Investigations into Methyl-coenzyme M Reductase Behavior: Expression in a Heterologous Host, Putative Post-Translational Modification Genes, and Molecular Dynamics of MCR Homologues Bound to F430 Variants

by

Chidinma Lucy Odili

A dissertation submitted to the Graduate Faculty of Auburn University

in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

Auburn, Alabama

May 4, 2024

Keywords: Methyl-coenzyme M reductase, coenzyme F430, heterologous expression, post-translational modifications, molecular dynamics

Approved by

Steven O. Mansoorabadi, Chair, Associate Professor of Chemistry and Biochemistry

Eduardus C. Duin, Professor of Chemistry and Biochemistry

Douglas C. Goodwin, Professor of Chemistry and Biochemistry

Holly Ellis, Professor of Chemistry and Biochemistry

Abstract

Methanogens are microorganisms widely found in wetlands and the digestive tracts of animals, which produce methane as a metabolic byproduct. The organism possesses the key methane-forming enzyme methyl-coenzyme M reductase (MCR) which is a dimer of heterotrimers comprised of McrA (α), McrB (β), and McrG (γ) subunits. The enzyme uses a unique nickel-containing coenzyme F430 for activity and contains several unprecedented post-translational modifications (PTMs). The six PTMs located in α subunit of MCR are 2-(*S*)-methylglutamine, 5-(*S*)-methylarginine, 3-methylhistidine, *S*-methylcysteine, didehydroaspartate, and thioglycine residues. Homologues of MCR have been identified in anaerobic methanotrophic archaea (ANME) which operates the anaerobic oxidation of methane (AOM), and in *Candidatus Ethanoperedens thermophilum* which is an ethane oxidizer. While the MCR homologues share the same structural composition, they differ in the composition of their PTMs, with a 17^2-methylthio F430 and a $17,17^2$-dimethyl F430 contained in ANME-1 and in *Ca E. thermophilum* respectively. The exact roles of these PTMs are unknown though several hypotheses have been proposed including their role in improving MCR stability under mesophilic conditions. Processes involved in the maturation and activation of the active enzyme are not yet fully understood. A better understanding of the assembly and activation of MCR may enable its application in natural gas conversion strategies and the development of inhibitors to reduce natural greenhouse gas emissions. Current investigation and progress in some of the PTMs and the expression of MCR in a heterologous, non-methanogenic host is described. Comparative genomics and homology studies were used to identify target genes suspected to be responsible for the PTMs. Putative genes reported (*mcmA*) and suspected (*mm4*) to be responsible for the methylcysteine and methylhistidine PTMs respectively, were studied using a combination of computational tools, *in vitro* and *in vivo* methods. Molecular dynamics (MD) methods and distance calculations were utilized to investigate the effects of the PTMs on MCR and its homologues. MD simulations were carried out on MCR homologues with and without their PTMs. Results show that MCR exhibits half sites reactivity, and that the PTMs may play a role in coordinating MCR catalytic activity within its two active sites. MCR homologues without PTMs exhibited less dynamism and this could be an indicator that the PTMs evolved as an adaptation of thermophilic proto-MCR to mesophilic growth, enabling proper enzyme dynamics at lower temperature. The requirement of zinc as an accessory factor for the activity of McmA in the methylcysteine modification was inferred from the results. Soluble McrA and McrG proteins were successfully expressed from *Escherichia coli* cell lines. McrG was coeluted with several coenzyme F430 biosynthetic (Cfb) proteins. Previous work by the Mansoorabadi group showed that CfbE interacts with McrD and could be activated by it. The coelution of McrG with Cfb proteins suggests that the Cfb proteins may form a larger coenzyme F430 synthase complex which interacts with McrG and McrD to coordinate the insertion of coenzyme F430 into the MCR active site. CfbE interaction with McrG indicates that CfbE might be the direct F430 chaperone that delivers the coenzyme to MCR. Considering MCR's strategic role in methane recycling, understanding these processes are pivotal for the enzyme's application in strategies for natural gas conversion and reduction.

Acknowledgements

I would like to start with acknowledging my supervisor, Dr. Steve O. Mansoorabadi for his excellent supervision and mentorship. Graduate school is fraught with lots of curve balls and unexpected turns. Having an understanding and invested supervisor/mentor is therefore very valuable towards maintaining a positive attitude and remaining motivated. Dr. Steve's ever calm demeanor and grace, his understated genius, and persistent positivity even in the face of daunting challenges not only taught me a life lesson in gentle resilience but kept me sane on the many days when I felt defeated and subdued. Dr. Steve's supervision was therefore very instrumental in the successful completion of my Ph.D.

I would also like to acknowledge Dr. Duin for sharing his insights on MCR as well as his laboratory and instruments with me. My research would never have progressed as well as it did without his input and the access I was afforded to his lab.

Special thanks to Dr. Goodwin, who as the then GPO for Chemistry department, ensured my smooth transition into the department as a new international student who was still learning the ropes. I am also very grateful for the use of his HPLC in the final stages of my research.

I would like to extend my gratitude to every member of my lab including Prosenjit, Chelsea, Luke, Joshua, and Dominic for making the lab a safe space for us all to work amicably. I would also like to appreciate the efforts of previous lab members such as Dr. Shuxin, who was always willing to answer my questions and listen to my stories, Dr. Patrick who nudged me into getting some computational experience, Dr. Shadi, who gladly shared her expertise with cloning plasmids with me, and our former Post-Doc Dr. Selam, who handled my training on basic lab concepts when I newly joined the lab. My Ph.D. journey will be incomplete without these people.

My sincere appreciation goes to members of the Duin lab, Dr. Robel, Theophila, Arielle and Chuks who gladly shared their lab space and instruments with me, as well as members of the Goodwin lab, Chidozie, Rejaul and Nana for their understanding every time I had to use the HPLC and work late into the night. I would also like to thank former Faculty Dr. Monica Raj and members of her lab, Dr. Ogonna and Dr. Victor for their training and help in synthesizing peptides for my experiments.

Special thanks to Melissa Boersma at the Auburn University Department of Chemistry and Biochemistry Mass Spectrometry Center and to the University of Georgia Proteomics and Mass Spectrometry Facility for the analysis of my samples. Finally, to the Alabama supercomputer for the generous grant of computer time.

This Ph.D. is dedicated to God Almighty!

Several times along the way I lost faith in myself and in my research and wanted to give up, run back to Nigeria and start a business. Only God's grace and comfort could have lifted me from depths of depression on those dark, lonely days and opened my eyes to new opportunities to explore. Only God's strength could have empowered me to finish well.

To my late Father, Michael Nwachukwu Odili, who inspired me to be and do all I was gifted for, irrespective of my gender, but did not live to see his girl graduate with a Ph.D.

To my sweet mother, Mrs. Chinyere Benardine Odili whose tireless encouragement and priceless words of wisdom were not only timely, but always helped me to put things in perspective, and to ensure that I strived for balance in every area of my life.

I come from a large family of seven, made even larger by the blessing of in-laws, nephews, and nieces. I have also been blessed to be adopted by the Denard and Brown families here in Alabama who gave me a home away from home. Space will not allow me to mention everyone by name. This work is dedicated to every member of my family who prayed with and supported me every step of the way. I am because we are!

Finally, to my husband, Mr. Olumide Babatope Falana for his unwavering support and gentle guidance, and for the many nights we stayed up together in the lab, willing my many cell lines to grow well. I love you babe.

Table of Contents

List of Tables

List of Figures

List of Schemes

List of Abbreviations

MFR Methanofuran

H_4MPT Tetrahydromethanopterin

CoM Coenzyme M

CoA Coenzyme A

CoB Coenzyme B

ANME Anaerobic methanotrophic archaea

AOM Anaerobic oxidation of methane

SRB Sulfate-reducing bacteria

MCR Methyl coenzyme M reductase

ECR Ethyl coenzyme M reductase

BPS 3-Bromopropanesulfonate

PTM Post translational modification

SAM S-adenosylmethionine

SAH S-adenosylhomocysteine

Mm10 Methanogenesis marker protein 10

ATP Adenosine triphosphate

ADP Adenosine diphosphate

Cfb Coenzyme F430 biosynthesis

TB Terrific broth

LB Luria-Bertani

HPLC High-performance liquid chromatography

MS Mass spectrometry

ESI Electrospray ionization

SDS-PAGE Sodium dodecyl-sulfate polyacrylamide gel electrophoresis

IPTG Isopropyl β-D-1-thiogalactopyranoside

NAD^+ Nicotinamide adenine dinucleotide

NADPH Nicotinamide adenine dinucleotide phosphate

PCR Polymerase chain reaction

MALDI-TOF Matrix-assisted laser desorption/ionization time-off-light

MD Molecular dynamics

WT Wildtype

DFT Density functional theory

IMAC Immobilized metal affinity chromatography

DMF Dimethylformamide

ACN Acetonitrile

Chapter One

Introduction

1 Chapter 1

1.1 The process of methanogenesis

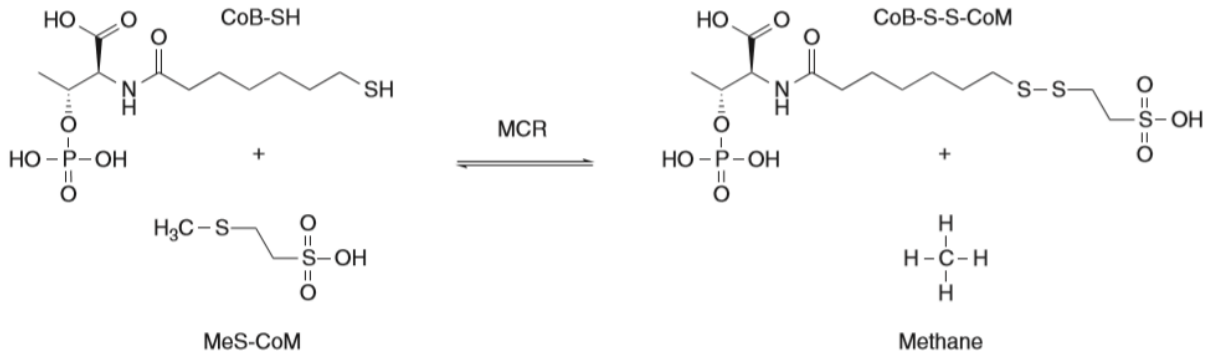
1.1.1 Methane and its significance

Methane, recognized as a potent greenhouse gas, plays a pivotal role in the ongoing issue of climate change due to its ability to trap solar radiation[1,2]. Methane also has 30 times the global warming potential of carbon dioxide[1]. Furthermore, an estimated over one billion tons of methane is produced each year by methanogens[3], part of which escapes to the atmosphere (20 % to 30 %) and contributes to greenhouse gases[4,5]. This makes it imperative that we delve into the intricate mechanisms underpinning methane production and consumption to address the environmental challenges we face.

Methanogens are a group of organisms that produce methane through a process known as methanogenesis. In 1933, Stephenson and Strickland reported the isolation of a pure culture of a methanogen[6], and this marked the beginning of a new era for research into the process of methanogenesis. Cell lysates of the isolated organism were found to catalyze the conversion of certain carbon sources such as carbon dioxide, carbon monoxide, methyl alcohol, formic acid, and formaldehyde into methane by means of molecular hydrogen[6].

1.1.2 Methanogenesis reaction

Studies into the processes of methanogenesis by methanogens has confirmed the role of the enzyme methyl-coenzyme M reductase (MCR) as the key enzyme in methanogenesis[5–7]. MCR catalyzes the last step of methanogenesis, that is, the reduction of methyl-coenzyme M (CH_3-S-CoM) with coenzyme B (CoB-SH) to produce methane and the heterodisulfide of CoM and CoB (CoB-S-S-CoM)[8,9] as shown in *Scheme 1*. This occurs under strictly anerobic *conditions.*



Scheme 1: Reaction catalyzed by MCR[10]

Most organisms exist in consortium. As a result, the biochemical pathway of methanogenesis is usually linked to other biochemical pathways[11]. In anaerobic freshwater sediments, glucose, which is mostly derived from cellulose is first converted to acetate, carbon dioxide and hydrogen or to acetate, formate, and hydrogen by strictly anaerobic bacteria[11]. This process is also referred to as glucose fermentation.

$$\text{Glucose} + 2H_2O \longrightarrow 2CH_3COOH + 2CO_2 + 4H_2$$

$$\text{Glucose} + 2H_2O \longrightarrow 2CH_3COOH + 2HCOOH + 2H_2$$

The products from glucose fermentation are converted to methane by methanogens.

$$CH_3COOH \longrightarrow CO_2 + CH_4$$

$$4H_2 + CO_2 \longrightarrow CH_4 + 2H_2O$$

$$4HCOOH \longrightarrow CH_4 + 3CO_2 + 2H_2O$$

Other carbon substrates utilized by methanogens for the process of methanogenesis include methylthiols, methylamines, and methanol.

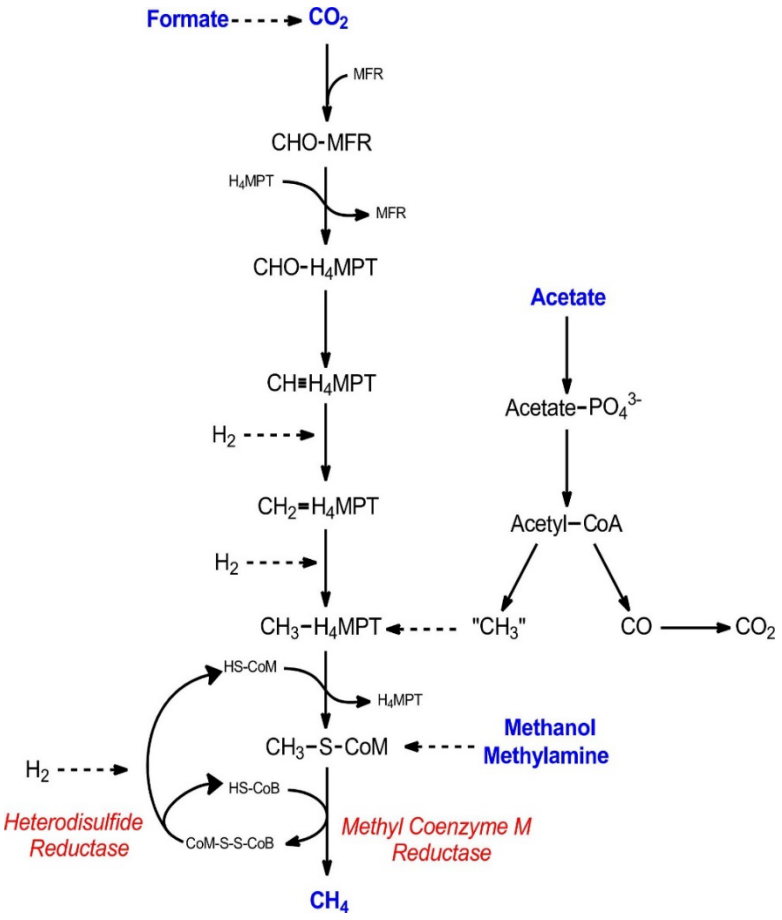
$$4CH_3OH + 2H_2O \longrightarrow 3CH_4 + CO_2 + 4H_2O$$

Based on the substrate being metabolized, there are three main branches for the methanogenesis process: acetoclastic, hydrogenotrophic, and methylotrophic methanogenesis. Irrespective of the branch, all methyl group donors transfer their methyl group to coenzyme M, but through different pathways. For acetoclastic organisms, acetate is used as the substrate. In a reaction driven by ATP, acetate reacts with coenzyme A to form acetyl-CoA. Acetyl-CoA is converted to CO_2 and a corrinoid protein-bound methyl group. The methyl group is transferred to tetrahydromethanopterin (H_4MPT) by the action of carbon monoxide dehydrogenase (CODH) and acetyl-CoA synthase (ACS) complex[12] to yield methyltetrahydromethanopterin (H_3C-H_4MPT), and the methyl group is transferred from H_4MPT to coenzyme M (HS-CoM) to form methyl-coenzyme M (CH_3-S-CoM) which is reduced to methane by the action of MCR (See Scheme 2), using coenzyme B (HS-CoB) as the electron donor[13].

For hydrogenotrophic organisms, CO_2 is utilized as a substrate with H_2 as an electron donor. CO_2 is first reduced and then combines with methanofuran (MFR) to form formylmethanofuran (CHO-MFR)[14]. Next, the formyl group is transferred to H_4MPT from MFR to produce methylene H_4MPT

(HC≡H_4MPT). HC≡H_4MPT is subsequently reduced twice to H_3C-H_4MPT. For methylotrophic organisms, methylated compounds such as methanol, methylthiols and methylamines are used as substrates[15]. By the action of specific methyltransferase complexes, methyl groups from the substrates are transferred first to a corrinoid protein and then to HS-CoM to form CH_3-S-CoM[15].

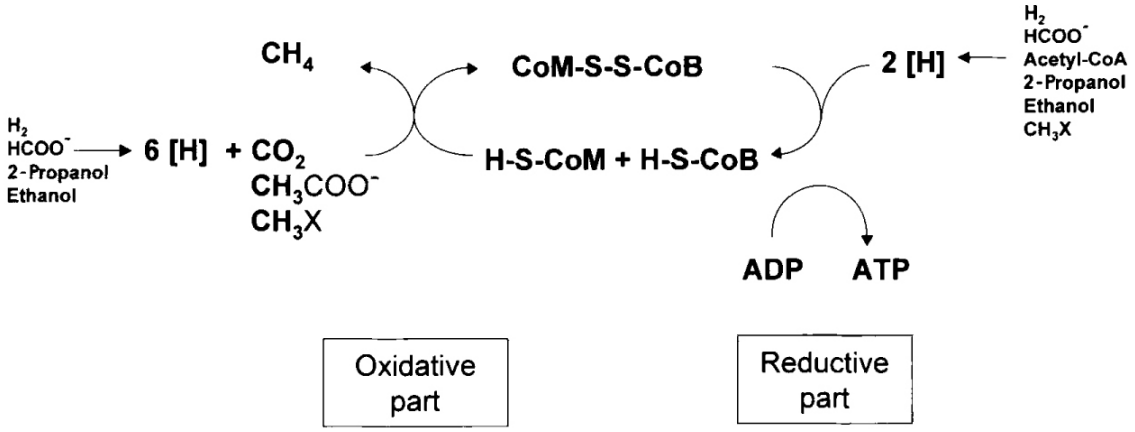
The last step in methanogenesis is the reaction between methyl-coenzyme M and coenzyme B, which results in the formation of methane and the heterodisulfide (CoM-S-S-CoB). This last step is catalyzed by the enzyme methyl-coenzyme M reductase (MCR).



Scheme 2: The process of methanogenesis[16]

1.1.3 Energy Metabolism in methanogenesis

Methanogenesis is essentially a respiratory process for methanogens, leading to the generation of ATP[5]. The energy metabolism in methanogens is divided into two parts: the oxidative part and the reductive part (Scheme 3). During the oxidative part, methyl-coenzyme M and coenzyme B are oxidized to the heterodisulfide CoM-S-S-CoB, with the release of methane[8,9]. This reaction is exergonic and is not coupled with energy conservation or ATP formation, as such the energy requirement for the growth of methanogens comes from the reductive part[17]. In the reductive part of methanogenesis, the heterodisulfide is reduced by heterodisulfide reductase (HDR) to regenerate coenzyme M and coenzyme B. This process is coupled with ADP phosphorylation and proton pumping through the hydrogenase-HDR complex. Electrons required by HDR for the reductive part is supplied by hydrogen oxidation. The free energy change associated with the heterodisulfide reduction sufficiently drives the phosphorylation of 1 mol of ADP (+31.8 kJ/mol)[5].



Scheme 3: The oxidative and reductive parts of methanogenesis[5]

1.2 Methanogenic archaea

The research by Stephenson and Strickland[6] also led to a marked increase in the study of the organisms that perform the metabolic process of methanogenesis. These organisms were initially thought to be bacteria, but they have been found to be members of a separate phylogenetic domain called archaea and were subsequently called methanogenic archaea[5]. Methanogens have been classified into seven orders, which are as follows: *Methanobacteriales*, *Methanococcales*, *Methanomicrobiales*, *Methanopyrales*, *Methanosarcinales*, *Methanocellales*, and *Methanomassiliicoccales*[18,19] . Methanogens are the only known organisms that produce methane as a catabolic end-product[5]. Of all seven orders, only *Methanosarcinales* can metabolize acetate as a carbon source to produce methane[11].

Methanogens are found in a wider range of anoxic environments[20], including wetlands, rice fields and peat bogs, for instance, *Methanoregula boonie* was isolated from an acidic peat bog[21]. They are also found in the digestive tract of herbivores mammals, for example *Methanobrevibacter ruminantium*[3].

Some others are extremophiles found in hot springs, glacial ice, hydrothermal vents, and saline lakes, for instance, *Methanohalophilus zhilinae* was cultured from an Egyptian saline lake[22]. They are mostly obligate anaerobes but species such as *Methanotrix paradoxum, Methanosarcina acetivorans*, and *Methanosarcina barkeri* have been observed to function within aerobic environments[13]. Studies of the process of methanogenesis have been performed using a wide range of methanogenic archaea such as *Methanococcus maripaludis*, *Methanosarcina barkeri*, *Methanosarcina acetivorans*, *Methanolobus tindarius*, *Methanopyrus kandleri*, *Methanocaldococcus janaschii*, *Methanothermobacter marbugensis*, *Methanothermococcus thermolithotrophicus* and a host of others[13,23].

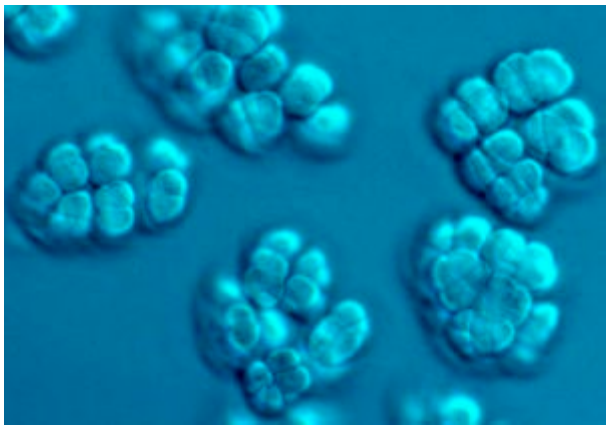


Figure 1: Electron micrograph of Methanosarcina acetivorans[24]

The *Methanosarcina spp.* seems to be more widely utilized for studies for a number of reasons: the organism uses all three methanogenesis pathways and has one of the largest known archaeal genomes[16]; all genes related to methanogenesis are conserved in the organism and the complete genome of *Methanosarcina acetivorans* (Figure 1) has been completely sequenced[16]; it is able to use acetate as a substrate[25], and acetate is produced cheaply and abundantly in nature[26]. This makes its cultivation in a laboratory setting relatively easier. The organism grows well at 37 °C, with its enzymes active at that temperature[27] and *Methanosarcina acetivorans* is also one of only two methanogens with established genetic systems. The other organism is *Methanococcus maripaludis.* This makes it possible to carry out processes such as knockout of genes and overexpression of proteins using these strains of methanogens.

1.3 Methyl-coenzyme M reductase structure and function

Initial studies by Ellefson and Wolfe[28] characterized MCR as a yellow protein consisting of three subunits arranged in an $\alpha_2\beta_2\gamma_2$ configuration with a molecular mass of approximately 300 kDa (Figure 2). Subsequent studies have revealed that the enzyme also contains a non-covalently bound coenzyme F430, which is a Ni-corphinoid required for enzyme activity[29,30]. In 1997, MCR from *Methanosarcina acetivoran* was successfully crystallized by Ermler *et al*[30]. The crystal structure confirmed the presence of the heterohexameric subunits, the coenzyme F430, and several post-translational modifications (PTMs)[30]. It also showed that the coenzyme F430 along with coenzyme M and B were embedded into a 30 Å long narrow channel extending into the interior of the protein from the protein surface, with the front face of the tetrapyrrole plane pointing towards the mouth of the channel and the rear face pointing towards the apex of the channel (Figure 3).

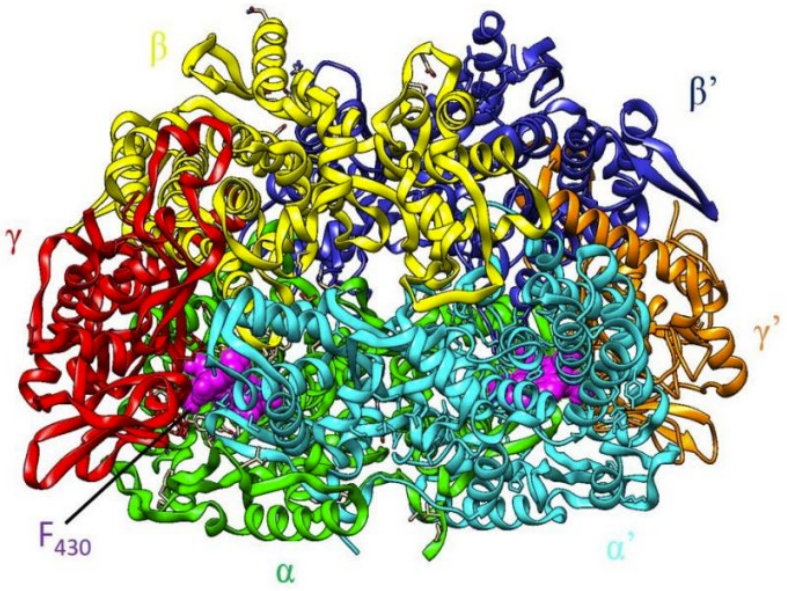


Figure 2: Structure of MCR from Methanosarcina acetivoran[30]

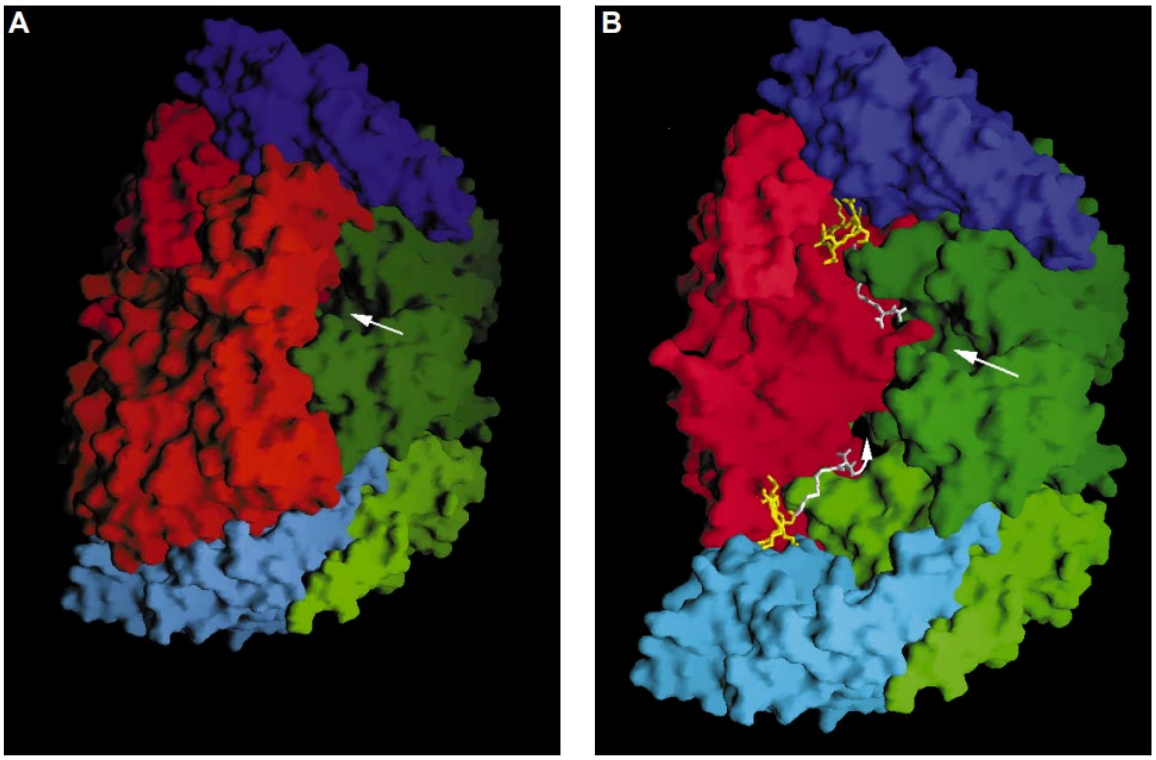


Figure 3: (a) Molecular surface representation of MCR showing the entrance of one of the channels. (b) Molecular surface of MCR without one α subunit, showing a view into the interior of the channel[30]

Several unprecedented post-translational modifications (PTMs) are found on the enzyme, some of which are conserved[31,32]. On the α-subunit of MCR is found 1-*N*-methylhistidine and thioglycine, which are conserved in all MCRs examined to date[33], as well as *S*-methylcysteine, 5-(*S*)-methylarginine, didehydroaspartate, and 2-(*S*)-methylglutamine present in some species (Figure 4). The β and γ-subunits have no post-translationally modified residues.

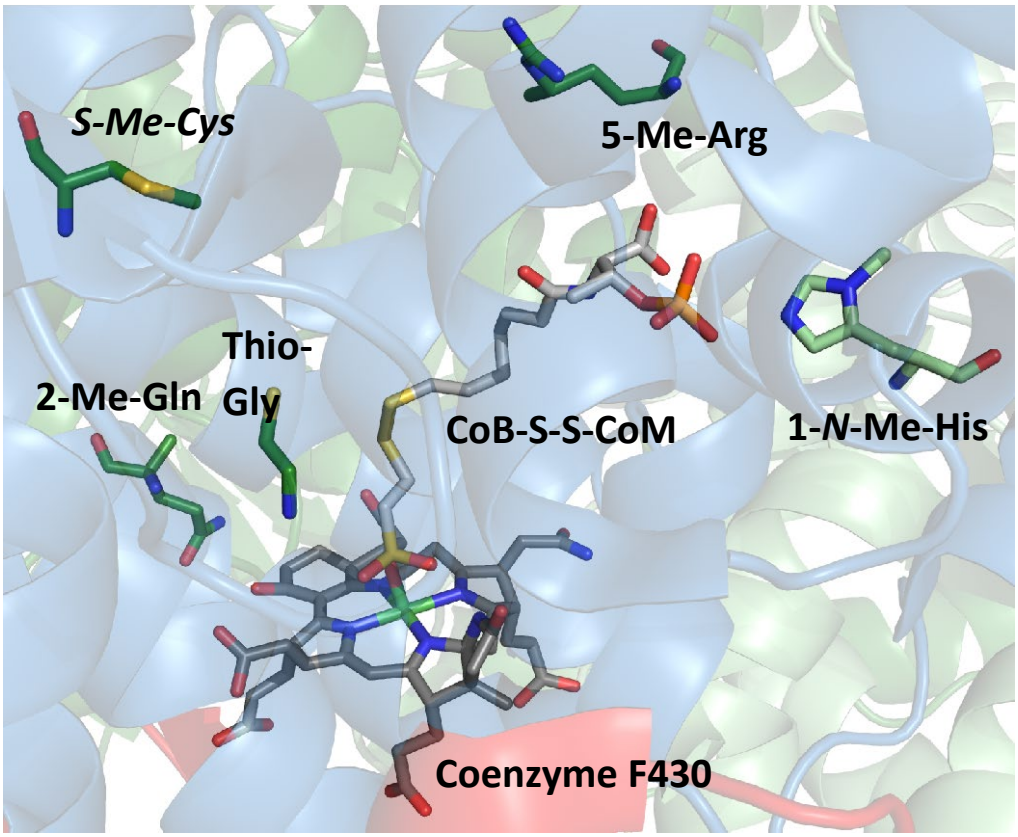


Figure 4: MCR active site highlighting the post-translational modifications, coenzyme F430, and the CoM-CoB disulfide[10]

The enzyme has also been found to exist in active and inactive forms depending on the oxidation state of the Ni center in coenzyme F430, which can be distinguished using electron paramagnetic resonance (EPR) spectroscopy. MCR from freshly harvested cells exhibits the inactive EPR silent state (MCR_{silent}) and is activated by reduction. The reduced MCR (MCR_{red1}) is the catalytically active enzyme and has a characteristic Ni(I) F_{430} EPR spectrum. Loss of enzyme activity occurs within a few hours (even under strictly anaerobic conditions) resulting in an inactive EPR-silent Ni(II) state of the enzyme[30]. An MCR_{ox1} state also exists which has a substantially different EPR spectrum than the MCR_{red1} state. It is usually obtained as a product of cells being gassed with CO_2-

N_2 before harvesting. The MCR_{ox1} state can be reduced with titanium citrate into the active MCR_{red1} state and can also revert into the MCR_{ox1}-silent state over time[30].

A better understanding of the assembly and activation process of MCR will be pivotal in future studies and discoveries bordering on the regulation of methane by environmental and industrial institutions. Studies into the *in vivo* assembly of MCR in a heterologous host will necessitate combining the right set of MCR genes in the heterologous host, and will facilitate the understanding of the genomic and environmental requirements for MCR assembly and maturation. This knowledge will find application in the development of engineered strains of organisms for methane production and/or conversion.

In chapter two, I describe the steps taken to express MCR in a heterologous non-methanogenic host. Putative *mcr* genes cloned from *Methanosarcina acetivoran* genome were introduced into *Escherichia coli* singly or in combination with other genes suspected to play a role in MCR maturation and assembly.

With the sequencing of the complete genome of *M. acetivorans*[16], several methanogenesis markers in the genome of the organism were identified. Methanogenesis markers are so named because they are genes that were identified as being unique to and conserved in all the initial genomes of methanogens. They include *mm* genes 1 – 17. The *A2* gene was also designated as a marker of methanogenesis, along with the *mcr* genes and is a part of the MCR activation complex[34]. *Mm7* is also a part of the MCR activation complex.

Many of the methanogenesis markers have been shown to be important for the function of MCR: *mm1* or the *YcaO homolog* (thioglycine), *mm10* or *mamA* (methylarginine), *mm13* (cfbD, F430 biosynthesis). *Mm3* is homologous to a peptidylprolyl cis-trans isomerase. *Mm2* is found

sometimes clustered with *mm3* (in *Methanopyrus kandleri*) or with *mm4* (in *Methanoculleus bourgensis*) and is annotated as an AIR synthase. The actual functions of the other methanogenesis markers are yet to be confirmed. Some other genes are known to act in association with methanogenic markers such as *tfuA* (with *mm1*); or suspected to be implicated in MCR maturation and assembly such as *thiI*. *ThiI* is a sulfur carrier protein suspected to be one of the sulfur donors for the thioglycine PTM. These genes, along with the coenzyme F430 (*cfbABCDE*) genes and the *mcrABCDG* genes were cloned into *E. coli* for studies into the heterologous expression of MCR.

1.4 Reverse methanogenesis

Considering the role of MCR both in methane production and consumption, it is not surprising to see the plethora of studies that have been carried out on the enzyme[13,16,30,32,35]. Reactive organic matter buried in soil is converted to methane (about 10 – 20%), and the annual rate of methanogenesis from the oceans is about 75 – 100 billion kg of methane per year[36,37]. However, 90% of methane produced in oceans is consumed by a group of organisms known as methanotrophs, which are found in sulfur-penetrated zones[38].

Homologs of MCR have been discovered in anaerobic methanotrophic archaea (ANME). MCR therefore plays a strategic role both in methane production and methane consumption in methanogenic and methanotrophic organisms, respectively. Methanotrophic archaea catalyze the anaerobic oxidation of methane (AOM), which operates as the reverse of methanogenesis[38]. Here, methane is oxidized with CoB-S-S-CoM as the first step of the pathway[39]. Methane oxidation is coupled to the reduction of electron acceptors such as nitrate, nitrite, Fe(III)[40,41]. More than 70 billion kg of methane is purported to be consumed annually through the process of AOM[42].

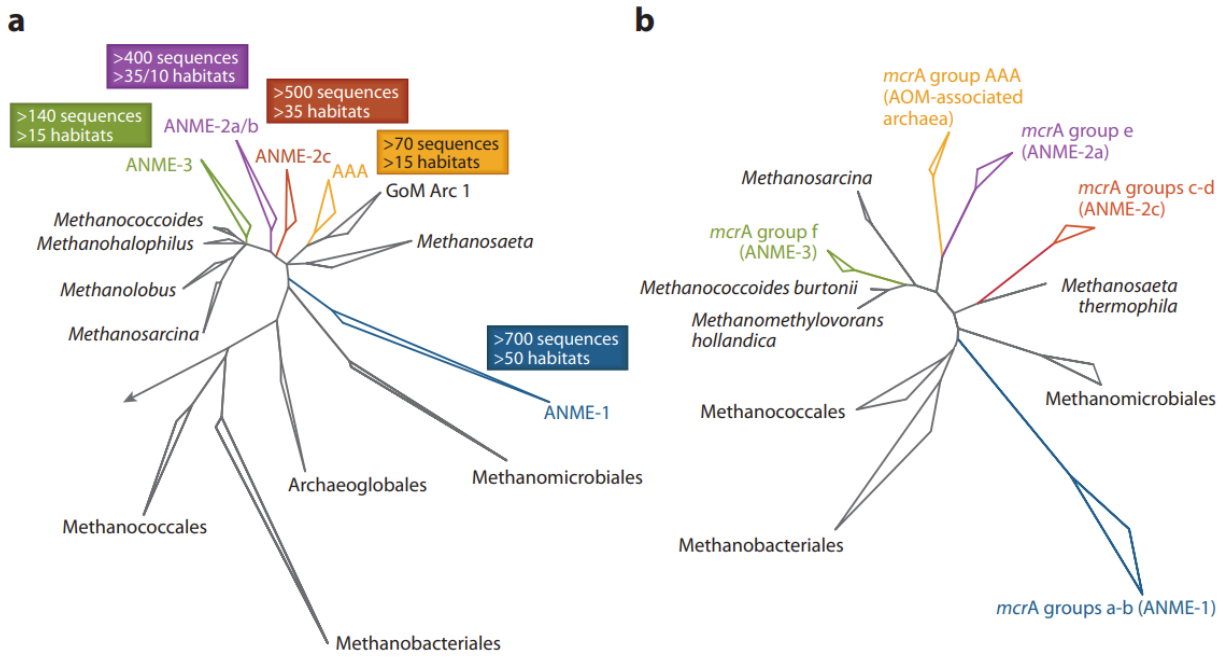


Figure 5: Phylogenetic trees of methanogenic and anaerobic methanotrophic archaea constructed using 16S rRNA (b) and the gene coding for the alpha subunit of methyl-coenzyme M reductase (mcrA)[38]

Three distinct clusters of ANME have been identified as follows: ANME-1, ANME-2, and ANME-3. ANME-2 is further divided into subgroups ANME-2a, -2b, -2c, and 2d (Figure 5). ANME-2d is also known as AOM-associated archaea (AAA). All clusters of ANME consume methane using the reverse methanogenesis pathway.

The existence of ANME was first established by Hinrichs et al [43], when they analyzed archaeal lipid biomarkers from sediments obtained from a methane seep in the Eel River Basin. Control sample was obtained from an area not near any active seep. ^{13}C Methane was bubbled into the water before extraction of organic matter. Organic matter extracted from the seep sediments

contained less ^{13}C than the control sample, pointing to methane oxidation by the organic matter. Comparative genomics showed that the rRNA of the organism constituted a novel clade, shown to be related to methanogens, and which the authors referred to as anaerobic methanotrophs[43].

ANMEs are usually found in consortiums with sulfate-reducing bacteria (SRB)[44]. Analysis of sediments from the Eel River Basin, Hydrate Ridge, and the Black Sea revealed that about 90% of the total microbial constitution was composed of archaea and sulfur reducing bacteria (SRB)[45–47].

Both ANMEs and their partner SRBs (for example, *Desulfosarcina spp.*, associated with ANME-1, *Desulfococcus spp.* associated with ANME-2, and *Desulfobulbus spp.* associated with ANME-3) are obligate anaerobes[48]. ANMEs are mostly found in the sulfate-methane transition zone (SMTZ) in the seabed[49]. SMTZs are found in all anoxic aquatic systems. Within this region, methane that is supplied from the reactive organic matter below the seafloor is consumed by ANMEs, while sulfate is supplied from the SRBs to provide the right niche for the anaerobic methanotroph[49].

