabolism. *Current Opinion in Plant Biology* **9**, 274-280 (2006). https://doi.org:10.1016/j.pbi.2006.03.002

356 King, B. M. *et al.* Orthonasal and Retronasal Perception of some Green Leaf volatiles used in Beverage Flavours. *Journal of Agricultural and Food Chemistry* **54**, 2664-2670 (2006).

357 Wache, Y., Husson, F., Feron, G. & Belin, J. M. Yeast as an efficient biocatalyst for the production of lipid-derived flavours and fragrances. *Antonie Van Leeuwenhoek* **89**, 405-416 (2006). https://doi.org:10.1007/s10482-005-9049-3

358 Myung, K., Hamilton-Kemp, T. R. & Archbold, D. D. Biosynthesis of *trans*-2-Hexenal in Response to Wounding in Strawberry Fruit. *Journal of Agricultural and Food Chemistry* **54**, 1442-1448 (2006).

359 ul Hassan, M. N., Zainal, Z. & Ismail, I. Green leaf volatiles: biosynthesis, biological functions and their applications in biotechnology. *Plant Biotechnology Journal* **13**, 727-739 (2015). https://doi.org:10.1111/pbi.12368

360 Deng, W. *et al.* Effects of Six-Carbon Aldehydes and Alcohols on Bacterial Proliferation. *Journal of Agricultural and Food Chemistry* **41**, 506-510 (1993). https://doi.org:10.1021/jf00027a030,

361 Kamdem, S. S., Belletti, N., Magnani, R., Lanciotti, R. & Gardini, F. Effects of carvacrol, (E)-2-hexenal, and citral on the thermal death kinetics of *Listeria monocytogenes*. *Journal of Food Protection* **74**, 2070-2078 (2011). https://doi.org:10.4315/0362-028X.JFP-11-064

362 Nakamura, S. & Hatanaka, A. Green-Leaf-Derived C6-Aroma Compounds with Potent Antibacterial Action That Act on Both Gram-Negative and Gram-Positive Bacteria. *Journal of Agricultural and Food Chemistry* **50**, 7639-7644 (2002).

363 Gardini, F., Lanciotti, R. & Guerzoni, M. E. Effect of *trans*-2-hexenal on the growth of *Aspergillus flavus* in relation to its concentration, temperature and water activity. *Letters in Applied Microbiology* **33**, 50-55 (2001).

364 Ma, W., Zhao, L., Zhao, W. & Xie, Y. (E)-2-Hexenal, as a Potential Natural Antifungal Compound, Inhibits *Aspergillus flavus* Spore Germination by Disrupting Mitochondrial Energy Metabolism. *Journal of Agricultural and Food Chemistry* **67**, 1138-1145 (2019). https://doi.org:10.1021/acs.jafc.8b06367

365 Neri, F., Mari, M., Menniti, A. M. & Brigati, S. Activity of trans-2-hexenal against *Penicillium expansum* in 'Conference' pears. *Journal of Applied Microbiology* **100**, 1186-1193 (2006). https://doi.org:10.1111/j.1365-2672.2006.02873.x

366 Shiojiri, K. *et al.* Changing green leaf volatile biosynthesis in plants: An approach for improving plant resistance against both herbivores and pathogens. *Proceedings of the National Academy of Sciences* **103**, 16672-16676 (2006).

367 Cheng, L. *et al.* Effects of *trans*-2-hexenal on reproduction, growth and behaviour and efficacy against the pinewood nematode, *Bursaphelenchus xylophilus*. *Pest Management Science* **73**, 888-895 (2017). https://doi.org:10.1002/ps.4360

368 Zhao, Y. *et al.* Effects of the plant volatile *trans*-2-hexenal on the dispersal ability, nutrient metabolism and enzymatic activities of *Bursaphelenchus xylophilus*. *Pesticide Biochemistry and Physiology* **143**, 147-153 (2017).

369 Derbassi, N. B. *et al.* Plant volatiles: Using Scented molecules as food additives. *Trends in Food Science & Technology* **122**, 97-103 (2022). https://doi.org:10.1016/j.tifs.2022.02.002

370 Hatanaka, A. & Harada, T. Formation of *Cis*-3-Hexenal, *Trans*-2-Hexenal and *Cis*-3-Hexenol in Macerated *Thea sinensis* Leaves. *Phytochemistry* **12** (1973).

371 Major, R. T. The Ginkgo, the Most Ancient Living Tree. *Science* **157**, 1270-1273 (1967).

372 Feussner, I. & Wasternack, C. The lipoxygenase pathway. *Annual Review of Plant Biology* **53**, 275-297 (2002). https://doi.org:10.1146/annurev.arplant.53.100301.135248

373 Matsui, K. Properties and Structures of Fatty Acid Hydroperoxide Lyase. *Belgian Journal of Botany* **131**, 50-62 (1998).

374 Kunishima, M. *et al.* Identification of (Z)-3:(E)-2-Hexenal Isomerases Essential to the Production of the Leaf Aldehyde in Plants. *Journal of Biological Chemistry* **291**, 14023-14033 (2016). https://doi.org:10.1074/jbc.M116.726687

375 Noordermeet, M. A., Veldink, G. A. & Vliegenthart, J. F. G. Alfafa contains substantial 9-hyderoperoxide lyase activity and a 3Z:2E-enal iosmerase. *FEBS Letters* **443**, 201-204 (1998).

376 Spyropoulou, E. A. *et al.* Identification and Characterization of (3Z):(2E)-Hexenal Isomerases from Cucumber. *Frontiers in Plant Science* **8**, 1342 (2017). https://doi.org:10.3389/fpls.2017.01342

377 Mano, J. *et al.* The NADPH:Quinone Oxidoreductase P1-zeta-crystallin in *Arabidopsis* Catalyses the alpha,beta-Hydrogenation of 2-Alkenals: Detoxication of the Lipid Peroxide-Derived Reactive Aldehydes. *Plant Cell Physiology* **43**, 1445-1455 (2002).

378 Hatanaka, A. & Harada, T. Purification and Properties of Alcohol Dehydrogenase from Tea Seeds. *Agricultural and Biological Chemistry* **36**, 2033-2035 (1972). https://doi.org:10.1080/00021369.1972.10860516

379 D'Auria, J. C., Pichersky, E., Schaub, A., Hansel, A. & Gershenzon, J. Characterization of a BAHD acyltransferase responsible for producing the green leaf volatile (Z)-3-hexen-1-yl acetate in *Arabidopsis thaliana*. *The Plant Journal* **49**, 194-207 (2007). https://doi.org:10.1111/j.1365-313X.2006.02946.x

380 Tawfik, M. M., Yamato, K. T., Kohchi, T., Koeduka, T. & Matsui, K. n-Hexanal and (Z)-3-hexenal are generated from arachidonic acid and linolenic acid by a lipoxygenase in *Marchantia polymorpha* L. *Bioscience, Biotechnology, and Biochemistry* **81**, 1148-1155 (2017). https://doi.org:10.1080/09168451.2017.1285688

381 Jacopini, S. *et al.* Olive Recombinant Hydroperoxide Lyase, an Efficient Biocatalyst for Synthesis of Green Leaf Volatiles. *Applied Biochemistry and Biotechnology* **179**, 671-683 (2016). https://doi.org:10.1007/s12010-016-2023-x

382 Kaur, I. *et al.* Biosynthesis of natural aroma compounds using recombinant whole-cell tomato hydroperoxide lyase biocatalyst. *Journal of Biosciences* **47** (2022). https://doi.org:10.1007/s12038-022-00269-4

383 Buchhaupt, M., Guder, J. C., Etschmann, M. M. & Schrader, J. Synthesis of green note aroma compounds by biotransformation of fatty acids using yeast cells coexpressing lipoxygenase and hydroperoxide lyase. *Applied Microbiology and Biotechnology* **93**, 159-168 (2012). https://doi.org:10.1007/s00253-011-3482-1

384 Małajowicz, J. & Kozłowska, M. Factors Affecting the Yield in Formation of Fat-Derived Fragrance Compounds by *Yarrowia lipolytica* Yeast. *Applied Sciences* **11** (2021). https://doi.org:10.3390/app11219843

385 Hatanaka, A., Kajiwara, T., Horino, H. & Inokuchi, K. Odor-Structure Relationships in n-Hexenols and n-Hexenals. *Zeitschrift für Naturforschung* **47**, 183-189 (1992).

386 Barber, C. E. *et al.* A novel regulatory system required for pathogenicity of Xanthomonas campestris is mediated by a small diffusible signal molecule. *Molecular Microbiology* **243**, 555-566 (1997).

387 Deng, Y., Wu, J., Tao, F. & Zhang, L. H. Listening to a New Language: DSF-Based Quorum Sensing in Gram-Negative Bacteria. *Chemical Reviews* **111**, 160-173 (2011).

388 Boon, C. *et al.* A novel DSF-like signal from *Burkholderia cenocepacia* interferes with *Candida albicans* morphological transition. *The ISME Journal* **2**, 27-36 (2008). https://doi.org:10.1038/ismej.2007.76

389 He, Y. W., Wu, J., Cha, J. S. & Zhang, L. H. Rice bacterial blight pathogen *Xanthomonas oryzae* pv. *oryzae* produces multiple DSF-family signals in regulation of virulence factor production. *BMC Microbiology* **10** (2010).

390 Deng, Y., Wu, J., Eberl, L. & Zhang, L. H. Structural and functional characterization of diffusible signal factor family quorum-sensing signals produced by members of the *Burkholderia cepacia* complex. *Applied and Environmental Microbiology* **76**, 4675-4683 (2010). https://doi.org:10.1128/AEM.00480-10

391 Davies, D. G. & Marques, C. N. A fatty acid messenger is responsible for inducing dispersion in microbial biofilms. *Journal of Bacteriology* **191**, 1393-1403 (2009). https://doi.org:10.1128/JB.01214-08

392 Vilchez, R. *et al. Streptococcus mutans* inhibits *Candida albicans* hyphal formation by the fatty acid signaling molecule trans-2-decenoic acid (SDSF). *Chembiochem* **11**, 1552-1562 (2010). https://doi.org:10.1002/cbic.201000086

393 Amari, D. T., Marques, C. N. & Davies, D. G. The putative enoyl-coenzyme A hydratase DspI is required for production of the *Pseudomonas aeruginosa* biofilm dispersion autoinducer cis-2-decenoic acid. *Journal of Bacteriology* **195**, 4600-4610 (2013). https://doi.org:10.1128/JB.00707-13

394 Rahmani-Badi, A., Sepehr, S., Fallahi, H. & Heidari-Keshel, S. Dissection of the *cis*-2-decenoic acid signaling network in *Pseudomonas aeruginos* using microarray technique. *Frontiers in Microbiology* **6** (2015). https://doi.org:10.3389/fmicb.2015.00383

395 Bi, H., Christensen, Q. H., Feng, Y., Wang, H. & Cronan, J. E. The *Burkholderia cenocepacia* BDSF quorum sensing fatty acid is synthesized by a bifunctional crotonase homologue having both dehydratase and thioesterase activities. *Molecular Microbiology* **83**, 840-855 (2012). https://doi.org:10.1111/j.1365-2958.2012.07968.x

396 Spadaro, F. *et al.* The Crystal Structure of *Burkholderia cenocepacia* DfsA Provides Insights into Substrate Recognition and Quorum Sensing Fatty Acid Biosynthesis. *Biochemistry* **55**, 3241-3250 (2016). https://doi.org:10.1021/acs.biochem.6b00178

397 Akhtar, M. K. & Jones, P. R. Deletion of *iscR* stimulates recombinant clostridial Fe-Fe hydrogenase activity and H2-accumulation in *Escherichia coli* BL21(DE3). *Applied Microbiology and Biotechnology* **78**, 853-862 (2008). https://doi.org:10.1007/s00253-008-1377-6

398 Hanahan, D. Studies on Transformation of *Escherichia coli* with Plasmids. *Journal of Molecular Biology* **166**, 557-580 (1893).

399 Stanier, R. Y., Kunisawa, R., Mandel, M. & Cohen-Bazar, G. Purification and Properties of Unicellular Blue-Green Algae (Order *Chroococcales*). *Bacteriological Reviews* **35**, 171-205 (1971).

400 Elhai, J. & Wolk, C. P. Conjugal Transfer of DNA to Cyanobacteria. *Methods in Enzymology* **167**, 747-754 (1987).

401 Elhai, J., Vepritskiy, A., Muro-Pastor, A. M., Flores, E. & Wolk, C. P. Reduction of Conjugal Transfer Efficiency by Three Restriction Activities of *Anabaenai* sp. Strain PCC 7120. *Journal of Bacteriology* **179** (1997).

402 Storch, M. *et al.* BASIC: A New Biopart Assembly Standard for Idempotent Cloning Provides Accurate, Single-Tier DNA Assembly for Synthetic Biology. *ACS Synth Biol* **4**, 781-787 (2015). https://doi.org:10.1021/sb500356d

403 Emanuelsson, O., Nielsen, H. & von Heijne, G. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Science* **8**, 978-984 (1999). https://doi.org:10.1110/ps.8.5.978

404 Madeira, F. *et al.* Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research* (2022). https://doi.org:10.1093/nar/gkac240

405 Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10**, 845-858 (2015). https://doi.org:10.1038/nprot.2015.053

406 The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

407 Dong, R., Pan, S., Peng, Z., Zhang, Y. & Yang, J. mTM-align: a server for fast protein structure database search and multiple protein structure alignment. *Nucleic Acids Research* **46**, W380-W386 (2018). https://doi.org:10.1093/nar/gky430

408 Boratyn, G. M. *et al.* BLAST: a more efficient report with usability improvements. *Nucleic Acids Research* **41**, W29-33 (2013). https://doi.org:10.1093/nar/gkt282

409 Akhtar, M. K., Dandapani, H., Thiel, K. & Jones, P. R. Microbial production of 1-octanol: A naturally excreted biofuel with diesel-like properties. *Metabolic Engineering Communications* **2**, 1-5 (2015). https://doi.org:10.1016/j.meteno.2014.11.001

410 Sarria, S., Bartholow, T. G., Verga, A., Burkart, M. D. & Peralta-Yahya, P. Matching Protein Interfaces for Improved Medium-Chain Fatty Acid Production. *ACS Synthetic Biology* **7**, 1179-1187 (2018). https://doi.org:10.1021/acssynbio.7b00334

411 Post-Beittenmiller, M. A., Schmid, K. M. & Ohlrogge, J. B. Expression of Holo and Apo Forms of Spinach Acyl Carrier Protein-I in Leaves of Transgenic Tobacco Plants. *The Plant Cell* **1**, 889-899 (1989).

412 Jha, J. K. *et al.* Functional expression of an acyl carrier protein (ACP) from *Azospirillum brasilense* alters fatty acid profiles in *Escherichia coli* and *Brassica juncea*. *Plant Physiology and Biochemistry* **45**, 490-500 (2007).

413 Brannen, J. K. Overexpression of Acyl Carrier Protein-1 Alters Fatty Acid Composition of Leaf Tissue in Arabidopsis. *Plant Physiology* **127**, 222-229 (2001).

414 De Marchis, F. *et al.* Overexpression of the olive acyl carrier protein gene (OeACP1) produces alterations in fatty acid composition of tobacco leaves. *Transgenic Research* **25**, 45-61 (2016). https://doi.org:10.1007/s11248-015-9919-z

415 Liu, X., Hicks, W. M., Silver, P. A. & Way, J. C. Engineering acyl carrier protein to enhance production of shortened fatty acids. *Biotechnology and Biofuels* **9**, 24 (2016). https://doi.org:10.1186/s13068-016-0430-4

416 Blatti, J. L. *et al.* Manipulating fatty acid biosynthesis in microalgae for biofuel through protein-protein interactions. *PLoS One* **7**, e42949 (2012). https://doi.org:10.1371/journal.pone.0042949

417 De Lay, N. R. & Cronan, J. E. In vivo functional analyses of the type II acyl carrier proteins of fatty acid biosynthesis. *Journal of Biological Chemistry* **282**, 20319-20328 (2007). https://doi.org:10.1074/jbc.M703789200

418 Dong, H. & Cronan, J. E. The Two Acycl Carrier Protein of *Enteroccus faecalis* Have Nonredundant Functions. *Journal of Bacteriology* (2022).

419 Kopka, J., Robers, M., Schuch, R. & Spener, F. Acyl carrier proteins from developing seeds of *Cuphea lanceolata* Ait. *Planta* **191**, 102-111 (1993).

420 Voetz, M., B., K., Schell, J. & Topfer, R. Three Different cDNAs Encoding Acyl Carrier Proteins from *Cuphea lanceolata*. *Plant Physiology* **106**, 785-786 (1994).

421 Schutt, B. S., Brummel, M., Schuch, R. & Spener, F. The role of acyl carrier protein isoforms from *Cuphea lanceolata* seeds in the de-novo biosynthesis of medium-chain fatty acids. *Planta* **205**, 263-268 (1998).

422 Deng, X. *et al.* Structure-guided reshaping of the acyl binding pocket of 'TesA thioesterase enhances octanoic acid production in *E. coli*. *Metabolic Engineering* **61**, 24-32 (2020). https://doi.org:10.1016/j.ymben.2020.04.010

423 Simoni, R. D., Criddle, R. S. & Stumpf, P. K. Fat Metabolism in Higher Plants XXXI. Purification and Properties of Plant and Bacterial Acyl Carrier Proteins. *Journal of Biological Chemistry* **242**, 573-581 (1967). https://doi.org:10.1016/s0021-9258(18)96243-5

424 National Centre for Biotechnology Information. *PubChem Annotation Record for METHYL DECANOATE, Source: Hazardous Substances Data Bank (HSDB)*, 2022).

425 Li, R.-J. *et al.* Enhancing the Catalytic Performance of a CYP116B Monooxygenase by Transdomain Combination Mutagenesis. *ChemCatChem* **10**, 2962-2968 (2018). https://doi.org:10.1002/cctc.201800054

426 Meng, S., Ji, Y., Liu, L., Davari, M. D. & Schwaneberg, U. Modulating the Coupling Efficiency of P450 BM3 by Controlling Water Diffusion through Access Tunnel Engineering. *ChemSusChem* **15**, e202102434 (2022). https://doi.org:10.1002/cssc.202102434

427 Bartasun, P. *et al.* The effect of modulating the quantity of enzymes in a model ethanol pathway on metabolic flux in *Synechocystis* sp. PCC 6803. *PeerJ* **7**, e7529 (2019). https://doi.org:10.7717/peerj.7529

428 Kozaeva, E., Mol, V., Nikel, P. I. & Nielsen, A. T. High-throughput colorimetric assays optimized for detection of ketones and aldehydes produced by microbial cell factories. *Microbial Biotechnology* **15**, 2426-2438 (2022). https://doi.org:10.1111/1751-7915.14097

429 Bräu, B. & Sahm, H. Cloning and expression of the structural gene for pyruvate decarboxylase of *Zymonas mobilis* in *Escherichia coli*. *Archives of Microbiology* **144**, 296-301 (1986).

430 Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research* **46**, W296-W303 (2018). https://doi.org:10.1093/nar/gky427

431 Yan, Q. & Pfleger, B. F. Revisiting metabolic engineering strategies for microbial synthesis of oleochemicals. *Metabolic Engineering* **58**, 35-46 (2020). https://doi.org:10.1016/j.ymben.2019.04.009

432 Kim, H. J., Park, H. S. & Lee, I. S. Microbial metabolism of *trans*-2-dodecenal. *Natural Product Sciences*, 19-22 (2011).

433 Kim, S., Cheong, S. & Gonzalez, R. Engineering *Escherichia coli* for the synthesis of short- and medium-chain alpha,beta-unsaturated carboxylic acids. *Metab Eng* **36**, 90-98 (2016). https://doi.org:10.1016/j.ymben.2016.03.005

434 Porcelli, C. & Steinhaus, M. Molecular characterisation of an atypical coconut-like odour in cocoa. *European Food Research and Technology* **248**, 1513-1523 (2022). https://doi.org:10.1007/s00217-022-03981-5

435 Jaiswal, D. *et al.* Genome Features and Biochemical Characteristics of a Robust, Fast Growing and Naturally Transformable Cyanobacterium *Synechococcus elongatus* PCC 11801 Isolated from India. *Scientific Reports* **8**, 16632 (2018). https://doi.org:10.1038/s41598-018-34872-z

436 Yu, J. *et al. Synechococcus elongatus* UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO(2). *Scientific Reports* **5**, 8132 (2015). https://doi.org:10.1038/srep08132

437 Jones, P. R. Genetic instability in cyanobacteria - an elephant in the room? *Frontiers in Bioengineering and Biotechnology* **2**, 12 (2014). https://doi.org:10.3389/fbioe.2014.00012

438 Lange, L. *et al.* Developing a Sustainable and Circular Bio-Based Economy in EU: By Partnering Across Sectors, Upscaling and Using New Knowledge Faster, and For the Benefit of Climate, Environment & Biodiversity, and People & Business. *Frontiers in Bioengineering and Biotechnology* **8**, 619066 (2020). https://doi.org:10.3389/fbioe.2020.619066

Appendices

Appendix A: *Escherichia coli* strains used for bioproduction and feeding experiments in this work

Strain No.	*E. coli* Strain	Plasmid 1	Plasmid 2
Chapter 4			
1	JW1794-1 (ΔfadD)	pET-Plac-CvFatB1 WT-Sfp-CAR-KanR	-
2	JW1794-1 (ΔfadD)	pET-Plac-CvT1-Sfp-CAR-KanR	-
3	JW1794-1 (ΔfadD)	pET-Plac-CvT2-Sfp-CAR-KanR	-
4	JW1794-1 (ΔfadD)	pET-Plac-CvT3-Sfp-CAR-KanR	-
5	JW1794-1 (ΔfadD)	pET-Plac-CvT4-Sfp-CAR-KanR	-
6	JW1794-1 (ΔfadD)	pET-Plac-CvT5-Sfp-CAR-KanR	-
7	JW1794-1 (ΔfadD)	pET-Plac-CvT6-Sfp-CAR-KanR	-
8	JW1794-1 (ΔfadD)	pET-Plac-CvT7-Sfp-CAR-KanR	-
9	JW1794-1 (ΔfadD)	pET-Plac-CvT8-Sfp-CAR-KanR	-
10	JW1794-1 (ΔfadD)	pET-Plac-CvT9-Sfp-CAR-KanR	-
11	JW1794-1 (ΔfadD)	pET-Plac-CvT10-Sfp-CAR-KanR	-
12	JW1794-1 (ΔfadD)	pET-Plac-CvT11-Sfp-CAR-KanR	-
13	JW1794-1 (ΔfadD)	pET-Plac-CvT12-Sfp-CAR-KanR	-
14	JW1794-1 (ΔfadD)	pET-Plac-CvT13-Sfp-CAR-KanR	-
15	JW1794-1 (ΔfadD)	pET-Plac-CvT14-Sfp-CAR-KanR	-
16	JW1794-1 (ΔfadD)	pET-Plac-CvT15-Sfp-CAR-KanR	-
17	JW1794-1 (ΔfadD)	pET-Plac-CvT16-Sfp-CAR-KanR	-
18	JW1794-1 (ΔfadD)	pET-Plac-CvT17-Sfp-CAR-KanR	-
19	JW1794-1 (ΔfadD)	pET-Plac-CvT18-Sfp-CAR-KanR	-
20	JW1794-1 (ΔfadD)	pET-Plac-CvT19-Sfp-CAR-KanR	-
21	JW1794-1 (ΔfadD)	pET-Plac-CvT20-Sfp-CAR-KanR	-
22	JW1794-1 (ΔfadD)	pET-Plac-CvT21-Sfp-CAR-KanR	-
23	JW1794-1 (ΔfadD)	pET-Plac-UaFatB1 WT-Sfp-CAR-KanR	-
24	JW1794-1 (ΔfadD)	pET-Plac-UaT1-Sfp-CAR-KanR	-

25	JW1794-1 (ΔfadD)	pET-Plac-UaT2-Sfp-CAR-KanR	-
26	JW1794-1 (ΔfadD)	pET-Plac-UaT3-Sfp-CAR-KanR	-
27	JW1794-1 (ΔfadD)	pET-Plac-UaT4-Sfp-CAR-KanR	-
28	JW1794-1 (ΔfadD)	pET-Plac-UaT5-Sfp-CAR-KanR	-
29	JW1794-1 (ΔfadD)	pET-Plac-UaT6-Sfp-CAR-KanR	-
30	JW1794-1 (ΔfadD)	pET-Plac-UaT7-Sfp-CAR-KanR	-
31	JW1794-1 (ΔfadD)	pET-Plac-UaT8-Sfp-CAR-KanR	-
32	JW1794-1 (ΔfadD)	pET-Plac-UaT9-Sfp-CAR-KanR	-
33	JW1794-1 (ΔfadD)	pET-Plac-CvT3-Sfp-CAR-KanR	pCDF-PrhaBAD-Sfp-CAR-GmR
34	JW1794-1 (ΔfadD)	pET-Plac-CvT4-Sfp-CAR-KanR	pCDF-PrhaBAD-Sfp-CAR-GmR
35	JW1794-1 (ΔfadD)	pET-Plac-CvT5-Sfp-CAR-KanR	pCDF-PrhaBAD-Sfp-CAR-GmR
36	JW1794-1 (ΔfadD)	pET-Plac-CvT6-Sfp-CAR-KanR	pCDF-PrhaBAD-Sfp-CAR-GmR
37	JW1794-1 (ΔfadD)	pET-Plac-CvT7-Sfp-CAR-KanR	pCDF-PrhaBAD-Sfp-CAR-GmR
38	JW1794-1 (ΔfadD)	pET-Plac-CvT8-Sfp-CAR-KanR	pCDF-PrhaBAD-Sfp-CAR-GmR
39	JW1794-1 (ΔfadD)	pET-Plac-CvT3-Sfp-CAR-KanR	pCDF-PrhaBAD-EcACP-Sfp-CAR-GmR
40	JW1794-1 (ΔfadD)	pET-Plac-CvT3-Sfp-CAR-KanR	pCDF-PrhaBAD-6803ACP-Sfp-CAR-GmR
41	JW1794-1 (ΔfadD)	pET-Plac-CvT3-Sfp-CAR-KanR	pCDF-PrhaBAD-ClACP2-Sfp-CAR-GmR
42	JW1794-1 (ΔfadD)	pET-Plac-CvT4-Sfp-CAR-KanR	pCDF-PrhaBAD-EcACP-Sfp-CAR-GmR
43	JW1794-1 (ΔfadD)	pET-Plac-CvT4-Sfp-CAR-KanR	pCDF-PrhaBAD-6803ACP-Sfp-CAR-GmR
44	JW1794-1 (ΔfadD)	pET-Plac-CvT4-Sfp-CAR-KanR	pCDF-PrhaBAD-ClACP2-Sfp-CAR-GmR
45	JW1794-1 (ΔfadD)	pET-Plac-CvT5-Sfp-CAR-KanR	pCDF-PrhaBAD-EcACP-Sfp-CAR-GmR
46	JW1794-1 (ΔfadD)	pET-Plac-CvT5-Sfp-CAR-KanR	pCDF-PrhaBAD-6803ACP-Sfp-CAR-GmR
47	JW1794-1 (ΔfadD)	pET-Plac-CvT5-Sfp-CAR-KanR	pCDF-PrhaBAD-ClACP2-Sfp-CAR-GmR
48	JW1794-1 (ΔfadD)	pET-Plac-CvT6-Sfp-CAR-KanR	pCDF-PrhaBAD-EcACP-Sfp-CAR-GmR
49	JW1794-1 (ΔfadD)	pET-Plac-CvT6-Sfp-CAR-KanR	pCDF-PrhaBAD-6803ACP-Sfp-CAR-GmR
50	JW1794-1 (ΔfadD)	pET-Plac-CvT6-Sfp-CAR-KanR	pCDF-PrhaBAD-ClACP2-Sfp-CAR-GmR
51	JW1794-1 (ΔfadD)	pET-Plac-CvT7-Sfp-CAR-KanR	pCDF-PrhaBAD-EcACP-Sfp-CAR-GmR
52	JW1794-1 (ΔfadD)	pET-Plac-CvT7-Sfp-CAR-KanR	pCDF-PrhaBAD-6803ACP-Sfp-CAR-GmR
53	JW1794-1 (ΔfadD)	pET-Plac-CvT7-Sfp-CAR-KanR	pCDF-PrhaBAD-ClACP2-Sfp-CAR-GmR

54	JW1794-1 (ΔfadD)	pET-Plac-CvT8-Sfp-CAR-KanR	pCDF-PrhaBAD-EcACP-Sfp-CAR-GmR
55	JW1794-1 (ΔfadD)	pET-Plac-CvT8-Sfp-CAR-KanR	pCDF-PrhaBAD-6803ACP-Sfp-CAR-GmR
56	JW1794-1 (ΔfadD)	pET-Plac-CvT8-Sfp-CAR-KanR	pCDF-PrhaBAD-ClACP2-Sfp-CAR-GmR
57	JW1794-1 (ΔfadD)	pET-Plac-CvT5-Sfp-CAR-KanR	pCDF-PrhaBAD-EcACP-Sfp-CAR-GmR
58	JW1794-1 (ΔfadD)	pET-Plac-CvT5 M90G-Sfp-CAR-KanR	pCDF-PrhaBAD-EcACP-Sfp-CAR-GmR
59	JW1794-1 (ΔfadD)	pET-Plac-CvT5 W122G-Sfp-CAR-KanR	pCDF-PrhaBAD-EcACP-Sfp-CAR-GmR
60	JW1794-1 (ΔfadD)	pET-Plac-CvT5 W140G-Sfp-CAR-KanR	pCDF-PrhaBAD-EcACP-Sfp-CAR-GmR
61	JW1794-1 (ΔfadD)	pET-Plac-CvT5 W122V-Sfp-CAR-KanR	pCDF-PrhaBAD-EcACP-Sfp-CAR-GmR
62	JW1794-1 (ΔfadD)	pET-Plac-CvT5 W122M-Sfp-CAR-KanR	pCDF-PrhaBAD-EcACP-Sfp-CAR-GmR
63	JW1794-1 (ΔfadD)	pET-Plac-CvT5 W122F-Sfp-CAR-KanR	pCDF-PrhaBAD-EcACP-Sfp-CAR-GmR
Chapter 5			
1	JW1794-1 (ΔfadD)	pET-Plac-GFP-KanR	-
2	JW1794-1 (ΔfadD)	pET-Plac-BM3 WT-KanR	-
3	JW1794-1 (ΔfadD)	pET-Plac-BM3 F87A-KanR	-
4	JW1794-1 (ΔfadD)	pET-Plac-BM3 L188Q-KanR	-
5	JW1794-1 (ΔfadD)	pET-Plac-BM3 V78A F87A L188Q-KanR	-
6	JW1794-1 (ΔfadD)	pET-Plac-RbCYP450-KanR	-
7	JW1794-1 (ΔfadD)	pET-Plac-TtCYP450-KanR	-
8	BL21 WT	pET-Plac-GFP-GmR	-
9	BL21 WT	pET-Plac-BM3 WT-GmR	-
10	BL21 WT	pET-Plac-BM3 V78A F87A L188Q-GnR	-
11	BL21 WT	pET-Plac-RbCYP450-GmR	-
12	BL21 WT	pET-Plac-TtCYP450-GmR	-
13	BL21 ΔiscR	pET-Plac-GFP-GmR	-
14	BL21 ΔiscR	pET-Plac-BM3 WT-GmR	-
15	BL21 ΔiscR	pET-Plac-BM3 V78A F87A L188Q-GnR	-
16	BL21 ΔiscR	pET-Plac-RbCYP450-GmR	-
17	BL21 ΔiscR	pET-Plac-TtCYP450-GmR	-
18	JW1794-1 (ΔfadD)	pET-Plac-BM3 F87A L188Q-KanR	-
19	JW1794-1 (ΔfadD)	pET-Plac-BM3 F87A L188Q I401P-KanR	-
20	JW1794-1 (ΔfadD)	pET-Plac-BM3 V78A F87A-KanR	-
21	JW1794-1 (ΔfadD)	pET-Plac-BM3 V78A F87A L188Q I401P-KanR	-
22	JW1794-1 (ΔfadD)	pET-Plac-BM3 F87G-KanR	-
23	JW1794-1 (ΔfadD)	pET-Plac-BM3 F87V-KanR	-
24	JW1794-1 (ΔfadD)	pET-Plac-BM3 F87I-KanR	-
25	JW1794-1 (ΔfadD)	pET-Plac-BM3 F87G L188Q-KanR	-
26	JW1794-1 (ΔfadD)	pET-Plac-BM3 F87G I401P-KanR	-
27	JW1794-1 (ΔfadD)	pET-Plac-BM3 F87G L188Q I401P-KanR	-

28	JW1794-1 (ΔfadD)	pET-Plac-BM3 V78A F87A L188Q-KanR	pCDF-Plac-GFP-SpecR
29	JW1794-1 (ΔfadD)	pET-Plac-BM3 V78A F87A L188Q-KanR	pCDF-Plac-PoAlkK-SpecR
30	JW1794-1 (ΔfadD)	pET-Plac-BM3 V78A F87A L188Q-KanR	pCDF-Plac-PoAlk-PpAAT1-SpecR
31	JW1794-1 (ΔfadD)	pET-Plac-BM3 V78A F87A L188Q-KanR	pCDF-Plac-Pp0763-SpecR
32	JW1794-1 (ΔfadD)	pET-Plac-BM3 V78A F87A L188Q-KanR	pCDF-Plac-Pp0763-PpAAT1-SpecR
33	JW1794-1 (ΔfadD)	pET-Plac-BM3 WT-GapN-KanR	-
34	JW1794-1 (ΔfadD)	pET-Plac-BM3 WT-BmGDH2-KanR	-
35	JW1794-1 (ΔfadD)	pET-Plac-BM3 WT-BmGDH3-KanR	-
36	JW1794-1 (ΔfadD)	pET-Plac-BM3 V78A F87A L188Q-GapN-KanR	-
37	JW1794-1 (ΔfadD)	pET-Plac-BM3 V78A F87A L188Q-BmGDH2-KanR	-
38	JW1794-1 (ΔfadD)	pET-Plac-BM3 V78A F87A L188Q-BmGDH3-KanR	-
39	JW1794-1 (ΔfadD)	pET-Plac-BM3 F87G L188Q-GapN-KanR	-
40	JW1794-1 (ΔfadD)	pET-Plac-BM3 F87G L188Q-BmGDH2-KanR	-
41	JW1794-1 (ΔfadD)	pET-Plac-BM3 F87G L188Q-BmGDH3-KanR	-
42	ΔfadD ΔiscR	pET-Plac-GFP-KanR	-
43	ΔfadD ΔiscR	pET-Plac-BM3 WT-KanR	-
44	ΔfadD ΔiscR	pET-Plac-BM3 V78A F87A L188Q-KanR	-
45	ΔfadD ΔiscR	pET-Plac-BM3 F87G L188Q-KanR	-
46	JW1794-1 (ΔfadD)	pET-Plac-BM3 V78A F87A L188Q D182S-KanR	-
47	JW1794-1 (ΔfadD)	pET-Plac-BM3 V78A F87A L188Q H266T-KanR	-
48	JW1794-1 (ΔfadD)	pET-Plac-BM3 V78A F87A L188Q A328V-KanR	-
49	JW1794-1 (ΔfadD)	pET-Plac-BM3 F87G L188Q D182S-KanR	-
50	JW1794-1 (ΔfadD)	pET-Plac-BM3 F87G L188Q H266T-KanR	-
51	JW1794-1 (ΔfadD)	pET-Plac-BM3 F87G L188Q A328V-KanR	-
Chapter 6			
1	JW1794-1 (ΔfadD)	pET-Plac-CvT5 W122F-Sfp-CAR-KanR	pCDF-PrhaBAD-GFP-GmR
2	JW1794-1 (ΔfadD)	pET-Plac-CvT5 W122F-Sfp-CAR-KanR	pCDF-PrhaBAD-EcACP-BM3 WT-GmR
3	JW1794-1 (ΔfadD)	pET-Plac-CvT5 W122F-Sfp-CAR-KanR	pCDF-PrhaBAD-EcACP-BM3 F87A L188Q I401P-GmR
4	JW1794-1 (ΔfadD)	pET-Plac-CvT5 W122F-Sfp-CAR-KanR	pCDF-PrhaBAD-EcACP-BM3 V78A F87A L188Q-GmR
5	JW1794-1 (ΔfadD)	pET-Plac-CvT5 W122F-Sfp-CAR-KanR	-

6	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBS3-CvT5 W122F-SpecR	pET-Plac-GFP-KanR
7	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-GFP-KanR
8	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBS2-CvT5 W122F-SpecR	pET-Plac-GFP-KanR
9	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA03-CvT5 W122F-SpecR	pET-Plac-GFP-KanR
10	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA08-CvT5 W122F-SpecR	pET-Plac-GFP-KanR
11	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA01-CvT5 W122F-SpecR	pET-Plac-GFP-KanR
12	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA09-CvT5 W122F-SpecR	pET-Plac-GFP-KanR
13	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBS1-CvT5 W122F-SpecR	pET-Plac-GFP-KanR
14	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-BM3 WT-KanR
15	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-BM3 V78A F87A L188Q-KanR
16	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-BM3 F87G L188Q-KanR
17	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-BM3 V78A F87A L188Q A328V-KanR
18	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-RbxBM3-KanR
19	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-TtxBM3-KanR
20	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-CvT5 W122F-RBS3-EcACP-SpecR	pET-Plac-GFP-KanR
21	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3- CvT5 W122F -RBA07-EcACP-SpecR	pET-Plac-GFP-KanR
22	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3- CvT5 W122F -RBS2-EcACP-SpecR	pET-Plac-GFP-KanR
23	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3- CvT5 W122F -RBSA03-EcACP-SpecR	pET-Plac-GFP-KanR
24	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3- CvT5 W122F -RBSA08-EcACP-SpecR	pET-Plac-GFP-KanR
25	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3- CvT5 W122F -RBSA01-EcACP-SpecR	pET-Plac-GFP-KanR
26	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3- CvT5 W122F -RBSA09-EcACP-SpecR	pET-Plac-GFP-KanR

27	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3- CvT5 W122F -RBS1-EcACP-SpecR	pET-Plac-GFP-KanR
28	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBS3-CvT5 W122F-SpecR	pET-Plac-GFP-KanR
29	JW1794-1 (ΔfadD)	pCDF-Plac-RBSA07-EcACP-RBS3-CvT5 W122F-SpecR	pET-Plac-GFP-KanR
30	JW1794-1 (ΔfadD)	pCDF-Plac-RBS2-EcACP-RBS3-CvT5 W122F-SpecR	pET-Plac-GFP-KanR
31	JW1794-1 (ΔfadD)	pCDF-Plac-RBSA03-EcACP-RBS3-CvT5 W122F-SpecR	pET-Plac-GFP-KanR
32	JW1794-1 (ΔfadD)	pCDF-Plac-RBSA08-EcACP-RBS3-CvT5 W122F-SpecR	pET-Plac-GFP-KanR
33	JW1794-1 (ΔfadD)	pCDF-Plac-RBSA01-EcACP-RBS3-CvT5 W122F-SpecR	pET-Plac-GFP-KanR
34	JW1794-1 (ΔfadD)	pCDF-Plac-RBSA09-EcACP-RBS3-CvT5 W122F-SpecR	pET-Plac-GFP-KanR
35	JW1794-1 (ΔfadD)	pCDF-Plac-RBS1-EcACP-RBS3-CvT5 W122F-SpecR	pET-Plac-GFP-KanR
36	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-RBSA07-BM3 F87G L188Q-KanR
37	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-RBS2-BM3 F87G L188Q-KanR
38	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-RBSA03-BM3 F87G L188Q-KanR
39	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-RBSA08-BM3 F87G L188Q-KanR
40	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-RBSA01-BM3 F87G L188Q-KanR
41	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-RBSA09-BM3 F87G L188Q-KanR
42	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-RBS1-BM3 F87G L188Q-KanR
43	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-RBSA07- BM3 V78A F87A L188Q A328V -KanR
44	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-RBS2- BM3 V78A F87A L188Q A328V -KanR
45	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-RBSA03- BM3 V78A F87A L188Q A328V -KanR
46	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-RBSA08-Plac- BM3 V78A F87A L188Q A328V -KanR
47	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-RBSA01- BM3 V78A F87A L188Q A328V -KanR
48	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-RBSA09- BM3 V78A F87A L188Q A328V -KanR

