Structure, Function and Metabolic Roles of IcmF-a Fusion Between the Radical B_{12} Enzyme and its G-protein Chaperone

by

Valentin F. Cracan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biological Chemistry)
in The University of Michigan
2012

Doctoral Committee:

Professor Ruma V. Banerjee, Chair
Professor Stephen W. Ragsdale
Professor David H. Sherman
Professor Janet L. Smith
Associate Professor Bruce A. Palfey

ACKNOWLEDGEMENTS

I am very fortunate to receive my PhD in biochemistry under the guidance of Dr. Ruma Banerjee. I would like to thank Dr. Banerjee for being my advisor and mentor for the lengthy duration of my graduate work and for her time and patience in helping me to complete my graduate studies. I am especially thankful to Dr. Banerjee for the gift of scientific freedom. My project took many unexpected turns over the years, and it was crucial that I was able to launch, with Dr. Banerjee's permission, multiple collaborations to supplement my understanding of the project.

When I joined the graduate program in biochemistry at the University of Nebraska-Lincoln in summer 2005 I didn't know that my graduate studies would have so many unexpected twists. After two years at Lincoln our laboratory was relocated to University of Michigan at Ann Arbor. This move resulted in the completion of a Masters Degree from University of Nebraska-Lincoln and my admittance to the PhD program at the University of Michigan.

I can say without a doubt that the educational atmosphere at the University of Michigan played a significant role in shaping me as a scientist. Ann Arbor is a town where every morning on the bus one is more likely to see people reading scientific articles than newspapers. I have found that different laboratories on campus are highly collaborative.

People at the university are extremely friendly and quite eager to discuss research everywhere on campus from the recreation center to the Hospital cafeteria.

I am also thankful to all members of our lab as well as Dr. Ragsdale's laboratories, past and present. For all these years the input from Dr. Ragsdale at our joint group meetings was irreplaceable and his ideas frequently helped me move forward in my research.

Especially, I am indebted to my colleagues who worked with me in the same sub-group on vitamin B_{12}-dependent enzymes and trafficking. Particular thanks go to Dr. Carmen Gherasim and Mike Lofgren.

I would like to separately mention Dr. Dominique Padovani who was a postdoctoral fellow when I joined the laboratory. Dr. Padovani' influence was always important, not only to me, but to all the members of our laboratory. For all the years that I interacted with Dr. Padovani he established high standards of laboratory etiquette and scientific rigor. Most particularly, I learned a lot from Dr. Padovani during the brief time we worked together on my project.

Another person from our laboratory whom I want to acknowledge separately is Dr. Victor Vitvitsky. Dr. Vitvitsky was very supportive in many aspects of my work. I learned much from Dr. Vitvitsky about laboratory equipment repairs. I must say that Dr. Vitvitsky and his wife Tatiana were very kind to me for all these years of study. Their kindness exceeded my life in the laboratory. I was always welcomed at their house where I ate a lot of delicious food.

I would also like to thank my committee members; Dr. Stephen Ragsdale, Dr. Janet Smith, Dr. David Sherman and Dr. Bruce Palfey for their advice and support over the course of the last four years. I am truly fortunate to have such an outstanding committee.

Dr. Palfey teaches BIOLCHEM 673 transient & steady-state kinetics, a class that is truly legendary at the University of Michigan. One cannot be considered a real enzymologist without having taken that class.

I would also like to thank the office personnel in the Department of Biological Chemistry, especially Beth, Julie and Prasanna for always being friendly and helpful.

I am also grateful to my collaborators from MIT, Dr. Catherine Drennan and Marco Jost. Dr. Drennan enabled the exciting collaboration to solve the structure of my protein and has been most helpful during my search for a post-doctorate placement.

I would like to mention another person who made my long stay in the US very pleasant. Dr. Wade Nelson, whom I met in summer 2004 when he was visiting Moldova, became my close friend. For all my years of graduate studies, Dr. Nelson and his family were very supportive of me.

Last but not least, I would like to express my heart felt thanks to my mother, Antonina Cracan, Aunt Alla Cracan and Uncle Pavel Epifanov. I thank my parents for their long distant but never wavering love, patience and faith in my ability to succeed in graduate education in the United States.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AdoCbi, 5'-deoxy-5'-adenosylcobinamide;

AdoCbl, 5'-deoxy-5'-adenosylcobalamin, adenosylcobalamin;

AMPPNP, adenosine 5'-(β, γ-imido)triphosphate;

ATR, ATP:Cob(I)alamin adenosyltransferase;

BDH, butyryl-CoA dehydrogenase;

Bkd, branched-chain ketoacid dehydrogenase complex;

Cbi, cobinamide;

Cbl, cobalamin;

CNCbl, cyanocobalamin;

CoA, coenzymeA;

CTT, casitone based complex medium used for *Myxococcus xanthus* growth;

DMB, dimethylbenzimidazole;

DTT, dithiothreitol;

ECM, ethylmalonyl-CoA mutase;

GAP, **G**TPase **a**ctivating **p**rotein;

GC, gas chromatography;

GEF, **g**uanine nucleotide **e**xchange **f**actor;

GMPPNP, guanosine 5'-(β,γ-imido)triphosphate;

HCM, 2-hydroxyisobutyryl-CoA mutase;

HPLC, high performance liquid chromatography;

ICM, isobutyryl-CoA mutase;

IcmA, substrate-binding, large subunit of ICM;

IcmB, B_{12}-binding, small subunit of ICM;

IcmF, **i**sobutyryl-**C**oA **m**utase **f**used;

IPTG, isopropyl-1-thio-β-D-galactopyranoside;

ITC, Isothermal Titration Calorimetry;

LIC, ligase independent cloning;

MCM, methylmalonyl-CoA mutase;

MMAA, for **m**ethy**m**alonic **a**ciduria linked to the cbl**A** complementation group;

OH_2Cbl, aquacobalamin;

OHCbl, hydroxocobalamin;

PCM, pivalyl-CoA mutase;

PCR, polymerase chain reaction;

PKSs, polyketide synthases;

SIMIBI, after **si**gnal recognition particle **Mi**nD and **Bi**oD;

STRING, **S**earch **T**ool for the **R**etrieval of **In**teracting **G**enes/Proteins;

TCEP, Tris(2-carboxyethyl)phosphine hydrochloride;

TIM, triosephosphate isomerase;

UV-vis, ultraviolet visible.

ASTRACT

Coenzyme B_{12} is a biologically active form of vitamin B_{12} and used in Nature as a radical reservoir to catalyze chemically challenging transformations. The loading of B_{12}-enzymes with the correct cofactor form is critically important for their function and is gated by chaperones that use the chemical energy of GTP hydrolysis to ensure the fidelity of the process. Two highly similar coenzyme B_{12}-dependent enzymes that catalyze carbon skeleton rearrangements, methylmalonyl-CoA mutase (MCM) and isobutyryl-CoA mutase (ICM), are widely distributed in bacteria. MCM catalyzes the isomerization of methylmalonyl-CoA to succinyl-CoA while ICM catalyzes the reversible interconversion of isobutyryl-CoA and n-butyryl-CoA. Curiously, a variant, IcmF, is found in >80 bacterial species. Although IcmF was misannotated as an MCM variant in the database, I have demonstrated by expressing four bacterial IcmFs that it is an active ICM, fused to its chaperone. This discovery expands the known distribution of ICM activity well beyond the genus *Streptomyces* where it is involved in polyketide biosynthesis.

Subsequently we have discovered that IcmF catalyzes a novel coenzyme B_{12}-dependent 1,2-rearrangement of isovaleryl-CoA and pivalyl-CoA (2,2-dimethylpropionyl-CoA). Biochemical experiments demonstrate that in an IcmF in which the base specificity loop motif, NKxD is modified to NKxE, catalyzes the hydrolysis of both GTP and ATP. IcmF is susceptible to rapid inactivation during turnover and GTP confers protection, but only during utilization of isovaleryl-CoA as substrate.

I have characterized the mutase and GTPase activities in intact and truncated versions of IcmF lacking the B_{12}- or the substrate-binding domains to investigate interactions between the domains. I have demonstrated that adenosyltransferase (ATR), which synthesizes coenzyme B_{12}, also transfers the cofactor to IcmF.

To gain insights into the metabolic role of IcmF, we have disrupted the *icm*F gene in *Myxococcus xanthus* in collaboration with Dr. Montserrat Elias-Arnanz (Universidad de Murcia, Spain) and in *Ralstonia eutropha* H16 in collaboration with Dr. Antony Sinskey (MIT) and currently are testing the resulting phenotypes under different growth conditions.

CHAPTER 1

Introduction

The focus of this thesis is on the enzymology of a newly discovered protein, IcmF, which represents a fusion between a coenzyme B_{12}-dependent isobutyryl-CoA mutase and a G protein. This fusion protein was incorrectly annotated in databases as a fusion between methylmalonyl-CoA mutase and its G protein chaperone. We have shown this annotation to be erroneous by demonstrating that this fusion protein does not catalyze the interconversion of methylmalonyl-CoA and succinyl-CoA but rather, the interconversion of isobutyryl-CoA and n-butyryl-CoA. Subsequently we have discovered that this protein catalyzes a novel coenzyme B_{12}-dependent 1,2-rearrangement of isovaleryl-CoA and pivalyl-CoA (2,2-dimethylpropionyl-CoA). In this thesis, I present an in-depth study of the IcmF-catalyzed reactions as well as the effect of its G protein domain on the mutase activity.

1.1 Isobutyryl-CoA mutase

Isobutyryl-CoA mutase (ICM) (EC 5.4.99.13) is a coenzyme B_{12} or AdoCbl-dependent enzyme, which catalyses the reversible interconversion of isobutyryl-CoA and n-butyryl-CoA (1-3) (Figure 1.1). ICM activity was discovered by John Robinson and colleagues in cell extracts of the gram-positive, filamentous soil bacterium, *Streptomyces cinnamonensis,* by monitoring the reaction using ^{1}H NMR (4). The reaction catalyzed by ICM is very similar to that catalyzed by methylmalonyl-CoA mutase (MCM), which is better studied and more widely distributed (Figure 1.1) (3). In both reactions, a 1,2-

rearrangement of the carbon skeleton takes place where the –(CO)S-CoA substituent and a H atom exchange positions on vicinal carbons (Figure 1.1) (*3, 5, 6*). Interestingly in their first publication on ICM, Robinson and colleagues predicted that this activity is catalyzed by an enzyme which is very similar in primary sequence to MCM, but where a basic amino acid residue (for binding of the carboxyl group in the substrate for MCM) is substituted by a hydrophobic one (for binding of the methyl group in the substrate for ICM) (4). This occurrence of subtle changes at key active site residues to accommodate different substrate specificities is central to the discussion in this thesis on the differences and similarities between members of this subfamily of carbon-skeleton rearranging AdoCbl-dependent mutases.

Besides several bacteria belonging to the *Streptomyces* genus, where the role of ICM in the synthesis of polyketides is well established (7), the presence of ICM-like proteins in other bacteria was neither biochemically nor genetically confirmed. Interestingly, ICM-like activity appears to be involved in the conversion of isobutyrate to butyrate in numerous anaerobic bacteria and cultures enriched in methanogenic bacteria (8-11). Oude Elferink and colleagues have shown, using ^{13}C-labelled butyrate, that several sulfate reducers (*Desulforhabdus amnigenus*, *Desulfobacterium vacuolatum*, *Desulfoarculus baarsii*, *Desulfotomaculum sp.*) are capable of isomerization of butyrate to isobutyrate (9). Also isomerization of butyrate and isobutyrate was reported in the anaerobe, *Pelospora glutarica* WoG13 (*12, 13*), and in a thermophilic bacterium, *Synthrophothermus lipocalidus* (14).

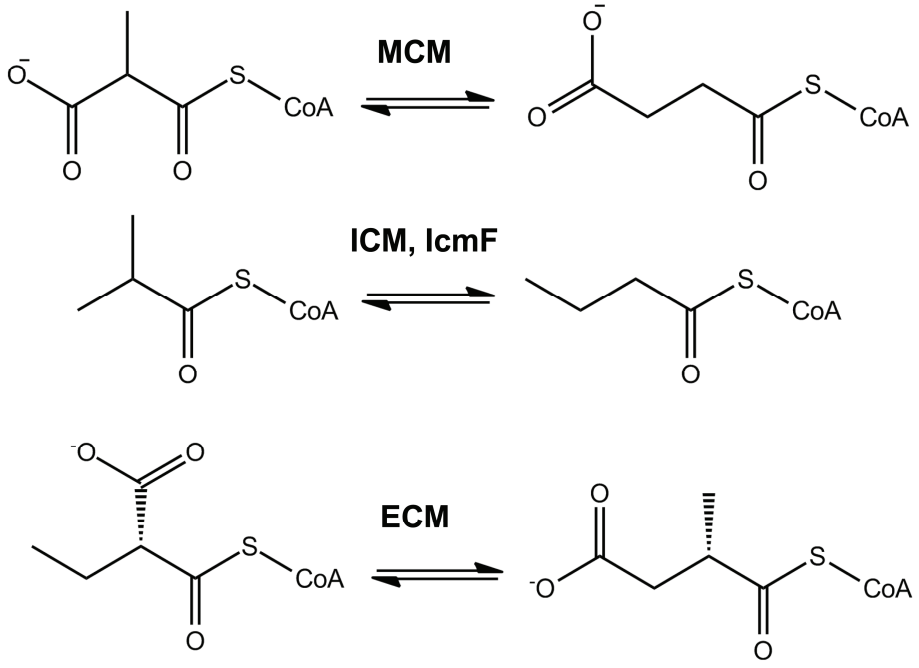


Figure 1.1: Reactions catalyzed by MCM, ICM, IcmF and ECM.

Unfortunately, for most of these bacteria, full genomic sequences are not available and therefore it is not possible to assess whether or not they actually encode ICM (See also section 5.3.1).

1.2 Properties of ICM from *S. cinnamonensis*

The genes encoding ICM were first cloned and sequenced from *S. cinnamonensis* (*1, 2*). ICM is an $\alpha_2\beta_2$-heterotetramer composed of two large subunits (IcmA) of 62.5 kDa and two small subunits (IcmB) of 14.3 kDa. The apparent native molecular weight of the AdoCbl-bound holo-enzyme is 152 kDa as estimated by gel filtration chromatography (1). Remarkably, subunits of ICM have very low affinity to each other in the absence of AdoCbl. When a protein sample lacking cofactor (apo-enzyme) is subjected to gel filtration chromatography, formation of the $\alpha_2\beta_2$-heterotetramer is not observed and the subunits elute separately (1).

The sequence of the large ICM subunit, IcmA from *Streptomyces cinnamonensis*, has 43% identity and 61% similarity with the large subunit of MCM from the same organism, and 43% identity and 65% similarity with MCM from *Propionibacterium shermanii* (1). IcmA subunit can be viewed as a truncated form of the large subunit of MCM lacking the C-terminal AdoCbl-binding domain (Figure 1.2).

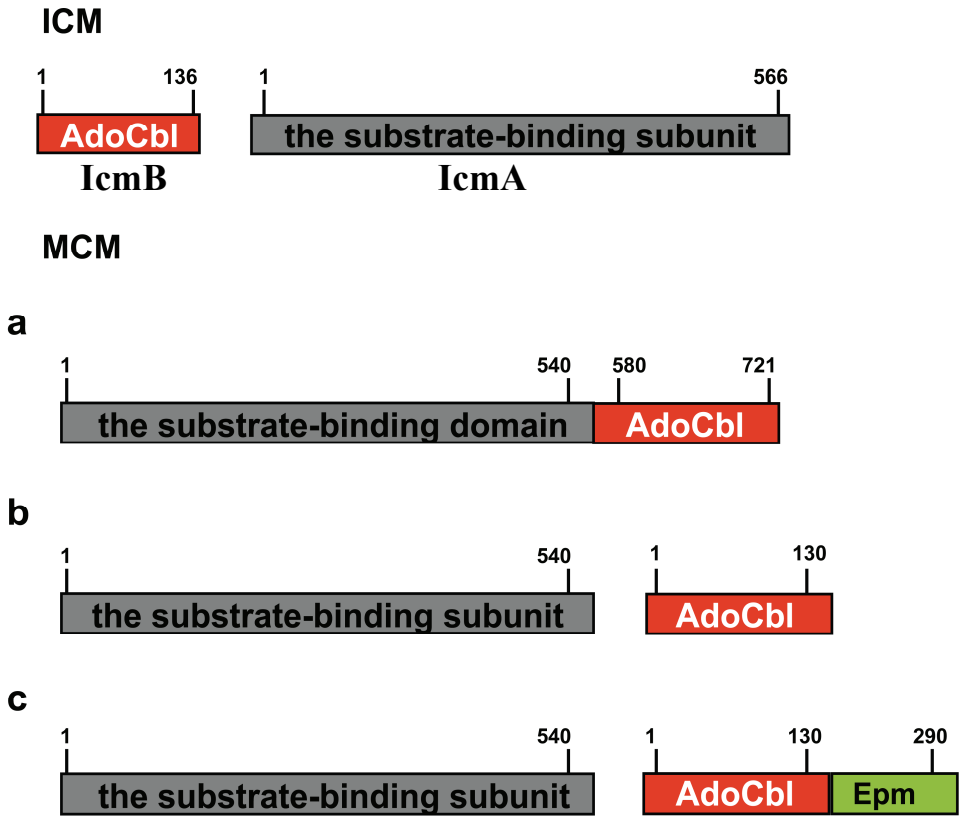


Figure 1.2: Domain organization of the genes encoding ICM and MCM. In ICMs the AdoCbl-binding and the substrate-binding domains are always found as two separate polypeptides (IcmB and IcmA). In vast majority of MCMs, the AdoCbl-binding and the substrate-binding domains are found on a single subunit (**a**). In several organisms, MCMs are organized similarly to ICM (**b**), where the AdoCbl-binding and the substrate-binding domains are found as two separate polypeptides and sometimes, the AdoCbl-binding domain is fused to methylmalonyl-CoA epimerase (Epm) (**c**). For the accession numbers see section 1.6.

The small subunit, IcmB has high sequence similarity to the AdoCbl-binding domains of the following coenzyme B_{12}-dependent enzymes: MCM from *P. shermanii* (58%),

glutamate mutase from *Clostridium cochlearum* (49%) and 2-methyleneglutarate mutase from *C. barkeri* (57%) (1). A striking organizational difference between ICM and MCM is that the genes encoding the ICM subunits, *icmA* and *icmB,* are not adjacent to each other in the genome of *S. cinnamonensis*. In contrast, the genes encoding MCM in many bacteria are usually located within the same operon, and sometimes their reading frames even overlap, making identification of putative MCM-like proteins a relatively easy task (1).

Initially, purification of ICM-like activity from *S. cinnamonensis* resulted in a partially purified protein with molecular weight of ~ 65 kDa (2). The specific activity of that preparation was very low (S.A.= 2.3 x 10^{-2} μmol min^{-1} mg^{-1} with n-butyryl-CoA as a substrate) (2). Authors used a GC-based assay to monitor the interconversion of n-butyryl-CoA and isobutyryl-CoA. That assay involved saponification of the thioesters, extraction of the corresponding acids (isobutyric and n-butyric) into ethylacetate and separation of free acids by GC (*2, 15*).

The ready loss of the small subunit of ICM during purification explained the low enzymatic activity. The initial study on ICM led to the cloning and purification of the large subunit (IcmA) (2). Subsequently, Robinson and colleagues identified and isolated the small subunit (IcmB) by adding purified 6xHis-tagged IcmA to cell extracts of *S. cinnamonensis* or *S. lividans* supplemented with AdoCbl. After a two-step purification, using affinity chromatography on Ni-NTA and anion exchange chromatography on a MonoQ column, the small subunit of ICM was isolated. IcmB has a molecular mass of ~ 17 kDa and, when mixed with recombinant IcmA and AdoCbl, yields active holo-protein (S.A.=1 μmol min^{-1} mg^{-1} with n-butyryl-CoA) (2). In subsequent studies, recombinant

IcmA and IcmB subunits were expressed in *E. coli* and purified separately (1). The activity of reconstituted protein was dependent on the ratio of IcmA:IcmB. The highest activity with n-butyryl-CoA (S.A.= 38 ± 3 μmol min^{-1} mg^{-1} of IcmA) was seen when IcmB was added in a 5-fold molar excess over IcmA. Assuming one active site per IcmA monomer, the k_{cat} was determined to be 39 ± 3 s^{-1} (1).

Reconstitution of active ICM made possible characterization of the other steady-state kinetic parameters for the enzyme. The K_M values for isobutyryl-CoA and butyryl-CoA were reported to be 57 ± 13 μM and 54 ± 12 μM, respectively (1). The equilibrium of this reaction is 1.7 in favor of isobutyryl-CoA. The K_{act} for AdoCbl was 12 ± 2 μM (1).

1.3. Organization of the *icm* genes

As mentioned above, ICM is a heterotetramer, comprised of two types of subunits, a large IcmA subunit and a small IcmB subunit (1). The sequences of all known IcmAs are very similar to the sequences of the large subunit of MCM, with the exception of the AdoCbl-binding region, which is missing in IcmA. Since IcmA does not have a C-terminal AdoCbl-binding domain with the canonical DxHxxG motif (16), this protein requires a separate small IcmB subunit to bind the cofactor (Figure 1.2). In this respect, ICM resembles other AdoCbl-dependent mutases that exhibit a similar organization. For example, glutamate mutase is also composed of two subunits: the large subunit MutE, which contains the substrate binding site and a small subunit MutS, with the AdoCbl-binding site (*17, 18*).

1.4. Insertional inactivation studies on ICM: Influence on polyketide antibiotic biosynthesis and the role of ethylmalonyl-CoA mutase (ECM)

Substantial effort has been dedicated to elucidation of the pathways involved in the biosynthesis of polyketides in the genus *Streptomyces* (7). These studies highlighted the importance of the ICM-catalyzed reaction for the supply of methylmalonyl-CoA to polyketide synthases (PKSs) in addition to its role in valine and fatty acid catabolism (7). PKSs, which catalyze polyketide biosynthesis, use several acyl thioesters as extender units, where the most abundant are malonyl-CoA, methylmalonyl-CoA and ethylmalonyl-CoA (*19, 20*). Ethylmalonyl-CoA and malonyl-CoA are products of carboxylation of butyryl-CoA and acetyl-CoA, whereas several routes for methymalonyl-CoA production are possible (*21*), as discussed below. Gene disruption studies coupled with NMR identification of ^{13}C-labeled precursors, have been widely used for the determination of the metabolic routes for biosynthesis of the polyketides monensin A and B in *S. cinnamonensis* (7). *S. cinnamonensis* is a useful bacterium for addressing the question about the preferred route for methylmalonyl-CoA formation, since both ICM and MCM in this organism were identified and cloned, and both mutases can be inactivated by inserting a hydromycin resistance gene into the large subunits of ICM or MCM (Δ*icmA::hygB*, Δ*mutB::hygB*) (7).

Until recently, three major pathways for the synthesis of methylmalonyl-CoA were known in *S. cinnamonensis* (*21*). In the first pathway, methylmalonyl-CoA is formed from succinyl-CoA in the isomerization reaction catalyzed by MCM (15). The second possibility is carboxylation of propionyl-CoA by propionyl-CoA carboxylase to (S)-methylmalonyl-CoA which in turn is converted to (R)-methylmalonyl-CoA by epimerase.

Methymalonyl-CoA has been shown to be formed in a multistep pathway, which involves interconversion of isobutyryl-CoA (*21*).

After it was shown that both mutants (Δ*mutB::hygB)* and (Δ*icmA::hygB)* were able to incorporate butyrate and acetoacetate into the propionate units in monensin A, it became evident that *S. cinnamonensis* possesses another route for forming methylmalonyl-CoA that is independent of both ICM and MCM (7). MeaA, a protein of unknown function at that time, with high sequence similarity to MCM, was proposed as the enzyme that allowed the formation of methylmalonyl-CoA independently of ICM and MCM (*7, 21*). When both ICM and MeaA genes were deactivated in *S. cinnamonensis*, the mutant was unable to produce monensin A and B, whereas the single mutants (Δ*mut::hygB* and Δ*icmA::hygB*) produced monensin at levels comparable to those in wild-type bacteria (7). It turned out that MeaA in *S. cinnamonensis* and *M. extorquens* is ethylmalonyl-CoA mutase (ECM), a recently discovered acyl-CoA mutase (Figure 1.1) (*22, 23*). Ethylmalonyl-CoA mutase catalyzes the interconversion of ethylmalonyl-CoA and methylsuccinyl-CoA and operates in a recently-discovered ethylmalonyl-CoA pathway for acetate assimilation (*22*). In this pathway, three molecules of acetyl-CoA, one molecule of CO_2 and one molecule of bicarbonate form malate and succinyl-CoA (*22*).

Since the ethylmalonyl-CoA pathway allows the formation of propionyl-CoA from acetoacetyl-CoA, previous findings from ^{13}C-labeling experiments in *S. cinnamonensis* can now be elegantly explained. Moreover ECMs are found in other bacteria belonging to the genus *Streptomyces*: *S. avermitilis* and *S. coelicolor* (*21, 22*).

1.5 Methylmalonyl-CoA mutase: Reaction and metabolic significance

MCM is an a coenzyme B_{12}-dependent enzyme that catalyzes the reversible interconversion between (2*R*)-methylmalonyl-CoA and succinyl-CoA (*3, 24-26*) (Figure 1.1). In fact, this is the only AdoCbl-dependent mutase that is found both in bacteria and animals, where this enzyme is located in the mitochondria.

In numerous studies, different aspects of the reaction mechanism of MCM have been elucidated (*3, 25*). The catalytic cycle of MCM is a good model for ICM since both enzymes are expected to share a similar reaction mechanism (*3, 5, 6*) (Figure 1.3).

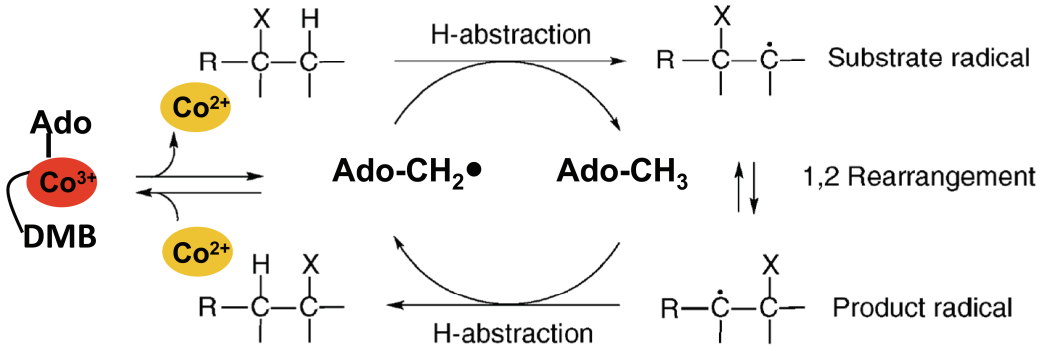


Figure 1.3: Generalized scheme for the 1,2-rearrangement catalyzed by AdoCbl-dependent mutases. The binding of substrate to the enzyme induces homolytic cleavage of the cobalt-carbon bond of AdoCbl which results in formation of a pair of radicals: 5'-deoxyadenosyl and cob(II)alamin. The 5'-deoxyadenosyl radical abstracts a specific H-atom from the substrate to form 5'-deoxyadenosine and a substrate radical, which subsequently undergoes isomerization to the product radical. The latter re-abstracts a H-atom from the methyl group of 5'-deoxyadenosine to form the product and regenerates the 5'-deoxyadenosyl radical. Finally, cob(II)alamin and the 5'-deoxyadenosyl radicals reform AdoCbl.

The reaction catalyzed by MCM starts with the homolytic cleavage of the cobalt-carbon bond in AdoCbl, which generates cob(II)alamin and a 5'-deoxyadenosyl radical. Following this, the highly reactive 5'-deoxyadenosyl radical abstracts a H atom from the substrate, which leads to the formation of a 5'-deoxyadenosine and a substrate radical. In

the last part of the reaction, rearrangement of a substrate radical via a cyclopropylcarbinyl intermediate leads to formation of a product radical, which after H atom re-abstraction from 5'-deoxyadenosine, gives product. Finally, the product is released and the two cofactor radicals, cob(II)alamin and 5'-deoxyadenosyl reform AdoCbl, completing the turnover cycle (*3, 26*) (Figure 1.3). The dramatic rate acceleration (~10^{12} -fold) of homolysis of the cobalt-carbon bond is induced upon substrate binding (*27*). During the reaction catalyzed by acyl-CoA mutases and other AdoCbl-dependent enzymes, like the aminomutases and diol dehydratase, the reactive radicals that are formed can undergo unproductive side reactions precluding the re-formation of AdoCbl and leading to enzyme inactivation (*28-33*).

MCM plays an indispensable role in the propionate pathway in mammals, where propionyl-CoA, derived from catabolism of branched-chain amino acids, odd-chain fatty acids and cholesterol, is converted to succinyl-CoA (*3, 24, 34*). In humans, genetic defects in MCM cause methylmalonic aciduria, an autosomal recessive metabolic disorder (*35, 36*). Methylmalonic aciduria also can result from mutations in genes that impair B_{12} trafficking, i.e. the assimilation and delivery of cofactor to its target enzymes, cytosolic methionine synthase and mitochondrial MCM (*35, 37*).

In bacteria on the other hand, the reaction catalyzed by MCM is important in the reverse metabolic direction, linking production of propionate and succinate (*38*). In the genus *Streptomyces,* the MCM-catalyzed reaction is important in the formation of methylmalonyl-CoA, a building block for polyketide biosynthesis as mentioned above (7)(Section 1.4).

1.6 Organization of the *mcm* genes

The vast majority of bacterial MCMs are αβ heterodimers, where the catalytic subunit binds the cofactor and the substrate (Figure 1.2). The two subunits of bacterial MCM are related in sequence suggesting that they are products of a gene duplication event (*25*). The α subunit has a C-terminal AdoCbl binding domain with the canonical "DxHxxG" cofactor-binding motif (16). Based on the crystal structure of the αβ heterodimer of MCM from *P. shermani* (*39*), the inactive subunit appears to serve a structural role. Genomic analyses show that heterodimeric MCMs are found in the vast majority but not all bacteria and usually, both MCM subunits are found in the same operon.

In *E. coli,* the *sbm* (sleeping beauty mutase) gene product was shown to encode MCM, which is a functional homodimeric enzyme (*40, 41*). Sbm is not the only example of a bacterial MCM with homodimeric organization. The enzyme from *Sinorhizobium meliloti* was shown to be a homodimer as well (*42*).

Notably, in some Archaea (for example in *Metallosphaera sedula* (YP_001190737, YP_001192119), *Sulfolobus solfataricus* (NP_343779, NP_343640), *S. tokodaii* (NP_376440, NP_378091), *S. islandicus* (YP_002913536, YP_002828266)) and in some bacteria (for example in *Petrotoga mobilis* (YP_001567606, YP_001567607), *Thermoanaerobacter sp.* (YP_001663471, YP_001663470), *Ilyobacter polytropus* (YP_003968833, YP_003968832), *Thermoanaerobacter tengcongensis* (NP_622843, NP_622844), *Clostridium sp. OhILAs* (YP_001511994, YP_001511995), *Veillonella parvula* (YP_003312208, YP_003312207), *Selenomonas sputigena* (ZP_05899584, ZP_05899583), *Desulfobacca acetoxidans* (YP_004369635, YP_004369636), *Aminobacterium colombiense* (YP_003553744, YP_003553743), *Thermanaerovibrio*

acidaminovorans (YP_003316807, YP_003316808), *Aminomonas paucivorans* (VIMSS IDs 11093272, 11093273), *Alkaliphilus metalliredigenes* (VIMSS IDs 6859005, 6859004), *Cloacibacillus evryensis* (VIMSS IDs 11182894, 11182893), *Thermotogales bacterium* (YP_002941070, YP_002941069)), the substrate-binding and the AdoCbl-binding domains of MCM are not found in a single polypeptide, but rather reside on separate subunits of ~63 kDa and ~15 kDa molecular mass (Figure 1.2). Sometimes, these two genes are not located close to each other (Figure 1.2). It is expected that the corresponding MCMs will be heterotetramers comprised of two substrate-binding subunits and two AdoCbl-binding subunits similarly to $\alpha_2\beta_2$-heterotetrameric organization of ICM. MCMs with such organization were shown to catalyze the interconversion of methylmalonyl-CoA to succinyl-CoA as a part of a 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway discovered in *Metallosphaera sedula* and in a number of *Sulfolobus* species (*43, 44*).

In addition to the arrangement of the small and large subunits of MCM described above, in several organisms the small B_{12}-binding subunit of MCM is fused to methylmalonyl-CoA epimerase, with a predicted molecular mass of ~30 kDa (Figure 1.2). The latter interconverts (*2S*)-methylmalonyl-CoA to (*2R*)-methylmalonyl-CoA. The latter is used as a substrate by MCM. This organization is found in *Caldalkalibacillus thermarum* (ZP_08532462, ZP_08532461), *Brevibacillus laterosporus* (ZP_08642783, ZP_08642782, ZP_08642783), *Bacillus tusciae* (YP_003589180, YP_003589181), *Ornithinibacillus sp.* (ZP_08783339, ZP_08783338), and *Bacillus selenitireducens* (ZP_02170815, YP_003700315).

In the photosynthetic coccolithophorid alga *Pleurochrysis carterae* (*45*), in round worm *Ascaris lumbricoides* (*46*), and in mammals, MCM is an α_2 homodimer, comprised of two ~75 kDa subunits. The crystal structure of human MCM (hMCM), reported recently (*47*) shows that each subunit houses an active site. Previously, based on genetic complementation studies, it was predicted that each active site of hMCM is formed by both subunits ("head-to-tail organization") *(47, 48).*

1.7 P-loop GTPases from the G3E family of metallochaperones

GTPases, proteins that bind and hydrolyze GTP, are crucial for many aspects of life (*49-51*). The main feature of P-loop GTPases and related ATPases is the presence of the mononucleotide-binding fold (called the G or GTPase domain) that is found in an estimated 10-18% of all gene products in most organisms (*50, 51*). G domains can be found fused to numerous protein domains, which makes members of this diverse group of proteins key players in countless cellular processes (*51*). According to the classification developed by Koonin and colleagues, P-loop GTPases and related ATPases can be divided into two large classes: TRAFAC (named after **tra**nslation **fac**tor-related) and SIMIBI (named after its 3 largest subgroups, the **si**gnal recognition particle, **Mi**nD and **Bi**oD superfamilies) (*50*).

In the past few decades, a number of P-loop GTPases from the SIMIBI class were ascribed as chaperones involved in the assembly of target metalloenzymes. Most of these chaperones belong to two families within the SIMIBI class: the Mrp/MinD and the G3E. The Mrp/MinD family is characterized by the consensus GKGGxGK[ST] Walker A motif which is also known as G1 motif (*50*). Here, lysine in the KGG portion of the motif was shown to be crucial for catalysis (*50*). Interestingly, the base specificity NKxD motif

(or G4 motif), which is conserved in many GTPases, is never found in the intact form in members of Mrp/MinD family. The aspartate in the NKxD motif provides specificity for guanine binding and none of these proteins has been shown to have GTPase activity (*51*). The most studied chaperones from the Mrp/MinD family are NifH and CooC. NifH is involved in an ATP-dependent maturation of the iron-molybdenum cofactor (FeMo-Co) of dinitrogenase (*52*). CooC proteins are ATPases, which participate in the incorporation of nickel into the active site of Ni,Fe-carbon monoxide (CO) dehydrogenase (*53-55*).

Phylogenetic analysis performed by Koonin and colleagues defined the G3E family, so termed because the members of this family have glutamic acid (E) in the Walker B (G3) motif and the GxxGxGK[ST] variant in the Walker A motif. Most importantly, all members of the G3E family have an intact base specificity NKxD motif, which makes GTP a preferable substrate[1]. The G3E family contains four relatively well-studied subfamilies: ArgK, HypB, UreG and CobW (Figure 1.4). Recently, studies in our laboratory have demonstrated that MeaB, a GTPase which belongs to ArgK subfamily, is involved in cofactor docking and protection of MCM from inactivation (*56, 57*).

HypB (hydrogenase pleiotropic B) is a metal-binding G protein that mediates and regulates, together with HypA and its homologue HybF, nickel incorporation into [NiFe]-hydrogenase (*58-60*). UreG is a metallochaperone which is highly homologous in sequence and function to HypB and is involved in nickel incorporation into urease (*61, 62*).

[1]In the G-protein domain of many fusion proteins IcmFs, the base-specificity loop motif NKxD is modified to NKxE (See section 3.4.7). The substitution of aspartate by glutamate results in relaxed substrate specificity for these IcmFs, which exhibit not only GTPase but also ATPase activity.

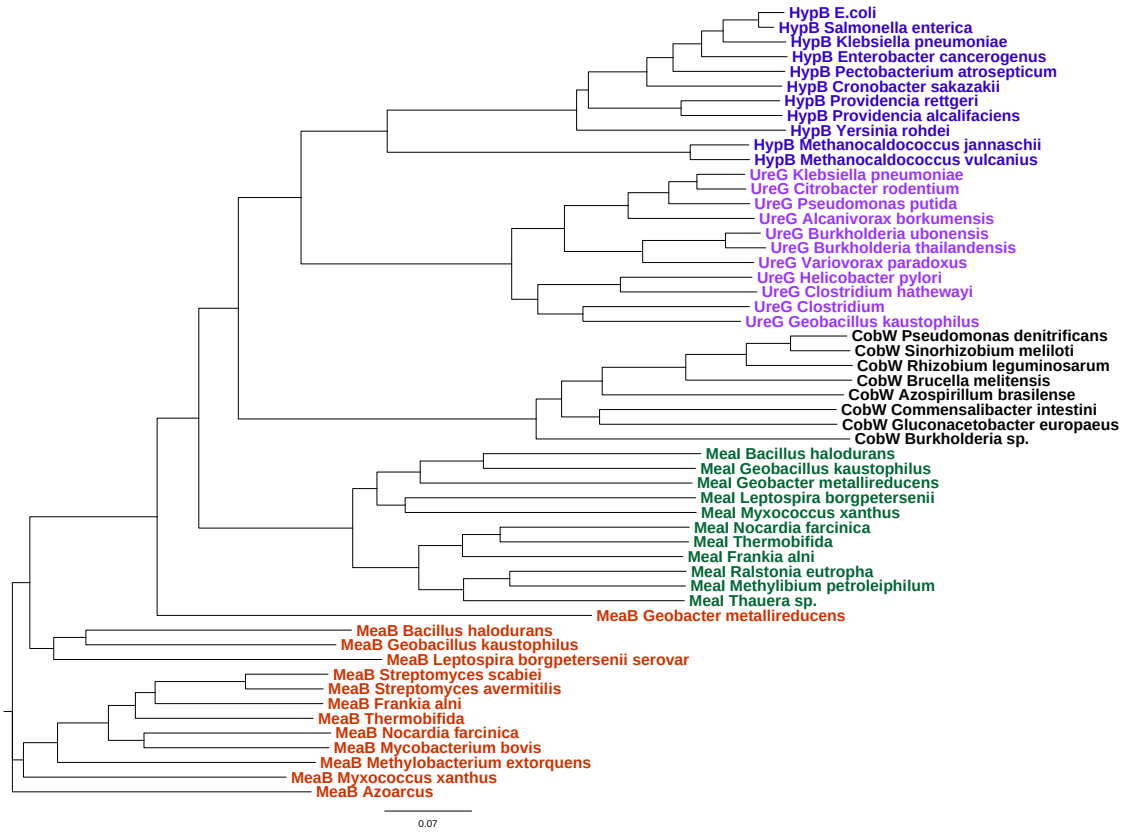


Figure 1.4: Summary of phylogenomic analysis of the G3E family. Tree was constructed using sequences from diverse members of the G3E family. It can be seen that all subfamilies are clearly defined: MeaB-like proteins are shown in red, MeaI-like domains extracted from IcmFs fusion proteins are shown in green, CobW-like proteins are shown in black, UreG-like proteins are shown in purple and HypB-like proteins are shown in blue.

1.8 Organization and common features of G proteins which belong to the G3E family

In thinking about how G proteins perform their chaperone functions, one must keep in mind that some chaperones bind metals and deliver them to the active site of the target enzyme while other chaperones do not bind cofactors/metals but regulate holo-enzyme maturation and perform other auxiliary functions, e.g. protecting target enzymes from inactivation during catalysis. Along these lines, Crecy-Lagard and colleagues proposed two general roles for members of the G3E family: 1) facilitation of cofactor incorporation

in an energy-dependent manner into target active sites (the insertase role) and, 2) storage and delivery of a metal/cofactor to a target metalloprotein (the metallochaperone role). In the light of our findings on the G protein chaperone for MCM (*57*), it is necessary to expand the suggested categories to include a third one: i.e. protection of the catalytically active form of the enzyme.

While MeaB, HypB and UreG have target metalloenzymes for which they serve as chaperones, not much is known about a large group of proteins which are classified under the name of CobW or COG0523 (Figure 1.4) (*63*). CobW from *Pseudomonas denitrificans* was the first member of this group to be described, and it was shown that disruption of the corresponding gene led to the inability to synthesize cobalamin (*64*). According to Crecy-Lagard and colleagues only 12.5% of COG0523 are orthologs of CobW from *Pseudomonas denitrificans* (*63*). Clearly, there is a substantial diversity among COG0523 members, which indicates that these chaperones support functions of several distinct metalloenzymes. For example, a member of CobW subfamily was shown to be a nitrile hydratase activator for the Fe-type nitrile hydratase from *Rhodococcus sp.* N-771 (*65*). CobW-like proteins are found in all kingdoms of life (*63*). In humans, two copies of CobW-like proteins are found: CBDW1 and CBDW2 and their function is not unknown. Interestingly, proteins belonging to COG0523 were linked to pathogenicity and ability of pathogens to overcome zinc deficiency induced by the host organism (*63*) and (Dr. Heran Darwin, NYU, personal communication).

Essentially, proteins from the G3E family can be seen as comprising a core G domain with the alpha-helical N- and C-terminal extensions. These extensions, which contain histidine-rich regions (or “histidine stretches”), are implicated in metal binding and are

found in different combinations together with the G domain in HypB, UreG and CobW proteins. Most HypB have "histidine stretches" at the N terminus. On the other hand, HypB from *E. coli* and some other bacteria have no "histidine stretches" at the N terminus but still bind 2 nickel ions per monomer (*58*). That metal binding is achieved via a CxxCGC motif at the N terminus in addition to another metal binding site located within the G-domain (*66*).

Interestingly, the absence of "histidine stretches" in UreGs in certain organisms is compensated by the presence of another protein UreE, which is capable of binding metals (*63*). Indeed, Crecy-Lagard and associates note that in organisms where UreEs are absent (for example in *Anaeromyxobacter sp.*, *Frankia sp.*, *Mycobacterium tuberculosis*, *Nocardia farcinica*, *Streptomyces coelicolor*), UreG-like proteins contain "histidine stretches" of different length at the N terminus (*63*).

CobW-like proteins have "histidine stretches" exclusively at the C termini, where the minimal motif is HxHxHxH (where x represents 0-4 residues). However some CobW-like proteins have many more histidines in this region, for example, a protein from *Anabaena variabilis* has a "histidine stretch" containing 29 histidines (*63*)! In summary, the variety in "histidine stretches" in HypB, UreG and CobW members of the G3E family of G proteins highlights the fact that these proteins are important in metal storage and transport to their target enzymes.

MeaB-like proteins have alpha-helical extensions both at the N- and C- termini but without any metal or cofactor binding motifs (Figure 1.5 and Figure 1.6). According to X-ray crystallographic studies, the C-terminal extension appears to be essential for dimerization of MeaB from *M. extorquens* (*67*) (Figure 1.5).

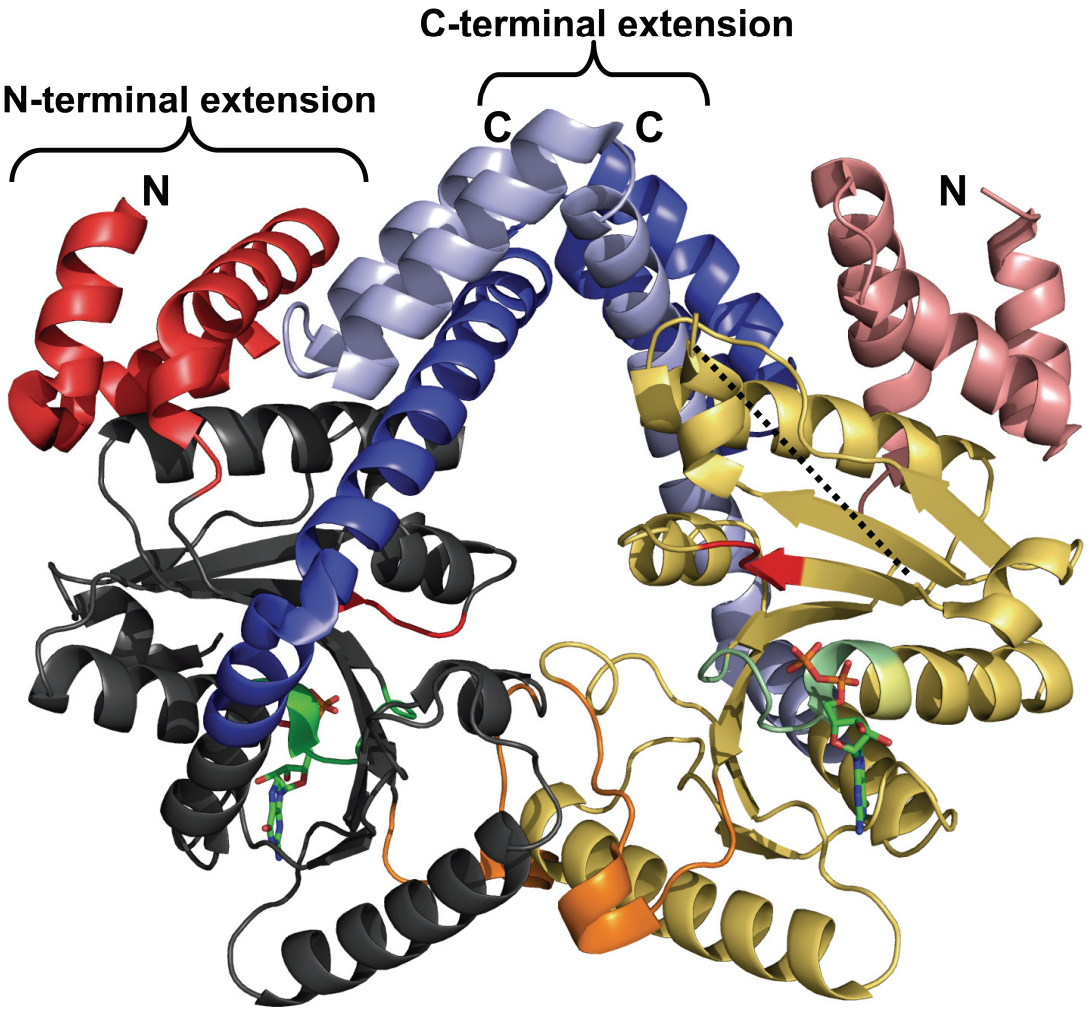


Figure 1.5: Crystal structure of MeaB. Homodimeric assembly of MeaB (PDB: 2QM7) with subunits shown in grey and yellow. The N- and C-terminal extensions are indicated in red and blue or in pale colors for the second subunit. The two GDP molecules are indicated as sticks representation. Important regions of the G domain are indicated: P-loop (G62-S69) in green, switch II (E154-G157) in red and switch III (L177-K188) in orange. Switch I (L103-M109) is not ordered in this structure and it is indicated as a black dotted line.

Another interesting feature of the G3E family is the connection between their function and oligomerization (Figure 1.6). Recently, Alfred Wittinghofer and colleagues have proposed a new classification for specific P-loop NTPases which is based on their

dimerization mode in the presence of ligands (*68*). GADs (**G**-proteins **a**ctivated by **d**imerization) are G proteins that have a relatively weak affinity for nucleotides and therefore do not require an external GEFs (guanine nucleotide exchange factors) (*68*). GADs dimerize only in a nucleotide-dependent manner following which, they can bind their corresponding GAPs (GTPase activating proteins) and perform their biological functions (*68*). In some cases, metalloenzymes function as GAPs for their corresponding chaperones. It has been known for a while that CooC and HypB are monomers in solution, but that they undergo nucleotide-dependent dimerization (*54, 55, 58*). It was proposed that nucleotide-dependent dimerization is a common feature of the SIMIBI class of G proteins or at least, those that bind nucleotides at the dimer interface (*51, 68*).

It has been reported recently that the *E. coli* HypB L242A/L246A mutant no longer dimerizes in the presence of nucleotides. On the other hand, the monomeric form of the mutant HypB has both GTPase and metal-binding behavior similar to the wild-type protein. Most importantly, the mutant protein can support maturation of [NiFe]-hydrogenase both *in vitro* and *in vivo* albeit only half of the activity is reconstituted as compared to the wild-type control (*59*).

MeaB from *M. extorquens* and the human ortholog MMAA or CblA are dimers in solution regardless of the presence of nucleotides (*47, 67*) (Figure 1.5 and 1.6). This challenges the proposed concept that G proteins from SIMIBI class are all GADs (*51, 68*).

1.9 MeaB is a chaperone for MCM

MeaB was recognized as an auxiliary protein for MCM by Mary Lidstrom and colleagues in their classical studies of glyoxylate regeneration pathway in *M. extrorquens* (*69, 70*).