1.5 Post-translational modifications of methyl-coenzyme M reductase

Crystal structures of MCR such as the one obtained from *Methanobacterium thermoautotrophicum* identified the presence of five modified amino acids in the active site region of the enzyme[30]. These modified amino acids include thioglycine, *S*-methylcysteine, 2-(*S*)-methylglutamine, 1-*N*-methylhistidine, and 5-(*S*)-methylarginine and were all found in the McrA subunit of the protein[30]. The didehydroaspartate modification was first reported by Wagner et al.[23], in a study analyzing the X-ray crystal structure of MCRs obtained from several methanogens. The modification was present in MCR from *Methanothermobacter marburgensis* and *Methanosarcina barkeri*. It was found adjacent to the thioglycine modification. The modification was absent in MCR from *Methanothermobacter wolfeii*.

Post-translational modifications that are not conserved across all species of methanogens and methanotrophs indicate that these modifications may be dispensable and not integral to the function and structure of the enzyme[23].

Available evidence obtained by labeling with L-(methyl-D_3)-methionine shows that the methylation modifications are post-translationally introduced by specific *S*-adenosylmethionine-dependent methylases[50], except for thioglycine and didehydroaspartate. McrA from methanotrophic archaea of the ANME-1 cluster possess a valine codon in place of the glutamine codon found in methanogenic archaea[51]. Most importantly, ANME McrA have a different subset of post translational modifications (see Figure 6). MCR structure obtained from ANME-1 revealed the presence of the 1-*N*-methylhistidine and the thioglycine modifications, as well as additional post-translational modifications such as 7-hydroxytryptophan, and *S*-oxymethionine[52] (Figure 6). It also featured a distinctly modified 17^2-methylthio-F430.

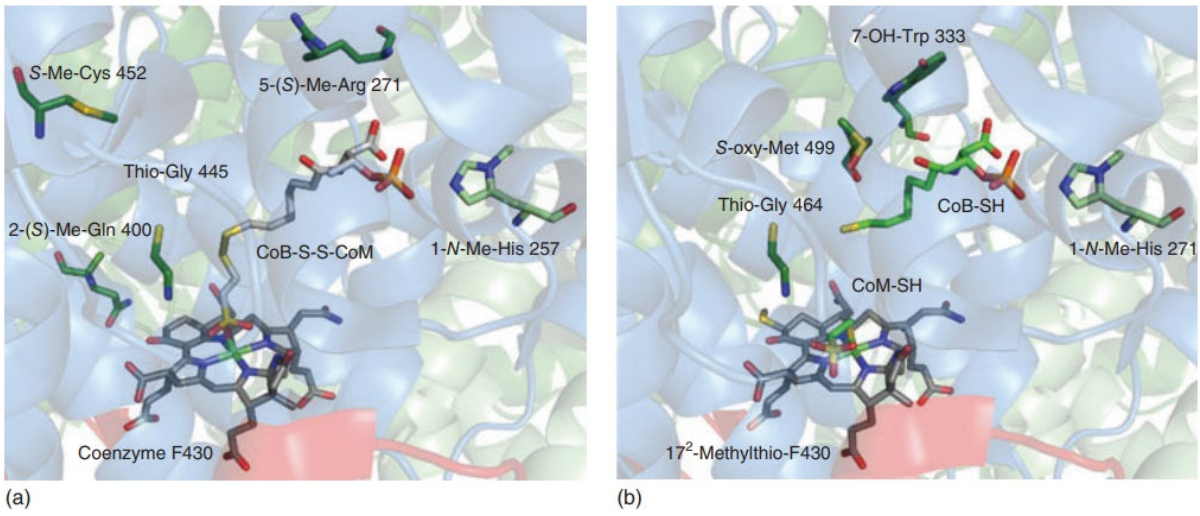


Figure 6: Post-translational modifications found in the MCR alpha subunit of (a) methanogens and (b) ANMEs MCR[10]

Work from Kahnt et al.[32], studying one representative of each of the five orders of methanogenic archaea and from two methanotrophic archaea revealed that the methylhistidine and the thioglycine modification was present in the MCR from all organisms studied including *Methanothermobacter marburgensis*, *Methanocaldococcus jannaschii*, *Methanopyrus kandleri*, and *Methanosarcina barkeri*, while the methylglutamine modification was absent in *Methanosarcina barkeri*. Also, the hyperthermophilic organisms *Methanocaldococcus jannaschii* and *Methanopyrus kandleri* did not feature the methylcysteine modification[32].

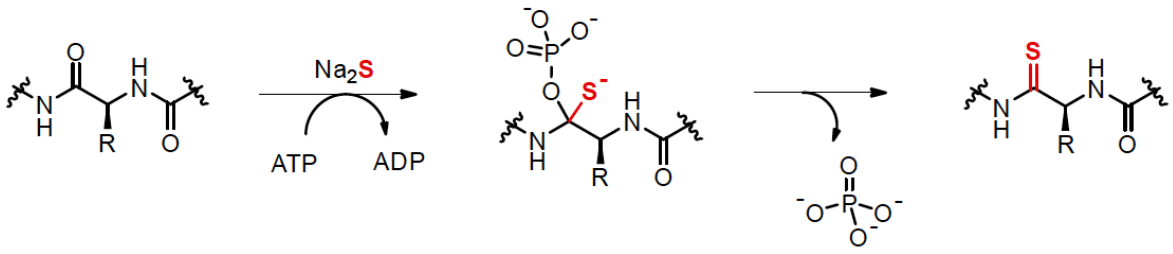
1.5.1 Thioglycine modification

Using genomic comparison and gene deletion strategies, the Metcalf group confirmed that a *tfuA*-associated *ycaO* homolog (methanogenesis marker 1) was responsible for the thioglycine PTM in the MCR from *Methanosarcina acetivorans*[53]. YcaO enzymes catalyze ATP-dependent backbone cyclodehydration of residues such as cys, ser and thr[54]. Mutant strains lacking these genes produced MCR with glycine lacking the modification. In comparison to the wildtype (WT), mutants did not grow at 45 °C when grown on methanol. Mutants also exhibited a low growth rate characterized by low cell yields and longer generation times when grown at 36 °C in acetate and dimethylsulfide (DMS). At lower temperatures (29 °C and 36 °C), mutant strains supplemented with methanol and trimethylamine did not record any significant decrease in cell yields, but experienced growth defects when grown at 39 °C and 42 °C[53].

The authors also investigated the influence of the thioglycine modification on MCR stability[53]. The melting point of MCR from the WT and the mutant strain was determined using the SYPRO Orange-based Thermofluor assay. WT MCR had a melting temperature of 67.3 $\pm$ 0.3 °C while the mutant had a melting temperature of 69.1 $\pm$ 0.2 °C, and it was consequently determined that the

thioglycine modification did not have any effect on MCR global stability but may influence the local stability within the active site region where it is buried[53].

The Mitchel group also investigated the role of YcaO and TfuA proteins on the thioamidation of McrA[54]. Due to the instability of the McrA protein, McrA peptides of 11-, 21- and 51- amino acid length was treated with YcaO and TfuA proteins in the presence of Na_2S and ATP. The [M + 16] ion expected for the thioamidation was detected using matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI_TOF_MS) for all treated peptides[54]. Tandem High Resolution Electrospray ionization mass spectrometry (HR-ESI-MS) localized the site of the thioamidation to Gly 465. They proposed a phosphorylation mechanism for the thioamidation of glycine shown in Scheme 4.



Scheme 4: Proposed mechanism for the thioamidation of glycine[54]

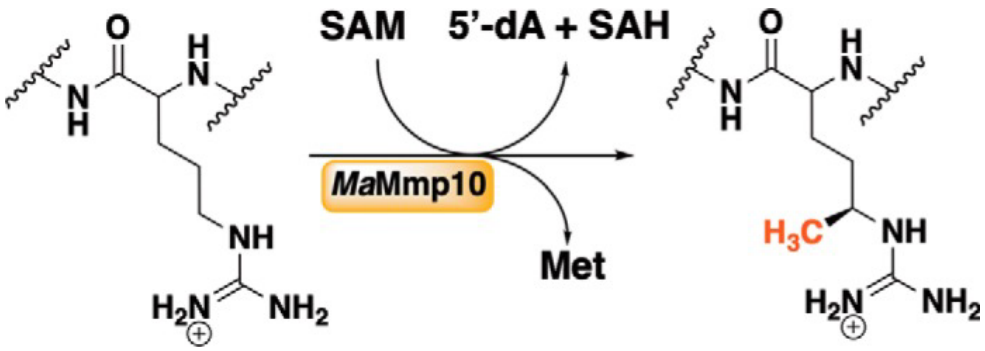
1.5.2 5-(*S*)-methylarginine modification

The role of the methanogenesis marker 10 (*mmp10*) gene (named *mamA* by the authors) in the 5-(*S*)-methylarginine modification was confirmed *in vitro* by Squire Booker and his group[33]. Prior to this, the Whitman and his colleagues had identified the mmp10 gene as the gene responsible for the modification[55]. The group deleted the mmp10 gene from the genome of *Methanococcus maripaludis*. Subsequent expression of McrA from the mutant strain and analysis using liquid

chromatography mass spectrometry (LC-MS) revealed the loss of the methylarginine modification (R275)[55]. Mutant strains of the organism also exhibited a 40 % to 60 % reduction in the rate of methane formation by whole cells.

The Booker group determined that the *mmp10* gene is adjacent to the *mcr* gene cluster and encodes a radical *S*-adenosyl-L-methionine (SAM) enzyme[33]. Radical SAM (RS) enzymes are the only family of enzymes capable of unactivated C-H bond methylation[56]. They function by the means of a $[4Fe\text{-}4S]^{1+}$ cluster coordinated to the enzyme by three conserved cysteine residues in the RS domain, and with S-adenosyl methionine (SAM) as the fourth ligand to the iron site[57]. SAM is cleaved to 5'-deoxyadenosyl 5'-radical (5'-dA) which abstracts a hydrogen atom from the substrate to initiate catalysis. They catalyze an array of reactions using substrate derived radical intermediates[58,59]. Methylases are a subgroup of the RS enzymes, and they are classified into five classes (A, B, C, D and E)[60,61]. Class B RS methylases use methylcobalamin (MeCbl) to donate a methyl group to the substrate and are able to carry out methylation reactions on sp2- and sp3-hybridized carbon centers[33]. Their domain architecture consists of a cobalamin-binding domain located next to the RS domain at the N-terminal.

The Booker group was able to show that Mmp10 protein is able to utilize cobalamin as the methyl donor, but its domain architecture does not feature the N-terminal cobalamin-binding domain characteristic of class B RS methylases[33]. Thus, a new subclass of the class B RS methylases was established by their study. In their studies, peptide substrates mimicking the primary structure of a section of MCR from *Methanosarcina acetivorans* containing the 5-(*S*)-methylarginine PTM were synthesized. *In vitro* assays with the peptides using purified Mmp10 were carried out. Results revealed methylation of the target residue on the peptide (Scheme 5) as well as the requirement of cobalamin and SAM for enzyme activity,



Scheme 5: Proposed mechanism for the methylarginine modification[33]

1.5.3 1-*S*-methylcysteine modification

In similar studies, Nayak et al. provided *in vivo* evidence that an *S*-adenosylmethionine (SAM)-dependent methyltransferase encoded by a *prmA* homolog (named *mcmA* gene by the authors) was responsible for the *S*-methylcysteine modification in *Methanosarcina acetivorans*[62]. The *mcmA* (*MA4545*) gene is adjacent to the *mcr* gene cluster in *M. acetivorans* and is conserved in organisms that have the methylcysteine PTM. Using cas9-based genome editing tools, mutant strains of *M. acetivorans* were created lacking the *MA4545* gene. McrA expressed from the mutant and subjected to MALDI-TOF-MS and tandem MS, was seen to lack the methylcysteine modification. The tryptic peptide containing the cys472 residue was 14 Da lighter than the corresponding peptide in the wildtype McrA. When grown on trimethylamine, a 12 % decrease in growth yield was observed and when grown on DMS, the mutant strain grew 30 % slower than the wildtype. This suggests that the modification might be important in the adaptation of the MCR to growth in mesophilic conditions. Interestingly, when grown on methanol, the mutant demonstrated a 20 % increase in growth rate at 42 °C compared to the wildtype.

While this study showed the requirement of the *MA4545* gene for the methylcysteine modification, it did not establish whether the *MA4545* gene alone was sufficient for the modification. It is possible that there might be other accessory factors necessary for the methylcysteine modification. Furthermore, there was no sufficient evidence to establish that the modification is post-translational rather than co-translational.

1.5.4 1-*N*-Methylhistidine

The 1-*N*-methylhistidine modification is present in all analyzed MCRs to date[32]. This may be an indication of the possibility that the modification plays an essential function in the enzyme. The modified residue is part of the coenzyme B binding site. Since the methylated residue has a slightly higher pKa when compared to the unmethylated residue (imidazole and N-methylimidazole differ by 0.1), Kahnt et al., postulates that the strength of the coenzyme B binding interaction is expected to be increased by this methylation[32]. It is also possible that the methylation results in an improved orientation of the histidine residue within the coenzyme B binding site pocket.

However, the gene(s) responsible for this methylation is yet to be discovered, impeding investigations into the potential impact of this modification on MCR structure and function.

Chapter three of this work details the investigations that were carried out into the methylhistidine and the methylcysteine modifications of MCR. The *MA4545* gene from *Methanosarcina acetivorans* was cloned into *Escherichia coli* cells and expressed under various conditions to determine optimum conditions for its activity. An McrA peptide (11-amino acid) mimicking the primary structure of the section of McrA containing the cys472 residue was used as a substrate for the assay. The *MA4545* gene along with his-tagged *mcrA* gene were also cloned into *E. coli* cells with subsequent expression and purification of the McrA protein for MALDI-TOF-MS analysis.

The *mtxX* (*mm4*) gene present in the genome of *Methanosarcina acetivoran* and annotated as a methyltransferase, was chosen as a possible gene candidate for the methylhistidine modification. The gene was also cloned into *E. coli* and the purified protein was used for methylation assays using an McrA peptide (11-amino acid length) mimicking the primary structure of the McrA section containing the his271 residue as the substrate.

Concerns about the enzyme interaction with the short peptide substrates necessitated the use of structural prediction analysis to determine whether the target proteins were interacting with their respective peptides.

1.6 Coenzyme F430

Each of the active sites of MCR binds to coenzyme F430, which is a unique nickel tetrahydrocorphin. Coenzyme F430 contains two exocyclic rings: a carbocyclic F ring as well as a γ-lactam E ring[10]. Of all tetrapyrroles found in nature, coenzyme F430 is the most highly reduced.

The active form of MCR contains an F430 that is in the Ni(I) oxidation state which is efficient in facilitating the catalysis of the MCR-catalyzed reaction. MCR catalysis is postulated to occur by the formation of Ni(II) thiolate and methyl radical intermediates[63].

While the mechanism of MCR catalysis has not been fully understood, three mechanisms have been proposed (Figure 7). In mechanism I, the nucleophile Ni(I) attacks the methyl group of methyl-CoM, leading to the formation of methyl-Ni(III) which receives an electron from coenzyme M to generate methyl-Ni(II) and a coenzyme M thiyl radical[64]. To form methane, the methyl group of methyl Ni(II) abstracts a proton from coenzyme B. To maintain the cycle, Ni(I) needs to be regenerated. To facilitate this, coenzyme M thiyl radical undergoes a reaction with CoB to form

the disulfide radical CoBS-SCoM$^{\bullet-}$[65]. This radical reduces Ni(II) to Ni(I) through electron transfer. The inhibition of MCR by 3-bromopropanesulfonate (BPS) supports this mechanism[66,67].

In mechanism II, the sulfur atom of methyl-S-CoM is attacked by the nucleophile Ni(I). This leads to the cleavage of the methyl-sulfur bond and the generation of a methyl radical, as well as a Ni(II)-S-CoM complex. The methyl-radical attacks HS-CoB, abstracting a hydrogen atom and generating methane[68–70], while the thiyl radical reacts with Ni(II)-S-CoM complex to form CoBS-SCoM$^{\bullet-}$ which transfers an electron to Ni(II) to generate Ni(I)[65].

Arguable limitations to the model is the long distance between the thiol group of HS-CoB and the methyl-SCoM. However, density functional theory (DFT) calculations, transient kinetics, magnetic circular dichroism (MCD) methods and rapid freeze-quench (RFQ) EPR corroborate this mechanism[23,64].

In the last mechanism (III), the reaction starts with a nucleophilic attack on the sulfur of methyl-CoM by Ni (I). This generates the Ni(III)-SCoM complex and a methyl anion. The methyl anion is protonated by HS-CoB to form methane and a CoBS^{-} anion[65]. The basis for this mechanism is the generation and characterization of the Ni(III)-SR intermediate by the Ragsdale group[65].

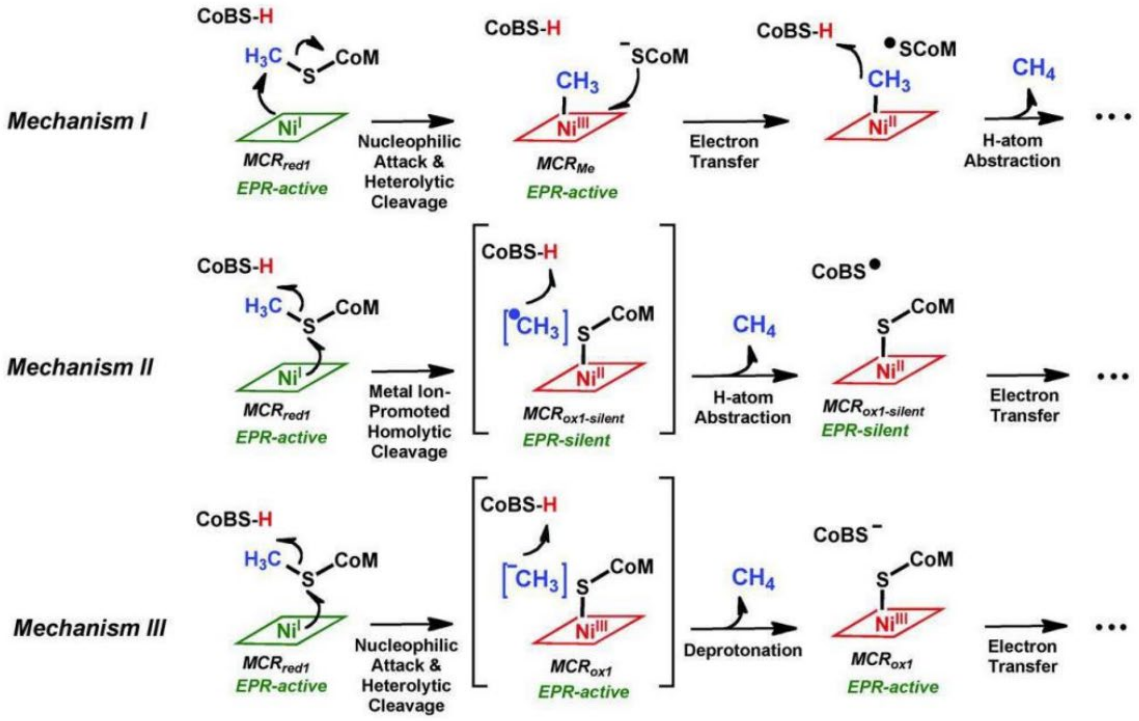
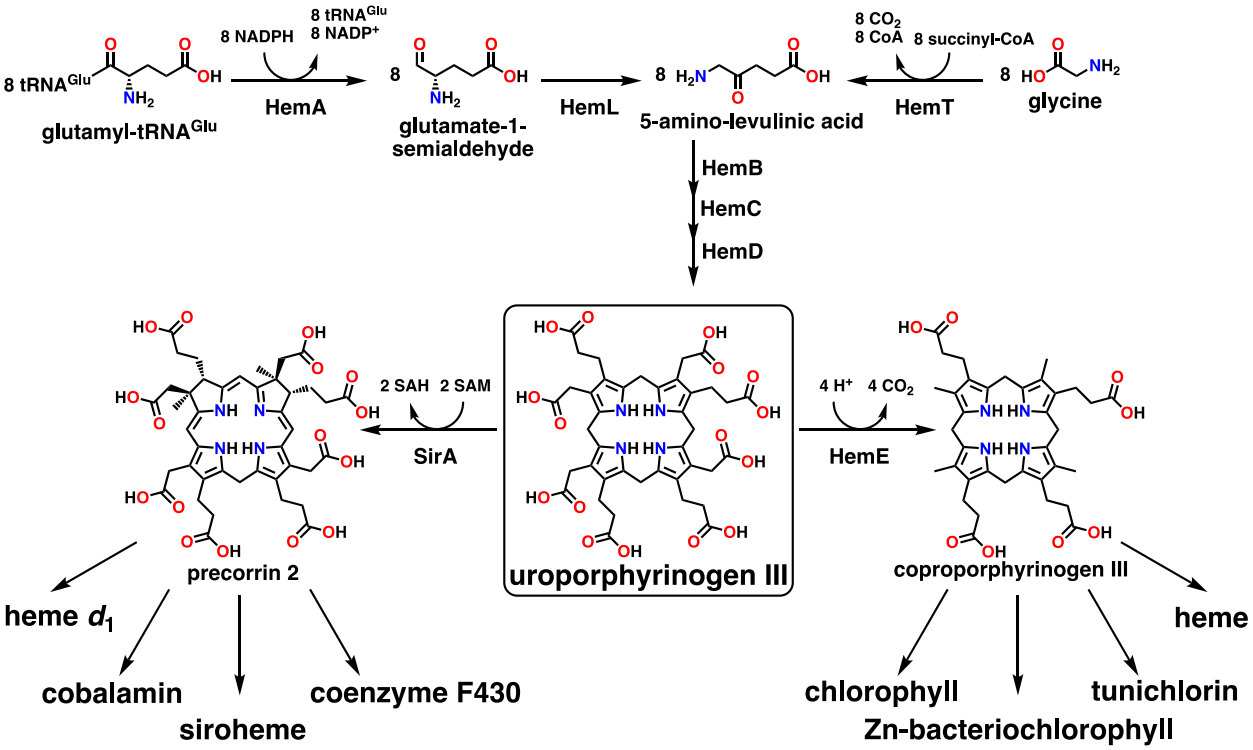


Figure 7: Proposed mechanisms for the initial steps of MCR catalysis[65]

1.6.1 Biosynthesis of tetrapyrroles

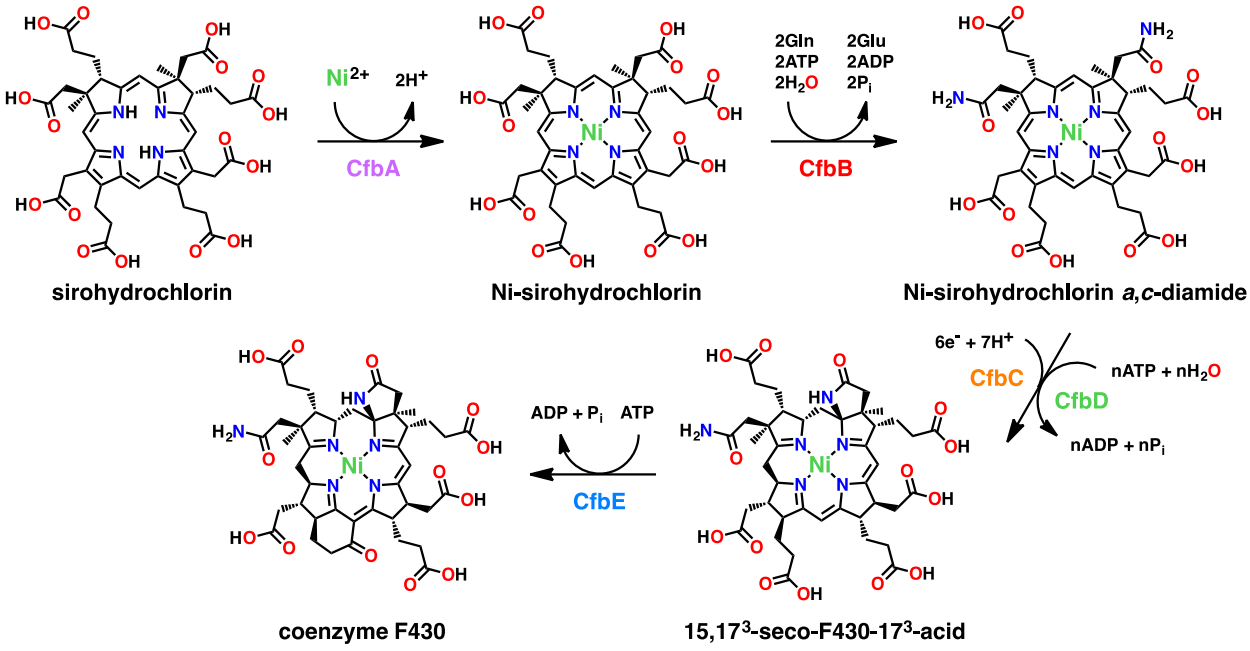
All tetrapyrroles are biosynthesized using either glycine or L-glutamate as the starting material. Both biosynthetic pathways lead to the formation of 5-aminolevulinate (ALA), which proceeds to the formation of uroporphyrinogen III, the last metabolite common to all tetrapyrroles[10].

Tetrapyrroles such as cobalamin and siroheme are derived from uroporphyrinogen III after it is methylated at C2 and C7 to produce precorrin 2, while tetrapyrroles such as heme or chlorophyll form from uroporphyrinogen III after it is decarboxylated to coproporphyrinogen III (Scheme 6). Coenzyme F430 biosynthesis follows the methylated tetrapyrrole branch and includes precorrin 2 as a biosynthetic pathway intermediate.



Scheme 6: The existing pathways for tetrapyrrole biosynthesis[10]

The biosynthetic pathway of coenzyme F430 was demonstrated *in vitro* by the Mansoorabadi group[71], after the group had identified the coenzyme F430 (*cfbABCDE*) genes using comparative genomics.



Scheme 7: In vitro biosynthetic pathway for coenzyme F430[71]

The group identified the genes responsible for coenzyme F430 biosynthesis (referred to as *cfb* genes) and conducted *in vitro* characterization of the corresponding enzymes. These enzymes include CfbA, a sirohydrochlorin cobaltochelatase homolog that facilitates specific Ni-chelation of sirohydrochlorin; CfbB, a cobyrinic acid a,c-diamide synthase homolog that catalyzes the ATP-dependent amidation of Ni-sirohydrochlorin's a- and c-acetate side chains using glutamine as an ammonia source; CfbCD, a nitrogenase homolog that converts Ni-sirohydrochlorin a,c-diamide to 15,17^3-secoF430-17^3-acid; and CfbE, a Mur-ligase homolog responsible for the ATP-dependent cyclization of the g-proprionate side chain, leading to the formation of the carbocyclic F ring. The *cfb* genes from *M. acetivoran* were cloned and expressed heterologously in *E. coli* (Scheme 7). Using *in vitro* assays, the recombinant proteins were shown to be capable of synthesizing coenzyme F430 with sirohydrochlorin as the starting substrate[71].

1.6.2 Modified F430s

MCR from ANME-1 featured a distinctly modified 17^2-methylthio F430 (F430-2). Due to the fact that there is limited information on the biochemical and genomic information of ANME-1, it is unclear what gene or enzyme is responsible for the formation of 17^2-methylthio F430, though Allen and his colleagues[72], opines that an enzyme of the radical SAM superfamily may be implicated in the F430 methylation. However, the concurrent presence of coenzyme F430 in ANME-1 suggests that it is a precursor to 17^2-methylthio F430 in the MCR of ANME-1[73].

Other derivatives of coenzyme F430 have been found in cell-free extracts of methanogens and ANME (Figure 8). These include a derivative containing a 3-mercaptopropionate moiety, proposed to be bound as a cyclic thioether (F430-3), 12,13-didehydro F430 (F430-4), vinyl F430 (F430-5), vinyl-17^2-methylthio F430 (F430-6), 17^2-methylsulfoxide F430 (F430-7), 17^2-keto F430 (F430-8), 17^2-hydroxy F430 (F430-9), and 17^1-17^2-dihydroxy F430 (F430-10)[72]. While the biological significance and biosynthetic routes of these derivatives are not yet clear, the possibility of them playing significant roles relevant to MCR function cannot be ruled out.

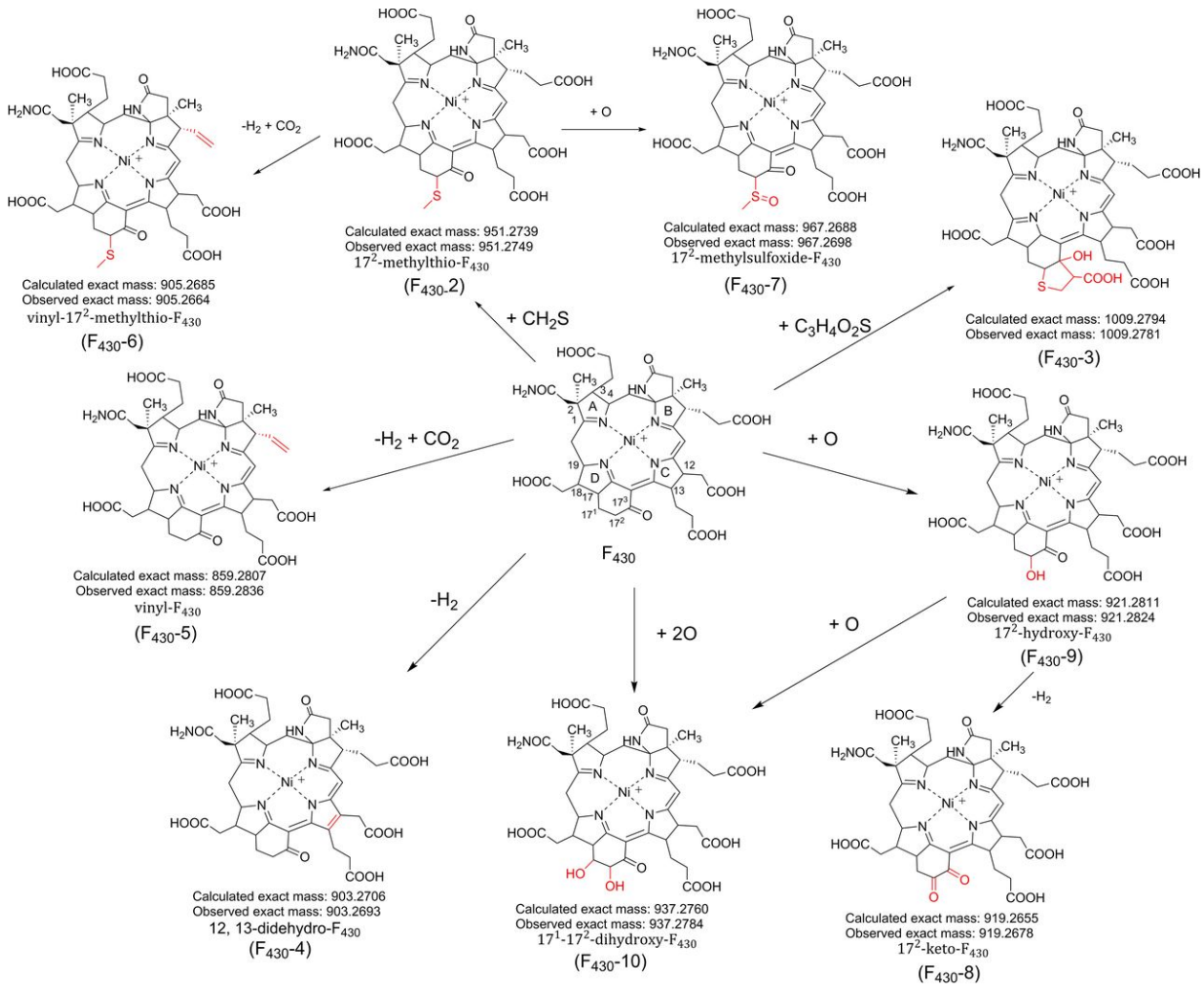


Figure 8: Proposed structures for Modified coenzyme F430 variants and their chemical relationship to F430[72]

In a recent study by the Wagner group[74], the crystal structure of an MCR homolog specific for ethane-activation was reported. The enzyme was found in the anaerobic archaea, *Candidatus Ethanoperedens thermophilim*, which is usually found in consortium with the sulfate-reducing bacterium *Candidatus Desulfofervidus auxilii*. The enzyme crystal structure resolved and refined to 0.99 Å revealed the enzyme to be structurally similar to MCR and was subsequently called ECR

(Ethyl-coenzyme M Reductase). However, the protein sequence in the MCR found in the ethanotroph featured some large insertions and substitutions of some canonical residues, making the enzyme different from methanogenic MCR. Due to the insertions α10-18, α90-111, α330-340, β70-95, and γ227-240, the MCR from *Ca. E. thermophilim* was found to be 20 kDa larger than the MCR from *M. marburgensis*.

The *Ca. E. thermophilim* MCR was also found to contain a 17,17^2-dimethyl coenzyme F430 (Figure 9). This wider chamber containing the dimethylated F430 is purported to be essential in adapting the enzyme to the reception of a two-carbon substrate (ethane)[74].

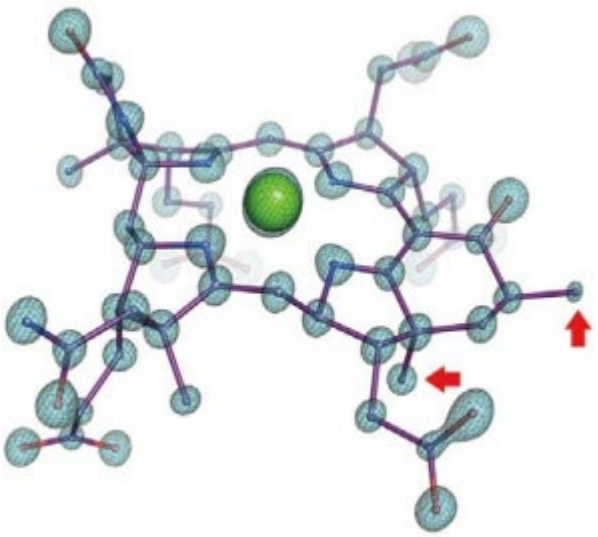


Figure 9: Structure of dimethyl coenzyme F430[74]

The biological significance of modified coenzyme F430 is yet unclear. Hahn et al., in reporting the crystal structure of ECR opine that the dimethylated F430 may be an adaptation of the enzyme to a two-carbon substrate[74]. While MCR from *M. marburgensis* can generate ethane from ethyl-CoM

and HS-CoB, the activity is much slower compared to its methane generation ability[75]. This informed the hypothesis that the affinity of MCR homologues to specific substrates may be impacted by the coenzyme F430 present in the enzyme.

To test this hypothesis, we carried out molecular dynamic (MD) simulations and distance calculations on MCR homologues including *M. acetivoran* MCR featuring regular F430, ANME-1 MCR featuring the 17^2-methylthio F430 and ECR featuring 17,17^2-dimethyl F430, in the presence and absence of the post-translationally modified residues. MD simulations were also carried out on ANME-1 MCR and ECR containing regular F430. The goal was to determine the impact of the F430 homologues on the enzyme active site in the presence of the substrate, and to see if the post translational modifications had any influence. Chapter four details the steps that were taken to test this hypothesis and the results of the simulations.

The intricate biochemistry surrounding MCR and its active sites, particularly its binding to coenzyme F430, is crucial for understanding its catalytic mechanisms. The multitude of derivatives of coenzyme F430 and their potential roles in MCR function add layers of complexity to this scientific discourse. Moreover, recent discoveries, like the structurally similar enzyme in *Ca. E. thermophilim* with its unique features, underscore the dynamism and diversity of these enzymes in nature. As scientific research continues to delve deeper into these mechanisms, it is anticipated that further nuances will be unraveled, paving the way for more comprehensive insights into MCR function, its derivatives, and their broader implications in the realm of biochemistry.

Chapter Two

The Heterologous Assembly of Methyl-coenzyme M Reductase (MCR) in a Non-methanogenic Host

2 Chapter 2

2.1 Introduction

Global warming has been a subject of scientific and political debate in recent times. It is defined as the increase in the global surface temperature of the earth attributed to the increased production of green-house gases[76]. These gases remain in the atmosphere and trap solar radiation that would otherwise reflect off the earth's surface[1]. It has been opined by NASA that each decade, the global surface temperature of the earth increases by 0.2 °C (0.36 °F) (Figure 10) and the forecast by the Intergovernmental Panel on Climate Change (IPCC) is that over the next century, there will be a temperature rise of up to 10 °F[77]. Proposed consequences of this rise in surface temperature of the earth include prolonged frost-free seasons, stronger and more intense hurricanes, sea level rise of up to 4 feet, more droughts and heat waves, change in precipitation patterns, and an ice-free artic[78].

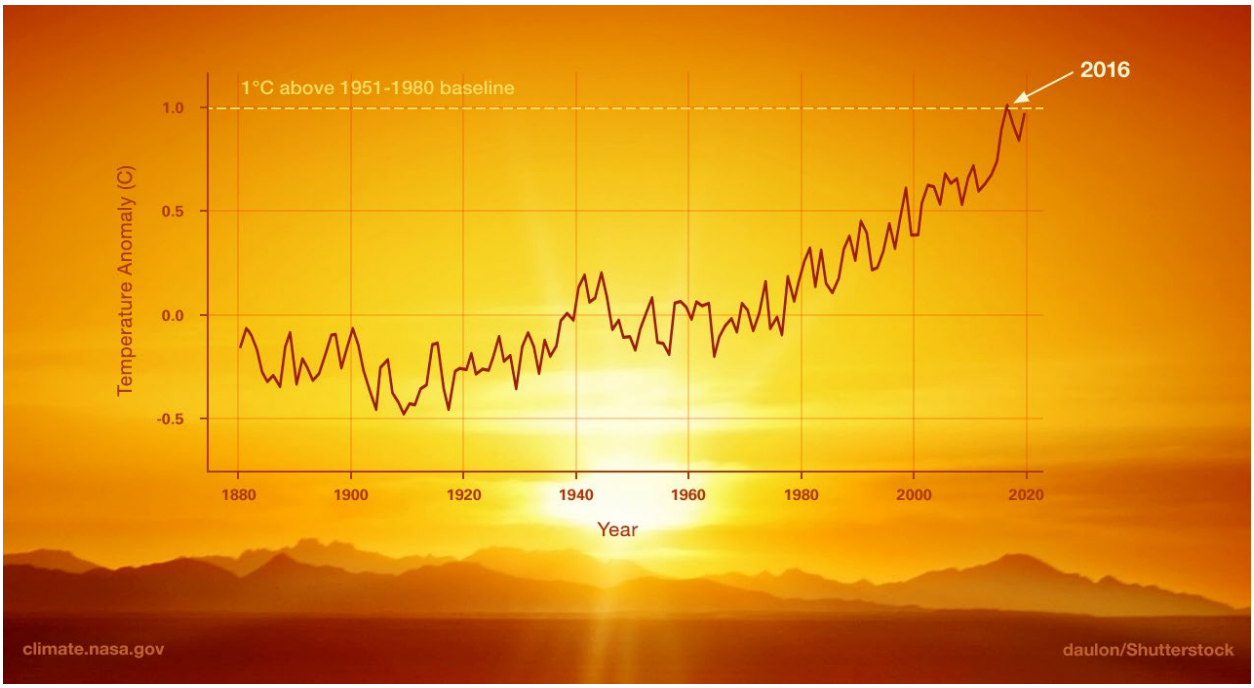


Figure 10: Change in global surface temperature relative to 1951-1980 average temperatures[77]

Examples of greenhouse gases include carbon monoxide (CO), carbon dioxide (CO_2), chlorofluorocarbons (CFCs), and methane (CH_4). About 10^9 tons of methane are produced each year by methanogenic archaea, part of which escapes into the atmosphere (20 % to 30 %) and contributes to greenhouse gases[4,5]. Two thirds of global methane emissions come from methanogenic archaea. Methane has approximately 30 times the global warming potential of CO_2, which makes it a potent gas worthy of study[79].

Methane is generated by methanogens through a process known as methanogenesis. In this pathway, various one-carbon (C1) compounds such as carbon dioxide, methanol, and methylamines serve as carbon sources for the production of methane. These substrates are transformed into an intermediate known as methyl-coenzyme M[35]. The final step in methanogenesis involves the enzyme methyl-coenzyme M reductase (MCR), which facilitates the conversion of methyl-coenzyme M into methane.