49	JW1794-1 (ΔfadD)	pCDF-Plac-RBS3-EcACP-RBSA07-CvT5 W122F-SpecR	pET-Plac-RBS1- BM3 V78A F87A L188Q A328V KanR
Chapter 7			
1	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	-
2	JW1794-1 (ΔfadD)	pET-Plac-BtTes-Sfp-CAR-KanR	-
3	JW1794-1 (ΔfadD)	pET-Plac-MfTes-Sfp-CAR-KanR	-
4	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pCDF-Plac-GFP-SpecR
5	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pCDF-Plac-Atf1-SpecR
6	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pCDF-Plac-CAT-SpecR
7	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pCDF-Plac-SAAT-SpecR
8	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pCDF-Plac-VAAT-SpecR
9	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pCDF-Plac-VpAAT1-SpecR
10	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-GFP-GmR
11	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-Atf1-GmR
12	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-CAT-GmR
13	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-SAAT-GmR
14	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-VAAT-GmR
15	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-VpAAT1-GmR
16	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-AtHPFT-GmR
17	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-CcVS-GmR
18	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-CsVS-GmR
19	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-LaAAT1-GmR
20	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-LaAAT2-GmR
21	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-LaAT1-GmR
22	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-LiAAT4-GmR
23	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-NbBBT-GmR
24	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-NtSHCT-GmR
25	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-RsVS-GmR
26	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-SsMMT-GmR
27	BL21 WT	pET-Plac-Tes4-Sfp-CAR-KanR	-

28	BL21 WT	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-GFP-GmR
29	BW25113 WT	pET-Plac-Tes4-Sfp-CAR-KanR	-
30	BW25113 WT	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-GFP-GmR
31	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	-
32	JW1794-1 (ΔfadD)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-GFP-GmR
33	JW5020-1 (ΔfadE)	pET-Plac-Tes4-Sfp-CAR-KanR	-
34	JW5020-1 (ΔfadE)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-GFP-GmR
35	JW5020-1 (ΔfadE)	pET-Plac-Tes4-Sfp-CAR-KanR	pCDF-Plac-GFP-SpecR
36	JW5020-1 (ΔfadE)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-Plac-GFP-GmR
37	JW5020-1 (ΔfadE)	pET-Plac-Tes4-Sfp-CAR-KanR	pCDF-PrhaBAD-GFP-GmR
38	JW5020-1 (ΔfadE)	pET-Plac-Tes4-Sfp-CAR-KanR	pACYC-PrhaBAD-GFP-EryR
39	JW1794-1 (ΔfadD)	pACYC-PrhaBAD-GFP-GmR	-
40	JW1794-1 (ΔfadD)	pACYC-PrhaBAD-Atf1-GmR	-
41	JW5020-1 (ΔfadE)	pACYC-PrhaBAD-CAT-GmR	-
42	JW5020-1 (ΔfadE)	pACYC-PrhaBAD-SAAT-GmR	-
43	JW5020-1 (ΔfadE)	pACYC-PrhaBAD-VAAT-GmR	-
44	JW5020-1 (ΔfadE)	pACYC-PrhaBAD-VpAAT1-GmR	-
45	JW5020-1 (ΔfadE)	pACYC-PrhaBAD-AtHPFT-GmR	-
46	JW5020-1 (ΔfadE)	pACYC-PrhaBAD-CcVS-GmR	-
47	JW1794-1 (ΔfadD)	pACYC-PrhaBAD-CsVS-GmR	-
48	JW1794-1 (ΔfadD)	pACYC-PrhaBAD-LaAAT1-GmR	-
49	JW5020-1 (ΔfadE)	pACYC-PrhaBAD-LaAAT2-GmR	-
50	JW5020-1 (ΔfadE)	pACYC-PrhaBAD-LaAT1-GmR	-
51	JW5020-1 (ΔfadE)	pACYC-PrhaBAD-LiAAT4-GmR	-
52	JW5020-1 (ΔfadE)	pACYC-PrhaBAD-NbBBT-GmR	-
53	JW5020-1 (ΔfadE)	pACYC-PrhaBAD-NtSHCT-GmR	-
54	JW5020-1 (ΔfadE)	pACYC-PrhaBAD-RsVS-GmR	-
55	JW5020-1 (ΔfadE)	pACYC-PrhaBAD-SsMMT-GmR	-
56	JW5020-1 (ΔfadE)	pET-Plac-Tes4-Sfp-CAR-KanR	pCDF-Plac-GFP-SpecR
57	JW5020-1 (ΔfadE)	pET-Plac-Tes4-Sfp-CAR-KanR	pCDF-Plac-Atf1-SpecR
58	JW5020-1 (ΔfadE)	pET-Plac-Tes4-Sfp-CAR-KanR	pCDF-Plac-CAT-SpecR

59	JW5020-1 (ΔfadE)	pET-Plac-Tes4-Sfp-CAR-KanR	pCDF-Plac-LiAAT4-SpecR
60	JW5020-1 (ΔfadE)	pET-Plac-Tes4-Sfp-CAR-KanR	pCDF-Plac-SAAT-SpecR
61	JW5020-1 (ΔfadE)	pET-Plac-Tes4-Sfp-CAR-KanR	pCDF-Plac-VAAT-SpecR
62	JW5020-1 (ΔfadE)	pET-Plac-Tes4-AtAAE7-SpecR	pCDF-Plac-GFP-SpecR
63	JW5020-1 (ΔfadE)	pET-Plac-Tes4-AtAAE7-SpecR	pCDF-Plac-Atf1-SpecR
60	JW5020-1 (ΔfadE)	pET-Plac-Tes4-AtAAE7-SpecR	pCDF-Plac-CAT-SpecR
61	JW5020-1 (ΔfadE)	pET-Plac-Tes4-AtAAE7-SpecR	pCDF-Plac-SAAT-SpecR
62	JW5020-1 (ΔfadE)	pET-Plac-Tes4-AtAAE7-SpecR	pCDF-Plac-VAAT-SpecR
63	JW5020-1 (ΔfadE)	pET-Plac-Tes4-AtAAE7-SpecR	pCDF-Plac-VpAAT1-SpecR
64	BW25113 WT	pET-Plac-Tes4-SpecR	pCDF-Plac-GFP-SpecR
65	BW25113 WT	pET-Plac-Tes4-AtAAE7-SpecR	pCDF-Plac-GFP-SpecR
66	BW25113 WT	pET-Plac-Tes4-AtAAE7-SpecR	pCDF-Plac-VAAT-SpecR
67	BW25113 WT	pET-Plac-Tes4-GaAAE7-SpecR	pCDF-Plac-GFP-SpecR
68	BW25113 WT	pET-Plac-Tes4-GaAAE7-SpecR	pCDF-Plac-VAAT-SpecR
69	BW25113 WT	pET-Plac-Tes4-TcAAE7-SpecR	pCDF-Plac-GFP-SpecR
70	BW25113 WT	pET-Plac-Tes4-TcAAE7-SpecR	pCDF-Plac-VAAT-SpecR
71	JW5020-1 (ΔfadE)	pET-Plac-Tes4-GaAAE7-SpecR	pCDF-Plac-VAAT-SpecR
72	JW5020-1 (ΔfadE)	pET-Plac-Tes4-TcAAE7-SpecR	pCDF-Plac-VAAT-SpecR
73	ΔyciA	pET-Plac-Tes4-AtAAE7-SpecR	pCDF-Plac-VAAT-SpecR
74	ΔyciA	pET-Plac-Tes4-GaAAE7-SpecR	pCDF-Plac-VAAT-SpecR
75	ΔyciA	pET-Plac-Tes4-TcAAE7-SpecR	pCDF-Plac-VAAT-SpecR
76	ΔyciA	pET-Plac-Tes4-AtAAE7-SpecR	pCDF-Plac-VAAT-PDC-SpecR
77	ΔyciA	pET-Plac-Tes4-AtAAE7-SpecR	pCDF-Plac-LiAAT4-SpecR
78	ΔyciA	pET-Plac-Tes4-AtAAE7-SpecR	pCDF-Plac-LiAAT4-PDC-SpecR
79	ΔyciA	pET-Plac-Tes4-GaAAE7-SpecR	pCDF-Plac-VAAT-PDC-SpecR
80	ΔyciA	pET-Plac-Tes4-GaAAE7-SpecR	pCDF-Plac-LiAAT4-SpecR
81	ΔyciA	pET-Plac-Tes4-GaAAE7-SpecR	pCDF-Plac-LiAAT4-PDC-SpecR
82	ΔyciA	pET-Plac-Tes4-TcAAE7-SpecR	pCDF-Plac-VAAT-PDC-SpecR
83	ΔyciA	pET-Plac-Tes4-TcAAE7-SpecR	pCDF-Plac-LiAAT4-SpecR
84	ΔyciA	pET-Plac-Tes4-TcAAE7-SpecR	pCDF-Plac-LiAAT4-PDC-SpecR

85	JW1794-1 (ΔfadD)	pET-Plac-GFP-KanR	-
86	JW1794-1 (ΔfadD)	pET-Plac-AtALT2-KanR	-
87	JW1794-1 (ΔfadD)	pET-Plac-MtALT2-KanR	-
88	BW25113 WT	pET-Plac-AtALT2-KanR	-
89	BW25113 WT	pET-Plac-AtALT2-YtpA-KanR	-
90	BW25113 WT	pET-Plac-YneP-KanR	-
91	BW25113 WT	pET-Plac-YneP-YtpA-KanR	-
92	JW1794-1 (ΔfadD)	pET-Plac-DfsA-Sfp-CAR-KanR	-
93	JW1794-1 (ΔfadD)	pET-Plac-DfsA F44W-Sfp-CAR-KanR	-
94	JW1794-1 (ΔfadD)	pET-Plac-DfsA F44Y-Sfp-CAR-KanR	-

Appendix B: *Synechocystis* Sp. PCC 6803 strains used for bioproduction experiments in this work

Strain No	*Synechocystis* strain	Plasmid
Chapter 4		
1	*Synechocystis* Δaas-Sfp-CAR	pRSF1010-PcoA-CvT3-EryR
2	*Synechocystis* Δaas-Sfp-CAR	pRSF1010-PcoA-CvT4-EryR
3	*Synechocystis* Δaas-Sfp-CAR	pRSF1010-PcoA-CvT5-EryR
4	*Synechocystis* Δaas-Sfp-CAR	pRSF1010-PcoA-CvT6-EryR
5	*Synechocystis* Δaas-Sfp-CAR	pRSF1010-PcoA-CvT7-EryR
6	*Synechocystis* Δaas-Sfp-CAR	pRSF1010-PcoA-CvT8-EryR
Chapter 6		
1	*Synechocystis* Δaas	pRSF1010-PcoA-CvT5-EryR
2	*Synechocystis* Δaas	pRSF1010-PcoA-CvT5-BM3 WT-EryR
3	*Synechocystis* Δaas	pRSF1010-PcoA-CvT5-BM3 F87G L188Q-EryR
4	*Synechocystis* Δaas	pRSF1010-PcoA-CvT5-BM3 V78A F87A L188Q A328V-EryR

Appendix C: Primers

Primers for genomic DNA amplification		
Primer	**Amplifying**	**Sequence**
JM125_EcACP_ampF	*Escherichia coli acpP*	TCTGGTGGGTCTCTGTCCATGAGCACTATC GAAGAACGCG
JM126_EcACP_ampR	*Escherichia coli acpP*	CGATAGGTCTCCCGAGCCTTACGCCTGGTG GCCGTTGATG
JM127_6803ACP_ampF	*Synechocystis* Sp. PCC 6803 *acp*	TCTGGTGGGTCTCTGTCCATGAATCAGGAA ATTTTTGAAA
JM128_6803ACP_ampR	*Synechocystis* Sp. PCC 6803 *acp*	CGATAGGTCTCCCGAGCCTTATTTACTTTC GATATGCTCA
JM166_Pp0763_AmpF	*Pseudomonas putida Pp0763*	TCTGGTGGGTCTCTGTCCATGTTGCAGACA CGCATCATCA
JM169_Pp0763_AmpR	*Pseudomonas putida Pp0763*	CGATAGGTCTCCCGAGCCTTACAACGTGGA AAGGAACGCG
Primers for truncating FatB thioesterases		
Primer	**Truncating**	**Sequence**
JM72_CvFatB1ampR	CvFatB1	CGATAGGTCTCCCGAGCCTTAACTAACACT GTTCCCATTG
JM73_CvFatB1_1ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGGCCAACGGT AGTGCAGTGA
JM74_CvFatB1_2ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGTTGCCCGAC TGGTCCATGC
JM92_CvFatB1_T3ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGCTACTAACC GCTATCACAA
JM93_CvFatB1_T4ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGGAAAAACAG TGGACCATGC
JM75_CvFatB1_3ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGCTCGACCGG AAATCCAAGC
JM94_CvFatB1_T6ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGCGACCTGAC ATGCTAGTAG
JM95_CvFatB1_T7ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGAAATCCATC GTGCGTGATG
JM111_CvFatB1_T8ampf	CvFatB1	TCTGGTGGGTCTCTGTCCATGCGCCACAGC TTCAGCATTCGAT
JM112_CvFatB1_T9ampf	CvFatB1	TCTGGTGGGTCTCTGTCCATGATCGGAGCT GACCGGACCG
JM113_CvFatB1_T10ampf	CvFatB1	TCTGGTGGGTCTCTGTCCATGACACTGATG AACCATTTACAAG
JM114_CvFatB1_T11ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGATCAATCAC TGTAAGAGCTTGG
JM115_CvFatB1_T12ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGAATGACGGA TTCGGCCGCACAC
JM116_CvFatB1_T13ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGTGTAAGAAT GACCTCATCT
JM117_CvFatB1_T14ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGAAGATGCAA ATTATGGTTAATC
JM118_CvFatB1_T15ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGACCTGGGGC GATACTGTAGAGA
JM119_CvFatB1_T16ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGTGGTTTAGT CAAAGTGGCAAGA
JM120_CvFatB1_T17ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGGCGTCCGAC TGGCTAATCTCTG
JM121_CvFatB1_T18ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGACCGGGGAA ATCCTGATCCGAG

JM122_CvFatB1_T19ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGGTGTGGGCCATGATGAACCAGA
JM123_CvFatB1_T20ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGCGCTTCTCTCGGCTCCCGTATG
JM124_CvFatB1_T21ampF	CvFatB1	TCTGGTGGGTCTCTGTCCATGCAAGAACTCACCCCGCACTTCG
JM133_UaFatB1ampR	UaFatB1	GATAGGTCTCCCGAGCCTTAGGAAGTCTTGGCAGGAACGT
JM134_UaFatB1_T1ampF	UaFatB1	TCTGGTGGGTCTCTGTCCATGCTAAATGGCTCTAATGTTG
JM135_UaFatB1_T2ampF	UaFatB1	TCTGGTGGGTCTCTGTCCATGCTACCCGATTGGTCTATGC
JM136_UaFatB1_T3ampF	UaFatB1	TCTGGTGGGTCTCTGTCCATGCTGCTCGCCGCTATTACGA
JM137_UaFatB1_T4ampF	UaFatB1	TCTGGTGGGTCTCTGTCCATGGAAAAGCAGTGGATGATGT
JM138_UaFatB1_T5ampF	UaFatB1	TCTGGTGGGTCTCTGTCCATGTTGGACTGGAAGCCAAAGC
JM139_UaFatB1_T6ampF	UaFatB1	TCTGGTGGGTCTCTGTCCATGCGACCCGATATGCTGGTCG
JM140_UaFatB1_T7ampF	UaFatB1	TCTGGTGGGTCTCTGTCCATGGGGCGATTTGTTCAGGACG
JM141_UaFatB1_T8ampF	UaFatB1	TCTGGTGGGTCTCTGTCCATGCGTAACAACTTTTCTATCC
JM142_UaFatB1_T9ampF	UaFatB1	TCTGGTGGGTCTCTGTCCATGATCGGAGCGGATCGGACTG
Primers for hybridising CYP450s		
Primer	**Role**	**Sequence**
JM191_RvHaem_Fw	Domain amplification + Final amplification	TCTGGTGGGTCTCTGTCCATG
JM192_RvHaem_Rv	Domain amplification	CTTGGGCTAGCATCGCGGCGTTCTGGATTCTGAGC
JM193_RvXBm3Reductase_Fw	Domain amplification	GCGATGCTAGCCCAAGCACTGAGCAGTCCGCGAAGA
JM194_TtHaem_Fw	Domain amplification + Final amplification	TCTGGTGGGTCTCTGTCCATG
JM195_TtHaem_Rv	Domain amplification	CTTGGGCTGGGGTCGCGACGCTCGGGATTTTGGGCT
JM196_TtxBM3Reductase_Fw	Domain amplification	GCGACCCCAGCCCAAGCACTGAGCAGTCCGCGAAGAA
JM197_BM3Reductase_Rv	Domain amplification + Final amplification	CGATAGGTCTCCCGAGCCTTA
Primers for site directed mutagenesis		
Primer	**Mutating**	**Sequence**
JM153_M90G	CvT5	TGACCTCATCTGGGTATTGACTAAGGGCCAAATTATGGTTAATCGCTATCCTA
JM151_W122G	CvT5	TGGCAAGATCGGTATGGCGTCCGACGGCCTAATCTCTGATTGCAACACCGGGG
JM152_W140G	CvT5	AATCCTGATCCGAGCAACTTCTGTGGGCGCCATGATGAACCAGAAGACCCGGC
JM161_W122V	CvT5	TGGCAAGATCGGTATGGCGTCCGACGTGCTAATCTCTGATTGCAACACCGGGG
JM162_W122M	CvT5	TGGCAAGATCGGTATGGCGTCCGACATGCTAATCTCTGATTGCAACACCGGGG

JM163_W122F	CvT5	TGGCAAGATCGGTATGGCGTCCGACTTTCT AATCTCTGATTGCAACACCGGGG
JM01_F87A	BM3	TCGTGATTTTGCGGGTGACGGTTTAGCC ACCAGTTGGACACATGAGAAAAATT
JM02_V78A	BM3	AAACTTGTCCCAGGCTCTCAAGTTTGCC CGTGATTTTGCGGGTGACGGTTTAG
JM176_F88G	BM3	TCGTGATTTTGCGGGTGACGGTTTAGGCAC CAGTTGGACACATGAGAAAAATT
JM177_F88V	BM3	TCGTGATTTTGCGGGTGACGGTTTAGTGAC CAGTTGGACACATGAGAAAAATT
JM178_F88I	BM3	TCGTGATTTTGCGGGTGACGGTTTAATTAC CAGTTGGACACATGAGAAAAATT
JM88_L188Q	BM3	TGCTCTCGACGAGGCCATGAACAAACAACA GCGTGCTAACCCAGACGACCCGG
JM91_I401P	BM3	TTTCGGCAACGGGCAACGTGCCTGCCCCGG ACAGCAATTTGCTTTACACGAAG
JM189_D182S_L188Q	BM3 L188Q	TATTACTAGCATGGTGCGTGCTCTCTCCGA GGCCATGAACAAACAACAGCGTG
JM190_H266T	BM3	AATCATCACGTTTTTGATTGCCGGGACCGA GACGACGAGTGGTTTGTTGAGTT
JM89_A328V	BM3	CGAAGCATTGCGGCTATGGCCAACGGTGCC GGCTTTTAGTTTGTATGCGAAGG
JM200_F44W	DfsA	GTTGCGGTCTGAGCCACGTCCATGTTGGAA TCAGCAATTAGTTACTGACATTA
JM201_F44Y	DfsA	GTTGCGGTCTGAGCCACGTCCATGTTATAA TCAGCAATTAGTTACTGACATTA
JM202_G135C	DfsA	CATGGTCGAGGGGAGCGCGCTGGGTTGTGG CTTTGAGGCGGCGTTGGCTCACC

Sequencing primers

Primer	Used to sequence	Sequence
JM34_pJETF	pJET backbones	AATACGACTCACTATAGGGAGAGC
JM35_pJETR	pJET backbones	AAGAACATCGATTTTCCATGGCAG
JM12_pUCIDT_F	pUCIDT backbones	GTAAAACGACGGCCAGTGCAAC
JM39_PrhaBADF	Backbones with PrhaBAD and Plac	TATCACGAGGCCCTTTCGTCTTC
JM158_PcoAF	Backbones with PcoA	CAATCCCAGGGTGTGGTAATGG
JM40_ColE1R	Backbones with ColE1	GATCTTTTCTACGGGGTCTGAC
JM41_CloDFR	Backbones with CloD13	TGCTTATCCACGGAATCTGTGG
JM43_p15aR	Backbones with p15a	TTGATCGGCACGTAAGAGGTTC
JM160_OriVRnew	Backbones with OriV	TACTCAGGAGAGCGTTCACC
JM185_KanR_R	Backbones with KanR	ACCTGATTGCCCGACATTATCG
JM186_SpecR_R	Backbones with SpecR	ATCAGCCCGTCATACTTGAAGC

JM198_BM3reductase_Rvseq	Hybridisation of CYP450 domains	TAGCATTATCTGGCGGGTGACC
Primers used for the knock-out of *iscR*		
Primer	**Used for**	**Sequence**
JM179_iscR_locus_F	Amplification of Frt-ChlR-Frt and incorporation of homology arms	TCCGTCCATCGTCATAAACTGC
JM180_iscR_locus_R	Amplification of Frt-ChlR-Frt and incorporation of homology arms	GACAACAGGTACAAACGCCACG
JM187_iscR_KO_F	Verifying integration of Frt-ChlR-Frt into the genome and subsequent removal of ChlR	CCGTGTTTACGGAGTATTTAGCACTCCGGC CTGATTCTGAATTCTTTTTACGATTGTGTA GGCTGGAGCTGCTTC
JM188_iscR_KO_R	Verifying integration of Frt-ChlR-Frt into the genome and subsequent removal of ChlR	TTTTACAATAAAAAACCCCGGGCAGGGGCG AGTTTGAGGTGAAGTAAGACGAATATCCTC CTTAGTTCCTATTCC

Appendix D: Preparation of partially double-stranded linkers for BASIC assembly

Final Linker	Adapter Oligonucleotide		Linker Oligonucleotide	
	Name	**Sequence**	**Name**	**Sequence**
L1-P	L1-P-A	TTTATTGAACTA	L1-P-L	GGACTAGTTCAATAAATACC CTCTGACTGTCTCGGAG
L1-S	L1-S-A	TGTCGTAAGTAA	L1-S-L	CTCGTTACTTACGACACTCC GAGACAGTCAGAGGGTA
U1-S	U1-S-A	GACGGTGTTCAA	U1-S-L	CTCGTTGAACACCGTCTCAG GTAAGTATCAGTTGTAA
U1-RBS3-P	U1-RBS3-P-A	AAAGAGGAGAAATA	U1-RBS3-P-L	GGACTATTTCTCCTCTTTTT ACAACTGATACTTACCTGA
U1-RBS5-P	U1-RBS5-P	ATTAGTGGAGGTTA	U1-RBS5-P-L	GGACTAACCTCCACTAATTT ACAACTGATACTTACCTGA
U2-S	U2-S-A	CCAATAGTAACA	U2-S-L	CTCGTGTTACTATTGGCTGA GATAAGGGTAGCAGAAA
U2-RBS3-P	U2-RBS3-P-A	AAAGAGGAGAAATA	U2-RBS3-P-L	GGACTATTTCTCCTCTTTTT TCTGCTACCCTTATCTCAG
U2-RBS5-P	U2-RBS5-P-A	ATTAGTGGAGGTTA	U2-RBS5-P-L	GGACTAACCTCCACTAATTT TCTGCTACCCTTATCTCAG

Appendix E: Partially double-stranded linker sequences

Linker	Sequence (ds region in red)
L1-S	CTCGTTACTTACGACACTCCGAGACAGTCAGAGGGTA
L1-P	GGACTAGTTCAATAAATACCCTCTGACTGTCTCGGAG
L2-S	CTCGATCGGTGTGAAAAGTCAGTATCCAGTCGTGTAG
L2-P	GGACAGGTAATAAGAACTACACGACTGGATACTGACT
L3-S	CTCGATCACGGCACTACACTCGTTGCTTTATCGGTAT
L3-P	GGACTCTGTAATAACAATACCGATAAAGCAACGAGTG
U1-XS	CTCGTTGAACACCGTCTCAGGTAAGTATCAGTTGTAA
U1-RBS1-P	GGACTAGTCCTGTGTGATTTACAACTGATACTTACCTGA
U1-RBS2-P	GGACTATTTCCCCTCTTTTTACAACTGATACTTACCTGA
U1-RBS3-P	GGACTATTTCTCCTCTTTTTACAACTGATACTTACCTGA
U1-RBS5-P	GGACTAACCTCCACTAATTTACAACTGATACTTACCTGA
U1-RBSA01-P	GGACTACCTCCTTGGATTTACAACTGATACTTACCTGA
U1-RBSA03-P	GGACTACCTCCTGGGATTTACAACTGATACTTACCTGA
U1-RBSA07-P	GGACTACCTCCTTAGATTTACAACTGATACTTACCTGA
U1-RBSA08-P	GGACTACCTCCCTAGATTTACAACTGATACTTACCTGA
U1-RBSA09-P	GGACTACCTCCTGAGATTTACAACTGATACTTACCTGA
U2-XS	CTCGTGTTACTATTGGCTGAGATAAGGGTAGCAGAAA
U2-RBS1-P	GGACTAGTCCTGTGTGATTTTCTGCTACCCTTATCTCAG
U2-RBS2-P	GGACTATTTCCCCTCTTTTTTCTGCTACCCTTATCTCAG
U2-RBS3-P	GGACTATTTCTCCTCTTTTTTCTGCTACCCTTATCTCAG
U2-RBS5-P	GGACTAACCTCCACTAATTTTCTGCTACCCTTATCTCAG
U2-RBSA01-P	GGACTACCTCCTTGGATTTTCTGCTACCCTTATCTCAG
U2-RBSA03-P	GGACTACCTCCTGGGATTTTCTGCTACCCTTATCTCAG
U2-RBSA07-P	GGACTACCTCCTTAGATTTTCTGCTACCCTTATCTCAG
U2-RBSA08-P	GGACTACCTCCCTAGATTTTCTGCTACCCTTATCTCAG
U2-RBSA09-P	GGACTACCTCCTGAGATTTTCTGCTACCCTTATCTCAG
U3-XS	CTCGGTATCTCGTGGTCTGACGGTAAAATCTATTGTA
U3-RBS1-P	GGACTAGTCCTGTGTGATTACAATAGATTTTACCGTCAG
U3-RBS2-P	GGACTATTTCCCCTCTTTTACAATAGATTTTACCGTCAG
U3-RBS3-P	GGACTATTTCTCCTCTTTTACAATAGATTTTACCGTCAG
U3-RBSA01-P	GGACTACCTCCTTGGATTACAATAGATTTTACCGTCAG
U3-RBSA03-P	GGACTACCTCCTGGGATTACAATAGATTTTACCGTCAG
U3-RBSA07-P	GGACTACCTCCTTAGATTACAATAGATTTTACCGTCAG
U3-RBSA08-P	GGACTACCTCCCTAGATTACAATAGATTTTACCGTCAG
U3-RBSA09-P	GGACTACCTCCTGAGATTACAATAGATTTTACCGTCAG

Appendix F: List of expression vectors used in this work

Plasmid	Sequence
pACYC-Plac-GFP-GmR	GCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAA AAATAGGCGTATCACGAGGCCCTTTCGTCTTCACCTCGAGAAATTTAT CAAAAAGAGTGTTGACTTGTGAGCGGATAACAATGATACTTAGATTCA ATTGTGAGCGGATAACAATTGGCTCGGGAGACCTATCGGTAATAACAG TCCAATCTGGTGTAACTTCGGAATCGTCCTTTACAGCTAGCTCAGTCC TAGGTATTATGCTAGCTACTAGAGTCACACAGGAAAGTACTAGATGCG TAAAGGCGAAGAACTGTTCACGGGCGTAGTTCCGATTCTGGTCGAGCT GGACGGCGATGTGAACGGTCATAAGTTTAGCGTTCGCGGTGAAGGTGA GGGCGACGCGACCAACGGCAAACTGACCCTGAAGTTCATCTGCACCAC CGGTAAACTGCCGGTGCCTTGGCCGACCTTGGTGACGACGTTGACGTA TGGCGTGCAGTGTTTTGCGCGTTATCCGGACCACATGAAACAACACGA TTTCTTCAAATCTGCGATGCCGGAGGGTTACGTCCAGGAGCGTACCAT TTCCTTCAAGGATGATGGCACTTACAAAACTCGCGCAGAGGTTAAGTT TGAAGGTGACACGCTGGTCAATCGTATCGAATTGAAGGGTATCGACTT TAAAGAGGATGGTAACATTCTGGGCCATAAACTGGAGTATAACTTCAA CAGCCATAATGTTTACATTACGGCAGACAAGCAAAAGAACGGCATCAA GGCCAATTTCAAGATTCGCCACAATGTTGAGGACGGTAGCGTCCAACT GGCCGACCATTACCAGCAGAACACCCCAATTGGTGACGGTCCGGTTTT GCTGCCGGATAATCACTATCTGAGCACCCAAAGCGTGCTGAGCAAAGA TCCGAACGAAAAACGTGATCACATGGTCCTGCTGGAATTTGTGACCGC TGCGGGCATCACCCACGGTATGGACGAGCTGTATAAGCGTCCGTAATA ATACTAGAGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACT GGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGA GTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAGGGGTA AGAACTCGCACTTCGTGGAAACACTATTATCTGGTGGGTCTCTGTCCC CAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCG TTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGG CTCACCTTCGGGTGGGCCTTTCTGCGTTTATAGGAGAAGTAGTGCCAC AGACAGTATTGCTTACGAGTTGATTTATCCTGTCCTTTAGCTTCCTTA GCTCCTGAAAATCTCGATAACTCAAAAAATACGCCCGGTAGTGATCTT ATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCT CATTTTCGCCAAAAGTTGGCCCAGGGCTTCCCGGTATCAACAGGGACA CCAGGATTTATTTATTCTGCGAAGTGATCTTCCGTCACAGGTATTTAT TCGGCGCAAAGTGCGTCGGGTGATGCTGCCAACTTACTGATTTAGTGT ATGATGGTGTTTTTGAGGTGCTCCAGTGGCTTCTGTTTCTATCAGCTG TCCCTCCTGTTCAGCTACTGACGGGGTGGTGCGTAACGGCAAAAGCAC CGCCGGACATCAGCGCTAGCGGAGTGTATACTGGCTTACTATGTTGGC ACTGATGAGGGTGTCAGTGAAGTGCTTCATGTGGCAGGAGAAAAAAGG CTGCACCGGTGCGTCAGCAGAATATGTGATACAGGATATATTCCGCTT CCTCGCTCACTGACTCGCTACGCTCGGTCGTTCGACTGCGGCGAGCGG AAATGGCTTACGAACGGGGCGGAGATTTCCTGGAAGATGCCAGGAAGA TACTTAACAGGGAAGTGAGAGGGCCGCGGCAAAGCCGTTTTTCCATAG GCTCCGCCCCCCTGACAAGCATCACGAAATCTGACGCTCAAATCAGTG GTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGG CGGCTCCCTCGTGCGCTCTCCTGTTCCTGCCTTTCGGTTTACCGGTGT CATTCCGCTGTTATGGCCGCGTTTGTCTCATTCCACGCCTGACACTCA GTTCCGGGTAGGCAGTTCGCTCCAAGCTGGACTGTATGCACGAACCCC CCGTTCAGTCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGT CCAACCCGGAAAGACATGCAAAAGCACCACTGGCAGCAGCCACTGGTA ATTGATTTAGAGGAGTTAGTCTTGAAGTCATGCGCCGGTTAAGGCTAA ACTGAAAGGACAAGTTTTGGTGACTGCGCTCCTCCAAGCCAGTTACCT CGGTTCAAAGAGTTGGTAGCTCAAAAAACCTTCGAAAAACCGCCCTGC AAGGCGGTTTTTTCGTTTTCAGAGCAAGAGATTACGCGCAGACCAAAA CGATCTCAAGAAGATCATCTTATTAATCAGATAAAATATTTCTAGATT TCAGTGCAATTTATCTCTTCAAATGTAGCACCTGAAGTCAGCCCCATA CGATATAAGTTGTAATTCTCATGTTAGTCATGCCCCGCGCCCACCGGA AGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCG

	GTGCCTAATGAGTGAGCTAACTTACATTAATTGCGTTGCGCTCAGGAT CGGTGTGAAAAGTCAGTATCCAGTCGTGTAGTTCTTATTACCTGTCCA CTAGTCTTGGACTCCTGTTGATAGATCCAGTAATGACCTCAGAACTCC ATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTATTG GTGAGAATCCAGGGGTCCCCAATAATTACGATTTAAATTTGACATAAG CCTGTTCGGTTCGTAAACTGTAATGCAAGTAGCGTATGCGCTCACGCA ACTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTGG CGGTTTTCATGGCTTGTTATGACTGTTTTTTTGTACAGCCTATGCCTC GGGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTCGATGTTTGATGT TATGGAGCAGCAACGATGTTACGCAGCAGCAACGATGTTACGCAGCAG GGCAGTCGCCCTAAAACAAAGTTAGGTGGCTCAAGTATGGGCATCATT CGCACATGTAGGCTCGGCCCTGACCAAGTCAAATCCATGCGGGCTGCT CTTGATCTTTTCGGTCGTGAGTTCGGAGACGTAGCCACCTACTCCCAA CATCAGCCGGACTCCGATTACCTCGGGAACTTGCTCCGTAGTAAGACA TTCATCGCGCTTGCTGCCTTCGACCAAGAAGCGGTTGTTGGCGCTCTC GCGGCTTACGTTCTGCCCAAGTTTGAGCAGCCGCGTAGTGAGATCTAT ATCTATGATCTCGCAGTCTCCGGAGAGCACCGGAGGCAGGGCATTGCC ACCGCGCTCATCAATCTCCTCAAGCATGAGGCCAACGCGCTTGGTGCT TATGTGATCTACGTGCAAGCAGATTACGGTGACGATCCCGCAGTGGCT CTCTATACAAAGTTGGGCATACGGGAAGAAGTGATGCACTTTGATATC GACCCAAGTACCGCCACCTAACAATTCGTTCAAGCCGAGATCGGCTTC CCGGCCGCGGAGTTGTTCGGTAAATTGGACAACGGTCGGATCACGGCA CTACACTCGTTGCTTTATCGGTATTGTTATTACAGAGTCCGACACCAT CGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGA GAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGT CGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAA CCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGC GATGGCGGAGCTGAATTACATTCCCAACCGCGTGGCACAACAACTGGC GGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCT GCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCA ACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAACGAAGCGGCGTCGA AGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGG GCTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGA AGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCA GACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACT GGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTT AGCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTG GCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGA AGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCT GAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGAT GGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGG TGCGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATG TTATATCCCGCCGTTAACCACCATCAAACAGGATTTTCGCCTGCTGGG GCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGT GAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCAC CCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTC ATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTG AGCGCAACGCAATTAATACTAGTGAGCGGATACATATTTGAATGTATT TAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGT
pACYC-PrhaBAD-GFP-EryR	TCGCTGTGGATTCAGCGACAGTAACCACTTTTCGACGCTTTTTCGCCG AGAGTTTAACTGGTCACCGCGTGATATTCGCCAGGGACGGGATGGCTT TCTGCAATAAGCGCAACGCAATTAATACTAGTGAGCGGATACATATTT GAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCC CGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTA ACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCACCTCGAG AAATTTATCAAAAAGAGTGCCACAATTCAGCAAATTGTGAACATCATC ACGTTCATCTTTCCCTGGTTGCCAATGGCCCATTTTCCTGTCAGTAAC GAGAAGGTCGCGTATTCAGGCGCTTTTTAGACTGGTCGTAATGAAGGG GAGACCTATCGGTAATAACAGTCCAATCTGGTGTAACTTCGGAATCGT CCTTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCTACTAGAGTCA CACAGGAAAGTACTAGATGCGTAAAGGCGAAGAACTGTTCACGGGCGT

	AGTTCCGATTCTGGTCGAGCTGGACGGCGATGTGAACGGTCATAAGTT TAGCGTTCGCGGTGAAGGTGAGGGCGACGCGACCAACGGCAAACTGAC CCTGAAGTTCATCTGCACCACCGGTAAACTGCCGGTGCCTTGGCCGAC CTTGGTGACGACGTTGACGTATGGCGTGCAGTGTTTTGCGCGTTATCC GGACCACATGAAACAACACGATTTCTTCAAATCTGCGATGCCGGAGGG TTACGTCCAGGAGCGTACCATTTCCTTCAAGGATGATGGCACTTACAA AACTCGCGCAGAGGTTAAGTTTGAAGGTGACACGCTGGTCAATCGTAT CGAATTGAAGGGTATCGACTTTAAAGAGGATGGTAACATTCTGGGCCA TAAACTGGAGTATAACTTCAACAGCCATAATGTTTACATTACGGCAGA CAAGCAAAAGAACGGCATCAAGGCCAATTTCAAGATTCGCCACAATGT TGAGGACGGTAGCGTCCAACTGGCCGACCATTACCAGCAGAACACCCC AATTGGTGACGGTCCGGTTTTGCTGCCGGATAATCACTATCTGAGCAC CCAAAGCGTGCTGAGCAAAGATCCGAACGAAAAACGTGATCACATGGT CCTGCTGGAATTTGTGACCGCTGCGGGCATCACCCACGGTATGGACGA GCTGTATAAGCGTCCGTAATAATACTAGAGCCAGGCATCAAATAAAAC GAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGT CGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGC CTTTCTGCGTTTATAGGGGTAAGAACTCGCACTTCGTGGAAACACTAT TATCTGGTGGGTCTCTGTCCCCAGGCATCAAATAAAACGAAAGGCTCA GTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGC TCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGT TTATAGGAGAAGTAGTGCCACAGACAGTATTGCTTACGAGTTGATTTA TCCTGTCCTTTAGCTTCCTTAGCTCCTGAAAATCTCGATAACTCAAAA AATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCT CTTACGTGCCGATCAACGTCTCATTTTCGCCAAAAGTTGGCCCAGGGC TTCCCGGTATCAACAGGGACACCAGGATTTATTTATTCTGCGAAGTGA TCTTCCGTCACAGGTATTTATTCGGCGCAAAGTGCGTCGGGTGATGCT GCCAACTTACTGATTTAGTGTATGATGGTGTTTTTGAGGTGCTCCAGT GGCTTCTGTTTCTATCAGCTGTCCCTCCTGTTCAGCTACTGACGGGGT GGTGCGTAACGGCAAAAGCACCGCCGGACATCAGCGCTAGCGGAGTGT ATACTGGCTTACTATGTTGGCACTGATGAGGGTGTCAGTGAAGTGCTT CATGTGGCAGGAGAAAAAAGGCTGCACCGGTGCGTCAGCAGAATATGT GATACAGGATATATTCCGCTTCCTCGCTCACTGACTCGCTACGCTCGG TCGTTCGACTGCGGCGAGCGGAAATGGCTTACGAACGGGGCGGAGATT TCCTGGAAGATGCCAGGAAGATACTTAACAGGGAAGTGAGAGGGCCGC GGCAAAGCCGTTTTTCCATAGGCTCCGCCCCCCTGACAAGCATCACGA AATCTGACGCTCAAATCAGTGGTGGCGAAACCCGACAGGACTATAAAG ATACCAGGCGTTTCCCCCTGGCGGCTCCCTCGTGCGCTCTCCTGTTCC TGCCTTTCGGTTTACCGGTGTCATTCCGCTGTTATGGCCGCGTTTGTC TCATTCCACGCCTGACACTCAGTTCCGGGTAGGCAGTTCGCTCCAAGC TGGACTGTATGCACGAACCCCCCGTTCAGTCCGACCGCTGCGCCTTAT CCGGTAACTATCGTCTTGAGTCCAACCCGGAAAGACATGCAAAAGCAC CACTGGCAGCAGCCACTGGTAATTGATTTAGAGGAGTTAGTCTTGAAG TCATGCGCCGGTTAAGGCTAAACTGAAAGGACAAGTTTTGGTGACTGC GCTCCTCCAAGCCAGTTACCTCGGTTCAAAGAGTTGGTAGCTCAAAAA ACCTTCGAAAAACCGCCCTGCAAGGCGGTTTTTTCGTTTTCAGAGCAA GAGATTACGCGCAGACCAAAACGATCTCAAGAAGATCATCTTATTAAT CAGATAAAATATTTCTAGATTTCAGTGCAATTTATCTCTTCAAATGTA GCACCTGAAGTCAGCCCCATACGATATAAGTTGTAATTCTCATGTTAG TCATGCCCCGCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCA AGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTGAGCTAACTTACAT TAATTGCGTTGCGCTCAGGATCGGTGTGAAAAGTCAGTATCCAGTCGT GTAGTTCTTATTACCTGTCCACTAGTATTCGATGAATTCGCGTGCTAT AATTATACTAATTTTATAAGGAGGAAAAAATATGGGCATTTTTAGTAT TTTTGTAATCAGCACAGTTCATTATCAACCAAACAAAAAATAAGTGGT TATAATGAATCGTTAATAAGCAAAATTCATATAACCAAATTAAAGAGG GTTATAATGAACGAGAAAAATATAAAACACAGTCAAAACTTTATTACT TCAAAACATAATATAGATAAAATAATGACAAATATAAGATTAAATGAA CATGATAATATCTTTGAAATCGGCTCAGGAAAAGGCCATTTTACCCTT GAATTAGTAAAGAGGTGTAATTTCGTAACTGCCATTGAAATAGACCAT AAATTATGCAAAACTACAGAAAATAAACTTGTTGATCACGATAATTTC