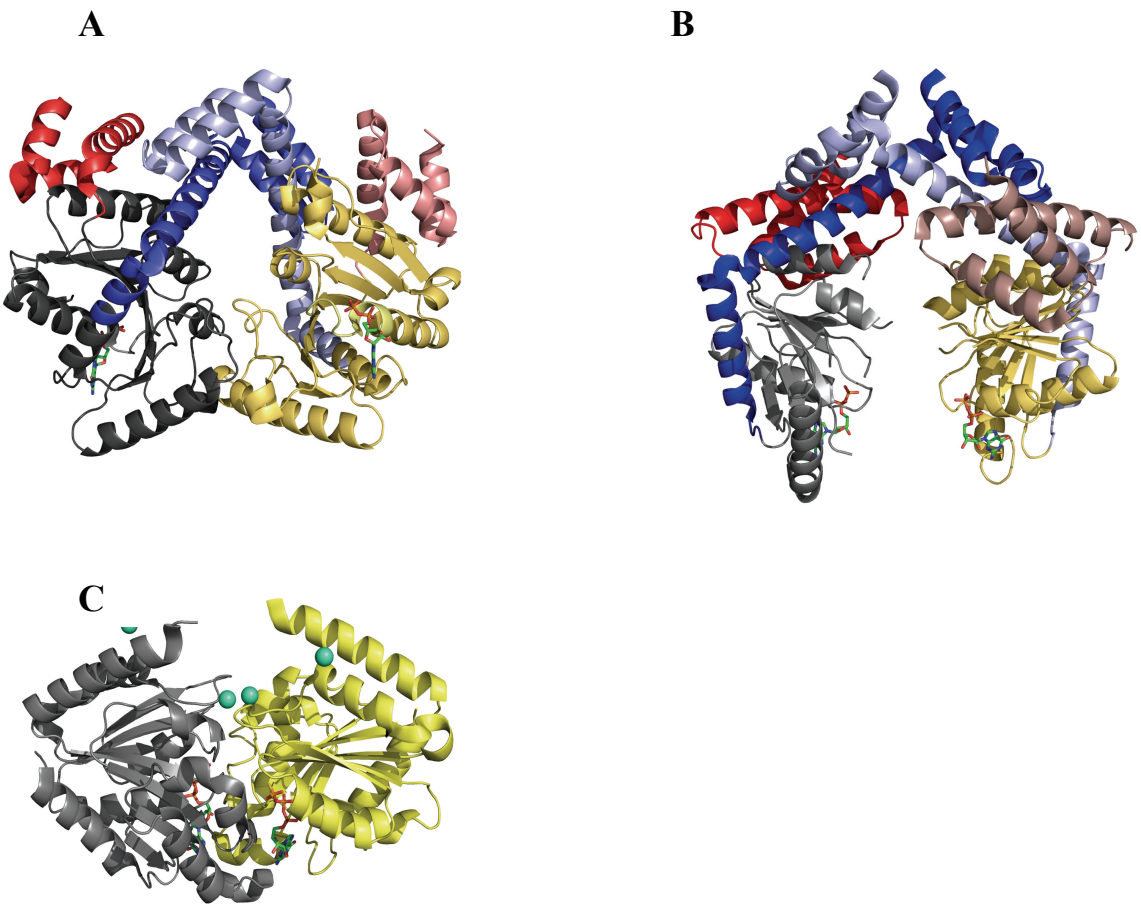


Figure 1.6: Dimmer assembly in the G3E family. Crystal structures of MeaB (**A**) (2QM7A), MMAA (**B**) (2WWW) and HypB (**C**) (2HF8) in complex with nucleotides (MeaB and MMAA with GDP, HypB with GTPγS). Nucleotides molecules are indicated as sticks representation. Subunits are shown in grey and yellow. The N- and C-terminal extensions are indicated in red and blue or in pale colors for the second subunit. Zink atoms bound to HypB are shown as green spheres. All three G3E family GTPases exhibit a modified Rossmann fold with a central parallel β-sheet flanked by α-helices, but they differ significantly in the arrangement of the dimer interface.

MCM activity was completely lost in *M. extorquens* when the *mea*B gene was disrupted (Δ*mea*B::*kan*) and the bacterium was unable to grow on C_1 and C_2 compounds (*69*). Interestingly, in strains in which *meaB* and *meaD* (PduO-type **a**denosyl**t**ransfe**r**ase or ATR) or *epm* (methylmalonyl-CoA epimerase) genes were inactivated

(*ΔepmΔmeaB::kan*, *ΔmeaDΔmeaB::kan*), MCM activity was restored by addition of exogenous AdoCbl to cell extracts (*70*). In cell extract of the strain where only the *meaB* gene (Δ*mea*B::*kan*) was disrupted, MCM activity could not be restored by addition of exogenous AdoCbl (*70*). While ATR catalyzes the formation of AdoCbl from cob(I)alamin and ATP, methylmalonyl-CoA epimerase catalyzes the conversion of S-methylmalonyl-CoA to R-methylmalonyl-CoA, i.e. it provides the correct stereoisomer of the substrate needed by MCM. The result from the Δ*mea*B::*kan* strain is important, as it clearly suggests that MCM is inactivated during catalysis and it cannot recover activity merely by addition of exogenous AdoCbl. Knowing the interplay between and functions of the MeaB, ATR and MCM trio of proteins, the results from the various gene disruption strains can be explained (*57*).

Bioinformatics pattern searches reveal that MeaB-like proteins are strongly associated with AdoCbl-dependent mutases (*63, 71*). In fact 63% of all MeaB-like proteins found in prokaryotic genomes are adjacent to or reside within the same operon with MCMs or the B_{12}-binding domains of other AdoCbl-dependent mutases (*63*). In fact, the analysis revealing a strong operonic association between MeaB and MCM, led to the identification of the corresponding human gene (MMAA) (*71*). MMAA (for methymalonic aciduria linked to the *cbl*A complementation group), is a mammalian ortholog of MeaB that supports the reaction catalyzed by human MCM and its gene maps to human chromosome 4q31 (*71*).

The crystal structures of MeaB (*67*) and MMAA (*47*) clearly show that their G domains are organized as in other members of the G3E family (Figure 1.5 and Figure 1.6). MeaB possesses the signature G1-G4 motifs of the G3E family: it has the GxxGxGK[ST]

Walker A and the DxxxxExxG Walker B motifs, a [V/I]xxD Mg^{2+}-binding motif and an NKxD GTP-binding motif. MeaB and MMAA are homodimers with a subunit molecular weight of ~34-46 kDa. The overall structure of the MeaB dimer is different from the structures of HypB and resembles a starfish, with a central G-domain comprised of parallel β-strands surrounded by α-helices (*67*) (Figure 1.6).

Like other members of the G3E family, the core G-domain of MeaB is flanked by alpha-helical extensions at the N- and C- termini. In both MeaB and MMAA, the C-terminal extension, which is called "dimerization arm" resembles a hook and is involved in dimerization (Figure 1.6). The dimer interface of MeaB is formed by the G domains in both subunits and the dimerization arms. In contrast, the subunit interface in MMAA is formed by the dimerization arms causing the protein to adopt an open "U-shaped" configuration (Figure 1.6). MeaB and MMAA exist as dimers in solution, and their oligomeric state is not driven by nucleotide binding. This challenges the notion that all members of the G3E family undergo nucleotide-dependent dimerization, which is crucial for their function (*58, 68*). The N-terminal extensions in MeaB are proposed to be important for interacting MCM. Unfortunately, numerous attempts to co-crystallize MeaB and MCM so far have been unsuccessful.

MeaB from *M. extorquens* possesses a low intrinsic GTPase activity (k_{cat} ~0.04 min^{-1}) that is enhanced ~ 100 fold in the presence of MCM. Thus MCM acts as a GAP for MeaB (*72*). MeaB exhibits a weak affinity for GMPPNP and GDP (7.3 ± 1.9 μM and 6.2 ± 0.7 μM, respectively), which is in agreement with previous findings that G-proteins from the G3E family do not require GEF proteins to cycle between the GDP- and GTP-bound forms (*68*). Direct protein-protein interaction between MeaB and MCM can be

demonstrated by native-PAGE and ITC (*72*). Homodimeric MeaB and heterodimeric MCM from *M. extorquens* form a tight 1:1 complex. In contrast, homodimeric MMAA and the homodimeric human MCM form a 2:1 complex, which is in agreement with MCM having two active sites, each of which needs to be loaded with cofactor (*47*). Interestingly, the K_D for binding of MCM and MeaB ranges from 34 ± 4 to 524 ± 66 nM depending on AdoCbl and the type of nucleotide present. The tightest K_D (34 ± 4 nM) was determined for binding of holo-MCM and MeaB/GMPPNP, whereas the weakest binding was demonstrated for holo-MCM and MeaB/GDP (524 ± 66 nM) (*72*).

Immunoprecipitation was used to demonstrate that YgfD, an ortholog of MeaB in *E. coli,* interacts with the endogenous MCM in cell lysates (*40*). When both YgfD and homodimeric *E. coli* MCM were cloned and overexpressed, both proteins formed a complex, as shown by gel-filtration and Western-blot analysis (*40*).

The presence of MeaB and nucleotides affects the activity of MCM. MeaB/GDP increased k_{cat} for the mutase reaction 1.8-fold (from 132 ± 16 min^{-1} to 237 ± 12 min^{-1}) while also increasing the K_M for methylmalonyl-CoA by 1.7-fold (*56*). As a result, the MeaB/GDP complex did not increase the catalytic efficiency (k_{cat}/K_M) of the mutase. Notably, MeaB/GDP decreased the K_{act} for AdoCbl ~4.6-fold, and no AdoCbl binding to MCM was observed in the presence of MeaB/GMPPMP, indicating that cofactor binding requires GTP hydolysis (*56*).

1.10 Biochemical properties of MMAA, a human ortholog of MeaB

Until recently, most of the biochemical data on MCM and its G-protein chaperone was available only for bacterial duo of proteins from *M. extorquens* since the human proteins were difficult to work with due to poor yields and instability during expression and

purification. More recently, the two human proteins were successfully overexpressed and purified in quantities sufficient both for crystallography and biochemical studies (*47*). Both proteins were overexpressed as N-terminally truncated variants (73-424 for MMAA and 12-750 for MCM). As with the bacterial proteins, the kinetic parameters of MMAA were modulated by the presence of MCM. In the presence of apo-MCM, the GTPase activity of MMAA increased 5-fold, whereas K_M for GTP decreased 16-fold (from 1210 ± 330 μM to 74 ± 8 μM). Consequently, the catalytic efficiency (k_{cat}/K_M) of the MMAA-catalyzed reaction increased 82-fold (*47*). The presence of holo-MCM only decreased the K_M for GTP and the catalytic efficiency of MMAA was increased only 8-fold (*47*).

The recently obtained crystal structures of MMAA and MCM with different ligands along with the structures of their bacterial counterparts set an important framework for elucidation of the mechanism by which the mutase and its G-protein chaperone interact and modulate each other's activities (Figure 1.6) (*47*).

1.11 The Chaperoning role of MeaB: The interplay between three proteins

According to the general classification of functions of metallochaperones in the G3E family, MeaB plays a role both as an insertase and as a chaperone (Section 1.8). The role of MeaB in the interplay between the three bacterial proteins MCM, its chaperone MeaB and ATR as been elucidated (*57*). ATR catalyses the transfer of 5'-deoxyadenosyl moiety from ATP to the cobalt atom of cob(I)alamin to form AdoCbl (*73*). This reaction is important not only in the *de novo* biosynthesis of AdoCbl in bacteria but also in higher animals in the assimilation of coenzyme B_{12} precursors (See section 4.1)(*73*).

It is interesting that there are three families of ATRs that are completely unrelated in their amino acid sequence: the PduO, CobA and EutT types. Human ATR belongs to the PduO

family (*74*). The PduO-type ATR from *M. extorquens* functions not just in the enzymatic synthesis of AdoCbl, but also as a chaperone for its direct delivery to MCM (*75*).

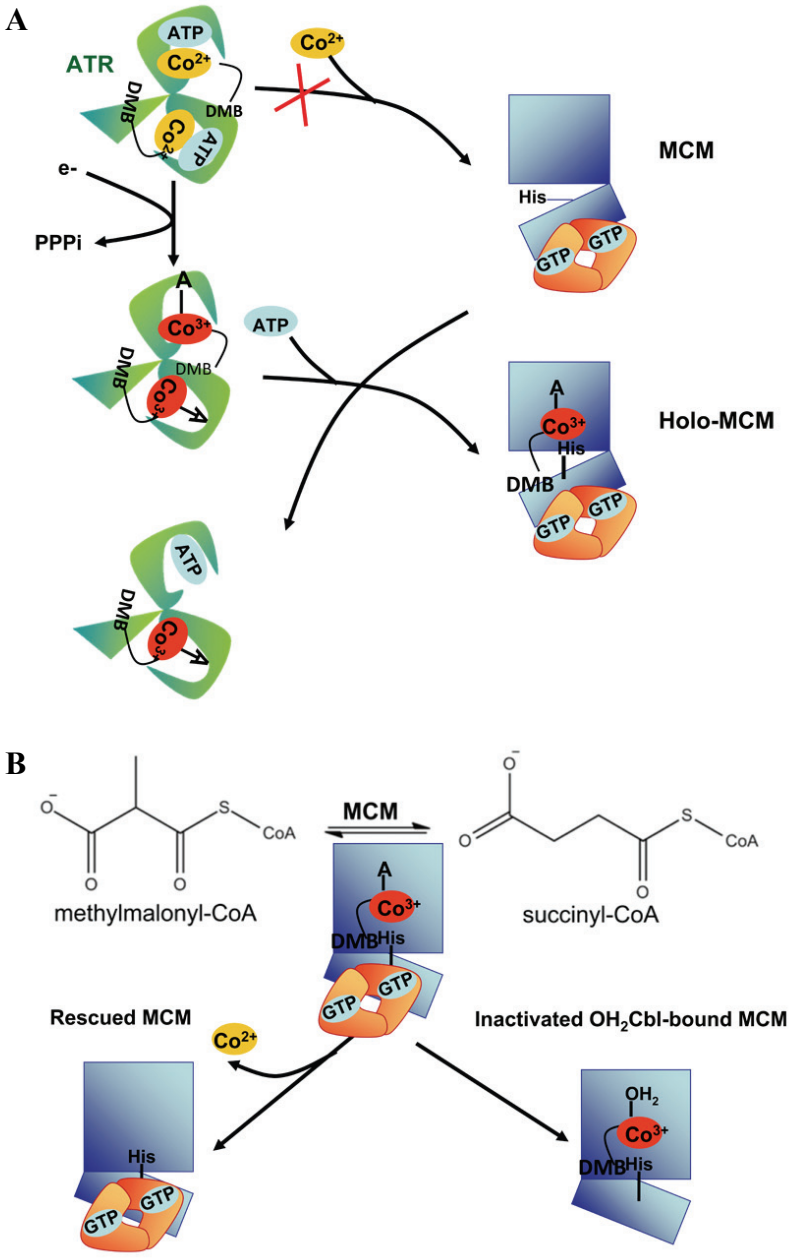


Figure 1.7: Schematic model for the editing and gating (A), and rescue functions (B) of MeaB. Adapted from (*57*).

This strategy is similar to substrate channeling when one enzyme transfers product of the reaction to the next enzyme in the pathway or to another active site within the same multifunctional protein, without releasing the product into solution.

The participation of these "molecular tunnels" helps to minimize unwanted side reactions of unstable intermediates and dilution of reactants in the cytosol. Substrate channeling is employed by several enzymes including pyruvate dehydrogenase complex (*76*), tryptophan synthase (*77*), dihydrofolate reductase-thymidylate synthase (*76*). Although coenzyme B_{12} can be synthesized only by some bacteria, many prokaryotes depend on cobalamin salvaging and transport to meet their needs for this rare cofactor. Thus, the strategy for sequestering AdoCbl following its synthesis by ATR and its direct transfer to MCM, ensures reconstitution of the active holo-enzyme (*75*).

Once the role of the PduO-type ATR in cofactor transfer to MCM was established, it immediately raised questions about the role of MeaB in the transfer process (*75*). Interestingly, when apo-MCM is complexed with MeaB/GTP, cofactor transfer is not seen in the presence of AdoCbl-loaded ATR (Figure 1.7 A) (*57*). In the presence of ATP, transfer of a single equivalent of AdoCbl from ATR to apo-MCM is observed (Figure 1.7 A) (*57*).

Remarkably, an inactive cofactor form, cob(II)alamin, can be transferred from holo-ATR to apo-MCM if it is not complexed to MeaB. In other words, the presence of MeaB prevents loading of MCM with cob(II)alamin which would lead to reconstitution of inactive MCM (*56, 57*). This editing function of MeaB uses the energy of GTP binding rather than hydrolysis (Figure 1.7 A).

The first chemical step in the reaction catalyzed by AdoCbl-dependent mutases is Co-carbon bond homolysis which results in the formation of two species: cob(II)alamin and 5'-deoxyadenosyl pair of radicals (Figure 1.3)(3). Occasionally during catalysis the 5'-deoxyadenosine moiety is lost from the active site. In this case AdoCbl can not be reformed and cob(II)alamin remaining in the active site is oxidized to OH_2Cbl (*56*). MCM binds OH_2Cbl tightly and renders the mutase inactive. Formation of OH_2Cbl can be monitored by UV-visible spectroscopy by following the increase in absorption at 350 nm, which is characteristic of OH_2Cbl. With MCM from *M. extroquens,* the observed rate constant for inactivation is k_{obs}= 0.0072 min^{-1}, and it is decreased 12-fold in the presence of MeaB/GTP (*56*). However, it was puzzling that when OH_2Cbl bound to MCM was mixed with MeaB/GTP, the damaged cofactor was not released, as it was shown with chaperones for other AdoCbl-dependent enzymes (*29, 30*). It turns out that during the reaction catalyzed by MCM, MeaB senses loss of the 5'-deoxyadenosine moiety and under these conditions, can expunge cob(II)alamin from the active site of MCM in a reaction that is dependent on its GTPase activity (Figure 1.7 B). If the 5'-deoxyadenosine moiety is lost and cob(II)alamin is oxidized to OH_2Cbl, it is "too late" for the chaperone to rescue the enzyme.

Based on the available data for human proteins, we expect that in mammals MMAB (the ATR ortholog in humans), MMAA and MCM interact in a similar way as their bacterial counterparts (*47*). Accordingly, MMAB converts cobalamin entering the mitochondrion to AdoCbl and subsequently transfers the cofactor to MCM in a process that is gated and edited by the chaperone, MMAB in the presence of GTP.

1.12 Fusion between MCM and MeaB is a misannotation

It was believed before that in addition to the strong operonic association between MeaB and MCM, in some bacteria MeaB is fused to the large subunit of MCM (*70*). Several genes with such an architecture were identified via BLAST: *Ralstonia solanacearun* (NP_518358), *R. metallidurans* (YP_582365), *Burkholderia xenovorans* (YP_556774), *Thermobifida fusca* (YP_290867), *Geobacter metallireducens* (YP_384678), *Leptospira interrogans serovar lai* (NP_713136), *Bacillus halodurans* (NP_244663) and the name McmC was designated to this fusion protein (*70*). Indeed, a *B. xenovorans* mutant, Δ*mcmC::kan,* was shown to lack MCM-like activity (*70*). The McmC protein from *B. xenovorans* is 1272 amino acids in length and can be divided into three domains: the N-terminal region (from residues 22-145) has very high sequence similarity to the AdoCbl-binding region of MCM, the middle region (extending from residues 293-607) aligns with MeaB, and the C-terminal domain (residues 691-1195) aligns with the large subunit of MCM. McmC was reported to be involved in propionate metabolism in *B. xenovorans*, since slow growth of the Δ*mcmC::kan* mutant was observed on propionate compared to the wild-type strain (*70*).

Our results however revealed that none of the identified McmC proteins represent fusions of MCM and MeaB (*78*). Using a combination of bioinformatics and biochemical approaches, we have demonstrated that the fused proteins in bacteria represent a fusion between ICM and MeaB. We propose a new name for this fusion, IcmF (**i**sobutyryl-**C**oA **m**utase **f**used (Figure 1.1).

1.13 References

(1) Ratnatilleke, A., Vrijbloed, J. W., and Robinson, J. A. (1999) Cloning and sequencing of the coenzyme B(12)-binding domain of isobutyryl-CoA mutase from Streptomyces cinnamonensis, reconstitution of mutase activity, and characterization of the recombinant enzyme produced in Escherichia coli. *J Biol Chem 274*, 31679-85.

(2) Zerbe-Burkhardt, K., Ratnatilleke, A., Philippon, N., Birch, A., Leiser, A., Vrijbloed, J. W., Hess, D., Hunziker, P., and Robinson, J. A. (1998) Cloning, sequencing, expression, and insertional inactivation of the gene for the large subunit of the coenzyme B12-dependent isobutyryl-CoA mutase from Streptomyces cinnamonensis. *J Biol Chem 273*, 6508-17.

(3) Banerjee, R. (2003) Radical carbon skeleton rearrangements: catalysis by coenzyme B12-dependent mutases. *Chem Rev 103*, 2083-94.

(4) G. Brendelberger, J. R. e., D.M. Ashworth, K. Reynolds, F. Willenbrock, J.A. Robinson. (1988) The enzymic interconversion of isobutyryl and N-butyrylcarba(dethia)-coenzyme-A - a coenzyme-B12-dependent carbon skeleton rearrangement. *Angew. Chem. Int. Ed. Engl. 27*, 1089–1091.

(5) Moore, B., Eisenberg, R., Weber, C., Bridges, A., Nanz, D., and Robinson, J. (1995) On the stereospecificity of the coenzyme B12-dependent isobutyryl-CoA mutase reaction. *J. Am. Chem. Soc. 117*, 11285-11291.

(6) Daublain, P., Horner, J. H., Kuznetsov, A., and Newcomb, M. (2004) Solvent polarity effects and limited acid catalysis in rearrangements of model radicals for the methylmalonyl-CoA mutase- and isobutyryl-CoA mutase-catalyzed isomerization reactions. *J Am Chem Soc 126*, 5368-9.

(7) Vrijbloed, J. W., Zerbe-Burkhardt, K., Ratnatilleke, A., Grubelnik-Leiser, A., and Robinson, J. A. (1999) Insertional inactivation of methylmalonyl coenzyme A (CoA) mutase and isobutyryl-CoA mutase genes in Streptomyces cinnamonensis: influence on polyketide antibiotic biosynthesis. *J Bacteriol 181*, 5600-5.

(8) Angelidaki, I., and Ahring, B. K. (1995) Isomerization of n- and i-butyrate in anaerobic methanogenic systems. *Antonie Van Leeuwenhoek 68*, 285-91.

(9) Oude Elferink, S. J., Lens, P. N., Dijkema, C. and Stams, A. J. . (1996) Isomerization of butyrate to isobutyrate by Desulforhabdus amnigenus. *FEMS Microbiology Letters 142*, 237–241.

(10) Tholozan, J.-L., Samain, E. and Grivet, J.-P. (1988) Isomerization between n-butyrate and isobutyrate in enrichment cultures. *FEMS Microbiology Letters 53*, 187–191.

(11) Wu, W. M., Jain, M. K., and Zeikus, J. G. (1994) Anaerobic degradation of normal- and branched-chain Fatty acids with four or more carbons to methane by a syntrophic methanogenic triculture. *Appl Environ Microbiol 60*, 2220-6.

(12) Matthies, C., Springer, N., Ludwig, W., and Schink, B. (2000) Pelospora glutarica gen. nov., sp. nov., a glutarate-fermenting, strictly anaerobic, spore-forming bacterium. *Int J Syst Evol Microbiol 50 Pt 2*, 645-8.

(13) Matthies, C., and Schink, B. (1992) Reciprocal isomerization of butyrate and isobutyrate by the strictly anaerobic bacterium strain WoG13 and methanogenic

isobutyrate degradation by a defined triculture. *Appl Environ Microbiol 58*, 1435-9.
(14) Sekiguchi, Y., Kamagata, Y., Nakamura, K., Ohashi, A., and Harada, H. (2000) Syntrophothermus lipocalidus gen. nov., sp. nov., a novel thermophilic, syntrophic, fatty-acid-oxidizing anaerobe which utilizes isobutyrate. *Int J Syst Evol Microbiol 50 Pt 2*, 771-9.
(15) Birch, A., Leiser, A., and Robinson, J. A. (1993) Cloning, sequencing, and expression of the gene encoding methylmalonyl-coenzyme A mutase from Streptomyces cinnamonensis. *J Bacteriol 175*, 3511-9.
(16) Drennan, C. L., Huang, S., Drummond, J. T., Matthews, R. G., and Lidwig, M. L. (1994) How a protein binds B12: A 3.0 A X-ray structure of B12-binding domains of methionine synthase. *Science 266*, 1669-74.
(17) Holloway, D. E., and Marsh, E. N. (1994) Adenosylcobalamin-dependent glutamate mutase from Clostridium tetanomorphum. Overexpression in Escherichia coli, purification, and characterization of the recombinant enzyme. *J Biol Chem 269*, 20425-30.
(18) Marsh, E. N. (2000) Review Article Coenzyme-B(12)-Dependent Glutamate Mutase. *Bioorg Chem 28*, 176-189.
(19) Weissman, K. J., and Leadlay, P. F. (2005) Combinatorial biosynthesis of reduced polyketides. *Nat Rev Microbiol 3*, 925-36.
(20) Chan, Y. A., Podevels, A. M., Kevany, B. M., and Thomas, M. G. (2009) Biosynthesis of polyketide synthase extender units. *Nat Prod Rep 26*, 90-114.
(21) Zhang, W., and Reynolds, K. A. (2001) MeaA, a putative coenzyme B12-dependent mutase, provides methylmalonyl coenzyme A for monensin biosynthesis in Streptomyces cinnamonensis. *J Bacteriol 183*, 2071-80.
(22) Erb, T. J., Retey, J., Fuchs, G., and Alber, B. E. (2008) Ethylmalonyl-CoA mutase from Rhodobacter sphaeroides defines a new subclade of coenzyme B12-dependent acyl-CoA mutases. *J Biol Chem 283*, 32283-93.
(23) Smith, L. M., Meijer, W. G., Dijkhuizen, L., and Goodwin, P. M. (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.
(24) Gruber, K., Puffer, B., and Krautler, B. (2011) Vitamin B12-derivatives-enzyme cofactors and ligands of proteins and nucleic acids. *Chem Soc Rev 40*, 4346-63.
(25) Matthews, R. G. (2009) Cobalamin- and corrinoid-dependent enzymes. *Met Ions Life Sci 6*, 53-114.
(26) Ludwig, M. L., and Matthews, R. G. (1997) Structure-based perspectives on B12-dependent enzymes. *Annu Rev Biochem 66*, 269-313.
(27) Padmakumar, R., Padmakumar, R., and Banerjee, R. (1997) Evidence that cobalt-carbon bond homolysis is coupled to hydrogen atom abstraction from substrate in methylmalonyl-CoA mutase. *Biochemistry 36*, 3713-8.
(28) Padovani, D., and Banerjee, R. (2006) Alternative pathways for radical dissipation in an active site mutant of B12-dependent methylmalonyl-CoA mutase. *Biochemistry 45*, 2951-9.

(29) Toraya, T. (2000) Radical catalysis of B12 enzymes: structure, mechanism, inactivation, and reactivation of diol and glycerol dehydratases. *Cell Mol Life Sci 57*, 106-27.
(30) Toraya, T., Tamura, N., Watanabe, T., Yamanishi, M., Hieda, N., and Mori, K. (2008) Mechanism-based inactivation of coenzyme B12-dependent diol dehydratase by 3-unsaturated 1,2-diols and thioglycerol. *J Biochem 144*, 437-46.
(31) Vlasie, M. D., and Banerjee, R. (2004) When a spectator turns killer: suicidal electron transfer from cobalamin in methylmalonyl-CoA mutase. *Biochemistry 43*, 8410-7.
(32) Tang, K. H., Chang, C. H., and Frey, P. A. (2001) Electron transfer in the substrate-dependent suicide inactivation of lysine 5,6-aminomutase. *Biochemistry 40*, 5190-9.
(33) Frey, P. A., and Reed, G. H. (2011) Pyridoxal-5'-phosphate as the catalyst for radical isomerization in reactions of PLP-dependent aminomutases. *Biochim Biophys Acta*.
(34) Willard, H. F., and Rosenberg, L. E. (1980) Inherited methylmalonyl CoA mutase apoenzyme deficiency in human fibroblasts: evidence for allelic heterogeneity, genetic compounds, and codominant expression. *J Clin Invest 65*, 690-8.
(35) Deodato, F., Boenzi, S., Santorelli, F. M., and Dionisi-Vici, C. (2006) Methylmalonic and propionic aciduria. *Am J Med Genet C Semin Med Genet 142C*, 104-12.
(36) Willard, H. F., and Rosenberg, L. E. (1977) Inherited deficiencies of human methylmalonyl CaA mutase activity: reduced affinity of mutant apoenzyme for adenosylcobalamin. *Biochem Biophys Res Commun 78*, 927-34.
(37) Banerjee, R., Gherasim, C., and Padovani, D. (2009) The tinker, tailor, soldier in intracellular B12 trafficking. *Curr Opin Chem Biol 13*, 484-91.
(38) Allen, S. H., Kellermeyer, R. W., Stjernholm, R. L., and Wood, H. G. (1964) Purification and Properties of Enzymes Involved in the Propionic Acid Fermentation. *J Bacteriol 87*, 171-87.
(39) Mancia, F., Keep, N. H., Nakagawa, A., Leadlay, P. F., McSweeney, S., Rasmussen, B., Bosecke, P., Diat, O., and Evans, P. R. (1996) How coenzyme B12 radicals are generated: the crystal structure of methylmalonyl-coenzyme A mutase at 2 A resolution. *Structure 4*, 339-50.
(40) Froese, D. S., Dobson, C. M., White, A. P., Wu, X., Padovani, D., Banerjee, R., Haller, T., Gerlt, J. A., Surette, M. G., and Gravel, R. A. (2009) Sleeping beauty mutase (sbm) is expressed and interacts with ygfd in Escherichia coli. *Microbiol Res 164*, 1-8.
(41) Haller, T., Buckel, T., Retey, J., and Gerlt, J. A. (2000) Discovering new enzymes and metabolic pathways: conversion of succinate to propionate by Escherichia coli. *Biochemistry 39*, 4622-9.
(42) Miyamoto, E., Watanabe, F., Charles, T. C., Yamaji, R., Inui, H., and Nakano, Y. (2003) Purification and characterization of homodimeric methylmalonyl-CoA mutase from Sinorhizobium meliloti. *Arch Microbiol 180*, 151-4.
(43) Ramos-Vera, W. H., Weiss, M., Strittmatter, E., Kockelkorn, D., and Fuchs, G. (2011) Identification of missing genes and enzymes for autotrophic carbon fixation in crenarchaeota. *J Bacteriol 193*, 1201-11.

(44) Berg, I. A., Kockelkorn, D., Buckel, W., and Fuchs, G. (2007) A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. *Science 318*, 1782-6.
(45) Miyamoto, E., Watanabe, F., Yamaguchi, Y., Takenaka, H., and Nakano, Y. (2004) Purification and characterization of methylmalonyl-CoA mutase from a photosynthetic coccolithophorid alga, Pleurochrysis carterae. *Comp Biochem Physiol B Biochem Mol Biol 138*, 163-7.
(46) Han, Y. S., Bratt, J. M., and Hogenkamp, H. P. (1984) Purification and characterization of methylmalonyl-CoA mutase from Ascaris lumbricoides. *Comp Biochem Physiol B 78*, 41-5.
(47) Froese, D. S., Kochan, G., Muniz, J. R., Wu, X., Gileadi, C., Ugochukwu, E., Krysztofinska, E., Gravel, R. A., Oppermann, U., and Yue, W. W. (2010) Structures of the human GTPase MMAA and vitamin B12-dependent methylmalonyl-CoA mutase and insight into their complex formation. *J Biol Chem 285*, 38204-13.
(48) Drennan, C. L., Matthews, R. G., Rosenblatt, D. S., Ledley, F. D., Fenton, W. A., and Ludwig, M. L. (1996) Molecular basis for dysfunction of some mutant forms of methylmalonyl-CoA mutase: deductions from the structure of methionine synthase. *Proc Natl Acad Sci U S A 93*, 5550-5.
(49) Verstraeten, N., Fauvart, M., Versees, W., and Michiels, J. (2011) The universally conserved prokaryotic GTPases. *Microbiol Mol Biol Rev 75*, 507-42.
(50) Leipe, D. D., Wolf, Y. I., Koonin, E. V., and Aravind, L. (2002) Classification and evolution of P-loop GTPases and related ATPases. *J Mol Biol 317*, 41-72.
(51) Wittinghofer, A., and Vetter, I. R. (2011) Structure-function relationships of the G domain, a canonical switch motif. *Annu Rev Biochem 80*, 943-71.
(52) Rangaraj, P., Shah, V. K., and Ludden, P. W. (1997) ApoNifH functions in iron-molybdenum cofactor synthesis and apodinitrogenase maturation. *Proc Natl Acad Sci U S A 94*, 11250-5.
(53) Jeon, W. B., Cheng, J., and Ludden, P. W. (2001) Purification and characterization of membrane-associated CooC protein and its functional role in the insertion of nickel into carbon monoxide dehydrogenase from Rhodospirillum rubrum. *J Biol Chem 276*, 38602-9.
(54) Jeoung, J. H., Giese, T., Grunwald, M., and Dobbek, H. (2009) CooC1 from Carboxydothermus hydrogenoformans is a nickel-binding ATPase. *Biochemistry 48*, 11505-13.
(55) Jeoung, J. H., Giese, T., Grunwald, M., and Dobbek, H. (2010) Crystal structure of the ATP-dependent maturation factor of Ni,Fe-containing carbon monoxide dehydrogenases. *J Mol Biol 396*, 1165-79.
(56) Padovani, D., and Banerjee, R. (2006) Assembly and protection of the radical enzyme, methylmalonyl-CoA mutase, by its chaperone. *Biochemistry 45*, 9300-6.
(57) Padovani, D., and Banerjee, R. (2009) A G-protein editor gates coenzyme B12 loading and is corrupted in methylmalonic aciduria. *Proc Natl Acad Sci U S A 106*, 21567-72.
(58) Gasper, R., Scrima, A., and Wittinghofer, A. (2006) Structural insights into HypB, a GTP-binding protein that regulates metal binding. *J Biol Chem 281*, 27492-502.

(59) Cai, F., Ngu, T. T., Kaluarachchi, H., and Zamble, D. B. (2011) Relationship between the GTPase, metal-binding, and dimerization activities of E. coli HypB. *J Biol Inorg Chem 16*, 857-68.
(60) Kaluarachchi, H., Chan Chung, K. C., and Zamble, D. B. (2011) Microbial nickel proteins. *Nat Prod Rep 27*, 681-94.
(61) Zambelli, B., Musiani, F., Benini, S., and Ciurli, S. (2011) Chemistry of Ni2+ in urease: sensing, trafficking, and catalysis. *Acc Chem Res 44*, 520-30.
(62) Boer, J. L., Quiroz-Valenzuela, S., Anderson, K. L., and Hausinger, R. P. (2010) Mutagenesis of Klebsiella aerogenes UreG to probe nickel binding and interactions with other urease-related proteins. *Biochemistry 49*, 5859-69.
(63) Haas, C. E., Rodionov, D. A., Kropat, J., Malasarn, D., Merchant, S. S., and de Crecy-Lagard, V. (2009) A subset of the diverse COG0523 family of putative metal chaperones is linked to zinc homeostasis in all kingdoms of life. *BMC Genomics 10*, 470.
(64) Crouzet, J., Levy-Schil, S., Cameron, B., Cauchois, L., Rigault, S., Rouyez, M. C., Blanche, F., Debussche, L., and Thibaut, D. (1991) Nucleotide sequence and genetic analysis of a 13.1-kilobase-pair Pseudomonas denitrificans DNA fragment containing five cob genes and identification of structural genes encoding Cob(I)alamin adenosyltransferase, cobyric acid synthase, and bifunctional cobinamide kinase-cobinamide phosphate guanylyltransferase. *J Bacteriol 173*, 6074-87.
(65) Lu J, Z. Y., Yamagishi H, Odaka M, Tsujimura M, Maeda M, Endo I. (2003) Motif CXCC in nitrile hydratase activator is critical for NHase biogenesis in vivo. *FEBS Letters 553*, 391-396.
(66) Leach, M. R., Sandal, S., Sun, H., and Zamble, D. B. (2005) Metal binding activity of the Escherichia coli hydrogenase maturation factor HypB. *Biochemistry 44*, 12229-38.
(67) Hubbard, P. A., Padovani, D., Labunska, T., Mahlstedt, S. A., Banerjee, R., and Drennan, C. L. (2007) Crystal structure and mutagenesis of the metallochaperone MeaB: insight into the causes of methylmalonic aciduria. *J Biol Chem 282*, 31308-16.
(68) Gasper, R., Meyer, S., Gotthardt, K., Sirajuddin, M., and Wittinghofer, A. (2009) It takes two to tango: regulation of G proteins by dimerization. *Nat Rev Mol Cell Biol 10*, 423-9.
(69) Korotkova, N., Chistoserdova, L., Kuksa, V., and Lidstrom, M. E. (2002) Glyoxylate regeneration pathway in the methylotroph Methylobacterium extorquens AM1. *J Bacteriol 184*, 1750-8.
(70) Korotkova, N., and Lidstrom, M. E. (2004) MeaB is a component of the methylmalonyl-CoA mutase complex required for protection of the enzyme from inactivation. *J Biol Chem 279*, 13652-8.
(71) Dobson, C. M., Wai, T., Leclerc, D., Wilson, A., Wu, X., Dore, C., Hudson, T., Rosenblatt, D. S., and Gravel, R. A. (2002) Identification of the gene responsible for the cblA complementation group of vitamin B12-responsive methylmalonic acidemia based on analysis of prokaryotic gene arrangements. *Proc Natl Acad Sci U S A 99*, 15554-9.

(72) Padovani, D., Labunska, T., and Banerjee, R. (2006) Energetics of interaction between the G-protein chaperone, MeaB, and B12-dependent methylmalonyl-CoA mutase. *J Biol Chem 281*, 17838-44.
(73) St Maurice, M., Mera, P. E., Taranto, M. P., Sesma, F., Escalante-Semerena, J. C., and Rayment, I. (2007) Structural characterization of the active site of the PduO-type ATP:Co(I)rrinoid adenosyltransferase from Lactobacillus reuteri. *J Biol Chem 282*, 2596-605.
(74) Erger, K. E., Conlon, T. J., Leal, N. A., Zori, R., Bobik, T. A., and Flotte, T. R. (2007) In vivo expression of human ATP:cob(I)alamin adenosyltransferase (ATR) using recombinant adeno-associated virus (rAAV) serotypes 2 and 8. *J Gene Med 9*, 462-9.
(75) Padovani, D., Labunska, T., Palfey, B. A., Ballou, D. P., and Banerjee, R. (2008) Adenosyltransferase tailors and delivers coenzyme B12. *Nat Chem Biol 4*, 194-6.
(76) Miles, E. W., Rhee, S., and Davies, D. R. (1999) The molecular basis of substrate channeling. *J Biol Chem 274*, 12193-6.
(77) Huang, X., Holden, H. M., and Raushel, F. M. (2001) Channeling of substrates and intermediates in enzyme-catalyzed reactions. *Annu Rev Biochem 70*, 149-80.
(78) Cracan, V., Padovani, D., and Banerjee, R. (2010) IcmF is a fusion between the radical B12 enzyme isobutyryl-CoA mutase and its G-protein chaperone. *J Biol Chem 285*, 655-66.

CHAPTER 2

IcmF Is a Fusion Between the Radical B_{12} Enzyme, Isobutyryl-CoA Mutase and its G-protein Chaperone[2,3]

2.1 Abstract

Coenzyme B_{12} is used by two highly similar radical enzymes, which catalyze carbon skeleton rearrangements: methylmalonyl-CoA mutase and isobutyryl-CoA mutase (ICM). ICM catalyzes the reversible interconversion of isobutyryl-CoA and n-butyryl-CoA and exists as a heterotetramer. In the present study, we have identified >70 bacterial proteins, which represent fusions between the subunits of ICM and a P-loop GTPase and are currently misannotated as methylmalonyl-CoA mutases. We designate this fusion protein as IcmF (**i**sobutyryl-**C**oA **m**utase **f**used). All IcmFs are comprised of three domains: the N-terminal AdoCbl binding region that is homologous to the small subunit of ICM (IcmB), a middle P-loop GTPase domain and a C-terminal part that is homologous to the large subunit of ICM (IcmA). The P-loop GTPase domain has very high sequence similarity to the *Methylobacterium extorquens* MeaB, which is a chaperone for methylmalonyl-CoA mutase. We have demonstrated that IcmF is an active ICM by cloning, expressing, and purifying the IcmFs from *Geobacillus kaustophilus*,

[2] The content of this chapter has been published in J Biol Chem. 2010 Jan 1;285(1):655-66: **Cracan V, Padovani D and Banerjee R**. "IcmF is a Fusion Between the Radical B_{12} Enzyme, Isobutyryl-CoA Mutase and its G-protein Chaperone".

[3] ITC and EPR experiments described in this work were performed by **Dr. Dominique Padovani.**We thank Dr. Bruce Palfey for providing purified porcine liver butyryl-CoA dehydrogenase and for help in the developing of coupled assay and Dr. Donald Becker (University of Nebraska-Lincoln) for providing the *M. elsdenii* BDH expression construct. This work was supported in part by a grant from the National Institutes of Health (DK45776).

Nocardia farcinica and *Burkholderia xenovorans*. This finding expands the known distribution of ICM activity well beyond the genus *Streptomyces* where it is involved in polyketides biosynthesis and suggests a role for this enzyme in novel bacterial pathways for amino acid degradation, myxalamid biosynthesis and acetyl-CoA assimilation.

2.2. Introduction

Isobutyryl-CoA mutase (ICM) (EC 5.4.99.13) is a coenzyme B_{12} (or 5′-deoxyadenosylcobalamin or AdoCbl)-dependent enzyme, which catalyses the rearrangement of isobutyryl-CoA to n-butyryl-CoA (1-3). This reaction is very similar to that catalyzed by methylmalonyl-CoA mutase (MCM), which is better studied and more widely distributed in Nature (4). In both reactions, carbon skeleton rearrangements take place where the carbonyl-CoA substituent and a hydrogen atom on neighboring carbon atoms exchange positions (2, 3) (Figure 2.1).

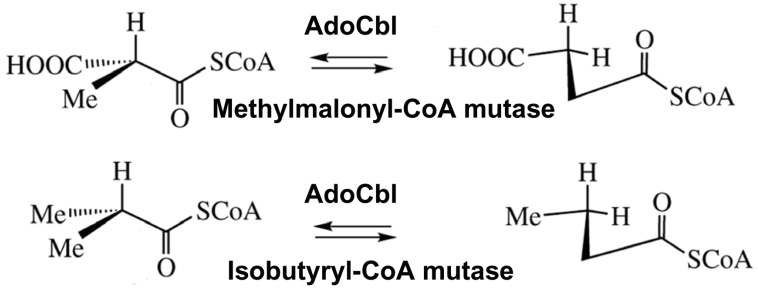


Figure 2.1: **Reactions catalyzed by MCM and ICM.**

The genes encoding ICM were first cloned and sequenced from the Gram-positive, filamentous soil bacterium *Streptomyces cinnamonensis.* ICM is an $\alpha_2\beta_2$-heterotetramer composed of two large subunits (IcmA) of 62.5 kDa and two small subunits (IcmB) of 14.3 kDa. The genes encoding the subunits of MCM, an αβ heterodimer in some

bacteria, are usually located in a single operon (Figure 2.2). In contrast, the *icmA* and *icmB* genes are distant from each other in the genome of *S. cinnamonensis* (2).

The sequence of IcmA is very similar to the sequences of the large subunit of MCM, with the exception of the AdoCbl-binding region, which is missing. Thus, IcmA lacks a C-terminal AdoCbl-binding domain containing the signature DxHxxG motif that is present instead in the small IcmB subunit, which binds the cofactor (5). In this respect, ICM resembles some other AdoCbl-dependent mutases that exhibit a similar organization (Figure 2.2).

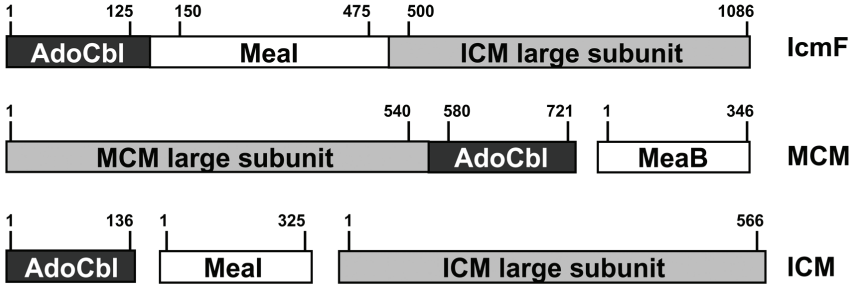


Figure 2.2: Comparison of domain and gene organizations of bacterial IcmF, MCM and ICM. MeaB and MeaI represent the P-loop GTPase chaperones for MCM and ICM respectively.

For example, glutamate mutase is also composed of two subunits of very different sizes: the large subunit MutE, which binds substrate and the small subunit MutS, which binds AdoCbl (6).

The ICM-catalyzed reaction plays an important role in polyketide biosynthesis in *Streptomyces*. In studies with ^{13}C-labeled isobutyrate, it was shown that this compound efficiently incorporates into monensin A, tylosin and leucomycin at positions derived from n-butyrate (1). While ICM was believed to have a rather limited distribution, its close sequence relative, MCM, is present in organisms ranging from bacteria to man (7).

A G-protein chaperone, MeaB, shows strong operonic association with MCM and mutations in the human ortholog, the product of the *cbl*A locus, result in methylmalonic aciduria due to dysfunctional MCM activity (8). MeaB from *Methylobacterium extorquens* has been characterized most extensively and is a P-loop GTPase (*9, 10*). Other members of this group of proteins include HypB, UreG which are important in the assembly of the metalloenzymes: nickel hydrogenases (11) and urease (12). MeaB has been proposed to function in the GTP-dependent assembly of holo-MCM and shown to protect the radical intermediates formed during MCM catalysis from oxidative interception (14). MCM in turn, influences the GTPase activity of MeaB increasing it by >100–fold. Hence, MCM exhibits GAP (GTPase activating protein) activity for MeaB (*14, 15*). In addition, MCM modulates the affinity of MeaB for nucleotides. The crystal structure of MeaB in the presence of GDP has been solved and confirms that it is a member of the G3E family of GTPases but differs from other family members in possessing N- and C-terminal extensions of unknown function (9). Structural insights into the interaction between MeaB and MCM are lacking.

In the present study, we show that in >70 bacteria ICM is fused to a P-loop GTPase, which is a paralog of MeaB. This fusion protein that we have named IcmF (for ICM-fused) is described as a putative MCM-like protein in the databases. The misannotation has led to the ascription of this gene product as representing a fusion between MCM and MeaB (16) and to its function in pathways that are unlikely to be correct (17). Using bioinformatics and biochemical approaches, we demonstrate that IcmF is an ICM with ICM and GTPase activities. IcmF represents an important paradigm for elucidating the crosstalk between a mutase and its auxiliary protein during the catalytic cycle.

2.3 Experimental procedures

2.3.1 Cloning and Expression of IcmF. IcmF from three organisms was cloned into the pET-30 Ek/LIC expression vector (Novagen, CA). The genomic DNA of *Geobacillus kaustophilus* and *Burkholderia xenovorans* (formerly known as *Burkholderia fungorum*) were generous gifts from Hideto Takami (Japan Agency for Marine-Earth Science and Technology, Kanagawa, Japan) (18). The genomic DNA clone KNL023_G20 in the pTS1 plasmid containing the *Nocardia facinica* IcmF gene was obtained from Jun Ishikawa (National Institute of Infectious Diseases, Tokyo, Japan) (19). *(i) IcmF from G. kaustophilus*. The IcmF gene from *G. kaustophilus* was amplified from genomic DNA using nested-primer PCR as attempts to amplify the full-length gene in a single PCR reaction were unsuccessful. The first round of PCR was performed with the following primers: Forward5′-TCTACCGATCTGCTAAAGTTCAACG-3′ and reverse 5′-GGATTATGGAGAAACAGCGAGTC-3′. The second round of amplification was performed with the following primers containing NheI and BamHI restriction sites (underlined): Forward 5′-TAG**GCTAGC**ATGGCGCACATTTACCGTCCG-3′ and reverse 5′-TAG**GGATCC**TTACATATTCCGCCGGTATTGTCC-3′. The resulting fragment was cloned into pGEM-T easy (Promega, WI) and subsequently used as a template for ligase-independent cloning (LIC). The insert was amplified with the following primers for LIC cloning (forward 5′-GACGACGACAAGATGGCGCACATTTACCGTCCGAAG-3′ and reverse 5′-GAGGAGAAGCCCGGTTTACATATTCCGCCGGTATTG-3′) and inserted in the pET30 Ek/LIC vector according to the manufacturer's protocol. *(ii) IcmF from B. xenovorans.* The first round of nested-primer PCR was performed on genomic DNA of *B.*

xenovorans with the following primers: forward 5′-TGTCGACTTCCTCGCTGAGCGGTT-3′ and reverse 5′-CGCGACGCGTTGTGGT TGTGCGTT-3′. The second round of nested PCRs was performed with primers for LIC: forward 5′-GACGACGACAAGATGACCGATCTGTCCACGCCG-3′ and reverse 5′-GAGGAGAAGCCCGGTTTACATATTGCGGCGGTACTG-3′. *(iii) IcmF from N. farcinica.* The IcmF gene was amplified from the pTS1 plasmid (genomic DNA clone KNL023_G20) with the following primers containing NdeI and HindIII restriction sites that are underlined: forward 5′-ATATATCATATGGCCGACAGTACGCTCCACCAA-3′ and reverse 5′-ATATCTAAGCTT TCACACGTTGCGCCGGTACTG-3′. The resulting PCR product was subcloned into the pGEM-T vector (Promega, WI) and used for LIC cloning with the following primers. Forward: 5′GACGACGACAAGATGGCCGACAGTACGCTCCAC-3′ and Reverse: 5′- GAGG AGAAGCCCGGTTCACACGTTGCGCCGGTA-3′. The sequences of all the resulting constructs were verified by nucleotide sequence determination at the University of Nebraska's Genomics Facility.

