Formaldehyde Metabolism in *Methylobacterium extorquens* AM1

Christopher James Marx

A dissertation submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2003

Program Authorized to Offer Degree: Department of Microbiology

UMI Number: 3091036

UMI Microform 3091036

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

In presenting this dissertation in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make copies freely available for inspection. I further agree that extensive copying of the dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to Proquest Information and Learning, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature 

Date 6/9/03

University of Washington

Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

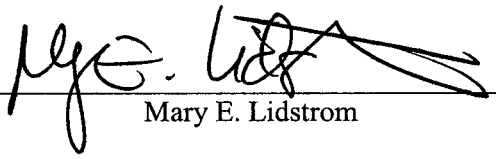
Christopher James Marx

and have found that it is complete and satisfactory in all respects,

and that any and all revisions required by the final

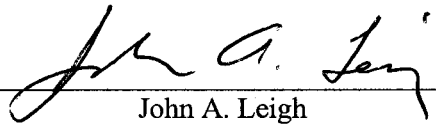
examining committee have been made.

Chair of Supervisory Committee:

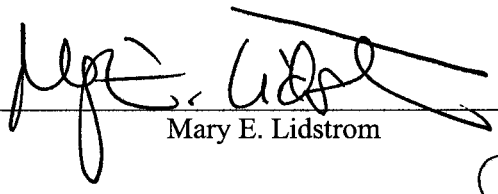


Mary E. Lidstrom

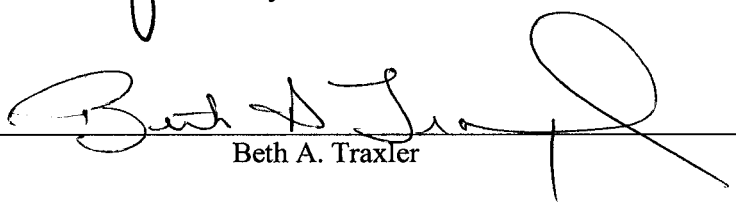
Reading Committee:



John A. Leigh



Mary E. Lidstrom



Beth A. Traxler

Date: JUNE 6th, 2003

University of Washington

Abstract

Formaldehyde Metabolism in *Methylobacterium extorquens* AM1

Christopher James Marx

Chairperson of the Supervisory Committee
Professor Mary E. Lidstrom
Departments of Microbiology and Chemical Engineering

Methylotrophic bacteria are capable of growth on single-carbon compounds as their sole source of carbon and energy. This metabolism entails the production of formaldehyde as a central intermediate. *Methylobacterium extorquens* AM1 possess two pterin-linked C_1 transfer pathways that are critical for methylotrophic growth. These utilize either tetrahydrofolate (H_4F) or tetrahydromethanopterin (H_4MPT) as C_1 carriers. Through genetic and physiological analyses, as well as ^{14}C label tracing studies, I have determined that the H_4MPT-linked pathway is required for formaldehyde oxidation and detoxification. The H_4F-linked pathway is also required for methylotrophic growth, but apparently not as a second formaldehyde oxidation pathway. In order to test whether the H_4F may actually play an assimilatory role, a deuterium label tracing method was devised. By examining the isotopomer distribution of serine-derived fragments I could determine whether the C_1 units that entered the assimilatory serine cycle came from methylene-H_4F generated directly from formaldehyde, or from the conversion of formate to methylene-H_4F by the H_4F-linked pathway. This technique led to the discovery that the H_4F pathway functions in an assimilatory role. Furthermore, the deuterium labeling experiments were combined with ^{14}C labeling to determine the flux through each branch of methylotrophic metabolism during the transition of cultures to and from methylotrophic growth. This has led to a dynamic picture of C_1 metabolism that has elucidated an elegant metabolic loop through which *M. extorquens* AM1 maintains formaldehyde below toxic levels while balancing the distribution of C_1 units to assimilatory and dissimilatory metabolism.

Acknowledgements

Many people have contributed immeasurably to my graduate school experience, only the foremost of which I have time and space to thank here. First, I must extend the deepest of gratitude to Mary for taking me into her laboratory and providing an environment ripe for success. The independence I have been granted from early on has given me the opportunity for many rewards, along with bountiful failures that do not appear in these pages. Both, however, have been critical in my development as a scientist and as a person. Second, I would like to thank the members of the Lidstrom lab. You are a tremendous group, but in particular Mila, Sergey, Yoko, Heather, and Brooke have contributed above and beyond the call of duty. In addition, the faculty, staff, and students from the Microbiology Department have been wonderful teachers, advisors, and friends, especially the members of my entering class, Molly, Heather, Vanya, and Carolyn. Finally, my collaborations with Madeline Rasche, Rolf Thauer, and Julia Vorholt, as well as my work with the help of Martin Sadilek, have been incredibly fruitful and enjoyable experiences. I wish all of you the very best in your professional and personal lives.

Most of all, it has been my family and friends to maintain have allowed me to maintain at least a small shred of sanity throughout the difficulties of graduate school. Thanks to my Mom, my Dad and Lee, Josiah and Annika, Gary and Ingrid, and Jon and Andrea. Finally, I must thank my wife, Maria. She has been there to raise a glass to share the celebration of my brightest moments, as well as helping me through what has been the most difficult of time of my life. Together we have made it through our time in Seattle, as we will throughout the years ahead.

Dedication

I dedicate this work to my grandparents, Vern and LaFaye, and Harpo and Susan.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

Methylotrophic bacteria can aerobically utilize compounds with no carbon-carbon bonds as their sole source of carbon and energy. These single-carbon (C_1) substrates include methane, methanol, halomethanes, methylated amines, and methylated sulfur species. Given the abundance and ubiquity of methylotrophs, they play a critical role in the biochemical cycle of these compounds and are the primary biological sink for methane and other C_1 greenhouse gasses (59). Interestingly, the primary source of some of these compounds, such as methane, is also of microbial nature. Methanogenic archaea account for majority of the methane released to the atmosphere. These diverse organisms generate methane in anaerobic environments from H_2-dependent reduction of CO_2, disproportionation of methanol, methylated amines, and methylated sulfur species, or cleavage of acetate (117).

Methylotrophy and methanogenesis compared:

Methylotrophy and methanogenesis are both examples of specialized metabolism, such that the largest carbon fluxes occur in pathways unique to these organisms, rather than widely conserved pathways such as glycolysis or the tricarboxylic acid cycle. Methylotrophy and methanogenesis result in the oxidation or reduction of C_1 compounds, respectively, however the metabolic pathways involved are not simply mirror images of each other (74). Hydrogenotrophic methanogenesis (Fig. 1.1), for example, which utilizes CO_2 and H_2 to produce methane, does not generate any free C_1 compounds of intermediate oxidation states: formate, formaldehyde, or methanol. Rather, specific C_1 carrier molecules, methanofuran (MFR), tetrahydromethanopterin (H_4MPT), and coenzyme M (CoM), are utilized (117). CO_2 is first activated by formyl-MFR dehydrogenase to form formyl-MFR. A formyltransferase then transfers the formyl group from MFR to H_4MPT, which is a pterin compound analogous to folate (34, 74). Formyl-H_4MPT is converted to methenyl-H_4MPT by methenyl-H_4MPT cyclohydrolase, and this compound is reduced to the methylene level by either a H_2- or F_{420}-dependent methylene-H_4MPT

dehydrogenase. The F_{420}-dependent methylene-H_4MPT reductase then generates methyl-H_4MPT, which is transferred to CoM by a methyltransferase. Finally, methyl-CoM reductase catalyzes the release of free methane and the concomitant generation of the CoM-coenzyme B heterodisulfide (117), which is then cleaved by heterodisulfide reductase. For those methanogens capable of utilizing acetate, methanol, methylated amines, or methylated sulfur species, methyl groups enter the core C_1 transfer pathway at the level of methyl-H_4MPT or methyl-CoM and are disproportionated to both methane and CO_2 (117). Methanogenesis has only been found in the Euryarchaeota, where they form a diverse but phylogenetically coherent group. Thus far, only isolated examples of non-methanogens, such as the sulfate-reducing archaeon *Archaeoglobus fulgidus* (105), have been found to belong to these clades. Similarly, the core methanogenic pathway of C_1 transfers described above is conserved amongst all of these organisms. This suggests that the core methanogenesis pathway was present prior to the divergence of the known Euryarchaeota, and thus that the common ancestor of this clade performed methanogenesis.

Methylotrophy differs from methanogenesis in several key aspects. First, the free C_1 compounds of intermediate oxidation states, methanol (for growth on methane), formaldehyde, and formate are pathway intermediates (Fig. 1.1). Methylotrophy consists of three stages: primary oxidation of reduced carbon compounds to formaldehyde, assimilation of formaldehyde into cell material, and oxidation of formaldehyde to formate and then CO_2. Unlike the methanogenic archaea, methylotrophic bacteria do not comprise a phylogenetically coherent group. Representatives are found in the α, β, and γ divisions of the proteobacteria and the Gram positive bacteria and are often closely related to non-methylotrophic organisms (69). Additionally, the various stages of methylotrophy are performed by different, and often multiple metabolic modules in various methylotrophs. Modules refer to groups of gene products that collectively contribute to a common function, such that theoretically, any two modules that perform the same function should be able to replace each other (46). Metabolic modules would thus be comprised

of the respective enzymes, cofactor biosynthesis enzymes, and regulatory proteins (18). The patchwork nature of methylotrophy modules and the scattered phylogeny of known methylotrophs suggest a more complex evolutionary history than for the methanogens, and likely involved multiple horizontal gene transfer events amongst the bacteria.

Primary oxidation of C_1 compounds by methylotrophs:

Methylotrophs contain unique enzymes dedicated to the primary oxidation of C_1 compounds to formaldehyde. The best-studied are those responsible for the oxidation of methane, methanol, and methylamine. The methylotrophs capable of utilizing methane, the methanotrophs, catalyze its oxidation to methanol through the action of one of two different types of methane monooxygenase (MMO), a membrane form (particulate MMO) or a cytoplasmic form (soluble MMO) (69). Gram-negative methane- and methanol-utilizing bacteria contain a periplasmic quinoprotein, methanol dehydrogenase (MDH), which oxidizes methanol to formaldehyde (40). These enzymes have an $\alpha_2\beta_2$ structure (encoded by *mxaF* and *mxaI*), contain Ca^{2+} and pyrroloquinoline quinone (PQQ), and utilize a specific cytochrome *c* to transfer electrons from PQQ to the electron-transfer chain. Gram-positive methylotrophs, on the other hand, contain one of two different decameric cytoplasmic methanol dehydrogenases (3, 15). Methylamine is generally oxidized to formaldehyde by another periplasmic quinoprotein, methylamine dehydrogenase (MaDH) (30). Rather than PQQ, MaDH contains tryptophan tryptophylquinone (TTQ) (82). Finally, a notable exception to primary oxidation pathways that lead to formaldehyde production from C_1 compounds is the case of chloromethane utilization by *Methylobacterium chloromethanicum* CM4. This organism metabolizes chloromethane through a methyl transferase to a corrinoid protein, and then again to tetrahydrofolate (H_4F) (121). Utilization of the resulting methyl-H_4F is described below.

C_1 assimilation by methylotrophs:

Three pathways are utilized by different methylotrophic bacteria to incorporate C_1 units into multi-carbon compounds (69) (Fig 1.2). Autotrophic methylotrophs first perform the complete oxidation of reduced C_1 compounds to CO_2 before assimilating CO_2 with the Calvin-Benson-Bassham cycle. Thus far, only some α-proteobacterial methylotrophs such as *Paracoccus denitrificans* and *Xanthobacter autotrophicus* have been found to employ this strategy (69). For the ribulose monophosphate (RuMP) cycle, formaldehyde is condensed with RuMP to generate the six-carbon compound hexulose phosphate (69). A series of carbohydrate interconversions regenerate the original five-carbon sugar-phosphate. The first two reactions of this pathway (hexulose-phosphate synthase and phosphohexulose isomerase) are uniquely methylotrophic, the remaining are part of general sugar metabolism. The RuMP pathway is found in all β- and γ-proteobacterial methylotrophs and the Gram positive methylotrophs.

The final pathway for C_1 assimilation in methylotrophs is the serine cycle (69)(Fig 1.3), which is found in non-autotrophic α-proteobacterial methylotrophs. The initial substrate of this pathway is not formaldehyde itself but methylene-H_4F that forms from the spontaneous reaction of H_4F and formaldehyde. Serine hydroxymethyltransferase (encoded by *glyA*) places the C_1 unit from methylene-H_4F onto glycine to generate serine and free H_4F. Through a series of steps serine is converted into phosphoenolpyruvate, which can be carboxylated to form malate, which can then be cleaved into a pair of two-carbon compounds that are converted to glyoxylate (62) in order to regenerate the starting substrate, glycine. For every two methylene-H_4F molecules incorporated, one three-carbon compound is generated and one CO_2 is incorporated to regenerate the two glycines. As with the RuMP cycle, only a few of the enzymes are specifically methylotrophic whereas most also have a role in central metabolism during growth on multi-carbon compounds.

Formaldehyde oxidation by methylotrophs:

A key issue for methylotrophs is the production of the toxic metabolite formaldehyde as a central intermediate. Formaldehyde is highly reactive electrophile that can nonspecifically react with proteins, nucleic acids, and other cell components (35, 41). A key challenge for these organisms is to maximize formaldehyde flux while preventing the intracellular pool of free formaldehyde from accumulating to toxic levels. It has been suggested for a typical methylotroph that, "If the consumption of formaldehyde were inhibited, the cytoplasmic formaldehyde concentration would increase to about 100 mM in less than 1 min" (5, 127). Although the problem is greatest during methylotrophic growth, facultative methylotrophs growing on multi-carbon compounds also need to maintain the ability to detoxify formaldehyde that may be produced from co-metabolism of C_1 substrates encountered in the environment.

Several formaldehyde oxidation modules are employed by various methylotrophs to meet this challenge (Fig 1.2), and often more than one pathway is present in a single organism. Dye- and NAD-dependent formaldehyde dehydrogenases that could potentially oxidize formaldehyde directly to formate have been found in methylotrophs, but these are generally active with a variety of aldehydes and are present at low, constitutive levels that cannot account for the rate of methylotrophic growth (77, 129). Due to these considerations, they are generally not assumed to participate in methylotrophic metabolism in a significant manner. Many RuMP methylotrophs, such as *Methylobacillus flagellatum* KT, have the capacity to generate energy through the RuMP cycle by funneling a portion of RuMP cycle intermediates through 6-phosphogluconate dehydrogenase (19).

The primary means of formaldehyde oxidation in methylotrophs is through cofactor-dependent pathways that oxidize C_1 units attached to the respective C_1 carrier molecules (Fig. 1.2) (124). These include thiol-dependent pathways that utilize glutathione (GSH) or mycothiol (MySH, (110)), and pterin-dependent pathways that utilize H_4F or H_4MPT (Figs. 1.4 and 1.5).

GSH-dependent formaldehyde oxidation has been found in the autotrophic methylotrophs and is present in many non-methylotrophs such as *Escherichia coli*, where it plays a role in formaldehyde detoxification (81). This pathway (Fig. 1.5) begins with the condensation of formaldehyde with the thiol group of glutathione to form *S*-hydroxymethyl-GSH. Although this reaction can occur spontaneously, and likely does so in non-methylotrophs, an enzyme (glutathione-dependent formaldehyde activating enzyme, Gfa) was recently discovered that catalyzes this reaction in *P. denitrificans* (39). *S*-hydroxymethyl-GSH is oxidized by an NAD- and GSH-dependent formaldehyde dehydrogenase to *S*-formyl-GSH (12, 102), which is then hydrolyzed by *S*-formyl-GSH hydrolase (45) to formate and free GSH. The MySH-dependent pathway proceeds similarly, with the participation of a NAD- and MySH-dependent formaldehyde dehydrogenase (85, 120). Currently, the mechanism underlying the condensation of formaldehyde with MySH, and the complete oxidation of *S*-formyl-MySH to release either formate or CO_2 is not known.

The most widely distributed formaldehyde oxidation pathway amongst methylotrophs is the H_4MPT-linked pathway (Fig. 1.5). H_4MPT-dependent C_1 transfers were thought to be unique to methanogenic and sulfate-reducing archaea until their unexpected discovery in *Methylobacterium extorquens* AM1 (20). Subsequently, this pathway has been found in all Gram-negative methylotrophs examined other than the autotrophic methylotrophs, but not in any Gram-positive methylotrophs (126). This pathway operates in the opposite direction in methylotrophs than described above for methanogenesis. In methylotrophic bacteria, the initial step in this pathway is the reaction of formaldehyde with H_4MPT to form methylene-H_4MPT. While this reaction can occur spontaneously, a specific formaldehyde-activating enzyme (Fae) has been shown to catalyze this condensation (127). This enzyme is present at a specific activity of 1.4 μmol min^{-1} mg^{-1} and has a K_m for formaldehyde of 0.2 mM (127). Homologs are found in methanogen genomes (127), but their role in these archaea is unknown. My contribution to this

discovery, demonstrating that this enzyme is required for methylotrophic growth, is described in Chapter 3. The resulting methylene-H_4MPT can be oxidized to methenyl-H_4MPT through the action of one of two methylene-H_4MPT dehydrogenases. MtdA, which also catalyzes the oxidation of methylene-H_4F, is strictly NADP-dependent (125), whereas MtdB can use either NAD^+ or $NADP^+$ but is specific for methylene-H_4MPT (43). These enzymes are not phylogenetically related to the F_{420}- or H_2-dependent enzymes that catalyze this reaction in methanogens (117). The final H_4MPT-dependent reactions are performed by enzymes homologous to those that participate in methanogenesis. Methenyl-H_4MPT is converted to formyl-H_4MPT by methenyl-H_4MPT cyclohydrolase (Mch, (97)). Finally, formate is released from formyl-H_4MPT through the action of the formyltransferase/hydrolase complex (Fhc, (95, 96)). Fhc contains formyltransferase activity that is active with methanofuran (MFR) purified from the methanogen *Methanothermobacter marburgensis* (96). The identity and function of the MFR-analogue present in methylotrophs has not yet been identified, however, so here this pathway will simply be referred to as the H_4MPT pathway.

Serine cycle methylotrophs also utilize H_4F-dependent pathways for methylotrophy (Fig 1.5). As stated above, *M. chloromethanicum* CM4 transfers the methyl group from chloromethane to H_4F to form methyl-H_4F, which is converted to methylene-H_4F by methylene-H_4F reductase (encoded by *metF*, (115, 121)). Methylene-H_4F is then oxidized to methenyl- and then formyl-H_4F through the action of the bifunctional NADP-dependent methylene-H_4F dehydrogenase/methenyl-H_4F cyclohydrolase (encoded by *folD*, (121)). Formyl-H_4F is then irreversibly hydrolyzed to free formate and H_4F by formyl-H_4F hydrolase (encoded by *purU*, (121)). All three of these genes, *metF*, *folD*, and *purU*, are found in many non-methylotrophs where they function to provide C_1 units at various oxidation states for biosynthetic reactions such as the synthesis of methionine, thymidylate, and purines (74). In *M. chloromethanicum* CM4, however, this pathway is specifically required to catabolize chloromethane (115, 121). Another

set of H_4F-dependent enzymes also exist in methylotrophs. These include two separate enzymes that perform NADP-dependent methylene-H_4F dehydrogenase (MtdA, (22, 125)) and methenyl-H_4F cyclohydrolase (Fch, (97)) activities, as well as a reversible, ATP-dependent formate-H_4F ligase (FtfL, alternatively called formyl-H_4F synthetase (99)). FtfL had been detected in methylotrophs (77), but until this work (Chapters 2 and 4), the gene encoding this activity had not been identified.

Finally, the formate produced from the cofactor-dependent formaldehyde oxidation pathways is oxidized to CO_2 through the action of formate dehydrogenase (FDH). A novel, tungsten-containing FDH has been found in *M. extorquens* AM1 (68), as well as two other FDHs (L. Chistoserdova, unpublished).

The model methylotroph *Methylobacterium extorquens* AM1:

The work described here was performed with the model methylotroph, *M. extorquens* AM1 (94). This α-proteobacterium is a pink-pigmented facultative methylotroph closely related to rhizobia that is capable of growth on the C_1 substrates methanol, methylamine, and formate, as well as a limited number of multi-carbon compounds such as succinate, pyruvate, ethylamine, and oxalate. *Methylobacterium* strains are found on plant leaf surfaces as epiphytes, but this interaction may be more than simple commensalism, for they have been shown to produce plant growth hormones such as *trans*-zeatin (61). It is assumed that the plant is the direct source of methanol in this environment. Plant cell wall growth leads to release of methanol from stomata (38). *M. extorquens* AM1 has emerged as a model organism for the genetic dissection of methylotrophic metabolism due to the ability to culture it on non-methylotrophic compounds and the ease of DNA transfer by either conjugation (132) or electroporation (118). Until this work, however, only a limited number of genetic tools were available for use with *M. extorquens* AM1 (71). Chapter 2 describes my work to develop a number of improved genetic tools for use with

this organism. Additionally, the *M. extorquens* AM1 genome sequencing project is currently in its final stages (www.integratedgenomics.com/genomereleases.html#list6), but already the data available have provided a rich database of information on metabolic potential (18).

Role of the H_4MPT- and H_4F-linked C_1 transfer pathways in *M. extorquens* AM1:

The methylotrophic metabolism of *M. extorquens* AM1 is depicted in Figure 1.6. This organism contains both the H_4MPT and H_4F pathways for potential formaldehyde oxidation and assimilates carbon through the serine cycle. Prior mutant analysis indicated that both pathways played a role in methylotrophy, but the role of each pathway had not been defined.

A number of mutants defective for known or suspected H_4MPT pathway functions had been generated. Null mutants lacking *mtdB* (one of the two methylene-H_4MPT dehydrogenases) were reported to be incapable of growth on methanol and inhibited by either methanol or formaldehyde during growth on succinate (43). This methanol-sensitive mutant phenotype was unique to this H_4MPT pathway mutant, and it was suggested that methanol-sensitivity may be a proxy for formaldehyde detoxification deficiency. Double mutants lacking both MDH activity and MtdB activity were no longer sensitive to methanol (43), consistent with this hypothesis. Interestingly, this mutant phenotype was observed despite the presence of the alternative methylene-H_4MPT dehydrogenase, *mtdA*. Null mutants lacking *orf4*, which encodes a homolog of the first enzyme in the H_4MPT biosynthesis pathway, β-ribofuranosylaminobenzene 5'-phosphate (β-RFAP) synthase (106), or *orf7* (unknown function) were also generated and were incapable of growth on methanol (20). Pathways leading to the biosynthesis of H_4MPT and MFR in methanogens have been proposed, and a few enzymes catalyzing activities proposed to be involved have been purified (42), however, no genetic analysis has been performed to validate these efforts.

Efforts to obtain null mutants lacking other H_4MPT pathway enzymes have been unsuccessful. Null mutants could not be isolated for *mtdA* (22), *mch*, and *fhcBADC*, as well as two orfs of unknown function proximal to the H_4MPT gene cluster, *orf9* and *orf17* (20). Mutants resulting from an incomplete allelic exchange event that maintained a wild-type copy of the gene separated from its native promoter by the integrated vector were obtained for these "essential" genes and, where examined, this led to reduced enzymatic activity (20, 22). In all cases, this class of mutants exhibited defective growth on methanol, indicating a specific role in methylotrophy in addition to apparent essentiality. MtdA activity is likely required to produce formyl-H_4F for biosynthetic needs (22). It has not been clear, however, why the other genes for which null mutants could not be obtained and have a known or predicted role in the H_4MPT pathway for formaldehyde oxidation would be required for growth on a multi-carbon substrate such as succinate.

Previous genetic analyses of H_4F-dependent C_1-transfer in *M. extorquens* AM1 have also been inconclusive. The high, C_1-inducible H_4F-dependent activities in serine cycle methylotrophs led to the hypothesis that this pathway functions as the primary formaldehyde oxidation pathway in these organisms (77), as well as the suggestion that this pathway channels carbon into the serine cycle during growth on formate (66). Null mutants of *mtdA* or *fch* could not be obtained, however, even during growth on multi-carbon substrates such as succinate (22, 23, 125). This was thought to be due to their critical role during growth on multi-carbon compounds, likely in producing formyl-H_4F for the biosynthesis of purines and other compounds (74), but other possibilities also existed. Mutants with a reduced activity of MtdA or Fch were obtained, however, and these strains were found to be defective for growth on C_1 compounds. These data confirmed a role for these enzymes in methylotrophy, but did not clarify whether that role was in formaldehyde oxidation or another function. Formate-H_4F ligase activity had been detected in serine cycle methylotrophs (77) but a candidate gene responsible for encoding its activity had not

been identified, nor had a mutant defective for FtfL been generated. Based on these reports, it was unclear whether the H_4F-pathway is a second formaldehyde oxidation pathway or whether it plays another role.

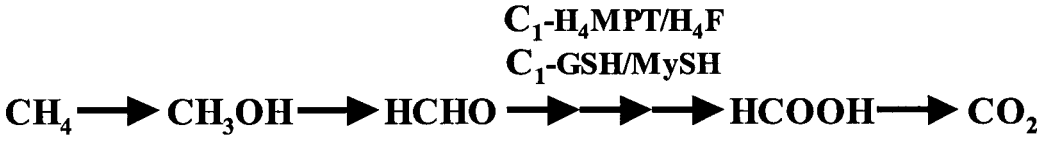
Significance of this work:

The large body of knowledge available regarding methylotrophic metabolism in *M. extorquens* AM1 makes it an attractive organism for studying the physiology of specialized metabolism. The majority of the components involved in the central methylotrophic pathways were known (Fig. 1.6), but their individual roles in formaldehyde metabolism were not well understood. In order to use *M. extorquens* AM1 as model for how bacteria switch between fundamentally different modes of metabolism, the roles of the H_4MPT and H_4F pathways must be understood. Another broad impact of studying methylotrophs is that they play a critical role in the biogeochemical cycling of C_1 compounds (44) and have the potential to convert alternate feedstocks such as methanol or methane into value-added chemicals and/or biopolymers. The fundamental processes that lead to the assimilation, oxidation, and detoxification of formaldehyde must be understood to enable efforts to engineer the metabolism of *M. extorquens* AM1 in a rational manner.

The fundamental question that inspired my project was to determine how *M. extorquens* AM1 efficiently deals with the toxic central metabolite formaldehyde such that it can efficiently direct flow into assimilatory and dissimilatory metabolism while preventing toxicity. My overarching hypothesis has been that specific enzymatic systems are required to direct formaldehyde into assimilatory and dissimilatory metabolism. Specifically, I hypothesized that: 1.) The H_4MPT pathway is responsible for formaldehyde oxidation and detoxification, and 2.) The H_4F pathway is required, not as a secondary formaldehyde oxidation route (77), but perhaps in an assimilatory role to convert formate into methylene-H_4F for the serine cycle. My approach

entailed an initial investment of time into the development of a series of genetic tools, which facilitated a genetic and physiological analysis of formaldehyde metabolism. These analyses led to the generation of a new model of methylotrophic metabolism. This model was verified through the use of analytical biochemical techniques, culminating in the ability to experimentally determine flux through every segment of C_1 metabolism. These efforts have solidified our understanding of the distinct roles of the H_4MPT and H_4F pathways in *M. extorquens* AM1 and provide a rich picture of C_1 fluxes that can serve as the backdrop for subsequent global approaches to understand the complex regulation of this specialized metabolism.

Methylotrophy



Methanogenesis

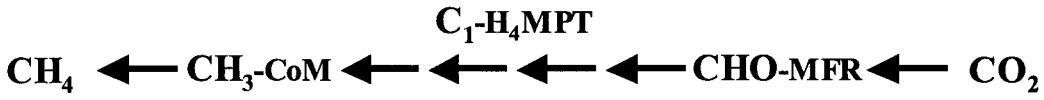


Figure 1.1: Comparison of C_1 flow in methylotrophy and methanogenesis. Methylotrophy (depicted starting with methane) involves the generation of the free C_1 intermediates methanol, formaldehyde, and formate. However, C_1 carriers are generally involved in formaldehyde oxidation (H_4MPT, tetrahydromethanopterin; H_4F, tetrahydrofolate; GSH, glutathione; MySH, mycothiol). In contrast, all C_1 intermediates in methanogenesis (depicted starting with CO_2) remain attached to C_1 carrier molecules (MFR, methanofuran; H_4MPT; CoM, coenzyme M).

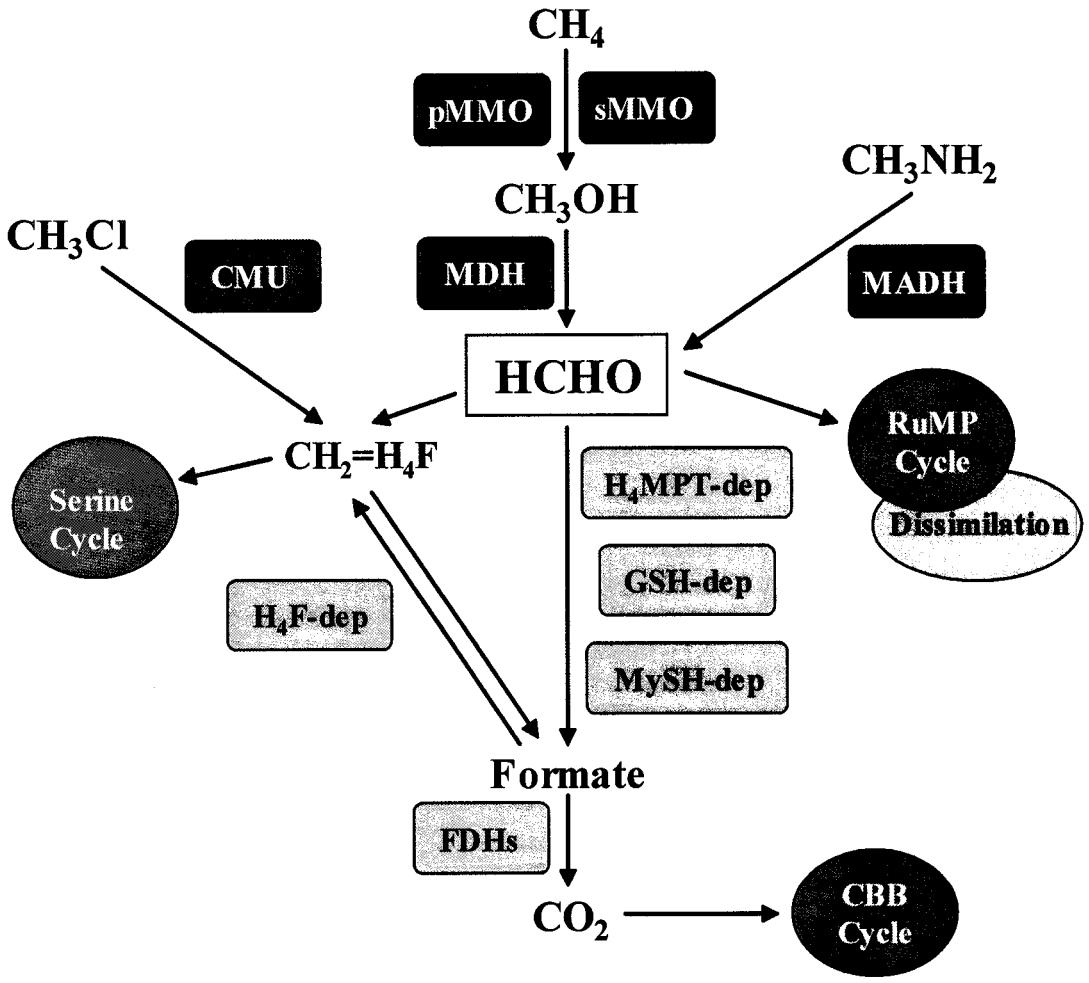


Figure 1.2: Methylotrophic modules. Primary oxidation modules are indicated in red, secondary oxidation of formaldehyde and formate in orange, and assimilatory modules in green. The modules indicated include: particulate and soluble methane monooxygenase (pMMO, sMMO), methanol dehydrogenase (MDH), methylamine dehydrogenase (MaDH), chloromethane utilization system (CMU), ribulose monophoshate cycle (RuMP, assimilatory and dissimilatory), serine cycle, Calvin-Benson-Bassham (CBB) cycle, cofactor-dependent formaldehyde oxidation pathways (tetrahydromethanopterin, H_4MPT; glutathione, GSH; mycothiol, MySH; tetrahydrofolate, H_4F), and formate dehydrogenases (FDHs).

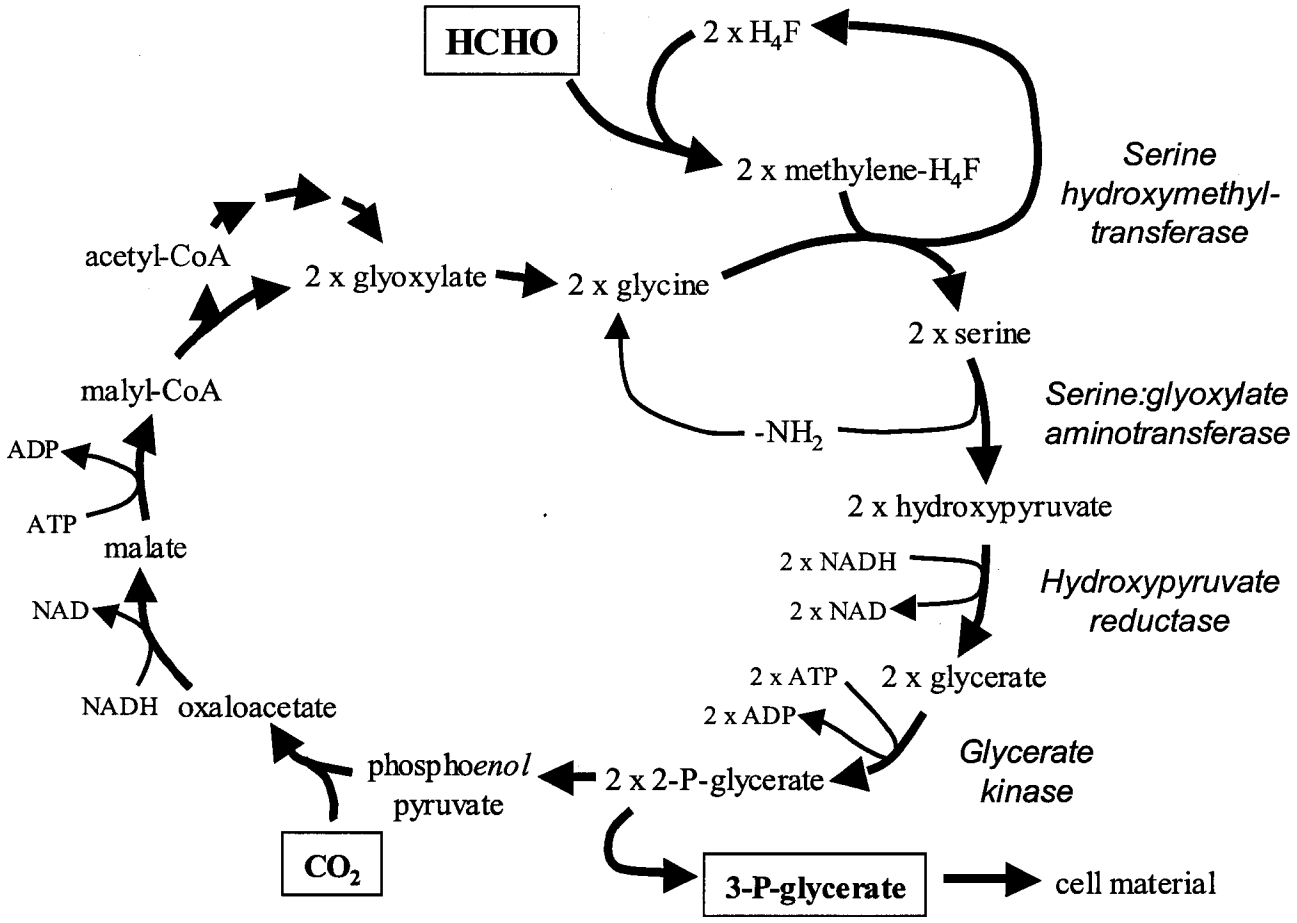


Figure. 1.3: Serine cycle for formaldehyde assimilation. Two formaldehydes molecules and one CO_2 result in the production of one C_3 compound for assimilation. The enzymes that are unique to the serine cycle (or highly divergent from non-methylotrophic homologs) are indicated in italics.

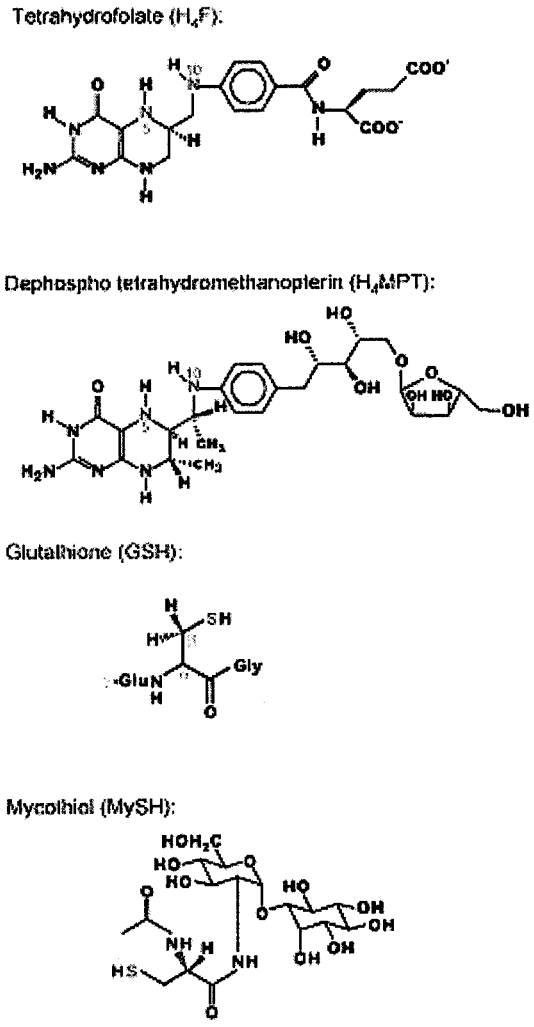


Figure. 1.4: Structures of C_1 carrier molecules. The amine and thiol atoms that covalently link to C_1 units are depicted in green. Adapted from Vorholt, 2003.

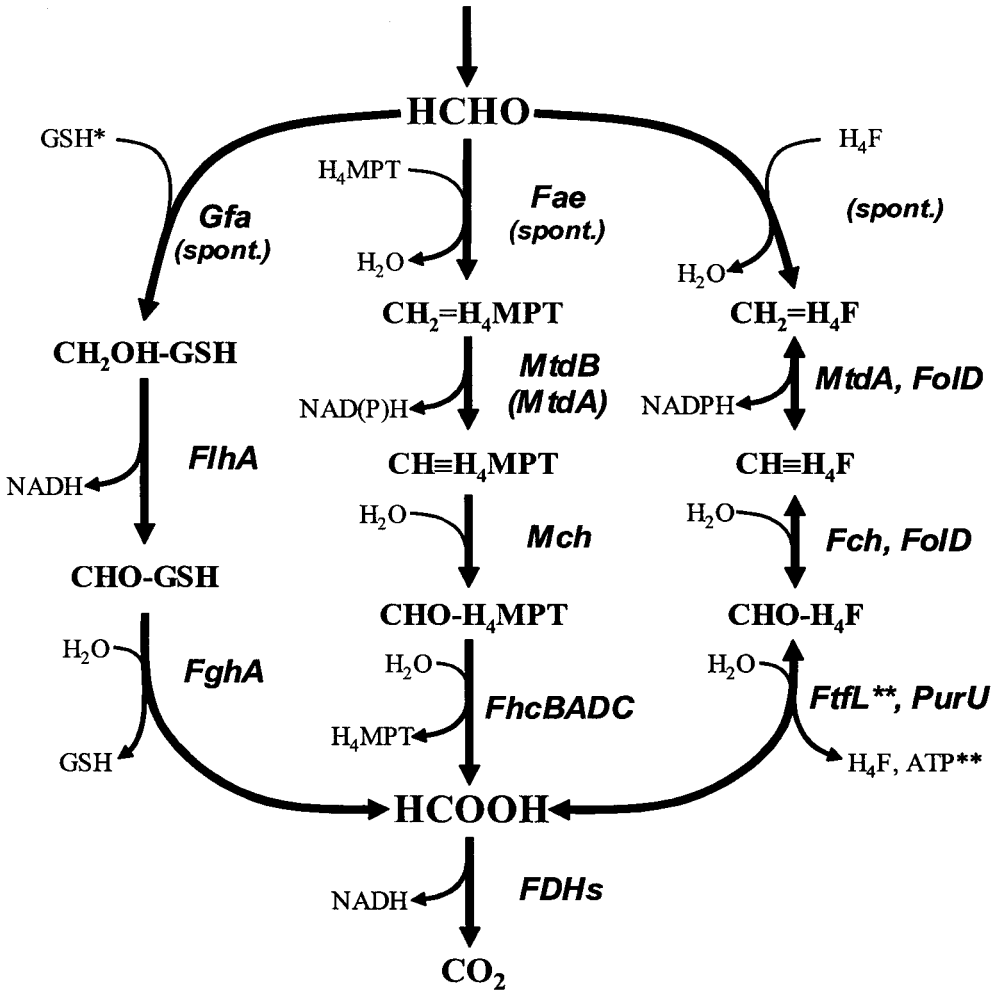


Figure. 1.5: Cofactor-dependent formaldehyde oxidation pathways. The thiol-dep. pathways are represented with the GSH-dep. enzymes Gfa (GSH-dep. formaldehyde activating enzyme), FlhA (NAD- and GSH-dep. formaldehyde dehydrogenase), and FghA (*S*-formyl-GSH hydrolase). *(An NAD- and mycothiol-dep. formaldehyde dehydrogenase is known, but the remainder of the mycothiol pathway is unknown.) The H_4MPT-dep. pathway involves the action of Fae (formaldehyde-activating enzyme), MtdB or MtdA (NAD(P)-dep. methylene-H_4MPT dehydrogenases), Mch (methenyl-H_4MPT cyclohydrolase), and FhcBADC (Formyltransferase/hydrolase complex). Two distinct H_4F pathways are known, one that involves MtdA (NADP-dependent methylene-H_4F dehydrogenase), Fch (methenyl-H_4F cyclohydrolase), and FtfL (formate-H_4F ligase, **reversible and ATP-dep.), or one that involves FolD (bifunctional NADP-dep. methylene-H_4F dehydrogenase/methenyl-H_4F cyclohydrolase) and PurU (formyl-H_4F hydrolase, irreversible). All condensations with formaldehyde can occur spontaneously.

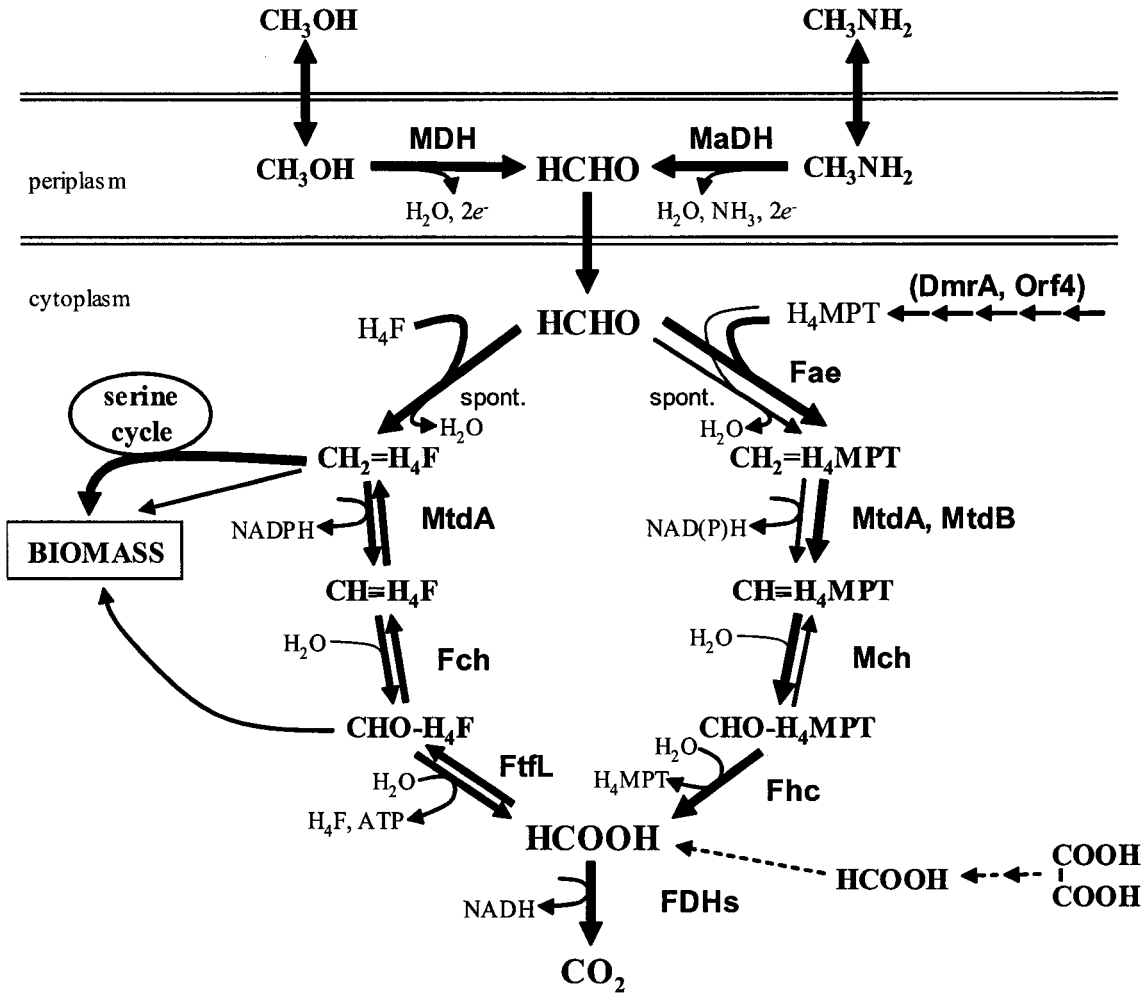


Figure 1.6: Methylotrophic metabolism in *M. extorquens* AM1. Those reactions that can occur spontaneously or catalyzed by two enzymes are indicated. The thin arrows leading from methylene-H_4F and formyl-H_4F to biomass represent biosynthetic reactions directly involving these two C_1-H_4F derivatives. Arrows to the right of H_4MPT indicate biosynthesis reactions for this cofactor; only the two indicated gene products (DmrA, putative H2MPT reductase; Orf4, putative βRFAP synthase) have been implicated to this process. The entry point of formate and oxalate are also shown in dashed lines. MDH (methanol dehydrogenase), MaDH (methylamine dehydrogenase), MtdA (NADP-dependent methylene-H_4F /methylene-H_4MPT dehydrogenase), Fch (methenyl-H_4F cyclohydrolase), FtfL (formate-H_4F ligase), FDHs (formate dehydrogenases), DmrA (putative dihydromethanopterin reductase), Orf4 (putative β-RFAP synthase), Fae (formaldehyde-activating enzyme), MtdB (NAD(P)-dependent methylene-H_4MPT dehydrogenase), Mch (methenyl-H_4MPT cyclohydrolase), Fhc (formyltransferase/hydrolase complex).