	CAAGTTTTAAACAAGGATATATTGCAGTTTAAATTTCCTAAAAACCAA TCCTATAAAATATATGGTAATATACCTTATAACATAAGTACGGATATA ATACGCAAAATTGTTTTTGATAGTATAGCTAATGAGATTTATTTAATC GTGGAATACGGGTTTGCTAAAAGATTATTAAATACAAAACGCTCATTG GCATTACTTTTAATGGCAGAAGTTGATATTTCTATATTAAGTATGGTT CCAAGAGAATATTTTCATCCTAAACCTAAAGTGAATAGCTCACTTATC AGATTAAGTAGAAAAAAATCAAGAATATCACACAAAGATAAACAAAAG TATAATTATTTCGTTATGAAATGGGTTAACAAAGAATACAAGAAAATA TTTACAAAAAATCAATTTAACAATTCCTTAAAACATGCAGGAATTGAC GATTTAAACAATATTAGCTTTGAACAATTCTTATCTCTTTTCAATAGC TATAAATTATTTAATAAGTAACTGCAGGGATCACGGCACTACACTCGT TGCTTTATCGGTATTGTTATTACAGAGTCCGACACCATCGAATGGTGC AAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTC AGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTAT GCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAGC CACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGACCGTA TTACATAGTGTGGATTTTTTTCCGTCTGGTAACGCGTCCGTGGCGATA GAACCCCGGCTCCCGCAGGCGGATTTTCCTGAACATCATCATGATTTT CATGAAATTGTGATTGTCGAACATGGCACGGGTATTCATGTGTTTAAT GGGCAGCCCTATACCATCACCGGTGGCACGGTCTGTTTCGTACGCGAT CATGATCGGCATCTGTATGAACATACCGATAATCTGTGTCTGACCAAT GTGCTGTATCGCTCGCCGGATCGATTTCAGTTTCTCGCCGGGCTGAAT CAGTTGCTGCCACAAGAGCTGGATGGGCAGTATCCGTCTCACTGGCGC GTTAACCACAGCGTATTGCAGCAGGTGCGACAGCTGGTTGCACAGATG GAACAGCAGGAAGGGGAAAATGATTTACCCTCGACCGCCAGTCGCGAG ATCTTGTTTATGCAATTACTGCTCTTGCTGCGTAAAAGCAGTTTGCAG GAGAACCTGGAAAACAGCGCATCACGTCTCAACTTGCTTCTGGCCTGG CTGGAGGACCATTTTGCCGATGAGGTGAATTGGGATGCCGTGGCGGAT CAATTTTCTCTTTCACTGCGTACGCTACATCGGCAGCTTAAGCAGCAA ACGGGACTGACGCCTCAGCGATACCTGAACCGCCTGCGACTGATGAAA GCCCGACATCTGCTACGCCACAGCGAGGCCAGCGTTACTGACATCGCC TA
pACYC-PrhaBAD-GFP-GmR	ATCGCTGTGGATTCAGCGACAGTAACCACTTTTCGACGCTTTTTCGCC GAGAGTTTAACTGGTCACCGCGTGATATTCGCCAGGGACGGGATGGCT TTCTGCAATAAGCGCAACGCAATTAATACTAGTGAGCGGATACATATT TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCC CCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATT AACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCACCTCGA GAAATTTATCAAAAAGAGTGCCACAATTCAGCAAATTGTGAACATCAT CACGTTCATCTTTCCCTGGTTGCCAATGGCCCATTTTCCTGTCAGTAA CGAGAAGGTCGCGTATTCAGGCGCTTTTTAGACTGGTCGTAATGAAGG CTCGGGAGACCTATCGGTAATAACAGTCCAATCTGGTGTAACTTCGGA ATCGTCCTTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCTACTAG AGTCACACAGGAAAGTACTAGATGCGTAAAGGCGAAGAACTGTTCACG GGCGTAGTTCCGATTCTGGTCGAGCTGGACGGCGATGTGAACGGTCAT AAGTTTAGCGTTCGCGGTGAAGGTGAGGGCGACGCGACCAACGGCAAA CTGACCCTGAAGTTCATCTGCACCACCGGTAAACTGCCGGTGCCTTGG CCGACCTTGGTGACGACGTTGACGTATGGCGTGCAGTGTTTTGCGCGT TATCCGGACCACATGAAACAACACGATTTCTTCAAATCTGCGATGCCG GAGGGTTACGTCCAGGAGCGTACCATTTCCTTCAAGGATGATGGCACT TACAAAACTCGCGCAGAGGTTAAGTTTGAAGGTGACACGCTGGTCAAT CGTATCGAATTGAAGGGTATCGACTTTAAAGAGGATGGTAACATTCTG GGCCATAAACTGGAGTATAACTTCAACAGCCATAATGTTTACATTACG GCAGACAAGCAAAAGAACGGCATCAAGGCCAATTTCAAGATTCGCCAC AATGTTGAGGACGGTAGCGTCCAACTGGCCGACCATTACCAGCAGAAC ACCCCAATTGGTGACGGTCCGGTTTTGCTGCCGGATAATCACTATCTG AGCACCCAAAGCGTGCTGAGCAAAGATCCGAACGAAAAACGTGATCAC ATGGTCCTGCTGGAATTTGTGACCGCTGCGGGCATCACCCACGGTATG GACGAGCTGTATAAGCGTCCGTAATAATACTAGAGCCAGGCATCAAAT AAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTG TTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGG

	TGGGCCTTTCTGCGTTTATAGGCTCGGGTAAGAACTCGCACTTCGTGG AAACACTATTATCTGGTGGGTCTCTGTCCCCAGGCATCAAATAAAACG AAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTC GGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCC TTTCTGCGTTTATAGGCTCGAGAAGTAGTGCCACAGACAGTATTGCTT ACGAGTTGATTTATCCTGTCCTTTAGCTTCCTTAGCTCCTGAAAATCT CGATAACTCAAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTG AAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATTTTCGCCAAAA GTTGGCCCAGGGCTTCCCGGTATCAACAGGGACACCAGGATTTATTTA TTCTGCGAAGTGATCTTCCGTCACAGGTATTTATTCGGCGCAAAGTGC GTCGGGTGATGCTGCCAACTTACTGATTTAGTGTATGATGGTGTTTTT GAGGTGCTCCAGTGGCTTCTGTTTCTATCAGCTGTCCCTCCTGTTCAG CTACTGACGGGGTGGTGCGTAACGGCAAAAGCACCGCCGGACATCAGC GCTAGCGGAGTGTATACTGGCTTACTATGTTGGCACTGATGAGGGTGT CAGTGAAGTGCTTCATGTGGCAGGAGAAAAAAGGCTGCACCGGTGCGT CAGCAGAATATGTGATACAGGATATATTCCGCTTCCTCGCTCACTGAC TCGCTACGCTCGGTCGTTCGACTGCGGCGAGCGGAAATGGCTTACGAA CGGGGCGGAGATTTCCTGGAAGATGCCAGGAAGATACTTAACAGGGAA GTGAGAGGGCCGCGGCAAAGCCGTTTTTCCATAGGCTCCGCCCCCCTG ACAAGCATCACGAAATCTGACGCTCAAATCAGTGGTGGCGAAACCCGA CAGGACTATAAAGATACCAGGCGTTTCCCCCTGGCGGCTCCCTCGTGC GCTCTCCTGTTCCTGCCTTTCGGTTTACCGGTGTCATTCCGCTGTTAT GGCCGCGTTTGTCTCATTCCACGCCTGACACTCAGTTCCGGGTAGGCA GTTCGCTCCAAGCTGGACTGTATGCACGAACCCCCCGTTCAGTCCGAC CGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGAAAGA CATGCAAAAGCACCACTGGCAGCAGCCACTGGTAATTGATTTAGAGGA GTTAGTCTTGAAGTCATGCGCCGGTTAAGGCTAAACTGAAAGGACAAG TTTTGGTGACTGCGCTCCTCCAAGCCAGTTACCTCGGTTCAAAGAGTT GGTAGCTCAAAAAACCTTCGAAAAACCGCCCTGCAAGGCGGTTTTTTC GTTTTCAGAGCAAGAGATTACGCGCAGACCAAAACGATCTCAAGAAGA TCATCTTATTAATCAGATAAAATATTTCTAGATTTCAGTGCAATTTAT CTCTTCAAATGTAGCACCTGAAGTCAGCCCCATACGATATAAGTTGTA ATTCTCATGTTAGTCATGCCCCGCGCCCACCGGAAGGAGCTGACTGGG TTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTG AGCTAACTTACATTAATTGCGTTGCGCTCAGGCTCGATCGGTGTGAAA AGTCAGTATCCAGTCGTGTAGTTCTTATTACCTGTCCACTAGTCTTGG ACTCCTGTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTT GTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTATTGGTGAGAATCC AGGGGTCCCCAATAATTACGATTTAAATTTGACATAAGCCTGTTCGGT TCGTAAACTGTAATGCAAGTAGCGTATGCGCTCACGCAACTGGTCCAG AACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTGGCGGTTTTCAT GGCTTGTTATGACTGTTTTTTTGTACAGCCTATGCCTCGGGCATCCAA GCAGCAAGCGCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAG CAACGATGTTACGCAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCC CTAAAACAAAGTTAGGTGGCTCAAGTATGGGCATCATTCGCACATGTA GGCTCGGCCCTGACCAAGTCAAATCCATGCGGGCTGCTCTTGATCTTT TCGGTCGTGAGTTCGGAGACGTAGCCACCTACTCCCAACATCAGCCGG ACTCCGATTACCTCGGGAACTTGCTCCGTAGTAAGACATTCATCGCGC TTGCTGCCTTCGACCAAGAAGCGGTTGTTGGCGCTCTCGCGGCTTACG TTCTGCCCAAGTTTGAGCAGCCGCGTAGTGAGATCTATATCTATGATC TCGCAGTCTCCGGAGAGCACCGGAGGCAGGGCATTGCCACCGCGCTCA TCAATCTCCTCAAGCATGAGGCCAACGCGCTTGGTGCTTATGTGATCT ACGTGCAAGCAGATTACGGTGACGATCCCGCAGTGGCTCTCTATACAA AGTTGGGCATACGGGAAGAAGTGATGCACTTTGATATCGACCCAAGTA CCGCCACCTAACAATTCGTTCAAGCCGAGATCGGCTTCCCGGCCGCGG AGTTGTTCGGTAAATTGGACAACGGTCGGCTCGATCACGGCACTACAC TCGTTGCTTTATCGGTATTGTTATTACAGAGTCCGACACCATCGAATG GTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCA ATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGA GTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAGGC CAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGAC

	CGTATTACATAGTGTGGATTTTTTTCCGTCTGGTAACGCGTCCGTGGC GATAGAACCCCGGCTCCCGCAGGCGGATTTTCCTGAACATCATCATGA TTTTCATGAAATTGTGATTGTCGAACATGGCACGGGTATTCATGTGTT TAATGGGCAGCCCTATACCATCACCGGTGGCACGGTCTGTTTCGTACG CGATCATGATCGGCATCTGTATGAACATACCGATAATCTGTGTCTGAC CAATGTGCTGTATCGCTCGCCGGATCGATTTCAGTTTCTCGCCGGGCT GAATCAGTTGCTGCCACAAGAGCTGGATGGGCAGTATCCGTCTCACTG GCGCGTTAACCACAGCGTATTGCAGCAGGTGCGACAGCTGGTTGCACA GATGGAACAGCAGGAAGGGGAAAATGATTTACCCTCGACCGCCAGTCG CGAGATCTTGTTTATGCAATTACTGCTCTTGCTGCGTAAAAGCAGTTT GCAGGAGAACCTGGAAAACAGCGCATCACGTCTCAACTTGCTTCTGGC CTGGCTGGAGGACCATTTTGCCGATGAGGTGAATTGGGATGCCGTGGC GGATCAATTTTCTCTTTCACTGCGTACGCTACATCGGCAGCTTAAGCA GCAAACGGGACTGACGCCTCAGCGATACCTGAACCGCCTGCGACTGAT GAAAGCCCGACATCTGCTACGCCACAGCGAGGCCAGCGTTACTGACAT CGCCT
pCDF-Plac-GFP-SpecR	ACACTGGAAGGACAGATTTGGTTGCTGTGCTCTGCGAAAGCCAGTTAC CACGGTTAAGCAGTTCCCCAACTGACTTAACCTTCGATCAAACCACCT CCCCAGGTGGTTTTTTCGTTTACAGGGCAAAAGATTACGCGCAGAAAA AAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACTGAACCGCTCTAG ATTTCAGTGCAATTTATCTCTTCAAATGTAGCACCTGAAGTCAGCCCC ATACGGCTCGATCGGTGTGAAAAGTCAGTATCCAGTCGTGTAGTTCTT ATTACCTGTCCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGG CATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATGTGAAAC CAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGA CCGTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGC GGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACC GCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTG CCACCTCCAGTCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGA TTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTCGATGG TAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTC TCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACC AGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTAT TTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCC ATGAAGACGGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTC ACCAGCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGC GTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTC AGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTC AACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGC TGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCG AGTCCGGGCTGCGCGTTGGTGCGGATATCTCGGTAGTGGGATACGACG ATACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACCATCAAAC AGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAAC TCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCAC TGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTC CCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCC GACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATACTAGTGAGCGG ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCG CACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTAT CATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCT TCACCTCGAGAAATTTATCAAAAAGAGTGTTGACTTGTGAGCGGATAA CAATGATACTTAGATTCAATTGTGAGCGGATAACAATTGGCTCGGGAG ACCTATCGGTAATAACAGTCCAATCTGGTGTAACTTCGGAATCGTCCT TTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCTACTAGAGTCACAC AGGAAAGTACTAGATGCGTAAAGGCGAAGAACTGTTCACGGGCGTAGT TCCGATTCTGGTCGAGCTGGACGGCGATGTGAACGGTCATAAGTTTAG CGTTCGCGGTGAAGGTGAGGGCGACGCGACCAACGGCAAACTGACCCT GAAGTTCATCTGCACCACCGGTAAACTGCCGGTGCCTTGGCCGACCTT GGTGACGACGTTGACGTATGGCGTGCAGTGTTTTGCGCGTTATCCGGA CCACATGAAACAACACGATTTCTTCAAATCTGCGATGCCGGAGGGTTA CGTCCAGGAGCGTACCATTTCCTTCAAGGATGATGGCACTTACAAAAC

	TCGCGCAGAGGTTAAGTTTGAAGGTGACACGCTGGTCAATCGTATCGA ATTGAAGGGTATCGACTTTAAAGAGGATGGTAACATTCTGGGCCATAA ACTGGAGTATAACTTCAACAGCCATAATGTTTACATTACGGCAGACAA GCAAAAGAACGGCATCAAGGCCAATTTCAAGATTCGCCACAATGTTGA GGACGGTAGCGTCCAACTGGCCGACCATTACCAGCAGAACACCCCAAT TGGTGACGGTCCGGTTTTGCTGCCGGATAATCACTATCTGAGCACCCA AAGCGTGCTGAGCAAAGATCCGAACGAAAAACGTGATCACATGGTCCT GCTGGAATTTGTGACCGCTGCGGGCATCACCCACGGTATGGACGAGCT GTATAAGCGTCCGTAATAATACTAGAGCCAGGCATCAAATAAAACGAA AGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGG TGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTT TCTGCGTTTATAGGCTCGGGTAAGAACTCGCACTTCGTGGAAACACTA TTATCTGGTGGGTCTCTGTCCCCAGGCATCAAATAAAACGAAAGGCTC AGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACG CTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCG TTTATAGGCTCGTTACTTACGACACTCCGAGACAGTCAGAGGGTATTT ATTGAACTAGTCCGCATTTGAGAAGCACACGGTCACACTGCTTCCGGT AGTCAATAAACCGGTAAACCAGCAATAGACATAAGCGGCTATTTAACG ACCCTGCCCTGAACCGACGACCGGGTCATCGTGGCCGGATCTTGCGGC CCCTCGGCTTGAACGAATTGTTAGACATTATTTGCCGACTACCTTGGT GATCTCGCCTTTCACGTAGTGGACAAATTCTTCCAACTGATCTGCGCG CGAGGCCAAGCGATCTTCTTCTTGTCCAAGATAAGCCTGTCTAGCTTC AAGTATGACGGGCTGATACTGGGCCGGCAGGCGCTCCATTGCCCAGTC GGCAGCGACATCCTTCGGCGCGATTTTGCCGGTTACTGCGCTGTACCA AATGCGGGACAACGTAAGCACTACATTTCGCTCATCGCCAGCCCAGTC GGGCGGCGAGTTCCATAGCGTTAAGGTTTCATTTAGCGCCTCAAATAG ATCCTGTTCAGGAACCGGATCAAAGAGTTCCTCCGCCGCTGGACCTAC CAAGGCAACGCTATGTTCTCTTGCTTTTGTCAGCAAGATAGCCAGATC AATGTCGATCGTGGCTGGCTCGAAGATACCTGCAAGAATGTCATTGCG CTGCCATTCTCCAAATTGCAGTTCGCGCTTAGCTGGATAACGCCACGG AATGATGTCGTCGTGCACAACAATGGTGACTTCTACAGCGCGGAGAAT CTCGCTCTCTCCAGGGGAAGCCGAAGTTTCCAAAAGGTCGTTGATCAA AGCTCGCCGCGTTGTTTCATCAAGCCTTACGGTCACCGTAACCAGCAA ATCAATATCACTGTGTGGCTTCAGGCCGCCATCCACTGCGGAGCCGTA CAAATGTACGGCCAGCAACGTCGGTTCGAGATGGCGCTCGATGACGCC AACTACCTCTGATAGTTGAGTCGATACTTCGGCGATCACCGCTTCCCT CATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTG TCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAAT AGCTAGCTCACTCGGTCGCTACGCTCCGGGCGTGAGACTGCGGCGGGC GCTGCGGACACATACAAAGTTACCCACAGATTCCGTGGATAAGCAGGG GACTAACATGTGAGGCAAAACAGCAGGGCCGCGCCGGTGGCGTTTTTC CATAGGCTCCGCCCCTCCTGCCAGAGTTCACATAAACAGACGCTTTTCC GGTGCATCTGTGGGAGCCGTGAGGCTCAACCATGAATCTGACAGTACG GGCGAAACCCGACAGGACTTAAAGATCCCCACCGTTTCCGGCGGGTCG CTCCCTCTTGCGCTCTCCTGTTCCGACCCTGCCGTTTACCGGATACCT GTTCCGCCTTTCTCCCTTACGGGAAGTGTGGCGCTTTCTCATAGCTCA CACACTGGTATCTCGGCTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCT GTAAGCAAGAACTCCCCGTTCAGCCCGACTGCTGCGCCTTATCCGGTA ACTGTTCACTTGAGTCCAACCCGGAAAAGCACGGTAAAACGCCACTGG CAGCAGCCATTGGTAACTGGGAGTTCGCAGAGGATTTGTTTAGCTAAA CACGCGGTTGCTCTTGAAGTGTGCGCCAAAGTCCGGCT
pCDF-PrhaBAD-GFP-GmR	GACCATTTTGCCGATGAGGTGAATTGGGATGCCGTGGCGGATCAATTT TCTCTTTCACTGCGTACGCTACATCGGCAGCTTAAGCAGCAAACGGGA CTGACGCCTCAGCGATACCTGAACCGCCTGCGACTGATGAAAGCCCGA CATCTGCTACGCCACAGCGAGGCCAGCGTTACTGACATCGCCTATCGC TGTGGATTCAGCGACAGTAACCACTTTTCGACGCTTTTTCGCCGAGAG TTTAACTGGTCACCGCGTGATATTCGCCAGGGACGGGATGGCTTTCTG CAATAAGCGCAACGCAATTAATACTAGTGAGCGGATACATATTTGAAT GTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAA AAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCT ATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCACCTCGAGAAAT

	TTATCAAAAAGAGTGCCACAATTCAGCAAATTGTGAACATCATCACGT TCATCTTTCCCTGGTTGCCAATGGCCCATTTTCCTGTCAGTAACGAGA AGGTCGCGTATTCAGGCGCTTTTTAGACTGGTCGTAATGAAGGGGAGA CCTATCGGTAATAACAGTCCAATCTGGTGTAACTTCGGAATCGTCCTT TACAGCTAGCTCAGTCCTAGGTATTATGCTAGCTACTAGAGTCACACA GGAAAGTACTAGATGCGTAAAGGCGAAGAACTGTTCACGGGCGTAGTT CCGATTCTGGTCGAGCTGGACGGCGATGTGAACGGTCATAAGTTTAGC GTTCGCGGTGAAGGTGAGGGCGACGCGACCAACGGCAAACTGACCCTG AAGTTCATCTGCACCACCGGTAAACTGCCGGTGCCTTGGCCGACCTTG GTGACGACGTTGACGTATGGCGTGCAGTGTTTTGCGCGTTATCCGGAC CACATGAAACAACACGATTTCTTCAAATCTGCGATGCCGGAGGGTTAC GTCCAGGAGCGTACCATTTCCTTCAAGGATGATGGCACTTACAAAACT CGCGCAGAGGTTAAGTTTGAAGGTGACACGCTGGTCAATCGTATCGAA TTGAAGGGTATCGACTTTAAAGAGGATGGTAACATTCTGGGCCATAAA CTGGAGTATAACTTCAACAGCCATAATGTTTACATTACGGCAGACAAG CAAAAGAACGGCATCAAGGCCAATTTCAAGATTCGCCACAATGTTGAG GACGGTAGCGTCCAACTGGCCGACCATTACCAGCAGAACACCCCAATT GGTGACGGTCCGGTTTTGCTGCCGGATAATCACTATCTGAGCACCCAA AGCGTGCTGAGCAAAGATCCGAACGAAAAACGTGATCACATGGTCCTG CTGGAATTTGTGACCGCTGCGGGCATCACCCACGGTATGGACGAGCTG TATAAGCGTCCGTAATAATACTAGAGCCAGGCATCAAATAAAACGAAA GGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGT GAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTT CTGCGTTTATAGGCTCGGGTAAGAACTCGCACTTCGTGGAAACACTAT TATCTGGTGGGTCTCTGTCCCCAGGCATCAAATAAAACGAAAGGCTCA GTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGC TCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGT TTATAGGAGAAGTAGTGCCACAGACAGTATTGCTTACGAGTTGATTTA TCCTGTCCTCACTCGGTCGCTACGCTCCGGGCGTGAGACTGCGGCGGG CGCTGCGGACACATACAAAGTTACCCACAGATTCCGTGGATAAGCAGG GGACTAACATGTGAGGCAAAACAGCAGGGCCGCGCCGGTGGCGTTTTT CCATAGGCTCCGCCCTCCTGCCAGAGTTCACATAAACAGACGCTTTTC CGGTGCATCTGTGGGAGCCGTGAGGCTCAACCATGAATCTGACAGTAC GGGCGAAACCCGACAGGACTTAAAGATCCCCACCGTTTCCGGCGGGTC GCTCCCTCTTGCGCTCTCCTGTTCCGACCCTGCCGTTTACCGGATACC TGTTCCGCCTTTCTCCCTTACGGGAAGTGTGGCGCTTTCTCATAGCTC ACACACTGGTATCTCGGCTCGGTGTAGGTCGTTCGCTCCAAGCTGGGC TGTAAGCAAGAACTCCCCGTTCAGCCCGACTGCTGCGCCTTATCCGGT AACTGTTCACTTGAGTCCAACCCGGAAAAGCACGGTAAAACGCCACTG GCAGCAGCCATTGGTAACTGGGAGTTCGCAGAGGATTTGTTTAGCTAA ACACGCGGTTGCTCTTGAAGTGTGCGCCAAAGTCCGGCTACACTGGAA GGACAGATTTGGTTGCTGTGCTCTGCGAAAGCCAGTTACCACGGTTAA GCAGTTCCCCAACTGACTTAACCTTCGATCAAACCACCTCCCCAGGTG GTTTTTTCGTTTACAGGGCAAAAGATTACGCGCAGAAAAAAAGGATCT CAAGAAGATCCTTTGATCTTTTCTACTGAACCGCTCTAGATTTCAGTG CAATTTATCTCTTCAAATGTAGCACCTGAAGTCAGCCCCATACGGATC GGTGTGAAAAGTCAGTATCCAGTCGTGTAGTTCTTATTACCTGTCCAC TAGTCTTGGACTCCTGTTGATAGATCCAGTAATGACCTCAGAACTCCA TCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTATTGG TGAGAATCCAGGGGTCCCCAATAATTACGATTTAAATTTGACATAAGC CTGTTCGGTTCGTAAACTGTAATGCAAGTAGCGTATGCGCTCACGCAA CTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTGGC GGTTTTCATGGCTTGTTATGACTGTTTTTTTGTACAGCCTATGCCTCG GGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTCGATGTTTGATGTT ATGGAGCAGCAACGATGTTACGCAGCAGCAACGATGTTACGCAGCAGG GCAGTCGCCCTAAAACAAAGTTAGGTGGCTCAAGTATGGGCATCATTC GCACATGTAGGCTCGGCCCTGACCAAGTCAAATCCATGCGGGCTGCTC TTGATCTTTTCGGTCGTGAGTTCGGAGACGTAGCCACCTACTCCCAAC ATCAGCCGGACTCCGATTACCTCGGGAACTTGCTCCGTAGTAAGACAT TCATCGCGCTTGCTGCCTTCGACCAAGAAGCGGTTGTTGGCGCTCTCG CGGCTTACGTTCTGCCCAAGTTTGAGCAGCCGCGTAGTGAGATCTATA

	TCTATGATCTCGCAGTCTCCGGAGAGCACCGGAGGCAGGGCATTGCCA CCGCGCTCATCAATCTCCTCAAGCATGAGGCCAACGCGCTTGGTGCTT ATGTGATCTACGTGCAAGCAGATTACGGTGACGATCCCGCAGTGGCTC TCTATACAAAGTTGGGCATACGGGAAGAAGTGATGCACTTTGATATCG ACCCAAGTACCGCCACCTAACAATTCGTTCAAGCCGAGATCGGCTTCC CGGCCGCGGAGTTGTTCGGTAAATTGGACAACGGTCGGATCACGGCAC TACACTCGTTGCTTTATCGGTATTGTTATTACAGAGTCCGACACCATC GAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAG AGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTC GCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAAC CAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCG ATGACCGTATTACATAGTGTGGATTTTTTTCCGTCTGGTAACGCGTCC GTGGCGATAGAACCCCGGCTCCCGCAGGCGGATTTTCCTGAACATCAT CATGATTTTCATGAAATTGTGATTGTCGAACATGGCACGGGTATTCAT GTGTTTAATGGGCAGCCCTATACCATCACCGGTGGCACGGTCTGTTTC GTACGCGATCATGATCGGCATCTGTATGAACATACCGATAATCTGTGT CTGACCAATGTGCTGTATCGCTCGCCGGATCGATTTCAGTTTCTCGCC GGGCTGAATCAGTTGCTGCCACAAGAGCTGGATGGGCAGTATCCGTCT CACTGGCGCGTTAACCACAGCGTATTGCAGCAGGTGCGACAGCTGGTT GCACAGATGGAACAGCAGGAAGGGGAAAATGATTTACCCTCGACCGCC AGTCGCGAGATCTTGTTTATGCAATTACTGCTCTTGCTGCGTAAAAGC AGTTTGCAGGAGAACCTGGAAAACAGCGCATCACGTCTCAACTTGCTT CTGGCCTGGCTGGAG
pET-Plac-GFP-EryR	CCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCC GACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATACTAGTGAGCGG ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCG CACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTAT CATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCT TCACCTCGAGAAATTTATCAAAAAGAGTGTTGACTTGTGAGCGGATAA CAATGATACTTAGATTCAATTGTGAGCGGATAACAATTGGGGAGACCT ATCGGTAATAACAGTCCAATCTGGTGTAACTTCGGAATCGTCCTTTAC AGCTAGCTCAGTCCTAGGTATTATGCTAGCTACTAGAGTCACACAGGA AAGTACTAGATGCGTAAAGGCGAAGAACTGTTCACGGGCGTAGTTCCG ATTCTGGTCGAGCTGGACGGCGATGTGAACGGTCATAAGTTTAGCGTT CGCGGTGAAGGTGAGGGCGACGCGACCAACGGCAAACTGACCCTGAAG TTCATCTGCACCACCGGTAAACTGCCGGTGCCTTGGCCGACCTTGGTG ACGACGTTGACGTATGGCGTGCAGTGTTTTGCGCGTTATCCGGACCAC ATGAAACAACACGATTTCTTCAAATCTGCGATGCCGGAGGGTTACGTC CAGGAGCGTACCATTTCCTTCAAGGATGATGGCACTTACAAAACTCGC GCAGAGGTTAAGTTTGAAGGTGACACGCTGGTCAATCGTATCGAATTG AAGGGTATCGACTTTAAAGAGGATGGTAACATTCTGGGCCATAAACTG GAGTATAACTTCAACAGCCATAATGTTTACATTACGGCAGACAAGCAA AAGAACGGCATCAAGGCCAATTTCAAGATTCGCCACAATGTTGAGGAC GGTAGCGTCCAACTGGCCGACCATTACCAGCAGAACACCCCAATTGGT GACGGTCCGGTTTTGCTGCCGGATAATCACTATCTGAGCACCCAAAGC GTGCTGAGCAAAGATCCGAACGAAAAACGTGATCACATGGTCCTGCTG GAATTTGTGACCGCTGCGGGCATCACCCACGGTATGGACGAGCTGTAT AAGCGTCCGTAATAATACTAGAGCCAGGCATCAAATAAAACGAAAGGC TCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAA CGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTG CGTTTATAGGCTCGGGTAAGAACTCGCACTTCGTGGAAACACTATTAT CTGGTGGGTCTCTGTCCCCAGGCATCAAATAAAACGAAAGGCTCAGTC GAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCT CTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTA TAGGTTACTTACGACACTCCGAGACAGTCAGAGGGTATTTATTGAACT AGTCCTAGGTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACT GAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTT TTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTAC CAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGA AGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAG TGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTA

	CATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCG ATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATA AGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCT TGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTAT GAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGG TAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGG GAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGAC TTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGA AAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGC CTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATA ACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAA CGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGA TGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCAATG GTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTA TACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACAC CCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCA TCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAG AGGTTTTCACCGTCATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGC TCATCAGCGTGGTCGTGAAGCGATTCACAGATGTCTGCCTGTTCATCC GCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTG ATAAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTGAT GCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATACCGATG AAACGAGAGAGGATGCTCACGATACGGGTTACTGATGATGAACATGCC CGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGG CGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAAT ACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGATGCAG ATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTT TACGAAACACGGAAACCGAAGACCATTCATGTTGTTGCTCAGGTCGCA GACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGTGAT TCATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTC AACGACAGGAGCACGATCATGCTAGTCATGCCCCGCGCCCACCGGAAG GAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCGGT GCCTAATGAGTGAGCTAACTTACATTAATTGCGTTGCGCGGATCGGTG TGAAAAGTCAGTATCCAGTCGTGTAGTTCTTATTACCTGTCCACTAGT ATTCGATGAATTCGCGTGCTATAATTATACTAATTTTATAAGGAGGAA AAAATATGGGCATTTTTAGTATTTTTGTAATCAGCACAGTTCATTATC AACCAAACAAAAAATAAGTGGTTATAATGAATCGTTAATAAGCAAAAT TCATATAACCAAATTAAAGAGGGTTATAATGAACGAGAAAAATATAAA ACACAGTCAAAACTTTATTACTTCAAAACATAATATAGATAAAATAAT GACAAATATAAGATTAAATGAACATGATAATATCTTTGAAATCGGCTC AGGAAAAGGCCATTTTACCCTTGAATTAGTAAAGAGGTGTAATTTCGT AACTGCCATTGAAATAGACCATAAATTATGCAAAACTACAGAAAATAA ACTTGTTGATCACGATAATTTCCAAGTTTTAAACAAGGATATATTGCA GTTTAAATTTCCTAAAAACCAATCCTATAAAATATATGGTAATATACC TTATAACATAAGTACGGATATAATACGCAAAATTGTTTTTGATAGTAT AGCTAATGAGATTTATTTAATCGTGGAATACGGGTTTGCTAAAAGATT ATTAAATACAAAACGCTCATTGGCATTACTTTTAATGGCAGAAGTTGA TATTTCTATATTAAGTATGGTTCCAAGAGAATATTTTCATCCTAAACC TAAAGTGAATAGCTCACTTATCAGATTAAGTAGAAAAAAATCAAGAAT ATCACACAAAGATAAACAAAAGTATAATTATTTCGTTATGAAATGGGT TAACAAAGAATACAAGAAAATATTTACAAAAAATCAATTTAACAATTC CTTAAAACATGCAGGAATTGACGATTTAAACAATATTAGCTTTGAACA ATTCTTATCTCTTTTCAATAGCTATAAATTATTTAATAAGTAACTGCA GGGATCACGGCACTACACTCGTTGCTTTATCGGTATTGTTATTACAGA GTCCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATA GCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAAC GTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTC CCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAA AGTGGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCGCGTGGC ACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTC CAGTCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATC

	TCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAACG AAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCA ACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAGGATGC CATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGA TGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGA CGGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCA AATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCG TCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGAT AGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAAC CATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGC CAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGG GCTGCGCGTTGGTGCGGATATCTCGGTAGTGGGATACGACGATACCGA AGACAGCTCATGTTATATCCCGCCGTTAACCACCATCAAACAGGATTT TCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCA GGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAA AAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTC
pET-Plac-GFP-KanR	TTACTTACGACACTCCGAGACAGTCAGAGGGTATTTATTGAACTAGTC CGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCG CCCGGAAGAGAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACGTT ATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCG CGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGT GGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCGCGTGGCACA ACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAG TCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCG CGCCGATCAACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAACGAAG CGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACG CGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAGGATGCCAT TGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGT CTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGG TACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAAT CGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCT GGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGC GGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCAT GCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAA CGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCT GCGCGTTGGTGCGGATATCTCGGTAGTGGGATACGACGATACCGAAGA CAGCTCATGTTATATCCCGCCGTTAACCACCATCAAACAGGATTTTCG CCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGG CCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAG AAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTT GGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAG CGGGCAGTGAGCGCAACGCAATTAATACTAGTGAGCGGATACATATTT GAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCC CGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTA ACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCACCTCGAG AAATTTATCAAAAAGAGTGTTGACTTGTGAGCGGATAACAATGATACT TAGATTCAATTGTGAGCGGATAACAATTGGCTCGGGAGACCTATCGGT AATAACAGTCCAATCTGGTGTAACTTCGGAATCGTCCTTTACAGCTAG CTCAGTCCTAGGTATTATGCTAGCTACTAGAGTCACACAGGAAAGTAC TAGATGCGTAAAGGCGAAGAACTGTTCACGGGCGTAGTTCCGATTCTG GTCGAGCTGGACGGCGATGTGAACGGTCATAAGTTTAGCGTTCGCGGT GAAGGTGAGGGCGACGCGACCAACGGCAAACTGACCCTGAAGTTCATC TGCACCACCGGTAAACTGCCGGTGCCTTGGCCGACCTTGGTGACGACG TTGACGTATGGCGTGCAGTGTTTTGCGCGTTATCCGGACCACATGAAA CAACACGATTTCTTCAAATCTGCGATGCCGGAGGGTTACGTCCAGGAG CGTACCATTTCCTTCAAGGATGATGGCACTTACAAAACTCGCGCAGAG GTTAAGTTTGAAGGTGACACGCTGGTCAATCGTATCGAATTGAAGGGT ATCGACTTTAAAGAGGATGGTAACATTCTGGGCCATAAACTGGAGTAT AACTTCAACAGCCATAATGTTTACATTACGGCAGACAAGCAAAAGAAC GGCATCAAGGCCAATTTCAAGATTCGCCACAATGTTGAGGACGGTAGC GTCCAACTGGCCGACCATTACCAGCAGAACACCCCAATTGGTGACGGT

	CCGGTTTTGCTGCCGGATAATCACTATCTGAGCACCCAAAGCGTGCTG AGCAAAGATCCGAACGAAAAACGTGATCACATGGTCCTGCTGGAATTT GTGACCGCTGCGGGCATCACCCACGGTATGGACGAGCTGTATAAGCGT CCGTAATAATACTAGAGCCAGGCATCAAATAAAACGAAAGGCTCAGTC GAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCT CTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTA TAGGCTCGGGTAAGAACTCGCACTTCGTGGAAACACTATTATCTGGTG GGTCTCTGTCCCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGA CTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTA GAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAGGCT CGATCGGTGTGAAAAGTCAGTATCCAGTCGTGTAGTTCTTATTACCTG TCCGTTCGATTTATTCAACAAAGCCACGTTGTGTCTCAAAATCTCTGA TGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTG TCTGCTTACATAAACAGTAATACAAGTGGTGTTATGAGCCATATTCAA CGGGAAACGTCTTGCTCCAGGCCGCGATTAAATTCCAACATGGATGCT GATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGT GCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTT CTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATG GTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAAG CATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATC CCCGGGAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGT GAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCG ATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTC GCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGAT TTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAGGAA ATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGT GATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGT TGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTT GCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAA CGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTG CAGTTTCATTTGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAAT TGGTTGTAACACTGGCAGAGGGCTCGATCACGGCACTACACTCGTTGC TTTATCGGTATTGTTATTACAGA
pRSF1010-PcoA-GFP-EryR	GTCCCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCC TTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACA AATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGGG TGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCAAATTAAGCAG AAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTTTT GTTTATTTTTCTAAATACATTCAAATATGTATCCGCCCTAGGAAAGAG GAGAAAGGTACACCGCTAGCTGAAAGCGACCAGGTGCTCGGCGTGGCA AGACTCGCAGCGAACCCGTAGAAAGCCATGCTCCAGCCGCCCGCATTG GAGAAATTCTTCAAATTCCCGTTGCACATAGCCCGGCAATTCCTTTCC CTGCTCTGCCATAAGCGCAGCGAATGCCGGGTAATACTCGTCAACGAT CTGATAGAGAAGGGTTTGCTCGGGTCGGTGGCTCTGGTAACGACCAGT ATCCCGATCCCGGCTGGCCGTCCTGGCCGCCACATGAGGCATGTTCCG CGTCCTTGCAATACTGTGTTTACATACAGTCTATCGCTTAGCGGAAAG TTCTTTTACCCTCAGCCGAAATGCCTGCCGTTGCTAGACATTGCCAGC CAGTGCCCGTCACTCCCGTACTAACTGTCACGAACCCCTGCAATAACT GTCACGCCCCCCTGCAATAACTGTCACGAACCCCTGCAATAACTGTCA CGCCCCCAAACCTGCAAACCCAGCAGGGGCGGGGGCTGGCGGGGTGTT GGAAAAATCCATCCATGATTATCTAAGAATAATCCACTAGGCGCGGTT ATCAGCGCCCTTGTGGGGCGCTGCTGCCCTTGCCCAATATGCCCGGCC AGAGGCCGGATAGCTGGTCTATTCGCTGCGCTAGGCTACACACCGCCC CACCGCTGCGCGGCAGGGGGAAAGGCGGGCAAAGCCCGCTAAACCCCA CACCAAACCCCGCAGAAATACGCTGGAGCGCTTTTAGCCGCTTTAGCG GCCTTTCCCCCTACCCGAAGGGTGGGGGCGCGTGTGCAGCCCCGCAGG GCCTGTCTCGGTCGATCATTCAGCCCGGCTCATCCTTCTGGCGTGGCG GCAGACCGAACAAGGCGCGGTCGTGGTCGCGTTCAAGGTACGCATCCA TTGCCGCCATGAGCCGATCCTCCGGCCACTCGCTGCTGTTCACCTTGG CCAAAATCATGGCCCCCACCAGCACCTTGCGCCTTGTTTCGTTCTTGC GCTCTTGCTGCTGTTCCCTTGCCCGCACCCGCTGAATTTCGGCATTGA

