

*Department of Chemical and
Environmental Engineering*

Andrew Yiakoumetti

The production of methacrylic acid in *Escherichia coli*

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

July 2016

Abstract

The global demand for methyl methacrylate, a derivative of methacrylic acid, exceeds 2 million tonnes per annum. Current production methods rely entirely upon crude oil and natural gas as feedstocks, use harsh conditions and generate hazardous waste, which has led to an increased demand for bio-derived Perspex. However, the lack of known metabolic pathways in which methyl methacrylate or methacrylic acid are intermediates has prevented Bio-Perspex from becoming a reality. Recently, the hydrolysis of methacrylyl-CoA, an intermediate of the valine degradation I pathway, was proposed as a candidate step for the biocatalytic production of methacrylic acid, but remained unsubstantiated. This project describes the whole-cell biocatalytic production of methacrylic acid *via* the hydrolysis of methacrylyl-CoA in both biotransformations of isobutyric acid and fermentations from glucose. To hydrolyse methacrylyl-CoA, thioesterases were screened for promiscuous activity on methacrylyl-CoA, and 4-hydroxybenzoyl-CoA thioesterase (4HBT) from *Arthrobacter sp.* strain SU was found to selectively hydrolyse methacrylyl-CoA over isobutyryl-CoA (the precursor to methacrylyl-CoA). As this pathway is not naturally found in *E. coli*, the preferred host, the oxidation of isobutyryl-CoA to methacrylyl-CoA was engineered into the organism. This step is usually performed by acyl-CoA dehydrogenases, but as these require a complex mechanism for electron transfer to the organism's electron transport chain, an acyl-CoA oxidase, ACX4 from *Arabidopsis thaliana*, was used instead. The oxidation of isobutyryl-CoA was confirmed *in vitro*, as was the production of methacrylic acid from isobutyryl-CoA in an enzyme coupled reaction between 4HBT and ACX4. For the biotransformation of isobutyric acid to methacrylic acid, 4HBT, ACX4 as well as an acyl-CoA synthetase, AcsA from *Pseudomonas chlororaphis* B23 were co-expressed in *E. coli*. Cultures supplemented with isobutyric acid produced 367 ± 28 μ M methacrylic acid in minimal medium. For the production of methacrylic acid from glucose, branched chain keto acid dehydrogenases from *Pseudomonas putida* KT2440 and *Bacillus subtilis* 168 were co-expressed with 4HBT and ACX4. The strain with the branched chain keto acid dehydrogenase from *P. putida* produced 171 ± 7.6 μ M methacrylic acid, also in a minimal medium. Overexpressing enzymes for the upregulation of 2-ketoisovalerate actually reduced MAA production, but also led to the production of 13mM isobutyric acid in a complex medium, suggesting a bottleneck in the pathway and the background hydrolysis of isobutyryl-CoA. The optimisation of gene expression and the disruption of enzymes that non-productively consume pathway intermediates is proposed, as is the testing of alternative routes to isobutyryl-CoA and the use of a host organism which naturally hosts the valine degradation I pathway.

Acknowledgments

I would like to first and foremost thank my two academic supervisors, Professor Gill Stephens and Professor Peter Licence for their continued guidance and support during this project, encouragement and help. I would also like to thank my industrial sponsors and supervisors at Lucite International UK, Dr Graham Eastham and Dr David Johnson, as well as Wataru Mizunashi San from Mitsubishi Rayon Company, for their invaluable advice and recommendations, which have helped to accelerate and expand this project. I must also express my gratitude to my laboratory supervisor, Dr Stephen Hall, without whose technical support and assistance, I would not have been able to start, let alone complete, this project. I also thank Samuel Wilson, for all his help as the laboratory technician.

Next, I would like to thank all my labmates for their help whenever it was needed, with special mention to the now Dr Pawel Mordaka, as well as Luca Rossoni and Joseph Webb. Pawel initially trained me in molecular biology techniques, and never grew impatient with my incessant requests for his knowledge, advice and experience! I started and finished my Ph.D. with Luca and Joe, and I have always been able to rely on their opinion and recommendations.

I have also been very fortunate to meet my very good friends Anna Andrejeva, Hugo Ferreira and Emily Kirk at the start of my PhD, with whom I have shared many fond memories.

Finally, I would like to thank my family for their unconditional patience and support, and to whom I dedicate this thesis. I especially dedicate this thesis to my grandmother, who sadly passed away during this project.

Table of Contents

1	Introduction	1
2	Literature Review	3
2.1	Current industrial routes to methacrylic acid and methyl methacrylate...	4
2.2	Biotechnology and biocatalysis for the sustainable production of useful chemicals	7
2.3	Possible direct biocatalytic routes to methacrylic acid	16
3	Aims and objectives	24
4	Materials and Methods.....	26
4.1	Materials.....	26
4.1.1	Electrophoresis reagents and buffers	26
4.1.2	Antibiotics, cofactors, vitamins nucleotides and substrates	26
4.1.3	Molecular biology enzymes.....	26
4.1.4	Molecular biology kits, protein extraction kits	26
4.1.5	Salts, acids, bases, buffers, trace elements and gases.	26
4.1.6	Solvents	27
4.1.7	Growth media and cryostock components	27
4.1.8	Tables of primers, bacterial strains, synthetic genes and synthetic operons	28
4.2	Methods.....	30
4.2.1	Preparation of frequently used buffers	30
4.3	Preparation of antibiotic, IPTG, glucose, cofactor and vitamin stocks	30
4.3.1	Glucose stock solutions.....	30
4.3.2	Carbenicillin stock (100 mg.ml ⁻¹)	30
4.3.3	Chloramphenicol stock (34 mg.ml ⁻¹).....	30
4.3.4	Isopropyl-β-D-thiogalactopyranoside (IPTG) stock (0.5 M)	31
4.3.5	Flavin adenine dinucleotide (FAD) stock (1 M)	31
4.3.6	Growth media.....	31
4.3.7	MSX	31
4.3.8	SM medium.....	32
4.3.9	Modified M9 medium	33
4.3.10	DSMZ medium 1	33
4.3.11	DSMZ medium 220	33
4.3.12	Luria Bertani agar plates.....	33
4.4	Cell maintenance, propagation and preservation.	34
4.4.1	Micro-organisms	34
4.4.2	Preparation of cryostocks.....	34
4.4.3	Cell maintenance and culture inoculations.....	35
4.5	Bacterial transformations	35
4.5.1	Transformation of <i>E. coli</i> JM107 cells.....	35
4.5.2	Transformation of <i>E. coli</i> BL21 (DE3) and <i>E. coli</i> BL21 (DE3) pLysS cells	36

4.5.3	Preparation of competent <i>E. coli</i> BL21 (DE3) pHAPP and <i>E. coli</i> BL21 (DE3) pHABS and transformation with pKIV.....	36
4.6	Analysis of protein extracts and purified proteins by Sodium Dodecyl-Sulphate Polyacrylamide Gel Electrophoresis.....	37
4.7	Molecular biology methods	37
4.7.1	Restriction digests of plasmids.....	37
4.7.2	Preparation of linear expression vector	38
4.7.3	Agarose gel electrophoresis and gel extraction	38
4.7.4	Ligation of inserts into linearised plasmids	39
4.7.5	Subcloning a polynucleotide into an expression vector	39
4.7.6	Analytical agarose gel electrophoresis	40
4.7.7	Preparation of genomic DNA extracts from <i>B. subtilis</i> 168 and <i>P. putida</i> KT2440. 40	
4.7.8	Polymerase chain reactions	41
4.8	Synthetic and analytical chemistry methods	43
4.8.1	Synthesis and purification of methacrylyl-CoA	43
4.8.2	Analysis of samples for methacrylic acid production by whole cell biotransformations and whole cell cultures on a C18 column.	44
4.8.3	Isocratic elution on ion exchange column.....	45
4.9	Enzyme expression, extraction and purification.....	45
4.9.1	Preparation of seed cultures	45
4.9.2	Culture inoculation from seed culture	45
4.9.3	Protein expression at 100 ml scale, biotransformations and shake flask fermentations.....	46
4.9.4	Protein expression at 500 ml scale and cell harvesting.....	47
4.9.5	Preparation of cell free protein extracts	47
4.9.6	Preparation of pure Carboxy terminal His-tagged 4HBT.....	48
4.10	Enzymatic assays.....	49
4.10.1	Thioesterase assays	49
4.10.2	Kinetic characterisation of pure His-Tagged 4HBT.	50
4.10.3	Crude acyl-CoA oxidase assay and enzyme coupled reaction.	50
4.11	Table of plasmids.....	52
4.12	Table of strains and supplements required in growth media.....	56
5	Results	58
5.1	The selective hydrolysis of methacrylyl-CoA by a suitable thioesterase .58	
5.1.1	The role of thioesterases in metabolism	58
5.1.2	Screening thioesterases for activity on methacrylyl-CoA	61
5.1.3	Database and literature search for thioesterases with known activity on structurally related substrates to methacrylyl-CoA.....	65
5.1.4	Sourcing the genes, gene cloning and gene expression.....	73
5.1.5	Preparation of methacrylyl-CoA for crude enzyme assays.....	75
5.1.6	Crude enzyme assay	77

5.1.7	Kinetic characterisation of 4-hydroxybenzoyl-CoA thioesterase	80
5.2	Oxidation of isobutyryl-CoA to methacrylyl-CoA	85
5.2.1	Enzymes catalysing the oxidation of acyl-CoA intermediates	85
5.2.2	Identification of candidate acyl-CoA dehydrogenases for the oxidation of isobutyryl-CoA with compatible ETFs and ETFQOs	94
5.2.3	Identification of an acyl-CoA oxidase for the oxidation of isobutyryl-CoA in <i>E. coli</i>	97
5.2.4	Expression of ACX4 and preparation of cell free extracts	98
5.2.5	ACX4 enzyme assays.....	99
5.3	The biotransformation of isobutyric acid to methacrylic acid.	103
5.3.1	Identification of an enzyme or enzyme system for the activation of isobutyric acid to isobutyryl-CoA.....	104
5.3.2	A whole cell biotransformation of isobutyric acid to methacrylic acid	108
5.4	The production of methacrylic acid from glucose	118
5.4.1	Selecting a branched chain keto acid dehydrogenase.....	121
5.4.2	Identifying the nucleotide sequences for the BCKD operons in <i>P. putida</i> KT2440 and <i>B. subtilis</i> 168.....	124
5.4.3	Construction of plasmids and strains for the co-expression of 4HBT, ACX4 and the BCKD genes from <i>P. putida</i> KT2440 and <i>B. subtilis</i> 168.	126
5.4.4	The production of methacrylic acid from glucose with an upregulated 2-ketoisovalerate biosynthesis pathway.....	134
5.4.5	Characterising the production of methacrylic acid over time, and comparison of the effect of riboflavin on host productivity.....	142
6	Discussion	144
6.1	Recent advances in the literature relevant to the biocatalytic production of methacrylic acid.	145
6.1.1	Metabolic pathway to acrylic acid.	145
6.1.2	Production of 3-hydroxyisobutyric acid, a late intermediate of the valine degradation I pathway	146
6.2	Further work	148
6.2.1	Improvements to experimental design.....	149
6.2.2	Improvements of existing metabolic routes for methacrylic acid production.	150
6.2.3	Exploration of alternative metabolic routes and host organisms ...	155
6.3	Conclusions.....	159
7	References.....	160
8	Appendices	I
8.1	Appendix 1 – polypeptide sequences	I
8.2	Appendix 2 – polynucleotide sequences	III
8.3	Appendix 3 – Additional figures.....	XV
8.4	Appendix 4: Derivations of equations	XXIII
8.4.1	Derivation of Equation 1	XXIII
8.4.2	Derivation of Equation 2	XXIV

List of figures

Figure 1	Methacrylic Acid, Methyl Methacrylate, Ethyl Methacrylate and <i>n</i> -Butyl Methacrylate.	3
Figure 2	Chemical steps to the synthesis of Methacrylic Acid and Methyl Methacrylate by the ACH route	4
Figure 3	A C3 route to methyl methacrylate using propyne, carbon monoxide and methanol using a palladium catalyst with a phosphine ligand	5
Figure 4	A C3 route utilising propylene and carbon monoxide as carbon feeds	5
Figure 5	C4 route to methacrylic acid from isobutylene	5
Figure 6	C2 routes to methacrylic acid from ethylene via propanal or propanoate.	6
Figure 7	A C2 route to methacrylic acid via methyl propionate, the Alpha process	6
Figure 8	Metabolic pathway to citrate from pyruvate	7
Figure 9	An example of a glycolytic pathway, the Embden-Meyerhof-Parnas pathway	8
Figure 10A	Glycerol fermentation pathway to 1,3-propanediol	11
Figure 10B	Pathway for the formation of glycerol from dihydroxyacetone phosphate	11
Figure 11A	The enzymatic production of isopentenyl-diphosphate by the concomitant dehydration and decarboxylation of mevalonate diphosphate by mevalonate diphosphate decarboxylase.	12
Figure 11B	The enzymatic production of isobutene by the concomitant dehydration and decarboxylation of 3-hydroxyisovalerate by mevalonate diphosphate decarboxylase.	12
Figure 12	Two routes used by Zhang et al. for the biocatalytic production of isobutyric acid.	13
Figure 13	The production of methacrylic acid from citric acid in near-critical and super-critical water.	14
Figure 14	The dehydration of 3-hydroxyisobutyric acid to methacrylic acid.	14
Figure 15	Routes to methacrylic acid and its esters, methyl methacrylate and butyl methacrylate, from candidate metabolic precursors.	16
Figure 16A	The valine degradation I pathway	19
Figure 16B	The valine biosynthesis pathway	19
Figure 17	Biocatalytic routes to citraconic, itaconic and mesaconic acids from widely available biological precursors	20
Figure 18	Route to methacrylic acid from pyruvate	22
Figure 19	Hydrolysis of a thioester bond	58
Figure 20	4-chlorobenzoate dehalogenation pathway	60
Figure 21	Structural features of methacrylyl-Coenzyme A	61

Figure 22	A comparison of the structural features of methacrylyl-CoA and 4-hydroxybenzoyl-CoA	62
Figure 23	A snapshot of a UniProt database search query to yield enzymes that belong to Enzyme Commission groupings that contain Acyl-CoA thioesterase	65
Figure 24	A snapshot of the results of a Microsoft Access query to show only acyl-CoA hydrolase enzymes that have been proven to exist at protein level	65
Figure 25	A cross section, hybrid mesh-cartoon view of FcoT.	70
Figure 26	Confirmation of the presence of the correct inserts in duplicated cryostocks of thioesterase expression strains	74
Figure 27	Expression of thioesterases	75
Figure 28	HPLC trace of purified methacrylyl-CoA	76
Figure 29	Reaction scheme for Ellman's reagent as indicator for acyl-CoA hydrolysis	77
Figure 30	Crude thioesterase assays using cell free extracts of <i>E. coli</i> BL21 (DE3) pLysS pET20b(+):4HBT, <i>E. coli</i> BL21 (DE3) pLysS pET20b(+):YciA, <i>E. coli</i> BL21 (DE3) pLysS pET20b(+):FcoT and <i>E. coli</i> BL21 (DE3) pLysS pET20b(+) negative control.	78
Figure 31	Expression of 4HBT, NtHS-4HBT and CtHIS-4HBT	81
Figure 32	Lineweaver Burk plots for kinetic characterisation of CtHIS-4HBT on methacrylyl-CoA and isobutyryl-CoA, with high and low estimates of substrate concentrations.	82
Figure 33	Enzymatic mechanism of 4-hydroxybenzoyl-CoA hydrolysis	84
Figure 34	4HBT catalytic mechanism with methacrylyl-CoA as substrate	84
Figure 35	Oxidation of isobutyryl-CoA to methacrylyl-CoA	85
Figure 36	The FAD reductive half reaction during the oxidation of acyl-CoA substrate to <i>trans</i> -2-enoyl-CoA	85
Figure 37	Oxidation of an acyl-CoA by an ACAD and regeneration of the ACAD cofactor	87
Figure 38	Oxidation of saturated acyl-CoA substrate (SH ₂) to an unsaturated <i>trans</i> -2-enoyl-CoA product by 2×ETF	87
Figure 39	Structure of ETFQO from pig liver mitochondria and inter-redox center distances	88
Figure 40	The acyl-CoA oxidase mediated oxidation of an acyl-CoA to its <i>trans</i> -2-enoyl-CoA analog by molecular oxygen and the concomitant production of hydrogen peroxide	92
Figure 41	ACX4 expression in both MSX media at LB media	98
Figure 42	Overlaid signals showing analytical reverse phase HPLC resolution of coenzyme A, methacrylic acid, flavin adenine dinucleotide, methacrylyl-CoA and isobutyryl-CoA	99
Figure 43	HPLC analysis of ACX4 crude extract incubated with isobutyryl-CoA	100

Figure 44	HPLC analysis of an isobutyryl-CoA spiked sample of methacrylyl-CoA that was generated by the incubation of ACX4 with isobutyryl-CoA	100
Figure 45	HPLC analysis of the product from the coincubation of crude ACX4 and pure CtHis-4HBT with isobutyryl-CoA. Major peaks appear for coenzyme A and methacrylic acid	101
Figure 46	No-substrate control HPLC analysis of acidified and clarified assay media containing ACX4 crude extract (0.8mg.ml ⁻¹), 4HBT (0.6mg.ml ⁻¹) and FAD (10μM)	102
Figure 47	Candidate routes for the biotransformation of isobutyric acid to methacrylic acid	103
Figure 48	Construction of pET20b(+):4HBT-ACX4, an expression vector for the co-expression of 4HBT and ACX4, containing a new cloning site downstream of the 4HBT and ACX4 genes	109
Figure 49	The end of the 4HBT gene (blue), the intergenic sequence (black) containing a new ribosome binding site for the ACX4 gene (underlined) and the start of the ACX4 gene (green)	110
Figure 50	The pET20b(+):4HBT-ACX4-AcsA plasmid	111
Figure 51	Coexpression of 4HBT, ACX4 and AcsA in MSX medium with riboflavin at 37°C	112
Figure 52	Production of methacrylic acid by <i>E. coli</i> BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-AcsA during biotransformation of isobutyric acid and growth of culture during the biotransformation (preliminary)	114
Figure 53	HPLC traces of whole cell live culture biotransformation supernatant (8h samples)	115
Figure 54	Growth of <i>E. coli</i> BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-AcsA on glucose and its conversion of isobutyric acid to methacrylic acid with time (triplicate)	116
Figure 55	Two candidate routes to methacrylic acid from 2-ketoisovaleric acid, a widely available intermediate of <i>Escherichia coli</i> .	118
Figure 56	The oxidative decarboxylation of 2-ketoisovalerate by branched chain keto acid dehydrogenases through three concerted enzymatic activities	119
Figure 57	Genes encoded in the operon containing the E3 (<i>lpdV</i>), E1α (<i>bkdAA</i>), E1β (<i>bkdAB</i>) and E2 (<i>BkdB</i>) components of the BCKD from <i>B. subtilis</i> 168	125
Figure 58	PCR amplification attempts for BCKD genes from <i>B. subtilis</i> 168 and <i>P. putida</i> KT2440 genomes.	126
Figure 59	The pET20b(+):4HBT-ACX4-bsBCKD plasmid (pHABS), for coexpression of 4HBT, ACX4 and the genes encoding the E3, E1α, E1β and E2 components of <i>B. subtilis</i> 168 BCKD.	127
Figure 60	The pET20b(+):4HBT-ACX4-ppBCKD plasmid (pHAPP), for the coexpression of 4HBT, ACX4 and the the <i>P. putida</i> KT2440 BCKD E1α, E1β, E2 and E3 components	128

Figure 61	The pET20b(+):pBCKD plasmid, for the coexpression of the genes encoding the E3, E1 α , E1 β and E2 components of <i>P. putida</i> KT2440 BCKD.	128
Figure 62	SDS-PAGE analysis of <i>E. coli</i> BL21 (DE3) pLysS pHAPP	131
Figure 63	Selection of HPLC traces of samples taken from cultures during experiment rapid screening for production of MAA from glucose by <i>E. coli</i> BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-ppBCKD and <i>E. coli</i> BL21 (DE3) pLysS pET20b(+) control strain.	133
Figure 64	Valine biosynthesis from pyruvate in <i>Escherichia coli</i> .	134
Figure 65	Polynucleotide containing <i>ilvD</i> , <i>alsS</i> and <i>ilvC</i> genes with ribosome binding sites preceeding each gene.	137
Figure 66	The pSTV28::ilvD-alsS-ilvC (pKIV) plasmid for the co-expression of <i>E. coli ilvD</i> , <i>E. coli ilvC</i> and <i>B. subtilis alsS</i> .	138
Figure 67	Growth of <i>E. coli</i> BL21 (DE3) pET20b(+):4HBT-ACX4-ppBCKD in MSX media supplemented with carbenicillin only, and its production of methacrylic acid with time. (triplicate)	142
Figure 68	An engineered pathway from L-threonine to acrylic acid. <i>Ec_tdcB</i> : Threonine ammonia lyase from <i>E. coli</i> .	145
Figure 69	Semi-engineered metabolic pathway to 3-hydroxyisobutyric acid in <i>Pseudomonas taiwanensis</i> VLB120 B83 T7.	147
Figure 70	Possible consumption of methacrylyl-CoA by hydration to 3-hydroxyisobutyryl-CoA and hydrolysis to 3HIBA in <i>E. coli</i> .	153
Figure 71	Summary of alternative routes for the production of methacrylic acid from glucose.	155
Figure 72	Alternative routes for the activation of isobutyric acid to isobutyryl-CoA, and overall ATP requirement of each.	156
Figure 73	Alternative gene disruptions for accumulation of methacrylyl-CoA in organisms containing the valine degradation I pathway, for the production of methacrylic acid by the expression of a suitable thioesterase.	158
Figure 74	DC assay calibration curve for thioesterase cell free extract concentrations in HEPES buffer using bovine serum albumin as the protein standard	XVIII
Figure 75	β -mercaptoethanol calibration curve for 500 μ M DTNB in HEPES buffer.	XX
Figure 76	β -mercaptoethanol calibration curve for 500 μ M DTNB in phosphate buffer.	XX
Figure 77	Methacrylic acid calibration curve for production by biotransformation of isobutyric acid.	XXII
Figure 78	Methacrylic acid calibration curve for production directly from glucose.	XXII

List of Equations

Equation 1	Equation to calculate the volume of water to add to a known mass of glycerol to prepare 80% (v/v) glycerol.	35
Equation 2	Equation to calculate the volume of resuspended pellet from seed culture required to add to culture media to reach a desired target starting OD.	46
Equation 3	Equation of the Lineweaver Burk plot	82
Equation 4	Specificity ratio equation for acetohydroxy acid synthase enzymes.	135

List of Tables

Table 1	Table of candidate transformations to yield methacrylic acid or its esters from known biological precursors	17
Table 2	Table of primers used throughout this study	28
Table 3	Table of bacterial strains used, and their genotypes	
Table 4	Table of synthesised polynucleotides	29
Table 5	Table of all plasmids used and constructed in this study	55
Table 6	Table of strains used in this study, and supplements required in growth media	57
Table 7	Grouping the thioesterases by Enzyme Commission number	59
Table 8	A list of putative and characterized, acyl-CoA hydrolases of prokaryotic origin, that have been expressed as proteins, listed on the UniProt database	66
Table 9	A summary table of candidate thioesterases for the hydrolysis of methacrylyl-CoA, identified by substrate similarity, from the UniProt and BRENDA databases	67
Table 10	Summary of enzyme kinetics of 4HBT from <i>Arthrobacter</i> sp. strain SU on various substrates.	68
Table 11	Summary of enzyme characterization results for FcoT on seven different substrates	69
Table 12	Substrates that are hydrolysed by YciA, including kinetic data of the hydrolysis reaction	71
Table 13	Summary of substrates that wild type and S49A mutant Rif ^R hydrolyse, with kinetic data for the hydrolysis reaction	72
Table 14	Table of thioesterase expression strains	74
Table 15	Expression conditions for the preparation of thioesterase cell free extracts	75
Table 16	Initial activity rates for crude thioesterase assays	78
Table 17	Summary kinetic data of 4HBT from <i>Arthrobacter</i> sp. strain SU with methacrylyl-CoA and isobutyryl-CoA as substrates.	83
Table 18	Acyl-CoA dehydrogenases found in mammalian cell mitochondria	86
Table 19	Candidate acyl-CoA dehydrogenases for the oxidation of isobutyryl-CoA	94
Table 20	Summary of ETF α , ETF β and ETFQO entries for organisms which contain acyl-CoA dehydrogenases capable oxidising isobutyryl-CoA to methacrylyl-CoA	94

Table 21	EC groups to consider for the activation of isobutyric acid to isobutyryl-CoA	104
Table 22	Database and literature search results for candidate enzymes for the activation isobutyric acid to isobutyryl-CoA	105
Table 23	Substrates activated to coenzyme A derivatives by AcsA from <i>P. chlororaphis</i>	105
Table 24	Kinetic data for butyryl-CoA synthetase for <i>P. varioti</i> .	106
Table 25	Kinetic data for branched chain fatty acid kinase. No k_{cat} or V_{MAX} data are available	106
Table 26	Forward and reverse primers used by Zhang <i>et al.</i> for the amplication of a polynucleotide containing genes encoding the E1 α , E1 β , E2 and E3 components of branched chain keto acid dehydrogenase from <i>Pseudomonas putida</i> KT2440.	124
Table 27	Summary of results for methacrylic acid production in cultures of <i>E. coli</i> BL21 (DE3) pLysS pHAPP and <i>E. coli</i> BL21 (DE3) pLysS pHABS, when cultured both with and without 2-ketoisovalerate supplementation.	132
Table 28	Summary of modifications made to remove <i>EcoRI</i> and <i>HindIII</i> sites from <i>lvD</i> , <i>alsS</i> and <i>ilvC</i> .	137
Table 29	DNA source and primer combinations used to confirm the presence of pSTV28::ilvD-alsS-ilvC and pET20b(+):4HBT-ACX4-ppBCKD in transformants of <i>E. coli</i> BL21 (DE3) pSTV28::ilvD-alsS-ilvC pET20b(+):4HBT-ACX4-ppBCKD and presence of pSTV28::ilvD-alsS-ilvC and pET20b(+):4HBT-ACX4-bsBCKD in transformants of <i>E. coli</i> BL21 (DE3) pSTV28::ilvD-alsS-ilvC pET20b(+):4HBT-ACX4-bsBCKD	138
Table 30	Summary of screening results for production of methacrylic acid by <i>E. coli</i> BL21 (DE3) pKIV pHAPP, <i>E. coli</i> BL21 (DE3) pHAPP and <i>E. coli</i> BL21 (DE3) pKIV pHABS in MSX minimal medium.	139
Table 31	Summary of methacrylic acid and isobutyric acid production by <i>E. coli</i> BL21 (DE3) pKIV pHAPP and <i>E. coli</i> BL21 (DE3) pKIV pHABS in modified SM medium	141
Table 32	Summary of results from 'Microorganisms and methods for producing acrylate and other products from propionyl-CoA' for the production of acrylic acid.	146
Table 33	Table of additional control plasmids to use in the biotransformation of isobutyric acid to methacrylic acid and the production of methacrylic acid from glucose	149
Table 34	List of reviewed acyl-CoA thioesterases in UniProt database with X-Ray crystal structures	XVII

Table 35	Summary table showing the overexpression conditions for each thioesterase gene (4HBT, FcoT, RifR, YciA) in pET20b(+) vector, transformed into <i>E. coli</i> BL21 (DE3) pLysS.	XVIII
Table 36	Table of total protein concentrations for cell free extracts of 4HBT, YciA and FcoT thioesterases, as well as the <i>E. coli</i> BL21 (DE3) pLysS pET20b(+) control strain lysate.	XIX
Table 37	Raw data for Lineweaver Burk plots during kinetic characterisation of CtHIS-4HBT on methacrylyl-CoA and isobutyryl-CoA as substrates.	XXI

List of abbreviations

2,5-DKG	2,5-diketo-D-gulonic acid
2KLG	2-keto-L-gulonic acid
3HIBA	3-hydroxyisobutyric acid
4HBT	4-Hydroxybenzoyl-CoA Thioesterase
ABE	Acetone-Butanol-Ethanol
ACAD	Acyl-CoA dehydrogenase
ACD	Acyl-CoA dehydrogenase
ACH	Acetone cyanohydrin
ACX	Acyl-CoA oxidase
ACX / ACOX	Acyl-CoA oxidase
ADP	Adenosine diphosphate
AHAIR	Acetohydroxyacid Isomeroeductase
AHAS	Acetohydroxyacid Synthase
ALS	Acetolactate synthase
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
BCKD	Branched chain keto acid dehydrogenase
CtHIS-4HBT	Carboxy terminal hexahistidine tagged 4-Hydroxybenzoyl-CoA Thioesterase
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
dGTP	Deoxyguanosine triphosphate
DHAD	Dihydroxyacid Dehydratase
DHAP	Dihydroxyacetone phosphate
DNA	Deoxyribosenucleic acid
DTNB	5'5'-dithio-bis-(2-nitrobenzoic acid)
dTTP	Thymidine triphosphate
EDTA	Ethylenediaminetetraacetic acid
ETC	Electron transport chain
ETF	Electron transfer flavoprotein
ETFQO	Electron transfer flavoprotein ubiquionone oxidoreductase
ETFSQ	Electron transfer flavoprotein semiquinone
FAD	Flavin adenine dinucleotide
FADH ₂ / FAD _{HQ}	Fully reduced flavin adenine dinucleotide / Flavin adenine dinucleotide hydroquinone
FAD _{SQ}	Flavin adenine dinucleotide semiquinone
GD	Glutaryl-CoA dehydrogenase
HCN	Hydrogen cyanide
IBA	Isobutyric acid
IBD	Isobutyryl-CoA dehydrogenase
IPTG	Isopropyl-β-D-thiogalactopyranoside
IVD	Isovaleryl-CoA dehydrogenase
LCAD	Long chain acyl-CoA dehydrogenase
MAA	Methacrylic acid
MCAD	Medium chain acyl-CoA dehydrogenase
MMA	Methyl methacrylate
NAD ⁺	Nicotinamide adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide
NADP ⁺	Nicotinamide adenine dinucleotide phosphate
NADPH	Reduced nicotinamide adenine dinucleotide phosphate

NtHIS-4HBT	Amino-terminal hexahistidine tagged 4-Hydroxybenzoyl-CoA Thioesterase
OE-PCR	Overlap extension polymerase chain reaction
PCR	Polymerase chain reaction
PD	Primary dehydrogenase
P _i	Inorganic phosphate
PMMA	Polymethyl methacrylate
PP _i	pyrophosphate
RBS	Ribosome Binding Site
RNA	Ribonucleic acid
SBD	Short branched chain acyl-CoA dehydrogenase
SCAD	Short chain acyl-CoA dehydrogenase
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
TAE	Tris-Acetate-EDTA
TMADH	Trimethylamine dehydrogenase
Tris	Tris(hydroxymethyl)aminomethane
UQ	Ubiquinone
UQ:H ₂	Ubiquinol
vLCAD	Very long chain acyl-CoA dehydrogenase

1 Introduction

The aim of this project was to biocatalytically produce methacrylic acid (MAA), a commodity chemical used for the manufacture of methyl-methacrylate (MMA) and similar esters, which polymerise to form polymethylmethacrylate (PMMA) and co-polymers thereof. PMMA and its co-polymers are extruded or cast into acrylic sheets, with trade names of Perspex[®], Lucite[®] and Plexiglas[®] [1].

The market demand for methyl methacrylate is 2 million tonnes per annum [2] and is currently produced solely through chemical means from intermediates of the petrochemicals industry. The chemical production of methacrylates is served predominantly by six different processes (summarised in literature survey), forming either methacrylic acid or methyl-methacrylate as end products, and requiring complex metal catalysts and corrosive acids, as well as other harsh conditions [3]. Fluctuating oil prices, as well as a general trend by chemicals companies to move towards more sustainable and environmentally friendly routes to their products, has culminated in the discipline of green chemistry [4]. Biotechnology and biocatalysis have the ability to achieve the 12 principles of green chemistry [5], with the potential to reduce hazardous waste production, to operate under mild conditions, lower energy consumption, reduce waste and operate from renewable feedstocks such as glucose [6].

Biocatalysis has long been used for the production of fine chemicals and active pharmaceutical ingredients due to the capacity of enzymes to catalyse reactions with chemo-, regio- and stereo-selectivity [7-9], but is rapidly gaining ground in the production of bulk chemicals [10, 11]. The rise of metabolic engineering has furthered the scope for bulk chemicals and fuels production by fermentation [7], but it is well recognised that one of the biggest hindrances to the broader application of biocatalysis to many of the products of the chemical industry is the availability of enzymes that catalyse the desired chemical reactions for non-natural products [10]. However, the use of enzymes demonstrating catalytic or substrate promiscuity has made the production of non-natural products possible in many cases [12-15].

Methacrylic acid is a non-natural product, with no known metabolic pathways that produce it as an intermediate. Fermentative routes to isobutyric acid (IBA) [16] and 3-hydroxyisobutyric acid (3HIBA) [17] have both been developed as precursors to existing chemical processes for the combined bio- and chemo- catalytic production of methacrylic acid, and a new chemical route for the production of

methacrylic acid from citric acid, a well established product of the biotechnology industry [18], has also been developed [19]. However, an entirely biocatalytic route to methacrylic acid remains appealing, to minimise the use of traditional chemical processes that use harsh conditions and produce copious waste. Recently, hypothetical biocatalytic transformations of naturally occurring intermediates to both methacrylic acid and esters of methacrylic acid were proposed [20], but none have been substantiated.

In this thesis, a biocatalytic route to methacrylic acid based on the hydrolysis of methacrylyl-CoA to methacrylic acid and coenzyme A has been developed, as previously proposed but unsubstantiated [20]. Enzyme databases were searched for candidate enzymes that could hydrolyse methacrylyl-CoA, and shortlisted enzymes were tested for this activity, as well as for their selectivity for methacrylyl-CoA over similar substrates in the pathway. Next, as methacrylyl-CoA is not a natural metabolite in *Escherichia coli* (a favourable microbial host for metabolic engineering), steps were taken to produce methacrylyl-CoA from readily available feedstocks. The first step was to identify an enzyme to catalyse the oxidation of isobutyryl-CoA to methacrylyl-CoA that would function in *E. coli*. Next, a step for the activation of isobutyric acid to isobutyryl-CoA was introduced, so as to create a strain for the biotransformation of isobutyric acid to methacrylic acid. Finally, the oxidative decarboxylation of 2-ketoisovaleric acid to isobutyryl-CoA was introduced into *E. coli*, to create a link with the natural metabolism of *E. coli*, such that methacrylic acid could be produced in fermentations directly from glucose. Therefore, this project established a proof of principle for the biocatalytic production of methacrylic acid using whole cell biocatalysts, and a patent based on the results of this project is currently pending (Patent application number: GB1508582.2).

2 Literature Review

Methacrylic acid, methyl methacrylate, ethyl methacrylate and *n*-butyl methacrylate (Figure 1), amongst others, are key components of the acrylics industry [3, 21]. With an annual production greater than 2 million tonnes [2], methyl methacrylate is by far the most important, and undergoes free radical polymerisation to polymethyl methacrylate (PMMA), otherwise known as acrylic glass [3, 21].

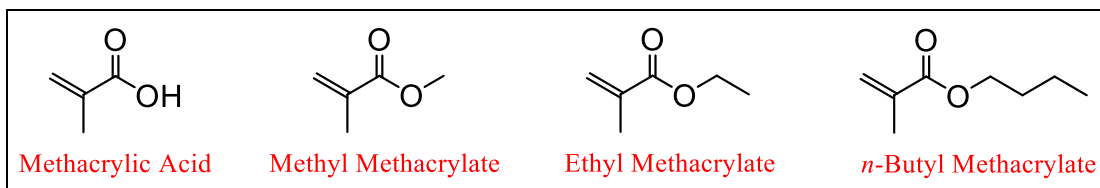


Figure 1: Methacrylic Acid, Methyl Methacrylate, Ethyl Methacrylate and *n*-Butyl Methacrylate.

PMMA is a colourless and hard material that is more rigid than, and softens at a higher temperature than polymers of corresponding acrylate esters [22]. This is explained by the bulky branched methyl group on the α -carbon, which prevents packing and rotation around the carbonyl- and α -carbon [21]. PMMA resins can be cast or extruded into acrylic sheets, injection moulded, and can be sawed and carved into desired shapes [23].

With superior optical clarity to glass, weather resistance, chemical resistance, toughness and resistance to scratching, demand for acrylic sheet grew during the Second World War for the canopies of military aircraft [21, 23], whilst the rise of commercial flights further increased demand after the war [23]. Contemporary usage includes, but is not limited to, the fronting of solar panels, signage, aquarium glass, greenhouse glass, shower glass and, being biocompatible, PMMA is also used in bone-cements [21, 23]. Current production methods of methacrylic acid and its esters (overleaf) focus on the direct production of either methacrylic acid or methyl methacrylate, with ethyl methacrylate, *n*-butyl methacrylate and other esters of methacrylic acid being produced by the direct esterification of methacrylic acid, or by the transesterification of methyl methacrylate [3].

2.1 Current industrial routes to methacrylic acid and methyl methacrylate

Today, methyl methacrylate and methacrylic acid are produced through 7 key routes. The first commercial route to methacrylic acid and methyl methacrylate was the acetone cyanohydrin (ACH) route (Figure 2) [3, 24].

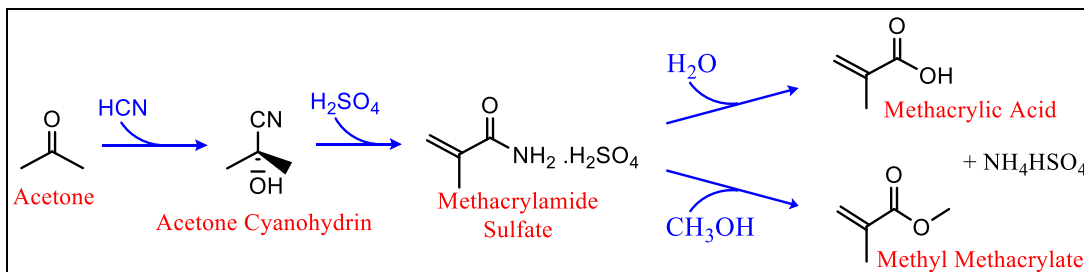


Figure 2: Chemical steps to the synthesis of Methacrylic Acid and Methyl Methacrylate by the ACH route.

In this route, acetone is reacted with hydrogen cyanide with a basic catalyst to form acetone cyanohydrin [3, 24]. Acetone cyanohydrin is then reacted with concentrated sulphuric acid, which is supplied in anhydrous form and in excess to form methacrylamide sulphate *via* a perceived α -sulfatoamide intermediate (not shown). Methacrylamide sulfate is then hydrolysed with water to yield methacrylic acid or with methanol dissolved in water to yield a mixture of methacrylic acid and methyl methacrylate [24]. Whilst some modifications to the acetone cyanohydrin route have been employed throughout the years, this method is still the main source for methacrylic acid and methyl methacrylate today [3, 24].

Despite high yield from ACH (80-90%) [24], the ACH route is characterised by an acetone cost that is linked to the phenols market, dependence on hazardous hydrogen cyanide, as well as the generation of a lot of waste. For example, 1.5 tonnes of ammonium bisulphate are generated for each tonne of methyl methacrylate. Whilst this can be treated to regenerate sulfuric acid, sulfuric acid is highly corrosive and requires exotic materials for its storage [25]. Furthermore, the production economics are highly linked to the cost of making hydrogen cyanide, which is produced catalytically from natural gas, ammonia and air, and so spikes in the wholesale cost of natural gas can render this route economically unviable [25]. Alternative routes to methyl methacrylate and methacrylic acid were developed, predominantly to avoid the need for the sulphuric acid regeneration step.

A less common yet direct route to methyl-methacrylate is achieved by reacting propyne with carbon monoxide and methanol over a palladium catalyst (Figure 3). Whilst this route has the advantage that the catalyst is highly selective (>99% selectivity), propyne is a byproduct of ethylene production, and there is insufficient

propyne byproduct from a single plant to sustain a methyl methacrylate plant with a capacity of over 45,000 tons per year, forcing a reliance on multiple ethylene plants [3].

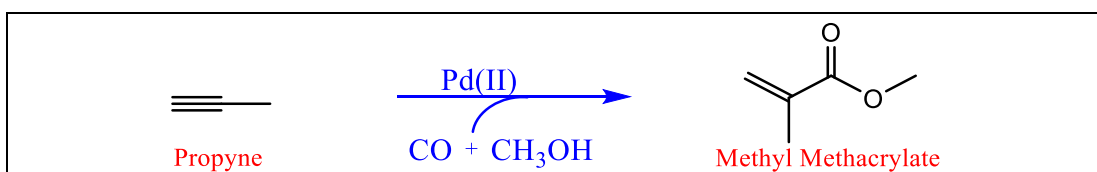


Figure 3: A C3 route to methyl methacrylate using propyne, carbon monoxide and methanol using a palladium catalyst with a phosphine ligand.

Another C3 feedstock, propylene, is also a byproduct of ethylene synthesis and is found in greater quantities than propyne. It can be used to produce methacrylic acid *via* a two-step process (Figure 4). However, a typical acid for the first step is hydrofluoric acid, which poses problems in its handling due to its highly corrosive nature. Typically, an iron-phosphorous or molybdenum phosphorous catalyst is used for the second step, at temperatures between 300 °C – 400 °C. The crude reaction mixture is then esterified, creating a mixture of methyl methacrylate and methyl isobutyrate. The methyl methacrylate is then separated from the methyl isobutyrate by distillation [3].

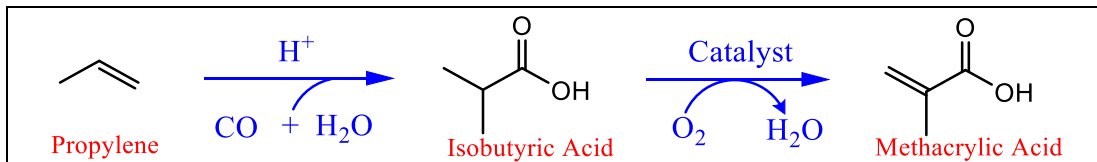


Figure 4: A C3 route utilising propylene and carbon monoxide as carbon feeds.

In the 1980s, a C4 route was also developed using isobutylene as a carbon feed. Isobutylene is oxidised to methacrylic acid *via* methacrolein (Figure 5). Isobutylene can be purchased directly, or obtained by the dehydration of *t*-butanol [3, 24, 25].

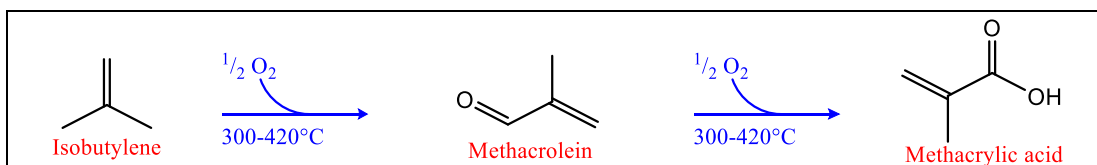


Figure 5: C4 route to methacrylic acid from isobutylene. The C4 route to methacrylic acid can take either isobutylene or *t*-butanol as feedstock as *t*-butanol converts to isobutylene under the physical conditions of the first oxidation step.

Isobutylene is oxidised to methacrolein using catalytic mixed metal oxides of molybdenum, bismuth and iron, although variations exist to improve process yield and catalyst lifetime. The second oxidation step typically uses a phosphomolybdic acid based catalyst, often doped with copper, vanadium or other heavy alkali

metals [3, 24]. The maximum claimed yield of this process is 86%, although typically only 65%-75% yield is achieved. Additionally, there is low selectivity of the catalyst, and other products are common. The catalysts are complex, expensive and have a limited life and the entire process is limited by scale. Finally, the process requires vast quantities of water, where methacrylic acid is only 20-40% of an aqueous solution before purification, the process generates copious contaminated water [3, 24, 25].

Methacrylic acid has also been produced through various C2 routes from ethylene. In 1990, an ethylene based plant began operation by BASF in Germany using a propanal intermediate (Figure 6A) [3]. Ethylene is catalytically hydroformylated to propanal while a liquid phase reaction of propanal with formaldehyde yields methacrolein and water [3, 24]. Methacrolein is then catalytically oxidized to methacrylic acid, as occurs in the C4 route from isobutylene (Figure 5). A plant operated by the Eastman Chemical Company at Kingsport, Tennessee, produced methacrylic acid from ethylene via propionic acid (Figure 6B). Ethylene is catalytically hydrocarbonylated to propionic acid and a condensation reaction between propionic acid and formaldehyde yields methacrylic acid and water [3, 24].

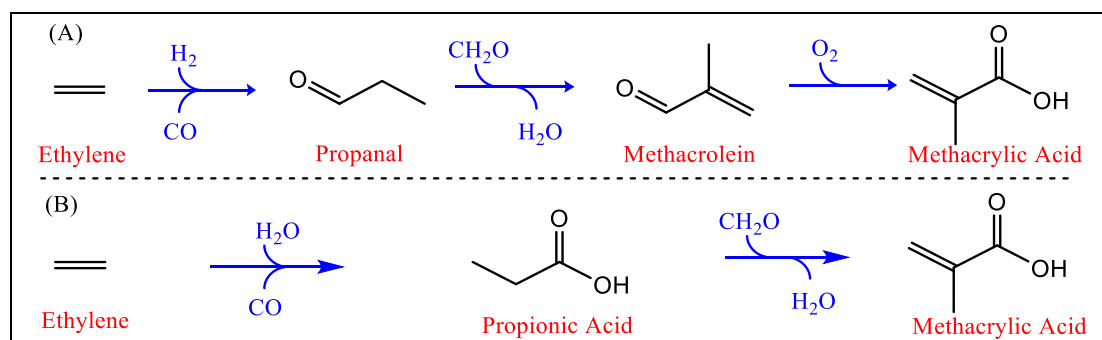


Figure 6: C2 routes to methacrylic acid from ethylene *via* propanal or propanoate. (A) The production of methacrylic acid from ethylene, *via* propanal and methacrolein. (B) The production of methacrylic acid from ethylene *via* propionic acid.

A third C2 route to methacrylic acid has been developed and patented by Lucite International as the Alpha process (Figure 7), and a commercial plant has begun operation in Singapore. This process also utilises ethylene, which reacts with carbon monoxide and methanol to form the intermediate methyl propionate. In a single step, methyl propionate is reacted with formaldehyde to yield methyl methacrylate.

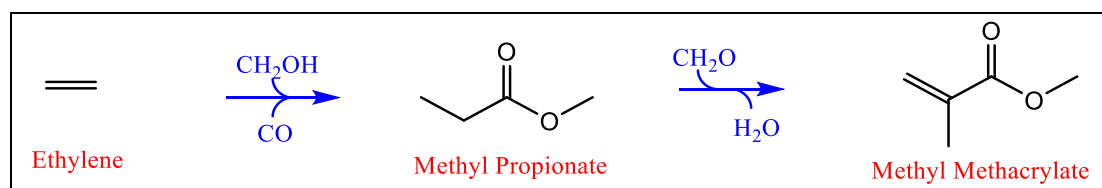


Figure 7: A C2 route to methacrylic acid via methyl propionate, the Alpha process

This process takes advantage of a highly selective, high turnover “enzyme-like” silica-zirconium-caesium catalyst. The operation occurs at high temperature, low pressure and has a complex but small byproduct spectrum. There is a reduced reliance on crude oil derived feedstocks and much of the heat required to drive the process can be derived from adjacent power stations due to the Alpha process being a low-grade heat sink [25].

All of the current routes to methacrylic acid and methyl methacrylate require carbon feedstocks that are ultimately derived from crude oil and natural gas, and often require exotic catalysts, corrosive acids and high operating temperatures. The recent Green Chemistry trends seek to reduce the environmental impact of such processes, whilst making them more sustainable [4]. Industrial biotechnology offers promise in this area [5], by enabling the production of chemicals under mild conditions and from renewable resources such as sugars or other feedstocks, either through the direct production of the target compound, or through the production of feedstocks or intermediates of established chemical routes from bio-based resources.

2.2 Biotechnology and biocatalysis for the sustainable production of useful chemicals

Modern industrial biotechnology typically involves the growth of a micro-organism which produces a compound of interest, under controlled conditions in a bioreactor. Perhaps the best examples for the biotechnological production of chemicals over the 20th century are citric acid fermentation, penicillin production and the acetone-butanol-ethanol (ABE) fermentation [26]. Citric acid is an intermediate of the tricarboxylic acid (TCA) cycle, and over 90% of global demand for citric acid (1.5 million tonnes) is produced *via* microbial fermentation [27]. *Aspergillus niger*, a filamentous fungus, is the predominant organism used in these microbial fermentations [27]. Citrate biosynthesis requires two pyruvates (Figure 8), which are both obtained from one glucose *via* glycolysis (Figure 9).

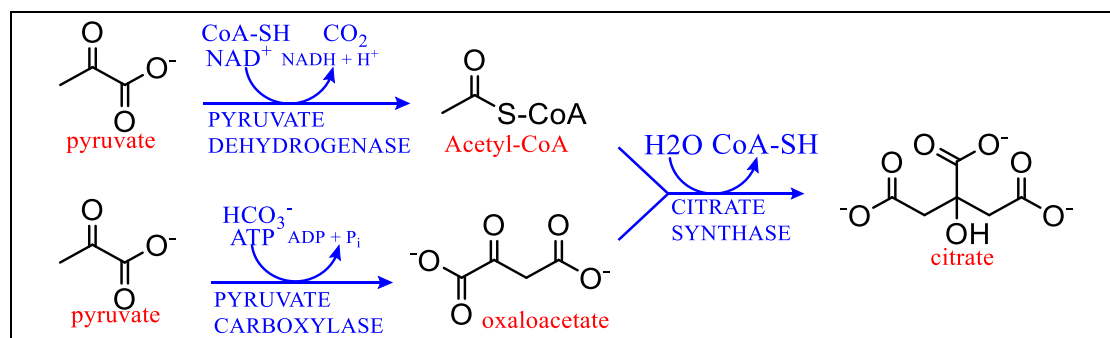


Figure 8: Metabolic pathway to citrate from pyruvate

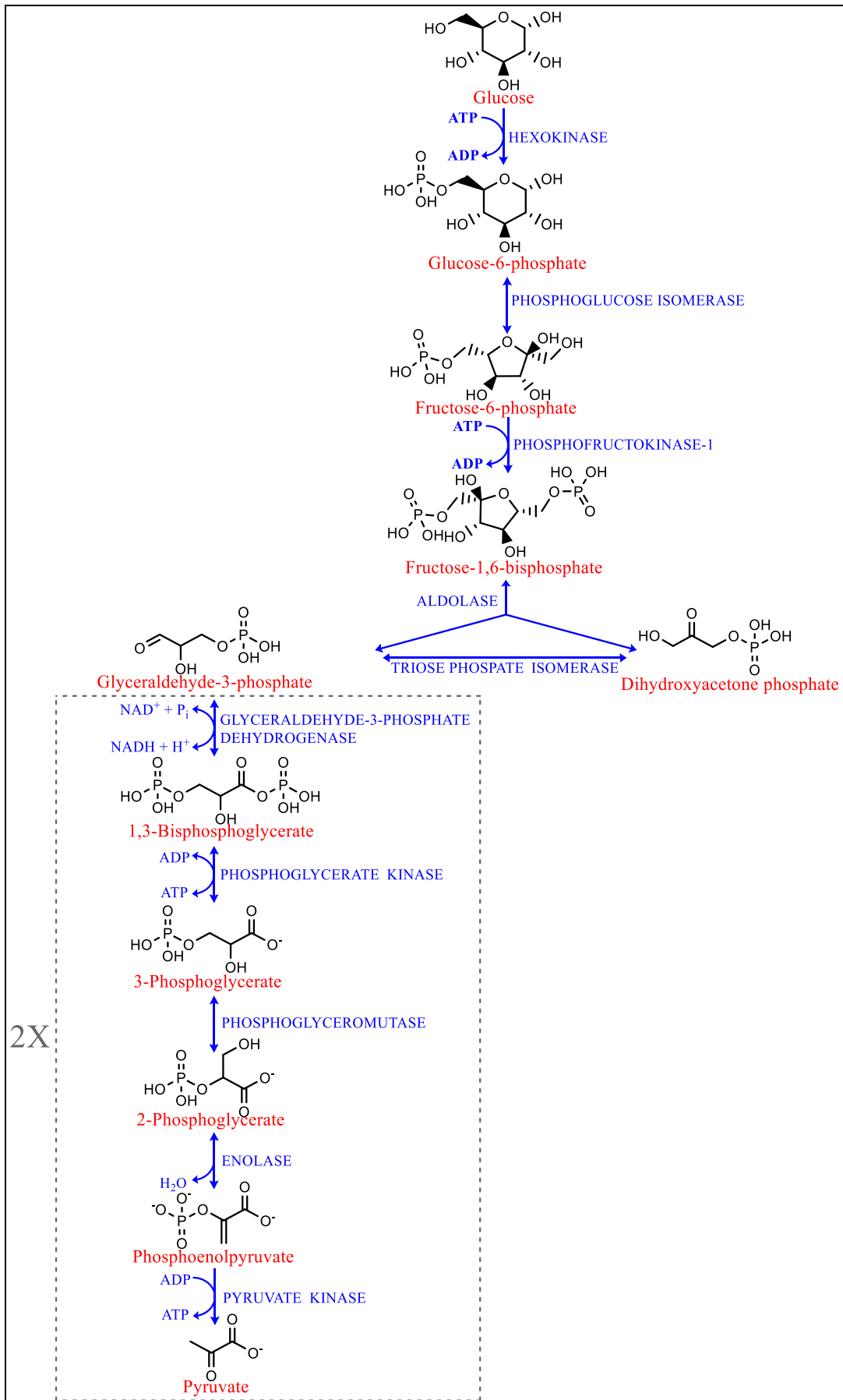


Figure 9: An example of a glycolytic pathway, the Embden-Meyerhof-Parnas pathway [28].

The production of penicillin by another filamentous fungus, *Penicillium notatum*, has also been of critical significance in the biotechnology industry. The productivity of penicillin biosynthesis by *Penicillium notatum* has been gradually improved through classical strain improvement techniques, notably random mutagenesis by techniques such as X-ray and ultra-violet irradiation, followed by strain selection and characterisation from just trace concentrations of penicillin to milligram per ml scale production [26].

Finally, though economically unable to compete with solvents production from crude oil since the late 1950s, ABE fermentation was once the predominant method for the production of acetone and butanol [29]. Demand for acetone from ABE fermentation grew rapidly from the onset of the First World War for the manufacture of the explosive cordite, as its traditional synthesis from acetate, which used to be delivered in a large part by Austria, was for obvious reasons, restricted. Though deemed an unwanted byproduct of acetone fermentation at the time, the butanol produced from ABE fermentation, and its acetate ester, was quickly put to use in the manufacture of lacquers [29]. Through a combination of increased competition for the feedstock of ABE fermentation, molasses, and competition from the synthetic production of solvents from petrochemicals, ABE fermentation quickly became uneconomical. However, with the energy and oil crises of the 1970s, as well as more recent concerns on climate change and oil supplies, interest in butanol fermentation has re-emerged, for its use as a carbon-neutral biofuel. This time, however, biotechnologists are aiming to take advantage of the advances in molecular biology, recombinant DNA technology and metabolic engineering, as well as systems and synthetic biology, to not only improve upon old processes such as butanol fermentation [30] and penicillin biosynthesis, but to create new and varied bioprocesses to drive the 21st century bioeconomy.

Whilst the discipline of biochemistry and molecular biology began with the understanding that enzymes were biological catalysts [31], the determination of the structure of DNA [32] and the description of the central dogma of molecular biology [33], the rational and targeted modification of an organism's metabolism began with the development of recombinant DNA technology [34]. The first significant step towards the development of recombinant DNA technology began with the discovery of DNA ligases in 1967 [35] followed by the first discovery of a type II restriction endonuclease [36, 37]. Soon after these discoveries, the first recombinant section of DNA was constructed from heterologous parts [38] which culminated in the first *in vitro* construction of a biologically functional plasmid [39]. Plasmids are extrachromosomal sections of DNA, typically but not necessarily

circular, and replicate in an autonomous manner using the host cell's replication machinery [40]. Plasmids occur naturally, with some genes in many organisms being plasmid borne and not genome borne [41]. Soon, molecular biologists and chemical engineers began to realise the possibilities of recombining genes into plasmids to encode heterologous metabolic pathways and the discipline of metabolic engineering was born [42], which is now defined as "the directed modulation of metabolic pathways using methods of recombinant technology for the purpose of overproducing fuels, chemicals and pharmaceutical products" [43].

In one of the first examples of metabolic engineering, molecular biologists sought to engineer a micro-organism for the production of 2-keto-L-gulonic acid (2KLG), the final precursor for the chemical synthesis of L-ascorbic acid, from glucose. A gene from *Corynebacterium sp.* encoding 2,5-diketo-D-gluconic acid reductase, which catalyses the conversion of 2,5-diketo-D-gluconic acid (2,5-DKG) to 2KLG, was cloned into *Erwinia herbicola*, which was known to produce 2,5-DKG from glucose. The molecular biologists were successful in engineering this organism to produce 2-KLG directly from glucose [44] and subsequent work using a combination of recombinant DNA technology as well as classical strain improvement techniques resulted in the the ability to ferment glucose to 2KLG to titers of $>120 \text{ g.L}^{-1}$ [45-47].

Since the advent of recombinant DNA technology, the invention of the polymerase chain reaction (PCR) in the early-mid 1980s [48] allowed molecular biologists to modify genes *in vitro*, as well as to string together multiple genes through methods such as overlap extension polymerase chain reaction [49]. In addition, techniques were developed for the targeted chromosomal gene deletion to remove non-productive reactions on metabolic pathway intermediates [50]. The more recent development of Gibson assembly has enabled the rapid assembly of long DNA molecules from shorter DNA parts, enabling the construction of synthetic operons and multiple expression cassettes on a single expression vector [51]. Finally, though not a new technology, the chemical synthesis of oligonucleotides and polynucleotides has decreased dramatically in cost over recent years, making custom gene synthesis cost competitive with cloning [52]. This has two significant implications. Firstly, the time savings achieved by the custom synthesis of a gene, artificial operon, or even multiple expression cassettes on a single polynucleotide, enable the scientist to spend time on fermentation characterisations and optimisations, and other strain improvements. Secondly, by having control over the gene sequence encoding a polypeptide, the gene can be codon optimised for the expression host, ensuring good levels of recombinant protein expression [53].

With these developments, more milestone examples of metabolic engineering have been observed. One of these is the metabolic synthesis of 1,3-propanediol directly from glucose [54]. Various organisms have been previously known to anaerobically ferment glycerol to 1,3-propanediol (Figure 10A) when glycerol was the only carbon source [55]. Whilst the market price of glycerol has dropped significantly over recent years, glucose was deemed a cheaper feedstock, and the ability to produce 1,3-propanediol in aerobic cultures was deemed likely to yield a bioprocess with higher productivity. Yeast has been widely known to accumulate glycerol under various conditions [56], and the enzymes responsible for its formation from dihydroxyacetone phosphate (Figure 10B), an intermediate of glycolysis, have been identified.

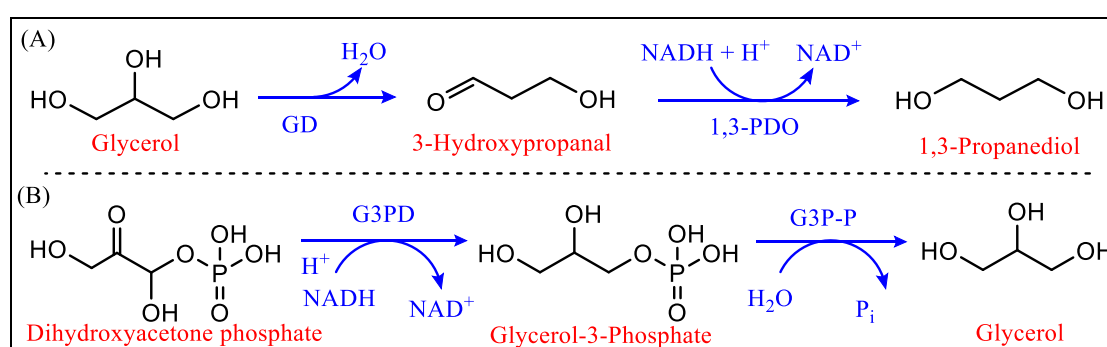


Figure 10: (A) Glycerol fermentation pathway to 1,3-propanediol. GD = Glycerol dehydratase. 1,3-PDO = 1,3-Propanediol oxidoreductase, as occurs in *Klebsiella aerogenes* [57]. (B) Formation of glycerol from dihydroxyacetone phosphate. G3PD = Glycerol-3-Phosphate Dehydrogenase. G3P-P = Glycerol-3-Phosphate Phosphatase.

For the production of 1,3-propanediol directly from glucose, metabolic engineers cloned a glycerol-3-phosphate dehydrogenase and a glycerol-3-phosphate phosphatase from *Saccharomyces cerevisiae* into *E. coli* along with three genes encoding the component polypeptides of glycerol dehydrogenase from *Klebsiella pneumoniae*. The engineers relied on an NADPH utilizing aldehyde oxidoreductase that was already present in *E. coli*, as opposed to cloning an NADH utilizing enzyme from another organism, to complete the pathway to 1,3-propane-diol [54].

Finally, the engineers performed a genetic knockout of the *tpi* gene in *E. coli* [54], which encodes triose phosphate isomerase of the glycolysis pathway (Figure 9). Though this prevented any replenishment of DHAP from glyceraldehyde-3-phosphate, and imposed a maximum theoretical yield of 42.5% (g/g) on glucose, it prevented the remainder of the glycolysis pathway from competing with the 1,3-propanediol pathway for dihydroxyacetone phosphate, and actually doubled the yield of 1,3-propanediol compared to the equivalent strain which did not have *tpi* deleted. The work performed on *E. coli* for the production of 1,3-propanediol

resulted in the commercial bioprocess for industrial 1,3-propanediol production, and with some further fine-tuning of enzyme expression levels and other non-pathway modifications, a final titre of 135 g.L⁻¹ was achieved, which is considerably higher than the 78 g.L⁻¹ titres achieved in fermentations from glycerol using non-engineered cells [54].

The above examples, however, describe the construction of heterologous metabolic pathways for intermediates that are known to exist in metabolism. Many industrially useful chemicals do not occur naturally in any known metabolic pathway, however, but this has not stopped the development of bio-catalytic routes to non-natural chemical products. Though enzymes are typically viewed as being both highly specific in the type of reaction that they catalyse, as well as the substrate that they accept, enzymes with catalytic and/or substrate promiscuity are known to exist in nature, and it is widely believed that it is from such enzymes that other enzymes with diverse activities and substrate ranges, have evolved [58, 59].

The potential to use enzymes with some substrate promiscuity has been highlighted very well through the development of a biocatalytic route to isobutene, a platform chemical with many uses [15], by Global Bioenergies, who are now scaling up their production to pilot scale. In this work, a mevalonate diphosphate decarboxylase (EC 4.1.1.33), which catalyses the concomitant decarboxylation and dehydration of mevalonate diphosphate to isopentenyl-diphosphate (Figure 11A), has shown activity on a different substrate, 3-hydroxyisovalerate, which is concomitantly decarboxylated and dehydrated to isobutene (Figure 11B).

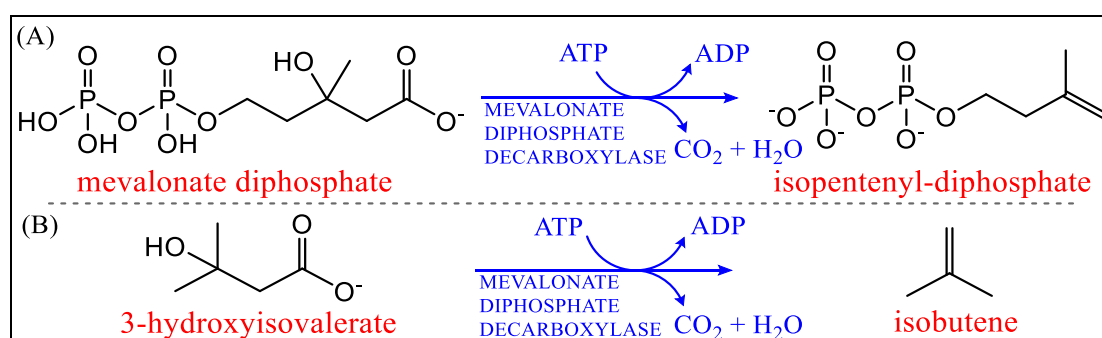


Figure 11: (A) The enzymatic production of isopentenyl-diphosphate by the concomitant dehydration and decarboxylation of mevalonate diphosphate by mevalonate diphosphate decarboxylase. (B) The enzymatic production of isobutene by the concomitant dehydration and decarboxylation of 3-hydroxyisovalerate by mevalonate diphosphate decarboxylase.

More recently, Zhang *et al.* published a patent in which *E. coli* was engineered to produce isobutyric acid [16]. Two routes to isobutyric acid were tested. In the first, 2-ketoisovalerate was oxidatively decarboxylated to isobutyryl-CoA by a branched chain keto acid dehydrogenase (BCKD), and isobutyryl-CoA was hydrolysed to

isobutyric acid by a broad specificity thioesterase (Figure 12A). Zhang *et al.* achieved shake-flask titres of 8.6 g.L⁻¹ isobutyrate from 40 g.L⁻¹ glucose, which is a yield of 0.22 g.g⁻¹ glucose, 44% of the maximum theoretical yield) with this route. In the second route by Zhang *et al.*, 2-ketoisovalerate was first decarboxylated to isobutyraldehyde, and this was then oxidised to isobutyric acid (Figure 12B). By then knocking out various alcohol dehydrogenases that were deemed to be responsible for the conversion of the intermediate isobutyraldehyde to isobutanol, and optimising the expression of the aldehyde dehydrogenase, a strain capable of producing isobutyric acid to a yield of 0.39 g.g⁻¹ glucose (80% of maximum theoretical yield) was constructed. By then culturing this strain in a bioreactor, an industrially relevant titre of 90 g.L⁻¹ was reached [16].

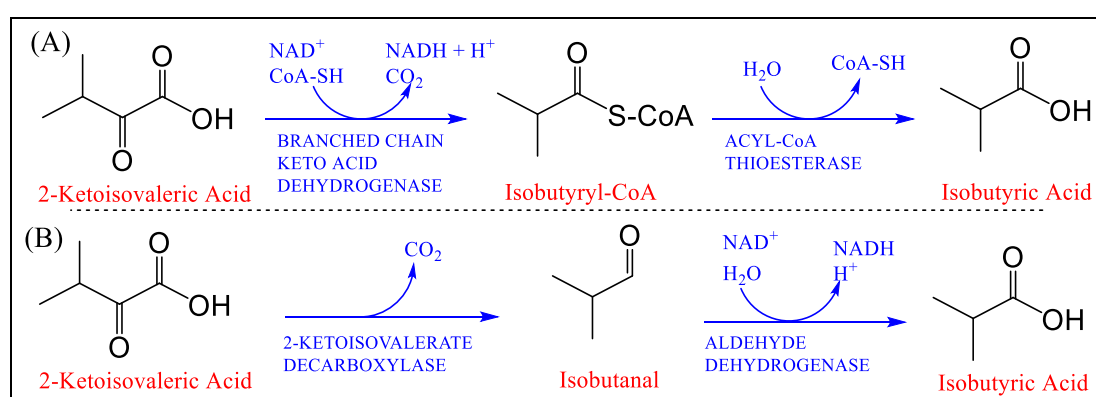


Figure 12: Two routes used by Zhang *et al.* for the biocatalytic production of isobutyric acid. (A) The production of isobutyric acid via an isobutyryl-CoA intermediate. (B) the production of isobutyric acid via an isobutyraldehyde intermediate.

The aim of Zhang *et al.* in this work was to produce sufficient isobutyric acid such that it could feed into the existing C3 route from propylene to methacrylic acid (Figure 4), whereby isobutyric acid is a chemical intermediate that is catalytically oxidised to methacrylic acid. This would create a combined bio- and chemo-catalytic route to methacrylic acid and onwards to methyl methacrylate whereby the carbon feedstock is produced from renewable resources.

Numerous other combined bio- and chemo- catalytic regimes for the production of methacrylic acid have also been proposed. Indeed, the C4 route to methacrylic acid from isobutene (Figure 5) could be fed from bio-derived isobutene, as produced by Global Bioenergies (Figure 11). Similarly, research is ongoing into bio-derived isobutanol production for transportation fuels [60, 61], which could be dehydrated to isobutene. For the ACH route to methacrylic acid and methyl methacrylate (Figure 2), sustainable routes to acetone have of course been in existence for over a century [29], and metabolic engineering activities could improve this production. The production of methanol from biomass is also being developed [62, 63] for the

esterification of methacrylic acid to methyl methacrylate, and hydrogen cyanide, which is ordinarily produced by the reaction of methane and ammonia with air [64], can be produced from bio-methane, ammonia and air. For C2 routes, routes to bio-ethylene are being developed [65] and the formaldehyde can be derived by the catalytic oxidation of the bio-derived methanol [66].

Furthermore, new chemical processes have been developed for the production of methacrylic acid from alternative bio-based resources in combined bio- and chemo-catalytic processes. For example, citric acid production is already one of the most successful examples of industrial biotechnology [27], and the conversion of citric acid to methacrylic acid has been demonstrated in near-critical and supercritical water through an itaconic acid intermediate [19] (Figure 13).

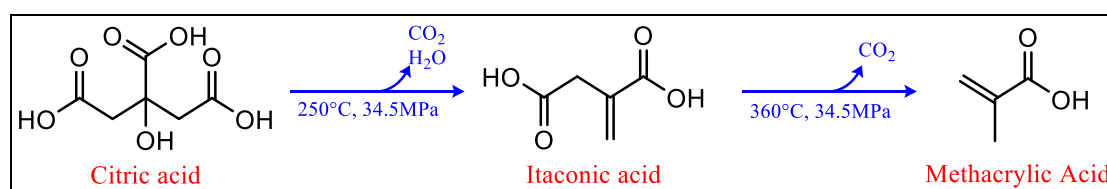


Figure 13: The production of methacrylic acid from citric acid in near-critical and supercritical water.

Finally, a chemical dehydration of biocatalytically derived 3-hydroxyisobutyric acid to methacrylic acid has been proposed. This approach is being pursued by Evonik Industries, however, and patents have already been filed for the biocatalytic production of 3-hydroxyisobutyric acid [67, 68].

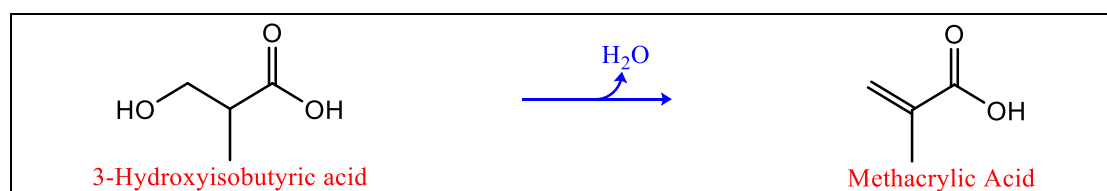


Figure 14: The dehydration of 3-hydroxyisobutyric acid to methacrylic acid.

With all processes however, the more steps involved, the greater the cost and the lower the overall yield. Thus, combined bio- and chemo- catalytic routes often have an additional infrastructure cost compared to a direct one-pot production of the same product. Furthermore, the requirement for exotic catalysts, corrosive acids, or extreme temperatures and conditions for the existing routes to methacrylic acid is not evaded. Whilst the new chemical routes to methacrylic acid, the dehydration of 3-hydroxyisovaleric acid and the supercritical water conversion of itaconic and citric acids, are perhaps more environmentally friendly, most of the market demand is for methyl-methacrylate, and so the need to esterify the methacrylic acid with bio-derived methanol still exists. This would add a further step to this process that

is already at least two steps. A direct metabolic route to methacrylic acid or one of its esters, which are non-naturally occurring, is therefore desired to minimise the number of steps to bio-derived perspex.

Identification, however, of enzymatic steps for the conversion of metabolic intermediates, catalysed by enzymes with substrate promiscuity, to either methacrylic acid or one of its esters is a challenging task. In synthetic chemistry, retrosynthetic analysis applies known chemical transformations in reverse on a target chemical compound, until a widely available compound is found [69]. A system to categorise enzymes by their function, the Enzyme Commission (EC) scheme, was developed by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB), whereby enzymes are grouped into one of six main activity types [70]. These EC groups are the oxidoreductases (EC group 1), transferases (EC group 2), hydrolases (EC group 3), lyases (EC group 4), isomerases (EC group 5), and ligases (EC group 6). Each parent group has three sublevels, further defining the functional nature of the activity, and eventually providing some level of substrate specificity at the fourth level, though this is not guaranteed. Furthermore, metabolic pathway databases such as MetaCyc [71] and KEGG [72] exist, which list known metabolic pathways and all known metabolic intermediates. It is therefore possible to envisage an analogous 'bioretrosynthetic analysis' whereby known enzymatic transformations are applied in reverse from a target compound, until a naturally occurring metabolite is reached. This would be an enormous task to undertake manually, but computational systems that work along similar approaches have been developed. Two such systems are SimPheny™ and Retropath.

SimPheny was developed by Genomatica Inc. recently, and was used to develop a metabolic pathway to the non-natural chemical 1,4-butanediol [73]. The SimPheny tool predicts pathways based on known transformations of functional groups, as opposed to known enzymatic transformations, and prioritises pathways with fewer metabolic steps to central metabolism, fewer numbers of un-natural metabolic steps, and thermodynamic favourability [73]. Retropath, however, recognises that without limitation on the deviation of substrate specificity from known natural reactions, there will be an "explosion in the combinatorial space," [74] and so limits enzymatic reactions to those whose natural substrates have molecular signatures similar to the hypothetical intermediates on putative pathways to the novel product.

Similar to the SimPheny™ tool, Retropath ranks the candidate metabolic routes by thermodynamic preference. Once candidate routes have been shortlisted for the

production of the target chemical, enzyme databases such as UniProt and BRENDA can be searched for enzymes in the appropriate enzyme commission groups that might exhibit sufficient substrate promiscuity to catalyse the transformation of the novel substrate to the target non-natural product.

2.3 Possible direct biocatalytic routes to methacrylic acid

Genomatica, who have previously used their SimPheny™ software for pathway prediction for the production of 1,4-butanediol [73, 75], have published a patent describing various single step transformations of known intermediates to methacrylic acid and its esters [20]. It is very likely that Genomatica used SimPheny™ to suggest these routes. The proposed routes and further details on the metabolic pathways that the substrates belong to and the class of enzyme transformation are summarised below (Figure 15) and overleaf (Table 1). However, none of these routes have been substantiated, and it is the aim of this project to substantiate one of these proposed routes.

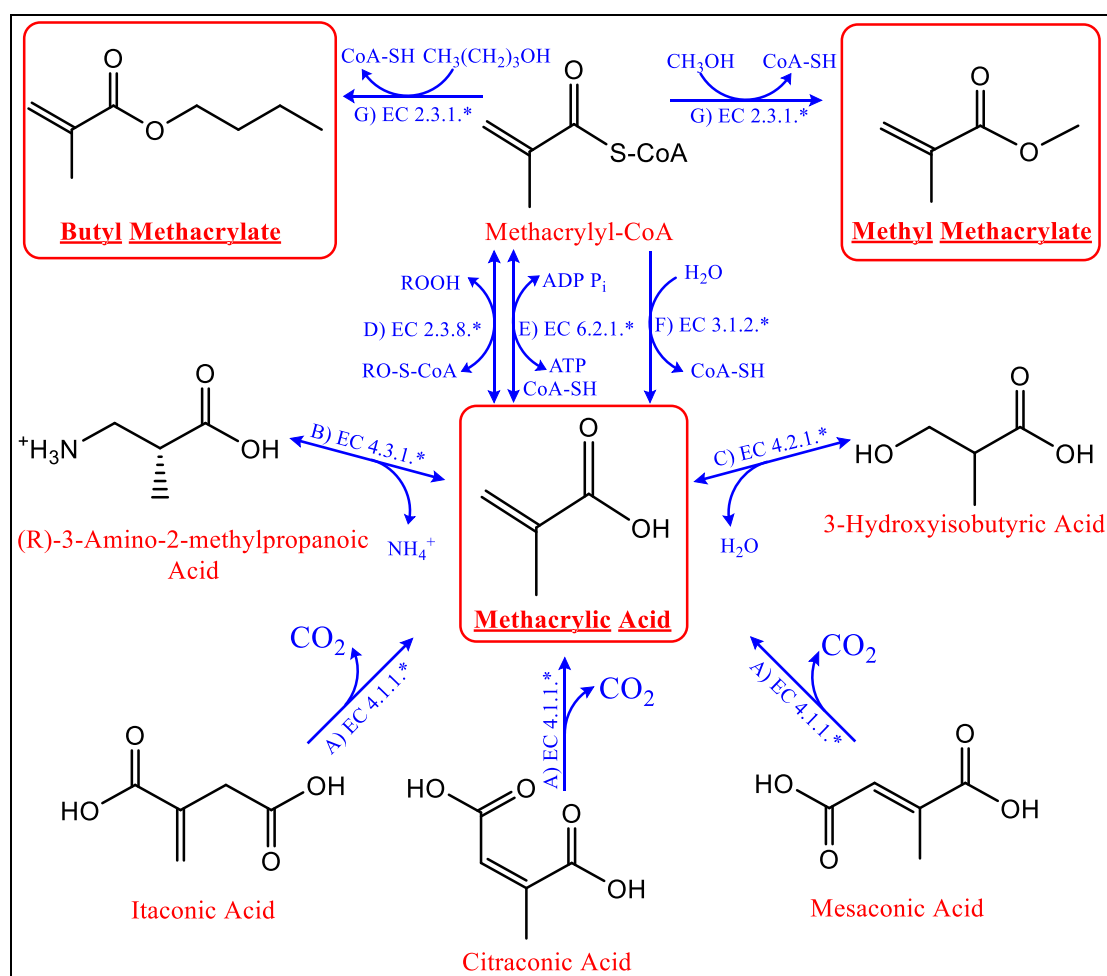


Figure 15: Routes to methacrylic acid and its esters, methyl methacrylate and butyl methacrylate, from candidate metabolic precursors. A) Decarboxylase enzyme. B) Ammonia lyase enzyme. C) Dehydratase enzyme. D) Acyl-CoA transferase E) Acid thiol ligase F) Thioesterase G) Acyl transferase transferring groups other than amino-acyl groups.

Target Product	Candidate substrate	Candidate Enzyme Class For transformation
Methacrylic Acid	(R)-3-amino-2-methylpropanoic acid Resides in: Thymine degradation pathway (MetaCyc: PWY-6430)	Ammonia lyases (EC 4.3.1.*)
	3-hydroxyisobutyric acid Resides in: Valine degradation I pathway (MetaCyc: VALDEG-PWY)	Dehydratase (EC 4.2.1.*)
	Itaconic acid Resides in: Itaconate biosynthesis pathway (MetaCyc: PWY-5750)	Decarboxylase (EC 4.1.1.*)
	Citraconic Acid Resides in: L-Isoleucine biosynthesis II pathway (MetaCyc: PWY-5101)	
	Mesaconic Acid Resides in: L-glutamate degradation VI pathway (MetaCyc: PWY-5087)	
	Methacrylyl-CoA Resides in: Valine degradation I pathway (MetaCyc: VALDEG-PWY)	Thioester hydrolase (EC 3.1.2.*) (Preferably Acyl-CoA hydrolase activity)
		Acid-thiol ligase (Preferably ADP or GDP forming acyl-CoA synthetase, must be reversible)
		Acyl-CoA transferase (EC 2.3.8.*)
Methyl-Methacrylate		Alcohol acyltransferase
<i>n</i> -Butyl-Methacrylate		

Table 1: Table of candidate transformations to yield methacrylic acid or its esters from known biological precursors

There is merit to investigating each candidate route to methacrylic acid as the different routes may have different advantages and different advantages over each other. Factors such as pathway thermodynamics, substrate availability, the number of steps to the product from central metabolism or a candidate biotransformation substrate, should be considered, as well as the likelihood of a promiscuous enzyme in the required EC group of the desired reaction being found. With regards to thermodynamic considerations, the hydrolysis of acyl-CoA thioesters (Figure 15F) is exergonic and irreversible with a typical free energy of -31 kJ.mol^{-1} [76]. Decarboxylations of dicarboxylic acids (Figure 15A) have a typical free energy change of -20 kJ.mol^{-1} [77] and are therefore also largely irreversible, but not quite as thermodynamically favourable as the hydrolysis of an acyl-CoA thioester. Similarly, as the ester bond is more stable than the thioester bond, the acyl transfer reaction, catalysed by the acyl alcohol transferases [78] to form the methacrylate esters (Figure 15G), are likely to favour the formation of the methacrylate ester, but quantitative free energy change data could not be identified for this type of reaction and so it is unclear exactly how favourable this reaction is. On the other hand, the ADP dependent hydrolysis of methacrylyl-CoA (Figure 15E) is very much reversible, as the free energy change of ATP hydrolysis to ADP and Pi is $-30.5 \text{ kJ.mol}^{-1}$, and with a free energy change of -31 kJ.mol^{-1} for the hydrolysis

of the acyl-CoA thioester, the free energy change of the reaction is -0.5 kJ.mol^{-1} . Similarly, the acyl-CoA transferase reactions (Figure 15D) are also reversible [79].

Deaminations of amino acids are typically reversible, which is the case for L-phenylalanine deamination [80], and so the proposed deamination of the non-proteinogenic amino acid (*R*)-3-amino-2-methyl-propanoate to methacrylic acid is also likely to be reversible. Similarly, many dehydrations are also reversible, such as is the case with the dehydration of 3-hydroxyisobutyryl-CoA to methacrylyl-CoA [81, 82], and so the dehydration of 3-hydroxyisobutyrate is also likely to be thermodynamically reversible.

Thus, from a thermodynamic perspective alone, the formation of methacrylic acid by the hydrolysis of methacrylyl-CoA and the decarboxylation of either mesaconic, citraconic or itaconic acids is preferred. The formation of a methacrylate ester from methacrylyl-CoA is also a possibility, but it is unclear exactly how thermodynamically favourable such a reaction would be.

As for substrate availability, methacrylyl-CoA is an intermediate of the valine degradation I (VDI) pathway (Figure 16A), which is present in mammalian cells [81], many plant cells [83, 84], and some bacteria [85]. The VDI pathway degrades valine to propionyl-CoA [86], which feeds back into the tricarboxylic acid cycle *via* methyl-malonyl-CoA and succinyl-CoA, as is the case with odd numbered fatty acid oxidation [87]. The VDI pathway is highly active, as bacteria containing the pathway are able to grow on valine [86]. In mammals and many bacteria, the oxidation of isobutyryl-CoA to methacrylyl-CoA is performed by acyl-CoA dehydrogenases [86, 88], while in some plants, it is performed by acyl-CoA oxidases [83, 84]. These are two different classes of enzyme [89]. Finally, the substrate 2-ketoisovalerate, an intermediate of the degradation I pathway, is also an intermediate of the valine biosynthesis pathway (Figure 16B) [90], and has of course been previously upregulated in *E. coli* for the production of isobutyric acid [16]. Thus the availability of methacrylyl-CoA for either its hydrolysis to methacrylic acid and coenzyme A, or its transesterification to a methacrylate ester, ought not to be a concern.

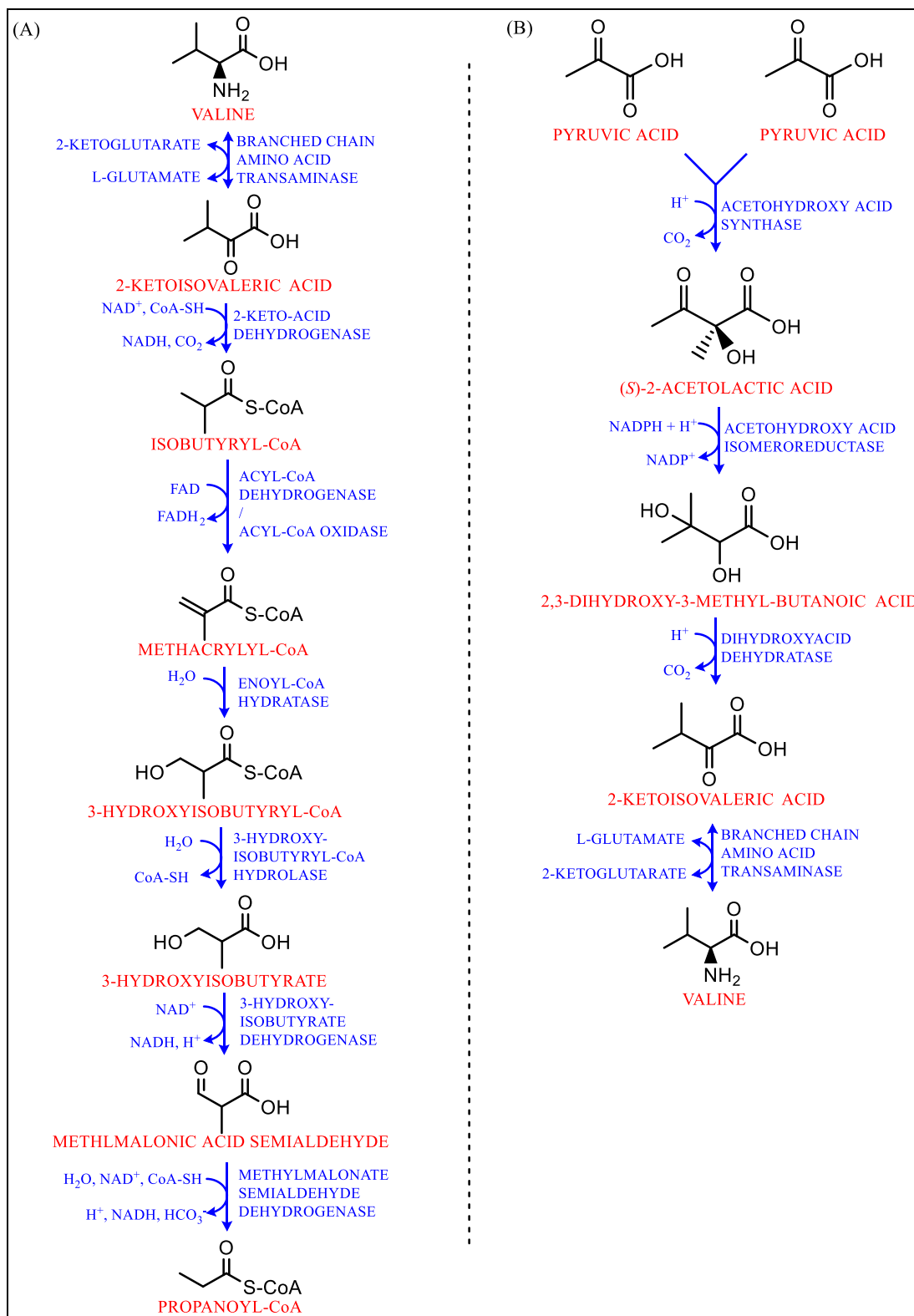


Figure 16: (A) The valine degradation I pathway. (B) The valine biosynthesis pathway

As for the substrate availability of mesaconic, citraconic and itaconic acids, according to the MetaCyc page for the L-isoleucine II biosynthesis pathway (MetaCyc PWY-5101), citraconic acid could be reached through just two enzymatic steps from pyruvate, through an (*R*)-citramalate synthase (EC 2.3.1.182) and a

citraconate forming (*R*)-citramalate-hydrolyase (EC 4.2.1.35). Similarly, the MetaCyc page for the itaconate biosynthesis pathway (MetaCyc: PWY-5750) shows that itaconate is just two enzymatic steps from citric acid, which of course, has been produced very successfully biotechnologically. As for mesaconic acid, the MetaCyc page for the L-glutamate degradation pathway (MetaCyc: PWY-5087) is also two enzyme catalysed steps away from L-glutamate, which is also produced to very high levels biotechnologically [91, 92].