CHAPTER 2

Development of Genetic Tools for use in *Methylobacterium extorquens* AM1

Abstract:

In order to perform a thorough genetic and physiological analysis of the role of the H_4MPT and H_4F pathways in the methylotrophic metabolism of *M. extorquens* AM1, a number of genetic tools were developed. First, a suite of small broad-host-range IncP vectors were created for cloning, expression, and promoter-probing. Second, a broad-host-range *cre-lox* based antibiotic marker exchange system was developed in order to generate unmarked mutant strains. Third, an efficient transposon mutagenesis system was applied to *M. extorquens* AM1 for the first time, allowing the identification of ten novel methylotrophy genes. Finally, an insertional expression system was developed to allow expression from a stable, unmarked chromosomal locus. The utility of these genetic tools were demonstrated and are described here.

Introduction:

The ever increasing availability of genome sequence data for microbial species has revolutionized and invigorated the field of microbiology. Bioinformatics approaches have allowed prediction of gene function and phylogenetic classification of up to 80% of the putative ORFs in these organisms (116). Simultaneously, functional genomics techniques yield global data regarding information such as the change in the transcript level of all of the genes in a given organism (27). Myriad hypotheses can be generated as to the functional role of various cell components based on these genome-enabled approaches; however, genetic analysis is required for testing these models/hypotheses. A major hurdle for studying the majority of these organisms is the paucity of sophisticated genetic tools. The capabilities and user-friendliness of broad-host-range genetic tools continue to lag far behind that available for use in model organisms such as *Escherichia coli* or *Bacillus subtilis*. In order to take full advantage of the wealth of information available from bacterial genome sequencing, significant advances in the quality of broad-host-range genetic tools are required.

M. extorquens AM1 is a prime example of an organism for which improved genetic tools have been needed. The ease of DNA transfer by either conjugation (132) or electroporation (118) and the ability to culture *M. extorquens* AM1 on both C_1 substrates and a limited number of multi-carbon compounds has made *M. extorquens* AM1 a model organism for the genetic dissection of methylotrophic metabolism. Although the genome sequence of this organism is still being completed at this time (www.integratedgenomics.com/genomereleases.html#list6), the sequence data available thus far have greatly expanded our knowledge about the central metabolism of this organism (18). The genetic tools available previous to this project included a few large (19-23 kb) IncP vectors with limited cloning sites (33, 60, 133) and a simple allelic exchange system (17), however, these were difficult to utilize on a routine basis.

The majority of broad-host-range vectors available are based upon either the IncP or IncQ replicons. Most IncP plasmids are derived from RK2, a 60.1 kb self-transmissible plasmid (93). RK2 has an estimated copy number in *E. coli* of 5 to 7 per chromosome (37), and plasmids containing the IncP origin of transfer, *oriT*, can be mobilized by IncP transfer proteins provided *in trans* (36). Currently available vector tools based on IncP replicons are low- to medium-copy in *E. coli*, and few are fully sequenced or have a large number of available restriction sites. Previously, only a few large IncP plasmids have been used with success in *M. extorquens* AM1, including the cloning vectors pRK310 (33) and pVK100 (60) and the promoter-probe vectors pHX200 (133), and pGD500 (33). No IncQ vectors or smaller IncP vectors have been found to be maintained in this strain (69). Therefore, the first set of genetic tools developed for use in *M. extorquens* AM1 were a suite of small broad-host-range vectors for cloning, overexpression, and promoter-probing.

A second key genetic tool previously unavailable for use in *M. extorquens* AM1 was the means to generate unmarked mutant strains. An antibiotic marker recycling system allows multiple genetic manipulations of organisms for which few antibiotic markers exist. In recent

years, a growing number of antibiotic marker recycling systems have been reported that utilize a variety of site-specific recombination systems and antibiotic markers (31). These include the utilization of the RP4 multimer resolution system (65) and both the yeast *Flp*/FRT (48) and P1 phage *cre/lox* (7) site-specific recombination systems. The *cre-lox* methodology relies upon the expression of a site-specific recombinase, Cre, which catalyzes *in vivo* excision of DNA regions flanked by co-directional *loxP* recognition sites (92). In order to take advantage of these capabilities for the study of *M. extorquens* AM1, a broad-host-range mobilizable allelic exchange vector with a *loxP*-flanked antibiotic resistance cassette and an IncP plasmid for expression of the Cre recombinase were developed.

Another basic genetic tool lacking for *M. extorquens* AM1 was a transposon mutagenesis system. Past attempts to identify methylotrophy genes using transposon mutagenesis have been problematic and found only previously known genes (131). During the course of this project, however, mutagenesis using miniTn*5* identified novel genes involved in chloromethane utilization in the closely related strain *Methylobacterium chloromethanicum* CM4 was reported (122). Therefore, a variety of Tn*5* derivatives were tested, and it was discovered that another miniTn*5* derivative, IS*phoA*/hah-Tc (10, 29), could be used efficiently. A transposon mutagenesis screen in *M. extorquens* AM1 to search for previously unknown methylotrophy genes was performed, leading to the discovery of ten novel methylotrophy genes.

Finally, no insertional vector systems were available for the overexpression of genes from a chromosomal locus. The chromosomal locus chosen for insertions was the *katA* (encodes a catalase) site because *M. extorquens* AM1 contains multiple active catalases and *katA* mutants were reported to have wild-type growth characteristics under the conditions tested (24). The insertional expression system developed here thereby allows genes to be expressed from a stable, single-copy, unmarked locus on the chromosome. This obviates antibiotic selection and enables further genetic manipulations in the background of the expression strain.

Materials and Methods:

Bacterial strains and growth conditions. Wild-type *M. extorquens* AM1 (94) was cultured at 30°C in a minimal salts medium (4) containing carbon sources at the following concentrations: 30 mM ethylamine, 35 mM formate, 125 mM methanol, 35 mM methylamine, 15 mM oxalate or 15 mM succinate. *Escherichia coli* strains were grown on LB medium (103). *M. extorquens* AM1 strains and plasmids used are described in Tables 2.1 and 2.2. Antibiotics were added to the following final concentrations: 50 μg/ml ampicillin, 50 μg/ml kanamycin, 50 μg/ml rifamycin, 35 μg/ml streptomycin, and 10 μg/ml tetracycline. Chemicals were obtained from Sigma. Nutrient agar and Bacto-agar were obtained from Difco.

Genetic procedures and recombinant DNA techniques. All DNA manipulations were performed according to standard techniques (103). Plasmids were transferred between *E. coli* and *M. extorquens* AM1 by conjugative transfer using an *E. coli* helper strain bearing pRK2013 or pRK2073 (36) or the helper strain S17-1 (C600::RP-4 2-(Tc::Mu) (Kn::Tn*7*) *thi pro hsdR hsdM*$^+$ *recA*) (107) overnight at 30°C on Nutrient agar platess. Primers were ordered from Invitrogen and are listed in Table 2.3. DNA sequencing was performed at the University of Washington Biochemistry Department DNA Sequencing Facility.

Construction of improved broad-host-range cloning vectors. The 2.0 kb *Nde*I-*Not*I region of pDN19X containing the *traJ'* allele that allowed efficient replication and/or transfer in *M. extorquens* AM1 was transferred into the corresponding region of pTJS75 (104), a parent vector of pDN19 (88), to create pCM48. pCM48 was then cut with *Mun*I and *Bln*I, and the DNA ends were blunted and self-ligated to produce pCM50. pCM50 was then digested with *Hin*dIII and *Xmn*I, and the DNA ends were blunted and self-ligated to produce the 5.3 kb minimal

transferable and selectable replicon pCM51 (Fig. 2.1). Two additional plasmids were created to confirm that the *traJ'* allele present in pDN19X is sufficient to maintain and/or transfer the plasmid into *M. extorquens* AM1. The 2.0 kb *Nde*I-*Not*I region surrounding the substitution in pDN19X was swapped into the corresponding region of pDN19 to create the plasmid pCM46. Similarly, the 2.0 kb *Nde*I-*Not*I region from pDN19 encoding the full length TraJ was cloned into pDN19X to create the plasmid pCM47.

In order to compare the reduced gene complement of the minimal selectable and transferable replicon pCM51 to the cloning vector previously used with *M. extorquens* AM1, pRK310 (33), the complete sequence of this large IncP vector was pieced together by determining the sequence over the junctions created during the partial digestion steps in its cloning history (Fig. 2.1).

The improved broad-host-range vector pCM62 was created by combining the functions present in the minimal transferable and selectable replicon pCM51 with the polylinker and ColE1 origin of replication (*ori*) of pUC19 (Fig. 2.2). This was accomplished by ligating the blunted 5.3 kb *Hin*dIII-*Xmn*I region of pCM50 (used to make pCM51) with the blunted 1.9 kb *Aat*II-*Ava*II region of pUC19 (134). During the construction of pCM62 an additional 356 bp inward from the site of *Xmn*I cleavage of pCM50 was lost, presumably caused by either non-specific cleavage or the exonuclease activity of T4 DNA polymerase. Loss of this region leads to a small alteration of the C-terminus of the TetR protein, from GDD-COOH to GAKKPLLS-COOH, but this mutation does not adversely affect the tetracycline resistance conferred by this plasmid in either *E. coli* or *M. extorquens* AM1. A derivative of pCM62 conferring resistance to kanamycin, pCM66, was constructed by inserting the 1.3 kb *Hinc*II fragment of pUC4K (123) between the *Eco*RV and *Nru*I sites of pCM62 (Fig. 2.2).

Construction of plasmids containing the reporter genes *xylE*, and *gfp*, and the P_{mxaF} and *mxaF* gene of *M. extorquens* AM1. *xylE* was amplified by PCR from the promoter-probe vector pHX200 (133) using CM-xylEf and CM-xylEr. The resulting PCR product was cloned into pCR2.1 (Invitrogen) to create pCM20. Two other plasmids were created to allow the introduction of *xylE* into various plasmids. The 1.0 kb *Hin*dIII-*Bam*HI region of pCM20 containing *xylE* was cloned between the *Hin*dIII and *Bam*HI sites of the cloning vector pMTL23 (16) to produce pCM22. Finally, the 1.0 kb *Hin*dIII-*Sac*I fragment of pCM22 containing *xylE* was inserted into pMTL23 cut with *Hin*dIII and *Sac*I to create pCM75. *lacZ* was amplified by PCR from pMUTIN2 (119) and cloned into pCR2.1 (Invitrogen) to create pLacZ2.1+ (R. Meima & M. E. L., unpublished). *gfp* was amplified by PCR from pGFPuv (Clonetech) using CM-gfpf and CM-gfpr. The resulting PCR product was cloned into pCR2.1 (Invitrogen) to create pCM21. An additional plasmid bearing *gfp*, pCM23, was created by inserting the 0.8 kb *Hin*dIII-*Nsi*I fragment of pCM21 into pMTL23 cut with *Hin*dIII and *Nsi*I. The P_{mxaF} of *M. extorquens* AM1 was amplified by PCR from pDN411 (89) using CM-PmxaFf and CM-PmxaFr. The resulting 324 bp PCR product was cloned into pCR2.1 (Invitrogen) to create pCM27. Similarly, a 1.9 kb region bearing the *mxaF* gene without its promoter was amplified by PCR from pDN411.(89) using CM-mxaFf and CM-mxaFr. This product was introduced into pCR2.1 (Invitrogen) to create pCM74.

Construction of low-background promoter-probe vectors. Two low-background broad-host-range promoter-probe vectors with different reporters were created using the cloning vectors pCM62 and pCM66 as their vector backbone. The 1.0 kb *Hin*dIII-*Nco*I fragment from pCM75 containing *xylE* was inserted into pCM62 cut with *Afl*III and *Hin*dIII to create pCM76. Similarly, the 3.2 kb *Eco*R1 fragment of pLacZ2.1+ (R. Meima & M. E. L., unpublished) containing *lacZ* was blunted and cloned into the *Bam*HI site of pCM66 which had been similarly blunted to create pCM66LacZ. pCM76 and pCM66LacZ were each found to have a high

background reporter gene activity in *M. extorquens* AM1 (data not shown). In order to reduce this background activity, an *E. coli* terminator, t_{rrnB}, was introduced upstream of the multiple cloning sites on the two vectors. The 0.5 kb *Xba*I-*Xmn*I fragment of pMTL20T$_1$T$_2$ (28) containing t_{rrnB} was excised, blunted, and ligated into the blunted *Eco*RI site of pCM76 to create pCM130 (Fig. 2.3). The 1.0 kb *Eco*RI-*Rsr*II fragment of pCM130 containing t_{rrnB} was then transferred into pCM66LacZ, cut with *Eco*RI and *Rsr*II, to create pCM132 (Fig. 2.3). In order to test the ability of these vectors to detect promoter activity, the P_{mxaF} of *M. extorquens* AM1 was introduced into both plasmids. The 0.4 kb *Eco*RI-*Bam*HI fragment from pCM27 was cloned into pCM130 cut with *Eco*RI and *Bam*HI to create pCM131. Similarly, the 0.4 kb *Eco*RI fragment from pCM27 was inserted into the *Eco*RI site of pCM132 to produce pCM133.

Construction of expression vectors for use in *M. extorquens* AM1. The first step toward the creation of the expression vector pCM80 was the introduction of the P_{mxaF} of *M. extorquens* AM1 into pCM62. This was accomplished by ligating the 0.4 kb *Hin*dIII-*Bam*HI fragment of pCM27 into pCM62 cut with *Hin*dIII and *Bam*HI to produce pCM64. In order to keep the *Hin*dIII site in the pCM80 polylinker unique, pCM64 was cut with *Hin*dIII, blunted with T4 DNA polymerase, and self-ligated to produce pCM79. Finally, the polylinker present in pCM62 including an intact *lacZα* was restored through the introduction of a 64 bp linker (Fig. 2.4). A pair of overlapping oligonucleotides with *Bam*HI compatible 5′ overhangs was designed for this purpose, CM-L64f and CM-L64r. These oligonucleotides were mixed to a final concentration of 100 μM each in distilled water, heated to 95°C for 5 minutes, and then allowed to cool at 70°C for 5 minutes followed by 50°C for 5 minutes. The linker was then added directly to a ligation mix containing pCM79 cut with *Bam*HI. Plates containing X-gal were used to screen for clones containing a linker region inserted in the desired orientation. Restriction analysis and sequence data across this region from one such clone, pCM80, confirmed the genuine insertion of

a single linker in the desired orientation (Fig. 2.4). A kanamycin resistant derivative, pCM160, was created by ligating together the 2.4 kb *Not*I-*Eco*RI region of pCM80 containing P_{mxaF} with the 5.6 kb *Not*I-*Eco*RI region of pCM66 containing the kanamycin resistance cassette (Fig. 2.4).

A number of genes were introduced into pCM80, and pCM62, in order to demonstrate the utility of pCM80 as an expression vector. The 0.8 kb *Hin*dIII-*Bgl*II fragment of pCM23 bearing *gfp* was inserted into both pCM62 and pCM80 cut with *Hin*dIII and *Bam*HI to create pCM87 and pCM88. Similarly, the 1.0 kb *Hin*dIII-*Bam*HI fragment of pCM22 containing *xylE* was introduced into both pCM62 and pCM80 cut with *Hin*dIII and *Bam*HI to produce pCM63 and pCM81, respectively. Finally, the 2.0 kb *Hin*dIII-*Xba*I fragment of pCM74 containing *mxaF* cloned into both pCM62 and pCM80 cut with *Hin*dIII and *Xba*I, generating pCM85 and pCM86. In addition to pCM80 and pCM160, an additional expression vector, pCM110, was created that would provide minimal expression in *E. coli* to allow toxic genes to be introduced into *M. extorquens* AM1 (Fig. 2.4). pCM110 was constructed by inserting the 0.4 kb *Hin*dIII-*Nsi*I fragment of pCM27 containing P_{mxaF} into pCM60 which had been cut with *Hin*dIII and *Nsi*I. In order to compare expression from the P_{mxaF} in pCM80 to that in pCM110, the 1.0 kb *Bam*HI-*Nsi*I fragment from pCM75 containing *xylE* was inserted into pCM110 cut with *Bam*HI and *Nsi*I to create pCM111.

Construction of a broad-host-range *cre-lox* system for antibiotic marker recycling. The allelic exchange vector pCM184 (Fig. 2.5) was created by inserting a *loxP*-bounded kanamycin resistance cassettes into a variant of the mobilizable suicide plasmid, pAYC61 (17). The 1.3 kb *Hin*cII fragment bearing the kanamycin resistance cassette from pUC4K (123) was inserted into pLox1 (92) which had been cut with *Xba*I and blunted, to create pCM161. In order to introduce convenient multiple cloning sites, the *loxP*-bounded kanamycin cassette of pCM161 was amplified with CM-ufkMCS and CM-dfkMCS. The resulting 1.4 kb PCR product was

purified and cloned into pCR2.1 (Invitrogen) to create pCM183. In order to preserve useful cloning sites, pAYC61 was cut with *Eco*RI and *Sma*I, blunted using T4 DNA polymerase, and self-ligated to produce pCM182. Finally, the 1.4 kb *Aat*II-*Spe*I fragment from pCM183 containing the *loxP*-flanked kanamycin cassette was ligated between the *Aat*II and *Xba*I sites of pCM182 to create pCM184 (GenBank accession number AY093429).

Two broad-host-range *cre* expression vectors, pCM157 and pCM158 (Fig. 2.6), were created. The 1.1 kb *Xba*I-*Eco*RI fragment from pJW168 (92) was cloned between the *Xba*I and *Eco*RI sites of pCM62 to generate the tetracycline-resistance conferring *cre* expression plasmid pCM157. A kanamycin-resistant version, pCM158, was generated by cloning the same *Xba*I-*Eco*RI fragment from pJW168 between the *Xba*I and *Eco*RI sites of pCM66. These plasmids contain *cre* behind the *E. coli lac* promoter, which only provides low constitutive activity (Table 2.4). Despite this low expression, the majority of cells obtained from the first passage onto plates lacking kanamycin were already kanamycin sensitive.

A construct to generate *mxaF* mutants was generated using pCM184. The *mxaF* upstream and downstream flanks were amplified by PCR using CM-DmxaF1 and CM-DmxaF2, and CM-DmxaF3 and CM-DmxaF4. The resulting products were purified and cloned into pCR2.1 (Invitrogen) to produce in pCM191 and pCM192. The 0.5 kb *Apa*I-*Sac*I fragment from pCM192 was introduced between the corresponding sites of pCM184 to produce pCM193. Subsequently, the 0.6 kb *Eco*RV-*Asp*718I fragment from pCM191 was ligated between the *Pvu*II and *Asp*718I sites of pCM193 to produce pCM194.

Construction of a *M. extorquens* AM1 insertional expression vector. A 1.2 kb *Nru*I-*Nco*I fragment containing the erythromycin cassette from pMTL23E (98) was blunted using T4 DNA polymerase and inserted into the *katA* region present in pLC11.28 (24), which had been cut with *Nru*I and *Pst*I and also blunted to generate pCM116. It should be noted that attempts to use

erythromycin as a selective marker in *M. extorquens* AM1 were unsuccessful. In order to reserve useful cloning sites in the final insertion vectors, pCM116 was cut with *Eco*RI and *Nde*I, blunted and self-ligated to produce pCM117, which subsequently was cut with *Hin*dIII, blunted and re-ligated to produce pCM118. A construct containing *tetAR* flanked by *loxP* sites was constructed by inserting the 2.3 kb *Xmn*I/*Stu*I fragment from pCM50 into the blunted *Xba*I site of pLox1 (92) to produce pCM159. Sequencing of the pCM159 construct revealed that an intact *Xba*I site remained on the *tetR* side of the cassette. Fortunately, the resulting sequence is also recognized by Dam methylase, leading to methylation that blocks *Xba*I cleavage, allowing use of the *Xba*I site in the polylinker, provided that the plasmid DNA was prepared from a Dam+ strain. The 2.6 kb *Nco*I-*Sca*I fragment from pCM159 was ligated between the *Nco*I and *Hin*cII sites of pCM118 to produce pCM165. To remove further restriction sites from the insertional vector, pCM165 was digested with *Nco*I and *Nde*I, blunted, and self-ligated to produce pCM166, which was cut with *Nsi*I, blunted, and self-ligated to produce pCM167.

The *E. coli rrnB* terminator (t_{rrnB}) from pCM130 and the T7 terminator (t_{T7}) from pET-3a (Novagen) were amplified by PCR using the primer pairs CM-trrnBf and CM-trrnBr, and CM-tT7f and CM-tT7r. The resulting products were cloned into pCR2.1 (Invitrogen) to generate pCM119 and pCM120, respectively. The 0.5 kb *Bam*HI-*Hin*cII fragment from pCM119 containing t_{rrnB} was ligated into the same sites of pMTL23 to generate pCM123. The 0.4 kb *Nru*I-*Xho*I fragment from pCM120 was inserted into the same sites of pCM123 to generate pCM124. A terminator-flanked cassette bearing P_{mxaF} was generated by inserting the 0.3 kb *Nru*I-*Hin*dIII fragment from pCM80 between the *Hin*cII and *Hin*dIII sites of pCM124 to generate pCM126.

The insertional vector backbone pCM167 contains unique *Bgl*II and *Stu*I sites between which terminator-flanked cassettes were inserted as *Bam*HI-*Nru*I fragments. The insertional cloning vector pCM168 contains the 0.9 kb fragment from pCM124, whereas the insertional expression vector pCM172 contains the 1.2 kb pCM126 fragment (Fig. 2.7). Additionally, a

construct was made to generate a *katA::kan* strain that allows the identification of those recombinants with a complete allelic exchange at the *katA* locus. The 3.4 kb *Eco*RI-*Sph*I fragment from pLC1128.Km containing the Km^r marker (24) was blunted and cloned into the *Sma*I site of pAYC61 (17) to generate pCM82.

In order to test the efficiency of transcription termination afforded by t_{rrnB} or t_{T7} in *M. extorquens* AM1, P_{mxaF} present in the 0.4 kb *Bam*HI-*Eco*RI fragment from pCM27 was introduced between the same sites upstream of the reporter gene *xylE* in pCM76 to generate pCM77. The 0.6 kb *Bam*HI-*Sph*I fragment from pCM119 and the 0.4 kb *Bam*HI-*Sph*I fragment from pCM120 were then ligated into the same sites of pCM77 between P_{mxaF} and *xylE* to generate pCM121 and pCM122, respectively. The expression level afforded by pCM172 was examined by inserting the 0.8 kb *Hin*dIII-*Nsi*I fragment from pCM21 with *gfp* into the same sites of pCM172 to generate pCM174.

Transposon mutagenesis and mutant screen. Bi-parental matings were performed to introduce pCM639 (29) from *E. coli* SM10 λ_{pir} (84) into wild-type *M. extorquens* AM1 (90) on nutrient agar (Difco) overnight at 30°C. Dilutions of the bi-parental mating mixtures were then plated onto minimal salts medium containing succinate for growth with rifamycin and tetracycline for selection. Analogous experiments using the related transposons Tn*phoA* and mini-*Tn*5 Tc resulted in a significantly lower number of Tc-resistant colonies. Fresh cultures of IS*phoA*/hah-Tc containing mutants were then screened on plates containing methanol for growth defects. Growth phenotypes on various substrates were assessed after 3 to 5 days of growth based on colony size relative to wild-type *M. extorquens* AM1 and placed into three categories: wild-type = "++", intermediate = "+" and trace to none = "-". Strains were also tested on medium containing both succinate and methanol to identify methanol-sensitive mutants.

A semi-random two-step PCR procedure (ST-PCR) to amplify the junction at the site of transposon insertion into the chromosome was performed essentially as described (26, 75). Modifications included using boiled colony preparations or chromosomal DNA preparations as template and supplementation of the reaction mixtures with either 0.1 mg/ml BSA or 5% DMSO. The primers utilized were CM-TnPCR1 and CM-TnPCR2b for the first round and CM-TnPCR3 and CM-TnPCR4 for the second round of amplification. PCR products were purified using QIAquick PCR purification columns (Qiagen, Germany) and sequenced using CM-TnSEQ. Sequence analyses were performed using the ERGO website (ergo.integratedgenomics.com/ERGO/CGI, Integrated Genomics, Chicago) and the NCBI website (www.ncbi.nlm.nih.gov).

Reporter gene assays and SDS-PAGE analysis. Cell extracts for enzyme assays and SDS-PAGE gels were prepared from mid-exponential cultures of *M. extorquens* AM1, as determined using 1.0 cm cuvettes in a Beckmann DU 640B spectrophotometer. XylE (catechol dioxygenase) and LacZ (β-galactosidase) assays are reported as the average and standard deviation of three replicates and were performed as described previously (57, 83, 136). XylE activities in *E. coli* JM109 were assayed in extracts prepared from cultures grown to an OD_{600} of 0.6-1.0 in LB. GFP (green fluorescent protein) production was analyzed *in vivo* by determining the relative fluorescence OD_{600}^{-1} using a Shimadzu RF-5301 PC spectrofluorophotometer with an excitation wavelength of 410 nm (with slit 5) and an emission wavelength of 509 nm (with slit 5). The total protein concentration of the extracts was determined either by direct spectrophotometric methods (54, 130) using a Beckmann DU 640B spectrophotometer for enzyme assays, or by the Bradford method using the Protein Assay Kit (Bio-Rad), using BSA as a standard, for SDS-PAGE analysis on 15% polyacrylamide gels. The relative intensities of selected bands on

Coomassie blue-stained SDS-PAGE gels were determined using Kodak 1D Image Analysis Software (Eastman Kodak).

Results:

Development of a suite of small, broad-host-range vectors for cloning, expression, and promoter-probing in *M. extorquens* AM1.

***The work described in this section has been published (79).**

Isolation of a spontaneous mutant of the small IncP plasmid pDN19 that could be maintained efficiently in M. extorquens *AM1.* The initial step toward the construction of improved genetic tools for use in *M. extorquens* AM1 was to examine a few small broad-host-range vectors for their ability to be transferred by conjugation and maintained efficiently relative to the large IncP cloning vector pRK310 (33). The plasmids tested included the small IncP vectors pDN19 (88) and pJB656 (14), the IncQ vector pMMB67HE.tet (52), and pBBR1MCS-2 from *Bordatella bronchiseptica* (64). Tetracycline resistant colonies were not obtained using pMMB67HE.tet, consistent with the previous observation that IncQ plasmids serve as suicide plasmids in *Methylobacterium* strains (13). The vectors pJB656 and pBBR1MCS-2 could be established in *M. extorquens* AM1, however, both plasmids were transferred at a low efficiency relative to pRK310 and caused *M. extorquens* AM1 to grow at a significantly reduced growth rate. Only the small IncP vector pDN19 was maintained at a growth rate comparable to strains carrying pRK310, and was thus chosen as the basis for further vector development.

Initially, it appeared that pDN19 was transferred from *Escherichia coli* to *M. extorquens* AM1 with an efficiency 1000 times less than that observed for pRK310. Following transfer of pDN19 from *M. extorquens* AM1 back into *E. coli*, however, the plasmid could be reintroduced into *M. extorquens* AM1 at the same high transfer efficiency as pRK310. This result suggested that the original pDN19 plasmid may have acquired a mutation that increased its transfer

efficiency or the ability of *M. extorquens* AM1 to maintain this plasmid. No difference was observed in the maintenance of this plasmid derivative in *E. coli*. This pDN19 derivative was designated pDN19X.

The exact sequence of the RK2-derived vector pDN19 (88) was not known. Therefore, single-strand sequence was obtained for pDN19. A map of pDN19 and the features present on this plasmid are presented (Fig. 2.1). Seven single-nucleotide differences were observed relative to the reported sequence of RK2. These included single-nucleotide deletions and insertions, and nucleotide replacements. The only changes present in coding regions disrupt *upf-16.5*, a putative ORF of unknown function (93). Single-strand sequence of pDN19X revealed a single nucleotide difference relative to the parent plasmid, pDN19, which was located within *traJ*. This C to A transversion results in an early stop codon in *traJ*, whose gene product recognizes the origin of transfer, *oriT*, initiating the DNA processing events required for conjugal transfer (135). This early stop codon creates a TraJ polypeptide that is missing the final 85 of 123 aa.

In order to determine whether the *traJ'* allele of pDN19X is sufficient to increase the maintenance and/or transfer of this plasmid in *M. extorquens* AM1, a 2.0 kb region encompassing *traJ* was swapped between pDN19X and the corresponding region of pDN19 to create pCM46. Similarly, the region from pDN19 encoding the full length TraJ was cloned into pDN19X to create pCM47. The plasmid pCM46 bearing the mutation was maintained and/or transferred as had been observed for pDN19X, whereas pCM47 lacking the mutation lost this capacity. This confirmed that the ability of pDN19X to be efficiently maintained and/or transferred in *M. extorquens* AM1 was due to the region containing *traJ'*, and not due to a distal region of the plasmid. The plasmid pRK310 is transferred and maintained in *M. extorquens* efficiently, and it has the same *traJ* allele as pDN19. In addition, all necessary *tra* functions including *traJ* are provided by the helper plasmid (pRK2013 or pRK2073) during tri-parental matings. Therefore, it is not clear why an alteration in TraJ would be required for efficient maintenance and/or transfer

of pDN19 into *M. extorquens* AM1. It is possible that the effect may be due to altered expression of downstream genes, such as the essential initiator gene *trfA*. TrfA has been shown to affect copy number and host range (47). In addition, our sequencing has shown that pDN19 lacks the native *trfA* promoter and the only known upstream promoter is that for *traJ*. It was not possible to compare expression of *trfA* or copy number between pDN19 and pDN19X, because pDN19 is apparently not maintained in *M. extorquens* AM1.

Facile broad-host-range cloning vectors. The nucleotide sequence of pDN19X was utilized to guide the removal of potentially nonessential plasmid regions in order to create a minimal transferable and selectable broad-host-range replicon that could be utilized as the backbone for further vector development. Three regions of the 7.8 kb pDN19X plasmid were sequentially excised to create the 5.3 kb minimal transferable and selectable replicon pCM51 (Fig. 2.1). The plasmid pCM51 consists solely of the IncP *oriV* and *oriT*, *traJ'*, *trfA*, and *tetA* and *tetR*. It behaved identically to pDN19X with regard to transfer and maintenance, indicating that the regions of the parent plasmid that had been removed did not affect plasmid transfer or maintenance in *M. extorquens* AM1. Furthermore, pCM51 possesses a considerably reduced gene complement relative to the 19.1 kb IncP cloning vector previously used with *M. extorquens* AM1, pRK310 (33) (Fig. 2.1). Achieving a small, functional replicon was critical in the development of improved broad-host-range cloning vectors, and the subsequent elaboration of these vectors into more sophisticated genetic tools.

The first step toward generating an improved set of genetic tools for use in *M. extorquens* AM1 was to develop a broad-host-range cloning vector for use in this organism based on the minimal transferable and selectable replicon pCM51. This was accomplished by fusing a portion of pUC19 containing its multiple-cloning site (MCS) and the ColE1 *ori* to the region of pCM50 present in the small IncP replicon pCM51, creating pCM62 (Fig. 2.2). A kanamycin resistant

derivative, pCM66, was constructed by inserting the kanamycin resistance cassette from pUC4K for use in bacteria such as methanotrophs in which tetracycline is a poor marker (Fig. 2.2). Both of these plasmids were maintained in and transferred into *M. extorquens* AM1 with efficiencies equal to pCM51. Routine alkaline lysis plasmid minipreps of pCM62 and pCM66 from *E. coli* cloning strains, however, indicated that the presence of the ColE1 *ori* raised their copy number in *E. coli* to that typical of pUC19 and related plasmids.

The broad-host-range cloning vectors pCM62 and pCM66 have a number of features that simplify their routine use: 1) high-copy in *E. coli*; 2) small size (7.2 and 8.0 kb, respectively); 3) complete sequences; 4) variety of unique restriction sites; 5) blue-white screening via *lacZα*; 6) conjugative mobilization between bacterial species, and; 7) readily adaptable into species-specific promoter-probe and expression vectors. Additionally, a number of proteobacterial species other than *E. coli* and *M. extorquens* AM1 have been found to maintain pCM62 and/or pCM66. These include the α-proteobacteria *Agrobacterium tumefaciens* (L. Chen & E. Nester, personal communication), *Methylobacterium* strains CM4 and DM4 (115), *Ralstonia eutropha* (O. Lenz & B. Friedrich, personal communication), and *Rhodobacter sphaeroides* (J. Hickman & T. Donohue, personal communication), and the γ-proteobacteria *Methylococcus capsulatus* Bath (114) and *Pseudomonas aeruginosa* (T. Motley & S. Lory, personal communication). These reports suggest that the improved broad-host range cloning vectors and the promoter-probe vectors described below will be generally applicable to a wide variety of gram-negative bacteria.

Low-background promoter-probe vectors. The second suite of broad-host-range tools generated were promoter-probe vectors for use in *M. extorquens* AM1, as well as other bacterial species. Two large IncP promoter probe vectors have been used in *M. extorquens* AM1, pHX200 bearing *xylE* as a reporter (133) and the *lacZ*-containing pGD500 (33). These plasmids are each greater than 20 kb in size, are not fully sequenced, and have limited cloning sites available upstream of

their reporter genes. In addition, pGD500 has a high background reporter activity in *M. extorquens* AM1 (86). In order to facilitate the identification and dissection of promoter regions in *M. extorquens* AM1, I created two promoter probe vectors, pCM130 and pCM132, that are based upon the cloning vectors pCM62 and pCM66, respectively (Fig. 2.3). These promoter-probe vectors were created from their respective cloning vectors in two cloning steps. The reporter genes *xylE* and *lacZ* were first introduced into one side of the polylinker. This was followed by the introduction of a transcriptional terminator into the opposite side of the polylinker, in order to reduce background activity. The resulting broad-host-range promoter-probe vectors have low background reporter gene activity in *M. extorquens* AM1 (Table 2.4) as compared to pHX200, facilitating their use in analyzing weak promoters as well as strong promoters.

In order to test the utility of pCM130 and pCM132 as promoter-probe vectors, I introduced the strong promoter upstream of the methanol dehydrogenase operon of *M. extorquens* AM1, P_{mxaF} (86), to produce pCM131 and pCM133, respectively. High reporter activities were obtained with both constructs (Table 2.4). pCM131 showed an approximately two-fold increase in XylE activity in methanol-grown cells as compared to succinate-grown.cells, similar to previously reported results (111). An analogous increase in β-galactosidase activity was not observed with pCM133. This result supports previous data suggesting the inability to use *lacZ* as a plasmid-borne reporter gene to monitor P_{mxaF} regulation (86). These results demonstrate the utility of pCM130 and pCM132 to locate, and potentially dissect, promoter regions. I expect that these low-background broad-host-range promoter-probe vectors, and variants developed from them, will prove similarly useful in locating promoter regions in a wide variety of bacterial species.

Plasmid-borne expression vectors. The development of the broad-host-range cloning vectors pCM62 and pCM66 opened the door for the generation of overexpression vectors for use in *M. extorquens* AM1. The three-step approach outlined below can be readily adapted to generate expression tools based on other promoters of *M. extorquens* AM1 or for other bacterial species that can maintain these plasmids. 1) Amplify by PCR a known promoter region and insert this fragment into the P_{lac}-proximal side of the pCM62 or pCM66 polylinker, taking advantage of sites introduced on the primers, if necessary. 2) Fill in the upstream site used for the insertion of the promoter fragment to preserve it for eventual use as a unique site in the polylinker. 3) Design and insert a linker with ends compatible to the downstream site used for insertion that can reconstitute the original polylinker within *lacZ*. This strategy allows the development of an expression vector that retains all of the advantages outlined for these cloning vectors.

For the conversion of pCM62 into an expression vector specifically designed for use in *M. extorquens* AM1, I chose to utilize P_{mxaF} (86). In the promoter-probe vectors described above, this promoter provides high-level expression during growth on both methylotrophic and non-methylotrophic substrates. The expression vectors pCM80 and the kanamycin resistant derivative, pCM160 (Fig. 2.4), were constructed using the strategy outlined above.

As a first test of the ability of pCM80 to express cloned genes in *M. extorquens* AM1, I used the gene encoding GFP from pGFPuv (Clontech). *gfp* was cloned into the cloning vector pCM62 (containing the P_{lac} promoter) and into pCM80 to produce pCM87 and pCM88, respectively. *In vivo* GFP activity of *M. extorquens* AM1 bearing pCM87, pCM88, or the 'empty' vector pCM80 was assessed on plates by observing the fluorescence of colonies during a brief exposure to UV light from a UV transilluminator (TFX-35M, Life Technologies). *M. extorquens* AM1 bearing pCM88 exhibited substantial fluorescence upon UV exposure, whereas colonies with pCM87 were indistinguishable from those carrying the 'empty' vector pCM80. These results suggested that substantial expression of a heterologous gene could be achieved in *M. extorquens*

AM1 using pCM80, while the P_{lac} promoter of pCM62 did not appear to direct significant expression.

In order to quantify gene expression in *M. extorquens* AM1 using pCM80, two constructs, pCM63 and pCM81, were made that contain *xylE* inserted into the polylinker of either pCM62 or pCM80, respectively. Only low-level XylE expression in *M. extorquens* AM1 was observed from the P_{lac} of pCM62 (Table 2.4). Similar results using plasmids bearing the P_{lac} derivatives P_{tac} and P_{spac} further suggest that P_{lac}-derived expression vectors will not be useful for high-level expression in *M. extorquens* AM1. In contrast, a high level of XylE activity was detected in extracts prepared from *M. extorquens* AM1 bearing pCM81, which expresses XylE from the P_{mxaF} of pCM80 (Table 2.4). Similar results were found for a pRK310-derived expression vector that bears P_{mxaF}. The modest induction observed using pCM81 during growth on methanol is similar to that observed from cells bearing the plasmid pCM131, however, the magnitude of XylE activity is roughly two-fold higher. pCM131 differs from pCM81 by the presence of the t_{rrnB} upstream of P_{mxaF} in pCM131 and the orientation of P_{mxaF}::*xylE* relative to the remainder of the plasmid (Figs. 2.3 and 2.4).

One concern about the use of pCM80 is its high copy-number in *E. coli* and the presence of P_{lac} upstream of P_{mxaF}. This could lead to significant expression of cloned genes in *E. coli*, preventing the use of this vector for the expression of genes in *M. extorquens* AM1 that are toxic in *E. coli*. Accordingly, XylE activity was measured in cell extracts prepared from *E. coli* strain JM109 bearing various plasmids that had been grown in LB and harvested at an OD_{600} of 0.6 – 1.0 (Table 2.4). Significant expression was observed from both pCM63 and pCM81 in JM109. An additional expression vector, pCM110, was constructed from an IncP plasmid that lacks both the ColE1 *ori* and P_{lac}. A construct containing *xylE* cloned into pCM110 (pCM111) was created to compare its XylE expression level to that of pCM81 in both organisms. pCM111 provided a high level of expression in *M. extorquens* AM1, 1.5-2 fold higher than that achieved with

pCM81, however, extracts from *E. coli* JM109 bearing pCM111 had a basal level of XylE activity, nearly at the background. These results indicate that the P_{mxaF} is expressed at very low levels in *E. coli.* Therefore, while facile expression vectors such as pCM80 are useful for the expression of most cloned genes, an alternative expression vector, such as pCM110, may be required to express genes that are toxic in *E. coli.*

In order to determine the percentage of the total cell protein that could be obtained by using pCM80, SDS-PAGE was used to estimate the relative content of the XylE and GFP polypeptides expressed from pCM81 and pCM88, respectively (Fig. 2.8). Protein bands could be identified specifically in the extracts from cells containing pCM81 and pCM88 that corresponded to the 35.1 and 26.8 kDa molecular masses predicted for XylE and GFP, respectively. Image analysis of the Coomassie stained SDS-PAGE gel indicated that pCM80 can express heterologous proteins at 5 to 9% of the total cell protein, with a modest increase in protein levels during growth on methanol relative to succinate. This level of induction corroborates the XylE activity data obtained from *M. extorquens* AM1 containing pCM81.

The final test of the utility of the expression vector pCM80 was to determine whether the expression level achieved is sufficient to complement a mutant of a highly expressed gene. For this, I chose to introduce a promoterless *mxaF* into pCM80 and examine the growth of an *mxaF* mutant strain bearing this plasmid on methanol. *mxaF* was amplified by PCR and cloned into both pCM62 and pCM80 to produce pCM85 and pCM86. These two plasmids were then introduced into the *mxaF* mutant UV26 (89) by conjugation. UV26 carrying pCM86, which expresses *mxaF* from the P_{mxaF} of pCM80, grew at the same rate on methanol plates as wt *M. extorquens* AM1 with the empty vector, pCM80. UV26 carrying pCM85, which expresses *mxaF* from the P_{lac} of pCM62, however, grew very slowly on methanol plates and was indistinguishable from UV26 carrying the empty vector pCM62. These results indicate that overexpression of

MxaF from pCM80, but not the low-level expression from the cloning vector pCM62, is sufficient to replace that from P_{mxaF} on the chromosome.

Development of a broad-host-range *cre-lox* based antibiotic marker recycling system.

***The work described in this section has been published (78).**

A broad-host-range antibiotic marker recycling system was developed for use in *M. extorquens* AM1 so that unmarked mutant strains could be generated and used to obtain strains bearing multiple genetic manipulations. This *cre-lox* based system consists of two parts. First, an allelic exchange vector with a *loxP*-flanked kanamycin resistance cassette, pCM184, was generated (Fig. 2.5). Second, a tetracycline-resistance-conferring broad-host-range plasmid, pCM157, was generated to express the Cre recombinase to allow the in vivo excision of the *loxP*-flanked *kan* cassette. The strategy outlined for this system is diagramed in Figure 2.9.

In order to test this system, a construct for the generation of *mxaF* (encodes the large subunit of MDH) mutants, pCM194, was generated and introduced into wild-type using the *E. coli* helper strain S17-1 (107). Kanamycin-resistant transconjugants obtained on succinate medium containing rifamycin were screened for tetracycline sensitivity to identify potential null (double crossover) mutants. Collectively, the work performed with pCM184-based constructs has led to a frequency of double-crossover events from 5% to 80%, depending on the construct. A Δ*mxaF::kan* mutant, CM194K.1, was chosen for further study. The plasmid pCM157 was then introduced by conjugation into CM194K.1 using the helper plasmid pRK2073 (36). Tetracycline-resistant strains were streaked for purity until the resulting strain produced only kanamycin-sensitive colonies (generally only two transfers). Subsequently, pCM157 was cured from the strain by two successive transfers on medium lacking tetracycline to produce the Δ*mxaF* strain CM194.1. Analytical PCR was performed with wild-type *M. extorquens* AM1, CM194K.1, and

CM194.1 for confirmation of allelic exchange, and subsequent deletion of the kanamycin cassette. Where examined, the sequence of the analytical PCR product indicated faithful recombination between the *loxP* sites. Diagnostic PCR confirmed replacement of each deleted gene with *kan*, and the subsequent excision of *kan* to produce the unmarked deletion. As expected, both the marked and unmarked *mxaF* mutants CM194K.1 and CM194.1 grew like wild-type on succinate but failed to grow on methanol, as reported for the *mxaF* mutant UV26 (89).

Transposon mutagenesis in *M. extorquens* AM1 using IS*phoA*/hah-Tc and identification of ten novel methylotrophy genes including *ftfL* and *dmrA*.

***The work described in this section has been published and was partially carried out by two undergraduate students under my direction, Brooke N. O'Brien and Jennifer Breezee (80).**

Eighty-six genes involved in C_1 metabolism in *M. extorquens* AM1 have been identified (24, 69). These include genes involved in methanol oxidation, methylamine oxidation, formaldehyde oxidation, formate oxidation, and the serine cycle, most of which are organized in a small number of large gene clusters. Since the known methylotrophy gene clusters have now been analyzed, a search for any remaining methylotrophy genes requires more global approaches. Although previous attempts to identify methylotrophy genes using transposon mutagenesis in *M. extorquens* AM1 have only revealed known genes (131), miniTn*5* was used successfully with the closely related strain *Methylobacterium chloromethanicum* CM4 (122). Therefore, the transposons Tn*phoA* (76), miniTn*5*-Tc (32), and IS*phoA*/hah-Tc (10, 29, 75) were tested for efficient recovery of transposon-containing mutants. Only IS*phoA*/hah-Tc was found to provide efficient transposition.

In order to test the utility of IS*phoA*/hah-Tc mutagenesis in *M. extorquens* AM1, a mutant screen was performed to isolate methanol-defective mutants. A pool of IS*phoA*/hah-Tc mutants

was plated onto succinate plates that contained both rifamycin and tetracycline for selection. From the 24,000 transposon insertion colonies obtained on succinate, 55 mutants with defective growth on methanol were isolated. To further classify the metabolic defect in each strain, the methanol-defective mutants were tested for growth defects on methylamine, formate, or ethylamine. Defects in methanol oxidation will show a growth defect only on methanol, while defects in the serine cycle and H_4F or H_4MPT pathways will show a growth defect on all C_1 compounds. A portion of C_1 and C_2 metabolism overlaps in the conversion of acetyl-CoA to glyoxylate, so defects in this part of metabolism will show growth defects on all C_1 and C_2 substrates (62, 63). Strains were also tested on medium containing both succinate and methanol to identify methanol-sensitive mutants, a phenotype specifically associated with defects in H_4MPT-dependent formaldehyde oxidation (43, 127). Of the 55 methanol-defective mutants, 22 were only defective on methanol, 11 were C_1-defective, 19 were C_1/C_2-defective, and three were C_1-defective/methanol-sensitive.

The site of IS*phoA*/hah-Tc insertion in each methanol-defective mutant strain was determined using a semi-random two-step PCR procedure (ST-PCR) with subsequent sequencing and analysis as described in Materials and Methods. The 55 independent IS*phoA*/hah-Tc mutant strains with a methanol growth defect contained insertions into 31 different genes, 12 of which were represented more than once. In those cases in which multiple insertions were identified in the same gene, the insertions were located in different sites, and the mutants exhibited a generally consistent mutant phenotype. Of the 86 previously characterized genes involved in methylotrophy, only 49 give clear growth defects on methanol. The others are either involved in the oxidation of methylamine (*mau* genes), are genes involved in redundant functions for which single mutants do not generate a growth phenotype, or carry out functions in both heterotrophic and methylotrophic metabolism so that null mutants cannot be isolated. The transposon mutagenesis identified 21 known methylotrophy genes, which is about half of the methylotrophy

genes for which methanol-defective mutants would be expected, indicating that the mutagenesis was not saturating. In addition to the 21 characterized genes, however, this screen also identified 10 previously uncharacterized methylotrophy genes (Table 2.5).

The mutants containing insertions into the 10 novel methylotrophy genes fell into four different mutant classes (Table 2.5). Only one methanol-defective strain contained an insertion into a novel gene. This *orf*, designated here as *mxdA* (methanol oxidation, cluster D), is unlinked to previously known methylotrophy genes. The predicted amino acid sequence of MxdA has 60% identity and 77% similarity to a putative dehydrogenase of *Mesorhizobium loti* (BAB50156) and 52% identity and 68% similarity to the putative zinc-binding alcohol dehydrogenase encoded by *yhdH* of *E. coli* (P26646). This gene will be studied by another researcher in the laboratory.

