Structural investigations of
adenosylcobalamin-dependent enzyme maturation

by

Francesca A. Vaccaro

B.A. Biochemistry
Loyola University New Orleans, 2016

Submitted to the Department of Chemistry
In Partial Fulfilment of the Requirements for the Degree of

Doctor of Philosophy in Chemistry

at the

Massachusetts Institute of Technology

June 2023

Signature of Author _____
Francesca A. Vaccaro
Department of Chemistry
May 12th, 2023

Certified by _____
Catherine L. Drennan
Professor of Biology and Chemistry
Howard Hughes Medical Institute Investigator and Professor
Thesis Supervisor

Accepted by _____
Adam P. Willard
Associate Professor of Chemistry
Graduate Officer

This doctoral thesis has been examined by a committee of the
Department of Chemistry as follows:

Professor Elizabeth M. Nolan _____
Thesis Committee Chair
Ivan R. Cottrell Professor of Immunology
Professor of Chemistry
Associate Head, Department of Chemistry

Professor Catherine L. Drennan _____
Thesis Supervisor
Professor of Biology and Chemistry
Howard Hughes Medical Institute Investigator and Professor

Professor Joseph H. Davis _____
Thesis Committee Member
Assistant Professor of Biology

Structural investigations of adenosylcobalamin-dependent enzyme maturation

By
Francesca A. Vaccaro

Submitted to the Department of Chemistry on May 12th, 2023 in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Chemistry

Abstract

Metalloenzymes utilize metallocofactors, ranging from single metal ions to complicated metallic clusters, to catalyze a wide range of challenging chemical reactions that are critical for life. Incorporation of these metallocofactors often relies on proteins known as metallochaperones that transport, modify, and/or insert the metallocofactor for its target metalloenzyme. This thesis focuses on the maturation of methylmalonyl-CoA mutase (MCM). In humans, MCM is the only known adenosylcobalamin (AdoCbl)-dependent enzyme; mutations or deletions of MCM or any metallochaperones involved in its maturation lead to methylmalonic aciduria, an inborn error of metabolism. The final step of MCM's maturation involves an adenosyltransferase (ATR), which catalyzes the adenylation reaction to form AdoCbl and then delivers AdoCbl to MCM, and a G-protein chaperone, which facilitates AdoCbl delivery by the ATR through GTP hydrolysis. In addition to the human system, there are two bacterial systems used to understand the maturation of MCM: a homologous three-component system from *Methylobacterium extorquens* and an analogous two-component system from *Cupriavidus metallidurans.* The two-component system consists of the natural fusion protein, IcmF, which contains the AdoCbl-dependent isobutyryl-CoA mutase and its corresponding G-protein chaperone, and the analogous ATR. This thesis provides structural and biochemical characterizations of these two model systems to understand how the metallochaperones, specifically the G-protein chaperones, enable efficient mutase maturation. We present the crystal structure of a minimal system consisting of the G-protein chaperone, MeaB, and the Cbl-binding domain of the MCM from *M. extorquens.* This structure trapped an active conformation of the G-protein chaperone, revealing the first snapshots of the 180° rotation of one protomer needed to complete the nucleotide binding site and perform GTP hydrolysis. We also present mutagenesis and solution state data for IcmF from *C. metallidurans* that characterize its nucleotide- and cofactor-state dependent oligomerization, important for cofactor loading and unloading. Using cryogenic electron microscopy, we obtain structural data on IcmF that show that the monomeric G-protein domains of IcmF dimerize to resemble the active conformation of dimeric MeaB, and that this "active" conformation of the G-protein domains physically props open the mutase domains to enable AdoCbl loading. Finally, we present the crystal structure of *C. metallidurans* ATR and use this structure in computational docking studies with *C. metallidurans* IcmF to probe the potential interfaces within a G-protein:ATR:mutase complex. Overall, this work deepens our understanding of the function of G-protein chaperones in the maturation of AdoCbl-dependent mutases and sets the stage for further studies of metallochaperones and their roles in metalloenzyme maturation. Collectively, these studies have implications in both human health and in biotechnology.

Thesis Supervisor: Catherine L. Drennan
Title: Professor of Biology and Chemistry; Howard Hughes Medical Institute Investigator and Professor

Acknowledgements

So many people have accompanied me on this PhD journey at MIT, without you all this would have not been possible. The communities I have had the honor of being a part of and the people that I have met inspired me and supported me as I dreamt big. This accomplishment is as much yours as it is mine, and I would have it no other way.

Before I ever set foot on MIT's campus, I was lucky to have an advisor that saw my research potential long before I did. Thank you, Dr. Koplitz, for convincing me to "just try out research" back in my first year of college at Loyola. Without you, I can definitively say that I would not be defending my dissertation at MIT. Thank you.

To Cathy, thank you. I showed up to MIT uncertain if I belonged here, uncertain of what it meant to be a graduate student, and uncertain of what I was interested in studying. You took a chance on me since I had never ever touched a protein. Thank you for creating a lab space that always allowed me to be a whole person in this space, not only a scientist. You always supported me as I pursued experiences that would take away from research output but were crucial to my development as a scientist. Thank you.

Also, a huge thank you to my committee. Liz Nolan, you were the person that first told me that I had been accepted to MIT all those years ago. Your scientific expertise and precision have helped me develop my scientific knowledge and intuition. I am grateful for our conversations about science and more; your advice and guidance has helped orient myself over and over again. Joey Davis, I have always appreciated our conversations about science and life, whether in building 68 or in front of your house as I walk the dogs. Your guiding questions and help, specifically in cryo-EM, were instrumental for my progress in answering my research questions. Thank you for all the support as I navigated this space.

To the Drennan lab, past and present, I could write a novel about how important you all have been to my growth as a scientist and a person. Thank you, David, for pitching the B_{12} project and teaching me everything that I needed to know about proteins to become independent. Manny, Paty, Emily, Dana, and Kelsey, you all were instrumental in the beginning (and middle and end) of my PhD. Thank you for answering all my constant questions about science and for providing a snapshot of life after graduate school. To Kenny, Percy, Tsehai, Rohan, Allena, and Dawson, thanks for cultivating a lab culture that showed me how to balance work and lots of fun as a graduate student; it got me off to a great start. To Lindsey, thanks for convincing me to join the Drennan Lab in those early conversations; how you portrayed the lab made me know it was the lab I wanted to join. To Steve, thank you for those morning speed work outs, keeping me up to date on all the chemistry gossip, and always providing critical feedback (with many hot takes sprinkled in). Sheena and Talya, I can't believe we are finally at the end; thanks for sticking with me throughout the ups and downs. To Gerardo, Dante, Alison, Andy, Sorin, and Andrea, you all have provided me with so many great moments that make me want to come to lab even when nothing is working; I am rooting for all of you as you continue your PhD journeys. To Gisele and Christa, thank you for being my pottery friends in addition to my lab friends. I have treasured our time with Darrell and can't wait to see all the things you do in the future (both in science and pottery). Becky, thanks for always listening to my frustrations regarding our communal spaces and for helping me figure out work arounds for all the sensitive needs of all my different proteins. To Naike, working

with you has been an adventure; thank you for never taking me too seriously and reminding me of how exciting research is. To Alex, thanks for taking up the B_{12} project; I'm excited for all the things you discover and promise to try not to hover too much. To Mary, there is no one else I would have wanted to start (and almost end) my time in the Drennan lab with. You have mentored me every step of the way. Thank you for the morning runs and constant conversation on topics ranging from how to plan bachelorette parties to how to facilitate an inclusive lab space. You are a brilliant scientist and an incredible mentor; your future lab is lucky to have you. Last but definitely not least, thank you Daphne. I could not have predicted what life would throw at both of us when you started working with me. Although chaotic most days, I could not have done this without your inquisitive mind, willingness to do complicated experiments (in the dark, in the box), and unwavering support. Mentoring you was one of the best parts of graduate school; I can't wait to see what you accomplish in the future.

Outside of the Drennan lab, there were so many people at MIT that shaped my time here. First off, to members of the 2017 cohort of MIT Chemistry, thanks for letting me be the unofficial social chair of our cohort and helping me create community in this space. Special shout out our very small chem bio cohort; the shenanigans of our first year made grad school so interesting even if that meant staying up all night to turn in proposals. Thank you, Ed, Soop, Dio, Tony, and Arun, for your friendship and continuing to invite me to hang out even when it was unlikely I would show up. To Grace, our long conversations in the hallway connecting building 18 to building 56 sustained me more than I can even explain. Thank you for helping me process this wild experience. I hope that we can still connect over very long lunches in the future. Meghann, thank you for all the space you have held for me as I navigated the struggles of graduate school and life. To my mentees in the chemistry department, Mani, Janet, and Shirley, the three of you accompanied me on this journey and helped remind me of what I love about science. Thank you for the conversations that inspired me and taught me over the years. To the founding members of CADI, you all helped me find my place and purpose in this space. I am so grateful for the community we shared and hope that CADI continues to be a place where members of the department can be whole people. To the members of DPhiE sorority that I had the privilege of knowing, thank you for picking me to be your GRA. You all welcomed me into your community with open arms and taught me all the required vocabulary. Thank you for always trying your best to not wake me up in the middle of the night and for giving me the space to be me. Being your GRA was not easy, but all of you made it so fun.

There are so many people outside of MIT that have sustained me through this journey; it has truly been a wild ride. To Rachel, Shravya, and Ayesha, thank you for reminding me that I like to have fun and forcing me to have it, despite my protesting. I am grateful that our OG Sisterhood has weathered the distance; I can't wait to see where life takes us all. To James, thank you for convincing me that Boston weather wasn't *that* bad. You offered up your family, when mine was thousands of miles away. You forced me to take breaks to go for runs and to concerts. Thank you. To Michael, no matter how hard it was to connect sometimes because of the time difference, our WhatsApp conversations helped me navigate MIT, empowered me to advocate for myself always, and reminded me to be true to myself and what I want. To the members of Casa Iggy, you all showed me what community and love looks like even if we are not physically close. Thank you, Shelly, for sending me dog tiktoks, coming to explore Boston with me, and holding space

for me to explore who I am outside of science. Thank you, Jacquelyn (and Alex), for making wherever we met up feel like home and providing a respite from the difficulties of graduate school. I can't wait to explore Durham with you, Banjo, and baby Mestre. Thank you, Rachel (and Kade), for always feeding me, watching bad reality tv with me, and letting me hang out with some of my favorite pups; I'm so glad that Talya didn't keep us apart because our friendship has been instrumental to surviving these past few years. Talya, we've been through a lot together, and I am so glad we hung on for dear life. Here's to making it to the end of our PhDs; I can't wait to see what is next for us. To the Andorfer-Gair family, thank you for being my Boston family. From officiating your wedding to picking Joanie up from daycare to crashing family vacations, you all have provided a respite from graduate school, literal days of processing my time at MIT, and so much laughter in between. A big shoutout to the various roommates I have had over the years, Isa, Trish, Rachel, Samira, and Lyla, that have reminded me to take a break and made Boston feel a little bit more like home. An extra special shoutout to my favorite roommate, Cleo, thanks for always cuddling with me, helping me improve my ball throwing skills, and being Ricky's best friend.

To my family, despite having no idea what I am doing most of the time, you all have provided the respite from graduate school and reminding me who I am and where I came from. Mother Teresa, you have provided a listening ear, reminding me to take breaks, eat, and sending pictures of Pepe. Dad, thank you for always helping me brainstorm ways to cheat if I needed to. Luckily, it never came to that. Giuliana, thanks for being my ally, letting me be a science fair judge, and showing me how to teach and mentor. You constantly inspire me and remind me of the joy in teaching and learning. Zia Carmela, thank you for always indulging my love for spoiling dogs. To Zia Enza, thank you for always checking in about the weather here. Titi Zuli, thank you for always reminding me that what I am doing is a big thing and that it is going to be hard sometimes. To Grandma, living closer to you these past few years have been a gift that I will treasure. Nonna and Nonno, grazie per amarmi sempre, per credere sempre in me, per inviarmi sempre pacchi pieni di tutte le cose che mi ricordano casa. Non avrei potuto farlo senza di te e le nostre videochiamate. Ho fatto questo per te; puoi finalmente chiamarmi dottore. Finally, a big thank you to Ricardo Cannoli aka Ricky, the pandemic pup that changed my life and melted my heart. You constantly reminded me to slow down, take in my surroundings, cuddle, and rest, all things that I definitely did not do enough of during my PhD.

Table of Contents

CHAPTER I: Introduction to metallochaperones and their roles in metalloenzyme maturation

This chapter is adapted from the following publication:

Vaccaro, F. A.; Drennan, C. L., The role of nucleoside triphosphate hydrolase metallochaperones in making metalloenzymes. *Metallomics* **2022,** *14* (6), mfac030.

Author contributions

Francesca A. Vaccaro wrote the original draft of all sections in this review. F. A. Vaccaro and Catherine L. Drennan contributed to editing of the manuscript.

Summary

Metalloenzymes catalyze a diverse set of challenging chemical reactions that are essential for life. These metalloenzymes rely on a wide range of metallocofactors, from single metal ions to complicated metallic clusters. Incorporation of metal ions and metallocofactors into apo-proteins often requires the assistance of proteins known as metallochaperones. Nucleoside triphosphate hydrolases (NTPases) are one important class of metallochaperones and are found widely distributed throughout the domains of life. NTPase metallochaperones use the binding and hydrolysis of nucleoside triphosphates, either adenosine triphosphate (ATP) or guanosine triphosphate (GTP), to carry out highly specific and regulated roles in the process of metalloenzyme maturation. Here, we present recent literature on NTPase metallochaperones and describe the current mechanistic proposals and available structural data. By using representative examples from each type of NTPase, we also illustrate the challenges in studying these complicated systems. We highlight open questions in the field and suggest future directions, focusing specifically on the delivery of adenosylcobalamin to its target mutase.

Introduction

Metal sites exist in about one third of all structurally characterized enzymes, and over half of all proteins are predicted to be metalloproteins[1]. These metal sites range in complexity from one metal ion to multi-metal or even organometallic clusters (*Figure I-1*)[2]. Metalloenzymes tend to catalyze the most challenging chemical transformations. Their roles include, but are not limited to, respiration[3], photosynthesis[4], regulation of transcription and translation[3], and nitrogen[5], carbon[6], and hydrogen fixation[7]. To produce a functioning metalloenzyme, the metallocofactor must be correctly installed, which often requires a metallochaperone. Metallochaperones are proteins that physically interact with apo-metalloprotein clients or intermediary proteins to assist in cofactor delivery or assembly as part of the process of metalloprotein maturation. The exact percentage of metalloproteins that require metallochaperones for cofactor biogenesis is not established. Quantification is complicated by the fact that the metallochaperones are often not needed for reconstitution in vitro but are in vivo, where metal ion concentrations are often limiting

due to toxicity issues. In the cell, metallochaperones balance out the cellular demand for metal-assisted reactivity with the toxicity associated with having too much metals[8].

Deletion or mutation to genes encoding metallochaperones can lead to deleterious biological consequences. In *Azobacter vinelandii*, deletion of the metallochaperones NifX and NafY decreases the amount of active nitrogenase by half, reducing levels of dinitrogen reduction and stunting organismal growth[9, 10]. In humans, the metabolic disorder methylmalonic aciduria is caused by mutations to or deletion of the gene for methylmalonic aciduria type A protein (MMAA), a metallochaperone required for the maturation of adenosylcobalamin-dependent methylmalonyl-CoA mutase (MCM). Without the adenosylcobalamin cofactor, methylmalonyl-CoA cannot be converted to succinyl-CoA, causing an accumulation of methylmalonic acid that alters the blood pH leading to disease[11, 12]. Because of the deleterious effects of metallochaperone impairment, bacterial metallochaperones are also potential drug targets. For example, the metallochaperone UreG is necessary for the maturation of the dinickel metal active site of urease, an enzyme that catalyzes the hydrolysis of urea to ultimately yield ammonia and carbon dioxide, providing buffering capacity necessary for *Heliobacter pylori* to live in the stomach. As such, urease is central to *H. pylori* metabolism and virulence[13, 14], making its metallochaperones potential drug targets.

The wide variety of metal centers employed by metalloproteins requires a diversity of metallochaperones. There are numerous types of metallochaperones that have been characterized, from the transporters of the FeMo-cofactor of nitrogenase[15, 16] to the copper metallochaperones involved in shuttling copper ions to the mitochondrial electron transport chain for cellular respiration[17]. This chapter focuses on one group of metallochaperones: nucleoside triphosphate hydrolases (NTPase) metallochaperones. NTPase metallochaperones use the binding and/or hydrolysis of nucleoside triphosphates (either GTP or ATP) for metalloenzyme maturation. Although NTPase metallochaperones[18] are ubiquitous, their exact roles in maturation processes are often not well understood. Known and postulated roles include: metal ion or metallocofactor transport to an intermediary protein in the metalloprotein maturation process; metallocofactor binding and direct insertion into apo-target protein; and induction of a conformational change in the target protein to allow for cofactor delivery. Here, we focus

on the recent advances in the study of NTPase metallochaperones, describing the current state of knowledge and highlighting open questions.

P-loop G3E GTPase metallochaperones

The first subclass of NTPase metallochaperones that we will consider are phosphate-binding loop (P-loop)-containing G3E GTPases (hydrolyzing guanosine triphosphate, GTP)[18]. The P-loop G3E GTPase family is named for a D-to-E substitution in the G3 motif compared to the canonical P-loop GTPase, Ras[18]. The best characterized subfamilies of the P-loop G3E GTPase metallochaperone family are the urease metallochaperone UreG[19, 20], the [NiFe]-hydrogenase metallochaperone HypB[21], and the cobalamin metallochaperone MeaB (MMAA in humans)[22-24]. In addition to UreG, HypB, and MeaB/MMAA, there is a large and diverse subfamily of G3E GTPases called cluster of orthologous groups (COG) 0523 proteins (*Table I-1*). COG0523 proteins are characterized by a conserved CxCC (C = Cys; x = any amino acid) motif that is implicated in high affinity metal binding[25, 26]. Known functions of COG0523 proteins include: zinc homeostasis (ZigA/ZagA)[27]; cobalamin cofactor biosynthesis (CobW)[28-30]; and nitrile hydratase maturation (Nha3)[31].

P-loop G3E GTPases employ a common protein fold

All the known structures of the G3E P-loop GTPases contain a G-domain (*Table I-2*) that is comprised of a very common protein fold[32]: regularly recurring α-β units with the β strands forming a central β-sheet surrounded on both sides by α-helices (*Figure I-2*). The typical G-domain contains 5 different conserved motifs, known as G1-5 that are involved in nucleotide binding and hydrolysis (*Figure I-2*). The G1 motif, also known as the Walker A motif, is a flexible loop in between a helix and sheet. This motif functions to position the triphosphate of the bound nucleotide. The G2 motif, also known as switch I, signals which nucleotide, if any, is bound. The G3 motif, also known as the Walker B motif loop, is often situated at the end of a strand and contains a conserved aspartate or, less commonly, glutamate residue that binds the water bridged Mg^{2+} ion used for NTP hydrolysis[32]. The G4 motif typically contains the sequence motif NKXD that is used in the recognition of the guanosine base[18]. Finally, the G5 motif is involved in nucleotide release[32]. The ubiquitous nature of P-loop NTPases has led to further classification of

various structural motifs that are beyond the scope of this review but have been previously detailed by Leipe et al.[18].

The mechanism of GTP hydrolysis is well established

The mechanism of nucleoside triphosphate hydrolysis by various GTPases has been extensively studied[32-36]. Briefly, after nucleotide triphosphate binding, the active site is completed by a Mg^{2+}-bound water molecule positioned by an aspartate (or less commonly glutamate) to perform the hydrolysis of the terminal phosphate moiety of the nucleoside triphosphate. After the cleavage of the bond to the γ phosphate, the resulting nucleotide is a diphosphate; when replaced with a nucleoside triphosphate, the cycle can proceed again. Oftentimes, this binding and/or cleavage occurs in the presence of another protein known as an activating factor that increases the intrinsic rate of hydrolysis of the GTPase. In the case of metalloprotein maturation, the activating factor can be the target enzyme or another protein involved in the maturation process. Similarly, a nucleotide exchange factor is sometimes needed to replace the nucleoside diphosphate with the corresponding triphosphate[32]. The changes throughout this cycle form the basis for the ability of the GTPases to perform their chaperone roles in the maturation of metalloproteins.

The well-studied family members UreG, HypB, and MeaB show mechanistic diversity

The most well studied family members of the G3E P-loop GTPase metallochaperones utilize different mechanisms for maturing their respective metalloenzymes[18]. As far as we know, the evolutionary factors that drove these mechanistic differences within this metallochaperone subfamily are not understood. These mechanistic differences are especially interesting given that both UreG and HypB are metallochaperones involved in the maturation of the nickel-dependent enzymes urease and [NiFe]-hydrogenase, respectively. UreG undergoes a conformational change when it binds GTP and accepts Ni^{2+} from UreE. Then, UreG forms a complex with apo-urease and the other accessory factors UreF, UreH, and UreD to directly insert the Ni^{2+} to its target enzyme, urease, through a tunnel created by the complex of all the accessory factors (*Figure 1-3A*)[14, 37, 38]. Although HypB is also conformationally gated by GTP binding and hydrolysis, it does not directly insert Ni^{2+} into the [NiFe]-hydrogenase active site. The Zamble lab showed that it is the GDP-loaded state of *Escherichia coli* HypB that

is “readied” for fast Ni^{2+} transfer to accessory protein HypA, which ultimately leads to Ni^{2+} insertion into the active site of [NiFe]-hydrogenase (*Figure I-3B*)[39]. However, not all GTPase metallochaperones directly bind the metallocofactor that they are involved in delivering. In the maturation of methylmalonyl-CoA mutase, the GTPase metallochaperone MeaB (in bacterial systems, MMAA in humans) is required for adenosylcobalamin delivery; however, there is no evidence of the chaperone interacting with the cofactor itself[24, 40-44].

COG0523 subfamily and its best characterized member, Nha3

Giedroc and coworkers recently provided a comprehensive comparative phylogenetic, biochemical, structural, and functional analysis of P-loop G3E GTPases in the COG0523 subgroup[45]. These proteins typically have two domains, a conserved G-domain at the N-terminus and a variable C-terminal domain that is predicted to be involved in target-specific protein interactions. The sequence analysis by Giedroc et al. suggests that there are multiple subfamilies within the COG0523 subgroup and that there is a considerable amount of unexplored sequence space. Many of the COG0523 sequence clusters have no characterized members. The lack of functional data is especially notable for COG0523 proteins from eukaryotic organisms. The best characterized COG053 member is Nha3, a Fe-type nitrile hydratase (NHase) maturase found in industrially useful microbes such as *Rhodococcus equi*[46]. Currently, it serves as the model system for this diverse GTPase metallochaperone protein family.

Although Co-type NHases do not appear to require a GTPase metallochaperone[31], Fe-type NHases do. NHases catalyze the hydrolysis of organic nitrile to the corresponding amide product (*Figure I-1*) using a metal ion (Co^{3+} or Fe^{3+}) that is buried at the interface between the α and β subunit[46, 47]. The maturation of Co-type nitrile hydratases relies on an α subunit swapping mechanism that requires a GTP-independent maturation protein[31]; whereas, the maturation protein for Fe-type NHases is a COG0532 GTPase metallochaperone (called ε or Nha3) (*Figure I-3C*). Briefly, the swapping mechanism for the Fe-type NHases appears to involve an Fe^{2+}-loaded GTP-bound Nha3 forming a complex with an apo-α subunit of Fe-type NHase[46, 47]. In a GTP-dependent manner, the α subunit receives the metal and is oxidized to Fe^{3+} in the α subunit. Consistent with other G3E P-loop GTPases, the presence of divalent metals bound increases the rate of GTP

hydrolysis of Nha3[46]. The apo-complex of the $\alpha_2\beta_2$ NHase swaps an apo-α subunit with a holo-α subunit to form the holo-NHase[46]. Substitutions of the conserved Lys or Thr of the P-loop of Nha3 from *Pseudomonas chlororaphis B23* result in an in vivo loss of detectable NHase activity similar to what has been observed for UreG and urease and HypB and [NiFe]-hydrogenase, further implicating GTPase activity of Nha3 in Fe-type NHase maturation[48].

ATPase metallotransporters and metallochaperones

Compared to GTPases, there is a wider variety in the types of characterized ATPases (hydrolysis of adenosine triphosphate, ATP) involved in metalloprotein maturation (*Figure I-4*, *Table I-3*)[49, 50]. The P1B class of P-type ATPases uses the energy of ATP hydrolysis to transport d-block metal ions across cell membranes. In plants, P1B class of P-type ATPases are key components in the maintenance of metal homeostasis, with characterized transporters for Zn^{2+}, Cu^{2+}, Cd^{2+}, Co^{2+}, and Pb^{2+} ions[49]. These transporters serve as the first step in metalloprotein maturation by acquiring d-block metal ions from the surrounding environments. Analogous to the P-loop GTPases described above, P-loop ATPases are also involved in transporting metal ions to the active site of their target protein. One well characterized P-loop ATPase metallochaperone is CooC[51], which is involved in the maturation of carbon monoxide dehydrogenases (CODHs)[52]. CODHs catalyze the reversible oxidation of carbon monoxide at a metallocofactor, the C-cluster, which consists of Ni, Fe, and S (*Figure I-1*). CooC is responsible for ATP-dependent Ni^{2+} insertion as part of the process of C-cluster maturation[51-53]. Additionally, ATPases containing a heat shock protein fold are known to be involved in the biogenesis and transport of various metallic clusters. For example, in the iron-sulfur cluster (ISC) pathway found in mammals, the ATPase HscA works with the co-chaperone HscB to stimulate the transfer of nascently synthesized [2Fe-2S] to various apo-enzymes in an ATP-dependent manner[50, 54].

Structures of P-loop ATPase CooC are similar to P-loop GTPase HypB

Comparisons of the structures of CooC, a P-loop ATPase, and HypB, a P-loop GTPase, reveal an incredibly similar overall architecture of α-β units, despite low sequence homology of around 25% (*Figure I-5A*). Both proteins contain the G1-G3

sequence motifs that are involved in binding the phosphates of the nucleotide and Mg^{2+} ion, although CooC contains the deviant G1 motif with the highly conserved lysine[55]. The deviant G1 motif has the highly conserved residue of a typical G1 motif in the second position instead of the second-to-last position in the sequence motif, which is a characteristic of the MinD/BioD class of P-loop NTPases[18]. The G4 nucleotide specificity motif in CooC is different from the conserved NKXD motif of HypB in that it only retains the first asparagine, and the CooC adenine base interacts with the residues of an alpha helix that is ordered upon binding the nucleotide (*Figure I-5B*)[55]. Additionally, these two proteins both have spatially separated but nucleotide coupled metal-binding sites that control the oligomeric state of the NTPases (*Figure I-5A*)[55].

Fe-S cluster biogenesis utilizes ATPases

Fe-S cluster biogenesis is a highly conserved yet complicated process that involves dedicated machineries to synthesize, transport and deliver Fe-S clusters. This dedicated machinery involves proteins that act as chaperones and/or "scaffolds" for the assembly and delivery of Fe-S clusters. Notably, many of these proteins have ATP binding sites and/or ATPase activity. There are four main Fe-S cluster biogenesis pathways: the SUF (sulfur fixation), the ISC (iron-sulfur cluster), the NIF (nitrogen fixation), and the CIA (cytosolic iron-sulfur cluster assembly) systems. The best characterized Fe-S cluster biogenesis pathway is the ISC system, which requires the HscA/HscB (also known as heat shock protein 70 (Hsp70)/J-protein) chaperone/co-chaperone ATPase complexes. Like most NTPases, these chaperone/co-chaperone proteins use nucleoside triphosphate binding and hydrolysis to regulate the transfer of a newly synthesized Fe-S cluster from a scaffold protein to an apo-recipient[56, 57]. The NIF system uses analogous chaperone/co-chaperone proteins but also requires an ATPase to deliver homocitrate-bound molybdenum during the process of assembling the FeMo-cofactor of nitrogenase; the roles of ATPases in nitrogenase maturation have been recently reviewed[5, 58]. In the SUF pathway, the ATP-binding cassette (ABC)-type ATPase SufC is necessary for Fe-S cluster formation[59]. SufB, which accepts sulfur, and SufD, which has been proposed to play a role in iron acquisition, have been shown to interact with each other in vitro[59, 60]. Along with SufB and SufD, deletions of SufC abolish SUF function in vivo. In vitro, the $SufBC_2D$ complex can serve as a scaffold for de novo Fe-S

cluster biogenesis[61]. Recent structural studies of the $SufBC_2D$ scaffold provide insight into how ATP binding to SufC may promote conformational changes that are necessary for formation of the cluster assembly site; however, a detailed mechanistic understanding of the ATPase cycle for cluster formation is not yet available[59, 62]. Finally, the CIA pathway uses a scaffold ATPase to assemble new Fe-S clusters; ATPase activity is required for both [2Fe-2S] cluster acquisition and transfer of the fully formed [4Fe-4S] cluster to the apo-client in the cytosol or nucleus[63-65]. Although many of the participants of these pathways have been identified, their biochemical characterizations are limited[57]. Recently, Perlstein and co-workers have provided a biochemical roadmap for exploring the roles of ATP binding sites and ATPase activity for proteins involved in Fe-S cluster biogenesis[66].

Nucleotide specificity of NTPase metallochaperones can be challenging to establish

Not all NTPase metallochaperones fall clearly into a GTPase subgroup or an ATPase subgroup. Although the better studied HypBs are GTPases, some organisms have versions of HypB that are ATPases[67, 68]. The HypB from the archaea *Thermococcus kodakarenis* is a P-loop ATPase containing the same overall fold as the characterized GTPase HypBs. Despite less biochemical characterization, there is evidence in some organisms that the HypB ATPases are functional homologs that replace the HypB GTPases with their activity also regulated by Ni^{2+}-binding, ATP hydrolysis, and interactions with HypA[68]. Also, the Feo iron transport system appears to utilize a NTPase called FeoB for which the nucleotide specificity has been unclear[69-75]. The Feo system is a dedicated Fe^{2+} ion transport system that is widely present in the archaeal and bacterial domains of life. It is composed of FeoA and FeoC, cytosolic proteins, and the transmembrane iron permease, FeoB. FeoB from the human pathogen *H. pylori* is thought to be an ATPase due to impaired Fe^{2+} ion transport when whole cells are treated with known inhibitors of ATP synthesis or hydrolysis[76]. In contrast, FeoB from *E. coli* is believed to be a GTPase because no ATP binding was observed in vitro[77]. The characterization of the FeoB from *Vibrio cholerae* showed both low intrinsic GTPase and ATPase activity that was not stimulated by any known factors[73, 74]. Research from Kim

and coworkers on various pathogenic bacterial FeoBs suggests that there are two classes of FeoBs: sole GTPases and promiscuous NTPases[75].

FeoB, and also HypB, highlight the difficulty in establishing nucleotide specificity for NTPase metallochaperones. It is clear in the case of HypB, and likely in the case of FeoB, that different organisms employ different NTPases (GTPases or ATPases). The reason(s) for this variation is not well established. However, it also seems to be the case that some metallochaperone systems have no variation; for example, there are no known MeaB metallochaperones that are ATPases. Yet in other systems, NTPases appear to be employed that are promiscuous. It is not known if this promiscuity is an in vitro or in vivo feature. Given that most NTPase metallochaperones are poor NTP hydrolases without their target protein and/or without the appropriate metal ion bound, one must be careful about drawing conclusions, both favorable and unfavorable, from low levels of NTP hydrolysis activity. Additional studies may reveal that all NTPases are specific under the right set of conditions or organismal diversity regarding the usage of a GTPase versus an ATPase. For example, a MeaB that is an ATPase may be discovered. It is currently not possible to predict nucleotide specificity from protein sequences, but more structural data will likely improve the prediction possibilities.

Metal specificity of NTPase metallochaperones can be challenging to establish

An important role of metallochaperones is to make sure that the correct metal ion is inserted into the correct apo-metalloprotein target, raising the question of how metal specificity of metallochaperones is established[78]. The biological challenge of correct protein metalation has been recently reviewed[79]. Briefly, one important factor in metalation is cellular metal availability, which is thought to be opposite of the Irving-Williams series, $Mg^{2+} < Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Cu^{2+} > Zn^{2+}$, with the weaker binding metals such as Mg^{2+}, Mn^{2+}, and Fe^{2+} more widely available and the tighter binding metals such as Ni^{2+}, Zn^{2+}, and Cu^{2+} less available[80-82]. The reducing/oxidizing environment in the cell is also a consideration in terms of availability. For example, Cu^{+} is believed to be the relevant form of copper in the reducing environment of the cytoplasm. Availability further depends on the ability of the organism to uptake metal ions, especially trace metals. In addition to availability, another factor affecting specificity is the metallochaperone's affinity

for different metal ions, a feature of the protein that can be modulated by GTP binding or hydrolysis. All of this points to the fact that establishing metal specificity of a metallochaperone can be challenging[83].

Establishing metal specificity for the GTPase metallochaperone CobW

The observation that *cobW* gene disruption in *Pseudomonas denitrificans* impairs aerobic cobalamin biosynthesis led to a proposal that CobW is a Co^{2+}-dependent metallochaperone[29, 30]. To test this proposal, affinities were measured for Co^{2+} and other metal ions for the GTPase-dependent CobW from *Rhodobacter capsulatus* in the presence and absence of nucleotide effectors[84]. Only weak interactions between CobW and Co^{2+} ions were observed in the absence of nucleotides. The addition of GTP or less hydrolyzable analogs promotes the tight coordination of two Co^{2+} ions in two different binding sites with different affinities. However, if Mg^{2+} ions are also present at physiological concentrations, the coordination of a Co^{2+} ion is observed in only one binding site. The other site is occupied by a Mg^{2+} ion. It is thought that the Mg^{2+} ion binds first in the weak affinity metal binding site, ordering the second binding site for high affinity binding of a Co^{2+} ion. When GDP is present instead of GTP, Co^{2+} binds CobW with a 1000-fold weaker affinity, indicating that an intact γ-phosphate is required for tight binding. The presence of bound GDP causes CobW to release the Co^{2+}, further tying the nucleotide state to the metalation state of CobW. Finally, when compared to other first row transition metals, Mg^{2+}GTP-CobW binds Zn^{2+} and Cu^{1+} more tightly[84]; however, using the idealized pool of bioavailable metals[85] and the free energy change for metal binding, calculations indicated that in vivo over 90% of Mg^{2+}GTP-CobW would be bound with Co^{2+}, as compared to less than 10% bound with Zn^{2+} due to the greater favorable free energy change for Co^{2+} binding[84]. This recent work firmly establishes Co^{2+} as the cognate metal for CobW, supporting its known involvement in cobalamin biosynthesis. Additionally, these studies confirm the presumed role of GTP binding and hydrolysis for metalation of the metallochaperone, like other G3E P-loop GTPases[84].

Final thoughts and future directions

The roles of NTPase metallochaperones continue to expand with new functions being established. However, there is much more work to be done. The COG0523

subfamily of G3E P-loop GTPases, for example, represents a poorly understood class of GTPase metallochaperones that are believed to bind and insert transition metals[45]. Many of these putative GTPases are uncharacterized and their target client proteins are unknown. Bioinformatic studies of the gene clusters containing these uncharacterized GTPases may allow for predictions of the target client proteins. Predictions should be followed by biochemical characterizations, exploring NTPase:target protein interactions, NTP specificity, and metal ion specificity. The latter can potentially be aided by the application of methods/calculations used in the CobW studies described above[84].

Genomic initiatives, and the wealth of sequence information that they generate, are leading to proposals of putative NTPase metallochaperones outside of the COG0523 family. For example, the protein MutS is well known for its involvement in DNA repair processes[86]; however, bioinformatics has revealed sequences of MutS-like proteins in operons associated with adenosylcobalamin-dependent enzymes, leading to a proposal that MutS variants might play a role in ATP-dependent metalloprotein complex assembly[87]. We hypothesize that there are likely other examples of ATPase families for which distinct cousins are involved in diverse biological processes. We are excited about how modern bioinformatics methods combined with genomic data will undoubtedly expand the NTPase metallochaperone field in the near future.

Among the already characterized NTPase metallochaperones, numerous questions remain. For example, it is often unclear whether the complete complement of stimulatory factors that increase NTP hydrolysis has been identified; *is another protein involved or maybe another metal ion*? Molecular mechanistic questions are also prevalent*; what exactly is the function of the nucleotide-state-dependent conformational change, and what type of conformational change occurs*? Structural information on GTPase metallochaperones is fairly limited (*Table I-2*), and even when structural methods have allowed for the capture of more than one nucleotide-bound state, it is often the case that crystal lattice contacts have prevented the conformational change from occurring[88]. Additionally, the absence of the target protein or other stimulatory factors may hinder the nucleotide-state-dependent conformational change of the NTPase from being fully realized[89]. In other words, obtaining the requisite structural snapshots of NTPases to understand their molecular mechanisms is not trivial. For the most part, these structural

data are missing for NTPase metallochaperones, leaving molecular mechanistic questions unanswered. For example, there is more to learn about the steps required to transfer nickel from the GTPase HypB to its partner metallochaperone HypA to complete the [NiFe] hydrogenase active site. The extent of the conformational changes that occur in HypB due to GTP binding and hydrolysis have not been visualized, limiting the molecular understanding of the mechanism[39, 90]. In the maturation of the adenosylcobalamin-dependent methylmalonyl-CoA mutase, the GTPase MMAA (MeaB in bacteria) is required for adenosylcobalamin insertion but does not bind the cofactor directly[24, 40-44, 91]. Although there are a number of structures available of MMAA/MeaB in various nucleotide-bound states[41, 89, 92] and even a structure of a MeaB-fusion protein in which the chaperone domain is covalently attached to the target enzyme[88], the molecular basis by which MMAA/MeaB facilitates adenosylcobalamin delivery is still in debate. Currently, none of the structural rearrangements observed explain the molecular basis of methylmalonyl-CoA mutase maturation[88, 92]. Cryo-electron microscopy (cryo-EM) represents a promising new direction for the obtainment of these requisite structural data as this method allows for structures of protein:protein complexes to be obtained and for multiple conformations of proteins to be more readily visualized. Although the NTPases by themselves are too small for cryo-EM, the protein:protein complexes involved in metalloprotein maturation should be sufficiently large[93], and ultimately, it is the structures of protein:protein complexes that are needed for a molecular understanding. Thus, the resolution revolution of cryo-EM, i.e. the revolutionary ability to determine near-atomic resolution protein structures by cryo-EM, represents an exciting prospect moving forward for the determination of structures of metalloprotein maturation machineries.

An improved understanding of metalloenzyme maturation has several possible industrial applications. For example, nitrogenases are attractive as environmentally friendly alternatives to the industrial Habor-Bosch process that is estimated to use ~1% of the world's energy. This alternative solution would be even more attractive if nitrogenases could be prepared at high levels with their complex metallocofactors correctly inserted (*Figure I-1*). Hydrogenases are appealing for use in biofuel cells, and carbon monoxide dehydrogenases for fixation of the greenhouse gas carbon dioxide. Again, these applications require that the metalloproteins be produced in high yield which

necessitates an understanding of cofactor biogenesis/delivery. Additionally, knowledge of metalloenzyme maturation processes could be exploited to deliver synthetic metallocofactors with altered reactivity to apo-target enzymes[94].

Regarding human health, understanding the molecular mechanisms of NTPase metallochaperones may provide novel solutions for therapeutics and treatments[11, 17, 95-97]. For example, methylmalonic aciduria, an inborn error of metabolism, is caused by deletions of or mutations to any of the numerous proteins that transport and/or insert adenosylcobalamin into methylmalonyl-CoA mutase, including MMAA. Since MMAA does not interact with the cofactor directly, understanding how MMAA uses GTP binding and hydrolysis to perform its gatekeeping role could provide a novel therapeutic strategy. Additionally, NTPase metallochaperones thought to be involved in virulence such as FeoB and UreG in *H. pylori* are additional drug targets[13, 76].

Recent developments in the NTPase metallochaperone field since the publication of this chapter

Since the publication of this chapter in early 2022, new information regarding the identity and mechanisms of action for NTPase metallochaperones has been published, expanding our understanding of their importance for metalloenzyme maturation both in vivo and in vitro. An ATPase with a novel nucleotide binding site has been identified for the Co-type NHase and a predicted eukaryotic Zn metallochaperone has been characterized[98, 99]. Here, we provide brief updates and elaborate on the importance of these new findings towards understanding NTPase metallochaperones.

Inspired by the identification and characterization of the GTPase activator for the Fe-type NHase from *Rhodococcus equi* TG328-2, an activator protein for the Co-type NHase from *Rhodococcus rhododchrous* J1 was sought after and identified[46]. This activator protein is an NTPase that increases the specific activity of the Co-type NHase in vitro[47, 98]. Despite the lack of a conserved NTPase binding site, residues impacting the rates of ATP hydrolysis were identified using sequence alignments and validated using site directed mutagenesis. Even with less than 20% of the wild type levels of ATP hydrolysis observed, the mutant activator proteins were still capable of increasing the specific activity of the Co-type NHase in vitro but at 64% percent of the wild type's levels.

Additionally, Co-type NHase activation was shown to be inhibited at ATP concentrations higher than 50 μM. Taken together, the characterization of a Co-type activator protein continues to expand our understanding of NHase maturation and highlights the challenges of identifying novel NTPases without complete knowledge of the nucleotide binding sites. However, the continual expansion of the motifs involved in NTPase hydrolysis will inform our ability to identify novel NTPase metallochaperones and maturation processes.

Despite their hypothesized existence, a vertebrate Zn-specific metallochaperone had never been biochemically validated until the characterization of the Zn-regulated GTPase metalloprotein activator (ZNG1), a G3E GTPase COG0523 protein[99]. A major hurdle to validating the mechanism of action of metallochaperones is lack of knowledge of interaction partners in vivo, preventing the ability to probe the maturation mechanism in vitro. Yeast-two-hybrid screens performed with ZNG1 in human, mouse and zebrafish identified methionine aminopeptidade 1 (METAP1) as a shared client across vertebrates. ZNG1 assists in METAP1 function in a Zn- and GTP-dependent manner, with METAP1 acting as a GTPase activating protein. The stimulation ZnZNG1 GTPase activity but not apo-ZNG1 GTPase activity by METAP1 supports the hypothesis that ZNG1 uses GTP hydrolysis to transfer Zn ions to the active site of METAP1 either through the energy generated by cleaving the phosphodiester bond of GTP or a conformational change. Additionally, the zebrafish and mouse models of *Zng1* deficiency link ZNG1 to mitochondrial function, organismal development, and Zn homeostasis, providing evidence for the importance of ZNG1 for survival during Zn deficiency. This study establishes ZNG1 as a Zn metallochaperone in vertebrates, expanding our understanding of the role of COG0523 G3E proteins in metalloprotein maturation.

Although these studies established ZNG1 as a GTP-dependent Zn metallochaperone and identified a novel NTPase involved in Co-type NHase activation, numerous questions remain regarding the molecular mechanisms of metal transfer and specific role of nucleotide binding and/or hydrolysis for maturation in these systems and other systems. These questions can begin to be answered by combining more structural data with biochemical characterizations. By capturing snapshots of metallochaperones in

various metalation and/or nucleotide states in the presence and absence of their target proteins, the molecular mechanisms of action can begin to be uncovered.

Summary of this thesis

In this thesis, I present novel structural and biochemical data that advances our understanding of how adenosylcobalamin (AdoCbl) is delivered to its target mutase metalloenzyme, with a specific emphasis on understanding the role of the G-protein metallochaperone that regulates successful maturation.

In Chapter II, I present the characterization of a minimal system consisting of the G-protein chaperone and the Cbl-binding domain of methylmalonyl-CoA mutase from *Methylobacterium extorquens* (*Me*MCM$_{cbl}$). *Me*MCM$_{cbl}$ activates GTP hydrolysis in the G-protein, MeaB, and forms a stable complex in the presence of a non-hydrolyzable GTP analog. This complex was characterized by crystallography, providing a snapshot of the conformation of the G-protein chaperone competent for GTP hydrolysis. Comparison of this active conformation of the G-protein chaperone to the inactive conformation of the G-protein chaperone revealed a conformational change is required to fully form the active site, with the conserved switch III motif located at the new dimer interface[92]. These studies explained the biochemical and physiological data from patients indicating the importance of the switch III region and its role in regulating the delivery and/or removal of AdoCbl for the mutase and allowed us to propose a mechanism for the G-protein's role in maturation of the mutase.

In Chapter III, I present the further characterization of AdoCbl-dependent isobutyryl-CoA mutase fused (IcmF) from *Cupriavidus metallidurans*. This AdoCbl-dependent system contains the G-protein chaperone and mutase on the same polypeptide, analogous to the two-component system found in humans and other bacteria[100]. Previous crystallographic studies were unable to explain if and how the conserved switch III region is involved in the maturation of the mutase in the fused system[88]. The work I describe here showed that substitutions of key conserved switch III residues does decrease observed GTP hydrolysis by IcmF, indicating a role for switch III in IcmF. Solution state data from negative stain EM and mass photometry indicated that oligomerization of IcmF is dependent on the identity of the G-nucleotide and Cbl cofactor,

consistent with what one would expect based on the MeaB studies presented in Chapter II. High-resolution cryo-EM analysis of the supramolecular complexes of IcmF revealed a novel IcmF structure. This structural data showed that an “active” G-protein domain structure is one in which two G-protein domains, each from a different IcmF molecule come together to form a dimeric interface. This interface wedges open the active site of the mutase domain for AdoCbl delivery. In this structure, the conserved switch III motif is found at the interface of the G-protein domain of the IcmF supramolecular complexes, supporting a similar role in regulating the delivery and/or removal of AdoCbl as with the non-fused system MeaB:MCM system. These data expand our understanding of the G-protein’s role in the fused system, supporting of the use of IcmF as a model system for mechanistic studies of AdoCbl delivery to mutase enzymes.

In Chapter IV, I discuss our characterization of AdoCbl transfer from the adenosyltransferase (ATR) to the target mutase with G-protein chaperone mediation. I present the crystal structure of ATR from *C. metallidurans* and use this structure to computationally model how ATR may interact with IcmF to transfer AdoCbl. These studies provide the foundation for understanding the molecular basis of Cbl transfer within a G-protein:ATR:mutase complex.

In Chapter V, I present my concluding thoughts on the final step in the delivery of AdoCbl to its target mutase. These future directions focus on the continued use of cryo-EM to capture complexes unamenable to crystallography. One avenue to pursue is the structural characterizations of the G-protein chaperone, MMAA, in complex with its target mutase, methylmalonyl-CoA mutase, from humans. Solution state analysis indicates the formation of higher order oligomers in the presence of a non-hydrolyzable GTP analog; preliminary negative stain EM analysis indicates the ability to trap these complexes moving forward. Another avenue to pursue is the structural characterizations of IcmF with its cognate ATR to elucidate how all three proteins work together to efficiently deliver the precious AdoCbl cargo. Overall, we have expanded our understanding of the GTPase metallochaperone molecular mechanism of action that regulates the loading and unloading of Cbl from its target mutase throughout my time in graduate school, setting the stage to answer even more questions regarding AdoCbl delivery by the ATR to its target mutase.

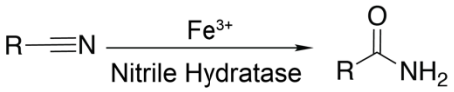
Funding

This work was supported by NIH grant R35 GM126982 to C.L.D. and an MIT Dean's Fellowship and an NIH F31 predoctoral fellowship GM131648 to F.A.V.. C.L.D. is a Howard Hughes Medical Institute Investigator and Senior Fellow for the Canadian Institute for Advanced Research Bio-Inspired Solar Energy program.

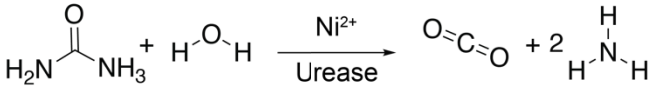
Figures

Metal Ions

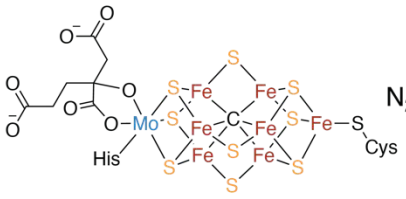
Fe^{3+}



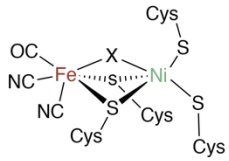
Ni^{2+}



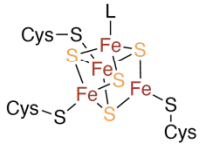
Metallic Clusters



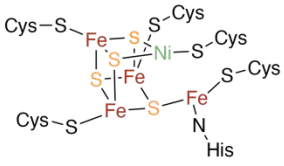
$N_2 + 8e^- + 8H^+ + 16ATP \xrightarrow[\text{Nitrogenase}]{\text{FeMo cofactor}} 2\ NH_3 + H_2 + 16ADP + 16P_i$



$H_2 \underset{\text{Hydrogenase}}{\overset{\text{[Ni-Fe] cluster}}{\rightleftharpoons}} 2H^+ + 2e^-$

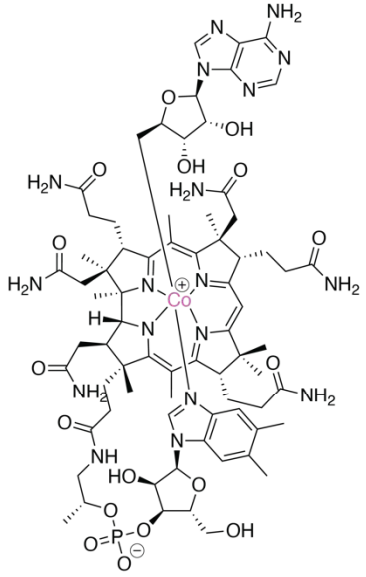


[4Fe-4S] clusters function in electron transfer when L=Cys and in radical chemistry when L=*S*-adenosylmethionine



$O=C=O + 2H^+ + 2e^- \underset{\text{Carbon Monoxide Dehydrogenase}}{\overset{\text{C-cluster}}{\rightleftharpoons}} C\equiv O + H-O-H$

Organometallic Cofactors



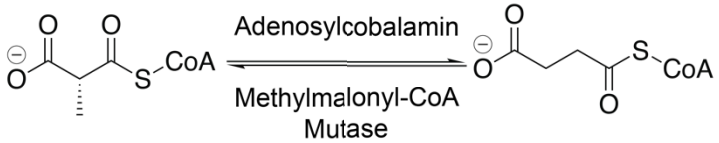


Figure I-1. Metallocofactors vary widely in structure and reactivity. In some reactions, metal ions such as Fe^{3+} or Ni^{2+} are ligated by residues of the metalloprotein but still require metallochaperones for proper maturation. Increasingly complex metallic clusters, such as the FeMo-cofactor and [4Fe-4S] clusters, are transported and inserted by metallochaperones; whereas, other metallic clusters such as the [Ni-Fe] cluster are assembled in situ. Additionally, organometallic cofactors are often metabolically expensive and are transported by metallochaperones to their target metalloenzyme to perform difficult chemistry.

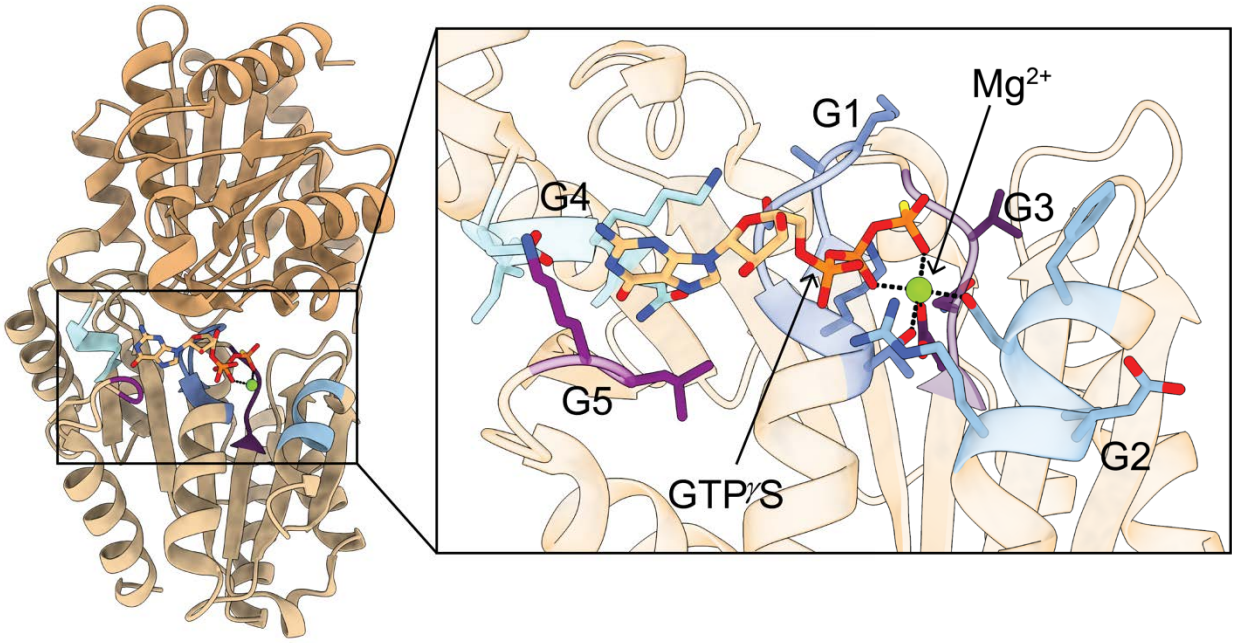


Figure I-2. Dimeric structure of P-loop-containing G3E GTPase metallochaperone with conserved motifs. Representative structure: HypB from *Methanocaldococcus jannaschii* (PDB 2HF8) with guanosine 5′-O-(3-thiotriphosphate) (GTPγS) and Mg^{2+} bound[101]. Structure is a dimer with each monomer containing a G-domain. **Insert:** The GTP binding site contains the G1-5 motifs, which are highlighted. G1 residues (dark blue), also known as the Walker A motif, interact with the α and β phosphates and the bound Mg^{2+} ion. The G2 residues (light blue), also known as switch I, change conformation based on the nucleotide state and the conserved aspartate residue coordinates the bound Mg^{2+} ion. The G3 residues (dark purple), also known as the Walker B motif, coordinate the γ-phosphate and the bound Mg^{2+} ion. The G4 residues (cyan) confer nucleotide specificity as they interact with the base of the nucleotide. The G5 residues (light purple) are involved in nucleotide dissociation.

