Technische Universität München
TUM School of Natural Sciences

Structural and Functional Characterizations of Proteins Involved in Pyrrolysine- and Cobalamin-dependent Methyl Transfer

Thomas Patrick Badmann

Vollständiger Abdruck der von der TUM School of Natural Sciences
der Technischen Universität München zur Erlangung eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

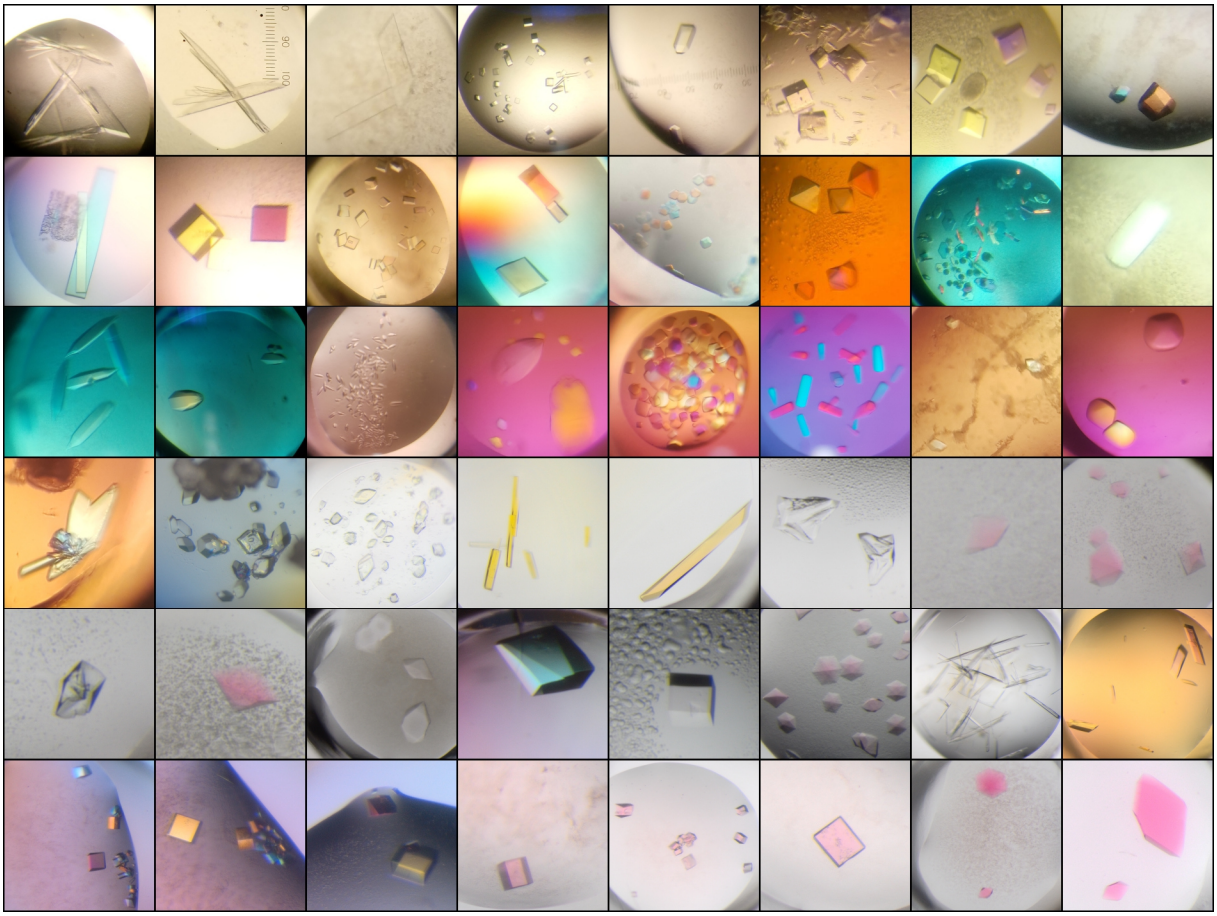
genehmigten Dissertation.

Vorsitz: Prof. Dr. Lukas Hintermann

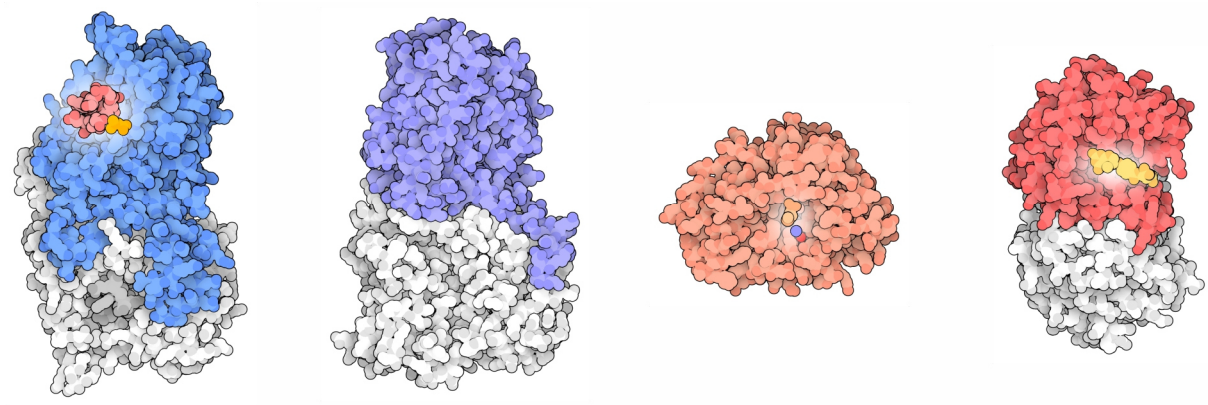
Prüfer der Dissertation:

1. Prof. Dr. Michael Groll
2. Prof. Dr. Stephan A. Sieber

Die Dissertation wurde am 17.04.2024 bei der Technischen Universität München eingereicht und durch die TUM School of Natural Sciences am 22.05.2024 angenommen.



Meiner Familie



Danksagung

Mein besonderer Dank gilt Prof. Dr. Michael Groll, der mir die Möglichkeit gegeben hat, meine Doktorarbeit über dieses interessante Thema an seinem Lehrstuhl zu schreiben. Seine Betreuung hat sich durch große Begeisterung für die Forschung, hilfreiche Ratschläge und viel wissenschaftlichen Freiraum für eigene Ideen ausgezeichnet. Ohne diese positive Arbeitsatmosphäre wäre diese Dissertation nicht möglich gewesen.

Des Weiteren möchte ich mich bei meinen Kooperationspartnern JProf. Dr. Silja Mordhorst, Dr. Anna Vagstad und Prof. Dr. Jörn Piel für die erfolgreiche Zusammenarbeit im Rahmen unseres gemeinsamen Projektes zur Landornamid-Biosynthese bedanken. Ich habe den wissenschaftlichen Austausch zwischen unseren Gruppen immer sehr genossen und freue mich über die interessante Publikation, die daraus entstanden ist.

Ich bedanke mich herzlich bei Dr. Marie-Lena Jokisch und Prof. Dr. Kathrin Lang, die durch die Bereitstellung von Material für die Amber Suppression Methode und durch die Einführung in die Gibson Assembly Methode sehr zum Fortschritt meiner Arbeit beigetragen haben. Des Weiteren möchte ich mich bei Prof. Dr. Lena Daumann für ihre hilfreichen Ratschläge über die Durchführung von anaeroben Aktivitätsassays bedanken. Bei der täglichen Laborarbeit konnte ich mich zudem auf Astrid König, Katrin Gärtner und Leon Kayser verlassen, die mich durch das Bereitstellen von Labormaterialien und das Pipettieren von Kristallisationsansätzen tatkräftig unterstützt haben. Außerdem möchte ich mich herzlich bei Laura Meier für das Anfertigen der MS-Messungen für diese Arbeit bedanken. Dem wissenschaftlichen und technischen Personal der Beamline S06SA am Paul-Scherrer-Institut in Villigen danke ich vielmals für die Unterstützung bei der Datenerfassung.

Darüber hinaus gilt mein Dank meinen Kolleginnen und Kollegen Wolfgang H., Eva, Astrid, Katrin, Leon, Felix, Max, Julia, Wolfgang K., Rachel, Annkathrin, Philip, Tim, Nick, Anna, Chengyang, Fiona und Johannes für die stets sehr angenehme Arbeitsatmosphäre und die große Hilfsbereitschaft. Zusätzlich möchte ich mich bei meinen Praktikantinnen und Praktikanten Simon, Viktor, Aziza, Lisa, Mika und Mona bedanken, die mit ihrem großen Arbeitseinsatz wesentlich zu den Ergebnissen dieser Arbeit beigetragen haben.

Ein großes Dankeschön geht schließlich an meinen Freundeskreis und meine Familie. Ohne eure stetige Unterstützung und Motivation während des Studiums und der Promotion wäre ich nicht so weit gekommen.

Table of Contents

Zusammenfassung

Pyrrolysin ist die 22. proteinogene Aminosäure, die in manchen methanogenen Archaeen sowie einigen Bakterien über Amber-Codons (TAG) in Proteine eingebaut wird. Nur drei Enzyme in der Natur verwenden Pyrrolysin für die Katalyse. Die Methyltransferasen MtmB, MtbB und MttB demethylieren ihre jeweiligen Substrate Monomethylamin, Dimethylamin oder Trimethylamin, um die gewonnenen Methylgruppen für die Energiegewinnung mittels Methanogenese bereitzustellen. Interessanterweise sind homologe Proteine der Methyltransferase MttB, die kein Pyrrolysin enthalten, weit verbreitet. Darunter befinden sich MtgB, MtpB, MtcB und MtyB, die als Demethylasen von verschiedenen quartären Aminen identifiziert wurden. In der Literatur fehlt es jedoch an strukturellen Details über ihren Mechanismus. Trägerproteine mit dem Cofaktor Cobalamin koordinieren die Methylabspaltung von methylierten Aminen und geben die Methylgruppe an assoziierte Methyltransferasen weiter, die die Methylierung von Cofaktoren wie Coenzym M oder Tetrahydrofolat katalysieren. Die an Cobalamin-vermitteltem Methyltransfer beteiligten Proteine weisen selbst in entfernt verwandten Organismen einen hohen Grad an Strukturverwandtschaft auf. Erkenntnisse über die Funktionsweise eines dieser Enzyme haben daher hohes Potential, auf eine Vielzahl von verwandten Proteinen übertragen werden zu können und helfen dabei, die immer noch nicht im Detail verstandenen Mechanismen des Cobalamin-vermittelten Methyltransfers aufzuklären.

Die Ergebnisse dieser Dissertation lassen sich in drei Hauptthemen gliedern: (i) Um Pyrrolysin in Proteine in *Escherichia coli* einzubauen, wurde die heterologe Produktion der für die Pyrrolysin-Biosynthese verantwortlichen Enzyme mit einem Protokoll für den Einbau nicht-natürlicher Aminosäuren kombiniert. Dieses System für Amber Suppression wurde in Zusammenarbeit mit der Gruppe von Prof. Kathrin Lang (ETH Zürich) am Lehrstuhl für Biochemie etabliert und kann nun projektübergreifend für andere nicht-natürliche Aminosäuren eingesetzt werden. (ii) Das Protokoll für molekularbiologische Arbeiten am Lehrstuhl wurde generalüberholt, um die Durchführung von klonierungsintensiveren Projekten zu erleichtern. Die Leistungsfähigkeit des neuen Arbeitsablaufes wurde anhand der Erstellung einer Reihe an Vektoren mit modifizierten Varianten des Gens der Methyltransferase MttB demonstriert. Das neu etablierte Protokoll zeichnet sich durch die Anwendung der Gibson Assembly®-Methode und eine digitale Plasmidbibliothek aus. Es wurde für die Erstellung einer Vielzahl von Vektoren für diese Arbeit verwendet und trägt wesentlich zu einem effizienteren Ablauf der Molekularbiologie am Lehrstuhl für Biochemie bei. (iii) Mit Hilfe der Röntgenkristallographie wurden die Strukturen von vier Polypeptiden des mikrobiellen Methylamin-Stoffwechsels, MtgB, MtcB, MtbA und MtgA, aufgeklärt. Substrat- und Cobalamin-

gebundene Strukturen von MtgB zeigen, wie Glycinbetain für die Methylübertragung auf Cobalamin ausgerichtet wird. Die Strukturen der Methyltransferasen MtgB und MtcB, zwei Homologe der MttB-Familie ohne Pyrrolysin, weisen deutliche Unterschiede zu ihren Verwandten auf. Das Fehlen der katalytischen Aminosäure Pyrrolysin wird durch grundlegend andere Aminosäurezusammensetzungen der aktiven Taschen kompensiert, die darauf abzielen, quartäre Amine in der richtigen Position zu komplexieren. Anhand einer Substrat- und einer Produktstruktur der Zink-abhängigen Coenzym M-Methylase MtbA lässt sich eine Inversion der tetraedrischen Metall-Koordinierung im Verlauf der Reaktion dieses Enzyms erkennen. Coenzym M geht durch die Methylierung als Interaktionspartner des Zink-Ions verloren und wird auf der anderen Seite durch ein Aspartat ersetzt. Substrat- und Produkt-gebundene Strukturen von MtgA wurden zusammen mit anaeroben Aktivitätsmessungen für eine detaillierte Charakterisierung dieser Methyltransferase herangezogen. Dieses Enzym fixiert seinen Cofaktor Tetrahydrofolat nach der Methylierung auf neuartige Weise und stabilisiert damit den protonierten Reaktionsübergangszustand. Die aus den Strukturen gewonnenen Informationen tragen zur Grundlagenforschung auf den Gebieten der Methanogenese, der Cofaktoren Cobalamin, Coenzym M und Tetrahydrofolat, sowie des Pyrrolysin-vermittelten Methyltransfers bei.

Abstract

Pyrrolysine is the 22nd proteinogenic amino acid that is incorporated into proteins in response to amber codons (TAG) in a subset of methanogenic archaea as well as some bacteria. Only three enzymes in nature use pyrrolysine for catalysis. The methyltransferases MtmB, MtbB, and MttB demethylate their respective substrates monomethylamine, dimethylamine, or trimethylamine to provide the obtained methyl groups for energy production via methanogenesis. Interestingly, homologous proteins of the methyltransferase MttB that do not contain pyrrolysine are widespread. These include MtgB, MtpB, MtcB, and MtyB, which have been identified as demethylases of various quaternary amines. However, literature is lacking structural details about their mode of action. Carrier proteins with the cofactor cobalamin coordinate the methyl cleavage from methylated amines and pass the methyl group onto associated methyltransferases that catalyze the methylation of cofactors such as coenzyme M or tetrahydrofolate. The proteins involved in cobalamin-mediated methyl transfer show a high degree of structural similarity even in distantly related organisms. Insights into the function of one of these enzymes therefore have great potential to be transferred to a large number of related proteins and help to elucidate the mechanisms of cobalamin-mediated methyl transfer, which are still not understood in detail.

The results of this dissertation can be divided into three main topics: (i) In order to incorporate pyrrolysine into proteins in *Escherichia coli*, the heterologous production of the enzymes responsible for pyrrolysine biosynthesis was combined with a protocol for the incorporation of non-natural amino acids. This amber suppression system was established at the Chair of Biochemistry in collaboration with the group of Prof. Kathrin Lang (ETH Zurich) and can now be used for other non-natural amino acids across projects. (ii) The protocol for molecular biology at the Chair has been updated to facilitate more cloning-intensive projects. The performance of the new workflow was demonstrated by creating a series of vectors with modified variants of the gene encoding the methyltransferase MttB. The newly established protocol is characterized by the use of the Gibson Assembly® method and a digital plasmid library. It was used to create a large number of vectors for this work and contributes significantly to a more efficient molecular biology workflow at the Chair of Biochemistry. (iii) The structures of four polypeptides of the microbial methylamine metabolism, MtgB, MtcB, MtbA, and MtgA, were elucidated using X-ray crystallography. Substrate- and cobalamin-bound structures of MtgB show how glycinebetaine is positioned for methyl transfer to cobalamin. The structures of the methyltransferases MtgB and MtcB, two homologs of the MttB family without pyrrolysine, show clear differences compared to their relatives. The lack of the catalytic amino acid pyrrolysine is compensated by fundamentally different amino acid

compositions of the active pockets, which are aimed at complexing quaternary amines in the correct position. Based on a substrate and a product structure of the zinc-dependent coenzyme M methylase MtbA, an inversion of the tetrahedral metal coordination can be observed during the course of the reaction of this enzyme. Upon methylation, coenzyme M is lost as an interaction partner of the zinc ion and is replaced by an aspartate on the other side. Substrate- and product-bound structures of MtgA were used in combination with anaerobic activity measurements for the detailed characterization of this methyltransferase. This enzyme locks its cofactor tetrahydrofolate after methylation in a novel way and thus stabilizes the protonated reaction intermediate state. The information obtained from the structures contributes to fundamental research in the fields of methanogenesis, the cofactors cobalamin, coenzyme M and tetrahydrofolate, as well as pyrrolysine-mediated methyl transfer.

Abbreviations

3MO	(3*R*)-3-methyl-*D*-ornithine
APS	Ammonium persulfate
β-ME	β-Mercaptoethanol
BocK	Boc-protected lysine
Cbl	Cobalamin or Vitamin B12
Cbl_{5OH}	5-Hydroxybenzimidazolecobamide
CH_4	Methane
CO	Carbon monooxide
CO_2	Carbon dioxide
CoFeSP	Corrinoid iron-sulfur protein
CoM	Coenzyme M
D. hafniense	*Desulfitobacterium hafniense*
DMA	Dimethylamine
dmb	Dimethylbenzimidazole
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
E. limosum	*Eubacterium limosum*
Fe-S	Iron-sulfur
GB	Glycinebetaine
GFP	Green fluorescent protein
H_2	Hydrogen
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IMAC	Immobilized-metal affinity chromatography
IPTG	Isopropyl-β-D-thiogalactoside
LB	Lysogeny Broth
LC	Liquid chromatography
M. barkeri	*Methanosarcina barkeri*
MeCbl	Methyl-cobalamin
$MeCbl_{5OH}$	Methyl-5-Hydroxybenzimidazolecobamide
MeCoM	Methyl-coenzyme M
MetH	Cobalamin-dependent methionine synthase
Methanogens	Methanogenic organisms
MeTHF	Methyl-tetrahydrofolate
MeTr	Methyltransferase component of the Wood-Ljungdahl pathway
MMA	Monomethylamine
MOPS	3-(*N*-morpholino)propanesulfonic acid
MS	Mass spectrometry
ncAA	Non-canonical amino acid
OD_{600}	Optical density at 600 nm
PAGE	Polyacrylamide gel electrophoresis
Pcl	Pyrroline-carboxy-lysine
RMSD	Root-mean-square deviation
RT	Room temperature
SAD	Single anomalous dispersion
SAM	*S*-adenosylmethionine
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography

SeMet	Selenomethionine
sfGFP	Superfolder GFP
TAE	Tris/Acetate/EDTA
TB	Terrific Broth
TEMED	Tetramethylethylenediamine
THF	Tetrahydrofolate
THMPt	Tetrahydromethanopterin
THSPt	Tetrahydrosarcinapterin
TMA	Trimethylamine
Tris	Tris(hydroxymethyl)aminomethane
$tRNA^{Pyl}$	Pyrrolysyl-tRNA
UROD	Uroporphyrinogen decarboxylase
WT	Wild Type

1 Introduction

1.1 The Exotic Biochemistry of Methanogenic Archaea

1.1.1 Methanogenesis

Methanogenic archaea (methanogens) play a pivotal role in the global carbon cycle. They produce an estimated 900 million metric tons of methane (CH_4) annually, making them a primary source of this greenhouse gas and a significant contributor to global warming.[1] Methanogens are obligate anaerobes and inhabit various oxygen-free environments in the biosphere. These include hydrothermal vents, sewage sludge, marine and freshwater sediments, rice fields, landfills, and the rumen and lower gut of livestock.[2-5] Being a remnant of the oxygen-free era, methanogenesis evolved as a unique form of anaerobic respiration, which made use of a variety of simple carbon compounds available in the primordial soup.[6] The carbon is used as a terminal electron acceptor and reduced to CH_4 to generate an electrochemical gradient for use in ATP synthesis.[7-9]

Three major biochemical pathways are responsible for methanogenesis. Hydrogenotrophic methanogenesis is conducted by reduction of carbon dioxide (CO_2) to CH_4 using hydrogen (H_2) as an electron donor (Fig. 1, yellow box). In acetoclastic methanogenesis, acetate is split into a methyl group and an enzyme-bound carbon monoxide, which is subsequently oxidized to CO_2 and funneled into the hydrogenotrophic pathway (Fig. 1, red box). The methylotrophic pathway is fueled by various methylated compounds, referred to as methylotrophic substrates, which are demethylated to produce CH_4 (Fig. 1, blue box). The three pathways converge in the conversion of 2-thioethanesulfonic acid (Coenzyme M or CoM, where "M" stands for methyl transfer) to 2-(methylthio)ethanesulfonic acid (methyl-coenzyme M or MeCoM).[2,10-13] This central intermediate is further reduced by MeCoM reductase to produce CH_4 (Fig. 1, grey box).[14]

The cofactor CoM is exclusively found in methanogenic biochemistry.[15] Its thiolate anion is a good nucleophile which can be easily alkylated,[16] making it an essential tool for methanogenic methyl transfer.[2] In addition to CoM, methanogenic archaea contain an abundance of unusual metabolites that function as cofactors or prosthetic groups.[2,4,5,17,18] Methanofuran and its structural derivatives are a unique class of cofactors that possess a furan moiety with a primary amine moiety. During hydrogenotrophic methanogenesis, this moiety undergoes formylation, which represents the initial conversion step from CO_2 to CH_4.[2,17] Methanopterin and its derivative sarcinapterin, which is only found in *Methanosarcina* species, are modified versions of folic acid and serve as carriers of single carbon groups of various oxidation stages (Fig. 1).[2,17] MeCoM reductase employs factor 430 to carry out the

terminal reductive demethylation of MeCoM to yield CH_4.[19] Factor 430 is a tetrapyrrole with a complexed nickel ion that undergoes different oxidation stages during the reaction.[2,17,20] The coenzyme 7-mercaptoheptanoylthreonine phosphate (HS-HTP) was established as the electron donor for this reductive de-methylation of MeCoM.[7,21] Factor 420 is a flavin analog that, among various other metabolic redox reactions, is involved in carbon reduction using electrons from H_2 during hydrogenotrophic methanogenesis.[2,17,22-24] Furthermore, methanogens contain extraordinarily high concentrations of cobamides (derivatives of Cbl/vitamin B12),[25] which are required as supernucleophiles for various methyl transfer reactions.[26,27] Cobamide-dependent reactions occur in all three branches of methanogenesis, underlining the importance of this group of cofactors.[11,28,29] Corrinoid proteins (the cobamide-harboring proteins) mediate the methyl transfer between a donor and an acceptor methyltransferase (Fig. 2).[26,27,30] For example, methylotrophic pathways share a characteristic three-step reaction scheme consisting of abstraction of the substrate methyl group, transfer

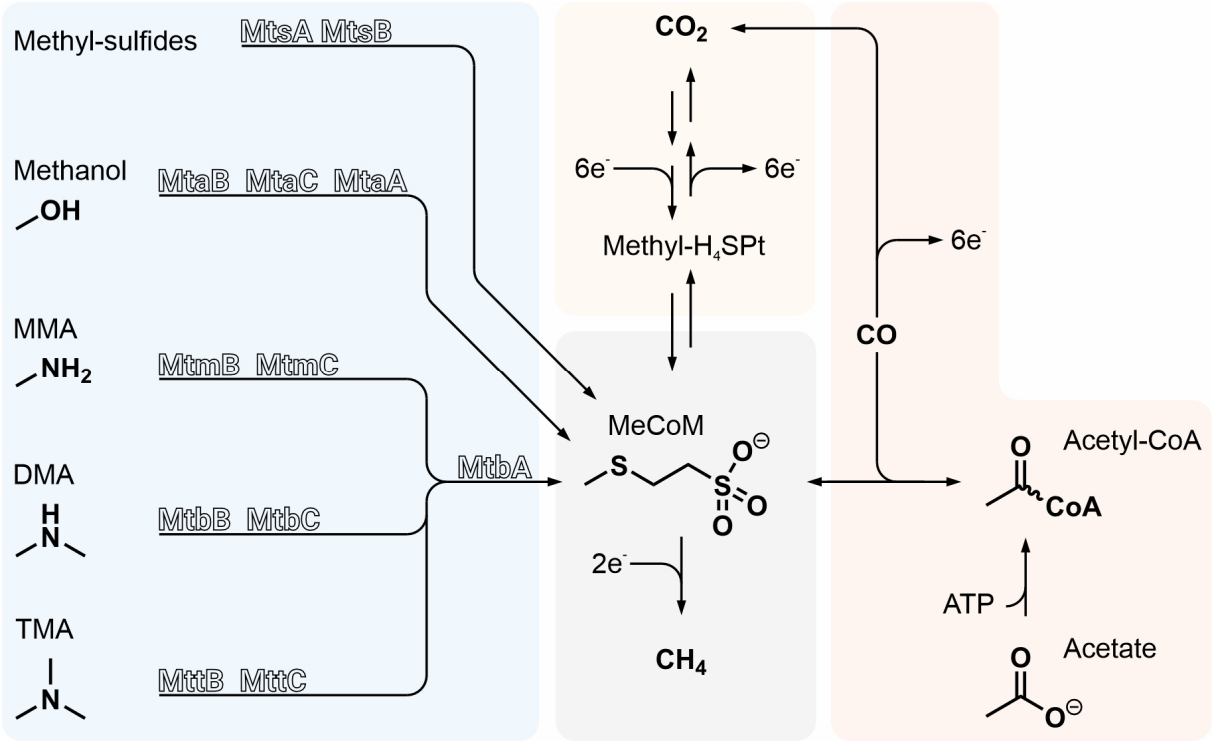


Figure 1 The three pathways of methanogenesis converge in the formation of MeCoM, which is reduced to CH_4 (grey box). Hydrogenotrophic methanogenesis (yellow box), a pathway performed by many members of methanogenic archaea, is characterized by the reduction from CO_2 to CH_4 with electrons derived from the oxidation of H_2 (H_4SPt or THSPt = tetrahydrosarcinapterin). Archaea that perform acetoclastic methanogenesis (red box), cleave acetate to obtain a methyl group and enzyme-bound carbon monoxide (CO). CO is oxidized to CO_2 to provide electrons for the reduction of the methyl group to CH_4. *Methanosarcina* species possess both the hydrogenotrophic and acetoclastic pathways and are additionally able to conduct methylotrophic methanogenesis (blue box) allowing them to transfer methyl groups from methylated thiols, methanol, and methylamines to CoM via a corrinoid protein (MtsB, MtaC, MtmC, MtbC, or MttC). The oxidation of one methyl group to CO_2 provides electrons for the reduction of three methyl groups to CH_4. All pathways assist in generating an electrochemical gradient used for ATP synthesis.

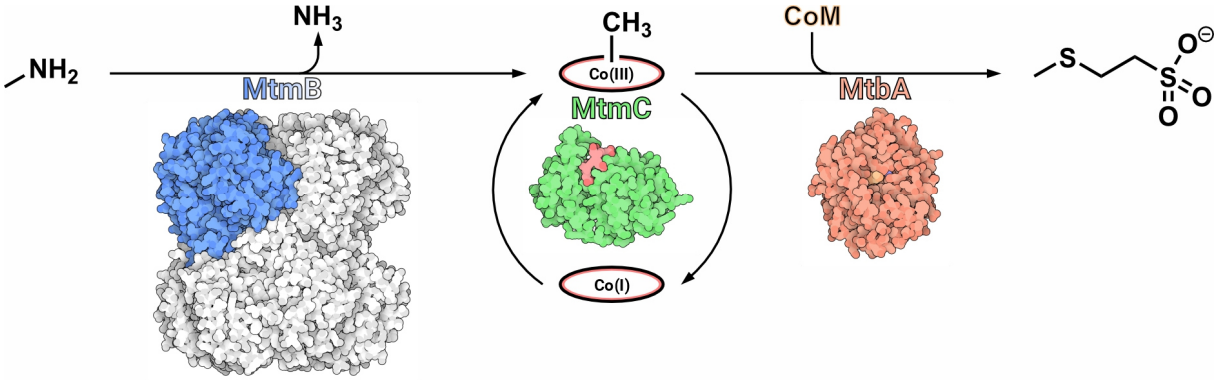


Figure 2 Scheme illustrating the corrinoid-dependent methyl transfer from MMA to CoM. MtmB abstracts a methyl group from MMA and transfers it to the corrinoid protein MtmC. The methyl group is then loaded onto CoM by the methyltransferase MtbA. During the methylation cycle, the cobalt ion of the corrinoid cofactor switches between oxidation states I and III. The structure depictions of MtmB and MtmC were generated from the PDB-IDs 1NTH and 3EZX, respectively. The structure of MtbA was solved in this work (MtbA:MeCoM).

to an associated corrinoid partner protein, and finally transfer of the methyl group to CoM to produce the central intermediate of methanogenesis, MeCoM (Fig. 2).[10]

The variety of different methanogenic cofactors is mirrored by the extensive range of substrates that can be utilized for methanogenesis. In particular, methylotrophic methanogenesis can be initiated by several methylated substrates including monomethylamine (MMA),[31,32] dimethylamine (DMA),[33,34] trimethylamine (TMA),[33,35,36] methanol,[13,37] as well as several methylated thiols (Fig. 1, blue box).[38-43] However, only some members of the order *Methanosarcinales* have evolved the ability to utilize methylotrophic methanogenesis in addition to hydrogenotrophic and acetoclastic methanogenesis, while most archaeal members are restricted to the latter two options.[44]

1.1.2 The 22nd Amino Acid

Methylotrophic methanogenesis is also responsible for the most noteworthy addition to the pool of methanogenic enzymatic tools. Examinations of the methylamine metabolism of *Methanosarcina* species led to the discovery of a novel proteinogenic amino acid. The discovery of this 22nd amino acid was first hinted at during genome analysis of the methanogenic archaeon *Methanosarcina barkeri* (*M. barkeri*). The open reading frame of *mtmB*, the gene encoding the MMA methyltransferase MtmB, was found to be interrupted by a single midframe canonical stop codon.[32,45] Initially dismissed as a sequencing or cloning error, repeated analysis with different methods and *M. barkeri* strains eventually and unambiguously confirmed the presence of an in-frame stop codon.[45] The 52 kDa MtmB was observed to be one of the most abundant proteins in MMA-grown cell extracts of *M. barkeri*, while the predicted 23 kDa truncated product of the UAG-interrupted *mtmB* gene was not noticeably abundant.[32] Hypotheses as to how translational termination was suppressed

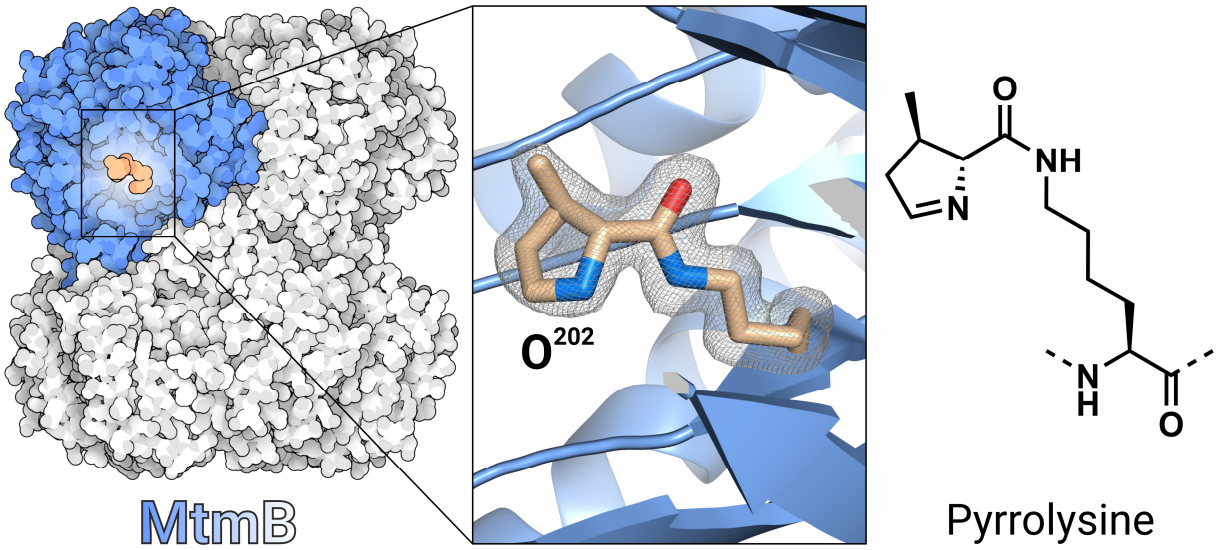


Figure 3 The methyltransferase MtmB (homohexamer, subunit A colored in blue, PDB-ID: 1NTH) and its active site catalytic residue, pyrrolysine (Pyl^{202}/O^{202} colored in beige).

during MtmB synthesis revolved around mechanisms already observed in other organisms, which included bypassing of the stop codon by translational hopping[46] or frameshifts during translation,[47] as well as usage of the codon for a canonical amino acid such as glutamine, which was observed in some protists.[48] Another scenario, which was later confirmed, suggested the co-translational insertion of a modified amino acid similar to the UGA-directed incorporation of selenocysteine in bacterial, eucaryotic, and archaeal proteins.[49-51] The proteins MtbB and MttB, which are responsible for the degradation of DMA and TMA, respectively, were found to carry a similar in-frame amber stop codon.[52] Given the lack of sequence similarity between the MMA, DMA, and TMA methyltransferase genes, the conservation of the mid-sequence stop codons was surprising.[52] In-frame amber codons have not been found in other genes of any member of the domain archaea, methanogenic or otherwise, suggesting that the amino acid encoded at the UAG site may function in a specialized regulatory or catalytic role regarding the methylamine metabolism.[45,52]

Sequencing of peptide fragments from MtmB and mass spectrometry (MS) postulated that the UAG codon was translated into a lysyl residue.[53] However, the determination of the MtmB crystal structure at atomic resolution revealed that the UAG-encoded residue was distinct from any of the 21 known proteinogenic amino acids. From the electron density the residue was identified as a lysine in linkage with (4*R*,5*R*)-4-substituted-pyrroline-5-carboxylate and named *L*-pyrrolysine (Fig. 3).[54,55] The occurrence of pyrrolysine at the amber codon position was subsequently confirmed for MtbB and MttB by MS, completing the list of naturally occurring pyrrolysine proteins.[56]

A tRNA with a CUA anticodon and an unusual tRNA-synthetase were found to be encoded near the *mtmB* genes on the genome of *M. barkeri*, which suggested tRNA loading of pre-

synthesized cytoplasmic pyrrolysine.[57-59] The production of recombinant MtmB together with the discovered pyrrolysyl-tRNA ($tRNA^{Pyl}$) and pyrrolysyl-tRNA synthetase (PylS) in *Escherichia coli* (*E. coli*), supplemented with chemically synthesized *L*-pyrrolysine, was sufficient to incorporate pyrrolysine at the UAG codon position. Canonical amino acids, including lysine, did not promote amber readthrough. MS analysis of produced proteins showed that pyrrolysine is recognized by the synthetase as a whole and incorporated into the growing polypeptide (Fig. 4).[60] This direct charging of pyrrolysine onto tRNA contrasts with selenocysteine, where a specialized tRNA is first loaded with serine, which is then converted to selenocysteine.[61-63]

The genes *pylB*, *pylC*, and *pylD* were identified near the *mtmB1* gene in *Methanosarcina* spp. and were suggested to play a role in pyrrolysine biosynthesis based on their sequence characteristics.[59,64,65] This was confirmed when recombinant expression of these genes in *E. coli* enabled pyrrolysine biosynthesis from metabolites available in *E. coli*.[66] Subsequent studies established that pyrrolysine was synthesized sequentially by PylB, PylC, and PylD from two equivalents of *L*-lysine.[67,68] The iron-sulfur (Fe-S) *S*-adenosylmethionine (SAM) enzyme PylB catalyzes the rearrangement of *L*-lysine to (3*R*)-3-methyl-*D*-ornithine (3MO).[69] PylC then ligates 3MO ATP-dependently to the ε-amino group of *L*-lysine producing the isopeptide *L*-lysine-Nε-(3*R*)-methyl-*D*-ornithine.[70] The biosynthesis is completed by PylD, which catalyzes the oxidative ring formation of the 3MO moiety.[71] The resulting *L*-pyrrolysine is then ready to be incorporated into methylamine methyltransferases via PylS and $tRNA^{Pyl}$ (Fig. 4).[60,72]

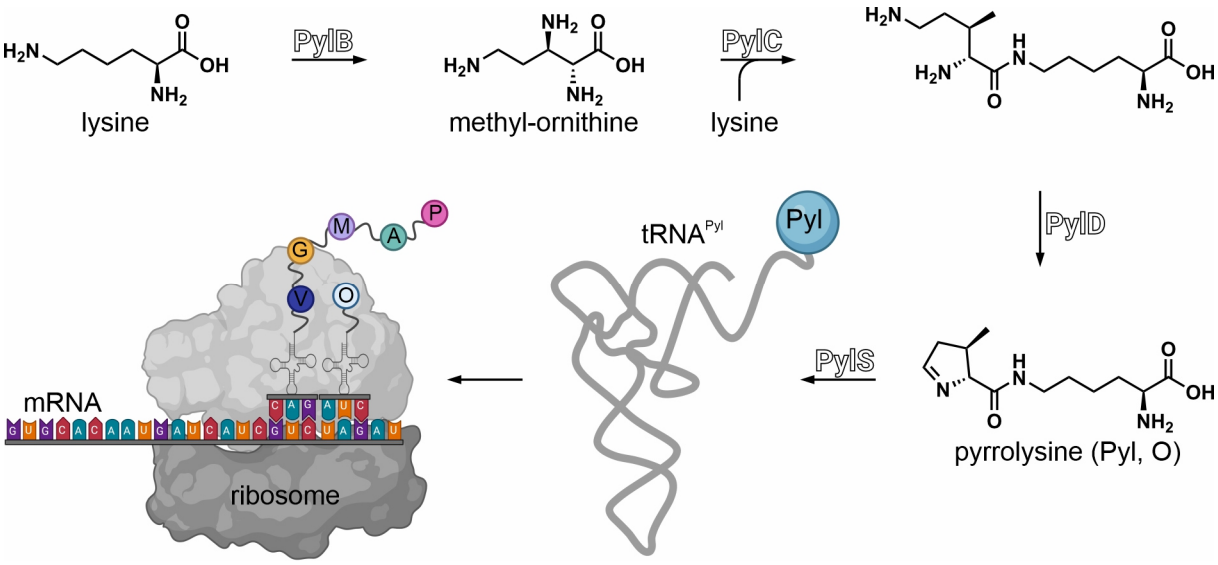


Figure 4 Scheme of the pyrrolysine (Pyl) biosynthesis and incorporation into proteins via PylB, PylC, PylD, the pyrrolysyl-tRNA synthetase (PylS), and the pyrrolysyl-tRNA ($tRNA^{Pyl}$). The tRNA and ribosome cartoons were created using BioRender.com.

1.2 Homologs without Pyrrolysine

Microbes that inhabit similar environments often attain a selection advantage through lateral gene transfer.[73-76] As a consequence, pyrrolysine usage is not unique to methanogenic archaea but was also detected in genomes of the distantly related bacterial classes Clostridia, Deltaproteobacteria, and Desulfovibrionia.[59,77] The presence of the pyrrolysine biosynthesis and incorporation machinery on a genome correlates with at least one encoded member of the methylamine methyltransferase families (homologs of MtmB, MtbB, or MttB) and vice versa. On the other hand, the use of the UAG codon for pyrrolysine incorporation does effectively not extend beyond MtmB, MtbB, and MttB.[77-79] Moreover, many genomes harbor homologs of these methyltransferases without amber-encoded pyrrolysine.[77] Relatives of MttB without an amber codon are particularly numerous. They are summarized as the COG5598 protein family and are widely distributed in many microorganisms as diverse as methanogenic archaea, marine bacterioplankton, nitrogen-fixing plant symbionts, or human gut microbes.[80] Interestingly, genes of pyrrolysine containing MttB members only encode a comparatively small branch within this large protein family.[56,64]

Since pyrrolysine is a key catalytic component of methylamine methyltransferases, the absence of this residue in MttB homologs raises questions about their functions.[77] Initial insights were gained by analyzing *mttB* relatives on the genome of the Clostridium *Desulfitobacterium hafniense* (*D. hafniense*), a strict anaerobe found both in the environment and in the human intestine.[59] The genome of *D. hafniense* encodes homologs of the pyrrolysine biosynthetic and translational machinery, *pylB*, *pylC*, *pylD*, *pylS*, and $tRNA^{Pyl}$, a 63 % similar homologous gene of the TMA methyltransferase *mttB* with a single in-frame amber codon, and three *mttB* homologs that don't contain amber codons but encode leucine or isoleucine at the amber codon position.[59] One of these homologs, MtgB, was found to allow *D. hafniense* to grow on the substrate glycinebetaine (*N*,*N*,*N*-trimethylglycine or GB).[80] Analogous to the TMA:corrinoid methyltransferase MttB, MtgB abstracts a methyl group from GB and transfers it to a dedicated corrinoid protein, MtgC. A third enzyme, MtgA, transfers the methyl group to tetrahydrofolate (THF) to form methyl-tetrahydrofolate (MeTHF).[80] Unlike methanogenic archaea, *D. hafniense* sequentially oxidizes the THF-bound methyl group to CO_2 to provide electrons for anaerobic respiration.[80] Following MtgB, the substrates of other non-pyrrolysine members of the MttB superfamily encoded on the genome of the anaerobic bacterium *Eubacterium limosum* (*E. limosum*) have been identified. As a common feature with MtgB, these demethylases also process trimethylated quaternary amine moieties. In total, three methyltransferases, designated MtpB, MtcB, and MtyB, were shown to allow growth on *L*-prolinebetaine, *L*-carnitine, and γ-butyrobetaine, respectively.[81-83]

The crystal structures of MtmB and MttB provided insights into the mode of action and the role of pyrrolysine in these enzymes.[54,55,84] However, structural data on the described non-amber MttB family members are scarce. The questions of how methyl transfer from the quaternary amine moieties is mechanistically accomplished and how the absence of pyrrolysine is compensated for in these demethylases remain unanswered. The analysis and comparison of some of these structures will therefore be a major subject of this work.