Seven of the C_1-defective mutants contained insertions into two previously uncharacterized genes that are unlinked to known methylotrophy genes (Table 2.5). Four independent mutants were obtained with insertions into an orf with significant similarity to *cbbR* (LysR-type transcriptional regulator of autotrophy). The predicted amino acid sequence of the *M. extorquens* AM1 CbbR homolog has 44% identity and 61% similarity to the CbbR in *Xanthobacter flavus* (P25545). This regulator, since renamed QscR (Quayle serine cycle regulator) has subsequently been shown to regulate the synthesis of the major serine cycle operon that also contains *mtdA* and *fch* (56). Three mutants were also found with insertions into an orf that is predicted to encode FtfL (formate-tetrahydrofolate ligase). The *M. extorquens* AM1 *ftfL* has a predicted amino acid sequence with 65% identity and 78% similarity to the FtfL of *M. loti* (BAB49812). The role of this gene product in methylotrophy is discussed in Chapter 4.

Eight C_1/C_2-defective mutants were isolated with insertions into 6 genes previously uncharacterized in *M. extorquens* AM1 and unlinked to previously known C_1 genes. Two C_1/C_2-defective mutants were found to contain insertions into genes apparently involved in molybdenum cofactor biosynthesis. Mutant S06-42 contains an insertion into an orf with

significant similarity to *moaA* (molybdenum cofactor biosynthesis, protein A), which encodes a protein with 73% identity and 83% similarity to the MoaA of *Brucella melitensis* (AAL52200). Mutant S165-21 contains an insertion into an orf with significant similarity to *mobA* (homolog of molybdopterin-guanine dinucleotide biosynthesis protein A), which has a predicted amino acid sequence with 56% identity and 64% similarity to the MobA from *Rhodobacter sphaeroides* (Y09560). Additionally, four C_1/C_2-defective mutants were obtained in genes of uncertain function, here designated as *mea* genes (methyl- and ethyl-assimilation defective). This designation has been used previously for genes required for C_1 and C_2 growth (109). Mutant S06-17 contains an insertion into an orf, here designated as *meaC,* which encodes a conserved hypothetical protein with 67% identity and 80% similarity to a MaoC-family protein of *Caulobacter crescentus* (AAK25618). Three mutants were isolated with an insertion into an orf, here designated as *meaD*, encoding another conserved hypothetical protein. The predicted amino acid sequence of *meaD* has 71% identity and 82% similarity to an ortholog in *B. melitensis* (AAL51247) and 43% identity and 56% similarity to the PduO protein of *Salmonella typhimurium* LT2 (AAD39014). Mutants S109-30 and S151-59 contain insertions into orfs, here designated as *meaE* and *meaF*, which encode small hypothetical proteins with no significant similarity to other known proteins. These genes are under study by Dr. N. Korotkova in the Lidstrom laboratory.

Finally, one C_1^-/methanol-sensitive mutant strain (S64-67) was identified in a previously uncharacterized orf unlinked to all known methylotrophy genes. Previously, this phenotype had only been observed for *M. extorquens* AM1 mutants defective for the H_4MPT-dependent enzymes *fae* (encodes formaldehyde-activating enzyme) (127) and *mtdB* (encodes a methylene-H_4MPT dehydrogenase) (43), and it is believed to be caused by the inability of these mutant strains to detoxify formaldehyde. This orf, here designated as *dmrA* (putative dihydromethanopterin reductase), encodes a protein (GenBank accession number AY093431) that

has 26% identity and 44% similarity to the dihydrofolate reductase (DHFR) of *Lactobacillus casei* (P00381). Further analysis of *dmrA* mutants is described in Chapter 3.

Development of an insertional expression vector system for use in *M. extorquens* AM1. A P_{mxaF}-based insertional expression system for *M. extorquens* AM1 was developed in the following manner. First, the insertional vector backbone was generated through a series of cloning steps that were performed in order to remove restriction sites for use in the final vectors. The resulting plasmid, pCM167, has a ColE1 replicon and a *loxP*-flanked tetracycline-resistance cassette (*tetAR*) inserted into the *M. extorquens* AM1 *katA* (which encodes a catalase) (24). Second, a construct bearing a terminator-bounded MCS cassette, pCM124, was generated by introducing the *E. coli* t_{rrnB} and the t_{T7} on opposite ends of the MCS cassette, and then the strong P_{mxaF} from *M. extorquens* AM1 was cloned into this to generate pCM126. Third, the terminator-bounded cassettes from pCM124 and pCM126 were introduced into pCM167 to generate the insertional cloning vector pCM168 and the insertional expression vector pCM172 (Fig. 2.7). Finally, a *katA::kan* strain, CM82.1 was generated in order to screen transformants for the desired insertion events. As had been reported previously for *M. extorquens* AM1 *katA* mutants (24), this strain exhibits no growth defects under all conditions tested.

Insertional vector constructs were introduced into the *katA::kan M. extorquens* AM1 strain CM82.1 by electroporation and selected on plates containing tetracycline. Tetracycline-resistant transformants can then be screened for kanamycin sensitivity to ensure complete allelic exchange occurred, rather than simply vector integration. This was performed with the empty vectors pCM168 and pCM172 to generate the *katA::(loxP-tetAR-loxP-*t_{rrnB}-MCS-t_{T7}*)* strain CM168T.1 and the *katA::(loxP-tetAR-loxP-*t_{rrnB}-P_{mxaF}-MCS-t_{T7}*)* strain CM172T.1. Additionally, in order to test the expression level afforded by this system, *gfp* was cloned behind the P_{mxaF} in pCM172, resulting in pCM174, and this construct was introduced into CM82.1 to generate

CM174T.1 (*katA::(loxP-tetAR-loxP-*t_{rrnB}*-*P_{mxaF}*-gfp-*t_{T7}*)*). Subsequently, the *cre*-expression plasmid pCM158 (78) was introduced into each of these strains to excise the tetracycline-resistance cassette. This resulted in the unmarked (tetracycline-sensitive) strains, CM168.1 (*katA::(loxP-*t_{rrnB}-MCS-t_{T7}*)*), CM172.1 (*katA::(loxP -*t_{rrnB}*-*P_{mxaF}-MCS-t_{T7}*)*), and CM174.1 (*katA::(loxP-*t_{rrnB}*-*P_{mxaF}*-gfp-*t_{T7}*)*). Diagnostic PCR indicated that all desired recombination events occurred as desired (data not shown). Finally, as with CM82.1, all of these strains grew identically to wild-type *M. extorquens* AM1, confirming that the introduction of the insertional system at the *katA* locus does not result in measurably altered growth under standard laboratory conditions.

The expression level afforded by the insertional expression vector was determined by comparing GFP fluorescence in CM174.1 to wild-type *M. extorquens* AM1 carrying the plasmid pCM88, which has *gfp* driven by P_{mxaF} of the plasmid expression vector pCM80. Succinate-grown wild-type with pCM88 had a fluorescence/OD_{600} of 320 or 130 for methanol- and succinate-grown cells, compared to a relative fluorescence/ OD_{600} of 680 versus 240 for CM174.1 grown under the same conditions (Table 2.4). Thus, for the case of GFP, the insertional expression vector provided two-fold higher expression than with the plasmid system but exhibited a similar expression ratio (2.6-fold induction on methanol) to that observed with pCM88 (2.4-induction). Additionally, the termination efficiency of t_{rrnB} and t_{T7} in *M. extorquens* AM1 was examined by inserting each of the terminators between P_{mxaF} and *xylE* (which encodes catechol 2,3-dioxygenase). The XylE activities of the parental plasmid, pCM77 (*P*$_{mxaF}$*-xylE*), were 800 and 190 mU in extracts prepared from methanol- and succinate-grown cultures (Table 2.4). These values dropped to 5 and 2 for pCM121 (*P*$_{mxaF}$*-*t_{rrnB}*-xylE*) and 290 and 95 for pCM122 (*P*$_{mxaF}$*-*t_{T7}*-xylE*) (Table 2.4). Therefore, the *E. coli* t_{rrnB} terminator provided a 99% reduction in activity, compared to only a 50 – 64% reduction by t_{T7}. Collectively, these data indicate that the insertional

expression vector pCM172 provides significant expression from a chromosomal locus that is largely transcriptionally isolated from the surrounding genes.

Discussion:

Here I have described the development of four sets of genetic tools that have proven critical for the genetic and physiological analysis of methylotrophy described in the following chapters. First, I have created improved broad-host-range cloning vectors and low-background promoter-probe vectors, and demonstrated a simple strategy to convert these plasmids into species-specific expression vectors. This suite of tools possesses a number of advantages over previously described broad-host-range genetic tools in terms of their ease of use and their ability to be adapted for multiple purposes. The observation that these vectors can be maintained in a wide variety of bacterial species suggests that these vectors, and future elaborations upon them, will help fulfill the need for more sophisticated broad-host-range genetic tools in a variety of bacteria. These vectors have been sent out to dozens of laboratories world-wide.

Second, the broad-host-range *cre-lox* antibiotic marker recycling system offers the possibility to create unmarked mutants in *M. extorquens* AM1 and a wide variety of Gram-negative bacteria. Utilization of allelic exchange with counter-selection against integrants, and an inherently unstable minimal IncP Cre expression plasmid, obviates the need for successful negative selection in the target organism, a feature of some previously developed marker recycling systems (48). Use of PCR to generate flanks for gene replacement allows for the facile generation of precise deletion mutants (Fig. 2.9), as well as truncations through the introduction of start or stop codons in the primers, as needed. Variants of this system can be readily developed to allow the construction of chromosomal transcriptional or translational fusions. Marker recycling systems such as the one described here offer a substantial advantage over standard allelic exchange methods due to the fact that these can be used iteratively to enable generation of

unmarked strains bearing multiple genetic modifications. For example, this system was utilized to generate a *M. extorquens* AM1 strain bearing four separate mutations (C. J. Marx, L. Chistoserdova, unpublished data). Finally, engineered strains generated with these tools are more acceptable for environmental release owing to the absence of introduced antibiotic resistance markers. This system should also be broadly applicable, and our lab is beginning to send these strains out to many laboratories.

Third, I have reported the first successful application of transposon mutagenesis to the study of *M. extorquens* AM1 and bring the total number of known methylotrophy genes in this organism to 96. Ten novel methylotrophy genes have been identified, over half of which have been the subject of intense investigation in our laboratory, including *ftfL* and *dmrA* in my own work (see Chapters 3 and 4). Since only about half of the known methylotrophy genes were identified in the mutagenesis screen in each of the four mutant categories, it is expected that several more methylotrophy genes are as yet undiscovered.

Finally, I have described the development of a system for the insertion and expression of genes from a stable, unmarked chromosomal locus. The P_{mxaF}-based insertional expression vector pCM172 provided significant expression during growth on either methanol or succinate. pCM168 can provide a starting point for either the development of other expression systems based on different endogenous or exogenous promoters, or insertional promoter-probe vectors generated through the introduction of reporter genes. Additionally, the terminator-bounded MCS cassettes and the *loxP*-flanked antibiotic resistance cassette used to generate this system could easily be inserted into a different backbone to generate insertional systems that recombine into other loci of *M. extorquens* AM1, or other bacteria. Insertional systems such as the one described here have a significant advantage over plasmid-borne systems in that antibiotic selection is not required for maintenance.

Table 2.1: *M. extorquens* AM1 strains described in Chapter 2

Strain	Description	Source
CM82.1	*katA::kan*	This chapter
CM168.1	*katA::(loxP-*t_{rrnB}*-*MCS*-*t_{T7}*)*	This chapter
CM168T.1	*katA::(loxP-tetAR-loxP-*t_{rrnB}*-*MCS*-*t_{T7}*)*	This chapter
CM172.1	*katA::(loxP-*t_{rrnB}*-*P_{mxaF}*-*MCS*-*t_{T7}*)*	This chapter
CM172T.1	*katA::(loxP-tetAR-loxP-*t_{rrnB}*-*P_{mxaF}*-*MCS*-*t_{T7}*)*	This chapter
CM174.1	*katA::(loxP-*t_{rrnB}*-*P_{mxaF}*-gfp-*t_{T7}*)*	This chapter
CM174T.1	*katA::(loxP-tetAR-loxP-*t_{rrnB}*-*P_{mxaF}*-gfp-*t_{T7}*)*	This chapter
CM194.1	Δ*mxaF*	This chapter
CM194K.1	Δ*mxaF::kan*	This chapter
M. extorquens AM1 Rif^R derivative		(89)
S06-17	*meaC::IS*phoA*/hah-Tc*	This chapter
S06-42	*moaA::IS*phoA*/hah-Tc*	This chapter
S09-09	*mxdA::IS*phoA*/hah-Tc*	This chapter
S33-90	*meaD::IS*phoA*/hah-Tc*	This chapter
S57-87	*ftfL::IS*phoA*/hah-Tc*	This chapter
S64-67	*dmrA::IS*phoA*/hah-Tc*	This chapter
S65-76	*ftfL::IS*phoA*/hah-Tc*	This chapter
S90-71	*cbbR::IS*phoA*/hah-Tc*	This chapter
S95-64	*ftfL::IS*phoA*/hah-Tc*	This chapter
S109-30	*meaE::IS*phoA*/hah-Tc*	This chapter
S151-59	*meaF::IS*phoA*/hah-Tc*	This chapter
S165-21	*mobA::IS*phoA*/hah-Tc*	This chapter
S203-07	*cbbR::IS*phoA*/hah-Tc*	This chapter
S203-14	*meaD::IS*phoA*/hah-Tc*	This chapter
S235-22	*meaD::IS*phoA*/hah-Tc*	This chapter
S235-75	*cbbR::IS*phoA*/hah-Tc*	This chapter
S252-39	*cbbR::IS*phoA*/hah-Tc*	This chapter
UV26	*mxaF* mutant	(89)

Table 2.2: Plasmids described in Chapter 2

Plasmid	Description	Source
pAYC61	Mobilizable allelic exchange vector	(17)
pBBR1MCS-2	Cloning vector from *Bordatella bronchiseptica*	(64)
pCM20	pCR2.1 with *xylE*	This chapter
pCM21	pCR2.1 with *gfp*	This chapter
pCM22	pMTL23 with *xylE*	This chapter
pCM23	pMTL23 with *gfp*	This chapter
pCM27	pCR2.1 with P_{mxaF} PCR product	This chapter
pCM46	pDN19 with *traJ'* allele from pDN19X	This chapter
pCM47	pDN19X with wt *traJ* from pDN19	This chapter
pCM48	pTJS75 with *traJ'* allele from pDN19X	This chapter
pCM50	Self-ligation of pCM48 *Mun*I-*Bln*I	This chapter
pCM51	Self-ligation of pCM50 *Hin*dIII-*Xmn*I	This chapter
pCM60	pCM50 derivative	This chapter
pCM62	Fusion of pUC19 and pCM51	This chapter
pCM63	pCM62 with *xylE*	This chapter
pCM64	pCM62 with P_{mxaF}	This chapter
pCM66	*kan* inserted into *tetA* of pCM62	This chapter
pCM66LacZ	pCM66 with *lacZ* (opposite orientation as P_{lac})	This chapter
pCM74	pCR2.1 with *mxaF* PCR product	This chapter
pCM75	pMTL23 with *xylE*	This chapter
pCM76	pCM62 with *xylE* replacing P_{lac}	This chapter
pCM77	pCM76 with P_{mxaF} upstream of *xylE*	This chapter
pCM79	pCM64 with *Hin*dIII filled-in	This chapter
pCM80	pCM79 with linker inserted	This chapter
pCM81	pCM80 with *xylE*	This chapter
pCM82	pAYC61 with *katA::kan* fragment from pLC1128.Km	This chapter
pCM85	pCM62 with *mxaF*	This chapter
pCM86	pCM80 with *mxaF*	This chapter
pCM87	pCM62 with *gfp*	This chapter
pCM88	pCM80 with *gfp*	This chapter

Table 2.2: Plasmids described in Chapter 2 (Continued)

Plasmid	Description	Source
pCM110	pCM60 with P_{mxaF} replacing P_{lac}	This chapter
pCM111	pCM110 with *xylE*	This chapter
pCM116	pLC11.28 with *katA::erm*	This chapter
pCM117	pCM116 with restriction sites removed	This chapter
pCM118	pCM117 with a restriction site filled-in	This chapter
pCM119	pCR2.1 with t_{rrnB}	This chapter
pCM120	pCR2.1 with t_{T7}	This chapter
pCM121	pCM77 with P_{mxaF}-t_{rrnB}-*xylE*	This chapter
pCM122	pCM77 with P_{mxaF}-t_{T7}-*xylE*	This chapter
pCM123	pMTL23 with t_{rrnB}	This chapter
pCM124	pMTL23 with t_{rrnB}-MCS-t_{T7} cassette	This chapter
pCM126	pMTL23 with t_{rrnB}-P_{mxaF}-MCS-t_{T7} cassette	This chapter
pCM130	pCM76 with t_{rrnB} of *E. coli*	This chapter
pCM131	pCM130 with P_{mxaF}	This chapter
pCM132	pCM66lacZ with t_{rrnB} from *E. coli*	This chapter
pCM133	pCM132 with P_{mxaF}	This chapter
pCM157	pCM62 with *cre* from pJW168	This chapter
pCM158	pCM66 with cre from pJW168	This chapter
pCM159	pLox1 with *tetAR*	This chapter
pCM160	*kan* inserted into *tetA* of pCM80	This chapter
pCM161	pLox1 with *kan* from pUC4K	This chapter
pCM165	pCM118 with *loxP-tetAR-loxP*	This chapter
pCM166	pCM165 with restriction sites removed	This chapter
pCM167	pCM166 with a restriction site filled in	This chapter
pCM168	pCM167 with t_{rrnB}-MCS-t_{T7}	This chapter
pCM172	pCM167 with t_{rrnB}-P_{mxaF}-MCS-t_{T7}	This chapter
pCM174	pCM172 with *gfp*	This chapter
pCM182	pAYC61 with sites removed	This chapter
pCM183	pCR2.1 with *kan* amplified from pCM161	This chapter
pCM184	pCM182 with *kan* from pCM183	This chapter

Table 2.2: Plasmids described in Chapter 2 (Continued)

Plasmid	Description	Source
pCM191	pCR2.1 with *mxaF* upstream flank	This chapter
pCM192	pCR2.1 with *mxaF* downstream flank	This chapter
pCM193	pCM184 with *mxaF* downstream flank	This chapter
pCM194	pCM193 with *mxaF* upstream flank	This chapter
pCM639	IS*phoA*/hah-Tc delivery vector	(29)
pCR2.1	PCR cloning vector	Invitrogen
pDN19	Small IncP cloning vector	(88)
pDN19X	spontatneous mutant of pDN19 (in *traJ*)	This chapter
pDN411	Cosmid containing 8.6 kb *Hin*dIII region of *mxa* operon	(89)
pET-3a	*E. coli* overexpression vector	Novagen
pGFPuv	Commercial vector containing *gfp*	Clontech
pHX200	Large IncP *xylE* promoter-probe vector	(133)
pJB656	Small IncP expression vector with *xylS*/P_m	(14)
pJW168	*cre* expression plasmid	(92)
pLacZ2.1+	pCR2.1 with *lacZ* PCR fragment	(R. Meima, unpub.)
pLC11.28	pUC19 with *katA* region	(24)
pLC1128.Km	pLC11.28 with *kan* inserted into *katA*	(24)
pLC410a	Large IncP plasmid containing *mtdA* and *fchA*	(22)
pLox1	Mobilizable suicide vector with *loxP* sites	(92)
pMMB67HE.tet	IncQ expression vector with *lacI*^{q}/P_{tac}	(52)
pMTL20T$_1$T$_2$	Cloning vector with t_{rrnB} from *E. coli*	(28)
pMTL23	Cloning vector with an extensive polylinker	(16)
pMTL23E	pMTL23 with erythromycin-resistance cassette	(98)
pRK310	Large IncP cloning vector	(33)
pRK2013	Helper plasmid expressing IncP *tra* functions	(36)
pRK2073	Helper plasmid expressing IncP *tra* functions	(36)
pTJS75	Small IncP plasmid	(104)
pUC4K	Vector with kanamycin cassette	(123)
pUC19	Universal cloning vector	(134)

Table 2.3: Primers described in Chapter 2

Primer	Sequence (5′-3′)
CM-dfkMCS	GACTAGTGAGCTCACCGGTTAACACGCGTACGTAGGGCCCGCGGTAT-CGATAAGCTGGATCC
CM-DmxaF1	GACTGCAGCATGCCAAGCTTACGCGCATCGTCTCCAAGTG
CM-DmxaF2	AAGTTATGCGGCCGCCATATGCATCTCGCGGTATCTCTCAGAC
CM-DmxaF3	GCTTATCGATACCGTCGACCTCGCTGACGGCAGATGCGAAC
CM-DmxaF4	CGGTACCTGAGCTCGAGATCTAGCCGTCCTTCATCGAGAG
CM-gfpf	TAGCTGCAGTAAGCTTGTCGACTCTAGAGGATCCCC
CM-gfpr	AGGGGATCCGAGCTCGGCGCTCAGTTGGAAT
CM-L64f	GATCTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCA-TGCCTGCAGGTCGACTCTAGAG
CM-L64r	GATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATG-GTCATAGCTGTTTCCTGTGTGA
CM-mxaFf	GGCATGCGAGGAGACGCAGGATG
CM-mxaFr	CGAATTCCGGCTTCAGACGTTAC
CM-PmxaFf	TAGATCTCGACAAGCTTCCCGCTTGG
CM-PmxaFr	AGGATCCGCGGTATCTCTCAGACG
CM-TnPCR1	GTGCAGTAATATCGCCCTGAGCA
CM-TnPCR2b	TGCCACGCGTCGACTAGTACNNNNNNNNNNACGCC
CM-TnPCR3	ATCCCCCTGGATGGAAAACGG
CM-TnPCR4	GGCCACGCGTCGACTAGTAC
CM-TnSEQ	AAACGGGAAAGGTTCCGTCCA
CM-trrnBf	CGGATCCCAGTCACGACGTTG
CM-trrnBr	CGTCGACCCTGCTTTCCTGATGC
CM-tT7f	GCTCGAGATCCGGCTGCTAACAAAG
CM-tT7r	GTCGCGATATCGTCCATTCCGACAG
CM-ufkMCS	TGACGTCTAGATCTGAATTCAGCTGTACAATTGGTACCATGGATGCA-TATGGCGGCCGCA
CM-xylEf	TAGCTGCAGTAAGCTTCAGGAGGTGACGTC
CM-xylEr	AGGGATCCGAGCTCCATCAGGTGAGCACGGTC

Table 2.4: Reporter gene activities demonstrating utility of genetic tools. XylE, LacZ, activity (nmol min^{-1} mg^{-1} protein), or GFP (relative fluorescence/OD_{600}) present in cell extracts prepared from wt *M. extorquens* AM1 and *E. coli* strain JM109 demonstrating the utility of the genetic tools developed for use in *M. extorquens* AM1.

			M. extorquens AM1 grown in:	
Reporter	Plasmid	Features	methanol	succinate
XylE	pCM63	pCM62 (P_{lac}) with *xylE*	32 ± 12	37 ± 5
	pCM77	pCM62 with P_{mxaF}-*xylE*	800 ± 140	190 ± 60
	pCM81	pCM80 (P_{lac}-P_{mxaF}) with *xylE*	1180 ± 60	680 ± 30
	pCM111	pCM110 (P_{mxaF}) with *xylE*	1580 ± 210	1170 ± 10
	pCM121	pCM62 with P_{mxaF}-t_{rrnB}-*xylE*	4.7 ± 1.1	2.4 ± 0.7
	pCM122	pCM62 with P_{mxaF}-t_{T7}-*xylE*	290 ± 60	100 ± 10
	pCM130	*xylE* promoter-probe vector	1.6 ± 0.4	3.4 ± 0.4
	pCM131	pCM130 with P_{mxaF}	580 ± 100	320 ± 10
	pHX200	large *xylE* promoter-probe vector	10 ± 1	10 ± 3
LacZ	pCM132	*lacZ* promoter-probe vector	2.4 ± 0.1	9.4 ± 0.2
	pCM133	pCM132 with P_{mxaF}	2920 ± 110	4140 ± 210
GFP	pCM88	pCM80 (P_{lac}-P_{mxaF}) with *gfp*	320 ± 20	130 ± 2
	CM174.1	pCM172 (P_{mxaF}) with *gfp*	680 ± 60	260 ± 6
			E. coli grown in LB	
XylE	pCM63	pCM62 (P_{lac}) with *xylE*	2300 ± 600	
	pCM81	pCM80 (P_{lac}-P_{mxaF}) with *xylE*	1270 ± 65	
	pCM111	pCM110 (P_{mxaF}) with *xylE*	1.1 ± 0.5	

Table 2.5: IS*phoA*/hah-Tc insertions into novel methylotrophy genes. The interrupted gene, corresponding mutant strain(s), and growth phenotype on methanol (M), methylamine (Ma), formate (F), ethylamine (Ea), and succinate (S) are indicated. Three categories of growth are reported: wild-type = "++", intermediate = "+" and trace to none = "-". *Additionally, mutant S165-21 grew slowly on ethanol (20 mM).

Gene	Mutants	M	Ma	F	Ea	S
cbbR	S90-71	+	+	+	++	++
	S203-07	+	+	+	++	++
	S235-75	+	++	++	++	++
	S252-39	+	+	+	++	++
dmrA	S64-67	-	-	+	+	+
ftfL	S57-87	-	-	-	++	++
	S65-76	-	-	-	++	++
	S95-64	+	-	-	++	++
meaC	S06-17	+	-	-	-	+
meaD	S33-90	-	-	+	+	++
	S203-14	+	-	+	+	++
	S235-22	+	+	+	++	++
meaE	S109-30	-	+	+	+	++
meaF	S151-59	+	+	+	+	++
moaA	S06-42	+	+	-	+	++
mobA	S165-21	-	+	-	++*	++
mxdA	S09-09	+	++	+	++	++

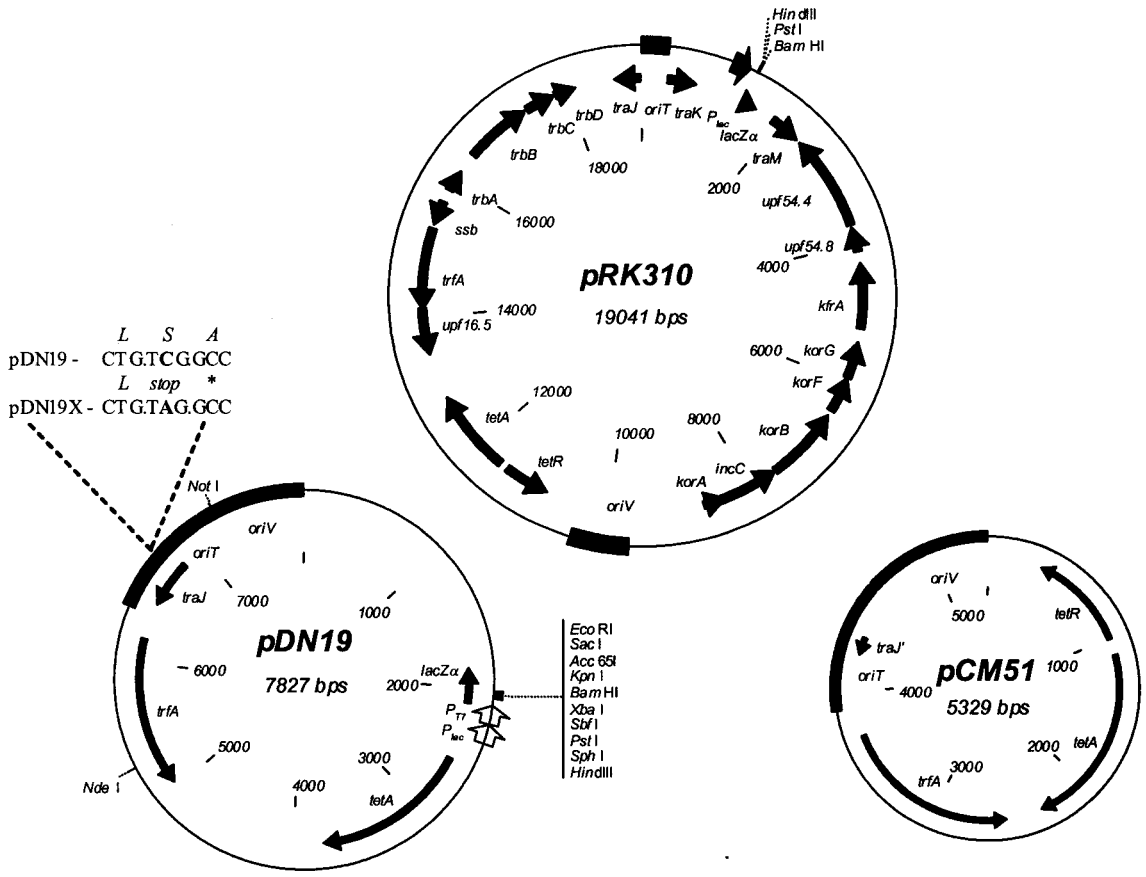


Figure 2.1: Plasmid maps of the IncP replicons pRK310, pDN19(X) and pCM51. The C to A transversion present in pDN19X that created the truncated *traJ'* allele is diagrammed. The GenBank accession numbers for these plasmids are AF327712, AF327711, and AF327713, respectively.

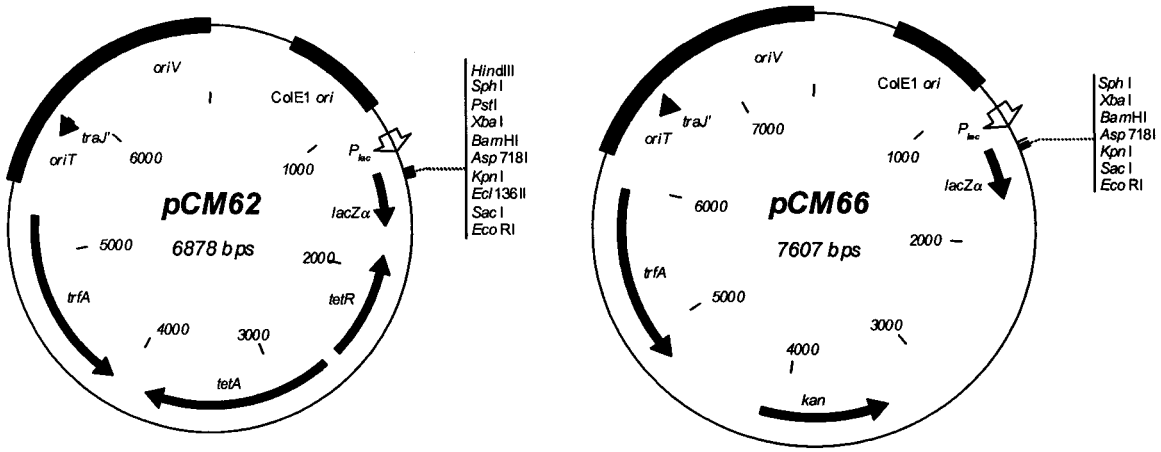


Fig. 2.2: Plasmid maps of the cloning vectors pCM62 and pCM66. The GenBank accession numbers for these plasmids are AF327714 and AF327715.

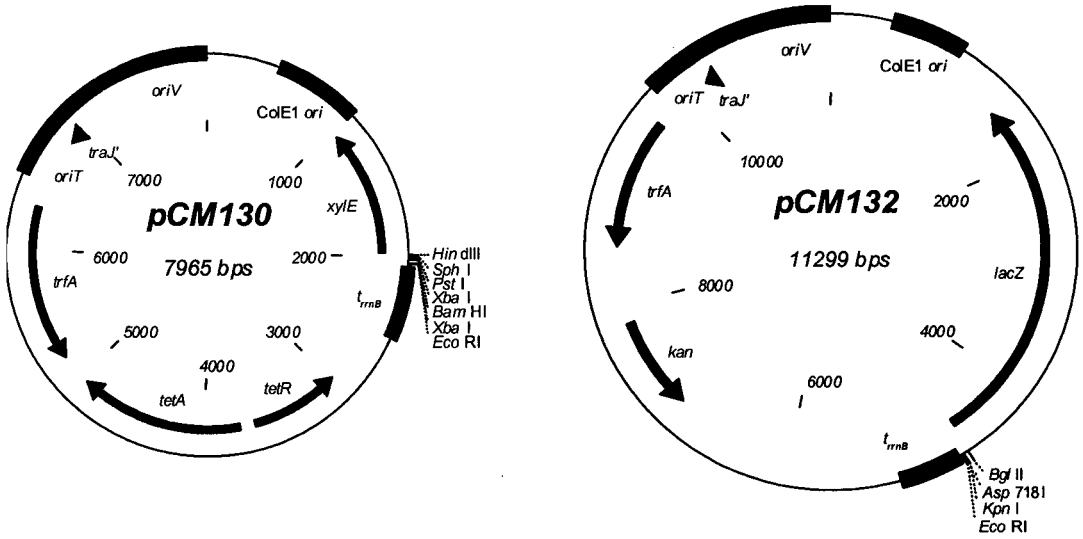


Fig. 2.3: Plasmid maps of the promoter-probe vectors pCM130 and pCM132. The GenBank accession numbers for these plasmids are AF327719 and AF327720.

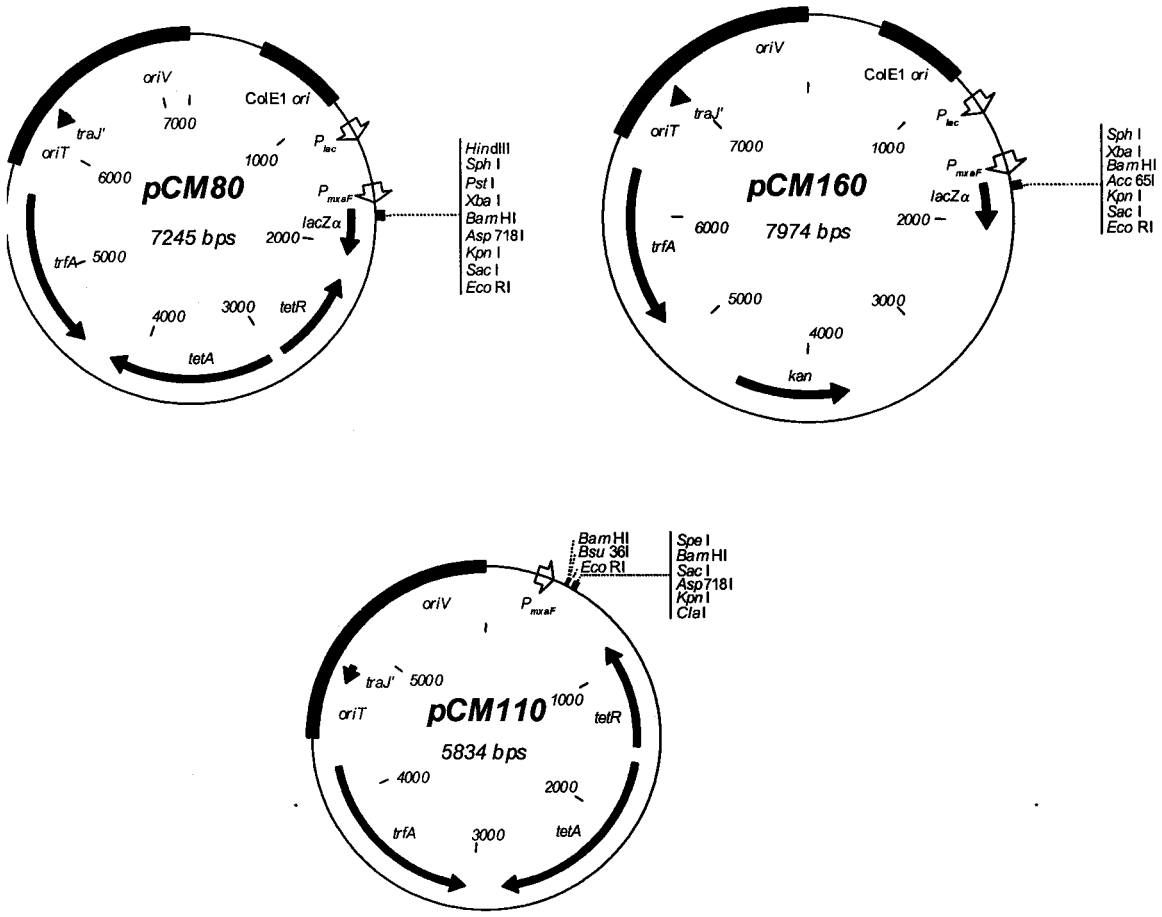


Fig. 2.4 - Plasmid maps of the expression vectors pCM80, pCM110 and pCM160. The GenBank accession numbers for these plasmids are AF327716, AF327718 and AF327717, respectively.

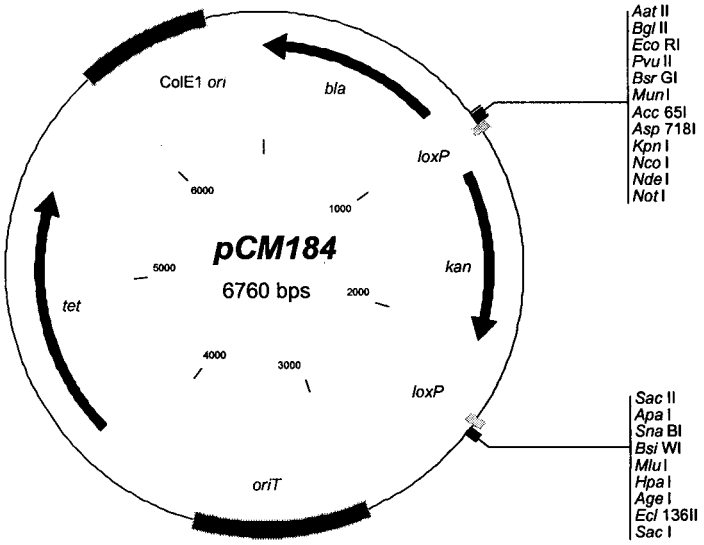


Figure 2.5: Plasmid map of the allelic exchange vector pCM184. The GenBank accession number for this plasmid is AY093429.

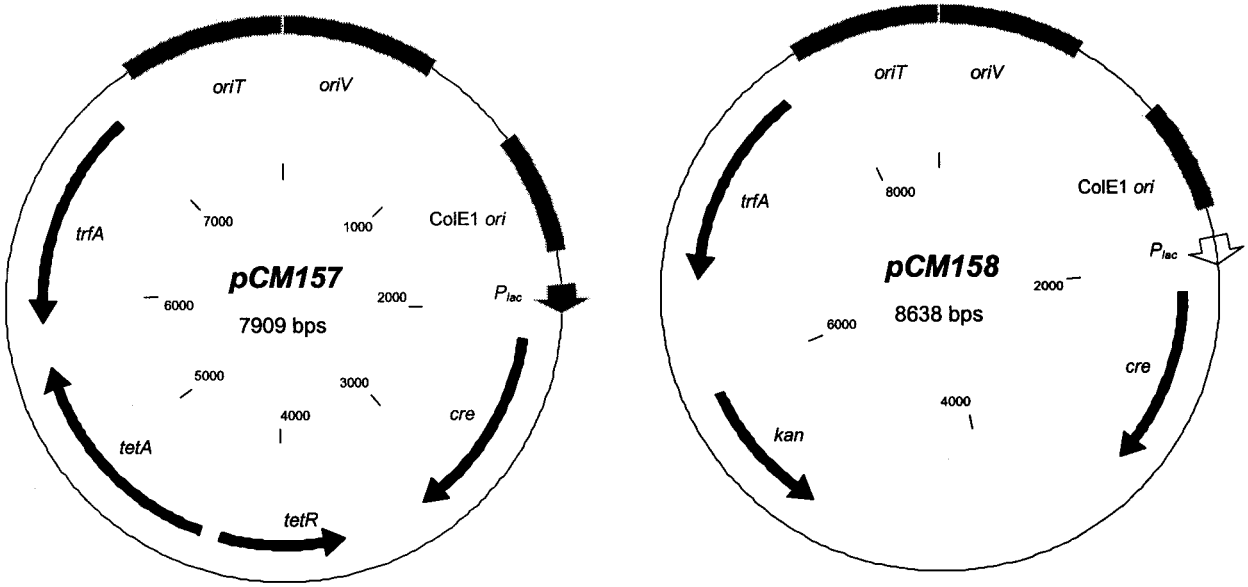


Figure 2.6: Plasmid map of the *cre*-expression plasmids pCM157 and pCM158.

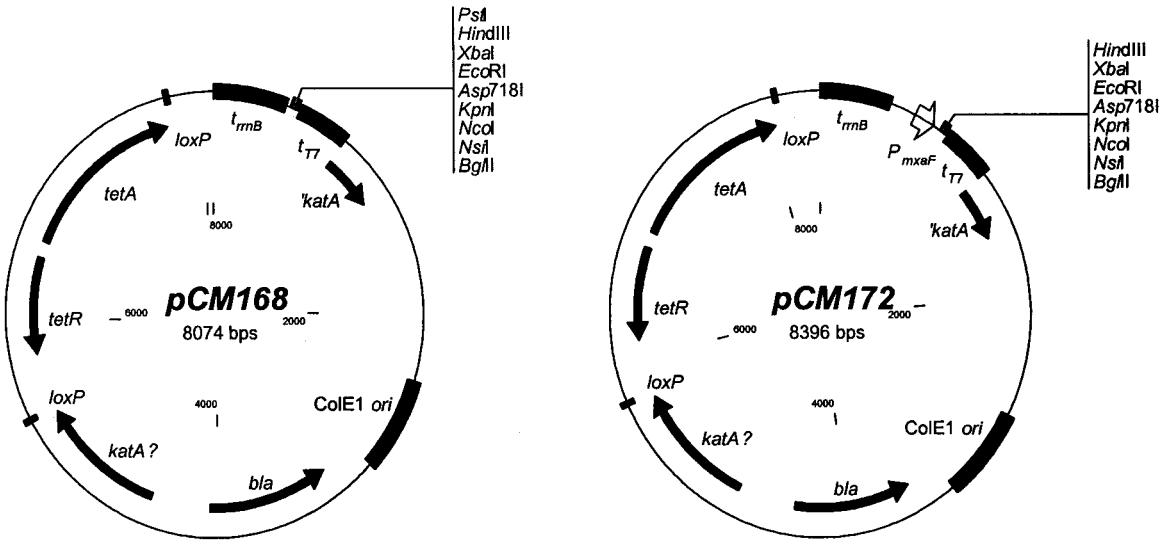


Figure 2.7: Plasmid maps of insertional vectors pCM168 and pCM172. GenBank accession numbers for these plasmids are AY307999 and AY308000, respectively.

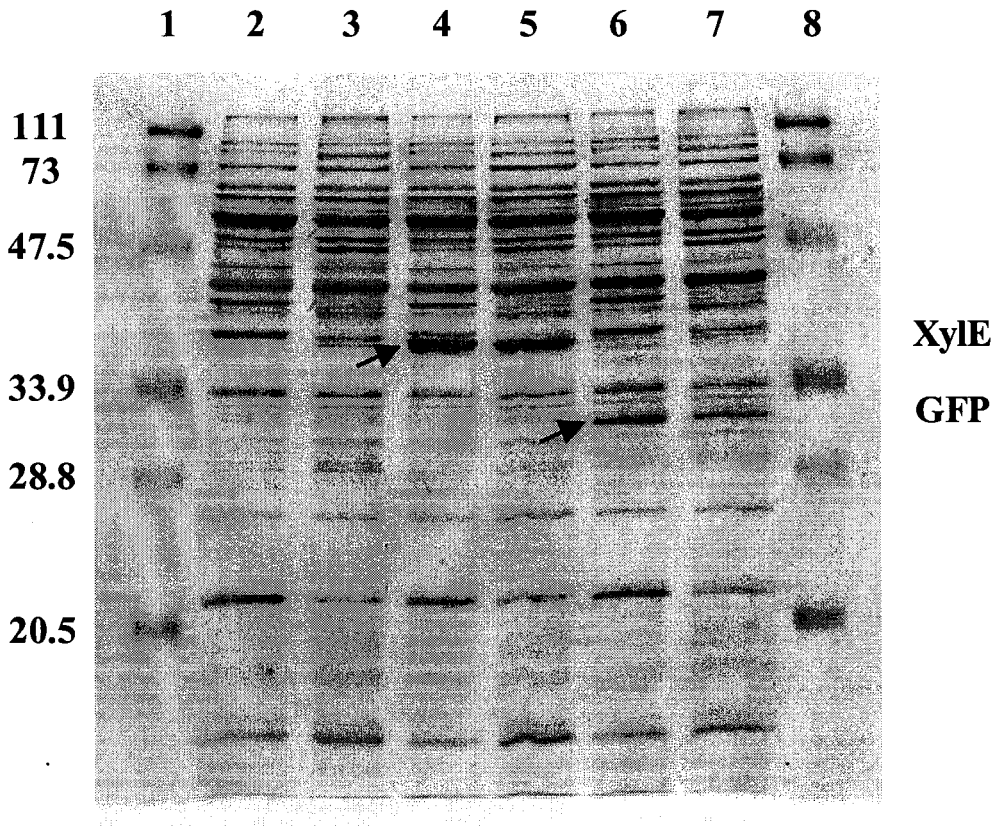


Figure 2.8: SDS-PAGE of cell extracts prepared from *M. extorquens* AM1 bearing various plasmids. Lanes 1 and 8, molecular weight markers, masses in kDa indicated to the left (Bio-Rad, Hercules, CA); lanes 2 and 3, pCM80; lanes 4 and 5, pCM81; lanes 6 and 7, pCM88. Lanes 2, 4, and 6 were loaded with extracts prepared from methanol-grown cultures, lanes 3, 5, and 7 contain extracts from succinate grown cells. Arrows indicate the position of XylE and GFP (predicted masses of 35.1 and 26.8 kDa, respectively).