2.3.2 Protein Expression and Purification. Recombinant *M. elsdenii* butyryl-CoA dehydrogenase expressed in *Escherichia coli* BL21 (DE3) was purified as previously described (20). The pET-30 Ek/LIC vector with the *G. kaustophilus* IcmF gene was transformed into *E. coli* BL21 (DE3) cells, which were grown at 37˚C in Luria Bertani (LB) medium containing 50 μg/mL kanamycin to an optical density at 600 nm of 0.6. Cells were grown for 14-16 h after induction with 0.5 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) at 15˚C. The cells were resuspended in ~150 mL lysis buffer (50 mM sodium phosphate buffer (NaPi), pH 8.0, 500 mM NaCl and 8 mM imidazole)

containing three protease inhibitor cocktail tablets (Roche) and disrupted by sonication on ice (output setting of 5.5 for 10 min with 30 sec bursts and 3 min breaks). Following centrifugation, the cell lysate was subjected to dilution to a final concentration of 3-5 mg/mL and loaded onto a 50 mL Ni-NTA Sepharose column. After washing with 10-20 column volumes of lysis buffer, the protein was eluted with a gradient of 8 to 250 mM imidazole in lysis buffer. Fractions containing IcmF were pooled, concentrated and dialyzed against 50 mM NaPi, pH 7.5 and applied to a 2.5 x 7.5 cm Source 15Q column equilibrated at a flow rate of 10 mL/min with Buffer A (50 mM NaPi, pH 7.5, 50 mM NaCl). The column was then washed at the same flow rate with 100 ml of Buffer A and eluted with a 500 mL gradient from 50 to 250 mM NaCl in 50 mM NaPi, pH 7.5 over 50 min at the same flow rate. The purified IcmF was concentrated and loaded on a 160 mL Superdex-200 column (GE Healthcare) equilibrated at a flow rate of 0.75 mL/min with 50 mM NaPi, pH 7.5, 250 mM NaCl. Under these conditions, IcmF eluted with a retention volume of ~82 mL. Fractions containing active IcmF were pooled, concentrated, flash-frozen in liquid nitrogen and stored at -80˚C until further use. Approximately 25-35 mg of recombinant *G. kaustophilus* IcmF was obtained from a 6L culture.

E. coli BL21 (DE3) cells transformed with recombinant *N. farcinica* or *B. fungorum* IcmF was grown and the cell extracts were prepared as described above for the *G. kaustophilus* enzyme using potassium phosphate (KPi) buffer. The cell extracts were loaded onto a 4 mL Ni-NTA column, washed with 50 mM KPi, pH 8, containing 50 mM imidazole and eluted with the same buffer containing 250 mM imidazole. IcmF-

containing fractions were pooled and the protein was obtained in ~40% purity from this one-step purification procedure.

2.3.3 GTPase activity of IcmF. The GTPase activity of IcmF (5 μM) was determined in the presence of varying concentrations of GTP (50–5000 μM) at 37°C in 0.4 mL of 50 mM KPi buffer, pH 7.5, 100 mM KCl and 5 mM $MgCl_2$. For each GTP concentration, aliquots (50 μl) were removed at varying time points (2-60 min), quenched with 2 M trichloroacetic acid (10% v/v), centrifuged and filtered through a 0.1 μm Ultrafree-MC filter (Millipore) to remove the precipitated protein.

The nucleotides were analyzed by ion exchange chromatography on a μBondapak NH_2 300 x 3.9 mm HPLC column (Waters, MA). Initial conditions were 100% Buffer B (50 mM monobasic KPi, pH 4.5) and 0% Buffer C (800 mM monobasic KPi, pH 4.5) and a flow rate of 1.0 mL/min. Between 5 and 20 min, Buffer C was increased to 80% and held at that concentration for 5 min. Between 25-26 min, Buffer C was decreased to 0% and held for 10 min at that composition to equilibrate the column between injections. Under these conditions, the retention time for GDP was 9.5 min and for GTP was 13.1 min respectively.

2.3.4 Enzyme assays. Initially, recombinant IcmF was assayed for MCM activity as previously described using the radioactive assay (21). To monitor IcmF activity, one of two assay methods was used. First, a fixed-time GC/MS-based assay was employed by a modification of a previously described method (*1, 22*). In this assay, normal- and isobutyryl-CoA thioesters were saponified, and the resulting free acids were extracted into ethylacetate. Product formation was followed in a 200 μl assay mixture containing 50 mM KPi pH 7.5, 100 mM KCl, n-butyryl-CoA or isobutyryl-CoA (0.1 to 1 mM), 50

μM AdoCbl and 0.5-5 μg of IcmF. The reaction was stopped by the addition of 100 μL of 2N KOH containing 0.18 mM valeric acid as an internal standard followed by addition of 100 μL of H_2SO_4 (15%, v/v). In the last step of sample preparation, the reaction mixture was saturated with NaCl and extracted with ethylacetate (250 μL). An aliquot of the extract (5 μL) was subjected to analysis by GC/MS using a DB-FFAP 30 m x 0.25 mm I.D., 0.25 μm capillary column (Agilent, CA). This column is especially designed for the separation of organic acids without derivatization.

A continuous assay was developed to determine the kinetic parameters for IcmF. In this assay, n-butyryl-CoA, which is produced from isobutyryl-CoA by IcmF is converted to crotonyl-CoA by butyryl-CoA dehydrogenase (BDH). BDH activity was followed by the decrease in absorbance at 300 nm over 1-2 min upon reduction of ferricenium hexafluorophosphate ($Fc^+PF_6^-$) ($\Delta\varepsilon$=4.3 $mM^{-1}cm^{-1}$) (23). BDH is able to use both isobutyryl-CoA and n-butyryl-CoA as substrates but with a preference for the latter. We found that with isobutyryl-CoA, K_M = 311 ± 26 μM and V_{max} = 1.66 ± 0.06 μmol/min/mg and with n-butyryl-CoA, K_M = 68 ± 4 μM and V_{max} = 33 ± 1 μmol/min/mg /mg. The reaction mixture for the coupled assay contained in a final volume of 200 μL: 2-4 μg IcmF, 50 μM AdoCbl, 0.2 μg BDH, varying concentrations of isobutyryl-CoA (10 to 1000 μM), 250 μM ($Fc^+PF_6^-$) ± 1-2 mM GDP, GTP or GMPPNP in 50 mM NaPi, pH 7.5, 250 mM NaCl. Under these conditions, the consumption of isobutyryl-CoA by BDH is negligible and similar to the background rate observed in the absence of BDH. The dye was preincubated for 3 min at 37°C before adding the substrate. After 1 min incubation, BDH was added and the reaction was started 1 min later by the addition of holo-IcmF.

2.3.5 UV-visible spectroscopy. UV-visible spectra were recorded on a Cary100 spectrophotometer (Agilent, formely Varian, CA). Holo-IcmF (10-12 μM) in 50 mM NaPi, pH 7.5, 0.25 M NaCl ± 5 mM $MgCl_2$ and 1-2 mM GDP, GTP or GMPPNP was incubated in the presence of 3-5 mM isobutyryl-CoA at 20°C. The spectra were acquired after 2-5 min incubation.

2.3.6 Isothermal Titration Calorimetry. The isothermal titration calorimetric experiments were performed as described previously (*14, 15*). Each experiment was performed in triplicate. IcmF was dialyzed for 10-12 h against 50 mM NaPi, pH 7.5, 0.25 M NaCl containing 1-2 mM TCEP (Buffer D) before use. The protein (8-24 μM) ± 1-2 mM GDP or GMPPNP in Buffer D was titrated with thirty to forty two 7-9.7 μL aliquots of a 15-20-molar excess solution of AdoCbl at 20°C. The calorimetric signals were integrated and the data were analyzed with Microcal ORIGIN software using a two-sites binding model to determine the thermodynamic parameters associated with AdoCbl binding to IcmF.

2.3.7 EPR Spectroscopy. EPR spectra were recorded on a Bruker EMX spectrometer (Bruker Biospin Corp., Billerica, MA), equipped with an Oxford ITC4 temperature controller, a Hewlett-Packard model 5340 automatic frequency counter, and a Bruker gaussmeter. Unless otherwise noted, the following parameters were used: temperature, 100 K; microwave power, 25 mW; microwave frequency, 9.38 GHz; receiver gain, 2×10^5; modulation amplitude, 10 G; modulation frequency, 100 kHz. Cob(II)alamin was generated by treating a solution of hydroxocobalamin (OHCbl) with 4-7 molar excess of TCEP. Formation of cob(II)alamin was followed by UV-visible spectroscopy and the concentration of the solution was estimated using $\varepsilon_{473nm} = 9.2\ mM^{-1}cm^{-1}$.

2.3.8 Bioinformatics analysis. STRING was used to find functional linkages for proteins of interest, as well as gene fusions and gene neighborhoods (string.embl.de) (24). A protein-protein blast search (www.ncbi.nlm.nih.gov/BLAST) was used to perform distant homologs searching. A multiple sequence alignment and phylogenetic tree were constructed using a stand-alone version of ClustalX v.1.8. Figures with multiple sequence alignments were generated using BOXSHADE 3.21 (http://www.ch.embnet.org/software/BOX_form.html). Phylogenetic analysis was carried out using default parameters in ClustalX. The trees were visualized using TreeView 1.6.6. Operon and regulon browsers on the Microbes Online web site were used for the elucidation of functional predictions for the genes of interest (http://www.microbesonline.org, http://www.microbesonline.org/operons/OperonList.html) (25).

2.4 Results and Discussion

2.4.1 Bioinformatics analysis of IcmF. *(i) Analysis of the mutase domains.* Based on bioinformatics analysis, it was previously concluded that MCM either colocalizes in the same operon with its chaperone, MeaB, or that MeaB is fused to the large subunit of MCM in some bacteria (16). Indeed, the putative fusion protein between MCM and MeaB in *B. xenovorans* was reported to possess MCM activity (16). Our laboratory has been elucidating the influence of MeaB and MCM on the substrate binding and catalytic activities of each other (14). Since the kinetics of a fusion protein are easier to characterize than the stand-alone versions of the component proteins, which interact with varying affinities depending on the ligand, we chose to focus on the putative MCM-MeaB fusion protein. A BLAST search using the fusion protein from *B. xenovorans* (YP_556774) as a query sequence resulted in the identification of >70 proteins in bacteria, including the seven proteins that were previously identified as examples of fusions between MCM and MeaB (Table 2.1) (16). In the databases, homologs of this fused protein are annotated as putative MCM-like proteins. However, a careful examination of the domain organization and sequence analysis of the substrate-binding site in the B_{12}-dependent isomerase component (Figure 2.2-2.4), suggested that this group of fusion proteins might in fact be misannotated.

Based on the high sequence similarity between MCM and ICM, Robinson and colleagues used the crystal structure of MCM from *P. shermanii* to identify residues which might be involved in specific substrate binding in ICM from *S. cinnamonensis* (2). They identified two key substitutions in the large subunit of MCM, Tyr89 and Arg207, which are replaced by Phe80 and Gln198 in the large subunit of ICM (Figure 2.3).

Name of the organism	Accession number
Anaeromyxobacter sp. Fw109-5	ZP_01670373
Algoriphagus sp.	ZP_01718212
Anoxybacillus flavithermus	YP_002317072
Azoarcus sp. BH72	YP_932200
Azoarcus sp. EbN1	YP_159147
Bacillus coagulans	ZP_01696637
Bacillus coahuilensis	ZP_03228090
Bacillus halodurans C-125	NP_244663
Bacillus selenitireducens	ZP_02169380
Bacillus sp. NRRL B-14911	ZP_01168476
Bdellovibrio bacteriovorus	NP_968715
Burkholderia graminis	ZP_02882806
Burkholderia phymatum STM815	ZP_01499709
Burkholderia phytofirmans	YP_001894080
Burkholderia sp. H160	ZP_03267069
Burkholderia xenovorans LB400	YP_556774
Candidatus Desulfococcus oleovorans Hxd3	ZP_01674860
Cellulophaga sp. MED134	ZP_01050856
Croceibacter atlanticus HTCC2559	ZP_00950324
Cupriavidus taiwanensis	YP_001795613
Dechloromonas aromatica RCB	YP_283472
Desulfatibacillum alkenivorans	ZP_02132705
Flavobacteriales bacterium	ZP_02182179
Flavobacterium sp. MED217	ZP_01059114
Flovobacterium johnsoniae UW101	ZP_01244537
Frankia alni ACN14a	YP_716016
Frankia sp. CcI3	YP_482733
Frankia sp. EAN1pec	ZP_00569662
Geobacillus kaustophilus HTA426	YP_149244
Geobacter bemidjiensis	YP_002139638
Geobacter metallireducens GS-15	YP_384678
Geobacter uraniumreducens	ZP_01142042
Gramella forsetii KTO0803	YP_860533
Heliobacterium modesticaldum	YP_001679677
Janibacter sp. HTCC2649	ZP_00993974
Kordia algicida	ZP_02163705
Leptospira biflexa serovar Patoc	YP_001964131
Leptospira borgpetersenii serovar Hardjo-bovis JB197	YP_801321
Leptospira interrogans serovar Lai str. 56601	NP_713136

Table 2.1: List of all identified IcmFs.

These differences in active site residues can be rationalized based on the structural difference between the respective substrates despite the very similar reactions catalyzed by the two enzymes (Figure 2.1).

Name of the organism	Accession number
Leptospira interrogans serovar copenhageni str. Fiocruz L1-130	YP_001075
Leptothrix cholodnii	YP_001792863
Limnobacter sp.	ZP_01914275
Lysinibacillus sphaericus C3-41	YP_001696718
marine gamma proteobacterium HTCC2207	ZP_01224856
Methylibium petroleiphilum	YP_001020100
Microscilla marina	ZP_01688358
Myxococcus xanthus DK 1622	YP_630482
Nitrosococcus oceani ATCC 19707	YP_343889
Nocardia farcinica IFM 10152	YP_117245
Nocardioides sp. JS614	YP_925327
Oceanobacter sp. RED65	ZP_01305689
Paenibacillus sp.	ZP_02848947
Polaribacter irgensii 23-P	ZP_01119326
Polaromonas naphthalenivorans CJ2	YP_983742
Polaromonas sp. JS666	YP_550845
Psychroflexus torquis ATCC 700755	ZP_01252245
Ralstonia eutropha H16	YP_724799
Ralstonia eutropha JMP134	YP_294479
Ralstonia metallidurans CH34	YP_582365
Ralstonia pickettii 12J	ZP_01661954
Ralstonia solanacearum GMI1000	NP_518358
Ralstonia solanacearum UM551	ZP_00945070
Reinekea sp. MED297	ZP_01115975
Rhodoferax ferrireducens T118	YP_524846
Robiginitalea biformata HTCC2501	ZP_01121607
Rubrivivax gelatinosus PM1	ZP_00242991
Salinibacter ruber	YP_445808
Stigmatella aurantiaca	ZP_01465914
Streptomyces sviceus	YP_002202957
Tenacibaculum sp. MED152	ZP_01052807
Thauera sp. MZ1T	ZP_02841697
Thermobifida fusca YX	YP_290867
unidentified eubacterium SCB49	ZP_01889691

Table 2.1: (Continued): **List of all identified IcmFs.**

In MCM, the carboxylate group of methylmalonyl-CoA is engaged in electrostatic interactions with the guanidinium group of Arg207 and the phenolic group of Tyr89 (Figure 2.3). The presence of a methyl group in the ICM substrates instead of the carboxylate is reflected in the loss of the hydrogen bond donating arginine and tyrosine residues. Instead, a glutamine and phenylalanine in ICM substitute for the arginine and tyrosine residues respectively in MCM (Figure 2.3). Apart from these two differences, the remaining residues in the active sites of both mutases are highly conserved.

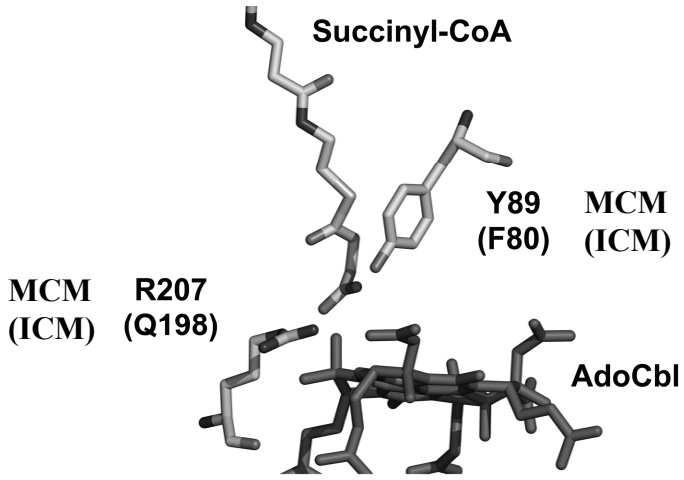


Figure 2.3: Comparison of the active site residues in *P. shermanii* MCM with those predicted for *S. cinnamonensis* ICM. The MCM structure was obtained from the pdb file, 4REQ. The two striking differences in the active site residues are the substitutions of Tyr89 and Arg207 in MCM to Phe80 and Gln198 in ICM.

Multiple sequence alignment of the predicted substrate-binding site in the C-termini of all the identified fusion proteins clearly reveals conservation of the Phe and Gln residues (Figure 2.4 and Figure S1 in (*26*)). This analysis strongly suggests that the substrate for the fusion protein is n-butyryl-CoA/isobutyryl-CoA and hence the fusion protein is predicted to be an ICM not an MCM. We thus designate this fusion protein as IcmF, for **i**sobutyryl-**C**oA **m**utase **f**used. The IcmF designation for this group of fusion proteins also distinguishes it from the "stand-alone" ICM described for the genus *Streptomyces*.

All IcmFs are predicted to be comprised of three domains: the N-terminal AdoCbl-binding region that is homologous to the small subunit of ICM, a middle P-loop GTPase domain and a C-terminal region that is homologous to the large substrate-binding subunit of ICM (Figure 2.2).

Clear sequence similarities are seen between the AdoCbl-binding regions of the large subunit of MCM, the small subunit of ICM (IcmB) and the N-terminal portion of IcmF (Figure 2.5).

Figure 2.4: Multiple sequence alignment of the C-terminal sequences of IcmFs, the large subunit of ICM (IcmA) from *S. cinnamonensis* (AAC08713), MCM from *M. extorquens* (YP_001642233) and MCM from *P. shermanii* (CAA33090). The C-terminal region of the IcmFs is homologous to IcmA and MCM. The two conserved residues in IcmF (Phe and Gln in IcmA) that are important for substrate binding are highlighted in grey and indicated with asterisks. In MCM, these residues are substituted by Tyr and Arg respectively.

The signature DxHxxG… SxL…GG motif (where x is any amino acid) used for binding of B_{12} in the "base-off/His-on" conformation is observed in all three proteins (5). However, in IcmF, this motif is similar but not identical to that seen in ICM and MCM. First, a G→A/S change is found in IcmFs in the conserved sequence: DxHxxA/S… SxY..GGGG. This substitution is not surprising since glycine is often replaced by alanine or serine and *vice versa* in sequences of orthologous proteins from different organisms.

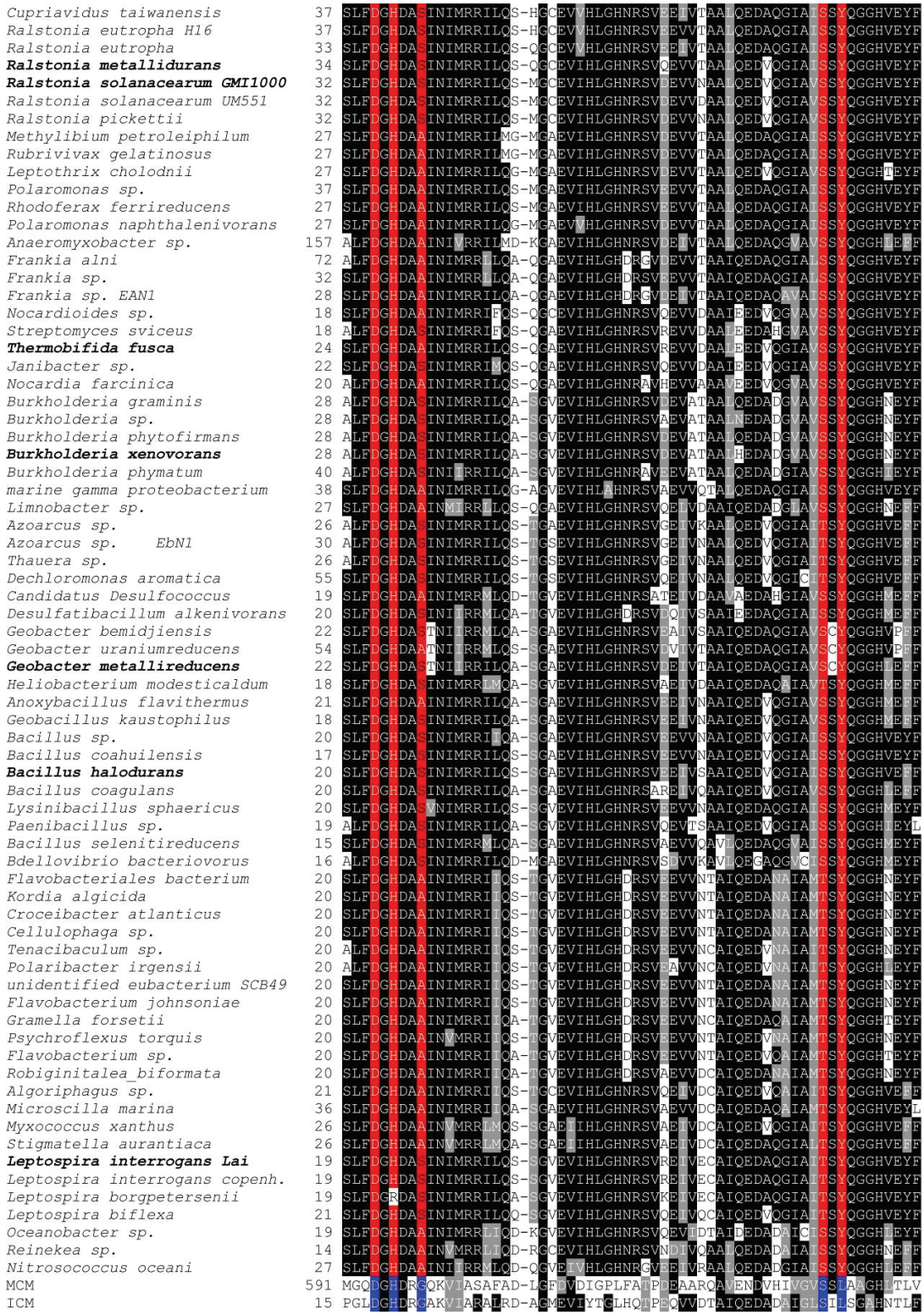


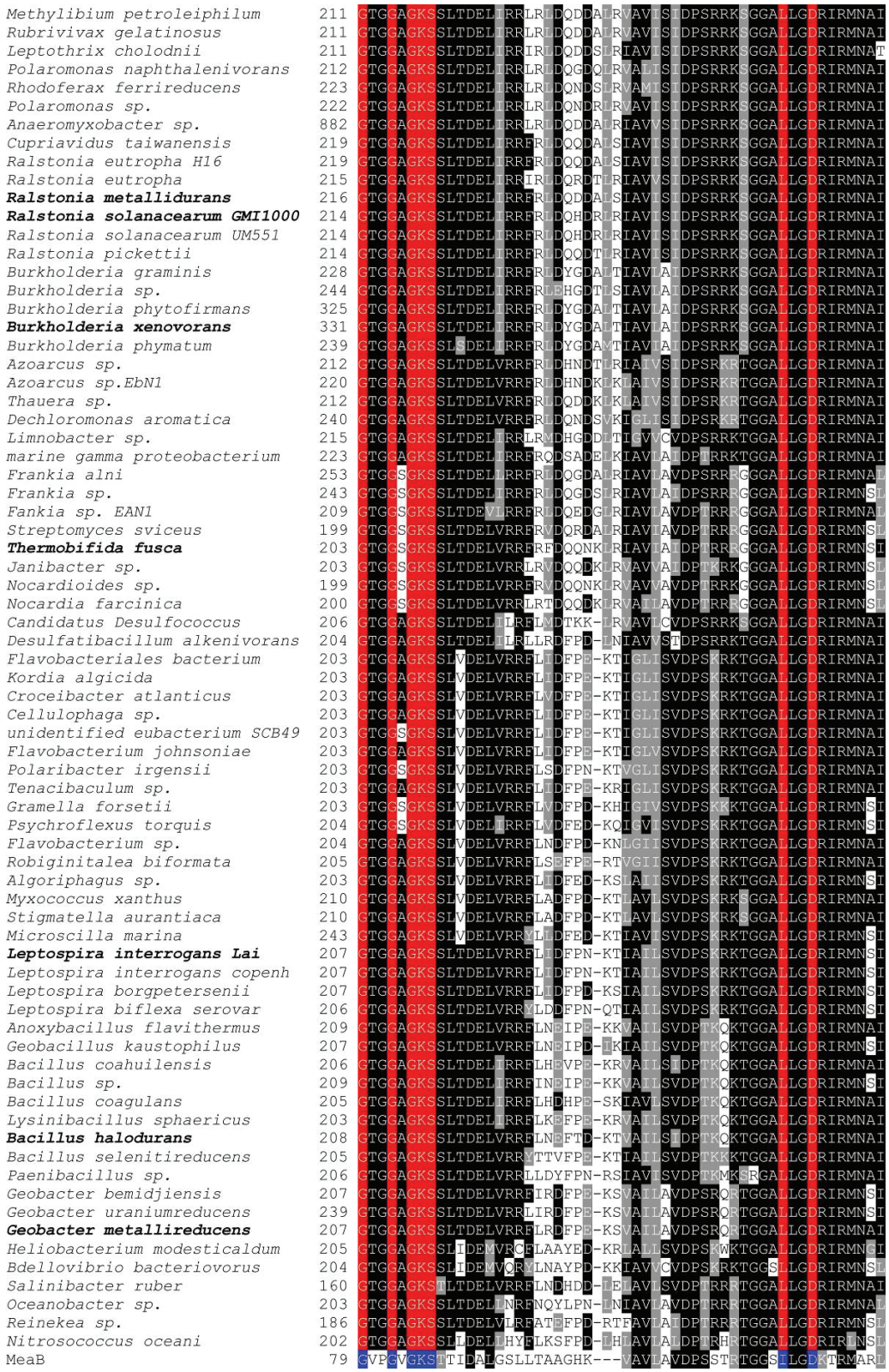
Figure 2.5: **Multiple sequence alignment of the N-terminal AdoCbl-binding domain of IcmFs, the small subunit of ICM (IcmB) from *S. cinnamonensis* (CAB59633) and MCM from *M. extorquens* (YP_001642233).** In B_{12} proteins that bind the cofactor in a "base-off/His-on" conformation, a signature DxHxxG… SxL…GG motif is found

(highlighted in blue). In IcmF, this motif is similar but not identical: DxHxxA/S… S/TxY…GGGG (highlighted in red). In *Leptospira borgpetersenii* the histidine which is predicted to coordinate to AdoCbl appears to be substituted by arginine. The *Salinibacter ruber* sequence is not included in the alignment since the AdoCbl region is truncated at the N-terminus (DxHxxG motif is missing). IcmFs, which were previously annotated as MCM-like enzymes are indicated in bold. For accession numbers see Table 2.1.

Second, an SxL→SxY substitution is seen in IcmF. The rationale for the leucine to tyrosine substitution and the insertion of two glycines at the end of the motif are not clear. Thus, it appears that the sequence of the AdoCbl-binding domain of IcmF has diverged from the corresponding sequences in IcmB and MCM.

(ii) Analysis of the G-protein domains. The P-loop GTPase domain has very high sequence similarity to MeaB from *M. extorquens* that was shown to be a chaperone for MCM (*14, 15*). We designated it as MeaI to distinguish it from the MeaB-like chaperone. The sequence encoding the MeaI domain of IcmF comprises ~250-300 amino acids in the middle of the protein (Figure. 2.6).

Sequence analysis strongly suggests that this domain belongs to the G3E family of P-loop GTPases (10). All four GTPase sequence fingerprints, the so-called G-domains, which define this family, are present in the MeaI domain of IcmF. These include: *(i)* the Walker A motif (G1), which binds the triphosphate moiety of GTP (as is typical for the G3E subfamily a slight modification of the GxxGxGK[ST] sequence to GxxGxGK[SS] is seen); *(ii)* the Mg^{2+} binding motif (G2) (LxxD in all IcmF sequences); *(iii)* the DxxxxExxG Walker B motif (G3) and *(iv)* the nucleotide specificity NKxD motif (G4). The replacement of Asp by Glu in NKxD motif in some proteins (Figure 2.6) might not affect the specificity for the guanine nucleotide and is also seen in other G proteins (*27*) (See also section 3.4.7).


Methylibium petroleiphilum 211
Rubrivivax gelatinosus 211
Leptothrix cholodnii 211
Polaromonas naphthalenivorans 212
Rhodoferax ferrireducens 223
Polaromonas sp. 222
Anaeromyxobacter sp. 882
Cupriavidus taiwanensis 219
Ralstonia eutropha H16 219
Ralstonia eutropha 215
Ralstonia metallidurans 216
Ralstonia solanacearum GMI1000 214
Ralstonia solanacearum UM551 214
Ralstonia pickettii 214
Burkholderia graminis 228
Burkholderia sp. 244
Burkholderia phytofirmans 325
Burkholderia xenovorans 331
Burkholderia phymatum 239
Azoarcus sp. 212
Azoarcus sp.EbN1 220
Thauera sp. 212
Dechloromonas aromatica 240
Limnobacter sp. 215
marine gamma proteobacterium 223
Frankia alni 253
Frankia sp. 243
Fankia sp. EAN1 209
Streptomyces sviceus 199
Thermobifida fusca 203
Janibacter sp. 203
Nocardioides sp. 199
Nocardia farcinica 200
Candidatus Desulfococcus 206
Desulfatibacillum alkenivorans 204
Flavobacteriales bacterium 203
Kordia algicida 203
Croceibacter atlanticus 203
Cellulophaga sp. 203
unidentified eubacterium SCB49 203
Flavobacterium johnsoniae 203
Polaribacter irgensii 203
Tenacibaculum sp. 203
Gramella forsetii 203
Psychroflexus torquis 204
Flavobacterium sp. 204
Robiginitalea biformata 205
Algoriphagus sp. 203
Myxococcus xanthus 210
Stigmatella aurantiaca 210
Microscilla marina 243
Leptospira interrogans Lai 207
Leptospira interrogans copenh 207
Leptospira borgpetersenii 207
Leptospira biflexa serovar 206
Anoxybacillus flavithermus 209
Geobacillus kaustophilus 207
Bacillus coahuilensis 206
Bacillus sp. 209
Bacillus coagulans 205
Lysinibacillus sphaericus 203
Bacillus halodurans 208
Bacillus selenitireducens 205
Paenibacillus sp. 206
Geobacter bemidjiensis 207
Geobacter uraniumreducens 239
Geobacter metallireducens 207
Heliobacterium modesticaldum 205
Bdellovibrio bacteriovorus 204
Salinibacter ruber 160
Oceanobacter sp. 203
Reinekea sp. 186
Nitrosococcus oceani 202
MeaB 79

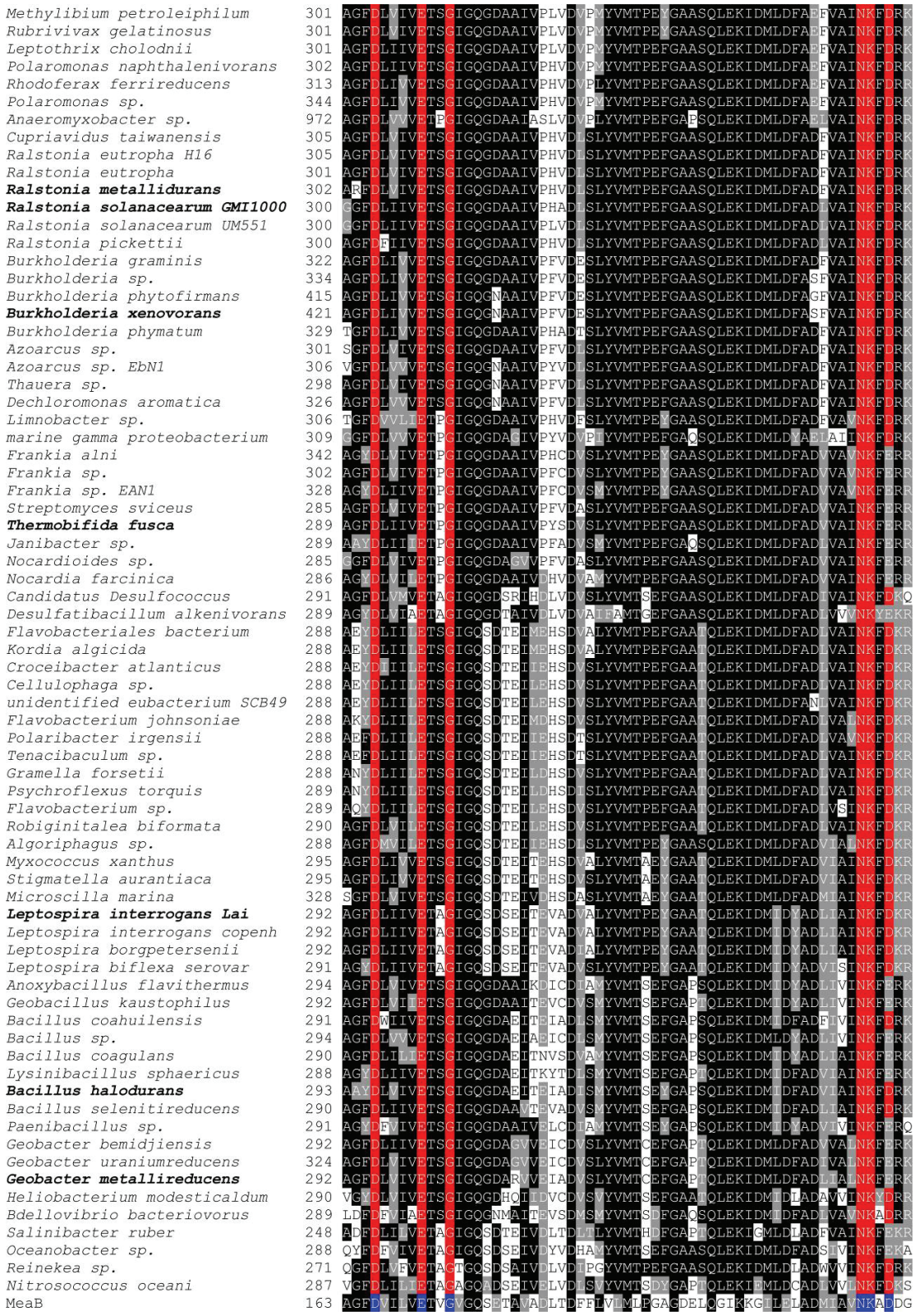


Figure 2.6: Multiple sequence alignment of the MeaI domain in IcmF sequences and MeaB from *M. extorquens* (YP_001637793). The MeaI domain of IcmF comprises

~250-300 amino acids in the middle of the protein. Sequence analysis strongly suggests that this domain belongs to the G3E family of P-loop GTPases. The G-domains, which defines this family, are present in the MeaI domain of IcmF and are highlighted in red. These include: *(i)* the Walker A GxxGxGK[SS] motif; *(ii)* the Mg^{2+} binding motif, which is typically V/IxxD in proteins in the G3E family and is LxxD in all IcmF sequences *(iii)* the DxxxxExG Walker B motif and *(iv)* the nucleotide specificity NKxD/E motif. IcmFs, which were previously annotated as MCM-like enzymes, are indicated in bold. For accession numbers see Table 2.1.

Based on this analysis, we conclude that the MeaI domain of IcmF, like MeaB, belongs to the SIMIBI subclass of G proteins since two key aspartate residues at the N-terminus of the Walker B motif and in the Mg^{2+}-binding motif are present (Figure 2.6). Within the SIMIBI subclass, the MeaI domain of IcmF belongs to the G3E family (the conserved glutamate residue in the Walker B motif is a signature of this family) as well as the intact nucleotide specificity motif (10). We predicted that the MeaI domain of IcmF functions like its paralog MeaB, i.e. as a chaperone for ICM in the fusion protein (14).

2.4.2 Expression and initial activity analysis of IcmF. In order to test the prediction from bioinformatics analysis that IcmF harbors ICM rather than MCM activity, IcmFs from three organisms, *G. kaustophilus*, *B. xenovorans* and *N. farcinica* were cloned into the expression vector pET30 Ek/LIC. Multiple IcmF encoding genes were subcloned and purified so that the enzymatic activity of the fusion protein from more than one organism could be assessed. Since the fusion protein from *B. xenovorans* was reported to have MCM activity (16), we initially tested the activity of all three IcmF proteins in the standard radiolabeled assay for MCM (21). However, none of the three IcmFs exhibited detectable MCM activity. On the other hand, all three IcmFs exhibited ICM activity. Two assays have been described for monitoring ICM activity and are based on either gas chromatography (GC) or NMR-based detection of the reactant and product (*28*). Since

the NMR-based method is not amenable for routine enzymatic assays, we used a modification of the previously described GC assay (*1, 22*) using mass spectrometry (MS) for detection of the reaction components. A specific activity of 0.6 ± 0.04 μmol $min^{-1}mg^{-1}$ protein at 37°C was obtained for the *G. kaustophilus* IcmF.

As an alternative to the GC-MS assay that depends on access to specialized instrumentation, a coupled spectrophotometric assay was developed to measure IcmF activity as described (Section 2.3.4). The specific activity determined in the coupled assay for the *N. farcinica* IcmF was 1.1±0.1 μmol $min^{-1}mg^{-1}$ and for the *B. xenovorans* IcmF was 0.34 ± 0.04 μmol $min^{-1}mg^{-1}$ at 37°C. A specific activity of 0.75 ± 0.01 μmol min^{-1} mg^{-1} protein at 37°C was measured for *G. kaustophilus* IcmF, which is comparable to the value from the GC-MS assay. In comparison, a V_{max} of 38 ± 3 μmol $min^{-1}mg^{-1}$ at 37˚C has been reported for purified stand-alone ICM from *S. cinnamonensis* (2).

The recombinant *G. kaustophilus* IcmF was the most stable and soluble of the three proteins and was further purified to ~95% purity as described (Section 2.3.2) to perform biochemical and biophysical characterizations. Based on its elution from a calibrated gel-filtration column, the *G. kaustophilus* IcmF appears to be a dimer with a native molecular mass of ~286 kDa (Figure 2.7).

2.4.4 Binding of AdoCbl to IcmF ± nucleotides. We investigated the energetics of AdoCbl binding to *G. kaustophilus* IcmF ± nucleotides by ITC (Figure 2.8, Table 2.2). These experiments revealed the presence of two non equivalent binding sites with an ~9-25-fold difference in affinity for AdoCbl that was influenced by the presence and identity of the guanine nucleotide (Table 2.2). Binding of AdoCbl to the high affinity site in the absence of nucleotides (K_D=81 ± 14 nM) is accompanied by a $\Delta G_1°$ of -9.5 ± 0.1

kcal/mol that is enthalpically favored whereas binding to the low affinity site (K_D=2.0 ± 0.4 μM) is entropically driven (Table 2.2).

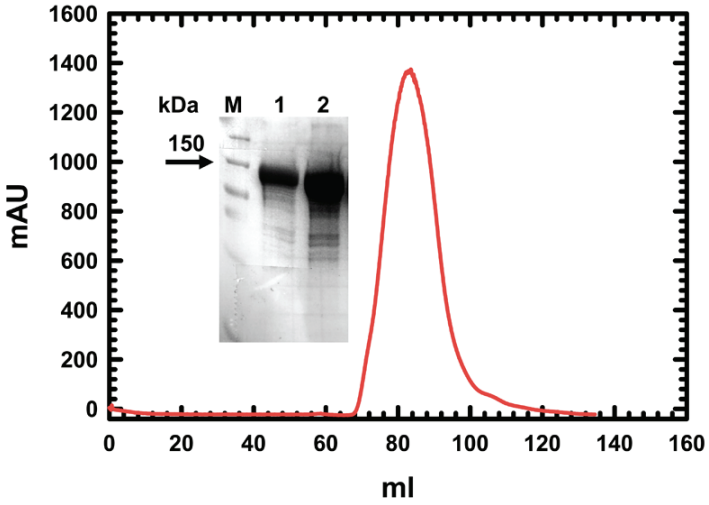


Figure 2.7: Gel-filtration of IcmF from *G.kaustophilus*. mAU, milliabsorbance units. Inset: The SDS-PAGE analysis of IcmF after gel-filtration (lane #1, 15 μg load; lane #2, 30 μg load).

These data suggest a possible difference in the flexibility of the two AdoCbl binding sites in IcmF. A K_{act} of 12 ± 2 μM for AdoCbl for ICM from *S. cinnamonensis* has been reported (2).

We next analyzed the influence of nucleotides on cofactor binding. The affinity for AdoCbl for the high affinity site was slightly increased in the presence of GDP (132 ± 9 nM), which resulted from changes in both enthalpic and entropic contributions ($\Delta\Delta H_1$~2.5 kcal/mol and $\Delta T\Delta S_1$~2.8 kcal/mol). GDP did not substantially influence binding of AdoCbl to the second site. Binding of AdoCbl in the presence of GMPPNP, a nonhydrolyzable analogue of GTP (Figure 2.8, Table 2.2) indicates that GTP hydrolysis is not required for binding of AdoCbl to IcmF•GTP. GMPPNP decreased by ~2-fold the affinity for AdoCbl to site 1 (K_{D1}=154 ± 71 nM) and slightly increased the affinity at site 2 (K_{D2}=1.3 ± 0.5 μM).

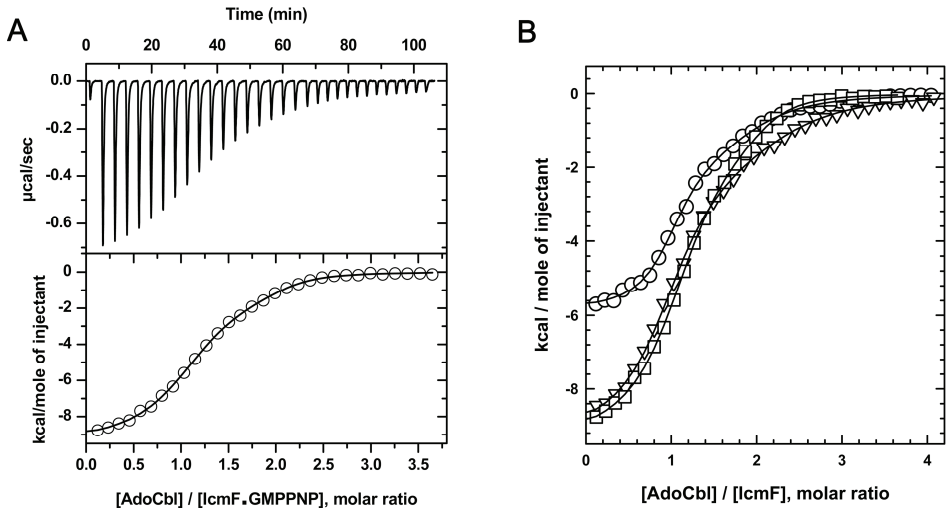


Figure 2.8: Binding isotherms for AdoCbl binding to IcmF. (**A**). Representative ITC data set for the binding of AdoCbl (250 µM stock solution) to 15.4 µM apo-IcmF in 50 mM sodium phosphate buffer, pH 7.5, 250 mM NaCl, 2 mM TCEP at 20 °C. (**B**). Titration curves for the binding of AdoCbl to apo-IcmF alone (○), in the presence of 1 mM GDP (▽) or 1 mM GMPPNP (□). Data were fitted to a two-site model and yielded the parameters reported in Table 2.2.

ligand	site	N	K_d, µM	ΔH, kcal/mol	TΔS, kcal/mol	ΔG, kcal/mol
none	**1**	0.9 ± 0.1	0.081 ± 0.014	-6.2 ± 0.2	+3.4 ± 0.3	-9.5 ± 0.1
	2	1.0 ± 0.2	1.98 ± 0.42	-2.0 ± 0.9	+5.6 ± 1.0	-7.7 ± 0.1
GDP	**1**	0.9 ± 0.1	0.132 ± 0.009	-8.7 ± 0.7	+0.6 ± 0.8	-9.2 ± 0.1
	2	1.2 ± 0.1	2.77 ± 1.27	-2.4 ± 0.3	+5.1 ± 0.4	-7.5 ± 0.2
GMPPNP	**1**	1.0 ± 0.1	0.154 ± 0.071	-9.6 ± 0.1	-0.4 ± 0.4	-9.2 ± 0.3
	2	1.0 ± 0.1	1.30 ± 0.51	-0.5 ± 0.7	+7.4 ± 0.5	-7.9 ± 0.2

Table 2.2: Thermodynamic parameters for the binding of AdoCbl to IcmF.

Changes in both the enthalpic and entropic terms contributed to this change. While cofactor binding to site 1 is enthalpically driven, it is almost entirely entropically driven at site 2.

2.4.5 IcmF is an active isobutyryl CoA mutase. Using the coupled assay, we further characterized the kinetic parameters for IcmF from *G. kaustophilus* (Figure 2.9 A, Table 2.3). The K_M for isobutyryl-CoA was determined to be 20 ± 1 μM and the k_{cat} 3.1 ± 0.1 s^{-1} in the absence of nucleotides. In comparison, a K_M for isobutyryl-CoA of 57 ± 13 μM has been reported for ICM from *S. cinnamonensis* (2). Thus the k_{cat}/K_M values for the stand-alone *S. cinnamonensis* ICM and *G. kaustophilus* IcmF are 6.8 ± 2.7 x10^5 $M^{-1}s^{-1}$ (2) and 1.5 ± 0.1 x10^5 $M^{-1}s^{-1}$ respectively. We note that the activity of the *G. kaustophilus* was measured at 37˚C in the coupled enzyme assay, which is significantly lower than the optimal growth temperature (60˚C) for the organism (*29*). Since purified recombinant IcmF was found to be unstable at higher temperatures, its activity at 60˚C could not be measured. Based on a coefficient of 2 for every 10˚C rise in temperature, we estimate that the k_{cat} for this enzyme might be ~4 fold higher at 60˚C.

Surprisingly, the presence of GDP or GTP affected both the k_{cat} and K_M values (Figure 2.9 A, Table 2.3). Thus, the presence of nucleotides decreased k_{cat} ~1.6-1.7-fold while increasing the K_M ~2-fold. Consequently, the catalytic efficiency k_{cat}/K_M, of IcmF decreased 3.5- and 4-fold respectively in the presence of GDP and GTP.

2.4.6 Absorption spectroscopy of IcmF under steady-state turnover conditions. As IcmF like ICM, is expected to deploy radical chemistry with AdoCbl (Figure 2.1), we analyzed whether the presence of nucleotides affected the cob(II)alamin levels under

steady-state turnover conditions (Figure 2.9 B). In the presence of isobutyryl-CoA, the spectrum of holo-IcmF was an ~1:2 mixture of cob(II)alamin:AdoCbl.