1.3 Cobalamin-dependent Methyl Transfer

Methyl transfer reactions are widespread in nature, but are carried out by only a few cofactors (Fig. 5). Methyltransferases, the enzymes that catalyze these reactions, can be divided into two classes that differ in their use of an activated versus an unactivated methyl group donor. The activated cofactor SAM is used for methyl transfer to various biomolecules such as DNA, proteins, and small-molecule secondary metabolites.[85] Its key characteristic is that the respective substrates are directly methylated (Fig. 5).[27] THF (Fig. 5) and its methanogenic analogues, tetrahydromethanopterin (THMPt) and tetrahydrosarcinapterin (THSPt), are unactivated methyl donors.[27] They are major one-carbon carriers in cells and are important for the nitrogen metabolism, as well as protein and DNA synthesis.[27,86] When unactivated methyl donors such as THF participate in methyltransferase reactions, they often require the assistance of the intermediate methyl carrier and strong nucleophile Cbl (Fig. 5).[27,87] Cbl is one of the most complex and versatile cofactors in nature. It consists of a tetrapyrrole ring that coordinates a central cobalt atom with its four equatorial pyrrole nitrogen

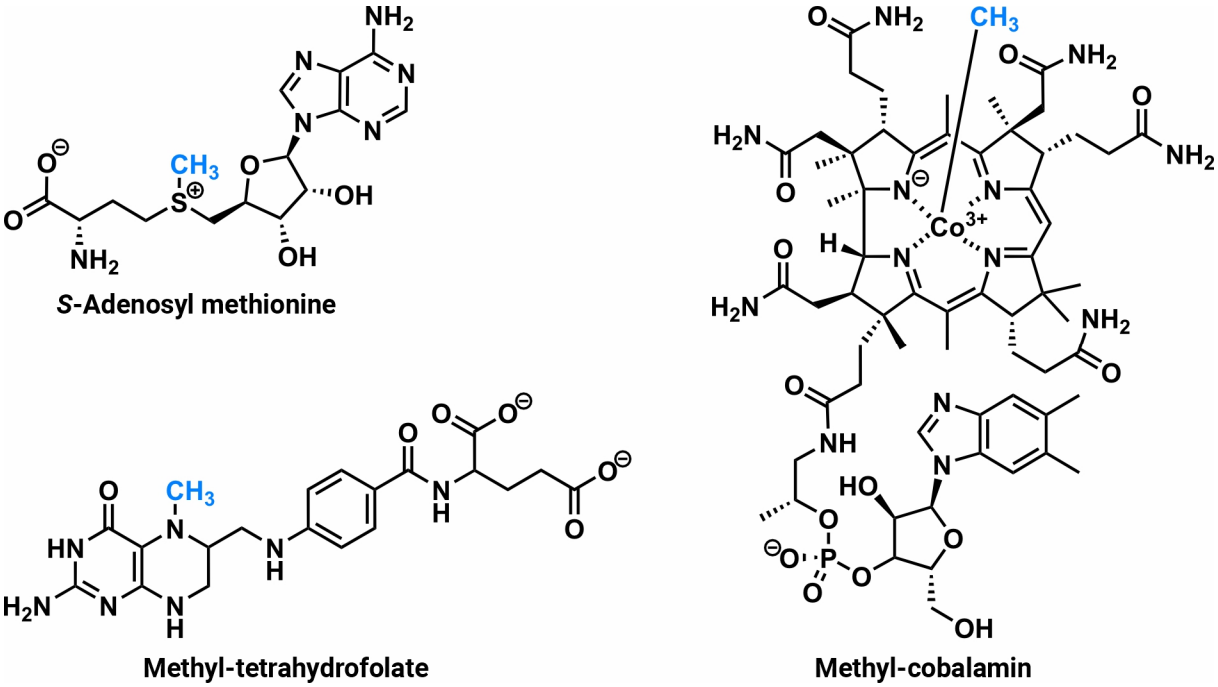


Figure 5 Structures of *S*-adenosylmethionine (SAM), methyl-tetrahydrofolate (MeTHF), and methyl-cobalamin (MeCbl). The methylation sites are shown in blue. The cobalt-methyl bond in MeCbl is stretched for better visibility.

atoms.[88] A pendant dimethylbenzimidazole (dmb) moiety extends from one side of the pyrrole ring and can serve as the lower axial ligand of the cobalt ion.[88] The conformations of the lower axial ligand are categorized as dmb-on and dmb-off. In the dmb-on position, the pendant dmb moiety binds to the cobalt ion. Conversely, the cofactor is categorized as dmb-off when the pendant dmb moiety is instead complexed by a corrinoid protein and the remaining coordination site of the cobalt ion is either left empty or replaced by a histidine residue of the protein.[27,87,89] Depending on the biological function, the upper axial ligand can vary between a cyano-, methyl-, or 5'-deoxyadenosyl group.[27,87] The versatility of this cofactor is reflected in its ability to catalyze both methyl transfer reactions as well as radical-intermediate mutase mechanisms.[90,91]

All Cbl-dependent methyltransferases are three-component enzymatic systems responsible for binding the methyl donor, the cofactor Cbl, and the methyl acceptor, respectively (Table 1). Cbl-dependent methionine synthase (MetH) occurs in both vertebrates as well as microbes and combines the components for methyl transfer from MeTHF via Cbl to homocysteine in a single peptide chain.[92] While MetH is the only Cbl-dependent methyltransferase in higher organisms, the microbial world contains an abundance of reactions that require Cbl or one of its numerous derivatives as a cofactor (Table 1).[93] Methanogens in particular rely heavily on the use of Cbl as an intermediate acceptor of unactivated methyl groups and as a precursor for MeCoM.[2,17] In the carbon fixating Wood-Ljungdahl pathway of acetogenic bacteria, the corrinoid iron-sulfur protein (CoFeSP) mediates methyl transfer from MeTHF to carbon monoxide to produce biomass in the form of acetyl-CoA.[94] Microbes have evolved related systems for the breakdown of various methylated compounds, such as methoxylated aromatics, chloromethane, and compounds with quaternary amine moieties (s. section 2.2) (Table 1).[80,95,96]

The domain rearrangements and mechanistic principles of Cbl-dependent methyl transfer of MetH and CoFeSP have been extensively studied.[97-103] Due to their variety, the many microbial Cbl-dependent methyltransferase systems have not been characterized to the same extent (Table 1). However, across all domains of life, cobalamin-dependent methyltransferases share a high degree of sequence and structural similarity regarding their corrinoid harboring domain, which allows for extensive knowledge transfer between all studies performed on these topics.[104,105] The donor and acceptor domains, on the other hand, vary greatly in sequence, size, and substrate choice. Exploring the similarities and differences of the wide range of microbial Cbl-dependent methyltransferases may help to uncover principles of transfer reactions not yet revealed by studies of MetH or CoFeSP and contribute to a better understanding of Cbl-dependent methyl transfer as a whole.

Table 1 Cbl-dependent methyltransferases

Methyltransferase components	Methyl Donor	Methyl acceptor	Biological system	Reference
MetH	MeTHF	Homocysteine	Methionine synthesis	Matthews *et al.*, 2001[106]
MtsA, MtsB	Methyl sulfides	CoM	Methanogenesis	Tallant *et al.*, 2001[107]
MtaB, MtaC, MtaA	Methanol	CoM	Methanogenesis	Hagemeier *et al.*, 2006[108]
MtmB, MtmC, MtbA	MMA	CoM	Methanogenesis	Krzycki, 2004[64]
MtbB, MtbC, MtbA	DMA	CoM	Methanogenesis	Soares *et al.*, 2005[56]
MttB, MttC, MtbA	TMA	CoM	Methanogenesis	Soares *et al.*, 2005[56]
MtrA-H	MeTHMPt	CoM	Methanogenesis	Gottschalk and Thauer, 2001[109]
AcsA-E (CoFeSP)	MeTHF	Ac-CoA synthase	Acetogenesis	Doukov *et al.*, 2007[100]
MtaABC	MeTHF	THF	Acetogenesis	Das *et al.*, 2007[110]
MtvB, MtvC, MtvA	Methylaromatics	THF	Acetogenesis	Naidu and Ragsdale, 2001[111]
CmuA, CmuB	Chloromethane	THF	Dehalorespiration	Studer *et al.*, 2001[112]
MtgB, MtgC, MtgA	Glycinebetaine	THF	Anaerobic respiration	Ticak *et al.*, 2014[80]
MtpB, MtqC, MtqA	Prolinebetaine	THF	Anaerobic respiration	Picking *et al.*, 2019[81]
MtcB, MtqC, MtqA	*L*-Carnitine	THF	Anaerobic respiration	Kountz *et al.*, 2020[82]
MtyB, MtqC, MtqA	γ-Butyrobetaine	THF	Anaerobic respiration	Ellenbogen *et al.*, 2021[83]

1.4 Objectives

This dissertation has three main objectives:

(i) The exotic amino acid pyrrolysine, which occurs in MtmB, MtbB, and MttB, is not incorporated naturally in response to amber codons in *E. coli*. In collaboration with the group of Prof. Kathrin Lang (ETH Zurich), this work aims to establish an amber suppression protocol for the incorporation of non-canonical amino acids (ncAAs) at the Chair of Biochemistry and use it for the incorporation of pyrrolysine into proteins. A further goal is to make this protocol applicable to other ncAAs in other projects of the group.

(ii) A number of projects requiring extensive manipulation of vector DNA are planned for this work. The heterologous co-production of the pyrrolysine biosynthesis gene cluster will be attempted to provide pyrrolysine for incorporation into proteins via amber suppression. To overcome problems such as low production or poor folding of proteins produced recombinantly in *E. coli*, a number of fusion-tagged as well as truncated variants of the enzymes in question will be tested. To conduct these cloning projects more efficiently, improvements to the Chair's molecular biology protocols are pursued. The molecular biology workflow will be based on the versatile Gibson Assembly® method and will include a digital library of the Chair's vectors, primers, and protocols.

(iii) The main objective of this work is to characterize the structural and biochemical interplay of archaeal and bacterial Cbl-dependent methyltransferases. In particular, the focus is on proteins involved in the methylamine metabolism of methanogenic archaea (Fig. 1, blue box). This group of metabolic pathways presents several intriguing research areas. It features the only three enzymes in nature that employ the exotic 22nd amino acid pyrrolysine. Additionally, there are numerous homologs without pyrrolysine found across microbial

organisms, raising questions about how these enzymes compensate for the absence of this catalytically important residue. Furthermore, the methyltransferases involved in this metabolism utilize various interesting cofactors for transferring or accepting substrate-derived methyl groups. Given the limited number of solved structures in the microbial methylamine metabolism, these factors motivate a comprehensive biochemical and structural analysis of the enzymes involved. The proteins of interest are sought to be heterologously produced in *E. coli* and purified using chromatographic methods. Crystallization trials of the prepared proteins are conducted to grow protein crystals suitable for 3D structure determinations via X-ray crystallography. Solved structures will be used for analyses of the proteins' reaction mechanisms and characteristic features. This process seeks to contribute to the fundamental research on the microbial methylamine metabolism.

2 Results and Discussion

2.1 Strategy for Heterologous Production of WT Methyltransferases in *E. coli*

The methyltransferases MtmB, MtbB, and MttB from *M. barkeri* were the first proteins investigated in this work. To purify the proteins, attempts were made to heterologously produce them in *E. coli*. Compared to purification from native organisms, which has been done successfully in some cases,[32,34,36,54,84] this approach promises a number of advantages. First, recombinant overexpression in *E. coli* can lead to high yields of the protein of interest, which is important for crystallography where high protein concentrations are required. In addition, the genetic material is much easier to manipulate, which facilitates affinity tag fusion, terminal truncation, and mutagenesis studies. Finally, members of the MttB superfamily stretch over a wide variety of organisms from different branches of life.[80] Rather than establishing appropriate cultivation conditions for each organism, adapting a purification protocol of recombinantly produced MttB may allow time-efficient access to structural analyses of a multitude of homologs.

Heterologous production in *E. coli* also has its drawbacks. On the one hand, protein yield and folding can be affected by the foreign cellular environment. In addition, archaeal methylamine methyltransferases require the incorporation of the 22nd amino acid pyrrolysine, which does not occur naturally in *E. coli*. In this work, these two issues are addressed separately. First, a protocol for heterologous biosynthesis and incorporation of pyrrolysine into green fluorescent protein (GFP) is presented. Then, recombinant protein production of alanine or lysine mutants of methylamine methyltransferases will be attempted. Ultimately, the goal is to combine both protocols to obtain the native methyltransferases.

Both protocols rely on the efficient generation of new vector constructs for protein production in *E. coli*. To improve this, the cloning workflow at the Chair of Biochemistry has been extensively updated. Several vector constructs encoding different variants of the methyltransferase MttB serve as a showcase to highlight the efficiency of the new setup.

2.1.1 Incorporation of Pyrrolysine

Heterologous production of wild type (WT) methylamine methyltransferases in *E. coli* is complicated by the absence of pyrrolysine production and incorporation in this organism. To the advantage of this project, PylS and tRNA^{Pyl}, due to their promiscuity and adaptability, have been widely established as an amber suppression method for heterologous incorporation of a wide variety of ncAAs including pyrrolysine itself.[66,68,113] Amber suppression protocols in *E. coli* for a specific ncAA require the production and supplementation of an appropriate tRNA-

synthetase, tRNA, ncAA, and a vector expressing the gene of interest with a specifically placed amber codon.[113] In the case of heterologous pyrrolysine incorporation, WT PylS and $tRNA^{Pyl}$ from *M. barkeri* can be used without further modifications.[66]

To test pyrrolysine incorporation independently of heterologous protein production of methylamine methyltransferases, superfolder GFP (sfGFP, UniProt-entry: P42212)[114] with an amber codon at residue position 149 was used as a reporter protein (Fig. 6). Successful incorporation of the ncAA can be verified via the characteristic green color of GFP as soon as the full-length protein is produced in sufficient amounts in the cell culture. Further experimental confirmation can be obtained via immobilized metal affinity chromatography (IMAC), size exclusion chromatography (SEC), sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), and liquid chromatography-mass spectrometry (LC-MS) analysis. The GFP construct was designed with a C-terminal His_6-tag to facilitate separation of gene products in which the amber codon led to termination rather than ncAA incorporation.

For the amber suppression experiments, a pRSF vector encoding WT PylS and $tRNA^{Pyl}$ from *M. barkeri* as well as the model protein of interest $GFP_{N149UAG}$ (vector construct 5b, Table 10) was kindly provided by colleagues from the research group of Prof. Kathrin Lang (Fig. 6) (ETH Zurich). However, the ncAA, pyrrolysine, is not commercially available and is often replaced by Boc-protected lysine (BocK) as a positive control for amber suppression experiments. One option is chemical synthesis of pyrrolysine, which has been successfully performed.[55,68] Alternatively, co-expression of the genes for the biosynthetic machinery for pyrrolysine, *pylB*, *pylC*, and *pylD*, has been shown to allow heterologous production of *L*-pyrrolysine from *L*-lysine and subsequent incorporation into the protein of interest.[66,68] Compared to chemical synthesis, co-translation of PylB, PylC, and PylD appears to be a more convenient approach to provide the quantities of pyrrolysine required during amber suppression in large scale *E. coli* cultures. Therefore, a pCDFDuet-1 vector encoding *pylB*, *pylC*, and *pylD* was cloned (Table 11, vector construct XYZ1-3c) (Fig. 6). *E. coli* BL21(DE3) were co-transformed with the vectors 5b as well as XYZ1-3c and cultivated. After induction with isopropyl-β-*D*-thiogalactoside (IPTG), all five genes for pyrrolysine biosynthesis and

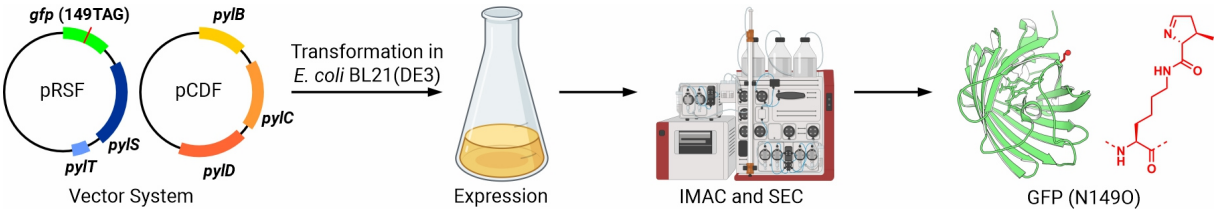


Figure 6 Scheme illustrating the co-expression of gfp_{149TAG}, *pylS*, *pylT* (encodes $tRNA^{Pyl}$), *pylB*, *pylC*, and *pylD* for the production and purification of GFP with incorporated pyrrolysine (vector constructs 5b and XYZ1-3c). The superfolderGFP PDB-entry (PDB-ID: 2B3P) was used for the cartoon illustration. The structural formula of pyrrolysine as well as the incorporation site are shown in red. Parts of this figure were created using BioRender.com.

incorporation were overexpressed together with the reporter gene *gfp*$_{149UAG}$. The full-length, C-terminally His$_6$-tagged GFP$_{N149UAG}$ was then purified in two steps by first isolating the tagged construct from the cell lysate and then performing SEC (Fig. 6 and S1). During purification, the distinct green color of GFP was a first indicator of successful pyrrolysine incorporation (Fig. 7A). Analysis via SDS-PAGE revealed a protein band at around 28 kDa for the full-length protein (Fig. 7B). The successful incorporation of pyrrolysine into GFP was confirmed via LC-MS analysis, which indicated a molecular mass of 27,817 Da (Fig. 7C). Taking into account the loss of 20 Da due to chromophore formation within the GFP barrel structure,[115] this mass is expected for the pyrrolysine-incorporated GFP construct encoded on vector 5b (Fig. 7C). The purification was repeated with a non-pyrrolysine-containing control of the GFP construct featuring an asparagine at position 149 (construct W1-1a, Table S1). As expected, LC-MS analysis for this GFP construct only indicated a molecular mass of 27,694 Da (Fig. 7C).

In summary, IPTG-induced pyrrolysine biosynthesis and incorporation into a model protein such as GFP in *E. coli* BL21(DE3) has been successfully emulated from literature.[66,68] Once recombinant soluble expression of MtmB, MtbB, or MttB in *E. coli* has been achieved, both protocols can be combined to produce the WT methyltransferases including pyrrolysine. A large library of PylS mutants exists that allows the incorporation of various ncAAs at the amber codon position.[116] The amber suppression system established at the Chair of Biochemistry can be adapted to incorporate other ncAAs, provided a suitable synthetase for the ncAA in question is available. To achieve this, *pylS* and *gfp*$_{149TAG}$ on vector 5b (Fig. 6) can be replaced by the synthetase mutant as well as the amber codon-interrupted gene encoding the protein of interest. The ncAA can be added during cultivation to allow production of the modified protein.

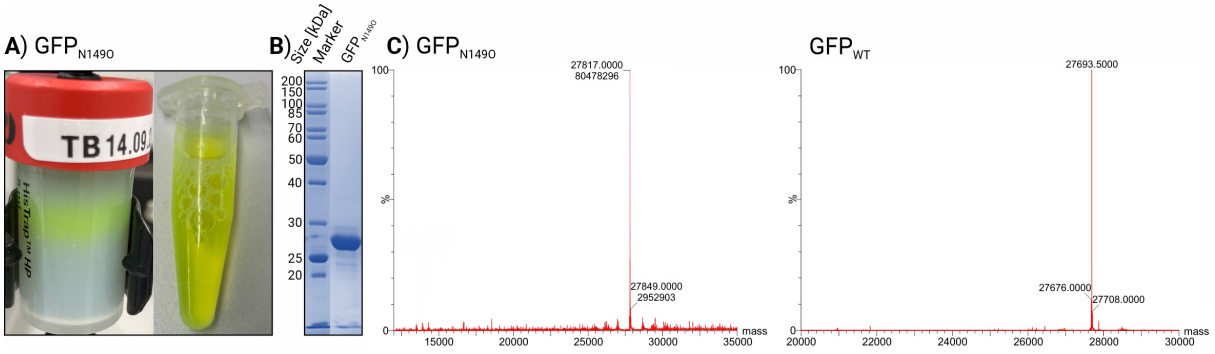


Figure 7 Purification and MS of pyrrolysine-containing and WT GFP. **A**) Pyrrolysine incorporation into GFP, rather than chain termination, was already evident during the purification of GFP$_{N149O}$ from the distinct green color of full-length GFP. **B**) The final purification product of GFP$_{N149O}$ ran at the correct size of 28 kDa during SDS-PAGE analysis. **C**) MS spectra of GFP$_{N149O}$ and GFP$_{WT}$. The determined masses of 27,817 and 27,694 Da are 123.5 Da apart, which is the difference between a pyrrolysine and an asparagine residue. Note that the expected total masses of both proteins are off by 20 Da due to a condensation and an oxidation reaction during the chromophore formation inside the β-barrel of GFP.

Pyrrolysine incorporation via co-translation of the biosynthetic machinery, as performed here, is more practical than supplementing large amounts of chemically synthesized pyrrolysine to the cultivation medium. Still, a considerable amount of optimization is required for large-scale pyrrolysine incorporation by biosynthetic co-expression. With the current protocol, the yield of pyrrolysine-containing GFP amounted to 3 mg of protein purified from 2 L of cell culture (Fig. 7A and S1). Repeating the same purification with WT GFP allowed the isolation of 180 mg (Fig. S2). Incorporation of the ncAA is not the bottleneck, as the addition of 2 mM of the pyrrolysine mimic BocK to the cultivation medium increased the yield of obtained amber-suppressed protein several-fold. In *E. coli*, common amber suppression protocols allow the production of high yields of protein[117] suggesting that biosynthesis via PylB, PylC, and PylD in *E. coli* does not produce enough pyrrolysine for efficient incorporation into GFP. Protein crystallization requires large amounts of protein making it worthwhile to try to optimize the cultivation setup for pyrrolysine incorporation.

Due to its susceptibility to oxygen, the Fe-S cluster-containing protein PylB has been proposed to be the rate limiting enzyme during heterologous pyrrolysine biosynthesis.[68] Previously demonstrated strategies such as reducing the oxygen stress on this enzyme by cultivating *E. coli* without flask agitation did not significantly improve amber suppressed protein production with the protocol used in this work.[68] Recombinant protein production of Fe-S proteins in *E. coli* has been shown to benefit from enhanced expression of an iron-sulfur assembly biosynthetic gene cluster.[118] Similarly, pyrrolysine production with the setup chosen here could be improved by enhanced production of functional PylB with a fully assembled Fe-S cluster.

Another approach could be to eliminate the need for PylB altogether by co-expressing *pylC* and *pylD* and adding the product of PylB, 3MO, to the culture. Like pyrrolysine, 3MO is not commercially available. Chemical synthesis has been performed before but requires a similar amount of effort than producing pyrrolysine itself.[68] As a compromise, *D*-ornithine can be used as the initial substrate for PylC and PylD, forming pyrroline-carboxy-lysine (Pcl, identical to pyrrolysine but without the methyl group attached to the pyrrole moiety) (Fig. S3).[68,119] The incorporation of this closely related pyrrolysine mimic into the methyltransferases may allow for easy production of large amounts of almost native methylamine methyltransferases, which may be sufficient for detailed characterization of the proteins.

Even if the attempts to optimize the protein yield fail, an adjustment of the crystallization strategy might still allow to obtain crystal structures of the methylamine methyltransferases from the small amounts produced in the experiment described above. Since mutation of one active site residue should not affect crystal contacts, crystallization conditions can be

optimized using mutant protein in which pyrrolysine is mutated to alanine or lysine. Once a suitable crystallization condition has been identified, crystal production can be reproduced on a small scale using WT (pyrrolysine-containing) protein.

2.1.2 Methylamine Methyltransferase Production in *E. coli*

In addition to amber suppressed pyrrolysine incorporation, protein production of the methyltransferases MtmB, MtbB, and MttB in *E. coli* has been extensively addressed in this work. To simplify the experimental conditions, expression and purification assays were performed with mutant protein versions in which pyrrolysine was mutated to alanine or lysine (Table 10). Typically, high amounts of recombinant methyltransferases were observed after overexpression in various strains of *E. coli* BL21(DE3). However, SDS-PAGE analysis showed that the proteins were located in the pellet fraction and could not be purified via NiNTA spin columns, indicating that they were not correctly folded during production in *E. coli* (Fig. 8A). Similarly, the addition of a SUMO-tag (Smt3p from *Saccharomyces cerevisiae*)[120] failed to produce soluble protein. Unless otherwise stated in this section, these results were consistent across all purifications of MtmB, MtbB, and MttB, regardless of the vector constructs, *E. coli* expression strains, and buffer conditions tested.

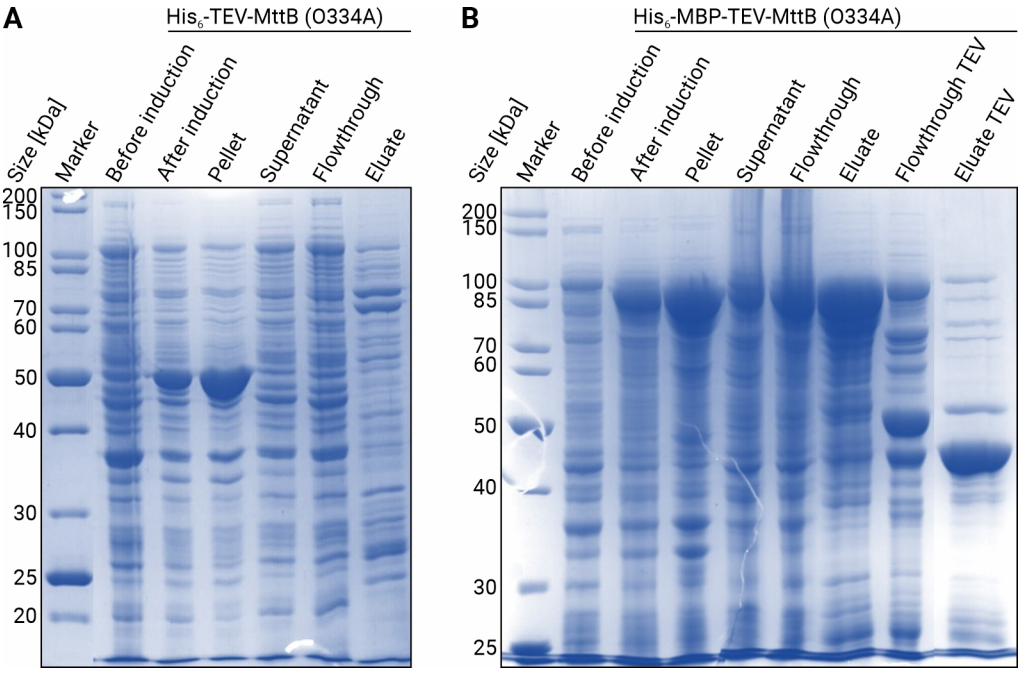


Figure 8 SDS-PAGE of exemplary purifications of methylamine methyltransferases produced in *E. coli*. Samples for SDS-PAGE were collected throughout the purification processes as indicated. **A**) MttB O334A mutant protein with an N-terminal His_6-tag (Vector construct F10-2a) was heterologously produced in *E. coli* BL21(DE3) Tuner cells and purified via a NiNTA spin column. Unless otherwise stated, the results shown here were consistent for all purifications of MtmB, MtbB, and MttB, regardless of the vector constructs, *E. coli* expression strains, and buffer conditions. **B**) MttB O334A mutant protein with an N-terminal His_6-TEV-MBP tag (vector construct F19-1c) was heterologously produced in *E. coli* BL21(DE3) cells, purified via a NiNTA spin column, digested overnight with TEV protease, and separated again via a NiNTA spin column.

MttB, the most interesting of the three proteins due to its large number of non-amber relatives, was analyzed more extensively than MtmB or MtbB. A notable finding was that the addition of an MBP-tag to the N-terminus of MttB was sufficient to allow protein pull-down via IMAC (Fig. 8B). Successful cleavage of the tag at the incorporated TEV site was verified via SDS-PAGE (Fig. 8B). However, MttB strongly precipitated in subsequent purification steps, suggesting that the solubility conferred by the correctly folded MBP tag was sufficient to keep the unfolded C-terminal MttB chain in solution.

2.1.3 Improving the Cloning Infrastructure for Advanced Methyltransferase Constructs

Modification of the DNA template encoding the protein of interest is one of the most powerful tools to potentially improve proper folding, solubility, as well as crystallization.[121] Therefore, after testing the production of several basic fusion-tagged versions of the full-length methyltransferases, the focus was shifted to the design of more specific constructs. To facilitate this project, the cloning infrastructure at the Chair of Biochemistry at the TUM was upgraded. The default method used was restriction and ligation-based cloning,[122] which was a reliable method for conventional full-length and fusion-tagged constructs, but proved cumbersome when used for more ambitious cloning projects. To improve this, Gibson Assembly®[123] was implemented as the new default method for generating vector constructs. This method produces a higher rate of positive *E. coli* colonies, is restriction site-independent, which is particularly useful when working with genomic DNA, and allows multiple DNA fragments to be assembled simultaneously. As a result, much time can be saved during vector production, allowing planning and execution of more challenging cloning projects.

Construct planning and the preparation of new vectors via Gibson Assembly® is complex and requires clear documentation of the conducted cloning operations. For this purpose, the program SnapGene® was implemented as the central resource for molecular biology. This program's ability to generate concise annotated vector maps was used to create a digital library of the group's most commonly used vectors and make them accessible via a shared

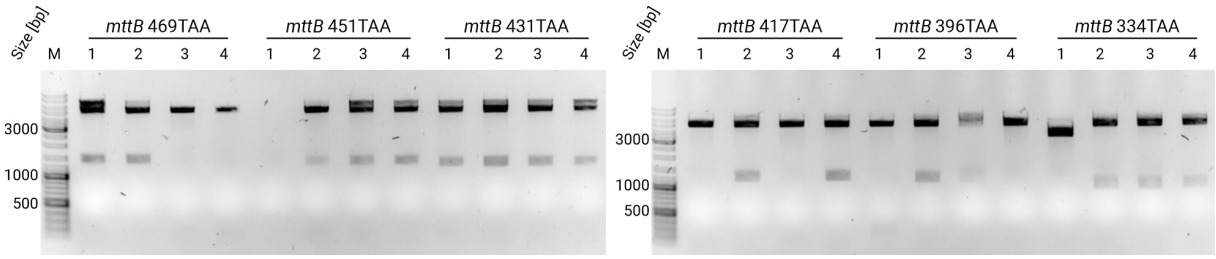


Figure 9 24 Minipreps were sufficient to complete the screening phase of the cloning process for all six MttB truncation variant vector constructs (F12-1f, F13-1f, F14-1f, F15-1f, F16-1f, and F17-1f). For each variant, four colonies were picked and cultivated in LB medium. The contained plasmid DNA was miniprepped, control digested with NcoI and XhoI, and visualized on a 1 % agarose gel. Positive (insert-containing) vectors were identified via the band at 1000 to 1500 bp.

network drive. The corresponding vectors were made available in the group's shared freezer. These vector maps formed the basis for new cloning operations using Gibson Assembly®. The program SnapGene® was used to plan new DNA modifications and to design the necessary primers. Finally, a library of standardized primers for Sanger sequencing and for the amplification of vector backbones was created to reduce the amount of primers necessary to be purchased.

One of the projects realized with this new molecular biology workflow was the construction of a series of vectors encoding MttB variants with C-terminal truncations designed to improve protein production. The execution and results of this project are presented here to demonstrate the capabilities of the new setup. The idea for MttB constructs with C-terminal truncations came from failed amber suppression experiments that were initially performed to test pyrrolysine incorporation into MttB (as performed with GFP in section 2.1.1). MttB variants in which translation was stopped at the UAG codon at residue position 334 appeared to be soluble under standard buffer conditions (Fig. 10A). Truncation at this position results in a version of MttB lacking the critical pyrrolysine residue. Structural information obtained from an AlphaFold prediction model of the structure of MttB indicated that extended variants of MttB ending at around residue 370 would form an intact active site.[124] Based on these results, a total of five vectors encoding C-terminal truncation variants

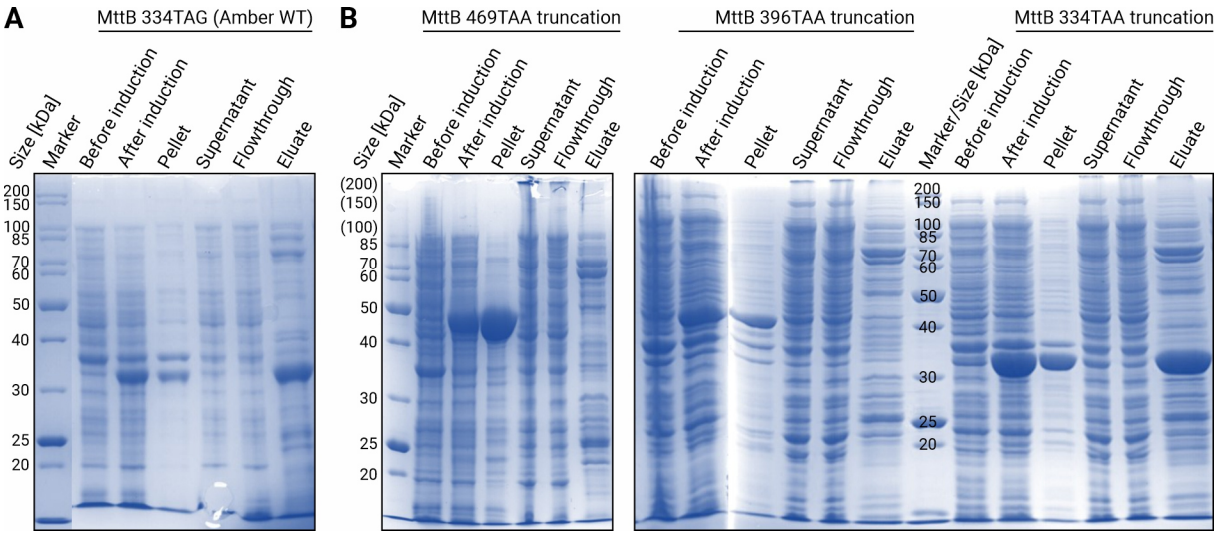


Figure 10 C-terminal truncation studies of MttB based on a failed amber suppression experiment. The respective MttB variants were heterologously produced in *E. coli* BL21(DE3) cells, purified via NiNTA spin columns, and analyzed via SDS-PAGE. **A**) This failed amber suppression experiment provided the idea for the C-terminal truncation variants. MttB (vector construct F8-5b) was terminated at amino acid position 334, but was eluted in soluble form from the NiNTA spin column at the expected size of 36 kDa. **B**) Results for three out of six MttB truncation variants are shown. Chain termination at residue position 469 (469TAA, vector construct F12-1f) and 396 (MttB 396TAA, vector construct F16-1f) resulted in insoluble protein in the pellet fraction at the corresponding size. Constructs with sizes in between behaved similarly and are not shown here. In contrast, the positive control, which was terminated at position 334 (MttB 334TAA, vector construct F17-1f), allowed the elution of soluble protein from the NiNTA spin column.

of MttB were created, each cropped by approximately 20 residues relative to the previous one. Using AlphaFold prediction, the truncations were placed in random coil regions.

Vector maps of the constructs F12-1f, F13-1f, F14-1f, F15-1f, F16-1f, and F17-1f were designed in SnapGene® (Table 10). The generated primers (Table 11) allowed PCR amplification of the six inserts of different lengths as well as the pETDuet vector backbone. The amplified DNA fragments were used in the assembly reactions. Assembled vectors were transformed into *E. coli* DH5α. Four colonies per construct were picked, cultivated, miniprepped, and control-digested to identify vectors with the correct insert size via agarose gel electrophoresis. The sample size was sufficient to obtain at least one positive representative in each case (Fig. 9). The correct sequence was confirmed via Sanger sequencing. Detailed overviews of the cloning steps can be found in the cloning history of the vector maps of the constructs in the Chair's shared network drive (\\europa.intern.ch.tum.de\BC\Archiv\Thomas Badmann). The expression of the generated vectors was tested similarly to section 2.1.2. Unfortunately, except for the specifically designed positive control (334TAA, vector construct F17-1f), which was truncated after position 333, none of the variants resulted in soluble protein (Fig. 10B). Still, the presented data highlight the efficiency of the updated cloning setup. Including cloning, cultivation, isolation, and SDS-PAGE analysis, they show the results of only two weeks of work.

Further efforts on this project were interrupted by the achievements of Li *et al*. who managed to obtain crystal structures of natively produced MttB (PDB-IDs: 7XCL, 7XCM, and

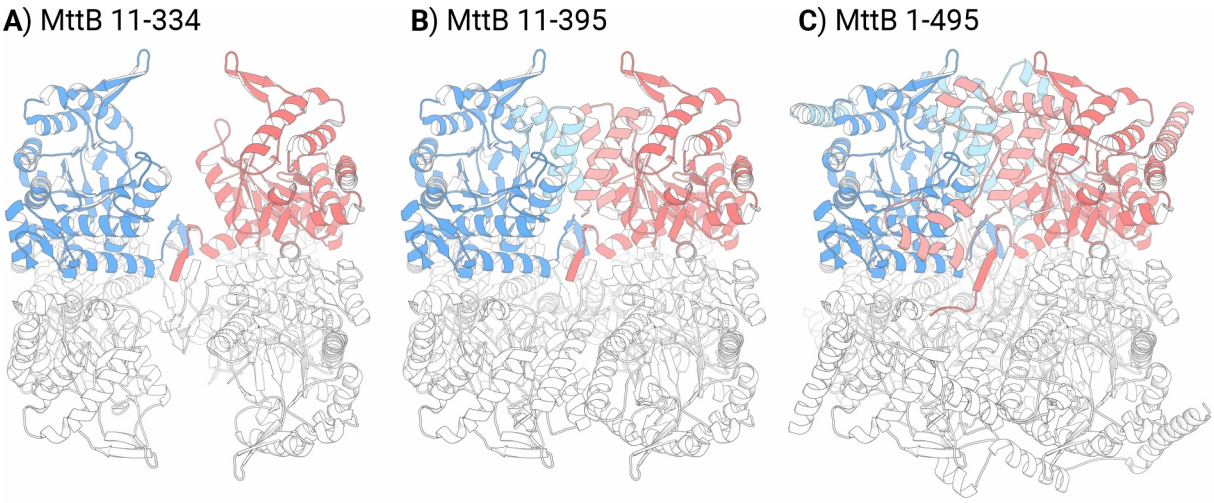


Figure 11 Structural analysis of the MttB truncation variants. Residue segments of each chain of the cartoon structure representations of MttB (PDB-ID: 7XCL) were shown and hidden as indicated in the headings. Subunit A is shown in blue and subunit B in red. The C-terminal region starting at position 334 (where all truncations were made) are highlighted in lighter shades of the respective color. The remaining subunits are shown in white. **A)** Positive control construct (MttB 334TAA, vector construct F17-1f) truncated at the pyrrolysine residue position. Note the lack of interaction interfaces compared to B. **B**) Truncation variant with the shortest C-terminus analyzed in this work (MttB 396TAA, vector construct F16-1f). A large part of the interface between the subunits is intact again. In contrast to A, the construct is not soluble after production in *E. coli*. **C**) Full-length cartoon representation of the MttB structure.

7XCN).[84] Nevertheless, the insights gained from the solved structures provided an interesting context to the performed experiments. MttB assembles as a homohexamer with the N- and C-termini involved in complex formation (Fig. 11A, B and C). Cropping the protein at position 333 results in loss of most of the interaction interface, which may explain the improved solubility of the shortest truncation construct (Fig. 10A, B and 11A). Longer variants recover the interaction surface, which could lead to the inability of the proteins to assemble correctly during production in *E. coli* (Fig. 10B and 11B). All constructs analyzed in this work carried N-terminal fusion tags. Since the immediate N-terminus of the protein is close to the oligomerization interface, working without an N-terminal fusion tag might support proper folding in *E. coli* (Fig. 11A, B and C). The *M. barkeri* TMA methyltransferase MttB belongs to a large protein family with a variety of interesting candidates for biochemical and structural characterization spread across a wide range of microorganisms.[80-83] The insights gained from the MttB structure may be applicable to other family members that display similar issues with heterologous production in *E. coli*. Although applicability to the methyltransferase project has yet to be shown, the improved cloning setup benefits all members of the Chair of Biochemistry allowing them to plan and perform more ambitious and complex experiments.

2.2 Structural Comparison of MtgB and MtcB with Pyrrolysine-containing Relatives

2.2.1 Structural Analysis of the Methyltransferase MtgB

MtgB was cloned and recombinantly produced in *E. coli*. The protein was purified via IMAC and SEC (Tables 10, 11, and 12, Fig. S5 and S6) and co-crystallized with different substrate combinations of GB, DMG, Cbl, and methyl-cobalamin (MeCbl). This approach led to eight structures of MtgB: MtgB:GB (2.15 Å), MtgB:DMG (2.0 Å), MtgB:Cbl (2.2 Å), MtgB:MeCbl (2.6 Å), MtgB:GB:Cbl (2.3 Å), MtgB:DMG:Cbl (2.1 Å), MtgB:DMG:MeCbl (2.35 Å), and MtgB:Bic (2.0 Å) (Tables S1, S2, and S3). The latter is the product of a crystallization condition containing 100 mM bicine pH 9.0, in which a bicine molecule was trapped in the active site of MtgB.

MtgB adopts a TIM-barrel fold consisting of an alternating series of β-sheets on the inside and α-helices on the outside (Fig. 12).[125-127] A detailed inspection of the structure as well as crystal packing analysis with the webtool PISA indicated that MtgB forms a homodimer (Fig. 12).[128] A search for structural homologs using the Protein Structure Comparison Server DALI[129] revealed a close relationship with two sequence homologs: a non-pyrrolysine member of the MttB family of unknown function (PDB-ID: 4YYC, Z-score: 59.5, RMSD (root-mean-square deviation): 1.3 Å, sequence identity: 43 %) and MttB from *M. barkeri* (PDB-ID: 7XCL, Z-score: 53.5, RMSD: 1.8 Å, sequence identity: 30 %).[84] MtmB (PDB-ID: 1NTH, Z-score: 25.0,

RMSD: 2.7 Å, sequence identity: 9 %)[54] was identified as a structural homolog with low sequence identity. Besides methylamine methyltransferases, other distantly related Cbl-dependent methyltransferases such as the THF-binding domain of MetH from *Thermotoga maritima* (PDB-ID: 1Q8J, Z-score: 17.2, RMSD: 3.1 Å, sequence identity: 9 %)[98] and the small subunit of CoFeSP from *Moorella thermoacetica* (PDB-ID: 4DJE, Z-score: 17.2, RMSD: 3.2 Å, sequence identity: 13 %)[102] can be superimposed on the TIM barrel structure of MtgB.