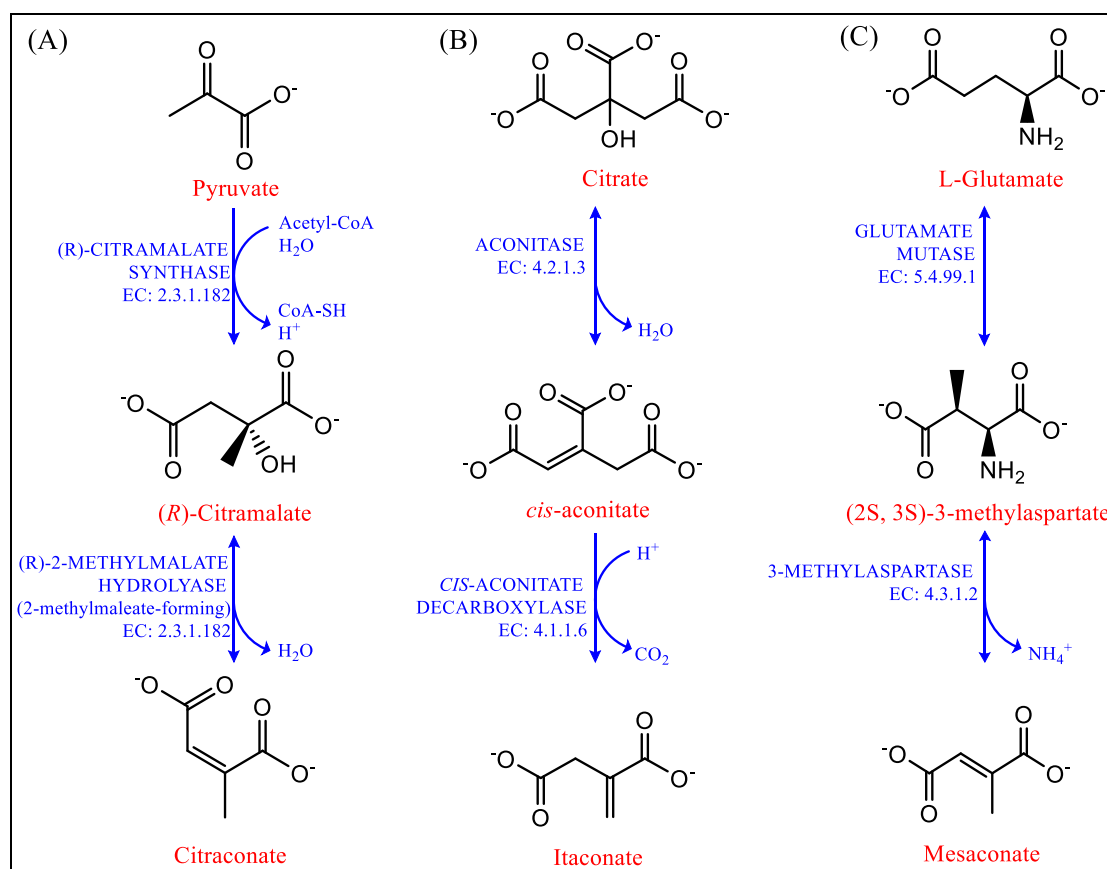


Figure 17: Biocatalytic routes to citraconic, itaconic and mesaconic acids from widely available biological precursors.

Thus, without further investigating each route, from both substrate availability and thermodynamics perspectives, the hydrolysis of methacrylyl-CoA and the decarboxylations of itaconic, mesaconic and citraconic acids as routes to methacrylic acid appear equally appealing. However, two factors swing the preference towards the hydrolysis of methacrylyl-CoA as the terminal metabolic step to methacrylic acid production. Firstly, a broad specificity thioesterase, *tesB* from *E. coli*, has already been identified as having activity on isobutyryl-CoA [16]. Whilst this is the precursor to methacrylyl-CoA, and should not be hydrolysed over methacrylyl-CoA, it demonstrates that thioesterases with activity on substrates with very similar carbon structures to methacrylyl-CoA do exist. As for the

promiscuous decarboxylation of itaconic, citraconic or mesaconic acids, it is possible to speculate that, being small, dicarboxylic acid without much hydrophobic character, an active site requiring substrate stabilisation through ionic or polar interactions would likely require precise polar or ionic functional groups in the correct places. Furthermore, the misplaced second carboxylate group of itaconic, citraconic and mesaconic acids compared to the natural substrate would destabilise hydrophobic interactions in an otherwise hydrophobic substrate binding site. By contrast, the coenzyme A moiety on methacrylyl-CoA would still bind effectively in the active site of an enzyme catalysing the hydrolysis of acyl-CoA molecules, partly compensating for sub-optimal binding of the methacrylyl-moiety. This is of course, speculative, however. With regards to the acyl alcohol transferase reaction, however, the enzyme would potentially have to show substrate promiscuity to two substrates, methacrylyl-CoA as well as a suitable alcohol, decreasing the chances of a suitable enzyme being found.

Thus, in this project, the hydrolysis of methacrylyl-CoA was chosen as a terminal metabolic step for the production of methacrylic acid, as the step is thermodynamically favourable and is driven to completion, has a precursor that is found in multiple organisms, including bacteria, and broad specificity thioesterases which accept short hydrophobic acyl-CoA substrates, are known to exist. These broad specificity thioesterases can be used as a back-up option should a more specific thioesterase not be found for the task.

With 2-ketoisovalerate acting as a common intermediate of both the valine degradation I pathway (Figure 16A) and the valine biosynthesis pathway (Figure 16B), methacrylic acid formed from the hydrolysis of methacrylyl-CoA is 6 steps from central metabolism. The pathway to methacrylic acid from central metabolism obtained by merging the relevant enzymes of the valine biosynthesis pathway, the valine degradation pathway and the acyl-CoA thioesterase for the hydrolysis of methacrylyl-CoA to methacrylic acid, is shown below (Figure 18).

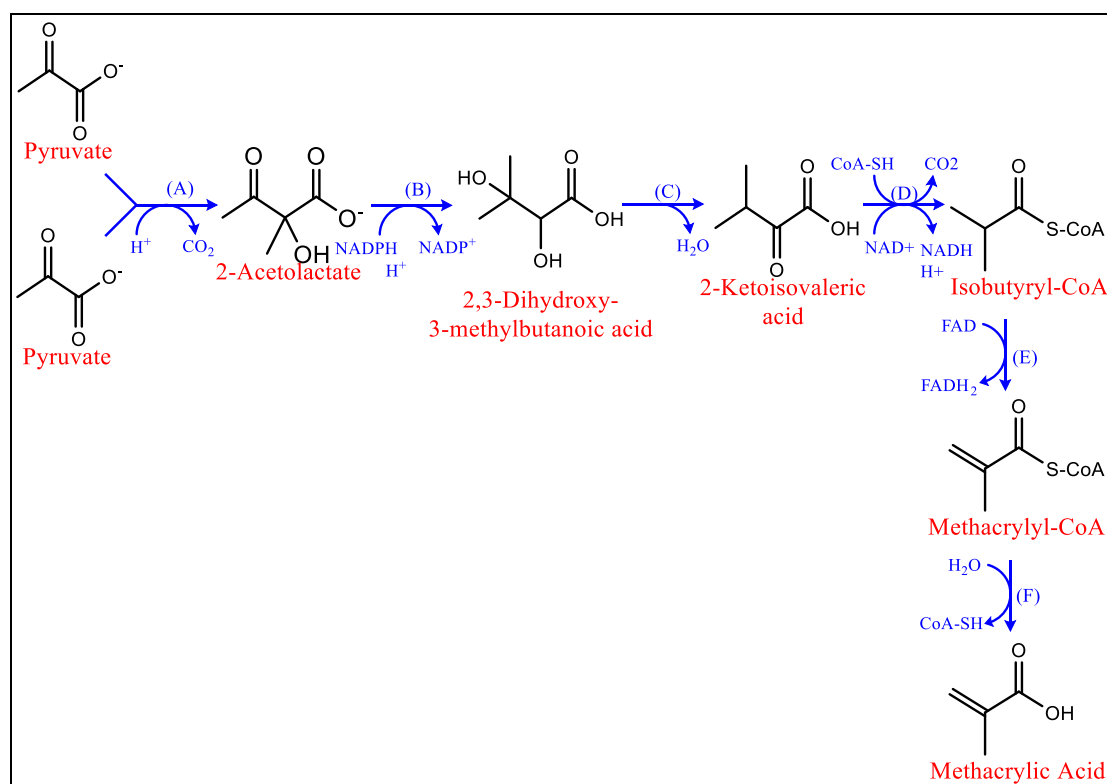


Figure 18: Route to methacrylic acid from pyruvate. A) Acetohydroxyacid synthase (AHAS). B) Acetohydroxy acid isomeroreductase (AHAIR). C) Di-hydroxy-acid dehydratase. D) Branched chain keto acid dehydrogenase. E) Acyl-CoA dehydrogenase or acyl-CoA oxidase. F) Acyl-CoA thioesterase.

E. coli, a tractable and industrially useful organism that has traditionally been the organism of choice for metabolic pathway engineering tasks [93], does not appear to have the VDI pathway, however. The merits of constructing a metabolic pathway to methacrylic acid in an organism that already contains the VDI pathway over one that does not, and *vice versa*, therefore deserve some discussion.

By using a host organism that already expresses the VDI pathway, the steps to methacrylyl-CoA already exist in the organism, and do not need to be introduced. Thus, only the candidate thioesterase needs to be cloned into the organism, with no additional introduction of heterologous genes. However, the VDI pathway has evolved to minimise the concentration of intracellular methacrylyl-CoA by hydrating

methacrylyl-CoA to 3-hydroxyisobutyryl-CoA and then hydrolysing this to 3-hydroxyisobutyric acid (3HIBA), driving the hydration to completion [81, 82]. Therefore, a heterologous enzyme for the hydrolysis of methacrylyl-CoA would have to compete with the remainder of the valine degradation pathway for methacrylyl-CoA, and as the heterologous enzyme would not be hydrolysing a natural substrate, it is unlikely to have a very low K_M for the substrate. Metabolic flux towards 3HIBA would likely be the preferred route unless either the hydration of 3-hydroxyisobutyryl-CoA or the hydrolysis of 3-hydroxyisobutyryl-CoA is disrupted to allow methacrylyl-CoA to accumulate.

In addition, intermediates of the VDI pathway could also act as branch points for other pathways in organisms that do contain the VDI pathway. For example, *Streptomyces avermitilis* uses isobutyryl-CoA to prime both avermectin biosynthesis and branched chain fatty acid synthesis for cell membrane fluidity [94, 95], which would compete for intermediates for the valine degradation pathway. On top of the disruption of the VDI pathway, these other pathways may also have to be disrupted.

Of course, it may also be necessary to disrupt genes in *E. coli*. Afterall, the *tesB* gene upregulated by Zhang *et al.* to hydrolyse isobutyryl-CoA for the production of isobutyryl-CoA is an *E. coli* gene [16]. However, if it is necessary to disrupt genes, it would be easier to do so in *E. coli*, as it is a genetically tractable organism [93], than doing so in another organism where less is known about the role of each gene, even if that requires building up the steps to methacrylyl-CoA in *E. coli*. During this project, therefore, *E. coli* was developed for the production of methacrylic acid.

3 Aims and objectives

The main aim of this project was to develop a whole cell biocatalytic route to methacrylic acid using a thioesterase to catalyse the hydrolysis of methacrylyl-CoA as the terminal metabolic step. Thus the objectives to the first part of this project were:

- Identify candidate thioesterases for the hydrolysis of methacrylyl-CoA.
- Screen the thioesterases for the hydrolysis of methacrylyl-CoA and selectivity for methacrylyl-CoA over isobutyryl-CoA.
- Obtain Michaelis Menten kinetics for any thioesterases that selectively hydrolyse methacrylyl-CoA over isobutyryl-CoA, to determine suitability for use in a whole cell biocatalyst.

In order to construct a whole cell biocatalyst, the host cell needed to be engineered to produce methacrylyl-CoA. The host organism for this project was *E. coli*, which does not naturally express the valine degradation I pathway and so the oxidation of isobutyryl-CoA needed to be engineered into the host. Two classes of enzyme exist for the oxidation of acyl-CoA molecules to their *trans*-2-enoyl-CoA analogues, the acyl-CoA dehydrogenases and the acyl-CoA oxidases. Thus, the specific objectives for the second part of the project were:

- Assess the differences, advantages and disadvantages of the two classes of enzyme and identify candidate enzymes for the successful oxidation of isobutyryl-CoA in *E. coli*, from both classes.
- Select an enzyme, or enzymes, based on the relative merits of the class of enzyme, as well as the proven ability to express well in *E. coli*.
- Clone and express the enzyme or enzymes, and confirm the oxidation of isobutyryl-CoA to methacrylyl-CoA *in vitro* after having overexpressed it in *E. coli*. Then confirm the conversion of isobutyryl-CoA to methacrylic acid and coenzyme A in an enzyme coupled reaction between the candidate enzyme and the thioesterase.

With enzymes for the oxidation of isobutyryl-CoA and the hydrolysis of methacrylyl-CoA, the formation of isobutyryl-CoA from available feedstocks needed to be engineered into *E. coli*. The biocatalytic production of isobutyric acid in *E. coli* has previously been demonstrated to high titres and it is known that certain strains of *P. putida* and *P. aeruginosa* grow on isobutyric acid as a sole carbon source [86, 96, 97]. *Peptostreptococcus elsdenii* and *Bacteroides ruminicola* can also produce valine from isobutyric acid [98]. As isobutyric acid is a fatty acid, and all fatty acid

degradation pathways, to my knowledge, occur through coenzyme A derivatives, it is very likely that isobutyric acid is activated to isobutyryl-CoA and is degraded through the valine degradation I pathway. Evidence that the growth of *P. putida* on isobutyrate induces non-constitutively expressed enzymes of the valine degradation pathway [96], supports this. The objectives for the third part of this project were therefore:

- Identify candidate enzymes for the activation of isobutyric acid to isobutyryl-CoA and base the selection on a known ability to express the enzyme/s in *E. coli*, as well as having favourable Michaelis Menten kinetic parameters.
- Co-express all pathway components required for the activation of isobutyric acid to isobutyryl-CoA, the oxidation of isobutyryl-CoA to methacrylyl-CoA and the hydrolysis of methacrylyl-CoA to methacrylic acid and coenzyme A in *E. coli*, and test the ability of the new strain to perform the biotransformation of isobutyric acid to methacrylic acid.

For the final part of the project, it was desired that methacrylic acid be produced directly from the organism's feedstock, such that methacrylic acid could be produced by direct fermentation. This required linking the existing enzymes for the oxidation of isobutyryl-CoA and the hydrolysis of methacrylyl-CoA to the organism's metabolism. Two routes were envisaged for this. These were to either clone a branched chain keto acid dehydrogenase for the direct oxidative decarboxylation of 2-ketoisovalerate to isobutyryl-CoA, or to copy Zhang *et al.*'s route to isobutyric acid and coupling the isobutyrate biotransformation pathway to it. Thus, the aims of the final part of the project were:

- Assess the suitability of cloning a branched chain keto acid dehydrogenase into the organism for the oxidative decarboxylation of 2-ketoisovaleric acid to isobutyryl-CoA relative to using a 2-ketoisovalerate decarboxylase and aldehyde dehydrogenase, as used by Zhang *et al.* for the production of 90 g.L⁻¹ isobutyrate (Figure 12B), and coupling this to the biotransformation pathway constructed for the previous objectives.
- Co-express the genes required for the production of isobutyryl-CoA from 2-ketoisovalerate, along with the genes required for the oxidation of isobutyryl-CoA to methacrylyl-CoA, and for the hydrolysis of methacrylyl-CoA to methacrylic acid.
- Assess the effect of upregulating the production of 2-ketoisovaleric acid in the production strain, on the production of methacrylic acid.

4 Materials and Methods

4.1 Materials

4.1.1 Electrophoresis reagents and buffers

Tris (Fisher); glycine (Fisher); 10X Tris-Glycine-SDS (Bio-Rad); bromophenol blue (Sigma-Aldrich); β -mercaptoethanol (Sigma-Aldrich).

4.1.2 Antibiotics, cofactors, vitamins nucleotides and substrates

Carbenicillin (Melford); chloramphenicol (Sigma-Aldrich); isopropyl- β -D-thiogalactopyranoside (IPTG) (Melford); coenzyme A trilithium salt (CoA-SH.3Li) (Sigma-Aldrich); isobutyryl-coenzyme A lithium salt (isobutyryl-CoA.xLi) (Sigma-Aldrich); flavin adenine dinucleotide disodium salt hydrate (Sigma-Aldrich); (-)-riboflavin (Sigma-Aldrich); thiamine (Sigma-Aldrich); deoxyribosenucleic acid (dNTP) mix containing deoxyguanosine triphosphate (dGTP) (10mM); deoxycytidine triphosphate (dCTP) (10mM); deoxyadenosine triphosphate (dATP) (10mM) and thymidine triphosphate (dTTP) (10mM) (Fermentas).

4.1.3 Molecular biology enzymes

*Nde*I FD (Fermentas), *Not*I FD (Fermentas), *Xho*I FD (Fermentas), *Eco*RI FD (Fermentas), *Hind*III FD (Fermentas), *Nhe*I FD (Fermentas), *Xba*I FD (Fermentas), Antarctic Phosphatase (New England Biolabs), T4 DNA ligase (Fermentas), JumpStart REDTaq DNA polymerase (Sigma-Aldrich), KOD DNA polymerase (Novagen), Phusion Hot Start II DNA polymerase (Fermentas), Benzonase[®] nuclease (Sigma-Aldrich).

4.1.4 Molecular biology kits, protein extraction kits

CloneJET[™] PCR cloning kit (Fermentas); Plasmid Midi Kit (QIAGEN); QIAprep Spin Miniprep Kit (QIAGEN); QIAquick Gel Extraction Kit (QIAGEN); Plasmid Maxi Kit (QIAGEN); GenElute[™] Plasmid Miniprep Kit (Sigma-Aldrich); GenElute[™] HP Plasmid Midiprep Kit (Sigma-Aldrich); GenElute[™] Gel Extraction Kit (Sigma-Aldrich); BugBuster[®] protein extraction kit (Novagen); cOmpleteMiniEDTA-free protease inhibitor (Roche); ChargeSwitch[®] gDNA Mini Bacteria Kit (Invitrogen[™]).

4.1.5 Salts, acids, bases, buffers, trace elements and gases.

Ammonium chloride (NH₄Cl) (Sigma-Aldrich), magnesium sulphate heptahydrate (MgSO₄.7H₂O) (Sigma-Aldrich), sodium chloride (NaCl) (Fisher), potassium dihydrogen orthophosphate (KH₂PO₄) (Sigma-Aldrich), di-potassium hydrogen orthophosphate (K₂HPO₄) (Sigma-Aldrich), sodium dihydrogen orthophosphate (NaH₂PO₄) (Sigma-Aldrich), di-sodium hydrogen orthophosphate (Na₂HPO₄)

(Fisons), potassium hydroxide (KOH) (Sigma-Aldrich), sodium hydroxide (NaOH) (Sigma-Aldrich), hydrochloric acid (HCl) (Fisher), ethylene diamine tetra-acetic acid disodium salt dehydrate (EDTA-disodium salt dihydrate) (Sigma-Aldrich), zinc sulphate heptahydrate ($\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$) (Acros), calcium chloride dihydrate ($\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$) (Fluka), manganese chloride tetrahydrate ($\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$) (Sigma-Aldrich), iron sulphate heptahydrate ($\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$) (Rectapur), ammonium molybdate tetrahydrate ($(\text{NH}_4)_6\text{Mo}_7\text{O}_{24} \cdot 4\text{H}_2\text{O}$) (BDH), copper sulphate pentahydrate ($\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$) (Sigma-Aldrich), cobalt (II) chloride hexahydrate ($\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$) (ACROS), magnesium sulphate (MgSO_4) (Fisher), aluminium chloride hexahydrate ($\text{AlCl}_3 \cdot 6\text{H}_2\text{O}$) (Sigma-Aldrich), copper chloride dihydrate ($\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$) (Sigma-Aldrich), sodium molybdate dihydrate ($\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$) (Sigma-Aldrich), manganese sulphate monohydrate ($\text{MnSO}_4 \cdot \text{H}_2\text{O}$) (Sigma-Aldrich), boric acid (H_3BO_3) (ACROS), di-sodium hydrogen phosphate dodecahydrate ($\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$) (Sigma-Aldrich), di-sodium hydrogen phosphate heptahydrate ($\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$) (Sigma-Aldrich), sodium dihydrogen phosphate dihydrate ($\text{NaH}_2\text{PO}_4 \cdot 2\text{H}_2\text{O}$) (Sigma-Aldrich), calcium carbonate (CaCO_3) (ACROS), magnesium chloride (MgCl_2) (BDH), isobutyric acid (Sigma-Aldrich), methacrylic acid (Sigma-Aldrich), 2-ketoisovaleric acid sodium salt (Sigma-Aldrich), N_2 (BOC), glacial acetic acid (Sigma-Aldrich).

4.1.6 Solvents

Ethanol (Fisher), isopropanol (Sigma-Aldrich), acetonitrile (VWR), diethyl ether (Sigma-Aldrich).

4.1.7 Growth media and cryostock components

Agar (Melford), D-glucose (Fisher), glycerol (ACROS), Luria Bertani (LB) granules (Melford), yeast extract (Fluka), meat extract (Fluka), peptone from casein (BD BBL™ Trypticase™), peptone from soymeal (BD BBL™ Phytone™), peptone (Formedium, PEP03).

4.1.8 Tables of primers, bacterial strains, synthetic genes and synthetic operons

4.1.8.1 Table of primers

Primer ID	Primer Name	Primer Sequence (5' to 3')
PID 1	T7_F	TAATACGACTCACTATAGGG
PID 2	T7_R	GCTAGTTATTGCTCAGCGG
PID 3	4HBT_CtHis_F	TATACATATGCACCGTACCTCTAACGGTTCTCACGC
PID 4	4HBT_CtHis_R	CTCGAGTCCGTCACGACGCGGACG
PID 5	OE_4HBT_F	ACATATGCACCGTACCTCTAACGGTTC
PID 6	OE_4HBT_R	CTGCCATATCTATATCTCCTGTTAGTCACGACGCGGACG
PID 7	OE_ACX4_F	GTGACTAACAGGAGATATAGATATGGCAGTTCTGAGCAGC
PID 8	OE_ACX4_R	CTCGAGATATTATAGCTAGCTTACAGACGGCTACGGGTTG
PID 9	pHAAC_S1	CAGAAATTGCACGTGTTGATG
PID 10	pHAAC_S2	GAAGTTACCGGTATTGCAAGC
PID 11	bsBCKD_F	GCTAGCTAGACAGACAGGAGTGAGTCACC
PID 12	bsBCKD_R	CTCGAGTTAGTAAACAGATGTCTTCTCGTCAATC
PID 13	ppBCKD_F	TGCATCGCTAGCAGGAGAAATTAACATATGAACGAGTACGCCCCCTGCGTTTGC
PID 14	ppBCKD_R	TGCATCCTCGAGTCAGATATGCAAGGCGTGCCCCAG
PID 15	bsBCKD_S1	GTTACCGGTCGTGAAGTTAC
PID 16	bsBCKD_S2	GCCGACTGAAGATCTAGAGATTTT
PID 17	bsBCKD_S3	CTTGAATGAGTACAAACCGACATC
PID 18	bsBCKD_S4	CGATTTCTTACCGCCTTACACC
PID 19	bsBCKD_S5	GCCATCAACACCATATGACG
PID 20	bsBCKD_S6	GTGGGAGAAGAAGGCCAAAC
PID 21	pSTV28_Phu_F	GGAAACAGCTATGACCATGATTACGAATTC
PID 22	pSTV28_Phu_R	CGATTAAGTTGGGTAACGCCAGG
PID 23	pET20b_Phu_F	CTCACTATAGGGAGACCACAACGGTTTC
PID 24	pET20b_Phu_R	GGCTTTGTTAGCAGCCGGATC
PID 25	pSTV_MOD_F	AATTAATCTCGAGTATAATAAGCTTGCTGGCGTAATAGCGAAGAGG
PID 26	pSTV_MOD_R	ATACATATGTATATCTCCTTCTTAAGAATTCGTGTGAAATTGTTATCCGCTCAC AATTCC

Table 2: Table of primers used throughout study.

4.1.8.2 Table of bacterial strains used, and their genotypes

Full Organism Name	Organism Genotype	Reference for Genotype	Organism purchase from
<i>Escherichia coli</i> JM107	endA1 glnV44 thi-1 relA1 gyrA96 Δ(lac-proAB) [F' traD36 proAB ⁺ lacI ^q lacZΔM15] hsdR17(R _K ⁻ m _K ⁺) λ ⁻	[99]	Fermentas and also AddGene
<i>Escherichia coli</i> BL21 (DE3) pLysS (purchased as competent cells)	F ⁻ ompT gal dcm lon hsdS _B (r _B ⁻ m _B ⁻) λ(DE3) pLysS(cm ^R)	[99]	Merck
<i>Escherichia coli</i> BL21 (DE3) (purchased as competent cells)	F ⁻ ompT gal dcm lon hsdS _B (r _B ⁻ m _B ⁻) λ(DE3)	[99]	
<i>Bacillus subtilis</i> 168 / DSM23778 / ATCC23857	trpC swrA sfp gudB ⁻	[100]	Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ)
<i>Pseudomonas putida</i> KT2440 / DSM6125 / ATCC47054	(not available)	N/A	

Table 3: Table of bacterial strains used, and their genotypes.

4.1.8.3 Table of synthesised polynucleotides

Polynucleotide ID	Polynucleotide name (in brackets: genes encoded on operons)	Sequence length (in base pairs)	Sequence on Appendix page
Polynucleotide sequence 1	4HBT	475	III
Polynucleotide sequence 2	FcoT	571	III
Polynucleotide sequence 3	YciA	484	III
Polynucleotide sequence 4	RifR	799	III
Polynucleotide sequence 5	ACX4	1320	III
Polynucleotide sequence 6	AcsA	1663	IV
Polynucleotide sequence 10	ppBCKD (contains: pp_bkdA1, pp_bkdA2, pp_bkdB, pp_lpdV)	4996	IX
Polynucleotide sequence 15	ilvD-alsS-ilvC (contains: ec_ilvD, bs_alsS, ec_ilvC)	5108	XIII

Table 4: Table of synthesised polynucleotides. All polynucleotides except that encoding ACX4 were synthesised by Biomatik. Polynucleotide encoding ACX4 was synthesised by Life Technologies.

4.2 Methods

4.2.1 Preparation of frequently used buffers

4.2.1.1 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer (50 mM)

Stocks of 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) were prepared by dissolving HEPES (Melford) to a final concentration of 50mM in distilled water, adjusting to pH 7.5 with KOH (10 M). HEPES buffer was filter sterilised directly into 50 ml sterile falcon tubes that were wrapped with foil to prevent formation of peroxides by the photolabile HEPES, by aseptic technique, and stored for up to 6 months at +4 °C.

4.2.1.2 Potassium phosphate buffer (100 mM)

Stocks of potassium phosphate buffer were prepared by titrating potassium dihydrogen orthophosphate (KH_2PO_4) (100 mM) into di-potassium hydrogen orthophosphate (K_2HPO_4) until a pH of 7.5 was reached. Potassium phosphate buffer was filter sterilised directly into 50 ml falcon tubes by aseptic technique, and stored for up to 6 months at +4 °C.

4.3 Preparation of antibiotic, IPTG, glucose, cofactor and vitamin stocks

4.3.1 Glucose stock solutions

Glucose stock solutions of 20% (m/v) and 12.5% (m/v) were prepared and filter sterilised directly into sterile 50 ml falcon tubes under aseptic conditions. Glucose stock solution of 15% (m/v) was prepared in a Duran bottle and sterilised by autoclave.

4.3.2 Carbenicillin stock (100 mg.ml⁻¹)

Stocks of carbenicillin (Mr: 378.401) were prepared to a concentration of 100 mg.ml⁻¹ by dissolving 5 g of carbenicillin disodium salt (Melford, Mr: 422.36 g.mol⁻¹) to a final volume of 44.8 ml distilled water and filter sterilising with a 0.22 µm filter directly into autoclaved microcentrifuge tubes in a laminar flow hood. Stocks of carbenicillin were stored at -20 °C.

4.3.3 Chloramphenicol stock (34 mg.ml⁻¹)

Stocks of chloramphenicol were prepared to a concentration of 34 mg.ml⁻¹ by dissolving 1 g of chloramphenicol (Sigma-Aldrich), into a final volume of 29.4 ml ethanol, and filter sterilised with a 0.22 µm filter directly into autoclaved microcentrifuged tubes in a laminar flow hood. Stocks of chloramphenicol were kept at -20 °C for long term storage, and on ice for short term storage whilst preparing media.

4.3.4 Isopropyl- β -D-thiogalactopyranoside (IPTG) stock (0.5 M)

Stocks of isopropyl- β -D-thiogalactopyranoside (IPTG) (Mr: 238.3 g.mol⁻¹) were prepared to a concentration of 0.5 M by dissolving 5 g of IPTG (Melford) into a final volume of 42 ml distilled water. Stocks were filter sterilised with a 0.22 μ M filter directly into autoclaved microcentrifuge tubes in a laminar flow hood. Stocks of IPTG were stored at -20 °C.

4.3.5 Flavin adenine dinucleotide (FAD) stock (1 M)

Stocks of flavin adenine dinucleotide (FAD) were prepared to a concentration of 1 M by dissolving 1.7 g FAD disodium salt hydrate into 2 ml of 50 mM HEPES buffer, pH 7.5. The FAD in HEPES buffer solution was aliquoted into thin walled PCR tubes, flash frozen in liquid nitrogen and stored at -80 °C.

4.3.6 Growth media

4.3.6.1 Luria-Bertani (LB) medium

Luria Bertani high salt medium (Melford) (25 g.L⁻¹) was used wherever LB medium is referred to, whereby 25 g of LB high salt medium in 1 L consists of peptone from casein digest (10 g.L⁻¹), yeast extract (5 g.L⁻¹) and sodium chloride (10 g.L⁻¹). LB medium was often supplemented with carbenicillin (50 μ g.ml⁻¹), chloramphenicol (34 μ g.ml⁻¹), riboflavin (1 mg.L⁻¹) and/or glucose (1% w/v). LB medium was sterilised by autoclave and allowed to cool to room temperature prior to the addition of supplements. LB supplemented with riboflavin was prepared by first preparing both LB medium as well as LB medium containing 5X riboflavin (5 mg.L⁻¹). The LB + 5X Riboflavin was filter sterilised, and LB + 1X riboflavin (100 ml) was then prepared by adding 20 ml LB + 5X riboflavin to 80 ml LB. Shake flasks for cultures supplemented with riboflavin were wrapped in foil to prevent damage to the photolabile riboflavin.

4.3.7 MSX

MSX medium was prepared by combining MSA medium (760 ml.L⁻¹), MSB medium (200 ml.L⁻¹) and a 12.5% (w/v) glucose stock solution (40 ml.L⁻¹) at room temperature. MSB medium was composed of NH₄Cl (15 g.L⁻¹) and MgSO₄.7H₂O (2.0 g.L⁻¹) and was sterilised by autoclave. MSA medium was composed of KH₂PO₄ (7.89 g.L⁻¹) and vishniac trace elements (2.63 ml.L⁻¹), and KOH solution (10 M) was added until the pH reached 7.0 before MSA was sterilised by autoclave. Vishniac trace elements were prepared as previously described [101] but using only 3.9 g.L⁻¹ of ZnSO₄.7H₂O. MSX medium was sometimes supplemented with riboflavin (1 mg.L⁻¹) and this was achieved by dissolving riboflavin into MSB medium first (5 mg.L⁻¹). This solution was then filter sterilised and the MSB + 5X

Riboflavin solution was then mixed with MSA and glucose in the same volumetric as for MSX medium alone, to form MSX supplemented with riboflavin. Both MSX and MSX supplemented with riboflavin were always supplemented with carbenicillin ($50 \mu\text{g}.\text{ml}^{-1}$) and sometimes also with chloramphenicol ($34 \mu\text{g}.\text{ml}^{-1}$). Shake flasks for cultures supplemented with riboflavin were wrapped in foil to prevent damage to the photolabile riboflavin.

4.3.8 SM medium

SM medium (1 L) was prepared by first adding SM salts (200 ml) from 5X sterile stock solution, glucose (100 ml) from 15% (w/v) sterile stock solution, and calcium chloride ($75 \mu\text{l}$) from 1 M sterile stock solution, to distilled water (498 ml) that was sterilised by autoclave. The components were mixed by swirling prior to the addition of magnesium sulfate (2 ml) from sterile stock solution and the components were mixed once more by swirling. SM trace elements ($100 \mu\text{l}$) were then added to the formulation from sterile stock solution, followed by thiamine ($200 \mu\text{l}$) from sterile stock solution and finally yeast extract (200 ml) from 2.5% (w/v) sterile stock solution.

Sterile stock solution of 5X SM salts was composed of $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$ ($18.9 \text{ g}.\text{L}^{-1}$), KH_2PO_4 ($3.75 \text{ g}.\text{L}^{-1}$), NaCl ($2.5 \text{ g}.\text{L}^{-1}$) and NH_4Cl ($35 \text{ g}.\text{L}^{-1}$), dissolved in distilled water, adjusted to pH 7.0 and sterilised by autoclave.

Individual sterile stock solutions of glucose ($150 \text{ g}.\text{L}^{-1}$), yeast extract ($25 \text{ g}.\text{L}^{-1}$), $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ ($246.5 \text{ g}.\text{L}^{-1}$), $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ ($147 \text{ g}.\text{L}^{-1}$) and thiamine ($10 \text{ g}.\text{L}^{-1}$) were prepared separately. Thiamine was filter sterilised whilst glucose, yeast extract, magnesium sulphate and calcium chloride stock solutions were each sterilised by autoclave.

SM trace elements stock solution was composed of $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ ($80 \text{ g}.\text{L}^{-1}$), $\text{AlCl}_3 \cdot 6\text{H}_2\text{O}$ ($10 \text{ g}.\text{L}^{-1}$), $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ ($2 \text{ g}.\text{L}^{-1}$), $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ ($1 \text{ g}.\text{L}^{-1}$), $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ ($2 \text{ g}.\text{L}^{-1}$), $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ ($10 \text{ g}.\text{L}^{-1}$), CoCl_2 ($4 \text{ g}.\text{L}^{-1}$) and H_3BO_3 ($0.5 \text{ g}.\text{L}^{-1}$), made up in 3 M HCl, and filter sterilised. Preparation of SM trace elements was kindly performed by Samuel Wilson, the laboratory technician.

SM medium was always supplemented with riboflavin ($1 \text{ mg}.\text{L}^{-1}$) and carbenicillin ($50 \mu\text{g}.\text{ml}^{-1}$), and often also supplemented with chloramphenicol ($34 \mu\text{g}.\text{ml}^{-1}$). Riboflavin was supplemented to the medium by preparing a 5X solution in the 5X yeast extract and filter sterilising, such that the yeast extract stock solution contained riboflavin ($5 \text{ mg}.\text{L}^{-1}$). Shake flasks were wrapped in foil to prevent damage to the photolabile riboflavin.

4.3.9 Modified M9 medium

Modified M9 medium with calcium carbonate buffering was described by Zhang *et al.* for the biosynthesis of isobutyric acid [16] and was further modified to contain riboflavin. For each litre of modified M9 medium with riboflavin required, 5X M9 salts (200 ml) were added to sterile distilled water (397.9 ml), followed by yeast extract from 2.5% sterile stock solution (200 ml) containing 5X riboflavin. Glucose was then added from 20% sterile stock solution (200 ml) to give a final concentration of 40 g.L⁻¹. To this, CaCl₂ was added from 1 M sterile stock solution (100 µl) and the medium mixed by swirling, before finally, MgSO₄ was added from 1 M sterile stock solution (2 ml). Sterile modified M9 medium (100 ml) was added aseptically to autoclaved shake flasks (500 ml), containing the calcium carbonate powder (10 g), that had been autoclaved *in situ* and allowed to cool to room temperature prior to the addition of the modified M9 medium. Shake flasks were wrapped in foil to prevent damage to the photolabile riboflavin.

Sterile stock solution of M9 salts (5X) contained Na₂HPO₄·7H₂O (64 g.L⁻¹), KH₂PO₄ (15 g.L⁻¹), NaCl (2.5 g.L⁻¹) and NH₄Cl (5 g.L⁻¹) and was sterilised by autoclave. Sterile stock solution of yeast extract (2.5%) with 5X riboflavin was prepared by first autoclaving yeast extract (25 g.L⁻¹) and allowing the yeast extract to cool to room temperature before dissolving riboflavin (5 mg.L⁻¹) into it as required, and filter sterilising. Excess 2.5% yeast extract containing 5X riboflavin was stored at -20 °C and thawed overnight at 4 °C prior to use. Sterile stock solutions of MgSO₄ (1 M) and CaCl₂ (1 M) contained MgSO₄ (120.36 g.L⁻¹) and CaCl₂·2H₂O (114.01 g.L⁻¹), respectively, and were autoclaved separately, whilst sterile stock solution of glucose (20% (m/v)) was filter sterilised using a 0.22 µm sterile disposable syringe filter.

4.3.10 DSMZ medium 1

DSMZ medium 1 was composed of peptone (5 g.L⁻¹) and meat extract (3 g.L⁻¹). The pH was adjusted to pH 7.0 and the growth medium was sterilised by autoclave.

4.3.11 DSMZ medium 220

DSMZ medium 220 was composed of peptone from casein (15 g.L⁻¹), peptone from soymeal (5 g.L⁻¹) and NaCl (5 g.L⁻¹). The pH was adjusted to pH 7.3 and the growth medium was sterilised by autoclave.

4.3.12 Luria Bertani agar plates

Luria Bertani agar plates were prepared using Luria Bertani high salt media (Melford) (25 g.L⁻¹) and agar (20 g.L⁻¹). The LB and agar mixture was sterilised by

autoclave and allowed to cool for 1 h in a 50 °C water bath prior to pouring. The LB agar plates were often supplemented with carbenicillin (50 µg.ml⁻¹), chloramphenicol (34 µg.ml⁻¹) and/or glucose (1% w/v) from sterile stock solutions. Glucose was added from the 20% (w/v) sterile stock solution.

4.4 Cell maintenance, propagation and preservation.

4.4.1 Micro-organisms

Escherichia coli JM107, *Escherichia coli* BL21 (DE3) pLysS, *Escherichia coli* BL21 (DE3), *Bacillus subtilis* 168 and *Pseudomonas putida* KT2440 were used in this project. From here on, *Escherichia coli* will be referred to as *E. coli*, *Bacillus subtilis* as *B. subtilis*, and *Pseudomonas putida* as *P. putida*, as is common notation.

LB-glycerol cryostocks of *E. coli* JM107 that was purchased from Fermentas were prepared by a previous PhD student in the laboratory, Pawel Mordaka. *E. coli* JM107 that was purchased from AddGene was supplied on an LB agar slant and initial cryostocks were prepared by Juliana Lebeau. Cells were collected from the slant using a sterile 10 µl disposable inoculating loop and aseptically streaked onto a fresh LB agar plate. The plate was incubated at 37 °C for 16 h and LB-glycerol cryostocks were prepared from a single colony of freshly grown *E. coli* JM107.

E. coli BL21 (DE3) and *E. coli* BL21 (DE3) pLysS competent cells were purchased from Merck in 200 µl aliquots. The cells were thawed on ice and redistributed into 20 µl aliquots in sterile 1.5 ml micro-centrifuge tubes under aseptic conditions for subsequent transformation steps, and these were stored at -80 °C.

P. putida KT2440 (DSM6125) and *B. subtilis* 168 (DSM23778) were purchased from Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ). Freeze dried cells were reconstituted in their respective media, DSMZ medium 1 for *P. putida* KT2440 and DSMZ medium 220 for *B. subtilis* 128. Reconstituted *P. putida* KT2440 and *B. subtilis* 168 were both incubated at 28 °C for 3 h before being plated onto LB agar plates and incubated at 30 °C for 16 h. Single colonies of each were used to inoculate 10 ml of LB media which was incubated at 30 °C until an OD₆₀₀ of 0.8 was reached. Cryostocks of the culture were prepared in 8% glycerol.

4.4.2 Preparation of cryostocks

Cryostocks were prepared by inoculating LB medium (10 ml) supplemented with either glucose, carbenicillin, chloramphenicol, or any combination thereof and only as required, from a single microbial colony. *E. coli* cultures were incubated at 37 °C with shaking at 250 rpm, whilst *P. putida* and *B. subtilis* cultures were incubated at 30 °C with shaking at 250 rpm. All cultures were incubated until an OD₆₀₀ of

between 0.6 and 1.0, and aliquoted (900 µl) into cryo-vials containing 100 µl of 80% (w/v) glycerol. Cells were then vortexed with the glycerol stock in the cryovial to create an 8% (w/v) glycerol stock, and then stored at -80 °C. Due to the viscosity of glycerol making accurate volume measurements very difficult, 80% (v/v) glycerol stock (100.8% w/v) was prepared by pipetting approximately 30 ml of glycerol into a duran bottle that was on a tared balance. From the known weight of the aliquoted glycerol, and assuming that the volume deviation from the target final volume after dissolving glycerol and water into each other was negligible, the required volume of water to add to the glycerol to prepare 80% (v/v) glycerol was calculated from Equation 1.

$$V_{WATER} = \frac{M_{GLYCEROL}}{5.04}$$

Equation 1: Equation to calculate the volume of water to add to a known mass of glycerol to prepare 80% (v/v) glycerol. V_{WATER} : Volume of water to add to the known mass of glycerol. $M_{GLYCEROL}$: Known mass of glycerol. For derivation, see section 8.4.1.

4.4.3 Cell maintenance and culture inoculations.

4.4.3.1 Plating micro-organisms onto an LB-agar plate.

Cells were plated onto LB agar plates from transformations by pipetting 50 µl of transformation mixture onto the centre of pre-warmed plates using an L-shaped sterile disposable cell spreader. Cells from viable liquid cultures or from cryostocks were plated by pipetting 5 µl of culture of thawed cryostock onto the edge of an LB-agar plate, and streaking along four segments of the circular plate using a 10 µl sterile disposable inoculation loop for each segment. *E. coli* strains were incubated at 37 °C whilst *P. putida* and *B. subtilis* strains were incubated at 30 °C. All strains were incubated for 16 h. All plating was performed aseptically, nearby a roaring blue Bunsen flame, or in a laminar flow hood. Plates containing colonies of *E. coli* JM107 transformed with pET20b(+) derivatives were stored at 4 °C for a maximum of 5 days, whilst plates containing colonies of *E. coli* JM107 strains transformed with pSTV28 derivatives, and *E. coli* BL21 (DE3) or *E. coli* BL21 (DE3) pLysS strains transformed with pET20b(+), pSTV28 or both pET20b(+) and pSTV28 derivatives were stored at 4 °C for a maximum of 24 h.

4.5 Bacterial transformations

4.5.1 Transformation of *E. coli* JM107 cells

E. coli JM107 cells were made competent using the Fermentas TransformAid bacterial transformation kit. *E. coli* JM107 cells were first streaked onto LB agar plates from cryostock and incubated at 37 °C for 16 h. Plates were kept at 4 °C for no more than 7 h. A single colony was used to inoculate C-medium (2 ml) and this

was then cultured for 16 h at 37 °C with shaking at 250 rpm. The *E. coli* JM107 cells grown up in C-media were used for up to one week for the remainder of this procedure. Overnight culture (150 µl) of *E. coli* JM107 cells in C-medium was used to inoculate fresh C-medium (1.5 ml) and this fresh culture was incubated for 20 min, and the remainder of the procedure was performed as subsequently described in the manual.

4.5.2 Transformation of *E. coli* BL21 (DE3) and *E. coli* BL21 (DE3) pLysS cells

Competent *E. coli* BL21 (DE3) and *E. coli* BL21 (DE3) pLysS cells were transformed with derivatives of the pET20b(+) plasmid by a heat shock method. Aliquots of competent cells were thawed on ice and plasmid DNA (1 ng) was added aseptically to the competent cells. The plasmid was mixed with the cells by stirring with the pipette tip, as opposed to pipetting, so as not to shear the competent cells. Cells were then incubated with the plasmid on ice for 5 min and were then heat shocked by submerging the base of the tubes in a water bath set to 42 °C for 30 s before returning the cells to ice for 2 min. *E. coli* BL21 (DE3) cells transformed with pET20b(+) derivatives were then plated on pre-warmed LB agar plates supplemented with glucose and carbenicillin, whilst *E. coli* BL21 (DE3) pLysS cells transformed with pET20b(+) derivatives were plated on pre-warmed LB plates supplemented with glucose, carbenicillin and chloramphenicol.

4.5.3 Preparation of competent *E. coli* BL21 (DE3) pHAPP and *E. coli* BL21 (DE3) pHABS and transformation with pKIV.

E. coli BL21 (DE3) pLysS pHAPP and *E. coli* BL21 (DE3) pLysS pHABS cells were streaked from cryostock onto fresh LB plates supplemented with carbenicillin and glucose and incubated overnight. A single colony from each plate was used to inoculate LB precultures (10 ml), supplemented with carbenicillin and glucose in a 50 ml Falcon tube. The preculture was incubated overnight (37 °C, 200 rpm). A new sterile LB liquid culture (100 ml) in a 250 ml conical flask, also supplemented with carbenicillin and glucose, was prepared for each preculture. These were inoculated with 1 ml of preculture, and incubated at 37 °C with shaking at 250 rpm until an OD₆₀₀ of 0.4 was reached. For each culture, cells were transferred to two 50 ml centrifuge tubes, kept on ice for 20 min and then centrifuged at 3000 x g, 4 °C for 10 min. The supernatant was removed and cells were resuspended on ice, in ice-cold (6 ml) 0.1 M CaCl₂ and incubated for 30 min on ice. The resuspended cells were then centrifuged again at 3000 x g, 4°C for a further 10 min. The supernatant was removed and cells were resuspended once more in ice-cold solution (1.6 ml) consisting of CaCl₂ (0.1 M, 1.6 ml). Aliquots of competent cells

(140 µl) were transferred to ice cold 1.5 ml microcentrifuge tubes. The pKIV plasmid (1 µl) was transferred to an aliquot of competent cells for each of *E. coli* BL21 (DE3) pHAPP and *E. coli* BL21 (DE3) pHABS, and was mixed into the cells by gently stirring the pipette tip so as to avoid shearing the competent cells. Cells were then incubated on ice for 5 min before being heat-shocked by submerging the base of the tubes in a water bath set to 42 °C for 30 s. Cells were then returned to ice and incubated for a further 2 min before being plated onto pre-warmed LB agar plates supplemented with glucose, chloramphenicol and carbenicillin.

4.6 Analysis of protein extracts and purified proteins by Sodium Dodecyl-Sulphate Polyacrylamide Gel Electrophoresis.

Samples (1.5 ml) were taken from cell cultures and cells were pelleted by centrifugation at 5000 x g for 10 min. Cell pellets were resuspended in cell lysis buffer, using 100 µl for each OD₆₀₀ unit the cell culture was at when the samples were taken. Cell lysis buffer contained potassium phosphate buffer (100 mM, pH 7.5), BugBuster cell lysis detergent (Merck-Millipore) (1X), Benzonase nuclease (Sigma Aldrich) (0.01%, v/v) and a dissolved cOmplete™, mini, EDTA-free protease inhibitor cocktail tablet. Resuspended cells were incubated for 20 min with shaking at 250 rpm and centrifuged at 18,000 x g for 20 min at 4 °C. The insoluble fraction was re-suspended in potassium phosphate buffer (100 mM, pH 7.5) using the same volume as that of the cell lysis buffer that was used to resuspend the cell pellet. Soluble and insoluble fractions were mixed in a 1:1 ratio with 2X Laemmli sample buffer (Bio-Rad), containing β-mercaptoethanol (5%, v/v). Samples were boiled at 100 °C for 5 min and loaded onto a Bio-Rad AnyKD TGX precast gel. Electrophoresis was performed at 200 V in Tris/Glycine/SDS running buffer (Bio-Rad). Gels were washed by soaking them in double distilled water at room temperature with gentle agitation (50 rpm) for 5 min. The water was removed and the wash procedure was repeated a further four times. Gels were then stained by soaking overnight (16 h) in EZBlue™ Gel Staining Reagent (Sigma-Aldrich) at room temperature with gentle agitation for 16 h.

4.7 Molecular biology methods

4.7.1 Restriction digests of plasmids

To isolate the gene or polynucleotide inserts from plasmids, restriction digests of plasmids were performed using the Fermentas FastDigest® series of restriction endonucleases, according to the instructions provided. All restriction digest reactions were performed in a water bath at 37 °C for 4 h. Inserts were then purified by agarose gel electrophoresis and gel extraction (Section 4.7.3).

4.7.2 Preparation of linear expression vector

Plasmids were linearised in restriction digestion reactions (100 µl) containing plasmid DNA (100 ng.µl⁻¹), Fermentas FastDigest® Green buffer (1X) and Fermentas FastDigest® restriction endonucleases, as described above. Linearised vectors were purified by agarose gel electrophoresis and gel extraction without first heat-inactivating the restriction endonucleases. The 5' phosphates of the linearised vectors were then hydrolysed by Antarctic phosphatase (New England Biolabs) according to the instructions provided. The dephosphorylated vector was concentrated by ethanol precipitation, which required the addition of 3 M sodium acetate, pH 5.2, to the sample (volume added: 1/10th of the original sample volume) followed by the addition of absolute ethanol (volume added: 2X the new sample volume). The precipitation mixture was incubated overnight at -20 °C. The precipitated DNA was then centrifuged in an Eppendorf miniSpin F-45-12-11 rotor at 13400 rpm for 30 min. The supernatant was discarded and the cell pellet was washed with 70% (v/v) ethanol. The DNA pellet was then centrifuged in the same rotor and at the same speed as before for a further 10 min. The supernatant was removed and the DNA pellet was air dried in a fume hood for 20 min before being resuspended in nuclease free water (100 µl). The concentration of the linearised vector was determined by analytical agarose gel electrophoresis (section 4.7.6).

4.7.3 Agarose gel electrophoresis and gel extraction

Linear DNA fragments from plasmid digestion or PCR amplification were purified by agarose gel electrophoresis. The digests or PCR products were loaded onto an agarose-TAE gel submerged in Tris-Acetate-EDTA (TAE) buffer. Agarose-TAE gels were prepared by adding agarose (1 g) to TAE buffer (100 ml) and microwaving until the agarose was fully dissolved and the agarose-TAE mix was hot and fluid. The agarose-TAE was allowed to cool for 5-15 min, or until it stopped steaming, and ethidium bromide (7 µl) was added before the agarose-TAE started to set. The ethidium bromide was mixed into the molten gel by swirling, and the gel was cast in a gel-mould with a well-comb in place. The gel was allowed to set for 1 h, before being submerged in TAE buffer in the electrophoresis tank, and the well-comb was removed. DNA samples were mixed with 6X loading dye in a 5:1 volumetric ratio, and loaded into the wells. A potential difference of 7 V per centimetre of gel length was applied across the gel to resolve the polynucleotides. DNA was visualised on the gel with a transilluminator, and a gel slice containing the polynucleotide of interest was excised with a scalpel. The polynucleotide was purified from the gel slice with the QIAquick® Gel extraction kit. The concentration of the polynucleotide was determined by analytical agarose gel electrophoresis.

TAE buffer (1X) was prepared by adding 20 ml of a 50X TAE buffer stock to 980 ml water, and mixing. TAE (50X) was composed of tris buffer (242 g.L^{-1}), EDTA disodium salt dihydrate (20.6 g.L^{-1}) and glacial acetic (57.1 ml.L^{-1}) in distilled water. The components were added to distilled water in the order displayed, and only after the preceding component was fully dissolved.

4.7.4 Ligation of inserts into linearised plasmids

Gene inserts were ligated into linearised plasmid in a ligation reaction (20 μl). Ligation reactions consisted of gene insert and linear vector (50 ng) in a 3:1 molar ratio for gene inserts less than 3 kb in length, or a 1:1 molar ratio for gene inserts greater than 3 kb in length, along with T4 DNA ligase (1 unit) and Fermentas buffer for T4 DNA ligase (1X). Ligations were performed at room temperature for 20 min and ligation mixture (2.5 μl) was used as the source of plasmid for the transformation into *E. coli* JM107, which was performed using the Fermentas TransformAid bacterial transformation kit, as before. Transformed cells were plated on prewarmed LB agar plates supplemented with carbenicillin and were incubated at 37 °C for 16 h.

4.7.5 Subcloning a polynucleotide into an expression vector

To sub-clone a polynucleotide into a new vector, the source vector was first amplified by transforming it (1 ng) into *E. coli* JM107 using the Fermentas TransformAid™ bacterial transformation kit and then preparing an LB culture (5 ml) supplemented with carbenicillin for each of 5 unique colonies on the agar plate. The 5 × 5 ml LB+Carbenicillin cultures were incubated at 37 °C with shaking at 250 rpm for 16 h and plasmids were purified out of the culture using the QIAGEN QIAprep® Spin miniprep kit as described in the manual. The polynucleotides of interest were digested out of each plasmid preparation in restriction digest reactions (20 μl) using 1 μl of the required restriction endonuclease for each of the desired 5' and 3' cloning sites. The restriction digests were pooled together and the polynucleotide to be subcloned into the new vector was purified from the remainder of the source vector by agarose gel electrophoresis and gel extraction. A ligation reaction was established to ligate the polynucleotide insert into a linearised expression vector previously prepared with complementary cohesive ends to the insert. The ligation reaction (2.5 μl) was used as the source of plasmid for transformation into *E. coli* JM107, as already described. The expression plasmid was amplified by inoculating an LB+Carbenicillin culture (100 ml) with a single colony from the plate of transformants, incubating the culture at 37 °C with shaking at 200 rpm for 16 h, and purifying the plasmid from the culture using the QIAGEN Plasmid MidiPrep kit.

To clone a PCR product into an expression vector, the PCR product was first purified by agarose gel electrophoresis and gel extraction, its concentration determined by analytical agarose gel electrophoresis, blunt end ligated into the pJET1.2 vector of the CloneJET™ PCR cloning kit and the ligation product was transformed into *E. coli* JM107 for amplification. This was the source plasmid for the remainder of the cloning procedure.

4.7.6 Analytical agarose gel electrophoresis

Linear DNA homogeneity and concentration was determined by analytical agarose gel electrophoresis. Agarose gels were cast as before (section 4.7.3). A fixed volume (5 µl) of DNA sample was loaded onto the gel alongside GeneRuler™ 1 kb Plus DNA ladder (Thermo Scientific) (5 µl). Samples (5 µl) of restriction digests were loaded directly onto the gel, as the FD Green buffer already contained a loading dye. DNA samples (5 µl) from preparations that did not include a loading dye were mixed with 6X Loading dye (1 µl) and all 6 µl was then loaded. To estimate the concentration DNA after gel extraction, three different sample preparations were made, using 1 µl, 2 µl and 3 µl of DNA in the different samples, along with 6X loading dye (1 µl) and nuclease-free water to make each sample up to 6 µl and all 6 µl were loaded onto the gel. The intensity of the sample band was visually compared to the intensities of bands of similar size and known mass in the ladder, as described in the manual for the GeneRuler™ 1 kb Plus DNA ladder.

4.7.7 Preparation of genomic DNA extracts from *B. subtilis* 168 and *P. putida* KT2440.

Fresh colonies of *P. putida* KT2440 and *B. subtilis* 168 were grown on LB agar plates from cryostocks of the organisms, with incubation at 30 °C for 16 h. Overnight liquid cultures of each organism were prepared by inoculating LB media (10 ml) with a single colony for each respective organism. Cultures were incubated at 30 °C with shaking at 200 rpm for 16 h. Genomic DNA from *P. putida* KT2440 and *B. subtilis* 168 was then extracted out of 0.5 ml of the bacterial culture using the Invitrogen™ ChargeSwitch® gDNA Mini Bacteria Kit according to the instructions provided. Briefly, 0.5 ml of each cell culture was pelleted in 1.5 ml sterile microcentrifuge tubes by centrifugation at 6000 rpm for 10 min, and resuspended into Resuspension Buffer (R4) containing RNase A and lysozyme. The resuspended cells were then incubated for 10 min at 37 °C before Lysis Buffer L14 (500 µl) containing proteinase K solution (10 µl) was added to each set of resuspended cells. The cells were then mixed by repeated inversion of the capped tubes (6-10 inversions). The cells were then incubated at 55 °C for 10 min before ChargeSwitch® Magnetic Beads (40 µl) were added to each set of lysed cells to bind

the genomic DNA. The beads were washed with the suitable Wash Buffer W12 (1 ml) and the tubes containing lysed cells and magnetic beads were placed in a magnetic rack to gather the magnetic beads before the supernatant was removed by pipette, and discarded. The tubes were removed from the rack, beads washed once more using Wash Buffer 12, and the tubes returned to the magnetic rack so that the wash buffer could be removed without disturbing the magnetic beads with bound genomic DNA. The tubes were removed from the magnetic rack, and Elution Buffer E5 (200 µl) was added to each tube and the beads were incubated in the elution buffer at 55 °C for 5 min. The beads were mixed by pipette halfway through the incubation with the elution buffer. The tubes were returned to the magnetic rack and the beads allowed to accumulate for 1 min before the genomic DNA in the elution buffer was removed from the tubes by pipette and the beads discarded. The concentration of the extracted genomic DNA from each organism was determined using the NanoDrop ND1000 spectrophotometer.

4.7.8 Polymerase chain reactions

4.7.8.1 Colony PCR

Colony PCR was performed to detect the presence of the pET20b(+) plasmid in colonies of transformed *E. coli*, and to confirm the size of the gene insert between the T7 promoter and T7 terminator of the plasmid. Individual colonies from streaked out cryostocks of transformants, or individual colonies taken directly from a plate of transformants, were selected with a sterile 200 µl plastic pipette tip. A copy of the selected colony was made by gently touching the pipette tip onto a fresh LB agar plate supplemented appropriately for the strain (see Table 6), and the remainder of the cells on the pipette tip were transferred to a PCR reaction mixture (50 µl) containing T7 Promoter primer (T7_F) (0.25 mM), T7 Terminator primer (T7_R) (0.25 mM), Sigma JumpStart™ REDTaq® DNA Polymerase (0.2 units.µl⁻¹), Sigma JumpStart buffer (1X), deoxyadenosinetriphosphate (dATP) (0.2 mM), thymidine triphosphate (dTTP) (0.2 mM), deoxycytidinetriphosphate (dCTP) (0.2 mM), deoxyguanosinetriphosphate (dGTP) (0.2 mM). Samples were loaded for a hot start in a thermocycler, programmed to start at 94 °C for 3 min, then to cycle through 30 iterations of 30 s melting at 94 °C, 30 s annealing at 50 °C, 80 s elongation at 72 °C, before ending with 5 min at 72 °C.

4.7.8.2 Modification of gene encoding 4HBT to introduce a C-terminal His-Tag.

Polymerase chain reaction mixtures contained the pET20b(+):4HBT plasmid as template DNA (50 pg.µl⁻¹), KOD DNA polymerase (1 unit), primer 4HBT_CtHis_F (0.4 µM), primer 4HBT_CtHis_R (0.4 µM), dATP (0.2 mM), dTTP (0.2 mM), dCTP

(0.2 mM), dGTP (0.2 mM), MgCl₂ (1 mM) and Novagen buffer #1 for KOD DNA polymerase (1X). PCR mixtures were loaded into a thermocycler programmed to start at 94 °C for 3 min, then to cycle through 30 iterations of 30 s melting at 94 °C, 30 s annealing at 55 °C, 80 s elongation at 72 °C, before ending with 5 min at 72 °C.

4.7.8.3 Overlap extension PCR to concatenate 4HBT and ACX4 genes

Two PCR reactions, A and B, were established in parallel. Polymerase chain reaction A (50 µl) contained pET20b(+):4HBT (10 pg.µl⁻¹) as the template DNA, KOD DNA polymerase (1 unit), primer OE_4HBT_F (0.4 µM), primer OE_4HBT_R (0.4 µM), dATP (0.2 mM), dTTP (0.2 mM), dCTP (0.2 mM), dGTP (0.2 mM), MgCl₂ (1 mM) and Novagen buffer for KOD DNA polymerase (1X). Polymerase chain reaction B (50µl) was composed of pET20b(+):ACX4 (20 pg.µl⁻¹) as template DNA, KOD DNA polymerase (1 unit), primer OE_ACX4_F (0.4 µM), primer OE_ACX4_R (0.4 µM), dATP (0.2 mM), dTTP (0.2 mM), dCTP (0.2 mM), dGTP (0.2 mM), MgCl₂ (1 mM) and Novagen buffer for KOD DNA polymerase (1X).

Both PCR reactions were performed in parallel and under the same conditions, commencing with an initial denaturation step at 95 °C for 3 min, followed by 25 cycles consisting of a 15 s denaturation step at 98 °C, a 2 s annealing step at 50 °C and a 20 s extension step at 72 °C. These 25 cycles were followed by a further 5 min extension step at 72 °C. The two double stranded PCR products, product A and product B, were purified by agarose gel electrophoresis and gel extraction, and their concentrations determined by analytical agarose gel electrophoresis.

PCR products A and B were then concatenated in a second polymerase chain reaction consisting of PCR product A (15 nM), PCR product B (15 nM), dATP (0.2 mM), dTTP (0.2 mM), dCTP (0.2 mM), dGTP (0.2 mM), MgCl₂ (1 mM), Novagen Buffer #1 for KOD DNA polymerase (1X), dimethylsulfoxide (5% (v/v)) and KOD DNA polymerase (8 nl.µl⁻¹). The reaction was performed in a thermocycler programmed to start at 95 °C for 3 min and continue with 15 cycles of a 15 s denaturation step at 98 °C, a 2 s annealing step at 50 °C and a 20 s extension step at 72 °C, and to finally end with a further extension step lasting 5 min at 72 °C.

After this PCR program, the outer forward primer, OE_4HBT_F, and the outer reverse primer, OE_ACX4_R, used in amplification reactions A and B, respectively, were added directly from a concentrated stock (50 µM) to a final concentration of 0.5 µM. The modified PCR reaction mixture was then subjected to another round of PCR, with the thermocycler programmed to commence with an initial 3 min melting step at 95 °C, and continue with 15 cycles consisting of a melting step lasting 15 s

at 98 °C, an annealing step lasting 2 s at 55 °C and an extension step lasting 20 s at 72 °C, before finishing with a further extension step lasting 5 min at 72 °C.

4.7.8.4 PCR to amplify branched chain keto acid dehydrogenase genes out of genomic DNA.

Polymerase chain reactions were composed of genomic DNA (2 ng.µl⁻¹), dATP (0.2 mM), dTTP (0.2 mM), dCTP (0.2 mM), dGTP (0.2 mM), Fermentas High Fidelity buffer for Phusion Hot Start II DNA polymerase (1X), primer bsBCKD_F (0.5 µM), primer bsBCKD_R (0.5 µM) and Fermentas Phusion Hot Start II DNA polymerase, made up to a final volume of 50 µl in a thin walled PCR tubes (200 µl). The thermocycler was programmed to commence with an initial denaturation step at 98 °C for 30 s, followed by 30 cycles of denaturation at 98 °C for 10 s, annealing at 55 °C for 15 s and extension at 72 °C for 90 s, followed finally by one extra extension step at 72 °C for 5 min.

4.8 Synthetic and analytical chemistry methods

4.8.1 Synthesis and purification of methacrylyl-CoA

The synthesis of methacrylyl-CoA was performed by the reaction of coenzyme A with methacrylic anhydride. Various starting amounts of coenzyme A were tested, ranging between 10 mg and 100 mg, but 30 mg of coenzyme A was found to be the most practical starting mass. Thus, the reaction consisted of coenzyme A (19.5 mM) and methacrylic anhydride (41.8 mM) in sodium phosphate buffer (100 mM, pH 8.5). The reaction was incubated on ice and vortexed every 2 min for 30 min and the final reaction mix was acidified to pH 3.5 with hydrochloric acid. Methacrylic acid byproduct and unreacted methacrylic anhydride were extracted four times with water saturated diethyl ether, using 0.4 ml for every 1 mg of coenzyme A in the reaction. Methacrylyl-CoA was purified by reverse phase high performance liquid chromatography (RP-HPLC) on an analytical scale column (Agilent Zorbax Eclipse XDB C18 column, 4.6 mm × 150 mm), using UV detection at 260 nm. Sample (75 µl) was injected onto the C18 column, and eluted over 40 min by a linear acetonitrile gradient (1.8% - 13.5%) in 0.1% trifluoroacetic acid (TFA), at a flowrate of 1 ml.min⁻¹. Methacrylyl-CoA was present in the major peak, and the eluent from this peak was collected. HPLC purification was repeated on the remaining sample, loading 75 µl at a time, and the purified methacrylyl-CoA collected from each HPLC run was pooled together. Acetonitrile was removed by rotary evaporation (21 °C, 3 kPa), leaving behind an aqueous solution of methacrylyl-CoA and trifluoroacetic acid. This solution of methacrylyl-CoA and trifluoroacetic acid was brought to pH 7 by sodium hydroxide and flash frozen with liquid nitrogen prior to lyophilisation. TFA was removed by re-dissolving the freeze

dried sample in nuclease free water (10 ml), and repeating the freeze-dry-redissolve cycle twice more, once in 5 ml, and finally in 1 ml nuclease free water. The concentration of methacrylyl-CoA was determined by absorbance at 260 nm with a molar extinction coefficient of $16800 \text{ M}^{-1} \cdot \text{cm}^{-1}$. The purity of methacrylyl-CoA was determined by HPLC. A sample of the purified methacrylyl-CoA preparation (10 μl) was diluted 20-fold by the addition of 190 μl water. The sample was acidified by the addition of 5M HCl (2 μl), and loaded onto the same Agilent Zorbax Eclipse XDB-C18 4.6 mm $\times$ 150 mm column that was used to purify the methacrylyl-CoA, but equilibrated instead with 1.8% acetonitrile in 75 mM KH_2PO_4 acidified to pH 2.5 with 5 M HCl. The methacrylyl-CoA was eluted by gradient elution with starting conditions of 1.8% acetonitrile in 75 mM KH_2PO_4 acidified to pH 2.5, rising to 13.5% acetonitrile over 40 min.

4.8.2 Analysis of samples for methacrylic acid production by whole cell biotransformations and whole cell cultures on a C18 column.

Samples were centrifuged in an Eppendorf miniSpin F-45-12-11 rotor for 10 min at 6000 rpm. The supernatants were filtered through 0.2 micron Sartorius RC 4 mm filters and then acidified to pH 2.5 with 5 M HCl, using 10 μl of acid per 1 ml of sample taken from cultures grown in MSX medium, and 60 μl of acid per 1 ml of sample taken from cultures grown in SM medium. The acidity was confirmed to be below pH 3 by aliquoting 20 μl of acidified sample over the pH sensitive areas of Fisher non-bleed pH 0 to pH 6 pH indicator paper sticks. Acidified supernatant was then injected onto an Agilent Zorbax Eclipse XDB C18 column (4.6 mm $\times$ 250 mm). HPLC was performed at $1 \text{ ml} \cdot \text{min}^{-1}$ and analytes were eluted by isocratic elution with 14% acetonitrile in 25 mM KH_2PO_4 that was adjusted to pH 2.5 with HCl. The column was equilibrated for 30 min under the starting conditions prior to the first sample, and between runs the column was washed by first switching the mobile phase to 14% acetonitrile in water and holding under those conditions for 10 min, prior to increasing the acetonitrile concentration to 90% within 1 min. This high acetonitrile concentration was held for 10 min, before returning to 14% acetonitrile in water in 1 min, and holding for a further 10 min. The column was then re-equilibrated for the next sample by switching back to 14% acetonitrile in KH_2PO_4 (25 mM) acidified to pH 2.5 with HCl and holding for 15 min. Detection was performed by UV absorbance at 215 nm. Methacrylic acid calibration curves were plotted for both the biotransformation of isobutyric acid to methacrylic acid (Figure 77, appendices, page XXII), and the production of methacrylic acid direct from glucose (Figure 78, appendices, page XXII).

4.8.3 Isocratic elution on ion exchange column

Additional analysis of whole cell cultures was performed on a Phenomenex® Rezex™ ROA column (7.8 mm × 300 mm) by isocratic elution with H₂SO₄ (0.01 N) for 35 min at 0.5 ml.min⁻¹. Both detection by UV absorbance at 215 nm and refractive index detection were used.

4.9 Enzyme expression, extraction and purification

4.9.1 Preparation of seed cultures

Seed cultures were prepared by selecting a single colony on an LB agar plate supplemented with the appropriate supplements, with a 10 ml sterile disposable inoculation loop. The colony was then transferred to 10-20 ml growth media in a 50 ml sterile falcon tube, supplemented with the appropriate supplements, and incubated at 37 °C with shaking at 250 rpm. Falcon tubes were sealed with loose lids that were secured in place with tape. Cultures that were grown in complex or semi-defined media containing yeast extract were incubated until an OD₆₀₀ of 1, which took approximately 8 h. Cultures grown in minimal media were grown for 16-24 h, until the desired OD was reached. Seed cultures were stored at 4 °C for a maximum of 24 h.

4.9.2 Culture inoculation from seed culture

In order to avoid diluting the growth media with spent media, the inoculation of growth media was performed from cell pellets of the seed culture that were re-suspended in fresh growth media. Seed cultures were first pelleted by centrifugation at 5000 g, 4 °C for 20 min. The supernatant was discarded and the pellets were re-suspended in fresh growth media (1 ml). If the growth media used for the seed culture and the culture to be inoculated were different, the same growth media as the one to be used for the main culture was used to re-suspend the cell pellet from the seed culture. To minimise errors arising from incomplete re-suspension of cell pellets or incomplete removal of supernatant, the volume of re-suspended cell pellets to use as the inoculum was calculated from the OD reading of the re-suspended cells and not of the seed-culture prior to its centrifugation. The OD₆₀₀ of the re-suspended cell pellet was measured by diluting seed culture (20 µl) fifty-fold into water or fresh media (980 µl), and measuring the OD₆₀₀ of this diluted sample. This reading was then multiplied by the dilution factor, to yield the OD₆₀₀ reading of the re-suspended cell pellet. The volume of the seed culture required to inoculate the main culture to a target OD was then calculated by the equation shown overleaf (Equation 2).

$$V_{RP} = \frac{OD_{TARGET} \cdot V_{MEDIA}}{OD_{RP} - OD_{TARGET}}$$

Equation 2: Equation to calculate the volume of resuspended pellet from seed culture required to add to culture media to reach a desired target starting OD. V_{RP} : Volume of resuspended pellet to add to culture media. V_{MEDIA} : Volume of culture media in shake flask. OD_{TARGET} : Desired target starting optical density. OD_{RP} : Optical density of the re-suspended cell pellet. For derivation, see section 8.4.2.

For example, if the OD_{600} reading of the cell pellet (OD_{RP}) was 25, volume of the culture media (V_{MEDIA}) to be inoculated was 100 ml, and the target starting OD_{600} of the culture (OD_{TARGET}) culture was to be inoculated to an OD_{600} reading of 0.2, the volume of re-suspended pellet to add was 806 μ l. The derivation of the above equation is described in the appendices (section 8.4.2). Whilst the percentage error that would result from assuming the final volume of the culture after inoculation is the same as that prior to inoculation is less than 1%, the equation above is simple and quick to use, and by taking into account volume changes, is more accurate.

4.9.3 Protein expression at 100 ml scale, biotransformations and shake flask fermentations

Expression at 100 ml scale was performed in 500 ml shake flasks. Flasks were stoppered with foam bungs to enable gas exchange with the environment and the bungs were loosely covered with foil. The flasks were then autoclaved, and sterile media (100 ml) was added to the flasks aseptically. The media was supplemented with chloramphenicol and carbenicillin, or as otherwise required (Table 6). Seed cultures of expression strains were prepared from cryostocks, as described above (section 4.9.1), and were used to inoculate the growth media (section 4.9.2). Cultures were incubated at 37 °C, or as otherwise required, with shaking at 250 rpm, and OD_{600} readings were taken regularly. At OD_{600} 0.5, a 1.5 ml sample was taken and expression was then induced at OD_{600} 0.5 with IPTG (0.4 mM). Further samples (1.5 ml) were then taken at intervals thereafter. For the detection of methacrylic acid production, a 0 h post induction sample was taken immediately after expression was induced. Samples were then centrifuged in an Eppendorf miniSpin F-45-12-11 rotor at 6000 rpm for 10 min. For the detection methacrylic acid production, supernatants were filtered through a 0.22 μ m filter and stored at -20 °C until required. For the detection of protein expression, the supernatant was removed entirely, and cell pellets were stored at -80 °C until analysed by SDS-PAGE. Cultures were sometimes supplemented with isobutyrate (5 mM, pH 7.0) or 2-ketoisovalerate (5 mM, pH 7.0). Isobutyrate supplement was prepared from isobutyric acid that was adjusted to pH 7.0 with KOH and then filter sterilised with a 0.22 μ m sterile filter, while 2-ketoisovalerate supplemented was prepared

from 2-ketoisovalerate sodium salt hydrate that was adjusted to pH 7.0 with HCl and then filtered with a 0.22 μm filter. The pH of the filtered 2-ketoisovalerate was rechecked and then filter sterilised with a 0.22 μm sterile filter.

4.9.4 Protein expression at 500 ml scale and cell harvesting.

Seed cultures of expression strains were prepared from cryostocks, as described (section 4.9.1), and were used to inoculate 100 ml LB or MSX growth media (section 4.9.2) supplemented with glucose, riboflavin, chloramphenicol and carbenicillin, or as otherwise required (Table 6) in 500ml shake flasks. Cultures were incubated at 37 °C, or as otherwise required, with shaking at 250 rpm, until an OD₆₀₀ of 1 was reached, and cultures were stored at +4 °C overnight. Cultures were then aseptically decanted into two sterile 50 ml falcon tubes, and cells were pelleted by centrifugation at 6000 $\times$ g in an Eppendorf F-34-6-38 rotor for 20 min at 4 °C. Cell pellets were then resuspended into 500 ml growth media in 2.5 L shake flasks to an OD₆₀₀ of ~0.2. Cultures were then incubated at 37 °C with shaking at 250 rpm for 5-7 h, or until stationary phase was reached. Samples (1.5 ml) were then taken for analysis by SDS-PAGE to confirm protein expression and the remaining culture was decanted evenly into 2 $\times$ 500 ml centrifuge tubes for the Beckman-Coulter JA10.500 rotor, and centrifuged at 6000 $\times$ g for 20 min at 4 °C. The supernatant was decanted off, and cell pellets in each tube were washed by resuspending them in 50 ml HEPES buffer (50 mM, pH 7.5) by gentle agitation. Resuspended cells were transferred to 50 ml Falcon tubes and centrifuged at 6000 $\times$ g for 20 min at 4 °C in an Eppendorf F-34-6-38 rotor. The supernatant was decanted once more, and cell pellets were stored at -80 °C until required.

4.9.5 Preparation of cell free protein extracts

Cell pellets from the 500 ml scale expression cultures were resuspended in 6 ml HEPES buffer (50 mM, pH7.5) containing 0.6 μl Benzonase nuclease. For ACX4, 10 μM FAD was added to HEPES buffer before cells were lysed, and HEPES buffer with 10 μM FAD was used throughout the entire process in place of HEPES buffer. The resuspended cells were lysed in a Constant Systems Single Shot cell disrupter, and lysed cells were transferred to 15 ml falcon tubes to be clarified by centrifugation at 18,000 $\times$ g for 20 min at 4 °C in an Eppendorf F-34-6-38 rotor. Supernatants of the clarified cell lysates were then transferred to Beckman-Coulter 10 ml polycarbonate tubes and further clarified by centrifugation at 57,750 $\times$ g for 60 min at 4 °C in a Beckman-Coulter JA25.50 rotor to remove the membrane fraction. The supernatants were then washed with a VivaSpin Viva6 10,000 Molecular Weight Cut-Off centrifugal concentrator, centrifuging at 10,000 $\times$ g at 18 °C in an Eppendorf F-34-6-38 rotor until a six-fold volume reduction, followed

by a six-fold re-dilution in the HEPES assay buffer and a further 6-fold volume reduction in the centrifugal concentrator by further centrifugation. The concentrated cell free extracts were then aliquoted (200 μ l) into thin walled PCR tubes and flash frozen. To flash freeze the cell free extracts, the PCR tubes containing the cell free extracts were dropped into a Dewar of liquid nitrogen and allowed to freeze. The liquid nitrogen was then decanted from the Dewar and the PCR tubes kept at -80 °C until required. The total protein concentration of the cell free extracts was determined by Bio-Rad DC assay, as described in the Bio-Rad DC assay manual, but scaled down 5-fold, using bovine serum albumin (BSA) as the protein standard. Calibration curves with BSA in HEPES (50 mM, pH 7.5) buffer were plotted in triplicate (Figure 74, Appendices, page XVIII), by first preparing three separate 1.5 mg.ml⁻¹ protein standards and preparing separate dilutions from each protein standard. No Bio-Rad reagent 'S' was used in these assays.

4.9.6 Preparation of pure Carboxy terminal His-tagged 4HBT

Cell free extracts were prepared for cell pellets from the 500 ml expression scale cultures of His-tagged 4HBT, but cells were resuspended into FPLC Nickel-Sepharose binding buffer, instead. Binding buffer consisted of NaH₂PO₄ (10 mM), Na₂HPO₄ (10 mM), NaCl (500 mM), imidazole (30 mM) and was adjusted to pH 7.4 with HCl. Binding buffer was filtered with 0.22 μ m nitrocellulose filters with a Buchner funnel under vacuum, and degassed under vacuum and sonication for 30 min. Cell lysates and membrane fractions were removed by centrifugation as before, but the clarified supernatant was loaded onto a GE Healthcare HisTrap™ FF Crude column (1 ml) that was equilibrated with binding buffer. Unbound proteins were washed off the column with five column volumes of binding buffer, and His-tagged protein was eluted 20 column volumes with a linear imidazole concentration gradient, from 30 mM imidazole in the binding buffer to 500 mM imidazole in the FPLC Nickel-Sepharose elution buffer. Elution buffer consisted of NaH₂PO₄ (10 mM), Na₂HPO₄ (10 mM), NaCl (500 mM), imidazole (500 mM) and was adjusted to pH 7.4 with HCl. Binding buffer was filtered with 0.22 μ m nitrocellulose filters with a Buchner funnel under vacuum, and degassed under vacuum and sonication for 30 min. Protein elution was monitored at 280 nm and fractions were checked for the presence and purity of carboxy-terminal His-tagged 4HBT by SDS-PAGE. Fractions containing pure CtHis-4HBT protein were pooled and a buffer exchange to replace the elution buffer with potassium phosphate assay buffer (100 mM, pH 7.5) was performed in a VivaSpin Viva6 10,000 Molecular Weight Cut-Off centrifugal concentrator. To perform the buffer exchange, the pooled fractions were centrifuged through the ultrafiltration membrane of the centrifugal concentrator at

10,000 × g at 18 °C until the volume of the pooled fractions was reduced to 1 ml. The remaining protein fraction was diluted six-fold in the potassium phosphate assay buffer and the samples concentrated by further centrifugation through the ultrafiltration membrane under the same conditions until a six-fold volume reduction was achieved. The latter dilution in potassium phosphate assay buffer and re-concentration was performed once more and the concentration of CtHis-4HBT protein was determined by UV₂₈₀ absorbance in a NanoDrop ND1000 spectrophotometer, using a molar extinction coefficient of 20970 M⁻¹.cm⁻¹ and a molecular weight of 17516.5 mg.mmol⁻¹ for the carboxy-terminal His-tagged 4HBT enzyme. The values for the molar extinction coefficient and molecular weight were determined using the ExPASy ProtParam tool, using the amino acid translation of the gene sequence encoding the carboxy-terminal his tagged 4HBT enzyme (Polypeptide sequence 6) in the pET20b(+):CtHis-4HBT plasmid. Purified His-tagged 4HBT was flash frozen by first loading a syringe with the pure enzyme in phosphate buffer, and drop freezing the small droplets through the syringe needle directly into a hemispherical Dewar of liquid nitrogen. Once the droplets had frozen, the liquid nitrogen was decanted off. The frozen enzyme droplets were collected into a cryovial with a loose lid to ensure that any residual liquid nitrogen could safely evaporate, and were then stored at -80 °C.

4.10 Enzymatic assays

4.10.1 Thioesterase assays

Crude thioesterase assays were performed in triplicate in Nunc 96 well plates. Wells contained DTNB (500 µM), cell-free enzyme extract (0.5 mg.ml⁻¹) and either methacrylyl-CoA or isobutyryl-CoA as substrate (200 µM) in HEPES buffer at a final volume of 200 µl. The assays were prepared by first creating a master-mix of DTNB (1 mM) from 2.5 mM freshly prepared stock solution and substrate (400 µM) from 10 mM stock solution, all in HEPES buffer. Cell free enzyme extract was also diluted to 1 mg.ml⁻¹ in HEPES buffer. The 96-well plate was pre-warmed to 30 °C in the Fluostar plate-reader, and the master mix and cell free enzyme extract were also pre-warmed to 30 °C in a water bath for 10 min. To each assay well, 100 µl master mix was added, and the reactions were started simultaneously by the addition of 100 µl of 1 mg.ml⁻¹ cell free enzyme extract using a multi-channel pipette. Reactions were monitored using a 410-420 nm UV filter. The DTNB stock (2.5 mM) was freshly prepared by dissolving 20 mg DTNB into 20.18 ml HEPES buffer and filtering the solution to ensure that no undissolved DTNB remained.

Calibration curves were prepared for the assay using β -mercaptoethanol as the sulfhydryl standard (Figure 75, appendices, page XX). A 10 mM stock solution of β -mercaptoethanol in water was prepared by three serial 1 in 10 dilutions from a 1 M stock solution of β -mercaptoethanol in water. This 1 M stock solution was prepared by adding 70.4 μ l of β -mercaptoethanol to 929.6 μ l water. Separate β -mercaptoethanol standards with final concentrations between 0 μ M and 300 μ M were prepared in assay wells also containing DTNB (500 μ M) to a final volume of 200 μ l. The calibration curves were performed in triplicate, starting from three separately prepared 1 M β -mercaptoethanol stock solutions.

4.10.2 Kinetic characterisation of pure His-Tagged 4HBT.

Kinetic characterisation of His-tagged 4HBT was performed with purified CtHis-4HBT in Nunc 96 well plates in triplicate, in a similar manner to the crude enzyme assays, for both isobutyryl-CoA and methacrylyl-CoA, but in phosphate buffer. Assay wells contained substrate (0 μ M - 500 μ M starting concentration), DTNB (500 μ M) and purified enzyme (75 μ g.ml⁻¹). Reactions were started by the addition of 100 μ l of enzyme at 2X concentration to a 100 μ l substrate-DTNB premix at 2X concentration. Calibration curves were again prepared in triplicate for 0 mM – 300 μ M β -mercaptoethanol, but in phosphate buffer (Figure 76, appendices, page XX).