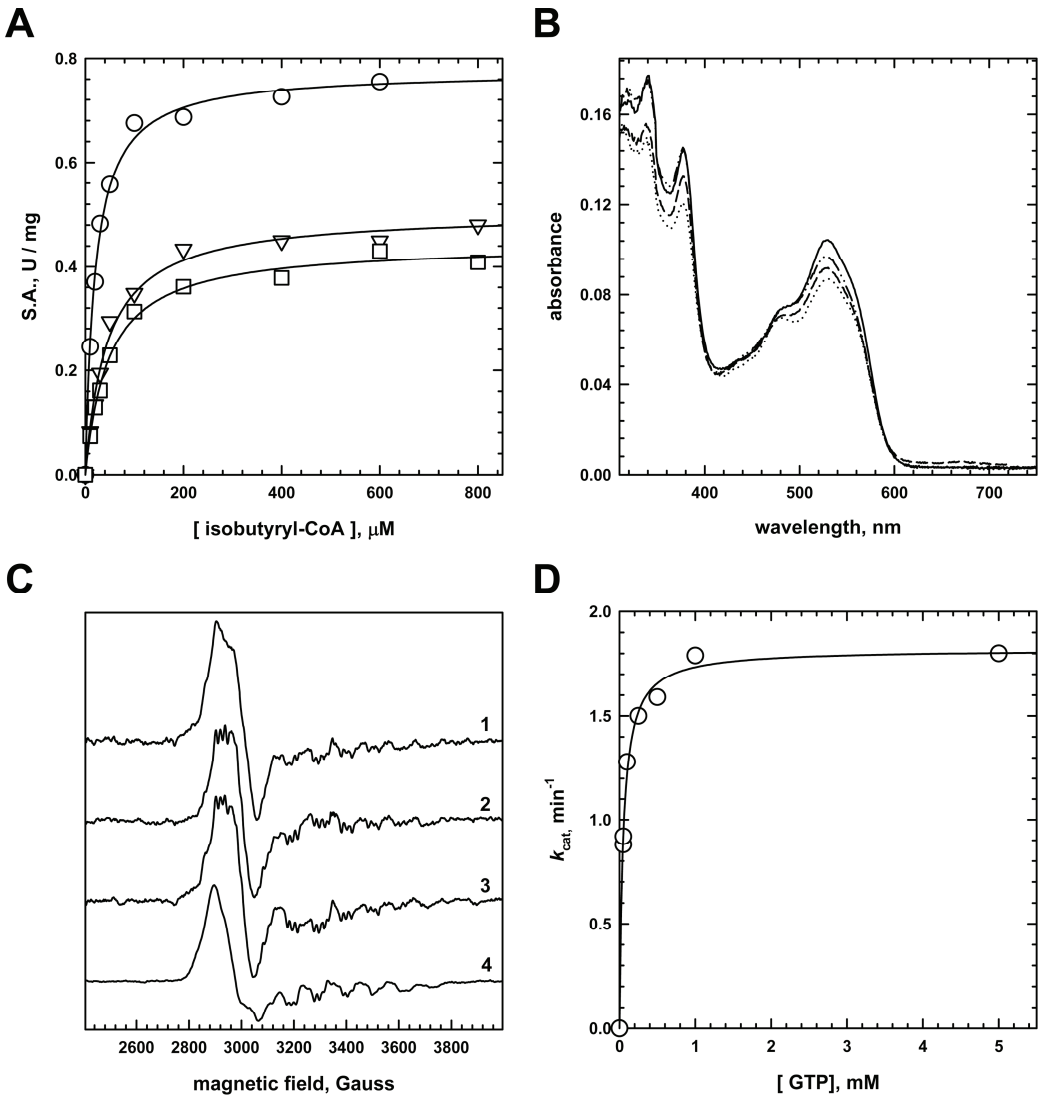


Figure 2.9: Kinetic and spectroscopic characterization of IcmF. (A). Michaelis-Menten analysis of the IcmF reaction as determined in the coupled enzyme assay. The specific activity of IcmF, analyzed alone (○) or in the presence of GDP (▽) or GTP (□), yielded the kinetic parameters reported in Table 2. **(B)**. UV-visible spectra of holo-IcmF under steady-state turnover conditions. Holo-IcmF (10.3 □M, solid line) was incubated at 20°C for 2 min with 3.4 mM isobutyryl-CoA (dotted line) in the presence of 2 mM GDP (dashed line) or GMPPNP (dashed-dotted line). **(C)**. EPR spectra of IcmF (28 □M) reconstituted with 20 □M cob(II)alamin (*spectrum* 1) in the presence of 10 mM 5′-deoxyadenosine (*spectrum* 2) and 8 mM isobutyryl-CoA (*spectrum* 3). An EPR spectrum of free "base-on" cob(II)alamin (*spectrum* 4) is shown for comparison. **(D)**. Michaelis-Menten analysis of the GTPase activity of IcmF determined as described under Experimental Procedures (section 2.3.3).

In the presence of nucleotides, accumulation of cob(II)alamin was diminished to ~1:4 (GDP) and ~1:9 (GMPPNP) (Figure 2.9 B). These results indicate that the nucleotides influence the steady-state distribution of intermediates, which might be related to their effects on k_{cat}.

Nucleotide	None	GDP	GTP
$K_{M(iBu\text{-}CoA)}$, μM	20.1 ± 1.3	45.3 ± 4.0	50.8 ± 2.6
k_{cat}, s^{-1}	3.10 ± 0.05	1.96 ± 0.04	1.88 ± 0.08
k_{cat}/K_M, $M^{-1}s^{-1}$	$(1.5 \pm 0.1)\times10^5$	$(4.3 \pm 0.5)\times10^4$	$(3.7 \pm 0.4)\times10^4$

Table 2.3: Kinetic parameters for IcmF. All experiments were performed in 50 mM NaPi, pH 7.5, 250 mM NaCl at 37°C as described under Experimental Procedures. The data represent the mean ± S.D. of three independent experiments.

2.4.7 EPR spectroscopy. The existence of a biradical intermediate has been demonstrated by EPR spectroscopy for MCM from *P. shermanii* with cob(II)alamin coupled to the product radical (*30*). However, an EPR spectrum was not observed when 40 μM holo-IcmF was mixed with 7 mM isobutyryl-CoA and frozen rapidly. Since the cob(II)alamin intermediate is observed by UV-visible spectroscopy (Figure 2.9 C), the lack of a paramagnetic signal suggests strong coupling between it and the organic radical species in the IcmF active site. This has also been observed with MCM from *M. extorquens* (Dominique Padovani and Ruma Banerjee, unpublished results).

The EPR spectrum of cob(II)alamin bound to IcmF was recorded (Figure 2.9 B). Binding of cob(II)alamin by IcmF yields an EPR spectrum that is diagnostic for the presence of an axial nitrogen ligand. Hyperfine coupling between the unpaired electron and the S=7/8

cobalt nucleus results in an eight line spectrum, which is further split into triplets due to superhyperfine coupling to the I=1 axial nitrogen ligand (*spectrum* 1). The spectrum of cob(II)alamin bound to IcmF differs from that of free cob(II)alamin (*spectrum* 4) particularly in the S-shaped absorption feature at $g \approx 2.3$ and probably results from immobilization of the cofactor in the active site. When IcmF was reconstituted with cob(II)alamin and 5′-deoxyadenosine in the presence or absence of isobutyryl-CoA (*spectra* 2 and 3), the spectra showed sharpening and resolution of additional hyperfine structure in the S-shaped feature was observed. These spectral differences suggest conformational changes that influence the electronic properties of the cob(II)alamin radical.

2.4.8 GTPase activity of IcmF. Since IcmF possesses a MeaI-like domain it was expected that this protein, like MeaB, can hydrolyze GTP. Hence, the kinetics of GTP hydrolysis catalyzed by apo-IcmF was characterized. A Michaelis-Menten analysis of the data yielded the following parameters: $K_{m(GTP)}=51 \pm 3$ μM and $k_{cat} = 1.8 \pm 0.05$ min^{-1} (Figure 2.9 D). In comparison, MeaB alone exhibits a lower intrinsic GTPase activity ($k_{cat}=0.039 \pm 0.003 min^{-1}$), which is increased ~100-fold in the presence of MCM (15).

2.4.9 The MeaI domain of IcmF is distinct from MeaB. The phylogenetic relationship between MeaB, the chaperone for MCM, and the MeaI domain of IcmF was evaluated. A dendrogram constructed from the analysis of MeaB and MeaI sequences found in the same organisms reveal that the two gene groups cluster separately (Figure 2.10). MeaB and MeaI are thus paralogs that have evolved to serve specific partner proteins, i.e. MCM and ICM.

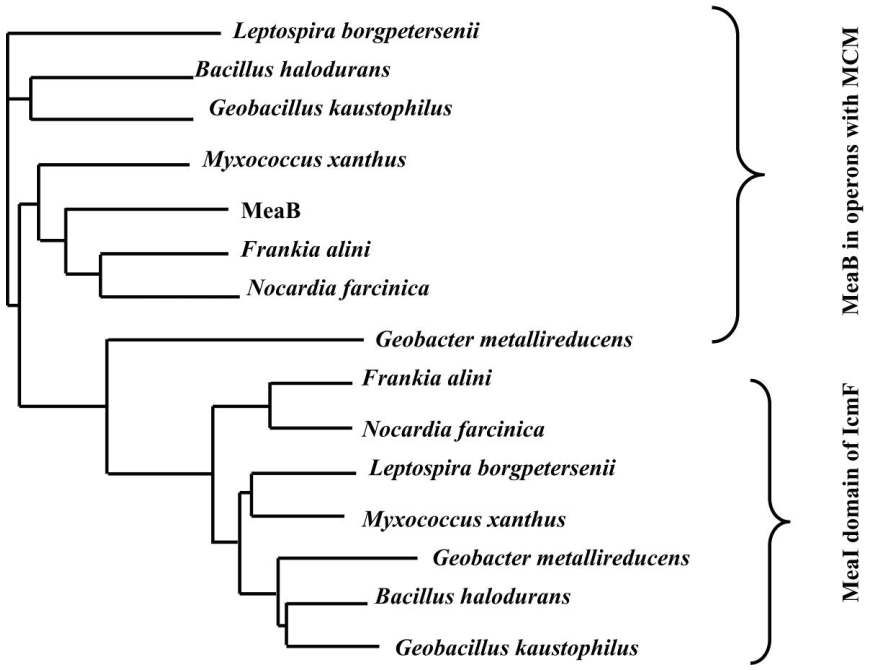


Figure 2.10: Phylogenetic tree of MeaBs that are located in operons with MCM and MeaIs that are fused to ICM (IcmF). MeaB-like sequences in the same operon with MCM: MeaB(AAL86727), *Bacillus halodurans* (NP_243820), *Frankia alini* (YP_715132), *Geobacillus kaustophilus* (YP_148222), *Geobacter metallireducens* (YP_385162), *Leptospira borgpetersenii* (YP_799393), *Myxococcus xanthus* (YP_630483), *Nocardia farcinica* (YP_119677). MeaI sequences which are part of IcmF: *Bacillus halodurans* (NP_244663), *Frankia alini* (YP_716016), *Geobacillus kaustophilus* (YP_149244), *Geobacter metallireducens* (YP_384678), *Leptospira borgpetersenii* (YP_801321), *Myxococcus xanthus* (YP_630482), *Nocardia farcinica* (YP_117245).

The observation of a MeaI domain in IcmF raises the obvious question of whether MeaI chaperones also exists for "stand-alone" ICMs. Indeed, as discussed below, analysis of genomic sequence reveals that two MeaB-like proteins are found in bacterial genomes, one associated with MCM (MeaB) and the other with ICM (MeaI). The diversification of the G-domain sequences within each subgroup strongly suggests that the MeaI-like domain of IcmF is evolutionary distinct from MeaB related to MCM.

2.4.10 Identification of stand-alone ICMs that do not belong to the genus *Streptomyces*. To investigate the relationships between the chaperones for ICM versus

IcmF, we analyzed other bacterial genomes for the presence of "stand-alone" ICMs and MeaIs. In our search we assumed that *(i)* the *icm*A and *icm*B genes are not necessarily located close to each other and that *(ii)* the amino acid substitutions corresponding to Phe80 and Gln198 in the *S. cinnamonensis* sequence are always found in the large subunit of ICM. A BLAST search using both subunits of the "stand-alone" ICM from *S. cinnamonensis* as the query sequence identified several "stand-alone" ICM sequences, primarily in thermophilic archaea but also in halophilic Archaea and in a limited number of bacteria (Figure 2.11 and Table 2.4). Furthermore, using the MeaI domain of IcmF as a query sequence revealed that genes encoding "stand-alone" MeaIs can be associated with either the large or the small subunit of ICM (Table 2.4). Interestingly, in several organisms, both ICM subunits are localized in the same operon or are in close proximity, e.g. in *Desulfibacterium hafniense*, *Archaeoglobus fulidus*, *Symbiobacterium thermophilum*. However, in other organisms, the two subunits are not close to each other in the genome e.g., in *Haloacrula marismortui*, *Halobacterium sp*., and in *Natronomonas pharaonis* (Table 2.4).

In some organisms, the gene order in an ICM-encoding operon is the following: MeaI, small subunit of ICM, large subunit of ICM whereas in others, the small subunit and MeaI co-localize in an operon while the large subunit is independently transcribed. A phylogenetic tree based on the alignment of "stand-alone" MeaIs, the MeaI domain of IcmF and the MeaBs associated with MCM reveals significant overall similarity between these proteins (Figure 2.12).

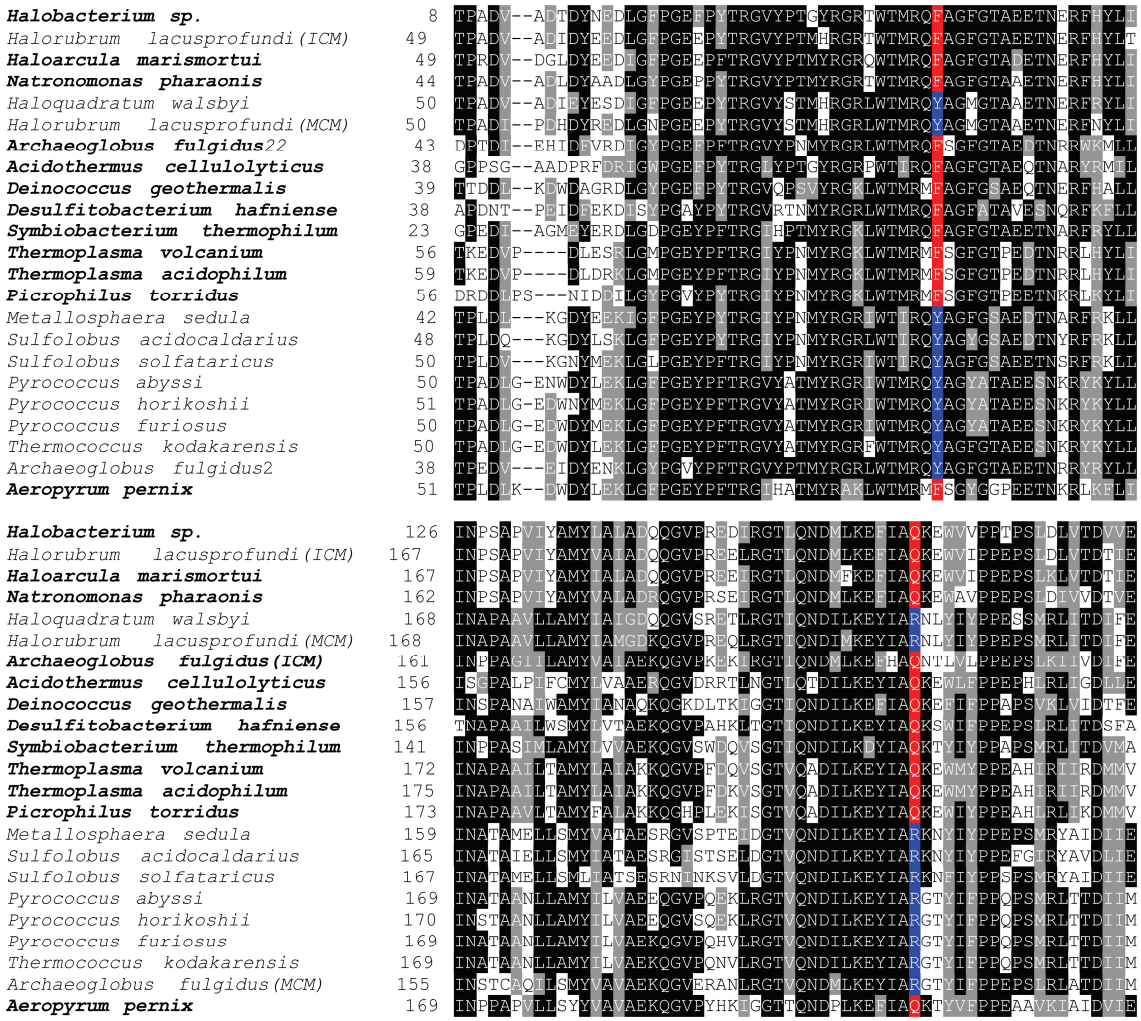


Figure 2.11: ICM and MCM sequences in Archaea, which are predicted to have some or all the key enzymes in the 3-hydroxypropionate/4-hydroxybutyrate cycle. "Stand-alone" ICMs identified in this study are indicated in bold (for the accession numbers see Table S2). Accession numbers for the rest of the organisms are as follows: *Halorubrum lacusprofundi* (ICM) (YP_002565846), *Halorubrum lacusprofundi* MCM)(YP_002565683), *Haloquadratum walsbyi* (YP_658133), *Metallosphaera sedula* (YP_001190737), *Sulfolobus acidocaldarius* (YP_255580), *Sulfolobus solfataricus* (NP_343779), *Pyrococcus abyssi* (NP_126525), *Pyrococcus horikoshii* (NP_143191), *Pyrococcus furiosus* (NP_579206), *Thermococcus kodakarensis* (YP_183562), *Archaeoglobus fulgidus* (MCM) (NP_071040).

Organism	IcmA	IcmB	icmA and icmB in the same operon	meaI in the same operon with icmA and/or icmB	MeaI	Description
Acidothermus cellulolyticus	YP_872433	YP_872137	no	yes	YP_872432	Thermophilic bacterium
Natronomonas pharaonis	YP_326811	YP_327005	no	yes	YP_327006	Haloalkaliphilic archaeon
Deinococcus geothermalis	YP_604333	YP_605005	no	no	YP_605077	Thermophilic bacterium
Desulfitobacterium hafniense	YP_517796	YP_517797	yes	yes	YP_517795	Anaerobic dehalogenating bacterium
Archaeoglobus fulgidus	NP_071268	NP_071269	yes	yes	NP_070118	Sulphate-reducing archaeon
Symbiobacterium thermophilum	YP_076295	YP_076294	yes	yes	YP_076293	Thermophilic bacterium
Thermoplasma volcanium	NP_111291	NP_111292	yes	yes	NP_111293	Thermophile archaeon
Thermoplasma acidophilum	NP_393940	NP_393941	yes	yes	NP_393942	Thermophile archaeon
Picrophilus torridus	YP_023047	YP_023048	yes	yes	YP_023049	Thermoacidophilic archaeon
Haloarcula marismortui	YP_135919	YP_135613	no	yes	YP_135612	Halophilic archaeon
Aeropyrum pernix	NP_148096	NP_148095	yes	yes	NP_148094	Hyperthermophilic archaeon
Halobacterium sp.	NP_279671	NP_279686	no	yes	NP_279687	Extremely halophilic archaeon

Table 2.4: List of identified “stand-alone” ICMs and MeaIs.

However, a careful examination of all available MeaI sequences reveals that this group is evolutionarily distinct from MeaB since they form two separate clusters in the dendrogram. In contrast, MeaIs associated with stand-alone ICMs are closely related to the corresponding domain in IcmF. Hence, the MeaI and MeaB are paralogs that probably evolved from a common ancestor and have diverged to support specific B_{12}-dependent isomerases.

2.4.11 Implications of the presence of IcmF. Gene fusion events occur during evolution resulting in the physical coupling of functionally coupled proteins. It is speculated that gene fusions that facilitate functional interactions between and/or co-regulation of proteins, might be maintained by selective pressure and are more common than gene fissions (*31, 32*). In the present study, we have characterized IcmF, a protein that likely arose by fusion of three genes encoding the large and small subunits of ICM and the chaperone, MeaI. Bioinformatics analysis has allowed identification of >70 IcmFs in bacterial and archaeal genomes. However, as noted earlier, all these proteins are incorrectly assigned as representing fusions between MCM and MeaB. There are several reasons that could have led to misannotation of IcmF in the databases. First, ICM activity was believed to be restricted to the genus *Streptomyces*, whereas the MCM-catalyzed reaction is important in secondary metabolism and widely distributed in bacteria. Second, the two signature active site substitutions in the *S. cinnamonensis* ICM was missed in the IcmF sequence (16).

The importance of these residues in substrate selectivity was previously demonstrated by mutagenesis studies in which the MCM double mutant, Tyr89Phe/Arg207Gln, was designed to mimic the active site residues in ICM (*33*).

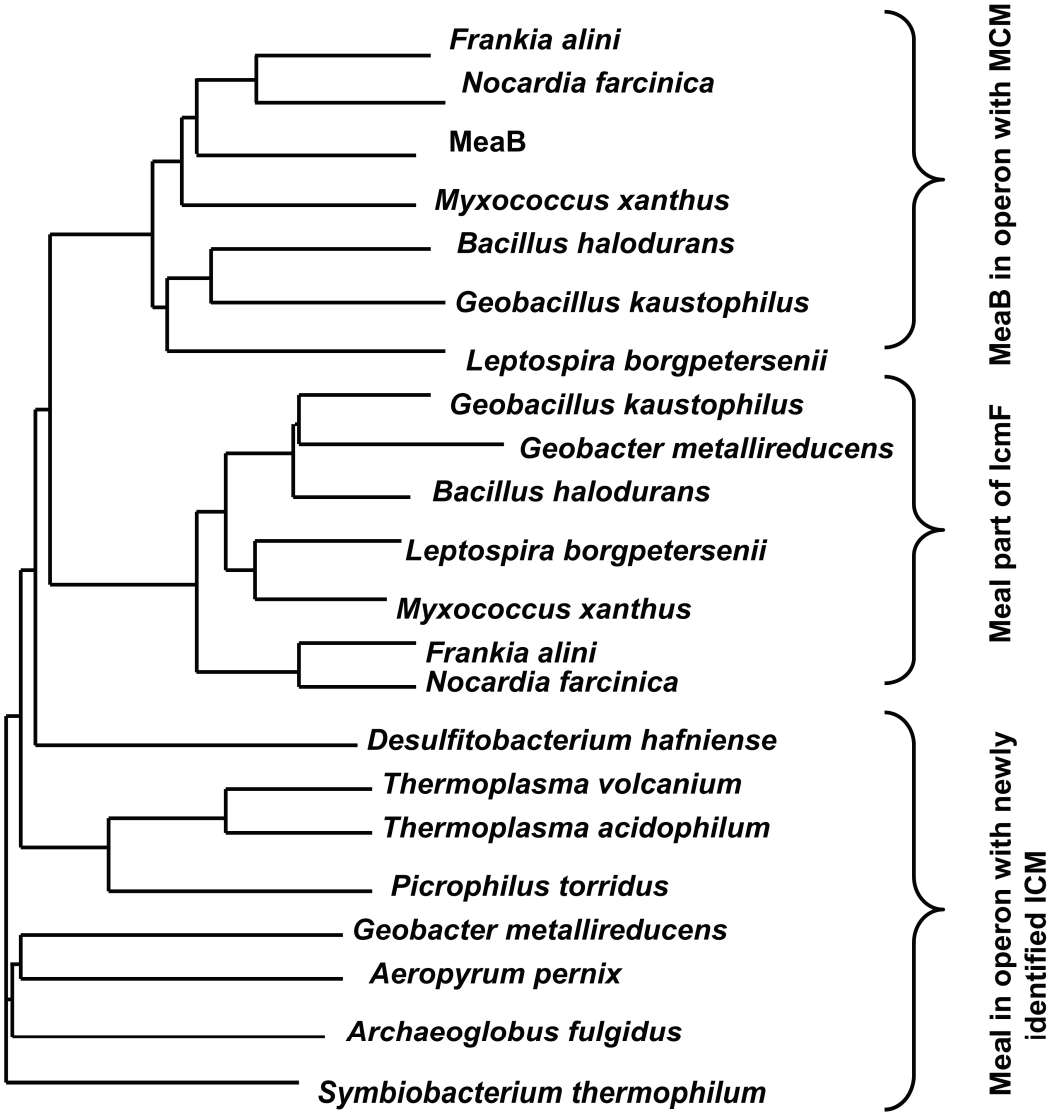


Figure 2.12: Dendrogram showing the phylogenetic relationships between MeaB-like proteins. MeaB-like sequences in the same operon with MCM: MeaB (AAL86727), *Bacillus halodurans* (NP_243820), *Frankia alini* (YP_715132), *Geobacillus kaustophilus* (YP_148222), *Geobacter metallireducens*(YP_385162), *Leptospira borgpetersenii* (YP_799393), *Myxococcus xanthus* (YP_630483), *Nocardia farcinica* (YP_119677). MeaB-like sequences part of IcmF: *Bacillus halodurans* (NP_244663), *Frankia alini* (YP_716016), *Geobacillus kaustophilus* (YP_149244), *Geobacter metallireducens* (YP_384678), *Leptospira borgpetersenii* (YP_801321), *Myxococcus xanthus* (YP_630482), *Nocardia farcinica* (YP_117245). MeaB-like proteins in the same operon with newly identified ICMs: *Desulfitobacterium hafniense* (YP_517795), *Archaeoglobus fulgidus* (NP_070118), *Symbiobacterium thermophilum* (YP_076293), *Thermoplasma volcanium* (NP_111293), *Thermoplasma acidophilum* (NP_393942), *Picrophilus torridus*(YP_023049), *Aeropyrum pernix* (NP_148094). For IcmFs accession numbers see Table 2.1

In contrast to wild-type MCM, the double mutant bound the ICM substrates, n-butyryl-CoA or isobutyryl-CoA, but instead of catalyzing an isomerization, led to inactivation via an internal electron transfer.

Third, the role for a MeaB like chaperone protein has only been described so far for MCM and could have contributed to the erroneous assignment. IcmF (previously described as McmC) was reported to have very low MCM activity (10-12x10^{-3} μmol $min^{-1}mg^{-1}$) in crude extracts of *B. xenovorans*, an organism that lacks the gene encoding a "stand-alone" MCM (16). However, when we cloned and purified *B. xenovorans* IcmF, we found it to be devoid of MCM activity, which in fact, prompted a closer inspection of the protein sequence and therefore its predicted activity.

Our findings expand our view of the distribution of B_{12} dependent mutases. We find that ICM activity is much more widely distributed in Nature than previously suspected and raises questions about the metabolic pathways in which this activity is involved.

In certain bacteria (e.g. *Butyrivibrio fibrisolvens*, *Streptomyces collinus*), acetyl-CoA is converted to butyryl-CoA via four reactions, involving acetyl-CoA acetyltransferase (thiolase), 3-hydroxybutyryl-CoA dehydrogenase, 3-hydroxybutyryl-CoA dehydratase (crotonase) and butyryl-CoA dehydrogenase (*34-36*).

Subjecting the IcmF sequences to operon analysis reveals that in eight bacterial genomes (*Lysinibacillus sphaericus*, *Bacillus* sp., *Bacillus halodurans*, *Bacillus coagulans*, *Bacillus selenitireducens*, *Bdellovibrio bacteriovorus*, *Geobacillus* sp., *Anoxybacillus flavithermus*), IcmF is located in the same operon with enzymes involved in formation of butyryl-CoA from acetyl-CoA. Based on this analysis, we posit that IcmF is involved in butyryl-CoA, rather than methylmalonyl-CoA metabolism in these bacteria.

Myxalamids are inhibitors of the eukaryotic electron transfer chain that are produced by the myxobacteria, *Myxococcus xanthus* and *Stigmatella aurantiaca* (*37*). In studies on *M. xanthus* and *S. aurantiaca* mutants in which the branched-chain ketoacid dehydrogenase was disrupted (*bkd* mutants), it was shown that isobutyryl-CoA is incorporated into the final product (*38*). These results were unexpected since the *bkd* mutants are unable to form isobutyryl-CoA starter units from valine. The authors suggested that fatty acid degradation by α- and β-oxidation of iso-odd fatty acid could be responsible for isobutyryl-CoA synthesis (*38*). We speculate that the ICM activity of IcmF found in both these bacteria might play a role in this process instead.

Another interesting implication of our study stems from the identification of "stand-alone" ICMs in a number of Archaea and bacteria (Figure 2.11 and Table 2.4). Recently, Fuchs and colleagues have reported the discovery of a novel CO_2-fixation pathway in several Archaea (*17, 39*). They have characterized the sixteen enzymes in the 3-hydroxypropionate/4-hydroxybutyrate pathway in *Metallosphaera sedula*. In this pathway, two CO_2 molecules are fixed with acetyl-CoA and reductively converted to succinyl-CoA. An intermediate step in this pathway is the conversion of methylmalonyl-CoA to succinyl-CoA, which is catalyzed by MCM. The majority of MCMs in bacteria are heterodimers, in which one of the subunit binds the substrate and the cofactor. Although *M. sedula* clearly encodes MCM in its genome, some of the putative mutases in other organisms that were identified as MCMs (Table S1 in (17)) are predicted to be "stand-alone" ICMs based on the Tyr→Phe/Arg→Gln substitutions in their active sites. Thus, in *Haloarcula marismortui*, *Halobacterium sp.* and *Natronomonas pharaonis* "stand-alone" ICMs rather than MCMs are predicted to exist, raising questions about the

presence of an intact 3-hydroxypropionate/4-hydroxybutyrate pathway in these organisms (Figure 2.11). In contrast, in *Archaeglobus fulgidus* and *Halorubrum lacusprofundi* both copies of the *mcm* and *icm* genes are present. In *M. sedula* only one copy of MCM is present (Figure 2.11). Fuchs and colleagues noted that in some organisms the enzymes from the first half of the cycle are missing and proposed that in this situation reversal of the second half of the pathway might be important for acetyl-CoA assimilation into succinyl-CoA (17). Interestingly, the first three reactions of this reverse sequence (acetoacetyl-CoA β-ketothiolase, 3-hydroxybutyryl-CoA dehydrogenase, and crotonyl-CoA hydratase) are identical to those in the acetyl-CoA assimilation pathway described for *Streptomyces collinus,* which converts acetyl-CoA to crotonyl-CoA (*34*). The latter, via the action of crotonyl-CoA reductase is converted to butyryl-CoA, which is isomerized to isobutyryl-CoA by the action of a "stand-alone" ICM. Isobutyryl-CoA can be converted to succinyl-CoA. Thus in organisms lacking enzymes in the first half of the 3-hydroxypropionate/4-hydroxybutyrate pathway, ICM may afford an alternative route for assimilation of acetyl-CoA.

It is interesting how MCM-like enzymes have evolved distinct substrate specificities by virtue of very limited changes in their active site residues. Muller and colleagues described a B_{12}-dependent enzyme which is involved in the pathway of degradation of fuel oxygenates (*40, 41*). This enzyme in *Methylibium petroleiphilum* PM1 was shown to convert 2-hydroxyisobutyryl-CoA into 3-hydroxybutyryl-CoA. The remarkable feature of this enzyme is that it resembles ICM and has two subunits, IcmA and IcmB. However in the active site of IcmA, Phe is substituted by Ile, whereas Gln is conserved (Figure 5 in (*40*)) (See also section 3.2). It is interesting that *M. petroleiphilum* also has a copy of the

*icm*F and *mcm* genes (based on amino acid substitutions in the active site sequences). Another example of subtle alterations in substrate specificity is seen in ethylmalonyl-CoA mutase (ECM) from *Rhodobacter sphaeroides* (*42*). This enzyme interconverts ethylmalonyl-CoA and methylsuccinyl-CoA. Like MCM, ECM is predicted to have Tyr and Arg residues in the active site. However in order to utilize the larger ethylmalonyl-CoA/methylsuccinyl-CoA substrates, it is speculated that a conserved His and Asn in MCM are substituted by Gly255 and Pro296 respectively in the *R. sphaeroides* ECM (Supplementary Figure S4 in (*42*)) (See also section 3.2).

The identification of ICM- and IcmF-encoding genes in a number of bacteria and Archaea should fuel studies aimed at identifying the metabolic contributions of the ICM activity in these organisms.

2.5 References

(1) Zerbe-Burkhardt, K., Ratnatilleke, A., Philippon, N., Birch, A., Leiser, A., Vrijbloed, J. W., Hess, D., Hunziker, P., and Robinson, J. A. (1998) Cloning, sequencing, expression, and insertional inactivation of the gene for the large subunit of the coenzyme B12-dependent isobutyryl-CoA mutase from Streptomyces cinnamonensis. *J Biol Chem 273*, 6508-17.

(2) Ratnatilleke, A., Vrijbloed, J. W., and Robinson, J. A. (1999) Cloning and sequencing of the coenzyme B(12)-binding domain of isobutyryl-CoA mutase from Streptomyces cinnamonensis, reconstitution of mutase activity, and characterization of the recombinant enzyme produced in Escherichia coli. *J Biol Chem 274*, 31679-85.

(3) Vrijbloed, J. W., Zerbe-Burkhardt, K., Ratnatilleke, A., Grubelnik-Leiser, A., and Robinson, J. A. (1999) Insertional inactivation of methylmalonyl coenzyme A (CoA) mutase and isobutyryl-CoA mutase genes in Streptomyces cinnamonensis: influence on polyketide antibiotic biosynthesis. *J Bacteriol 181*, 5600-5.

(4) Banerjee, R., and Chowdhury, S. (1999) *Methylmalonyl-CoA mutase*, John Wiley and Sons, New York.

(5) Drennan, C. L., Huang, S., Drummond, J. T., Matthews, R. G., and Lidwig, M. L. (1994) How a protein binds B12: A 3.0 A X-ray structure of B12-binding domains of methionine synthase. *Science 266*, 1669-74.

(6) Holloway, D. E., and Marsh, E. N. (1994) Adenosylcobalamin-dependent glutamate mutase from Clostridium tetanomorphum. Overexpression in Escherichia coli, purification, and characterization of the recombinant enzyme. *J Biol Chem 269*, 20425-30.

(7) Banerjee, R. (2003) Radical carbon skeleton rearrangements: catalysis by coenzyme B_{12}-dependent mutases. *Chem Rev 103*, 2083-94.

(8) Dobson, C. M., Wai, T., Leclerc, D., Wilson, A., Wu, X., Dore, C., Hudson, T., Rosenblatt, D. S., and Gravel, R. A. (2002) Identification of the gene responsible for the cblA complementation group of vitamin B12-responsive methylmalonic acidemia based on analysis of prokaryotic gene arrangements. *Proc Natl Acad Sci U S A 99*, 15554-9.

(9) Hubbard, P. A., Padovani, D., Labunska, T., Mahlstedt, S. A., Banerjee, R., and Drennan, C. L. (2007) Crystal structure and mutagenesis of the metallochaperone MeaB: insight into the causes of methylmalonic aciduria. *J Biol Chem 282*, 31308-16.

(10) Leipe, D. D., Wolf, Y. I., Koonin, E. V., and Aravind, L. (2002) Classification and evolution of P-loop GTPases and related ATPases. *J Mol Biol 317*, 41-72.

(11) Gasper, R., Scrima, A., and Wittinghofer, A. (2006) Structural insights into HypB, a GTP-binding protein that regulates metal binding. *J Biol Chem 281*, 27492-502.

(12) Zambelli, B., Musiani, F., Savini, M., Tucker, P., and Ciurli, S. (2007) Biochemical studies on Mycobacterium tuberculosis UreG and comparative modeling reveal structural and functional conservation among the bacterial UreG family. *Biochemistry 46*, 3171-82.

(13) Jeon, W. B., Cheng, J., and Ludden, P. W. (2001) Purification and characterization of membrane-associated CooC protein and its functional role in the insertion of nickel into carbon monoxide dehydrogenase from Rhodospirillum rubrum. *J Biol Chem 276*, 38602-9.
(14) Padovani, D., and Banerjee, R. (2006) Assembly and protection of the radical enzyme, methylmalonyl-CoA mutase, by its chaperone. *Biochemistry 45*, 9300-6.
(15) Padovani, D., Labunska, T., and Banerjee, R. (2006) Energetics of interaction between the G-protein chaperone, MeaB, and B12-dependent methylmalonyl-CoA mutase. *J Biol Chem 281*, 17838-44.
(16) Korotkova, N., and Lidstrom, M. E. (2004) MeaB is a component of the methylmalonyl-CoA mutase complex required for protection of the enzyme from inactivation. *J Biol Chem 279*, 13652-8.
(17) Berg, I. A., Kockelkorn, D., Buckel, W., and Fuchs, G. (2007) A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. *Science 318*, 1782-6.
(18) Takami, H., Takaki, Y., Chee, G. J., Nishi, S., Shimamura, S., Suzuki, H., Matsui, S., and Uchiyama, I. (2004) Thermoadaptation trait revealed by the genome sequence of thermophilic Geobacillus kaustophilus. *Nucleic Acids Res 32*, 6292-303.
(19) Ishikawa, J., Yamashita, A., Mikami, Y., Hoshino, Y., Kurita, H., Hotta, K., Shiba, T., and Hattori, M. (2004) The complete genomic sequence of Nocardia farcinica IFM 10152. *Proc Natl Acad Sci U S A 101*, 14925-30.
(20) Becker, D. F., Fuchs, J. A., Banfield, D. K., Funk, W. D., MacGillivray, R. T., and Stankovich, M. T. (1993) Characterization of wild-type and an active-site mutant in Escherichia coli of short-chain acyl-CoA dehydrogenase from Megasphaera elsdenii. *Biochemistry 32*, 10736-42.
(21) Taoka, S., Padmakumar, R., Lai, M. T., Liu, H. W., and Banerjee, R. (1994) Inhibition of the human methylmalonyl-CoA mutase by various CoA-esters. *J Biol Chem 269*, 31630-4.
(22) Birch, A., Leiser, A., and Robinson, J. A. (1993) Cloning, sequencing, and expression of the gene encoding methylmalonyl-coenzyme A mutase from Streptomyces cinnamonensis. *J Bacteriol 175*, 3511-9.
(23) Lehman, T. C., Hale, D. E., Bhala, A., and Thorpe, C. (1990) An acyl-coenzyme A dehydrogenase assay utilizing the ferricenium ion. *Anal Biochem 186*, 280-4.
(24) Snel, B., Lehmann, G., Bork, P., and Huynen, M. A. (2000) STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res 28*, 3442-4.
(25) Alm, E. J., Huang, K. H., Price, M. N., Koche, R. P., Keller, K., Dubchak, I. L., and Arkin, A. P. (2005) The MicrobesOnline Web site for comparative genomics. *Genome Res 15*, 1015-22.
(26) Cracan, V., Padovani, D., and Banerjee, R. (2010) IcmF is a fusion between the radical B12 enzyme isobutyryl-CoA mutase and its G-protein chaperone. *J Biol Chem 285*, 655-66.
(27) Pandit, S. B., and Srinivasan, N. (2003) Survey for g-proteins in the prokaryotic genomes: prediction of functional roles based on classification. *Proteins 52*, 585-97.

(28) Banerjee, R. (1999) *Chemistry and Biochemistry of B_{12}*, John Wiley and Sons.
(29) Montoro-Garcia, S., Martinez-Martinez, I., Navarro-Fernandez, J., Takami, H., Garcia-Carmona, F., and Sanchez-Ferrer, A. (2009) Characterization of a novel thermostable carboxylesterase from Geobacillus kaustophilus HTA426 shows the existence of a new carboxylesterase family. *J Bacteriol 191*, 3076-85.
(30) Mansoorabadi, S. O., Padmakumar, R., Fazliddinova, N., Vlasie, M., Banerjee, R., and Reed, G. H. (2005) Characterization of a succinyl-CoA radical-cob(II)alamin spin triplet intermediate in the reaction catalyzed by adenosylcobalamin-dependent methylmalonyl-CoA mutase. *Biochemistry 44*, 3153-8.
(31) Enright, A. J., Iliopoulos, I., Kyrpides, N. C., and Ouzounis, C. A. (1999) Protein interaction maps for complete genomes based on gene fusion events. *Nature 402*, 86-90.
(32) Snel, B., Bork, P., and Huynen, M. (2000) Genome evolution. Gene fusion versus gene fission. *Trends Genet 16*, 9-11.
(33) Vlasie, M. D., and Banerjee, R. (2004) When a spectator turns killer: suicidal electron transfer from cobalamin in methylmalonyl-CoA mutase. *Biochemistry 43*, 8410-7.
(34) Akopiants, K., Florova, G., Li, C., and Reynolds, K. A. (2006) Multiple pathways for acetate assimilation in Streptomyces cinnamonensis. *J Ind Microbiol Biotechnol 33*, 141-50.
(35) Asanuma, N., Ishiwata, M., Yoshii, T., Kikuchi, M., Nishina, Y., and Hino, T. (2005) Characterization and transcription of the genes involved in butyrate production in Butyrivibrio fibrisolvens type I and II strains. *Curr Microbiol 51*, 91-4.
(36) Miller, T. L., and Jenesel, S. E. (1979) Enzymology of butyrate formation by Butyrivibrio fibrisolvens. *J Bacteriol 138*, 99-104.
(37) Gerth, K., Jansen, R., Reifenstahl, G., Hofle, G., Irschik, H., Kunze, B., Reichenbach, H., and Thierbach, G. (1983) The myxalamids, new antibiotics from Myxococcus xanthus (Myxobacterales). I. Production, physico-chemical and biological properties, and mechanism of action. *J Antibiot (Tokyo) 36*, 1150-6.
(38) Bode, H. B., Meiser, P., Klefisch, T., Cortina, N. S., Krug, D., Gohring, A., Schwar, G., Mahmud, T., Elnakady, Y. A., and Muller, R. (2007) Mutasynthesis-derived myxalamids and origin of the isobutyryl-CoA starter unit of myxalamid B. *Chembiochem 8*, 2139-44.
(39) Alber, B. E., Kung, J. W., and Fuchs, G. (2008) 3-Hydroxypropionyl-coenzyme A synthetase from Metallosphaera sedula, an enzyme involved in autotrophic CO2 fixation. *J Bacteriol 190*, 1383-9.
(40) Rohwerder, T., Breuer, U., Benndorf, D., Lechner, U., and Muller, R. H. (2006) The alkyl tert-butyl ether intermediate 2-hydroxyisobutyrate is degraded via a novel cobalamin-dependent mutase pathway. *Appl Environ Microbiol 72*, 4128-35.
(41) Muller, R. H., Rohwerder, T., and Harms, H. (2008) Degradation of fuel oxygenates and their main intermediates by Aquincola tertiaricarbonis L108. *Microbiology 154*, 1414-21.

(42) Erb, T. J., Retey, J., Fuchs, G., and Alber, B. E. (2008) Ethylmalonyl-CoA mutase from Rhodobacter sphaeroides defines a new subclade of coenzyme B12-dependent acyl-CoA mutases. *J Biol Chem 283*, 32283-93.

CHAPTER 3

A Novel IcmF Activity Interconverts Isovaleryl-CoA and Pivalyl-CoA[4,5]

3.1 Abstract

Adenosylcobalamin (AdoCbl)-dependent isomerases catalyze carbon skeleton rearrangements using radical chemistry. We have recently characterized a fusion protein IcmF, which is comprised of the two subunits of the AdoCbl-dependent isobutyryl-CoA mutase flanking a G-protein chaperone. IcmF catalyzes the interconverion of isobutyryl-CoA and n-butyryl-CoA while the GTPase activity is associated with its G-protein domain. In this study, we report a novel activity associated with IcmF, i.e. the interconversion of isovaleryl-CoA and pivalyl-CoA. Kinetic characterization of IcmF yielded the following values: K_M for isovaleryl-CoA = 62 ± 8 μM and V_{max} = 0.021 ± 0.004 μmol/min/mg at 37 ˚C. Biochemical experiments show that in an IcmF in which the base specificity loop motif, NKxD is modified to NKxE, catalyzes the hydrolysis of both GTP and ATP. IcmF is susceptible to rapid inactivation during turnover and GTP conferred modest protection during utilization of isovaleryl-CoA as substrate. Interestingly, there was no protection from inactivation when either isobutyryl-CoA or n-butyryl-CoA was used as substrate. Detailed kinetic analysis indicated that inactivation is

[4] The content of this chapter has been published in J Biol Chem. 2011 Dec 13: **Cracan V and Banerjee R**. "A novel coenzyme B_{12} –dependent interconversion of isovaleryl-CoA and pivalyl-CoA".

[5] We gratefully acknowledge Dr. Thore Rohwerder (Helmholtz Center for Environmental Research-UFZ) for helpful discussions on the distribution of acyl-CoA mutases in Nature and their potential roles in microbial metabolism. A pET28a vector expressing IcmF from *Cupriavidus metallidurans* was a generous gift from the laboratory of Dr. Catherine Drennan (MIT). This work was supported by a grant from the National Institutes of Health (DK45776).

associated with loss of the 5´-deoxyadenosine moiety from the active site, precluding re-formation of AdoCbl at the end of the turnover cycle. Under aerobic conditions, oxidation of the cob(II)alamin radical in the inactive enzyme results in accumulation of aquacobalamin. Since pivalic acid found in sludge can be used as a carbon source by some bacteria and isovaleryl-CoA is an intermediate in leucine catabolism, our discovery of a new isomerase activity associated with IcmF, expands its metabolic potential.

3.2. Introduction

Methylmalonyl-CoA mutase (MCM) and Isobutyryl-CoA mutase (ICM) are two closely related 5´-deoxyadenosylcobalamin (AdoCbl)-dependent isomerases, which catalyze 1,2-rearrangements of methylmalonyl-CoA to succinyl-CoA and isobutyryl-CoA to n-butyryl-CoA, respectively (Figure 3.1) (1,2). MCMs are widely distributed in Nature, ranging from bacteria and Archaea to animals, including humans. ICMs were initially believed to be restricted to the genus *Streptomyces,* which belongs to the *Actinobacteria* phylum *(1,2)*. With our discovery of IcmF, the fusion protein between ICM and its G-protein chaperone, the known distribution of ICM has expanded to include four more phyla: *Proteobacteria*, *Bacteroidetes*, *Firmicutes* and *Spirochaetes* (3). The only known function of ICM in the genus *Streptomyces* is its participation in polyketide biosynthesis (4). In contrast the relatively wide distribution of IcmF in diverse organisms points to a broader range of roles in bacterial metabolism, which remain to be elucidated.

A family of enzymes that are similar in their primary sequence to MCM catalyze AdoCbl-dependent carbon-skeleton rearrangements and include, in addition to ICM, 2-hydroxyisobutyryl-CoA mutase (HCM) (*5*) and ethylmalonyl-CoA mutase (ECM) (*6*) (Figure 3.1). The three-dimensional structure of these proteins is expected to resemble the

overall structure of MCM (7,8). The B_{12}-binding domain of MCM exhibits a typical Rossmann-like fold and the AdoCbl cofactor is bound in a "base-off/His-on" conformation (9). The substrate-binding domain is a TIM barrel comprised of a core of eight (αβ)-repeats (*8*).

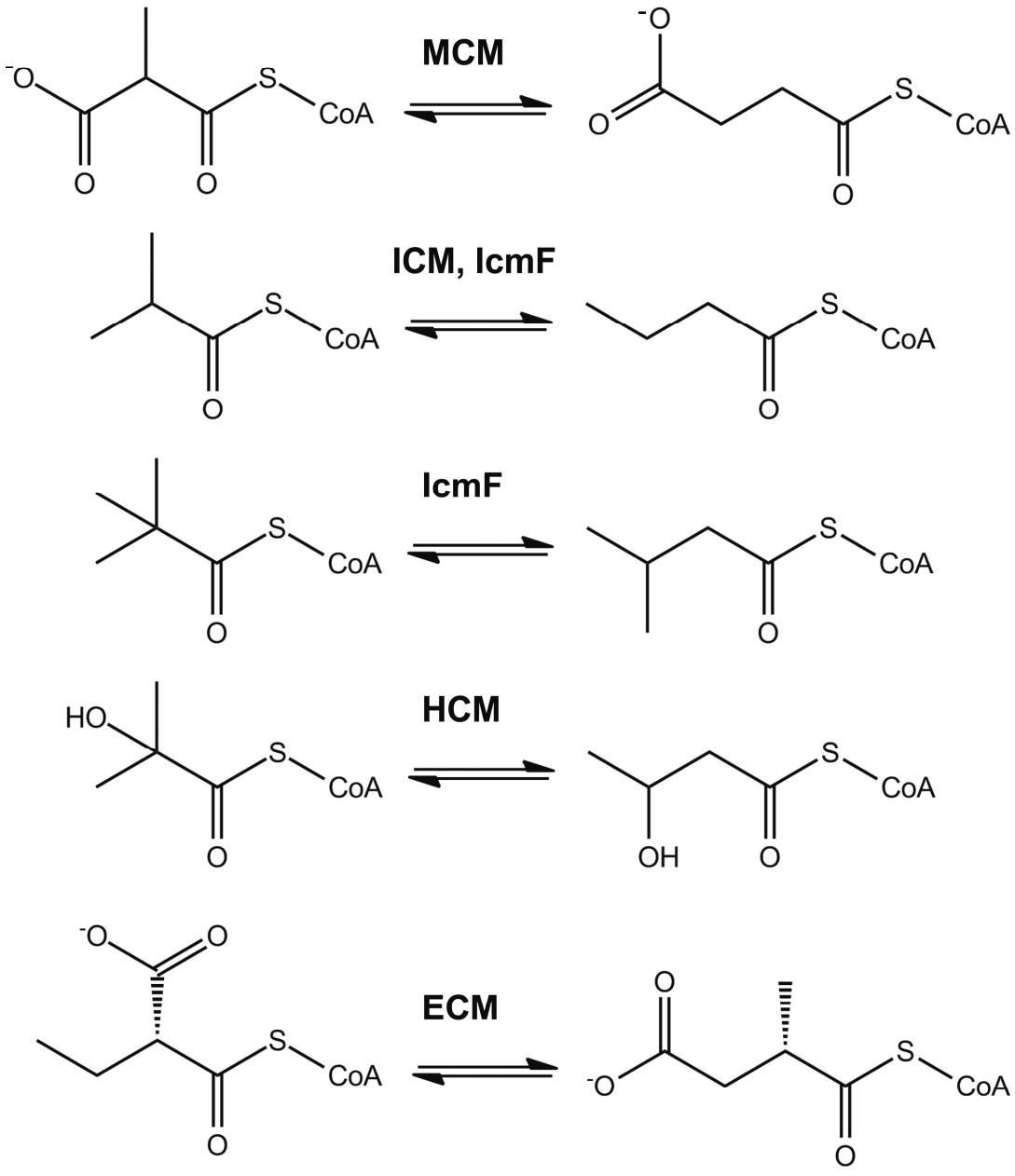


Figure 3.1: Reactions catalyzed by MCM, ICM, IcmF, HCM and ECM.