The active site of MtgB is located within the cavity formed by the barrel structure. The net negative charge of the cavity is contrasted by two positively charged residues, Arg312 and His348, located directly at the bottom (Fig. 13). Asn202, Thr275, Arg312, and His348 form hydrogen bonds with the carboxy group of the substrate (Fig. 13). Leu200 and Leu232 are positioned to engage in lipophilic interactions with the backbone and methyl groups of the substrate (Fig. 13). Similar to the crystal structures of unrelated GB binding enzymes, the aromatic residues, Tyr97 and Phe356, are involved in π-stacking interactions with the methyl groups of GB (Fig. 13).[130-135] One of the three methyl groups of the substrate is directed toward the solvent and is perfectly oriented for methyl transfer from GB to Cbl.

The Cbl-dependent methyltransferase MtgB is predicted to transfer a methyl group to the Cbl of its associated corrinoid partner protein MtgC (Table 1). MtgB and MtgC are

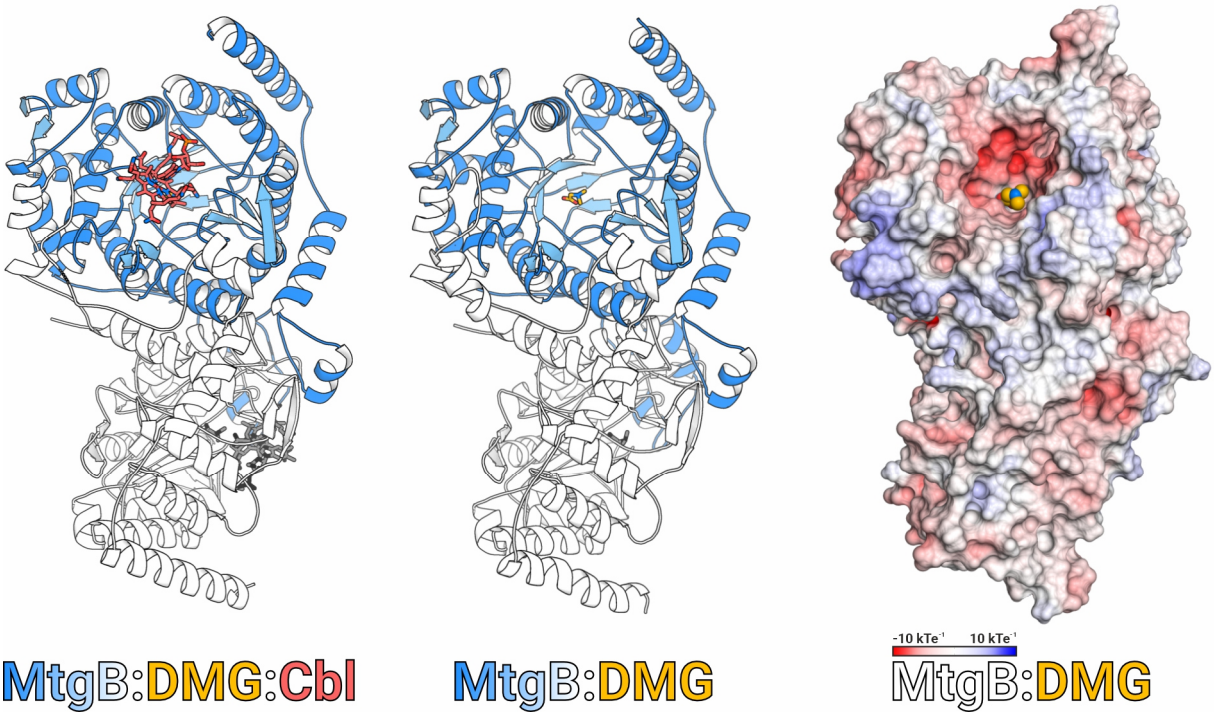


Figure 12 Cartoon and surface charge representations of MtgB. Cartoons are shown for MtgB:DMG:Cbl and MtgB:DMG. In both structures, subunit A is colored blue and β-sheets are highlighted in light blue. Dimethylglycine (DMG) and cobalamin (Cbl) are shown as sticks (C-atoms in yellow and coral, respectively; O-atoms and Co: red, N-atoms: blue, P-atoms: orange, corresponding molecules in subunit B are depicted in grey and black). On the right, a surface charge representation of MtgB:DMG is displayed with DMG represented as spheres (color coding analogous to sticks).

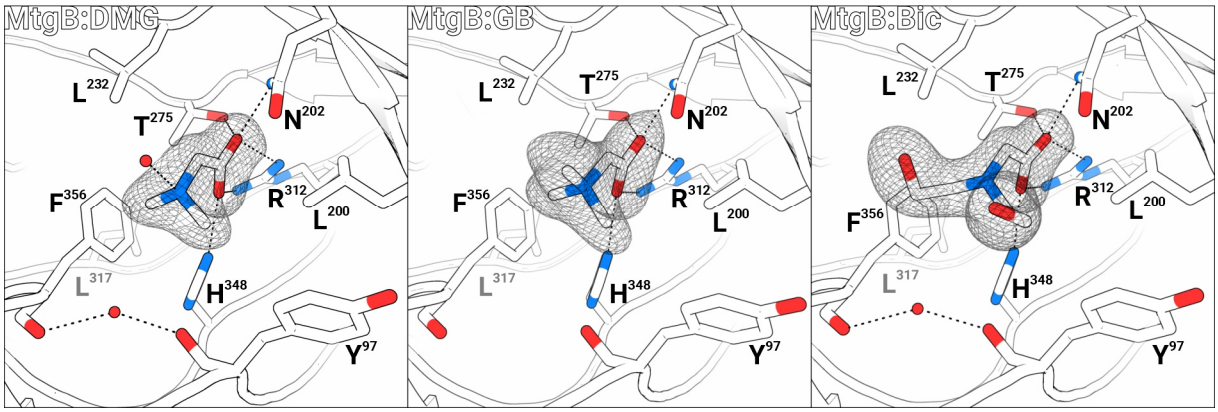


Figure 13 Cartoon and stick representation of the active site of MtgB in complex with DMG (2.0 Å), GB (2.15 Å), and bicine (2.0 Å). The cartoon and C-atoms are displayed in white, oxygen in red, and nitrogen in blue. Waters are depicted as red spheres, hydrogen bonds as dashed lines. The $2F_O$-F_C density is shown for the bound molecules. Note that MtgB:GB is not a finalized structure, it only serves to highlight the electron density of the additional methyl group.

produced from the same gene cluster on the *D. hafniense* genome[136,137] and are upregulated in the presence of the methyl source GB,[80] strongly suggesting their metabolic connection. Crystallization trials containing MtgB, MtgC, and supplemented Cbl produced pink protein crystals under various conditions with $(NH_4)_2SO_4$ (Table S1 and Fig. S13). However, data acquisitions and evaluations of these crystals revealed a complex structure in which only Cbl was bound to MtgB. To obtain the MtgB structures presented in this work, the experiment was repeated without MtgC, but with several different combinations of the substrate and the cofactor with or without the methyl group (Tables S1, S2, and S3). The obtained structures of MtgB display a trapped Cbl or MeCbl in dmb-on position (Fig. 14 and 15). They allow detailed insights into the interplay between the active site of MtgB and free Cbl without its corrinoid carrier protein MtgC (Fig. 14 and 15). This configuration will be referred to as exogenous Cbl. Structures of Cbl-dependent methyltransferases usually show the cofactor in dmb-off conformation, where Cbl is buried within the Rossman fold of the corrinoid carrier protein (endogenous Cbl).[84,97,99,101,102,138-145] The structures of MtgB in complex with Cbl or MeCbl stand out from this ensemble. They are the first solved structures of a methyltransferase bound to exogenous Cbl. Notably, Cbl-dependent methyltransferases, including MtgB, are usually active with both endogenous and exogenous Cbl.[37,80-82,146-150] Even MetH, which combines all methyltransferase domains for methyl transfer from MeTHF via Cbl to homocysteine in a single peptide chain, performs efficient catalysis with exogenous Cbl and without the help of the corrinoid Cbl-carrier component.[151] Nevertheless, binding of the Cbl-carrier component influences the reaction. Studies suggest that binding the cofactor dmb-off with an associated corrinoid carrier protein allows fine tuning of reaction parameters such as conversion rate and the direction of the methyl transfer reaction through conformational shifts as well as pH.[27,43,97,152,153] In MetH, binding and release of the catalytically important histidine

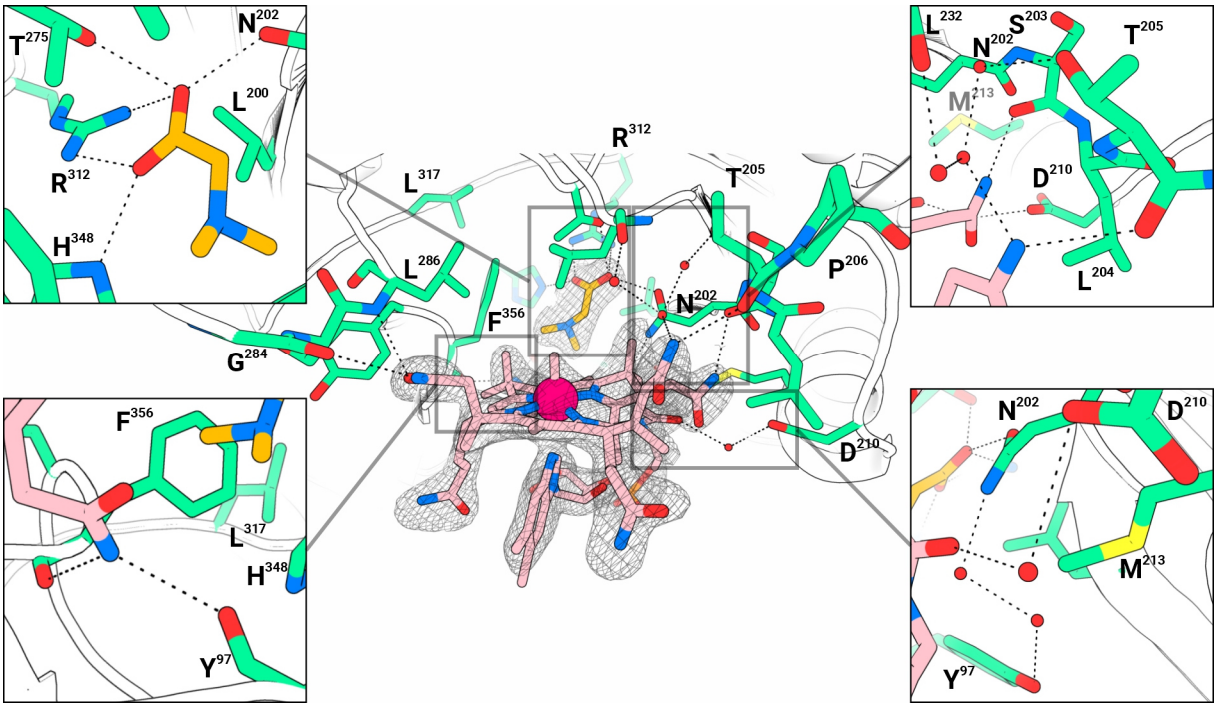


Figure 14 Cartoon and stick representation of the active site of MtgB:DMG:MeCbl. For DMG and MeCbl, the $2F_O$-F_C density is shown in the central figure. Carbon atoms of MtgB, DMG, and MeCbl are colored green, gold, and pink, oxygens in red, nitrogens in blue. Waters are depicted as red spheres, hydrogen bonds as dashed lines, and the cobalt ion as a large pink sphere. The cartoon is drawn in white. Selected amino acid residues are labeled by one-letter-code. The central figure shows the active site in its entirety. For better visualization, four areas (rectangles) are displayed separately in a zoomed view and in different orientations. Only some residues are labeled in the central figure for better clarity.

residue of the Cbl-carrier component is associated with inducing conformational changes between different domains of the protein.[141]

The obtained structures of MtgB display the upper face of the corrin ring of the cofactor sitting on the active site. The pendant benzimidazole moiety is exposed to the solvent and distally bound dmb-on to the cobalt ion of the corrin ring. The cofactor is coordinated by MtgB through an extensive hydrogen-bonding network involving Tyr97, Ser203, Thr205, Asp210, Leu232, Gly284, Leu286, and Phe356. Furthermore, lipophilic interactions with the corrin ring are exerted via Leu204, Leu232, Leu286, and Phe356. It is striking that hydrogen-bond interactions with Cbl are mainly formed via the carbonyl backbone of the protein. Of the residues mentioned, only Asp210 interacts exclusively via its side chain. Ser203 and Thr205 exercise both side chain and backbone interactions with Cbl. In addition, several water molecules are responsible for hydrogen bonds between the cofactor and the active site residues (Fig. 14).

In the MtgB structures containing both the substrate and the cofactor with or without methyl group, the nitrogen atom of the substrate (demethylation site) and the cobalt ion (methylation site) are 8.0 Å apart. This distance is too large to allow methyl transfer between both moieties suggesting that the Cbl cofactor must undergo a structural rearrangement within the active site to allow catalysis. The active site water involved in binding of Cbl is

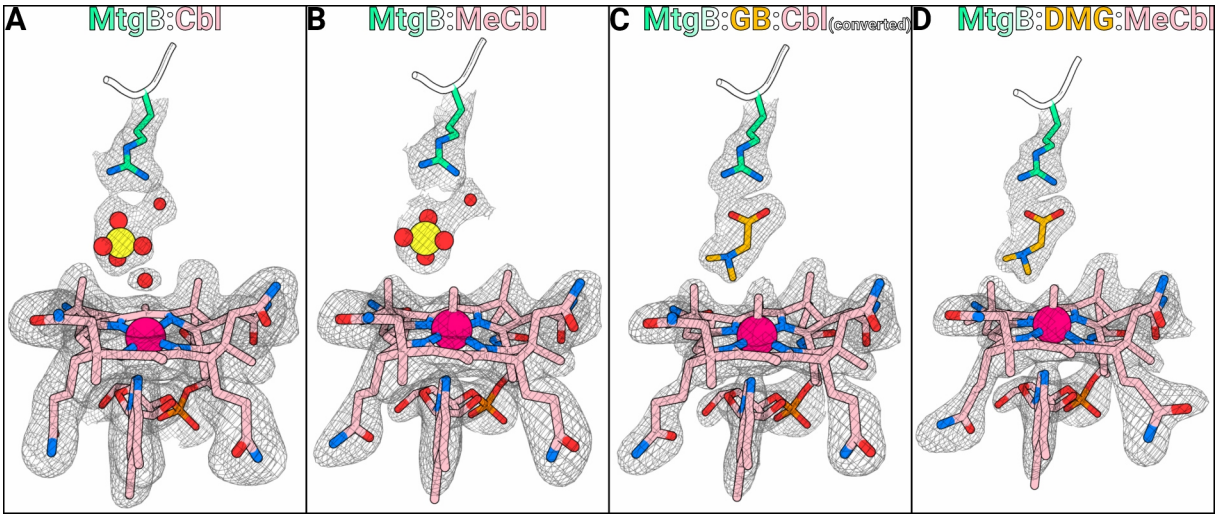


Figure 15 Comparison of the active site $2F_O$-F_C electron densities of the structures MtgB:Cbl (**A**), MtgB:MeCbl (**B**), MtgB:GB:Cbl (**C**), and MtgB:DMG:MeCbl (**D**). Arg312 is shown as the only residue of MtgB. Carbon atoms of MtgB, DMG, and Cbl/MeCbl are shown in green, gold, and pink, respectively. Oxygens are depicted in red, nitrogens in blue. Waters are shown as small red spheres, the cobalt ion as a large pink sphere. Sulfates are depicted as a combination of mid-sized yellow and red spheres.

located exclusively on one side of the tetrapyrrole ring. It may be that this water can be pushed out of the active site to allow tilting of the ring further into the active site to bring the substrate and the Cbl cobalt ion into methyl transfer distance (Fig. 14).

In the MtgB:Cbl structure, the active site is occupied by a sulfate molecule derived from the crystallization buffer (Fig. 15A) (Table S1). The electron density shows that the cobalt ion of Cbl binds a water molecule at its methylation site. Reproduction of the same structure co-crystallized with MeCbl results in a similar structure, but clearly shows a methyl group covalently bound to the cobalt ion (Fig. 15B). Interestingly, attempts at co-crystallization with GB and Cbl resulted in a structure containing DMG and MeCbl, indicating that the protein was active and successfully performed the methyl transfer reaction during the crystallization process (Fig. 15C). The co-crystallization of MtgB, DMG, and MeCbl served as a control experiment and produced a dataset with a similar electron density (Fig. 15D).

The gathered data allow to propose a comprehensive reaction mechanism for MtgB. The substrate GB is coordinated by Asn202, Thr275, Arg312, and His248 through interactions with its carboxy group (Fig. 19C). In this position, the methyl groups of the substrate are presented to the incoming Cbl. In all cofactor-bound MtgB structures, Cbl or MeCbl is located at a similar position 8.0 Å away from the demethylation site (Fig. 15). It is possible that Cbl displaces excess water from the active pocket to approach the substrate. Once the methyl group has been transferred, the remaining DMG can be exchanged for a new GB to restart the reaction cycle (Fig. 19C).

2.2.2 Structural Analysis of the Methyltransferase MtcB

The methyltransferase MtcB was obtained by cloning and recombinant expression in *E. coli* followed by purification using IMAC and SEC (Tables S1 and S4, Fig. S7). Successful crystallization of the protein resulted in an apo structure with a resolution of 1.85 Å (Tables S1 and S2). MtcB displays a TIM barrel fold and assembles as a homodimer (Fig. 16).[128] The protein is a homolog of MtgB and a search for structural homologs using DALI returned similar results.[129]

The active site was identified by comparison with MtgB (Fig. 13 and 18B). Like MtgB, the active site has a negative net charge and a prominent central arginine residue ($Arg318_{MtcB}$) (Fig. 16). Analogously to its counterpart $Arg312_{MtgB}$, $Arg318_{MtcB}$ is probably responsible for complexing the carboxy group of the substrate *L*-carnitine. Other residues involved in substrate binding are difficult to identify from the apo structure. $Ser282_{MtcB}$ is similar to its counterpart $Thr275_{MtgB}$ and could analogously be involved in the complexation of one of the functional groups of the substrate. Otherwise, the proteins show few specific similarities with respect to their active site composition. These differences probably accommodate the two extra carbon atoms and the extra hydroxy group that *L*-carnitine has compared to GB (Fig. 16, 18A and B).

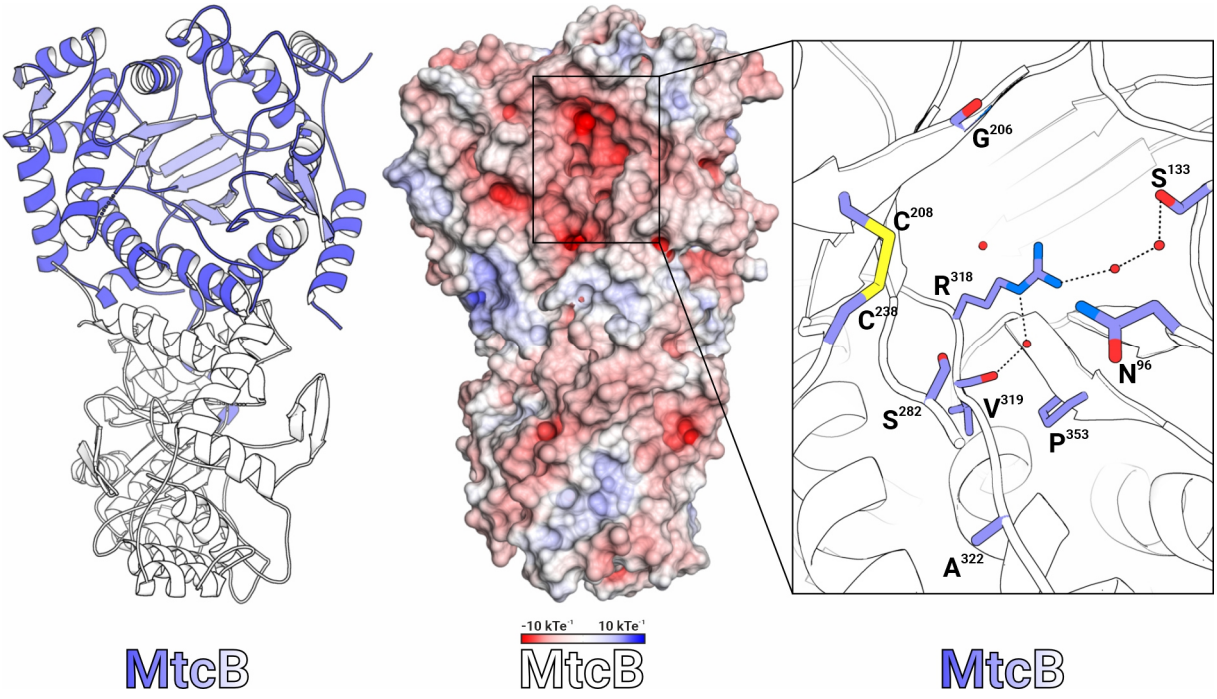


Figure 16 Cartoon, surface charge and stick representations of MtcB. In the cartoon, subunit B is colored in purple with β-sheets highlighted in light purple. Subunit A is colored in white. On the right, the active site is shown as white cartoon with colored sticks for selected residues (carbon: purple, oxygen: red, nitrogen: blue, sulfur: yellow). Residues are labeled by one-letter code. Waters are depicted as red spheres. Hydrogen bonds are drawn as dashed lines. All three figures display the protein in the same orientation.

2.2.3 Structural Comparison of Methyltransferases with and without Pyrrolysine

MtmB and MttB share pyrrolysine as the main catalytic component but lack sequence homology. On the other hand, MtgB and MtcB show sequence similarity to MttB, but do not share a pyrrolysine residue. The structures of MtmB[54,55] and MttB[84] have been solved but have not yet been compared with non-pyrrolysine quaternary amine methyltransferases. The structures of MtgB and MtcB solved in this work allow such a comparison and show how the absence of pyrrolysine is compensated for in these enzymes.

Besides the absence of pyrrolysine, the oligomeric assemblies of MtgB and MtcB mark a major difference to the archaeal methylamine methyltransferases MtmB and MttB. Unlike MtgB and MtcB, MtmB and MttB each consist of a homohexamer (Fig. 17).[36,54,55,84] Like MtgB and MtcB, the subunits of MtmB and MttB display a TIM barrel fold that forms the active site cavity. The pyrrolysine residue is located in the center of this cavity (Fig. 3, 18C, D, and E). Interestingly, isolated MtbB was analyzed to be composed of two 50 kDa subunits, thus diverging from the hexameric assemblies of the other two pyrrolysine-containing methyltransferases.[154]

For MtmB, an apo and an amine-bound structure were solved at 1.55 and 1.70 Å, respectively.[54] The structures display two different orientations of the pyrrolysine residue, which are characterized by a 180° turn of the pyrrole ring. While apo pyrrolysine (PDB-ID: 1NTH) shows more occupancy within the first orientation (Fig. 18D), amine-bound pyrrolysine (PDB-ID: 1L2Q), which represents a late reaction intermediate after abstraction of the methyl group from methylamine, occupies the second orientation more extensively (Fig. 18E).[54] Mechanistically, this orientation shift was proposed to position the added and activated methylamine moiety for efficient transfer to the corrinoid cosubstrate of MtmC.[54,64]

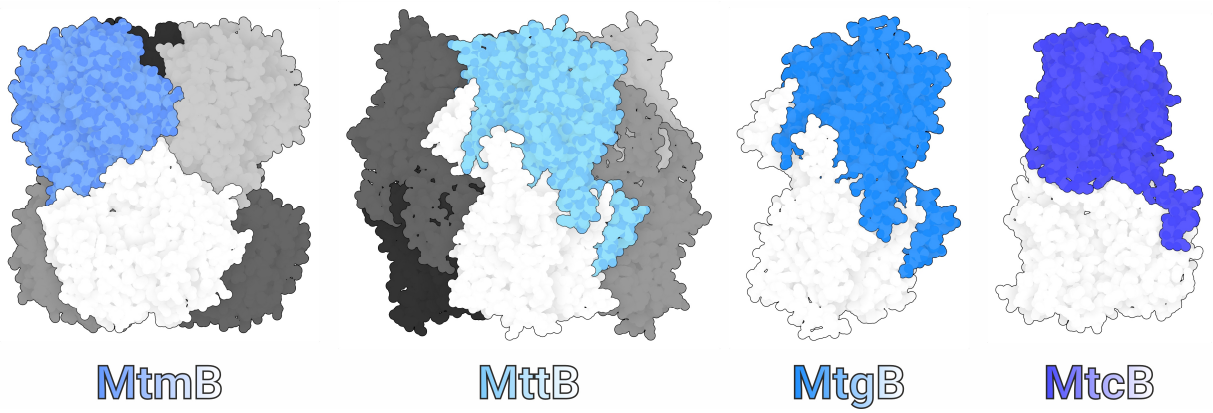


Figure 17 Sphere representations of the oligomerization of the methyltransferases MtmB (PDB-ID: 1NTH), MttB (PDB-ID: 7XCL), MtgB (this work), and MtcB (this work). MtmB and MttB form a homohexamer, whereas MtgB and MtcB assemble as homodimers. In each structure, subunit A is shown in a different shade of blue or purple. Remaining subunits are depicted in different shades of grey. MttB, MtgB and MtcB are homologs. MtmB and MttB are not homologous but contain the amino acid pyrrolysine in their active site.

Structures of MttB were solved in an apo form at 2.5 Å resolution (Fig. 18C) as well as in a state where a non-physiological sulfite was added to the pyrrolysine imine moiety (3.2 Å resolution). Additionally, a complex structure with the associated corrinoid protein MttC was solved at 2.7 Å resolution (Fig. 20). In contrast to MtmB, where two distinct orientations of the pyrrolysine residue with varying occupancy levels were observed, all solved structures of MttB display pyrrolysine in the same orientation (Fig. 18F).[84]

Due to their homology, MtgB, MtcB, and MttB can be overlayed with $RMSD_{C\alpha}$ values around 1.5 Å. The comparison between MtgB and MtcB has already been discussed in section 2.2.2. In this discussion, MtgB is primarily chosen for comparison with MttB due to the availability of a substrate-bound structure. Despite sharing 47 % sequence similarity, MtgB and MttB exhibit significant differences in their active sites. Only $Leu200_{MtgB}$ in the DMG-binding pocket matches $Leu217_{MttB}$. $Asn202_{MtgB}$ and $Thr275_{MtgB}$ have similar counterparts in MttB, $Cys219_{MttB}$ and $Ser292_{MttB}$, respectively (Fig. 18A and C). $Leu232_{MtgB}$ is replaced by the bulkier $Met249_{MttB}$, which occupies the space where DMG is bound in MtgB. This makes sense given the smaller size of the MttB substrate TMA. $Arg312_{MtgB}$ and $His348_{MtgB}$ are replaced by tyrosines ($Tyr329_{MttB}$ and $Tyr364_{MttB}$), since in MttB there is no need to stabilize a substrate

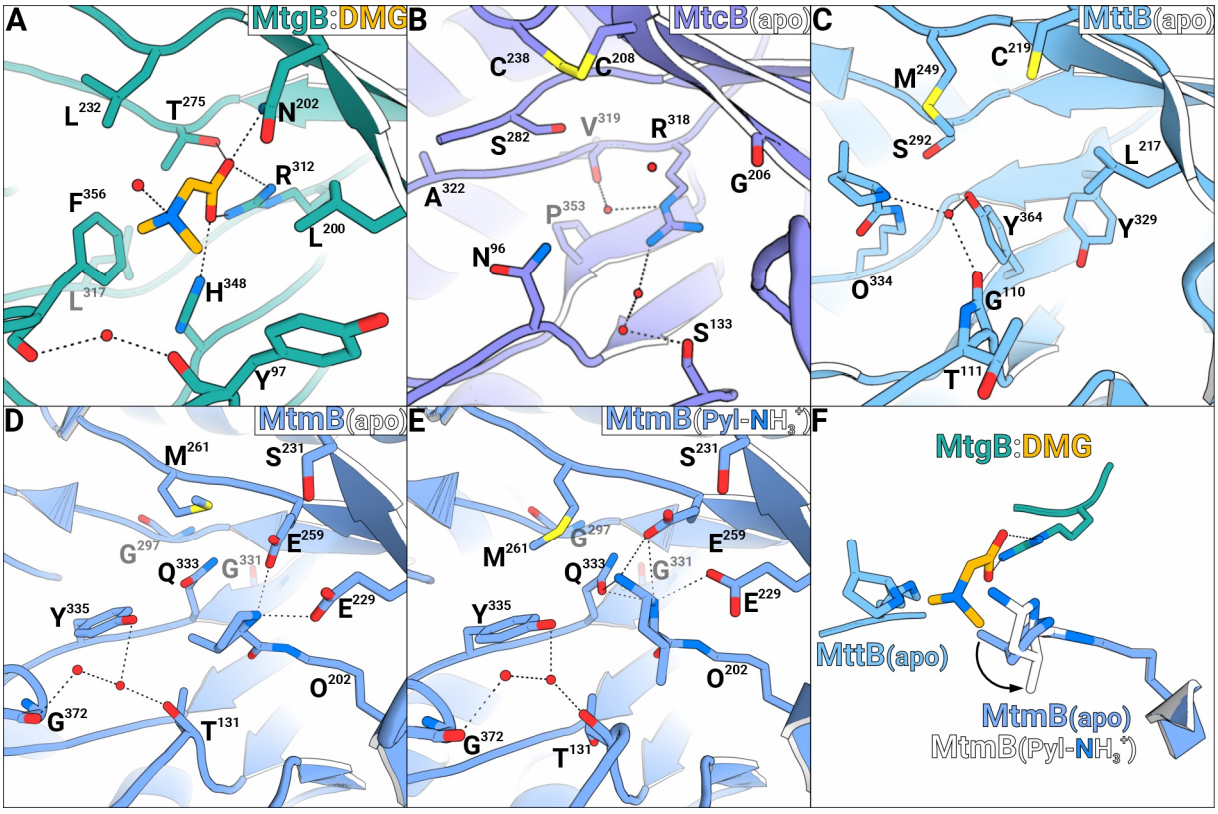


Figure 18 Active site comparison of MtgB, MtcB, MttB, and MtmB. Cartoon and stick representations of the active sites of MtgB:DMG (**A**), MtcB (**B**), MttB (**C**), MtmB (**D**), and amine-bound MtmB (**E**) are shown in similar orientation based on their superposition with MttB. Oxygens are depicted in red, nitrogens in blue. Carbon atoms are colored in the corresponding cartoon color. Waters are shown as red spheres. Hydrogen bonds are drawn as dashed lines. Residues are labeled by one-letter code. **F**) Superposition of key residues and substrates from A, C, D, and E. An arrow indicates the movement between the two pyrrolysine orientations in the MtmB structures.

carboxy moiety such as the one in GB. Two lipophilic amino acids occupy the space of $Pyl334_{MttB}$, $Leu317_{MtgB}$ and $Phe356_{MtgB}$ (Fig. 18A, C and S11). These differences in residue environment can be attributed to the need of MtgB to accommodate the larger substrate GB and to compensate for the absence of pyrrolysine.

Despite the lack of sequence homology, MtmB and MttB can be superimposed with an $RMSD_{C\alpha}$ of 4.2 Å due to their similar TIM barrel topology. Although both proteins contain pyrrolysine as a catalytic residue, their active sites exhibit notable differences. Among the observed residues, only $Met261_{MtmB}$ directly corresponds to $Met249_{MttB}$ (Fig. 18C, D, E, and F), but it remains unclear whether these residues fulfil a role in the reaction mechanisms of the enzymes. The pyrrolysine residue originates from different sides of the cavity in MtmB and MttB (Fig. 18C, D, E, and F). Taken together, these differences imply significant deviations in the reaction trajectories of the two proteins.

Based on the two observed conformations of the catalytic pyrrolysine within the active site of MtmB, a detailed reaction mechanism has been proposed (Fig. 19A).[54,55,64] MMA is added to the imine carbon moiety of $Pyl202_{MtmB}$ moving the electron pair of the double bond to the secondary amine of the pyrrole ring. Throughout the process, $Glu229_{MtmB}$ and $Glu259_{MtmB}$

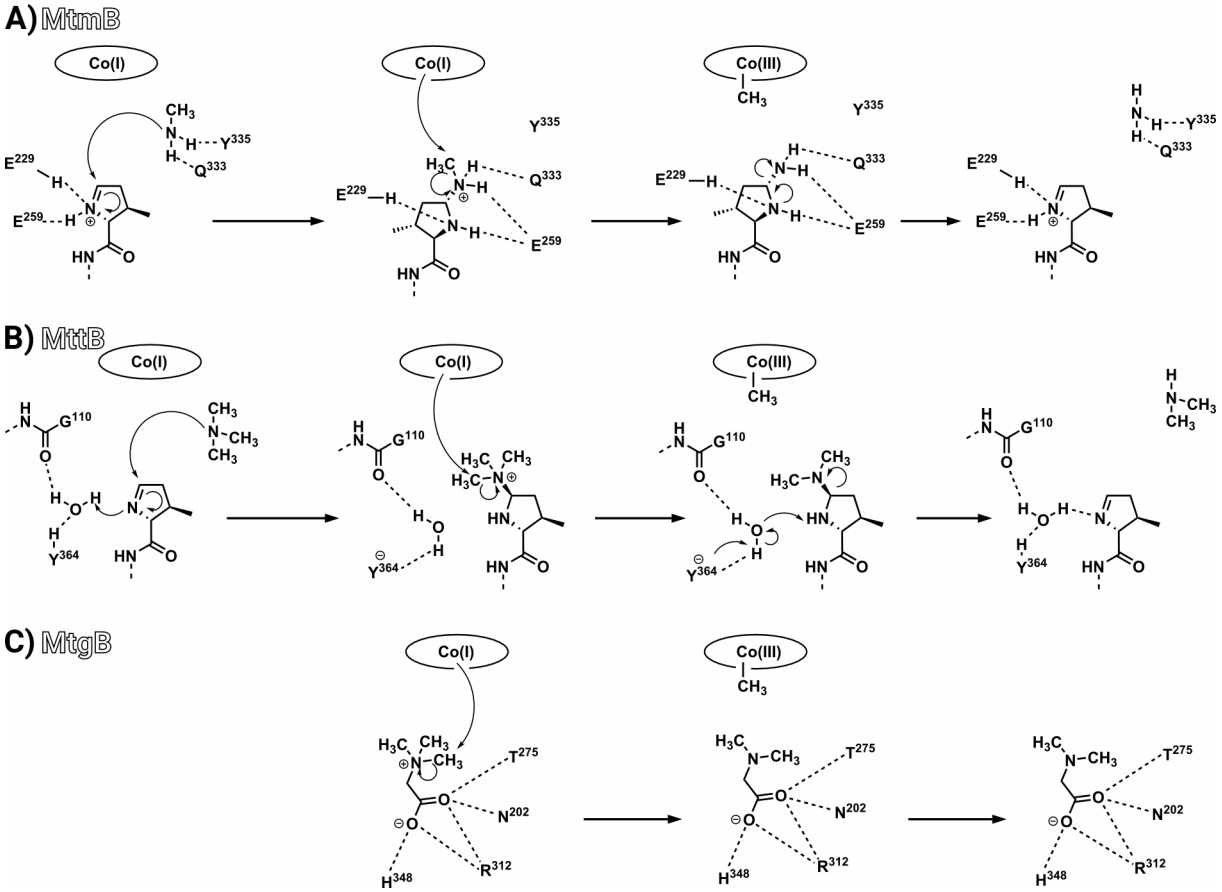


Figure 19 Proposed reaction mechanisms of MtmB, MttB, and MtgB. **A** and **B**) MtmB and MttB activate their respective substrates MMA and TMA with their specialized amino acid pyrrolysine to abstract a methyl group and transfer it to the corrinoid cofactor. **C**) MtgB uses the quaternary amine GB as substrate and skips the activation step.

help stabilize the secondary amine by acting as proton donors and acceptors as needed (Fig. 18D, E, and 19A). Upon addition of MMA, the pyrrole ring rotates 90° from orientation one to orientation two (arrow in Fig. 18F and 19A). In this way, the methyl group of the added MMA is brought closer to the corrin ring of the methyl group accepting corrinoid protein MtmC. $Gln333_{MtmB}$ and $Tyr335_{MtmB}$ stabilize the free and bound substrate (and later product) while it resides in the active pocket (Fig. 18E and 19A). After methyl transfer, the remaining amine is still bound to pyrrolysine, which is the state observed in orientation two (Fig. 18E and 19A). Ammonia is abstracted after protonation by an unknown base, allowing the pyrrole ring to revert to its original orientation one and start a new reaction cycle (Fig. 18D and 19A).[54,55,64]

Assuming that pyrrolysine behaves analogously in different enzymes, a similar mechanism seems likely for MttB (Fig. 19B). First, TMA is added to the pyrrolysine imine moiety, where one of the methyl groups is presented to the incoming Cbl (Fig. 19B). Upon methyl abstraction, DMA is released and pyrrolysine returns to its initial state (Fig. 19B). However, the available crystal structures of MttB reveal some mechanistic differences from MtmB. In MttB, no alternative conformations of pyrrolysine were observed even after the addition of a substrate analog to the pyrrolysine imine group, suggesting that pyrrolysine remains in one orientation throughout the reaction. It is possible that the trimethylated substrate intermediate of MttB has a methyl group facing the receiving Cbl by default and does not require repositioning (Fig. 18C, D, E, F, 19A and B). Stabilization of the pyrrole ring amine moiety, as occurs in MtmB via $Glu229_{MtmB}$ and $Glu259_{MtmB}$, is likely conferred in MttB via $Gly110_{MttB}$, $Thr111_{MttB}$, and $Tyr364_{MttB}$ (Fig. 18C, D, E, 19A and B). This is suggested by the strong conservation of these residues in (pyrrolysine-containing) TMA methyltransferases. In MtmB, $Gln333_{MtmB}$ and $Tyr335_{MtmB}$ interact with the amine moiety of the substrate through hydrogen bonds. MttB does not possess analogous counterparts for these residues due to the lack of hydrogen-bond donors in the trimethylated substrate TMA (Fig. 18E and 19A).

Regarding the reaction mechanism of MtgB, it has been postulated that GB-bound MtgB is reminiscent of TMA-bound pyrrolysine in MttB.[80] The placement of the quaternary amine moieties as suggested by the solved structures is consistent with this. Superposition of the sulfite-bound MttB structure with the GB-bound MtgB structure shows that the respective demethylation sites are only 2.5 Å apart. Both enzymes are able to prepare and position a quaternary amine moiety in a similar position to prime it for transfer to Cbl. Similar to MttB, the trimethylated amine moiety of GB cannot interact with hydrogen bond acceptors within MtgB. While MttB fixes TMA via its pyrrolysine residue, the substrate of MtgB is held in place by interactions with the carboxy group of GB with $Asn202_{MtgB}$, $Thr275_{MtgB}$, $Arg312_{MtgB}$, and $His248_{MtgB}$ (Fig. 18A and 19C). One of the methyl groups of the substrate is presented to the

incoming Cbl. After methyl abstraction, the remaining DMG can be exchanged for a new GB to restart the reaction cycle (Fig. 19C).

Methylated compounds such as methylamines and methanol are highly inert species. Heterolytic cleavage of the methyl group is accomplished by the combination of the supernucleophile Co(I) and a strong electrophile.[147,155] In practice, all Cbl-dependent methyltransferases achieve this by donating a positive charge to the heteroatom attached to the methyl group.[27] Depending on the substrate, this activation process is performed by a wide variety of different enzymatic mechanisms.[54,102,108] Access to non-activated methylated amines as an energy and carbon source functioned as a strong evolutionary driving force for *Methanosarcinales* to introduce the specialized amino acid pyrrolysine into their repertoire. This electrophilic residue, a property otherwise found only in protein prosthetic groups or post translational modifications, is capable of forming covalent bonds with methylamines and activating them by installing a positive charge at the N-atom.[156] The activated quaternary amine is then presented to Cbl for demethylation (Fig. 19C).[54,55,64,84] Quaternary amine methyltransferases, such as MtgB, operate without the need for additional substrate activation.[157-159] Their substrates already resemble the positively charged intermediate state encountered during pyrrolysine-mediated TMA activation in MttB (Fig. 19B and C). Their purpose is primarily to position the activated methyl moieties for transfer to Cbl (Fig. 19C).

Although the protein structures of MtmB, MttB, and MtgB are related either by the characteristic amino acid pyrrolysine or by sequence homology, their active site environments differ significantly. In MttB and MtmB, pyrrolysine originates from different sides of the active pocket (Fig. 18C, D, E, and F). MtgB is a homolog of MttB; however, this is not reflected in their active site compositions. Nevertheless, the basic features of a uniform reaction mechanism can be recognized in all three enzymes (Fig. 19A, B, and C). However, the execution becomes increasingly simpler. MtmB activates its substrate MMA and rotates the pyrrole ring to enable the transfer of the methyl group to Cbl (Fig. 19A). MttB functions similarly but omits the rotation of the pyrrole ring (Fig. 19B). MtgB forgoes the activation step and directly positions the substrate GB for demethylation (Fig. 19C). The positioning step is where most of the similarities between the enzymes can be seen. The methyl groups of the respective substrates are fixed in a position accessible to the cofactor of the associated corrinoid protein (Fig. 19A, B, and C). In the structural superposition, all demethylation sites are located within 4 Å of each other, highlighting this conserved feature of the reaction mechanism (Fig. 18F).

The molecular compositions of MtgB and MtcB help in better understanding the peculiarity that the only three pyrrolysine-utilizing enzymes in nature, MtmB, MtbB, and MttB, do not share any sequence similarity. In the case of MttB, favorable traits from both the

quaternary amine methyltransferases and the pyrrolysine methyltransferases must have been combined via horizontal gene transfer to create an enzyme that enables the degradation of the otherwise metabolically inaccessible compound TMA. The sequence scaffold of quaternary amine methyltransferases allows positioning of the quaternary amine moiety for efficient methyl transfer. The acquired amber-encoded pyrrolysine, in turn, allows activation of TMA to a quaternary amine. MtbB is the least studied member of the three pyrrolysine enzymes. Although no non-pyrrolysine relatives of MtbB are known, it is likely that MtbB evolved similarly to MttB by combining the pyrrolysine residue from MtmB with a pre-existing protein scaffold. Structural studies of MtbB and other non-pyrrolysine relatives of pyrrolysine methyltransferases will provide further insight into the driving forces behind the entry of pyrrolysine into the genetic code of methanogenic archaea.

