Metabolic control of the ethylmalonyl-CoA pathway in *Methylobacterium extorquens* AM1

by

Nathan M. Good

A dissertation submitted in partial fulfillment of the

requirements for the Degree of

Doctor of Philosophy

University of Washington

2014

Reading committee:
Mary E. Lidstrom (chair)
Carrie S. Harwood
John A. Leigh

Program Authorized to Offer Degree: Microbiology

University of Washington

Abstract

Metabolic control of the ethylmalonyl-CoA pathway in *Methylobacterium extorquens* AM1

Nathan M. Good

Chair of Supervisory Committee:
Professor Mary E. Lidstrom
Microbiology

Methylobacterium extorquens AM1 is a facultative methylotroph capable of growth on reduced carbon compounds with no carbon-carbon bonds, as well as multi-carbon compounds. *M. extorquens* AM1 is widespread in nature and is often associated with the plant phyllosphere, a heterogeneous environment with consistently changing substrate availability. During growth on one-carbon substrates, such as the renewable feedstock methanol, *M. extorquens* AM1 utilizes the ethylmalonyl-CoA pathway that involves several rare four- and five carbon CoA-esters of great biotechnological interest. *M. extorquens* AM1 also incorporates carbon dioxide as half of the fixed carbon during biomass production, making it a desirable organism to develop into a platform for the production of value-added chemicals. A better understanding of the regulation of the ethylmalonyl-CoA pathway is vital to future platform development of *M. extorquens* AM1.

RNA-seq transcriptomics was developed and optimized for *M. extorquens* AM1. This method provided better replication and higher sensitivity than traditional microarrays for investigating gene expression, a primary level of regulation. With this approach validated, the

ethylmalonyl-CoA pathway was investigated for a metabolic control point taking advantage of growth on ethylamine, a two-carbon substrate that is converted to biomass precursors through the ethylmalonyl-CoA pathway. Using a two-carbon substrate successfully bypassed identified regulatory mechanisms that restrict carbon flux into the assimilatory pathways during growth on methanol. The approach used involved challenging *M. extorquens* AM1 to switch from growth on succinate to growth on ethylamine and generating a timecourse of response measured by targeted metabolomics, RNA-seq transcriptomics, and enzyme activities. From these experiments, ethylmalonyl-CoA mutase was identified as a metabolic control point. To confirm the role of this enzyme as a metabolic control point, over-expression of ethylmalonyl-CoA mutase overrode the control and caused a change in growth phenotype. In this thesis, I demonstrated the utility of systems-level approaches to investigate a directed hypothesis and answer fundamental physiological questions, and have identified a candidate for future strain manipulation.

Acknowledgments

I would like to thank my advisor, Dr. Mary Lidstrom. She welcomed me into her lab when we first moved to Seattle and gave me the opportunity to change careers and pursue a life-long passion for science and education. She has taught me how to be a researcher, an educator and a professional. Her understanding of my need to be a scientist and a father will forever be appreciated.

I also want to thank my committee: Dr. Carrie Harwood, Dr. John A. Leigh, Dr. Matthew Parsek and Dr. Robert Morris for all of their insight, comments, criticisms, and interest in my work.

I want to acknowledge past and current members of the Lidstrom lab, without whom my work would not have been possible.

I want to thank my family for their unwavering support. To my parents, Brian and Nanette Good, from the bottom of my heart: you have imparted to me the importance of education since I was very young, inspiring me to always do my best even if it was more difficult. You have taught me to set goals and never give up on achieving them. To my siblings, Danika, Jon, Aaron and Nick, thank you grounding me. Your visits and phone calls have meant more than you can possibly know. To Aaron and Nick, thank you for helping us when we needed it the most.

Finally, I want to dedicate this dissertation to wife, Ceci, and our children, Lilu and Niko. Ceci, none of this would have been possible without you. Your example as a wife, a mother, and a scientist, is inspiring. Your unwavering love, support, and guidance has gotten me where I am today. To Lilu, you were the motivation for all of this. With your first breath you rekindled my

love of learning, as I took in the world anew right by your side. You are my little scientist. To Niko, your laughter has kept me young in heart and mind, and your boundless energy is better than any cup of coffee.

Table of Contents

Abstract iii

Acknowledgments v

Table of Contents vii

List of Tables xi

List of Figures xii

Chapter 1. Introduction 1

1.1. Metabolic flexibility of *Methylobacterium extorquens* AM1 1

1.1.1. Growth on multi-carbon compounds 2

1.1.2. Methylotrophy in M. extorquens AM1 2

1.1.3. Carbon assimilation during Methylotrophic growth 4

1.1.4. Carbon assimilation during growth on C2 compounds 8

1.2. Comparison of C1 and C2 growth in *M. extorquens* AM1 10

1.3. C1 and C2 growth involves high flux through toxic intermediates 10

1.4. Metabolic control points in the assimilatory pathways of *M. extorquens* AM1 11

1.5. Regulation of the EMC pathway 13

1.6. Studying methylotrophy in *M. extorquens* AM1 using systems-level approaches 14

1.7. Dissertation Outline 14

Chapter 2. Implementation of RNA-seq for *M. extorquens* AM1 17

2.1. Abstract 17

2.2. Introduction 17

2.3. Methods 18

2.3.1. Growth conditions 18

2.3.2. Sample collection and processing 18

2.3.3. Sequencing and data processing 19

2.4. Results 19

2.4.1. Greater rRNA depletion results in greater mapping to protein-coding sequences 19

2.4.2. Effects of depletion on protein-coding sequence counts 20

2.4.3. Reproducibility of RNA-seq datasets 22

2.4.4. Comparison of RNA-seq to microarray for succinate vs. methanol growth 25

2.5. Discussion 28

2.6. Acknowledgments 30

Chapter 3. Ethylmalonyl-CoA mutase operates as a metabolic control point in *M. extorquens* AM1 32

3.1. Abstract 32

3.2. Introduction 33

3.3. Materials and methods 38

3.3.1. Materials 38

3.3.2. Bacterial strains and growth conditions 39

3.3.3. Chemostat cultivation 39

3.3.4. LC-MS and GC-MS metabolomics 41

3.3.5. RNA-seq transcriptomics 42

3.3.6. Enzyme assays 42

3.3.7. Calculation of hypothetical minimum enzyme activity 43

3.3.8. Cloning and heterologous expression of ccr, ecm, and epi from M. extorquens AM1 in E. coli 43

3.3.9. Purification of Ccr, Ecm and Epi 45

3.3.10. Generation of expression mutant constructs 45

3.4. Results and discussion 46

3.4.1. Growth response to the switch from succinate to ethylamine 46
3.4.2. Gene expression response of ethylamine oxidation and central metabolic pathways 48
3.4.3. Response of EMC pathway metabolites and glyoxylate 55
3.4.4. Response of key C2 enzymes 61
3.4.5. Over-expression of ecm removes control in the EMC Pathway 66
3.5. Summary 71
3.6. Acknowledgments 71
Chapter 4. Conclusions and future directions 72
4.1. Conclusions 72
4.1.1. RNA-seq for gene expression analysis 73
4.1.2. Metabolic control point in the EMC 74
4.2. Future Directions 74
4.3. Final Comments 75
References 77
Appendix A 83
Introduction 83
Methods 83
Initial Results 84
Discussion 89
Appendix B 92
Introduction 92
Methods 92
Initial Results 93
Discussion 99

List of Tables

Table 2.1 RNA-seq reads mapped for two substrates and three depletion methods.....20

Table 2.2 Depletion effects on read counts of methylotrophy genes.....21

Table 3.1 *Methylobacterium extorquens* strains and plasmids used in this study 40

List of Figures

Figure 1.1 Metabolic modules of methylotroph in *M. extorquens* AM1 3

Figure 1.2 Methylotrophic Metabolism in *M. extorquens* AM1.....5

Figure 1.3 C2 Metabolism in *M. extorquens* AM1 7

Figure 2.1 Replication of RNA-seq expression data for growth on methanol using three depletion methods 23

Figure 2.2 Concordance of succinate vs. methanol growth experiments by RNA-seq 26

Figure 2.3 Comparison of succinate vs. methanol gene expression by RNA-seq and microarray 27

Figure 3.1 Metabolic pathway for two-carbon metabolism in *M. extorquens* AM1 36

Figure 3.2 Growth response of *M. extorquens* AM1 after the switchover from succinate-growth to ethylamine growth. 47

Figure 3.3 Transcriptional profile of EMC pathway genes during the transition from succinate growth to ethylamine growth 51

Figure 3.4 Transcriptional profile of the middle EMC pathway genes during the transition from succinate growth to ethylamine growth 53

Figure 3.5 Metabolite profile of the mid-EMC pathway during the transition from succinate growth to ethylamine growth 58

Figure 3.6 Degradation of ethylmalonyl-CoA to butyryl-CoA by hot water extraction ...60

Figure 3.7 Ecm limits ethylmalonyl-CoA to methylsuccinyl-CoA activity 64

Figure 3.8 Switchover from succinate to ethylamine-growth over-expressing *ecm*.....69

Chapter 1. Introduction

1.1. Metabolic flexibility of *Methylobacterium extorquens* AM1

Methylobacterium extorquens AM1 is an alpha-proteobacterium and a facultative methylotroph capable of utilizing reduced single-carbon (C1) compounds such as methanol, methylamine and formate, and multi-carbon compounds such as ethylamine (C2), pyruvate (C3), and succinate (C4) as a sole source of carbon and energy (1–4). *M. extorquens* AM1 is widespread in nature and is associated with the leaf surfaces of plants where methanol is released in the early morning hours from the leaf stomata (5, 6). Methanol is abundant in high concentrations as a growth substrate only for short durations. Therefore, *M. extorquens* AM1 must acclimate to rapidly changing substrate availability in order to survive and thrive in its natural habitat (7). The biochemical machinery of the metabolic network must quickly reset to accommodate the bursts of methanol while preventing the lethal accumulation of the toxic methanol oxidation product formaldehyde and the downstream toxic metabolites glyoxylate and glycine (8, 9). Responses of this type might be expected to involve establishing new metabolic set-points, such as increased or decreased activities of key pathway enzymes.

Methylotrophy is an interesting mode of growth for study because of high flux through toxic intermediates, and the necessity to balance production and consumption reactions, along with the regulatory mechanisms that govern these processes. The ecology of *M. extorquens* AM1 implies that it is not only capable of acclimatizing to rapidly changing growth substrate availability, but that is it biochemically adapted to do so (7). Methylotrophic metabolism in *M.*

extorquens AM1 also involves production of several compounds of commercial interest, such as methylsuccinyl-CoA, mesaconyl-CoA, succinate, ethylmalonyl-CoA, and methylmalonyl-CoA (10–12).

1.1.1. Growth on multi-carbon compounds

M. extorquens AM1 has been studied for decades as a model organism for the biochemical machinery necessary for methylotrophic metabolism (13). When *M. extorquens* AM1 grows on multi-carbon substrates, it utilizes a typical aerobic, heterotrophic metabolic network including the TCA cycle, NADH dehydrogenase and an electron transport chain (1, 14).

1.1.2. Methylotrophy in M. extorquens AM1

Growth on C1 compounds, however, is very different. The TCA cycle and NADH dehydrogenase are repressed, and over 100 methylotrophy-specific genes are up-regulated (15). Methylotrophic metabolism in *M. extorquens* AM1 can be compartmentalized into several pathways linked to a particular metabolic process that operate as a functional unit, referred to as metabolic modules [Fig. 1.1].

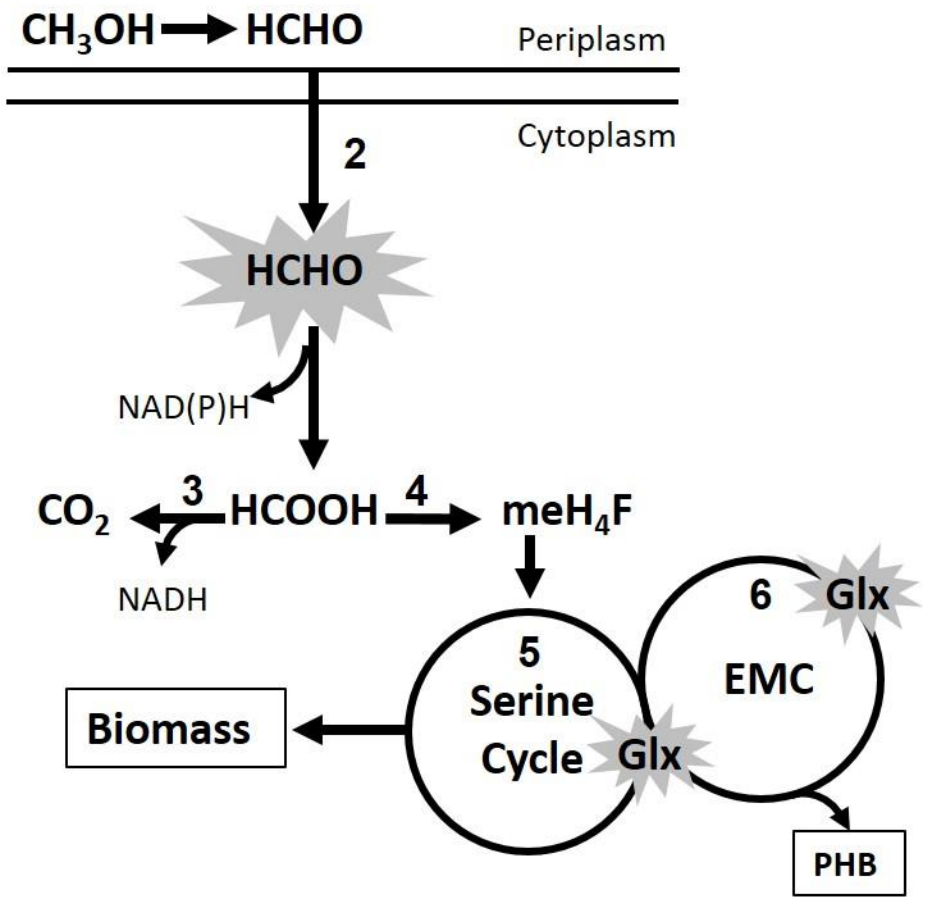

Figure 1.1 Metabolic modules of methylotroph in *M. extorquens* AM1

Methylotrophy metabolic modules: oxidation of methanol to formaldehyde in the periplasm by methanol dehydrogenase (1); oxidation of cytoplasmic formaldehyde to formate, H_4MPT pathway (2); oxidation of formate to CO_2 by formate dehydrogenases (3); conversion of formate to methylene-H_4F, H_4F pathway (4), incorporation of methylene-H_4F into the Serine Cycle (5) and the Ethylmalonyl-CoA (6) assimilatory pathways. H_4MPT, tetrahydromethanopterin; H_4F, tetrahydrofolate; me-H_4F, methylene-tetrahydrofolate; Serine, Serine Cycle; EMC, Ethylmalonyl-CoA pathway.

Methylotrophy-specific modules include: oxidation of methanol to formaldehyde in the periplasm by methanol dehydrogenase; oxidation of formaldehyde to formate in the cytoplasm linked to tetrahydromethanopterin (the H4MPT pathway); oxidation of formate to CO2 by formate dehydrogenases; conversion of formate to methylene tetrahydrofolate (H4F pathway); and incorporation of methylene-H4F into the Serine Cycle and the Ethlymalonyl-CoA pathway (EMC) for assimilation into biomass (13, 16, 17).

1.1.3. Carbon assimilation during Methylotrophic growth

During growth on methanol, C1-units as methylene H4F are condensed with glycine to form serine by serine-hydroxymethyltransferase (GlyA) (18). Serine and glyoxylate are trans-aminated by serine-glyoxylate aminotransferase (Sga), generating hydroxypyruvate and glycine. Hydroxypyruvate is then converted to glycerate via hydroxypyruvate reductase (Hpr) (19, 20). Glycerate is phosphorylated by glycerate kinase (Gck) to 2-phosphoglycerate (2PG) consuming a molecule of adenosine triphosphate (ATP). 2PG can be converted to phosphoenolpyruvate (PEP) by an enolase (Eno). PEP is carboxylated, incorporating a molecule of carbon dioxide (CO2), by PEP carboxylase (Ppc), to form oxaloacetate (OAA). OAA can be drawn off for biosynthesis or reduced to malate. Malate is ligated with CoA to form malyl-CoA by malyl-CoA thiokinase (Mtk), which consumes another molecule of ATP (21). Malyl-CoA is split into one molecule of glyoxylate and one molecule of acetyl-CoA, which feeds into the EMC pathway. Under methylotrophic growth conditions, three- and four-carbon compounds are drawn off for biosynthesis and continued flux through the Serine Cycle requires the addition of a two-carbon intermediate to replenish the glycine pool (22, 23) [Fig. 1.2].

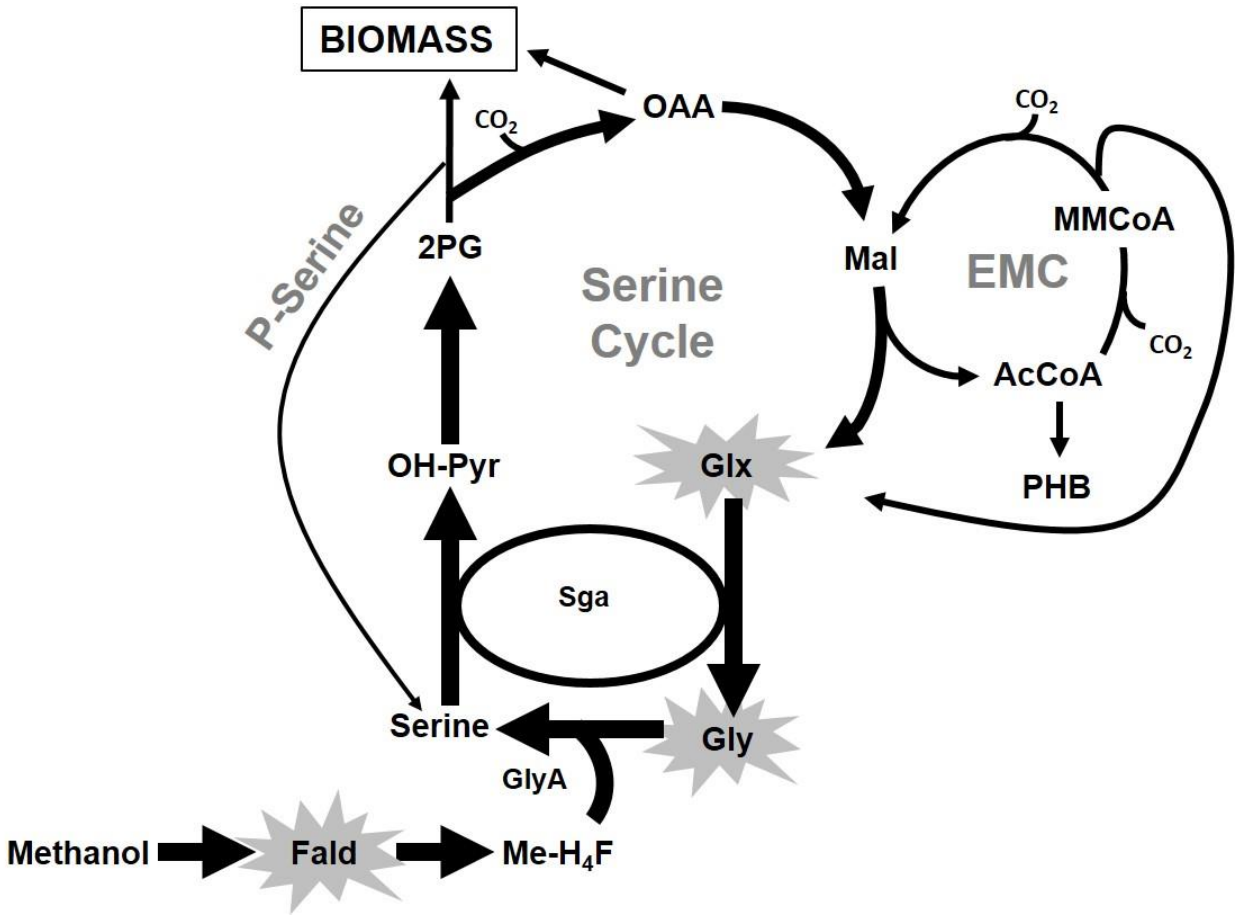

Figure 1.2 Methylotrophic Metabolism in *M. extorquens* AM1

AcCoA, acetyl-CoA; CO2, carbon dioxide; EMC, ethylmalonyl-CoA pathway; Fald, formaldehyde; Gly, glycine; GlyA, serine-hydroxymethyltransferase (SHMT); Glx, glyoxylate; Mal, malate; MMCoA, methylmalyl-CoA; OAA, oxaloacetate; Me-H_4F, methylene-tetrahydrofolate; 2PG, 2-phosphoglycerate; OH-Pyr, hydroxypyruvate; PHB, poly-β-hydroxybutyrate; P-Serine, phosphoserine pathway; Sga, serine-glyoxylate aminotransferase. Relative carbon fluxes are depicted by arrow width.