	TTCGCGCTCGTTGTTCTTCGAGCTTGGCCAGCCGATCCGCCGCCTTGT TGCTCCCCTTAACCATCTTGACACCCCATTGTTAATGTGCTGTCTCGT AGGCTATCATGGAGGCACAGCGGCGGCAATCCCGACCCTACTTTGTAG GGGAGGGCGCACTTACCGGTTTCTCTTCGAGAAACTGGCCTAACGGCC ACCCTTCGGGCGGTGCGCTCTCCGAGGGCCATTGCATGGAGCCGAAAA GCAAAAGCAACAGCGAGGCAGCATGGCGATTTATCACCTTACGGCGAA AACCGGCAGCAGGTCGGGCGGCCAATCGGCCAGGGCCAAGGCCGACTA CATCCAGCGCGAAGGCAAGTATGCCCGCGACATGGATGAAGTCTTGCA CGCCGAATCCGGGCACATGCCGGAGTTCGTCGAGCGGCCCGCCGACTA CTGGGATGCTGCCGACCTGTATGAACGCGCCAATGGGCGGCTGTTCAA GGAGGTCGAATTTGCCCTGCCGGTCGAGCTGACCCTCGACCAGCAGAA GGCGCTGGCGTCCGAGTTCGCCCAGCACCTGACCGGTGCCGAGCGCCT GCCGTATACGCTGGCCATCCATGCCGGTGGCGGCGAGAACCCGCACTG CCACCTGATGATCTCCGAGCGGATCAATGACGGCATCGAGCGGCCCGC CGCTCAGTGGTTCAAGCGGTACAACGGCAAGACCCCGGAGAAGGGCGG GGCACAGAAGACCGAAGCGCTCAAGCCCAAGGCATGGCTTGAGCAGAC CCGCGAGGCATGGGCCGACCATGCCAACCGGGCATTAGAGCGGGCTGG CCACGACGCCCGCATTGACCACAGAACACTTGAGGCGCAGGGCATCGA GCGCCTGCCCGGTGTTCACCTGGGGCCGAACGTGGTGGAGATGGAAGG CCGGGGCATCCGCACCGACCGGGCAGACGTGGCCCTGAACATCGACAC CGCCAACGCCCAGATCATCGACTTACAGGAATACCGGGAGGCAATAGA CCATGAACGCAATCGACAGAGTGAAGAAATCCAGAGGCATCAACGAGT TAGCGGAGCAGATCGAACCGCTGGCCCAGAGCATGGCGACACTGGCCG ACGAAGCCCGGCAGGTCATGAGCCAGACCCAGCAGGCCAGCGAGGCGC AGGCGGCGGAGTGGCTGAAAGCCCAGCGCCAGACAGGGGCGGCATGGG TGGAGCTGGCCAAAGAGTTGCGGGAGGTAGCCGCCGAGGTGAGCAGCG CCGCGCAGAGCGCCCGGAGCGCGTCGCGGGGGTGGCACTGGAAGCTAT GGCTAACCGTGATGCTGGCTTCCATGATGCCTACGGTGGTGCTGCTGA TCGCATCGTTGCTCTTGCTCGACCTGACGCCACTGACAACCGAGGACG GCTCGATCTGGCTGCGCTTGGTGGCCCGATGAAGAACGACAGGACTTT GCAGGCCATAGGCCGACAGCTCAAGGCCATGGGCTGTGAGCGCTTCGA TATCGGCGTCAGGGACGCCACCACCGGCCAGATGATGAACCGGGAATG GTCAGCCGCCGAAGTGCTCCAGAACACGCCATGGCTCAAGCGGATGAA TGCCCAGGGCAATGACGTGTATATCAGGCCCGCCGAGCAGGAGCGGCA TGGTCTGGTGCTGGTGGACGACCTCAGCGAGTTTGACCTGGATGACAT GAAAGCCGAGGGCCGGGAGCCTGCCCTGGTAGTGGAAACCAGCCCGAA GAACTATCAGGCATGGGTCAAGGTGGCCGACGCCGCAGGCGGTGAACT TCGGGGGCAGATTGCCCGGACGCTGGCCAGCGAGTACGACGCCGACCC GGCCAGCGCCGACAGCCGCCACTATGGCCGCTTGGCGGGCTTCACCAA CCGCAAGGACAAGCACACCACCCGCGCCGGTTATCAGCCGTGGGTGCT GCTGCGTGAATCCAAGGGCAAGACCGCCACCGCTGGCCCGGCGCTGGT GCAGCAGGCTGGCCAGCAGATCGAGCAGGCCCAGCGGCAGCAGGAGAA GGCCCGCAGGCTGGCCAGCCTCGAACTGCCCGAGCGGCAGCTTAGCCG CCACCGGCGCACGGCGCTGGACGAGTACCGCAGCGAGATGGCCGGGCT GGTCAAGCGCTTCGGTGATGACCTCAGCAAGTGCGACTTTATCGCCGC GCAGAAGCTGGCCAGCCGGGGCCGCAGTGCCGAGGAAATCGGCAAGGC CATGGCCGAGGCCAGCCCAGCGCTGGCAGAGCGCAAGCCCGGCCACGA AGCGGATTACATCGAGCGCACCGTCAGCAAGGTCATGGGTCTGCCCAG CGTCCAGCTTGCGCGGGCCGAGCTGGCACGGGCACCGGCACCCCGCCA GCGAGGCATGGACAGGGGCGGGCCAGATTTCAGCATGTAGTGCTTGCG TTGGTACTCACGCCTGTTATACTATGAGTACTCACGCACAGAAGGGGG TTTTATGGAATACGAAAAAAGCGCTTCAGGGTCGGTCTACCTGATCAA AAGTGACAAGGGCTATTGGTTGCCCGGTGGCTTTGGTTATACGTCAAA CAAGGCCGAGGCTGGCCGCTTTTCAGTCGCTGATATGGCCAGCCTTAA CCTTGACGGCTGCACCTTGTCCTTGTTCCGCGAAGACAAGCCTTTCGG CCCCGGCAAGTTTCTCGGTGACTGATATGAAAGACCAAAAGGACAAGC AGACCGGCGACCTGCTGGCCAGCCCTGACGCTGTACGCCAAGCGCGAT ATGCCGAGCGCATGAAGGCCAAAGGGATGCGTCAGCGCAAGTTCTGGC TGACCGACGACGAATACGAGGCGCTGCGCGAGTGCCTGGAAGAACTCA GAGCGGCGCAGGGCGGGGGTAGTGACCCCGCCAGCGCCTAACCACCAA CTGCCTGCAAAGGAGGCAATCAATGGCTACCCATAAGCCTATCAATAT

	TCTGGAGGCGTTCGCAGCAGCGCCGCCACCGCTGGACTACGTTTTGCC CAACATGGTGGCCGGTACGGTCGGGGCGCTGGTGTCGCCCGGTGGTGC CGGTAAATCCATGCTGGCCCTGCAACTGGCCGCACAGATTGCAGGCGG GCCGGATCTGCTGGAGGTGGGCGAACTGCCCACCGGCCCGGTGATCTA CCTGCCCGCCGAAGACCCGCCCACCGCCATTCATCACCGCCTGCACGC CCTTGGGGCGCACCTCAGCGCCGAGGAACGGCAAGCCGTGGCTGACGG CCTGCTGATCCAGCCGCTGATCGGCAGCCTGCCCAACATCATGGCCCC GGAGTGGTTCGACGGCCTCAAGCGCGCCGCCGAGGGCCGCCGCCTGAT GGTGCTGGACACGCTGCGCCGGTTCCACATCGAGGAAGAAAACGCCAG CGGCCCCATGGCCCAGGTCATCGGTCGCATGGAGGCCATCGCCGCCGA TACCGGGTGCTCTATCGTGTTCCTGCACCATGCCAGCAAGGGCGCGGC CATGATGGGCGCAGGCGACCAGCAGCAGGCCAGCCGGGGCAGCTCGGT ACTGGTCGATAACATCCGCTGGCAGTCCTACCTGTCGAGCATGACCAG CGCCGAGGCCGAGGAATGGGGTGTGGACGACGACCAGCGCCGGTTCTT CGTCCGCTTCGGTGTGAGCAAGGCCAACTATGGCGCACCGTTCGCTGA TCGGTGGTTCAGGCGGCATGACGGCGGGGTGCTCAAGCCCGCCGTGCT GGAGAGGCAGCGCAAGAGCAAGGGGGTGCCCCGTGGTGAAGCCTAAGA ACAAGCACAGCCTCAGCCACGTCCGGCACGACCCGGCGCACTGTCTGG CCCCCGGCCTGTTCCGTGCCCTCAAGCGGGGCGAGCGCAAGCGCAGCA AGCTGGACGTGACGTATGACTACGGCGACGGCAAGCGGATCGAGTTCA GCGGCCCGGAGCCGCTGGGCGCTGATGATCTGCGCATCCTGCAAGGGC TGGTGGCCATGGCTGGGCCTAATGGCCTAGTGCTTGGCCCGGAACCCA AGACCGAAGGCGGACGGCAGCTCCGGCTGTTCCTGGAACCCAAGTGGG AGGCCGTCACCGCTGAATGCCATGTGGTCAAAGGTAGCTATCGGGCGC TGGCAAAGGAAATCGGGGCAGAGGTCGATAGTGGTGGGGCGCTCAAGC ACATACAGGACTGCATCGAGCGCCTTTGGAAGGTATCCATCATCGCCC AGAATGGCCGCAAGCGGCAGGGGTTTCGGCTGCTGTCGGAGTACGCCA GCGACGAGGCGGACGGGCGCCTGTACGTGGCCCTGAACCCCTTGATCG CGCAGGCCGTCATGGGTGGCGGCCAGCATGTGCGCATCAGCATGGACG AGGTGCGGGCGCTGGACAGCGAAACCGCCCGCCTGCTGCACCAGCGGC TGTGTGGCTGGATCGACCCCGGCAAAACCGGCAAGGCTTCCATAGATA CCTTGTGCGGCTATGTCTGGCCGTCAGAGGCCAGTGGTTCGACCATGC GCAAGCGCCGCCAGCGGGTGCGCGAGGCGTTGCCGGAGCTGGTCGCGC TGGGCTGGACGGTAACCGAGTTCGCGGCGGGCAAGTACGACATCACCC GGCCCAAGGCGGCAGGCTGACCCCCCCCACTCTATTGTAAACAAGACA TTTTTATCTTTTATATTCAATGGCTTATTTTCCTGCTAATTGGTAATA CCATGAAAAATACCATGCTCAGAAAAGGCTTAACAATATTTTGAAAAA TTGCCTACTGAGCGCTGCCGCACAGCTCCATAGGCCGCTTTCCTGGCT TTGCTTCCAGATGTATGCTCTTCTGCTCCTGCCGGCCGGGCTCGTTAC TTACGACACTCCGAGACAGTCAGAGGGTATTTATTGAACTAGTCCACT AGTATTCGATGAATTCGCGTGCTATAATTATACTAATTTTATAAGGAG GAAAAAATATGGGCATTTTTAGTATTTTTGTAATCAGCACAGTTCATT ATCAACCAAACAAAAAATAAGTGGTTATAATGAATCGTTAATAAGCAA AATTCATATAACCAAATTAAAGAGGGTTATAATGAACGAGAAAAATAT AAAACACAGTCAAAACTTTATTACTTCAAAACATAATATAGATAAAAT AATGACAAATATAAGATTAAATGAACATGATAATATCTTTGAAATCGG CTCAGGAAAAGGCCATTTTACCCTTGAATTAGTAAAGAGGTGTAATTT CGTAACTGCCATTGAAATAGACCATAAATTATGCAAAACTACAGAAAA TAAACTTGTTGATCACGATAATTTCCAAGTTTTAAACAAGGATATATT GCAGTTTAAATTTCCTAAAAACCAATCCTATAAAATATATGGTAATAT ACCTTATAACATAAGTACGGATATAATACGCAAAATTGTTTTTGATAG TATAGCTAATGAGATTTATTTAATCGTGGAATACGGGTTTGCTAAAAG ATTATTAAATACAAAACGCTCATTGGCATTACTTTTAATGGCAGAAGT TGATATTTCTATATTAAGTATGGTTCCAAGAGAATATTTTCATCCTAA ACCTAAAGTGAATAGCTCACTTATCAGATTAAGTAGAAAAAAATCAAG AATATCACACAAAGATAAACAAAAGTATAATTATTTCGTTATGAAATG GGTTAACAAAGAATACAAGAAAATATTTACAAAAAATCAATTTAACAA TTCCTTAAAACATGCAGGAATTGACGATTTAAACAATATTAGCTTTGA ACAATTCTTATCTCTTTTCAATAGCTATAAATTATTTAATAAGTAACT GCAGGGCTCGATCGGTGTGAAAAGTCAGTATCCAGTCGTGTAGTTCTT ATTACCTGTCCCTGCACTAAAGACAAGTGAGATAGCAGTGGCAATCTG

	GCTTTGCAATCAATGTTTCCACTAAAGCGTTTAGCGTTACTGCGGCTA GAAGTCCTCCACCGAGGCTCCCCTGAATGGTGATATGGGGAATGGGAC TGGTCATCAGTCGTCGTTTTGCCCCCGGAGCATGACTAAAACCGATCG GCATTCCGATCACAAGAGCCGGCTGAATATGTTGTTGCTCTATCAGCT TACAGGCAGTGAGTAAAACAGAAGGGGCATAGCCGATCGCCAGCACAC ATCCTTGGGGAATCTGTTGTAACCGCTGTTGCCAATGGTCATGGTGCC AAAAAGCTTGCTCGGCTTCCCTAAGCCCTGTGATGTGAGGGTCGTCAA TCAGCGTTTTAACCGTACATCCTAAATGAGCTAACCGAGTTTGATCAA GAGCCGCAGCCACAACCGGAACATCGGTGACGACTGGACACCCTGCTT TCAGTGCATCTCGTGCCGAGGCGATCGCTCCCTGACTCAATCGAACGG CGTTTACCAAGCTAACATCACCACAGGCCAGCACTAATTGATGTAGTA AGTGAATGGTAATTTCAGAGTAAGCCGATAAATCCGGTAGCAGGTGTT TGAGGGATTCCTGAAAGGCTTCTGGATGAGTTGTTGTCTCCGCATCTA GGTTCGTCCACAACTGATCGAGTTTTCCTAACCCCTCCTGGACATCCA CATCAAGCTGTTTCAGTTGGGCCAGAGCTTCCGCTTGGGTAATCTGGC AACTCTGGTCGCGTCCCAGTAATCCTTCTAAAGCAGATGCGGTTTGGC GGAGTCGAGTAATCTGCTGAATCACAGCCTGATATTGCTGTTGCAACT GCACCATTAGGGTGGGATCAAGGCTCTCTTCAGAATGGCTATCCAGCA GTTGCCGAATATGAGACAACTGAAAGCCCTGCTGTTTGAGGGCAATGA CTCGTTGGAGCCGTTGTACGTCCTGCTGAGTATAAAGGCGGTAGTTGC CCTCTGAGCGTTGAACGGGGGGAAGCAATCCCAGGGTGTGGTAATGGC GCACCATGCGAGGCGTAACGCCACCTCCCACTGCATCTGTGAGTTCTT TAATCGTTAAGTGATTAGTCTTCATCCCTTTAGTTTACTCAAAACCTT GACATTGACACTAATGTTAAGGTTTAGGCTGAGAAGGTAAAAATCCAA GTTAAAAAGCGGCTCGGGAGACCTATCGGTAATAACAGTCCAATCTGG TGTAACTTCGGAATCGTCCTTTACAGCTAGCTCAGTCCTAGGTATTAT GCTAGCTACTAGAGTCACACAGGAAAGTACTAGATGCGTAAAGGCGAA GAACTGTTCACGGGCGTAGTTCCGATTCTGGTCGAGCTGGACGGCGAT GTGAACGGTCATAAGTTTAGCGTTCGCGGTGAAGGTGAGGGCGACGCG ACCAACGGCAAACTGACCCTGAAGTTCATCTGCACCACCGGTAAACTG CCGGTGCCTTGGCCGACCTTGGTGACGACGTTGACGTATGGCGTGCAG TGTTTTGCGCGTTATCCGGACCACATGAAACAACACGATTTCTTCAAA TCTGCGATGCCGGAGGGTTACGTCCAGGAGCGTACCATTTCCTTCAAG GATGATGGCACTTACAAAACTCGCGCAGAGGTTAAGTTTGAAGGTGAC ACGCTGGTCAATCGTATCGAATTGAAGGGTATCGACTTTAAAGAGGAT GGTAACATTCTGGGCCATAAACTGGAGTATAACTTCAACAGCCATAAT GTTTACATTACGGCAGACAAGCAAAAGAACGGCATCAAGGCCAATTTC AAGATTCGCCACAATGTTGAGGACGGTAGCGTCCAACTGGCCGACCAT TACCAGCAGAACACCCCAATTGGTGACGGTCCGGTTTTGCTGCCGGAT AATCACTATCTGAGCACCCAAAGCGTGCTGAGCAAAGATCCGAACGAA AAACGTGATCACATGGTCCTGCTGGAATTTGTGACCGCTGCGGGCATC ACCCACGGTATGGACGAGCTGTATAAGCGTCCGTAATAATACTAGAGC CAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCG TTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGG CTCACCTTCGGGTGGGCCTTTCTGCGTTTATAGGCTCGGGTAAGAACT CGCACTTCGTGGAAACACTATTATCTGGTGGGTCTCT

Appendix G: List of gene sequences used in this work

Gene	Sequence
Sfp	ATGAAGATCTACGGCATATACATGGACCGCCCCCTAAGCCAGGAAGAGAACGAGC GCTTTATGAGCTTCATCTCCCCGGAGAAGCGTGAAAAATGTCGCCGTTTCTACCA CAAAGAAGATGCCCATCGCACTTTACTCGGTGATGTTTTGGTCCGATCTGTCATT AGCCGTCAATACCAACTCGATAAAAGCGACATTCGTTTCTCGACTCAAGAATACG GCAAACCCTGCATCCCAGACCTACCGGATGCCCACTTTAACATTTCCCACAGTGG CCGGTGGGTGATCTGTGCCTTTGACTCCCAACCGATTGGAATTGACATTGAAAAG ACCAAACCCATTAGTTTAGAAATAGCTAAACGTTTCTTCTCTAAGACAGAATATA GCGATCTCCTGGCCAAAGACAAGGACGAGCAAACTGACTACTTCTATCATTTGTG GAGCATGAAAGAATCCTTTATCAAACAAGAGGGCAAAGGTCTTTCACTGCCTTTA GACTCTTTTTCTGTCCGTTTACATCAGGACGGCCAGGTCAGTATTGAGCTTCCGG ATTCGCATTCCCCCTGTTATATCAAAACGTACGAAGTAGATCCGGGCTACAAAAT GGCGGTCTGCGCAGCGCACCCCGATTTCCCGGAAGATATTACGATGGTTTCTTAC GAAGAACTACTCTAA
CAR	ATGAGCCCGATTACCCGTGAAGAACGTCTGGAACGTCGTATTCAGGATCTGTATG CGAACGATCCGCAGTTCGCAGCAGCCAAACCGGCGACCGCGATTACCGCGGCGAT TGAACGTCCGGGTCTGCCGCTGCCGCAGATCATCGAAACGGTGATGACCGGCTAT GCGGATCGTCCGGCACTGGCACAACGTAGCGTGGAATTTGTGACCGATGCGGGCA CCGGTCATACCACCCTGCGTCTGCTGCCGCATTTTGAAACCATTAGCTATGGCGA ACTGTGGGATCGTATTAGCGCGCTGGCCGATGTTCTGAGCACCGAACAGACCGTG AAACCGGGCGATCGTGTGTGCCTGCTGGGCTTTAACAGCGTGGATTATGCGACCA TTGATATGACCCTGGCACGTCTGGGTGCTGTCGCTGTCCCGCTGCAGACCTCTGC TGCGATTACCCAGCTGCAGCCGATTGTGGCGGAAACCCAGCCGACCATGATTGCG GCGAGCGTGGATGCCCTGGCCGATGCGACCGAACTGGCACTGAGTGGTCAAACGG CTACGCGTGTGCTGGTGTTTGATCATCATCGTCAGGTGGATGCGCATCGTGCGGC GGTTGAAAGCGCGCGTGAACGTCTGGCCGGTAGCGCGGTGGTTGAAACCCTGGCC GAAGCGATTGCGCGTGGTGATGTGCCGCGTGGTGCGAGCGCGGGTAGCGCACCGG GCACCGATGTGAGCGATGATAGCCTGGCCCTGCTGATTTATACCTCTGGTAGTAC GGGTGCGCCGAAAGGCGCCATGTATCCGCGTCGTAACGTGGCGACCTTTTGGCGT AAACGTACCTGGTTTGAAGGCGGCTATGAACCGAGCATTACCCTGAACTTTATGC CGATGAGCCATGTGATGGGCCGTCAGATTCTGTATGGCACCCTGTGCAACGGCGG CACCGCGTATTTTGTGGCGAAAAGCGATCTGAGCACCCTGTTTGAAGATCTGGCC CTGGTGCGTCCGACCGAACTGACCTTCGTCCCGCGTGTTTGGGATATGGTGTTCG ATGAATTTCAGAGCGAAGTGGATCGTCGTCTGGTGGATGGCGCGGATCGTGTTGC GCTGGAAGCGCAGGTGAAAGCGGAAATTCGTAACGATGTGCTGGGCGGTCGTTAT ACCTCTGCTCTGACGGGTTCTGCTCCGATTAGCGATGAAATGAAAGCGTGGGTGG AAGAACTGCTGGATATGCATCTGGTGGAAGGCTATGGCAGCACCGAAGCGGGCAT GATTCTGATTGATGGCGCGATTCGTCGTCCGGCGGTGCTGGATTATAAACTGGTG GATGTTCCGGATCTGGGCTATTTTCTGACCGATCGTCCGCATCCGCGTGGCGAAC TGCTGGTGAAAACCGATAGCCTGTTTCCGGGCTATTATCAGCGTGCGGAAGTGAC CGCGGATGTGTTTGATGCGGATGGCTTTTATCGCACCGGCGATATTATGGCGGAA GTGGGCCCGGAACAGTTTGTGTATCTGGATCGTCGTAACAACGTGCTGAAACTGA GCCAGGGCGAATTTGTTACCGTGAGCAAACTGGAAGCGGTGTTTGGCGATAGCCC GCTGGTGCGTCAGATTTATATTTATGGCAACAGCGCGCGTGCGTATCTGCTGGCC GTGATTGTGCCGACCCAGGAAGCGCTGGACGCGGTCCCGGTTGAAGAACTGAAAG CGCGTCTGGGTGACTCTCTGCAGGAAGTGGCGAAAGCGGCGGGTCTGCAGAGCTA TGAAATTCCGCGCGATTTTATTATCGAAACCACCCCGTGGACCCTGGAAAACGGC CTGCTGACGGGTATTCGTAAACTGGCCCGTCCGCAGCTGAAAAAACATTATGGTG AACTGCTGGAACAAATTTATACCGATCTGGCCCACGGCCAGGCGGATGAACTGCG TAGCCTGCGTCAGAGCGGTGCGGATGCGCCGGTGCTGGTGACCGTTTGTCGTGCG GCTGCAGCTCTGCTGGGTGGTAGCGCGAGCGATGTGCAGCCGGATGCGCATTTCA CCGATCTGGGTGGTGATAGCCTGAGCGCCCTGAGCTTTACCAACCTGCTGCATGA AATCTTTGATATTGAAGTGCCGGTGGGCGTGATTGTGAGCCCGGCGAACGATCTG CAGGCGCTGGCCGATTATGTGGAAGCGGCGCGTAAACCGGGTAGCAGCCGTCCGA

	CCTTTGCGAGCGTGCATGGCGCGAGCAACGGCCAGGTGACCGAAGTGCATGCGGG CGATCTGAGCCTGGATAAATTTATTGATGCGGCGACCCTGGCCGAAGCCCCGCGT CTGCCGGCTGCAAATACCCAGGTGCGTACCGTGCTGCTGACCGGTGCGACCGGCT TTCTGGGCCGTTACCTGGCCCTGGAATGGCTGGAACGTATGGATCTGGTTGATGG CAAACTGATTTGCCTGGTGCGTGCCAAAAGCGATACCGAAGCGCGTGCGCGTCTG GATAAAACCTTTGATAGCGGCGATCCGGAACTGCTGGCCCATTATCGTGCGCTGG CCGGCGATCATCTGGAAGTGCTGGCCGGTGATAAAGGCGAAGCGGATCTGGGCCT GGATCGTCAGACCTGGCAACGCCTGGCAGATACCGTGGATCTGATTGTTGACCCG GCTGCCCTGGTGAATCATGTGCTGCCGTATAGCCAGCTGTTTGGCCCGAATGCGC TGGGCACCGCTGAACTGCTGCGCCTGGCTCTGACCAGCAAAATTAAACCGTATAG CTACACCAGCACCATTGGCGTGGCGGATCAGATTCCGCCGAGCGCGTTTACCGAA GATGCGGATATTCGTGTGATTAGCGCGACCCGTGCGGTGGATGATAGCTATGCGA ACGGCTATAGCAACAGCAAATGGGCGGGTGAAGTGCTGCTGCGTGAAGCGCATGA TCTGTGCGGTCTGCCGGTGGCGGTGTTTCGTTGCGATATGATCCTGGCAGACACG ACCTGGGCGGGTCAGCTGAACGTGCCGGATATGTTTACCCGTATGATTCTGTCTC TGGCAGCTACGGGTATCGCACCGGGTAGCTTTTATGAACTGGCCGCGGATGGTGC GCGTCAGCGTGCGCATTATGATGGCCTGCCGGTGGAATTTATTGCGGAAGCGATT AGCACCCTGGGCGCGCAGAGCCAGGATGGCTTTCATACCTATCATGTGATGAATC CGTATGATGATGGCATTGGCCTGGATGAATTTGTGGATTGGCTGAACGAAAGCGG CTGCCCGATTCAGCGTATTGCGGATTATGGCGATTGGCTGCAGCGTTTTGAAACC GCGCTGCGCGCTCTGCCGGATCGTCAGCGTCATAGCAGCCTGCTGCCGCTGCTGC ATAACTATCGTCAGCCGGAACGTCCGGTGCGTGGTAGCATTGCGCCGACCGATCG CTTTCGTGCGGCCGTGCAGGAAGCGAAAATTGGCCCGGATAAAGATATTCCGCAT GTGGGTGCGCCGATTATTGTGAAATATGTGAGCGATCTGCGCCTGCTGGGCCTGC TGTAA
CvFatB1 (Wild-Type)	ATGGTGGCCGCAGCAGCAACGAGTGCTTTTTTTCCAGTGCCAGCGCCAGGCACCA GCCCCAAGCCTGGAAAAAGTGGGAATTGGCCCAGCAGCCTCTCTCCGACTTTCAA GCCAAAAAGTATCCCTAATGGCGGGTTTCAAGTGAAGGCGAATGCTTCTGCACAC CCCAAAGCCAACGGTAGTGCAGTGAACCTCAAGAGCGGCTCCCTCAACACGCAGG AAGATACTTCCAGCAGCCCGCCTCCCCGAGCGTTTTTGAATCAATTGCCCGACTG GTCCATGCTACTAACCGCTATCACAACGGTGTTTGTGGCAGCAGAAAAACAGTGG ACCATGCTCGACCGGAAATCCAAGCGACCTGACATGCTAGTAGATTCCGTTGGAC TGAAATCCATCGTGCGTGATGGCTTGGTCAGCCGCCACAGCTTCAGCATTCGATC CTACGAGATCGGAGCTGACCGGACCGCCAGTATTGAGACACTGATGAACCATTTA CAAGAGACAACTATCAATCACTGTAAGAGCTTGGGATTACACAATGACGGATTCG GCCGCACACCCGGCATGTGTAAGAATGACCTCATCTGGGTATTGACTAAGATGCA AATTATGGTTAATCGCTATCCTACCTGGGGCGATACTGTAGAGATCAACACCTGG TTTAGTCAAAGTGGCAAGATCGGTATGGCGTCCGACTGGCTAATCTCTGATTGCA ACACCGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGGCCATGATGAACCAGAA GACCCGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCAC TTCGTCGACAGTCCCCATGTTATCGAGGATAATGACCAGAAGTTGCGGAAATTCG ATGTCAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCGCGGTGGAATGATCT CGACGTAAATCAACATGTTTCTAACGTCAAGTACATCGGATGGATTTTGGAATCC ATGCCGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCTAACAGTTGAGTACC GACGCGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTAG TGAAAATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGGAAGATGGGACA GATATCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGGTACAAATGGGG CTATCTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTAA
CvT1	ATGGCCAACGGTAGTGCAGTGAACCTCAAGAGCGGCTCCCTCAACACGCAGGAAG ATACTTCCAGCAGCCCGCCTCCCCGAGCGTTTTTGAATCAATTGCCCGACTGGTC CATGCTACTAACCGCTATCACAACGGTGTTTGTGGCAGCAGAAAAACAGTGGACC ATGCTCGACCGGAAATCCAAGCGACCTGACATGCTAGTAGATTCCGTTGGACTGA AATCCATCGTGCGTGATGGCTTGGTCAGCCGCCACAGCTTCAGCATTCGATCCTA CGAGATCGGAGCTGACCGGACCGCCAGTATTGAGACACTGATGAACCATTTACAA GAGACAACTATCAATCACTGTAAGAGCTTGGGATTACACAATGACGGATTCGGCC GCACACCCGGCATGTGTAAGAATGACCTCATCTGGGTATTGACTAAGATGCAAAT TATGGTTAATCGCTATCCTACCTGGGGCGATACTGTAGAGATCAACACCTGGTTT AGTCAAAGTGGCAAGATCGGTATGGCGTCCGACTGGCTAATCTCTGATTGCAACA CCGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGGCCATGATGAACCAGAAGAC CCGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCACTTC GTCGACAGTCCCCATGTTATCGAGGATAATGACCAGAAGTTGCGGAAATTCGATG

	TCAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCGCGGTGGAATGATCTCGA CGTAAATCAACATGTTTCTAACGTCAAGTACATCGGATGGATTTTGGAATCCATG CCGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCTAACAGTTGAGTACCGAC GCGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTAGTGA AAATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGGAAGATGGGACAGAT ATCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGGTACAAATGGGGCTA TCTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTAA
CvT2	ATGTTGCCCGACTGGTCCATGCTACTAACCGCTATCACAACGGTGTTTGTGGCAG CAGAAAAACAGTGGACCATGCTCGACCGGAAATCCAAGCGACCTGACATGCTAGT AGATTCCGTTGGACTGAAATCCATCGTGCGTGATGGCTTGGTCAGCCGCCACAGC TTCAGCATTCGATCCTACGAGATCGGAGCTGACCGGACCGCCAGTATTGAGACAC TGATGAACCATTTACAAGAGACAACTATCAATCACTGTAAGAGCTTGGGATTACA CAATGACGGATTCGGCCGCACACCCGGCATGTGTAAGAATGACCTCATCTGGGTA TTGACTAAGATGCAAATTATGGTTAATCGCTATCCTACCTGGGGCGATACTGTAG AGATCAACACCTGGTTTAGTCAAAGTGGCAAGATCGGTATGGCGTCCGACTGGCT AATCTCTGATTGCAACACCGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGGCC ATGATGAACCAGAAGACCCGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAG AACTCACCCCGCACTTCGTCGACAGTCCCCATGTTATCGAGGATAATGACCAGAA GTTGCGGAAATTCGATGTCAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCG CGGTGGAATGATCTCGACGTAAATCAACATGTTTCTAACGTCAAGTACATCGGAT GGATTTTGGAATCCATGCCGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCT AACAGTTGAGTACCGACGCGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACA GCCGTGGATCCTAGTGAAAATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTT TGGAAGATGGGACAGATATCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGC GGGTACAAATGGGGCTATCTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGT GTTAGTTAA
CvT3	ATGCTACTAACCGCTATCACAACGGTGTTTGTGGCAGCAGAAAAACAGTGGACCA TGCTCGACCGGAAATCCAAGCGACCTGACATGCTAGTAGATTCCGTTGGACTGAA ATCCATCGTGCGTGATGGCTTGGTCAGCCGCCACAGCTTCAGCATTCGATCCTAC GAGATCGGAGCTGACCGGACCGCCAGTATTGAGACACTGATGAACCATTTACAAG AGACAACTATCAATCACTGTAAGAGCTTGGGATTACACAATGACGGATTCGGCCG CACACCCGGCATGTGTAAGAATGACCTCATCTGGGTATTGACTAAGATGCAAATT ATGGTTAATCGCTATCCTACCTGGGGCGATACTGTAGAGATCAACACCTGGTTTA GTCAAAGTGGCAAGATCGGTATGGCGTCCGACTGGCTAATCTCTGATTGCAACAC CGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGGCCATGATGAACCAGAAGACC CGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCACTTCG TCGACAGTCCCCATGTTATCGAGGATAATGACCAGAAGTTGCGGAAATTCGATGT CAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCGCGGTGGAATGATCTCGAC GTAAATCAACATGTTTCTAACGTCAAGTACATCGGATGGATTTTGGAATCCATGC CGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCTAACAGTTGAGTACCGACG CGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTAGTGAA AATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGGAAGATGGGACAGATA TCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGGTACAAATGGGGCTAT CTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTAA
CvT4	ATGGAAAAACAGTGGACCATGCTCGACCGGAAATCCAAGCGACCTGACATGCTAG TAGATTCCGTTGGACTGAAATCCATCGTGCGTGATGGCTTGGTCAGCCGCCACAG CTTCAGCATTCGATCCTACGAGATCGGAGCTGACCGGACCGCCAGTATTGAGACA CTGATGAACCATTTACAAGAGACAACTATCAATCACTGTAAGAGCTTGGGATTAC ACAATGACGGATTCGGCCGCACACCCGGCATGTGTAAGAATGACCTCATCTGGGT ATTGACTAAGATGCAAATTATGGTTAATCGCTATCCTACCTGGGGCGATACTGTA GAGATCAACACCTGGTTTAGTCAAAGTGGCAAGATCGGTATGGCGTCCGACTGGC TAATCTCTGATTGCAACACCGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGGC CATGATGAACCAGAAGACCCGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAA GAACTCACCCCGCACTTCGTCGACAGTCCCCATGTTATCGAGGATAATGACCAGA AGTTGCGGAAATTCGATGTCAAGACAGGAGACAGTATCCGCAAGGGGCTCACACC GCGGTGGAATGATCTCGACGTAAATCAACATGTTTCTAACGTCAAGTACATCGGA TGGATTTTGGAATCCATGCCGATTGAAGTCCTAGAGACACAAGAATTATGCAGTC TAACAGTTGAGTACCGACGCGAATGTGGCATGGACTCCGTATTAGAGTCTGTAAC AGCCGTGGATCCTAGTGAAAATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGT TTGGAAGATGGGACAGATATCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATG

	CGGGTACAAATGGGGCTATCTCTACCAGCACGGCGAAGACAAGCAATGGGAACAG TGTTAGTTAA
CvT5	ATGCTCGACCGGAAATCCAAGCGACCTGACATGCTAGTAGATTCCGTTGGACTGA AATCCATCGTGCGTGATGGCTTGGTCAGCCGCCACAGCTTCAGCATTCGATCCTA CGAGATCGGAGCTGACCGGACCGCCAGTATTGAGACACTGATGAACCATTTACAA GAGACAACTATCAATCACTGTAAGAGCTTGGGATTACACAATGACGGATTCGGCC GCACACCCGGCATGTGTAAGAATGACCTCATCTGGGTATTGACTAAGATGCAAAT TATGGTTAATCGCTATCCTACCTGGGGCGATACTGTAGAGATCAACACCTGGTTT AGTCAAAGTGGCAAGATCGGTATGGCGTCCGACTGGCTAATCTCTGATTGCAACA CCGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGGCCATGATGAACCAGAAGAC CCGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCACTTC GTCGACAGTCCCCATGTTATCGAGGATAATGACCAGAAGTTGCGGAAATTCGATG TCAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCGCGGTGGAATGATCTCGA CGTAAATCAACATGTTTCTAACGTCAAGTACATCGGATGGATTTTGGAATCCATG CCGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCTAACAGTTGAGTACCGAC GCGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTAGTGA AAATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGGAAGATGGGACAGAT ATCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGGTACAAATGGGGCTA TCTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTAA
CvT6	ATGCGACCTGACATGCTAGTAGATTCCGTTGGACTGAAATCCATCGTGCGTGATG GCTTGGTCAGCCGCCACAGCTTCAGCATTCGATCCTACGAGATCGGAGCTGACCG GACCGCCAGTATTGAGACACTGATGAACCATTTACAAGAGACAACTATCAATCAC TGTAAGAGCTTGGGATTACACAATGACGGATTCGGCCGCACACCCGGCATGTGTA AGAATGACCTCATCTGGGTATTGACTAAGATGCAAATTATGGTTAATCGCTATCC TACCTGGGGCGATACTGTAGAGATCAACACCTGGTTTAGTCAAAGTGGCAAGATC GGTATGGCGTCCGACTGGCTAATCTCTGATTGCAACACCGGGGAAATCCTGATCC GAGCAACTTCTGTGTGGGCCATGATGAACCAGAAGACCCGGCGCTTCTCTCGGCT CCCGTATGAAGTTCGTCAAGAACTCACCCCGCACTTCGTCGACAGTCCCCATGTT ATCGAGGATAATGACCAGAAGTTGCGGAAATTCGATGTCAAGACAGGAGACAGTA TCCGCAAGGGGCTCACACCGCGGTGGAATGATCTCGACGTAAATCAACATGTTTC TAACGTCAAGTACATCGGATGGATTTTGGAATCCATGCCGATTGAAGTCCTAGAG ACACAAGAATTATGCAGTCTAACAGTTGAGTACCGACGCGAATGTGGCATGGACT CCGTATTAGAGTCTGTAACAGCCGTGGATCCTAGTGAAAATGGTGGGCGCAGTCA GTATAAACATCTCCTCCGTTTGGAAGATGGGACAGATATCGTTAAAAGCCGGACT GAATGGCGCCCAAAGAATGCGGGTACAAATGGGGCTATCTCTACCAGCACGGCGA AGACAAGCAATGGGAACAGTGTTAGTTAA
CvT7	ATGAAATCCATCGTGCGTGATGGCTTGGTCAGCCGCCACAGCTTCAGCATTCGAT CCTACGAGATCGGAGCTGACCGGACCGCCAGTATTGAGACACTGATGAACCATTT ACAAGAGACAACTATCAATCACTGTAAGAGCTTGGGATTACACAATGACGGATTC GGCCGCACACCCGGCATGTGTAAGAATGACCTCATCTGGGTATTGACTAAGATGC AAATTATGGTTAATCGCTATCCTACCTGGGGCGATACTGTAGAGATCAACACCTG GTTTAGTCAAAGTGGCAAGATCGGTATGGCGTCCGACTGGCTAATCTCTGATTGC AACACCGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGGCCATGATGAACCAGA AGACCCGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCA CTTCGTCGACAGTCCCCATGTTATCGAGGATAATGACCAGAAGTTGCGGAAATTC GATGTCAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCGCGGTGGAATGATC TCGACGTAAATCAACATGTTTCTAACGTCAAGTACATCGGATGGATTTTGGAATC CATGCCGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCTAACAGTTGAGTAC CGACGCGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTA GTGAAAATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGGAAGATGGGAC AGATATCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGGTACAAATGGG GCTATCTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTAA
CvT8	ATGCGCCACAGCTTCAGCATTCGATCCTACGAGATCGGAGCTGACCGGACCGCCA GTATTGAGACACTGATGAACCATTTACAAGAGACAACTATCAATCACTGTAAGAG CTTGGGATTACACAATGACGGATTCGGCCGCACACCCGGCATGTGTAAGAATGAC CTCATCTGGGTATTGACTAAGATGCAAATTATGGTTAATCGCTATCCTACCTGGG GCGATACTGTAGAGATCAACACCTGGTTTAGTCAAAGTGGCAAGATCGGTATGGC GTCCGACTGGCTAATCTCTGATTGCAACACCGGGGAAATCCTGATCCGAGCAACT TCTGTGTGGGCCATGATGAACCAGAAGACCCGGCGCTTCTCTCGGCTCCCGTATG AAGTTCGTCAAGAACTCACCCCGCACTTCGTCGACAGTCCCCATGTTATCGAGGA TAATGACCAGAAGTTGCGGAAATTCGATGTCAAGACAGGAGACAGTATCCGCAAG