A limited number of key amino acid substitutions in the active site distinguish these closely related enzymes and accommodate chemical differences in the substrate (6,10). Previous studies from our laboratory had pointed out that the two key substitutions that accommodate the switch from a carboxylate to a methyl group in the substrates of MCM versus ICM are Tyr→Phe and Arg→Gln respectively (Figure 3.2 and Figure 3.3) (*3*). The Phe and Arg substitutions are conserved in all ICM/IcmFs (Figure 3.2 and Figure 3.3).

In contrast, in HCM, the corresponding amino acids that interact with substrate are Ile and Gln, respectively. Thus while the Gln residue is conserved in both ICM/IcmF and in HCM, the Phe is substituted by Ile to accommodate the bulkier 2-hydroxyisobutyryl-CoA substrate (Figure 3.2 and Figure 3.3).

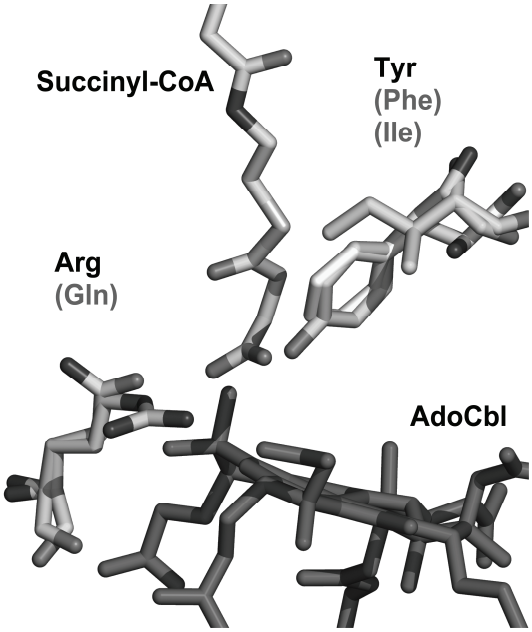


Figure 3.2: Comparison of active site residues in related AdoCbl-dependent mutases. The MCM structure from *P. shermanii* (4REQ) was used as a template to show that Tyr and Arg in MCM/ECM correspond to Phe and Gln in ICM/IcmF. In HCM, this pair of residues is Ile and Gln.

In contrast to ICM, ECM, which catalyzes the interconversion of ethylmalonyl-CoA and methylsuccinyl-CoA is very similar to MCM (*6*). It has been proposed that two key substitutions in the active site of ECM dictate specificity for the bulkier ethylmalonyl-CoA substrate: a His and an Asn residue in MCM are replaced by Gly and Pro, respectively (*6*) (Figure 3.3). Notably, the Tyr and Arg residues in the active site of MCM that interact with the carboxylate moiety of the substrate are also conserved in ECM (Figure 3.3). In B_{12}-dependent isomerases, AdoCbl serves as a radical reservoir and a common first step that initiates the isomerization reactions is homolytic cleavage of the cobalt-carbon bond leading to formation of the 5′-deoxyadenosyl radical and a paramagnetic cob(II)alamin species (1,11).

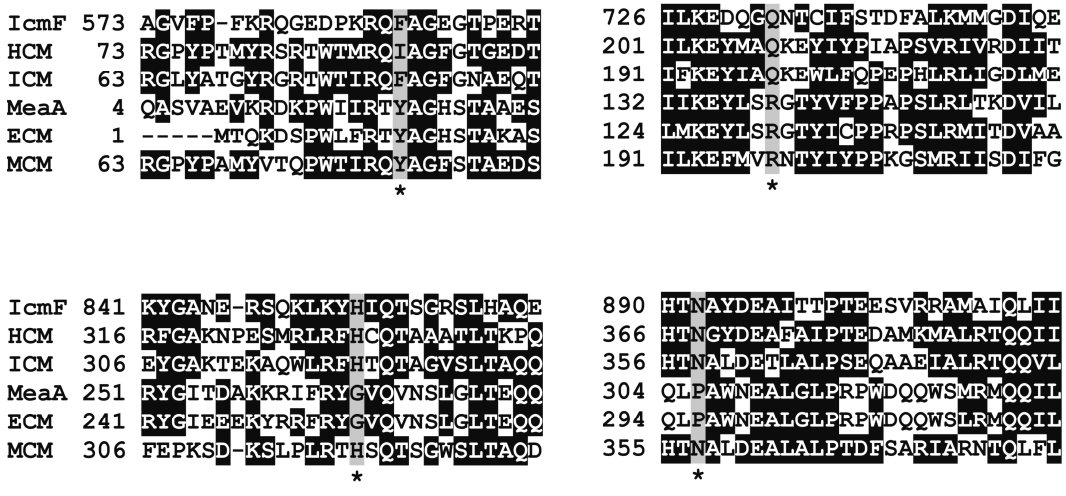


Figure 3.3: Multiple sequence alignment of the substrate-binding domain of different AdoCbl-dependent mutases. IcmF from *G.kaustophilus* (YP_149244), ICM from *S. cinnamonensis* (AAC08713), MCM from *M. extorquens* (YP_001642233), MeaA from *M. extorquens* (YP_002961419), ECM from *R. sphaeroides* (YP_354045) and HCM from *M. petroleiphilum* (YP_001023546). Four residues, which were recognized to be important for a substrate binding are highlighted in gray and indicated with asterisks. Two residues Phe and Gln, found in ICM and IcmF, are substituted by Tyr and Arg in MCM and ECM or Ile and Gln in HCM. ECM is different from other acyl-CoA mutases by the substitution of His and Asp to Gly and Pro, respectively. ECM was previously known as MeaA in *M.extorquens* (Section 1.4).

Inadvertent side reactions of the reactive radical intermediates render AdoCbl-dependent enzymes susceptible to inactivation (12). Alternatively, inactivation can result from escape of the 5′-deoxyadenosine intermediate during the catalytic cycle (13). In both cases, inactivation results from the inability to re-form AdoCbl from the 5′-deoxyadenosyl and cob(II)alamin radicals, at the end of the turnover cycle. MeaB, the G-protein chaperone of MCM, protects against inactivation in the presence of GTP (13,14). A similar role for the homologous G protein chaperone of ICM, MeaI has not been studied. In IcmF, the MeaI domain is sandwiched between the AdoCbl- and substrate-binding domains (3). In this study, we report a novel AdoCbl-dependent 1,2-rearrangement reaction catalyzed by IcmF and demonstrate that in the presence of GTP, the isovaleryl-CoA mutase activity of IcmF is partially protected from the inactivation.

3.3 Experimental procedures

3.3.1 DNA manipulations *(i) Cloning of IcmF.* In a previous study, we used IcmF from *G. kaustophilus* cloned into pET30 Ek/LIC expression vector (3). The S-tag which is located just downstream of the N-terminal His-tag in pET30 Ek/LIC was removed by subcloning the full-length IcmF DNA into the ligation independent cloning vector, pMCSG7(15). The insert was amplified with the following primers: (forward: 5′-TACTTCCAATCCAATGCCATGGCGCACATTTACCGTCC-3′ and reverse: 5′-TTATCCACTTCCAATGCTATTACATATTCGCCGGTATTGTCC-3′). In pMCSG7, which is a derivative of the pET21 vector, the N-terminal His-tag can be cleaved using TEV protease. (*ii*) *Construction of the K213A mutant of IcmF from G. kaustophilus* (*Gk*). The K213A mutation was created using the QuikChange XL Site-Directed Mutagenesis Kit (Agilent, CA) and the following sense primer:

5′-CCGGCACAGGCGGAGCTGGGGCAAGCTCGCTCACCGATG-3′. The sequence of the reverse primer was complementary to the sequence of the forward primer. The *Gk* IcmF cloned in pMCSG7 was used as a template. All constructs were confirmed by nucleotide sequence determination at The University of Michigan DNA sequencing Core.

3.3.2 Protein expression and purification. *(i) Purification of Gk wild-type and K213A IcmF. Escherichia coli* BL21 (DE3) cells containing plasmids expressing the IcmF constructs were grown at 37 ˚C in Luria-Bertani (LB) media supplemented with 100 μg/ml ampicillin to an absorbance at 600 nm of 0.5-0.6. Cells were grown for 12-14 h at 15 ˚C after induction with 0.5 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG). Proteins were purified as previously described for wild-type IcmF (3) with the following difference. Gel-filtration chromatography was performed with 50 mM HEPES pH 7.5,

containing 100 mM NaCl (Buffer A). *(ii) Purification of Cupriavidus metallidurans (Cm) IcmF.* *E. coli* BL21 (DE3) cells containing plasmid expressing IcmF were grown in LB media supplemented with 50 μg/mL kanamycin and induced at 15˚ C with 0.1 mM IPTG. Cells were harvested 12-14 h after induction. *Cm* IcmF was purified as previously described for *Gk* IcmF (3). Gel-filtration chromatography was also performed on this enzyme using Buffer A.

3.3.3 ATPase/GTPase assays. NTPase activity of IcmF was measured by HPLC as previously described (*3*) or by a modification of the malachite green assay, involving inclusion of citrate in the reaction mixture (16). Briefly, the color reagent was prepared by mixing 4.2 % ammonium molybdate in 4 N HCl with 0.045 % malachite green hydrochloride (1:3 v/v). The resulting solution was shaken for 30 min at 250 RPM in a falcon tube and filtered through a 0.2 μm Millipore filter. After that, 10% Tween 20 solution was added (200 μL for every 100 mL of color reagent). This solution was stable for a week at 4 ºC, with only a minor increase in the background absorbance at 650 nm. The reaction was performed in Buffer A in a total volume of 0.6-1 ml containing: 0.5-2.5 μM IcmF, 10 mM $MgCl_2$ and nucleotides (0.02-1.2 mM GTP or 0.02-6 mM ATP). At the desired time points, 200 μL aliquots were removed and the reaction was quenched with 20 μL of 2 N trichloroacetic acid (TCA). After centrifugation, 150 μL of supernatant was added to 750 μL of color reagent. After 3 min, 100 μL of 34% sodium citrate was added to the sample. The color was allowed to develop for 30 min at room temperature and the absorbance was measured at 650 nm. The calibration curve for phosphate was prepared using 5-200 μM solution of monobasic potassium phosphate, which was dried in an oven (to remove traces of moisture) for 4 h at 120 ºC.

3.3.4 IcmF assay. A GC-based assay was used to measure both ICM activity and the new, isovaleryl-CoA/pivalyl-CoA mutase activity (3). In all assays, apo-enzyme was preincubated with AdoCbl ± nucleotides, and the reaction was started by addition of substrate. *(i) Pivalyl-CoA mutase activity.* The reaction was performed in Buffer A in a total volume of 0.8-1.4 ml containing: 600-2500 μg of *Gk* IcmF or *Cm* IcmF, 100 μM AdoCbl, 20-2000 μM isovaleryl-CoA, 15 mM $MgCl_2$ ± 5 mM GTP. For every concentration of substrate, 2-6 aliquots (200 μl each) were removed and treated as described below. *(ii) Isobutyryl-CoA mutase activity.* The reaction was performed in Buffer A in a total volume of 0.8-1.4 ml containing: 10-60 μg of *Gk* IcmF or *Cm* IcmF, 100 μM AdoCbl, saturating concentration of substrates (600-2000 μM isobutyryl-CoA or n-bytyryl-CoA), 10 mM $MgCl_2$ ± 3-6 mM GTP or ATP. At various time points (0.5-30 min), 200 μL aliquots were removed and quenched with 100 μL of 2 N KOH containing 0.18 mM valeric acid used as an internal standard. Following addition of 100 μL of H_2SO_4 (15%, v/v), the reaction mixture was saturated with NaCl and extracted with ethyl acetate (250 μL). The extract was analyzed directly by GC using a DB-FFAP (30 m x 0.25 mm I.D., 0.25 μm) capillary column (Agilent, CA). A 5 μl sample was injected in the pulsed split-less mode. The oven temperature was initially at 80 ºC. Following sample injection, the temperature was raised to 150 ºC at a rate of 10 ºC/min and maintained at 150 ºC for 2 min. Retention times for the compounds of interest was as follows: isobutyric acid 5.85 min, pivalic acid: 5.99 min, n-butyric acid: 6.5 min, isovaleric acid: 6.96 min, valeric acid: 7.78 min.

3.3.5 IcmF assays with alternative substrates. We used an HPLC-based assay to evaluate whether 3-hydroxybutyryl-CoA can be converted by IcmF to 2-

hydroxyisobutyryl-CoA. The assay mixture contained in a total volume of 0.5 mL: Buffer A, 10 mM $MgCl_2$, 100 μM AdoCbl, 2 mM GTP, 0.4-1 mM D,L-β-hydroxybutyryl-CoA and 0.2-0.4 mg *Gk* IcmF. The reaction was initiated by addition of enzyme and the reaction mixture was incubated at 37 ºC. At different time points (0.5-10 min), 60 μL aliquots were removed, quenched with 2N TCA (10% v/v), centrifuged and subjected to HPLC analysis.

The acyl-CoA esters were separated using an HPLC system equipped with an Alltima HP 5 μm C18 (250 x 4.6 mm) column (Grace, IL). The detector was set at 254 nm. Solvent A was 50 mM monobasic potassium phosphate, pH 5.4. Solvent B was prepared as follows: 500 mL of MeOH was added to 500 mL of 100 mM monobasic potassium phosphate, pH 5.4 (to give a final concentration of 50 mM potassium phosphate and 50% MeOH). Initial conditions used for separation were: 10% solvent B; a flow rate of 1.0 mL/min. Between 5 and 30 min, solvent B was increased to 100% and then held at 100% solvent B for 5 min. At 36 min, solvent B was decreased to 10% and held for 10 min at that composition to equilibrate the column between injections. Under these conditions, the retention time for 3-hydroxybutyryl-CoA was 20.1 min. An authentic standard of 2-hydroxyisobutyryl-CoA was not available, but based on the known performance of the C18 column, branched acyl-CoAs elute earlier than linear isomers.

3.3.6 Enzyme-monitored turnover of IcmF. Changes in the spectra of *Gk* IcmF-bound AdoCbl were monitored by UV-visible spectroscopy at 24 °C in Buffer A containing 5-15 mM $MgCl_2$. Substrates (final concentration of 0.5-4 mM) were added to 20-65 μM apo-IcmF loaded with one or two equivalents of AdoCbl. To check the influence of the MeaI domain on catalytic turnover, the reaction was also supplemented with 1-5 mM

GTP or ATP. The amount of cob(II)alamin formed under steady-state turnover conditions was calculated from the decrease in absorbance at 525 nm upon substrate addition using a value of $\Delta\varepsilon_{525nm}$= -4.8 mM^{-1} cm^{-1} (14). During the course of the reaction, AdoCbl was gradually converted to enzyme-bound aquacobalamin (OH_2Cbl), as indicated by the appearance of a 350 nm absorption peak. The increase at 350 nm was fitted to a single-exponential function: $A= A_0+A1(1-e^{-bt})$, where A is the absorbance at 350 nm, b is the observed rate constant for inactivation, A_0 is the initial absorbance at 350 nm, t is time in minutes and A1 is the amplitude.

3.3.7 Enzyme-monitored turnover of *Gk* IcmF under anaerobic conditions. To assess the effect of oxygen on enzyme inactivation, enzyme-monitored turnover experiments were performed under anaerobic conditions. For this, Buffer A containing 10 mM $MgCl_2$ was bubbled with N_2 for 3 h before use and introduced into an anaerobic chamber (containing <0.3 ppm O_2). Stock solutions of AdoCbl, n-butyryl-CoA, isobutyryl-CoA and isovaleryl-CoA were prepared in the chamber using anaerobic buffer. The enzyme solution was deoxygenated by blowing a stream of N_2 over its surface at 4 ºC for 40 min. UV-visible spectra were collected using Mikropack DH-2000 UV-visible light source connected with fiber optics to a cuvette holder inside the glove box.

3.3.8 HPLC characterization of inactivation products. A solution containing 64 μM *Gk* holo-IcmF (containing two equivalents of AdoCbl) in Buffer A and 10 mM $MgCl_2$ ± 5 mM GTP was incubated with 1 mM isobutyryl-CoA at 37 °C in the dark. At various times (0 to 60 min), 15 μL aliquots were removed and immediately quenched with 60 μL of 0.5% trifluoroacetic acid (TFA). The aerobic inactivation products of AdoCbl, 5′-deoxyadenosine and OH_2Cbl, were monitored by HPLC using an Alltima HP 5 μm C18

(250 x 4.6mm) column (Grace, IL). All steps of sample preparation and HPLC analysis were performed in the dark. Initial buffer conditions were 92% Solvent C (0.1%TFA in water) and 8% Solvent D (0.1%TFA in acetonitrile) at a flow rate of 1.0 mL/min. Between 10 and 35 min, Solvent D was increased to 32%. Between 35 and 36 min, Buffer D was decreased to 8 % and held for 5 min at that composition to equilibrate the column between injections. 50 μL of the sample was injected and elution was monitored at 254 nm and 350 nm. Under these conditions, the following retention times were obtained: 6.92 min for 5′-deoxyadenosine, 22.77 min for OH_2Cbl and 29.27 min for AdoCbl. The control reaction was performed in the absence of isobutyryl-CoA. Calibration curves were generated for all three compounds prepared and treated similarly as the assay samples. An extinction coefficient of ε_{260nm}= 15 mM^{-1} cm^{-1} was used to estimate the concentration of 5′-deoxyadenosine (17). The data were well fitted by a single exponential function for the disappearance of AdoCbl and appearance of 5′-deoxyadenosine and OH_2Cbl. To improve recovery of OH_2Cbl, proteinase K (Roche) was used to digest the protein sample for 1-2 h at 37 ºC before precipitation with acid.

3.3.9 Bioinformatics analysis. Operon and regulon browsers on the Microbes Online web site were used for the elucidation of functional predictions for the genes of interest (18).

3.4. Results

3.4.1 Gene Neighborhood Analysis for *icmF*. In several bacteria, the *icmF* gene is located in the same operon or in close proximity to genes encoding enzymes involved in fatty acid metabolism (3). For example, enzymes found in the operon with *icmF* in several bacteria (*Bacillus selenitireducens*, *Lysinibacillus sphaericus*, *Anoxybacillus flavithermus*, *B.megaterium*, *B.halodurans*, *B.pseudofirmus*, *B.coagulans* and *Bacillus sp.* NRRL B-14911, *Geobacillus sp.* WCH70 and *Brevibacillus brevis*) are annotated as enzymes in the mother cell metabolic gene (*mmg*) operon, which has been described in *B.subtilis* (19) (Figure 3.4). Three genes in this operon are annotated as: *mmg*A (acetyl-CoA transferase), *mmg*B (3-OH-butyryl-CoA dehydrogenase) and *mmg*C (acetyl-CoA dehydrogenase) (19). A similar set of enzymes in *Ralstonia eutropha* encoded by the H16_A0459-H16_A0464 operon, allows growth on long chain fatty acids. The absence of this operon together with the H16_A1526-H16_A1532 operon, renders *R. eutropha* unable to grow on plant oils or long chain fatty acids as a carbon source (20). The gene *acd*H from *Streptomyces coelicolor* and *Streptomyces avermitilis,* which is homologous to *mmg*C, is an acyl-CoA dehydrogenase and plays a role in the catabolism of branched-chain amino acids (21). Finally, the presence of the *rpo*E gene just down stream of *icm*F, which encodes the sigma subunit of RNA polymerase, strongly suggests that IcmF is linked to fatty acid metabolism (22) (Figure 3.4).

3.4.2 Alternative substrates for IcmF. Based on the above gene neighborhood analysis, we sought to assess whether IcmF plays a role in the metabolism of branched fatty acids and can catalyze the isomerization of substrates other than isobutyryl-CoA and n-butyryl-CoA. We decided to use isovaleryl-CoA, a building block for iso-odd branched fatty

acids, and 2-methylbutyryl-CoA, a building block for anteiso branched fatty acids as potential substrates for IcmF. Isovaleryl-CoA and 2-methylbutyryl-CoA are synthesized from the branched-chain amino acids, leucine and isoleucine, respectively in reactions catalyzed by the branched-chain ketoacid dehydrogenase complex (21).

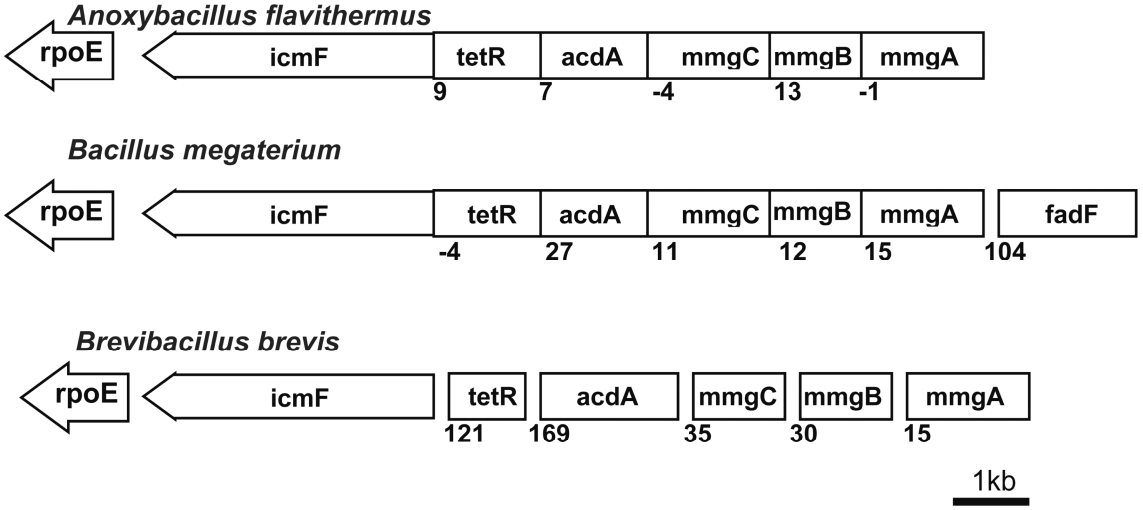


Figure 3.4. Organization of genes in the *mmg*ABC operon harboring the *icm*F gene. (**A**) *Anoxybacillus flavithermus,* (**B**) *Bacillus megaterium and* (**C**) *Brevibacillus brevis.* The following genes are found in the same operon with *icm*F: tetR (TetR/Acr-like family transcriptional regulator), *acd*A (acyl-CoA dehydrogenase), *mmg*A (acetyl-CoA transferase), *mmg*B (3-OH-butyryl-CoA dehydrogenase), *mmg*C (acetyl-CoA dehydrogenase) and *fad*F (medium-/long-chain fatty acyl-CoA dehydrogenase). *rpo*E gene which encodes the sigma subunit of RNA polymerase, is found just downstream of *icm*F. The numbers below the genes indicate the distance in nucleotides between two adjacent genes. Negative numbers indicate overlapping genes. For a complete list of bacteria showing similar operonic organization see the text (Section 3.4.1).

Gk IcmF was mixed with isovaleryl-CoA in the presence of AdoCbl and GTP/$MgCl_2$ and the products were hydrolyzed and analyzed by GC. A time-dependent decrease in the isovaleric acid peak at 6.96 min was accompanied by the appearance of a peak at 5.99 min, suggesting conversion of isovaleryl-CoA to a new compound (Figure 3.5). The expected product of AdoCbl-dependent 1,2-rearrangement of isovaleryl-CoA (or 3-methyl-butyryl-CoA) is pivalyl-CoA (or 2,2-dimethylpropionyl-CoA) (Figure 3.1) (1).

The retention time of a standard pivalic acid sample (5.99 min) exactly coincided with that of the product formed from isovaleryl-CoA (Figure 3.5). To confirm our findings we also purified IcmF from *C.metallidurans* as described under Experimental Procedures (section 3.3.2) and (Figure 3.6). The isomerization of isovaleryl-CoA to pivalyl-CoA by both *Gk* IcmF and *Cm* IcmF is slow in comparison to the isomerization of n-butyryl-CoA to isobutyryl-CoA (Table 3.1). The K_M value for isovaleryl-CoA for *Gk* IcmF is 62 ± 8 μM and the specific activity of the *Gk* IcmF with isovaleryl-CoA is 0.021 ± 0.004 μmol min^{-1} mg^{-1}, which is ~150-fold lower than the activity with n-butyryl-CoA (3.25 ± 0.35 μmol min^{-1} mg^{-1}). With the *Cm* IcmF, an ~2200 fold difference in the specific activities was obtained with the two substrates (Table 3.1).

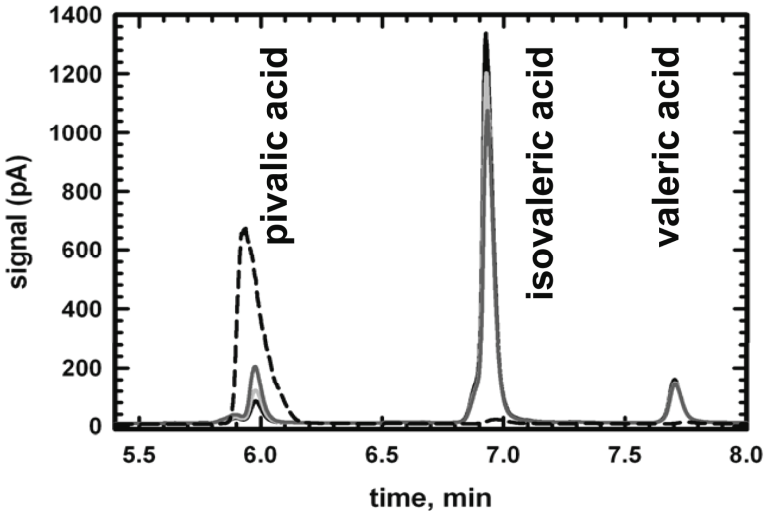


Figure 3.5 Pivalyl-CoA mutase activity of IcmF. A representative GC chromatogram showing separation of isovaleric and pivalic acids following enzymatic conversion of isovaleryl-CoA to pivalyl-CoA. The reaction mixture contained in Buffer A, 15 mM $MgCl_2$, 2000 μg *Gk* IcmF, 100 μM AdoCbl, 5 mM GTP and 1.56 mM isovaleryl-CoA. At different time points, aliquots were removed, the esters hydrolyzed and the corresponding acids separated by GC as described under Experimental Procedures. The traces represent: black line (1 min), light gray line (5 min), dark gray line (25 min), dashed line (pivalic acid standard). Valeric acid was used as an internal standard.

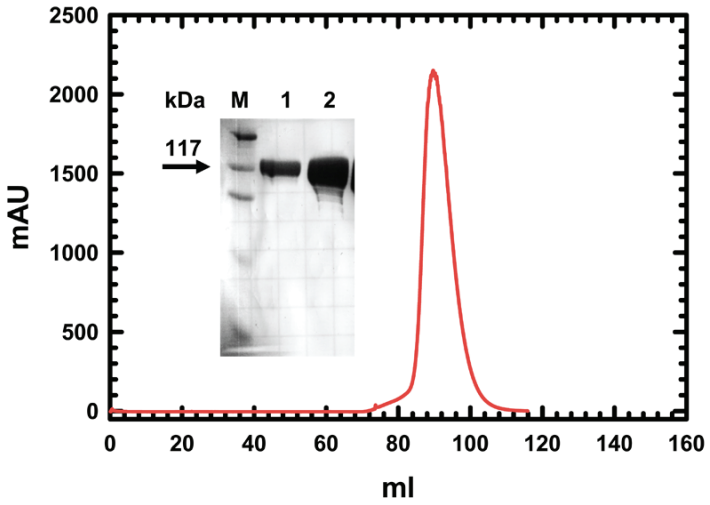


Figure 3.6: Gel-filtration of IcmF from *C.metallidurans*. mAU, milliabsorbance units. Inset: The SDS-PAGE analysis of IcmF after gel-filtration (lane #1, 10 µg load; lane #2, 30 µg load).

Isomerization of valeryl-CoA is expected to produce 2-methylbutyryl-CoA. However, consumption of valeryl-CoA was not observed in the presence of IcmF, AdoCbl and GTP/$MgCl_2$ in the GC-based assay (data not shown).

Substrate	*Gk* IcmF	*Cm* IcmF
	S.A., µmol min^{-1} mg^{-1}	S.A., µmol min^{-1} mg^{-1}
isovaleryl-CoA	0.021 ± 0.004	0.015 ± 0.004
isobutyryl-CoA	1.2 ± 0.1	13.8 ± 1.1
n-butyryl-CoA	3.3 ± 0.4	33.0 ± 1.4

Table 3.1: Isobutyryl-CoA mutase and pivalyl-CoA mutase activities of recombinant IcmFs. Values represent the average of at least 3 independent experiments. In all experiments, saturating concentration of substrate was used.

Since HCM (5,23) from *M. petroleiphilum,* is very similar to the stand-alone ICM from *S. cinnamonensis* and catalyses the interconversion of 3-hydroxybutyryl-CoA and 2-hydroxyisobutyryl-CoA (5,24), we decided to examine their substrate specificities

overlap. However, conversion of D,L-3-hydroxybutyryl-CoA to 2-hydroxyisobutyryl-CoA was not observed as judged by an HPLC-based assay (data not shown).

3.4.3 Absorption spectrum of *Gk* IcmF during steady-state turnover. The binding-sites for AdoCbl in the *Gk* IcmF dimer are nonidentical and exhibit an ~25-fold difference in affinities (K_{D1}=0.08±0.01 μM, K_{D2}=1.98 ± 0.4 μM) (3). For enzyme-monitored turnover experiments, *Gk* IcmF was reconstituted with one equivalent of AdoCbl to avoid the presence of free cofactor. The absence of free AdoCbl was confirmed by analyzing the spectrum of the flow-through obtained upon concentrating the reaction mixture using a Microcon 30K concentrator. Addition of substrate (isobutyryl-CoA or n-butyryl-CoA) to holo-IcmF resulted in cob(II)alamin formation as evidenced by the decrease in absorbance at 525 nm and increase at 480 nm (Figure 3.7 A). Using $\Delta\varepsilon_{525nm}$= -4.8 mM^{-1} cm^{-1} (14), the ratio of AdoCbl:cob(II)alamin under steady-state turnover conditions was estimated to be ~2:1 when isobutyryl-CoA was used as a substrate. With isovaleryl-CoA, the ratio of AdoCbl:cob(II)alamin under steady-state turnover conditions was 9:1 (Figure 3.7 B).

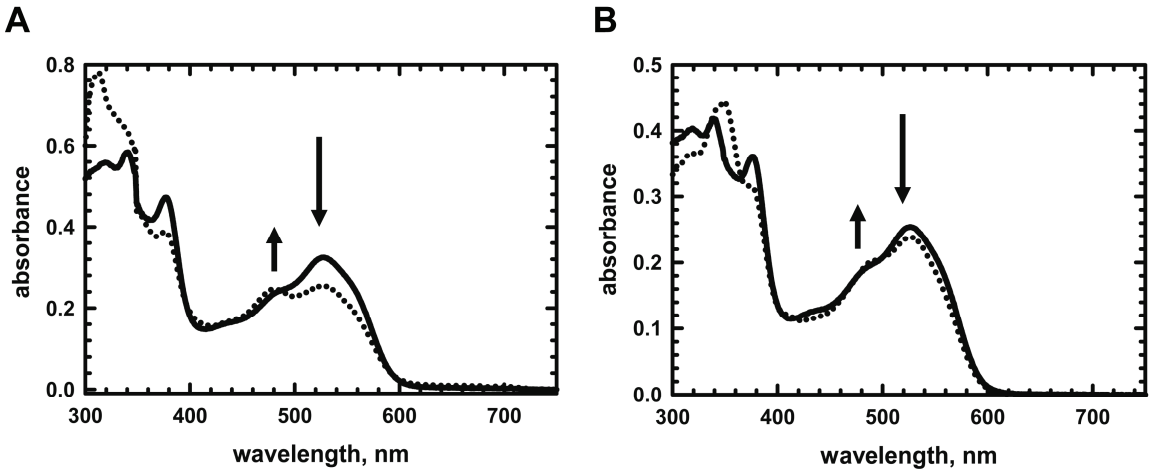


Figure 3.7: Spectral changes in *Gk* holo-IcmF in the presence of isobutyryl-CoA and isovaleryl-CoA. (**A**) Spectra of holo-IcmF (40 μM) in buffer A at 24 ˚C (solid line) and

after addition of 4.1 mM isobutyryl-CoA which resulted in formation of cob(II)alamin (dotted line). (**B**) Spectra of holo-IcmF (31 μM) in Buffer A with 10 mM $MgCl_2$ at 24 ˚C (solid line) and after addition of 1.5 mM isovaleryl-CoA (dotted line).

3.4.4 Inactivation of IcmF and the effect of nucleotides. Cobalamin spectra under turnover conditions were used for monitoring the effect of nucleotides on the IcmF-catalyzed reaction (Figure 3.8). Addition of isobutyryl-CoA results in formation of cob(II)alamin, which is converted over time to OH_2Cbl (Figure 3.8). The increase in absorption at 350 nm (k_{obs}=0.1 ± 0.01 min^{-1}, (Figure 3.8, inset)) and the distinctive double peaks at 505 and 540 nm indicated formation of OH_2Cbl, a hallmark of inactivation for AdoCbl-dependent enzymes (12,25,26). The amplitude but not the rate of OH_2Cbl formation was diminished in the presence of GTP (k_{obs}= 0.063 ± 0.002 min^{-1}) and ATP (k_{obs}= 0.10 ± 0.01 min^{-1}) (Figure 3.8, inset).

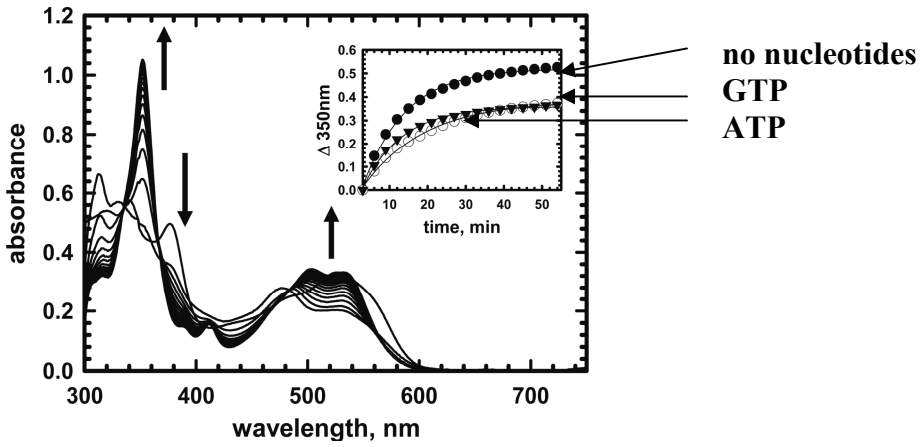


Figure 3.8: Inactivation of *Gk* IcmF during turnover with isobutyryl-CoA. Rapid oxidation of cob(II)alamin to OH_2Cbl was observed during reaction of holo-IcmF (with one equivalent, 41 μM AdoCbl, bound) with 1.4 mM isobutyryl-CoA in Buffer A with 5 mM $MgCl_2$ at 24 ºC in the dark. The spectra were recorded between 0 and 60 min. *Inset*: The time-dependent increase at 350 nm was fitted to a single exponential function in the absence of nucleotides (filled circles) (k_{obs}=0.096 ± 0.008min^{-1}), or in the presence of 5 mM GTP (open circles) (k_{obs}=0.062 ± 0.002 min^{-1}) or 5 mM ATP (triangles) (k_{obs}=0.10 ± 0.006 min^{-1}).

Under standard *in vitro* assay conditions, *Gk* IcmF exhibits a linear reaction time course for only 40-60 s before the activity plateaus off (Figure 3.9 A). The specific activity with isobutyryl-CoA (1.1 ± 0.1 μmol $min^{-1}mg^{-1}$) was calculated using the linear portion of the curve and was comparable to values obtained previously in the coupled enzyme assay (3). When n-butyryl-CoA was used as a substrate, a specific activity of 3.25 ± 0.35 μmol $min^{-1}mg^{-1}$ was obtained. With both isobutyryl-CoA and n-butyryl-CoA, inactivation of *Gk* IcmF resulted in termination of the reaction in ~6-10 min, during which time, only ~7% of substrate was consumed (Figure 3.9 A). Very similar behavior was also observed with *Cm* IcmF (data not shown). Protein stability, the presence of metal ions (K^+, Ca^{2+}), reductants (DTT or TCEP) and different buffers (50 mM sodium phosphate pH 7.5/100 mM NaCl; 50 mM potassium phosphate pH 7.5/100mM KCl and 50 mM HEPES pH 7.5/100mM NaCl) were assessed, and ruled out as contributors to the observed inactivation (data not shown). Nucleotides (ATP and GTP) decreased the activity of *Gk* IcmF (Figure 3.9 A) as reported previously with GDP and GTP (3).

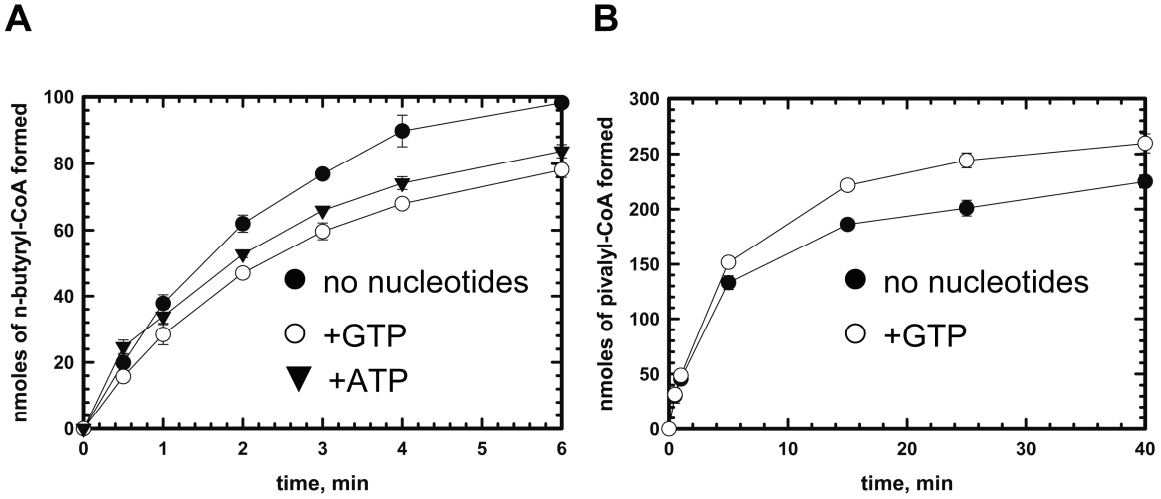


Figure 3.9: Effect of nucleotides on the time course of reactions catalyzed by IcmF. (**A**) Time course of the isobutyryl-CoA mutase reaction catalyzed by Gk IcmF at 37 ˚C. The assay mixture in Buffer A with 10 mM $MgCl_2$ contained 40 μg of holo-IcmF, 100

μM AdoCbl, 1.5 mM isobutyryl-CoA and either no nucleotides (black circles), 6 mM ATP (triangles), or 3 mM GTP (open circles). (**B**) Time course of the isovaleryl-CoA mutase reaction catalyzed by *Gk* IcmF. The reaction mixture in Buffer A with 15 mM $MgCl_2$ contained 2500 μg of IcmF, 100 μM AdoCbl, 1.5 mM isovaleryl-CoA with (white circles) or without 5 mM GTP (black circles) at 37 °C. Aliquots of the reactions were removed at different time points and analyzed by GC as described under Experimental Procedures. Data represent the average of three independent experiments.

With isovaleryl-CoA, the *Gk* IcmF reaction was almost completely inhibited after 25-30 min (Figure 3.9 B). However, in the presence of GTP, the enzyme was inactivated to a lesser extent in comparison to the reaction without GTP (Figure 3.9 B). Similar behavior was seen for *Cm* IcmF (data not shown). Interestingly, for the first 2 min, no difference was seen in the isovaleryl-CoA/pivalyl-CoA mutase activity ± GTP (Figure 3.9 B).

3.4.5 Loss of 5′-deoxyadenosine leads to inactivation of IcmF. The steady formation of OH_2Cbl during IcmF turnover under aerobic conditions suggests that inactivation might result from oxidative interception of cob(II)alamin (27). An alternative possibility is that inactivation is associated with loss of 5′-deoxyadenosine from the active site and the uncoupled cob(II)alamin subsequently oxidizes to OH_2Cbl. To distinguish between these possibilities, enzyme-monitored turnover experiments were conducted under anaerobic conditions (Figure 3.10 A). Addition of n-butyryl-CoA to holo-IcmF resulted in formation of cob(II)alamin, which was not further converted to OH_2Cbl even after 60 min (Figure 3.10 A).

Next, we compared the time courses of the reaction catalyzed by *Gk* IcmF with n-butyryl-CoA under aerobic and anaerobic conditions (Figure 3.10 B). Under anaerobic conditions, the reaction exhibited a linear dependence for a longer time period (2 versus 1 min), but eventually ceased after 10 min. Since spectroscopic analysis of the reaction

time course shows that OH_2Cbl is not formed under these conditions, it demonstrates that inactivation is not directly correlated with OH_2Cbl formation.

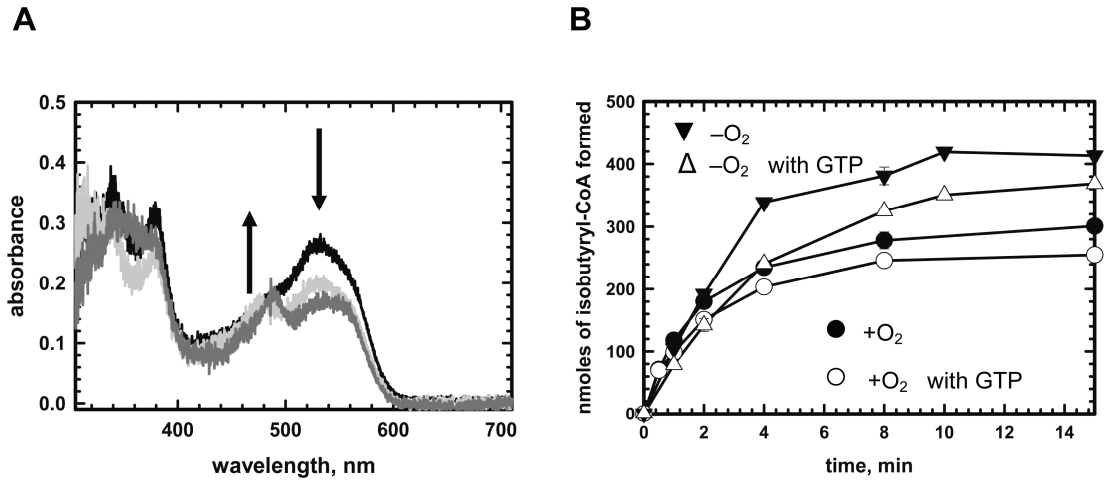


Figure 3.10: Inactivation of *Gk* IcmF under anaerobic conditions. (**A**) Spectral changes upon incubation of 33 μM holo-IcmF (containing one equivalent of bound AdoCbl) with 4.8 mM n-butyryl-CoA in Buffer A with 5 mM $MgCl_2$ under anaerobic conditions at 24 ºC in the dark. Spectra were recorded at time = 0 (black), immediately after addition of substrate (light gray) and 60 min after addition of substrate (dark grey). Formation of cob(II)alamin (480 nm peak) without further conversion to OH_2Cbl was observed. (**B**) Comparison of time courses for the reaction catalyzed by *Gk* IcmF at 37 °C under aerobic (circles) and anaerobic (triangles) conditions in the presence of GTP. The aerobic and anaerobic assay mixtures in Buffer A with 15 mM $MgCl_2$ contained 40 μg of IcmF and 2 mM n-butyryl-CoA and either no nucleotides (solid symbols) or 4.3 mM GTP (open symbols).

As seen under aerobic conditions, GTP also led to decreased OH_2Cbl formation under anaerobic conditions (Figure 3.10 B).

3.4.6 Characterization of inactivation products by HPLC. The reaction of 64 μM *Gk* holo-IcmF (containing two equivalents of AdoCbl) with 4 mM isobutyryl-CoA was monitored as described under Experimental Procedures (Figure 3.11). Under these conditions, AdoCbl disappeared with a k_{obs} = 0.40 ± 0.04min^{-1} (Table 3.2). The rate of appearance of 5′-deoxyadenosine (k_{obs}=0.32 ± 0.01 min^{-1}) corresponded to the rate of AdoCbl disappearance. The concentration of 5′-deoxyadenosine recovered (20.6 ± 0.8

μM) was equal to the concentration of AdoCbl consumed (23 ± 1 μM). In contrast, the rate of OH_2Cbl formation (k_{obs}= 0.23 ± 0.02 min^{-1}) was slower and only 4.34 ± 0.01 μM OH_2Cbl was recovered (Table 3.2). Treatment of inactivated enzyme with proteinase K increased OH_2Cbl recovery only marginally (6.3 ± 0.2 μM). The reason for the low yield of OH_2Cbl is presently not clear. To ensure that AdoCbl is not cleaved non-enzymatically to 5´-deoxyadenosine during sample preparation and/or chromatographic separation, a control reaction was performed in which substrate was omitted from the reaction mixture (Figure 3.11).

IcmF appears to show half of sites activity since reconstitution with 1 versus 2 equivalents of AdoCbl resulted in the same steady-state concentration of cob(II)alamin (data not shown). Surprisingly, in the inactivation experiments, not all the AdoCbl bound to *Gk* IcmF was converted to OH_2Cbl and 5´-deoxyadenosine.

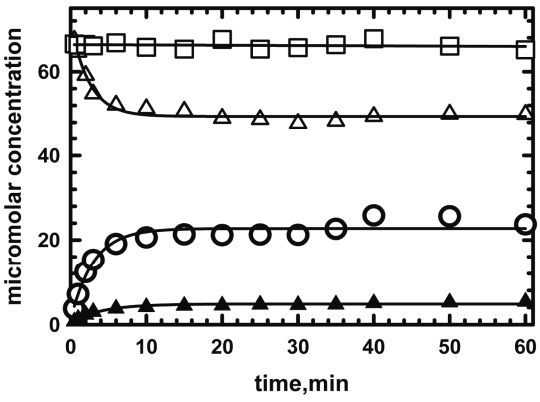


Figure 3.11: Formation of OH_2Cbl and 5'-deoxyadenosine during enzyme-monitored turnover. Holo-IcmF (64 μM IcmF active site concentration containing two equivalents of AdoCbl) was mixed with 1.5 mM isobutyryl-CoA in Buffer A at 37 ºC. All manipulations with the samples and HPLC were performed in the dark. The decay of AdoCbl (open triangles), appearance of OH_2Cbl (solid triangles) and of 5'-deoxyadenosine (open circles) was monitored over 50 min. As a control for AdoCbl stability during samples handling, the analysis was repeated without addition of isobutyryl-CoA (open squares).

Instead, ~ 66 % of the AdoCbl at one site (i.e. 32 μM) was converted to 5′-deoxyadenosine, which corresponded to the concentration of cob(II)alamin (~21 μM) under steady-state turnover conditions. This provides further evidence for half-of-sites activity in *Gk* IcmF. It is unclear why complete inactivation at one of the two active sites was not observed.

3.4.7 ATPase activity of IcmF. In the MeaI domains of many IcmFs, the base-specificity loop motif NKxD, is modified to NKxE (3). In the *Gk* IcmF, the NKxE sequence is present whereas in the *Cm* IcmF, the sequence is NKxD (Figure 3.12). The Lys residue in the NKxD motif interacts via hydrophobic interactions with the plane of the guanine ring while the Asp coordinates two nitrogen atoms in the purine ring. *Gk* IcmF catalyses the hydrolysis of both GTP and ATP (Table 3.3). The kinetic parameters for IcmF measured in the presence of various nucleotides are comparable: k_{cat} with ATP is 19 ± 1 min^{-1} and k_{cat} with GTP is 10 ± 1 min^{-1}. Although the K_M for ATP is high (1290 ± 300 μM) relative to that for GTP (51 ± 3 μM), the concentration of ATP (3-5 mM) is higher than of GTP (1 mM) in bacterial cells (28).

	AdoCbl Consumption		OH_2Cbl Production		5'-deoxyadenosine Production	
nucleotide	k_{obs}, min^{-1}	concentration, μM	k_{obs}, min^{-1}	concentration, μM	k_{obs}, min^{-1}	concentration, μM
-	0.4 ± 0.04	23 ± 1	0.23 ± 0.02	4.34 ± 0.01	0.32 ± 0.01	20.6 ± 0.8
GTP	0.37 ± 0.04	20.65 ± 1.02	0.25 ± 0.04	5.9 ± 1.4	0.35 ± 0.05	19.9 ± 1.3

Table 3.2: Kinetics of OH_2Cbl and 5´-deoxyadenosine formation. Values represent the average of at least two independent experiments.

To rule out the possibility that the ATPase activity of *Gk* IcmF is due to a contaminant that co-purifies with the mutase, first, we tested the effect of introducing the K213A mutation in the GtgGaGKSS sequence of the P-loop motif, which is important for phosphate-binding (29). The K213A *Gk* IcmF mutant was devoid of both GTPase and ATPase activities. Notably, the mutase activity of the K213A mutant was similar to wild-type protein (0.6 μmol min^{-1} mg^{-1} with isobutyryl-CoA). Furthermore, in the presence of 3 mM AMPPNP, the k_{cat} for the GTPase activity was inhibited ~3.6-fold to 2.7 ± 0.14 min^{-1}. Taken together, the above results are consistent with the ATPase and GTPase activities being intrinsic to *Gk* IcmF.