Glyoxylate is regenerated by the EMC pathway to support production of biomass precursors via the serine cycle [Fig. 1.3] (24). The EMC performs this anapleurosis to replenish the serine cycle with glyoxylate, which is converted to glycine, the C1-unit (methylene-H4F) acceptor molecule (22, 23). The pathway is partly the amalgamation of familiar metabolic steps from multiple standard pathways including the TCA cycle, the poly-ß-hydroxybutyrate (PHB) cycle, and fatty acid synthesis. In addition, the EMC pathway also utilizes three enzymes unique to the pathway. The first reaction of the EMC pathway effects the conversion of two molecules of acetyl-CoA to acetoacetyl-CoA via β-ketothiolase (PhaA), condensing two two-carbon molecules into a four-carbon intermediate. Then acetoacetyl-CoA is reduced to ß-hydroxybutyryl-CoA (BHB-CoA) via acetoacetyl-CoA reductase (PhaB) consuming a molecule of NADPH. Both PhaA and PhaB are enzymes from fatty acid synthesis. ß-hydroxybutyryl-CoA (BHB-CoA) is the branch point between PHB synthesis and the remaining enzymatic steps of the EMC pathway. BHB-CoA is converted to crotonyl-CoA by a crotonase (CroR), which is then reduced and carboxylated to (2S)-ethylmalonyl-CoA, the namesake of the pathway, by NADPH-dependent crotonyl-CoA reductase/carboxylase (Ccr) (22). (2S)-ethylmalonyl-CoA is rearranged to (2R)-ethylmalonyl-CoA by ethylmalonyl-CoA/methylmalonyl-CoA epimerase (Epi). This enzyme performs a similar reaction later in the pathway. (2R)-ethylmalonyl-CoA is rearranged to form (2S)-methylsuccinyl-CoA via ethylmalonyl-CoA mutase (Ecm).

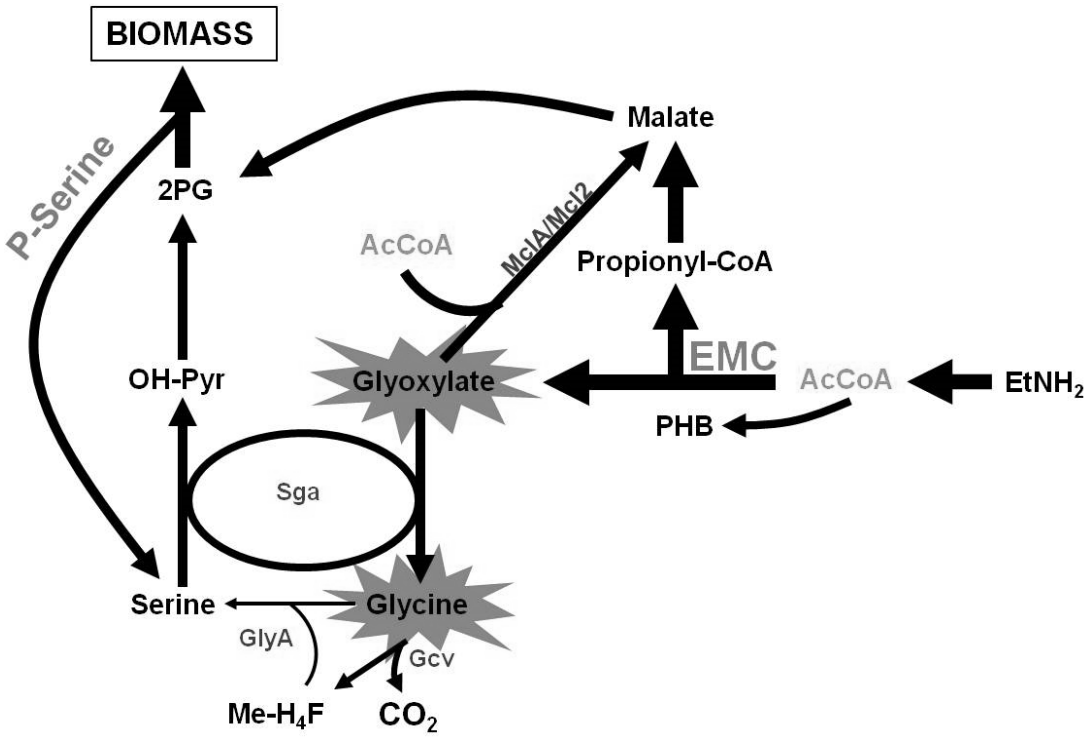

Figure 1.3 C2 Metabolism in *M. extorquens* AM1

AcCoA, acetyl-CoA; CO_2, carbon dioxide; EMC, ethylmalonyl-CoA pathway; $EtNH_2$, ethylamine; Gcv, glycine cleavage system; GlyA, serine-hydroxymethyltransferase (SHMT); Glx, glyoxylate; Mal, malate; MclA/Mcl2, malate synthase via malyl-CoA lyase and malyl-CoA thioesterase; MMCoA, methylmalyl-CoA; OAA, oxaloacetate; Me-H_4F, methylene-tetrahydrofolate; 2PG, 2-phosphoglycerate; OH-Pyr, hydroxypyruvate; PHB, poly-β-hydroxybutyrate; P-Serine, phosphoserine pathway; Sga, serine-glyoxylate aminotransferase. Relative carbon fluxes are depicted by arrow width.

Ecm is one of two cobalt-dependent mutases in the pathway. (2S)-methylsuccinyl-CoA is oxidized to mesaconyl-CoA by methylsuccinyl-CoA dehydrogenase (Msd) (25). Mesaconyl-CoA is converted to ß-methylmalyl-CoA by mesaconyl-CoA hydratase (Mcd) (26).

At a second branch point in the pathway, ß-methylmalyl-CoA is split into propionyl-CoA and one molecule of glyoxylate by a lyase (MclA). Glyoxylate is converted to hydroxypyruvate with glycine and serine as intermediates. Propionyl-CoA is carboxylated to methylmalonyl-CoA by propionyl-CoA carboxylase (PccAB). Methylmalonyl-CoA is converted to malate by succinyl-CoA transferase (Sct), succinate dehydrogenase (Sdh) and fumarase (Fum), enzymes shared with the TCA cycle. Malate can either be drawn off for biosynthesis or converted to a second molecule of glyoxylate, which replenishes the serine cycle with the two-carbon intermediate needed to keep the cycle operational.

1.1.4. Carbon assimilation during growth on C2 compounds

M. extorquens AM1 is capable of growth on C2 compounds such as ethylamine, acetate and ethanol (27). In our hands, *M. extorquens* AM1 grows the most robustly on the C2 substrate ethylamine. During growth on this substrate ethylamine is first oxidized to acetaldehyde by methylamine dehydrogenase (2, 18). Acetaldehyde is oxidized to acetate by an aldehyde dehydrogenase or an aldehyde oxidase, or both. Acetate is activated to acetyl-CoA by acetyl-CoA synthase (28). Genes encoding the enzymes for both the acetaldehyde oxidation and acetate activation steps have not been confirmed by phenotypic analysis, but RNA expression data have shown that genes encoding putative enzymes with these functions are up-regulated during growth on ethylamine (28). Oxidation of acetyl-CoA for energy and reducing power occurs via

the TCA cycle. Assimilation occurs through the EMC pathway, into which acetyl-CoA can enter directly (28, 29). The EMC operates as a linear pathway under these growth conditions, producing one molecule of glyoxylate and one molecule of malate, both of which are then converted to intermediates for biosynthesis [Fig. 1.3]. However, unlike the glyoxylate produced from methylotrophic growth, which is converted to hydroxypyruvate with glycine and serine as intermediates, glyoxylate produced during C2 growth has several fates. *in vitro* assays have shown that during growth on C2 compounds glyoxylate is converted to malate by an apparent malate synthase activity derived from malyl-CoA/ß-methylmalyl-CoA lyase (MclA) and malyl-CoA thioesterase (Mcl2) (30). Malate can be converted to 2PG and pyruvate by C4-C3 interconverting enzymes which drive carbon flow opposite to the direction of the serine cycle. PEP carboxykinase decarboxylates OAA to generate PEP. NADP+-dependent malic enzyme can decarboxylate OAA to pyruvate. *In vitro* activities suggest this route should be capable of carrying the full glyoxylate flux for *M. extorquens* AM1, but other *in vivo* experiments indicate that the malate synthase activity must be insufficient to carry the full flux generating a bottleneck for carbon assimilation (28, 31). ^{13}C-labeling studies have shown that glyoxylate is also converted to C3 compounds via serine. This conversion involves a coupled glyoxylate/glycine transamination step that requires one serine for every glyoxylate [Fig. 1.2]. It also requires methylene H_4F, which is generated from glycine, reducing the amount of glycine available to generate the serine amino donor. Production of serine occurs either by the phosphoserine pathway (~75%) or by the glycine cleavage system (Gcv) (~25%) (29). In the latter scenario Gcv splits glycine into methylene-H_4F and CO_2, then GlyA condenses methylene-H_4F with another molecule of glycine to form serine (28). Unlike during methylotrophic growth, the EMC pathway operates as a linear assimilation pathway during growth on C2 compounds.

1.2. Comparison of C1 and C2 growth in *M. extorquens* AM1

Differences in the metabolic networks for C1 and C2 growth result in measurable physiological differences. During growth on methanol, carbon flux is split at formate between oxidation for energy production by formate dehydrogenases, of which four have been reported (32–34), and assimilation via the H_4F-pathway to generate me-H_4F for incorporation in to the serine cycle. Approximately 15% of methanol-derived carbon flows into the serine cycle (14, 35). One-third of the assimilatory flux is carried by the EMC pathway to regenerate glyoxylate. In batch culture with minimal media and methanol, *M. extorquens* AM1 grows with a doubling time of approximately 6 hours.

In contrast, during growth on ethylamine, all of the assimilatory flux is carried by the ethylmalonyl-CoA pathway, while oxidation of acetyl-CoA occurs via the TCA cycle. Approximately 35% of the C2-derived carbon goes to the EMC pathway for assimilation (29) However, the assimilatory flux in this condition is about 60% of the assimilatory flux through EMC pathway during growth on methanol. The assimilatory flux during growth on ethylamine can support a doubling time of approximately 12 hours (28).

1.3. C1 and C2 growth involves high flux through toxic intermediates

Both C1 and C2 growth involve high flux through toxic intermediates. During growth on methanol, the cell must maintain metabolic homeostasis preventing an imbalance of production and consumption of formaldehyde, glyoxylate and glycine. Formaldehyde and glyoxylate are highly reactive aldehydes, and intracellular pools must be maintained below millimolar concentrations (9, 36). The precise nature of the toxicity of glycine is unknown, but the amino acid has been reported to chelate magnesium, manganese and copper ions (37, 38). Metabolism

of C2 substrates, like ethylamine, does not involve production of formaldehyde. Rather, glyoxylate is the primary toxic intermediate involved, since nearly half (44%) of the assimilatory flux flows through it (29). Glyoxylate has been shown to inhibit growth above millimolar concentrations (27). Glycine has been reported to inhibit growth at a concentration of 2 mM and arrest growth at 5 mM (28). *M. extorquens* AM1 must precisely balance production and consumption of these toxic intermediates in order to maintain metabolic balance.

1.4. Metabolic control points in the assimilatory pathways of *M. extorquens* AM1

When exposed to an environmental perturbation, such as a sudden switch in growth substrate, *M. extorquens* AM1 must re-establish metabolic balance to prevent lethal accumulation of a toxic intermediate and to resume growth. Recently, an experiment was reported transitioning the cell from succinate growth to methanol growth, during which enzyme activities, transcripts, metabolites and nucleotides were measured in a time course. The analysis of this dataset generated several hypotheses regarding specific strategies used by the cell to acclimate to a new growth substrate (39). The cell passed through a transient phase of metabolic imbalance, then the metabolic network acclimated, resetting to a new metabolic set point. After the substrate switch, carbon flowed through the formaldehyde and formate oxidation modules; however, flux was blocked at the assimilation step for C1 units (methylene-H_4MPT dehydrogenase, MtdA), resulting in a cessation of growth. When methanol was added to the culture *mtdA* transcripts increased immediately, reaching a 4-fold increase after ten minutes, and continuing to increase reaching ~15 fold after 1 hour. MtdA activity, however, did not increase until between 1 to 2 hours after the addition of methanol, suggesting a post-translational mechanism regulating the C1 entry point. It is now known that MtdA is regulated by competitive

inhibition by methylene-H_4F (40). Like *mtdA* transcripts, serine cycle and EMC pathway genes were up-regulated; however, enzyme activities also increased, priming the assimilatory pathways for high flux. While the restriction of carbon flow into the assimilatory pathways was active, β-hydroxybutyryl-CoA (BHB-CoA) and ethylmalonyl-CoA accumulated (4-fold and 3-fold respectively), while downstream metabolites did not accumulate. Once the block at MtdA was removed, carbon flowed into the assimilatory network and growth resumed. The majority of EMC pathway metabolites accumulated, consistent with higher flux through the pathway. Glyoxylate and glycine levels did not increase relative to initial measurements, suggesting that a regulatory mechanism could function to balance production and consumption of these toxic intermediates. The changes in relative metabolite concentrations during the time course suggest a probable metabolic control in the middle of the EMC pathway, which could operate to prevent glyoxylate and/or glycine from accumulating to toxic concentrations.

Dynamic labeling flux studies have shown that the EMC pathway is utilized by *M. extorquens* AM1 for assimilation during growth on the C2 substrate acetate (29). Comparison of flux measurements for the EMC pathway to *in vitro* enzyme activities (41) identified ethylmalonyl-CoA mutase (Ecm) as the only enzyme with an *in vitro* activity that was not significantly greater than the calculated flux [see Fig. 3.1]. This finding suggests a hypothesis that Ecm could serve as a metabolic control point. However, because three of the four enzymes responsible for converting (2S)-ethylmalonyl-CoA to β-methylmalyl-CoA are so-called "reversible" enzymes, a single flux value was given to this four-step reaction sequence (29, 35). Furthermore, *in vitro* activity of ethylmalonyl-CoA/methylmalonyl-CoA epimerase in cell-free extracts had not been determined until the studies reported here in Chapter three. Therefore,

other enzymes in this portion of the EMC pathway could not be ruled out as potential metabolic control points.

1.5. Regulation of the EMC pathway

Relatively little is known about how the EMC pathway enzymes are regulated, and how carbon flow is controlled through the pathway. The genes encoding EMC pathway enzymes are located at several different loci in the *M. extorquens* AM1 genome (42). This is unsurprising, as the pathway is comprised of portions of several other pathways. Nonetheless, it makes the study of regulation of the EMC pathway more challenging than if these enzymes were located in a single operon, as is true of the serine cycle. Only *ccr* and *ecm* are located near one another (within 2.1 kb), but they are encoded on opposite strands. Upstream of *ecm* is a TetR-family regulator that was recently reported as the only known transcriptional regulator of the EMC pathway (43). This regulator was named CcrR because of its positive regulatory interaction with *ccr*. Promoter-fusion assays demonstrated that CcrR is required for full expression of the promoter of *ccr* but is not required for expression of the promoter region of *ecm*. Another putative transcriptional regulator, of the ArsR-family of transcriptional regulators, is located 4 kb downstream of *ecm*. However, a mutant of the ArsR-family regulator was also tested by promoter-fusion analysis and the results indicated that it did not regulate *ccr* or *ecm* (43). The promoter region of ethylmalonyl-CoA/methylmalonyl-CoA epimerase (*epi*) was not tested in this study.

To this date, no universal EMC pathway regulator has been discovered. The serine cycle regulator, QscR (Quayle serine cycle regulator), a LysR-family regulator, positively regulates three key operons for serine cycle function with the co-inducer formyl-H_4F (44, 45). A *qscR* null

mutant does not grow on methanol and has a severely reduced growth rate on ethylamine. The growth defect on ethylamine could be due to lack of induction of the serine cycle genes necessary for C2 growth, or reduced induction of the EMC pathway, or both. Thus far, no regulatory role of the EMC pathway by QscR has been identified.

1.6. Studying methylotrophy in *M. extorquens* AM1 using systems-level approaches

Systems-level approaches are often thought of as methods for hypothesis generation; however, with careful experimental design and a clear hypothesis these tools can be used to answer fundamental scientific questions. The modularity of metabolism in *M. extorquens* AM1, and its metabolic flexibility, allow for the study of substrate-specific metabolic networks, and their regulation. Systems-level approaches provide large datasets often encompassing multiple metabolic modules. These datasets can provide clear information as to how individual modules and metabolic modes interact. Metabolomics can demonstrate the induction of an entire pathway or restriction of carbon through a pathway (24, 46). Transcriptomics can provide detailed information about the up- or down-regulation of entire metabolic modules (15, 39). These approaches are limited, however, in their capacity to validate metabolic details. Nevertheless, the effectiveness of systems-level approaches to study methylotrophy and fundamental metabolic questions has been demonstrated using *M. extorquens* AM1.

1.7. Dissertation Outline

The objective of this dissertation was to investigate the ethylmalonyl-CoA pathway for a metabolic control point. Previous work had suggested that a set of reactions in the EMC pathway were likely to constitute a control point. By transitioning *M. extorquens* AM1 from growth on

limiting succinate to ethylamine, and using several systems-level approaches, I identified one of these enzymes, ethylmalonyl-CoA mutase (Ecm), as a control point. Over-expression of *ecm* successfully overrode the control. The over-expression strain did not accumulate ethylmalonyl-CoA to the same relative concentrations as the control strains, and it was able to acclimate to the substrate switch more quickly.

Chapter 2 reports the first implementation of RNA-seq transcriptomics for *M. extorquens* AM1. We investigated reproducibility, rRNA depletion, and identified biases in read counts. We validated RNA-seq by comparing our results to reported microarray data, updating the toolkit available for investigating *M. extorquens* AM1.

Chapter 3 describes the investigation of the EMC pathway for a metabolic control point using a substrate switchover from succinate growth to ethylamine growth. Systems-level approaches, including targeted metabolomics, RNA-seq transcriptomics, and enzymatic assays, were used to assess the metabolic network. Comparison of data from these three levels, analyzed within the context of the growth response of the culture, allowed for the identification of potential control points. Metabolomics data signaled potential bottlenecks in the pathway leading to accumulation of a metabolite. RNA-seq transcript levels were compared to enzyme activities for control targets to determine the possible level of regulation, either transcriptional or post-transcriptional. Through these studies we identified Ecm, ethylmalonyl-CoA mutase, as a metabolic control point in the EMC pathway. We also partially characterized the control of Ecm as likely occurring at the transcriptional level. Removal of the control allowed the strain to grow more rapidly after a substrate switchover.

Earlier growth suffered a metabolic cost, however, as growth was severely diminished initially. After a period of slow growth the culture shifted to a faster growth rate than wild-type and the control strain, indicating a further metabolic acclimation.

Chapter 4 summarizes the completed work in this dissertation and highlights outstanding questions that could be addressed in future work.

Appendix A presents the metabolite analysis of the substrate switchover experiment with the strain over-expressing *ecm*.

Appendix B is a report of preliminary work studying the regulatory role of ethylmalonyl-CoA epimerase in the EMC pathway. Studies investigated the effect on growth of over-expressing *epi* during C1 and C2 metabolism, and the interaction between CcrR and the *epi* promoter.

Chapter 2. Implementation of RNA-seq for *M. extorquens* AM1

2.1. Abstract

Global gene expression datasets are an important tool for studying basic physiology and metabolism. RNA-seq is a powerful method which utilizes next-generation sequencing to provide a read count for expressed genes, rather than a fluorescence signal as in traditional microarrays. Several studies have described the many benefits of RNA-seq over microarrays for studying global gene expression. RNA-seq was tested and validated for *M. extorquens* AM1. Three ribosomal RNA depletion methods were tested, and inherent biases for each method were investigated. Comparison of RNA-seq expression values to microarray expression values in a succinate-growth vs. methanol-growth differential expression experiment demonstrated that RNA-seq is a viable replacement for the study of global gene expression in *M. extorquens* AM1.

2.2. Introduction

As noted in Chapter 1, *M. extorquens* AM1 is one of the best understood methylotrophs and is a model of study for prokaryotic methylotrophy. A variety of tools are available for the study of *M. extorquens* AM1 including a genome sequence, allelic exchange and expression plasmids, optimized methods for metabolite extraction and quantification, and DNA microarrays for transcriptome analysis (15, 42, 46–48). With the advent of high-throughput sequencing, it has become possible to sequence the entire transcriptome using RNA-seq (49). RNA-seq has several

advantages over traditional DNA microarrays, and we tested the method to update the current tools available to study methylotrophs (49, 50). We investigated reproducibility of RNA-seq methods and concordance with existing microarray data. We also investigated ribosomal RNA depletion methods and report expression biases. Since many of the genes involved in succinate and methanol growth are known, we chose to use these growth conditions for our study. This investigation is the first report of RNA-seq technology used for *M. extorquens* AM1 and provides a powerful tool for in depth transcriptome analysis for this model methylotroph, and a reference for studying gene expression in other methylotrophs.

2.3. Methods

2.3.1. Growth conditions

Cultures of wild-type *M. extorquens* AM1 were grown as reported (16) with succinate or methanol added as the growth substrate.

2.3.2. Sample collection and processing

Cultures were grown to an OD_{600} of 0.6-0.8 and samples were taken. Sample harvesting, RNA isolation, and chromosomal DNA removal were performed as reported (15).