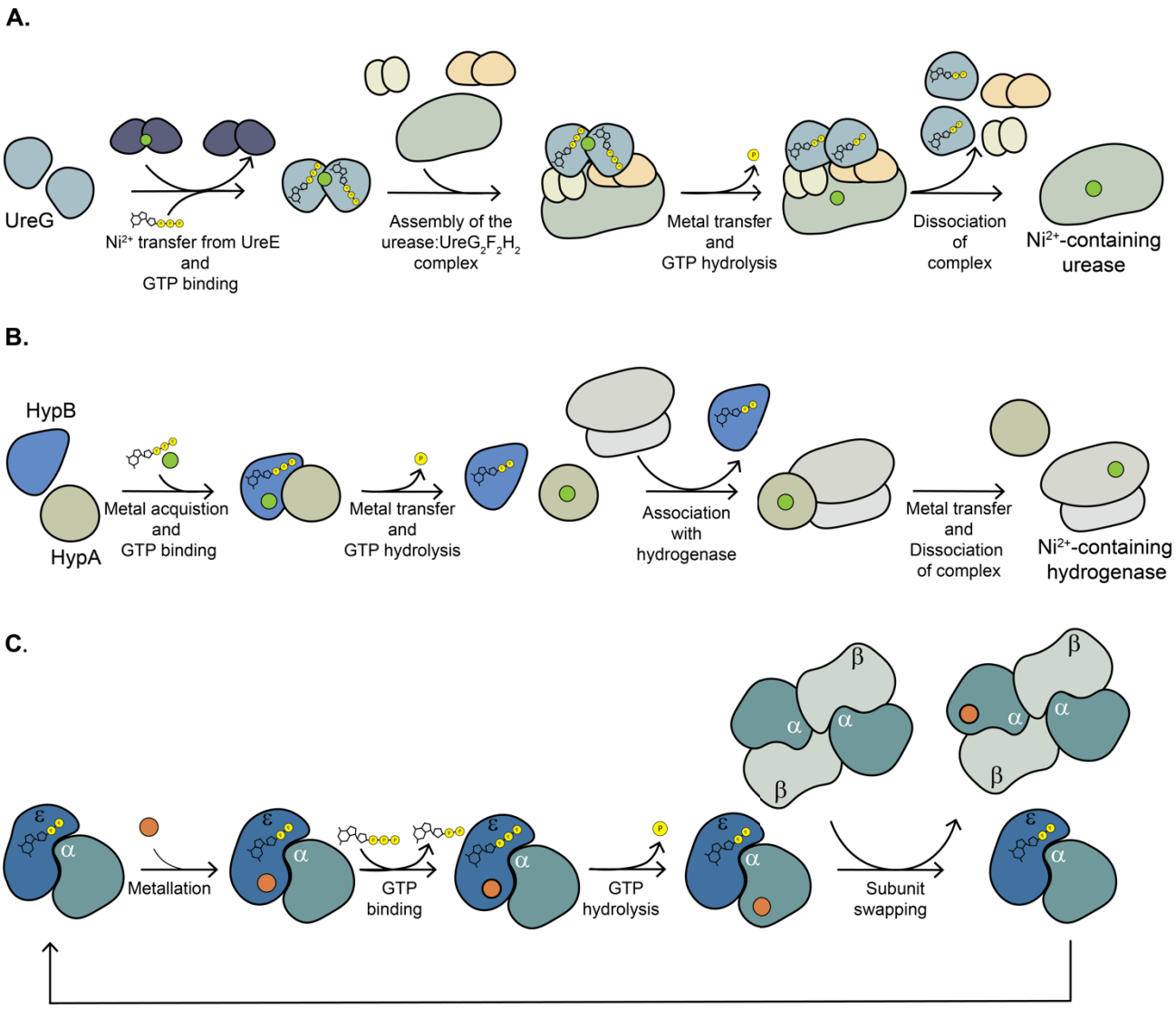


Figure I-3. Postulated roles of G3E P-loop GTPases in metalloenzyme maturation. **A.** Urease (pale green) requires the GTPase metallochaperone UreG (light blue) for insertion of the required Ni^{2+} ion. UreG receives the Ni^{2+} ion from UreE (dark blue) and binds GTP. The GTP binding event allows for the assembly of the apo-urease: $UreG_2F_2H_2$ complex. GTP hydrolysis triggers a conformational change that facilitates metal transfer to urease. After GTP hydrolysis occurs, the complex dissociates, and the urease now contains the needed Ni^{2+} ion for a fully activate enzyme[14, 19, 37]. **B.** [Ni-Fe]-hydrogenases (light grey) require the metallochaperone HypB (blue) for Ni^{2+} ion insertion. HypB binds GTP and acquires Ni^{2+} which is transferred to HypA (pale green). HypA then interacts with the inactive hydrogenase to transfer the metal to the active site of the [NiFe]-hydrogenases for activation[21, 38, 39, 90, 102]. **C.** Fe-type NHases maturation with Fe ions

(orange)[46, 47] involves a COG0532 GTPase metallochaperone, called ε of Nha3 (blue), which forms a complex with the α subunit of NHase (gray). The holo α subunit of NHase is swapped for an apo α subunit of NHase from the NHase $\alpha_2\beta_2$ complex. The process is repeated to fully mature the Fe-type NHases using the activating protein Nha3[46].

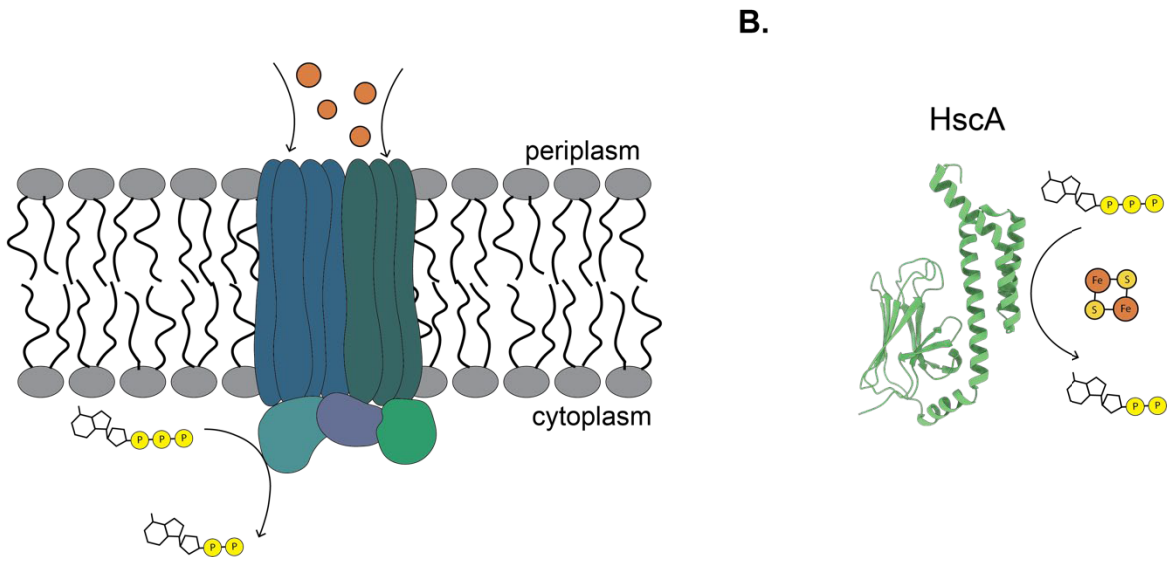


Figure I-4. Some representative ATPases involved in metalloprotein maturation. **A.** P1B-type ATPases are involved in the transport of heavy metals (orange spheres) into the cells to maintain metal homeostasis. The protein contains two domains: a transmembrane domain and a cytoplasmic ATPase domain. The flux of heavy metals is coupled to hydrolysis of ATP[49]. **B.** The heat shock protein fold ATPases such as HscA from *Escherichia coli* (PDB 1U00) interact with co-chaperones in a nucleotide dependent manner to transfer newly synthesized [2Fe-2S] clusters in the ISC biogenesis pathway[103].

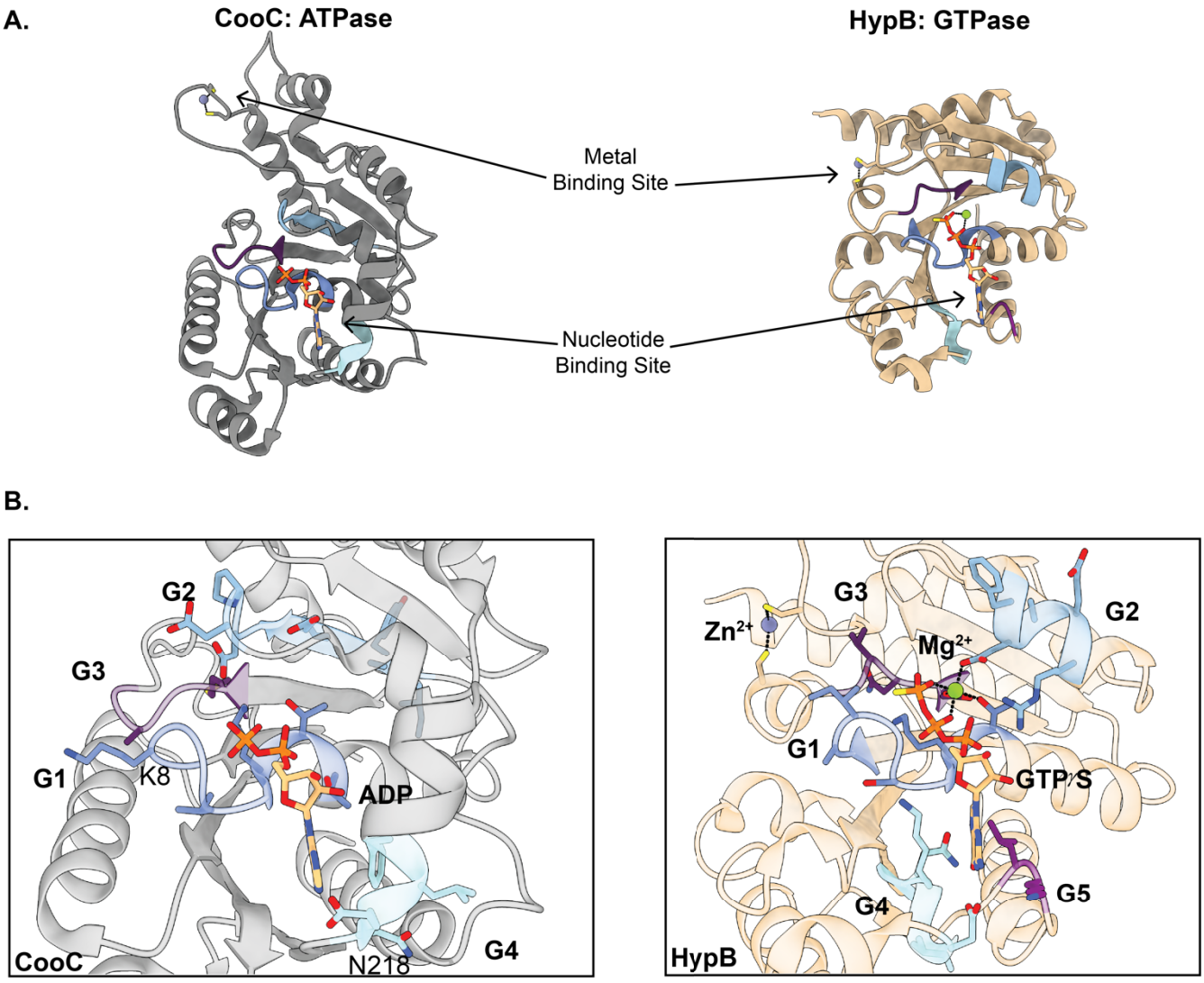


Figure I-5. Comparison of the ATPase CooC and GTPase HypB. **A.** The ATPase metallochaperone CooC from *C. hydrogenoformans* (PBD 3KJI) (Left, grey)[55] and the GTPase metallochaperone HypB from *M. jannaschii* (PBD 2HF8) (Right, tan)[101] contain the conserved overall fold found in P-loop NTPases consisting of repeating α-β units. The metal binding site and the nucleotide binding site are spatially separated but involved in controlling the oligomeric states of the NTPases. Only one monomer for each metallochaperone is shown for simplicity. **B.** (Left) A closer view of the ATP binding site of CooC bound to ADP indicates that it contains the conserved G1-G3 motifs implicated in interacting with the phosphates of the nucleotide and the Mg^{2+} ion if present. G1 residues (dark blue), also known as the Walker A motif, interact with the α and β phosphates and the bound Mg^{2+} ion. CooC contains the deviant Walker A motif with a

highly conserved lysine residue (K8 in CooC) in the second position of the sequence instead of the second to last position. The G2 residues (light blue), also known as switch I, change conformation based on the nucleotide state and the conserved aspartate residue coordinates the bound Mg^{2+} ion. The G3 residues (dark purple), also known as the Walker B motif, coordinate the γ-phosphate and the Mg^{2+} ion if present. CooC only retains the first asparagine residue (N218 in CooC) of the G4 motif (cyan) that confers nucleotide specificity. An alpha helix, which is only present when a nucleotide is bound, interacts with the adenine base. (Right) A closer view of the GTP binding site of HypB bound to GTPγS and a Mg^{2+} ion reveals the canonical G1-G5 motifs described in *Figure I-2*. Unlike CooC, HypB contains the conserved G4 motif (cyan) that confers nucleotide specificity by interacting with the guanosine base. The G5 residues (light purple) are involved with nucleotide dissociation and are not conserved in CooC.

Tables

Table I-1. Subfamilies of the G3E P-Loop GTPases

Subfamily	Known Roles	Known Metals/Metallocofactor
UreG	Maturation of ureases	Nickel
HypB	Maturation of [Ni-Fe] hydrogenases	Nickel
MeaB	Maturation of adenosylcobalamin-dependent mutases	Cobalamin
COG0523	Zinc homeostasis (ZigA/ZagA), Nitrile hydratase maturation (Nha3), Cobalamin cofactor biosynthesis (CobW)	Zinc Iron Cobalt

Table I-2. Example Structures of Members of the G3E P-loop GTPases

Subfamily	Protein	Organism	PDB IDs: Ligand(s) bound
MeaB	MeaB	*Methylobacterium extorquens*	2QM8: no ligands bound[89] 2QM7: GDP* bound[89] 4JYB: GMPPNP** bound[92]
	MMAA	*Homo sapiens*	2WWW: GDP bound[41]
	MeaB fusion protein (IcmF)	*Cupriavidus metallidurans*	4XC7: no ligands bound[88] 4XC8: GDP and Mg^{2+} bound[88]
HypB	HypB	*Methanocaldococcus jannaschii*	2HF9: GTPγS‡ and Mg^{2+} bound[101]
		Heliobacter pylori	4LPS: GDP, Mg^{2+}, Ni^{2+} bound[104]
UreG	UreG	*Klebsiella pneumoniae*	5XKT: GMPPNP, Ni^{2+} bound[37]
	UreG/UreF/UreH	*Helicobacter pylori*	4HI0: GDP bound[14]
COG0523	YjiA	*Escherichia coli*	4IXM: Zn^{2+} bound[26]

* GDP: guanosine diphosphate

** GMPPNP: Guanosine 5′-[β,γ-imido]triphosphate

‡ GTPγS: guanosine 5′-*O*-(3-thiotriphosphate)

Table I-3. Subtypes of ATPase Metallochaperones

Subtype	Known Roles	Known Metals/Metallocofactor
P1B class	Transporting metal ions across membranes	d-block metals
P-loop	Maturation of metalloenzyme active sites	Ni^{2+}
Heat Shock Fold Proteins	Biogenesis of Fe-S clusters	Fe-S clusters

References

(1) Degtyarenko, K. Metalloproteins. In *Online Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics*; Jorde, L. B., Little, P. F. R., Dunn, M. J., Subramaniam, S., Eds.; John Wiley & Sons, Ltd 2005; DOI: 10.1002/047001153X.g306204.

(2) Hausinger, R. P. New metal cofactors and recent metallocofactor insights. *Current Opinion in Structural Biology* **2019**, *59*, 1-8. DOI: 10.1016/j.sbi.2018.12.008.

(3) Wachnowsky, C.; Fidai, I.; Cowan, J. A. Iron-sulfur cluster biosynthesis and trafficking - impact on human disease conditions. *Metallomics* **2018**, *10* (1), 9-29. DOI: 10.1039/c7mt00180k PubMed.

(4) Yruela, I. Transition metals in plant photosynthesis. *Metallomics* **2013**, *5* (9), 1090-1109. DOI: 10.1039/c3mt00086a.

(5) Lee, C.-C.; Stiebritz, M. T.; Hu, Y.; Ribbe, M. W. Assembly and Function of Nitrogenase. In *Enzymes for Solving Humankind's Problems: Natural and Artificial Systems in Health, Agriculture, Environment and Energy*, Moura, J. J. G., Moura, I., Maia, L. B. Eds.; Springer International Publishing, 2021; pp 155-184.

(6) Ragsdale, S. W. Nickel and the carbon cycle. *Journal of Inorganic Biochemistry* **2007**, *101* (11), 1657-1666. DOI: 10.1016/j.jinorgbio.2007.07.014.

(7) Shafaat, H. S.; Rüdiger, O.; Ogata, H.; Lubitz, W. [NiFe] hydrogenases: A common active site for hydrogen metabolism under diverse conditions. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*
2013, *1827* (8), 986-1002. DOI: 10.1016/j.bbabio.2013.01.015.

(8) Capdevila, Daiana A.; Edmonds, Katherine A.; Giedroc, David P. Metallochaperones and metalloregulation in bacteria. *Essays In Biochemistry* **2017**, *61* (2), 177-200. DOI: 10.1042/ebc20160076.

(9) Homer, M. J.; Dean, D. R.; Roberts, G. P. Characterization of the γ Protein and Its Involvement in the Metallocluster Assembly and Maturation of Dinitrogenase from *Azotobacter vinelandii*. *Journal of Biological Chemistry* **1995**, *270* (42), 24745-24752. DOI: 10.1074/jbc.270.42.24745.

(10) Rubio, L. M.; Rangaraj, P.; Homer, M. J.; Roberts, G. P.; Ludden, P. W. Cloning and Mutational Analysis of the γ Gene from *Azotobacter vinelandii* Defines a New Family of Proteins Capable of Metallocluster Binding and Protein Stabilization. *Journal of Biological Chemistry* **2002**, *277* (16), 14299-14305. DOI: 10.1074/jbc.m107289200.

(11) Lerner-Ellis, J. P.; Dobson, C. M.; Wai, T.; Watkins, D.; Tirone, J. C.; Leclerc, D.; Doré, C.; Lepage, P.; Gravel, R. A.; Rosenblatt, D. S. Mutations in the MMAA gene in patients with the cblA disorder of vitamin B_{12} metabolism. *Human Mutation* **2004**, *24* (6), 509-516.

(12) Dobson, C. M.; Wai, T.; Leclerc, D.; Kadir, H.; Narang, M.; Lerner-Ellis, J. P.; Hudson, T. J.; Rosenblatt, D. S.; Gravel, R. A. Identification of the gene responsible for the cblB complementation group of vitamin B12-dependent methylmalonic aciduria. *Human Molecular Genetics* **2002**, *11* (26), 3361-3369. DOI: 10.1093/hmg/11.26.3361.

(13) Yang, X.; Koohi-Moghadam, M.; Wang, R.; Chang, Y.-Y.; Woo, P. C.; Wang, J.; Li, H.; Sun, H. Metallochaperone UreG serves as a new target for design of urease inhibitor: A novel strategy for development of antimicrobials. *PLoS Biology* **2018**, *16* (1), e2003887.

(14) Fong, Y. H.; Wong, H. C.; Yuen, M. H.; Lau, P. H.; Chen, Y. W.; Wong, K.-B. Structure of UreG/UreF/UreH complex reveals how urease accessory proteins facilitate maturation of *Helicobacter pylori* urease. *PLoS biology* **2013**, *11* (10), e1001678.

(15) Phillips, A. H.; Hernandez, J. A.; Payá-Tormo, L.; Burén, S.; Cuevas-Zuviría, B.; Pacios, L. F.; Pelton, J. G.; Wemmer, D. E.; Rubio, L. M. Environment and coordination of FeMo–co in the nitrogenase metallochaperone NafY. *RSC Chemical Biology* **2021**, *2* (5), 1462-1465.

(16) Nonaka, A.; Yamamoto, H.; Kamiya, N.; Kotani, H.; Yamakawa, H.; Tsujimoto, R.; Fujita, Y. Accessory Proteins of the Nitrogenase Assembly, NifW, NifX/NafY, and NifZ, Are Essential for Diazotrophic Growth in the Nonheterocystous Cyanobacterium Leptolyngbya boryana. *Front Microbiol* **2019**, *10*, 495-495. DOI: 10.3389/fmicb.2019.00495 PubMed.

(17) Robinson, N. J.; Winge, D. R. Copper Metallochaperones. *Annual Review of Biochemistry* **2010**, *79* (1), 537-562. DOI: 10.1146/annurev-biochem-030409-143539.

(18) Leipe, D. D.; Wolf, Y. I.; Koonin, E. V.; Aravind, L. Classification and evolution of P-loop GTPases and related ATPases. *Journal of Molecular Biology* **2002**, *317* (1), 41-72. DOI: 10.1006/jmbi.2001.5378.

(19) Nim, Y. S.; Wong, K.-B. The maturation pathway of nickel urease. *Inorganics* **2019**, *7* (7), 85.

(20) Farrugia, M. A.; Macomber, L.; Hausinger, R. P. Biosynthesis of the urease metallocenter. *Journal of Biological Chemistry* **2013**, *288* (19), 13178-13185.

(21) Lacasse, M. J.; Zamble, D. B. [NiFe]-Hydrogenase Maturation. *Biochemistry* **2016**, *55* (12), 1689-1701. DOI: 10.1021/acs.biochem.5b01328.

(22) Banerjee, R.; Gherasim, C.; Padovani, D. The tinker, tailor, soldier in intracellular B_{12} trafficking. *Current Opinion in Chemical Biology* **2009**, *13* (4), 484-491. DOI: 10.1016/j.cbpa.2009.07.007.

(23) Banerjee, R. B_{12} Trafficking in Mammals: A Case for Coenzyme Escort Service. *ACS Chemical Biology* **2006**, *1* (3), 149-159. DOI: 10.1021/cb6001174.

(24) Korotkova, N.; Lidstrom, M. E. MeaB is a component of the methylmalonyl-CoA mutase complex required for protection of the enzyme from inactivation. *Journal of Biological Chemistry* **2004**, *279* (14), 13652-13658.

(25) Jordan, M. R.; Wang, J.; Weiss, A.; Skaar, E. P.; Capdevila, D. A.; Giedroc, D. P. Mechanistic insights into the metal-dependent activation of ZnII-dependent metallochaperones. *Inorganic Chemistry* **2019**, *58* (20), 13661-13672.

(26) Sydor, A. M.; Jost, M.; Ryan, K. S.; Turo, K. E.; Douglas, C. D.; Drennan, C. L.; Zamble, D. B. Metal binding properties of Escherichia coli YjiA, a member of the metal homeostasis-associated COG0523 family of GTPases. *Biochemistry* **2013**, *52* (10), 1788-1801.

(27) Lonergan, Z. R.; Skaar, E. P. Nutrient Zinc at the Host–Pathogen Interface. *Trends in Biochemical Sciences* **2019**, *44* (12), 1041-1056. DOI: 10.1016/j.tibs.2019.06.010.

(28) Rodionov, D. A.; Vitreschak, A. G.; Mironov, A. A.; Gelfand, M. S. Comparative genomics of the vitamin B_{12} metabolism and regulation in prokaryotes. *Journal of Biological Chemistry* **2003**, *278* (42), 41148-41159.

(29) Osman, D.; Cooke, A.; Young, T. R.; Deery, E.; Robinson, N. J.; Warren, M. J. The requirement for cobalt in vitamin B_{12}: A paradigm for protein metalation. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **2020**, *1868* (1), 1-25.

(30) Crouzet, J.; Levy-Schil, S.; Cameron, B.; Cauchois, L.; Rigault, S.; Rouyez, M.; Blanche, F.; Debussche, L.; Thibaut, D. 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. *Journal of Bacteriology* **1991**, *173* (19), 6074-6087.

(31) Zhou, Z.; Hashimoto, Y.; Cui, T.; Washizawa, Y.; Mino, H.; Kobayashi, M. Unique biogenesis of high-molecular mass multimeric metalloenzyme nitrile hydratase: intermediates and a proposed mechanism for self-subunit swapping maturation. *Biochemistry* **2010**, *49* (44), 9638-9648.

(32) Sprang, S. R. G PROTEIN MECHANISMS: Insights from Structural Analysis. *Annual Review of Biochemistry* **1997**, *66*, 639-678.

(33) Calixto, A. R.; Moreira, C.; Pabis, A.; Kötting, C.; Gerwert, K.; Rudack, T.; Kamerlin, S. C. L. GTP Hydrolysis Without an Active Site Base: A Unifying Mechanism for Ras and Related GTPases. *Journal of the American Chemical Society* **2019**, *141* (27), 10684-10701. DOI: 10.1021/jacs.9b03193.

(34) Carvalho, A. T. P.; Szeler, K.; Vavitsas, K.; Åqvist, J.; Kamerlin, S. C. L. Modeling the mechanisms of biological GTP hydrolysis. *Archives of Biochemistry and Biophysics* **2015**, *582*, 80-90. DOI: 10.1016/j.abb.2015.02.027.

(35) Wittinghofer, A.; Vetter, I. R. Structure-Function Relationships of the G Domain, a Canonical Switch Motif. *Annual Review of Biochemistry* **2011**, *80* (1), 943-971. DOI: 10.1146/annurev-biochem-062708-134043.

(36) Li, G.; Zhang, X. C. GTP Hydrolysis Mechanism of Ras-like GTPases. *Journal of Molecular Biology* **2004**, *340* (5), 921-932. DOI: 10.1016/j.jmb.2004.06.007.

(37) Yuen, M. H.; Fong, Y. H.; Nim, Y. S.; Lau, P. H.; Wong, K.-B. Structural insights into how GTP-dependent conformational changes in a metallochaperone UreG facilitate urease maturation. *Proceedings of the National Academy of Sciences* **2017**, E10890-E10898.

(38) Alfano, M.; Cavazza, C. Structure, function, and biosynthesis of nickel-dependent enzymes. *Protein Science* **2020**, *29* (5), 1071-1089. DOI: 10.1002/pro.3836.

(39) Douglas, C. D.; Ngu, T. T.; Kaluarachchi, H.; Zamble, D. B. Metal transfer within the Escherichia coli HypB–HypA complex of hydrogenase accessory proteins. *Biochemistry* **2013**, *52* (35), 6030-6039.

(40) Takahashi-Iñiguez, T.; González-Noriega, A.; Michalak, C.; Flores, M. E. Human MMAA induces the release of inactive cofactor and restores methylmalonyl-CoA mutase activity through their complex formation. *Biochimie* **2017**, *142*, 191-196. DOI: 10.1016/j.biochi.2017.09.012.

(41) Froese, D. S.; Kochan, G.; Muniz, J. R. C.; Wu, X.; Gileadi, C.; Ugochukwu, E.; Krysztofinska, E.; Gravel, R. A.; Oppermann, U.; Yue, W. W. Structures of the human GTPase MMAA and vitamin B_{12}-dependent methylmalonyl-CoA mutase and insight into their complex formation. *Journal of Biological Chemistry* **2010**, *285* (49), 38204-38213.

(42) Padovani, D.; Banerjee, R. A G-protein editor gates coenzyme B_{12} loading and is corrupted in methylmalonic aciduria. *Proceedings of the National Academy of Sciences* **2009**, *106* (51), 21567-21572. DOI: 10.1073/pnas.0908106106.

(43) Padovani, D.; Banerjee, R. Assembly and protection of the radical enzyme, methylmalonyl-CoA mutase, by its chaperone. *Biochemistry* **2006**, *45* (30), 9300-9306. DOI: 10.1021/bi0604532.

(44) Padovani, D.; Labunska, T.; Banerjee, R. Energetics of interaction between the G-protein chaperone, MeaB, and B_{12}-dependent methylmalonyl-CoA mutase. *Journal of Biological Chemistry* **2006**, *281* (26), 17838-17844. DOI: 10.1074/jbc.M600047200.

(45) Edmonds, K. A.; Jordan, M. R.; Giedroc, D. P. COG0523 proteins: A functionally diverse family of transition metal-regulated G3E P-loop GTP hydrolases from bacteria to man. *Metallomics* **2021**, *13* (8), mfab046.

(46) Lankathilaka, K. P. W.; Bennett, B.; Holz, R. C. The Fe-type nitrile hydratase from *Rhodococcus equi TG*328-2 forms an alpha-activator protein complex. *Journal of*

Biological Inorganic Chemistry **2020**, *25* (6), 903-911. DOI: 10.1007/s00775-020-01806-y.

(47) Gumataotao, N.; Lankathilaka, K. P. W.; Bennett, B.; Holz, R. C. The iron-type nitrile hydratase activator protein is a GTPase. *Biochemical Journal* **2017**, *474* (2), 247-258. DOI: 10.1042/bcj20160884.

(48) Hashimoto, Y.; Ube, Y.; Doi, S.; Kumano, T.; Kobayashi, M. Metal chaperone, NhpC, involved in the metallocenter biosynthesis of nitrile hydratase. *The Journal of General and Applied Microbiology* **2021**, *67* (1), 24-32. DOI: 10.2323/jgam.2020.03.003.

(49) Keeran, N. S.; Usha, B.; Ganesan, G. P-type ATPases and their role in metal homeostasis in plants. In *Metal and Nutrient Transporters in Abiotic Stress*, Elsevier: 2021; pp 33-54.

(50) Dutkiewicz, R.; Nowak, M.; Craig, E. A.; Marszalek, J. Fe–S Cluster Hsp70 Chaperones: The ATPase Cycle and Protein Interactions. *Methods in Enzymology* **2017**, *595*, 161-184.

(51) Jeon, W. B.; Cheng, J.; Ludden, P. W. Purification and characterization of membrane-associated CooC protein and its functional role in the insertion of nickel into carbon monoxide dehydrogenase from *Rhodospirillum rubrum*. *Journal of Biological Chemistry* **2001**, *276* (42), 38602-38609.

(52) Jeoung, J.-H.; Goetzl, S.; Hennig, S. E.; Fesseler, J.; Wörmann, C.; Dendra, J.; Dobbek, H. The extended reductive acetyl-CoA pathway: ATPases in metal cluster maturation and reductive activation. *Biological Chemistry* **2014**, *395* (5), 545-558. DOI: doi:10.1515/hsz-2013-0290.

(53) Wittenborn, E. C.; Cohen, S. E.; Merrouch, M.; Léger, C.; Fourmond, V.; Dementin, S.; Drennan, C. L. Structural insight into metallocofactor maturation in carbon monoxide dehydrogenase. *Journal of Biological Chemistry* **2019**, *294* (35), 13017-13026. DOI: 10.1074/jbc.RA119.009610.

(54) Silberg, J. J.; Tapley, T. L.; Hoff, K. G.; Vickery, L. E. Regulation of the HscA ATPase reaction cycle by the co-chaperone HscB and the iron-sulfur cluster assembly protein IscU. *Journal of Biological Chemistry* **2004**, *279* (52), 53924-53931.

(55) Jeoung, J.-H.; Giese, T.; Grünwald, M.; Dobbek, H. Crystal Structure of the ATP-Dependent Maturation Factor of Ni,Fe-Containing Carbon Monoxide Dehydrogenases. *Journal of Molecular Biology* **2010**, *396* (4), 1165-1179. DOI: 10.1016/j.jmb.2009.12.062.

(56) Kim, J. H.; Alderson, T. R.; Frederick, R. O.; Markley, J. L. Nucleotide-Dependent Interactions within a Specialized Hsp70/Hsp40 Complex Involved in Fe–S Cluster Biogenesis. *Journal of the American Chemical Society* **2014**, *136* (33), 11586-11589. DOI: 10.1021/ja5055252.

(57) Maio, N.; Jain, A.; Rouault, T. A. Mammalian iron–sulfur cluster biogenesis: recent insights into the roles of frataxin, acyl carrier protein and ATPase-mediated transfer to recipient proteins. *Current Opinion in Chemical Biology* **2020**, *55*, 34-44.

(58) Burén, S.; Jiménez-Vicente, E.; Echavarri-Erasun, C.; Rubio, L. M. Biosynthesis of nitrogenase cofactors. *Chemical Reviews* **2020**, *120* (12), 4921-4968.

(59) Saini, A.; Mapolelo, D. T.; Chahal, H. K.; Johnson, M. K.; Outten, F. W. SufD and SufC ATPase activity are required for iron acquisition during in vivo Fe-S cluster formation on SufB. *Biochemistry* **2010**, *49* (43), 9402-9412. From Unbound Medicine PRIME.

(60) Layer, G.; Gaddam, S. A.; Ayala-Castro, C. N.; Ollagnier-de Choudens, S.; Lascoux, D.; Fontecave, M.; Outten, F. W. SufE Transfers Sulfur from SufS to SufB for Iron-Sulfur Cluster Assembly. *Journal of Biological Chemistry* **2007**, *282* (18), 13342-13350. DOI: 10.1074/jbc.M608555200.

(61) Chahal, H. K.; Dai, Y.; Saini, A.; Ayala-Castro, C.; Outten, F. W. The SufBCD Fe−S Scaffold Complex Interacts with SufA for Fe−S Cluster Transfer. *Biochemistry* **2009**, *48* (44), 10644-10653. DOI: 10.1021/bi901518y.

(62) Hirabayashi, K.; Yuda, E.; Tanaka, N.; Katayama, S.; Iwasaki, K.; Matsumoto, T.; Kurisu, G.; Outten, F. W.; Fukuyama, K.; Takahashi, Y.; et al. Functional Dynamics Revealed by the Structure of the SufBCD Complex, a Novel ATP-binding Cassette (ABC) Protein That Serves as a Scaffold for Iron-Sulfur Cluster Biogenesis *Journal of Biological Chemistry* **2015**, *290* (50), 29717-29731. DOI: 10.1074/jbc.M115.680934.

(63) Camire, E. J.; Grossman, J. D.; Thole, G. J.; Fleischman, N. M.; Perlstein, D. L. The yeast Nbp35-Cfd1 cytosolic iron-sulfur cluster scaffold is an ATPase. *Journal of Biological Chemistry* **2015**, *290* (39), 23793-23802.

(64) Netz, D. J. A.; Pierik, A. J.; Stümpfig, M.; Mühlenhoff, U.; Lill, R. The Cfd1–Nbp35 complex acts as a scaffold for iron-sulfur protein assembly in the yeast cytosol. *Nature Chemical Biology* **2007**, *3* (5), 278-286. DOI: 10.1038/nchembio872.

(65) Lill, R.; Dutkiewicz, R.; Freibert, S. A.; Heidenreich, T.; Mascarenhas, J.; Netz, D. J.; Paul, V. D.; Pierik, A. J.; Richter, N.; Stümpfig, M. The role of mitochondria and the CIA machinery in the maturation of cytosolic and nuclear iron–sulfur proteins. *European Journal of Cell Biology* **2015**, *94* (7-9), 280-291.

(66) Molé, C. N.; Dave, K.; Perlstein, D. L. Methods to Unravel the Roles of ATPases in Fe-S Cluster Biosynthesis. In *Fe-S Proteins: Methods and Protocols*, Dos Santos, P. C., Ed.; Springer US: New York, NY, 2021; pp 155-171.

(67) Sasaki, D.; Watanabe, S.; Matsumi, R.; Shoji, T.; Yasukochi, A.; Tagashira, K.; Fukuda, W.; Kanai, T.; Atomi, H.; Imanaka, T.; et al. Identification and Structure of a Novel Archaeal HypB for [NiFe] Hydrogenase Maturation. *Journal of Molecular Biology* **2013**, *425* (10), 1627-1640. DOI: 10.1016/j.jmb.2013.02.004.

(68) Watanabe, S.; Kawashima, T.; Nishitani, Y.; Kanai, T.; Wada, T.; Inaba, K.; Atomi, H.; Imanaka, T.; Miki, K. Structural basis of a Ni acquisition cycle for [NiFe] hydrogenase by Ni-metallochaperone HypA and its enhancer. *Proceedings of the National Academy of Sciences* **2015**, *112* (25), 7701. DOI: 10.1073/pnas.1503102112.

(69) Sestok, A. E.; Linkous, R. O.; Smith, A. T. Toward a mechanistic understanding of Feo-mediated ferrous iron uptake. *Metallomics* **2018**, *10* (7), 887-898. DOI: 10.1039/c8mt00097b.

(70) Lau, C. K. Y.; Krewulak, K. D.; Vogel, H. J. Bacterial ferrous iron transport: the Feo system. *FEMS Microbiology Reviews* **2015**, *40* (2), 273-298. DOI: 10.1093/femsre/fuv049.

(71) Ash, M.-R.; Maher, M. J.; Guss, J. M.; Jormakka, M. The Initiation of GTP Hydrolysis by the G-Domain of FeoB: Insights from a Transition-State Complex Structure. *PLOS ONE* **2011**, *6* (8), e23355. DOI: 10.1371/journal.pone.0023355.

(72) Guilfoyle, A. P.; Deshpande, C. N.; Vincent, K.; Pedroso, M. M.; Schenk, G.; Maher, M. J.; Jormakka, M. Structural and functional analysis of a FeoB A143S G5 loop mutant explains the accelerated GDP release rate. *The FEBS Journal* **2014**, *281* (9), 2254-2265. DOI: 10.1111/febs.12779 (acccessed 2022/01/07).

(73) Shin, M.; Mey, A. R.; Payne, S. M. *Vibrio cholerae* FeoB contains a dual nucleotide-specific NTPase domain essential for ferrous iron uptake. *Proceedings of the National Academy of Sciences* **2019**, *116* (10), 4599. DOI: 10.1073/pnas.1817964116.

(74) Gómez-Garzón, C.; Payne, S. M. Vibrio cholerae FeoB hydrolyzes ATP and GTP in vitro in the absence of stimulatory factors†. *Metallomics* **2020**, *12* (12), 2065-2074. DOI: 10.1039/d0mt00195c.

(75) Shin, M.; Park, J.; Jin, Y.; Kim, I. J.; Payne, S. M.; Kim, K. H. Biochemical characterization of bacterial FeoBs: A perspective on nucleotide specificity. *Archives of Biochemistry and Biophysics* **2020**, *685* (108350), 1-25. DOI: 10.1016/j.abb.2020.108350.

(76) Velayudhan, J.; Hughes, N. J.; McColm, A. A.; Bagshaw, J.; Clayton, C. L.; Andrews, S. C.; Kelly, D. J. Iron acquisition and virulence in *Helicobacter pylori*: a major role for FeoB, a high-affinity ferrous iron transporter. *Molecular Microbiology* **2000**, *37* (2), 274-286. DOI: 10.1046/j.1365-2958.2000.01987.x (acccessed 2021/12/22).

(77) Marlovits, T. C.; Haase, W.; Herrmann, C.; Aller, S. G.; Unger, V. M. The membrane protein FeoB contains an intramolecular G protein essential for Fe(II) uptake in bacteria. *Proceedings of the National Academy of Sciences* **2002**, *99* (25), 16243. DOI: 10.1073/pnas.242338299.

(78) Waldron, K. J.; Robinson, N. J. How do bacterial cells ensure that metalloproteins get the correct metal? *Nature Reviews Microbiology* **2009**, *7* (1), 25-35.

(79) Foster, A. W.; Osman, D.; Robinson, N. J. Metal preferences and metallation. *The Journal of Biological Chemistry* **2014**, *289* (41), 28095-28103. DOI: 10.1074/jbc.R114.588145 PubMed.

(80) Rae, T. D.; Schmidt, P. J.; Pufahl, R. A.; Culotta, V. C.; O'Halloran, T. V. Undetectable Intracellular Free Copper: The Requirement of a Copper Chaperone for Superoxide Dismutase. *Science* **1999**, *284* (5415), 805-808. DOI: doi:10.1126/science.284.5415.805.

(81) Outten, C. E.; O'Halloran, T. V. Femtomolar Sensitivity of Metalloregulatory Proteins Controlling Zinc Homeostasis. *Science* **2001**, *292* (5526), 2488-2492. DOI: doi:10.1126/science.1060331.

(82) Dann, C. E.; Wakeman, C. A.; Sieling, C. L.; Baker, S. C.; Irnov, I.; Winkler, W. C. Structure and Mechanism of a Metal-Sensing Regulatory RNA. *Cell* **2007**, *130* (5), 878-892. DOI: 10.1016/j.cell.2007.06.051.

(83) Xiao, Z.; Wedd, A. G. The challenges of determining metal–protein affinities. *Natural Product Reports* **2010**, *27* (5), 768-789.

(84) Young, T. R.; Martini, M. A.; Foster, A. W.; Glasfeld, A.; Osman, D.; Morton, R. J.; Deery, E.; Warren, M. J.; Robinson, N. J. Calculating metalation in cells reveals CobW acquires Co II for vitamin B_{12} biosynthesis while related proteins prefer Zn II. *Nature Communications* **2021**, *12* (1), 1-15.

(85) Osman, D.; Martini, M. A.; Foster, A. W.; Chen, J.; Scott, A. J. P.; Morton, R. J.; Steed, J. W.; Lurie-Luke, E.; Huggins, T. G.; Lawrence, A. D.; et al. Bacterial sensors define intracellular free energies for correct enzyme metalation. *Nature Chemical Biology* **2019**, *15* (3), 241-249. DOI: 10.1038/s41589-018-0211-4.

(86) Junop, M. S.; Obmolova, G.; Rausch, K.; Hsieh, P.; Yang, W. Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair. *Molecular Cell* **2001**, *7* (1), 1-12.

(87) Krishnan, A.; Burroughs, A. M.; Iyer, L. M.; Aravind, L. Comprehensive classification of ABC ATPases and their functional radiation in nucleoprotein dynamics and biological conflict systems. *Nucleic Acids Research* **2020**, *48* (18), 10045-10075. DOI: 10.1093/nar/gkaa726.

(88) Jost, M.; Cracan, V.; Hubbard, P. A.; Banerjee, R.; Drennan, C. L. Visualization of a radical B_{12} enzyme with its G-protein chaperone. *Proceedings of the National Academy of Sciences* **2015**, *112* (8), 2419-2424.

(89) Hubbard, P. A.; Padovani, D.; Labunska, T.; Mahlstedt, S. A.; Banerjee, R.; Drennan, C. L. Crystal structure and mutagenesis of the metallochaperone MeaB: Insight into the causes of methylmalonic aciduria. *Journal of Biological Chemistry* **2007**, *282* (43), 31308-31316. DOI: 10.1074/jbc.M704850200.

(90) Khorasani-Motlagh, M.; Noroozifar, M.; Kerman, K.; Zamble, D. B. Complex formation between the *Escherichia coli* [NiFe]-hydrogenase nickel maturation factors. *BioMetals* **2019**, *32* (3), 521-532. DOI: 10.1007/s10534-019-00173-9.

(91) Takahashi-Íñiguez, T.; García-Arellano, H.; Trujillo-Roldán, M. A.; Flores, M. E. Protection and reactivation of human methylmalonyl-CoA mutase by MMAA protein. *Biochemical and Biophysical Research Communications* **2011**, *404* (1), 443-447. DOI: 10.1016/j.bbrc.2010.11.141.

(92) Lofgren, M.; Padovani, D.; Koutmos, M.; Banerjee, R. A switch III motif relays signaling between a B_{12} enzyme and its G-protein chaperone. *Nature Chemical Biology* **2013**, *9* (9), 535-541. DOI: 10.1038/nchembio.1298.

(93) Wu, M.; Lander, G. C. How low can we go? Structure determination of small biological complexes using single-particle cryo-EM. *Current Opinion in Structural Biology* **2020**, *64*, 9-16. DOI: 10.1016/j.sbi.2020.05.007.

(94) Oldroyd, G.; Dixon, R. Biotechnological solutions to the nitrogen problem. *Current opinion in biotechnology* **2014**, *26C*, 19-24. DOI: 10.1016/j.copbio.2013.08.006.

(95) Dempsey-Nunez, L.; Illson, M. L.; Kent, J.; Huang, Q.; Brebner, A.; Watkins, D.; Gilfix, B. M.; Wittwer, C. T.; Rosenblatt, D. S. High resolution melting analysis of the MMAA gene in patients with cblA and in those with undiagnosed methylmalonic aciduria. *Molecular Genetics and Metabolism* **2012**, *107* (3), 363-367.

(96) Beilschmidt, L. K.; Puccio, H. M. Mammalian Fe–S cluster biogenesis and its implication in disease. *Biochimie* **2014**, *100*, 48-60. DOI: 10.1016/j.biochi.2014.01.009.

(97) Mandal, S. K.; Nayak, S. G.; Kanaujia, S. P. Identification and characterization of metal uptake ABC transporters in *Mycobacterium tuberculosis* unveil their ligand specificity. *International Journal of Biological Macromolecules* **2021**, *185*, 324-337. DOI: 10.1016/j.ijbiomac.2021.06.126.

(98) Xia, Y.; Cheng, Z.; Hou, C.; Peplowski, L.; Zhou, Z.; Chen, X. Discovery of the ATPase Activity of a Cobalt-Type Nitrile Hydratase Activator and Its Promoting Effect on Enzyme Maturation. *Biochemistry* **2022**. DOI: 10.1021/acs.biochem.2c00167.

(99) Pasquini, M.; Grosjean, N.; Hixson, K. K.; Nicora, C. D.; Yee, E. F.; Lipton, M.; Blaby, I. K.; Haley, J. D.; Blaby-Haas, C. E. Zng1 is a GTP-dependent zinc transferase needed for activation of methionine aminopeptidase. *Cell Reports* **2022**, *39* (7), 110834. DOI: 10.1016/j.celrep.2022.110834.

(100) Cracan, V.; Banerjee, R. Novel coenzyme B_{12}-dependent interconversion of isovaleryl-CoA and pivalyl-CoA. *Journal of Biological Chemistry* **2012**, *287* (6), 3723-3732. DOI: 10.1074/jbc.M111.320051.

(101) Gasper, R.; Scrima, A.; Wittinghofer, A. Structural insights into HypB, a GTP-binding protein that regulates metal binding. *Journal of Biological Chemistry* **2006**, *281* (37), 27492-27502. DOI: 10.1074/jbc.M600809200.

(102) Chan, K. H.; Li, T.; Wong, C. O.; Wong, K. B. Structural basis for GTP-dependent dimerization of hydrogenase maturation factor HypB. *PLoS ONE* **2012**, *7* (1). DOI: 10.1371/journal.pone.0030547.

(103) Cupp-Vickery, J. R.; Peterson, J. C.; Ta, D. T.; Vickery, L. E. Crystal Structure of the Molecular Chaperone HscA Substrate Binding Domain Complexed with the IscU Recognition Peptide ELPPVKIHC. *Journal of Molecular Biology* **2004**, *342* (4), 1265-1278. DOI: 10.1016/j.jmb.2004.07.025.

(104) Sydor, A. M.; Lebrette, H.; Ariyakumaran, R.; Cavazza, C.; Zamble, D. B. Relationship between Ni(II) and Zn(II) Coordination and Nucleotide Binding by the *Helicobacter pylori* [NiFe]-Hydrogenase and Urease Maturation Factor HypB *Journal of Biological Chemistry* **2014**, *289* (7), 3828-3841. DOI: 10.1074/jbc.M113.502781 (acccessed 2022/01/03).

CHAPTER II: A minimal system traps the active conformation of the G-protein chaperone

This chapter is reformatted from the following publication:

Vaccaro, F.A.; Born, D. A.;Drennan, C. L. Structure of metallochaperone in complex with the cobalamin-binding domain of its target mutase provides insight into cofactor delivery. *Proceedings of the National Academy of Science* **2023**, *120* (8), e2214085120.

Author contributions

D. A. Born and C. L. Drennan conceptualized the project. F. A. Vaccaro and D. A. Born performed all the experiments. F. A. Vaccaro and C. L. Drennan analyzed the data. F. A. Vaccaro and C. L. Drennan wrote the original draft.

Summary

The G-protein metallochaperone MMAA (MeaB in bacteria) is responsible for facilitating the delivery of adenosylcobalamin (AdoCbl) to methylmalonyl-CoA mutase (MCM), the only AdoCbl-dependent enzyme in humans. Genetic defects in the switch III region of MMAA lead to the genetic disorder methylmalonic aciduria in which the body is unable to process certain lipids. Here we present a crystal structure of *Methylobacterium extorquens* MeaB bound to a non-hydrolyzable GTP analog (GMPPCP) with the Cbl-binding domain of its target mutase enzyme (*Me*MCM_{cbl}). This structure provides an explanation for the stimulation of GTPase activity of MeaB afforded by target protein binding. We find that upon MCM_{cbl} association, one protomer of the MeaB dimer rotates ~180°, such that the inactive state of MeaB is converted to an active state in which the nucleotide substrate is now surrounded by catalytic residues. Importantly, it is the switch III region that undergoes the largest change, rearranging to make direct contacts with the terminal phosphate of GMPPCP. These structural data additionally provide insight into the molecular basis by which this metallochaperone contributes to AdoCbl delivery without directly binding the cofactor. Our data suggest a model in which GTP-bound MeaB stabilizes a conformation of MCM that is open for AdoCbl insertion, and GTP hydrolysis, as signaled by switch III residues, allows MCM to close and trap its cofactor. Substitutions of switch III residues destabilize the active state of MeaB through loss of protein:nucleotide and protein:protein interactions at the dimer interface, thus uncoupling GTP hydrolysis from AdoCbl delivery.

Significance statement

Guanine triphosphate hydrolyase (GTPase) metallochaperones enable delivery of metal cofactors to client proteins in host organisms that range from the pathogen *Helicobacter pylori* to humans. The combination of metal with protein facilitates critical cellular reactions. For example, nickel-dependent urease provides the requisite buffering capacity that allows for *H. pylori* to survive in the acidic human stomach. Human adenosylcobalamin-dependent methylmalonyl-CoA mutase is a vital metabolic enzyme whose impairment leads to disease. Overexpression and maturation of mutases and metal-dependent hydrogenases is of interest for industrial applications such as production of carbon fixation cycles and biofuel cells. In short, understanding the molecular basis of

metallochaperone function has numerous applications that extend from the development of *H. pylori* therapies to human health to metalloenzyme overexpression.

Introduction

Metalloproteins are ubiquitous in biology, accounting for about 30-50% of all proteins. The catalysis of chemically-challenging reactions in biology often requires a metallocofactor. This catalytic prowess can come at a price, however, as metal ions can be toxic to the cell. To prevent toxicity and to afford proper metalloprotein maturation, specialized proteins known as metallochaperones are employed for metallocofactor transport and/or delivery[1-3]. Guanine nucleotide-binding proteins (G-proteins) belonging to the SIMIBI (signal recognition particle, MinD, and BioD) class of P-loop NTPases are one important class of metallochaperones. [4] These proteins include: UreG[5, 6], which is involved in the assembly of the nickel metallocofactor of urease; HypB[7, 8], which is involved in incorporating Ni^{2+} ions into hydrogenase; and MeaB (MMAA in humans), which is involved in the delivery of coenzyme B_{12} (5′-deoxyadenosylcobalamin or AdoCbl) to methylmalonyl-CoA mutase (MCM) and related mutases[9-14]. Without these metallochaperones, metalloprotein function is impaired[15]. In humans, mutations to MCM, MMAA or any of the other B_{12} trafficking proteins result in inborn errors of metabolism[16].

This study focuses on MeaB from the bacteria *Methylobacterium extorquens*[17], which facilitates the delivery of AdoCbl to apo MCM and also assists in the removal of damaged cob(II)alamin from holo MCM. MCM catalyzes the 1,2-rearrangement of (*R*)-methylmalonyl-CoA to succinyl-CoA (*Figure II.1A*), a necessary step in the metabolism of odd-chain fatty acids, cholesterol, and branched amino acids[18]. The AdoCbl cofactor is essential for this chemically challenging carbon skeletal rearrangement. Homolytic cleavage of the carbon–cobalt bond of AdoCbl generates cob(I)alamin and a highly reactive 5'-deoxyadenosyl radical species, which abstracts a hydrogen atom from substrate to initiate this radical-based reaction. Upon reaction completion, the product radical abstracts a hydrogen atom back from 5'-deoxyadenosine and the AdoCbl is reformed. If, however, the cob(II)alamin is oxidized to cob(III)alamin, AdoCbl cannot be reformed, and the cob(III)alamin must be removed from the enzyme and replaced with a new AdoCbl. MeaB facilitates both the removal of the damaged cofactor and the insertion

of new cofactor. Importantly, MeaB does so without directly binding the cobalamin (Cbl) cofactor. In a process that requires GTP hydrolysis, MeaB enables the AdoCbl, which was generated by an adenosyltransferase (ATR) from cob(II)alamin and ATP, to be transferred from ATR to MCM.

To understand the molecular basis by which MeaB performs this chaperoning function, our lab and others have determined structures of bacterial MeaB and its human counterpart MMAA (methylmalonic aciduria type A)[12, 19-21] as well as structures of bacterial and human MCMs[22, 23]. Human MCM is a homodimer whereas bacterial MCMs are commonly heterodimers with one inactive subunit and one active subunit[12, 23]. Active subunits, as depicted in the structure of MCM from *Propionibacterium shermanii,* have a characteristic Rossmann domain that binds AdoCbl and a triose-phosphate isomerase (TIM) barrel domain to bind the substrate (***Figure II.1B***)[23-25]. Enzyme activity requires a 'closed' state in which the Ado moiety of AdoCbl is positioned within the substrate-binding TIM barrel, whereas cofactor delivery would require an 'open' state in which the Cbl-binding Rossmann domain is displaced from the TIM barrel. Presumably, MeaB assists in either or both the opening and closing of MCM.

The structure determination of MeaB revealed a common G-domain fold comprised of regularly recurring α-β units with the β-strands forming a central β-sheet surrounded on both sides by α-helices (***Figure II.1C***)[19]. Consistent with its classification as a P-loop NTPase, MeaB contains a P-loop (residues 62-70), a base-specificity loop (residues 200-207), and two conserved switch regions for signal transduction: switch I (residues 92-108) and switch II (residues 154-158) (***Figure II.1C***)[19, 20, 26-28]. MeaB and MMAA also contain an N-terminal extension domain[12, 19], a C-terminal extension domain involved in dimerization, and a switch III region (residues 177-188) (***Figure II.1C***)[21]. The switch III region was identified by the Banerjee lab through investigation of residues whose mutation is associated with methylmalonic aciduria, an inborn error of metabolism[29]. In vitro, they found that the substitution of switch III residues Lys188, Gln185, or Asp182 with alanine led to an uncoupling of GTP hydrolysis from AdoCbl transfer[21]. Thus, prior work has informed our understanding of which regions of MeaB (switch I, II, and III) are important for function. All structures, regardless of whether MeaB is bound to GDP, GMPPNP, (a non-hydrolyzable analog of GTP) or no nucleotide,

showed that the switch I and II regions were buried at the MeaB dimer interface and thus inaccessible, and that the switch III residues were more than 10 Å away from the nucleotide phosphates, too far to explain how they might be involved in the uncoupling of GTP hydrolysis from cofactor insertion (*Figure II.1C)*[19, 21]. The structures of MeaB alone thus indicated that additional conformational states of the G-protein must exist.

Most G-protein metallochaperones are poor GTPases (for example, MeaB alone has a k_{cat} of 0.039 +/- 0.003 min^{-1})[9] in the absence of their target protein. In fact, many GTPases have GTPase-accelerating proteins (GAP) that stimulate GTP hydrolysis upon binding. In the case of MeaB, GTP hydrolysis is increased ~100-fold in the presence of MCM, again suggesting that MeaB may undergo a conformational change upon binding MCM[9, 12]. No structure has been obtained of MeaB bound to MCM, but a structure of a natural fusion protein of a MeaB-like G-protein domain with a mutase enzyme has been solved (*Figure II.1D*). This protein, isobutyryl-CoA mutase fused (IcmF), interconverts isobutyryl-CoA and *n*-butyryl-CoA (*Figure II.1A*), as well as pivalyl-CoA and isovaleryl-CoA[30]. The IcmF structure revealed a monomeric G-protein domain wrapped around the Cbl-binding domain of the mutase with the switch I and II regions contacting the nucleotide binding site and the switch I region making direct contacts to the Cbl-binding domain[31]. However, it was unclear how this fused protein structure could be interpreted in terms of a dimeric standalone MeaB[31]. Additionally, the switch III region was solvent exposed in this structure and did not make any protein:protein contacts or protein:nucleotide contacts[31], indicating that a structure of the catalytically relevant state of MeaB was still missing.

Since the large 123 kDa fusion protein did not provide all the requisite structural data, we turned to a minimal system. We engineered a minimal mutase system consisting of solely the Cbl-binding domain of *Methylobacterium extorquens* MCM (*Me*MCM_{cbl}), removing the substrate-binding TIM barrel domain. Using this minimal mutase system and a non-hydrolyzable GTP analog (GMPPCP), we have trapped a conformation of a G-protein chaperone that has not been previously observed. This conformation reveals an ordered nucleotide binding site with switch III residues interacting across the dimer interface, providing a molecular route for signaling.

Results

The Cbl-binding domain of MeMCM provides a minimal model for investigation of MeaB's GTPase activity in the presence of MeaB's target

To assess the putative GTPase activating properties of the *Me*MCM$_{cbl}$, we incubated MeaB with apo *Me*MCM$_{cbl}$ and 500 µM GTP. We monitored the amount of inorganic phosphate (P_i) produced. In the same amount of time, MeaB in the presence of *Me*MCM$_{cbl}$ produced 25 times more P_i than MeaB alone (*Figure II.2A*). This observation validates the ability of the Cbl-binding domain alone to activate GTPase activity of MeaB. To test whether apo *Me*MCM$_{cbl}$ and MeaB form a stable complex, analytical size exclusion chromatography was used. Both MeaB and *Me*MCM$_{cbl}$ migrated independently with apparent molecular weights of 75 kDa and 40 kDa, respectively, consistent with both proteins forming homodimers. A mixture of MeaB and apo *Me*MCM$_{cbl}$ preincubated with GMPPCP eluted with a higher molecular weight species in addition to the individual protein dimer peaks (*Figure II.2B*). The higher molecular weight species had an apparent molecular weight of 110 kDa, consistent with one MeaB dimer and one *Me*MCM$_{cbl}$ dimer. SDS-PAGE analysis confirmed the co-elution of MeaB and *Me*MCM$_{cbl}$ from this peak (*Figure II.3*), indicating complex formation. The presence of the individual dimer peaks eluting after the high molecular weight species suggests an equilibrium between the complex and its protein components. Interestingly, the complex requires prebound GMPPCP for its formation. No complex was observed with GDP, GMPPNP, or in the absence of nucleotide.

MeaB uses a novel conformation to bind the apo form of MCM's Cbl-binding domain

To investigate how MeaB interacts with an apo form of the Cbl-binding domain of *Me*MCM in the absence of the substrate-binding TIM barrel, we solved the structure of MeaB in the presence of apo *Me*MCM$_{cbl}$ and GMPPCP to 2.72 Å resolution using molecular replacement (*Table II.1*). The final model contained a MeaB dimer (chains A and B) and two *Me*MCM$_{cbl}$ protomers (chains C and D) in the asymmetric unit (*Figure II.2C*). As reported previously, MeaB has a characteristic G-domain fold that is typical of G-proteins[19], and the mutase Cbl-binding domain has the characteristic Rossmann fold[23]. Only the first 18 residues of each of the *Me*MCM$_{cbl}$ protomers are disordered, which is not surprising given that these residues represent a linker region between the domains when

*Me*MCM$_{cbl}$ is part of a full-length enzyme. There is one Mg^{2+} ion and one GMPPCP molecule associated with each of the MeaB protomers (*Figure II.2C, II.4*). The Mg^{2+} ion is coordinated by six oxygens: an oxygen atom from the β- and γ-phosphate groups of GMPPCP, one water, and the side chains of Ser69, Asp105, and Glu154 (*Figure II.2C*).