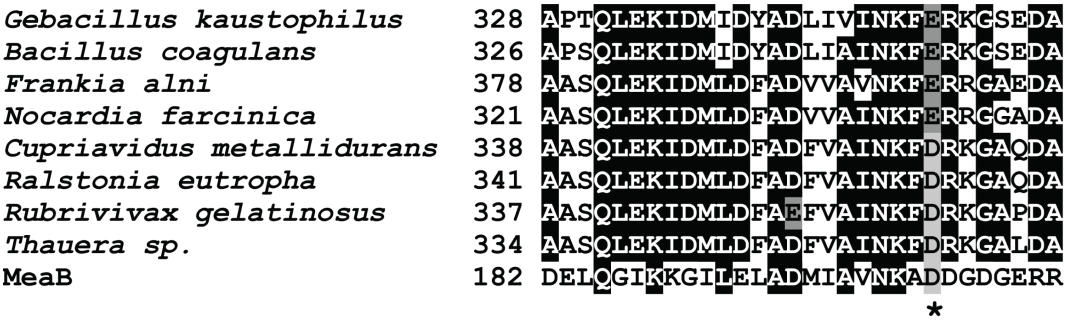


Figure 3.12: Multiple sequence alignment of IcmFs and MeaB from *M. extorquens* showing base specificity loop NKxD/E. Accession numbers are: *Gebacillus kaustophilus* (YP_149244), *Cupriavidus metallidurans* CH34 (YP_582365), *Ralstonia eutropha* H16 (YP_724799), *Frankia alni* (YP_716016), *Nocardia farcinica* (YP_117245), *Bacillus coagulans* (ZP_01696637), *Thauera* sp. (ZP_02841697), *Rubrivivax gelatinosus* (ZP_00242991) and MeaB (AAL86727).

Although it has been suggested that replacement of Asp by Glu in the NKxD motif has no effect on nucleotide specificity (30), our results suggest the contrary. Thus, in *Cm* IcmF, the base specificity loop sequence is NKFD and while this protein does not exhibit ATPase activity, it is an active GTPase (k_{cat} = 18 ± 1.3 min^{-1}, K_{GTP} = 40 ± 8 μM). These

results provide an interesting illustration of a G-protein losing its specificity due to a single substitution in the NKxD motif.

	GTP[a]			ATP[b]		
enzyme	k_{cat}, min^{-1}	K_{GTP}, μM	k_{cat}/K_{GTP}, M^{-1} min^{-1}	k_{cat}, min^{-1}	K_{ATP}, μM	k_{cat}/K_{ATP}, M^{-1} min^{-1}
wild type	10 ± 2	51 ± 3	$(1.96 \pm 0.37)x10^5$	19 ± 1	1290 ± 300	$(1.47 \pm 0.03)x10^4$
K213A	N.D.	-	-	N.D.	-	-

Table 3.3: Comparison of the GTPase and ATPase Activities of *Gk* IcmF. [a,b]All experiments were performed in Buffer A with 20 mM $MgCl_2$ at 37 °C as described under Experimental Procedures. Values represent the average of at least five independent experiments. N.D. is not detected.

3.5 Discussion

Previously, it was believed that ICM-like activity was restricted to the *Streptomyces* genus where it is involved in monensin A and B production (*4*). Our discovery of ICM activity in the IcmF fusion protein, which is widely distributed in bacteria, suggests its involvement in metabolic processes beyond polyketide synthesis (*3*). In our efforts to elucidate the possible roles of IcmF in bacterial metabolism, we noticed that *icm*F genes co-localize with genes encoding enzymes involved in β-oxidation of fatty acids. Subsequently, we probed branched organic acids, which are synthesized from the branched amino acids, valine, leucine and isoleucine, as substrates for IcmF. This in turn, led to the discovery of a new IcmF activity, i.e. isomerization of isovaleryl-CoA/pivalyl-CoA (Figure 3.1).

Although AdoCbl-dependent pivalyl-CoA mutase activity has been predicted to exist (10,31), an enzyme with this catalytic activity has not been previously identified. In this study, we report that IcmFs from *G. kaustophilus* and *C. metallidurans* can convert isovaleryl-CoA to pivalyl-CoA (Figure 3.1). Depending on the organism, the isovaleryl-CoA/pivalyl-CoA mutase activity of IcmF was ~150-2,200 fold lower than the conversion of n-butyryl-CoA to isobutyryl-CoA (Table 3.1). Recently it was shown that pivalic acid can be incorporated as a starter unit in fatty acids in several bacteria (32). Thus, the production of pivalyl-CoA catalyzed by IcmF might be important in bacteria which use this starter unit for the biosynthesis of branched fatty acids containing a quaternary carbon. Our discovery of the isovaleryl-CoA/pivalyl-CoA mutase activity adds to the growing list of carbon-skeleton rearrangements catalyzed by AdoCbl-dependent isomerases (Figure 3.1). Thus, it appears that the "mutase core", is quite

versatile for two reasons. First, substitutions of a limited number of key active site residues alter substrate specificity and second, relaxed substrate specificity allows alternative reactions to be catalyzed by the same active site as exemplified by IcmF.

We also report relaxed substrate specificity in the MeaI domain of *Gk* IcmF. It has been reported that some G proteins have either lost or switched nucleotide specificity (33). For example in centaurin gamma-1 GTPase, where the sequence of the base specificity loop NKxD, is modified to GTQD(R) (34), nucleotide specificity is lacking and it is described as a general NTPase (34). Proteins belonging to the YchF subfamily of the Obg family of G proteins harbor the NxxE motif instead of the NKxD (35). For example human OLA1, which belongs to the Obg family, hydrolyzes ATP more efficiently than GTP.

In our study, we show that in a subset of the MeaI chaperone domains of IcmF, the NKxD motif is modified to NKxE (Figure 3.12). This constitutes an interesting example where only some members of a protein family have lost specificity for GTP. We note that while the MeaI domain of *Gk* IcmF exhibits ATPase activity, it is functionally distinct from the ATP-dependent chaperones for AdoCbl-dependent eliminases e.g. diol dehydratases (36,37).

A remarkable feature of AdoCbl-dependent enzymes is that they catalyze reactions involving radical intermediates under aerobic conditions. This however, comes with a price, i.e. their susceptibility to inactivation (12). Under standard in vitro assay conditions, IcmF inactivates quite rapidly with either isobutyryl-CoA or isolvaleryl-CoA as substrate (Figure 3.9). In a subclass of AdoCbl-dependent enzymes, reactivating factors mediate an ATP-dependent exchange of enzyme-bound OH_2Cbl with free AdoCbl (36-39). These reactivases have sequence similarity to DnaK and to other members of the

Hsp70 family of molecular chaperones and lower sequence similarity to the large subunits of corresponding mutases (37). In contrast, the G-protein chaperones associated with AdoCbl-dependent mutases belong to the SIMIBI subclass of the G3E family of P-loop metallochaperones (3,13,29). In MCM, where the role of the G-protein chaperone, MeaB, is best characterized (14), the chaperone uses GTP hydrolysis to power the expulsion of cob(II)alamin when 5'-deoxyadenosine is lost from the active site (13). Unlike the ATP-dependent reactivases, MeaB is unable to release OH_2Cbl bound to MCM (13,14). Instead, MeaB exerts a protective effect on the MCM reaction, by reducing the inactivation rate in the presence of nucleotides. In numerous bacterial genomes MeaB-like proteins are found in the same operon as the mutases (40). Mutations in the MeaB ortholog in humans results in methylmalonic aciduria, an inborn error of metabolism (41,42), pointing to the important role of this auxillary protein in maintaining MCM function.

The presence of nucleotides has virtually no effect on the inactivation kinetics of IcmF with either isobutyryl-CoA or n-butyryl-CoA as substrate (Figure 3.9 A). In contrast, when isovaleryl-CoA is employed as substrate, protection, albeit modest, was seen in the presence of GTP (Figure 3.9 B). The activity of *Gk* IcmF with isobutyryl-CoA is reduced in the presence of nucleotides (3). This is unexpected since MeaB, increases the k_{cat} of MCM 1.8-fold in addition to protecting it from inactivation (14). The rate of IcmF inactivation (0.1 ± 0.01 min^{-1}) in the presence of isobutyryl-CoA, is ~14-fold higher than inactivation of MCM in the presence of methylmalonyl-CoA (0.0072 min^{-1}) (14).

Under anaerobic conditions, the IcmF reaction with n-butyryl-CoA was linear for a longer duration than in the presence of oxygen (Figure 3.10 B). Since OH_2Cbl formation

is observed only under aerobic conditions, it argues against an internal electron transfer from cob(II)alamin to the substrate being responsible for inactivation, as described for lysine 5,6-aminomutase (26). Since IcmF is also inactivated under anaerobic conditions where OH_2Cbl is not formed, we conclude that inactivation is primarily signaled by the loss of 5´-deoxyadenosine from the active site.

In conclusion, we report that IcmF catalyzes the formation pivalyl-CoA from isovaleryl-CoA. There is virtually no information on bacterial metabolism of pivalic acid, which has a mostly anthropogenic origin (10). We suggest that pivalyl-CoA mutase activity of IcmF might be important for biodegradation of branched compounds, where pivalic acid is a central intermediate (10,43). The activity of IcmF would reduce branching of compounds with a quaternary carbon. It will be important to follow the fate of pivalyl-CoA in IcmF-containing bacteria and test whether this compound is incorporated in fatty acids and/or other compounds.

3.6 References

(1) Banerjee, R. (2003) Radical carbon skeleton rearrangements: catalysis by coenzyme B12-dependent mutases. *Chem Rev 103*, 2083-94.

(2) Ratnatilleke, A., Vrijbloed, J. W., and Robinson, J. A. (1999) Cloning and sequencing of the coenzyme B(12)-binding domain of isobutyryl-CoA mutase from Streptomyces cinnamonensis, reconstitution of mutase activity, and characterization of the recombinant enzyme produced in Escherichia coli. *J Biol Chem 274*, 31679-85.

(3) Cracan, V., Padovani, D., and Banerjee, R. (2010) IcmF is a fusion between the radical B12 enzyme isobutyryl-CoA mutase and its G-protein chaperone. *J Biol Chem 285*, 655-66.

(4) Vrijbloed, J. W., Zerbe-Burkhardt, K., Ratnatilleke, A., Grubelnik-Leiser, A., and Robinson, J. A. (1999) Insertional inactivation of methylmalonyl coenzyme A (CoA) mutase and isobutyryl-CoA mutase genes in Streptomyces cinnamonensis: influence on polyketide antibiotic biosynthesis. *J Bacteriol 181*, 5600-5.

(5) Rohwerder, T., Breuer, U., Benndorf, D., Lechner, U., and Muller, R. H. (2006) The alkyl tert-butyl ether intermediate 2-hydroxyisobutyrate is degraded via a novel cobalamin-dependent mutase pathway. *Appl Environ Microbiol 72*, 4128-35.

(6) Erb, T. J., Retey, J., Fuchs, G., and Alber, B. E. (2008) Ethylmalonyl-CoA mutase from Rhodobacter sphaeroides defines a new subclade of coenzyme B12-dependent acyl-CoA mutases. *J Biol Chem 283*, 32283-93.

(7) Mancia, F., and Evans, P. (1998) Conformational changes on substrate binding to methylmalonyl CoA mutase and new insights into the free radical mechanism. *Structure 6*, 711-720.

(8) Mancia, F., Keep, N. H., Nakagawa, A., Leadlay, P. F., McSweeney, S., Rasmussen, B., Bosecke, P., Diat, O., and Evans, P. R. (1996) How coenzyme B12 radicals are generated: the crystal structure of methylmalonyl-coenzyme A mutase at 2 A resolution. *Structure 4*, 339-50.

(9) Drennan, C. L., Huang, S., Drummond, J. T., Matthews, R. G., and Lidwig, M. L. (1994) How a protein binds B12: A 3.0 A X-ray structure of B12-binding domains of methionine synthase. *Science 266*, 1669-74.

(10) Rohwerder, and Müller, H. (2007) New bacterial cobalamin-dependent CoA-carbonyl mutases involved in degradation pathways, in *Vitamin B: new research* (CM, E., Ed.) pp 81–98, Nova Science Publishers.

(11) Gruber, K., Puffer, B., and Krautler, B. (2011) Vitamin B(12)-derivatives-enzyme cofactors and ligands of proteins and nucleic acids. *Chem Soc Rev 40*, 4346-63.

(12) Toraya, T. (2000) Radical catalysis of B12 enzymes: structure, mechanism, inactivation, and reactivation of diol and glycerol dehydratases. *Cell Mol Life Sci 57*, 106-27.

(13) Padovani, D., and Banerjee, R. (2009) A G-protein editor gates coenzyme B12 loading and is corrupted in methylmalonic aciduria. *Proc Natl Acad Sci U S A 106*, 21567-72.

(14) Padovani, D., and Banerjee, R. (2006) Assembly and protection of the radical enzyme, methylmalonyl-CoA mutase, by its chaperone. *Biochemistry 45*, 9300-6.

(15) Stols, L., Gu, M., Dieckman, L., Raffen, R., Collart, F. R., and Donnelly, M. I. (2002) A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. *Protein Expr Purif 25*, 8-15.

(16) Lanzetta, P. A., Alvarez, L. J., Reinach, P. S., and Candia, O. A. (1979) An improved assay for nanomole amounts of inorganic phosphate. *Anal Biochem 100*, 95-7.

(17) Babior, B. M., Carty, T. J., and Abeles, R. H. (1974) The mechanism of action of ethanolamine ammonia-lyase, a B12-dependent enzyme. The reversible formation of 5'-deoxyadenosine from adenosylcobalamin during the catalytic process. *J Biol Chem 249*, 1689-95.

(18) Alm, E. J., Huang, K. H., Price, M. N., Koche, R. P., Keller, K., Dubchak, I. L., and Arkin, A. P. (2005) The MicrobesOnline Web site for comparative genomics. *Genome Res 15*, 1015-22.

(19) Reddick, J. J., and Williams, J. K. (2008) The mmgA gene from Bacillus subtilis encodes a degradative acetoacetyl-CoA thiolase. *Biotechnol Lett 30*, 1045-50.

(20) Brigham, C. J., Budde, C. F., Holder, J. W., Zeng, Q., Mahan, A. E., Rha, C., and Sinskey, A. J. (2010) Elucidation of beta-oxidation pathways in Ralstonia eutropha H16 by examination of global gene expression. *J Bacteriol 192*, 5454-64.

(21) Zhang, Y. X., Denoya, C. D., Skinner, D. D., Fedechko, R. W., McArthur, H. A., Morgenstern, M. R., Davies, R. A., Lobo, S., Reynolds, K. A., and Hutchinson, C. R. (1999) Genes encoding acyl-CoA dehydrogenase (AcdH) homologues from Streptomyces coelicolor and Streptomyces avermitilis provide insights into the metabolism of small branched-chain fatty acids and macrolide antibiotic production. *Microbiology 145 (Pt 9)*, 2323-34.

(22) Matsuoka, H., Hirooka, K., and Fujita, Y. (2007) Organization and function of the YsiA regulon of Bacillus subtilis involved in fatty acid degradation. *J Biol Chem 282*, 5180-94.

(23) Rohwerder, T., and Muller, R. H. (2010) Biosynthesis of 2-hydroxyisobutyric acid (2-HIBA) from renewable carbon. *Microb Cell Fact 9*, 13.

(24) Muller, R. H., Rohwerder, T., and Harms, H. (2007) Carbon conversion efficiency and limits of productive bacterial degradation of methyl tert-butyl ether and related compounds. *Appl Environ Microbiol 73*, 1783-91.

(25) Toraya, T., Tamura, N., Watanabe, T., Yamanishi, M., Hieda, N., and Mori, K. (2008) Mechanism-based inactivation of coenzyme B12-dependent diol dehydratase by 3-unsaturated 1,2-diols and thioglycerol. *J Biochem 144*, 437-46.

(26) Tang, K. H., Chang, C. H., and Frey, P. A. (2001) Electron transfer in the substrate-dependent suicide inactivation of lysine 5,6-aminomutase. *Biochemistry 40*, 5190-9.

(27) Maiti, N., Widjaja, L., and Banerjee, R. (1999) Proton transfer from histidine 244 may facilitate the 1,2 rearrangement reaction in coenzyme B(12)-dependent methylmalonyl-CoA mutase. *J Biol Chem 274*, 32733-7.

(28) Gaal, T., Bartlett, M. S., Ross, W., Turnbough, C. L., Jr., and Gourse, R. L. (1997) Transcription regulation by initiating NTP concentration: rRNA synthesis in bacteria. *Science 278*, 2092-7.

(29) Leipe, D. D., Wolf, Y. I., Koonin, E. V., and Aravind, L. (2002) Classification and evolution of P-loop GTPases and related ATPases. *J Mol Biol 317*, 41-72.
(30) Pandit, S. B., and Srinivasan, N. (2003) Survey for g-proteins in the prokaryotic genomes: prediction of functional roles based on classification. *Proteins 52*, 585-97.
(31) Probian, C., Wulfing, A., and Harder, J. (2003) Anaerobic mineralization of quaternary carbon atoms: isolation of denitrifying bacteria on pivalic acid (2,2-dimethylpropionic acid). *Appl Environ Microbiol 69*, 1866-70.
(32) Rezanka, T., Siristova, L., Schreiberova, O., Rezanka, M., Masak, J., Melzoch, K., and Sigler, K. (2011) Pivalic acid acts as a starter unit in a fatty acid and antibiotic biosynthetic pathway in Alicyclobacillus, Rhodococcus and Streptomyces. *Environ Microbiol 13*, 1577-89.
(33) Wittinghofer, A., and Vetter, I. R. (2011) Structure-function relationships of the G domain, a canonical switch motif. *Annu Rev Biochem 80*, 943-71.
(34) Soundararajan, M., Yang, X., Elkins, J. M., Sobott, F., and Doyle, D. A. (2007) The centaurin gamma-1 GTPase-like domain functions as an NTPase. *Biochem J 401*, 679-88.
(35) Koller-Eichhorn, R., Marquardt, T., Gail, R., Wittinghofer, A., Kostrewa, D., Kutay, U., and Kambach, C. (2007) Human OLA1 defines an ATPase subfamily in the Obg family of GTP-binding proteins. *J Biol Chem 282*, 19928-37.
(36) Kajiura, H., Mori, K., Shibata, N., and Toraya, T. (2007) Molecular basis for specificities of reactivating factors for adenosylcobalamin-dependent diol and glycerol dehydratases. *Febs J 274*, 5556-66.
(37) Mori, K., Bando, R., Hieda, N., and Toraya, T. (2004) Identification of a reactivating factor for adenosylcobalamin-dependent ethanolamine ammonia lyase. *J Bacteriol 186*, 6845-54.
(38) Zelder, O., Beatrix, B., Leutbecher, U., and Buckel, W. (1994) Characterization of the coenzyme-B12-dependent glutamate mutase from Clostridium cochlearium produced in Escherichia coli. *Eur J Biochem 226*, 577-85.
(39) Chang, C. H., and Frey, P. A. (2000) Cloning, sequencing, heterologous expression, purification, and characterization of adenosylcobalamin-dependent D-lysine 5, 6-aminomutase from Clostridium sticklandii. *J Biol Chem 275*, 106-14.
(40) Dobson, C. M., Wai, T., Leclerc, D., Wilson, A., Wu, X., Dore, C., Hudson, T., Rosenblatt, D. S., and Gravel, R. A. (2002) Identification of the gene responsible for the cblA complementation group of vitamin B12-responsive methylmalonic acidemia based on analysis of prokaryotic gene arrangements. *Proc Natl Acad Sci U S A 99*, 15554-9.
(41) Hubbard, P. A., Padovani, D., Labunska, T., Mahlstedt, S. A., Banerjee, R., and Drennan, C. L. (2007) Crystal structure and mutagenesis of the metallochaperone MeaB: insight into the causes of methylmalonic aciduria. *J Biol Chem 282*, 31308-16.
(42) Lerner-Ellis, J. P., Dobson, C. M., Wai, T., Watkins, D., Tirone, J. C., Leclerc, D., Dore, C., Lepage, P., Gravel, R. A., and Rosenblatt, D. S. (2004) Mutations in the MMAA gene in patients with the cblA disorder of vitamin B12 metabolism. *Hum Mutat 24*, 509-16.

(43) Solano-Serena, F., Marchal, R., Heiss, S., and Vandecasteele, J. P. (2004) Degradation of isooctane by Mycobacterium austroafricanum IFP 2173: growth and catabolic pathway. *J Appl Microbiol 97*, 629-39.

CHAPTER 4

Adenosyltransferase Synthesizes and Delivers Coenzyme B_{12} to IcmF: Insights into the Function of the G-protein Domain

4.1 Introduction

Cobalamin is an important cofactor in bacteria, Archaea and eukaryotes. Since cobalamin is biosynthesized only by some bacteria and Archaea, organisms that require this cofactor must rely on B_{12} uptake from the environment or their diet (*1-3*). AdoCbl or coenzyme B_{12}, is a derivative of cobalamin in which a 5'-deoxyadenosyl ligand is bound at the "upper" axial coordination site of the corrin ring. Six carbon skeleton isomerases are known to use AdoCbl as a cofactor (4-6). Most of AdoCbl-dependent carbon skeleton isomerases are disproportionally distributed in bacteria and Archaea. The only AdoCbl-dependent enzyme which distribution ranges from bacteria to mammals is methylmalonyl-CoA mutase (MCM). MCM catalyzes the interconversion of (2R)-methylmalonyl-CoA and succinyl-CoA (6).

ATR or ATP:Cob(I)alamin Adenosyltransferase (ATR) is needed for formation of the AdoCbl form of the cofactor. ATR catalyses the transfer of a 5'-deoxyadenosyl moiety from ATP to the cobalt atom of cob(I)alamin to form AdoCbl. Notably, this reaction is important not only in the *de novo* biosynthesis of AdoCbl in bacteria but also in the assimilation of vitamin B_{12} precursors (Figure 4.1 A).

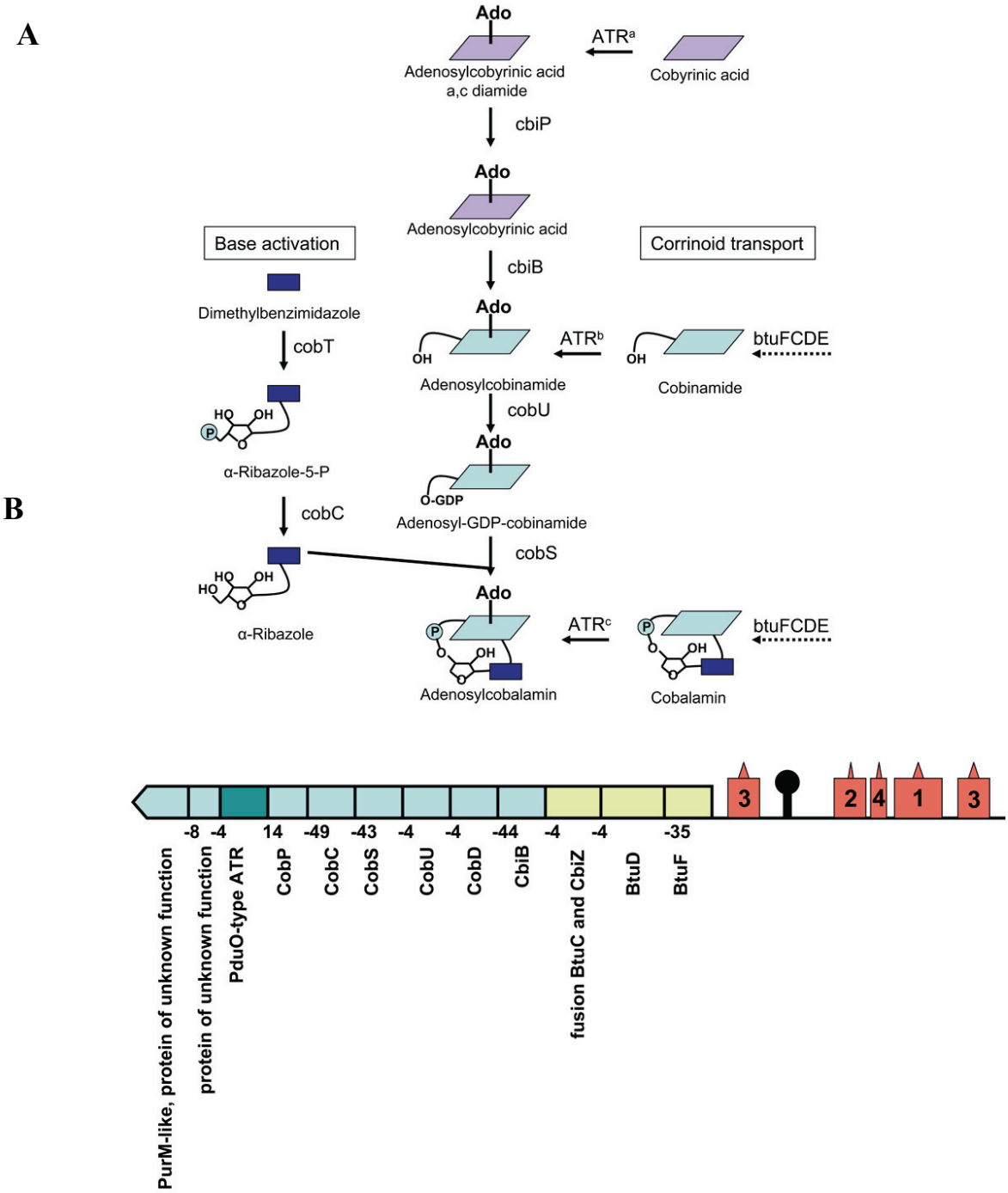


Figure 4.1: Cobalamin scavenging. (**A**) Schematic representation of reactions catalyzed by enzymes involved in cobalamin scavenging. [a,b]-reaction is catalyzed by CobA-like ATRs. [b,c]- reaction is catalyzed by PduO-like and EutT-like ATRs. Cobyrinic acid and its derivatives, which are part of *de novo* cobalamin synthesis are designated by purple boxes. Cobinamide and cobalamin and their derivatives, which can be imported in the

cell by the BtuCD-BtuF transport system are designated as light blue boxes. Dimethylbenzimidazole (DMB) is designated as a dark blue box and 5'-deoxyadenosyl as Ado. (**B**) Organization of genes in the operon harboring the *atr* gene in *G.kaustophilus.* Proteins and accession numbers: PurM-like, protein of unknown function which contains ATP-binding site (YP_148108), protein of unknown function (YP_148109), PduO-type ATR (YP_148110), CobP (YP_148111), CobC (YP_148112), CobS (YP_148113), CobU (YP_148114), CobD (YP_148115), CbiB (YP_148116), fusion between BtuC and CbiZ(YP_148117), BtuD (YP_148118), BtuF (YP_148119). BtuCD-BtuF is a putative cobalamin transporter. The numbers below the genes indicate the distance in nucleotides between two adjacent genes. Negative numbers indicate overlapping genes. Genes and regulatory elements are not drawn to scale. Immediately upstream of BtuF, four motifs that comprise the cobalamin riboswitch (red boxes) can be seen, together with a transcriptional attenuator (black lollipop).

There are three families of ATRs that are completely unrelated in their amino acid sequence: CobA-type, PduO-type, and EutT-type (7-10). On the surface, CobA (in some organisms designated as BtuR or CobO) is involved in *de novo* cobalamin biosynthesis where in both aerobic and anaerobic pathways it catalyzes preferentially the adenosylation of cobyrinic acid a,c-diamide (Figure 4.1 A).

In contrast, PduO-like and EutT-like ATRs are associated with assimilation of corrinoids and adenosylate both cobinamide (Cbi) and cobalamin (Cbl) (7,11) (Figure 4.1 A). Similarly, the PduO-type ATR from *Lactobacillus reuteri* adenosylates both Cbi (k_{cat}= (2.4 ± 0.1) x 10^{-2} min^{-1})) and Cbl (k_{cat} = (2.0 ± 0.2) x 10^{-2} min^{-1})) (Figure 4.1 A) (11).

Since the K_M values for both substrates are very similar K_{Cbl}= (13 ± 1) x 10^{-2} μM and K_{Cbi}= (9.6 ± 1.4) x 10^{-2} μM, the catalytic efficiencies are comparable. EutT-like ATR from *Salmonella enterica* adenosylates Cbi with only 21% of the activity reported for Cbl (12).

As noted by Bobik and colleagues, in bacterial operons, PduO-type ATRs are often associated with AdoCbl-dependent mutases (7). ATR binds AdoCbl in a “base-off” state where the lower axial ligand, dimethylbenzimidazole (DMB) is swung away from the

cobalt (Figure 4.2 A) (5). On the other hand, MCM and other similar mutases bind cofactor in "base-off/His-on" state, where a histidine in the canonical DxHxxG motif, coordinates the cobalt atom (Figure 4.2 B) (13). The "base-off" mode of binding of AdoCbl in ATR, an enzyme which synthesizes the cofactor, is mirrored in the mutase, which utilizes the cofactor and is unlikely to be coincidental. It was therefore proposed that this similarity in "base-off" binding mode but the difference in the coordination environment (i.e. 5-coordinate in ATR versus 6-coordinate in MCM) is exploited for direct delivery of AdoCbl from ATR to MCM (14,15).

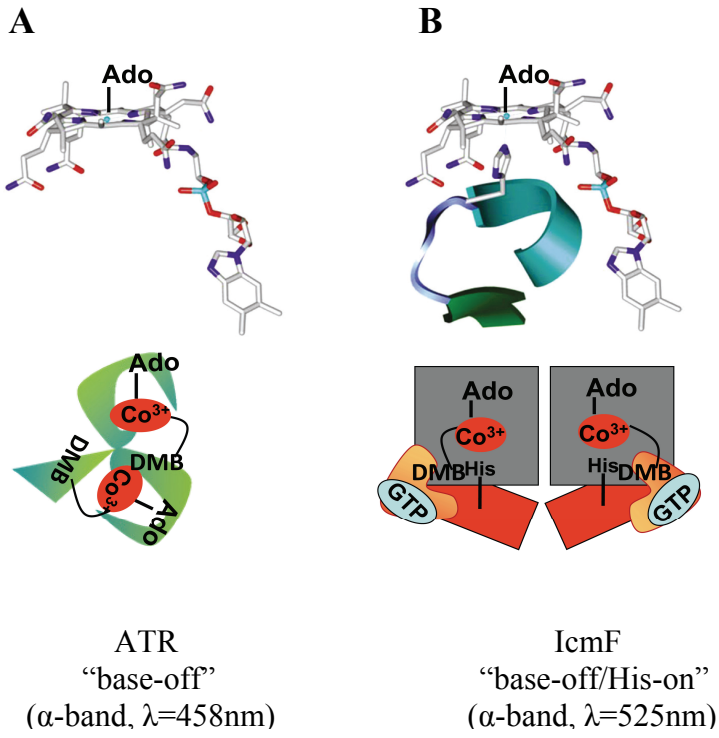


Figure 4.2: Schematic representation of two modes of AdoCbl binding to proteins. ATR is a trimer which binds AdoCbl in the "base-off" state (**A**), IcmF is a homodimer which binds AdoCbl in a "base-off/His-on" state (**B**) where a histidine residue from the DXHXXG motif replaces the endogenous DMB ligand. Distinguishing spectral features are indicated.

Indeed, it was recently shown that the PduO-type ATR from *Methylobacterium extorquens* not only synthesizes AdoCbl, but also delivers the active form of the cofactor to its target enzyme, MCM, without releasing it into solution (16).

Importantly, direct transfer of AdoCbl from ATR to MCM in which the histidine in the DxHxxG motif is mutated to alanine, was impaired (16). This result is consistent with the proposed role of the histidine residue in cofactor transfer to the mutase active site (14,15). Isobutyryl-CoA mutase (ICM), another AdoCbl-dependent enzyme is very similar to MCM both at the primary sequence level and catalytic mechanism (5,17-19).The recently discovered AdoCbl-dependent enzyme, IcmF is a fusion between the two subunits of ICM and its G-protein chaperone (20). IcmF like MCM, binds AdoCbl in "base-off/His-on" state (Figure 4.2 B). Since many IcmF-containing bacteria also contain a copy of MCM, it raised the intriguing question as to whether the PduO-type ATR delivers AdoCbl not only to MCM but also to IcmF.

MeaB, a G-protein chaperone for MCM, was shown to play a key role in ensuring the fidelity of the cofactor-docking process. It was demonstrated that ATR can transfer cob(II)alamin, the cofactor precursor, to MCM (21). Subsequent oxidation of cob(II)alamin in the MCM active site leads to formation of aquacobalamin (OH_2Cbl) (22). Remarkably, complexation of MeaB and MCM allows discrimination between cob(II)alamin and AdoCbl and permits transfer only of the active form of the cofactor (21).

Since IcmF is a fusion between the two subunits of ICM and MeaI, a G-protein similar to MeaB, we were interested in the investigating the possible gating function of the MeaI domain. Given that ATR can catalyze adenosylation of both Cbi and Cbl, it is also

possible that AdoCbi or other incomplete cofactors can be transferred to the target mutase. Previously, it was demonstrated that MCM is completely inactive when AdoCbi is used as a cofactor and k_{cat} for the enzyme is reduced 4-fold when AdoCbi-GDP is used instead of AdoCbl (23). Thus, we hypothesize that G-protein chaperone protects IcmF from binding not only of cob(II)alamin but other incomplete forms of the cofactor. In this study, we show that ATR from *G.kaustophilus* directly transfers AdoCbl to IcmF. We also provide evidence that the MeaI domain possesses a "gating" function, which is similar to the corresponding function described for MeaB (21).

4.2 Experimental procedures

4.2.1 DNA manipulations. *(i) Cloning of ATR from G.kaustophilus.* ATR was amplified from the genomic DNA obtained from the Bacillus Genetic Stock Center, The Ohio State University (Columbus, OH) using the following primers for LIC cloning in pMCSG7 vector (forward 5'-TACTTCCAATCCAATGCCGTGAAATTGTATACGCGAACAGG -3' and the reverse 5'- TTATCCACTTCCAATGCTATCATTGTTTCTCCTCCTTGT CGCGG -3'). *(ii) Cloning of ATR from Cupriavidus metallidurans.* ATR was amplified from the genomic DNA obtained from the ATCC biological resource center using the following primers for LIC cloning in pMCSG7 vector (forward 5'- TACTTCCAATCC AATGCCATGGGTAATCGCCTGTCCAAAATTGC -3' and the reverse 5'- TTATCC ACTTCCAATGCTATCAGGACTCCCGCTCACGCTGCCAG -3'). *(iii) Cloning of the N-terminal part of IcmF*. The N-terminal part of IcmF was amplified from vector pET30 Ek/LIC containing *icmF* from *G.kaustophilus* using the following primers for LIC cloning in pMCSG7 vector (forward 5′-TACTTCCAATCCAATGCCATGGCGCAC ATTTACCGTCC-3′ and reverse 5'-TTATCCACTTCCAATGCTACTAACCGGACA ACGTT-3′).*(iv) Cloning of the C-terminal part of IcmF*. The C-terminal part of IcmF was amplified from vector pET30 Ek/LIC containing *icmF* from *G.kaustophilus* using the following primers for LIC cloning in pMCSG7 vector (forward 5' - TACTTCCAATCC AATGCCACCGTGACGGTGGTGACCGATG -3' and reverse 5'- TTATCCACTTCC AATGCTACTAACCGGACAACGTT-3′).

4.2.2 Enzyme expression and purification. *(i) Recombinant IcmFs from G.kaustophilus (Gk) and Cupriavidus metallidurans (Cm) expression and purification.* Both proteins were expressed in *Escherichia coli* BL21 (DE3) cells (Invitrogen) and were purified as

previously described (20). *(ii) Gk* ATR *expression and purification.* The pMCSG7 vector with the *Gk* ATR gene was transformed into *E. coli* BL21 (DE3) cells (Invitrogen), which were grown at 37˚C in Luria Bertani (LB) medium containing 100 μg/ml ampicilin to an absorbance at 600 nm of 0.5-0.6. Cells were grown for 12h after induction with 0.5mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) at 15˚C. *E.coli* cells (~25 g wet weight obtained from 6 L of culture) were lysed in lysis buffer (50 mM NaPi, pH 8.0, 500 mM NaCl, 30 mM Imidazole) supplemented with protease inhibitor cocktail (Roche Applied Science), lysozyme (Sigma) and benzonase (Novagen) using a sonicator (Ultrasonic XL Misonix). The resulted cell homogenate after centrifugation was subjected to dilution to a final concentration of 3-5 mg/ml and loaded onto a 30 ml Ni-Sepharose column (GE Life Sciences). After washing with 10-20 column volumes of lysis buffer, the protein was eluted with a gradient of 30 to 300 mM imidazole in 50 mM NaPi, pH 8.0, 500 mM NaCl buffer. Fractions containing ATR were pooled, concentrated and applied to a 160 ml Superdex 200 column connected to the AKTApurifier and equilibrated with 50 mM HEPES pH 7.5, 100 mM NaCl buffer (buffer A). Under these conditions, ATR eluted with a retention volume of ~114 ml, which corresponds to an apparent molecular weight of 75 kDa. This is in good agreement with the expected molecular weight of a trimer of ATR. Fractions containing ATR were pooled, concentrated, flash-frozen in liquid nitrogen and stored at -80°C until further use. Approximately 80-100 mg of recombinant ATR was obtained from a 6 L culture. *(iii) Cm ATR expression and purification. Cm* ATR was expressed and purified similarly to *Gk* ATR. *(iv) The N-terminal part and the C-terminal parts of Gk IcmF expression and purification*. Both truncated proteins were overexpressed in *E. coli* BL21 (DE3) cells (Invitrogen), which were grown at 37˚C in LB

medium containing 100 μg/ml ampicilin to an absorbance at 600 nm of 0.5-0.6. Cells were usually grown for 10-14 h after induction with 0.5 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) at 15˚C. Proteins were purified similarly to *Gk* full length IcmF where gel-filtration was performed in buffer A (20).

4.2.3 Determination of molecular weights by gel-filtration. To determine the oligomeric state of the N-terminal part of IcmF, the C-terminal part of IcmF and *Gk* ATR in solution, protein samples were loaded on a 160 ml Superdex 200 column in buffer A at a flow rate of 0.75 ml/min. Prior to loading, the protein sample was filtered through a 0.2 μm syringe filter (Fisherbrand). The column was calibrated using gel-filtration standards from Sigma. The molecular weights of the standards were: β-amylase (200 kDa), alcohol dehydrogenase (150 kDa), albumin (66 kDa), carbonic anhydrase (29 kDa) and cytochrome c (12.4 kDa).

4.2.4 UV-Visible Spectroscopy. *(i) AdoCbl binding to ATR.* AdoCbl solution was titrated with ATR to determine cofactor binding. Apo-ATR (0-80 μM) was added to a solution of AdoCbl (50-60 μM) in 50 mM NaPi pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$ at 25 °C. After each addition of ATR solution was allowed to equilibrate for 2-3 min. The binding of the cofactor to ATR was readily visible by the increase at 458 nm and decrease at 525 nm, which corresponds to "base-on" to "base-off" shift (Figure 4.2). *(ii) Effect of ATP on holo-ATR.* A solution of ATP (0.125-8.6 mM) was added to 8-30 μM ATR reconstituted with two equivalents of AdoCbl (holo-ATR) in 50 mM NaPi pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$ at 25 °C. The data were acquired at 2-3 min after mixing, and the concentration of AdoCbl released after each ATP additions was estimated from the increase in absorbance at 525nm, using $\Delta\varepsilon_{525nm}$=6.69 mM^{-1} cm^{-1} (24). Spectra of the

filtrate obtained by centrifuging the mixture using a Microcon 30K concentrator (Millipore) were obtained to confirm the release of AdoCbl. *(iii) AdoCbl transfer between ATR and IcmF.* The transfer of AdoCbl between the two proteins was monitored by UV-Visible spectroscopy at 25 °C in buffer A with 10mM $MgCl_2$ as it was previously described for MCM and ATR (16). Briefly, for the forward transfer, 4-8 μM holo-ATR (8-16 μM bound AdoCbl) was added to a final concentration of 2-200 μM apo-IcmF. For the reverse transfer, 30-40 μM holo-IcmF (30-80 μM bound AdoCbl) was added to a final concentration of 2-210 μM apo-ATR. In all experiments, the solution containing IcmF was blanked first, before cofactor or ATR was added. Each reaction was given 3 min to equilibrate and absorption spectra were recorded. Holo-ATR concentration was measured using the extinction coefficient $\Delta\varepsilon_{458nm}$= 8 mM^{-1} cm^{-1}. The amount of AdoCbl transferred was estimated from the increase (for the forward transfer) or decrease (for the reverse transfer) in absorbance at 525 nm ($\Delta\varepsilon_{525nm}$ = 7.75 mM^{-1} cm^{-1}) (16). Nucleotides (GDP/GTP/GMPPNP/ATP or AMPPNP) were added usually to a final concentration of 2-3 mM. *(iv) AdoCbl transfer between ATR and the N-terminal part of IcmF.* The transfer of AdoCbl between two proteins was monitored by UV-Visible spectroscopy at 25 °C in buffer A supplemented with 10 mM $MgCl_2$ similarly to transfer experiments between ATR and full length IcmF. Briefly, for the forward transfer, 50-60 μM apo-N-terminal part of IcmF was blanked first before holo-ATR (10-15 μM AdoCbl bound) was added ± 2 mM ATP. In the reverse transfer experiment, apo-ATR was added to 20-30 μM holo-N-terminal part of IcmF (one equivalent of AdoCbl bound) ± 2mM GMPPNP.

4.2.5 ATR assay. The specific activity of ATR activity was monitored as described previously under saturating concentrations of hydroxycobalamin (OHCbl) and ATP (7).

4.2.6 Determination of the NTPase activity. The ATPase and GTPase activities of IcmF were determined using HPLC and Malachite-Green assays as previously described (20) (See also section 2.3.3).

4.2.7 Fluorescence Stopped-flow Spectroscopy. N-methylanthraniloyl (mant)-labeled nucleotides are widely used as probes for stopped-flow analysis of nucleotide-protein interactions (25). The mant moiety is an environmentally-sensitive fluorophore whose fluorescence intensity increases with increasing hydrophobicity of its environment (25-27). Our rapid reaction kinetics studies were performed using a Hi-Tech Scientific SF-61DX2 stopped-flow spectrophotometer. All solutions were filtered through a 0.2 μm syringe filter (Fisherbrand) and transferred to loading syringes. All experiments were performed at 22 ºC. The excitation wavelength for mant-GDP was 355 nm and the emission was recorded using a 408 nm cut-off-filter. Binding experiments were performed in 50 mM NaPi pH 7.5, 150 mM NaCl supplemented with 5 mM $MgCl_2$ with a final concentration of 0.25 μM mant-GDP and increasing concentrations of IcmF (0.6-5.8 μM) (after mixing). Kinetic traces of binding of mant-GDP to IcmF were fit to a two phase exponential function, where the observed rate constants from each phase (k_{on1}, k_{on2}) had a liner dependence on protein concentration. The dissociation rates were obtained from a ligand displacement experiment. Addition of a large excess of GDP (200-350 μM) to a pre-formed complex of 0.25 μM mant-GDP•5μM IcmF, led to a decrease in signal intensity that was fit to a biphasic exponential function to obtain the dissociation rate constants, $k_{off1\_dis.}$ and $k_{off2\_dis.}$ for mant-GDP release from the two binding sites of IcmF.

4.2.8 Isothermal Titration Calorimetry. *(i)* Nucleotide binding was determined by MicroCal isothermal titration calorimeter. ITC experiments were performed as described

previously (20,28). IcmF was dialyzed for 10–12 h against 50 mM HEPES pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$ and 1–2 mM TCEP (Buffer B) before use. *Gk* apo-IcmF (18-25 μM) in Buffer B was titrated with 30–42 7–9.7-μl injections of a 10-15 molar excess of GDP/GMPPNP at 20 °C. The calorimetric signals were integrated, and the data were analyzed with Microcal ORIGIN software using a single-site binding model to obtain the thermodynamic parameters associated with binding of nucleotides to IcmF. *(ii)* The N-terminal part of IcmF was dialyzed for 10-12 h against buffer B before use. The protein (25-30 μM) in Buffer B was titrated with 35-38 8 μL injections of a 10-15 molar excess of AdoCbl at 20°C. The calorimetric signals were integrated and the data were analyzed with Microcal ORIGIN software using a two-sites binding model to determine the thermodynamic parameters associated with AdoCbl binding to protein.

4.2.9 Bioinformatics analysis. Operon and regulon browsers on the Microbes Online web site were used for the elucidation of functional predictions for the genes of interest (29) (Section 2.3.8). RibEx (riboswitch explorer) server was used for searching for potential cobalamin riboswitches (30).

4.3 Results

4.3.1 PduO-type ATR gene from *Geobacilus kaustophilus.* In the genome of *G.kaustophilus* the gene encoding the PduO-type ATR is localized within an operon harboring 8 genes that are involved in the late stages of cobalamin biosynthesis (CobP, CobC, CobS, CobU, CobD, CbiB and others) and 3 genes which encode the BtuCD-BtuF cobalamin transporter system (Figure 4.1B). In the genome of *G. kaustophilus,* this cobalamin transporter is missannotated as a three subunit ATP-binding cassette transporter (fepBCD) which allows enterobactin, a high affinity iron siderophore, to cross the inner membrane (31-33). Interestingly in this operon, BtuC is fused with CbiZ (Figure 4.1B). It was recognized previously that CbiZ is fused with BtuD in *Bacillus halodurans* and *Bacillus subtilis*. However, the exact role of this fusion remains obscure (34).

Additionally, we have identified a cobalamin riboswitch (35,36) immediately upstream of the ORF encoding for BtuF (Figure 4.1B). Not all bacteria possess the complete set of enzymes that are necessary for *de novo* synthesis of cobalamin. Instead, they assimilate cobalamin into the biologically active cofactor forms. This process is called "cobalamin salvaging" (Figure 4.1A). Clustering of genes encoding enzymes involved in cobalamin salvage and transport is not surprising. It is interesting that a PduO-type and not a CobA-type ATR is a part of the above mentioned operon since it was shown in *Ralstonia metallidurans* (37), *Methylobium petroleophilum* (38) and other bacteria (38) CobA-type ATRs are found in similar operons.

4.3.2 Properties of recombinant *Gk* ATR. The recombinant *Gk* ATR was purified using a combination of affinity and size-exclusion chromatography as described under

Experimental Procedures. As seen by SDS-PAGE analysis, the gel-filtration step separates an active trimer of ATR (elution volume ~ 114 ml corresponding to an apparent molecular mass of 75 kDa) from an aggregated protein, which elutes in the void volume of the column (~71 ml) (Figure 4.3A). The expected molecular mass of the *Gk* ATR trimer of is 72 kDa.

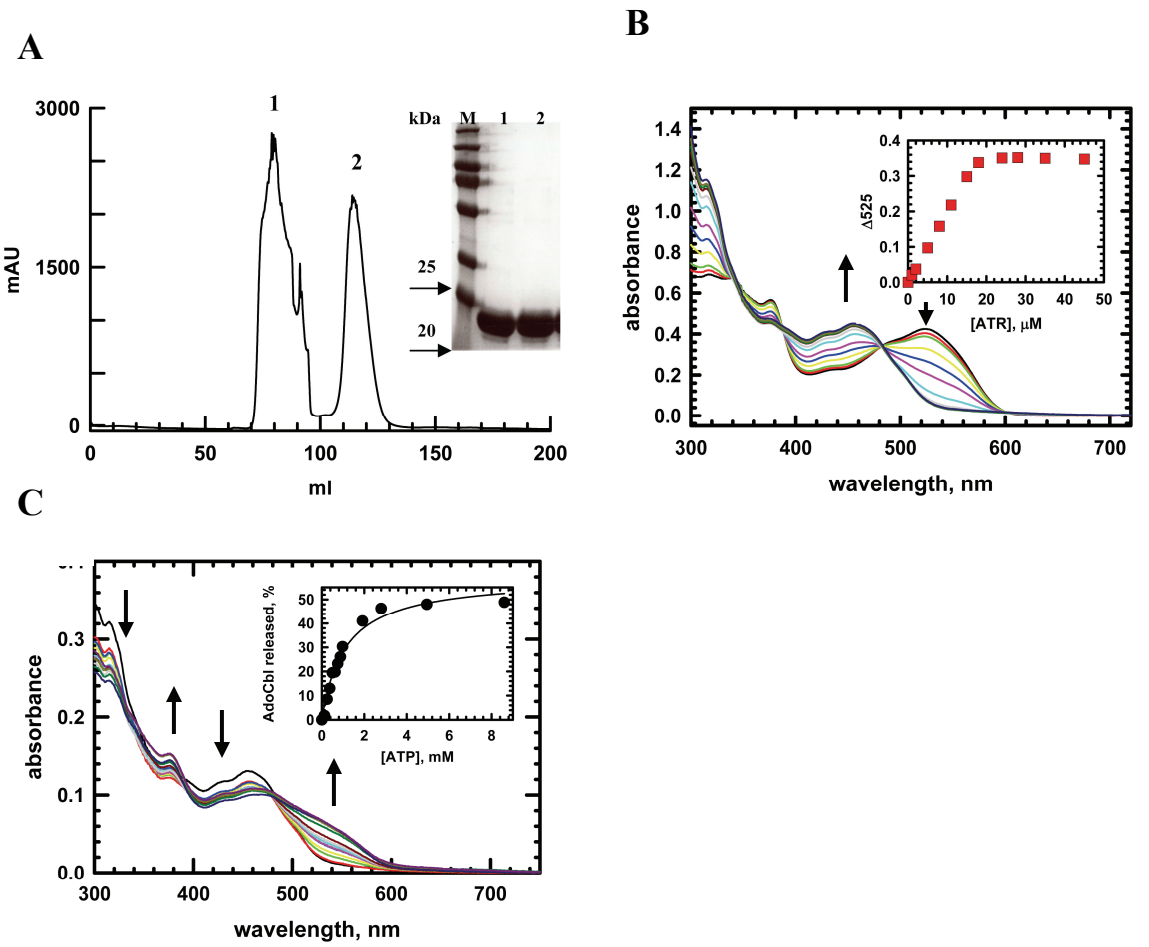


Figure 4.3: Properties of ATR from *G.kaustophilus.* (**A**) Gel-filtration chromatography of ATR. Inset: The SDS-PAGE analysis of both peaks. (**B**) ATR binds two equivalents of AdoCbl per trimer. Apo-ATR (0-45 μM) was added to 53 μM AdoCbl in 50 mM NaPi pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$ and transition from "base-on" to "base-off" state was followed. Inset: Dependence of decrease at 525 nm on ATR concentration. (**C**) Addition of ATP (0.125-8.6 mM) to holo-ATR (15 μM bound AdoCbl) led to the partial conversion of AdoCbl from the "base-off" to "base-on" state. Inset: Dependence of AdoCbl release on ATP concentration. Only one of the two equivalents of AdoCbl bound per ATR trimer is released into solution.