	GGGCTCACACCGCGGTGGAATGATCTCGACGTAAATCAACATGTTTCTAACGTCA AGTACATCGGATGGATTTTGGAATCCATGCCGATTGAAGTCCTAGAGACACAAGA ATTATGCAGTCTAACAGTTGAGTACCGACGCGAATGTGGCATGGACTCCGTATTA GAGTCTGTAACAGCCGTGGATCCTAGTGAAAATGGTGGGCGCAGTCAGTATAAAC ATCTCCTCCGTTTGGAAGATGGGACAGATATCGTTAAAAGCCGGACTGAATGGCG CCCAAAGAATGCGGGTACAAATGGGGCTATCTCTACCAGCACGGCGAAGACAAGC AATGGGAACAGTGTTAGTTAA
CvT9	ATGATCGGAGCTGACCGGACCGCCAGTATTGAGACACTGATGAACCATTTACAAG AGACAACTATCAATCACTGTAAGAGCTTGGGATTACACAATGACGGATTCGGCCG CACACCCGGCATGTGTAAGAATGACCTCATCTGGGTATTGACTAAGATGCAAATT ATGGTTAATCGCTATCCTACCTGGGGCGATACTGTAGAGATCAACACCTGGTTTA GTCAAAGTGGCAAGATCGGTATGGCGTCCGACTGGCTAATCTCTGATTGCAACAC CGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGGCCATGATGAACCAGAAGACC CGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCACTTCG TCGACAGTCCCCATGTTATCGAGGATAATGACCAGAAGTTGCGGAAATTCGATGT CAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCGCGGTGGAATGATCTCGAC GTAAATCAACATGTTTCTAACGTCAAGTACATCGGATGGATTTTGGAATCCATGC CGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCTAACAGTTGAGTACCGACG CGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTAGTGAA AATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGGAAGATGGGACAGATA TCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGGTACAAATGGGGCTAT CTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTAA
CvT10	ATGACACTGATGAACCATTTACAAGAGACAACTATCAATCACTGTAAGAGCTTGG GATTACACAATGACGGATTCGGCCGCACACCCGGCATGTGTAAGAATGACCTCAT CTGGGTATTGACTAAGATGCAAATTATGGTTAATCGCTATCCTACCTGGGGCGAT ACTGTAGAGATCAACACCTGGTTTAGTCAAAGTGGCAAGATCGGTATGGCGTCCG ACTGGCTAATCTCTGATTGCAACACCGGGGAAATCCTGATCCGAGCAACTTCTGT GTGGGCCATGATGAACCAGAAGACCCGGCGCTTCTCTCGGCTCCCGTATGAAGTT CGTCAAGAACTCACCCCGCACTTCGTCGACAGTCCCCATGTTATCGAGGATAATG ACCAGAAGTTGCGGAAATTCGATGTCAAGACAGGAGACAGTATCCGCAAGGGGCT CACACCGCGGTGGAATGATCTCGACGTAAATCAACATGTTTCTAACGTCAAGTAC ATCGGATGGATTTTGGAATCCATGCCGATTGAAGTCCTAGAGACACAAGAATTAT GCAGTCTAACAGTTGAGTACCGACGCGAATGTGGCATGGACTCCGTATTAGAGTC TGTAACAGCCGTGGATCCTAGTGAAAATGGTGGGCGCAGTCAGTATAAACATCTC CTCCGTTTGGAAGATGGGACAGATATCGTTAAAAGCCGGACTGAATGGCGCCCAA AGAATGCGGGTACAAATGGGGCTATCTCTACCAGCACGGCGAAGACAAGCAATGG GAACAGTGTTAGTTAA
CvT11	ATGATCAATCACTGTAAGAGCTTGGGATTACACAATGACGGATTCGGCCGCACAC CCGGCATGTGTAAGAATGACCTCATCTGGGTATTGACTAAGATGCAAATTATGGT TAATCGCTATCCTACCTGGGGCGATACTGTAGAGATCAACACCTGGTTTAGTCAA AGTGGCAAGATCGGTATGGCGTCCGACTGGCTAATCTCTGATTGCAACACCGGGG AAATCCTGATCCGAGCAACTTCTGTGTGGGCCATGATGAACCAGAAGACCCGGCG CTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCACTTCGTCGAC AGTCCCCATGTTATCGAGGATAATGACCAGAAGTTGCGGAAATTCGATGTCAAGA CAGGAGACAGTATCCGCAAGGGGCTCACACCGCGGTGGAATGATCTCGACGTAAA TCAACATGTTTCTAACGTCAAGTACATCGGATGGATTTTGGAATCCATGCCGATT GAAGTCCTAGAGACACAAGAATTATGCAGTCTAACAGTTGAGTACCGACGCGAAT GTGGCATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTAGTGAAAATGG TGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGGAAGATGGGACAGATATCGTT AAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGGTACAAATGGGGCTATCTCTA CCAGCACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTAA
CvT12	ATGAATGACGGATTCGGCCGCACACCCGGCATGTGTAAGAATGACCTCATCTGGG TATTGACTAAGATGCAAATTATGGTTAATCGCTATCCTACCTGGGGCGATACTGT AGAGATCAACACCTGGTTTAGTCAAAGTGGCAAGATCGGTATGGCGTCCGACTGG CTAATCTCTGATTGCAACACCGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGG CCATGATGAACCAGAAGACCCGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCA AGAACTCACCCCGCACTTCGTCGACAGTCCCCATGTTATCGAGGATAATGACCAG AAGTTGCGGAAATTCGATGTCAAGACAGGAGACAGTATCCGCAAGGGGCTCACAC CGCGGTGGAATGATCTCGACGTAAATCAACATGTTTCTAACGTCAAGTACATCGG ATGGATTTTGGAATCCATGCCGATTGAAGTCCTAGAGACACAAGAATTATGCAGT CTAACAGTTGAGTACCGACGCGAATGTGGCATGGACTCCGTATTAGAGTCTGTAA

	CAGCCGTGGATCCTAGTGAAAATGGTGGGCGCAGTCAGTATAAACATCTCCTCCG TTTGGAAGATGGGACAGATATCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAAT GCGGGTACAAATGGGGCTATCTCTACCAGCACGGCGAAGACAAGCAATGGGAACA GTGTTAGTTAA
CvT13	ATGTGTAAGAATGACCTCATCTGGGTATTGACTAAGATGCAAATTATGGTTAATC GCTATCCTACCTGGGGCGATACTGTAGAGATCAACACCTGGTTTAGTCAAAGTGG CAAGATCGGTATGGCGTCCGACTGGCTAATCTCTGATTGCAACACCGGGGAAATC CTGATCCGAGCAACTTCTGTGTGGGCCATGATGAACCAGAAGACCCGGCGCTTCT CTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCACTTCGTCGACAGTCC CCATGTTATCGAGGATAATGACCAGAAGTTGCGGAAATTCGATGTCAAGACAGGA GACAGTATCCGCAAGGGGCTCACACCGCGGTGGAATGATCTCGACGTAAATCAAC ATGTTTCTAACGTCAAGTACATCGGATGGATTTTGGAATCCATGCCGATTGAAGT CCTAGAGACACAAGAATTATGCAGTCTAACAGTTGAGTACCGACGCGAATGTGGC ATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTAGTGAAAATGGTGGGC GCAGTCAGTATAAACATCTCCTCCGTTTGGAAGATGGGACAGATATCGTTAAAAG CCGGACTGAATGGCGCCCAAAGAATGCGGGTACAAATGGGGCTATCTCTACCAGC ACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTAA
CvT14	ATGAAGATGCAAATTATGGTTAATCGCTATCCTACCTGGGGCGATACTGTAGAGA TCAACACCTGGTTTAGTCAAAGTGGCAAGATCGGTATGGCGTCCGACTGGCTAAT CTCTGATTGCAACACCGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGGCCATG ATGAACCAGAAGACCCGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAAC TCACCCCGCACTTCGTCGACAGTCCCCATGTTATCGAGGATAATGACCAGAAGTT GCGGAAATTCGATGTCAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCGCGG TGGAATGATCTCGACGTAAATCAACATGTTTCTAACGTCAAGTACATCGGATGGA TTTTGGAATCCATGCCGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCTAAC AGTTGAGTACCGACGCGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACAGCC GTGGATCCTAGTGAAAATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGG AAGATGGGACAGATATCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGG TACAAATGGGGCTATCTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGTGTT AGTTAA
CvT15	ATGACCTGGGGCGATACTGTAGAGATCAACACCTGGTTTAGTCAAAGTGGCAAGA TCGGTATGGCGTCCGACTGGCTAATCTCTGATTGCAACACCGGGGAAATCCTGAT CCGAGCAACTTCTGTGTGGGCCATGATGAACCAGAAGACCCGGCGCTTCTCTCGG CTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCACTTCGTCGACAGTCCCCATG TTATCGAGGATAATGACCAGAAGTTGCGGAAATTCGATGTCAAGACAGGAGACAG TATCCGCAAGGGGCTCACACCGCGGTGGAATGATCTCGACGTAAATCAACATGTT TCTAACGTCAAGTACATCGGATGGATTTTGGAATCCATGCCGATTGAAGTCCTAG AGACACAAGAATTATGCAGTCTAACAGTTGAGTACCGACGCGAATGTGGCATGGA CTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTAGTGAAAATGGTGGGCGCAGT CAGTATAAACATCTCCTCCGTTTGGAAGATGGGACAGATATCGTTAAAAGCCGGA CTGAATGGCGCCCAAAGAATGCGGGTACAAATGGGGCTATCTCTACCAGCACGGC GAAGACAAGCAATGGGAACAGTGTTAGTTAA
CvT16	ATGTGGTTTAGTCAAAGTGGCAAGATCGGTATGGCGTCCGACTGGCTAATCTCTG ATTGCAACACCGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGGCCATGATGAA CCAGAAGACCCGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACC CCGCACTTCGTCGACAGTCCCCATGTTATCGAGGATAATGACCAGAAGTTGCGGA AATTCGATGTCAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCGCGGTGGAA TGATCTCGACGTAAATCAACATGTTTCTAACGTCAAGTACATCGGATGGATTTTG GAATCCATGCCGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCTAACAGTTG AGTACCGACGCGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGA TCCTAGTGAAAATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGGAAGAT GGGACAGATATCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGGTACAA ATGGGGCTATCTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTA A
CvT17	ATGGCGTCCGACTGGCTAATCTCTGATTGCAACACCGGGGAAATCCTGATCCGAG CAACTTCTGTGTGGGCCATGATGAACCAGAAGACCCGGCGCTTCTCTCGGCTCCC GTATGAAGTTCGTCAAGAACTCACCCCGCACTTCGTCGACAGTCCCCATGTTATC GAGGATAATGACCAGAAGTTGCGGAAATTCGATGTCAAGACAGGAGACAGTATCC GCAAGGGGCTCACACCGCGGTGGAATGATCTCGACGTAAATCAACATGTTTCTAA CGTCAAGTACATCGGATGGATTTTGGAATCCATGCCGATTGAAGTCCTAGAGACA CAAGAATTATGCAGTCTAACAGTTGAGTACCGACGCGAATGTGGCATGGACTCCG

	TATTAGAGTCTGTAACAGCCGTGGATCCTAGTGAAAATGGTGGGCGCAGTCAGTA TAAACATCTCCTCCGTTTGGAAGATGGGACAGATATCGTTAAAAGCCGGACTGAA TGGCGCCCAAAGAATGCGGGTACAAATGGGGCTATCTCTACCAGCACGGCGAAGA CAAGCAATGGGAACAGTGTTAGTTAA
CvT18	ATGACCGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGGCCATGATGAACCAGA AGACCCGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCA CTTCGTCGACAGTCCCCATGTTATCGAGGATAATGACCAGAAGTTGCGGAAATTC GATGTCAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCGCGGTGGAATGATC TCGACGTAAATCAACATGTTTCTAACGTCAAGTACATCGGATGGATTTTGGAATC CATGCCGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCTAACAGTTGAGTAC CGACGCGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTA GTGAAAATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGGAAGATGGGAC AGATATCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGGTACAAATGGG GCTATCTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTAA
CvT19	ATGGTGTGGGCCATGATGAACCAGAAGACCCGGCGCTTCTCTCGGCTCCCGTATG AAGTTCGTCAAGAACTCACCCCGCACTTCGTCGACAGTCCCCATGTTATCGAGGA TAATGACCAGAAGTTGCGGAAATTCGATGTCAAGACAGGAGACAGTATCCGCAAG GGGCTCACACCGCGGTGGAATGATCTCGACGTAAATCAACATGTTTCTAACGTCA AGTACATCGGATGGATTTTGGAATCCATGCCGATTGAAGTCCTAGAGACACAAGA ATTATGCAGTCTAACAGTTGAGTACCGACGCGAATGTGGCATGGACTCCGTATTA GAGTCTGTAACAGCCGTGGATCCTAGTGAAAATGGTGGGCGCAGTCAGTATAAAC ATCTCCTCCGTTTGGAAGATGGGACAGATATCGTTAAAAGCCGGACTGAATGGCG CCCAAAGAATGCGGGTACAAATGGGGCTATCTCTACCAGCACGGCGAAGACAAGC AATGGGAACAGTGTTAGTTAA
CvT20	ATGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCACTTCG TCGACAGTCCCCATGTTATCGAGGATAATGACCAGAAGTTGCGGAAATTCGATGT CAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCGCGGTGGAATGATCTCGAC GTAAATCAACATGTTTCTAACGTCAAGTACATCGGATGGATTTTGGAATCCATGC CGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCTAACAGTTGAGTACCGACG CGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTAGTGAA AATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGGAAGATGGGACAGATA TCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGGTACAAATGGGGCTAT CTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTAA
CvT21	ATGCAAGAACTCACCCCGCACTTCGTCGACAGTCCCCATGTTATCGAGGATAATG ACCAGAAGTTGCGGAAATTCGATGTCAAGACAGGAGACAGTATCCGCAAGGGGCT CACACCGCGGTGGAATGATCTCGACGTAAATCAACATGTTTCTAACGTCAAGTAC ATCGGATGGATTTTGGAATCCATGCCGATTGAAGTCCTAGAGACACAAGAATTAT GCAGTCTAACAGTTGAGTACCGACGCGAATGTGGCATGGACTCCGTATTAGAGTC TGTAACAGCCGTGGATCCTAGTGAAAATGGTGGGCGCAGTCAGTATAAACATCTC CTCCGTTTGGAAGATGGGACAGATATCGTTAAAAGCCGGACTGAATGGCGCCCAA AGAATGCGGGTACAAATGGGGCTATCTCTACCAGCACGGCGAAGACAAGCAATGG GAACAGTGTTAGTTAA
CvT5 M90G	ATGCTCGACCGGAAATCCAAGCGACCTGACATGCTAGTAGATTCCGTTGGACTGA AATCCATCGTGCGTGATGGCTTGGTCAGCCGCCACAGCTTCAGCATTCGATCCTA CGAGATCGGAGCTGACCGGACCGCCAGTATTGAGACACTGATGAACCATTTACAA GAGACAACTATCAATCACTGTAAGAGCTTGGGATTACACAATGACGGATTCGGCC GCACACCCGGCATGTGTAAGAATGACCTCATCTGGGTATTGACTAAGGGCCAAAT TATGGTTAATCGCTATCCTACCTGGGGCGATACTGTAGAGATCAACACCTGGTTT AGTCAAAGTGGCAAGATCGGTATGGCGTCCGACTGGCTAATCTCTGATTGCAACA CCGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGGCCATGATGAACCAGAAGAC CCGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCACTTC GTCGACAGTCCCCATGTTATCGAGGATAATGACCAGAAGTTGCGGAAATTCGATG TCAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCGCGGTGGAATGATCTCGA CGTAAATCAACATGTTTCTAACGTCAAGTACATCGGATGGATTTTGGAATCCATG CCGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCTAACAGTTGAGTACCGAC GCGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTAGTGA AAATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGGAAGATGGGACAGAT ATCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGGTACAAATGGGGCTA TCTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTAA
CvT5 W122G	ATGCTCGACCGGAAATCCAAGCGACCTGACATGCTAGTAGATTCCGTTGGACTGA AATCCATCGTGCGTGATGGCTTGGTCAGCCGCCACAGCTTCAGCATTCGATCCTA

	CGAGATCGGAGCTGACCGGACCGCCAGTATTGAGACACTGATGAACCATTTACAA GAGACAACTATCAATCACTGTAAGAGCTTGGGATTACACAATGACGGATTCGGCC GCACACCCGGCATGTGTAAGAATGACCTCATCTGGGTATTGACTAAGATGCAAAT TATGGTTAATCGCTATCCTACCTGGGGCGATACTGTAGAGATCAACACCTGGTTT AGTCAAAGTGGCAAGATCGGTATGGCGTCCGACGGCCTAATCTCTGATTGCAACA CCGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGGCCATGATGAACCAGAAGAC CCGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCACTTC GTCGACAGTCCCCATGTTATCGAGGATAATGACCAGAAGTTGCGGAAATTCGATG TCAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCGCGGTGGAATGATCTCGA CGTAAATCAACATGTTTCTAACGTCAAGTACATCGGATGGATTTTGGAATCCATG CCGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCTAACAGTTGAGTACCGAC GCGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTAGTGA AAATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGGAAGATGGGACAGAT ATCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGGTACAAATGGGGCTA TCTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTAA
CvT5 W140G	ATGCTCGACCGGAAATCCAAGCGACCTGACATGCTAGTAGATTCCGTTGGACTGA AATCCATCGTGCGTGATGGCTTGGTCAGCCGCCACAGCTTCAGCATTCGATCCTA CGAGATCGGAGCTGACCGGACCGCCAGTATTGAGACACTGATGAACCATTTACAA GAGACAACTATCAATCACTGTAAGAGCTTGGGATTACACAATGACGGATTCGGCC GCACACCCGGCATGTGTAAGAATGACCTCATCTGGGTATTGACTAAGATGCAAAT TATGGTTAATCGCTATCCTACCTGGGGCGATACTGTAGAGATCAACACCTGGTTT AGTCAAAGTGGCAAGATCGGTATGGCGTCCGACTGGCTAATCTCTGATTGCAACA CCGGGGAAATCCTGATCCGAGCAACTTCTGTGGGCGCCATGATGAACCAGAAGAC CCGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCACTTC GTCGACAGTCCCCATGTTATCGAGGATAATGACCAGAAGTTGCGGAAATTCGATG TCAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCGCGGTGGAATGATCTCGA CGTAAATCAACATGTTTCTAACGTCAAGTACATCGGATGGATTTTGGAATCCATG CCGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCTAACAGTTGAGTACCGAC GCGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTAGTGA AAATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGGAAGATGGGACAGAT ATCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGGTACAAATGGGGCTA TCTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTAA
CvT5 W122V	ATGCTCGACCGGAAATCCAAGCGACCTGACATGCTAGTAGATTCCGTTGGACTGA AATCCATCGTGCGTGATGGCTTGGTCAGCCGCCACAGCTTCAGCATTCGATCCTA CGAGATCGGAGCTGACCGGACCGCCAGTATTGAGACACTGATGAACCATTTACAA GAGACAACTATCAATCACTGTAAGAGCTTGGGATTACACAATGACGGATTCGGCC GCACACCCGGCATGTGTAAGAATGACCTCATCTGGGTATTGACTAAGATGCAAAT TATGGTTAATCGCTATCCTACCTGGGGCGATACTGTAGAGATCAACACCTGGTTT AGTCAAAGTGGCAAGATCGGTATGGCGTCCGACGTGCTAATCTCTGATTGCAACA CCGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGGCCATGATGAACCAGAAGAC CCGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCACTTC GTCGACAGTCCCCATGTTATCGAGGATAATGACCAGAAGTTGCGGAAATTCGATG TCAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCGCGGTGGAATGATCTCGA CGTAAATCAACATGTTTCTAACGTCAAGTACATCGGATGGATTTTGGAATCCATG CCGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCTAACAGTTGAGTACCGAC GCGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTAGTGA AAATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGGAAGATGGGACAGAT ATCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGGTACAAATGGGGCTA TCTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTAA
CvT5 W122M	ATGCTCGACCGGAAATCCAAGCGACCTGACATGCTAGTAGATTCCGTTGGACTGA AATCCATCGTGCGTGATGGCTTGGTCAGCCGCCACAGCTTCAGCATTCGATCCTA CGAGATCGGAGCTGACCGGACCGCCAGTATTGAGACACTGATGAACCATTTACAA GAGACAACTATCAATCACTGTAAGAGCTTGGGATTACACAATGACGGATTCGGCC GCACACCCGGCATGTGTAAGAATGACCTCATCTGGGTATTGACTAAGATGCAAAT TATGGTTAATCGCTATCCTACCTGGGGCGATACTGTAGAGATCAACACCTGGTTT AGTCAAAGTGGCAAGATCGGTATGGCGTCCGACATGCTAATCTCTGATTGCAACA CCGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGGCCATGATGAACCAGAAGAC CCGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCACTTC GTCGACAGTCCCCATGTTATCGAGGATAATGACCAGAAGTTGCGGAAATTCGATG TCAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCGCGGTGGAATGATCTCGA CGTAAATCAACATGTTTCTAACGTCAAGTACATCGGATGGATTTTGGAATCCATG

	CCGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCTAACAGTTGAGTACCGAC GCGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTAGTGA AAATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGGAAGATGGGACAGAT ATCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGGTACAAATGGGGCTA TCTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTAA
CvT5 W122F	ATGCTCGACCGGAAATCCAAGCGACCTGACATGCTAGTAGATTCCGTTGGACTGA AATCCATCGTGCGTGATGGCTTGGTCAGCCGCCACAGCTTCAGCATTCGATCCTA CGAGATCGGAGCTGACCGGACCGCCAGTATTGAGACACTGATGAACCATTTACAA GAGACAACTATCAATCACTGTAAGAGCTTGGGATTACACAATGACGGATTCGGCC GCACACCCGGCATGTGTAAGAATGACCTCATCTGGGTATTGACTAAGATGCAAAT TATGGTTAATCGCTATCCTACCTGGGGCGATACTGTAGAGATCAACACCTGGTTT AGTCAAAGTGGCAAGATCGGTATGGCGTCCGACTTTCTAATCTCTGATTGCAACA CCGGGGAAATCCTGATCCGAGCAACTTCTGTGTGGGCCATGATGAACCAGAAGAC CCGGCGCTTCTCTCGGCTCCCGTATGAAGTTCGTCAAGAACTCACCCCGCACTTC GTCGACAGTCCCCATGTTATCGAGGATAATGACCAGAAGTTGCGGAAATTCGATG TCAAGACAGGAGACAGTATCCGCAAGGGGCTCACACCGCGGTGGAATGATCTCGA CGTAAATCAACATGTTTCTAACGTCAAGTACATCGGATGGATTTTGGAATCCATG CCGATTGAAGTCCTAGAGACACAAGAATTATGCAGTCTAACAGTTGAGTACCGAC GCGAATGTGGCATGGACTCCGTATTAGAGTCTGTAACAGCCGTGGATCCTAGTGA AAATGGTGGGCGCAGTCAGTATAAACATCTCCTCCGTTTGGAAGATGGGACAGAT ATCGTTAAAAGCCGGACTGAATGGCGCCCAAAGAATGCGGGTACAAATGGGGCTA TCTCTACCAGCACGGCGAAGACAAGCAATGGGAACAGTGTTAGTTAA
UaFatB1 (Wild-Type)	ATGGGATCCGGCGCACTGCAAGTTAAGGCCAGCAGCCAAGCTCCCCCTAAACTAA ATGGCTCTAATGTTGGCCTCGTCAAGAGCAGCCAGATTGTTAAGAAGGGAGATGA TACGACAAGTCCCCCTGCTCGTACTTTCATTAATCAGCTACCCGATTGGTCTATG CTGCTCGCCGCTATTACGACTCTATTCCTAGCCGCCGAAAAGCAGTGGATGATGT TGGACTGGAAGCCAAAGCGACCCGATATGCTGGTCGATCCATTCGGATTAGGGCG ATTTGTTCAGGACGGTCTGGTCTTTCGTAACAACTTTTCTATCCGAAGCTATGAG ATCGGAGCGGATCGGACTGCGAGTATTGAGACTCTGATGAATCATTTACAAGAGA CTGCCTTGAATCATGTTAAATCCGTCGGGCTCCTAGAAGATGGGTTGGGGAGCAC TCGAGAAATGTCCTTGCGGAACCTCATTTGGGTTGTAACGAAAATGCAAGTAGCG GTTGACCGGTATCCGACGTGGGGGGATGAGGTTCAGGTTAGCAGCTGGGCTACCG CCATCGGAAAAAATGGGATGCGACGGGAGTGGATCGTAACAGACTTCCGCACGGG TGAGACTCTCTTGCGTGCGACTAGTGTTTGGGTAATGATGAACAAATTGACACGT CGGATCAGCAAAATTCCAGAAGAAGTCTGGCACGAAATCGGTCCGAGCTTCATCG ATGCACCTCCTCTACCTACGGTAGAAGATGACGGACGTAAGCTCACACGGTTCGA CGAGAGTTCCGCGGATTTTATTCGTAGTGGCCTGACGCCCCGCTGGTCTGACTTA GACATCAATCAGCATGTTAATAACGTAAAATATATTGGATGGCTACTAGAAAGCG CTCCTCCCGAAATTCACGAATCTCACGAGATTGCTAGTCTAACACTCGAATACCG CCGAGAGTGCGGTCGTGATAGTGTCCTAAATTCCGCAACTAAAGTCAGTGACTCC AGTCAGCTCGGTAAGAGCGCGGTTGAGTGCAATCACTTAGTACGACTGCAAAACG GAGGAGAGATTGTGAAGGGTCGTACTGTATGGCGGCCAAAACGGCCACTATATAA CGATGGTGCAGTAGTGGACGTTCCTGCCAAGACTTCCTAA
UaT1	ATGCTAAATGGCTCTAATGTTGGCCTCGTCAAGAGCAGCCAGATTGTTAAGAAGG GAGATGATACGACAAGTCCCCCTGCTCGTACTTTCATTAATCAGCTACCCGATTG GTCTATGCTGCTCGCCGCTATTACGACTCTATTCCTAGCCGCCGAAAAGCAGTGG ATGATGTTGGACTGGAAGCCAAAGCGACCCGATATGCTGGTCGATCCATTCGGAT TAGGGCGATTTGTTCAGGACGGTCTGGTCTTTCGTAACAACTTTTCTATCCGAAG CTATGAGATCGGAGCGGATCGGACTGCGAGTATTGAGACTCTGATGAATCATTTA CAAGAGACTGCCTTGAATCATGTTAAATCCGTCGGGCTCCTAGAAGATGGGTTGG GGAGCACTCGAGAAATGTCCTTGCGGAACCTCATTTGGGTTGTAACGAAAATGCA AGTAGCGGTTGACCGGTATCCGACGTGGGGGGATGAGGTTCAGGTTAGCAGCTGG GCTACCGCCATCGGAAAAAATGGGATGCGACGGGAGTGGATCGTAACAGACTTCC GCACGGGTGAGACTCTCTTGCGTGCGACTAGTGTTTGGGTAATGATGAACAAATT GACACGTCGGATCAGCAAAATTCCAGAAGAAGTCTGGCACGAAATCGGTCCGAGC TTCATCGATGCACCTCCTCTACCTACGGTAGAAGATGACGGACGTAAGCTCACAC GGTTCGACGAGAGTTCCGCGGATTTTATTCGTAGTGGCCTGACGCCCCGCTGGTC TGACTTAGACATCAATCAGCATGTTAATAACGTAAAATATATTGGATGGCTACTA GAAAGCGCTCCTCCCGAAATTCACGAATCTCACGAGATTGCTAGTCTAACACTCG AATACCGCCGAGAGTGCGGTCGTGATAGTGTCCTAAATTCCGCAACTAAAGTCAG TGACTCCAGTCAGCTCGGTAAGAGCGCGGTTGAGTGCAATCACTTAGTACGACTG

	CAAAACGGAGGAGAGATTGTGAAGGGTCGTACTGTATGGCGGCCAAAACGGCCAC TATATAACGATGGTGCAGTAGTGGACGTTCCTGCCAAGACTTCCTAA
UaT2	ATGCTACCCGATTGGTCTATGCTGCTCGCCGCTATTACGACTCTATTCCTAGCCG CCGAAAAGCAGTGGATGATGTTGGACTGGAAGCCAAAGCGACCCGATATGCTGGT CGATCCATTCGGATTAGGGCGATTTGTTCAGGACGGTCTGGTCTTTCGTAACAAC TTTTCTATCCGAAGCTATGAGATCGGAGCGGATCGGACTGCGAGTATTGAGACTC TGATGAATCATTTACAAGAGACTGCCTTGAATCATGTTAAATCCGTCGGGCTCCT AGAAGATGGGTTGGGGAGCACTCGAGAAATGTCCTTGCGGAACCTCATTTGGGTT GTAACGAAAATGCAAGTAGCGGTTGACCGGTATCCGACGTGGGGGGATGAGGTTC AGGTTAGCAGCTGGGCTACCGCCATCGGAAAAAATGGGATGCGACGGGAGTGGAT CGTAACAGACTTCCGCACGGGTGAGACTCTCTTGCGTGCGACTAGTGTTTGGGTA ATGATGAACAAATTGACACGTCGGATCAGCAAAATTCCAGAAGAAGTCTGGCACG AAATCGGTCCGAGCTTCATCGATGCACCTCCTCTACCTACGGTAGAAGATGACGG ACGTAAGCTCACACGGTTCGACGAGAGTTCCGCGGATTTTATTCGTAGTGGCCTG ACGCCCCGCTGGTCTGACTTAGACATCAATCAGCATGTTAATAACGTAAAATATA TTGGATGGCTACTAGAAAGCGCTCCTCCCGAAATTCACGAATCTCACGAGATTGC TAGTCTAACACTCGAATACCGCCGAGAGTGCGGTCGTGATAGTGTCCTAAATTCC GCAACTAAAGTCAGTGACTCCAGTCAGCTCGGTAAGAGCGCGGTTGAGTGCAATC ACTTAGTACGACTGCAAAACGGAGGAGAGATTGTGAAGGGTCGTACTGTATGGCG GCCAAAACGGCCACTATATAACGATGGTGCAGTAGTGGACGTTCCTGCCAAGACT TCCTAA
UaT3	ATGCTGCTCGCCGCTATTACGACTCTATTCCTAGCCGCCGAAAAGCAGTGGATGA TGTTGGACTGGAAGCCAAAGCGACCCGATATGCTGGTCGATCCATTCGGATTAGG GCGATTTGTTCAGGACGGTCTGGTCTTTCGTAACAACTTTTCTATCCGAAGCTAT GAGATCGGAGCGGATCGGACTGCGAGTATTGAGACTCTGATGAATCATTTACAAG AGACTGCCTTGAATCATGTTAAATCCGTCGGGCTCCTAGAAGATGGGTTGGGGAG CACTCGAGAAATGTCCTTGCGGAACCTCATTTGGGTTGTAACGAAAATGCAAGTA GCGGTTGACCGGTATCCGACGTGGGGGGATGAGGTTCAGGTTAGCAGCTGGGCTA CCGCCATCGGAAAAAATGGGATGCGACGGGAGTGGATCGTAACAGACTTCCGCAC GGGTGAGACTCTCTTGCGTGCGACTAGTGTTTGGGTAATGATGAACAAATTGACA CGTCGGATCAGCAAAATTCCAGAAGAAGTCTGGCACGAAATCGGTCCGAGCTTCA TCGATGCACCTCCTCTACCTACGGTAGAAGATGACGGACGTAAGCTCACACGGTT CGACGAGAGTTCCGCGGATTTTATTCGTAGTGGCCTGACGCCCCGCTGGTCTGAC TTAGACATCAATCAGCATGTTAATAACGTAAAATATATTGGATGGCTACTAGAAA GCGCTCCTCCCGAAATTCACGAATCTCACGAGATTGCTAGTCTAACACTCGAATA CCGCCGAGAGTGCGGTCGTGATAGTGTCCTAAATTCCGCAACTAAAGTCAGTGAC TCCAGTCAGCTCGGTAAGAGCGCGGTTGAGTGCAATCACTTAGTACGACTGCAAA ACGGAGGAGAGATTGTGAAGGGTCGTACTGTATGGCGGCCAAAACGGCCACTATA TAACGATGGTGCAGTAGTGGACGTTCCTGCCAAGACTTCCTAA
UaT4	ATGGAAAAGCAGTGGATGATGTTGGACTGGAAGCCAAAGCGACCCGATATGCTGG TCGATCCATTCGGATTAGGGCGATTTGTTCAGGACGGTCTGGTCTTTCGTAACAA CTTTTCTATCCGAAGCTATGAGATCGGAGCGGATCGGACTGCGAGTATTGAGACT CTGATGAATCATTTACAAGAGACTGCCTTGAATCATGTTAAATCCGTCGGGCTCC TAGAAGATGGGTTGGGGAGCACTCGAGAAATGTCCTTGCGGAACCTCATTTGGGT TGTAACGAAAATGCAAGTAGCGGTTGACCGGTATCCGACGTGGGGGGATGAGGTT CAGGTTAGCAGCTGGGCTACCGCCATCGGAAAAAATGGGATGCGACGGGAGTGGA TCGTAACAGACTTCCGCACGGGTGAGACTCTCTTGCGTGCGACTAGTGTTTGGGT AATGATGAACAAATTGACACGTCGGATCAGCAAAATTCCAGAAGAAGTCTGGCAC GAAATCGGTCCGAGCTTCATCGATGCACCTCCTCTACCTACGGTAGAAGATGACG GACGTAAGCTCACACGGTTCGACGAGAGTTCCGCGGATTTTATTCGTAGTGGCCT GACGCCCCGCTGGTCTGACTTAGACATCAATCAGCATGTTAATAACGTAAAATAT ATTGGATGGCTACTAGAAAGCGCTCCTCCCGAAATTCACGAATCTCACGAGATTG CTAGTCTAACACTCGAATACCGCCGAGAGTGCGGTCGTGATAGTGTCCTAAATTC CGCAACTAAAGTCAGTGACTCCAGTCAGCTCGGTAAGAGCGCGGTTGAGTGCAAT CACTTAGTACGACTGCAAAACGGAGGAGAGATTGTGAAGGGTCGTACTGTATGGC GGCCAAAACGGCCACTATATAACGATGGTGCAGTAGTGGACGTTCCTGCCAAGAC TTCCTAA
UaT5	ATGTTGGACTGGAAGCCAAAGCGACCCGATATGCTGGTCGATCCATTCGGATTAG GGCGATTTGTTCAGGACGGTCTGGTCTTTCGTAACAACTTTTCTATCCGAAGCTA TGAGATCGGAGCGGATCGGACTGCGAGTATTGAGACTCTGATGAATCATTTACAA GAGACTGCCTTGAATCATGTTAAATCCGTCGGGCTCCTAGAAGATGGGTTGGGGA

	GCACTCGAGAAATGTCCTTGCGGAACCTCATTTGGGTTGTAACGAAAATGCAAGT AGCGGTTGACCGGTATCCGACGTGGGGGGATGAGGTTCAGGTTAGCAGCTGGGCT ACCGCCATCGGAAAAAATGGGATGCGACGGGAGTGGATCGTAACAGACTTCCGCA CGGGTGAGACTCTCTTGCGTGCGACTAGTGTTTGGGTAATGATGAACAAATTGAC ACGTCGGATCAGCAAAATTCCAGAAGAAGTCTGGCACGAAATCGGTCCGAGCTTC ATCGATGCACCTCCTCTACCTACGGTAGAAGATGACGGACGTAAGCTCACACGGT TCGACGAGAGTTCCGCGGATTTTATTCGTAGTGGCCTGACGCCCCGCTGGTCTGA CTTAGACATCAATCAGCATGTTAATAACGTAAAATATATTGGATGGCTACTAGAA AGCGCTCCTCCCGAAATTCACGAATCTCACGAGATTGCTAGTCTAACACTCGAAT ACCGCCGAGAGTGCGGTCGTGATAGTGTCCTAAATTCCGCAACTAAAGTCAGTGA CTCCAGTCAGCTCGGTAAGAGCGCGGTTGAGTGCAATCACTTAGTACGACTGCAA AACGGAGGAGAGATTGTGAAGGGTCGTACTGTATGGCGGCCAAAACGGCCACTAT ATAACGATGGTGCAGTAGTGGACGTTCCTGCCAAGACTTCCTAA
UaT6	ATGCGACCCGATATGCTGGTCGATCCATTCGGATTAGGGCGATTTGTTCAGGACG GTCTGGTCTTTCGTAACAACTTTTCTATCCGAAGCTATGAGATCGGAGCGGATCG GACTGCGAGTATTGAGACTCTGATGAATCATTTACAAGAGACTGCCTTGAATCAT GTTAAATCCGTCGGGCTCCTAGAAGATGGGTTGGGGAGCACTCGAGAAATGTCCT TGCGGAACCTCATTTGGGTTGTAACGAAAATGCAAGTAGCGGTTGACCGGTATCC GACGTGGGGGGATGAGGTTCAGGTTAGCAGCTGGGCTACCGCCATCGGAAAAAAT GGGATGCGACGGGAGTGGATCGTAACAGACTTCCGCACGGGTGAGACTCTCTTGC GTGCGACTAGTGTTTGGGTAATGATGAACAAATTGACACGTCGGATCAGCAAAAT TCCAGAAGAAGTCTGGCACGAAATCGGTCCGAGCTTCATCGATGCACCTCCTCTA CCTACGGTAGAAGATGACGGACGTAAGCTCACACGGTTCGACGAGAGTTCCGCGG ATTTTATTCGTAGTGGCCTGACGCCCCGCTGGTCTGACTTAGACATCAATCAGCA TGTTAATAACGTAAAATATATTGGATGGCTACTAGAAAGCGCTCCTCCCGAAATT CACGAATCTCACGAGATTGCTAGTCTAACACTCGAATACCGCCGAGAGTGCGGTC GTGATAGTGTCCTAAATTCCGCAACTAAAGTCAGTGACTCCAGTCAGCTCGGTAA GAGCGCGGTTGAGTGCAATCACTTAGTACGACTGCAAAACGGAGGAGAGATTGTG AAGGGTCGTACTGTATGGCGGCCAAAACGGCCACTATATAACGATGGTGCAGTAG TGGACGTTCCTGCCAAGACTTCCTAA
UaT7	ATGGGGCGATTTGTTCAGGACGGTCTGGTCTTTCGTAACAACTTTTCTATCCGAA GCTATGAGATCGGAGCGGATCGGACTGCGAGTATTGAGACTCTGATGAATCATTT ACAAGAGACTGCCTTGAATCATGTTAAATCCGTCGGGCTCCTAGAAGATGGGTTG GGGAGCACTCGAGAAATGTCCTTGCGGAACCTCATTTGGGTTGTAACGAAAATGC AAGTAGCGGTTGACCGGTATCCGACGTGGGGGGATGAGGTTCAGGTTAGCAGCTG GGCTACCGCCATCGGAAAAAATGGGATGCGACGGGAGTGGATCGTAACAGACTTC CGCACGGGTGAGACTCTCTTGCGTGCGACTAGTGTTTGGGTAATGATGAACAAAT TGACACGTCGGATCAGCAAAATTCCAGAAGAAGTCTGGCACGAAATCGGTCCGAG CTTCATCGATGCACCTCCTCTACCTACGGTAGAAGATGACGGACGTAAGCTCACA CGGTTCGACGAGAGTTCCGCGGATTTTATTCGTAGTGGCCTGACGCCCCGCTGGT CTGACTTAGACATCAATCAGCATGTTAATAACGTAAAATATATTGGATGGCTACT AGAAAGCGCTCCTCCCGAAATTCACGAATCTCACGAGATTGCTAGTCTAACACTC GAATACCGCCGAGAGTGCGGTCGTGATAGTGTCCTAAATTCCGCAACTAAAGTCA GTGACTCCAGTCAGCTCGGTAAGAGCGCGGTTGAGTGCAATCACTTAGTACGACT GCAAAACGGAGGAGAGATTGTGAAGGGTCGTACTGTATGGCGGCCAAAACGGCCA CTATATAACGATGGTGCAGTAGTGGACGTTCCTGCCAAGACTTCCTAA
UaT8	ATGCGTAACAACTTTTCTATCCGAAGCTATGAGATCGGAGCGGATCGGACTGCGA GTATTGAGACTCTGATGAATCATTTACAAGAGACTGCCTTGAATCATGTTAAATC CGTCGGGCTCCTAGAAGATGGGTTGGGGAGCACTCGAGAAATGTCCTTGCGGAAC CTCATTTGGGTTGTAACGAAAATGCAAGTAGCGGTTGACCGGTATCCGACGTGGG GGGATGAGGTTCAGGTTAGCAGCTGGGCTACCGCCATCGGAAAAAATGGGATGCG ACGGGAGTGGATCGTAACAGACTTCCGCACGGGTGAGACTCTCTTGCGTGCGACT AGTGTTTGGGTAATGATGAACAAATTGACACGTCGGATCAGCAAAATTCCAGAAG AAGTCTGGCACGAAATCGGTCCGAGCTTCATCGATGCACCTCCTCTACCTACGGT AGAAGATGACGGACGTAAGCTCACACGGTTCGACGAGAGTTCCGCGGATTTTATT CGTAGTGGCCTGACGCCCCGCTGGTCTGACTTAGACATCAATCAGCATGTTAATA ACGTAAAATATATTGGATGGCTACTAGAAAGCGCTCCTCCCGAAATTCACGAATC TCACGAGATTGCTAGTCTAACACTCGAATACCGCCGAGAGTGCGGTCGTGATAGT GTCCTAAATTCCGCAACTAAAGTCAGTGACTCCAGTCAGCTCGGTAAGAGCGCGG TTGAGTGCAATCACTTAGTACGACTGCAAAACGGAGGAGAGATTGTGAAGGGTCG