RNA purity was analyzed by rtPCR using iScript (Bio-Rad) reverse transcriptase. RNA integrity was verified by 2100 BioAnalyzer (Agilent). MICROBExpress (Life Technologies, NY, USA) rRNA depletion was performed by myself as recommended by the manufacturer. Ribo-Zero (Epicentre, WI, USA) rRNA depletion, library prep and Illumina Hi-Seq sequencing

were carried out by the UW High Throughput Genomics Center (HTGC) according to the manufacturer's instructions. For each sample condition, two biological replicates were used.

2.3.3. Sequencing and data processing

Each replicate was sequenced to a target depth of 20 million single-ended 36 bp reads using a multiplex strategy on an Illumina Hi-Seq 2000 (Illumina, San Diego, CA, USA). Quality filtering of the raw read counts was performed using Illumina's CASAVA. The filtered data were processed as follows. The raw reads were aligned to the *M. extorquens* AM1 [MAGE, March 2nd, 2012] genome using BWA version 0.7.4-r385 with default parameters (51). The alignments were post-processed into sorted BAM files with SAMTools version 0.1.19-44428cd (52). Reads were attributed to open reading frames (ORFs) using the htseq-count tool from the 'HTSeq' framework version 0.5.4p5 in the 'intersection-nonempty' mode (53). The final abundance was measured by reads per kilobase per million mapped reads (RPKM) (50). T-tests were used to assign significance with post-hoc adjustment by the method of Storey (54).

2.4. Results

2.4.1. Greater rRNA depletion results in greater mapping to protein-coding sequences

The *M. extorquens* AM1 chromosome has 5345 ORFs in the primary chromosome, 1216 ORFs on a mega-plasmid and another 125 ORFs on three plasmids (42). As mentioned, samples were either untreated to deplete rRNA, or treated by commercially available methods. Samples were sequenced to a depth of 100-210 million reads. In all samples, over 97.9-99.4% of the reads

were mapped (Table 2.1). The total number of reads mapped was higher for the non-depleted samples (ND); however 97% of these reads mapped to rRNA genes. In the samples that were not depleted, only 2% of mapped reads were to protein-coding sequences. For MICROBExpress-treated samples (ME) the percentage mapped to protein-coding sequences increased to 7.7-11.2%. Samples that were treated by Ribo-Zero (RZ) were the most depleted, with > 84% of reads mapping to protein-coding sequences. There was also a strong positive correlation between the amount of depletion and the reads mapped per kilobase CDS across the samples.

Table 2.1 RNA-seq reads mapped for two substrates and three depletion methods

			fraction of reads mapped			
sample	**depletion**	**total reads**	**total**	**rRNA**	**CDS†**	**reads/kb to CDS**
me1	MICROBExpress	169212783	0.9941	0.9001	0.0773	2787.2
me2	MICROBExpress	209592329	0.9946	0.8896	0.0834	3788.3
me3	Ribo-Zero	101946127	0.9872	0.0029	0.8487	19653.9
me4	Ribo-Zero	166290203	0.9792	0.0034	0.8443	31844.5
me5	none	199043434	0.9916	0.9731	0.0214	924.6
me6	none	194779447	0.9905	0.9719	0.0225	948.6
su1	MICROBExpress	210731305	0.9931	0.8559	0.1120	4812.2
su2	MICROBExpress	200958802	0.9937	0.8983	0.0777	3188.7

†protein-coding sequences

2.4.2. Effects of depletion on protein-coding sequence counts

Read counts were compared to determine the effects of rRNA depletion on reads of individual genes. 4-6 protein-coding sequences (CDS) were identified in the no depletion datasets that had no reads, whereas these CDS did have a small number of reads in the depleted datasets. All these CDS were for conserved proteins of unknown function found on the mega-plasmid (data not shown). Next, CDS with significant differences in read counts between no

depletion and depletion datasets were identified. A difference of 2-fold or more was considered to be significant. For the MICROBExpress-depleted datasets, 15% of genes had read counts that were significantly different. For the Ribo-Zero-depleted datasets 1.4% of genes had read counts that differed by 2-fold or more from the non-depleted samples (data not shown). From these subsets, CDS known to be involved in methylotrophy were targeted and compared across the datasets (Table 2.2). Read counts of eight genes involved in methylotrophy were found to be significantly different. Of these, six had significantly reduced read counts in the Ribo-Zero depleted dataset. One had significantly reduced reads in the MICROBExpress-depleted dataset. Read counts for the small subunit of methanol dehydrogenase (*mxaI*) increased with deeper rRNA depletion.

Table 2.2 Depletion effects on read counts of methylotrophy genes

locus tag	gene	meND*	meME†	meRZ‡	ratio ND/ME	ratio ND/RZ
META1_1751	pqqA	47970.4	29578.5	8208.5	1.6	5.8
META1_0178	ccr	1959.2	2142.3	564.2	0.9	3.5
META1_1733	mclA1	5143.8	2927.6	1464.0	1.8	3.5
META1_4538	mxaF	31141.9	31736.9	10746.7	1.0	2.9
META1_2772	mauD	237.3	168.8	91.6	1.4	2.6
META1_2771	mauE	150.3	102.3	63.3	1.5	2.4
META1_4295	mclA2	377.5	345.9	184.5	1.1	2.0
META1_2093	fdh4B	124.4	51.0	198.9	2.4	0.6
META1_4535	mxaI	31404.3	30518.4	39545.2	1.0	0.8

† depleted by MICROBExpress, ‡ depleted by Ribo-Zero, * not depleted

2.4.3. Reproducibility of RNA-seq datasets

We compared the concordance of replicate datasets for methanol grown cultures treated by each depletion method [Fig. 2.1]. Non-depleted samples had a correlation coefficient of r^2 = 0.990. MICROBExpress-treated samples had a correlation coefficient of r^2 = 0.991). Ribo-zero-depleted samples had a correlation coefficient of r^2 = 0.988). The non-depleted samples showed the highest variance between the two replicates across expression levels. Ribo-Zero-treated samples had the highest correlation at low expression levels (<100 reads), but showed discordance at moderate expression levels (100-1000 reads). MICROBExpress-treated samples had the highest correlation overall, at all expression levels.

The most differentially expressed genes between replicates were analyzed for the non-depleted and Ribo-zero-treated samples. The majority of the genes that were significantly differentially expressed between replicates (> 2-fold difference in read count) were conserved proteins of unknown function. Three annotated genes identified as significantly different were: *cycA*, encoding a cytochrome c2 precursor (2.9 and 3.2 fold difference for ND and RZ-treated samples respectively), *katA* encoding a catalase (2.3 and 2.0 fold difference for ND and RZ treated samples, respectively), and *fdh2B* encoding a subunit of formate dehydrogenase (2.2 and 2.1 fold difference for ND and RZ treated samples, respectively). ME treated replicates showed a very high correlation for these genes (data not shown).

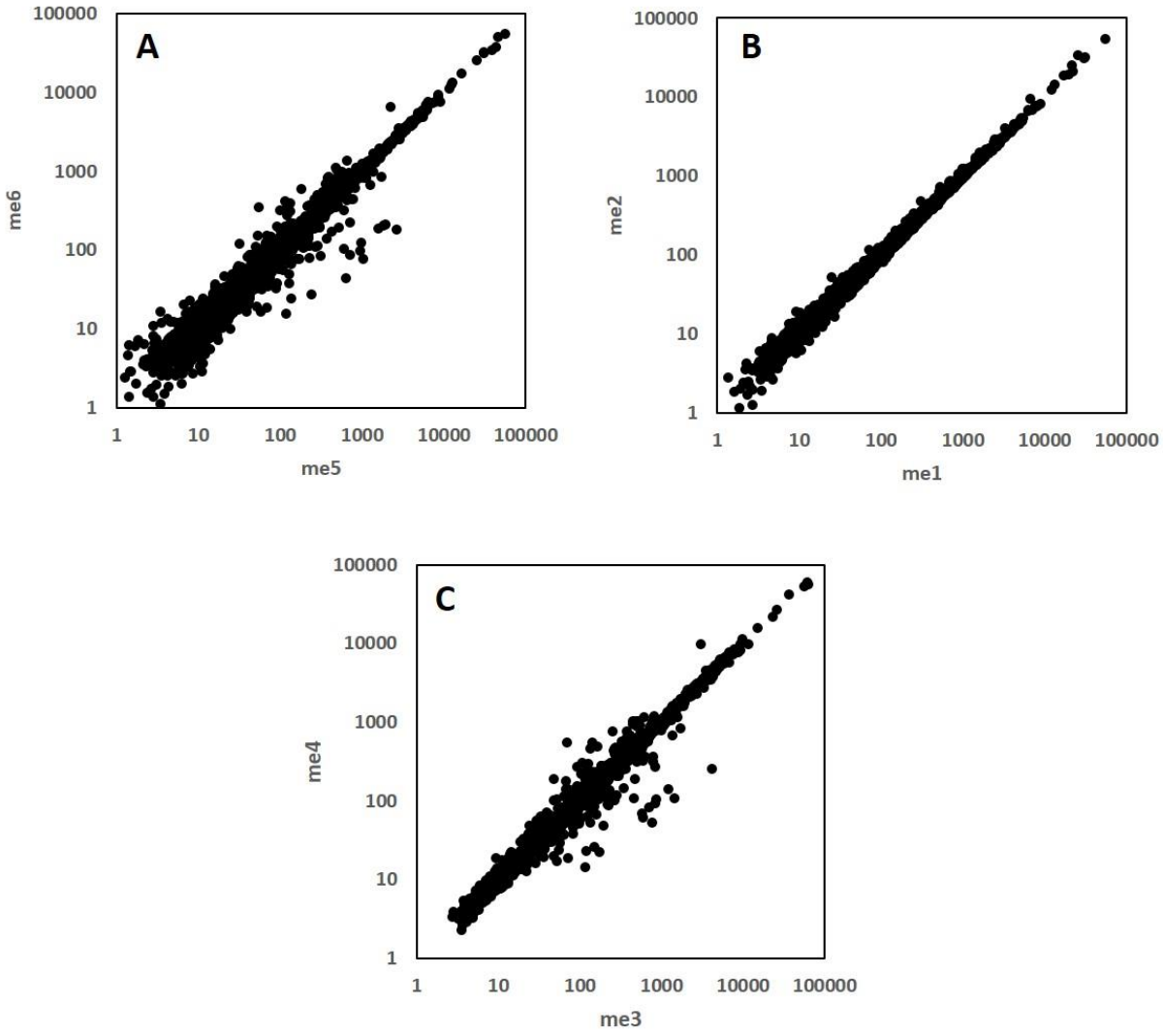

Figure 2.1 Replication of RNA-seq expression data for growth on methanol using three depletion methods

Scatter plot of replication of gene expression during growth on methanol using RNA-seq. Plot points are the log2 RPKM values for each ORF. me1, methanol-grown, MICROBExpress depletion, replicate 1; me2, methanol-grown, MICROBExpress depletion, replicate 2; me3, methanol-grown Ribo-Zero depletion, replicate 1; me4, methanol-grown Ribo-Zero depletion, replicate 2; me5, methanol-grown, no depletion, replicate 1; me6, methanol-grown, no depletion, replicate 2.

2.4.4. Comparison of RNA-seq to microarray for succinate vs. methanol growth

Next, we compared expression levels reported for microarrays investigating succinate versus methanol growth to those generated from the RNA-seq experiments (15). For this comparison we chose to use ME treated samples for the RNA-seq data sets because of the high correlation between them for both methanol growth and succinate growth ($r^2 = 0.999$, succinate growth). Replicate experiments comparing the shift in gene expression of succinate growth against methanol growth were very similar ($r^2 = 0.906$) [Fig. 2.2]. RNA-seq results were compared to microarray results by plotting only genes with a change in expression of 2-fold or greater that was statistically significant ($p < 0.05$) to reduce noise, resulting in a correlation coefficient of $r^2 = 0.833$ [Fig. 2.3]. Clusters of genes that were significantly different in expression between the two methods were investigated. The gene encoding the C subunit of fumarase, *fumC*, was up-regulated 9-fold in the array datasets, but only 2 -fold in the RNA-seq datasets. Expression of *katA*, encoding a catalase, was reversed in the two experiments, with a 4-fold up-regulation in the microarrays compared to a 3-fold down-regulation in the RNA-seq experiments. Two other clusters of a combined fourteen genes had opposed expression values. These clusters were comprised of conserved proteins of unknown function, a putative peptide-synthetase, a putative sideraphore synthetase, a putative acetyltransferase, cobalamin synthase (*cobV*), and a putative transporter (*ygfu*).

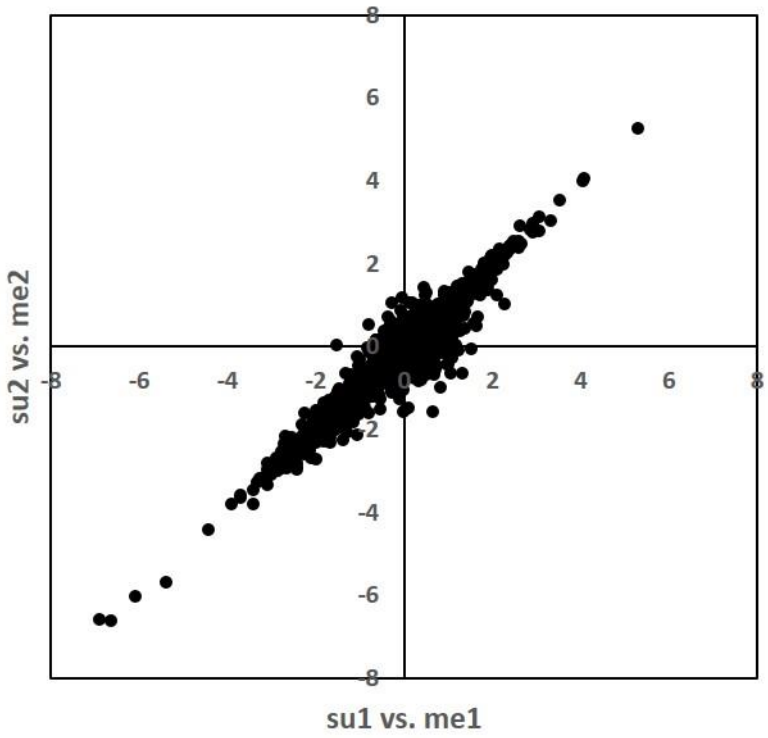

Figure 2.2 Concordance of succinate vs. methanol growth experiments by RNA-seq

Scatter plot of RPKM for succinate vs. methanol growth experiments. The log2 fold change of the RPKM value is shown for each experiment.

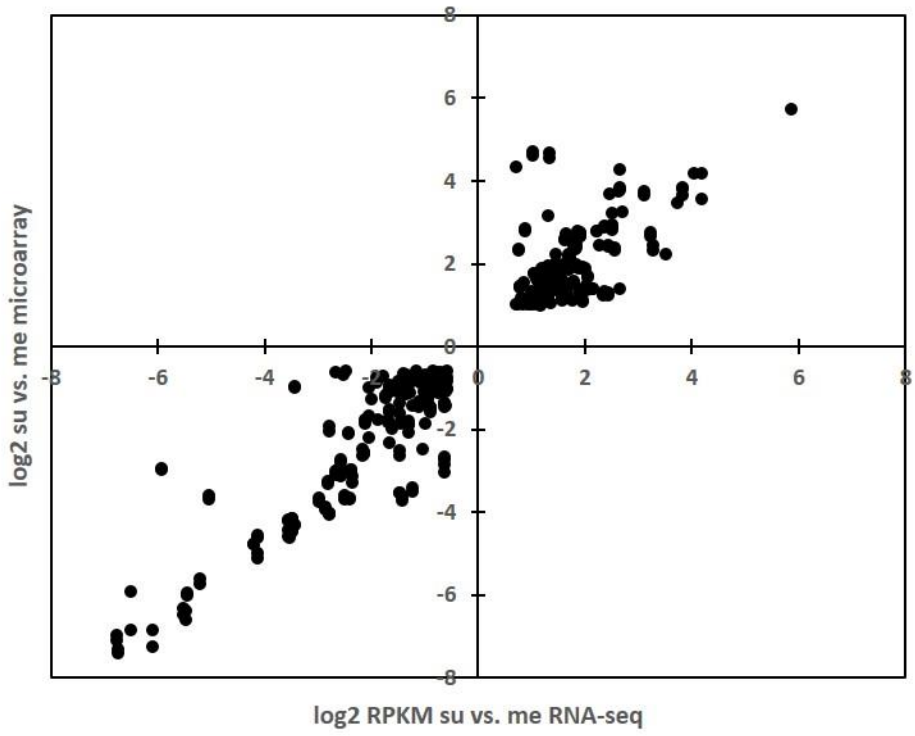

Figure 2.3 Comparison of succinate vs. methanol gene expression by RNA-seq and microarray

Scatter plot comparing succinate vs. methanol expression levels for RNA-seq and microarray experiments. Only genes with statistically significant ($p < 0.05$) expression of 1.5-fold or greater are included. Plot points are the log2 fold change in expression on succinate vs. expression on methanol. Plot is representative of two experiments for each method.

2.5. Discussion

High-throughput sequencing has made it possible to sequence the entire transcriptome of an organism, providing the option of RNA-seq to replace traditional DNA microarrays. *M. extorquens* AM1 is the best understood methylotroph with a variety of tools already available for its study. We investigated RNA-seq as a method for analyzing gene expression in *M. extorquens* AM1 and validated the method by comparing the results to existing microarray data.

By using reported RNA isolation methods optimized for *M. extorquens* AM1, high quality sequencing data were obtained using an Illumina platform. A strong, positive correlation between the depth of rRNA depletion and the amount of sequencing reads that mapped to protein-coding sequences across the genome was observed. Even with the highest level of rRNA depletion, over 100 million reads were mapped to protein-coding sequences. This opens up the possibility of multiplexing samples without sacrificing coverage. Replicate datasets from growth on two different substrates were compared, and the datasets had higher correlation than previous microarray experiments (15), demonstrating that RNA-seq is a highly reproducible method for gene expression analysis. Comparison of replicated experiments of succinate vs. methanol expression were also highly reproducible, and had a higher correlation than replicate microarray experiments. These results demonstrate that RNA-seq can out-perform traditional microarrays in terms of reproducibility.

Because rRNA sequences are the most abundant transcripts and their prevalence leads to a much lower fraction of reads mapped to protein-coding sequences, measures are often taken to reduce the amount of rRNA in the sample. However, removal of rRNA transcripts is neither completely efficient nor completely selective and can introduce biases. For this reason we investigated three rRNA depletion methods. We tested a No Depletion method to determine what

sequencing depth would be needed to obtain an accurate reflection of expression for all expressed protein-coding sequences. Even with nearly 200 million reads, and 99% of those reads mapped, a number of protein-coding sequences were found to have zero reads. However, in the depleted samples these same protein-coding sequences did have reads indicating that the sequencing depth was not sufficient in the No-Depletion samples. The genes that dropped out were for conserved proteins of unknown function. Of the two commercial depletion methods tested, Ribo-Zero was more effective at removing rRNA than MICROBExpress. Greater rRNA depletion correlated with greater loss of protein-coding transcripts, resulting in significant losses in total read counts; however, general trends in read counts across the transcriptome were unchanged, with a few exceptions. Investigation of depleted transcripts showed that several methylotrophy genes were affected to varying degrees. These results provide further evidence that comparing read counts within a sample is ill-advised, since total read counts for a particular protein-coding sequence can be significantly altered by the depletion method. These results also demonstrate the importance of studies like this one, which can identify specific biases that should be taken into consideration when designing future experiments.

When comparing replicates, it was observed that non-depleted and Ribo-Zero depleted samples had higher discordance than MICROExpress-treated samples. A pattern was identified between non-depleted and Ribo-Zero-depleted samples in which many of the genes that showed significantly different expression between replicates treated with Ribo-Zero were also significantly different in the non-depleted samples. These same genes had highly similar expression in the MICROBExpress-treated samples. This result suggests that there was a technical issue during processing of the non-depleted and Ribo-Zero-depleted samples, which were processed at a later time than the MICROBExpress-depleted samples. Removal of these

read counts from the correlation plot would give a much higher concordance. Therefore, most likely a technical issue, and not a problem with the depletion method, is the cause of the lower correlations for non-depleted and Ribo-Zero-depleted replicates.

To validate our RNA-seq approach, succinate vs. methanol datasets were compared to previous microarray datasets from the same experimental design. A high concordance was observed between the two methods. However, a number of genes with divergent expressions were observed as well. In the case of *fumC* and other genes that had larger differences in the microarrays, it is possible that there were issues in probe design resulting in abnormally high expression. These results demonstrate that RNA-seq is a viable method for transcriptomics in *M. extorquens* AM1 and that the data generated is comparable to previous microarray experiments.

The results of this study demonstrate the utility of RNA-seq as a next generation transcriptomics approach for *M. extorquens* AM1. Previously, DNA microarrays had been used to study gene expression, and we are the first to report RNA-seq transcriptomics for this model methylotroph. Furthermore, we demonstrated that the results are highly reproducible. We also showed that caution and consideration should be used when determining which rRNA depletion method will be used. Finally, we identified depletion biases, which can be used for analyzing current and future transcriptomics datasets.