The MCR complex, a substantial 300 kDa metalloenzyme, comprises a heterohexamer structure consisting of three different pairs of subunits: two α, two β, and two γ[32]. This enzyme features two active sites, with each MCR trimer (α', β, γ) bound firmly to one equivalent of a prosthetic group called coenzyme F430 which is a nickel-containing tetrapyrrole that is required for MCR activity[29–31]. The Ni(I)-MCR is the active form of MCR. Several mechanisms have been proposed for MCR catalysis, all of which involves a nucleophilic attack by the Ni(I) on the MCR substrate[63].

Crystal structure of MCR from *Methanobacterium thermoautotrophicum*, revealed five distinct modified amino acids in the enzyme's active site[30]. The McrA subunit of the protein contained the following post-translationally modified amino acid residues: thioglycine, S-methylcysteine, 2-(S)-methylglutamine, 1-N-methylhistidine, and 5-(S)-methylarginine[30]. The requirement of these post translational modifications (PTMs) for MCR activity has not been established but based on studies

by the Metcalf group[53], the group hypothesized that the thioglycine modification may influence the local stability within the active site region of MCR where it is found.

At present, there is significant enthusiasm surrounding the heterologous expression of holo MCR. A successful heterologous assembly of MCR should result in the expression of an active, recombinant MCR. An active MCR is one that would have properly folded and assembled subunits, all its PTMs in the right positions, the coenzyme F430 properly inserted into the MCR structure and the Ni of F430 reduced to Ni(I) which is necessary for MCR activation. This process will entail detailing the complete set of genes required for the MCR genes (*mcrABG*), the coenzyme F430 genes (*cfbABCDE*) as outlined by the Mansoorabadi group[80], the post-translational modification genes and other genes necessary for the folding of MCR subunits and the insertion of coenzyme F430 into the MCR complex. The *mcrCD* genes are implicated in F430 insertion into the MCR complex, McrC has been shown to participate in the MCR activation complex[34], while the Mansoorabadi group has postulated that McrD may facilitate the insertion of coenzyme F430 into the MCR complex[80]. These genes will then be transferred into a heterologous host by means of genetic engineering for the expression of mature MCR. Achieving this would not only establish the complete set of genes and accessories required for MCR assembly and activation, it would also facilitate the engineering of an industrial host microbe for either the production or conversion of methane.

Studies into the enzyme, methyl-coenzyme M reductase (MCR), have been slowed by several factors such as the strictly anaerobic conditions it takes to grow methanogenic archaea as opposed to a facultative organism like *Escherichia coli* (*E. coli*)[81]. Furthermore, the entire machinery required for MCR assembly and activation, is yet to be fully elucidated or understood [82–84]. Genetic

information for the 1-*N*-methylhistidine PTM remains unresolved[83,85]. The successful expression of active MCR in a heterologous non-methanogenic host has yet to be reported.

Whitman and his colleagues successfully cloned MCR genes from *Methanothermococcus okinawensis* into *Methanococcus maripaludis*[86]. Recombinant MCR was successfully expressed and purified from the modified organism, but with very little activity. McrA was tagged on the C-terminus, and two populations of MCR were observed, one that was bound to F430 and another lacking F430 but attached to McrD. Neither population of MCR had much activity. Poor activity could have been due to the incorporation of the C-terminal His-tag on McrA which prevented complete F430 incorporation into MCR, which would also explain the two populations of MCR observed. It is also possible that the rate of recombinant MCR overexpression was too fast for the amount of coenzyme F430 produced by the cell. In another study by Dipti's group, the mcrG protein was tagged and the resulting strains were viable, suggesting that tagging mcrG might result in better outcomes for MCR expression[62].

The complete genomic sequencing of *Methanosarcina acetivorans* revealed the presence of various methanogenesis markers within the organism's genome. These markers, termed as such due to their unique and conserved nature across the initial methanogen genomes analyzed, encompass genes known as methanogenesis marker (*mm*) genes 1 - 17, along with the *A2* gene, which is a part of the MCR activation complex[86]. Additionally, the *mm7* gene is a part of this activation complex. While *mm1* have been found to be involved in the insertion of the thioglycine PTM[53], *mm10* responsible for the methylarginine PTM[55] , and *mm13* involved in coenzyme F430 biosynthesis[80], the exact role of the other methanogenesis markers are yet to be uncovered. The *mm3*, *mm5*, *mm6*, *mm7*, *mm15* and *mm17* genes are clustered with *A2* gene, it is possible that the

other *mm* genes may be a part of the MCR activation complex. *mm2* may also be a part of the activation complex due to its proximity to *mm3*.

Some other genes are known to act in association with methanogenic markers and have been implicated in MCR maturation and assembly, such as *tfuA* and *thiI* (with *mm1*). *TfuA*-associated *mm1* was found to be responsible for the thioglycine PTM in the MCR from *M. acetivorans*[53], while ThiI is a sulfur carrier protein that could be one of the sulfur donors for the thioglycine PTM.

The successful assembly and expression of MCR in a heterologous, non-methanogenic host will require that all genes necessary for the enzyme assembly be accurately identified, characterized, and introduced into the heterologous non-methanogenic host. Successfully expressing active MCR in a heterologous host will open new frontiers for MCR studies and a better understanding of the complete apparatus required for its assembly and activation[62,82,87].

In this chapter, I describe the steps taken to transfer *mcr* genes, *cfb* genes, *mm* genes and some of the post-translational modification genes from *M. acetivorans* into *E. coli* cells for the expression of MCR. *E. coli* cells are commonly used as heterologous hosts for protein expression. Several vectors were utilized to introduce *M. acetivorans* genes into *E. coli* including the pETDuet-1, pACYCDuet-1, pRSFDuet-1 and the pCDFDuet-1 vectors.

Initial studies utilized the following plasmids; pETDuet-1:*mcrA-mm3* consisting of a codon optimized *mcrA* gene which contains a C-terminal His-tag, and the *mm3* gene, pRSFDuet-1:*mcrBCDG* plasmid containing the rest of the *mcr* genes, pCDFDuet-1:*cfbABCDE-nixA* which contains the genes responsible for the biosynthesis of coenzyme F430, and the pACYCDuet-1:*mm1-mm10-tfuA-thiI-prmA-mm4* plasmid containing known PTM genes and one gene (*mm4*) suspected to be responsible for the methylhistidine modification. Only McrA was observed in the

insoluble fraction of the cell post expression and purification. The study by Whitman and his colleagues[86] discussed above which recorded low activity of MCR when McrA was tagged, informed the need to transfer the his-tag to McrG. McrA is a bigger protein (63 kDa) than McrG (27 kDa). It is possible that attaching a tag to McrA resulted in a distortion of its secondary structure, thereby leading to its expression in the insoluble fraction of the cell. A smaller protein like McrG may not experience as much distortion to its secondary structure.

Methanogenesis markers such as *mm1*, *mm10*, and *mm15*, have been demonstrated to contribute to MCR maturation/assembly. The *A2* protein is essential for the *in vivo* activation of MCR through the activation complex[34]. It is co-located with *mm7*, which is also a component of the activation complex, along with *mcrC*, as well as other methanogenesis markers. These markers may have roles in either MCR folding/assembly (with several putative ATPases that could function as chaperones) or in the delivery and reductive activation of MCR, potentially serving as a large reductase complex that aids in reducing the nickel of F430 to its active $^{1+}$ oxidation state. This informed the decision to utilize the pETDuet-1:*mm2-mm3-mm5-mm6-mm7-mm15-mm17-A2* plasmid in creating subsequent MCR cell lines.

Subsequent studies utilized the pCDFDuet-1:*cfbABCDE-nixA* plasmid, the pETDuet-1:*mm2-mm3-mm5-mm6-mm7-mm15-mm17-A2* plasmid, and the pRSFDuet-1:*mcrABCDG* plasmid containing the *mcr* genes with an N-terminal his-tag on *mcrG*.

Previously, plasmids containing *mcr* related genes had been successfully constructed by the Mansoorabadi lab using the Gibson Assembly protocol (See appendix). These plasmids include: pETDuet-1:*mm2-mm3-mm5-mm6-mm7-mm15-mm17-A2* containing the methanogenesis marker genes, pCDFDuet-1:*mm1-mm10-tfuA-thiI*[88] containing some post translational modification genes, pETDuet-1:*mcrA- mm3*[88] containing a codon-optimized *mcrA* gene with a C-terminal His-

tag attached to the *mcrA* gene, pRSFDuet-1:*mcrBCDG* containing the *mcr* genes except for *mcrA*[88] and pETSUMO:*mcrA* also containing a codon optimized *mcrA* gene[89] (see Figure 11).

From these existing plasmids, two more plasmids were constructed: pACYCDuet-1:*mm1-mm10-tfuA-thiI-prmA-mm4* and pRSFDuet-1:*mcrABCDG* plasmid with an N-terminal tag on *mcrG*. The pACYCDuet-1:*mm1-mm10-tfuA-thiI-prmA-mm4* plasmid was created with the aim of adding more PTM genes (*mcmA* and *mtxX*) to the existing pCDFDuet-1:*mm1-mm10-tfuA-thiI* plasmid. The pRSFDuet-1:*mcrABCDG* plasmid was constructed to create an *mcr* plasmid with the his-tag on *mcrG*. For the purpose of this section, I will be interchanging the literature established gene names with the names that the genes are called in the Mansoorabadi lab; *mcmA* gene will be *prmA* and *mtxX* will be *mm4*.

Figure 11: Previously constructed MCR-related plasmids in the Mansoorabadi lab.

Genes ligated into these vectors are strictly controlled by the bacteriophage T7 promoter[90]. The induction of the target gene with isopropyl-β-D-thiogalactopyranoside (IPTG) relies on the expression of the T7 RNA polymerase. For *E. coli* BL21(DE3) cells, there is a chromosomal copy of T7 RNA polymerase gene which is under the control of the *lacUV5* promoter, so there is still some expression of the target protein without IPTG induction[90]. For *E. coli* BL21(DE3)pLysS cells, target protein expression is lowered or eliminated without IPTG. This is due to the expression of a T7 lysozyme by the cell, which inhibits T7 RNA polymerase[91].

We chose both the BL21(DE3) and the BL21(DE3)pLysS cells as the heterologous host for the study due to the following reasons: both cells were available and are compatible with the chosen vectors, the BL21(DE3)pLysS cell also encodes the T7 lysozyme protein that suppresses the basal expression of toxic target proteins[91], thus stabilizing target gene expression by ensuring that prior to induction, protein expression is suppressed.

Transformed *E. coli* cells were grown under varying conditions to determine optimum conditions for the expression of soluble MCR. The successful expression of MCR can be determined by the generation of appropriately sized protein bands on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gels. Protein bands can be excised from the gel and subjected to mass spectrometric analysis to identify the protein more accurately.

2.2 Methods

2.2.1 Construction of the plasmid containing MCR PTM genes (pACYCDuet-1:*mm1-mm10-tfuA-thiI-prmA-mm4*)

The genes *mcmA* (*prmA*), and *mtxX* (*mm4*) were amplified using Polymerase Chain Reaction (PCR) from the genomic DNA of *M. acetivoran* C2A purchased from DSMZ in Germany (DSM-2834). The gene cluster *mm1-mm10-tfuA-thiI* was amplified from previously constructed pCDFDuet-1:*mm1-mm10-tfuA-thiI*. The plasmid pACYCDuet-1 purchased from Novagen EMD Millipore of Germany was amplified using PCR. Gibson Assembly primers used in the PCR process to generate sticky ends are outlined in Table 1. Primers were synthesized by Sigma-Aldrich and Phusion High-Fidelity DNA polymerase (New England Biolabs) was used for all PCR reactions following the manufacturer's protocol.

pACYCDuet-1:*prmA* was created using the Gibson Assembly protocol (see Appendix). *prmA* gene and an empty pACYCDuet-1 vector were amplified by PCR. The PCR products were purified via a Lonza FlashGel^{TM} DNA Cassette (Basel, Switzerland), and recovered using the OMEGA Bio-tek E.Z.N.A.® Gel Extraction Kit (Norcross, GA). Purified PCR gene and vector fragments were incubated with the Gibson Assembly Master Mix at 50 °C for 15 minutes. The *prmA* gene was inserted into the first multiple cloning site of pACYCDuet-1 (see Figure 12). The product of the incubation was used to transform *E. coli* NEB 5α cells and plated on LB agar plates containing 35 μg/mL chloramphenicol and left to grow overnight. Colonies were handpicked from the agar plate, grown in liquid LB medium and plasmids were extracted from the colonies using the Omega Bio-tek E.Z.N.A. ® Plasmid Mini Kit. The sequence of the pACYCDuet-1:*prmA* was verified by Eurofins Genomics.

pACYCDuet-1:*prmA-mm4* was created by ligating *mm4* gene to the second multiple cloning site of pACYCDuet-1:*prmA* (See Figure 12). *Mm4* gene and pACYCDuet-1:*prmA* vector were

amplified and the PCR products were purified and recovered and subjected to the Gibson Assembly protocol. The ligation mix was incubated at 50 °C for 30 minutes and the incubation product was used to transform *E. coli* NEB 5α cells and plated on LB agar plates containing chloramphenicol overnight. Grown colonies were picked and grown in LB medium. The pACYCDuet-1:*prmA-mm4* plasmid was extracted and sequence verified by Eurofins Genomics.

Finally, to create pACYCDuet-1:*mm1-mm10-tfuA-thiI-prmA-mm4*, a *Swa1* restriction enzyme site was introduced approximately 100 bp upstream of the first multiple cloning site of pACYCDuet-1:*prmA-mm4* plasmid. The plasmid was treated with SwaI restriction enzymes to create a linear plasmid. The *mm1-mm10-tfuA-thiI* gene was amplified using PCR and was purified and recovered. Purified PCR product *mm1-mm10-tfuA-thiI* was incubated with the *Swa1* cut and linearized pACYCDuet-1:*prmA-mm4* plasmid for one hour at 50 °C using the Gibson Assembly method. The resultant product was used to transform *E. coli* NEB 5α cells and was plated on LB agar plates containing 35 μg/mL chloramphenicol. The plasmid was extracted and sequence verified by Eurofins Genomics.

Table 1: Sequence of primers used for the creation of pACYCDuet-1:mm1-mm10-tfuA-thiI-prmA-mm4 plasmid

Plasmid Name		Primers
pACYCDuet-1 *prmA*	Gene	Fwd: 5'-GGAGATATACCatggaaataagatgtaggtgtg-3' Rev: 5'-CGGATCCtcaaatcacaacaacttttc-3'
	Vector	Fwd: 5'-ttgtgatttgaGGATCCGAATTCGAG-3' Rev: 5'-cctacatcttatttcCATGGTATATCTCCTTATTAAAG-3'
pACYCDuet-1 *prmA-mm4*	Gene	Fwd: 5'-AGTATAAGAAGGAGATATACATatgtatgccctcctcgag-3' Rev: 5'-AGACTCGAtcagcgaatcacctttttc-3'
	Vector	Fwd: 5'-attcgctgaTCGAGTCTGGTAAAGAAAC-3' Rev:5'-atacatATGTATATCTCCTTCTTATACTTAACTAATATAC-3'
pACYCDuet-1 *mm1-mm10-tfuA-thiI-prmA-mm4*	Gene	Fwd: 5'-atggccgagataaaaattgacaggtcac-3' Rev: 5'-ctgaaagctgtaaaaatcctcaagctctga-3'

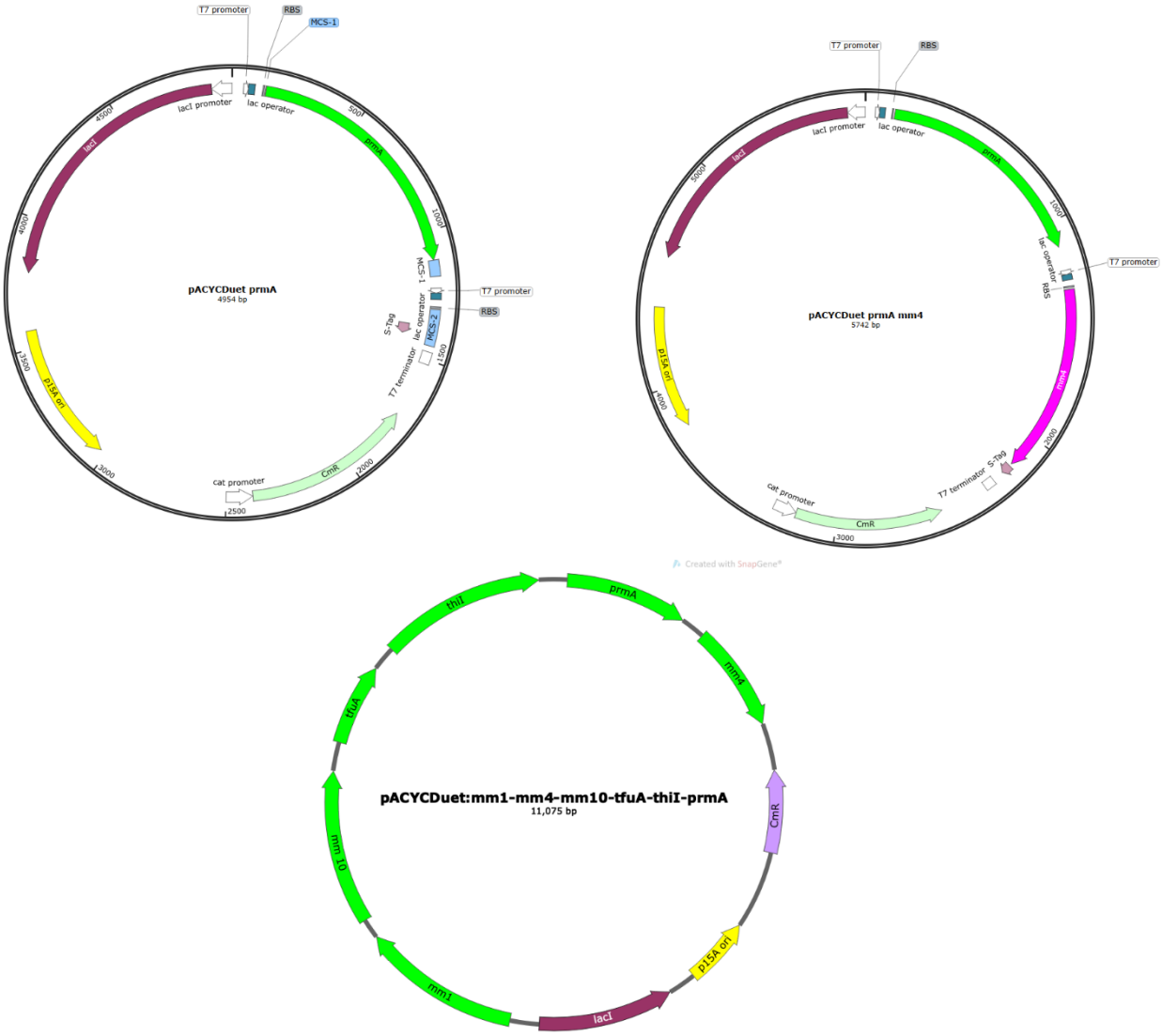


Figure 12: Plasmids used in the construction of the pACYCDuet-1: mm1-mm10-tfuA-thiI-prmA-mm4 plasmid

2.2.2 Creation of the pRSFDuet-1:*mcrABCDG-NTH* plasmid

The gene *mcrG* was amplified from pRSFDDuet-1:*mcrGBDC* template using Gibson Assembly primers specific to the Nde1 site of the vector pET-28b(+). The plasmid pET-28b(+) was also amplified via PCR. The gene and plasmid PCR products were purified using Lonza FlashGel ™ DNA Cassette (Basel, Switzerland) and recovered using the OMEGA Bio-tek E.Z.N.A.® Gel Extraction Kit (Norcross, GA). Purified gene and vector fragments were subjected to the Gibson Assembly ligation protocol to create pET-28(+):*mcrG-NTH* (Figure 13). Newly created plasmid was used to transform *E. coli* NEB 5α cells and was plated on LB agar plates containing 100 µg/mL ampicillin and left to grow overnight. Grown colonies were picked and the plasmids were extracted using the Omega Bio-tek E.Z.N.A. ® Plasmid Mini Kit. The plasmid was sent to Eurofins Genomics where it was subsequently sequence verified.

Using (pET-28b(+):*mcrG-NTH*) as a template, *mcrG-NTH* gene was amplified with Gibson Assembly primers. pRSFDuet-1:*mcrBDC* vector was amplified using pRSFDuet-1:*mcrBCDG* as the template. PCR fragments were purified and recovered. Purified PCR fragments were subjected to the Gibson Assembly ligation protocol to create pRSFDuet-1:*mcrG-NTH-mcrBDC* plasmid (Figure 13) which was used to transform *E.coli* NEB 5α cells. Cells were plated on LB agar plates containing 100 µg/mL of ampicillin. Colonies were picked from cells and the plasmid was extracted and sent to Eurofins Genomics for sequence verification.

The *mcrA* gene was amplified using pETSUMO:*mcrA* as template and empty pRSFDuet-1 vector was amplified with Gibson Assembly primers. PCR fragments were purified and recovered. The purified fragments were ligated using Gibson Assembly protocol as per the manufacturer's instructions. The resulting pRSFDuet-1:*mcrA* plasmid (Figure 13) was used to transform *E.coli*

NEB 5α cells and plated on agar cells containing 50 µg/mL of kanamycin. Colonies were picked for plasmid extraction and the extracted plasmids were sequenced verified by Eurofins Genomics.

Finally, using pRSFDuet-1:*mcrG-NTH-mcrBDC* as template, the gene cluster *mcrG-NTH-mcrBDC* was amplified with Gibson Assembly primers. The vector pRSFDuet-1:*mcrA* was also amplified with Gibson Assembly primers. PCR fragments were purified and recovered. Purified PCR fragments were subjected to the Gibson Assembly ligation protocol to create pRSFDuet-1:*mcrABCDG-NTH* plasmid (Figure 14) by incubating for 1 hour, 35 minutes at 50 °C and with the addition of 10% DMSO to the Gibson Assembly incubation mix. *E. coli* NEB 5α cells were transformed with the plasmid and plated on agar plates containing 50 µg/mL of kanamycin. The plasmid was extracted from the resulting colonies and was sequenced verified by Eurofins Genomics. Table 2 contains a list of Gibson Assembly primers used in the creation of this plasmid.

Table 2: Sequence of primers used in creating pRSFDuet-1:mcrABCDG-NTH

pET-28(+):*mcrG-NTH*	Gene	Fwd: 5'-GGCAGCCATatggcatacgaagcacag-3' Rev: 5'-GTGCTCGAGtcatttcggctggaatcc-3'
	Vector	Fwd: 5'-gccgaaatgaCTCGAGCACCACCACCAC-3' Rev: 5'-cttcgtatgccatATGGCTGCCGCGCG-3'
pRSFDuet-1: *mcrBCDG-NTH*	Gene	Fwd: 5'-GATTTAAATCTCGATCCCGCGAAATTAATACGAC-3' Rev: 5'-gtgtacccattcaatgacttctgcgtc-3'
	Vector	Fwd: 5'-agtcattgaatgggtacacagagtattcgacc-3' Rev: 5'-GCGGGATCGAGATTTAAATCGATCCTCTACG-3'
pRSFDuet-1:*mcrA*	Gene	Fwd: 5'-AGGAGATATACCatggcagcagatatcttcg-3' Rev: 5'-TCGACTTAAGCActattttgcaggaataaccaggc-3'
	Vector	Fwd: 5'-tgcaaaatagTGCTTAAGTCGAACAGAAAGTAATCG-3' Rev: 5'-ctgctgccatGGTATATCTCCTTATTAAAGTTAAACAAAATTATTTC-3'
pRSFDuet-1: *mcrABCDG-NTH*	Gene	Fwd: 5'-CCGGCAAAATAGAAATCTCGATCCCGCGAAATTAATAC-3' Rev: 5'-TAATTTCGCGGGATCGAGATTTCTATTTTGCCGGGATGacg-3'
	Vector	Fwd: 5'-CCGGCGTAGAGGATCGATTTAAATTAATCGTATTGTACACGG-3' Rev: 5'-CAATACGATTAATTTAAATCGATCCTCTACGCCGGACG-3'

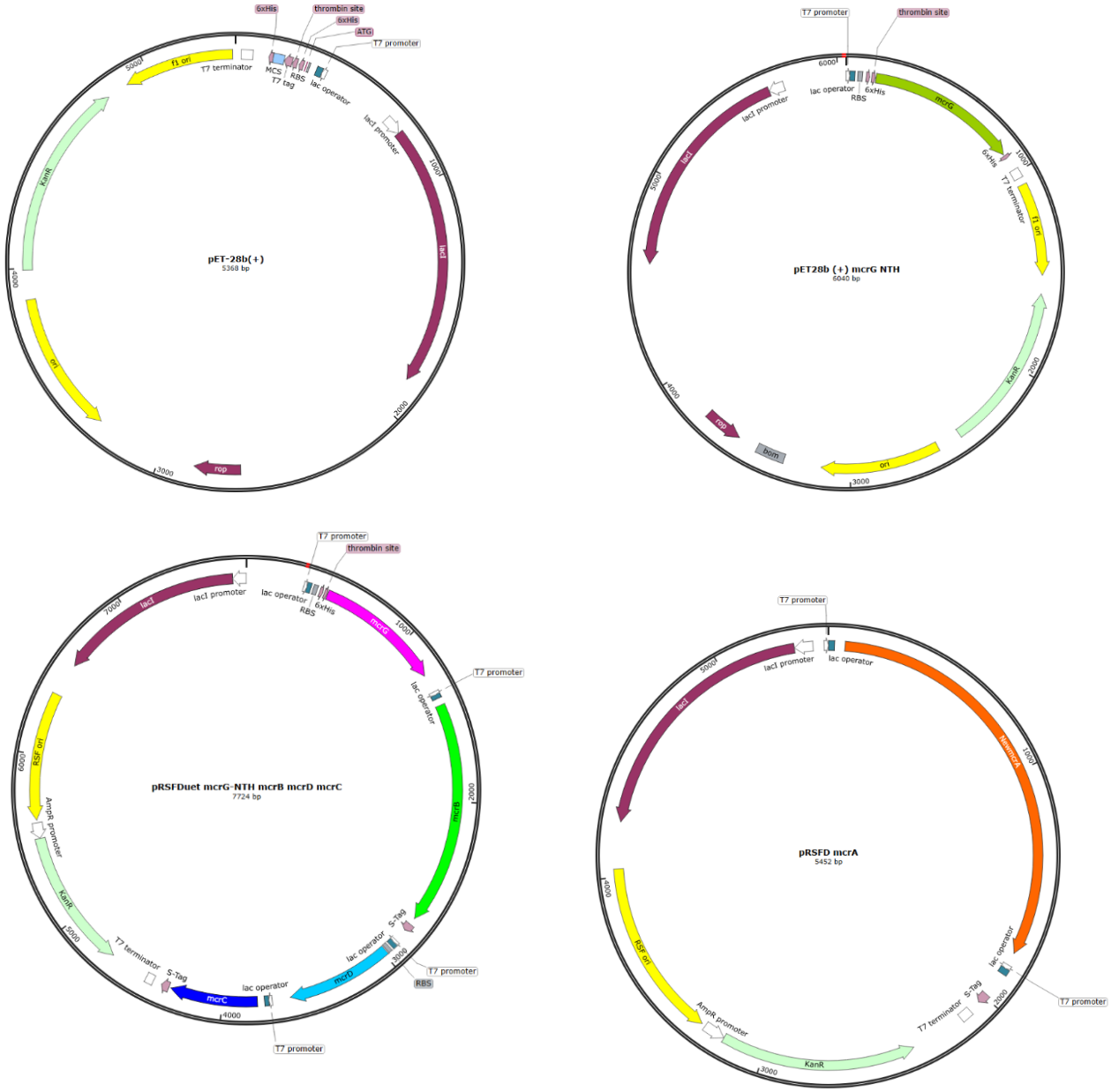


Figure 13: Set of plasmids used in creating pRSFDuet-1:mcrABCDG-NTH

2.2.3 Creation of MCR cell lines

To study the heterologous expression of MCR, four MCR cell lines were created using different combination of the existing and created plasmids (Figure 14).

MCR cell line 1 was created with the plasmids pETDuet-1:*optmcrA-CTH-mm3*, pRSFDuet-1:*mcrBCDG*, pCDFDuet-1:*mm1-mm10-tFuA-ThiI* and pACYCDuet-1:*prmA-mm4*. The plasmids were used to transform *E. coli* BL21(DE3) cells.

MCR cell line 2 consisted of the *mcr* and *cfb* genes containing plasmids. *E. coli* BL21(DE3)pLysS cells were transformed with the plasmids pRSFDuet-1:*mcrABCDG-NTH* and pCDFDuet-1:*cfbABDCE-nixA*.

MCR cell line 3 consisted of the *mcr* and *mm* genes containing plasmids. It was created by transforming *E. coli* BL21(DE3)pLysS cells with the plasmids pRSFDuet-1:*mcrABCDG-NTH* and pETDuet-1:*mm2-mm3-mm5-mm6-mm7-mm15-mm17-A2*.

MCR cell line 4 consisted of the *mcr*, *cfb* and *mm* genes containing plasmids. It was created by transforming *E. coli* BL21(DE3)pLysS cells with the plasmids pRSFDuet-1:*mcrABCDG-NTH*, pETDuet-1:*mm2-mm3-mm5-mm6-mm7-mm15-mm17-A2*, and pCDFDuet-1:*cfbABDCE-nixA*.

Cells were first made chemically competent using 0.1M Calcium Chloride solution (see Appendix) and the plasmids were added individually to the cell. Transformed cells were made chemically competent in between each transformation step. After each step of transformation, the plasmids were extracted from transformed cells using Omega Bio-tek E.Z.N.A. ® Plasmid Mini Kit. The identity of each plasmid extracted from the cell line was verified by PCR verification. Using gene specific primers, PCR amplification was carried out on the extracted plasmids. The PCR

amplification product was injected into the Lonza FlashGel™ DNA Cassette for separation of DNA by electrophoresis. Genes present in the extracted plasmid could be identified based on size.

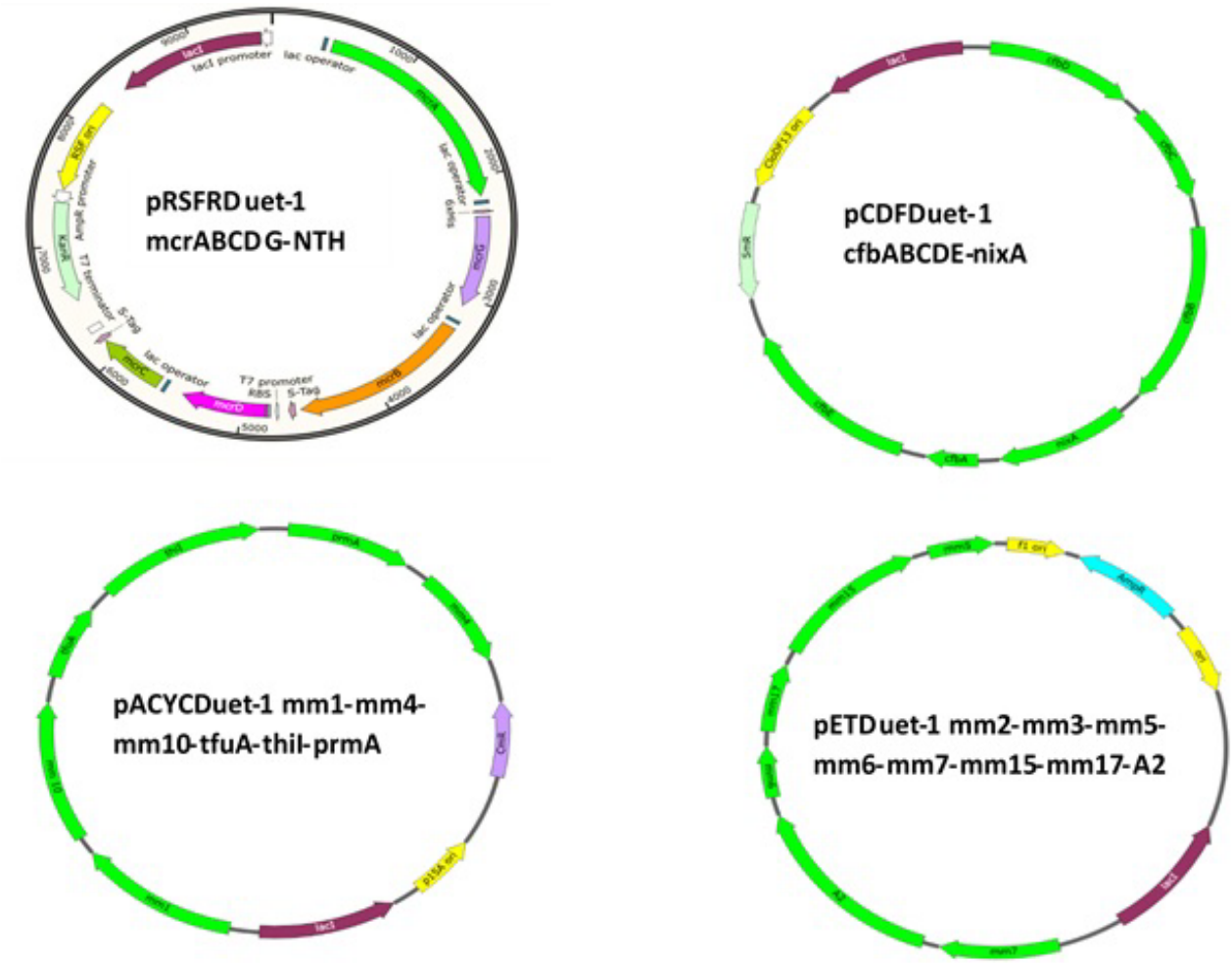


Figure 14: Complete set of genes used in creating MCR cell lines for MCR heterologous assembly studies

2.2.4 Aerobic and anaerobic expression of MCR cell lines

The newly constructed MCR cell lines were grown under aerobic and anaerobic conditions to study MCR expression under various conditions.

MCR cell line 1 (*E. coli* BL2{DE3}) was grown in terrific broth supplemented with kanamycin (50 µg/mL final concentration), ampicillin (100 µg/mL final concentration), spectinomycin (25 µg/mL final concentration) and chloramphenicol (35 µg/mL final concentration) antibiotics with shaking at 37 °C. The broth was also supplemented with 0.2 mM iron II sulphate, 0.15 mM cysteine, 0.1 mM cyanocobalamin, 0.2 mM riboflavin and 0.2 mM methionine. After 6 h, the cells were induced using β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 400 µM. The temperature was lowered to 18 °C and the cells were left to shake for an additional 18 h before harvesting.

MCR cell lines 2 – 4 were expressed aerobically for the first 24 hours and anaerobically for another 24 hours.

MCR cell line 4 (*E.coli* BL21{DE3}pLysS) was grown in broth consisting of sodium chloride (5 g/L), tryptone (10 g/L), and yeast extract (5 g/L) as well as salts consisting of 0.17 M potassium phosphate monobasic and 0.72 M potassium phosphate dibasic salts. The broth was supplemented with kanamycin (50 µg/mL final concentration), ampicillin (100 µg/mL final concentration), spectinomycin (25 µg/mL final concentration) and chloramphenicol (35 µg/mL final concentration) antibiotics and cells were left to shake at 37 °C for 6 h. At induction with 300 µM IPTG, cells were further supplemented with 0.2 mM iron II sulphate, 0.15 mM cysteine, 0.1 mM cyanocobalamin, 0.2 mM riboflavin and 0.2 mM methionine, the temperature was lowered to 18 °C and shaking continued for 18 h. After 18 h, the cells were transferred to anaerobic flasks, placed in an ice-bath and subjected to argon pressurization for 2 h to remove all traces of oxygen from

the broth. De-oxygenated cells were left overnight (12 – 18 h) standing at 4 °C and then harvested anaerobically.

MCR cell line 2 was grown in broth supplemented with kanamycin (50 µg/mL final concentration), spectinomycin (25 µg/mL final concentration) and chloramphenicol (35 µg/mL final concentration) antibiotics and 200 µM IPTG while MCR cell-line 3 was grown in broth supplemented with kanamycin (50 µg/mL final concentration), ampicillin (100 µg/mL final concentration), and chloramphenicol (35 µg/mL final concentration) antibiotics and 200 µM IPTG. Both cell lines were subjected to the same set of conditions as MCR cell line 4.

The cell pellets of all harvested cell-lines were stored at -80 °C.

2.2.5 Aerobic and anaerobic purification of MCR cell lines using affinity chromatography

MCR cell line 1 was purified under aerobic conditions. The cell pellets were thawed at room temperature and re-suspended in lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, lysozyme (1 mg/ml). The cell suspension was lysed using sonication and centrifuged. The resulting supernatant was subjected to purification using Bio-Rad Econo-Pac column packed with Ni^{2+}-charged Profinity IMAC Resin. The column was washed with 50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, pH 8.0 buffer. The protein was then eluted with 50 mM sodium phosphate, 300 mM NaCl, 500 mM imidazole, pH 8.0 buffer, and buffer exchanged with 100 mM Tris-HCl, pH 8.0 buffer. Protein was stored at -80 °C.

Cell pellets from MCR cell lines 2-4 were purified under anaerobic conditions. Cells were transferred to the anaerobic tent and thawed at room temperature. It was re-suspended in anaerobic lysis buffer and the cell suspension was sonicated and centrifuged. The supernatant was applied to Bio-Rad Econo-Pac column packed with Ni^{2+}-charged Profinity IMAC Resin. The column was

washed with anaerobic 50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, pH 8.0 buffer. The protein was then eluted with anaerobic 50 mM sodium phosphate, 300 mM NaCl, 500 mM imidazole, pH 8.0 buffer, and buffer exchanged with anaerobic 100 mM Tris-HCl, 20% glycerol, pH 8.0 buffer. It was stored at -80 °C. Purification fractions were used for sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and western blot. Bands on the SDS gel corresponding to the size of the expected proteins were excised and used for mass spectrometry analysis.

2.2.6 LC-MS analysis

The analysis of the assays was conducted at the Mass Spectrometry Center within the Department of Chemistry and Biochemistry at Auburn University. This analysis utilized an ultra-performance LC system (ACQUITY, Waters Corp., USA) coupled with a quadrupole time-of-flight mass spectrometer (Q-TOF Premier, Waters Corp., USA) featuring electrospray ionization (ESI) in positive mode and controlled by MassLynx software (version 4.1).

Sample injections were directly made into the mass spectrometer or onto a C18 column (Aeris™ 3.6 µm C4 200 Å, 50 × 2.1 mm, Phenomenex) at a flow rate of 300 µL/min, using mobile phase solutions A (0.1 % formic acid in 95 % water and 5 % acetonitrile) and B (0.1 % formic acid in 95 % acetonitrile and 5 % water). The gradient used for elution was as follows: 0 % B for 2 minutes, followed by a linear gradient to 100 % B over 11 minutes, maintaining 100 % B for 1 minute, another linear gradient back to 0 % B over 1 minute, and a final 3-minute equilibration at 0 % B. The column temperature was held at 40 °C, and UV absorption was monitored at 280 nm. Operating parameters for the mass spectrometer included a capillary voltage of 3.1 kV, a sample cone voltage of 15 V, and an extraction cone voltage of 4.0 V. The source and desolvation temperatures were maintained at 95 and 300 °C, respectively, with a desolvation gas flow rate of

600 L/h. The mass spectrometry scan had a duration of 0.5 seconds, covering the range from 50 to 2,000 m/z, with a 0.02 s interscan delay, and data were recorded in the centroid format.

2.3 Results

2.3.1 The heterologous assembly of MCR in MCR cell line 1

Successfully created plasmids containing *mcr* genes, *cfb* genes, *mm* genes and post translational genes were used to construct the MCR cell lines outlined in section 2.2.3. Transformed cells were grown in broth, induced with IPTG, harvested, and lysed to obtain soluble and insoluble fractions of the cell. Flowthrough (FT), wash (W), and elution (E) fractions are collected during the purification process, as well as the insoluble (IN) fraction and these fractions are subjected to sodium dodecyl polyacrylamide gel electrophoresis (SDS-PAGE) and western blot. Detection of a protein size matching the calculated size of the protein (calculation performed using the amino acid sequence of protein on https://expasy.com), suggests the presence of the tagged protein which can be subjected to further analysis. The size of McrA protein is 63 kDa.