The comparison of the overall conformation of the MeaB dimer as it exists in the MeaB:*Me*MCM$_{cbl}$:GMPPCP complex to the previous structures of apo MeaB or GDP-bound MeaB[19, 21] reveals a novel arrangement of MeaB protomers (*Figure II.5A, B*). When aligning one of the MeaB protomers of each structure, there is almost a 180° rotation of the other protomer needed to alter the conformation of MeaB that is observed on its own to the conformation observed in the MeaB:*Me*MCM$_{cbl}$ complex structure (*Figure II.5C, D*). This movement relocates the switch III region such that it forms new interactions with the other protomer of MeaB and creates a new interface that buries the MeaB nucleotide binding site.

Novel conformation of MeaB creates a fully ordered nucleotide binding site using both protomers

A comparison of this new MeaB structure with previous MeaB structures and the IcmF structure shows both previously observed and novel interactions between the nucleotide and chaperone domain[19, 21, 31]. In particular, the base-specificity loop and guanine base are the same between structures (*Figure II.6, II.7*). The interactions of the backbone atoms of the P-loop and the nucleotide phosphate groups are also similar (*Figure II.6, II.7*). Previous structures of MeaB do not have a Mg^{2+} ion bound (*Figure II.8A-C*), but from comparison with the Mg^{2+}-GDP bound IcmF structure (*Figure II.9*)[31] and the structure of HypB[8], three residues, Ser69, Asp105, and Glu154 were predicted to coordinate the Mg^{2+} ion in MeaB[19]. Our MeaB:*Me*MCM$_{cbl}$:GMPPCP structure confirms this prediction; as mentioned above, side chains of Ser69, Asp105 and Glu154 provide three coordinating oxygens, which together with oxygen atoms from the β- and γ-phosphate groups of GMPPCP and one water complete the Mg^{2+} ion's coordination sphere (*Figure II.2, II.8D*). Comparison with previous MeaB structures show that both Asp105 and Glu154 become ordered and/or repositioned to coordinate the bound Mg^{2+} ion (*Figure II.8*). This movement of Glu154 to coordinate the Mg^{2+} ion and to interact with Lys188B leads to a series of conformational rearrangements. First, Glu154 movement

breaks the interaction between Glu154 and Arg108, freeing the Arg108 side chain (*Figure II.8*). The side chain of Arg108 undergoes a substantial rearrangement such that it can coordinate the α-phosphate group of GMPPCP (*Figure II.8D*). In this conformation, Arg108 also forms a salt bridge with Asp182 from the second protomer (Asp182B of switch III), playing a role similar to that of an Arg finger, a common functional motif in the protein superfamily that includes GTPases[32, 33].

An important consequence of the ~180° MeaB protomer rearrangement described above is that the switch III residues of chain B are now positioned to contact the nucleotide binding site of chain A. In addition to Asp182, switch III residue Gln185 of chain B now contacts the β-phosphate of GMPPCP in chain A (*Figure II.8D*). Previously, Gln185 was either disordered (*Figure II.8A*) or made no observable interactions (*Figure II.8B, C*). Finally, switch III residue Lys188 from chain B rearranges to coordinate the γ-phosphate group of the nucleotide and interacts with switch I residue Asp92 and switch II residue Glu154 (*Figure II.8D*). In previous MeaB structures[19, 21], Lys188 formed hydrogen bonds with Gln160 and Glu162 but showed no nucleotide interactions and no interactions with other switch residues (*Figure II.8A-C*). Thus, this structure provides a molecular explanation for the involvement of switch III in nucleotide-dependent signaling. It also provides the previously undetermined snapshot of MeaB prior to GTP hydrolysis, revealing formerly unknown interactions between the protein and the Mg^{2+} ion and the protein and the nucleotide.

Due to the increased number of interactions made by the protein to the GMPPCP in this novel MeaB conformation (*Figure II.8*), it would appear that this MeaB conformer represents the “active state” of this GTPase and the structure of the conformer previously solved represents the “inactive state.” The “active conformer” is stabilized by interactions across the dimer interface (Asp92A-Lys188B; Arg108A-Asp182B; Glu154A-Lys188B; Gln160A-Gln185B) and by interactions of switch III residues with the GTP (Asp182B-β-phosphate; Gln185B-γ-phosphate; Lys188B-γ-phosphate) (*Table II.2*).

Switch I region makes up the MeaB:MeMCM$_{cbl}$ interface

Using our structure of Mg^{2+}-GMPPCP-bound MeaB:*Me*MCM$_{cbl}$, we were able to investigate how the binding of *Me*MCM$_{cbl}$ can shift the conformational equilibrium of MeaB

from an “inactive conformer” to an “active conformer.” We find that one protomer of MeaB contacts the edge of the Rossmann domain of *Me*MCM$_{cbl}$ in an analogous fashion as was observed in the structure of IcmF (***Figure II.10A***)[31]. Residues (99-107) of the switch I region that were disordered in the initial MeaB structure[19] become ordered and form the majority of the interface with *Me*MCM$_{cbl}$, explaining the molecular basis of switch I signaling. Notably, MeaB switch I residues 100-103 form a β-strand that runs parallel to, and hydrogen bonds with, a β-strand of the Rossmann domain of *Me*MCM$_{cbl}$ (residues 612-616). This interaction extends the Rossmann domain’s β-sheet from five strands to six (***Figure II.10, Table II.3***). With *Me*MCM$_{cbl}$ now forming a continuous β-sheet with MeaB, the chaperone and target protein would be expected to move as a rigid body, which was observed for IcmF previously[31]. Further anchoring MeaB and *Me*MCM$_{cbl}$ together are five salt bridges, Lys189B:Asp609, Lys106:Asp632, Arg33:Glu577, Arg25:Asp634, and Arg20:Asp632 (***Figure II.10B, Table II.4*****)**; all but one of them are formed by chain A of MeaB. The one notable exception is the salt bridge consisting of Lys189B (a switch III residue) of MeaB and Asp609 of *Me*MCM$_{cbl}$.

Although it was unclear at the time how relevant the structure of IcmF would be to a dimeric standalone MeaB system[31], we now see that much of the information provided by the IcmF structure was relevant. We find that it is one protomer of MeaB (chain A) that makes most of the interactions to *Me*MCM$_{cbl}$ (1150 Å^2 buried by chain A and 225 Å^2 buried by chain B), and that the interactions made, including those made by switch I residues, are the same in MeaB:MCM as in IcmF (***Figure II.10A,B***). Where IcmF, and its monomeric chaperone domain, was unsuccessful as a model system in providing insight into the importance of the switch III residues, and we now see why. It is the second protomer of MeaB (chain B) that contributes switch III residues to the chaperone:target interface and to the nucleotide binding site.

Superimposition of the minimal system on full-length mutase structures reveals MeaB:mutase clashes when the Cbl-binding domain is positioned for catalysis

To investigate how the “active conformation” of MeaB, as observed in our minimal system, would interact with a full-length mutase, we superimposed *Me*MCM$_{cbl}$ onto the Cbl-binding domains of previously solved structures of full-length mutases *P. shermanii*

MCM (*Ps*MCM)[23] and IcmF[31]. These Cbl-binding domains have high structural similarity: RMSD of 1.195 Å (Cα only) for *Me*MCM$_{cbl}$ with the *Ps*MCM and 1.193 Å and 1.180 Å (Cα only) for *Me*MCM$_{cbl}$ with dimeric IcmF (*Figure II.11*). When MeaB:*Me*MCM$_{cbl}$ is superimposed on the full length *Ps*MCM using the Cbl-binding domains for alignment, a clash is present between one MeaB helix (residues 206-228) and the substrate-binding TIM barrel domain of *Ps*MCM (*Figure II.12A*). This same clash is also present when MeaB:*Me*MCM$_{cbl}$ is superimposed onto an IcmF structure in which the Cbl-binding domain is sitting on top of the substrate-binding TIM barrel domain with the Cbl positioned for catalysis (*Figure II.12B*)[31]. In contrast, no clash is present when MeaB:*Me*MCM$_{cbl}$ is superimposed onto an IcmF protomer structure that displays a Cbl-binding domain that is positioned away from the substrate-binding domain, i.e. an open mutase conformation (*Figure II.12C*)[31]. Instead of clashing, the MeaB helix (residues 206-228) appears wedged between the Cbl-binding and the substrate-binding domains of the open mutase (*Figure II.12C*). Thus, the active conformation of MeaB appears to be structurally compatible with inactive (open) states of mutases and not with active (closed) mutase states. The inactive conformation of MeaB[19], on the other hand, is structurally compatible with the active (closed) state of IcmF[31]; superimposition of these structures is possible without any clashes (*Figure II.13*). There is a biochemical logic to these observations; the binding of the active conformation of MeaB should be restricted to inactive mutases that need Cbl.

Discussion

Metallochaperones assist in the maturation of metalloenzymes, ensuring that valuable metallocofactors are delivered efficiently and with minimal toxicity[34]. The molecular basis of metallochaperone function is an active area of research with many open questions[15]. In this study, we investigate the molecular mechanism of the AdoCbl maturase MeaB using a minimal system: MeaB in complex with an apo Cbl-binding domain. This minimal system has allowed us to probe the molecular basis of GTP binding and hydrolysis in MeaB, the basis by which MCM$_{cbl}$ binding enhances MeaB's GTPase activity, and the basis for the conformational opening and closing of the mutase that allows for AdoCbl insertion and holoenzyme maturation.

The structure of GMPPCP-bound MeaB:*Me*MCM$_{cbl}$ provided the first view of a fully formed active site for MeaB, allowing us to further consider previous proposals regarding the molecular mechanism of GTP hydrolysis. Asp92 of the switch I region has been proposed to activate and position a water molecule for an in-line nucleophilic attack of the γ-phosphate group[19]. In the presence of Mg^{2+}-GMPPCP, we find that Asp92 is ~4.5 Å (~4.1 Å in the other protomer) away from the oxygen of the γ-phosphate with an approximately in-line orientation (see *Figure II.8D*). Thus, Asp92 is a strong candidate to position and activate water for nucleophilic attack. No water is present in the structure near the terminal phosphate of GMPPCP, however. This lack of a water molecule may be since GMPPCP is not a perfect mimic of GTP (P-C-P bonds: 112^{o} versus P-O-P bonds: 121^{o}). The structure of MeaB:*Me*MCM$_{cbl}$ also reveals another possible player in catalyzing GTP hydrolysis; Lys188B of the switch III region is ~2.1 Å (~2.4 Å in the other protomer) away from the oxygen of γ-phosphate and ~3.4 Å (~3.0 Å in the other protomer) from Asp92 (see *Figure II.8D*). Substitutions of both Asp92 and Lys188 lead to MeaB variants that are impaired in their GAP activity[21, 28].

A key question for many GTPases is how the GTPase is designed such that its activity can be stimulated by the binding of a GAP or target protein. Here the structure of this minimal system provides insight into the mechanism by which the binding of MeaB to *Me*MCM$_{cbl}$ enhances GTP hydrolysis by 25-fold. We find that MeaB's molecular mechanism involves "inactive" and "active" conformers with MCM binding causing a shift of the conformational equilibrium towards the "active MeaB conformer." The "active conformer" is stabilized by interactions across the MeaB dimer interface (Asp92A-Lys188B; Arg108A-Asp182B; Glu154A-Lys188B; Gln160A-Gln185B) and by interactions of switch III residues with the GTP (Asp182B-α-phosphate; Gln185B-γ-phosphate; Lys188B-γ-phosphate). Almost all of these residues have been subject to mutagenesis previously (D92A, D92N, K188A, K188E, D182A, E154A, Q160A, Q185A) with substitutions shown to decrease the stimulatory effect of MCM binding on GTP hydrolysis, consistent with a role in stabilizing an "active" conformer[20, 21, 28]. Q185A and K188A, the residues of switch III that directly coordinate the terminal phosphate, have the largest effects on GAP activity[21], and substitutions of the equivalent residues in MMAA, the human homolog of MeaB, have been reported to cause methylmalonic aciduria[29].

Notably, our minimal model is not reporting on all of the stabilizing interactions: binding the full mutase stimulates GTPase activity of MeaB by 100-fold, considerably more than the 25-fold of the Cbl-binding domain alone[9]. We believe that this higher GAP activity is due to additional interactions between MeaB and the substrate-binding TIM barrel that further help to stabilize the active conformation of MeaB.

Additionally, these structural data provide insight into how MeaB facilitates the delivery of AdoCbl to a target mutase protein, either as part of a mutase maturation or a mutase repair process. We observe that when Mg^{2+}-GTP binds MeaB (*Figure II.14, state I*), the second protomer of MeaB rearranges to form an "active conformer." Structural superimpositions (*Figure II.12C*) suggest that a helix of the second protomer of "active" MeaB acts as a wedge between the Cbl-binding and substrate-binding domains of the mutase, presumably opening up the mutase for Cbl delivery or stabilizing an open conformation, which is formed by an apo- or damaged mutase (*Figure II.14, state II).* An open mutase conformation would allow for AdoCbl transfer from ATR to MCM (*Figure II.14, state II to IV).* Once AdoCbl is transferred, the mutase must close down to trap AdoCbl inside and position it for catalysis. Our structural data suggest that GTP hydrolysis will dramatically destabilize the "active conformer," transitioning MeaB back to the "inactive" conformation (*Figure II.14, state IV to V*). Our structure informs us that hydrolysis of GTP would result in the loss of the stabilizing interactions made to the terminal phosphate, to Lys188, and to the Mg^{2+} ion, resulting in the repositioning of Arg108, which in turn should break cross-dimer interactions. In other words, GTP hydrolysis would result in the loss of all the stabilizing interactions that we have identified for the "active conformer." When MeaB undergoes its 180° rotation back to the "inactive conformer", the mutase closes and is ready for catalysis (*Figure II.14, state V*).

Previously, the Banerjee lab showed that substitutions of switch III residues Asp182, Gln185, and Lys188 not only reduced in vitro GAP activity of MCM, but also uncoupled AdoCbl transfer from GTP hydrolysis[21]. In these experiments, authors measured AdoCbl transfer from ATR to MCM in presence of MeaB and a non-hydrolyzable analog of GTP, GMPPNP. For wild type MeaB, their data indicated that in the absence of GTP hydrolysis, AdoCbl is transferred from ATR but is not captured by MCM. Presumably without GTP hydrolysis, the mutase cannot close, so AdoCbl is

released into solution. In contrast to wild type MeaB, variants of MeaB that have a less stable “active conformation,” including switch III variants D182A, Q185A, K188A and switch I variants D92A, D105A, can capture more AdoCbl[21, 28]. AdoCbl is transferred by ATR and much more of it is captured by MCM. Presumably, in these cases, GTP hydrolysis was not necessary to convert these MeaB variants back to their inactive conformations, releasing the molecular wedge that was holding MCM open. These variants were unstable enough to transition back to the inactive states on their own, allowing for increased AdoCbl capture. Thus, in these MeaB variants, mutase maturation was uncoupled from GTP hydrolysis.

The proposed process of mutase maturation shown in *Figure II.14* requires the formation of a MeaB dimer that has one protomer attached to the Cbl-binding domain of the mutase via switch I and the other inserted between Cbl-binding and substrate-binding domains to stabilize an open conformation of the mutase for AdoCbl delivery. This model for MeaB begs the question of whether the already dimeric IcmF needs to form transient higher order oligomers for AdoCbl insertion, i.e. whether the same “active” and “inactive” protomer arrangements observed for MeaB also exist for the chaperone domain of IcmF, and if so, what is the benefit (if any) of the fused IcmF system. This MeaB model also begs the question of how many other dimeric metallochaperones (if any) will employ a similar molecular mechanism. There is considerable structural similarity between NTPase metallochaperones[15], leaving open the possibility that target protein binding will accelerate GTPase activity of other metallochaperones via 180° rotations. Additional structural data will be invaluable for answering these questions.

Sometimes smaller is better. This minimal system has provided a wealth of information about AdoCbl-dependent mutase maturation. Using this minimal system, we were finally able to provide a molecular explanation for the importance of the switch III region and the basis for methylmalonic aciduria when switch III residues Lys188 and Gln185 are substituted. It is always intriguing to consider the molecular function of metallochaperones that do not directly bind their metallocofactor. What do these chaperones do exactly and what is the role of GTP hydrolysis? Active sites in metalloenzymes are often buried to protect the reactive intermediates and/or to protect the metallocofactor itself from oxidation or other unwanted chemistry. In all mutases,

AdoCbl is buried[35], presumably for both reasons. Here we find that the ability to open up a mutase to insert AdoCbl and then close the structure down again is key to capturing this metabolically expensive cofactor (approximately 30 enzymes are needed for Cbl biogenesis). Metalloenzyme maturation processes can be as fascinating as the metalloenzymes are themselves, and understanding maturation is essential for industrial applications of metalloenzymes and for an understanding of human disease. We hope that the studies presented here will benefit metalloenzyme applications in industry and in medicine by providing insights into metalloprotein maturation.

Materials

All chemicals, solvents, and reagents were purchased from Sigma-Aldrich unless otherwise noted below. Restriction enzymes (*NdeI*, *NcoI*, and *XhoI*), *E. coli* Bl21 T7 Express competent cells and the Gibson Assembly® kit were purchased from New England Biolabs (NEB). Sanger sequence verification was performed by Genewiz®. LB medium components were purchased from Fisher BioReagents. Kanamycin was purchased from GoldBio and used at a concentration of 50 μg/mL. Isopropyl-β-D-thiogalactopyranoside (IPTG) was purchased from GoldBio. The NaCl for purification buffer components was purchased from Fisher Chemical. The EDTA-free protease inhibitor cocktail tablets were purchased from Roche. The Ni-NTA 1 mL columns and Superdex75 16/60, Superdex200 16/60, and Superdex200 10/300 Increase GL size exclusion columns (SEC) were purchased from GE Healthcare. The gel filtration standards, polyacrylamide gels, sodium dodecyl sulfate solution, Bradford protein assay dye and bovine serum albumin were purchased from BioRad. The $MgCl_2$ was purchased from CalBiochem. The crystallization solution (PEG 3350, LiCl) was purchased as one condition in the 96-condition PEG/Ion HT screen from Hampton Research.

Methods

The methods provided here have been described previously in the doctoral dissertation of Dr. David Born[36] and reproduced here with updates and modifications.

Cloning

The genes encoding MeaB (WP_003597297.1) and the Cbl-binding domain (residues 545-712) of the alpha subunit of methylmalonyl-CoA mutase (*Me*MCM$_{cbl}$) (WP_015857646.1) from *Methylobacterium extorquens*, each with an N-terminal hexahistidine affinity tag (HisTag), were synthesized individually by GenScript. Each individual gene was inserted into separate pET28a expression vectors at the *NdeI* and *XhoI* restriction sites.

Preliminary solubility tests of the *Me*MCM$_{cbl}$ with an N-terminal HisTag indicated low solubility, so a pET28a vector containing *Me*MCM$_{cbl}$ with a C-terminal HisTag was constructed to improve *Me*MCM$_{cbl}$ solubility. The *Me*MCM$_{cbl}$ gene was amplified from the pET28a vector with appropriate extension primers to facilitate Gibson assembly[37] into a pET28a vector cut with restriction enzymes *NcoI* and *XhoI*. The sequences were verified by Sanger sequencing.

Primers for Gibson Assembly

Name	**Sequence (5′ to 3′)**
pET28a_*Me*MCM_Cbl_Fw_1 (PCR amplification)	ATGCGTGCGCAGATCCGTAG
pET28a_*Me*MCM_Cbl_Rv_1 (PCR amplification)	CAGACGGGTGTTCAGTTCACCC
pET28a_*Me*MCM_Cbl_Fw_2 (installation of C-terminal HisTag)	GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGCGTGCGCAGATCCGTAG
pET28a_*Me*MCM_Cbl_Rv_2 (installation of C-terminal HisTag)	GATCTCAGTGGTGGTGGTGGTGGTGCTCGAGCAGACGGGTGTTCAGTTCAC

Plasmids

Name	**Features**	**Source**

pET28a-MeaB	KAN^R, *NdeI* and *XhoI* restriction sites	This study
pET28a-*me*MCM_{cbl}-N-TerminalHisTag	KAN^R, *NdeI* and *XhoI* restriction sites	This study
pET28a-*me*MCM_{cbl}-C-terminalHisTag	KAN^R, *NcoI* and *XhoI* restriction sites	This study

Protein expression and purification

Cell growth and purifications of MeaB and *Me*MCM_{cbl}, were conducted following the same procedure described here. An overnight starter culture of 100 mL LB supplemented with 50 µg/L kanamycin was inoculated from a single colony of *E. coli* Bl21 T7 Express competent cells transformed with the plasmid containing the appropriate gene and grown at 37 ºC with shaking. The overnight starter culture was used to inoculate 1 L of LB supplemented with 50 µg/L kanamycin at 37 ºC with shaking. The 1 L culture was induced with 0.5 mM IPTG when OD_{600} reached ~0.4-0.6 and grown for 16 h at 18 ºC with shaking. Cells were harvested by centrifugation (4,000 x g, 15 min, 4 ºC) and flash frozen in liquid N_2 before being stored in a -80 ºC freezer for future use.

Cells from 1 L of culture were resuspended in 80 mL lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 25 mM imidazole) supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF) and 1 Roche EDTA-free protease inhibitor cocktail tablet. Cells were lysed by ultrasonication, and cell lysates were clarified by centrifugation (25,000 x g, 30 min, 4 ºC). Clarified lysate was passed through a 0.2 µm filter before being loaded onto a 1 mL Ni-NTA column equilibrated with lysis buffer. Protein was eluted with a linear gradient of elution buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 500 mM imidazole) using an FPLC (Amersham Biosciences AKTA FPLC System). Elution fractions were concentrated in either a 30 kDa molecular weight cutoff (MWCO) centrifugal filter (for MeaB), or a 10 kDa MWCO centrifugal filter (for *Me*MCM_{cbl}). The concentrated fractions of MeaB were loaded onto a Superdex 200 16/60 SEC equilibrated with SEC buffer (50 mM HEPES pH 7.5, 500 mM NaCl). The concentrated fractions of *Me*MCM_{cbl} were loaded onto a Superdex75 16/60 SEC equilibrated with SEC buffer. Elution fractions from SEC were concentrated in a 30 kDa or 10 kDa MWCO centrifugal filter for MeaB or *Me*MCM_{cbl}

respectively. Purity was assessed by 4-20% (w/v) sodium dodecyl sulfate polyacrylamide gel electrophoresis. The concentration of MeaB monomer was determined to be 994 μM (36.4 mg/mL) by UV/Vis absorbance at 280 nm using an extinction coefficient of 23,490 $M^{-1}cm^{-1}$, determined using the ProtParam tool[38]. The concentration of *Me*MCM_{cbl} monomer was determined to be 307 μM by Bradford assay, using bovine serum albumin as a standard[39]. Protein samples in SEC buffer were flash frozen in liquid N_2 and stored in a -80 ºC freezer for future use.

GTPase assays

The GTPase activity of MeaB (2 μM) was determined in the presence of 500 μM GTP. The effect of *Me*MCM_{cbl} (4 μM) in SEC buffer was determined by preincubating the complex (2 μM MeaB) before initiating the GTPase assay with 500 μM GTP. The EnzCheck™ phosphate assay kit was used for all GTPase assays following manufacturer's instructions with the following modifications (Molecular Probes). The assay reactions (200 μL) were prepared excluding the GTP and incubated at room temperature for 30 minutes before initiating the assay to control for contaminating phosphate in the various components. Assays were performed in triplicate. The absorbance at 360 nm for each assay reaction was recorded using a SpectraMax Plus 384 microplate reader (Molecular Dimensions). The absorbances were converted into concentration of inorganic phosphate using the standard curve generated according to the manufacturer's directions. The no enzyme condition was subtracted from the reaction mixtures to control for background GTP hydrolysis.

Analytical size-exclusion chromatography

A Superdex200 Increase 10/300 GL SEC column equilibrated with SEC buffer was used to assess complex formation between MeaB and *Me*MCM_{cbl}. A 300 μL volume of the following five samples were injected onto the SEC column. Sample 1 contained 50 μM MeaB in SEC buffer. Sample 2 contained 50 μM MeaB and 100 μM *Me*MCM_{cbl} in SEC buffer. Samples 3-5 contained 50 μM MeaB and 100 μM *Me*MCM_{cbl} supplemented with 500 μM of the appropriate nucleotide (guanosine diphosphate (GDP), guanosine-5′-[(β,γ)-imido]triphosphate trisodium salt hydrate (GMPPNP), or guanosine-5′-[(β,γ)-

methyleno]triphosphate sodium salt (GMPPCP)) and 1 mM $MgCl_2$. The identity of each elution peak was determined by SDS-PAGE analysis (*Figure II.3*). The estimated molecular weights of elution peaks were calculated by comparison to the gel filtration standards containing thyroglobulin, bovin γ-globulin, chicken ovalbumin, equine myoglobin, and vitamin B_{12}, molecular weight from 1,3500-670,000 kDa.

Crystallography

The complex between MeaB and *Me*MCM_{cbl} was prepared by incubating equimolar amounts of MeaB and *Me*MCM_{cbl} (207 µM of each protein monomer) in SEC buffer with 1 mM GMPPCP and 2 mM $MgCl_2$ for 1 h on ice. The MeaB:*Me*MCM_{cbl} complex was purified by SEC on a Superdex200 16/60 column equilibrated with SEC buffer. The complex eluted with a molecular weight of approximately 110 kDa, representing a complex of two MeaB protomers and two *Me*MCM_{cbl} protomers. Following SEC, the complex was concentrated using a 30 kDa MWCO centrifugal filter to 182 µM as determined by UV/Vis absorbance at 280 nm using an extinction coefficient of 27,960 $M^{-1}cm^{-1}$, which was calculated by adding the extinction coefficients of MeaB (23,490 $M^{-1}cm^{-1}$) and *Me*MCM_{cbl} (4,470 $M^{-1}cm^{-1}$). The extinction coefficients were estimated using the ProtParam tool[38].

A 96-well plate of sitting-drops was set up with a Phoenix liquid handling robot (Art Robbins Instruments) with a Hampton Research PEG/Ion HT screen. The sitting drops were stored and imaged using the Formulatrix® Rock Imager 1000 at 18 ºC. Crystals were obtained by the vapor diffusion method by mixing 150 nL of protein solution (182 µM (20.2 mg/mL) complex in SEC buffer supplemented with 182 µM GMPPCP and 364 µM $MgCl_2$) with 230 nL precipitant solution (20% (w/v) PEG 3350 and 200 mM LiCl) incubated over 70 µL precipitant solution. A single thick plate-like crystal formed within 9 days and grew to a maximum size of approximately 250 µm within two weeks. The crystal was broken, and the resulting smaller crystals were transferred stepwise through three drops of increasing glycerol concentration into a cryogenic solution containing 20% (w/v) PEG 3350, 200 mM LiCl, 200 µM GMPPCP, 400 µM $MgCl_2$, and 20% (v/v) glycerol and flash frozen in liquid N_2.

Data collection, processing, structure determination, and refinement

A preliminary dataset for MeaB:*Me*MCM$_{cbl}$ was collected on an in-house Cu-K$_{\alpha}$ rotating anode source (Rigaku) with a Saturn 944 CCD detector at a temperature of 100 K in a single 180º wedge with 0.5º per image. Data were indexed, integrated, and scaled in XDS[40].

The structure was solved by molecular replacement with the Phenix implementation of Phaser[41] using data trimmed to 3.1 Å resolution. Two MeaB (PDB 2QM7)[19] protomers with the C-terminal dimerization helices removed (residues 5-293 out of 329) were first placed individually to create a partial solution which was used as a starting point for molecular replacement to place the Cbl-binding domains. The Phenix implementation of Sculptors[42], which removed the sidechains of divergent sequences, was used to create a homology model of *Me*MCM$_{cbl}$ using the Cbl-binding domain of methylmalonyl-CoA mutase (residues 545 - 712) from *Propionibacterium freudenreichii subsp. shermanii* (PDB 4REQ)[25] (70.9% identical). The final molecular replacement solution with LLG of 210 and TFZ of 12.0 identified four protomers in the asymmetric unit, corresponding to two MeaB protomers (chains A and B) and two *Me*MCM$_{cbl}$ protomers (chains C and D).

Applying non-crystallographic symmetry (NCS) restraints to each of the chains of the same protomer, the final solution identified from molecular replacement was refined using phenix.refine[43] with two rounds of simulated annealing at 5000 K to minimize model bias. Additionally, rigid body, positions, and group *B*-factors were refined with coordinate restraints (wc) equal to 4 and *B*-factor restraints (wu) equal to 3 and the weighting of the crystallographic refinement target (wxc_scale) equal to 0.1 in order to optimize the weight of geometry restraints compared to X-ray data. Subsequent iterative rounds of model building and refinement were performed in Coot[44] and Phenix, respectively. Sidechains were added to residues with clear electron density, and one molecule of GMPPCP was modeled in for each MeaB protomer. Iterative refinement including positional and group *B*-factor refinement with NCS restraints continued until the *R*-factors were 32.5% and 33.4% for the working *R*-factor (R_{work}) and the free *R*-factor (R_{free}) representing 5% of reflections, respectively. This model was not refined to completion and only used as a starting model for the additional dataset (*Table II.1*).

An additional dataset for the same MeaB:*Me*MCM$_{cbl}$ crystal used for the preliminary model was collected at the Advanced Photon Source (Argonne, Illinois, USA) on beamline 24ID-E using an Eiger-16M pixel array detector at a temperature of 100 K. Data were collected on the MeaB:*Me*MCM$_{cbl}$ crystal at a wavelength of 0.9791 Å in a single 360º wedge with 0.25º per image with 5% transmission. The MeaB:*Me*MCM$_{cbl}$ crystal belongs to space group $P2_12_12_1$. The MeaB:*Me*MCM$_{cbl}$ data were indexed, integrated and scaled to a 2.72-Å resolution in XDS. The same *R*-free flags from the preliminary dataset were used and extended in CCP4[45].

With NCS restraints for each of the chains of the same protomers, the model was refined with one round of simulated annealing at 2000 K to minimize model bias. Additionally, rigid body, positional and group *B*-factor refinement were used. Side chains were added to residues with clear electron density and one Mg^{2+} ion was placed with each MeaB protomer, restraining the distances from Mg^{2+} to its ligands to 2.1 Å. Waters were placed manually into regions with $2mF_o$-DF_c composite omit density, $2F_o$-F_c density, and 3 σ F_o-F_c density. Restraints from the CCP4 monomer library were used for the GMPPCP. The placement of GMPPCP, glycerol, and waters were confirmed using a $2mF_o$-DF_c composite omit map. The crystallographic refinement target for coordinates (wxc_scale), the crystallographic refinement target for *B*-factors (wxu_scale), coordinate restraints (wc) and *B*-factor restraints (wu) were optimized on each round of refinement. Subsequent iterative rounds of model building and refinement were performed in Coot and Phenix, respectively. The final model has a R_{work} of 21.0% and R_{free} of 23.6%. In chain A, residues Thr4-Gly294 and Val298-Ile327 of 329 were modeled into the density. In chain B, residues Met1-Pro228, Trp233-Leu262, and Asp276-Leu329 of 329 were modeled into the density. In chain C, residues Ser563-Gly592, Gly595-Leu712 of 712 were modeled into the density along with two residues of the C-terminal HisTag (713-714). In chain D, residues Ser563-Thr710 of 712 were modeled into the density. There was no density observed for the affinity tags on chains A, B, or D. Crystallographic software packages were compiled by SBGrid[46] (*Table II.1*). Structural figures were made in PyMOL 2.3.3 (The PyMOL Molecular Graphics System Version 2.3.3 Schrodinger, LLC).

Funding

This work was support by National Institutes of Health (NIH) grant R35 GM126982 (C.L.D.), a Ruth L. Kirschstein Predoctoral Individual National Research Service Award from NIH F31 GM131648 (F.A.V.), and a NIH Molecular Biophysics Training Grant (T32 GM008313) that supported D.A.B. F.A.V. was also supported by an MIT School of Science Fellowship. C.L.D is a Howard Hughes Medical Investigator. This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P41 GM103403). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. The structure of MeaB:*Me*MCM$_{cbl}$ is available through the Protein Data Bank under accession code 8DPB.

Figures

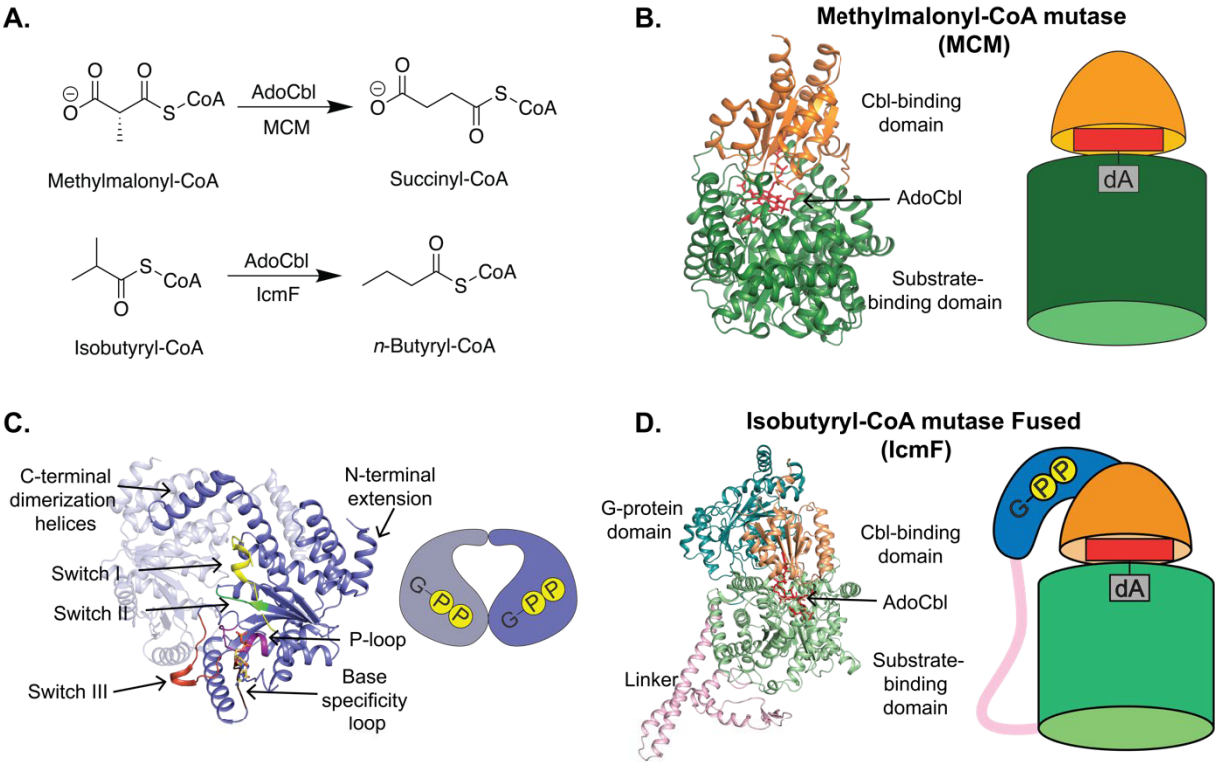


Figure II.1. Adenosylcobalamin-Dependent Mutases Require G-protein Chaperones for Maturation. **A.** Reaction catalyzed by methylmalonyl-CoA mutase (top) and a reaction catalyzed by isobutyryl-CoA mutase fused (bottom). **B.** The overall structure of the monomer of methylmalonyl-CoA mutase (MCM) from *Propionibacterium shermanii* bound to AdoCbl (PDB 4REQ)[23]. The AdoCbl (red sticks) is bound at the interface of the cobalamin (Cbl)-binding domain (light orange) and substrate binding domain (green). The cartoon representation of MCM has the same coloring as the ribbon structure. **C.** Signature motifs of the G3E P-loop GTPases shown on MeaB (slate) from *Methylobacterium extorquens* bound to GDP (PDB 2QM7)[19]. The P-loop (purple, residues (res. 62-70) interacts with the phosphates of the nucleotide. The base specificity loop (brown, res. 200-207) interacts with the guanosine base. The switch I (yellow, res. 92-108), switch II (green, res. 154-158) and switch III (red orange, res. 177-188) function as switches signaling the GTP hydrolysis event. **D.** The overall structure of one monomer of the dimeric isobutyryl-CoA mutase Fused (IcmF) from *Cupriavidius metallidurans* bound to GDP and AdoCbl (PDB 4XC6)[31] contains the Cbl-binding domain (light orange) and substrate-binding domain (light green) of the mutase on same polypeptide chain as the

G-protein domain (teal) connected by a polypeptide linker (pink). AdoCbl (red sticks) is bound in the active site at the interface of the Cbl-binding and substrate-binding domains. The cartoon representation of IcmF has the same coloring as the ribbon structure.

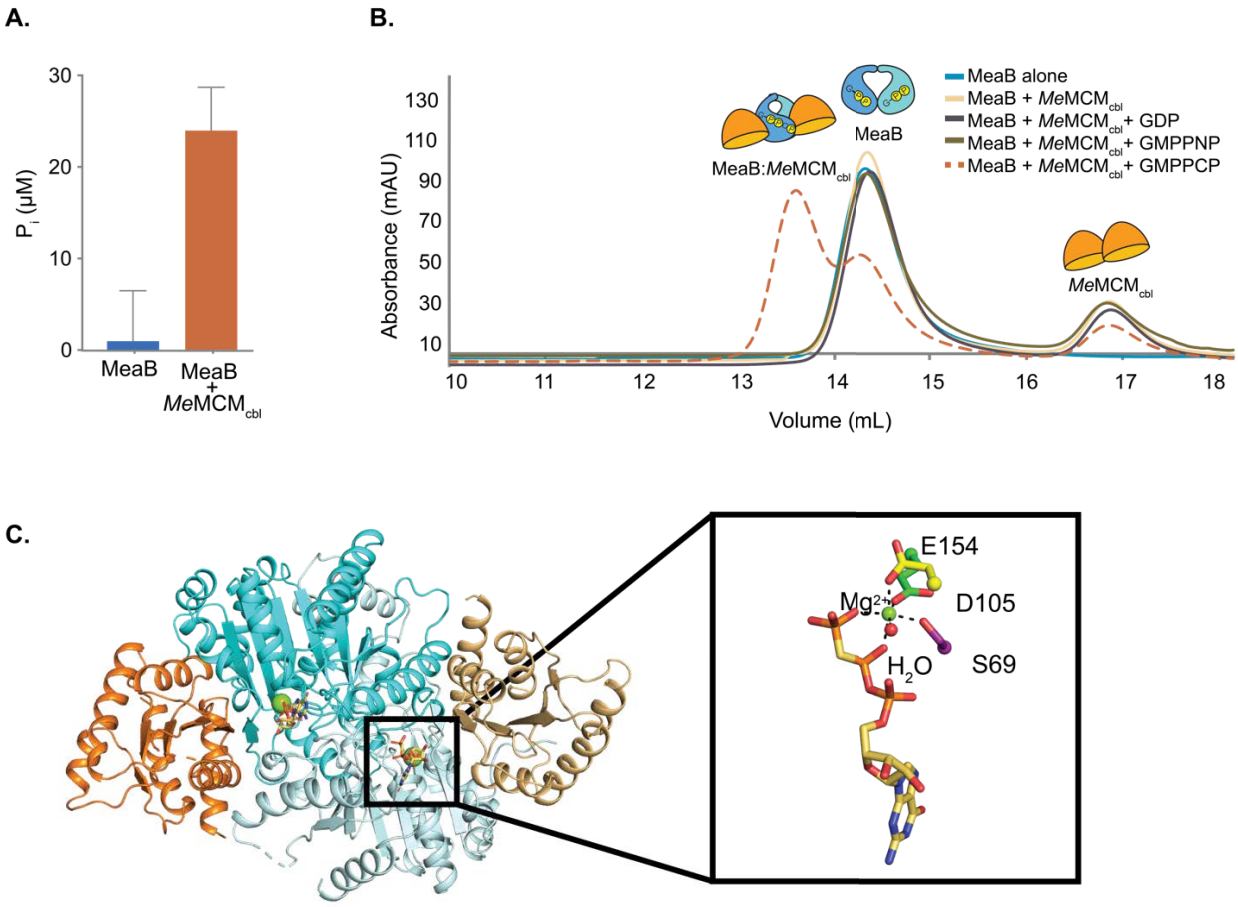


Figure II.2. MeMCM$_{cbl}$ Increases GTP Hydrolysis of MeaB and Can Form a Stable Complex with MeaB Allowing for Structure Determination. **A.** MeaB alone (blue) or MeaB with *Me*MCM$_{cbl}$ (orange) was incubated with 500 μM of GTP for 30 minutes at 22 ˚C. MeaB alone produced 0.95 μM of P$_i$ and MeaB in the presence of *Me*MCM$_{cbl}$ produced 24.0 μM of P$_i$. The amount of inorganic phosphate produced by MeaB in the presence of *Me*MCM$_{cbl}$ is 25-fold greater. The no enzyme control was subtracted both conditions. The reactions were performed in triplicate; error bars represent the 95% confidence interval. **B.** The S200 10/300 size exclusion chromatograms indicate the formation of a stable complex only occurs in the presence of GMPPCP (dashed orange). The traces of MeaB alone (blue), MeaB and *Me*MCM$_{cbl}$ with no G-nucleotide (yellow), MeaB and *Me*MCM$_{cbl}$ with GDP (black) and MeaB and *Me*MCM$_{cbl}$ with GMPPNP (green) indicate no stable complex formation. **C.** The asymmetric unit of the MeaB:*Me*MCM$_{cbl}$:GMPPCP structure

contains the MeaB homodimer (chain A and chain B, cyan and pale cyan, respectively,) bound to two molecules of GMPPCP (yellow sticks) and two Mg^{2+} ions (green spheres) and two apo *Me*MCM_{cbl} protomers (orange and tan). **Insert**: Mg^{2+} is coordinated by a water (red sphere), the oxygen atoms of Ser69, Asp105, Glu154, and the β- and γ-phosphate groups of GMPPCP.

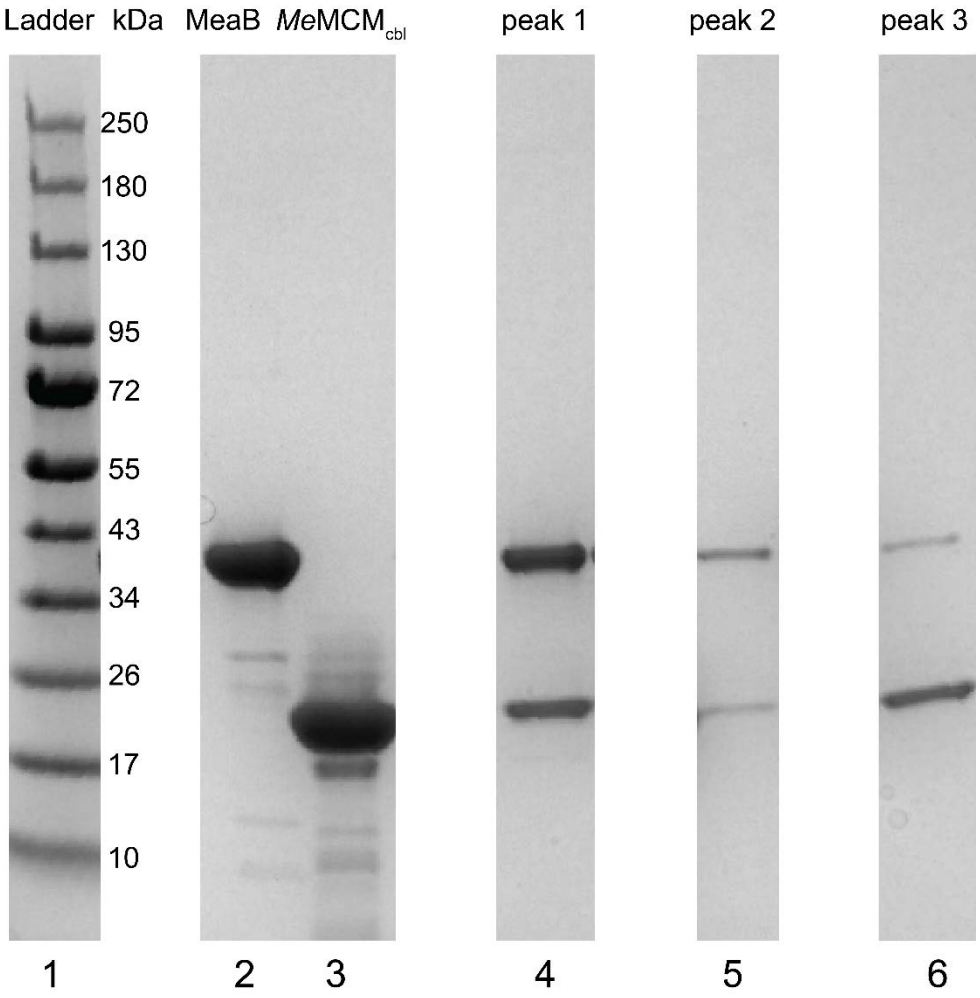


Figure II.3. SDS-PAGE Analysis Confirms the Presence of the Complex of MeaB with MeMCM$_{cbl}$. Each peak from the analytical SEC of MeaB, *Me*MCM$_{cbl}$ and GMPPCP was loaded onto an SDS-PAGE gel and compared to samples of MeaB and *Me*MCM$_{cbl}$. Lane 1, Ladder. Lanes 2 and 3, MeaB and *Me*MCM$_{cbl}$, respectively. Lane 4, elution peak 1 at 13.5 minutes which contains both MeaB and *Me*MCM$_{cbl}$ in equal amounts. Lane 5, elution peak 2 at 14.5 minutes which contains primarily MeaB. Lane 6, elution peak 3 at 16.5 minutes which contains primarily *Me*MCM$_{cbl}$.

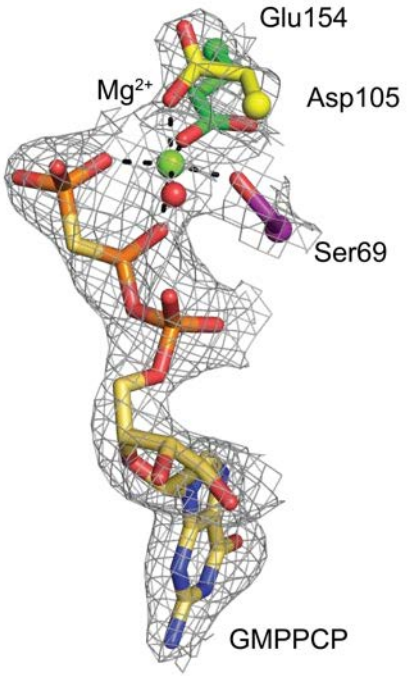


Figure II.4. Composite Omit Density Confirms the Presence of a Mg^{2+} ion in MeaB in the MeaB:MeMCM$_{cbl}$ structure. The $2mF_o$-DF_c composite omit maps contoured to 1σ confirm that Mg^{2+} ion (green sphere) is coordinated by a water (red sphere), the oxygen atoms of Ser69, Asp105, and Glu154 from MeaB, and the β- and γ-phosphate groups of GMPPCP.

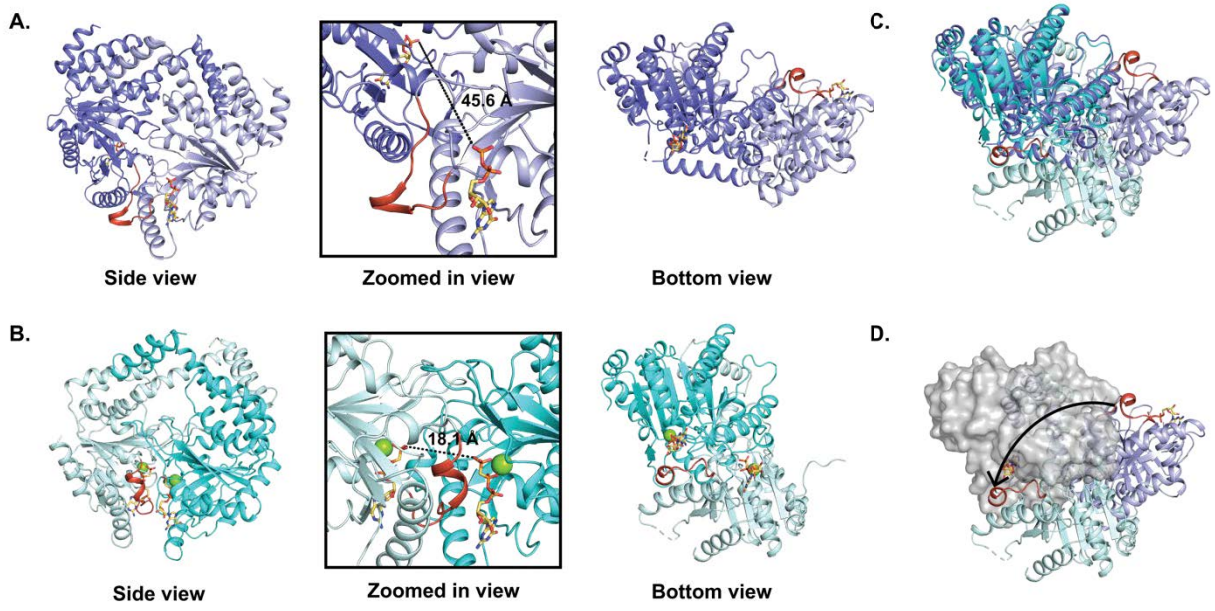


Figure II.5. The Structure of the MeaB:MeMCM$_{cbl}$:GMPPCP Complex Reveals a Novel MeaB Conformation. **A.** MeaB bound to GDP (PDB 2QM7)[19] from side view (same view shown in *Figure II.1*) and bottom view. Zoomed in panel shows distances between bound nucleotides in this previously observed MeaB conformation. The switch III region is in red orange. **B.** The MeaB dimer rearrangement from MeaB:*Me*MCM$_{cbl}$:GMPPCP shown in side view and bottom view. Relationships between side and bottom view are shown in *Figure II.15*. The upper protomers of both bottom views (here and in A) are oriented identically. Zoomed in panel shows distances between bound nucleotides in this novel MeaB conformation. **C.** Bottom views superimposed. Colored as in A and B. **D.** Bottom view in which both upper protomers are aligned and are shown as a gray surface, the lower protomer of MeaB:*Me*MCM$_{cbl}$:GMPPCP is in pale cyan, and lower protomer of MeaB bound to GDP is in light purple. The switch III region (red orange) moves from being solvent exposed when MeaB is bound to GDP to being at the interdimer interface of MeaB in the MeaB:*Me*MCM$_{cbl}$:GMPPCP complex. The arrow indicates the conformational change required to form the MeaB:*Me*MCM$_{cbl}$ complex.

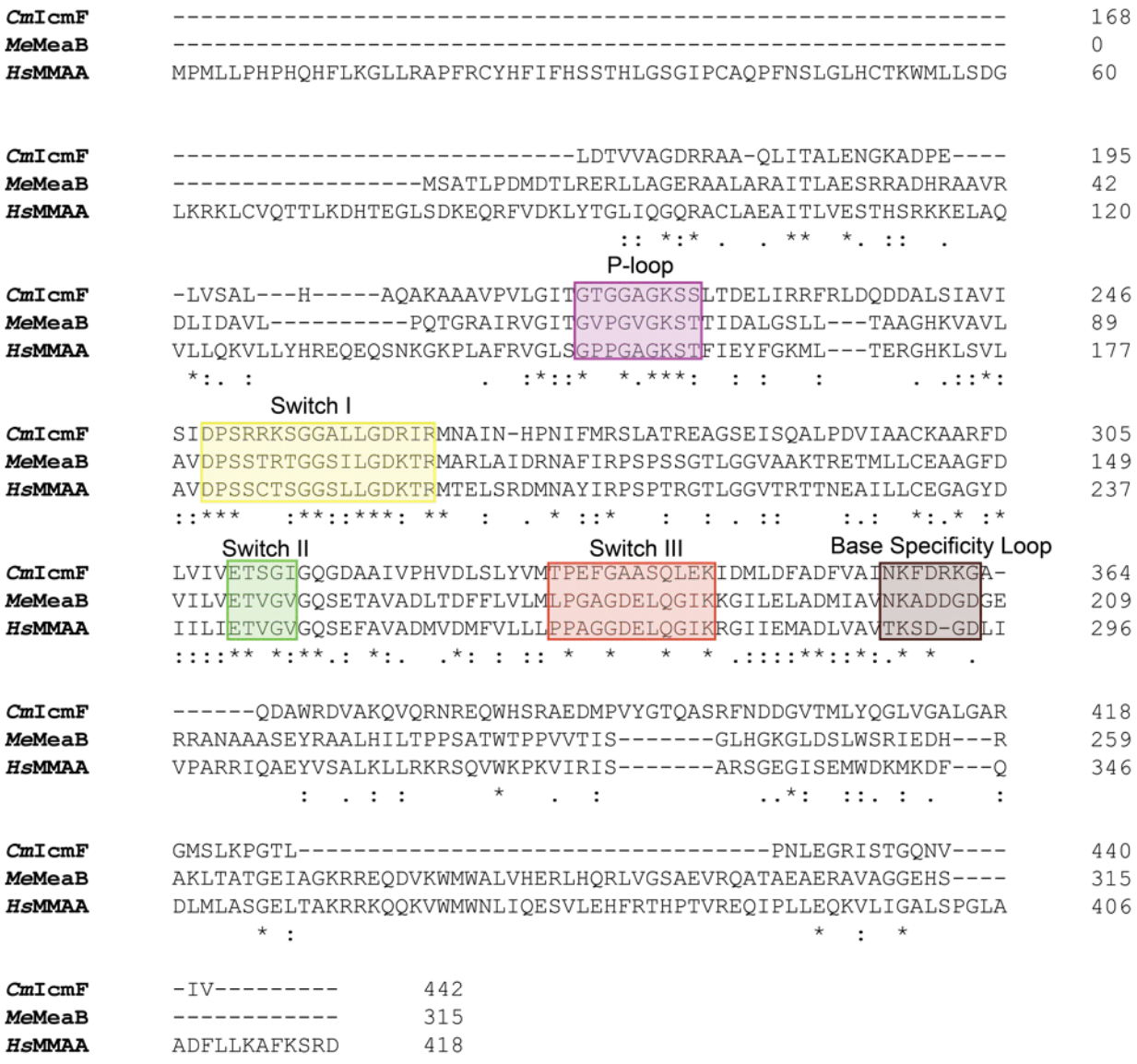


Figure II.6. Sequence Alignment of MeaB with MMAA and the G-protein Domain of CmIcmF. The regions which participate in binding and catalysis of the nucleotide substrate are highlighted in boxes corresponding to their coloring in the structure of MeaB in *Figure II.1*. Asterisks (*) denote positions with conserved residues. Colons (:) denote positions with conservation with strongly similar properties. Periods (.) denote positions with conservation with weakly similar properties. Alignments performed using ClustalW[47].