2.6. Acknowledgments

I performed all the RNA-seq experiments reported in this chapter and analyzed the RPKM data. David A.C. Beck processed the sequencing data and performed the statistical analysis.

The results presented are in preparation for submission.

Chapter 3. Ethylmalonyl-CoA mutase operates as a metabolic control point in *M. extorquens* AM1

3.1. Abstract

Metabolism of one- and two-carbon compounds by the methylotrophic bacterium *Methylobacterium extorquens* AM1 involves high carbon flux through the ethylmalonyl-CoA (EMC) pathway. During growth on ethylamine, the EMC pathway operates as a linear pathway carrying the full assimilatory flux to produce glyoxylate, malate and succinate. Assimilatory carbon enters the ethylmalonyl-CoA pathway directly as acetyl-CoA, bypassing pathways for formaldehyde oxidation/assimilation and the regulatory mechanisms controlling them, making ethylamine growth a useful condition to study regulation of the EMC pathway. Wild-type *M. extorquens* was grown at steady state on limiting succinate and then the growth substrate was switched to ethylamine, a condition when the cell must make a sudden switch from utilizing the TCA cycle to the ethylmalonyl-CoA pathway for assimilation, which has been an effective strategy for identifying metabolic control points. A 9 hour lag in growth was observed during which butyryl-CoA, a degradation product of ethylmalonyl-CoA, accumulated suggesting a metabolic imbalance. Ethylmalonyl-CoA mutase activity increased to a level sufficient for the observed growth rate at 9 hours, which correlated with up-regulation of RNA transcripts for *ecm* and a decrease in the levels of ethylmalonyl-CoA When the wild-type strain over-expressing *ecm* was tested under the same substrate switchover experiment ethylmalonyl-CoA did not accumulate, growth resumed earlier, and after a transient period of slow growth, the culture grew

at a faster rate than the control. These findings demonstrate that ethylmalonyl-CoA mutase is a metabolic control point in the EMC pathway, expanding our understanding of its regulation.

3.2. Introduction

The function of central carbon metabolism is to generate essential metabolic intermediates and the energy and reducing power necessary to convert them to biomass. Metabolism must be fine-tuned to maintain cell homeostasis, balancing production and consumption of metabolic intermediates and preventing lethal accumulation of toxic compounds. Regulation of enzymes that control this balance is vital to the cell, particularly when challenged with an environmental perturbation such as a sudden change in growth substrate. Identifying metabolic control points leads to a better understanding of fundamental physiology and can suggest physiological significance, such as preventing a toxic intermediate from accumulating or balancing growth rate and growth yield. In addition, characterizing the metabolic strategies employed to effectively navigate an environmental perturbation provides valuable insight for strain manipulation and applied goals.

Methylobacterium extorquens AM1 is a facultative methylotroph capable of growth on single-carbon (C1) compounds such as methanol and methylamine, and multi-carbon compounds such as succinate (C4), and two-carbon compounds such as ethylamine and acetate (C2) (1–3). *M. extorquens* is widespread in nature and is associated with the leaf surfaces of plants, a heterogeneous and dynamic growth environment (5, 6). In this niche, *M. extorquens* must acclimate to sudden changes in growth substrate availability while preventing the lethal accumulation of the toxic methanol oxidation product formaldehyde and the downstream toxic metabolites glyoxylate and glycine (9, 36). Recent studies have demonstrated that *M. extorquens*

utilizes specific regulatory strategies when challenged with a sudden change in growth substrate, from succinate growth to methanol growth (39). Under these conditions, *M. extorquens* restricts carbon flux into the assimilatory pathways until the necessary enzymes are up-regulated sufficiently to accommodate the full assimilatory flux. These strategies allow *M. extorquens* to acclimate to the change in growth substrate more efficiently and prevent a metabolic imbalance. While regulatory check points have been identified in the methanol oxidation pathways (40), it was not possible to assess control of the assimilatory pathways, since carbon flux into assimilation was restricted, limiting the magnitude of observable change.

When *M. extorquens* grows on multi-carbon substrates like succinate, it utilizes a typical aerobic, heterotrophic metabolic network including oxidation of acetyl-CoA via the TCA cycle, NADH dehydrogenase and an electron transport chain (1, 14). During growth on C2 compounds like ethylamine and acetate, the primary growth substrate is oxidized to acetyl-CoA, which has two discrete inputs to central metabolism. Oxidation of acetyl-CoA to CO_2 occurs primarily through the TCA cycle, while assimilation of acetyl-CoA occurs through the ethylmalonyl-CoA pathway (EMC) (29). The EMC pathway has been recently described as an alternative pathway for the regeneration of glyoxylate in many isocitrate lyase-negative microbes (23, 29, 55). When *M. extorquens* is growing on a C2 substrate, the EMC pathway converts two molecules of acetyl-CoA plus two molecules of CO_2 into one molecule of glyoxylate and one molecule of succinate (23, 56). Succinate is converted to malate via succinate dehydrogenase. Glyoxylate can be condensed with one molecule of acetyl-CoA to generate one molecule of malate or be aminated to glycine (28). Glycine is converted to 2-phosphoglycerate for biomass production (29) [Fig.3.1].

Because acetyl-CoA enters directly into the EMC pathway, C2 growth provides a useful model for studying the EMC pathway independent of other C1 pathways and the mechanisms that regulate them. Furthermore, *M. extorquens* must balance production and consumption of glyoxylate, a toxic intermediate that inhibits growth at mM concentrations (15). We hypothesized that the EMC pathway might harbor a metabolic control point to restrict the production of glyoxylate. Relatively little is known about control of the EMC pathway enzymes, with only a single transcriptional regulator (CcrR, regulating *ccr*) having been reported (43). Most of the genes encoding the EMC pathway enzymes are not co-located in the *M. extorquens* genome, complicating the study of their regulation (42). Identifying a metabolic control point and characterizing its control would provide insight into how *M. extorquens* regulates carbon flux through the EMC pathway. It would also provide evidence concerning how cells acclimate to an environmental perturbation, such as a switch in growth substrate, and how cells achieve metabolic balance to restore growth.

Figure 3.1 Metabolic pathway for two-carbon metabolism in *M. extorquens* AM1

Three known fates of β-methylmalyl-CoA are shown. Genes encoding pathway enzymes are shown in gray. *mauFBEDACJGLMN*, methylamine dehydrogenase; *aldB*, aldehyde dehydrogenase; *alo*, aldehyde oxidase; *acsA*, acetyl-CoA synthase; *phaA*, β-ketothiolase; *phaB*, acetoacetyl-CoA reductase; *croR*, crotonase; *ccr*, crotonyl-CoA reductase/carboxylase; *epi*, ethylmalonyl-CoA/methylmalonyl-CoA epimerase; *ecm*, ethylmalonyl-CoA mutase; *msd*, methylsuccinyl-CoA dehydrogenase; *mcd*, mesaconyl-CoA hydratase; *mclA1*, *mclA2*, β-methylmalyl-CoA lyase; *mcl2*, malyl-CoA thioesterase; *pccAB*, propionyl-CoA carboxylase; *mcm*, methylmalonyl-CoA mutase; *sct*, succinyl-CoA transferase; *sga*, serine-glyoxylate aminotransferase; *gcv*, glycine cleavage system; *glyA*, serine-hydroxymethyltransferase. *, genes for the conversion of serine to 2-phosphoglycerate are not shown. Parentheses denote a predicted function not confirmed by mutant phenotype. amcy, amicyanin; me-H_4F, methylene tetrahydrofolate; PHB, poly-β-hydroxybutyrate cycle.

In this study we report ethylmalonyl-CoA mutase (Ecm), a reversible enzyme, as a regulatory control point in the EMC pathway that allows *M. extorquens* to efficiently restore metabolic balance when challenged with a sudden change in growth substrate. Our multi-layered analysis indicates that Ecm levels could be controlled at the transcriptional level.

Since its full description in 2007, the EMC pathway has been of interest for biotechnological applications due the variety of CoA-esters produced as metabolic intermediates in the pathway (10–12). Identifying a regulatory control point in the EMC pathway will assist future engineering efforts to develop *M. extorquens* as a platform for the production of value-added chemicals.

3.3. Materials and methods

3.3.1. Materials

Chemicals and chemical standards were obtained from Sigma-Aldrich (St. Louis, MO, USA) unless noted otherwise. Analytical-reagent grade potassium dihydrogen phosphate and sodium dihydrogen phosphate were purchased from Fisher Scientific (Hampton, NH, USA). HPLC-grade methanol and absolute ethanol (Sigma, St. Louis, MO, USA) were used for extraction procedures and buffer preparation for HPLC analysis. Nanopure (Barnstead Life Sciences model, Thermo Fisher Scientific, Waltham, MA) water was used for the preparation of all media, buffer, standard, and samples solutions.

3.3.2. Bacterial strains and growth conditions

M. extorquens AM1 strains and plasmids used in this work are described in Table 1. *M. extorquens* AM1 was grown in shake flasks as reported with succinate or ethylamine as the growth substrate (28, 39, 40).

3.3.3. Chemostat cultivation

Cultures were grown as reported (16, 39) with minor changes as described here in brief: For the initial time point (T = 0 min) of the substrate transition experiment, *M. extorquens* AM1 was grown to steady state in chemostat culture with limiting succinate. Cultivation was in minimal medium (39) with 3.7 mM succinate as a growth-limiting nutrient in a 2.2-liter bench-top BioFlo/Celligen 115 Modular BioReactor-Fermentor (New Brunswick Scientific) with a working volume of 2.0 L, resulting in an OD_{600} of ~0.63. The dilution rate was sustained at 0.163 h^{-1}. Once the culture reached steady-state, defined as maintaining constant OD_{600} and a balanced flow rate for a minimum of three regenerations of the chemostat volume, the flows were stopped and ethylamine was added to the succinate-limited steady-state culture to a final concentration of 20 mM. Cells were harvested at several time points as described in the text for enzyme activity measurements, RNA-seq transcriptomics and metabolomics analyses as described below. For the substrate switchover experiment over-expressing *ecm* via pCM80, 1.5 ug/ml tetracycline was added to retain the plasmid and 13 *u*M cobalt chloride was added to the culture medium to negate inhibitory effects of pCM80. The same concentration of cobalt chloride was added to the WT culture (57).

Table 3.1 *Methylobacterium extorquens* strains and plasmids used in this study

Strain or plasmid	Description	Reference
Strain		
AM1	Rif[§] derivative	Nunn, 1986
Plasmid		
pCM80	*M. extorquens* AM1 expression vector (PmxaF)	Marx, 2001
pET-16b	N-terminal deca-histidine tag expression vector	EMD Millipore
pNG109	pCM80 with *ecm*	This study
pNG221	pET-16b with *ccr*	This study
pNG235	pET-16b with *ecm*	This study
pNG253	pET-16b with *epi*	This study

3.3.4. LC-MS and GC-MS metabolomics

Samples for GC-MS and LC-MS were collected by fast filtration using 20 mL of cell culture. Cells in the filter were frozen immediately with liquid nitrogen into a pre-cooled 50 mL-plastic tube. Samples were lyophilized for two hours. Hot extraction of the metabolites was achieved by adding 20 mL of boiling water and by boiling the samples for 10 minutes. Samples were then cooled on ice for 30 min and centrifuged for 20 min at 4,800 x g, 4^0C. The supernatant was transferred to a new 50 mL tube and lyophilized for 2 days. The lyophilized product was resuspended in 1 mL of water and lyophilized for one day. Samples were resuspended in 60 *u*L of water and divided in two. One half was used for metabolomics analysis using two dimensional gas chromatography. Samples were derivatized using TBDMS. GC-MS measurements were used to analyze amino acid and organic acid content as described (39). The second half of the sample was analyzed by LC-MS to determine changes in levels of CoA-derivatives using a Waters Xevo LC–MS system consisting of an Acquity UPLC system and a Xevo triple-quadruple mass spectrometer (Milford, MA). A CSH-C18 column (130Å, 1.7 µm, 2.1 mm X 100 mm) was used for the liquid chromatography with a mobile phase A consisting of 25 mM ammonium acetate, 2 % acetic acid (v:v) , 1 % formic acid (v:v) in water, and a mobile phase B consisting of 2 % acetic acid (v:v) , 1 % formic acid (v:v) in acetonitrile. The gradient was as follows: Start A = 100 %, set A = 80 % at 2.5 min, A= 55 % at 5 min, A= 5 % at 6 min, A=100 % at 7 min, A= 100 % at 8 min at a flow rate of 0.3 ml/min. All data were analyzed using the Masslynx Quanlynx Applications Manager software.

3.3.5. RNA-seq transcriptomics

Samples (50 ml) of cells were collected and RNA was immediately extracted and digested with DNaseI (Ambion) as described (15). RNA purity was analyzed by rtPCR using iScript (Bio-Rad) reverse transcriptase. RNA integrity was verified by 2100 BioAnalyzer (Agilent). rRNA depletion, library prep and Illumina Hi-Seq sequencing were carried out by the UW High Throughput Genomics Center (HTGC). For each sample condition, two biological replicates were used. Ribosomal RNA was depleted using Ribo-Zero (Epicentre, Madison, WI, USA). Each replicate was sequenced to a target depth of 20 million single-ended 36 bp reads using a multiplex strategy on an Illumina Hi-Seq 2000 (Illumina, San Diego, CA, USA). Quality filtering of the raw read counts was performed using Illumina's CASAVA. The filtered data were processed as follows. The raw reads were aligned to the *M. extorquens* AM1 [MAGE, March 2nd, 2012] genome using BWA version 0.7.4-r385 with default parameters (51). The alignments were post-processed into sorted BAM files with SAMTools version 0.1.19-44428cd (52). Reads were attributed to open reading frames (ORFs) using the htseq-count tool from the 'HTSeq' framework version 0.5.4p5 in the 'intersection-nonempty' mode (53). The final abundance was measured by RPKM (50). T-tests were used to assign significance with post-hoc adjustment by the method of Storey (54).

3.3.6. Enzyme assays

Cells for enzyme activities (50 ml) were harvested and processed as reported (39). A minimum of 2 biological replicates were assayed for each enzyme. Crotonyl-CoA reductase (Ccr) activity was measured as reported (22) with 2mM NADPH (39). Ecm and Epi activities

were measured as reported (55) with minor changes. Briefly, the assays were quenched by boiling at 100°C for 5 minutes. Methylsuccinate was quantified by GC-MS as described above. When pure enzyme is indicated, 45 ug of Ecm was added and 3-10 ug of Epi was added to the reaction. Ccr activity was measured as reported (39).

3.3.7. Calculation of hypothetical minimum enzyme activity

The standard approved equation $dS/dt = (\mu/Y) \cdot X$ was used as reported (41) with minor changes. For growth yield, Y, the value of 48% of bacterial cell dry mass as carbon was used (58). For X, the value of 55% cell dry weight as protein was used. Activity was calculated accounting for 5 carbon atoms since 2 molecules of acetyl-CoA are condensed and then carboxylated to generate ethylmalonyl-CoA.

3.3.8. Cloning and heterologous expression of ccr, ecm, and epi from M. extorquens AM1 *in E. coli*

The gene encoding crotonyl-CoA reductase was amplified from *M. extorquens* AM1 chromosomal DNA using the forward primer 5'-AT*CATATG*GCTGCAAGCGCAGCAC-3' inserting and NdeI restriction site and the reverse primer 5'-AT*GGATCC*TCACATCGCCTTGAGCGGGat-3' inserting a BamHI restriction site. PCR product was generated using Herculase II Fusion DNA polymerase (Agilent Technologies, Santa Clara, CA, USA) as per the manufacturer's protocol, including 30 cycles, denaturation for 20 seconds, primer annealing at 65°C for 20 seconds, and extension at 72°C for 1.5 minutes. The PCR product was isolated and cloned into pET16B (EMD Millipore, Darmstadt, Germany)

producing pNG221 for expression of *ccr* and production of N-terminal deca-His tagged fusion protein. Competent E. coli (BL21-AI) was transformed with pNG221 and grown at 37°C in LB medium containing 50 $ug \cdot ml^{-1}$ ampicillin. When the culture reached an OD_{600} of 0.6-0.7, expression was induced by the addition of 1.0 mM isopropyl thiogalactopyranoside (IPTG) and 0.2% (w/v) L-Arabinose (Sigma). The cultures were grown for an additional 3.5 hours at 30°C.

The gene encoding ethylmalonyl-CoA mutase was amplified using the forward primer 5'-A*CATATG*AGCGCGCAAGCGAG -3' inserting an NdeI restriction site and the reverse primer 5'-AT*GGATCC*GAAGACCTGCGCCC-3' inserting a BamHI restriction site. PCR was performed for 30 cycles with 20 seconds denaturation, 20 seconds annealing at 65°C, and extension at 72°C for 2 minutes. The PCR product was isolated and cloned into pET16b to obtain pNG235 for expression of *ecm* and production of an N-terminal deca-His tagged fusion protein.

The gene encoding ethylmalonyl-CoA/methylmalonyl-CoA epimerase was PCR amplified using the forward primer 5'-TCACAGCAGCGGCCATATCGAAGGTCGTCAATGATCGGACGGCTCAAT-3' and the reverse primer 5'-CAGCTTCCTTTCGGGCTTTGTTAGCAGCCGTCAGACCTGCTCCAGCTC-3' generating the insert with ends overlapping the pET16b insertion site. The pET16b backbone fragment was amplified using the forward primer 5'- CGGCTGCTAACAAAGCCCGAAAGGAAGCTGA-3' and the reverse primer 5'-TGACGACCTTCGATATGGCCGCTGCTGTG-3'. The backbone and insert fragments were assembled using Gibson Assembly (NEB) to generate pNG252. The plasmid was transformed into BL21-AI for protein expression.

Cultures for expression of *ecm* and *epi* were grown at 37°C in LB medium containing 100 ug.ml-1 ampicillin. For expression of *ecm* 1 ml of 20 mg/L vitamin B12 was added per liter of LB medium (12, 32). Cultures were induced with 0.5 mM IPTG and 0.2% (w/v) L-Arabinose (Sigma) when the OD_{600} was 0.6-0.9. The cultures were grown for another 3 hours at 30°C. All cells were harvested by ultracentrifugation using a Sorvall RC-5B centrifuge (Thermo Fisher Scientific) at 7,500 rpm, 4°C and stored at -80°C until needed.

3.3.9. Purification of Ccr, Ecm and Epi

All enzymes were purified as reported (40). Buffer A was 50 mM Tris-Hcl, pH 7.8, 5 mM imidazole. Buffer B was 50 mM Tris-Hcl, pH 7.8, 200 mM NaCl, 5 mM imidazole. Buffer C was 50 mM Tris-Hcl, pH 7.8, 500 mM imidazole. Glycerol (15% vol/vol) was added to the purified enzyme which was stored at -80°C until needed.

3.3.10. Generation of expression mutant constructs

ecm was pcr amplified using the forward primer 5'-ATAAGCTTCGGACAGGGAGTCGAGTG -3' introducing a HindIII restriction site and the reverse primer 5'- ATTCTAGAGGCGGCCCTCAGAAGACC -3' introducing an XbaI restriction site. The pcr product was inserted into pCM80, generating pNG109 which was mated into *M. extorquens* AM1 as reported (47).

3.4. Results and discussion

A metabolic perturbation strategy was used to assess the response of the assimilatory pathways. Chemostat cultures of wild-type *M. extorquens* AM1 were grown on limiting succinate at a growth rate of 0.163 h^{-1}. After the culture reached steady-state, the in and out flows were stopped and 20 mM ethylamine was added. This method achieves a rapid substrate switch, since a succinate-limited culture has very low succinate (39), forcing the metabolic network to acclimate to metabolize the new, available substrate. Samples were taken immediately before the substrate switch (Time 0) and at several time points after (2, 4.5, 9, 18, and 22.5 hours). Culture density (OD_{600}), metabolites, RNA transcripts, and specific enzyme activities were measured at these time points. These analyses were compared to generate a hypothesis for a metabolic control point in the EMC pathway.

3.4.1. Growth response to the switch from succinate to ethylamine

Immediately after the switchover in growth substrate the culture reproducibly entered a transient lag period during which the culture density decreased from OD_{600} 0.065±0.01 to 0.060±0.02 [Fig. 3.2]. After nine hours the culture resumed growth at a rate of 0.051 h^{-1} that is similar to batch culture in shake flasks, demonstrating the transient nature of the lag.