The heterologous production and assembly of MCR from cell-line 1 was first examined. It consisted of *E. coli* BL21(DE3) cells containing the pETDuet-1:*optmcrA-CTH-mm3*, pRSFDuet-1:*mcrBCDG*, pCDFDuet-1:*mm1-mm10-tFuA-ThiI*, and pACYC:*prmA* plasmids. A band corresponding approximately to the size of McrA (63 kDa) was found in the insoluble fraction (IN) as shown in the western blot (Figure 15). Another distinct band was observed at the top of the western. This was interesting because the band may be as a result of the formation of McrA dimer (~126 kDa) or the McrA, McrB and McrG trimer (~300 kDa), and would mean that either McrA dimer or MCR subunits were assembling together.

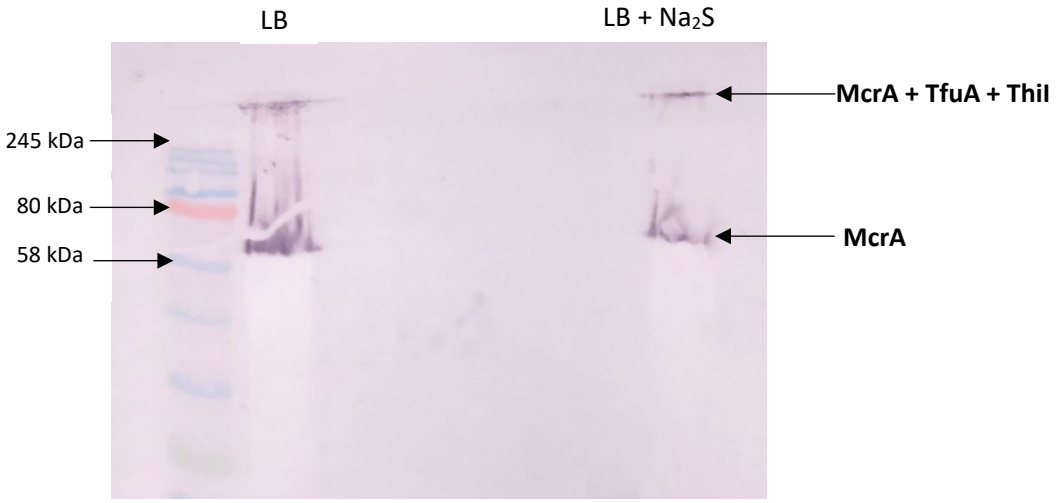


Figure 15: Western blot membrane result for MCR cell line 1

The protein bands were cut out of the gel and sent for mass spectrometry. Mass spectrometry confirmed the presence of McrA in the 63 kDa protein band. Mass spectrometry also confirmed the presence of an McrA complex consisting of McrA, TfuA and ThiI proteins in the protein band at the top of the gel. While this was unexpected, it was nonetheless interesting because *tfuA*-associated *ycaO* homolog (*mm1*) has been found to be responsible for the thioglycine PTM in the MCR from *Methanosarcina acetivorans* [53], while ThiI is a sulfur carrier protein that could be one of the sulfur donors for the thioglycine PTM. ThiI has a disulfide bond that can react with sulfides to make a persulfide that can then be donated to other proteins. TfuA plays a role in the release of thioglycine by hydrolyzing the thioester after it has been formed (Scheme 4). It is unclear if ThiI can also donate sulfur to MCR and whether TfuA will be involved in this process. The presence of the three proteins in the band may be indicative of such an interaction. It was also possible that

they were precipitated proteins which remained at the top of the gel, and not that the three proteins were forming a complex.

The association of McrA, TfuA and ThiI proteins informed a repeat of the cell expression and protein purification process with sodium sulfide (0.15 mM) supplementation. This is because we wanted to see if the complex (McrA+TfuA+ThiI) formed was due to insufficient sulfide, preventing the formation and release of thioglycine by TfuA. By supplementing the broth with 0.15 mM sodium sulfide at the point of IPTG induction, we expected to drive the formation of the thioglycine and see a complete resolution of the proteins. Supplementation with sodium sulfide did not result in the resolution of the McrA, TfuA and ThiI proteins in the complex (Figure 15), which leads to a strong suggestion that the proteins are not forming a complex. They may have been precipitated proteins left over in the SDS-PAGE well.

McrA protein was found only in the insoluble fraction of the cell. The insoluble fraction was obtained by dissolving the pellet in urea and centrifuging to extract the supernatant which was loaded unto the SDS gel. The presence of McrA in the insoluble fraction suggests that the protein is folding improperly. Furthermore, McrA was the only subunit containing the histidine tag, and thus, the only subunit that can be detected by western blot. If McrA was assembling with McrB and McrG, all three proteins should remain bound to the column and be detected. However, since only McrA was detected by western blot and by mass spec, we can assume that it did not assemble with McrB and McrG. Subsequent attempts to study the assembly of MCR with this cell line did not yield different results.

The role of the *cfb* genes and the *mm* gene cluster in MCR assembly (described section 2.1) was explored leading to the creation of cell lines 2 – 4.

2.3.2 The heterologous assembly of MCR in MCR cell lines 2 and 3

MCR cell line 2 contains the plasmids pRSFDuet-1:*mcrABCDG-NTH* and pCDFDuet-1:*cfbABDCE-nixA* and MCR cell line 3 contains the plasmids pRSFDuet:*mcrABCDG-NTH* and pETDuet-1:*mm2-mm3-mm5-mm6-mm7-mm15-mm17-A2.* Both cell lines were *E. coli* BL21(DE3)pLysS cells.

The two cell lines were constructed to study the effect of the corresponding genes on MCR assembly in a heterologous host. Cell line 2 consists of *mcr* genes and the *cfb* genes while cell line 3 contains *mcr* genes and the *mm* genes. The his-tagged McrG protein has a mass of 27 kDa.

Gel slices corresponding to the bands highlighted in Figure 16 were excised and subjected for mass spectrometry analysis. While SDS bands looked promising in the MCR cell line 3 containing *mcr* and *mm* genes, McrG was not detected by mass spectrometry when the SDS bands were analyzed. This may be attributed to the degradation of the samples over time. It is also possible that the *mm* genes are toxic and do not express well in non-methanogenic hosts, therefore interfering with the expression of MCR.

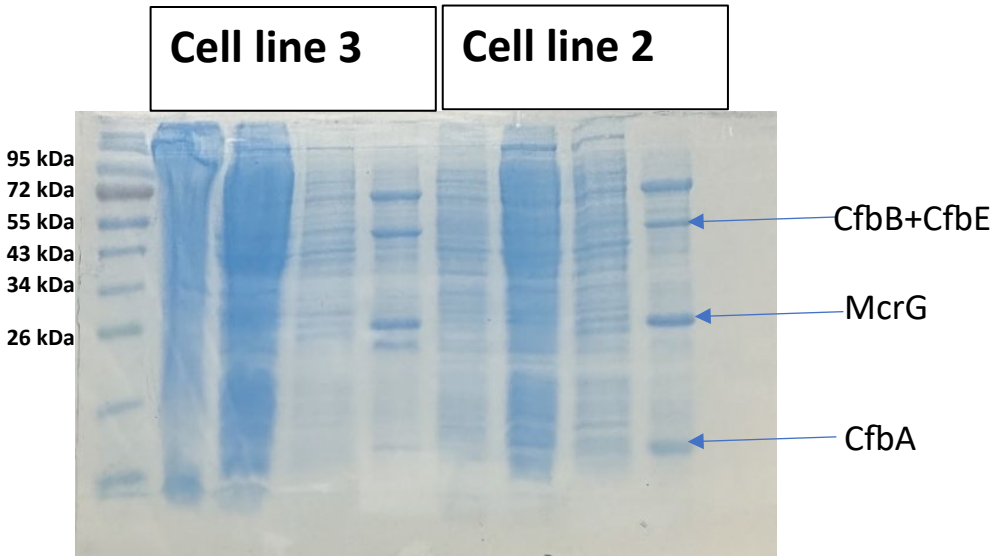


Figure 16: SDS PAGE gel result for MCR heterologously expressed in MCR cell lines 2 and 3.

Lane 1 represents the molecular weight standard, lanes 2 - 5 represents the insoluble, flowthrough, wash and elution fractions from cell line 3. Lanes 5 - 8 represents the insoluble, flowthrough, wash and elution fractions from cell line 2.

For cell line 2 containing *cfb* genes and *mcr* genes with a his tag on McrG, the presence of McrG in the soluble fraction of the cell was confirmed by SDS-PAGE and mass spec (Table 3). Mass spec analysis also detected the presence of the CfbABE proteins in the bands as shown in Figure 16. The presence of the CfbABE proteins in the elution fraction indicates that McrG coeluted with the CfbA, CfbB and CfbE proteins.

Table 3: Mass spec results for MCR cell line 2

Description	Coverage	# Peptides	# Amino acids	MW (kDa)	Gene Symbol
Sirohydrochlorin cobaltochelatase	96	15	130	13.7	CfbA
UDP-N-acetylmuramoylalanine-D-glutamate ligase	19	5	472	50.4	CfbE
Cobyrinate a,c-diamide synthase	14	4	497	53.5	CfbB
Coenzyme-B sulfoethylthiotransferase subunit gamma	11	2	248	27.6	McrG

The low peptide score of McrG suggests that the protein is poorly expressed, unstable, or not easily ionizable by MALDI-TOS-MS. This also provides some explanation for our inability to detect McrA and McrB in the bands corresponding to their sizes. McrA was detected by mass spec when it was found in the insoluble fraction in MCR cell line 1. This suggests that its formation in inclusion bodies, enhanced its stability, leading to its detection by mass spec in MCR cell line 1.

2.3.3 The heterologous assembly of MCR in MCR cell line 4

MCR cell line 4 contains the plasmids pRSFDuet-1:*mcrABCDG-NTH*, pETDuet-1:*mm2-mm3-mm5-mm6-mm7-mm15-mm17-A2*, and pCDFDuet-1:*cfbABDCE-nixA*. The effect of the *cfb* genes and the *mm* genes on MCR assembly in a heterologous host was explored with this cell line.

The MCR cell line 4 containing *mcr*, *cfb* and *mm* genes was expressed and purified, and the purification fractions were subjected to SDS-PAGE and western blot. The soluble fraction of the cell was also subjected to mass spec analysis.

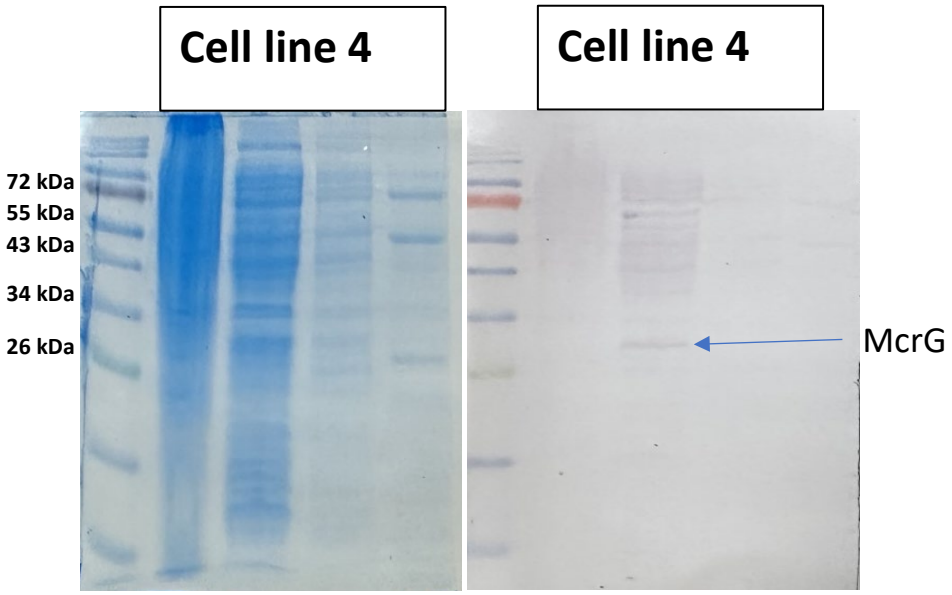


Figure 17: SDS PAGE gel and western blot membrane results for MCR heterologously expressed in MCR cell line 4 containing mcr genes, cfb genes and mm genes.

Lane 1 represents the molecular weight standard, lanes 2 - 5 represents the insoluble, flowthrough, wash, and elution fractions from cell line 4.

The presence of McrG in the soluble fraction of the cell was confirmed by SDS-PAGE and western blot (Figure 17). Mass spec result also revealed the presence of CfbB, CfbE and CfbC proteins (Table 4).

Table 4: Mass spec results for MCR cell line 4

Description	Coverage	# Peptides	# Amino acids	MW (kDa)	Gene Symbol
Ni-sirohydrochlorin a,c-diamide reductive cyclase complex	82	24	265	28.6	CfbC
Coenzyme F430 synthetase	34	12	472	50.4	CfbE
Cobyrinate a,c-diamide synthase	26	10	497	53.5	CfbB

Table 5: Summary of MCR cell line description and results

Description	**Genes present**	**Aerobic or Anaerobic expression/purification**	**MCR protein detected**	**Other proteins detected**
MCR cell line 1	*mcr* and PTM	Aerobic	McrA insoluble	TfuA and ThiI
MCR cell line 2	*mcr* and *cfb*	Anaerobic	McrG soluble	CfbABE
MCR cell line 3	*mcr* and *mm*	Anaerobic	McrG soluble	-
MCR cell line 4	*mcr, cfb* and *mm*	Anaerobic	McrG soluble	CfbBCE

2.4 Discussion

Initial expression of MCR with cell line 1, an *E. coli* BL21(DE3) cell containing the *mcr* genes with a his-tagged *mcrA* and the post translational modification genes, resulted in McrA being detected in the insoluble fraction of the cell. For an *E. coli* BL21(DE3) cell, induction with IPTG upregulates protein expression, which can lead to a faster rate of protein expression than the cell folding machinery can handle. This may lead to the expressed proteins forming inclusion bodies which are detected in the insoluble fraction of the cell. To counter this challenge, different concentrations of IPTG such as 200 μM, 100 μM, 50 μM and 50 μM were used to induce the cells during expression. None of these concentrations appeared to increase the solubility of McrA or its assembly with McrB and McrG subunits.

Data from Whitman's group[86] showing poor incorporation of F430 into MCR when McrA was his-tagged in *Methanococcus maripaludis* and data from Dipti's group[62] showing complete assembly of MCR when McrG was his-tagged in *M. acetivorans*, provided the basis for the decision to switch the his-tag from McrA to McrG in this study. In both studies, recombinant MCR was expressed in archaea hosts. McrG is a smaller protein and does not feature any post translational modification and so the attachment of a his-tag to McrG may not result in a serious distortion of its structure and may improve its assembly with other subunits. For MCR cell lines 2, 3 and 4, the his-tag was placed on McrG and the *E. coli* BL21(DE3)PLysS cell line was used for the studies. Interestingly, McrG was detected in the soluble fraction of the cell (see Figure 16 and Figure 17). This confirms that hypothesis that placing the his-tag on McrG may result in a more efficient folding of the protein and impacted the protein's solubility.

The aerobic expression and purification of the MCR cell line 1 may have also contributed to the protein being formed in the insoluble fraction of the cell. MCR is found in methanogens which are

strictly anaerobic organisms. It is therefore very plausible that the heterologous assembly of MCR in *E. coli* requires strictly anaerobic conditions. This observation informed the choice to harvest *E. coli* cells anaerobically and to carry out the process of MCR purification anaerobically in subsequent cell lines. This was successfully implemented in experiments with cell lines 2-4.

Anaerobic purification of the MCR cell lines 2-4 resulted in the expression of McrG protein in the soluble fraction of the cell. The presence of soluble McrG protein in MCR cell line 2 containing *mcr* and *cfb* genes was confirmed by SDS-PAGE and mass spectrometry (Figure 16). Soluble McrG protein was also present in MCR cell line 4 containing *mcr, cfb* and *mm* genes as confirmed by SDS-PAGE and western blot (Figure 17). In addition to McrG, the proteins CfbA, CfbB, CfbC and CfbE were also detected in the soluble fraction of the cell.

The coelution of CfbE protein with McrG is significant because CfbE is the terminal enzyme in the F430 biosynthetic pathway (Scheme 7). Its coelution with McrG is indicative that it interacts directly with McrG to deliver F430 into MCR. Previous work by the Mansoorabadi group[80] showed that CfbE interacts with McrD and could be activated by it. McrD has also been shown to be a part of the apo/semi-apo forms of MCR that is poised to receive F430[86]. Consequently, CfbE interaction with McrG indicates that CfbE might be the direct F430 chaperone that delivers the coenzyme to MCR. The interaction of CfbA, CfbB and CfbC with McrG is both unexpected and interesting, as this could be an indication that the Cfb enzymes form a larger coenzyme F430 synthase complex.

Unfortunately, none of the methods employed was successful in assembling MCR in a heterologous host (Table 5). However, we were able to establish the importance of carrying out the purification process under strictly anaerobic conditions, which improved the solubility of the

recombinant proteins. We also established that the McrG tagged protein could be expressed in the soluble fraction of the cell, and that the protein co-elutes with the Cfb proteins.

The inability to heterologously express MCR using the methods employed may be because of the absence of yet undetermined archaea genes which are integral to MCR assembly and function. It is also possible that the *E. coli* cell is an unfit host for the expression of MCR and other bacteria hosts may be explored in future studies.

Finally, MCR is a large enzyme complex consisting of six sub-units. The exact mechanism for the assembly of all required sub-units is yet to be understood. Cell conditions, proteomic and metabolite concentrations required for the accurate transcription and translation of the genes and proteins necessary for MCR assembly need to be determined in the archaea host firstly, before replication in a non-methanogenic/methanotrophic host.

Chapter Three

Heterologous expression of putative MCR post-translational modification genes and in vivo McrA maturation

3 Chapter 3

3.1 Introduction

Methane is a highly potent greenhouse gas[1,2] that is produced and utilized as part of the energy metabolism processes of methanogenic archaea, commonly known as methanogens and anaerobic methanotrophic archaea (ANME), respectively. Annually, approximately 1 billion tons of methane are generated each year by methanogens[3].

At the heart of methane metabolism in both methanogens and ANME lie the enzyme methyl-coenzyme M reductase (MCR)[6]. MCR catalyzes the conversion of coenzyme B and methyl-coenzyme M into the mixed heterodisulfide CoB-S-S-CoM and methane, playing a pivotal role in methane generation[5,6,88]. Anaerobic methanotrophic archaea (ANME) employ reverse methanogenesis to oxidize methane under anaerobic conditions with MCR acting as the first enzyme in that metabolic process[47,72,73]. Studies on MCR crystal structure revealed that the protein is a hexameric, 300 kD protein consisting of three different subunits in an α2β2γ2 configuration[23,30,52]. Tightly bound to the enzyme and relevant to its catalytic function is an unusual Ni tetrapyrrole, coenzyme F430 or 17^2-methylthio F430 in ANME-1[30,52,73].

On the α-subunit of MCR is found several unprecedented post-translational modifications (PTMs). The 1-*N*-methylhistidine and thioglycine PTMs have been observed consistently across all studied MCRs[23,30,31,50]. Additionally, certain species exhibit the presence of *S*-methylcysteine, 5-(*S*)-methylarginine, didehydroaspartate, and 2-(*S*)-methylglutamine in their α-subunit[23,30]. In contrast, the β- and γ-subunits lack any post-translational modifications.

Kahnt and coworkers evaluated the MCR structure from one member each of five methanogenic archaeal orders and two from methanotrophic archaea[50]. They found the 1-*N*-methylhistidine and thioglycine modifications in MCR from all analyzed organisms, which included

Methanothermobacter marburgensis, *Methanocaldococcus jannaschii*, *Methanopyrus kandleri*, and *Methanosarcina barkeri*. Furthermore, the hyperthermophilic species *Methanocaldococcus jannaschii* and *Methanopyrus kandleri* were devoid of the methylcysteine modification[32].

The MCR structure derived from ANME-1 had other post-translational changes like 7-hydroxytryptophan and *S*-oxymethionine[52]. Additionally, the ANME-1 MCR possessed a uniquely modified F430, 17^2-methylthio F430, which was designated F430-2[51].

The exact role of these modifications in MCR assembly, structure, and function is not yet clear. However, studies using gene deletion strategies on *Methanosarcina acetivorans* identified the methanogenesis marker 1, which is a *tfuA*-associated *ycaO* homolog, as the gene responsible for the thioglycine modification[53]. Mutant strains lacking these genes exhibited low cell yields at increased temperatures when compared to the wild-type[53]. Similarly, in studies where the *mcmA* gene which is responsible for the methylcysteine modification was deleted[62], evaluations of the phenotypic traits in mutants without the *mcmA* gene hinted at the potential significance of this PTM for the adaptation of methanogens to mesophilic environments. Experiments by Layer and his colleagues[92] with the deletion of the *mmp10* gene responsible for methyl-arginine modification indicates that the modification may be useful for the enzyme's stability under stress conditions such as elevated temperature and the presence of 0.2 mM hydrogen peroxide.

The importance of these post-translational modifications in the heterologous assembly of MCR in a non-methanogenic host is yet to be determined. However, none of these modifications have been found to be essential to the catalytic function of MCR expressed from methanogens. In studies by the Metcalf group[62], all three genes responsible for the methylcysteine, thioglycine and methyl arginine modifications were deleted and the mutant MCR was found to be active. The authors suggested that the PTMs might play a role in finetuning MCR activity and efficiency[62].

Research by the same group[62] presented *in vivo* findings indicating that a *S*-adenosylmethionine (SAM)-dependent methyltransferase, encoded by a 50S ribosomal protein L11 methyltransferase (*mcmA*) homolog (referred to as the *prmA* gene by the Mansoorabadi lab), was responsible for the *S*-methylcysteine modification in *M. acetivorans*. The *mcmA* gene is positioned next to the *mcr* gene cluster in *M. acetivorans* and remains consistent in species exhibiting the *S*-methylcysteine PTM. MCR expressed from mutant strains lacking the *mcmA* gene was found to have an unmethylated cysteine residue (cys472)[62] , as shown in Figure 18.

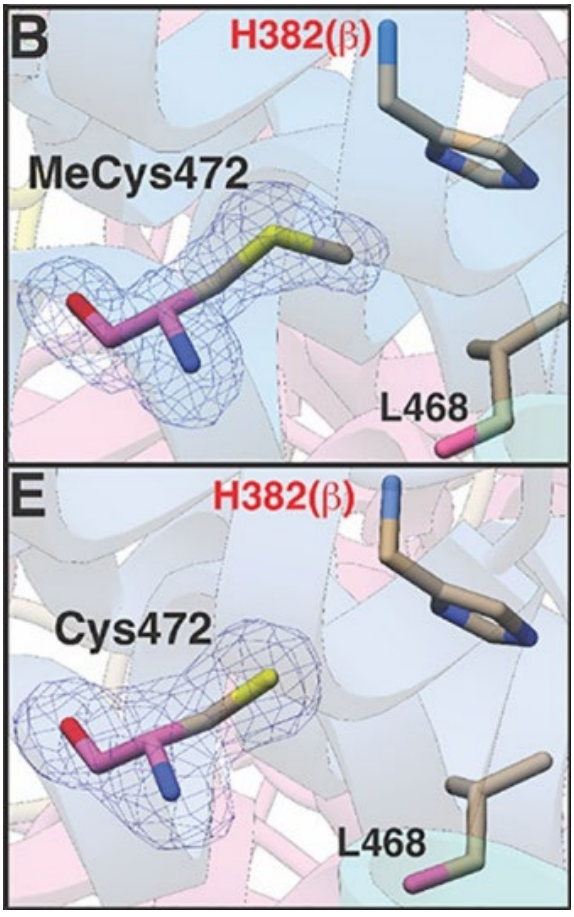


Figure 18: Methylated (B) and unmethylated (E) cys472 residue in MCR expressed from native and mutant Methanosarcina acetivorans[62]

While the study established the importance of the *mcmA* gene in the methycysteine PTM, it did not provide details of the methylation mechanism, whether the gene worked in isolation or with other accessory molecules, and if the methylation was post translational or co-translational. My study was designed to further investigate the function of the *mcmA* gene. Most studies on MCR are conducted in a native methanogen, which would already possess MCR relevant genes. For instance, in the first reported case of MCR heterologous assembly, Lyu and his colleagues[86] effectively transferred *mcr* genes from one methanogen, *Methanothermococcus okinawensis*, into another methanogen, *Methanococcus maripaludis*. While the recombinant MCR was expressed and isolated from the altered organism, its activity was notably low. Studying the implicated post translational modification genes of MCR in a heterologous host such as *Escherichia coli* (*E. coli*) may provide useful insight on the post-translational mechanism, and the impact of the modification on MCR expression in a heterologous host.

As previously noted, all MCRs studied so far exhibit the 1-*N*-methylhistidine alteration. This consistency suggests that this modification might hold a crucial role in the enzyme's functionality. However, the gene responsible for this modification remains unknown. Identifying the gene responsible for the 1-*N*-methylhistidine post-translational modification will take us a step closer to unravelling the entire set of genes required for MCR maturation/activation and getting a better understanding of the MCR assembly process/requirements.

Several genetic markers associated with methanogenesis have been identified in the genome of *M. acetivorans*[16]. These markers earned their designation because they represent genes that are unique to and conserved across all initial methanogen genomes. They encompass the methanogenesis marker (*mm*) genes 1 through 17, along with the *A2* gene, which is a part of the MCR activation complex of MCR. Additionally, the *mm7* gene is integral to this activation complex. The *mm1*

gene, a homolog of *ycaO*, is involved in the insertion of the thioglycine PTM[53], while *mm10* (*mamA*) encodes a radical S-adenosyl-L-methionine (SAM) enzyme responsible for introducing the 5-(S)-methylarginine PTM[33]. While many of these methanogenesis markers are known to play essential roles in the function of MCR, the functions of some others remain unidentified.

The *mm4* (*mtxX*) gene which is a methanogenesis marker is also annotated as a methyltransferase. Some homologs of *mm4* operate as SAM-dependent methyltransferases. This informed the interest in the gene either as a sole culprit in the methylation of histidine, or as one of the genes involved in the formation of the methylhistidine.

Approximately 10^9 tons of methane are produced each year by methanogens, part of which contributes to the global methane load[3,93], therefore unveiling the secrets of MCR is a pivotal endeavor. Unraveling the genes responsible for MCR post-translational modifications and the effect of these modifications on MCR assembly and maturation is important to this overarching goal.

This chapter explores the steps that were taken to study the *mm4 (mtxX)* gene suspected to be responsible for the methylhistidine modification and to confirm the role of the *mcmA* gene in the methylation of cysteine using *in vivo* and *in vitro* approaches, as well as some computational analysis. Work on *mcmA* gene had been ongoing in the Mansoorabadi prior to the identification of the gene function by Nayak and his colleagues[62]. The *mcmA* gene was assigned the name *prmA* in the Mansoorabadi lab, and for the purpose of this work, both names (*mcmA* and *prmA*) will be used interchangeably. *Mm4* and *mcmA* genes were cloned into *E. coli* and the expressed and purified protein was used for *in vitro* assays to establish the enzyme function. McmA was also co-expressed with McrA protein to determine its influence on McrA expression and modification.

3.2 Methods

3.2.1 Construction of pETSUMO:*mcmA* and pETSUMO:*mm4* plasmids

The *MA1806* (*mm4*) and *MA4545* (*mcmA*) genes were amplified through polymerase chain reaction (PCR) using the genomic DNA extracted from *Methanosarcina acetivorans* C2A purchased from DSMZ in Germany (DSM-2834) as the gene template. Additionally, the pETSUMO vector was amplified using PCR. Gibson Assembly primers provided by Sigma-Aldrich were utilized in the PCR process to generate sticky ends, as outlined in Table 6. All PCR reactions were conducted using Phusion High-Fidelity DNA polymerase from New England Biolabs according to the manufacturer's instructions. Purification of the PCR products was carried out using a Lonza FlashGel^{TM} DNA Cassette from Basel, Switzerland, followed by recovery using the OMEGA Bio-tek E.Z.N.A.® Gel Extraction Kit from Norcross, GA.

For the construction of pETSUMO:*mm4* (Figure 19), the purified *mm4* gene and pETSUMO vector fragments were incubated with the Gibson Assembly Master Mix at 50 °C for 15 minutes. The resulting product was used to transform *E. coli* NEB 5α cells and plated on LB agar plates containing 50 µg/mL kanamycin, allowing overnight growth. Colonies were selected from the agar plate, cultured in liquid LB medium, and plasmids were extracted from the colonies using the Omega Bio-tek E.Z.N.A. ® Plasmid Mini Kit. The plasmid sequence was subsequently verified by Eurofins Genomics.

Similarly, pETSUMO:*mcmA* (Figure 19) was constructed following similar procedures, and the plasmid sequence was confirmed by Eurofins Genomics.

pETSUMO tagged genes are designed to incorporate a SUMO tag to the N-terminal amino acid residue of the protein to be expressed. Additionally, a 6-histidine tag is attached to the SUMO tag (and to the protein) which facilitates the purification of the protein using the nickel immobilized

metal affinity chromatography (Ni-IMAC) column. SUMO tags are efficiently cleaved off the recombinant protein by the action of sumo protease upon incubation with the SUMO-tagged protein. The sumo protease recognizes the Ubiquitin-like protein (SUMO) and cleaves in a very specific manner[94]. The resultant reaction mix is a combination of the cleaved recombinant protein, the sumo tag, some residuals of the uncleaved protein and the sumo protease protein. The recombinant protein without the tag passes through the Ni-IMAC column without being bound to it, while the other components bind to the column and are only eluted by addition of the elution buffer.

pACYCDuet-1:*prmA* creation was reported in Chapter 2 (see 2.2.1) and pETSUMO:*mcrA* had been constructed by a previous member of the Mansoorabadi lab[89].

Table 6: Sequence of primers used for construction of pETSUMO:mcmA and pETSUMO:mm4

pETSUMO:*mcmA*	Gene	Fwd: 5′-GATTGGTGGTatggaaataagatgtaggtgtg-3′ Rev: 5′-GCTTGTCTtcaaatcacaacaacttttctg-3′
	Vector	Fwd: 5′-ttgtgatttgaAGACAAGCTTAGGTATTTATTCG-3′ Rev: 5′-cttatttccatACCACCAATCTGTTCTCTG-3′
pETSUMO:*mm4*	Gene	Fwd: 5′-GATTGGTGGTatgtatgccctcctcg-3′ Rev: 5′-GCTTGTCTtcagcgaatcacctttttc-3′
	Vector	Fwd: 5′-ttcgctgaAGACAAGCTTAGGTATTTATTCG-3′ Rev: 5′-ggcatacatACCACCAATCTGTTCTCTG-3′

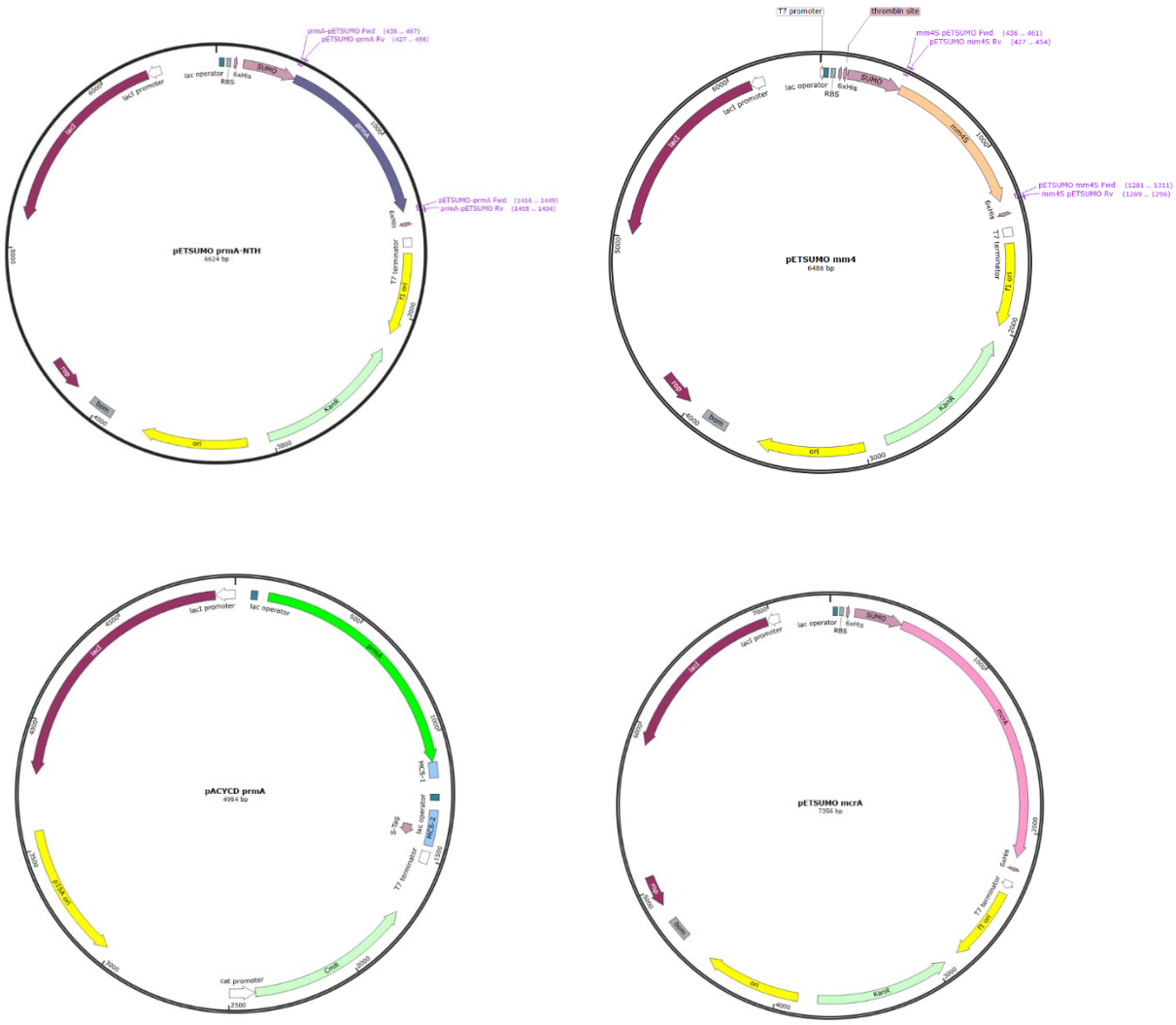


Figure 19: Plasmids used in the study of the mm4 and mcmA gene functions.

3.2.2 Aerobic expression of Mm4 and McmA proteins

Sequence verified pETSUMO:*mm4* was used to transform chemically competent *E. coli* BL21(DE3) cells. *E. coli* cells were made chemically competent using calcium chloride (see Appendix). The transformed cells were grown in LB supplemented with kanamycin (50 µg/mL final concentration) with shaking at 37 °C and 200 rpm. After 6 h, at OD_{600nm} of 0.6, cells were induced with 100 µM isopropyl β-D-1-thiogalactopyranoside (IPTG) and the temperature was lowered to 18 °C. Cells were left to shake at 90 rpm for 18 h and were harvested by centrifugation at 5000 rpm. The cell pellets were stored at -80 °C.

The pETSUMO:*prmA* plasmid was used to transform chemically competent *E. coli* BL21(DE3) cells. Transformed cells were grown in kanamycin (50 µg/mL final concentration) supplemented lysogeny broth (LB) and with shaking at 200 rpm and 37 °C for 6 hours. Cells were induced with 100 µM IPTG at OD_{600nm} of 0.6 and the temperature was lowered to 18 °C. Cells were left to shake at 90 rpm for 18 hours and were harvested.

3.2.3 Creation of McrA cell lines

Two McrA cell lines were created:

- *E. coli* BL21(DE3) cell line containing pETSUMO:*mcrA* and
- *E. coli* BL21(DE3) cell line containing pETSUMO:*mcrA* and pACYCDuet-1:*prmA*

Chemically competent *E. coli* BL21(DE3) cells were transformed first with pETSUMO *mcrA*. The transformed cells were plated on LB agar plates supplemented with 50 µg/mL of kanamycin and left to grow overnight. Colonies were selected from the agar plate, cultured in liquid LB medium containing kanamycin, and plasmids were extracted from the colonies using the Omega Bio-tek E.Z.N.A. ® Plasmid Mini Kit. The plasmid was subsequently verified by amplification using *mcrA*

PCR primers, and identification of the *mcrA* gene sized bands on the Lonza FlashGel^{TM} DNA Cassette.

PCR verified *E. coli* BL21(DE3) cell line containing pETSUMO:*mcrA* plasmid was made chemically competent using calcium chloride (see Appendix) and transformed with the pACYCDuet-1:*prmA* plasmid. Again, transformed cells were plated on agar plates containing 50 µg/mL of kanamycin and 35 µg/mL of chloramphenicol and left to grow overnight. Colonies were picked and cultured in LB medium containing kanamycin and chloramphenicol and plasmids were extracted from the colonies. The presence of the *mcrA* and *prmA* genes were verified by PCR amplification.

3.2.4 Anaerobic expression of McrA cell lines

The *E. coli* BL21(DE3) cell line containing pETSUMO:*mcrA* and pACYCDuet-1:*prmA* was grown in broth supplemented with kanamycin (50 µg/mL final concentration) and chloramphenicol (35 µg/mL final concentration) with shaking at 37 °C and 200 rpm. At OD_{600nm} of 0.6, cells were induced with 200 µM IPTG and supplemented with 0.2 mM iron (II) sulfate, 0.15 mM cysteine, 0.1 mM cyanocobalamin, 0.2 mM riboflavin, and 0.2 mM methionine, the temperature was lowered to 18 °C, and shaking continued at 90 rpm for 18 h. After 18 h, the cells were transferred to anaerobic flasks, placed in an ice-bath, and subjected to argon pressurization for 2 – 4 h to remove all traces of oxygen from the broth. Deoxygenated cells were left overnight (12 – 18 h) standing at 4 °C and then harvested anaerobically. The cells were transferred into an anaerobic glove box and measured into centrifugation tubes for centrifugation.

E. coli BL21(DE3) cells transformed with only pETSUMO:*mcrA* and grown in broth supplemented with kanamycin (50 µg/mL final concentration) at 37 °C and 200 rpm, and subjected

to the expression protocol outlined above. Cell pellets were harvested in the glove box and stored in -80 °C until use.

3.2.5 Protein purification using Nickel Immobilized Metal Affinity Chromatography (Ni-IMAC)

Harvested cell pellets were thawed at room temperature. They were resuspended in lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole). The cells were lysed by sonication and the resultant solution was centrifuged at 7, 500 × g for 2 h. The decanted supernatant was applied to Bio-Rad Econo-Pac column packed with Ni^{2+}-charged Profinity IMAC Resin. The column was then washed with 50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, pH 8.0 buffer. The protein was eluted with 50 mM sodium phosphate, 300 mM NaCl, 500 mM imidazole, pH 8.0 buffer, and buffer exchanged with 100 mM Tris-HCl, pH 8.0 buffer. All purification fractions including flowthrough (FT), wash (W), elution (E) and insoluble (IN) fractions were stored for further analysis. Eluted proteins were stored at -80 °C. Purification fractions were subjected to sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and western blot for protein identification.

Aerobically purified Mm4 proteins were subjected to a SUMO-tag cleavage reaction by incubation with a histidine tagged SUMO protease and DTT for 18 h. Protein and SUMO protease were incubated in a 1:1 ratio and 1% dithiothreitol (DTT) was added to the reaction mixture. The reaction mixture was applied to a Bio-Rad Econo-Pac column packed with Ni^{2+}-charged Profinity IMAC Resin. The column was then washed with the wash buffer (50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, pH 8.0 buffer), and tagged proteins still bound to the column were eluted with the elution buffer (50 mM sodium phosphate, 300 mM NaCl, 500 mM imidazole, pH

8.0 buffer). Wash and elution fractions were buffer exchanged using 100 mM Tris-HCl, pH 8.0 buffer. Aliquots of all collected fractions were used for SDS-PAGE analysis.