4.10.3 Crude acyl-CoA oxidase assay and enzyme coupled reaction.

The activity test of ACX4 were performed in 1.5 ml micro-centrifuge tubes. Crude enzyme reactions consisted of cell free ACX4 protein extract (0.8 mg.ml⁻¹), isobutyryl-CoA (500 μ M) and FAD (10 μ M) in HEPES buffer (50 mM, pH 7.5). The reaction was incubated at 30 °C in the 1.5 ml micro-centrifuge tube, with shaking at 250 rpm for 30 min before the reaction was stopped by the acidification to pH ~1 using 5 M HCl. Enzyme coupled reactions were performed with CtHis-4HBT (0.6 mg.ml⁻¹), that was prepared and purified in the same manner as for the kinetic characterization, but with a buffer exchange into HEPES buffer performed instead. The analysis of crude enzymatic assays was performed by high performance liquid chromatography (HPLC) on an Agilent 1200 HPLC system, using an Agilent Zorbax eclipse XDB C18 column, 4.6 mm $\times$ 150 mm. The column was equilibrated under starting conditions for 30 min prior to the first sample, whereby starting conditions were 1.8% acetonitrile in potassium phosphate (75 mM) that was acidified to pH 2.5 with HCl. As the enzymatic reactions were stopped by acidification, no further acidification was required for analysis on the C18 column. Sample (20 μ l) was injected onto the column, and was eluted over a 40 min acetonitrile gradient, starting at 1.8% and ending at 13.5%. The column was washed between each run

by first switching to 13.5% acetonitrile in water, and washing off phosphate salts for 10 min, before the acetonitrile concentration in the mobile phase was increased to 90% in 1 min, and held at 90% for 10 min. The acetonitrile concentration was then reduced back to 1.8% in 1 min and held for 10 min before switching the mobile phase to 1.8% acetonitrile in potassium phosphate (50 mM) acidified to pH 2.5 with HCl and held for 10 min for the column to re-equilibrate.

For authentic standards and methacrylic acid calibration curves, standards were dissolved in water and acidified with 10 μ l of 6 M HCl for every 1 ml of sample. HPLC was performed over a 40 min elution gradient as before, but without washing the column between standards. Mobile phase was returned to 1.8% acetonitrile in potassium phosphate (50 mM) that was acidified to pH 2.5 with HCl and the column was re-equilibrated under those conditions for 10 min prior to the next sample.

4.11 Table of plasmids

PLASMID REFERENCE	SOURCE	DESCRIPTION
pBMH::4HBT	Biomatik Corporation	A cloning vector conferring resistance to ampicillin and containing a codon optimised gene encoding 4HBT flanked by the <i>NdeI</i> and <i>NotI</i> restriction sites.
pBMH::YciA	Biomatik Corporation	A cloning vector conferring resistance to ampicillin and containing a codon optimised gene encoding YciA flanked by the <i>NdeI</i> and <i>NotI</i> restriction sites.
pBMH::FcoT	Biomatik Corporation	A cloning vector conferring resistance to ampicillin and containing a codon optimised gene encoding FcoT flanked by the <i>NdeI</i> and <i>NotI</i> restriction sites.
pBMH::RifR	Biomatik Corporation	A cloning vector conferring resistance to ampicillin and containing a codon optimised gene encoding RifR flanked by the <i>NdeI</i> and <i>NotI</i> restriction sites.
pBMH::AcsA	Biomatik Corporation	A cloning vector conferring resistance to ampicillin and containing a codon optimised gene encoding AcsA flanked by a 3' <i>XhoI</i> restriction site and a 5' sequence containing an <i>NheI</i> restriction site, a ribosome binding site, a spacer sequence and an <i>NdeI</i> restriction site.
pBSK::ppBCKD	Biomatik Corporation	A cloning vector with the four genes encoding the branched chain keto acid dehydrogenase complex of <i>Pseudomonas putida</i> KT2440, flanked by a 3' <i>XhoI</i> restriction site and a 5' sequence composed of an <i>XbaI</i> restriction site, a spacer sequence, an <i>NheI</i> restriction site, a ribosome binding site and a second spacer sequence.
pMA-RQ::ACX4	Life Technologies	A cloning plasmid conferring resistance to ampicillin, and containing a codon optimised gene

		encoding ACX4 from <i>Arabidopsis thaliana</i> , flanked by a 5' <i>NdeI</i> and a 3' <i>XhoI</i> restriction site.
pJET1.2	Thermo Scientific	The pJET1.2 linearised blunt end cloning vector by Thermo Scientific.
pJET1.2::CtHIS-4HBT	This work	The pJET1.2 cloning vector containing the product of a polymerase chain reaction to replace the stop codon of 4HBT with a 5'-GGA-3' sequence followed by a <i>XhoI</i> site, such that on sub-cloning the insert into pET20b(+), the gene formed encodes 4HBT with a C-terminal His-tag appended to the end, with a tripeptide linker in between.
pJET1.2::4HBT-ACX4	This work	The pJET1.2 cloning vector containing the product of an overlap extension polymerase chain reaction that linked the 4HBT and ACX4 genes into a single polynucleotide with a new ribosome binding site in between the two genes.
pET20b(+)	Novagen	The pET20b(+) expression vector by Novagen.
pET20b(+)::4HBT	This work	The pET20b(+) expression vector containing the codon optimised gene encoding 4HBT from <i>Arthrobacter sp.</i> strain SU between the <i>NdeI</i> and <i>NotI</i> restriction sites.
pET20b(+)::YciA	This work	The pET20b(+) expression vector containing the codon optimised gene encoding YciA from <i>Haemophilus influenzae</i> between the <i>NdeI</i> and <i>NotI</i> restriction sites.
pET20b(+)::FcoT	This work	The pET20b(+) expression vector containing the codon optimised gene encoding FcoT from <i>Mycobacterium tuberculosis</i> between the <i>NdeI</i> and <i>NotI</i> restriction sites.
pET20b(+)::RifR	This work	The pET20b(+) expression vector containing the codon optimised gene encoding RifR from <i>Amycolatopsis mediterranei</i> between the <i>NdeI</i> and <i>NotI</i> restriction sites.

pET20b(+)::CtHIS-4HBT	This work	The pET20b(+) expression vector with the insert from pJET1.2::CtHIS-4HBT subcloned in between the <i>NdeI</i> and <i>XhoI</i> restriction sites, for the expression of 4HBT with a Gly-Leu-Glu-His-His-His-His-His peptide appended to the carboxy-terminus of each subunit.
pET20b(+)::ACX4	This work	The pET20b(+) expression vector containing the codon optimised gene encoding short chain acyl-CoA oxidase from <i>Arabidopsis thaliana</i> , cloned in between the <i>NdeI</i> and <i>XhoI</i> restriction sites.
pET20b(+)::AcsA	This work	The pET20b(+) expression vector containing the codon optimised gene encoding AcsA, cloned in between the <i>NdeI</i> and <i>XhoI</i> restriction sites.
pET20b(+)::4HBT-ACX4	This work	The pET20b(+) expression vector containing the 4HBT and ACX4 genes subcloned as a single polynucleotide from the pJET1.2::4HBT-ACX4 plasmid into the pET20b(+) plasmid in between the <i>NdeI</i> and <i>XhoI</i> restriction sites, such that an expression unit is formed for the co-expression of both 4HBT and ACX4, and capable of accepting a second insert between the newly introduced <i>NheI</i> site and the <i>XhoI</i> site.
pET20b(+)::4HBT-ACX4-AcsA Abbreviated to: pHAAC	This work	The pET20b(+)::4HBT-ACX4 expression vector containing the AcsA gene subcloned from the pBMH::AcsA plasmid in between its <i>NheI</i> and <i>XhoI</i> sites.
pET20b(+)::ppBCKD	This work	The pET20b(+) expression vector containing the four genes encoding the ppBCKD complex, subcloned from the pBSK::ppBCKD plasmid in between the <i>XbaI</i> and <i>XhoI</i> restriction sites.
pET20b(+)::4HBT-ACX4-ppBCKD	This work	The pET20b(+)::4HBT-ACX4 expression vector containing the four genes

Abbreviated to: pHAPP		encoding the ppBCKD complex, subcloned from the pBSK::ppBCKD plasmid into it in between its <i>NheI</i> and <i>XhoI</i> restriction sites.
pJET1.2::bsBCKD	This work	The pJET1.2 cloning vector containing the four genes encoding the branched chain keto acid dehydrogenase from <i>B. subtilis</i> 168.
pET20b(+):4HBT-ACX4-bsBCKD Abbreviated to: pHABS	This work	The pET20b(+):4HBT-ACX4 expression vector containing the four genes encoding the bsBCKD complex, subcloned from the pJET1.2::bsBCKD plasmid into it in between its <i>NheI</i> and <i>XhoI</i> restriction sites.
pSTV28::ilvD-alsS-ilvC Abbreviated to : pKIV	This work	The pSTV28 expression vector containing the alsS, ilvC and ilvD genes between the <i>EcoRI</i> and <i>HindIII</i> restriction sites.

Table 5: Table of all plasmids used and constructed in this study.

4.12 Table of strains and supplements required in growth media

Strain ID	Strain Description	Supplements required			
		glucose	Carbenicillin	Chloramphenicol	Riboflavin
Strain 1	<i>E. coli</i> JM107	No	No	No	No
Strain 2	<i>E. coli</i> JM107 pBMH::4HBT	No	Yes	No	No
Strain 3	<i>E. coli</i> JM107 pBMH::YciA	No	Yes	No	No
Strain 4	<i>E. coli</i> JM107 pBMH::FcoT	No	Yes	No	No
Strain 5	<i>E. coli</i> JM107 pET20b(+):4HBT	No	Yes	No	No
Strain 6	<i>E. coli</i> JM107 pET20b(+):YciA	No	Yes	No	No
Strain 7	<i>E. coli</i> JM107 pET20b(+):RifR	No	Yes	No	No
Strain 8	<i>E. coli</i> JM107 pET20b(+):FcoT	No	Yes	No	No
Strain 9	<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):4HBT	Yes	Yes	Yes	No
Strain 10	<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):FcoT	Yes	Yes	Yes	No
Strain 11	<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):YciA	Yes	Yes	Yes	No
Strain 12	<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):RifR	Yes	Yes	Yes	No
Strain 13	<i>E. coli</i> JM107 pET20PM::NtHIS-4HBT	No	Yes	No	No
Strain 14	<i>E. coli</i> BL21 (DE3) pLysS pET20PM::NtHIS-4HBT	Yes	Yes	Yes	No
Strain 15	<i>E. coli</i> JM107 pJET1.2::CtHIS-4HBT	No	Yes	No	No
Strain 16	<i>E. coli</i> JM107 pET20b(+):CtHIS-4HBT	No	Yes	No	No
Strain 17	<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):CtHIS-4HBT	Yes	Yes	Yes	No
Strain 18	<i>E. coli</i> JM107 pMA-RQ::ACX4	No	Yes	No	No
Strain 19	<i>E. coli</i> JM107 pET20b(+):ACX4	No	Yes	No	No
Strain 20	<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):ACX4	Yes	Yes	No	As Required
Strain 21	<i>E. coli</i> JM107 pBMH::AcsA	No	Yes	No	No
Strain 22	<i>E. coli</i> JM107 pET20b(+):AcsA	No	Yes	No	No
Strain 23	<i>E. coli</i> JM107 pCLONING::4HBT-ACX4	No	Yes	No	No
Strain 24	<i>E. coli</i> JM107 pET20b(+):4HBT-ACX4	No	Yes	No	No
Strain 25	<i>E. coli</i> JM107 pET20b(+):4HBT-ACX4-AcsA	No	Yes	No	No
Strain 26	<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):AcsA	Yes	Yes	Yes	No
Strain 27	<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):4HBT-ACX4	Yes	Yes	Yes	As Required
Strain 28	<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-AcsA	Yes	Yes	Yes	As Required
Strain 29	<i>Bacillus subtilis</i> 168	No	No	No	No
Strain 30	<i>Pseudomonas putida</i> KT2440 (DSM)	No	No	No	No
Strain 31	<i>E. coli</i> JM107 pJET1.2::bsBCKD	No	Yes	No	No

Strain 32	<i>E. coli</i> JM107 pBSK::ppBCKD	No	Yes	No	No
Strain 33	<i>E. coli</i> JM107 pET20b(+):ppBCKD	No	Yes	No	No
Strain 34	<i>E. coli</i> JM107 pET20b(+):4HBT-ACX4-ppBCKD	No	Yes	No	No
Strain 35	<i>E. coli</i> JM107 pET20b(+):4HBT-ACX4-bsBCKD	No	Yes	No	No
Strain 36	<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):ppBCKD	Yes	Yes	Yes	As Required
Strain 37	<i>E. coli</i> BL21 (DE3) pLysS pHAPP	Yes	Yes	Yes	As Required
Strain 38	<i>E. coli</i> BL21 (DE3) pLysS pHABS	Yes	Yes	Yes	As Required
Strain 39	<i>E. coli</i> BL21 (DE3) pHABS	Yes	Yes	No	As Required
Strain 40	<i>E. coli</i> BL21 (DE3) pHAPP	Yes	Yes	No	As Required
Strain 41	<i>E. coli</i> JM107 pKIV	Yes	No	Yes	No
Strain 42	<i>E. coli</i> BL21 (DE3) pKIV	Yes	No	Yes	As Required
Strain 43	<i>E. coli</i> BL21 (DE3) pKIV pHAPP	Yes	Yes	Yes	As Required
Strain 44	<i>E. coli</i> BL21 (DE3) pKIV pHABS	Yes	Yes	Yes	As Required

Table 6: Table of strains used and constructed in this study, and supplements required in growth media. Supplement concentrations as described in media recipes (Glucose in MSX: 5 g.L⁻¹, glucose in LB: 10 g.L⁻¹, glucose in SM: 15 g.L⁻¹, glucose in modified M9: 40 g.L⁻¹, Carbenicillin: 50 µg.ml⁻¹, Chloramphenicol: 34 µg.ml⁻¹). Riboflavin only added as required to expression cultures, when explicitly stated in results. Riboflavin added 'as required' depended on the particular experiment, and specific details are provided in the appropriate results chapter. Typically, seed cultures and intermediate cultures in minimal media were not supplemented with riboflavin, while seed cultures of complex media were supplemented with riboflavin. MSX and SM media are supplemented with 5 g.L⁻¹ and 15 g.L⁻¹ glucose, respectively, by default. LB medium was only supplemented with glucose (10 g.L⁻¹) when specifically stated.

5 Results

5.1 The selective hydrolysis of methacrylyl-CoA by a suitable thioesterase

5.1.1 The role of thioesterases in metabolism

As already discussed, the hydrolysis of a thioester is an exergonic process and occurs spontaneously in an aqueous environment [76] (Figure 19), but the uncatalysed rate of this hydrolysis is slow.

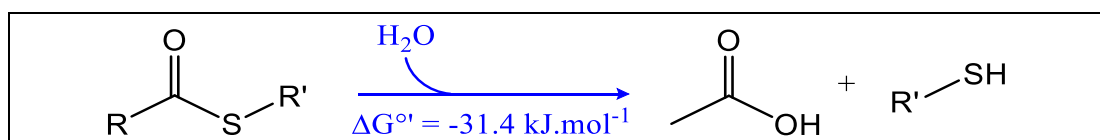


Figure 19: Hydrolysis of a thioester bond. The hydrolysis of a thioester produces a free carboxylate and a free thiol. The Gibbs free energy change of the reaction is $-31.4 \text{ kJ.mol}^{-1}$.

Thioesterases are known to catalyse this hydrolysis and belong to the enzyme commission (EC) groups 3.1.2.1 to 3.1.2.30, as well as the unclassified reactions covered by the miscellaneous group 3.1.2.- (Table 7, overleaf). Thioesterases have a plethora of metabolic functions in both eukaryotes and prokaryotes, as is highlighted by the range of subgroups in the thioesterase EC grouping. Of the 27 active EC groups pertaining to thioesterase activity, 18 groups are dedicated to thioesterases that hydrolyse esters of coenzyme A, 4 to thioesterases that hydrolyse thioesters of proteins such as acyl-carrier proteins, 4 to thioesterases that hydrolyse thioesters of glutathione, and 1 group, the miscellaneous EC 3.1.2.- group, to various general activities (Table 7) [71].

Thioesterases hydrolysing acyl-ACPs are typically involved in fatty acid synthesis (FAS), polyketide synthesis (PKS) and non-ribosomal peptide synthesis (NRPS), during which the biosynthesis of fatty acids, polyketides and peptides is employed on large protein complexes (megasynthases), whereby the fatty acid, polyketide or peptide is attached to an acyl carrier protein (ACP) or peptidyl carrier protein (PCP) through the phosphopantetheine arm of the ACP or PCP. Type I thioesterases are complex-associated thioesterase subunits that cleave the final product from the complex whilst type II thioesterases are discrete enzymes that, in the case of PKS and NRPS, hydrolyse aberrant intermediates that stall the biosynthesis of polyketides and peptides, and in the case of FAS, can act as alternatives to the type I thioesterases for the hydrolysis of the mature product [102, 103].

The glutathione thioester hydrolases are most likely involved in various detoxification pathways, due to the various cytoprotective roles that glutathione plays [104, 105]. For example, S-formyl-glutathione is formed by the formaldehyde

dehydrogenase catalysed reaction between toxic formaldehyde and glutathione. S-formyl-glutathione would then be hydrolysed to formic acid and glutathione by S-formyl-glutathione hydrolase. Similarly, various acyl-CoA thioesters are toxic due to their tendency to acylate nucleophilic amino acid residues such as cysteines, in proteins. Acyl-CoA thioesters are often also reactive with glutathione [106] and the glutathione-S-transferase catalysed nucleophilic attack of glutathione on acyl-CoA molecules mops excess acyl-CoAs up before they can exhibit toxic effects. The product of such a reaction with glutathione would be an S-acyl-glutathione, which would be hydrolysed to the less toxic carboxylic acid and glutathione by S-acyl-glutathione thioesterase.

EC Number	Enzyme activity	General activity	Thioesterase
3.1.2.1	Acetyl-CoA hydrolase	Coenzyme A thioester	
3.1.2.2	Palmitoyl-CoA hydrolase	Coenzyme A thioester	
3.2.1.3	Succinyl-CoA hydrolase	Coenzyme A thioester	
3.1.2.4	3-Hydroxyisobutyryl-CoA hydrolase	Coenzyme A thioester	
3.1.2.5	Hydroxymethylglutaryl-CoA Hydrolase	Coenzyme A thioester	
3.1.2.6	S-Hydroxyacyl-glutathione hydrolase	Glutathione thioester	
3.1.2.7	S-Acyl-glutathione Thioesterase	Glutathione thioester	
3.1.2.8	DELETED, Formerly S-acetoglutathione hydrolase	-	
3.1.2.9	DELETED, Formerly S-acetoacetylhydrolipoate	-	
3.1.2.10	Formyl-CoA Hydrolase	Coenzyme A thioester	
3.1.2.11	Acetoacetyl-CoA Hydrolase	Coenzyme A thioester	
3.1.2.12	S-formyl-glutathione hydrolase	Glutathione thioester	
3.1.2.13	S-succinyl-glutathione hydrolase	Glutathione thioester	
3.1.2.14	Oleyl-[Acyl Carrier Protein] hydrolase	Protein-thioester	
3.1.2.15	TRANSFERRED TO 3.4.19.12	-	
3.1.2.16	Citrate-lyase deacetylase	Protein-thioester	
3.1.2.17	(S)-methylmalonyl-CoA hydrolase	Coenzyme A thioester	
3.1.2.18	ADP dependent short chain acyl-CoA hydrolase	Coenzyme A thioester	
3.1.2.19	ADP dependent medium chain acyl-CoA hydrolase	Coenzyme A thioester	
3.1.2.20	Acyl-CoA hydrolase	Coenzyme A thioester	
3.1.2.21	Dodecanoyl-ACP hydrolase	Protein-thioester	
3.1.2.22	Palmitoyl-protein hydrolase	Protein-thioester	
3.1.2.23	4-Hydroxybenzoyl-CoA thioesterase	Coenzyme A thioester	
3.1.2.24	TRANSFERRED to 3.13.1.3	-	
3.1.2.25	Phenylacetyl-CoA hydrolase	Coenzyme A thioester	
3.1.2.26	Bile acid CoA Hydrolase	Coenzyme A thioester	
3.1.2.27	Choloyl-CoA hydrolase	Coenzyme A thioester	
3.1.2.28	1,4-dihydroxy-2-naphthoyl-CoA hydrolase	Coenzyme A thioester	
3.1.2.29	Fluoroacetyl-CoA Thioesterase	Coenzyme A thioester	
3.1.2.30	(3S)-malyl-CoA thioesterase	Coenzyme A thioester	
3.1.2.-	Miscellaneous	Various	

Table 7: Grouping the thioesterases by Enzyme Commission number. The parent group for the thioesterases is 3.1.2. Thioesterases are divided into 27 currently active groups, as three of the original EC thioesterase subdivisions are no longer in use. 15 active thioesterase groups are specific for thioesterases that hydrolase various Acyl-CoA substrates as their main substrate. The miscellaneous category also contains thioesterases that hydrolyse Acyl-CoA substrates, which don't, however, match the more specific substrate descriptions.

The metabolic roles that acyl-CoA thioesterases play are much more varied than those of both the acyl-protein hydrolases and the glutathione thioester hydrolases. Acyl-CoA thioesterases have been implicated in the regulation of lipid metabolism by controlling the level of free fatty acids, coenzyme A and acyl-CoA thioesters in the cell [107-109], and have also been implicated in cell-membrane formation [76, 87, 110]. One of the first acyl-CoA thioesterases to be identified is mammalian thioesterase II, which has been implicated in the termination of fatty acid biosynthesis in some circumstances [111-113], a task normally performed by acyl-ACP thioesterases. Similarly, thioesterase II of *E. coli*, otherwise known as tesB, a broad specificity thioesterase, has been shown to be involved in 3-hydroxyalkanoic acid synthesis, whereby 3-hydroxyalkanoic acid moiety of 3-hydroxyalkanoyl-ACP is transferred to coenzyme A before either joining polyhydroxyalkanoate biosynthesis, or being hydrolysed to the free acid [114]. In another example, 3-hydroxyisobutyryl-CoA hydrolase (EC 3.1.2.4) in valine metabolism ensures a metabolic driving force for the hydration of the toxic methacrylyl-CoA [82], whereby the ability of coenzyme A to make the α -carbon of the acyl moiety more acidic likely enables the oxidation of isobutyryl-CoA to methacrylyl-CoA, whilst the free energy release from its hydrolysis acts as a thermodynamic driver to keep the methacrylyl-CoA concentrations to a minimum. In a further example, 4-hydroxybenzoyl-CoA thioesterase (EC 3.1.2.23) is involved in the dehalogenation of 4-chlorobenzoate, an industrial waste product, which proceeds to an activation step to 4-chlorobenzoyl-CoA, a dehalogenation step to 4-hydroxybenzoyl-CoA, and a hydrolysis step to 4-chlorobenzoate (Figure 20) [115].

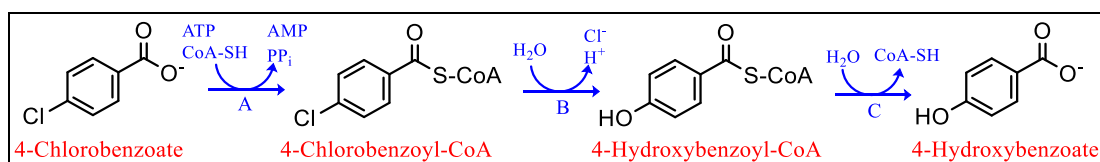


Figure 20: 4-chlorobenzoate dehalogenation pathway. A) 4-chlorobenzoyl-CoA synthetase. B) 4-chlorobenzoyl-CoA dehalogenase. C) 4-hydroxybenzoyl-CoA thioesterase.

In this example, the activation step likely assists the dehalogenation step, by enabling resonance stabilisation of the 4-hydroxy group which can more readily contribute its lone electrons to the benzoyl groups delocalised π -system, compared to the free acid. The thioesterase then hydrolyses the 4-hydroxybenzoyl-CoA to 4-hydroxybenzoate, which is then degraded *via ortho* and *meta* cleavage pathways [116]. This hydrolysis will add an additional thermodynamic driving force to the dehalogenation of 4-chlorobenzoate, especially considering that the activation step utilised an AMP-forming acyl-CoA synthetase, an irreversible step in itself due to the hydrolysis of pyrophosphate. Thus, in many cases, an acyl-CoA

thioester will have been activated to that intermediate to either enable an otherwise unfavourable metabolic step to take place, which then requires the product of that step to be hydrolysed so that the rest of the metabolic pathway can proceed, or the hydrolysis of the thioester may provide a thermodynamic driving force to a metabolic pathway, especially when coupled to a previous highly exergonic step, such as an activation step involving an AMP forming acyl-CoA synthetase.

5.1.2 Screening thioesterases for activity on methacrylyl-CoA

The proposed hydrolysis of methacrylyl-CoA to methacrylic acid and coenzyme A is not a naturally occurring process. Whilst it is already known that tesB from *E. coli* already hydrolyses isobutyryl-CoA, which is structurally similar to methacrylyl-CoA, it is not known that it does indeed also hydrolyse methacrylyl-CoA, with what selectivity over isobutyryl-CoA. Thus, candidates from the repertoire of known thioesterases must be screened for fortuitous activity on methacrylyl-CoA. However, on the UniProt database alone, there are 11233 entries in the EC groups, shown in Table 7, that pertain to thioesterases that hydrolyse esters of coenzyme A, inclusive of the miscellaneous EC 3.1.2.- group. Of these, 607 entries have been reviewed. However, this is still a very high number of enzymes to screen experimentally for fortuitous activity on methacrylyl-CoA, and so a method to narrow down the number of candidate enzymes that are likely to show fortuitous activity on methacrylyl-CoA, based on the structure of methacrylyl-CoA (Figure 21), was required.

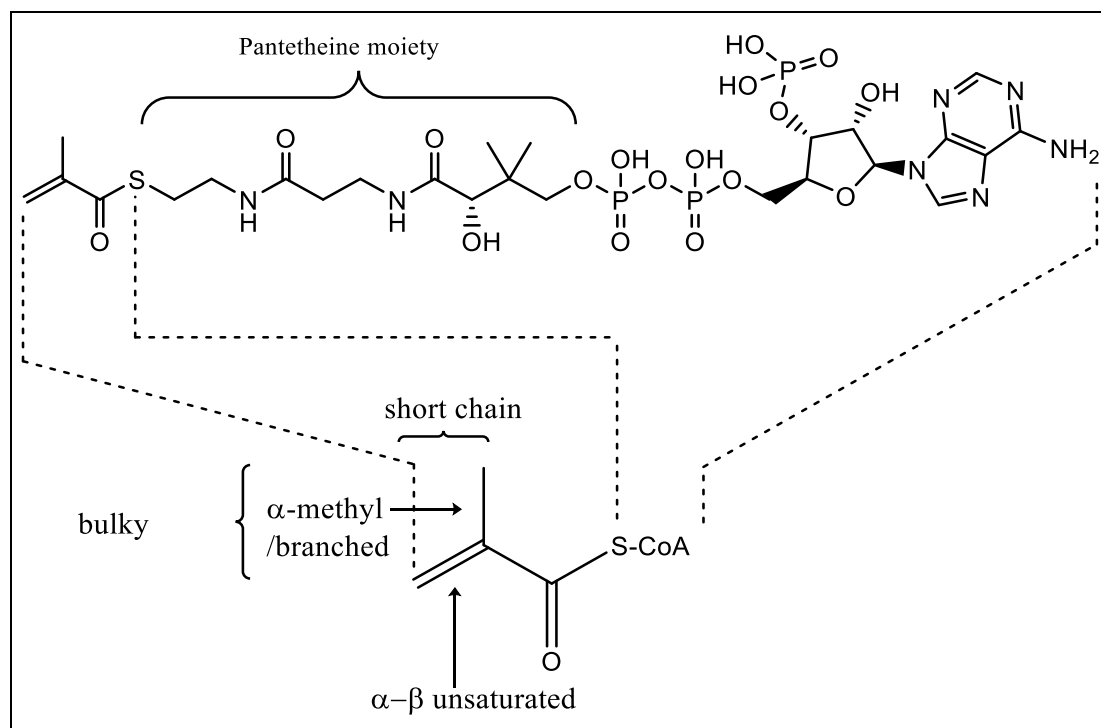


Figure 21: Structural features of methacrylyl-Coenzyme A.

The acyl group on methacrylyl-CoA is short (3-carbons), α - β -unsaturated, and with a 2-methyl branch, bulky compared to straight chained equivalents. The coenzyme A group contains a pantetheine moiety, which is also held by acyl carrier proteins.

Two approaches were initially envisaged as to how to narrow down total candidates for the hydrolysis of methacrylyl-CoA. The first was to examine the substrate binding pocket and active sites of the enzyme from its crystal structure, to ensure that firstly, the active site cavity is large enough in all three dimensions to accept the novel substrate; secondly, whether enough of the residues that stabilise the binding of the wild-type substrate stabilising residues for the wild-type substrate are positioned in such a way that they can also stabilise the binding of the novel substrate; and finally, whether all of the catalytic residues, including those which simply polarise bonds but don't necessarily participate in electron transfer, are located in positions such that they are able to participate in the hydrolysis of the novel substrate. For example, the enzyme 4-hydroxybenzoyl-CoA thioesterase, which participates in the 4-chlorobenzoate dehalogenation pathway described above (Figure 20), accepts a substrate whose acyl moiety is relatively short, compared to substrates such as palmitoyl-CoA, for palmitoyl-CoA hydrolase (EC 3.1.2.2), and is also bulky, as with methacrylyl-CoA. Indeed, the methacryloyl moiety of methacrylyl-CoA is shorter than, and no bulkier than, the 4-hydroxybenzoyl moiety of 4-hydroxybenzoyl-CoA, and so there should be no difficulty for this substrate to fit within the active site of a 4-hydroxybenzoyl-CoA thioesterase.

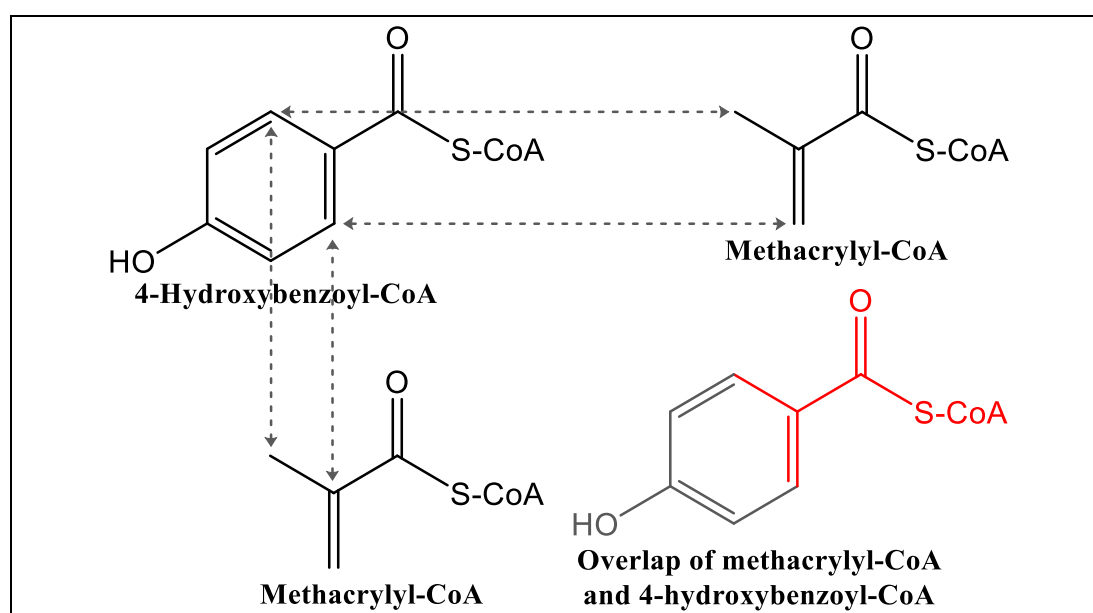


Figure 22: A comparison of the structural features of methacrylyl-CoA and 4-hydroxybenzoyl-CoA. Methacrylyl-CoA shorter than, and no bulkier than, 4-hydroxybenzoyl-CoA.

However, it is unclear whether or not the 4-hydroxy group of the acyl-moiety of 4-hydroxybenzoyl-CoA needs stabilisation, and whether the absence of such a group on methacrylyl-CoA would affect the K_M so drastically that the substrate does not bind. The crystal structure of the protein would enable molecular docking experiments and energy minimisation calculations to help ascertain whether or not methacrylyl-CoA would make a good substrate for the enzyme. A knowledge of the enzyme mechanism would also help to ascertain whether any feature of the substrate other than the thioester bond itself, participate in the reaction mechanism, and again, whether or not methacrylyl-CoA would likely make a good substrate.

However, this method relies on the presence of crystal structures for the enzymes being screened, and preferably for the catalytic mechanism to have been determined, and herein lies two issues. Firstly, of the 607 reviewed entries in the UniProt databases for acyl-CoA thioesterases, only 37 have crystal structures (Table 34, appendices, page XVII). Thus, a huge number of enzymes cannot be considered, let alone to either be pursued or excluded for wet-lab screening. Secondly, the analysis of the active site and the performance of energy minimisation calculations on the unnatural substrate is a laborious and time intensive task. Therefore, on the one hand, with only a few crystal structures and resolved mechanisms lies the problem of potential candidates being excluded for no reason other than that they cannot be analysed, whilst on the other hand, were more crystal structures to be available, an overly time-consuming shortlisting process would have ensued.

However a less sophisticated but reliable approach is to compare the structures of the substrates that are known to be hydrolysed by the enzyme with that of the target substrate. This can give an idea of the likelihood of an enzyme catalyzing the hydrolysis of a certain substrate, if another chemical with a similar chemistry has been accepted as a substrate. This relies on a thorough kinetic characterisation of the enzyme having been previously performed and documented on a range of substrates with various structural features. Therefore, a database search to identify enzymes with known activity on substrates with similar structural features to coenzyme A ought to include the BRENDA database. This is a comprehensive, dedicated and manually curated enzyme activity database, is based on published literature, and contains known kinetic data for a given enzyme. The publically accessible BRENDA database cannot be automatically queried however, and must be searched manually. By contrast, UniProt entries are not necessarily manually curated, and enzyme activities can be computationally predicted rather based on

published literature, and many entries in UniProt do not list kinetic data, even when kinetic data is publically available. However, the UniProt database interface provides advanced querying tools, the ability to filter out unreviewed entries, and the ability to download result sets for further querying in dedicated database packages such as Microsoft Access. Thus, the additional time requirement to search the UniProt database in addition to the BRENDA database is very small compared to the added benefit of combining the coverage of both the BRENDA and UniProt databases, and so to maximise the chances of identifying a suitable candidate thioesterase for the hydrolysis of methacrylyl-CoA, both databases ought to be searched.

Thus, the first objective of this experiment is to perform a bioinformatics search in the UniProt and BRENDA databases to identify candidate thioesterases for the role of hydrolyzing methacrylyl-CoA. Candidate enzymes will show proven activity on substrates that are chemically similar to methacrylyl-CoA (Figure 21), such that acyl groups of tested substrates are ideally short (~ 3 -carbons), bulky (preferably α -branched), and/or α - β unsaturated. In order to maximise the chances that these heterologous proteins would express in *E. coli*, enzymes from prokaryotic sources are preferred. Once candidate thioesterases have been identified, genes encoding them will be cloned into an expression host, overexpressed and purified for crude enzymatic assays on methacrylyl-CoA as a rapid screen for activity on methacrylyl-CoA, using a single substrate concentration. As the metabolic pathway to methacrylic acid would contain both methacrylyl-CoA and isobutyryl-CoA as intermediates, the candidate thioesterases will also be tested using isobutyryl-CoA as a substrate. Any candidate thioesterases with demonstrated selectivity for methacrylyl-CoA over isobutyryl-CoA in the crude enzyme assays will be kinetically characterised, to determine its usefulness in an *in vivo* metabolic pathway to methacrylic acid.

5.1.3 Database and literature search for thioesterases with known activity on structurally related substrates to methacrylyl-CoA

The UniProt and BRENDA databases were searched for candidate enzymes for the hydrolysis of methacrylyl-CoA. The initial search was performed on the queryable UniProt database, where the search query consisted of each of the above enzyme commission (EC) numbers, separated by 'OR' statements (Figure 23).

Figure 23: A snapshot of a UniProt database search query to yield enzymes that belong to Enzyme Commission groupings that contain Acyl-CoA thioesterases.

This search generated over 50,000 results. The search results were customized to show only the Entry Accession Number, the Protein name, gene name, organism, protein existence, and EC number, and the result set was then narrowed to just 108 enzymes by downloading the result set, importing it into Microsoft Access, and querying for proteins with evidence at protein level. This method is illustrated in the example below (Figure 24).

Entry	Protein name	Gene names	Organism	Protein existence
P32316	Acetyl-CoA hydr	ACH1 YBL015W	Saccharomyces	Evidence at protein level
Q8WXI4	Acyl-coenzyme A	ACOT11 BFIT KI	Homo sapiens (	Evidence at protein level
Q8WYK0	Acyl-coenzyme A	ACOT12 CACH C	Homo sapiens (	Evidence at protein level
Q99NB7	Acyl-coenzyme A	Acot12 Cach Ca	Rattus norvegic	Evidence at protein level
Q9NPJ3	Acyl-coenzyme A	ACOT13 THEM2	Homo sapiens (	Evidence at protein level
Q9CQR4	Acyl-coenzyme A	Acot13 Them2	Mus musculus (	Evidence at protein level
Q86TX2	Acyl-coenzyme A	ACOT1 CTE1	Homo sapiens (	Evidence at protein level
O55137	Acyl-coenzyme A	Acot1 Cte1	Mus musculus (	Evidence at protein level
O88267	Acyl-coenzyme A	Acot1 Cte1	Rattus norvegic	Evidence at protein level

Figure 24: A snapshot of the results of a Microsoft Access query to show only Acyl-CoA Hydrolase enzymes that have been proven to exist at protein level.

As one of the requirements was that these proteins needed to be from a prokaryotic source, it was possible to further narrow down the results by looking only at proteins whose host organism is prokaryotic. There are 27 putative and characterized, coenzyme A derivative thioester hydrolases of prokaryotic origin that have been expressed as proteins, listed on the UniProt database (Table 8).

Entry	Protein names
Q04416	4-hydroxybenzoyl-CoA thioesterase (4-HBA-CoA thioesterase) (EC 3.1.2.23)
P56653	4-hydroxybenzoyl-CoA thioesterase (EC 3.1.2.23)
P19413	Bile acid-CoA hydrolase (EC 3.1.2.26) (Bile acid-coenzyme A hydrolase) (Bile acid-inducible operon protein F)
Q7V4A7	1,4-dihydroxy-2-naphthoyl-CoA hydrolase (DHNA-CoA hydrolase) (EC 3.1.2.28) (DHNA-CoA thioesterase)
Q55777	1,4-dihydroxy-2-naphthoyl-CoA hydrolase (DHNA-CoA hydrolase) (EC 3.1.2.28) (DHNA-CoA thioesterase)
P0A8Y8	Proofreading thioesterase EntH (EC 3.1.2.-) (Enterobactin synthase component H) (p15)
Q88QS0	Esterase EstP (EC 3.1.1.1) (Autotransporter esterase EstP) (Palmitoyl-CoA hydrolase) (EC 3.1.2.2)
P77712	Long-chain acyl-CoA thioesterase FadM (EC 3.1.2.-) (Acyl-CoA thioester hydrolase) (Thioesterase 3)
P0AC84	Hydroxyacylglutathione hydrolase (EC 3.1.2.6) (Glyoxalase II) (Glx II)
Q8ZRM2	Hydroxyacylglutathione hydrolase (EC 3.1.2.6) (Glyoxalase II) (Glx II)
Q93MW7	Lipase (EC 3.1.1.3) (Diheptanoyl glycerophosphocholine esterase) (EC 3.1.1.-) (Extracellular lipase) (GDSL-like lipase) (Palmitoyl-CoA hydrolase) (EC 3.1.2.2) (SRL)
D3JV05	(3S)-maly-CoA thioesterase (EC 3.1.2.n2) ((3S)-maly-CoA thioesterase)
P76084	Acyl-coenzyme A thioesterase PaaI (EC 3.1.2.-) (Phenylacetic acid degradation protein PaaI)
P51025	S-formylglutathione hydrolase frmB (FGH) (EC 3.1.2.12)
P33018	S-formylglutathione hydrolase yeiG (FGH) (EC 3.1.2.12)
P44556	S-formylglutathione hydrolase (FGH) (EC 3.1.2.12)
A1AXZ2	S-formylglutathione hydrolase (FGH) (EC 3.1.2.12)
Q08788	Surfactin synthase thioesterase subunit (EC 3.1.2.-) (Cold shock protein CSI16)
P0ADA1	Acyl-CoA thioesterase I (EC 3.1.2.-) (Lecithinase B) (Lysophospholipase L1) (EC 3.1.1.5) (Protease I)
P0AGG2	Acyl-CoA thioesterase 2 (EC 3.1.2.-) (Acyl-CoA thioesterase II) (TEII)
P45083	Putative esterase HI_1161 (EC 3.1.2.-)
P95162	Putative esterase Rv1847/MT1895 (EC 3.1.2.-)
P94842	Uncharacterized protein HP_0496 (EC 3.1.2.-)
P44886	Uncharacterized acyl-CoA thioester hydrolase HI_0827 (EC 3.1.2.-)
P0A8Z3	Acyl-CoA thioester hydrolase YbgC (Acyl-CoA thioesterase) (EC 3.1.2.-)
P46130	Putative acyl-CoA thioester hydrolase ybhC (EC 3.1.2.-)
Q81EE4	Acyl-CoA hydrolase (EC 3.1.2.20)

Table 8: A list of putative and characterized, Acyl-CoA hydrolases of prokaryotic origin, that have been expressed as proteins, listed on the UniProt database

Any kinetic data associated with the enzymes listed above (Table 8) on the UniProt database was then analysed to determine whether any of the listed enzymes hydrolysed substrates with chemical similarity to methacrylyl-CoA. The BRENDA database, which is based entirely on the published literature, was manually searched. Each page corresponding to acyl-CoA thioesterase EC groups was visually inspected for substrates with structural similarity to methacrylyl-CoA, and shortlisted enzymes were further narrowed by only selecting enzymes from prokaryotic sources. In total, four enzymes from the UniProt and BRENDA databases were shortlisted for screening (Table 9), though the same search performed now may well yield more candidate thioesterase enzymes as more enzymes are added to each database with time.

Enzyme Name	Organism	Function	Pdb ID	Database identified from
4Hydroxybenzoyl-CoA thioesterase (4HBT)	Arthrobacter sp. strain SU	Xenobiotic degradation	1Q4T 1Q4S	UniProt
Fatty Acyl-CoA thioesterase (FcoT)	Mycobacterium tuberculosis	Synthesis of PDIMs	2PFC	BRENDA
YciA	Haemophilus influenzae	Perceived involvement in cell membrane synthesis	1YLI	BRENDA
RifR	Amycolatopsis mediterranei	Rifamycin synthesis housekeeping	3FLA	BRENDA

Table 9: A summary table of candidate thioesterases for the hydrolysis of Methacrylyl-CoA, identified by substrate similarity, from the UniProt and BRENDA databases.

Any literature associated with these enzymes was analysed for useful information which may make any of the above enzymes particularly interesting candidates, and these are summarised overleaf, and the following four subsections provide more detail on each of the thioesterases selected, detailing why they were shortlisted as candidates for screening.

5.1.3.1 4-hydroxybenzoyl-CoA thioesterase from *Arthrobacter* sp. strain SU

The 4-hydroxybenzoyl-CoA thioesterase (4HBT) enzyme from *Arthrobacter* sp. strain SU was also identified during the bioinformatics search, based on its activity on substrates that are chemically similar to methacrylyl-CoA. The enzyme is, as described previously, present in a bioremediation pathway that degrades chlorobenzoates that have been introduced into the environment since the start of the industrial revolution [117]. The enzyme is a homotetramer and its crystal structure is available, with protein databank IDs 1Q4S and 1Q4T [118]. The component polypeptide (Polypeptide sequence 1, appendices, page I) is 16.4 kDa. Enzyme characterization studies by Zhuang *et al.* [117] found that enzyme had reasonably broad specificity (Table 10).

Substrate	Mean k_{cat} (s^{-1}) ±SD	Mean K_m (μM) ±SD	K_{cat}/K_m ($M^{-1}s^{-1}$)
Benzoyl-CoA	1.86 ± 0.06	(2.43±0.16)×10 ²	7.6×10 ³
4-hydroxybenzoyl-CoA	6.7 ± 0.1	1.24±0.06	5.4×10 ⁶
3-hydroxybenzoyl-CoA	15.8 ± 0.2	(2.2±0.1)×10 ¹	7.2×10 ⁵
2,5-dihydroxybenzoyl-CoA	(8.0±0.1)×10 ⁻²	(2.9±0.3)×10 ¹	2.7×10 ³
4-chlorobenzoyl-CoA	(2.5±0.1)×10 ⁻¹	(1.13±0.08)×10 ²	2.2×10 ³
Phenylacetyl-CoA	(9.0±0.2)×10 ⁻⁴	(1.4±0.1)×10 ²	6.5
Acetyl-CoA	(1.08±0.02)×10 ⁻²	(2.2±0.1)×10 ³	5.0
Propionyl-CoA	(2.6 ± 0.1)×10 ⁻²	(1.1±0.09)×10 ³	2.4×10 ¹
Butyryl-CoA	(3.19±0.05)×10 ⁻²	(9.3±0.5)×10 ²	3.4×10 ¹
Crotonyl-CoA	(7.5 ± 0.1)×10 ⁻²	(8.1±0.3)×10 ²	9.2×10 ¹
Octanoyl-CoA	7.3 ×10 ⁻⁴	ND	-

Table 10: Summary of enzyme kinetics of 4HBT from *Arthrobacter* sp. strain SU on various substrates. Adapted from Zhuang *et al.* [117].

The enzyme shows most activity on substrates with a benzene ring, but also shows activity on short, straight chain (propionyl-CoA) and α - β -unsaturated (crotonyl-CoA) substrates. Greater activity on bulkier substrates that contain bulky benzoyl-derivative groups provides confidence that the enzyme ought to hydrolyse branched substrates. This known activity on short, α - β -unsaturated and bulky substrates lends hope that this enzyme may show activity on methacrylyl-CoA. There is also some evidence of preference towards α - β -unsaturated substrates over their saturated analogues, with the catalytic efficiency on crotonyl-CoA (92 $M^{-1}.s^{-1}$) being almost three times that of its saturated analogue, butyryl-CoA (34 $M^{-1}.s^{-1}$).

5.1.3.2 Fatty Acyl-CoA thioesterase

Fatty acyl-CoA thioesterase (FcoT) from *Mycobacterium tuberculosis* is a homotrimer of homodimers, and therefore a homohexamer [119]. FcoT is involved in the synthesis of pthiocerol dimycocerosates (PDIMs) which are important for the survival of *Mycobacterium tuberculosis* in macrophages [120]. The crystal structure of FcoT is available, and the protein databank ID is 2PFC. The component polypeptide (Polypeptide sequence 2) is 20.5 kDa.

Enzyme characterisation studies performed by Wang *et al.* [119], the results of which are summarised below (Table 11), showed that the enzyme had preference for long and narrow acyl-CoAs, with mildly better activity on saturated acyl chains than unsaturated acyl chains. Nevertheless, the enzyme does have activity on benzoyl-CoA which is a comparatively short and bulky substrate, as well as crotonyl-CoA, a short, α - β unsaturated acyl-CoA molecule, though the turnover numbers are comparatively low. The k_{cat} of 4HBT for crotonyl-CoA, for example, is 10 times greater than that of FcoT. However, all the Michaelis constants are very low for the substrates, which could compensate in a metabolic pathway whereby the concentrations of substrates may not be otherwise saturating.

Substrate	k_{cat} (s^{-1})	K_M (μM)	K_M (M)	k_{cat}/K_M ($M^{-1}.s^{-1}$)
Crotonyl-CoA	7.8×10^{-3}	87.35	8.7×10^{-5}	89
Dodecanoyl-CoA	1×10^{-2}	19.13	1.9×10^{-5}	522
Hexanoyl-CoA	1.2×10^{-2}	63.18	6.3×10^{-5}	190
Lauroyl-CoA	1.5×10^{-2}	38.52	3.9×10^{-5}	389
Palmitoyl-CoA	3.7×10^{-2}	11.87	1.2×10^{-5}	3117
Stearoyl-CoA	1.9×10^{-2}	10.28	1.0×10^{-5}	1848
Benzyl-CoA	4.0×10^{-3}	272.13	2.7×10^{-4}	15

Table 11: Summary of enzyme characterization results for FcoT on seven different substrates [119].

What is most appealing about this enzyme, however, is that the substrate binding pocket (Figure 25) has numerous potential sites for directed evolution, which may make this an interesting enzyme from the point of introducing or improving hydrolytic activity on methacrylyl-CoA.

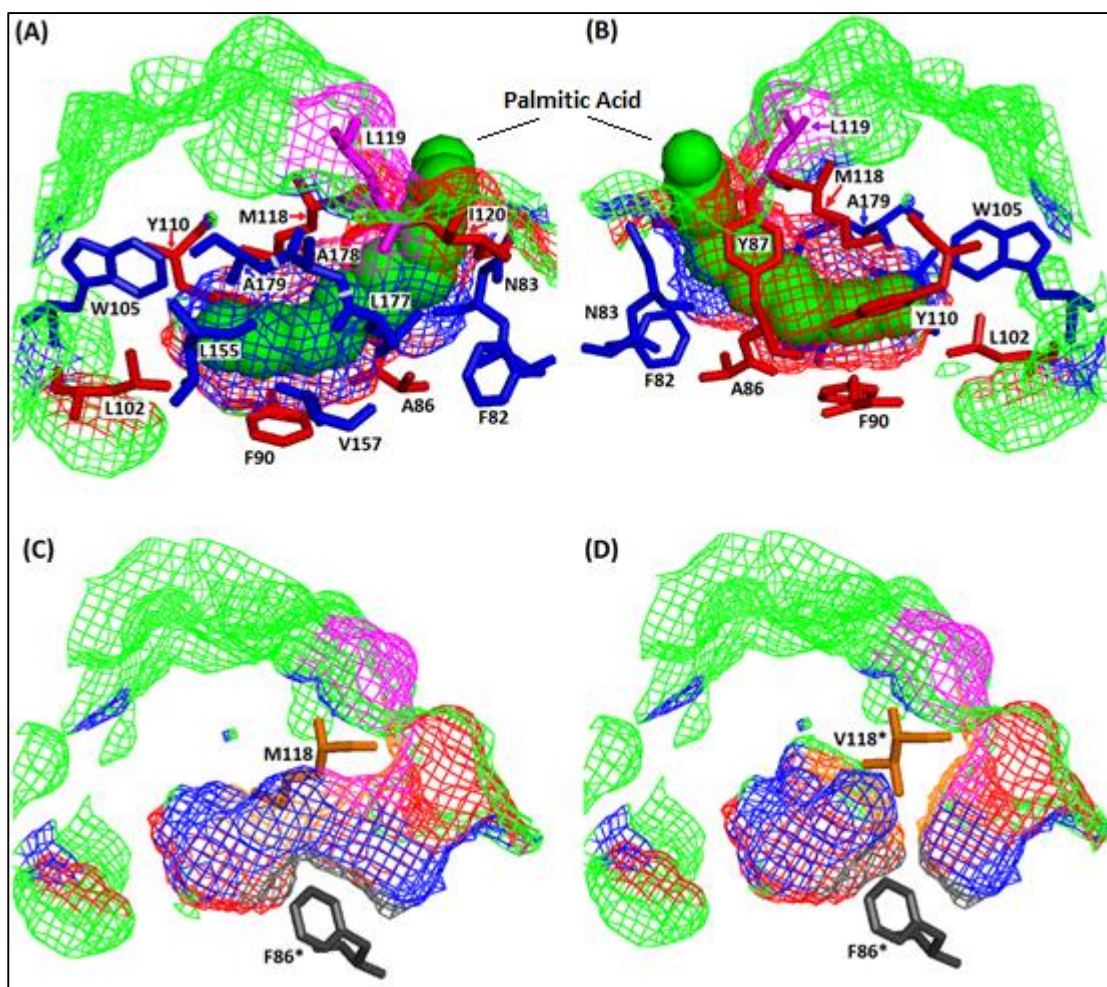


Figure 25: The substrate binding site of FcoT, from two opposing views (A&B), showing residues described by Wang *et al.* to contribute to the substrate binding site in red, and further residues identified during a further analysis of the crystal structure during this project to contribute to the substrate binding site, in blue. An A86F mutation could constrict the substrate binding site (C), and a further M118V mutation could then fully shorten the substrate binding site (D).

Mutation studies by Wang *et al.* have identified 3 residues crucial to the catalytic activity of this thioesterase: Y87, E77 and H72, which would help in directed evolution exercises if it is necessary. In addition, a further study of the crystal structure, performed during this project, shows that an A86F mutation could constrict the substrate binding site (Figure 25C), and a further M118V mutation could then fully shorten the substrate binding site (Figure 25D).

5.1.3.3 YciA from *Haemophilus influenzae*

YciA from *Haemophilus influenzae* is a homohexamer (trimer of dimers) and has a perceived role in the formation of biomembranes [121]. The crystal structure of YciA is known, and that of the homodimer is provided in the protein databank (PDB ID 1YLI) [122]. The single polypeptide (Polypeptide sequence 3) weighs 16.8 kDa. YciA is a very broad specificity enzyme that can hydrolyse many acyl-CoA substrates (Table 12), including some with similarity to methacrylyl-CoA.

Substrate	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($M^{-1}.s^{-1}$)
acetyl-CoA	1.3 ± 0.1	4.7 ± 0.9	2.7×10^5
propionyl-CoA	25 ± 10	6 ± 1	4.2×10^6
propionyl-s-N-acetylcysteamine	<0.0001	-	-
n-butyryl-CoA	22 ± 1	3.5 ± 0.2	6.2×10^6
n-hexanoyl-CoA	15 ± 1	4.1 ± 0.5	3.6×10^6
n-octanoyl-CoA	4.6 ± 0.2	1.9 ± 0.3	2.4×10^6
n-decanoyl-CoA	13 ± 1	4.5 ± 0.5	2.8×10^6
lauroyl-CoA	9.2 ± 0.3	6.6 ± 0.7	1.4×10^6
myristoyl-CoA	6.9 ± 0.1	2.5 ± 0.1	2.8×10^6
palmitoyl-CoA	2.7 ± 0.1	17 ± 1	1.6×10^5
stearoyl-CoA	0.77 ± 0.02	10 ± 1	7.5×10^4
arachidoyl-CoA	0.27 ± 0.01	39 ± 1	6.9×10^3
palmitoleoyl-CoA	5.7 ± 0.2	6.4 ± 0.8	8.9×10^5
oleoyl-CoA	1.7 ± 0.1	11 ± 1	1.5×10^5
linoleoyl-CoA	2.0 ± 0.1	4.2 ± 0.1	4.7×10^5
D,L- γ -hydroxybutyryl-CoA	8.3 ± 0.2	31 ± 2	2.7×10^5
isobutyryl-CoA	140 ± 10	13 ± 1	1.1×10^7
n-butyryl-pantethienephosphate	0.8 ± 0.02	1400 ± 100	5.7×10^2
isovaleryl-CoA	30 ± 1	14 ± 2	2.1×10^6
tiglyl-CoA	2.6 ± 0.1	2.1 ± 0.2	1.3×10^6
β -methylcrotonyl-CoA	0.54 ± 0.03	9 ± 1	6.0×10^4
crotonyl-CoA	0.92 ± 0.03	5.9 ± 0.4	1.5×10^5
methylmalonyl-CoA	1.1 ± 0.1	170 ± 0.1	6.3×10^3
malonyl-CoA	3.3 ± 0.2	140 ± 20	2.4×10^4
succinyl-CoA	190 ± 10	20 ± 3	9.5×10^6
glutaryl-CoA	100 ± 10	8.1 ± 0.6	1.3×10^7
D,L-3-hydroxy-3-methylglutaryl-CoA	20 ± 2	4.5×10^5	-
acetoacetyl-CoA	35 ± 1	37 ± 2	9.4×10^5
4-HBA-CoA	0.078 ± 0.003	9.1 ± 0.9	8.6×10^3
phenylacetyl-CoA	130 ± 10	1.6 ± 0.1	8.3×10^7
4-hydroxyphenylacetyl-CoA	75 ± 3	7.6 ± 0.8	9.9×10^6
3-hydroxyphenylacetyl-CoA	140 ± 0.10	3.7 ± 0.5	3.7×10^7

Table 12: Substrates that are hydrolysed by YciA, including kinetic data of the hydrolysis reaction

Of all candidate thioesterases, YciA is perhaps the most likely to hydrolyse methacrylyl-CoA by reason of its exceptionally broad substrate range, including with short, α - β -unsaturated and α -methyl branched chain substrates. However, the enzyme is most active on isobutyryl-CoA, a key intermediate in the proposed pathway to methacrylic acid and may therefore hydrolyse isobutyryl-CoA before it can be converted to methacrylyl-CoA for its hydrolysis of methacrylic acid. The preference, if any, for hydrolysis of methacrylyl-CoA over isobutyryl-CoA must therefore be tested.

5.1.3.4 RifR from *Amycolatopsis mediterranei* S699

RifR is a homodimer from *Amycolatopsis mediterranei* S699, and is a type II thioesterase with a housekeeping role in rifamycin synthesis [123]. Acyl-ACPs are its main substrates, but the enzyme has shown activity on Acyl-CoAs. The RifR monomer (Polypeptide sequence 4) is 28.2 kDa. Of particular interest with RifR, is a single mutation, S49A, in its component polypeptide that significantly narrows the substrate specificity of RifR to isobutyryl-CoA, which is structurally very similar to methacrylyl-CoA.

	Substrate	k_{cat}/K_M / $M^{-1}s^{-1}$	
		Wild Type RifR	S49A RifR
Coenzyme substrates	Decanoyl-CoA	160 ± 18	<0.16
	Octanoyl-CoA	31 ± 2.5	<0.16
	Propionyl-CoA	25 ± 0.5	0.96±0.37
	Butyryl-CoA	13 ± 3.2	<0.04
	Acetyl-CoA	11 ± 0.2	<0.03
	Isobutyryl-CoA	9.6 ± 0.08	4.5±0.08
	Hexanoyl-CoA	5.9 ± 0.33	<0.04
	Methylmalonyl-CoA	1.8 ± 0.17	<0.03
	Malonyl-CoA	1.5 ± 0.08	<0.07
Acyl-ACP substrates	Propionyl-RifM1	210 ± 20	-
	Acetyl-RifM1	150 ± 38	-
	Methylmalonyl-RifM1	54 ± 6.3	-

Table 13 : Summary of substrates that wild type and S49A mutant RifR hydrolyse, with kinetic data for hydrolysis reaction.

Though the catalytic efficiency of this enzyme on isobutyryl-CoA is incredibly low compared to that of YciA, it is possible that with the overexpression of this enzyme, sufficient activity on methacrylyl-CoA may be observed for its use in a proof of concept pathway to methacrylic acid.

5.1.4 Sourcing the genes, gene cloning and gene expression

Four enzymes, 4HBT, FcoT, YciA and RifR were identified as candidate thioesterases for the hydrolysis of methacrylyl-CoA to methacrylic acid and coenzyme A. In order to test these enzymes for such novel activity on methacrylyl-CoA, they needed to be overexpressed. This required that genes encoding these enzymes be cloned into expression vectors. For their kinetic characterisation, 4HBT and YciA had both been previously expressed using the pET23b vector [117, 121], FcoT using the pET28b vector [119], and RifR with the pET21 vector [123]. As it was impractical to use three different expression vectors for an enzyme screening activity, the pET20b(+) expression vector, a very similar expression vector to the above, was used instead for all of the above enzymes.

The required genes encoding 4HBT, FcoT, YciA and RifR were chemically synthesized and not cloned from genomic DNA, as some of the organisms from which the genes originate, *Mycobacterium tuberculosis* for FcoT, for example, posed biological hazards, and the cost of custom gene synthesis was comparable with the cost of cloning from genomic DNA. Furthermore, custom synthesis permitted codon optimisation of the genes for expression in *E. coli*, which ought to ensure maximal expression of an enzyme. This is useful when cloning enzymes to test for activity on novel substrates, as high levels of enzyme are desirable.

Genes encoding proteins 4HBT, FcoT, YciA and RifR were therefore optimized for overexpression in *E. coli*, by BioMatik Corporation USA. Sequences 5' -TATACAT-3' and 5' -TAAGCGGCCGCACTC-3' were appended to the 5' and 3' ends of each codon optimised genes respectively. The 5' appended sequence contained a 5' -CAT-3' sequence, such that together with the 5' -ATG-3' start codon of the gene, an *NdeI* restriction site (5' -CATATG-3') was included. The 3' appended sequence contained a 5' -TAA-3' stop codon, followed immediately by a *NotI* restriction site (underlined). The *NdeI* and *NotI* restriction sites were included in the gene sequences to enable cloning into the pET20b(+) vector. *NdeI* was specifically chosen as a cloning site so that the start codon of the gene was placed 8bp downstream from the ribosome binding site (RBS) of the pET20b(+) vector, well within the optimal distance range from the RBS [124]. Thus, the polynucleotides containing the codon optimised genes for 4HBT (Polynucleotide sequence 1), FcoT (Polynucleotide sequence 2), YciA (Polynucleotide sequence 3) and RifR (Polynucleotide sequence 4), which included the 5' and 3' appended sequences, were synthesised by Biomatik, and were each delivered in a pBMH cloning vectors, a derivative of the pBluescript(II)SK+ vector. The inserts from each cloning vector

were then subcloned from their cloning plasmids into the pET20b(+) expression vector to form pET20b(+):4HBT, pET20b(+):FcoT, pET20b(+):YciA and pET20b(+):RifR. In turn, the expression plasmids were then transformed into the expression host *E. coli* BL21 (DE3) pLysS (Table 14).

Cloning plasmid	Expression plasmid	Expression strain	Strain ID
pBMH::4HBT	pET20b(+):4HBT	<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):4HBT	Strain 9
pBMH::FcoT	pET20b(+):FcoT	<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):FcoT	Strain 10
pBMH::YciA	pET20b(+):YciA	<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):YciA	Strain 11
pBMH::RifR	pET20b(+):RifR	<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):RifR	Strain 12

Table 14: Table of thioesterase expression strains.

Cryostocks were prepared from two transformant colonies for each strain listed above, and the presence of the pET20b(+) derived expression plasmids was confirmed in each cryostock by colony PCR (Figure 26).

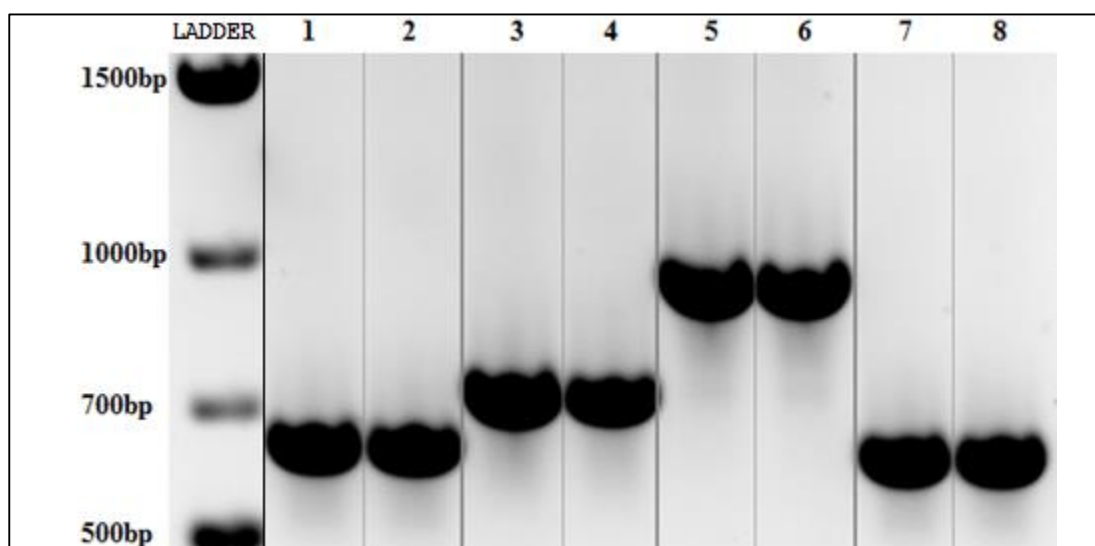


Figure 26: Confirmation of the presence of the correct inserts in duplicated cryostocks of thioesterase expression strains. Expression hosts: *E. coli* BL21 (DE3) pLysS. Transformed plasmids: pET20b(+) containing gene inserts for 4HBT (lanes 1&2), FcoT (lanes 3&4), RifR (lanes 5&6) or YciA (lanes 7&8), cloned at the *Nde*I and *Not*I restriction sites. Ladder is Fermentas GeneRuler 1kb Plus DNA ladder.

Expression of each protein was then tested for at 100 ml scale under various conditions (Table 35, appendices, page XVIII), with glucose in the growth media to prevent leaky expression prior to induction. RifR did not appear to express at all under any of the conditions tested. For 4HBT, FcoT and YciA, there was no noticeable difference in either total expression levels or soluble protein levels between the various conditions. 4HBT, FcoT and YciA all expressed well, with 4HBT and YciA present predominantly in the soluble fraction, and FcoT predominantly in the insoluble fraction, and so expression for these proteins was scaled to 500 ml scale using the conditions summarised below (Table 15). These are the same

conditions as those used in the literature for the respective enzymes. A culture of *E. coli* BL21 (DE3) pLysS containing an empty pET20b(+) vector was also grown as a negative control strain.

Expression Strain	Incubation temperature	Induction point (OD ₆₀₀)	Literature reference
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):4HBT	37 °C	1.0	[117]
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):FcoT	32 °C	0.8	[119]
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):YciA	37 °C	0.8	[121]
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+) (control)	37 °C	0.8	N/A

Table 15: Expression conditions for the preparation of thioesterase cell free extracts.

The 500ml scale cultures were incubated for 5h post induction, after which samples were taken for analysis by SDS-PAGE to confirm the expression patterns observed during small scale cultures expression (Figure 27). Cells were then harvested by centrifugation, cell free extracts were prepared for each culture and the total protein concentrations of those cell free extracts determined (Table 36, appendices, page XIX) using bovine serum albumin as the protein standard.

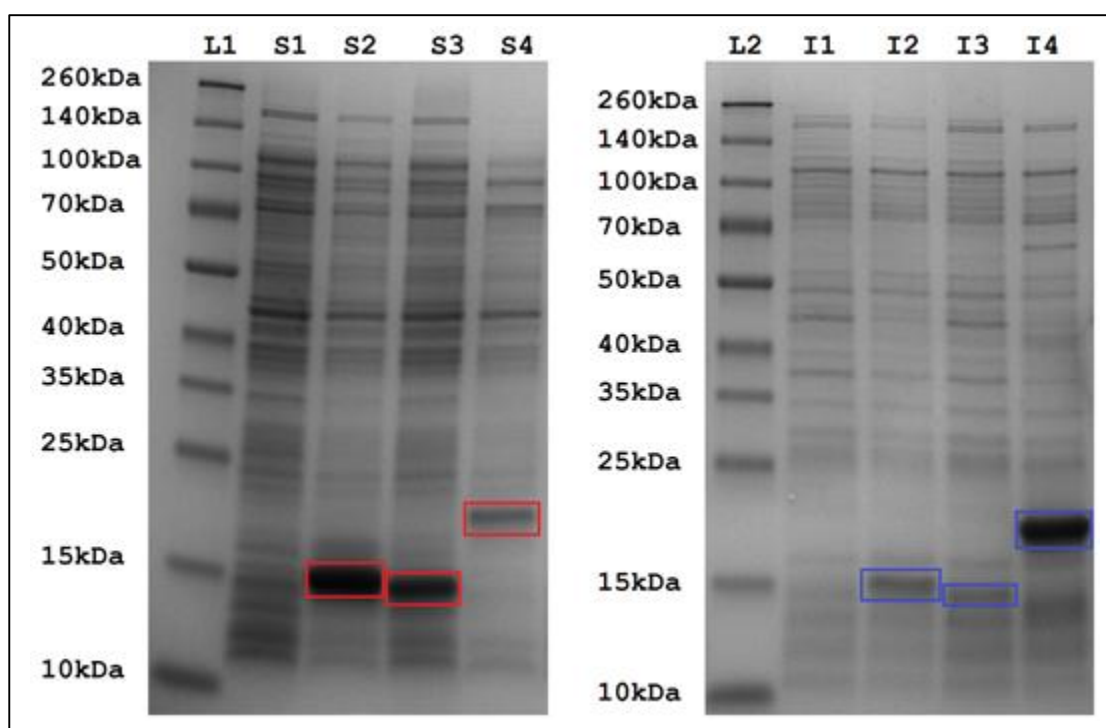


Figure 27: Expression of thioesterases. L1: Spectra BR protein ladder. S1: Control strain soluble fraction. S2: 4HBT soluble. S3: YciA soluble. S4: FcoT soluble. L2: Spectra BR protein ladder. I1: Control strain insoluble fraction I2: 4HBT insoluble. I3: YciA insoluble. I4: FcoT insoluble.

5.1.5 Preparation of methacrylyl-CoA for crude enzyme assays

Methacrylyl-CoA was synthesised by reacting methacrylic anhydride and coenzyme A, using methods described previously [125] with some modifications that had been suggested (Wataru Mizunashi, Mitsubishi Chemical Group, personal

communication). The synthesised methacrylyl-CoA was then purified by HPLC and lyophilised, as described in the materials and methods. The purity was determined by HPLC (Figure 28).

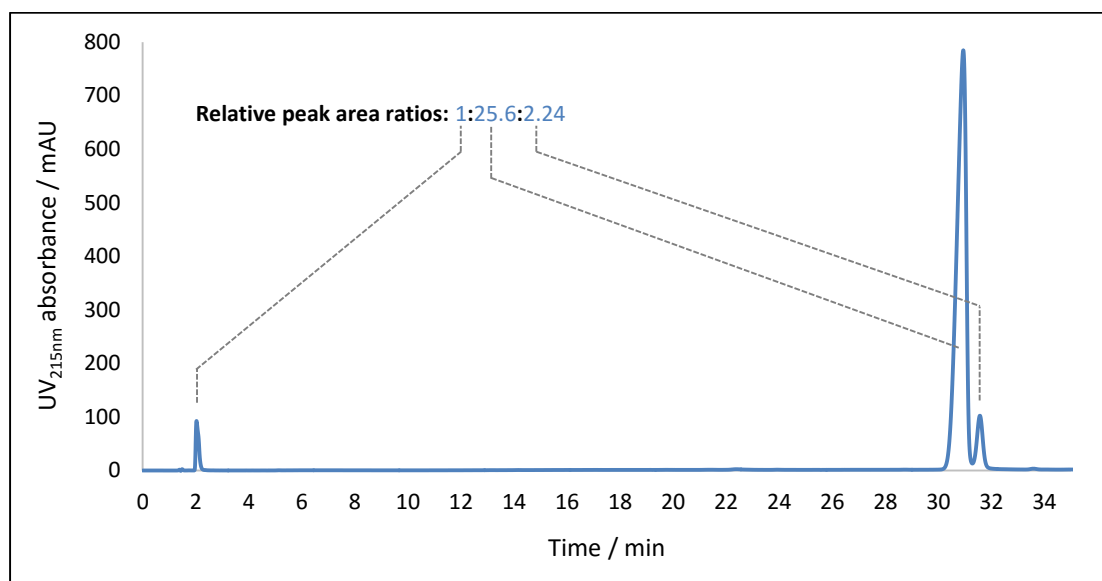


Figure 28: HPLC trace of purified methacrylyl-CoA. Method: Gradient elution. Starting conditions 1.8% acetonitrile in 75 mM KH_2PO_4 acidified to pH 2.5, rising to 13.5% acetonitrile over 40 min.

The major peak observed at 31 min was methacrylyl-CoA. The small tail peak at 31.6min is most likely associated with methacrylyl-CoA, as a coenzyme A standard (not shown) had a similar tail peak. The small peak at 2min could not be explained.

5.1.6 Crude enzyme assay

Crude enzyme assays using the soluble fraction of the cell free extracts of 4HBT, YciA and FcoT as the enzyme source were performed to rapidly screen the thioesterases for activity on methacrylyl-CoA. As isobutyryl-CoA was the precursor for methacrylyl-CoA, any activity on isobutyryl-CoA would consume this intermediate unproductively. Thus, enzymes were also tested for activity on isobutyryl-CoA to ensure that the thioesterases were selective for methacrylyl-CoA and not isobutyryl-CoA. The activity assay was based on Ellman's reagent (5'5'-dithio-bis-(2-nitrobenzoic acid)) (DTNB) which acted as an indicator for the reaction progression. On hydrolysis of the acyl-CoA substrate, the free sulfhydryl group of the liberated coenzyme A moiety reacts with DTNB to form thio-nitrobenzoic acid (TNB) and a mixed disulphide (Figure 29). TNB has a molar extinction coefficient of $14150 \text{ M}^{-1}.\text{cm}^{-1}$ at 412 nm compared a molar extinction coefficient of $211 \text{ M}^{-1}.\text{cm}^{-1}$ and $210 \text{ M}^{-1}.\text{cm}^{-1}$ for DTNB and the mixed disulphide, respectively, and so the reaction can be monitored spectrophotometrically [126].

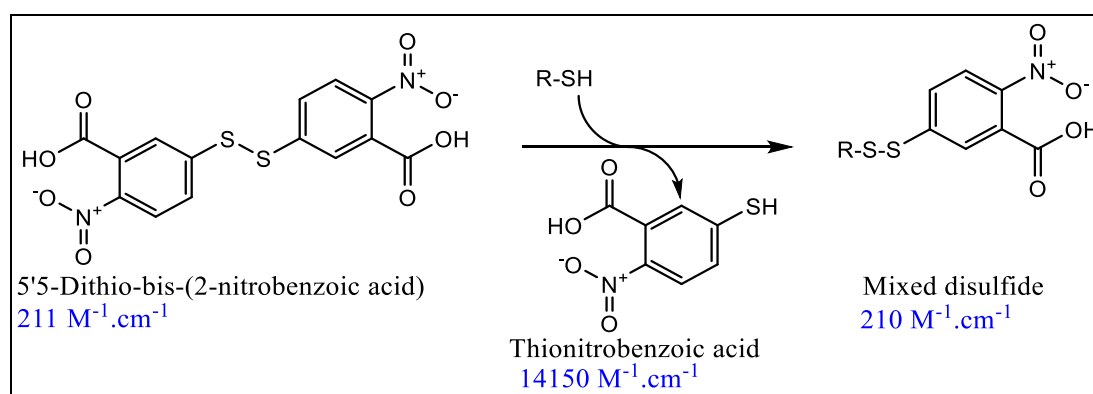


Figure 29: Reaction scheme for Ellman's reagent as indicator for acyl-CoA hydrolysis.

The assays (Figure 30) showed cell free extracts of 4HBT to preferentially hydrolyse methacrylyl-CoA over isobutyryl-CoA, with initial rates of $54 \text{ nmol}.\text{min}^{-1}.\text{mg}^{-1}$ for methacrylyl-CoA and $2 \text{ nmol}.\text{min}^{-1}.\text{mg}^{-1}$ for isobutyryl-CoA (Table 16). Cell free extracts of YciA hydrolysed both methacrylyl-CoA ($54 \text{ nmol}.\text{min}^{-1}.\text{mg}^{-1}$) and isobutyryl-CoA ($40 \text{ nmol}.\text{min}^{-1}.\text{mg}^{-1}$) well, whilst cell free extracts of FcoT were predominantly inactive on both isobutyryl-CoA and methacrylyl-CoA. Cell free extracts *E. coli* with no overexpressed protein showed significant background activity on isobutyryl-CoA ($19 \text{ nmol}.\text{min}^{-1}.\text{mg}^{-1}$), which corroborates similar observations in the literature [16, 127]. It is quite striking that cell free extracts of the *E. coli* BL21 (DE3) pLysS pET20b(+) control strain exhibited higher activity on isobutyryl-CoA than the cell free extracts of 4HBT or FcoT did, considering that the same enzymes that are responsible for this activity in the *E. coli* control ought to still be expressed in the cell free extracts of 4HBT and YciA. One obvious explanation

for why the extracts of 4HBT did not exhibit high isobutyryl-CoA activity is that it is a highly soluble enzyme that expressed very well (Figure 27), and as the total protein concentrations that were used for the 4HBT cell free assay and the *E. coli* cell free assay were the same, there would be significantly less of the *E. coli* enzyme responsible for the hydrolysis of isobutyryl-CoA in the 4HBT assay compared to in the control.

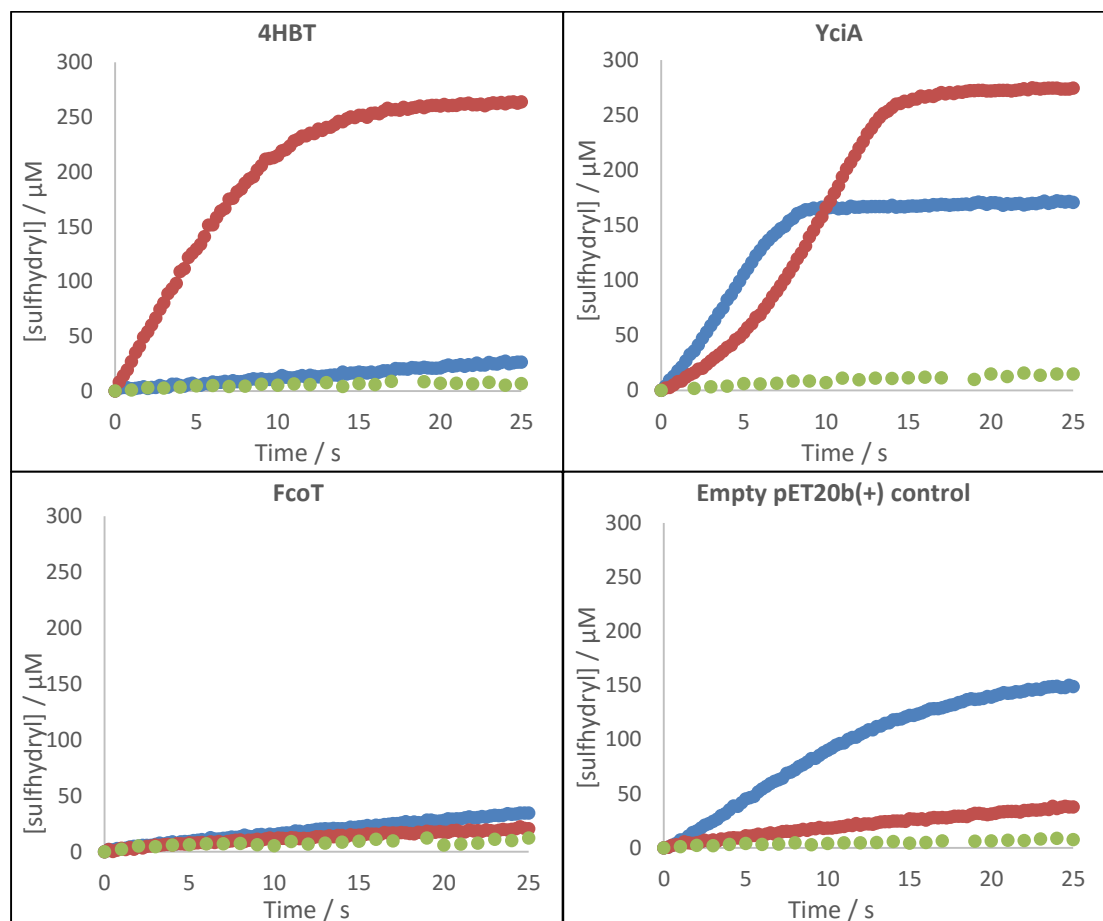


Figure 30: Crude thioesterase assays using cell free extracts of *E. coli* BL21 (DE3) pLysS pET20b(+):4HBT, *E. coli* BL21 (DE3) pLysS pET20b(+):YciA, *E. coli* BL21 (DE3) pLysS pET20b(+):FcoT and *E. coli* BL21 (DE3) pLysS pET20b(+) negative control. Red: Methacrylyl-CoA; Blue: Isobutyryl-CoA; Green: No substrate control. Assays were performed in triplicate and averages are shown on graphs. All crude enzyme concentrations were normalised at 0.5mg.ml⁻¹ total protein content.

Enzyme expressed by <i>E. coli</i> BL21 (DE3) pLysS chassis	Specific rate of sulfhydryl release (nmol.min ⁻¹ .mg ⁻¹ total protein content)	
	Isobutyryl-CoA as substrate	Methacrylyl-CoA as substrate
4HBT	2.0	54
YciA	40	17 (low), 54 (high)
FcoT	2.4	1.2
None (pET20b(+)) negative control strain	19	3.6

Table 16: Initial activity rates for crude thioesterase assays

It is harder to apply this theory to the observation that cell free extracts of FcoT, which was predominantly insoluble with very little protein in the soluble fraction,

did not hydrolyse isobutyryl-CoA to the same extent that the cell free extracts of the *E. coli* control did, considering that only the soluble fractions of each protein were used. It is possible that the response of FcoT to the protein concentration assay (DC assay) was on average stronger than the other proteins in *E. coli*'s soluble fraction, meaning that the true total protein concentration in the cell free extract was lower than that indicated by the protein assay which used BSA as the standard. The fact that the specific activity of the FcoT cell free extract on methacrylyl-CoA was three times lower than that of the *E. coli* control cell free extract, adds some weight to this idea. Perhaps a combination of the two theories may explain this observation.

Two further observations have been made. Firstly, though a starting concentration of 200 μM for both methacrylyl-CoA and isobutyryl-CoA was supposedly used for all assays, it is clear from the 4HBT and the YciA catalysed reactions, which appear to have proceeded to completion, that the final sulfhydryl concentration is approximately 260 μM for when methacrylyl-CoA was used as a substrate. Likewise, from the YciA assay, when isobutyryl-CoA was used as a substrate, the final sulfhydryl concentration is approximately 160 μM . Whilst it is conceivable that the hydrolysis of isobutyryl-CoA stops before reaching completion, it is also possible that the isobutyryl-CoA, which was commercially purchased, was not 100% pure by weight, or was incompletely redissolved. This concentration of isobutyryl-CoA should therefore be considered as a minimum estimate. As for methacrylyl-CoA, which was synthesised in the laboratory, it is possible that the published molar extinction coefficient for the substrate is inaccurate, or that the spectrophotometer was miscalibrated, and the concentration measured by DTNB assay should therefore be considered as the high estimate for methacrylyl-CoA.

The final observation is that for methacrylyl-CoA hydrolysed by YciA, the specific rate of hydrolysis increases approximately 5 min after the reaction has commenced from 17 $\text{nmol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ to 54 $\text{nmol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$. This strange observation has been reported previously for this enzyme [121], and is due to initial inhibition of the enzyme by coenzyme A that co-purifies with the enzyme. It was reported that a linear reaction can be obtained by incubating the enzyme in the assay mixture for 5 min prior to adding the substrate. This was attempted, but was unsuccessful. Nevertheless, YciA hydrolyses both methacrylyl-CoA and isobutyryl-CoA at a similar rate and is therefore not sufficiently selective compared to 4HBT.

From this experiment, it appears that a thioesterase, 4HBT, which is selective for methacrylyl-CoA over isobutyryl-CoA, has been identified. Background hydrolytic

activity does appear to exist, however, in cell free extracts of the control strain, on both isobutyryl-CoA and methacrylyl-CoA. Thus, a kinetic characterisation of purified 4HBT was deemed desirable, to both confirm its preference for methacrylyl-CoA over isobutyryl-CoA, and to obtain kinetic parameters for the enzyme. This would provide an insight into the usefulness of this enzyme for use in a metabolic pathway to methacrylic acid which progresses via methacrylyl-CoA and isobutyryl-CoA.

5.1.7 Kinetic characterisation of 4-hydroxybenzoyl-CoA thioesterase

The kinetic characterisation of 4-hydroxybenzoyl-CoA thioesterase was performed using purified hexahistidine tagged (His-tagged) 4HBT. Initially, an attempt to purify an N-terminal His tagged 4HBT enzyme was made, but the protein was unstable. The 4HBT insert from the pBSK::4HBT cloning vector was subcloned into a modified pET20b(+) plasmid, the pET20b(+)PM plasmid prepared by Pawel Mordaka of the same laboratory [128], whereby the *NdeI* restriction site appears downstream of a start codon, itself downstream of the ribosome binding site. However, when the N-terminal-His-tagged 4HBT enzyme (NtHis-4HBT) (Polypeptide sequence 5) was purified by FPLC, and a buffer exchange into HEPES buffer performed, considerable precipitation began to occur during the buffer exchange process, and the enzyme completely precipitated when placed on ice. The effect was reduced when the buffer exchange was performed with phosphate buffer, but there was still some precipitation.

This issue was solved when a carboxy-terminal His-tagged 4HBT enzyme (CtHis-4HBT) was used, instead. To create the CtHis-4HBT gene, the gene encoding 4HBT was modified by polymerase chain reaction. The forward primer, 4HBT_CtHis_F (PID 3) was designed to maintain the 5' end of the 4HBT gene as its wild-type sequence, as well as to keep the *NdeI* restriction site in place. The reverse primer, 4HBT_CtHis_R (PID 4), however, was designed to introduce a tripeptide between the last amino acid of the wild-type enzyme, and the hexahistidine tag encoded by the pET20b(+) vector. Importantly, the first amino acid of this tripeptide was designed to be glycine, as the side chain of this amino acid, a simple hydrogen atom, provides little to no steric obstruction to free rotation of the amino group relative to the carboxy group, and thus acts as a secondary and tertiary structure disrupter, minimising any chance that the His-tag would have a detrimental effect on enzyme solubility. To include this glycine immediately after the 4HBT gene, the reverse primer was designed to replace the stop codon with 5'-GGA-3'. As a restriction site was required to allow the modified insert into the pET20b(+) vector,

the *Xho*I restriction site, encoded by 5'-CTCGAG-3', was chosen. Ligating the PCR product into the pET20b(+) vector at this site would introduce just two amino acids, leucine and glutamate, between the glycine and the His-tag itself. Thus, the 4HBT enzyme was modified by PCR using the 4HBT_CtHis_F and the 4HBT_CtHis_R primer. The PCR product was cloned into pET20b(+) to form pET20b(+):CtHIS-4HBT, and this plasmid was transformed into *E. coli* BL21 (DE3) pLysS to form *E. coli* BL21 (DE3) pLysS pET20b(+):CtHIS-4HBT (Strain 17). Expression studies showed that the enzyme was highly soluble, with very little insoluble protein, as was the case with non-tagged 4HBT (Figure 31).