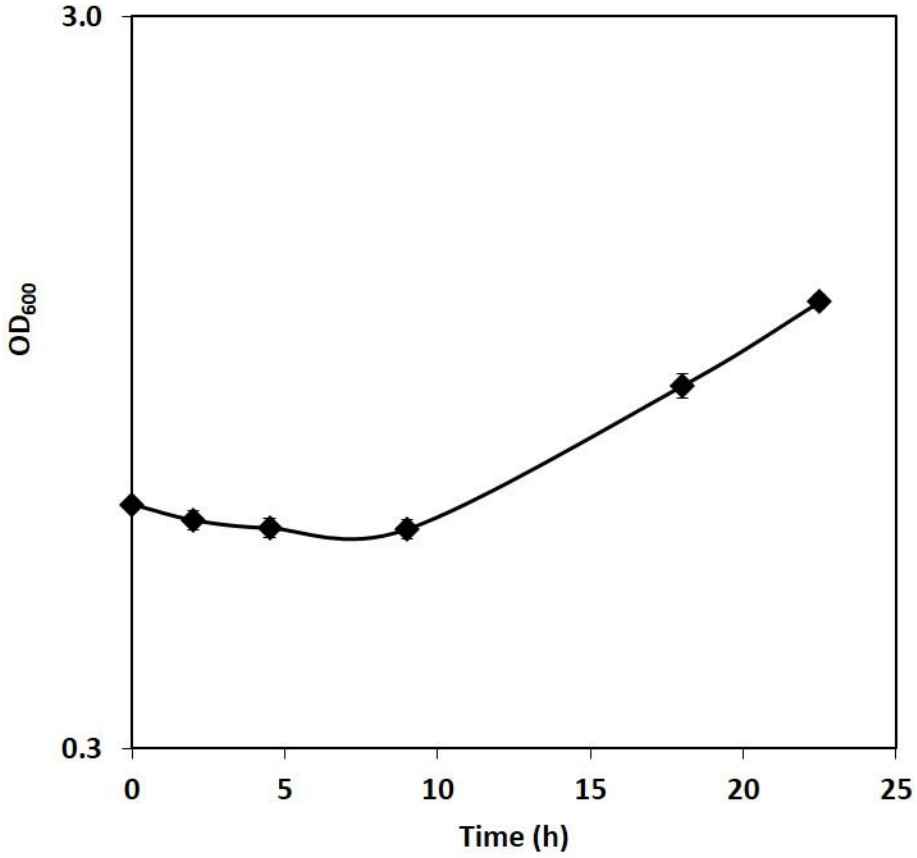

Figure 3.2 Growth response of *M. extorquens* AM1 after the switchover from succinate-growth to ethylamine growth.

Time zero indicates when chemostat flows were stopped and 20 mM ethylamine was added to the bioreactor. Optical density (OD_{600}) was measured at the plotted time points. Error bars are the standard deviation of 6 biological replicates.

3.4.2. Gene expression response of ethylamine oxidation and central metabolic pathways

RNA-seq transcriptomics was used to assess response of *M. extorquens* to the change in growth substrate at the transcriptional level. Oxidation of ethylamine to acetaldehyde occurs via methylamine dehydrogenase (Mau), conversion of acetaldehyde to acetate is proposed to be carried out by an aldehyde dehydrogenase and/or aldehyde oxidase activity, and conversion of acetate to acetyl-CoA is proposed to be carried out by acetyl-CoA synthetase activity [Fig. 3.1;(28)]. Transcription of the genes encoding these enzymes began to increase by 2 h and peaked at 9 h, remaining highly expressed for the duration of the experiment (data not shown). Peak increases ranged from 8- to 59-fold compared to the time zero transcripts. At the later time points the transcriptomic profile for ethylamine oxidation is consistent with transcripts reported for batch culture (28).

The TCA cycle is the primary assimilatory pathway utilized by *M. extorquens* while growing on succinate, and carbon flux supports a growth rate of 0.163 h^{-1}.(16, 35). While growing on ethylamine, however, *M. extorquens* splits acetyl-CoA flux between the TCA cycle (oxidation) and the EMC pathway (assimilation) (29) and the growth rate is much slower (0.051 h^{-1}). As expected due to the downshift in growth rate, TCA cycle genes were immediately down-regulated (data not shown) between 2- and 8-fold at two hours after the substrate switch. Nine hours after the switch in substrate, several TCA cycle genes were up-regulated again but did not return to the same levels of expression as on succinate. On average TCA cycle genes were down-regulated about 4-fold.

The EMC pathway genes, on the other hand, did not respond as a uniform module and showed multiple regulation patterns. The genes encoding the first two steps, β-ketothiolase (*phaA*) and acetoacetyl-CoA reductase (*phaB*), were initially up-regulated 2- to 3-fold (2 hours)

but by 9 hours transcripts of both genes decreased to near-initial levels [Fig. 3.3]. The genes encoding crotonase (*croR*), the third enzyme of the EMC pathway, and crotonyl-CoA reductase/carboxylase (*ccr*), the fourth enzyme, showed a pattern of initial down-regulation, with subsequent up-regulation above the time zero point [Fig. 3.3 and Fig. 3.4]. The gene encoding the next enzyme, ethylmalonyl-CoA/methylmalonyl-CoA epimerase (*epi*) was the most differentially regulated of the EMC pathway genes. At 2 hours it was down-regulated 6-fold, and 9-fold by 4.5 hours. At 9 hours transcripts increased 2-fold, and by 18 hours transcript levels had returned to succinate-grown levels but the response was never greater than the time zero point [Fig. 3.4]. A similar pattern was observed for the gene encoding the 8th enzyme, mesaconyl-CoA dehydratase, [*mcd*; Fig. 3.4]. However, the two steps between *epi* and *mcd* showed very different responses. Transcripts of ethylmalonyl-CoA mutase (*ecm*) were up-regulated initially, peaked at 9 h, then decreased to 2-fold relative to initial levels [Fig. 3.4]. Expression of methylsuccinyl-CoA dehydrogenase (*msd*) did not change significantly over the course of the switchover. Likewise, the genes encoding the *L*-malyl-CoA/β-methylmalyl-CoA lyases (*mclA1* and *mclA2*) displayed disparate expression patterns. Transcripts of *mclA1* initially increased nearly 4-fold (2 hours), remained up-regulated at 4.5 hours (2-fold), but by 9 hours had returned to succinate-growth levels. Transcript levels of *mclA2*, on the other hand, decreased slightly initially, then at 9 hours increased over 2-fold and remained up-regulated for the duration of the experiment. Expression of the gene encoding *L*-malyl-CoA thioesterase (*mcl2*) was initially down-regulated 2-fold and remained at this level until 18 hours after the substrate switch, when levels returned to succinate-growth levels. The genes encoding the two subunits of propionyl-CoA carboxylate (*pccAB*) both displayed the same expression pattern, being down-regulated 4-fold at 2 hours, 5-fold at 4.5 hours, and being up-regulated 2-fold at 9 hours. By 18

hours transcripts had returned to succinate-growth levels. Transcripts of the methylmalonyl-CoA mutase gene *mcmA* responded in a manner similar to *ecm*, while transcripts of *mcmB* showed a pattern similar to *mcl2*. Expression of the genes encoding succinyl-CoA transferase (*sct*) and succinate dehydrogenase (*sdh*) was down-regulated until 18 hours after the substrate switch, and expression levels returned to near succinate-growth levels by the end of the time course. The transcriptomic analysis suggests that *M. extorquens* undergoes substantial remodeling of the metabolic network to switch from succinate-growth to ethylamine-growth, and that the genes of the EMC pathway are not uniformly regulated.

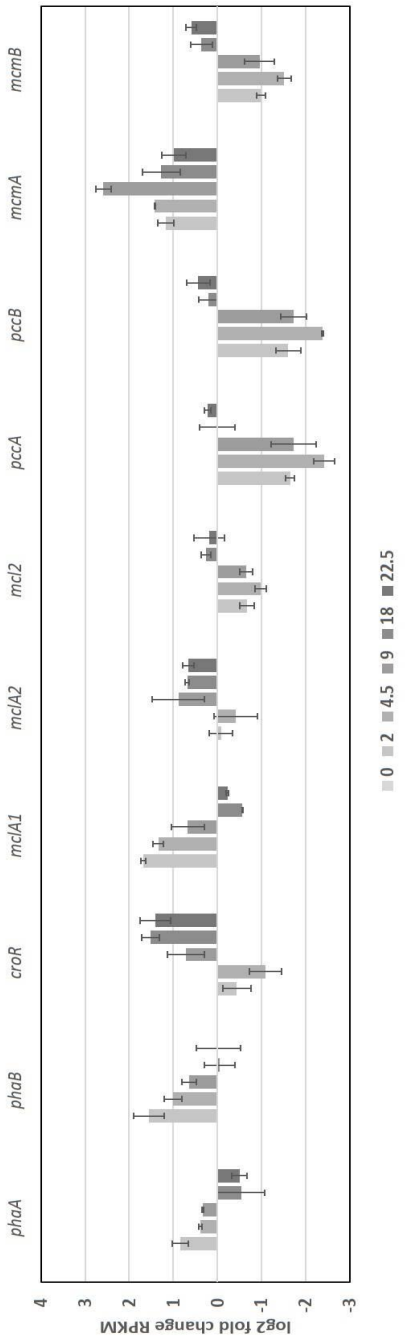

Figure 3.3 Transcriptional profile of EMC pathway genes during the transition from succinate growth to ethylamine growth

Genes encoding enzymes for the conversion of acetyl CoA to crotonyl-CoA and β-methylmalyl-CoA to succinyl-CoA or glyoxylate are shown. Time in hours after the addition of ethylamine is indicated by grayscale. The RPKM was normalized to the read count at time zero (0 h). Error bars are the standard deviation of two biological replicates.

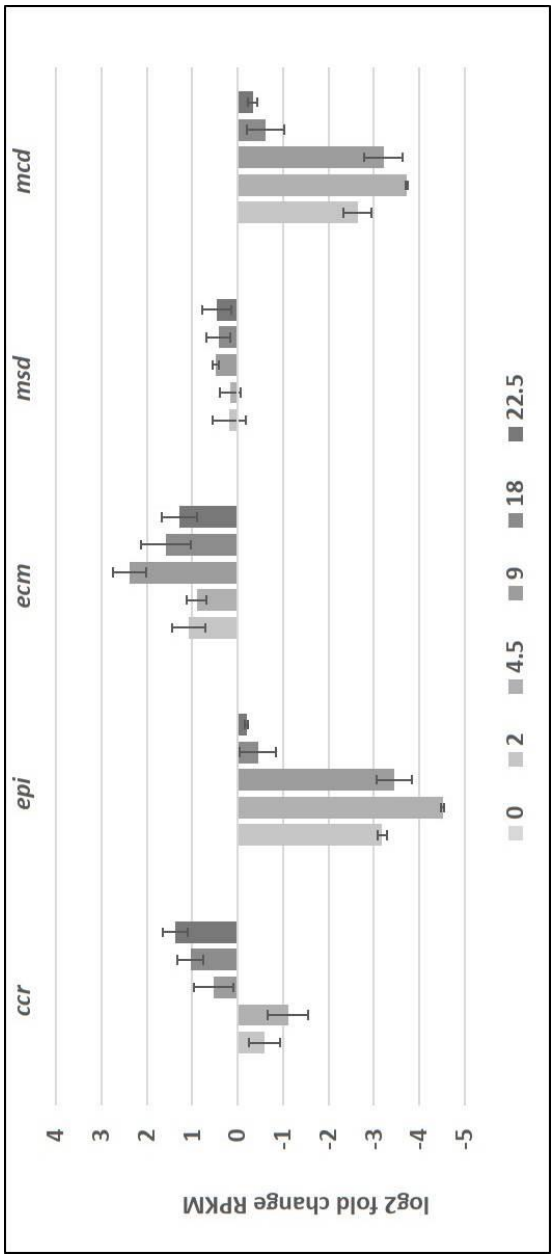

Figure 3.4 Transcriptional profile of the middle EMC pathway genes during the transition from succinate growth to ethylamine growth

Genes encoding enzymes for the conversion of crotonyl-CoA to β-methylmalyl-CoA are shown. Time in hours after the addition of ethylamine is indicated by grayscale. The RPKM was normalized to the read count at time zero (0 h). Error bars are the standard deviation of two biological replicates.

3.4.3. Response of EMC pathway metabolites and glyoxylate

Metabolic intermediates of the EMC pathway were measured in samples from the switchover time course to determine whether any accumulated during the transient growth lag. Accumulation of a metabolite during the growth lag would indicate an imbalance in production and consumption of the compound. The metabolites of the EMC pathway were measured either directly via LC-MS as CoA-esters or in the case of methylsuccinate by hydrolyzing the ester and measuring the organic acid via GC-MS. In the latter case, the CoA derivative was poorly detectable but the acid form was detected in a more robust manner. By this method we were able to determine relative concentrations of most EMC pathway intermediates as well as several other compounds of interest, such as glyoxylate. Relative concentrations were normalized to time zero and compared across the substrate switchover time points.

Ethylamine is oxidized to acetaldehyde, which is converted to acetyl-CoA, which is the entry molecule for the EMC pathway [Fig. 3.1]. Acetyl-CoA levels did not change significantly over the course of the switchover experiment from the initial levels during growth on succinate [Fig. 3.5]. Relative levels of β-hydroxybutyryl-CoA did not change within the first two hours after the switch from succinate growth to ethylamine growth, but by 4.5 hours had increased over 3-fold, suggesting an imbalance of carbon flux at the branch point between the poly-β-hydroxybutyrate (PHB) cycle and the EMC pathway [Fig. 3.1]. However, by 9 hours β-hydroxybutyryl-CoA levels had returned to their initial relative concentration, suggesting that the imbalance had been alleviated even though the culture had not yet resumed growth. Relative levels of crotonyl-CoA decreased slightly after the switch from succinate to ethylamine.

The relative concentration of butyryl-CoA increased the most during the time course. Butyryl-CoA is produced from crotonyl-CoA by crotonyl-CoA reductase/carboxylase (Ccr) at a

very low rate and only when bicarbonate/CO_2 is absent (22). When bicarbonate/CO_2 is present ethylmalonyl-CoA is formed exclusively. Since air was constantly flowing into the bioreactors during the substrate switchover experiment and the cells produce CO_2 during metabolism, it is unlikely that CO_2 was limiting. Instead, it seemed possible that butyryl-CoA was formed via spontaneous decarboxylation of ethylmalonyl-CoA during metabolite extraction. No signal above background for ethylmalonyl-CoA was reliably detected during the substrate switchover, also suggesting that it was unstable during sample processing. A sample of pure ethylmalonyl-CoA was treated by the same hot water extraction process (see methods) and both compounds were measured by LC-MS. After hot water treatment ethylmalonyl-CoA decreased 4-fold whereas butyryl-CoA increased 4-fold, demonstrating that ethylmalonyl-CoA was unstable during the extraction process, and the degradation generated butyryl-CoA [Fig. 3.6]. Based on these results, we interpreted the butyryl-CoA signal as indicative of ethylmalonyl-CoA.

By 4.5 hours after the growth substrate change, ethylmalonyl-CoA-derived butyryl-CoA began to accumulate (over 2-fold) and by 9 hours had increased over 3.5-fold, suggesting a metabolic imbalance in the EMC pathway after the ethylmalonyl-CoA production step within the first 9 hours. Relative levels of ethylmalonyl-CoA decreased at 18 and 22.5 hours, concurrent with the culture resuming growth, but still remained higher than succinate-grown levels (Time 0).

Two EMC pathway intermediates that follow ethylmalonyl-CoA are methylsuccinyl-CoA and propionyl-CoA [Fig. 3.1]. As noted above, methylsuccinyl-CoA was hydrolyzed and measured as methylsuccinate. Both intermediates decreased in concentration relative to time 0, and only returned to the initial levels at later time points. Another intermediate, mesaconyl-CoA, was not detected by our method, suggesting that the compound did not increase greatly in

relative concentration above background for any time point of the experiment. Glyoxylate decreased roughly 2-fold at 2 hours, and relative levels did not increase significantly during the time course. This result corroborates the concept that the metabolic network is set in a way such that this toxic intermediate does not accumulate during perturbations (39).

Accumulation of ethylmalonyl-CoA and decreased levels of the downstream metabolites methylsuccinate and propionyl-CoA suggest that carbon flux was restricted at ethylmalonyl-CoA during the growth lag.

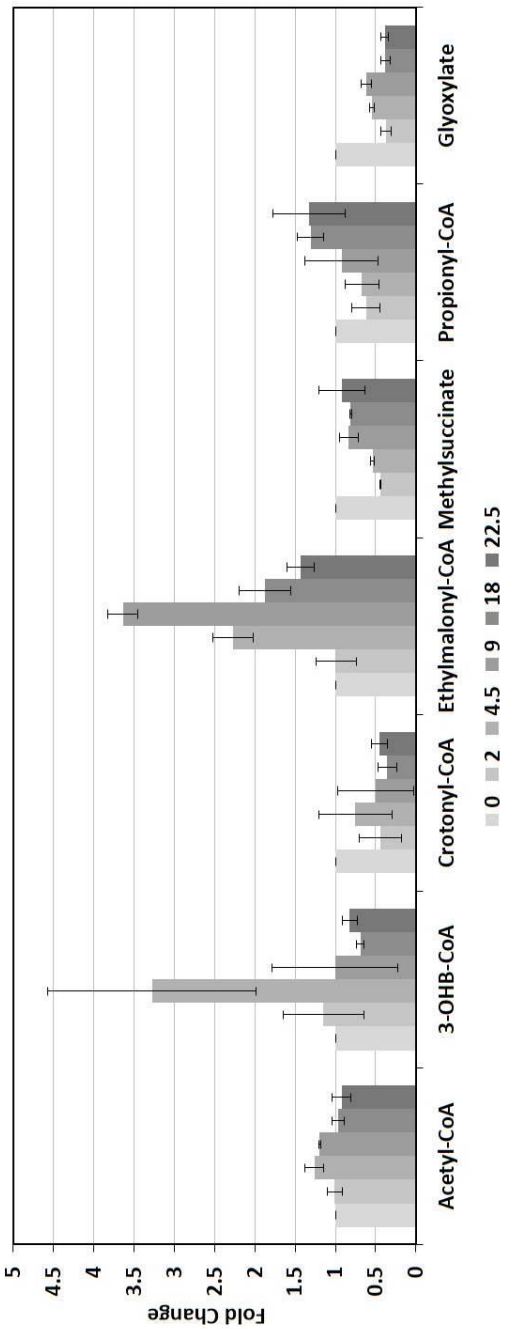

Figure 3.5 Metabolite profile of the mid-EMC pathway during the transition from succinate growth to ethylamine growth

CoA intermediates were measured by LC-MS. Organic acids were measured by GC-MS. Time in hours after the addition of ethylamine is indicated by grayscale. 3-OHB-CoA; 3-hydroxybutyryl-CoA. Ethylmalonyl-CoA, ethylmalonyl-CoA-derived butyryl-CoA (see text). Error bars are the standard deviation of two biological replicates. Each sample represents a minimum of two injections.

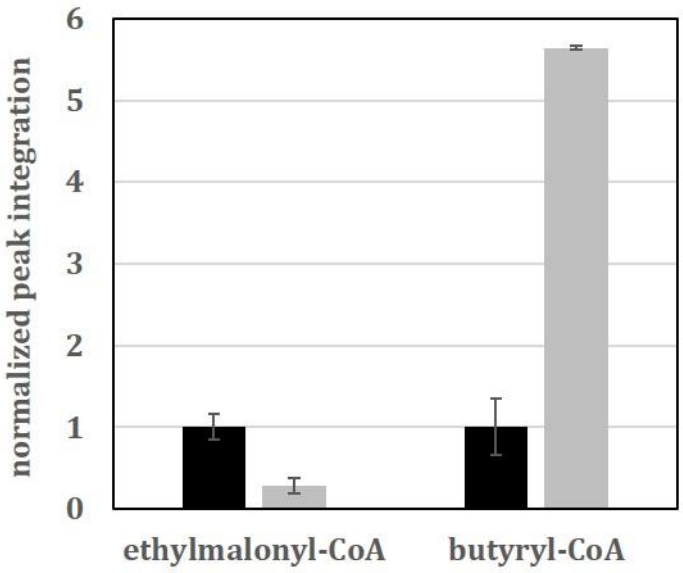

Figure 3.6 Degradation of ethylmalonyl-CoA to butyryl-CoA by hot water extraction

Ethylmalonyl-CoA was boiled at 100 C for 10 minutes. Relative concentrations of ethylmalonyl-CoA and butyryl-CoA were measured before and after treatment by LC-MS/MS. The untreated sample is shown in black bars. The treated control sample is shown in the gray bars. A minimum of two injections were made for each sample. Error bars represent the standard deviation of two replicates.