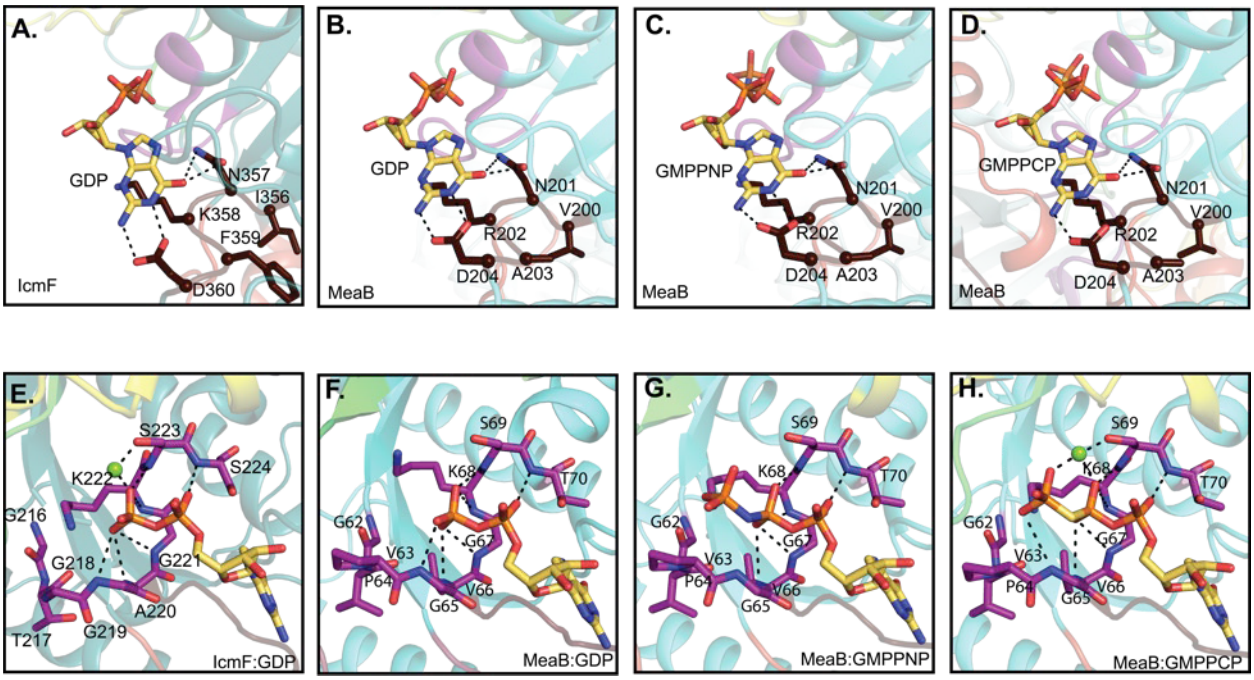


Figure II.7. The Base Specificity Loop and the P-loop Make Similar Interactions in all MeaB structures and in the IcmF structure. **A.** GDP-bound IcmF structure (PDB 4XC6)[31] with base specificity loop (brown sticks) of IcmF (cyan ribbons) providing stabilizing hydrogen bonds to the nucleotide. Asp360 coordinates both the amine at position 2 and the nitrogen at position 1, and Asn357 coordinates the carbonyl at position 6 to specifically bind only guanine bases**. B.** GDP-bound MeaB structure (PDB 2QM7)[19] with base specificity loop of MeaB providing stabilizing hydrogen bonds to the nucleotide. Asp 204 corresponds to Asp360 in IcmF; Asn201 corresponds to Asn357 in IcmF. Colored as in A. **C.** GMPPNP-bound MeaB structure (PDB 4JYB)[21]. Colored as in A. **D.** GMPPCP-bound MeaB:*Me*MCM$_{cbl}$ structure (this work). Colored as in A. **E.** GDP-bound IcmF structure (PDB 4XC6)[31] with the amines of the backbone of the P-loop (purple sticks) providing stabilizing hydrogen bonds to the negatively charged phosphate groups. **F.** GDP-bound MeaB structure (PDB 2QM7)[19] with the corresponding P-loop residues in MeaB shown as sticks. Colored as in E. **G.** GMPPNP-bound MeaB structure (PDB 4YJB)[21]. Colored as in E. **H.** GMPPCP-bound MeaB:*Me*MCM$_{cbl}$ structure (this work). Colored as in E. The nucleotides are shown as yellow sticks in all panels.

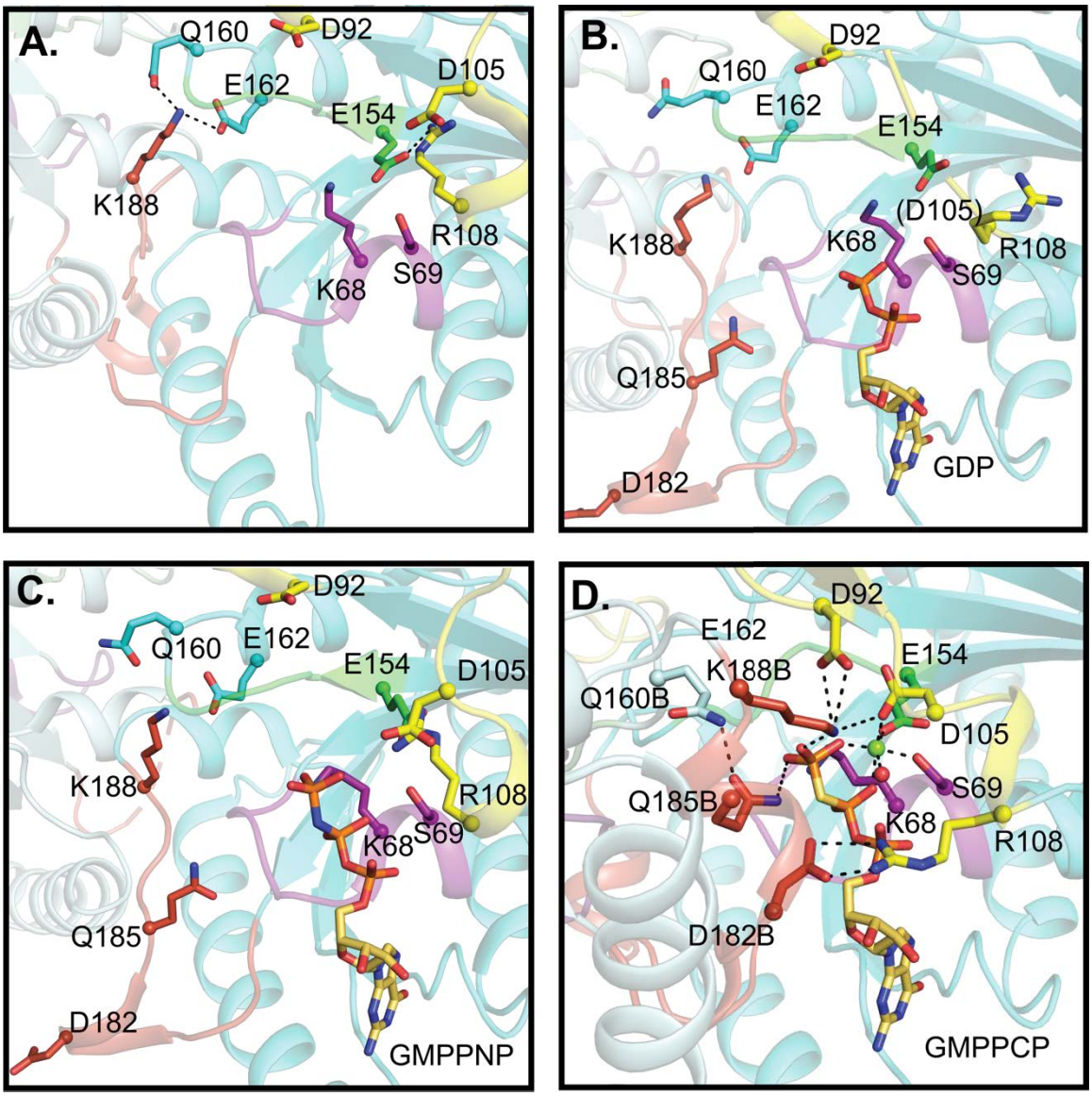


Figure II.8. Comparison of the Interactions of Nucleotide-binding Sites in MeaB Structures with MeaB:MeMCM$_{cbl}$:GMPPCP Structure. **A.** Nucleotide-free MeaB (PDB 2QM8)[19]. P-loop (purple), switch I (yellow), switch II (green), switch III (red orange), additional residues that are outside of the motifs that undergo conformational rearrangements (cyan). Base specificity loop interactions are shown in *Figure II.7* instead of here for clarity. **B.** GDP-bound MeaB (PDB 2QM7)[19]. **C.** GMPPNP-bound MeaB (PDB 4JYB)[21]. **D.** GMPPCP-bound MeaB in the presence of *Me*MCM$_{cbl}$ (this work). The residues shown in sticks undergo nucleotide-dependent interactions that contribute to the conformational changes necessary to increase GTP hydrolysis and communicate with the

mutase. Lys188B and Gln185B are ~2.1 Å and ~3.5 Å from the γ-phosphate of GMPPCP, respectively. Lys188B is also ~3.4 Å away from Asp92. The “B” label indicates residues from MeaB chain B.

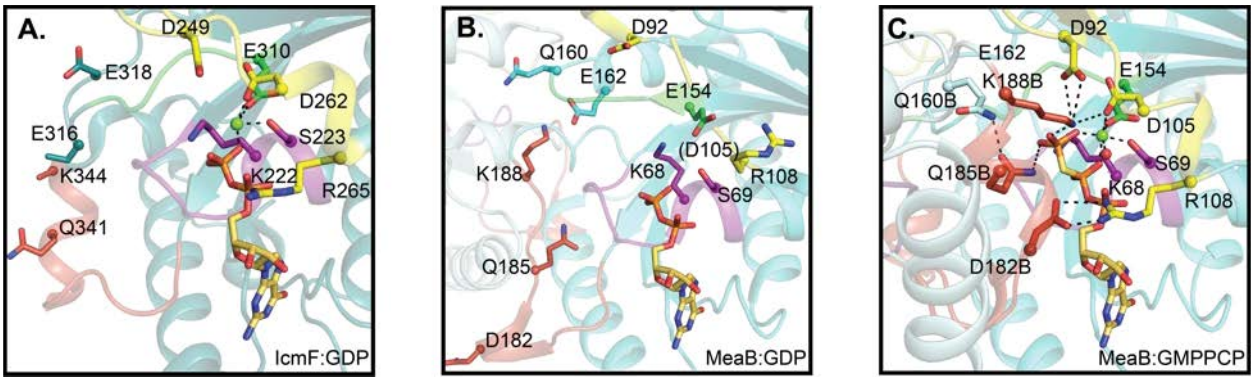


Figure II.9. The Nucleotide Binding Site of IcmF Compared to MeaB. **A.** The GDP-bound G-protein domain of IcmF (PDB 4XC6)[31]. Coloring: P-loop (purple), switch I (yellow), switch II (green), switch III (red orange), additional residues that are outside of the motifs that undergo conformational rearrangements (teal for IcmF, cyan for MeaB). Base specificity loop interactions are shown in *Figure II.7* instead of here for clarity. Mg^{2+} ion in green. **B.** MeaB with GDP (PDB 2QM7)[19] as shown in *Figure II.8*. **C.** GMPPCP-bound MeaB in the presence of *Me*MCM$_{cbl}$ (this work, also shown in *Figure II.8*). The "B" label indicates residues from MeaB chain B.

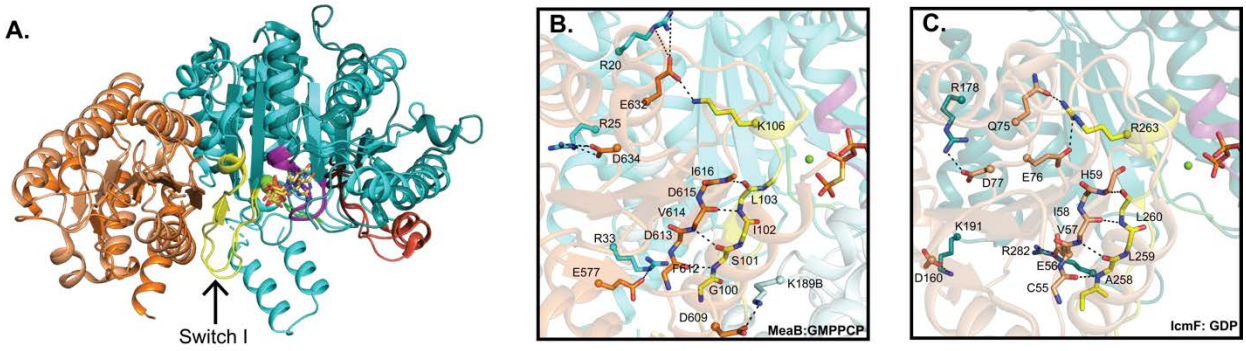


Figure II.10. Interactions at the Interface of the MeaB and MeMCM$_{cbl}$ in the MeaB:MeMCM$_{cbl}$:GMPPCP Complex. **A.** Superimposition of *Me*MCM$_{cbl}$ (orange) and one protomer of MeaB (cyan) from the MeaB:*Me*MCM$_{cbl}$:GMPPCP complex on the Cbl-binding domain (light orange) and G-protein domain (teal) from IcmF (PDB 4XC6)[31] shows a similar interface between domains. Additionally, the conserved regions of the G-protein are in similar orientations. Coloring: Base loop (dark brown), P-loop (purple), switch I (yellow and labeled with an arrow), switch II (green), switch III (red orange). **B.** The interface between MeaB (cyan with switch I region in yellow) and *Me*MCM$_{cbl}$ (orange) primarily consists of the hydrogen bonding of two β-strands, one from MeaB (switch I residues 100 to 103 in yellow) and one from *Me*MCM$_{cbl}$ (residues 612-616 in orange). The side chains of the β-strands are omitted for simplicity. Additionally, the interface has five observed salt bridges, K189B:D609, K106:D632, R33:E577, R25:D634, and R20:D632. The "B" indicates residues from MeaB chain B. **C.** The interface between the G-protein domain (teal with switch I in yellow) and the Cbl-binding domain (light yellow) of IcmF (PDB 4XC6)[31] primarily consists of hydrogen bonding to two β-strands, one from the G-protein domain (switch I residues 258-260 in yellow) and one from the Cbl-binding domain (residues 55-59 also yellow), shown as sticks. The side chains of the β-strands are omitted for simplicity. Additionally, the interface has five observed salt bridges, R178:D77, K191:D160, R263:Q75, R263:E76, and R282:E56.

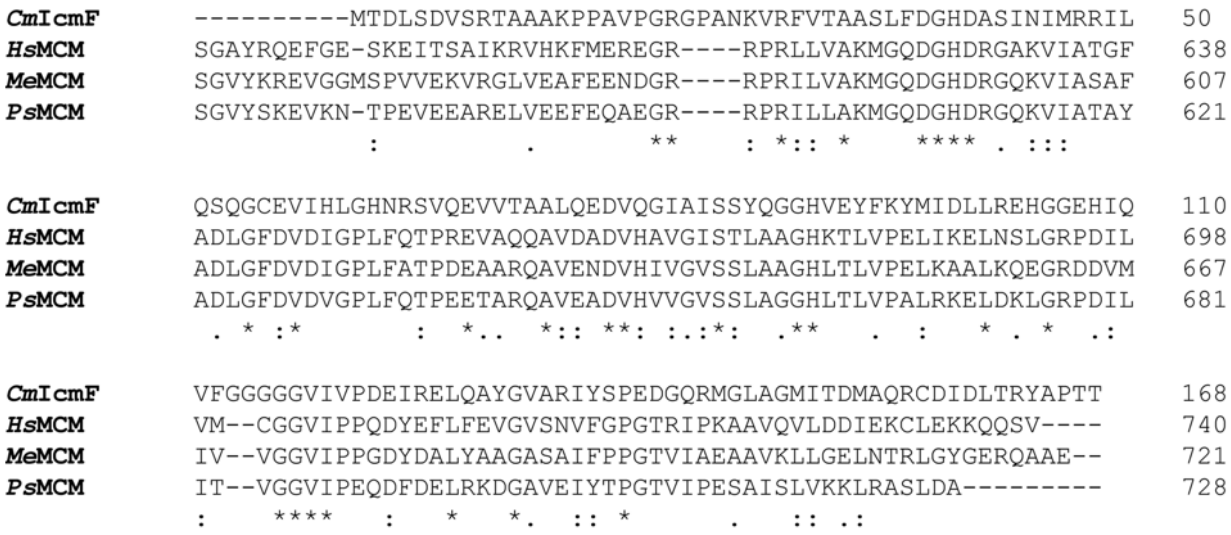


Figure II.11. Sequence alignment of the Cbl-binding domain of CmIcmF, HsMCM, MeMCM, and PsMCM. The sequences aligned are residues 1-168 from *Cm*IcmF, residues 584-740 from *Hs*MCM, residues 552-721 from *Me*MCM active subunit, and residues 567-728 from *Ps*MCM active subunit. Asterisks (*) denote positions with conserved residues. Colons (:) denote positions with conservation with strongly similar properties. Periods (.) denote positions with conservation with weakly similar properties. Alignments performed using ClustalW[47].

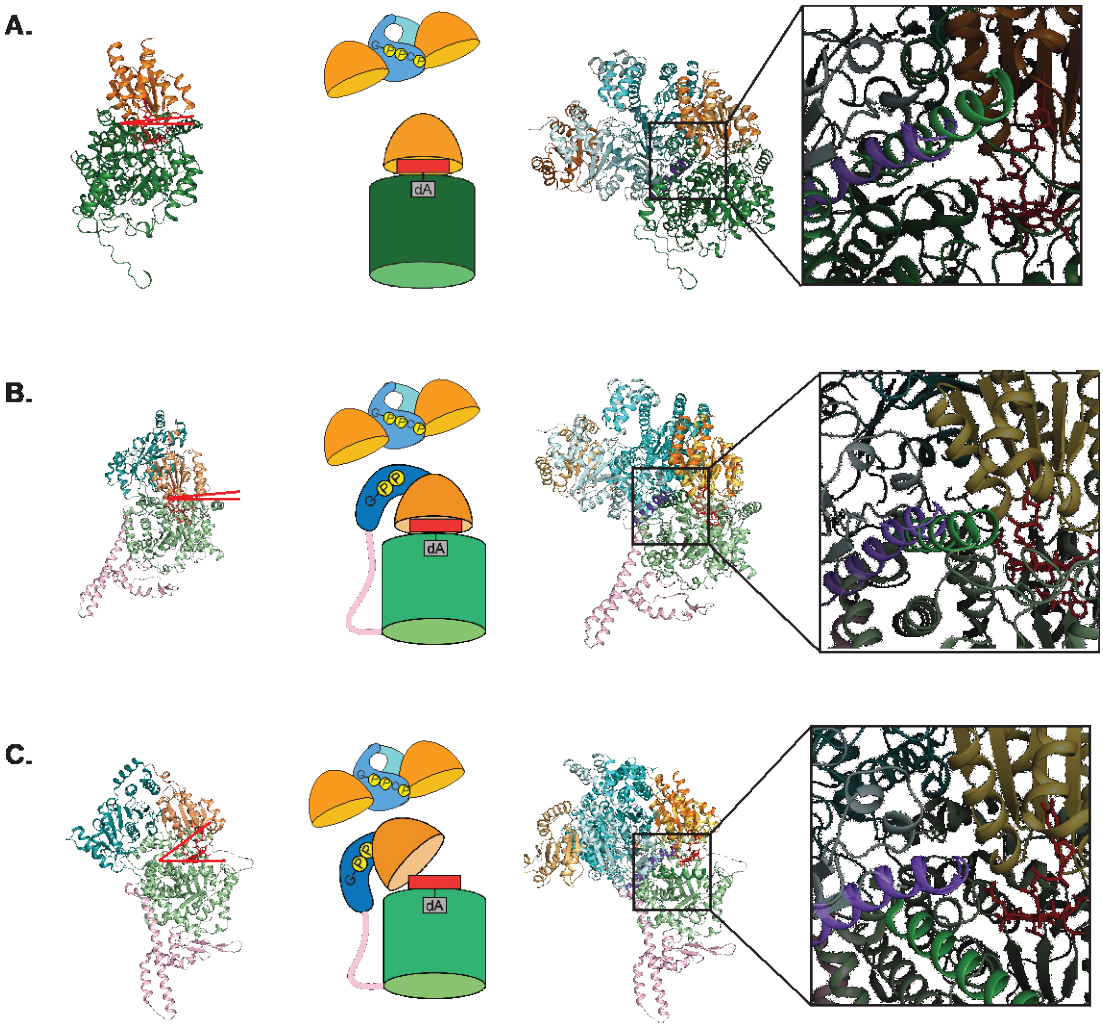


Figure II.12. Superimposition of MeaB:MeMCM$_{cbl}$:GMPPCP onto Full-length Mutase Structures. **A.** (left) Ribbon drawing of *Ps*MCM (PDB 2REQ)[25] with substrate-binding domain (dark green) and Cbl-binding domain (light orange). Red lines indicate that there is no gap between domains; the Cbl is positioned into the barrel for catalysis. (middle) Cartoon of MeaB:*Me*MCM$_{cbl}$ and an active protomer of *Ps*MCM. (right) Overlay of MeaB:*Me*MCM$_{cbl}$ with the *Ps*MCM structure with inset showing clash between an alpha helix of MeaB (residues 206-228 in purple) and the substrate binding domain (residues 441-459 in dark green) of *Ps*MCM. **B.** (left) IcmF (PDB 4XC6)[31] with substrate-binding domain (light green), Cbl-binding domain (light orange), G-protein domain (teal), and linker (pink). Red lines indicate that there is no gap between domains; the Cbl is positioned into the barrel for catalysis. (middle) Cartoon of MeaB:*Me*MCM$_{cbl}$ and an active IcmF protomer. (right) Overlay of MeaB:*Me*MCM$_{cbl}$ with the closed IcmF structure with

insert showing a clash between an alpha helix of MeaB (residues 206-228 in purple) and the substrate binding domain (residues 975-995 in dark green) of IcmF. **C.** (left) Ribbon drawing of open IcmF (PDB 4XC6)[31] with colors indicated in B. Red lines indicate a gap between domains is available for Cbl insertion. (middle) Cartoon of MeaB:*Me*MCM$_{cbl}$ and an inactive protomer of IcmF. (right) Overlay of MeaB:*Me*MCM$_{cbl}$ with the open protomer of IcmF with insert showing that the helix of the substrate binding domain of IcmF (dark green) does not clash with the helix of MeaB (purple) when the mutase is “open”.

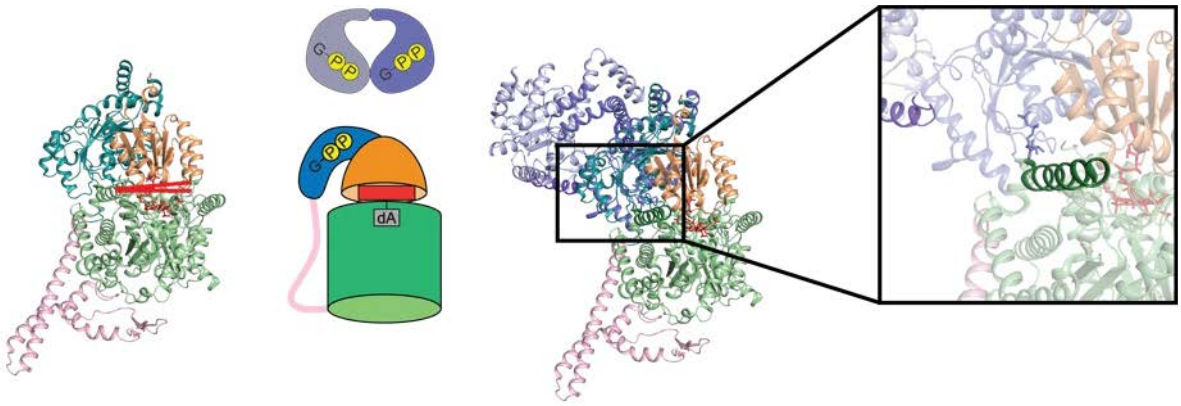


Figure II.13. Structural Superimposition of the Previous MeaB Structure with Full-length IcmF. (left) IcmF (PDB 4XC6)[31] with substrate-binding domain (light green), Cbl-binding domain (light orange), G-protein domain (teal), and linker (pink). Red lines indicate that there is no gap between domains for Cbl insertion, i.e. that IcmF is in the active, closed state. (middle) Cartoon of "inactive conformer" of MeaB bound to GDP and a closed protomer of IcmF. (right) Overlay of the GDP-bound "inactive conformer" of MeaB (purple and light purple) with the closed IcmF protomer structure. Insert shows that the "inactive conformer" of MeaB can bind to the active IcmF protomer with no clashes.

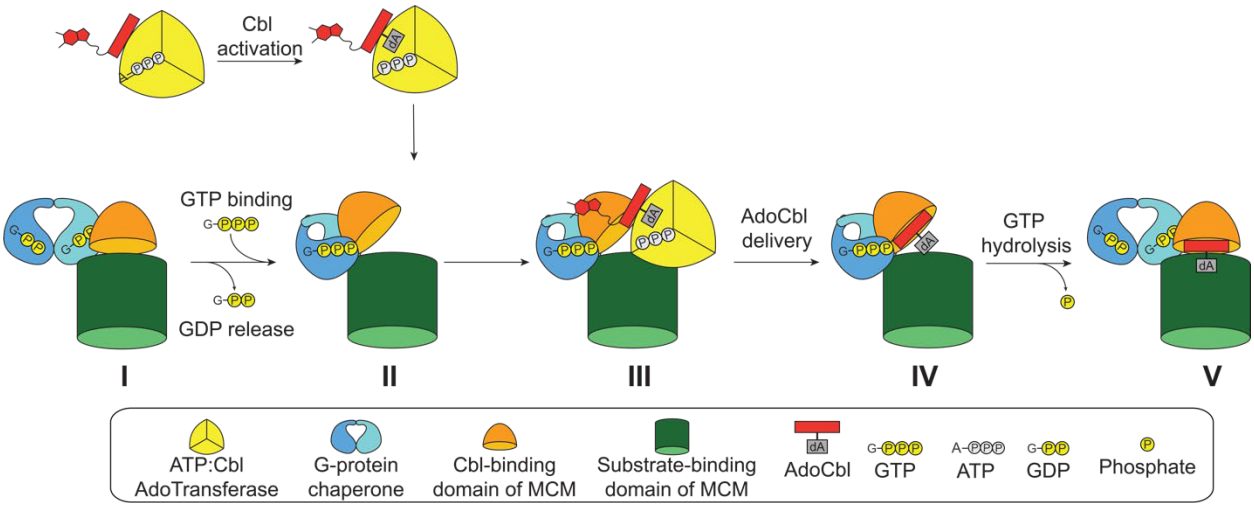


Figure II.14. Proposed Steps for Loading of Mutase Active Site with AdoCbl. In solution, *Me*MCM associates with GDP bound MeaB. The binding of GTP and the Mg^{2+} ion displace GDP. In the presence of both the mutase and Mg^{2+}-GTP bound, MeaB adopts an "active" conformation that is proposed to wedge open the active site of the mutase. The ATR prepares the AdoCbl and inserts it into the wedged open active site of the mutase. After GTP hydrolysis occurs, the conformation of the MeaB returns to the "inactive" state and the mutase closes, trapping the AdoCbl. For simplicity, the inactive subunit of the *Me*MCM, which is not predicted to interact with the G-domain (*Figure II.16*), is omitted.

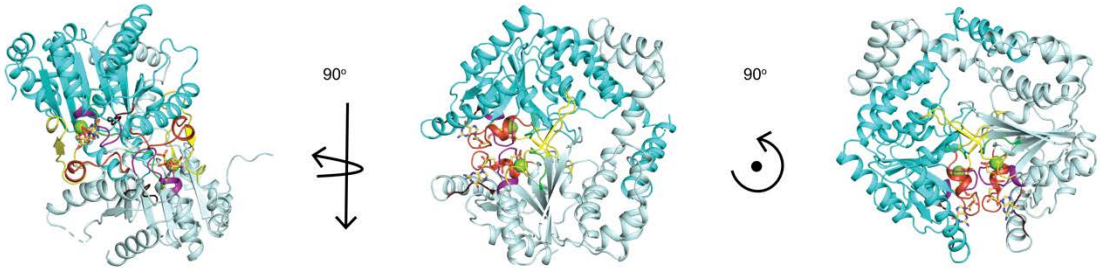


Figure II.15. Comparison of the Views of MeaB. The active conformation of MeaB (chain A: light blue, chain B: cyan) has a new interface comprised of the switch III region (red orange). The base specificity loop (dark brown), P-loop (purple), switch I (yellow), switch II (green) regions interact with the Mg^{2+} ions (green spheres) and GMPPCP (yellow sticks) but are buried from the surface and inaccessible to the mutase. Rotating the active conformation (left) around the y-axis 90˚ (middle) and then another 90˚ around the z-axis (right) yields the side view of MeaB shown in *Figure II.1C, II.5A* and *II.5B*.

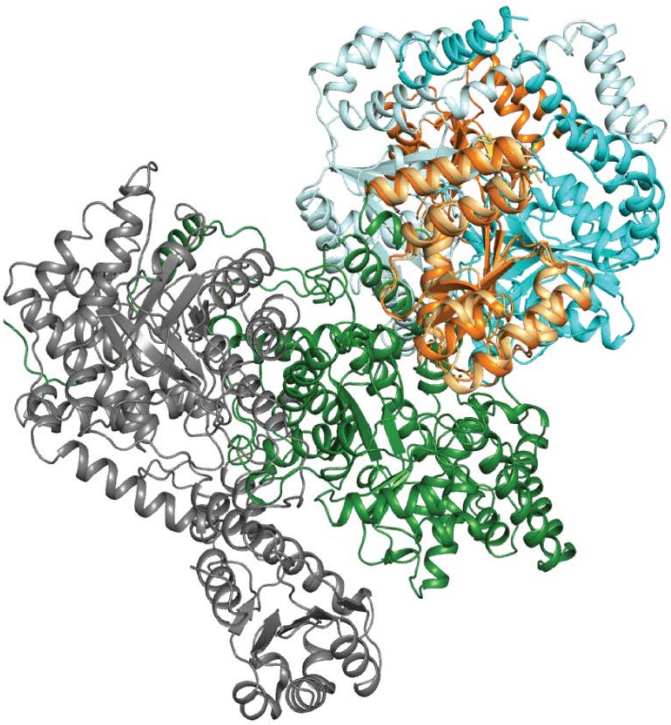


Figure II.16. Structural Superimposition of PsMCM with MeaB:MeMCM$_{cbl}$:GMPPCP Suggests that MeaB Does Not Interact with the Inactive Subunit of MCM. When Cbl-binding domains of *Ps*MCM (PDB 1REQ)[23] and MeaB:*Me*MCM$_{cbl}$:GMPPCP (orange) are superimposed, MeaB (cyan) contacts the active substrate-binding domain of *Ps*MCM (dark green) but not the inactive subunit of *Ps*MCM (gray).

Tables

Table II.1. Data collection and refinement statistics for MeaB:*Me*MCM$_{cbl}$

Data Collection		
Beamline	Cu_{α} Home Source*	APS 24-ID-E
Wavelength (Å)	1.54	0.9792
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions, *a, b, c* (Å)	66.1, 81.0, 166.2	66.2, 81.0, 166.4
Resolution (Å)	45.72-3.10	45.76-2.72
R_{meas} † (%)	27. 4 (103.7)	15.8 (68.9)
</>/<σI>	7.5 (2.2)	10.9 (2.5)
Completeness (%)	94.2 (100.0)**	99.6 (99.0)
Redundancy	7.1 (8.0)**	8.4 (8.4)
CC1/2	98.7 (80.1)	99.6 (80.4)
No. Total Reflections	112637 (8942)	206565 (15753)
No. Unique Reflections	15835 (1120)	24685 (3897)
Refinement		
Resolution (Å)	16.28-3.10	48.99-2.72
No. Reflections used	15696 (1626)	24625 (2366)
R_{work}/R_{free}‡	29.8/31.1	21.0/23.6
No. molecules in asymmetric unit	4	4
MeaB protomers	2	2
*Me*MCM$_{cbl}$ protomers	2	2
No. atoms		
Protein	6464	6501
GMPPCP	64	64
Glycerol	-	12
Mg^{2+}	-	2
Water	-	67
B-factors (Å^2)		
Protein		
Chain A	39.4	48.8
Chain B	40.5	49.7
Chain C	50.2	59.4
Chain D	39.8	51.3
GMPPCP	36.5	36.3
Glycerol	-	53.8
Mg^{2+}	-	33.8
Water	-	39.2
RMS deviations		
Bond lengths (Å)	0.027	0.012
Bond angles (°)	2.64	0.90
Rotamer outliers (%)	26.5	1.66
Ramachandran plot (%)		
Most favored	90.0	98.8
Additionally allowed	6.6	1.2

Disallowed	3.4	0.0

***This model was not refined to completion. **Data are more complete and more redundant in high resolution bin than overall due to loss of data because of ice rings**

† Values in parentheses are for the highest resolution shell.

‡5% of reflections were set aside for a test set

Table II.2. Interactions of the switch III residues in the MeaB:*Me*MCM$_{cbl}$:GMPPCP complex

	MeaB Chain B	Dist. [Å]	MeaB Chain A		MeaB Chain A	Dist. [Å]	MeaB Chain B
1	Asp182[OD1]	2.5	Arg108 [NH2]	1	Asp182[OD1]	2.6	Arg108 [NH2]
2	Asp182[OD2]	3.5	Arg108 [NH1]	2	Asp182[OD2]	3.5	Arg108 [NH1]
3	Gln185[NE2]	3.5	GMPPCP[O1G]	3	Gln185[NE2]	3.1	GMPPCP[O1G]
4	Lys188[NZ]	2.1	GMPPCP[O1G]	4	Lys188[NZ]	2.4	GMPPCP[O1G]
5	Lys188[NZ]	3.4	Asp92[OD1]	5	Lys188[NZ]	3.0	Asp92[OD1]

Table II.3. Hydrogen bonds between MeaB and *Me*MCM$_{cbl}$ in the MeaB:*Me*MCM$_{cbl}$:GMPPCP complex

	MeaB Chain B	Dist. [Å]*	*Me*MCM$_{cbl}$
1	Arg20 [NH2]	2.9	D:Glu632 [OE2]
2	Ala21 [O]	3.7	D:Arg663 [NH2]
3	Arg25 [NH2]	3.3	D:Asp634 [OD1]
4	Arg25 [NE]	3.0	D:Asp634 [OD2]
5	Thr28 [OG1]	2.8	D:Asp634 [OD2]
6	Thr28 [OG1]	3.7	D:Arg582 [NH1]
7	Ser32 [OG]	3.8	D:Gly581 [O]
8	Arg33 [N]	3.0	D:Gly581 [O]
9	Thr98 [O]	3.7	D:Arg583 [NH1]
10	Thr98 [OG1]	3.1	D:Ala608 [O]
11	Ser101 [N]	3.1	D:Phe612 [O]
12	Ser101 [N]	3.9	D:Asp613 [OD1]
13	Ser101 [OG]	2.7	D:Asp613 [OD1]
14	Ser101 [O]	3.0	D:Val614 [N]
15	Ler103 [N]	2.8	D:Val614 [O]
16	Leu103 [O]	3.4	D:Ile616 [N]
17	Lys106 [NZ]	2.8	D:Glu632 [OE1]
18	Lys106 [NZ]	3.0	D:Asn633 [OD1]
19	Thr107 [O]	3.8	D:His596 [NE2]
20	Thr107 [OG1]	2.5	D:Ile616 [O]
21	Ala113 [O]	2.7	D:Arg628 [NH1]
22	Asp115 [O]	3.3	D:Arg628 [NH1]
23	Asp115 [O]	3.5	D:Arg628 [NH2]
24	Arg121 [NH1]	3.3	D:Glu632 [O]
25	Pro122 [O]	3.0	D:Arg585 [NH2]
26	Lys189 [NZ]	2.8	C:Asp609 [OD1]
27	Lys189 [NZ]	3.6	C:Ser605 [OG]

***** Interactions calculated using PDBe PISA v1.52[48]

Table II.4. Salt bridges between MeaB and *Me*MCM$_{cbl}$ in the MeaB:*Me*MCM$_{cbl}$:GMPPCP complex

	MeaB Chain B	Dist. [Å]*	*Me*MCM$_{cbl}$
1	Arg20 [NE]	3.8	D:Glu632 [OE2]
2	Arg20 [NH2]	2.9	D:Glu632 [OE2]
3	Arg25 [NH2]	3.3	D:Asp634 [OD1]
4	Arg25 [NE]	3.3	D:Asp634 [OD1]
5	Arg25 [NE]	3.0	D:Asp634 [OD2]
6	Arg33 [NE]	3.2	D:Glu577 [OE2]
7	Lys106 [NZ]	2.8	D:Glu632 [OE1]
8	Lys189 [NZ]	2.8	C:Asp609 [OD1]

***** Interactions calculated using PDBe PISA v1.52[48]

References

(1) Rosenzweig, A. C. Metallochaperones: bind and deliver. *Chemistry & Biology* **2002**, *9* (6), 673-677. DOI: 10.1016/S1074-5521(02)00156-4.

(2) O'Halloran, T. V.; Culotta, V. C. Metallochaperones, an Intracellular Shuttle Service for Metal Ions *. *Journal of Biological Chemistry* **2000**, *275* (33), 25057-25060. DOI: 10.1074/jbc.R000006200.

(3) Capdevila, Daiana A.; Edmonds, Katherine A.; Giedroc, David P. Metallochaperones and metalloregulation in bacteria. *Essays In Biochemistry* **2017**, *61* (2), 177-200. DOI: 10.1042/ebc20160076.

(4) Leipe, D. D.; Wolf, Y. I.; Koonin, E. V.; Aravind, L. Classification and evolution of P-loop GTPases and related ATPases. *Journal of Molecular Biology* **2002**, *317* (1), 41-72. DOI: 10.1006/jmbi.2001.5378.

(5) Lee, M. H.; Mulrooney, S. B.; Renner, M. J.; Markowicz, Y.; Hausinger, R. P. *Klebsiella aerogenes* urease gene cluster: sequence of *ureD* and demonstration that four accessory genes (*ureD*, *ureE*, *ureF*, and *ureG*) are involved in nickel metallocenter biosynthesis. *Journal of Bacteriology* **1992**, *174* (13), 4324-4330.

(6) Fong, Y. H.; Wong, H. C.; Yuen, M. H.; Lau, P. H.; Chen, Y. W.; Wong, K.-B. Structure of UreG/UreF/UreH complex reveals how urease accessory proteins facilitate maturation of *Helicobacter pylori* urease. *PLoS biology* **2013**, *11* (10), e1001678.

(7) Maier, T.; Jacobi, A.; Sauter, M.; Böck, A. The product of the *hypB* gene, which is required for nickel incorporation into hydrogenases, is a novel guanine nucleotide-binding protein. *Journal of Bacteriology* **1993**, *175* (3), 630-635.

(8) Gasper, R.; Scrima, A.; Wittinghofer, A. Structural insights into HypB, a GTP-binding protein that regulates metal binding. *The Journal of Biological Chemistry* **2006**, *281* (37), 27492-27502. DOI: 10.1074/jbc.M600809200.

(9) Padovani, D.; Labunska, T.; Banerjee, R. Energetics of interaction between the G-protein chaperone, MeaB, and B_{12}-dependent methylmalonyl-CoA mutase. *Journal of Biological Chemistry* **2006**, *281* (26), 17838-17844. DOI: 10.1074/jbc.M600047200.

(10) Takahashi-Iñiguez, T.; González-Noriega, A.; Michalak, C.; Flores, M. E. Human MMAA induces the release of inactive cofactor and restores methylmalonyl-CoA mutase activity through their complex formation. *Biochimie* **2017**, *142*, 191-196. DOI: 10.1016/j.biochi.2017.09.012.

(11) Takahashi-Íñiguez, T.; García-Arellano, H.; Trujillo-Roldán, M. A.; Flores, M. E. Protection and reactivation of human methylmalonyl-CoA mutase by MMAA protein. *Biochemical and Biophysical Research Communications* **2011**, *404* (1), 443-447. DOI: 10.1016/j.bbrc.2010.11.141.

(12) Froese, D. S.; Kochan, G.; Muniz, J. R.; Wu, X.; Gileadi, C.; Ugochukwu, E.; Krysztofinska, E.; Gravel, R. A.; Oppermann, U.; Yue, W. W. Structures of the human GTPase MMAA and vitamin B_{12}-dependent methylmalonyl-CoA mutase and insight into their complex formation. *Journal of Biological Chemistry* **2010**, *285* (49), 38204-38213.

(13) Padovani, D.; Labunska, T.; Palfey, B. A.; Ballou, D. P.; Banerjee, R. Adenosyltransferase tailors and delivers coenzyme B_{12}. *Nature chemical biology* **2008**, *4* (3), 194.

(14) Yamanishi, M.; Vlasie, M.; Banerjee, R. Adenosyltransferase: An enzyme and an escort for coenzyme B_{12}? **2005**, *30*, 304-308. DOI: 10.1016/j.tibs.2005.04.008.

(15) Vaccaro, F. A.; Drennan, C. L. The role of nucleoside triphosphate hydrolase metallochaperones in making metalloenzymes. *Metallomics* **2022**, *14* (6), mfac030.

(16) Gherasim, C.; Lofgren, M.; Banerjee, R. Navigating the B_{12} road: Assimilation, delivery, and disorders of cobalamin. *Journal of Biological Chemistry* **2013**, *288* (19), 13186-13193. DOI: 10.1074/jbc.R113.458810.

(17) Korotkova, N.; Lidstrom, M. E. MeaB is a component of the methylmalonyl-CoA mutase complex required for protection of the enzyme from inactivation. *Journal of Biological Chemistry* **2004**, *279* (14), 13652-13658.

(18) Banerjee, R. Radical carbon skeleton rearrangements: Catalysis by coenzyme B_{12} -dependent mutases. *Chemical Reviews* **2003**, *103* (6), 2083-2094. DOI: 10.1021/cr0204395.

(19) Hubbard, P. A.; Padovani, D.; Labunska, T.; Mahlstedt, S. A.; Banerjee, R.; Drennan, C. L. Crystal structure and mutagenesis of the metallochaperone MeaB: Insight into the causes of methylmalonic aciduria. *Journal of Biological Chemistry* **2007**, *282* (43), 31308-31316. DOI: 10.1074/jbc.M704850200.

(20) Lofgren, M.; Koutmos, M.; Banerjee, R. Autoinhibition and signaling by the switch II motif in the G-protein chaperone of a radical B_{12} enzyme. *Journal of Biological Chemistry* **2013**, *288* (43), 30980-30989. DOI: 10.1074/jbc.M113.499970.

(21) Lofgren, M.; Padovani, D.; Koutmos, M.; Banerjee, R. A switch III motif relays signaling between a B_{12} enzyme and its G-protein chaperone. *Nature Chemical Biology* **2013**, *9* (9), 535-541. DOI: 10.1038/nchembio.1298.

(22) Froese, D. S.; Kochan, G.; Muniz, J. R. C.; Wu, X.; Gileadi, C.; Ugochukwu, E.; Krysztofinska, E.; Gravel, R. A.; Oppermann, U.; Yue, W. W. Structures of the human GTPase MMAA and vitamin B_{12}-dependent methylmalonyl-CoA mutase and insight into their complex formation. *Journal of Biological Chemistry* **2010**, *285* (49), 38204-38213.

(23) Mancia, F.; Keep, N. H.; Nakagawa, A.; Leadlay, P. F.; McSweeney, S.; Rasmussen, B.; secke, P. B.; Diat, O.; Evans, P. R. How coenzyme B_{12} radicals are

generated: the crystal structure of methylmalonyl-coenzyme A mutase at 2 Å resolution. *Structure* **1996**, *4* (3), 339-350. DOI: 10.1016/s0969-2126(96)00037-8.

(24) Mancia, F.; Smith, G. A.; Evans, P. R. Crystal structure of substrate complexes of methylmalonyl-CoA mutase. *Biochemistry* **1999**, *38* (25), 7999-8005. DOI: 10.1021/bi9903852.

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

(26) Vetter, I. R.; Wittinghofer, A. The Guanine Nucleotide-Binding Switch in Three Dimensions. *Science* **2001**, *294* (5545), 1299-1304. DOI: doi:10.1126/science.1062023.

(27) Wittinghofer, A.; Vetter, I. R. Structure-Function Relationships of the G Domain, a Canonical Switch Motif. *Annual Review of Biochemistry* **2011**, *80* (1), 943-971. DOI: 10.1146/annurev-biochem-062708-134043.

(28) Campanello, G. C.; Lofgren, M.; Yokom, A. L.; Southworth, D. R.; Banerjee, R. Switch I-dependent allosteric signaling in a G-protein chaperone–B_{12} enzyme complex. *Journal of Biological Chemistry* **2017**, *292* (43), 17617-17625.

(29) Dempsey-Nunez, L.; Illson, M. L.; Kent, J.; Huang, Q.; Brebner, A.; Watkins, D.; Gilfix, B. M.; Wittwer, C. T.; Rosenblatt, D. S. High resolution melting analysis of the MMAA gene in patients with cblA and in those with undiagnosed methylmalonic aciduria. *Molecular Genetics and Metabolism* **2012**, *107* (3), 363-367.

(30) Cracan, V.; Banerjee, R. Novel coenzyme B_{12}-dependent interconversion of isovaleryl-CoA and pivalyl-CoA. *Journal of Biological Chemistry* **2012**, *287* (6), 3723-3732. DOI: 10.1074/jbc.M111.320051.

(31) Jost, M.; Cracan, V.; Hubbard, P. A.; Banerjee, R.; Drennan, C. L. Visualization of a radical B_{12} enzyme with its G-protein chaperone. *Proceedings of the National Academy of Sciences* **2015**, *112* (8), 2419-2424.

(32) Sprang, S. R. G PROTEIN MECHANISMS: Insights from Structural Analysis. *Annual Review of Biochemistry* **1997**, *66*, 639-678.

(33) Gerwert, K.; Mann, D.; Kötting, C. Common mechanisms of catalysis in small and heterotrimeric GTPases and their respective GAPs. *Biological Chemistry* **2017**, *398* (5-6), 523-533. DOI: doi:10.1515/hsz-2016-0314.

(34) Banerjee, R.; Gherasim, C.; Padovani, D. The tinker, tailor, soldier in intracellular B_{12} trafficking. *Current Opinion in Chemical Biology* **2009**, *13* (4), 484-491. DOI: 10.1016/j.cbpa.2009.07.007.

(35) Dowling, D. P.; Croft, A. K.; Drennan, C. L. Radical Use of Rossmann and TIM Barrel Architectures for Controlling Coenzyme B 12 Chemistry. *Annual Review of Biophysics* **2012**, *41* (1), 403-427. DOI: 10.1146/annurev-biophys-050511-102225.

(36) Born, D. A. Crystallographic Snapshots of Functional Motifs in Cobalamin Maintenance and Methylphosphonate Production. Doctoral dissertation, Harvard University, 2018.

(37) Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **2009**, *6* (5), 343-345, OriginalPaper. DOI: doi:10.1038/nmeth.1318.

(38) Wilkins, M. R.; Gasteiger, E.; Bairoch, A.; Sanchez, J.-C.; Williams, K. L.; Appel, R. D.; Hochstrasser, D. F. Protein Identification and Analysis Tools in the ExPASy Server. In *2-D Proteome Analysis Protocols*, Link, A. J. Ed.; Humana Press, 1999; pp 531-552.

(39) Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **1976**, *72* (1), 248-254. DOI: 10.1016/0003-2697(76)90527-3.

(40) Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallographica Section D* **2010**, *66* (2), 133-144. DOI: doi:10.1107/S0907444909047374.

(41) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *Journal of Applied Crystallography* **2007**, *40* (4), 658-674.

(42) Bunkóczi, G.; Read, R. J. Improvement of molecular-replacement models with Sculptor. *Acta Crystallographica Section D: Biological Crystallography* **2011**, *67* (4), 303-312. DOI: 10.1107/s0907444910051218.

(43) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66* (2), 213-221.

(44) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66* (4), 486-501. DOI: 10.1107/s0907444910007493.

(45) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A. Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography* **2011**, *67* (4), 235-242.

(46) Morin, A.; Eisenbraun, B.; Key, J.; Sanschagrin, P. C.; Timony, M. A.; Ottaviano, M.; Sliz, P. Cutting Edge: Collaboration gets the most out of software. *elife* **2013**, *2*. DOI: doi:10.7554/eLife.01456.

(47) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **2011**, *7* (1), 539. DOI: 10.1038/msb.2011.75.

(48) Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology* **2007**, *372* (3), 774-797.

Chapter III: The role of G-protein domain dimerization in the maturation of a fused isobutyryl-CoA mutase

Author Contributions

Francesca A. Vaccaro conceptualized and performed all experiments with assistance from Daphne A. Faber. Gisele A. Andree collected the negative stain micrographs. David A. Born, Gyunghoon K. Kang, Dallas Fonseca, and Marco Jost assisted with preliminary experiments. F. A. Vaccaro and Catherine L. Drennan analyzed all the data and wrote and edited this chapter.

Summary

G-protein metallochaperones are essential for the correct maturation of numerous metalloenzymes. The human G-protein chaperone MMAA is responsible for facilitating the delivery of adenosylcobalamin (AdoCbl) to the only known AdoCbl-dependent enzyme in humans: methylmalonyl-CoA mutase. It is unknown how the G-protein chaperone dimer prepares the active site of the mutase for cofactor delivery. Here, we present the biochemical and structural investigations of IcmF, a natural fusion protein found in bacteria bearing AdoCbl-dependent isobutyryl-CoA mutase and its corresponding G-protein chaperone, which is homologous to MMAA (MeaB in bacteria). Mutating the conserved residues of the switch III region in the G-protein domain of IcmF, which are implicated in disease states in MMAA, decreases the GTPase activity of the fused IcmF system. Additionally, mass photometry and negative stain electron microscopy (EM) data indicate that the oligomeric state of IcmF in solution is altered by the identity of the G-nucleotide and the identity of the Cbl cofactor in the active site of the mutase. Cryogenic EM reconstructions of IcmF reveal the formation of a G-protein interface in a conformation consistent with the active state conformation of MeaB. This conformation positions the switch III region to interact across the protomer interface, such that it can participate in GTP hydrolysis. The formation of this transient G-protein interface physically props open the active site of the mutase domain for AdoCbl delivery. Notably, these data indicate that the importance of the G-protein dimer to prepare the active site of the mutase for maturation, signifying a consensus mechanism for the role of the G-protein chaperone domain in a fused bacterial system and non-fused systems found in bacteria and humans.

Introduction

With 30-50% of the proteome predicted to be metalloproteins, proper maturation of metalloproteins is a nontrivial biological process. Metallochaperones are essential for the correct maturation of metalloproteins, ensuring that valuable metallocofactors are delivered efficiently and with minimal toxicity and degradation[1-3]. One important class of metallochaperones is the guanine nucleotide-binding proteins (G-proteins), which belong to the SIMIBI (signal recognition particle, MinD, and BioD) class of P-loop NTPases[4]. In

the absence of their target protein, G-protein metallochaperones tend to have low GTP hydrolysis activity. However, target protein binding simulates GTPase by a factor of ~100-fold[5, 6]. Some members of this class, such as the metallochaperones UreG and HypB, directly bind their metallocofactors and then use GTP hydrolysis to facilitate the maturation of their targets, urease, and hydrogenase, respectively[7-10]. Others, like MeaB (MMAA in humans), do not bind directly to their metallocofactor, which is coenzyme B_{12} (5′-deoxyadenosylcobalamin or AdoCbl) in the case of MeaB/MMAA, but still use GTP hydrolysis to deliver AdoCbl to the enzyme target, methylmalonyl-CoA mutase (MCM)[6, 11]. In particular, an adenosyltransferase (ATR) adenylates cob(II)alamin using ATP, and MeaB/MMAA facilitates the transfer of AdoCbl from ATR to the cobalamin(Cbl)-binding domain of MCM[12, 13]. Typically, G-protein metallochaperones are standalone proteins, with the notable exception of AdoCbl-dependent isobutyryl-CoA mutase fused (IcmF), in which the G-protein metallochaperone exists as a domain of the target enzyme[14]. Mutations or deletions in the genes encoding metallochaperones can impair metalloprotein function in vivo and lead to disease in humans. For example, in humans, mutations to the genes for MCM, MMAA or any other chaperones involved in B_{12} trafficking results leads to the disease phenotype methylmalonic aciduria, an inborn error of metabolism[13, 15].

The AdoCbl cofactor of a matured mutase is essential for the chemically challenging carbon skeletal rearrangements that are performed[16]. For MCM, the radical reservoir of the cobalt—carbon bond of AdoCbl catalyzes the 1,2-rearrangement of (*R*)-methylmalonyl-CoA to succinyl-CoA (*Figure III.1A*)[17]. IcmF catalyzes the 1,2-rearrangement of isobutyryl-CoA to *n*-butyryl-CoA, as well as pivalyl-CoA and isovaleryl-CoA, using AdoCbl (*Figure III.1A*)[14, 18, 19]. The homolytic cleavage of the cobalt—carbon bond of AdoCbl generates cob(II)alamin and a highly reactive 5'-deoxyadenosyl radical species[20]. After catalysis, these species must come together to regenerate AdoCbl (*Figure III.2*); however, the deoxyadenosine moiety can be lost and/or the cob(II)alamin species is susceptible to oxidation to hydroxocobalamin (OHCbl), preventing reformation of AdoCbl and inactivating the enzyme[19]. AdoCbl is also susceptible to inactivation by photolysis of the cobalt—carbon bond of AdoCbl[21]. To restore activity, the G-protein metallochaperone facilitates the removal of the damaged cofactor and the insertion of

new cofactor from ATR[11, 22, 23]. In an active conformation of all AdoCbl-dependent mutases, AdoCbl is positioned at the interface of a Cbl-binding Rossmann domain and a substrate-binding TIM barrel to afford the generation of substrate-radical species (*Figure III.1B, Figure III.2*)[24-26]. In this enzyme active state, AdoCbl is sequestered. Thus, AdoCbl delivery to the mutase requires a transient opening and closing of the mutase structure, a conformational change that is believed to be facilitated by MeaB/MMAA[23, 27, 28] .

Consistent with their classification as P-loop NTPases, MeaB/MMAA and the G-protein domain of IcmF all have a P-loop, a base-specificity loop, and two conserved switch regions for signal transduction: switch I (residues 92-108 in MeaB) and switch II (residues 154-158 in MeaB) (*Figure III.1C*)[5, 29-31]. Unique to MeaB and MMAA is a so-called switch III region[30], which was identified through investigation of residues associated with methylmalonic aciduria[15]. In vitro, the substitution of MeaB switch III residues Lys188, Gln185, or Asp182 with alanine (*Figure III.1C*) reduces the stimulatory effect in GTP hydrolysis afforded by mutase binding, i.e. the GAP (GTPase-accelerating protein) activity, and also leads to an uncoupling of GTP hydrolysis from AdoCbl transfer[30]. In the absence of its target mutase, the switch III region of MeaB points toward solution (*Figure III.1D*). However, following the binding of MeaB to the Cbl-binding domain of MCM in the presence of a nonhydrolyzable analog of GTP, Guanosine-5'-[(β,γ)-methyleno]triphosphate (GMPPCP), MeaB undergoes a 180° rotation that results in the re-positioning of switch III residues directly into the GTP binding site (*Figure III.1D*) (Chapter II)[32]. Importantly, it is the switch III residues of the neighboring protomer that contact GTP in this active conformation, indicating the importance of the dimeric structure of MeaB to its function (*Figure III.1C,D*)[32].

The recent structure of GMPPCP-MeaB bound to the Cbl-binding domain of MCM (GMPPCP-MeaB:MCM_{Cbl}) described in Chapter II led to a proposed molecular mechanism for MeaB function (*Figure III.3A*)[32]. In this proposal, the association of GTP-MeaB with the Cbl-binding domain of MCM leads to a conformational change of MeaB from an inactive to active state. The MeaB active state stabilizes an open MCM conformation, allowing MCM to receive AdoCbl from ATR. GTP hydrolysis causes MeaB to undergo a conformational change back to the inactive state. This conformational

change of MeaB is proposed to destabilize the open MCM conformation, causing MCM to close and capture AdoCbl inside of the enzyme (*Figure III.3A*).

This proposed mechanism (*Figure III.3A*) is supported by crystal structures of GDP-MeaB alone and GMPPCP-MeaB in the presence of MCM_{Cbl} that together reveal the conformational gymnastics involved in conversion of inactive and active MeaB states (*Figure III.1D*)[31, 32]. The connection between GTP hydrolysis and AdoCbl capture, displayed in *Figure III.3A* (IV to V), is also supported by biochemical and structural data. Briefly, MCM is unable to capture AdoCbl in the presence of a non-hydrolyzable GTP analog, consistent with the proposal in *Figure III.3A* that a GTP-hydrolysis-driven MeaB conformational change is needed for MCM to close and thus capture AdoCbl[18, 33]. Furthermore, when the active state of MeaB is destabilized by substitutions of switch III residues, such that MeaB can convert to the inactive state independent of GTP hydrolysis, AdoCbl capture is uncoupled from MeaB's GTPase activity[27, 30]. What is missing in terms of experimental support of *Figure III.3A* are structures of MCM and MeaB showing that an active MeaB conformation stabilizes an open MCM state in which the Cbl-binding domain of the mutase is positioned away from the substrate-binding domain (II, III, IV in *Figure III.3A*). There is no structure of MeaB bound to an intact MCM; only a structure of MeaB bound to the Cbl-binding domain of MCM (MCM_{Cbl}). Thus, part of the mechanistic proposal in *Figure III.3A* is based on structural superpositions of GMPPCP-MeaB:MCM_{Cbl} with the crystal structure of IcmF from *Cupriavidus metallidurans*[28], which contains all relevant structural units: the metallochaperone G-protein domain, the Cbl-binding domain, and the substrate-binding TIM barrel and shares a high level of structural similarity (*Figure III.1E,4,5*). There is, however, the question as to whether IcmF is a good model system for a dimeric MeaB:MCM complex given that the G-protein domains of IcmF are monomeric and located on opposite sides of this large enzyme structure (*Figure III.1E*). IcmF is a dimer, but the dimeric interface is not composed of the G-protein domains (*Figure III.1E*). Thus, if dimerization of the G-protein domains is essential for AdoCbl delivery in IcmF as it is for MeaB, then IcmF would need to form transient higher-order oligomers (*Figure III.3B*). If oligomers do not form, then IcmF must use a distinct molecular mechanism from MeaB and not what is shown in *Figure III.3B*. Also, if IcmF is

a good model system for MeaB:MCM, then IcmF should also employ a switch III, a question that has not been investigated.