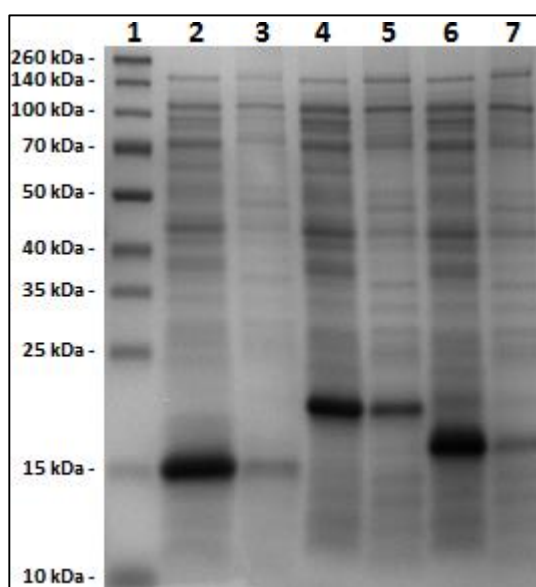


Figure 31: Expression of 4HBT (~16.4 kDa), NtHS-4HBT (~18.6 kDa) and CtHIS-4HBT (~17.5 kDa). Lane 1) Ladder. 2) 4HBT Soluble, 3) 4HBT Insol. 4) NtHis-4HBT Sol. 5) NtHis 4HBT Insol. 6) Ct-His4HBT Sol. 7) CtHIS-4HBT Insol. Molecular weights calculated from polypeptide sequences 1, 5 & 6, using the Expasy ProtParam tool.

The expression was scaled to 500 ml cultures and the enzyme was purified by FPLC. The buffer exchange for the purified protein was performed with phosphate buffer, which had resulted in lower precipitation of NtHis-4HBT than HEPES buffer. No precipitation was observed for this enzyme during buffer exchange. The purified enzyme was then flash frozen in liquid nitrogen prior to kinetic characterisation. For the kinetic characterisation, the enzyme was thawed on ice. Some small level of precipitation was cleared by warming the enzyme to 30 °C once fully thawed, and the concentration of the thawed enzyme determined by UV₂₈₀ absorbance on a NanoDrop ND1000 spectrophotometer using the predicted molar extinction coefficient for the enzyme of 20970 M⁻¹.cm⁻¹ and a molecular weight of 17516.5 mg.mmol⁻¹. These values were calculated with the Expasy ProtParam tool, using the CtHIS-4HBT polypeptide sequence (Polypeptide sequence 6), assuming that all cysteins were reduced. A kinetic characterisation was then performed for

purified CtHIS-4HBT with both methacrylyl-CoA, in triplicate, and isobutyryl-CoA, in duplicate. A preliminary experiment showed that methacrylyl-CoA was hydrolysed much more rapidly than isobutyryl-CoA, and so for the kinetic characterisation with methacrylyl-CoA, 75 $\mu\text{g}.\text{ml}^{-1}$ of the CtHIS-4HBT enzyme was used, whilst 1 $\text{mg}.\text{ml}^{-1}$ was used for the characterisation with isobutyryl-CoA. As it was not certain what the exact stock concentration of methacrylyl-CoA and isobutyryl-CoA were, both high and low substrate concentration estimates were used when plotting the Lineweaver Burk graphs (Figure 32).

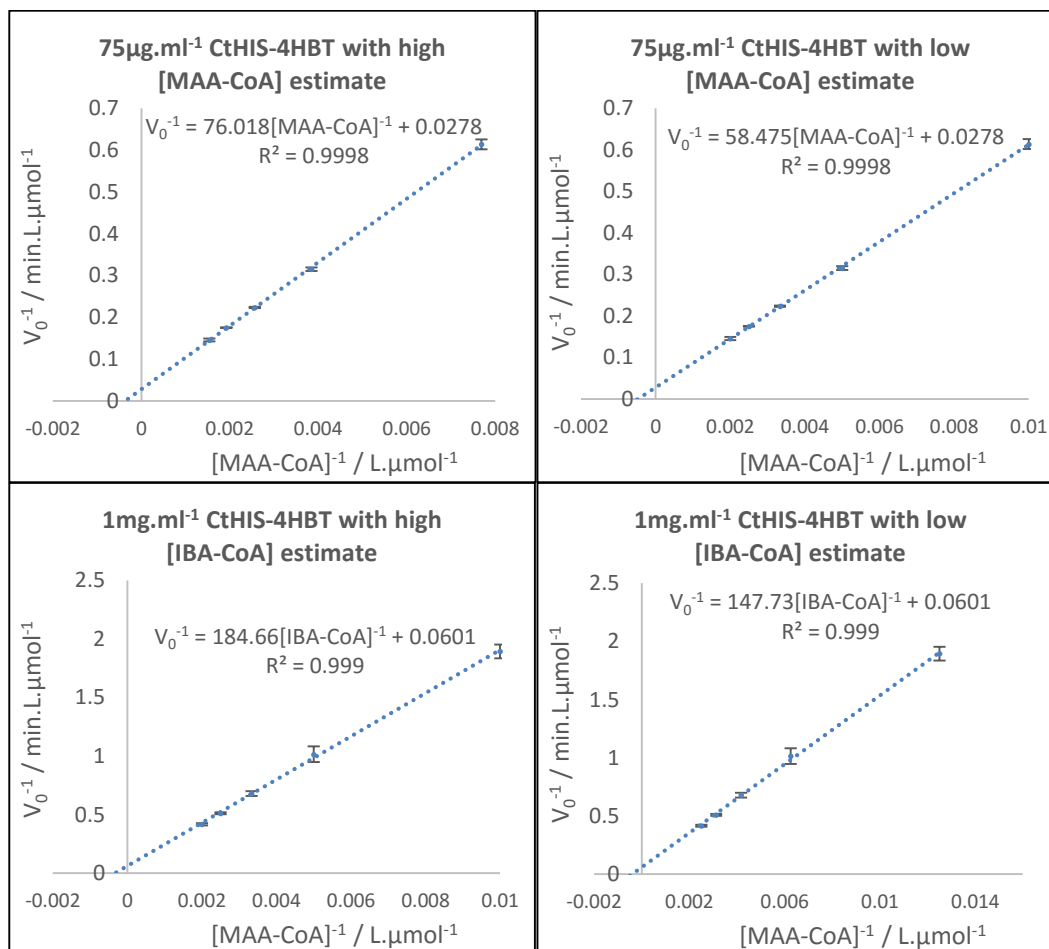


Figure 32: Lineweaver Burk plots for kinetic characterisation of CtHIS-4HBT on methacrylyl-CoA and isobutyryl-CoA, with high and low estimates of substrate concentrations. Raw data for initial velocities in appendices: Table 37, page XXI.

The kinetic parameters for the reactions were then calculated from the Lineweaver Burk plots using the equation of the Lineweaver Burk plot (Equation 3).

$$\frac{1}{V_0} = \frac{K_M}{V_{MAX}} \cdot \frac{1}{[S]} + \frac{1}{V_{MAX}}$$

Equation 3: Equation of the Lineweaver Burk plot

The maximum specific activity of CtHIS-4HBT with both methacrylyl-CoA and isobutyryl-CoA was calculated from the V_{MAX} and the enzyme concentration in each kinetic characterisation experiment. Both high and low estimates of the Michaelis constants were also calculated (Table 17).

Substrate	Maximum specific activity / $\text{nmol.min}^{-1}.\text{mg}^{-1}$	K_M / mM	
		(high estimate)	(low estimate)
Methacrylyl-CoA	480	2.7	2.1
Isobutyryl-CoA	16.4	3.1	2.5

Table 17: Summary kinetic data of 4HBT from *Arthrobacter sp.* strain SU with methacrylyl-CoA and isobutyryl-CoA as substrates.

The affinity of CtHIS-4HBT for both methacrylyl-CoA and isobutyryl-CoA appeared to be similar, with the high estimate for methacrylyl-CoA marginally higher, at 2.7 mM, than the low estimate for isobutyryl-CoA at 2.5 mM. Though these are not particularly good Michaelis constants for an enzyme catalysed reaction, they are not so high that the enzyme be discounted from engineering it into a metabolic pathway for the production of methacrylic acid. As these K_M values are similar, the binding affinity of methacrylyl-CoA compared to that of isobutyryl-CoA is not the reason for the selectivity of 4HBT for methacrylyl-CoA over isobutyryl-CoA. With a maximum specific activity of $480 \text{ nmol.min}^{-1}.\text{mg}^{-1}$ for CtHIS-4HBT with methacrylyl-CoA, an adequate activity for an engineered metabolic pathway, compared to a maximum specific activity of $16.4 \text{ nmol.min}^{-1}.\text{mg}^{-1}$ with isobutyryl-CoA as a substrate, the catalytic activity of the enzyme appears to be the main reason for selectivity.

An explanation for the observation that 4HBT from *Arthrobacter sp.* strain SU accepts methacrylyl-CoA as a substrate, as well as its selectivity for methacrylyl-CoA over isobutyryl-CoA may lie in its crystal structure [118] and the catalytic mechanism [129]. The coenzyme A moiety of 4-hydroxybenzoyl-CoA is stabilised by hydrogen bonding interactions with non-catalytic residues of 4HBT, Ser120, Thr121, Arg102 and Arg150 (Figure 33). The 4-hydroxybenzoyl moiety itself is stabilised through direct hydrogen bond interactions between its hydroxy group and Glu78, and through indirect, water mediated, hydrogen bonding interactions between the same hydroxyl group, and Asp31 [129]. Catalytic residues include Glu73 and Thr77. Glu73 first initiates nucleophilic attack on the substrate's carbonyl carbon, forming an anhydride and displacing the coenzyme A moiety. The transition state of this is stabilised by Gly65 and Gln58. Thr77 then co-ordinates a water molecule for nucleophilic attack on the 4-hydroxybenzoyl carbon of the anhydride, hydrolysing the anhydride and releasing 4-hydroxybenzoic acid [129].

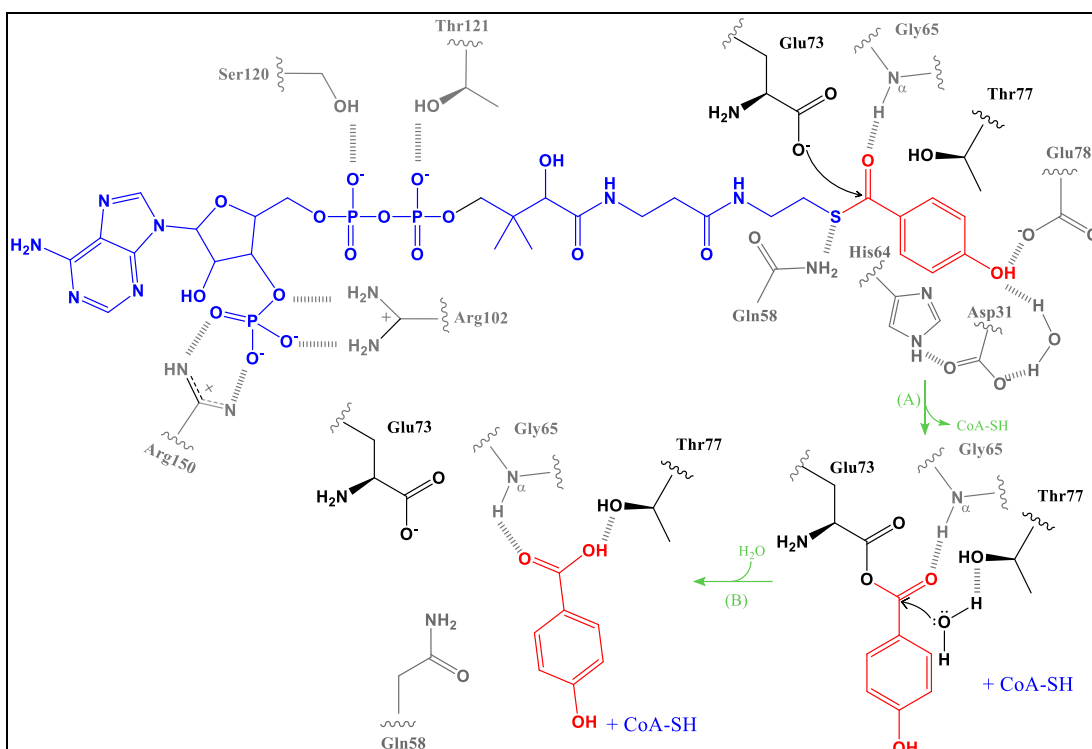


Figure 33: Enzymatic mechanism of 4-hydroxybenzoyl-CoA hydrolysis. (Adapted from [129])

It is therefore easy to envisage how, with methacrylyl-CoA anchored to the enzyme through the interactions between the coenzyme A moiety and residues Ser120, Thr121, Arg102 and Arg150, an acyl moiety no larger or bulkier than the natural substrate, catalytic residues which only attack the carbonyl carbon of the substrate and residues which stabilise the carbonyl oxygen and thioester sulphur in transition state, 4HBT could accept methacrylyl-CoA with a similar mechanism (Figure 34).

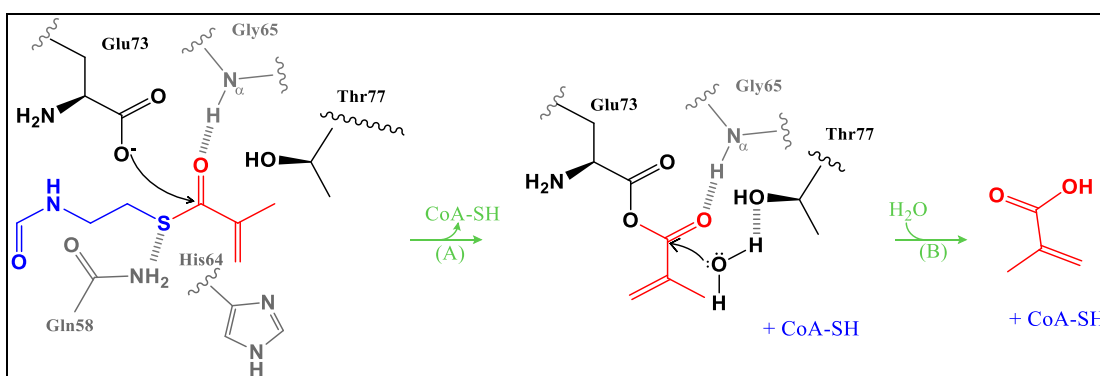


Figure 34: 4HBT catalytic mechanism [129], envisaged with methacrylyl-CoA as substrate.

The answer for the selectivity of methacrylyl-CoA over isobutyryl-CoA could lie with the fact that the methacrylyl-group is electron withdrawing [81], making its carbonyl more electropositive than the carbonyl of the isobutyryl anhydride, and thus more susceptible to nucleophilic attack by the water molecule [69].

5.2 Oxidation of isobutyryl-CoA to methacrylyl-CoA

Having demonstrated the selective hydrolysis of methacrylyl-CoA, the next step of this project was to introduce the enzymatic oxidation of isobutyryl-CoA to methacrylyl-CoA (Figure 35) into *E. coli*.

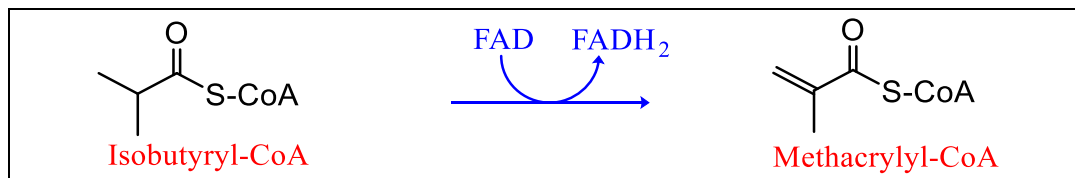


Figure 35: Oxidation of isobutyryl-CoA to methacrylyl-CoA.

5.2.1 Enzymes catalysing the oxidation of acyl-CoA intermediates

The only enzymes listed in the MetaCyc and KEGG pathway databases for the oxidation of isobutyryl-CoA to methacrylyl-CoA in the valine degradation I pathway are the acyl-CoA dehydrogenases. Furthermore, previous designs for the production of either methacrylic acid or methacrylate esters [20], as well as for the potential precursor to methacrylic acid, 3-hydroxyisobutyric acid [130], have also only considered the acyl-CoA dehydrogenases for the oxidation of isobutyryl-CoA to methacrylyl-CoA. However, in some plants, acyl-CoA oxidases are responsible for the oxidation of isobutyryl-CoA [83]. The acyl-CoA dehydrogenases and the acyl-CoA oxidases are two different classes of enzyme which are both flavoproteins that utilise a tightly associated flavin adenine dinucleotide (FAD) cofactor as a temporary electron acceptor for the oxidation. They have very similar mechanisms for the FAD reductive half-reaction [89] (Figure 36), but differ in the mechanism for the re-oxidation of FADH₂. The relative merits of both classes were investigated in order to make an informed decision.

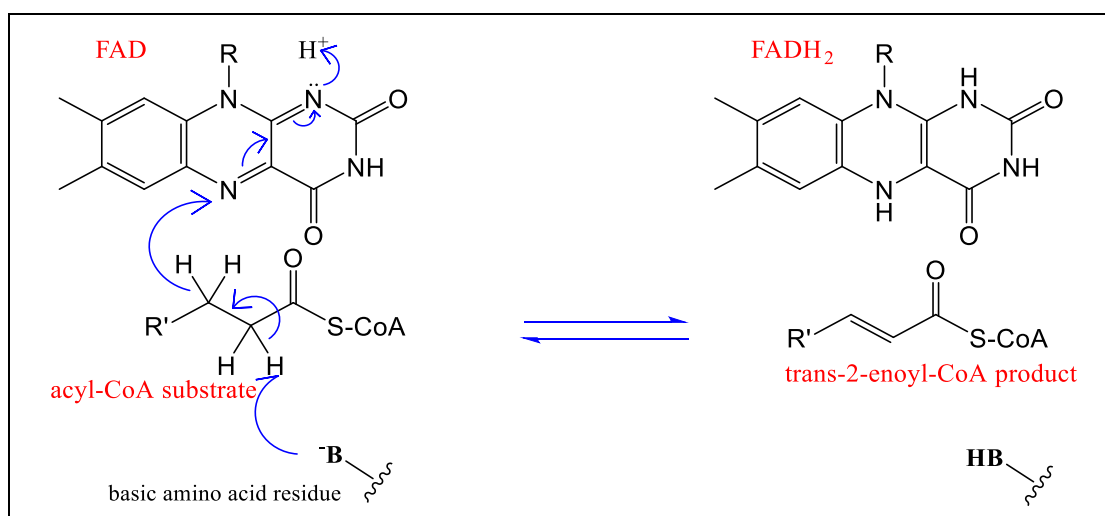


Figure 36: The FAD reductive half reaction during the oxidation of acyl-CoA substrate to *trans*-2-enoyl-CoA. Involves proton abstraction from α -carbon of substrate followed by hydride transfer to FAD. Adapted from: [131].

5.2.1.1 The acyl-CoA dehydrogenases

The acyl-CoA dehydrogenases catalyse the α - β -dehydrogenation of an acyl-CoA substrate to its *trans*-2-enoyl-CoA analogue, and have been most widely studied in mammalian cell mitochondria, as defects in the acyl-CoA dehydrogenases are associated with various human diseases [132, 133]. In mammalian cell mitochondria, nine acyl-CoA dehydrogenases (ACADs) are involved in the degradation of acyl-CoA molecules from both fatty acids and amino acids, with five ACADs involved in fatty acid degradation, and four ACADs involved in amino acid degradation (Table 18). These primary acyl-CoA dehydrogenases typically have between 35% and 45% sequence homology [89] and also have conserved structural features [134, 135], suggesting common ancestry.

Short name	Full Name	Role
SCAD	Short chain acyl-CoA dehydrogenase	Fatty acid β -oxidation
MCAD	Medium chain acyl-CoA dehydrogenase	
LCAD	Long chain acyl-CoA dehydrogenase	
vLCAD	Very long chain acyl-CoA dehydrogenase	
vLCAD2 (ACAD9)	Very long chain acyl-CoA dehydrogenase 2 (acyl-CoA dehydrogenase 9)	
GD	Glutaryl-CoA dehydrogenase	Proline & Tryptophan Degradation
IBD	Isobutyryl-CoA dehydrogenase	Valine degradation
IVD	Isovaleryl-CoA dehydrogenase	Leucine degradation
SBD	Short/Branched chain acyl-CoA dehydrogenase (2-methyl-butyryl-CoA dehydrogenase)	Isoleucine degradation

Table 18: Acyl-CoA dehydrogenases found in mammalian cell mitochondria

With FAD acting as the temporary electron acceptor for the oxidation of the acyl-CoA substrate, the physiological electron acceptor is a type I electron transfer flavoprotein (ETF) [136]. In mammals, there is a single ETF acting as the physiological acceptor to all nine acyl-CoA dehydrogenases, and it also acts as the electron acceptor for other primary dehydrogenases, including sarcosine and dimethylglycine dehydrogenases [137]. In more metabolically diverse prokaryotes, ETFs can accept and even donate electrons from/to other dehydrogenases such as trimethylamine dehydrogenase (TMADH) [138]. There are up to three types of ETF that may be expressed by an organism, and though they are structurally related, serve different roles. It is the type I ETFs that re-oxidise the flavin cofactor of the acyl-CoA dehydrogenases, and subsequent reference to ETFs will imply type I ETFs.

The ETF reoxidises the reduced flavin of the primary dehydrogenases by electron transfer from that flavin onto its own FAD molecule. For kinetic reasons, only a single electron transfer to the ETF's flavin occurs physiologically [139, 140], and so two ETFs are required for the full reoxidation of the ACAD's flavin, with each ETF's flavin being reduced to a semiquinone [131, 141]. Electrons liberated from the acyl-CoA dehydrogenases by the ETF are shuttled to the inner mitochondrial/cell

membrane and transferred to a matrix/cytosol facing, membrane associated electron transfer flavoprotein ubiquinone oxidoreductase (ETFQO), first identified in bovine heart mitochondria [142]. The ETFQO is an iron-sulfur flavoprotein that accepts a single electron from each of two ETF_{SQ} , and in turn reduces ubiquinone (UQ) to ubiquinol (UQ:H_2) [143] (Figure 37), linking the oxidation of an acyl-CoA to the generation of ATP at the electron transport chain [144, 145].

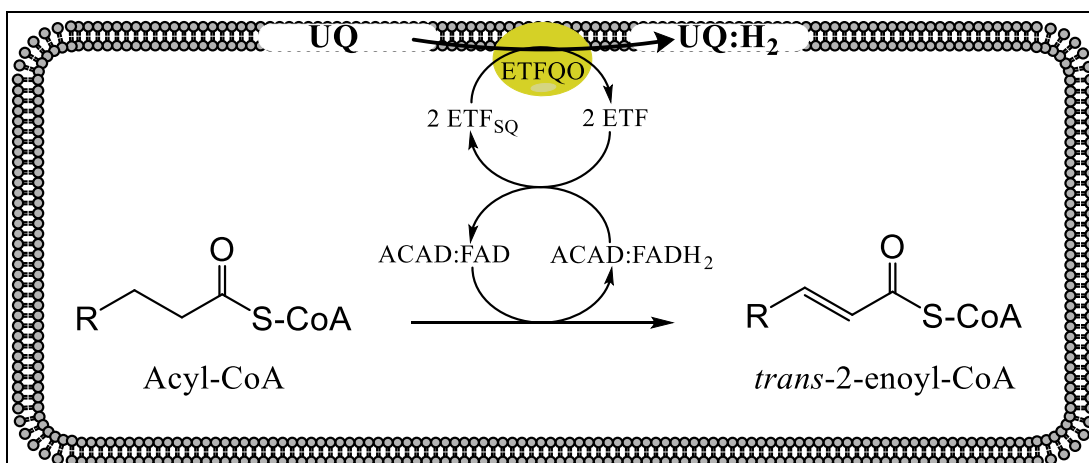


Figure 37: Oxidation of an acyl-CoA by an ACAD and regeneration of the ACAD cofactor.

The acyl-CoA dehydrogenase-catalysed oxidation of the acyl-CoA substrate happens as a ternary complex between the substrate, the acyl-CoA dehydrogenase and the ETF (Figure 38). In this ternary complex, the *trans*-2-enoyl-CoA product remains bound to the active site until the FADH₂ is reoxidised by the two ETFs. With the product remaining bound, the FAD is solvent inaccessible, preventing it from being reoxidised instead by molecular oxygen, ensuring that the liberated electrons are transferred to the electron transport chain [131], though mutant dehydrogenases with oxidase activity have been developed [89].

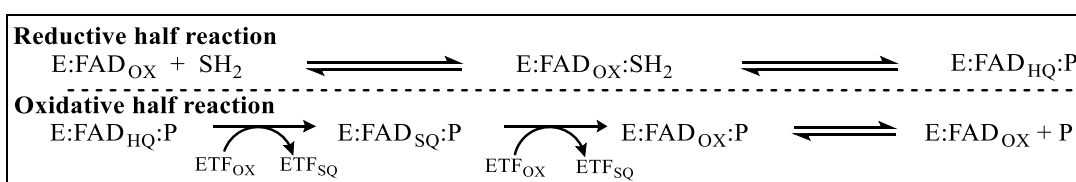


Figure 38: Oxidation of saturated acyl-CoA substrate (SH_2) to an unsaturated *trans*-2-enoyl-CoA product by $2 \times \text{ETF}$. Adapted from [131]. Reaction is mediated by two redox half reactions at an enzyme associated FAD molecule, through a ternary complex enzymatic reaction mechanism (adapted from [131]). (FAD_{OX} : oxidised FAD. FAD_{HQ} : FADH_2 . ETF_{OX} : ETF with its FAD cofactor fully oxidised. ETF_{SQ} : ETF with FAD semiquinone.

The ETFQO has a 4Fe4S iron sulphur cluster and an FAD molecule as redox centers. The structure of ETFQO from pig liver mitochondria (Figure 39) [146], showed that the ETFQO, which is composed of a single polypeptide [147], is associated to the

inner mitochondrial/cell membrane by a small hydrophobic patch, through which the membrane soluble ubiquinone binds.

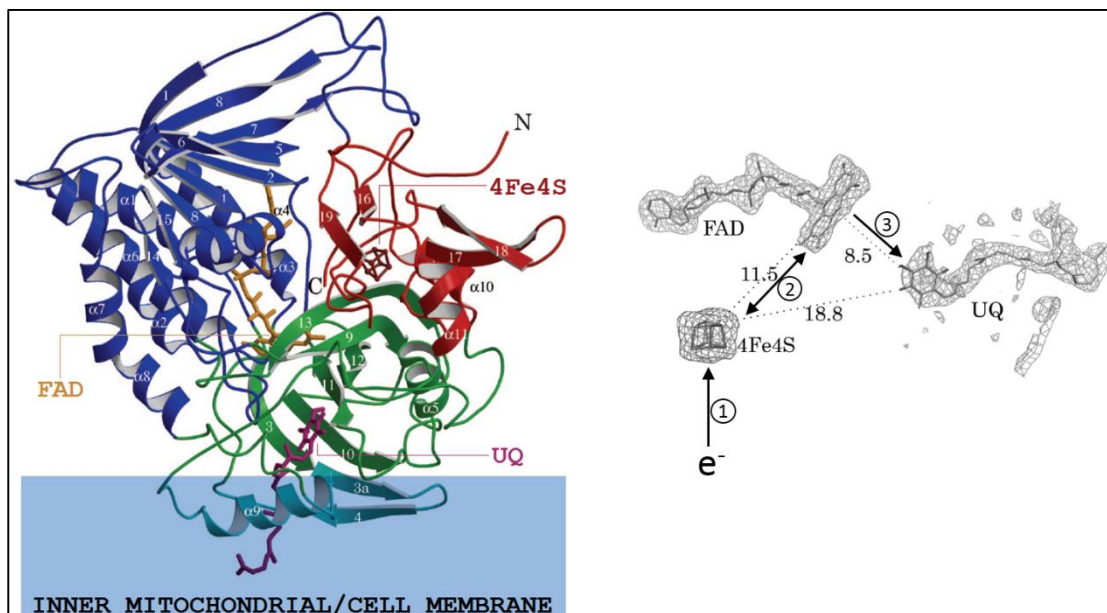


Figure 39: Structure of ETFQO from pig liver mitochondria and inter-redox center distances (Adapted from [146] and [148]). Studies from *R. sphaeroides* [149] suggest that: 1) electrons enter the ETFQO via the iron sulphur cluster. 2) The electron supplied by the first ETF_{SQ} equilibrates between the isopotential 4Fe₄S cluster and the FAD. On the next electron transfer from the second ETF_{SQ}, one electron is held in the cluster, and the other on the Flavin. 3) Electrons are transferred to the ubiquinone via the Flavin [148]. This may change between ETFQOs of different organisms.

In order to accept electrons from the primary dehydrogenases, the FAD redox center of the ETF must be brought into close proximity with that of the primary dehydrogenases's redox center, requiring complex protein-protein interactions in order to do so [150, 151]. This is also the case for the transfer of electrons to the ETFQO. Whilst the ACADs share some structural similarities, they are not structurally identical, and sarcosine dehydrogenase, dimethylglycine dehydrogenase and the ETFQO are all structurally distinct [135, 146, 152]. The ETF must therefore balance relatively broad specificity to its partner complexes such that it can accept electrons from multiple dehydrogenases and donate them to an ETFQO, without participating in undesired redox reactions. Unique structural features of the ETF and its partner complexes enable the balance to be achieved between promiscuity and specificity, in spite of these broad structural differences.

5.2.1.2 Structural properties of ETFs balance specificity and promiscuity

Electron transfer flavoproteins are heterodimers and are composed of two subunits, ETF α and ETF β , which together form three domains, domain I, II and III. Domains I and II are predominantly composed of the α -subunit, whilst domain III entirely of the β -subunit [89, 150, 153]. ETFs typically contain a purely structural ATP

molecule in domain III, and a tightly associated FAD compound in domain II. Domain II sits in a cleft formed by the α and β subunits, joined to domains I and III by a flexible linker. Domain III of the ETF contains a highly conserved recognition loop which interacts with a hydrophobic patch on the primary dehydrogenase, and is critical to successful complex formation between the ETF and its partner complexes [135, 138, 150]. Cross-linking experiments between ETF and ETFQO have shown that the ETF β -subunits are also involved in complex formation with the ETFQO [154], but the mechanisms have not been as thoroughly studied. On complex formation, a conformational shift undocks domain II from the well ordered position that it is in when the ETF is in an uncomplexed state, enabling it to sample the available (thermodynamically permitted) conformational space between the PD and domains I and III of the ETF [138], in a manner described as induced disorder [155]. If in one of the conformations that is sampled, the oxidised FAD molecule of the ETF's domain II, and the reduced redox center of the PD are brought within 14Å, then so long as the concomitant oxidation of the ACAD's FAD and reduction of the ETF's FAD is thermodynamically feasible, an electron transfer competent state is achieved, and electron transfer from one complex to the other may proceed by electron tunneling [135, 138, 150, 151, 155, 156].

Thus, with a hydrophobic recognition loop, the ETF can specifically complex primary dehydrogenases, and with a disordered flavin domain, the ETF can promiscuously form electron transfer competent states with structurally diverse partners, so long as that initial specific interaction via the recognition loop has occurred.

However, there is sufficient evidence to suggest that the above story is rather more complicated. Outside of the highly conserved recognition loop on the ETF, which is essential for complex formation, ionic interactions as well as both direct and indirect (mediated through water molecules) hydrogen bonding interactions are believed to further stabilise complex formation, though the residues which participate in these interactions are not conserved between acyl-CoA dehydrogenases [135, 138]. Furthermore, electron transfer competent conformations between the primary dehydrogenase and the disordered flavin domain on the ETF are perceived to be transiently "frozen" [155] through intermolecular interactions. By transiently freezing the electron transfer competent states, the half-life of the electron-transfer competent state is increased, making actual successful electron transfer more likely [138]. In both human and *Methylophilus methylotrophus* ETFs, which are two very well studied ETFs, an arginine residue in the ETF domain II is involved in such a stabilisation, with a corresponding glutamate (Glu212) in human MCAD interacting with the arginine from human ETF *via* ionic interactions, and a tyrosine (Tyr442)

from TMAD_{Mm} interacting with the R α 237 from ETF_{Mm} via hydrogen bonding interactions, though other residues may be involved in such stabilisations in similar systems from other organisms. Disruption of these interactions between TMAD and ETF_{Mm} diminished electron transfer [138, 157-159]. Similarly, disruption of the initial hydrophobic interaction between the recognition loop of the ETF and the hydrophobic patch on the acyl-CoA dehydrogenase has also shown similar impedance of complex formation and successful electron transfer [135]. With non-conserved salt bridges further stabilising initial complex formation, and an unclear contribution of hydrogen bonding interactions and ionic interactions transiently stabilising electron transfer competent states with the disordered flavin domain, it is quite likely that complex formation may be rather more specific than having a disordered flavin domain may imply. In addition, Toogood *et al.* predict that a different family of conformations are compatible with electron transfer between the human ETF:MCAD complex and the *M. methylotrophus* ETF:TMAD complex [135], suggesting that it would be unlikely for an ETF from *M. methylotrophus* to form an electron transfer competent state with human MCAD, even if there was successful complex formation, let alone have the electron transfer competent state transiently stabilised.

5.2.1.2.1 Thermodynamic contributions to ETF specificity

An electron transfer competent state is only achieved if electron transfer is thermodynamically favourable. Both the electron transferring flavoproteins and the majority of the primary dehydrogenases utilise an FAD molecule as the redox centre. Whilst the standard reduction potential for the full reduction of FAD to FADH₂ (FAD_{ox} to FAD_{HQ}) is -207 mV when free in solution, the midpoint potential of protein bound FAD to FADH₂ can vary from -400 mV to +60 mV [160]. Similarly, the redox potentials of the iron sulphur clusters in both hydrogen bonding interactions that stabilise the iron sulphur clusters [161] of the ETFQO, as well as the flavins [146, 162], of different ETFQOs can alter their reduction potentials. Considering this wide range of variability in midpoint potentials of protein bound redox centers, it is not unreasonable to consider that ETFs, ETFQOs and primary dehydrogenases from one organism could have one set of midpoint potentials, and the ETF, ETFQO and primary dehydrogenases from another organism having another set, whereby electron transfer between the ETF from one organism and primary dehydrogenases or ETFQOs from other organisms, is thermodynamically unfavorable.

Thus, in summary, the association of ETF and ACAD is temporary and requires both conserved hydrophobic interactions and non-conserved ionic, and hydrogen bonding interactions, the precise nature of which may vary between organisms, to

transiently stabilise complex formation; the available conformational space differs between complexes from different sources, and electron transfer competent states do not depend on simply there being 14 Å between the two redox centers. Electron transfer must be thermodynamically favourable and this is a factor which is affected by the protein environment of the specific redox partner. Finally, electron transfer competent states are transiently stabilised through interactions that are not necessarily conserved between redox partners. This feature is highlighted by the differing effects of the disruption of one residue in human ETF on the electron transfer from the different primary dehydrogenases of the same organism [163].

It should, therefore, not be assumed that cloning a heterologous acyl-CoA dehydrogenase into a host organism will necessarily create a functional system *in vivo* for the oxidation of its natural acyl-CoA substrate to the corresponding *trans*-2-enoyl-CoA product, without expressing a compatible electron transport system at the same time. Some evidence to support this concern comes from experiments between the ETF, ETFQO and primary dehydrogenases from *Paracoccus denitrificans* as well as from *Sus scrofa* (pig). Whilst *Paracoccus* ETF oxidised *Paracoccus* glutaryl-CoA dehydrogenase, porcine butyryl-CoA dehydrogenase and porcine octanoyl-CoA dehydrogenase, *Paracoccus* ETF does not oxidise porcine glutaryl-CoA dehydrogenase, nor does it reduce porcine ETFQO, despite *paracoccus* ETFQO oxidising porcine ETF [164, 165].

In order to therefore clone an acyl-CoA dehydrogenase into *E. coli* for the *in vivo* oxidation of isobutyryl-CoA, some evidence that it is compatible with the electron transport system of *E. coli* must exist, or a compatible electron transport system must also be engineered into the *E. coli* host. Interestingly, it is believed that *E. coli* only has a single acyl-CoA dehydrogenase, encoded by *fadE*. This enzyme is responsible for the α - β -dehydrogenation oxidation of all acyl-CoA intermediates in fatty acid β -oxidation to their *trans*-2-enoyl-CoA analogues [166]. Furthermore, despite *E. coli* being the best studied micro-organism, a database and literature search failed to identify a type I electron transfer flavoprotein in this organism, and Campbell and Cronan suggest that the acyl-CoA dehydrogenase from *E. coli* may also serve an electron transferring function to the respiratory chain [166]. The implication of this would be that an exogenous acyl-CoA dehydrogenase could never function in *E. coli*. Whether this is correct or not has not been proven, but previous attempts to clone just an isobutyryl-CoA dehydrogenase into *E. coli* for the *in vivo* oxidation of isobutyryl-CoA have indeed failed (Wataru Mizunashi, personal communication). A compatible type I ETF and ETFQO would therefore also have to

be cloned into the *E. coli* host organism along with the exogenous acyl-CoA dehydrogenase, in order for an isobutyryl-CoA dehydrogenase to function *in vivo*.

5.2.1.3 The acyl-CoA oxidases

The acyl-CoA oxidases catalyse the α - β -dehydrogenation of acyl-CoA thioesters to their *trans*-2-enoyl-CoA analogues, with the concomitant reduction of molecular oxygen to hydrogen peroxide (Figure 40), without the need for an ETF mediated electron transfer to the respiratory chain, and at rates higher than the analogous oxidase activity in mutant dehydrogenases that would be ETF-independent [89].

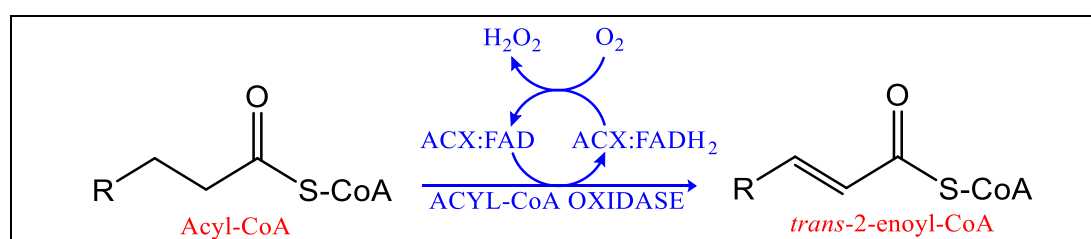


Figure 40: The acyl-CoA oxidase (ACX) mediated oxidation of an acyl-CoA to its *trans*-2-enoyl-CoA analog by molecular oxygen and the concomitant production of hydrogen peroxide.

These acyl-CoA oxidases participate in peroxisomal β -oxidation, in a pathway that runs parallel to mitochondrial β -oxidation in many eukaryotes. Interestingly, whilst both peroxisomal and mitochondrial β -oxidation occurs in mammalian cells, there is no detectable short chain acyl-CoA oxidase activity ($C < 8$) in mammalian peroxisomes, and so mammalian peroxisomal β -oxidation does not proceed to completion [167]. On the other hand, in most yeasts and plants, the peroxisomes appear to be the sole location for β -oxidation [168].

Whilst peroxisomes have been established as an alternative site for β -oxidation of fatty acids in eukaryotes, the analogous pathways for the degradation of the branched chain amino acids appear to occur solely in the mitochondria of mammalian cells [169, 170]. However, the location for branched chain amino acid degradation in plants, on the other hand, is less clear cut, with evidence for both mitochondria and peroxisomes playing a role [171].

An acyl-CoA oxidase with activity on isobutyryl-CoA could therefore exist, and so long as that oxidase can be expressed in *E. coli*, it ought to catalyse the oxidation of isobutyryl-CoA to methacrylyl-CoA *in vivo*, without any complications associated with electron transfer to the electron transport chain.

The acyl-CoA oxidases do produce, however, hydrogen peroxide as a product, which is toxic [172]. The expression of an acyl-CoA oxidase for the oxidation of

isobutyryl-CoA may therefore be detrimental to the health of the organism, especially if the pathway to methacrylic acid is highly active. *E. coli* does have an oxidative stress response system, which acts to scavenge hydrogen peroxide [172, 173]. Whether such a system is able to cope with the H₂O₂ produced by the metabolic pathway would depend on how active that pathway was, and can only be determined during characterisation experiments.

A search for both acyl-CoA dehydrogenases, with a compatible ETF and ETFQO, and acyl-CoA oxidases, was therefore performed. If an isobutyryl-CoA dehydrogenase and compatible ETF and ETFQO could be identified, this would mitigate the risk from hydrogen peroxide toxicity. An acyl-CoA oxidase with activity on isobutyryl-CoA would be far simpler to clone, however, and ought to function *in vivo* without the additional electron transfer infrastructure that the acyl-CoA dehydrogenases require.

5.2.2 Identification of candidate acyl-CoA dehydrogenases for the oxidation of isobutyryl-CoA with compatible ETFs and ETFQOs

A database and literature search was performed to identify candidate acyl-CoA dehydrogenases for the oxidation of isobutyryl-CoA to methacrylyl-CoA, and six candidate acyl-CoA dehydrogenases were identified (Table 19). In addition to these, bovine and mouse isobutyryl-CoA dehydrogenases were also listed in the UniProt databases, but the cited literature was inconsistent with the database, so the entries were deemed unreliable, and therefore not further considered.

Name of enzyme	Source	Uniprot accession number	Ref.
Isobutyryl-CoA dehydrogenase	<i>Homo sapiens</i>	Q9UKU7	[88]
Isovaleryl-CoA dehydrogenase	<i>Arabidopsis thaliana</i>	Q9SWG0	[174]
Isovaleryl-CoA dehydrogenase	<i>Solanum tuberosum</i>	Q9FS87	[175]
2-methylacyl-CoA dehydrogenase	<i>Rattus norvegicus</i>	P70584	[176]
2-methylacyl-CoA dehydrogenase	<i>Streptomyces avermitilis</i>	N/A	[95]
2-methylacyl-CoA dehydrogenase	<i>Streptomyces coelicolor</i>	N/A	[95]

Table 19: Candidate acyl-CoA dehydrogenases for the oxidation of isobutyryl-CoA

A search was then made for the type I ETF and the ETFQO proteins from the same organisms to ensure that the ETF and ETFQO would be compatible with that dehydrogenase. A database search was therefore performed on the UniProt database using "Electron transfer flavoprotein" as the search term and the result-set was screened for ETF α , ETF β and ETFQO proteins associated with the organisms from which the acyl-CoA dehydrogenases listed above are derived. The results are summarised below (Table 20).

Organism	Protein	Entry for protein in UniProt?	Status	UniProt accession number	Most relevant refs.
Homo sapiens	ETF α	Yes	Reviewed	P13804	[177, 178]
	ETF β	Yes	Reviewed	P38117	[178, 179]
	ETFQO	Yes	Reviewed	Q16134	[180]
Arabidopsis thaliana	ETF α	Yes	Reviewed	Q9C6I6	[181]
	ETF β	Yes	Reviewed	Q9LSW8	[181]
	ETFQO	Yes	Reviewed	O22854	[182]
Solanum tuberosum	ETF α	No	N/A	N/A	
	ETF β	No	N/A	N/A	
	ETFQO	No	N/A	N/A	
Rattus norvegicus	ETF α	Yes	Reviewed	P13803	[183]
	ETF β	Yes	Reviewed	Q68FU3	[184]*
	ETFQO	Yes	Reviewed	Q6UPE1	[185]
Streptomyces avermitilis	ETF α	Yes	Unreviewed	Q82N23	
	ETF β	Yes	Unreviewed	Q82N22	
	ETFQO	No	N/A	N/A	
Streptomyces coelicolor	ETF α	Yes	Unreviewed	Q9K418	
	ETF β	Yes	Unreviewed	Q9K417	
	ETFQO	No	N/A	N/A	

Table 20: Summary of ETF α , ETF β and ETFQO entries for organisms which contain acyl-CoA dehydrogenases capable oxidising isobutyryl-CoA to methacrylyl-CoA. * - reference does not show experimental verification of ETF β sequence.

Of all the organisms for which ETF α , ETF β and ETFQO proteins were searched for, only *Homo sapiens*, *Arabidopsis thaliana* and *Rattus norvegicus* had reviewed entries for the ETF α , ETF β and ETFQO proteins. No reviewed entries for the ETF α , ETF β and ETFQO proteins found for *Solanum tuberosum*, *Streptomyces avermitilis* or *Streptomyces coelicolor*. As for the set of human enzymes, gene sequences for the ETF α [177] and ETF β [179] components of ETF have been identified, and have both been co-expressed into a functional ETF in *E. coli* [178]. However, the gene sequence of human ETFQO was only successfully expressed in yeast [180], with no published evidence for successful expression of this enzyme in *E. coli*. In addition, the expression of human IBD in *E. coli* required the coexpression of GroES/EL chaperonin [88], suggesting that even the cytosol soluble IBD is difficult to express in *E. coli*.

The entry for rat ETF β was deemed unreliable, as the literature cited here does not show experimental verification of the gene sequence listed. More importantly, however, there appears to be no evidence for the expression of recombinant rat ETF in any organism, let alone *E. coli*. Furthermore, as with human ETFQO, rat ETFQO was not expressed in *E. coli*, but instead in yeast and also in COS-7 mammalian cell lines [185].

Finally, as for the *Arabidopsis* ETF α , ETF β and ETFQO genes, compelling if not definitive evidence is provided that the At2g43400 gene identified by the *Arabidopsis* genome initiative, does indeed encode an ETFQO [182], and that genes At1g50940 and At5g43430 encode the ETF α and ETF β subunits, respectively. However, the genes encoding *Arabidopsis* ETF α and ETF β do not appear to have yet been co-expressed into a functional ETF in *E. coli*, nor does the gene encoding the *Arabidopsis* ETFQO appear to have yet been functionally expressed in *E. coli*.

The first choice for the expression of an acyl-CoA dehydrogenase with known activity on isobutyryl-CoA along with a compatible electron transport system composed of an ETF and an ETFQO was to use enzymes all from the same organism. However, none of the identified ETFQO enzymes from any of the organisms in which an acyl-CoA dehydrogenase with known activity for the oxidation of isobutyryl-CoA have been demonstrated to be functionally expressed in *E. coli*. During the literature search for this chapter, however, it was observed that porcine ETF accepts electrons from human isobutyryl-CoA dehydrogenase [88, 186]. In addition to this, ETFQO from *Rhodobacter sphaeroides*, which has been successfully expressed in *E. coli*, accepts electrons from porcine ETF [149]. These two observations have been made in isolated experiments, *in vitro*, and it is unclear whether such a system

would function, as a whole, *in vivo*. Another caveat is that unlike human ETF, porcine ETF does not appear to have been expressed in *E. coli* before. However, considering that porcine ETF has high sequence similarity to human ETF, successful expression is likely.

Finally, it is known that the valine degradation I pathway exists in both *Pseudomonas putida* [85, 86, 96] and *Pseudomonas aeruginosa* [85, 97]. In *Pseudomonas putida* it is known that an acyl-CoA dehydrogenase is responsible for the oxidation of isobutyryl-CoA to methacrylyl-CoA, and the same is likely with *Pseudomonas aeruginosa*. After using "Isobutyryl-CoA dehydrogenase *Pseudomonas*" as the search term on the UniProt database, two unreviewed entries were returned. These were for a putative isobutyryl-CoA dehydrogenase from *Pseudomonas aeruginosa* (no strain designated) (UniProt accession number: A0A0A8RMU3) as well as one from *Pseudomonas aeruginosa* MH38 (UniProt accession number: W0WLM7). The two proteins differ by just two amino acids, and a protein BLAST search showed that both have 99.7% sequence identity to an enzyme in *Pseudomonas aeruginosa* PAO1, a well studied strain of *P. aeruginosa*, annotated as a probable acyl-CoA dehydrogenase with UniProt accession number Q9I5I3.

A BLAST comparison between the probable acyl-CoA dehydrogenase from *P. aeruginosa* PAO1 listed above and human isobutyryl-CoA dehydrogenase using the human enzyme as the search query, showed that the enzyme from *P. aeruginosa* PAO1 has 55% sequence identity and 72% sequence similarity to the human isobutyryl-CoA dehydrogenase, making it quite likely that this is indeed an isobutyryl-CoA dehydrogenase, considering that the human ACADs only share 35%-45% sequence homology between themselves. For comparison, human IBD shares 33% sequence identity and 55% sequence similarity with human Medium Chain Acyl-CoA Dehydrogenase (MCAD), whilst the probable isobutyryl-CoA dehydrogenase from *P. aeruginosa* shares 37% sequence identity and 57% sequence similarity with human MCAD. If the probable isobutyryl-CoA dehydrogenase from *P. aeruginosa* can be experimentally verified, then an isobutyryl-CoA dehydrogenase, an ETF and an ETFQO all from the same prokaryotic organism could be identified and co-expressed in *E. coli* for the *in vivo* oxidation of isobutyryl-CoA to methacrylyl-CoA, as the probable sequences for the *Pseudomonas aeruginosa* ETFQO (UniProt: Q9HZP5), ETF α (UniProt: Q9HZP7) and ETF β (UniProt: Q9HZP6) polypeptides are known. These probable ETFQO, ETF α and ETF β sequences share 54%, 63% and 58% sequence identity, and 67%, 75% and 73% sequence similarity, with their human counterparts.

Cloning components for the transfer of electrons from an ACAD to ubiquinone, including a membrane protein, into *E. coli* is risky. For the eukaryotic proteins that have been experimentally confirmed as being isobutyryl-CoA dehydrogenases, ETFs or ETFQOs, there is the risk that the proteins won't express well in *E. coli*. On the other hand, for the putative system from *P. aeruginosa*, none of the enzymes have been experimentally proven to perform the function that they are annotated to perform.

5.2.3 Identification of an acyl-CoA oxidase for the oxidation of isobutyryl-CoA in *E. coli*.

As previously discussed, fatty acid β -oxidation in plants is a predominantly, if not entirely, peroxisomal affair, with H_2O_2 producing acyl-CoA oxidases being responsible for the oxidation of a fatty acyl-CoA to its *trans*-2-enoyl-CoA product [187]. The analogous peroxisomal degradation of branched chain amino acids in plants has previously been suggested [83, 84], whereby the peroxisomes of *Vigna radiata* have been shown to produce intermediates of the catabolism of all three branched chain amino acids, when fed with the 2-keto-acid derivatives. Most importantly to this project, when fed with 2-ketoisovalerate, the peroxisomes from this organism were shown to produce methacrylyl-CoA, indicating that an acyl-CoA oxidase with activity on isobutyryl-CoA exists in this organism. However, no amino acid or nucleotide sequence could be identified for the enzyme responsible for this activity. Interestingly, Gerbling also makes reference to an acyl-CoA oxidase from *Candida sp.* with activity on isobutyryl-CoA [84], but the exact oxidase being referred to could not be identified.

Branched chain amino acid degradation in *Arabidopsis thaliana*, a model plant organism, however, appears to be slightly different, with leucine degradation being an entirely mitochondrial affair. Mitochondrial branched chain keto acid dehydrogenase and isovaleryl-CoA dehydrogenase from this organism do not only show activity on the derivatives of leucine catabolism, but also on those of valine degradation, giving some indication that valine could be physiologically degraded in the mitochondria. A short chain acyl-CoA oxidase, ACX4, has been identified in *Arabidopsis* peroxisomes, which on top of its C4-C8 activity, has also demonstrated activity on isobutyryl-CoA [188]. In addition, ACX4 has been previously successfully expressed in *E. coli* [189, 190]. Therefore, despite the risks associated with the production of hydrogen peroxide, as *E. coli* is a catalase positive organism [173], cloning this enzyme would be considerably less risky than cloning any of the potential isobutyryl-CoA, ETF and ETFQO combinations described previously.

5.2.4 Expression of ACX4 and preparation of cell free extracts

The enzyme, short chain acyl-CoA oxidase (ACX4), from *Arabidopsis thaliana* (Genbank accession number AB017643.1, Uniprot accession number Q96329, EC number 1.3.3.6, Polypeptide sequence 7) was codon optimised for expression in *E. coli* and synthesised with an *Nde*I restriction site integrated at the 5' end and a *Xho*I restriction site at the 3' end. The synthesised polynucleotide (Polynucleotide sequence 5) was delivered in the pMA-RQ::ACX4 plasmid and the gene insert was subcloned into a previously linearised pET20b(+) vector, at the *Nde*I and *Xho*I restriction sites, to form pET20b(+):ACX4. The newly constructed pET20b(+):ACX4 plasmid was then transformed into *E. coli* BL21 (DE3) pLysS to form *E. coli* BL21 (DE3) pLysS pET20b(+):ACX4. The expression of ACX4 was determined in both LB media and MSX media, at both 28 °C and 37 °C for each.

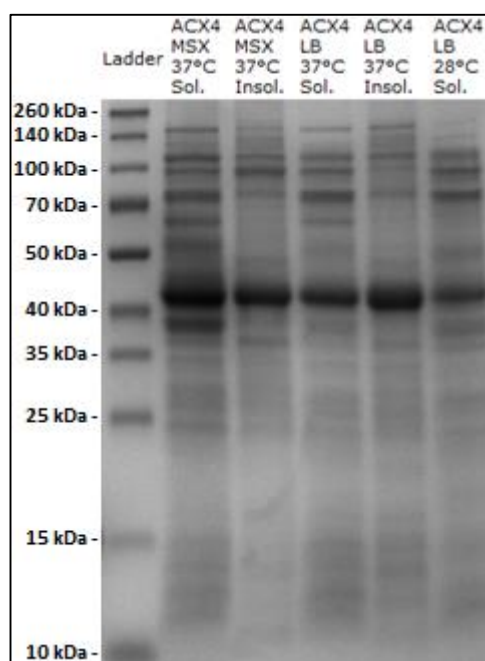


Figure 41: ACX4 expression in both MSX media at LB media. Lane 1: Ladder. Lane 2: ACX4 in MSX+riboflavin at 37°C, lane 3: ACX4 in LB+riboflavin, 37°C soluble. Lane 4: ACX4 in LB+riboflavin, 37°C Insoluble. Lane 5: ACX4 in LB+riboflavin, 28°C soluble. 20ml seed cultures were used. LB media seed cultures were supplemented with riboflavin, glucose (1%), chloramphenicol and carbenicillin. MSX seed cultures were supplemented with chloramphenicol and carbenicillin. 100ml Cultures were all supplemented with riboflavin, chloramphenicol and carbenicillin, with LB cultures supplemented with additional glucose (1%). All cultures were incubated with shaking at 250rpm and induced at OD₆₀₀ 0.7.

The expression of ACX4 at 37°C in both MSX and LB media showed similar soluble expression levels (Figure 41), with perhaps slightly more in MSX. However, far less soluble protein was expressed in LB media at 28 °C, but similar expression levels were observed in MSX medium at both 37 °C and 28 °C (not shown). For crude enzyme assays, expression was therefore scaled up to 500 ml scale using MSX medium supplemented with riboflavin, chloramphenicol and carbenicillin.

Expression was induced at OD₆₀₀ 0.7, and the culture was incubated for a further 7 h post induction before cells were harvested and cell free extracts prepared in HEPES buffer supplemented with FAD (10 μ M).

5.2.5 ACX4 enzyme assays

Before the enzyme assays were performed, an analytical HPLC method was first developed to resolve isobutyryl-CoA (IB-CoA), methacrylyl-CoA (MAA-CoA), FAD, methacrylic acid (MAA) and coenzyme A (CoA-SH). Coenzyme A, methacrylyl-CoA and isobutyryl-CoA eluted at 11.0 min, 30.8 min and 32.2 min respectively. Coenzyme A, methacrylyl-CoA and isobutyryl-CoA each had a small tailing (minor) peak associated with the main peak at 12 min, 31.6 min and 32.9 min, respectively (Figure 42). Methacrylic acid eluted at 13.5 min, and FAD, used in the crude assays of ACX4 and as an internal standard for the isobutyryl-CoA and methacrylyl-CoA standards, eluted at 27.8 min. Some impurities in the commercially purchased coenzyme A and isobutyryl-CoA samples were detected by HPLC, with small peaks in the coenzyme A trace also being observed at 5 min, 7.14 min, 13.7 min and 15.6 min, as well as a shallow and broad peak in the isobutyryl-CoA trace at 23.4 min.

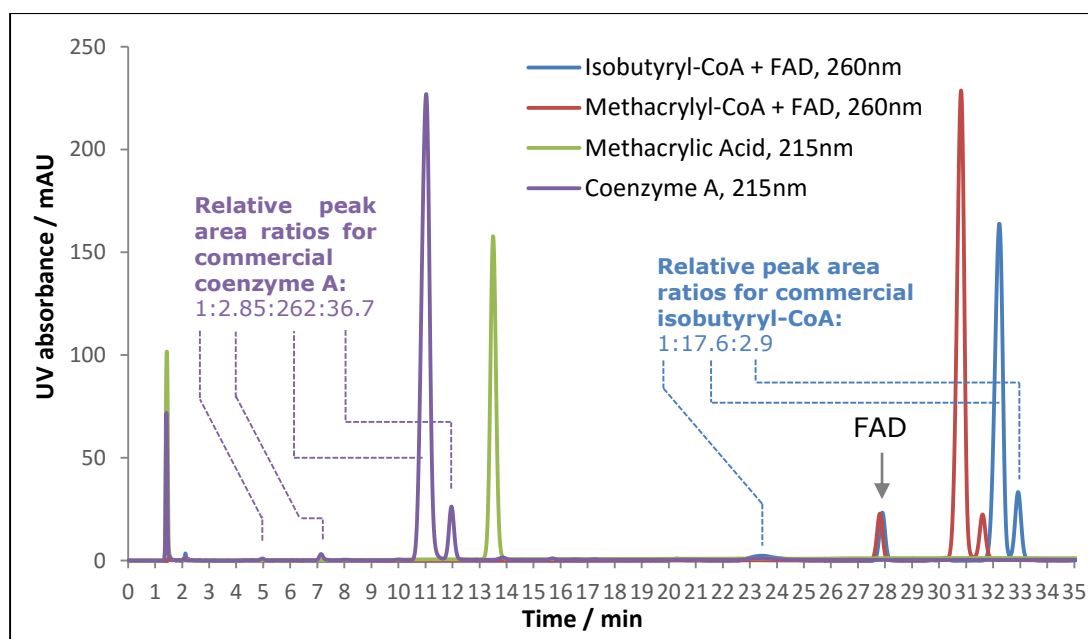


Figure 42: Overlaid signals showing analytical reverse phase HPLC resolution of coenzyme A, methacrylic acid, flavin adenine dinucleotide, methacrylyl-CoA and isobutyryl-CoA. Methacrylic acid and coenzyme A were detected at 215 nm, whereas FAD, methacrylyl-CoA and isobutyryl-CoA were detected at 260 nm.

To test the activity of ACX4, enzyme assays composed of ACX4 cell free extract (0.8 mg.ml⁻¹), FAD (10 μ M) and isobutyryl-CoA (500 μ M) were performed in HEPES buffer. The sample was incubated at 30 °C with shaking at 250 rpm for 30 min, and analysed by HPLC (Figure 43). Methacrylyl-CoA was the major product, with a

peak at 30.7 min. No isobutyryl-CoA was detected, though some methacrylyl-CoA hydrolysis was observed, and this could be attributed to background thioesterases. The concentration of methacrylic acid formed during the crude assay was 74 μM .

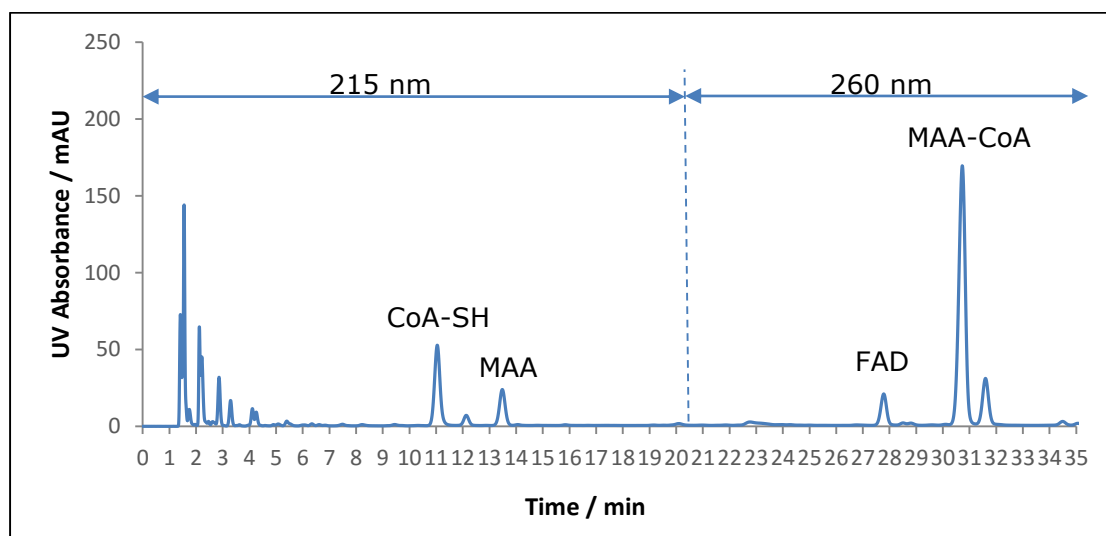


Figure 43: HPLC analysis of ACX4 crude extract incubated with isobutyryl-CoA. A strong peak eluting at the same time as methacrylyl-CoA appears along with smaller peaks for coenzyme A and methacrylic acid.

In order to confirm that the methacrylyl-CoA peak was genuine, and that it was not an isobutyryl-CoA peak with a shifted elution time, the sample was spiked with isobutyryl-CoA and analysed by HPLC again. ACX4 was inactivated when the sample was acidified to stop the reaction, and this ensured that any additional isobutyryl-CoA would not be converted to methacrylyl-CoA. Indeed, the isobutyryl-CoA spiked sample showed not only the original methacrylyl-CoA peak, but also an additional peak at 32 min with the characteristic tail peak of isobutyryl-CoA (Figure 44).

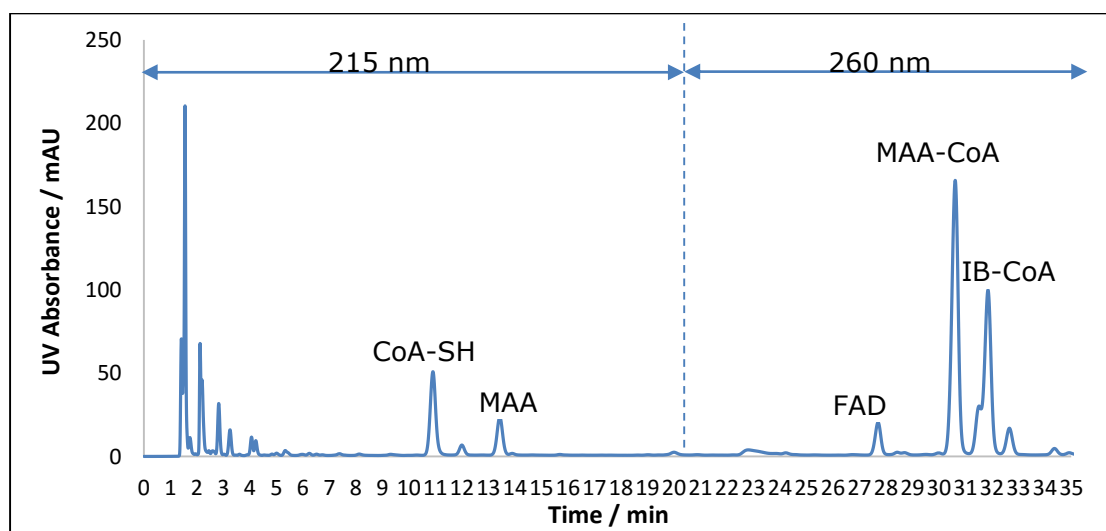


Figure 44: HPLC analysis of an isobutyryl-CoA spiked sample of methacrylyl-CoA that was generated by the incubation of ACX4 with isobutyryl-CoA.

To determine whether methacrylic acid could be produced from isobutyryl-CoA in an enzyme coupled reaction, an experiment was established whereby crude ACX4 was incubated with purified CthHis-4HBT enzyme. The same cell free extract of ACX4 was used as the protein source for the ACX4, though pure CthHis-4HBT was prepared again, in the same manner as for its kinetic characterisation, but performing a buffer exchange into HEPES buffer (50 mM, pH 7.5) at the end, instead of the previously used phosphate buffer. Thus, crude ACX4 (0.8 mg.ml^{-1}) was co-incubated with pure CthHis-4HBT (0.6 mg.ml^{-1}) along with FAD ($10 \text{ }\mu\text{M}$) and isobutyryl-CoA ($500 \text{ }\mu\text{M}$) in HEPES buffer. The sample was incubated at $30 \text{ }^{\circ}\text{C}$ with shaking at 250 rpm for 30 min, and analysed by HPLC. Methacrylic acid and coenzyme A were the major products. The methacrylic acid peak was observed at 13.45 min whilst the coenzyme A major and minor peaks were observed at 11.1 min and 12.1 min, respectively. The concentration of methacrylic acid generated during the coupled enzyme reaction (Figure 45) was $345 \text{ }\mu\text{M}$, 4.7 times greater than that during the crude ACX4 assay alone.

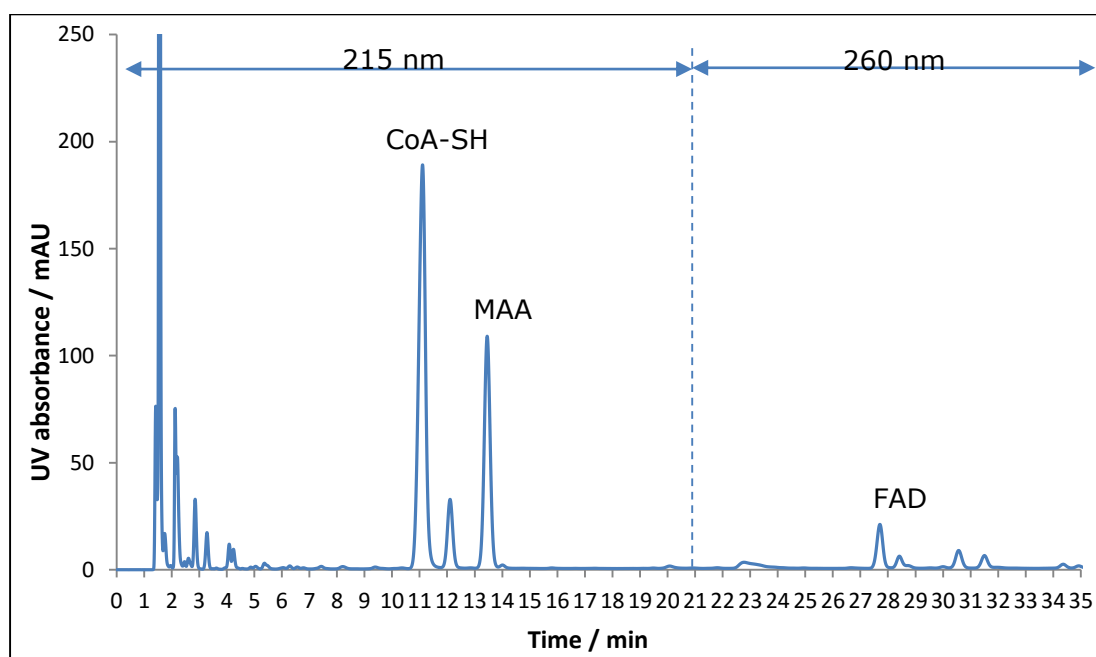


Figure 45: HPLC analysis of the product from the coincubation of crude ACX4 and pure CthHis-4HBT with isobutyryl-CoA. Major peaks appear for coenzyme A and methacrylic acid.

The isobutyryl-CoA used in this experiment was from a different preparation as that used during the thioesterase assays. It is therefore unclear whether the $500 \text{ }\mu\text{M}$ isobutyryl-CoA as a starting concentration is indeed accurate, or whether it is a high estimate, as was the case during the thioesterase assays. Thus, the conversion of isobutyryl-CoA to methacrylic acid is 69% when assuming that the starting concentration of isobutyryl-CoA is accurate. The conversion could be higher, if as

was the case during the thioesterase assays, the actual concentration of isobutyryl-CoA in the new preparation was lower than expected.

To ensure that coenzyme A, methacrylic acid, isobutyryl-CoA and methacrylyl-CoA were not be confused with peaks from the cell free extract, a no-substrate control was performed, using ACX4 cell free extract (0.8 mg.ml⁻¹), purified CtHis-4HBT (0.6 mg.ml⁻¹) and FAD (10 µM) in HEPES buffer (Figure 46). No peaks were observed eluting at the corresponding times for coenzyme A, methacrylic acid, methacrylyl-CoA or isobutyryl-CoA.

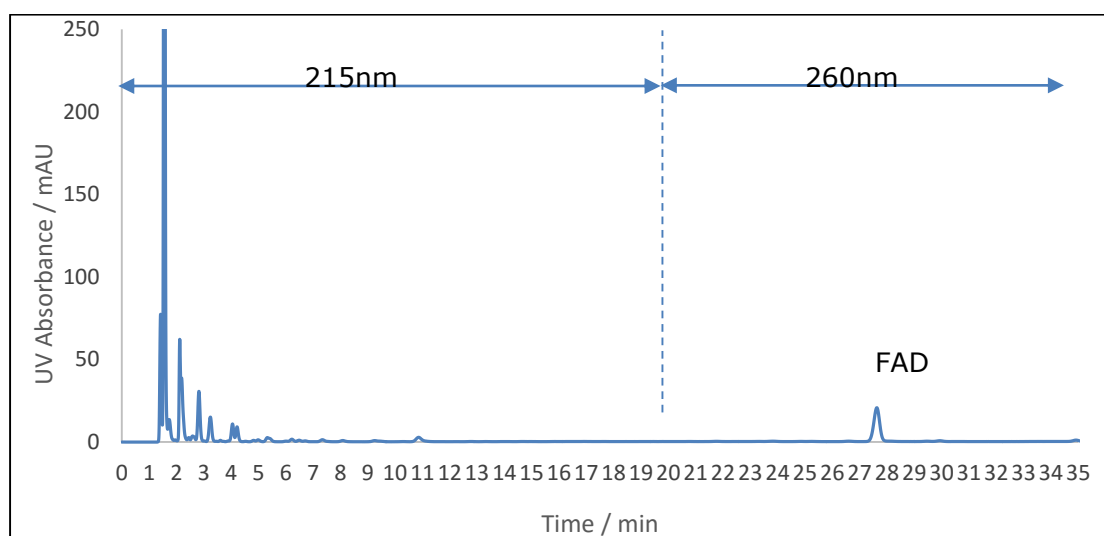


Figure 46: No-substrate control HPLC analysis of acidified and clarified assay media containing ACX4 crude extract (0.8 mg.ml⁻¹), 4HBT (0.6 mg.ml⁻¹) and FAD (10 µM).

This confirms that ACX4 oxidises isobutyryl-CoA when overexpressed in *E. coli*, and that an enzyme couple reaction between ACX4 and 4HBT can convert isobutyryl-CoA to methacrylic acid and coenzyme A. The next stage was thus to test the activity of these two enzymes *in vivo* for the production of methacrylic acid from renewable resources.

5.3 The biotransformation of isobutyric acid to methacrylic acid.

The successful construction of a metabolic pathway for the production of methacrylic acid required the linking of the oxidation of isobutyryl-CoA by ACX4 and the hydrolysis of methacrylyl-CoA by 4HBT to the host organism's metabolic network. The obvious choice to achieve this would have been to engineer the oxidative decarboxylation of 2-ketoisovaleric acid to isobutyryl-CoA into cells expressing ACX4 and 4HBT. However, Zhang *et al.* had produced isobutyric acid to titres of 90 g.L^{-1} from glucose using a pathway that decarboxylated 2-ketoisovalerate to isobutyraldehyde, and then oxidised the isobutyraldehyde to isobutyrate. It is also believed that isobutyric acid is activated to isobutyryl-CoA in various organisms [86, 96-98]. Thus, a biotransformation of isobutyric acid to methacrylic acid was proposed as a simple biocatalytic pathway to methacrylic acid from a bio-based feedstock, in which isobutyric acid would first be activated to isobutyryl-CoA, before ACX4 and 4HBT converted the isobutyryl-CoA to methacrylic acid. The most likely candidates for the activation of isobutyric acid to isobutyryl-CoA in nature are the AMP forming acyl-CoA synthetases, as is the case in fatty acid activation for fatty acid β -oxidation [87], or the concerted action of an isobutyrate kinase and phosphate isobutyryl transferase [98]. Thus, two candidate biotransformation routes were proposed with the two alternative routes for the activation of isobutyric acid (Figure 47A&B).

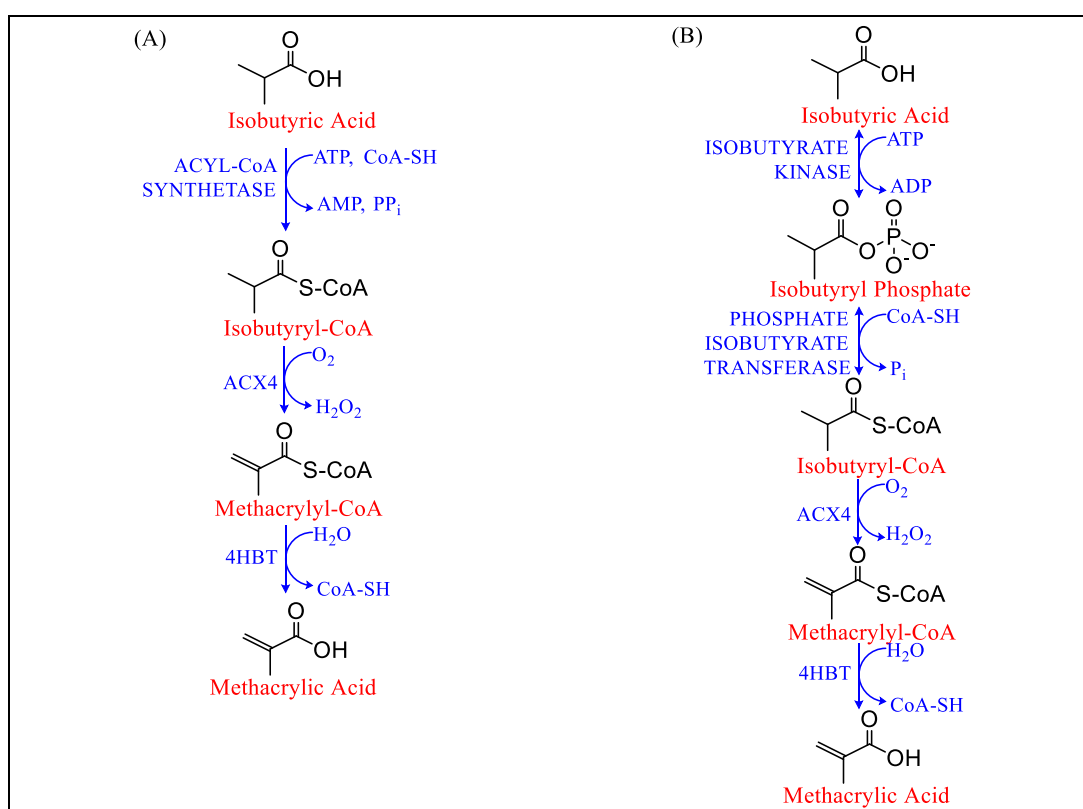


Figure 47: Candidate routes for the biotransformation of isobutyric acid to methacrylic acid.

The route in which the activation step is catalysed by an AMP-forming acyl-CoA synthetase may well be the favoured route, as the pyrophosphate is hydrolysed *in vivo*, which is a highly exergonic step, driving the activation step to completion [87]. As phosphate butyrate transferases and butyrate kinase enzymes catalyse reversible reactions [191], it is very likely that the phosphorylation of isobutyrate and the transfer of coenzyme A to isobutyryl-CoA will both also be reversible steps, and so won't be as thermodynamically favourable as the AMP forming acyl-CoA synthetase catalyse reaction. However, as a suitable isobutyryl-CoA ligase had not, until this stage, been identified, all biocatalytic routes were considered, and so a database and literature search was performed to identify suitable acyl-CoA synthetases, as well as suitable fatty acid kinases and phosphate acyl transferases.

5.3.1 Identification of an enzyme or enzyme system for the activation of isobutyric acid to isobutyryl-CoA

A database and literature search was performed to identify candidate enzymes for the activation of isobutyric acid to isobutyryl-CoA, whether directly or *via* isobutyryl-phosphate. For the activation of isobutyric acid to isobutyryl-phosphate, various fatty acid kinases belong to the subsections of EC parent group 2.7.2, with a branched chain fatty acid kinase specifically belonging to EC 2.7.2.14. Other carboxylic acid kinases such as butyrate kinase, propionate kinase and acetate kinase could also show fortuitous activity on isobutyric acid and so these EC groups were also considered in the database and literature search (Table 21). For the phosphate acyl-transferase type enzymes, phosphate butyryl transferases, phosphate acetyl transferases and phosphate propanoyl transferases were considered. As for AMP-forming acyl-CoA synthetases, these span various level 4 groups of the EC 6.2.1 parent group. A further study of enzyme commission ontology on the MetaCyc database, searching for subgroups of the EC 6.2.1 family, revealed that in addition to AMP-forming acyl-CoA synthetases, both ADP- and GDP-forming acyl-CoA synthetases exist, and the search was expanded to include these too.

Role		EC group	EC group name
Activation of isobutyric acid to isobutyryl-CoA via a fatty acid kinase and phosphate acyl transferase.	Fatty acid kinase	2.7.2.14	Branched chain fatty acid kinase
		2.7.2.7	Butyrate kinase
		2.7.2.1	Acetate kinase
		2.7.2.15	Propionate kinase
	Phosphate acyl transferase	2.3.1.19	Phosphate butyryl transferase
		2.3.1.8	Phosphate acetyl transferase
		2.3.1.222	Phosphate propanoyl transferase
Activation of isobutyric acid to isobutyryl-CoA by an acyl-CoA ligase		6.2.1.1	Acetate-CoA ligase (AMP forming)
		6.2.1.2	Butyrate-CoA ligase (AMP forming)
		6.2.1.10	Carboxylic acid-CoA ligase (GDP forming)
		6.2.1.13	Acetate-CoA ligase (ADP forming)
		6.2.1.17	Propionate-CoA ligase (AMP forming)

Table 21: EC groups to consider for the activation of isobutyric acid to isobutyryl-CoA

The database search was therefore performed to identify candidate enzymes in the above enzyme commission groups (Table 21) for enzymes that could activate isobutyric acid to isobutyryl-CoA, either directly or *via* isobutyryl-phosphate. The results of the database search are summarised below (Table 22). The literature verification of the database results showed that there was an additional enzyme, AcsA from *P. chlororaphis* B23 [192], that was not listed in the databases under the EC numbers that were searched.

Enzyme type	Enzyme Name	Source	Enzyme identified by
Acyl-CoA ligase	AcsA	<i>Pseudomonas chlororaphis</i> B23	Literature search [192]
	Butyryl-CoA ligase	<i>Paecilomyces varioti</i> AHU 9417	BRENDA (EC: 6.2.1.2) [193]
	Butyryl-CoA ligase	Bovine heart mitochondria	BRENDA (EC: 6.2.1.2) [194]
Fatty acid kinase	Branched chain fatty acid kinase	<i>Spirochete</i> MA-2	BRENDA (EC: 2.7.2.14) [195]
	Butyrate kinase	<i>Clostridium butyricum</i>	BRENDA (EC: 2.7.2.7) [196]
Phosphate acyl transferase	Phosphate butyryl-transferase	<i>Clostridium acetobutylicum</i> ATCC824	BRENDA (EC: 2.3.1.19) [197]

Table 22: Database and literature search results for candidate enzymes for the activation isobutyric acid to isobutyryl-CoA.

AcsA is an AMP forming, short-chain acyl-CoA ligase from *Pseudomonas chlororaphis* [192]. When overexpressed in *E. coli* at 28°C for 12h and assayed in a cell free extract, AcsA showed greatest catalytic efficiency (k_{cat}/k_m) for isobutyrate out of the substrates tested (Table 23).

Substrate	K_m mM	V_{max} / $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$	$K_{cat}\cdot K_M$ / $\text{s}^{-1}\cdot\text{mM}^{-1}$
Acetic acid	10.6 $\pm$ 1.9	1.78 $\pm$ 0.22	0.168
Propionic acid	1.10 $\pm$ 0.21	12.4 $\pm$ 1.6	11.3
<i>n</i> -Butyric acid	0.32 $\pm$ 0.04	15.1 $\pm$ 0.8	47
<i>n</i> -Valeric acid	0.42 $\pm$ 0.06	8.59 $\pm$ 0.25	20
Hexanoic acid	0.88 $\pm$ 0.10	0.427 $\pm$ 0.015	0.49
Heptanoic acid	2.80 $\pm$ 0.53	0.0083 $\pm$ 0.0003	0.0030
Isovaleric acid	2.32 $\pm$ 0.85	1.16 $\pm$ 0.10	0.500
DL-3-Hydroxybutyric acid	5.86 $\pm$ 2.03	0.135 $\pm$ 0.015	0.0230
DL-2-Hydroxybutyric acid	1.98 $\pm$ 0.35	0.674 $\pm$ 0.008	0.340
<i>trans</i> -Crotonic acid	0.075 $\pm$ 0.019	0.185 $\pm$ 0.006	2.5
Isobutyric acid	0.14 $\pm$ 0.07	10.7 $\pm$ 0.20	76
Acrylic acid	2.34 $\pm$ 0.65	3.06 $\pm$ 0.20	1.31
Benzoic acid	1.33 $\pm$ 0.15	0.255 $\pm$ 0.002	0.192

Table 23: Substrates activated to coenzyme A derivatives by AcsA from *P. chlororaphis*. Adapted from [192].

Butyryl-CoA ligase from *P. varioti* also showed activity on isobutyric acid [193], with a V_{MAX} 30% of that when using butyrate as a substrate. The Michaelis constant was 1.6 mM compared to 0.14 mM for butyrate (Table 24).

Substrate	Vmax*	K _M / mM
Acetate	0	-
Propionate	76	3.3
Butyrate	100	0.14
Isobutyrate	30	1.6
Valerate	71	0.38
Isovalerate	16	0.47
Caproate	0	-
Octanoate	0	-

Table 24: Kinetic data for butyryl-CoA synthetase for *P. varioti*. * = V_{MAX} given as a percentage of activity compared to when using butyric acid as a substrate. Adapted from [193].

Butyryl-CoA synthetase from bovine heart mitochondria [194] shows an activity on isobutyrate to be 20% that of its activity on butyrate, but no Michaelis constant is provided for this enzyme.

The activation of isobutyrate to isobutyryl-phosphate has been shown by branched chain fatty acid kinase from *Spirochete* MA-2 [195]. The enzyme is active on isobutyrate, 2-methylbutyrate, isovalerate, butyrate, valerate and propionate, but not with acetate. The enzyme was half as active with the branched chain substrates as with the straight chain substrates, but the lowest K_m value was for isobutyrate.

Substrate	Apparent K _M / mM
Propionate	14.3
Isobutyrate	4.3
Isovalerate	9.5
2-Methylbutyrate	10.8
Butyrate	16.9
Valerate	12.5

Table 25: Kinetic data for branched chain fatty acid kinase. [195]. No K_{cat} or V_{max} data are available.

A similar carboxylic acid kinase from *Clostridium butyricum* is active on propionate and butyrate to similar extents, with 50% of that activity shown using isobutyrate as a substrate. Again, no kinetic data was provided. As for an enzyme to catalyse the exchange of the phosphate on isobutyryl-phosphate for a coenzyme A moiety, such that isobutyryl-phosphate is formed, a phosphotransbutyrylase from *Clostridium acetobutylicum* ATCC824 [197] shows activity with isobutyryl-CoA as a substrate. This activity is 30% that of the activity with butyryl-CoA.

From the above data, AcsA from *P. chlororaphis* B23 was chosen over the other AMP forming acyl-CoA synthetases as well as any of the possible phosphotransbutyrylase/butyrate kinase combinations primarily due to its very low

K_M value for isobutyric acid. At just 0.14 mM, the enzyme is likely to be saturated in a metabolic pathway which upregulates isobutyric acid, or in a biotransformation whereby isobutyric acid is supplemented to the media and will therefore be functioning at maximum capacity. The higher K_M values of the AMP-forming acyl-CoA synthetase from *P. varioti* as well as the fatty acid kinase from *Spirochete* MA-2 for isobutyric acid, means that they are unlikely to function at full capacity, and are more likely to participate in side reactions. AcsA is also the only enzyme with quantitative V_{MAX} data. At a V_{MAX} of $10 \mu\text{mol}.\text{min}.\text{mg}^{-1}$, the maximum specific activity is over 20 times that of 4HBT on methacrylyl-CoA, and so long as there is good expression of this enzyme, there should be sufficient activity to drive the metabolic pathway. Finally, there is empirical evidence that AcsA from *P. chlororaphis* B23 can be expressed in *E. coli* [192], which is not the case for many of the other enzymes.

5.3.2 A whole cell biotransformation of isobutyric acid to methacrylic acid

5.3.2.1 Construction of an expression vector for the co-expression of 4HBT, ACX4 and AcsA

To perform the biotransformation of isobutyric acid to methacrylic acid, a single vector capable of expressing 4HBT, ACX4 and AcsA from a single operon was constructed. Whilst the alternative would have been to use separate plasmids with compatible origins of replication and different antibiotic resistance genes [198], there is plenty of precedent for the use of single expression operons for the co-expression of multiple genes. In addition, it was foreseen that additional genes may need to be cloned and co-expressed in future experiments in order to meet the next objective of producing methacrylic acid directly from glucose. As there is a practical limit on the number of inducible promoters and the number of plasmids present in a host organism, collecting common genes of different metabolic pathways together into single expression cassettes at an early stage was deemed advantageous.