3.4.4. Response of key C2 enzymes

Since ethylmalonyl-CoA accumulated during the transitory period between the substrate switchover and resumption of growth, we measured activity of the two enzymes that utilize ethylmalonyl-CoA as a substrate. As shown in Fig. 3.1, the epimerase converts (2S)-ethylmalonyl-CoA to (2R)-ethylmalonyl-CoA. Then the ethylmalonyl-CoA mutase converts (2R)-ethylmalonyl-CoA to (2S)-methylsuccinyl-CoA in an adenosylcobalamin (AdoCbl)-dependent manner. By measuring activities we could identify which enzyme was limiting, and compare activity to transcripts, which would suggest if the enzyme levels were regulated transcriptionally or post-transcriptionally. Since the substrates of these enzymes are not available commercially, (2*S*)-ethylmalonyl-CoA was produced using commercially purchased crotonyl-CoA and purified recombinant Ccr (see Methods). We first measured the rate of conversion of (2*S*)-ethylmalonyl-CoA to (2*S*)-methylsuccinyl-CoA, a reaction encompassing both the epimerase and mutase activities, in cell-free extracts from the substrate switchover experiment. The product was followed by a GC-MS assay (see Methods). The conversion activity measured should reflect the rate-limiting enzyme at the time the sample was taken. The activity profile over time [Fig. 3.7A] is similar to the RNA expression profile of *ecm*, suggesting that Ecm is the rate limiting enzyme. To verify this result, purified recombinant Epi was added in excess to the 4.5 h cell-free extract and activity did not increase [Fig. 3.7B]. When purified recombinant Ecm was added in excess to the 4.5 h cell-free extract the activity increased nearly 2-fold from 9 $nmol \cdot min^{-1} \cdot mg^{-1}$ to 17 $nmol \cdot min^{-1} \cdot mg^{-1}$. When both recombinant Epi and Ecm were added to the extract, the activity increased over 2-fold to 14 $nmol \cdot min^{-1} \cdot mg^{-1}$ which is not significantly different from the result when adding Ecm alone ($p = 0.1$, Student's T-test). Because β-hydroxybutyryl-CoA accumulated during the early time points when activity was low, we also

tested 5-400 uM concentrations of β-hydroxybutyryl-CoA for inhibition of the reaction but no effect on activity was observed (data not shown). This result does not, however, eliminate the possibility of another small molecule regulating the reaction, and more studies would be needed to identify such an effector.

Since the conversions of (2*S*)-ethylmalonyl-CoA to (2*R*)-ethylmalonyl-CoA and (2*R*)-ethylmalonyl-CoA to (2*S*)-methylsuccinyl-CoA are catalyzed by reversible enzymes, ethylmalonyl-CoA and methylsuccinyl-CoA are assumed to be at equilibrium during steady-state (59, 60). However, the switch from succinate to ethylamine growth creates a transient metabolic imbalance, as evidenced by the long growth lag and change in the metabolite pools. A similar response was seen during the substrate switchover from succinate-growth to methanol-growth (39). In this condition, when carbon begins to flow suddenly into the middle portion of the EMC pathway, the metabolite pools are no longer at equilibrium. Our data suggest that Ccr generates (2*S*)-ethylmalonyl-CoA much more rapidly than the combined activity of Epi/Ecm can consume it. In support of this hypothesis, Ccr activity measured in the switchover samples was high (150 to 240 $nmol \cdot min^{-1} \cdot mg^{-1}$; data not shown) compared to the Epi/Ecm activity measured from 4.5 h extracts (9 $nmol \cdot min^{-1} \cdot mg^{-1)}$. It is suggestive from Fig. 3.7B that the in vitro Epi activity is higher than Ecm, since added Epi does not increase the activity. Msd activity is expected to be on the order of five times higher than Ecm (41), which would draw down the (2*S*)-methylsuccinyl-CoA pool and drive the Ecm reaction towards product. Over time the substrate and product pools should begin to reach a new equilibrium sufficient for normal growth on ethylamine. Since Ecm activity is lower than Epi and Msd, it should be the limiting step in the reaction sequence. As the cells reach a new metabolic setpoint while up-regulating genes encoding necessary enzymes to drive metabolism, the metabolite pools would return to equilibrium. Consistent with this model,

ethylmalonyl-CoA pools return to baseline levels as the cells resume growth, beginning at 9 hours after the switchover. The correlation of a growth lag with accumulation of ethylmalonyl-CoA and resumed growth with decreased ethylmalonyl-CoA coupled to the enzyme assay results suggested Ecm as a metabolic control point.

Based on the growth rate observed, the minimal Ecm activity needed for the growth rate observed was calculated to be 12 $nmol \cdot min^{-1} \cdot mg^{-1}$ (see methods). Our enzyme activity results show that at the time this level of activity was reached (9 hour time point), the culture growth resumes. Since the Epi activity measured was greater than this level, the results support the model that Ecm activity, not Epi activity, was restricting consumption of ethylmalonyl-CoA and preventing growth during the growth lag.

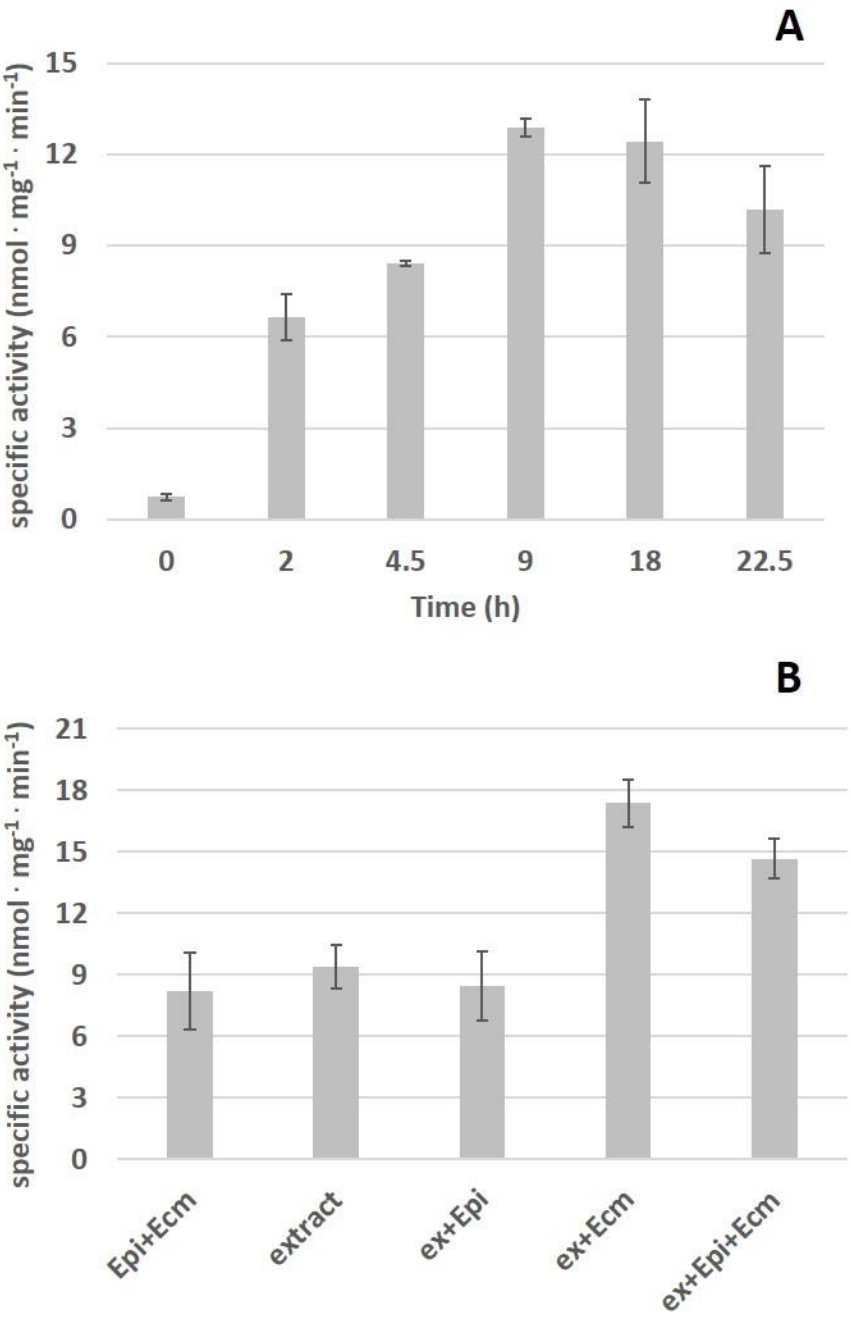

Figure 3.7 Ecm limits ethylmalonyl-CoA to methylsuccinyl-CoA activity

Specific activity was calculated using the amount of methylsuccinate (hydrolyzed methylsuccinyl-CoA) produced from (2S)-ethylmalonyl-CoA. A) Activity measured for each time point of the succinate-growth to ethylamine-growth substrate switchover experiment. B) Activity measured in a pure component system with only purified Epi and Ecm (+Epi+Ecm) and no extract, extract alone from the 4.5 h time point (ex), extract with Epi (ex+Epi), Ecm (ex+Ecm) or both (ex+Epi+Ecm) added. GC-MS measurements are from a minimum of two injections per sample. Error bars are the standard deviation of at least two biological replicates.

3.4.5. Over-expression of ecm removes control in the EMC Pathway

The previous results suggested that Ecm was functioning as a control point during the transition from succinate to ethylamine growth, restricting consumption of ethylmalonyl-CoA and production of downstream intermediates during the lag period of the transition. To test this model, we overrode the control point by over-expressing *ecm* from a strong promoter (the *mxaF* promoter) that is highly expressed in both succinate and ethylamine growth (15, 61). Ecm activity increased 4-fold (data not shown). We then repeated the switchover from succinate growth to ethylamine growth with the over-expression strain. The concentration of cobalt chloride in the medium was increased to 13 *u*M, which is known to reduce a vector-induced growth defect (57). In these higher cobalt growth conditions the wild-type strain responded to the substrate switchover as we had observed with lower cobalt in the medium (data not shown). The growth response of the strain harboring the empty plasmid was similar to the response of the wild-type strain, with a lag in growth following the substrate switchover for 9 h and resumed growth corresponding to a 12-13 hour doubling time [Fig. 3.8A]. However, the strain over-expressing *ecm* reproducibly began growing after the 4.5 hour time point [Fig. 3.8A]. When growth resumed, the initial rate was 0.024 h^{-1}. Between 8 and 9 hours after the switchover, the growth rate of the culture increased to 0.071 h^{-1}, 21% faster than the wild-type and control strains. In the control strain ethylmalonyl-CoA accumulated over time, peaking at about 5-fold at 9 hours, and then decreased when the culture resumed growth [Fig. 3.8B], a pattern similar to the wild type strain [Fig. 3.2]. In contrast, ethylmalonyl-CoA did not accumulate significantly in the strain over-expressing *ecm*.

Under these conditions, the growth lag decreased and earlier onset of growth in the over-expression strain correlated with a lack of accumulation of ethylmalonyl-CoA compared to the

control. In addition, the pattern of growth in the switchover was different, with initial growth of the *ecm*-over-expression strain slower than both the wild-type and the strain containing the empty vector. Later, growth of the *ecm*-over-expression strain increased to a rate faster than the controls. This shift in growth might indicate transient accumulation of a toxic intermediate, but the later, faster growth suggests that Ecm activity limits flux through the EMC pathway after growth occurs. This finding for C2 metabolism in a switchover experiment is consistent with the previous suggestion that the EMC pathway could contain a metabolic control point (29, 39).

These results also have implications for why activity of Ecm is low in the wild type strain during the transition. One explanation is that if the concentration of the required cofactor AdoCbl was at sub-saturating levels, Ecm would catalyze its reaction at a slower rate. However, overexpression of *ecm* results in increased Ecm activity, suggesting that AdoCbl is not limiting.

We hypothesized that a metabolic control point in the EMC pathway could be in place to restrict production of the toxic intermediate glyoxylate. However, metabolite analysis from the substrate switchover with the *ecm* over-expressing strain did not show accumulation of glyoxylate (data not shown). Another possibility is that Ecm functions to restrict carbon into the EMC pathway until downstream consumption enzymes are induced sufficiently to handle the full carbon flux. In support of this hypothesis, the genes encoding phosphoenolpyruvate carboxykinase (*pckA)*, enolase (*eno*), pyruvate dehydrogenase (*pdhABCD*), the glycine cleavage system (*gvcP*), the phosphoserine pathway (*serCAB*), and malyl-CoA lyase (*mclA2*), all of which encode enzymes utilized in C2 metabolism, are not up-regulated until 9 h after the substrate switchover. By limiting Ecm activity, the cell undergoes a longer growth lag than when this mechanism is overridden, but grows more quickly after the lag is finished. It is possible that in

the dynamic growth environment of leaf surfaces and/or soil (normal niches for this bacterium), the metabolic trade-off may provide an overall growth advantage.

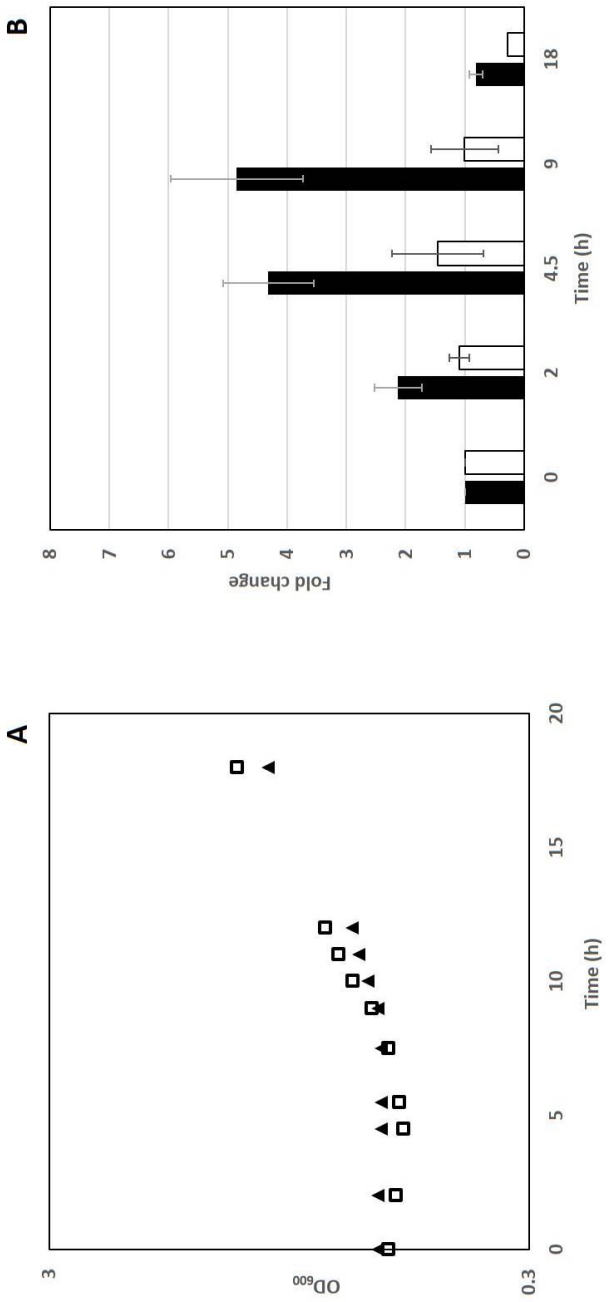

Figure 3.8 Switchover from succinate to ethylamine-growth over-expressing *ecm*

A) Growth response of the *ecm* over-expressing strain. Open squares are WT/pCM80_*ecm*. Black triangles are WT/pCM80. Time zero indicates when chemostat flows were stopped and 20 mM ethylamine was added to the bioreactor. Optical density (OD_{600}) was measured at the plotted time points. Growth curves are representative of a minimum of two biological replicates. B) Relative concentration of ethylmalonyl-CoA during the substrate switchover. Ethylmalonyl-CoA was measured by LC-MS. Measurements are normalized to time zero (0 h). Time is shown in hours. Black bars are WT/pCM80. Open bars are WT/pCM80_*ecm*. Measurements were calculated from a minimum of two injections each. Error bars are the standard deviation of two biological replicates.

3.5. Summary

Determining how metabolic balance is restored after a metabolic perturbation in the ethylmalonyl-CoA pathway is fundamental to understanding the physiology of EMC pathway-containing bacterial strains, such as the model methylotroph *M. extorquens*. Furthermore, insight into regulatory mechanisms of the EMC pathway could be vital to the engineering of such bacteria as a platform for producing value-added chemicals such as methylmalonyl-CoA, ethylmalonyl-CoA, mesaconic acid, methylsuccinic acid and butanol (10–12). It is difficult to study control of the EMC pathway during C1 metabolism, due to the regulation of the upstream steps (39, 40, 62). However, the experimental approach used here, involving C2 metabolism, successfully bypassed these reported regulatory mechanisms for C1 oxidation and assimilation and allowed us to focus specifically on the EMC pathway. The results presented here indicate that during a transition from succinate growth to growth on ethylamine, ethylmalonyl-CoA mutase transiently restricts carbon flux through the EMC pathway and the regulatory mechanism is likely transcriptional.

3.6. Acknowledgments

I performed all the experiments reported in this chapter except for extracting metabolites from the wild-type strain, which was carried out by Ceci Martinez-Gomez.

The work has been submitted to the Journal of Bacteriology for publication.

This work was funded by a grant from the DOE (DE-SC0006871).

We thank Mila Chistoserdova and Frances Chu for critical reading of the manuscript; Lisa Yuki-Bloch for her assistance with experiments; and Martin Sadilek for assistance with mass spectrometry.

Chapter 4. Conclusions and future directions

4.1. Conclusions

The work presented in this dissertation is a contribution within a larger framework to develop *M. extorquens* AM1 into a platform for the production of value-added chemicals using renewable feedstocks. Using systems-level approaches, we are investigating fundamental physiological questions, primarily how carbon flux through the assimilatory pathways for C1 and C2 metabolism are regulated. The primary goals of this study were to develop RNA-seq as a method for gene expression analysis and to identify possible metabolic control points in the EMC pathway.

As part of our suite of systems-level approaches, we developed a platform for RNA-seq transcriptomics analysis (Chapter 2). We compared previously obtained microarray gene expression datasets with newly acquired RNA-seq gene expression datasets for cultures grown under the same conditions. We found that although differences existed for specific genes, in general datasets from the two methods showed very similar trends and were highly comparable at that level. We also identified a subset of genes that had distinct expression levels depending on the method of data acquisition. We compared different rRNA depletion methods, and determined that higher depletion led to higher mapping of reads to protein-coding sequences, but caused biases in some read counts. However, we determined this was preferable to not removing rRNA sequences given that we can now predict how read counts will be affected and have the potential to multiplex while maintaining sufficient sequencing depth.

To identify possible metabolic control points in the EMC pathway, *M. extorquens* AM1 was challenged with a metabolic perturbation, a sudden change in growth substrate, and the response of the EMC pathway was investigated using systems level approaches (Chapter 3), including RNA-seq. We observed that after a transient period of metabolic imbalance, the cell acclimates to re-establish metabolic equilibrium. By this approach, ethylmalonyl-CoA mutase (Ecm) was identified as a metabolic control point in the EMC pathway. Ecm was shown to limit growth during the period of metabolic imbalance, and contribute to growth rate after homeostasis was re-established. The finding of Ecm as a control point is unusual, since reversible enzymes are not normally targets for regulation. However, our experimental approach was able to identify a transient, uncommon regulatory control. This work demonstrates the effectiveness of our experimental strategy, to perturb a system at metabolic steady-state and study acclimation of the culture over time, to discover transient regulatory strategies. Furthermore, it highlights the power of systems-level approaches to address directed hypotheses.

4.1.1. RNA-seq for gene expression analysis

RNA-seq leverages the power of high-throughput sequencing to quantify gene expression, providing a direct count of transcripts rather than fluorophore output. We experimentally demonstrated the high reproducibility of sequencing methods for assessing mRNA expression levels. We also identified biases in multiple rRNA depletion strategies, demonstrating a trade-off between total higher mapping to protein-coding sequences and loss of reads for particular genes. Finally, we validated RNA-seq as a method of transcriptome analysis

for *M. extorquens* AM1, updating the suite of tools available to study this model methylotroph for many years to come.

4.1.2. Metabolic control point in the EMC

The EMC pathway functions as the primary assimilatory pathway during growth on C2 compounds to generate glyoxylate, malate and succinate. We challenged a culture of wild-type *M. extorquens* AM1 with a sudden switch in growth substrate. Targeted metabolomics identified butyryl-CoA, which our data suggested is derived from ethylmalonyl-CoA, as accumulating during a transient lag phase after the switchover. Enzyme assays using cell-free extracts from the switchover experiment with purified Ecm and Epi demonstrated that Ecm, and not Epi, was the limiting activity. RNA-seq transcriptomics data for the expression of *ecm* suggested the control of Ecm expression is at least in part transcriptional. Over-expression of Ecm generated a severe, but temporary, growth defect in the same switchover experiment. Metabolite analysis showed that neither glyoxylate nor glycine accumulated during this time. Later, over-expressing *ecm* increased the growth rate. These results build on previous studies that proposed a metabolic control point in the EMC pathway and identify the actual control point (29, 39).

4.2. Future Directions

The control mechanism of Ecm expression is not yet completely defined. The RNA-seq dataset generated is the starting point for identifying candidate regulators. A promoter-fusion screen is a more holistic approach to identify regulators, and should be carried out. In addition, it is possible that regulation occurs at the level of small molecule effectors. Further activity assays

with other intermediates of the upstream steps to Ecm would rule out or identify potential inhibitors or activators.

The observation of two distinct growth rates for the Ecm over-expression strain in the substrate switchover is intriguing and should be followed up. Focused, targeted metabolite analysis during the period of slow growth could identify an accumulating intermediate. RNA-seq analysis in the over-expression strain could provide clues into the induction of other potential pathways, such as another glyoxylate consumption route or up-regulation of another branch of the C2 assimilation pathway. Testing growth phenotypes with mutants for genes in this pathway, most of which are already available, would provide supporting evidence for this model.