In this study, we investigate whether IcmF, a fused G-protein system, uses the same molecular mechanism as a standalone G-protein system. By employing site-directed mutagenesis and enzyme assays, we establish that switch III residues are relevant in IcmF from *C. metalliduran*s, and by employing negative stain electron microscopy (EM) and mass photometry, we show that IcmF oligomerizes as one would expect for a conserved molecular mechanism. With support for a conserved mechanism, we go on to use IcmF to obtain the missing structural snapshot, a mutase bound to a G-protein in a state competent for AdoCbl transfer. With these new data, we describe a consensus molecular mechanism for metallochaperone-assisted AdoCbl-dependent mutase maturation.

Results

Substitutions of the switch III regions of IcmF decrease GTPase activity, establishing the relevance of switch III in the IcmF system

To assess if the residues in IcmF that are analogous to the switch III residues of MeaB affect GTPase activity of IcmF as they do in MeaB, we substituted Q341 (Q185 in MeaB) and K344 (K188 in MeaB) with alanine residues and measured GTPase activity (*Table III.1*, *Figure III.5,6*)[30]. The GTPase activity of *C. metallidurans* IcmF wt (k_{cat} = 3.13 $\pm$ 0.18 min^{-1}) is comparable to previously reported values for IcmF from *C. metallidurans* and *Geobacillus kaustophilus* (*Table III.1*)[14, 19]. Substituting the switch III residue Q341 with alanine lowers the catalytic efficiency 2.6–fold mainly through a decrease in k_{cat} (*Table III.1*). Substituting the switch III residue K344 with alanine lowers the GTPase activity to undetectable levels (*Table III.1*). The observed decrease in GTPase activity of the Q341A and K344A variants is consistent with the observed decrease in the GTPase activity of MeaB switch III variants, validating the importance of the switch III residues in the fused system in addition to the non-fused G-protein:mutase system[30].

IcmF forms higher order oligomers in the presence of GTP and non-hydrolyzable GTP analogs

To understand if the fused system also utilizes a G-protein dimer arrangement similar to MeaB and other members of the SIMIBI G-proteins[4], we analyzed the oligomeric state of *C. metallidurans* IcmF wt in solution in the presence and absence of various G-nucleotides. If the fused system's active state is comparable to the non-fused system, IcmF should form a higher order oligomeric state in the presence of GTP or a non-hydrolyzable analog due to the association of one G-protein domain of an IcmF molecule with the G-protein domain of another. GTP hydrolysis should break apart these higher order oligomeric states, returning IcmF to the dimeric state that was visualized in the crystal structures (*Figure III.1E*)[28, 34]. Thus, we would expect to see more higher order oligomers with GMPPCP, a non-hydrolyzable analog of GTP. Additionally, we would expect to see only dimers under inactive GTPase conditions, i.e. with no nucleotide or with GDP. Importantly, our negative stain EM and mass photometry data are consistent with these predictions (*Figure III.7*). Negative stain EM analysis indicates that with no nucleotide or in the presence of 500 μM GDP, IcmF does not form any observable higher order oligomers and remains dimeric (*Figure III.7A,B*). In the presence of 500 μM GTP, the percentage of IcmF wt in a dimeric state decreases and higher order oligomers are observed (*Figure III.7C*). With 500 μM of the non-hydrolyzable GTP analog, GMPPCP, long chains of IcmF protomers are observed, with all visible IcmF protomers comprising these filamentous-like chains (*Figure III.7D)*. These supramolecular structures are consistent with IcmF dimers interacting with other IcmF dimers through the surface exposed G-protein domains on either end of the long IcmF molecule (see *Figure III.1E*).

Using mass photometry to evaluate the oligomers, we also find that in the presence of 500 μM GDP, the primary state of IcmF is dimeric (80%), consistent with the negative stain EM (*Figure III.7E*) and the crystal structure (*Figure III.1E*)[28]. In the presence of 500 μM GTP, the percentage of IcmF in a dimeric state is decreased from 80% to 60% and higher order oligomers are observed (22% of particles are tetramers and 7.5% of particles are hexamers). The dimensions and heterogeneity in length of the GMPPCP-generated IcmF supramolecular complexes precluded accurate quantification by mass photometry.

Substitutions of the switch III residues uncouple oligomerization from nucleotide-bound state

With the finding that substituting switch III residues K344 and Q341 in IcmF with alanine decrease GTPase activity, we investigated the effect of switch III residue substitutions on IcmF oligomeric state by negative stain EM and mass photometry. In the absence of any nucleotides, switch III residue variants' primary oligomeric state is dimeric, the same as IcmF wt, and in the presence of 500 μM non-hydrolyzable GTP analog, GMPPCP, long chains of IcmF molecules are formed, the same as IcmF wt (*Figure III.8A,B*). The main difference is that the correlation is broken between the activity state of the G-protein domain (active GTP or inactive GDP) and the oligomeric state of IcmF.

The identity of the cobalamin cofactor influences IcmF's oligomeric state

According to the mechanism shown in *Figure III.3B*, we would expect that the G-protein domain of IcmF would only dimerize when IcmF needs to open for removal of a damaged cofactor and installation of a new one. Thus, we investigated whether a correlation exists between the identity of the Cbl cofactor (active AdoCbl or inactive cob(II)alamin or OHCbl) and the oligomeric state of IcmF. We employed negative stain EM exclusively for our studies with Cbl, as the Cbl visual spectra interfered with the light scattering method utilized in mass photometry. We carried out the preparation of negative stain EM both under red light ("dark") and under white light ("light"). Samples prepared under red light should have intact AdoCbl, and experiments conducted in the light will have some amount of inactivated cob(II)alamin cofactor due to photolysis of AdoCbl and loss of the adenosyl moiety. We also used OHCbl in these studies, which represents the inactivated and oxidized AdoCbl degradation product.

We find that with no nucleotide or with GDP that the identity of the Cbl present (AdoCbl in dark or light or OHCbl) does not alter to any appreciable extent *C. metallidurans* IcmF oligomerization. IcmF appears to be dimeric in all cases (*Figure III.9A,B*). When GMPPCP is used, supramolecular complexes are observed in all cases. However, more supramolecular complexes are apparent for samples prepared in the light, when more AdoCbl has been photolyzed and in the presence of OHCbl, than for AdoCbl samples prepared in the dark (*Figure III.9C*). We have previously noted that GMPPCP is more effective at inducing complex formation between MeaB and MCM_{Cbl} in comparison

to another non-hydrolyzable GTP analog, guanosine-5'-[(β,γ)-imido]triphosphate (GMPPNP) (*Figure III.10*)[32]. Thus, we wondered if usage of the weaker oligomerization-agonist GMPPNP would allow us to observe smaller effects on oligomerization that are due to the nature of the Cbl cofactor. We find that there are very few, if any, IcmF supramolecular complexes in the presence of GMPPNP when AdoCbl is intact (dark sample) (*Figure III.9D*). Photolysis of AdoCbl leads to more supramolecular complex formation but not as many as in the GMPPNP + OHCbl sample (*Figure III.9D*). Together, these data indicate that IcmF oligomerization depends more strongly on the identity of the G-nucleotide than the identity of the Cbl cofactor, but IcmF with a damaged Cbl is more prone to oligomerization compared to IcmF with an intact AdoCbl. Thus, both the G-nucleotide identity and the Cbl identity shift the oligomeric state equilibrium, but to different degrees.

Cryogenic EM data reveal a G-protein dimer and an open conformation of the mutase

To further investigate the structures of IcmF's supramolecular complexes, we prepared cryogenic EM grids and collected datasets of *C. metallidurans* IcmF wt with GMPPCP and butyryl-CoA (IcmF wt + GMPPCP) and IcmF Q341A with GTP (IcmF Q341A + GTP). For the IcmF wt + GMPPCP dataset, the supramolecular complexes were first manually picked on the micrographs, extracted using helical parameters, and processed as a single particle dataset. Signal subtraction was utilized to remove the extra density that was due to the extraction of single particles from a continuous helical particle (*Figure III.11*)[35]. Local B factor post-processing was performed to yield a 5.0 Å resolution reconstruction (*Figure III.12A, Table III.2*)[36]. For the IcmF Q341A + GTP dataset, we performed 3D single particle reconstructions, using repeated classifications for both 2D and 3D steps to separate out heterogeneity (*Figure III.13*). Local B factor post-processing was also performed to yield a 4.6 Å resolution reconstruction (*Figure III.13C*, *Table III.2*). For each reconstruction, the previously solved crystal structure of *C. metallidurans* IcmF with GDP and AdoCbl bound was used as the initial docking model. In particular, residues 22 - 1093 from chain B of IcmF from PDB 4XC6[28], was modeled into the EM maps in two pieces: fragment one consisted of the Cbl-binding domain and the G-protein domain (residues 22 - 442); and fragment two consisted of substrate-binding domain and the

linker region (residues 443 - 1093). For the IcmF wt + GMPPCP reconstruction, four complete protomers (eight fragments) were docked into the density and subjected to rigid body real-space refinement (*Figure III.12A*). For the IcmF Q341A + GTP reconstruction, three complete protomers (six fragments) were docked into the density and subjected to rigid body real-space refinement with minimization (*Figure III.12C*).

In both reconstructions, one IcmF dimer contacts the other IcmF dimer through intermolecular interactions made by the G-protein domains (cyan in *Figure III.12B,D*). The G-protein domain interface adopts the same conformation as observed for MeaB in the GMPPCP-MeaB:MCM_{Cbl} complex structure. Again, we observe switch III residues positioned at the interface adjacent to the GTP binding site (*Figure III.14A*,*15*). For the IcmF Q341A + GTP reconstruction, there is clear density for the guanosine and at least two phosphates of GTP in all the nucleotide binding sites (*Figure III.12D*). However, due to the lower resolution of the IcmF wt + GMPPCP reconstruction, there is no clear density for GMPPCP in the nucleotide binding site. For both reconstructions, we observed an open conformation of the mutase active site (*Figure III.12B,D*).

Overlaying previously determined closed (*Figure III.16A*) and open (*Figure III.16C*) conformations of IcmF onto the structure of IcmF Q341A + GTP reveals that the G-protein interdimer interface clashes with the closed conformation (*Figure III.16B*). Specifically, the switch III region (residues 333-344) and a helix (purple in *Figure III.16B*, residues 363-380) of one protomer clash with two helices of the substrate binding domain of the other protomer (helix 1, residues 942-965, and helix 2, residues 974-995 in *Figure III.16B*); these clashes are not observed when the mutase is in an open conformation (*Figure III.16D*). Thus, in addition to switch III residues residing at the interface created by the G-protein domains, they also participate in propping open the mutase (*Figure III.16B,D*). These structures are consistent with the biochemical solution state data, showing that supramolecular complexes form under conditions that require loading or unloading of Cbl. Overall, the cryo-EM reconstructions have trapped the conformation of G-protein domain that wedges open the active site of the mutase domain for cofactor to be loaded and unloaded.

Discussion

Metallochaperones ensure that valuable, and often highly reactive, metallocofactors are delivered efficiently to their target enzymes. The molecular basis for successful delivery is often enigmatic, especially for metallochaperones like MeaB that do not bind the metallocofactor directly. A major hurdle in this research is that the complexes formed by the metallochaperone and target enzyme are transient and can have more than one conformational state. Therefore, understanding the molecular basis of metalloenzyme maturation and/or cofactor repair, can involve the difficult task of trapping transient protein:protein complexes in multiple conformational states. In this study, we use cryo-electron microscopy to capture a long-awaited structure of the active conformation of an AdoCbl metallochaperone in complex with a target mutase and use mutagenesis and mass photometry to understand the assembly/disassembly of active cofactor-transfer complexes. Our data support the existence of a conserved molecular mechanism for AdoCbl transfer between the fused IcmF system and the standalone MeaB:MCM system and provide the snapshots needed to understand the molecular basis of metallochaperone-assisted AdoCbl transfer.

Previous studies showed that the bacterial AdoCbl metallochaperone MeaB are always dimeric and uses a dramatic 180° conformational change to switch between active (GTP-bound) and inactive (GDP-bound) states (*Figure III.1D*) (Chapter II)[31, 32]. In contrast, we previously showed that the fused metallochaperone:mutase system IcmF has a chaperone G-protein domain that is monomeric in the inactive state (*Figure III.1E*) [28] and show here that IcmF uses oligomerization to switch between its inactive (GDP-bound) states and the active (GTP-bound) state. Evidence for the coupling of GTP binding to an IcmF oligomeric state change comes from negative stain EM, cryo-EM, and mass photometry (*Figure III.7,12*). Importantly, regardless of whether the active conformation of the G-protein domain is formed through oligomerization (IcmF) or through a conformational change of an obligate dimer (MeaB/MMAA), the interface between protomers is the same (*Figure III.15*).

The G-protein:G-protein interface is formed in both systems by switch III residues of one protomer, and the G-nucleotide and switch I region of the other protomer (*Figure III.1C, inset*, and *Figure III.12D, inset*) (Chapter II). Notably, substitutions of switch III

residues uncouple the G-nucleotide state from oligomerization (*Figure III.8*), consistent with a juxtaposition of switch III residues to the G-nucleotide binding site. This G-protein:G-protein interface is small, which allows for modest changes, such as loss of a phosphate group due to GTP hydrolysis, to easily shift the conformational (or oligomeric state) equilibrium between active and inactive G-protein states. Insight into the molecular basis for this conformational/oligomeric state shift comes from structural comparisons of GMPPCP-bound MeaB:MCM_{Cbl} (Chapter II)[32] and GDP-bound MeaB[31] which indicate that GTP hydrolysis would bring about the loss of the contacts made by chain-B switch III residues K188 and Q185 (MeaB numbering) to the chain-A terminal phosphate of GMPPCP, and loss of the interchain salt bridge made by switch III residue D182 (chain-B) and switch I residue R108 (chain-A, MeaB numbering) (*Figure III.1C, inset*). In terms of the salt bridge, structural comparisons suggest that the re-positioning of intradomain MeaB residues D92, E154, and R108 (*Figure III.1C, inset*) as a result of GTP hydrolysis and Mg^{2+} loss, in turn re-positions R108, breaking the interchain salt bridge (see *Figure II.8*). Collectively, the loss of the terminal phosphate of GTP results in a loss of all contacts that stabilize the active state of MeaB. Although the resolution of cryo-EM IcmF is too low for a detailed analysis of protein:G-nucleotide interactions, we do find that the switch III residues, including K344 (K188 in MeaB) and Q341 (Q185 in MeaB) of one IcmF protomer are in direct contact with the G-nucleotide bound to another IcmF protomer. This arrangement in IcmF is just what one would expect for a conserved mechanism in which the binding/orientation of one G-protein protomer is based on the G-nucleotide-bound state of another G-protein protomer.

Despite the modest resolution of the cryo-EM structures determined in this study, establishing the positioning of the individual domains of IcmF was straightforward (*Figure III.12*). The resulting model shows that the active G-protein:G-protein interface stabilizes an open conformation of the mutase that is ready for AdoCbl delivery. It was previously proposed through structural superimpositions that the active G-protein conformation would stabilize the Cbl-binding domain away from the substrate-binding domain for cofactor loading in MeaB:MCM (*Figure III.3A*) and similarly in IcmF (*Figure III.3B*), and here we see that this is in fact the case (*Figure III.12*). The second protomer of the G-domain dimer is wedged between the Cbl-binding domain and the substrate-binding

domain, holding the mutase open (*Figure III.16D, left inset*). The closed conformation of the mutase appears incapable of interacting with the active G-protein state based on the structural superimpositions of IcmF structures (*Figure III.16B, left inset*). We previously predicted that the formation of the active state of MCM-bound MeaB would lead to a clash between a helix of MeaB (equivalent to the IcmF helix shown in purple in *Figure III.16B, left inset*) and a helix of the substrate-binding domain (helix 1 in *Figure III.16B, left inset*) if MCM remained in a closed state (Chapter II)[32]. This prediction is consistent with our current structures. Additionally, structure superimpositions suggest that the switch III region (red in *Figure III.16B, left inset*) also would make unfavorably close interactions with a substrate-binding domain helix (helix 2 in *Figure III.17B, left inset*). This observation provides an additional role for switch III; acting as a molecular wedge that is sensitive to the identity of the G-nucleotide state. The wedge is secured with GTP bound, allowing AdoCbl to be delivered to an open structure, and the wedge is loosened by GTP hydrolysis, allowing the Cbl-binding domain to close, trapping a delivered AdoCbl (*Figure III.3B*).

Chemical logic dictates that a metallochaperone delivery/repair system be designed such that the target enzyme structure would only be opened for cofactor delivery in the apo-state of enzyme and/or following cofactor damage; it is wasteful to replace a working cofactor. The structure of IcmF suggests that the Cbl itself may regulate the open/closed equilibrium of the mutase (*Figure III.17*). In particular, IcmF structures show that the adenosyl moiety of AdoCbl reaches across the boundary between the Cbl-binding domain to the substrate-binding domain, thereby securing the domains together[28] and shifting the conformational equilibrium toward a closed mutase state. Loss of the adenosyl moiety to due to photolysis or oxidative damage is expected loosen the connection between domains, facilitating the movement of the Cbl-binding domain and shifting the conformational equilibrium toward the open mutase state[24-26, 28]. The cryo-EM structure presented here shows that an open conformation of IcmF creates a pocket that can be filled by a neighboring IcmF's G-protein domain, and in particular, its wedge helix (purple residues 363-380) and its switch IIII region (*Figure III.16D*). Thus, loss of AdoCbl or damage to AdoCbl should shift the conformational equilibrium of IcmF to an open state, facilitating the binding of a G-domain of a neighboring IcmF molecule (*Figure III.17*) and

starting the repair process. In addition to data showing that switch III residues play a role in the removal of damaged Cbl[30], our negative stain EM data using GMPPNP and GMPPCP show more extensive IcmF oligomerization when AdoCbl is subjected to photolysis or replaced with OHCbl than for intact AdoCbl (*Figure III.9C,D*). We expect that in vivo, the conformational equilibrium shifts toward the open mutase structure due to Cbl damage or AdoCbl loss will be necessary for oligomerization, limiting the wasteful replacement of a working cofactor. In terms of maturation, the formation of the supramolecular complexes made up of apo-IcmF molecules with GMPPCP (*Figure III.7A*, *right*) invokes an assembly line model in which a chain of open mutases can be filled in succession by an ATR that moves down the chain, delivering AdoCbl as it is synthesized. A structure of these IcmF supramolecular complexes in the presence of ATR would be most informative in this regard.

Collectively, the structural and biochemical data obtained previously[5, 28, 31] and presented here indicate that the only significant difference between the standalone system (*Figure III.3A*) and the fused IcmF (*Figure III.3B*) is whether the active G-protein state is formed via a conformational change or via oligomerization. It is not unprecedented for signaling proteins to employ both methods (conformational change and oligomerization) to switch between active and inactive states. For example, members of the CAP family of transcription factors can use either ligand binding to induce dimerization that affords DNA association or ligand binding to induce a conformational change that affords DNA association[37, 38]. The use of two different methods for creating an active G-protein state makes sense if one considers that G-protein chaperones are designed to be poor GTPases in the absence of their target protein to prevent unwanted GTP hydrolysis. Since the G-protein and target protein are fused in IcmF, unwanted GTP hydrolysis would be a large problem if the active G-protein state was easily formed. For MeaB, MCM is its GAP (GTPase-accelerating protein). For the G-protein domain of IcmF, a second IcmF protomer is the GAP, our cryo-EM data show that residues required for GTP hydrolysis (K344 for example) are on the second protomer (*Figure III.12*). Thus, the active site for GTP hydrolysis is missing in IcmF's resting homodimeric state, which is the state of IcmF wt that is visible in the absence of GTP or GTP analogs (*Figure III.7*). If the G-protein domains of IcmF were present as an obligate dimer, allowing a conformational change to

create an active site capable of GTP hydrolysis, chemical logic suggests that the unwanted GTP hydrolysis would be a more substantial problem. IcmF is a poor GTPase because its active site residues are not part of its resting structure.

The presence of fused systems has not been reported for other members of the SIMIBI class of P-loop G-proteins, such as UreG and HypE which also rely on binding a G-nucleotide to form the active conformation of the G-protein. Instead UreG and HypE resemble MeaB[39]; however, the degree to which these metallochaperones utilize the same molecular mechanisms is unclear. Each metallochaperone system appears to have a different number of accessory proteins, which are commonly of unknown function and unknown structure, and if structures exist, the structures are often of inactive states or isolated states, not of the protein:protein complex that is responsible for maturation. Our results here suggest that cryo-EM is likely to be crucial for obtaining structures of transient protein complexes, affording new snapshots of the metallochaperone delivery and repair processes. These new snapshots will expand the industrial uses of metalloenzymes by increasing our understanding of the necessary components for maximal efficiency. Industrial uses of metalloenzymes have the potential to be more environmentally friendly, as in the use of hydrogenase-biofuel cells for example, but such applications require an improved understanding of metalloenzyme maturation. We hope that the studies presented here will benefit metalloenzyme applications in industry and in medicine by providing structural and mechanistic insight into one system in detail. With the cryo-EM resolution revolution, we hope that this work is the beginning of what will be an exciting decade for this field.

Materials

All chemicals, solvents and reagents were purchased from Sigma-Aldrich unless otherwise noted.

Methods

Plasmids

Name	Features	Source
pET28a_*cm*IcmF	N-terminal His-tag, thrombin cleavage site, T7 protomer, Kan^R	Ref. [28]

pET28a_*cm*IcmF_Q341A	N-terminal His-tag, thrombin cleavage site, T7 protomer, Kan^R	This study
pET28a_*cm*IcmF_K344A	N-terminal His-tag, thrombin cleavage site, T7 protomer, Kan^R	This study

Cloning

The plasmid containing the wild type *Cupriavidus metallidurans icmF* gene as described previously[28] was used as the template for site directed mutagenesis. Site-directed variants were generated using a Quikchange II XL site-directed mutagenesis kit (Agilent) using the following primers:

Name	**Sequence (5′ to 3′)**
IcmF_Q341A_F	GCGCGGCCAGCGCGCTCGAGAAGATCGAC
IcmF_Q341A_R	GTCGATCTTCTCGAGCGCGCTGGCCGCGC
IcmF_K344A_F	GCGAGCCAGCTCGAGGCGATCGACATGCTCGACTTCG C
IcmF_ K344A_R	GCGAAGTCGAGCATGTCGATCGCCTCGAGCTGGCTGG C

Mutations generated according to the Quikchange protocol were confirmed by Sanger sequencing (Genewiz ®) using the following primers for full sequencing overlap:

Name	**Direction**	**Sequence (5′ to 3′)**
IcmF_seq_1	Forward	GCGCAACTGATTACCGCG
IcmF_seq_2	Forward	CAAGCAGGTGCAGCGCAA
IcmF_seq_3	Forward	CGTGTTCGCGTTCAAGCG
IcmF_seq_4	Forward	GAAGCCGGTGCGAATCCG
IcmF_seq_5	Reverse	CCACATCGCCAGCAGCTTG

Protein expression and purification

Cell growth and purifications of IcmF wild type (wt), IcmF Q341A and IcmF K344A from *Cupriavidus metallidurans* were conducted following the same procedure described here. An overnight culture of 100 mL of Luria Broth (LB) medium (Fisher BioReagents) supplemented with 50 µg/L kanamycin (GoldBio) was inoculated from a single colony of *E. coli* Bl21 T7 Express competent cells (New England Biolabs) transformed with the appropriate gene and grown at 37 ^oC with shaking. The overnight starter culture was used to inoculate 1 L of LB supplemented with 50 µg/L kanamycin at 37 ^oC. The 1 L culture

was placed at 16 °C with shaking when OD_{600} reached ~0.5-0.6. After 2 hours, the 1 L culture was induced with a final concentration of 0.1 mM IPTG (GoldBio) and grown for 10 h to 12 h at 16 °C with shaking. Cells were harvested by centrifugation (5,000 x g, 4 °C, 20 min) and flash frozen in liquid N_2 before being stored in a -80 °C freezer for future use.

Cells from 2 L of cell culture were resuspended in 80 mL of lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 20 mM imidazole) supplemented with 1 cOmplete EDTA-free protease inhibitor tablet (Roche), 1 mM of phenylmethylsulfonyl fluoride, 1 mM of benzamidine HCl, and benzonase nuclease. Cells were lysed by ultrasonification, and cell lysates were clarified by centrifugation (28,000 x g, 30 min, 4 °C). Clarified lysate was passed through a 0.2 µm filter before being loaded onto a 5 mL Ni-NTA column (GE Healthcare) equilibrated with lysis buffer using an FPLC (BioRad NGC System). The column was washed with 10 column volumes of lysis buffer and 10 column volumes of 50 mM HEPES pH 7.5, 500 mM NaCl, and 40 mM imidazole using an FPLC. Protein was eluted with 50 mM HEPES pH 7.5, 500 mM NaCl, 200 mM imidazole using an FPLC with a flow rate of 4 mL/min. Elution fractions were buffer exchanged into 50 mM HEPES pH 7.5 and then concentrated in a 50 kDa MWCO centrifugal filter. The concentrated fractions were loaded onto a MonoQ 10/100 anion exchange chromatography column (Cytivia) prepped with 50 mM HEPES pH 7.5 and 5 mM NaCl. Protein was eluted with a linear gradient from 10% to 80% of 50 mM HEPES pH 7.5 and 500 mM NaCl with a flow rate of 2 mL/min. The protein eluted in a sharp peak at around 250 mM to 300mM NaCl. There will be a second peak immediately following the first peak, corresponding to aggregated/truncated protein. The eluted protein was concentrated in a 50 kDa MWCO centrifugal filter. The concentrated fractions of full-length IcmF were loaded onto a Superdex 200 16/60 size exclusion chromatography (SEC) column (GE Healthcare) equilibrated with SEC buffer (20 mM HEPES pH 8, 50 mM NaCl) and eluted with a flow rate of 1 mL/min. Elution fractions from SEC were concentrated in a 50 kDa MWCO centrifugal filter. Purity was assessed by a 4 - 20% (w/v) sodium dodecyl sulfate polyacrylamide gel electrophoresis (BioRad). The concentration of the IcmF monomer was determined by UV/Vis absorbance at 280 nm using an extinction coefficient of 84,600 $M^{-1}cm^{-1}$, determined using the ProtParam tool[40]. Protein samples at a concentration of

~10 mg/mL (80 μM) in SEC buffer were flash frozen in liquid N_2 and stored in a -80 °C freezer for future use.

GTPase assays

The GTPase activity of IcmF wt, IcmF Q341A and IcmF K344A was determined in the presence of various concentrations of GTP (2 - 1500 μM) (Roche). The Enzcheck phosphate assay kit (Molecular Probes) was used for all GTPase assays following the manufacturer's instructions with the following modifications. The assay reactions (200 μL) were prepared excluding the enzyme (IcmF wt, IcmF Q341A, K344A, or no enzyme SEC buffer control) and incubated at room temperature (23 °C) for 5 minutes before initiating the assay to control for contaminating inorganic phosphate in the GTP. After incubation, the assays were initiated with the addition of 0.5 μM enzyme or 10 μL of IcmF SEC buffer. After 7 s of initial mixing, the absorbance at 360 nm for each assay reaction was recorded every 8 s with 2 s of mixing in between readings for 15 min total using a SpectraMax Plus 384 microplate reader (Molecular Dimensions). The absorbances were converted into concentration of inorganic phosphate using the standard curve generated according to the manufacturer's directions. The initial rates were calculated for each assay reaction and subtracted from the average initial rates of the no enzyme control of the corresponding concentration of GTP. The Michaelis-Menton parameters (k_{cat}, K_m, and V_{max}) were generated using a non-linear regression by Prism v9.4.1 (Graphpad). Reported values ±s.d. are the results of at least three independent experiments (*Table III.1*).

Mass photometry

Mass photometry (interferometric scattering mass spectrometry) was performed on a Refeyn instrument using AcquireMP v2.4.1 and DiscoverMP v2.4.2. All movies were taken for a length of 60 s using the default parameters. The contrasts were converted into molecular weights using the standard curve generated from a sample of NativeMark™ Unstained Protein Standard (Novex by life Technologies). Gaussian curves were fit to each histogram distribution and the mass (kDa), sigma (kDa) and normalized counts were determined using the PhotoMol software[41]. Each percentage of total counts ± s.d.

reported are the results of at least three independent experiments. All samples that contained any additives also contained 500 μM $MgCl_2$. Samples incubated with 500 μM GDP or no nucleotide for 15 minutes on ice at a concentration of 80 ng/μL prior to the final dilution to 8 ng/μL on the instrument and data recording. Samples incubated with 500 μM β,γ-methyleneguanosine 5′-triphosphate (GMPPCP) for 15 minutes on ice at a concentration of 160 ng/μL prior to the final dilution to 16 ng/μL on the instrument and data recording. Samples containing 500 μM GTP were initially diluted to 160 ng/μL before the final dilution to 16 ng/μL in SEC buffer (20 mM HEPES pH 8, 50 mM NaCl) supplemented with 500 μM GTP on the instrument and data recording. Samples containing the substrate butyryl-CoA (500 μM) and/or various nucleotides (GDP or GMPPCP) were incubated for 15 minutes on ice at a concentration of 200 ng/μL prior to the final dilution to 20 ng/μL on the instrument and data recording. Reported values ± s.d. are the results of at least three independent experiments.

Negative stain electron microscopy specimen preparation and imaging

IcmF wt, IcmF Q341A, or IcmF K344A was thawed on ice and diluted to 20 ng/μL in IcmF SEC buffer (20 mM HEPES pH 8, 50 mM NaCl). Each of the samples except samples with GTP were incubated for 30 minutes on ice with the corresponding nucleotide (GDP or GMPPCP) before grid preparation. The final concentration of any additives (nucleotides: GDP, GTP or GMPPCP; cofactor: AdoCbl or OHCbl; $MgCl_2$) in the samples were 500 μM. For the samples containing GTP, the GTP was added, and the protein solution immediately applied to the grid. For the samples containing AdoCbl and exposed to light, after incubation for 30 minutes in the dark, the protein solution was then exposed to a white light for 15 minutes before application on the grid.

Carbon-coated 300 mesh copper electron microscopy grids (Electron Microscopy Services) were glow discharged for 1 min at -15 mA. 5 μL of the protein solution was applied to the grid, after approximately 1 min the solution was blotted and immediately replaced with solution of 2% uranyl acetate (VWR). The stain solution was blotted and replaced twice, then allowed to stand for 1 min before the final blot, and then was dried. All blotting was done manually using filter paper (Whatman, grade 40). The specimens

were imaged with an AMT Nanosprint5 camera on a FEI Morgagni electron microscope operated at 80 kV. Images were collected at 18,000 x magnification.

Cryogenic EM grid preparation

The grid used to investigate the supramolecular complex of IcmF in the presence of GMPPCP (IcmF wt + GMPPCP) were prepared as follows: 0.2 mg/mL graphene oxide suspension was prepared using molecular grade water. The suspension was centrifuged at 300 x g for 1 min to remove large aggregates. A Quantifoil 1.2-1.3 Cu 300 mesh holey-carbon grid (Electron Microscopy Services) was glow discharged at -40 mA at 0.1 bar for 2 min before application of 3 µL of the graphene oxide suspension. The suspension was incubated for 1 min before excess suspension was blotted away (Whatman, grade 40). The grid was washed twice on the graphene oxide suspension side and once on the backside of the grid in molecular grade water and dried. The graphene oxide covered grid was plunged on a Thermo Fisher Scientific Vitrobot (Mk IV) cryo-plunger. The final protein solution contained 2 µM IcmF wt, 500 µM GMPPCP and 500 µM butyryl-CoA incubated in SEC buffer for at least 15 minutes before application. 5 µL sample was applied to the grids that were blotted for 5 s with a blot force of 10 (Whatman filter paper #1) before plunging into liquid ethane and transferring to storage grids. The temperature and humidity inside the Vitrobot chamber were set to 8 °C and 95%, respectively.

The grid used to investigate the conformation of IcmF Q341A in the presence of GTP (IcmF Q341A + GTP) were prepared as followed. A Quantifoil R 1.2-1.3 Cu 300 mesh holey-carbon grid (Electron Microscopy Services) was glow discharged at -15 mA at 0.039 bar for 1 min before application of the protein solution. The final protein solution contained 2 µM IcmF Q341A, 500 µM GTP in SEC buffer. 5 µL sample was applied to the grids that were blotted for 3 s with a blot force of 10 (Whatman filter paper #1) before plunging into liquid ethane and transferring to storage grids. The temperature and humidity inside the Vitrobot chamber were set to 8 °C and 95%, respectively.

Cryogenic EM data collection

Data were collected at the MIT.nano Center for Automated Cryogenic Electron Microscopy at the Massachusetts Institute of Technology on a FEI Talos Arctica G2 Cryo

200 kV transmission electron microscope equipped with a Falcon 3EC camera. The data collection parameters for the IcmF + GMPPCP grid were as follows: 73000 x magnification resulting a pixel size of 2.143 Å, 14 frames, 10.5 e^- /Å^2 /frame dose, and defocus range 1.2 - 3.1 µm. The data set contained 602 movies. The data collection parameters for the IcmF Q341A + GTP grid were as follows: 92000 x magnification resulting a pixel size of 1.5998 Å, 14 frames, 14.7 e^- /Å^2 /frame dose, and defocus range 1.2-3.1 µm. The data set contained 673 movies. These parameters are summarized in *Table III.2*.

Cryogenic EM data processing and model refinement

Cryo-EM data processing of the datasets were carried out using a combination of Relion 4.0-beta[42] and CryoSPARC v3.3.2[35] and is summarized in *Figure III.11*. For the IcmF wt + GMPPCP dataset, individual frames of dose-fractionated exposure were aligned and summed using Relion's implementation of MotionCor2[43] and the defocus of the summed frames was estimated using Relion's implementation of CTTFind4[44]. The start-end coordinates for the locations of the helical supramolecular complexes were manually determined for all micrographs. Using these coordinates, 510941 particles were extracted using a box size of 128 pixels (257.8 Å), tube diameter of 200 Å, one asymmetric unit and a helical rise of 10 Å (*Figure III.11*). These particles were imported in CryoSPARC for the rest of the data processing. The particles were subjected to one round of initial reference-free 2D classification with a mask of 240 Å to generate 100 class averages. After removing the classes that visually did not look like particles, another round of reference-free 2D classification was performed with the remaining 248494 particles to generate 100 2D class averages. The 159463 particles selected after the second round of 2D classification were used to generate one *ab initio* initial reference-free model using no imposed symmetry. The initial model was subjected to homogenous refinement. After homogenous refinement, the aligned particles were subjected to signal subtraction to remove signal from the ends of the supramolecular complexes that were artifacts of processing helical data as single particles using a manually generated mask. These subtracted particles were then subjected to a local refinement using CryoSPARC's new implementation. CryoSPARC's dynamic masking was utilized for all refining steps,

excluding the signal subtraction. Combination of the two half-maps along with local B-factor adjustment was performed on the COSMIC[2] server's implementation[45] of LocSpiral[36] with a low pass filter of 15 Å, bandwidth of band pass filter of 8 Å, and an initial binarization threshold of 0.58. The FSC plots were generated by CryoSPARC. The final masked resolution at FSC = 0.143 was 5.0 Å (*Figure III.18*).

For the IcmF Q341A + GTP dataset, all data processing was performed in CryoSPARC v3.3.2 [35] and is summarized in *Figure III.13*. First, individual frames of dose-fractionated exposure were aligned and summed using patch motion correction. Next, the defocus of the summed frames was estimated using CryoSPARC's patch CTF estimation. To generate the coordinates of the particles, the blob picker was used with minimum particle diameter of 200 Å and maximum particle diameter of 350 Å. 180271 particles with an NCC score above 0.160 and local power between -8352 and 242864 were extracted a box size of 196 pixels. These particles were subjected to one round of initial reference-free 2D classification to generate 200 2D class averages. After removing classes that visually did not look like particles, another round of reference-free 2D classification was performed with the remaining 74767 particles to generate 100 2D class averages. The 70582 particles selected after the second round of 2D classification were used to generate three *ab initio* initial reference-free models using no imposed symmetry. For the following heterogenous refinement, two of the *ab initio* models (one containing intact complex and one containing junk particles) were supplied. Only the 59335 particles associated with the model containing intact complex were subject to a homogenous refinement. These particles were then subjected to a local refinement using CryoSPARC. CryoSPARC's dynamic masking was utilized for all refining steps. Combination of the two half-maps along with local B-factor adjustment was performed on the COSMIC[2] server's implementation[45] of LocSpiral[36] with a low pass filter of 15 Å, bandwidth of band pass filter of 8 Å, and an initial binarization threshold of 0.878. The FSC plots were generated by CryoSPARC. The final masked resolution at FSC = 0.143 was 4.6 Å (*Figure III.18*).

For model building and refinement, one protomer of the dimeric IcmF (chain B from PDB 4XC6, residues 22 - 1093) was segmented into two fragments that were manually docked into the maps resulting from the reconstruction of the IcmF wt + GMPPCP and IcmF Q341A + GTP datasets, respectively[28]. All the ligands and water molecules were

removed from the fragments. Despite the low to modest resolution, side chains were retained in the fragments. The first fragment consisted of the Cbl-binding domain and the G-protein domain (residues 22 - 442); the second fragment consisted of substrate-binding domain and the linker region (residues 443 - 1093). For the IcmF wt + GMPPCP reconstruction, four complete protomers (eight fragments) of IcmF were manually docked in the map in ChimeraX v1.4[46]. Using Phenix real-space refinement, the resulting model was subjected to one round of refinement consisting of rigid body refinement, with each fragment refined as an individual rigid body. There was no clear density for substrate or GMPPCP and, therefore, these ligands were not modeled in. For the IcmF Q341A + GTP reconstruction, three copies of a complete protomer of IcmF (six fragments) were manually docked into the map in ChimeraX v1.4. Although the map appeared to contain two IcmF homodimers, the density was only good enough to place three protomers (six fragments) and no atoms were modeled into a fourth protomer. Within the modelled protomers, however, there was density present for residues 1011-1018, which were previously disordered in the crystal structures, and those residues were added manually by model building in Coot[47]. Additionally, there was clear density for GDP at each nucleotide binding site. In the nucleotide binding site for chain A, there was enough density to model in a magnesium ion in addition to the GDP. Using Phenix real-space refinement[48], the docked models were subjected to one round of refinement consisting first of simulated annealing and then rigid body refinement, with each copy of each fragment of IcmF defined as an individual rigid body. The second round of real-space refinement was carried out on the model with rigid body refinement and minimization. Molecular building and refining software packages were compiled by SBGrid[49].

Figures

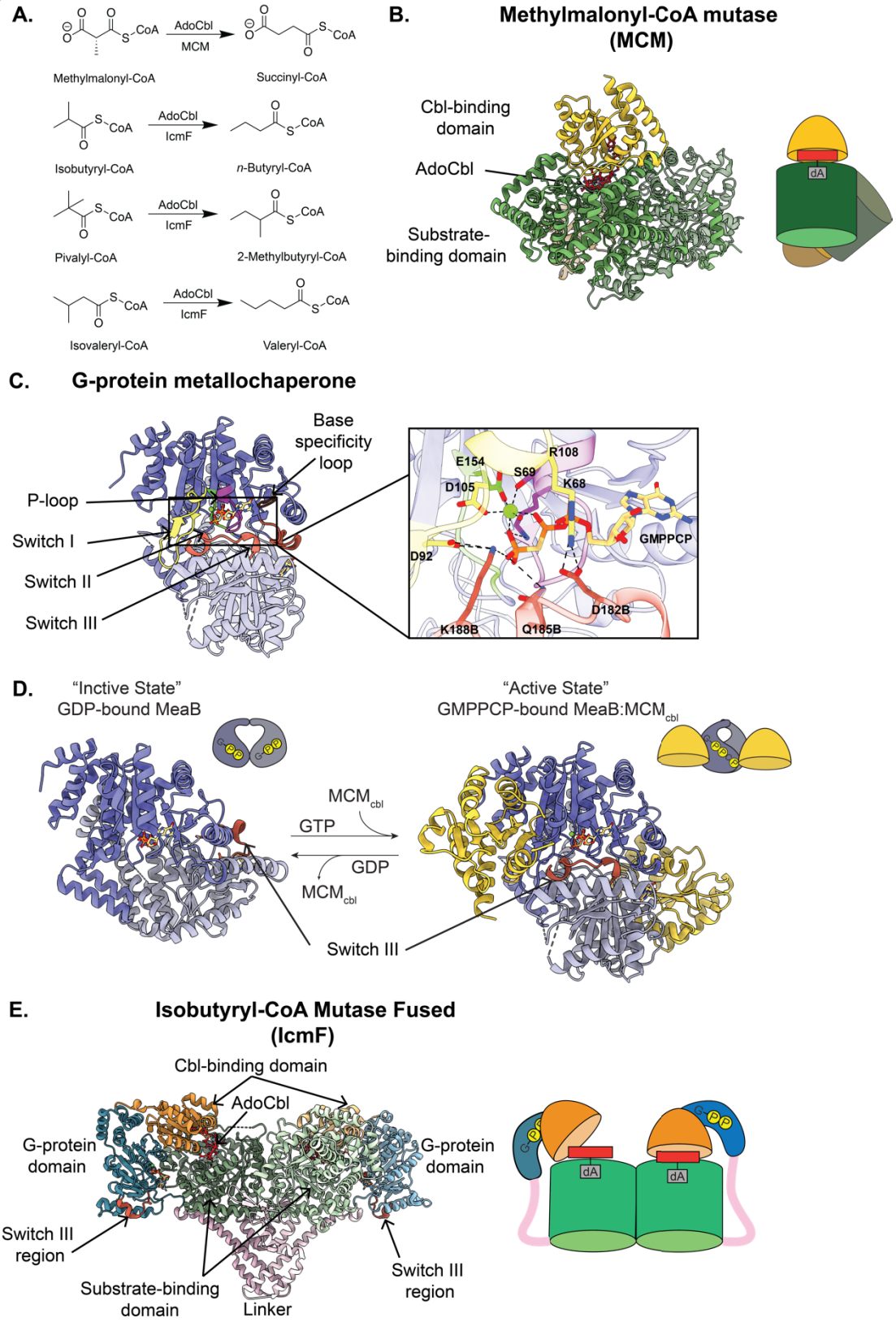

A.
AdoCbl
MCM
Methylmalonyl-CoA
Succinyl-CoA
AdoCbl
IcmF
Isobutyryl-CoA
n-Butyryl-CoA
AdoCbl
IcmF
Pivalyl-CoA
2-Methylbutyryl-CoA
AdoCbl
IcmF
Isovaleryl-CoA
Valeryl-CoA
B.
Methylmalonyl-CoA mutase (MCM)
Cbl-binding domain
AdoCbl
Substrate-binding domain
dA
C.
G-protein metallochaperone
Base specificity loop
P-loop
Switch I
Switch II
Switch III
E154
R108
S69
D105
K68
GMPPCP
D92
D182B
K188B
Q185B
D.
"Inctive State"
GDP-bound MeaB
"Active State"
GMPPCP-bound MeaB:MCM$_{cbl}$
MCM$_{cbl}$
GTP
GDP
MCM$_{cbl}$
Switch III
E.
Isobutyryl-CoA Mutase Fused (IcmF)
Cbl-binding domain
AdoCbl
G-protein domain
G-protein domain
Switch III region
Substrate-binding domain
Linker
Switch III region
dA
dA

Figure III.1. Adenosylcobalamin-dependent mutase reactions and structures. **A.** AdoCbl-dependent mutase reactions**. B.** Methylmalonyl-CoA mutase (PDB 4REQ)[25] consists of a Cbl-binding domain (yellow ribbons) and substrate binding domain (green ribbons) with the AdoCbl cofactor (red sticks) binding at the interface. The inactive subunit consists of a substrate binding domain (light green ribbons) and Cbl-binding domain (tan ribbons) which cannot perform catalysis. Cartoon representation of MCM is colored according to the ribbon drawings. **C**. The G-protein chaperone (PDB 8DPB) (Chapter II) [32] (purple ribbons) has conserved motifs indicative of its class of P-loop GTPases as labeled and an additional switch III motif (red orange) implicated in GTP hydrolysis which is necessary for its maturation role. **Inset**: The nucleotide binding site with GMPPCP and Mg^{2+} ion bound interacting with residues from the phosphate binding loop (purple sticks: K68 and S69), switch I region (yellow sticks: D92, D105, D108), switch II region (green sticks: E154) and the switch III region from the other protomer (red-orange sticks: K188, Q185, D182). Residues labeled with "B" are from other protomer. **D.** MeaB conformational change. Inactive state: PDB 2QM7[31]; Active state: PDB 8DPB (Chapter II)[32]. Switch III region of the light purple MeaB protomer (chain B) rearranges from being solvent-exposed in "inactive state" to contacting GMPPCP in the "active state". **E.** *C. metallidurans* IcmF (PDB 4XC6)[28] contains two protomers each consisting of a G-protein domain (blue) and mutase domain (substrate-binding domain (green) and Cbl-binding domain (orange)) connected by a linker (pink). AdoCbl (red sticks) is bound at the active site which is located at the interface of the Cbl-binding domain and substrate-binding domain. The switch III residues are highlighted in red-orange.

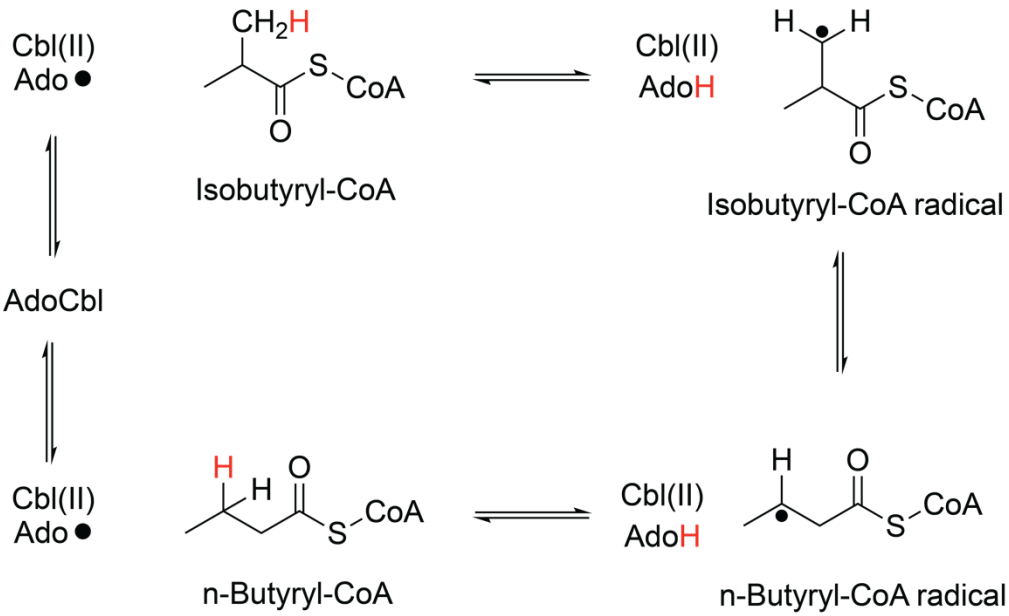


Figure III.2. Scheme showing proposed catalytic mechanism for IcmF. The rearrangement of isobutyryl-CoA to *n*-butyryl-CoA requires the formation of a 5′-deoxyadenosyl (Ado•) radical from the homolytic cleavage of AdoCbl to perform a hydrogen atom abstraction from substrate. The product n-butryryl-CoA radical re-abstracts the hydrogen atom from 5′-deoxyadenosine (AdoH), regenerating Ado•. Cob(II)alamin (Cbl(II)) and Ado• reform AdoCbl for the next round of turnover[50].

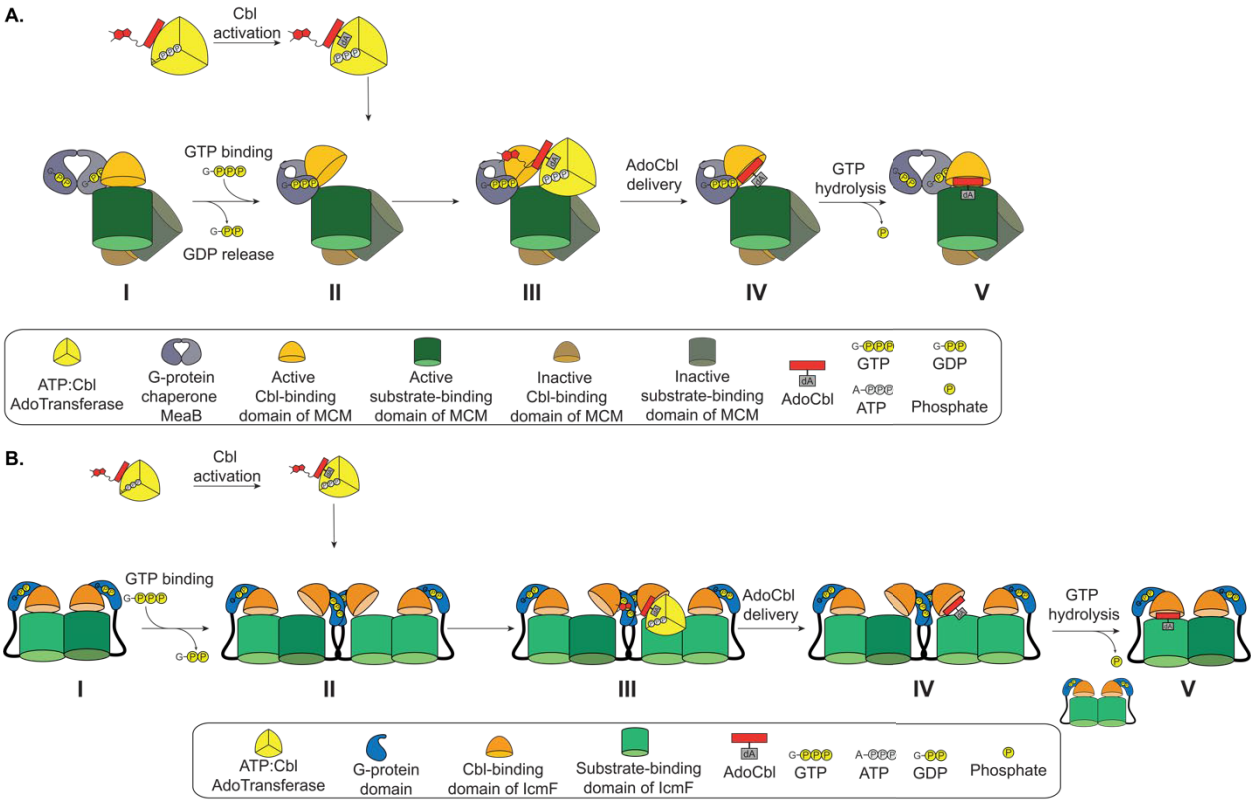


***Figure III.3. Proposed steps for loading AdoCbl into the mutase active site*. A.** GTP binding results to MeaB leads to the formation of the MeaB "active state". This conformational change of MeaB opens up one subunit of the MCM heterodimer for AdoCbl delivery from ATR. GTP hydrolysis returns MeaB to the "inactive state", closing the MCM subunit and capturing AdoCbl. **B.** GTP binding to the G-protein domain of IcmF leads to the formation of a higher order oligomeric state of IcmF, generating a dimer interface analogous to the "active state" of MeaB. Oligomerization of IcmF opens up one promoter of the IcmF homoodimer for AdoCbl delivery from ATR. GTP hydrolysis breaks apart the higher order oligomer of IcmF, closing the IcmF protomer, trapping the AdoCbl.

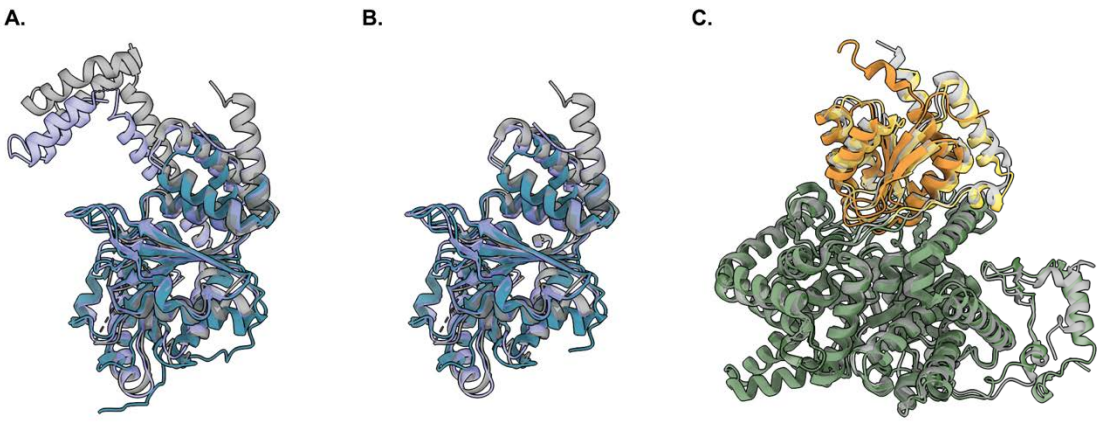


Figure III.4. The domains of IcmF are similar to the corresponding two component analogs. **A**. The G-protein domain of IcmF (PDB 4XC6)[28] (blue ribbons) overlayed with one protomer of MeaB (RMSD for C∝: 0.956 Å) (PDB 8DPB) (Chapter II)[32] (purple ribbons) and one protomer of MMAA (RMSD for C∝: 1.065 Å) (PDB 2WWW)[5] (grey ribbons). **B**. The same overlay in (A) with the C-terminal dimerization helices of MeaB and MMAA removed to highlight the similarity of the core G-domain. **C.** The Cbl-binding domain (orange ribbons) and substrate binding domain (green ribbons) of IcmF (PDB 4XC6)[28] overlayed with the active subunit (Cbl-binding domain: yellow; substrate-binding domain: dark green) of methylmalonyl-CoA mutase from *P. shermanii* (RMSD for C∝: 1.209 Å) (PDB 4REQ)[25] and one protomer of methylmalonyl-CoA mutase from *H. sapiens* (RMSD for C∝: 0.948 Å) (PDB 2XIQ)[5] (grey ribbons).

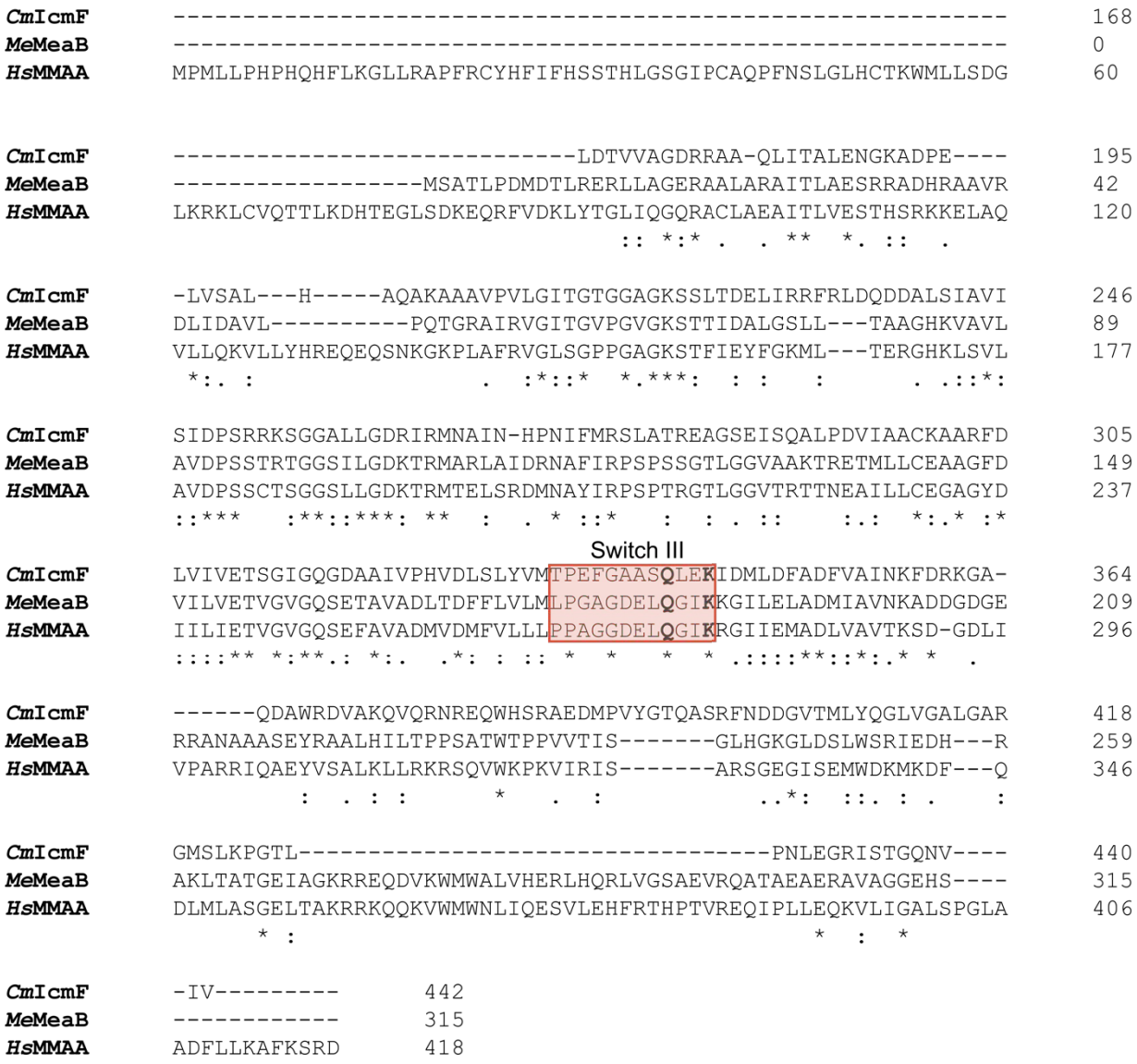


Figure III.5. Sequence alignment of the G-protein domain of IcmF from C. metallidurans with MeMeaB and HsMMAA. The switch III region which participates in binding and catalysis of the nucleotide substrate is highlighted. The conserved residues of the switch III region that were substituted to alanine are bolded. Asterisks (*) denote positions with conserved residues. Colons (:) denote positions with conservation with strongly similar properties. Periods (.) denote positions with conservation with weakly similar properties. Alignments performed using Clustal Omega[51].