	TACTGTATGGCGGCCAAAACGGCCACTATATAACGATGGTGCAGTAGTGGACGTT CCTGCCAAGACTTCCTAA
UaT9	ATGATCGGAGCGGATCGGACTGCGAGTATTGAGACTCTGATGAATCATTTACAAG AGACTGCCTTGAATCATGTTAAATCCGTCGGGCTCCTAGAAGATGGGTTGGGGAG CACTCGAGAAATGTCCTTGCGGAACCTCATTTGGGTTGTAACGAAAATGCAAGTA GCGGTTGACCGGTATCCGACGTGGGGGGATGAGGTTCAGGTTAGCAGCTGGGCTA CCGCCATCGGAAAAAATGGGATGCGACGGGAGTGGATCGTAACAGACTTCCGCAC GGGTGAGACTCTCTTGCGTGCGACTAGTGTTTGGGTAATGATGAACAAATTGACA CGTCGGATCAGCAAAATTCCAGAAGAAGTCTGGCACGAAATCGGTCCGAGCTTCA TCGATGCACCTCCTCTACCTACGGTAGAAGATGACGGACGTAAGCTCACACGGTT CGACGAGAGTTCCGCGGATTTTATTCGTAGTGGCCTGACGCCCCGCTGGTCTGAC TTAGACATCAATCAGCATGTTAATAACGTAAAATATATTGGATGGCTACTAGAAA GCGCTCCTCCCGAAATTCACGAATCTCACGAGATTGCTAGTCTAACACTCGAATA CCGCCGAGAGTGCGGTCGTGATAGTGTCCTAAATTCCGCAACTAAAGTCAGTGAC TCCAGTCAGCTCGGTAAGAGCGCGGTTGAGTGCAATCACTTAGTACGACTGCAAA ACGGAGGAGAGATTGTGAAGGGTCGTACTGTATGGCGGCCAAAACGGCCACTATA TAACGATGGTGCAGTAGTGGACGTTCCTGCCAAGACTTCCTAA
Tes4	ATGAGCGATGACAAAAAAATTGGGAGTTACAAATTTATTGCCGAACCCTTTCACG TTGACTTTAATGGGCGGTTGACTATGGGCGTTCTGGGGAATCATTTGTTAAACTG TGCCGGATTTCACGCTTCTGAACGGGGGTTTGGAATTGCCACCTTAAATGAAGAT AACTATACTTGGGTGTTGTCCCGCTTAGCTATTGACCTGGAAGAAATGCCCTATC AATACGAAGAATTTACCGTTCAGACTTGGGTGGAAAATGTTTATCGTTTATTTAC CGATCGGAACTTTGCCATTATTGATAAAGACGGCAAGAAAATTGGTTACGCCCGC TCCGTGTGGGCTATGATTAATCTGAACACCCGTAAACCCGCCGATTTGTTGACTT TGCATGGCGGTAGTATTGTGGATTACGTTTGTGACGAACCCTGCCCCATTGAAAA ACCCAGCCGTATTAAAGTGGCCACCGATCAACCCTGTGCCAAATTGACTGCTAAA TATTCTGATATTGACATTAACGGCCATGTGAACTCCATTCGGTACATTGAACACA TTTTGGATTTGTTTCCCATTGACTTGTACAAAAGTAAACGGATTCAGCGCTTTGA AATGGCCTACGTGGCTGAATCCTATTACGGGGATGAATTGTCCTTTTTCGAAGAA GAAGTGAGCGAAAACGAATACCACGTTGAAATTAAGAAAAATGGTAGCGAAGTGG TTTGCCGCGCCAAAGTGAAATTTGTTTAA
BtTes	ATGTCCGAGGAAAATAAAATTGGAACTTACCAGTTTGTAGCCGAACCATTTCACG TAGACTTCAACGGTCGACTGACAATGGGTGTTCTGGGCAACCACCTACTAAACTG CGCTGGTTTCCACGCCTCCGACCGTGGATTTGGTATCGCGACTCTCAACGAGGAT AATTACACATGGGTTCTATCCCGCCTAGCGATTGAGTTGGATGAAATGCCCTATC AGTACGAAAAATTCTCCGTCCAGACATGGGTCGAAAATGTATATCGCCTGTTCAC GGACCGGAATTTCGCGGTGATTGATAAAGATGGCAAAAAAATTGGTTACGCCCGA TCTGTTTGGGCCATGATCAACCTCAACACCCGCAAACCAGCGGACCTCCTGGCTC TCCACGGAGGTAGTATCGTGGATTACATTTGTGACGAACCCTGCCCTATCGAGAA GCCTTCCCGGATTAAGGTGACATCCAACCAACCAGTAGCAACGCTAACCGCTAAA TACTCTGATATTGATATTAACGGTCATGTGAATAGCATTCGGTATATCGAACATA TCCTCGACTTGTTCCCTATTGAGTTGTATCAAACAAAACGAATCCGCCGGTTCGA AATGGCATACGTCGCTGAATCTTACTTTGGTGATGAGCTAAGCTTCTTCTGCGAT GAAGTGTCTGAGAATGAATTTCATGTTGAGGTTAAAAAGAATGGTAGTGAGGTTG TATGTCGTTCCAAGGTCATCTTCGAGTAA
MfTes	ATGATCTACATGGCCTACCAGTATCGAAGCCGGATTCGATATTCTGAAATCGGCG AAGACAAAAAACTGACTCTACCGGGATTGGTAAACTATTTCCAGGACTGTTCCAC GTTTCAGAGCGAAGCCCTGGGGATCGGCCTGGATACCTTGGGCGCGCGGCAACGG GCGTGGTTGCTCGCCAGCTGGAAGATTGTTATTGACCGACTACCGCGTTTGGGTG AGGAAGTGGTCACAGAGACATGGCCATACGGGTTCAAGGGATTCCAGGGGAACCG TAACTTCCGAATGCTGGATCAAGAGGGGCACACGCTGGCTGCAGCCGCGTCTGTA TGGATTTATCTAAACGTGGAGTCTGGTCACCCTTGTCGGATTGATGGCGATGTAC TCGAAGCGTATGAGTTAGAAGAGGAACTGCCTCTCGGTCCTTTCTCTCGTAAAAT TCCAGTACCGGAGGAGTCCACGGAACGCGATAGTTTTTTAGTCATGCGAAGTCAT TTGGACACCAACCATCATGTTAATAACGGTCAGTATATTCTCATGGCAGAGGAAT ACCTGCCGGAGGGCTTTAAGGTTAAGCAAATTCGGGTGGAATATCGTAAAGCTGC CGTACTCCACGACACCATCGTTCCTTTCGTGTGCACGGAACCGCAACGGTGCACC GTGTCTTTGTGCGGGTCCGATGAAAAACCCTTCGCCGTCGTGGAATTCTCTGAAT AA
EcACP	ATGAGCACTATCGAAGAACGCGTTAAGAAAATTATCGGCGAACAGCTGGGCGTTA AGCAGGAAGAAGTTACCAACAATGCTTCTTTCGTTGAAGACCTGGGCGCGGATTC

	TCTTGACACCGTTGAGCTGGTAATGGCTCTGGAAGAAGAGTTTGATACTGAGATT CCGGACGAAGAAGCTGAGAAAATCACCACCGTTCAGGCTGCCATTGATTACATCA ACGGCCACCAGGCGTAA
6803ACP	ATGAATCAGGAAATTTTTGAAAAAGTAAAAAAAATCGTCGTGGAACAGTTGGAAG TGGATCCTGACAAAGTGACCCCCGATGCCACCTTTGCCGAAGATTTAGGGGCTGA TTCCCTCGATACAGTGGAATTGGTCATGGCCCTGGAAGAAGAGTTTGATATTGAA ATTCCCGATGAAGTGGCGGAAACCATTGATACCGTGGGCAAAGCCGTTGAGCATA TCGAAAGTAAATAA
ClACP2	ATGGCCGCGAAACCTGAAACAGTTAAGAAGGTCTGTGAGATCGTGAAGAAGCAAC TGGCACTACCAGATGACTCTGCGGTCACGGGCGCGAGCAAATTTAGCGCACTCGG AGCCGATAGCTTGGATACGGTAGAAATCGTGATGGGCTTAGAAGAAGAATTTGGG ATTTCTGTGGAAGAGGAGAGTGCCCAATCCATCCAAACCGTACAGGATGCGGCCG ACCTCATCGAATAA
BM3 (Wild-Type)	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGTTCGTGATTTTGCGGGTGACGGTTTATTTACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACTACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCATTGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT

	TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 F87A	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGTTCGTGATTTTGCGGGTGACGGTTTAGCCACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACTACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCATTGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC

	GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 F87G	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGTTCGTGATTTTGCGGGTGACGGTTTAGGCACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACTACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCATTGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC

	TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 F87V	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGTTCGTGATTTTGCGGGTGACGGTTTAGTGACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACTACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCATTGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC

	GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 F87I	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGTTCGTGATTTTGCGGGTGACGGTTTAATTACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACTACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCATTGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA

BM3 V78A F87A	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGCCCGTGATTTTGCGGGTGACGGTTTAGCCACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACTACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCATTGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 L188Q	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT

	CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGTTCGTGATTTTGCGGGTGACGGTTTATTTACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACAACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCATTGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 F87A L188Q	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGTTCGTGATTTTGCGGGTGACGGTTTAGCCACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG

	CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACAACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCATTGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 F87A L188Q I401P	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGTTCGTGATTTTGCGGGTGACGGTTTAGCCACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC

	TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACAACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCCCCGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 V78A F87A L188Q	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGCCCGTGATTTTGCGGGTGACGGTTTAGCCACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACAACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC

	ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCATTGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 V78A F87A L188Q I401P	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGCCCGTGATTTTGCGGGTGACGGTTTAGCCACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACAACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT

	TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCCCCGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 F87G L188Q	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGTTCGTGATTTTGCGGGTGACGGTTTAGGCACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACAACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG

	GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCATTGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 F87G I401P	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGTTCGTGATTTTGCGGGTGACGGTTTAGGCACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACTACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG

	ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCCCCGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 F87G L188Q I401P	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGTTCGTGATTTTGCGGGTGACGGTTTAGGCACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACAACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCCCCGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG

	AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 V78A F87A L188Q D182S	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGCCCGTGATTTTGCGGGTGACGGTTTAGCCACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCTCCG AGGCCATGAACAAACAACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCATTGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG

	GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 V78A F87A L188Q H266T	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGCCCGTGATTTTGCGGGTGACGGTTTAGCCACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACAACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGACCGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCATTGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC

	TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 V78A F87A L188Q A328V	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGCCCGTGATTTTGCGGGTGACGGTTTAGCCACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACAACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGTGCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCATTGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT

	GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 F87G L188Q D182S	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGTTCGTGATTTTGCGGGTGACGGTTTAGGCACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCTCCG AGGCCATGAACAAACAACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCATTGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC

	GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 F87G L188Q H266T	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGTTCGTGATTTTGCGGGTGACGGTTTAGGCACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACAACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGACCGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGCCCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCATTGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC

	ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
BM3 F87G L188Q A328V	ATGACCATCAAGGAAATGCCGCAGCCGAAGACATTTGGTGAGCTGAAGAACCTCC CGCTATTGAATACGGATAAGCCCGTGCAGGCCCTAATGAAAATCGCGGATGAACT CGGTGAGATTTTTAAATTCGAAGCACCCGGACGGGTGACGCGATACCTAAGTAGT CAGCGCCTCATTAAGGAAGCCTGCGATGAAAGCCGATTCGACAAAAACTTGTCCC AGGCTCTCAAGTTTGTTCGTGATTTTGCGGGTGACGGTTTAGGCACCAGTTGGAC ACATGAGAAAAATTGGAAAAAAGCTCATAATATCCTCCTACCGAGCTTTTCCCAG CAAGCAATGAAAGGGTATCATGCTATGATGGTTGACATCGCCGTTCAGCTCGTCC AAAAATGGGAACGCCTAAATGCCGACGAGCACATTGAGGTACCCGAAGATATGAC GCGACTAACACTAGATACTATTGGCCTCTGCGGTTTTAACTATCGGTTTAATAGC TTCTATCGCGATCAGCCACATCCGTTTATTACTAGCATGGTGCGTGCTCTCGACG AGGCCATGAACAAACAACAGCGTGCTAACCCAGACGACCCGGCCTACGATGAGAA TAAGCGCCAATTTCAAGAAGACATCAAAGTCATGAACGACCTGGTCGATAAGATC ATTGCCGACCGTAAGGCTTCTGGCGAACAGAGCGACGATCTCTTGACCCACATGT TAAACGGGAAAGATCCCGAGACAGGCGAGCCGTTAGATGATGAAAACATTCGCTA CCAAATCATCACGTTTTTGATTGCCGGGCATGAGACGACGAGTGGTTTGTTGAGT TTCGCACTGTATTTCTTGGTAAAAAACCCACATGTTTTACAAAAGGCGGCGGAGG AAGCCGCTCGCGTATTGGTAGACCCAGTGCCTTCTTACAAACAGGTCAAGCAATT GAAATACGTAGGGATGGTTCTCAACGAAGCATTGCGGCTATGGCCAACGGTGCCG GCTTTTAGTTTGTATGCGAAGGAGGACACAGTCCTGGGAGGAGAATACCCATTGG AGAAAGGTGACGAATTGATGGTATTAATTCCACAACTACACCGAGACAAAACTAT CTGGGGTGATGATGTGGAAGAGTTTCGACCCGAACGTTTTGAAAACCCGAGCGCG ATTCCGCAACATGCTTTTAAACCTTTCGGCAACGGGCAACGTGCCTGCATTGGAC AGCAATTTGCTTTACACGAAGCGACACTGGTGTTGGGCATGATGTTGAAGCATTT CGACTTCGAGGACCACACAAACTATGAACTAGACATCAAGGAGACATTAACTTTG AAGCCAGAAGGTTTCGTTGTGAAAGCGAAGTCCAAAAAAATCCCATTAGGGGGAA TCCCAAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGGCAGAGAA TGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAACGGCAGAG GGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCCCCCCAAG TCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTGTTTTGAT TGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTTCGTCGAC TGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCCGTGTTTG GGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGTTTATCGA TGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGAAGCAGAT GCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATGTGGAGCG ACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATAAGTCCAC GTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAAGATGCAC GGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCCGGCAGTG CCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTTACCAAGA AGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAACCGCGTC ACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCAGAAGAGG AAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGCTACTGCA ATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGATGGCTGCT

	AAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAGAAACAAG CATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTATTAGAAAA ATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCCTTCCATC CGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAGCAAGCTT CCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGGAATACAA GGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTATTACTTGT TTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAGACTCCAC TCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCGTTCAGGC ACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCTGTATTTC GGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAAAATGCAC AGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGAATCAGCC TAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCATCGAACTC TTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATGGCTCCAG CGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGTCCGAGGC GGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGCCAAAGAT GTATGGGCAGGTTAA
RbCYP450	ATGGAGCGTGAATTGCGTGAAACCGTAACGGCTCGATGCCCCGTTGCTCACGATG CATCCTCCATCAACGGATGTCCTGTTCAGCGGATGGCTCAAGAATTTGACCCCTT TCAGCACGACTACATGGCCGACCCTGCGCAGTTCGTACGTTGGGCACGAGAACAG GCTCCCATCTTCTACGCTCCTAAACTCGATTACTGGGTCGTTACCCGGTATGATA CCATCAAACAGATCTTCCGTGACCCAGTGACTTTTAGCCCAGCAAATGTTTTGCA AAGCTTCGCGCAACCAAGTGCGGAAGTGCGACAAGTACTCGAACGTTACGGATAT GCGTTTAATCGCACCCTCGTGAACGAGGACGAACCAATGCACCTAGAGCGCCGCC GCGTGCTAATGGAACCCTTCACACCCGAACACCTAGCCGAGCATGAGCCGATGGT GCGGGAGCTAGTGCGACGAGCAGTAGATCGCTTTATCGACGCGGGCCGCGCGGAT TTAGTGGATCAGATGATCTGGGAAGTCCCATTCACTGTCGCCTTACATTTTCTGG GGGTGGACGATGATGATCGGGAAAAGATGCGACGGTTCGCTATCGCTCACACGGT GAATGCGTTTGGTCGACCCAGTCCTGAAGAGCAACTGGCTGTTGCGGAGACGGTC GGACAGTTCTGGCAGTTCTGCGGTGAAGTCTTGGAGAAAATGCGACGCACAGCGG ATGGGCCCGGCTGGATGCGGTATAGCATCCGCCAGCAAAAGTTGTACCCAGAGGT TGTTACAGACTCTTACCTCCACAGTATGATGCAGGCGATTATCGTAGCTGCACAT GAGACGACCGTATTTGCCACGACAAACGCTTTAAAGACGCTATTGGAGCATGAGG CAGTATGGCGCGAAATTTGCGCCGACCCCTCCTTGATTCCTGCCGCTGCAGAAGA ATGCTTACGCTATAACGGCCCTGTGGCTGGATGGCGTCGGCGGACCACGCGAGAG GTTGAAGTCGAGGGTGTTCGTCTACCAACCGGAGCGAATATTCTAATGATCGTCG CGTCTGCAAATCATGATCCAGCTCACTTCGACGATCCTGAGTTTTTCGATATCGG GCGAGCAAACGCGTCTGAACATTTAAACTTCGGATACGGTGCTCATCAATGCTTG GGCAAAAATTTAGGCCGTATGGAGATGCAGATCATGATCGAGGAATTGTCTCGTC GTTTACCCCACATGCGGCTAGCGGAACAGCGTTTCGACTACCTGCATAATGTCAG TTTCCGAGCTCCTCGACACTTATGGGTTGAGTGGGACCCTGCTCAGAATCCAGAA CGCCGCGATGCTAGTATTTTACATCTACGACAACCGGTGCGTATTGGGCCTCCGC GCGCAAAAGACGTCGTACGTACAATGGAAGTGGCCGCTGTTGAGCGCCCGGCTGA AAATATCGTAGTGCTGCATCTGACCCGTCCAGATCGGCGTCCCCTCCCGCGTTGG AGTCCGGGGGCACACATTGATATTGAGTGCGGGGAGCCGGAGCGGAGCCGGCAAT ACAGCCTGTGCGGTGACCCGGAAAACCGAAATGAATGGCGTGTGGCTGTTCAACG TGATCCGGCTTCCCGCGGTGGCAGTCGATGGATCCATGAACAGGTTCGCCCCGGT ATGCTACTACGAGTTCGTGGCCCGCGTAACTCCTTCCGCTTAGACGAAAACGCTC CCCGTTATCTGTTCTTAGCCGGCGGCATCGGCATTACCCCCATTATGACAATGGC GGCTCGCGCAAAGGAACTCGGCGCTGATTACGAATTACACTATTCTGCCCGTAGT CGAGCAAGCCTAATTTTTGTGGACGAACTCCGACAAAGTCATGGAGATCGCCTAC ATGTCTACGTATCTGAGGAGGGGGTACGTAATGATTTAGCAGCGCTCATCGGACG GGCGTCCGCTGACACGCAGATTTATGCGTGTGGACCTCAGCGAATGCTAGAGGCT TTAGAGCGTTTGGTCGAGAACCGCCCCGAAGTGACACTCCGCGTGGAACACTTCT TCGGAGAAGCCAGTCGTTTGGATCCGGCTAAGGAACGGTCCTTCCAAGTCGTACT GCGCAATTCTGGCTTGACGCTAGAAGTACCTGCTGATAAAACCTTATTGGAAGTG CTACGCGCATACAATATCGAAGTACAGAGTGACTGTGAAGAGGGACTCTGTGGTA CATGCGAAGTCAGTGTTGTGGACGGTGAAGTTGATCACCGAGACTCCGTCTTAAC GCGTGCCGAACGACGCGAGAACCGCCGAATGATGTGCTGCTGCTCTCGGGCGAAA ACTGAACGTCTCGTTTTAGATCTGTGA
TtCYP450	ATGGAAACTGAGCTAAAGGAGACTGCTCGGGGTACGTGTCCAGTTGCCCACGGGG GACAGTCCTCTGTAGGCGGATGTCCGGTCCATCGCCTAGCGGAGGATTTCGATCC

	GTTTCAAGATGCCTATATGGCAGACCCTGCACAATTCGTACGGTGGGCCCGAGAG CAGGTTCCTATTTTCTACGCACCGAAATTAAACTACTGGGTAGTGACCCGTTACG ACACAATTAAGCAAATTTTTCGCGACCCGGTCACTTTTAGCCCGAGCAACGTGCT CCAATCCTTTGCCCAACCTAGCGCCGAGGTGCGCCAGGTCCTGGAGCGTTACGGA TATGCCTTTAACCGTACTTTGGTGAACGAGGATGAACCTATGCATCTGGAACGCC GCCGAGTATTAATGGAACCGTTTGCCTCCGAGCACCTGGCAGAGCACGAGCCGAT GGTTCGTGAGCTAGTGCGTCGAGCCGTCAACCGTTTCATTGATACGGGCCGGGCA GACCTCGTGGACCAAATGATCTGGGAAGTGCCTTTTACGGTGGCTCTACATTTCC TGGGAGTAGACGACGATGATCGTGAGAAGATGCGTCGGTTCGCAATCGCACATAC GGTTAATGCATTCGGACGTCCATCTCCAGAAGAACAATTAGCCGTAGCAGAAACC GTGGGCCAGTTCTGGCAATTTTGCGGTGAAGTCTTAGAGAAAATGCGCCGTACCG CGGACGGCCCTGGATGGATGCGCTATAGTATTCGTCAACAGAAGCTATACCCCGA CGTCGTCACTGACTCCTATCTGCATTCCATGATGCAAGCAATTATCGTTGCTGCA CACGAAACGACCGTTTTTGCCACCACAAATGCACTCAAGACTCTACTAGAGCACG AAACTGTATGGCGCGAAATTTGCGCCGATCCATCCCTGATCCCCGCAGCAGCCGA GGAGTGTTTACGCTACAATGGTCCTGTTGCAGGGTGGCGACGCCGGACCACACGA GAGGTCGAGGTCGAGGGCGTACGTCTGCCAGTGGGGGCAAACATCTTGATGGTTG TAGCCTCCGCAAACCACGACTCTGCTCATTTCGACGATCCTGAATTTTTTGATAT TGGACGTTCCAACGCTTCTGAGCATCTAAATTTTGGTTACGGCGCACATCAGTGC CTAGGGCGAAACCTAGGACGAATGGAGATGCAGATTATGATTGAAGAGCTGTCTC GACGCCTCCCCCATATGCGGCTAGCCGAGCAGCGATTTGATTATTTACATAACGT TTCCTTCCGGGCGCCTCGGCATCTGTGGGTTGAGTGGGATCCAGCCCAAAATCCC GAGCGTCGCGACCCCGACATTTTACGACTACGCCAGCCAGTTCGTATTGGCCCCC CGCGAGCTAAGGACGTGGTCCGTACTATGGAGGTGGCTGCAGTAGAACGCCCCTC CGAGGACATTGTCGTGTTGCACTTGACGCGACCTGATCGACGCCCGCTGCCACGT TGGTCACCTGGCGCGCATATTGACATCGAGTGCGGCGAACCGGATCGCAGCCGTC AATATAGTCTATGTTCCGATCCGGAAAACCGAGACGCTTGGCGCGTCGCTGTGCA GCGTGATCCGGCTAGCCGCGGAGGTAGTCGCTGGATCCACGAGGAAGTGCGCCCT GGCATGTTACTACGTGTTCGTGGACCGCGCAATAGCTTCCGTTTAGACGAACACG CTCCGCGTTATTTGTTCCTGGCTGGGGGAATTGGCATCACCCCGATTATGACCAT GGCGGCTCGTGCAAAGGAGCTGGGTACGGACTACGAACTACACTACAGCGTCCGT AGCCGAACAAGTTTAATCTTCGTCGACGAACTACGACAGATTCATGGCGATCGGC TACACGTGTACGTCTCCGAGGAGGGGGTCCGCAATGATCTGGCTGCGCTGATTCG TCGCGCCTCTGCCGGCACGCAAATTTACGCATGCGGTCCACAACGTATGCTCGAC ACATTGGAGCGTCTGATTGAAAATCGTCCCGAAGTCACCCTCCGTGTCGAACATT TTTTCGGCGAACCTTCTCACCTAGACCCCGCGAAGGAGCGTCCATTCCAGGTTGT ACTCCGTAATTCCGGATTGACAGTAGAAGTACCCGCGGATAAAACTCTCCTAGAG GTACTCCGTGCGTACAATATCGAGGTACAGAGCGACTGTGAGGAAGGGTTATGCG GGACGTGTGAGGTTTCTGTCGTGGAGGGGGAGGTTGATCACCGAGATTCTGTCTT AACCCGGGCCGAACGGCGCGAAAATCGCCGTATGATGTGTTGTTGCTCCCGGGCT AAGACTGAGCGGTTAGTCTTAGATTTATGA
RbxBM3CYP450	ATGGAGCGTGAATTGCGTGAAACCGTAACGGCTCGATGCCCCGTTGCTCACGATG CATCCTCCATCAACGGATGTCCTGTTCAGCGGATGGCTCAAGAATTTGACCCCTT TCAGCACGACTACATGGCCGACCCTGCGCAGTTCGTACGTTGGGCACGAGAACAG GCTCCCATCTTCTACGCTCCTAAACTCGATTACTGGGTCGTTACCCGGTATGATA CCATCAAACAGATCTTCCGTGACCCAGTGACTTTTAGCCCAGCAAATGTTTTGCA AAGCTTCGCGCAACCAAGTGCGGAAGTGCGACAAGTACTCGAACGTTACGGATAT GCGTTTAATCGCACCCTCGTGAACGAGGACGAACCAATGCACCTAGAGCGCCGCC GCGTGCTAATGGAACCCTTCACACCCGAACACCTAGCCGAGCATGAGCCGATGGT GCGGGAGCTAGTGCGACGAGCAGTAGATCGCTTTATCGACGCGGGCCGCGCGGAT TTAGTGGATCAGATGATCTGGGAAGTCCCATTCACTGTCGCCTTACATTTTCTGG GGGTGGACGATGATGATCGGGAAAAGATGCGACGGTTCGCTATCGCTCACACGGT GAATGCGTTTGGTCGACCCAGTCCTGAAGAGCAACTGGCTGTTGCGGAGACGGTC GGACAGTTCTGGCAGTTCTGCGGTGAAGTCTTGGAGAAAATGCGACGCACAGCGG ATGGGCCCGGCTGGATGCGGTATAGCATCCGCCAGCAAAAGTTGTACCCAGAGGT TGTTACAGACTCTTACCTCCACAGTATGATGCAGGCGATTATCGTAGCTGCACAT GAGACGACCGTATTTGCCACGACAAACGCTTTAAAGACGCTATTGGAGCATGAGG CAGTATGGCGCGAAATTTGCGCCGACCCCTCCTTGATTCCTGCCGCTGCAGAAGA ATGCTTACGCTATAACGGCCCTGTGGCTGGATGGCGTCGGCGGACCACGCGAGAG GTTGAAGTCGAGGGTGTTCGTCTACCAACCGGAGCGAATATTCTAATGATCGTCG CGTCTGCAAATCATGATCCAGCTCACTTCGACGATCCTGAGTTTTTCGATATCGG

	GCGAGCAAACGCGTCTGAACATTTAAACTTCGGATACGGTGCTCATCAATGCTTG GGCAAAAATTTAGGCCGTATGGAGATGCAGATCATGATCGAGGAATTGTCTCGTC GTTTACCCCACATGCGGCTAGCGGAACAGCGTTTCGACTACCTGCATAATGTCAG TTTCCGAGCTCCTCGACACTTATGGGTTGAGTGGGACCCTGCTCAGAATCCAGAA CGCCGCGATGCTAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGAAGG CAGAGAATGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGGAAC GGCAGAGGGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTCGCC CCCCAAGTCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAGCTG TTTTGATTGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACAGTT CGTCGACTGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTACTCC GTGTTTGGGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCGCGT TTATCGATGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGGGGA AGCAGATGCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCACATG TGGAGCGACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATAATA AGTCCACGTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGCCAA GATGCACGGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAGCCC GGCAGTGCCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAAGTT ACCAAGAAGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGTCAA CCGCGTCACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAAGCA GAAGAGGAAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGGAGC TACTGCAATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGCGAT GGCTGCTAAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTCGAG AAACAAGCATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAATTAT TAGAAAAATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACTCCC TTCCATCCGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAGAAG CAAGCTTCCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATGGGG AATACAAGGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATACTAT TACTTGTTTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCGGAG ACTCCACTCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGTTCG TTCAGGCACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACACCT GTATTTCGGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTGGAA AATGCACAGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGCCGA ATCAGCCTAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCTCAT CGAACTCTTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAGATG GCTCCAGCGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAGTGT CCGAGGCGGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTATGC CAAAGATGTATGGGCAGGTTAA
TtxBM3CYP450	ATGGAAACTGAGCTAAAGGAGACTGCTCGGGGTACGTGTCCAGTTGCCCACGGGG GACAGTCCTCTGTAGGCGGATGTCCGGTCCATCGCCTAGCGGAGGATTTCGATCC GTTTCAAGATGCCTATATGGCAGACCCTGCACAATTCGTACGGTGGGCCCGAGAG CAGGTTCCTATTTTCTACGCACCGAAATTAAACTACTGGGTAGTGACCCGTTACG ACACAATTAAGCAAATTTTTCGCGACCCGGTCACTTTTAGCCCGAGCAACGTGCT CCAATCCTTTGCCCAACCTAGCGCCGAGGTGCGCCAGGTCCTGGAGCGTTACGGA TATGCCTTTAACCGTACTTTGGTGAACGAGGATGAACCTATGCATCTGGAACGCC GCCGAGTATTAATGGAACCGTTTGCCTCCGAGCACCTGGCAGAGCACGAGCCGAT GGTTCGTGAGCTAGTGCGTCGAGCCGTCAACCGTTTCATTGATACGGGCCGGGCA GACCTCGTGGACCAAATGATCTGGGAAGTGCCTTTTACGGTGGCTCTACATTTCC TGGGAGTAGACGACGATGATCGTGAGAAGATGCGTCGGTTCGCAATCGCACATAC GGTTAATGCATTCGGACGTCCATCTCCAGAAGAACAATTAGCCGTAGCAGAAACC GTGGGCCAGTTCTGGCAATTTTGCGGTGAAGTCTTAGAGAAAATGCGCCGTACCG CGGACGGCCCTGGATGGATGCGCTATAGTATTCGTCAACAGAAGCTATACCCCGA CGTCGTCACTGACTCCTATCTGCATTCCATGATGCAAGCAATTATCGTTGCTGCA CACGAAACGACCGTTTTTGCCACCACAAATGCACTCAAGACTCTACTAGAGCACG AAACTGTATGGCGCGAAATTTGCGCCGATCCATCCCTGATCCCCGCAGCAGCCGA GGAGTGTTTACGCTACAATGGTCCTGTTGCAGGGTGGCGACGCCGGACCACACGA GAGGTCGAGGTCGAGGGCGTACGTCTGCCAGTGGGGGCAAACATCTTGATGGTTG TAGCCTCCGCAAACCACGACTCTGCTCATTTCGACGATCCTGAATTTTTTGATAT TGGACGTTCCAACGCTTCTGAGCATCTAAATTTTGGTTACGGCGCACATCAGTGC CTAGGGCGAAACCTAGGACGAATGGAGATGCAGATTATGATTGAAGAGCTGTCTC GACGCCTCCCCCATATGCGGCTAGCCGAGCAGCGATTTGATTATTTACATAACGT TTCCTTCCGGGCGCCTCGGCATCTGTGGGTTGAGTGGGATCCAGCCCAAAATCCC

	GAGCGTCGCGACCCCAGCCCAAGCACTGAGCAGTCCGCGAAGAAAGTTCGCAAGA AGGCAGAGAATGCGCACAATACGCCATTACTAGTTCTCTACGGCTCCAATATGGG AACGGCAGAGGGTACAGCACGCGATTTAGCTGACATTGCTATGTCTAAAGGATTC GCCCCCCAAGTCGCCACCTTAGACTCCCATGCAGGAAACTTGCCCCGCGAGGGAG CTGTTTTGATTGTGACAGCCTCCTACAACGGTCACCCGCCAGATAATGCTAAACA GTTCGTCGACTGGTTGGACCAAGCTAGTGCGGACGAAGTAAAGGGAGTACGCTAC TCCGTGTTTGGGTGCGGCGATAAAAATTGGGCCACGACCTACCAAAAGGTACCCG CGTTTATCGATGAGACGCTCGCGGCTAAGGGTGCCGAAAACATCGCGGATCGGGG GGAAGCAGATGCGAGCGACGATTTTGAGGGTACGTATGAAGAGTGGCGGGAGCAC ATGTGGAGCGACGTTGCGGCTTACTTTAACTTAGACATCGAGAATTCTGAAGATA ATAAGTCCACGTTAAGTCTACAATTCGTAGATAGCGCTGCGGACATGCCTCTCGC CAAGATGCACGGCGCATTTAGTACAAATGTAGTGGCCAGTAAAGAACTACAACAG CCCGGCAGTGCCCGGTCCACGCGACATCTCGAGATCGAGTTGCCTAAGGAGGCAA GTTACCAAGAAGGAGATCACTTAGGGGTAATCCCACGGAACTACGAAGGTATCGT CAACCGCGTCACGGCACGATTCGGTTTAGATGCAAGCCAGCAAATTCGTCTCGAA GCAGAAGAGGAAAAGTTAGCACATCTACCACTAGCAAAAACAGTATCCGTTGAGG AGCTACTGCAATATGTAGAACTACAGGATCCGGTTACGCGGACCCAGCTGCGCGC GATGGCTGCTAAGACGGTGTGTCCTCCTCACAAGGTCGAGCTAGAGGCACTCCTC GAGAAACAAGCATATAAGGAGCAAGTCTTGGCAAAGCGCCTGACGATGTTGGAAT TATTAGAAAAATACCCAGCATGTGAGATGAAATTCAGCGAATTTATTGCGTTACT CCCTTCCATCCGACCCCGTTATTACAGCATTAGCTCCTCCCCACGTGTTGATGAG AAGCAAGCTTCCATCACGGTGTCCGTAGTTAGTGGTGAAGCGTGGAGCGGATATG GGGAATACAAGGGAATCGCTAGCAATTATTTGGCAGAACTACAAGAGGGCGATAC TATTACTTGTTTCATTAGCACGCCCCAATCCGAGTTCACGCTCCCAAAGGATCCG GAGACTCCACTCATTATGGTCGGACCCGGTACGGGTGTCGCGCCTTTTCGTGGGT TCGTTCAGGCACGGAAGCAACTGAAGGAGCAGGGCCAGTCTTTGGGAGAAGCACA CCTGTATTTCGGCTGTCGTAGTCCACACGAGGATTATTTATATCAGGAAGAGTTG GAAAATGCACAGTCTGAGGGAATTATCACACTCCATACAGCTTTTAGCCGGATGC CGAATCAGCCTAAGACCTATGTACAACACGTTATGGAGCAGGATGGGAAAAAGCT CATCGAACTCTTAGATCAGGGGGCACACTTCTACATCTGCGGCGATGGTTCCCAG ATGGCTCCAGCGGTAGAGGCTACACTAATGAAATCTTATGCCGACGTACATCAAG TGTCCGAGGCGGATGCTCGCCTATGGCTCCAGCAGCTCGAAGAAAAAGGTCGCTA TGCCAAAGATGTATGGGCAGGTTAA
GapN	ATGTTCGAGAATATTTCGTCGAATGGGGTATACAAGAACTTATTTGACGGTAAGT GGGTAGAAAGTAAGACGAACAAAACTATTGAAACGCACTCCCCGTACGACGGTTC CTTAATTGGAAAGGTACAGGCACTGAGTAAAGAAGAGGTCGATGAGATTTTTAAA TCGTCGCGCACTGCGCAAAAAAAATGGGGCGAGACTCCCATCAATGAGCGTGCGC GTATTATGCGCAAGGCCGCCGACATTCTGGACGATAATGCAGAGTATATTGCGAA GATCTTAAGCAATGAAATCGCAAAAGACTTAAAGTCGAGCCTGTCTGAGGTTAAA CGCACGGCGGATTTTATTCGTTTCACCGCCAACGAAGGTACGCACATGGAAGGGG AGGCCATTAACTCGGACAATTTTCCCGGCTCTAAAAAAGACAAACTGTCGTTGGT TGAACGCGTACCCTTGGGAATTGTGCTTGCGATTAGTCCGTTCAATTATCCGGTG AATCTGAGCGGATCTAAAGTAGCTCCGGCCTTAATTGCCGGTAACAGTGTAGTAC TTAAGCCAAGTACCACAGGGGCCATCAGTGCATTACATCTTGCGGAGATTTTTAA CGCAGCCGGTCTTCCAGCAGGGGTCTTGAATACAGTTACGGGGAAGGGCTCAGAG ATTGGCGATTACTTAATCACGCATGAGGAAGTCAATTTTATCAATTTTACGGGAA GTAGTGCTGTGGGCAAACATATTTCCAAAATTGCAGGTATGATCCCTATGGTTCT TGAGCTTGGAGGAAAAGACGCTGCGATCGTTCTGGAGGATGCCAACTTGGAGACT ACAGCGAAATCCATCGTATCGGGCGCCTACGGCTATAGCGGTCAGCGCTGTACTG CTGTAAAGCGCGTTCTGGTGATGGACAAAGTGGCGGATGAACTTGTAGAATTAGT AACAAAAAAGGTTAAAGAATTAAAGGTCGGTAATCCTTTCGACGACGTCACAATT ACTCCACTTATCGACAATAAAGCTGCCGATTACGTTCAGACACTGATTGACGACG CGATCGAAAAGGGCGCGACTTTGATTGTAGGCAACAAGCGCAAAGAGAATTTGAT GTACCCGACATTATTTGATAATGTGACGGCGGACATGCGCATTGCGTGGGAAGAG CCCTTCGGGCCGGTGTTGCCTATCATTCGTGTCAAATCAATGGACGAGGCTATTG AATTAGCAAATCGCAGCGAATACGGATTACAAAGCGCAGTCTTCACAGAAAATAT GCACGATGCGTTCTACATTGCGAACAAATTGGATGTAGGGACGGTCCAGGTAAAC AACAAGCCAGAACGTGGCCCCGACCATTTTCCCTTTTTGGGCACAAAAAGCTCGG GTATGGGGACTCAAGGGATTCGTTATTCTATCGAGGCTATGACACGTCACAAAAG TATCGTCCTTAACCTTTAA