Purification of proteins from all cell lines containing the pETSUMO:*mcrA* plasmid was carried out in the anaerobic chamber using anaerobic buffer solutions.

3.2.6 Design and construction of Mm4 and McmA peptide substrates

The Mm4 peptide substrate was designed to mimic the 1-*N*-methylhistidine methylation site (his271) of McrA. Five amino acids to the left and to the right of the active histidine residue were added to create the peptide. The peptide SFAAK**H**AALVS (with a molecular weight of 1100 Da) was synthesized using a rink amide resin on a 0.25 mmol scale. The fluorenylmethoxycarbonyl (Fmoc)-protected resin (0.5 g of 0.5 mM resin) was swollen with dichloromethane for 2 h and subsequently deprotected by shaking with 20% piperidine in dimethylformamide (DMF) for 5 min.

Amino acids were added sequentially starting from the last amino acid (C-terminal) to the first one (N-terminal). 1.25 mmol of the first Fmoc protected amino acid (S) was dissolved in DMF along with coupling agents 1-hydroxy-7-azabenzotriazole (HOAt) (1.25 mmol) and dissolved inorganic carbon (DIC) (1.25 mmol), added to the resin, and left to stir on a wrist action shaker for 15 min. The rest of the Fmoc protected amino acids (0.75 mmol) in the sequence were sequentially coupled to the resin using 0.75 mmol hexafluorophosphate benzotriazole tetramethyl uronium (HBTU) and 1.50 mmol N,N-diisopropylethylamine (DIEA), which were added to the resin and stirred for 5 min. All Fmoc protected amino acids underwent deprotection by shaking with 20% piperidine in dimethylformamide (DMF) for 5 min before addition of the next amino acid.

Amino acids ser, his, and lys had side groups that were protected by acid labile trityl (Tr), Tr, and tert-butyloxycarbonyl (Boc) groups, respectively.

Peptide was cleaved from the resin using a cocktail mix of trifluoroacetic acid (92.5 %), triethylsilane (2.5 %), dimethoxybenzene (2.5 %), and water (2.5 %). Cleaved peptide was air dried, dissolved in a 1:1 ratio mixture of water and acetonitrile (ACN) and purified. The purification was performed using high performance liquid chromatography (HPLC) on an Agilent 1100/1200 series HPLC equipped with an Eclipse Plus C18 (4.6 x 150 mm, 3.5 µm particle size) reversed phase column. All separations involved a mobile phase of 0.1% formic acid (v/v) in water (solvent A) and 0.1% formic acid (v/v) in acetonitrile (solvent B). The HPLC method employed a linear gradient of 0–80% solvent B over 30 minutes at ambient temperature with a flow rate of 1.0 mL min^{-1}. The separation was monitored by UV absorbance at 220 nm. The peak containing the peptide eluted at 8 min and was collected using 10 mL tubes.

Purified peptide was characterized by liquid chromatography-mass spectrometry (LC-MS). It was lyophilized with a Labconco lyophilizer and stored at 4 °C.

Lyophilized McmA (PrmA) peptide with sequence DLQDQ**C**GATNV was purchased from Biomatik. The McmA peptide was designed to mimic the methylcysteine methylation site (cys472) of McrA.

3.2.7 Mm4 and McmA structural prediction studies

Using Alphafold2 v1.5.2[95,96], the structures of the Mm4 and McmA (PrmA) proteins, as well as their complexes with their respective peptide substrates were generated. The PyMOL molecular visualization system[97] was used to analyze the resulting structures.

Two putative metal binding sites consisting of four cysteine residues each were identified in the McmA structure. Zinc (Zn^{2+}) and iron (Fe^{2+}) ions were manually inserted into the structure by editing the PDB file of McmA. Structural homologs of McmA were derived through a Dali search.

The search algorithm assigns a Z-score to structural homologs, and a Z-score lower than 2 signifies that the structural homology of that protein is spurious. The structure of *S*-adenosyl-L-homocysteine (SAH) was sourced from the crystal structure of N^2,N^2-dimethylguanosine tRNA methyltransferase in complex with S-adenosyl-L-homocysteine (SAH) (PDB ID: 2eju-A)[98]. The protein was chosen because it is a structural homolog to McmA, and it is a SAM-dependent methyltransferase which possesses a C-terminal domain containing a typical alpha/beta nucleotide-binding fold found in methyltransferases[99]. This fold participates in SAM binding and was detected in the alpha-fold generated structure of McmA. The McmA-SAH complex was generated by using the alignment function in PyMOL to align McmA to the SAH bound 2eju-A. Interactions between McmA and SAH were then determined.

3.2.8 Mm4 activity assay

Methylation activity of Mm4 was carried out aerobically. 100 μM of the Mm4 peptide was incubated with 100 μM of SAM and 40 μM of the Mm4 protein in 50 mM HEPES buffer pH 8.0 for 18 h. The reaction was quenched with equal volume of methanol incubation for 20 minutes and centrifuged for 20 minutes to separate precipitated Mm4 enzyme from solution. The supernatant was decanted and used for HPLC analysis.

3.2.9 McmA activity assay

The methylation activity of McmA (PrmA) was first investigated using aerobically purified McmA protein. 10 μM of protein was incubated with 100 μM of the peptide and 100 μM of SAM for 18 hours in 100 mM Tris-HCl buffer, pH 8.0. Reaction was quenched with equal volume of methanol for 20 minutes and centrifuged for 20 minutes. The supernatant was decanted and used for mass spec and HPLC analysis.

Anaerobically purified McmA (20 μM) was then incubated for 18 hours with the peptide (100 μM) and SAM (100 μM) with the addition of DTT (1%) and zinc sulphate (50 μM) in 100 mM Tris-HCl buffer, pH 8.0. A no-enzyme control was also set up without McmA. The reaction mixture was quenched with equal volumes of methanol. Each aliquot of the quenched reaction was subjected to centrifugation and the supernatant was collected and analyzed using HPLC and MS.

3.2.10 McmA activity assays using adenosylhomocysteinase

Adenosylhomocysteinase *E. coli* BL21(DE3) cell line was a generous gift from Prof. William Metcalf's laboratory (University of Illinois Urbana-Champaign). The cell line was grown, harvested, and purified by Ni-IMAC method using the aerobic expression and aerobic purification protocols described above.

McmA activity assays were repeated with the addition of adenosylhomocysteinase. The reactions were carried out anaerobically and contained the following: anaerobically purified McmA (20 μM), McmA peptide (100 μM), SAM (200 μM), adenosylhomocysteinase (20 μM), DTT (1%), and zinc sulphate (50 μM) in 100 mM Tris buffer, pH 8.0. McmA peptide, SAM, DTT, zinc sulphate and aliquots of adenosylhomocysteinase to be used for the reaction were made anaerobic by transferring them in 1.5 mL reaction tubes into the glove box 6 hours before the reaction and leaving the tubes open. A control group was also set up without McmA protein. Aliquots of the reaction mixture were quenched with equal volumes of methanol at 1 h, 2 h, and 18 h time intervals. Each aliquot of the quenched reaction was subjected to centrifugation and the supernatant was collected and analyzed using HPLC and mass spectrometry.

3.2.11 HPLC analysis

HPLC analysis was performed on an Agilent 1100/1200 Infinity Quaternary LC System with Diode Array Detector (DAD) VL+, equipped with an Agilent Poroshell 120 EC-C18 (4.6 x 150

mm, 2.7 μm particle size) reverse phase column. All separations involved a mobile phase of 0.5 % formic acid (v/v) in water (solvent A) and 0.5 % formic acid (v/v) in acetonitrile (solvent B). The HPLC method employed a linear gradient of 0–24 % solvent B over 24 min, 24 – 100 % solvent B over 1 min, 100 % solvent B for 8 min, 100 – 0 % solvent B over 1 min, and 0 % B for 5 min. Flow was at ambient temperature with a flow rate of 1.0 mL min^{-1}. The separation was monitored by UV-vis absorbance at 256 nm.

3.2.12 Mass spectrometry analysis

LC-MS analyses on the Ni-IMAC elution fractions from McrA cell lines were performed on a Thermo-Fisher LTQ Orbitrap Elite Mass Spectrometer coupled with a Proxeon Easy NanoLC system (Waltham, MA) located in the Proteomics and Mass Spectrometry Facility at the University of Georgia.

The tryptic peptides were resuspended in 0.1% formic acid/2% acetonitrile. A self-packed column with reversed phase resin (Dr. Maisch ReproSil-pur C18-AQ, 3 μm particle size, 120 Å pore size, 100 μm id ~120 mm) was coupled to the mass spectrometer. Peptides were separated and eluted into the electrospray ionization source at a flow rate of 300 nL/min. Briefly, the two-buffer gradient elution (water with 0.1% formic acid as buffer A and acetonitrile with 0.1% formic acid as buffer B) starts with 0% B for 2 min, linearly increases to 12% B over 25 min, to 30% B over 25 min, to 50% B over 10 min, and finally to 95% B over 10 min.

The data-dependent acquisition (DDA) method was used to acquire MS data. A list of expected precursor ions was generated for monitoring possible N-terminal peptides. A survey MS scan was acquired first (350-1500 *m*/*z*), and then the top 10 ions in the precursor list were selected for CID MS/MS analysis with an isolation width of 2 *m*/*z*. If no peptide ions from the precursor list were

found, the most abundant ions are chosen for MS/MS analysis. Both MS and MS/MS scans were acquired by Orbitrap at resolutions of 120,000 and 15,000, respectively.

Data were acquired using Xcalibur software (version 3.0, Thermo Fisher Scientific). Protein identification and modification characterization were performed using Thermo Proteome Discoverer (version 3.0) with the Mascot (Matrix Science, London, UK) and Uniprot Salmonella databases. The search parameters include:

- Precursor Mass Tolerance: 10 ppm
- Fragment Mass Tolerance: 0.02 Da
- Modification: Oxidation of methionine
- Validated with Percolator (decoy database)
- Targeted FDR (restrict/relax): 0.01/0.05

In the Xcalibur QaulBroswer, the extracted ion chromatograms of monitored N-terminal peptides were plotted by combining peptide ions at various charges within the mass tolerance of 5 ppm in MS1. ICIS peak detection in Xcalibur is the default method to integrate the peak area and peak intensity at the apex.

3.3 Results

3.3.1 McmA and Mm4 purification

E. coli BL21(DE3) cells transformed with pETSUMO:*mcmA* and *E. coli* BL21(DE3) cells transformed with pETSUMO:*mm4* were expressed, harvested, lysed and purified using a Ni-IMAC column. Purified McmA (PrmA) protein (46 kDa) was detected in the elution fraction by SDS-PAGE and western blot (Figure 20). A band matching the size of the purified Mm4 protein (41 kDa) was also detected in the elution fraction of the SDS-gel and western blot (see Figure 21).

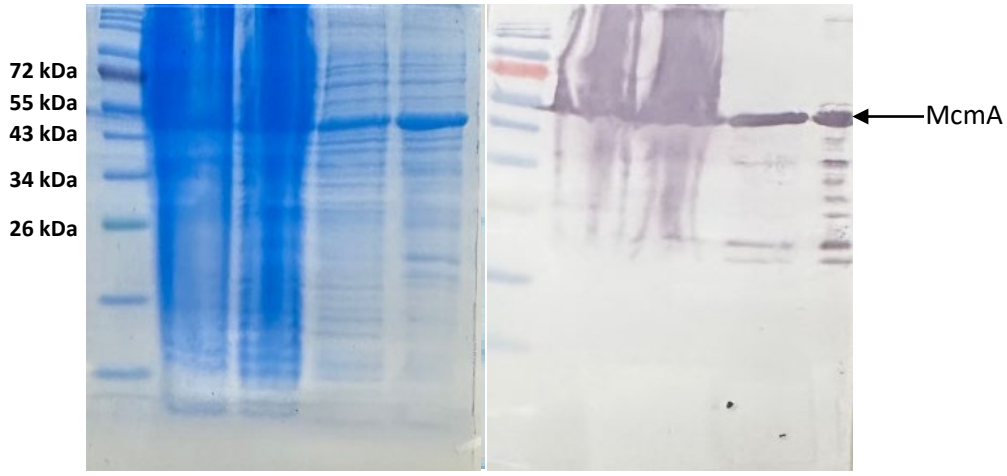


Figure 20: SDS-PAGE gel and western blot of purified fractions of McmA protein

Lanes 1 - 5 represent the molecular weight standards, insoluble, flowthrough, wash and elution fractions of the McmA protein

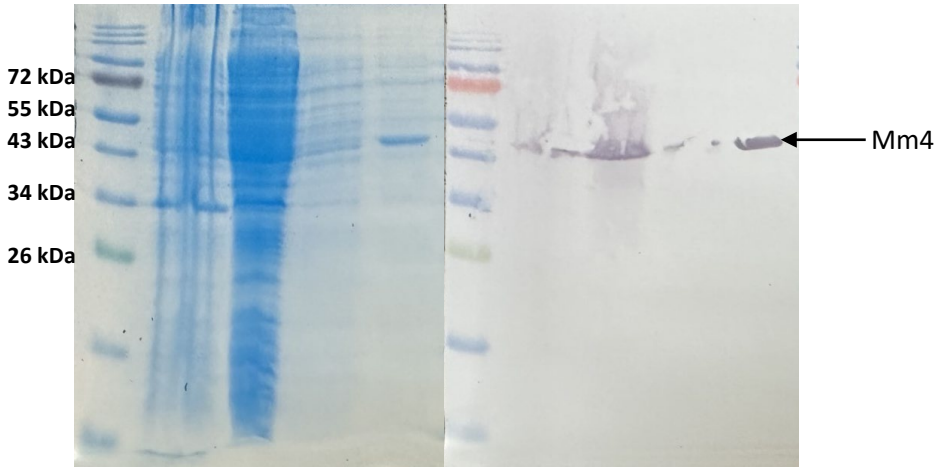


Figure 21: SDS-PAGE and western blot of purified fractions of Mm4 protein

Lanes 1 - 5 represent the molecular weight standards, insoluble, flowthrough, wash and elution fractions of the Mm4 protein

3.3.2 Mass spec result of McmA peptide purchased from Biomatix

The peptide purchased from Biomatix (DLQDQCGATNV) was analyzed using mass spectrometry and the *m/z* value was detected as shown in Figure 22.

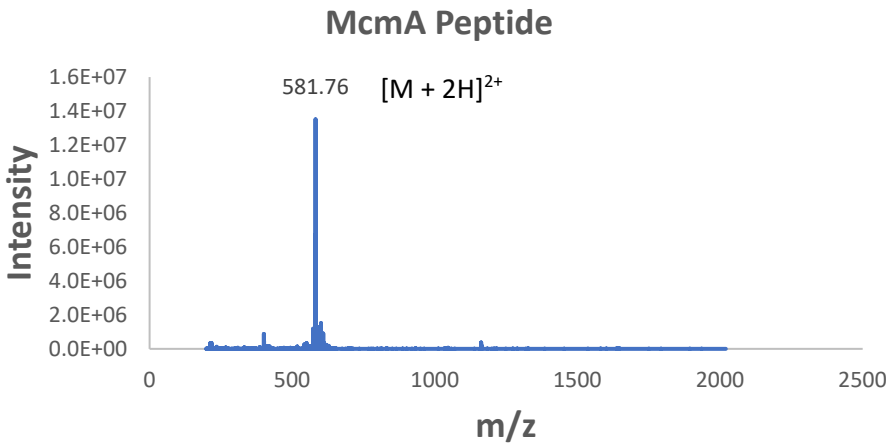


Figure 22: Mass spec data of McmA peptide (1163.52 m/z value)

3.3.3 McmA structure shows zinc binding sites

Initial McmA assays were carried out and the methylated peptide was not detected by mass spectrometry. It was surmised that the protein might require accessory co-factors for its activity. McmA structural prediction studies were carried out to glean insight on enzyme function as it relates to its structure. McmA structure was determined using AlphaFold and analysis of the McmA structure revealed the presence of several cysteine residues. The presence of 16 cysteine residues on the McmA structure suggests that the protein will be very sensitive to oxygen and may purify better anaerobically. In particular, some of the cysteine residues had their sulfur residues oriented towards an imaginary center as though forming a metal binding site (Figure 23). The possibility of a metal binding to this imaginary center was tested first with zinc and then with iron. The zinc

metal seemed to be a better fit than the iron in the metal binding site. Two molecules of the metal were required for each molecule of the protein (Figure 23).

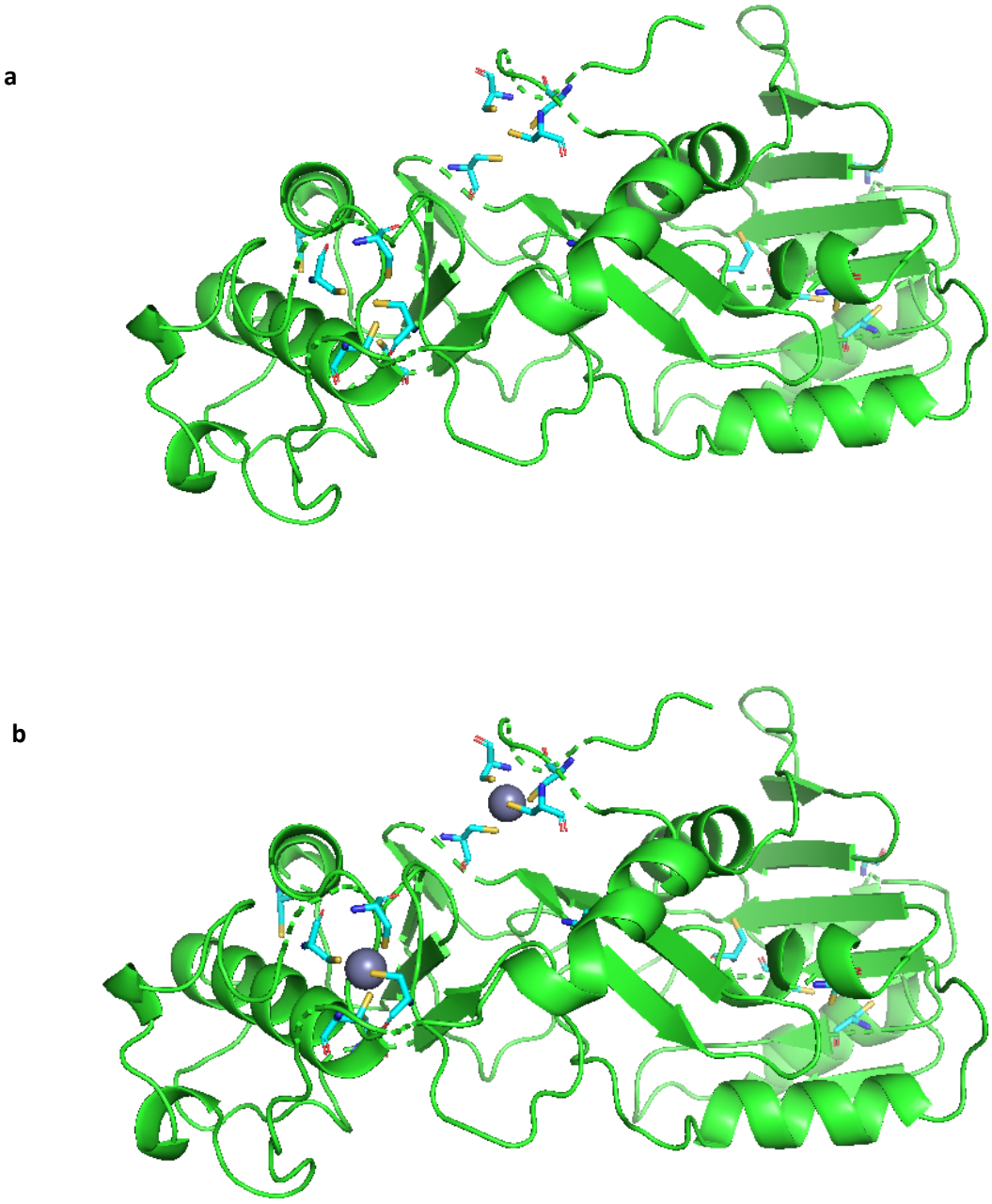


Figure 23: Alphafold structures of McmA protein showing (a) its cysteine residues shown as sticks and (b) McmA binding to zinc

3.3.4 McmA expression with zinc sulfate and anaerobic purification

Prior to McmA structural analysis, McmA purification was good but not great. Truncated versions of the protein were observed in its SDS-gel and western blot membrane. It is possible that metal binding was required to stabilize the protein and improve purification yield. Data derived from McmA structural analysis informed the expression of McmA in LB supplemented with zinc sulfate (Zn_2SO_4) at induction, as well as McmA anaerobic purification. *E. coli* BL21(DE3) cell line containing pETSUMO:*mcmA* was also expressed with the addition of iron sulfate and purified anaerobically. Eluted protein was cleaved with sumo protease and the cleaved fractions from both cell lines were subjected to SDS-PAGE and western blot.

The result of the SDS-gel suggested that McmA protein was expressed with more yield when supplemented with zinc sulfate (Figure 24). The molecular weight of McmA is 36 kDa.

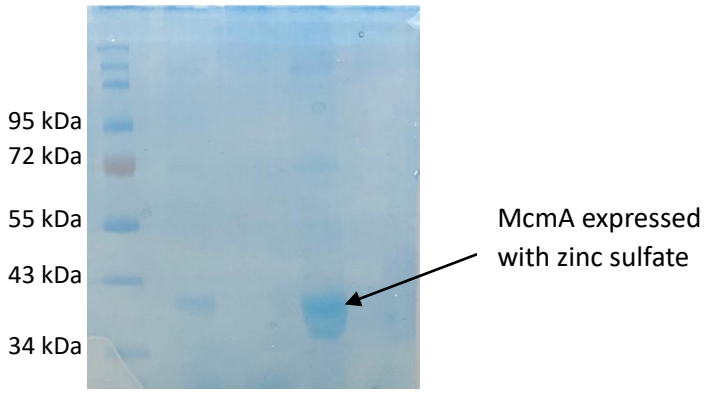


Figure 24: SDS-PAGE gel of McmA expressed with zinc sulfate and with iron sulfate

Lanes 1 - 5 represent the molecular weight standards, cleaved McmA supplemented with iron, sumo cleavage elution fraction of McmA supplemented with iron, cleaved prmA supplemented with zinc, sumo cleavage elution fraction of prmA supplemented with zinc.

3.3.5 McmA activity assays

McmA activity assays were repeated with the anaerobically purified McmA expressed with zinc sulfate supplementation. The assay was based on the detection of SAH (Figure 25) and the methylated peptide (1177.52 m/z) by HPLC and by mass spectrometry respectively. The protein was incubated with SAM, Zn_2SO_4 and the peptide overnight and the sample was subjected to methanol extraction with subsequent decanting of the supernatant for HPLC and mass spec analysis. Neither the methylated peptide nor SAH were detected by mass spectrometry (Figure 28 and Figure 29) or by HPLC (Figure 26 and Figure 27).

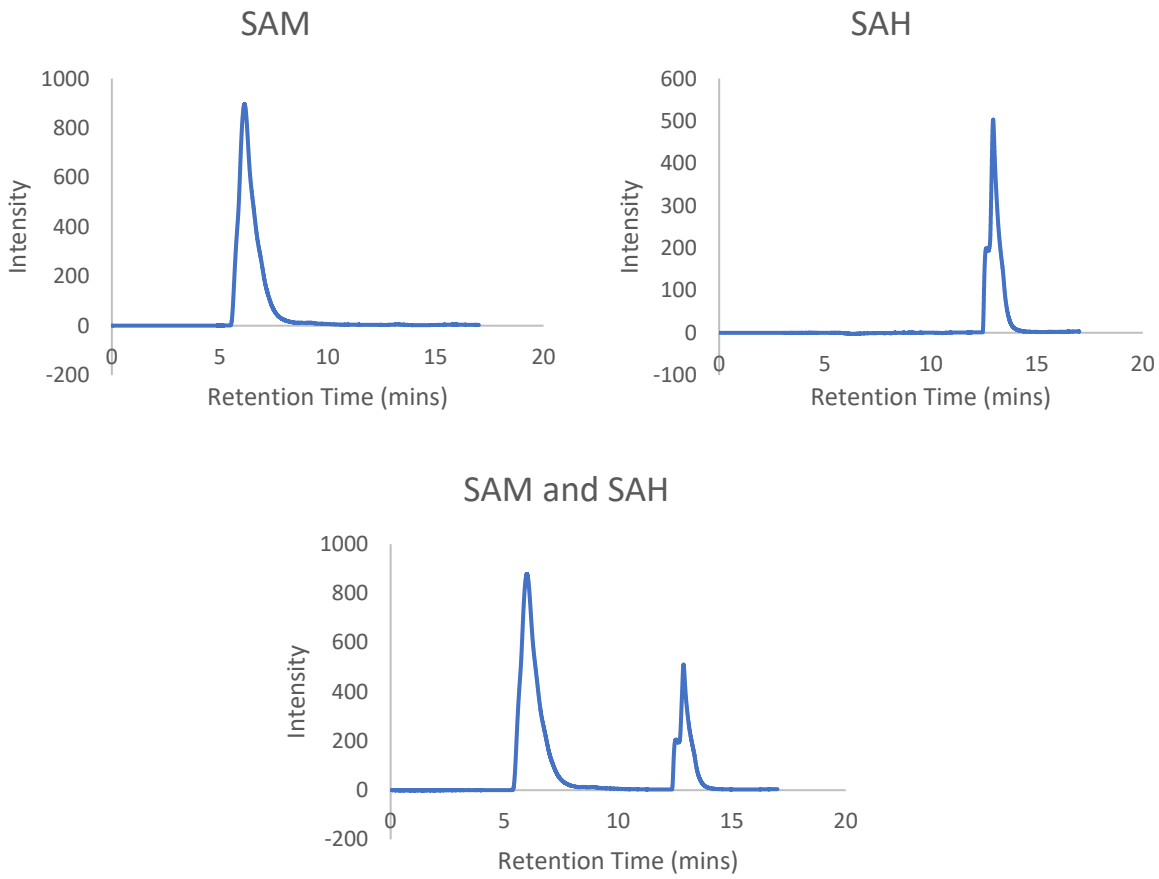


Figure 25: HPLC analysis with SAM and SAH standards showing their separation profiles

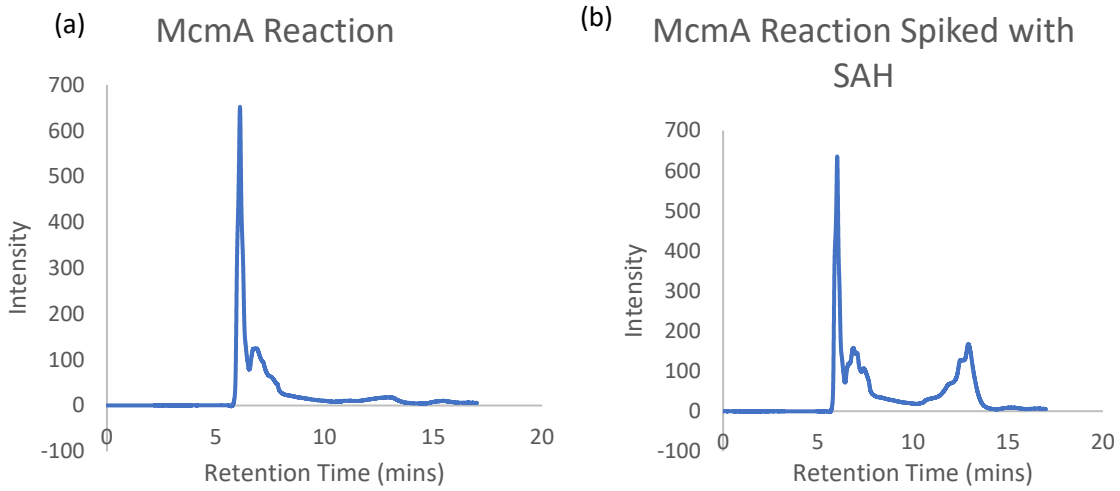


Figure 26: HPLC analysis of McmA reaction (a) and McmA reaction spiked with SAH (b). Reaction contained McmA (20 µM), McmA peptide (100 µM), SAM (100 µM), DTT (1%) and zinc sulphate (50 µM) in 100 mM Tris buffer, pH 8.0, and was incubated for 18 hours. Reaction was spiked with 100 µM SAH in (b)

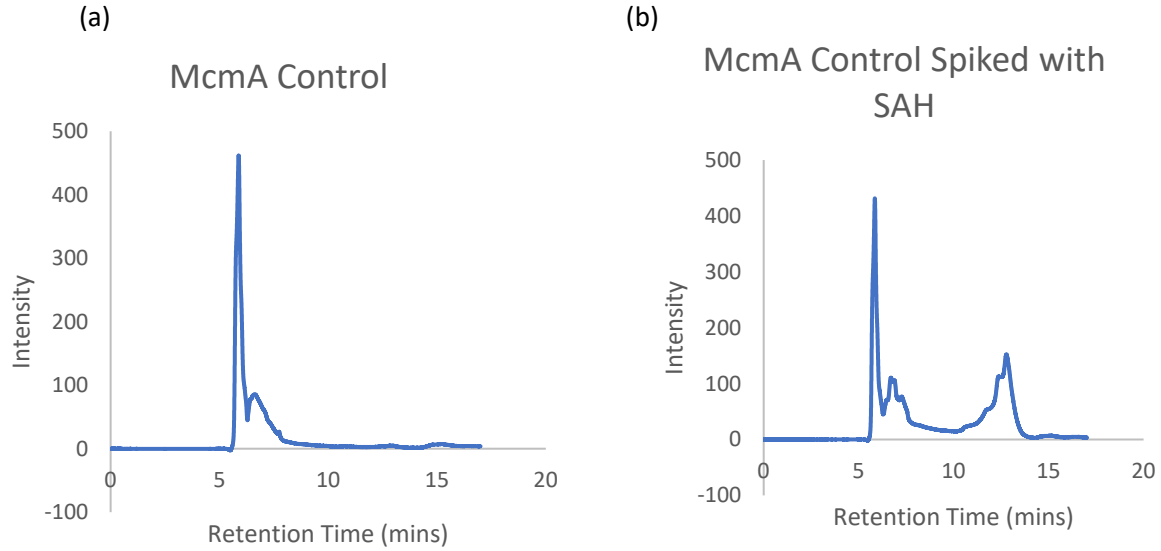


Figure 27: HPLC analysis of McmA control (a) and McmA control spiked with SAH (b). Reaction contained McmA peptide (100 µM), SAM (100 µM), DTT (1%) and zinc sulphate (50 µM) in 100 mM Tris buffer, pH 8.0, and was incubated for 18 hours. Reaction was spiked with 100 µM SAH in (b).

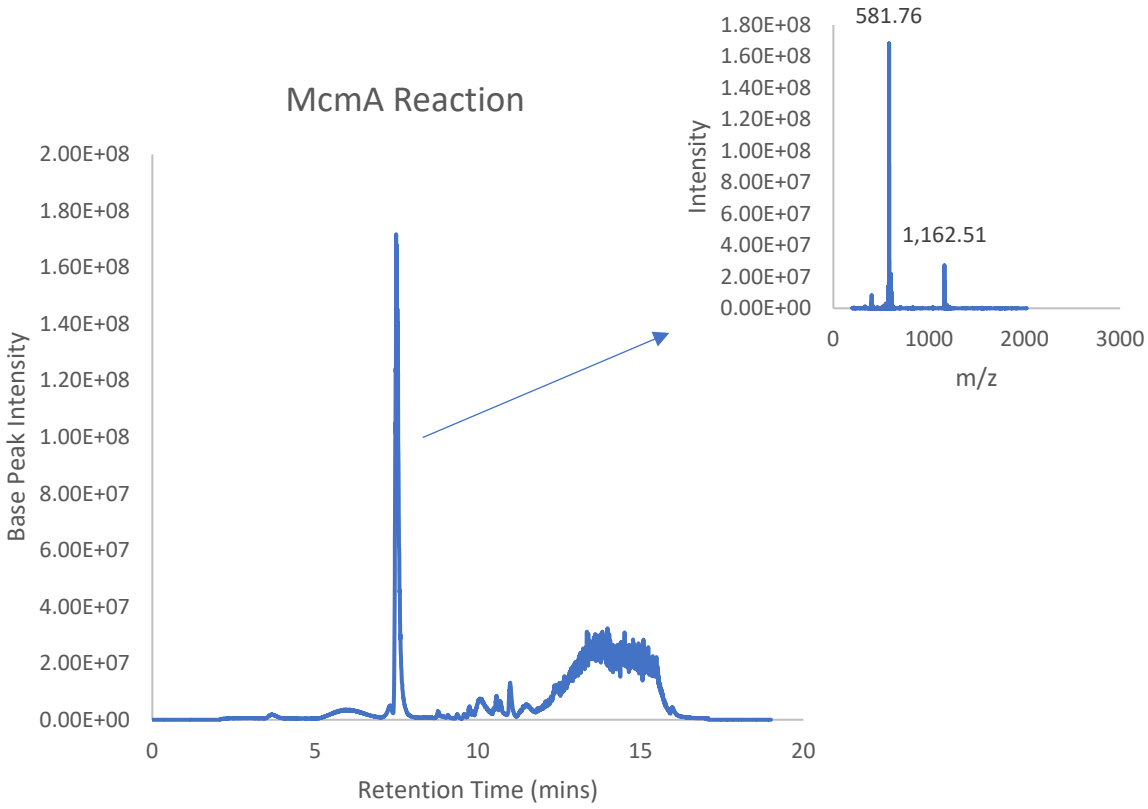


Figure 28: Mass spec analysis of McmA reaction. Reaction contained McmA (20 µM), McmA peptide (100 µM), SAM (100 µM), DTT (1%) and zinc sulphate (50 µM) in 100 mM Tris buffer, pH 8.0, and was incubated for 18 hours.

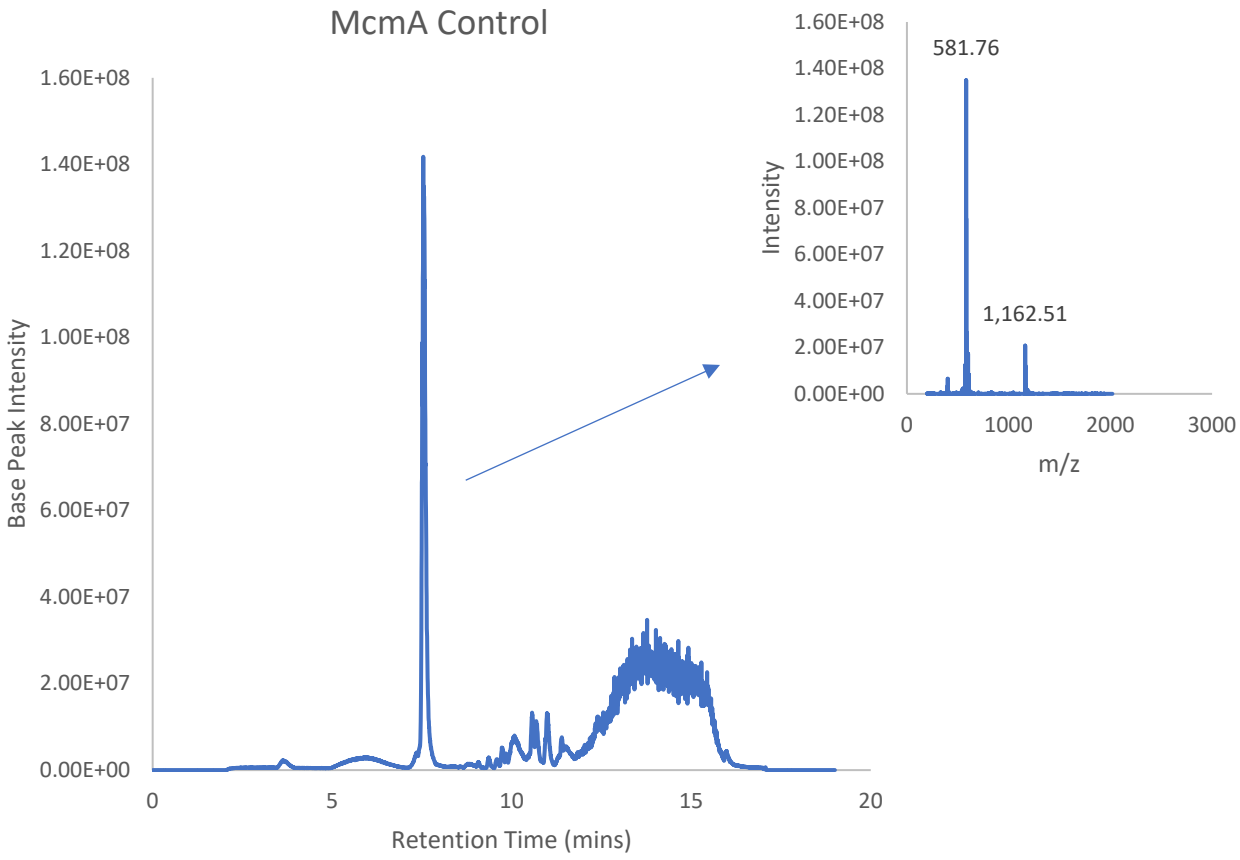


Figure 29: Mass spec analysis of McmA control group. Reaction contained McmA peptide (100 µM), SAM (100 µM), DTT (1%) and zinc sulphate (50 µM) in 100 mM Tris buffer, pH 8.0, and was incubated for 18 hours

3.3.6 McmA activity assays with adenosylhomocysteinase

Activity assays were repeated with the addition of adenosylhomocysteinase. 20 µM of purified adenosylhomocysteinase was added to the McmA reaction with the aim of combating possible SAH inhibition of McmA, which is common in methyltransferase[100]. The reaction was conducted anaerobically. Mass spectrometry results did not detect the methylated peptide (Figure 30 and Figure 31).