As 4HBT and ACX4 had both been successfully expressed under the T7 promoter of the pET20b(+) vector in *E. coli* BL21 (DE3) pLysS, and AcsA has been reported to be functionally expressed under a similar expression system [192], the same pET20b(+) expression system was kept for the co-expression of the three genes. As such, a single operon in the pET20b(+) vector, containing the 4HBT, ACX4 and AcsA genes, each preceded by their own ribosome binding sites, was constructed, and this was performed in two stages.

The first stage was the construction of a pET20b(+) based vector for just 4HBT and ACX4, pET20b(+):4HBT-ACX4, and the second was the ligation of AcsA into this vector to form pET20b(+):4HBT-ACX4-AcsA. To construct pET20b(+):4HBT-ACX4, an overlap extension polymerase chain reaction was performed to join the 4HBT gene and the ACX4 gene together into a single, double-stranded polynucleotide and to introduce a new ribosome binding site before the ACX4 gene. To perform the overlap extension PCR, 4HBT and ACX4 were first amplified out of the previously prepared pET20b(+):4HBT and pET20b(+):ACX4 vectors in two separate PCR reactions, using primers OE_4HBT_F and OE_4HBT_R for 4HBT in one reaction, and primers OE_ACX4_F and OE_ACX4_R for ACX4 in another reaction (Figure 48A).

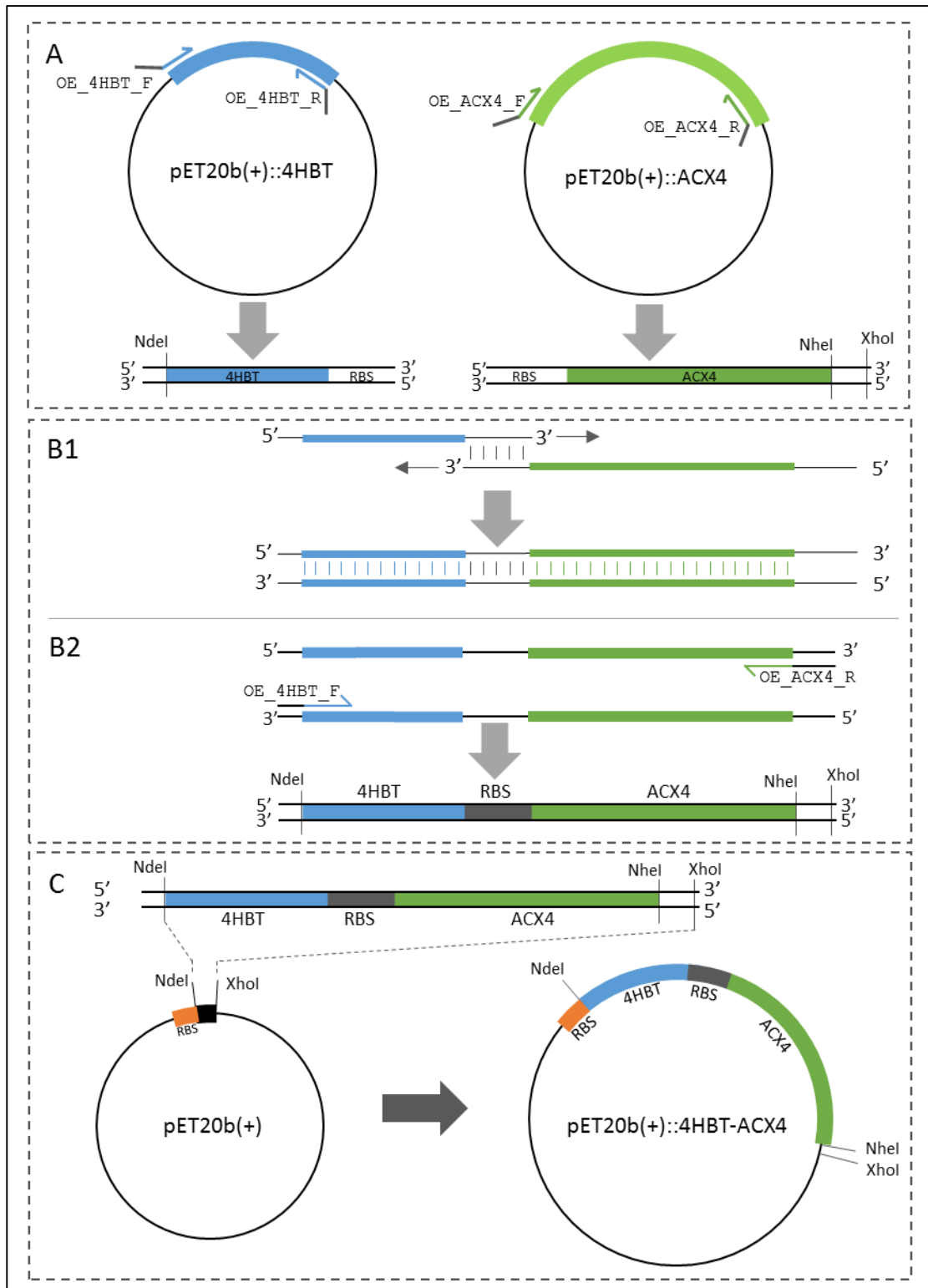


Figure 48: Construction of pET20b(+):4HBT-ACX4, an expression vector for the co-expression of 4HBT and ACX4, containing a new cloning site downstream of the 4HBT and ACX4 genes. A) Amplification of the 4HBT and ACX4 genes with OE_4HBT_F primer containing *NdeI* restriction site, OE_ACX4_R containing *NheI* and *XhoI* restriction sites, and primers OE_4HBT_R and OE_ACX4_F containing overlapping sequences with a new ribosome binding site for the expression of ACX4. B1) Overlap extension phase 1 – anneal complementary sequences of product A and product B and extend. B2) Add primers OE_4HBT_F and OE_ACX4_R and amplify the single polynucleotide encoding 4HBT and ACX4. C) Clone the 4HBT-ACX4 polynucleotide (Polynucleotide sequence 7) into empty pET20b(+) vector between *NdeI* and *XhoI* sites to construct pET20b(+):4HBT-ACX4.

Primer OE_4HBT_F maintained the *NdeI* restriction site immediately upstream of the 4HBT gene and primer OE_ACX4_R introduced a short downstream sequence containing *NheI* and *XhoI* restriction sites, separated by an adenine and thymine rich spacer sequence. The spacer sequence was added to ensure efficient double digestion with *NheI* and *XhoI*, in order to facilitate the insertion of AcsA. The spacer sequence also helped lower the primer melting temperature in this otherwise guanine and cytosine rich region. Primers OE_4HBT_R and OE_ACX4_F, on the other hand, introduced complementary sequences in the downstream region of PCR product A and the upstream region of PCR product B, to enable the self-priming required between the melted single strands of PCR products A and B in the initial phase of the overlap extension PCR (Figure 48B1), the overlap extension phase. A significant portion of the complementary sequence included a new intergenic sequence between 4HBT and ACX4, containing the new ribosome binding site (RBS) for ACX4 and an 8 base-pair spacer sequence between the RBS and the ACX4 start-codon for efficient translation initiation (Figure 49).

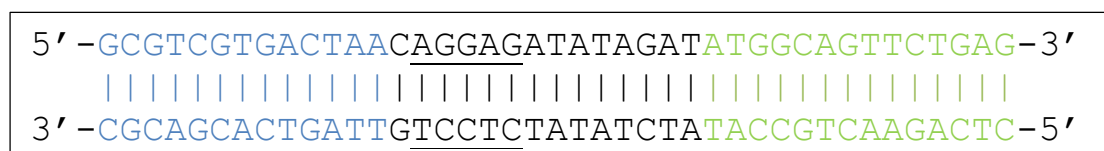


Figure 49: The end of the 4HBT gene (blue), the intergenic sequence (black) containing a new ribosome binding site for the ACX4 gene (underlined) and the start of the ACX4 gene (green). Top strand = coding (sense) strand. Bottom strand = template (anti-sense) strand.

The product of the overlap extension phase was then amplified by the addition of the OE_4HBT_F and OE_ACX4_R primers in the second phase of the actual overlap extension PCR, the chain reaction phase (Figure 48B2). The final product of the overlap extension polymerase chain reaction was then cloned into the pET20b(+) plasmid between the *NdeI* and *XhoI* restriction sites to form pET20b(+):4HBT-ACX4 (Figure 48C). The pET20b(+) vector provided the RBS for 4HBT.

Acyl-CoA synthetase (AcsA) from *Pseudomonas chlororaphis* was codon optimised and synthesised by BioMatik corporation with a *XhoI* restriction site at the 3' end and a (5'-GCTAGC**AGGAG**ATATACAT-3') sequence just prior to its start codon. This sequence contained an *NheI* restriction site (underlined), a ribosome binding site (bold), and a spacer sequence for efficient translation initiation just prior to the start codon. The 5' -CAT-3' nucleotides (double underline) of the spacer sequence and the 5' -ATG-3' start codon of AcsA combine to form an *NdeI* restriction site (5' -CATATG-3').

Thus, AcsA was cloned between the *NheI* and *XhoI* restriction sites of pET20b(+):4HBT-ACX4 to form pET20b(+):4HBT-ACX4-AcsA (Figure 50). It was also cloned between the *NdeI* and *XhoI* sites of the pET20b(+) vector, to form pET20b(+):AcsA for the expression of just the AcsA protein.

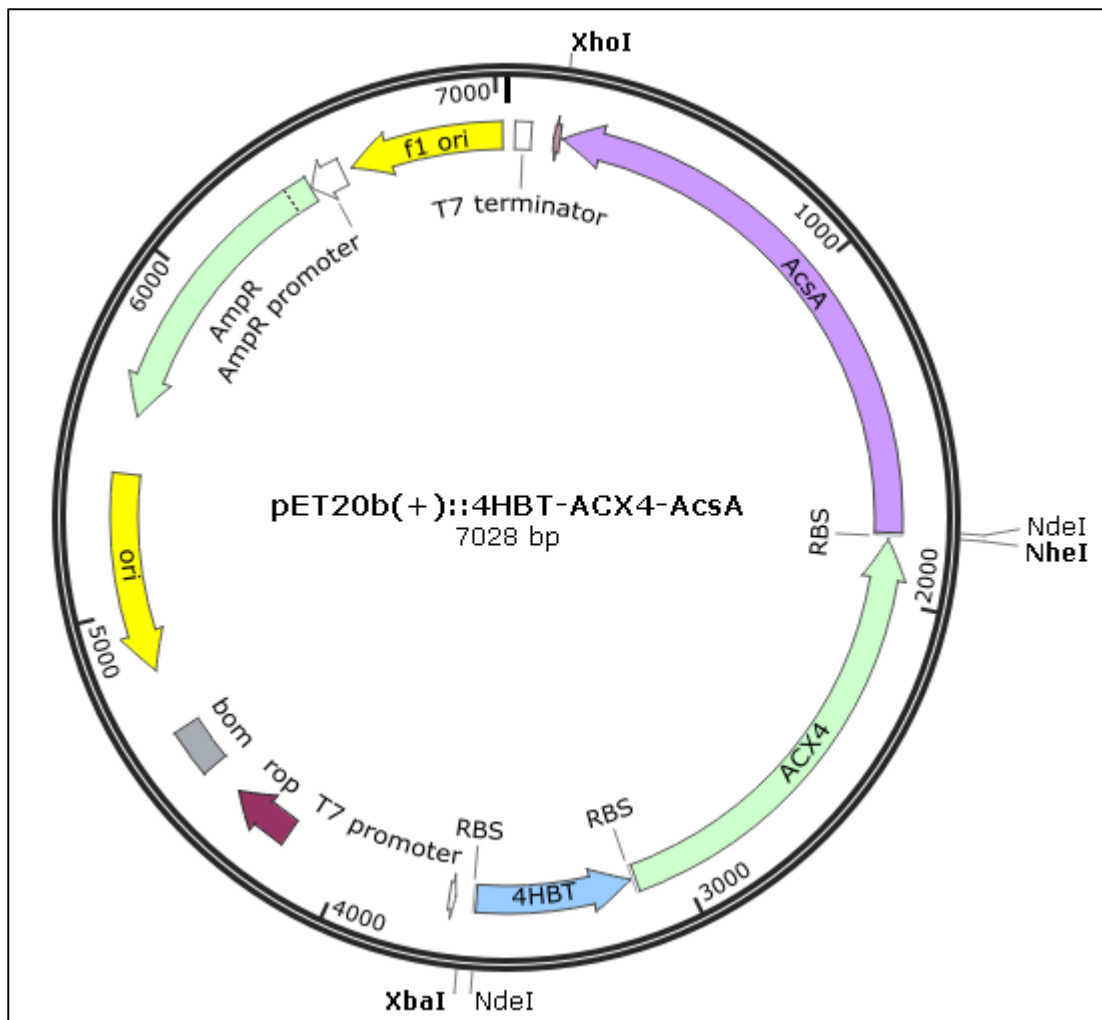


Figure 50: The pET20b(+):4HBT-ACX4-AcsA plasmid. The expression plasmid was amplified by culturing *E. coli* JM107 in LB (400ml), and the plasmid was purified from this culture using the QIAGEN Plasmid MaxiPrep kit.

The sequence of the operon was confirmed using sequencing primers *T7_F*, *pHAAc_S1*, *pHAAc_S2* and *T7_R*, and the pET20b(+):4HBT-ACX4-AcsA vector was then transformed into *E. coli* BL21 (DE3) pLysS, forming *E. coli* BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-AcsA. The same was done for the pET20b(+):AcsA plasmid, to form *E. coli* BL21 (DE3) pLysS pET20b(+):AcsA. The presence of the plasmids in the transformed colonies was confirmed in PCR reactions using KOD DNA polymerase. Plasmid mini preps (1µl) of the transformants were used as the DNA template, with the *T7_F* and *T7_R* primers, an annealing temperature of 50 °C and extension time of 80 s.

5.3.2.2 Confirming the coexpression of 4HBT, ACX4 and AcsA in *E. coli*.

The co-expression of 4HBT, ACX4 and AcsA by *E. coli* BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-AcsA was confirmed by incubating cultures (100ml) of *E. coli* BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-AcsA in MSX media supplemented with riboflavin (1 mg.L^{-1}), to ensure the expression of active ACX4, at two different test temperatures, 37 °C and 28 °C. Cultures were grown to an OD₆₀₀ of 0.5 and expression was induced with the addition of IPTG (0.4 mM). The best expression was observed for the culture grown at 37 °C (Figure 51), as expression at 28 °C yielded significantly less ACX4 protein, with no improvements in protein solubility for either 4HBT, ACX4 or AcsA.

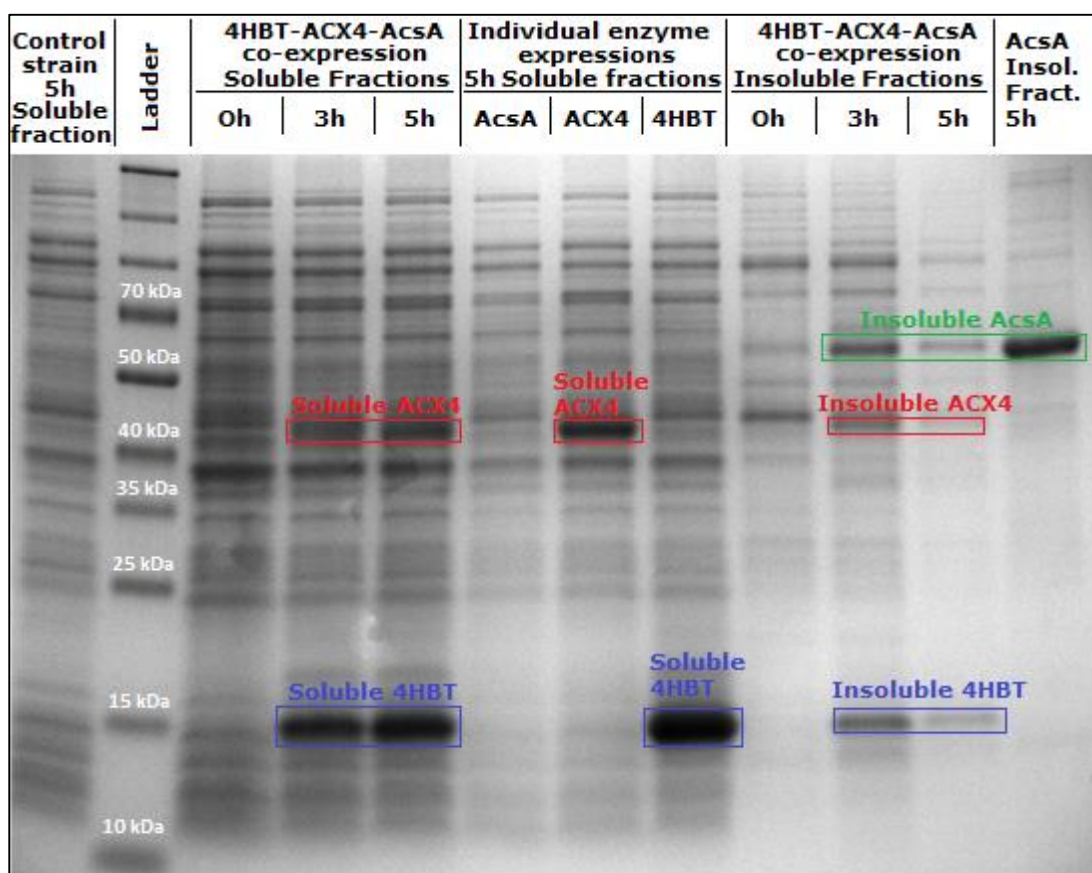


Figure 51: Coexpression of 4HBT, ACX4 and AcsA in MSX medium with riboflavin at 37 °C.

Analysis by SDS-PAGE showed that when expression was performed at 37 °C, high levels of 4HBT (blue boxes) were observed in the soluble fractions at 5 h post induction, and good levels of ACX4 (red boxes) were detected with very little insoluble protein detected. However, high levels of insoluble AcsA (green box) were detected with no AcsA protein visible in the soluble fractions. Nevertheless, AcsA is reported to have a good maximum specific activity for isobutyric acid as a substrate ($10 \text{ } \mu\text{mol.min}^{-1}.\text{mg}^{-1}$). It was therefore postulated that AcsA may have sufficiently high activity to catalyse the activation of isobutyric acid to isobutyryl-CoA even at

protein levels below the detection limit of the SDS-PAGE procedure. Therefore, a preliminary experiment was performed to test the ability of *E. coli* BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-AcsA to convert isobutyric acid to methacrylic acid, in which three separate negative controls were performed, but in which the entire experiment was not performed in triplicate.

5.3.2.3 Production of methacrylic acid in cultures supplemented with isobutyric acid

An overnight seed culture of *E. coli* BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-AcsA was prepared, and used to inoculate MSX medium (100 ml) supplemented with riboflavin, chloramphenicol and carbenicillin. The culture was incubated at 37 °C with shaking at 200 rpm and induced with IPTG (0.4 mM) at an OD₆₀₀ of 0.5. The culture was incubated for 1 h and then potassium isobutyrate (pH 7.0) was added from a sterile concentrated stock solution of 500 mM to a final concentration of 5 mM. Samples were taken immediately after the addition of isobutyric acid and at intervals thereafter, and samples were analysed by reverse phase HPLC. The *E. coli* BL21 (DE3) pLysS pET20b(+) strain, which did not contain the 4HBT-ACX4-AcsA co-expression operon, was also cultured with isobutyric acid supplemented to the culture, as above, as a control. In the second control, *E. coli* BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-AcsA was cultured with no isobutyric acid, and, in the final control, the *E. coli* BL21 (DE3) pLysS pET20b(+) control strain was also cultured without isobutyric acid.

Methacrylic acid production in the *E. coli* BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-AcsA culture that was supplemented with isobutyric acid increased with time, but methacrylic acid production increased only marginally between the 8 h sample and the 19.5 h sample. A final concentration of ~0.25 mM was reached (Figure 52).

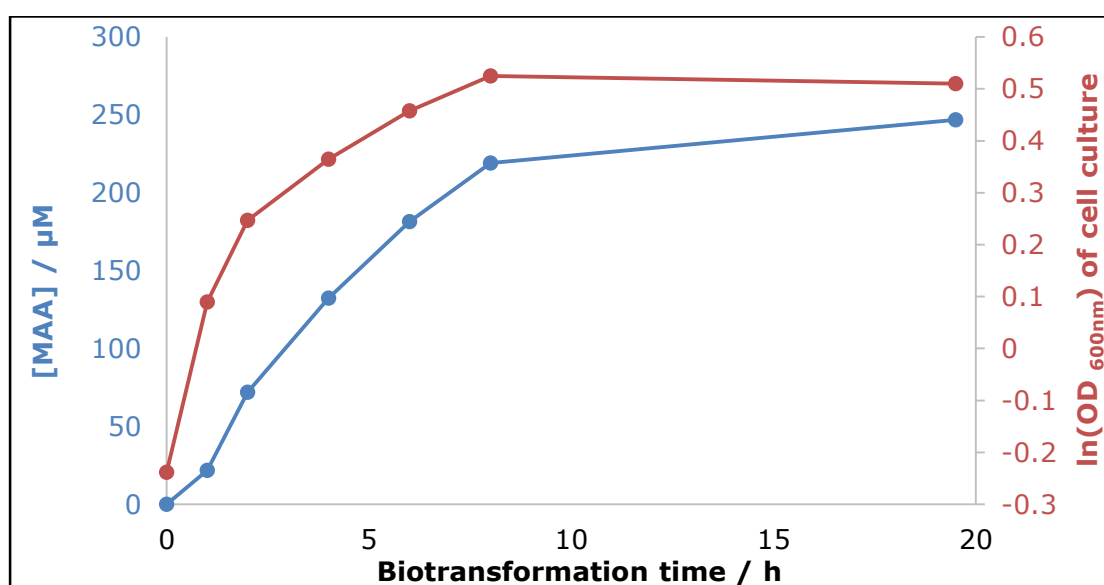


Figure 52: Production of methacrylic acid by *E. coli* BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-AcsA during biotransformation of isobutyric acid and growth of culture during the biotransformation expressed as natural logarithm of the culture's OD₆₀₀ reading.

By contrast, no methacrylic was produced in any of the control cultures at any of the timepoints. The HPLC traces for the samples taken at 8 hours from all cultures

and controls are summarised below (Figure 53). An important observation is that whilst only approximately 0.25 mM methacrylic acid was produced in the biotransformation, almost all of the isobutyric acid remained unconverted.

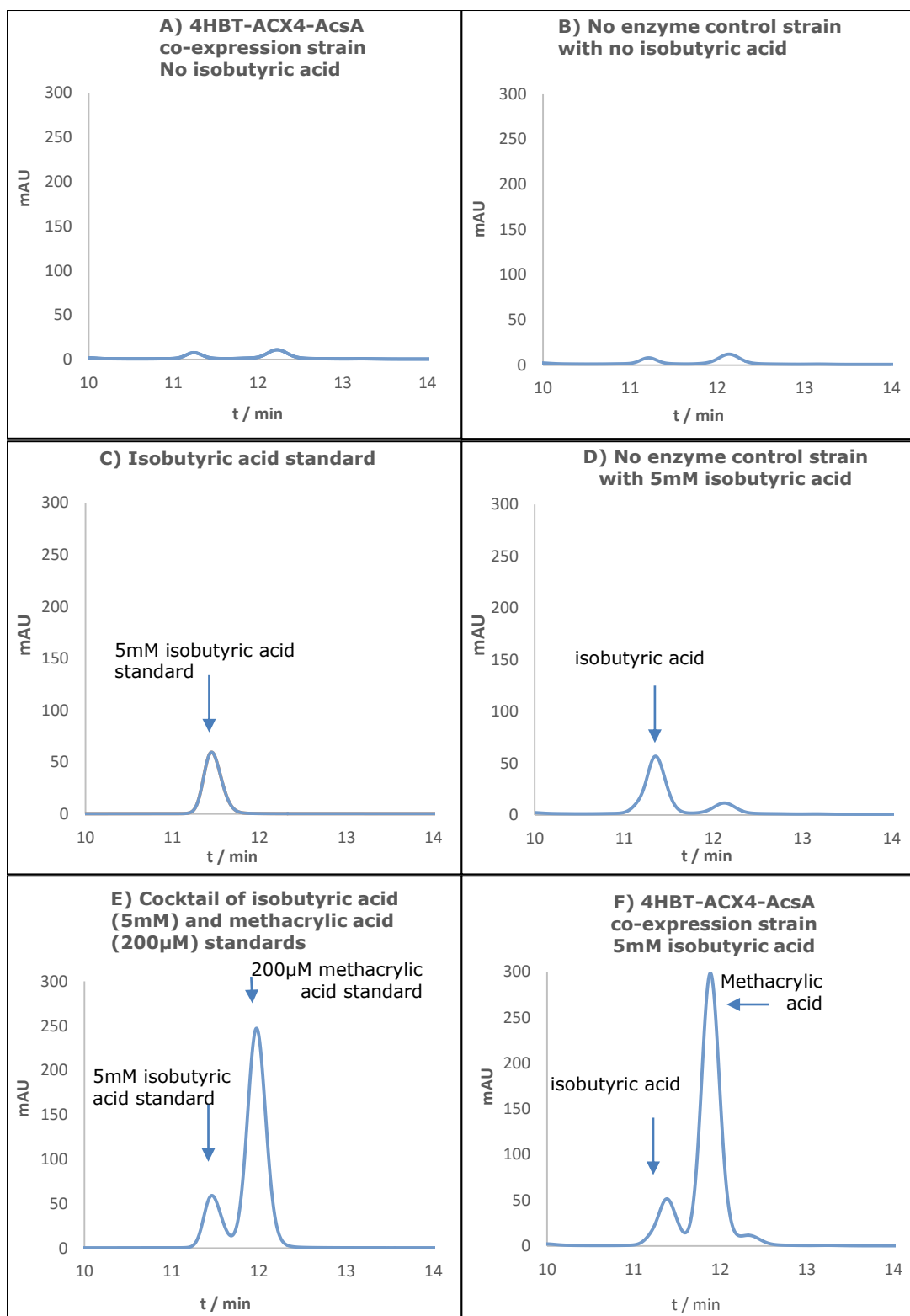


Figure 53: HPLC traces of whole cell live culture biotransformation supernatant (8h samples). HPLC conditions: Isocratic elution using 14% Acetonitrile in 50mM KH_2PO_4 acidified to pH 2.5 with HCl, $0.4 \text{ ml} \cdot \text{min}^{-1}$ on Agilent Zorbax Eclipse XDB C18 150 mm $\times$ 4.6 mm.

The biotransformation of isobutyric acid was repeated in triplicate, using cultures of *E. coli* BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-AcsA that were supplemented with isobutyric acid. Cultures were again incubated at 37 °C but with shaking at 250 rpm.

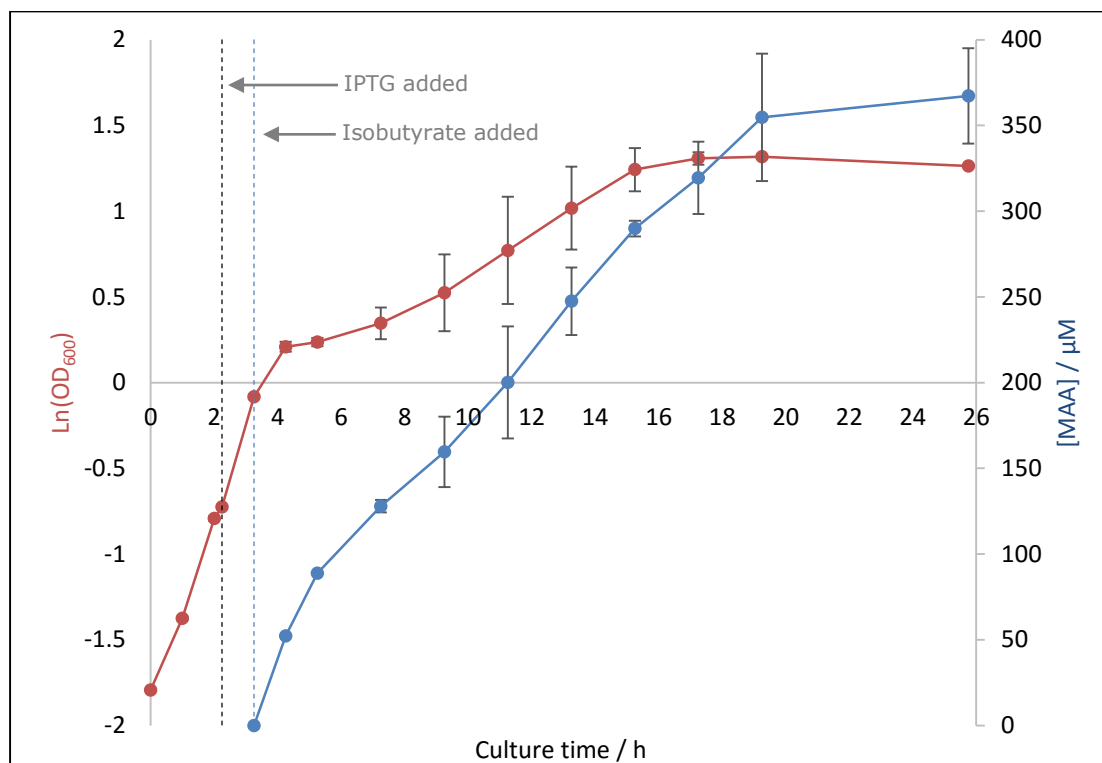


Figure 54: Growth of *E. coli* BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-AcsA on glucose and its conversion of isobutyric acid to methacrylic acid with time (triplicate). ● = Ln(OD₆₀₀), ● = [MAA]/μM. Three separate seed cultures were prepared and each used to inoculate biotransformation scale (100 ml) cultures to a target starting OD₆₀₀ of 0.16. HPLC conditions: Isocratic elution using 14% Acetonitrile in 25 mM KH₂PO₄ acidified to pH 2.5 with HCl, 1 mL.min⁻¹ on Agilent Zorbax eclipse XDB-C18, 250 mm × 4.6 mm.

In these cultures, methacrylic acid production began immediately after the addition of isobutyric acid, peaked at an average of 367 μM ± 28 μM, and began to plateau 2 h after the cell culture entered stationary phase (Figure 54). This biotransformation demonstrates a functional metabolic pathway to methacrylic acid. However, as in the preliminary experiment, the majority of the isobutyric acid remained unreacted.

The presence of unreacted isobutyric acid may be due to the low levels of soluble AcsA expressed, resulting in slow activation of isobutyric acid to isobutyryl-CoA. It may also, however, be due to the hydrolysis of isobutyryl-CoA that was observed previously by cell free extracts of the *E. coli* control strain, as well as indeed by Zhang *et al* [16], causing the isobutyryl-CoA to be hydrolysed back to isobutyric acid and coenzyme A. If it is due to the former, then optimisation of AcsA expression

needs to be performed. Indeed attempts were made to optimise the expression of AcsA on its own using the *E. coli* BL21 (DE3) pLysS pET20b(+):AcsA host, by replicating the conditions published by Hashimoto *et al.* [192], who had managed to express soluble AcsA in 2xYT medium at 28 °C. Expression at 18 °C in 2xYT medium was also attempted. However, neither attempt was successful at expressing soluble AcsA. The previous attempt to co-express 4HBT, ACX4 and AcsA in MSX medium at 28 °C had resulted in lower ACX4 expression without any effect on AcsA solubility. One difference between the AcsA gene used by Hashimoto *et al.* and the AcsA gene used in this experiment, however, is that the gene used in this experiment was codon optimised for maximal protein expression in *E. coli*. This could have been counterproductive, making AcsA express more rapidly, and leaving less time for the protein to fold correctly [199]. Thus, any further experimentation with this pathway may require using the wild-type AcsA gene. Furthermore, alternatives to AcsA for the activation of isobutyric acid to isobutyryl-CoA exist, which may be easier to express in *E. coli*. These alternatives may either be other AMP-forming acyl-CoA ligases, or the combination of an isobutyrate kinase and a phosphate butyryl transferase. Indeed, Cary *et al.* were able to enable growth of *E. coli* on butyrate by cloning a butyrate kinase and phosphate butyryl transferase from *C. acetobutylicum* [200], suggesting that the pathway, though for butyrate and not necessarily isobutyrate, is active enough to sustain growth of an organism, and so ought also be active enough to drive a biotransformation. However, if the main reason for unreacted isobutyrate is that isobutyryl-CoA is being hydrolysed back to isobutyric acid, then the enzymes that are responsible would need to be disrupted. Three broad specificity thioesterases that exist in *E. coli* are known to hydrolyse isobutyryl-CoA. These are *E. coli* YciA [121], tesA and tesB [16], and the effect of deleting these genes on the biotransformation productivity ought to be investigated. However, the priority of this project was to develop a route to methacrylic acid from glucose, and so reaching that goal quickly was prioritised over further optimisations to this biotransformation pathway.

5.4 The production of methacrylic acid from glucose

Having demonstrated the successful *in vivo* biotransformation of isobutyric acid to methacrylic acid, the next step was to produce methacrylic acid directly from glucose. Two routes were considered to achieve this. In the first route, the AcsA-ACX4-4HBT catalysed biotransformation pathway from isobutyric acid to methacrylic acid would be coupled to the metabolism of *E. coli* through Zhang *et al.*'s optimised route to isobutyric acid (titre: 90 g.L⁻¹) [16], which involved the decarboxylation of 2-ketoisovaleric acid to isobutyraldehyde and the oxidation of isobutyraldehyde to isobutyric acid (Figure 55A). In the second pathway, the ACX4- and 4HBT- catalysed conversion of isobutyryl-CoA to methacrylic acid and coenzyme A would be coupled to the metabolism of *E. coli* through the oxidative decarboxylation of 2-ketoisovaleric acid to isobutyryl-CoA (Figure 55B), as occurs in the valine degradation I pathway.

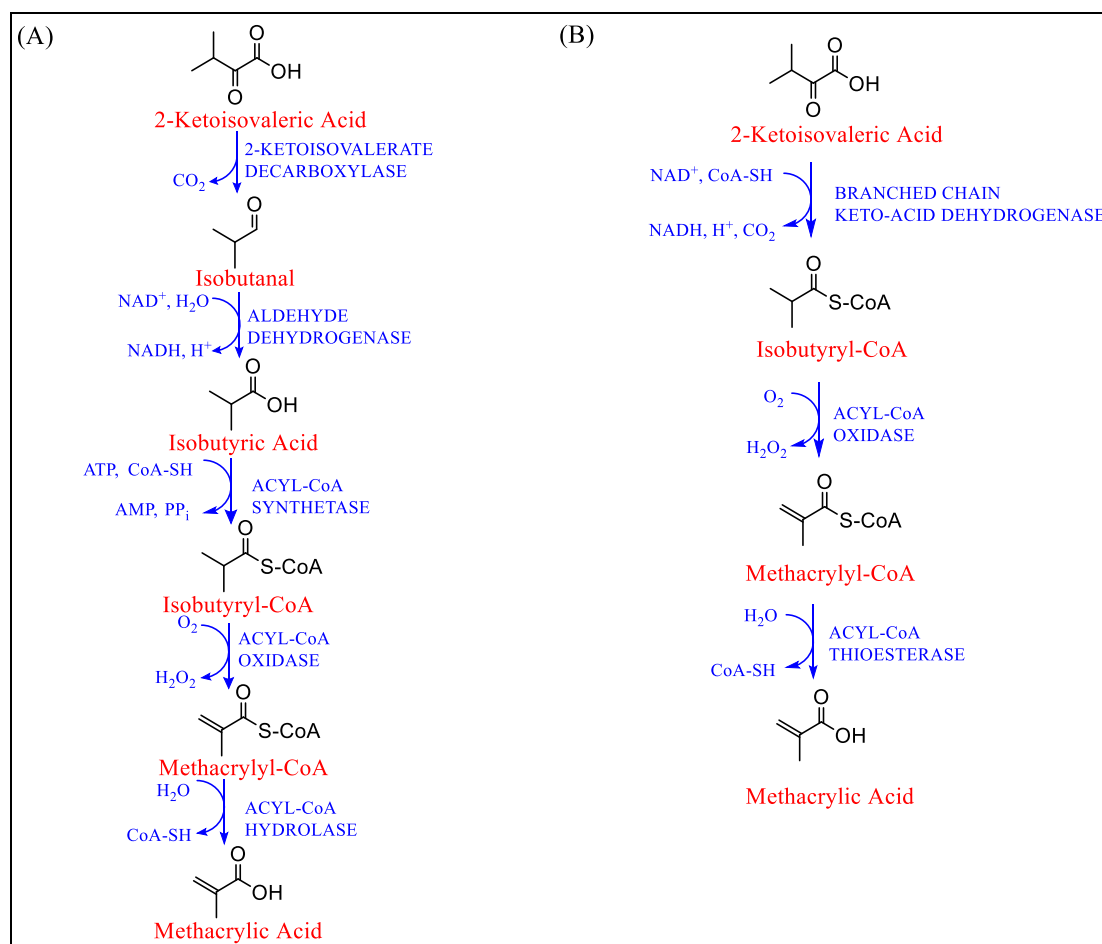


Figure 55: Two candidate routes to methacrylic acid from 2-ketoisovaleric acid, a widely available intermediate of *Escherichia coli*. Route A) 2-ketoisovalerate is first decarboxylated to isobutanal before this is oxidised to isobutyric acid. Isobutyric acid is then oxidised to methacrylic acid by the concerted action of an acyl-CoA synthetase, acyl-CoA oxidase and acyl-CoA thioesterase, as occurred in the isobutyrate to methacrylate biotransformation pathway. Route B) 2-ketoisovaleric acid is oxidatively decarboxylated to isobutyryl-CoA, which is then oxidised to methacrylyl-CoA by an acyl-CoA oxidase, and methacrylyl-CoA is in turn hydrolysed to methacrylic acid and coenzyme A by an acyl-CoA thioesterase.

In the first candidate route, three discrete enzymes catalyse the conversion of 2-ketoisovaleric acid to isobutyryl-CoA. In the second route, however, what is apparently a single enzymatic step is catalysed by a large heteromeric protein complex, with three enzymatic activities (Figure 56). This combined enzymatic activity is analogous to the oxidative decarboxylation of pyruvate by the pyruvate dehydrogenase complex (PDC), as well as the oxidative decarboxylation of 2-ketoglutarate by the 2-ketoglutarate dehydrogenase complex (KDC). The BCKDs, PDCs and KDCs are collectively termed the 2-ketoacid dehydrogenases [201].

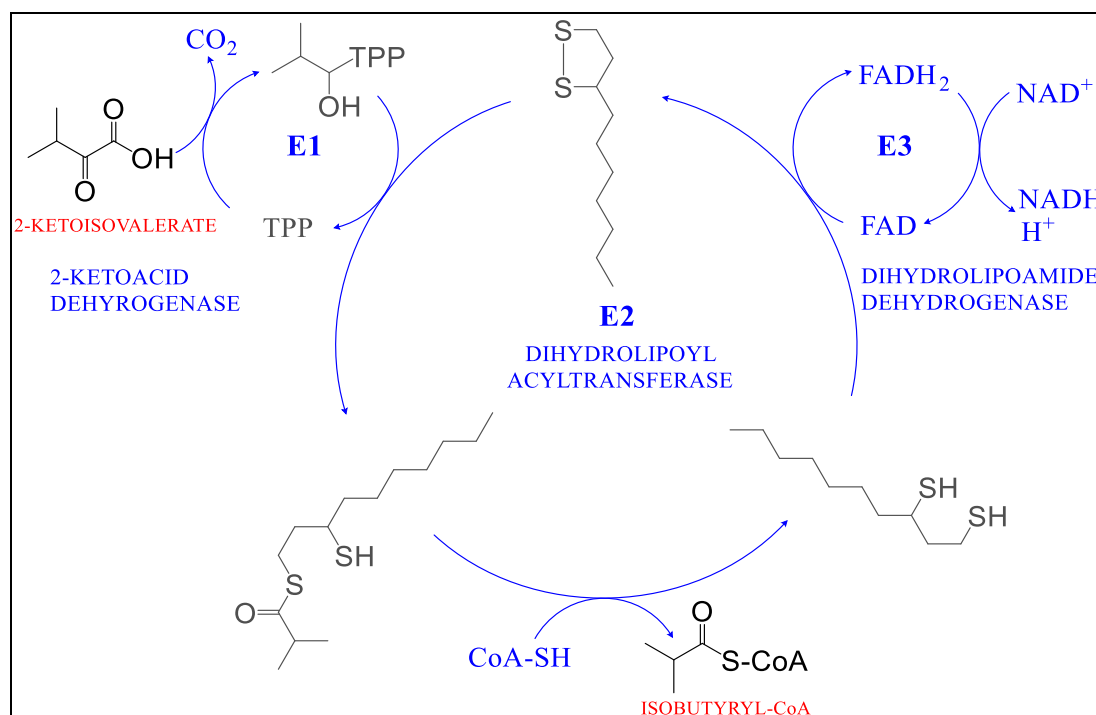


Figure 56: The oxidative decarboxylation of 2-ketoisovalerate by branched chain keto acid dehydrogenases through three concerted enzymatic activities (Adapted from [201]).

The three activities catalysed by the 2-keto acid dehydrogenase complexes are a 2-ketoacid dehydrogenase activity, catalysed by an E1 subunit, dihydrolipoyl acyltransferase activity (E2), and a dihydrolipoyl dehydrogenase activity (E3), with multiple genes expressing the component polypeptides of this heteromer. For the oxidative decarboxylation of 2-ketoisovalerate, subunit E1 catalyses the decarboxylation of 2-ketoisovalerate, using thiamine pyrophosphate (TPP) to form 2-methyl-hydroxypropionyl-TPP. The E2 subunit catalyses the transfer for the acyl group on the TPP conjugate to lipoic acid, which is anchored to the E2 component through an amide bond to a lysine residue, and then onto coenzyme A (CoA-SH), forming isobutyryl-CoA and leaving behind a dihydrolipoyl group. The E3 group reoxidises the dihydrolipoyl group to a lipoyl group, recruiting a flavin adenine dinucleotide (FAD) cofactor as the electron acceptor, which is itself reoxidised by

the mobile electron carrier, nicotinamide adenine dinucleotide (NAD^+). By forming a protein complex in which the product of one component is channeled to the next component as opposed to being released into the cytosol [201], the overall enzyme activity of the protein complex is believed to be greater than the equivalent reactions catalysed by distinct and separate enzymes, and minimises the risk of intermediates being consumed non-productively by other enzymes in the cell [202, 203]. Indeed, prior to Zhang *et al.*'s optimisation of their route to isobutyric acid which proceeded *via* the isobutanal intermediate, the route *via* isobutyryl-CoA, in which a branched chain keto acid dehydrogenase was expressed, produced greater titres of isobutyric acid [16]. It was only when various alcohol dehydrogenases were knocked out, and the aldehyde dehydrogenase further upregulated, that the route *via* isobutanal was more productive. Furthermore, though both pathways release NADH, which is coupled to the electron transport chain to yield 3 ATP molecules, the route *via* isobutyraldehyde and isobutyric acid requires ATP for the reactivation of isobutyric acid to isobutyryl-CoA, and thus conserves less energy for the host organism. Finally, during the biotransformation of isobutyric acid to methacrylic acid, considerable quantities of isobutyric acid remained unconsumed. The causes for this were perceived to be a combination of insufficient soluble acyl-CoA synthetase expression, and the presence of broad specificity thioesterases hydrolysing isobutyryl-CoA.

A pathway to methacrylic acid which utilises a BCKD ought to therefore produce more methacrylic acid without needing to perform genetic knockouts of alcohol dehydrogenases and optimising the expression of the aldehyde dehydrogenase and acyl-CoA synthetase, and in a more energetically favourable manner, thus reaching a proof of principle production of methacrylic acid from glucose more rapidly. Furthermore, any isobutyric acid that is produced from this pathway would be attributed to non-productive consumption of isobutyryl-CoA and not non-optimal acyl-CoA synthetase expression, and would be able to be addressed accordingly.

As such, the route to methacrylic acid *via* BCKD (Figure 55B) was pursued for the production of methacrylic acid from glucose. As Zhang *et al.* successfully cloned and expressed the genes encoding the components of BCKD from *Pseudomonas putida* KT2440, the enzyme from this source was an obvious choice for this experiment. It is useful, however, to review alternative sources of BCKD, such that any features that may make one BCKD more favourable over the other, can be compared.

5.4.1 Selecting a branched chain keto acid dehydrogenase

BCKDs have been reported in mammals [204-208], plants [83, 84, 171, 209] and bacteria [16, 86, 96, 127, 210-223], and in all of these sources, the E1 component is composed of two parts, E1 α and E1 β , whilst the E2 and E3 components are each composed of one part. Thus, four genes in total encode the component polypeptides of BCKDs. Mammalian BCKDs have been identified in rat kidney [204, 205], rabbit liver [206], bovine liver [207] and bovine kidney [208]. The BCKDs from all these sources have demonstrated *in vitro* activity for the oxidative decarboxylation of 2-ketoisovaleric acid to isobutyryl-CoA. The work performed on rat kidney, rabbit liver and bovine liver BCKDs showed that these enzymes are subject to inactivation by a protein kinase, and are also in some cases affected by product inhibition. Furthermore, all of the reported mammalian BCKDs appear to be of mitochondrial origin and any reports for the expression of mammalian BCKD in *E. coli* have focused on just the E1 component. Successful expression required fusing the proteins to maltose binding protein, as well as co-expression with the GroES/EL chaperonin [224, 225]. Due to this difficulty to express mammalian BCKDs in bacteria, the ability for the enzymes to be inactivated by phosphorylation, and the feed-back inhibition that some mammalian BCKDs are subjected to, they are not desirable candidates for a metabolic pathway to MAA, especially in a bacterial host.

In plants, branched chain amino acid catabolism has perhaps been best studied in *V. radiata* [83, 84] and *A. thaliana* [171, 209]. Gene sequences for the BCKD from *V. radiata* were not identified in the literature and database search for this project, whilst putative sequences for all components of *A. thaliana* BCKD genes have been identified [209, 226, 227]. The E2 component was successfully expressed in *E. coli* [228], but it is unclear whether the other components have been successfully expressed in *E. coli*. Due to this lack of sufficient evidence of recombinant BCKD expression from plant sources in *E. coli*, plant BCKDs are also not ideal.

In bacteria, BCKDs have been identified in *P. putida* [16, 86, 96, 127, 210-212], as well as in *Pseudomonas aeruginosa* [213], *Bacillus subtilis* [214-217], *Proteus vulgaris* [218], *Streptomyces avermitilis* [219-221] and *Enterococcus faecalis* [222, 223]. The BCKDs of *P. putida* and *P. aeruginosa* are components of the branched chain amino acid degradation pathways [85, 86, 96, 97, 229, 230]. In *B. subtilis*, however, the branched chain acyl-CoA products of the BCKD catalysed reaction are believed to prime branched chain fatty acid biosynthesis. In *S. avermitilis*, the BCKD provides the fatty acyl-CoA primers for both branched chain fatty acid (BCFA) biosynthesis, and in the case of isobutyryl-CoA and (S)-2-methyl-butyryl-CoA, for avermectin biosynthesis [231]. *S. avermitilis* also degrades the branched chain

amino acids in a similar manner to *P. putida* and *P. aeruginosa*, with its BCKD catalysing the first committed step. In both *B. subtilis* and *S. avermitilis*, the branched chain fatty acids increase membrane fluidity [232-236], and, whilst BCFA synthesis is essential in *B. subtilis*, production of unsaturated straight chain fatty acids compensates for such a loss in *S. avermitilis* [94]. In *E. faecalis*, the physiological role of BCKD is less clear. Unlike mammalian BCKDs, there is no evidence of enzyme inactivation by phosphorylation in any of the bacterial BCKDs.

There is evidence, though, to suggest that Pseudomonad BCKDs are activated by the branched chain amino acids [211, 213], with valine being the strongest effector [210]. The K_M values of the BCKDs from *P. putida* PpG2/ATCC23287 and *P. aeruginosa* were reported to be reduced by up to 10-fold by the addition of 5 mM valine in the assay mixture, with a reported 1.7-fold increase in V_{MAX} for the *P. putida* BCKD. However, both the *P. putida* and *P. aeruginosa* BCKDs have V_{MAX} values approaching $10 \mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ when fully activated, which is comparable to that of AcsA. As the V_{MAX} is only increased 1.7-fold by the branched chain amino acids, the V_{MAX} values of the Pseudomonad BCKDs for 2-ketoisovalerate is still quite high when not activated. In addition, the K_M values of *P. putida* BCKD and *P. aeruginosa* BCKD for 2-ketoisovalerate when not activated are 0.46 mM and 0.3 mM respectively, which are still reasonably low. Thus, whether or not there is a sufficient intracellular valine concentration in *E. coli* for maximum activation of the Pseudomonad BCKDs, they are still reasonably active. Interestingly though, it has been observed that valine can affect the constitutive expression of the BCKD from *P. putida* PpG2/ATCC23287 in *E. coli* [127], with an improved BCKD specific activity in cell free extracts from cultures grown in the presence of 25 mM valine, than from those grown without valine present in the medium.

In *S. avermitilis* ATCC 31272, two similar gene clusters encoding the $E1\alpha$, $E1\beta$ and $E2$ components of BCKD were identified, *bkdABC* [220] and *bkdFGH* [219]. Whilst both gene clusters appear to encode functional proteins, only the *bkdFGH* cluster is believed to function *in vivo*. While the $E1\alpha/\beta$ complex encoded by *bkdAB* was functionally expressed in *E. coli*, and gene sequences for the *bkdABC* genes from this organism are known, the gene encoding the lipoamide dehydrogenase from that complex is not known, and the gene sequences of the *bkdFGH* genes are not known for that organism. The complete genome sequence of a related organism, *S. avermitilis* MA-4680 (ATCC31267) is known (GenBank BA000030.3), and also has *bkdABC* and *bkdFGH* gene clusters. However, again, the gene for the appropriate lipoamide dehydrogenase is not known, and it seems that the *bkdFGH* genes have neither been experimentally confirmed in this organism, nor been

previously expressed in *E. coli*. In *E. faecalis* 10C1, an operon containing genes encoding E1 α , E1 β , E2 and E3 components along with phosphotransbutyrylase and butyrate kinase-like enzymes was identified, and the sequences of these genes are known [223]. However, as a category 2 pathogen on all major cell banks (ATCC, DSMZ, NCIB), handling this organism for amplification of the operon by PCR is difficult. Furthermore, there is no evidence of these genes having been functionally expressed in *E. coli* before, therefore having them custom synthesised would be risky.

In *B. subtilis*, a dedicated BCKD as well as a dual function pyruvate dehydrogenase-BCKD enzyme [237-239] contribute to the oxidative decarboxylation of branched chain keto acids. Some confusion as to the exact genes which encode the BCKD exist with *B. subtilis*, as strain 168s appears to contain the structural genes for the E1 and E2 components adjacent to each other [217], but no structural gene for the E3 component, whereas in strain 168, an operon exists which contains genes encoding the E1 α , E1 β , E2 and E3 components along with a postulated phosphotransbutyrylase and butyrate kinase type gene [216] arranged in a very similar manner to the BCKD of *E. faecalis* 10C1 [223]. In a recent paper, codon-optimised genes encoding BCKD proteins from an unspecified *B. subtilis* strain were expressed in *E. coli* for the synthesis of branched chain alkanes for biofuels synthesis [240]. The codon optimised gene sequences were not published, but the GenBank accession numbers for the amino acid sequences that were codon optimised, were. The Genbank sequences were NP_390285.1, NP_390284.1, NP_390283.1, and ZP_03600867.1, encoding the BCKD E1 α , E1 β and E2 components from *B. subtilis* 168, and the BCKD E3 component from *B. subtilis* JH642, respectively. However, when a BLAST search was performed on the amino acid sequence published under the Genbank ZP_03600867.1 entry, it was found to share 100% sequence identity with the E3 component of *B. subtilis* 168.

From the above information, it was decided that the genes encoding BCKD from both *P. putida* KT2440 (ppBCKD) and *B. subtilis* 168 (bsBCKD) would be cloned and expressed for the biosynthesis of methacrylic acid from glucose. The reasoning for this was that, whilst there was evidence for the BCKDs from the *Pseudomonas* organisms to require branched chain amino acids for maximal activation, there is no such evidence for the BCKDs from the other bacterial sources, and so the BCKDs from these organisms may be more active in *E. coli*. Cloning and expressing both ppBCKD and bsBCKD ought to maximise the chances for detectable methacrylic acid production, though the ppBCKD was considered a priority over that from

B. subtilis, as its function had been demonstrated in a pathway that yielded good titres of a product similar to methacrylic acid.

5.4.2 Identifying the nucleotide sequences for the BCKD operons in *P. putida* KT2440 and *B. subtilis* 168.

Although the BCKD from *P. putida* was previously cloned [16], the sequences of the cloned genes were not published. In order to ensure that the required restriction sites for the cloning of the BCKD genes into the pET20b(+):4HBT-ACX4 plasmid were not already in the sequence of the BCKD genes, the sequence encoding the BCKD genes was determined. To do this, the complete genome sequence for *P. putida* KT2440 (GenBank: AE015451.1) was first downloaded in FASTA format. New line characters were removed using the Microsoft Word replace tool, using “^p” in the ‘Replace’ field, and leaving empty the ‘with’ field. The sequence was then screened for nucleotide sequences matching the 3’ ends of the primers used by Zhang *et al.* [16], BKD-F-Zhang and BKD-R-Zhang, to amplify the operon out of the organism’s genome (Table 26).

Primer type	Primer sequence
BKD-F-Zhang	5’-TGCATCGAATTCAGGAGAAATTAAGTATGAACGAGTACGCCCCCTGCGTTTGC-3’
BKD-R-Zhang	5’-TGCATCAAGCTTTCAGATATGCAAGCGTGGCCAG-3’

Table 26: Forward and reverse primers used by Zhang *et al.* [16] for the amplification of a polynucleotide containing genes encoding the E1 α , E1 β , E2 and E3 components of branched chain keto acid dehydrogenase from *Pseudomonas putida* KT2440. BKD-F-Zhang is the forward primer, BKD-R-Zhang is the reverse primer. The forward primer contains an *Eco*RI site (dotted underline), an apparent ribosome binding site (bold underline) and a 5’-ATG-3’ start codon (underlined) starting 9 nucleotides after the apparent ribosome binding site. The reverse primer contains a *Hind*III restriction site (dotted underline) and the reverse complement of an apparent 5’-TGA-3’ stop codon.

One match was found when screening the complete genome sequence for the 3’ end of the BKD-F-Zhang primer, and another match was found when screening the sequence for the reverse complement of the 3’ end of the sequence to the apparent TGA stop codon in the BKD-R-Zhang primer. The sequence length between and including these two matches was 4,947 base pairs. This entire sequence between the apparent start codon at the beginning, and the apparent stop codon, was processed by the ExPASy Translate tool [241]. Two putative polypeptides were identified in frame 1 of the translated sequence, two putative polypeptides were identified in frame two of the sequence and one putative polypeptide was identified in frame three of the sequence. A BLAST search was performed on the putative polypeptides with the NCBI BLAST+ service from the European Bioinformatics Institute, using the blastp program and the default settings, and four genes were identified with 100% homology to genes in *P. putida* KT2440. These were *bkdA1*, encoding 2-oxoisovalerate dehydrogenase alpha subunit (E1 α); *bkdA2*, encoding

2-oxoisovalerate dehydrogenase beta subunit (E1 β); *bkdB*, encoding 2-oxoisovalerate dehydrogenase lipoamide acyltransferase component (E2) and *lpdV*, encoding 2-oxoisovalerate dehydrogenase dihydrolipoyl dehydrogenase (E3). The primers used by Zhang *et al.* [16] to amplify the *bkdA1*, *bkdA2*, *bkdB* and *lpdV* genes replaced the sequence immediately upstream of the *bkdA1* gene to introduce a new, synthetic RBS for the *bkdA1* gene, whilst the wild-type RBSs for the downstream genes would have been preserved during the amplification. The amplified polynucleotide sequence that Zhang *et al.* [16] would have obtained was determined (Polynucleotide sequence 8, appendices, Page VI).

Similarly, although the amino acid sequences of the E1 α , E1 β , E2 and E3 components of *B. subtilis* 168 were published [240], and the individual proteins could be codon optimised and chemically synthesised, as was performed by Howard *et al.* [240], it was considerably cheaper to amplify the genes out of the genome of *B. subtilis* 168. The collective length of the genes was approaching 5 kb, and they were clustered together within an operon, meaning that they could be amplified in a single PCR reaction without needing to subsequently concatenate the genes into a single polynucleotide. Therefore, to identify the sequence of DNA within which the *B. subtilis* 168 BCKD E1 α , E1 β , E2 and E3 components were encoded on, a BLAST search was performed using the published amino acid sequences, and the genes that were identified were *lpdV*, *bkdAA*, *bkdAB* and *bkdB* (Figure 57). The sequence of DNA encompassing those genes was then downloaded in FASTA format (Polynucleotide sequence 9, appendices, page VII).

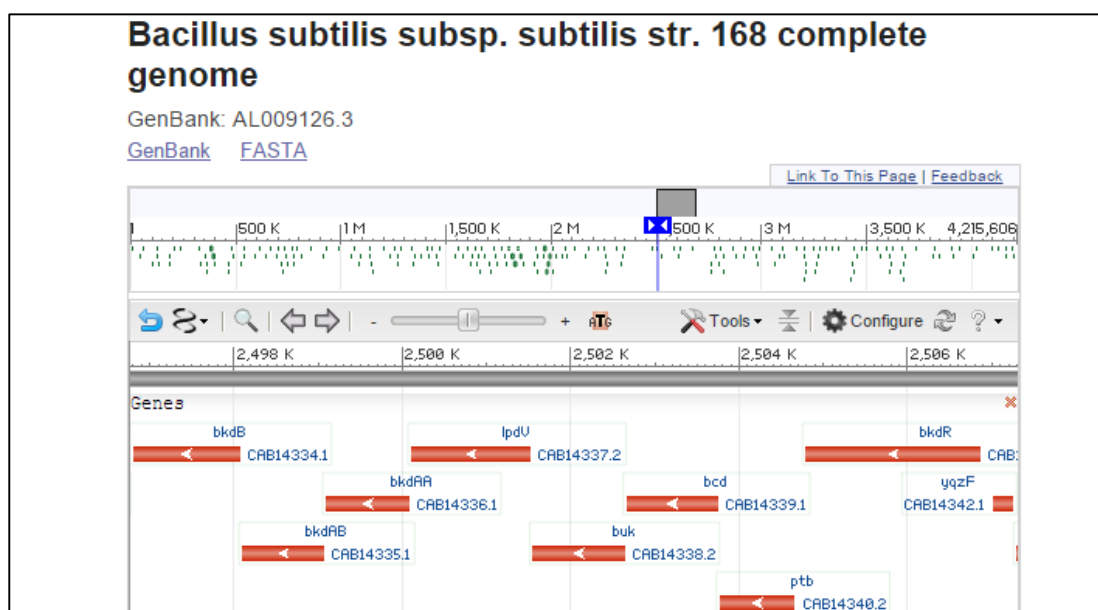


Figure 57: Genes encoded in the operon containing the E3 (*lpdV*), E1 α (*bkdAA*), E1 β (*bkdAB*) and E2 (*BkdB*) components of the BCKD from *B. subtilis* 168.

5.4.3 Construction of plasmids and strains for the co-expression of 4HBT, ACX4 and the BCKD genes from *P. putida* KT2440 and *B. subtilis* 168.

An attempt to amplify the ppBCKD and bsBCKD genes from the genomic DNA of *P. putida* KT2440 and *B. subtilis* 168 was performed using primers bsBCKD_F and bsBCKD_R for the genes from *B. subtilis* 168, and primers ppBCKD_F and ppBCKD_R for the genes from *P. putida* KT2440 (Table 2).

Primer bsBCKD_F (5'-GCTAGCTAGACAGACAGGAGTGAGTCACC-3') was designed to anneal just upstream of the *lpdV* gene from *B. subtilis* 168, and to include the wild-type RBS (double underline). The primer's 5' end contained an *NheI* restriction site (underline). The bsBCKD_R primer (5'-CTCGAGTTAGTAAACAGATGTCTTCTCGTCAATC-3') was designed to anneal at the 3' end of the *bkdB* gene (dotted underline) of *B. subtilis* 168, and its 5' end contained the *XhoI* restriction site (underline). The *NheI* and *XhoI* restriction sites were selected to facilitate cloning of the PCR product into the pET20b(+):4HBT-ACX4 vector, to yield pET20b(+):4HBT-ACX4-bsBCKD.

Primer ppBCKD_F (5'-TGCATCGCTAGCAGGAGAAATTAACATATGAACGAGTACGCCCCCTGCGTTGC-3') and primer ppBCKD_R (5'-TGCATCCTCGAGTCAGATATGCAAGGCGTGGCCCAG-3') were modelled after those used by Zhang *et al.* (Table 26), who used Phusion DNA polymerase to amplify the genes [16]. Primer ppBCKD_F was therefore intended to anneal to the 5' end of the *P. putida* KT2440 BCKD *bkdA1* gene (dotted underline), and replace the natural RBS and spacer sequence with the same as those used by Zhang *et al.*, but to include an *NheI* site (bold underline) instead of the original *EcoRI* site. Similarly, the ppBCKD_R primer was intended to anneal to the 3' end of the *P. putida* *lpdV* gene (dotted underline), and include a *XhoI* site (underlined) instead of the *HindIII* site included by Zhang *et al.* [16]. The *NheI* and *XhoI* restriction sites were selected to facilitate cloning of the PCR product into the pET20b(+):4HBT-ACX4 vector.

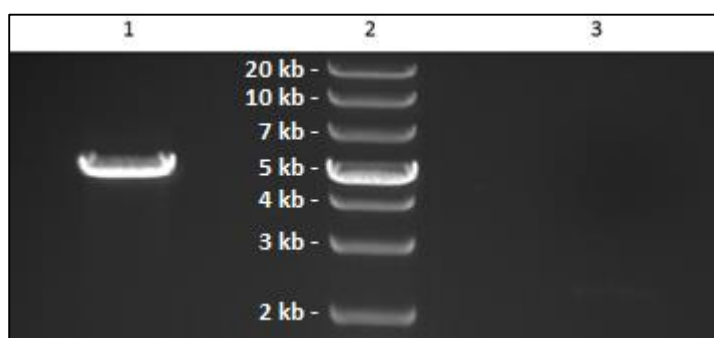


Figure 58: PCR amplification attempts for BCKD genes from *B. subtilis* 168 and *P. putida* KT2440 genomes. Lane 1: PCR product from the amplification of the BCKD genes out of *B. subtilis* 168 genomic DNA. Lane 2: GeneRuler 1kb plus DNA ladder. Lane 3: PCR product from the attempted amplification of the BCKD genes out *P. putida* KT2440 genomic DNA.

Whilst there was a successful amplification of the genes from *B. subtilis* 168 (Figure 58, lane 1), the PCR amplification of the genes from *P. putida* KT2440 was unsuccessful (Figure 58, lane 3), and attempts to modify the PCR conditions for the amplification of ppBCKD led to non-specific PCR products. Thus, the genes were instead custom synthesised by Biomatik Corporation as part of a single polynucleotide, (Polynucleotide sequence 10, appendices, page IX), to avoid further complications in amplifying the genes for this BCKD, which was considered a priority over that from *B. subtilis*. As the genes were custom synthesised, in addition to the *NheI* cloning site and ribosome binding site that was designed into the primers for the amplification of the genes, an *XbaI* cloning site, followed by a spacer sequence, was also included upstream of the *NheI* site. This was to enable cloning of the *P. putida* BCKD genes into the pET20b(+) vector, without the 4HBT and ACX4 genes, to yield pET20b(+):ppBCKD. The spacer sequence was included between the *XbaI* site and the *NheI* side to maintain the same number of base-pairs between the T7 promoter and the first RBS as exists in the pET20b(+) plasmid.

The PCR product containing the *B. subtilis* 168 BCKD genes (Polynucleotide sequence 11, appendices, page X) was then subcloned into the pET20b(+):4HBT-ACX4 plasmid between the *NheI* and *XhoI* restriction sites, to form pET20b(+):4HBT-ACX4-bsBCKD (Figure 59) (pHABS). The sequence of the polynucleotide containing the *bs_lpdV*, *bs_bkdAA*, *bs_bkdAB* and *bs_bkdB* genes was then verified by DNA sequencing directly from pHABS using primers *bsBCKD_S1* to *bsBCKD_S6*.

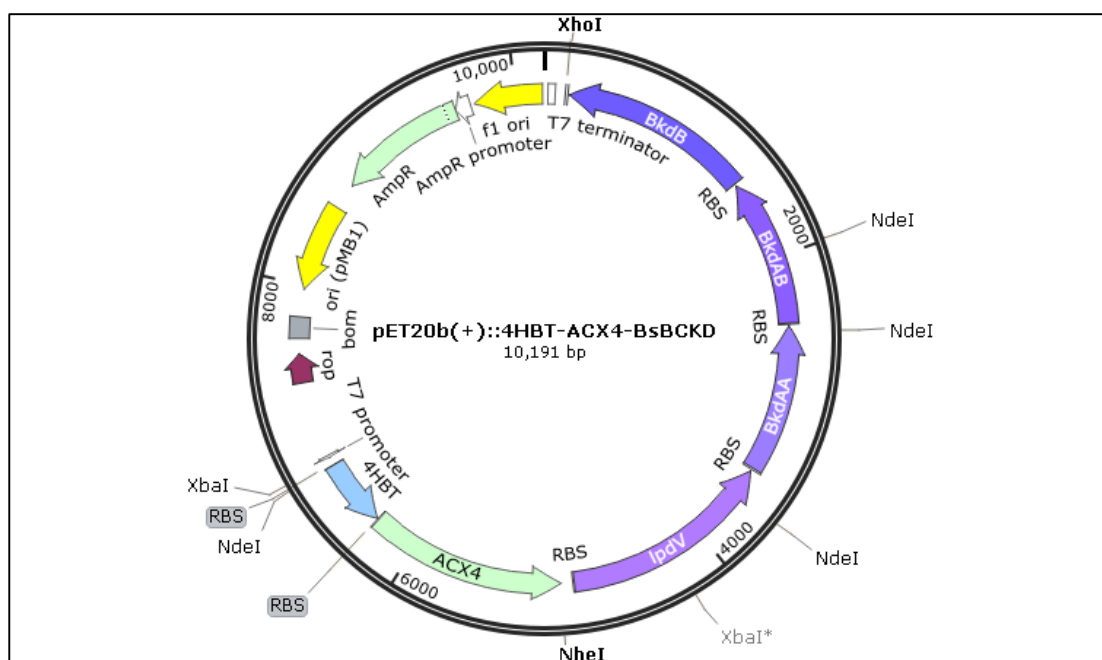


Figure 59: The pET20b(+):4HBT-ACX4-bsBCKD plasmid (pHABS), for coexpression of 4HBT, ACX4 and the genes encoding the E3, E1 α , E1 β and E2 components of *B. subtilis* 168 BCKD.

The single polynucleotide containing the genes of the *P. putida* KT2440 BCKD was delivered in the pBSK plasmid (pBSK::ppBCKD), and was sub-cloned into the pET20b(+):4HBT-ACX4 plasmid, in between the *Nhe*I and *Xho*I restriction sites, to form pET20b(+):4HBT-ACX4-ppBCKD (Figure 60) (pHAPP). Similarly, the polynucleotide was sub-cloned into the pET20b(+) plasmid between the *Xba*I and *Xho*I restriction sites, forming pET20b(+):ppBCKD (Figure 61).

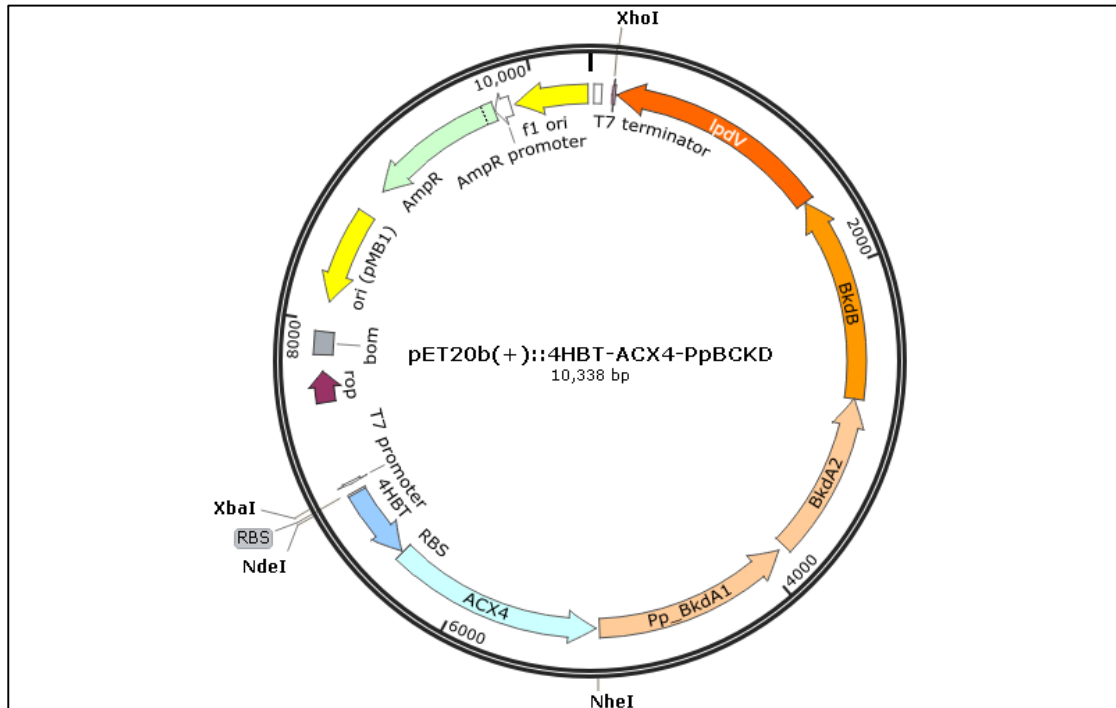


Figure 60: The pET20b(+):4HBT-ACX4-ppBCKD plasmid (pHAPP), for the coexpression of 4HBT, ACX4 and the the *P. putida* KT2440 BCKD E1 α , E1 β , E2 and E3 components.

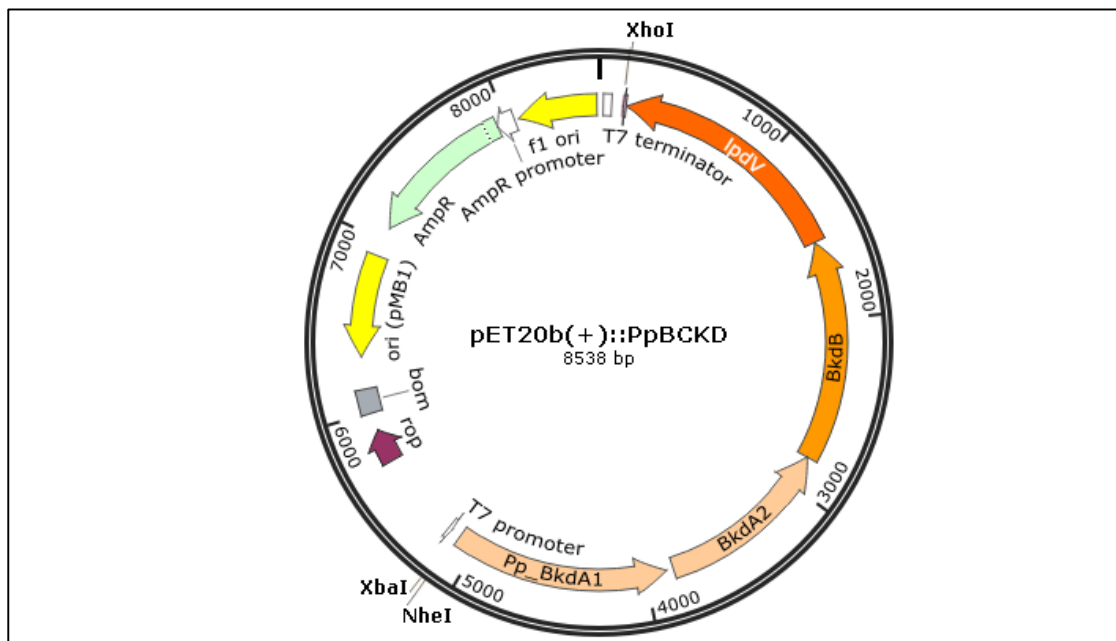


Figure 61: The pET20b(+):ppBCKD plasmid, for the coexpression of the genes encoding the E3, E1 α , E1 β and E2 components of *P. putida* KT2440 BCKD.

The pET20b(+):ppBCKD, pHAPP and pHABS plasmids were then transformed into *E. coli* BL21 (DE3) pLysS strains to produce *E. coli* BL21 (DE3) pLysS pET20b(+):ppBCKD, *E. coli* BL21 (DE3) pLysS pHAPP and *E. coli* BL21 (DE3) pLysS pHABS. In order to confirm the presence of the pET20b(+):ppBCKD, pHAPP and pHABS plasmids in the above strains, cryostocks of each strain were first plated on LB agar plates with the appropriate supplements (Table 6) and plasmid mini-preps were prepared for each strain from single colonies. Due to the presence of the pLysS plasmid in these extracts, restriction mapping of the plasmid extracts was not considered a viable method to confirm the presence of the pET20b(+):ppBCKD, pHAPP and pHABS plasmids in these extracts. Primers pET20b_Phu_F and pET20b_Phu_R were therefore designed to bind to the pET20b(+) vector backbone, to amplify whichever insert was cloned between the *Xba*I and *Xho*I cloning sites, using the Phusion HotStart II DNA polymerase. However, only the PCR reaction using the plasmid extract from *E. coli* BL21 (DE3) pLysS pHABS resulted in the successful amplification. No PCR product was observed when using the plasmid extracts from *E. coli* BL21 (DE3) pLysS pET20b(+):ppBCKD or from *E. coli* BL21 (DE3) pLysS pHAPP as the plasmid source. As the pET20b_Phu_F and pET20b_Phu_R primers were designed to bind to the pET20b(+) vector skeleton, and not to the insert, and the fact that the PCR regime was successful with the plasmid extract from *E. coli* BL21 (DE3) pLysS pHABS, the primer design and the PCR regime were not the issue. The ppBCKD genes appeared to be difficult to amplify. This was of course also the case when an unsuccessful attempt was made to amplify the ppBCKD genes out of the genomic DNA of *P. putida* KT2440 using primers ppBCKD_F and ppBCKD_R. To explore this, the pET20b(+):ppBCKD plasmid that was prepared from *E. coli* JM107 pET20b(+):ppBCKD during the sub-cloning of the chemically synthesised genes, was used as the DNA source in a separate PCR reaction. Again, no successful PCR product was achieved. An alternative to amplifying the entire insert was thus found, and instead, primers pET20b_Phu_F and OE_ACX4_R were used in PCR reactions using plasmid DNA extract from *E. coli* BL21 (DE3) pLysS pHAPP. Though this bypassed the ppBCKD genes, it was treated as sufficient confirmation that the pHAPP plasmid was successfully transformed into *E. coli* BL21 (DE3) pLysS pHAPP.

5.4.3.1 A functional pathway for the production of methacrylic acid from 2-ketoisovaleric acid and from glucose.

In order to determine whether methacrylic acid could be produced using either *E. coli* BL21 (DE3) pLysS pHAPP or *E. coli* BL21 (DE3) pHABS, whether in cultures where the only carbon source was glucose, or in cultures that were supplemented with 2-ketoisovaleric acid, an initial screen was performed. A negative control using the empty vector *E. coli* BL21 (DE3) pLysS pET20b(+) control strain, from which no methacrylic acid production was expected, was included in this screen. In this screen, overnight seed cultures of these strains were prepared in duplicate, such that one repeat could be used to inoculate a culture supplemented with both glucose and 2-ketoisovaleric acid, and the other with glucose only. However, it was found that only the seed cultures of *E. coli* BL21 (DE3) pLysS pHAPP and *E. coli* BL21 (DE3) pLysS pET20b(+) grew overnight in minimal media, with no apparent growth for *E. coli* BL21 (DE3) pLysS pHABS. Further incubation resulted in growth of the latter strain after over 48 h. Experimentation with different volumes of seed culture demonstrated that an MSX seed culture of 4ml was the optimal volume in a 50 ml falcon tube for overnight growth of a colony of *E. coli* BL21 (DE3) pLysS pHABS. Thus, whilst the seed cultures from *E. coli* BL21 (DE3) pLysS pHAPP and *E. coli* BL21 (DE3) pLysS pET20b(+) were used to inoculate 100ml scale cultures in MSX media supplemented with carbenicillin, chloramphenicol and riboflavin, intermediate MSX cultures (20 ml) of *E. coli* BL21 (DE3) pLysS pHABS were prepared in a 100 ml shake flask for each repeat of the duplicated pre-cultures, and were incubated at 37 °C with shaking at 250 rpm until the cultures reached stationary phase. These were then used to inoculate expression-scale cultures of MSX media (100 ml) that were also supplemented with carbenicillin, chloramphenicol and riboflavin.

The 100 ml scale cultures were incubated at 37 °C with shaking at 250 rpm and expression was induced at OD readings between 0.5 and 0.7 by the addition of IPTG (0.4 mM). After 3 h, one of the two cultures for each expression strain was supplemented with 2-ketoisovalerate (5 mM), from a sterile stock (500 mM, pH 7.0). Cultures were incubated for a total of 17 h post induction, and samples from each culture were then analysed by HPLC (Table 27).

Samples had also been taken for analysis by SDS-PAGE from the cultures of *E. coli* BL21 (DE3) pLysS pHAPP in MSX medium that was supplemented with riboflavin, carbenicillin and chloramphenicol, along with samples from a culture *E. coli* BL21 (DE3) pLysS pET20b(+):ppBCKD, expressing just the ppBCKD genes, under identical conditions (Figure 62).

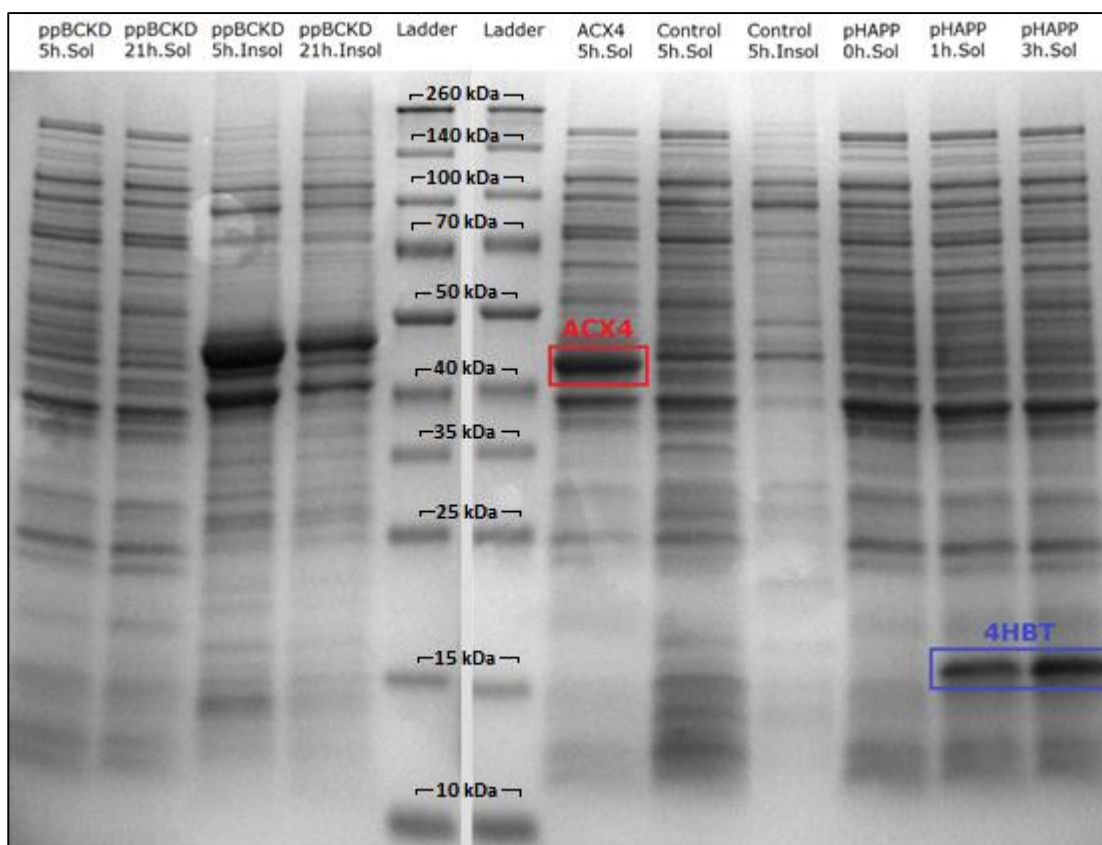


Figure 62: SDS-PAGE analysis of *E. coli* BL21 (DE3) pLysS pHAPP (right gel), with samples taken from cultures of *E. coli* BL21 (DE3) pLysS pET20b(+):ppBCKD as control for ppBCKD genes and from *E. coli* BL21 (DE3) pLysS pET20b(+):ACX4 as control for ACX4 gene expression only.

The SDS-PAGE analysis showed that 4HBT expressed well in the *E. coli* BL21 (DE3) pLysS pHAPP culture, but expression of ACX4 and the ppBCKD genes in the same culture could not be reliably detected. Soluble ppBCKD gene expression could not even be reliably detected in the culture of *E. coli* BL21 (DE3) pLysS pET20b(+):ppBCKD, with a lot of protein detected in the insoluble fractions.

Nevertheless, in the HPLC analysis for methacrylic acid production for both the *E. coli* BL21 (DE3) pLysS pHAPP and the *E. coli* BL21 (DE3) pLysS pHABS cultures showed that methacrylic acid was produced directly from glucose to final concentrations of 110 μ M MAA and 58 μ M MAA respectively, without the need to supplement the cultures with 2-ketoisovalerate. However, when the cultures were supplemented with 2-ketoisovalerate, those concentrations were boosted to 240 μ M and 86 μ M for the strains expressing the ppBCKD genes and the bsBCKD genes, respectively. This is only a small increase, considering that in both cases, all of the 2-ketoisovalerate (5 mM) was consumed. Very small peaks for isobutyric acid seemed to appear for the cultures that were supplemented with 2-ketoisovaleric acid, with peak areas representing isobutyric acid concentrations

in the order of 1 mM. However, as isobutyric acid has very low ultraviolet (UV) absorbance, the peaks were not large enough to accurately measure. Curiously, in the culture of the *E. coli* BL21 (DE3) pLysS, only 3 mM 2-ketoisovaleric acid remained in the culture at the end. The 2 mM 2 ketoisovalerate that was consumed in the control culture can be explained by the production of valine, leucine or pantothenate, from it, as 2-ketoisovalerate is a precursor to these [242]. However, considering that 3 mM 2-ketoisovalerate remained in the control culture that was supplemented with 2-ketoisovaleric acid, and that very little methacrylic acid and isobutyric acid, if at all, was produced in the cultures that were supplemented with 2-ketoisovalerate, it is strange that no 2-ketoisovalerate was detected in either of the *E. coli* BL21 (DE3) pLysS pHAPP or the *E. coli* BL21 (DE3) pLysS pHABS cultures. Perhaps methacrylyl-CoA was hydrated to 3-hydroxyisobutyric acid by an enoyl-CoA hydratase and consumed non-productively. Further repeats are required to confirm this observation, and further work is required to investigate it.

Culture of <i>E. coli</i> BL21 (DE3) pLysS			Final concentrations	
Plasmid	Induced at OD ₆₀₀ of:	Supplemented with [2-KIV] / mM	2 ketoisovaleric acid / mM	Methacrylic acid / μ M
pHAPP	0.61	5	0	240
pHAPP	0.61	0	0	110
pHABS	0.5	5	0	86
pHABS	0.51	0	0	58
pET20b(+) (control)	0.67	5	2.9	0
pET20b(+) (control)	0.66	0	0	0

Table 27: Summary of results for methacrylic acid production in cultures of *E. coli* BL21 (DE3) pLysS pHAPP and *E. coli* BL21 (DE3) pLysS pHABS, when cultured both with and without 2-ketoisovalerate supplementation.

Finally, in order to confirm that the HPLC peaks for methacrylic acid from the cultures of *E. coli* BL21 (DE3) pLysS pHAPP cultures, both with and without 2-ketoisovalerate supplementation (Figure 51A&B), were genuine, the samples were spiked with an additional 240 μ M of authentic methacrylic acid. The spiked samples were re-analysed by HPLC to ensure that the peaks did co-elute, and they did (Figure 51C&D).