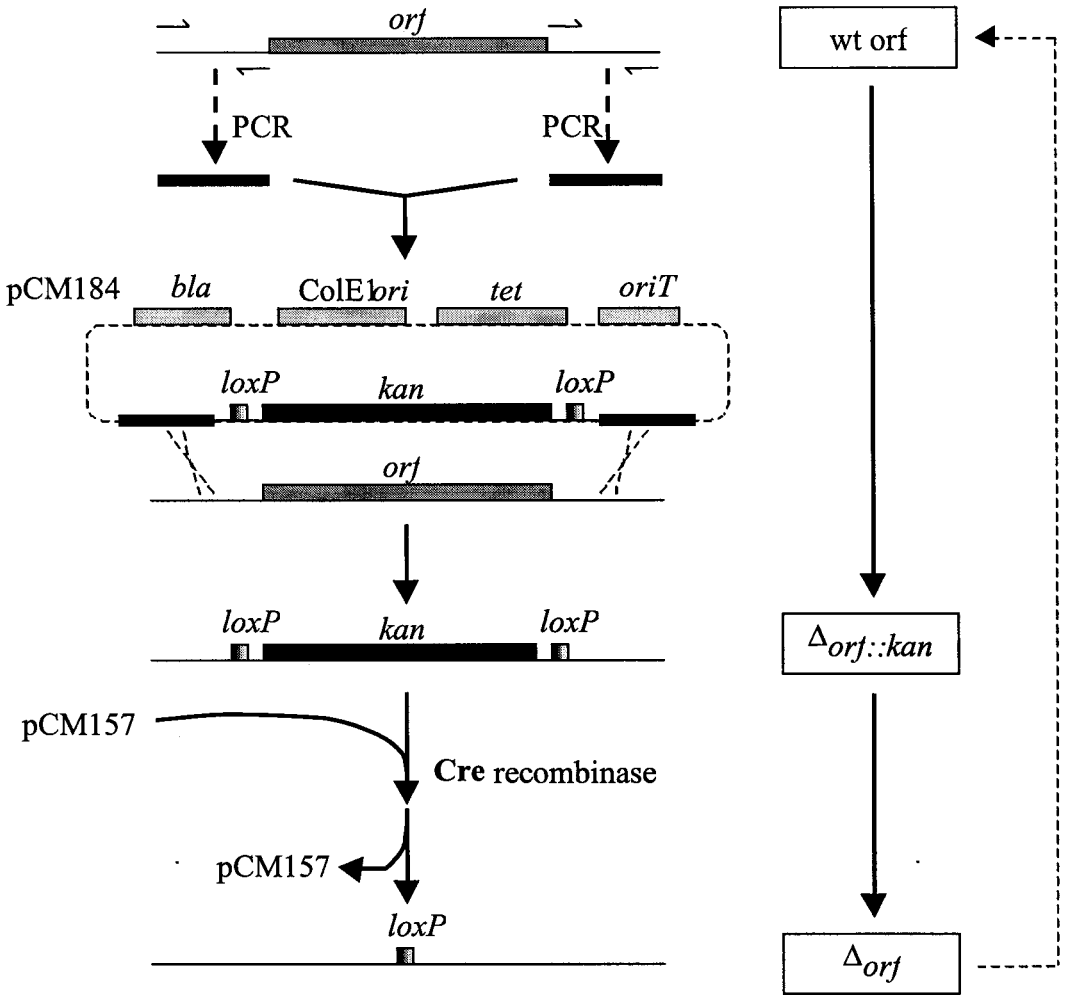


Figure 2.9: Strategy for antibiotic marker recycling. DNA flanks upstream and downstream of the target gene are amplified by PCR and cloned into pCM184. Allelic exchange leads to a *kan* insertion mutant, which can then be unmarked through the introduction of the *cre* expression plasmid pCM157. The process can then be repeated with a second target gene to generate a strain bearing multiple genetic manipulations.

CHAPTER 3

Genetic and Physiological Analysis of the Role of the Tetrahydromethanopterin-linked Formaldehyde Oxidation Pathway in *Methylobacterium extorquens* AM1

Abstract:

The facultative methylotroph *Methylobacterium extorquens* AM1 possesses two pterin-dependent pathways for C_1 transfer between formaldehyde and formate, the tetrahydrofolate (H_4F)-linked pathway and the tetrahydromethanopterin (H_4MPT)-linked pathway. Both pathways are required for growth on C_1 substrates, however, mutants defective for the H_4MPT pathway reveal a unique phenotype of being inhibited by methanol during growth on multi-carbon compounds such as succinate. It has been previously proposed that this methanol-sensitive phenotype is due to the inability to effectively detoxify formaldehyde produced from methanol. Here I present mutant analysis of two new genes of the H_4MPT pathway, *fae* (encodes formaldehyde-activating enzyme) and *dmrA* (encodes putative dihydromethanopterin reductase), which suggests a role for these gene products in formaldehyde oxidation and detoxification. A comparative physiological characterization of these and other deletion mutants defective in the H_4MPT pathway places them into three different phenotypic classes that are concordant with the biochemical roles of the respective enzymes. I demonstrate that the analogous H_4F pathway present in *M. extorquens* AM1 cannot fulfill the formaldehyde detoxification function, while a heterologously expressed pathway linked to glutathione and NAD can successfully substitute for the H_4MPT pathway. Additionally, null mutants were generated in genes previously thought to be essential, indicating that the H_4MPT pathway is not absolutely required during growth on multi-carbon compounds. Combined, these data clarify the role of the H_4MPT pathway as the primary formaldehyde oxidation and detoxification pathway in *M. extorquens* AM1.

Introduction:

Subsequent to the initial discovery of the H_4MPT-dependent formaldehyde oxidation pathway in *M. extorquens* AM1 (20) and other bacteria (126), attention has turned to a thorough analysis of the enzymes involved, the biosynthesis of the H_4MPT cofactor, and the role of the

H_4MPT pathway in methylotrophic metabolism. My project has contributed to these efforts by the discovery of two novel genes of the H_4MPT pathway, *fae* and *dmrA*, and through a genetic and physiological analysis of these and other H_4MPT pathway mutants. My work described in this chapter has addressed two key questions regarding the H_4MPT pathway: 1.) What is the role of the pathway in formaldehyde detoxification? and 2.) Why is this pathway apparently essential for growth on multi-carbon compounds?

The first part of this work involved the search for previously unidentified genes involved in the pathway. One of these was hypothesized to exist due to the key role of formaldehyde during methylotrophic growth. The condensation of formaldehyde with either H_4F (55) or H_4MPT (34) has been demonstrated to occur nonenzymatically. The question arose, however, whether this would occur rapidly enough to accommodate the high rate of formaldehyde production in methylotrophy. Although no formaldehyde-H_4F condensing activity was found in cell-free extracts, an enzyme was purified that catalyzed the reaction of formaldehyde with H_4MPT (127). This enzyme, Fae (formaldehyde-activating enzyme), was purified and the N-terminal amino acid sequence was determined. The sequence and molecular mass matched that predicted for the gene product of *orf18* (20), which was thus renamed *fae*. In this chapter I describe my role in that project, which was to generate a *fae* mutant strain and determine the role of this gene in methylotrophy by examining the mutant phenotype.

Another approach for identification of genes in the pathway was the random transposon mutagenesis screen described in Chapter 2. As noted in that chapter, we isolated an IS*phoA*/hah-Tc strain, S64-67 (80), that had the same phenotype as the *fae* (127)and *mtdB* (43) mutants. Namely, this mutant failed to grow on C_1 compounds and was sensitive to methanol during growth on multi-carbon compounds. The mutated gene was identified as a previously uncharacterized *orf*, now designated as *dmrA* (putative dihydromethanopterin reductase) unlinked to all known methylotrophy genes. This gene encodes a protein that has 26% identity and 44%

similarity to the dihydrofolate reductase (DHFR) of *Lactobacillus casei* (P00381) (Fig. 3.1). In this chapter I present a genetic analysis of *dmrA*.

The next part of this work involved the role of the H_4MPT pathway, with specific emphasis on the methanol-sensitive phenotype. The methanol-sensitive mutant phenotype is thus far unique to the mutants of *M. extorquens* AM1 defective for the H_4MPT pathway and has been proposed to be due to the inability to detoxify the formaldehyde produced from methanol. Double mutants lacking both MDH activity and MtdB activity were no longer sensitive to methanol (43), lending further support to the concept that methanol-sensitivity is a proxy for formaldehyde detoxification deficiency. Null mutants lacking Orf4, a homolog of the first enzyme in the H_4MPT biosynthesis pathway, β-RFAP synthase (106), and Orf7 (unknown function) have also been generated and these were incapable of growth on methanol (20). In this chapter I present a comparative physiological analysis of four H_4MPT pathway mutants, uncovering three mutant classes that are consistent with their roles in the pathway and providing further evidence that the methanol-sensitive mutant phenotype is due to defective formaldehyde detoxification.

The final part of this work involved the apparent essentiality of the pathway, even during growth on multi-carbon compounds. Efforts to obtain null mutants have been unsuccessful for some known genes of this pathway, including *mtdA* (22), *mch*, and *fhcBADC*, as well as two orfs of unknown function proximal to the H_4MPT gene cluster, *orf9* and *orf17* (20). Mutants resulting from an incomplete allelic exchange event that maintained a wild-type copy of the gene separated from its native promoter by the integrated vector were obtained for these genes and, where examined, this led to reduced enzymatic activity (20, 22). In all cases, this class of mutants exhibited defective growth on methanol, indicating a specific role in methylotrophy in addition to apparent essentiality. MtdA activity is likely required to produce formyl-H_4F for biosynthetic needs (22). It has not been clear, however, why the other genes for which null mutants could not be obtained and have a known or predicted role in the H_4MPT pathway for formaldehyde

oxidation would be required for growth on a multi-carbon substrate such as succinate. In this chapter I demonstrate the generation of null mutants lacking these genes, thus rejecting the notion that the H_4MPT pathway is essential. Collectively, these genetic and physiological analyses have led to the discovery of new genes involved in the H_4MPT pathway of *M. extorquens* AM1 and have demonstrated that the H_4MPT pathway functions as the primary formaldehyde oxidation and detoxification route during methylotrophic growth.

Materials and Methods:

Standard laboratory procedures. The growth of bacterial strains, genetic procedures, and recombinant DNA techniques are performed as described in Chapter 2. *M. extorquens* AM1 strains, plasmids, and primers used in this chapter are described in Tables 3.1, 3.2, and 3.3.

Generation of a *fae* mutant and complementing plasmid. In order to generate a *fae::kan* mutant, a 2.0 kb chromosomal region of *M. extorquens* AM1 containing *fae* was amplified using the primers CM-faef and CM-faer. The resulting PCR product was cloned into pCR2.1 (Invitrogen) to yield pCM112. The 2.1 kb *Bam*HI-*Sph*I region of pCM112 was subcloned into pUC19 (134) to generate pCM113. The *kan* cassette from pUC4K (123) was then inserted between the two *Hin*cII sites found in *fae* to generate pCM114. Finally, the 3.1 kb *Bam*HI-*Sph*I fragment from pCM114 was excised, blunted, and inserted into the *Sma*I site of pAYC61 (17) to yield pCM115. Subsequently, pCM115 was introduced into *M. extorquens* AM1 using the helper strain S17-1 (107), and a *fae::kan* exconjugant was isolated, CM115.1.

A plasmid containing *fae* under the expression of its own putative promoter was generated to complement CM115.1. A 0.8 kb region containing *fae* and its putative promoter region was amplified by PCR using CM-Pfaef and CM-4010r and cloned into pCR2.1

(Invitrogen) to yield pCM138. This region was excised with *Bam*HI and *Sph*I and inserted into pCM62 to yield pCM139.

Construction of plasmids to generate deletion mutants of *fae*, *dmrA*, *orf4*, *mtdB*, *mch*, and *fhcBADC*. *M. extorquens* AM1 deletion mutants lacking *fae*, *dmrA*, *orf4*, *mtdB*, *mch*, or the *fhcBADC* cluster were generated using pCM184. Approximately 0.5 kb regions upstream and downstream of these genes or gene clusters were amplified by PCR and cloned into pCR2.1 (Invitrogen) as follows. The *fae* flanks were amplified with CM-Dfae1 and CM-Dfae2, and CM-Dfae3 and CM-Dfae4, and were cloned into pCR2.1 (Invitrogen) to generate pCM195 and pCM196. The *dmrA* flanks were amplified with CM-dmrAuf and CM-dmrAur, and CM-dmrAdf and CM-dmrAdr, and resulted in pCM207 and pCM208. The *orf4* flanks were amplified with CM-orf4uf and CM-orf4ur, and CM-orf4df and CM-orf4dr, and resulted in pCM250 and pCM251. The *mtdB* flanks were amplified with CM-mtdBuf and CM-mtdBur, and CM-mtdBdf and CM-mtdBdr, and resulted in pCM255 and pCM256. The *mch* flanks were amplified with CM-mchuf and CM-mchur, and CM-mchdf and CM-mchdr, and resulted in pCM260 and pCM261. The *fhcC* downstream flank was amplified with CM-fhcCdf and CM-fhcCdr and resulted in pCM264.

The construct to generate Δ*fae::kan* mutants was generated by introducing the 0.6 kb *Eco*RI-*Not*I fragment from pCM195 between the corresponding sites of pCM184 to produce pCM197, and subsequently, the 0.6 kb *Apa*I-*Sac*I fragment from pCM196 was ligated between the same sites of pCM197 to produce pCM198. The construct to generate Δ*orf4::kan* mutants was generated by introducing the 0.5 kb *Eco*RI-*Asp*718I fragment from pCM250 into the same sites of pCM184 to produce pCM252, and subsequently, the 0.7 kb *Apa*I-*Sac*I fragment from pCM251 was ligated into the same sites of pCM252 to produce pCM253. The construct to generate Δ*mtdB::kan* mutants was generated by introducing the 0.6 kb *Sac*II-*Sac*I fragment from pCM256

into the same sites of pCM184 to produce pCM257, and subsequently, the 0.5 kb *Aat*II-*Asp*718I fragment from pCM255 was ligated into the same sites of pCM257 to produce pCM258. The construct to generate Δ*mch::kan* mutants was generated by introducing the 0.6 kb *Aat*II-*Asp*718I fragment from pCM260 into the same sites of pCM184 to produce pCM262, and subsequently, the 0.7 kb *Apa*I-*Sac*I fragment from pCM261 was ligated into the same sites of pCM262 to produce pCM263. Finally, the construct to generate Δ*fhcBADC::kan* mutants was generated by introducing the 0.5 kb *Eco*RI-*Nco*I fragment from pCM250 into the same sites of pCM184 to produce pCM265, and subsequently, the 0.5 kb *Sac*II-*Age*I fragment from pCM264 was ligated into the same sites of pCM265 to produce pCM266. The generation of marked and unmarked mutant strains was performed as described in Chapter 2.

Generation of plasmids expressing *dmrA* or *mtdA*. Plasmids expressing *dmrA* and *mtdA* were generated as follows. The coding regions of *dmrA* and *mtdA* were amplified by PCR using CM-dmrAf1 and CM-dmrAr, and CM-mtdAf and CM-mtdAr. The resulting PCR products were cloned into pCR2.1 (Invitrogen) to produce pCM185 and pCM254. The 0.4 kb *Hin*dIII-*Bgl*II fragment of pCM185 was cloned between the *Hin*dIII and *Bam*HI sites of pCM62 to produce pCM186. The 1.0 kb *Hin*dIII-*Xba*I fragment of pCM254 was cloned between the same sites of the expression plasmid pCM80 to generate pCM259.

Construct for the heterologous expression of the GSH-dependent formaldehyde oxidation pathway from *P. denitrificans*. The two primary genes comprising the GSH-dependent formaldehyde oxidation pathway of *P. denitrificans*, *flhA* (102) and *fghA* (45), were amplified by PCR using pWRxox451 (101) as a template and cloned into pCR2.1 (Invitrogen) to produce pCM102 and pCM103. The 1.0 kb *Eco*RI fragment of pCM103 was introduced into the

*Eco*RI site of pCM80 to generate pCM104, into which the 1.4 kb *Xba*I fragment from pCM102 was inserted into the corresponding site to produce pCM106.

Sequencing of the *mtdA* coding region and P_{sgaA}. The coding region of *mtdA* and the promoter that expresses the operon containing *mtdA* (P_{sgaA}) (56) were amplified by PCR from MR1, MR2, MR3, and MR4 using the primers mtdAfor and mtdArev, and PsgaAf and PsgaAr. The resulting PCR products were purified using QIAquick PCR Purification Kit (Qiagen) and sequenced (University of Washington Biochemistry Department DNA Sequencing Facility).

Phenotypic analyses of mutant strains. In order to compare the growth of wild-type *M. extorquens* AM1 with mutants in liquid medium, cultures were grown to mid-exponential phase, centrifuged, and then resuspended into fresh medium containing the carbon source described. To test for sensitivity to methanol, methanol was added to one set of succinate flasks to the reported final concentration after two hours. Mutant phenotypes were also assessed on solid medium by comparing the relative rate of colony formation. Sensitivity to methanol or formaldehyde was assayed using succinate medium to which methanol or formaldehyde was added immediately before pouring plates at the following tested concentrations: 125, 10, 1 mM, and 100, 10, 1, and 0.1 μM for methanol, and 1, 0.5, 0.1, 0.05, 0.01, 0.005 mM for formaldehyde. Because an undetermined fraction of the methanol will volatilize, the reported MIC is a maximum value. All phenotypic analyses were performed at least twice.

Enzyme assays. The activities of MtdA (125), FlhA (102), and FghA (45) were assayed in 2-3 replicates as described using cell extracts prepared using a French press from cell material harvested from exponential phase cultures. Variation in enzyme activities between cultures was less than 20%. The presence of MDH was determined using activity gels as described (129).

Total protein concentration in the extracts was assayed spectrophotometrically (54, 130) using a Beckmann DU 640B spectrophotometer.

Results:

Identification and mutant analysis of new genes encoding functions of the H_4MPT pathway.

Phenotypic analysis of a fae *mutant and complementation.*

***The work described in this section has been published as part of the paper describing the purification and characterization of Fae (127).**

In order to assess the physiological importance of the formaldehyde-H_4MPT condensation activity of Fae, a null mutant strain was generated by allelic exchange. The *fae::kan* mutant strain CM115.1 lacked detectable Fae activity and was unable to grow on methanol but exhibited wild-type growth on succinate. Furthermore, growth of the *fae::kan* mutant strain on succinate plates was inhibited by the addition of methanol or formaldehyde at MICs of 10-50 and 100-200 μM, respectively. These results suggested that the Fae-H_4MPT-dependent pathway is the primary formaldehyde detoxification system during growth on both C_1 and multi-carbon compounds. This suggestion is examined in greater detail below. Additionally, introduction of pCM139, a plasmid bearing *fae* with its putative promoter region, restored wild-type growth to CM115.1, thus confirming that the phenotype is not due to a polar effect on downstream genes.

*Phenotypic analysis of dmrA::IS*phoA*/hah-Tc mutant and complementation.*

***The work described in this and the next section has been published, as part of the paper on the transposon mutagenesis study described in Chapter 2 (80).**

As reported in Chapter 2, a C_1^-/methanol-sensitive mutant was isolated in an *orf* with homology to dihydrofolate reductases (DHFR) that was named *dmrA* (putative dihydromethanopterin reductase, see below). The inhibitory effect of methanol and formaldehyde

was determined to be 1 and 10 μM during growth on succinate plates. This suggests that the *dmrA* mutant has an even more severe defect in formaldehyde oxidation than the *fae* mutant described above. In order to confirm that this defect was solely due to the transposon insertion into *dmrA*, the plasmid pCM186 that contains *dmrA* behind P_{lac} was introduced into S64-67. This led to full complementation of the mutant phenotype, thereby confirming that *dmrA* is required for methylotrophy.

Suggested role for DmrA in methylotrophy. If *dmrA* encoded DHFR, interruption by mutation should be lethal due to the central role of DHFR in biosynthetic reactions involving C_1 units. However, the genome sequence contains three orfs with similarity to DHFR. One of these orfs, here designated as *dfrA*, has a predicted amino acid sequence significantly more similar to other DHFRs than the other two (Fig. 3.1). The DfrA sequence (GenBank accession number AY093432) has 50% identity and 63% similarity to the DHFR present in the closely related α-proteobacterium *Mesorhizobium loti* (BAB48827). Additionally, *dfrA* is located directly downstream of an apparent thymidylate synthase, a gene organization that is conserved across a wide variety of bacteria. These sequence analyses suggest that *dfrA* encodes the general DHFR activity.

Antibiotics such as trimethoprim and methotrexate inhibit bacteria through the inhibition of DHFR activity. Resistance to trimethoprim is most often accomplished by expression of an alternate antibiotic-resistant DHFR (108). Because *Methylobacterium* strains generally exhibit resistance to this class of antibiotics (49), I determined whether *dmrA* was involved. Wild-type *M. extorquens* AM1 and S64-67 were grown on plates containing up to 400 μg/mL trimethoprim. No differences in growth were observed, indicating that the *dmrA* gene product is not required for trimethoprim resistance. The third *orf* with identity to DHFR, here designated as *dfrB*, has a predicted amino acid sequence (GenBank accession number AY093433) with 25% identity and

47% similarity to the trimethoprim-resistant DHFR found on the *E. coli* plasmid pCJ001 (51) (Fig. 1). Mutational and biochemical analyses will be required to assess the role of *dfrA* and *dfrB* in *M. extorquens* AM1.

The *dmrA* mutants have a C_1^- defective/methanol-sensitive phenotype that has been associated with defects in H_4MPT-dependent formaldehyde oxidation (43, 127). Given the similarity of the predicted DmrA protein sequence to DHFR, I suggest that *dmrA* encodes dihydromethanopterin reductase. If so, this would represent a novel enzymatic activity analogous to DHFR, but functioning with H_4MPT rather than H_4F:

$$H_2MPT + NADPH \rightleftharpoons H_4MPT + NADP^+$$

If this is the function of DmrA, it would explain the severe methanol sensitivity of the mutant, since the entire H_4MPT pathway would be inoperable with no H_4MPT in the cell. Consistent with this hypothesis, an apparent *dmrA* ortholog was found in the genome sequence of *Methylococcus capsulatus* Bath (www.jgi.doe.gov/JGI_microbial/html/index.html), a methanotrophic bacterium that contains the H_4MPT pathway, but not in any other available genome sequences (L. Chistoserdova, unpublished). Furthermore, it has recently been determined by our collaborator, Dr. M. E. Rasche (U. Florida), that *M. extorquens* AM1 *dmrA* mutants fail to synthesize H_4MPT and mutants lack an apparent dihydromethanopterin reductase activity found in the wild-type (S. Wyles, M. Caccamo, and M. E. Rasche, personal communication). Detailed biochemical characterization will be required to confirm our hypothesis that *dmrA* encodes a dihydromethanopterin reductase.

Mutants defective for the H_4MPT pathway have varying degrees of sensitivity to methanol and formaldehyde. *M. extorquens* AM1 mutants defective for *mtdB* (43), *fae*, and *dmrA* of the H_4MPT pathway are all unable to grow on C_1 compounds as their sole source of carbon and energy and exhibit sensitivity to methanol or formaldehyde during growth on multi-

carbon compounds such as succinate. It has been hypothesized that this unique phenotype is due to an inability to detoxify formaldehyde (43, 80, 127). A prediction of this hypothesis is that formaldehyde or a toxic derivative should accumulate to toxic levels in the cell when such mutants are exposed to methanol or formaldehyde. For the case of formaldehyde accumulation, currently it is not possible technically to measure formaldehyde inside the cell. Formaldehyde is produced in the periplasm and enters the cytoplasm, where the consumption reactions occur. Formaldehyde outside the cell is not as toxic as methanol, and it is clear that the important compartment for formaldehyde toxicity is the cytoplasm. Although other groups have measured formaldehyde accumulating in methylotrophic cultures (11, 128), this represents formaldehyde leaking out of the periplasm and the relationship between that level and the formaldehyde inside the cell is uncertain. In addition, the proteins in the cell represent a sink for formaldehyde, and so as formaldehyde begins to rise inside the cell, a large reservoir consumes it, albeit with damage to the cell. Therefore, my approach to this problem in this section has been to use physiological and genetic approaches to infer the role of the H_4MPT pathway with regard to formaldehyde detoxification. Our laboratory is currently collaborating with a colleague at Los Alamos National Laboratories (Dr. C. Unkefer) to develop novel NMR- and MS-based methods to distinguish between cytoplasmic and periplasmic formaldehyde, but that effort will take some years to carry out. In addition, it is possible that the actual toxicity is mediated by accumulation of an intermediate downstream of formaldehyde in metabolism, for instance, a H_4MPT pathway intermediate. There are currently no methods sufficiently sensitive to measure these unstable intermediates either in whole cells or cell extracts, so once again, my approach is to use genetic methods to address this issue.

In order to test this hypothesis and to better understand the role of the H_4MPT pathway in *M. extorquens* AM1, the phenotypes of these mutants were examined in more detail. For this work, mutants were employed in which the genes in question were deleted from the chromosome

via allelic exchange. Some of these deletions were unmarked using the *cre-lox* system, and the resulting strains were used to construct new strains bearing more than one mutation. New mutants were generated for *fae* (CM198K.1, Δ*fae::kan*), *dmrA* (CM212K.1, Δ*dmrA::kan*), *orf4* (CM253K.1, Δ*orf4::kan*) and *mtdB* (CM258K.1, Δ*mtdB::kan*).

All four of these H_4MPT pathway mutants grew with wild-type characteristics in liquid medium containing succinate, indicating that the respective functions are not required for general heterotrophic growth (Fig. 3.2), but were unable to grow in medium containing methanol (Fig.3.2). Analogous results were observed for growth on solid medium. Additionally, the H_4MPT pathway mutant strains grew like wild-type rates formate, indicating that they are not required for the metabolism of the more oxidized C_1 compound formate. Interestingly, trace growth of H_4MPT pathway mutants that was significantly slower than wild-type was observed on methylamine plates. In this case, the mutants fell into three classes; CM258K.1 grew better than CM198K.1, which itself grew better than either CM212K.1 or CM253K.1.

In order to compare the inhibitory effect of methanol on the growth of mutants defective for the H_4MPT pathway on succinate, methanol was added to a set of succinate flasks after two hours to either 1 or 125 mM final concentration (Fig. 3.2). Under these conditions, the *mtdB* mutant CM258K.1 grew like wild-type. However, the three other mutants were inhibited at both methanol concentrations. Addition of methanol at 1 mM caused a more severe inhibition of the *dmrA* and *orf4* mutant strains relative to the *fae* mutant strain, whereas 125 mM methanol caused cessation of growth in all three strains. Preliminary results suggested that over 24 hours, however, no significant loss of viability occurred (C. J. Marx, unpublished).

These four H_4MPT pathway mutants were again observed by examining the minimal inhibitory concentrations (MICs) of methanol or formaldehyde during growth on succinate plates. On solid medium, a distinct inhibitory effect of methanol was observed for the *mtdB* mutant CM258K.1 with an MIC of 10 mM, and a MIC of formaldehyde of 0.5 mM. In comparison, wild-

type had an MIC of formaldehyde of 1 mM and was not inhibited by 125 mM methanol. The other mutants were observed to be significantly more sensitive, with MICs for methanol or formaldehyde of 10 μM and 100 μM for the *fae* mutant CM198K.1, and 1 μM and 10 μM for the *dmrA* and *orf4* mutants CM212K.1 and CM253K.1.

Overexpression of *mtdA* provides partial complementation of the *mtdB* mutant phenotype. I hypothesized that the relatively moderate sensitivity of the *mtdB* mutant strain CM258K.1 to methanol may be due to the presence of another enzyme, MtdA, whose substrate specificity overlaps with that of MtdB. Even though the presence of MtdA is insufficient for wild-type resistance to methanol, it may contribute to the removal of formaldehyde by converting methylene-H_4MPT to methenyl-H_4MPT. To test this hypothesis, the region encoding *mtdA* was cloned and introduced into the expression vector pCM80 to allow for overexpressed levels of MtdA. The plasmid containing *mtdA* expressed from the strong promoter P_{mxaF} resulted in nearly an order of magnitude increase in MtdA activity from 270 to 1970 mU during growth on methanol. Neither CM258K.1 bearing the empty vector pCM80 or pCM259 was capable of growth on methanol plates, however, there was both a moderate increase in the ability to grow on methylamine and the MIC for methanol was 125 mM for CM258K.1 with pCM259, compared to 10 mM for CM258K.1 with pCM80. Therefore, a substantial increase in MtdA activity only provides partial complementation of the *mtdB* mutant phenotype.

Spontaneous methanol-resistant mutants are easily obtained in the *mtdB* mutant background. While examining the methanol-sensitivity of the *mtdB* mutant CM258K.1 I noted that mutants with significant resistance to methanol arose spontaneously, with a frequency of about 10^{-5}, while no spontaneous methanol-resistant mutants were observed in the *fae*, *dmrA*, or *orf4* backgrounds. To investigate this phenomenon, four of the spontaneous methanol-resistant

mtdB mutants, named MR1 through MR4, were selected for further study. MR2, MR3 and MR4 exhibited wild-type resistance to methanol, whereas for MR1 the MIC for methanol was 125 mM compared to 10 mM for CM258K. I hypothesized that these mutants either 1.) Contain mutations that increase MtdA expression and/or alter its activity, or 2.) Block the ability of the cell to oxidize methanol, or 3.) Activate a previously silent (and unknown) formaldehyde oxidation pathway. To test the first hypothesis, the coding sequence of *mtdA* and the promoter that drives its expression, P_{sgaA}, (56) from each of the four methanol-resistant *mtdB* mutants were amplified by PCR and sequenced. No mutations were detected, however, in any of the four strains in either the *mtdA* coding region or the promoter responsible for its expression, P_{sgaA}. To test the second hypothesis, MDH activity was assessed in cell extracts of wild-type, CM258K.1, and the methanol-resistant strains MR1-MR4 grown on succinate, via specific activity staining in isoelectrofocusing gels. MDH activity was detected in wild-type, CM258K.1 and MR1, while the mutants MR2, MR3 and MR4 lacked activity. In order to further investigate this phenomenon, a plasmid bearing *mtdB* under its own apparent promoter (pLC310BgBm) (20) was introduced into CM258K.1 and the four methanol-resistant mutants. CM258K.1 bearing pLC310BgBm recovered the ability to grow on both methanol and methylamine like wild-type. MR2, MR3 and MR4 with pLC310BgBm recovered the ability to grow on methylamine, but remained incapable of growth on methanol despite the presence of a wild-type *mtdB* allele. These data are consistent with the hypothesis that these three strains bear a mutation in one of the 25 genes required for MDH activity. The MR1 mutant which exhibited MDH activity *in vitro* was neither complemented for growth on methanol or methylamine when carrying pLC310BgBm, a phenotype previously shown for mutants defective in the synthesis of cytochrome *c* (91), suggesting that it may have acquired a defect in this system. None of the mutants appeared to have activated an unknown formaldehyde detoxification system.

Methanol-sensitivity of H_4MPT pathway mutants requires formaldehyde production. Previous work (43) on *mtdB* mutants indicated that the sensitivity to methanol could be alleviated if a *mtdB::kan* mutant was generated in a strain that contained a mutation in the gene (*mxaF*) encoding the large subunit of MDH (89). This demonstrated that the sensitivity of this strain required the production of formaldehyde, and was not simply a consequence of methanol itself. I extended this analysis to characterize the other three mutants with greater sensitivity to methanol. A series of double mutants were constructed in the Δ*mxaF* strain CM194.1. The resulting strains were resistant to the addition of 125 mM methanol to succinate cultures, with only the Δ*mxaF*, Δ*dmrA::kan* and Δ*mxaF*, Δ*orf4::kan* mutants showing even a slight growth inhibition (Fig. 3.3). Similar results were obtained using solid media. The Δ*mxaF*, Δ*mtdB::kan* strain CM194-258K.1 was not inhibited by 125 mM, compared to a MIC of 10 mM for the Δ*mtdB::kan* strain CM258K.1. The Δ*mxaF*, Δ*fae::kan* strain CM194-198K.1 had a MIC for methanol of 10 mM, compared to 10 μM for the Δ*fae::kan* strain CM198K.1. Finally, the Δ*mxaF*, Δ*dmrA::kan* strain CM194-212K.1 and the Δ*mxaF*, Δ*orf4::kan* strain CM194-253K.1 exhibited an MIC for methanol of 100 μM compared to 1 μM for the corresponding MDH^+ strains. This residual sensitivity to methanol suggests either a low-level alternate methanol oxidation activity or a direct effect of methanol at higher levels. However, these data show that the extreme sensitivity to methanol observed in all tested H_4MPT pathway mutants is not due to methanol itself, but rather, requires the production of formaldehyde.

Methanol-sensitivity of H_4MPT pathway mutants is alleviated and growth on C_1 compounds is achieved by expressing a heterologous GSH-dependent formaldehyde oxidation pathway. The data presented thus far suggest that the methanol-sensitive phenotype observed for H_4MPT pathway mutants is due to an inability to detoxify the formaldehyde

produced from methanol, either directly or as one of its derivatives. As a final test of this hypothesis, a heterologous formaldehyde oxidation system was cloned and expressed in H_4MPT pathway mutants. The two primary genes of the GSH-dependent formaldehyde oxidation pathway of *Paracoccus denitrificans*, *flhA* (encodes GSH- and NAD-dependent formaldehyde dehydrogenase, (102)) and *fghA* (encodes *S*-formyl-GSH hydrolase, (45)) were cloned by PCR amplification, and introduced together into the expression vector pCM80 to generate the plasmid pCM106. Introduction of pCM106 resulted in activities of 2500 and 2300 mU for FlhA and FghA, respectively, whereas these activities were undetectable in wild-type *M. extorquens* AM1 carrying pCM80 without an insert. All mutants bearing pCM106 were resistant to 125 mM methanol added to succinate growth medium (Fig.3.4). Similarly, the mutants bearing pCM106 were insensitive to 125 mM methanol present in succinate plates. Additionally, the presence of pCM106 allowed the *mtdB* mutant to grow like wild-type in the presence of 1 mM formaldehyde, and raised the MIC of formaldehyde to 0.5 mM for *fae*, *dmrA* and *orf4* mutant strains. Finally, beyond alleviating methanol-sensitivity the expression of the GSH pathway in the H_4MPT pathway mutants allowed growth in methanol liquid medium (Fig. 3.4), and on methanol or methylamine plates, albeit the complementation of the *dmrA* and *orf4* mutants was less robust than for the other two mutants. The ability of the heterologous GSH-dependent formaldehyde oxidation system to alleviate the methanol sensitivity of H_4MPT pathway mutants provides strong evidence that the cause of the methanol-sensitive phenotype is the inability to detoxify intracellular formaldehyde produced from methanol.

Null mutants lacking *mch* or *fhcBADC* can only be obtained in a H_4MPT biosynthesis-negative background. One important question regarding the role of the H_4MPT pathway in *M. extorquens* AM1 is why it has not been possible to obtain null mutations in genes that encode the two enzymes catalyzing the final reactions of the pathway, Mch or Fhc (20). A

few scenarios may be suggested to explain this phenomenon: 1.) A C_1-H4MPT intermediate is required for growth on multi-carbon compounds, 2.) These mutants are even more sensitive to methanol, such that ambient concentrations are lethal, or 3.) Accumulation of an intermediate of the H_4MPT pathway causes a toxic effect or a regulatory problem. To test these hypotheses, I attempted constructing deletion versions of these mutations in various backgrounds. No mutants were obtained in wild-type using the deletion constructs for Mch (Δ*mch::kan*) or FhcBADC (Δ*fhcBADC::kan*), in agreement with previous results for insertion mutants (20). Likewise, no deletion mutants were obtained in the backgrounds lacking MDH (Δ*mxaF* strain CM194.1), Fae (Δ*fae* strain CM198.1), or MtdB (Δ*mtdB* strain, CM258.1). Both deletions were readily generated, however, in the Δ*orf4* strain CM253.1. The Δ*orf4*, Δ*mch::kan* strain CM253-263K.1 and a Δ*orf4*, Δ*fhcBADC::kan* strain CM253-266K.1 grew normally on succinate and formate, but exhibited defective growth on methanol and methylamine as had been observed for the *orf4* mutant CM253K.1. Furthermore, the MICs for methanol and formaldehyde during growth on succinate were the same as those observed for CM253K.1. These data suggest that, although none of the H_4MPT pathway enzymes are essential for metabolism of multi-carbon compounds, accumulation of an intermediate(s) of the pathway somehow affects multi-carbon metabolism.

Discussion:

The genetic analysis presented for *fae* and *dmrA* suggested that the novel functions encoded by these novel genes are critical to methylotrophic growth and detoxification of formaldehyde during growth on multi-carbon compounds. First, this suggests that the nonenzymatic condensation of formaldehyde and H_4MPT is insufficient to accommodate high formaldehyde flux. Second, the data presented for *dmrA* are inconsistent with a role as either the housekeeping DHFR, or as an alternative DHFR conferring trimethoprim resistance. Rather, based on the dmrA mutant phenotype and sequence similarity to DHFR I have presented the

hypothesis that *dmrA* encodes dihydromethanopterin reductase, the enzyme responsible for catalyzing the final step in H_4MPT biosynthesis. Consistent with this suggestion, *dmrA* mutants have been found to lack detectable H_4MPT (S. Wyles and M. E. Rasche, personal communication). Thus far, no dihydromethanopterin reductases have been characterized from either methanogenic archaea or other methylotrophic bacteria. Preliminary work indicates that DmrA can perform H_2MPT reductase activity using NADPH (M. Caccamo and M. E. Rasche). The discovery of these two genes involved in the H_4MPT pathway has been a critical step toward understanding how this pathway functions in methylotrophic bacteria.

In order to gain further insights into formaldehyde oxidation and detoxification in *M. extorquens* AM1, I have conducted a physiological analysis of mutants defective for the H_4MPT-linked pathway. These mutants fall into three phenotypic classes that correlate with the biochemical roles of the respective enzymes in the pathway. The most severe defect is found for mutants defective for *dmrA* and *orf4*. Both mutants are predicted to lack H_4MPT (106), and would thus lack both the H_4MPT cofactor and any C_1-intermediates linked to this cofactor. Therefore, the flux of formaldehyde through the H_4MPT pathway should be zero and the full burden of formaldehyde production would fall on these mutants, upon exposure to methanol.

Mutants defective for *fae* exhibit an intermediate level of sensitivity to methanol or formaldehyde. Fae catalyzes the condensation of formaldehyde with H_4MPT, but this reaction also proceeds non-enzymatically at a lower rate (127). The fact that the *fae* mutant has a less severe phenotype than the H_4MPT biosynthesis mutants is consistent with the concept that the nonenzymatic condensation of formaldehyde with H_4MPT occurs at sufficient levels to allow a low level of formaldehyde oxidation through this pathway in the absence of Fae activity, but not enough to handle the full formaldehyde flux of methylotrophic growth.

Mutants lacking MtdB activity have the least severe phenotype of the H_4MPT pathway mutants investigated in this work. Two methylene-H_4MPT dehydrogenases are present in *M.*

extorquens AM1, MtdA (NADP-dependent, but also utilizes methylene-H_4F) and MtdB (H_4MPT-specific, but utilizes either NAD^+ or $NADP^+$). The sensitivity of *mtdB* mutants to methanol or formaldehyde and the inability to grow on methanol indicated that this enzyme plays a critical role in formaldehyde oxidation (43). The relatively moderate sensitivity of the *mtdB* mutant compared to either the *fae* or H_4MPT biosynthesis mutants indicates that, despite being insufficient for growth on C_1 compounds or complete resistance to methanol or formaldehyde, MtdA activity can support a moderate formaldehyde flux in the absence of MtdB. To further address this question, *mtdA* was cloned and overexpressed to levels 7.4-fold higher than in wild-type. This level of MtdA activity was insufficient to allow growth on methanol, however, there was a modest increase in the ability to grow on methylamine and it largely alleviated the sensitivity to methanol. These data suggest that despite the normal high level of MtdA activity in the wild-type, the enzyme level is limiting in the absence of MtdB activity. It has been suggested that the requirement for MtdA to use $NADP^+$, rather than NAD^+, in methylene-H_4MPT reduction limits its *in vivo* activity (125). This idea is also consistent with the observation of spontaneous mutants, most of which were defective in methanol oxidation. The ease of obtaining methanol-resistant *mtdB* mutants relative to the other H_4MPT pathway mutants is likely due to the modest inhibition of methanol on *mtdB* defective strains. Particularly for *mtdB* mutants, methanol-sensitivity is not lethal selection, therefore, mutations providing resistance would be expected occur during colony outgrowth and appear at higher than expected frequency, in contrast to the requirement for preexisting mutations with antibiotic selections (73).

The mutant phenotypes discussed above are correlated with the magnitude of the decreased formaldehyde flux through the H_4MPT pathway. For the mutants with the greatest defect, the H_4MPT biosynthesis mutants (*dmrA* and *orf4*), the impact is remarkable considering that the MIC drops at least 5 orders of magnitude compared to the wild-type. Our demonstration that this phenotype can be at least partially compensated with an alternate NAD- and GSH-linked

formaldehyde oxidation system demonstrates that this H_4MPT pathway not only serves as the main energy–generating pathway during methylotrophic growth, it also must be the major pathway through which formaldehyde flows. It is notable that an analogous methanol-sensitive phenotype has been observed for *P. denitrificans* mutants lacking *flhA* (102) or *fghA* (45), which demonstrates the widespread importance for methylotrophic bacteria to maintain the capacity for formaldehyde detoxification. Whether the growth inhibition observed for *M. extorquens* AM1 H_4MPT pathway mutants is due to a direct effect of formaldehyde accumulation itself, or indirectly due to a reactive conjugate of formaldehyde with another compound (58), or perhaps even a regulatory circuit poised to sense an imbalance of formaldehyde production and utilization is unclear at this time.

While this work and a previous report (43) have implicated the H_4MPT pathway in formaldehyde oxidation, a cellular function not generally thought of as essential during growth on multi-carbon compounds, null mutants had not been obtained for genes encoding Mch and Fhc, as well as a number of genes of unknown function suspected to be involved in the H_4MPT pathway (20). Among the possible explanations for this are that, 1.) A C_1-H4MPT intermediate is required for biosynthesis, 2.) These mutants are even more sensitive to methanol, such that ambient concentrations inhibit growth, or 3.) Mutations that create a complete block of the H_4MPT pathway cause C_1-H_4MPT intermediates to accumulate leading to either toxicity or regulatory problems. Null mutants were not obtained in wild-type or any mutant background except the *orf4* mutant defective for H_4MPT biosynthesis, and were no more methanol-sensitive than the mutant only lacking *orf4*. These data indicate that the H_4MPT pathway is not absolutely required for growth due to either biosynthetic needs or an even more extreme sensitivity to methanol. Rather, the ability to remove Mch or Fhc activity only in the absence of H_4MPT suggests it is the accumulation of a C_1-H_4MPT intermediate(s) that are either toxic and/or cause global regulatory problems. Further work will be required to distinguish between these possibilities.

The work presented here demonstrates that *M. extorquens* AM1 relies on the H_4MPT pathway to oxidize formaldehyde both during growth on C_1 substrates and to detoxify formaldehyde during growth on multi-carbon compounds. Remarkably, the heterologous GSH-dependent pathway from *P. denitrificans* is able to largely replace this function. This result indicates that these pathways comprise analogous metabolic modules (18, 46). Although they use entirely different enzymes and cofactors they can fulfill the same cellular function, namely, the NAD(P)-dependent oxidation of formaldehyde to formate. This work has led to a clearer understanding of the role of the H_4MPT pathway in *M. extorquens* AM1, providing further evidence for the importance of the H_4MPT pathway in formaldehyde oxidation, as well as further support for the hypothesis that the analogous H_4F pathway operates mostly in the reductive direction, and thus plays little or no role in formaldehyde oxidation. This hypothesis is developed further in Chapter 4 and tested directly in Chapter 5.