BmGDH2	ATGTACACGGACCTGAAAGATAAGGTCGTTGTCGTCACTGGGGGGAGTAAAGGGC TGGGTCGAGCTATGGCAGTTCGTTTCGGGCAAGAGCAAAGCAAGGTTGTGGTTAA TTACCGGTCCAATGAAGAAGAAGCCTTAGAGGTGAAAAAAGAGATCGAACAGGCT GGCGGTCAGGCTATCATCGTCCGTGGAGATGTAACAAAGGAGGAAGATGTAGTCA ATTTGGTAGAAACCGCAGTGAAAGAGTTTGGCACCCTGGATGTAATGATCAATAA CGCTGGCGTCGAGAATCCTGTGCCCAGTCACGAATTGTCTCTCGAAAACTGGAAT CAAGTTATCGATACAAACCTGACGGGGGCATTCTTGGGATCTCGGGAGGCTATTA AGTATTTTGTCGAGAACGACATTAAAGGAAATGTTATTAACATGAGCAGCGTCCA CGAAATGATCCCCTGGCCGTTATTTGTTCACTACGCCGCTAGTAAGGGAGGTATG AAGTTAATGACCGAGACGCTGGCGTTAGAGTATGCACCTAAAGGTATCCGGGTTA ATAATATCGGACCAGGTGCGATCGACACACCAATTAACGCGGAGAAGTTTGCTGA CCCTGAACAACGTGCCGACGTAGAAAGTATGATTCCAATGGGCTATATCGGTAAT CCGGAGGAGATTGCATCTGTCGCAGCCTTTTTAGCTTCTTCCCAGGCGTCTTACG TAACGGGTATTACACTATTTGCTGATGGCGGAATGACCAAATATCCCTCCTTCCA AGCAGGGCGCGGATGA
BmGDH3	ATGTACACGGACCTGAAAGATAAGGTCGTTGTCATTACTGGGGGGAGTACCGGGC TGGGTCGAGCTATGGCAGTTCGTTTCGGGCAAGAGGAAGCCAAGGTTGTGATTAA TTACTATAATAATGAAGAAGAAGCCTTAGATGCCAAAAAAGAGGTGGAACAGGCT GGCGGTCAGGCTATCATCGTCCGTGGAGATGTAACAAAGGAGGAAGATGTAGTCA ATTTGGTACAAACCGCAATTAAAGAGTTTGGCACCCTGGATGTAATGATCAATAA CGCTGGCGTCGAGAATCCTGTGCCCAGTCACGAATTGTCTCTCGATAACTGGAAT AAAGTTATCGATACAAACCTGACGGGGGCATTCTTGGGATCTCGGGAGGCTATTA AGTATTTTGTCGAGAACGACATTAAAGGAAATGTTATTAACATGAGCAGCGTCCA CGAAATGATCCCCTGGCCGTTATTTGTTCACTACGCCGCTAGTAAGGGAGGTATG AAGCAAATGACCGAGACGCTGGCGTTAGAGTATGCACCTAAAGGTATCCGGGTTA ATAATATCGGACCAGGTGCGATGAATACACCAATTAACGCGGAGAAGTTTGCTGA CCCTGAACAACGTGCCGACGTAGAAAGTATGATTCCAATGGGCTATATCGGTAAT CCGGAGGAGGTGGCAGCCGTCGCAGCCTTTTTAGCTTCTTCCCAGGCGTCTTACG TAACGGGTATTACACTATTTGCTGATGGCGGAATGACCAAATATCCCTCCTTCCA AACCGGGCGCGGATGA
PoAlkK	ATGTTAGGACAAATGATGCGCAATCAGCTCGTAATCGGGTCCCTCGTTGAACACG CGGCCCGATATCATGGGGCACGCGAGGTAGTTTCCGTAGAAACGTCCGGAGAAGT CACACGCTCTTGCTGGAAAGAAGTAGAGCTACGAGCGCGGAAATTAGCGTCTGCG CTCGGCAAAATGGGACTGACGCCAAGCGACCGCTGTGCGACTATCGCATGGAATA ACATTCGACATCTCGAGGTGTACTATGCTGTGTCTGGGGCGGGGATGGTGTGCCA TACGATCAATCCCCGCCTGTTCATTGAGCAAATCACATATGTAATTAACCACGCA GAGGATAAAGTGGTTCTGTTAGACGACACCTTCCTCCCCATTATCGCTGAGATCC ATGGCAGTCTGCCGAAAGTCAAAGCTTTTGTATTGATGGCGCATAATAACAGCAA TGCTTCCGCCCAGATGCCGGGGCTCATCGCCTACGAAGATTTAATTGGACAAGGG GATGATAATTACATTTGGCCGGACGTCGATGAAAACGAGGCGTCCTCCTTATGCT ATACTAGCGGGACTACGGGGAACCCGAAAGGGGTCTTATACAGCCACCGCTCCAC CGTACTACATTCCATGACCACCGCGATGCCAGACACCCTAAATCTCAGCGCCCGA GATACCATTCTACCCGTGGTCCCCATGTTTCATGTAAACGCCTGGGGCACTCCAT ATTCCGCGGCGATGGTGGGTGCTAAGTTAGTACTCCCAGGTCCTGCCTTGGATGG AGCATCTCTAAGTAAATTAATTGCGAGTGAGGGAGTGAGCATCGCACTGGGAGTT CCGGTAGTGTGGCAGGGTTTATTGGCCGCTCAGGCGGGGAACGGAAGTAAGTCTC AATCTCTGACTCGGGTCGTGGTCGGAGGTTCCGCCTGCCCGGCTAGCATGATTCG TGAATTTAACGATATCTATGGGGTCGAGGTGATCCATGCTTGGGGCATGACCGAG CTGAGCCCATTTGGCACTGCCAACACCCCATTGGCCCACCACGTAGACTTATCTC CAGACGAGAAACTGTCTTTACGCAAATCCCAGGGCCGTCCACCTTATGGGGTTGA GCTAAAGATTGTGAATGACGAAGGGATCCGGTTACCCGAAGACGGTCGTAGCAAA GGTAATCTCATGGCGCGCGGACACTGGGTAATTAAGGACTATTTTCACAGTGACC CTGGTAGCACCTTATCCGACGGTTGGTTTAGCACGGGTGACGTCGCAACTATTGA TTCTGATGGATTTATGACCATTTGTGACCGGGCCAAGGACATTATTAAATCCGGT GGCGAGTGGATTAGTACTGTCGAATTAGAGAGCATTGCAATCGCTCACCCACACA TTGTGGACGCCGCCGTAATCGCCGCGCGCCACGAAAAGTGGGACGAGCGGCCCCT CTTAATCGCTGTCAAGTCTCCAAACAGCGAACTGACTTCTGGGGAGGTCTGCAAC TACTTCGCGGACAAAGTAGCTCGCTGGCAGATCCCAGATGCTGCCATTTTTGTGG AGGAGCTACCTCGTAATGGCACAGGCAAGATTCTGAAAAATCGTCTCCGGGAGAA GTACGGTGACATTTTACTCCGATCCTCCTCTTCCGTTTGCGAATAA

Pp0763	ATGTTGCAGACACGCATCATCAAGCCCGCCGAGGGCGCCTATGCCTATCCATTGC TGATCAAGCGCCTGCTGATGTCCGGCAGCCGCTATGAAAAGACCCGGGAAATCGT CTACCGCGACCAGATGCGGCTGACGTATCCACAGCTCAACGAGCGCATTGCCCGC CTGGCCAACGTGCTGACCGAGGCCGGGGTCAAGGCCGGTGACACCGTGGCGGTGA TGGACTGGGACAGCCATCGCTACCTGGAATGCATGTTCGCCATCCCGATGATCGG CGCTGTGGTGCACACCATCAACGTGCGCCTGTCGCCCGAGCAGATCCTCTACACC ATGAACCATGCCGAAGACCGCGTGGTGCTGGTCAACAGCGACTTCGTCGGCCTGT ACCAGGCCATCGCCGGGCAGCTGACCACTGTCGACAAGACCCTGCTACTGACCGA TGGCCCGGACAAGACTGCCGAACTGCCCGGTCTGGTCGGCGAGTATGAGCAGCTG CTGGCTGCTGCCAGCCCGCGCTACGACTTCCCGGATTTCGACGAGAATTCGGTGG CCACTACCTTCTACACCACTGGCACCACCGGTAACCCCAAGGGCGTGTATTTCAG TCACCGCCAGCTGGTGCTGCACACCCTGGCCGAGGCCTCGGTCACCGGCAGTATC GACAGCGTGCGCCTGCTGGGCAGCAACGATGTGTACATGCCCATCACCCCGATGT TCCACGTGCATGCCTGGGGCATCCCCTACGCTGCCACCATGCTCGGCATGAAGCA GGTGTACCCAGGGCGCTACGAGCCGGACATGCTGGTCAAGCTTTGGCGTGAAGAG AAGGTCACTTTCTCCCACTGCGTGCCGACCATCCTGCAGATGCTGCTCAACTGCC CGAACGCCCAGGGGCAGGACTTCGGCGGCTGGAAGATCATCATCGGCGGCAGCTC GCTCAACCGTTCGCTGTACCAGGCCGCCCTGGCGCGCGGCATCCAGCTGACCGCC GCGTATGGCATGTCGGAAACCTGCCCGCTGATCTCCGCGGCACACCTGAACGATG AACTGCAGGCCGGCAGCGAGGATGAGCGCGTCACTTACCGTATCAAGGCCGGTGT GCCGGTGCCGTTGGTCGAAGCGGCCATCGTCGACGGCGAAGGCAACTTCCTGCCC GCCGATGGTGAAACCCAGGGCGAGCTGGTACTGCGTGCGCCGTGGCTGACCATGG GCTACTTCAAGGAGCCGGAGAAGAGCGAGGAGCTGTGGCAGGGCGGCTGGCTGCA CACCGGTGACGTCGCCACCCTCGACGGCATGGGCTACATCGACATCCGCGACCGC ATCAAGGATGTGATCAAGACCGGTGGCGAGTGGGTTTCCTCGCTCGACCTGGAAG ACCTGATCAGCCGCCACCCGGCCGTGCGCGAAGTGGCGGTGGTGGGGGTGGCCGA CCCGCAGTGGGGTGAGCGCCCGTTTGCCCTGCTGGTGGCACGTGACGGCCACGAT ATCGACGCCAAGGCGCTGAAGGAACACCTCAAGCCATTCGTCGAGCAAGGTCATA TCAACAAGTGGGCGATTCCAAGCCAGATCGCCCTTGTTACTGAAATTCCCAAGAC CAGTGTCGGCAAGCTCGACAAGAAACGCATTCGCCAGGACATCGTCCAGTGGCAG GCCAGCAACAGCGCGTTCCTTTCCACGTTGTAA
PpAAt1	ATGGGATCCTTGTGCCCACTCCTATTTCCTGTGAACCGATTCGAACCTGAGCTAA TTACCCCCGCAAAACCTACCCCCATCGAGACGAAGCAACTGAGTGATATTGATGA TCAAGACGGGTTGCGATTTCACTTCCCTGTTATTATTAGCTATAAAAACAATCCG AGTATGAAGGGCAACGACGCAGTGATGGTAATCCGCGAGGCTTTATCTCGTGCAC TCGTGTATTACTACCCACTGGCGGGGCGCCTACGTGAAGGGCCTAATCGGAAGCT CATGGTTGAGTGTAACGGGGAGGGCGTGCTGTTCATCGAGGCGAATGCCGATGTA ACGCTAGAGCAACTAGGTGACCGGATCCTACCGCCGTGCCCCGTTTTAGAAGAGT TTCTAAGTAATCCACCAGGTTCCGACGGTATTTTGGGATGCCCTTTACTGCTAGT ACAGGTTACCCGTCTCACGTGTGGCGGCTTCATCTTCGGTTTACGTATCAATCAT GCGATGTGTGACGCAGTCGGCCTAGCAAAATTTCTGAATGCGATCGGTGAAATGG CACAAGGGGCTGATAGTTTAAGTGTGCCGCCGGTATGGGCGCGGGAGCTACTCAA CGCACGTGACCCACCGACGGTGACCCGTTGGCACTACGAATACGATCAACTACTC GATTCTCAGGGGAGTTTTATTGCCGCCATCGACCAAAGCAACATGGCACAGCGTT CCTTCTACTTCGGTCCTCAGCAAATTCGGGCGTTGCGTAAGCATTTGCCGCCGCA CTTGTCCACCTGTAGCAGTTTTGAGTTGATCACCGCATGTGTGTGGCGATGTCGA ACTTTAAGCCTACGGTTGAATCCTAAAGATACCGTGCGTATCTCCTGTGCTGTCA ATGCCCGGGGTAAATCTATCAACGATTTGTGTCTACCCAGCGGCTTTTACGGTAA TGCTTTTTCTATTCCTACTGCAGTTTCCACCGTGGAGTTGCTCTGTGCATCTCCA CTAGGCTACGGTGTCGAGCTAGTGCGCAAAAGTAAGGCTCAAATGGATAAAGAGT ATATGCAGTCCTTAGCTGATTTTTTCGTTATTCGGGGACGACCCCCACTGCCTAT GGGATGGAACGTCTTTATCGTCTCCGATAACCGACATACCGGGTTCGGGGAATTC GACGTAGGTTGGGGTCGGCCACTATTTGCTGGGCTAGCACGAGCTTTCAGTATGA TTTCCTTCTATGTCCGGGATAACAATCAAGAAGAGGAATTTGGAACGCTGGTCCC GATCTGTCTACCGAGTACAAGTCTGGAACGGTTCGAGGAAGAACTGAAAAAAATG ACATTGGAACACGTAGAGGAAATCAGTAAGTGA
Atf1	ATGAACGAGATCGATGAGAAGAATCAAGCACCGGTCCAGCAAGAGTGTTTGAAGG AGATGATTCAAAATGGGCATGCACGTCGCATGGGTTCTGTCGAAGATTTGTATGT AGCACTGAACCGTCAGAACTTATACCGTAATTTTTGCACCTACGGGGAGTTATCT GACTATTGTACCCGTGATCAATTGACTCTTGCCTTACGTGAAATCTGCCTGAAGA ATCCCACATTGTTACACATCGTACTGCCAACCCGTTGGCCAAATCACGAGAATTA

	CTATCGCTCCAGTGAGTACTACTCACGCCCTCATCCCGTGCACGATTATATCTCT GTGCTTCAAGAACTGAAATTATCAGGCGTAGTCCTTAATGAACAACCGGAGTATT CTGCGGTTATGAAGCAGATTTTGGAGGAGTTTAAGAATAGTAAGGGCTCTTATAC AGCGAAAATCTTTAAATTAACTACTACTTTAACAATTCCGTATTTCGGTCCTACG GGGCCCTCATGGCGTCTGATTTGCTTACCGGAAGAGCACACTGAAAAATGGAAAA AGTTCATCTTCGTTTCTAATCATTGTATGTCCGATGGCCGCTCAAGTATCCATTT CTTCCACGACCTGCGTGATGAATTGAATAATATTAAGACACCACCTAAAAAGTTG GATTACATCTTTAAGTATGAGGAAGACTACCAGTTACTGCGTAAATTGCCCGAGC CTATCGAGAAAGTAATTGATTTTCGTCCCCCGTACCTTTTCATCCCGAAATCTCT TTTAAGCGGTTTTATCTATAATCACCTGCGTTTCAGTTCGAAGGGAGTGTGTATG CGTATGGACGATGTGGAGAAGACGGACGATGTTGTCACCGAAATTATCAACATCA GCCCTACGGAATTTCAAGCTATCAAGGCCAATATTAAATCTAATATTCAAGGCAA ATGCACAATCACTCCGTTTTTGCATGTTTGCTGGTTTGTCTCACTGCACAAATGG GGAAAATTCTTTAAACCATTAAATTTCGAGTGGCTTACTGACATTTTCATTCCAG CCGACTGCCGTTCTCAGTTACCTGACGATGATGAGATGCGTCAGATGTACCGTTA TGGAGCAAACGTTGGATTCATCGACTTCACACCGTGGATTTCCGAGTTCGACATG AATGACAACAAGGAAAATTTTTGGCCCCTTATCGAACATTACCATGAAGTAATTT CGGAAGCACTGCGCAATAAGAAGCACCTGCATGGGTTAGGCTTCAATATTCAGGG GTTCGTCCAGAAATATGTAAATATCGATAAGGTGATGTGTGATCGTGCCATTGGT AAGCGTCGTGGCGGAACACTGCTGTCCAACGTCGGTTTGTTCAACCAGTTAGAAG AACCTGATGCGAAATATTCAATCTGTGACTTAGCTTTCGGGCAATTTCAGGGATC GTGGCATCAGGCTTTTAGCTTGGGCGTATGTTCAACCAATGTAAAAGGCATGAAC ATCGTTGTTGCTAGTACAAAAAATGTGGTTGGCTCTCAAGAAAGTCTTGAGGAAC TTTGCTCTATTTATAAAGCCCTGTTGTTGGGGCCCTAA
CAT	ATGAGCAGCGGCATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCC AATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTA TAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAAT AAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTC ATCCGGAGTTCCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGT TCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGG AGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTCGCAAGATGTGG CGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTT TTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGTGGCC AATATGGACAACTTCTTCGCCCCCGTTTTCACTATGGGCAAATATTATACGCAAG GCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTCTGTGATGG CTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAG GGCGGGGCGTAA
SAAT	ATGGAGAAGATTGAAGTGTCGATTAATTCGAAACACACCATCAAGCCGAGCACCT CCTCCACCCCTCTTCAACCGTACAAACTCACGTTACTGGACCAATTAACGCCGCC AGCATATGTTCCAATTGTGTTTTTTTATCCGATTACAGATCACGACTTTAACCTT CCGCAGACTTTGGCCGATCTGCGTCAAGCGCTGAGCGAAACACTGACGCTGTATT ATCCACTGAGCGGCCGCGTCAAGAACAACCTGTACATTGACGATTTTGAAGAGGG GGTGCCGTACCTGGAAGCTCGTGTGAATTGCGACATGACGGATTTCCTCCGTCTG CGTAAGATCGAGTGCTTAAACGAATTTGTCCCTATTAAGCCGTTTAGTATGGAAG CAATTTCGGATGAGCGCTATCCGCTGCTTGGGGTTCAAGTAAACGTATTTGATTC CGGTATTGCGATCGGGGTCAGTGTCTCCCATAAACTGATCGACGGCGGGACCGCT GATTGCTTCCTGAAGTCGTGGGGAGCGGTGTTCCGCGGATGCCGCGAGAACATTA TCCATCCGTCCCTCTCGGAAGCAGCATTGCTTTTTCCGCCACGTGACGACCTCCC GGAAAAATATGTAGATCAAATGGAAGCGTTATGGTTCGCGGGTAAAAAGGTTGCG ACCCGTCGTTTCGTGTTTGGTGTGAAAGCTATCTCATCTATTCAGGATGAAGCAA AATCTGAGAGCGTACCAAAGCCGAGTCGCGTGCACGCCGTCACGGGTTTTTTGTG GAAGCATCTGATCGCGGCGAGTCGTGCACTGACCTCAGGTACTACTTCTACTCGC CTGAGTATCGCCGCCCAAGCGGTTAATTTGCGCACGCGCATGAATATGGAAACCG TACTGGATAACGCGACGGGTAATCTGTTTTGGTGGGCGCAGGCGATCCTCGAGTT AAGCCATACCACTCCAGAGATCAGCGATCTGAAACTGTGTGATCTGGTGAACCTG TTGAACGGCAGCGTTAAGCAGTGCAATGGGGATTATTTTGAAACCTTTAAAGGTA AAGAAGGTTATGGCCGTATGTGCGAGTACCTTGATTTTCAGCGGACCATGAGTTC CATGGAACCTGCACCGGATATTTACCTGTTTTCTTCTTGGACGAACTTTTTCAAC CCCTTGGATTTTGGCTGGGGCCGCACCTCTTGGATTGGGGTTGCGGGCAAGATTG AAAGCGCATCTTGCAAATTTATTATCCTTGTCCCGACACAGTGTGGCTCAGGCAT

	CGAAGCGTGGGTAAATCTGGAGGAAGAAAAAATGGCGATGCTTGAACAGGACCCC CATTTCCTGGCTCTTGCGTCTCCCAAAACGTTGATTTAA
VAAT	ATGGAAAAAATTGAAGTTTCGATTATCTCAAAACATACGATCAAACCGAGTACCT CTTCGAGCCCCCTGCAGCCTTACAAGTTAACCCTGCTGGACCAGCTTACCCCGCC GAGTTACGTCCCTATGGTGTTTTTCTATCCCATCACTGGGCCTGCAGTTTTTAAC CTGCAGACATTGGCCGACCTTCGCCATGCCTTGTCAGAGACGTTAACCCTTTATT ATCCCTTGTCTGGGCGCGTGAAGAATAATCTGTACATCGACGACTTCGAGGAAGG CGTGCCTTACTTAGAAGCACGTGTTAATTGCGATATGAATGACTTTCTGCGCTTG CCTAAAATCGAATGTCTGAACGAATTCGTCCCCATTAAACCATTTAGCATGGAGG CCATTTCGGATGAACGCTATCCTTTATTAGGTGTGCAGGTTAATATTTTTAATTC AGGAATCGCAATTGGCGTATCGGTTTCACATAAGCTTATCGACGGGCGTACTTCA GATTGCTTCCTGAAGTCCTGGTGTGCAGTTTTCCGTGGGTCGCGCGACAAAATCA TCCATCCGAACTTGTCTCAAGCGGCCTTACTTTTTCCTCCACGCGACGATTTGCC CGAGAAATACGCGCGTCAGATGGAGGGGTTATGGTTCGTAGGTAAGAAAGTTGCC ACACGCCGTTTTGTCTTTGGAGCTAAAGCGATCTCAGTGATTCAAGACGAGGCTA AGAGCGAAAGCGTACCCAAGCCATCCCGTGTACAGGCAGTCACTTCATTTTTATG GAAACACTTAATTGCTACCTCTCGCGCACTGACGTCTGGCACTACAAGCACGCGT CTTAGTATTGCCACACAGGTAGTAAACATTCGCTCTCGTCGCAATATGGAAACGG TCTGGGACAACGCGATCGGCAACTTGATTTGGTTTGCTCCTGCGATCTTAGAGCT GTCCCACACAACTCTTGAGATTTCTGACTTAAAACTGTGCGATTTGGTTAATTTA CTTAATGGCAGCGTTAAGCAATGCAATGGAGACTACTTTGAGACATTCATGGGAA AGGAGGGGTATGGAAGCATGTGTGAATACTTAGACTTTCAGCGCACGATGTCGTC TATGGAACCCGCTCCAGAAATCTACTTATTTACGTCGTGGACAAATTTTTTCAAC CAACTTGATTTCGGTTGGGGGCGTACATCATGGATTGGGGTTGCAGGGAAGATTG AGTCGGCCTTTTGTAACCTTACTACATTGGTCCCGACCCCATGCGACACAGGTAT CGAGGCTTGGGTAAATCTGGAAGAAGAAAAAATGGCAATGCTGGAGCAGGACCCG CAATTTCTTGCTCTTGCATCGCCGAAGACCCTGATCAGCCGTTATTAA
VpAAT1	ATGGCAGAAAAGGCGTCTTCTCTTATGTTTAACGTGCGCCGTCATGAGCCAGAGC TTATCACACCTGCCAAGCCAACACCGCGTGAGATCAAACTGTTGTCGGATATCGA TGACCAAGATGGCCTGCGTTTCCAGGTACCCATTATCCAGTTTTACAAGAATAAT TCCTCGATGCAGGGTAAGAACCCTGCCAAAATTATCAAATCAGCCTTAGCGGAGA CTTTGGTACACTATTATCCACTTGCAGGGCGCTTACGCGAAGGGTTCGGTCGCAA GTTAATGGTCGAGTGTACGGGGGAAGGTATCCTTTTTATTGAAGCTGATGCGGAC GTTACATTACACGAATTCGGCGATGATTTGCCCCCACCTTTTCCGTGCCTTGTGG AGTTGCTGTATGACGTGCCTGGCTCTAGTGGCATTATCGATACACCGCTTTTGTT GATCCAGGTAACACGCTTAAAATGCGGAGGTTTCATTTTCGCCTTGCGCTTAAAC CATACCATGAGCGATGCGAGTGGGCTGGTGCAATTCATGACGGCAGTCGGCGAGA TGGCTCGTGGACAACGTTCTTTGTCCATTCAGCCAGTATGGGAGCGTCATTTATT AAATGCGCGTGATCCACCGCGTGTAACCCATATTCACCACGAATACGATGATTTG GAGGACACTAAAGGTACCATTATTCCATTAGACGATATGGTGCATCGCTCGTTCT TTTTCGGCCCATCTGAGATGGCAGCAATTCGTCGCCTGGTACCGGCGCACTTCCA CCGCTCTACGACTAGCGAGGTTTTAACAGCATATTTGTGGCGCTGCTACACTATT GCGCTTCAACCAGACCCAGAGGAGGAAATGCGTGTAATCTGCGTTGTCAATTCCC GTACCAAACTTAACCCACCCCTTCCTACAGGGTTTTACGGTAATGGAATTGCTTT CCCCGCTGCTATCAGCCAGGCGAAGAAAATTTGTGAAAATCCGTTTGGATATACC TTACAGCTGGTAAAGCAGACCAAAGTTGATGTGACCGAGGAATATATGCGCAGCG CGGCCGATTTGATGGCTATGAAGGGACGTCCGCACTTCACTGTTGTTCGTCGCTA CATGGTAAGTGACGTAACACGTGCAGGGTTTGGATTGGTGGACTTTGGATGGGGC CGTCCTGAACCGGTGTACGGTGGGCCTGCAAAAGGGGGAGTGGGGCCCATTCCGG GAGTAACCTCTTTTTTTGTACCATTCAAAAATCGTAAGGGCGAAAAAGGCATTGT CGTGCCAACATGCCTGCCTACGCCGGCGATGGAACGTTTTGCTAAACTTATGAAT GAAATTCTTCAGAATCAATTACTGGTTTCCGCCGAAGAAAACAAGTCAGTGTTTA TCGTCAGCGCGATCTAA
AtHPFT	ATGCTTGTTTCCCCCTTGCATCCCACGCCAAAACGTTCATTCTTCCTGAGTAACA TTGATCGCATGCTGAATTATAACATTCCCACTGTGTACTTCTTCGCTGCCAATCC TGAATACCCGCCCCCCGTAGCCGCCGCTAATTTGAAACTGGCTCTTCAAAAGTTG CTTGTCCCTTACGATTTTATGGCCGGCCGCTTGAAGATGAACGCTGACAGTGGTC GCTTAGAAATCGACTGCAACGGAGCAGGGGCAGGCTTCGTAGTGGCCTCATCGGC TTTATCACTTGAGCAAATCGGTGACTTAGTGCGCCCTAACTTAGGATTTCGCGAT CTTGCTGTGCAAACTATGATTAAGAAAGATGACCACGACCCCATGTTCATTCTTC AGTTGACTTCGTTTGCCTGTGGAGGTTTCGCGATCGGCTTATGTGTGAATCATAT

	CTTGCTTGATGGTATGAGTGCGAAAGCTTTTAATCAAAATCTGGCATCCCAAGCA TTCCACGACGACCGCCCTCTTGCAGTGGTGCCTTGTTTCGATCGCCGCTTGATGG CCGCCCGCAGTCCGCCGCGCCCTGCTTTCGACCACCCAGAATTCTTCAAGCCCGA CCTGTGGGGGCCAGCGGTTTTTGACTGTGAACGCGAGGAATTAGAATACCGCGTC TTTCAGCTTGATCCCACGCACATTACGTTACTGAAAGAAAAAGCGTTACATTTGC AGGAGTCCGGCAGCCGCATTTCCTCACTGACCGTTGCGGCAGCACTTATCTGGAA GTGTAAGGCCTTGAGCAAGGAGTACAAGGATGATAAGGACAAAGTATCAACGTTG CTTAACGTTATGGACCTGCGTGCTCGCTTAAAACCACCTTTACCTGCAGACTACT GTGGCAATGCGTTGTTGGTCGCTTATGCAACCGGCCGCTGCGGCGAACTGGAATT TTGGGAGGTTGCGAAAATGGTTGGCGAAGGGCCCCGCCGCGTTACCGATGAATAC GCAAAGTCTGCCATTGACTGGTTGGAGATTAACAAGGGCAAAGGCGTGCCGCGTG GAGATTACATGGTGAGTTCTTGGTTACGCTTGGGGTTTGAGGACGTTGTTTTCCC CTGGGGAAAGGCTTTGCACTCAGGCCCATTGGTCAGCCATCGCAAAGACATTTGC TGGTTATTTCCAACCCCTCACGGGTAA
CcVS	ATGGAGATCGGTATTGTATCGCGCGAAGTGGTCCGCCCCTCTTCATTGAATGTAC ACTTGCTGAAGCCTTTCCGCATTTCCTTACTTGACCAGCTTACGCCGACTACGTT TGCTCCTCTGGTCCTTTTCTACCCGATGAAGTCAACACACTTAAAAGGAACCCAG ATCTCGACGCAGCTGAAGGAGTCGTTGAGTAAGACCCTTGAACACTTCTACCCAT TGGCGGGTCGTGTCCGCGACAACTTAATTATTAACGATTATGATGAGGGCGTGCC ATACATTGAAACTCGTGTTAACACGCATCTTTTTGAGTTTTTACAAAATCCCCCC ATGGAGCTTTTAAATCAATGCCTGCCGTATCCACCATTCCGTTACCAGCCTAACC CTGACCGTGTTCCGCAGGTAGCAGTACAACTTAACACTTTTGATTGCGGCGGTAT CGCCTTGGGACTGAGTTTTTCCCACAAAATCAACGATGGCGCGACCACTTCGGCG TTTTTACGTTCATGGGCAGCTAACAGCCGTGGAGCCTGTCACAAAGCCGTAAAGT ACAAAAACTTATCGGAGGCATCTATGATTTTCCCCCCACAAAACCCGAGTCCCAA CCACCACCTGTCGGTTATGGAAAAGATTTGGTTCCGTGAGGCGAAATACAAGACC CGCCGCTTCGTATTTGACGCAAAAGCGATTGCGACCTTGCGCTCAGAATGCAAGG GAGAACGTGTTCCGAATCCAACGCGTATCGAAGCTCTGTCTGCGTTCATTCTTAA ATCCGCAATGCTTGCGTCTCGCTCTACAGCTTCATCACGTTTTGTTCTGCACCAA GCTGTGAACTTACGTCGCCTGACCGAACCCCGCCTTTCTCCCTGTAGCGTGGGTA ACCTGTTCTTGTGGGCGACTGCAGCATATAACATGGAGCACGCCGCCGAAATGGA GTTGCACGGCTTGGTAGCCCGCATGAAACAGGCAGTAGGAAAGATCAATTCTGAG TATCTGAAGACTTTACACGGGGACGAAGGCTTCCCAAAAGTTTGTGAATATATCA AACGCATCGAAGAGGTAAGTGCCCACAAAAATCTTGAGGCATTCACGTTTTCATC CTGGGTTAAGTTCGGCTTCAACGAGGTTGACTTTGGGTGGGGAAACCCCATTTGG TCAGGAATTTTCGGGGAGGTCGGGAGTAACAGCTTTCGCAACTTAACCTTTTTTA AGGAGACACGCTCGGCGAACTACGACAACGCCGTTGAGGCTTGGGTGACTCTTGA CGAAAAAATCATGTCGTTACTGGAACATGATCCCCAGTTTTTAGCTTTTGCATCA CCTAACCCCAGTATTCTGTTATTTGATGTCTGCCGTTTCTAA
CsVS	ATGGAGATCGGAATTGTCTCACGTGAAGTCGTGCGTCCTTCTTCGTTAAACGTCC ATTTGCTTAAACCCTTCAAGATTTCCCTTTTAGACCAACTGACGCCGACGACGTT CGCGCCTTTGGTATTGTTTTACCCCATGAAGAGCACCCATTTAAAAGGGACACAA ATTTCCACTCAATTGAAGGAGAGTCTTAGCAAGACGCTGGAGCACTTCTACCCTT TGGCAGGCCGCGTTCGCGATAATCTTATCATTAACGATTATGATGAAGGCGTCCC GTACATCGAGACGCGCGTGAATACACATTTGTTTGAATTTTTACAAAACCCTCCA ATGGAGTTATTAAACCAATGTTTACCATACCCCCCCTTCTCGTATCAGCCTAATC CCGATCGCGTACCCCAGGTCGCTGTTCAGCTTAATACTTTTGATTGTGGCGGGAT CGCGTTAGGGTTATGTTTCTCCCATAAGATTAACGATGGCGCGACAACGTCCGCC TTTCTTCGCAGTTGGGCCGCCTTTTCTCGCGGGGCTTACCATAAAGCAGTGAAGT ACAAGAACATTTCCGAAGCGTCTATGATTTTCCCCCCACAAAACCCTTCCCCGAA TCACCGCTTGTCCGTGATGGAGAAAATCTGGTTTCGTGAGGCGAAATATAAAACC CGTCGTTTCGTTTTCGACGCGAAGGCTATCGCGACACTGCGCAGCGAGTGCAAAG GCGAACGCGTCCCGAATCCAACGCGCATTGAAGCACTGTCGGCGTTTATTTTAAA GAGCGCAATGTTGGCCTCCCGCTCGACAGCCTCTTCTCGTTTCGTACTGCACCAG GCCGTCAATCTTCGCCGTCTTACTGAACCTCGTCTTTCCCCGTGTTCTGTTGGGA ATTTATTCTTGTGGGCAACAGCGGCATACAACATGGAACATGCGGCCGAGATGGA ACTGCATGGCCTTGTAGCACGTATGAAGCAGGCAGTAGGGAAAATTAATAGTGAA TACTTGAAAACCCTGCATGGCGATGAAGGCTTCCCCAAGGTTTGCGAGTATATTA AGCGCATCGAAGAAGTGTCCGCACACAAGAACTTGGAAGCATTTACATTTTCCTC CTGGGTGAAGTTCGGGTTTAATGAAGTGGACTTTGGATGGGGCAATCCGATTTGG AGTGGTATTTTTGGAGAAGTCGGAAGTAATTCCTTTCGTAACTTAACATTTTTTA

	AGGAGACACGTTCTGCGAATTACGATAATGCGGTCGAGGCATGGGTGACGTTGGA CGAGAAGATCATGTCGGTCTTGGAACACGATCCGCAGTTTCTTGCCTTTGCCTCC CCCAATCCCAGTATCCTTTTGTTTGACGTTTGCGGGTTCTAA
LaAAT1	ATGAAAATTGAAATCAAAGAATCAACCATGGTGCGTCCCGCCGCCGAAACCCCTA GCGGATCACTGTGGCTTAGTAATCTTGATCTGTTGAGTCCTGCTAATTATCACAC TTTGAGCGTACATTTCTACTCTCATGACGGATCTGCCAATTTCTTTGACGCGACT GCGCTTAAAGAGGCGTTATCCCGCGCACTGGTAGACTTCTACCCCTACGCCGGCC GCTTAAAGTTAAACAAGGAGAACCGCCTTGAAATCGAATGTAATGGCGAAGGAAT TCTTCTGGTCGAAGCAGAGTGTAGTGGAGCACTTGACGAATTGGGTGACTTCACC CCACGCCCAGAGCTGAACCTTATTCCCAAGGTGGATTACAGTAAAGGTATGAGCA CATATCCCTTGATGCTTTTTCAGATTACGCGCTTTAAGTGCGGTGGCGTCGCGCT TGGCGTGGCCAACGAACACCACTTAAGTGACGGTGTGGCAGCGCTGCACTTTATT AACACTTGGGCACACTATAGTCGCGGGGTGCCGGCTCCCTCCCCTCCTCCTCATT TTGATCGTACTGCCTTGTCTGCCCGTAACCCTCCCCAACCTCAGTTCTCACACGC CGAGTATCAGCCACCGCCGACTCTGGAGAACCCTCTTCCTGCAACCGATATCGCA CACTCGAAATTCAAATTAACCCGTGCTCAGCTTAATTCCTTAAAGGCGAAGTGTG CGGCAGGAGACTCTGATGGGCATACGAACGGGACCGCAAATGGTAAATCCGATGC GAATGGCACAGCCGATGGCAAATCAGATGCGAACGGCACCGCGAACGGCAAATCC GCGGCCAAGCGTTACAGCACATTTGAGGTCCTTGCCGGACACATTTGGCGTTCAG TGTGCACTGCCCGCGGGCTTCCGGCGGAACAAGAAACCAAGCTGCATATCCCCTT TGACGGTCGTAGTCGCTTAAATTTACCACCAGGATACTTTGGCAACGCGATCTTT TTCGCTACACCGATCGCTACGTGCGGAGAGATTGAATCGAATTCATTGTCCTATG CGGTCCGCCGCGTAGGTGATGGAATCGCACGTCTGGATGAGGAATACCTGAAAAG TTCCCTTGATTTTCTTGAGCTTCAGCCTGATATCTCGAAACTGGCTCAAGGCGCT CACTCCTTTCGTTGTCCGAATCTGTGGGTCATTAGCTGGGTTTGGCTGCCGATTT ACGAGCCTGATTTTGGCTGGGGTAAGGCCGTTCATATGGGACCGTGGGCAGCGCC GTTCGAGGGCAAGTCTTATTTACTTCCTAATCCTGAGAACGACGGATCCTTGTTC GTTTCCATTACCTTGCATAAACAACATATGGAGCGTTTCCAAAAACTTTTTTACG AGATCTAA
LaAAT2	ATGGGAGAGGTAGCGAATGATGAGAAGAAAGTAGGAAGCGTGAAAACTTTTAATC CCACATATGTGAAACCGAAGAAACCCATTGGACGTAAAGAGTGTCAATTGGTCAC CTTTGATTTACCCTATTTAGCGTTTTACTATAACCAAAAGTTGCTTATCTACAAC GGTGGAGATGATTTCTACGGTGCTGTCGAGAAATTGAAAGACGGATTAGCGGTCG TTCTTGAGGAGTTCCATCAGCTGGCTGGCAAGTTAGAGAAGGATGAGGATGGGGT CTTCAAAGTCGTCTATGACGATGACATGGAGGGTGTTGAGGTTGTGGAAGCGGCC GCAGAGGGTGTGGAGGTAGCGGATATTACGGCGGAGGAGGGTTTCAGTAAATTTA AGGAACTTCTGCCATATAACGGCGTTCTTAACCTTGAAGGACTGCAGCGTCCGTT ACTTGCAGTACAACTTACTAAGTTGAAGGATGGCATGGTAATGGGGTGTGCTTTC AATCATGCGATCCTTGATGGTACAAGCACATGGCATTTTATGTCATCGTGGGCTG CTATTTGTTCAGGAGCAACGAGTGTTTCCGTTCCACCATTTCTGGAGCGCACCAA AGCTCGTAACACGCGCGTCAAGCTGGATCTGTCGCAGCCAAGCGACGCACCAGAG CATGCCAATACTGCGTCCAACGGTGATACGCCTGTTAATCCCCTGCTGCGCGGCA AAGTGTTCAAGTTCTCAGAATCAGTGATCGACCAGATCAAAAGTAAGGTAAATGC AGGTTCTGACAGCTCTAAACCATTCTCTACTTTTCAGAGTCTTTCCGCCCATGTG TGGCAAGCCGTGACTCGTGCACGTGAGTTGGGGCCAACTGACTATACTGTATTTA CAGTATTCGCAGATTGCCGTAAGCGTGTTGATCCACCCATGCCGGAGTCGTATTT TGGTAACCTTATTCAGGCGATTTTCACCGTAACGGGAGCTGGCTTGATCTTAGCG AATCCAGTAGAGTTCGGGGCCGGCTTAATCCGTGGAGCTATCGAATCACATAACG CTGAGGCTATTAATAAGCGCAACGAAGAGTGGGAATCGAAGCCAGTAATTTTTCA GTACAAAGATGCTGGGGTAAACTGCGTCGCGGTCGGTAGTTCTCCGCGTTTCCAG GTCTACGGTGTGGACTTTGGATGGGGCAGTCCTGAGTCCGTTCGTAGCGGGCTTA ATAATCGTTTCGACGGTATGGTATACTTATACCCAGGGAAGTCTGGAGGGCGTAG CATCGACGTAGAACTGTCTTTGGAGGCTAAATGCATGGAGAAGTTAGAGAAAGAC AAAGAGTTCTTAATGGAGGCCTAA
LaAT1	ATGAAAATTGAGATTAAGGAAAGCACCATGGTGCGCCCAGCGGCAGAAACACCGT CGGGGTCGTTGTGGTTGTCCAATCTTGACCTTCTTTCGCCGGCAAACTATCACAC ACTGTCAGTACATTTTTATAGTCACGATGGGAGCGCGAATTTCTTTGGTGCTGCC GCATTGAAAGAGGCTCTTTCACGCGCCTTGGTGGACTTTTACCCATATGCTGGAC GCTTGAAATTGAACAAGGAAAATCGTTTGGAGATTGAATGTAATGGGGAGGGTAT TCTGTTGGTGGAGGCTGAGTGCGGAGGAGCATTAGATGAGCTGGGTGACTTTACG CCGCGCCCCCAGCTGAACTTGATTCCTAAAGTTGATTATTCGAAAGGGATGAGCG