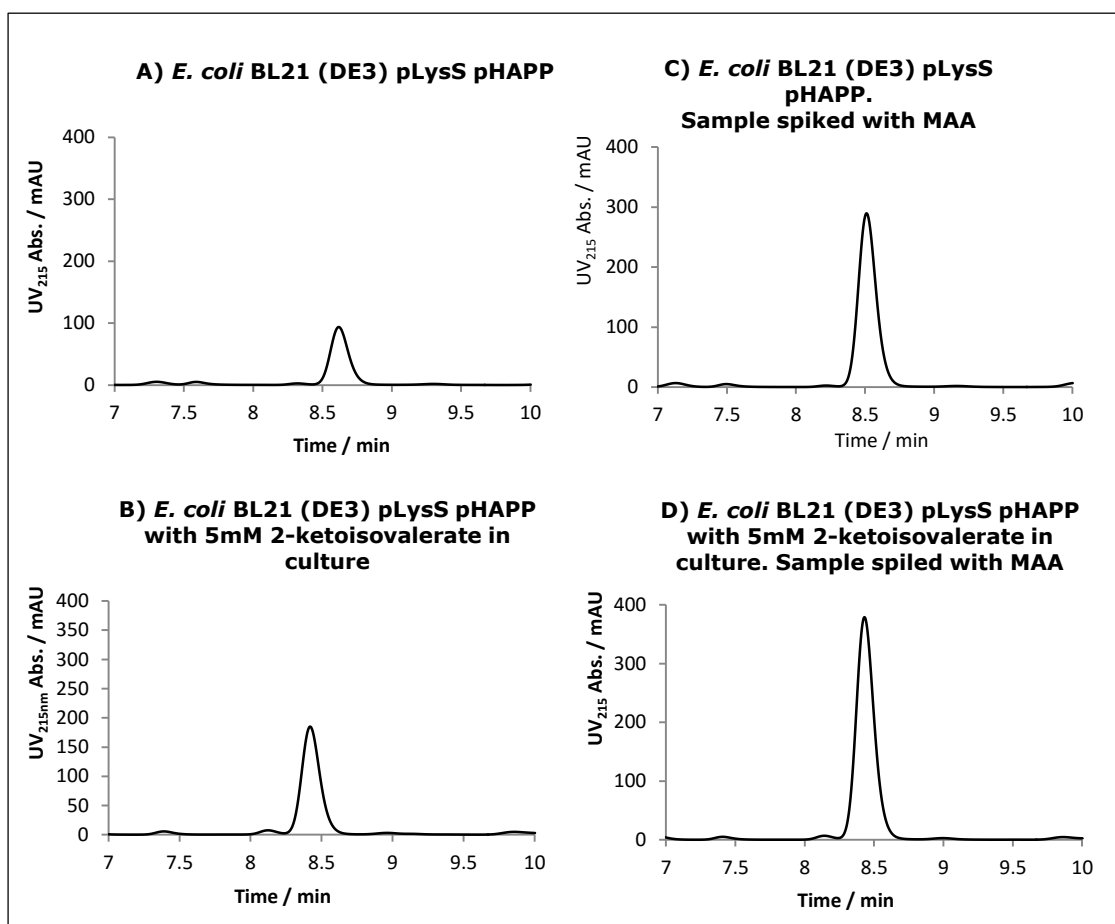


Figure 63: Selection of HPLC traces of samples taken from cultures during experiment rapid screening for production of MAA from glucose by *E. coli* BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-ppBCKD and *E. coli* BL21 (DE3) pLysS pET20b(+) control strain.

This preliminary screen has therefore shown that both *E. coli* BL21 (DE3) pLysS pHAPP and *E. coli* BL21 (DE3) pLysS pHABS were able to produce methacrylic acid directly from glucose, and that in both cases, supplementing cultures with 2-ketoisovaleric acid increased methacrylic acid production, though most of the additional 2-ketoisovalerate was consumed non-productively.

Ultimately, the aim of this series of experiments was to identify a host strain for the production of methacrylic acid directly from glucose, such that suitable optimisations such as genetic knock-outs could be made and their effects investigated. Thus, maximising the methacrylic acid production in the otherwise unmodified *E. coli* BL21 (DE3) chassis and identifying any side-products was deemed important for this objective. An investigation was therefore next made into the effect of upregulating 2-ketoisovaleric acid production by the host strain itself, and to determine whether the methacrylic acid production matched or exceeded that by the strains used in this experiment, when they were supplemented with 2-ketoisovaleric acid during growth.

5.4.4 The production of methacrylic acid from glucose with an upregulated 2-ketoisovalerate biosynthesis pathway.

In, *E. coli*, 2-ketoisovalerate is a natural intermediate of the valine biosynthesis pathway, whereby three enzyme-catalysed reactions yield 2-ketoisovalerate from pyruvate (Figure 64) [90]. First, acetohydroxyacid synthase (AHAS) catalyses the TPP-dependent decarboxylation of pyruvate and transfers the activated aldehyde to another pyruvate to form 2-acetolactate. Acetohydroxyacid isomeroreductase (AHAIR) catalyses the reduction of the 2-acetolactate β -keto group to a hydroxyl group and the carbon skeleton rearrangement of the α -methyl group to the beta carbon, to form 2,3-dihydroxy-3-methylbutanoate. Finally, dihydroxyacid dehydratase (DHAD) catalyses the dehydration of 2,3-dihydroxy-3-methylbutanoic acid to 2-ketoisovaleric acid [90].

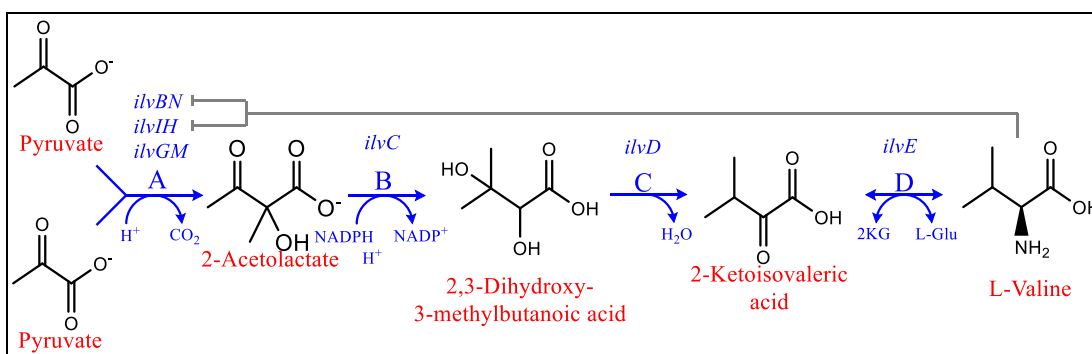


Figure 64: Valine biosynthesis from pyruvate in *Escherichia coli*. A) Acetohydroxyacid Synthase (AHAS). B) Acetohydroxy acid Isomeroreductase (AHAIR), C) Dihydroxyacid dehydratase. D) Branched chain amino acid transaminase. Three isozymes exist for AHAS. These are encoded for by *ilvBN*, *ilvIH*, *ilvGM*. The AHAS enzymes encoded for by *ilvBN* and *ilvIH* are both subject to feedback inhibition by L-valine. *ilvGM* is not subject to feedback inhibition by L-valine, but is inactive in *E. coli* K12 strains due to a frameshift mutation.

In *E. coli*, these enzymes also catalyse the formation of (S)-3-methyl-2-oxopentanoate of the isoleucine biosynthesis pathway [90]. The starting intermediates for the pathway to isoleucine are pyruvate and 2-ketobutyrate, which are converted to (S)-2-aceto-2-hydroxybutanoate as opposed to 2-acetolactate by AHAS [243]. While there is only one isozyme in *E. coli* for each of AHAIR and DHAD, encoded by *ilvC* and *ilvD*, respectively, three isozymes of AHAS exist in *E. coli*, AHASI, AHASII and AHASIII. These are encoded by *ilvBN*, *ilvGM* and *ilvIH*, respectively [90].

All three isozymes are composed of two large subunits and two small subunits. AHASI and AHASIII are subject to feedback inhibition by valine, whilst AHASII is not [90]. However, AHASII does have a frameshift mutation in *E. coli* K12, rendering it inactive [244] in this strain. The three isozymes also have different preferences for the formation of either 2-acetolactate acid or (S)-2-aceto-2-

hydroxy-butanoic acid, towards valine or isoleucine biosynthesis, respectively, and also have different preferences for acetolactate formation or (S)-2-aceto-2-hydroxybutanoate formation. A specificity ratio, 'R' (Equation 4) has been described [243] to quantify this preference for either 2-acetolactic or (S)-2-aceto-2-hydroxybutanoic acid formation, whereby the larger its value, the greater the enzyme's preference for (S)-2-aceto-2-hydroxybutanoate formation. *E. coli* AHAS I has a specificity ratio of 2, whilst AHAS III has a specificity ratio of 40 [245].

$$R = \frac{V_{AHB}/V_{AL}}{[\alpha - \text{ketobutyrate}]/[\text{pyruvate}]}$$

Equation 4: Specificity ratio equation for acetohydroxy acid synthase enzymes. V_{AHB} = rate of (S)-2-aceto-2-hydroxybutanoate production. V_{AL} = rate of 2-acetolactate production.

The production of 2-ketoisovalerate by *E. coli* has previously been upregulated for the biosynthesis of valine [246, 247], isobutanol [61, 248, 249] and isobutyric acid [16]. It has also been performed in *Corynebacterium glutamicum* for the biosynthesis of valine, isobutanol, isobutyric acid and also 2-ketoisovalerate in its own right [60, 250-255]. For the biosynthesis of isobutanol in *E. coli*, Atsumi *et al.* upregulated *E. coli* AHAS III, *E. coli* AHAS I and *E. coli* DHAD for the upregulation of 2-ketoisovalerate, as well as a keto-acid decarboxylase and an alcohol dehydrogenase for the conversion of 2-ketoisovalerate to isobutanol [61]. Cultures produced 23 mM isobutanol, which was approximately five-fold higher than when the *ilvIHCD* genes were not upregulated. Chromosomal deletions of genes encoding enzymes which could compete for isobutanol precursors, mainly pyruvate, boosted production to 30 mM. However, a further 1.7-fold improvement in final isobutanol concentration was observed when the *ilvIH* genes, encoding the large (*ilvI*) and small (*ilvH*) subunits of *E. coli* AHAS were replaced with a single gene, *alsS*, encoding degradative acetolactate synthase (ALS) from *B. subtilis* 168, first characterized by Holtzclaw and Chapman [256].

Degradative ALSs were first shown to be involved in acetoin formation, and catalyse the formation of 2-acetolactate, which is subsequently decarboxylated to acetoin [257]. The degradative ALSs have distinct roles from the biosynthetic AHAS enzymes [258], and whilst the degradative ALS from *B. subtilis* has a relatively high K_M for pyruvate of 13 mM, its specificity ratio was found to be 0.06 [245], demonstrating high preference for 2-acetolactate formation and not (S)-2-aceto-2-hydroxy-butanoate. Furthermore, degradative ALS from *Bacillus subtilis* is not subject to feedback inhibition by either isoleucine, leucine or valine.

This degradative ALS was also used by Zhang *et al.* to upregulate 2-ketoisovalerate production for the biosynthesis of isobutyric acid [16], during which titres of 90 g.L⁻¹ were achieved in a bioreactor. There is therefore good precedent to use degradative ALS from *B. subtilis* over any of the *E. coli* AHAS enzymes for the upregulation of 2-ketoisovalerate. However, while Atsumi *et al.* also upregulated both AHAIR (*ilvC*) and DHAD (*ilvD*) from *E. coli* to improve 2-ketoisovalerate production [61], Zhang *et al.* only upregulated DHAD and not AHAIR, as *ilvC* is purportedly upregulated under high acetolactate concentrations [16].

In this project, the degradative ALS from *B. subtilis* 168 was upregulated along with both AHAIR and DHAD from *E. coli*, in order to ensure that the pathway to ketoisovalerate was maximally upregulated. As *E. coli* BL21 (DE3) pLysS pHAPP and *E. coli* BL21 (DE3) pLysS pHABS both produced methacrylic acid from glucose via 2-ketoisovalerate, it was deemed undesirable to further modify the pHAPP and pHABS plasmids. Therefore, *alsS*, *ilvC* and *ilvD* were expressed from a separate plasmid. The pSTV28 plasmid has previously been used in the laboratory, and as it contains a chloramphenicol resistance gene and a p15a origin of replication, which is in a different incompatibility group to the pBR322 origin [198], it is compatible with the pET20b(+) plasmid which contains an ampicillin resistance gene and the pBR322 origin. This plasmid was therefore chosen for the coexpression of *alsS*, *ilvC* and *ilvD*.

5.4.4.1 Design and construction of the 2-ketoisovaleric acid upregulation plasmid pKIV

As the wild-type sequences for *E. coli ilvD*, *E. coli ilvC* and *B. subtilis alsS* were known to overexpress well in *E. coli*, the genes were not codon-optimised, and wild-type sequences were taken from the annotated complete genome sequences of *E. coli* MG1655 (genbank accession number: NC_000913.3) and *B. subtilis* 168 (genbank accession number: AL009126.3). Thus, to design the overexpression operon for 2-ketoisovalerate upregulation, the wild type sequences for *ilvD* (Polynucleotide sequence 12), *alsS* (Polynucleotide sequence 13) and *ilvC* (Polynucleotide sequence 14) were first arranged in that order (Figure 65). Prior to each gene, short sequences were inserted containing ribosome binding sites and spacer sequences modelled after those used by Atsumi *et al.* and Zhang *et al.* for the overexpression of the *ilvD*, *alsS* and *ilvC* genes. The sequence upstream of *ilvD* contained an *EcoRI* site, and a *HindIII* site was appended to the 3' end of the designed polynucleotide, to enable insertion into the pSTV28 vector. However, the genes within the polynucleotide had multiple *HindIII* and *EcoRI* sites which had to be disrupted to enable the insertion into the pSTV28 plasmid.

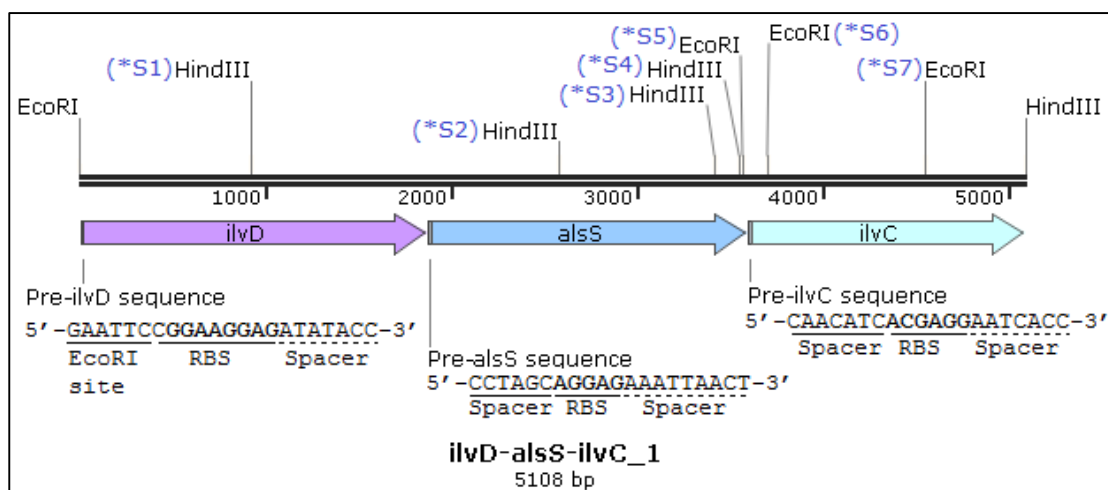


Figure 65: Polynucleotide containing *ilvD*, *alsS* and *ilvC* genes with ribosome binding sites preceeding each gene. The polynucleotide contains an *EcoRI* site at the start, and a *HindIII* site at the end. *HindIII* and *EcoRI* sites within the polynucleotide, sites S1 to S7, were disrupted with samesense mutations.

To disrupt restriction sites S1 to S7 (Figure 65), samesense mutations were made in the design of the polynucleotide (Table 28), using a codon frequency table for *E. coli* [259] to ensure that rare codons were not used.

Original sequence	RE site	Modified sequence	Amino acid codon frequencies		
			AA	Codon	Codon Frequency
GATAAGCTTCC AspLysLeuSer	<i>HindIII</i> (*S1)	GATAAGCTCTCC AspLysLeuSer	Leu	CTT CTC	0.12 0.10
GCGGAAAGCTTCGGA AlaGluSerPheGly	<i>HindIII</i> (*S2)	GCGGAAAGTTTCGGA AlaGluSerPheGly	Ser	AGC AGT	0.25 0.16
GACAAGCTTCCG AspLysLeuPro	<i>HindIII</i> (*S3)	GACAAGCTCCCG AspLysLeuPro	Leu	CTT CTC	0.12 0.10
AAAGAATTTCGGG LysGluPheGly	<i>EcoRI</i> (*S4)	AAAGAATTTCGGG LysGluPheGly	Phe	TTC TTT	0.42 0.58
GATGAATTCGCC AspGluPheAla	<i>EcoRI</i> (*S5)	GATGAATTCGCC AspGluPheAla	Phe	TTC TTT	0.42 0.58
GGCGAATTCCTCT GlyGluPheSer	<i>EcoRI</i> (*S6)	GGCGAATTCCTCT GlyGluPheSer	Phe	TTC TTT	0.42 0.58
TATAAAAGCTTAAGC TyrLysSerLeuSer	<i>HindIII</i> (*S7)	TATAAAAGCTTAAGC TyrLysSerLeuSer	Ser	AGC AGT	0.25 0.16

Table 28: Summary of modifications made to remove *EcoRI* and *HindIII* sites from *lvD*, *alsS* and *lvC*. Underlined: Restriction site. Yellow highlight: Affected codons. Dotted underline: Targeted codon for same-sense mutation. Bold red: mutated nucleotide.

Finally, all stop codons were replaced with the strong 5'-TAA-3' stop codon, and as the wild-type start codon for the *alsS* gene was 5'-TTG-3', this was replaced with the stronger 5'-ATG-3' start codon. Thus, the final design for the polynucleotide containing the *ilvD alsS* and *ilvC* genes (Polynucleotide sequence 15) was synthesised by Biomatik Corporation and delivered in the pBSK::ilvD-alsS-ilvC vector. The polynucleotide containing the *ilvD*, *alsS* and *ilvC* genes was then

subcloned into the pSTV28 vector in between the *EcoRI* and *HindIII* restriction sites, to form pSTV28::ilvD-alsS-ilvC (pKIV).

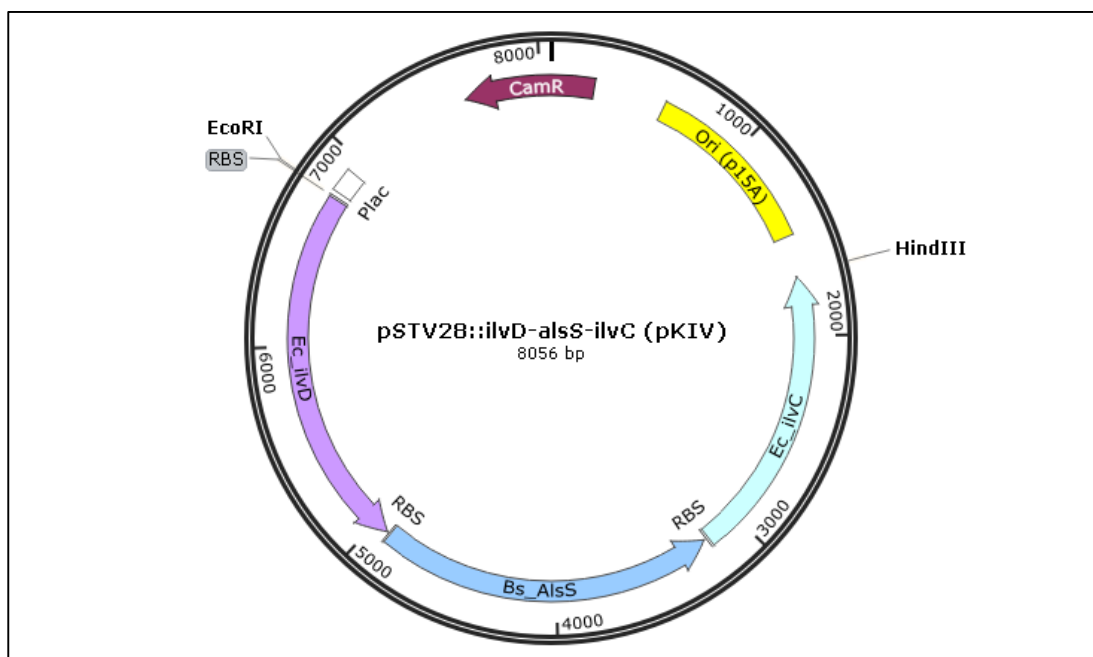


Figure 66: The pSTV28::ilvD-alsS-ilvC (pKIV) plasmid for the co-expression of *E. coli* *ilvD*, *E. coli* *ilvC* and *B. subtilis* *alsS*.

5.4.4.2 Construction and testing of *E. coli* BL21 (DE3) pKIV pHAPP and *E. coli* BL21 (DE3) pKIV pHABS

To construct *E. coli* BL21 (DE3) pKIV pHAPP and *E. coli* BL21 (DE3) pKIV pHABS, *E. coli* BL21 (DE3) cells not harbouring the pLysS plasmid were first transformed with the pHAPP and the pHABS plasmids to form *E. coli* BL21 (DE3) pHAPP and *E. coli* BL21 (DE3) pHABS, respectively. These cells were then made chemically competent and both transformed with the pKIV plasmid. To confirm the presence of the pKIV, pHAPP and pHABS in the two strains, plasmid mini-preps were prepared from the transformants, and four polymerase chain reactions were performed on the mini-preps of the two strains as summarised below (Table 29).

	Reaction A	Reaction B	Reaction C	Reaction D
Strain	<i>E. coli</i> BL21 (DE3) pKIV pHAPP	<i>E. coli</i> BL21 (DE3) pKIV pHABS	<i>E. coli</i> BL21 (DE3) pKIV pHAPP	<i>E. coli</i> BL21 (DE3) pKIV pHABS
Confirm plasmid:	pKIV	pHAPP	pKIV	pHABS
Forward primer	pSTV28_Phu_F	pET20b_Phu_F	pSTV28_Phu_F	pET20b_Phu_F
Reverse primer	pSTV28_Phu_R	OE_ACX4_R	pSTV28_Phu_R	pET20b_Phu_R

Table 29: DNA source and primer combinations used to confirm the presence of pSTV28::ilvD-alsS-ilvC and pET20b(+):4HBT-ACX4-ppBCKD in transformants of *E. coli* BL21 (DE3) pSTV28::ilvD-alsS-ilvC pET20b(+):4HBT-ACX4-ppBCKD and presence of pSTV28::ilvD-alsS-ilvC and pET20b(+):4HBT-ACX4-bsBCKD in transformants of *E. coli* BL21 (DE3) pSTV28::ilvD-alsS-ilvC pET20b(+):4HBT-ACX4-bsBCKD

To rapidly test the production of methacrylic acid by strains overproducing 2-ketoisovalerate, overnight MSX seed cultures of *E. coli* BL21 (DE3) pKIV pHAPP

(10 ml) and *E. coli* BL21 (DE3) pKIV pHABS (4 ml) were prepared. For comparison, overnight seed cultures of *E. coli* BL21 (DE3) pHAPP and *E. coli* BL21 (DE3) pLysS pET20b(+) (control) were also prepared. Strains were cultured with the appropriate antibiotics (Table 6). The overnight seed culture of *E. coli* BL21 (DE3) pKIV pHAPP grew more slowly than the *E. coli* BL21 (DE3) pLysS pHAPP strain had grown in the previous experiment, and did not reach stationary phase after 16 h incubation. Thus, the 100 ml scale MSX cultures of *E. coli* BL21 (DE3) pKIV pHAPP, *E. coli* BL21 (DE3) pHAPP and *E. coli* BL21 (DE3) pLysS pET20b(+) were inoculated to OD₆₀₀ 0.1 as the seed cultures were not strong enough to inoculate to OD₆₀₀ 0.2. The seed culture of *E. coli* BL21 (DE3) pKIV pHABS, however, was used to inoculate an intermediate culture (20 ml) and was incubated until stationary phase as for *E. coli* BL21 (DE3) pLysS pHABS in the previous experiment. The stationary phase *E. coli* BL21 (DE3) pKIV pHABS culture was stored at 4 °C until the next day, when it was then used to inoculate a 100 ml MSX culture supplemented with riboflavin, chloramphenicol and carbencillin. All expression scale cultures (100 ml) were incubated at 37 °C with shaking at 250 rpm and induced at OD₆₀₀ readings of between 0.4 and 0.8 (Table 30).

Culture of <i>E. coli</i> BL21 (DE3)				Final concentrations	
Plasmid #1	Plasmid #2	Excpresion cultures Supplemented with	Induced at OD ₆₀₀ of:	2 ketoisovaleric acid / mM	Methacrylic acid / μM
pKIV	pHAPP	Riboflavin, Cam, C	0.44	1.3	51
-	pHAPP	Riboflavin, C	0.5	0	123
pKIV	pHABS*	Riboflavin, Cam, C	0.78	0	66
pLysS	pET20b(+)	Riboflavin, Cam, C	0.58	0	0

Table 30. Summary of screening results for production of methacrylic acid by *E. coli* BL21 (DE3) pKIV pHAPP, *E. coli* BL21 (DE3) pHAPP and *E. coli* BL21 (DE3) pKIV pHABS in MSX minimal medium. Abbreviations: Riboflavin: Riboflavin, 1μg.ml⁻¹; Cam: Chloramphenicol, 34 μg.ml⁻¹; C: Carbenicillin, 50 μg.ml⁻¹. *Intermediate culture stored at 4 °C overnight before inoculation of main culture.

The best producer of methacrylic acid was the culture of *E. coli* BL21 (DE3) pHAPP, producing a final concentration of 123 μM, which was marginally higher than, but in the same order of magnitude as the 110 μM methacrylic acid produced by *E. coli* BL21 (DE3) pLysS pHAPP in the previous experiment (Table 27). In this experiment, however, *E. coli* BL21 (DE3) pKIV pHABS showed only marginal improvement over the concentration of methacrylic acid produced by *E. coli* BL21 (DE3) pLysS pHABS, and *E. coli* BL21 (DE3) pKIV pHAPP produced less than half the methacrylic acid produced by *E. coli* BL21 (DE3) pLysS pHAPP and *E. coli* BL21 (DE3) pHAPP. It is unclear whether this was caused by the reduction in the starting OD, from 0.2 to 0.1, as insufficient controls and repeats were used. It is probably more likely that the pathway to methacrylic acid was already saturated, and that overexpressing more genes in the form of *alsS*, *ilvC* and *ilvD* could have reduced the overall

expression of the enzymes catalysing the pathway from 2-ketoisovalerate to methacrylate. There are no repeats to confirm this, however, and this does not explain why *E. coli* BL21 (DE3) pKIV pHABS produced more methacrylic acid than *E. coli* BL21 (DE3) pLysS pHABS in the previous experiment (Table 27). Nevertheless, in the case of *E. coli* BL21 (DE3) pKIV pHAPP, a peak with the same retention time as 2-ketoisovaleric acid appeared during the HPLC analysis, correlating to 1.3 mM 2-ketoisovaleric acid, implying that upregulation of 2-ketoisovalerate is occurring in this strain. However, the HPLC sample needs to be spiked with 2-ketoisovalerate to confirm this. Again, it is unclear why this was not observed with *E. coli* BL21 (DE3) pKIV pHABS, and more repeats, under stricter control of expression conditions need to be performed to investigate that.

However, in the case of both *E. coli* BL21 (DE3) pKIV pHAPP and *E. coli* BL21 (DE3) pKIV pHABS, some isobutyric acid in the order of 1 mM appeared to be produced, but the peak areas were again not large enough to be accurately quantified. If these apparent peaks did represent genuine isobutyric acid, however, then this would suggest that the ACX4 catalysed oxidation of isobutyryl-CoA to methacrylyl-CoA is bottlenecked, and that broad specificity thioesterases in *E. coli* are hydrolysing isobutyryl-CoA to isobutyrate. Indeed, during the development of a strain for the production of isobutyric acid, Zhang *et al.* had previously obtained 5.6 g.L⁻¹ (64 mM) isobutyrate in shake-flask cultures of *E. coli* with upregulated *alsS*, *ilvD* and the ppBCKD genes. The *E. coli* was cultured in M9 minimal medium supplemented with 5 g.L⁻¹ yeast extract, 40 g.L⁻¹ glucose, and 100 g.L⁻¹ solid calcium carbonate as an additional buffer system. This prompted an investigation as to whether similar titres of isobutyric acid would be achieved in either *E. coli* BL21 (DE3) pKIV pHAPP or *E. coli* BL21 (DE3) pKIV pHABS when grown under these conditions. Attempts to recreate these conditions from this experiment for the above strains were not always successful, in part due to the culture turning alkaline from the calcium carbonate before the cells started producing any acetate to counter the calcium carbonate. Instead of optimising the culturing conditions for this medium, a modified SM medium, which has previously been used in our laboratory with 15 g.L⁻¹ glucose, was used instead.

Overnight seed cultures (10 ml) of *E. coli* BL21 (DE3) pKIV pHAPP, *E. coli* BL21 (DE3) pKIV pHABS, *E. coli* BL21 (DE3) pHAPP and *E. coli* BL21 (DE3) pHABS were therefore prepared in SM media supplemented with glucose (15 g.L⁻¹), riboflavin and the appropriate antibiotics (Table 6). Expression scale cultures (50 ml) were then inoculated to a starting OD₆₀₀ of 0.1 and incubated at 37 °C with shaking at 250 rpm and were induced at OD₆₀₀ readings of between 0.6 and 0.8 (Table 31).

Cultures were incubated for a further 17 h, and samples were taken and analysed by HPLC.

Culture of <i>E. coli</i> BL21 (DE3)				Final concentrations	
Plasmid #1	Plasmid #2	Excpresion cultures Supplemented with	Induced at OD ₆₀₀ of:	Isobutyric acid / mM	Methacrylic acid / μ M
pKIV	pHAPP	Riboflavin, Cam, C	0.6	13	33
-	pHAPP	Riboflavin, C	0.64	2.3	86
pKIV	pHABS	Riboflavin, Cam, C	0.67	13	25
	pHABS	Riboflavin, C	0.76	2.4	14
pLysS	pET20b(+)	Riboflavin, Cam, C	0.7	0	0

Table 31. Summary of methacrylic acid and isobutyric acid production by *E. coli* BL21 (DE3) pKIV pHAPP and *E. coli* BL21 (DE3) pKIV pHABS in modified SM medium. Abbreviations: Cam: Chloramphenicol; C: Carbenicillin.

In this case, methacrylic acid production was even lower than in MSX minimal media. It wasn't possible to determine 2-ketoisovalerate production accurately, as the peak for this did not resolve from the media components. It is possible that the complex media was not as good for the expression of either ACX4 or 4HBT as minimal medium. It is worth remembering the observation that mostly insoluble ACX4 was expressed in LB medium at 28 °C, compared to when it was expressed in LB medium at 37 °C and MSX medium at both 37 °C and 28 °C.

Regardless, with both *E. coli* BL21 (DE3) pKIV pHAPP and *E. coli* BL21 (DE3) pKIV pHABS, peaks with retention times for isobutyric acid did appear and approximated to 13 mM isobutyric acid (1.13 g.L⁻¹). To confirm that the peaks were indeed isobutyrate, the sample from *E. coli* BL21 (DE3) pKIV pHAPP was also analysed by HPLC using a Phenomenex REZEX ROA ion exchange column, and a peaks with retention time corresponding to that of isobutyrate was obtained, with a peak area corresponding again to approximately 13 mM, with no such peak appearing in the sample from the control strain. Clearly, therefore, attempts to optimise the expression of the genes for the production of methacrylic acid must be made.

There was insufficient time in this project to optimise the expression of *alsS*, *ilvC*, *ilvD* and the genes encoding ACX4, 4HBT, ppBCKD and bsBCKD. However, in order to do so, reliable and reproducible benchmarks are required, with strict control of expression conditions, starting ODs and induction points. There was insufficient time to re-test each culture in triplicate, and so instead, the best methacrylate producing strain and culturing conditions, *E. coli* BL21 (DE3) pHAPP in MSX minimal media supplemented with riboflavin, carbenicillin and chloramphenicol was chosen.

5.4.5 Characterising the production of methacrylic acid over time, and comparison of the effect of riboflavin on host productivity.

In this experiment, the objective was to characterise methacrylic acid production over time by the best strain yet tested for methacrylic acid production, *E. coli* BL21 (DE3) pHAPP, both with and without riboflavin.

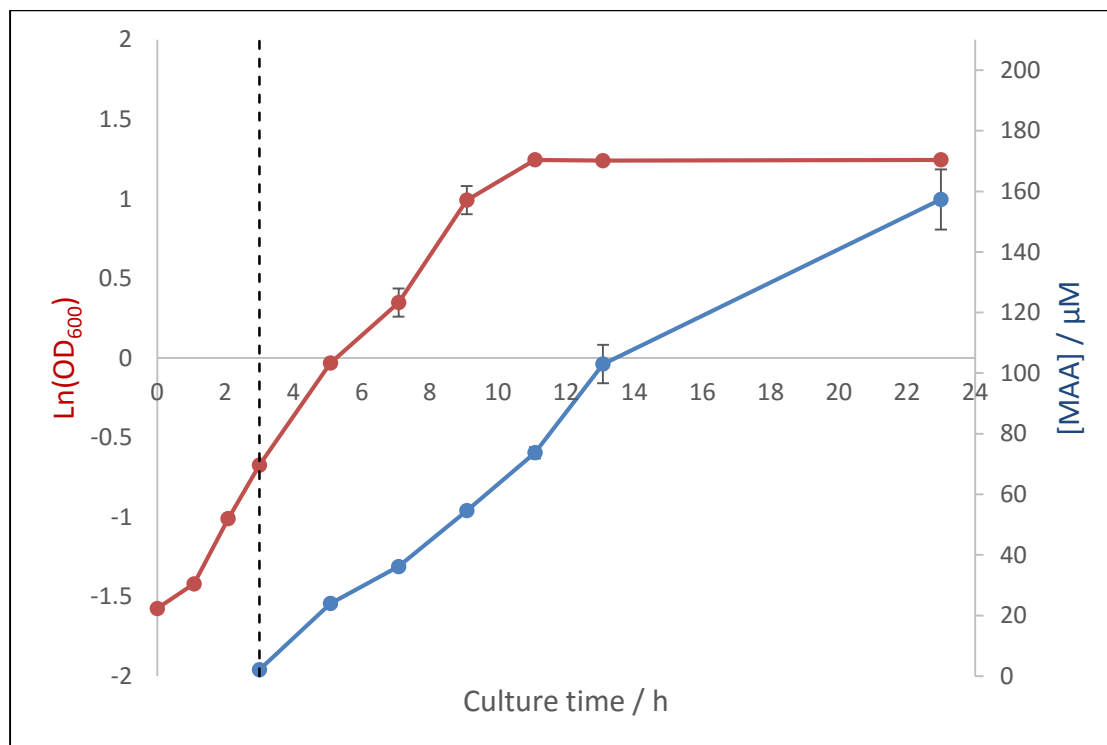


Figure 67: Growth of *E. coli* BL21 (DE3) pET20b(+):4HBT-ACX4-ppBCKD in MSX media supplemented with carbenicillin only, and its production of methacrylic acid with time. ● = $\ln(\text{OD}_{600})$, ● = [MAA]/ μM , ---- = IPTG added. (Triplicate data). No isobutyrate or 2-ketoisovalerate added. Glucose (5 g.L^{-1}) is the only carbon source.

Thus, six MSX seed cultures (10 ml) supplemented with riboflavin were prepared, each from single colonies of *E. coli* BL21 (DE3) pHAPP and incubated overnight (16 h) at 37 °C with shaking at 250 rpm. Each seed culture was used to inoculate an MSX culture (100 ml) to an OD_{600} of 0.2, three of which were supplemented with just carbenicillin, and three of which with both carbenicillin and riboflavin. These cultures were incubated at 37 °C with shaking at 250 rpm, expression was induced at an OD_{600} of 0.5, and samples were taken at intervals thereafter. One of the cultures that was supplemented with riboflavin, however, grew much slower than the other two cultures, and so was discarded.

The time-plot for methacrylic acid production for cultures of *E. coli* BL21 (DE3) pHAPP in MSX media not supplemented with riboflavin shows that methacrylic acid production continues after the culture has reached stationary phase. It is unclear exactly when the culture stops being productive, as the time between the final

sample and the penultimate sample is 10 h, whilst samples were taken every 2 h before that. This continued production is perhaps unsurprising as firstly, the pathway releases NADH and does not require any ATP dependent activation steps, unlike the biotransformation of isobutyric acid to isobutyryl-CoA. The glucose content of the samples was not analysed, and in the future it would be interesting to determine whether the production of methacrylic acid stops due to a lack of glucose, or for other reasons, such as inactivation of the enzymes.

Finally, the final samples taken at 20 h post induction for the cultures not supplemented with riboflavin had methacrylic acid concentrations of $157 \mu\text{M} \pm 9.9 \mu\text{M}$. By contrast, in the two surviving cultures in MSX medium supplemented with riboflavin, samples taken 17 h post induction had methacrylic acid concentrations of $171 \pm 7.6 \mu\text{M}$, the highest level of methacrylic acid production from glucose, yet. Whilst this is a 9% higher concentration than the former culture, the two cultures should not be directly compared, as individual seed cultures were prepared for each culture. A fair comparison should strictly require that the same seed culture be used to inoculate one repeat for each condition, one whereby the media is supplemented with riboflavin, and another which is not, such that 3 seed cultures are used to inoculate six cultures.

Nevertheless, this series of experiments has shown that methacrylic acid can be produced directly from glucose, to a concentration of 0.17 mM. Whilst this is a very small concentration, the fact that in some cultures 13 mM isobutyric acid was produced, indicates that not only are there bottlenecks in the pathway, but there is also unproductive consumption of intermediates. Optimisation of enzyme expression levels and the disruption of enzymes which compete with the key intermediates ought to increase methacrylic acid production.

6 Discussion

In this project, a thioesterase selective for methacrylyl-CoA was identified, as was an acyl-CoA oxidase capable of oxidising isobutyryl-CoA to methacrylyl-CoA, without the need for the associated electron transport system required by acyl-CoA dehydrogenases. When these enzymes were combined with an overexpressed acyl-CoA synthetase or a branched chain keto acid dehydrogenase, methacrylic acid was produced to titres of $367 \mu\text{M} \pm 28 \mu\text{M}$ in a biotransformation of isobutyric acid, and $171 \mu\text{M} \pm 7.6 \mu\text{M}$ direct from glucose, respectively. Though these are not very high titres, this is the first successful biocatalytic production of methacrylic acid reported. Other attempts to sustainably produce methacrylic acid have focused on the biocatalytic production of the chemical precursors 3-hydroxyisobutyric acid [68], 2-hydroxyisobutyric acid [260] and isobutyric acid [16], for combined bio- and chemo- catalytic processes. The only designs for the direct production of methacrylic acid or its esters [20], had not been experimentally substantiated. It is unclear whether any attempts by others to implement any of the designs for direct methacrylic acid production were even made at all, were made but were unsuccessful, or have simply not yet reached the public domain.

Perhaps the two most important achievements of this project were the identification of a thioesterase, 4HBT, which selectively hydrolysed methacrylyl-CoA over its precursor, isobutyryl-CoA, and the identification in the literature of an acyl-CoA oxidase, ACX4, which displayed activity on isobutyryl-CoA. With regards to 4HBT, there is some argument that with sufficient rates of isobutyryl-CoA oxidation to methacrylyl-CoA, this selectivity may not be critical. High rates of isobutyryl-CoA oxidation would result in higher steady state concentrations of methacrylyl-CoA over isobutyryl-CoA, and would keep the enzyme saturated with methacrylyl-CoA. However, the results of this project indicated strongly that there was a bottleneck in the oxidation of isobutyryl-CoA to methacrylyl-CoA when methacrylic acid was produced from glucose, resulting in the hydrolysis of isobutyryl-CoA to isobutyric acid. Thus, a less selective thioesterase than 4HBT may have exacerbated this problem and led to yet less methacrylic acid being produced. With regards to using ACX4, an acyl-CoA oxidase, for the oxidation of isobutyryl-CoA over an acyl-CoA dehydrogenase, the need for an ETF and ETFQO to transfer the liberated electrons to the organism's electron transport chain (ETC) was avoided. This is significant, as there is considerable evidence that the individual primary dehydrogenases, ETFs and ETFQOs from different organisms are not necessarily compatible with each other, and so the cloning of just an isobutyryl-CoA dehydrogenase into *E. coli* may not be guaranteed to succeed. Indeed, there is evidence that *E. coli* fadB has

combined acyl-CoA dehydrogenase and ETF function [166], meaning that there may well not be a distinct ETF in *E. coli*, let alone one which would be functional with a recombinant isobutyryl-CoA dehydrogenase. It is not immediately obvious in the metabolic pathway and enzyme activity databases that the acyl-CoA dehydrogenases are dependent on the presence of a compatible ETF and ETFQO. Indeed, in the patent published by Genomatica for the design of pathways to methacrylic acid [20], and even in a peer reviewed paper for the production of 3-hydroxyisobutyric acid in *Pseudomonas taiwanensis* [17], an organism which does express the valine degradation I pathway, it was assumed that the acyl-CoA dehydrogenase for the oxidation of isobutyryl-CoA was NAD⁺ dependent, and not FAD dependent. This project almost fell into the same trap, and whilst it was at least recognised early on that the acyl-CoA dehydrogenases were FAD dependent, it was not recognised that these enzymes needed an electron transport system for their cofactors to be re-oxidised. It was only after a personal communication with Wataru Mizunashi of Mitsubishi Rayon Company, who own Lucite International, the sponsors of this project, that this issue was discovered. This should act as a warning against over-reliance on databases and automated pathway design algorithms in the design of metabolic pathways, and as guides for the selection of suitable enzymes for those pathways. They are clearly useful tools for the identification of metabolic routes and enzymes, but are not yet mature enough to be depended on.

6.1 Recent advances in the literature relevant to the biocatalytic production of methacrylic acid.

6.1.1 Metabolic pathway to acrylic acid.

A patent for the production of acrylic acid was recently published whereby a metabolic pathway from threonine was described [261].

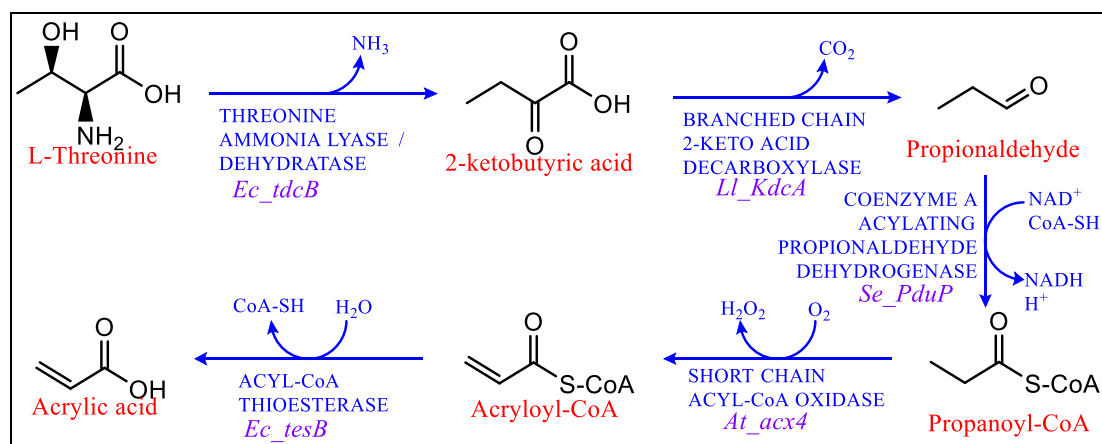


Figure 68: An engineered pathway from L-threonine to acrylic acid [261]. *Ec_tdcB*: Threonine ammonia lyase from *E. coli*. *Ll_KdcA*: Branched chain keto acid decarboxylase from *L. lactis*. *Se_PduP*: Coenzyme A acylating propionaldehyde dehydrogenase from *S. enteritica*. *At_acx4*: short chain acyl-CoA oxidase from *A. thaliana*. *Ec_tesB*: Thioesterase II from *E. coli*.

Cultures of *E. coli* expressing the pathway outlined above (Figure 68) were cultured in minimal media supplemented with various vitamins. The cell culture broth was analysed, as well as the cells themselves for intracellular levels of propionyl-CoA and acryloyl-CoA (Table 32).

Expressed Heterologous genes	Threonine added (g.L ⁻¹)	2-ketobutyrate in broth (μg.ml ⁻¹)	Propionyl-CoA (ng.ml ⁻¹)	Acryloyl-CoA (ng.ml ⁻¹)	Acrylic acid in broth (μg.ml ⁻¹)
<i>tdcB, kdA, pduP, acx4, tesB</i>	0	<0.25	204	7.3	0.44
<i>tdcB, kdcA, pduP, acx4</i>	0	5.1	415	75	0.21
None	0	<0.25	9.3	<2.5	0.03
<i>tdcB, kdA, pduP, acx4, tesB</i>	1	14.7	317	9.1	0.62
<i>tdcB, kdcA, pduP, acx4</i>	1	31.0	425	85	0.27
None	1	1.0	8.8	1.9	0.08

Table 32: Summary of results from 'Microorganisms and methods for producing acrylate and other products from propionyl-CoA' for the production of acrylic acid. (Taken from [261])

The most acrylic acid obtained was 0.62 μg.ml⁻¹ (8 μM), and this was when 1 g.L⁻¹ threonine was supplemented to the culture. Though the pathway to methacrylic acid developed in this project was considerably more productive than the pathway to acrylic acid described here, it is perhaps unfair to make direct comparisons as they are separate pathways. However, comparisons can be made with the use of the acyl-CoA oxidase ACX4 from *A. thaliana*. The higher intracellular levels of propionyl-CoA than acrylyl-CoA suggest inefficient oxidation of the propionyl-CoA to acrylyl-CoA. Though isobutyryl-CoA is bulkier than propionyl-CoA, it has the same acyl-chain length, and so there may be unfavourable ACX4 enzyme kinetics for isobutyryl-CoA and propionyl-CoA. Alternatively, there may also be some inherent difficulty to express ACX4, the only enzyme from a eukaryotic source used in the pathway to acrylic acid [261], as well as to methacrylic acid in this project.

6.1.2 Production of 3-hydroxyisobutyric acid, a late intermediate of the valine degradation I pathway

Recently, 3-hydroxyisobutyric acid, an intermediate of the valine degradation I pathway, was produced to a final concentration of 22 mM (2.3 g.L⁻¹) from glucose in a repeated batch culture, to a yield of 0.138 g.g⁻¹ glucose (24% of maximum) [17]. The strain used in this pathway, *Pseudomonas taiwanensis* VLB120 B83 T7, was developed from *P. taiwanensis* VLB120 in pursuit of a strain for the overproduction of isobutyric acid [242], whereby the BCKD was disrupted to prevent loss of 2-ketoisovalerate to the valine degradation I pathway, and cells were subjected to a regime of random mutagenesis to attenuate growth on

isobutyrate. One of the mutants, B83, had a mutated 3-hydroxyisobutyrate dehydrogenase enzyme in the valine degradation I pathway, and an indigenous isobutyrate-CoA ligase activated the isobutyrate to isobutyryl-CoA of the valine degradation I pathway (Figure 69).

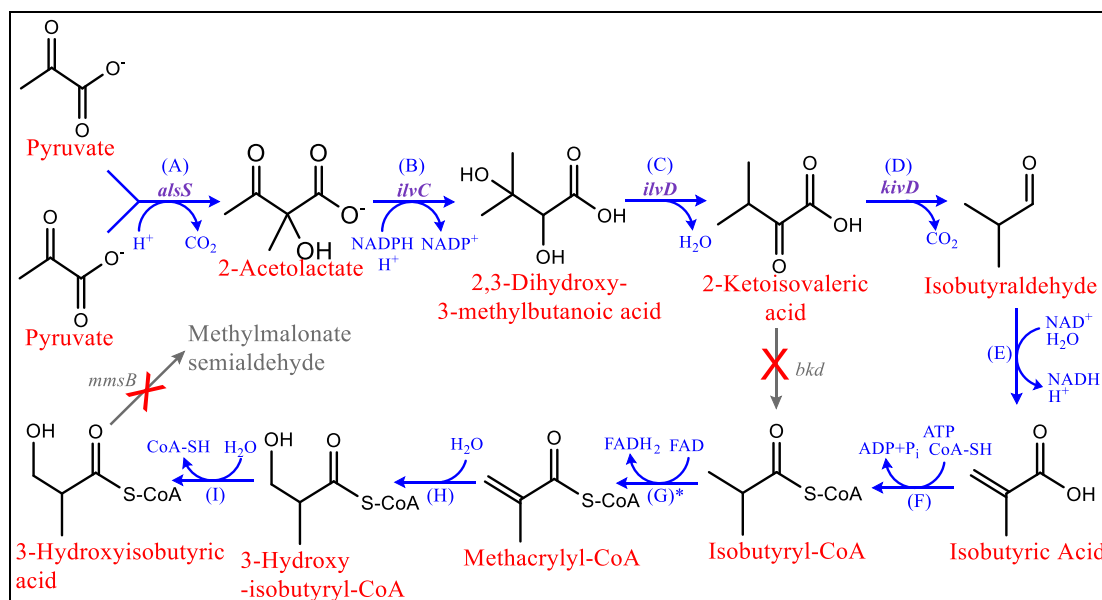


Figure 69: Semi-engineered metabolic pathway to 3-hydroxyisobutyric acid in *Pseudomonas taiwanensis* VLB120 B83 T7. (A) Acetohydroxyacid/acetolactate synthase. (B) Acetohydroxyacid isomeroreductase. (C) Dihydroxyacid dehydratase. (D) 2-Keto-acid decarboxylase. (E) Aldehyde dehydrogenase. (F) Acyl-CoA synthetase. (G) Acyl-CoA dehydrogenase. (H) Enoyl-CoA hydratase. (I) 3-hydroxyisobutyryl-CoA hydrolase. Steps A to D were upregulated by overexpression of *alsS*, *ilvC*, *ilvD* and *kivD*. Remaining steps are natural steps of host strain. * = This is an FAD dependent enzyme. Authors of original paper mislabelled it as an NAD⁺ dependent enzyme.

Therefore, 3-hydroxyisobutyrate accumulated in screening experiments. This was recognised as an industrially useful product, and so *alsS*, *ilvC*, *ilvD* and *kivD*, a 2-keto-acid decarboxylase, were upregulated. The same aldehyde dehydrogenases were then relied upon to convert the isobutyrylaldehyde to isobutyric acid and the same indigenous isobutyrate-CoA ligase and valine degradation I enzymes (Figure 69) were relied upon for the production of 3HIBA. In the repeated batch culture of this organism that yielded 22mM 3HIBA, isobutyric acid also accumulated (~17.5 mM). The strain was originally constructed for that purpose, however, and this accumulation was attributed to inefficient activation of the isobutyrate to isobutyryl-CoA by the indigenous ligase, which was not upregulated.

Despite the various inefficiencies of this strain, 22 mM (2.3 g.L⁻¹) 3-hydroxyisobutyrate, with a yield of 0.138 g.g⁻¹ glucose, far outweighs the concentrations of methacrylic acid produced in this project. While direct comparison to the production of 3HIBA is unfair due to the highly active nature of the enoyl-CoA hydratase and 3-hydroxyisobutyryl-CoA hydrolase enzymes in the valine

degradation I pathway [81, 82, 262] compared to the relatively low V_{MAX} and high K_M of 4HBT for methacrylyl-CoA in the *E. coli* pathway, it is possible to envisage expressing 4HBT or another thioesterase in this *Pseudomonas* strain and disrupting the enoyl-CoA hydratase that hydrates methacrylyl-CoA to 3-hydroxyisobutyryl-CoA. Such a route to methacrylic acid in this organism could yield similar titres of methacrylic acid without needing to further upregulate any of the enzymes.

Indeed, product titres could be improved further, as Lang *et al.* did not upregulate the aldehyde dehydrogenase and isobutyrate-CoA ligase, and instead relied upon indigenous enzymes for these activities. Zhang *et al.* [16] have already identified an ideal aldehyde dehydrogenase, and the isobutyrate-CoA ligase that was identified in this project, AcsA, would be useful for the activation of isobutyrate. When considering that ~ 18 mM isobutyrate (~ 1.5 g.L⁻¹) accumulated at the end of the culture, total acids (IBA + 3HIBA) production was ~ 3.8 g.L⁻¹ from less than 40 g.L⁻¹ total glucose in the repeated batch culture. If the aldehyde dehydrogenase and isobutyrate-CoA ligase enzymes were upregulated, the final 3HIBA titre could reach this value, which compares to the 4.8 g.L⁻¹ of isobutyrate that Zhang *et al.* achieved from 40 g.L⁻¹ glucose in shake flasks, before further optimising the pathway that led to 90 g.L⁻¹ isobutyrate in a bioreactor. Thus, it is very likely that with similar optimisations, including deletions for enzymes that compete for 2-ketoisovalerate, such as *ilvE* for the production of valine, *LeuA* for the production of leucine, and *PanB* for the production of pantothenate, methacrylic acid could be produced to high titres in a *Pseudomonad*.

6.2 Further work

The aim of this project was to demonstrate the proof of principle biocatalytic production of methacrylic acid, and this was achieved. During the process, however, some opportunities for further improvement of methacrylic acid production using the metabolic routes and enzymes that were chosen, were noticed, and these opportunities should be addressed. Some improvements to the experimental design can also be made to clear some uncertainties, such as the inclusion of additional controls to ensure that each enzyme in the metabolic pathway is indeed playing the required role. The possible implications of an improved metabolic pathway for methacrylic acid production on the accumulation of toxic intermediates and by-products should also be considered, and ways to address such toxic accumulation should be explored. In addition to the above, the alternative metabolic routes and enzymes that were not investigated during the short time-frame of this project should be explored. Finally, there is an opportunity to broaden the scope of this

project. By developing an *E. coli* host that is capable of accumulating methacrylyl-CoA, the possibility of using this chassis to screen acyl alcohol transferases for the production of methacrylate esters should be investigated.

6.2.1 Improvements to experimental design

During the production of methacrylic acid, either by a biotransformation of isobutyric acid or directly from glucose, control strains were used whereby the empty expression vector was used instead of the production vector, and this showed that any peaks on the HPLC at the same retention time as methacrylic acid, 2-ketoisovaleric acid and isobutyric acid were indeed caused by the production strains, and not by the control strains. However, controls were not performed to determine the contribution of each enzyme in the metabolic pathway for the production of methacrylic acid. For example, one or more indigenous enzymes of *E. coli* may have been compensating for, or contributing to, individual metabolic steps. Thus, additional control plasmids should be used for the experiments, as summarised below.

Experiment	Control plasmid	Purpose
Biotransformation of isobutyric acid to methacrylic acid	pET20b(+):4HBT-ACX4	Determine contribution of AcsA to biotransformation
	pET20b(+):4HBT-AcsA	Determine contribution of ACX4 to biotransformation
	pET20b(+):ACX4-AcsA	Determine contribution of 4HBT to biotransformation
Production of methacrylic acid from glucose	pET20b(+):4HBT-ppBCKD	Determine contribution of ACX4 in pHAPP
	pET20b(+):4HBT-bsBCKD	Determine contribution of ACX4 in pHABS
	pET20b(+):ACX4-ppBCKD	Determine contribution of 4HBT in pHAPP
	pET20b(+):ACX4-bsBCKD	Determine contribution of 4HBT in pHABS

Table 33: Table of additional control plasmids to use in the biotransformation of isobutyric acid to methacrylic acid and the production of methacrylic acid from glucose.

The contribution that both the ppBCKD and bsBCKD genes played to the production of methacrylic acid from glucose was already demonstrated during the biotransformation of isobutyric acid to methacrylic acid, in the control experiment when *E. coli* BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-AcsA was cultured with no isobutyric acid added to the culture media. If there was endogenous oxidative decarboxylation of 2-ketoisovalerate to isobutyryl-CoA, then methacrylic acid would have been produced in the above control. However, the contribution of ACX4 needed to be demonstrated, as an acyl-CoA dehydrogenase activity could have resulted in that same activity. The *in vitro* no enzyme controls that were used during oxidation of isobutyryl-CoA by cell free extracts would not have detected acyl-CoA dehydrogenase activity, as without an intact cell membrane, the electron transport chain of the organism would have not been functional, and thus the electron transport system transferring electrons from the acyl-CoA dehydrogenase to the electron transport system would also have not been functional, thus rendering the acyl-CoA dehydrogenase non-functional *in vitro*. The contribution of ACX4 during

both the biotransformation of isobutyric acid to methacrylic acid as well as the production of methacrylic acid from glucose was also important to determine, as different rates of production of the isobutyryl-CoA intermediate by AcsA, ppBCKD or bsBCKD may have led to different steady state concentrations of isobutyryl-CoA in the two experiments, resulting in different metabolic flux contributions. For similar reasons, the contribution of 4HBT in both the biotransformation and the production directly from glucose needed to be determined. Broad specificity thioesterases, such as tesB and YciA do exist in *E. coli*, and whilst the upregulation of 4HBT was shown to hydrolyse methacrylyl-CoA *in vitro*, it is unclear whether the steady state concentrations of methacrylyl-CoA were sufficient enough for 4HBT, with its K_M of between 2.1mM and 2.7mM for methacrylyl-CoA, to exhibit the same activity *in vivo*. It is possible that at low steady state concentrations of methacrylyl-CoA, broad specificity thioesterases were the major contributors to the hydrolysis of methacrylyl-CoA. Thus, the above controls are important to perform. To rapidly construct all of these control plasmids, however, a more sophisticated approach to plasmid construction, such as Gibson assembly [51], is required.

Finally, while ACX4 was shown in this project to oxidise isobutyryl-CoA to methacrylyl-CoA, a kinetic characterisation of this enzyme was not performed. Whilst this is a further experiment rather an improvement to the experimental design, knowledge of its K_M for, and V_{MAX} with, isobutyryl-CoA, would provide useful information for future experimental designs. Thus, a kinetic characterisation of ACX4 with isobutyryl-CoA as a substrate may be beneficial.

6.2.2 Improvements of existing metabolic routes for methacrylic acid production.

6.2.2.1 Improvement of enzyme expression

6.2.2.1.1 Improve expression of ACX4 for the production from glucose.

In SDS-PAGE gels for both *E. coli* BL21 (DE3) pLysS pHAPP (Figure 62) and *E. coli* BL21 (DE3) pLysS pHABS (not shown), no ACX4 protein could be detected. In addition, in experiments whereby *E. coli* BL21 (DE3) pKIV pHAPP and *E. coli* BL21 (DE3) pKIV pHABS were cultured in SM media, approximately 13 mM isobutyric acid was detected, which was orders of magnitude higher than the concentrations of methacrylic acid detected in any of the cultures during this project. This strongly indicates that the oxidation of isobutyryl-CoA to methacrylyl-CoA is inefficient due to the low levels of ACX4 being expressed, allowing isobutyryl-CoA to accumulate, and the indigenous thioesterases in *E. coli* to hydrolyse it. It is also possible that with the accumulation of high intracellular concentrations of isobutyryl-CoA and the similar K_M values of 4HBT for both isobutyryl-CoA and

methacrylyl-CoA, that 4HBT is being saturated by isobutyryl-CoA, in effect acting as a competitive inhibitor for the methacrylyl-CoA hydrolysis activity due to the far slower turnover of isobutyryl-CoA than methacrylyl-CoA by 4HBT.

It is therefore very important to investigate the effect of improving the expression of ACX4 on methacrylic acid production. Indeed, further work was in progress to optimise ACX4 expression, by modifying the pSTV28 vector using primers pSTV_MOD_F and pSTV_MOD_R to introduce *Nde*I and *Xho*I restriction sites into the pSTV28 plasmid, such that ACX4 could be cloned in to create pSTV28::ACX4. The plan was to co-transform this plasmid with pHAPP into *E. coli* BL21 (DE3) to create *E. coli* BL21 (DE3) pSTV28::ACX4 pHAPP, whereby two copies of the ACX4 gene existed, with one of the ACX4 copies being under the control of its own promoter in the pSTV28 plasmid. This would not have enabled the *in vivo* upregulation of 2-ketoisovalerate using the pKIV vector, but the cells could have been fed with 2-ketoisovalerate, and any increases in methacrylic acid production would have been compared to the same experiment being performed with *E. coli* BL21 (DE3) pHAPP. However, an unforeseen problem with the cloning had occurred, and by the time it was solved, there was no longer enough time for the actual experiment. Future work, therefore, should seek to optimise ACX4 expression and with more time available, a better solution than what was in progress can be implemented. For example, a plasmid with three promoters could be used, with one promoter for the ppBCKD or bsBCKD genes, another for 4HBT, and the third for ACX4. This plasmid could then still be contrasformed with pKIV for the upregulation of 2-ketoisovalerate. The sponsors of this project have indeed contracted Ingenza Ltd to optimise the expression of all genes in the pathway to methacrylic acid using their proprietary InAble™ technology, which puts each gene under the control of its own constitutive promoter.

6.2.2.1.2 Improve expression of AcsA for the biotransformation

In SDS-PAGE gels for *E. coli* BL21 (DE3) pLysS pET20b(+):4HBT-ACX4-AcsA, no soluble AcsA protein could be detected, and during biotransformations of isobutyric acid to methacrylic acid, the majority of the isobutyric acid remained unconsumed. This was attributed to either inefficient activation of isobutyric acid to isobutyryl-CoA, the hydrolysis of isobutyryl-CoA back to isobutyric acid and coenzyme A, or a combination of the two. Thus, the improvement of AcsA expression should be investigated. As the protein did express, but as an insoluble protein, it may be wise to express the protein from its wild-type gene sequence as opposed to the codon-optimised sequence, and it is also advisable to attempt to express the protein from its own promoter. A softer promoter than the T7 promoter

such as the arabinose P_{BAD} promoter [263] or a lac promoter could be beneficial for soluble protein expression. Having already modified the pSTV28 vector to include *NdeI* and *XhoI* restriction sites as described above, the *AcsA* gene could easily be inserted into this plasmid to yield pSTV28::*AcsA*. This plasmid could then be cotransformed with either the pET20b(+>::4HBT-ACX4-*AcsA* plasmid or the pET20b(+>::4HBT-ACX4 plasmid into *E. coli* BL21 (DE3) to yield *E. coli* BL21 (DE3) pSTV28::*AcsA* pET20b(+>::4HBT-ACX4-*AcsA* or *E. coli* BL21 (DE3) pSTV28::*AcsA* pET20b(+>::4HBT-ACX4, and the biotransformation repeated with these strains.

6.2.2.2 Addressing the non-productive consumption of pathway intermediates

6.2.2.2.1 Non-productive consumption of isobutyryl-CoA

There is strong evidence throughout this project that isobutyryl-CoA is being hydrolysed to isobutyric acid by broad specificity thioesterases. There are three thioesterases in *E. coli* with known activity on isobutyryl-CoA, and these are *E. coli* YciA [121], TesA [264] and TesB [265]. TesA is a periplasmic protein and is thus unlikely to contribute to isobutyryl-CoA hydrolysis. *E. coli* BL21 (DE3) Δ YciA and *E. coli* BL21 (DE3) Δ TesB single knockout strains, as well as an *E. coli* BL21 (DE3) Δ YciA Δ TesB double knock-out strain should also be prepared using the method of Datsenko and Wanner [50]. The single knockout strains of *E. coli* K12 should also be available from the Keio collection [266]. Cell free extracts of the single and double knockout strains could therefore be prepared, and screened for activity on isobutyryl-CoA, as well as methacrylyl-CoA. The optimised plasmids for construction of the production of methacrylic acid from glucose, or for the biotransformation of isobutyric acid to methacrylic acid can then be transformed into the single or double knockout strains, and the effect of these knockouts on methacrylic acid production investigated.

6.2.2.2.2 Non-productive consumption of 2-ketoisovalerate

The leucine biosynthetic pathway [90], the valine biosynthetic pathway [90] and the pantothenate biosynthetic pathway [267] are all metabolic pathways which consume 2-ketoisovalerate in prokaryotes. Lang *et al.* investigated the effect of disruption of the *leuA*, *ilvE* and *panB* genes on 3HIBA production [242]. These genes encode the enzymes catalysing the first metabolic steps of the respective pathways above, and Lang *et al.* found the disruption of *ilvE* and *panB* to be particularly effective in boosting 3HIBA production [242]. Thus, it is worth constructing single, double and triple knockouts of the above genes, and investigating the effect on MAA production. It would be necessary to grow the organism in complex, or otherwise suitably supplemented, media, as the disrupted metabolic pathways are

essential for cell growth in minimal media. However, with the development of synthetic biology tools such as bistable toggle switches [268], it may be possible in the future to keep the essential gene expression switched on during a growth phase in minimal media, and to switch gene expression off during a production phase.

6.2.2.2.3 Non-productive consumption of pyruvate

Pyruvate carboxylase as well as pyruvate dehydrogenase have been previously believed to be responsible for non-productive consumption of pyruvate in metabolic pathways that involve the upregulation of 2-ketoisovalerate [242, 269]. Disruptions of these activities could benefit future methacrylic acid production. Again, the use of bistable toggle switches [268] could be used wisely to switch off the expression of these genes during the production phase.

6.2.2.2.4 Possible non-productive consumption of methacrylyl-CoA

It is possible that methacrylyl-CoA is being hydrated to 3-hydroxyisobutyryl-CoA by the *E. coli* enoyl-CoA hydratase for the fatty acid β -oxidation pathway, and that 3-hydroxyisobutyryl-CoA is then hydrolysed by *E. coli* broad specificity thioesterases (Figure 70).

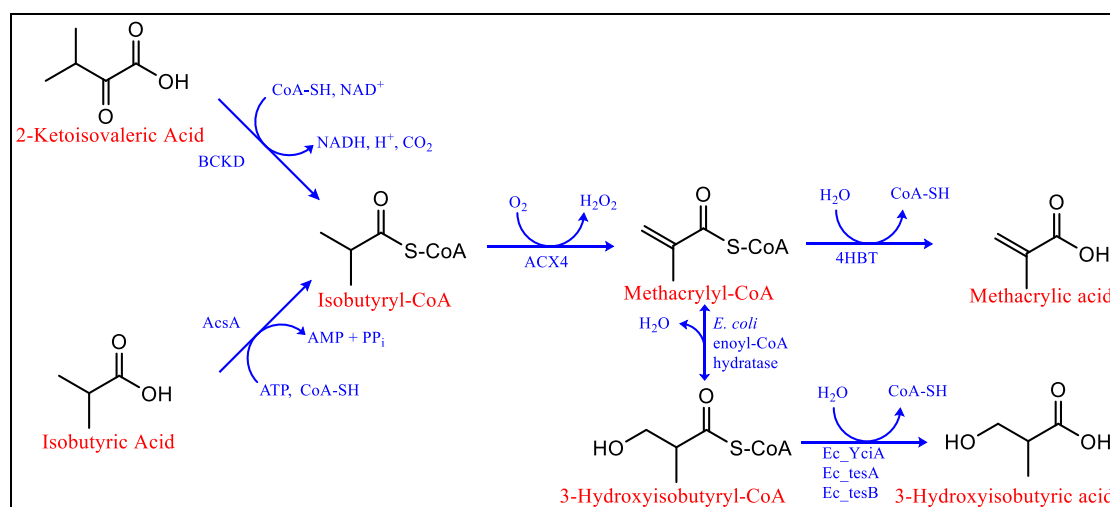


Figure 70: Possible consumption of metharylyl-CoA by hydration to 3-hydroxyisobutyryl-CoA and hydrolysis to 3HIBA in *E. coli*.

This could be investigated by analysing the culture media for 3-hydroxyisobutyrate. A more detailed metabolomics analysis could be performed for 3-hydroxyisobutyryl-CoA production. Finally, methacrylyl-CoA could be reacting with glutathione [82]. It may be unwise to disrupt glutathione detoxification mechanisms, but non-productive consumption of methacrylyl-CoA may otherwise be best addressed by ensuring that its concentration is kept low by ensuring that the terminal metabolic step is efficient.

6.2.2.3 Enzyme improvements and reducing toxic intermediates.

Improvements in the kinetic properties of the enzymes catalysing each step in the metabolic pathways to methacrylic acid would have big advantages for the production of methacrylic acid. Improvements in the V_{MAX} /turnover number of an enzyme for a particular substrate will allow lower levels of enzyme expression for the same metabolic flux, leaving a smaller metabolic burden on the host and greater flexibility in the promoters used, and thus less risk for insoluble protein expression. Greater steady state metabolic flux as a result of improvements to the V_{MAX} /turnover number will reduce non-productive consumption of pathway intermediates relative to productive consumption. Improvements to K_M values will mean that whatever level of protein is expressed will be more effective at lower steady state substrate concentrations. As such, this can prevent a build-up of toxic intermediates, and lower the risk of non-productive consumption of pathway intermediates. Perhaps the two key enzymes to improve the kinetic properties of are 4HBT and ACX4. Crystal structures are available for both enzymes [118, 190], facilitating rational design and directed evolution. An improvement of the K_M of 4HBT would ensure lower accumulation of the toxic intermediate methacrylyl-CoA. Though there was no considerable evidence for methacrylyl-CoA toxicity, the improvements outlined above for the increase in the rate of isobutyryl-CoA oxidation could lead to a build-up of methacrylyl-CoA, and so an increased rate of methacrylyl-CoA hydrolysis would be advantageous. There is some merit to testing a broader specificity thioesterase than 4HBT, such as *YciA* from *H. influenzae* and *tesB* from *E. coli*, which may have naturally higher turnover numbers and lower K_M values for methacrylyl-CoA than 4HBT, so long as the rate of isobutyryl-CoA oxidation is rapid enough. This would of course have to be tested experimentally. Improvements in the V_{MAX} and K_M values of ACX4 for isobutyryl-CoA would result in higher rates of isobutyryl-CoA oxidation at lower concentrations, leave a lower steady state concentration of isobutyryl-CoA, and in turn alleviate any effects of competitive inhibition on 4HBT for methacrylyl-CoA hydrolysis, as well as resulting in more isobutyryl-CoA being oxidised to methacrylyl-CoA than being hydrolysed to isobutyric acid and coenzyme A, should there be no luck in disrupting such background activity. Improvements in the rate of isobutyryl-CoA oxidation would also result in proportional increases in hydrogen peroxide formation, and the *E. coli* oxidative stress response mechanisms may be unable to cope with this increase in H_2O_2 production. It may therefore be wise to increase catalase upregulation or NADH peroxidase activity, which reduces H_2O_2 to $2\text{H}_2\text{O}$ at the expense of one NADH. Alternatively, an acyl-CoA dehydrogenase could be used instead, which does not produce hydrogen peroxide when its FADH_2 is re-oxidised.

6.2.3 Exploration of alternative metabolic routes and host organisms

6.2.3.1 Alternative metabolic routes

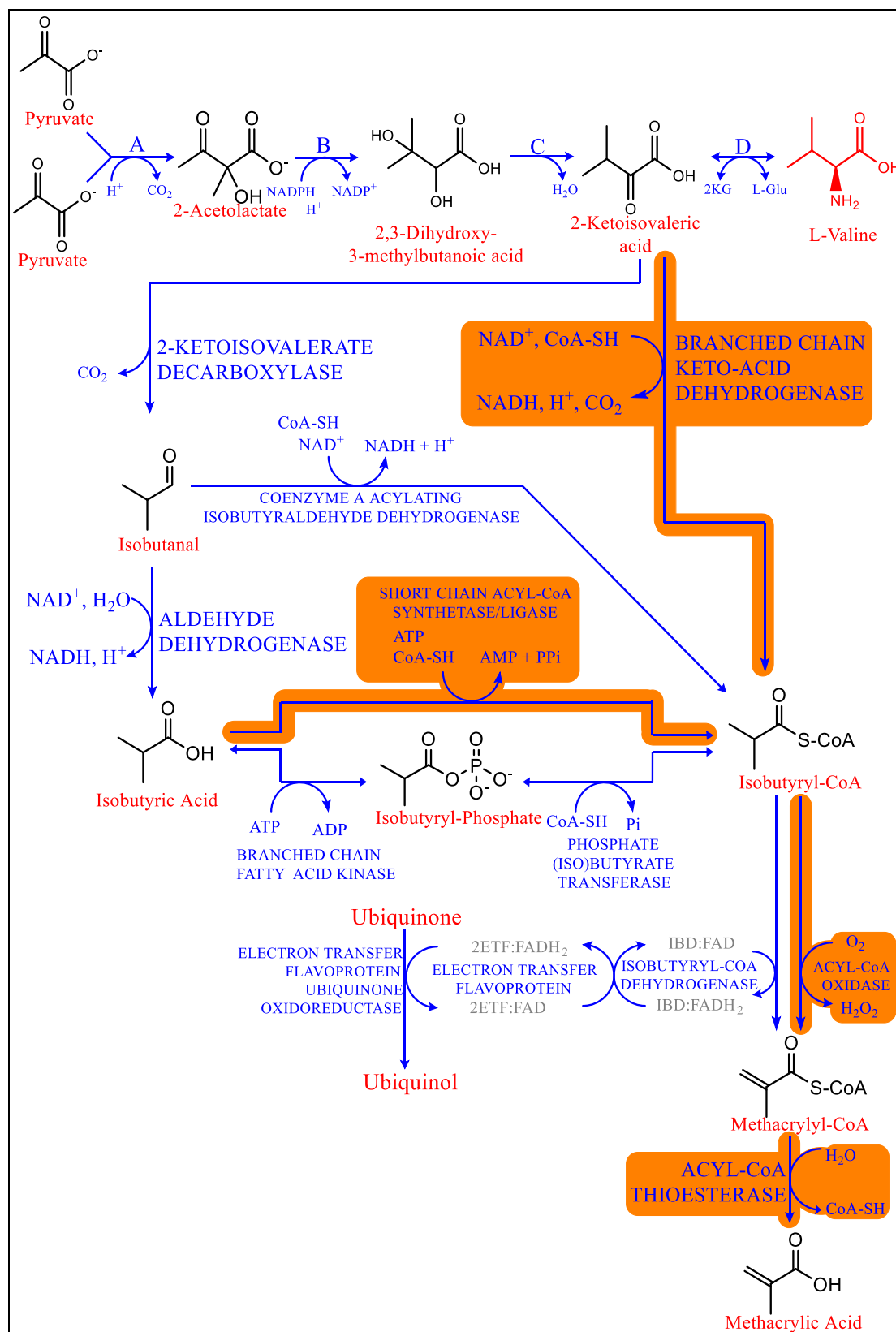


Figure 71: Summary of alternative routes for the production of methacrylic acid from glucose. Highlighted: Routes investigated in this project.

6.2.3.1.1 Producing isobutyryl-CoA from 2-ketoisovalerate via isobutyraldehyde and isobutyrate as an alternative to BCKD

As an alternative to the direct oxidative decarboxylation of 2-ketoisovalerate to isobutyryl-CoA, 2-ketoisovalerate could be first decarboxylated to isobutyraldehyde by a 2-ketoacid decarboxylase, the isobutyraldehyde oxidised to isobutyrate, and the isobutyrate activated to isobutyryl-CoA. The activation step would consume ATP, but as each step is catalysed by discrete enzymes, the optimisation of enzyme expression levels should be easier than for a heteromeric protein complex as is the case with BCKD, and so higher metabolic flux rates could be achieved. This should therefore be considered if the oxidative decarboxylation of 2-ketoisovalerate is found to be a bottleneck.

For the activation of isobutyric acid to isobutyryl-CoA, during the biotransformation of isobutyric acid to methacrylic acid, the AMP forming acyl-CoA synthetase was used. While this had the advantage of catalysing an irreversible reaction, this in effect consumes two ATP molecules, as for AMP to be rephosphorylated back to ATP, it must first be phosphorylated to ADP with another ATP by adenylate cyclase (Figure 72A). The activation of isobutyrate to isobutyryl phosphate only consumes 1ATP, as this is an ADP forming step (Figure 72B). For the biotransformation of isobutyric acid to methacrylic acid, this may well be an advantage, as the biotransformation was found to stop shortly after the cells entered stationary phase. The biotransformation might proceed for longer if more energy is conserved during the biotransformation, and so this route should also be tested.

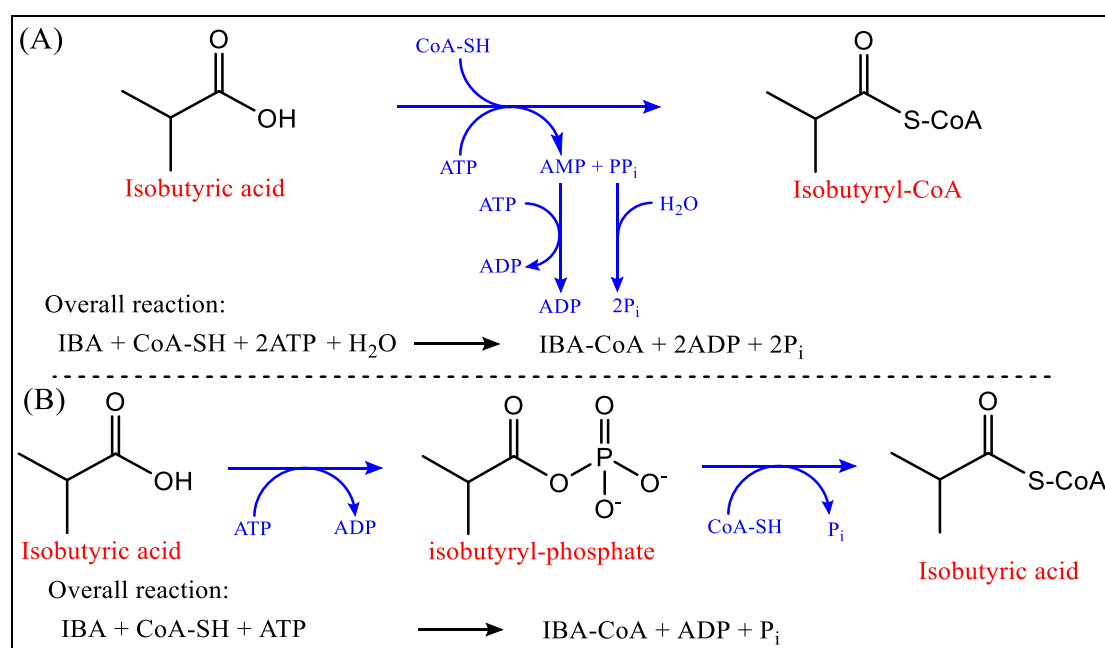


Figure 72: Alternative routes for the activation of isobutyric acid to isobutyryl-CoA, and overall ATP requirement of each.

6.2.3.1.2 Producing isobutyryl-CoA from 2-ketoisovalerate via just isobutyraldehyde

As an alternative to first oxidising isobutyraldehyde to isobutyric acid before this is then activated to isobutyryl-CoA, as described in the above example, isobutyraldehyde could be directly activated to isobutyryl-CoA using a Coenzyme A acylating aldehyde dehydrogenase type enzyme. A similar enzyme was used for the activation of propionaldehyde to propanoyl-CoA in the previously described pathway to acrylic acid (Figure 68, page 145) [261]. Thus, just a 2-ketoisovalerate decarboxylase and a suitable Coenzyme A acylating isobutyraldehyde dehydrogenase enzyme would be required for the formation of isobutyryl-CoA from 2-ketoisovalerate. This has all of the advantages of generating NADH, using discrete enzymes as opposed to an enzyme complex, and perhaps more importantly, not requiring ATP to activate a carboxylic acid to the coenzyme A thioester. Another concern with the proposed pathway that proceeded *via* isobutyric acid was that the enzyme(s) required for the activation of isobutyric acid to methacrylic acid could also re-activate any methacrylic acid product back to methacrylyl-CoA, creating a futile cycle. Such an issue should not occur with the use of a coenzyme A acylating isobutyraldehyde dehydrogenase enzyme. However, though it is known that enzymes catalysing this class of reaction do exist, it is unclear whether any such enzymes with activity on isobutyraldehyde as a substrate exist, and so a search for a suitable candidate would need to be performed.

6.2.3.1.3 Using an acyl-CoA dehydrogenase as an alternative to ACX4

To mitigate the risk of hydrogen peroxide toxicity, an isobutyryl-CoA dehydrogenase and an associated ETF and ETFQO can be used instead. Furthermore, by coupling the re-oxidation of the enzyme's FADH₂ to the electron transport chain, two ATP molecules would be generated, which would be beneficial to the health of the organism. In addition, for the biotransformation of isobutyryl-CoA to methacrylyl-CoA, by supplying the ATP required to activate isobutyric acid to isobutyryl-CoA, this could continue to drive the biotransformation even after all glucose in the medium has been consumed. Finally, a dedicated isobutyryl-CoA dehydrogenase ought to have more favourable kinetic parameters than ACX4, but without a kinetic characterisation of ACX4, this cannot be said for certain.

6.2.3.2 Use of an organism already expressing the valine degradation I pathway

Numerous micro-organisms including *P. putida* and *P. taiwanensis* are known to contain the valine degradation I pathway, and these organisms therefore already have a BCKD and isobutyryl-CoA dehydrogenase with the associated ETF and

ETFQO. The production of methacrylic acid could therefore be achieved in these organisms by the expression of 4HBT. The expression of BCKD is induced by 2-ketoisovalerate in *P. putida*, and the isobutyryl-CoA dehydrogenase is constitutively expressed [86]. As the organisms are capable of growing on valine, expression levels ought to be sufficiently optimised to ensure high metabolic flux to reach respectable titres of product, as recently demonstrated by Lang *et al.* for 3HIBA [17]. The levels of 2-ketoisovalerate required to induce BCKD expression may be achieved by upregulating 2-ketoisovalerate production *via* the co-expression of *alsS*, *ilvC* and *ilvD*, and if this is insufficient, the BCKD could be overexpressed, or a 2-ketoisovalerate decarboxylase, aldehyde dehydrogenase and acyl-CoA synthetase could be upregulated. The 2-ketoisovalerate decarboxylase used by Lang *et al.* [17], the aldehyde dehydrogenase used by Zhang *et al.* [16] as well as the acyl-CoA synthetase *AcsA* identified in this project, should be sufficient for the task. To ensure that sufficient flux is directed to methacrylic acid, however, disruptions in either the enoyl-CoA hydratase (Figure 73Aiii), or the 3-hydroxyisobutyryl-CoA hydrolase (Figure 73Biv) of the pathway would have to be made.

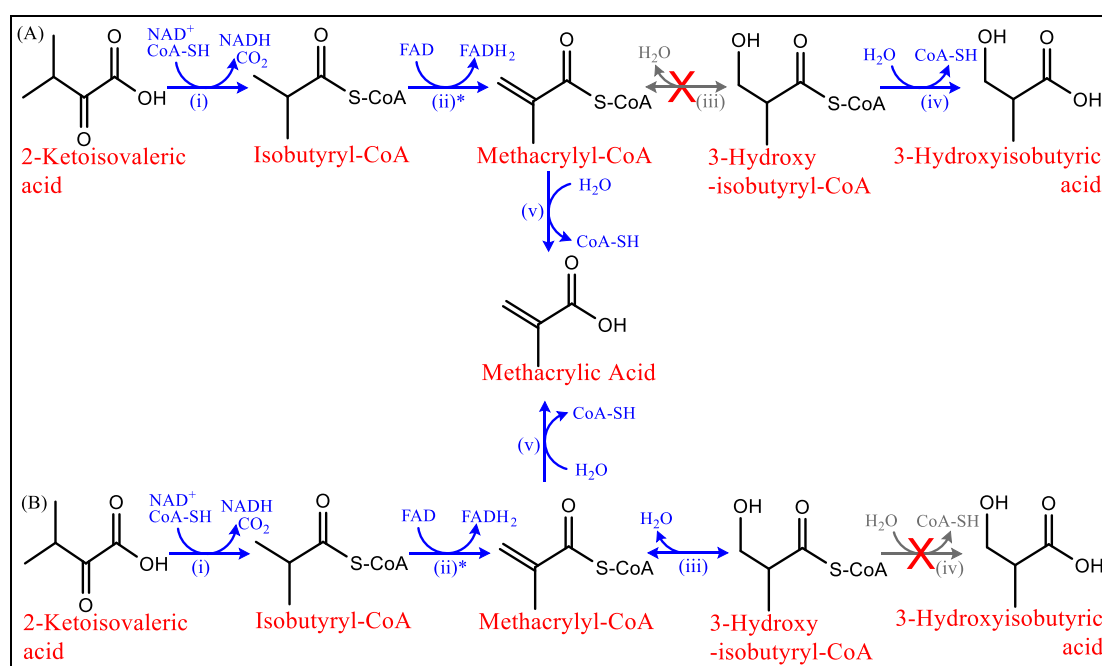


Figure 73: Alternative gene disruptions for accumulation of methacrylyl-CoA in organisms containing the valine degradation I pathway, for the production of methacrylic acid by the expression of a suitable thioesterase. A) Disruption of the enoyl-CoA hydratase gene. B) disruption of the 3-hydroxyisobutyryl-CoA hydrolase gene. i = Branched Chain Keto Acid Dehydrogenase. ii = Acyl-CoA Dehydrogenase. iii) Enoyl-CoA Hydratase. iv) 3-hydroxyisobutyryl-CoA hydrolase. v) Acyl-CoA Thioesterase.

Without these disruptions, the metabolic flux would most likely continue down the valine degradation I pathway, with very little diverting to methacrylic acid, as the

enoyl-CoA hydratase and 3-hydroxyisobutyryl-CoA hydrolase have evolved to rapidly consume the methacrylyl-CoA, such that very low steady state concentrations of methacrylyl-CoA would exist [82]. While a disruption of the enoyl-CoA hydratase responsible for the hydration of methacrylyl-CoA (Figure 73Aiii) ought to ensure that all metabolic flux is diverted to methacrylic acid production, a disruption of the 3-hydroxyisobutyryl-CoA hydrolase would also ensure a build up of methacrylic acid as the hydration of methacrylyl-CoA is reversible. With no other routes for the 3-hydroxyisobutyryl-CoA to follow, a steady state equilibrium will be reached with methacrylyl-CoA, and carbon flux will be diverted to methacrylic acid, enabling similar if not equal yields of methacrylic acid. The disruption of any of the enzymes further downstream of 3-hydroxyisobutyryl-CoA hydrolase instead, would most likely lead to an accumulation of 3HIBA as the major product, as was observed with Lang. *et al.* [17], and not methacrylic acid.