Table 3.1: *M. extorquens* AM1 strains described in Chapter 3

Strain	Description	Source
CM115.1	*fae ::kan*	This chapter
CM194.1	Δ*mxaF*	Chapter 2
CM194K.1	Δ*mxaF ::kan*	Chapter 2
CM194-198K.1	Δ*mxaF,* Δ*fae::kan*	This chapter
CM194-212K.1	Δ*mxaF,* Δ*dmrA::kan*	This chapter
CM194-253K.1	Δ*mxaF,* Δ*orf4::kan*	This chapter
CM194-258K.1	Δ*mxaF,* Δ*mtdB::kan*	This chapter
CM198.1	Δ*fae*	This chapter
CM198K.1	Δ*fae::kan*	This chapter
CM212K.1	Δ*dmrA::kan*	This chapter
CM253.1	Δ*ofr4*	This chapter
CM253K.1	Δ*orf4::kan*	This chapter
CM253-263K.1	Δ*orf4,* Δ*mch::kan*	This chapter
CM253-266K.1	Δ*orf4,* Δ*fhcBADC*	This chapter
CM258.1	Δ*mtdB*	This chapter
CM258K.1	Δ*mtdB::kan*	This chapter
M. extorquens AM1 Rif^R derivative		(89)
MR1	Spontaneous methanol-resistant mutant of CM258K.1	This chapter
MR2	Spontaneous methanol-resistant mutant of CM258K.1	This chapter
MR3	Spontaneous methanol-resistant mutant of CM258K.1	This chapter
MR4	Spontaneous methanol-resistant mutant of CM258K.1	This chapter
S64-67	*dmrA ::Is*phoA*/hah-Tc*	Chapter 2

Table 3.2: Plasmids described in Chapter 3

Plasmid	Description	Source
pAYC61	Allelic exchange vector	(17)
pCM62	Broad-host-range cloning vector (P_{lac})	Chapter 2
pCM80	*M. extorquens* AM1 expression vector (P_{mxaF})	Chapter 2
pCM102	pCR2.1 with *flhA* from *P. denitrificans*	This chapter
pCM103	pCR2.1 with *fghA* from *P. denitrificans*	This chapter
pCM104	pCM80 with *fghA*	This chapter
pCM106	pCM80 with *flhA-fghA*	This chapter
pCM112	pCR2.1 with 2.0 kb *fae* region	This chapter
pCM113	pUC19 with 2.0 kb *fae* region	This chapter
pCM114	pUC19 with *fae ::kan*	This chapter
pCM115	pAYC61 with *fae ::kan*	This chapter
pCM138	pCR2.1 with P_{fae}-*fae*	This chapter
pCM139	pCM62 with P_{fae}-*fae*	This chapter
pCM157	Broad-host-range *cre*-expression vector	Chapter 2
pCM184	Broad-host-range allelic exchange vector	Chapter 2
pCM185	pCR2.1 with *dmrA*	This chapter
pCM186	pCM62 with *dmrA*	This chapter
pCM195	pCR2.1 with *fae* upstream flank	This chapter
pCM196	pCR2.1 with *fae* downstream flank	This chapter
pCM197	pCM184 with *fae* upstream flank	This chapter
pCM198	pCM197 with *fae* downstream flank	This chapter
pCM207	pCR2.1 with *dmrA* upstream flank	This chapter
pCM208	pCR2.1 with *dmrA* downstream flank	This chapter
pCM211	pCM184 with *dmrA* upstream flank	This chapter
pCM212	pCM211 with *dmrA* downstream flank	This chapter
pCM250	pCR2.1 with *orf4* upstream flank	This chapter
pCM251	pCR2.1 with *orf4* downstream flank	This chapter
pCM252	pCM184 with *orf4* upstream flank	This chapter
pCM253	pCM252 with *orf4* downstream flank	This chapter
pCM254	pCR2.1 with *mtdA*	This chapter

Table 3.2: Plasmids described in Chapter 3 (Continued)

Plasmid	Description	Source
pCM255	pCR2.1 with *mtdB* upstream flank	This chapter
pCM256	pCR2.1 with *mtdB* downstream flank	This chapter
pCM257	pCM184 with *mtdB* downstream flank	This chapter
pCM258	pCM257 with *mtdB* upstream flank	This chapter
pCM259	pCM80 with *mtdA*	This chapter
pCM260	pCR2.1 with *mch* upstream flank	This chapter
pCM261	pCR2.1 with *mch* downstream flank	This chapter
pCM262	pCM184 with *mch* upstream flank	This chapter
pCM263	pCM262 with *mch* downstream flank	This chapter
pCM264	pCR2.1 with *fhcC* downstream flank	This chapter
pCM265	pCM184 with *orf4* upstream (*fhcB* upstream) flank	This chapter
pCM266	pCM265 with *fhcC* downstream flank	This chapter
pCR2.1	PCR cloning vector	Invitrogen
pLC310BgBm	pRK310 with *Bgl*II-*Bam*HI fragment containing *mtdB*	(20)
pRK2073	Helper plasmid expressing IncP *tra* functions	(36)
pRK310	19 kb IncP cloning vector	(33)
pWRxox451	Plasmid with *P. denitrificans* region containing *flhA* and *fghA*	(101)
pUC4K	Vector with kanamycin cassette	(123)
pUC19	Universal cloning vector	(134)

Table 3.3 – Primers described in Chapter 3

Primer	Sequence (5′-3′)
CM-4010r	TGACTGCCTCCGATCTAAG
CM-Dfae1	CGGGTTTCGTGACCTGTTC
CM-Dfae2	GTTATGCGGCCGCCATATGCATGGAAGCCATCCTTGTTTGC
CM-Dfae3	GCTTATCGATACCGTCGACCTCGAGGCAGTCCTGGGCAGA
CM-Dfae4	CGGGCATCGAGCGTTTCAC
CM-dmrAdf	ACCGCGGAGCTGATTCCTATTTTGCAC
CM-dmrAdr	TGAGCTCGGCGATGGTGTAGCTGTC
CM-dmrAf1	CAAGCTTCAGGTTTCGCGCCATGATC
CM-dmrAr	GAGATCTGTGGGTGCAAAATAGGAATC
CM-dmrAuf	ACTGCAGATCTACGCGATGGTCGGCAATG
CM-dmrAur	GAAGCTTGGTACCATGGTCGATCATGGCGCGAAAC
CM-faef	GTCCCAAATCGATGACGAAG
CM-faer	GGTTCACGCGATGTCTCAC
CM-fhcCdf	ACCGCGGCATATGCAAGGGCGAACTGTTCGTG
CM-fhcCdr	ACCATGGACCGGTGGCGAACCGGACTCTCAG
CM-fghAf	GTCTAGAGTTCCACGACTTGACAGAAG
CM-fghAr	TGAGCTCGCCTAGCCCAAGGCTC
CM-flhAf	GAGGCCTACATGAAGGAGAAGAC
CM-flhAr	GTCTAGATAAGCAGCCCTTGACGATC
CM-mchdf	TGGGCCCGTTCGTCTAAGGCAACGATC
CM-mchdr	AGAGCTCGAAACAACGAAGGTGTC
CM-mchuf	TGACGTCTGCGGCCGTTCCGCTATG
CM-mchur	TGGTACCTAGTCAGCGACGAGGCTCTC
CM-mtdAf	CCGGCCAGAGGAATCATGTC
CM-mtdAr	CCTCAGGCCATTTCCTTG
CM-mtdBdf	ACCGCGGAGAGGATCGCCGGATGAGTCC
CM-mtdBdr	AGAGCTCCATCGGACAGAAAGTTCAG
CM-mtdBuf	ACTGCAGACGTCGCGAGACCGAAGCGTTC
CM-mtdBur	TGGTACCGAGCGGGCCATGGTCTTC
CM-orf4df	AGGGCCCGGTTGCGGAAGTTTAGAG

Table 3.3 – Primers described in Chapter 3 (Continued)

Primer	Sequence (5′-3′)
CM-orf4dr	AGAGCTCCGAAGCGCCACTTCATG
CM-orf4uf	GTGTACAGCTAGCCATGGCCCAGGCTGCCATGCTTC
CM-orf4ur	GTCTAGAGGTACCCTTCCGCCGACGTTG
CM-Pfaef	GGATCCTGAGCCTTGGTCCAG
mtdAfor	CCGGCCAGAGGAATCATG
mtdArev	CCTCAGGCCATTTCCTTG
PsgaAf1	AAGAGCGTCCGAACGCGGAT
PsgaAr1	TACGGGCGAGCGTGTTGGACA

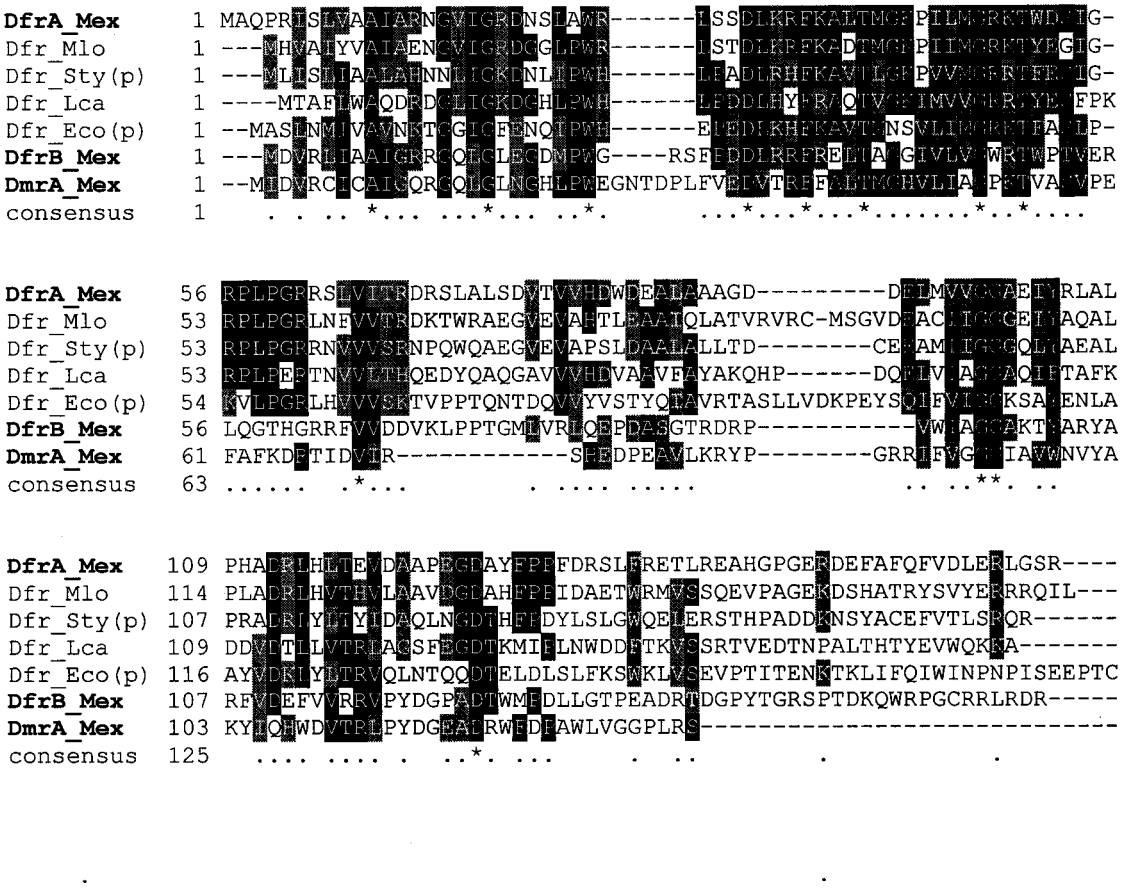


Figure 3.1: Multiple sequence alignment of DHFR amino acid sequences. DHFR homologs encoded by *M. extorquens* AM1 are indicated in bold. Sequences used include DfrA_Mex (*M. extorquens* AM1, AY093432), Dfr_Mlo (*Mesorhizobium loti*, BAB48827), Dfr_Sty(p) (*Salmonella typhimurium* plasmid pAZ1, RDEBDT), Dfr_Lca (*Lactobacillus casei*, P00381), Dfr_Eco(p) (*Escherichia coli* plasmid pCJ001, A49788), DfrB_Mex (*M. extorquens* AM1, AY093433), and DmrA_Mex (putative dihydromethanopterin reductase of *M. extorquens* AM1, AY09341). Alignment was constructed using Clustal W (dot.imgen.bcm.tmc.edu:9331/multi-align/Options/clustalw.html) and Shadebox (www.ch.embnet.org/software/BOX_form.html).

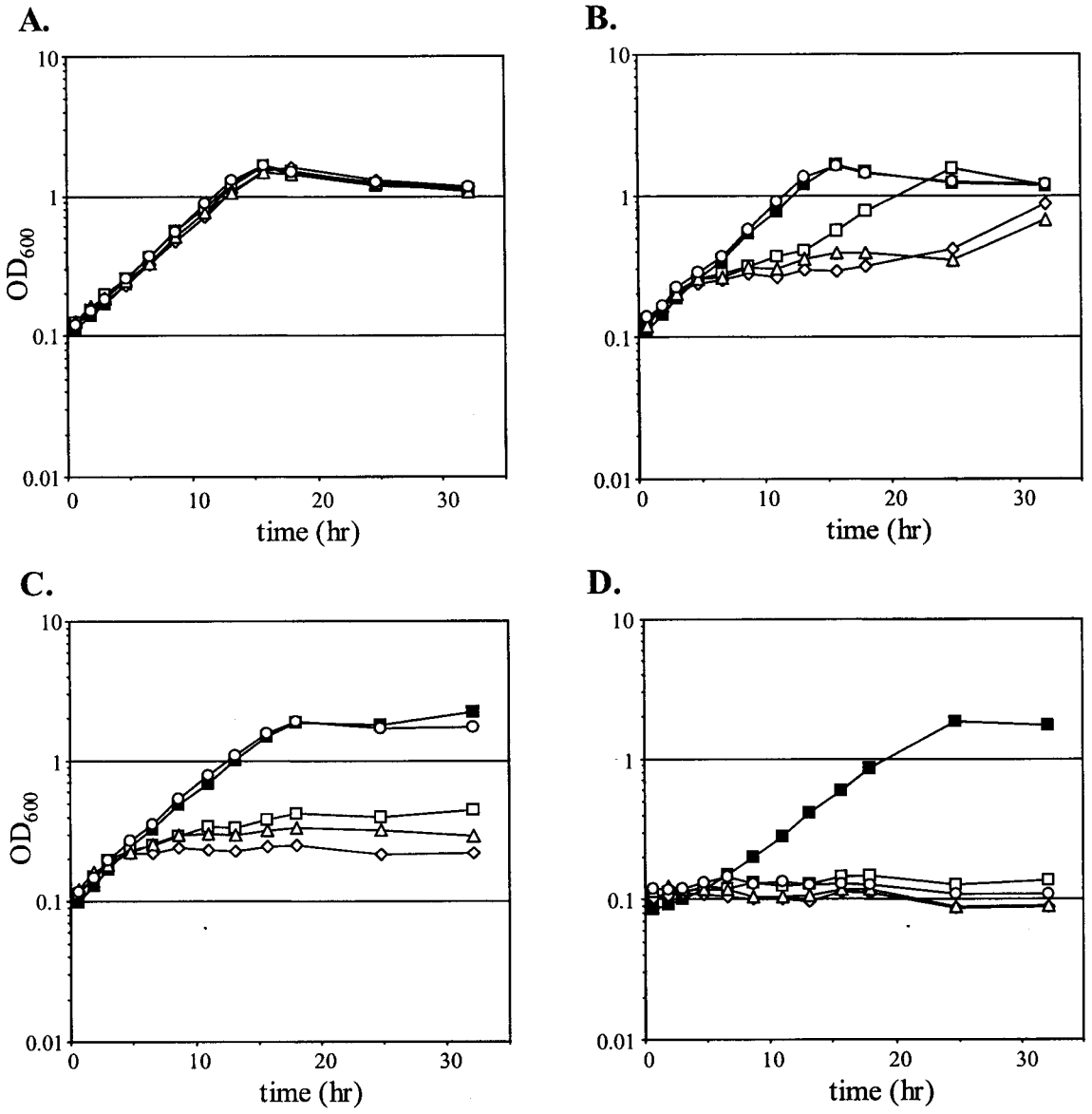


Figure 3.2: Growth of H_4MPT pathway mutants. Wild-type *M. extorquens* AM1 and mutant strains pre-grown on succinate, harvested, and resuspended in medium containing succinate (A.), succinate with methanol added to 1 mM (B) or 125 mM (C) at two hours, or methanol (D). The strains represented are wild-type (filled squares), the *fae* mutant CM198K.1 (open squares), the *dmrA* mutant CM212K.1 (open diamonds), the *orf4* mutant CM253K.1 (open triangles), and the *mtdB* mutant CM258K.1 (open circles).

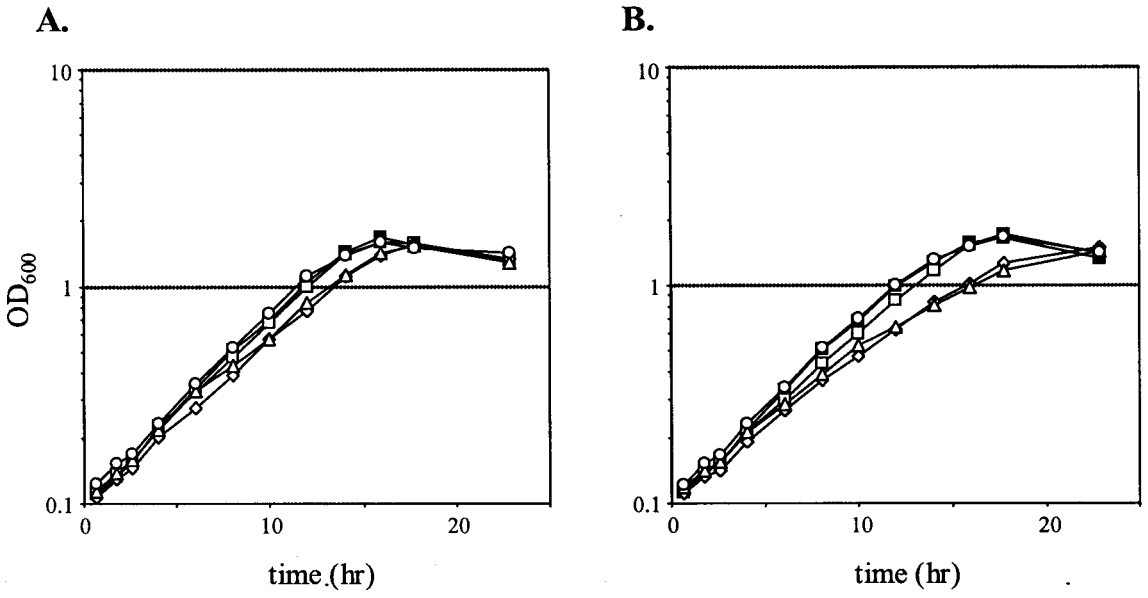


Figure 3.3: Growth of H_4MPT pathway mutants generated in Δ*mxaF* background. Mutant strains pre-grown on succinate, harvested, and resuspended in medium containing succinate (A.) or succinate with methanol added to 125 mM at two hours (B). The strains represented the *mxaF* mutant CM194.1 (filled squares), the *mxaF, fae* mutant CM194-198K.1 (open squares), the *mxaF, dmrA* mutant CM194-212K.1 (open diamonds), the *mxaF, orf4* mutant CM194-253K.1 (open triangles), and the *mxaF, mtdB* mutant CM194-258K.1 (open circles).

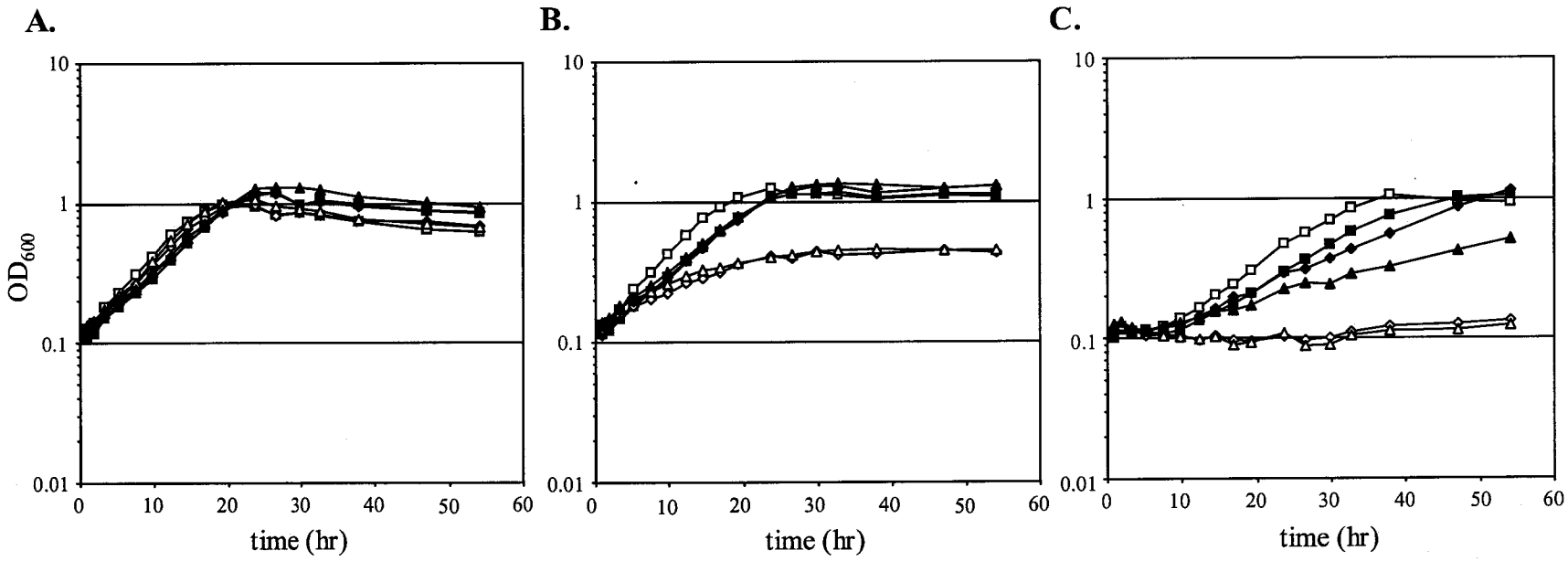


Figure 3.4: Complementation of H_4MPT pathway mutants by GSH pathway . Wild-type *M. extorquens* AM1 and mutant strains with plasmids were pre-grown on succinate, harvested, and resuspended in medium containing succinate (A.), succinate with methanol added to 125 mM at two hours (B), or methanol (C). All medium also contained tetracycline for plasmid maintenance. The strains represented are wild-type (squares), the *fae* mutant CM198K.1 (diamonds), and the *orf4* mutant CM253K.1 (triangles) with the empty vector pCM80 (open symbols) and the pCM106 plasmid expressing *flhA-fghA* (filled symbols).

CHAPTER 4

Genetic and Physiological Analysis of the Role of the Tetrahydrofolate-linked C_1 Transfer Pathway in *Methylobacterium extorquens* AM1

Abstract:

This chapter focuses on the role of the tetrahydrofolate (H_4F)-linked pathway for C_1 transfers in the serine cycle methylotroph *Methylobacterium extorquens* AM1. Previously, this pathway had been proposed to be responsible for the oxidation of formaldehyde to formate for energy metabolism (77). However, the results presented in Chapter 3 called that role into question. The purification and characterization of formate-tetrahydrofolate ligase (FtfL) confirmed that this enzyme is encoded by the *ftfL* homolog identified previously through transposon mutagenesis in Chapter 2. Physiological analysis of *ftfL* mutant strains demonstrated that FtfL activity is required for growth on C_1 compounds. The *ftfL* mutant strains do not exhibit phenotypes indicative of defective formaldehyde oxidation, unlike mutants defective for the H_4MPT pathway (Chapter 3). This represented the first mutant with a complete block of the H_4F-linked C_1 transfer pathway, for it had not been possible to generate null mutants lacking either *mtdA* (encodes an NADP-dependent methylene-H_4F/Methylene-H_4MPT dehydrogenase) or *fch* (encodes methenyl-H_4F cyclohydrolase). An unmarked strain was generated that expressed the analogous *folD* from *M. chloromethanicum* CM4, which encodes a bifunctional NADP-dependent methylene-H_4F dehydrogenase/methenyl-H_4F cyclohydrolase, however, and in this strain, null mutants of *mtdA* and/or *fch* could be obtained. These strains grew normally on multi-carbon substrates but were defective for growth on C_1 substrates. Additionally, null mutants of *mtdA* and/or *fch* were obtained in wild-type by supplementing the medium with formate. These approaches have demonstrated that the apparent essentiality of *mtdA* and *fch* is due to the need for formyl-H_4F for biosynthesis of purines and other compounds. Collectively, these data demonstrate that the H_4F pathway is not the key formaldehyde oxidation pathway in *M. extorquens* AM1. Rather, the data suggest an alternative model for the role of the H_4F pathway in this organism, in which it functions to convert formate to methylene H_4F for assimilatory metabolism.

Introduction:

The enzymes of the H_4F pathway are found at high specific activities in serine cycle methylotrophs and are generally present at 3-4 fold higher levels during growth on C_1 compounds than on multi-carbon compounds (22, 77, 97, 125). As the H_4F-dependent C_1 transfer reactions are reversible, it has been postulated that this pathway is responsible for channeling carbon into the serine cycle during growth on formate (66), but this has never been demonstrated by mutant analysis. Additionally, the absence of significant levels of NAD- (53) or dye-linked (77) formaldehyde dehydrogenase activities led to the suggestion that the H_4F-linked C_1 transfer pathway might be the key formaldehyde oxidation pathway in these organisms (77) (Fig. 6A). The surprising discovery of the H_4MPT pathway in *M. extorquens* AM1 (20) and other methylotrophs (126) refocused attention on the role of the H_4F C_1 transfer pathway in methylotrophic metabolism. My work presented in Chapter 3 clearly demonstrates a critical role for the H_4MPT pathway in formaldehyde oxidation and detoxification, bringing into question the previously suggested role of the H_4F pathway.

Results of the previous genetic analyses of H_4F-dependent C_1-transfer in *M. extorquens* AM1 have been somewhat inconclusive. Null mutants of *mtdA* or *fch* could not be obtained even selecting on multi-carbon substrates such as succinate (22, 23, 125). This was assumed to be due to their critical role during growth on multi-carbon compounds, likely in producing formyl-H_4F for the biosynthesis of purines and other compounds. Mutants with a reduced activity of MtdA or Fch were obtained, however, and these were found to be defective for growth on C_1 compounds. These data established a role for these enzymes in methylotrophy, but did not clarify whether that role was in formaldehyde oxidation or another function. Formate-H_4F ligase activity had been detected in serine cycle methylotrophs (67, 77) but a candidate gene responsible for encoding its activity had not been identified. Therefore, a high priority of my work was to identify this gene,

generate a mutant, and characterize it phenotypically, to determine whether a complete H_4F pathway was required for methylotrophy. As noted in Chapter 2, the transposon mutagenesis screen identified a candidate for this gene.

Our collaborators purified FtfL from cell extracts of *M. extorquens* AM1 (Marx et al., submitted) and determined that it is encoded by the *ftfL* homolog identified through transposon mutagenesis (Chapter 2). Amino acid residues predicted to participate in binding ATP, H_4F, and formyl phosphate (100) are conserved within FtfL from *M. extorquens* AM1. Therefore, the highly conserved sequence of FtfL from *M. extorquens* AM1 suggests similar chemical, physical, and enzymatic properties as for the well-studied enzymes present in Gram-positive bacteria. As noted in Chapter 2, the putative *ftfL* mutant had the phenotype of being unable to grow on C_1 compounds. Here I present a physiological analysis of *ftfL* mutant strains, which establishes that this enzymatic activity, and thus a complete H_4F pathway, is required for growth on C_1 compounds. I have extended my genetic analysis of the role of the H_4F pathway in methylotrophy by using an unmarked strain expressing *folD* from *M. chloromethanicum* CM4, which encodes a bifunctional NADP-dependent methylene-H_4F dehydrogenase/methenyl-H_4F cyclohydrolase (115, 121) to generate for the first time null mutants lacking *mtdA* and/or *fch*. Additionally, I found that null mutants of *mtdA* and/or *fch* could be generated in wild-type by supplementing the medium with formate. These approaches have demonstrated that the apparent essentiality of *mtdA* and *fch* is due to the need for formyl-H_4F and have clarified the role of the H_4F pathway in methylotrophy. My data are inconsistent with the model in which the main role of this pathway is in formaldehyde oxidation (77). Rather, my physiological data led me to present the hypothesis that there was an unknown mechanism by which C_1 units from the H_4MPT pathway entered the H_4F pathway, and could then be converted into methylene-H_4F, the starting substrate for the serine cycle. Several months later, I learned that our collaborators in the lab of Dr. J. A. Vorholt discovered that Fhc releases formate, rather than CO_2, as the end product of the H_4MPT pathway

(95). The remarkable result from this was that formate connects the two pathways, providing the biochemical mechanism necessary for my hypothesis that had been based on genetic analysis. In Chapter 5 I describe my experiments that directly confirmed this new model of methylotrophy.

Materials and Methods:

Standard laboratory procedures. The growth of bacterial strains, genetic procedures, and recombinant DNA techniques are performed as described in Chapter 2. Phenotypic analyses of mutant strains are performed as described in Chapter 3. *M. extorquens* AM1 strains, plasmids, and primers used in Chapter 4 are described in Tables 4.1, 4.2, and 4.3.

Generation of *ftfL* mutant strains and complementing plasmid. *M. extorquens* AM1 mutants defective for *ftfL* were generated using pCM184. The regions immediately flanking *ftfL* were amplified by PCR using CM-ftfLuf and CM-ftfLur, and CM-ftfLdf and CM-ftfLdr. The resulting products were cloned into pCR2.1 (Invitrogen) to produce pCM213 and pCM214. The 0.6 kb *Bgl*II-*Nco*I fragment from pCM213 was then introduced between the corresponding sites of pCM184 to produce pCM215. Subsequently, the 0.5 kb *Sac*II-*Sac*I fragment from pCM214 was ligated into the same sites of pCM215 to produce pCM216. In addition to the ftfL mutant strains CM216K.1 (Δ*ftfL::kan*) and CM216.1 (Δ*ftfL*), two double mutant strains were obtained by introducing pCM216 into CM198.1 to generate the Δ*fae*, Δ*ftfL ::kan* mutant CM198-216K.1, and by introducing pCM212 into CM216.1 to generate the Δ*ftfL*, Δ*dmrA ::kan* mutant CM216-212K.1.

In order to construct a plasmid to complement *ftfL* defective strains, a 2.7 kb region containing the *ftfL* coding region and putative promoter was amplified by PCR and cloned into pCR2.1 (Invitrogen) to produce pCM217. The entire 2.7 kb region of pCM217 was sequenced (University of Washington Biochemistry Department DNA Sequencing Facility) to confirm the

sequence present on the ERGO website (www.integratedgenomics.com/genomereleases.html#list6). The 2.7 kb *Hin*dIII-*Bam*HI fragment of pCM217 was cloned into the same sites of pCM62 to produce pCM218.

Generation of constructs containing *folD* and *purU* from *M. chloromethanicum* CM4. The coding regions of *purU* and *folD* were amplified from a chromosomal DNA preparation of *M. chloromethanicum* CM4 by PCR using CM-purUf and CM-purUr, and CM-folDf and CM-folDr. The resulting products were cloned into pCR2.1 (Invitrogen) to produce pCM201 and pCM202, respectively. Both constructs were sequenced to confirm no errors had been introduced. The 0.9 kb *Xba*I-*Kpn*I fragment from pCM201 was cloned into the same sites of pCM80 to generate pCM203, and subsequently the 1.0 kb *Kpn*I-*Sac*I fragment from pCM202 was introduced between the same sites of pCM203 to generate pCM205. The 1.9 kb *Xba*I-*Nsi*I fragment from pCM205 containing *purU-folD* was then inserted into the same sites of pCM172 to generate pCM206. A construct for expression of just *folD* was made by self-ligating the 9.3 kb blunted pCM206 *Xba*I-*Asp*718I fragment to produce pCM219. Similarly, a *purU* expression construct was generated by self-ligating the 9.3 kb pCM206 *Asp*718I-*Nsi*I fragment to produce pCM240. Strains carrying insertion vectors were generated by electroporating the appropriate constructs into the *katA::kan* strain CM82.1 as described (118). Tetracycline-resistant transformants were then screened for kanamycin sensitivity. Unmarked (tetracycline-sensitive) insertion strains were generated using the *cre*-expressing plasmid pCM158 as described in Chapter 2.

Construction of donor plasmids to generate mutants defective for *mtdA* and/or *fch*. *M. extorquens* AM1 deletion mutants lacking *mtdA* and/or *fch* were generated using the allelic exchange vector pCM184. Approximately 0.5 kb regions upstream and downstream of each of

these genes were amplified by PCR using the primer pairs CM-mtdAuf and CM-mtdAur, CM-mtdAdf and CM-mtdAdr, CM-fchuf and CM-fchur, and CM-fchdf and CM-fchdr. The resulting *mtdA* flanks were introduced into pCR2.1 (Invitrogen) to generate pCM272 and pCM273, and the *fch* flanks are contained in pCM276 and pCM277. The construct to generate Δ*mtdA::kan* mutants was made by introducing the 0.5 kb *Sac*II-*Age*I fragment from pCM273 between the corresponding sites of pCM184 to produce pCM274, and subsequently, the 0.5 kb *Bgl*II-*Nde*I fragment from pCM272 was ligated into the same sites of pCM274 to produce pCM275. The construct to generate Δ*fch::kan* mutants was made by introducing the 0.6 kb *Apa*I-*Sac*I fragment from pCM277 into the same sites of pCM184 to produce pCM278, and subsequently, the 0.5 kb *Eco*RI-*Nde*I fragment from pCM276 was ligated into the same sites of pCM278 to produce pCM279. Finally, a construct to make Δ*mtdA-fch::kan* mutants was generated by introducing the 0.5 kb *Bgl*II-*Nde*I fragment from pCM272 into the same sites of pCM278 to produce pCM280.

Enzyme assays. NADP-dependent methylene-H_4F dehydrogenase (22), methenyl-H_4F cyclohydrolase (97), formate-H_4F ligase (99) and formyl-H_4F hydrolase (87) activities were assayed were assayed as described with extracts prepared from cell material that was harvested from exponential phase cultures. Between-culture variability in enzyme activities was less than 20%. Total protein content of the extracts was determined spectrophotometrically (54, 130) using a Beckmann DU 640B spectrophotometer.

Results:

Generation of a Δ*ftfL::kan* mutant by allelic exchange and phenotypic analysis. In order to confirm the C_1^- phenotype observed for *ftfL::IS*phoA/*hah-Tc* mutant strains (Chapter 2) a Δ*ftfL::kan* mutant, CM216K.1, was generated using the allelic exchange vector pCM184. Cell

extracts of the resulting Δ*ftfL::kan* strain CM216K.1 lacked detectable FtfL activity. The CM216K.1 mutant grew like wild-type *M. extorquens* AM1 on solid medium containing succinate, but showed no growth on plates containing methanol or methylamine. The mutant strain containing a plasmid with the *ftfL* gene grew normally, demonstrating that the defect in the mutant was due to the loss of FtfL. Growth experiments in liquid medium containing either succinate or methanol confirmed these results (Fig. 4.1). Furthermore, no growth was observed on plates containing either formate or oxalate, which is catabolized through formate in organisms such as *Oxalobacter formigenes* (1). Addition of either methanol or formaldehyde to succinate plates only slightly inhibited CM216K.1, in contrast to the severe inhibition effect observed for mutant strains defective for the H_4MPT pathway (Chapter 3) (43, 80, 127). The MIC of methanol or formaldehyde was found to be 125 and 0.5 mM, respectively, whereas the H_4MPT pathway mutant defective for Fae, for example, is sensitive to 0.05 to 0.1 and 0.1 to 0.2 mM (Chapter 3) (127). Similarly, growth in liquid medium was not affected by the addition of 125 mM methanol (Fig. 4.1). These data are consistent with the preliminary analysis of the *ftfL::IS*phoA/*hah* mutants (Chapter 2) and are the first demonstration that FtfL activity is required for growth on C_1 compounds. Furthermore, these data suggest that the H_4F pathway may play a minor role, if any, in formaldehyde oxidation or detoxification.

The FtfL-deficient mutant is not complemented by the expression of the GSH-dependent formaldehyde oxidation pathway. The C_1^- and methanol-sensitive phenotype of mutants defective for the H_4MPT pathway in *M. extorquens* AM1 can be complemented by the heterologous expression of the enzymes for the GSH-dependent formaldehyde oxidation pathway of *P. denitrificans* (Chapter 3). This result suggested that the H_4MPT pathway is required for formaldehyde oxidation and detoxification, and called into question whether the endogenous H_4F pathway significantly contributes to these functions. In order to determine whether the C_1^- mutant

phenotype of the *ftfL* mutant is also due to defective formaldehyde oxidation, the pCM106 plasmid that expresses the GSH pathway was introduced into CM216K.1. The presence of pCM106 did not alter the mutant phenotypes. Again, this result is inconsistent with the hypothesis that the H_4F pathway may be required for the oxidation of formaldehyde to formate.

Mutants defective for both the H_4F and H_4MPT pathways are not more sensitive to methanol or formaldehyde than mutants solely lacking the H_4MPT pathway. As a second physiological test of the hypothesis that the H_4F pathway is required for a role other than net formaldehyde oxidation to formate, mutants were generated that were defective for both the H_4F and the H_4MPT pathways to determine whether the double mutants would exhibit a more severe physiological defect than either single mutant alone. The H_4MPT pathway was interrupted at two levels, at *fae,* which encodes the enzyme that generates methylene H_4MPT, and at *dmrA*, which encodes the putative dihydromethanopterin reductase. The *fae* mutant was shown to possess limited ability for oxidation of formaldehyde through the H_4MPT pathway, via the non-enzymatic condensation (127), while the *dmrA* mutants do not produce H_4MPT (S. Wyles and M. E. Rasche, personal communication) and therefore should possess no H_4MPT pathway activity. Double Δ*fae*, Δ*ftfL::kan* (CM198-216K.1) and a Δ*ftfL*, Δ*dmrA::kan* (CM216-212K.1) mutants were generated. Phenotypes of these mutants were compared to the single mutants CM198K.1 and CM212K.1 defective for *fae* and *dmrA*, respectively, on solid succinate medium containing a range of methanol or formaldehyde concentrations. In all cases, CM198K.1 and CM198-216K.1 were found to be equally sensitive, and the same was true for the pair CM212K.1 and CM216-212K.1. These data provide additional evidence that the H_4F pathway does not play a significant role in formaldehyde oxidation.

***M. extorquens* AM1 mutants lacking *mtdA* and/or *fch* can be generated in a strain expressing *folD* from *M. chloromethanicum* CM4.** In order to better understand the role of the *M. extorquens* AM1 H_4F pathway in methylotrophy and the apparent essentiality of *mtdA* and *fch*, mutants defective for these H_4F pathway activities were generated in strains expressing analogous but non-orthologous enzymes from the related methylotroph *M. chloromethanicum* CM4. In order to accomplish this, *folD* (encodes a bifunctional NADP-dependent methylene-H_4F dehydrogenase/methenyl-H_4F cyclohydrolase) and *purU* (putative formyl-H_4F hydrolase) from *M. chloromethanicum* CM4 (115, 121)were cloned and introduced into the insertional expression vector pCM172. These constructs were introduced into CM82.1 and tetracycline-resistant, kanamycin-sensitive transformants were isolated. This resulted in three antibiotic-resistance-free insertion strains that contain either *folD*, *purU*, or both behind P_{mxaF}. Enzymatic assays confirmed that FolD was expressed in active form, with 81 and 39 mU of NADP-dependent methylene-H_4F dehydrogenase activity on methanol and succinate, respectively. Repeated attempts to detect formyl-H_4F hydrolase activity indicative of expression of active PurU were unsuccessful. Even the prototypic *E. coli* PurU is difficult to assay, however, such that extracts from *E. coli* overexpressing PurU to 35% of the total protein contained only 30 mU of activity, and that value dropped by 96% if the allosteric activator of the *E. coli* PurU, methionine, is omitted (87). Although the inability to detect activity in this case is not entirely surprising, conclusions cannot be drawn from experiments involving the *purU*-containing strains and thus will not be described further here. The *folD*-expressing strain CM219T.1 was unmarked using the *cre*-expression vector pCM158 to generate the antibiotic-resistance-free strain CM219.1 (*katA::(loxP-*t_{rrnB}*-*P_{mxaF}*-folD-*t_{T7}*)*) for further experiments.

Constructs based on pCM184 were generated to delete *mtdA*, *fch*, or both, and these were introduced into both wild-type *M. extorquens* AM1 and the *folD*-expressing strain CM219.1. As had been reported previously, null mutants were not obtained in wild-type on succinate medium

(22, 97), but were readily obtained in CM219.1. The resulting strains CM219-275K.1 (*katA::(loxP-*t_{rrnB}*-*P_{mxaF}*-folD-*t_{T7}*)*, Δ*mtdA::kan*), CM219-279K.1 (*katA::(loxP-*t_{rrnB}*-*P_{mxaF}*-folD-*t_{T7}*)*, Δ*fch::kan*), and CM219-280K.1 (*katA::(loxP-*t_{rrnB}*-*P_{mxaF}*-folD-*t_{T7}*)*, Δ*mtdA-fch::kan*) and CM219.1 grew like wild-type in medium containing succinate (Fig. 4.2). Additionally, the addition of methanol to the medium did not inhibit growth (Fig. 4.2), unlike the mutants defective for the H_4MPT pathway (Chapter 3), (43, 80, 127). The *folD*-expressing strain CM219.1 grew more slowly on methanol than wild-type, however, and the *mtdA* and *fch* mutants generated in this strain are unable to grow on methanol (Fig. 4.2). Similar results were obtained on solid medium, and additionally CM219-275K.1, CM219-279K.1, and CM219-280K.1 failed to grow on methylamine, formate, or oxalate.

Null mutants in *mtdA* and/or *fch* can be obtained in wild-type *M. extorquens* AM1 by supplementing the media with formate or another C_1 compound. Given that the presence of an alternative enzyme to convert methylene-H_4F to formyl-H_4F, FolD, eliminated the need for MtdA and Fch under standard heterotrophic conditions, I hypothesized that FtfL activity could similarly obviate these enzymes if the medium was supplemented with formate. Formate or compounds that are metabolized to formate (methanol, methylamine, or oxalate) were added to the medium throughout the conjugation procedure and to the medium used to select transconjugants. Under these conditions, the mutant strains CM275K.1 (Δ*mtdA::kan*), CM279K.1 (Δ*fch::kan*), and CM280K.1 (Δ*mtdA-fch::kan*) lacking *mtdA*, *fch*, or both were obtained. As expected, MtdA and/or Fch activity was undetectable in these strains (<1 mU), whereas wild-type exhibited 270 and 70 mU MtdA activity, and 250 and 240 mU of Fch activity on methanol and succinate, respectively. Growth of these mutants is nearly wild-type in succinate medium supplemented with 7 mM methylamine, but they fail to grow when transferred to succinate medium without methylamine (Fig. 4.3). Similarly, no growth is observed on succinate plates

unless they contain formate (7 mM), oxalate (4 mM), methanol (1 or 10 mM), or methylamine (7 mM), with the latter supplement supporting the most vigorous growth. Attempts to generate *mtdA* and/or *fch* mutants in the Δ*ftfL* strain CM216.1 produced no null mutants, confirming that FtfL activity is required to provide formyl-H_4F from formate under these conditions.

Mutants lacking *mtdA* and/or *fch* are unable to grow on C_1 compounds including formate but are not methanol-sensitive. In order to examine the role of MtdA and Fch in methylotrophy, cultures of wild-type, CM275K.1, CM279K.1, and CM280K.1 grown on succinate plus methylamine were transferred to medium containing succinate plus methylamine (7 mM), succinate plus methanol (added to 125 mM after two hours), and methanol alone (Fig. 5). No growth was observed in liquid medium containing methanol or on plates containing methanol, methylamine, formate, or oxalate as substrates. Introduction of the plasmid pLC410a, which contains *mtdA* and *fch,* completely restored wild-type growth to all mutant strains. Unlike mutants defective for the H_4MPT pathway (43, 80, 127), however, the addition of methanol did not inhibit growth. Attempts to generate *mtdA* null mutants in the Δ*mtdB* strain CM258.1 in the presence of formate failed, however, suggesting that the elimination of both methylene-H_4MPT dehydrogenases is not tolerated.

Discussion:

I have presented genetic and physiological analysis regarding the role of the H_4F-linked C_1 transfer pathway in the methylotrophic metabolism of *M. extorquens* AM1. *ftfL* mutants, which are blocked in the H_4F-linked interconversion of methylene-H_4F and formate, are incapable of growth on C_1 compounds, but are neither sensitive to formaldehyde-producing substrates nor are they complemented by the expression of the heterologous GSH pathway for formaldehyde

oxidation. Furthermore, even in the absence of the H_4MPT pathway, the degree of sensitivity to methanol was the same independent of the presence or absence of an intact H_4F pathway. This represents the first identification of an *ftfL* gene in a methylotroph and the first demonstration that FtfL activity is required for growth on C_1 compounds.

I have extended the genetic analysis of the role of the H_4F pathway by using both genetic and biochemical complementation to obtain null mutants lacking *mtdA* and/or *fch* in order to more fully understand their role in methylotrophy. First, *mtdA* and/or *fch* were deleted in an unmarked strain expressing *folD* from *M. chloromethanicum* CM4 from a chromosomal locus, and these strains grew indistinguishably from wild-type on succinate. This indicates that the bifunctional FolD enzyme (NADP-dependent methylene-H_4F dehydrogenase/methenyl-H_4F cyclohydrolase) can functionally replace MtdA and/or Fch under conditions in which the H_4F pathway is only required for the generation of formyl-H_4F for biosynthetic reactions. However, these strains were incapable of growth on C_1 compounds. Given that the level of FolD activity assayed *in vitro* was only three times lower than that for MtdA it came as a surprise that these strains exhibited such a dramatic growth defect. FolD from *M. chloromethanicum* CM4 is required for growth on chloromethane, a C_1 substrate which is not oxidized to formaldehyde, but rather, is catabolized through C_1-H_4F pathway intermediates by MetF (methylene-H_4F reductase), FolD and PurU (115, 121). The work presented in this chapter supports a previous suggestion (95, 124) that the H_4F pathway of *M. extorquens* AM1 (MtdA, Fch, and FtfL) functions in the assimilatory direction during growth on methanol to supply methylene-H_4F for the serine cycle from some fraction of the formate that is produced from formaldehyde by the H_4MPT pathway. Therefore, it is appears that the net flux through the two H_4F pathways are in opposite directions in these two *Methylobacterium* strains. In this regard, it is interesting to note that *M. chloromethanicum* CM4 actually contains both *folD-purU* and *mtdA-fch* (and likely *ftfL*) (115). I suggest that the inability of FolD to functionally replace MtdA and Fch in *M. extorquens* AM1 is

indicative that this enzyme has evolved to fulfill a fundamentally different role. This may be reflected, for example, in different affinities for substrates and/or the effect and identity of potential effector molecules that may modulate flow through C_1-H_4F intermediates. Further work with the purified enzymes would be required to test this hypothesis.

Second, the ability to obtain *mtdA* and/or *fch* null mutants by supplementing the medium with formate provides strong support for the concept that the apparent essentiality of these genes was simply due to the biosynthetic need for formyl-H_4F. This conclusion was further supported by the inability to generate nulls for these genes in a Δ*ftfL* strain. Mutants lacking *mtdA* and/or *fch* exhibit wild-type growth in succinate medium supplemented with methylamine, and are not inhibited by the presence of methanol. The insensitivity of *mtdA* and/or *fch* mutants to methanol provides additional support that the H_4F pathway does not contribute significantly to formaldehyde oxidation. In the case of *mtdA*, in particular, this suggests that MtdB activity alone is sufficient for the detoxification of formaldehyde. These results confirm the previous suggestion (97) that the MtdB is the primary methylene-H_4MPT dehydrogenase *in vivo*. Attempts to generate a *mtdA*, *mtdB* double mutant were unsuccessful, as has been reported for attempts to obtain null mutants lacking *mch* or *fhcBADC* (Chapter 3) (20). Null mutants lacking *mch* or *fhcBADC* were generated in an unmarked mutant strain defective for H_4MPT biosynthesis, however (Chapter 3). Therefore, it appears that, although the pathway can be completely removed and is thus not essential, any mutation, or pair of mutations, in the case of *mtdA* and *mtdB*, that leads to a block within the H_4MPT pathway is not tolerated. The inability to generate a *mtdA*, *mtdB* double mutant is therefore consistent with the suggestion made previously (Chapter 3) that the lack of viability of such mutants is due to the accumulation of C_1-H_4MPT intermediate(s) that may be either toxic and/or cause regulatory defects. Further work is required to distinguish between these possibilities and identify which intermediate(s) are responsible.

With the exception of the requirement for supplementation with formate to grow on succinate, or a chemical which can be converted into formate (methanol, methylamine, or oxalate), the growth phenotype of *mtdA* and/or *fch* mutants is consistent with that of *ftfL* mutants, thus providing a consistent phenotype associated with a defective H_4F pathway: no growth on C_1 compounds including formate, but lack of inhibition by methanol during growth on succinate. One role that has been suggested for the H_4F pathway in serine cycle methylotrophs is to function in the reductive direction, generating methylene-H_4F during growth on formate, thereby providing the means to assimilate carbon during growth on this substrate (Fig. 1), (66). In contrast to strains defective for the H_4MPT pathway, *ftfL* mutant strains failed to grow on formate, confirming the role of this pathway in formate utilization. Additionally, *ftfL* mutants were defective for growth on oxalate, which is converted to formate in other organisms that grow on this compound through the action of oxalyl-CoA decarboxylase (9) and formyl-CoA transferase (8). Consistent with this model for growth of *M. extorquens* AM1 on oxalate, mutants lacking one of the two putative formyl-CoA transferases found in the genome sequence (www.integratedgenomics.com/genomereleases.html#list6) fail to grow on oxalate (C. J. Marx, unpublished data). Interestingly, the initial assimilatory reactions during growth of serine cycle methylotrophs on formate mirror the initial steps of the Wood-Ljungdahl pathway utilized by acetogenic bacteria (72), with both classes of organisms utilizing FtfL to activate formate for further assimilation. Therefore, it is possible these two pathways share evolutionary roots in these C_1 interconversions.

The data presented here clearly demonstrate that the complete H_4F pathway is required for methylotrophic growth of *M. extorquens* AM1, but contradict the previous suggestion that serine cycle methylotrophs may oxidize formaldehyde via the H_4F-linked C_1 transfer pathway (77). My data support the alternative hypothesis independently put forth by myself and our collaborators (95, 124), that a portion of the formate produced by the H_4MPT pathway may be

assimilated via the reductive H_4F pathway. In accordance with this hypothesis, the H_4F pathway would function as a second route for the production of methylene-H_4F, the starting substrate for the serine cycle, in addition to the non-enzymatic condensation of formaldehyde with H_4F. This hypothesis is tested directly in Chapter 5.