2.2.4 Structural Comparison of Cobalamin-bound Structures of MtgB with Related Methyltransferases

As discussed in section 2.2.1, several structures of MtgB in complex with Cbl or MeCbl have been solved. Specifically, MtgB:Cbl (2.2 Å), MtgB:MeCbl (2.6 Å), MtgB:GB:Cbl (2.3 Å), MtgB:DMG:Cbl (2.1 Å), and MtgB:DMG:MeCbl (2.35 Å) featured Cbl and MeCbl within their active sites (Tables S1 and S2, Fig. 12, 14 and 15). These complex structures allow a detailed comparison with the closely related MttB:MttC complex (PDB-ID: 7XCN) as well as other related methyltransferases regarding the positioning, binding and putative mechanism revolving around the corrinoid cofactor.[84]

The MttB:MttC structure displays a large complex in which the hexameric MttB core is expanded to a MttB-MttC hetero-dodecamer (Fig. 20D). In the MttB:MttC complex, the corrinoid cofactor (5-hydroxybenzimidazole-cobamide, Cbl_{5OH}; compared to Cbl, two methyl groups of the benzimidazole moiety are replaced by one hydroxy group) is the central element responsible for most of the interactions between the two different polypeptide chains (Fig. 20F). The pendant benzimidazole moiety is complexed by the Rossman fold of MttC (Fig. 20F). Both the lower and upper faces of the corrin ring form an extensive hydrogen bond network with MttC as well as MttB, respectively (Fig. 20F). Direct contacts between the enzymes are rare, suggesting that the proteins would dissociate without the presence of Cbl_{5OH} (Fig. 20F).

Due to their homology, one would expect MtgB and MttB to bind very similarly to the corrin ring interface of the corrinoid cofactor. However, analysis of both proteins from a similar orientation shows that in MtgB the corrin ring is rotated by 180° and slightly tilted to the side (Fig. 20B, C, E and F). Therefore, a direct comparison of the residues involved in cofactor binding in MtgB and MttB is not possible. However, some general principles regarding cofactor

binding highlighted in the description of MtgB-Cbl interactions described in section 2.2.1 can however also be recognized in the MttB:MttC structure. First, the interactions with the corrin ring are predominantly exerted by carbonyl groups of the MttB backbone (Fig. 20F). In addition, the region around the start of the pendant benzimidazole arm is characterized by a multitude of hydrogen bonds bridged by water molecules (Fig. 20C and F).

Another common feature of the two protein structures is the significant distance between the demethylation and methylation sites. The imine group of Pyl334MttB and the cobalt ion of Cbl_{50H} are 7 Å apart. Similar to MtgB (Section 2.2.1), this distance is too large to allow methyl transfer. Similar characteristics can be observed in the structures of other corrinoid methyltransferases. CoFeSP structures show distances of about 13 Å between the reaction sites of the cofactors Cbl and THF.[102] In the MtaB:MtaC structure (PDB-ID: 2I2X), there is a distance of 7.7 Å between the cobalt ion of Cbl_{50H} and the zinc ion within the MtaB active site.[108] The hypothetical addition of the methyl source methanol to the active site of MtaB would imply a distance of 4.5 Å between the methyl group and the cobalt ion, which

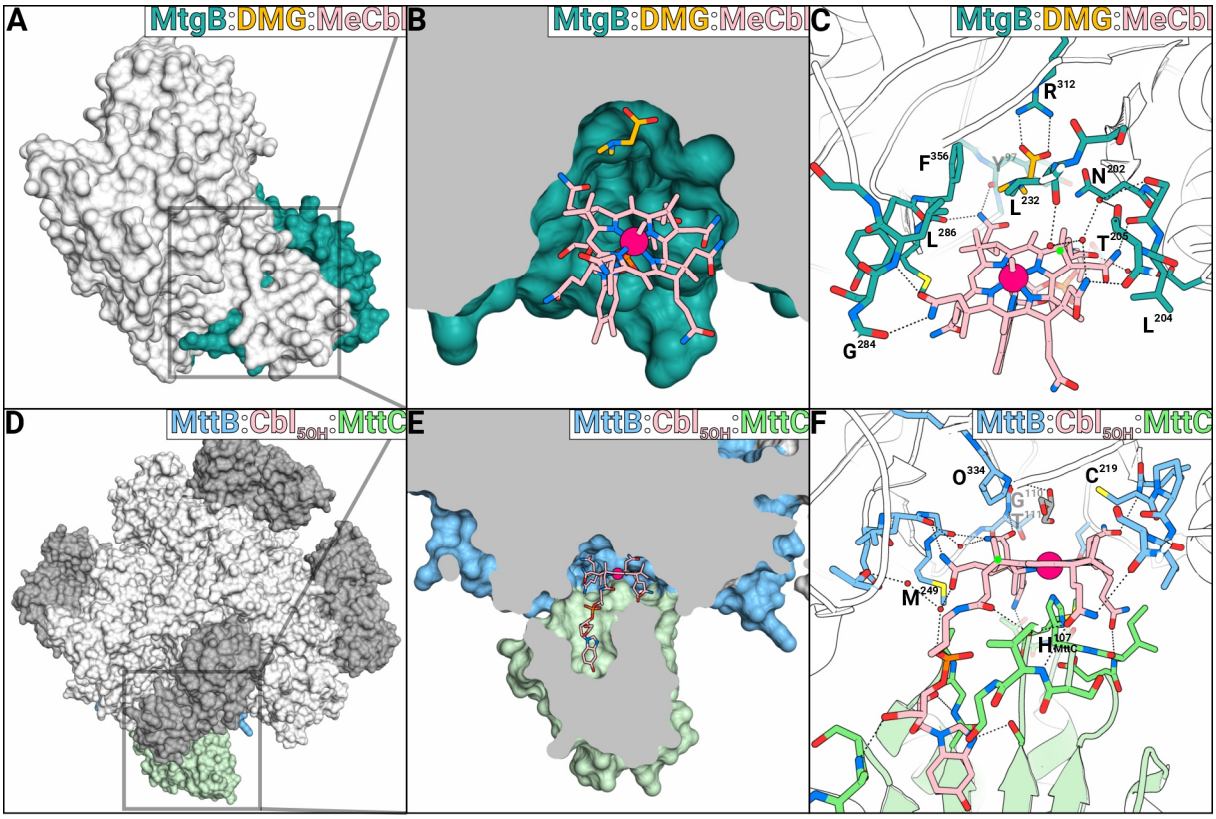


Figure 20 Surface, cross-section, and cartoon/stick representations of MtgB (top) and the MttB:MttC complex (PDB-ID: 7XCN) (bottom). The two structures were aligned based on the homologous TIM barrel topology of MtgB and MttB. Each depiction is shown from a similar angle. Subunit A of the MtgB homodimer is shown in teal, subunit B in white. Subunit A of the MttB homohexamer is shown in light blue, while the interacting MttC subunit is shown in green. The remaining MttB and MttC chains of the 6MttB:6MttC complex are shown in white and grey, respectively. **A** and **D**) Full surface views of both structures. **B** and **E**) Zoomed-in cross-sections of A and D. **C** and **F**) Zoomed-in cartoon- and stick representations of B and E. Cartoons of the A subunits of MtgB and MttB are shown in white for clarity, and only the sticks are displayed in teal and light blue, respectively. Hydrogen bonds are indicated as dashed lines. To highlight the rotations of the cofactors, the origins of the pendant benzimidazole moiety are marked with a green dot.

again was considered too long to allow methyl transfer.[108]. Interestingly, the simulation of the molecular movement of the active site of the MttB:MttC structure by molecular modeling showed that substrate-bound pyrrolysine can manage to move its methyl group as close as 4.6 Å toward the cobalt ion.[84] This distance is still not sufficient for methyl transfer catalysis and suggests that additional structural rearrangements of both proteins are required to push Cbl_{5OH} further into the active site.[84] Such a rearrangement was proposed for MtaB:MtaC, postulating the existence of an open and a closed state of the protein complex. In the observed open state, MtaC was found to be too far away from the associated MtaB subunit for efficient methyl transfer. Furthermore, MtaC was characterized by high temperature factors and slightly different conformations within the $(MtaB:MtaC)_2$ tetramer, indicating high flexibility for this construction. It was suggested that the transition to a closed state, induced by a 1.0 Å approach, would bring the methyl donor and acceptor sites into catalytic distance, rigidify the structure, and displace bulk solvent from the active site. Reduction of Co(III) to Co(I) has been proposed as a possible mechanism for this conformational change. Unlike hexa-coordinated (methyl-)cob(III)alamin, cob(I)alamin is tetra-coordinated and thus lacks the axial ligand histidine of MtaC, which pushes the ring into the active site.[160,161] Conversely, methylated Co(III) would facilitate the switch back to the open conformation again.[108]

Similarly, MtgB exhibits high temperature factors for bound Cbl or MeCbl, suggesting a high degree of flexibility. However, in the MtgB structure, MtgC is not present to exert a closing mechanism to push Cbl into the MtgB active site. Still, the reaction takes place within the formed crystals (Section 2.2.1), suggesting that without the constraints of a corrinoid protein complexing the pendant benzimidazole arm, Cbl can diffuse into the active site to accept the methyl group from GB. The bulk water in the Cbl-bound active site of MtgB indicates that there is still room for Cbl to move further into the pocket toward the demethylation site (Fig. 14). Based on the Cbl-bound MtgB structures, molecular modeling experiments could provide great insights into how this process is carried out.

Structures of Cbl-dependent methyltransferases show remarkable similarities across all domains of life.[30,104] The corrinoid carrier component, which complexes the pendant benzimidazole arm of the cofactor, even shows a high degree of sequence similarity throughout all Cbl-dependent methyltransferase systems.[45] A wealth of structural data on the individual components and domains of these methyltransferase systems is available, allowing a thorough analysis of the associated enzymatic mechanisms. In this work, several 3D structures of individual components of Cbl-dependent methyltransferases have been solved, allowing proposals for detailed mechanistic analyses. A much more challenging task is to analyze the complex formation and cofactor interactions in which these components

engage are engaged. Cbl-dependent methyltransferases are highly complex enzymatic systems that combine a large number of reaction steps and conformational rearrangements.[30,104] While these domain shifts have been biochemically characterized quite extensively for common proteins such as MetH and CoFeSP, structural data on specific domain interactions are rare. Structures of MetH often depict the interaction between the Cbl-carrier domain and the SAM-carrying reactivation domain.[97,141,162-164] However, no MetH structures that show the specific interactions of Cbl with either THF or homocysteine are available. In this regard, the structures of MtgB, MttB:MttC, and MtaB:MtaC discussed above give more insights into how the interaction between the substrate and Cbl might be achieved. This provides the main motivation to further investigate structures of microbial Cbl-dependent methyltransferases such as MtgB, MtgC, and MtgA from *D. hafniense*. Each new complex structure of Cbl-dependent enzymes might feature the cofactors and substrates in a different conformation and exhibit different interactions between the subunits. Due to the similarities of the underlying peptides, each new structure has the potential to provide insight into the conformational maneuvering of Cbl-dependent methyltransferases as a whole.

2.3 MtbA and MtgA - Structures of the Acceptor Proteins

In methyltransferases, Cbl acts as a vehicle to transfer a methyl group from a donor molecule to an acceptor molecule. In the methanogenic pathway of methylamine degradation, MtbA is the methyl group accepting protein that catalyzes the methylation of CoM (Fig. 1 and 2). GB degradation in *D. hafniense* leads to the methylation of THF by the structurally related enzyme MtgA.[80] This section illustrates the 3D structures of both proteins in complex with their respective substrates and products. Both proteins share the TIM barrel topology typical of Cbl-associated methyltransferases.[54,98,108,165] Despite their similar shape and size, they each manage to embed a cofactor of very different size and chemical properties. Here, both proteins are characterized to illustrate how methyl transfer from the preceding corrinoid protein is achieved.

2.3.1 The Active Site of MtbA Differs from Its Isozyme MtaA

MtbA was cloned and heterologously expressed in *E. coli* (Tables 10, 11, and 12, Fig. S8). The enzyme was purified to homogeneity via IMAC and SEC (Table 12) and co-crystallized with CoM or MeCoM (Table S1). The crystal structures of MtbA:CoM and MtbA:MeCoM were solved at resolutions of 1.25 Å and 2.2 Å, respectively (Table S5). Crystal packing analysis using the PISA webserver indicated that MtbA occurs as a monomer.[128]

The protein adopts a TIM barrel fold with the active site located in the center of the barrel structure (Fig. 21A). In MtbA:CoM, CoM and Zn^{2+} are located horizontally in the opening of this

barrel structure, with the thiol group of CoM facing Zn^{2+}. Several residues help coordinate the sulfonate group and the carbon backbone of CoM. Arg70 forms hydrogen bonds with two oxygens of the deprotonated sulfonate group of CoM. Bending of the guanidine toward the sulfonate is assisted by hydrogen bonds with the hydroxyl group of Ser144 and the carbonyl oxygen of Pro72, which is positioned by a cis-peptide bond that occurs between Val71 and Pro72. Furthermore, the distances between the carbon in the δ-position of Ile75 and the primary methyl group in the γ-position of Ile200 indicate van der Waals interactions with the two carbon atoms of CoM (Fig. 22A and C). The water molecule between Arg70 and the sulfonate group of CoM is tetrahedrally coordinated by four partners (Fig. 22A and C). In addition to forming hydrogen bonds with the guanidine of Arg70, the sulfonate group, and the carbonyl backbone of Asp74, a proton relay system involving a series of water molecules is initiated by Glu78. This system may allow protonation of the sulfonate group installing re-deprotonation as a driving force for ejection of methylated CoM.

The Zn^{2+} complexed in the active site of MtbA exhibits tetrahedral coordination. In the MtbA:CoM structure, it is coordinated by His239, Cys241, Cys316, and the thiol group of CoM. This configuration may assist in stabilizing the thiol group of CoM in its deprotonated form to increase the nucleophilicity for the addition of the incoming methyl group. Interestingly, the coordination of Zn^{2+} is inverted in the MtbA:MeCoM structure (Fig. 22B). The superposition of the two structures shows that after the methyl transfer, the zinc ion has moved 1.1 Å away from the substrate (Fig. 22C). While MeCoM no longer binds to Zn^{2+}, the fourth zinc coordination is now provided by a water molecule complexed by Asp262 (Fig. 22B and C). With this, MtbA conforms to the zinc binding motif H-X-C-X_n-E/D-X_m-C, which is prevalent in many zinc binding methyltransferases.[107,110,148,166-169] Although the reaction product MeCoM is still complexed within the active site of MtbA, its thiomethyl group is now directed away from Zn^{2+}

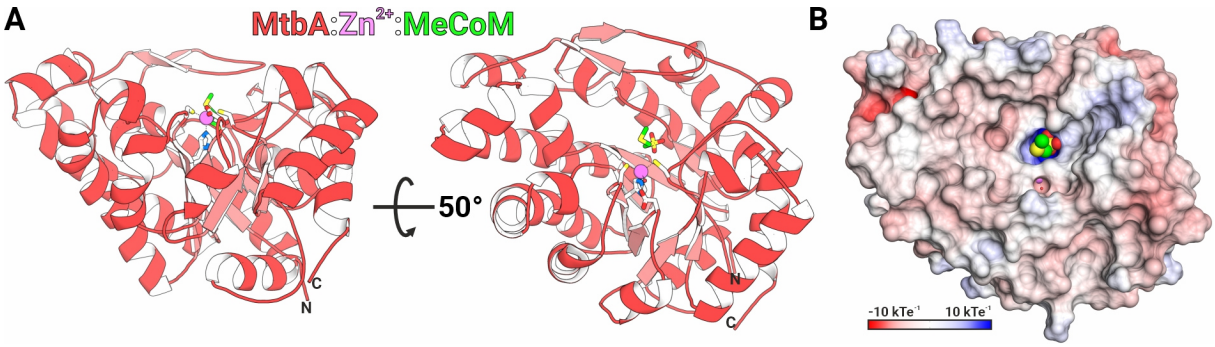


Figure 21 Cartoon and surface charge depictions of MtbA in complex with MeCoM. **A**) Top and side views of MtbA to highlight its TIM barrel fold. The methyltransferase is depicted in complex with its product, MeCoM, in stick representation. The carbon atoms of the cofactor and the methyl group are colored in green. The zinc ion is colored in violet. The zinc-interacting residues Asp262, His239, Cys241, and Cys316 are shown as white sticks. **B**) Surface charge representation of MtbA revealing the solvent exposure of the active site. The colors highlight negative and positive electrostatic potentials, which are contoured from -10 kT/e (red) to +10 kT/e (blue). MeCoM and zinc are shown as spheres and colored similarly to A.

toward the solvent, displaying a possible conformational state immediately after methyl transfer from MeCbl.

MtbA belongs to a group of isozymes within the methylotrophic pathway system of methanogenic archaea. The other two isozymes, MtaA and MtsA, analogously bind zinc as their prosthetic group, allowing them to activate the methyl group for transfer to the thiol group of CoM.[107,146,170] However, these enzymes obtain the methyl group from different precursors (Fig. 1). MtsA forms a 480 kDa oligomer with MtsB, which simultaneously demethylates various methylated thiol compounds and transfers the methyl group to the $Cbl_{(5OH)}$ of MtsB.[107] In the final step, MtsA again catalyzes the demethylation of $Cbl_{(5OH)}$ and methylates CoM.[107] It shares 47.7 % sequence similarity with MtbA. Still, the oligomeric structure of the MtsA:MtsB complex differs significantly from that of monomeric MtbA. A crystal structure of MtsA is not available, but an AlphaFold prediction of the monomeric MtsA subunit showed high agreement with the overall structure as well as the catalytic residues of the solved MtbA structures. The only distinguishing feature is a 30 residue long helix at the C-terminus of MtaA. This structural element is missing in MtbA and may be responsible for complex formation in the MtsA:MtsB oligomer.[124] MtaA, together with its associated methyltransferases MtaB and MtaC, mediates methyl transfer from methanol to CoM.[171] Similar to MtbA, MtaA occurs as a monomer.[172,173] The proteins are so similar that cross-activity has been reported where MtaA can receive the methyl group from both the methanol and TMA pathway.[174] The 3D structure of MtaA from *Methanosarcina mazei* (*M. mazei*) has been solved with Zn^{2+} (1.8 Å, PDB-ID: 4AY7, here referred to as “apo”) and in complex with Zn^{2+} and CoM (2.1 Å, PDB-ID: 4AY8, here referred to as “MtaA:CoM”).[173] MtbA from *M. barkeri* and MtaA from *M. mazei* share 54.3 % sequence

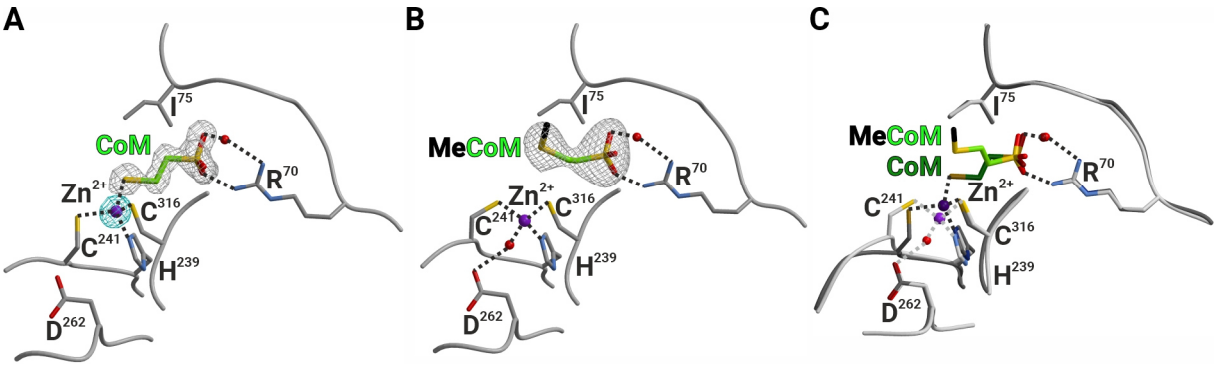


Figure 22 Active site close-up views of the MtbA structures **A**) MtbA in complex with its substrate CoM (1.25 Å resolution) showing the tetrahedral coordination of Zn^{2+} by the $HXCX_nC$ motif. The electron density (grey mesh) of the ligand is represented by a $2F_o$-F_c-map (contoured at 1.0 σ). The anomalous electron density (cyan) of the heavy metal atom is contoured at 15 σ. Residues engaged in cofactor binding are depicted as sticks and labeled by one-letter-code. Hydrogen and Zn^{2+}-coordinating bonds are shown as dashed black lines. Water molecules are shown as red spheres. **B**) MtbA in complex with its product MeCoM (2.2 Å resolution). **C**) Structural superposition of MtbA:CoM (dark) with MtbA:MeCoM (bright). The nucleophilicity of the thiol group of CoM is generated by its coordination with the Lewis acid Zn^{2+}. During methyl transfer, the S-Zn bond is broken and the metal atom is displaced by 1.1 Å. Although the tetrahedral configuration is retained, the geometry is inverted. A water molecule is captured during the conversion of the zinc atom and is H-bonded to Asp262. It is well-defined but only seen in the MeCoM complex structure.

similarity, which is reflected in their nearly identical structural topology (Fig. 23A and B). The available apo and CoM-bound structures of MtaA are complemented by the CoM- and MeCoM-bound structures of MtbA presented here. Although both enzymes are similar, interesting nuances can be found in the respective active sites. Furthermore, the MeCoM-bound MtbA structure provides insights into a late reaction stage of the enzyme, which was not possible with the available MtaA structures.

The CoM-bound structures of both enzymes display virtually identical coordination of the active site components (Fig. 23A). As described for MtbA, the zinc ion is tetrahedrally coordinated by $His240_{MtaA}$, $Cys242_{MtaA}$, and $Cys319_{MtaA}$. The fourth zinc coordination site is occupied by the thiol group of CoM. The sulfonate moiety of CoM is stabilized by $Arg73_{MtaA}$ and $Gln97_{MtaA}$. Interestingly, the zinc ion in the apo structure of MtaA shows a similar behavior as in the MeCoM-bound structure of MtbA. In MtaA apo structure, the Zn^{2+} shifts by 1.0 Å, inverts its coordination, and interacts with $Glu263_{MtaA}$ on the opposite side (Fig. 23B). This glutamate residue also represents the main difference between the two active sites. In MtbA:MeCoM, the fourth coordination site of Zn^{2+} is occupied by a water molecule, which forms a hydrogen bond with $Asp262_{MtbA}$ (Fig. 22B, C, and 23B). Strikingly, across methanogenic organisms, MtaA features a glutamate in this position, while MtbA and MtsA both possess aspartates.[173] It seems that this change does not affect the reaction mechanisms of the enzymes, since the measured distances and residue positions in the two available MtbA structures match their counterparts in the MtaA structures. However, the characteristic change in this position may be important in regulating the reaction through pH

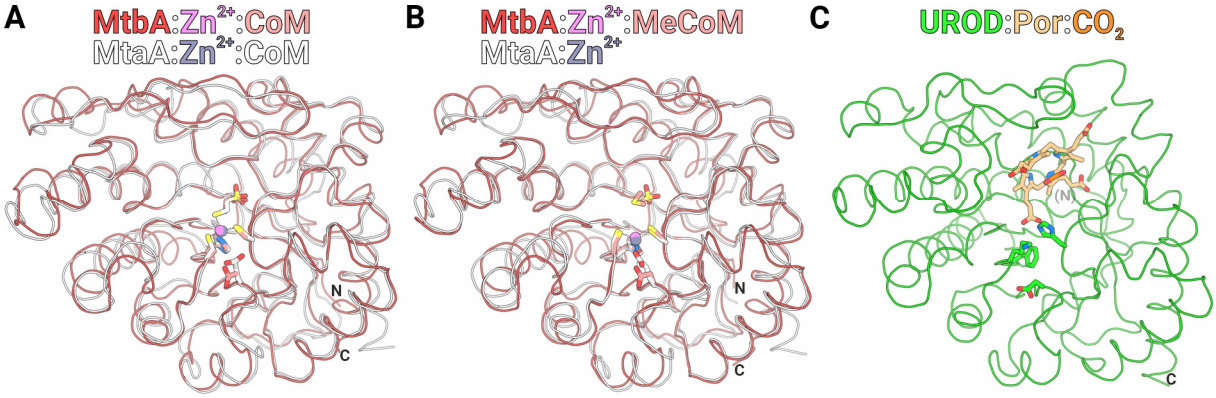


Figure 23 Ribbon superpositions and comparisons of MtbA with homolog proteins. **A**) Superposition of MtbA:Zn^{2+}:CoM (this work) and MtaA: Zn^{2+}:CoM (PDB-ID: 4AY8). Residues involved in the complexation of zinc and CoM are shown as sticks. MtbA and MtaA share a high degree of sequence similarity, which is reflected in the overlay of the structures. The N- and C-termini of MtbA are labeled. **B**) Superposition of MtbA:Zn^{2+}:MeCoM (this work) and MtaA: Zn^{2+} (PDB-ID: 4AY7). Note that both structures feature a similar inversion of the zinc binding mode. **C**) Ribbon depiction of uroporphyrinogen decarboxylase (UROD) in complex with coproporphyrinogen I (Por) and CO_2. The overall sequence and fold of this protein is similar to MtbA and MtaA. However, the enzyme catalyzes a completely different reaction. The corresponding residues of the structures shown in A and B are shown as sticks. In UROD, these are not necessarily involved in catalysis, but serve as an aid for better orientation in this figure. The similarity between the compounds Por and Cbl helps to better understand how Cbl enters the active sites of MtbA and MtaA.

shifts or domain interactions with the preceding methyltransferase components. If the aspartate-bound water in MtbA (or MtsA, in which the active site looks identical to MtbA according to the AlphaFold prediction) is kept out of the active site, inversion of the zinc ion coordination and thus CoM methylation could be prevented. This mechanism for inhibiting the reaction would not be possible in MtaA, where direct interaction with $Glu263_{MtaA}$ is always possible. While MtaA is found in methanogens grown on methanol, MtbA and MtsA are preferentially produced in methanogens grown on acetate.[175,176] This provides a rationale for a possible dependence of the enzymatic activities of MtbA and MtsA on low pH, while MtaA may be more pH-independent. MtbA interacts with MtmC, MtbC, as well as MttC in the methylamine:CoM methyl transfer pathways.[36,172,174] This promiscuity may also require additional potential to regulate the methyl transfer reaction.

The DALI webserver identifies MtaA from *M. mazei* as the closest available structural relative of MtbA (PDB-ID: 4AY8, Z-score: 50.2, RMSD: 1.4 Å, sequence identity: 36 %),[129] whereas the second hit is the functionally different enzyme uroporphyrinogen decarboxylase (UROD) (PDB-ID: 1R3Q, Z-score: 36.8, RMSD: 2.3 Å, sequence identity: 20 %, sequence similarity: 31.8 %). UROD does not contain zinc in its active site and catalyzes a different reaction.[177] Despite the sequence similarities to MtbA and MtaA, UROD accommodates a large tetrapyrrole substrate in its active pocket.[178] Like $Cbl_{(5OH)}$, which delivers the methyl group to MtbA and MtaA, uroporphyrinogen is a porphyrin derivative. Therefore, it was hypothesized that the binding site of uroporphyrinogen in UROD can serve as a template for $Cbl_{(5OH)}$ binding in MtaA.[173,179] While the configuration of uroporphyrinogen within the TIM barrel opening of UROD is not directly transferable to MtbA or MtaA due to steric clashes, structural data of this sort provide clues as to how a large cofactor like Cbl can enter a TIM barrel shaped active site to perform a methyl transfer reaction (Fig. 23C). Furthermore, combining the information of such structures can serve as valuable input for modeling complex rearrangements of the cofactor Cbl within the active site of methyltransferases.

2.3.2 MtgA Displays a Novel THF Locking Mode During Methylation

WT and mutant versions of MtgA were cloned and heterologously produced in *E. coli* (Tables 10, 11, and 12, Fig. S9 and S10), purified via IMAC and SEC (Table 12, Fig. S9 and S10), and co-crystallized with THF or MeTHF (Table S1). The enzymatic activity of WT and mutant MtgA was analyzed to evaluate the role of specific residues in the reaction mechanism of the protein (Fig. 25). Apo, THF-, and MeTHF-bound structures of MtgA were solved at resolutions of 1.85 Å (PDB-ID: 6SJK), 1.35 Å (PDB-ID: 6SJ8), and 1.55 Å (PDB-ID: 6SK4), respectively. Furthermore, the structures of Apo (1.90 Å), THF- (1.75 Å), and MeTHF-bound (1.80 Å, PDB-ID: 6SJS) MtgA N227A mutant protein were determined. Finally, the structure of MtgA D102A in

complex with MeTHF was determined at 1.95 Å (PDB-ID: 6SJO) (Tables S6 and S7). PISA analysis indicated that MtgA assembles as a homodimer with a contact area of 1870 Å^2.[128] Each subunit adopts a TIM-barrel fold. The active site is located within the negatively charged cavity of each TIM barrel (Fig. 24A). A peculiarity of MtgA is that the characteristic $(\beta/\alpha)_8$ architecture of the TIM barrel fold is broken from residues 225 to 238. Starting at Pro225, the β-sheet on the inside of the barrel structure is interrupted and replaced by two small helices, which, as discussed below, may be important for catalysis (Fig. 24A, green region).

Analysis of the electron density of the MtgA:THF structure revealed that only the pterin portion of THF has defined electron density. The lack of density around the glutamyl-p-aminobenzoate moiety of the cofactor (THF tail) indicates that this region is flexible and not tightly complexed by the protein (Fig. 26A). The aromatic pterin ring system of THF on the other hand is stabilized via hydrogen bonds with Ser32, Asp75, Asp102, Asn129, and Asp194. Consistent with other THF methylating enzymes, the negative charges of Asp75 and Asp102 appear to be responsible for attracting the hydrogen atom at position N^8 of the pterin ring, thereby increasing the negative charge within the ring system and priming N^5 for the addition of the incoming Cbl-bound methyl group (Fig. 26A).[98,100,102,180] The enzymatic activity of

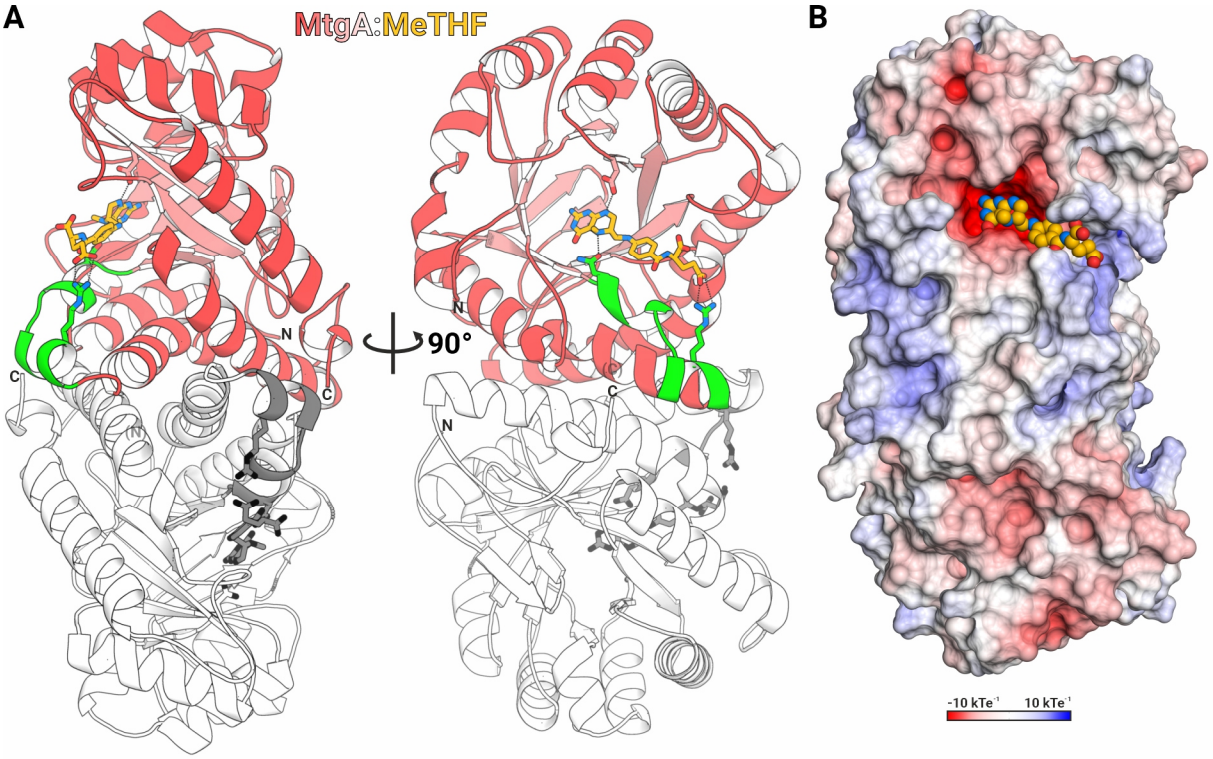


Figure 24 Crystal structure of MtgA. **A**) Cartoon representation of MtgA:MeTHF in two different orientations. The homodimer adopts a TIM barrel fold. Subunit A is colored in red, subunit B in white and shades of grey. MeTHF (yellow) and the catalytically important residues Asp102, Asn227, and Arg236 are depicted as sticks. The TIM barrel of MtgA is interrupted between residues 225 and 238, where the expected β-sheet is replaced by two small helices (green region). **B**) Surface charge distribution of MtgA. Cosubstrate coordination takes place in a solvent exposed binding cavity.

$MtgA_{D102A}$ mutant protein is greatly reduced compared to WT protein, demonstrating the importance of this electron-donating residue for the reaction (Fig. 25C).

In contrast to the substrate-bound structure, MtgA:MeTHF revealed fully defined electron density for MeTHF (Fig. 26B). The methylated cofactor interacts entirely with the active site environment. While the overall binding motif of the pterin ring is similar to that of MtgA:THF, it is tilted by approximately 35° into the active site cavity in the MeTHF-bound structure (Fig. 26C). The MeTHF benzoate moiety forms a π-stacking network with Tyr230, Phe243, Phe251, and Trp272 on one side, and Tyr35 and His38 on the other (Fig. 26B). The functional groups of the THF tail engage in water-coupled hydrogen bonding interactions with functional groups of the protein backbone, either directly or bridged by water molecules. The most pronounced interaction of the THF tail occurs between its terminal γ-carboxylate, which forms salt-bridges with Arg236. Because of the important role of this residue in the coordination of MeTHF in its extended conformation, the enzymatic activity of an $MtgA_{R236A}$ mutant was examined. $MtgA_{R236A}$ displayed a moderate reduction in enzymatic activity (Fig. 25C). Although the fixation of the THF tail is organized by more residues than just Arg236, this influence on the enzymatic activity suggests that fixation of the THF tail is a relevant factor during the methylation stage of the catalytic cycle of the enzyme.

The overlay of the substrate- and product-bound structures indicates that the benzoate moiety of MeTHF displaces the side chain of Asn227 by 2 Å during fixation of the THF tail (Fig. 26C, grey arrow). The resulting helix distortion is compensated by the formation of additional H-bonds between Asn227, Ser226, and N^5 of the pterin ring (Fig. 26C, yellow dotted line). Coincidentally, this region of the enzyme occurs in the two-helix motif that deviates from the canonical TIM barrel architecture (Fig. 24A, green cartoon area). A β-sheet in this position of the enzyme would probably not allow for a similarly wide-ranged swinging motion of Asn227, suggesting that the small helix may be a necessary evolutionary adaptation to allow for efficient catalysis in MtgA.

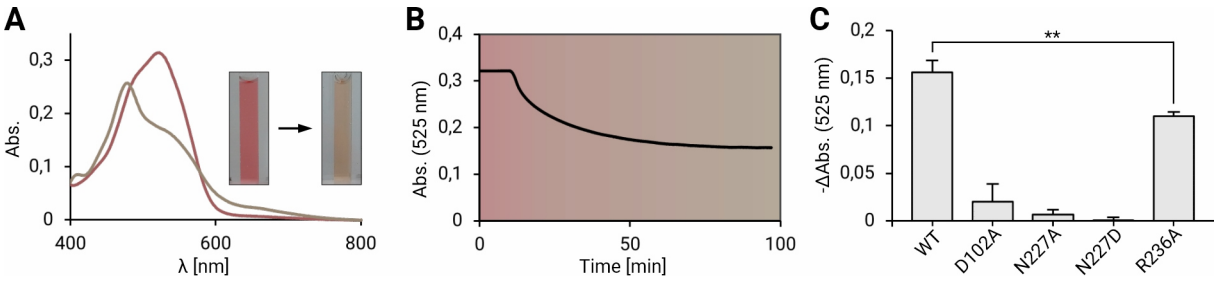


Figure 25 Enzymatic activity assays of WT and mutant MtgA. **A**) Spectra in the wavelength range between 400 nm and 800 nm before and after the methyl transfer reaction. Photos show the cuvette before and after the experiment. **B**) The conversion from $Co^{(III)}$ (MeCbl) to $Co^{(I)}$ (Cbl) accompanying the methyl transfer reaction can be monitored at 525 nm. **C**) Comparison of catalysis (mean ± SD; n = 3) between WT and mutant MtgA. The reduced activity of $MtgA_{R236A}$ was verified by an independent-samples t-test; $t(4) = 5.217$, $p = 0.0064$. Figure adapted from Badmann and Groll, 2020 (s. section 7).

In the MtgA:MeTHF complex structure, the methylated amine N^5 exhibits tetrahedral sp^3-hybridization rather than the planar sp^2-hybridization that would be expected for the final product MeTHF. This suggests that the structure depicts a trapped protonated reaction intermediate of MeTHF (Fig. 26B). The swinging motion that Asn227 undergoes during the reaction positions it for H-bond formation with the protonated N^5 (Fig. 26C, yellow dotted line). As previously observed for other THF-methyltransferases,[100] mutation of Asn227 to alanine or aspartate severely impaired enzymatic activity, suggesting that proper stabilization of the tetrahedral N^5-configuration is essential for catalyzing the methyl transfer from MeCbl to THF (Fig. 25C).

Structures of the MtgA mutants $MtgA_{D102A}$ and $MtgA_{N227A}$ in complex with MeTHF provided further insight into the reaction. In the $MtgA_{D102A}$:MeTHF complex structure, the electron density around the N^5 methyl group is undefined, suggesting a fluctuation between sp^3- and sp^2-hybridization (Fig. 27A). The N^5 of MeTHF in the $MtgA_{N227A}$:MeTHF complex structure is sp^2-hybridized and complexes an additional water molecule together with Ser226 (Fig. 27B). Thus, both mutant structures mimic late reaction states in which MeTHF is either fully deprotonated in the case of $MtgA_{N227A}$ or switches between a protonated and deprotonated state in the case of $MtgA_{D102A}$. This inability to properly stabilize the protonated intermediate state is reflected in the lack of enzymatic activity of both mutants (Fig. 25C). In contrast, active WT MtgA complexes MeTHF in the protonated intermediate state that forms immediately after methyl transfer from MeCbl to THF. By stabilizing this crucial point in the reaction, the WT enzyme is able to accomplish efficient catalysis (Fig. 25A, B, and C).

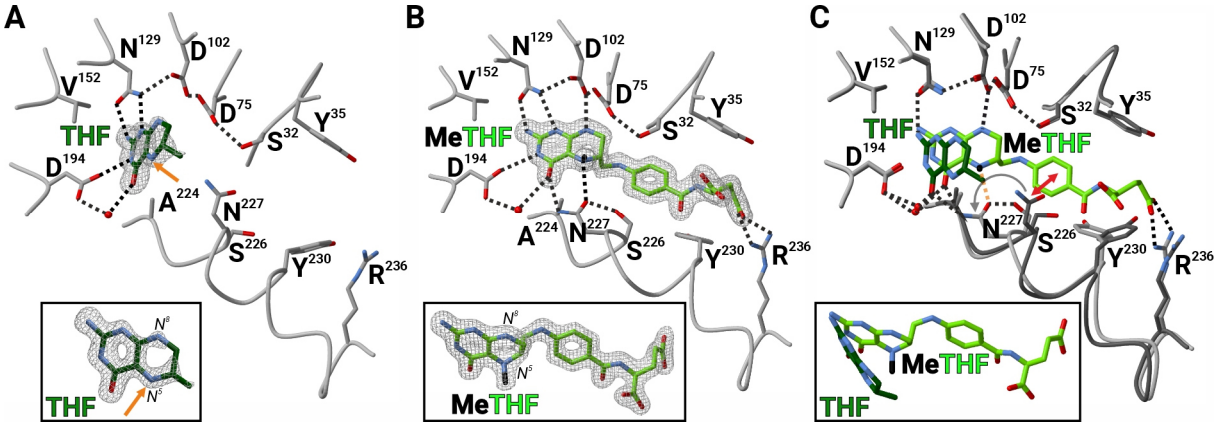


Figure 26 Active site of MtgA with complexed THF and MeTHF. **A**) MtgA in complex with its substrate THF. Residues engaged in cofactor binding are depicted as grey sticks and labeled by one-letter-code. The $2F_o$-F_c electron density map (grey meshes, contoured to 1.0 σ) is shown for the ligand in two orientations (icon below). Only defined, non-flexible moieties are depicted in the THF stick model. H-bonds are drawn as black dotted lines. The arrow (orange) points to the N^5 atom where methylation occurs. **B**) MtgA bound to MeTHF highlights a trapped reaction intermediate state (MeTHF is sp^3 hybridized). **C**) Structural superposition of MtgA:THF (dark) with MtgA:MeTHF (bright) reveals a 35° tilt of the ligand. MeTHF binding shifts Asn227 of MtgA by 2 Å (grey arrow) to avoid a clash with the THF tail (red double arrow). Figure adapted from Badmann and Groll, 2020 (s. section 7).