6.3 Conclusions

This project has demonstrated the proof of principle biocatalytic production of methacrylic acid. Though the final titres that were achieved were, at $171 \mu\text{M} \pm 7.6 \mu\text{M}$ ($14.7 \text{ mg.L}^{-1} \pm 0.65 \text{ mg.L}^{-1}$), low, this is considerably better than other proof of principle productions of similar products [261]. There is plenty of scope for improving this metabolic pathway, through improvements in enzyme expression, disruption of non-productive enzymatic activities on pathway intermediates, as well as considering alternative metabolic routes and host organisms. By developing these routes until an optimised host for the production of methacrylic acid is achieved, the organism could be used as a chassis for the production of methacrylate esters by swapping the thioesterase for a suitable acyl alcohol transferase enzyme, enabling the future production of methacrylate esters directly from glucose in a single bioprocess, and this would be an important step forward in the production of bio-perspex.

7 References

1. Platt, D., *Engineering and High Performance Plastics (Rapra Market Report)*. 2003: Smithers Rapra Technology.
2. Witczak, T., et al., *Liquid-Phase Esterification of Methacrylic Acid with Methanol Catalyzed by Heteropolyacid*. International Journal of Chemical Reactor Engineering, 2010. **8**(1).
3. Bauer, W., *Methacrylic Acid and Derivatives*. Kirk-Othmer Encyclopedia of Chemical Technology. 2000: John Wiley & Sons, Inc.
4. Anastas, P. and N. Eghbali, *Green Chemistry: Principles and Practice*. Chemical Society Reviews, 2010. **39**(1): p. 301-312.
5. Gavrilescu, M. and Y. Chisti, *Biotechnology—a sustainable alternative for chemical industry*. Biotechnology Advances, 2005. **23**(7-8): p. 471-499.
6. Wenda, S., et al., *Industrial biotechnology-the future of green chemistry?* Green Chemistry, 2011. **13**(11): p. 3007-3047.
7. Keasling, J.D., *Manufacturing Molecules Through Metabolic Engineering*. Science, 2010. **330**(6009): p. 1355-1358.
8. Straathof, A.J.J., S. Panke, and A. Schmid, *The production of fine chemicals by biotransformations*. Current Opinion in Biotechnology, 2002. **13**(6): p. 548-556.
9. Schulze, B. and M.G. Wubbolts, *Biocatalysis for industrial production of fine chemicals*. Current Opinion in Biotechnology, 1999. **10**(6): p. 609-615.
10. Schmid, A., et al., *Industrial biocatalysis today and tomorrow*. Nature, 2001. **409**(6817): p. 258-268.
11. Schmid, A., et al., *The use of enzymes in the chemical industry in Europe*. Current Opinion in Biotechnology, 2002. **13**(4): p. 359-366.
12. Bornscheuer, U.T. and R.J. Kazlauskas, *Catalytic Promiscuity in Biocatalysis: Using Old Enzymes to Form New Bonds and Follow New Pathways*. Angewandte Chemie International Edition, 2004. **43**(45): p. 6032-6040.
13. Hult, K. and P. Berglund, *Enzyme promiscuity: mechanism and applications*. Trends in Biotechnology, 2007. **25**(5): p. 231-238.
14. Copley, S.D., *Enzymes with extra talents: moonlighting functions and catalytic promiscuity*. Current Opinion in Chemical Biology, 2003. **7**(2): p. 265-272.
15. van Leeuwen, B.N., et al., *Fermentative production of isobutene*. Appl Microbiol Biotechnol, 2012. **93**(4): p. 1377-87.
16. Zhang, K., M. Xiong, and A.P. Woodruff, *Cells and Methods for Producing Isobutyric Acid*. 2012, Regents of the University of Minnesota.
17. Lang, K., K. Buehler, and A. Schmid, *Multistep Synthesis of (S)-3-Hydroxyisobutyric Acid from Glucose using Pseudomonas taiwanensis VLB120 B83 T7 Catalytic Biofilms*. Advanced Synthesis & Catalysis, 2015. **357**(8): p. 1919-1927.
18. Papagianni, M., *Advances in citric acid fermentation by Aspergillus niger: Biochemical aspects, membrane transport and modeling*. Biotechnology Advances, 2007. **25**(3): p. 244-263.
19. Carlsson, M., et al., *Study of the sequential conversion of citric to itaconic to methacrylic acid in near-critical and supercritical water*. Industrial & Engineering Chemistry Research, 1994. **33**(8): p. 1989-1996.
20. Burgard, A.P., et al., *Microorganisms for producing methacrylic acid and methacrylate esters and methods related thereto*. Patent number WO2012/135789A2. 2012, Genomatica, Inc.
21. Ali, U., K.J.B.A. Karim, and N.A. Buang, *A Review of the Properties and Applications of Poly (Methyl Methacrylate) (PMMA)*. Polymer Reviews, 2015. **55**(4): p. 678-705.

22. Salkind, M., E.H. Riddle, and R.W. Keefer, *Acrylates and Methacrylates. Raw Materials, Intermediates, Plant Integration*. Industrial & Engineering Chemistry, 1959. **51**(10): p. 1232-1238.
23. Salkind, M., E.H. Riddle, and R.W. Keefer, *Acrylates and Methacrylates. Ester Manufacture and Markets*. Industrial & Engineering Chemistry, 1959. **51**(11): p. 1328-1334.
24. Bauer, W., *Methacrylic Acid and Derivatives*. Ullman's Encyclopedia of Industrial Chemistry. 2011: Wiley-VCH.
25. Eastham, G., *Presentation on Lucite International's industrial activities*. 2011.
26. Demain, A.L., *History of Industrial Biotechnology*, in *Industrial Biotechnology*. 2010, Wiley-VCH Verlag GmbH & Co. KGaA. p. 17-77.
27. Max, B., et al., *Biotechnological production of citric acid*. Braz J Microbiol, 2010. **41**(4): p. 862-75.
28. Bar-Even, A., et al., *Rethinking glycolysis: on the biochemical logic of metabolic pathways*. Nat Chem Biol, 2012. **8**(6): p. 509-517.
29. Jones, D.T. and D.R. Woods, *Acetone-butanol fermentation revisited*. Microbiological Reviews, 1986. **50**(4): p. 484-524.
30. Zheng, J., et al., *Recent advances to improve fermentative butanol production: Genetic engineering and fermentation technology*. Journal of Bioscience and Bioengineering, 2015. **119**(1): p. 1-9.
31. SOURKES, T.L., *Moritz Traube, 1826-1894: His Contribution to Biochemistry*. Journal of the History of Medicine and Allied Sciences, 1955. **X**(4): p. 379-391.
32. Watson, J.D. and F.H.C. Crick, *A Structure for Deoxyribose Nucleic Acid*. Nature, 1953. **171**: p. 737-738.
33. Crick, F.H.C., *On protein synthesis*. Symposia of the Society for Experimental Biology, 1958. **12**: p. 138-163.
34. Berg, P. and J.E. Mertz, *Personal Reflections on the Origins and Emergence of Recombinant DNA Technology*. Genetics, 2010. **184**(1): p. 9-17.
35. Lehnman, I.R., *DNA Ligase: Structure, Mechanism, and Function*. Science, 1974. **186**(4166): p. 790-797.
36. Smith, H.O. and K.W. Welcox, *A Restriction enzyme from Hemophilus influenzae: I. Purification and general properties*. Journal of Molecular Biology, 1970. **51**(2): p. 379-391.
37. Kelly Jr, T.J. and H.O. Smith, *A restriction enzyme from Hemophilus influenzae: II. Base sequence of the recognition site*. Journal of Molecular Biology, 1970. **51**(2): p. 393-409.
38. Jackson, D.A., R.H. Symons, and P. Berg, *Biochemical Method for Inserting New Genetic Information into DNA of Simian Virus 40: Circular SV40 DNA Molecules Containing Lambda Phage Genes and the Galactose Operon of Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America, 1972. **69**(10): p. 2904-2909.
39. Cohen, S.N., et al., *Construction of Biologically Functional Bacterial Plasmids In Vitro*. Proceedings of the National Academy of Sciences of the United States of America, 1973. **70**(11): p. 3240-3244.
40. del Solar, G., et al., *Replication and control of circular bacterial plasmids*. Microbiol Mol Biol Rev, 1998. **62**(2): p. 434-64.
41. Summers, D.K., *The biology of plasmids*. 1st ed. 1996: Wiley-Blackwell.
42. Nielsen, J., et al., *Engineering synergy in biotechnology*. Nat Chem Biol, 2014. **10**(5): p. 319-322.
43. Stephanopoulos, G., *Synthetic Biology and Metabolic Engineering*. ACS Synthetic Biology, 2012. **1**(11): p. 514-525.
44. Anderson, S., et al., *Production of 2-Keto-L-Gulonate, an Intermediate in L-Ascorbate Synthesis, by a Genetically Modified Erwinia herbicola*. Science, 1985. **230**(4722): p. 144-149.

45. Bailey, J., *Toward a science of metabolic engineering*. Science, 1991. **252**(5013): p. 1668-1675.
46. Lazarus, R., et al., *A Biocatalytic Approach to Vitamin C Production*, in *Biocatalysis*, D. Abramowicz, Editor. 1990, Springer Netherlands. p. 135-155.
47. Lazarus, R.A. and E. al. Proceedings of the Sixth International Symposium on Genetics of Industrial Microorganisms. Vol. 2. 1990, Strasbourg: Société Française de Microbiologie.
48. Saiki, R., et al., *Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia*. Science, 1985. **230**(4732): p. 1350-1354.
49. Higuchi, R., B. Krummel, and R.K. Saiki, *A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions*. Nucleic Acids Research, 1988. **16**(15): p. 7351-7367.
50. Datsenko, K.A. and B.L. Wanner, *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products*. Proc Natl Acad Sci U S A, 2000. **97**(12): p. 6640-5.
51. Gibson, D.G., et al., *Enzymatic assembly of DNA molecules up to several hundred kilobases*. Nat Meth, 2009. **6**(5): p. 343-345.
52. Reese, C.B., *Oligo- and poly-nucleotides: 50 years of chemical synthesis*. Organic & Biomolecular Chemistry, 2005. **3**(21): p. 3851-3868.
53. Puigbò, P., et al., *OPTIMIZER: a web server for optimizing the codon usage of DNA sequences*. Nucleic Acids Research, 2007. **35**(suppl 2): p. W126-W131.
54. Nakamura, C.E. and G.M. Whited, *Metabolic engineering for the microbial production of 1,3-propanediol*. Current Opinion in Biotechnology, 2003. **14**(5): p. 454-459.
55. Deckwer, W.-D., *Microbial conversion of glycerol to 1,3-propanediol*. FEMS Microbiology Reviews, 1995. **16**(2-3): p. 143-149.
56. Gancedo, C., J.M. Gancedo, and A. Sols, *Glycerol Metabolism in Yeasts. Pathways of Utilization and Production*. European Journal of Biochemistry, 1968. **5**(2): p. 165-172.
57. Streekstra, H., et al., *Overflow metabolism during anaerobic growth of Klebsiella aerogenes NCTC 418 on glycerol and dihydroxyacetone in chemostat culture*. Archives of Microbiology, 1987. **147**(3): p. 268-275.
58. Jensen, R.A., *Enzyme Recruitment in Evolution of New Function*. Annual Review of Microbiology, 1976. **30**(1): p. 409-425.
59. O'Brien, P.J. and D. Herschlag, *Catalytic promiscuity and the evolution of new enzymatic activities*. Chemistry & Biology, 1999. **6**(4): p. R91-R105.
60. Smith, K.M., K.M. Cho, and J.C. Liao, *Engineering Corynebacterium glutamicum for isobutanol production*. Appl Microbiol Biotechnol, 2010. **87**(3): p. 1045-55.
61. Atsumi, S., T. Hanai, and J.C. Liao, *Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels*. Nature, 2008. **451**(7174): p. 86-9.
62. Hishida, M., et al., *Biomass Syngas Production Technology by Gasification for Liquid Fuel and Other Chemicals*. Mitsubishi Heavy Industries Technical Review, 2011. **48**(3): p. 37-41.
63. *Bio-methanol Production*. IRENA-IEA-ETSAP Technology Briefs 2013 [cited 2015].
64. Pirie, J.M., *The Manufacture of Hydrocyanic Acid by the Andrussow Process*. Platinum Metals Review, 1958. **2**(1): p. 7-11.
65. *Bio-ethylene Production*. IRENA-IEA-ETSAP Technology Briefs 2013 [cited 2015].
66. Reuss, G., et al., *Formaldehyde*, in *Ullmann's Encyclopedia of Industrial Chemistry*. 2000, Wiley-VCH Verlag GmbH & Co. KGaA.

67. Marx, A., et al., *Microbiological production of 3-hydroxyisobutyric acid*. 2010.
68. Haas, T., et al., *Biotechnological production of 3-hydroxyisobutyric acid*. 2013, Google Patents.
69. Clayden, J., et al., *Organic Chemistry*. 2001: Oxford University Press.
70. *Enzyme Nomenclature: Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes*. 2nd ed. 1992: Academic Press.
71. Caspi, R., et al., *The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases*. Nucleic Acids Research, 2014. **42**(D1): p. D459-D471.
72. Kanehisa, M. and S. Goto, *KEGG: kyoto encyclopedia of genes and genomes*. Nucleic Acids Res, 2000. **28**(1): p. 27-30.
73. Yim, H., et al., *Metabolic engineering of Escherichia coli for direct production of 1,4-butanediol*. Nat Chem Biol, 2011. **7**(7): p. 445-452.
74. Carbonell, P., et al., *A retrosynthetic biology approach to metabolic pathway design for therapeutic production*. BMC Syst Biol, 2011. **5**: p. 122.
75. BURK, M., J., et al., *Microorganisms for production of 1,4-butanediol via in silico modeling of metabolic pathways and genetic engineering of enzymes*. 2010, Genomatica, Inc.
76. Cronk, J.D. Lecture 26 - Introduction to metabolism: Metabolism and thermodynamics. Metabolic flux and its control. "High energy" compounds. Oxidation-reduction reactions. 2014; Available from: <http://guweb2.gonzaga.edu/faculty/cronk/CHEM440pub/L26-index.cfm?L26resource=thioesters>.
77. Dimroth, P. and B. Schink, *Energy conservation in the decarboxylation of dicarboxylic acids by fermenting bacteria*. Arch Microbiol, 1998. **170**(2): p. 69-77.
78. Dewick, P.M., *Essentials of Organic Chemistry: For Students of Pharmacy, Medical Chemistry and Biological Chemistry*. 2006: Wiley-Blackwell.
79. Heider, J., *A new family of CoA-transferases*. FEBS Letters. **509**(3): p. 345-349.
80. Jones, A.R., et al., *Probing Reversible Chemistry in Coenzyme B(12)-Dependent Ethanolamine Ammonia Lyase with Kinetic Isotope Effects*. Chemistry (Weinheim an Der Bergstrasse, Germany), 2015. **21**(24): p. 8826-8831.
81. Ishigure, K., et al., *Human liver disease decreases methacrylyl-CoA hydratase and β -hydroxyisobutyryl-CoA hydrolase activities in valine catabolism*. Clinica Chimica Acta, 2001. **312**(1-2): p. 115-121.
82. Shimomura, Y., et al., *Purification and partial characterization of 3-hydroxyisobutyryl-coenzyme A hydrolase of rat liver*. Journal of Biological Chemistry, 1994. **269**(19): p. 14248-14253.
83. Gerbling, H. and B. Gerhardt, *Oxidative decarboxylation of branched-chain 2-oxo Fatty acids by higher plant peroxisomes*. Plant Physiol, 1988. **88**(1): p. 13-5.
84. Gerbling, H. and B. Gerhardt, *Peroxisomal degradation of branched-chain 2-oxo acids*. Plant Physiol, 1989. **91**(4): p. 1387-92.
85. Massey, L.K., J.R. Sokatch, and R.S. Conrad, *Branched-chain amino acid catabolism in bacteria*. Bacteriol Rev, 1976. **40**(1): p. 42-54.
86. Marshall, V.D. and J.R. Sokatch, *Regulation of valine catabolism in Pseudomonas putida*. J Bacteriol, 1972. **110**(3): p. 1073-81.
87. Lehninger, A.L., D.L. Nelson, and M.M. Cox, *Lehninger Principles of Biochemistry*. 2005: W. H. Freeman.
88. Nguyen, T.V., et al., *Identification of isobutyryl-CoA dehydrogenase and its deficiency in humans*. Mol Genet Metab, 2002. **77**(1-2): p. 68-79.

89. Kim, J.-J.P. and R. Miura, *Acyl-CoA dehydrogenases and acyl-CoA oxidases. Structural basis for mechanistic similarities and differences*. European Journal of Biochemistry, 2004. **271**(3): p. 483-493.
90. Umbarger, H.E., *Biosynthesis of the Branched-chain Amino Acids*. 2 ed. Escherichia coli and Salmonella: Cellular and Molecular biology. Vol. 1. 1995.
91. Eggeling, L. and H. Sahm, *L -Glutamate and L -lysine: traditional products with impetuous developments*. Applied Microbiology and Biotechnology, 1999. **52**(2): p. 146-153.
92. Shimizu, H. and T. Hirasawa, *Production of Glutamate and Glutamate-Related Amino Acids: Molecular Mechanism Analysis and Metabolic Engineering*. 2007. **5**: p. 1-38.
93. Woolston, B.M., S. Edgar, and G. Stephanopoulos, *Metabolic Engineering: Past and Future*. Annual Review of Chemical and Biomolecular Engineering, Vol 4, 2013. **4**: p. 259-288.
94. Cropp, T.A., et al., *Fatty-acid biosynthesis in a branched-chain alpha-keto acid dehydrogenase mutant of Streptomyces avermitilis*. Can J Microbiol, 2000. **46**(6): p. 506-14.
95. Zhang, Y.X., et al., *Genes encoding acyl-CoA dehydrogenase (AcdH) homologues from Streptomyces coelicolor and Streptomyces avermitilis provide insights into the metabolism of small branched-chain fatty acids and macrolide antibiotic production*. Microbiology-Sgm, 1999. **145**: p. 2323-2334.
96. Martin, R.R., et al., *Common enzymes of branched-chain amino acid catabolism in Pseudomonas putida*. J Bacteriol, 1973. **115**(1): p. 198-204.
97. Sokatch, J.R., *Alanine and Aspartate Formation During Growth on Valine-C14 by Pseudomonas aeruginosa*. Journal of Bacteriology, 1966. **92**(1): p. 72-75.
98. Allison, M.J. and J.L. Peel, *The biosynthesis of valine from isobutyrate by peptostreptococcus elsdenii and Bacteroides ruminicola*. Biochem J, 1971. **121**(3): p. 431-7.
99. *E. coli Genotypes*. 11/01/2016]; Available from: [http://openwetware.org/wiki/E. coli_genotypes](http://openwetware.org/wiki/E._coli_genotypes).
100. Zeigler, D.R., et al., *The Origins of 168, W23, and Other Bacillus subtilis Legacy Strains*. Journal of Bacteriology, 2008. **190**(21): p. 6983-6995.
101. Vishniac, W. and M. Santer, *The Thiobacilli*. Bacteriological Reviews, 1957. **21**(3): p. 195-213.
102. Heathcote, M.L., J. Staunton, and P.F. Leadlay, *Role of type II thioesterases: evidence for removal of short acyl chains produced by aberrant decarboxylation of chain extender units*. Chem Biol, 2001. **8**(2): p. 207-20.
103. Kotowska, M. and K. Pawlik, *Roles of type II thioesterases and their application for secondary metabolite yield improvement*. Applied Microbiology and Biotechnology, 2014. **98**(18): p. 7735-7746.
104. Sies, H., *Glutathione and its role in cellular functions*. Free Radical Biology and Medicine, 1999. **27**(9-10): p. 916-921.
105. Wu, G., et al., *Glutathione Metabolism and Its Implications for Health*. The Journal of Nutrition, 2004. **134**(3): p. 489-492.
106. Sidenius, U., et al., *In Vitro Reactivity of Carboxylic Acid-CoA Thioesters with Glutathione†*. Chemical Research in Toxicology, 2004. **17**(1): p. 75-81.
107. Hunt, M.C. and S.E. Alexson, *The role Acyl-CoA thioesterases play in mediating intracellular lipid metabolism*. Prog Lipid Res, 2002. **41**(2): p. 99-130.
108. Hunt, M.C., M.I. Siponen, and S.E.H. Alexson, *The emerging role of acyl-CoA thioesterases and acyltransferases in regulating peroxisomal lipid metabolism*. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 2012. **1822**(9): p. 1397-1410.

109. Ellis, J.M., G.W. Wong, and M.J. Wolfgang, *Acyl Coenzyme A Thioesterase 7 Regulates Neuronal Fatty Acid Metabolism To Prevent Neurotoxicity*. Molecular and Cellular Biology, 2013. **33**(9): p. 1869-1882.
110. Angelini, A., et al., *Structural and enzymatic characterization of HP0496, a YbgC thioesterase from Helicobacter pylori*. Proteins, 2008. **72**(4): p. 1212-21.
111. Knudsen, J., S. Clark, and R. Dils, *Acyl-CoA hydrolase(s) in rabbit mammary gland which control the chain length of fatty acids synthesised*. Biochemical and Biophysical Research Communications, 1975. **65**(3): p. 921-926.
112. Knudsen, J., S. Clark, and R. Dils, *Purification and some properties of a medium-chain acyl-thioester hydrolase from lactating-rabbit mammary gland which terminates chain elongation in fatty acid synthesis*. Vol. 160. 1976. 683-691.
113. Libertini, L.J. and S. Smith, *Purification and properties of a thioesterase from lactating rat mammary gland which modifies the product specificity of fatty acid synthetase*. Journal of Biological Chemistry, 1978. **253**(5): p. 1393-1401.
114. Zheng, Z., et al., *Thioesterase II of Escherichia coli Plays an Important Role in 3-Hydroxydecanoic Acid Production*. Applied and Environmental Microbiology, 2004. **70**(7): p. 3807-3813.
115. Chang, K.H., et al., *Isolation and characterization of the three polypeptide components of 4-chlorobenzoate dehalogenase from Pseudomonas sp. strain CBS-3*. Biochemistry, 1992. **31**(24): p. 5605-5610.
116. Dunaway-Mariano, D. and P. Babbitt, *On the origins and functions of the enzymes of the 4-chlorobenzoate to 4-hydroxybenzoate converting pathway*. Biodegradation, 1994. **5**(3-4): p. 259-276.
117. Zhuang, Z., et al., *Characterization of the 4-Hydroxybenzoyl-Coenzyme A Thioesterase from Arthrobacter sp. Strain SU*. Applied and Environmental Microbiology, 2003. **69**(5): p. 2707-2711.
118. Thoden, J.B., et al., *The structure of 4-hydroxybenzoyl-CoA thioesterase from arthrobacter sp. strain SU*. J Biol Chem, 2003. **278**(44): p. 43709-16.
119. Wang, F., et al., *Identification of a type III thioesterase reveals the function of an operon crucial for Mtb virulence*. Chem Biol, 2007. **14**(5): p. 543-51.
120. Astarie-Dequeker, C., et al., *Phthiocerol dimycocerosates of M. tuberculosis participate in macrophage invasion by inducing changes in the organization of plasma membrane lipids*. PLoS Pathog, 2009. **5**(2): p. e1000289.
121. Zhuang, Z., et al., *Divergence of function in the hot dog fold enzyme superfamily: the bacterial thioesterase YciA*. Biochemistry, 2008. **47**(9): p. 2789-96.
122. Willis, M.A., et al., *Structure of YciA from Haemophilus influenzae (HI0827), a hexameric broad specificity acyl-coenzyme A thioesterase*. Biochemistry, 2008. **47**(9): p. 2797-805.
123. Claxton, H.B., et al., *Structure and functional analysis of RifR, the type II thioesterase from the rifamycin biosynthetic pathway*. J Biol Chem, 2009. **284**(8): p. 5021-9.
124. Kozak, M., *Regulation of translation via mRNA structure in prokaryotes and eukaryotes*. Gene, 2005. **361**: p. 13-37.
125. Hawes, J.W. and E.T. Harper, *Synthesis of methacrylyl-CoA and (R)- and (S)-3-hydroxyisobutyryl-CoA*. Methods Enzymol, 2000. **324**: p. 73-9.
126. Riddles, P.W., R.L. Blakeley, and B. Zerner, *Ellman's reagent: 5,5' -dithiobis(2-nitrobenzoic acid) —a reexamination*. Analytical Biochemistry, 1979. **94**(1): p. 75-81.
127. Sykes, P.J., et al., *Molecular cloning of genes encoding branched-chain keto acid dehydrogenase of Pseudomonas putida*. J Bacteriol, 1987. **169**(4): p. 1619-25.

128. Mordaka, P.M., *Reductions using Clostridium sporogenes*, in *Department of Chemical and Environmental Engineering*. 2014, The University of Nottingham.
129. Song, F., et al., *The catalytic mechanism of the hotdog-fold enzyme superfamily 4-hydroxybenzoyl-CoA thioesterase from Arthrobacter sp. strain SU*. *Biochemistry*, 2012. **51**(35): p. 7000-16.
130. Haas, T., et al., *BIOTECHNOLOGICAL PRODUCTION OF 3-HYDROXYISOBUTYRIC ACID*. 2013, EVONIK DEGUSSA GMBH.
131. Thorpe, C. and J.J. Kim, *Structure and mechanism of action of the acyl-CoA dehydrogenases*. *The FASEB Journal*, 1995. **9**(9): p. 718-25.
132. Gregersen, N., P. Bross, and B.S. Andresen, *Genetic defects in fatty acid β -oxidation and acyl-CoA dehydrogenases*. *European Journal of Biochemistry*, 2004. **271**(3): p. 470-482.
133. Bartlett, K. and S. Eaton, *Mitochondrial β -oxidation*. *European Journal of Biochemistry*, 2004. **271**(3): p. 462-469.
134. Kim, J.-J.P. and K.P. Battaile, *Burning fat: the structural basis of fatty acid β -oxidation*. *Current Opinion in Structural Biology*, 2002. **12**(6): p. 721-728.
135. Toogood, H.S., et al., *Extensive Domain Motion and Electron Transfer in the Human Electron Transferring Flavoprotein-Medium Chain Acyl-CoA Dehydrogenase Complex*. *Journal of Biological Chemistry*, 2004. **279**(31): p. 32904-32912.
136. Crane, F.L. and H. Beinert, *ON THE MECHANISM OF DEHYDROGENATION OF FATTY ACYL DERIVATIVES OF COENZYME A: II. THE ELECTRON-TRANSFERRING FLAVOPROTEIN*. *Journal of Biological Chemistry*, 1956. **218**(2): p. 717-731.
137. Steenkamp, D.J. and M. Husain, *The effect of tetrahydrofolate on the reduction of electron transfer flavoprotein by sarcosine and dimethylglycine dehydrogenases*. *Biochemical Journal*, 1982. **203**(3): p. 707-715.
138. Leys, D., et al., *Extensive conformational sampling in a ternary electron transfer complex*. *Nat Struct Biol*, 2003. **10**(3): p. 219-25.
139. Gorelick, R.J. and C. Thorpe, *Electron-transferring flavoprotein from pig kidney: flavin analog studies*. *Biochemistry*, 1986. **25**(22): p. 7092-7098.
140. Beckmann, J.D. and F.E. Frerman, *Reaction of electron-transfer flavoprotein with electron transfer flavoprotein-ubiquinone oxidoreductase*. *Biochemistry*, 1985. **24**(15): p. 3922-3925.
141. Gorelick, R.J., et al., *Interflavin oxidation-reduction reactions between pig kidney general acyl-CoA dehydrogenase and electron-transferring flavoprotein*. *Biochemistry*, 1985. **24**(24): p. 6830-6839.
142. Ruzicka, F.J. and H. Beinert, *A new iron-sulfur flavoprotein of the respiratory chain. A component of the fatty acid beta oxidation pathway*. *Journal of Biological Chemistry*, 1977. **252**(23): p. 8440-5.
143. Frerman, F.E., *Reaction of electron-transfer flavoprotein ubiquinone oxidoreductase with the mitochondrial respiratory chain*. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1987. **893**(2): p. 161-169.
144. Cox, G.B., et al., *The function of ubiquinone in Escherichia coli*. *Biochem J*, 1970. **117**(3): p. 551-62.
145. Lenaz, G., *A critical appraisal of the mitochondrial coenzyme Q pool*. *FEBS Letters*. **509**(2): p. 151-155.
146. Zhang, J., F.E. Frerman, and J.J. Kim, *Structure of electron transfer flavoprotein-ubiquinone oxidoreductase and electron transfer to the mitochondrial ubiquinone pool*. *Proc Natl Acad Sci U S A*, 2006. **103**(44): p. 16212-7.
147. Beckmann, J.D. and F.E. Frerman, *Electron transfer flavoprotein-ubiquinone oxidoreductase from pig liver: purification and molecular, redox, and catalytic properties*. *Biochemistry*, 1985. **24**(15): p. 3913-3921.

148. Watmough, N.J. and F.E. Frerman, *The electron transfer flavoprotein: ubiquinone oxidoreductases*. Biochim Biophys Acta, 2010. **1797**(12): p. 1910-6.
149. Usselman, R.J., et al., *Impact of mutations on the midpoint potential of the [4Fe-4S]^{+1,+2} cluster and on catalytic activity in electron transfer flavoprotein-ubiquinone oxidoreductase (ETF-QO)*. Biochemistry, 2008. **47**(1): p. 92-100.
150. Toogood, H.S., D. Leys, and N.S. Scrutton, *Dynamics driving function: new insights from electron transferring flavoproteins and partner complexes*. FEBS J, 2007. **274**(21): p. 5481-504.
151. Page, C.C., C.C. Moser, and P.L. Dutton, *Mechanism for electron transfer within and between proteins*. Current Opinion in Chemical Biology, 2003. **7**(5): p. 551-556.
152. Leys, D., J. Basran, and N.S. Scrutton, *Channelling and formation of 'active' formaldehyde in dimethylglycine oxidase*. The EMBO Journal, 2003. **22**(16): p. 4038-4048.
153. Roberts, D.L., F.E. Frerman, and J.J. Kim, *Three-dimensional structure of human electron transfer flavoprotein to 2.1-Å resolution*. Proc Natl Acad Sci U S A, 1996. **93**(25): p. 14355-60.
154. Steenkamp, D.J., *Cross-linking of the electron-transfer flavoprotein to electron-transfer flavoprotein-ubiquinone oxidoreductase with heterobifunctional reagents*. Biochemical Journal, 1988. **255**(3): p. 869-876.
155. Leys, D. and N.S. Scrutton, *Electrical circuitry in biology: emerging principles from protein structure*. Curr Opin Struct Biol, 2004. **14**(6): p. 642-7.
156. Page, C.C., et al., *Natural engineering principles of electron tunnelling in biological oxidation-reduction*. Nature, 1999. **402**(6757): p. 47-52.
157. Talfourier, F., et al., *α Arg-237 in Methylophilus methylotrophus (sp. W3A1) Electron-transferring Flavoprotein Affords ~200-Millivolt Stabilization of the FAD Anionic Semiquinone and a Kinetic Block on Full Reduction to the Dihydroquinone*. Journal of Biological Chemistry, 2001. **276**(23): p. 20190-20196.
158. Chohan, K.K., et al., *Protein dynamics enhance electronic coupling in electron transfer complexes*. J Biol Chem, 2001. **276**(36): p. 34142-7.
159. Jones, M., et al., *Electron Transfer and Conformational Change in Complexes of Trimethylamine Dehydrogenase and Electron Transferring Flavoprotein*. Journal of Biological Chemistry, 2002. **277**(10): p. 8457-8465.
160. Mattevi, A., *To be or not to be an oxidase: challenging the oxygen reactivity of flavoenzymes*. Trends in Biochemical Sciences, 2006. **31**(5): p. 276-283.
161. Stephens, P.J., D.R. Jollie, and A. Warshel, *Protein Control of Redox Potentials of Iron-Sulfur Proteins*. Chemical Reviews, 1996. **96**(7): p. 2491-2514.
162. Yin, Y., et al., *The Presence of a Hydrogen Bond between Asparagine 485 and the π System of FAD Modulates the Redox Potential in the Reaction Catalyzed by Cholesterol Oxidase^{†,‡}*. Biochemistry, 2001. **40**(46): p. 13779-13787.
163. Parker, A., *A single arginine residue is required for the interaction of the electron transferring flavoprotein (ETF) with three of its dehydrogenase partners*. Molecular and Cellular Biochemistry, 2003. **254**(1-2): p. 91-100.
164. Watmough, N.J., J. Kiss, and F.E. Frerman, *Structural and redox relationships between Paracoccus denitrificans, porcine and human electron-transferring flavoproteins*. European Journal of Biochemistry, 1992. **205**(3): p. 1089-1097.
165. Husain, M. and D.J. Steenkamp, *Partial purification and characterization of glutaryl-coenzyme A dehydrogenase, electron transfer flavoprotein, and*

- electron transfer flavoprotein-Q oxidoreductase from *Paracoccus denitrificans*. *Journal of Bacteriology*, 1985. **163**(2): p. 709-15.
166. Campbell, J.W. and J.E. Cronan Jr, *The Enigmatic Escherichia coli fadE Gene Is yafH*. *Journal of Bacteriology*, 2002. **184**(13): p. 3759-3764.
 167. Hiltunen, J.K. and Y.-M. Qin, *β -Oxidation – strategies for the metabolism of a wide variety of acyl-CoA esters*. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 2000. **1484**(2-3): p. 117-128.
 168. Poirier, Y., et al., *Peroxisomal beta-oxidation--a metabolic pathway with multiple functions*. *Biochim Biophys Acta*, 2006. **1763**(12): p. 1413-26.
 169. Hutson, S.M., A.J. Sweatt, and K.F. LaNoue, *Branched-Chain Amino Acid Metabolism: Implications for Establishing Safe Intakes*. *The Journal of Nutrition*, 2005. **135**(6): p. 1557S-1564S.
 170. Harper, A.E., R.H. Miller, and K.P. Block, *Branched-chain amino acid metabolism*. *Annu Rev Nutr*, 1984. **4**: p. 409-54.
 171. Binder, S., *Branched-Chain Amino Acid Metabolism in Arabidopsis thaliana*. *Arabidopsis Book*, 2010. **8**: p. e0137.
 172. Imlay, J.A. and S. Linn, *Mutagenesis and stress responses induced in Escherichia coli by hydrogen peroxide*. *Journal of Bacteriology*, 1987. **169**(7): p. 2967-2976.
 173. Schellhorn, H.E., *Regulation of hydroperoxidase (catalase) expression in Escherichia coli*. *FEMS Microbiology Letters*, 1995. **131**(2): p. 113-119.
 174. Däschner, K., I. Couée, and S. Binder, *The Mitochondrial Isovaleryl-Coenzyme A Dehydrogenase of Arabidopsis Oxidizes Intermediates of Leucine and Valine Catabolism*. *Plant Physiology*, 2001. **126**(2): p. 601-612.
 175. Faivre-Nitschke, S.E., et al., *Purification, characterization and cloning of isovaleryl-CoA dehydrogenase from higher plant mitochondria*. *Eur J Biochem*, 2001. **268**(5): p. 1332-9.
 176. Ikeda, Y. and K. Tanaka, *Purification and characterization of 2-methyl-branched chain acyl coenzyme A dehydrogenase, an enzyme involved in the isoleucine and valine metabolism, from rat liver mitochondria*. *J Biol Chem*, 1983. **258**(15): p. 9477-87.
 177. Finocchiaro, G., et al., *Molecular cloning and nucleotide sequence of cDNAs encoding the alpha-subunit of human electron transfer flavoprotein*. *Journal of Biological Chemistry*, 1988. **263**(30): p. 15773-80.
 178. Herrick, K.R., et al., *Expression and characterization of human and chimeric human-Paracoccus denitrificans electron transfer flavoproteins*. *J Biol Chem*, 1994. **269**(51): p. 32239-45.
 179. Finocchiaro, G., et al., *cDNA cloning and mitochondrial import of the β -subunit of the human electron-transfer flavoprotein*. *European Journal of Biochemistry*, 1993. **213**(3): p. 1003-1008.
 180. Goodman, S.I., et al., *Molecular cloning and expression of a cDNA encoding human electron transfer flavoprotein-ubiquinone oxidoreductase*. *Eur J Biochem*, 1994. **219**(1-2): p. 277-86.
 181. Ishizaki, K., et al., *The mitochondrial electron transfer flavoprotein complex is essential for survival of Arabidopsis in extended darkness*. *Plant J*, 2006. **47**(5): p. 751-60.
 182. Ishizaki, K., et al., *The critical role of Arabidopsis electron-transfer flavoprotein:ubiquinone oxidoreductase during dark-induced starvation*. *Plant Cell*, 2005. **17**(9): p. 2587-600.
 183. Shinzawa, K., et al., *Molecular cloning of a cDNA for α -subunit of rat liver electron transfer flavoprotein*. *Biochemical and Biophysical Research Communications*, 1988. **155**(1): p. 300-304.
 184. *The Status, Quality, and Expansion of the NIH Full-Length cDNA Project: The Mammalian Gene Collection (MGC)*. *Genome Research*, 2004. **14**(10b): p. 2121-2127.

185. Huang, S., *cDNA cloning, functional expression and cellular localization of rat liver mitochondrial electron-transfer flavoprotein-ubiquinone oxidoreductase protein*. Science in China Series C, 2005. **48**(4): p. 357.
186. Vockley, J., et al., [24] - *Mammalian Branched-Chain Acyl-CoA Dehydrogenases: Molecular Cloning and Characterization of Recombinant Enzymes*, in *Methods in Enzymology*, A.H. Robert and R.S. John, Editors. 2000, Academic Press. p. 241-258.
187. Graham, I.A. and P.J. Eastmond, *Pathways of straight and branched chain fatty acid catabolism in higher plants*. Prog Lipid Res, 2002. **41**(2): p. 156-81.
188. Hayashi, H., *A Novel Acyl-CoA Oxidase That Can Oxidize Short-chain Acyl-CoA in Plant Peroxisomes*. Journal of Biological Chemistry, 1999. **274**(18): p. 12715-12721.
189. Pedersen, L. and A. Henriksen, *Expression, purification and crystallization of two peroxisomal acyl-CoA oxidases from Arabidopsis thaliana*. Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 6): p. 1125-8.
190. Mackenzie, J., et al., *Controlling electron transfer in Acyl-CoA oxidases and dehydrogenases: a structural view*. J Biol Chem, 2006. **281**(41): p. 31012-20.
191. Valentine, R.C. and R.S. Wolfe, *Purification and Role of Phosphotransbutyrylase*. Journal of Biological Chemistry, 1960. **235**(7): p. 1948-1952.
192. Hashimoto, Y., et al., *Nitrile pathway involving acyl-CoA synthetase: overall metabolic gene organization and purification and characterization of the enzyme*. J Biol Chem, 2005. **280**(10): p. 8660-7.
193. Takao, S., T. Ito, and M. Tanida, *Purification and Kinetic-Properties of Butyryl-CoA Synthetase from Paecilomyces-Varioti*. Agricultural and Biological Chemistry, 1987. **51**(1): p. 145-152.
194. Webster, L.T., Jr., L.D. Gerowin, and L. Rakita, *Purification and Characteristics of a Butyryl Coenzyme a Synthetase from Bovine Heart Mitochondria*. J Biol Chem, 1965. **240**(1): p. 29-33.
195. Harwood, C.S. and E. Canale-Parola, *Properties of acetate kinase isozymes and a branched-chain fatty acid kinase from a spirochete*. J Bacteriol, 1982. **152**(1): p. 246-54.
196. Twarog, R. and R.S. Wolfe, *Enzymatic phosphorylation of butyrate*. J Biol Chem, 1962. **237**: p. 2474-7.
197. Wiesenborn, D.P., F.B. Rudolph, and E.T. Papoutsakis, *Phosphotransbutyrylase from Clostridium acetobutylicum ATCC 824 and its role in acidogenesis*. Appl Environ Microbiol, 1989. **55**(2): p. 317-22.
198. Sambrook, J. and D. Russel, *Molecular Cloning, A laboratory Manual*. 3rd ed. 2001: Cold Spring Harbor Laboratory Press.
199. Rosano, G. and E. Ceccarelli, *Rare codon content affects the solubility of recombinant proteins in a codon bias-adjusted Escherichia coli strain*. Microbial Cell Factories, 2009. **8**(1): p. 41.
200. Cary, J.W., et al., *Cloning and expression of Clostridium acetobutylicum phosphotransbutyrylase and butyrate kinase genes in Escherichia coli*. J Bacteriol, 1988. **170**(10): p. 4613-8.
201. Mooney, B.P., J.A. Miernyk, and D.D. Randall, *The complex fate of alpha-ketoacids*. Annu Rev Plant Biol, 2002. **53**: p. 357-75.
202. Miles, E.W., S. Rhee, and D.R. Davies, *The Molecular Basis of Substrate Channeling*. Journal of Biological Chemistry, 1999. **274**(18): p. 12193-12196.
203. Zhang, Y.H.P., *Substrate channeling and enzyme complexes for biotechnological applications*. Biotechnology Advances, 2011. **29**(6): p. 715-725.

204. Odessey, R., *Direct evidence for the inactivation of branched-chain oxo-acid dehydrogenase by enzyme phosphorylation*. FEBS Letters, 1980. **121**(2): p. 306-308.
205. Odessey, R., *Purification of rat kidney branched-chain oxo acid dehydrogenase complex with endogenous kinase activity*. Biochemical Journal, 1982. **204**(1): p. 353-356.
206. Paxton, R. and R.A. Harris, *Isolation of rabbit liver branched chain alpha-ketoacid dehydrogenase and regulation by phosphorylation*. J Biol Chem, 1982. **257**(23): p. 14433-9.
207. Danner, D.J., et al., *Purification and characterization of branched chain alpha-ketoacid dehydrogenase from bovine liver mitochondria*. Journal of Biological Chemistry, 1979. **254**(12): p. 5522-6.
208. Pettit, F.H., S.J. Yeaman, and L.J. Reed, *Purification and characterization of branched chain alpha-keto acid dehydrogenase complex of bovine kidney*. Proc Natl Acad Sci U S A, 1978. **75**(10): p. 4881-5.
209. Taylor, N.L., et al., *Lipoic acid-dependent oxidative catabolism of alpha-keto acids in mitochondria provides evidence for branched-chain amino acid catabolism in Arabidopsis*. Plant Physiol, 2004. **134**(2): p. 838-48.
210. Roberts, C.M. and J.R. Sokatch, *Branched-Chain Amino-Acids as Activators of Branched-Chain Ketoacid Dehydrogenase*. Biochemical and Biophysical Research Communications, 1978. **82**(3): p. 828-833.
211. Sokatch, J.R., V. McCully, and C.M. Roberts, *Purification of a branched-chain keto acid dehydrogenase from Pseudomonas putida*. J Bacteriol, 1981. **148**(2): p. 647-52.
212. Sykes, P.J., et al., *Conjugative mapping of pyruvate, 2-ketoglutarate, and branched-chain keto acid dehydrogenase genes in Pseudomonas putida mutants*. Journal of Bacteriology, 1985. **162**(1): p. 203-208.
213. McCully, V., G. Burns, and J.R. Sokatch, *Resolution of branched-chain oxo acid dehydrogenase complex of Pseudomonas aeruginosa PAO*. Biochem J, 1986. **233**(3): p. 737-42.
214. Namba, Y., et al., *Coenzyme A- and Nicotinamide Adenine Dinucleotide-dependent Branched Chain α -Keto Acid Dehydrogenase: I. PURIFICATION AND PROPERTIES OF THE ENZYME FROM BACILLUS SUBTILIS*. Journal of Biological Chemistry, 1969. **244**(16): p. 4437-4447.
215. Willecke, K. and A.B. Pardee, *Fatty Acid-requiring Mutant of Bacillus subtilis Defective in Branched Chain α -Keto Acid Dehydrogenase*. Journal of Biological Chemistry, 1971. **246**(17): p. 5264-5272.
216. Debarbouille, M., et al., *Role of bkdR, a transcriptional activator of the sigL-dependent isoleucine and valine degradation pathway in Bacillus subtilis*. J Bacteriol, 1999. **181**(7): p. 2059-66.
217. Wang, G.-F., et al., *The primary structure of branched-chain alpha-oxo acid dehydrogenase from Bacillus subtilis and its similarity to other alpha-oxo acid dehydrogenases*. European Journal of Biochemistry, 1993. **213**(3): p. 1091-1099.
218. SASAKI, S., *On the Decarboxylase Operating in the Degradative Pathway of L-Leucine by Proteus vulgaris*. Journal of Biochemistry, 1962. **51**(5): p. 335-344.
219. Denoya, C.D., et al., *A second branched-chain alpha-keto acid dehydrogenase gene cluster (bkdFGH) from Streptomyces avermitilis: its relationship to avermectin biosynthesis and the construction of a bkdF mutant suitable for the production of novel antiparasitic avermectins*. J Bacteriol, 1995. **177**(12): p. 3504-11.
220. Skinner, D.D., et al., *Cloning and sequencing of a cluster of genes encoding branched-chain alpha-keto acid dehydrogenase from Streptomyces avermitilis and the production of a functional E1 [alpha beta] component in Escherichia coli*. Journal of Bacteriology, 1995. **177**(1): p. 183-190.

221. Hafner, E.W., et al., *Branched-chain fatty acid requirement for avermectin production by a mutant of Streptomyces avermitilis lacking branched-chain 2-oxo acid dehydrogenase activity*. The Journal of Antibiotics, 1991. **44**(3): p. 349-356.
222. Rüdiger, H.W., U. Langenbeck, and H.W. Goedde, *Oxidation of Branched Chain α -Ketoacids in Streptococcus faecalis and its Dependence on Lipoic Acid*. Hoppe-Seyler's Zeitschrift für physiologische Chemie, 1972. **353**(1): p. 875-882.
223. Ward, D.E., et al., *Catabolism of Branched-Chain α -Keto Acids in Enterococcus faecalis: the bkd Gene Cluster, Enzymes, and Metabolic Route*. Journal of Bacteriology, 1999. **181**(17): p. 5433-5442.
224. Max Wynn, R., et al., [18] - *Expression of E1 Component of Human Branched-Chain α -Keto Acid Dehydrogenase Complex in Escherichia coli by Cotransformation with Chaperonins GroEL GroES*, in *Methods in Enzymology*, A.H. Robert and R.S. John, Editors. 2000, Academic Press. p. 179-191.
225. Davie, J.R., et al., *Expression and assembly of a functional E1 component (α 2 β 2) of mammalian branched-chain α -ketoacid dehydrogenase complex in Escherichia coli*. Journal of Biological Chemistry, 1992. **267**(23): p. 16601-6.
226. Fujiki, Y., *Isolation and Characterization of cDNA Clones for the E1 β and E2 Subunits of the Branched-chain α -Ketoacid Dehydrogenase Complex in Arabidopsis*. Journal of Biological Chemistry, 2000. **275**(8): p. 6007-6013.
227. *The Electronic Plant Gene Register*. Plant Physiology, 1998. **118**(2): p. 711-711.
228. Mooney, B.P., et al., *The dihydrolipoyl acyltransferase (BCE2) subunit of the plant branched-chain α -ketoacid dehydrogenase complex forms a 24-mer core with octagonal symmetry*. Protein Sci, 2000. **9**(7): p. 1334-9.
229. Sokatch, J.R., L.E. Sanders, and V.P. Marshall, *Oxidation of Methylmalonate Semialdehyde to Propionyl Coenzyme A in Pseudomonas aeruginosa Grown on Valine*. Journal of Biological Chemistry, 1968. **243**(10): p. 2500-2506.
230. Bannerjee, D., L.E. Sanders, and J.R. Sokatch, *Properties of Purified Methylmalonate Semialdehyde Dehydrogenase of Pseudomonas aeruginosa*. Journal of Biological Chemistry, 1970. **245**(7): p. 1828-1835.
231. Ikeda, H. and S. Ōmura, *Chapter 5 - Biosynthesis, Regulation, and Genetics of Macrolide Production*, in *Macrolide Antibiotics (Second Edition)*, S. Ōmura, Editor. 2003, Academic Press: San Diego. p. 285-326.
232. Kaneda, T., *Fatty Acids in the Genus Bacillus. I. Iso- and Anteiso-Fatty Acids as Characteristic Constituents of Lipids in 10 Species*. Journal of Bacteriology, 1967. **93**(3): p. 894-903.
233. Kaneda, T., *Fatty acids of the genus Bacillus: an example of branched-chain preference*. Bacteriological Reviews, 1977. **41**(2): p. 391-418.
234. Kaneda, T., *Iso- and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic significance*. Microbiological Reviews, 1991. **55**(2): p. 288-302.
235. Butterworth, P.H.W. and K. Bloch, *Comparative Aspects of Fatty Acid Synthesis in Bacillus subtilis and Escherichia coli*. European Journal of Biochemistry, 1970. **12**(3): p. 496-501.
236. Kaneda, T., *Biosynthesis of Branched Chain Fatty Acids: II. MICROBIAL SYNTHESIS OF BRANCHED LONG CHAIN FATTY ACIDS FROM CERTAIN SHORT CHAIN FATTY ACID SUBSTRATES*. Journal of Biological Chemistry, 1963. **238**(4): p. 1229-1235.
237. Hemilä, H., et al., *Secretory S complex of Bacillus subtilis: sequence analysis and identity to pyruvate dehydrogenase*. Journal of Bacteriology, 1990. **172**(9): p. 5052-5063.

238. Hodgson, J.A., P.N. Lowe, and R.N. Perham, *Wild-type and mutant forms of the pyruvate dehydrogenase multienzyme complex from Bacillus subtilis*. Biochemical Journal, 1983. **211**(2): p. 463-472.
239. Lowe, P.N., J.A. Hodgson, and R.N. Perham, *Dual role of a single multienzyme complex in the oxidative decarboxylation of pyruvate and branched-chain 2-oxo acids in Bacillus subtilis*. Biochemical Journal, 1983. **215**(1): p. 133-140.
240. Howard, T.P., et al., *Synthesis of customized petroleum-replica fuel molecules by targeted modification of free fatty acid pools in Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(19): p. 7636-7641.
241. Gasteiger, E., et al., *ExPASy: the proteomics server for in-depth protein knowledge and analysis*. Nucleic Acids Research, 2003. **31**(13): p. 3784-3788.
242. Lang, K., et al., *Metabolic engineering of Pseudomonas sp. strain VLB120 as platform biocatalyst for the production of isobutyric acid and other secondary metabolites*. Microbial Cell Factories, 2014. **13**(1): p. 2.
243. Barak, Z., D.M. Chipman, and N. Gollop, *Physiological implications of the specificity of acetohydroxy acid synthase isozymes of enteric bacteria*. Journal of Bacteriology, 1987. **169**(8): p. 3750-3756.
244. Lawther, R.P., et al., *DNA sequence fine-structure analysis of ilvG (IlvG+) mutations of Escherichia coli K-12*. Journal of Bacteriology, 1982. **149**(1): p. 294-298.
245. Gollop, N., et al., *Physiological implications of the substrate specificities of acetohydroxy acid synthases from varied organisms*. Journal of Bacteriology, 1990. **172**(6): p. 3444-3449.
246. Park, J.H., et al., *Metabolic engineering of Escherichia coli for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation*. Proceedings of the National Academy of Sciences, 2007. **104**(19): p. 7797-7802.
247. Park, J.H., et al., *Escherichia coli W as a new platform strain for the enhanced production of L-valine by systems metabolic engineering*. Biotechnol Bioeng, 2011. **108**(5): p. 1140-7.
248. Savrasova, E.A., et al., *Use of the valine biosynthetic pathway to convert glucose into isobutanol*. J Ind Microbiol Biotechnol, 2011. **38**(9): p. 1287-94.
249. Blombach, B. and B.J. Eikmanns, *Current knowledge on isobutanol production with Escherichia coli, Bacillus subtilis and Corynebacterium glutamicum*. Bioeng Bugs, 2011. **2**(6): p. 346-50.
250. Blombach, B., et al., *Corynebacterium glutamicum tailored for efficient isobutanol production*. Appl Environ Microbiol, 2011. **77**(10): p. 3300-10.
251. Nishio, Y., et al., *Evolutionary process of amino acid biosynthesis in Corynebacterium at the whole genome level*. Mol Biol Evol, 2004. **21**(9): p. 1683-91.
252. Hasegawa, S., et al., *Improvement of the redox balance increases L-valine production by Corynebacterium glutamicum under oxygen deprivation conditions*. Appl Environ Microbiol, 2012. **78**(3): p. 865-75.
253. Radmacher, E., et al., *Linking Central Metabolism with Increased Pathway Flux: L-Valine Accumulation by Corynebacterium glutamicum*. Applied and Environmental Microbiology, 2002. **68**(5): p. 2246-2250.
254. Ruklisha, M., L. Paegle, and I. Denina, *L-Valine biosynthesis during batch and fed-batch cultivations of Corynebacterium glutamicum: Relationship between changes in bacterial growth rate and intracellular metabolism*. Process Biochemistry, 2007. **42**(4): p. 634-640.

255. Krause, F.S., B. Blombach, and B.J. Eikmanns, *Metabolic engineering of Corynebacterium glutamicum for 2-ketoisovalerate production*. Appl Environ Microbiol, 2010. **76**(24): p. 8053-61.
256. Holtzclaw, W.D. and L.F. Chapman, *Degradative acetolactate synthase of Bacillus subtilis: purification and properties*. J Bacteriol, 1975. **121**(3): p. 917-22.
257. Juni, E., *MECHANISMS OF FORMATION OF ACETOIN BY BACTERIA*. Journal of Biological Chemistry, 1952. **195**(2): p. 715-726.
258. Halpern, Y.S. and H.E. Umbarger, *Evidence for Two Distinct Enzyme Systems Forming Acetolactate in Aerobacter aerogenes*. Journal of Biological Chemistry, 1959. **234**(12): p. 3067-3071.
259. GenScript Codon Frequency Table. Available from: http://www.genscript.com/cgi-bin/tools/codon_freq_table.
260. Essen, L.R., et al., *Recombinant cell producing 2-hydroxyisobutyric acid*. 2011, Evonik Roehm GMBH.
261. XU, J., et al., *Microorganisms and methods for producing acrylate and other products from propionyl-CoA*. 2014, The Procter & Gamble Company.
262. Brown, G.K., et al., *β -Hydroxyisobutyryl Coenzyme A Deacylase Deficiency: A Defect in Valine Metabolism Associated with Physical Malformations*. Pediatrics, 1982. **70**(4): p. 532-538.
263. Guzman, L.M., et al., *Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter*. Journal of Bacteriology, 1995. **177**(14): p. 4121-4130.
264. Cho, H. and J.E. Cronan, Jr., *Escherichia coli thioesterase I, molecular cloning and sequencing of the structural gene and identification as a periplasmic enzyme*. J Biol Chem, 1993. **268**(13): p. 9238-45.
265. Naggert, J., et al., *Cloning, sequencing, and characterization of Escherichia coli thioesterase II*. J Biol Chem, 1991. **266**(17): p. 11044-50.
266. Baba, T., et al., *Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection*. Mol Syst Biol, 2006. **2**: p. 2006 0008.
267. Jones, C.E., et al., *Cloning and sequencing of the Escherichia coli panB gene, which encodes ketopantoate hydroxymethyltransferase, and overexpression of the enzyme*. Journal of Bacteriology, 1993. **175**(7): p. 2125-2130.
268. Gardner, T.S., C.R. Cantor, and J.J. Collins, *Construction of a genetic toggle switch in Escherichia coli*. Nature, 2000. **403**(6767): p. 339-342.
269. Blombach, B., et al., *L-Valine Production with Pyruvate Dehydrogenase Complex-Deficient Corynebacterium glutamicum*. Applied and Environmental Microbiology, 2007. **73**(7): p. 2079-2084.

8 Appendices

8.1 Appendix 1 – polypeptide sequences

Polypeptide sequence 1: 4-hydroxybenzoyl-CoA thioesterase from *Arthrobacter* sp. strain SU

MHRTSNGSHATGGNLPDVASHYPVAYEQTLTGTVGFVIDEMTPERATASVEVTDTLRQRWGLVHGGAYC
ALAEMLATEATVAVVHEKGMMAVGQSNHTSFFRPVKEGHVRAEAVRIHAGSTTWFDVSLRDDAGRLCA
VSSMSIAVRPRRD

Polypeptide sequence 2: Fatty Acyl-CoA thioesterase from *Mycobacterium tuberculosis*

MSHTDLTPCTRVLASSGTVPIAEELLARVLEPYSCCKGCRYLIDAQYSATEDSVLAYGNFTIGESAYIRS
TGHFNAVELILCFNQLAYSAPAVLNEEIRVLRGWSIDDYCQHQQLSSMLIRKASSRFRKPLNPQKFS
RLLCRDLQVIERTWRYLKVPCVIEFWDENGGAASGEIELAALNIP

Polypeptide sequence 3: YciA from *Haemophilus influenzae*

MSANFTDKNGRQSKGVLLRLTAMPSTNANGDIFGGWIMSQMDMGGAILAKEIAHGRVTVAVESMNF
IKPISVGDVVCYGCQCLKVGRSSIKIKVEVWVKVASEPIGERYCVTDAVFTFVAVDNNGRSRTIPREN
NQELEKALALISEQPL

Polypeptide sequence 4: RifR from *Amycolatopsis mediterranea*

MHRPEAEKWLRRFERAPDARARLVCLPHAGGSASFFFPLAKALAPAVEVLAVQYPGRQDRRHEPPVDSI
GGLTNRLLEVLRFPGDRPLALFGHSMGAIIGYELALRMPEAGLPAPVHLFASGRRAPSRYRDDDVRGAS
DERLVAELRKLGSDAAMLADPELLAMVLPAIRSDYRAVETRYRHEPGRRVDCPVTVFTGDHDPVSVGE
ARAWEEHTTGPADLRVLPGGHFFLVDQAAPMIATMTEKLAGPALTGSTGNS

Polypeptide sequence 5: N-terminal His-tagged 4HBT

MGSSHHHHHSSGLVPRGSHMRTSNGSHATGGNLPDVASHYPVAYEQTLTGTVGFVIDEMTPERATAS
VEVTDTLRQRWGLVHGGAYCALAEMLATEATVAVVHEKGMMAVGQSNHTSFFRPVKEGHVRAEAVRIHA
GSTTWFDVSLRDDAGRLCAVSSMSIAVRPRRD

Polypeptide sequence 6: C-terminal His-tagged 4HBT

MHRTSNGSHATGGNLPDVASHYPVAYEQTLTGTVGFVIDEMTPERATASVEVTDTLRQRWGLVHGGAYC
ALAEMLATEATVAVVHEKGMMAVGQSNHTSFFRPVKEGHVRAEAVRIHAGSTTWFDVSLRDDAGRLCA
VSSMSIAVRPRRDGLEHHHHHH

Polypeptide sequence 7: Short chain acyl-CoA oxidase (ACX4)

MAVLSSADRASNEKKVKSSYFDLPPMEMSVAFPQATPASTFPPCTSDYYHFNDLLTPEEQAIRKKVREC
MEKEVAPIMTEYWEKAEFFPHITPKLGAMGVAGGSIKGYGCPGLSITANAIATAEIRVDASCSTFILV
HSSLGMLTIALCGSEAQKEKYLPSLAQLNTVACWALTEPDNGSDASGLGTTATKVEGGWKINGQKRWIG
NSTFADLLIIFARNTTTNQINGFIVKKDAPGLKATKIPNKIGLRMVQNGDILLQNVFVPDEDRLPGVNS
FQDTSKVLAVSRVMVAWQPIGISMGYDMCHRYLKERKQFGAPLAAFQLNQKLVQMLGNVQAMFLMGW
RLCKLYETGQMTPGQASLGKAWISSKARETASLGRELLGGNGILADFLVAKAFCDLEPIYTYEGTYDIN
TLVTGREVTGIASFKPATRSRL

Polypeptide sequence 8: AcsA from *Pseudomonas chlororaphis* B23

MRDYEHVVESFDYLQSATQDLHGELTALNACVECCDRHAHGEAVALYCEAQDGHAEYRFRDLQRQAAR
FGNFLREQGVKPGDRVAGLMPRTVELLIAILGTWRIGAVYQPLFTAFGPKAIEQRLNCSNARWIVTDPH
NRPKLDDVTDCPSIVVTGGAPQNPADHHFWSALNRQADDCAVLLDASAPFLMCTSGTTGPAKPLEVP
LSAILAFKGYMRDAIDLRADDRFWNLADPGWAYGLYYAVTGPLACGYATLFYDGPFTVESTRHIIAKYA
INNLAGSPTAYRFLIAAGAEFADAVRGRLRAVSSAGEPLNPQVVRWFAEQLGVVIHDHYGQTEIGMVL
NHHGLRHPVREGSAGYAVPGYRIVVLDKAHRELPAQPGVLAVDRERSPLCWFDGYLGMPQTQAFAGRYY
LSGDIVELNDDGSISFVGRNDDLITTSgyRVGPFdVESALIEHPAVVEAAVIGKPDQPRTelikAFVVL
NTPYLPSPELAEELRLHVRQRLAAHAYPREMEFVDHLPKTPSGKLQRFILRNQEIAKQQALG

Polypeptide sequence 9: BCKD E1 α from *B. subtilis* 168s

MSTNRHQALGLTDQEAVDMYRTMLLARKIDERMWLLDRSGKIPFVISCQGQEEAAQVGAAAFALDREMDYV
LPYYRDMGVVLAFGMTAKDLMMSGFAKAADPNSSGGRQMPGHFGQKKNRIVTGSSPVTQVPHAVGIALA
GRMEKKDIAAFVTFEGSSNQDGFHEGANFAAVHKLPIVFCENNKYAISVPYDKQVACENISDRAIGY
GMPGVTVNGNDPLEVYQAVKEARERARRGEGPTLIETISYRLTPHSSDDDDSSYRGREEVEEAKKSDPL
LTYQAYLKETGLLSDEIEQTMLDEIMAIVNEATDEAENAPYAAPESALDYVYAK

Polypeptide sequence 10: BCKD E1 β from *B. subtilis* 168s

MSVMSYIDAINLAMKEEMERDSRVFVLGEDVGRKGGVFKATAGLYEQFGEERVMDTPLAESAIAGVGIG
AAMYGMRPIAEMQFADFIMPAVNQII SEAAKIRYRSNNDWSCPVVVRAPYEEVHGALYHSQSVEAIFA
NQPGLKIVMPSTPYDAKGLLKA AVRDEDPVLFFEHKRAYRLIKGEVPADDYVLP I GKADVKREGDDITV
ITYGLCVHFALQAAERLEKDGISAHVVDLRTVYPLDKEAIEAASKTGKVLVLTEDTKEGSIMSEVAAI
ISEHCLFDLDAPIKRLAGPDIPAMPYAPTMEKYFMVNPDKVEAAMRELAEF

Polypeptide sequence 11: BCKD E2 from *B. subtilis* 168s

MAIEQMTMPQLGESVTEGTISKWLVPAGDKVNKYDPVIAEVMTDKVNAEVPSSFTGTITELVGEEGQTLQ
VGEMICKIETEGANPAEQKQEQPAASEAAENPVAKSAGAADQPNKKRYSPAVLRLAGEHGIDLDQVTGT
GAGGRITRKDIQRLIETGGVQEQNPPEELKTAAPAPKSASKPEPKEETSYPASAAGDKEIPVTGVRKAIA
SNMKRSKTEIPHAWTMMEVDVTNMVAYRNSIKDSFKKTEGFNLTFFAFFVKAVAQALKEFPQMNSMWAG
DKIIQKKDINISIAVATEDSLFVPVIKNADEKTIKGI AKDITGLAKKVRDGKLTADDMQGGTFTVNNTG
SFGSVQSMGIINYPQAAILQVESIVKRPPVMDNGMIAVRDMVNLCLSLDHRVLDGLVCGRFLGRVKQIIL
ESIDEKTSVY

8.2 Appendix 2 – polynucleotide sequences

Polynucleotide sequence 1: 4HBT

TATACATATGCACCGTACCTCTAACGGTTCTCACGCGACCGGTGGTAACCTGCCGGATGTTGCTTCTCA
TTACCCGGTTGCGTACGAACAGACCCCTGGACGGCACCGTGGGTTTCGTTATCGACGAAATGACCCCGGA
ACGTGCGACCGCGTCTGTTGAAGTTACCGACACCCCTGCGTCAGCGTTGGGGTCTGGTTCACGGTGGTGC
TTACTGCGCACTGGCGGAGATGCTGGCGACCGAAGCGACCGTTGCGGTTGTGCACGAAAAGGGTATGAT
GGCGGTTGGCCAGTCTAACCACACCTCTTTCTTCCGTCCGGTTAAAGAAGGTCACGTTGCGCGGGAAGC
TGTTTCGCATCCACGCGGGTTCTACCACCTGGTTCTGGGACGTTTCCCTGCGCGATGATGCGGGTCGCCT
GTGTGCGGTGTCTTCTATGTCTATCGCGGTGCGTCCGCGTCTGTGACTAAGCGGCCGCACTC

Polynucleotide sequence 2: FcoT

TATACATATGTCTCACACCGACCTGACCCCGTGCACTCGTGTGCTGGCTTCCTCTGGTACCGTTCCGAT
CGCGGAAGAACTGCTGGCGCGTGTCTGGAACCGTATTCTCGCAAAGGTTGCCGTTATCTGATCGACGC
TCAGTACAGCGCGACCGAAGATAGCGTTCTGGCGTACGGTAACCTCACTATCGGTGAATCTGCGTACAT
CCGTTCCACCGGTCACCTCAACGCGGTTGAACTGATTCTGTGCTTCAACCAGCTGGCATACTCTGCATT
CGCACCGGCGGTTCTGAACGAAGAAATTCGCGTGCTGCGTGGTTGGTCTATCGACGACTACTGTGAGCA
TCAGCTGTCTCTATGCTGATCCGTAAAGCGTCTTCTCGTTTTCCGCAAGCCGCTGAACCCGCAGAAGTT
TTCTGCGCGCCTGCTGTGCCGTGATCTGCAGGTTATTGAACGCACTTGGCGCTACCTGAAAGTTCCGTG
CGTTATCGAGTTCTGGGACGAAAACGGTGGTGACGCTCTGGTGAAATCGAACTGGCGGCTCTGAACAT
CCCGTAAGCGGCCGCACTC

Polynucleotide sequence 3: YciA

TATACATATGAGCGCAAACCTTACCGACAAAAACGGTCGTCAGTCCAAAGGTGTTCTGCTGCTGCGTAC
CCTGGCAATGCCGTCTGACACCAACGCTAACGGTGACATCTTCGGTGGTTGGATCATGTCTCAGATGGA
CATGGGTGGTGCGATCCTGGCGAAAGAAATCGCGCACGGTCGCGTTGTTACCGTGGCGGTGGAGTCTAT
GAACTTCATCAAGCCGATCTCTGTGGGTGACGTGGTTTTGCTGCTACGGTCAGTGCCCTGAAAGTTGGTCG
CTCCAGCATCAAGATCAAAGTTGAGGTGTGGGTTAAGAAGGTTGCGTCTGAACCGATTGGCGAGCGTTA
CTGCGTTACCGACGCGGTGTTTACCTTCGTTGCGGTTGACAACAACGGCCGAGCCGTACCATCCCGCG
TGAAAACAACCGGAACCTGGA AAAAGCGCTGGCGCTGATTTCTGAACAGCCGCTGTAAGCGGCCGCACTC
C

Polynucleotide sequence 4: RifR

TATACATATGCATCGTCCGGAAGCGGAAAAAGTGGCTGCGCCGCTTTGAACGTGCTCCGGATGCGCGTG
TCGCCTGGTTTGTCTGCCGCACGCGGGTGGCTCCGCGTCTTTCTTTTCCCGCTGGCTAAGGCACTGGC
ACCGGTGTGGAGGTTCTGGCAGTGCAGTATCCGGGCCGTGAGGACCGTCGTCATGAACCGCCGGTTGA
CTCTATCGGTGGTCTGACCAACCGTCTGCTGGAAGTTCTGCGCCCGTTTGGCGACCGTCCGCTGGCACT
GTTCCGTCACCTGATGGGCGCGATTATCGGTTATGAACTGGCGCTGCGCATGCCGAGGCTGGCCTGCC
GGCTCCGGTTACCTGTTTCGCATCTGGTCGTGCTGCGCCGTCTCGTTATCGTGACGATGATGTGCGTGG
TGCGTCTGACGAACGTCTGGTGGCTGAACTGCGTAACTGGGTGGTTCCGACGCGGCGATGCTGGCAGA
TCCGGAACGTGCTGGCGATGGTGCTGCCGGCGATTGCTTCTGACTACCGTGCGGTTGAGACCTACCGCCA
CGAACCGGGTCGCCGTGTTGACTGCCCGGTTACCGTTTCACTGGTGACCATGATCCGCGTGTGAGCGT
TGGCGAAGCTCGTGCTTGGGAGGAACACACTACTGGCCCGGCTGACCTGCGCGTTCTGCCGGGTGGTCA
TTTCTTCTCGTTGATCAGGCGGCACCGATGATCGCTACTATGACCGAGAACTGGCAGGTCCGGCGCT
GACTGGTTCCACTGGCGGCAACTCTTAAGCGGCCGCACTC

Polynucleotide sequence 5: ACX4

CATATGGCAGTTCTGAGCAGCGCAGATCGTGCAAGCAATGAAAAAAAAGTGAAAAGCAGCTATTTTCGAC
CTGCCTCCGATGGAAATGAGCGTTGCATTTCCGCAAGCAACACCGGCAAGCACCTTTCCGCCTTGTACC
AGCGATTATTATCATTTTAATGATCTGCTGACCCCTGAAGAACAGGCCATTTCGTAAAAAAGTTTCGTGAG
TGCATGGAAAAAGAAAGTGGCACCGATTATGACCGAATATTGGGAAAAAGCAGAAATTCGGTTTCATATT
ACCCCGAACTGGGTGCAATGGGTGTTGCCGGTGGTAGCATTAAGGTTATGGTTGTCCGGGTCTGAGC

ATTACCGCAAATGCAATTGCAACCGCAGAAATTGCACGTGTTGATGCAAGCTGTAGCACCTTTATTCTG
GTTTCATAGCAGCCTGGGTATGCTGACCATTGCACTGTGTGGTAGCGAAGCACAGAAAGAAAAATATCTG
CCGAGCCTGGCACAGCTGAATACCGTTGCATGTTGGGCACTGACCGAACCGGATAATGGTAGTGATGCA
AGCGGTCTGGGCACCAACCGCAACCAAAAGTTGAAGGTGGTTGGAAAAATTAACGGTCAGAAACGTTGGATT
GGCAATAGCACCTTTGCAGATCTGCTGATTATCTTTGCACGTAATACCACCACCAATCAGATTAAACGGC
TTCATCGTTAAAAAAGATGCACCGGGTCTGAAAGCAACCAAAATTCGGAACAAAATTGGTCTGCGTATG
GTGCAGAATGGTGATATTCTGCTGCAGAATGTTTTGTGCCGGATGAAGATCGTCTGCCCTGGTGTTAAT
AGCTTTCAGGATACCAGCAAAGTTCTGGCAGTTAGCCGTGTTATGGTTGCATGGCAGCCGATTGGTATT
AGCATGGGTATTTATGATATGTGCCACCGCTATCTGAAAGAACGTAAACAGTTTGGTGCACCGCTGGCA
GCATTTTCAGTGAATCAGCAGAACTGGTTCAGATGCTGGGTAATGTTTCAGGCAATGTTTCTGATGGGT
TGGCGTCTGTGTAAACTGTATGAAACCGGTCAGATGACACCGGGTCAGGCAAGCCTGGGTAAAGCATGG
ATTAGCAGCAAAGCACGTGAAACCGCCAGCCTGGGTCTGTGAACTGCTGGGTGGTAATGGTATTCTGGCA
GATTTTCTGGTTGCAAAAGCATTTTGTGATCTGGAACCGATCTATACCTATGAAGGCACCTATGATATC
AATACCCTGGTTACCGGTCTGTGAAGTTACCGGTATTGCAAGCTTTAAACCGGCAACCCGTAGCCGTCTG
TAACTCGAG

Polynucleotide sequence 6: AcsA

GCTAGCAGGAGATATACATATGCGCGACTACGAACACGTTGTGGAATCTTTCGACTACCTGCAGAGCGC
AACCAGGACCTGCATGGCGAACTGACTGCTCTGAACCGGTGTGTGGAGTGTGTGATCGTCACGCGCA
CGGCGAGGCTGTTGCTCTGTATTGCGAAGCGCAGGATGGCCACGCGGAACGCTACCGTTTCCGCGATCT
GCAGCGCCAGGCTGCACGTTTGGCAACTTCCTGCGTGAGCAGGGCGTGAAACCGGGCGACCGTGTTGC
GGGCTGATGCCGCGTACCGTGGAAGTCTGATCGCAATCCTGGGCACTTGGCGCATCGGTGCGGTTTA
TCAGCCGCTGTTTACTGCGTTCGGTCCGAAAGCAATTGAGCAGCGTCTGAACTGTTCTAACGCACGCTG
GATTGTGACCGATCCGCACAACCGCCGAAACTGGACGACGTTACCGACTGCCCCTCTATTGTGGTTAC
CGGTGGCGCACCGCAGAACCCGGCAGACCATCACTTTTGGTCCGCACTGAACCGTCAGGCGGATGACTG
TGCGCCGGTGCTGCTGGACGCATCTGCACCGTTTCTGCTGATGTGCACCTCCGGCACCCTGGTCCGGC
AAAGCCGCTGGAAGTTCCGCTGTCTGCGATTCTGGCATTAAAGGTTACATGCGTGATGCAATCGACCT
GCGCGCTGATGATCGCTTCTGGAACCTGGCAGATCCGGGTGGGCATACGGTCTGTACTACGCGGTGAC
TGGCCCCGCTGGCGTGTTGGCTACGCGACCCCTGTTTTACGACGGTCCGTTCACTGTTGAATCCACCCGTCA
CATCATCGCGAAATACGCGATTAACAACCTGGCGGGTAGCCCCGACCGCTTACCGCTTTCTGATTGCTGC
AGGCGCGGAATTTGCGGACGCTGTTTCGTGGTCTGCTGCGTGACGTTAGCAGCGCGGGTGAACCGCTGAA
CCCGCAGGTGGTTTCGTTGGTTTGCTGAACAGCTGGGTGTTGTTATCCATGACCACTATGGTCAGACCGA
AATTGGTATGGTTCTGTGCAACCACCGCCTGCGTCACCCGGTTCGCGAGGGTAGCGCTGGCTATGC
AGTGCCGGGCTATCGTATCGTGGTGCTGGATAAAGCACACCGTGAGCTGCCGGCTGGTCAGCCGGGTGT
TCTGGCGGTGGACCGTGAACGCTCCCCGCTGTGCTGGTTTGACGGCTACCTGGGTATGCCGACTCAGGC
GTTTCGCGGGTCGTTATTACCTGTCTGGTGATATCGTTGAACTGAACGACGATGGCAGCATCTCTTTCGT
GGGTGCGCAACGATGATCTGATCACCACCTCTGGTTATCGTGTTGGGCCCCGTTGACGTTGAGTCTGCACT
GATTGAACATCCGGCAGTTGTTGAAGCTGCTGTTATCGGTAAACCGGACCCGCAGCGTACTGAGCTGAT
TAAAGCGTTTGTGTTCTGAACACTCCGTACCTGCCGAGCCCGGAACTGGCGGAAGAAGTGCCTCTGCA
TGTTTCGTGAGCGCTGGCTGCGCACGCATATCCGCGCGAGATGGAGTTTGTGGACCATCTGCCGAAAAC
CCCGTCTGGCAAACCTGCAGCGTTTCATTCTGCGTAACCAGGAAATCGCTAAGCAGCAGGCTCTGGGTTA
ACTCGAG

Polynucleotide sequence 7: 4HBT-ACX4 construct containing the gene encoding 4HBT (blue), the gene encoding ACX4 (green), a spacer sequence encoding a ribosome binding site for ACX4 (**bold underline**), an upstream NdeI restriction site (single underline), and downstream NheI (double underline) and XhoI (dotted underline) restriction sites.

CATATGCACCGTACCTCTAACGGTTCTCACGCGACCGGTGGTAACCTGCCGGATGTTGCTTCTCATTAC
 CCGGTTGCGTACGAACAGACCCTGGACGGCACCGTGGGTTTCGTTATCGACGAAATGACCCCGAACGT
 GCGACCGCGTCTGTTGAAGTTACCGACACCCTGCGTCAGCGTTGGGGTCTGGTTCACGGTGGTGCTTAC
 TGCGCACTGGCGGAGATGCTGGCGACCGAAGCGACCGTTGCGGTTGTGCACGAAAAGGTATGATGGCG
 GTTGGCCAGTCTAACCACACCTCTTTCTTCCGTCCGGTTAAAGAAGGTCACGTTGCGCGGGAAGCTGTT
 CGCATCCACGCGGGTTCTACCACCTGGTTCTGGGACGTTTCCCTGCGCGATGATGCGGGTCGCCTGTGT
 GCGGTGTCTTCTATGTCTATCGCGGTGCGTCCGCGTCGTGACTAA**CAGGAGATATAGAT**ATGGCAGTTC
 TGAGCAGCGCAGATCGTGCAAGCAATGAAAAAAAAAGTGAAAAGCAGCTATTTTCGACCTGCCCTCCGATGG
 AAATGAGCGTTGCATTTCCGCGAGGCAACACCGGCAAGCACCTTTCCGCCTTGTAACAGCGATTATTATC
 ATTTTAATGATCTGCTGACCCCTGAAGAACAGGCCATTTCGTAAAAAGTTCGTGAGTGCAATGGAAAAAG
 AAGTGGCACCGATTATGACCGAATATTGGGAAAAAGCAGAATTCCCGTTTCATATTACCCGAAACTGG
 GTGCAATGGGTGTTGCCGGTGGTAGCATTAAAGGTTATGGTTGTCCGGGTCTGAGCATTACCGCAAATG
 CAATTGCAACCGCAGAAATTGCACGTGTTGATGCAAGCTGTAGCACCTTTATTCTGGTTCATAGCAGCC
 TGGGTATGCTGACCATTGCACTGTGTGGTAGCGAAGCACAGAAAGAAAAATATCTGCCGAGCCTGGCAC
 AGCTGAATACCGTTGCATGTTGGGCACTGACCGAACCGGATAATGGTAGTGATGCAAGCGGTCTGGGCA
 CCACCGCAACCAAAGTTGAAGGTGGTTGGAAAAATTAACGGTCAGAAACGTTGGATTGGCAATAGCACCT
 TTGCAGATCTGCTGATTATCTTTGCACGTAATACCACCACCAATCAGATTAAACGGCTTCATCGTTAAAA
 AAGATGCACCGGGTCTGAAAGCAACCAAAATTCGAACAAATTTGGTCTGCGTATGGTGCAGAATGGTG
 ATATTCTGCTGCAGAATGTTTTTGTGCCGGATGAAGATCGTCTGCCTGGTGTTAATAGCTTTCAGGATA
 CCAGCAAAGTTCTGGCAGTTAGCCGTGTTATGGTTGCATGGCAGCCGATTGGTATTAGCATGGGTATTT
 ATGATATGTGCCACCGCTATCTGAAAGAACGTAAACAGTTTGGTGCACCGCTGGCAGCATTTTCAGCTGA
 ATCAGCAGAACTGGTTCAGATGCTGGGTAATGTTTCAGGCAATGTTTCTGATGGGTGGCGTCTGTGTA
 AACTGTATGAAACCGGTGAGATGACACCGGGTCAGGCAAGCCTGGGTAAAGCATGGATTAGCAGCAAAG
 CACGTGAAACCGCCAGCCTGGGTGCTGAACTGCTGGGTGGTAATGGTATTCTGGCAGATTTTCTGGTTG
 CAAAAGCATTTTGTGATCTGGAACCGATCTATACCTATGAAGGCACCTATGATATCAATACCTGGTTA
 CCGGTGCTGAAGTTACCGGTATTGCAAGCTTTAAACCGGCAACCCGTAGCCGTCTGTAA**GCTAGCTATA**
 ATATCTCGAG

Polynucleotide sequence 8: Section of *P. putida* KT2440 genome amplified by Zhang et al. Contains: bkdA1 (red), bkdA2 (orange), bkdB (green), lpdV (blue).

TGCATCGAATTCAGGAGAAATTAACTATGAACGAGTACGCCCCCTGCGTTTGCATGTGCCCGAGCCCA
CCGGCCGGCCAGGCTGCCAGACCGATTTTTCTACCTGCGCCTCAACGATGCAGGTCAAGCCCGTAAAC
CCGCGATCGATGTGCATGCTGCCGACACTGCCGACCTGTCTACAGCCTGGTCCGCGTGTCTCGACGAGC
AAGGCGATGCGCAAGGCCCTGGGCGAAGACATCGACCCGAGATCCTCCGTCAAGGCATGCGCGCCA
TGCTCAAGACGCGGATCTTCGACAGCCGATGGTGGTTGCCAGCGCCAGAAGAAGATGTCCTTCTACA
TGCAAAGCCTGGGCGAAGAAGCCATCGGCAGCGGCCAGGCGCTGGCGCTGAACCGCACCGACATGTGCT
TCCCAGCTACCGCCAGCAAAGCATCCTGATGGCCCGCAGCTGTGCTGGTTCGAGATGATCTGCCAAC
TGCTGTCCAACGAGCGCGACCCCTCAAGGGCCGCCAGTTGCCGATCATGTATTCCGTGCGCGAAGCCG
GCTTCTTACCATCAGCGGCAACCTGGCGACCCAGTTTCGTGCAGGCGGTTCGGCTGGGCCATGGCATCGG
CGATCAAGGGCGATACCAAGATCGCCTCGGCATGGATCGGTGACGGCGCTACTGCCGAGTCGGACTTCC
ACACCGCCCTTACCTTTGCCACGTATACCGCGCCCCGGTTCATCCTCAACGTGGTCAACAACCAATGGG
CCATTTCCACCTTCCAGGCCATCGCCGGTGGCGAGTCGACCACCTTTGCCGGCCGTGGCGTGGGTGCG
GTATTGCTTCGCTGCGGGTTGACGGCAACGACTTCGTGCGCGTGACGCTGCCTCGCGCTGGGCGGCCG
AGCGCGCCCGCGCGGCTGGGCCCCAAGCCTGATCGAGTGGGTACCTACCGTGCCGGCCCGCACTCGA
CGTCGGACGACCCCTCCAAGTACCGCCCTGCCGACGACTGGAGCCACTTCCCCTGGGTGACCCGATCG
CCCGCTGAAGCAGCACCTGATCAAGATCGGCCACTGGTCCGAGGAAGAACCAGGCCGTACAGGCCG
AGCTCGAAGCTGCGGTGATTGCCGCACAGAAAGGCCGAGCAGTACGGCACCTGGCCAAACGGGCACA
TCCCAGCGCCGCTCGATGTTTCGAGGATGTGTACAAGGAAATGCCCGACCACTGCGCCGTCAACGCC
AGGAACTGGGGTTTTGAGATGAACGACCACAACAACAGCATCAACCCGGAAAACCGCCATGGCCACCACT
ACCATGACCATGATCCAGGCCCTGCGCTCGGCCATGGATGTTCATGCTTGAGCGCGACGACAATGTGGTG
GTGTACGGCCAGGACGTGCGTTACTTCGGCGGCGTGTTCGCTGCACCGAAGGCCTGCAGAACAAAGTAC
GGCAAATCGCGCGTGTTCGACGCGCCATCTCCGAGAGCGGCATCGTCGGTACCGCCGTGGGCATGGGT
GCCTATGGCCTGCGCCCGGTGGTGGAGATCCAGTTTCGCCGACTACTTCTACCGGCCCTCCGACCAGATC
GTCTCCGAGCTGGCCCGCCTGCGTTACCGTTTCGGCCGCGAGTTCAATTGCCCGCTGACCTGCGCATG
CCTTGCGGCGGCGGCATCTATGGCGGCCAGACTCACAGCCAGAGCCCGGAAGCGATGTTACCCAGGTG
TGCGGCCTGCGCACCGTGATGCCGTCCAACCTTATGACGCCAAAGGCCTGTTGATTGCCTCGATCGAA
TGCGACGACCCGTAATCTTCTGGAGCCCCAACGCCTGTACAACGGCCCCGTTCGATGGCCACCACGAC
CGCCCTGTAACCCCGTGGTGAAGCACCCGCACAGCGCCGTGCCCGACGGTTATTACACCGTACCGCTG
GACAAGGCCGCCATTACCCGCCCTGGCAATGACGTGACCGTGCTGACCTACGGCACCAAGGTGTACGTG
GCCCAGGTGGCCGCCGAAGAAAGCGGCGTCGATGCCGAAGTGATCGACCTGCGCAGCCTGTGGCCGCTG
GACCTGGACACTATCGTCGAGTCGGTGAAAAAGACCGGCCGTTCGCTGGTGGTGACAGAGGCCACCCGC
ACCTGCGGCTTCGGTGCCGAGCTGGTGTGCTGGTGACGAGCACTGCTTCCACCACTGGAGGCGCCG
ATCGAACGCTCACCGGCTGGGACACCCCTACCTACCGCACAGGAATGGGCTTACTTCCAGGCCCT
TCGCGGGTAGGTGCGGCACTGAAAAAGGTCATGGAGTCTGAATGGGACGCGACGTATCAAGATGCCG
GACATTGGCGAAGGCATCGCGCAGGTGAGTTGGTGGAATGGTTGTCGAAGTTCGGCGACATCATCGCC
GAGGACCAGGTGGTGGCCGACGTATGACCGACAAGGCCACCGTGGAATCCCCTCGCCGGTCAGCGGC
AAGGTGTTGGCCCTGGGTGGCCAGCCGGGGAAGTGATGGCGGTTCGGTAGCGAACTGATCCGCATCGAA
GTGGAAGGCAGCGGCAACCATGTGGATGTGCTCAGCCAAAACCGGTAGAGGCCCGGCTGCCCCATT
GCAGCCAAGCCGGAACCGCAGAAAGACGTAAAACCCGCCGTGTACCAGGCGCCCGCAACCACGAAGCT
GCGCCCATCGTGCCGCGCCAGCCGGGCGACAAGCCGCTGGCTCGCCTGCCGTGCGCAAACGCGCCCTG
GACGCCGATATCGAACTGCGTTATGTGCATGGTAGCGGCCCGCGCGGCGTATTCTGCACGAAGACCTC
GATGCCTTCATGAGCAAGCCGCAAAGCAATGCCGGGCAAGCACCTGATGGTTATGCCAAGCGCACCGAC
AGCGAGCAGGTGCCAGTGATCGGCCTGCGCCGCAAGATTGCCAGCGCATGCAGGACGCCAAACGCCCG
GTCGCGCACTTCAGTTATGTGAGGAAATCGACGTACCCGCCCTGGAGGCCCTGCGCCAGCAACTCAAC
AGCAAGCACGGCGACAGCCGGGGCAAGCTGACCTTGCTGCCATTCTTGGTACGCGCCCTGGTCTGGCG
CTGCGTGACTTCCCGCAGATCAACCGACCTACGATGACGAAGCGCAGATCATCACCCGCCATGGCGCG
GTGCATGTGGGCATTGCCACCAAGGTGACAACGGCCTGATGGTGCCCGTGCTGCGCCACGCCGAAGCG
GGCAGCCTGTGGGCAATGCCGGCGAGATTTGCGCCTGGCCAACGCTGCACGTAACAACAAGGCCAGC
CGTGAAAGAGCTGTCCGGCTGACCATCACCTGACCGGCTTGGCGCCCTGGGTGGCATTTGTGACGACG
CCGGTGGTCAACACCCCGGAAGTGGAATCGTCGGGTCAACCGCATGGTCAACCGCATGGTGGTGATC
GACGGCCAGATCGTCGTGCGCAAGATGATGAACCTGTCCAGCTCGTTGACCAACCGCGTGGTTCGATGCG
ATGGATGCCGCCCTGTTTCATCCAGGCCGTACGTGGCCTGCTCGAACAACCCGCCCTGCCTGTTCTGGAG
TGAGCATGCAACAGATTATCCAGACTACCTGTTGATCATCGGCGGCGGCCCTGGCGGCTATGTAGCAG
CCATCCGCGCCGGGCAACTGGGCATTCTACCGTACTGGTGGAAAGGCCAGGCACTGGGCGGCACCTGCC
TGAACATCGGCTGCATCCCGTCCAAGGCGCTGATCCACGTGGCCGAGCAGTTTCACCAGGCCCTCGCGCT
TTACCGAACCCCTCGCCGCTGGGCATCAGCGTGGCTTCGCCGCGCCTGGACATCGGCCAGAGCGTCACCT
GGAAGGACGGCATTTGTCGACCGCTGACCACAGGTGTTGCCGCCCTGCTGAAAAAGCACGGGTGAAAG

TGGTGCATGGTTGGGCCAAGGTACTGGACGGCAAGCAGGTGAGGTGATGGCCAGCGTATCCAGTGCG
AGCATCTGTTGCTGGCGACCGGTTCCAGCAGTGTGCAACTGCCTATGCTGCCGCTGGGTGGCCCGGTGA
TTTCCTCGACCGAAGCCCTGGCGCCGAAAAACCTGCCGCAACACCTGGTGGTGGTCGGCGGTGGCTATA
TCGGCCTGGAGCTGGGCATTGCCTATCGCAAGCTGGGTGCACAGGTGAGTGTGGTGAAGCGCGGGAGC
GCATCCTGCCGACCTACGACAGCGAATTGACCGCCCCGGTGGCCGAGTCGCTGAAGAACTGGGCATAG
CGTTGCACCTGGGCCACAGCGTCGAGGGCTACGAAAAATGGCTGCCTGCTGGCCAGCGACGGCAAGGGTG
GGCAACTGCGCCTTGAGGCCGACCAGGTACTGGTGGCCGTGGGACGCCGGCCACGCACCAAGGGCTTCA
ACCTGGAATGCCTGGACCTGAAGATGAACGGCGCCGCCATTGCCATCGACGAGCGTGTACACACAGCA
TGCACAACGTCTGGGCCATCGGCGACGTGCTGGCGAACCAGTGTGGCGCACCGGGCCATGGCCAGG
GCGAGATGGTCGCGAAATCATCGCCGGCAAGCCCCGACGCTTTGAACCGACAGCGATTGCCCGCGTGT
GCTTTACCGACCCGGAAGTGGTGGTGGTGGTGGCAAGACCCCGGAACAAGCCAGCCAGCGGCGCTGGACT
GCATCGTCGCGCAGTTCCCGTTTGCCGCCAATGGCCGGGCCATGAGCCTGGAATCGAAAAGCGTTTCG
TGCGGGTGGTGGCGCGCCGTGACAACCACCTGATCGTGGGTGGCAGGCGGTTGGCGTGGCGGTCTCCG
AGCTATCCACCGCGTTTGCCCAATCGCTGGAGATGGGCGCGTGCCTGGAAGATGTGGCCGGTACCATT
ATGCCACCCACGCTCGGTGAAGCGGTACAGGAAGCCGCACTGCGCGCCCTGGGCCACGCCTTGCCATA
TCTGAAAGCTTGATGCA

Polynucleotide sequence 9: *B. subtilis* 168 genome section containing *lpdV*, *bfmBAA*, *bfmBAB* and *bfmBB*.