Table 4.1: *M. extorquens* AM1 strains described in Chapter 4

Strain	Description	Source
CM82.1	*katA::kan*	Chapter 2
CM198.1	Δ*fae*	Chapter 3
CM198K.1	Δ*fae::kan*	Chapter 3
CM198-216K.1	Δ*fae*, Δ*ftfL ::kan*	This chapter
CM206T.1	*katA::(loxP-tetAR-loxP-*t_{rrnB}*-*P_{mxaF}*-purU-folD-*t_{T7}*)*	This chapter
CM212K.1	Δ*dmrA::kan*	Chapter 3
CM216.1	Δ*ftfL*	This chapter
CM216K.1	Δ*ftfL ::kan*	This chapter
CM216-212K.1	Δ*ftfL*, Δ*dmrA ::kan*	This chapter
CM219.1	*katA::(loxP-*t_{rrnB}*-*P_{mxaF}*-folD-*t_{T7}*)*	This chapter
CM219T.1	*katA::(loxP-tetAR-loxP-*t_{rrnB}*-*P_{mxaF}*-folD-*t_{T7}*)*	This chapter
CM219-275K.1	*katA::(loxP-*t_{rrnB}*-*P_{mxaF}*-folD-*t_{T7}*)*, Δ*mtdA::kan*	This chapter
CM219-279K.1	*katA::(loxP-*t_{rrnB}*-*P_{mxaF}*-folD-*t_{T7}*)*, Δ*fch::kan*	This chapter
CM219-280K.1	*katA::(loxP-*t_{rrnB}*-*P_{mxaF}*-folD-*t_{T7}*)*, Δ*mtdA-fch::kan*	This chapter
CM240T.1	*katA::(loxP-tetAR-loxP-*t_{rrnB}*-*P_{mxaF}*-purU-*t_{T7}*)*	This chapter
CM258.1	Δ*mtdB*	Chapter 3
CM275K.1	Δ*mtdA::kan*	This chapter
CM279K.1	Δ*fch::kan*	This chapter
CM280K.1	Δ*mtdA-fch::kan*	This chapter
M. extorquens AM1 Rif^R derivative		(89)

Table 4.2: Plasmids described in Chapter 4

Plasmid	Description	Source
pCM62	Broad-host-range cloning vector	Chapter 2
pCM80	*M. extorquens* AM1 expression vector (P_{mxaF})	Chapter 2
pCM106	pCM80 with *flhA-fghA* from *P. denitrificans*	Chapter 3
pCM157	Broad-host-range *cre*-expression vector	Chapter 2
pCM158	Broad-host-range *cre*-expression vector	Chapter 2
pCM172	*M. extorquens* AM1 insertional expression vector (P_{mxaF})	Chapter 2
pCM184	Broad-host-range allelic exchange vector	Chapter 2
pCM201	pCR2.1 with *purU* from *M. chloromethanicum* CM4	This chapter
pCM202	pCR2.1 with *folD* from *M. chloromethanicum* CM4	This chapter
pCM203	pCM80 with *purU*	This chapter
pCM205	pCM80 with *purU-folD*	This chapter
pCM206	pCM172 with *purU-folD*	This chapter
pCM212	Donor to generate Δ*dmrA ::kan* mutation	Chapter 3
pCM213	pCR2.1 with *ftfL* upstream flank	This chapter
pCM214	pCR2.1 with *ftfL* downstream flank	This chapter
pCM215	pCM184 with *ftfL* upstream flank	This chapter
pCM216	pCM215 with *ftfL* upstream flank	This chapter
pCM217	pCR2.1 with 2.7 kb *ftfL* region	This chapter
pCM218	pCM62 with *ftfL* region	This chapter
pCM219	pCM172 with *folD*	This chapter
pCM240	pCM172 with *purU*	This chapter
pCM272	pCR2.1 with *mtdA* upstream flank	This chapter
pCM273	pCR2.1 with *mtdA* downstream flank	This chapter
pCM274	pCM184 with *mtdA* downstream flank	This chapter
pCM275	pCM274 with *mtdA* upstream flank	This chapter
pCM276	pCR2.1 with *fch* upstream flank	This chapter
pCM277	pCR2.1 with *fch* downstream flank	This chapter
pCM278	pCM184 with *fch* downstream flank	This chapter
pCM279	pCM278 with *fch* upstream flank	This chapter
pCM280	pCM278 with *mtdA* upstream flank	This chapter

Table 4.2: Plasmids described in Chapter 4 (Continued)

Plasmid	Description	Source
pCR2.1	PCR cloning vector	Invitrogen
pLC410a	Large IncP plasmid containing *mtdA* and *fchA*	(22)
pRK2073	Helper plasmid expressing IncP *tra* functions	(36)

Table 4.3 – Primers described in Chapter 4

Primer	Sequence (5′-3′)
CM-fchdf	AGGGCCCGAAGTCCAAGGTAAACTGAG
CM-fchdr	AGAGCTCCGCATAGACCGCTTGATC
CM-fchuf	GGGCGCAAGCTCGACAAG
CM-fchur	TCATATGTTATGTTTCGATCGTCTCGTTG
CM-folDf	TGGTACCTCTACACATCAGGTTTAACATG
CM-folDr	AGAGCTCACGCACGGTCAGACATC
CM-ftfLdf	TCCGCGGTGAGATCATGACCATG
CM-ftfLdr	AGGATCCGAGCTCGACGTCGTCGCCTTC
CM-ftfLuf	GAAGCTTAGATCTAGGTTTCTCCCGTTCATTC
CM-ftfLur	AGGTACCATGGCACCGTCGTTTCTTGTTCAC
CM-mtdAdf	ACCGCGGCCAAGGAAATGGCCTGAG
CM-mtdAdr	TACCGGTCCGAGATGACGTTGAGATTG
CM-mtdAuf	AAGATCTGACGCAATCCGACGTCATC
CM-mtdAur	TCATATGTTACAGCTTCTTGGACATGATTC
CM-purUf	GTCTAGAGGGGAATCGCTCGATG
CM-purUr	TGGTACCGGGATTAGAGCGCATTAGTC

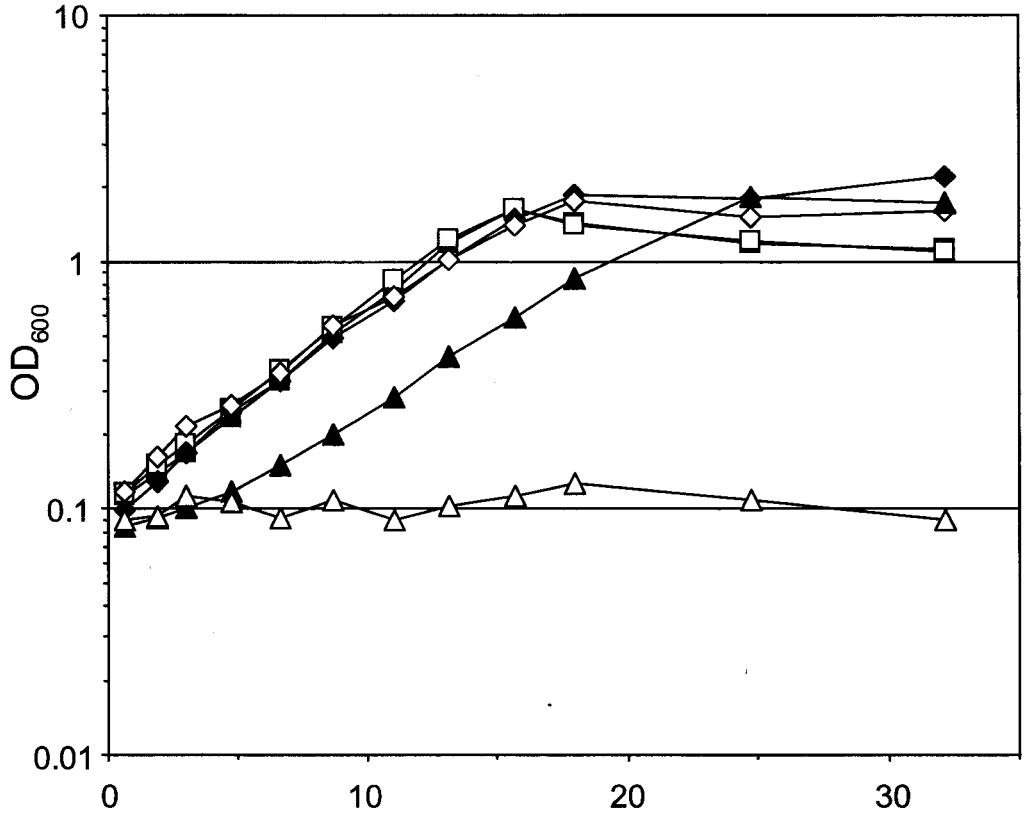


Figure 4.1 : Growth of *ftfL* mutant. Wild-type *M. extorquens* AM1 (filled symbols) and the *ftfL* mutant CM216K.1 (open symbols) pre-grown in succinate, harvested, and resuspended in media containing succinate (squares), succinate with methanol added to 125 mM at two hours (triangles), or methanol (circles).

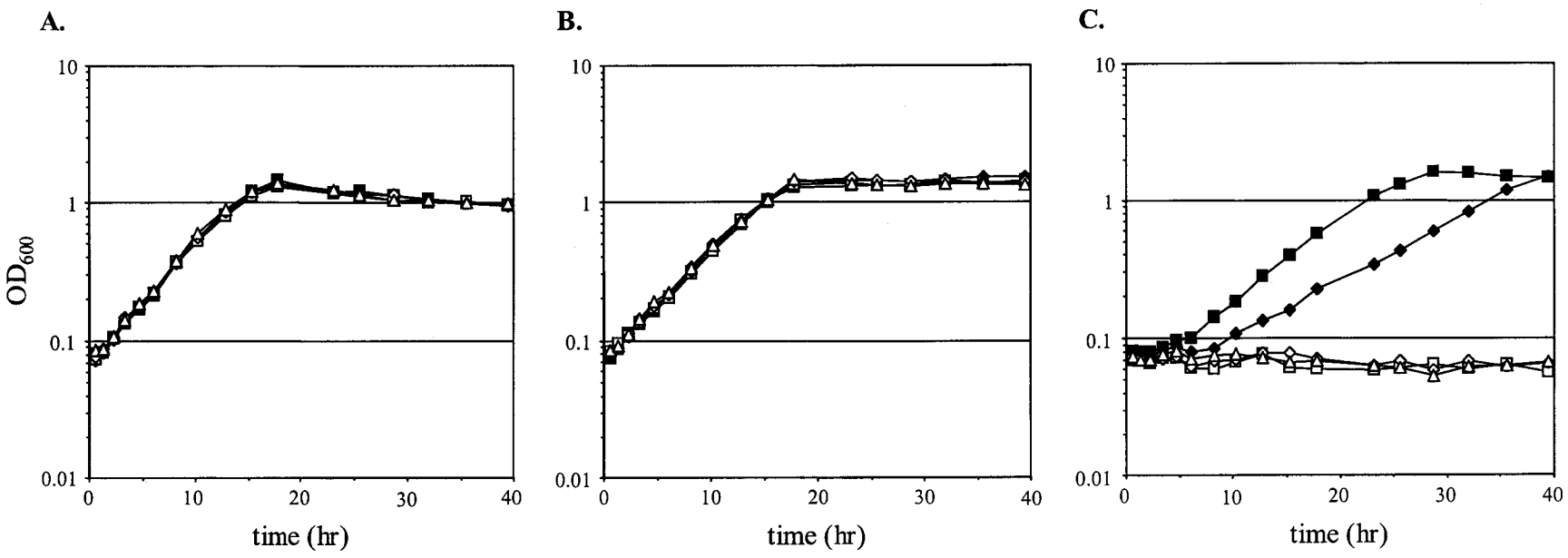


Figure 4.2: Growth of *mtdA* and *fch* mutants generated in a *folD*-expressing strain. Wild-type *M. extorquens* AM1 and mutant strains pre-grown on succinate, harvested, and resuspended in medium containing succinate (A.), succinate with methanol added to 125 mM at two hours (B), or methanol (C). The strains represented are wild-type (filled squares), the *folD*-expressing strain CM219.1 (filled diamonds), and three mutants generated in the CM219.1 background, the *mtdA* mutant CM219-275K.1 (open squares), the *fch* mutant CM219-279K.1 (open diamonds), and the *mtdA*, *fch* mutant CM219-280K.1 (open triangles).

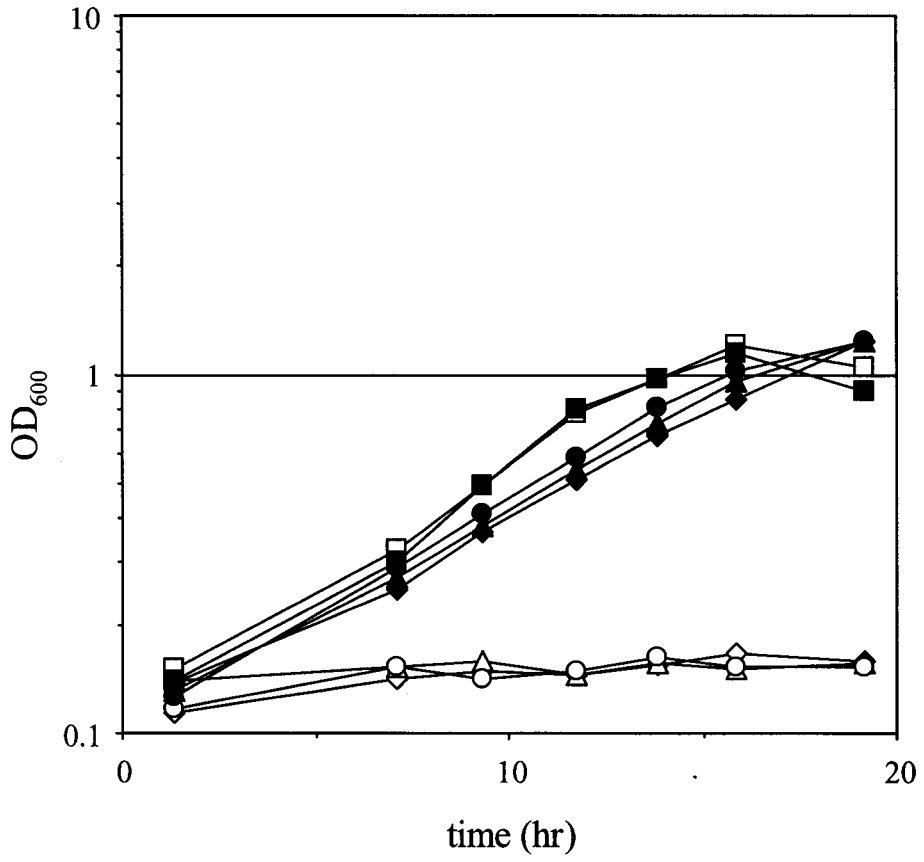


Figure 4.3 : Growth of *mtdA* and *fch* mutants on succinate with or without methylamine. Wild-type *M. extorquens* AM1 and mutant strains pre-grown on succinate with 7 mM methylamine, harvested, and resuspended in medium containing just succinate (open symbols) or succinate with 7 mM methylamine (filled symbols). The strains represented are wild-type (squares), the *mtdA* mutant CM275K.1 (diamonds), the *fch* mutant CM279K.1 (triangles), and the *mtdA*, *fch* mutant CM280K.1 (circles).

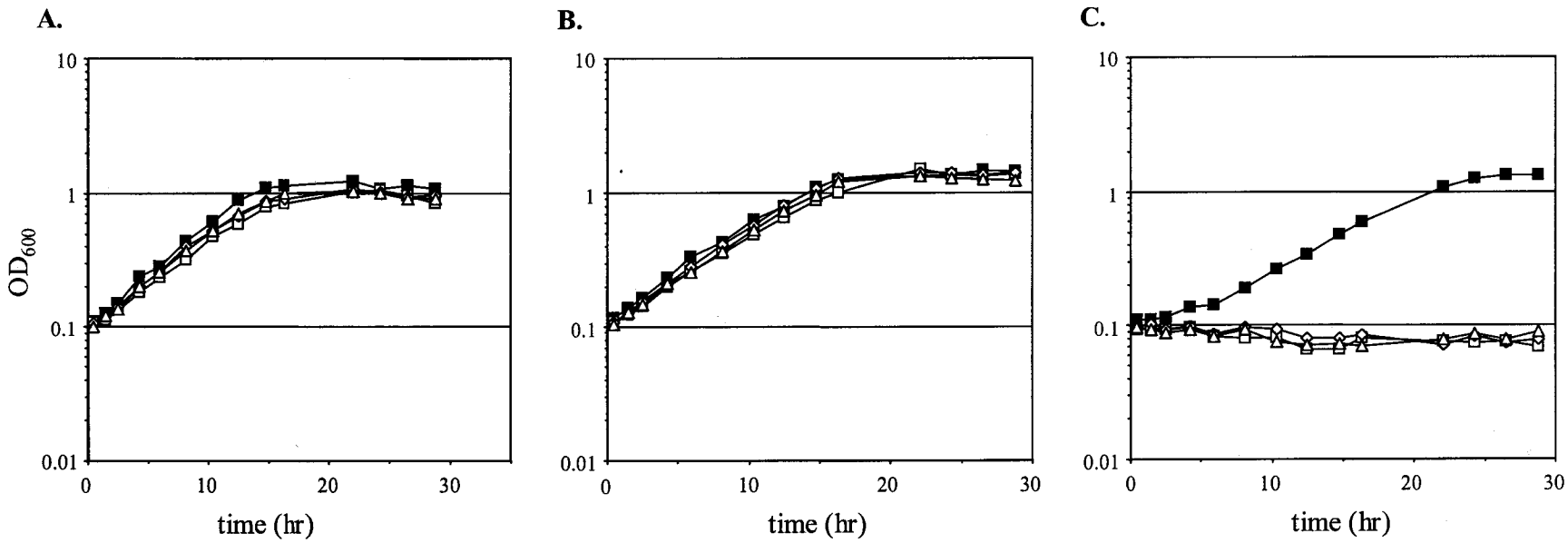


Figure 4.4 : Growth of *mtdA* and *fch* mutants on methanol or succinate with methanol. Wild-type *M. extorquens* AM1 and mutant strains pre-grown on succinate with 7 mM methylamine, harvested, and resuspended in medium containing succinate with 7 mM methylamine (A), succinate with methanol added to 125 mM at two hours (B), or methanol (C). The strains represented are wild-type (squares), the *mtdA* mutant CM275K.1 (diamonds), the *fch* mutant CM279K.1 (triangles), and the *mtdA*, *fch* mutant CM280K.1 (circles).

CHAPTER 5

Analytical Determination of the Flux of C_1 Compounds through the Methylotrophic Metabolism of *Methylobacterium extorquens* AM1

Abstract:

Genetic and physiological analyses of the H_4MPT- and H_4F-dependent C_1 transfer pathways of *M. extorquens* AM1 suggested that the former is required for formaldehyde oxidation and detoxification (Chapter 3), whereas the latter is required, but apparently not as a secondary formaldehyde oxidation pathway (Chapter 4). It was hypothesized that the role of the H_4F pathway in methylotrophy was to supply methylene-H_4F for assimilation by the serine cycle. Thus, C_1 units could enter the serine cycle through two routes, either through the direct condensation of formaldehyde with H_4F, or formaldehyde could be oxidized to formate by the H_4MPT pathway and then converted by the reductive H_4F pathway to methylene-H_4F. In this chapter I present a combination of analytical approaches that confirms this hypothesis. Carrying out labeling experiments with ^{14}C-methanol, I found that a H_4MPT pathway mutant is defective for dissimilation to CO_2, whereas a H_4F pathway mutant is not. A gas-chromatography-mass spectrometry method was then developed that allowed the deuterium labeling from CD_3OD to be followed into serine. Methylene-H_4F generated by direct condensation with formaldehyde contains two deuteriums that are donated to serine, whereas the indirect pathway through formate leads to labeling with a single deuterium. In succinate-grown cells the majority of methylene-H_4F is made from formate, whereas direct condensation dominates in methanol-grown cultures. Cultures shifted between these two substrates show a smooth transition between these two states. With samples of the same cultures, I determined the rate of methanol utilization, CO_2 production, and assimilation of C_1 units from methanol using ^{14}C-methanol labeling. This allowed a complete calculation of C_1 fluxes through the various branches of methylotrophy. This demonstrated that, although the proportion of methylene-H_4F made through the reductive H_4F pathway decreased during the transition to growth on methanol, the actual flux through the pathway increased. The rate of the direct condensation of H_4F with formaldehyde increased to an even greater degree, however, to over 100-fold the level in succinate-grown cells. These data therefore verify that the

H_4F pathway converts formate into methylene-H_4F for assimilation and provide a dynamic picture of flux through all branches of methylotrophy during the transition to and from methylotrophic growth.

Introduction:

The genetic and physiological characterization of mutants defective for either the H_4MPT- or H_4F-dependent C_1 transfer pathways clearly demonstrated that both pathways are required for methylotrophy (Chapters 3 and 4). The methanol-sensitivity of H_4MPT pathway mutants and rescue of these mutants by an alternative formaldehyde oxidation pathway suggested that this pathway is required for formaldehyde oxidation and detoxification. In contrast, H_4F pathway mutants are unable to grow on C_1 compounds but do not exhibit sensitivity to formaldehyde. Furthermore, double mutants lacking both pterin-linked pathways were no more sensitive to methanol than single mutants defective solely for the H_4MPT pathway. These data suggested that the requirement for the H_4F pathway is not simply as a second formaldehyde oxidation route. Rather, I hypothesized that the H_4F pathway functions in the reductive direction to assimilate formate produced from formaldehyde by the H_4MPT pathway, thereby converting it into methylene-H_4F. This would provide a second route for entry of C_1 units into the serine cycle in addition to the direct condensation of formaldehyde with H_4F.

In this chapter I present direct biochemical evidence for the roles of the two C_1 transfer pathways of *M. extorquens* AM1 in methylotrophy. First, the dissimilation of ^{14}C-methanol to ^{14}C-CO_2 was monitored in wild-type and mutants in order to determine if either the H_4MPT or H_4F pathways are required for formaldehyde dissimilation. Second, a gas chromatography-mass spectrometry (GC-MS) method was developed that could detect free serine, the first assimilatory intermediate of the serine cycle. With this method, I can follow the incorporation of deuteriums from fully deuterated methanol (CD_3OD) into serine. The ratio of serine containing two

deuteriums to those containing only one is dependent on whether methylene-H_4F was generated from the direct condensation of deuterated formaldehyde with H_4F, or from the proposed indirect pathway involving the conversion of formate to methylene-H_4F. The CD_3OD label tracing assay was verified by comparing mutants blocked in various branches of methylotrophy to wild-type. Subsequently, both the CD_3OD and ^{14}C-methanol tracing studies were performed on samples from cultures shifted between growth on succinate or methanol. This allowed the complete determination of flux through every branch of methylotrophy, thereby providing a dynamic picture of the response of *M. extorquens* AM1 during these transitions. The resulting data confirm that the H_4MPT pathway is the primary formaldehyde oxidation pathway and that the H_4F pathway is required to provide C_1 units for assimilation.

Materials and Methods:

Standard laboratory procedures. The growth of bacterial strains, genetic procedures, and recombinant DNA techniques are performed as described in Chapter 2. *M. extorquens* AM1 strains, plasmids, and primers used in this chapter are described in Table 5.1.

Generation of a Δ*glyA::kan* mutant. As a control for labeling studies, a mutant lacking the initial enzyme of the serine cycle, serine hydroxymethyltransferase (encoded by *glyA*) (21), was generated. The regions upstream and downstream of *glyA* were amplified by PCR using CM-glyAuf and CM-glyAur, and CM-glyAdf and CM-glyAdr. The resulting products were cloned into pCR2.1 (Invitrogen) to generate pCM236 and pCM237, respectively. The 0.6 kb *Bgl*II-*Nco*I fragment from pCM236 was introduced into the same sites of pCM184 to generate pCM238, into which the 0.6 kb *Apa*I-*Sac*I fragment of pCM237 was inserted into the same sites to generate pCM239. A Δ*glyA::kan* mutant, CM239K.1, was generated as described in Chapter 2.

Whole-cell ^{14}C-CO_2 production assay. In order to determine whether mutants defective for the H_4F and/or H_4MPT pathways were competent for the complete oxidation of methanol to CO_2, the rate of ^{14}C-CO_2 production from ^{14}C-methanol was determined using a variation of previous methods (6, 70). Assays were performed at room temperature with cultures of wild-type *M. extorquens* AM1 and appropriate mutants in three independent experiments. Consistent kinetics was observed in all cases. Cultures grown to mid-exponential phase on succinate were centrifuged, washed, and resuspended to 1.0 OD_{600}. The assays were initiated by the addition of methanol to a final concentration of 1 mM, with 3.3 μCi/μL ^{14}C-methanol (Sigma). Aliquots of 0.3 mL of the cell suspension containing methanol were then immediately dispensed into 2.0 mL autosampling vials (Kimble) and sealed with black phenolic screw caps tops (Kimble) and red PTFE-faced/white silicone septa (Kimble). Every five minutes 0.3 mL of 0.1 M NaOH was added with a syringe to a set of samples to stop growth and trap CO_2 as bicarbonate. The samples were equilibrated for one hour, then they were centrifuged to remove cell material, and 0.4 mL of the cell-free spent medium was placed into a tube in a 80°C heat block to allow complete evaporation to eliminate the ^{14}C-methanol. The samples were then resuspended in 0.4 mL distilled H_20 and transferred into 20 mL serum vials (Kimble). Truncated 1.7 mL eppendorf tubes containing 0.2 mL phenylethylamine were placed into the serum vials, and the vials were capped with one-piece aluminum seal crimp-tops (Kimble) and teflon-lined grey butyl septa (Wheaton). Bicarbonate was released as CO_2 through the addition of 0.2 - 0.3 mL of 0.3 M HCl to each of the stoppered vials by syringe. After one hour for equilibration to trap the CO_2 in the phenylethylamine, the vials were opened and the phenylethylamine was transferred into scintillation vials and assayed. Controls were performed with ^{14}C-labelled C_1 compounds to confirm >99% retention of bicarbonate, 85% loss of formate, and >99% loss of both methanol and formaldehyde. Given that formate does not accumulate in the cell medium to appreciable amounts (Laukel et al., unpublished results) and that formate production requires

formaldehyde oxidation, the minor retention of formate in this protocol does not significantly alter the results, nor their interpretation. The data represent nmol CO_2 mL^{-1} OD_{600}^{-1}.

CD_3OD labeling and preparation of ECF-TFAA derivatized metabolites. CD_3OD was added to cultures in order to label cell metabolites with deuterium for analysis by GC-MS. CD_3OD assays were performed at room temperature with cultures of wild-type *M. extorquens* AM1 and appropriate mutants in three independent experiments. Cultures grown to mid-exponential phase on succinate or methanol were centrifuged, washed, and resuspended to 2.5 mL of 6.0 OD_{600}. The assays were initiated by the addition of 99.8% CD_3OD (Cambridge Isotope Laboratories) to a final concentration of 1 mM. The cultures were then shaken at room temperature for 20 seconds before 2.0 mL of the suspension was added to 6.0 mL of boiling 100% ethanol (90°C) for lysis. After three minutes at 90°C, the mixture was placed at 50°C for 10 minutes. This mixture was then centrifuged for two minutes to remove the components insoluble in 75% ethanol. The supernatant was retained and dried at 90°C. The dried material was resuspended into 200 μL distilled H_2O, which was then centrifuged for two minutes to remove the H_20-insoluble fraction. Thin-layer chromatography and size-exclusion centrifugation were attempted to partially purify the serine present in the H_2O-soluble fraction, but the best results were achieved by simply derivatizing the H_2O-soluble fraction without further treatment (data not shown).

The H_2O-soluble small molecule fraction was then derivatized in a manner similar to that described previously (25, 50). First, 20 μL of 0.25 M HCl was added to the water-soluble fraction, after which 150 μL of pyridine:ethanol (1:4) was added. The mixture was derivatized by addition of 20 μL of ECF and incubated until all CO_2 evolution had ceased (about 30 minutes). This solution was then extracted with 300 μL dichloromethane. Finally, 20 μL of TFAA was added and the solution was vented in a hood to concentrate three- to five-fold.

Gas chromatography-mass spectrometry. Experiments were performed using a Agilent 6890 gas chromatograph/Agilent 5973 quadrupole mass selective detector (electron impact ionization) operated at 70 eV equipped with an Agilent 7683 autosampler/injector (Hewlett-Packard). The column utilized was a HP-5MS. The MS was operated in selected ion monitoring (SIM) mode to detect M/z = 175/176/177 after a 5.4 minute solvent delay, then switching to 156/157/158/228/229/230 from 7 minutes to the end of the method. The oven temperature started at an initial temperature of 60°C, and ramping at 20°C min^{-1} to 130°C, 4°C min^{-1} to 155°C, and then 120°C min^{-1} to a final temperature of 300°C that was held for 5 minutes. Flow through the column was held constant at 1 mL min^{-1}. The injection volume was 1 μL and the machine was run in split-less mode. The temperature of the inlet was 230°C, the interface temperature was 270°C, and the quadrupole temperature was 150°C.

The GC-MS data was analyzed using Agilent Enhanced ChemStation G1701CA (Hewlett-Packard). The two mass clusters for serine, M/z = 156/157/158, and 228/229/230, represent fragments of ECF-TFAA derivatized serine ($C_{10}H_{14}O_6NF_3$) that have lost one or both of the carboxyl ethyl esters and have the molecular formulas $C_4H_5O_2NF_3$ and $C_7H_9O_4NF_3$, respectively. Increased mass fragments from glycine were also analyzed. The glycine mass cluster used for this analysis was 175/176/177, which represents analogous fragments as above with the molecular formula of $C_7H_{13}O_4N$. The data were corrected for the natural abundance of heavy isotopes in the derivatized serine fragments using proportions calculated with the Isoform 1.02 program (NIST). For each sample, the ratio of $\Delta(+1)/\Delta(+2)$ was calculated for both mass clusters and averaged. The mean and standard error for these data were then calculated for the three replicates of each experiment.

Labeling experiments with ^{14}C-methanol and CD_3OD with cultures shifted between growth on succinate and methanol. In order to experimentally determine the fluxes through the various branches of methylotrophy as cultures transition from growth on methanol to growth on succinate, or vise versa, simultaneous labeling experiments were performed with ^{14}C-methanol and CD_3OD. Mid-exponential phase cultures of wild-type *M. extorquens* AM1 that had been grown on either methanol or succinate were centrifuged, washed, and resuspended into medium containing the other substrate. The growth characteristics were monitored at OD_{600} (Beckmann DU 640B spectrophotometer) every hour and are presented as an average with error bars depicting the standard error. Samples were obtained at five time points for each transition. The first sample for each was taken one hour prior to the transition to the other substrate, thereby representing the flux characteristics of cultures grown on that substrate. For methanol to succinate, samples were then taken 1, 3, 5, and 7 hours after the transition. For succinate to methanol, which results in a longer lag-time, samples were taken 1, 5, 7.5, and 10 hours after the transition. The CD_3OD labeling was performed as described above. The ^{14}C-methanol labeling was performed as above with slight modifications. NaOH was added at 2.5, 5, 7.5, and 10 minutes. Of the resulting 0.6 mL of cell suspension, 0.2 mL was used for detection of ^{14}C-CO_2, and 0.3 mL was filtered (0.2 μm PVDF, Millipore) to obtain the cell material, the filter was washed with distilled H_2O, and placed directly into scintillation fluid. The entire experiment was repeated on three separate occasions. All measured and calculated fluxes were determined using the data from each replicate and then analyzed to determine the mean and standard error for each flux.

Results

A H_4F pathway mutant generates ^{14}C-CO_2 from ^{14}C-methanol with wild-type rates whereas a H_4MPT pathway mutant shows a reduced capacity. Genetic and physiological data

presented in Chapters 3 and 4 suggested that the H_4MPT pathway is responsible for formaldehyde oxidation, whereas the H_4F pathway does not contribute to this process. Production of ^{14}C-CO_2 from ^{14}C-methanol was determined as a proxy for *in vivo* formaldehyde oxidation. For a control, the Δ*mxaF::kan* strain CM194K.1 (Chapter 2) was analyzed and found to produce no detectable ^{14}C-CO_2 (Fig. 5.1), consistent with the inability to generate formaldehyde from methanol. The *ftfL* mutant CM216K.1 produced CO_2 at a rate similar to the wild-type, however, there was a significant lag of 10-15 minutes for the *dmrA* mutant CM212K.1 before CO_2 was produced. In order to determine whether the H_4F pathway was responsible for the residual CO_2 production that occurred after a time lag in CM212K.1, the strain CM216-212K.1, mutated in *ftfL* and *dmrA*, was investigated. This strain exhibited similar CO_2 production to CM212K.1. These data do not support a significant role for the H_4F pathway in formaldehyde oxidation, even in the absence of the H_4MPT pathway. The likely source(s) of the remaining formaldehyde oxidation capacity is discussed below.

Detection of serine using GC-MS. A CD_3OD label tracing strategy was devised to directly determine what fraction of the methylene-H_4F that enters the serine cycle was formed from the direct condensation of formaldehyde and H_4F ("direct pathway"), or through the proposed pathway involving oxidation of formaldehyde by the H_4MPT pathway to formate, followed by assimilation through the H_4F pathway ("long pathway", Fig. 5.2). Oxidation of CD_3OD results in the production of formaldehyde with two deuteriums (CD_2O), neither of which originates from the hydroxyl deuterium of methanol that can exchange with the bulk solvent, H_2O. Direct condensation of CD_2O with H_4F results in CD_2=H_4F. Alternatively, CD_2O oxidized by the H_4MPT pathway to formate, and then assimilated through the H_4F pathway to methylene-H_4F loses one deuterium, thereby producing CDH=H_4F. Since both hydrogens from CH_2=H_4F are incorporated into serine, labeling with CD_3OD would result in serine containing two or one

deuteriums. Therefore, the ratio of serine with one or two deuteriums provides an assay for whether the C_1 unit incorporated formed methylene-H_4F from the proposed long pathway involving conversion of formate by the reductive H_4F pathway, or the direct condensation with formaldehyde.

In order for this label tracing method to be successful, the ratio of serine isotopomers containing one or two deuteriums from CD_3OD must be detected. For preliminary analysis, standard methanol (CH_3OH) was added to cell suspensions of wild-type *M. extorquens* AM1 to a concentration of 1 mM and incubated at room temperature for 10 to 30 seconds. The suspensions were then directly added to three volumes of boiling ethanol for lysis. This mixture was centrifuged to remove the components that were insoluble in 75% ethanol, dried, resuspended in distilled H_2O, and centrifuged a second time to obtain a H_2O-soluble small molecule preparation. Efforts to further simplify this mixture by either size-exclusion filtration or by thin-layer chromatography resulted in significant loss of the desired analytes (data not shown). The H_2O-soluble preparations were then derivatized with ECF-TFAA (25, 50) for analysis by GC-MS. These samples, as well as a derivatized serine standard, were analyzed by GC-MS as described in Materials and Methods. Consistent with a derivatized serine standard and previous work (25, 50), a small peak was observed at ~8.6 minutes that contained two major ions with M/z of 156 and 228 (Fig. 5.3). The proportion of (+1) and (+2) M/z ions detected (Table 5.2) were within 1.1 ± 1.7% and -0.7 ± 0.5% of the predicted distribution (Isoform 1.02, NIST) of heavy isotopomers for these fragments. These data indicated that serine isotopomers could be detected with this GC-MS method.

Incorporation of deuteriums into serine from CD_3OD establishes that the H_4F pathway produces methylene-H_4F from formate. As described above, labeling cell cultures with CD_3OD will result in serine with an increased mass of 1 or 2 M/z. The incorporation of

deuteriums into serine was first investigated with cell-suspensions of wild-type *M. extorquens* AM1 harvested from succinate-grown cultures. CD_3OD (99.8%) was added and the cell suspensions were incubated with shaking for 20 seconds before addition to the boiling ethanol. Analysis of the derivatized H_20-soluble small molecule preparation from these samples indicated over 35% of the isotopomers had an increased mass (Table 5.2). After correcting for the predicted distribution of naturally occurring heavy isotopes, the ratio of serine containing one deuterium to those with two deuteriums was 8.0 ± 0.6 (Table 5.2).

Two classes of mutants were analyzed to determine whether the (+1) ions represented serine that had been synthesized from methylene-H_4F via the proposed "long" pathway (H_4F pathway-dependent assimilation of formate produced by the H_4MPT pathway-dependent oxidation of CD_2O formaldehyde). The first mutant class included the *glyA* mutant CM239K.1 that lacks the initial serine cycle enzyme, serine hydroxymethyltransferase and is therefore completely blocked in the serine cycle. As predicted, no significant increases in the (+1) or (+2) fragments occurred (1.1 ± 0.7% and 0.2 ± 0.5%, respectively), indicating no serine was produced from methanol by either route. This confirms that the increased masses for the serine fragments are dependent upon the serine cycle. The second mutant class analyzed contained lesions in the proposed long pathway for methylene-H_4F formation. These were the *ftfL* mutant CM216K.1, blocked for the H_4F pathway, and the *dmrA* (encodes a putative dihydromethanopterin reductase) mutant CM212K.1 (Chapter 3), which has been shown to lack H_4MPT (S. Wyles and M. E. Rasche, personal communication), and as described earlier, is defective for conversion of ^{14}C-methanol to ^{14}C-CO_2. Consistent with their proposed role, these mutants resulted in a negligible increase in the (+1) fragments occurred (1.6 ± 1.4%, compared to 29.2 ± 7.6% for wild-type), but an increase in the (+2) fragments remained (6.7 ± 2.8%, compared to 3.8 ± 1.1% for wild-type) (Table 5.2). These data indicate that both the H_4F and H_4MPT pathways affect labeling of serine, and are required to generate the large increase in (+1) isotopomers seen with wild-type. These

data also indicate that exchange reactions that could eliminate the deuteriums, which might occur with MtdA (NADP-dependent methylene-H_4F dehydrogenase), do not account for the generation of the (+1) ions. Additionally, given the multiple enzymatic steps involved in generating methylene-H_4F from formate versus the single nonenzymatic reaction of formaldehyde with H_4F, any potential isotope discrimination effects against deuterated compounds would, if anything, cause an underestimation of the contribution of the H_4F pathway in supplying C_1 units for the serine cycle. However, the large enrichment of CD_3OD (99.8%) would minimize any isotope discrimination.

As a final test that fragments containing a single deuterium represent incorporation via the proposed long pathway of methylene-H_4F generation, the mass distribution of glycine was examined. Through the action of the serine cycle, some fraction of the deuteriums incorporated into serine could end up in glyoxylate and then glycine, which would skew the interpretation of the observed mass distributions (Fig. 1.3). The glyoxylate generated directly from malyl-CoA (Fig. 1.3) results from the part of the molecule without any potential deuteriums. Deuterium(s) present in malyl-CoA remain in the acetyl-CoA that is generated during cleavage. Given that this acetyl-CoA molecule undergoes a significant number of transformations through the.glyoxylate regeneration pathway (62), some of which remain uncharacterized, it is predicted that only a small fraction could potentially remain to provide a glycine with a single deuterium on the α-carbon. The mass fragments from ECF-TFAA derivatized glycine (retention time of ~6.5 minutes) present in the samples were also analyzed in a manner analogous to that described for serine. In cultures grown on either methanol or succinate, a trivial amount of glycine was found to have an increased mass indicating a deuterium label (1.6 ± 0.9%), so this can not account for the occurrence of (+1) serine fragments (29.2 ± 7.6%). Collectively, these data indicate that the (+1) and (+2) serine mass fragments can serve as an accurate proxy for methylene-H_4F generated through the long (through formate) or direct (from direct condensation with formaldehyde)

pathways. Furthermore, they indicate that in cultures pre-grown in succinate-containing medium, the long pathway through formate dominates.

Direct condensation is the dominant methylene-H_4F generating pathway during growth on methanol. Cell-suspensions of wild-type *M. extorquens* AM1 harvested from cultures grown in medium containing methanol were used to assess the proportion of methylene-H_4F made throught the long or direct pathways during methylotrophic growth. Under these conditions, there was a significant increase in both elevated mass fragments, but the majority were present in the (+2) forms (Table 5.2). Incubations with both methanol- and succinate-grown cell suspensions were performed for times varying from 10 seconds to 1 minute and over this period no significant changes in the ratios were detected (data not shown), so 20 seconds was chosen as a standard incubation period. Correcting for the abundance of naturally occurring heavy isotopes, the (+1)/(+2) ratio was found to be .065 ± .006, equivalent to a (+2)/(+1) ratio of 15.7 ± 1.4 (Table 5.2). Therefore, although both pathways for methylene-H_4F production operate in cultures grown on either substrate, the relative contribution shifts over 100-fold from primarily production from the long pathway in succinate-grown cultures to dominance of the direct pathway in methanol-grown cells.

Change of the ratio of indirect/direct formation of methylene-H_4F, and the rate of methanol utilization, CO_2 formation, and assimilation into cell material during shifts between methanol and succinate. From the two order of magnitude change in the ratio of the formation of methylene-H_4F from the long versus direct pathways, it would appear that the long pathway operates in succinate-grown cultures, and is shut down in favor of activation of the direct pathway during growth on methanol. The direct pathway, condensation of formaldehyde and H_4F, occurs spontaneously (127), thus the rate of this reaction will depend solely on the relative

concentrations of products and reactants. The long pathway depends upon the action of both the H_4MPT and H_4F pathways. Both of these pterin-dependent pathways are induced three- to four-fold during growth on methanol (22, 43, 67, 77, 97, 125, 127). The induction of the two branches of the long pathway during growth on methanol seems contradictory with the 100-fold decrease in the relative contribution of the long pathway during growth on methanol. Furthermore, the fundamental difference in the relative contribution of the two methylene-H_4F formation pathways between succinate- and methanol-grown cells suggested a dynamic balance that would shift during transition of cultures between growth on C_1 and multi-carbon compounds. In order to understand the dynamics of the utilization of the long and direct pathways for methylene-H_4F formation, and ultimately to determine flux of C_1 units through the different branches of methylotrophic metabolism, further label tracing experiments were performed throughout the shift of *M. extorquens* AM1 to and from methylotrophic growth.

Mid-exponential phase cultures of wild-type *M. extorquens* AM1 grown in either succinate- or methanol-containing medium were harvested, washed, and resuspended into medium containing either methanol, or succinate, respectively. Both transitions resulted in a growth lag, but the transition from methanol to succinate occurred more quickly than the switch from succinate to methanol (Fig. 5.4). In both cases, cultures were sampled one hour prior to the transition in order to establish the "steady-state" flux distributions on each substrate. Following the switch in growth substrates, samples were taken at various intervals throughout the growth curve. For each sample, both CD_3OD label tracing to determine the ratio of the methylene-H_4F formation pathways and ^{14}C-methanol labeling to determine the rate of methanol utilization, CO_2 formation from methanol, and assimilation of C_1 units from methanol into biomass were performed. The ratio of the contribution of the long pathway for methylene-H_4F formation to the direct pathway varied in a continuous fashion during the transition from succinate to methanol, or methanol to succinate (Fig. 5.5). Similarly, the rate of methanol utilization, CO_2 production, and

assimilation of methanol-derived C_1 units into biomass varied significantly during these transitions (Fig. 5.6). The total methanol utilization for methanol-grown cells was ten-fold higher than for succinate grown cells, and the percentage of carbon from methanol assimilated into biomass was approximately three-fold higher.

Dynamics of C_1 fluxes during transitions between succinate and methanol. The data generated above allowed the C_1 fluxes of every branch of methylotrophy in *M. extorquens* AM1 to be determined. The other values incorporated into the calculated fluxes are the stoichiometry of the serine cycle, in which two C_1 units from methylene-H_4F and one CO_2 are incorporated for every C_3 compound assimilated, and the amount of CO_2 incorporated by the serine cycle that originated from the oxidation of methanol (63.3% of the total CO_2 incorporated), which had been determined previously (Van Dien et al., *in press*). Since this value was not re-determined in my experiments, the sensitivity of the calculated fluxes to a 2-fold increase or decrease in the determined ratio of 1.73:1.00 internal:external CO_2 incorporated into the serine cycle was examined. Besides the direct effect on relative fluxes of internal and external CO_2 into the serine cycle, the calculated incorporation of C_1 units from methylene-H_4F would vary no more than 7%, which would be balanced by a change in the dissimilatory flux through the H_4MPT pathway and formate dehydrogenase of less than 6% (data not shown). Therefore, deviations in the ratio of methanol-derived and external CO_2 incorporation from the reported work (Van Dien et al., *in press*) should not greatly affect the determined fluxes.

The ten C_1 fluxes (arbitrarily labeled "A" through "J") calculated using the CD_3OD and ^{14}C-methanol are presented in Table 5.3 and are presented graphically in Figures 5.7 and 5.8. These data indicate that, although the relative contribution of the long pathway (flux E) to methylene-H_4F formation decreases (Fig. 5.5), the flux through the long pathway increased significantly during the transition to growth on methanol (Fig. 5.7). The flux through this

pathway (flux E) peaked five hours after the transition to methanol, when the flux reached a value 8-fold higher than succinate grown cells, and dropped somewhat afterwards. The flux through the direct pathway (flux F) also increased during the transition to methanol growth, but increased to a far greater degree (150-fold) and continued to increase throughout the transition to growth on methanol (Fig. 5.7). It is also notable that the majority of formaldehyde oxidation occurred through the action of the H_4MPT pathway (flux B), as previously suggested (Chapter 3). This percentage (flux A/flux B) was 67.7 ± 1.2% for methanol-grown cells and 98.6 ± 0.1% in succinate-grown cells. The C_1 fluxes during the transition from methanol to succinate were not a precise reversal of the transition from succinate to methanol (Fig. 5.8). A sharp decrease in methanol oxidation (flux A) during the first hour was followed by a relatively slow decrease. The decrease in assimilation (flux J) initially did not fall as fast as the methanol oxidation rate, which led to percentages of total carbon assimilated to methanol utilized greater than 60%. The direct pathway for methylene-H_4F formation (flux F) also fell substantially in a relatively continuous fashion, whereas the flux through the long pathway (flux E) actually rose three-fold during the first five hours after the transition to succinate before settling to a rate 1.5-fold higher than the initial methanol-grown cells. Finally, it was apparent that the flux pattern of cultures 7 or 10 hours after the transition to the new substrate did not match the precise flux distribution of the original cultures grown on that substrate. This suggests either an incomplete induction of systems required for growth on the new substrate and/or incomplete degradation of enzymes utilized before the transition.

Attempts to eliminate FtfL activity during growth on C_1 compounds. The combination of CD_3OD and ^{14}C-methanol label tracing studies clearly demonstrate that the long pathway contributes methylene-H_4F to the serine cycle, and that the flux through this pathway (flux E) increases significantly during the transition to growth on methanol. Cells growing on

methanol, however, generate 93.9 ± 0.5% of methylene-H_4F from the direct pathway (flux F). It therefore seemed possible that the minor contribution of the H_4F pathway during growth on C_1 compounds is not strictly required, such that the inability of H_4F pathway mutants to grow on C_1 compounds may only be due to the role of the long pathway in the transition from multi-carbon to methylotrophic metabolism. Given that no regulated expression systems are available for use in *M. extorquens* AM1, it is not possible to simply examine the growth of an *ftfL* mutant with a complementing plasmid on C_1 compounds after removal of inducer (or addition of a repressor) to prevent expression of *ftfL*. Therefore, attempts were made to establish *ftfL* mutants during growth on C_1 compounds. First, I attempted to obtain an *ftfL* null mutant by allelic exchange during growth of wild-type on methanol or methylamine. As expected, null mutants were obtained on succinate medium, but no null mutants were obtained after screening over 100 colonies on medium containing methanol or methylamine. Second, cultures of the Δ*ftfL::kan* mutant CM216K.1 bearing the complementing plasmid pCM218 (Chapter 4) were grown in medium containing methanol without tetracycline for plasmid maintenance. No plasmid-free isolates were obtained for CM216K.1 with pCM218 during growth on methanol despite screening over 100 isolates, however, they were obtained for wild-type with pCM218 in methanol, or CM216K.1 with pCM218 grown in succinate. Therefore, using the tools currently available, it appears that the H_4F pathway plays an essential role in methylotrophy even after cells have already begun to grow on C_1 compounds.