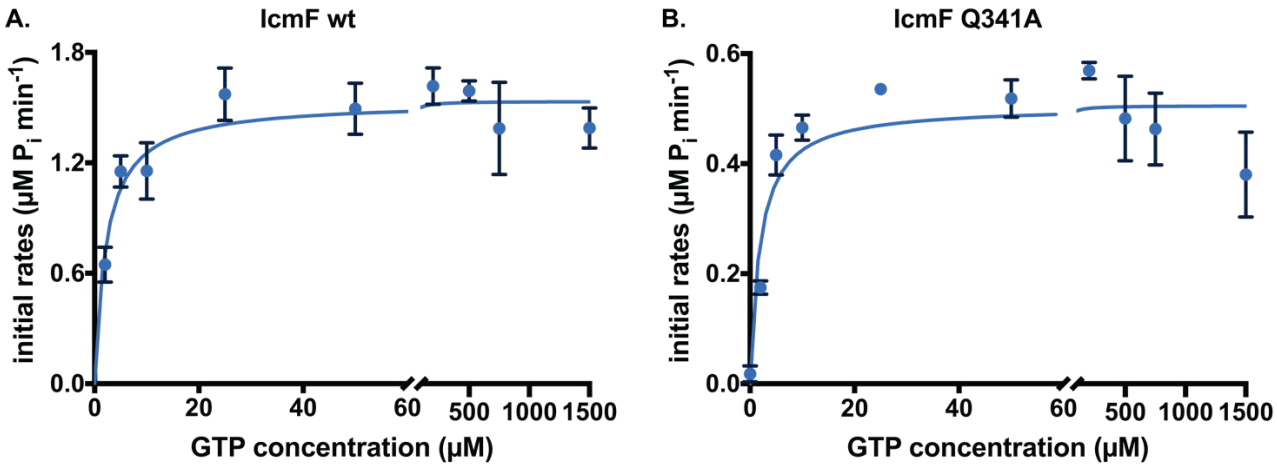


Figure III.6. The Michaelis-Menton kinetic analysis of IcmF wt and switch III variant IcmF Q341A. **A.** The GTP hydrolysis activity of IcmF wt **B.** The GTP hydrolysis activity of IcmF Q341A variant is lower than wild type. Each reaction was performed with 0.5 μM of the enzyme. The data for each curve represent the average ± S.D of at least three replicates.

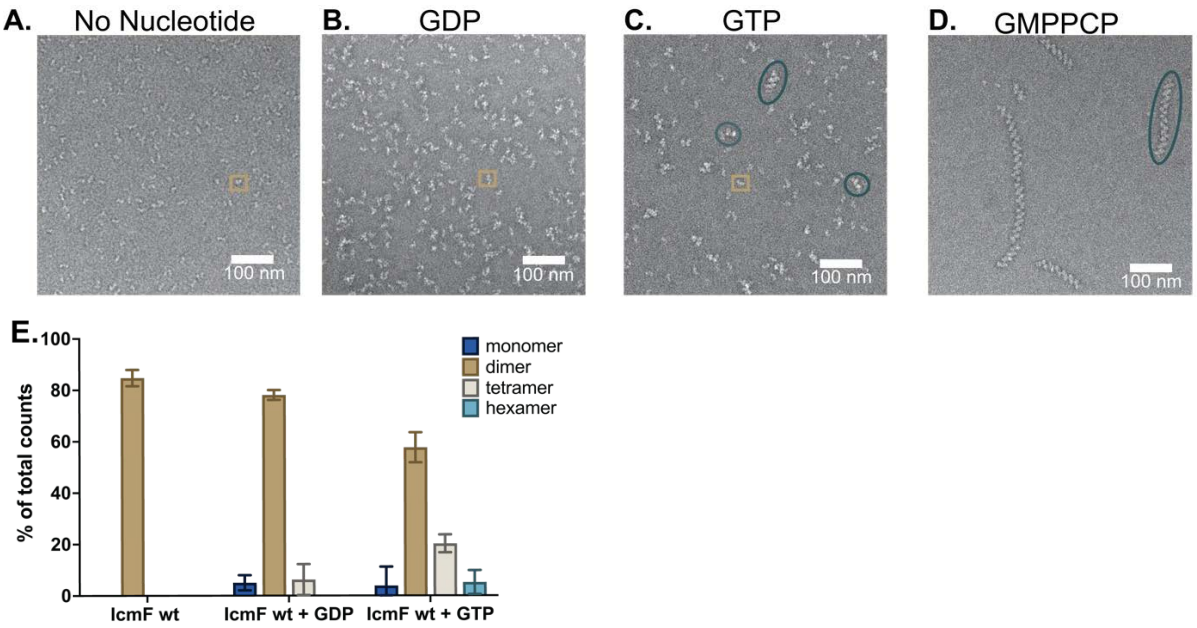


Figure III.7. Solution state analysis of the oligomeric state of IcmF wt in the absence and presence of G-nucleotides. **A**. Negative stain EM image of 20 ng/µL IcmF wt without any nucleotide. Representative IcmF dimer is boxed in tan. **B**. 20 ng/µL IcmF wt in the presence of 500 µM of GDP. Representative IcmF dimer is boxed in tan. **C**. 20 ng/µL IcmF wt in the presence of 500 µM GTP. Representative IcmF dimers are boxed in tan and supramolecular complexes are circled in teal. **D**. 20 ng/µL IcmF wt in the presence of 500 µM GMPPCP. Representative IcmF dimers are boxed in tan and supramolecular complexes are circled in teal. **E.** Mass photometry analysis of IcmF wt without any nucleotide or in the presence of 500 µM GDP or 500 µM GTP indicate the presence of more higher order oligomers in the presence of GTP. The data for each condition represents the mean ± S.D of at least four replicates.

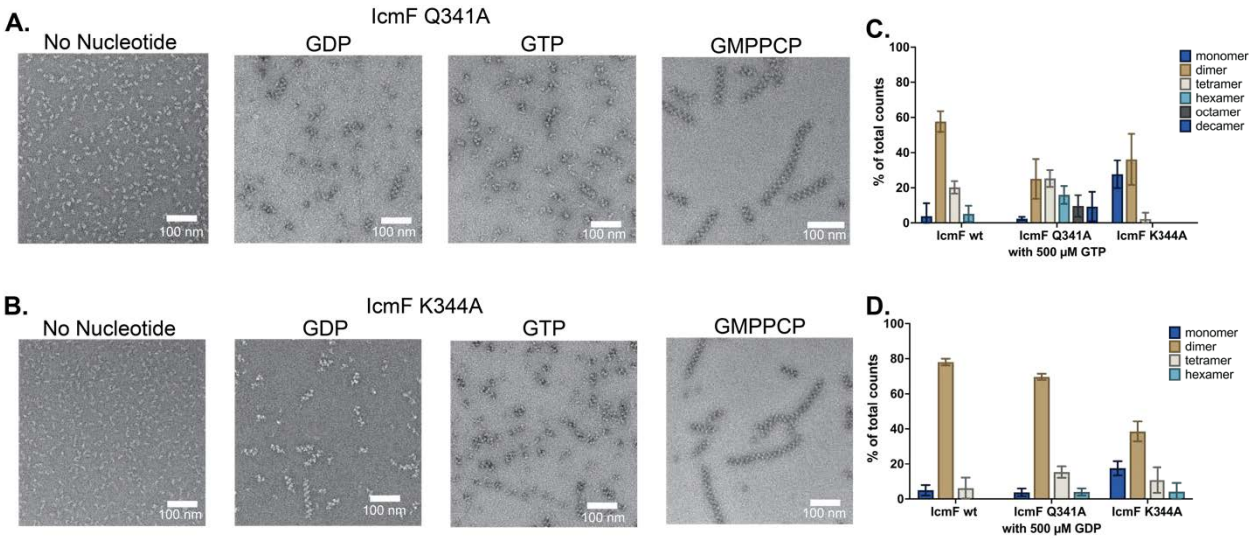


***Figure III.8. Solution state analysis of the oligomeric state of IcmF variants in the absence and presence of G-nucleotides*. A.** Representative negative stain images of 20 ng/µL IcmF Q341A without any nucleotide or in the presence of 500 µM of GDP, GTP, or GMPPCP. Without any nucleotide, only IcmF dimers are observed. In the presence of GDP, GTP, or a non-hydrolyzable GTP analog, supramolecular complexes are observed. **B.** Representative negative stain images of 20 ng/µL IcmF K344A without any nucleotide or in the presence of 500 µM of GDP, GTP, or GMPPCP. Without any nucleotide, only IcmF dimers are observed. In the presence of GDP, GTP, or a non-hydrolyzable GTP analog, supramolecular complexes are observed. **C.** Mass photometry analysis of IcmF wt, IcmF Q341A, or IcmF K344A in the presence of GTP indicate the presence of supramolecular complexes. **D.** Mass photometry analysis of IcmF wt, IcmF Q341A, or IcmF K344A in the presence of 500 µM GDP indicate the presence of supramolecular complexes only in the switch III variants.

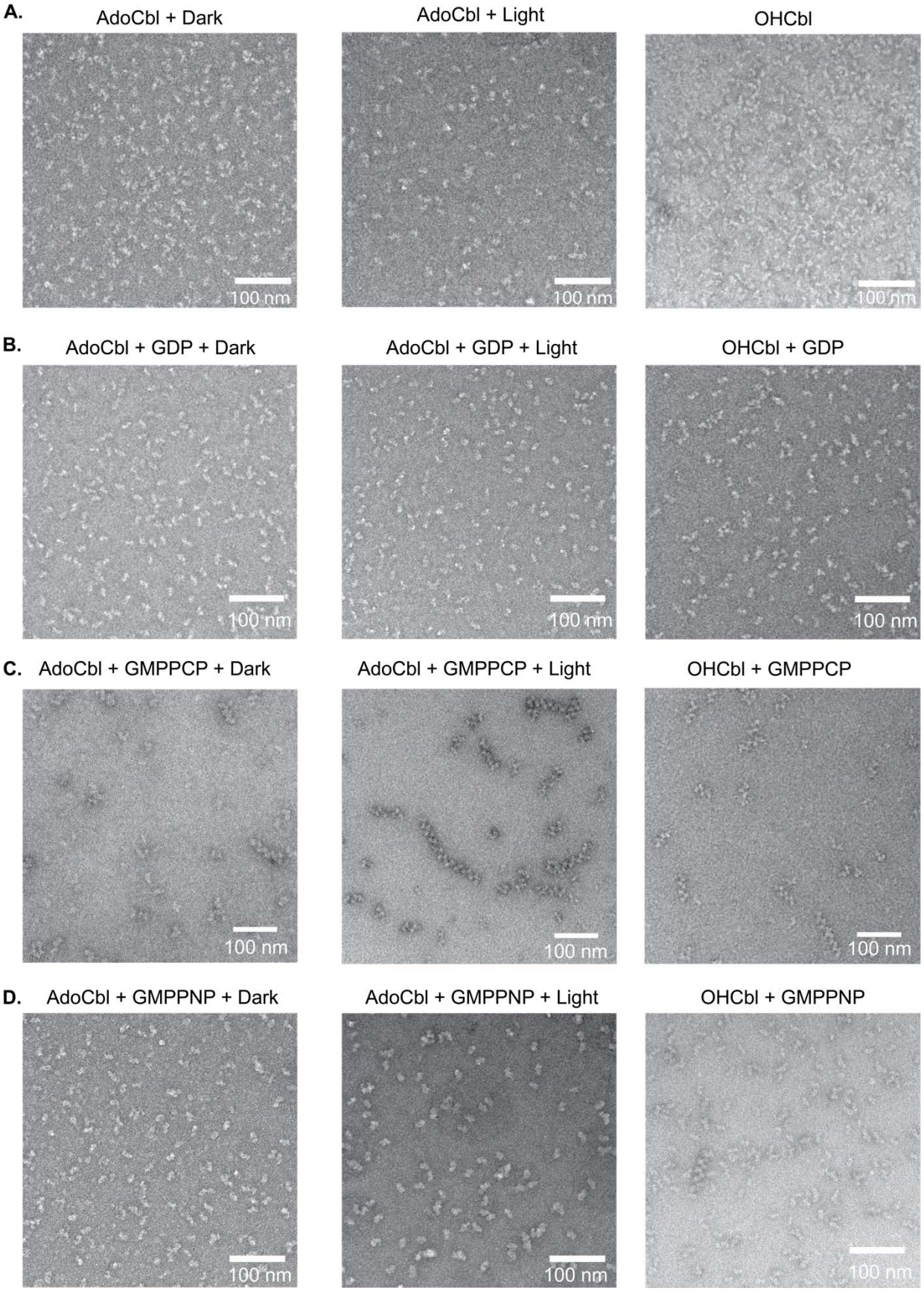

A.
AdoCbl + Dark
AdoCbl + Light
OHCbl
B.
AdoCbl + GDP + Dark
AdoCbl + GDP + Light
OHCbl + GDP
C.
AdoCbl + GMPPCP + Dark
AdoCbl + GMPPCP + Light
OHCbl + GMPPCP
D.
AdoCbl + GMPPNP + Dark
AdoCbl + GMPPNP + Light
OHCbl + GMPPNP
100 nm

Figure III.9. Negative stain electron microscopy analysis of the oligomeric state of IcmF when bound to various states of the cobalamin cofactor in the presence and absence of G-nucleotides. **A.** Representative negative stain images of 20 ng/μL IcmF wt in the presence of AdoCbl incubated in the dark (left), AdoCbl exposed to light (middle) and OHCbl (right). None of the conditions form supramolecular complexes. **B.** Representative negative stain images of 20 ng/μL IcmF wt in the presence of AdoCbl and GDP incubated in the dark (left), AdoCbl and GDP exposed to light (middle) and OHCbl and GDP (right). None of the conditions form supramolecular complexes. **C.** Representative negative stain images of 20 ng/μL IcmF wt in the presence of AdoCbl and GMPPCP incubated in the dark (left), AdoCbl and GMPPCP exposed to light (middle) and OHCbl and GMPPCP (right). All the states that have inactivated cofactor contain supramolecular complexes; however, the states with inactivated cofactor (AdoCbl exposed to light or OHCbl) have more supramolecular complexes. The concentration of all cofactors and nucleotides added was 500 μM. **D.** Representative negative stain images of 20 ng/μL IcmF wt in the presence of AdoCbl and GMPPNP incubated in the dark (left), AdoCbl and GMPPNP exposed to light (middle) and OHCbl and GMPPNP (right). Only the states that have inactivated cofactor (AdoCbl exposed to light or OHCbl) and a non-hydrolyzable analog contain supramolecular complexes. The concentration of all cofactors and nucleotides added was 500 μM.

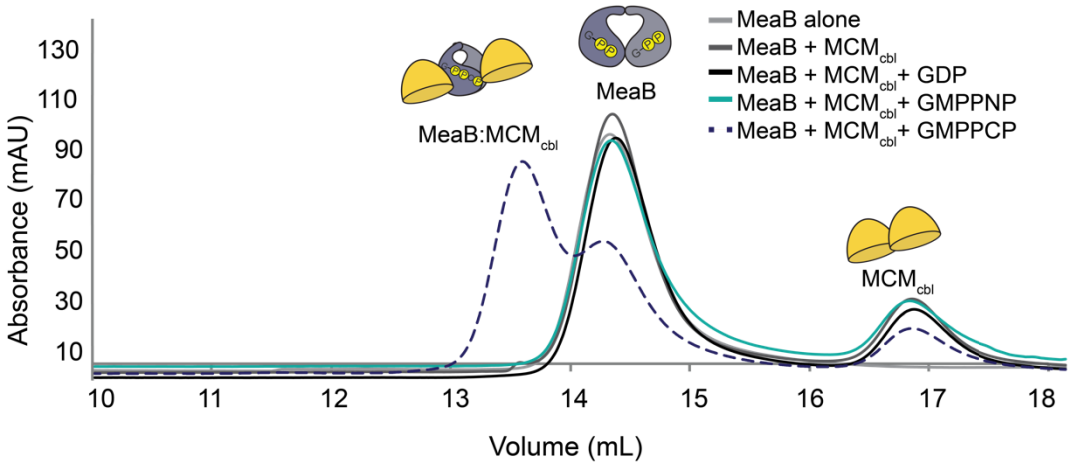


Figure III.10. MeaB:MCM$_{Cbl}$ complex formation is only observed in the presence of GMPPCP. The S200 10/300 size exclusion chromatograms indicate the formation of a stable complex between MeaB and the Cbl-binding domain of MCM (MCM$_{cbl}$) only occurs in the presence of GMPPCP (dashed blue). The traces of MeaB alone (light gray), MeaB and MCM$_{cbl}$ with no G-nucleotide (dark gray), MeaB with MCM$_{cbl}$ and GDP (black) and MeaB and MCM$_{cbl}$ with GMPPNP (cyan) indicate no stable complex formation. Figure reproduced from Chapter II.

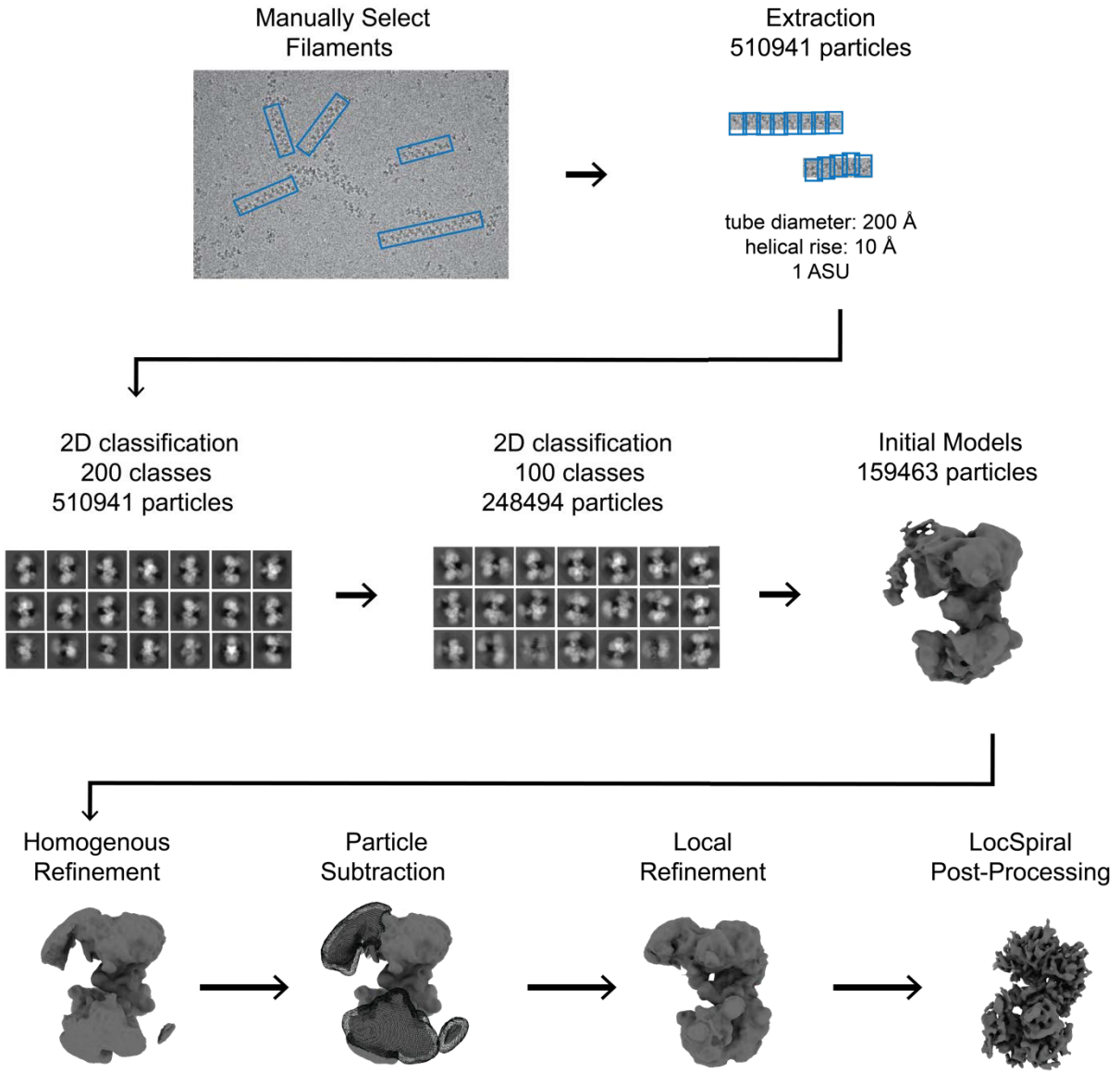


Figure III.11. Data processing workflow for the dataset of IcmF wt in the presence of GMPPCP and butyryl-CoA. Drift assessment and initial CTF assessment was performed, filaments were manually traced, and particles were extracted in Relion 4.0. All subsequent analysis was except for post-processing B-factor sharpening, which was performed using COSMIC^2's implementation of LocSpiral, was performed using CryoSPARC.

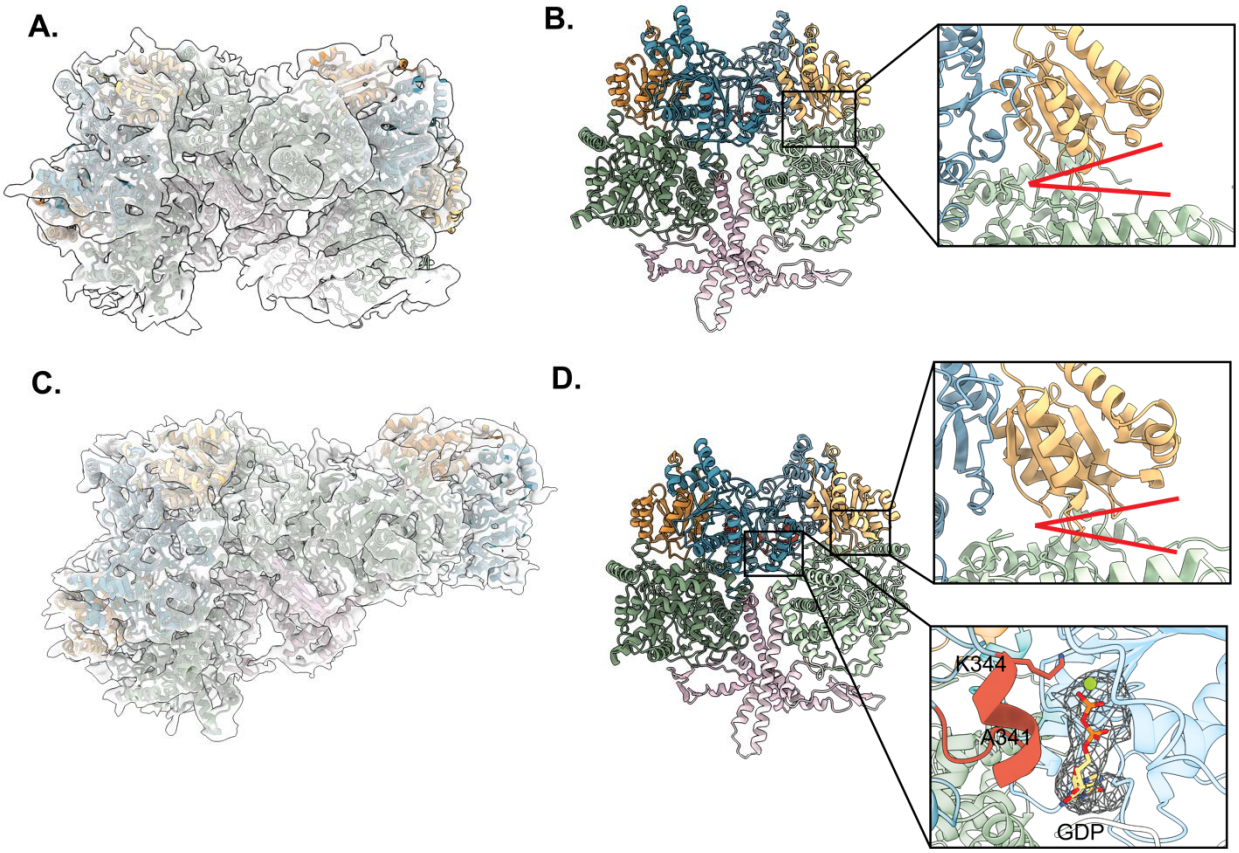


Figure III.12. The 3D cryogenic EM helical reconstruction of the supramolecular complexes of IcmF indicate the formation of a G-protein interdimer interface in solution. **A.** The 5.0 Å resolution single particle reconstruction of the supramolecular complexes of IcmF wt in the presence of 500 μM GMPPCP and 500 μM butyryl-CoA (IcmF wt + GMPPCP) with four IcmF protomers modeled into the EM map. **B**. Two of the four protomers from IcmF wt + GMPPCP structure are shown. Both protomers are from a different IcmF homodimer (the homodimer is not shown for simplicity). The majority of the intermolecular contacts are made by the G-protein domains. Coloring: G-protein domain in teal, the substrate-binding domain in green, and the linker region in pink. **Inset**: The mutase active site is open indicated by the red lines. **C.** The 4.6 Å resolution single particle reconstruction of the complexes of IcmF Q341A in the presence of 500 μM GTP has three protomers of IcmF modeled into the map. **D.** Two of the three protomers from IcmF Q341A + GTP structure are shown. Both protomers are from a different IcmF homodimer (the homodimer is not shown for simplicity). As in B, the majority of the intermolecular contacts are made by the G-protein domains. Colored as described in B.

Top Inset: The mutase active site is open indicated by the red lines. **Bottom Inset**: Switch III residues (red) from one protomer contact the GDP bound to the second protomer across the G-protein domain: G-protein domain interface. GDP and a Mg^{2+} ion are shown against the EM map.

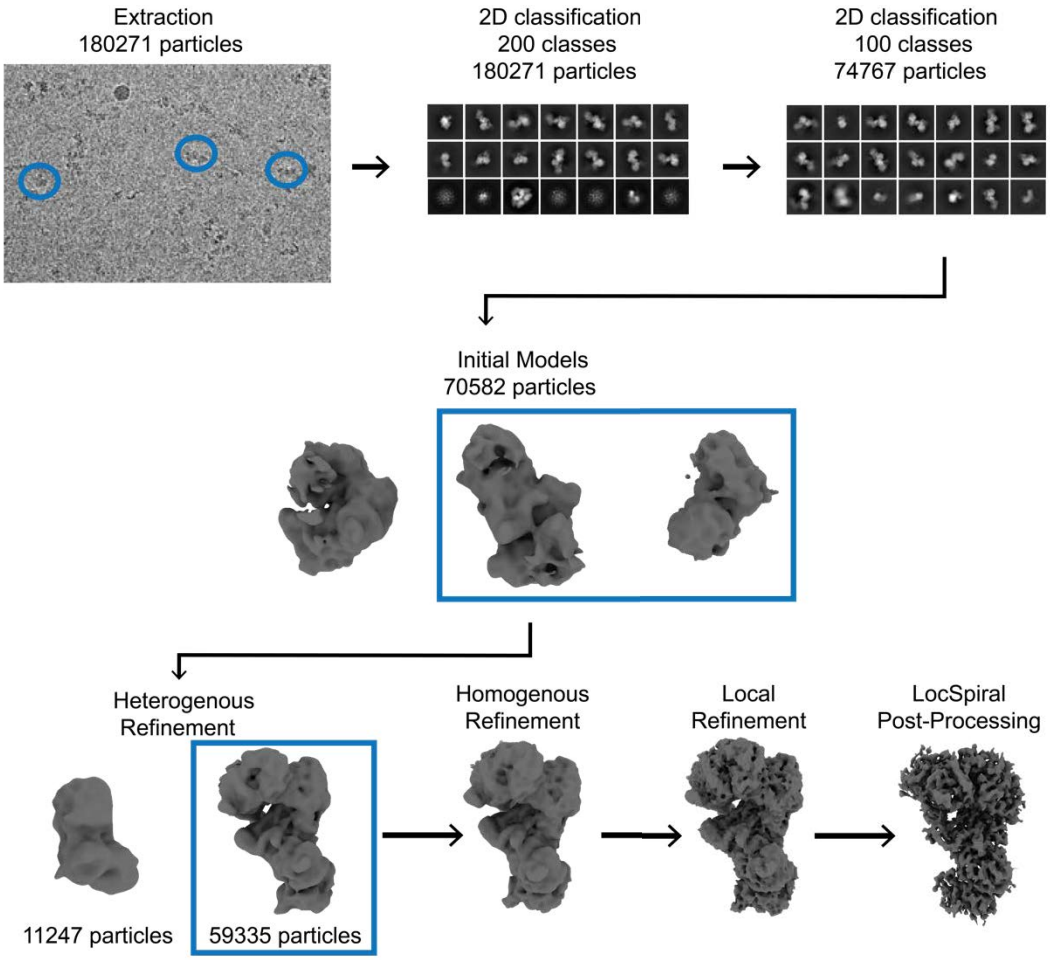


Figure III.13. The processing workflow for the dataset of the IcmF Q341 variant in the presence of GTP. All analysis was except for post-processing B-factor sharpening, which was performed using COSMIC[2]'s implementation of LocSpiral, was performed using CryoSPARC.

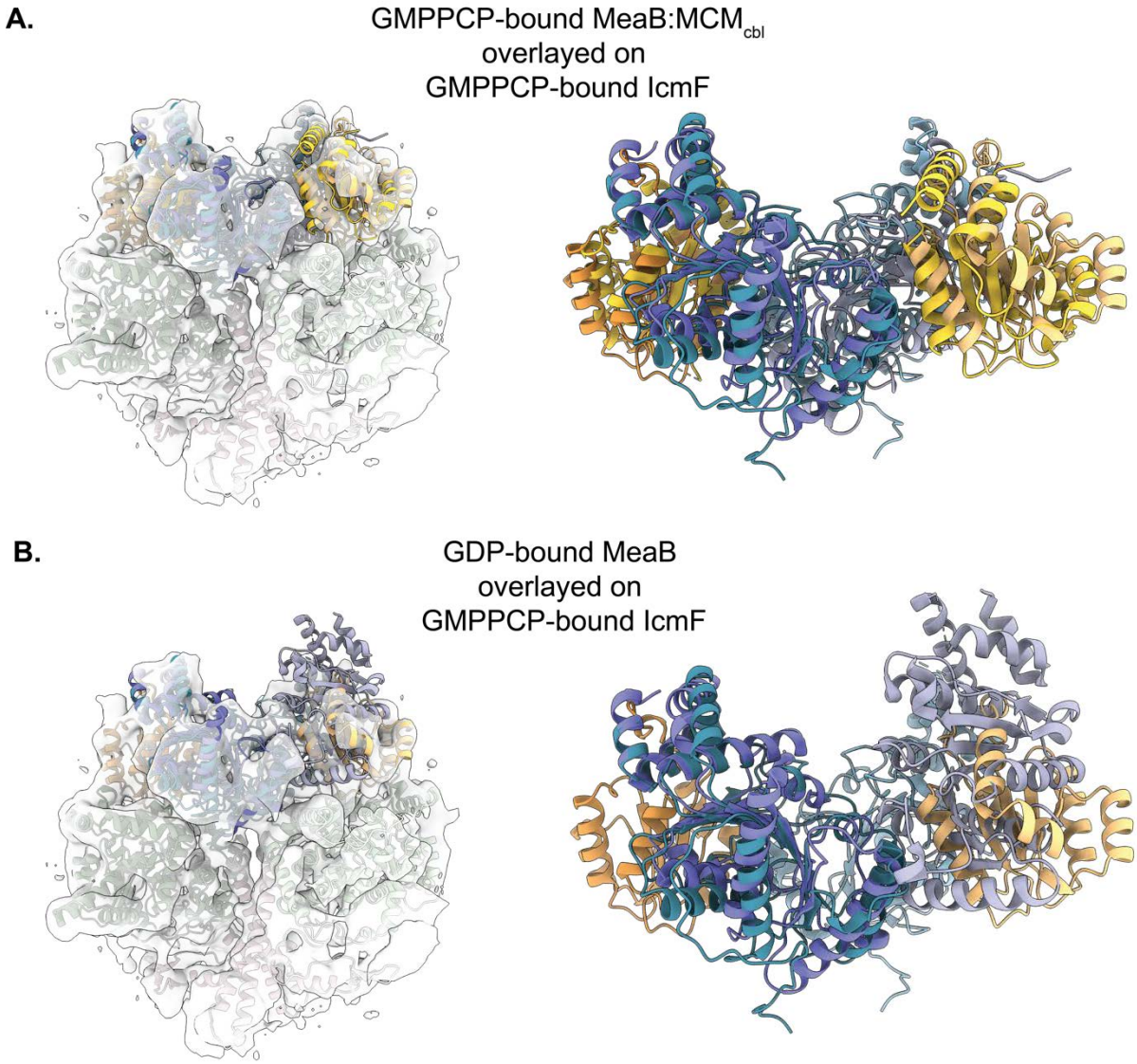


Figure III.14. The active conformation of MeaB has good agreement with IcmF cryo-EM data. **A.** Superimposition of MeaB:MCM$_{Cbl}$ complex (MeaB in purple/light purple; MCM$_{Cbl}$ in yellow/orange) (PDB 8DPB) (Chapter II)[32] aligned to IcmF wt + GMPPCP structure (colored as in *Figure III.13B*) with cryo-EM map (left) and without map (right). Dimerization domain of MeaB is not shown for clarity in either panel, and IcmF substrate-binding domains and linker regions are not shown on right for clarity. **B.** Superimposition of GDP-bound MeaB (purple/light purple) (PDB 2QM7)[31] aligned to the IcmF wt + GMPPCP structure with cryo-EM map (left) and without map (right). Dimerization domain

of MeaB is not shown for clarity in either panel, and IcmF substrate-binding domains and linker regions are not shown on right for clarity.

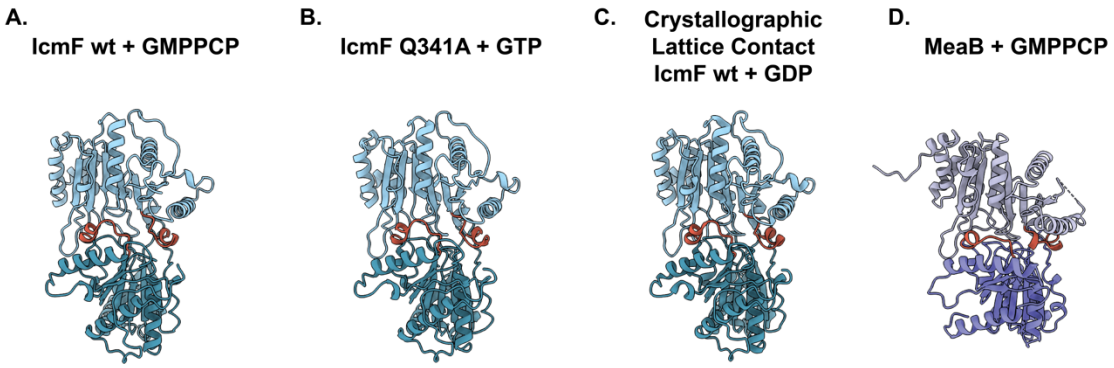


Figure III.15. Comparisons of the G-protein domain interface from the IcmF cryo-EM structures to previous structures. **A.** The arrangement of the G-protein domains from the IcmF wt + GMPPCP reconstruction. The switch III residues are colored red-orange and are present at the interface. **B.** The arrangement of the G-protein domains from the IcmF Q341A + GTP reconstruction. The switch III residues are colored red-orange and are present at the interface. **C.** The arrangement of the G-protein domains of IcmF wt from the crystallographic lattice contact (PDB 4XC6) [28] of IcmF wt in the presence of GDP. The switch III residues are colored red-orange and are present at the interface. **D.** The arrangement of the G-protein dimer of MeaB from *M. extorquens* in the presence of the Cbl-binding domain of its target mutase and GMPPCP (PDB 8DPB) [32] (Chapter II). The corresponding switch III residues are colored orange and are present at the interface. The C-terminal dimerization helices are not shown for clarity.

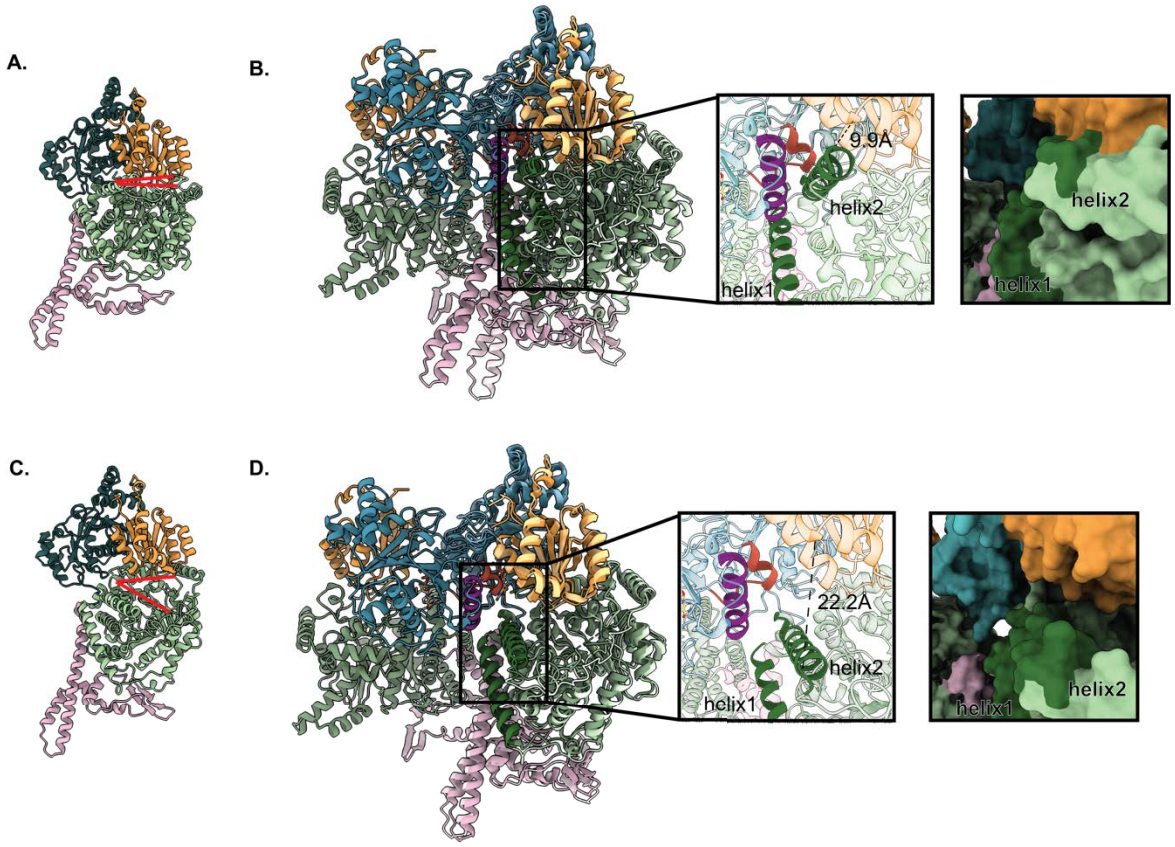


Figure III.16. The G-protein domain interface props open the active site of the mutase domain of IcmF. **A.** Ribbon drawing of the closed conformation of IcmF (PDB 4XC6)[28] with the substrate-binding domain (green) and Cbl-binding domain (orange), G-protein domain (blue) and linker (pink). Red lines indicate there is no gap between the domains. **B.** Overlay of the closed conformation of IcmF (PDB 4XC6)[28] with the IcmF Q341A + GTP structure. **Left Inset**: overlay reveals clash between the switch III region (red-orange, residues: 333-344) and a helix (purple, residues 363-380) from the neighboring IcmF protomer with two helices of the substrate-binding domain of the closed conformation of IcmF (dark green, helix 1 residues: 942-965, helix 2 residues: 974-995). The distance between helix 2 and the Cbl-binding domain is 9.9 Å. **Right Inset**: the same orientation as on left as surface representation but without the crashing G-protomer. **C.** Ribbon drawing of the open conformation of IcmF (PDB 4XC6)[28] colored as in A. Red lines indicate there is a gap between the domains. **D.** Overlay of the open conformation

of IcmF (PDB 4XC6) [28] with the IcmF Q341A + GTP structure. **Left Inset**: overlay shows no clashing between the helices of the substrate-binding domain of IcmF (dark green, helix 1 residues: 942-965, helix 2 residues: 974-995) with the switch III region (red-orange, residues 333-344) or helix (purple, residues 363-380) from the neighboring IcmF protomer. The distance between helix 2 and the Cbl-binding domain is 22.2 Å. **Right Inset**: the same orientation as on left as surface representation but without the wedging G-protomer.

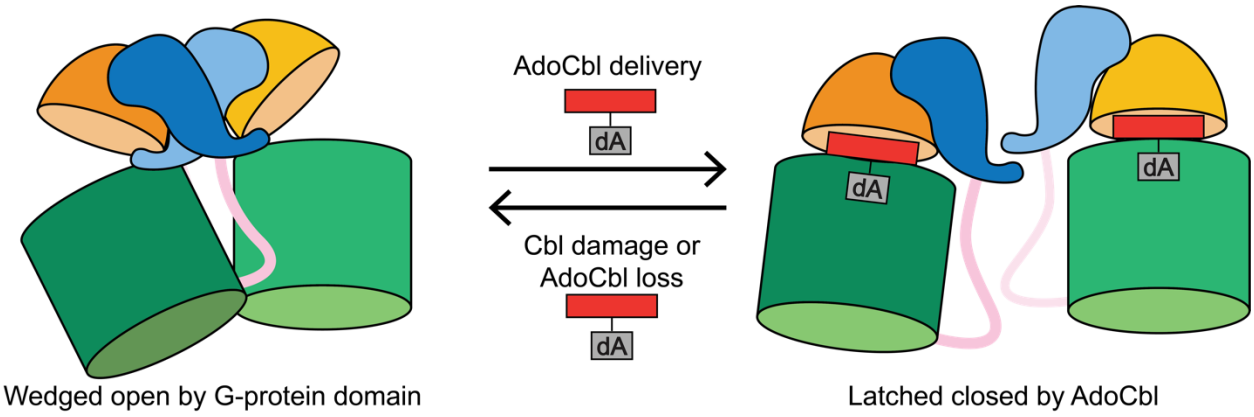


Figure III.17. Cartoon depicting equilibrium between open and closed states of IcmF mutase. The equilibrium is expected to shift to the right in the presence of a bound AdoCbl due to the adenosyl moiety of AdoCbl, which guts into the substrate-binding domain (green), securing the Cbl-binding domain (orange) against the substrate-binding domain. The equilibrium is expected to shift to the left in the case of AdoCbl damage or loss. Without the adenosyl moiety gutting across the mutase interface, the open conformation is more favorable, which in turn would facilitate IcmF oligomerization as the open mutase conformation has a pre-formed binding pocket for the G-protein "wedge." The second protomer of each IcmF molecule is not shown for simplicity.

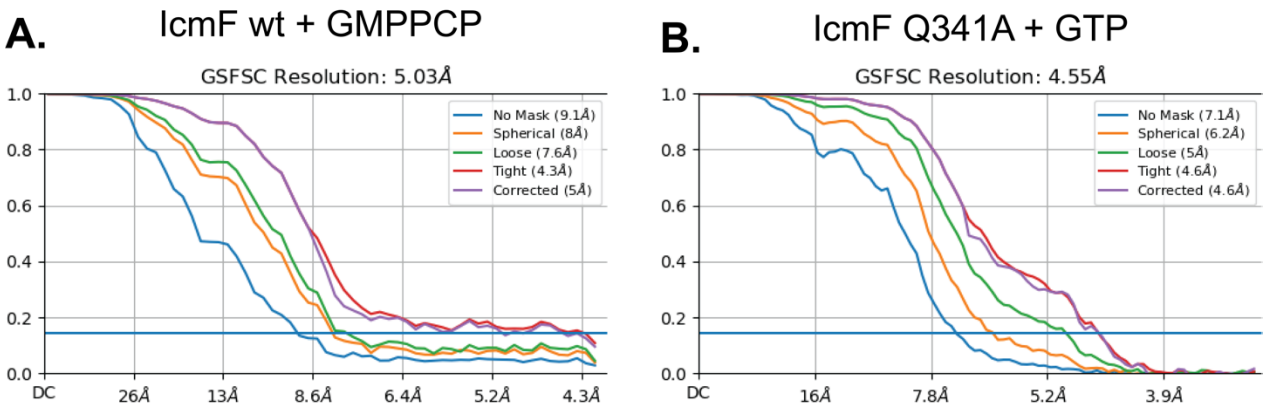


Figure III.18. FSC analysis of the two cryo EM reconstructions. **A.** FSC plots for the IcmF wt + GMPPCP model indicating the Fourier shell correlation as a function resolution with no mask (blue), a spherical mask (orange), a loose mask (green), a tight mask (red) and the corrected mask (purple). **B.** IcmF Q341A + GTP model indicating the Fourier shell correlation as a function of resolution with no mask (blue), a spherical mask (orange), a loose mask (green), a tight mask (red) and the corrected mask (purple). The masks used in the refinements and in the plots were generated by CryoSPARC.

Tables

Table III.1. Kinetic parameters for GTPase activity of wild type and switch III variants of IcmF and MeaB

	K_m GTP *μM*	V_{max} *μM/min*	k_{cat} min^{-1}	Catalytic efficiency k_{cat}/K_m	
IcmF*	2.4 ± 0 .9	1.568	3.13 ± 0.18	1.3	This study
	40 ± 8	N.R.**	18 ± 1.3	0.45	[19]
G. kaustophilus IcmF	51 ± 3	N. R.	10 ± 1	0.19	[19]
MeaB	N.R.	N.R.	0.039 ± 0.003	N. R.	[30]
MeaB + MCM	N.R.	N.R	4.20 ± 0.21	N.R.	[30]
IcmF Q341A	2.2 ± 0.8	0.533	1.07 ± 0.07	0.49	This study
MeaB Q185A + MCM	N.R.	N.R.	0.17 ± 0.03	N.R.	[30]
IcmF K344A	N.D.†	N.D.	N.D.	N.D.	This study
MeaB K188A + MCM	N.R.	N.R.	0.14 ± 0.02	N.R.	[30]

* Unless noted, all IcmF GTPase activity data is from *C. metallidurans*

** N.R.: Not reported

† N.D.: Not detected

Table III.2. Imaging parameters and 3D reconstruction

Dataset	**IcmF wt + GMPPCP**	**IcmF Q341A + GTP**
Microscope	FEI Talos Artica G2	FEI Talos Artica G2
Camera	Falcon 3EC	Falcon 3EC
Acceleration voltage (kV)	200	200
Magnification (x)	73000	92000
Pixel size (Å)	2.0143	1.5998
Defocus range (μm)	1.2 - 3.1	1.2 - 3.1
Number of frames	14	14
Exposure time (s)	3.99	3.99
Total exposure (e^-/Å)	41.81	58.71
Total micrographs collected	602	673
Automation software	EPU	EPU
Particles	167425	59335
Symmetry imposed	n/a	n/a
Unmasked resolution at 0.5/0.143 FSC (Å)	13/9.1	8.2/7.1
Masked resolution at 0.5/0.143 FSC (Å)	8.6/5.0	6.3/4.6

References

(1) Capdevila, Daiana A.; Edmonds, Katherine A.; Giedroc, David P. Metallochaperones and metalloregulation in bacteria. *Essays In Biochemistry* **2017**, *61* (2), 177-200. DOI: 10.1042/ebc20160076.

(2) Rosenzweig, A. C. Metallochaperones: bind and deliver. *Chemistry & Biology* **2002**, *9* (6), 673-677. DOI: 10.1016/S1074-5521(02)00156-4.

(3) O'Halloran, T. V.; Culotta, V. C. Metallochaperones, an intracellular shuttle service for metal ions. *Journal of Biological Chemistry* **2000**, *275* (33), 25057-25060. DOI: 10.1074/jbc.R000006200.

(4) Leipe, D. D.; Wolf, Y. I.; Koonin, E. V.; Aravind, L. Classification and evolution of P-loop GTPases and related ATPases. *Journal of Molecular Biology* **2002**, *317* (1), 41-72. DOI: 10.1006/jmbi.2001.5378.

(5) Froese, D. S.; Kochan, G.; Muniz, J. R.; Wu, X.; Gileadi, C.; Ugochukwu, E.; Krysztofinska, E.; Gravel, R. A.; Oppermann, U.; Yue, W. W. Structures of the human GTPase MMAA and vitamin B_{12}-dependent methylmalonyl-CoA mutase and insight into their complex formation. *Journal of Biological Chemistry* **2010**, *285* (49), 38204-38213.

(6) Padovani, D.; Labunska, T.; Banerjee, R. Energetics of interaction between the G-protein chaperone, MeaB, and B_{12}-dependent methylmalonyl-CoA mutase. *Journal of Biological Chemistry* **2006**, *281* (26), 17838-17844. DOI: 10.1074/jbc.M600047200.

(7) Lee, M. H.; Mulrooney, S. B.; Renner, M. J.; Markowicz, Y.; Hausinger, R. P. *Klebsiella aerogenes* urease gene cluster: sequence of *ureD* and demonstration that four accessory genes (*ureD*, *ureE*, *ureF*, and *ureG*) are involved in nickel metallocenter biosynthesis. *Journal of Bacteriology* **1992**, *174* (13), 4324-4330.

(8) Fong, Y. H.; Wong, H. C.; Yuen, M. H.; Lau, P. H.; Chen, Y. W.; Wong, K.-B. Structure of UreG/UreF/UreH complex reveals how urease accessory proteins facilitate maturation of *Helicobacter pylori* urease. *PLoS biology* **2013**, *11* (10), e1001678.

(9) Gasper, R.; Scrima, A.; Wittinghofer, A. Structural insights into HypB, a GTP-binding protein that regulates metal binding. *The Journal of Biological Chemistry* **2006**, *281* (37), 27492-27502. DOI: 10.1074/jbc.M600809200.

(10) Maier, T.; Jacobi, A.; Sauter, M.; Böck, A. The product of the *hypB* gene, which is required for nickel incorporation into hydrogenases, is a novel guanine nucleotide-binding protein. *Journal of Bacteriology* **1993**, *175* (3), 630-635.

(11) Takahashi-Iñiguez, T.; González-Noriega, A.; Michalak, C.; Flores, M. E. Human MMAA induces the release of inactive cofactor and restores methylmalonyl-CoA mutase activity through their complex formation. *Biochimie* **2017**, *142*, 191-196. DOI: 10.1016/j.biochi.2017.09.012.

(12) Padovani, D.; Labunska, T.; Palfey, B. A.; Ballou, D. P.; Banerjee, R. Adenosyltransferase tailors and delivers coenzyme B_{12}. *Nature Chemical Biology* **2008**, *4* (3), 194.

(13) Padovani, D.; Banerjee, R. A G-protein editor gates coenzyme B_{12} loading and is corrupted in methylmalonic aciduria. *Proceedings of the National Academy of Sciences* **2009**, *106* (51), 21567-21572. DOI: 10.1073/pnas.0908106106.

(14) Cracan, V.; Padovani, D.; Banerjee, R. IcmF is a fusion between the radical B_{12} enzyme isobutyryl-CoA mutase and its G-protein chaperone. *Journal of Biological Chemistry* **2010**, *285* (1), 655-666. DOI: 10.1074/jbc.M109.062182.

(15) Dempsey-Nunez, L.; Illson, M. L.; Kent, J.; Huang, Q.; Brebner, A.; Watkins, D.; Gilfix, B. M.; Wittwer, C. T.; Rosenblatt, D. S. High resolution melting analysis of the MMAA gene in patients with cblA and in those with undiagnosed methylmalonic aciduria. *Molecular Genetics and Metabolism* **2012**, *107* (3), 363-367.

(16) Banerjee, R. Radical carbon skeleton rearrangements: Catalysis by coenzyme B_{12} -dependent mutases. *Chemical Reviews* **2003**, *103* (6), 2083-2094. DOI: 10.1021/cr0204395.

(17) Korotkova, N.; Chistoserdova, L.; Kuksa, V.; Lidstrom, M. E. Glyoxylate regeneration pathway in the methylotroph *Methylobacterium extorquens* AM1. *Journal of Bacteriology* **2002**, *184* (6), 1750-1758. DOI: doi:10.1128/JB.184.6.1750-1758.2002.

(18) Li, Z.; Kitanishi, K.; Twahir, U. T.; Cracan, V.; Chapman, D.; Warncke, K.; Banerjee, R. Cofactor Editing by the G-protein Metallochaperone Domain Regulates the Radical B12 Enzyme IcmF*♦. *Journal of Biological Chemistry* **2017**, *292* (10), 3977-3987. DOI: 10.1074/jbc.M117.775957.

(19) Cracan, V.; Banerjee, R. Novel coenzyme B_{12}-dependent interconversion of isovaleryl-CoA and pivalyl-CoA. *Journal of Biological Chemistry* **2012**, *287* (6), 3723-3732. DOI: 10.1074/jbc.M111.320051.

(20) Mansoorabadi, S. O.; Padmakumar, R.; Fazliddinova, N.; Vlasie, M.; Banerjee, R.; Reed, G. H. Characterization of a succinyl-CoA radical-Cob(II)alamin spin triplet intermediate in the reaction catalyzed by adenosylcobalamin-dependent methylmalonyl-CoA mutase. *Biochemistry* **2005**, *44* (9), 3153-3158. DOI: 10.1021/bi0482102.

(21) Sension, R. J.; Cole, A. G.; Harris, A. D.; Fox, C. C.; Woodbury, N. W.; Lin, S.; Marsh, E. N. G. Photolysis and recombination of adenosylcobalamin bound to glutamate mutase. *Journal of the American Chemical Society* **2004**, *126* (6), 1598-1599. DOI: 10.1021/ja0396910.

(22) Takahashi-Íñiguez, T.; García-Arellano, H.; Trujillo-Roldán, M. A.; Flores, M. E. Protection and reactivation of human methylmalonyl-CoA mutase by MMAA protein. *Biochemical and Biophysical Research Communications* **2011**, *404* (1), 443-447. DOI: 10.1016/j.bbrc.2010.11.141.

(23) Padovani, D.; Banerjee, R. Assembly and protection of the radical enzyme, methylmalonyl-CoA mutase, by its chaperone. *Biochemistry* **2006**, *45* (30), 9300-9306. DOI: 10.1021/bi0604532.

(24) Mancia, F.; Smith, G. A.; Evans, P. R. Crystal structure of substrate complexes of methylmalonyl-CoA mutase. *Biochemistry* **1999**, *38* (25), 7999-8005. DOI: 10.1021/bi9903852.

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

(26) Mancia, F.; Keep, N. H.; Nakagawa, A.; Leadlay, P. F.; McSweeney, S.; Rasmussen, B.; secke, P. B.; Diat, O.; Evans, P. R. How coenzyme B_{12} radicals are generated: the crystal structure of methylmalonyl-coenzyme A mutase at 2 Å resolution. *Structure* **1996**, *4* (3), 339-350. DOI: 10.1016/s0969-2126(96)00037-8.

(27) Ruetz, M.; Campanello, G. C.; McDevitt, L.; Yokom, A. L.; Yadav, P. K.; Watkins, D.; Rosenblatt, D. S.; Ohi, M. D.; Southworth, D. R.; Banerjee, R. Allosteric regulation of oligomerization by a B_{12} trafficking G-protein is corrupted in methylmalonic aciduria. *Cell Chemical Biology* **2019**, *26* (7), 960-969.e964. DOI: 10.1016/j.chembiol.2019.03.014.