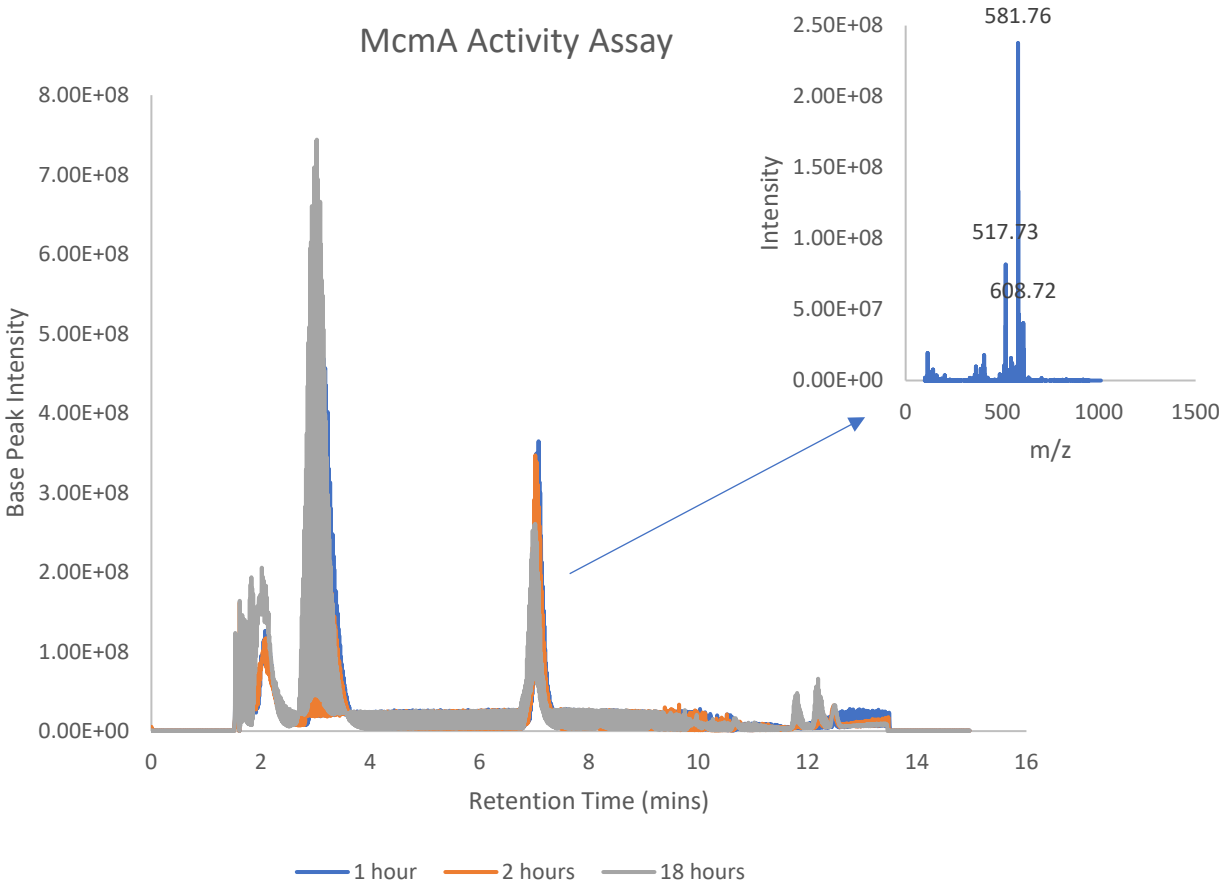


Figure 30: HPLC and mass spec result of the reaction group of McmA time course activity assay. Reaction contained McmA (20 µM), McmA peptide (100 µM), SAM (200 µM), adenosylhomocysteinase (20 µM), DTT (1%), and zinc sulphate (50 µM) in 100 mM Tris buffer, pH 8.0

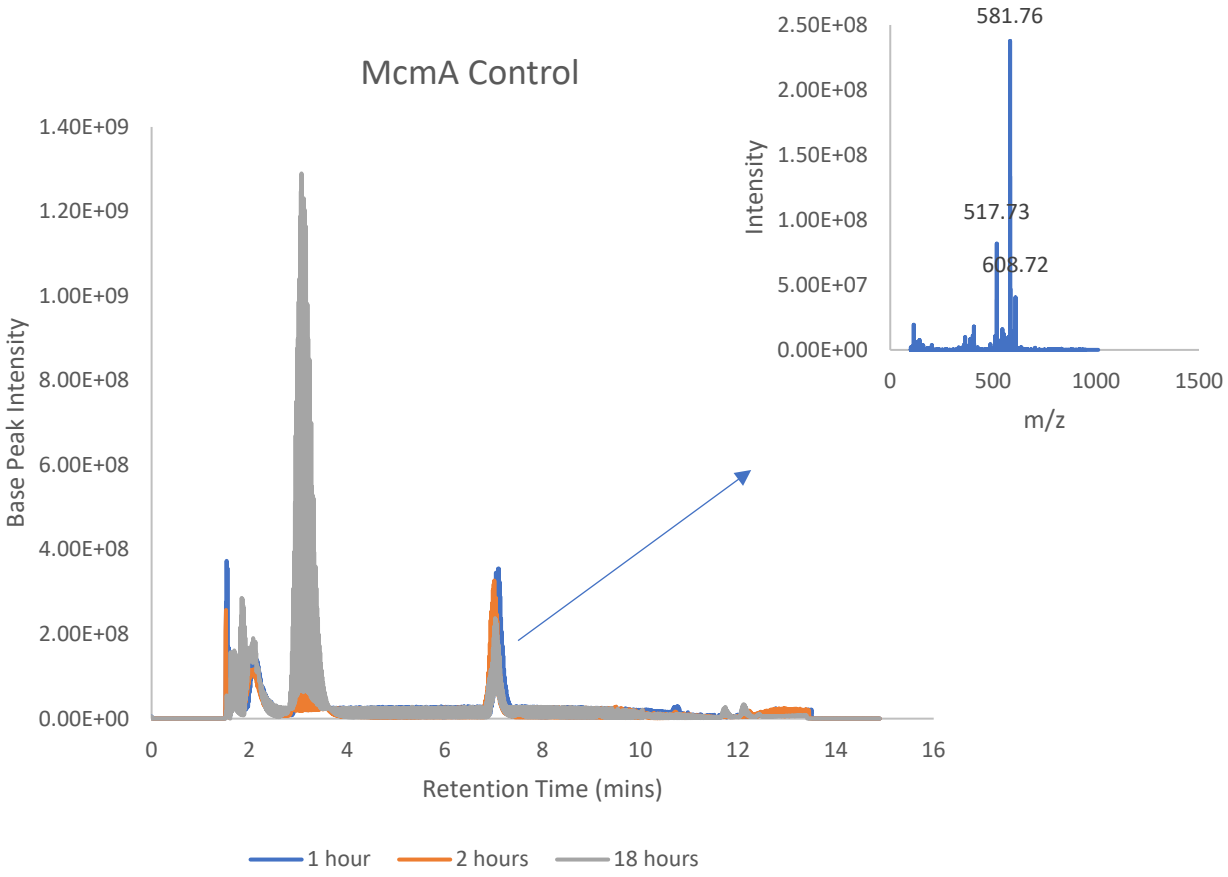


Figure 31: HPLC and mass spec result of the control group of McmA time course activity assay. Reaction contained McmA peptide (100 µM), SAM (200 µM), adenosylhomocysteinase (20 µM), DTT (1%), and zinc sulphate (50 µM) in 100 mM Tris buffer, pH 8.0

3.3.7 McmA structural prediction studies

The inability to detect either SAH or the methylated peptide by the methods employed so far prompted structural prediction studies to determine whether the peptide was interacting with the McmA active site. The peptide substrate is a short analog of the actual substrate (McrA) and may not be binding properly to the enzyme active site.

McmA structure was determined using AlphaFold and Dali search[101] was carried out to identify structurally homologous proteins to McmA (Table 7). The structural homolog dimethylguanosine tRNA methyltransferase (PDB ID: 2eju-A) was chosen because it was co-crystallized with S-adenosylhomocysteine (SAH). Alignment of the McmA structure (Figure 32) with 2eju-A, revealed that McmA contains a putative SAM binding site which is the enzyme's active site. The cys472 residue of the peptide substrate was found to be at a distance of 17.8 Å to the sulfur of SAH in the McmA structure (Figure 33). This shows that the peptide may not be within a close distance for catalysis. Another McrA peptide of 61 amino acid length was found within a distance of 12.8 Å. Further studies might explore using the longer peptide for McmA activity assays.

Analyzing the structural prediction of the 61 amino acid length McrA peptide provided evidence for another possibility – that the methylation may be co-translational rather than post-translational. The hypothesis of co-translational modification was strengthened by comparing residues around the cys472 residue in both the 61 amino acid length peptide and in McrA. Residues after cys472 in the peptide were less tightly coiled than that of McrA (Figure 34). It is possible that within the cell, cys472 was methylated before the complete folding of McrA.

To test this hypothesis, in-vivo assays were carried out using two cell lines – one cell line harbored only an *mcrA* containing plasmid while the other cell line harbored an *mcrA* and *mcmA* containing plasmid.

Table 7: Dali search result showing proteins with structural similarity to McmA

No	Chain	Z-score	Description
1	2ytz-B	15.8	Dimethylguanosine tRNA
2	2eju-A	15.7	Dimethylguanosine tRNA methyltransferase
3	2ytz-A	14.9	Dimethylguanosine tRNA
4	7t39-A	14.9	Protein arginine N-methyltransferase 9
5	6arj-A	14.8	Histone-arginine methyltransferase CARMI
6	4m38-A	14.8	Protein arginine N-methyltransferase 7
7	6p7i-B	14.8	Protein arginine N-methyltransferase 6
8	4c03-B	14.7	Protein arginine N-methyltransferase 6
9	5dx0-A	14.6	Histone-arginine methyltransferase CARMI
10	4m36-A	14.6	Protein arginine N-methyltransferase 7

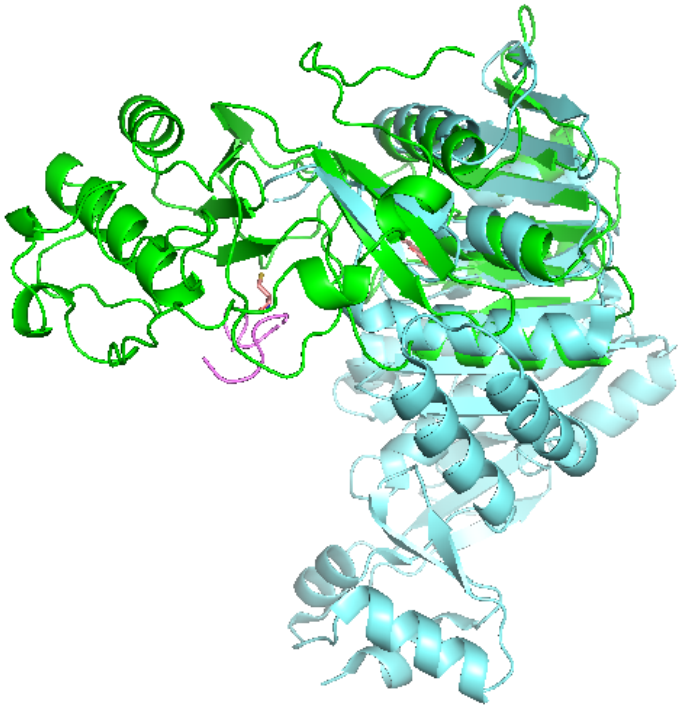


Figure 32: Alignment of McmA (Green) and dimethylguanosine tRNA methyltransferase {2eju-A} (cyan)

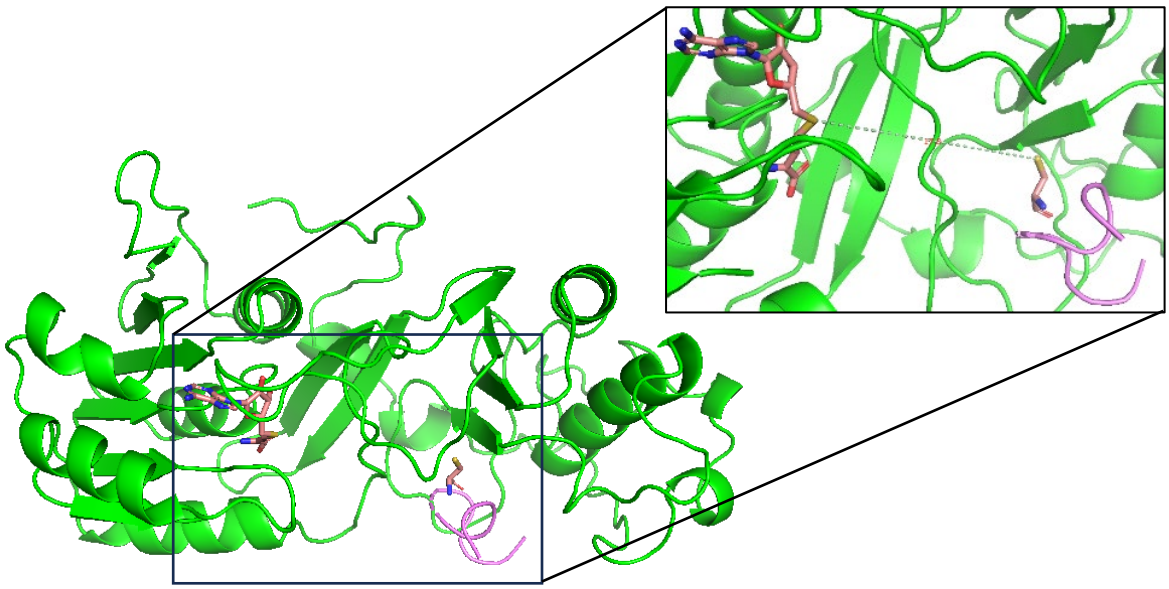


Figure 33: AlphaFold structure of McmA protein showing the peptide and the protein active site where SAH (shown in sticks) is bound. Inset shows a distance of 17.8 Å between the sulfur of the Cys472 (shown in sticks) and sulfur of SAH.

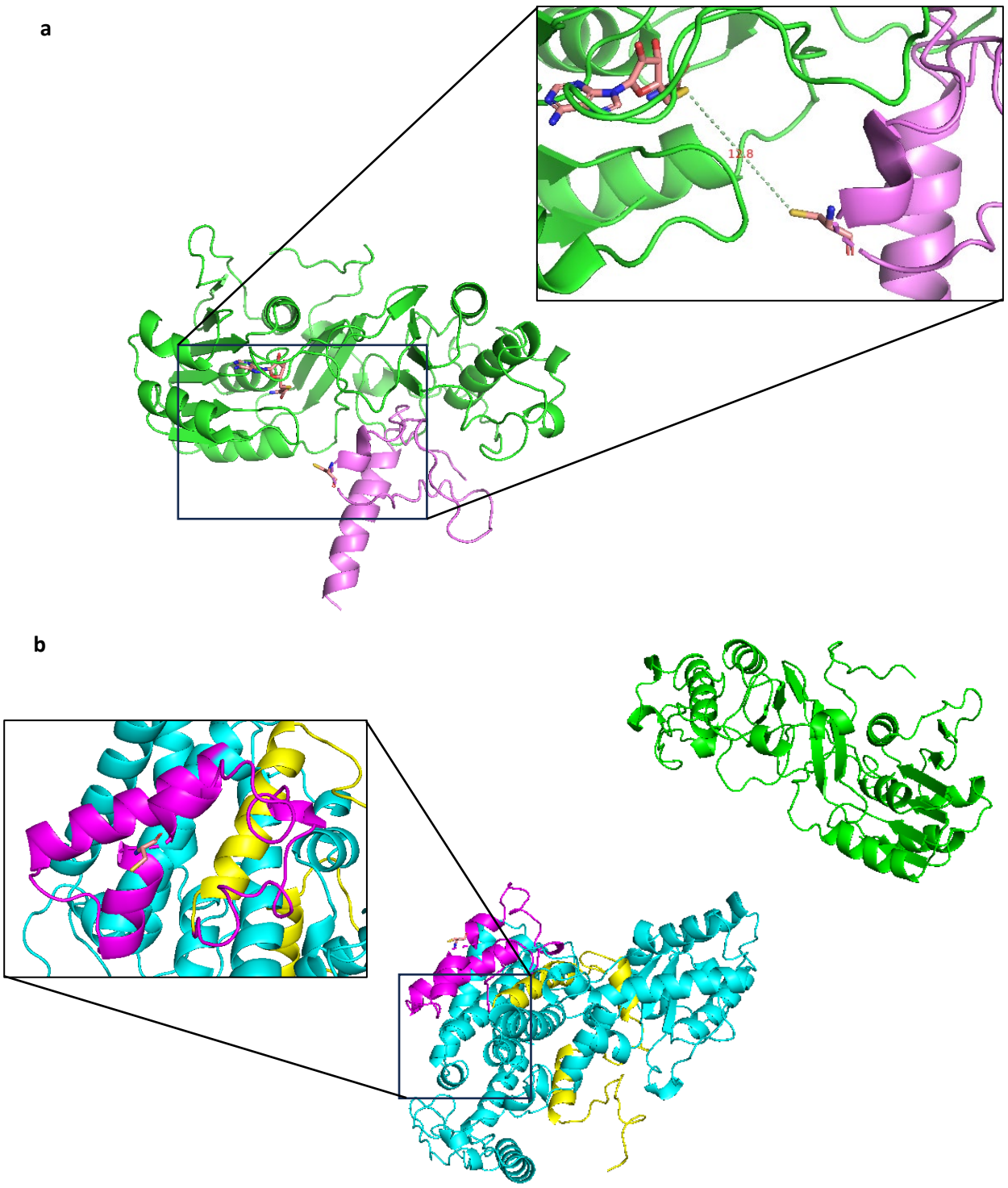


Figure 34: AlphaFold structure of McmA showing Cys472 of a 61 amino acid length McrA peptide within the active site of McmA. Inset shows a 12.8 Å distance of the sulfur of Cys472 to sulfur of active site SAH. (b) shows McrA not interacting with McmA (green); residues around Cys472 are tightly coiled. Cys472 and SAH shown in sticks.

3.3.8 McrA expression with and without McmA

McrA was expressed from the cell lines *E. coli* BL21(DE3) harboring pETSUMO:*mcrA* and *E. coli* BL21(DE3) harboring pETSUMO:*mcrA* and pACYCduet-1:*prmA*. Terrific broth was supplemented with 200 µM iron II sulphate, 200 µM cysteine, 2 µM cyanocobalamin, 50 µM riboflavin and 50 µM methionine at the start of the expression. McrA (74 kDa) was detected in the insoluble fraction of the cell line harboring pETSUMO:*mcrA* (Figure 35).

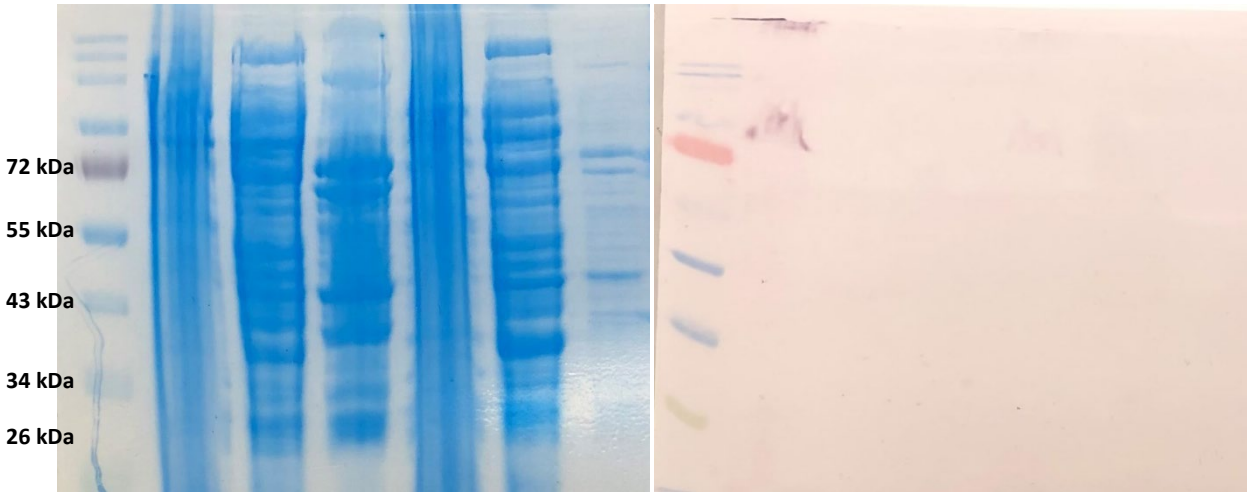


Figure 35: SDS-PAGE gel and western blot result of McrA expressed with and without McmA

Lanes 1 - 4 represent the molecular weight standards, insoluble, flowthrough and elution fraction of the cell line *E. coli* BL21(DE3) harboring pETSUMO:*mcrA*

Lanes 5-7 represent the insoluble, flowthrough and elution fraction of the cell line *E. coli* BL21(DE3) harboring pETSUMO:*mcrA*, and pACYCDuet-1:*prmA*

3.3.9 McrA expression from *E. coli* BL21(DE3) cell lines in terrific broth containing glycerol

McrA expression was carried out anaerobically in terrific broth containing 10 % glycerol, with the addition of 0.2 mM zinc sulfate to the broth after induction. This was based on the data derived from the structural analysis of McmA protein and its interaction with zinc.

The result showed that for both cell lines, McrA (74 kDa) was detected in the insoluble fraction of the cell (Figure 36).

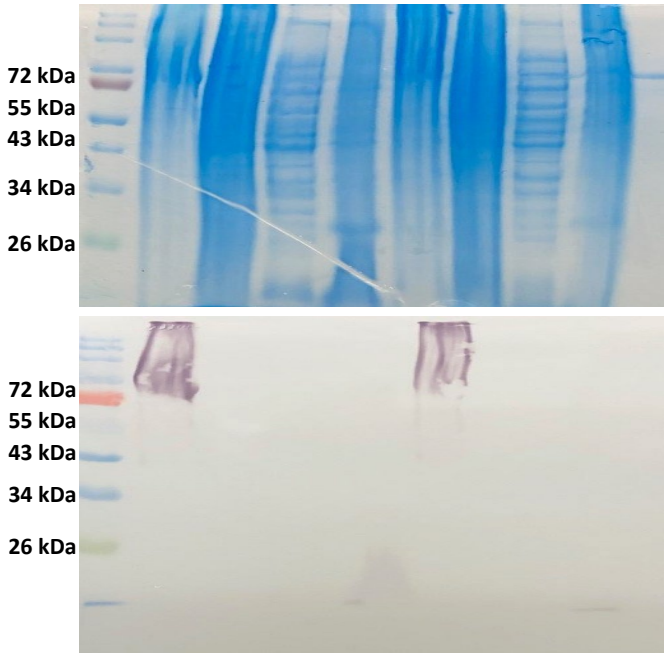


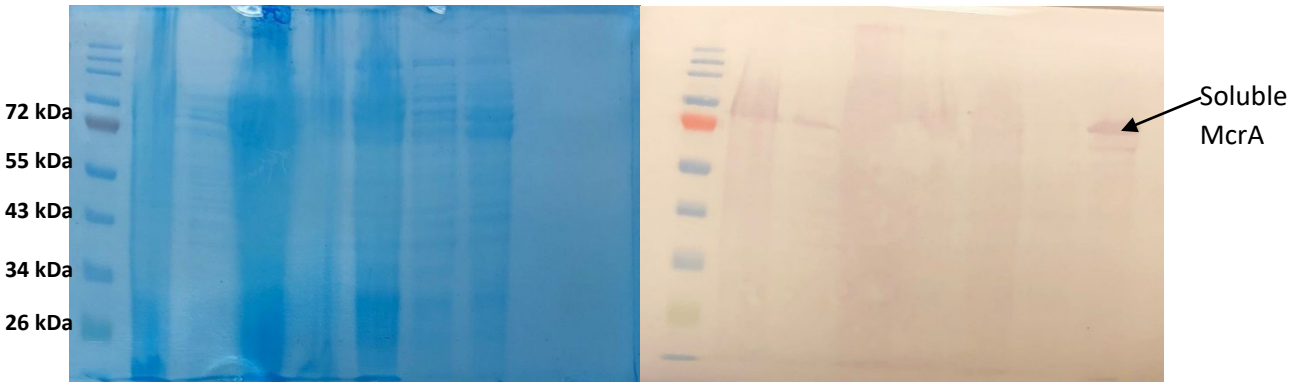
Figure 36: SDS-PAGE gel and western blot result of McrA expressed with terrific broth containing 10 % glycerol.

Lanes 1 - 5 represent the molecular weight standards, insoluble, flowthrough, wash and elution fraction of the cell line *E. coli* BL21(DE3) harboring pETSUMO:*mcrA*

Lanes 6-9 represent the insoluble, flowthrough, wash and elution fraction of the cell line *E. coli* BL21(DE3) harboring pETSUMO:*mcrA*, and pACYCDuet-1:*prmA*

3.3.10 Expression of soluble McrA from *E. coli* BL21(DE3) cell line harboring pETSUMO:*mcrA*

The successful expression of McrA from the *E. coli* BL21(DE3) pETSUMO:*mcrA* cell line in the soluble fraction of the cell (Figure 37) was carried out using broth made up of 10 g of tryptone, 5 g of sodium chloride and 5 g of yeast extract in 10 % glycerol (100 mL) and to a total volume of 1 L. Cells were supplemented with 200 µM iron II sulphate, 200 µM cysteine, 2 µM cyanocobalamin, 50 µM riboflavin and 50 µM methionine at the point of induction. Cells were subjected to argon pressurization to remove all traces of oxygen, prior to harvesting and were purified anaerobically.



3.3.11 Expression of soluble McrA with and without McmA from *E. coli* BL21(DE3) cell lines

Figure 37: SDS-PAGE gel and western blot results of soluble McrA expressed from E. coli BL21(DE3) cell line harboring pETSUMO:mcrA

Lanes 1 - 8 represent the molecular weight standards, whole cell, sonicated cell, cell free extract, insoluble, flowthrough, wash and elution fraction of the cell line *E. coli* BL21(DE3) harboring pETSUMO:*mcrA*

The cell lines *E. coli* BL21(DE3) harboring pETSUMO:*mcrA* and *E. coli* BL21(DE3) harboring pETSUMO:*mcrA* and pACYCDuet-1:*prmA* were expressed and purified under the new set of conditions that generated soluble McrA. Again, McrA (74 kDa) was detected by SDS-PAGE and western blot in the soluble fraction of the cell. It was also present in the insoluble fraction (Figure 38).

Gel bands were excised and sent for mass spectrometry analysis. McmA was detected by mass spec (Table 8) in the insoluble and elution fraction of the *E. coli* BL21(DE3) harboring pETSUMO:*mcrA*; pACYCDuet-1:*prmA* (Figure 39). McmA detection in the elution fraction even though it was not tagged suggests that it was properly folded and bound to McrA, hence its co-elution with it. McrA was not detected in the insoluble or elution fraction by MS, so further analysis was not possible.

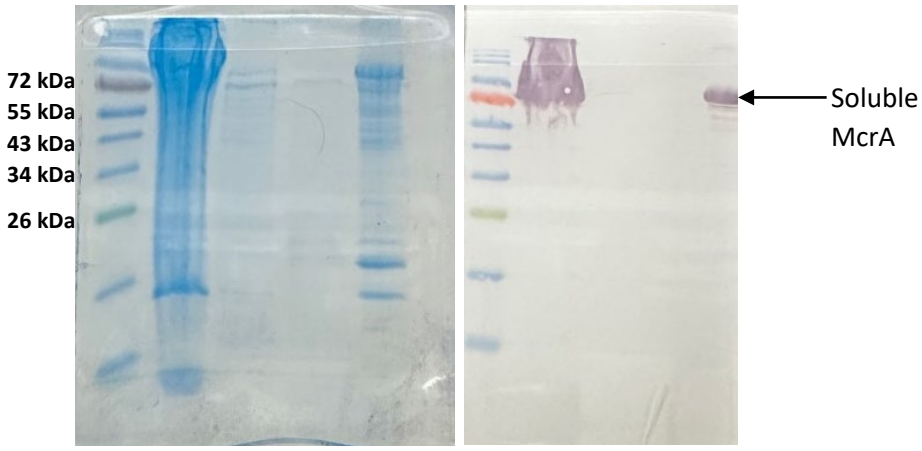


Figure 38: SDS-PAGE gel and western blot result of soluble McrA expressed from the E. coli BL21(DE3) cell line harboring pETSUMO:mcrA

Lanes 1 - 5 represent the molecular weight standards, insoluble, flowthrough, wash and elution fraction of the cell line *E. coli* BL21(DE3) harboring pETSUMO:*mcrA*.

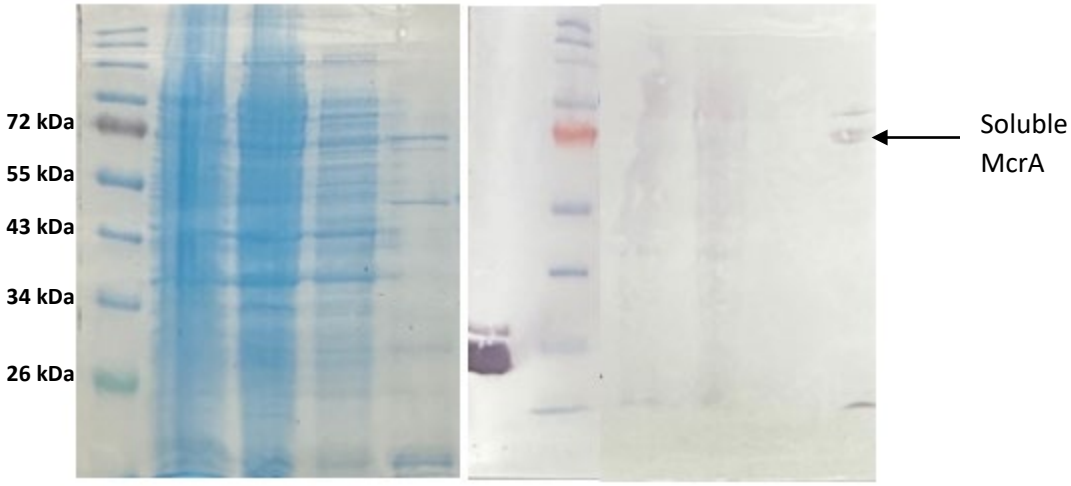


Figure 39: SDS-PAGE gel and western blot result of soluble McrA expressed from the cell line E. coli BL21(DE3) harboring pETSUMO:mcrA and pACYCDuet-1:prmA

Lanes 1 - 5 represent the molecular weight standards, insoluble, flowthrough, wash and elution fraction of the cell line.

Table 8: Mass spec result showing the McmA protein in the elution fraction of the cell line E. coli BL21(DE3) harboring pETSUMO:mcrA and pACYCDuet-1:prmA

	Description	Coverage	No. of peptides	Sum PEP score	Score Mascot: A2 Mascot
Elution fraction	Uncharacterized protein OS=Methanosarcina acetivorans (strain ATCC 35395 / DSM 2834 / JCM 12185 / C2A) OX=188937 GN=MA_4545 PE=4 SV=1	17	5	13.819	292
Insoluble fraction	Uncharacterized protein OS=Methanosarcina acetivorans (strain ATCC 35395 / DSM 2834 / JCM 12185 / C2A) OX=188937 GN=MA_4545 PE=4 SV=1	17	5	9.278	153

3.3.12 Mass spec result of synthesized Mm4 peptide

Synthesized Mm4 peptide with sequence SFAAK**H**AALVS was lyophilized as detailed in section 3.2.6. The lyophilized peptide (1 mg) was dissolved in 1 mL of deionized water and subjected to mass spec analysis. The peptide which has a molecular weight of 1100.62 Da was detected with its corresponding *m/z* value (See Figure 40).

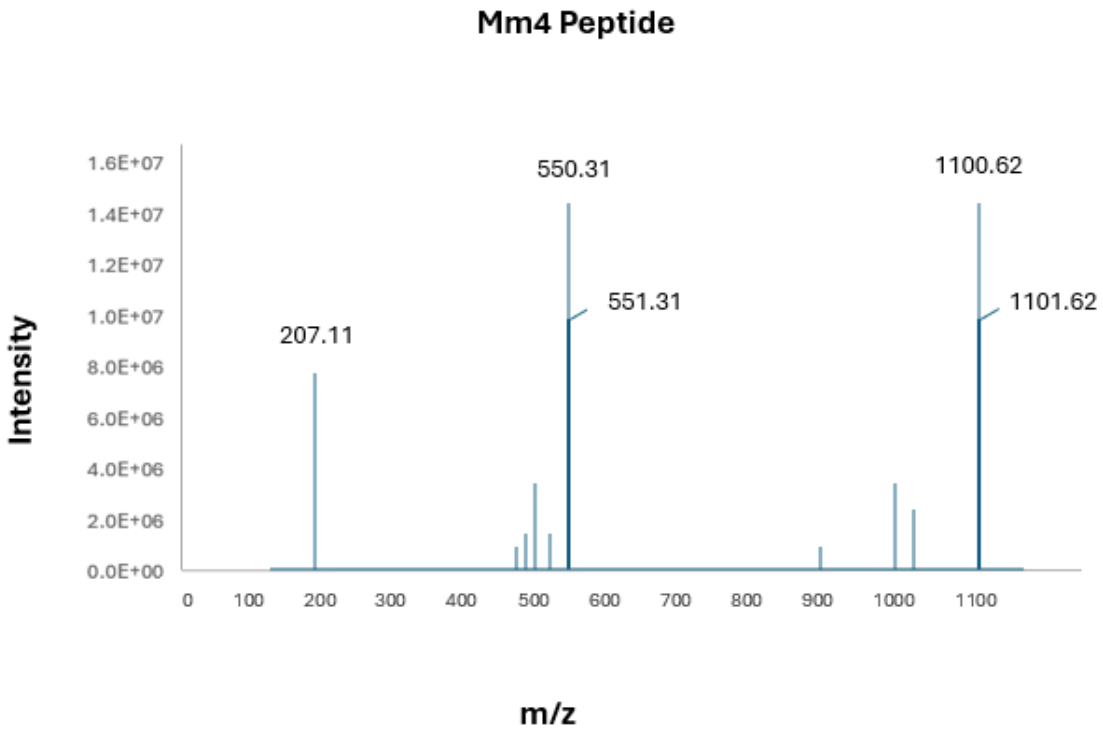


Figure 40: Mass spec data of Mm4 peptide (1100.62 m/z value)

3.3.13 Mm4 activity assay

Purified Mm4 protein was incubated with SAM and the Mm4 peptide for 18 h and the sample was subjected to methanol extraction with subsequent decanting of the supernatant for HPLC and mass spec analysis. The methyltransferase activity of Mm4 can be confirmed by the detection of the methylated peptide (mass of 1114.62) using mass spec, and by the conversion of S-adenosylmethionine (SAM) to S-adenosylhomocysteine (SAH) which can be detected using HPLC (Figure 42).

Similar to what was observed with McmA protein, mass spectrometry result did not show the presence of the methylated peptide (Figure 41), and HPLC did not detect SAH (Figure 43). Upon spiking the reaction group with SAH, a new peak for SAH was detected by HPLC (Figure 43 and Figure 44).

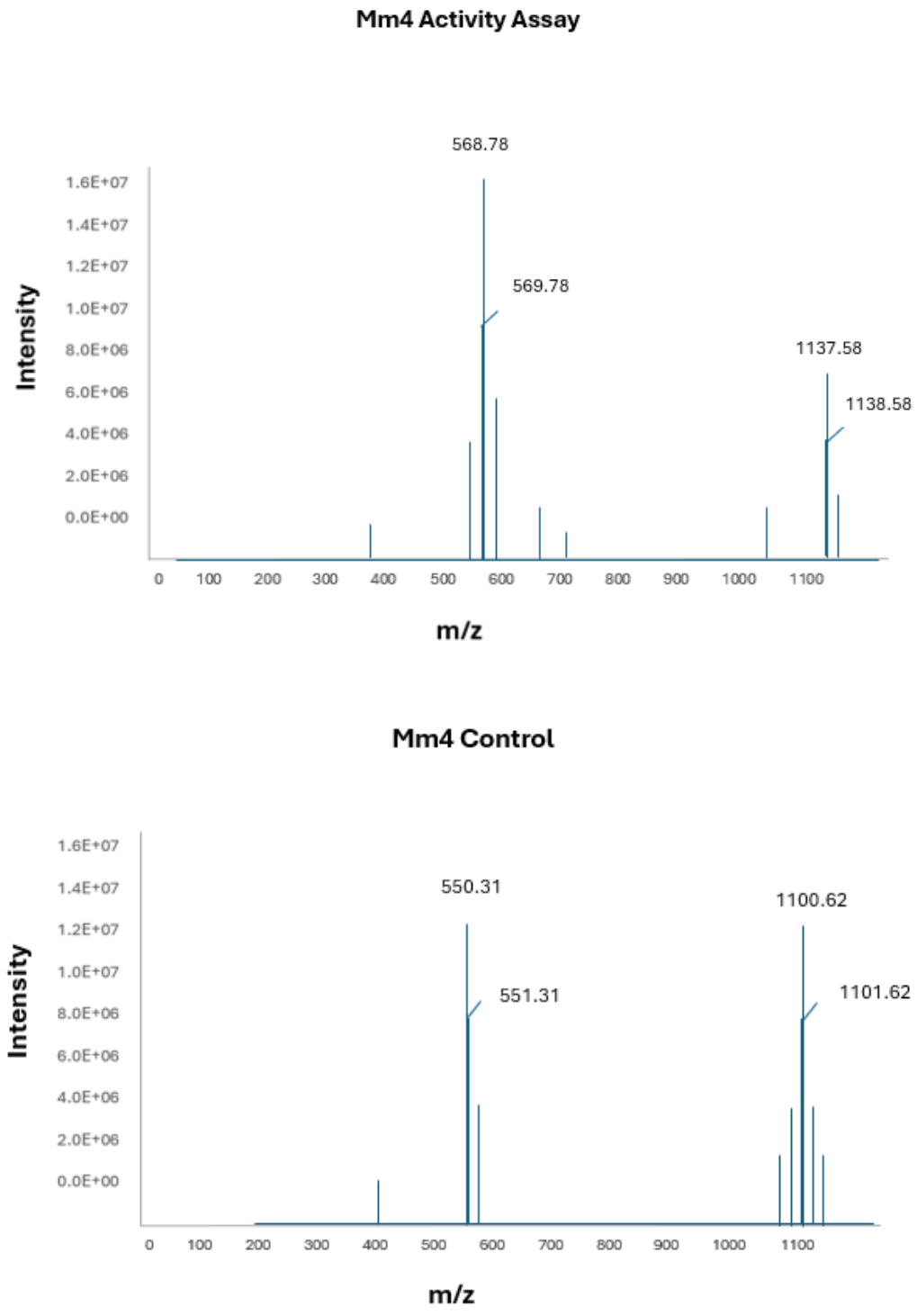


Figure 41: Mass spec data showing the reaction and control groups of the Mm4 activity assay. Reaction contained 100 µM of the Mm4 peptide, 100 µM of SAM and 40 µM of the Mm4 protein in 50 mM HEPES buffer pH 8.0 for 18 h. Mm4 protein was not added to the control group.