Over-expression of Ecm during methanol growth generated a growth defect, suggesting that it may be important for regulation of the EMC pathway during C1 growth as well. Testing the over-expression strain in a substrate switchover to methanol would demonstrate if Ecm functions as a control during C1 growth. Because of reported regulatory mechanisms in the methanol oxidation steps, it may be necessary to over-express Ecm in tandem with genes encoding enzymes necessary for overriding these mechanisms. These upstream mechanisms are currently being defined in our laboratory. By switching to methanol growth and higher flux through the EMC pathway, it may be possible to observe changes in the EMC pathway.

4.3. Final Comments

The results in this dissertation confirmed the hypothesis that the EMC pathway contains a metabolic control point. RNA-seq, a new tool for studying *M. extorquens* AM1, was developed and utilized in the investigation of the EMC pathway. The success of these studies represent another example of how systems-level approaches can be used to test directed hypotheses, and

detailed genetic and biochemical studies can validate conclusions drawn from them. This work has provided a confirmed target for strain manipulation and provided a powerful new tool for studying global gene expression, both of which will contribute to ongoing efforts to develop *M. extorquens* AM1 into a platform for biotechnology.

References

1. **Anthony C**. 1982. The biochemistry of methylotrophs. Academic Press, London.

2. **Chistoserdov AY, Tsygankov YD, Lidstrom ME**. 1991. Genetic organization of methylamine utilization genes from *Methylobacterium extorquens* AMI **173**.

3. **Green P**. 1992. The genus *Methylobacterium*, p. 2342–2349. *In* Balows, Truper A, Dworkin HG, Harder M, SK (ed.), The prokaryotes, 2nd ed. Springer, Berlin.

4. **Lidstrom ME**. 2006. Aerobic methylotrophic prokaryotes, p. 618–634. *In* Dworkin, M, Falkow, S, Schleifer, K-H, Stackebrandt, E (eds.), The Prokaryotes, 3rd ed. Springer Science-Business Media, New York, NY.

5. **Corpe W, Rheem S**. 1989. Ecology of the methylotrophic bacteria on living leaf surfaces. FEMS Microbiol. Lett. **62**:243–249.

6. **Hüve K, Christ MM, Kleist E, Uerlings R, Niinemets U, Walter a, Wildt J**. 2007. Simultaneous growth and emission measurements demonstrate an interactive control of methanol release by leaf expansion and stomata. J. Exp. Bot. **58**:1783–93.

7. **Sy A, Timmers ACJ, Knief C, Vorholt JA**. 2005. Methylotrophic metabolism is advantageous for *Methylobacterium extorquens* during colonization of *Medicago truncatula* under competitive conditions. Appl Environ Microbiol **71**:7245-7252.

8. **Harder W, Quayle JR**. 1971. Aspects of glycine and serine biosynthesis during growth of Pseudomonas AM1 on C compounds. Biochem. J. **121**:763–9.

9. **Salem, R, Hacking AJ, Quayle J**. 1974. Lack of malyl-CoA lyase in a mutant of *Psuedomonas* AM1. J. Gen. Microbiol. **81**:525–527.

10. **Alber BE**. 2011. Biotechnological potential of the ethylmalonyl-CoA pathway. Appl. Microbiol. Biotechnol. **89**:17–25.

11. **Schrader J, Schilling M, Holtmann D, Sell D, Filho MV, Marx A, Vorholt JA**. 2009. Methanol-based industrial biotechnology: current status and future perspectives of methylotrophic bacteria. Trends Biotechnol. **27**:107–15.

12. **Sonntag F, Buchhaupt M, Schrader J**. 2014. Thioesterases for ethylmalonyl-CoA pathway derived dicarboxylic acid production in *Methylobacterium extorquens* AM1. Appl. Microbiol. Biotechnol. **98**:4533–44.

13. **Chistoserdova L**. 2011. Modularity of methylotrophy, revisited. Environ. Microbiol. **13**:2603–22.

14. **Van Dien SJ, Strovas T, Lidstrom ME**. 2003. Quantification of central metabolic fluxes in the facultative methylotroph *Methylobacterium extorquens* AM1 using ^{13}C-label tracing and mass spectrometry. Biotechnol. Bioeng. **84**:45–55.

15. **Okubo Y, Skovran E, Guo X, Sivam D, Lidstrom ME**. 2007. Implementation of microarrays for *Methylobacterium extorquens* AM1. OMICS **11**:325–40.

16. **Guo X, Lidstrom ME**. 2006. Physiological analysis of *Methylobacterium extorquens* AM1 grown in continuous and batch cultures. Arch. Microbiol. **186**:139–49.

17. **Dien SJ Van, Lidstrom ME**. 2002. Stoichiometric model for evaluating the metabolic capabilities of the facultative methylotroph *Methylobacterium extorquens* AM1 , with application to reconstruction of C3 and C4 metabolism. Microbiology. **149**:601-609.

18. **Chistoserdov A**. 1994. Genetic organization of the *mau* gene cluster in *Methylobacterium extorquens* AM1: complete nucleotide sequence and generation and characteristics of *mau* mutants. J. Bacteriol. **176**:4052–4065.

19. **Chistoserdova L V, Lidstrom ME**. 1991. Purification and characterization of hydroxypyruvate reductase from the facultative methylotroph *Methylobacterium extorquens* AM1. J. Bacteriol. **173**:7228–32.

20. **Chistoserdova L V, Lidstrom ME**. 1994. Genetics of the serine cycle in *Methylobacterium extorquens* AM1: identification of *sgaA* and *mtdA* and sequences of *sgaA, hprA,* and *mtdA*. J. Bacteriol. **176**:1957–68.

21. **Chistoserdova L V, Lidstrom ME**. 1994. Genetics of the serine cycle in *Methylobacterium extorquens* AM1: identification, sequence, and mutation of three new genes involved in C1 assimilation, orf4, *mtkA,* and *mtkB*. J. Bacteriol. **176**:7398–404.

22. **Erb TJ, Berg IA, Brecht V, Müller M, Fuchs G, Alber BE**. 2007. Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway. Proc. Natl. Acad. Sci. U. S. A. **104**:10631–6.

23. **Korotkova N, Chistoserdova L, Kuksa V, Lidstrom ME**. 2002. Glyoxylate regeneration pathway in the methylotroph *Methylobacterium extorquens* AM1. J. Bacteriol. **184**:1750–1758.

24. **Peyraud R, Kiefer P, Christen P, Massou S, Portais J, Vorholt JA**. 2009. Demonstration of the ethylmalonyl-CoA pathway by using ^{13}C metabolomics. Proc. Natl. Acad. Sci. U.S.A. **106**:4846-4851.

25. **Erb TJ, Fuchs G, Alber BE**. 2009. (2S)-Methylsuccinyl-CoA dehydrogenase closes the ethylmalonyl-CoA pathway for acetyl-CoA assimilation. Mol. Microbiol. **73**:992–1008.

26. **Zarzycki J, Schlichting A, Strychalsky N, Müller M, Alber BE, Fuchs G**. 2008. Mesaconyl-coenzyme A hydratase, a new enzyme of two central carbon metabolic pathways in bacteria. J. Bacteriol. **190**:1366–74.

27. **Dunstan BPM, Anthony C, Drabble WT**. 1972. Microbial Metabolism of C_1 and C_2 Compounds Biochem. J. **128**:107–115.

28. **Okubo Y, Yang S, Chistoserdova L, Lidstrom ME**. 2010. Alternative route for glyoxylate consumption during growth on two-carbon compounds by *Methylobacterium extorquens* AM1. J. Bacteriol. **192**:1813–23.

29. **Schneider K, Peyraud R, Kiefer P, Christen P, Delmotte N, Massou S, Portais J-C, Vorholt J a**. 2012. The ethylmalonyl-CoA pathway is used in place of the glyoxylate cycle by *Methylobacterium extorquens* AM1 during growth on acetate. J. Biol. Chem. **287**:757–66.

30. **Erb TJ, Frerichs-Revermann L, Fuchs G, Alber BE**. 2010. The apparent malate synthase activity of *Rhodobacter sphaeroides* is due to two paralogous enzymes, (3S)-malyl-coenzyme A (CoA)/ β-methylmalyl-CoA lyase and (3S)-malyl-CoA thioesterase. J. Bacteriol. **192**:1249–1258.

31. **Korotkova N, Lidstrom ME**. 2005. Identification of genes involved in the glyoxylate regeneration cycle in *Methylobacterium extorquens* AM1 ,including two new genes , *meaC* and *meaD*. J. Bacteriol. **187**:1523-1526.

32. **Chistoserdova L, Laukel M, Vorholt JA, Lidstrom ME, Portais J**. 2004. Multiple formate dehydrogenase enzymes in the facultative methylotroph *Methylobacterium extorquens* AM1 are dispensable for growth on methanol. J. Bacteriol. **186**:22-28.

33. **Chistoserdova L, Crowther GJ, Vorholt J a, Skovran E, Portais J-C, Lidstrom ME**. 2007. Identification of a fourth formate dehydrogenase in *Methylobacterium extorquens* AM1 and confirmation of the essential role of formate oxidation in methylotrophy. J. Bacteriol. **189**:9076–81.

34. **Laukel M, Chistoserdova L, Lidstrom ME, Vorholt J a.** 2003. The tungsten-containing formate dehydrogenase from *Methylobacterium extorquens* AM1: Purification and properties. Eur. J. Biochem. **270**:325–333.

35. **Peyraud R, Schneider K, Kiefer P, Massou S, Vorholt J a, Portais J-C**. 2011. Genome-scale reconstruction and system level investigation of the metabolic network of *Methylobacterium extorquens* AM1. BMC Syst. Biol. **5**:189.

36. **Harder W, Quayle JR**. 1971. The biosynthesis of serine and glycine in *Pseudomonas* AM1 with special reference to growth on carbon sources other than C1 compounds. Biochem. J. **121**:753–62.

37. **Eisenhut M, Bauwe H, Hagemann M**. 2007. Glycine accumulation is toxic for the cyanobacterium *Synechocystis* sp. strain PCC 6803, but can be compensated by supplementation with magnesium ions. FEMS Microbiol. Lett. **277**:232–7.

38. **Murphy CB, Martell AE**. 1957. Article: Metal chelates of glycine and glycine peptides J. Biol. Chem. **226**:37–50.

39. **Skovran E, Crowther GJ, Guo X, Yang S, Lidstrom ME**. 2010. A systems biology approach uncovers cellular strategies used by *Methylobacterium extorquens* AM1 during the switch from multi- to single-carbon growth. PLoS One **5**:e14091.

40. **Martinez-Gomez NC, Nguyen S, Lidstrom ME**. 2013. Elucidation of the role of the methylene-tetrahydromethanopterin dehydrogenase MtdA in the tetrahydromethanopterin-dependent oxidation pathway in *Methylobacterium extorquens* AM1. J. Bacteriol. **195**:2359–67.

41. **Smejkalová H, Erb TJ, Fuchs G**. 2010. Methanol assimilation in *Methylobacterium extorquens* AM1: demonstration of all enzymes and their regulation. PLoS One **5**:e13001.

42. **Vuilleumier S, Chistoserdova L, Lee M-C, Bringel F, Lajus A, Zhou Y, Gourion B, Barbe V, Chang J, Cruveiller S, Dossat C, Gillett W, Gruffaz C, Haugen E, Hourcade E, Levy R, Mangenot S, Muller E, Nadalig T, Pagni M, Penny C, Peyraud R, Robinson DG, Roche D, Rouy Z, Saenampechek C, Salvignol G, Vallenet D, Wu Z, Marx CJ, Vorholt J a, Olson M V, Kaul R, Weissenbach J, Médigue C, Lidstrom ME**. 2009. *Methylobacterium* genome sequences: a reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources. PLoS One **4**:e5584.

43. **Hu B, Lidstrom M**. 2012. CcrR, a TetR family transcriptional regulator, activates the transcription of a gene of the ethylmalonyl coenzyme A pathway in *Methylobacterium extorquens* AM1. J. Bacteriol. **194**:2802–8.

44. **Kalyuzhnaya MG, Lidstrom ME**. 2003. QscR, a LysR-type transcriptional regulator and CbbR homolog , is Involved in regulation of the serine cycle genes in *Methylobacterium extorquens* AM1. J. Bacteriol. **185**:1229–1235.

45. **Kalyuzhnaya MG, Lidstrom ME**. 2005. QscR-mediated transcriptional activation of serine cycle genes in *Methylobacterium extorquens* AM1. J. Bacteriol. **187**:7511–7517.

46. **Yang S, Sadilek M, Synovec RE, Lidstrom ME**. 2009. Liquid chromatography-tandem quadrupole mass spectrometry and comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry measurement of targeted metabolites of *Methylobacterium extorquens* AM1 grown on two different carbon sources. J. Chromatogr. A **1216**:3280–9.

47. **Marx CJ, Lidstrom ME**. 2001. Development of improved versatile broad-host-range vectors for use in methylotrophs and other Gram-negative bacteria. Microbiology **147**:2065–75.

48. **Guo X, Lidstrom ME**. 2008. Metabolite profiling analysis of *Methylobacterium extorquens* AM1 by comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry. Biotechnol. and Bioeng. **99**:929–940.

49. **Wang Z, Gerstein M, Snyder M**. 2010. RNA-Seq : a revolutionary tool for transcriptomics. Nat. Rev. Genet. **10**:57–63.

50. **Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B**. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods **5**:621–8.

51. **Li H, Durbin R**. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25**:1754–60.

52. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R**. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**:2078–9.

53. **Anders S, Pyl PT, Huber W**. 2014. HTSeq - A Python framework to work with high-throughput sequencing data. Bioinformatics 1–4.

54. **Storey JD, Tibshirani R**. 2003. Statistical significance for genomewide studies. Proc. Natl. Acad. Sci. U. S. A. **100**:9440–5.

55. **Erb TJ, Rétey J, Fuchs G, Alber BE**. 2008. Ethylmalonyl-CoA mutase from *Rhodobacter sphaeroides* defines a new subclade of coenzyme B12-dependent acyl-CoA mutases. J. Biol. Chem. **283**:32283–93.

56. **Korotkova N, Lidstrom ME**. 2004. MeaB is a component of the methylmalonyl-CoA mutase complex required for protection of the enzyme from inactivation. J. Biol. Chem. **279**:13652–8.

57. **Kiefer P, Buchhaupt M, Christen P, Kaup B, Schrader J, Vorholt JA**. 2009. Metabolite profiling uncovers plasmid-induced cobalt limitation under methylotrophic growth conditions. PLoS One **4**:e7831.

58. **Niedhardt F, Ingraham J, Schaecter M**. 1990. Physiology of the Bacterial Cell. A Molecular Approach. Sinauer Associates, Sunderland, MA.

59. **Kellermeyer RW, Allen SHG, Wood HG**. 1964. Methylmalonyl-CoA Isomerase: IV. Purification and propterties of the enzyme from *Propionibacterium*. J. Biol. Chem. **239**:2562–2569.

60. **Mancia F, Smith GA, Evans PR**. 1999. Crystal structure of substrate complexes of methylmalonyl-CoA mutase. Biochemistry **38**:7999–8005.

61. **Zhang M, FitzGerald K a, Lidstrom ME**. 2005. Identification of an upstream regulatory sequence that mediates the transcription of *mox* genes in *Methylobacterium extorquens* AM1. Microbiology **151**:3723–8.

62. **Marx CJ, Van Dien SJ, Lidstrom ME**. 2005. Flux analysis uncovers key role of functional redundancy in formaldehyde metabolism. PLoS Biol. **3**:e16.

63. **Schneider K, Skovran E, Vorholt JA**. 2012. Oxalyl-CoA reduction to glyoxylate is the preferred route of oxalate assimilation in *Methylobacterium extorquens* AM1. J. Bacteriol. **194**:3144–3155.

64. **Blackmore M a, Quayle JR**. 1970. Microbial growth on oxalate by a route not involving glyoxylate carboligase. Biochem. J. **118**:53–9.

65. **Chou H-H, Berthet J, Marx CJ**. 2009. Fast growth increases the selective advantage of a mutation arising recurrently during evolution under metal limitation. PLoS Genet. **5**:e1000652.

66. **Gibson DG**. 2011. Enzymatic assembly of overlapping DNA fragments. Methods Enzymol. **498**:349–61.

67. **Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO**. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods **6**:343–5.

68. **Nunn DN, Lidstrom ME**. 1986. Isolation and complementation analysis of 10 methanol oxidation mutant classes and identification of the methanol dehydrogenase structural gene of *Methylobacterium* sp. strain AM1. J. Bacteriol. **166**:581–90.

Appendix A

The following are initial studies defining the response of metabolites measured by GC-MS and LC-MS during a substrate switchover from succinate-growth to ethylamine growth while over-expressing *ecm*. I performed all of the experiments reported here.

Introduction

In this thesis, ethylmalonyl-CoA mutase was shown to operate as a metabolic control point during the switch from succinate-growth to ethylamine-growth. In addition, this control was overridden by over-expressing *ecm* and repeating the substrate switchover experiment. Our working hypothesis was that having a metabolic control point in the EMC pathway could be an advantage, by restricting carbon flux to glyoxylate, which is a toxic metabolic intermediate (8, 9). By overriding the Ecm checkpoint, we anticipated that flux through the middle of the EMC pathway (from crotonyl-CoA to β-methylmalyl-CoA) might increase, causing glyoxylate and/or glycine to accumulate. Accumulation of a metabolite(s) downstream of ethylmalonyl-CoA during the lag phase after the substrate switchover would indicate an imbalance between production and consumption metabolic intermediates suggesting that an additional metabolic control point was operating.

Methods

Growth conditions

M. extorquens AM1 was grown in a 2.2-liter bench-top BioFlo/Celligen 115 Modular BioReactor-Fermentor (New Brunswick Scientific) with a starting volume of 2.0 L with

succinate, methanol or ethylamine as the growth substrate (16, 28, 39, 40). The following concentrations of antibiotics were used: kanamycin (50 ug/ml), tetracycline (1.5 ug/ml). 13 uM cobalt chloride was added to the culture medium to negate inhibitory effects of pCM80 (57).

Chemostat cultivation

The substrate switchover from succinate-growth to ethylamine growth was executed as reported (Chapter 2).

Metabolite analysis by LC-MS and GC-MS

Samples were taken immediately before the substrate switchover (time zero) and at a series of times after the switchover (2, 4.5, 9, and 18 hours) for analysis via LC-MS and GC-MS. Metabolite extraction and data analysis was conducted as reported [Chapter 2].

Initial Results

GC-MS analysis of amino and organic acids

Relative concentrations of the EMC pathway metabolites methylsuccinate and mesaconate, along with the organic acids malate and glyoxylate, and the amino acids glycine and serine were measured [Fig. A.1]. In both the WT/pCM80 (vector control) and WT/pCM80_*ecm* strains methylsuccinate levels remained relatively unchanged throughout the time course. Mesaconate measurements had a similar pattern in WT/pCM80; whereas, in WT/pCM80_*ecm* mesaconate increased 4-fold at 18 hours. Malate measurements had the same trend for both strains, increasing incrementally over the time course, reaching a maximum relative

concentration of around 3-fold by 18 hours. Glyoxylate, glycine and serine measurements all showed the same general pattern. In WT/pCM80, the relative concentrations of these compounds changed very little. In WT/pCM80_*ecm*, however, all three compounds decreased at 2 hours, and by 4.5 hours had decreased further to 2-fold. After 4.5 hours the relative concentrations increased slightly.