(28) Jost, M.; Cracan, V.; Hubbard, P. A.; Banerjee, R.; Drennan, C. L. Visualization of a radical B_{12} enzyme with its G-protein chaperone. *Proceedings of the National Academy of Sciences* **2015**, *112* (8), 2419-2424.

(29) Lofgren, M.; Koutmos, M.; Banerjee, R. Autoinhibition and signaling by the switch II motif in the G-protein chaperone of a radical B_{12} enzyme. *Journal of Biological Chemistry* **2013**, *288* (43), 30980-30989. DOI: 10.1074/jbc.M113.499970.

(30) Lofgren, M.; Padovani, D.; Koutmos, M.; Banerjee, R. A switch III motif relays signaling between a B_{12} enzyme and its G-protein chaperone. *Nature Chemical Biology* **2013**, *9* (9), 535-541. DOI: 10.1038/nchembio.1298.

(31) Hubbard, P. A.; Padovani, D.; Labunska, T.; Mahlstedt, S. A.; Banerjee, R.; Drennan, C. L. Crystal structure and mutagenesis of the metallochaperone MeaB: Insight into the causes of methylmalonic aciduria. *Journal of Biological Chemistry* **2007**, *282* (43), 31308-31316. DOI: 10.1074/jbc.M704850200.

(32) Vaccaro, F. A.; Born, D. A.; Drennan, C. L. Structure of metallochaperone in complex with the cobalamin-binding domain of its target mutase provides insight into cofactor delivery. *Proceedings of the National Academy of Sciences* **2023**, *120* (8), e2214085120. DOI: 10.1073/pnas.2214085120 From NLM Medline.

(33) Mascarenhas, R.; Ruetz, M.; McDevitt, L.; Koutmos, M.; Banerjee, R. Mobile loop dynamics in adenosyltransferase control binding and reactivity of coenzyme B_{12}. *Proceedings of the National Academy of Sciences* **2020**, *117* (48), 30412-30422. DOI: doi:10.1073/pnas.2007332117.

(34) Jost, M.; Born, D. A.; Cracan, V.; Banerjee, R.; Drennan, C. L. Structural basis for substrate specificity in adenosylcobalamin-dependent isobutyryl-CoA mutase and related acyl-CoA mutases. *Journal of Biological Chemistry* **2015**, *290* (45), 26882-26898.

(35) Punjani, A.; Rubinstein, J. L.; Fleet, D. J.; Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* **2017**, *14* (3), 290-296. DOI: 10.1038/nmeth.4169.

(36) Kaur, S.; Gomez-Blanco, J.; Khalifa, A. A. Z.; Adinarayanan, S.; Sanchez-Garcia, R.; Wrapp, D.; McLellan, J. S.; Bui, K. H.; Vargas, J. Local computational methods to improve the interpretability and analysis of cryo-EM maps. *Nature Communications* **2021**, *12* (1), 1240. DOI: 10.1038/s41467-021-21509-5.

(37) Townsend, P. D.; Jungwirth, B.; Pojer, F.; Bußmann, M.; Money, V. A.; Cole, S. T.; Pühler, A.; Tauch, A.; Bott, M.; Cann, M. J.; et al. The crystal structures of apo and cAMP-bound GlxR from *Corynebacterium glutamicum* reveal structural and dynamic changes upon cAMP binding in CRP/FNR family transcription factors. *PLOS ONE* **2014**, *9* (12), e113265. DOI: 10.1371/journal.pone.0113265.

(38) Lazazzera, B. A.; Beinert, H.; Khoroshilova, N.; Kennedy, M. C.; Kiley, P. J. DNA binding and dimerization of the FeS-containing FNR protein from *Escherichia coli* are regulated by oxygen. *Journal of Biological Chemistry* **1996**, *271* (5), 2762-2768. DOI: 10.1074/jbc.271.5.2762 (acccessed 2023/01/29).

(39) Vaccaro, F. A.; Drennan, C. L. The role of nucleoside triphosphate hydrolase metallochaperones in making metalloenzymes. *Metallomics* **2022**, *14* (6), mfac030.

(40) Wilkins, M. R.; Gasteiger, E.; Bairoch, A.; Sanchez, J.-C.; Williams, K. L.; Appel, R. D.; Hochstrasser, D. F. Protein identification and analysis tools in the ExPASy server. *Methods in Molecular Biology* **1999**, *112*, 531-552. DOI: 10.1385/1-59259-584-7:531.

(41) Niebling, S.; Veith, K.; Vollmer, B.; Lizarrondo, J.; Burastero, O.; Schiller, J.; Struve García, A.; Lewe, P.; Seuring, C.; Witt, S.; et al. Biophysical Screening Pipeline for Cryo-EM Grid Preparation of Membrane Proteins. *Frontiers in Molecular Biosciences* **2022**, *9*, Methods. DOI: 10.3389/fmolb.2022.882288.

(42) Kimanius, D.; Dong, L.; Sharov, G.; Nakane, T.; Scheres, S. H. W. New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochemical Journal* **2021**, *478* (24), 4169-4185. DOI: 10.1042/BCJ20210708 (acccessed 10/21/2022).

(43) Zheng, S. Q.; Palovcak, E.; Armache, J.-P.; Verba, K. A.; Cheng, Y.; Agard, D. A. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods* **2017**, *14* (4), 331-332. DOI: 10.1038/nmeth.4193.

(44) Rohou, A.; Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology* **2015**, *192* (2), 216-221. DOI: https://doi.org/10.1016/j.jsb.2015.08.008.

(45) Cianfrocco, M. A.; Wong-Barnum, M.; Youn, C.; Wagner, R.; Leschziner, A. COSMIC2: A science gateway for cryo-electron microscopy structure determination. In Proceedings of the Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact, New Orleans, LA, USA; 2017.

(46) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Meng, E. C.; Couch, G. S.; Croll, T. I.; Morris, J. H.; Ferrin, T. E. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science* **2021**, *30* (1), 70-82. DOI: 10.1002/pro.3943.

(47) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66* (4), 486-501. DOI: 10.1107/s0907444910007493.

(48) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66* (2), 213-221.

(49) Morin, A.; Eisenbraun, B.; Key, J.; Sanschagrin, P. C.; Timony, M. A.; Ottaviano, M.; Sliz, P. Collaboration gets the most out of software. *elife* **2013**, *2*, e01456.

(50) Dowling, D. P.; Croft, A. K.; Drennan, C. L. Radical Use of Rossmann and TIM Barrel Architectures for Controlling Coenzyme B 12 Chemistry. *Annual Review of Biophysics* **2012**, *41* (1), 403-427. DOI: 10.1146/annurev-biophys-050511-102225.

(51) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **2011**, *7* (1), 539. DOI: 10.1038/msb.2011.75.

CHAPTER IV: The role of the adensosyltransferase in mutase maturation

Author contributions

Francesca A. Vaccaro conceptualized the experiments and performed them with Daphne A Faber. F. A. Vaccaro wrote the original draft of all sections in this chapter, excluding the methods for the crystallography which was written by D.A. Faber. F. A. Vaccaro and Catherine L. Drennan contributed to the editing of the chapter.

Summary

Metallochaperones transport, modify and deliver essential metallocofactors in the process of metalloenzyme maturation. To mature the adenosylcobalamin (AdoCbl)-dependent methylmalonyl-CoA mutase (MCM), a G-protein chaperone (MeaB in bacteria and MMAA in humans) and an adenosyltransferase (ATR) work in concert to prepare and deliver AdoCbl. The ATR catalyzes the synthesis of AdoCbl from cob(II)alamin in a reductive adenosylation reaction before releasing it to the MCM. The G-protein uses GTP binding and hydrolysis to prepare the mutase to receive the AdoCbl from ATR. Chapter II and III of this thesis investigate G-protein:mutase interactions using two systems: a minimal system composed of a G-protein MeaB and a MCM cobalamin-binding domain from *Methylobacterium extorquens*; and a natural fusion system containing a G-protein fused to AdoCbl-dependent isobutyryl-CoA mutase from *Cupriavidus metallidurans* (*Cm*IcmF). This chapter focuses on an ATR for which no structures of ATR:mutase or ATR:G-protein complexes are available. Here we present the crystal structure of apo-ATR from *C. metallidurans* to 2.35 Å resolution and compare this structure to previously determined structures of ATRs from other organisms that have product and/or substrate bound. Our structural analyses indicate high structural similarity overall and suggest that the active sites of ATRs are pre-formed for Cbl binding with conserved residues in *Cm*ATR already positioned to interact the Cbl substrate. We further use this *Cm*ATR structure and computational docking to begin to investigate how *Cm*ATR may interacts with *Cm*IcmF, the natural G-protein:mutase fusion, for cofactor handoff. Alphafold2 multimer fails to predict an ATR:IcmF complex of any kind, and RosettaDocking4.0 generates a series of solutions that sample a wide range of orientations of the *Cm*ATR with respect to *Cm*IcmF, none of which are clearly superior to the others. We discuss the pros and cons of the possible RosettaDocking4.0 docking solutions in light of the biochemical data that are available and suggest further experiments aimed at understanding G-protein-assisted AdoCbl handoff between an ATR and a target mutase.

Introduction

Metallocofactors are important for a wide range of metabolic processes, expanding the diversity and complexity of chemical reactions performed by enzymes[1, 2]. Cobalamin (Cbl) is an essential metallocofactor for humans, obtained through their diet. Humans

ingest Cbl as cyanocobalamin (CNCbl), also known as vitamin B_{12}, and utilize a complex series of metallochaperones responsible for the transport, modification, and delivery of the active Cbl species to their final metalloenzymes[3]. Humans have only two enzymes that need Cbl to perform their chemistry: 1) methionine synthase, which requires methylcobalamin (MeCbl) to catalyze the formation of methionine from homocysteine[4], and 2) methylmalonyl-CoA mutase (MCM), which requires adenosylcobalamin (AdoCbl) to catalyze the conversion of methylmalonyl-CoA to succinyl-CoA[5, 6]. MCM's catalytic activity is important for efficient metabolism of certain amino acids, fatty acids and other metabolites[7], and disruptions of MCM and/or any of the other metallochaperones involved in the maturation of MCM results in methylmalonic acidemia, a disease state associated with an inborn error of metabolism[8].

The AdoCbl that MCM requires is synthesized from inactive Cbl by the adenosyltransferase (ATR). Briefly, the ATR binds cob(II)alamin, which has been previously reduced from the cob(III)alamin ingested and trafficked[9]. If the ATP is bound to the ATR, cob(II)balamin binding facilitates the formation of a four-coordinate cob(II)alamin which then becomes reduced to cob(I)alamin[10-12]. The supernucleophilic cob(I)alamin then attacks the C5′ atom of the bound ATP, releasing PPP_i and forming the cobalt—carbon bond of AdoCbl. The final step in MCM's maturation is the delivery of AdoCbl from the ATR to the Cbl-binding domain of MCM, a step that is facilitated by a G-protein metallochaperone, methylmalonic aciduria type A (MMAA)[13]. The G-protein metallochaperone utilizes GTP hydrolysis to regulate the loading of AdoCbl by the ATR and the unloading of damaged Cbl to the ATR (*Figure IV.1*)[14-16].

In addition to the human system, two model bacterial systems have been used to study AdoCbl transfer to MCM, yielding extensive structural and biochemical information on the interactions of the final step (*Figure IV.2*). An orthologous three-component system from *Methylobacterium extorquens* contains a G-protein chaperone, known as MeaB, an ATR and an MCM. One notable difference is the oligomerization of MCM; the MCM from *M. extorquens* is a heterodimer with one active and one inactive subunit whereas the *H. sapiens* MCM is a homodimer with two active subunits. Recent work indicated the conformation of MeaB in the presence of the Cbl-binding domain of MCM is active for GTP hydrolysis (Chapter II). In this conformation, the conserved switch III motif, a motif

also present in MMAA which harbors causative mutations for methylmalonic aciduria in humans, is positioned to participate in GTP hydrolysis due to a ~30 Å conformational shift from the previously characterized inactive conformation of the G-protein chaperone (Chapter II). There is also a two-component system from *Cupriavidus metallidurans* (*Figure IV.2*), which contains isobutyryl-CoA fused (IcmF), an AdoCbl-dependent mutase that naturally contains the mutase domains and G-protein chaperone domain on the same polypeptide chain[17]. The crystal structure of IcmF provided the first view of an AdoCbl-dependent mutase in complex with its G-protein chaperone[18]. Recent work (Chapter III) has indicated that maturation of IcmF involves the formation of a transient G-protein domain dimer that has a similar conformation to the active conformation of MeaB from *M. extorquens* (Chapter II). The same active conformation is predicted to form in human MMAA as well (Chapter III). This conformation of the G-protein dimer appears to wedge open the active site of the mutase domain of IcmF, facilitating the state of the mutase that is competent for cofactor loading and unloading (Chapter III). These systems have provided numerous insights into the molecular interactions of mutase and G-protein chaperone.

Despite vast biochemical information regarding the individual enzymes required for the delivery of AdoCbl to its target mutase, numerous questions remain regarding the molecular interactions of the G-protein chaperone:ATR:mutase complex and how disruptions to those interactions result in disease phenotypes. Biochemical studies suggest that the formation of a G-protein chaperone:ATR:mutase complex is required to efficiently load and unload the Cbl cofactor, limiting the inactivation or loss of Cbl to solution[19]. There are three classes of ATRs found in nature; however the most extensively studied is the PduO-type ATR which is the only type found in humans[20]. Numerous structures of ATRs from various organisms that have been previously solved with AdoCbl, cob(II)alamin, ATP, and/or PPP_i bound in the three binding sites of this trimeric protein. These structures have provided information regarding residues that are important for substrate binding and catalysis, but these structures failed to explain how ATR's product, AdoCbl, is transferred (*Table IV.1*)[11, 21-27]. Additionally, a lack of structural information regarding the interfaces and important residues in the G-protein chaperone:mutase:ATR

complex limits our ability to understand how Cbl is transferred between the ATR and the mutase.

Given the recent work indicating the similar mechanism of action of the G-protein chaperone domains of two-component system compared to the G-protein chaperones in the three-component systems (Chapter III), it appears that the fused system provides a suitable model to investigate the molecular interactions of the G-protein chaperone:mutase:ATR complex as a basis for understanding the human system. Here we present the crystal structure of a PduO-type ATR from *C. metallidurans.* This structure contains conserved motifs found in other characterized PduO-type ATRs, such as the human ATR and *M. tuberculosis* ATR. We use this structure to predict the interactions between *C. metallidurans* ATR and *C. metallidurans* IcmF, using various computational methods to generate models. These models provide insight into the potential interfaces and molecular interactions that facilitate efficient Cbl delivery and removal by the ATR.

Results

CmATR is a PduO-type with an overall fold consistent with other PduO-type ATRs

To determine if *Cm*ATR (MW of 19.9 kDa) is analogous to the other PduO-type ATRs, the oligomeric state in solution and structure were determined. Size exclusion chromatography was used to determine the oligomeric state of *Cm*ATR, showing that *Cm*ATR migrates with an apparent molecular weight of 65 kDa, consistent with the formation of a trimer, as previously reported for other PduO-type ATRs (*Figure IV.3A*). To investigate the molecular structure, we crystallized and solved the structure of *Cm*ATR to 2.35 Å resolution using molecular replacement (*Table IV.2*). The final model contained six copies of the *Cm*ATR protomer in the asymmetric unit, with three of the protomers forming a trimer as observed in solution (*Figure IV.3B,C*). *Cm*ATR has the characteristic ferritin-like fold observed in other PduO-type ATRs (*Figure IV.3B).* Residues 1-29 of the N-terminus of each protomer are disordered, which is expected in the absence of substrate because of these residues are predicted to be involved in completing the substrate binding site[21].

To determine if the apo-*Cm*ATR structure is analogous to other characterized PduO-type ATRs, the trimer of *Cm*ATR was aligned on the ATR from *Homo sapiens* (*Hs*ATR) bound to ATP, AdoCbl, and Cbl (II) (PDB 6D5X)[22] and the ATR from

Mycobacterium tuberculous (*Mt*ATR) bound to AdoCbl and PPP_i (PDB 6WGV)[21] (*Figure IV.3D*). Despite having ATP and/or Cbl bound in the active site of *Hs*ATR or *Mt*ATR, the root mean squared deviation (RMSD) values for all atoms was 1.7 $Å^2$ and 2.6 $Å^2$, respectively. The low RMSD values for both structures compared to the *Cm*ATR structure indicates the overall structure of the *Cm*ATR trimer is not predicted to undergo widespread conformational changes upon ATP and/or AdoCbl binding. The alignment of the *Cm*ATR trimer to the *Mt*ATR and *Hs*ATR structures illustrates the similarity in the overall organization and folds for all PduO-type ATRs reported.

The predicted cobalamin binding site of CmATR is conserved

To determine which residues are important for the function of the *Cm*ATR, a sequence alignment analysis and structural analysis was completed. First, the sequences of the ATRs from *H. sapiens, C. metallidurans, Burkholderia thailandensis, M. extorquens*, *M. tuberculosis,* and *Propionibacterium freudenreichii subsp. shermanii* were aligned using Clustal(O) to reveal conserved residues (*Figure IV.4*)[28]. The alignment indicated numerous conserved regions known to be involved in ATP, Cbl(II) and AdoCbl binding. Because ATR needs to deliver its product, AdoCbl, to another enzyme, the binding site of *Cm*ATR, like other PduO-type ATRs, was predicted to be solvent exposed. Aligning the structure of *Mtb*ATR bound to AdoCbl and PPP_i on the apo *Cm*ATR structure suggests that these binding sites will be solvent exposed in *Cm*ATR (*Figure IV.5A*). A surface representation of *Mtb*ATR bound to AdoCbl and PPP_i aligned on *Cm*ATR further indicates the binding pocket is solvent exposed. However, the location of ATP and PPP_i binding is more protected by the N-terminus than the corrin ring and dimethylbenzamide (DMB) tail of AdoCbl (*Figure IV.5B*). Closer inspection of the alignment of the AdoCbl binding site of *Mtb*ATR to *Cm*ATR reveals that conserved residues are in similar locations despite the lack of substrate or product bound. Before binding Cbl(II), ATRs must bind ATP [11]. Based on the sequence alignment and previous structural data, N-terminal residues are known to bind the triphosphate of ATP, Mg^{2+} ions, and water molecules *(Figure IV.4, Table IV.1)*; however, those residues are disordered in the *Cm*ATR structure, limiting the comparison of the position of those residues in *Cm*ATR to previous structures.

After binding ATP, ATRs bind inactive cobalamin in order to form AdoCbl[29]. Previous structural characterizations of various ATRs emphasized the importance of the hydrophobic pocket for maintaining a four-coordinate cobalt center required for the formation of AdoCbl when bound to the ATR[10-12]. The sequence alignment indicates the hydrophobic pocket characterized in other organisms is conserved in *Cm*ATR (*Figure IV.4*). Similarly, the structural alignment of *Cm*ATR on *Mtb*ATR indicates that despite lacking bound substrates/products, the orientation of the C-terminal residues involved in creating the hydrophobic pocket is largely conserved when compared to the *M. tuberculosis* structure with AdoCbl bound. Specifically, the conserved phenylalanine residue (F110 in *Cm*ATR and F117 in *Mtb*ATR) maintains the orientation needed to block the DMB tail of Cbl or a water molecule from coordinating the cobalt center of Cbl (*Figure IV.5C*).

After the formation of AdoCbl by the ATR, the binding site of ATR contains PPP_i and AdoCbl, the latter of which is delivered to its target mutase. Aligning the structure of *Cm*ATR onto the structure of *Mtb*ATR with bound PPP_i and AdoCbl predicts a similar mode of interactions (*Figure IV.5D*). As was the case for the hydrophobic pocket, many of the residues of *Cm*ATR predicted to interact with the PPP_i after the formation of AdoCbl are in similar locations to their corresponding residues in *Mtb*ATR. Once again, the comparison of the binding modes of PPP_i and AdoCbl is limited by the disordered N-terminus of *Cm*ATR, which may be disordered because PPP_i is not present.

A structure of *Mtb*ATR bound to Cbl(II) and PPP_i has also been previously reported[21]. Although this state does not represent a step in the catalytic cycle of ATR, this structure captured a novel conformation of the DMB tail of the Cbl in which it is extended and interacts with C-terminal residues (*Figure IV.5E*). Structural superimpositions show that core helices of the *Mtb*ATR and *Cm*ATR are in good agreement with the only difference being in the N-terminal residues.

Modeling of the CmATR: IcmF complex predicts various complexes that lack a consensus interface

To understand how cofactor transfer between the ATR and mutase occurs, we used computational modeling to predict the structure of the complex formed between *Cm*ATR and *Cm*IcmF. First, alphafold2 multimer was used to model the complex using

three copies of the *Cm*ATR sequence representing the trimer and two copies of the *Cm*IcmF sequence representing the dimer[30]. However, alphafold2 multimer did not predict an interface between the two proteins, instead folding them into their correct respective oligomeric states (*Figure IV.6*). Second, RosettaDocking4.0[31] was used to find docking solutions for *Cm*ATR with *Cm*IcmF. In particular, the structures used were apo-*Cm*ATR trimer and the *Cm*IcmF dimer in the "open state," in which the Cbl-binding domain is displaced from the substrate-binding domain, generating an opening for AdoCbl insertion (Chapter III). To generate the input models, apo-*Cm*ATR and *Cm*IcmF were manually placed near each other (*Figure IV.7*). The first input model placed *Cm*ATR near the Cbl-binding domain to minimize the transfer distance between the ATR and the IcmF. The second input model placed *Cm*ATR near the G-protein chaperone domain interface of IcmF. RosettaDocking4.0 generated over 400 docking solutions for each input model; the top ten docking solutions for each input model were manually inspected. Analysis of the top 20 docking solutions indicated that there was no consensus orientation for *Cm*ATR relative to *Cm*IcmF, with many models placing the Cbl binding sites away from the mutase active site or an interface between the *Cm*ATR and the substrate binding domain of *Cm*IcmF. Four docking solutions, which are representative of the most probable modes of interacting, are shown in *Figure IV.8*. We found that the docking solution that provided the maximum contacts between ATR and the G-protein domain dimer of IcmF also had the longest transfer distance for Cbl, over 50 Å from the phenylalanine involved in forming the hydrophobic pocket (F110) on the ATR to the histidine (His39) that coordinates the cobalt of the Cbl-binding domain of IcmF (*Figure IV.8A*)[18]. Three docking solutions were fairly similar to each other (*Figure IV.8B-D*), showing ATR at different orientations with respect the Cbl-binding domain. These docking solutions have closer Cbl transfer distances between the F110 from ATR and His39 from IcmF with distances between 20 Å and 30 Å. For these three solutions, ATR has little to no contact with the G-protein domain of IcmF.

Discussion

As a metallochaperone, ATR must bind various forms of cobalamin, modify the inactive forms, and then deliver it to its target mutase. There have been numerous

investigations into the mechanism of modifying inactive Cbl to form active AdoCbl[10-12, 32, 33]. However, little is known regarding how ATR interacts with its target mutase to deliver AdoCbl and remove inactivated Cbl. Previous biochemical work has established the requirements for cofactor transfer between the ATR and mutase, indicating the importance of the DMB tail of Cbl and the necessity of a G-protein chaperone[19, 21, 22]. Recent structural work has established that the fused system utilizes the same active conformation and oligomeric state of the G-protein observed in the *M. extorquens* system (Chapters II & III), allowing the fused system from *C. metallidurans* to serve as a simpler model to characterize the structure of the ATR:G-protein:mutase complex. However, the ATR from *C. metallidurans* has not been structurally characterized previously. Solving the structure of apo-*Cm*ATR by crystallization revealed a trimeric protein with a conserved overall ferritin-like fold as observed for all previously characterized PduO-type ATRs (*Figure IV.3B, Table IV.1*). The alignment of apo-*Cm*ATR on *Hs*ATR and *Mt*ATR indicates high structural similarity in the overall folds of the trimers (*Figure IV.3D)*. The main differences are localized at the N-terminal region that is involved in binding substrate ATP and product PPP_i. In many other structures with substrate and/or product bound, the N-terminal region is ordered but it is disordered in the apo-*Cm*ATR structure, preventing comparison (*Figure IV.5*)[21, 26]. However, the sequence alignment of PduO-type ATRs indicates a high level of conservation in the N-terminal region, supporting a similar mode of binding of ATP/PPP_i and a conserved role of the N-terminal region for *Cm*ATR (*Figure IV.4*).

The alignment of apo-*Cm*ATR on *Mtb*ATR with various substrate and/or products bound allows for structure-based active site comparisons. For the regions that are not disordered, the similarity of the positions of the residues involved in binding Cbl, ATP, and/or PPP_i for *Cm*ATR to *Mtb*ATR suggests that the binding sites, except for mobile the N-terminal region, are pre-positioned (*Figure IV.5*). This pre-positioning allows for the use of the apo-*Cm*ATR for computational modeling of the ATR:IcmF complex with high confidence that the binding of substrate and/or product during the cofactor transfer and complex formation does not rely on large changes in the conformation of the ATR as a whole. However, the disorder of the N-terminal residues in the apo-*Cm*ATR structure,

which may be important for controlling the coordination of the Cbl and cofactor transfer, is a caveat with respect to an conclusion drawn from the apo-*Cm*ATR modeling[21].

The attempt to computationally predict the structure of ATR:G-protein:mutase complex provided numerous potential interfaces, implicating various residues as being important for cofactor transfer. The alphafold2 multimer model accurately folded both the structure of the *Cm*ATR trimer and the *Cm*IcmF dimer but did not predict formation of a complex (*Figure IV.6*). Because the competent oligomer for cofactor transfer to and from *Cm*IcmF requires the formation of a G-protein domain interface, future attempts to use alphafold2 multimer to predict the ATR:G-protein:mutase complex should supply four copies of the sequence of IcmF to allow for the formation of the tetramer (Chapter III). Using RosettaDocking-4.0, numerous models were predicted, with heavy biasing from the manually created input models (*Figure IV.7-9*). The docking solutions only moved the ATR trimer relative to the IcmF dimer, rotating it around IcmF with no clear solution (*Figure IV.9*). These rotations often led to the positioning of the binding site of *Cm*ATR away from the opening of the *Cm*IcmF active site (*Figure IV.8*). With no obvious best solution for the docking of the ATR:G-protein:mutase complex from these attempts, future docking experiments could employ the transient *Cm*IcmF tetramer consisting of the G-protein domain and substrate-binding domain interface that is believed to be the competent state for AdoCbl transfer (Chapter III).

We expect that an ATR:G-protein:mutase complex that is competent for AdoCbl delivery will have the active site of the ATR pointing toward the active site of IcmF, will have a close distance between the two active sites, and will have the DMB tail of the ATR-bound AdoCbl available to the mutase. In *M. tuberculosis*, the DMB tail of AdoCbl has been shown to be important for efficient transfer of AdoCbl from the ATR to the mutase; transfer of 5′-deoxyadenosyl cobinamide, a form of Cbl which lacks the DMB tail, is not observed in either direction by UV-Vis spectroscopy[21]. In the *Mtb*ATR structures, any bound Cbl has the DMB tail solvent exposed, priming it to interact with mutase for cofactor transfer (*Figure IV.5B,C,E*). Additionally, the DMB tail can sample numerous conformations. The *Mtb*ATR structure bound to Cbl(II) and PPP_i shows an extended DMB tail, physically displaced by the N-terminal residues that are disordered in all other structures (*Figure IV.5E*). Thus, as long as the active site of ATR is pointing toward the

IcmF active site, the DMB tail should be available for the mutase to bind. The next consideration is a short distance. The docking solutions place the AdoCbl bound to ATR more than 20 Å away from its binding site in the Cbl-binding domain of IcmF (*Figure IV.8*). This distance may be too distant for a direct handoff of the cofactor; however, there could be additional Cbl binding sites and a delivery process that involves multiple steps. Given the small size of ATR, an additional Cbl binding site on IcmF is more probable. Extra AdoCbl-binding sites on IcmF could be elucidated using crystallography; apo-IcmF crystals could be soaked with excess AdoCbl to resolve sites that could then be validated using mutagenesis and cofactor transfer assays[18, 19, 21]. Another possibility is that IcmF could undergo a further conformational change that would move the Cbl-binding domain closer to ATR. Large swinging movements of Cbl-binding domains have been observed previously[34, 35], but in those cases, the Cbl-binding domain was not tethered to a G-protein domain restricting the Cbl-binding domain's movements. Previous snapshots of IcmF show that the Cbl-binding domain and the G-protein domain form a continuous β-sheet and the two domains moves as a rigid body[18], restricting the flexibility of motion of the Cbl-binding domain to swing closer to ATR. However, some degree of motion of the Cbl-binding domain could occur, shortening the ATR-IcmF distance. Finally, there may be a docking solution not discovered yet that has a closer distance without any conformational changes needed or multiple step processes required.

In the *C. metallidurans* system, removal of inactivated cob(II)alamin from IcmF does not occur without the presence of ATR and GTP in vitro with no cob(II)alamin ejected into solution[19]. These requirements support a mechanism of metallochaperone mediated cofactor removal and imply the presence of an ATR:IcmF complex; however, the stability of the complex is still unknown. Ultimately, the questions regarding the molecular mechanism could be answered by the structural characterizations of the ATR:IcmF complex in either the cofactor loading or unloading state. The use of the fused system reduces the complexity of the structural characterizations and simultaneously provides valuable information to inform our understanding of how the non-fused systems found in humans function, expanding our understanding of the mutase maturation and the molecular interactions that govern disease states in humans.

Materials

All chemicals, solvents and reagents were purchased from Sigma-Aldrich unless otherwise noted. *E. coli* Bl21 T7 Express competent cells were purchased from New England Biolabs (NEB). LB medium components were purchased from Fisher BioReagents. Ampicillin and isopropyl-β-D-thiogalactopyranoside (IPTG) was purchased from GoldBio. The NaCl for purification buffer components was purchased from Fisher Chemical. The EDTA-free protease inhibitor cocktail tablets were purchased from Roche. The Ni-NTA 1 mL columns and Superdex200 16/60 size exclusion columns (SEC) were purchased from GE Healthcare. The gel filtration standards, polyacrylamide gels, sodium dodecyl sulfate solution and were purchased from BioRad. The original crystallization solution was purchased as one condition in the 96-Index HT screen from Hampton Research. The crystallization solutions used for optimizations were purchased individually from Hampton Research.

Methods

Cell growth, protein expression and purification of CmATR

Cell growth and purifications of *CmATR* were conducted according to the following protocol. An overnight starter culture of 25 mL LB supplemented with 100 mg/mL ampicillin was inoculated with a single colony of *E. coli* Bl21 T7 Express competent cells transformed with a pMCSG7 plasmid containing the *cobO* gene from *C. metallidurans* with an N-terminal his tag and grown at 37 ºC with shaking at 220 RPM. The overnight starter culture was used to inoculate 1 L of LB supplemented with 100 mg/mL ampicillin LB in a ratio of 2 mL starter culture to 1 L LB at 37 ºC with shaking at 220 RPM. The 1 L culture was induced with 0.1 mM IPTG when OD_{600} reached ~0.6 and grown for 18 h at 15 ºC with shaking at 220 RPM. Cells were harvested by centrifugation (5,000 x g, 20 min, 4 ºC) and flash frozen in liquid N_2 then stored in a -80 ºC freezer for future use.

Cells from 1 L of culture were resuspended in 45 mL lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 25 mM imidazole) supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF) and one-half of an EDTA-free protease inhibitor cocktail tablet. Cells were lysed by ultrasonication, and cell lysates were clarified by centrifugation (28,000 x g, 30 min, 4 ºC). Clarified lysate was passed through a 0.22 µm syringe filter and loaded onto a 1 mL Ni-NTA column equilibrated with lysis buffer. Protein was eluted with a linear

gradient of elution buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 500 mM imidazole) using an FPLC (Amersham Biosciences AKTA FPLC System). Elution fractions were concentrated in a 30 kDa MWCO centrifugal filter. The concentrated fractions of *CmATR* were loaded onto a Superdex 200 16/60 SEC equilibrated with SEC buffer (50 mM HEPES pH 7.5, 500 mM NaCl). Elution fractions from SEC were concentrated in a 30 kDa MWCO centrifugal filter. Purity was assessed by 4-20% (w/v) sodium dodecyl sulfate polyacrylamide gel electrophoresis. The concentration of *CmATR* monomer was determined to be 2.6 mM (53 mg/mL) by UV/Vis absorbance at 280 nm using an extinction coefficient of 15,595 $M^{-1}cm^{-1}$. Protein samples in SEC buffer were flash frozen in liquid N_2 and stored in the -80 ºC freezer for future use.

Crystallography

*Cm*ATR was buffer exchanged into 100 mM HEPES pH 7.5, 100 mM NaCl and concentrated using a 30 kDa MWCO centrifugal filter to 618 µM as determined by UV/Vis absorbance at 280 nm using an extinction coefficient of 15,595 $M^{-1}cm^{-1}$ and a molecular weight of 19.9 kDa. Initial crystallization conditions for *Cm*ATR were identified using crystallization screens dispensed by Mosquito liquid-handling robot (SPT Labtech) at room temperature. Initial crystallization was obtained in numerous conditions from the Hampton Research Index HT screen.

The identified crystallization conditions were further optimized and screened for diffraction. Apo-*Cm*ATR crystals with the best morphology and diffraction were obtained via the hanging-drop vapor diffusion method using a 1 µL of protein (9.8 mg/mL, 493 µM) to 1 µL of well solution (25% PEG 3350, 0.2 M lithium sulfate, and 0.1 M BIS-TRIS pH 5.5) in a 2 µL drop with a reservoir of 500 µL. Several stacks of thin crystal plates developed over two weeks. For data collection, the stacked crystals were separated, and the resulting crystals were stepped into well solution containing 20% glycerol as a cryoprotectant and flash frozen in liquid N_2.

Data collection, processing, structure determination, and refinement

A dataset for apo-*CmATR* was collected at the Advanced Photon Source (Argonne, Illinois, USA) on beamline 24ID-E using an Eiger-16M pixel array detector at a

temperature of 100 K. Data were collected at a wavelength of 0.979 Å in a single 360º wedge with 0.2º per image with 5% transmission.

The apo-*CmATR* crystal belongs to space group *C*2. The data were indexed, integrated, and scaled to a 2.35-Å resolution in XDS[36]. The structure was solved by molecular replacement with the Phenix implementation of Phaser using data trimmed to 2.35 Å resolution[37]. The Phenix implementation of Sculptor, which removed the sidechains of divergent sequences, was used to create a homology model of *CmATR* using the analogous ATR from *Burkholderia thailandensis* reported to 1.80 Å resolution (PDB 2ZHY) (75.27% identical)[23]. The final molecular replacement solution with LLG of 7918 and TFZ of 58.5 identified six monomers of apo-*CmATR* in the asymmetric unit.

With NCS restraints for each of the chains, the model was first refined with iterative rounds of simulated annealing to minimize model bias, positional refinement, and rigid body refinement using Phenix[38]. Between rounds of refinement, the model was manually adjusted using Coot[39]. Then the model was refined with more iterative rounds of simulated annealing and positional and individual B-factor refinements using Phenix[38]. Sides chains were added to residues with clear electron density. Waters were placed manually into regions with $2mF_o$-DF_c composite omit density, $2F_o$-F_c density, and 3 σ F_o-F_c density. Subsequent iterative rounds of model building and refinement were performed in Coot[39] and Phenix[38], respectively. The final model has a R_{work} of 23.7% and R_{free} of 25.2%. In chains A and B, residues Lys28-Gln179 of 184 residues total were modeled into the density. In chains C, D, E, and F, residues Asn29-Gln179 of 184 residues total were modeled into the density. There was no density observed for the affinity tags on any of the chains. The positions of the atoms were confirmed by comparison to the $2mF_o$-DF_c composite omit density. Crystallographic software packages were compiled by SBGrid (*Table II.2*)[40]. Structural figures were made in ChimeraX version 1.5[41].

Modeling of the ATR and IcmF complex

Alphafold2 multimer was first used to try to predict the structure of the complex consisting of *Cm*IcmF and *Cm*ATR[30]. Three copies of the sequence of *Cm*ATR and two copies of the sequence of *Cm*IcmF were inputted with all default parameters on the COSMIC^2 server[42]. The resulting five protein structure models were manually examined

in ChimeraX v1.5 for complex formation. No complex was predicted for the *Cm*ATR trimer with the *Cm*IcmF dimer (*Figure IV.6*). RosettaDocking4.0 was then utilized to computationally predict the structure of the complex of *Cm*IcmF and *Cm*ATR[43]. The two input structures were prepared manually in ChimeraX v1.5 by placing the apo-*Cm*ATR trimer near the open-active site protomer of the *Cm*IcmF dimer (Chapter III) (*Figure IV.7*). The only difference in the input models was the location of the apo-*Cm*ATR trimer relative to the Cbl-binding domain of *Cm*IcmF, with the positions of apo-*Cm*ATR on opposite sides of the Cbl-binding domains of *Cm*IcmF. The default parameters were used on all input models[31]. Over four hundred models were generated, and the resulting top ten output models were analyzed in ChimeraX v1.5 by aligning the *Hs*ATR structure with AdoCbl, PPP_i, and ATP bound (PDB 6D5X)[22] on the *Cm*ATR trimer in the output model[41]. The models output oriented the AdoCbl binding site of *Cm*ATR near the opening of the active site of the mutase domain of *Cm*IcmF. Structural figures were made in ChimeraX v1.5[41].

Figures

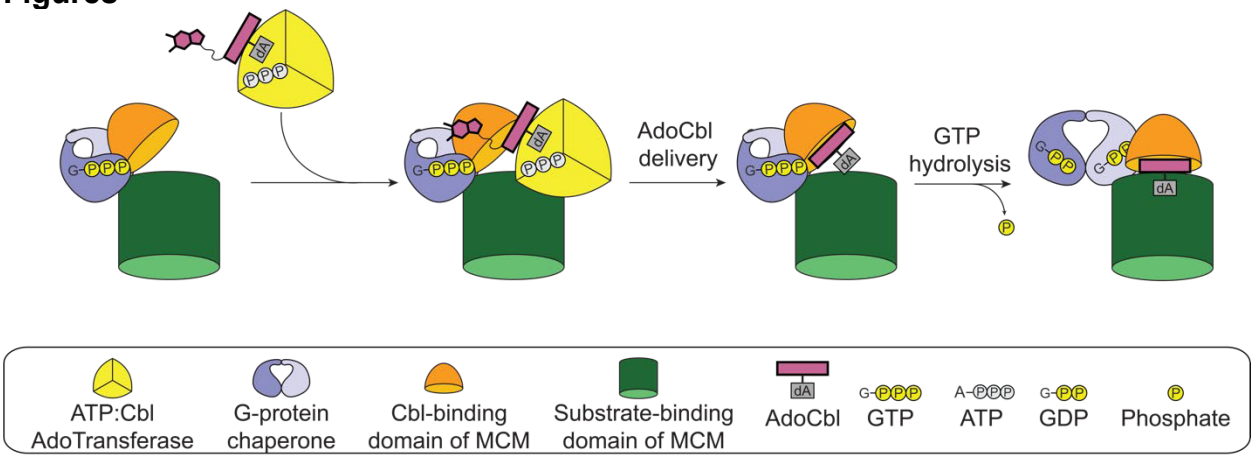


Figure IV.1. Proposed molecular mechanism showing final step in the delivery of AdoCbl to its target mutase. The G-protein chaperone prepares the active site of the mutase for AdoCbl delivery by wedging it open. The ATR prepares the AdoCbl and inserts it into the wedged open active site of the mutase. After GTP hydrolysis occurs, the conformation of the G-protein chaperone returns to the “inactive” state and the mutase closes, trapping the AdoCbl.

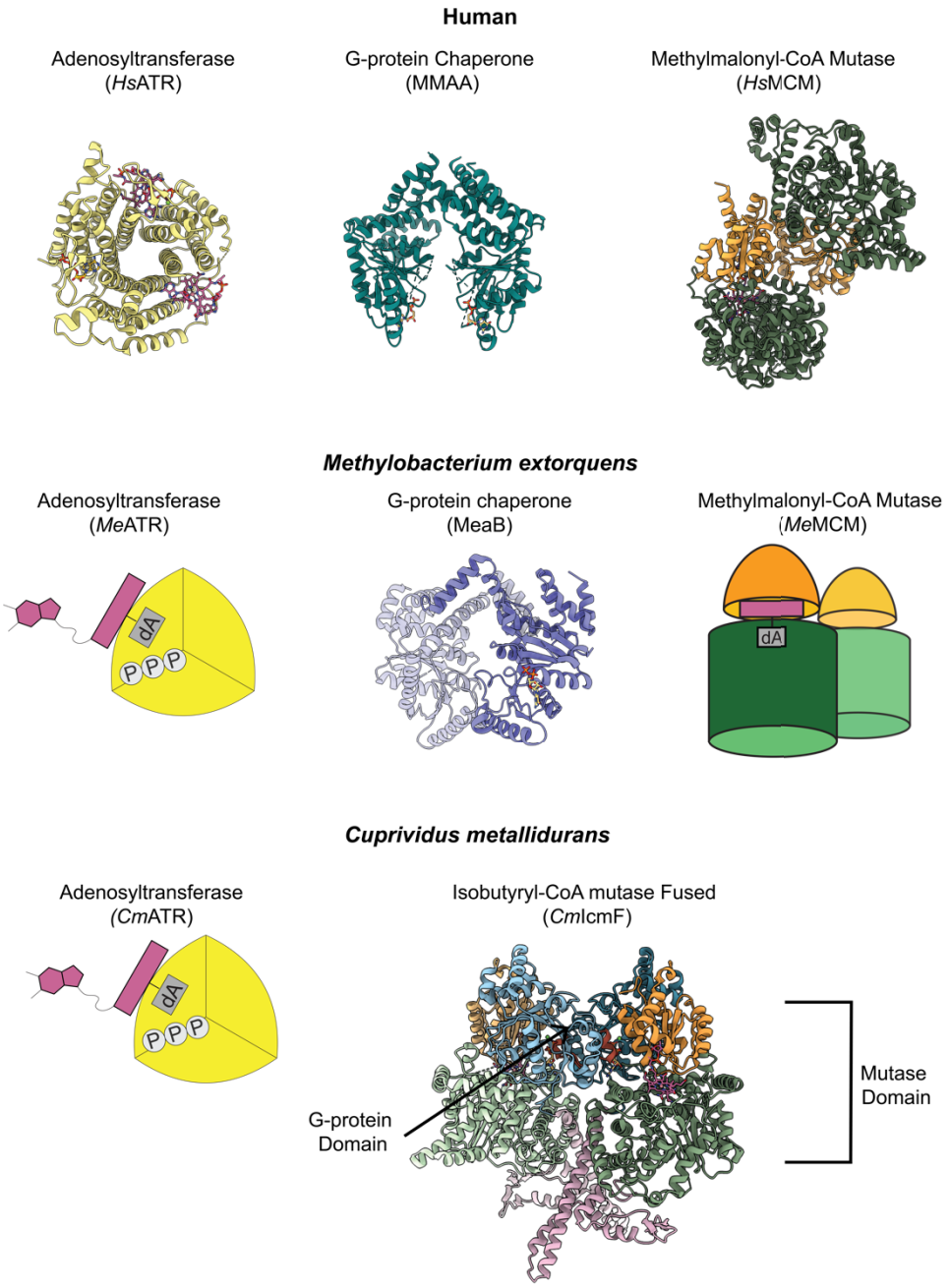


Figure IV.2. The different systems used to investigate the final delivery of AdoCbl to its target mutase. The human system requires three enzymes for the final delivery of AdoCbl to the target mutase: an ATR trimer (PDB 6XD5)[22] (yellow ribbons) that can bind AdoCbl (sticks, carbon in pink), AdoCbl and PPP_i (sticks colored by heteroatom) and ATP (sticks, carbon in yellow), a G-protein chaperone (PDB 2WWW)[44] (blue ribbons) that can bind various G-nucleotides such as GDP (sticks, carbon in yellow), and an AdoCbl-dependent methylmalonyl-CoA mutase (PDB 2XIQ)[44] that exists as a homodimer

(substrate-binding domain: green ribbons; Cbl-binding domain: orange ribbons) with AdoCbl bound (sticks, carbon in pink)[15]. The homologous system from *M. extorquens* also requires three enzymes for the final delivery of AdoCbl to its target: an ATR trimer (yellow cartoon, structurally uncharacterized) that binds and delivers AdoCbl (pink cartoon), a G-protein chaperone (purple ribbons) that can bind various G-nucleotides such as the nonhydrolyzable GTP analog, GMPPCP, (sticks, carbon in yellow) (PDB 8DPB)[45] and an AdoCbl-dependent methylmalonyl-CoA mutase that exists as heterodimer with one active and one inactive subunit (substrate binding domain: green cartoon; Cbl-binding domain: orange cartoon)[16, 46]. The final system to investigate the delivery of AdoCbl to its target mutase is the fused system from *C. metallidurans.* isobutyryl-CoA mutase fused (IcmF) that contains two domains: a mutase domain comprising of the Cbl-binding domain (orange ribbons) and the substrate binding domain (green ribbons) and the G-protein domain (blue ribbons) (Chapter III). The adenosyltransferase (yellow cartoon) delivers AdoCbl (pink cartoon) to IcmF. The structure of the ATR from *C. metallidurans* has not been previously reported, and the interfaces important for cofactor transfer have not been elucidated in any system.

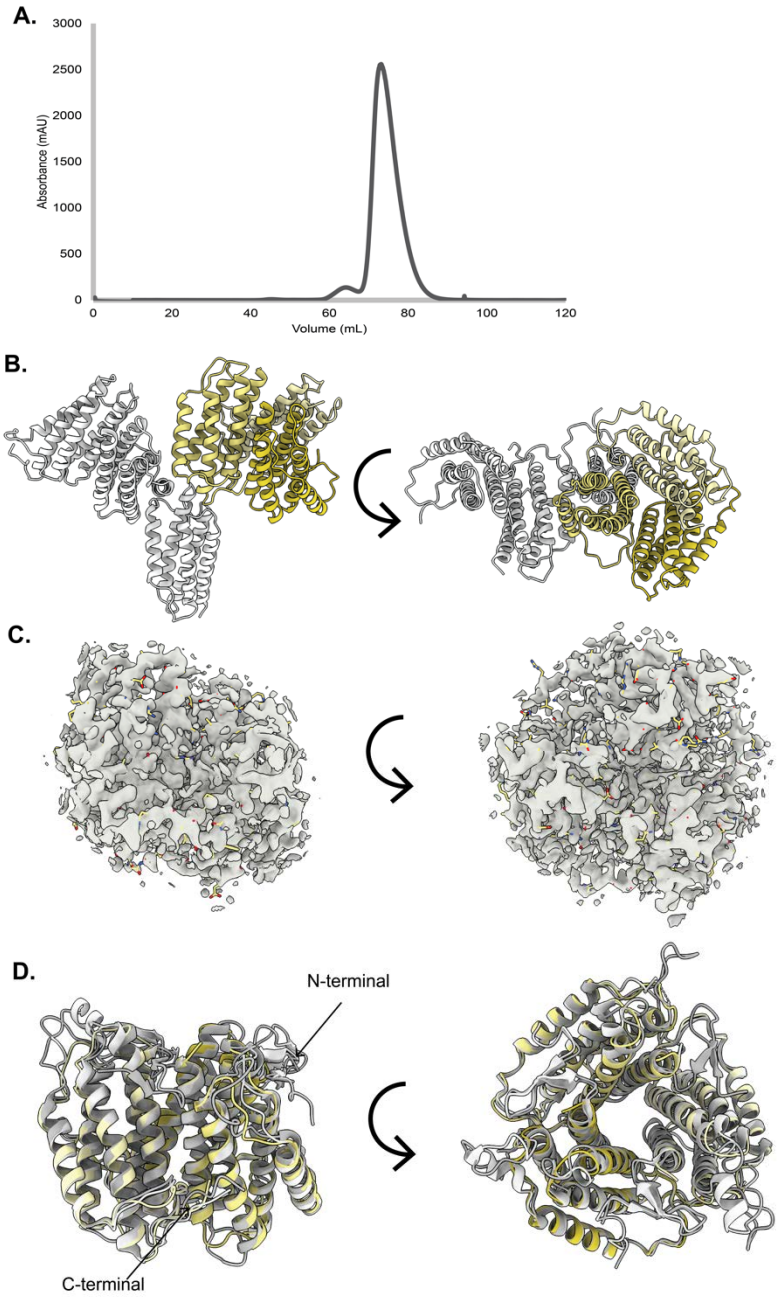


Figure IV.3. The structure of CmATR is analogous to other PduO-type ATRs. **A.** The S200 16/60 size exclusion chromatogram indicates a single species representing the formation of an apparent *Cm*ATR trimer in solution, consistent with other previously characterized PduO-type ATRs[19, 23, 24, 26]. **B.** The asymmetric unit of the *Cm*ATR structure contains six copies of the ATR protomer, each with a ferritin-like fold characteristic of this family of enzymes[33]. The relevant ATR trimer is colored in yellow and the other three ATR

protomers of the asymmetric unit are colored in light grey. **C**. The $2mF_o$-DF_c omit density contoured at 1 σ (grey surface) for the *Cm*ATR trimer (carbon in yellow). **D.** The alignment of apo-*Cm*ATR (yellow ribbons) on the trimer of ATR from *M. tuberculosis* with AdoCbl and PPP_i bound (light grey ribbons) (PDB 6WGV)[21] and from *H. sapiens* with AdoCbl, Cbl and ATP bound (grey ribbons) (PDB 6D5X)[22] indicates an overall good agreement, with differences localized mainly in the N- and C-terminal regions. The bound ligands are hidden for simplicity.

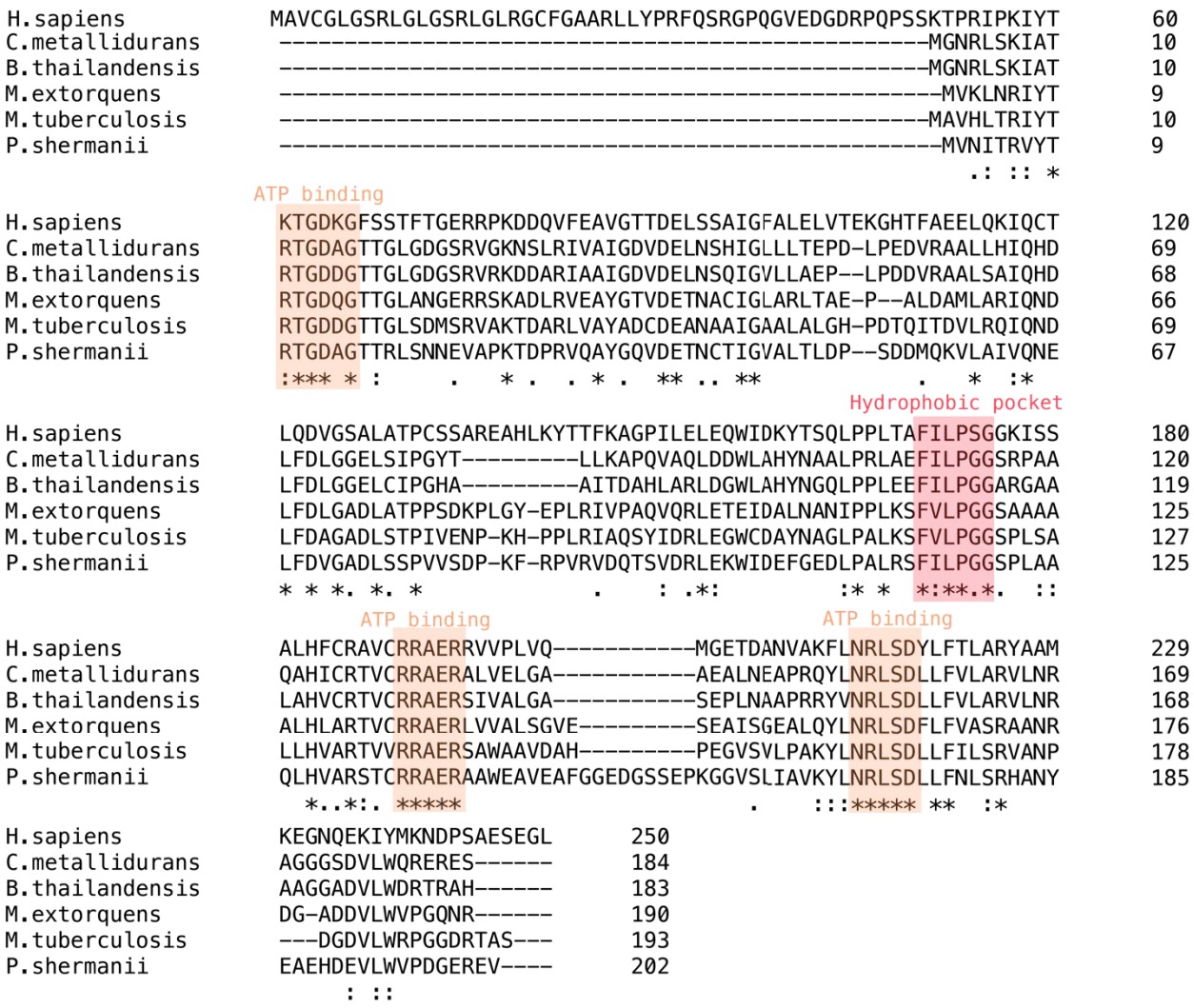


Figure IV.4. Sequence alignment of selected PduO-type adenosyltransferases. The sequence alignment of the PduO-type ATRs from *H. sapiens, C. metallidurans*, *B. thailandensis*, *M. extorquens*, *M. tuberculosis* and *P. shermanii* indicates numerous conserved regions. The regions that participate in binding of ATP, AdoCbl, and/or Cbl are highlighted in boxes. Orange boxes indicate residues involved in binding ATP, triphosphate, and/or adenosine. The red box indicates residues involved in creating the hydrophobic pocket to facilitate a four-coordinate cob(II)alamin species[11]. Asterisks (*) denote positions with conserved residues. Colons (:) denote positions of conservation with strongly similar properties. Periods (.) denote positions of conservation with weakly similar properties. Alignments performed using ClustalW [28].

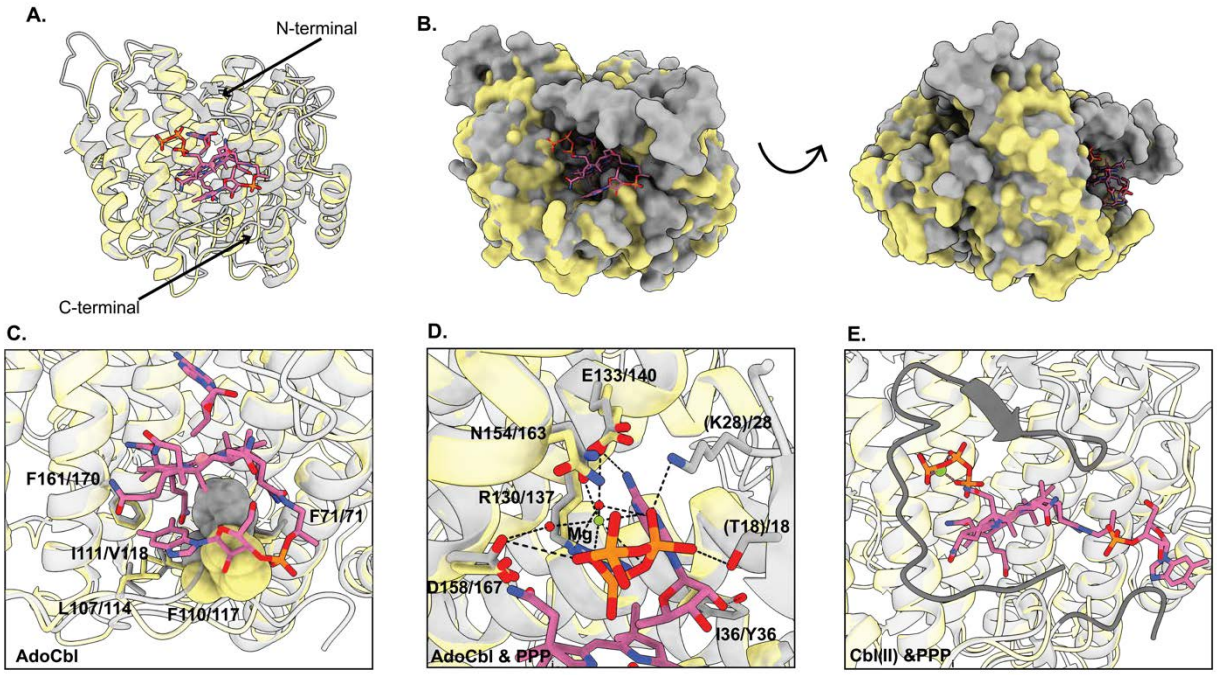


Figure IV.5. Structural comparison of CmATR to MtbATR suggests similar binding modes for AdoCbl. **A.** The alignment of the structure of apo-*Cm*ATR (yellow ribbons) on the structure of *Mtb*ATR (grey ribbons) with AdoCbl bound (sticks, carbon in pink) and PPP_i (sticks, colored by heteroatom) (PDB 6WGV)[21] indicates the predicted binding site of AdoCbl for *Cm*ATR. The alignment indicates that the N-terminal residues are involved in forming the binding site but are disordered in the *Cm*ATR structure. **B.** The AdoCbl (sticks, carbon in pink) binding site of *Mtb*ATR (grey surface) (PDB 6WGV)[21] aligned on *Cm*ATR (yellow surface) is surface exposed but tucked into a recessed pocket. **C.** A close view of the alignment of *Mtb*ATR (grey ribbons) bound to AdoCbl (sticks, carbon in pink) (PDB 6WGS)[21] on *Cm*ATR (yellow ribbons) to show the hydrophobic pocket underneath the corrin ring that is formed by the C-terminal residues of *Mtb*ATR (grey sticks). Equivalent residues predicted to interact with AdoCbl from *Cm*ATR are shown as yellow sticks. The conserved Phe (110 in *Cm*ATR; 117 in *Mtb*ATR) is shown as space-filling spheres to highlight how it prevents the DMB tail from coordinating the Cbl cobalt. The residue numbers are indicated for the *Cm*ATR and the *Mtb*ATR sequences, respectively. **D.** A close view of the PPP_i (sticks colored by heteroatom), Mg^{2+} ions (green spheres) and AdoCbl (sticks, carbon in pink) in binding site of *Mtb*ATR (grey ribbons) aligned to *Cm*ATR (yellow ribbons). The similarity between bound and unbound states suggests that

the binding pockets are pre-formed. Equivalent residues predicted to interact with PPP_i and AdoCbl from *Cm*ATR are shown as yellow sticks. The residue numbers are indicated for the *Cm*ATR and the *Mtb*ATR sequences, respectively, with disordered residues from *Cm*ATR indicated by parentheses. **E.** The binding site of *Mtb*ATR (grey ribbons) with PPP_i (sticks colored by heteroatom) and Cbl(II) (sticks, carbon in pink) bound (PDB 6WH5)[21] aligned on *Cm*ATR (yellow ribbons) shows an extended conformation of the DMB tail which is displaced by the N-terminal residues and stabilized by the C-terminal residues (dark grey ribbons).