MtgA does not share sequence similarities with solved protein structures deposited in the PDB. However, a search for structural homologs via the DALI webserver revealed a couple of Cbl-dependent THF methyltransferases with similar folds and THF-binding motifs.[129] The closest homolog is the desulfitobacterial methyl group acceptor protein MT2DH (RMSD: 2.9 Å, sequence similarity: 32.6 %, Z-score: 20.2), which plays a role in the catabolism of aromatic methyl ethers.[180] Similarly, MtgA is related to the THF-binding domains of MetH as well as the methyltransferase MeTr, which occurs in the carbon fixating Wood-Ljungdahl pathway in acetogenic bacteria and accepts the methyl group from CoFeSP.[98,100-102] In all cases, the residues involved in THF-binding are conserved. Arginine residues assist in the binding of the glutamate tail of bound THF or MeTHF. Furthermore, a protonated THF intermediate seems to be common in Cbl-dependent THF methyltransferases[181-183] and is present in many MeTHF-bound structures.[98,100-102,180] On the other hand, conformational rearrangements of THF, as observed in the THF- and MeTHF-bound structures of MtgA, are unique among solved protein methyltransferase structures. MT2DH shows no differences between the THF- and MeTHF-bound structures.[180] The swinging motion of $Asn227_{MtgA}$ as a result of the fixation of the THF tail during methylation is so far unique to MtgA. However, structural studies on CoFeSP:MeTr suggest that this asparagine residue ($Asn199_{MeTr}$) is responsible for a variation of this locking mechanism within this methyltransferase complex.[100,102,165] While $Asn199_{MeTr}$ has the same orientation in the THF- and MeTHF-bound CoFeSP:MeTr structures, the apo structure of CoFeSP:MeTr features a displacement of $Asn199_{MeTr}$ similar to $Asn227_{MtgA}$ within the MtgA active site. For CoFeSP:MeTr, it was suggested that $Asn199_{MeTr}$ acts as a gatekeeper

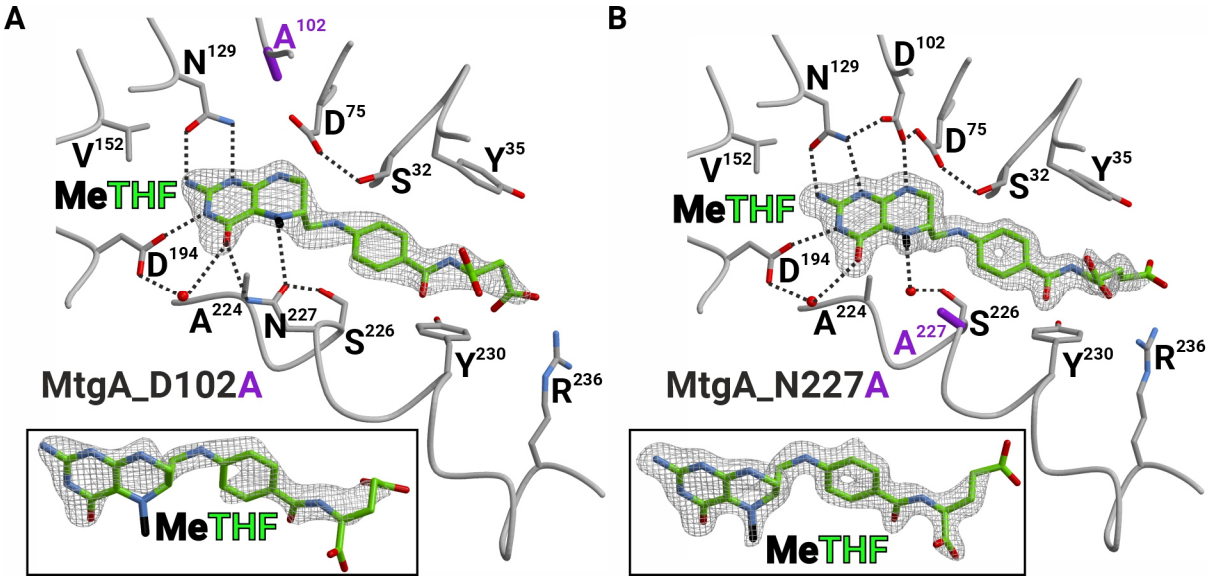


Figure 27 Active sites of MtgA D102A (**A**) and N227A (**B**) in complex with MeTHF. Residues engaged in cofactor binding are depicted as sticks and labeled by one-letter-code. The $2F_o$-F_c electron density map (grey meshes, contoured to 1.0 σ) is shown for the ligand in two orientations (icon below). H-bonds are drawn as black dotted lines. Color coding, contouring, and labeling are as in Fig. 26. Figure adapted from Badmann and Groll, 2020 (s. section 7).

to block a closer position of Cbl as long as MeTHF is absent from the protein.[100,102] MtgC and MtgA catalyze methyl transfer from MeCbl to THF. The CoFeSP:MeTr complex, on the other hand, mediates methyl transfer from MeTHF to Cbl to ultimately produce acetyl-CoA. The asparagine-mediated locking mechanism within the active sites of both proteins represents an illustrative case where a conformational mechanism is conserved across enzymes, but has been adapted to serve a completely different purpose as a necessity of the opposite reaction direction.

3 Conclusion and Outlook

Several of the initial goals of this dissertation have been accomplished. Pyrrolysine was heterologously produced in *E. coli* via the proteins PylB, PylC, and PylD. In combination with the protocol for amber suppression adapted from the group of Prof. Kathrin Lang (ETH Zurich), this amino acid was then successfully incorporated into the model protein GFP. Provided a suitable PylS mutant is available, the established amber suppression system can now be used for other ncAAs across projects at the Chair of Biochemistry.

The resources and protocols for molecular biology at the Chair of Biochemistry have been extensively updated. The new workflow revolves around the Gibson Assembly® method and is documented with the program SnapGene®. Primers, vector maps, and protocols are stored in a digital library that can be accessed from any computer in the laboratory. The improved efficiency of this setup was demonstrated by the production of several vector constructs encoding variants of the methyltransferase MttB. In addition, the implementation of this workflow allowed the design of many new vector constructs for this work and positively impacted molecular biology projects throughout the group.

3D structures of MtgB, MtcB, MtbA, and MtgA were determined, providing new insights into Cbl-dependent methyl transfer. Substrate- and product-bound structures of MtgB were used to propose a comprehensive reaction mechanism for this enzyme. Furthermore, Cbl-bound structures of MtgB provided insight into how the methyl group is transferred from GB to Cbl. The active sites of MtgB and MtcB showed little similarity to their pyrrolysine-containing relatives, MttB and MtmB. The different residue environments found in these enzymes are designed to accommodate the quaternary amine substrates. They skip the activation step that occurs in pyrrolysine-containing methyltransferases and only serve only to position the methyl group for transfer to Cbl.

The structures of substrate- and product-bound MtbA allowed a detailed mechanistic description of CoM methylation, where the coordination of the catalytic zinc ion is inverted during the reaction. Furthermore, the determined MtbA structures completed the available structural data on the enzymes involved in the pathways for methanogenic MMA and TMA degradation.[54,84] Structures of MtgA were solved in complex with THF and MeTHF, allowing the proposal of a detailed mechanism for the catalyzed methyl transfer reaction. MtgA differs from other THF methylases by locking the THF tail region during methylation in a novel way to stabilize the protonated reaction intermediate. The proposed mechanism was further supported by monitoring the activity of WT and mutant MtgA via an anaerobic assay.

Microbial Cbl-dependent methyl transfer pathways offer many other interesting topics for structural analysis. The DMA demethylase MtbB is one of only three proteins that carries

the exotic amino acid pyrrolysine in its active site, but shows no sequence similarity to its two counterparts. The structure of MtbB remains to be determined and may provide additional insight into the origin and function of pyrrolysine. Although complete structural information is now available for the individual components of the MMA, TMA, and, with the exception of MtgC, GB demethylation pathways, interactions between the different proteins have rarely been visualized. To achieve methyl transfer from the donor to the acceptor peptide, an intricate interplay of domains is required to move the large cofactor Cbl from one active site to the next. This movement requires a series of snapshot-like structural images to understand its principles. Adding new structural information to the knowledge of these processes will be the defining goal of follow-up research on this topic.

4 Materials and Methods

4.1 Materials

4.1.1 Instruments

Table 2 Instruments

Product	Supplier	Supplier Location
3-30k centrifuge, Rotors 12150-H and 12154-H	Sigma Laborzentrifugen GmbH	Osterode am Harz, DE
4k15 centrifuge, Rotor 11150	Sigma Laborzentrifugen GmbH	Osterode am Harz, DE
6-16ks centrifuge, Rotor 12500-H	Sigma Laborzentrifugen GmbH	Osterode am Harz, DE
8k centrifuge, Rotor 11805	Sigma Laborzentrifugen GmbH	Osterode am Harz, DE
Acquity Premier HPLC System	Waters	Milford MA, US
Acquity UPLC Protein BEH C4 Column	Waters	Milford MA, US
ÄKTA Prime	Cytiva	Marlborough MA, US
ÄKTA Pure, Fraction collector F9-R	Cytiva	Marlborough MA, US
ÄKTA Purifier, Frac-920	Cytiva	Marlborough MA, US
ÄKTA Start, Frac30	Cytiva	Marlborough MA, US
Analytical Balance, LA 124i Classic	VWR	Radnor PA, US
Autoclave Prestige Medical 2100 classic	Soleni	Schenkendöbern, DE
Autoclave Varioklav Classic 500	HP Medizintechnik	Oberschleißheim, DE
Branson Digital Sonifier® 250	G. Heinemann	Schwäbisch Gmünd, DE
C1000 Touch™ Thermal Cycler	Bio-Rad	Hercules CA, US
Camera SC50	Olympus	Tokyo, JP
CFX™ Real-Time System	Bio-Rad	Hercules CA, US
Cooled Incubator Series 3000	RUMED	Laatzen, DE
Electrophoresis Power Supply EPS600	Pharmacia Biotech	Uppsala, SE
Electroporation Cuvette (2 mm gap)	Peqlab	Erlangen, DE
Electroporator Eporator®	Eppendorf	Hamburg, DE
GeneAmp PCR System 2400	PerkinElmer	Waltham MA, US
Incubator ED53	Binder	Tuttlingen, DE
Incubator Multitron Standard	INFORS HT	Bottmingen, CH
InoLab pH 720 pH-Meter	WTW	Weilheim, DE
Magnetic Stirrer MR Hei-Standard	Heidolph	Schwabach, DE
Microscope SZX10/KL 1600 LED	Olympus	Tokyo, JP
Microscope SZX10/KL1500LCD	Olympus	Tokyo, JP
MJ Mini Personal Themal Cycler	Bio-Rad	Hercules CA, US
MyCycler™ Thermal Cycler	Bio-Rad	Hercules CA, US
M-power Toploading Balance AZ4101	Satorius	Göttingen, DE
Nanodrop 2000c	Thermo Fisher Scientific	Waltham MA, US
Orbital Shaker VXR basic Vibrax®	IKA	Staufen, DE
Power Supply PowerPac	Bio-Rad	Hercules CA, US
Precision Balance TE4101	Satorius	Göttingen, DE
Quantum CX5 Edge	Vilber	Collégien, FR
Quick Combi Sealer Plus	HJ-Bioanalytik	Mönchengladbach, DE
Robot Crystal Gryphon LCP	Art Robbins Instruments	Sunnyvale CA, US
Robot Microlab STARlet	Hamilton	Reno NV, US
Robot Oryx4	Douglas Instruments	Berkshire, GB
Robot Phoenix Liquid Handling System	Art Robbins Instruments	Sunnyvale CA, US
T100 Thermal Cycler	Bio-Rad	Hercules CA, US
Table centrifuge 1-14, Rotor 12097	Sigma Laborzentrifugen GmbH	Osterode am Harz, DE
Table centrifuge 1-14k, Rotor 12084	Sigma Laborzentrifugen GmbH	Osterode am Harz, DE
Thermomixer comfort (1.5 mL)	Eppendorf	Hamburg, DE
Ultrospec 7000 photometer	Cytiva	Marlborough MA, US
Ultrospec 10 Cell Density Meter	Cytiva	Marlborough MA, US
UVEX (UV microscope)	JAN Scientific	Seattle WA, US

4.1.2 Materials and Consumables

Table 3 Materials and consumables

Product	Supplier	Supplier Location	Catalog #
4-15 % Mini-PROTEAN TGX Precast Gel (BioRad)	Bio-Rad	Hercules CA, US	4561086
Amicon Ultra-15 Centrifugal Filter Unit (3kDa, 10kDa, 30kDa, or 50kDa)	Merck	Darmstadt, DE	UFC900324, UFC901024, UFC903024, UFC905024
Antarctic Phosphatase	New England Biolabs	Ipswich MA, US	M0289S
Antarctic Phosphatase Reaction Buffer	New England Biolabs	Ipswich MA, US	B0289S
BamHI-HF®	New England Biolabs	Ipswich MA, US	R3136S
Chamber and tray for agarose gels	Appligene	Heidelberg, DE	
CryoLoop	Hampton	Aliso Viejo CA, USA	HR4-336, HR4-338, HR8-102, HR8-072, HR8-104, HR8-106, HR8-108
CrystalCap HT for CryoLoop	Hampton	Aliso Viejo CA, USA	HR4-779
CrystalWand Magnetic	Hampton	Aliso Viejo CA, USA	HR4-729
Dewar Taylor Wharton CX100	tec-lab	Idstein, DE	CX100
Dewar Taylor Wharton CX20	tec-lab	Idstein, DE	CX20
Dialysis Tubing Membra-Cell™ cellulose 14000	Roth	Karlsruhe, DE	0655.1
DNA Stain G	SERVA	Heidelberg, DE	39803.01
DNA-Ladder (peqGOLD)	VWR	Radnor PA, US	N3200S
DpnI	New England Biolabs	Ipswich MA, US	R0176S
Gel Loading Dye, Purple (6x)	New England Biolabs	Ipswich MA, US	B7025S
Glassware	VWR	Radnor PA, US	
HiLoad 16/600 Superdex 200 pg preparative SEC column	Cytiva	Marlborough MA, US	28989335
HiLoad 16/600 Superdex 75 pg preparative SEC column	Cytiva	Marlborough MA, US	28989333
HisTrap HP (5 ml)	Cytiva	Marlborough MA, US	17524801
Intelli-Plates® 96-3-Well	Art Robbins Instruments	Sunnyvale CA, US	HR3-183
KpnI	New England Biolabs	Ipswich MA, US	R3142S
Microcentrifuge tubes	VWR	Radnor PA, US	LPIT111540
Mini-PROTEAN® Tetra System	Bio-Rad	Hercules CA, US	1658001FC
NcoI	New England Biolabs	Ipswich MA, US	R3193S
NdeI	New England Biolabs	Ipswich MA, US	R0111S
NEBuilder® HiFi DNA Assembly Master Mix	New England Biolabs	Ipswich MA, US	E2621L
NiNTA Spin Column Kit	Qiagen	Venlo, NL	31314
pACYCDuet-1	Merck	Darmstadt, DE	71147-3
pCDFDuet-1	Merck	Darmstadt, DE	71340-3
pET-28b(+)	Merck	Darmstadt, DE	69865-3
pETDuet-1	Merck	Darmstadt, DE	71146-3
pRSETHis6TEV	Thermo Fisher Scientific	Waltham MA, US	V35120
pRSFDuet-1	Merck	Darmstadt, DE	71341-3
PstI-HF®	New England Biolabs	Ipswich MA, US	R3140S
Q5® High-Fidelity DNA Polymerase	New England Biolabs	Ipswich MA, US	M0491S
Q5® Reaction Buffer Pack	New England Biolabs	Ipswich MA, US	B9027S, B9027SVIAL
QuikChange II Site-Directed Mutagenesis Kit	Agilent	Santa Clara CA, US	200523
rCutSmart™ Buffer	New England Biolabs	Ipswich MA, US	B6004S
Rothilabo Syringe filters, CME	Roth	Karlsruhe, DE	KH 54.1, KH55.1
SacI	New England Biolabs	Ipswich MA, US	R3156S
SalI	New England Biolabs	Ipswich MA, US	R3138S
Sealable cell 117 (1400 µL)	Hellma	Mühlheim, DE	117-104-10-40
Sonication Rosette	G. Heinemann	Schwäbisch Gmünd, DE	
Spectra/Por® Dialysis Membran Standard RC Tubing 3.5kDa	Repigen	Waltham MA, US	132720
Storage Dewar HC20	Worthington Industries	Worthington OH, US	MD7-73
T4 DNA Ligase	New England Biolabs	Ipswich MA, US	M0202S
T4 DNA Ligase Reaction Buffer	New England Biolabs	Ipswich MA, US	M0202S, B0202SVIAL
T4 Polynucleotide Kinase	New England Biolabs	Ipswich MA, US	M0201S

Unstained Protein Standard, Broad Range (10-200 kDa)	New England Biolabs	Ipswich MA, US	P7704
Wizard® Plus SV Minipreps DNA Purification Systems	Promega	Madison WI, US	A1460
Wizard® SV Gel and PCR Clean-Up System	Promega	Madison WI, US	A9282

Electrocompetent cells, SUMO protease, and TEV protease were prepared or purified by the Chair's technician Katrin Gärtner.

4.1.3 Chemicals

Unless stated otherwise, all chemicals were purchased from one of the following companies: AppliChem (Darmstadt, DE), Biomol (Hamburg, DE), Merck (Darmstadt, DE), Roth (Karlsruhe, DE), TCI (Zwijndrecht, BE), and VWR (Darmstadt, DE).

4.1.4 *E. coli* Strains

Table 4 *E. coli* strains

Antibiotic	Genotype	Supplier
BL21(DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	Merck (Darmstadt, DE)
SolubBL21(DE3)	B F^- *ompT* $hsdS_B(r_B^- m_B^-)$ *gal dcm lon* λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*])*	Genlantis (San Diego CA, US)
BL21(DE3) Tuner	F^- *ompT* $hsdS_B$ (r_B^- m_B^-) *gal dcm lacY1*(DE3)	Merck (Darmstadt, DE)
DH5α	*fhuA2Δ(argF-lacZ)U169 phoA glnV44 Φ80Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	NEB (Ipswich MA, US)

4.1.5 Antibiotic Concentrations

Table 5 Antibiotic concentrations added to cultivation media

Antibiotic	Concentration (final)
Ampicillin	100 µg/mL
Chloramphenicol	25 µg/mL
Kanamycin	50 µg/mL
Streptomycin	50 µg/mL

4.1.6 Recipes for Buffers and Media

Table 6 Recipes for buffers and media. All media were sterile filtered or autoclaved immediately after preparation.

Product	Components
Coomassie staining solution	0.05 % (*w/v*) Coomassie Brilliant Blue R250, 10 % (*v/v*) acetic acid, 25 % (*v/v*) isopropanol
Destaining solution	10 % (*v/v*) acetic acid
Laemmli sample buffer (2x)	65.8 mM Tris/HCl pH 6.8, 26.2 % (*v/v*) glycerol, 2.1 % (*w/v*) SDS, 0.01 % (*w/v*) bromophenol blue, mixed 19:1 (*v/v*) with β-mercaptoethanol (β-ME) directly before use
LB medium	1 % (*w/v*) peptone, 0.5 % (*w/v*) yeast extract, 0.5 % (*w/v*) NaCl
M9 medium	33.7 mM Na_2HPO_4, 22.0 mM KH_2PO_4, 9.4 mM NH_4Cl, 8.6 mM NaCl, 1.0 mM $MgSO_4$, 0.3 mM $CaCl_2$, 0.4 % (*w/v*) glucose, 1 µg/mL biotin, 1 µg/mL niacinamide, 1 µg/mL pyridoxine, 1 µg/mL thiamin, 134 µM EDTA, 31.0 µM $FeCl_3$, 6.2 µM $ZnCl_2$, 1.62 µM H_3BO_3, 0.76 µM $CuCl_2$, 0.42 µM $CoCl_2$, 0.01 µM $MnCl_2$, 0.01 µM molybdate, 0.01 µM Ni_2SO_4
SDS running buffer	25 mM Tris base, 0.1 % (*w/v*) SDS, 192 mM glycine
SDS separating gel buffer	1.5 M Tris/HCl pH 8.8, 0.4 % (*w/v*) SDS
SDS stacking gel buffer	0.5 M Tris/HCl pH 6.8, 0.4 % (*w/v*) SDS
SOC medium	2 % (*w/v*) peptone, 0.5 % (*w/v*) yeast extract, 20 mM glucose, 10 mM MgSO4, 10 mM MgCl2, 10 mM NaCl, 2.5 mM KCl
TAE buffer	40 mM Tris base, 1 mM EDTA, 20 mM acetic acid, pH 8.2
TB medium	1.2 % (*w/v*) peptone, 2.4 % (*w/v*) yeast extract, 0.4 % (*w/v*) glycerol
TB salts (10x)	170 mM KH_2PO_4, 720 mM K_2HPO_4

Table 7 Crystallization broad screens were purchased from Qiagen (Venlo, NL).

Screen (with 96 conditions each)	**Name**	**Catalog #**
1	pHClear Suite	1348-01 to 96
2	pHClear II Suite	1349-01 to 96
3	Classics Suite[184]	1340-01 to 96
4	Classics II Suite	1361-01 to 96
5	JCSG+ Suite[185,186]	1359-01 to 96
6	ComPAS Suite[187]	1356-01 to 96
7	Protein Complex Suite[187,188]	1354-01 to 96
8	PEGs Suite	1343-01 to 96
9	PEGs II Suite	1355-01 to 96
10	PACT Suite[186]	1357-01 to 96
11	MPD Suite	1345-01 to 96
12	$AmSO_4$ Suite	1344-01 to 96
13	Cations Suite	1347-01 to 96
14	Anions Suite	1346-01 to 96
15	Cryos Suite[184]	1342-01 to 96
16	Classics Lite Suite[184]	1341-01 to 96
17	MbClass Suite	1350-01 to 96
18	MbClass II Suite	1351-01 to 96

4.1.7 Software, Databases, and Bioinformatic Tools

Table 8 Software

Product	Supplier	Supplier Location	Purpose
BioRender	BioRender	Toronto, CA	Graphical design
CFX Maestro 2.3	Bio-Rad	Hercules CA, US	Real-Time PCR
ChemDraw Professional v. 22.2.0.3300	PerkinElmer	Waltham MA, US	Chemical structure schemes
CorelDRAW 2023 v. 24.2	Allundo	Ottawa, CA	Graphical design
GraphPad Prism v. 5	GraphPad Software	Boston MA, US	T-Test
Gryphon software v. 1.5.0.1	Art Robbins Instruments	Sunnyvale CA, US	Robot control
Hamilton Run Control v. 4.2.0.6424	Hamilton	Reno NV, US	Robot control
Hamilton Screen Designer v. 1.3.109	Hamilton	Reno NV, US	Fine screen design
MassLynx V4.2	Waters	Milford MA, US	Mass spectrometry
Microsoft Office 365	Microsoft	Redmond WA, US	Text editing and diagram illustrations
NanoDrop2000c v. 1.6.198	Thermo Fisher Scientific	Waltham MA, US	NanoDrop control
Phoenix software	Art Robbins Instruments	Sunnyvale CA, US	Robot control
PrimeView 5.31	Cytiva	Marlborough MA, US	ÄKTA control and evaluation
PyMOL v. 2.5.7	Schrödinger	New York NY, US	Structure illustrations
SnapGene v. 7.1.0	GSL Biotech LLC	Boston MA, USA	Vector map design
Unicorn 7	Cytiva	Marlborough MA, US	ÄKTA control of Pure and Purifier systems
Unicorn 1.3	Cytiva	Marlborough MA, US	ÄKTA control of Prime systems
Wasp Run v. 3.01.19	Douglas Instruments	Berkshire, GB	Robot control

Table 9 Bioinformatic tools, data bases, and open-source software

Product	Purpose	Website
AlphaFold2 Colab[124]	Structure prediction	colab.research.google.com
ARP/wARP solvent[189]	Solvent structure building	ccp4.ac.uk
BLAST[190]	Local Alignment Search	blast.ncbi.nlm.nih.gov/Blast.cgi
CATH[127]	Protein structure topology categorization	cathdb.info
CCP4i[191]	Structure building	ccp4.ac.uk
CCP4i2[192]	Structure building	ccp4.ac.uk
Coot v. 0.8.6[193]	Structure building and visualization	mrc-lmb.cam.ac.uk
CRANK2[194]	Structure solution for experimental phasing	ccp4.ac.uk
DALI[129]	Protein structure comparison	ekhidna2.biocenter.helsinki.fi/dali
DM[195]	Solvent flattening	ccp4.ac.uk
EMBOSS Needle	Pairwise Sequence Alignment	ebi.ac.uk
ExPASy ProtParam[196]	Peptide properties	web.expasy.org/protparam
fft[197]	Electron density depictions	ccp4.ac.uk
Illustrate[198]	Structure illustrations	ccsb.scripps.edu
Jpred 4[199]	Secondary Structure Prediction	compbio.dundee.ac.uk/jpred
MolProbity v. 4.5.2[200]	Structure evaluation	http://molprobity.biochem.duke.edu/
Molscript v. 2.1[201]	Structure illustrations	www-structmed.cimr.cam.ac.uk
PDBe[202]	Protein structure database	ebi.ac.uk/pdbe
PDBePISA v. 1.52[128]	Crystal packing analysis	ebi.ac.uk/msd-srv/prot_int/pistart.html
Phaser v. 2.8.3[203]	Phasing macromolecular crystal structures	phenix-online.org/documentation
Phenix[204]	Structure building	phenix-online.org
QuikChange Primer Design[205]	Primer design for mutagenesis	agilent.com/store
RCSB PDB[206]	Protein structure database	rcsb.org
REFMAC5[207,208]	Macromolecular refinement	ccp4.ac.uk
SHARP-SAD[209]	Experimental phasing	ccp4.ac.uk
SHELXD[210]	Structure solution program	ccp4.ac.uk
UniProt[211]	Protein database	uniprot.org
XDS program package[212]	Diffraction data processing	xds.mr.mpg.de

4.1.8 Vector Constructs, Primers, Cultivation Conditions, and Purification Buffers

WT gene sequences were either obtained from genomic DNA of *M. barkeri* (DSM 800) or *D. hafniense* (DSM 10664), or purchased as synthetic genes codon-optimized for *E. coli*. Genomic DNA was purchased from the German Collection of Microorganisms and Cell Cultures GmbH (DSMZ, Braunschweig, DE). Synthetic genes were purchased from Eurofins Genomics (Ebersberg, DE). The Uniprot-IDs of the respective gene products are provided in Table 10.

Table 10 Gene name (with Uniprot-ID of the gene product and DSMZ number of the genomic DNA; sg = synthetic gene), construct code, vector, and notable construct features are listed for all vector constructs that were used in this work. The exact vector maps and sequences can be accessed in the common local network drive of the Chair of Biochemistry under \\europa.intern.ch.tum.de\BC\Archiv\Thomas Badmann; vector backbone maps can be found under \\europa.intern.ch.tum.de\BC\Gruppe\Laborорganisation\Klonierung, Primer, Ecoli-Stämme\sonstiges - Klonierung\Vektor-Collection\DNA Files. The vector constructs are stored at -20 °C in the archive storage drawer.

Gene (Uniprot-ID, DSM number)	**Construct code**	**Vector**	**Construct features**
mtmB | (O30642, sg)	B1-6a	pET28bSUMOSer	O202K mutant
	B1-7a	pRSETHis6TEV	O202K mutant
	B2-6a	pRSETHis6TEV	O202A mutant
	B2-7a	pET28bSUMOSer	O202A mutant
mtbB | (O93661, sg)	D1-7a	pET28bSUMOSer	O356K mutant
	D2-6a	pRSETHis6TEV	O356A mutant
	D2-7a	pET28bSUMOSer	O356A mutant
mttB | (O93658, sg)	F1-7a	pET28bSUMOSer	O334K mutant
	F2-6a	pRSETHis6TEV	O334A mutant
	F2-7a	pET28bSUMOSer	O334A mutant
	F8-5b	pRSFDuet-Amber	WT
	F10-2a	pACYCDuet-1	O334A mutant
	F12-1f	pETDuet-1-TEV	O334A mutant, N-terminal deletion, C-terminal truncation 469TAA
	F13-1f	pETDuet-1-TEV	O334A mutant, N-terminal deletion, C-terminal truncation 451TAA
	F14-1f	pETDuet-1-TEV	O334A mutant, N-terminal deletion, C-terminal truncation 431TAA
	F15-1f	pETDuet-1-TEV	O334A mutant, N-terminal deletion, C-terminal truncation 417TAA
	F16-1f	pETDuet-1-TEV	O334A mutant, N-terminal deletion, C-terminal truncation 396TAA
	F17-1f	pETDuet-1-TEV	O334A mutant, N-terminal deletion, C-terminal truncation 334TAA
	F19-1c	pETDuet-MBP	O334A mutant, N-terminal MBP-tag
pylB (Q8NKQ5), pylC (Q8NKQ4), pylD (A0A832TJ77) | (DSM 800)	XYZ1-3c (or X2Y1Z2-3c)	pCDFDuet-1	Co-expression of *pylB*, *pylC*, and *pylD*
pylB (Q8NKQ5), pylC (Q8NKQ4) | (DSM 800)	YZ1-3c (or Y1Z2-3c)	pCDFDuet-1	Co-expression of *pylC* and *pylD*
*gfp*149TAG (ALP48449.1)	5b	pRSFDuet-Amber	
gfp (ALP48449.1)	W1-1a	pETDuet-1	
mtgB | (Q24SP7, DSM 10664)	I1-7a	pET28bSUMOSer	
	I2-1a	pETDuet-1	C-terminal His_6-tag
mtcB | (A0A317RMF0, sg)	O1-7a	pET28bSUMOSer	
mtbA | (O30640, sg)	A1-6a	pRSETHis6TEV	
mtgA | (Q24SP6, sg)	L1-7a	pET28bSUMOSer	
	L2-7a	pET28bSUMOSer	D102A mutant
	L7-7a	pET28bSUMOSer	N227A mutant
	L8-7a	pET28bSUMOSer	N227D mutant
	L9-6a	pRSETHis6TEV	R236A mutant

Table 11 Primers used for cloning of the constructs in Table 11. All primers were purchased from Eurofins Genomics (Ebersberg, DE).

Restriction and ligation cloning			
Gene	**Construct Code**	**Primer**	**Sequence (5' → 3')**
mtmB	B1-6a, B1-7a	MtmB_BamHI_for	CCAGGATCCACATTTCGTAAATCCTTTG
		MtmB_PstI_rev	CCACTGCAGTTAGAACACCAGACCCAGATC
mtbB	D1-7a	MtbB_BamHI_for	CCAGGATCCGCCACGGAGTATGCCCTC
		MtbB_PstI_rev	CCACTGCAGTTAGGCAATCTGTTTGCGAAAC
mttB	F1-7a	MttB_BamHI_for	CCAGGATCCGCGAAGAATAACGCGGTGG
		MttB_PstI_rev	CCACTGCAGTTACATGCCGCGAAAGGC
	F8-5b	MttB_PstI_for	TAAGCACTGCAGATGGCGAAGAATAACGCGGTG
		MttB_XhoI_rev	TAAGCACTCGAGTTAGTGGTGATGGTGATGATGCATGCCGCGAAAGGCC
	F10-2a	MttB_BamHI_for-2	TAAGCAGGATCCGGCGAAGAATAACGCGGTG
		MttB_PstI_rev-2	TAAGCACTGCAGTTACATGCCGCGAAAGGC
mtgB	I1-7a	MtgB_SpeI_for	TAAGCAACTAGTGGTCAATTGTTACCTAAGTAC
		MtgB_XhoI_rev	TAAGCACTCGAGTTACTCTTCTCCAACTTCC
mtcB	O1-7a	MtcB_BamHI_for	CCAGGATCCATCCGCAACTCACTGACTG
		MtcB_PstI_rev	CCACTGCAGTTAAATGTCGAACTTGTACTGTTCCG
mtbA	A1-6a	MtbA_BamHI_for	CCAGGATCCGCCGAATACACCCCGAAAG
		MtbA_XhoI_rev	CCACCACTCGAGTTAATAAGTATGCGATTTGGC
mtgA	L1-7a	MtgA_BamHI_for	CCAGGATCCTTCAAGTTTACTGCCCAACAAC
		MtgA_PstI_rev	CCACTGCAGTTAAAAGATTTTCAGTAACGGATGTTC
pylB, pylC, pylD	XYZ1-3c (or X2Y1Z2-3c)	PylB_NcoI_for	GCTCCATGGCACTTGATGAATTTGACAG
		PylB_BamHI_rev	CCAGGATCCTTACAGGACTGCCTCAAAGTC
		PylC_PstI_for	CCACTGCAGATGAAAACGATCTGTCTTGTGG
		PylC_SalI_rev	CCAGTCGACTTATAAAACCGCTCCAAAACG
		PylD_NdeI_for	CCACATATGGCACTCTTAACTCCCGATG
		PylD_KpnI_rev	CCAGGTACCTTATAGCACAGAGTACAACATTG
pylC, pylD	YZ1-3c (or Y1Z2-3c)	PylC_PstI_for, PylC_SalI_rev, PylD_NdeI_for, PylD_KpnI_rev	s. above
Gibson Assembly			
Gene	**Construct Code**	**Primer**	**Sequence (5' → 3')**
	-	Vector primer 1	CAGAGCATTTGACTGGAAGTAGAGATTCTCGTGGTGATGATGGTGATGGC
	-	Vector primer 2	CTCGAGTCTGGTAAAGAAACCGCT
	-	Vector primer 3 (ga7)	TGACTGGAAGTAGAGATTCTCGATCCC
	-	Vector primer 4 (ga5)	CCCTCGAGTCTGGTAAAGAAACCG
	-	Vector primer 5 (ga13)	GGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGG
	-	Vector primer 6 (ga1)	TGCTTAAGTCGAACAGAAAGTAATCGTATTG
	-	Vector primer 7 (ga2)	GTGGTGATGATGGTGATGGCT
mttB	F12-1f (vector primers 1 and 2)	MttB_for-1	ATCACCACGAGAATCTCTACTTCCAGTCAAATGCTCTGAATGGCGTTGAA
		MttB_rev-1	GTTTCTTTACCAGACTCGAGCTTTAATGATTCTTTAAGACGTCCTCAACCT
	F13-1f (vector primers 1 and 2)	MttB_for-1	s. above
		MttB_rev-2	GTTTCTTTACCAGACTCGAGCTTTACGACCCCGCAGCG
	F14-1f (vector primers 1 and 2)	MttB_for-1	s. above
		MttB_rev-3	GTTTCTTTACCAGACTCGAGCTTTAGTAATCGACTAATTGCCTCGTTTGT
	F15-1f (vector primers 1 and 2)	MttB_for-1	s. above
		MttB_rev-4	GTTTCTTTACCAGACTCGAGCTTTAGTTGCCAATACCAACTTTCTGAATAG ATTCT
	F16-1f (vector primers 1 and 2)	MttB_for-1	s. above
		MttB_rev-5	GTTTCTTTACCAGACTCGAGCTTTACTGCATCGCTTTTTTGACCATCG
	F17-1f (vector primers 1 and 2)	MttB_for-1	s. above
		MttB_rev-6	GTTTCTTTACCAGACTCGAGCTTTATGAGCCAGCCACGTAACTGG
	F19-1c (vector primers 1 and 2)	MttB_for-2	AGAATCTCTACTTCCAGTCAAATGCTCTGAATGGCGTTGAATTAAAC
		MttB_rev-7	TTTACCAGACTCGAGGGTTACATGCCGCGAAAGGC
gfp	W1-1a (vector primers 4 and 5)	GFP_for	TTTAACTTTAAGAAGGAGATATACCATGCCGAGCAAAGGTGAAGAACTGT TTACCGG
		GFP_rev	TTTACCAGACTCGAGGGTTAATGGTGATGATGATGGTGGCTGCCTTTATA CAGTTCATCCATACC
mtgB	I2-1a (vector primers 5 and 6)	MtgB_for	TTTAACTTTAAGAAGGAGATATACCATGGGTCAATTGTTACCTAAGTACAA TATCCTGACC
		MtgB_rev	TTTCTGTTCGACTTAAGCATTAGTGGTGATGGTGATGATGGCCGCCCTCTT CGCCAACTTCCTTACC
mtgC	K3-1a (vector primers 6 and 7)	MtgC_for	CATCACCATCATCACCACATAATGTCCTTGTTAGATGAACTAAAACAAGC C
		MtgC_rev	TTTCTGTTCGACTTAAGCATTATTTCCCGAGAATTTTGTTCGCTAGC

Mutagenesis			
Gene	**Construct Code**	**Primer**	**Sequence (5′ → 3′)**
mtmB	B2-6a, B2-7a	MtmB_202A_for	CGCCCAGGCATGGGAGTAGCGGGTCCGGAAAC
		MtmB_202A_rev	GTTTCCGGACCCGCTACTCCCATGCCTGGGCG
mtbB	D2-6a, D2-7a	MtbB_356A_for	CGGGAGTTGATGGGATCGCGATTGGCGTAGGCGATC
		MtbB_356A_rev	GATCGCCTACGCCAATCGCGATCCCATCAACTCCCG
mttB	F2-6a, F2-7a	MttB_334A_for	GTTACGTGGCTGGCTCAGCGTCAGACGCCAAGGTGC
		MttB_334A_rev	GCACCTTGGCGTCTGACGCTGAGCCAGCCACGTAAC
mtgA	L2-7a	MtgA_D102A_for	GCTCCATTTCTCCTGGCTTCGATCTCTCCCGAT
		MtgA_D102A_rev	ATCGGGAGAGATCGAAGCCAGGAGAAATGGAGC
	L7-7a	MtgA_N227A_for	GCGGTTGCGCGCCGAGTGCCGCCGTGTA
		MtgA_N227A_rev	TACACGGCGGCACTCGGCGCGCAACCGC
	L8-7a	MtgA_N227D_for	GGTTGCGCGCCGAGTGACGCCGTG
		MtgA_N227D_rev	CACGGCGTCACTCGGCGCGCAACC
	L9-6a	MtgA_R236A_for	TGTACCTGTGGAAAAAAATGGCCAGCAAAGGCACCCC
		MtgA_R236A_rev	GGGGTGCCTTTGCTGGCCATTTTTTTCCACAGGTACA
Sequencing primers			
Primer Name		**Primer Code**	
T7		s1	TAATACGACTCACTATAGGG
pRSET-RP		s2	ATGCTAGTTATTGCTCAGC
SUMO-Seq		s3	GATTTGGACATGGAGGATAACG
pET-RP		s4	CTAGTTATTGCTCAGCGG
pET_Upstream_#69214-3		s5	ATGCGTCCGGCGTAGA
DuetUP2 _#71180-3		s6	TTGTACACGGCCGCATAATC
DuetDOWN1_#71179-3		s7	GATTATGCGGCCGTGTACAA
T7 Terminator_#69337-3		s8	GCTAGTTATTGCTCAGCGG
DuetSUMO-Seq		s9	GACCTGGACATGGAAGACAAC
ACYCDuetUP1_#71178-3		s10	GGATCTCGACGCTCTCCCT

Table 12a Protein name, construct code of the parent vector, *E. coli* expression strain, cultivation medium, fusion tag, buffers used during purification, and associated figures with diagrams and gels of every large-scale purification occurring in this work. (Abbreviations: nc = tag not cleavable, SD75/200 = HiLoad 16/600 Superdex 75 or 200 pg column)

Protein	**Construct**	**Expression strain**	**Medium**	**Tag removal**	**SEC**	**Diagrams and gels**
GFP_{N149O}	5b (+XYZ1-3c)	BL21(DE3)	TB	nc	SD75	Figure S1
GFP	W1-1a	BL21(DE3)	TB	nc	SD75	Figure S2
GFP_{N149O*}	5b (+XYZ1-3c)	BL21(DE3)	TB	nc	-	Figure S3
MttB	F19-1c	BL21(DE3)	TB	yes, His_6-TEV	SD200	Figure S4
MtgB	I1-7a	BL21(DE3)	TB	yes, SUMO	SD200	Figure S5
MtgB	I2-1a	BL21(DE3)	TB	nc	SD200	Figure S6
MtcB	O1-7a	BL21(DE3)	TB	yes, SUMO	SD200	Figure S7
MtbA	A1-6a	BL21(DE3)	LB	yes, His_6-TEV	SD75	Figure S8
MtgA	L1-7a	BL21(DE3)	LB	yes, SUMO	SD200	Figure S9
MtgA (SeMet)	L1-7a	BL21(DE3)	M9	yes, SUMO	SD200	Figure S10
$MtgA_{D102A}$	L2-7a	BL21(DE3)	LB	yes, SUMO	SD200	
$MtgA_{N227A}$	L7-7a	BL21(DE3)	LB	yes, SUMO	SD200	
$MtgA_{N227D}$	L8-7a	BL21(DE3)	LB	yes, SUMO	SD200	
$MtgA_{R236A}$	L9-7a	BL21(DE3)	LB	yes, His_6-TEV	SD200	

Table 13b Protein name, construct code of the parent vector, and buffers used during purification of the proteins listed in Table 13a.