Discussion:

I have presented several label tracing studies that clearly and independently verify the role of the H_4MPT and H_4F pathways suggested from my earlier work (Chapters 3 and 4). The aim of the first set of experiments was to use the conversion of ^{14}C-methanol to ^{14}C-CO_2 as a proxy for *in vivo* formaldehyde oxidation. Although this approach is less direct than monitoring

the conversion of formaldehyde to formate directly, it is technically feasible and more physiologically relevant. First, CO_2 production is simpler to analyze than formate because the pH of the medium can be manipulated to either trap it as bicarbonate or release it as a free gas as needed. Also, three active formate dehydrogenases are present in *M. extorquens* AM1 (68)(L. Chistoserdova, unpublished), therefore, in order to attempt to analyze formate directly in the ^{14}C assay, mutants defective for one or both of the pterin-linked C_1 transfer pathways would have to have been generated in an unmarked triple mutant strain defective for all three formate dehydrogenases. Finally, wild-type *M. extorquens* AM1 is sensitive to modest levels of external formaldehyde (MIC of 1 mM) and cannot grow on medium containing formaldehyde as a carbon and energy source despite the fact that growth on methanol or methylamine necessitates a high flux of internally produced formaldehyde. Therefore, the conversion of ^{14}C-methanol to ^{14}C-CO_2 was assayed in mutant strains blocked for one, or both, of the pterin-linked pathways in order to identify defects in the dissimilation of C_1 compounds. The significant delay in CO_2 production observed for the *dmrA* mutant represents the first biochemical demonstration that the H_4MPT pathway is required for formaldehyde dissimilation. In contrast, the *ftfL* mutant strain exhibited wild-type conversion of methanol to CO_2 and the kinetics of dissimilation by the *ftfL*, *dmrA* double mutant was very similar to the *dmrA* single mutant. This provides clear evidence that the H_4F pathway does not contribute in a measurable way to formaldehyde oxidation, even in the absence of an operational H_4MPT pathway.

The likely source(s) of the remaining formaldehyde oxidation capacity in the mutant blocked for both the H_4F and H_4MPT pathways are general aldehyde dehydrogenases present in *M. extorquens* AM1 (53, 129). These are neither specific for formaldehyde nor induced during growth on methanol and in the one case in which an enzyme was purified it had a high K_m for formaldehyde (53), suggesting that these enzymes are not specific for methylotrophy. As stated before, it has been calculated that the intracellular concentration of this toxic intermediate would

rise to 100 mM in less than a minute if methanol oxidation proceeded in the absence of subsequent formaldehyde oxidation (5, 127). It is possible, therefore, that the lag observed in CO_2 production by strains lacking the H_4MPT pathway corresponds to the time required for the intracellular formaldehyde to rise to a sufficiently high concentration to allow the low-affinity general aldehyde dehydrogenases to function in a measurable way. A variant of this explanation is that formaldehyde leaks from the cells and enters the periplasm where it can be oxidized by MDH (2). Further work is required to test these hypotheses.

A larger challenge was to develop a label tracing method that would allow me to test my hypothesis that the role of the H_4F pathway during growth on C_1 compounds is to supply methylene-H_4F from formate. Preliminary efforts examining labeled metabolites via thin-layer chromatography following the addition of ^{14}C-methanol suggested that less was incorporated in H_4F or H_4MPT pathway mutants than wild-type, but the results were inconclusive. Because labeling on the C_1 carbon cannot distinguish between the two pathways (long and direct) for methylene-H_4F formation, this method could only be used to compare mutants to wild-type. Therefore, the relative contribution of these two pathways could not be determined in wild-type. I surmised that, although the carbon from formaldehyde will end up in serine via either route, the hydrogens have a different fate depending on the route from which the methylene-H_4F was formed. Therefore, a GC-MS method was developed that allows labeling with CD_3OD to trace which methylene-H_4F formation pathway was used by monitoring the ratio of increased mass isotopomers of serine.

Under all conditions, wild-type *M. extorquens* AM1 generates methylene-H_4F from each of the two pathways, however, the ratio changes dramatically. Succinate-grown cultures generated 89.0 ± 0.4% of methylene-H_4F from the long pathway involving assimilation of formate by the H_4F-dependent enzymes. These results confirmed the hypothesis developed from genetic and physiological data regarding the role of the H_4F pathway during methylotrophy. The

utilization of the two pathways was found to be substantially different in methanol-grown cells, where the direct pathway of formaldehyde condensing with H_4F accounted for 93.9 ± 0.5% of the methylene-H_4F production. The ratio of the utilization of the two pathways changed in a continuous manner, but it would have been misleading to have therefore concluded that the flux through the H_4F pathway decreased as the flux through the direct condensation rose. Rather, by coupling the CD_3OD labeling to ^{14}C-methanol tracing a complete picture of C_1 fluxes during transitions to and from methylotrophic growth was determined. This analysis indicated that, in fact, flux through the long pathway increased despite a decrease in the proportion of the total methylene-H_4F formed via this route. This is consistent with the induction of H_4F-dependent enzyme activities during growth on methanol (22, 77, 97, 125).

The flux through the long pathway rises significantly early in the transition from growth on succinate to growth on methanol, but contributes only a minor fraction (6.1 ± 0.5%) of the methylene-H_4F during growth on methanol. It seemed possible that the requirement for the H_4F pathway on C_1 compounds may only involve the transition from multi-carbon growth to methylotrophy rather than for methylotrophic growth itself. Two approaches were taken to obtain mutants lacking *ftfL* during growth on C_1 compounds, however, neither was successful, suggesting that this pathway is indispensable during growth on C_1 compounds. In addition to supplying methylene-H_4F for assimilation, the H_4F pathway is also responsible for generating the compound that induces synthesis of serine cycle enzymes. Dr. Kalyuzhnaya in our laboratory recently discovered that the inducer of QscR, a LysR-family positive regulator of most serine cycle genes (56) that was discovered in the IS*phoA*/hah-Tc mutagenesis (Chapter 2), is formyl-H_4F (M. G. Kalyuzhnaya, unpublished). Therefore, the H_4F pathway may be required during methylotrophy to maintain a sufficient pool of formyl H_4F such that the serine cycle remains induced. Interestingly, one of the operons regulated by QscR (56) contains genes encoding serine cycle enzymes as well as two genes encoding enzymes of the H_4F pathway enzymes, *mtdA* and

fch. This genetic organization makes sense in light of the work presented here; Fch and MtdA supply methylene-H_4F for assimilation through the serine cycle from formyl-H_4F, which is generated from formate through the action of FtfL. Thus, the demonstration of the role of the H_4F pathway in supplying methylene-H_4F for assimilation allows a clear interpretation of the recent findings regarding the regulation of methylotrophy.

Condensation of formaldehyde with H_4F to form methylene-H_4F occurs non-enzymatically, and as such cannot be controlled directly through regulation of enzyme synthesis of activity. As with any chemical reaction, the rate will be determined by the relative concentrations of reactants and products. It is known that the equilibrium constant for the condensation is 3.2×10^{-4} (55). According to Le Chatelier's principle, flux will only occur if either the concentrations of the reactants formaldehyde and/or H_4F rise above the equilibrium concentration, or utilization of methylene-H_4F is sufficient to keep the pool of this metabolite below the equilibrium concentration. At this time, it is not technically feasible to measure the intracellular concentrations of free formaldehyde, H_4F, nor the methylene derivative of H_4F. Changes in H_4F concentration are unlikely to be large enough to contribute significantly to the two order of magnitude rate increase. Formaldehyde concentration may rise somewhat, but the high activity of Fae (1.4 μmol min^{-1} mg^{-1})(127) and the subsequent H_4MPT-dependent enzymes likely maintain a formaldehyde pool at levels at or below the Fae K_m for formaldehyde of 0.2 mM (127). Consistent with this suggestion, the flux through the H_4MPT pathway was found to increase in a linear fashion and does not plateau over the entire physiological range of methanol oxidation rates (Fig 5.10). The most likely explanation for high flux through the non-enzymatic condensation of formaldehyde and H_4F would be draw-off of the product (methylene-H_4F) by the serine cycle. A prediction of this hypothesis is that the direct condensation of formaldehyde with H_4F would increase with increasing flux through the serine cycle. While the flux through the H_4F pathway-dependent methylene-H_4F production pathway plateaued with increasing flux into the

serine cycle, direct condensation continued to increase (Fig. 5.10). This result is consistent with control of the flux of direct condensation of formaldehyde with H_4F by the flux through the serine cycle. This provides a shunt for formaldehyde to directly enter assimilatory metabolism when the assimilatory flux is high.

Table 5.1: *M. extorquens* AM1 strains, plasmids, and primers described in Chapter 5.

Strain	Description	Source
CM194K.1	Δ*mxaF::kan*	Chapter 2
CM212K.1	Δ*dmrA::kan*	Chapter 3
CM216K.1	Δ*ftfL::kan*	Chapter 4
CM216-212K.1	Δ*ftfL*, Δ*dmrA::kan*	Chapter 4
CM239K.1	Δ*glyA::kan*	This chapter
M. extorquens AM1	Rif^R derivative	(89)

Strain	Description	Source
pCM184	Broad-host-range allelic exchange vector	Chapter 2
pCM216	Donor to generate Δ*ftfL::kan* mutation	Chapter 4
pCM218	pCM62 with *ftfL* region	Chapter 4
pCM236	pCR2.1 with *glyA* upstream flank	This chapter
pCM237	pCR2.1 with *glyA* downstream flank	This chapter
pCM238	pCM184 with *glyA* upstream flank	This chapter
pCM239	pCM238 with *glyA* downstream flank	This chapter
pCR2.1	PCR cloning vector	Invitrogen

Primer	Sequence (5′-3′)
CM-glyAdf	AGGGCCCGATCTACGCCTGAG
CM-glyAdr	AGAGCTCGAAGGCCCATTTGAGAC
CM-glyAuf	ACTGCAGATCTCGCCTCAACGATTTCGTTG
CM-glyAur	AGGTACCATGGCGCTCATGAGCGTAG

Table 5.2: Serine isotopomer data and calculations. The observed percentage of each isotopomer and the adjusted numbers representing the number of deuteriums (D) present in the analytes for the predicted distribution (Isoform 1.02, NIST), succinate-grown wild-type (wt) labeled with standard methanol (CH_3OH) or with CD_3OD, and methanol-grown wild-type labeled with CD_3OD. The ratio shown represents the average of the 1D/2D ratio from individual experiments.

	Observed%			**Adjusted%**			**Ratio**
Sample	**+0**	**+1**	**+2**	**0**	**1D**	**2D**	**1D/2D**
Predicted	**94.5**	**5.0**	**0.5**	**100**	**0**	**0**	
wt (S) – CH_3OH	**92.7** ± 1.8	**6.8** ± 1.8	**0.5** ± 0.3	**99.5** ± 1.7	**1.1** ± 1.7	**-0.7** ± 0.5	
wt (S)	**62.7** ± 8.1	**31.4** ± 6.6	**5.8** ± 1.5	**66.9** ± 8.7	**29.2** ± 7.6	**3.8** ± 1.1	**8.0** ± 0.6
CM239K.1	**92.0** ± 1.1	**6.7** ± 1.0	**1.3** ± 0.3	**98.7** ± 1.1	**1.1** ± 0.7	**0.2** ± 0.5	
CM212K.1/CM216K.1	**85.9** ± 1.5	**6.8** ± 1.4	**7.4** ± 2.6	**91.7** ± 1.6	**1.6** ± 1.4	**6.7** ± 2.8	
wt (M)	**8.3** ± 2.7	**6.0** ± 0.5	**85.7** ± 2.8	**8.4** ± 2.8	**5.5** ± 0.6	**86.0** ± 2.8	**0.061** ± 0.007

Table 5.3: Calculated C_1 fluxes during transitions between succinate and methanol. All values are reported in nmol min^{-1} mL^{-1} OD_{600}^{-1}

Succinate to methanol

Flux	-1	1	5	7.5	10
A	9.4 ± 0.3	18.7 ± 0.5	66.8 ± 5.2	72.6 ± 5.0	96.1 ± 1.2
B	9.3 ± 0.3	18.1 ± 0.6	57.8 ± 4.1	59.0 ± 4.3	76.6 ± 1.1
C	8.2 ± .03	16.3 ± 0.5	48.5 ± 5.9	51.3 ± 5.1	71.1 ± 1.2
D	7.9 ± 0.3	15.5 ± 0.5	43.0 ± 6.1	44.6 ± 5.1	63.2 ± 1.2
E	1.0 ± .01	1.8 ± 0.1	8.6 ± 1.9	7.7 ± 1.1	5.4 ± 0.5
F	0.1 ± < 0.1	0.7 ± 0.1	8.7 ± 1.1	13.6 ± 1.1	19.5 ± 1.1
G	1.2 ± 0.1	2.4 ± 0.1	17.3 ± 0.8	21.3 ± < 0.1	25.0 ± 0.6
H	0.4 ± < 0.1	0.9 ± < 0.1	5.5 ± 0.2	6.7 ± < 0.1	7.9 ± 0.2
I	0.2 ± < 0.1	0.4 ± < 0.1	3.2 ± 0.1	3.9 ± < 0.1	4.6 ± 0.1
J	1.7 ± 0.1	3.6 ± 0.1	26.0 ± 1.1	31.9 ± < 0.1	37.5 ± 0.9

Methanol to succinate

Flux	-1	1	3	5	7
A	102.8 ± 5.0	51.2 ± 6.3	45.7 ± 2.8	37.4 ± 2.1	23.8 ± 1.7
B	69.7 ± 4.6	31.8 ± 3.5	32.1 ± 3.9	29.2 ± 2.9	21.4 ± 1.5
C	67.6 ± 4.7	28.2 ± 2.9	27.4 ± 2.7	22.2 ± 2.5	17.5 ± 1.5
D	56.4 ± 4.5	20.9 ± 2.0	21.5 ± 2.6	17.4 ± 2.6	15.5 ± 1.5
E	2.1 ± 0.2	3.6 ± 0.7	4.8 ± 1.3	7.0 ± 1.0	3.8 ± 0.2
F	33.1 ± 0.7	19.4 ± 3.3	13.6 ± 1.5	8.3 ± 1.2	2.4 ± 0.1
G	35.2 ± 0.5	23.1 ± 3.6	18.3 ± 0.4	15.2 ± 0.4	6.3 ± 0.3
H	11.2 ± 0.2	7.3 ± 1.2	5.8 ± 0.1	4.8 ± 0.1	2.0 ± 0.1
I	6.4 ± 0.1	4.2 ± 0.7	3.4 ± 0.1	2.8 ± 0.1	1.1 ± 0.1
J	52.8 ± 0.8	34.6 ± 5.5	27.5 ± 0.6	22.8 ± 0.6	9.4 ± 0.4

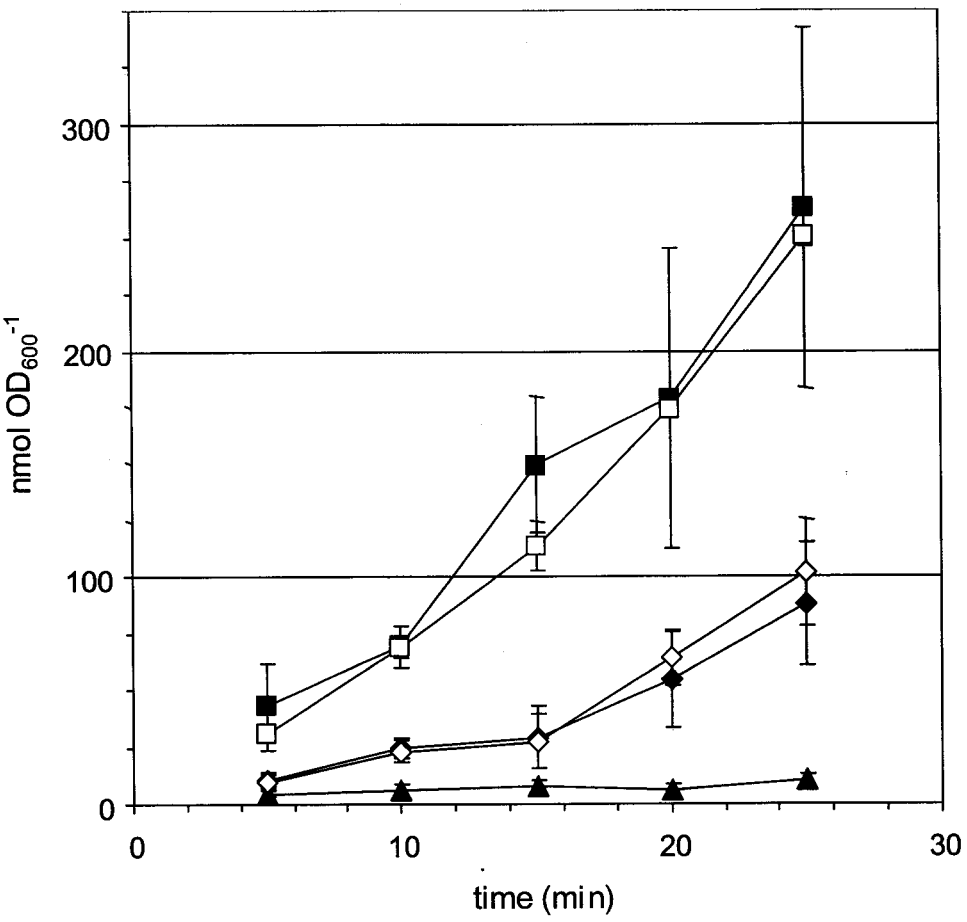


Figure 5.1: Whole-cell production of ^{14}C-CO_2 from ^{14}C-methanol. Strains examined are wild-type (filled squares), the *dmrA* mutant CM212K.1 (filled diamonds), the *ftfL* mutant CM216K.1 (open squares), the *ftfL, dmrA* double mutant CM216-212K.1 (open diamonds), and the *mxaF* mutant CM194K.1 (filled triangles). Consistent kinetics were observed for each of the three replicates. For further experimental detail see Materials and Methods.

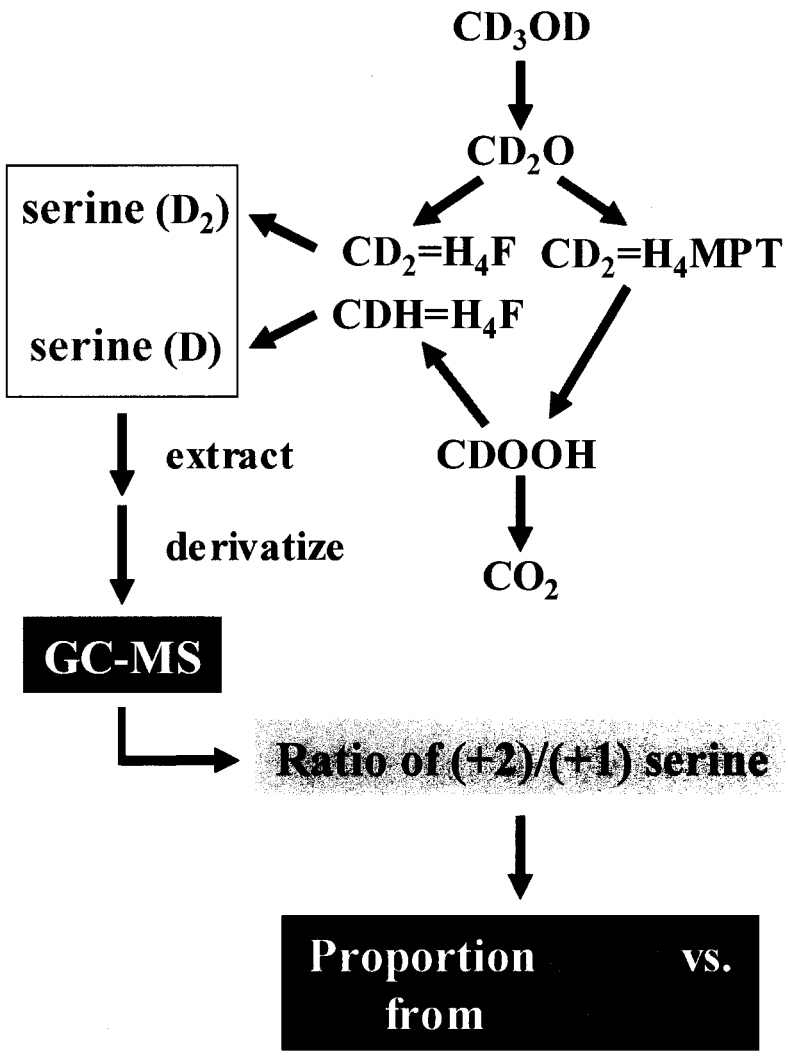


Figure 5.2: CD_3OD label tracing strategy. Oxidation of deuterated methanol (CD_3OD) leads to the production of formaldehyde with two deuteriums. Direct condensation with H_4F denoted in blue generates serine that bears both of these deuteriums. Methylene-H_4F made from formate contains only one of the original deuteriums, thereby generating serine with a single deuterium. A small molecule preparation was extracted and derivatized for analysis by GC-MS. Therefore, determining the ratio of (+2)/(+1) serine isotopomers, the proportion of methylene-H_4F generated directly or from formate can be obtained.

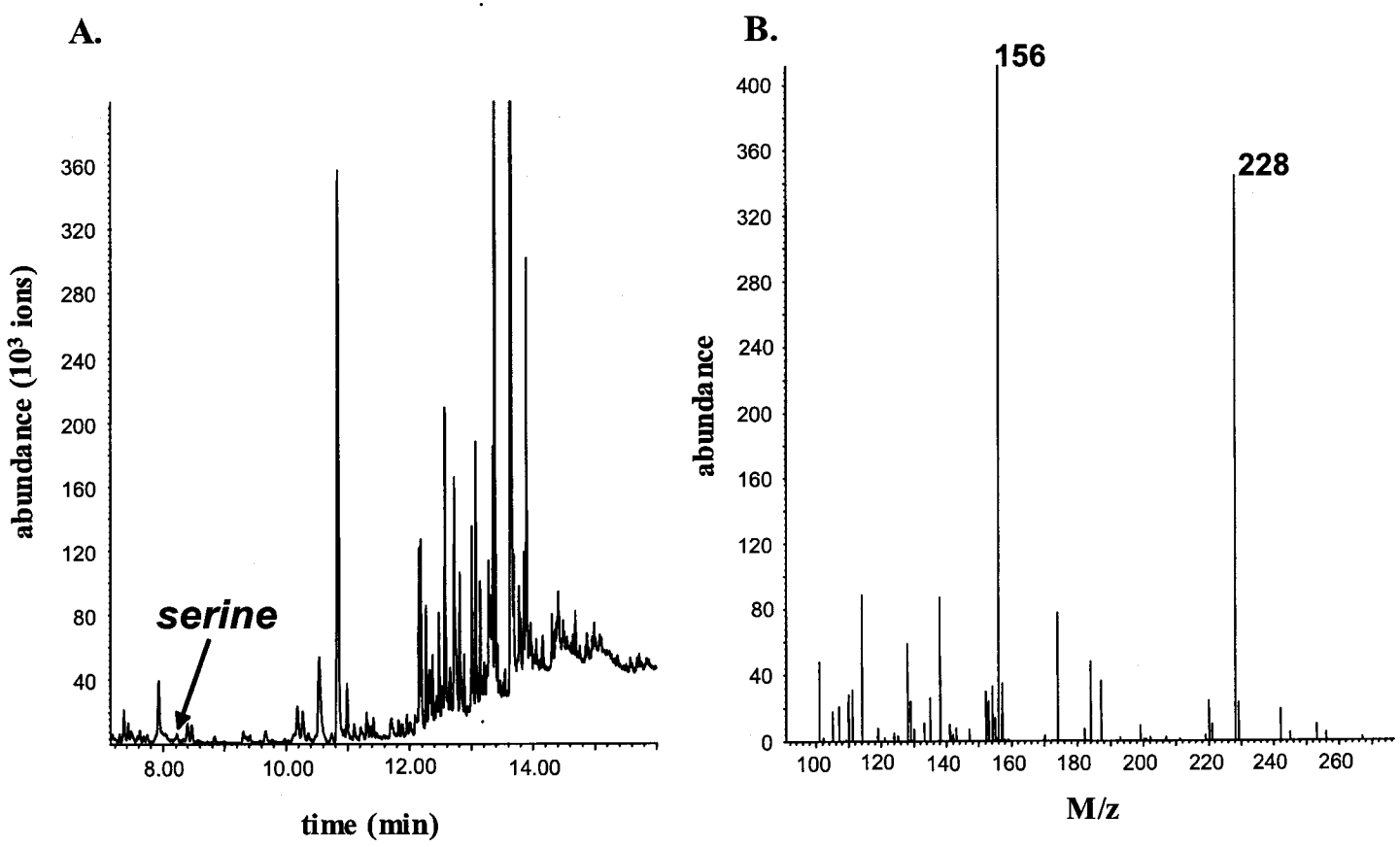


Figure 5.3: Detection of serine by GC-MS. The small peak in total ion abundance detected by the MS denoted by the arrow represents serine (A). Analysis of the mass fragments present in this peak revealed the presence of ions with a M/z of 156 and 228, which are diagnostic for ECF-TFAA derivatized serine (B).

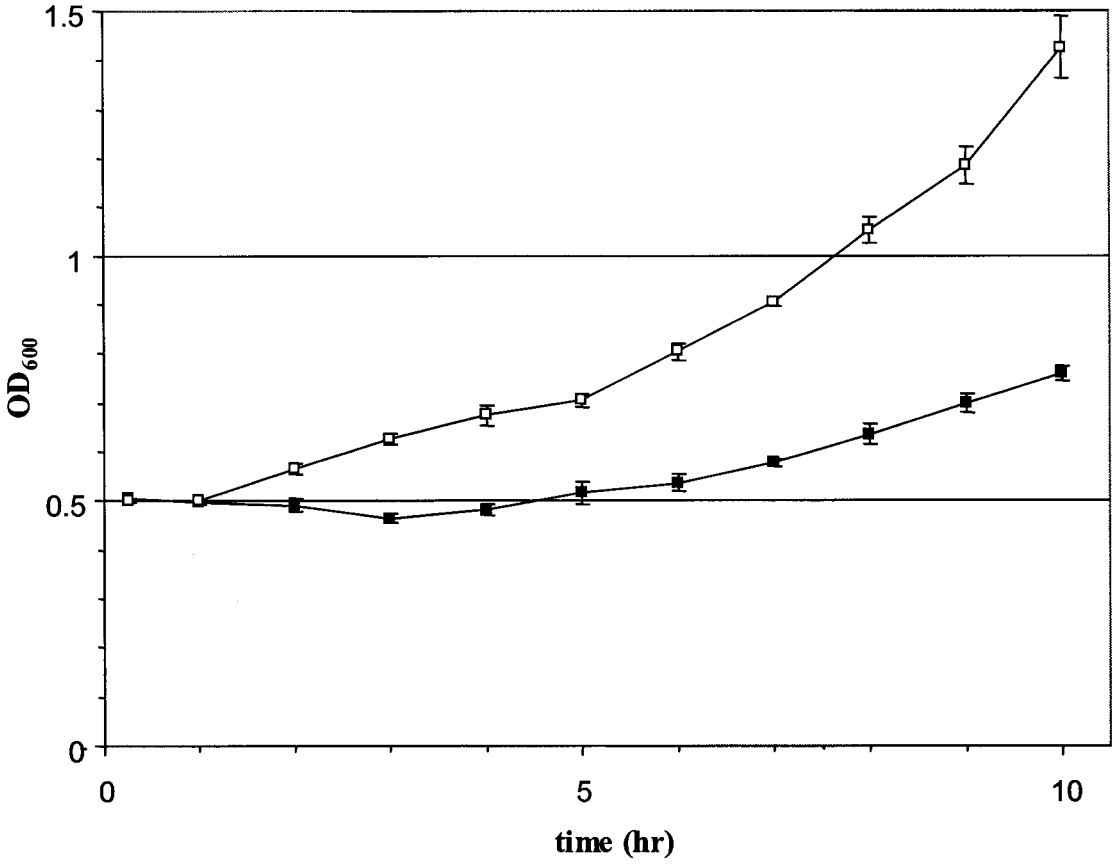


Figure 5.4: Growth of wild-type *M. extorquens* AM1 during transitions. Cultures grown on succinate were shifted to methanol (filled squares), or from methanol to succinate (open squares) at 0 hrs.

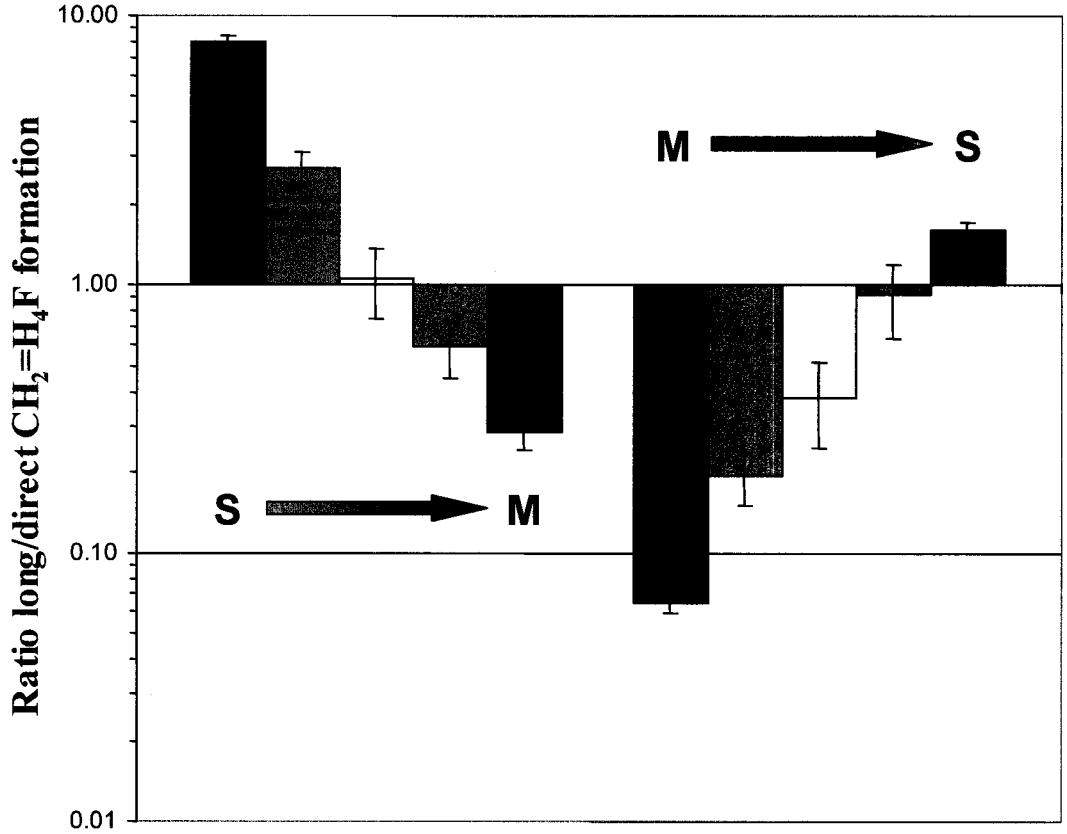


Figure 5.5: Ratio of flux through methylene-H_4F formation pathways. The ratio of the long pathway through formate versus the direct condensation with formaldehyde determined by GC-MS analysis of serine isotopomers during shifts between succinate- and methanol-containing medium. For the succinate to methanol transition, the bars represent -1, 1, 5, 7.5, and 10 hours after the switch to methanol. For the methanol to succinate transition, the bars represent -1, 1, 3, 5, and 7 hours after the switch to succinate.

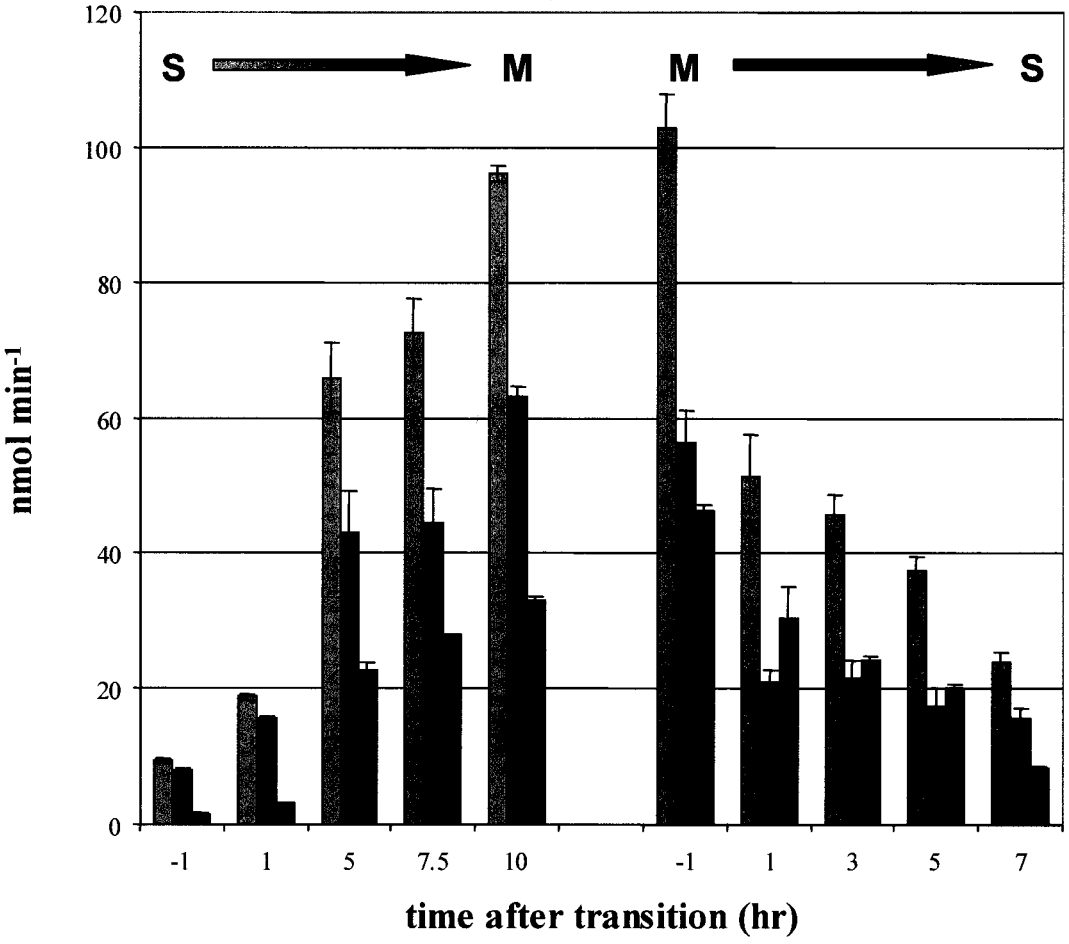


Figure 5.6: Input and output fluxes as determined with ^{14}C labeling. Total methanol oxidation (green), CO_2 production from methanol (orange), and assimilation of C_1 units from methanol (blue) are depicted for both growth transitions.

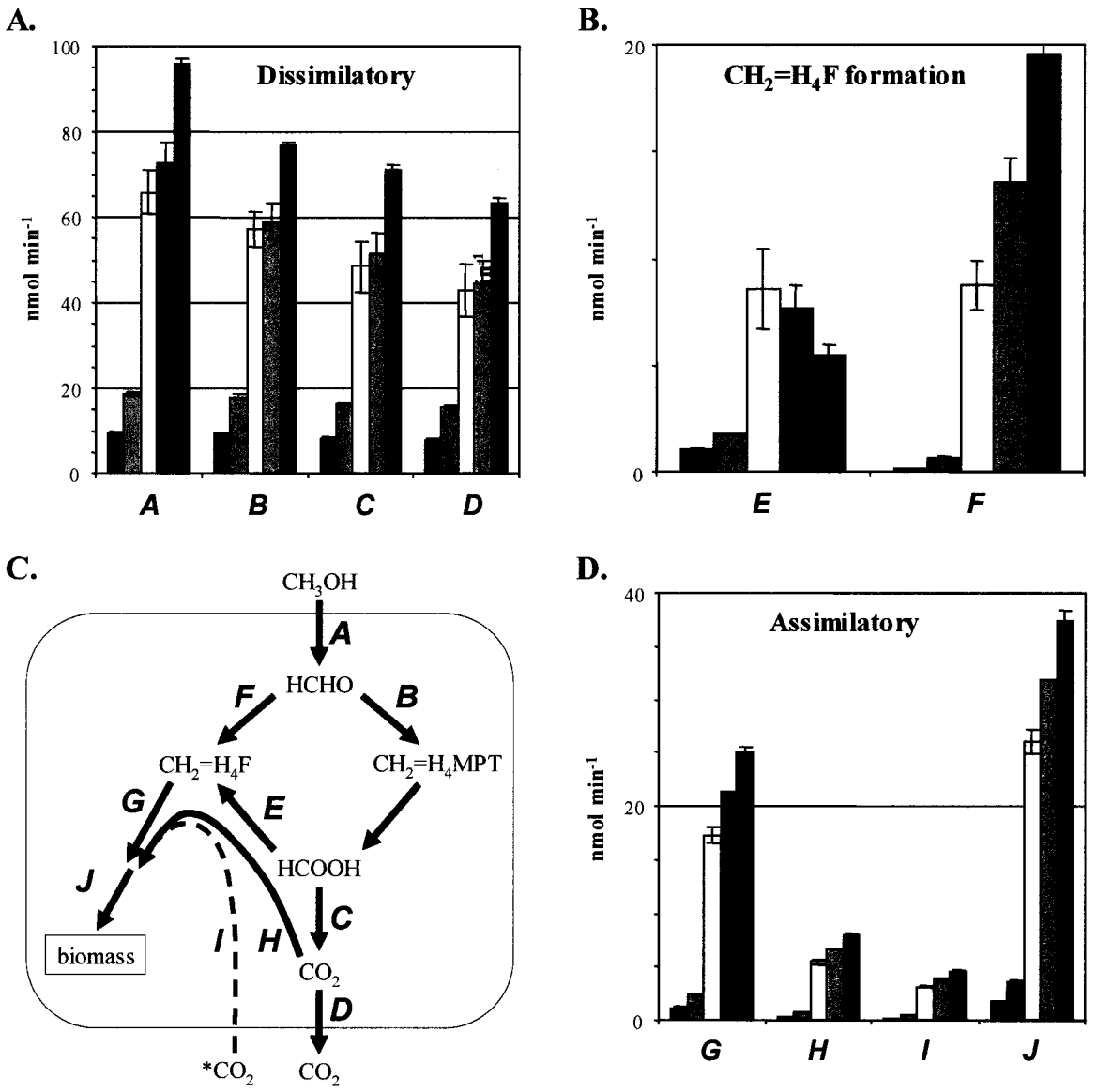


Figure 5.7: C_1 fluxes during transition from succinate to methanol. The fluxes determined are represented schematically (C). The other panels present flux for each branch (labelled A through J) 1 hour prior to the transition to methanol (blue), and 1 (green), 5 (yellow), 7.5 (orange), and 10 (red) hours after the switch. Dissimilatory (A), methylene-H_4F formation (B), and assimilatory (D) fluxes are presented separately with different scales for clarity.

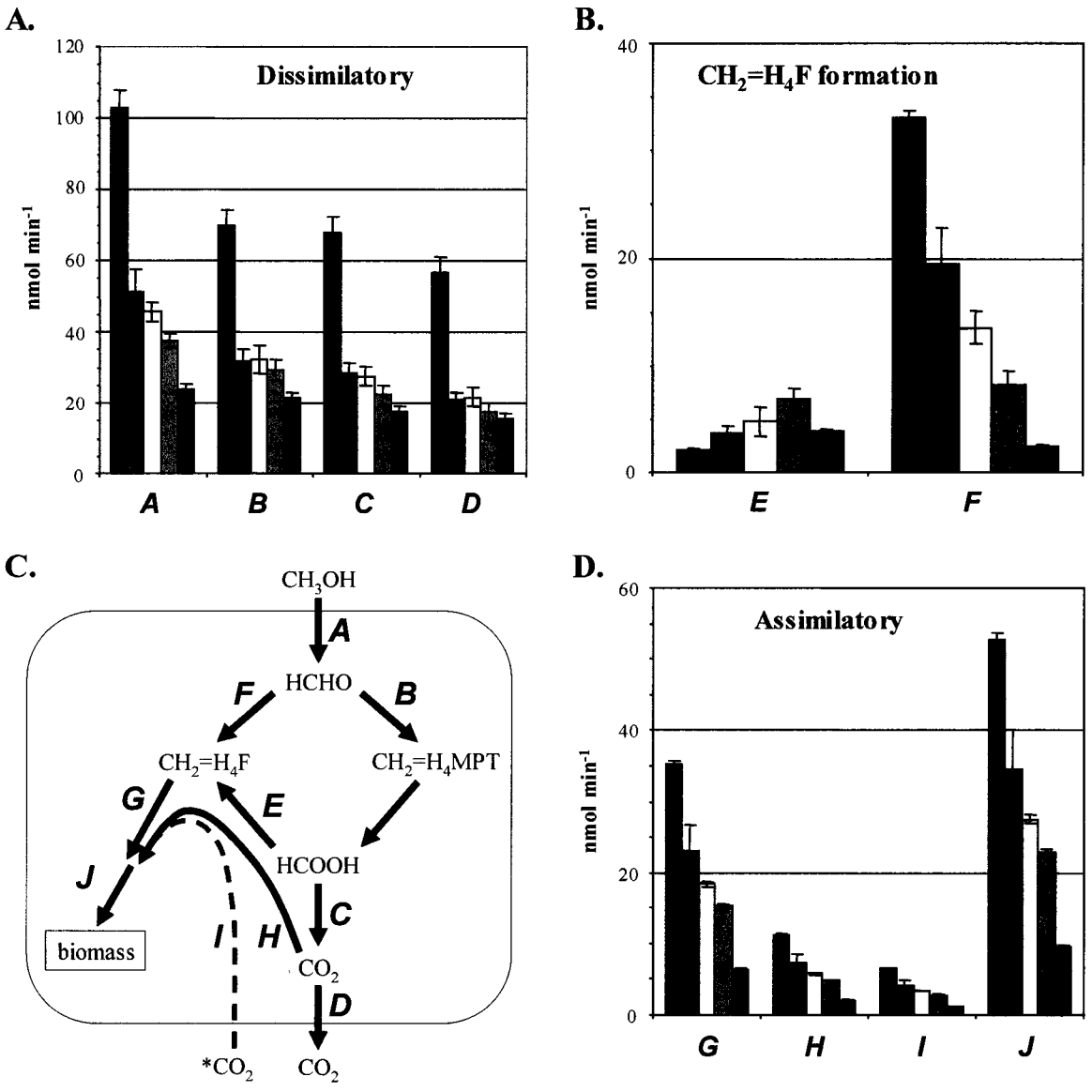


Figure 5.8: C_1 fluxes during transition from methanol to succinate. The fluxes determined are represented schematically (C). The other panels present flux for each branch (labelled A through J) 1 hour prior to the transition to succinate (red), and 1 (orange), 3 (yellow), 5 (green), and 7 (blue) hours after the switch. Dissimilatory (A), methylene-H_4F formation (B), and assimilatory (D) fluxes are presented separately with different scales for clarity.

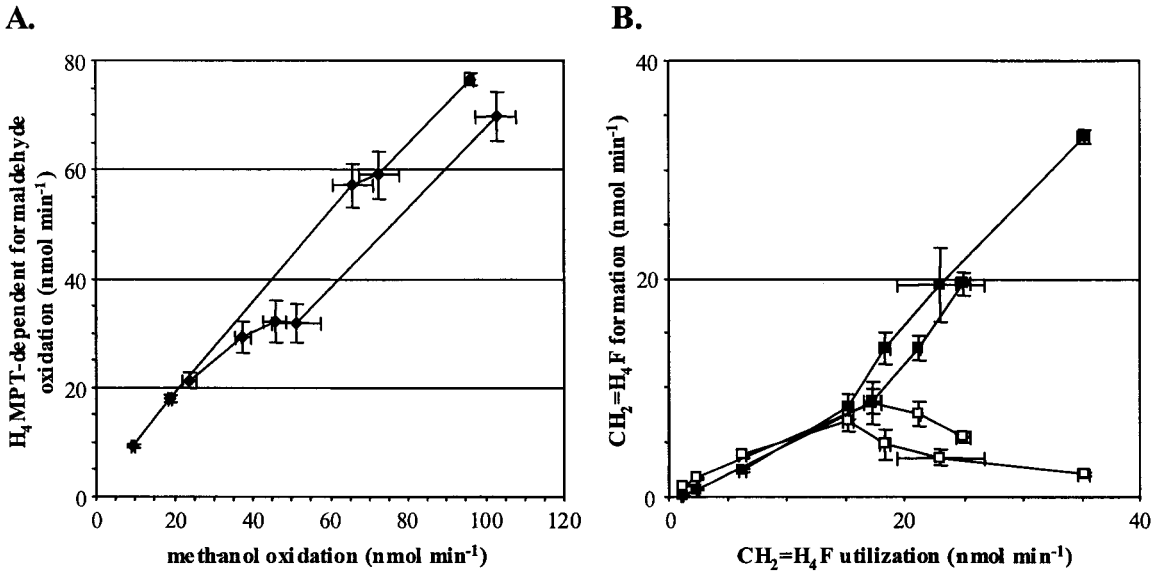


Figure 5.9: Relationships between C_1 fluxes. H_4MPT-dependent formaldehyde oxidation versus methanol oxidation (A). Rate of formation of methylene-H_4F via the H_4F pathway (filled squares) or via direct condensation with formaldehyde (open squares) plotted versus flux of C_1 units into the serine cycle (B). The lines connecting data points indicate consecutive samples for each of the growth transitions.

CHAPTER 6

Conclusions and Future Directions

This thesis has addressed a fundamental question in bacterial physiology: How do cells cope with the production of a toxic central metabolite? Formaldehyde is highly toxic, but must be produced at a high rate in methylotrophs for efficient growth. This poses a special challenge because both detoxification and efficient regulation of carbon flow into assimilatory and dissimilatory metabolism are important for increased fitness. I have examined the formaldehyde metabolism of *M. extorquens* AM1 as a model system for addressing this question. The larger hypothesis of my project is that specific systems are necessary to direct carbon into assimilatory and dissimilatory metabolism while preventing formaldehyde from accumulating to harmful levels. From the start of my project, the key to unraveling this mystery appeared to lie in the fact that this organism possesses two C_1 transfer systems that can handle C_1 units from formaldehyde, the H_4F- and H_4MPT-dependent pathways. The majority of genes involved in these two pathways were identified at the initiation of my project, and the mutant analysis available at the time indicated that both were required for methylotrophy (20, 22, 23). Mutants in *mtdB* of the H_4MPT pathway exhibited a unique phenotype, inhibition by methanol during growth on succinate (43). Combined with the requirement for methanol oxidation for this phenotype, it was suggested that methanol sensitivity is indicative of defective formaldehyde detoxification (43). However, this mutant was not completely blocked for H_4MPT-dependent formaldehyde oxidation (due to the presence of *mtdA*) and further analysis of the role of the H_4MPT pathway was not possible due to the inability to generate null mutants defective for the genes encoding the other enzymes of the pathway *mtdA*, *mch*, and *fhcBADC* (20, 22). Similar issues were manifest with genetic analyses of the H_4F pathway, for null mutants were not obtained for either *mtdA* or *fch* (22, 23), thus a clear interpretation of the role of the H_4F pathway was also not possible. These preliminary analyses led to two specific hypotheses that I tested in this project, 1.) The H_4MPT pathway is the key formaldehyde oxidation and detoxification route, and 2.) The H_4F pathway is required for methylotrophy, but rather than functioning as a second formaldehyde

oxidation route, it plays another role such converting formate into methylene-H_4F for assimilation via the serine cycle. I approached this problem first with genetic and physiological analysis of both mutants defective for one or both C_1 transfer pathways, which was followed by analytical label tracing studies to verify the predicted roles.