	CCTACCCGCTTATGCTTTTCCAAATTACGCGCTTTAAGTGTGGCGGTGTCGCCTT GGGAGTTGCTAACGAACATCACTTATCAGATGGCGTGGCAGCCTTACACTTCATT AACACCTGGGCACATTACTCACGCGGTGTACCAGCGCCGAGTCCCCCGCCCCACT TTGACCGTACGGCCTTGTCAGCTCGCAACCCAAGTAAACCGCAGTTTTCGCATGC TGAATATCAGCCCCCGCCGACAATGGAAAACCCTCTGCCTTATACAGATATTGCC CACTCTAAATTTAAACTGACGCGTGCACAATTAAATTCTTTAAAGGCCAAGTGCG CGGCAGCAGATTCAAATGCTCACACAAATTCGATTGCCAACGGCAAGTCGGACGC AAATGGGACCGCAAATGGTAAATCAGACGCCAACCGCACCGCAAATTCCGCAGCA GCTGCGAAACGCTACTCGACGTTTGAAGTTCTGGCCGGGCATATTTGGCGCTGCG TGTGCACAGCTCGCGGTCTGCCGGCTGAACAGGAGACTAAATTGCATATTCCTTT TGACGGCCGTTCGCGCCTTAACCTGCCTCCAGGGTATTTTGGAAACGCTATTTTC TTTGCGACCCCCATCGCCACATGCGGGGAGATCGAGAGTAATTCCTTGTCATATG CCGTACGTCGTGTGGGCGACGGAATCGCGCGTCTGGATGAAGAATATTTAAAAAG CAGTTTAGATTTTTTAGAATTACAGCCAGATATTTCAAAGCTGGCTCAGGGTGCG CATAGCTTCCGCTGTCCTAATCTGTGGGTTATTTCATGGGTTTGGTTACCCATCT ATGAGCCTGACTTCGGGTGGGGAAAGGCTGTGCATATGGGCCCCTGGGCGGCGCC GTTCGAGGGGAAGTCGTATCTGCTTCCGAATCCCGAGAACGATGGCTCACTTTTC GTCTCAATTACTTTACACAAGCAACACATGGAACGCTTCCAGAAGCTGTTTTATG AAATCTAA
LiAAT4	ATGGCTATGATTATTACCAAGCAAATCTTACGCCCGAGTTCACCTACTCCGCAGG CGTTTAAAAACCACAAGTTGAGCTATCTGGATCAGATTCAGGCGCCAATCTATAT CCCTCTGTTGTTCTTCTATAAAAACGAGGAATCGAAGTACCCGGACCAAATTTCC CAGCGTTTTAAGCAGTCTTTGTCAGAAATTCTGACCATTTTCTATCCCTTGGCGG GTACGATGCGTCACAACAGCTTCGTGGATTGTAACGACCGCGGGGTCGAGTTTGT TGAGGTCCGTGTACATGCTCGTTTAGCCCAGTTCATCCAAGACCCTAAAATGGAA GAACTGAAACAGTTGATCCCTGTCGATTGCATCTCACATACAGACGACGACTTTT TATTACTGGTCAAGATTTCTTACTTTGACTGCGGAGAAGTGGTAGTGGGCGTGTG CATGAGCCATAAGATCGGGGACGGAATTTCTTTGGCAGCGTTCATGAATGCATGG GCTGCCACATGTCGTGGTGAAAGTAGTTCCGAGATTATTCACCCCTCGTTTGATT TAGCATTACACTTTCCTCCAAAGGACCATCTTTCTTCGGCGTCATCTTTTCGTGT TGCTATTGCTCAGGAAAACATCATGACTAAACGTCTTGTATTCGACCGTGAGAAG CTGGAAAAGCTTCGTAAGCGCATCGCTGCATCAAGCGACGGTGTGCGCGATCCCT CACGCGTTGAGGCCGTGAGCGTCTTCATTTGGAAGAGTTTAATTGAGGCCCATAA GGCTGAAAGCCATATGACGGAAACTCCTGCAGTCTCTATCGCATCCCATGCAGTC AATCTGCGCCCCCGCACAGTGCCACAGATGGATCAAACATTTGGGAACTGCTACG CCCCTGCGAGCGCGGTCGTCAGCTGGGATGAAGATTACGTGCACCATTCTCGCTT ACGCGCGGCGCTTCGTGAAATCGACGACGACTACATCAACAAAGTGCTTAAAGCC GATAATAATTATCTGACACAAGATCAAATTGGGGACTTGTTTAAGCCCGAAAATA GCGTTCTGAGTTCGTGGTGGCGCTTCCCAGTGTACAAAGTTGACTTCGGTTGGGG AAAACCCGTTTGGGTAAGTACTACAACAATTCAGTACATGAACTTGATCATCTTT ACAAGTACACCCTCTGAAGACGGGATTGAAGCCTGGGTAACAACTACGCACAATT TCTTCCAGGTGTTGCAGGCGAACTATAATAAGCTGGACACTTAA
NtBBT	ATGTTAATCCCACCCGCTAAGCCAACTCCGCATGAGTTTAAGTTCTTGAGTGACA TTGATGATCAACAGAGTTTGCGTTTTCAGATCCCCCTTATTCAGTTCTACCGCCG CAATCCTTCCGTGGAAGGCAAAGACCAGGTGAAGGTTATTCGCGACGCAATCGCT AAAGCCCTGGTATTTTATTATCCCTTCGCCGGCCGCTTGCGCGAACGCGCTACCT CGAAACTTGTCGTCGAATGTACTGGGGAGGGAGTCGTGTTCATTGAGGCCGATGC TGACGTGGCATTGGAACAGTTTGGAGACGAGTTATATCCACCGTTCCCTTGTCTT GGTGAGCTGTTATACGACGTCCCGGGCACCAGTGGTATCATCAACTGCCCTCTGC TTTTGATTCAGGTCACGCGTCTTCGCTGTGGCGGTTTCGTGTTCGCTATCCGCTT CAACCATGCTATGTCGGATGCTGCGGGCATGTTCCAATTTATGTCAGCAGTAGCA GAACTGGCACGCGGGGCGGAGGCGCCAACTGTTCTTCCAGTCTGGCGTCGCGACC TTTTGTCGGCGCGTGATCCACCCCGTGTAACATGTACCCATCACGAATTCGGTGT GGTCCCTGTAACCACCTTCACCCCACCTGCCAACATGGTTGAGCGCGGGTTTTTT TTCAGTCCGGCAGATATCGCGGCCCTGCGTAGCACCTTACCACCACACTTACGTC GCCGCACCTCTGCCTTCCAGATCGCGGTCGCCTGTGCTTGGCGCTGTCGCGTAAT CGCTCTGAGCCCGGACCCTAGTGAAGAAATCCGTATCTCCTGCATGGTAAATTGC CTGAACCGCTTTGACCCTCCTCTTCCGGAGGGATACTACGGAAATGCAGTCGTCT ACCCGGCCGCCGTAGCAGCGGCAGGACGTTTGTGTGCGAGTCCTTTGGAGTACGC AGTTGAGTTGGTCCGTTCAGCTAAGTCGCAGGCTACAGAAGAGTATGTGAAGAGC GTAGCGGACTTAATGGTAATGCGCGGTCGTCCGTTGTTCAAGGCGGCGGGCACCT

	TTTTGGCCTCCGACGTAACTCGCGCAGGCTTTGAACAAGTTGATTTCGGCTGGGG GGAATAA
NtSHCT	ATGTTAACCTCAGCATTGGCTCGCACGTTATCTGAGTTTTATCCAGTTGCAGGGC GTTTGAAGAAAGATGGGAACGGACGTGTGGAAATCGATTGTTCGGGGGAGGGAGC CGTATTTGTCGAAGCTGAGGCGGACGGGGAGATTGACGACTTGGGCGATTTTTCC CCTAACCCTAATATCTGTTTGGCTCCCAAAGTAGATTATTCACAAGGCATCTCCA GCTTTCCACTGCTGCTGGTCCAGGTCACGCGCTTCAAATGTGGGGGGGTTTGTCT GGGGGTGGCAATGGATCATCAAGTAAAAGATGGAATCTCTGCGCTGCATATCATT CATACATGGTGCGACATTGCCCGCGGGCTGGATATCGCAGTACCGCCGTACATGG ATCGTCGTGTATTAGCGGCCCGCGAGCCTCCACAACCCAAGTTTGAGCATGTAGA GTTCCATCCTCCGCCCCCGTTGAAGAATAGTGAGGCTCACACGAATGGGTCCGAG ACGAAATTTGCTGTCTTGAAGTTGACACGCGAACAGCTTATCATTTTAAAGGGCA ATTCTCAAGAGGACGATGGTAAACGCAGTCCCTATTCATCCTTTGAGGCATTGAC TGGACACGTATGGCGCTGCATTTGTAAAGCACGTCGTTTACCAGAAGACCAGGAG ACTAAGTTGACTATTATCATTGACGGTCGTTCACGCCTTCGCCCACCACTGCCTC CGGGTTACTTTGGCAATGCTGTATTTAAGGCAACACATATTGCTCTGTCCCGCGA AGTGGAATCGAATCCTCTTAAGTATGCGGTCTGTAAAGTACATGAGGCACTTACC CGCATGGACGACGAGTATTTACGCAGTGCTATTGATTACCTGGAGGTCGAGGGGG GGTTGGGTCCAAATGCACGTGGCACCGGGTTGTACAAATCACCGAATTTGGGAAT TACTTCCTGGGCCCGCCTGCCCTTCTACGAGGCTGATTTTGGTTGGGGACGCCCC TTCCACGTGGGGTTAGGCGCCATCCCCGCAGAGGGCCACTTAATCGTTATGCCGA GCCCGACCAACGACGGAGGGCTGTCAATGGCCATTGCGTTGCCCGAAGAACAAAT GAAGATGTTCGAGAAGTAA
RsVS	ATGGCTCAGATCCTTGCCAGCCATTTGATCAAGCCAAGTTCACCCACTCCCAACA CCTTTAAGAAACACAAACTTTCAGTCTTAGACCAAATTTCACCCCCGGCTTATTT AACACTTATTTTCTTTTACCAAGACCTGGAGAGTAACCAGCATGAGGAAATCAGT CGCCGCCTGAAACAGTCCTTGAGTGAAATTCTGACTATTTTTTATCCTCTTGCGG GAACGGTACATCGTAATTCTTTCGTAGATTGTAACGACCGCGGGGCAGAGTTTGT TGAGGCACGTGTTCACGGAGGTTTGTCAAAATTTGTACAAAATCCTAAAATGGAG GAGTTGGAACAGCTTCTTCCTGCGGACTTCTCTTCTCACACTGATAACCCGATTC TGTCCGTGCGTATCAGCTATTTTGACTGCGGAGGAATTGCTGTTGGGGTTTGTTT CCCACACAAGATTGGAGACACTTCCTCTTTCGCAACTTTCATGAATGCGTGGGCT GCAACATGCCGTGGAGAGGCTAGTCGCATTACACCTCCCTCGTTCGATTTAGCAT TGCGTTTTCCTCCCCGCGAGTCATTAGCTTCTGGTTTTTACTTGGGCATCAGTGG AGAGAAAATCGTGACCCGCCGCCTGGTTTTCGAACGTGAGAAAGTTGAGAAGATT CGCAAAGAGGCGTCGCGTAACCACGAGGTAAAGGACCCCAGTCGCGTCGAGGCAA TTAGTTCACTTTTATGGCGCAGTTTTATCGAAGCCCATAAAAAGGTAGATAAGGA GACTACAAGCTTTCCTACTAGTCATATGGTATCTATGCGTCGCCGTGCCGTGCCC CCTGTCCCCGACCACGGTTTTGGAAATTGTTTCACACTGGCGCCAGCTATGGCTA GCAATGAAGAAGAGGCGGAGGAGGACGGTGTGCTTGTTTCTCGTCTTCGTGCGAC AATCCGCGGAGTGGACGAGGACTATATCAAGGCCATTAGCGATGACGAATTCATC AAAGAGGTCTTAGGCCAAATCGGAGACTTTTTCAAACCTGGTAACTGTATTTTCA CGTCATGGCTGCGCTTCCCCCTTTATGAAGTGGACTTTGGTTGGGGCAAGCCAGC CCGCGTGTGTACGGCGACCATGCCATGCATGAACTTGGTGATTTTAATGCCTACG CCGCCTTCCCGCGATGACGGAGGTGTGGAGGCATGGGTGAATGTAGCAGACGAGC AATAA
SsMMT	ATGTTGGTTTTCTTCGACATGCCCTTTTACAATTACGCGACTGATAGTTTGCTGT TTTTTGAGTTTCCCTACAGTAAGCCTCATTTCTTGAATGCGATTGTCCCACATCT TAAGAAATCTTTGTCTTTAGCGCTGTCGCATTTTTTACCGCTGGCCGGAAATATT ATCCACCCCCTGAATCCTCAGGACATGCCATTATTACGCTATGTGTCCACAGACT CTGTCCCGTTCACAGTTTCGGAGTCCGAGGCAGATTTTAGTCGTTTGACTGGGAA CCAACCACGCGACTGTGATGAATACTACGCATTTGCCCCTTTACTTCCCCGTGCT GCGCCATCAGGGACATCGATCAGCTGTCCTGTACTTGCTTTACAGGTAACGCTGT TTCCGGGCAAGGGACTTTCAATTGGAATTGTTACCCGTCACGCAGTCGCAGATGG ATCTACTACCGTATCCTTTATCCGTGCGTGGGCTATGATCTCCGAGATCGGACCC GATGAAGGTTCAGTAGAGGCCCGTTTGACAGAAGCGGGCCTTCTGCCGTTCTTAG ATCGCAAAACAGTCACCAACATTAATGGGCTTGATTCTGTACACTGGGAAAAAAA AGTGAAGAGTTCTTGCTCGGCGGAACCGCTTCCCATGAAATTTCCTATCAACAAG GTGCGCGCCACGTTCGTCTTACGTCGCGATGAAATTGAAAAACTGAAGAATTTGG TCGTCGCTAAGAGTGACGGCGGTGTGGCTCATGTGTCGAGCTTTACCGTTGTCTG TGCTCTGGTATGGGCCTGTTCCGCTCGTGCAGCCGCGAGTGGAGGGGAAGAGGTC

	GCAGAGGAGGAAGAGGAGTATTTCATGTTCGTTGCTGACTTTCGCGGCCGTTTGA GTCCGGCATTGCCAGCTAACTACTTCGGAAATTGTGTTGCGCCTGCAAAGGCTGT TTTGAAACATGGGCAAGTCAAAGGAGATGACGGCTTCGTCACAGCCGCAGAGGCT ATCGGGGCAGCGGTAAAGGGGCTTATCATGTCATCGGATAAGAAGGGAATTCTGG ATGTGGCGGAAGACTGGCTGGAAGAGTATGGCAATCTTGCCGGAAAGCGTCGTCA CGGTGTCGCAGGCTCTCCCCGTTTTGACTTTTACGAGGCTGACTTTGGTTGGGGG AAGCCTTAA
AtAAE7	ATGGCTGCTACAAAATGGCGAGACATCGATGACCTGCCAAAAATCCCTGCCAACT ACACGGCTCTAACGCCGTTGTGGTTCCTAGACCGTGCTGCTGTCGTTCATCCTAC ACGAAAGTCCGTGATTCACGGCAGTCGCGAATACACTTGGCGTCAAACTTACGAT CGATGCCGCCGATTAGCTTCTGCCTTAGCAGACCGATCCATCGGTCCTGGTTCTA CGGTTGCAATTATTGCGCCAAACATTCCTGCAATGTATGAAGCCCATTTCGGGGT TCCCATGTGCGGCGCTGTTCTGAACTGCGTTAATATTCGCTTAAATGCGCCCACC GTCGCCTTCTTACTAAGTCATTCCCAAAGCAGTGTAATCATGGTCGATCAAGAGT TCTTTACTCTAGCCGAGGACTCCCTCCGCCTGATGGAGGAGAAGGCCGGCAGCAG CTTCAAACGTCCATTACTAATCGTTATCGGGGATCACACCTGCGCTCCAGAGTCT CTCAATCGGGCTCTCAGTAAAGGGGCAATCGAGTACGAGGATTTTCTGGCGACTG GTGACCCCAACTATCCTTGGCAGCCACCAGCTGACGAGTGGCAAAGTATTGCTCT CGGTTACACCAGCGGGACCACTGCTAGCCCGAAAGGTGTCGTTTTGCATCATCGT GGGGCATACATTATGGCATTAAGTAATCCCTTGATCTGGGGGATGCAGGATGGGG CCGTATACCTATGGACCCTGCCAATGTTTCACTGTAATGGTTGGTGCTTCCCATG GAGTCTCGCTGTTTTATCTGGCACTTCTATCTGTCTACGACAAGTTACCGCTAAG GAGGTATACAGCATGATTGCCAAGTATAAGGTCACCCATTTCTGCGCAGCCCCGG TAGTCCTGAATGCAATCGTAAATGCCCCTAAGGAAGATACCATCCTACCCCTCCC TCACACTGTACATGTTATGACGGCAGGGGCTGCTCCGCCCCCCAGTGTTCTCTTT AGCATGAACCAGAAAGGCTTCCGGGTAGCGCATACGTATGGATTAAGTGAAACTT ACGGTCCGAGTACCGTATGTGCATGGAAACCAGAATGGGACTCTCTGCCTCCAGA GACACAAGCAAAACTGAACGCCCGCCAAGGAGTCCGTTACACAGGAATGGAACAA TTAGATGTGATCGATACGCAAACTGGGAAGCCAGTCCCGGCTGACGGAAAAACCG CCGGCGAGATTGTATTCCGAGGAAATATGGTGATGAAGGGGTACTTGAAAAACCC TGAAGCAAATAAAGAAACCTTTGCAGGGGGATGGTTTCACAGTGGGGATATTGCA GTAAAACACCCGGATAATTATATTGAAATTAAGGATCGAAGTAAAGATGTGATTA TCTCTGGTGGCGAGAACATTTCTTCCGTAGAAGTAGAGAATGTCGTGTATCACCA CCCTGCTGTACTGGAGGCATCTGTAGTCGCCCGACCCGATGAACGCTGGCAAGAA AGCCCTTGCGCGTTCGTTACGTTGAAAAGTGACTACGAGAAGCATGATCAGAATA AATTAGCACAAGACATCATGAAGTTTTGCCGAGAGAAGTTGCCTGCGTACTGGGT GCCAAAGAGTGTAGTCTTCGGGCCCTTGCCAAAGACGGCAACTGGCAAAATCCAG AAACACATTTTGCGAACCAAGGCTAAGGAGATGGGCCCGGTGCCCCGCAGTCGCC TGTAA
GaAAE7	ATGGAAATGCGTGACATCGATGACCTCCCCAAGAACGCTGCCAACTATATGGCTC TAACCCCGTTGTGGTTCTTAGAGCGTGCTGCTACTGTGCACCCAACGCGTAAAGC TGTGATCCACGGTAGTCGTACATATACGTGGCTGCAAACCTACCAGCGGTGTCGA CGCTTAGCATCCGCCTTGTCCAAAAGTTCTATCGGAGTCGGGAGTTCCGTGGCAT TGATCGCGCCTAACATGCCGGCATTATATGAAGCTCATTTCGGGATCCCTATGAG CGGAGCTGTCGTGAATACTATCAATATCCGCCTAAACGCCTCTACAATCGCTTTT CTATTAGGTCATTCCCAGTCTACCGTGGTCATTGTGGATCAAGAATTTTTTAATT TAGCGGAGGATGCCCTACGGATCATGAAGGAAAAAAGCCAGGGTAATTTTACACC CCCTTTAGTAATTGTAATTGCCGATGAATCCTGCAATCCAGAGACTCTCAAGTAC GCATTGGGTCAAGGGGCTATTGAGTACGAAAAGTTCCTCGAAAGTGGCGACCCCG ATTTTGCGTGGAAGCCGCCACAAGACGAATGGCAGTCCATTGCGCTGGGTTACAC GAGTGGAACTACAGCCAGCCCCAAAGGGGTCGTCCTACATCACCGGGGCGCTTAT TTAATGGCGCTGTCCAACCCTATTATCTGGGAGATGAATGAGGGAGCGGTTTATC TATGGACTTTGCCGATGTTTCACTGTAATGGCTGGTGCTTCCCTTGGACCCTGGC CGCAGTCTGCGGGACAAACATTTGCCTGCGCCAGGTGACAGCTAAGGGCGTGTAC TCCGCTATTGCTAAATATCGGGTGACACATTTCTGCGCGGCGCCTGTTGTCCTGA ATAGCATCGTGAACGCGCCACCTGAAGAAACTATTTTGCCATTGCCACACGTGGT ACGCGTTATGACGGCAGGCGCCGCACCACCGCCTAGCGTTCTCGCTGGCATGTCC CAAAAAGGGTTCCGAGTAACACACACTTATGGCCTATCCGAGACTTACGGTCCCA GTACAGTATGCGCATGGAAACCAGAATGGGACAGCCTCCCTCCTGAAACACAGGC TCGTCTGAACGCACGGCAGGGCGTACGTTACGTCGGCCTAGAATGTCTAGACGTT ATTAGCACACAGACTGGACAACCGGTCCCTGCGGATGGAAAGACGATTGGAGAGA

	TCGTATTCCGGGGAAACATTGTAATGAAAGGATATCTGAAGAACCCGAAAGCCAA CGAGGAAACATTCGCCAATGGGTGGTTTCACAGCGGGGACCTGGGTGTGAAGAAC CCCGACGGCTATATTGAAATCAAAGATCGATCCAAGGACATTATCATCTCTGGGG GAGAAAACATTTCCAGCGTGGAAGTAGAGAACTCTTTGTACCTACATCCTGCTGT CATGGAGGCAAGCGTGGTAGCACGGGCGGACGAACGGTGGGGTGAGAGTCCTTGC GCCTTCGTTACCTTAAAACCCGGCGTTGACAAGAGCGATGAACAACGATTAGCCG AGGACATCATGAAATTTTGTCGGAGCAAGATGCCTGCTTATTGGGTTCCTAAGTC CGTCATTTTCGGGCCGTTACCTAAGACCGCGACCGGAAAAATTCAAAAACACGTG CTCCGCGCCAAGGCTAAGGAACTGGGACCCGTTCGCATGTCTAAACTCTAA
TcAAE7	ATGCGTGACATCGACGATTTGCCCAAAAACGCAGCGAATTACATGGCCTTAACGC CGCTCTGGTTTCTCGAGCGCGCCGCGACAGTACACCCCGCACGAAAGGCAGTAGT CCATGCTAGTCGGACTTACACGTGGCTACAAACTTACCGACGGTGCTGTTGTTTG GCTTCTGCATTGTCTCGTCGGTCCGTAGGGGTAGGGTCTACCGTAGCTATTATTG CGCCAAACGTCCCCGCGCTGTACGAAGCTCATTTCGGTATCCCGATGAGTGGCGC GGTACTGAACACCGTTAACATTCGCTTAAATGCTAGCACTATCGCTTTCCTGCTG GGACACTCCCAGTCTGCTGTAGTGATTGTGGATCAGGAATTCTTCACGTTAGCAG AGGACGCTCTGAAGATTATGAAAGAGAAGAGCCAGGGAAAGTTTAAACATCCTCT GTTGATCGTTATTGCCGACGAATCTTGTGATCCGAAGGCATTACGTTATGCCCTA GGTAAGGGAGCTATCGAGTACGAGAAATTTCTAGAGAGTGGCGATCCCGAATTTG CTTGGAAACCACCGCAAGACGAATGGCAATCTATTGCTCTGGGCTATACCAGCGG GACAACAGCGAGTCCAAAAGGAGTGGTACTCCACCATCGTGGCGCTTACCTCATG AGCTTGAGTAACCCTTTGATCTGGGGAATGACTGAAGGCGCCGTATACCTGTGGA CGCTACCAATGTTCCATTGTAATGGATGGTGTTTTACATGGAGTTTAGCTGCGTT GTGTGGGACGAATATCTGTTTGCGTCAAGTCACGGCGAAGGGGGTGTATTCTGCC ATTGCCAAGTACGGAGTGACCCACTTCTGCGGAGCACCGGTAGTTCTGAATACGA TCGTTAACGCACCACCGAAGGACACTATTTTACCACTCCCCCACGTGGTACATGT GATGACAGCGGGTGCTGCACCCCCGCCAGCGGTATTGAGCGCTATGAGTCAAAAA GGGTTTCGGGTAACACACACATATGGCTTAAGCGAAACATATGGCCCCAGCACAG TTTGTGCTTGGAAGCCGGAGTGGGATTCCCTCCCCCCCGAAACCCAGGCCCGACT AAACTCTCGACAAGGCGTCCGATATATTGGACTAGAAGGTCTAGACGTAATTTCC ACGAAGACTGGACAGCCTGTACCAGCCGATGGGAAAACCATCGGCGAAATCGTAA TCCGTGGTAATGTCGTCATGAAGGGGTACCTGAAAAACCCCAAGGTCAATGAGGA GACATTTGCTAACGGTTGGTTCCACTCCGGGGACCTAGGCGTGAAGCATAGCGAC GGTTATATCGAAATCAAGGACCGGTCCAAAGACATTATCATCTCCGGGGGAGAGA ATATTTCTTCCGTCGAAGTCGAGAACAACTTATACCTACACCCTGCAGTGCTCGA GGCGAGTGTTGTCGCTCGGGTTGATGAGCGGTGGGGAGAGTCCCCCTGCGCTTTC GTGACGCTCAAGCCGGGTGTGGACAAAAGCGACGAGCAGCAGCTGGCGGAAGATA TCATGCGTTTCTGTCGGAGCAAGATGCCCGCATACTGGGTGCCGAAATCTGTAGT TTTCGGTCCCCTGCCAAAGACAGCTACCGGTAAGATTCAAAAGCACGTACTCCGG GCTAAGGCGAAAGAATTGGGACCCGTCAAGCTCTCCAAGCTCTAA
PDC	ATGAGTTATACTGTCGGTACCTATTTAGCGGAGCGGCTTGTCCAGATTGGCCTCA AGCATCACTTCGCAGTCGCGGGCGACTACAACCTCGTCCTTCTTGACAACCTGCT TTTGAACAAAAACATGGAGCAGGTTTATTGCTGTAACGAACTGAACTGCGGTTTC AGTGCAGAAGGTTATGCTCGTGCCAAAGGCGCAGCAGCAGCCGTCGTTACCTACA GCGTCGGTGCGCTTTCCGCATTTGATGCTATCGGTGGCGCCTATGCAGAAAACCT TCCGGTTATCCTGATCTCCGGTGCTCCGAACAACAATGATCACGCTGCTGGTCAC GTGTTGCATCACGCTCTTGGCAAAACCGACTATCACTATCAGTTGGAAATGGCCA AGAACATCACGGCCGCAGCTGAAGCGATTTACACCCCAGAAGAAGCTCCGGCTAA AATCGATCACGTGATTAAAACTGCTCTTCGTGAGAAGAAGCCGGTTTATCTCGAA ATCGCTTGCAACATTGCTTCCATGCCCTGCGCCGCTCCTGGACCGGCAAGCGCAT TGTTCAATGACGAAGCCAGCGACGAAGCTTCTTTGAATGCAGCGGTTGAAGAAAC CCTGAAATTCATCGCCAACCGCGACAAAGTTGCCGTCCTCGTCGGCAGCAAGCTG CGCGCAGCTGGTGCTGAAGAAGCTGCTGTCAAATTTGCTGATGCTCTCGGTGGCG CAGTTGCTACCATGGCTGCTGCAAAAAGCTTCTTCCCAGAAGAAAACCCGCATTA CATCGGTACCTCATGGGGTGAAGTCAGCTATCCGGGCGTTGAAAAGACGATGAAA GAAGCCGATGCGGTTATCGCTCTGGCTCCTGTCTTCAACGACTACTCCACCACTG GTTGGACGGATATTCCTGATCCTAAGAAACTGGTTCTCGCTGAACCGCGTTCTGT CGTCGTTAACGGCGTTCGCTTCCCCAGCGTTCATCTGAAAGACTATCTGACCCGT TTGGCTCAGAAAGTTTCCAAGAAAACCGGTGCTTTGGACTTCTTCAAATCCCTCA ATGCAGGTGAACTGAAGAAAGCCGCTCCGGCTGATCCGAGTGCTCCGTTGGTCAA CGCAGAAATCGCCCGTCAGGTCGAAGCTCTTCTGACCCCGAACACGACGGTTATT

	GCTGAAACCGGTGACTCTTGGTTCAATGCTCAGCGCATGAAGCTCCCGAACGGTG CTCGCGTTGAATATGAAATGCAGTGGGGTCACATCGGTTGGTCCGTTCCTGCCGC CTTCGGTTATGCCGTCGGTGCTCCGGAACGTCGCAACATCCTCATGGTTGGTGAT GGTTCCTTCCAGCTGACGGCTCAGGAAGTCGCTCAGATGGTTCGCCTGAAACTGC CGGTTATCATCTTCTTGATCAATAACTATGGTTACACCATCGAAGTTATGATCCA TGATGGTCCGTACAACAACATCAAGAACTGGGATTATGCCGGTCTGATGGAAGTG TTCAACGGTAACGGTGGTTATGACAGCGGTGCTGGTAAAGGCCTGAAGGCTAAAA CCGGTGGCGAACTGGCAGAAGCTATCAAGGTTGCTCTGGCAAACACCGACGGCCC AACCCTGATCGAATGCTTCATCGGTCGTGAAGACTGCACTGAAGAATTGGTCAAA TGGGGTAAGCGCGTTGCTGCCGCCAACAGCCGTAAGCCTGTTAACAAGCTCCTCT AG
AtALT2	ATGTTCCATGAAATTGAATTGAAGGTACGTGACTACGAACTAGATCAATTTGGCG TTGTGAACAATGCGGTTTATGCAAATTACTGTCAGCATGGTCGCCATGAATTTAT GGACAGCATCGGCATTAACTGTAACGAGGTCAGTCGCTCCGGTGGTGCCCTCGCT ATCCCCGAATTGACTATTAAATTTTTGGCTCCCTTGCGGAGTGGATGTCGTTTTG TCGTGAAGACGCGCATTAGTGGTATCAGCTTGGTCCGAATTTACTTCGAGCAATT CATCTTTAAGTTGCCGAACCAGGAGCCAATCCTAGAGGCAAAGGGAACTGCGGTC TGGCTCGACAACAAGTACCGTCCTACTCGTGTACCTTCCCACGTTCGCTCCTACT TCGGCCACTTCCAATGCCAGCACCTCGTAGATTGA
MtALT1	ATGTTCTGCGACGTCGAACTAAAGGTGCGCGACTACGAACTCGATCAGTTTGGAG TAGTGAACAACAGTGTTTACGCCGGTTATTGCCAACATGGGCGTCACGAGTTTTT AGAATCTATTGGAATTAACTGCGACGCCGTTGCCCGATGTGGGGACGCTCTGGCG TTGAGTGAATTGAGTTTCAAATTTCTAGCCCCCTTACGCTCTGGCGACCGGTTCG TCGTGAAGGTACGCGTCAGTGGAAGTAGCGCGGCTCGCATTTATTTCGACCACTT CATCTACAAACTGCCCAACCAGGAGCCGATCTTGGAGGCAAAGGCAACGGCTGTG TGGTTGGACAAGAACTACCGTCCTATTCGTATCCCTGCTGACATCAAGTCCAAAT TCGTCAAATTTATTCGGAATGAGGATTCCTGA
YneP	ATGCACGTCTCCAAGAAAGAAATCGAGGTTCGGTACGCAGAAACAGACCAGATGG GTATTGTCTATCACGCAAATTATTTAGTTTGGATGGAGGTAGGGCGAACTGCTCT AATTAAAGAGTTAGGGTTCCTATATAAGGACATGGAGGATCGTGGTGTTCTAAGT CCGGTACTCGACATTTCCATCTCCTATAAGAAACCGCTACGATACGGAGAAACAG CAGTGGTACACACATGGATTGAGGAGTACAATGGTTTTAAGACAGTCTACGGTTA TCATATCTACAATCCAGACCAGGAACTCGCGATCAAAGCTACTAGTAGCCACATT TGTGTGGATAAGGAGTCCTTTAAGCCAATCCAGTTTCGGAAAGCTTTCCCGGATT GGCACGCTGCCTACGAAAAAGCAAAAAAGTGA
YtpA	ATGTGGACCTGGAAGGCAGACCGCCCGGTGGGTACTATTGTAATTATTCACGGTG CCTCTGAATATCATGGTCGTTACAAGTGGCTAGTAGAGATGTGGCGAAGCTCCGG GTACAACGTCGTGATGGGGGATTTGCCAGGGCAGGGCACATCCACACGAGCGCGA GGACATATCCGTTCCTTCCAAGAATACATTGACGAGGTGGATATTTGGATTGACA AGGCGCGTACTCTCGAAGCACCCGTGTTCTTGCTAGGACACTCCATGGGAGGGCT AATCGCGATCGAGTGGTTCAAACAACAACGCAATCCGCGGATCACCGCTTTGATT TTATCTAGTCCATGCCTAGGACTGCAAATTAAAGTTAATAAAGTCCTCGATTTTG CATCTAAAGGGCTGAATGTCCTAGCGCCGTCTTTGCGAGTGGATAGCGGGCTAAG CCCTGACATGGCAACGCGTAACGCCGACATGATTGAAGCTGATCAGAATGATTCT CTCTATGTGACCAAGGTGTCCGTCCGATGGTATCGCGAACTACTAAAAACCATTG ATGCAGCGATGGTACCCACAGACGCGTTCTTAAAGGTGCCACTCCTACTAATGCA AGGTGGAGACGACAAAATCGTTGATAAGATGAAAGTGCGGAAGTGGTTTAGTGGT GTCGCTTCCCATAATAAGAGTTATCGAGAATGGGAGGGGTTGTATCACGAGATTT TCAACGAGCCTGAACGGGAAACAGTGTTCAAAGCCGCGAAGGCTTTCACCGATCA ATATATCTGA
DfsA (Wild-Type)	ATGCAACTACAATCCCATCCCGCTTGCCGTCCGTTTTACGAGGCAGGTGAGTTAT CCCAACTCACTGCCTTTTACGAGGAGGGGCGGAATGTCATGTGGATGATGTTGCG GTCTGAGCCACGTCCATGTTTTAATCAGCAATTAGTTACTGACATTATCCATTTG GCGCGGGTCGCTCGCGACAGTGGCCTGACCTTCGATTTTTGGGTCACTGGAAGTC TGGTCCCCGAACTATTCAATGTTGGTGGTGACCTGTCCTTTTTTGTTGACGCCAT TCGTAGCGGACGGCGCGACCAACTGATGGCCTATGCGCGGTCTTGTATTGATGGC GTCTATGAGATCTATACCGGCTTTGGAACAGGTGCAATCTCCATCGCCATGGTCG AGGGGAGCGCGCTGGGTGGGGGCTTTGAGGCGGCGTTGGCTCACCATTATGTACT GGCGCAAAAGGGCGTGAAGCTGGGCTTCCCAGAAATTGCCTTCAATCTGTTCCCT GGAATGGGTGGTTACTCTTTAGTAGCCCGAAAGGCCAATCGAGGGCTGGCGGAAT CTCTGATCGCAACCGGTGAGGCGCATGCGGCGGAATGGTACGAGGATTGTGGTTT

	GATTGACGAGACATTTGACGCGGGCGACGCTTATCTGGCGACCCGGACGTTTATC GATGTCACCAAACCCAAGCTGAATGGCATTCGGGCTATGCTGCGCGCTCGGGAGC GAGTTTTTCAACTGAGTCGCAGTGAGTTAATGGACATCACCGAAGCCTGGGTACA TGCCGCGTTCACTATTGAACCAAAAGATTTGGCATATATGGAGCGTCTAGTAATG CTCCAAAACCGACGAGTGAGTAAATTACGTACAGTATGA
DfsA F44W	ATGCAACTACAATCCCATCCCGCTTGCCGTCCGTTTTACGAGGCAGGTGAGTTAT CCCAACTCACTGCCTTTTACGAGGAGGGGCGGAATGTCATGTGGATGATGTTGCG GTCTGAGCCACGTCCATGTTGGAATCAGCAATTAGTTACTGACATTATCCATTTG GCGCGGGTCGCTCGCGACAGTGGCCTGACCTTCGATTTTTGGGTCACTGGAAGTC TGGTCCCCGAACTATTCAATGTTGGTGGTGACCTGTCCTTTTTTGTTGACGCCAT TCGTAGCGGACGGCGCGACCAACTGATGGCCTATGCGCGGTCTTGTATTGATGGC GTCTATGAGATCTATACCGGCTTTGGAACAGGTGCAATCTCCATCGCCATGGTCG AGGGGAGCGCGCTGGGTGGGGGCTTTGAGGCGGCGTTGGCTCACCATTATGTACT GGCGCAAAAGGGCGTGAAGCTGGGCTTCCCAGAAATTGCCTTCAATCTGTTCCCT GGAATGGGTGGTTACTCTTTAGTAGCCCGAAAGGCCAATCGAGGGCTGGCGGAAT CTCTGATCGCAACCGGTGAGGCGCATGCGGCGGAATGGTACGAGGATTGTGGTTT GATTGACGAGACATTTGACGCGGGCGACGCTTATCTGGCGACCCGGACGTTTATC GATGTCACCAAACCCAAGCTGAATGGCATTCGGGCTATGCTGCGCGCTCGGGAGC GAGTTTTTCAACTGAGTCGCAGTGAGTTAATGGACATCACCGAAGCCTGGGTACA TGCCGCGTTCACTATTGAACCAAAAGATTTGGCATATATGGAGCGTCTAGTAATG CTCCAAAACCGACGAGTGAGTAAATTACGTACAGTATGA
DfsA F44Y	ATGCAACTACAATCCCATCCCGCTTGCCGTCCGTTTTACGAGGCAGGTGAGTTAT CCCAACTCACTGCCTTTTACGAGGAGGGGCGGAATGTCATGTGGATGATGTTGCG GTCTGAGCCACGTCCATGTTATAATCAGCAATTAGTTACTGACATTATCCATTTG GCGCGGGTCGCTCGCGACAGTGGCCTGACCTTCGATTTTTGGGTCACTGGAAGTC TGGTCCCCGAACTATTCAATGTTGGTGGTGACCTGTCCTTTTTTGTTGACGCCAT TCGTAGCGGACGGCGCGACCAACTGATGGCCTATGCGCGGTCTTGTATTGATGGC GTCTATGAGATCTATACCGGCTTTGGAACAGGTGCAATCTCCATCGCCATGGTCG AGGGGAGCGCGCTGGGTGGGGGCTTTGAGGCGGCGTTGGCTCACCATTATGTACT GGCGCAAAAGGGCGTGAAGCTGGGCTTCCCAGAAATTGCCTTCAATCTGTTCCCT GGAATGGGTGGTTACTCTTTAGTAGCCCGAAAGGCCAATCGAGGGCTGGCGGAAT CTCTGATCGCAACCGGTGAGGCGCATGCGGCGGAATGGTACGAGGATTGTGGTTT GATTGACGAGACATTTGACGCGGGCGACGCTTATCTGGCGACCCGGACGTTTATC GATGTCACCAAACCCAAGCTGAATGGCATTCGGGCTATGCTGCGCGCTCGGGAGC GAGTTTTTCAACTGAGTCGCAGTGAGTTAATGGACATCACCGAAGCCTGGGTACA TGCCGCGTTCACTATTGAACCAAAAGATTTGGCATATATGGAGCGTCTAGTAATG CTCCAAAACCGACGAGTGAGTAAATTACGTACAGTATGA
DfsA G135C	ATGCAACTACAATCCCATCCCGCTTGCCGTCCGTTTTACGAGGCAGGTGAGTTAT CCCAACTCACTGCCTTTTACGAGGAGGGGCGGAATGTCATGTGGATGATGTTGCG GTCTGAGCCACGTCCATGTTTTAATCAGCAATTAGTTACTGACATTATCCATTTG GCGCGGGTCGCTCGCGACAGTGGCCTGACCTTCGATTTTTGGGTCACTGGAAGTC TGGTCCCCGAACTATTCAATGTTGGTGGTGACCTGTCCTTTTTTGTTGACGCCAT TCGTAGCGGACGGCGCGACCAACTGATGGCCTATGCGCGGTCTTGTATTGATGGC GTCTATGAGATCTATACCGGCTTTGGAACAGGTGCAATCTCCATCGCCATGGTCG AGGGGAGCGCGCTGGGTTGTGGCTTTGAGGCGGCGTTGGCTCACCATTATGTACT GGCGCAAAAGGGCGTGAAGCTGGGCTTCCCAGAAATTGCCTTCAATCTGTTCCCT GGAATGGGTGGTTACTCTTTAGTAGCCCGAAAGGCCAATCGAGGGCTGGCGGAAT CTCTGATCGCAACCGGTGAGGCGCATGCGGCGGAATGGTACGAGGATTGTGGTTT GATTGACGAGACATTTGACGCGGGCGACGCTTATCTGGCGACCCGGACGTTTATC GATGTCACCAAACCCAAGCTGAATGGCATTCGGGCTATGCTGCGCGCTCGGGAGC GAGTTTTTCAACTGAGTCGCAGTGAGTTAATGGACATCACCGAAGCCTGGGTACA TGCCGCGTTCACTATTGAACCAAAAGATTTGGCATATATGGAGCGTCTAGTAATG CTCCAAAACCGACGAGTGAGTAAATTACGTACAGTATGA