Protein	Construct	Buffer A	Buffer B	Buffer C	Buffer D
GFP_{N149O}	5b (+XYZ1-3c)	100 mM Tris pH 7.5 500 mM NaCl 20 mM Imidazole pH 7.5 5 mM β-ME	100 mM Tris pH 7.5 500 mM NaCl 500 mM Imidazole pH 7.5 5 mM β-ME	20 mM Tris pH 7.5 100 mM NaCl 1 mM β-ME	20 mM Tris pH 7.5 100 mM NaCl 1 mM DTT
GFP	W1-1a	100 mM Tris pH 7.5 500 mM NaCl 20 mM Imidazole pH 7.5 1 mM β-ME	100 mM Tris pH 7.5 500 mM NaCl 500 mM Imidazole pH 7.5 1 mM β-ME	20 mM Tris pH 7.5 100 mM NaCl 1 mM β-ME	20 mM Tris pH 7.5 100 mM NaCl 1 mM DTT
GFP_{N149O*}	5b (+YZ1-3c)	100 mM Tris pH 7.5 500 mM NaCl 20 mM Imidazole pH 7.5 5 mM β-ME	100 mM Tris pH 7.5 500 mM NaCl 500 mM Imidazole pH 7.5 5 mM β-ME	-	-
MttB	F19-1c	100 mM Tris pH 7.5 500 mM NaCl 10 % Glycerol 20 mM Imidazole pH 7.5 5 mM β-ME	100 mM Tris pH 7.5 500 mM NaCl 10 % Glycerol 500 mM Imidazole pH 7.5 5 mM β-ME	20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol 5 mM β-ME	20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol 5 mM DTT
MtgB	I1-7a	20 mM Sodium phosphate pH 7.4 500 mM NaCl 20 mM Imidazole pH 7.4 1 mM β-ME	20 mM Sodium phosphate pH 7.4 500 mM NaCl 500 mM Imidazole pH 7.4 1 mM β-ME	20 mM Sodium phosphate pH 7.4 100 mM NaCl 1 mM β-ME	20 mM Tris pH 7.5 100 mM NaCl 1 mM DTT
MtgB	I2-1a	20 mM Sodium phosphate pH 7.4 500 mM NaCl 10 % Glycerol 20 mM Imidazole pH 7.4 1 mM β-ME	20 mM Sodium phosphate pH 7.4 500 mM NaCl 10 % Glycerol 500 mM Imidazole pH 7.4 1 mM β-ME	-	20 mM Sodium phosphate pH 7.4 150 mM NaCl 5 % Glycerol 1 mM DTT
MtcB	O1-7a	20 mM Sodium phosphate pH 7.4 500 mM NaCl 10 % Glycerol 30 mM Imidazole pH 7.4 3 mM β-ME	20 mM Sodium phosphate pH 7.4 500 mM NaCl 10 % Glycerol 500 mM Imidazole pH 7.4 3 mM β-ME	50 mM MOPS pH 7.2 200 mM NaCl 10 % Glycerol 3 mM β-ME	50 mM MOPS pH 7.2 200 mM NaCl 10 % Glycerol 3 mM DTT
MtbA	A1-6a	100 mM Tris pH 7.5 500 mM NaCl 20 mM Imidazole pH 7.5 5 mM β-ME	100 mM Tris pH 7.5 500 mM NaCl 500 mM Imidazole pH 7.5 5 mM β-ME	20 mM Tris pH 7.5 100 mM NaCl 5 mM β-ME	20 mM Tris pH 7.5 100 mM NaCl 5 mM DTT
MtgA	L1-7a	100 mM Tris pH 7.5 500 mM NaCl 10 % Glycerol 20 mM Imidazole pH 7.5	100 mM Tris pH 7.5 500 mM NaCl 10 % Glycerol 500 mM Imidazole pH 7.5	20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol	20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol
MtgA (SeMet)	L1-7a	100 mM Tris pH 7.5 500 mM NaCl 10 % Glycerol 20 mM Imidazole pH 7.5 5 mM β-ME	100 mM Tris pH 7.5 500 mM NaCl 10 % Glycerol 500 mM Imidazole pH 7.5 5 mM β-ME	20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol 5 mM β-ME	20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol 5 mM DTT
$MtgA_{D102A}$	L2-7a	100 mM Tris pH 7.5 500 mM NaCl 10 % Glycerol 20 mM Imidazole pH 7.5	100 mM Tris pH 7.5 500 mM NaCl 10 % Glycerol 500 mM Imidazole pH 7.5	20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol	20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol
$MtgA_{N227A}$	L7-7a	100 mM Tris pH 7.5 500 mM NaCl 10 % Glycerol 20 mM Imidazole pH 7.5	100 mM Tris pH 7.5 500 mM NaCl 10 % Glycerol 500 mM Imidazole pH 7.5	20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol	20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol
$MtgA_{N227D}$	L8-7a	100 mM Tris pH 7.5 500 mM NaCl 10 % Glycerol 20 mM Imidazole pH 7.5	100 mM Tris pH 7.5 500 mM NaCl 10 % Glycerol 500 mM Imidazole pH 7.5	20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol	20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol
$MtgA_{R236A}$	L9-7a	100 mM Tris pH 7.5 500 mM NaCl 10 % Glycerol 20 mM Imidazole pH 7.5	100 mM Tris pH 7.5 500 mM NaCl 10 % Glycerol 500 mM Imidazole pH 7.5	20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol	20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol

4.2 Methods

4.2.1 Polymerase Chain Reaction

Polymerase chain reaction (PCR)[213] was used to amplify gene inserts and vectors for cloning (Primers s. Table 11). Reactions were performed using Q5® High-Fidelity DNA Polymerase according to manual instructions. PCR-products were analyzed via agarose gel electrophoresis and purified using the Wizard® SV Gel and PCR Clean-Up System according to manual instructions.

4.2.2 Agarose Gel Electrophoresis

DNA fragments were analyzed and separated according to their size by agarose gel electrophoresis. 1 % (*w/v*) agarose gels in TAE buffer, supplemented with 2 µL DNA Stain G, were prepared. DNA samples were mixed with gel loading dye and loaded alongside 5 µL of DNA ladder to evaluate fragment sizes. The electrophoresis was carried out in TAE buffer for 50 min at 120 V. The UV fluorescence of Stain G-stained DNA bands was visualized using the Quantum CX5 Edge gel documentation system.

4.2.3 DNA Restriction and Dephosphorylation

4.2.3.1 In Preparation for Ligation

In preparation for ligation, vector and insert DNA were digested with appropriate restriction enzymes (Table 11). In a total volume of 50 µL, 1 µg DNA was mixed with 10 U of each restriction enzyme in rCutSmart™ buffer. The reaction was incubated at 37 °C for 1 h. Digested vector DNA was additionally dephosphorylated by adding 5 U Antarctic Phosphatase and 6 µL Antarctic Phosphatase Reaction Buffer to the digested mixture. The total volume was adjusted to 60 µL by adding water. The reaction mixture was incubated at 37 °C for 30 min. All resulting DNA fragments were purified using the Wizard® SV Gel and PCR Clean-Up System according to manual instructions.

4.2.3.2 For Control Digestion of Cloned Vectors

Screening for successfully cloned vectors was performed by digestion with selected restriction enzymes and agarose gel analysis of the resulting fragments. In a total volume of 10 µL, 100 to 200 ng DNA were mixed with 4 U of each restriction enzyme in rCutSmart™ buffer. The reaction was incubated at 37 °C for 1 h and subsequently analyzed via agarose gel electrophoresis.

4.2.4 Restriction-based Ligation

Digested insert and vector DNA fragments were ligated using T4 ligase. 50 ng of digested vector DNA were mixed with a compatible insert in a molar ratio of 3 to 1 in a total

volume of 8.5 µL. The mixture was centrifuged for 1 min, incubated at 55 °C for 10 min, and incubated on ice for 5 min. T4 Ligase Buffer and 0.5 U T4 Ligase were added, resulting in a final reaction volume of 10 µL. The mixture was incubated at room temperature (RT) for 1 h.

4.2.5 Gibson Assembly

As an alternative to restriction-based ligation, several constructs were generated using Gibson Assembly®.[123] Suitable primers for each reaction were designed using the SnapGene® cloning software and the instructions provided in the NEBuilder® HiFi DNA Assembly Master Mix manual (Table 11). Amplified fragments of linear vector DNA and insert DNA with compatible overlaps were generated via PCR and purified using the Wizard® SV Gel and PCR Clean-Up System. 40 ng of amplified vector DNA was mixed with a 2:1 molar ratio of amplified insert DNA. The reaction mixture was completed by adding water and 2xNEBuilder® HiFi DNA Assembly Master Mix to achieve a total volume of 8 µL. The reaction was incubated at 50 °C for 1 h in a PCR cycler.

4.2.6 Mutagenesis

Mutagenesis of cloned vectors was performed using the QuikChange II Site-Directed Mutagenesis Kit. Primers were designed using the QuikChange Primer Design Tool available at agilent.com. The reaction was performed according to the manual instructions provided by Agilent.

4.2.7 Concentration Determination of DNA and Protein Solutions

The concentrations of DNA and protein solutions were determined using a nano-photometer. 1 µL of solution was applied to the photometer and used for absorbance determination at 260 nm (DNA) or 280 nm (proteins). Concentrations were calculated by applying the Lambert-Beer law. For DNA, an average extinction coefficient of 0.020 $(\mu g/\mu L)^{-1}$ cm^{-1} was employed by the Nanodrop software. Protein extinction coefficients were calculated manually for each individual protein using ExPASy ProtParam.

4.2.8 Transformation

Electroporation[214] was used to introduce vector DNA into electrocompetent *E. coli* strains. 45 µL of electrocompetent cells were thawed from storage at -80 °C and mixed with either 50 pg of purified vector DNA or 1 µL of the finished ligation, Gibson Assembly, or mutagenesis reactions. For transformation with more than one vector, 5 ng of each vector were used. The mixture was transferred to a 2 mm electroporation cuvette and exposed to an electric pulse (5.0 - 6.0 ms and 2.5 kV) using an electroporator. Transformed cells were resuspended in 1 mL SOC medium and incubated for 1 h at 37 °C while shaking at 750 rpm.

Bacterial cultures were centrifuged at 5,000 × g for 2 min, resuspended in 200 µL SOC medium, and plated on LB/agar supplemented with the vector selecting antibiotic. Agar plate cultures were incubated at 37 °C over night.

4.2.9 Cultivation of *E. coli*

4.2.9.1 Media Preparation

Media were prepared using the components listed in Table 6 and autoclaved at 121 °C for 20 min. TB medium and TB salts were autoclaved separately to prevent salt precipitation. Similarly, several components in the M9 medium were sterilized separately to avoid precipitation or degradation (Table 13).

Table 13 M9 stock solutions

Category	Stock concentration	Sterilization
Na_2HPO_4, KH_2PO_4, NH_4Cl, NaCl (pH adjusted to 7.2 with NaOH)	20x	autoclaved
$MgSO_4$	1000x	autoclaved
$CaCl_2$	1000x	autoclaved
Glucose	50x	autoclaved
Biotin, niacinamide, pyridoxine, thiamin	1000x	sterile-filtered
EDTA and trace elements	100x	sterile-filtered

4.2.9.2 For Cloning Experiments

For cloning experiments, *E. coli* were cultivated by transferring a single LB/agar-grown colony into 5 mL LB medium supplemented with an antibiotic selecting for the transformed vector. Cells were incubated at 37 °C over night while shaking at 130 rpm.

4.2.9.3 For Protein Expression

For protein expression, a single LB/agar-grown *E. coli* colony, transformed with one or several different expression vectors, was transferred to 20 mL TB medium (completed with TB salts and vector selecting antibiotic(s)). Cells were grown over night at 37 °C while shaking at 130 rpm. The complete culture was transferred to 2 L TB medium (completed with TB salts and antibiotic(s)) and incubated during shaking at 130 rpm at 37 °C until an OD_{600} of 1.5 to 2.0 (TB) was reached. After cooling for 30 min at 4 °C, protein expression was induced with 1 mM IPTG, and the cells were incubated over night (approximately 20 h) at 20 °C during constant shaking at 130 rpm. Grown cells were centrifuged at 5,000 × g for 30 min, re-suspended in 30 mL 0.9 % (*w/v*) NaCl, and centrifuged again at 5000 × g for 30 min. The resulting cell pellet was stored at -20 °C.

Expression in LB medium was performed similarly, except that expression was induced at an OD_{600} of 0.6. Amber suppression experiments were performed as described, but incubated at 30 °C after induction. Project-specific information on cultivation parameters is listed in Table 12a.

4.2.9.4 Selenomethionine (SeMet) Labeling of Proteins

SeMet-labeled protein was produced using a modified version of a previously described protocol.[215] Expression was performed in *E. coli* BL21(DE3). A single colony, transformed with the expression vector, was transferred to 5 mL LB medium and grown at 37 °C for 8 h. During all incubation steps, the cultures were shaken at 130 rpm. 300 mL M9 medium, supplemented with the vector selecting antibiotic, were inoculated with 300 µL pre-culture and incubated at 37 °C over night. From this second pre-culture, 3 L antibiotic-supplemented M9 medium were inoculated and incubated at 37 °C until an OD_{600} of 0.6 was reached. 0.3 g of each, lysine, threonine, and phenylalanine, as well as 0.15 g of each, leucine, isoleucine, valine, and SeMet were added. After incubation for 15 min at 37 °C, expression was induced by adding 1 mM IPTG. Cells were incubated at 20 °C over night (approximately 20 h), harvested at 5,000 × g for 30 min, washed in 0.9 % (*w/v*) NaCl, pelleted at 5,000 × g for 15 min, and stored at -20 °C. SeMet-labeled proteins were purified analogously to their unlabeled counterparts (4.2.13), except that all purification buffers contained either 5 mM β-ME or dithiothreitol (DTT).

4.2.10 Plasmid DNA Preparation

Plasmid DNA from grown *E. coli* cultures was isolated using the Wizard® Plus SV Minipreps DNA Purification System.

4.2.11 DNA Sequencing

DNA constructs, pre-selected by control digestion, were sequenced by Eurofins Genomics (Ebersberg, GER) via Sanger sequencing.[216]

4.2.12 Expression Tests

Small-scale expression tests were conducted to evaluate the influence of different expression parameters on the protein production of transformed *E. coli*. Standard cultivations were performed similar to 4.2.9.3, using 5 mL pre- and 50 mL main cultures. Variations included different media, expression temperatures, *E. coli* strains, and vector constructs. They are specifically mentioned as they occur in the results section. Samples before and after induction were analyzed via SDS-PAGE to assess expression levels.

Optionally, further protein purifications were performed using NiNTA spin columns to evaluate the solubility of the produced His_6-tagged proteins. Buffers A and B contained 100 mM Tris pH 7.5, 500 mM NaCl, 5 mM β-ME, and either 20 or 500 mM imidazole pH 7.5. After resuspension of the pellets in 1 mL buffer A, the cells were lysed by sonication (10 % amplitude, total on-time: 1 min, pulse on/off: 10 s) and centrifuged at 40,000 × g at 4 °C for 20 min. Purification was performed according to the manufacturer's instructions. Samples were collected and analyzed via SDS-PAGE.

4.2.13 Protein Purification

All proteins examined in this work were produced as a His_6-tag fusion construct and purified using a combination of IMAC and SEC (Table 12a and b). After initial separation from the lysate via IMAC, dialysis into buffer C was performed (Table 12b). If applicable, tag cleavage was performed during dialysis by adding a 1:100 molar ratio of TEV or SUMO protease. The cleaved protein, tag, and protease were separated using a reverse IMAC protocol. The general purification protocol is described below. Specific buffers and conditions used for the purification of each individual protein are listed in Tables 12a and b.

4.2.13.1 Immobilized Metal Affinity Chromatography (IMAC)

10 g of cell mass, typically obtained form 2 L of cell culture, were re-suspended in buffer A (Table 12b). 5 mg of each, PefablocSC and DNAse were added. Cells were lysed by ultrasonication (80 % amplitude, total on-time: 4 min, pulse on/off: 1 sec) in a glass rosette, cooled in an ice bath. The resulting cell lysate was centrifuged at 40,000 × g for 20 min at 4 °C. The supernatant was applied to a 5 mL HisTrap HP column equilibrated with buffer A (flow rate: 5 mL/min) using an ÄKTA Start system. Unbound or loosely associated proteins were removed by washing with 50 mL buffer A. Bound proteins were eluted by applying a 100 mL linear gradient from buffer A to buffer B (Table 12b). Protein-containing fractions were combined and dialyzed overnight at 4 °C against 2 L buffer C (Table 12b).

4.2.13.2 Reverse IMAC

The dialyzed and cleaved protein was applied to a 5 mL HisTrap HP column, which had been pre-equilibrated with buffer A (Table 12b). Cleaved proteins were eluted with buffer A. The His_6-tagged protease (if added), cleaved tags, and uncleaved proteins were retained on the column until elution with buffer B (Table 12b).

4.2.13.3 Size Exclusion Chromatography (SEC)

Protein-containing fractions from the IMAC purification were pooled and concentrated to a volume of approximately 2 mL using an Amicon Ultra Centrifugal Filter Unit. Precipitated protein was removed by centrifugation at 20,000 × g for 10 min at 4 °C. SEC was performed using a HiLoad 16/600 Superdex 75 or 200 pg column on an ÄKTA Pure or Purifier system (Table 12a). Buffer C was used for equilibration and protein separation (Table 12b). A flow rate of 1.0 mL/min was used. Protein-containing fractions were pooled and concentrated using an Amicon Ultra Centrifugal Filter Unit.

4.2.14 SDS-PAGE

4.2.14.1 Gel Preparation

For SDS-PAGE[217] separating gels, 5:6:9 (12 % gels) or 2:3:3 (15 % gels) mixtures (*v/v*) of SDS separating gel buffer, acrylamide, and water were prepared. SDS-PAGE stacking gels were prepared by mixing SDS stacking gel buffer, acrylamide, and water in a 5:1:4 ratio (*v/v*). Gel polymerizations were initiated by adding 50 µL 10 % (*w/v*) ammonium persulfate (APS) and 5 µL (*w/v*) tetramethylethylenediamine (TEMED) to 10 mL of the separating gel mixture, and 10 µL 10 % (*w/v*) APS and 1 µL (*w/v*) TEMED to 1 mL of stacking gel mixture. The separating gel mixture was added to the casting system of the Mini-PROTEAN Tetra System (s. manufacturer's instructions), polymerized for 30 min, and topped off with the stacking gel mixture. After addition of the comb, the stacking gel was polymerized for 30 min. Gels were sealed airtight and stored at 4 °C for up to two weeks.

4.2.14.2 Sample Preparation and Electrophoresis

5 µL samples were mixed 1:1 (*v/v*) with 2x Laemmli sample buffer, heated at 95 °C for 5 min, and applied to an SDS gel alongside 5 µL Unstained Protein Standard, Broad Range (10-200 kDa). Electrophoresis was performed for 30 min at 35 mA in SDS running buffer (s. manufacturer's instructions). Protein bands were stained by incubating the gel in Coomassie staining solution for 15 min followed by destaining for 2 h in 10 % (*v/v*) acetic acid.

4.2.15 Liquid Chromatography - Mass Spectrometry (LC-MS)

For LC, an Acquity UPLC Protein BEH C4 column was operated on an Acquity Premier HPLC system. The mobile phases consisted of water with 0.1 % formic acid and acetonitrile with 0.1 % formic acid. After loading the protein sample, the column was run at a flow rate of 0.4 mL/min.

For MS, a Synapt XS instrument was used. Measurements were conducted in positive ion mode with the analyzer set to resolution mode. The detection mass range was set between 400 Da and 4000 Da. A cone voltage of 70 V was applied for ionization. Accurate mass calibration was achieved by locking Glufibrinopeptide B (785.8427 Da) as a mass reference.

The MassLynx V4.2 software was used for data analysis.

4.2.16 Photometric Activity Assay

The enzymatic activities of WT and mutant MtgA were determined using a modified version of established protocols for Cbl-dependent methyltransferases.[151,180] Methyl transfer from methylcob(III)alamin to THF is accompanied by conversion of Co(III) to Co(I), which can be monitored photometrically. To avoid oxidation of the oxygen-sensitive Co(I), the assay was performed under anaerobic conditions. A sealable cuvette was closed under N_2-atmosphere.

Buffers, substances, enzymes were added via syringe injections through the septum in the lid of the cuvette. The syringes were washed with N_2 beforehand. After each addition, the mixture inside the cuvette was mixed with an N_2-washed syringe. The buffer (100 mM Tris/HCl pH 7.5) was degassed by bubbling with N_2.[218] Photometric measurements were performed in an Ultrospec 7000. At the start of the assay, 1 mL degassed buffer and 500 µM THF (in DMSO) were added to the cuvette. After subtracting the absorbance of the solution as background, 50 µM MeCbl (in DMSO) were injected.

For the quantification of the enzymatic activity, the absorption at 525 nm was measured every second for 10 min before and 2 h after the addition of 1 µM WT or mutant MtgA (in 20 mM Tris/HCl pH 7.5, 100 mM NaCl). To avoid oxygen contamination, stock concentrations of non-degassed THF, MeCbl, and protein solutions were adjusted to achieve the final reaction mixture by adding 1 µL each. To assess the reduction in enzymatic activity of the R236A mutant compared to WT MtgA, an unpaired, two-tailed t-test was performed and evaluated using GraphPad Prism. To visualize the activity of WT MtgA, two spectra were recorded from 400 nm to 800 nm immediately and 2 h after the addition of 1 µM WT MtgA.

4.2.17 Crystallization

4.2.17.1 Broad Screens

Crystallization trials were performed using the sitting drop vapor diffusion method. The reservoirs of Intelli-Plates® 96-3-Well (96 reservoirs with three sitting drop spots each) were filled with 50 µL reservoir solutions (Table 7) using the Phoenix LHS or Gryphon LCP pipetting robots. Droplets consisting of 0.2 + 0.2 µL, 0.2 + 0.1 µL, and 0.3 + 0.1 µL of protein and reservoir solution, respectively, were prepared by using the Phoenix LHS, Gryphon LCP, or Oryx4 pipetting robots. Plates were sealed and incubated at 20 °C. After incubation for at least one day, the plates were inspected for crystals using an SZX10/KL transmission microscope.

4.2.17.2 Fine Screens

Successful broad screening was often followed up by fine screening specific conditions to improve crystal size and quality. Fine screens were designed using the Hamilton Screen Designer software. The Intelli-Plates® 96-3-Well reservoirs were filled with 50 µL of the designed fine screen conditions using the Microlab Starlet pipetting robot. Droplet preparation, sealing, and incubation were performed analogously to the broad screen crystallization protocol (4.2.17.1).

4.2.17.3 Crystal Handling

The conditions that produced the crystals that ultimately led to the determined structures presented in this work are listed in Table S1. In preparation for data acquisition,

crystals were cryoprotected with a 7:3-mixture of mother liquor and 100 % (*v/v*) glycerol prior to vitrification in liquid nitrogen.

4.2.18 Data Collection and Structure Determination

High resolution native datasets of the protein crystals were recorded using synchrotron radiation of λ = 1.0 Å at the beamline X06SA of the Swiss Light Source (SLS, Paul Scherrer Institute, Villigen, CH). The occurrence of anomalous diffraction of crystals of zinc-containing MtbA as well as selenium-derivatized MtgA was verified via fluorescence scans. Anomalous datasets of MtbA and MtgA were recorded at 1.284 Å (Zn) and 0.979 Å (Se), respectively. The obtained reflection intensities were evaluated using the program package XDS.[212] Data reductions were performed using XSCALE (Tables S2 - S7).[212]

Considering the molecular weight and the oligomeric state of the proteins, as well as the crystal parameters (Tables S2 - S7), the number of protein chains within the asymmetric unit as well as the solvent content were estimated based on the calculation of the Matthews coefficient.[219] Phasing was accomplished either by experimental methods (MtbA and MtgA) or by molecular replacement (MtgB and MtcB). For MtbA, experimental phases were obtained by single anomalous dispersion (SAD)[220] methods using the peak absorption wavelength of the zinc ion bound to the active site of the enzyme (λ = 1.284 Å, f' = -6.8; f'' = 4.6). CRANK2 was used for the automatic X-ray structure solution of the anomalous dataset with 1.5 Å resolution (Table S5).[194] For MtgA, experimental phases were obtained by SAD methods using the peak absorption wavelength of a selenium derivatized crystal (λ = 0.979 Å, f' = -6.7; f'' = 5.5). Using a 1.9 Å resolution dataset, 10 heavy atom sites were located with SHELXD (Table S6).[210] Subsequent SHARP-SAD-phasing[209] and solvent flattening with the program DM[195] resulted in an electron density map with phases at 2.8 Å resolution. The quality was sufficient to model secondary structure elements by polyalanine residues. For MtgB and MtcB, the deposited apo structure of MtgB (PDB-ID: 2QNE) was used as a search model to generate electron density maps with phase information by Patterson search calculations.[203]

The initial models were transferred to the other datasets of the respective proteins by applying rigid body- and positional-refinement using REFMAC5.[207,208] Subsequent datasets of WT and mutant MtgA were refined using the high resolution initial MtgA coordinates as starting model. The models were completed in iterative rounds of restrained refinements in REFMAC5 and model building with the 3D graphics program Coot.[193] After model building was completed, water molecules were automatically placed with ARP/wARP solvent.[189] Temperature factors were refined anisotropically with restraints between bonded atoms using translation/liberation/screw motion parameters. The resulting X-ray structures were validated using the online tool MolProbity.[200] All statistics are summarized in Tables S2 through S7.

5 References

1. Schlesinger, W.H. *Biogeochemistry : an analysis of global change*, xiii, 588 pages : illustrations (Academic Press, San Diego, Calif, 1997).
2. DiMarco, A.A., Bobik, T.A. & Wolfe, R.S. Unusual coenzymes of methanogenesis. *Annu Rev Biochem* **59**, 355-94 (1990).
3. Whitman, W.B., Coleman, D.C. & Wiebe, W.J. Prokaryotes: the unseen majority. *Proc Natl Acad Sci U S A* **95**, 6578-83 (1998).
4. Deppenmeier, U. The unique biochemistry of methanogenesis. *Prog Nucleic Acid Res Mol Biol* **71**, 223-83 (2002).
5. Ferry, J.G. Fundamentals of methanogenic pathways that are key to the biomethanation of complex biomass. *Curr Opin Biotechnol* **22**, 351-7 (2011).
6. Zinder, S.H. Physiological Ecology of Methanogens. in *Methanogenesis: Ecology, Physiology, Biochemistry & Genetics* (ed. Ferry, J.G.) 128-206 (Springer US, Boston, MA, 1993).
7. Ellermann, J., Hedderich, R., Bocher, R. & Thauer, R.K. The final step in methane formation. Investigations with highly purified methyl-CoM reductase (component C) from Methanobacterium thermoautotrophicum (strain Marburg). *Eur J Biochem* **172**, 669-77 (1988).
8. Deppenmeier, U., Blaut, M., Mahlmann, A. & Gottschalk, G. Reduced coenzyme F420: heterodisulfide oxidoreductase, a proton- translocating redox system in methanogenic bacteria. *Proceedings of the National Academy of Sciences* **87**, 9449-9453 (1990).
9. Galagan, J.E. et al. The genome of M. acetivorans reveals extensive metabolic and physiological diversity. *Genome Res* **12**, 532-42 (2002).
10. Ferry, J.G. Biochemistry of methanogenesis. *Crit Rev Biochem Mol Biol* **27**, 473-503 (1992).
11. van der Meijden, P., Jansen, L.P.J.M., van der Drift, C. & Vogels, G.D. Involvement of corrinoids in the methylation of coenzyme M (2-mercaptoethanesulfonic acid) by methanol and enzymes from Methanosarcina barkeri. *FEMS Microbiology Letters* **19**, 247-251 (1983).
12. McBride, B.C. & Wolfe, R.S. A new coenzyme of methyl transfer, coenzyme M. *Biochemistry* **10**, 2317-24 (1971).
13. Shapiro, S. & Wolfe, R.S. Methyl-coenzyme M, an intermediate in methanogenic dissimilation of C1 compounds by Methanosarcina barkeri. *J Bacteriol* **141**, 728-34 (1980).
14. Thauer, R.K. Biochemistry of methanogenesis: a tribute to Marjory Stephenson. 1998 Marjory Stephenson Prize Lecture. *Microbiology (Reading)* **144 (Pt 9)**, 2377-2406 (1998).
15. Balch, W.E. & Wolfe, R.S. Specificity and biological distribution of coenzyme M (2-mercaptoethanesulfonic acid). *J Bacteriol* **137**, 256-63 (1979).
16. Jocelyn, P.C. *Biochemistry of the SH group the occurrence, chemical properties, metabolism and biological function of thiols and disulphides*, XX, 404 S. (Acad. Press, London u.a., 1972).

17. Ferry, J.G. Enzymology of one-carbon metabolism in methanogenic pathways. *FEMS Microbiol Rev* **23**, 13-38 (1999).

18. Thauer, R.K., Kaster, A.-K., Seedorf, H., Buckel, W. & Hedderich, R. Methanogenic archaea: ecologically relevant differences in energy conservation. *Nature Reviews Microbiology* **6**, 579-591 (2008).

19. Hartzell, P.L. & Wolfe, R.S. Requirement of the nickel tetrapyrrole F430 for in vitro methanogenesis: reconstitution of methylreductase component C from its dissociated subunits. *Proc Natl Acad Sci U S A* **83**, 6726-30 (1986).

20. Rouviere, P.E., Bobik, T.A. & Wolfe, R.S. Reductive activation of the methyl coenzyme M methylreductase system of Methanobacterium thermoautotrophicum delta H. *J Bacteriol* **170**, 3946-52 (1988).

21. Noll, K.M. & Wolfe, R.S. The role of 7-mercaptoheptanoylthreonine phosphate in the methylcoenzyme M methylreductase system from Methanobacterium thermoautotrophicum. *Biochem Biophys Res Commun* **145**, 204-10 (1987).

22. Yamazaki, S. et al. Stereochemical studies of a selenium-containing hydrogenase from Methanococcus vannielii: determination of the absolute configuration of C-5 chirally labeled dihydro-8-hydroxy-5-deazaflavin cofactor. *Proc Natl Acad Sci U S A* **82**, 1364-6 (1985).

23. Yamazaki, S., Tsai, L. & Stadtman, T.C. Analogues of 8-hydroxy-5-deazaflavin cofactor: relative activity as substrates for 8-hydroxy-5-deazaflavin-dependent NADP+ reductase from Methanococcus vannielii. *Biochemistry* **21**, 934-9 (1982).

24. Yamazaki, S., Tsai, L., Stadtman, T.C., Jacobson, F.S. & Walsh, C. Stereochemical studies of 8-hydroxy-5-deazaflavin-dependent NADP+ reductase from Methanococcus vannielii. *J Biol Chem* **255**, 9025-7 (1980).

25. Krzycki, J. & Zeikus, J.G. Quantification of corrinoids in methanogenic bacteria. *Current Microbiology* **3**, 243-245 (1980).

26. Banerjee, R. & Ragsdale, S.W. The many faces of vitamin B12: catalysis by cobalamin-dependent enzymes. *Annu Rev Biochem* **72**, 209-47 (2003).

27. Ragsdale, S.W. Catalysis of methyl group transfers involving tetrahydrofolate and B(12). *Vitam Horm* **79**, 293-324 (2008).

28. Eikmanns, B. & Thauer, R.K. Catalysis of an isotopic exchange between CO2 and the carboxyl group of acetate by Methanosarcina barkeri grown on acetate. *Archives of Microbiology* **138**, 365-370 (1984).

29. Poirot, C.M. et al. Formation of methylcoenzyme M from formaldehyde by cell-free extracts of Methanobacterium thermoautotrophicum. Evidence for the involvement of a corrinoid-containing methyltransferase. *FEMS Microbiology Letters* **40**, 7-13 (1987).

30. Matthews, R.G., Koutmos, M. & Datta, S. Cobalamin-dependent and cobamide-dependent methyltransferases. *Curr Opin Struct Biol* **18**, 658-66 (2008).

31. Weimer, P.J. & Zeikus, J.G. One carbon metabolism in methanogenic bacteria. *Archives of Microbiology* **119**, 49-57 (1978).

32. Burke, S.A. & Krzycki, J.A. Reconstitution of Monomethylamine:Coenzyme M methyl transfer with a corrinoid protein and two methyltransferases purified from Methanosarcina barkeri. *J Biol Chem* **272**, 16570-7 (1997).

33. Hippe, H., Caspari, D., Fiebig, K. & Gottschalk, G. Utilization of trimethylamine and other N-methyl compounds for growth and methane formation by Methanosarcina barkeri. *Proc Natl Acad Sci U S A* **76**, 494-8 (1979).

34. Ferguson, D.J., Jr., Gorlatova, N., Grahame, D.A. & Krzycki, J.A. Reconstitution of dimethylamine:coenzyme M methyl transfer with a discrete corrinoid protein and two methyltransferases purified from Methanosarcina barkeri. *J Biol Chem* **275**, 29053-60 (2000).

35. Naumann, E., Fahlbusch, K. & Gottschalk, G. Presence of a Trimethylamine - Hs-Coenzyme-M Methyltransferase in Methanosarcina-Barkeri. *Archives of Microbiology* **138**, 79-83 (1984).

36. Ferguson, D.J., Jr. & Krzycki, J.A. Reconstitution of trimethylamine-dependent coenzyme M methylation with the trimethylamine corrinoid protein and the isozymes of methyltransferase II from Methanosarcina barkeri. *J Bacteriol* **179**, 846-52 (1997).

37. van der Meijden, P. et al. Methyltransferases involved in methanol conversion by Methanosarcina barkeri. *Arch Microbiol* **134**, 238-42 (1983).

38. Zinder, S.H. & Brock, T.D. Methane, carbon dioxide, and hydrogen sulfide production from the terminal methiol group of methionine by anaerobic lake sediments. *Appl Environ Microbiol* **35**, 344-52 (1978).

39. Kiene, R.P., Oremland, R.S., Catena, A., Miller, L.G. & Capone, D.G. Metabolism of reduced methylated sulfur compounds in anaerobic sediments and by a pure culture of an estuarine methanogen. *Appl Environ Microbiol* **52**, 1037-45 (1986).

40. Oremland, R.S., Kiene, R.P., Mathrani, I., Whiticar, M.J. & Boone, D.R. Description of an estuarine methylotrophic methanogen which grows on dimethyl sulfide. *Appl Environ Microbiol* **55**, 994-1002 (1989).

41. Finster, K., Tanimoto, Y. & Bak, F. Fermentation of Methanethiol and Dimethylsulfide by a Newly Isolated Methanogenic Bacterium. *Archives of Microbiology* **157**, 425-430 (1992).

42. van der Maarel, M., Jansen, M. & Hansen, T.A. Methanogenic conversion of 3-s-methylmercaptopropionate to 3-mercaptopropionate. *Appl Environ Microbiol* **61**, 48-51 (1995).

43. Tallant, T.C. & Krzycki, J.A. Methylthiol:coenzyme M methyltransferase from Methanosarcina barkeri, an enzyme of methanogenesis from dimethylsulfide and methylmercaptopropionate. *J Bacteriol* **179**, 6902-11 (1997).

44. Boone, D.R., Whitman, W.B. & Rouvière, P. Diversity and Taxonomy of Methanogens. in *Methanogenesis: Ecology, Physiology, Biochemistry & Genetics* (ed. Ferry, J.G.) 35-80 (Springer US, Boston, MA, 1993).

45. Burke, S.A., Lo, S.L. & Krzycki, J.A. Clustered genes encoding the methyltransferases of methanogenesis from monomethylamine. *J Bacteriol* **180**, 3432-40 (1998).

46. Weiss, R.B. Ribosomal frameshifting, jumping and readthrough. *Curr Opin Cell Biol* **3**, 1051-5 (1991).

47. Farabaugh, P.J. Programmed translational frameshifting. *Microbiol Rev* **60**, 103-34 (1996).

48. Fox, T.D. Natural variation in the genetic code. *Annu Rev Genet* **21**, 67-91 (1987).

49. Zinoni, F., Birkmann, A., Stadtman, T.C. & Böck, A. Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase

(formate-hydrogen-lyase-linked) from Escherichia coli. *Proc Natl Acad Sci U S A* **83**, 4650-4 (1986).

50. Zinoni, F., Birkmann, A., Leinfelder, W. & Bock, A. Cotranslational insertion of selenocysteine into formate dehydrogenase from Escherichia coli directed by a UGA codon. *Proc Natl Acad Sci U S A* **84**, 3156-60 (1987).
51. Bock, A. et al. Selenocysteine: the 21st amino acid. *Mol Microbiol* **5**, 515-20 (1991).
52. Paul, L., Ferguson, D.J., Jr. & Krzycki, J.A. The trimethylamine methyltransferase gene and multiple dimethylamine methyltransferase genes of Methanosarcina barkeri contain in-frame and read-through amber codons. *J Bacteriol* **182**, 2520-9 (2000).
53. James, C.M., Ferguson, T.K., Leykam, J.F. & Krzycki, J.A. The amber codon in the gene encoding the monomethylamine methyltransferase isolated from Methanosarcina barkeri is translated as a sense codon. *J Biol Chem* **276**, 34252-8 (2001).
54. Hao, B. et al. A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science* **296**, 1462-6 (2002).
55. Hao, B. et al. Reactivity and chemical synthesis of L-pyrrolysine- the 22(nd) genetically encoded amino acid. *Chem Biol* **11**, 1317-24 (2004).
56. Soares, J.A. et al. The residue mass of L-pyrrolysine in three distinct methylamine methyltransferases. *J Biol Chem* **280**, 36962-9 (2005).
57. Atkins, J.F. & Gesteland, R. Biochemistry. The 22nd amino acid. *Science* **296**, 1409-10 (2002).
58. Polycarpo, C. et al. Activation of the Pyrrolysine Suppressor tRNA Requires Formation of a Ternary Complex with Class I and Class II Lysyl-tRNA Synthetases. *Molecular Cell* **12**, 287-294 (2003).
59. Srinivasan, G., James, C.M. & Krzycki, J.A. Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* **296**, 1459-62 (2002).
60. Blight, S.K. et al. Direct charging of tRNA(CUA) with pyrrolysine in vitro and in vivo. *Nature* **431**, 333-5 (2004).
61. Leinfelder, W., Zehelein, E., Mandrand-Berthelot, M.A. & Böck, A. Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. *Nature* **331**, 723-5 (1988).
62. Stadtman, T.C. Selenocysteine. *Annu Rev Biochem* **65**, 83-100 (1996).
63. Commans, S. & Böck, A. Selenocysteine inserting tRNAs: an overview. *FEMS Microbiol Rev* **23**, 335-51 (1999).
64. Krzycki, J.A. Function of genetically encoded pyrrolysine in corrinoid-dependent methylamine methyltransferases. *Curr Opin Chem Biol* **8**, 484-91 (2004).
65. Zhang, Y., Baranov, P.V., Atkins, J.F. & Gladyshev, V.N. Pyrrolysine and selenocysteine use dissimilar decoding strategies. *J Biol Chem* **280**, 20740-51 (2005).
66. Longstaff, D.G. et al. A natural genetic code expansion cassette enables transmissible biosynthesis and genetic encoding of pyrrolysine. *Proc Natl Acad Sci U S A* **104**, 1021-6 (2007).
67. Gaston, M.A., Zhang, L., Green-Church, K.B. & Krzycki, J.A. The complete biosynthesis of the genetically encoded amino acid pyrrolysine from lysine. *Nature* **471**, 647-650 (2011).

68. Cellitti, S.E. et al. D-Ornithine coopts pyrrolysine biosynthesis to make and insert pyrroline-carboxy-lysine. *Nat Chem Biol* **7**, 528-30 (2011).

69. Quitterer, F., List, A., Eisenreich, W., Bacher, A. & Groll, M. Crystal structure of methylornithine synthase (PylB): insights into the pyrrolysine biosynthesis. *Angew Chem Int Ed Engl* **51**, 1339-42 (2012).

70. Quitterer, F., List, A., Beck, P., Bacher, A. & Groll, M. Biosynthesis of the 22nd genetically encoded amino acid pyrrolysine: structure and reaction mechanism of PylC at 1.5A resolution. *J Mol Biol* **424**, 270-82 (2012).

71. Quitterer, F., Beck, P., Bacher, A. & Groll, M. Structure and Reaction Mechanism of Pyrrolysine Synthase (PylD). *Angewandte Chemie International Edition* **52**, 7033-7037 (2013).

72. Krzycki, J.A. The path of lysine to pyrrolysine. *Curr Opin Chem Biol* **17**, 619-25 (2013).

73. Deppenmeier, U. et al. The genome of Methanosarcina mazei: evidence for lateral gene transfer between bacteria and archaea. *J Mol Microbiol Biotechnol* **4**, 453-61 (2002).

74. Woese, C. The universal ancestor. *Proceedings of the National Academy of Sciences* **95**, 6854-6859 (1998).

75. Nelson, K.E. et al. Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of Thermotoga maritima. *Nature* **399**, 323-329 (1999).

76. Ng, W.V. et al. Genome sequence of Halobacterium species NRC-1. *Proceedings of the National Academy of Sciences* **97**, 12176-12181 (2000).

77. Gaston, M.A., Jiang, R. & Krzycki, J.A. Functional context, biosynthesis, and genetic encoding of pyrrolysine. *Curr Opin Microbiol* **14**, 342-9 (2011).

78. Heinemann, I.U. et al. The appearance of pyrrolysine in tRNAHis guanylyltransferase by neutral evolution. *Proceedings of the National Academy of Sciences* **106**, 21103-21108 (2009).

79. Krzycki, J.A. Translation of UAG as Pyrrolysine. in *Recoding: Expansion of Decoding Rules Enriches Gene Expression* (eds. Atkins, J.F. & Gesteland, R.F.) 53-77 (Springer New York, New York, NY, 2010).

80. Ticak, T., Kountz, D.J., Girosky, K.E., Krzycki, J.A. & Ferguson, D.J., Jr. A nonpyrrolysine member of the widely distributed trimethylamine methyltransferase family is a glycine betaine methyltransferase. *Proc Natl Acad Sci U S A* **111**, E4668-76 (2014).

81. Picking, J.W., Behrman, E.J., Zhang, L. & Krzycki, J.A. MtpB, a member of the MttB superfamily from the human intestinal acetogen Eubacterium limosum, catalyzes proline betaine demethylation. *J Biol Chem* **294**, 13697-13707 (2019).

82. Kountz, D.J., Behrman, E.J., Zhang, L. & Krzycki, J.A. MtcB, a member of the MttB superfamily from the human gut acetogen Eubacterium limosum, is a cobalamin-dependent carnitine demethylase. *J Biol Chem* **295**, 11971-11981 (2020).

83. Ellenbogen, J.B., Jiang, R., Kountz, D.J., Zhang, L. & Krzycki, J.A. The MttB superfamily member MtyB from the human gut symbiont Eubacterium limosum is a cobalamin-dependent gamma-butyrobetaine methyltransferase. *J Biol Chem* **297**, 101327 (2021).

84. Li, J. et al. Insights into pyrrolysine function from structures of a trimethylamine methyltransferase and its corrinoid protein complex. *Commun Biol* **6**, 54 (2023).

85. Struck, A.-W., Thompson, M.L., Wong, L.S. & Micklefield, J. S-Adenosyl-Methionine-Dependent Methyltransferases: Highly Versatile Enzymes in Biocatalysis, Biosynthesis and Other Biotechnological Applications. *ChemBioChem* **13**, 2642-2655 (2012).

86. Maden, B.E. Tetrahydrofolate and tetrahydromethanopterin compared: functionally distinct carriers in C1 metabolism. *Biochem J* **350 Pt 3**, 609-29 (2000).

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

88. Hodgkin, D.C. et al. Structure of vitamin B12. *Nature* **178**, 64-6 (1956).

89. Ragsdale, S.W., Lindahl, P.A. & Münck, E. Mössbauer, EPR, and optical studies of the corrinoid/iron-sulfur protein involved in the synthesis of acetyl coenzyme A by Clostridium thermoaceticum. *Journal of Biological Chemistry* **262**, 14289-14297 (1987).

90. Barker, H.A., Weissbach, H. & Smyth, R.D. A COENZYME CONTAINING PSEUDOVITAMIN B(12). *Proc Natl Acad Sci U S A* **44**, 1093-7 (1958).

91. Guest, J.R., Friedman, S., Woods, D.D. & Smith, E.L. A methyl analogue of cobamide coenzyme in relation to methionine synthesis by bacteria. *Nature* **195**, 340-2 (1962).

92. Banerjee, R.V. & Matthews, R.G. Cobalamin-dependent methionine synthase. *The FASEB Journal* **4**, 1450-1459 (1990).

93. Roth, J.R., Lawrence, J.G. & Bobik, T.A. Cobalamin (coenzyme B12): synthesis and biological significance. *Annu Rev Microbiol* **50**, 137-81 (1996).

94. Barker, H. & Kamen, M.D. Carbon dioxide utilization in the synthesis of acetic acid by Clostridium thermoaceticum. *Proceedings of the National Academy of Sciences* **31**, 219-225 (1945).