CAGATCATACATAGATTGGATTTTCAGACGTTTTAGTTTTATCCGGGTGAAAATGAACTGCAGTCTCTTGCC
CCAAGGGGCGCTAAGAGTATTGCAGGGCGAAGAGCAATCCAAACAATATCGTATTGAATAGACAGACAG
GAGTGAGTCAACATGGCAACTGAGTATGACGTAGTCATTCTGGGCGGCGGTACCGGCGGTTATGTTGCG
GCCATCAGAGCCGCTCAGCTCGGCTTAAAAACAGCCGTTGTGGAAAAAGGAAAACTCGGGGGAACATGT
CTGCATAAAGGCTGTATCCCGAGTAAAGCGCTGCTTAGAAGCGCAGAGGTATACCGGACAGCTCGTGAA
GCCGATCAATTTCGAGTGGAACCGCTGGCGTGTCCCTCAACTTTGAAAAAGTGCAGCAGCGTAAGCAA
GCCGTTGTTGATAAGCTTGCAGCGGGTGTAATCATTTAATGAAAAAGGAAAAATTGACGTGTACACC
GGATATGGACGTATCCTTGGACCGTCAATCTTCTCTCCGTGCCGGGAACAATTTCTGTTGAGCGGGGA
AATGGCGAAGAAAATGACATGCTGATCCCGAAACAAGTGTATGCAACAGGATCAAGACCGAGAATG
CTTCCGGGTCTTGAAGTGGACGGTAAGTCTGTACTGACTTCAGATGAGGCGCTCCAAATGGAGGAGCTG
CCACAGTCAATCATCATTGTGCGCGGAGGGGTTATCGGTATCGAATGGGCGTCTATGCTTCATGATTTT
GGCGTTAAGGTAACGGTTATTGAATACGCGGATCGCATATTGCCGACTGAAGATCTAGAGATTTCAAAA
GAAATGGAAAGTCTTCTTAAGAAAAAAGGCATCCAGTTCATAACAGGGGCAAAAGTGTGCTGACACA
ATGACAAAAACATCAGACGATATCAGCATACAAGCGGAAAAAGACGGAGAAACCGTTACCTATTCTGCT
GAGAAAATGCTTGTTTCCATCGGCAGACAGGCAAATATCGAAGGCATCGGCCAGAGAACACCGATATT
GTTACTGAAAATGGCATGATTTTCAGTCAATGAAAGCTGCCAAACGAAGGAATCTCATATTTATGCAATC
GGAGACGTAATCGGTGGCCTGCAGTTAGCTCAGCTGCTTACATGAGGGAATTATTGCTGTTGAGCAT
TTTGCAAGTCTCAATCCGCATCCGCTTGATCCGACGCTTGTGCGGAAGTGCATTTACTCAAGCCCTGAA
GCTGCCAGTGTGCGCTTAACCGAAGACGAAGCAAAGGCGAAGCGGCATAATGTCAAAATCGGCAAGTTC
CCATTTATGGCGATTGGAAGCGCTTGTATACGGTGAAAGCGACGGTTTTGTCAAATCGTGGCTGAC
CGAGATACAGATGATATTCTCGGCGTTCATATGATTGGCCCGCATGTCACCGACATGATTTCTGAAGCG
GGTCTTGCCAAAGTGTGACGCAACACCGTGGGAGGTGCGGCAAACGATTACCCGCATCCAACGCTT
TCTGAAGCAATTGGAGAAGCTGCGCTTGCCGAGATGGCAAAGCCATTCATTTTTAAAGCATAAAGGA
GGGCTTGAATGAGTACAAACCGACATCAAGCACTAGGGCTGACTGATCAGGAAGCCGTTGATATGTAT
AGAACCATGCTGTAGCAAGAAAAATCGATGAAAGAATGTGGCTGTTAAACCGTCTTGCCAAAATTCCA
TTTGTAATCTCTTGTCAGGACAGGAAGCAGCAGGTAGGAGCGGCTTTCGCACCTGACCGTGAAATG
GATTATGTATTGCCGTACTACAGAGACATGGGTGTGCTGCTCGCGTTTGGCATGACAGCAAAGGACTTA
ATGATGTCCGGGTTTGCAAAGCAGCAGATCCGAACTCAGGAGGCCGCCAGATGCCGGGACATTTCCGGA
CAAAAGAAAAACCGCATTTGTGACGGGATCATCTCCGGTTACAACGCAAGTGCCGCACGCAGTCGGTATT
GCGCTTGCGGGACGTATGGAGAAAAAGGATATCGCAGCCTTTGTTACATTCGGGGGAAGGGTCTTCAAAC
CAAGGCGATTTCCATGAAGGGGCAAACTTTGCCGCTGTCCATAAGCTGCCGGTTATTTTCATGTGTGAA
AACAACAAATACGCAATCTCAGTGCCTTACGATAAGCAAGTCGCATGTGAGAACATTTCCGACCGTGCC
ATAGGCTATGGGATGCCTGGCGTAACTGTGAATGAAATGATCCGCTGGAAGTTTATCAAGCGGTTAA
GAAGCAGCGAAAGGGCACGCAGAGGAGAAGGCCCGACATTAATTGAAACGATTTCTTACCGCCTTACA
CCACATTCCAGTGATGACGATGACAGCAGTACAGAGGCCGTGAAGAAGTAGAGGAAGCGAAAAAAGT
GATCCCCTGCTTACTTATCAAGCTTACTTAAAGGAAACAGGCCCTGCTGTCCGATGAGATAGAACAAC
ATGCTGGATGAAATTATGGCAATCGTAAATGAAGCGACGGATGAAGCGGAGAACGCCCCATATGCAGCT
CCTGAGTCAGCGCTTGATTATGTTTATGCGAAGTAGGGAGGAAGAACAAATGTCAGTAATGTCATATAT
TGATGCAATCAATTTGGCGATGAAAGAAGAAATGGAACGAGATTCTCGCGTTTTCTGCTCTTGGGGAAGA
TGTAGGAAGAAAAGCGGTGTGTTTAAAGCGACAGCGGGACTCTATGAACAATTTGGGGAAGAGCGCT

TATGGATACGCCGCTTGCTGAATCTGCAATCGCAGGAGTCGGTATCGGAGCGGCAATGTACGGAATGAG
ACCGATTGCTGAAATGCAGTTTGTCTGATTTTCATTATGCCGGCAGTCAACCAAATTTATTTCTGAAGCGGC
TAAAATCCGCTACCGCAGCAACAATGACTGGAGCTGTCCGATTGTCTGTCAGAGCGCCATACGGCGGAGG
CGTGACAGGAGCCCTGTATCATTCTCAATCAGTCGAAGCAATTTTCGCCAACCCAGCCCGGACTGAAAAAT
TGTCATGCCATCAACACCATATGACGCGAAAGGGCTCTTAAAAAGCCGAGTTCGTGACGAAGACCCCGT
GCTGTTTTTTGAGCACAAGCGGGCATAACCGTCTGATAAAGGGCGAGGTTCCGGCTGATGATTATGTCCT
GCCAATCGGCAAGGCGGACGTAAAAAGGGAAGGCGACGACATCACAGTGATCACATACGGCCTGTGTGT
CCACTTCGCCTTACAAGCTGCAGAACGTCTCGAAAAAGATGGCATTTTCAGCGCATGTGGTGGATTTAAG
AACAGTTTACCCGCTTGATAAAGAAGCCATCATCGAAGCTGCGTCCAAAACCTGGAAAGGTTCTTTTGGT
CACAGAAGATACAAAAGAAGGCAGCATCATGAGCGAAGTAGCCGCAATTATATCCGAGCATTGTCTGTT
CGACTTAGACGCGCCGATCAAACGGCTTGCAGGTCTGATATTCCGGCTATGCCTTATGCGCCGACAAT
GGAAAAATACTTTATGGTCAACCCTGATAAAGTGGAAAGCGGCGATGAGAGAATTAGCGGAGTTTTAAAG
ACGTAAGCGAGGATACAATCATGGCAATTGAACAAATGACGATGCCGCAGCTTGGAGAAAGCGTAACAG
AGGGGACGATCAGCAAATGGCTTGTGCCCCCGGTGATAAAGTGAACAAATACGATCCGATCGCGGAAG
TCATGACAGATAAGGTAAATGCAGAGGTTCCGTCTTCTTTTACTGGTACGATAACAGAGCTTGTGGGAG
AAGAAGGCCAAACCCTGCAAGTCGGAGAAATGATTTGCAAAATTGAAACAGAAGGCGCGAATCCGGCTG
AACAAAAACAAGAACAGCCAGCAGCATCAGAAGCCGCTGAGAACCCTGTTGCAAAAAGTGCTGGAGCAG
CCGATCAGCCCAATAAAAAAGCGCTACTCGCCAGCTGTTCTCCGTTTGGCCGGAGAGCACGGCATTGACC
TCGATCAAGTGACAGGAAGTGGTGCCGGCGGGCGCATCACACGAAAAGATATTCAGCGCTTAATTGAAA
CAGGCGGCGTGCAAGAACAGAATCCTGAGGAGCTGAAAACAGCAGCTCCTGCACCGAAGTCTGCATCAA
AACCTGAGCCAAAAGAAGAGACGTCATATCCTGCGTCTGCAGCCGGTGATAAAGAAATCCCTGTCACAG
GTGTAAGAAAAGCAATTGCTTCCAATATGAAGCGAAGCAAAAACAGAAATTCGCGCATGCTTGGACGATGA
TGGAAGTCGACGTCACAAATATGGTTGCATATCGCAACAGTATAAAAGATTCTTTTAAGAAGACAGAAG
GCTTTAATTTAACGTTCTTCGCCTTTTTTGTAAAAGCGGTGCTCAGGCGTTAAAAGAATTCCCGCAA
TGAATAGCATGTGGGCGGGGGACAAAATTATTGAAAAAGGATATCAATATTTCAATTGCAGTTGCCA
CAGAGGATTCTTTATTTGTTCCGGTGATTAAAAACGCTGATGAAAAACAATTAAAGGCATTGCGAAAG
ACATTACCGCCTAGCTAAAAAAGTAAGAGACGGAAGTCACTGCAGATGACATGCAGGGAGGCACGT
TTACCGTCAACAACACAGGTTTCGTTTCGGGTCTGTTTCAGTCGATGGGCATTATCAACTACCCTCAGGCTG
CGATTCTTCAAGTAGAATCCATCGTCAAACGCCCCGGTTGTCATGGACAATGGCATGATTGCTGTCAGAG
ACATGGTTAATCTGTGCCTGTCATTAGATCACAGAGTGCTTGACGGTCTCGTGTGCGGACGATTCCCTCG
GACGAGTGAAACAAATTTTAGAATCGATTGACGAGAAGACATCTGTTTACTAAATAAGCAAAAAGAGCA
TTTTTTGAAGTTTTGTTTCAAAAAATGCTCTTTTCTATGCTTTATTATTTCAGCGAT

Polynucleotide sequence 10: ppBCKD operon synthesised by Biomatik corporation.

TCTAGAAATAATTTTGTTTAACTGCTAGCAGGAGAAATTAACATATGAACGAGTACGCCCCCTGCGTTT
GCATGTGCCCCGAGCCCACCGGCCGGCCAGGCTGCCAGACCGATTTTTTCTACCTGCGCCTCAACGATGC
AGGTCAAGCCCGTAAACCCGCGATCGATGTGATGCTGCCGACACTGCCGACCTGTCTACAGCCTGGT
CCGCGTGCTCGACGAGCAAGGCGATGCGCAAGGCCCTGGGCCGAAGACATCGACCCGCGAGATCCTCCG
TCAAGGCATGCGCGCCATGCTCAAGACGCGGATCTTCGACAGCCGCATGGTGGTTGCCAGCGCCAGAA
GAAGATGTCTTTCTACATGCAAAGCCTGGGCGAAGAAGCCATCGGCAGCGGCCAGGCGCTGGCGCTGAA
CCGCACCGACATGTGCTTCCCGACCTACCGCCAGCAAAGCATCTGATGGCCCGCGACGTGTGCTGGT
CGAGATGATCTGCCAACTGCTGTCCAACGAGCGCGACCCCCCTCAAGGGCCGCCAGTTGCCGATCATGTA
TTCGGTGCGCGAAGCCGGCTTCTTCAACATCAGCGGCAACCTGGCGACCCAGTTTCGTGCAGGCGGTGCG
CTGGGCCATGGCATCGGCGATCAAGGGCGATACCAAGATCGCCTCGGCATGGATCGGTGACGGCGCTAC
TGCCGAGTCGGACTTCCACACCGCCCTTACCTTTGCCCACGTATACCGCGCCCCGGTCATCCTCAACGT
GGTCAACAACCAATGGGCCATTTCCACCTTCCAGGCCATCGCCGGTGGCGAGTCGACCACCTTTGCCGG
CCGTGGCGTGGGTGGCGGTATTGCTTCGCTGCGGGTTGACGGCAACGACTTCGTGCGCGTGTACGCTGC
CTCGCGCTGGGCGGCCGAGCGCGCCCCGCCGCGGCTGGGCCCAAGCCTGATCGAGTGGGTACCTACCG
TGCCGGCCCCGACTCGACGTGCGACGACCCCTCCAAGTACCGCCCTGCCGACGACTGGAGCCACTTCCC
GCTGGGTGACCCGATCGCCCCGCTGAAGCAGACCTGATCAAGATCGGCCACTGGTCCGAGGAAGAACA
CCAGGCCGTACGGCCGAGCTCGAAGCTGCGGTGATTGCCGACAGAAAGAACGCCGAGCATACGGCAC
CCTGGCCAACGGGCACATCCCGAGCGCCGCTCGATGTTCGAGGATGTGTACAAGGAAATGCCCGACCA
CCTGCGCCGTCAACGCCAGGAACTGGGGTTTTGAGATGAACGACCACAACAACAGCATCAACCCGAAAA
CCGCCATGGCCACCACTACCATGACCATGATCCAGGCCCTGCGCTCGGCCATGGATGTGATGCTTGAGC
GCGACGACAATGTGGTGGTGTACGGCCAGGACGTGCGTTACTTCGGCGGCGTGTTCGCTGCACCGAAG
GCCTGCAGAACAAGTACGGCAAATCGCGCGTGTTCGACGCGCCCATCTCCGAGAGCGGCATCGTCGGTA
CCGCCGTGGGCATGGGTGCCTATGGCCTGCGCCCGGTGGTGGAGATCCAGTTCGCCGACTACTTCTACC
CGGCCTCCGACCAGATCGTCTCCGAGCTGGCCCGCTGCGTTACCGTTCGGCCGCGAGTTCATTGCC
CGCTGACCCTGCGCATGCCTTGCGGCGGCGGCATCTATGGCGGCCAGACTCACAGCCAGAGCCCGGAAG
CGATGTTTACCCAGGTGTGCGGCCCTGCGCACCGTGATGCCGTCCAACCTTATGACGCCAAAGGCCTGT
TGATTGCCTCGATCGAATGCGACGACCCGGTAATCTTCTGGAGCCCAAACGCCTGTACAACGGCCCCGT
TCGATGGCCACCACGACCGCCCTGTAACCCCGTGGTTCGAAGCACC CGCACAGCGCCGTGCCCGACGGTT
ATTACACCGTACCGCTGGACAAGGCCGCCATTACCCGCCCTGGCAATGACGTGACCGTGTGACCTACG
GCACCACGGTGTACGTGGCCAGGTGGCCGCCGAAGAAAGCGGCGTCGATGCCGAAGTGATCGACCTGC
GCAGCCTGTGGCCGCTGGACCTGGACACTATCGTCGAGTCGGTGAAAAAGACCGGCCGTTGCGTGGTGG
TGCACGAGGCCACCCGCACCTGCGGCTTCGGTGGCGAGCTGGTGTGCTGGTGCAGGAGCACTGCTTCC
ACCACCTTGAGCGCCGATCGAACGCGTCACCGGCTGGGACACCCCCCTACCTCACGCACAGGAATGGG
CTTACTTCCCAGGCCCTTCGCGGGTAGGTGCGGCACTGAAAAAGGTGATGGAGGTCTGAATGGGACGC
ACGTATCAAGATGCCGGACATTGGCGAAGGCATCGCGCAGGTCGAGTTGGTGGAATGGTTTCGTCGAAG
TCGGCGACATCATCGCCGAGGACCAGGTGGTGGCCGACGTCATGACCGACAAGGCCACCGTGGAATCC
CCTCGCCGGTCAGCGGCAAGGTGTTGGCCCTGGGTGGCCAGCCCGGGGAAGTGATGGCGGTGCGTAGCG
AACTGATCCGCATCGAAGTGGAAGGCAGCGGCAACCATGTGGATGTGCCTCAGCCAAAACCGGTAGAGG
CCCCGGCTGCCCCCATTGACGCCAAGCCGGAACCGCAGAAAAGACGTAAAACCCGCCGTGTACCAGGCGC
CCGCCAACCACGAAGCTGCGCCCATCGTGCCGCGCCAGCCGGGCGACAAGCCGCTGGCCTCGCCTGCCG
TGCGCAAACGCGCCCTGGACGCCGGTATCGAACTGCGTTATGTGCATGGTAGCGGCCCGGCCGGCGTA
TTCTGCACGAAGACCTCGATGCCTTCATGAGCAAGCCGCAAAGCAATGCCGGGCAAGCACCTGATGGTT
ATGCCAAGCGCACCGACAGCGAGCAGGTGCCAGTGATCGGCCTGCGCCGCAAGATTGCCAGCGCATGC
AGGACGCCAAACGCCGGGTGCGCGCACTTCAGTTATGTGAGGAAATCGACGTCACCGCCCTGGAGGCC
TGCGCCAGCAACTCAACAGCAAGCACGGCGACAGCCGGGGCAAGCTGACCTTGCTGCCATTCTTGGTAC
GCGCCCTGGTTCGTGGCGCTGCGTGACTTCCCGCAGATCAACGCGACCTACGATGACGAAGCGCAGATCA
TCACCCGCCATGGCGCGGTGCATGTGGGCATTGCCACCCAAGGTGACAACGGCCTGATGGTGCCCGTGC
TGCGCCACGCCGAAGCGGGCAGCCTGTGGGCCAATGCCGGCGAGATTCGCGCCTGGCCAACGCTGCAC
GTAACAACAAGGCCAGCCGTGAAGAGCTGTCCGGCTCGACCATCACCTGACCAGCCTTGGCGCCCTGG
GTGGCATTGTGACGACGCCGGTGGTCAACACCCCGGAAGTGGAATCGTCGGGGTCAACCGCATGGTGC
AACGGCCAGTGGTGATCGACGGCCAGATCGTGTGCGCAAGATGATGAACCTGTCCAGCTCGTTCGACC
ACCGCGTGGTCGATGGCATGGATGCCGCCCTGTTCATCGACGGCCGTACGTGGCCTCGAACAACCCG
CCTGCTGTTTCGTGGAGTGAGCATGCAACAGATTATCCAGACTACCTGTTGATCATCGGCGGCGGCC
TGGCGGCTATGTAGCAGCCATCCGCGCCGGGCAACTGGGCATTCTTACCGTACTGGTGGAAAGGCCAGGC
ACTGGGCGGCACCTGCCTGAACATCGGCTGCATCCCGTCCAAGGCGCTGATCCACGTGGCCGAGCAGTT
TCACCAGGCCTCGCGCTTTACCGAACCCTCGCCGCTGGGCATCAGCGTGGCTTCGCCGCGCCTGGACAT
CGGCCAGAGCGTCACCTGGAAGGACGGCATTGTGACCGCCTGACCACAGGTGTTGCCGCCCTGCTGAA
AAAGCACGGGGTGAAAGTGGTGCATGGTTGGGCCAAGGTACTGGACGGCAAGCAGGTTCGAGGTTCGATGG

CCAGCGTATCCAGTGCAGCATCTGTTGCTGGCGACCGGTTCCAGCAGTGTGCGAACTGCCTATGCTGCC
GCTGGGTGGCCCCGGTGATTTCTCGACCGAAGCCCTGGCGCCGAAAACCCCTGCCGCAACACCTGGTGGT
GGTGGCGGGTGGCTATATCGGCCTGGAGCTGGGCATTGCCATATCGCAAGCTGGGTGCACAGGTGAGTGT
GGTGAAGCGCGGGAGCGCATCCTGCCGACCTACGACAGCGAATTGACCGCCCCGGTGGCCGAGTCGCT
GAAGAACTGGGCATAGCGTTGCACCTGGGCCACAGCGTCGAGGGCTACGAAAAATGGCTGCCCTGCTGGC
CAGCGACGGCAAGGGTGGGCAACTGCGCCTTGAGGCCGACCAGGTACTGGTGGCCGTGGGACGCCGGCC
ACGCACCAAGGGCTTCAACCTGGAATGCCTGGACCTGAAGATGAACGGCGCCGCCATTGCCATCGACGA
GCGCTGTACACCAGCATGCACAACGTCTGGGCCATCGGCGACGTCGCTGGCGAACCAGTGTGGCGCA
CCGGGCCATGGCCCAGGGCGAGATGGTCGCCGAAATCATCGCCGGCAAGGCCCGACGCTTTGAACCGAC
AGCGATTGCCCGGTGTGCTTTACCGACCCGGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
CCAGCAGGGCCTGGACTGCATCGTCGCGCAGTTCCCGTTTGGCCGCAATGGCCGGGCCATGAGCCTGGA
ATCGAAAAGCGGTTTTCGTGCGGGTGGTGGCGCGCCGTGACAACACCTGATCGTGGGTGGCAGGCGGT
TGGCGTGGCGGTCTCCGAGCTATCCACCGCGTTTGGCCAAATCGCTGGAGATGGGCGCGTGCCTGGAAGA
TGTGGCCGGTACCATTTCATGCCCACCAACGCTCGGTGAAGCGGTACAGGAAGCCGCACTGCGCGCCCT
GGGCCACGCCTTGCATATCTGACTCGAG

Polynucleotide sequence 11: Sequence of the PCR product when amplifying *B. subtilis* 168 genes using primers bsBCKD_F and bsBCKD_R

GCTAGCTAGACAGACAGGAGTGAGTCACCATGGCAACTGAGTATGACGTAGTCATTCTGGGCGGCGGTA
CCGGCGGTTATGTTGCGGCCATCAGAGCCGCTCAGCTCGGCTTAAAAACAGCCGTTGTGGAAAAGGAAA
AACTCGGGGGAACATGTCTGCATAAAGGCTGTATCCCGAGTAAAGCGCTGCTTAGAAGCGCAGAGGTAT
ACCGGACAGCTCGTGAAGCCGATCAATTTCGGAGTGGAAACGGCTGGCGTGTCCCTCAACTTTGAAAAAG
TGCAGCAGCGTAAGCAAGCCGTTGTTGATAAGCTTGCAGCGGGTGTAAATCATTTAATGAAAAAGGAA
AAATTGACGTGTACACCGGATATGGACGTATCCTTGGACCGTCAATCTTCTCTCCGCTGCCGGGAACAA
TTTCTGTGTAGCGGGGAAATGGCGAAGAAATGACATGCTGATCCCGAAACAAGTATGATCATTTGCAACAG
GATCAAGAGCGAGAATGCTTCCGGGTCTTGAAGTGGACGGTAAGTCTGTACTGACTTTCAGATGAGCGC
TCCAAATGGAGGAGCTGCCACAGTCAATCATCATTGTTCGGCGGAGGGGTTATCGGTATCGAATGGCGGT
CTATGCTTCATGATTTTGGCGTTAAGGTAACGGTTATTGAATACGCGGATCGCATATTGCCGACTGAAG
ATCTAGAGATTTCAAAAGAAATGGAAAGTCTTCTTAAAGAAAAAGGCATCCAGTTTCATAACAGGGGCAA
AAGTGTCTGCCTGACACAATGACAAAAACATCAGACGATATCAGCATAACAAGCGAAAAAGACGGAGAAA
CCGTTACCTATTCTGCTGAGAAAATGCTTGTTCATCGGCAGACAGGCAAAATATCGAAGGCATCGGCC
TAGAGAACACCGATATTGTTACTGAAAATGGCATGATTTTCAGTCAATGAAAGCTGCCAAACGAAGGAAT
CTCATATTTTATGCAATCGGAGACGTAATCGGTGGCTGCAGTTAGCTCACGTTGCTTCACATGAGGGAA
TTATTGCTGTTGAGCATTTTGCAGGTCTCAATCCGCATCCGCTTGATCCGACGCTTGTGCCGAAGTGCA
TTTACTCAAGCCCTGAAGCTGCCAGTGTTCGGCTTAACCGAAGACGAAGCAAAGGCGAACGGGCATAATG
TCAAAATCGGCAAGTTCCCATTTATGGCGATTGGAAAAGCGCTTGTATACGGTGAAAGCGACGGTTTTG
TCAAAATCGTGGCTGACCGAGATACAGATGATATTCTCGGCGTTTCATATGATTGGCCCGCATGTCACCG
ACATGATTTCTGAAGCGGGTCTTGCCAAAGTGTGGACGCAACACCGTGGGAGGTCGGGCAAACGATTC
ACCCGCATCCAACGCTTTCTGAAGCAATTGGAGAAGCTGCGCTTGCCGAGATGGCAAAGCCATTTCATT
TTTAAAAGCATAAAGGAGGGGCTTGAATGAGTACAAACCGACATCAAGCACTAGGGCTGACTGATCAGG
AAGCCGTTGATATGTATAGAACCATGCTGTTAGCAAGAAAAATCGATGAAAGAATGTGGCTGTTAAACC
GTTCTGGCAAAATTCCATTTGTAATCTCTTGTCAAGGACAGGAAGCAGCACAGGTAGGAGCGGCTTTTCG
CACTTGACCGTGAATGGATTATGTATTGCGGTACTACAGAGACATGGGTGTCGTGCTCGCTTTGGCA
TGACAGCAAAGGACTTAATGATGTCCGGGTTTGCAAAAGCAGCAGATCCGAACCTCAGGAGGCCGCCAGA
TGCCGGGACATTTTCGGACAAAAGAAAAACCGCATTGTGACGGGATCATCTCCGGTTACAACGCAAGTGC
CGCACGAGTCGGTATTGCGCTTGCGGGACGTATGGAGAAAAAGGATATCGCAGCCTTTGTTACATTTCG
GGGAAGGGTCTTCAAACCAAGGCGATTTCCATGAAGGGGCAAACTTTGCCGCTGTCCATAAGCTGCCGG
TTATTTTCATGTGTGAAAACAACAAATACGCAATCTCAGTGCCTTACGATAAGCAAGTCGCATGTGAGA
ACATTTCCGACCGTGCCATAGGCTATGGGATGCCTGGCGTAACTGTGAATGGAAATGATCCGCTGGAAG
TTTATCAAGCGGTTAAAGAAGCACGCGAAAGGGCACGCAGAGGAGAAGGCCCCGACATTAATTGAAACGA
TTTCTTACCGCTTACACCACATTCCAGTGATGACGATGACAGCAGCTACAGAGGCCGTGAAGAAGTAG
AGGAAGCGAAAAAAGTGATCCCTGCTTACTTATCAAGCTTACTTAAAGGAAACAGGCCTGCTGTCCG
ATGAGATAGAACAAACCATGCTGGATGAAATTATGGCAATCGTAAATGAAGCGACGGATGAAGCGGAGA
ACGCCCCATATGCAGCTCCTGAGTCAGCGCTTGATTATGTTTATGCGAAGTAGGGAGGAAGAACAAATG
TCAGTAATGTCATATATTGATGCAATCAATTTGGCGATGAAAGAAGAAATGGAACGAGATTCTCGCGTT
TTCGTCTTGGGGAAGATGTAGGAAGAAAAGGCGGTGTGTTTAAAGCGACAGCGGGACTCTATGAACAA
TTTGGGGAAGAGCGCGTTATGGATACGCCGCTTGCTGAATCTGCAATCGCAGGAGTCGGTATCGGAGCG

GCAATGTACGGAATGAGACCGATTGCTGAAATGCAGTTTGCTGATTTCAATTATGCCGGCAGTCAACCAA
ATTATTTTCTGAAGCGGCTAAAATCCGCTACCGCAGCAACAATGACTGGAGCTGTCCGATTGTGCTCAGA
GCGCCATACGGCGGAGGCGTGCACGGAGCCCTGTATCATTCTCAATCAGTCGAAGCAATTTTCGCCAAC
CAGCCCCGACTGAAAATTGTCATGCCATCAACACCATATGACGCGAAAAGGGCTCTTAAAAGCCGCAGTT
CGTGACGAAGACCCCGTGCTGTTTTTTGAGCACAAAGCGGGCATAACCGTCTGATAAAGGGCGAGGTTCCG
GCTGATGATTATGTCTTGCCAATCGGCAAGGCGGACGTAAAAAGGGAAGGCGACGACATCACAGTGATC
ACATACGGCCTGTGTGTCCACTTCGCCTTACAAGCTGCAGAACGTCTCGAAAAAGATGGCATTTCAGCG
CATGTGGTGGATTTAAGAACAGTTTACCCGCTTGATAAAGAAGCCATCATCGAAGCTGCGTCCAAAAC
GGAAAGGTTCTTTTGGTCACAGAAGATACAAAAGAAGGCAGCATCATGAGCGAAGTAGCCGCAATTATA
TCCGAGCATTGTCTGTTTCGACTTAGACGCGCCGATCAAAACGGCTTGCAAGTTCCTGATATTCCGGCTATG
CCTTATGCGCCGACAATGGAAAAATACTTTATGGTCAACCCCTGATAAAGTGGAAGCGGCGATGAGAGAA
TTAGCGGAGTTTTTAAAGACGTAAGGGAGGATACAATCATGGCAATTGAACAAATGACGATGCCGCAGCT
TGGAGAAAGCGTAACAGAGGGGACGATCAGCAAAATGGCTTGTCGCCCCCGGTGATAAAGTGAACAAATA
CGATCCGATCGCGGAAGTCATGACAGATAAGGTAAATGCAGAGGTTCCGTCTTCTTTTACTGGTACGAT
AACAGAGCTTGTGGGAGAAGAAGGCCAAACCCCTGCAAGTCGGAGAAATGATTTGCAAAATTGAAACAGA
AGGCGCGAATCCGGCTGAACAAAAACAAGAACAGCCAGCAGCATCAGAAGCCGCTGAGAACCCTGTTGC
AAAAAGTGCTGGAGCAGCCGATCAGCCCAATAAAAAGCGCTACTCGCCAGCTGTTCTCCGTTTGGCCGG
AGAGCACGGCATTGACCTCGATCAAGTGACAGGAAGTGGTGCCGGCGGGCGCATCACACGAAAAGATAT
TCAGCGCTTAATTGAAACAGGCGGCGTGCAAGAACAGAATCCTGAGGAGCTGAAAACAGCAGCTCCTGC
ACCGAAGTCTGCATCAAAACCTGAGCCAAAAGAAGAGACGTCATATCCTGCGTCTGCAGCCGGTGATAA
AGAAATCCCTGTACAGGTGTAAGAAAAGCAATTGCTTCCAATATGAAGCGAAGCAAAACAGAAATTC
GCATGCTTGGACGATGATGGAAGTCGACGTCACAAAATATGGTTGCATATCGCAACAGTATAAAAAGATT
TTTTAAGAAGACAGAAGGCTTTAATTTAACGTTCTTCGCCCTTTTTTGTAAGGCGGTCGCTCAGGCGTT
AAAAGAATTCCCGCAAATGAATAGCATGTGGGCGGGGGACAAAATTATTCAGAAAAAGGATATCAATAT
TTCAATTGCAGTTGCCACAGAGGATTCTTTATTTGTTCCGGTGATTAAAAACGCTGATGAAAAACAAT
TAAAGGCATTGCGAAAGACATTACCGGCCTAGCTAAAAAAGTAAGAGACGGAAAACTCACTGCAGATGA
CATGACGGGAGGCAGCTTTACCGTCAACAACAGAGTTCTGTTTCGGGTCTGTTTCAGTCGATGGGCATTAT
CAACTACCCTCAGGCTGCGATTCTTCAAGTAGAATCCATCGTCAAAACGCCCGGTTGTCATGGACAATGG
CATGATTGCTGTGACAGACATGGTTAATCTGTGCCTGTCAATTAGATCACAGAGTGCTTGACGGTCTCGT
GTGCGGACGATTCTCGGACGAGTGAAACAAAATTTTAGAATCGATTGACGAGAAGACATCTGTTTACTA
ACTCGAG

Polynucleotide sequence 12: Wild-type *ilvD* from *E. coli* MG1655

ATGCCTAAGTACCGTTCCGCCACCACCACTCATGGTCGTAATATGGCGGGTGCTCGTGCGCTGTGGCGC
GCCACCGGAATGACCGACGCCGATTTCGGTAAGCCGATTATCGCGGTTGTGAACCTCGTTCACCCAATTT
GTACCGGGTCACGTCCATCTGCGCGATCTCGGTAAACTGGTCGCCGAACAAATTGAAGCGGCTGGCGGC
GTTGCCAAAGAGTTCAACACCATTTGCGGTGGATGATGGGATTGCCATGGGCCACGGGGGGATGCTTTAT
TCACTGCCATCTCGCGAACTGATCGCTGATTCCGTTGAGTATATGGTCAACGCCCACTGCGCCGACGCC
ATGGTCTGCATCTCTAACTGCGACAAAATCACCCCGGGGATGCTGATGGCTTCCCTGCGCCTGAATATT
CCGGTGATCTTTGTTTCCGGCGGCCCGATGGAGGCGGGGAAAAACCAAATTTCCGATCAGATCATCAAG
CTCGATCTGGTTGATGCGATGATCCAGGGCGCAGACCCGAAAGTATCTGACTCCAGAGCGATCAGGTT
GAACGTTCCGCGTGTCCGACCTGCGGTTCTGTCTCCGGGATGTTTACCGCTAACTCAATGAAGTGCCTG
ACCGAAGCGCTGGGCCTGTGCGAGCCGGGCAACGGCTCGCTGCTGGCAACCCACGCCGACCGTAAGCAG
CTGTTCCCTTAATGCTGGTAACGCATTGTTGAATTGACCAACGTTATTACGAGCAAAACGACGAAAGT
GCACTGCCGCGTAATATCGCCAGTAAGGCGGCGTTTTGAAAACGCCATGACGCTGGATATCGCGATGGGT
GGATCGACTAACACCGTACTTCACCTGCTGGCGGCGGCGCAGGAAGCGGAAATCGACTTCACCATGAGT
GATATCGATAAGCTTTCCCGCAAGGTTCCACAGCTGTGTAAGTTGCGCCGAGCACCCAGAAATACCAT
ATGGAAGATGTTACCGTGCTGGTGGTGTTATCGGTATTCTCGGCGAACTGGATCGCGCGGGGTTACTG
AACCGTGATGTGAAAAACGTACTTGGCCTGACGTTGCCGCAAACGCTGGAACAATACGACGTTATGCTG
ACCCAGGATGACGCGGTAAAAAATATGTTCCGCGCAGGTCCTGCAGGCATTTCGTACCACACAGGCATT
TCGCAAGATTGCCGTTGGGATACGCTGGACGACGATCGCGCCAATGGCTGTATCCGCTCGCTGGAACAC
GCCTACAGCAAAGACGGCGGCCCTGGCGGTGCTCTACGGTAACTTTGCGGAAAACGGCTGCATCGTGAAA
ACGGCAGGCGTCGATGACAGCATCCTCAAATTCACCGGCCCGGCGAAAAGTGACGAAAGCCAGGACGAT
GCGGTAGAAGCGATTCTCGGCGGTAAAGTTGTGCGCCGAGATGTGGTAGTAATTTCGCTATGAAGGCCCG
AAAGGCGGTCCGGGGATGCAGGAAATGCTCTACCCAACCAGCTTCTGAAATCAATGGGTCTCGGCAAA
GCCTGTGCGCTGATCACCGACGGTCGTTTCTCTGGTGGCACCTCTGGTCTTTCCATCGGCCACGTCCTCA
CCGGAAGCGGCAAGCGGCGGCAGCATTGGCCTGATTGAAGATGGTGACCTGATCGCTATCGACATCCCG

AACCGTGGCATTTCAGTTACAGGTAAGCGATGCCGAACTGGCGGCGCGTCTGTGAAGCGCAGGACGCTCGA
GGTGACAAAGCCTGGACGCCGAAAAATCGTGAACGTCAGGTCTCCTTTGCCCTGCGTGCTTATGCCAGC
CTGGCAACCAGCGCCGACAAAGGCGCGGTGCGCGATAAATCGAAACTGGGGGGTTAA

Polynucleotide sequence 13: Wild-type *alsS* from *B. subtilis* 168

TTGACAAAAGCAACAAAAGAACAAAAATCCCTTGTGAAAAACAGAGGGGCGGAGCTTGTTGTTGATTGC
TTAGTGAGCAAGGTGTCACACATGTATTTGGCATTCCAGGTGCAAAAAATTGATGCGGTATTTGACGCT
TTACAAGATAAAGGACCTGAAATTATCGTTGCCCGGCACGAACAAAACGCAGCATTTCATGGCCCAAGCA
GTCGGCCGTTTAACTGGAAAACCGGGAGTCGTGTTAGTCACATCAGGACCGGGTGCCTCTAACTTGGCA
ACAGGCCTGCTGACAGCGAACACTGAAGGAGACCCTGTCGTTGCGCTTGCTGGAAACGTGATCCGTGCA
GATCGTTTAAACGGACACATCAATCTTTGGATAATGCGGCGCTATTCCAGCCGATTACAAAATACAGT
GTAGAAGTTCAAGATGTAAAAAATATACCGGAAGCTGTTACAAATGCATTTAGGATAGCGTCAGCAGGG
CAGGCTGGGGCCGCTTTTGTGAGCTTTCCGCAAGATGTTGTGAATGAAGTACAAAATACGAAAAACGTG
CGTGCTGTTGACGCGCCAAAACCTCGGTCTGACGAGATGATGCAATCAGTGCGGCCATAGCAAAAATC
CAAACAGCAAAACTTCCTGTCTGTTTGGTCCGCATGAAAGGCGGAAGACCGGAAGCAATTAAAGCGGTT
CGCAAGCTTTTGAAAAAGGTTTCAGCTTCCATTTGTTGAAACATATCAAGCTGCCGGTACCCTTTCTAGA
GATTTAGAGGATCAATATTTTGGCCGTATCGGTTTGTTCGCAACCAGCCTGGCGATTTACTGCTAGAG
CAGGCAGATGTTGTTCTGACGATCGGCTATGACCCGATTGAATATGATCCGAAATTCTGGAATATCAAT
GGAGACCGGACAAATTATCCATTTAGACGAGATTATCGCTGACATTGATCATGCTTACCAGCCTGATCTT
GAATTGATCGGTGACATTTCCGTCCACGATCAATCATATCGAACACGATGCTGTGAAAGTGGAATTTGCA
GAGCGTGAGCAGAAAAATCCTTTCTGATTTAAACAATATATGCATGAAGGTGAGCAGGTGCCTGCAGAT
TGGAATCAGACAGAGCGCACCTCTTGAAATCGTTAAAGAGTTGCGTAATGCAGTCGATGATCATGTT
ACAGTAACTTGCGATATCGGTTTCGCACGCCATTTGGATGTCACGTTATTTCCGCAGCTACGAGCCGTTA
ACATTAATGATCAGTAACGGTATGCAACACTCGGCGTTGCGCTTCCTTGGGCAATCGGCGCTTCATTG
GTGAAACCGGGAGAAAAAGTGGTTTCTGTCTCTGGTGACGGCGGTTTCTTATTCTCAGCAATGGAATTA
GAGACAGCAGTTGACTAAAAGCACCAATTGTACACATTGTATGGAACGACAGCACATATGACATGGTT
GCATTCCAGCAATTGAAAAAATATAACCGTACATCTGCGGTGATTTTCGAAATATCGATATCGTGAAA
TATGCGGAAAGCTTCGGAGCAACTGGCTTGCGCGTAGAATCACCAGACCAGCTGGCAGATGTTCTGCGT
CAAGGCATGAACGCTGAAGGTCCTGTCTATCATCGATGTCCCGGTTGACTACAGTGATAACATTAATTTA
GCAAGTGACAAGCTTCCGAAAGAATTCGGGGAACTCATGAAAACGAAAGCTCTCTAG

Polynucleotide sequence 14: Wild-type *ilvC* from *E. coli* MG1655

ATGGCTAACTACTTCAATACACTGAATCTGCGCCAGCAGCTGGCACAGCTGGGCAAATGTCGCTTTATG
GGCCGCGATGAATTCGCCGATGGCGCGAGCTACCTTCAGGGTAAAAAAGTAGTCATCGTCGGCTGTGGC
GCACAGGGTCTGAACCAGGGCCTGAACATGCGTGATTCTGGTCTCGATATCTCTACGCTCTGCGTAAA
GAAGCGATTGCCGAGAAGCGCGCGTCTGGCGTAAAGCGACCGAAAAATGGTTTTAAAGTGGGTACTTAC
GAAGAACTGATCCACAGGCGGATCTGGTGATTAACCTGACGCCGACAAAGCAGCACTCTGATGTAGTG
CGCACCGTACAGCCACTGATGAAAGACGGCGCGGCGCTGGGCTACTCGCACGGTTTCAACATCGTCGAA
GTGGGCGAGCAGATCCGTAAAGATATCACCGTAGTGATGGTTGCGCCGAAATGCCCAGGCACCGAAGTG
CGTGAAGAGTACAAACGTGGGTTTCGGCGTACCGACGCTGATTGCCGTTCACCCGGAAAACGATCCGAAA
GGCGAAGGCATGGCGATTGCCAAAGCCTGGGCGGCTGCAACCGGTGGTCACCGTGCGGGTGTGCTGGAA
TCGTCTTTCGTTGCGGAAGTGAAATCTGACCTGATGGGCGAGCAAACCATCCTGTGCGGTATGTTGCAG
GCTGGCTCTCTGCTGTGCTTCGACAAGCTGGTGGAAGAAGGTACCGATCCAGCATACGCAGAAAAACTG
ATTCAGTTTCGGTTGGGAAACCATCACCGAAGCACTGAAACAGGGCGGCATCACCTGATGATGGACCGT
CTCTCTAACCCGGCGAAACTGCGTGCTTATGCGCTTTCTGAACAGCTGAAAGAGATCATGGCACCCCTG
TTCCAGAAACATATGGACGACATCATCTCCGGCGAATTCTCTTCCGGTATGATGGCGGACTGGGCCAAC
GATGATAAGAAACTGCTGACCTGGCGTGAAGAGACCGGCAAAAACCGCGTTTGAAACCGCGCCGAGTAT
GAAGGCAAAATCGGCGAGCAGGAGTACTTCGATAAAGGCGTACTGATGATTGCGATGGTGAAAGCGGGC
GTTGAACTGGCGTTTCGAAACCATGGTCGATTCCGGCATCATTTGAAGAGTCTGCATATTATGAATCACTG
CACGAGCTGCCGCTGATTGCCAACACCATCGCCCGTAAGCGTCTGTACGAAATGAACGTGGTTATCTCT
GATACCGCTGAGTACGGTAACATCTGTCTCTTACGCTTGTGTGCCGTTGCTGAAACCGTTTTATGGCA
GAGCTGCAACCGGGCGACCTGGGTAAAGCTATTCCGGAAGGCGCGGTAGATAACGGGCAACTGCGTGAT
GTGAACGAAGCGATTTCGACCCATGCGATTGAGCAGGTAGGTAAGAAACTGCGCGGCTATATGACAGAT
ATGAAACGTATTGCTGTTGCGGGTTAA

Polynucleotide sequence 15: ilvD-alsS-ilvC

GAATTCCGGAAGGAGATATACCATGCCTAAGTACCGTTCCGCCACCACCCTCATGGTCGTAATATGGC
GGGTGCTCGTGCGCTGTGGCGCGCCACCGGAATGACCGACGCCGATTTCCGGTAAGCCGATTATCGCGGT
TGTGAACTCGTTACCCAATTTGTACCGGGTCACGTCCATCTGCGCGATCTCGGTAAACTGGTCGCCGA
ACAAATTGAAGCGGCTGGCGGCGTTGCCAAAGAGTTCAACACCATTGCGGTGGATGATGGGATTGCCAT
GGGCCACGGGGGGATGCTTTATTTCACTGCCATCTCGGAACTGATCGCTGATTCCGTTGAGTATATGGT
CAACGCCCCTGCGCCGACGCCATGGTCTGCATCTCTAACTGCGACAAAATCACCCCGGGGATGCTGAT
GGCTTCCCTGCGCCTGAATATTCCGGTGATCTTTGTTTCCGGCGGCCCGATGGAGGCCGGGAAAACCAA
ACTTTCCGATCAGATCATCAAGCTCGATCTGGTTGATGCGATGATCCAGGGCGCAGACCCGAAAGTATC
TGACTCCCAGAGCGATCAGGTTGAACGTTCCGCGTGTCCGACCTGCGGTTCCCTGCTCCGGGATGTTTAC
CGCTAACTCAATGAACTGCCTGACCGAAGCGCTGGGCCTGTGCGAGCCGGGCAACGGCTCGCTGCTGGC
AACCCACGCCGACCGTAAGCAGCTGTTCCCTTAATGCTGGTAAACGCATTGTTGAATTGACCAAACGTTA
TTACGAGCAAAACGACGAAAGTGCACTGCCCGTAATATCGCCAGTAAGGCGGCGTTTGAACCGCCAT
GACGCTGGATATCGCGATGGGTGGATCGACTAACACCGTACTTCACCTGCTGGCGGCGGCGCAGGAAGC
GGAAATCGACTTCACCATGAGTGATATCGATAAGCTCTCCCGCAAGGTTCCACAGCTGTGTAAAGTTGC
GCCGAGCACCCAGAAATACCACATGGAAGATGTTACCGTGCTGGTGGTGTTATCGGTATTCTCGGCGA
ACTGGATCGCGCGGGGTTACTGAACCGTGATGTGAAAAACGTACTTGGCCTGACGTTGCCGCAACCGCT
GGAACAATACGACGTTATGCTGACCCAGGATGACGCGGTAAAAAATATGTTCCGCGCAGGTCCCTGCAGG
CATTCGTACCACACAGGCATTCTCGCAAGATTGCCGTTGGGATACGCTGGACGACGATCGCGCCAATGG
CTGTATCCGCTCGCTGGAACACGCCTACAGCAAAAGACGGCGGCTGGCGGTGCTCTACGGTAACTTTGC
GGAAACCGCTGCATCGTGAAAAACGGCAGGCGTGATGACAGCATCCTCAAATTCACCGGCCCGGCGAA
AGTGTAAGCAAGCCAGGACGATGCGGTAGAAGCGATTCTCGGCGGTAAAGTTGTCGCCGGAGATGTGGT
AGTAATTGCTATGAAGGCCCGAAAGGCGGTCCGGGGATGCAGGAAATGCTCTACCCAACAGCTTCCT
GAAATCAATGGGTCTCGGCAAGCCTGTGCGCTGATCACCGACGGTCGTTTCTCTGGTGGCACCTCTGG
TCTTTCCATCGGCCACGTCTCACCGGAAGCGGCAAGCGGCGGCAGCATTGGCCTGATTGAAGATGGTGA
CCTGATCGCTATCGACATCCCGAACCGTGGCATTTCAGTTACAGGTAAAGCGATGCCGAACCTGGCGGCGCG
TCGTGAAGCGCAGGACGCACGAGGTGACAAAGCCTGGACGCCGAAAAATCGTGAACGTCAGGTCTCCTT
TGCCCTGCGTGCTTATGCCAGCCTGGCAACCAGCGCCGACAAAGGCGCGGTGCGCGATAAATCGAAACT
GGGGGGTTAACTAGCAGGAGAAATTAATATGACAAAAAGCAACAAAAAGCAAAAAATCCCTTGTGAAA
AACAGAGGGGGCGGAGCTTGTGTGTTGATTGCTTAGTGAGCAAGGTGTCACACATGTATTTGGCATTCCA
GGTGCAAAATTTGATGCGGTATTTGACGCTTTACAAGATAAAGGACCTGAAATTTATCGTTGCCCGGCAC
GAACAAAACGCAGCATTTCATGGCCCAAGCAGTCGGCCGTTTAACTGGAAAAACCGGGAGTCGTGTTAGTC
ACATCAGGACCGGGTGCTCTAACTTGGCAACAGGCCTGCTGACAGCGAACACTGAAGGAGACCCGTGTC
GTTGCGCTTGCTGGAACCGTGATCCGTGCAGATCGTTTAAACCGGACACATCAATCTTTGGATAATGCG
GCGCTATTTCAGCCGATTACAAAATACAGTGTAAGTTCAAGATGTAAAAAATATACCGGAAGCTGTT
ACAAATGCATTTAGGATAGCGTCAGCAGGGCAGGCTGGGGCCGCTTTTGTGAGCTTTCCGCAAGCTGTT
GTGAATGAAGTCACAAATACGAAAAACGTGCGTGCTGTTGCAGCGCCAAAACCTCGGTCTGCAGCAGAT
GATGCAATCAGTGCGGCCATAGCAAAAATCCAAACAGCAAAAACCTTCCTGTCGTTTTTGGTCGGCATGAAA
GGCGGAAGACCGGAAGCAATTAAGCGGTTTCGCAAGCTCTTGAAAAAGGTTTCAGCTTCCATTTGTTGAA
ACATATCAAGCTGCCGGTACCCCTTTCTAGAGATTTAGAGGATCAATATTTTGGCCGTATCGGTTTTGTTT
CGCAACCAGCCTGGCGATTTACTGCTAGAGCAGGCAGATGTTGTTCTGACGATCGGCTATGACCCGATT
GAATATGATCCGAAATTTCTGGAATATCAATGGAGACCGGACAATTATCCATTTAGACGAGATTATCGCT
GACATTGATCATGCTTACCAGCCTGATCTTGAATTGATCGGTGACATTCCGTCCACGATCAATCATATC
GAACACGATGCTGTGAAAGTGGAATTTGCAGAGCGTGAGCAGAAAAATCCTTTCTGATTTAAACAATAT
ATGCATGAAGGTGAGCAGGTGCCTGCAGATTGGAAATCAGACAGAGCGCACCCCTCTTGAAATCGTTAAA
GAGTTGCGTAATGCAGTCGATGATCATGTTACAGTAACTTGCGATATCGGTTTCGCACGCCATTTGGATG
TCACGTTATTTCCGCGAGCTACGAGCCGTTAACATTAATGATCAGTAACGGTATGCAAAACACTCGGCGTT
GCGCTTCCTTGGGCAATCGGCGCTTCATTGGTGAAACCGGGGAGAAAAAGTGTTTCTGTCTCTGGTGAC
GGCGGTTTCTTATTCTCAGCAATGGAATTAGAGACAGCAGTTTCGACTAAAAGCACCATTGTACACATT
GTATGGAACGACAGCACATACGACATGGTTGCATTCCAGCAATTGAAAAAATATAACCGTACATCTGCG
GTCGATTTTCGAAATATCGATATCGTGAAATATGCGGAAAGTTTTCGGAGCAACTGGCTTGCGCGTAGAA
TCACAGACACAGCTGGCAGATGTTCTGCGTCAAGCTGAACGCTGAAGGTCCGTGTCATCATCGATGTC
CCGGTTGACTACAGTGATAACATTAATTTAGCAAGTGACAAGCTCCCGAAAGAAATTTGGGGAACCTCATG
AAAACGAAAGCTCTCTAAACAACATCACGAGGAATCACCATGGCTAACTACTTCAATACACTGAATCTGC
GCCAGCAGCTGGCACAGCTGGGCAAAATGTCGCTTTATGGGCGCGGATGAATTTGCCGATGGCGCGAGCT
ACCTTCAGGGTAAAAAAGTAGTCATCGTCCGCTGTGGCGCACAGGGTCTGAACCAGGGCCTGAACATGC
GTGATTCTGGTCTCGATATCTCCTACGCTCTGCGTAAAGAAGCGATTGCCGAGAAGCGCGCGTCTTGGC
GTAAAGCGACCGAAAAATGGTTTTAAAGTGGGTACTTACGAAGAACTGATCCACAGGCGGATCTGGTGA
TTAACTGACGCCGACAAGCAGCACTCTGATGTAGTGCGCACCGTACAGCCACTGATGAAAGACGGCG

CGGCGCTGGGCTACTCGCACGGTTTCAACATCGTCGAAGTGGGCGAGCAGATCCGTAAAGATATCACCG
TAGTGATGGTTGCGCCGAAATGCCCAGGCACCGAAGTGC GTGAAGAGTACAAACGTGGGTTCGGCGTAC
CGACGCTGATTGCCGTTTACCCGGAAAAACGATCCGAAAGGCGAAGGCATGGCGATTGCCAAAGCCTGGG
CGGCTGCAACCGGTGGTCACCGTGCGGGTGTGCTGGAATCGTCCTTCGTTGCGGAAGTGAAATCTGACC
TGATGGGCGAGCAAACCATCCTGTGCGGTATGTTGCAGGCTGGCTCTCTGCTGTGCTTCGACAAGCTGG
TGGAAGAAGGTACCGATCCAGCATACGCAGAAAAACTGATTTCAGTTCGGTTGGGAAACCATCACCGAAG
CACTGAAACAGGGCGGCATCACCTGATGATGGACCGTCTCTCTAACCCGGCGAAACTGCGTGCTTATG
CGCTTTCTGAACAGCTGAAAGAGATCATGGCACCCCTGTTCCAGAAACACATGGACGACATCATCTCCG
GCGAATTTTCTTCCGGTATGATGGCGGACTGGGCCAACGATGATAAGAACTGCTGACCTGGCGTGAAG
AGACCGGCAAAACCGCGTTTGAACCCGCGCCGAGTATGAAGGCAAAATCGGCGAGCAGGAGTACTTCG
ATAAAGGCGTACTGATGATTGCGATGGTGAAAGCGGGCGTTGAACTGGCGTTCGAAACCATGGTCGATT
CCGGCATCATTGAAGAGTCTGCATATTATGAATCACTGCACGAGCTGCCGCTGATTGCCAACACCATCG
CCCGTAAGCGTCTGTACGAAATGAACGTGGTTATCTCTGATACCGCTGAGTACGGTAACTATCTGTTCT
CTTACGCTTGTGTGCCGTTGCTGAAACCGTTTATGGCAGAGCTGCAACCGGGCGACCTGGGTAAAGCTA
TTCCGGAAGGCGCGGTAGATAACGGGCAACTGCGTGATGTGAACGAAGCGATTTCGCAGCCATGCGATTG
AGCAGGTAGGTAAGAAACTGCGCGGCTATATGACAGATATGAAACGTATTGCTGTTGCGGGTTAAAAGC
TT

8.3 Appendix 3 – Additional figures

Entry	E.C number	Protein names	Organism	Protein DataBank ID
Q1EMV2	3.1.2.29	Fluoroacetyl-CoA thioesterase	Streptomyces cattleya	3KUV;3KUW;3KV7;3KV8;3KVI;3KVU;3KVZ;3KW1;3KX7;3KX8;3P2Q;3P2R;3P2S;3P3F;3P3I;
Q04416	3.1.2.23	4-hydroxybenzoyl-CoA thioesterase	Arthrobacter sp.	1Q4S;1Q4T;1Q4U;3R32;3R34;3R35;3R36;3R37;3R3A;3R3B;3R3C;3R3D;3R3F;3TEA;
P56653	3.1.2.23	4-hydroxybenzoyl-CoA thioesterase	Pseudomonas sp. (strain CBS-3)	1BVQ;1LO7;1LO8;1LO9;
Q8N1Q8	3.1.2.2	Acyl-coenzyme A thioesterase THEM5	Homo sapiens	4AE7;
P49753	3.1.2.2	Acyl-coenzyme A thioesterase 2	Homo sapiens	3HLK;
Q8N9L9	3.1.2.2	Acyl-coenzyme A thioesterase 4	Homo sapiens	3K2I;
Q8WXI4	3.1.2.-	Acyl-coenzyme A thioesterase 11	Homo sapiens	3FO5;
Q8WYK0	3.1.2.1	Acyl-coenzyme A thioesterase 12	Homo sapiens	3B7K;4MOB;4MOC;
Q9CQR4	3.1.2.-	Acyl-coenzyme A thioesterase 13	Mus musculus	2CY9;
Q9NPJ3	3.1.2.-	Acyl-coenzyme A thioesterase 13	Homo sapiens	2FOX;2H4U;3F5O;
O00154	3.1.2.2	Cytosolic acyl coenzyme A thioester hydrolase	Homo sapiens	2QQ2;
Q91V12	3.1.2.2	Cytosolic acyl coenzyme A thioester hydrolase	Mus musculus	2Q2B;2V1O;4ZV3;
Q9SX65	3.1.2.-	1,4-dihydroxy-2-naphthoyl-CoA thioesterase 1	Arabidopsis thaliana	4K02;
Q7V4A7	3.1.2.28	1,4-dihydroxy-2-naphthoyl-CoA hydrolase	Prochlorococcus marinus (strain MIT 9313)	2HX5;
Q55777	3.1.2.28	1,4-dihydroxy-2-naphthoyl-CoA hydrolase	Synechocystis sp. (strain PCC 6803 / Kazusa)	4K00;
P0A8Y8	3.1.2.-	Proofreading thioesterase Enth	Escherichia coli (strain K12)	1VH9;4K4C;4K4D;
P10768	3.1.2.12	S-formylglutathione hydrolase (FGH)	Homo sapiens (Human)	3FCX;
P77712	3.1.2.-	Long-chain acyl-CoA thioesterase FadM	Escherichia coli (strain K12)	1NJK;
Q6NVY1	3.1.2.4	3-hydroxyisobutyryl-CoA hydrolase	Homo sapiens (Human)	3BPT;
Q75608	3.1.2.-	Acyl-protein thioesterase 1 (APT-1)	Homo sapiens (Human)	1FJ2;
Q5VWZ2	3.1.2.-	Lysophospholipase-like protein 1	Homo sapiens (Human)	3U0V;
P77781	3.1.2.28	1,4-dihydroxy-2-naphthoyl-CoA hydrolase	Escherichia coli (strain K12)	1SBK;1VH5;1VI8;4K49;4K4A;4K4B;
P76084	3.1.2.-	Acyl-coenzyme A thioesterase Paal	Escherichia coli (strain K12)	1PSU;2FS2;
Q9UMR5	3.1.2.-	Lysosomal thioesterase PPT2 (PPT-2)	Homo sapiens (Human)	1PJA;

P41903	3.1.2.2	Peroxisomal acyl-coenzyme A thioester hydrolase 1	Saccharomyces cerevisiae (strain ATCC 204508 / S288c)	1TBU;
P40363	3.1.2.12	S-formylglutathione hydrolase (FGH)	Saccharomyces cerevisiae (strain ATCC 204508 / S288c)	1PV1;3C6B;4FLM;4FOL;
P0ADA1	3.1.2.-	Acyl-CoA thioesterase I (Lecithinase B)	Escherichia coli (strain K12)	1IVN;1J00;1JRL;1U8U;1V2G;
P0AGG2	3.1.2.-	Acyl-CoA thioesterase 2	Escherichia coli (strain K12)	1C8U;
Q5T1C6	3.1.2.2	Acyl-coenzyme A thioesterase THEM4	Homo sapiens (Human)	4AE8;4GAH;
P45083	3.1.2.-	Putative esterase HI_1161	Haemophilus influenzae (strain ATCC 51907 / DSM 11121 / KW20 / Rd)	1SC0;2B6E;3LZ7;
Q9I3A4	3.1.2.-	Putative esterase PA1618	Pseudomonas aeruginosa (strain ATCC 15692 / PAO1 / 1C / PRS 101 / LMG 12228)	4QD7;4QD8;4QD9;4QDA;4QDB;
P9WIM3	3.1.2.-	Putative esterase Rv1847	Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv)	3S4K;
P44886	3.1.2.-	Uncharacterized acyl-CoA thioester hydrolase HI_0827	Haemophilus influenzae (strain ATCC 51907 / DSM 11121 / KW20 / Rd)	1YLI;3BJK;
Q46948	3.1.2.-	Probable protein deglycase (EC 3.1.2.-)	Escherichia coli (strain K12)	2AB0;
P0A8Z3	3.1.2.-	Acyl-CoA thioester hydrolase YbgC (Acyl-CoA thioesterase)	Escherichia coli (strain K12)	1S5U;
P94842	3.1.2.-	Acyl-CoA thioesterase YbgC	Helicobacter pylori (strain ATCC 700392 / 26695) (Campylobacter pylori)	2PZH;
P46130	3.1.2.-	Putative acyl-CoA thioester hydrolase YbhC	Escherichia coli (strain K12)	3GRH;

Table 34: List of reviewed acyl-CoA thioesterases in UniProt database with X-Ray crystal structures

Expression strain	Incubation temperature (°C)	Induction point (OD ₆₀₀)
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+) (control)	21°C	0.46
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):4HBT		0.60
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):FcoT		0.40
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):RifR		0.64
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):YciA		0.44
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+) (control)	30°C	0.48
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):4HBT		0.67
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):FcoT		0.67
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):RifR		0.60
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):YciA		0.48
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):4HBT	37°C	1.0
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):FcoT	37°C	0.8
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):3HIBCH	32°C	0.8

Table 35: Summary table showing the overexpression conditions for each thioesterase gene (4HBT, FcoT, RifR, YciA) in pET20b(+) vector, transformed into *E. coli* BL21 (DE3) pLysS.

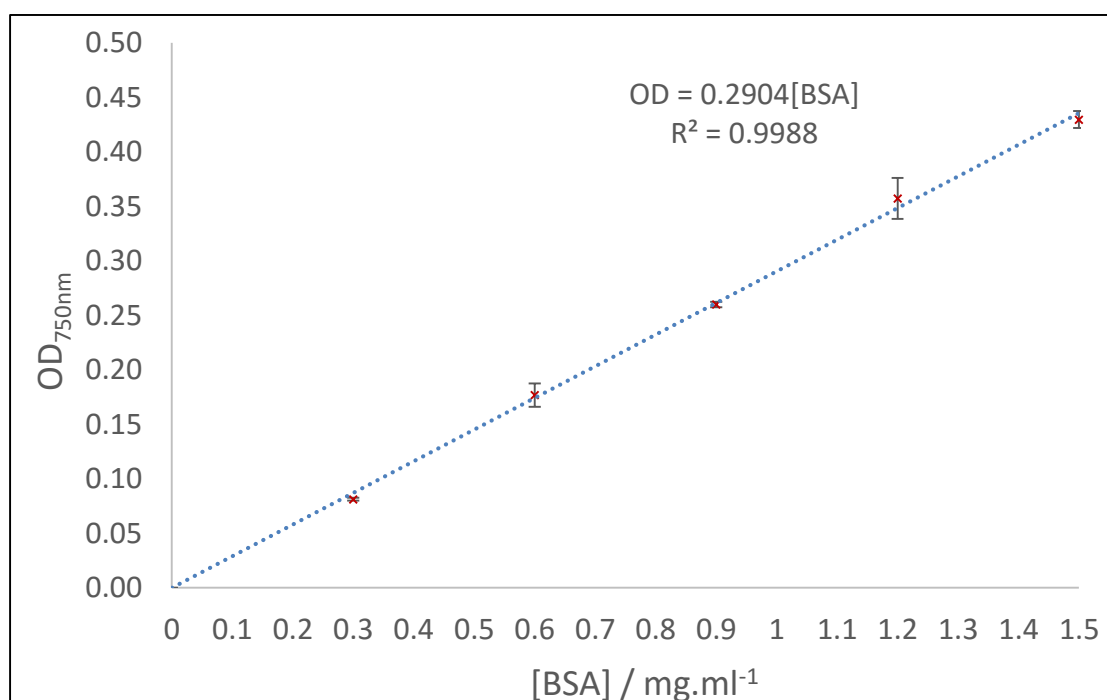


Figure 74: DC assay calibration curve for thioesterase cell free extract concentrations in HEPES buffer using bovine serum albumin as the protein standard

Strain that cell free extract was prepared from	Buffer	Repeat	OD ₇₅₀	Total protein concentration of diluted cell free extract mg.ml ⁻¹	Dilution	Total protein concentration of cell free extract mg.ml ⁻¹	Average total protein concentration mg.ml ⁻¹	Standard Deviation mg.ml ⁻¹
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+)	50 mM HEPES, pH 7.5	1	0.22	0.75	50	38	36	1.2
		2	0.20	0.69	50	35		
		3	0.21	0.72	50	36		
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):FcT	50 mM HEPES, pH 7.5	1	0.30	1.0	20	21	22	1.2
		2	0.30	1.0	20	21		
		3	0.34	1.2	20	23		
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):YciA	50 mM HEPES, pH 7.5	1	0.20	0.67	50	34	36	2.0
		2	0.22	0.75	50	37		
		3	0.22	0.76	50	38		
<i>E. coli</i> BL21 (DE3) pLysS pET20b(+):4HBT	50 mM HEPES, pH 7.5	1	0.32	1.1	50	55	57	1.6
		2	0.33	1.1	50	57		
		3	0.35	1.2	50	60		

Table 36: Table of total protein concentrations for cell free extracts of 4HBT, YciA and FcoT thioesterases, as well as the *E. coli* BL21 (DE3) pLysS pET20b(+) control strain lysate.

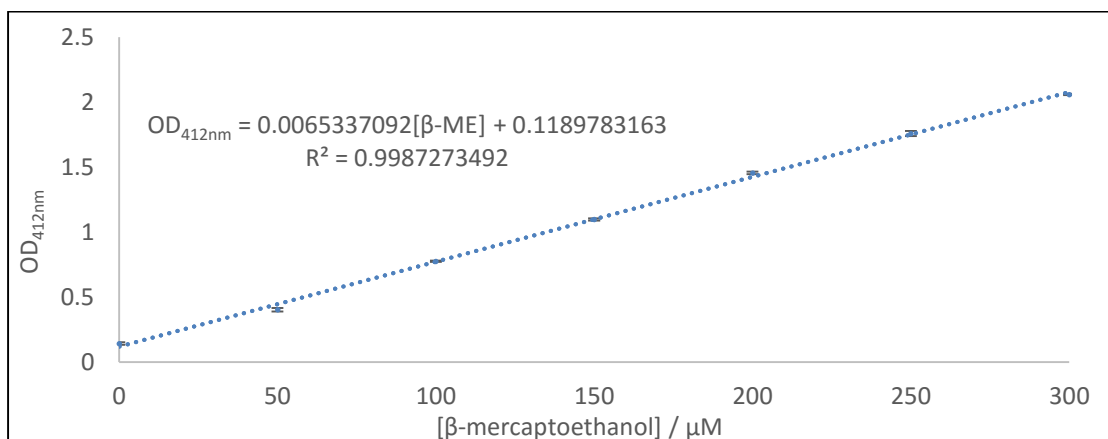


Figure 75: β -mercaptoethanol calibration curve for 500 μM DTNB in HEPES buffer. Average of triplicates with standard deviations.

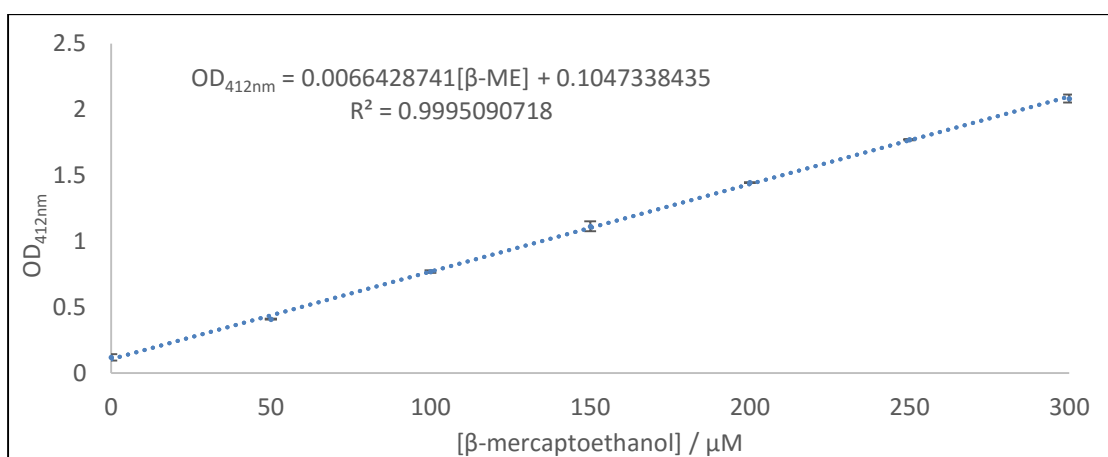


Figure 76: β -mercaptoethanol calibration curve for 500 μM DTNB in phosphate buffer. Average of triplicates with standard deviations.

Raw data for CtHIS-4HBT (75 µg.ml ⁻¹) with methacrylyl-CoA as substrate							Raw data for CtHIS-4HBT (1 mg.ml ⁻¹) with methacrylyl-CoA as substrate						
[MAA-CoA] / mM		Initial reaction rate (V ₀) / µM.min ⁻¹					[IBA-CoA] / mM		Initial reaction rate (V ₀) / µM.min ⁻¹				
(high)	(low)	R1	R2	R3	AVERAGE	S.D.	(high)	(low)	R1	R2	AVERAGE	S.D.	
0.13	0.10	1.6	1.6	1.7	1.6	0.031	0.10	0.080	0.54	0.51	0.53	0.016	
0.26	0.20	3.2	3.1	3.2	3.2	0.045	0.20	0.16	0.92	1.1	0.99	0.065	
0.39	0.30	4.4	4.5	4.5	4.5	0.034	0.30	0.24	1.4	1.5	1.5	0.045	
0.52	0.40	5.7	5.7	5.7	5.7	0.027	0.40	0.32	1.9	2.0	2.0	0.032	
0.65	0.50	6.6	7.0	7.0	6.9	0.17	0.50	0.40	2.3	2.5	2.4	0.052	
[MAA-CoA] ⁻¹ / L.µmol ⁻¹		Inverse Initial reaction rate (V ₀ ⁻¹) / µM.min ⁻¹					[MAA-CoA] ⁻¹ / L.µmol ⁻¹		Initial reaction rate (V ₀ ⁻¹) / µM.min ⁻¹				
(high)	(low)	R1	R2	R3	AVERAGE	S.D.	(high)	(low)	R1	R2	AVERAGE	S.D.	
0.0077	0.010	0.63	0.61	0.60	0.61	0.012	0.010	0.013	1.8	2.0	1.9	0.059	
0.0038	0.0050	0.31	0.32	0.32	0.32	0.0044	0.0050	0.0063	1.1	0.95	1.0	0.067	
0.0026	0.0033	0.23	0.22	0.22	0.22	0.0017	0.0033	0.0042	0.70	0.66	0.68	0.021	
0.0019	0.0025	0.17	0.17	0.18	0.18	0.00083	0.0025	0.0031	0.52	0.50	0.51	0.0082	
0.0015	0.0020	0.15	0.14	0.14	0.15	0.0036	0.0020	0.0025	0.43	0.41	0.42	0.0090	

Table 37: Raw data for Lineweaver Burk plots during kinetic characterisation of CtHIS-4HBT on methacrylyl-CoA and isobutyryl-CoA as substrates.

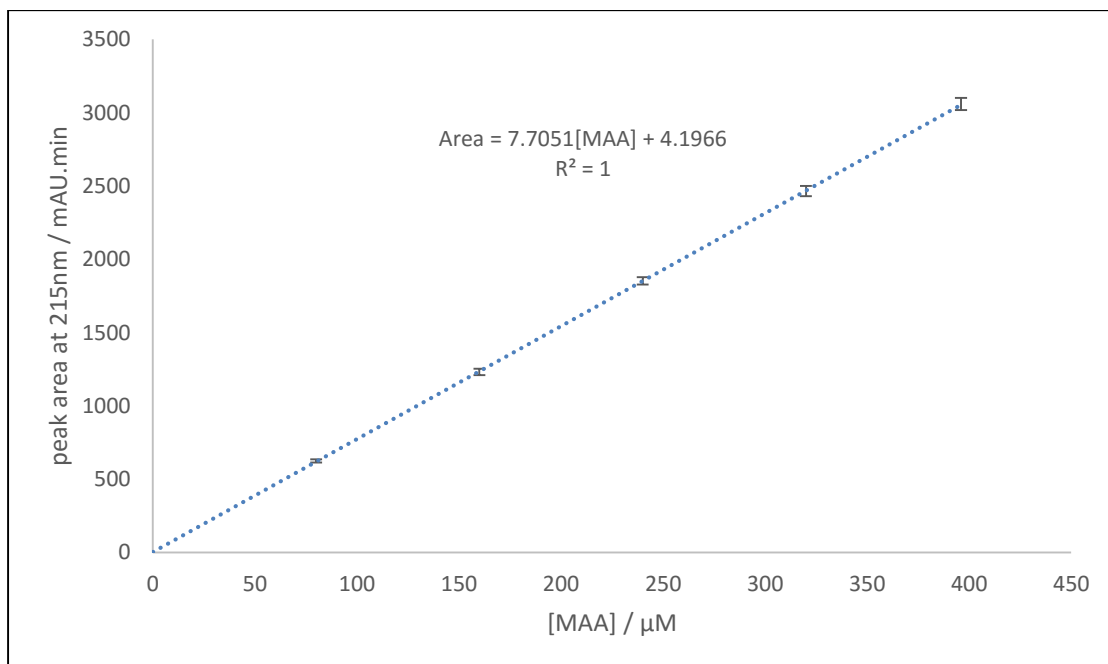


Figure 77: Methacrylic acid calibration curve for production by biotransformation of isobutyric acid.

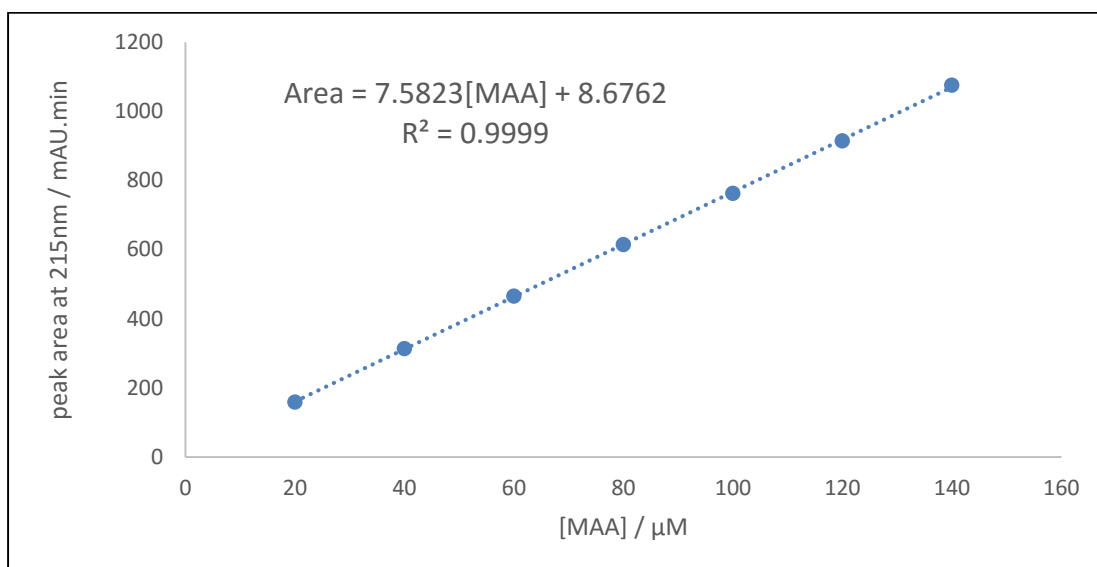


Figure 78: Methacrylic acid calibration curve for production directly from glucose.

8.4 Appendix 4: Derivations of equations

8.4.1 Derivation of Equation 1

Let:

V_{GLYCEROL}	= Volume of glycerol.	(Unit: ml)
V_{TOTAL}	= Final volume, of glycerol and water solution.	(Unit: ml)
V_{WATER}	= Volume of water.	(Unit: ml)
M_{GLYCEROL}	= Known mass of glycerol.	(Unit: g)
ρ_{GLYCEROL}	= Density of glycerol at room temperature.	(Unit: g.ml ⁻¹)

And let:

$$\frac{V_{\text{GLYCEROL}}}{V_{\text{TOTAL}}} = 0.8$$

Where: $V_{\text{TOTAL}} = V_{\text{GLYCEROL}} + V_{\text{WATER}}$

Thus, substituting V_{TOTAL} :

$$\frac{V_{\text{GLYCEROL}}}{V_{\text{GLYCEROL}} + V_{\text{WATER}}} = 0.8$$

Where: $V_{\text{GLYCEROL}} = M_{\text{GLYCEROL}} / \rho_{\text{GLYCEROL}}$

Thus, substituting V_{GLYCEROL} :

$$\frac{\frac{M_{\text{GLYCEROL}}}{\rho_{\text{GLYCEROL}}}}{\frac{M_{\text{GLYCEROL}}}{\rho_{\text{GLYCEROL}}} + V_{\text{WATER}}} = 0.8$$

Rearranging:

$$\frac{M_{\text{GLYCEROL}}}{\rho_{\text{GLYCEROL}}} = 0.8 \left(\frac{M_{\text{GLYCEROL}}}{\rho_{\text{GLYCEROL}}} + V_{\text{WATER}} \right)$$

$$\frac{\frac{M_{\text{GLYCEROL}}}{\rho_{\text{GLYCEROL}}}}{0.8} = \frac{M_{\text{GLYCEROL}}}{\rho_{\text{GLYCEROL}}} + V_{\text{WATER}}$$

$$V_{\text{WATER}} = \frac{\frac{M_{\text{GLYCEROL}}}{\rho_{\text{GLYCEROL}}}}{0.8} - \frac{M_{\text{GLYCEROL}}}{\rho_{\text{GLYCEROL}}}$$

$$V_{\text{WATER}} = 1.25 \left(\frac{M_{\text{GLYCEROL}}}{\rho_{\text{GLYCEROL}}} \right) - \frac{M_{\text{GLYCEROL}}}{\rho_{\text{GLYCEROL}}}$$

$$V_{\text{WATER}} = \frac{M_{\text{GLYCEROL}}}{\rho_{\text{GLYCEROL}}} (1.25 - 1)$$

$$V_{\text{WATER}} = 0.25 \left(\frac{M_{\text{GLYCEROL}}}{\rho_{\text{GLYCEROL}}} \right)$$

Since $\rho_{\text{GLYCEROL}} = 1.26 \text{ g.ml}^{-1}$ at room temperature:

$$V_{\text{WATER}} = \frac{1}{4} \left(\frac{M_{\text{GLYCEROL}}}{1.26} \right)$$

$$V_{\text{WATER}} = \frac{M_{\text{GLYCEROL}}}{5.04}$$

8.4.2 Derivation of Equation 2:

Let:

V_{RP}	= Volume of re-suspended cell pellet	(Unit: ml)
V_{MEDIA}	= Volume of culture media in shake flask prior to addition of resuspended cells.	(Unit: ml)
V_{TOTAL}	= Final volume of inoculated culture media	(Unit: ml)
OD_{TARGET}	= Target OD_{600} of inoculated culture media	(No units)
OD_{RP}	= OD_{600} of resuspended cell pellet.	(No units)

And let:

$$OD_{TARGET} = OD_{RP} \left(\frac{V_{RP}}{V_{TOTAL}} \right)$$

Where: $V_{TOTAL} = V_{MEDIA} + V_{RP}$

Thus, substituting V_{TOTAL} :

$$OD_{TARGET} = OD_{RP} \left(\frac{V_{RP}}{V_{MEDIA} + V_{RP}} \right)$$

Rearrange to make V_{RP} the subject:

$$OD_{TARGET} = OD_{RP} \left(\frac{V_{RP}}{V_{MEDIA} + V_{RP}} \right)$$

$$OD_{TARGET}(V_{MEDIA} + V_{RP}) = OD_{RP} \cdot V_{RP}$$

$$OD_{TARGET} \cdot V_{MEDIA} + OD_{TARGET} \cdot V_{RP} = OD_{RP} \cdot V_{RP}$$

$$\frac{OD_{TARGET} \cdot V_{MEDIA} + OD_{TARGET} \cdot V_{RP}}{V_{RP}} = OD_{RP}$$

$$\frac{OD_{TARGET} \cdot V_{MEDIA}}{V_{RP}} + \frac{OD_{TARGET} \cdot V_{RP}}{V_{RP}} = OD_{RP}$$

$$\frac{OD_{TARGET} \cdot V_{MEDIA}}{V_{RP}} + OD_{TARGET} = OD_{RP}$$

$$\frac{OD_{TARGET} \cdot V_{MEDIA}}{V_{RP}} = OD_{RP} - OD_{TARGET}$$

$$OD_{TARGET} \cdot V_{MEDIA} = V_{RP} (OD_{RP} - OD_{TARGET})$$

$$V_{RP} = \frac{OD_{TARGET} \cdot V_{MEDIA}}{OD_{RP} - OD_{TARGET}}$$