The specific activity of *Gk* ATR was 0.38 μmol min^{-1} mg^{-1}, which is very similar to the value reported for ATR from *M. extorquens* (0.74 ± 0.01 μmol min^{-1} mg^{-1}), when both activities measured at 22 ºC (39).

ATR binds AdoCbl in a "base-off" state (Figure 4.2A and Figure 4.3B). In solution at physiological pH, AdoCbl exists in the "base-on" state. Upon addition of increasing concentrations of *Gk* ATR to a solution of AdoCbl, apo-ATR is converted from the "base-on" to "base-off" state as evidenced by its characteristic decrease in absorbance at 525 nm and the concomitant increase at 458 nm (Figure 4.2A and Figure 4.3B).

4.3.3 Effect of ATP on holo-ATR. Partial conversion of the cofactor from the "base-off" to "base-on" state is seen when holo-ATR reconstituted with two equivalents of AdoCbl is mixed with ATP (Figure 4.3C). The "base-on" species represent free AdoCbl released into solution as judged by analysis of the filtrate following concentration of the reaction mixture in a Microcon 30K concentrator (Millipore). The UV-visible spectrum of the filtrate corresponds to ~50% of AdoCbl originally bound to ATR being released into solution (Figure 4.3C). This result is consistent with the rotary mechanism proposed for ATR from *M. extorquens*, where at any given time, only two of the three active sites can be occupied by AdoCbl. In the proposed rotary mechanism, binding of ATP to the vacant active site induces the release of one equivalent of AdoCbl (24) and is proposed as a mechanism for powering transfer of the cofactor from the active site of ATR to MCM.

4.3.4 Transfer of AdoCbl between ATR and IcmF. Reversible transfer of AdoCbl between the active sites of ATR and MCM from *M. extorquens* has been demonstrated (16). We examined whether ATR can also deliver AdoCbl to IcmF. Indeed, when holo-ATR was mixed with increasing concentrations of apo-IcmF, the characteristic spectral

change for conversion of AdoCbl from the "base-off " to the "base-off/His-on" state was observed with isosbestic points at 340, 390 and 485 nm (Figure 4.4A). Addition of either substrate, isobutyryl-CoA or n-butyryl-CoA, to the transfer mixture led to a formation of cob(II)alamin indicating that IcmF was productively loaded with AdoCbl (data not shown) (See also section 3.4.3). The possible release of AdoCbl into solution during the transfer between ATR and IcmF was monitored by analysis of the filtrate obtained from concentrating the reaction mixture using a Microcon 30K concentrator (Millipore). Free AdoCbl was not detected in the filtrate in the absence of nucleotides, indicating a direct transfer.

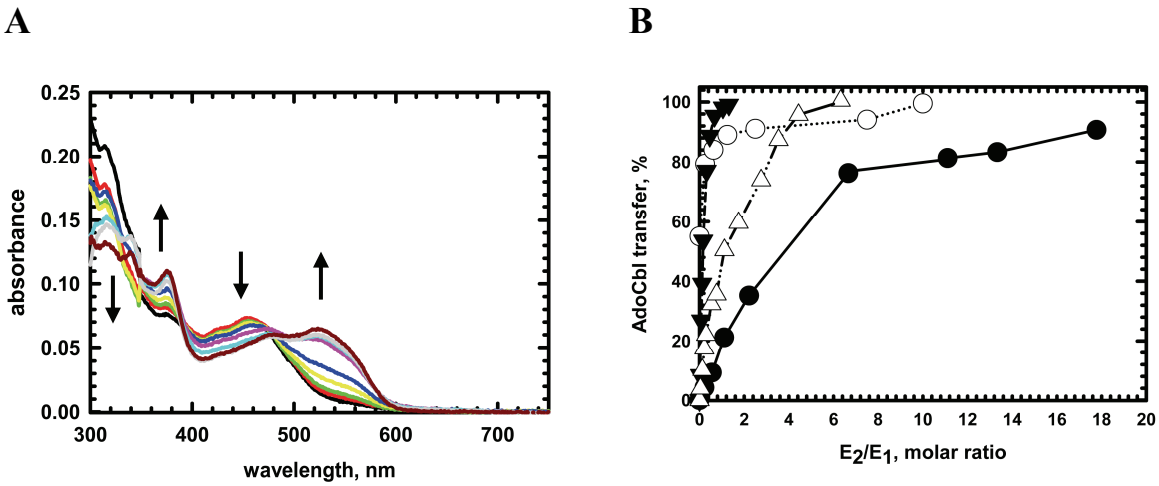


Figure 4.4: Transfer of AdoCbl between ATR and IcmF. (**A**) Dependence of AdoCbl transfer from holo-ATR (9 µM bound AdoCbl) on the concentration of apo-IcmF (2–160 µM) in buffer A supplemented with 5 mM $MgCl_2$. at 24 °C. (**B**) Dependence of cofactor transfer efficiency between holo-ATR and apo-IcmF (filled circles), holo-ATR and apo-IcmF with ATP (filled triangles), holo-IcmF and apo-ATR (open circles) and holo-IcmF and apo-ATR with ATP (open triangles). E_1 and E_2 are defined as the enzymes donating and accepting the cofactor, respectively.

Organizationally, IcmF is different from heterodimeric MCM from *M. extorquens* since IcmF contains two AdoCbl-binding sites (20). Nevertheless, in the absence of nucleotides almost complete cofactor transfer was achieved only at a very high excess of the

acceptor, apo-IcmF, over the donor, holo-ATR (Figure 4.4B). Under these conditions, an IcmF:ATR ratio of 6:1 or higher is needed to achieve ~80% transfer from ATR (Figure 4. 4B).

The absence of the complete transfer is explained by the reversibility of the process and the reverse transfer being favored at high IcmF concentrations. Remarkably, the reverse transfer from holo-IcmF to apo-ATR is favored over the forward transfer and ~90% transfer is achieved when the ATR:IcmF ratio is 1:1. These results are very similar to the equilibrium for AdoCbl transfer between ATR and MCM favoring the reverse transfer in the absence of nucleotides (16).

Cm holo-ATR also directly transfers AdoCbl to *Cm* apo-IcmF as demonstrated by the characteristic spectral shift from the "base-off" to "base-off/His-on" form when both proteins are mixed (data not shown).

4.3.5 The effect of ATP on cofactor transfer. Next, we studied how ATP affects the AdoCbl transfer reaction. When the reaction was supplemented with 3 mM ATP, the forward transfer was favored over the reverse (Figure 4.4B), even when the donor:acceptor ratio was 1:1 (Figure 4.4B). In contrast, the reverse transfer goes to completion in the presence of ATP at a donor:acceptor ratio of 5:1 (where IcmF is loaded with 2 equivalents of AdoCbl). Thus, in the presence of ATP, the equilibrium favors the forward transfer, i.e. from ATR to IcmF. These data suggest that under physiological conditions i.e. in the presence of high concentrations of ATP, the cofactor is transferred from ATR to IcmF.

4.3.6 The effect of GTP and GMPPNP on cofactor transfer. To establish how the MeaI domain of IcmF affects transfer, we performed the transfer experiments in the

presence of GTP and GMPPNP (a non-hydrolyzable analog of GTP). Remarkably, when holo-ATR was added to apo-IcmF (donor:acceptor ratio of 7:1) preincubated with GMPPNP, no cofactor transfer was detected (Figure 4.5). When the same experiment was repeated in the presence of GTP, the absorption spectrum showed a partial (~45%) transfer of AdoCbl immediately after addition of holo-ATR to IcmF•GTP (Figure 4.5, brown spectrum). After 10 min, almost complete transfer was seen (Figure 4.5, blue line). Taking into account that the experiment had 80 μM IcmF and that the k_{cat} for GTP hydrolysis by the MeaI domain is ≈ 6 ± 0.7 min^{-1} at 22 ˚C, 14% of GTP was converted to GDP in one min. Subsequently all GTP was hydrolyzed to GDP within 8 min. Since no gating of transfer was seen with GDP (data not shown) it explains why full transfer was observed after 10 min (Figure 4.5, blue spectrum).

Interestingly, the reverse transfer from holo-IcmF to apo-ATR was virtually unaffected by the presence of GMPPNP (Figure 4.6), and >90% of transfer was observed ± GMPPNP.

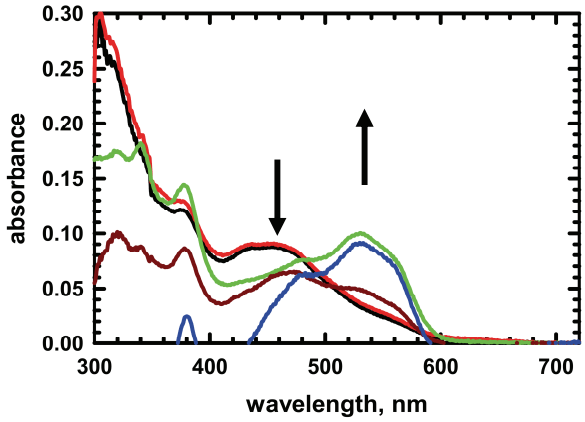


Figure 4.5: Gating by MeaI of AdoCbl transfer from ATR to IcmF. Addition of holo-ATR (11 μM) to apo-IcmF (80 μM) in buffer A supplemented with 10 mM $MgCl_2$ at 24 °C without nucleotides (green line), with 3 mM GMPPNP (red line), with 3 mM GTP immediately after addition (brown line) and 10 min after addition (blue line) of nucleotide. Holo-ATR (11 μM) alone (black line). The negative drift in the spectra (brown and blue lines) is always seen when GTP is present in the reaction mixture.

Since ATP is also a substrate for the MeaI domain of the *Gk* IcmF (Table 4.3), we studied the forward transfer in the presence of AMPPNP (a non-hydrolyzable ATP analog). We observed that cofactor transfer was not inhibited in the presence of AMPPNP. Essentially there was no difference between transfer with ATP or AMPPNP i.e. 100% of forward transfer was observed. AMPPNP, like ATP, triggers release of one equivalent of AdoCbl from fully loaded holo-ATR (data not shown).

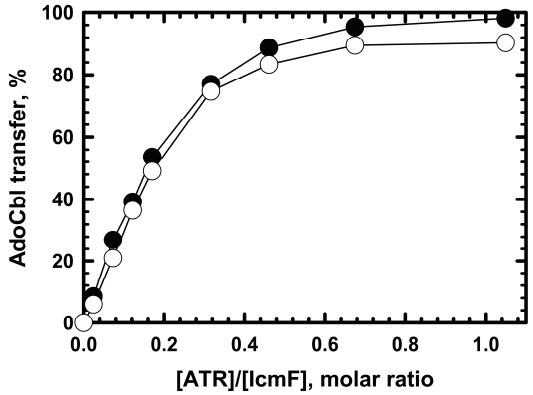


Figure 4.6: The effect of GMPPNP on the reverse transfer from holo-IcmF to apo-ATR. Apo-ATR was added to holo-IcmF (59 μM bound AdoCbl, two equivalents of AdoCbl) in Buffer A supplemented with 5 mM $MgCl_2$ at 24 °C with 2 mM GMPPNP (open circles) or without GMPPNP (filled circles).

4.3.7 Analysis of nucleotide binding to IcmF. To further study the role of the MeaI domain, we measured the affinities of nucleotide to IcmF using both ITC and stopped-flow fluorescence spectroscopy and compared our results with data described for other G-proteins. Chaperones from the G3E family of SIMIBI class bind nucleotides with K_D in low μM range (28,40,41).

Based on the ITC results, IcmF binds one mole of nucleotide per dimer (N = 0.89 ± 0.11 for GDP and 0.78 ± 0.01 for GMPPNP, respectively) (Table 4.1). The binding isotherms

for GDP and GMMPPNP were best fit to a one-site binding model (Figure 4.7). IcmF binds GMPPNP with a ~3-fold higher affinity (K_D=0.95 ± 0.07 μM) than GDP (K_D= 3.3 ± 0.5 μM) (Table 4.1). Binding of GDP is accompanied by a ΔGº of -7.3 ± 0.4 kcal/mol with comparable enthalpic and entropic contributions. Binding of GMPPNP is mostly enthalpically driven with a ΔGº of -8.07 ± 0.1 kcal/mol (Table 4.1).

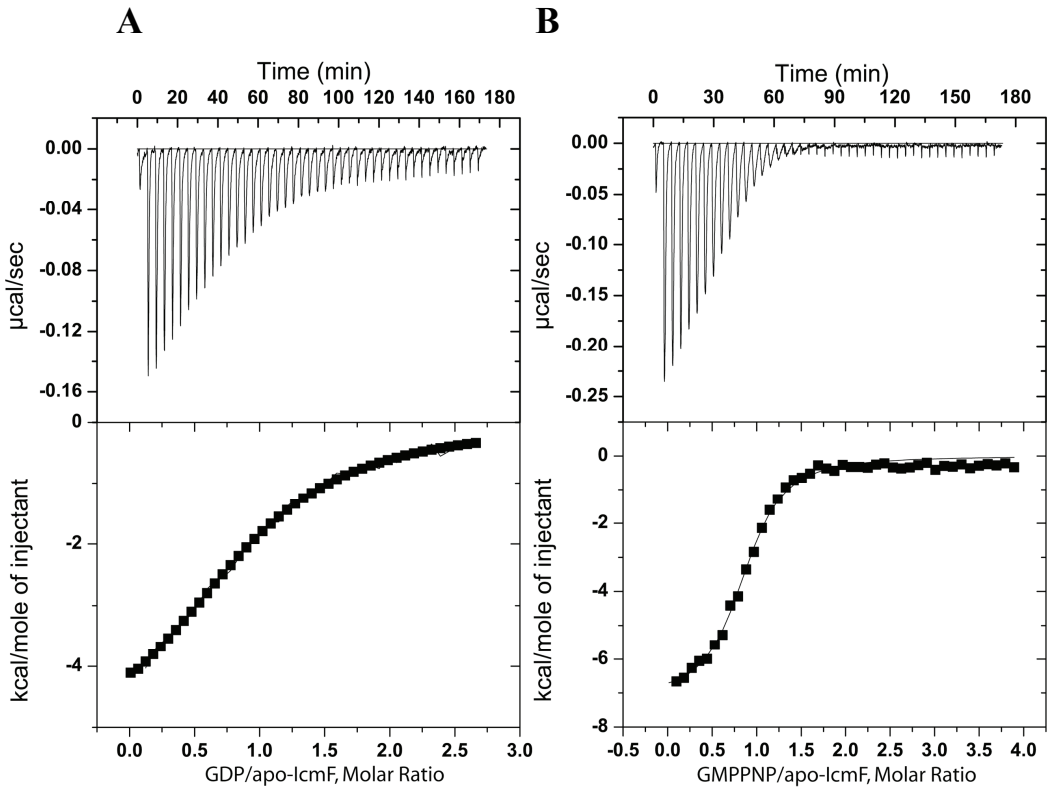


Figure 4.7: Binding isotherms for GDP and GMPPNP binding to IcmF. (**A**) Representative ITC data set for the binding of GDP (255 μM stock solution) to 22.2 μM IcmF in Buffer B at 20 °C. (**B**) Representative ITC data set for the binding of GMPPNP (366 μM stock solution) to 18 μM IcmF in Buffer B at 20 °C. The top panel shows the raw data versus time. The bottom panel shows the integrated areas normalized to the moles of nucleotides added with each injection. Data were fitted to a one-site binding model and yielded the parameters reported in Table 4.1.

Stopped-flow fluorescence spectroscopy was used as an independent method for monitoring binding of fluorescently labeled GDP (mant-GDP) to IcmF (Figure 4.8). Stopped-flow experiments were carried out under pseudo-first order conditions: 0.25 μM

mant-GDP (after mixing) was mixed with increasing concentrations of apo-IcmF (0.6-5.8 μM after mixing) (Figure 4.8A). Stopped-flow fluorescence traces were fit to a double-exponential function to obtain the observed rate constants, k_{obs1} and k_{obs2}, for mant-GDP binding to subunits 1 and 2, respectively (Figure 4.8B, Table 4.2). The amplitudes associated with the two phases (ΔA1=2ΔA2) were unequal. The two rate constants exhibited a linear dependence on protein concentration consistent with a simple binding mechanism (Figure 4.8 B).

Thermodynamic parameters	GDP[a]	GMPPNP[b]
K_D, μM	3.28 ± 0.5	0.95 ± 0.07
ΔH, kcal/mol	-3.7 ± 0.4	-7.02 ± 0.09
TΔS, kcal/mol	+3.6 ± 0.3	+1.05 ± 0.1
ΔGº, kcal/mol	-7.3 ± 0.4	-8.07 ± 0.10
N	0.89 ± 0.11	0.78 ± 0.01

Table 4.1: Thermodynamic parameters for the binding of nucleotides to apo-IcmF. [a,b] The isotherms for binding of GDP and GMPPNP to IcmF were best fit to a one-site binding model. All experiments were performed in buffer B at 20ºC described under Experimental Procedures. The data represent the mean ± S.D. of two independent experiments.

To obtain the dissociation rate constant, GDP (200-300 μM, after mixing) was mixed with 0.25 μM mant-GDP•5 μM IcmF (after mixing) and traces from the displacement experiment were fitted to a double exponential function. Values of $k_{off1\text{-}dis.}$=0.85 ± 0.09 s^{-1} and $k_{off2\text{-}dis.}$=0.11 ± 0.01 s^{-1} were obtained (Figure 4.8C, Table 4.2). In our calculations, we used $k_{off1\text{-}dis.}$ and $k_{off2\text{-}dis.}$ determined from the displacement experiment, to calculate K_{D1}=1.57 ± 0.08 μM and K_{D2}=6.55 ± 0.39 μM (Table 4.2). The IcmF dimer contains two MeaI domains, and hence, two sites for nucleotide binding. In all our stopped-flow

experiments, the fluorescence traces were clearly biphasic and fits were not of acceptable quality when a single exponential function was used for fitting. This is in a good agreement with the two binding sites for nucleotides being unequal in IcmF dimmer. Like the nucleotide binding in the MeaI domain, the two binding sites for AdoCbl in IcmF also exhibit different affinities (20).

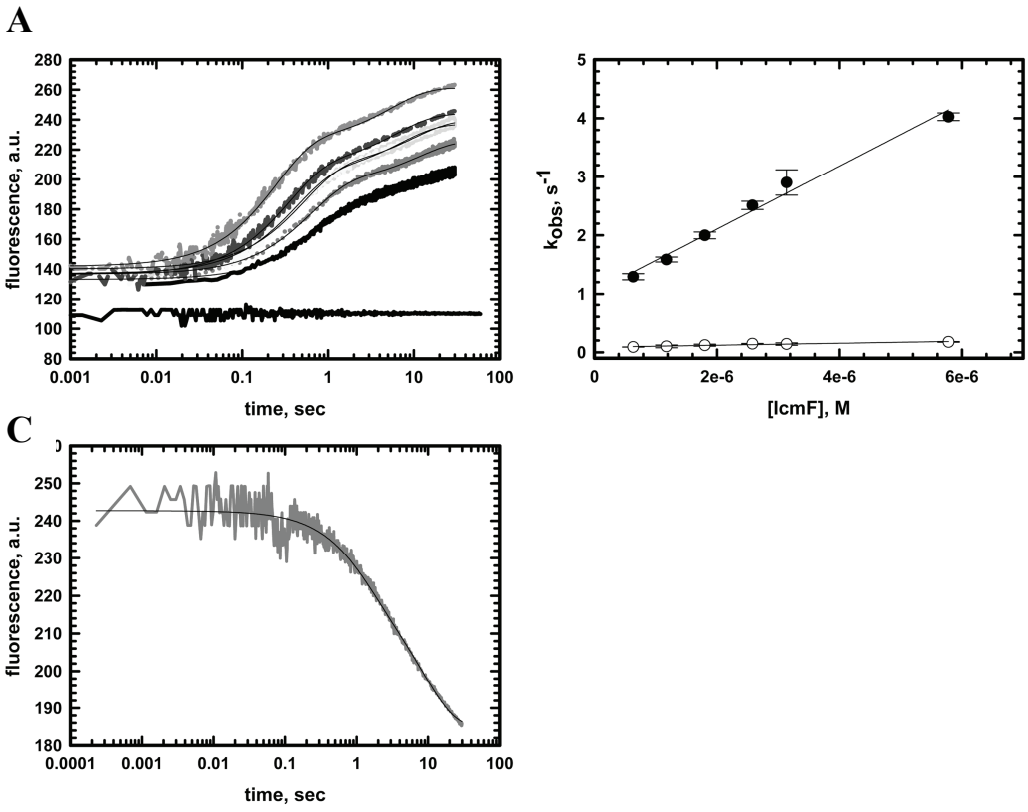


Figure 4.8: Binding of mant-GDP to IcmF. (**A**) Representative fluorescent traces observed by mixing 0.25 μM mant-GDP and 0.6-5.8 μM IcmF in buffer 50 mM NaPi pH 7.5, 150 mM NaCl supplemented with 5 mM $MgCl_2$ at 20°C. Black lines represent the fit with a double exponential function. (**B**) Dependence of two sets of observed rate constants (k_{obs1},● , and k_{obs2}, ○) on protein concentration. Black lines represent the liner fit (Table 3). (**C**) Dissociation of 0.25 μM mant-GDP from 5 μM IcmF was monitored in a displacement experiment using 250 μM GDP. The kinetic trace is best fit to a double exponential function (black line) with $k_{obs1\text{-}diss.}$=0.85 ± 0.09 s^{-1} and $k_{obs2\text{-}diss.}$=0.11 ± 0.01 s^{-1} (Table 4.2).

	mant-GDP[a]	
Kinetic parameter	Site 1	Site 2
k_{on}, $\mu M^{-1}s^{-1}$	0.54 ± 0.03	0.0168 ± 0.0025
k_{off}, s^{-1}	1.03 ± 0.09	0.087 ± 0.007
[b]$k_{off\text{-}dis.}$, s^{-1}	0.85 ± 0.09	0.11 ± 0.01
K_D, μM	1.57 ± 0.08	6.55 ± 0.39

Table 4.2: Kinetic Parameters for binding of mant-GDP to apo-IcmF. [a]Stopped-flow traces were fit to double exponential function. [b]Dissociation constants determined in displacement experiment. All stopped-flow experiments were performed in buffer A with 5 mM $MgCl_2$ as described under Experimental Procedures.

4.3.8 Properties of *Gk* IcmF truncation constructs: the N-terminal part of IcmF and the C-terminal part of IcmF. The results described above suggested that MeaI domain of IcmF modulates AdoCbl transfer from ATR to IcmF. To further study the role of MeaI, we created two truncated IcmF variants: the N-terminal part containing the B_{12}-binding and MeaI domains and the C-terminal part containing the MeaI and the substrate-binding domains (Figure 4.9). When the N-terminal part of IcmF was subjected to gel-filtration chromatography, it eluted as a symmetric peak with an apparent molecular mass of ~158 kDa (Figure 4.10, red trace). The C-terminal part of IcmF truncation variant eluted as a peak with an apparent molecular mass of 251 kDa (Figure 4.10, green trace). Based on this analysis, both truncated variants behave as dimers in solution. The predicted monomeric molecular mass of the N-terminal part is 60 kDa and for the C-terminal part is 106 kDa (Figure 4.10). Since full-length IcmF is a dimer (Figure 4.10, black trace), the simplest interpretation of the oligomeric state of the two truncated

variants is that the MeaI domain is involved in dimerization[6]. This seemed like a reasonable conclusion since MeaB, the stand-alone chaperone for MCM, is a dimer (42).

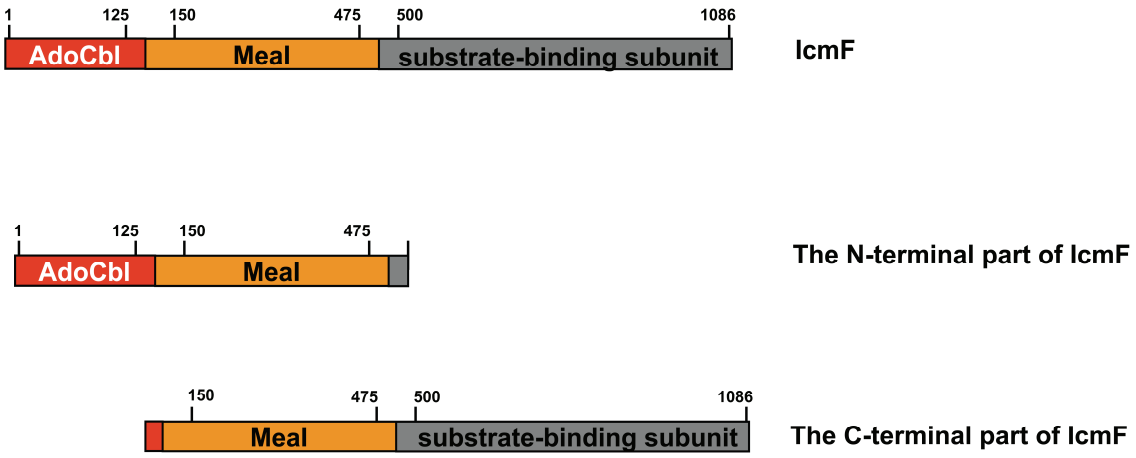


Figure 4.9: Truncated IcmFs variants generated in this study: The N-terminal part containing the B_{12}-binding and the MeaI domains and the C-terminal part containing the MeaI and the substrate-binding domains.

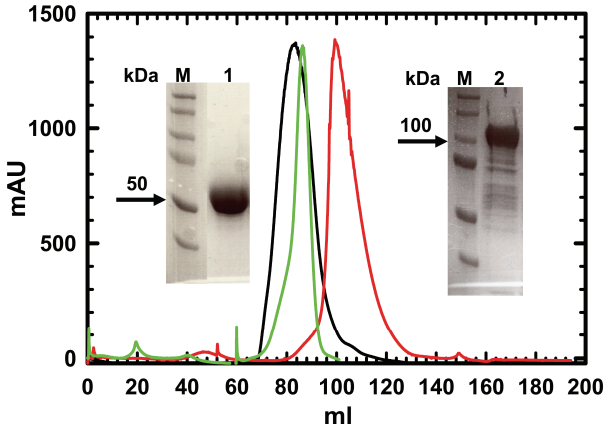


Figure 4.10: Gel-filtration of full-length IcmF and the truncated constructs. Full-length IcmF, (molecular mass = 240 kDa, black trace), the N-terminal part of IcmF (molecular mass =120 kDa, red trace) and the C-terminal part of IcmF (molecular mass = 212 kDa, green trace). mAU, milli absorbance units. The N- and C-terminal parts of IcmF migrated during size-exclusion chromatography on Superdex 200 with apparent molecular masses of ~158 kDa and ~251 kDa, respectively, consistent with homodimeric assembly in each case. The SDS-PAGE analysis of the N-terminal part of IcmF (lane #1) and of the C-terminal part of IcmF (lane #2).

[6] We now know from the crystal structure of full-length IcmF that dimerization does not involve the MeaI domain and hence, a different domain interface is involved in dimerization of the N-terminal part of IcmF.

4.3.9 NTPase activity of the truncated IcmF variants. We have shown recently that the G4 signature motif NKxD, which is responsible for the specificity for guanine in the MeaI domain, is modified to NKxE in *Gk* IcmF (20) (See also section 3.4.7). As a result, *Gk* IcmF has relaxed substrate specificity and can accept both GTP and ATP (Table 4.3). For the N-terminal part of *Gk* IcmF, the following kinetic parameters were obtained: k_{cat} with ATP = 12 ± 1 min^{-1} and k_{cat} with GTP = 23 ± 2 min^{-1} (Table 4.3). Since the K_M for ATP is (910 ± 310 μM) relative to that for GTP (130 ± 56 μM) enzyme is more efficient with GTP as a substrate based on the k_{cat}/K_M value (Table 4.3).

Kinetic parameter	Wild-type IcmF	The N-terminal part of IcmF
k_{cat} with GTP, min^{-1}	10 ± 2	23 ± 2
K_{GTP}, μM	51 ± 3	130 ± 56
$k_{cat\ with\ GTP}/K_{GTP}$, $M^{-1}\ min^{-1}$	$(1.96 \pm 0.37)x10^5$	$(1.77 \pm 0.88)x10^5$
k_{cat} with ATP, min^{-1}	19 ± 1	12 ± 1
K_{ATP}, μM	1290 ± 300	910 ± 310
$k_{cat\ with\ ATP}/K_{ATP}$, $M^{-1}\ min^{-1}$	$(1.47 \pm 0.03)x10^4$	$(1.32 \pm 0.34)x10^4$

Table 4.3. GTPase and ATPase Activity of wild-type IcmF and the N-terminal part of IcmF. All experiments were performed in buffer A with 20mM $MgCl_2$ at 37°C as described under Experimental Procedures. Values represent the results of at least five independent experiments.

In the presence of 0.1 mM GMPPNP, the ATPase activity of the N-terminal part of IcmF was inhibited ~60% (higher concentrations of GMPPNP could not be employed in the malachite-green assay due to the high background). In contrast, very low GTPase activity

(0.034 ± 0.005 min^{-1}) was detected with the C-terminal part of IcmF, despite the presence of the MeaI domain.

4.3.10 Binding of AdoCbl to the N-terminal part of IcmF. We investigated AdoCbl binding to the N-terminal variant of IcmF by ITC (Figure 4.11). These experiments revealed the presence of two non equivalent binding sites with an ~44-fold difference in affinity for AdoCbl (K_{D1}= 4 ± 0.2 μM and K_{D2} = 176 ± 4 μM). These dissociation constants are considerably higher than for the full-length IcmF (K_{D1}=0.081 ± 0.014 μM and K_{D2}= 1.98 ± 0.42 μM) (20).

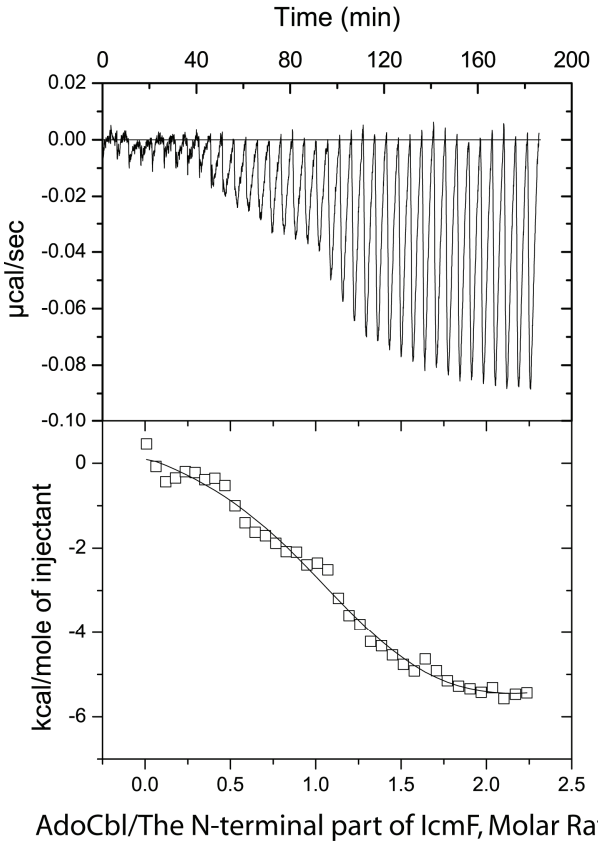


Figure 4.11. Binding isotherms for AdoCbl binding to the N-terminal part of IcmF. Representative ITC data for the binding of AdoCbl (273 μM stock solution) to 30 μM protein in 50 mM HEPES pH 7.5, 50 mM NaCl, 2 mM TCEP and 5 mM $MgCl_2$ at 20 °C. The top panel shows the raw data versus time and the bottom panel shows the integrated areas normalized to the moles of nucleotides added with each injection. Data were fitted to a two-site binding model.

4.3.11 Cofactor transfer between ATR and the N-terminal part of IcmF. Next, we examined whether AdoCbl could be transferred from ATR to the N-terminal part of IcmF. However, cofactor transfer was not observed even at a 1:8 ratio of holo-ATR: the N-terminal part of IcmF (Figure 4.12A, dotted red line). Addition of ATP to the same reaction mixture caused transfer of ~50 % of AdoCbl (Figure 4.12 A, dashed green line). Next we investigated the reverse transfer from the holo-N-terminal part of IcmF to apo-ATR (Figure 4.12B). We used < one equivalent of cofactor in our experiment because of the low affinity of the N-terminal part of IcmF for AdoCbl. The reverse transfer goes to completion when ATR: N-terminal part of IcmF ratio is ~2:0 (Figure 4.12 B, inset). Our result suggest the possibility that the organization of the N-terminal part of IcmF variant might be significantly different than of the parent protein although both exist as dimers in solution.

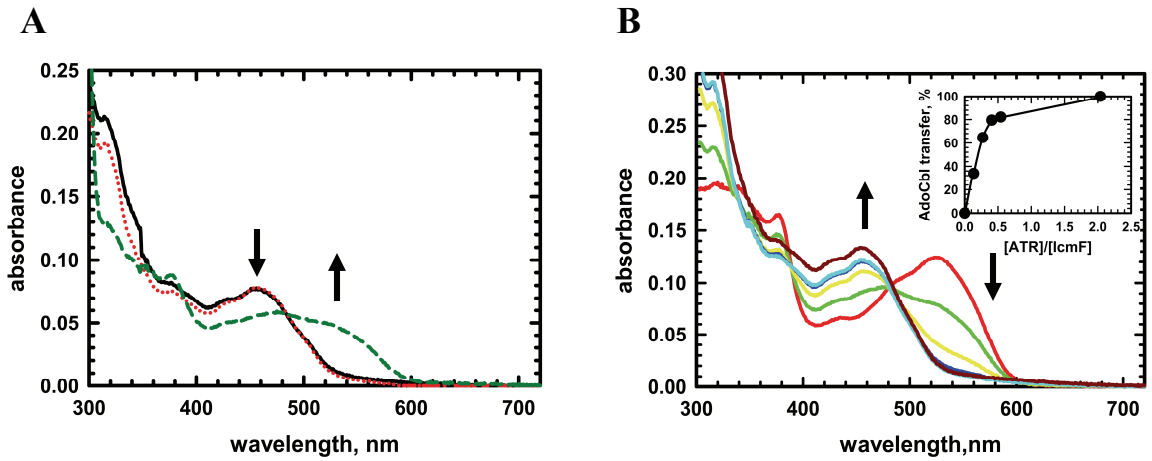


Figure 4.12: Transfer of AdoCbl between ATR and the N-terminal part of IcmF. (**A**) Forward transfer from holo-ATR to apo-N-terminal part of IcmF. In this experiment 10 µM AdoCbl-loaded ATR was added to 50 µM apo-N-terminal part of IcmF (dotted red line). The same reaction mixture after addition of 2 mM ATP (dashed green line). Holo-ATR (11 µM) alone (solid black line). Cofactor transfer was not observed without ATP. In the presence of ATP, only ~ 50% of the cofactor was transferred, whereas for ATR and WT IcmF with ATP the forward transfer goes to full completion. (**B**) Reverse transfer between the holo-N-terminal part of IcmF and apo-ATR. The reaction mixture contained 27 µM of the N-terminal part of IcmF (16 µM bound AdoCbl) and 0-60 µM apo-ATR.

4.4 Discussion

Characterization of ATR and IcmF from *G. kaustophilus* reveals that the PduO-type ATR can transfer AdoCbl directly to IcmF (Figure 4.4). This transfer resembles the process described for ATR and MCM from *M. extorquens* (16). We speculate that transfer of AdoCbl by ATR to the target AdoCbl-dependent isomerase might represent a general strategy for ensuring cofactor delivery, particularly in organisms where cobalamin is a rare cofactor.

The difference in the coordination state of AdoCbl bound to the active sites of ATR and IcmF allows for ready monitoring of AdoCbl transfer by UV-visible absorption spectroscopy (Figure 4.2, Figure 4.3B and Figure 4.4A).

AdoCbl transfer is affected by the presence of nucleotides in the reaction mixture. In the absence of ATP, the reverse transfer, i.e., the transfer from holo-IcmF to apo-ATR, is favored (Figure 4.4 B). Notably in the presence of ATP, which is present at high concentrations inside cells, the forward transfer is greatly favored and ~ 90% transfer is achieved at a 1:1 ratio of ATR: IcmF (Figure 4.4 B).

In the presence of GMPPNP, the forward transfer is inhibited, which indicates that hydrolysis of GTP is required and when hydrolysis is blocked, the transfer is precluded (Figure 4.5). Hence, GTP hydrolysis is required for facilitating transfer as it was reported previously for transfer of AdoCbl from ATR to MCM in the presence of MeaB (21) (See also section 1.11). In contrast, binding of free AdoCbl by IcmF is practically unaffected by the presence or absence of GMPPNP (20). In contrast, complexation of MeaB and MCM in the presence of GMPPNP prevents MCM from AdoCbl binding (22).

To obtain more insight into the mode of transfer between ATR and IcmF, we created two truncated variants of IcmF protein lacking either the B_{12}-binding or the substrate-binding domains but retaining MeaI domains (Figure 4.9). Characterization of the N-terminal part of IcmF confirmed our previous findings that the NKxE sequence in the nucleotide specificity loop relaxes nucleotide specificity and allows the MeaI domain to accept both GTP and ATP (Table 4.3). Since the C-terminal part of IcmF containing the substrate-binding domain has very low intrinsic GTPase activity (0.034 ± 0.005 min^{-1}), it suggests that that the B_{12}-domain stimulates the NTPase activity of the MeaI domain, i.e. the B_{12} domain functions as a GAP domain (See section 5.3.5).

Interestingly, no forward transfer of AdoCbl was detected from ATR to the N–terminal part of IcmF (Figure 4.12A). In the presence of ATP, which triggers the release of one equivalent of AdoCbl, transfer was only 50% (Figure 4.12A).

A hallmark of the G3E family metallochaperones is their relatively weak affinity for nucleotides (in low μM range), which allows for facile release of GDP without the need for GEFs (41,43). IcmF also binds nucleotides with similar affinity (Table 4.1 and Table 4.2). Stopped-flow fluorescence studies with mant-GDP revealed that MeaI domains in IcmF dimer bind mant-GDP with non-equivalent affinities (Figure 4.8, Table 4.2).

Since the bacterial PduO-type and EutT-type ATRs can adenosylate Cbi in addition to Cbl (Figure 4.1A) (11,12), it would be interesting to evaluate whether a truncated cofactor, i.e. 5'-deoxyadenosyl cobinamide, can be transferred to the mutase or whether the G-protein chaperone exerts an "editing" function, preventing this from happening.

In summary, we have demonstrated that ATR transfers AdoCbl to IcmF in a process that is gated by GTP hydrolysis. ATP also shifts the equilibrium of the transfer process in the

forward direction, i.e. from ATR to IcmF. Since there is only one copy of PduO-type ATR in the *G. kaustophilus* genome, it appears that this ATR serves two client mutases, MCM and IcmF. In fact PduO-type ATR was originally discovered in 1,2-propanediol utilization (*pdu*) operon where it is found together with AdoCbl-dependent diol dehydratase (PduCDE) in *Salmonella enterica* (7). Most probably ATR is also involved in AdoCbl delivery to diol dehydratase.

4.5 References

(1) Martens, J. H., Barg, H., Warren, M. J., and Jahn, D. (2002) Microbial production of vitamin B12. *Appl Microbiol Biotechnol 58*, 275-85.

(2) Raux, E., Schubert, H. L., and Warren, M. J. (2000) Biosynthesis of cobalamin (vitamin B12): a bacterial conundrum. *Cell Mol Life Sci 57*, 1880-93.

(3) Warren, M. J., Raux, E., Schubert, H. L., and Escalante-Semerena, J. C. (2002) The biosynthesis of adenosylcobalamin (vitamin B12). *Nat Prod Rep 19*, 390-412.

(4) Krautler, B. (2009) Organometallic chemistry of b(12) coenzymes. *Met Ions Life Sci 6*, 1-51.

(5) Banerjee, R. (2003) Radical carbon skeleton rearrangements: catalysis by coenzyme B12-dependent mutases. *Chem Rev 103*, 2083-94.

(6) Matthews, R. G. (2009) Cobalamin- and corrinoid-dependent enzymes. *Met Ions Life Sci 6*, 53-114.

(7) Johnson, C. L., Pechonick, E., Park, S. D., Havemann, G. D., Leal, N. A., and Bobik, T. A. (2001) Functional genomic, biochemical, and genetic characterization of the Salmonella pduO gene, an ATP:cob(I)alamin adenosyltransferase gene. *J Bacteriol 183*, 1577-84.

(8) Buan, N. R., Suh, S. J., and Escalante-Semerena, J. C. (2004) The eutT gene of Salmonella enterica Encodes an oxygen-labile, metal-containing ATP:corrinoid adenosyltransferase enzyme. *J Bacteriol 186*, 5708-14.

(9) Fonseca, M. V., Buan, N. R., Horswill, A. R., Rayment, I., and Escalante-Semerena, J. C. (2002) The ATP:Co(I)rrinoid adenosyltransferase (CobA) enzyme of Salmonella enterica requires the 2'-OH group of ATP for function and yields inorganic triphosphate as its reaction byproduct. *J Biol Chem 277*, 33127-31.

(10) Mera, P. E., and Escalante-Semerena, J. C. (2010) Multiple roles of ATP:cob(I)alamin adenosyltransferases in the conversion of B12 to coenzyme B12. *Appl Microbiol Biotechnol 88*, 41-8.

(11) Park, K., Mera, P. E., Escalante-Semerena, J. C., and Brunold, T. C. (2008) Kinetic and spectroscopic studies of the ATP:corrinoid adenosyltransferase PduO from Lactobacillus reuteri: substrate specificity and insights into the mechanism of Co(II)corrinoid reduction. *Biochemistry 47*, 9007-15.

(12) Buan, N. R., and Escalante-Semerena, J. C. (2006) Purification and initial biochemical characterization of ATP:Cob(I)alamin adenosyltransferase (EutT) enzyme of Salmonella enterica. *J Biol Chem 281*, 16971-7.

(13) Drennan, C. L., Huang, S., Drummond, J. T., Matthews, R. G., and Lidwig, M. L. (1994) How a protein binds B12: A 3.0 A X-ray structure of B12-binding domains of methionine synthase. *Science 266*, 1669-74.

(14) Yamanishi, M., Vlasie, M., and Banerjee, R. (2005) Adenosyltransferase: an enzyme and an escort for coenzyme B12? *Trends Biochem Sci 30*, 304-8.

(15) Yamanishi, M., Labunska, T., and Banerjee, R. (2005) Mirror "base-off" conformation of coenzyme B12 in human adenosyltransferase and its downstream target, methylmalonyl-CoA mutase. *J Am Chem Soc 127*, 526-7.

(16) Padovani, D., Labunska, T., Palfey, B. A., Ballou, D. P., and Banerjee, R. (2008) Adenosyltransferase tailors and delivers coenzyme B12. *Nat Chem Biol 4*, 194-6.
(17) Daublain, P., Horner, J. H., Kuznetsov, A., and Newcomb, M. (2004) Solvent polarity effects and limited acid catalysis in rearrangements of model radicals for the methylmalonyl-CoA mutase- and isobutyryl-CoA mutase-catalyzed isomerization reactions. *J Am Chem Soc 126*, 5368-9.
(18) Moore, B., Eisenberg, R., Weber, C., Bridges, A., Nanz, D., and Robinson, J. (1995) On the stereospecificity of the coenzyme B12-dependent isobutyryl-CoA mutase reaction. *J. Am. Chem. Soc. 117*, 11285-11291.
(19) Ratnatilleke, A., Vrijbloed, J. W., and Robinson, J. A. (1999) Cloning and sequencing of the coenzyme B(12)-binding domain of isobutyryl-CoA mutase from Streptomyces cinnamonensis, reconstitution of mutase activity, and characterization of the recombinant enzyme produced in Escherichia coli. *J Biol Chem 274*, 31679-85.
(20) Cracan, V., Padovani, D., and Banerjee, R. (2010) IcmF is a fusion between the radical B12 enzyme isobutyryl-CoA mutase and its G-protein chaperone. *J Biol Chem 285*, 655-66.
(21) Padovani, D., and Banerjee, R. (2009) A G-protein editor gates coenzyme B12 loading and is corrupted in methylmalonic aciduria. *Proc Natl Acad Sci U S A 106*, 21567-72.
(22) Padovani, D., and Banerjee, R. (2006) Assembly and protection of the radical enzyme, methylmalonyl-CoA mutase, by its chaperone. *Biochemistry 45*, 9300-6.
(23) Chowdhury, S., Thomas, M. G., Escalante-Semerena, J. C., and Banerjee, R. (2001) The coenzyme B12 analog 5'-deoxyadenosylcobinamide-gdp supports catalysis by methylmalonyl-CoA mutase in the absence of trans-ligand coordination. *J Biol Chem 276*, 1015-9.
(24) Padovani, D., and Banerjee, R. (2009) A rotary mechanism for coenzyme B(12) synthesis by adenosyltransferase. *Biochemistry 48*, 5350-7.
(25) Ahmadian, M. R., Wittinghofer, A., and Herrmann, C. (2002) Fluorescence methods in the study of small GTP-binding proteins. *Methods Mol Biol 189*, 45-63.
(26) Marlovits, T. C., Haase, W., Herrmann, C., Aller, S. G., and Unger, V. M. (2002) The membrane protein FeoB contains an intramolecular G protein essential for Fe(II) uptake in bacteria. *Proc Natl Acad Sci U S A 99*, 16243-8.
(27) Hemsath, L., and Ahmadian, M. R. (2005) Fluorescence approaches for monitoring interactions of Rho GTPases with nucleotides, regulators, and effectors. *Methods 37*, 173-82.
(28) Padovani, D., Labunska, T., and Banerjee, R. (2006) Energetics of interaction between the G-protein chaperone, MeaB, and B12-dependent methylmalonyl-CoA mutase. *J Biol Chem 281*, 17838-44.
(29) Alm, E. J., Huang, K. H., Price, M. N., Koche, R. P., Keller, K., Dubchak, I. L., and Arkin, A. P. (2005) The MicrobesOnline Web site for comparative genomics. *Genome Res 15*, 1015-22.
(30) Abreu-Goodger, C., and Merino, E. (2005) RibEx: a web server for locating riboswitches and other conserved bacterial regulatory elements. *Nucleic Acids Res 33*, W690-2.

(31) Ozenberger, B. A., Nahlik, M. S., and McIntosh, M. A. (1987) Genetic organization of multiple fep genes encoding ferric enterobactin transport functions in Escherichia coli. *J Bacteriol 169*, 3638-46.
(32) Sprencel, C., Cao, Z., Qi, Z., Scott, D. C., Montague, M. A., Ivanoff, N., Xu, J., Raymond, K. M., Newton, S. M., and Klebba, P. E. (2000) Binding of ferric enterobactin by the Escherichia coli periplasmic protein FepB. *J Bacteriol 182*, 5359-64.
(33) Wyckoff, E. E., Valle, A. M., Smith, S. L., and Payne, S. M. (1999) A multifunctional ATP-binding cassette transporter system from Vibrio cholerae transports vibriobactin and enterobactin. *J Bacteriol 181*, 7588-96.
(34) Woodson, J. D., and Escalante-Semerena, J. C. (2004) CbiZ, an amidohydrolase enzyme required for salvaging the coenzyme B12 precursor cobinamide in archaea. *Proc Natl Acad Sci U S A 101*, 3591-6.
(35) Nahvi, A., Barrick, J. E., and Breaker, R. R. (2004) Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes. *Nucleic Acids Res 32*, 143-50.
(36) Fowler, C. C., Brown, E. D., and Li, Y. (2010) Using a riboswitch sensor to examine coenzyme B(12) metabolism and transport in E. coli. *Chem Biol 17*, 756-65.
(37) Pohlmann, A., Fricke, W. F., Reinecke, F., Kusian, B., Liesegang, H., Cramm, R., Eitinger, T., Ewering, C., Potter, M., Schwartz, E., Strittmatter, A., Voss, I., Gottschalk, G., Steinbuchel, A., Friedrich, B., and Bowien, B. (2006) Genome sequence of the bioplastic-producing "Knallgas" bacterium Ralstonia eutropha H16. *Nat Biotechnol 24*, 1257-62.
(38) Ewing, J. (2005) Characterization of the Cobalamin and Fep Operons in Methylobium petrolphilum PM1 (online publication, accession number: UCRL-TR-215218).
(39) Lofgren, M., and Banerjee, R. (2011) Loss of allostery and coenzyme B12 delivery by a pathogenic mutation in adenosyltransferase. *Biochemistry 50*, 5790-8.
(40) Gasper, R., Scrima, A., and Wittinghofer, A. (2006) Structural insights into HypB, a GTP-binding protein that regulates metal binding. *J Biol Chem 281*, 27492-502.
(41) Gasper, R., Meyer, S., Gotthardt, K., Sirajuddin, M., and Wittinghofer, A. (2009) It takes two to tango: regulation of G proteins by dimerization. *Nat Rev Mol Cell Biol 10*, 423-9.
(42) Hubbard, P. A., Padovani, D., Labunska, T., Mahlstedt, S. A., Banerjee, R., and Drennan, C. L. (2007) Crystal structure and mutagenesis of the metallochaperone MeaB: insight into the causes of methylmalonic aciduria. *J Biol Chem 282*, 31308-16.
(43) Wittinghofer, A., and Vetter, I. R. (2011) Structure-function relationships of the G domain, a canonical switch motif. *Annu Rev Biochem 80*, 943-71.