95. Kaufmann, F., Wohlfarth, G. & Diekert, G. O-demethylase from Acetobacterium dehalogenans--cloning, sequencing, and active expression of the gene encoding the corrinoid protein. *Eur J Biochem* **257**, 515-21 (1998).

96. Vannelli, T., Messmer, M., Studer, A., Vuilleumier, S. & Leisinger, T. A corrinoid-dependent catabolic pathway for growth of a Methylobacterium strain with chloromethane. *Proc Natl Acad Sci U S A* **96**, 4615-20 (1999).

97. Bandarian, V. et al. Domain alternation switches B(12)-dependent methionine synthase to the activation conformation. *Nat Struct Biol* **9**, 53-6 (2002).

98. Evans, J.C. et al. Structures of the N-terminal modules imply large domain motions during catalysis by methionine synthase. *Proc Natl Acad Sci U S A* **101**, 3729-36 (2004).

99. Watkins, M.B., Wang, H., Burnim, A. & Ando, N. Conformational switching and flexibility in cobalamin-dependent methionine synthase studied by small-angle X-ray scattering and cryoelectron microscopy. *Proc Natl Acad Sci U S A* **120**, e2302531120 (2023).

100. Doukov, T.I., Hemmi, H., Drennan, C.L. & Ragsdale, S.W. Structural and kinetic evidence for an extended hydrogen-bonding network in catalysis of methyl group transfer. Role of an active site asparagine residue in activation of methyl transfer by methyltransferases. *J Biol Chem* **282**, 6609-6618 (2007).

101. Goetzl, S., Jeoung, J.H., Hennig, S.E. & Dobbek, H. Structural basis for electron and methyl-group transfer in a methyltransferase system operating in the reductive acetyl-CoA pathway. *J Mol Biol* **411**, 96-109 (2011).

102. Kung, Y. et al. Visualizing molecular juggling within a B12-dependent methyltransferase complex. *Nature* **484**, 265-9 (2012).

103. Bandarian, V., Ludwig, M.L. & Matthews, R.G. Factors modulating conformational equilibria in large modular proteins: a case study with cobalamin-dependent methionine synthase. *Proc Natl Acad Sci U S A* **100**, 8156-63 (2003).

104. Ludwig, M.L. & Matthews, R.G. Structure-based perspectives on B12-dependent enzymes. *Annu Rev Biochem* **66**, 269-313 (1997).

105. Drennan, C.L., Matthews, R.G. & Ludwig, M.L. Cobalamin-dependent methionine synthase: the structure of a methylcobalamin-binding fragment and implications for other B12-dependent enzymes. *Curr Opin Struct Biol* **4**, 919-29 (1994).

106. Matthews, R.G. Cobalamin-dependent methyltransferases. *Acc Chem Res* **34**, 681-9 (2001).

107. Tallant, T.C., Paul, L. & Krzycki, J.A. The MtsA subunit of the methylthiol:coenzyme M methyltransferase of Methanosarcina barkeri catalyses both half-reactions of corrinoid-dependent dimethylsulfide: coenzyme M methyl transfer. *J Biol Chem* **276**, 4485-93 (2001).

108. Hagemeier, C.H., Krer, M., Thauer, R.K., Warkentin, E. & Ermler, U. Insight into the mechanism of biological methanol activation based on the crystal structure of the methanol-cobalamin methyltransferase complex. *Proc Natl Acad Sci U S A* **103**, 18917-22 (2006).

109. Gottschalk, G. & Thauer, R.K. The Na(+)-translocating methyltransferase complex from methanogenic archaea. *Biochim Biophys Acta* **1505**, 28-36 (2001).

110. Das, A. et al. Characterization of a corrinoid protein involved in the C1 metabolism of strict anaerobic bacterium Moorella thermoacetica. *Proteins* **67**, 167-76 (2007).

111. Naidu, D. & Ragsdale, S.W. Characterization of a three-component vanillate O-demethylase from Moorella thermoacetica. *J Bacteriol* **183**, 3276-81 (2001).

112. Studer, A., Stupperich, E., Vuilleumier, S. & Leisinger, T. Chloromethane: tetrahydrofolate methyl transfer by two proteins from Methylobacterium chloromethanicum strain CM4. *Eur J Biochem* **268**, 2931-8 (2001).

113. Brabham, R. & Fascione, M.A. Pyrrolysine Amber Stop-Codon Suppression: Development and Applications. *Chembiochem* **18**, 1973-1983 (2017).

114. Pedelacq, J.-D. & Cabantous, S. Development and Applications of Superfolder and Split Fluorescent Protein Detection Systems in Biology. *International Journal of Molecular Sciences* **20**, 3479 (2019).

115. Cody, C.W., Prasher, D.C., Westler, W.M., Prendergast, F.G. & Ward, W.W. Chemical structure of the hexapeptide chromophore of the Aequorea green-fluorescent protein. *Biochemistry* **32**, 1212-8 (1993).

116. Wan, W., Tharp, J.M. & Liu, W.R. Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. *Biochim Biophys Acta* **1844**, 1059-70 (2014).

117. Wals, K. & Ovaa, H. Unnatural amino acid incorporation in E. coli: current and future applications in the design of therapeutic proteins. *Front Chem* **2**, 15 (2014).

118. Corless Elliot, I., Mettert Erin, L., Kiley Patricia, J. & Antony, E. Elevated Expression of a Functional Suf Pathway in Escherichia coli BL21(DE3) Enhances Recombinant Production of an Iron-Sulfur Cluster-Containing Protein. *Journal of Bacteriology* **202**, 10.1128/jb.00496-19 (2020).

119. Quitterer, F., Beck, P., Bacher, A. & Groll, M. The formation of pyrroline and tetrahydropyridine rings in amino acids catalyzed by pyrrolysine synthase (PylD). *Angew Chem Int Ed Engl* **53**, 8150-3 (2014).

120. Peroutka III, R.J., Orcutt, S.J., Strickler, J.E. & Butt, T.R. SUMO Fusion Technology for Enhanced Protein Expression and Purification in Prokaryotes and Eukaryotes. in *Heterologous Gene Expression in E.coli: Methods and Protocols* (eds. Evans, J.T.C. & Xu, M.-Q.) 15-30 (Humana Press, Totowa, NJ, 2011).

121. Rupp, B. & Kantardjieff, K. *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*, (Garland Science, 2010).

122. Cohen, S.N., Chang, A.C., Boyer, H.W. & Helling, R.B. Construction of biologically functional bacterial plasmids in vitro. *Proceedings of the National Academy of Sciences* **70**, 3240-3244 (1973).

123. Gibson, D.G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343-5 (2009).

124. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nature Methods* **19**, 679-682 (2022).

125. Banner, D.W. et al. Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 angstrom resolution using amino acid sequence data. *Nature* **255**, 609-14 (1975).

126. Orengo, C.A. et al. CATH--a hierarchic classification of protein domain structures. *Structure* **5**, 1093-108 (1997).

127. Sillitoe, I. et al. CATH: increased structural coverage of functional space. *Nucleic Acids Res* **49**, D266-D273 (2021).

128. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-97 (2007).

129. Holm, L., Laiho, A., Toronen, P. & Salgado, M. DALI shines a light on remote homologs: One hundred discoveries. *Protein Sci* **32**, e4519 (2023).

130. Schiefner, A. et al. Cation-π interactions as determinants for binding of the compatible solutes glycine betaine and proline betaine by the periplasmic ligand-binding protein ProX from Escherichia coli. *Journal of Biological Chemistry* **279**, 5588-5596 (2004).

131. Horn, C. et al. Molecular determinants for substrate specificity of the ligand-binding protein OpuAC from Bacillus subtilis for the compatible solutes glycine betaine and proline betaine. *Journal of molecular biology* **357**, 592-606 (2006).

132. Wolters, J.C. et al. Ligand binding and crystal structures of the substrate-binding domain of the ABC transporter OpuA. *PLoS One* **5**, e10361 (2010).

133. Salvi, F., Wang, Y.-F., Weber, I.T. & Gadda, G. Structure of choline oxidase in complex with the reaction product glycine betaine. *Acta Crystallographica Section D: Biological Crystallography* **70**, 405-413 (2014).

134. Li, C.-Y. et al. Mechanistic insight into trimethylamine N-oxide recognition by the marine bacterium Ruegeria pomeroyi DSS-3. *Journal of Bacteriology* **197**, 3378-3387 (2015).

135. Creighbaum, A.J., Ticak, T., Shinde, S., Wang, X. & Ferguson, D.J. Examination of the Glycine Betaine-Dependent Methylotrophic Methanogenesis Pathway: Insights Into Anaerobic Quaternary Amine Methylotrophy. *Frontiers in Microbiology* **10**(2019).

136. Nonaka, H. et al. Complete genome sequence of the dehalorespiring bacterium Desulfitobacterium hafniense Y51 and comparison with Dehalococcoides ethenogenes 195. *J Bacteriol* **188**, 2262-74 (2006).

137. Kim, S.H. et al. Genome sequence of Desulfitobacterium hafniense DCB-2, a Gram-positive anaerobe capable of dehalogenation and metal reduction. *BMC Microbiol* **12**, 21 (2012).

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

139. Svetlitchnaia, T., Svetlitchnyi, V., Meyer, O. & Dobbek, H. Structural insights into methyltransfer reactions of a corrinoid iron-sulfur protein involved in acetyl-CoA synthesis. *Proc Natl Acad Sci U S A* **103**, 14331-6 (2006).

140. Datta, S., Koutmos, M., Pattridge, K.A., Ludwig, M.L. & Matthews, R.G. A disulfide-stabilized conformer of methionine synthase reveals an unexpected role for the histidine ligand of the cobalamin cofactor. *Proc Natl Acad Sci U S A* **105**, 4115-20 (2008).

141. Koutmos, M., Datta, S., Pattridge, K.A., Smith, J.L. & Matthews, R.G. Insights into the reactivation of cobalamin-dependent methionine synthase. *Proc Natl Acad Sci U S A* **106**, 18527-32 (2009).

142. Sjuts, H., Dunstan, M.S., Fisher, K. & Leys, D. Structure of the cobalamin-binding protein of a putative O-demethylase from Desulfitobacterium hafniense DCB-2. *Acta Crystallogr D Biol Crystallogr* **69**, 1609-16 (2013).

143. Hennig, S.E. et al. ATP-induced electron transfer by redox-selective partner recognition. *Nat Commun* **5**, 4626 (2014).

144. Wagner, T., Ermler, U. & Shima, S. MtrA of the sodium ion pumping methyltransferase binds cobalamin in a unique mode. *Sci Rep* **6**, 28226 (2016).

145. Mendoza, J., Purchal, M., Yamada, K. & Koutmos, M. Structure of full-length cobalamin-dependent methionine synthase and cofactor loading captured in crystallo. *Nat Commun* **14**, 6365 (2023).

146. Grahame, D.A. Different isozymes of methylcobalamin:2-mercaptoethanesulfonate methyltransferase predominate in methanol- versus acetate-grown Methanosarcina barkeri. *J Biol Chem* **264**, 12890-4 (1989).

147. Sauer, K., Harms, U. & Thauer, R.K. Methanol:coenzyme M methyltransferase from Methanosarcina barkeri. Purification, properties and encoding genes of the corrinoid protein MT1. *Eur J Biochem* **243**, 670-7 (1997).

148. Sauer, K. & Thauer, R.K. Methanol:coenzyme M methyltransferase from Methanosarcina barkeri. Zinc dependence and thermodynamics of the methanol:cob(I)alamin methyltransferase reaction. *Eur J Biochem* **249**, 280-5 (1997).

149. Sauer, K. & Thauer, R.K. Methanol:coenzyme M methyltransferase from Methanosarcina barkeri--identification of the active-site histidine in the corrinoid-

harboring subunit MtaC by site-directed mutagenesis. *Eur J Biochem* **253**, 698-705 (1998).

150. Sauer, K. & Thauer, R.K. Methanol:coenzyme M methyltransferase from Methanosarcina barkeri -- substitution of the corrinoid harbouring subunit MtaC by free cob(I)alamin. *Eur J Biochem* **261**, 674-81 (1999).

151. Goulding, C.W., Postigo, D. & Matthews, R.G. Cobalamin-dependent methionine synthase is a modular protein with distinct regions for binding homocysteine, methyltetrahydrofolate, cobalamin, and adenosylmethionine. *Biochemistry* **36**, 8082-91 (1997).

152. Dorweiler, J.S., Finke, R.G. & Matthews, R.G. Cobalamin-dependent methionine synthase: probing the role of the axial base in catalysis of methyl transfer between methyltetrahydrofolate and exogenous cob(I)alamin or cob(I)inamide. *Biochemistry* **42**, 14653-62 (2003).

153. Stich, T.A. et al. Spectroscopic studies of the corrinoid/iron-sulfur protein from Moorella thermoacetica. *J Am Chem Soc* **128**, 5010-20 (2006).

154. Wassenaar, R.W., Keltjens, J.T., Van Der Drift, C. & Vogels, G.D. Purification and characterization of dimethylamine : 5-hydroxybenzimidazolyl-cobamide methyltransferase from Methanosarcina barkeri Fusaro. *European Journal of Biochemistry* **253**, 692-697 (1998).

155. Schnyder, A., Darbre, T. & Keese, R. Methyl Transfer from Methanol to Co-cobyrinate: A model for the Coenzyme B(12) Dependent Methyltransferase? *Angew Chem Int Ed Engl* **37**, 1283-1285 (1998).

156. Retey, J. Discovery and role of methylidene imidazolone, a highly electrophilic prosthetic group. *Biochim Biophys Acta* **1647**, 179-84 (2003).

157. Hilhorst, E., Iskander, A.S., Chen, T.B. & Pandit, U.K. Alkyl transfer from quaternary ammonium salts to cobalt (I): model for the cobalamin-dependent methionine synthase reaction. *Tetrahedron* **50**, 8863-8870 (1994).

158. Zheng, D., Darbre, T. & Keese, R. Methanol and dimethylaniline as methylating agents of Co (I) corrinoids under acidic conditions. *Journal of inorganic biochemistry* **73**, 273-275 (1999).

159. Wedemeyer-Exl, C., Darbre, T. & Keese, R. A model for the cobalamin-dependent methionine synthase. *Helvetica chimica acta* **82**, 1173-1184 (1999).

160. Van Doorslaer, S. et al. Axial solvent coordination in “base-off” cob (II) alamin and related Co (II)-corrinates revealed by 2D-EPR. *Journal of the American Chemical Society* **125**, 5915-5927 (2003).

161. Fasching, M. et al. Coα-(1H-Imidazolyl)-Coβ-methylcob (III) amide: Model for Protein-Bound Corrinoid Cofactors. *Helvetica Chimica Acta* **83**, 2295-2316 (2000).

162. Dixon, M.M., Huang, S., Matthews, R.G. & Ludwig, M. The structure of the C-terminal domain of methionine synthase: presenting S-adenosylmethionine for reductive methylation of B12. *Structure* **4**, 1263-75 (1996).

163. Wolthers, K.R. et al. Crystal structure and solution characterization of the activation domain of human methionine synthase. *The FEBS Journal* **274**, 738-750 (2007).

164. Fick, R.J. et al. Water-Mediated Carbon–Oxygen Hydrogen Bonding Facilitates S-Adenosylmethionine Recognition in the Reactivation Domain of Cobalamin-Dependent Methionine Synthase. *Biochemistry* **57**, 3733-3740 (2018).

165. Doukov, T., Seravalli, J., Stezowski, J.J. & Ragsdale, S.W. Crystal structure of a methyltetrahydrofolate- and corrinoid-dependent methyltransferase. *Structure* **8**, 817-30 (2000).

166. Gencic, S. et al. Zinc-thiolate intermediate in catalysis of methyl group transfer in Methanosarcina barkeri. *Biochemistry* **40**, 13068-78 (2001).

167. Pejchal, R. & Ludwig, M.L. Cobalamin-independent methionine synthase (MetE): a face-to-face double barrel that evolved by gene duplication. *PLoS Biol* **3**, e31 (2005).

168. Ferrer, J.L., Ravanel, S., Robert, M. & Dumas, R. Crystal structures of cobalamin-independent methionine synthase complexed with zinc, homocysteine, and methyltetrahydrofolate. *J Biol Chem* **279**, 44235-8 (2004).

169. Fu, T.M. et al. Crystal structures of cobalamin-independent methionine synthase (MetE) from Streptococcus mutans: a dynamic zinc-inversion model. *J Mol Biol* **412**, 688-97 (2011).

170. LeClerc, G.M. & Grahame, D.A. Methylcobamide:coenzyme M methyltransferase isozymes from Methanosarcina barkeri. Physicochemical characterization, cloning, sequence analysis, and heterologous gene expression. *J Biol Chem* **271**, 18725-31 (1996).

171. Sauer, K. & Thauer, R.K. Methyl-coenzyme M formation in methanogenic archaea. Involvement of zinc in coenzyme M activation. *Eur J Biochem* **267**, 2498-504 (2000).

172. Burke, S.A. & Krzycki, J.A. Involvement of the "A" isozyme of methyltransferase II and the 29-kilodalton corrinoid protein in methanogenesis from monomethylamine. *J Bacteriol* **177**, 4410-6 (1995).

173. Hoeppner, A. et al. Structure of the corrinoid:coenzyme M methyltransferase MtaA from Methanosarcina mazei. *Acta Crystallographica Section D* **68**, 1549-1557 (2012).

174. Ferguson, D.J., Jr., Krzycki, J.A. & Grahame, D.A. Specific roles of methylcobamide:coenzyme M methyltransferase isozymes in metabolism of methanol and methylamines in Methanosarcina barkeri. *J Biol Chem* **271**, 5189-94 (1996).

175. Yeliseev, A., Gartner, P., Harms, U., Linder, D. & Thauer, R.K. Function of methylcobalamin: coenzyme M methyltransferase isoenzyme II in Methanosarcina barkeri. *Arch Microbiol* **159**, 530-6 (1993).

176. Kremer, J.D., Cao, X. & Krzycki, J. Isolation of two novel corrinoid proteins from acetate-grown Methanosarcina barkeri. *J Bacteriol* **175**, 4824-33 (1993).

177. Kappas, A. The porphyrias. *The Metabolic and Molecular Basis of Inherited Disease, 7^< th> Edition*, 2103-2159 (1995).

178. Phillips, J.D., Whitby, F.G., Kushner, J.P. & Hill, C.P. Structural basis for tetrapyrrole coordination by uroporphyrinogen decarboxylase. *The EMBO Journal* **22**, 6225-6233 (2003).

179. Paul, L. & Krzycki, J.A. Sequence and transcript analysis of a novel Methanosarcina barkeri methyltransferase II homolog and its associated corrinoid protein homologous to methionine synthase. *J Bacteriol* **178**, 6599-607 (1996).

180. Sjuts, H., Dunstan, M.S., Fisher, K. & Leys, D. Structures of the methyltransferase component of Desulfitobacterium hafniense DCB-2 O-demethylase shed light on methyltetrahydrofolate formation. *Acta Crystallogr D Biol Crystallogr* **71**, 1900-8 (2015).

181. Seravalli, J., Shoemaker, R.K., Sudbeck, M.J. & Ragsdale, S.W. Binding of (6R,S)-methyltetrahydrofolate to methyltransferase from Clostridium thermoaceticum: role of protonation of methyltetrahydrofolate in the mechanism of methyl transfer. *Biochemistry* **38**, 5736-45 (1999).

182. Seravalli, J., Zhao, S. & Ragsdale, S.W. Mechanism of transfer of the methyl group from (6S)-methyltetrahydrofolate to the corrinoid/iron-sulfur protein catalyzed by the methyltransferase from Clostridium thermoaceticum: a key step in the Wood-Ljungdahl pathway of acetyl-CoA synthesis. *Biochemistry* **38**, 5728-35 (1999).

183. Smith, A.E. & Matthews, R.G. Protonation state of methyltetrahydrofolate in a binary complex with cobalamin-dependent methionine synthase. *Biochemistry* **39**, 13880-90 (2000).

184. Jancarik, J. & Kim, S.-H. Sparse matrix sampling: a screening method for crystallization of proteins. *Journal of Applied Crystallography* **24**, 409-411 (1991).

185. Page, R. et al. Shotgun crystallization strategy for structural genomics: an optimized two-tiered crystallization screen against the Thermotoga maritima proteome. *Acta Crystallogr D Biol Crystallogr* **59**, 1028-37 (2003).

186. Newman, J. et al. Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy. *Acta Crystallogr D Biol Crystallogr* **61**, 1426-31 (2005).

187. Gilliland, G.L., Tung, M., Blakeslee, D.M. & Ladner, J.E. Biological Macromolecule Crystallization Database, Version 3.0: new features, data and the NASA archive for protein crystal growth data. *Acta Crystallogr D Biol Crystallogr* **50**, 408-13 (1994).

188. Radaev, S. & Sun, P.D. Crystallization of protein-protein complexes. *Journal of Applied Crystallography* **35**, 674-676 (2002).

189. Perrakis, A., Sixma, T.K., Wilson, K.S. & Lamzin, V.S. wARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models. *Acta Crystallogr D Biol Crystallogr* **53**, 448-55 (1997).

190. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410 (1990).

191. Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr D Biol Crystallogr* **59**, 1131-7 (2003).

192. Potterton, L. et al. CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallographica Section D* **74**, 68-84 (2018).

193. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).

194. Pannu, N.S. et al. Recent advances in the CRANK software suite for experimental phasing. *Acta Crystallographica Section D: Biological Crystallography* **67**, 331-337 (2011).

195. Cowtan, K.D. & Main, P. Phase combination and cross validation in iterated density-modification calculations. *Acta Crystallogr D Biol Crystallogr* **52**, 43-8 (1996).

196. Duvaud, S. et al. Expasy, the Swiss Bioinformatics Resource Portal, as designed by its users. *Nucleic Acids Research* **49**, W216-W227 (2021).

197. Read, R.J. & Schierbeek, A.J. A phased translation function. *Journal of Applied Crystallography* **21**, 490-495 (1988).

198. Goodsell, D.S., Autin, L. & Olson, A.J. Illustrate: Software for Biomolecular Illustration. *Structure* **27**, 1716-1720.e1 (2019).

199. Drozdetskiy, A., Cole, C., Procter, J. & Barton, G.J. JPred4: a protein secondary structure prediction server. *Nucleic Acids Research* **43**, W389-W394 (2015).

200. Williams, C.J. et al. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Science* **27**, 293-315 (2018).

201. Kraulis, P.J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *Journal of Applied Crystallography* **24**, 946-950 (1991).

202. Armstrong, D.R. et al. PDBe: improved findability of macromolecular structure data in the PDB. *Nucleic Acids Research* **48**, D335-D343 (2019).

203. McCoy, A.J. et al. Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).

204. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica Section D: Structural Biology* **75**, 861-877 (2019).

205. Novoradovsky, A. et al. Computational principles of primer design for site directed mutagenesis. in *Technical Proceedings of 2005 NSTI Nanotechnology Conference and Trade Show* Vol. 1 532-535 (Nano Science and Technology Institute Danville, CA, 2005).

206. Berman, H.M. et al. The Protein Data Bank. *Nucleic Acids Research* **28**, 235-242 (2000).

207. Murshudov, G., Vagin, A. & Dodson, E. Application of maximum likelihood refinement in the refinement of protein structures. *Proceedings of Daresbury Study Weekend*, 93-104 (1996).

208. Murshudov, G.N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-67 (2011).

209. Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr D Biol Crystallogr* **59**, 2023-30 (2003).

210. Sheldrick, G.M. A short history of SHELX. *Acta Crystallogr A* **64**, 112-22 (2008).

211. Consortium, T.U. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research* **51**, D523-D531 (2022).

212. Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-32 (2010).

213. Mullis, K. et al. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol* **51 Pt 1**, 263-73 (1986).

214. Dower, W.J., Miller, J.F. & Ragsdale, C.W. High efficiency transformation of E. coli by high voltage electroporation. *Nucleic Acids Res* **16**, 6127-45 (1988).

215. Van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L. & Clardy, J. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol* **229**, 105-24 (1993).

216. Sanger, F., Nicklen, S. & Coulson, A.R. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463-7 (1977).

217. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-5 (1970).

218. Butler, I.B., Schoonen, M.A. & Rickard, D.T. Removal of dissolved oxygen from water: A comparison of four common techniques. *Talanta* **41**, 211-5 (1994).

219. Matthews, B.W. Solvent content of protein crystals. *Journal of Molecular Biology* **33**, 491-497 (1968).

220. Hendrickson, W.A. Anomalous diffraction in crystallographic phase evaluation. *Q Rev Biophys* **47**, 49-93 (2014).

6 Supplement

6.1 Protein Purifications

The following figures show purification data obtained during all major protein purifications presented in this work. If a protein construct was purified more than once, one representative is shown. The purification conditions are listed in Table 12a and b. Throughout the cultivation and purification processes, samples for SDS-PAGE analysis were collected and labeled as follows:

Before induction: Cell culture sample right before the addition of IPTG

After induction: Cell culture sample before harvesting

Pellet: Sample of the cell debris after lysis and centrifugation

Supernatant: Sample of the supernatant after lysis and centrifugation

Flowthrough (1 or 2): Flowthrough of IMAC (1) or reverse IMAC (2)

Eluate (1 or 2): Eluate of IMAC (1) or reverse IMAC (2)

After dialysis: Samples obtained after dialysis

SEC/Peak fractions: Samples obtained during SEC

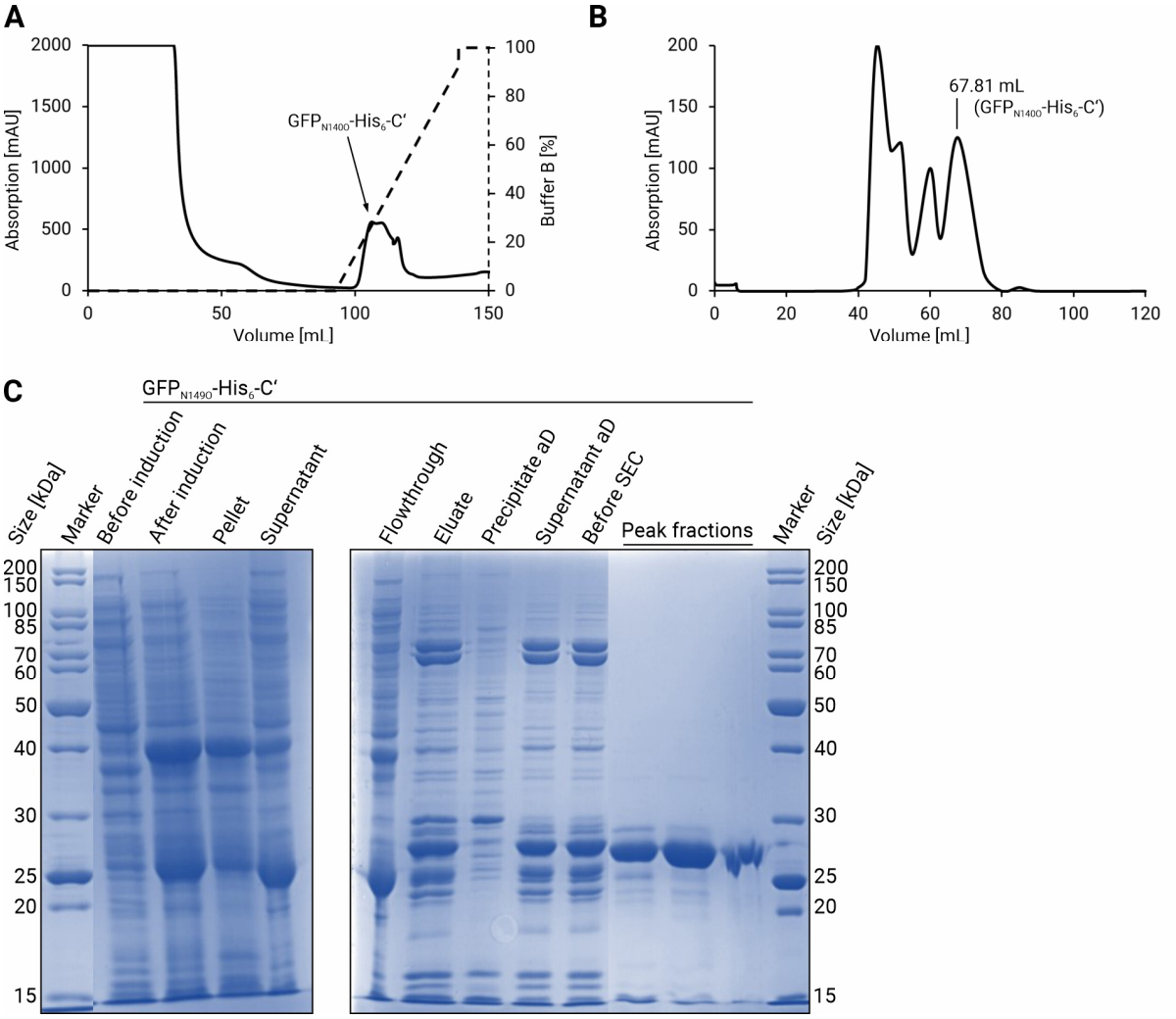


Figure S1 IMAC (**A**), SEC (**B**), and SDS-PAGE analysis (**C**) of C-terminally His$_6$-tagged GFP$_{N149O}$. To incorporate pyrrolysine via amber suppression, the parent gene was co-expressed with *pylB*, *pylC*, *pylD*, *pylS*, and *pylT* (vector constructs 5b and XYZ1-3c, s. section 2.1.1). The correct SEC fractions were identifiable via their green color (67.81 mL). SDS-PAGE of these fractions showed prominent bands at the correct size of 28 kDa. (aD = after Dialysis).

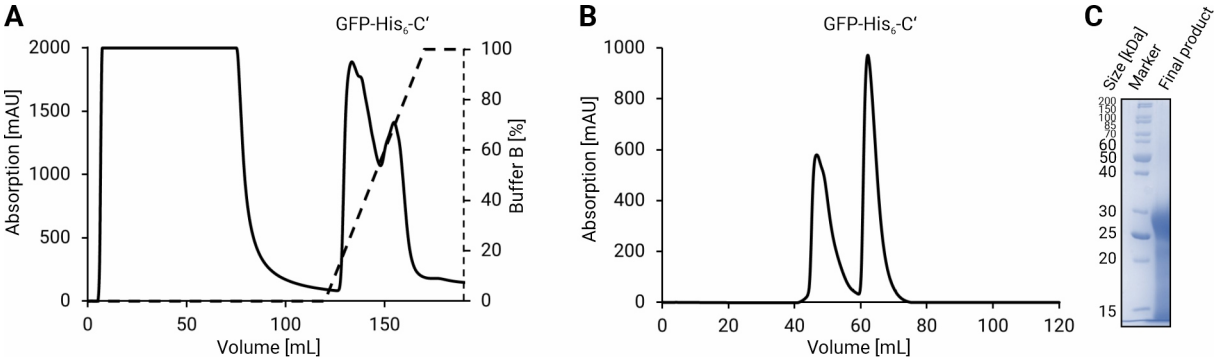


Figure S2 IMAC (**A**), SEC (**B**), and SDS-PAGE analysis (**C**) of GFP (vector construct W1-1a). To allow comparison with the amber suppressed protein version (GFP$_{N149O}$), the used construct was completely identical but with an asparagine at position 149.

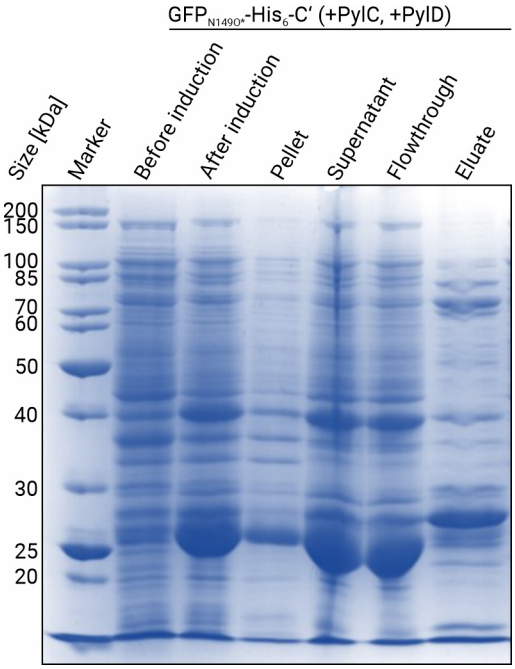


Figure S3 SDS-PAGE analysis of the NiNTA spin column purification of Pcl-incorporated (GFP$_{N1490*}$-His$_6$-C', (vector constructs 5b and YZ1-3c, s. section 2.1.1). Full-length C-terminally tagged GFP can be observed in the final eluate at the correct size of 28 kDa.

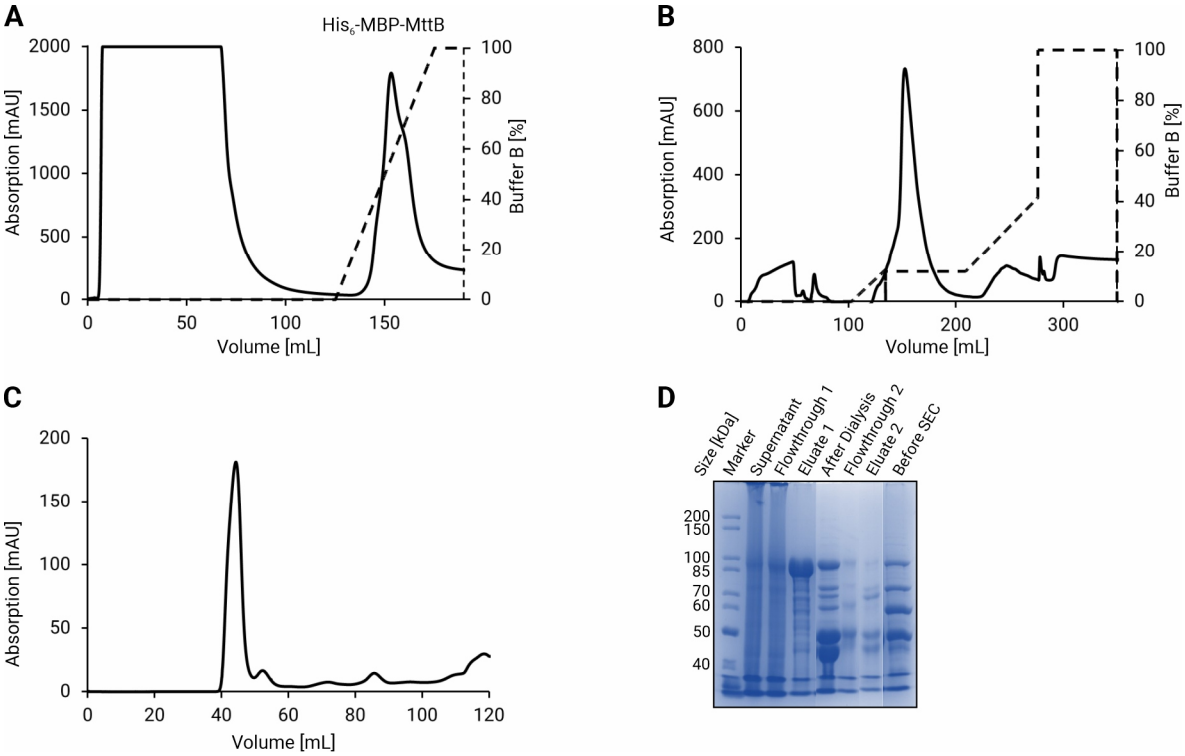


Figure S4 IMAC (**A**), reverse IMAC (**B**), SEC (**C**), and SDS-PAGE analysis (**D**) of His$_6$-MBP-tagged MttB (vector construct F19-1c). MBP-fused MttB was soluble but precipitated after tag removal (s. section 2.1.2).

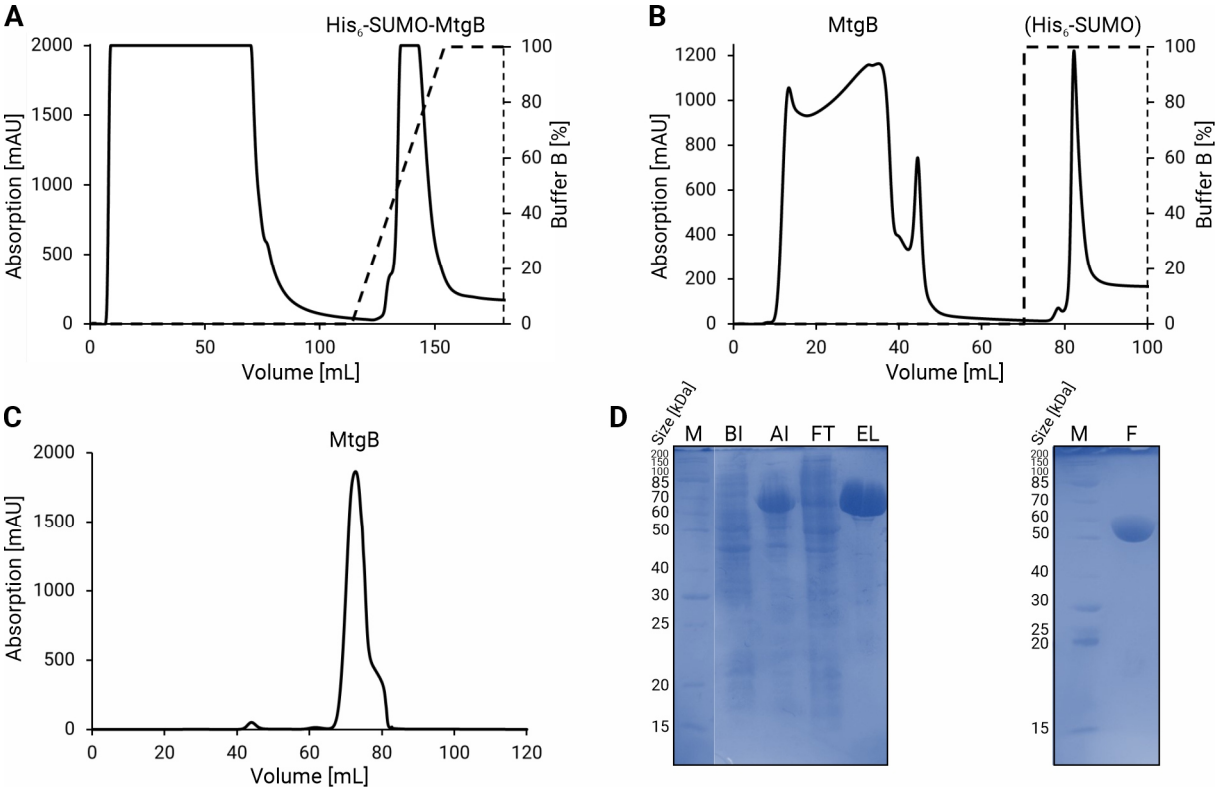


Figure S5 IMAC (**A**), reverse IMAC (**B**), SEC (**C**), and SDS-PAGE analysis (**D**) of N-terminally His_6-SUMO-tagged MtgB (vector construct I1-7a). (BI = Before induction, AI = After induction, FT = Flowthrough, EL = Eluate, F = Final product)

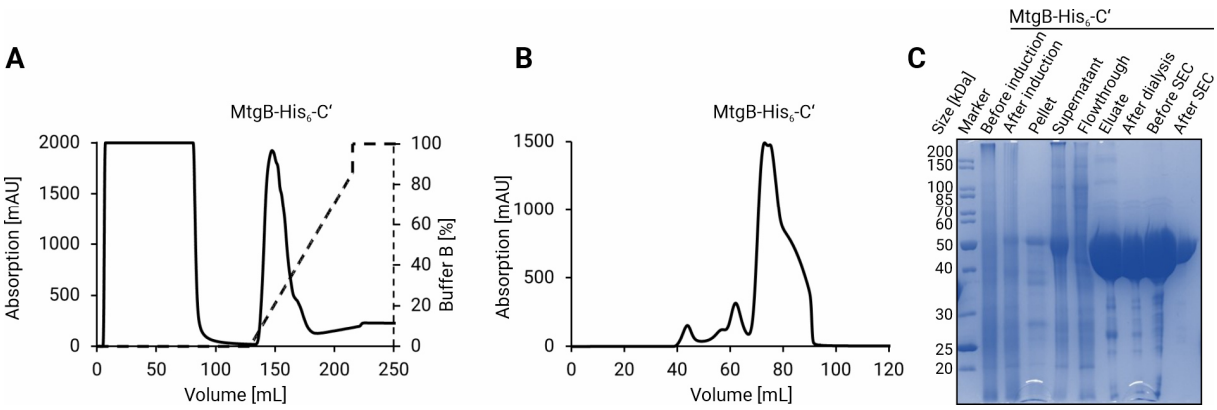


Figure S6 IMAC (**A**), SEC (**B**), and SDS-PAGE analysis (**C**) of C-terminally His_6-tagged MtgB (vector construct I1-7a).