In order to perform physiological analyses of formaldehyde metabolism in *M. extorquens* AM1 a series of genetic tools were developed (Chapter 2). This involved a significant time investment that yielded improved plasmid vectors, a *cre-lox*-based allelic exchange system, implementation of transposon mutagenesis, and the construction of an insertional expression system. With these advances and the availability of genome sequence data (www.integratedgenomics.com/genomereleases.html#list6) *M. extorquens* AM1 has emerged as an attractive model organism for physiological studies. The tools I have generated form the basis for all genetic studies in the laboratory and have also served as the backbones for other researchers who have developed improved tools for specific purposes. Furthermore, we have sent these plasmids to dozens of laboratories worldwide where they are being tested with an array of other Gram-negative bacteria for which sophisticated genetic tools are scarce.

The suite of genetic tools developed in Chapter 2 proved to be invaluable for my efforts to demonstrate the role of the H_4MPT and H_4F pathways. First, the application of transposon mutagenesis unexpectedly resulted in the discovery of two new genes involved in the C_1 transfer pathways, *ftfL* and *dmrA*. Whereas the gene encoding *ftfL* would have been identified later through analysis of the genome sequence data, the role of *dmrA* (encodes a putative dihydromethanopterin reductase) would not have been suspected without the data regarding the mutant phenotype. Not only is *dmrA* the only gene identified that is required for the H_4MPT pathway outside of the large gene cluster that contains those that encode the enzymes of the pathway and *orf4* (homolog of β-RFAP synthase, (106)), no enzyme catalyzing this reaction has been described for any organism. Furthermore, the location of *dmrA* and its sequence relationship

to bacterial dihydrofolate reductases, rather than archaeal genes, suggest that it evolved independently of the majority of the genes encoding the H_4MPT pathway. Future studies examining the distribution and phylogeny of *dmrA* homologs will provide critical information for understanding the evolution of the H_4MPT pathway.

Two critical breakthroughs in our understanding have occurred through the genetic analyses of the H_4MPT and H_4F pathways in *M. extorquens* AM1 (Chapters 3 and 4). The first was the demonstration that methanol-sensitivity indicates defective formaldehyde detoxification, and this phenotype was therefore used to classify mutants into different phenotypic classes that correspond to their biochemical role in formaldehyde metabolism. As part of this work, I have demonstrated that a heterologous formaldehyde oxidation system dependent upon GSH can functionally replace the H_4MPT pathway, but not the H_4F pathway. This establishes that the H_4MPT pathway and the GSH pathway are analogous metabolic modules that perform the same functions for the cell, which accounts for their interchangeability. An implication of this result is that non-orthologous formaldehyde oxidation modules (different enzymes and cofactor) can be exchanged across different clades of methylotrophic bacteria to result in functional methylotrophs. The phylogenetically incoherent nature of formaldehyde oxidation modules may be a result of this interchangeability. The second fundamental result was the demonstration that null mutants could be obtained for every gene encoding C_1 transfer functions. The ability to obtain nulls mutants for *mtdA* and *fch* in the presence of either formate or *folD*-expression demonstrated that the only requirement for these genes during multi-carbon growth was to provide formyl-H_4F for biosynthesis. Mutations that result in a complete block of the H_4MPT pathway were also obtained, but only in a strain defective for H_4MPT biosynthesis. This result demonstrated that the H4MPT pathway is not essential during multi-carbon growth but that the accumulation of a C_1 intermediate(s) is not tolerated, either due to toxicity or regulatory effects.

Uncovering the reasons behind the inability to obtain null mutants for these genes in wild-type during multi-carbon growth eliminated speculation as to other potential roles for the cell.

The genetic and physiological analysis of the H_4MPT and H_4F pathways led me to suggest a new model for methylotrophic growth by *M. extorquens* AM1. I postulated that the H_4MPT pathway represents the sole route for oxidation of formaldehyde, whereas the H_4F pathway converts some of the resulting formate into methylene-H_4F for assimilation by the serine cycle, thus providing a second mechanism for methylene-H_4F production in addition to the direct, non-enzymatic condensation with formaldehyde. The first version of this model included an arrow representing a hypothetical mechanism for channeling carbon from the H_4MPT pathway into the H_4F pathway at the level of a formyl group, for at the time it had been assumed that the H_4MPT pathway dissimilated formaldehyde completely to CO_2. Remarkably this arrow did not remain hypothetical for long, for several months later our collaborators in the Vorholt laboratory discovered that Fhc releases formate, rather than CO_2, as the end product of the H_4MPT pathway (95), providing the direct connection between the two C_1 transfer pathways that I had suspected. With this information in hand, I decided to implement a combination of ^{14}C-methanol and CD_3OD label tracing methods to test this model (Chapter 5). Using ^{14}C-methanol, I provided direct biochemical evidence for the first time that the H_4MPT pathway is required for complete dissimilation to CO_2, whereas the H_4F pathway is not required. Using CD_3OD and the GC-MS method I developed to identify free serine I demonstrated that methylene-H_4F is produced both from direct condensation with formaldehyde and from formate through the activity of the H_4F pathway.

Cultures transferred between medium containing succinate or methanol were examined to understand the relative contribution of the two methylene-H_4F pathways under different physiological conditions. This revealed that succinate-grown cultures shifted to methanol initially generate the majority of methylene-H_4F from formate, but that the direct condensation with

formaldehyde eventually dominates. In order to move beyond an examination of the ratio of these two pathways, ^{14}C label tracing was performed concurrently, which allowed the calculation of the actual flux through each branch of methylotrophic metabolism. This analysis provided for the first time a comprehensive, dynamic picture of methylotrophic metabolism, focused on the C_1 transfer and incorporation steps. It nicely complements another approach used in the laboratory, a steady-state flux analysis performed using ^{13}C-methanol that addressed the fluxes through central metabolism after carbon enters the serine cycle, metabolically downstream of the C_1 transfer pathways I have studied (Van Dien et al., *in press*). These two projects have provided a flux "roadmap" that will serve as the basis for nascent projects in the laboratory aimed at understanding the coordinate regulation between the various branches of methylotrophic metabolism. With respect to formaldehyde metabolism, the critical result from my flux analyses was that, although the proportion of methylene-H_4F generated from formate decreases during the shift to growth on methanol, the flux through the "long" pathway increases. This is consistent with previous work demonstrating that the H_4F-dependent enzymes are induced during growth on methanol (22, 67, 77, 97, 125). The direct condensation route, however, increases by over two orders of magnitude during the shift to methylotrophic growth.

Finally, the new insights available from the genetic, physiological, and analytical approaches I have employed suggest the following model for how *M. extorquens* AM1 has evolved to handle the toxic compound formaldehyde as a central metabolite. When *M. extorquens* AM1 encounters C_1 compounds in the environment that can be oxidized to formaldehyde, the serine cycle is uninduced, and thus only trace condensation with H_4F can occur (see Fig. 5.9). Instead, the high basal levels of H_4MPT enzymes and the presence of Fae to catalyze the condensation with formaldehyde allow formaldehyde to be readily detoxified by its conversion to formate, and also involve energy metabolism by producing NAD(P)H. A small fraction of this formate can be converted to methylene-H_4F by the H_4F pathway. In the process, formyl-H_4F is

generated, which induces QscR to activate transcription of serine cycle genes (M. G. Kalyuzhnaya, unpublished). Although methylene-H_4F generation from formate costs an extra ATP per formate as compared to the direct condensation route, it provides a means to initiate the serine cycle without accumulating the formaldehyde necessary to drive the non-enzymatic synthesis of methylene-H_4F. Without the contribution of the long pathway, the carbon assimilated per methanol utilized would drop from 18% to 2% for cells just beginning to transition from succinate to methanol. Only after the serine cycle is induced and assimilatory flux rises can it serve as a sink for methylene-H_4F to "pull" the reaction of formaldehyde with H_4F toward methylene-H_4F production, thereby reducing the dependence on the energy-expensive route from formate. If all methylene-H_4F were made from the long pathway (i.e. no direct condensation) the ATP utilization for generating methylene-H_4F would rise 15-fold to nearly one ATP for every three methanol molecules oxidized. Thus, the two C_1 transfer pathways of *M. extorquens* AM1 provide a dynamic formaldehyde buffer while elegantly and efficiently balancing carbon flow into assimilatory and dissimilatory metabolism, allowing the cell to respond to a wide range of formaldehyde production rates. The general strategy that *M. extorquens* AM1 employs to deal with formaldehyde may be true for other organisms that produce toxic central metabolites: always maintain an efficient means to detoxify while using one or more separate pathways to shunt the metabolite into assimilatory metabolism.

In order to build upon the model of methylotrophic metabolism in *M. extorquens* AM1 I have presented here, three primary areas of work are needed. The first is to uncover the regulation of the various methylotrophic modules at the level of enzyme synthesis and activity. Transcriptional regulators of methanol oxidation (86, 111-113) and the serine cycle (56) are characterized, but transcriptional regulation of the H_4MPT pathway or *ftfL* remain unknown and post-transcriptional control has not been addressed. The potential effects of allosteric regulation of enzyme activity also have not been examined. The second area of work that will further our

knowledge of methylotrophy is to use analytical techniques to assay additional layers of information besides flux. These include genomics techniques such as DNA microarrays and proteomics, assaying the activity of key methylotrophic enzymes over a variety of conditions, and developing sensitive analytical techniques to determine the actual concentration of key intermediates. Whereas the net result of my project has been the determination of what fluxes occur throughout methylotrophic metabolism, it is my expectation that these future efforts will allow us to understand how this remarkable biological system operates. Finally, flux through the branches of methylotrophic metabolism need to be investigated in cells growing in their natural niche, the surface of leaves. Methylotrophs in nature are likely exposed to multiple growth substrates and do not generally, if ever, experience the luxuriant, stable conditions provided in the laboratory setting that promote fast growth and high metabolic fluxes. My work suggests that the ability to direct carbon into assimilatory metabolism through either of the two methylene-H_4F production pathways provides a mechanism for *Methylobacterium* strains in the natural environment to maximize growth while limiting the toxic effects of formaldehyde.

REFERENCES

1. **Anantharam, V., M. J. Allison, and P. C. Maloney.** 1989. Oxalate:formate exchange, the basis for energy coupling in *Oxalobacter*. J Biol Chem **264:**7244-7250.

2. **Anthony, C., and L. J. Zatman.** 1964. The microbial oxidation of methanol. 2. The methanol-oxidizing enzyme of *Pseudomonas* sp. M27. Biochem J **92:**614-621.

3. **Arfman, N., E. M. Watling, W. Clement, R. J. van Oosterwijk, G. E. de Vries, W. Harder, M. M. Attwood, and L. Dijkhuizen.** 1989. Methanol metabolism in thermotolerant methylotrophic *Bacillus* strains involving a novel catabolic NAD-dependent methanol dehydrogenase as a key enzyme. Arch Microbiol **152:**280-8.

4. **Attwood, M. M., and W. Harder.** 1972. A rapid and specific enrichment procedure for *Hyphomicrobium* spp. Antonie Van Leeuwenhoek **38:**369-77.

5. **Attwood, M. M., and J. R. Quayle.** 1984. Formaldehyde as a central intermediary metabolite of methylotrophic metabolism, p. 315-323. *In* R. L. Crawford and R. S. Hanson (ed.), Microbial growth on C1 compounds. American Society for Microbiology, Washington, D. C.

6. **Auman, A. J., S. Stolyar, A. M. Costello, and M. E. Lidstrom.** 2000. Molecular characterization of methanotrophic isolates from freshwater lake sediment. Appl Environ Microbiol **66:**5259-66.

7. **Ayres, E. K., V. J. Thomson, G. Merino, D. Balderes, and D. H. Figurski.** 1993. Precise deletions in large bacterial genomes by vector-mediated excision (VEX). The *trfA* gene of promiscuous plasmid RK2 is essential for replication in several gram-negative hosts. J Mol Biol **230:**174-85.

8. **Baetz, A. L., and M. J. Allison.** 1990. Purification and characterization of formyl-coenzyme A transferase from *Oxalobacter formigenes*. J Bacteriol **172:**3537-3540.

9. **Baetz, A. L., and M. J. Allison.** 1989. Purification and characterization of oxalyl-coenzyme A decarboxylase from *Oxalobacter formigenes*. J Bacteriol **171:**2605-2608.

10. **Bailey, J., and C. Manoil.** 2002. Genome-wide internal tagging of bacterial exported proteins. Nat Biotechnol **20:**839-42.

11. **Barber, R. D., and T. J. Donohue.** 1998. Function of a glutathione-dependent formaldehyde dehydrogenase in *Rhodobacter sphaeroides* formaldehyde oxidation and assimilation. Biochemistry **37:**530-7.

12. **Barber, R. D., M. A. Rott, and T. J. Donohue.** 1996. Characterization of a glutathione-dependent formaldehyde dehydrogenase from *Rhodobacter sphaeroides*. J Bacteriol **178:**1386-93.

13. **Biville, F., P. Mazodier, E. Turlin, and F. Gasser.** 1989. Mutants of *Methylobacterium organophilum* unable to synthesize PQQ. Antonie Van Leeuwenhoek **56:**103-7.

14. **Blatny, J. M., T. Brautaset, H. C. Winther-Larsen, P. Karunakaran, and S. Valla.** 1997. Improved broad-host-range RK2 vectors useful for high and low regulated gene expression levels in gram-negative bacteria. Plasmid **38:**35-51.

15. **Bystrykh, L. V., N. I. Govorukhina, L. Dijkhuizen, and J. A. Duine.** 1997. Tetrazolium-dye-linked alcohol dehydrogenase of the methylotrophic actinomycete *Amycolatopsis methanolica* is a three-component complex. Eur J Biochem **247:**280-7.

16. **Chambers, S. P., S. E. Prior, D. A. Barstow, and N. P. Minton.** 1988. The pMTL *nic*$^-$ cloning vectors. I. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing. Gene **68:**139-49.

17. **Chistoserdov, A. Y., L. V. Chistoserdova, W. S. McIntire, and M. E. Lidstrom.** 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:**4052-65.

18. **Chistoserdova, L., S. W. Chen, A. Lapidus, and M. E. Lidstrom.** 2003. Methylotrophy in *Methylobacterium extorquens* AM1 from a Genomic Point of View. J Bacteriol **185:**2980-7.

19. **Chistoserdova, L., L. Gomelsky, J. A. Vorholt, M. Gomelsky, Y. D. Tsygankov, and M. E. Lidstrom.** 2000. Analysis of two formaldehyde oxidation pathways in *Methylobacillus flagellatus* KT, a ribulose monophosphate cycle methylotroph. Microbiology **146 (Pt 1):**233-8.

20. **Chistoserdova, L., J. A. Vorholt, R. K. Thauer, and M. E. Lidstrom.** 1998. C1 transfer enzymes and coenzymes linking methylotrophic bacteria and methanogenic Archaea. Science **281:**99-102.

21. **Chistoserdova, L. V., and M. E. Lidstrom.** 1994. Genetics of the serine cycle in *Methylobacterium extorquens* AM1: cloning, sequence, mutation, and physiological effect of *glyA*, the gene for serine hydroxymethyltransferase. J Bacteriol **176:**6759-62.

22. **Chistoserdova, L. V., and M. E. Lidstrom.** 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.

23. **Chistoserdova, L. V., and M. E. Lidstrom.** 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.

24. **Chistoserdova, L. V., and M. E. Lidstrom.** 1996. Molecular characterization of a chromosomal region involved in the oxidation of acetyl-CoA to glyoxylate in the isocitrate-lyase-negative methylotroph *Methylobacterium extorquens* AM1. Microbiology **142 (Pt 6):**1459-68.

25. **Christensen, B., and J. Nielsen.** 1999. Isotopomer analysis using GC-MS. Metab Eng **1:**282-90.

26. **Chun, K. T., H. J. Edenberg, M. R. Kelley, and M. G. Goebl.** 1997. Rapid amplification of uncharacterized transposon-tagged DNA sequences from genomic DNA. Yeast **13:**233-40.

27. **Conway, T., and G. K. Schoolnik.** 2003. Microarray expression profiling: capturing a genome-wide portrait of the transcriptome. Mol Microbiol **47:**879-89.

28. **Cordes, C., R. Meima, B. Twiest, B. Kazemier, G. Venema, J. M. van Dijl, and S. Bron.** 1996. The expression of a plasmid-specified exported protein causes structural plasmid instability in *Bacillus subtilis*. J Bacteriol **178:**5235-42.

29. **D'Argenio, D. A., L. A. Gallagher, C. A. Berg, and C. Manoil.** 2001. *Drosophila* as a model host for *Pseudomonas aeruginosa* infection. J Bacteriol **183:**1466-71.

30. **Davidson, V. L.** 1999. Methylamine dehydrogenase: structure and function of electron transfer complexes. Biochem Soc Trans **27:**201-6.

31. **Davison, J.** 2002. Genetic tools for pseudomonads, rhizobia, and other Gram-negative bacteria. BioTechniques **32:**386-8, 390, 392-4, passim.

32. **de Lorenzo, V., M. Herrero, U. Jakubzik, and K. N. Timmis.** 1990. Mini-Tn*5* transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. J Bacteriol **172:**6568-72.

33. **Ditta, G., T. Schmidhauser, E. Yakobson, P. Lu, X. W. Liang, D. R. Finlay, D. Guiney, and D. R. Helinski.** 1985. Plasmids related to the broad host range vector, pRK290, useful for gene cloning and for monitoring gene expression. Plasmid **13:**149-53.

34. **Escalante-Semerena, J. C., K. L. Rinehart, Jr., and R. S. Wolfe.** 1984. Tetrahydromethanopterin, a carbon carrier in methanogenesis. J Biol Chem **259:**9447-55.

35. **Feldman, M. Y.** 1973. Reactions of nucleic acids and nucleoproteins with formaldehyde. Prog Nucleic Acid Res Mol Biol **13:**1-49.

36. **Figurski, D. H., and D. R. Helinski.** 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans*. Proc Natl Acad Sci U S A **76:**1648-52.

37. **Figurski, D. H., R. J. Meyer, and D. R. Helinski.** 1979. Suppression of Co1E1 replication properties by the Inc P-1 plasmid RK2 in hybrid plasmids constructed *in vitro*. J Mol Biol **133:**295-318.

38. **Galbally, I. E., and W. Kirstine.** 2002. The production of methanol by flowering plants and the global cycle of methanol. J Atmospheric Chemistry **43:**195-229.

39. **Goenrich, M., S. Bartoschek, C. H. Hagemeier, C. Griesinger, and J. A. Vorholt.** 2002. A glutathione-dependent formaldehyde-activating enzyme (Gfa) from *Paracoccus denitrificans* detected and purified via two-dimensional proton exchange NMR spectroscopy. J Biol Chem **277:**3069-72.

40. **Goodwin, P. M., and C. Anthony.** 1998. The biochemistry, physiology and genetics of PQQ and PQQ-containing enzymes. Adv Microb Physiol **40:**1-80.

41. **Grafstrom, R. C., A. J. Fornace, Jr., H. Autrup, J. F. Lechner, and C. C. Harris.** 1983. Formaldehyde damage to DNA and inhibition of DNA repair in human bronchial cells. Science **220:**216-8.

42. **Graham, D. E., and R. H. White.** 2002. Elucidation of methanogenic coenzyme biosyntheses: from spectroscopy to genomics. Nat Prod Rep **19:**133-47.

43. **Hagemeier, C. H., L. Chistoserdova, M. E. Lidstrom, R. K. Thauer, and J. A. Vorholt.** 2000. Characterization of a second methylene tetrahydromethanopterin dehydrogenase from *Methylobacterium extorquens* AM1. Eur J Biochem **267:**3762-9.

44. **Hanson, R. S., and T. E. Hanson.** 1996. Methanotrophic bacteria. Microbiol Rev **60:**439-71.

45. **Harms, N., J. Ras, W. N. Reijnders, R. J. van Spanning, and A. H. Stouthamer.** 1996. *S*-formylglutathione hydrolase of *Paracoccus denitrificans* is homologous to human esterase D: a universal pathway for formaldehyde detoxification? J Bacteriol **178:**6296-9.

46. **Hartwell, L. H., J. J. Hopfield, S. Leibler, and A. W. Murray.** 1999. From molecular to modular cell biology. Nature **402:**C47-52.

47. **Haugan, K., P. Karunakaran, A. Tondervik, and S. Valla.** 1995. The host range of RK2 minimal replicon copy-up mutants is limited by species-specific differences in the maximum tolerable copy number. Plasmid **33:**27-39.

48. **Hoang, T. T., R. R. Karkhoff-Schweizer, A. J. Kutchma, and H. P. Schweizer.** 1998. A broad-host-range *Flp*-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants. Gene **212:**77-86.

49. **Holloway, B. W.** 1984. Genetics of methylotrophs, p. 87-106. *In* C. Hou (ed.), Methylotrophs: microbiology, biochemistry and genetics. CRC Press, Boca Raton, FL.

50. **Husek, P.** 1991. Amino acid derivatization and analysis in five minutes. FEBS Lett **280:**354-6.

51. **Jansson, C., and O. Skold.** 1991. Appearance of a new trimethoprim resistance gene, *dhfrIX*, in *Escherichia coli* from swine. Antimicrob Agents Chemother **35:**1891-9.

52. **Jiang, B., and S. P. Howard.** 1992. The *Aeromonas hydrophila exeE* gene, required both for protein secretion and normal outer membrane biogenesis, is a member of a general secretion pathway. Mol Microbiol **6:**1351-61.

53. **Johnson, P. A., and J. R. Quayle.** 1964. Microbial growth on C_1 compounds. 6. Oxidation of methanol, formaldehyde and formate by methanol-grown *Pseudomonas* AM1. Biochemical Journal **93:**281-290.

54. **Kalb, V. F., and R. W. Bernlohr.** 1977. A new spectrophotometric assay for protein in cell extracts. Anal Biochem **82:**362-371.

55. **Kallen, R. G., and W. P. Jencks.** 1966. The mechanism of the condensation of formaldehyde with tetrahydrofolic acid. J Biol Chem **241:**5851-63.

56. **Kalyuzhnaya, M. G., and M. E. Lidstrom.** 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:**1229-35.

57. **Kataeva, I. M., and L. A. Golovleva.** 1990. Catechol 2,3-dioxygenases from Pseudomonas aeruginosa 2x. Methods Enzymol **188:**115-121.

58. **Kayser, M. F., and S. Vuilleumier.** 2001. Dehalogenation of dichloromethane by dichloromethane dehalogenase/glutathione *S*-transferase leads to formation of DNA adducts. J Bacteriol **183:**5209-12.

59. **King, G. M.** 1992. Ecological aspects of methane oxidation, a key determinant of global methane dynamics. Adv Microb Ecol **12:**431-474.

60. **Knauf, V. C., and E. W. Nester.** 1982. Wide host range cloning vectors: a cosmid clone bank of an *Agrobacterium* Ti plasmid. Plasmid **8:**45-54.

61. **Koenig, R. L., R. O. Morris, and J. C. Polacco.** 2002. tRNA is the source of low-level *trans*-zeatin production in *Methylobacterium* spp. J Bacteriol **184:**1832-42.

62. **Korotkova, N., L. Chistoserdova, V. Kuksa, and M. E. Lidstrom.** 2002. Glyoxylate regeneration pathway in the methylotroph *Methylobacterium extorquens* AM1. J Bacteriol **184:**1750-8.

63. **Korotkova, N., L. Chistoserdova, and M. E. Lidstrom.** 2002. Poly-beta-hydroxybutyrate biosynthesis in the facultative methylotroph *Methylobacterium extorquens* AM1: identification and mutation of *gap11*, *gap20*, and *phaR*. J Bacteriol **184:**6174-81.

64. **Kovach, M. E., P. H. Elzer, D. S. Hill, G. T. Robertson, M. A. Farris, R. M. Roop, 2nd, and K. M. Peterson.** 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene **166:**175-6.

65. **Kristensen, C. S., L. Eberl, J. M. Sanchez-Romero, M. Givskov, S. Molin, and V. De Lorenzo.** 1995. Site-specific deletions of chromosomally located DNA segments with the multimer resolution system of broad-host-range plasmid RP4. J Bacteriol **177:**52-8.

66. **Large, P. J., D. Peel, and J. R. Quayle.** 1961. Microbial growth on C_1 compounds. 2. Synthesis of cell constituents by methanol- and formate-grown *Pseudomonas* AM1, and methanol-grown *Hyphomicrobium vulgare*. Biochemical Journal **81:**470-479.

67. **Large, P. J., and J. R. Quayle.** 1963. Microbial growth on C_1 compounds. 5. Enzyme activities in extracts of *Pseudomonas* AM1. Biochem J **87:**383-396.

68. **Laukel, M., L. Chistoserdova, M. E. Lidstrom, and J. A. Vorholt.** 2003. The tungsten-containing formate dehydrogenase from *Methylobacterium extorquens* AM1: purification and properties. Eur J Biochem **270:**325-33.

69. **Lidstrom, M. E.** 2 November 2001, posting date. Aerobic metylotrophic prokaryotes. *In* M. Dworkin (ed.), The procaryotes. link.springer.de/link/service/books/10125/. [Online.]

70. **Lidstrom, M. E., and L. Somers.** 1984. Seasonal study of methane consumption in Lake Washington. Appl Environ Microbiol **47:**1255-1260.

71. **Lidstrom, M. E., and Y. D. Tsygankov.** 1991. Molecular genetics of methylotrophic bacteria. Biotechnology **18:**273-304.

72. **Ljungdahl, L. G.** 1986. The autotrophic pathway of acetate synthesis in acetogenic bacteria. Annu Rev Microbiol **40:**415-450.

73. **Luria, S., and M. Delbruck.** 1943. Mutation of bacteria from virus sensitive to virus resistant. Genetics **28:**491-511.

74. **Madsen, E. H.** 2000. Tetrahydrofolate and tetrahydromethanopterin compared: functionally distinct carriers in C_1 metabolism. Biochem J **350:**609-629.

75. **Manoil, C.** 2000. Tagging exported proteins using *Escherichia coli* alkaline phosphatase gene fusions. Methods Enzymol **326:**35-47.

76. **Manoil, C., and J. Beckwith.** 1985. Tn*phoA*: a transposon probe for protein export signals. Proc Natl Acad Sci U S A **82:**8129-33.

77. **Marison, I. W., and M. M. Attwood.** 1982. A possible alternative mechanism for the oxidation of formaldehyde to formate. J Gen Microbiol **128:**1441-1446.

78. **Marx, C. J., and M. E. Lidstrom.** 2002. Broad-host-range *cre-lox* system for antibiotic marker recycling in gram-negative bacteria. BioTechniques **33:**1062-7.

79. **Marx, C. J., and M. E. Lidstrom.** 2001. Development of improved versatile broad-host-range vectors for use in methylotrophs and other Gram-negative bacteria. Microbiology **147:**2065-75.

80. **Marx, C. J., B. N. O'Brien, J. Breezee, and M. E. Lidstrom.** 2003. Novel methylotrophy genes of *Methylobacterium extorquens* AM1 identified by using transposon mutagenesis including a putative dihydromethanopterin reductase. J Bacteriol **185:**669-73.

81. **Mason, R. P., J. K. Sanders, A. Crawford, and B. K. Hunter.** 1986. Formaldehyde metabolism by *Escherichia coli*. Detection by in vivo ^{13}C NMR spectroscopy of S-(hydroxymethyl)glutathione as a transient intracellular intermediate. Biochemistry **25:**4504-7.

82. **McIntire, W. S., D. E. Wemmer, A. Chistoserdov, and M. E. Lidstrom.** 1991. A new cofactor in a prokaryotic enzyme: tryptophan tryptophylquinone as the redox prosthetic group in methylamine dehydrogenase. Science **252:**817-24.

83. **Miller, J. H.** 1972. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

84. **Miller, V. L., and J. J. Mekalanos.** 1988. A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins

and virulence determinants in *Vibrio cholerae* requires *toxR*. J Bacteriol **170:**2575-83.

85. **Misset-Smits, M., P. W. van Ophem, S. Sakuda, and J. A. Duine.** 1997. Mycothiol, 1-O-(2'-[*N*-acetyl-L-cysteinyl]amido-2'-deoxy-alpha-D-glucopyranosyl)-D- myo-inositol, is the factor of NAD/factor-dependent formaldehyde dehydrogenase. FEBS Lett **409:**221-2.

86. **Morris, C. J., and M. E. Lidstrom.** 1992. Cloning of a methanol-inducible *moxF* promoter and its analysis in *moxB* mutants of *Methylobacterium extorquens* AM1rif. J Bacteriol **174:**4444-9.

87. **Nagy, P. L., A. Marolewski, S. J. Benkovic, and H. Zalkin.** 1995. Formyltetrahydrofolate hydrolase: a regulatory enzyme that functions to balance pools of tetrahydrofolate and one-carbon tetrahydrofolate adducts in *Escherichia coli*. J Bacteriol **177:**1292-1298.

88. **Nunn, D., S. Bergman, and S. Lory.** 1990. Products of three accessory genes, *pilB*, *pilC*, and *pilD*, are required for biogenesis of *Pseudomonas aeruginosa* pili. J Bacteriol **172:**2911-9.

89. **Nunn, D. N., and M. E. Lidstrom.** 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.

90. **Nunn, D. N., and M. E. Lidstrom.** 1986. Phenotypic characterization of 10 methanol oxidation mutant classes in *Methylobacterium* sp. strain AM1. J Bacteriol **166:**591-7.

91. **Page, M. D., and S. J. Ferguson.** 1993. Mutants of *Methylobacterium extorquens* and *Paracoccus denitrificans* deficient in *c*-type cytochrome biogenesis synthesise the methylamine-dehydrogenase polypeptides but cannot assemble the tryptophan-tryptophylquinone group. Eur J Biochem **218:**711-7.

92. **Palmeros, B., J. Wild, W. Szybalski, S. Le Borgne, G. Hernandez-Chavez, G. Gosset, F. Valle, and F. Bolivar.** 2000. A family of removable cassettes designed to obtain anibiotic-resistance-free genomic modifications of *Escherichia coli* and other bacteria. Gene **247:**255-264.

93. **Pansegrau, W., E. Lanka, P. T. Barth, D. H. Figurski, D. G. Guiney, D. Haas, D. R. Helinski, H. Schwab, V. A. Stanisich, and C. M. Thomas.** 1994. Complete nucleotide sequence of Birmingham IncP alpha plasmids. Compilation and comparative analysis. J Mol Biol **239:**623-63.

94. **Peel, D., and J. R. Quayle.** 1961. Microbial growth on C_1 compounds: 1. Isolation and characterization of *Pseudomonas* AM1. Biochem J **81:**465-469.

95. **Pomper, B. K., O. Saurel, A. Milon, and J. A. Vorholt.** 2002. Generation of formate by the formyltransferase/hydrolase complex (Fhc) from *Methylobacterium extorquens* AM1. FEBS Lett **523:**133-7.

96. **Pomper, B. K., and J. A. Vorholt.** 2001. Characterization of the formyltransferase from *Methylobacterium extorquens* AM1. Eur J Biochem **268:**4769-75.

97. **Pomper, B. K., J. A. Vorholt, L. Chistoserdova, M. E. Lidstrom, and R. K. Thauer.** 1999. A methenyl tetrahydromethanopterin cyclohydrolase and a methenyl tetrahydrofolate cyclohydrolase in *Methylobacterium extorquens* AM1. Eur J Biochem **261:**475-80.

98. **Purdy, D., T. A. O'Keeffe, M. Elmore, M. Herbert, A. McLeod, M. Bokori-Brown, A. Ostrowski, and N. P. Minton.** 2002. Conjugative transfer of clostridial shuttle vectors from *Escherichia coli* to *Clostridium difficile* through circumvention of the restriction barrier. Mol Microbiol **46:**439-52.

99. **Rabinowitz, J. C., and J. Pricer, W. E.** 1963. Formyltetrahydrofolate synthetase. Methods Enzymol **6:**375-379.

100. **Radfar, R., R. Shin, G. M. Sheldrick, W. Minor, C. R. Lovell, J. D. Odom, R. B. Dunlap, and L. Lebioda.** 2000. The crystal structure of N^{10}-formyltetrahydrofolate synthetase from *Moorella thermoacetica*. Biochemistry **39:**3920-6.

101. **Ras, J., W. N. Reijnders, R. J. Van Spanning, N. Harms, L. F. Oltmann, and A. H. Stouthamer.** 1991. Isolation, sequencing, and mutagenesis of the gene encoding cytochrome c_{553i} of *Paracoccus denitrificans* and characterization of the mutant strain. J Bacteriol **173:**6971-9.

102. **Ras, J., P. W. Van Ophem, W. N. Reijnders, R. J. Van Spanning, J. A. Duine, A. H. Stouthamer, and N. Harms.** 1995. Isolation, sequencing, and mutagenesis of the gene encoding NAD- and glutathione-dependent formaldehyde dehydrogenase (GD-FALDH) from *Paracoccus denitrificans*, in which GD-FALDH is essential for methylotrophic growth. J Bacteriol **177:**247-51.

103. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular Cloning: a Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

104. **Schmidhauser, T. J., and D. R. Helinski.** 1985. Regions of broad-host-range plasmid RK2 involved in replication and stable maintenance in nine species of gram-negative bacteria. J Bacteriol **164:**446-55.

105. **Schworer, B., J. Breitung, A. R. Klein, K. O. Stetter, and R. K. Thauer.** 1993. Formylmethanofuran: tetrahydromethanopterin formyltransferase and N_5,N_{10}-

methylenetetrahydromethanopterin dehydrogenase from the sulfate-reducing *Archaeoglobus fulgidus*: similarities with the enzymes from methanogenic Archaea. Arch Microbiol **159:**225-32.

106. **Scott, J. W., and M. E. Rasche.** 2002. Purification, overproduction, and partial characterization of beta-RFAP synthase, a key enzyme in the methanopterin biosynthesis pathway. J Bacteriol **184:**4442-8.

107. **Simon, R., U. Priefer, and A. Puhler.** 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram negative bacteria. Bio/Technology **1:**784-791.

108. **Skold, O., and A. Widh.** 1974. A new dihydrofolate reductase with low trimethoprim sensitivity induced by an R factor mediating high resistance to trimethoprim. J Biol Chem **249:**4324-5.

109. **Smith, L. M., W. G. Meijer, L. Dijkhuizen, and P. M. Goodwin.** 1996. A protein having similarity with methylmalonyl-CoA mutase is required for the assimilation of methanol and ethanol by *Methylobacterium extorquens* AM1. Microbiology **142 (Pt 3):**675-84.

110. **Spies, H. S., and D. J. Steenkamp.** 1994. Thiols of intracellular pathogens. Identification of ovothiol A in *Leishmania donovani* and structural analysis of a novel thiol from *Mycobacterium bovis*. Eur J Biochem **224:**203-13.

111. **Springer, A. L., A. J. Auman, and M. E. Lidstrom.** 1998. Sequence and characterization of *mxaB*, a response regulator involved in regulation of methanol oxidation, and of *mxaW*, a methanol-regulated gene in *Methylobacterium extorquens* AM1. FEMS Microbiol Lett **160:**119-24.

112. **Springer, A. L., H. H. Chou, W. H. Fan, E. Lee, and M. E. Lidstrom.** 1995. Methanol oxidation mutants in *Methylobacterium extorquens* AM1: identification of new genetic complementation groups. Microbiology **141 (Pt 11):**2985-93.

113. **Springer, A. L., C. J. Morris, and M. E. Lidstrom.** 1997. Molecular analysis of *mxbD* and *mxbM*, a putative sensor-regulator pair required for oxidation of methanol in *Methylobacterium extorquens* AM1. Microbiology **143 (Pt 5):**1737-44.

114. **Stolyar, S., M. Franke, and M. E. Lidstrom.** 2001. Expression of individual copies of *Methylococcus capsulatus* bath particulate methane monooxygenase genes. J Bacteriol **183:**1810-2.

115. **Studer, A., C. McAnulla, R. Buchele, T. Leisinger, and S. Vuilleumier.** 2002. Chloromethane-induced genes define a third C_1 utilization pathway in *Methylobacterium chloromethanicum* CM4. J Bacteriol **184:**3476-3484.

116. **Tatusov, R. L., M. Y. Galperin, D. A. Natale, and E. V. Koonin.** 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res **28:**33-6.

117. **Thauer, R. K.** 1998. Biochemistry of methanogenesis: a tribute to Marjory Stephenson. 1998 Marjory Stephenson Prize Lecture. Microbiology **144 (Pt 9):**2377-406.

118. **Toyama, H., C. Anthony, and M. E. Lidstrom.** 1998. Construction of insertion and deletion *mxa* mutants of *Methylobacterium extorquens* AM1 by electroporation. FEMS Microbiol Lett **166:**1-7.

119. **Vagner, V., E. Dervyn, and S. D. Ehrlich.** 1998. A vector for systematic gene inactivation in *Bacillus subtilis*. Microbiology **144 (Pt 11):**3097-104.

120. **van Ophem, P. W., J. Van Beeumen, and J. A. Duine.** 1992. NAD-linked, factor-dependent formaldehyde dehydrogenase or trimeric, zinc-containing, long-chain alcohol dehydrogenase from *Amycolatopsis methanolica*. Eur J Biochem **206:**511-8.

121. **Vannelli, T., M. Messmer, A. Studer, S. Vuilleumier, and T. Leisinger.** 1999. A corrinoid-dependent catabolic pathway for growth of a *Methylobacterium* strain with chloromethane. Proc Natl Acad Sci USA **96:**4615-4620.

122. **Vannelli, T., A. Studer, M. Kertesz, and T. Leisinger.** 1998. Chloromethane metabolism by *Methylobacterium* sp. Strain CM4. Appl Environ Microbiol **64:**1933-1936.

123. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19:**259-68.

124. **Vorholt, J. A.** 2002. Cofactor-dependent pathways of formaldehyde oxidation in methylotrophic bacteria. Arch Microbiol **178:**239-49.

125. **Vorholt, J. A., L. Chistoserdova, M. E. Lidstrom, and R. K. Thauer.** 1998. The NADP-dependent methylene tetrahydromethanopterin dehydrogenase in *Methylobacterium extorquens* AM1. J Bacteriol **180:**5351-6.

126. **Vorholt, J. A., L. Chistoserdova, S. M. Stolyar, R. K. Thauer, and M. E. Lidstrom.** 1999. Distribution of tetrahydromethanopterin-dependent enzymes in methylotrophic bacteria and phylogeny of methenyl tetrahydromethanopterin cyclohydrolases. J Bacteriol **181:**5750-7.

127. **Vorholt, J. A., C. J. Marx, M. E. Lidstrom, and R. K. Thauer.** 2000. Novel formaldehyde-activating enzyme in *Methylobacterium extorquens* AM1 required for growth on methanol. J Bacteriol **182:**6645-50.

128. **Weaver, C. A., and M. E. Lidstrom.** 1987. Isolation, complementation and partial characterization of mutants of the methanol autotroph *Xanthobacter* H4-14 defective in methanol dissimilation. J Gen Microbiol **133 (Pt 7):**1721-31.

129. **Weaver, C. A., and M. E. Lidstrom.** 1985. Methanol dissimilation in *Xanthobacter* H4-14: activities, induction and comparison to *Pseudomonas* AM1 and *Paracoccus denitrificans*. J Gen Microbiol **131 (Pt 9):**2183-97.

130. **Whitaker, J. R., and P. E. Granum.** 1980. An absolute method for protein determination based on the difference at 235 and 280 nm. Anal Biochem **109:**156-159.

131. **Whitta, S., M. I. Sinclair, and B. W. Holloway.** 1985. Transposon mutagenesis in *Methylobacterium* AM1 (*Pseudomonas* AM1). J Gen Microbiol **131:**1547-1549.

132. **Windass, J. D., M. J. Worsey, E. M. Pioli, D. Pioli, P. T. Barth, K. T. Atherton, E. C. Dart, D. Byrom, K. Powell, and P. J. Senior.** 1980. Improved conversion of methanol to single-cell protein by *Methylophilus methylotrophus*. Nature **287:**396-401.

133. **Xu, H. H., M. Viebahn, and R. S. Hanson.** 1993. Identification of methanol-regulated promoter sequences from the facultative methylotrophic bacterium *Methylobacterium organophilum* XX. J Gen Microbiol **139 (Pt 4):**743-52.

134. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103-19.

135. **Ziegelin, G., J. P. Furste, and E. Lanka.** 1989. TraJ protein of plasmid RP4 binds to a 19-base pair invert sequence repetition within the transfer origin. J Biol Chem **264:**11989-94.

136. **Zukowski, M. M., D. F. Gaffney, D. Speck, M. Kauffmann, A. Findeli, A. Wisecup, and J. P. Lecocq.** 1983. Chromogenic identification of genetic regulatory signals in *Bacillus subtilis* based on expression of a cloned *Pseudomonas* gene. Proc Natl Acad Sci U S A **80:**1101-5.

Vita

CHRISTOPHER JAMES MARX

ADDRESS University of Washington
Department of Microbiology
Box 357242
Seattle, WA 98195
(206) 543-0355
tofer@u.washington.edu

DATE OF BIRTH March 31, 1976 (San Luis Obispo, CA)

EDUCATION

1997 - 2003 Ph.D. Microbiology (May 30, 2003)
University of Washington

1993 – 1997 S.B. Biology, S.B. Environmental Engineering (GPA = 5.0/5.0)
Massachusetts Institute of Technology

GRANTS, FELLOWSHIPS AND AWARDS

National Science Foundation Microbial Biology Postdoctoral Fellowship (2003-2005)

Helen Riaboff Whiteley Graduate Fellowship Award, Microbiology, U. Washington (2002)

Howard Hughes Medical Institute, Predoctoral Fellowships in Biological Sciences –
Alternate: Microbiology (1998)

National Science Foundation, Graduate Research Fellowship –
Honorable Mention: Microbiology (1998), Engineering (1997)

Richard Lee Russel Award, Civil and Environmental Engineering, M.I.T. (1997)

Ward and Dorothy Melville Summer Fellow, Marine Sciences Research Center,
State University of New York, Stonybrook (1996)

Sea Grant Research Fellowship, M.I.T. Sea Grant (1994, 1995)

Member: American Society for Microbiology, Sigma Xi Scientific Research Society, Union of Concerned Scientists, American Association for the Advancement of Science

POSITIONS HELD

Graduate Research Assistant. Formaldehyde metabolism in *Methylobacterium extorquens* AM1. Dr. Mary E. Lidstrom, Departments of Chemical Engineering and Microbiology, University of Washington. (Spring 1998 – Spring 2003)

Teaching Assistant: General Microbiology Laboratory (Fall, 1998), Bacterial Genetics (Winter, 1998)

Research Assistant/Undergraduate Thesis. Effects of light deprivation on the survival and recovery of two *Prochlorococcus* strains. Dr. Sallie W. Chisholm, Departments of Civil and Environmental Engineering and Biology, M.I.T. (Fall, 1994 – Spring, 1997)

Research Assistant. Aromatic amine dehydrogenase from *Alcaligenes faecalis*. Dr. Andrei Y. Chistoserdov, Marine Sciences Research Center, S.U.N.Y. Stonybrook. (Summer, 1996)

Research Assistant. Modeling populations of benthic macroinvertebrates as an indicator of watershed health. Dr. Graham Harris and Dr. Richard Norris, Cooperative Research Centre for Freshwater Ecology, University of Canberra, Australia. (Summer, 1995)

PRESENTATIONS

Posters: American Society for Microbiology, General Meeting. (2001, 2002)
Molecular Basis of Microbial One-Carbon Metabolism, GRC (2000, 2002)

Short talks: West Coast Bacterial Physiologists Asilomar Conference. (2001, 2002)
Interfacing Microbiology and Biotechnology, U. Florida. (2001)

PUBLICATIONS

Marx, C. J., and M. E. Lidstrom. Flux analysis reveals a metabolic loop critical for controlling formaldehyde toxicity and carbon partitioning in a methylotrophic bacterium. *In preparation.*

Marx, C. J., and M. E. Lidstrom. Development of an insertional vector system for *Methylobacterium extorquens* AM1 and generation of null mutants lacking *mtdA* and/or *fch. In preparation.*

Marx, C. J., M. Laukel, J. A. Vorholt, and M. E. Lidstrom. Purification of the formate-tetrahydrofolate ligase from *Methylobacterium extorquens* AM1 and demonstration of its requirement for methylotrophic growth. *In preparation.*

Marx, C. J., L. Chistoserdova, and M. E. Lidstrom. The formaldehyde detoxifying role of the tetrahydromethanopterin-linked pathway in *Methylobacterium extorquens* AM1. *In preparation.*

Marx, C. J., J. A. Miller, M. E. Lidstrom, and L. Chistoserdova. Redundancy of formaldehyde oxidation pathways in *Burkholderia fungorum* LB400. *In preparation.*

Van Dien, S. V., **C. J. Marx**, B. N. O'Brien, and M. E. Lidstrom. Genetic characterization of the carotenoid biosynthetic pathway in *Methylobacterium extorquens* AM1, and the isolation of a colorless mutant. *In preparation.*

Chistoserdova, L., C. Jenkins, M. G. Kalyuzhnaya, **C. J. Marx**, A. Lapidus, J. T. Staley, and M.E. Lidstrom. The enigmatic Planctomyces may hold a key to the origins of methanogenesis and methylotrophy. *Submitted.*

Marx, C. J., B. N. O'Brien, J. Breezee, and M. E. Lidstrom. 2003. Novel methylotrophy genes of *Methylobacterium extorquens* AM1 identified by using transposon mutagenesis including a putative dihydromethanopterin reductase. J. Bacteriol. 185:669-673.

Marx, C. J., and M. E. Lidstrom. 2002. Broad-host-range *cre-lox* system for antibiotic marker recycling in Gram-negative bacteria. BioTechniques. 33:1062-1067.

Marx, C. J., and M. E. Lidstrom. 2001. Development of improved versatile broad-host-range vectors for use in methylotrophs and other Gram-negative bacteria. Microbiology 147:2065-2075.

Vorholt, J. A., **C. J. Marx**, M. E. Lidstrom, and R. K. Thauer. 2000. Novel formaldehyde-activating enzyme in *Methylobacterium extorquens* AM1 required for growth on methanol. J. Bacteriol. 182:6645-6650.

Cohen-Kupiec, R., **C. J. Marx**, and J. A. Leigh. 1999. Function and regulation of *glnA* in the methanogenic archaeon *Methanococcus maripaludis*. J. Bacteriol. 181:256-261