CHAPTER 5

Ongoing Work and Future Directions[7]

5.1 Introduction

The main goal of my dissertation was to dissect, using various biophysical and biochemical methods, the mechanism of cross-talk between the MeaI domain and the two isobutyryl-CoA mutase domains (IcmA and IcmB) in the fusion protein, IcmF (See Chapter 2). Additionally, I examined whether ATR delivers the active form of the cofactor to IcmF as is the case for the ATR/MCM duo of proteins (1). To this end, I developed robust expression systems for both IcmF (see Chapter 1) and ATR from *Geobacillus kaustophilus*. Our working hypothesis implies that conformational changes in MeaI induced by nucleotide binding, regulate access of AdoCbl to the IcmF active site (See Chapter 4).

We also discovered a novel pivalyl-CoA mutase activity of IcmF and demonstrated that only this activity is protected from inactivation in a GTP-dependent manner (See Chapter 3).

Many aspects of IcmF fusion protein remain to be elucidated. My current studies are focused on elucidation of the role of the reaction catalyzed by IcmF *in vivo*.

[7] We are extremely indebted to Dr. Greg Stephanopoulos and Deepak Dugar (MIT) for providing pivalyl-CoA and a pET30Ek/LIC vector expressing IcmF from *Thauera sp.* Also we are very thankful to Dr. Antony Sinskey and Dr. Chris Brigham for metabolome analysis of *Ralstonia eutropha H16*

5.2 Experimental procedures

5.2.1 Construction of IcmF mutants. All mutants were created using the QuikChange XL Site-Directed Mutagenesis Kit (Agilent) and the following sense primers: 5'-CCG AAGCATCACGTCTGCTTTGTGACTGCATCGAGC-3' for R12C; 5'- ATCAACATT ATGGCCCGCATTTTGCAGGCG -3' for R31A; 5'- ATCAACATTATGCGCGCC ATTTTGCAGGCG -3' for R32A; 5'- CCATTTAGGCCATAACGCATCGGTGGA GGAAATCG -3' for R48A.The sequences of the reverse primers were complementary to the corresponding sequences of the forward primers. IcmF from *G. kaustophilus* cloned into pET30 Ek/LIC vector was used as a template. All constructs and mutations were confirmed by nucleotide sequence determination at the University of Michigan DNA sequencing Core.

5.2.2 Protein expression and purification. *(i) Purification of G. kaustophilus (Gk) IcmF mutants.* Proteins were purified as previously described for wild-type *Gk* IcmF (2). *(ii) Purification of IcmF from Thauera sp. (Th).* *E. coli* BL21 (DE3) cells containing plasmid expressing *Th* IcmF were grown in LB media supplemented with 50 μg/mL kanamycin to an absorbance at 600 nm of 0.5-0.6 and induced at 15˚ C with 0.1 mM IPTG. Cells were harvested 12-14 h after the induction. Protein was purified as previously described (2). Gel-filtration chromatography was performed in 50 mM HEPES pH 7.5 containing 100 mM NaCl (Buffer A).

5.2.3 GTPase activity of IcmF. GTPase activity of *Gk* IcmF mutants was measured using an HPLC-based assay as previously described (2).

5.2.4 Mutase activity of IcmF. A GC-based assay was used to measure both the isobutyryl-CoA mutase and a novel pivalyl-CoA mutase activities of *Th* IcmF (2).

(i) Determination of K_M values for isovaleryl-CoA and pivalyl-CoA. The reaction was performed in Buffer A in a total volume of 0.8-1.4 ml containing: 2.2 mg of *Th* IcmF, 100 μM AdoCbl, 25-2000 μM isovaleryl-CoA or pivalyl-CoA, 15 mM $MgCl_2$ and 3 mM GTP. At various time points, 200 μL aliquots were removed and subjected to GC analysis as previously described (2). *(ii) Isobutyryl-CoA mutase activity.* The reaction was performed in Buffer A in a total volume of 0.8-1.4 ml containing: 15 μg of *Th* IcmF, 100 μM AdoCbl, 1.4-2.5 mM isobutyryl-CoA or n-bytyryl-CoA, 10 mM $MgCl_2$ ± 1-3 mM GTP. At various time points, 200 μL aliquots were removed and subjected to GC analysis as previously described (2).

5.2.5 Analysis of myxochromides production in *M. xanthus* by HPLC. *M. xanthus* DK1622 (WT), DK1878 (Δ*icmF*) and Δ*bkd* strains (a strain in which the branched-chain ketoacid dehydrogenase complex is disrupted) were grown in CTT medium (casitone 10 g/L, 1 M Tris pH 7.6 10 ml/L, 1 M K_2HPO_4 pH 7.6 1 ml/L, 0.8 M $MgSO_4$ 10 ml/L ± 10 μM CNCbl) supplemented with the adsorber resin XAD-16 (1%). The culture was inoculated with preculture and incubated for 3-5 days at 30°C on a rotary shaker (160 rpm). Cells and XAD-16 resin were harvested by centrifugation and stored at -80 ºC. Usually, 0.1-0.05 g of frozen sample (cell pellet and XAD resin) was extracted with 0.8-1 ml of methanol. The extract was filtered through 0.2 μm filter and analyzed directly by HPLC. Extracts (100-150 μL injected sample) were separated using an HPLC system equipped with an Alltima HP 5 μm C18 (250 x 4.6 mm) column (Grace, IL). The UV-vis. detector was set up to monitor multiple wavelengths simultaneously (254, 337, 363, 381, 400 and 422 nm). Solvent A was 0.1 % formic acid in water. Solvent B was 0.1 % formic acid in acetonitrile. Initial conditions used for separation were: 5% solvent B; at flow rate

of 1.0 mL/min. Between 5 and 14 min, solvent B was increased to 40 %, and then was held at 40 % for 10 min. Between 24 and 35 min, solvent B was increased to 100 % and held there after for 5 min. At 41 min, solvent B was decreased to 5% and held for 10 min at that composition to equilibrate the column between injections. Under these conditions, the retention time of myxochromides was ~17-19 min and myxalamids A-C was ~36-38 min. Myxochromides were identified by comparison of UV-visible spectra of collected fractions (0.2 mL) with published spectra of myxochromides (*3-5*). Samples under investigation were diluted as required to be have an optical density of <1.0. Myxalamids A-C were identified by comparing published spectra (3, 7) and the LC-MS analysis performed in the laboratory of Dr. Rolf Muller (Department of Pharmaceutical Biotechnology, Saarland University, Germany).

5.2.6 Sample preparation of *Ralstonia eutropha* H16. *R. eutropha* Re2302 (WT) and Re2303 (Δ*icmF*) strains were grown on acetate and bicarbonate ± OHCbl. Samples were submitted to Metabolon Inc. (Durham, NC) for metabolic profiling.

5.3. Results, discussion and future directions

5.3.1 Distribution of stand-alone ICMs. As mentioned previously, a number of bacteria from the *Firmicutes* and *Proteobacteria* phyla were shown to assimilate isobutyrate (Section 1.1). We examined available bacterial genomic sequences for MCM-like proteins where the interconversion of isobutyrate to butyrate has been documented. We have found "stand-alone" ICMs in several bacteria (Figure 5.1 and Table 5.1). Our findings expand the known distribution of "stand-alone" ICMs beyond the genus *Streptomyces*. In *D. baarsii,* IcmF, but not the "stand-alone" ICM, is present and suggests that in this bacterium, IcmF is the only isobutyryl-CoA mutase.

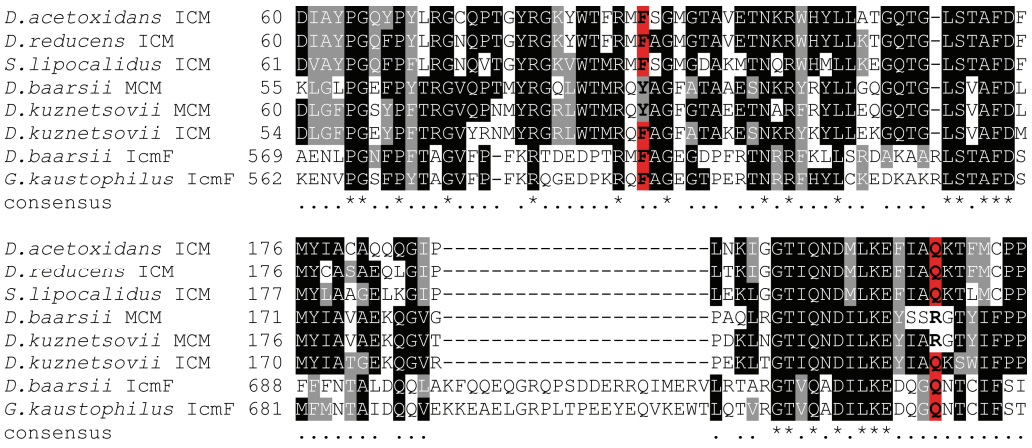


Figure 5.1: "Stand-alone" ICMs identified in this study. Multiple sequence alignment of IcmF from *G. kaustophilus* (YP_149244), MCMs from *Desulfarculus baarsii* (YP_003806846), *Desulfotomaculum kuznetsovii* (YP_004516746) and "stand-alone" ICMs from *Syntrophothermus lipocalidus* (YP_003702103), *D. acetoxidans* (YP_003191373), *D. kuznetsovii* (YP_004516591) and *D. reducens* (YP_001113118). Two residues in ICM and IcmF, Phe and Gln are substituted to Tyr and Arg in MCM.

5.3.2 The role of IcmF in assimilation of pivalic acid. Characterization of IcmF from *Thauera sp.* Dr. Greg Stephanopoulos and colleagues from MIT indentified a couple of bacterial strains that can use pivalic acid as the sole carbon source (personal communication). Subsequently, our laboratories have worked collaboratively to establish

that the pivalyl-CoA mutase activity of IcmF allows *Thauera sp.* to grow on pivalic acid. To this end, we have characterized recombinant IcmF from *Thauera sp.* The kinetic parameters for *Th* IcmF with pivalyl-CoA and isovaleryl-CoA are summarized in Table 5.2 and Figure 5.2. This enzyme is more active with isovaleryl-CoA (S.A. = 0.036 ± 0.0007 μmol/min/mg) compared to pivalyl-CoA (S.A. = 0.0111 ± 0.0008 μmol/min/mg). However, since the K_M for isovaleryl-CoA is 2.8-fold higher than for pivalyl-CoA, the catalytic efficiency of IcmF with both substrates is similar (Table 5.2). *Th* IcmF is much more active with n-butyryl-CoA (S.A. = 42 ± 0.9 μmol/min/mg) and isobutyryl-CoA (S.A. = 24 ± 0.3 μmol/min/mg). In other words, the kinetic parameters of *Th* IcmF are very similar to the parameters measured for other IcmFs (See Section 3.4.2 and Table 2.1).

Organism	ICM	MCM	conversion of isobutyrate to butyrate was detected
Syntrophothermus lipocalidus	+	-	+
Desulfarculus baarsii	+[a]	+	+
Desulfotomaculum acetoxidans	+	-	+
Desulfotomaculum kuznetsovii	+	+	+
Desulfotomaculum reducens	+	-	+

Table 5.1: List of bacteria known to convert isobutyrate to butyrate. [a]-IcmF but not the "stand-alone" ICM is present.

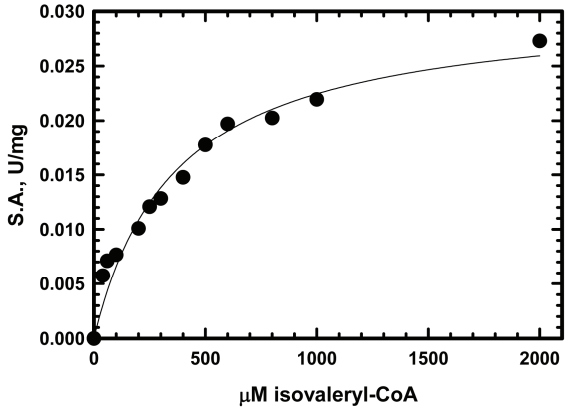


Figure 5.2:Michaelis-Menten analysis of the reaction catalyzed by IcmF from *Thauera sp.* as determined by the GC-based assay. Reaction mixture in Buffer A contained: 10 mM $MgCl_2$, 2.27 mg *Th* IcmF, 100 μM AdoCbl, 2 mM GTP and 20-2000 μM isovaleryl-CoA. The kinetic parameters are reported in Table 5.2.

	Substrate	
	Isovaleryl-CoA	Pivalyl-CoA
K_M, μM	396 ± 13	139 ± 32
S.A., μmol/min/mg	0.036 ± 0.0007	0.0111 ± 0.0008
k_{cat}, s^{-1}	0.144 ± 0.003	0.044 ± 0.003
k_{cat}/ K_M, $M^{-1} s^{-1}$	$(3.64 \pm 0.04)x10^2$	$(3.17 \pm 0.51)x10^2$

Table 5.2: Kinetic parameters of IcmF from *Thauera sp*. Measured in the presence of GTP.Values represent the average of at least 3 independent experiments. S.A. with n-butyryl-CoA is 42 ± 0.9 U/mg and with isovaleryl-CoA is 24 ± 0.3 U/mg.

When isovaleryl-CoA was used as a substrate, the reaction catalyzed by *Th* IcmF was linear only for the first 2 min and a complete cessation of the reaction was observed after 5 min (Figure 5.3). In contrast, when the reaction mixture was supplemented with GTP, the enzyme was still active even after 15 min (Figure 5.3). When isobutyryl-CoA/n-butyryl-CoA, were employed as substrates, no protective effect was seen. The same

effect of nucleotides on the enzyme activity was seen for IcmF from *G. kaustophilus* and *C. metallidurans* (Section 3.4.4).

Dr. Stephanopoulos' group at MIT is currently working on a reconstructing a pathway in *E. coli*, which will allow IcmF-dependent mineralization of pivalic acid. Pivalic acid, which is found in sludge, is mostly of anthropogenic origin and its clearance is an important target of bioremediation. Previously MCM was successfully integrated in an engineered pathway in *E.coli* (*6*). We are planning to use the same strategy to have functional AdoCbl-loaded IcmF in the pathway (6).

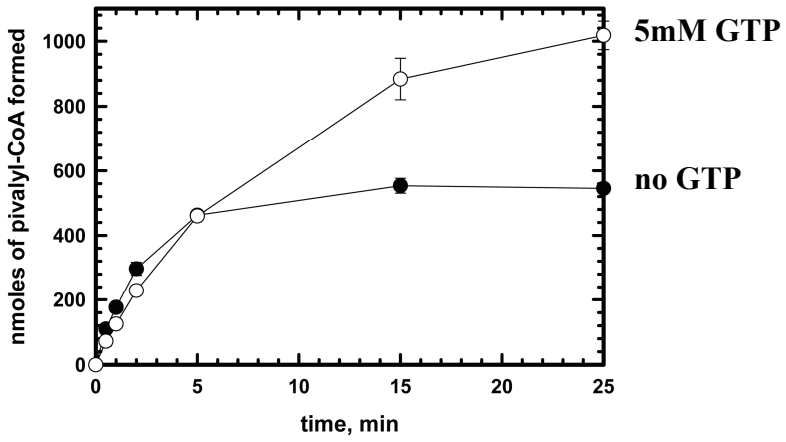
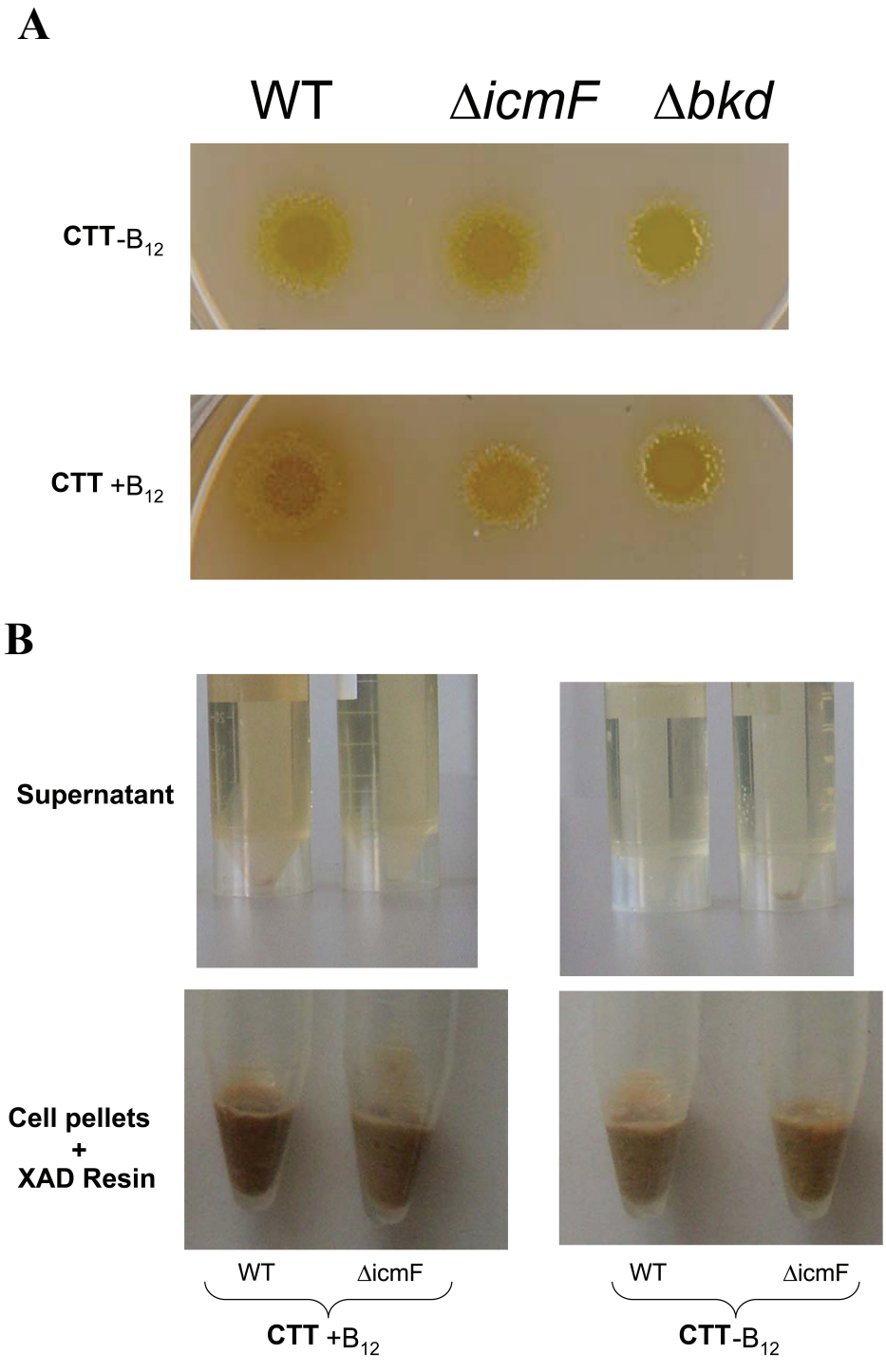


Figure 5.3: Effect of GTP on the time course of the pivalyl-CoA mutase reaction catalyzed by *Th* IcmF. The reaction mixture in Buffer A contained: 10 mM $MgCl_2$, 4 mg *Th* IcmF, 100 μM AdoCbl, 1.6 mM isovaleryl-CoA without GTP (black circles) and with 5 mM GTP (white circles) at 37 °C. Aliquots of the reaction were removed at: 0.5, 1, 2, 5, 15 and 25 min and analyzed by GC as described under Methods. Error bars represent the mean ± SD of three independent experiments.

5.3.3 Role of IcmF catalyzed reaction in the metabolism of *M.xanthus.* The role of the ICM catalyzed reaction in the biosynthesis of polyketides is well documented (Section 1.4). However, the only IcmF-containing organism that is known to produce polyketides

is myxobacterium *Myxococcus xanthus*[8]. The latter produces myxalamids A-D, which are inhibitors of the eukaryotic electron transport chain (*3, 7*).



[8] *Stigmatella aurantiaca,* another bacterium which contains IcmF, is almost identical to *M. xanthus* from a genetic point of view.

Figure 5.4: "Brown" phenotype of WT *M. xanthus* grown in the presence of CNCbl. (**A**) CTT plates ± CNCbl with WT *M. xathus* DK1622, Δ*icm*F and Δ*bkd*. (**B**) WT *M. xanthus* DK1622 and mutants (Δ*icm*F and Δ*bkd*) grown in liquid CTT medium ± CNCbl. Supernatant and cell pellets with XAD resin are shown after centrifugation. These data were generated in the laboratory of Dr. Montserrat Elias-Arnanz (Universidad de Murcia, Spain).

In collaboration with Dr. Montserrat Elias-Arnanz (Universidad de Murcia, Spain), the *icmF* gene in *M. xanthus* DK1622 has been deleted, as previously described (*8*). Briefly, ~1 Kb of genomic DNA upstream and downsream of *icm*F was cloned. Then, the construct was introduced into *M. xanthus* DK1622 and haploid colonies lacking the *icmF* allele were identified by PCR. *M. xanthus* DK1878 designation was given to the strain with the deleted *icm*F gene.

We noticed that when *M. xanthus* DK1622 age on plates (CTT rich medium in the presence of CNCbl), cells and the medium around them become brownish (Figure 5.4A). The appearance of the brownish color was not seen in either DK1878 and Δ*bkd* strains grown in the presence of CNCbl (Figure 5.4 A). When CNCbl was omitted from the medium, the brown color was not seen.

Next, we confirmed our results in liquid culture in CTT medium containing the XAD resin (which is routinely added to *M. xanthus* cultures to bind various polyketides which are secreted by bacteria) (Figure 5.4B). Finally, we note that in DK1050 (an *M. xanthus* strain widely used in many laboratories) the brownish phenotype seen with DK1622 is not so obvious (data not shown).

HPLC was used to fractionate a methanol extract of the XAD resin (DK1622 cultures were grown + CNCbl for 3 days) into two major fractions, one eluted at 17-19 min and the other eluted at 36-38 min (Figure 5.5). Interestingly, the peaks that eluted at 17-19

min are decreased in intensity in DK1878 (~8-fold reduction in absorbance at 400 nm, data not shown). We collected 0.15 mL fractions in this region and analyzed them by UV-visible spectroscopy (Figure 5.5, inset). Analysis of individual fractions clearly shows spectra with maxima at 405 and 420 nm, which provides an explanation for the brownish color of the observed phenotype. Similar spectra were attributed to myxochromides isolated from *M. xanthus* DK1050 and *Stigmatella aurantiaca* (*4, 5, 9, 10*). Myxochromides are compounds that are comprised of a peptidic core (usually 5-6 amino acids) and polyunsaturated side chains (Figure 5.6) Myxochromides differ in the composition and order of amino acids in the peptidic core and both the length and the number of double bonds in the polyunsaturated tail (Figure 5.6).

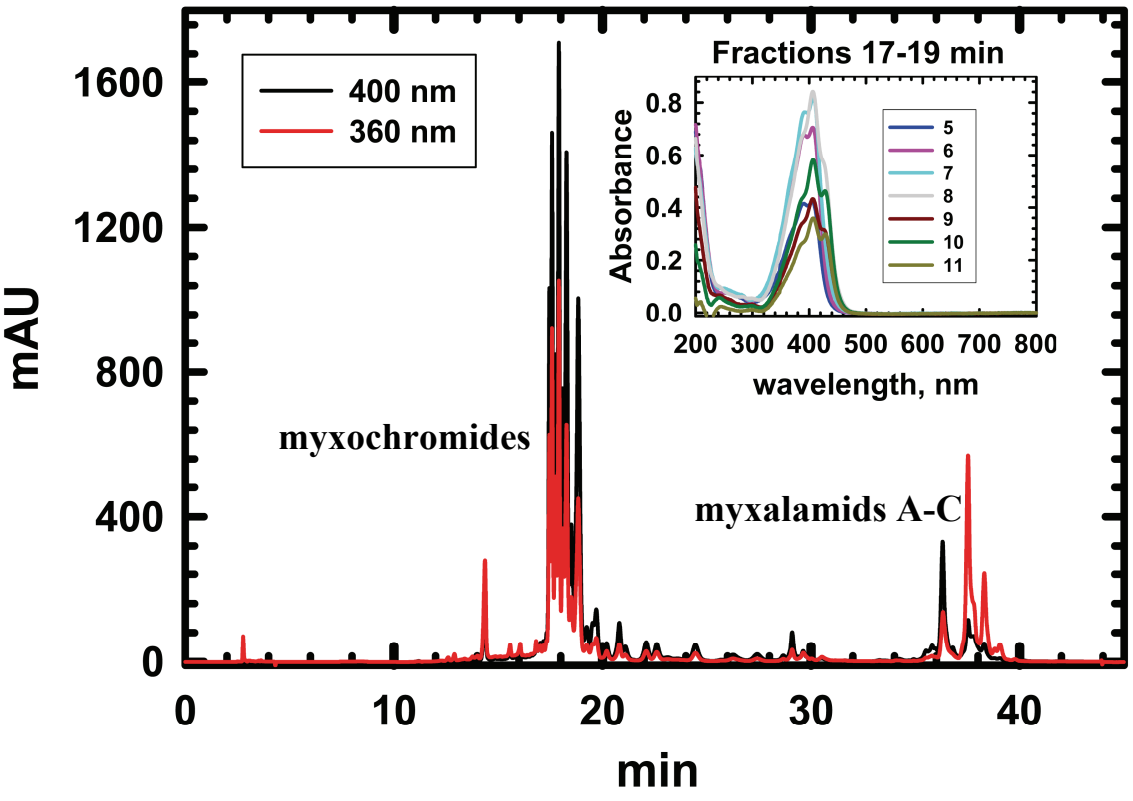


Figure 5.5: HPLC analysis of wild-type and Δ*icm*F mutant of *M. xanthus*. HPLC traces depict separation of methanol extracts of cell pellets/XAD resin monitored at 360

nm (absorbance maximum of myxalamids A-C) and 400 nm (absorbance maximum of myxochromides). *Inset*: UV-visible spectra of individual fractions collected at 17-19 min.

Currently, in collaboration with Dr. Rolf Muller (Department of Pharmaceutical Biotechnology, Saarland University, Germany), we are using LC-MS/HPLC analysis for further characterization of the partially purified compounds that we believe belong to myxochromides.

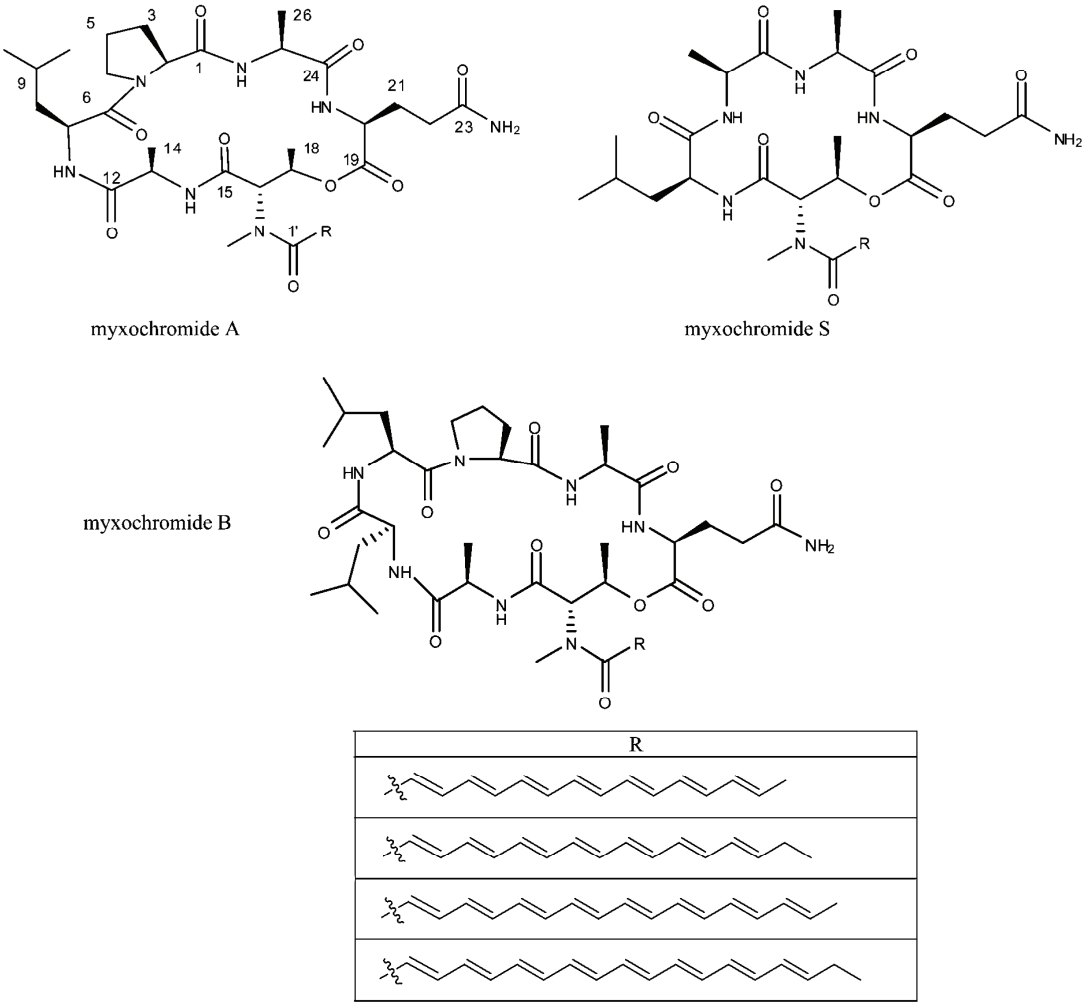


Figure 5.6: **Overall structures of myxochromides.** Myxochromides are comprised of a peptidic core (A, B, or S) and a polyunsaturated side chain (R). Adapted from (*4*).

We are also considering using the reported strategy (*10*) of feeding experiments employing [$^{13}C_4$, $^{15}N_1$] threonine to determine the structure of myxochromides. Since production of myxochromides in the Δ*bkd* strain is also diminished, it supports our hypothesis that IcmF is involved in metabolism of branched-chain amino acids.

We initially proposed that the isobutyryl-CoA starter unit for myxalamide B is synthesized from n-butyryl-CoA in a reaction catalyzed by IcmF. In studies on the *M. xanthus* Δ*bkd* strain (DK1878), isobutyryl-CoA was found to be incorporated into myxalamide B (*7*). These results were surprising since inactivation of branched-chain ketoacid dehydrogenase complex was expected to preclude formation of the isobutyryl-CoA starter unit (from valine). When we compared production of myxalamide B by HPLC in DK1622 and DK1878 grown in the presence of CNCbl, almost no difference was seen (data not shown). This indicated that the primary role of IcmF in DK1622 is not the production of isobutyryl-CoA for myxalamide B synthesis.

Interestingly, tert-butyl group found in pivalic acid is a part of apratoxins (Figure 5.7) and laingolides isolated from different strains of cyanobacterium *Lyngbya bouillonii (12-14)*. Although it was suggested that tert-butyl thioester (pivalyl-CoA) is coming from decarboxylation of malonyl-CoA where three methyl groups are donated by S-adenosyl-L-methionine (SAM) (*13*), it is also possible that in complex microbial assemblage pivalyl-CoA is a direct product of pivalyl-CoA mutase activity (see Chapter 3).

5.3.4 Role of the IcmF-catalyzed reaction in metabolism of *R. eutropha.* In collaboration with the group of Dr. Antony Sinskey (MIT), we are using *Ralstonia eutropha* as a model organism, to test the hypothesis that IcmF is involved in fatty acid metabolism. Currently, *icm*F has been knocked out and initial feeding studies are being

performed. Samples containing WT *Ralstonia eutropha* (Re2302) and *icm*F knockout strain (Re2303) were subjected to metabolome analysis in which >300 compounds are detected using LC-MS/MS and GC-MS platforms. Preliminary results show that when the Re2303 strain is compared to the WT strain, significant accumulation of different dipeptides containing leucine, isolucine and valine is detected (~1.8-7.7 fold of change) (Figure 5.8).

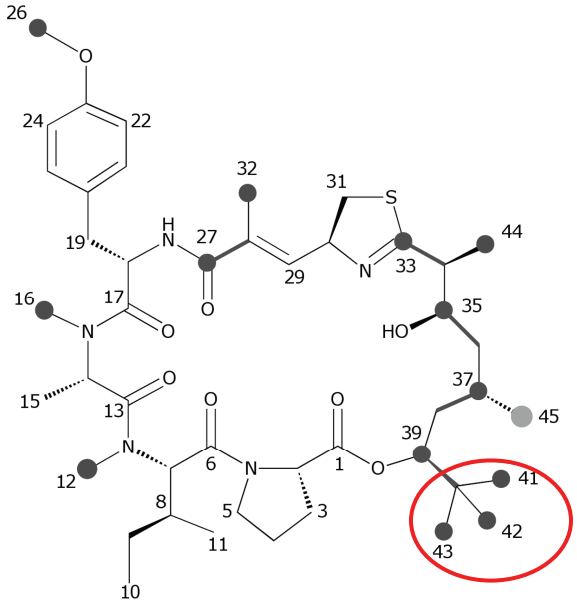


Figure 5.7: Overall structure of Aprtatoxin A. Tert-butyl group is marked with a red oval**.**Taken from (13).

In bacteria, branched-chain amino acids (leucine, isoleucine and valine) are converted by the action of the branched-chain ketoacid dehydrogenase complex to isovaleryl-CoA, 2-methylbutyryl-CoA and isobutyryl-CoA, which are precursors of branched chain fatty acids (*7*). Isovaleryl-CoA is also a substrate for IcmF, which converts it to pivalyl-CoA (Section 3.4.2). Thus, it is tempting to speculate that IcmF is involved in the metabolism of branched fatty acids. We are planning to perform fatty acid profiling both in WT and

the Re2303 mutant strain and hope that this analysis will shed more light on the function of IcmF in the bacterial cell.

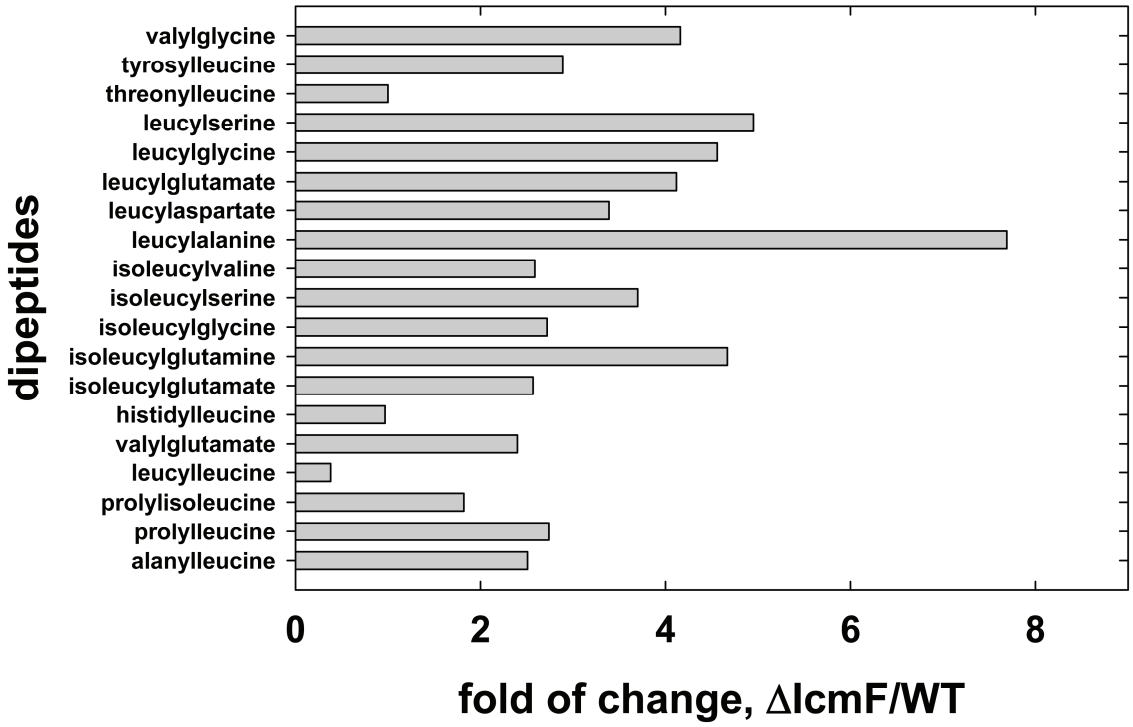


Figure 5.8: Accumulation of dipeptides in ΔicmF strain of *R. eutropha*. Statistical analysis of the data represents the fold change in each dipeptide in the Δ*icm*F strain compared to the wild-type strain ($0.001 < p < 0.0285$). It can be seen that in the Δ*icmF* strain, dipeptides containing branched chain amino acids (leucine, isoleucine and valine) accumulate several fold. These data were generated in the laboratory of Dr. Antony Sinskey by Dr. Chris Brigham (MIT).

5.3.5 How do the mutase domains of IcmF signal to MeaI? Prior to obtaining structural insights into the surface of the two mutase domains (IcmA and IcmB) that interact with MeaI in the IcmF fusion, we used the published data on MCM and MeaB from *M. extorquens* to guide our studies (*11*). The R585C mutation impairs the GAP activity of MCM and leads to loss of the proofreading activity of MeaB (*11*). Based on these effects, it was speculated that R585 might function as an arginine finger residue to "complete" the MeaB active site (*11*). R585 resides in a cluster of four solvent exposed

arginines at the tip of the B_{12}-binding domain of MCM with the others being R582, R583 and R663 (Figure 5.9).

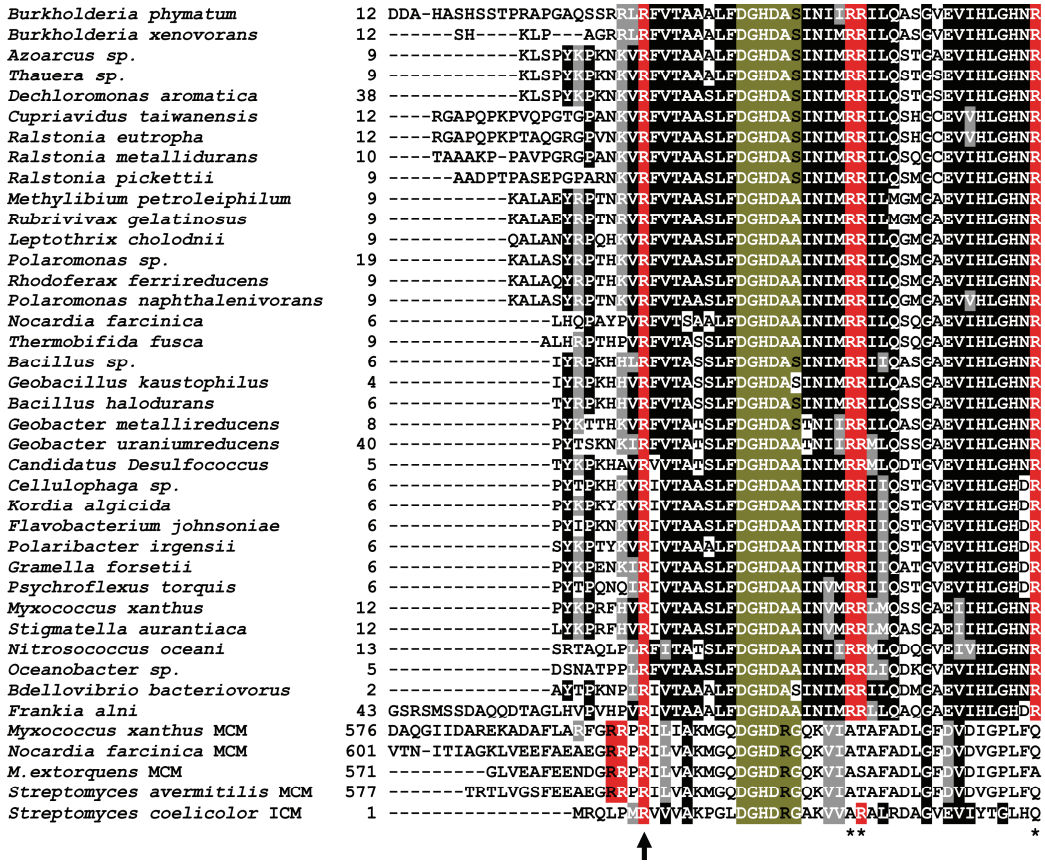


Figure 5.9: Conserved arginine residues in the B_{12}-binidng domain of IcmFs. Multiple sequence alignment of IcmFs, MCMs from *M. extorquens* (YP_001642233), *Streptomyces avermitilis* (NP_823216), *Myxococcus xanthus* (YP_630484), *Nocardia farcinica* (YP_119678) and ICM from *Streptomyces coelicolor*. For accession number of IcmFs see Table 2.1.

This tetra-arginine cluster is highly conserved in MCMs but does not align well when the MCM sequence is compared to the B_{12}-binding domain of different IcmFs (Figure 5.9). Nevertheless, the B_{12}-binding domain of IcmF also has a cluster of four conserved arginines (R12, R31, R32 and R48, numbering of *Gk* IcmF). To test our hypothesis that an arginine residue in the B_{12}-binding domain activates the G-protein domain, all four arginine residues were mutated. Surprisingly, the GTPase activities of all four mutants

were almost identical to that of WT *Gk* IcmF (Table 5.3). With the availability of the IcmF structures, the mutagenesis results can be explained. All four arginine residues are distant from the MeaI active site and are unlikely to participate in communication between domains[9].

protein	k_{cat}, min^{-1}
WT	10 ± 2
R12C	11.6 ± 3
R31A	12 ± 1
R32A	9.9 ± 1
R48A	11 ± 3

Table 5.3: GTPase activity of wild-type *Gk* IcmF and of the arginine mutants.
All experiments were performed in Buffer A with 10 mM $MgCl_2$ at 37 °C as described under Experimental Procedures. Values represent the average of at least three independent experiments.

5.3.6 IcmF structure determination. In collaboration with Dr. Catherine Drennan's laboratory at MIT, we recently obtained the first crystal structure of IcmF from *C. metallidurans,* which provides unprecedented insights into how the MeaI and the ICM domains interact. Currently, we have a structure of holo-IcmF with bound GDP at 3.5 Å resolution. Additionally, we have crystals of apo-IcmF and holo-IcmF with isobutyryl-CoA/n-butyryl-CoA and isovaleryl-CoA bound. These structures will provide a useful framework for interpreting biochemical studies designed to probe conformational

[9] Currently we are using the structure of full length IcmF to guide our studies.

dynamics and the structural basis for the cross-talk between the mutase and GTPase domains of IcmF.

We plan to analyze in more detail the structure of IcmF with different substrates bound (pivalyl-CoA, isovaleryl-CoA, n-butyryl-CoA and isobutyryl-CoA). Differences between structures of holo-IcmF with substrates could provide information on the inactivation observed with the butyrate thioesters (Section 3.4.4 and Section 5.3.2). Unfortunately, up to date, apo-IcmF/apo-IcmF with nucleotides crystals diffract very badly >20Å so we are planning to further optimize the conditions.

We recently succeeded in cloning ATR from *C. metallidurans* (Section 4.2.1), which delivers AdoCbl to IcmF. We will therefore attempt to co-crystallize ATR and IcmF to generate a complex that contains all proteins necessary for AdoCbl delivery to a B_{12}-dependent mutase.

5.3.7 Final remarks. My thesis research significantly contributed to broadening our understanding of vitamin B_{12}-dependent enzymes that support bacterial metabolism in specific niches and natural products synthesis. Our studies provided insights into the complex structures of a cobalamin-dependent mutase and its G protein chaperone that is pertinent to other B_{12}-mutase/chaperone duos including those found in humans and associated with disease. The G protein chaperone in IcmF is a member of a larger family of proteins involved in the assembly of several important metalloenzymes such as nickel hydrogenases and urease. However, the mechanisms of metal delivery and cluster assembly are largely unknown. Our studies on IcmF, which are simplified by the fusion between the G-protein domain and the client enzyme, serve as a paradigm for understanding intracellular metal trafficking and chaperone function. Since dysfunction

of these processes is involved in a number of diseases, a deeper understanding of IcmF function has the potential to lead to therapeutic strategies. To our knowledge the IcmF structure we obtained is the first example of a crystal structure of a fusion between an enzyme and its chaperone.

5.4 References

(1) Padovani, D., Labunska, T., Palfey, B. A., Ballou, D. P., and Banerjee, R. (2008) Adenosyltransferase tailors and delivers coenzyme B12. *Nat Chem Biol 4*, 194-6.

(2) Cracan, V., Padovani, D., and Banerjee, R. (2010) IcmF is a fusion between the radical B12 enzyme isobutyryl-CoA mutase and its G-protein chaperone. *J Biol Chem 285*, 655-66.

(3) Gerth, K., Jansen, R., Reifenstahl, G., Hofle, G., Irschik, H., Kunze, B., Reichenbach, H., and Thierbach, G. (1983) The myxalamids, new antibiotics from Myxococcus xanthus (Myxobacterales). I. Production, physico-chemical and biological properties, and mechanism of action. *J Antibiot (Tokyo) 36*, 1150-6.

(4) Ohlendorf, B., Kehraus, S., and Konig, G. M. (2008) Myxochromide B3, a new member of the myxochromide family of secondary metabolites. *J Nat Prod 71*, 1708-13.

(5) Wenzel, S. C., Kunze, B., Hofle, G., Silakowski, B., Scharfe, M., Blocker, H., and Muller, R. (2005) Structure and biosynthesis of myxochromides S1-3 in Stigmatella aurantiaca: evidence for an iterative bacterial type I polyketide synthase and for module skipping in nonribosomal peptide biosynthesis. *Chembiochem 6*, 375-85.

(6) Dayem, L. C., Carney, J. R., Santi, D. V., Pfeifer, B. A., Khosla, C., and Kealey, J. T. (2002) Metabolic engineering of a methylmalonyl-CoA mutase-epimerase pathway for complex polyketide biosynthesis in Escherichia coli. *Biochemistry 41*, 5193-201.

(7) Bode, H. B., Meiser, P., Klefisch, T., Cortina, N. S., Krug, D., Gohring, A., Schwar, G., Mahmud, T., Elnakady, Y. A., and Muller, R. (2007) Mutasynthesis-derived myxalamids and origin of the isobutyryl-CoA starter unit of myxalamid B. *Chembiochem 8*, 2139-44.

(8) Ortiz-Guerrero, J. M., Polanco, M. C., Murillo, F. J., Padmanabhan, S., and Elias-Arnanz, M. (2011) Light-dependent gene regulation by a coenzyme B12-based photoreceptor. *Proc Natl Acad Sci U S A 108*, 7565-70.

(9) Perlova, O., Gerth, K., Kuhlmann, S., Zhang, Y., and Muller, R. (2009) Novel expression hosts for complex secondary metabolite megasynthetases: Production of myxochromide in the thermopilic isolate Corallococcus macrosporus GT-2. *Microb Cell Fact 8*, 1.

(10) Wenzel, S. C., Meiser, P., Binz, T. M., Mahmud, T., and Muller, R. (2006) Nonribosomal peptide biosynthesis: point mutations and module skipping lead to chemical diversity. *Angew Chem Int Ed Engl 45*, 2296-301.

(11) Padovani, D., and Banerjee, R. (2009) A G-protein editor gates coenzyme B12 loading and is corrupted in methylmalonic aciduria. *Proc Natl Acad Sci U S A 106*, 21567-72.

(12) Matthew, S., Salvador, L. A., Schupp, P. J., Paul, V. J., and Luesch, H. (2010) Cytotoxic halogenated macrolides and modified peptides from the apratoxin-producing marine cyanobacterium Lyngbya bouillonii from Guam. *J Nat Prod 73*, 1544-52.

(13) Grindberg, R. V., Ishoey, T., Brinza, D., Esquenazi, E., Coates, R. C., Liu, W. T., Gerwick, L., Dorrestein, P. C., Pevzner, P., Lasken, R., and Gerwick, W. H.

(2011) Single cell genome amplification accelerates identification of the apratoxin biosynthetic pathway from a complex microbial assemblage. *PLoS One 6*, e18565.

(14) Tidgewell, K., Engene, N., Byrum, T., Media, J., Doi, T., Valeriote, F. A., and Gerwick, W. H. (2010) Evolved diversification of a modular natural product pathway: apratoxins F and G, two cytotoxic cyclic depsipeptides from a Palmyra collection of Lyngbya bouillonii. *Chembiochem 11*, 1458-66.