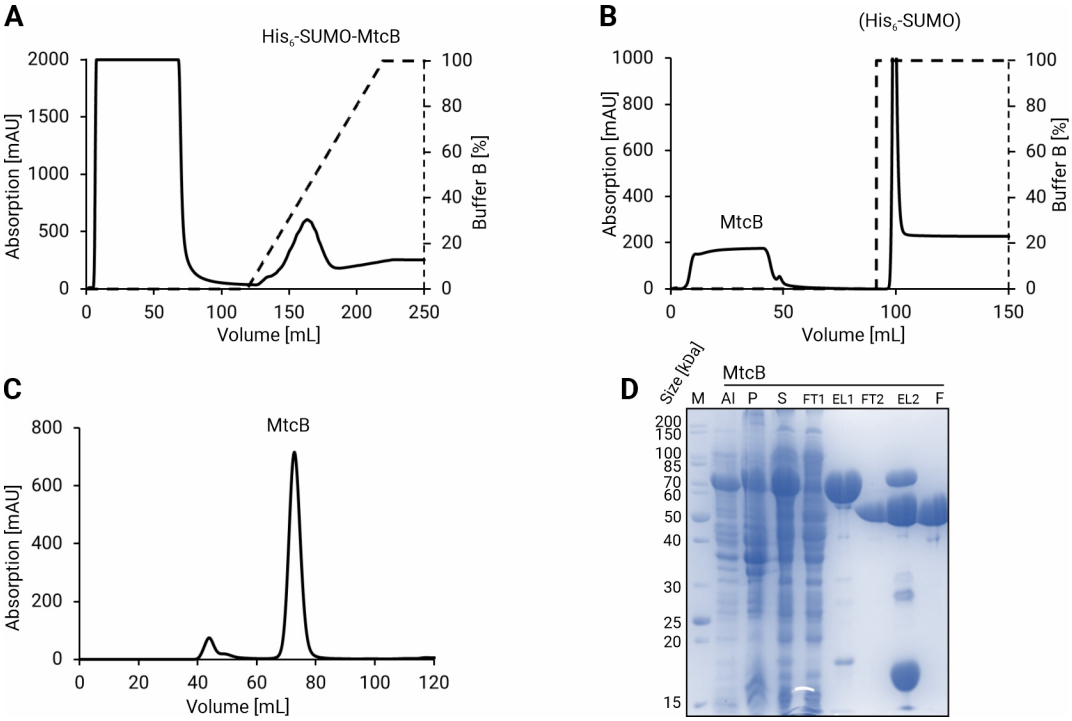


Figure S7 IMAC (**A**), reverse IMAC (**B**), SEC (**C**), and SDS-PAGE analysis (**D**) of N-terminally His_6-SUMO-tagged MtcB (vector construct O1-7a). (Al = After induction, P = Pellet, S = Supernatant, FT1 = Flowthrough 1, EL1 = Eluate 1, FT2 = Flowthrough 2, EL2 = Eluate 2, F = Final product)

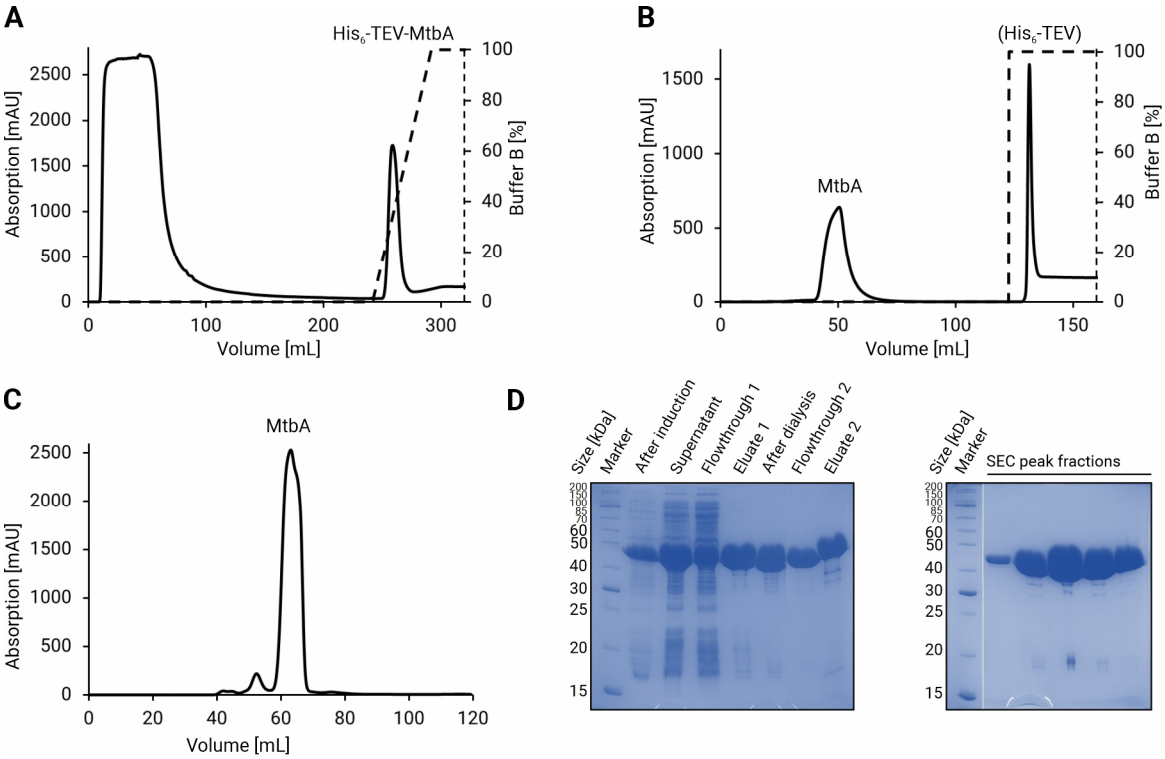


Figure S8 IMAC (**A**), reverse IMAC (**B**), SEC (**C**), and SDS-PAGE analysis (**D**) of N-terminally His_6-TEV-tagged MtbA (vector construct A1-6a).

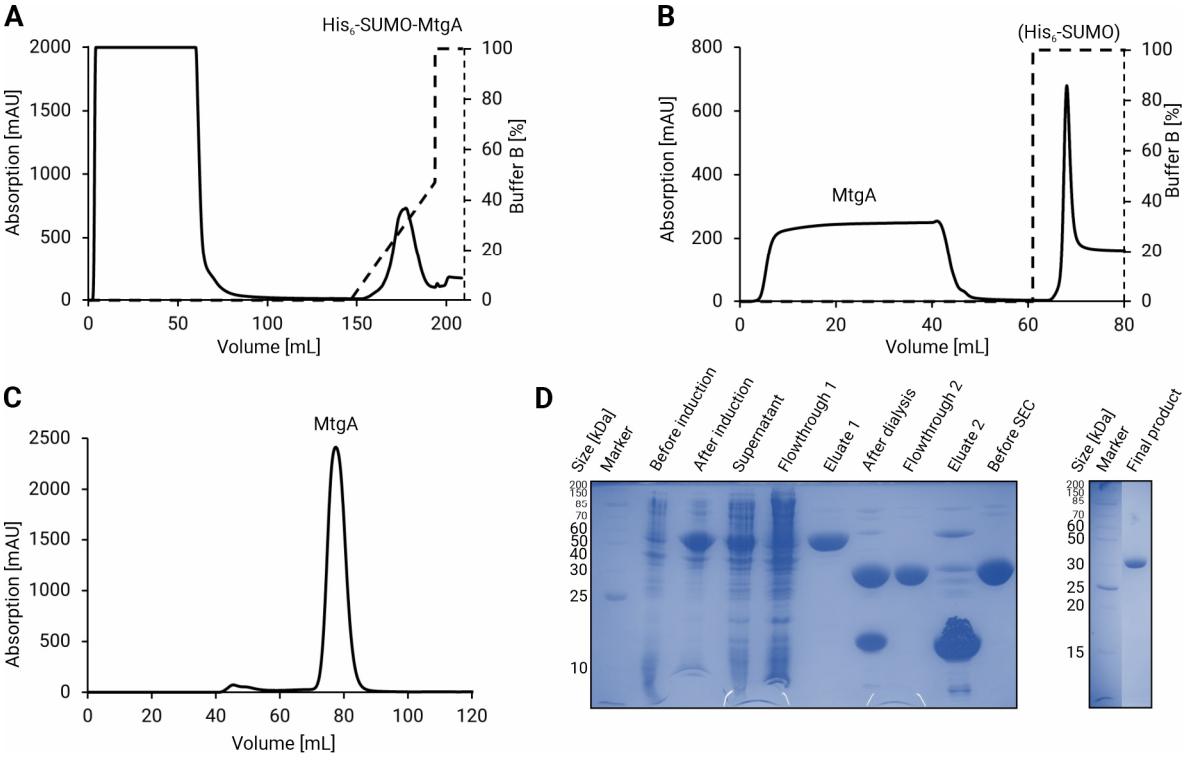


Figure S9 IMAC (**A**), reverse IMAC (**B**), SEC (**C**), and SDS-PAGE analysis (**D**) of N-terminally His_6-SUMO-tagged MtgA (vector construct L1-7a).

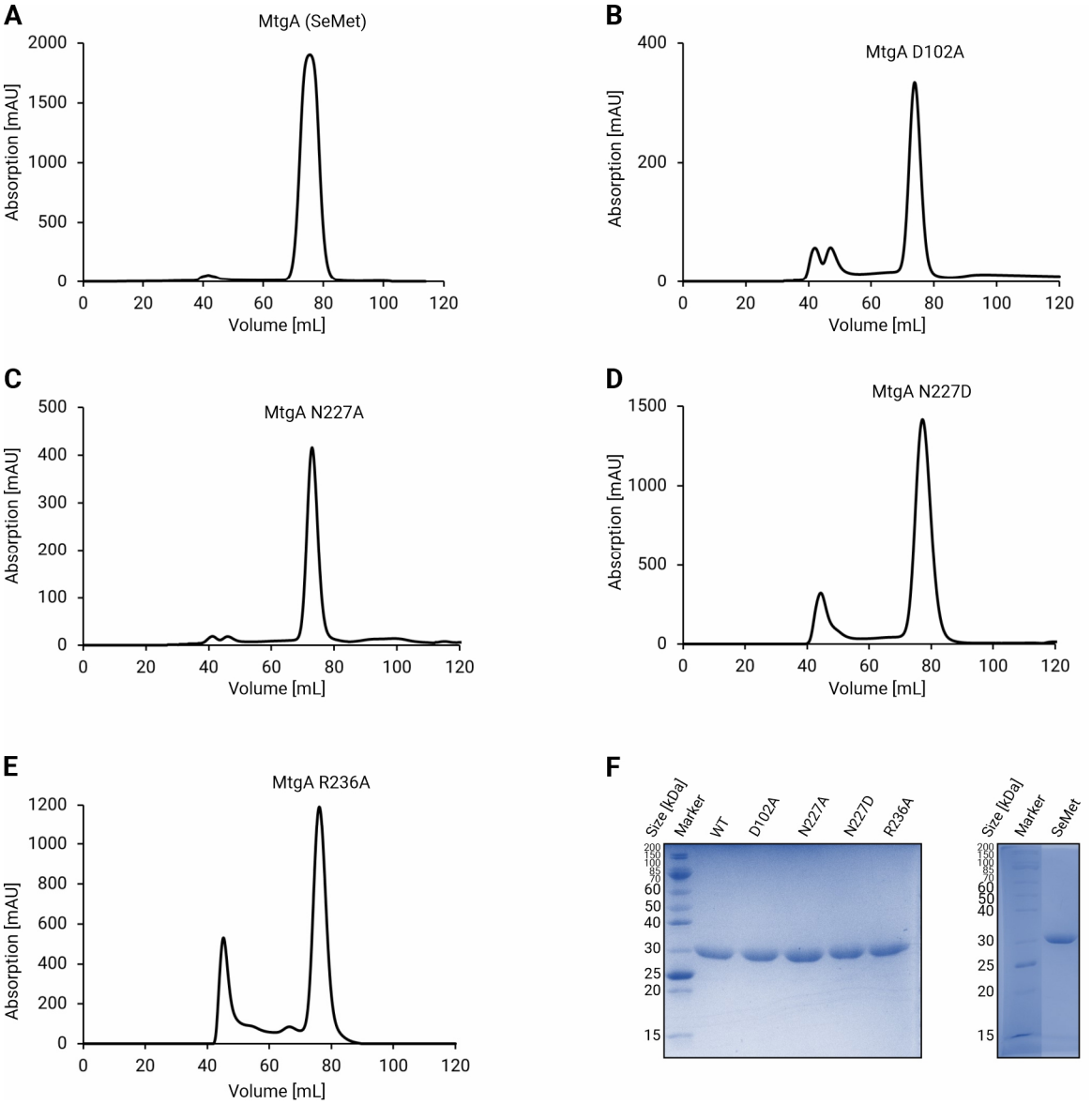


Figure S10 SEC diagrams and SDS-PAGE analysis of the purifications of SeMet-labeled MtgA (**A** and **F**) as well as several mutants (**B** – **F**).

6.2 Primary Sequence Alignments and Supplementary Figures

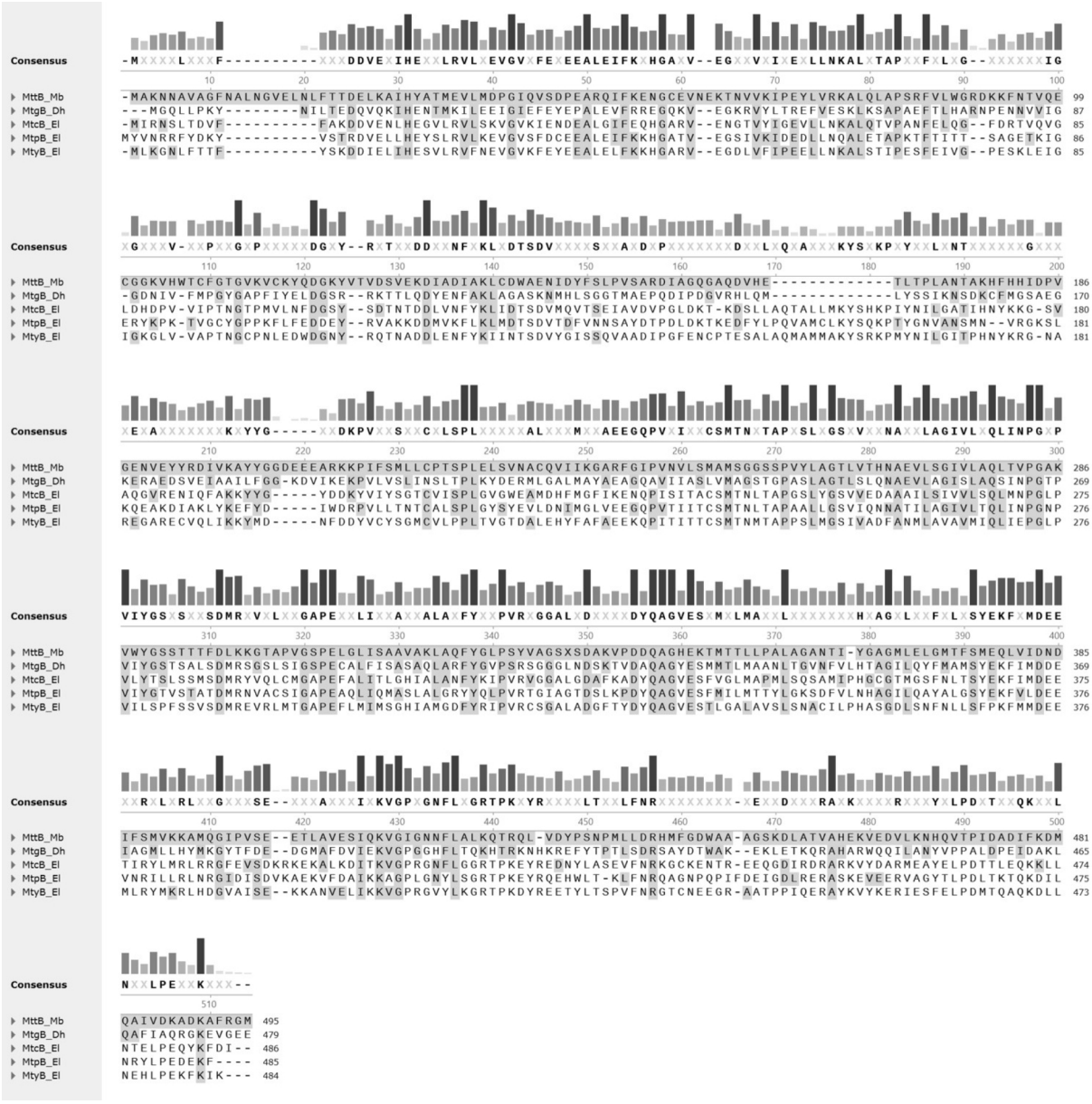


Figure S11 Sequence alignment of several characterized members of the MttB superfamily. MttB from *M. barkeri* has a pyrrolysine residue at position 334 (“X”, position 349 in consensus alignment). MtgB, MtcB, MtpB, and MtyB share around 40 - 50 % sequence similarity with MttB, but do not contain pyrrolysine.

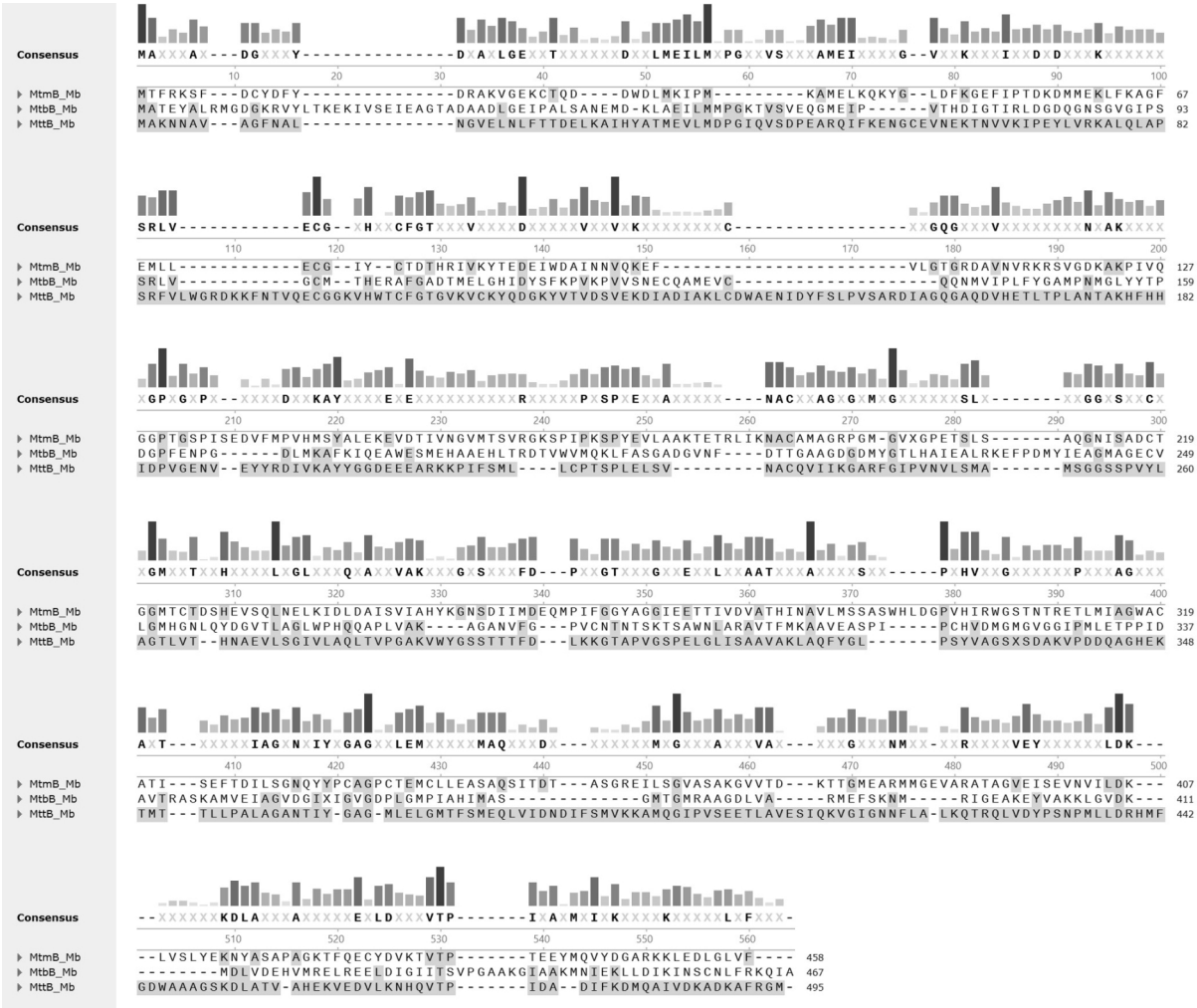
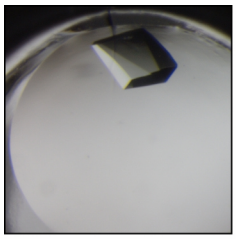
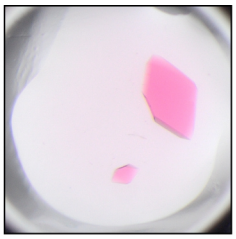


Figure S12 Sequence alignment of MtmB, MtbB, and MttB from *M. barkeri*. The three methyltransferases each contain a pyrrolysine residue ("X", consensus sequence positions 276, 386, and 419) but do not share significant sequence similarities.



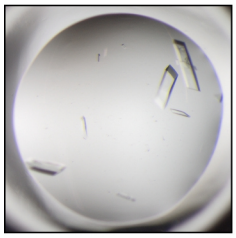
40 mg/mL MtgB (I1-7a)
with 100 mM DMG

Reservoir:
0.1 M Lithium sulfate
0.1 M Sodium acetate
0.1 M Tris pH 9.0
18 % (w/v) PEG 5000MME



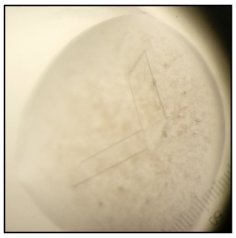
15 mg/mL MtgB (I2-1a)
with 1 mM Cbl

Reservoir:
2.2 M $(NH_4)_2SO_4$



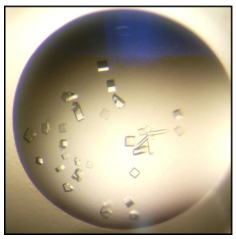
20 mg/mL MtcB (O1-7a)

Reservoir:
0.2 M Lithium citrate
20 % (w/v) PEG 3350



20 mg/mL MtbA (A1-6a)
with 2 mM CoM

Reservoir:
0.1 M Tris pH 8.0
19 % (w/v) PEG 3350
50 mM Zinc acetate



15 mg/mL MtgA (L1-7a)
with 5 mM THF

Reservoir:
0.1 M HEPES pH 7.5
25 % (w/v) PEG 3350

Figure S13 Exemplary protein crystals and growing conditions. A list of the exact conditions that led to the determined 3D structures in this work can be found in Table S1. Note that all pictures were taken without a polarizer. The cofactors Cbl and THF are responsible for the pink and yellow color of their respective protein crystals.

6.3 Crystallization Conditions

Table S1 Vector construct codes, structure identifiers, PDB-ID, and crystallization conditions (protein/reservoir ratio and components of the protein and reservoir solutions) of all determined crystal structures featured in this work.

Structure	Vector construct	Structure identifier, PDB-ID	Protein/reservoir ratio [µL/µL]	Protein solution	Reservoir solution
MtgB:Cbl	I2-1a	i2	0.3/0.1	15 mg/mL MtgB 1 mM Cbl 50 mM MOPS pH 7.2 0.5 mM DTT	2.2 M $(NH_4)_2SO_4$
MtgB:GB:Cbl	I2-1a	i3	0.2/0.2	15 mg/mL MtgB 5 mM GB 1 mM Cbl 20 mM Tris pH 7.5 100 mM NaCl 1 mM DTT	2.2 M $(NH_4)_2SO_4$
MtgB:DMG:Cbl	I2-1a	i4	0.3/0.1	15 mg/mL MtgB 5 mM DMG 1 mM Cbl 50 mM MOPS pH 7.2 0.5 mM DTT	2.2 M $(NH_4)_2SO_4$
MtgB:MeCbl	I2-1a	i5	0.2/0.2	15 mg/mL MtgB 1 mM MeCbl 20 mM Tris pH 7.5 100 mM NaCl 1 mM DTT	2.2 M $(NH_4)_2SO_4$ 0.2 M Cs_2SO_4
MtgB:DMG:MeCbl	I2-1a	i6	0.2/0.2	15 mg/mL MtgB 5 mM DMG 1 mM MeCbl 20 mM Tris pH 7.5 100 mM NaCl 1 mM DTT	2.2 M $(NH_4)_2SO_4$ 0.2 M Na_2SO_4
MtgB:GB	I1-7a	i7	0.2/0.2	30 mg/mL MtgB 2 mM GB 20 mM Tris pH 7.5 100 mM NaCl 1 mM DTT	0.2 M Li_2SO_4 20 mM Tris pH 8.5 18 % (*w/v*) PEG 8000
MtgB:DMG	I1-7a	i9	0.2/0.2	40 mg/mL MtgB 100 mM DMG 20 mM Tris pH 7.5 100 mM NaCl 1 mM DTT	0.1 M Li_2SO_4 0.1 M Sodium acetate 20 mM Tris pH 9.0 18 % (*w/v*) PEG 5000MME
MtgB:Bicine	I1-7a	i10	0.2/0.2	40 mg/mL MtgB 5 mM GB 20 mM Tris pH 7.5 100 mM NaCl 1 mM DTT	0.1 M Li_2SO_4 0.1 M Sodium acetate 20 mM Bicine pH 9.0 23 % (*w/v*) PEG 5000MME
MtcB (apo)	O1-7a	o1	0.2/0.2	20 mg/mL MtcB 1 mM *L*-Carnitine 50 mM MOPS pH 7.2 200 mM NaCl 10 % Glycerol 3 mM DTT	0.2 M Lithium citrate 20 % PEG3350
MtbA:CoM	A1-6a	a1 and a3	0.3/0.1	20 mg/mL MtbA 2 mM CoM 20 mM Tris pH 7.5 100 mM NaCl 5 mM DTT	100 mM Tris pH 8.0 19 % (*w/v*) PEG 3350 50 mM Zinc acetate
MtbA:MeCoM	A1-6a	a2	0.2/0.2	20 mg/mL MtbA 2 mM MeCoM 20 mM Tris pH 7.5 100 mM NaCl 5 mM DTT	100 mM Bicine pH 9.0 16 % (*w/v*) PEG 3350 20 mM Zinc acetate
MtgA:THF	L1-7a	L1, 6SJ8	0.2/0.1	15 mg/mL MtgA 5 mM THF 20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol	100 mM Bis-Tris pH 6.5 26 % (*w/v*) PEG 3350
MtgA:MeTHF	L1-7a	L2, 6SJN	0.2/0.2	15 mg/mL MtgA 5 mM MeTHF 20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol	100 mM HEPES pH 7.5 26 % (*w/v*) PEG 3350
MtgA:MeTHF ($P2_1$)	L1-7a	L3, 6SK4	0.3/0.1	15 mg/mL MtgA 10 mM MeTHF 20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol	100 mM Tris pH 8.0 29 % (*w/v*) PEG 3350

MtgA (apo)	L1-7a	L4, 6SJK	0.2/0.1	15 mg/mL MtgA 20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol	100 mM HEPES pH 7.5 27 % (*w/v*) PEG 3350
$MtgA_{D102A}$:MeTHF	L2-7a	L5, 6SJO	0.2/0.2	15 mg/mL $MtgA_{D102A}$ 5 mM MeTHF 20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol	100 mM Bis-Tris pH 6.2 29 % (*w/v*) PEG 3350
$MtgA_{N227A}$ (apo)	L7-7a	L6, 6SJP	0.3/0.1	15 mg/mL $MtgA_{N227A}$ 20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol	100 mM HEPES pH 7.5 19 % (*w/v*) PEG 3350
$MtgA_{N227A}$:THF	L7-7a	L7, 6SJR	0.2/0.1	15 mg/mL $MtgA_{N227A}$ 10 mM THF 20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol	100 mM HEPES pH 7.5 24 % (*w/v*) PEG 3350
$MtgA_{N227A}$:MeTHF	L7-7a	L8, 6SJS	0.2/0.1	15 mg/mL $MtgA_{N227A}$ 10 mM MeTHF 20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol	100 mM HEPES pH 7.8 22 % (*w/v*) PEG 3350
MtgA (SeMet)	L1-7a	L9	0.2/0.2	15 mg/mL MtgA 20 mM Tris pH 7.5 100 mM NaCl 10 % Glycerol 5 mM DTT	100 mM HEPES pH 7.8 21 % (*w/v*) PEG 3350

6.4 Crystallographic Data Collection und Refinement Statistics

The following remarks apply to all crystallographic data collection and refinement statistics tables:

[a] Asymmetric unit

[b] The values in parentheses for resolution range, completeness, R_{merge} and I/σ (I) correspond to the highest resolution shell

[c] Data reduction was carried out with XDS and from a single crystal. #Friedel pairs were treated as individual reflections

[d] $R_{merge}(I) = \Sigma_{hkl}\Sigma_j \mid I(hkl)_j - \langle I(hkl)\rangle \mid / \Sigma_{hkl}\, \Sigma_j\, I(hkl)_j$, where $I(hkl)_j$ is the j^{th} measurement of the intensity of reflection hkl and $\langle I(hkl)\rangle$ is the average intensity

[e] $R = \Sigma_{hkl} \mid |F_{obs}| - |F_{calc}| \mid / \Sigma_{hkl} |F_{obs}|$, where R_{free} is calculated without a sigma cut off for a randomly chosen 5% of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections

[f] Deviations from ideal bond lengths/angles

[g] Number of residues in favored region / allowed region / outlier region

Table S2 Crystallographic data collection and refinement statistics of the obtained MtgB datasets

	MtgB:Cbl	**MtgB:GB:Cbl**	**MtgB:DMG:Cbl**	**MtgB:MeCbl**
Crystal parameters				
Space group	$P6_222$	$P6_222$	$P6_222$	$P6_222$
Cell constants (Å)	a = 145.1	a = 145.1	a = 144.2	a = 144.6
	b = 145.1	b = 145.1	b = 144.2	b = 144.6
	c = 236.0	c = 235.5	c = 234.2	c = 234.9
Subunits / AU[a]	1	1	1	1
Data collection				
Beam line	X06SA, SLS	X06SA, SLS	X06SA, SLS	X06SA, SLS
Wavelength (Å)	1.0	1.0	1.0	1.0
Resolution range (Å)[b]	30-2.20	30-2.30	30-2.10	30-2.60
	(2.30-2.20)	(2.40-2.30)	(2.20-2.10)	(2.70-2.60)
No. observations	475287	859325	543579	375261
No. unique reflections[c]	72887	62868	84041	44757
Completeness (%)[b]	97.5 (99.4)	96.0 (97.3)	99.9 (100.0)	98.9 (99.7)
R_{merge} (%)[b, d]	11.3 (67.2)	10.4 (66.7)	9.7 (69.1)	11.4 (65.0)
I/σ(I)[b]	10.9 (2.7)	19.4 (3.8)	11.3 (2.4)	14.1 (4.1)
Refinement (REFMAC5)				
Resolution range (Å)	30-2.20	30-2.30	30-2.10	30-2.60
No. refl. working set	69213	59693	79804	42486
No. refl. test set	3643	3142	4200	2236
No. non hydrogen	4250	4382	4351	4179
No. of ligand atoms (DMG, Cbl, MeCbl)	96	99	98	97
Solvent (H_2O, ions, glycerol, Cbl, MeCbl)	532	592	586	387
R_{work} / R_{free} (%)[e]	14.2/16.1	12.7/15.3	14.2/16.2	14.7/17.0
r.m.s.d. bond (Å) / (°)[f]	0.004/1.7	0.003/1.8	0.003/1.7	0.003/1.8
Average B-factor (Å^2)	37.0	39.1	36.6	44.1
Ramachandran Plot (%)[g]	97.0/3.0/0	97.5/2.5/0	97.5/2.5/0	97.1/2.9/0
Structure identifier	**i2**	**i3**	**i4**	**i5**
PDB-ID	-	-	-	-

Table S3 Crystallographic data collection and refinement statistics of the obtained MtgB datasets

	MtgB:DMG:MeCbl	**MtgB:GB** (not finalized)	**MtgB:DMG**	**MtgB:Bic**
Crystal parameters				
Space group	$P6_222$	$P3_221$	$P3_221$	$P3_221$
Cell constants (Å)	a = 144.5	a = 123.4	a = 123.7	a = 123.4
	b = 144.5	b = 123.4	b = 123.7	b = 123.4
	c = 234.8	c = 124.8	c = 124.0	c = 124.2
Subunits / AU[a]	1	2	2	2
Data collection				
Beam line	X06SA, SLS	X06SA, SLS	X06SA, SLS	X06SA, SLS
Wavelength (Å)	1.0	1.0	1.0	1.0
Resolution range (Å)[b]	30-2.35 (2.45-2.35)	30-2.15 (2.25-2.15)	30-2.00 (2.10-2.00)	30-2.00 (2.10-2.00)
No. observations	795603	250122	306102	370525
No. unique reflections[c]	58645	57776	73657	73848
Completeness (%)[b]	96.4 (97.7)	96.2 (98.3)	99.1 (99.4)	99.7 (99.9)
R_{merge} (%)[b, d]	9.9 (65.8)	6.2 (64.3)	6.3 (68.4)	6.8 (65.6)
I/σ(I)[b]	18.3 (4.0)	12.9 (2.3)	12.0 (2.1)	11.9 (2.0)
Refinement (REFMAC5)				
Resolution range (Å)	30-2.35		30-2.00	30-2.00
No. refl. working set	55678		69957	70135
No. refl. test set	2931		3682	3691
No. non hydrogen	4319		7446	7582
No. of ligand atoms (DMG, GB, MeCbl, Bicine)	99		14	22
Solvent (H_2O, ions, glycerol, ethylene glycol, MeCbl)	525		77	226
R_{work} / R_{free} (%)[e]	14.1/16.3		19.4/23.7	16.3/19.7
r.m.s.d. bond (Å) / (°)[f]	0.003/1.8		0.003/1.2	0.002/1.2
Average B-factor (Å^2)	39.7		64.2	50.4
Ramachandran Plot (%)[g]	97.5/2.5/0		96.1/3.9/0	97.4/2.6/0
Structure identifier	**i6**	**i7**	**i9**	**i10**
PDB-ID	**-**	**-**	**-**	**-**

Table S4 Crystallographic data collection and refinement statistics of the obtained MtcB dataset

	MtcB apo
Crystal parameters	
Space group	C2
Cell constants (Å)	a = 111.0
	b = 100.9
	c = 91.3
	β = 111.0°
Subunits / AU[a]	2
Data collection	
Beam line	X06SA, SLS
Wavelength (Å)	1.0
Resolution range (Å)[b]	30-1.85
	(1.95-1.85)
No. observations	275856
No. unique reflections[c]	78493
Completeness (%)[b]	98.0 (97.6)
R_{merge} (%)[b, d]	4.4 (60.1)
I/σ(I)[b]	14.7 (2.3)
Refinement (REFMAC5)	
Resolution range (Å)	30-1.85
No. refl. working set	74552
No. refl. test set	3924
No. non hydrogen	6959
No. of ligand atoms	-
Solvent (H_2O, glycerol)	383
R_{work} / R_{free} (%)[e]	14.8/18.6
r.m.s.d. bond (Å) / (°)[f]	0.002/1.2
Average B-factor (Å^2)	35.9
Ramachandran Plot (%)[g]	98.1/1.8/0.1

Structure identifier	**o1**
PDB-ID	**-**

Table S5 Crystallographic data collection and refinement statistics of the obtained MtbA datasets

	MtbA:CoM	**MtbA:CoM** [Zn^{2+} (peak)]	**MtbA:MeCoM**
Crystal parameters			
Space group	$P2_12_12_1$	$P2_12_12_1$	P1
Cell constants (Å)	a = 48.0 b = 49.4 c = 129.5	a = 48.0 b = 49.4 c = 129.5	a = 48.7 b = 51.4 c = 67.9 α = 92.7° β = 90.4° γ = 89.8°
Subunits / AU[a]	1	1	2
Data collection			
Beam line	X06SA, SLS	X06SA, SLS	X06SA, SLS
Wavelength (Å)	1.0	1.284	1.0
Resolution range (Å)[b]	30-1.25 (1.35-1.25)	30-1.50 (1.60-1.50)	30-2.20 (2.30-2.20)
No. observations	436353	690565	85406
No. unique reflections[c]	85904	47620[#]	30845
Completeness (%)[b]	99.7 (99.7)	99.9 (100.0)	92.5 (91.6)
R_{merge} (%)[b, d]	9.3 (58.4)	11.0 (47.7)	10.9 (58.8)
I/σ(I)[b]	9.3 (2.6)	10.5 (3.8)	6.6 (2.1)
Refinement (REFMAC5)			
Resolution range (Å)	15-1.25		30-2.20
No. refl. working set	81543		29291
No. refl. test set	4292		1542
No. non hydrogen	2982		5286
No. of ligand atoms (CoM, Zn^{2+})	8		18
Solvent (H_2O, ions, glycerol)	412		170
R_{work} / R_{free} (%)[e]	12.3/14.3		18.6/22.5
r.m.s.d. bond (Å) / (°)[f]	0.013/1.7		0.002/0.7
Average B-factor ($Å^2$)	13.8		37.5
Ramachandran Plot (%)[g]	97.6/2.4/0		95.8/3.9/0.3
Structure identifier	**a1**	**a3**	**a2**
PDB-ID	**-**	**-**	**-**

Table S6 Crystallographic data collection and refinement statistics of the obtained MtgA datasets

	MtgA [SeMet]	**MtgA apo**	**MtgA:THF**	**MtgA:MeTHF**	**MtgA:MeTHF** ($P2_1$)
Crystal parameters					
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_1$
Cell constants (Å)	a = 76.2 b = 83.8 c = 86.7	a = 76.1 b = 84.1 c = 86.8	a = 75.8 b = 83.9 c = 86.5	a = 75.9 b = 83.8 c = 86.6	a = 83.4 b = 73.1 c = 89.8 β = 113.4°
Subunits / AU[a]	2	2	2	2	4
Data collection					
Beam line	X06SA, SLS	X06SA, SLS	X06SA, SLS	X06SA, SLS	X06SA, SLS
Wavelength (Å)	0.979	1.0	1.0	1.0	1.0
Resolution range (Å)[b]	30-1.90 (2.00-1.90)	30-1.85 (1.95-1.85)	30-1.35 (1.45-1.35)	30-1.75 (1.85-1.75)	30-1.55 (1.65-1.55)
No. observations	401906	195313	493740	227947	427254
No. unique reflections[c]	82696[#]	46728	120092	55629	139692
Completeness (%)[b]	97.9 (95.8)	96.9 (99.2)	99.1 (99.6)	98.7 (99.4)	97.2 (96.3)
R_{merge} (%)[b, d]	9.4 (47.1)	8.5 (57.4)	4.8 (64.5)	6.3 (51.5)	7.1 (55.7)
I/σ(I)[b]	10.8 (3.3)	9.4 (2.3)	13.8 (2.1)	13.9 (2.7)	8.1 (2.2)
Refinement (REFMAC5)					
Resolution range (Å)		30-1.85	30-1.35	30-1.75	30-1.55
No. refl. working set		44379	113987	52746	132691
No. refl. test set		2336	5999	2776	6984
No. non hydrogen		4898	5351	5144	9768
No. of ligand atoms (THF)		-	26	33	132
Solvent (H_2O, ions, glycerol)		212	617	425	408
R_{work} / R_{free} (%)[e]		17.3/20.4	11.9/15.7	13.7/17.3	20.9/24.3
r.m.s.d. bond (Å) / (°)[f]		0.002/1.2	0.01/1.5	0.007/1.2	0.004/1.3
Average B-factor (Å^2)		26.8	19.4	23.2	21.9
Ramachandran Plot (%)[g]		98.7/1.3/0	99.3/0.7/0	99.3/0.7/0	99.3/0.7/0
Structure identifier	**L9**	**L4**	**L1**	**L2**	**L3**
PDB-ID	**-**	**6SJK**	**6SJ8**	**6SJN**	**6SK4**

Table S7 Crystallographic data collection and refinement statistics of the obtained MtgA datasets

	$MtgA_{D102A}$:MeTHF	**$MtgA_{N227A}$ apo**	**$MtgA_{N227A}$:THF**	**$MtgA_{N227A}$:MeTHF**
Crystal parameters				
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell constants (Å)	a = 77.0	a = 76.0	a = 76.0	a = 76.2
	b = 84.3	b = 83.9	b = 84.0	b = 84.3
	c = 87.6	c = 86.6	c = 86.4	c = 86.7
Subunits / AU[a]	2	2	2	2
Data collection				
Beam line	X06SA, SLS	X06SA, SLS	X06SA, SLS	X06SA, SLS
Wavelength (Å)	1.0	1.0	1.0	1.0
Resolution range (Å)[b]	30-1.95 (2.05-1.95)	30-1.90 (2.00-1.90)	30-1.75 (1.85-1.75)	30-1.80 (1.90-1.80)
No. observations	214249	172347	218088	191359
No. unique reflections[c]	41877	43506	55277	51389
Completeness (%)[b]	99.2 (99.2)	98.1 (98.9)	97.9 (95.4)	98.1 (99.2)
R_{merge} (%)[b, d]	9.2 (57.9)	12.6 (57.0)	9.2 (56.4)	9.3 (59.5)
I/σ(I)[b]	11.5 (2.6)	7.5 (3.0)	11.2 (2.4)	8.6 (2.2)
Refinement (REFMAC5)				
Resolution range (Å)	30-1.95	30-1.90	30-1.75	30-1.80
No. refl. working set	39770	41319	52502	48810
No. refl. test set	2093	2175	2763	2569
No. non hydrogen	4818	4867	4996	4890
No. of ligand atoms (THF)	33	-	26	46
Solvent (H_2O, ions, glycerol)	106	165	268	164
R_{work} / R_{free} (%)[e]	18.5/21.7	18.9/22.4	16.9/18.9	17.9/21.9
r.m.s.d. bond (Å) / (°)[f]	0.004/1.3	0.004/1.2	0.002/1.2	0.003/1.2
Average B-factor ($Å^2$)	31.9	23.4	18.7	22.3
Ramachandran Plot (%)[g]	99.0/1.0/0	99.0/1.0/0	98.8/1.2/0	99.0/1.0/0
Structure identifier	**L5**	**L6**	**L7**	**L8**
PDB-ID	**6SJO**	**6SJP**	**6SJR**	**6SJS**

7 Publications

Authored publications in peer-reviewed journals or manuscripts in preparation that are part of this thesis:

2020 **Badmann T.**, Groll M., "Structures in Tetrahydrofolate Methylation in Desulfitobacterial Glycine Betaine Metabolism at Atomic Resolution"
ChemBioChem, 21 (6), 776-779

2024 **Badmann T.**, Groll M., "High-Resolution Structures of the Methanogenic Methyltransferase MtbA Give Detailed Insights into the Methylation of Coenzyme M"
Manuscript in preparation

2024 **Badmann T.**, Groll M., "Structures of a Desulfitobacterial Homolog of the Pyrrolysine Protein Family Depict the Methyl Transfer from Glycinebetaine to Cobalamin at Atomic Resolution"
Manuscript in preparation

Results that are not subject of this thesis have been published in the peer-reviewed journal ACS Chemical Biology:

2023 Mordhorst S.*, **Badmann T.***, Bösch N. M., Morinaka B. I., Rauch H., Piel J., Groll M., Vagstad A. L., "Structural and Biochemical Insights into Post-Translational Arginine-to-Ornithine Peptide Modifications by an Atypical Arginase",
**these authors contributed equally*
ACS Chemical Biology, 18 (3), 528-536

8 Declaration of Authorship

I, Thomas Patrick Badmann, hereby declare that the submitted thesis is my own unaided work. All direct or indirect sources used are acknowledged as references.

This thesis has not yet been submitted to any examination board. Parts of this work have been or will be published in scientific journals.

Garching, April 15, 2024

_____

Thomas Patrick Badmann