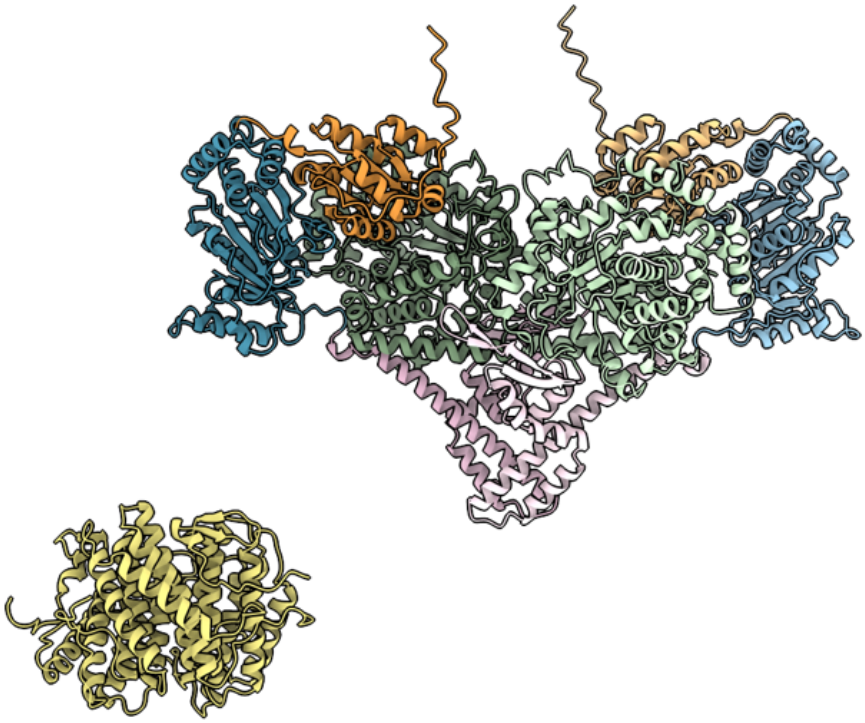


Figure IV.6. The alphafold2 multimer model did not predict the complex between the ATR and IcmF. The models computed by alphafold2 multimer[30] folded the individual proteins correctly with the correct oligomeric states. The *Cm*ATR (yellow ribbons) was computed to be a trimer. The *Cm*IcmF (G-protein domain: blue ribbons; Cbl-binding domain: orange ribbons; substrate-binding domain: green ribbons; linker: pink ribbons) was computed to be a dimer with a substrate-binding domain interface.

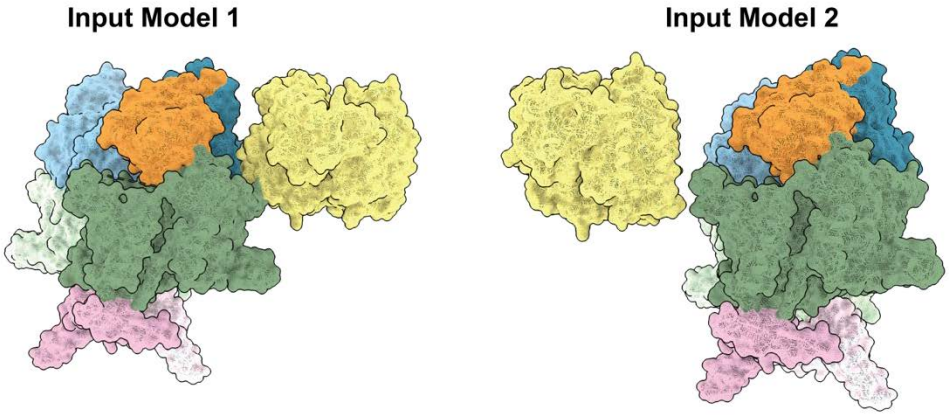


Figure IV.7. The input models of the IcmF:ATR used for RosettaDocking4.0. The manually prepared input models of *Cm*IcmF dimer in the open conformation of the active site (Chapter III) (G-protein domain: blue surface; Cbl-binding domain: orange surface; substrate-binding domain: green surface; linker: pink surface) and apo-*Cm*ATR trimer (yellow surface).

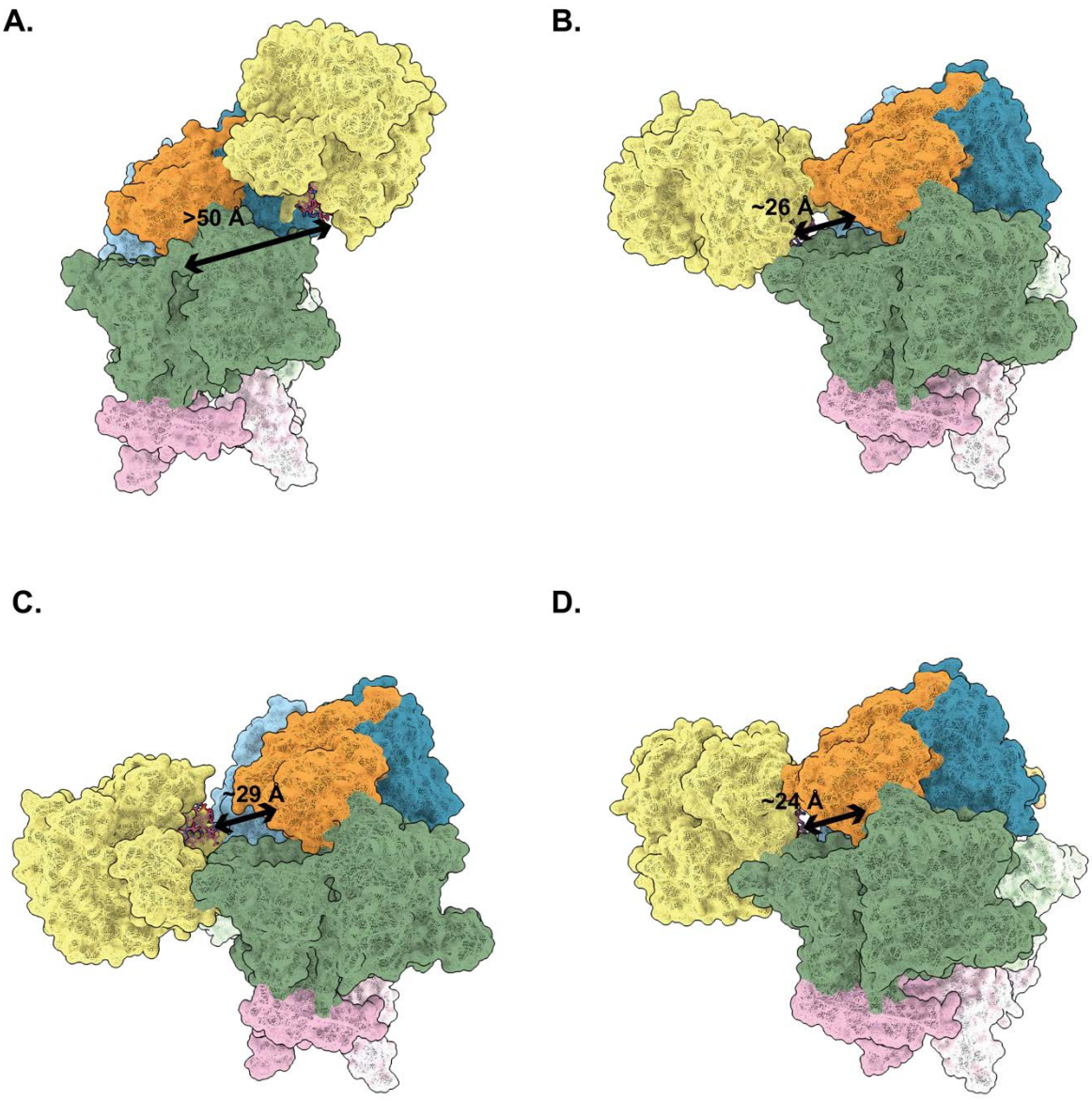


Figure IV.8. **Four docking solutions for *Cm*IcmF-*Cm*ATR that were generated in RosettaDocking4.0 show considerable variation**. **A.** The *Cm*IcmF-*Cm*ATR docking solution that has the largest buried surface area between ATR and G-protein dimer region is shown. In this model, the AdoCbl binding site of ATR more than 50 Å away from the binding site of Cbl-binding domain of IcmF. *Cm*IcmF dimer is shown in the open conformation of the active site (Chapter III) (G-protein domain: blue surface; Cbl-binding domain: orange surface; substrate-binding domain: green surface; linker: pink surface). Apo-*Cm*ATR trimer (yellow surface) is shown with one molecule of AdoCbl (sticks, carbon in pink) modeled into the binding site[43]. **B.** *Cm*IcmF-*Cm*ATR docking solution that has

fewer contacts between ATR and G-protein domain than in (A) and a closer distance between AdoCbl on ATR and Cbl-binding domain of IcmF. **C.** *Cm*IcmF-*Cm*ATR docking solution that is similar to that shown in (B); AdoCbl on ATR is more accessible for transfer to the Cbl-binding domain of IcmF, but the distance is farther. **D**. *Cm*IcmF-*Cm*ATR docking solution that is similar to that shown in (B and C), but the orientation of the Cbl-binding site of ATR is positioned away from the mutase active site compared to the docking solution shown in (C).

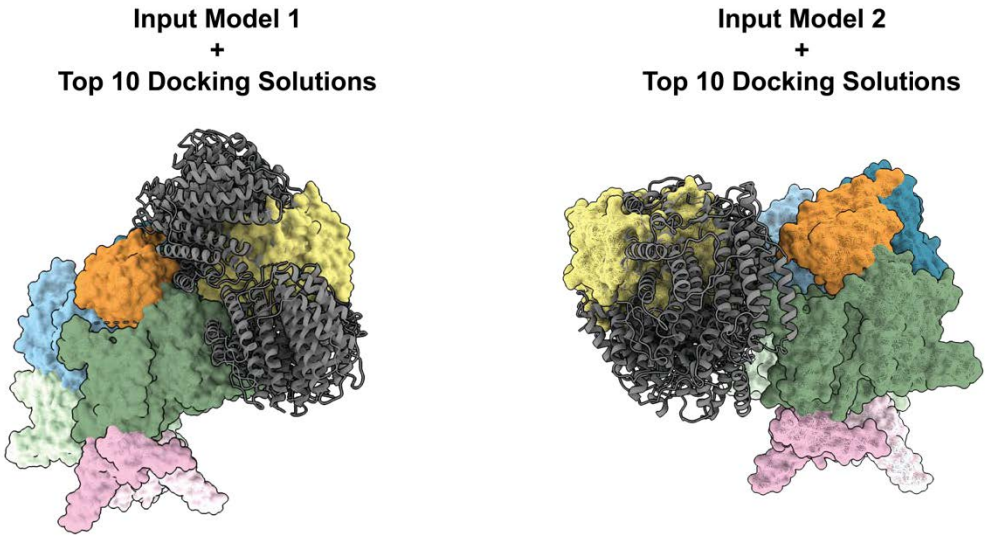


Figure IV.9. The input models bias the final docking solutions for RosettaDocking4.0. The manually prepared input models of *Cm*IcmF dimer in the open conformation (Chapter III) (G-protein domain: blue surface; Cbl-binding domain: orange surface; substrate-binding domain: green surface; linker: pink surface) and apo-*Cm*ATR trimer (yellow surface) with the top ten docking solutions shown as grey ribbons.

Tables

Table IV.1. Summary of the structures of PduO-type adenosyltransferases deposited in the PDB

Organism	Ligands Bound to Trimer			PDB ID
	Site 1	**Site 2**	**Site 3**	
Homo sapiens	ATP	ATP		2IDX[26]
	AdoCbl	AdoCbl, PPP_i	ATP	6D5K[22]
Lactobacillus reuteri	ATP	ATP	ATP	2NT8[24]
	Cob(II)inamide, ATP	Cob(II)inamide, ATP	Cob(II)inamide, ATP	3CI4[11]
	Cob(II)alamin, ATP	Cob(II)alamin, ATP	Cob(II)alamin, ATP	3CI1[11]
	AdoCbl, PPP_i	AdoCbl, PPP_i	AdoCbl, PPP_i	3CI3[11]
Burkholderia thailandensis	n/a			2ZHY[23]
	ATP	ATP		2ZHZ[23]
Bacillus cereus	n/a			3KE4*
	ATP	ATP	ATP	3KE5*
Mycobacterium tuberculosis	n/a			6WGU[21], 2G2D*
	AdoCbl	AdoCbl	AdoCbl	6WGS[21]
	AdoCbl, PPP_i	AdoCbl, PPP_i	AdoCbl, PPP_i	6WGV[21]
	Cob(II)alamin, PPP_i	Cob(II)alamin, PPP_i	Cob(II)alamin, PPP_i	6WH5[21]
Bacillus subtilis	n/a			1RTY[25]
Sulfolobus tokadaii ST2180	n/a			1WVT[47]
Pyrococcus horikoshi OT PH0671	n/a			1WY1*
Thermoplasma acidophilum	n/a			1NOG[33]
Cupriavidus metallidurans	n/a			This study

*No primary literature associated with the PDB deposition.

Table IV.2. Data collection and refinement statistics for apo-*Cm*ATR

Beamline	APS 24-ID-C
Wavelength (Å)	0.979
Space group	C2
Cell dimensions	
a, b, c (Å)	158.9, 91.7, 102.8
α, β, γ (°)	90, 119.57, 90
Resolution† (Å)	45.86-2.35 (2.39-2.35)
R_{meas}† (%)	13.2 (68.3)
$\langle I \rangle / \langle \sigma I \rangle$†	8.2 (2.2)
Completeness† (%)	96.8 (96.4)
Redundancy†	5.53 (5.23)
CC1/2†	99.6 (85.9)
No. Total Reflections†	287267 (43231)
No. Unique Reflections†	51913 (8267)
Refinement	
Resolution (Å)	45.86-2.35
No. Reflections used	51834
R_{work}/R_{free}‡	23.7/25.2
No. molecules in asymmetric unit	6
No. atoms	
Protein	7003
Water	65
B-factors (Å^2)	
Protein	
Chain A	57.4
Chain B	57.3
Chain C	57.0
Chain D	57.3
Chain E	57.1
Chain F	57.3
Water	55.0
RMS deviations	
Bond lengths (Å)	0.008
Bond angles (°)	0.908
Rotamer outliers (%)	0.97
Ramachandran plot (%)	
Most favored	97.0
Additionally allowed	2.2
Disallowed	0.8

† Values in parentheses are for the highest resolution shell.
‡5% of reflections were set aside for a test set

References

(1) Capdevila, Daiana A.; Edmonds, Katherine A.; Giedroc, David P. Metallochaperones and metalloregulation in bacteria. *Essays In Biochemistry* **2017**, *61* (2), 177-200. DOI: 10.1042/ebc20160076.

(2) Hausinger, R. P. New metal cofactors and recent metallocofactor insights. *Current Opinion in Structural Biology* **2019**, *59*, 1-8. DOI: 10.1016/j.sbi.2018.12.008.

(3) Banerjee, R.; Gherasim, C.; Padovani, D. The tinker, tailor, soldier in intracellular B_{12} trafficking. *Current Opinion in Chemical Biology* **2009**, *13* (4), 484-491. DOI: 10.1016/j.cbpa.2009.07.007.

(4) Banerjee, R. V.; Matthews, R. G. Cobalamin-dependent methionine synthase. *The FASEB journal* **1990**, *4* (5), 1450-1459.

(5) Banerjee, R. Radical carbon skeleton rearrangements: Catalysis by coenzyme B_{12} -dependent mutases. *Chemical Reviews* **2003**, *103* (6), 2083-2094. DOI: 10.1021/cr0204395.

(6) Banerjee, R. B_{12} trafficking in mammals: A for coenzyme escort service. *ACS chemical biology* **2006**, *1* (3), 149-159. DOI: 10.1021/cb6001174.

(7) Willard, H. F.; Mellman, I. S.; Rosenberg, L. E. Genetic complementation among inherited deficiencies of methylmalonyl-CoA mutase activity: evidence for a new class of human cobalamin mutant. *American journal of human genetics* **1978**, *30* (1), 1-13.

(8) Rosenblatt, D. S.; Cooper, B. A. Inherited disorders of vitamin B12 utilization. *Bioessays* **1990**, *12* (7), 331-334.

(9) Leal, N. A.; Olteanu, H.; Banerjee, R.; Bobik, T. A. Human ATP:Cob(I)alamin Adenosyltransferase and Its Interaction with Methionine Synthase Reductase *Journal of Biological Chemistry* **2004**, *279* (46), 47536-47542. DOI: 10.1074/jbc.M405449200 (acccessed 2022/12/14).

(10) Stich, T. A.; Yamanishi, M.; Banerjee, R.; Brunold, T. C. Spectroscopic Evidence for the Formation of a Four-Coordinate Co^{2+} Cobalamin Species upon Binding to the Human ATP:Cobalamin Adenosyltransferase. *Journal of the American Chemical Society* **2005**, *127* (21), 7660-7661. DOI: 10.1021/ja050546r.

(11) St. Maurice, M.; Mera, P.; Park, K.; Brunold, T. C.; Escalante-Semerena, J. C.; Rayment, I. Structural Characterization of a Human-Type Corrinoid Adenosyltransferase Confirms That Coenzyme B_{12} Is Synthesized through a Four-Coordinate Intermediate. *Biochemistry* **2008**, *47* (21), 5755-5766. DOI: 10.1021/bi800132d.

(12) Park, K.; Mera, P. E.; Escalante-Semerena, J. C.; Brunold, T. C. Spectroscopic Characterization of Active-Site Variants of the PduO-type ATP:Corrinoid Adenosyltransferase from Lactobacillus reuteri: Insights into the Mechanism of Four-

Coordinate Co(II)corrinoid Formation. *Inorganic Chemistry* **2012**, *51* (8), 4482-4494. DOI: 10.1021/ic202096x.

(13) Padovani, D.; Banerjee, R. A G-protein editor gates coenzyme B_{12} loading and is corrupted in methylmalonic aciduria. *Proceedings of the National Academy of Sciences* **2009**, *106* (51), 21567-21572. DOI: 10.1073/pnas.0908106106.

(14) Takahashi-Iñiguez, T.; González-Noriega, A.; Michalak, C.; Flores, M. E. Human MMAA induces the release of inactive cofactor and restores methylmalonyl-CoA mutase activity through their complex formation. *Biochimie* **2017**, *142*, 191-196. DOI: 10.1016/j.biochi.2017.09.012.

(15) Takahashi-Íñiguez, T.; García-Arellano, H.; Trujillo-Roldán, M. A.; Flores, M. E. Protection and reactivation of human methylmalonyl-CoA mutase by MMAA protein. *Biochemical and Biophysical Research Communications* **2011**, *404* (1), 443-447. DOI: 10.1016/j.bbrc.2010.11.141.

(16) Padovani, D.; Banerjee, R. Assembly and protection of the radical enzyme, methylmalonyl-CoA mutase, by its chaperone. *Biochemistry* **2006**, *45* (30), 9300-9306. DOI: 10.1021/bi0604532.

(17) Cracan, V.; Banerjee, R. Novel coenzyme B_{12}-dependent interconversion of isovaleryl-CoA and pivalyl-CoA. *Journal of Biological Chemistry* **2012**, *287* (6), 3723-3732. DOI: 10.1074/jbc.M111.320051.

(18) Jost, M.; Cracan, V.; Hubbard, P. A.; Banerjee, R.; Drennan, C. L. Visualization of a radical B_{12} enzyme with its G-protein chaperone. *Proceedings of the National Academy of Sciences* **2015**, *112* (8), 2419-2424.

(19) Li, Z.; Kitanishi, K.; Twahir, U. T.; Cracan, V.; Chapman, D.; Warncke, K.; Banerjee, R. Cofactor Editing by the G-protein Metallochaperone Domain Regulates the Radical B12 Enzyme IcmF*♦. *Journal of Biological Chemistry* **2017**, *292* (10), 3977-3987. DOI: 10.1074/jbc.M117.775957.

(20) Padovani, D.; Banerjee, R. A Rotary Mechanism for Coenzyme B_{12} Synthesis by Adenosyltransferase. *Biochemistry* **2009**, *48* (23), 5350-5357. DOI: 10.1021/bi900454s.

(21) Mascarenhas, R.; Ruetz, M.; McDevitt, L.; Koutmos, M.; Banerjee, R. Mobile loop dynamics in adenosyltransferase control binding and reactivity of coenzyme B_{12}. *Proceedings of the National Academy of Sciences* **2020**, *117* (48), 30412-30422. DOI: doi:10.1073/pnas.2007332117.

(22) Campanello, G. C.; Ruetz, M.; Dodge, G. J.; Gouda, H.; Gupta, A.; Twahir, U. T.; Killian, M. M.; Watkins, D.; Rosenblatt, D. S.; Brunold, T. C.; et al. Sacrificial Cobalt-Carbon Bond Homolysis in Coenzyme B_{12} as a Cofactor Conservation Strategy. *Journal of the American Chemical Society* **2018**, *140* (41), 13205-13208. DOI: 10.1021/jacs.8b08659.

(23) Moon, J. H.; Park, A. K.; Jang, E. H.; Kim, H. S.; Chi, Y. M. Crystal structure of a PduO-type ATP:cobalamin adenosyltransferase from Burkholderia thailandensis. *Proteins: Structure, Function, and Bioinformatics* **2008**, *72* (3), 1066-1070. DOI: 10.1002/prot.22084.

(24) Maurice, M. S.; Mera, P. E.; Taranto, M. P.; Sesma, F.; Escalante-Semerena, J. C.; Rayment, I. Structural Characterization of the Active Site of the PduO-Type ATP:Co(I)rrinoid Adenosyltransferase from *Lactobacillus reuteri*. *Journal of Biological Chemistry* **2007**, *282* (4), 2596-2605. DOI: 10.1074/jbc.M609557200 (acccessed 2022/12/13).

(25) Forouhar, F.; Kuzin, A.; Seetharaman, J.; Lee, I.; Zhou, W.; Abashidze, M.; Chen, Y.; Yong, W.; Janjua, H.; Fang, Y.; et al. Functional insights from structural genomics. *Journal of Structural and Functional Genomics* **2007**, *8* (2), 37-44. DOI: 10.1007/s10969-007-9018-3.

(26) Schubert, H. L.; Hill, C. P. Structure of ATP-Bound Human ATP:Cobalamin Adenosyltransferase. *Biochemistry* **2006**, *45* (51), 15188-15196. DOI: 10.1021/bi061396f.

(27) Saridakis, V.; Yakunin, A.; Xu, X.; Anandakumar, P.; Pennycooke, M.; Gu, J.; Cheung, F.; Lew, J. M.; Sanishvili, R.; Joachimiak, A.; et al. The Structural Basis for Methylmalonic Aciduria: The Crystal Structure of Archaeal ATP:Cobalamin Adenosyltransferase. *Journal of Biological Chemistry* **2004**, *279* (22), 23646-23653. DOI: 10.1074/jbc.M401395200 (acccessed 2022/12/13).

(28) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **2011**, *7* (1), 539. DOI: 10.1038/msb.2011.75.

(29) Padovani, D.; Labunska, T.; Palfey, B. A.; Ballou, D. P.; Banerjee, R. Adenosyltransferase tailors and delivers coenzyme B_{12}. *Nature Chemical Biology* **2008**, *4* (3), 194.

(30) Evans, R.; O'Neill, M.; Pritzel, A.; Antropova, N.; Senior, A.; Green, T.; Žídek, A.; Bates, R.; Blackwell, S.; Yim, J. Protein complex prediction with AlphaFold-Multimer. *BioRxiv* **2022**, 2021.2010. 2004.463034.

(31) Lyskov, S.; Chou, F.-C.; Conchúir, S. Ó.; Der, B. S.; Drew, K.; Kuroda, D.; Xu, J.; Weitzner, B. D.; Renfrew, P. D.; Sripakdeevong, P.; et al. Serverification of Molecular Modeling Applications: The Rosetta Online Server That Includes Everyone (ROSIE). *PLOS ONE* **2013**, *8* (5), e63906. DOI: 10.1371/journal.pone.0063906.

(32) Yamanishi, M.; Vlasie, M.; Banerjee, R. Adenosyltransferase: An enzyme and an escort for coenzyme B_{12}? **2005**, *30*, 304-308. DOI: 10.1016/j.tibs.2005.04.008.

(33) Saridakis, V.; Yakunin, A.; Xu, X.; Anandakumar, P.; Pennycooke, M.; Gu, J.; Cheung, F.; Lew, J. M.; Sanishvili, R.; Joachimiak, A.; et al. The Structural Basis for Methylmalonic Aciduria: THE CRYSTAL STRUCTURE OF ARCHAEAL ATP:COBALAMIN ADENOSYLTRANSFERASE *. *Journal of Biological Chemistry* **2004**, *279* (22), 23646-23653. DOI: 10.1074/jbc.M401395200.

(34) Kung, Y.; Ando, N.; Doukov, T. I.; Blasiak, L. C.; Bender, G.; Seravalli, J.; Ragsdale, S. W.; Drennan, C. L. Visualizing molecular juggling within a B_{12}-dependent methyltransferase complex. *Nature* **2012**, *484* (7393), 265-269. DOI: 10.1038/nature10916.

(35) Berkovitch, F.; Nicolet, Y.; Wan, J. T.; Jarrett, J. T.; Drennan, C. L. Crystal Structure of Biotin Synthase, an S-Adenosylmethionine-Dependent Radical Enzyme. *Science* **2004**, *303* (5654), 76-79. DOI: 10.1126/science.1088493.

(36) Kabsch, W. Xds. *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66* (2), 125-132.

(37) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *Journal of Applied Crystallography* **2007**, *40* (4), 658-674.

(38) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66* (2), 213-221.

(39) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66* (4), 486-501. DOI: 10.1107/s0907444910007493.

(40) Morin, A.; Eisenbraun, B.; Key, J.; Sanschagrin, P. C.; Timony, M. A.; Ottaviano, M.; Sliz, P. Collaboration gets the most out of software. *elife* **2013**, *2*, e01456.

(41) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Meng, E. C.; Couch, G. S.; Croll, T. I.; Morris, J. H.; Ferrin, T. E. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science* **2021**, *30* (1), 70-82. DOI: 10.1002/pro.3943.

(42) Cianfrocco, M. A.; Wong-Barnum, M.; Youn, C.; Wagner, R.; Leschziner, A. COSMIC2: A science gateway for cryo-electron microscopy structure determination. In Proceedings of the Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact, New Orleans, LA, USA; 2017.

(43) Lyskov, S.; Gray, J. J. The RosettaDock server for local protein–protein docking. *Nucleic Acids Research* **2008**, *36*, W233-W238. DOI: 10.1093/nar/gkn216.

(44) Froese, D. S.; Kochan, G.; Muniz, J. R.; Wu, X.; Gileadi, C.; Ugochukwu, E.; Krysztofinska, E.; Gravel, R. A.; Oppermann, U.; Yue, W. W. Structures of the human GTPase MMAA and vitamin B_{12}-dependent methylmalonyl-CoA mutase and insight into their complex formation. *Journal of Biological Chemistry* **2010**, *285* (49), 38204-38213.

(45) Vaccaro, F. A.; Born, D. A.; Drennan, C. L. Structure of metallochaperone in complex with the cobalamin-binding domain of its target mutase provides insight into cofactor delivery. *Proceedings of the National Academy of Sciences* **2023**, *120* (8), e2214085120. DOI: doi:10.1073/pnas.2214085120.

(46) Padovani, D.; Labunska, T.; Banerjee, R. Energetics of interaction between the G-protein chaperone, MeaB, and B_{12}-dependent methylmalonyl-CoA mutase. *Journal of Biological Chemistry* **2006**, *281* (26), 17838-17844. DOI: 10.1074/jbc.M600047200.

(47) Tanaka, Y.; Sasaki, T.; Kumagai, I.; Yasutake, Y.; Yao, M.; Tanaka, I.; Tsumoto, K. Molecular properties of two proteins homologous to PduO-type ATP:cob(I)alamin adenosyltransferase from Sulfolobus tokodaii. *Proteins: Structure, Function, and Bioinformatics* **2007**, *68* (2), 446-457. DOI: 10.1002/prot.21303.

Chapter V: Concluding thoughts on and future directions for the final step of adensosylcobalamin delivery to its target mutase

Author Contributions

Chapter V was written by Francesca A. Vaccaro, and Catherine L. Drennan contributed to the editing of the chapter.

Closing thoughts and future directions

An expansive variety of metallocofactors leads to diverse metallochaperone mechanisms that protect the metallocofactors during transport and delivery and simultaneously minimize toxicity from unwanted off-target reactions[1-3]. Chapter I of this thesis focuses on one family of metallochaperones that relies on nucleoside triphosphates for their function: the NTPases[4]. Despite the similarities in nucleotide requirements, these metallochaperones have diverse targets. Therefore, precise molecular mechanisms are needed to maintain the necessary specificity in a cell with numerous competing functional needs. This thesis focuses on a structural approach to understanding the subfamily of the GTPase metallochaperones known as the MeaB subfamily[5]. These GTPases are involved in the delivery of adenosylcobalamin (AdoCbl) to mutases involved in carbon skeleton rearrangements. In humans, disruptions to this process inactivate methylmalonyl-CoA mutase (MCM) and lead to a disease phenotype known as methylmalonic aciduria[6]. Methylmalonic aciduria has numerous complementation groups that are associated with mutations or deletions in the different proteins involved in the maturation of MCM; these different complementation groups lead to varying levels of severity of the disease, from physical and mental impairment to death. Therefore, understanding the molecular basis of each player's role in the maturation would further inform treatments[7-9]. Additionally, a deeper understanding of the role of metallochaperones, and the various mechanisms that these proteins employ, can inform not only how maturation of metalloproteins occurs in vivo but also allow for the manipulations of these systems in vitro for various applications.

The fundamental underpinning the work in this thesis centers around how a G-protein chaperone is involved in the proper maturation of AdoCbl-dependent mutases without ever binding the Cbl cofactor. This role varies from the other members of the GTPase subfamily that directly bind the desired cofactor as part of their mechanism[10-12]. Answering this question has implications for understanding the molecular basis of methylmalonic aciduria and therefore will influence the development of treatments for this disease. Both the homodimeric human MCM and the corresponding MeaB G-protein, methylmalonic aciduria type A (MMAA), also a homodimer, have been structural characterized independently[13]. However, our initial attempts to obtain a structure of a

MMAA-MCM complex using the human enzymes was less successful than for the bacterial counterparts. Therefore, this work has relied on the structural characterizations of two bacterial model systems to investigate the molecular role of the G-protein in the maturation of the AdoCbl-dependent mutase. The first model system employed was from *Methylobacterium extorquens* and utilized a G-protein homodimer and an MCM heterodimer with one active and one inactive subunit. The second model system was from *Cupriavidius metallidurans* and is known as isobutyryl-CoA mutase fused (IcmF). IcmF contains the G-protein chaperone and the mutase on the same polypeptide chain, a fused system analogous to the two-component human and *M. extorquens* systems[14].

This work began by first structurally characterizing a minimal system from *M. extorquens* consisting of the G-protein, MeaB, and the Cbl-binding domain of MCM (*Me*MCM$_{Cbl}$). The minimal system was used because previous attempts to capture the complex with the full-length mutase from human or *M. extorquens* failed. Additionally, the structural characterization of IcmF failed to reveal an intact active site that was capable of catalyzing GTP hydrolysis, and was unable to link GTP binding and hydrolysis to a molecular change that could explain the G-protein domains' function[15]. The structure also failed to explain the significance of the switch III motif that was identified as harboring disease causative mutations in patients with methylmalonic aciduria and biochemically characterized to be important GTP hydrolysis in the non-fused systems[16-18]. The minimal system described in Chapter II provided the first snapshot of an AdoCbl metallochaperone in an active state; the GTP active site had the non-hydrolyzable GTP analog, GMPPCP, bound and was fully formed, and the function of switch III residues could be visualized. This active state of MeaB was formed through a 180° movement of one protomer of the G-protein dimer, creating an interdimer active site with switch III residue of one protomer contacting the G-nucleotide of the other protomer. This MeaB active state was proposed in Chapter II to prop open the active site of the mutase for AdoCbl delivery, but validation of this proposal required more structural data since the minimal system described in Chapter II lacked a complete mutase.

A deeper structural understanding of the role of the G-protein chaperone led to the revisitation of the fused system, IcmF. The asymmetric unit of the crystal structure of IcmF did not explain the role of the switch III residues, indicating that either switch III is not

important in the fused system, which we now know is not the case, or that we were missing a structural snapshot, which we now know is the case[15]. A hint as to the missing snapshot was discovered as a lattice contact. We noticed the presence of a G-protein dimer formed through a lattice contact, which suggested the potential for the formation of a G-protein dimer in solution (see *Figure III.16C*)[15]. Chapter III presents the solution state biochemical data and electron microscopy (EM) analysis that supports the relevance of a G-protein dimer in the fused system. Biochemical analysis indicates the importance of the switch III residues for GTP hydrolysis, consistent with the non-fused systems. The cryogenic (cryo) EM reconstructions of IcmF in the presence of GTP or a non-hydrolyzable analog reveal the same conformation of the G-protein chaperone domains as seen in the active state of MeaB. Analysis of the IcmF structures confirm the proposed role of the G-protein chaperone dimer in propping open the active site of the mutase. Interestingly, the switch III region, in addition to interacting with the nucleotide bound in the opposite chain of IcmF, also appears to have a role in propping open the mutase active site, firmly cementing its importance for the G-protein's role.

The work presented here provides structural snapshots that clarify our understanding of the G-protein:mutase interactions and allows for the demonstration of a consensus mechanism for the G-protein in the fused and non-fused system. This work has benefitted greatly from the advances in cryo-EM, allowing for the determination of structures that were too transient to be captured previously by crystallography. Future research can utilize cryo-EM to continue to trap complexes that will provide insight into the molecular mechanisms of AdoCbl delivery and damaged Cbl removal from its target mutases. Although obtainment of high resolution structures of human G-protein:mutase complexes has historically been challenging in our hands and in the hands of others[16], I now have preliminary negative stain EM data that indicates that these complexes can be captured by EM characterization (*Figure AI.1*). Cryo-EM analyses of the distinct complexes that are formed for loading and unloading of Cbl will allow for protein-protein interaction surfaces to be characterized. High resolution or even modest resolution cryo-EM data should inform as to whether MMAA, like the G-protein domain of IcmF, props open the active site of the mutase to allow the adenosyltransferase (ATR) to insert and/or remove a Cbl cofactor.

Collectively the work described in Chapter II and III has allowed for the elucidation of the structure-function relationship of the G-protein chaperone, has explained the molecular basis of switch III signaling and has provided structural insight into the question of how a G-protein facilitates cofactor transfer without binding the cofactor. However, the G-protein:mutase interactions do not represent the complete picture. The final step of the AdoCbl delivery also relies on the ATR interacting in concert with the G-protein to facilitate the maturation of the mutase. Chapter IV presented the structure of the ATR from *C. metallidurans*, comparing it to the previous structures of human and *M. tuberculosis* ATRs, which have both been extensively characterized. Our recently determined *Cm*ATR structure was used to computationally predict the structure of an *Cm*ATR:*Cm*IcmF complex. Variations in predicted docking solutions lead to numerous potential ways that cofactor transfer could occur, none of which seeming all that likely due to long distances, a lack of contacts between the ATR and G-protein domains of IcmF, and what would appear to be non-productive orientations of the AdoCbl binding site in ATR with respect to the IcmF Cbl-binding domain. Although some biochemical information is available about features important for AdoCbl transfer from ATR to a mutase, such as the importance of the dimethylbenzimidazole tail of the AdoCbl cofactor for AdoCbl transfer [19], for the most part residues or motifs important for cofactor transfer have not yet been elucidated, limiting our ability to validate of any docking solutions. Once again, cryo-EM could provide access to these complexes. The method of sample preparation for cryo-EM allows for the trapping of transient complexes, as observed with the reconstruction of the IcmF Q341A + GTP dataset (Chapter III). The ATR is involved in both delivery of intact AdoCbl and removal of the damaged cofactor, providing two different steps in which G-protein:mutase:ATR complexes form and two different steps to possibly characterize using cryo EM. The establishment of the fused system as comparable to the non-fused system sets the stage for using this ‘simpler’ system. Structurally characterizing cofactor transfer in the fused system will allow for understanding regarding how these three enzymes work in concert to mature the AdoCbl dependent mutase, deepening our understanding of metallochaperones and their roles in metalloenzyme maturation.

References

(1) Capdevila, Daiana A.; Edmonds, Katherine A.; Giedroc, David P. Metallochaperones and metalloregulation in bacteria. *Essays In Biochemistry* **2017**, *61* (2), 177-200. DOI: 10.1042/ebc20160076.

(2) Rosenzweig, A. C. Metallochaperones: bind and deliver. *Chemistry & Biology* **2002**, *9* (6), 673-677. DOI: 10.1016/S1074-5521(02)00156-4.

(3) O'Halloran, T. V.; Culotta, V. C. Metallochaperones, an intracellular shuttle service for metal ions. *Journal of Biological Chemistry* **2000**, *275* (33), 25057-25060. DOI: 10.1074/jbc.R000006200.

(4) Vaccaro, F. A.; Drennan, C. L. The role of nucleoside triphosphate hydrolase metallochaperones in making metalloenzymes. *Metallomics* **2022**, *14* (6), mfac030.

(5) Leipe, D. D.; Wolf, Y. I.; Koonin, E. V.; Aravind, L. Classification and evolution of P-loop GTPases and related ATPases. *Journal of Molecular Biology* **2002**, *317* (1), 41-72. DOI: 10.1006/jmbi.2001.5378.

(6) Oberholzer, V.; Levin, B.; Burgess, E.; Young, W. F. Methylmalonic aciduria. An inborn error of metabolism leading to chronic metabolic acidosis. *Archives of disease in childhood* **1967**, *42* (225), 492.

(7) Froese, D. S.; Gravel, R. A. Genetic disorders of vitamin B_{12} metabolism: eight complementation groups--eight genes. *Expert reviews in molecular medicine* **2010**, *12*, e37-e37. DOI: 10.1017/s1462399410001651.

(8) Dobson, C. M.; Wai, T.; Leclerc, D.; Kadir, H.; Narang, M.; Lerner-Ellis, J. P.; Hudson, T. J.; Rosenblatt, D. S.; Gravel, R. A. Identification of the gene responsible for the cblB complementation group of vitamin B12-dependent methylmalonic aciduria. *Human Molecular Genetics* **2002**, *11* (26), 3361-3369. DOI: 10.1093/hmg/11.26.3361.

(9) Dobson, C. M.; Wai, T.; Leclerc, D.; Wilson, A.; Wu, X.; Doré, C.; Hudson, T.; Rosenblatt, D. S.; Gravel, R. A. Identification of the gene responsible for the *cblA* complementation group of vitamin B_{12}-responsive methylmalonic acidemia based on analysis of prokaryotic gene arrangements. *Proceedings of the National Academy of Sciences* **2002**, *99* (24), 15554. DOI: 10.1073/pnas.242614799.

(10) Yuen, M. H.; Fong, Y. H.; Nim, Y. S.; Lau, P. H.; Wong, K.-B. Structural insights into how GTP-dependent conformational changes in a metallochaperone UreG facilitate urease maturation. *Proceedings of the National Academy of Sciences* **2017**, 201712658.

(11) Cai, F.; Ngu, T. T.; Kaluarachchi, H.; Zamble, D. B. Relationship between the GTPase, metal-binding, and dimerization activities of E. coli HypB. *JBIC Journal of Biological Inorganic Chemistry* **2011**, *16* (6), 857-868.

(12) Gasper, R.; Scrima, A.; Wittinghofer, A. Structural insights into HypB, a GTP-binding protein that regulates metal binding. *The Journal of Biological Chemistry* **2006**, *281* (37), 27492-27502. DOI: 10.1074/jbc.M600809200.

(13) Froese, D. S.; Kochan, G.; Muniz, J. R.; Wu, X.; Gileadi, C.; Ugochukwu, E.; Krysztofinska, E.; Gravel, R. A.; Oppermann, U.; Yue, W. W. Structures of the human GTPase MMAA and vitamin B_{12}-dependent methylmalonyl-CoA mutase and insight into their complex formation. *Journal of Biological Chemistry* **2010**, *285* (49), 38204-38213.

(14) Cracan, V.; Padovani, D.; Banerjee, R. IcmF is a fusion between the radical B_{12} enzyme isobutyryl-CoA mutase and its G-protein chaperone. *Journal of Biological Chemistry* **2010**, *285* (1), 655-666. DOI: 10.1074/jbc.M109.062182.

(15) Jost, M.; Cracan, V.; Hubbard, P. A.; Banerjee, R.; Drennan, C. L. Visualization of a radical B_{12} enzyme with its G-protein chaperone. *Proceedings of the National Academy of Sciences* **2015**, *112* (8), 2419-2424.

(16) Ruetz, M.; Campanello, G. C.; McDevitt, L.; Yokom, A. L.; Yadav, P. K.; Watkins, D.; Rosenblatt, D. S.; Ohi, M. D.; Southworth, D. R.; Banerjee, R. Allosteric regulation of oligomerization by a B_{12} trafficking G-protein is corrupted in methylmalonic aciduria. *Cell Chemical Biology* **2019**, *26* (7), 960-969.e964. DOI: 10.1016/j.chembiol.2019.03.014.

(17) Lofgren, M.; Padovani, D.; Koutmos, M.; Banerjee, R. A switch III motif relays signaling between a B_{12} enzyme and its G-protein chaperone. *Nature Chemical Biology* **2013**, *9* (9), 535-541. DOI: 10.1038/nchembio.1298.

(18) Dempsey-Nunez, L.; Illson, M. L.; Kent, J.; Huang, Q.; Brebner, A.; Watkins, D.; Gilfix, B. M.; Wittwer, C. T.; Rosenblatt, D. S. High resolution melting analysis of the MMAA gene in patients with cblA and in those with undiagnosed methylmalonic aciduria. *Molecular Genetics and Metabolism* **2012**, *107* (3), 363-367.

(19) Mascarenhas, R.; Ruetz, M.; McDevitt, L.; Koutmos, M.; Banerjee, R. Mobile loop dynamics in adenosyltransferase control binding and reactivity of coenzyme B_{12}. *Proceedings of the National Academy of Sciences* **2020**, *117* (48), 30412-30422. DOI: doi:10.1073/pnas.2007332117.

Appendix AI: Materials and methods for the preliminary structural characterizations of the MMAA:MCM complex

Author Contributions

Appendix AI was written by Francesca A. Vaccaro; the negative stain electron microscopy experiments were performed with Daphne A. Faber. Talya S. Levitz collected the dataset.

Materials

All chemicals, solvents and reagents were purchased from Sigma-Aldrich unless otherwise noted below. *E. coli* Bl21 T7 Express competent cells were purchased from New England Biolabs. LB medium components were purchased from Fisher BioReagents. Kanamycin was purchased from GoldBio and used at a concentration of 50 µg/mL. Isopropyl-β-D-thiogalactopyranoside (IPTG) was purchased from GoldBio. The NaCl for purification buffer components was purchased from Fisher Chemical. The EDTA-free protease inhibitor cocktail tablets were purchased from Roche. The Ni-NTA 1 mL columns and Superdex200 16/60, and Superdex200 10/300 Increase GL size exclusion columns (SEC) were purchased from GE Healthcare. The gel filtration standards, polyacrylamide gels and sodium dodecyl sulfate solution were purchased from BioRad. The $MgCl_2$ was purchased from CalBiochem. The carbon-coated 300 mesh copper electron microscopy grids were purchased from Electron Microscopy Sciences.

Methods

Protein expression and purification of MMAA

The cell growth and purification of MMAA from *Homo sapiens* was conducted with the following procedure. An overnight culture of 10 mL of LB supplemented with 50 µg/mL of kanamycin was inoculated from a single colony of T7 express cells transformed with the plasmid containing the optimized *mmaa* gene from *H. sapiens* with an N-terminal his tag in a pET28a vector (Genescript ©) and grown at 37 ºC with shaking at 220 RPM. The overnight starter culture was used to inoculate 1 L of LB supplemented with 50 µg/mL kanamycin at 37 ºC with shaking at 220 RPM. The 1 L culture was induced with 0.5 mM IPTG when OD_{600} reached ~0.4 - 0.5 and grown for 17 h at 18 ºC with shaking at 220 RPM. Cells were harvested by centrifugation (5,000 x g, 15 min, 4 ºC) and flash frozen in liquid N_2 before being stored in a -80 ºC freezer for future use.

Cells from 2 L of growth were resuspended in 40 mL lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 25 mM imidazole) supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF) and one-half of an EDTA-free protease inhibitor cocktail tablet. Cells were lysed by ultrasonication, and cell lysates were clarified by centrifugation (28,000 x g, 30 min, 4 ºC). Clarified lysate was passed through a 0.2 µm filter before being loaded onto a 1 mL Ni-NTA column equilibrated with lysis buffer. Protein was eluted with a step

gradient of elution buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 500 mM imidazole) using an FPLC (BioRad NGC Chromatography system). The column was washed with 5% of elution buffer, then 12% elution buffer and finally the protein was eluted with 80% of elution buffer. Elution fractions were concentrated in a 30 kDa MWCO centrifugal filter. The concentrated fractions of MMAA were loaded onto a Superdex 200 16/60 SEC equilibrated with SEC buffer (50 mM HEPES pH 7.5, 500 mM NaCl). Elution fractions from SEC were concentrated in a 30 kDa molecular weight cutoff (MWCO) spin concentrator. Purity was assessed by 4-20% (w/v) sodium dodecyl sulfate polyacrylamide gel electrophoresis. The concentration of the MMAA monomer was determined by UV/Vis absorbance at 280 nm using a molecular weight of 41.3 kDa and an extinction coefficient of 30,940 $M^{-1}cm^{-1}$, determined using the ProtParam tool [1]. Protein samples in SEC buffer were flash frozen in liquid N_2 and stored in a -80 ºC freezer for future use.

Protein expression and purification of HsMCM

The cell growth and purification of MCM from *Homo sapiens* was conducted with the following procedure. An overnight culture of 10 mL of LB supplemented with 50 µg/mL of kanamycin was inoculated from a single colony of T7 express cells transformed with the plasmid containing the optimized *mut* gene from *H. sapiens* with an N-terminal his tag in a pET28a vector and grown at 37 ºC with shaking at 220 RPM. The overnight starter culture was used to inoculate 1 L of LB supplemented with 50 µg/mL kanamycin at 37 ºC with shaking at 220 RPM. The 1 L culture was induced with 0.1 mM IPTG when OD_{600} reached ~0.6 - 0.7 and grown for 17 h at 18 ºC with shaking at 220 RPM. Cells were harvested by centrifugation (3,500 x g, 15 min, 4 ºC) and flash frozen in liquid N_2 before being stored in a -80 ºC freezer for future use.

Cells from 2 L of growth were resuspended in 35 mL lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 25 mM imidazole) supplemented with 1 mM PMSF and one-half of an EDTA-free protease inhibitor cocktail tablet. Cells were lysed by ultrasonication, and cell lysates were clarified by centrifugation (28,000 x g, 30 min, 4 ºC). Clarified lysate was passed through a 0.2 µm filter before being loaded onto a 1 mL Ni-NTA column equilibrated with lysis buffer. Protein was eluted with a linear gradient of from 40 % to 100% of the elution buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 500 mM imidazole)

using an FPLC (BioRad NGC Chromatography system). Elution fractions were concentrated in a 30 kDa MWCO centrifugal filter. The concentrated fractions of MMAA were loaded onto a Superdex 200 16/60 SEC equilibrated with SEC buffer (50 mM HEPES pH 7.5, 500 mM NaCl). Elution fractions from SEC were concentrated in a 30 kDa MWCO spin concentrator. Purity was assessed by 4-20% (w/v) sodium dodecyl sulfate polyacrylamide gel electrophoresis. The concentration of MCM monomer was determined by UV/Vis absorbance at 280 nm using a molecular weight of 81.6 kDa and an extinction coefficient of 66,320 $M^{-1}cm^{-1}$, determined using the ProtParam tool [1]. Protein samples in SEC buffer were flash frozen in liquid N_2 and stored in a -80 ºC freezer for future use.

Negative stain electron microscopy specimen preparation

Frozen samples of MMAA and MCM were thawed on ice. MMAA was diluted to 50 µM with SEC buffer and incubated with 1 mM GMPPCP for 30 minutes before adding MCM at a final concentration of 25 µM. This sample was incubated on ice for 2 minutes and then was spun at max speed for 3 minutes. 100 µL of the MMAA:MCM:GMPPCP sample was loaded onto an S200 increase 10/300 column equilibrated with SEC buffer; the sample was eluted with 1 CV of SEC buffer. The concentration of the elution fractions was determined by UV/Vis absorbance at 280 nm normalizing 1 Abs to 1 mg/mL of protein concentration because the stoichiometry of the complex was unknown. A sample representing the first elution fraction was diluted to 20 ng/µL. Carbon-coated 300 mesh copper electron microscopy grids were glow discharged for 1 min at -15 mA. 5 µL of the protein solution was applied to the grid. After approximately 1 min, the solution was blotted and immediately replaced with solution of 2% uranyl acetate. The stain solution was blotted and replaced twice, then allowed to stand for 1 min before the final blot, and then was dried. All blotting was done manually using filter paper.

Negative stain electron microscopy dataset collection and processing

Data were collected at the Electron Microscopy Facility at Brandeis University. The specimen was imaged with a Falcon2 camera on a FEI Tecnai F30 electron microscope operated at 300 kV using SerialEM. 512 micrographs were collected at 59,000 x magnification as 0.8923 s (19.28 $e^{-}/Å^2$) exposures with a pixel size corresponding to 1.83

Å. Data processing was carried out using Relion 3.0[2]. After importing the micrographs, the defocus of the images was estimated using Relion's implementation of Gctf. Approximately 1000 particles were manually picked and extracted with a box size of 540 pixels. These extracted particles were subjected to a round of 2D classification to generate templates for automatic picking of particles. Before automatically picking particles, the micrographs were visually inspected for quality, resulting in 496 micrographs. Relion autopicking with an inter-particle distance of 200 Å and threshold of 0.7 yielded 38,721 particles. These particles were subjected to numerous rounds of 2D classifications with a mask diameter of 420 Å. 10,722 particles were supplied to generate an ab initio initial model with a mask diameter of 420 Å. The crystal structure of MMAA (PDB 2WWW)[3] and MCM (PDB 2XIQ)[3] were manually docked into the initial model using ChimeraX v1.5. Further refinements of the map were unsuccessful. The figure of the reconstructed density was made using ChimeraX v1.5[4].

Figures

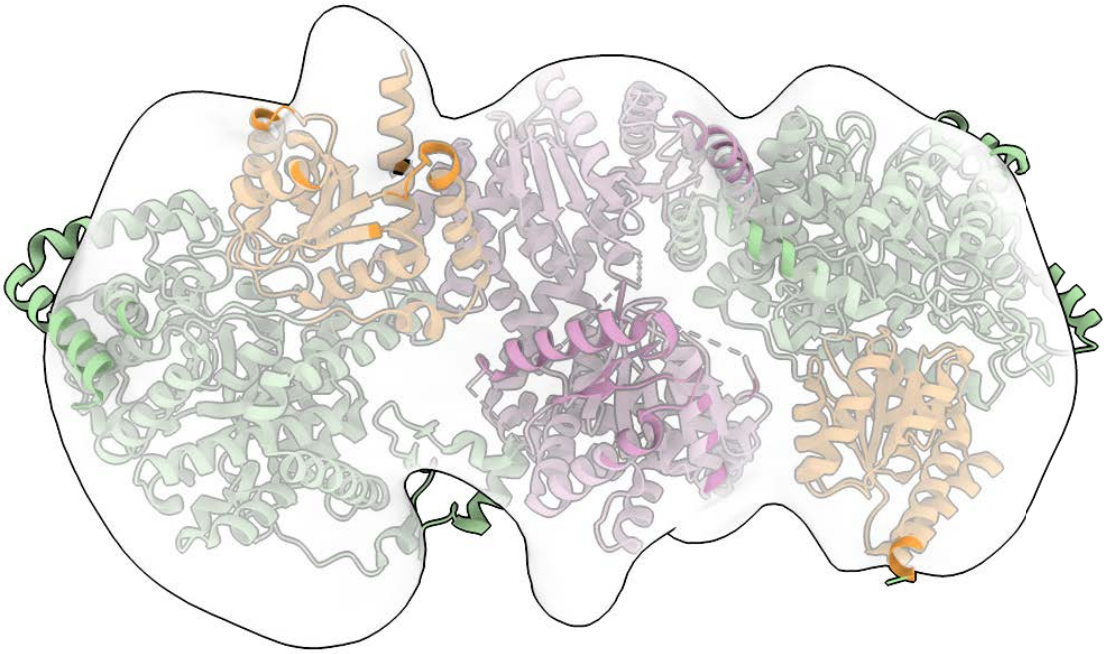


***Figure AI.1*. The initial 3D model from the negative stain sample of the MMAA and MCM in the presence of GMPPCP**. The initial model reconstructed from the MCM:MMAA:GMPPCP negative stain dataset contained density for three different lobes. Two copies of MCM (Cbl-binding domain: orange; substrate binding domain: green; PDB 2XIQ)[3] and one copy of MMAA dimer (pink ribbons; PDB 2WWW)[3] were manually docked into the reconstruction with good agreement indicating the potential formation of a linear complex competent for cofactor loading.

References

(1) Wilkins, M. R.; Gasteiger, E.; Bairoch, A.; Sanchez, J.-C.; Williams, K. L.; Appel, R. D.; Hochstrasser, D. F. Protein Identification and Analysis Tools in the ExPASy Server. In *2-D Proteome Analysis Protocols*, Link, A. J. Ed.; Humana Press, 1999; pp 531-552.

(2) Zivanov, J.; Nakane, T.; Forsberg, B. O.; Kimanius, D.; Hagen, W. J.; Lindahl, E.; Scheres, S. H. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *elife* **2018**, *7*, e42166.

(3) Froese, D. S.; Kochan, G.; Muniz, J. R.; Wu, X.; Gileadi, C.; Ugochukwu, E.; Krysztofinska, E.; Gravel, R. A.; Oppermann, U.; Yue, W. W. Structures of the human GTPase MMAA and vitamin B_{12}-dependent methylmalonyl-CoA mutase and insight into their complex formation. *Journal of Biological Chemistry* **2010**, *285* (49), 38204-38213.

(4) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Meng, E. C.; Couch, G. S.; Croll, T. I.; Morris, J. H.; Ferrin, T. E. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science* **2021**, *30* (1), 70-82. DOI: 10.1002/pro.3943.