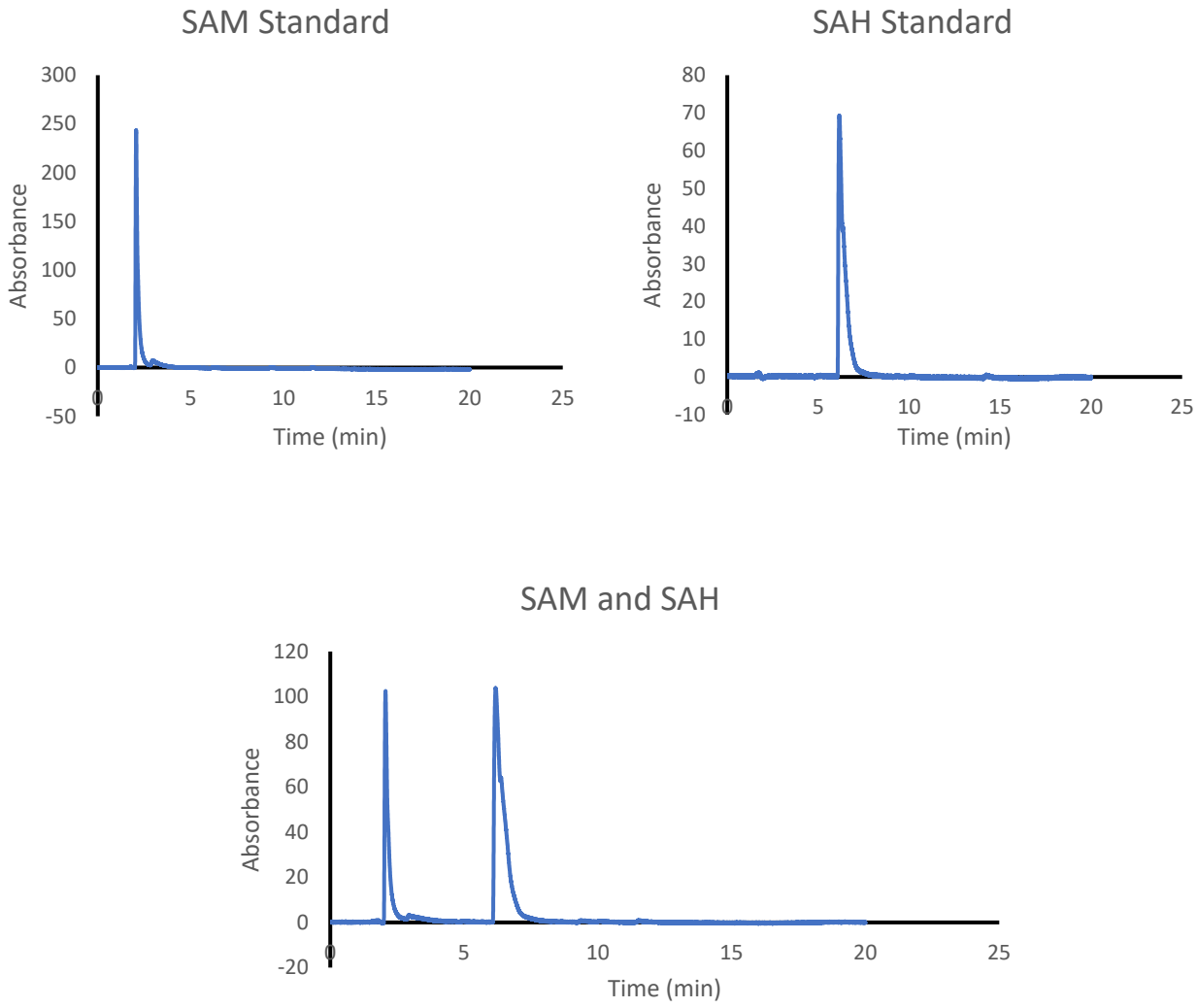


Figure 42: HPLC analysis with SAM and SAH standards showing their separation profiles

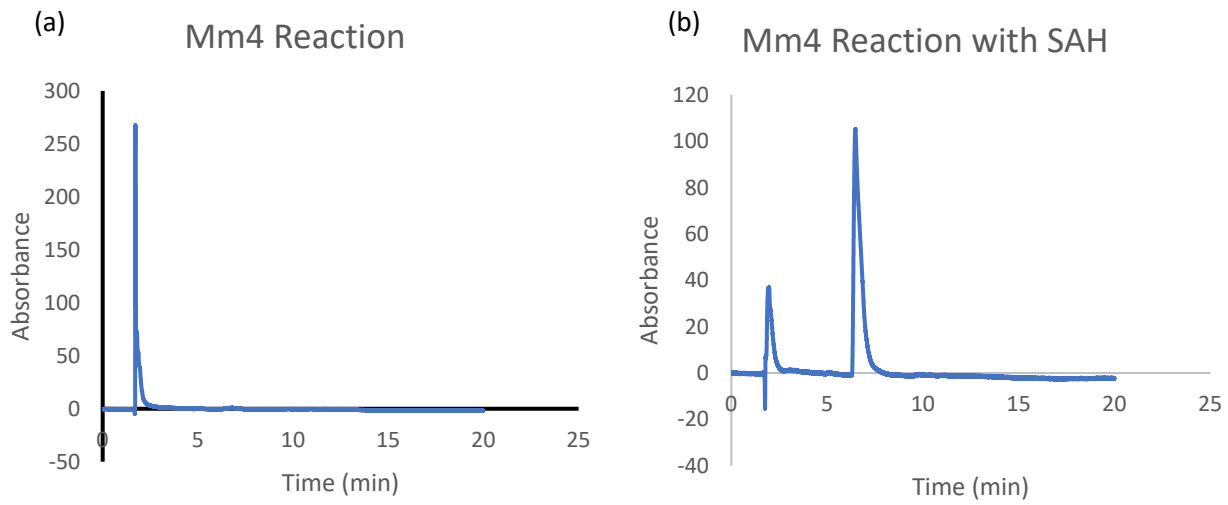


Figure 43: HPLC analysis of Mm4 reaction (a) and of Mm4 reaction spiked with SAH (b). Reaction contained 100 µM of the Mm4 peptide, 100 µM of SAM and 40 µM of the Mm4 protein in 50 mM HEPES buffer pH 8.0 for 18 h. Reaction was spiked with 100 µM SAH in (b)

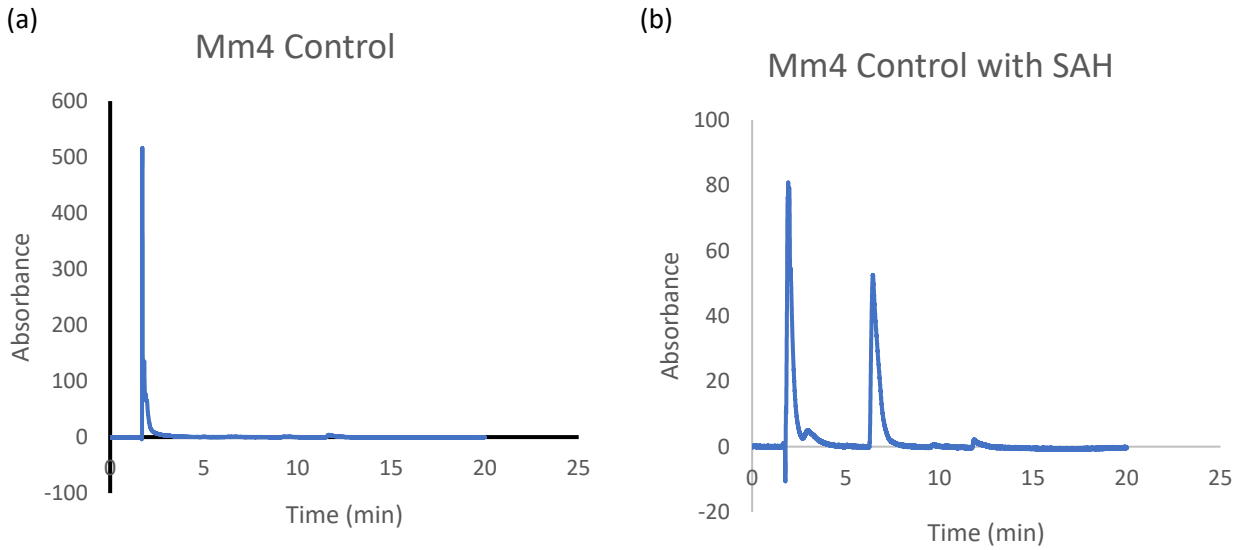


Figure 44: HPLC analysis of Mm4 control (a) and Mm4 control spiked with SAH (b). Reaction contained 100 µM of the Mm4 peptide, and 100 µM of SAM in 50 mM HEPES buffer pH 8.0 for 18 h. Control group was spiked with 100 µM SAH in (b)

3.3.14 Mm4 structural prediction studies

Neither the methylated peptide nor SAH were detected using the assay. This raised the question of the possibility of the peptide not binding to the enzyme due to its small size. Structural prediction of the Mm4:peptide complex was carried out using the AlphaFold software[95,96], and visualized using PyMOL[97]. It showed that the Mm4 peptide was not binding inside or near the active site of the Mm4 protein. The exact position of the Mm4 active site (Figure 46) was predicted by comparing the AlphaFold derived structure of Mm4 to that of structural homologs identified through Dali search[101]. Phosphate butaryl transferase was identified as one of the structural homologs (Table 9). A crystal structure of phosphate butaryl transferase (3u9e-B) co-crystallized with coenzyme A (CoA) was aligned with Mm4 using PyMOL (Figure 45) and the active site of Mm4 was predicted.

Table 9: Dali search result showing proteins with structural similarity to Mm4

No	Chain	Z-score	Description
1	7vg9-C	24.5	Phosphate butyryl transferase
2	3u9e-B	23.5	Phosphate butaryl transferase
3	1vil-A	20.8	Fatty acid/phospholipid synthesis protein PLSX
4	6e85-A	17.7	D-Threonate 4-phosphate dehydrogenase
5	6znj-A	16.4	NADP-dependent malate dehydrogenase
6	3tsn-A	15.5	4-hydroxythreonine-4-phosphate dehydrogenase
7	2d1c-A	10.9	Isocitrate dehydrogenase
8	61kz-C	10.7	Isocitrate dehydrogenase
9	6tdw-B	9.4	ATPTB1

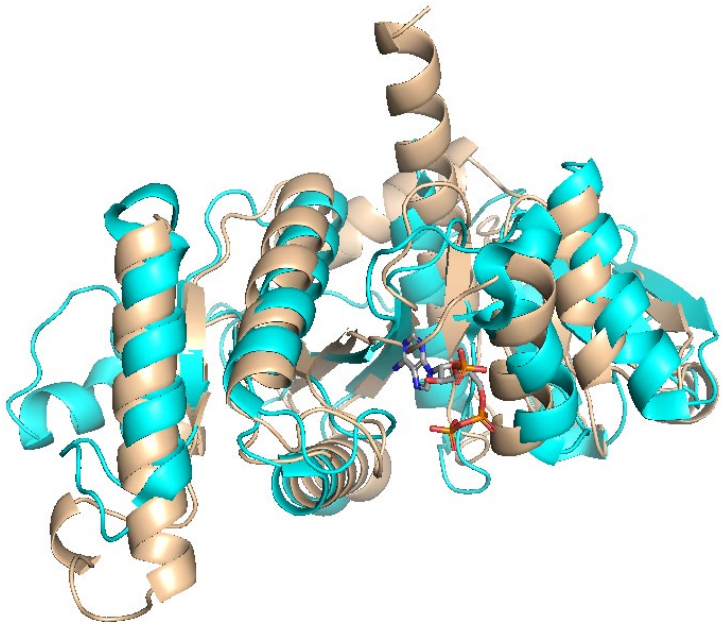


Figure 45: Structural alignment of Mm4 (wheat) to phosphate butaryl transferase {3u9e-B} (cyan)

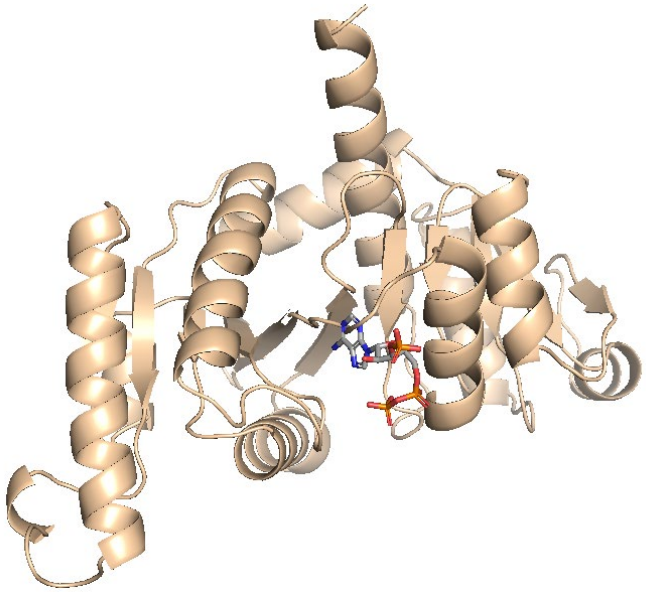


Figure 46: Mm4 protein structure showing the predicted enzyme active site. CoA shown in sticks.

Subsequent extension of the Mm4 peptide by addition of amino acids to the C- and N-terminals (15 amino acid length peptides) did not result in the peptide binding to the active site of the Mm4 protein (Figure 47). Upon extension to 51 amino acid length, the peptide appeared to be extending into the active site of Mm4 though the his271 residue remained outside the active site. With the 101 amino acid length peptide, its extension into Mm4 active site was not observed and with the full length McrA, Mm4 protein was too far away to be interacting with its substrate in any way (Figure 48).

This observation was unexpected but not surprising considering the results of the Dali search (Table 9). It can be inferred that Mm4 was not accurately labeled as a methyltransferase. The protein structure is more closely related to that of transferases and dehydrogenases. It should be noted that the Mm4 structure used in this study is a predicted structure, and conclusions on whether or not it is correctly annotated as a methyltransferase will be more apt when this analysis is conducted with the enzyme crystal structure.

However, the data supports the conclusion that neither McrA nor its peptide fragments is a substrate for Mm4. Therefore, the *mm4* gene is not responsible for the methylhistidine PTM in MCR. Other gene candidates that can be explored include *MA0673* which is also annotated as a methyltransferase. *MA0673* homolog is clustered with *mcrC* and *A2* genes in ANME-1, suggesting a role in MCR maturation. The MA0673 protein also exhibited more promising structural interactions with the McrA protein.

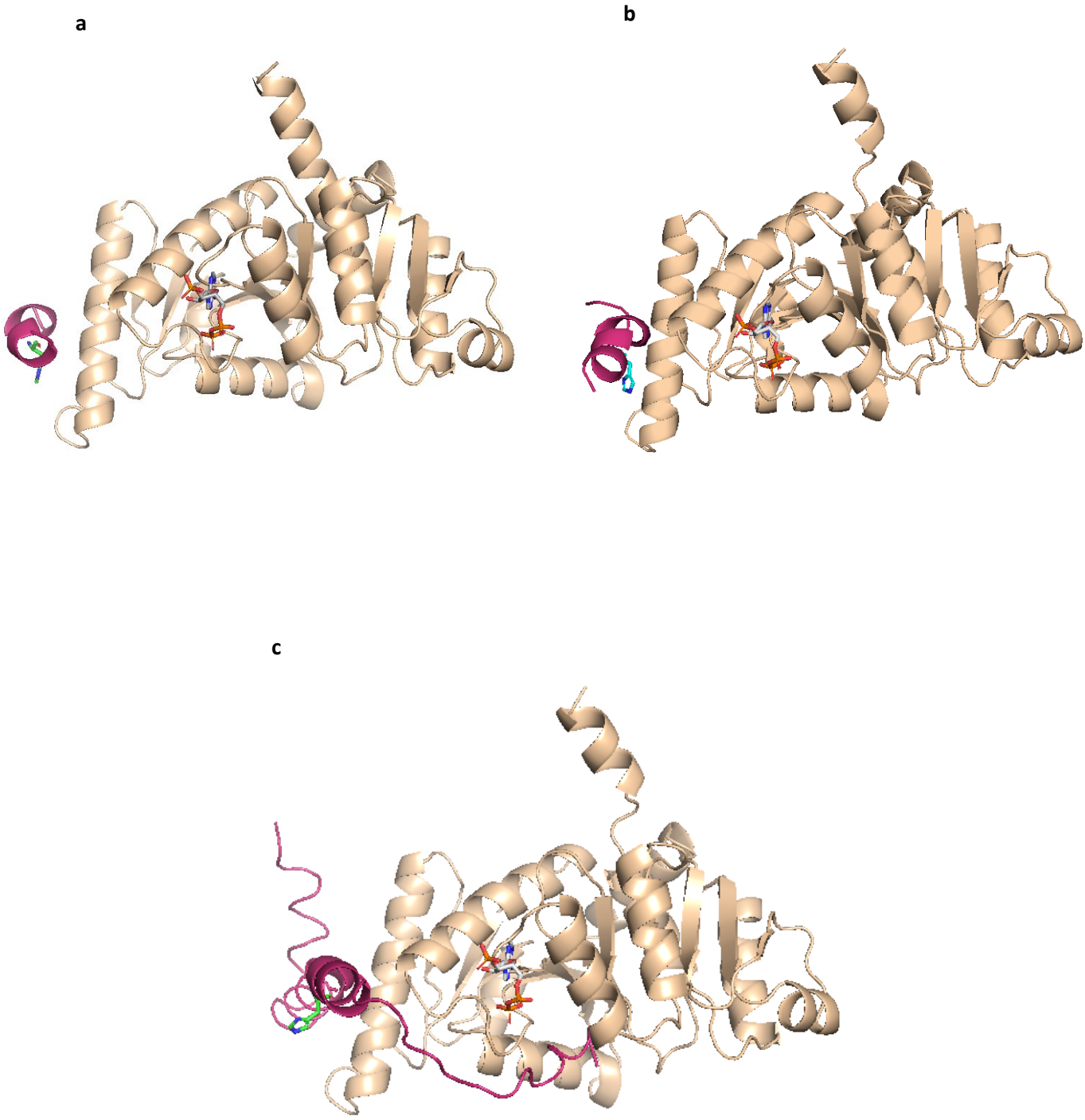


Figure 47: AlphaFold structure of Mm4 protein with (a) 11 amino acid length peptide, (b) 15 amino acid length peptide and (c) 51 amino acid length peptide. His271 shown as sticks

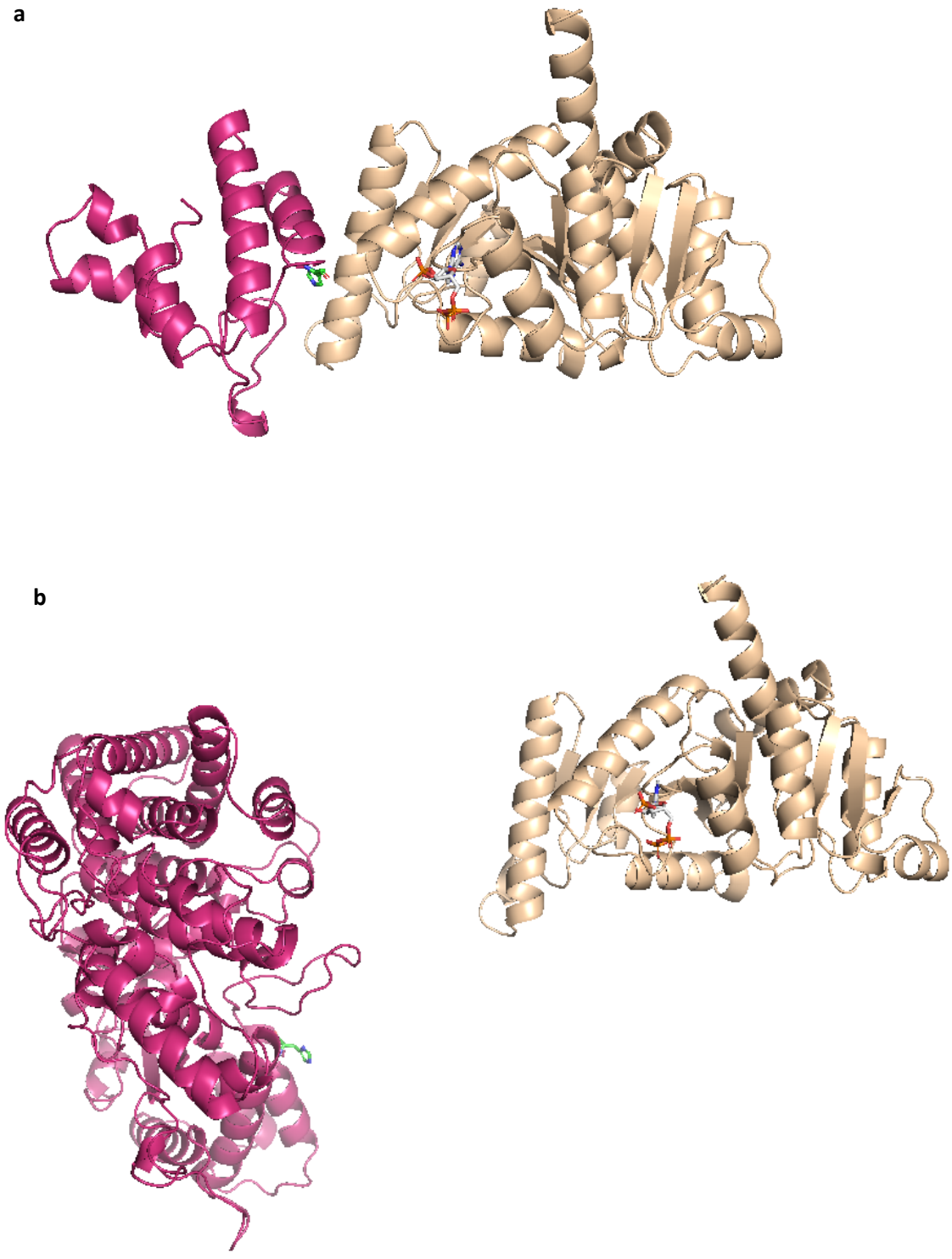


Figure 48: AlphaFold structure of Mm4 protein with (a) 101 amino acid length peptide, (b) full length McrA (rose pink). His271 shown in sticks

3.4 Discussion

Both Mm4 and McmA (PrmA) proteins were successfully expressed and purified aerobically. Both proteins were detected in the elution fraction after purification as shown on the SDS-gel and the western blot results (Figure 21 and Figure 20).

However, results for the initial activity assays for McmA did not show the methylation of its peptide substrate. The peptides are tiny analogs of the actual substrate (McrA), so there was always the possibility that they would not be accepted as a substrate by the enzyme. The HPLC analysis detected only SAM in both the reaction and control groups and the mass spec analysis showed only the peptide mass in both the reaction and the control groups.

Analysis of the AlphaFold generated structure of the McmA (PrmA) protein revealed that it contained 16 cysteine residues, some of which appeared to be oriented towards an imaginary metal binding center (Figure 23). Both zinc and iron were inserted into the metal binding center and zinc appeared to be a better fit pictorially. McmA was also better expressed from the cell line *E. coli* BL21(DE3) pETSUMO:*prmA* when the cell line was supplemented with zinc rather than with iron.

Subsequent McmA expression was carried out with zinc supplementation and McmA purification was carried out anaerobically to account for the cysteine residues which may be sensitive to oxygen. McmA activity assays were also carried out anaerobically, with the addition of 1% DTT to keep the reaction under reducing conditions and with the addition of zinc sulphate to the reaction group. However, despite these adjustments, Mass spec and HPLC analysis of the McmA activity assays did not detect the presence of the methylated peptide (Figure 28) or the formation of SAH (Figure 26) respectively.

The observed results led to the exploration of the possibility of product inhibition phenomenon. It is not uncommon for methyltransferases to exhibit strong product inhibition by SAH[100], preventing product formation in quantities substantial enough to be detected by mass spec or HPLC. To address this, the addition of the enzyme adenosylhomocysteinase to the McmA reaction was explored. Adenosylhomocysteinase catalyzes the reversible hydration of S-adenosyl-L-homocysteine into adenosine and homocysteine. With SAH conversion into adenosine, the SAH-inhibited McmA will be available for enzymatic methylation of the peptide.

Mass spectrometry results of the McmA methylation assay with addition of adenosylhomocysteinase did not detect the methylated peptide (See Figure 30). These results suggest that the McmA enzyme may require the intact McrA protein for catalysis to occur. Structural prediction studies of the McmA enzyme with the peptide showed that the peptide was within the active site of McmA but at a distance of 17.8 Å from SAH (Figure 33). This, along with the research by Nayak, et al.[62], showing that mutant strains of *M. acetivorans* lacking the *mcmA* gene produced MCR without the methylcysteine modification, informed the decision to continue investigation into the methylation activity of the enzyme.

The possibility that the methylation of McrA by McmA may be a co-translational modification rather than a post-translational one was explored next. In co-translational modification, the modification of the protein occurs while the protein is being translated on the ribosome. An example of a co-translational system is the N-terminal acetyltransferases (NATs) system[102]. A member of the NATs family, NatA, is composed of a catalytic NAA10 subunit and an auxiliary NAA15 subunit[103]. The NAA15 subunit acts as an anchor connecting the catalytic NAA10 subunit to the ribosome during translation for co-acetylation of the nascent peptide[104]. It is possible that McmA may function in a similar manner and may act in tandem with an anchor protein.

An *in-vivo* approach was therefore employed to study the methylation of cysteine by McmA. Two *E. coli* BL21(DE3) cell lines were created. One cell line contained the pETSUMO:*mcrA* plasmid and was expected to express McrA without the methylcysteine modification. The other cell line contained pETSUMO:*mcrA* and pACYCDuet-1:*prmA* plasmids and was expected to express McrA with the methylcysteine modification. Initial McrA expression from both cell lines produced McrA in the insoluble fraction of the cell (Figure 35). Experimentation with broth composition was carried out to determine optimum conditions for soluble McrA expression.

Upon reduction of the salt content of the broth media used for expression, supplementation with some nutrients, and using anaerobic purifications protocols, a band corresponding to the size of McrA was found on the SDS-gel and western blot membrane of the purification fractions. The McrA band was detected in the elution fraction of the cell (see Figure 37).

The method for soluble McrA expression and purification was repeated in purifying the two cell lines *E. coli* BL21(DE3) harboring pETSUMO:*mcrA* and *E. coli* BL21(DE3) harboring pETSUMO:*mcrA* and pACYCDuet-1:*prmA*, under the same set of conditions and at the same time. McrA was found in the soluble fraction in both cell lines (Figure 38 and Figure 39).

This sets the stage for future work, expressing large volumes of soluble McrA for use in in vitro methylation experiments. However, McrA degrades very quickly so care must be taken to detect the methylated product through MS as fast as possible. McmA protein was detected through MS in the soluble fraction of the E. coli BL21(DE3) cell line harboring pETSUMO:mcrA and pACYCDuet-1:prmA (Table 8) while McrA was not detected. Considering that only the McrA protein was tagged and should be present in the elution fractions, the presence of McmA in the fractions indicates that McrA co-eluted with McmA, but while McmA could be detected, McrA could not be detected most likely due to the instability of McrA.

It is possible that soluble McrA is not easily ionizable, or it is easily degraded in the absence of its co-subunits. It is also possible that McrA requires binding to some other elements for stability. Every MCR crystal structure obtained to date has composed of the complete hexamer, usually bound with F430 and/or some of the substrates. This raises the question of the suitability of the *E .coli* cell as a viable host for the expression of archaeal proteins. While both organisms are single-celled prokaryotes, variations exist in the chemical composition of their respective DNAs, which might influence how archaeal proteins expressed in *E. coli* cells might behave post expression.

Expression of stable, soluble McrA for methylation activity assays such as this one may require the use of the native archaea (*M.acetivorans*), after deleting the post-translational modification genes of interest, as well as the *mcrB* and *mcrG* genes from *M. acetivorans*. Alternatively, only the post-translational modification genes can be deleted, and the intact MCR can be expressed and used for the activity assays.

A similar challenge was encountered when testing for the methylation activity of Mm4 on its peptide. The mass spec result for Mm4 activity assay showed the intact peptide (*m/z* = 1100.62) in the control group lacking the Mm4 enzyme (Figure 41). In the reaction group, the unmethylated peptide was detected as a potassium adduct $[M+H+K]^+$ and $[M+H+K]^{2+}$. HPLC analysis of the Mm4 activity assay corroborated the mass spec result, with only SAM detected by HPLC in both the reaction and control groups (Figure 44 and Figure 43). Subsequent spiking of both the reaction and control groups with SAH led to the formation of a new peak which corresponded to the SAH standard, confirming that SAH was not formed in the reaction.

Structural prediction studies of the Mm4 protein and the peptide provided some explanation for the mass spec and HPLC results observed. The generated structures showed that the peptide did not bind within the active site region of the mm4 protein (Figure 47). Peptides with amino acid

lengths 11, 15, 51, 71, and 101 were used for the analysis to ascertain their interaction with the Mm4 protein *in silico*. Extension of the peptide length did not improve active site binding of the peptide to the Mm4 protein. Enzyme catalysis is only possible when the substrate interacts with the active site of the enzyme, leading to product formation. Because the peptide is only a fragment of the McrA substrate, non-interaction between the peptide and the Mm4 enzyme could be because of the short peptide length which may make it impossible for the mm4 protein to initiate other interactions with the missing residues of the peptide that will facilitate active site binding. In this case, only the presence of a full-length McrA protein as a substrate can negate the potential of non-interaction between the Mm4 enzyme and its substrate.

However, subsequent analysis revealed that the full length McrA did not bind to Mm4 and the his271 residue on full length McrA was not situated anywhere close to Mm4 active site (Figure 48). Structural comparison of Mm4 homologs using Dali search (Table 9) showed that the protein was structurally similar to transferases and dehydrogenases, rather than methyltransferases. This strongly suggests not only that *mm4* is the wrong gene candidate for the methylation of the his 271, but also that Mm4 was wrongly annotated as a methyltransferase. In general, these results highlight the knowledge gap yet to be covered in the quest to fully understand and elucidate the mechanism by which active MCR is assembled and activated.

The *M. acetivoran* genome is replete with several other genes annotated as methyltransferases, any of which could be responsible for the methylation of histidine. Another gene currently being investigated by the Mansoorabadi lab for histidine methylation is the *MA0673* gene.

Dali search of *MA0673* shows its structural similarity to methyltransferases (Table 10). Structural prediction analysis showed that the his271 residue on McrA was in a position close enough to interact with the SAH binding region of MA0673 protein (Figure 49), supporting the hypothesis

that the gene is a promising target for his271 methylation. The SAM binding site in MA0673 was deduced by aligning MA0673 with a structural homolog (PDB ID: 1wzn-A).

Table 10: Dali search result showing proteins with structural similarity to MA0673

No	Chain	Z-score	Description
1	3sm3-A	16.6	SAM-dependent methyltransferase
2	1ve3-B	16.6	Hypothetical protein PHO226
3	1wzn-A	16.4	SAM-dependent methyltransferase
4	3mgg-B	16.1	Methyltransferase
5	3tm5-A	16.0	Cystal structure of TRM14
6	5m58-A	15.9	C-methyltransferase COUO
7	2o57-A	15.9	Putative sarcosine dimethylglycine methyltransferase
8	3dli-A	15.8	Methyltransferase
9	5f8c-B	15.8	Methyltransferase
10	7v6h-A	15.7	Cyclopropane fatty-acyl-phospholipid synthase-like

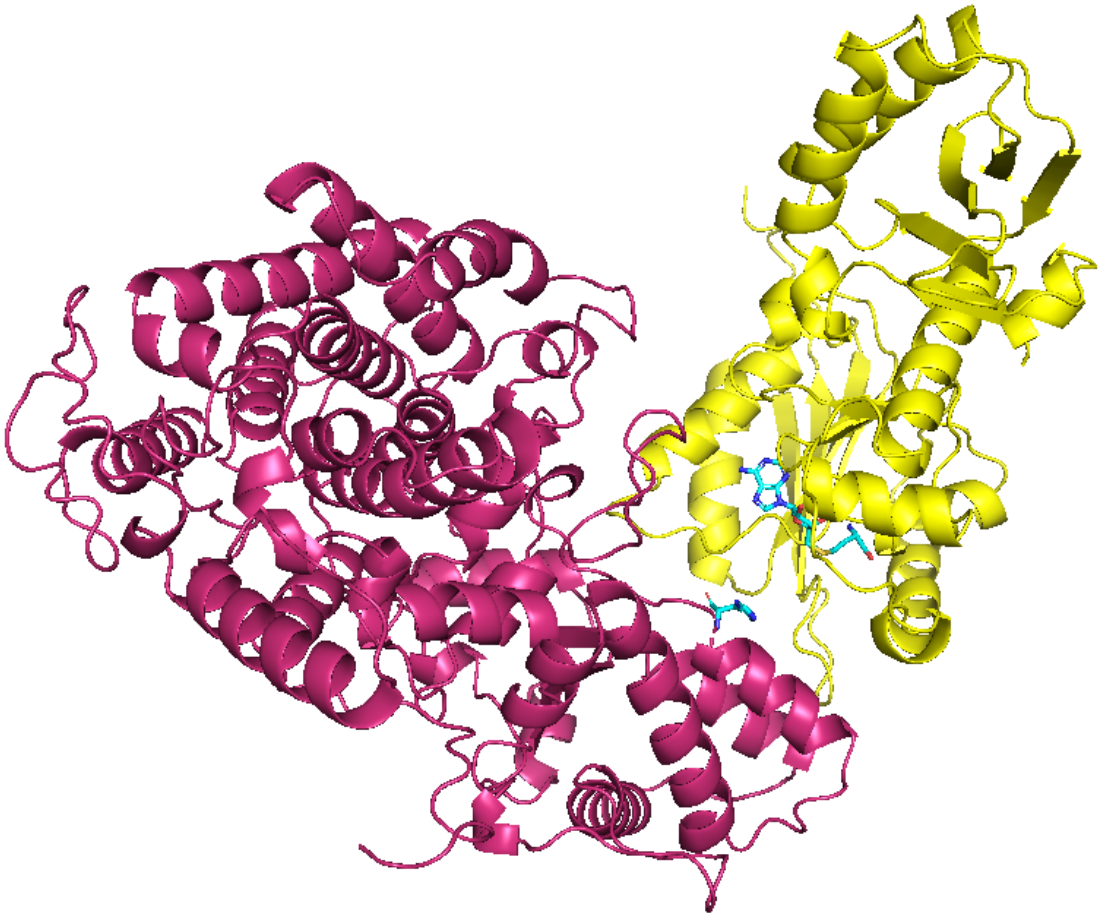


Figure 49: AlphaFold structure of MA0673 (yellow) interacting with McrA (Rose pink), SAH and his271 shown as sticks

Chapter 4

Molecular dynamics studies of the effects of F430 and post translational modifications on methyl-coenzyme M reductase and homologs catalyzing anaerobic alkane oxidation

4 Chapter 4

4.1 Introduction

Tetrapyrroles are one of the most abundant pigments on earth and exist in nature in different forms[105]. The biosynthetic pathway of tetrapyrroles usually starts with either glycine or L-glutamate[106]. Irrespective of the starting material, both pathways lead to the formation of 5-aminolevulinic acid, which proceeds to the formation of uroporphyrinogen III[106] (See Scheme 6) in **Chapter 1**.

4.1.1 Tetrapyrroles in MCR

The enzyme methyl coenzyme M reductase (MCR) catalyzes the production of methane in methanogens, through a process called methanogenesis. MCR is typically isolated as a yellow-hued protein composed of three subunits organized in an α2β2γ2 configuration, with a molecular weight of about 300 kDa[28]. Intrinsic to the activity of MCR is the highly reduced tetrapyrrole - coenzyme F430 (*Figure 50*) which is a unique nickel tetrahydrocorphin[29,30]. Coenzyme F430 is found in the active site of MCR and contains two exocyclic rings: a carbocyclic F ring and a γ-lactam E ring[10].

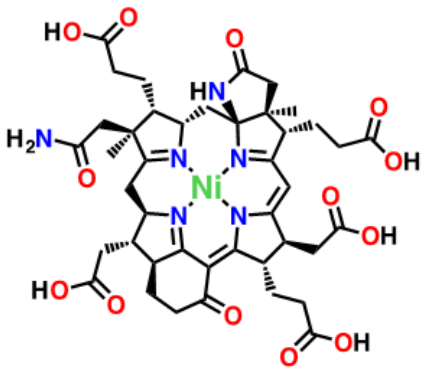


Figure 50: Structure of coenzyme F430

MCR catalyzes the conversion of coenzyme B (CoB-SH) and methyl coenzyme M (MeS-CoM) to methane and the heterodisulphide (CoB-S-S-CoM)[30]. This reaction occurs in methanogens where MCR functions as the key methane forming enzyme. Additionally, MCR plays a pivotal role in initiating the anaerobic oxidation of methane (AOM) in methanotrophic archaea (ANME)[38,107]. Methanotrophs operate via a reverse methanogenesis pathway leading to the oxidation of methane, with MCR catalyzing the first step[38,39,107].

The α subunit of MCR contains several post translational modifications (PTMs)[23,30,52,74]. MCR from *Methanobacterium thermoautotrophicum* features 1-*N*-methylhistidine, 2-(*S*)-methylarginine, and thioglycine PTMs which appear to be conserved in all methanogenic MCR analyzed to date[30–32,108], as well as 2-(*S*)-methylglutamine, *S*-methylcysteine PTMs[30] and didehydroaspartate[23]. MCR from ANME-1 lacks the 2-(*S*)-methylarginine, 2-(*S*)-methylglutamine, *S*-methylcysteine and didehydroaspartate PTMs, but in addition to 1-*N*-methylhistidine and thioglycine, it also features 7-hydroxytrptophan and *S*-oxymethionine PTMs[52]. In contrast, the ethane oxidizing homologue of MCR, ethyl-coenzyme M reductase (ECR) from *Candidatus Ethanoperedens thermophilum* contains all of the PTMs of *M. thermoautotrophicum* MCR except for didehydroaspartate, as well as additional 3-*N*-methylhistidine and 3-methylisoleucine PTMs[74].

All MCRs studied to date show a requirement of coenzyme F430 for activity[30,38,52]. However, variations exist in the F430 found in homologues of MCR. The MCR structure derived from ANME-1[52] possesses a uniquely modified F430, 17^2-methylthio F430, and ECR was found to possess 17,17^2-dimethyl F430[74] (See Figure 51). While the exact functionality of these F430 modifications and PTMs are unclear, several hypotheses have been proffered. Nayak and colleagues have suggested that the PTMs may be needed to finetune catalytic MCR activity[53,62], while Deobald and colleagues opine that the methylglutamine modification may be necessary for

maintaining the ideal shape of the MCR substrate channel for efficient binding[92]. It is possible that for ECR, the methyl groups added to F430 is necessary to make the space where ethane binds more greasy, and thus improve ethane binding. The positioning of the regular F430 may be impaired by the wider ethane binding catalytic chamber present in ECR (Figure 52)[74].

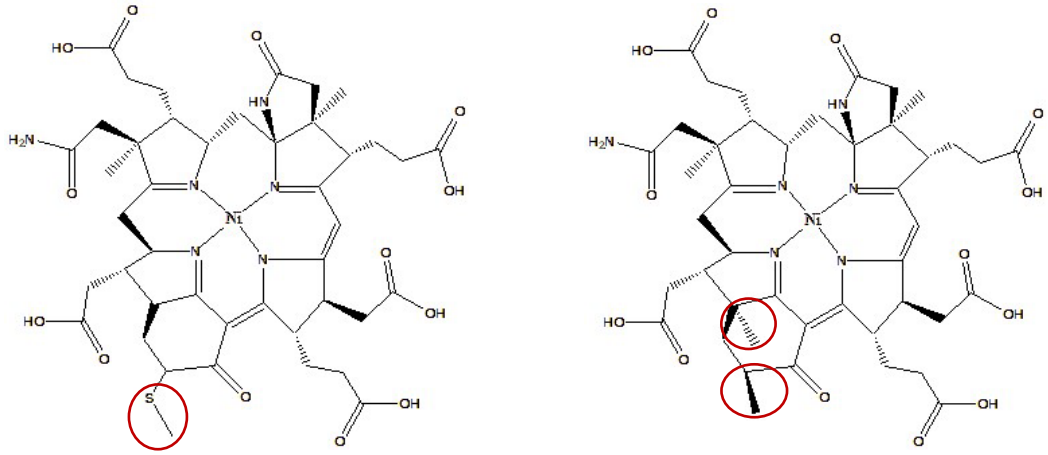


Figure 51: Structure of 17²-methylthio F430 and 17,17²-dimethyl F430

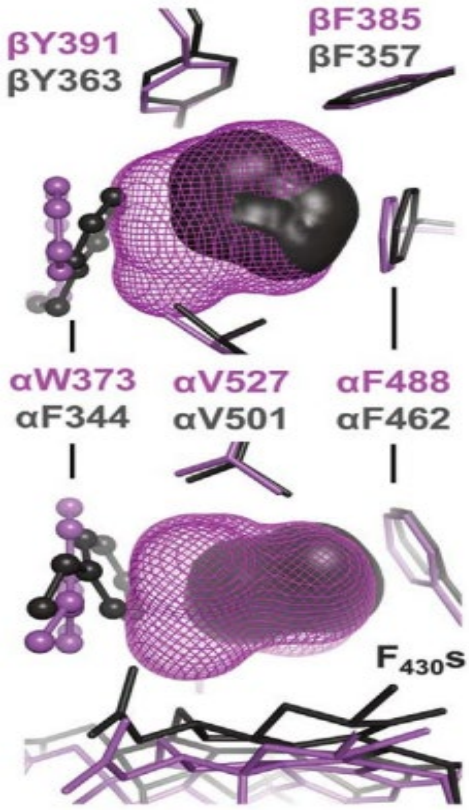


Figure 52: Structure comparing the catalytic cavity volume of ANME-1 MCR (black) and ECR from Ca. E. thermophilum (purple mesh)[74]

While the functions of the coenzyme F430 and post-translational modifications found in MCR are still unclear, it is highly unlikely that nature would invest energy into such rare modifications without an overarching significance. Their presence near the enzyme active site suggests that they play yet undetermined roles in MCR catalytic function. In this chapter, the effect of coenzyme F430 and post-translational modifications on active site features such as active site cooperativity and methane/ethane binding, of three homologues of MCR in complex with their substrates/products were explored. MCR from *Methanothermobacter marburgensis* containing regular coenzyme F430, MCR from ANME-1 containing 17^2-methylthio F430 and ECR from *Ca.*

E. thermophilum containing 17,17^2-dimethyl F430 were investigated using molecular dynamics (MD) simulations. Simulations on MCR from ANME-1 and on ECR were also carried out with regular F430, rather than their native coenzymes. Additionally, all MCR simulations were carried out with and without their native PTMs.

MCR from methanogens such as *M. marburgensis* and *M. acetivorans* catalyzes the conversion of CoB-SH and MeS-CoM to HDS and methane, while the analogous MCR-catalyzed reaction runs in reverse in ANME-1[52]. For the ECR-catalyzed reaction, the substrates are ethane and HDS and the products are CoB-SH and EtS-CoM[74]. The active form of MCR contains an F430 that is in the Ni(I) oxidation state[109]. Since the reaction catalyzed is a reversible one, calculating the distance of a carbon atom in an MCR substrate/product like methane to the Ni(I) atom will provide information on active site behavior of each MCR analyzed and illuminate the influence of the coenzyme F430 analogues on MCR. A clear picture of the influence of F430 homologues on MCR active site can potentially provide insights into the exact role of F430 modifications in MCR.