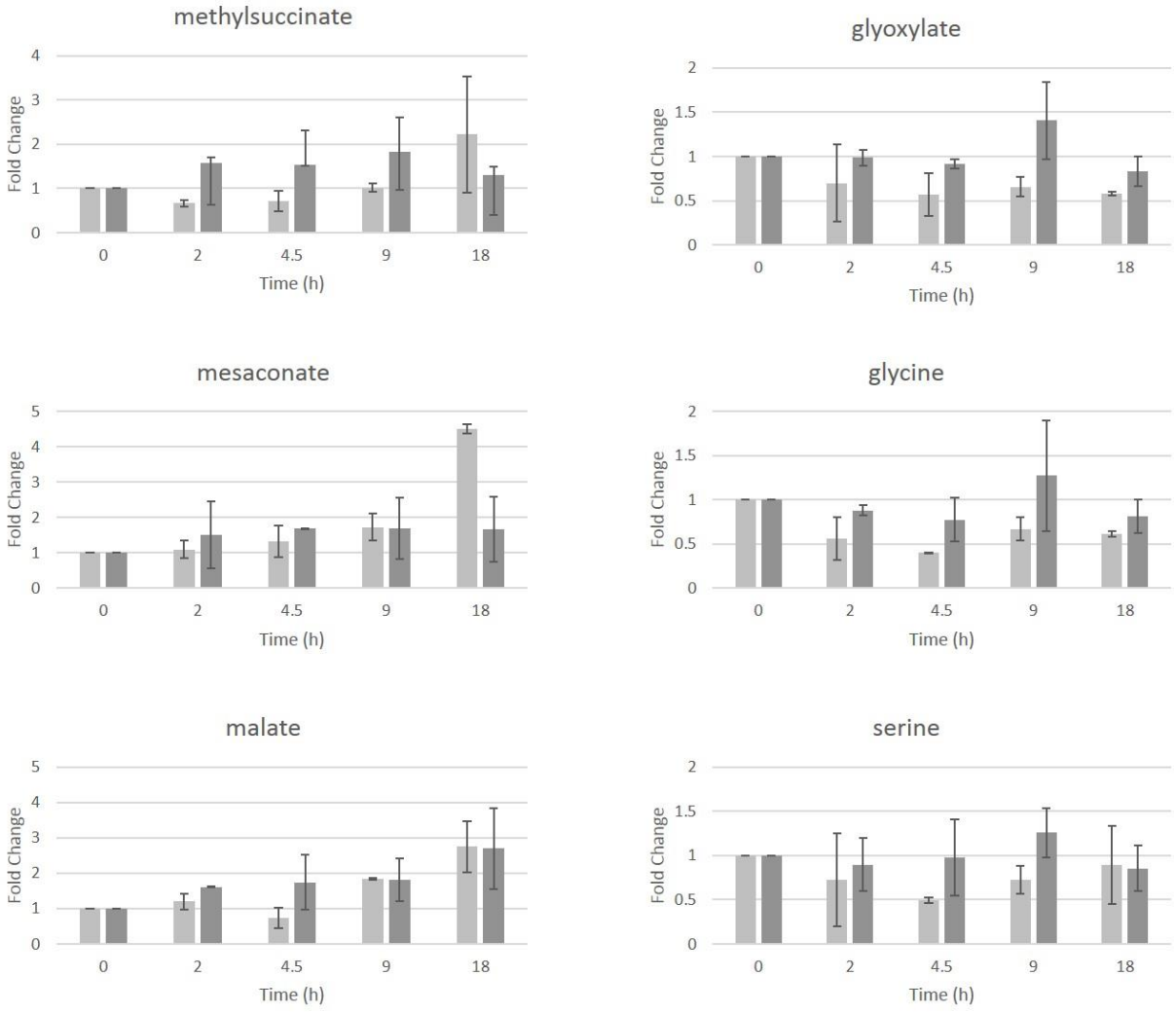

Figure A.1 GC-MS analysis of metabolite concentrations during the switch from succinate growth to ethylamine growth over-expressing *ecm*

The relative concentration was normalized to the relative concentration at time zero. Metabolites for WT/pCM80 are shown in dark gray. Metabolites for WT/pCM80_*ecm* are shown in light gray. All values shown are from two biological replicates. Each replicate is from a minimum of two injections. Error bars represent the standard deviation.

LC-MS analysis of CoA intermediates

CoA intermediates of the EMC pathway were measured by LC-MS [Fig. A.2]. The relative concentrations of acetyl-CoA changed very little, in either strain, over the course of the experiment. The relative concentrations of 3-hydroxybutyryl-CoA in the control strain dropped slightly at 2, 4.5 and 9 hours, then returned to baseline levels at 18 hours. For the over-expression strain, 3-hydroxybutyryl-CoA decreased more than 2-fold by 4.5 hours and nearly 4-fold by 18 hours. Relative concentrations of propionyl-CoA were decreased at 2 and 4.5 hours in both strains; however, the decrease was 2-3-fold greater in the over-expression strain. At 9 hours propionyl-CoA increased above initial levels in the strain over-expressing *ecm*. Propionyl-CoA did not increase above baseline in the control strain until 18 hours.

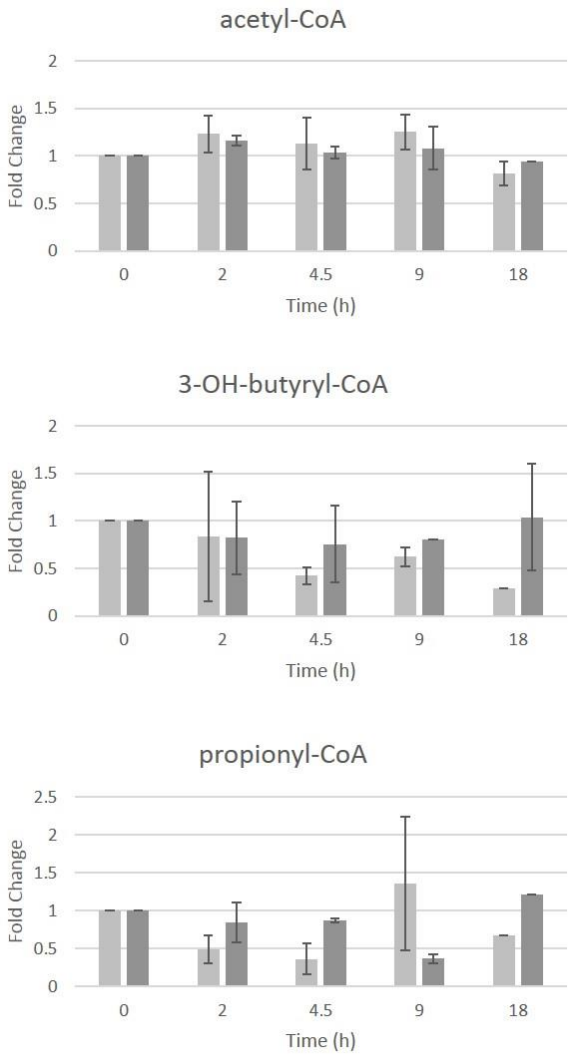

Figure A.2 LC-MS analysis of metabolites during the switch from succinate growth to ethylamine growth over-expressing *ecm*

The relative concentration was normalized to the relative concentration at time zero. Metabolites for WT/pCM80 are shown in dark gray. Metabolites for WT/pCM80_*ecm* are shown in light gray. All values shown are from two biological replicates. Each replicate is from a minimum of two injections. Error bars represent the standard deviation.

Discussion

Targeted metabolomics of the EMC and downstream assimilation pathways was used to investigate how the cell acclimates to a switch from succinate to ethylamine growth when the Ecm control point is no longer operating. By monitoring fluctuations in metabolite pools, we hoped to gain insight into why the growth phenotype of WT/pCM80_*ecm* was different from the controls. Acetyl-CoA concentrations did not change significantly in either strain over the course of the experiment, indicating that oxidation of ethylamine to acetate and the subsequent activation to acetyl-CoA was not influenced by over-expressing *ecm*. 3-hydroxybutyryl-CoA pools did not mirror what was observed in the substrate switchover experiment with the wild-type strain, as the compound did not accumulate in either strain. This suggests that this compound may not be a metabolic control point as was proposed (39). By 18 hours 3-hydroxybutyryl-CoA beings to accumulate again in the control strain, while in the over-expression strain it continues to decrease. This indicates that there could be higher consumption of 3-hydroxybutyryl-CoA in the overexpression strain, which correlates with reported evidence that Ecm is a metabolic control point.

The intermediates directly downstream of ethylmalonyl-CoA were analyzed to determine if another metabolic checkpoint was active in EMC prior to the split of β-methylmalyl-CoA to glyoxylate and propionyl-CoA. Neither mesaconate nor methylsuccinate accumulated during the first 4.5 hours after the substrate switchover, indicating that Msd and Mcd are not restricting consumption during this time. Relative concentrations of both compounds did not change significantly, with the exception of mesaconate at 18 hours. This accumulation corresponds with the time of highest expression of *mcd*, encoding mesaconyl-CoA dehydratase, the consuming enzyme. Presumably, at this time point Mcd activity would be highest. It is possible that

accumulation is a reflection of the downstream metabolite, β-methylmalyl-CoA, accumulating and feeding back to mesaconyl-CoA since Mcd is a reversible enzyme. Relative concentrations of propionyl-CoA were decreased in the over-expression strain until 9 hours. The increase in the pool size at 9 hours could suggest more rapid production of the compound due to higher Ecm activity and/or lower propionyl-CoA carboxylase activity.

The relative concentrations of glyoxylate, glycine and serine initially decreased (at 2 and 4.5 hours) and then slowly increased as the culture resumed growth. Neither glyoxylate nor glycine, both toxic intermediates, accumulated between 4.5 and 9 hours, when the overexpression strain was growing with a 29 hour doubling time, which suggests that accumulation of either compound was not the cause of slow growth. As the culture changed to the faster growth rate (9 hours), these compounds, along with serine, began to accumulate but only by small relative amounts. A possible explanation for the lack of accumulation of these toxic compounds is that another pathway capable of consuming glyoxylate was induced. A pathway for the metabolism of oxalate to glyoxylate was recently reported for *M. extorquens* AM1 (63, 64). Activity of oxalyl-CoA reductase (PanE2), a key enzyme in the pathway that converts oxalyl-CoA to glyoxylate, is induced by glyoxylate (Skovran, personal communication). Oxalyl-CoA can be decarboxylated to formyl-CoA, which can be converted to formate. Formate can be oxidized to CO_2 by formate dehydrogenase. Currently, a detailed genetic analysis of this pathway is being completed [Skovran et al., in prep]. Involvement of this pathway in the substrate switchover would suggest that it may be operating to rescue the cell from glyoxylate accumulation.

Lastly, there was no observable difference in the response of malate between the strain over-expressing *ecm* and the control strain. In both strains the relative malate concentration

steadily increased, albeit very slowly. This suggests that production and consumption of malate is not differentially affected enough to cause a metabolic imbalance. This also implies that activities of the enzymes for the conversion of propionyl-CoA to malate are sufficient to convert increased substrate to product and prevent accumulation in the *ecm* over-expressing strain, which corroborate *in vivo* the *in vitro* activities reported for this part of the pathway (41).

Together the metabolite measurements show evidence of the effects of overriding the Ecm control point, but they do not provide clear insight into how metabolic equilibrium is restored more quickly and how this affects growth rate. Many of the metabolite profiles are supporting evidence of higher Ecm activity, and suggest higher flux through the EMC pathway. However, downstream metabolites do not accumulate. It is possible that accumulation, particularly, of the toxic intermediate glyoxylate, occurs, but if so our time points did not capture it. Because the over-expression strain grows very poorly during for the first few hours after it resumes growth, then grows better than both the wild-type and control strains, it is possible that glyoxylate or another toxic intermediate accumulate transiently and then induce another consumption pathway until the primary consumption pathways can be fully induced. Further investigation is needed to test these hypotheses.

Appendix B

The following are initial studies investigating the regulation of Epi. I performed all of the experiments reported here.

Introduction

Epi catalyzes the isomerization of (2S)-ethylmalonyl-CoA to (2R)-ethylmalonyl-CoA. It was identified as a candidate control point after initial examination of the metabolic response to a change in growth substrate, as mentioned. Here I present preliminary results and observations investigating the regulation of Epi and a potential role in regulation of the EMC pathway.

Methods

Growth conditions

M. extorquens AM1 was grown in a 2.2-liter bench-top BioFlo/Celligen 115 Modular BioReactor-Fermentor (New Brunswick Scientific) with a starting volume of 2.0 L with succinate, methanol or ethylamine as the growth substrate (16, 39, 40). The following concentrations of antibiotics were used: kanamycin (50 ug/ml).

Construction of Epi over-expression and epi promoter fusion strains using Gibson Assembly

The gene encoding ethylmalonyl-CoA/methylmalonyl-CoA epimerase (*epi*) was PCR amplified using Phusion High Fidelity DNA Polymerase (New England Biolabs, MA, USA) with primers designed with 30 base pair overlapping regions homologous to regions flanking the multiple cloning site of the expression plasmid pHC61 (65). The vector backbone was amplified

as linear DNA with overlapping regions to the insert. pHC61_*epi* was generated by Gibson Assembly (Table 5.1) (66, 67) using Gibson Assembly Master Mix (New England Biolabs, MA, USA) to generate pNG243. The construct was transformed into *E. coli* via electroporation to generate strain NG243. NG243 was mated into the wild-type *M. extorquens* AM1 and *ccrR* backgrounds as reported (47). Similarly, the 150 bp intergenic region upstream of *epi* was cloned into the GFP-reporter plasmid pHC42 using Gibson assembly (65). All constructs were verified by Sanger sequencing (High Throughput Genomics Center, WA, USA).

Table B.1 *M extorquens* AM1 strains and plasmids

Strain or plasmid	Description	Reference
Strain		
AM1	Rif^{S} derivative	(68)
Plasmid		
pHC42	GFP-containing reporter plasmid	(65)
pHC61	*M. extorquens* AM1 expression vector (Mtac)	(65)
pNG243	pCM42 with epi promoter region	This study
pNG208	pHC61 with *epi*	This study

Initial Results

Over-expression of Epi generates a growth defect in both methanol and ethylamine

Epi was over-expressed in pHC61 under a modified tac promoter shown to generate high expression in *M. extorquens* AM1 (65). Previous attempts to grow cultures on ethylamine in

shake flasks failed in our hands, with the strains flocculating at an OD_{600} of 0.4 and forming large flakes as the cultures continue to grow resulting in variable optical density measurements (data not shown). Because acetaldehyde is produced as an oxidation product of ethylamine, it was suspected to be the cause of flocculation. For this reason, growth curves in which ethylamine was the growth substrate were conducted in a bioreactor flushing 1.5 Standard Liter Per Minute (SLPM) of air into the headspace to reduce acetaldehyde vapors. By this method, growth curves on ethylamine were executed in a highly reproducible manner. Cultures grown in a bioreactor on ethylamine did not flocculate until reaching an OD_{600} of 1-1.2. Cultures grown on methanol in a bioreactor grew normally, albeit with a slightly faster doubling time.

The strain over-expressing *epi* had a growth defect in batch culture in a bioreactor with either methanol or ethylamine as the growth substrate when compared to the wild-type strain harboring the empty vector. The control strain grew with a specific growth rate of 0.135 h^{-1} on methanol and 0.087 h^{-1} on ethylamine, while the strain harboring pHC61_*epi* grew with a specific growth rate of 0.101 h^{-1} on methanol and 0.066 h^{-1} on ethylamine [Fig. B1].

Analysis of the promoter region and transcriptional profile of epi

The RNAseq profiles of *epi* and *ccrR* from the substrate switchover experiment were examined and found to be strikingly similar [Fig. B2]. CcrR is a TetR-family transcriptional regulator that activates expression from *ccr* (encoding crotonyl-CoA reductase), so the upstream intergenic region of *epi* was examined by eye for a palindromic sequence that is characteristic of this type of regulator. A palindrome (underlined) was located 113 bp upstream of the translational start of *epi*: *CGCGCC*CGCTTCGCCACGGCAAGCG*GGCGCG*. The palindrome contained six base pairs sequences at the 5' and 3' ends with 100% sequence identity to the ends

of the palindrome identified as the CcrR binding site in the *ccr* promoter region (italics) (43).

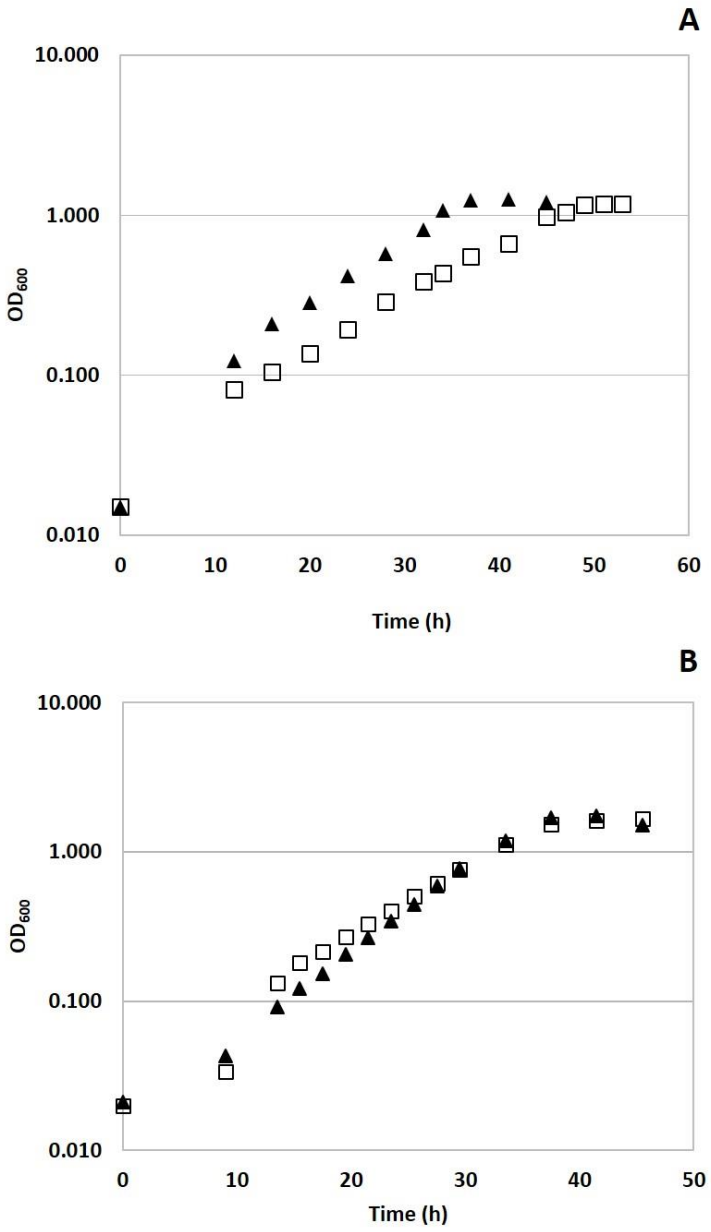

Figure B.1 Over-expressing *epi* generates a growth defect on methanol and ethylamine

Wild-type *M. extorquens* AM1 harboring pHC61_*epi* was grown in a bioreactor with either ethylamine (A) or methanol (B) as the growth substrate. Optical density (OD_{600}) was measured at the time points shown. Open squares are WT/pHC61_*epi* and filled triangles are WT/pHC61. Each growth curve is representative of two biological replicates.

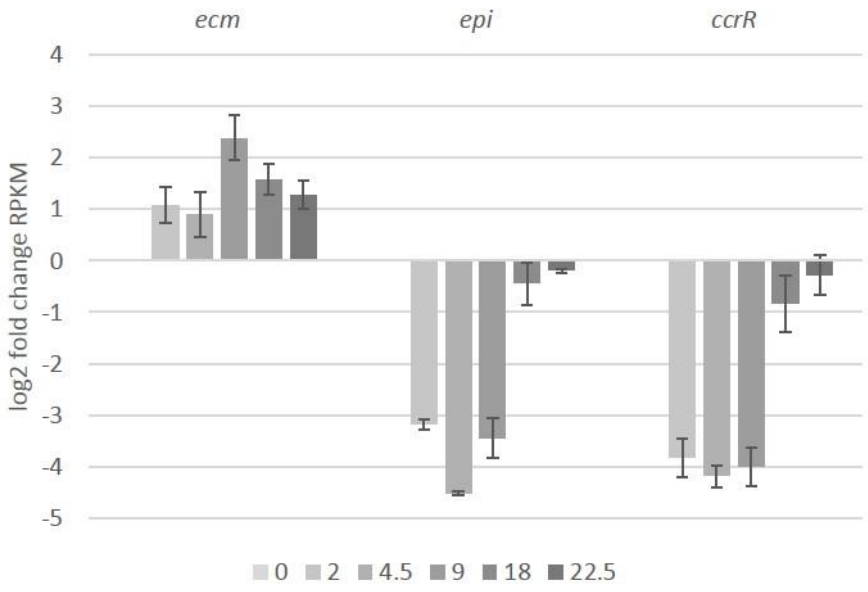

Figure B.2 RNA-seq profile of *ecm*, *epi*, and *ccrR*

RPKM for genes encoding enzymes for the conversion of ethylmalonyl-CoA to methylsuccinyl-CoA are shown. Time in hours after the addition of ethylamine is indicated by grayscale. The RPKM was normalized to the read count at time zero (0 h). Error bars are the standard deviation of two biological replicates.

Discussion

Analysis of the RNA-seq dataset from the substrate switchover showed that the pattern of changes to transcript levels of *epi* during the time course was highly similar to the pattern of changes for *ccrR*. The promoter region of *epi* was found to contain a palindromic sequence that could be a putative TetR-family regulator binding site. Preliminary promoter fusion assays to determine if CcrR interacts with the *epi* promoter were inconclusive. Even though it has been shown that during the switchover Epi is not a limiting enzyme for the conversion of (2S)-ethylmalonyl-CoA to (2R)-ethylmalonyl-CoA, it is possible that Epi activity regulates in part another reaction in the pathway and/or is important for control of carbon flux through the pathway during batch growth. Epi is also responsible for the interconversion of (2S)-methylmalonyl-CoA to (2R)-methylmalonyl-CoA, and methylmalonyl-CoA mutase activity was not measured during the switchover experiment. It has been shown that Epi has similar K_m and V_{max} values towards both substrates, suggesting that it is a promiscuous enzyme (3). The severity of the growth defects observed for both ethylamine growth and methanol growth with the Epi over-expression strain were the same, suggesting that the defects could be due to the same mechanism. The EMC pathway is fully functional during growth on either substrate, and Epi plays a dual role in both cases. One possible explanation for the growth defects observed is that over-expressing *epi* disrupts a balance between the ethylmalonyl-CoA and methylmalonyl-CoA stereoisomerization reactions. Another possibility is that Epi forms a complex with either or both of the mutases. Over-expressing *epi* in this case could generate an imbalance of enzymes, leading to overall decreased activity through these biochemical steps. Further biochemical analysis would be needed to verify this however, as there is no evidence of any protein-protein interaction at this time.