Furthermore, the crystal structure of MCR from *M. thermoautotrophicum* showed that the Ni atom in coenzyme F430 of MCR coordinated to six ligands in an octahedral configuration: four equatorial nitrogen atoms from the tetrapyrrole ring, a lower axial ligand from the side chain oxygen of α Gln 147, and an upper axial ligand from the thiol group of CoM or the sulfonate oxygen of HDS [23,30,74]. Shima and coworkers reported that ANME-1 MCR was virtually identical to methanogenic MCR in overall structure and in the conformation of the substrates HS-CoM and CoB-SH in the active site[52]. As shown in Figure 53, the ECR from *Ca. E. thermophilum* also shares a similar active site configuration with the notable exception of the lower axial ligand being an α′ Met 181 residue[74].

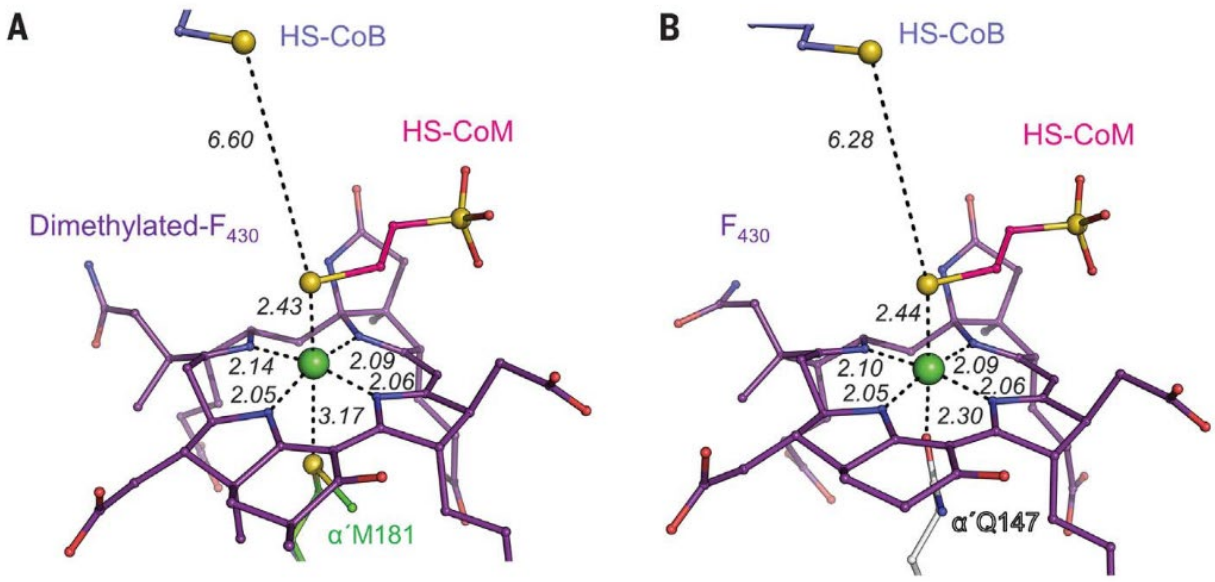


Figure 53: Comparison of coenzyme F430 coordination to its ligands in (A) ECR from Ca. E. thermophilum and (B) MCR from Methanobacterium thermoautotrophicum[74]

4.1.2 Simplified background on the process of molecular dynamics

Molecular dynamics (MD) are computational methods that simulate systems following Newton's second law of motion.

$$\vec{F} = m\vec{a}$$

Where:

$\vec{F}$ = Force applied to an object

m = Mass of the object

$\vec{a}$ = acceleration of the object

MD simulations utilize a generalized model of interatomic interactions to predict the motion of each atom within a molecular system (e.g., an enzyme) over time[110]. These simulations can depict numerous significant biomolecular phenomena, such as conformational transitions, binding of ligands, and the folding of proteins[110,111]. They provide detailed atomic-level insights, including the precise positions of all atoms, with a temporal resolution as fine as femtoseconds. Notably, MD

simulations can also anticipate how biomolecules will react at the atomic level to perturbations like mutations, phosphorylation, protonation, or the introduction or removal of a ligand[112,113].

The MD force field which is a derivative of the potential energy of the system is given as:

$$V(r^N) = \sum_{i\in \text{bonds}} k_{bi}\left(l_i - l_i^0\right)^2 + \sum_{i\in \text{angles}} k_{ai}\left(\theta_i - \theta_i^0\right)^2 + \sum_{i\in \text{torsions}} \sum_{n} \frac{1}{2} V_i^n [1 + \cos\left(n\omega_i - \gamma_i\right)] + \sum_{j=1}^{N-1} \sum_{i=j+1}^{N} f_{ij} \left\{ \epsilon_{ij} \left[\left(\frac{r_{ij}^0}{r_{ij}}\right)^{12} - 2\left(\frac{r_{ij}^0}{r_{ij}}\right)^{6} \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\}$$

where the first term on the right-hand side of the equation represents the energy between covalently bonded atoms, the second term (summing over angles) represents the energy due to the geometry of electron orbitals involved in covalent bonding, the third term (summing over torsions) represents the energy for twisting a bond due to bond order, and the last term represents the non-bonded energy between all atom pairs.

In general, MD simulations of proteins follow an established protocol usually starting with an experimentally determined crystal structure of the protein or a computationally generated model such as that derived using the Alphafold software[95,113,114]. To prepare the protein for MD, solvents, polyethylene glycol (PEG) and other unnecessary molecules that are a part of the crystal structure are stripped off the protein, also potential ligands not part of the original structure, such as products, substrates and substrate analogues, and hydrogen are added to the protein[112,113]. The next stage called parameterization is carried out if non-standard residues such as unusual PTMs, F430, CoM and CoB are present. This involves using a program like Gaussian to obtain compatible force field

parameters so that calculations can be carried out on the entire system. After all components of the system are parameterized, the minimization step is carried out to alleviate potential stearic clashes and to ensure that the atoms of the system (e.g., protein plus ligands) are placed in energetically favorable positions. Next, the minimized system is heated to the desired temperature using a thermostat. This serves to raise the temperature of the minimized system slowly to the specific biologically relevant temperature (e.g., 300 or 310 K). After the heating comes the equilibration step. During the equilibration phase, the system undergoes a brief simulation akin to the setup for the production or primary MD calculation. The main purpose of the equilibration step is to ensure that the system does not initiate a production run in a phase space region (coordinates of positions and velocities) that is significantly distant from equilibrium, as this could introduce biases into subsequent analyses. Data derived from MD simulation is analyzed in a number of ways, including root mean square deviation (RMSD) and root mean square fluctuations (RMSF) analyses and distance calculations[113].

RMSD is used in the assessment of two coordinate sets and provides a measurement of the difference between the two static structures. It is typically used in comparing the starting coordinates of a simulation with the coordinate at a chosen time point[113]. The RMSD between two proteins is typically reported as the minimum value between backbone atoms in the aligned structures.

While RMSD assesses two complete sets of coordinates, RMSF evaluates the coordinates of a specific portion of the system (e.g., an amino acid residue of an enzyme) in relation to its average position throughout the entire trajectory. RMSF serves as a valuable tool for identifying areas of conformational flexibility in enzyme MD simulations[113].

Distance calculations are used in determining the distance between any two atoms, such as the distance between a substrate atom and a catalytically relevant residue side chain.

4.2 Methods

4.2.1 Starting MCR crystal structures

Protein Data Bank (PDB) structures of MCR from *M. marburgensis* (5A0Y)[23] containing regular F430, MCR from ANME-1 (3SQG)[52] containing 17^2-methylthio-F430 and ECR from *Ca. E. thermophilum* (7BIS)[74] containing dimethylated F430 were obtained.

4.2.2 Parameterization of nonstandard residues

Thioglycine, 5-(S)-methylarginine, 1-N-methylhistidine, and S-methylcysteine were parameterized using MRP.py[115]. Restrained electrostatic potential (RESP) calculations were conducted in Gaussian using the B3LYP hybrid functional and the 6-31G* basis set. The resulting modified force field parameters were generated via the AM1-BCC[116] method implemented by the antechamber program within AMBER Molecular Dynamics software package[117–119]. For the post-translational modifications didehydroaspartate, 2-(*S*)-methylglutamine, 7-hydroxytryptophan, S-oxymethionine, 3-methylisoleucine, and the thioglycine/2-N-methylhistidine dipeptide, RESP calculations were performed in the gas phase using the Hartree-Fock level of theory with the 6-31G* basis set. Subsequently, the AM1-BCC method was utilized in antechamber to generate the necessary files for creating amber prep files through the prepgen tool within AmberTools20[119].

Cofactor F430 and its variants underwent parameterization using the MCPB.py program included in AmberTools20[120]. For coenzyme B, methyl- and ethyl-coenzyme M, CoB-S-S-CoM, methane, and ethane, parameterization was achieved using the antechamber and parmchk2 programs within AmberTools20, applying the AM1-BCC charge derivation method.

All nonstandard residues were parameterized according to the general amber force field (GAFF), except for Ni(I) atoms and their ligands within F430, which were assigned custom atom types through MCPB.py. Standard residues were handled using the amber ff14SB force field[121] and were explicitly solvated. Each residue was situated at least 12 Å away from the solvation box edge, employing the TIP3P force field[121]. The charge of the system was neutralized using Na^+ counterions.

4.2.3 Molecular dynamics simulations

The MD simulations of the MCR homologs began with 2500 steps of steepest descent optimization followed by 2500 steps of conjugate gradient minimization. Subsequently, a 2 ns heating phase was conducted, gradually increasing the temperature from 10 to 310 K. This heating step employed a Langevin thermostat and constant volume periodic boundary conditions[122]. During both the minimization and heating phases, 2 kcal mol^{-1} positional restraints were applied to backbone atoms.

Afterward, a 0.2 ns density equilibration phase was performed under constant pressure periodic boundary conditions, utilizing a 10 Å cutoff and a 1 ps pressure relaxation time. During this phase, 2 kcal mol^{-1} restraints were still imposed on backbone atoms. Subsequently, the simulations underwent further equilibration under constant pressure periodic boundary conditions for 2 ns with the same cutoff, but with a 2 ps pressure relaxation time and without restraints on backbone atoms. All heating, density, and equilibration steps were performed with a collision frequency of 5 ps^{-1}.

Following these equilibration steps, production simulations were conducted for 25 ns under similar conditions, but with a collision frequency of 2 ps^{-1}. To assess the impact of PTMs and/or F430 modifications on substrate binding, calculations for RMSD to the initial frame and per-residue

RMSF analyses were performed, along with atom-atom distance calculations, using the cpptraj program in Amber20[123].

4.3 Results

4.3.1 Evidence of active site cooperativity during MCR catalysis

Previously, Thauer in his review[81] suggested that though MCR can bind to both substrates independently, only one binding complex is productive, while the other is inhibitory. Evidence for this assertion was provided by structural studies[124] which indicates that one MCR active site was bound to methyl-coenzyme M and was productive while the other site was bound to coenzyme B and was inhibitory. This provided initial evidence of MCR exhibiting active site asymmetry and half site reactivity. Ermler et al.[30] determined from the crystal structure of MCR from *M. thermoautotrophicum* that the active site cavity was accessible through one channel, which was only wide enough for molecules up to a diameter of 6 Å. The crystal structure of MCR from *Methanothermobacter marburgensis*[23] show that in both active sites, coenzyme F430 was coordinating with its substrate (coenzyme M) in a symmetrical fashion as shown in Figure 54. In calculating the methane/ethane distance from the active site Ni-F430, we reason that as a substrate (in ANME-1 MCR and in ECR) or as a product (in *M. marburgensis* MCR), methane/ethane is expected to be in close proximity to Ni for catalysis to occur. The substrates/products in the relevant MCR structures used for this study were symmetrically bound to its active sites, the MD simulations were started with this symmetrical starting point, and then became asymmetric with methane/ethane close to Ni in one active site, and further away in the other site, providing evidence of cooperativity between both active sites (Figure 55).

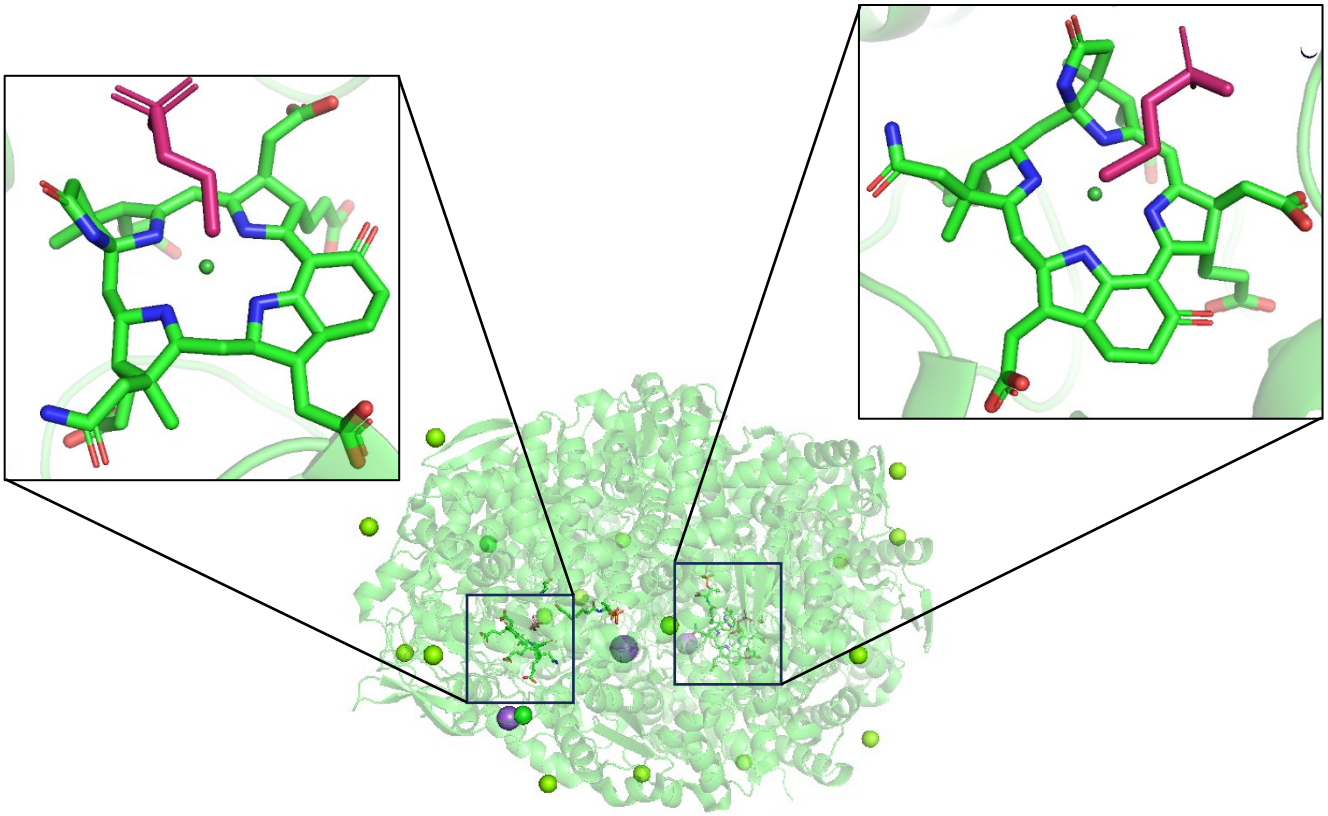


Figure 54: Crystal structure of MCR (5A0Y) showing both coenzyme F430 in close proximity with magenta colored CoM

The first evidence of the half site reactivity exhibited by MCR active sites was provided by distance calculations between the methane/ethane substrate and Ni using the entire trajectory of the MD run as shown in Figure 56.

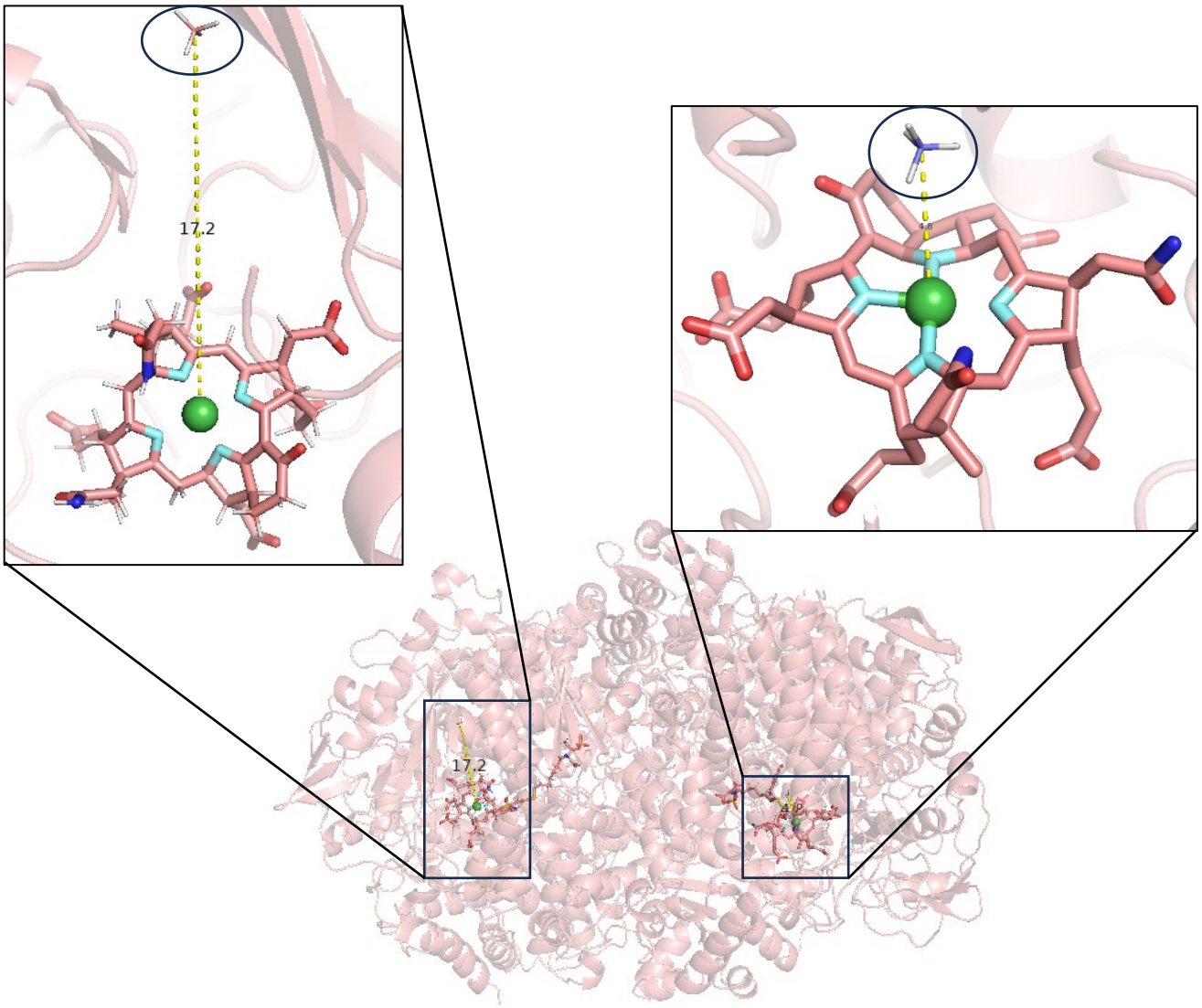


Figure 55: Closest to average structure of wild type MCR from M. marburgensis showing evidence of active site asymmetry/half sites reactivity between both active sites.

As shown in Figure 55 which is a closest to average structure of MCR, methane is within an average distance of 4.8 Å from the Ni in only one of the active sites and is further away (17.2 Å average distance) in the other active site. A distance analysis comparing the distance of both methane molecules to the Ni atom confirmed the observation that only one methane was coordinating to the active site per time. It is worthwhile to note that the other methane remained at the same position for the entire duration of the production run (Figure 56), which suggests that it is bound to a distal methane binding pocket. Both active site Ni atoms were properly coordinated to their respective tetrapyrrole nitrogen, and to the lower axial ligand (Gln 147) (data in Appendix).

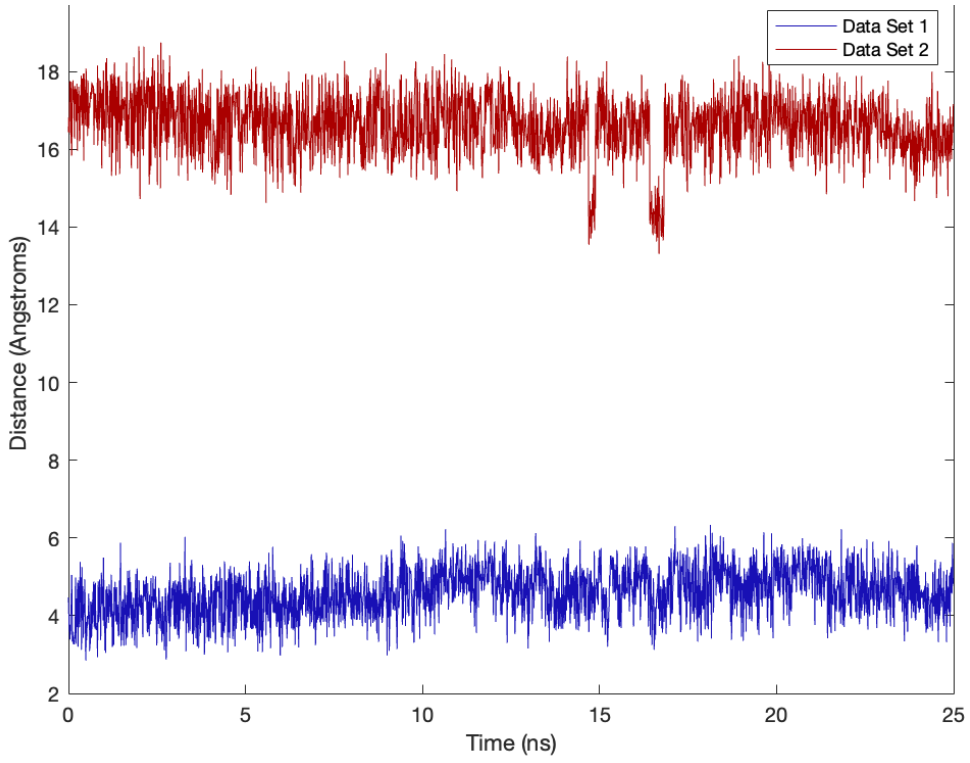


Figure 56: Distance plot showing methane distance to Ni in MCR. Methane remains closely bound to Ni within a distance of 4.8 Å on average in only one of the active sites. Data sets 1 and 2 refer MCR active sites 1 and 2

Similarly, distance plots for ANME-1 MCR show that one methane remains bound/close to the active site Ni, while the other is at a considerable distance from Ni. A similar result was observed when production run for ECR was carried out, only one ethane was in a close distance to Ni while the other ethane is in a distal pocket (Figure 57).

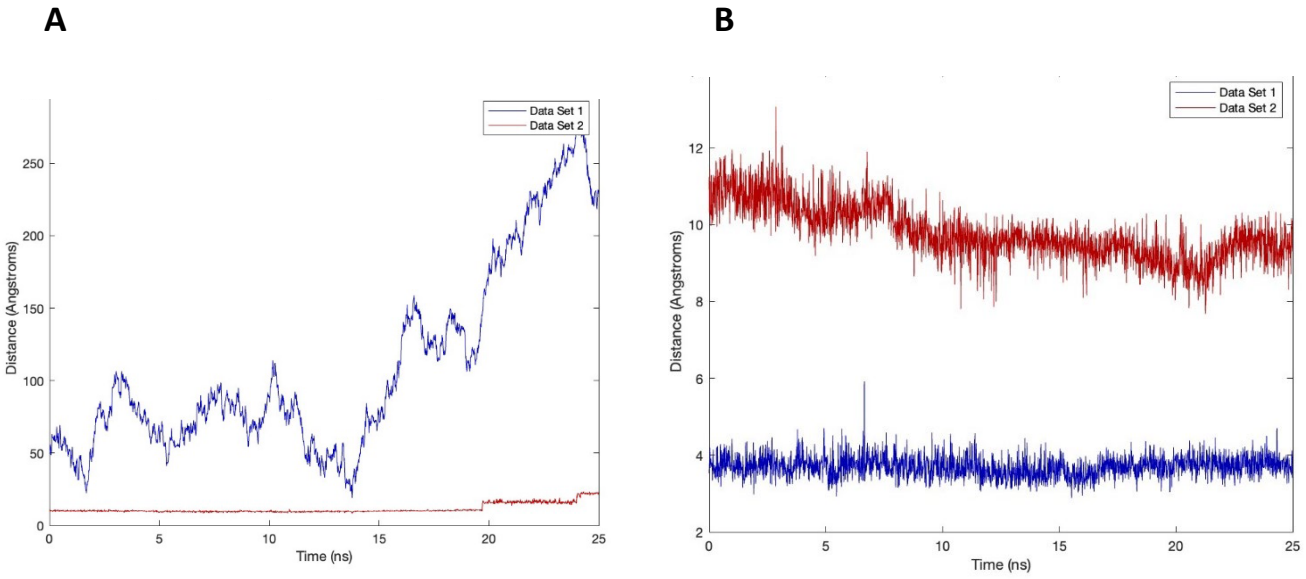


Figure 57: Distance plots of MCR from ANME-1 and ECR showing (A) methane distance to Ni, (B) ethane distance to Ni. Data sets 1 and 2 refer to MCR/ECR active sites 1 and 2

A study by Duin and McKee[125] on the mechanism of methane production from MCR using density functional calculations, proposed that in one of the intermediate steps of catalysis, Ni moves out of the plane of the four tetrapyrrole nitrogen upon protonation of a C-ring nitrogen atom. The protonated C-ring oxidatively adds CH_3SCoM^- to Ni, giving rise to a 4-coordinate Ni center with Ni bound to two tetrapyrrole nitrogen, methane, and CoM. This was not observed in this study.

4.3.2 Effect of post-translational modification on the active site of MCR and on ECR

MD runs were carried out with and without the presence of the PTMs found in MCR. The crystal structure of MCR reported by Ermler[30] revealed that F430 was coordinated to six ligands including the four tetrapyrrole nitrogen atoms (approximately 2.2 Å distance), the thiol group of CoM (2.4 Å distance) and the oxygen atom of a side chain α Gln 147 (2.3 Å distance). Hahn and colleagues[74] in reporting the crystal structure of ECR, showed that the Gln 147 residue in MCR is replaced by Met 181 which was at a 3.17 Å distance from the Ni atom.

4.3.2.1 Effect of PTMs on MCR

The absence or presence of the PTMs made no significant difference on the tetrapyrrole nitrogen distance to Ni. Ni remained bound to its tetrapyrrole nitrogen unlike what was proposed by Duin and McKee[125]. The Gln 147 distance to Ni (See Appendix) also remained the same in WT-MCR and in MCR without PTMs. However, for methane distance to Ni (Figure 58), the data shows a loss of active site asymmetry in MCR without PTMs, which suggests that the PTMs may be playing a role in synchronizing active site cooperativity through some form of allosteric communication. The data suggests that the absence of PTMs resulted in the loss of allosteric communication as shown in Figure 59.

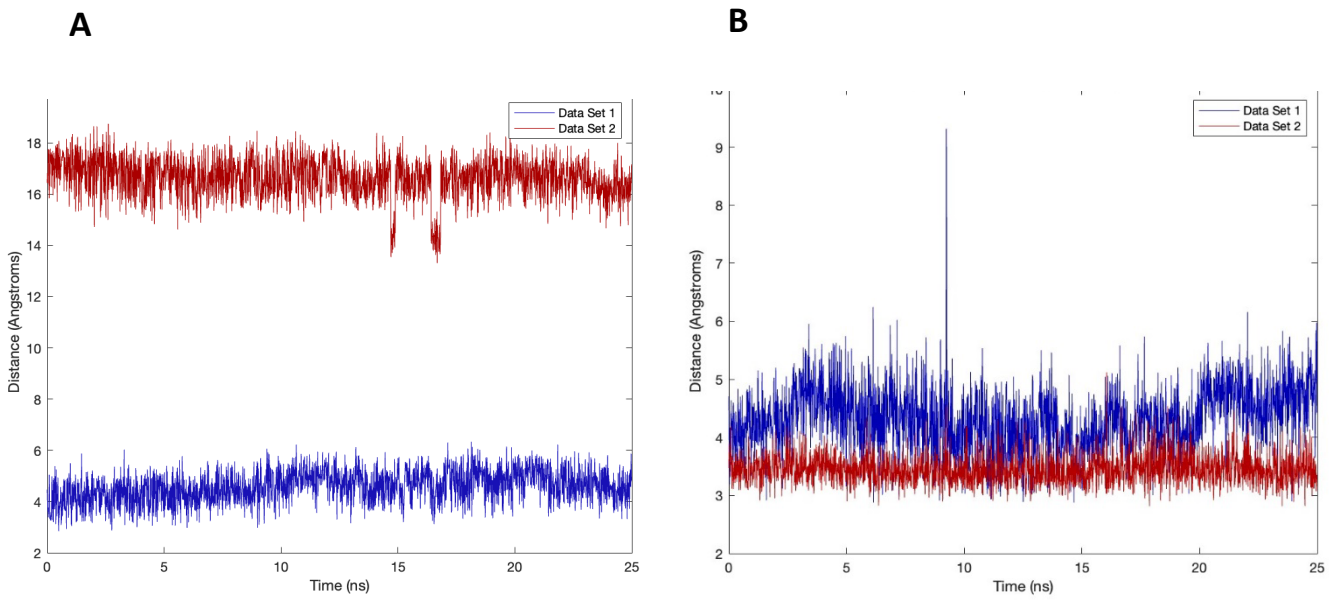


Figure 58: Distance plots of MCR showing (A) methane to Ni distance with WT MCR and (B) methane to Ni when MCR without PTMs was used for MD. Data sets 1 and 2 refer to MCR active sites 1 and 2

Unlike the wildtype where methane was close to Ni in only one active site of MCR, without the PTMs, both methane molecules were close to both active sites of MCR, showing a loss of half site reactivity.

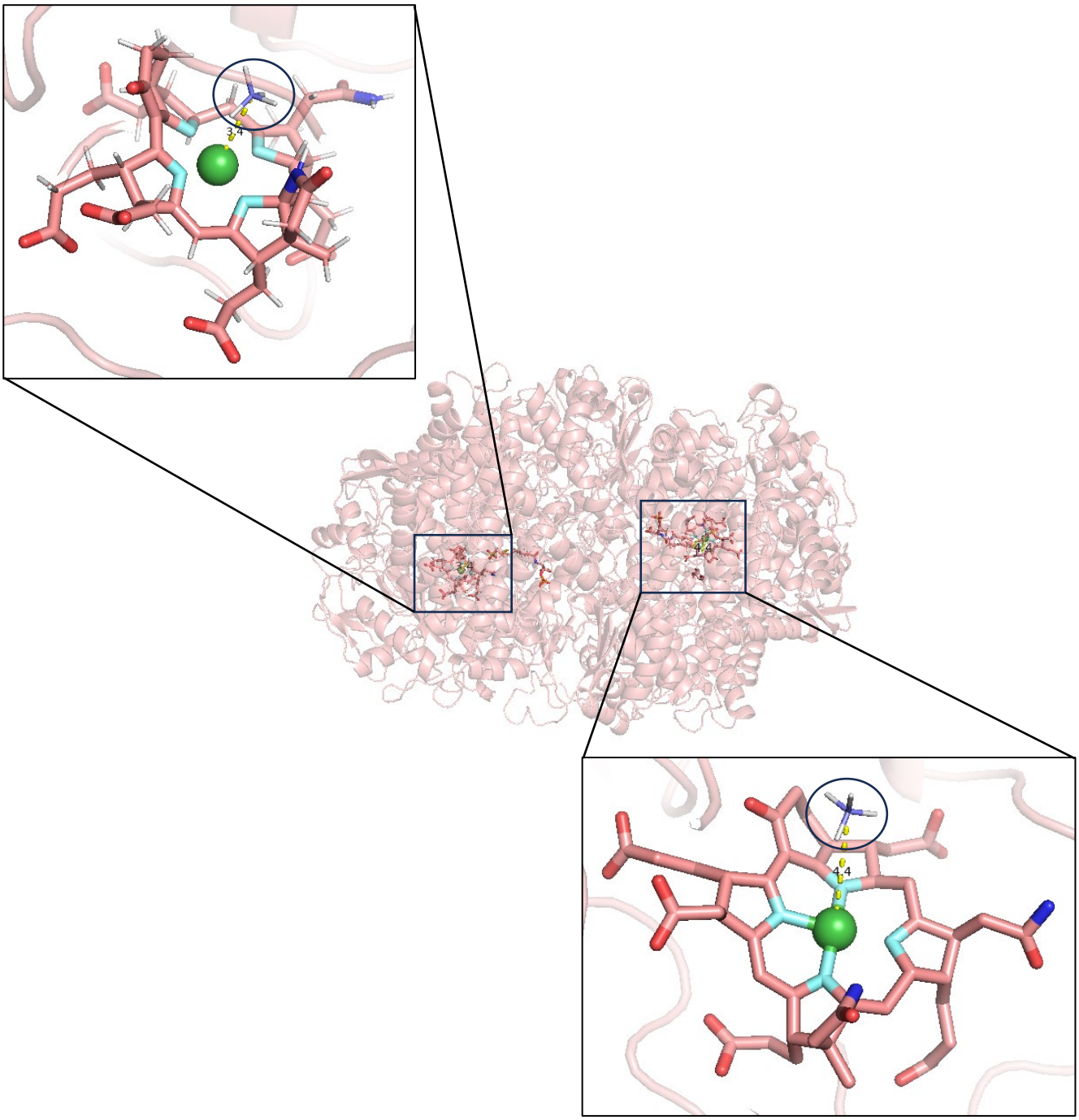


Figure 59: Closest to average structure of MCR from M. marburgensis without PTMs showing a loss of active site asymmetry.

4.3.2.2 Effect of PTMs on ECR

For ECR, no difference was observed in the tetrapyrrole nitrogen distance to Ni (Figure 61) when comparing the WT ECR to the ECR without PTMs. The distance between the lower axial methionine ligand distance to Ni was also within the range observed by Hahn and his group[74] (Figure 60) and did not change significantly in the WT ECR and ECR without PTMs.

Again, in observing the ethane distance to Ni distance for the WT and for the ECR without PTMs, inference can be made that the asymmetry between the active sites was affected by the absence of the PTMs. One ethane stayed within a close distance to the active site, while the other ethane alternated between staying close to the active site and moving away from it (Figure 62).

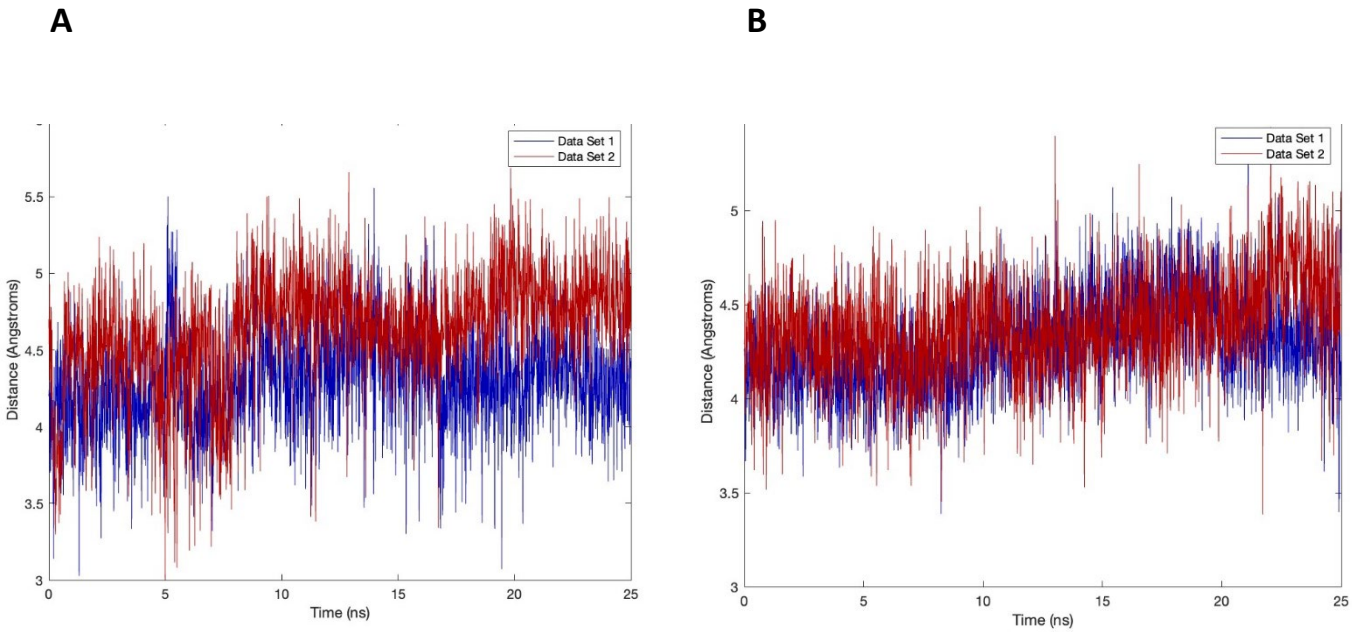


Figure 60: Distance plots of ECR showing (A) Met 181 distance to Ni with WT ECR and (B) Met 181 distance to Ni when ECR without PTMs was used for MD. Data sets 1 and 2 refer to ECR active sites 1 and 2

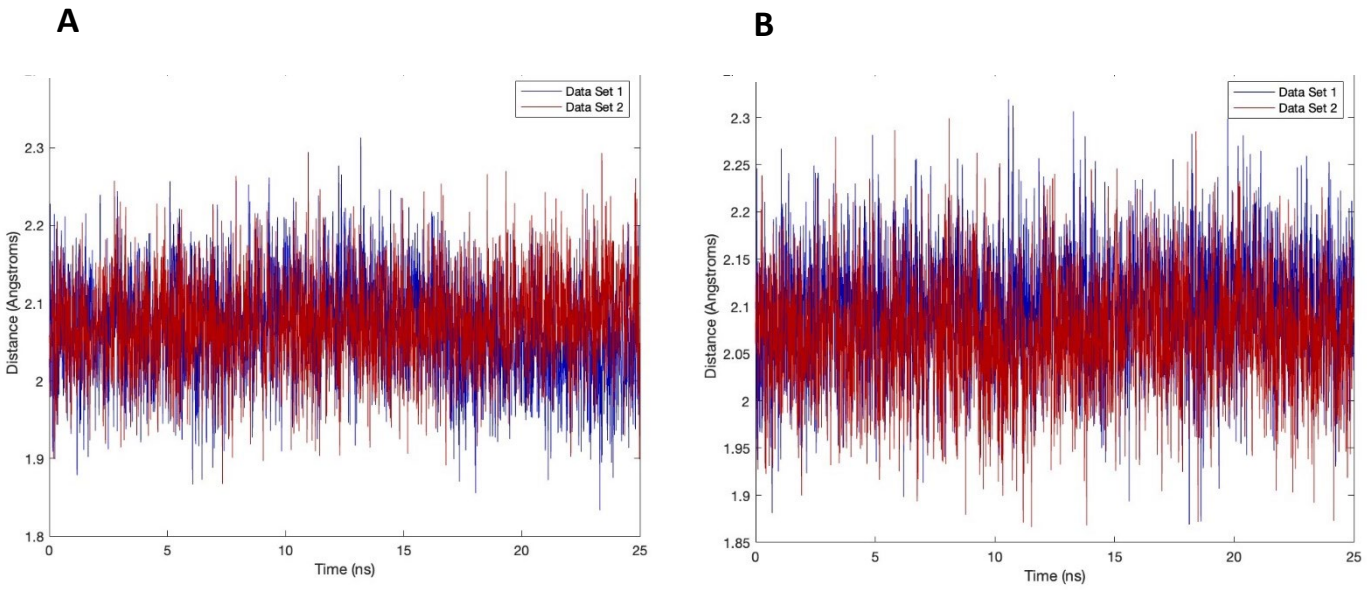


Figure 61: Distance plots of ECR showing (A) tetrapyrrole nitrogen distance to Ni with WT ECR and (B) tetrapyrrole nitrogen distance to Ni when ECR without PTMs was used for MD. Data sets 1 and 2 refers to ECR active sites 1 and 2

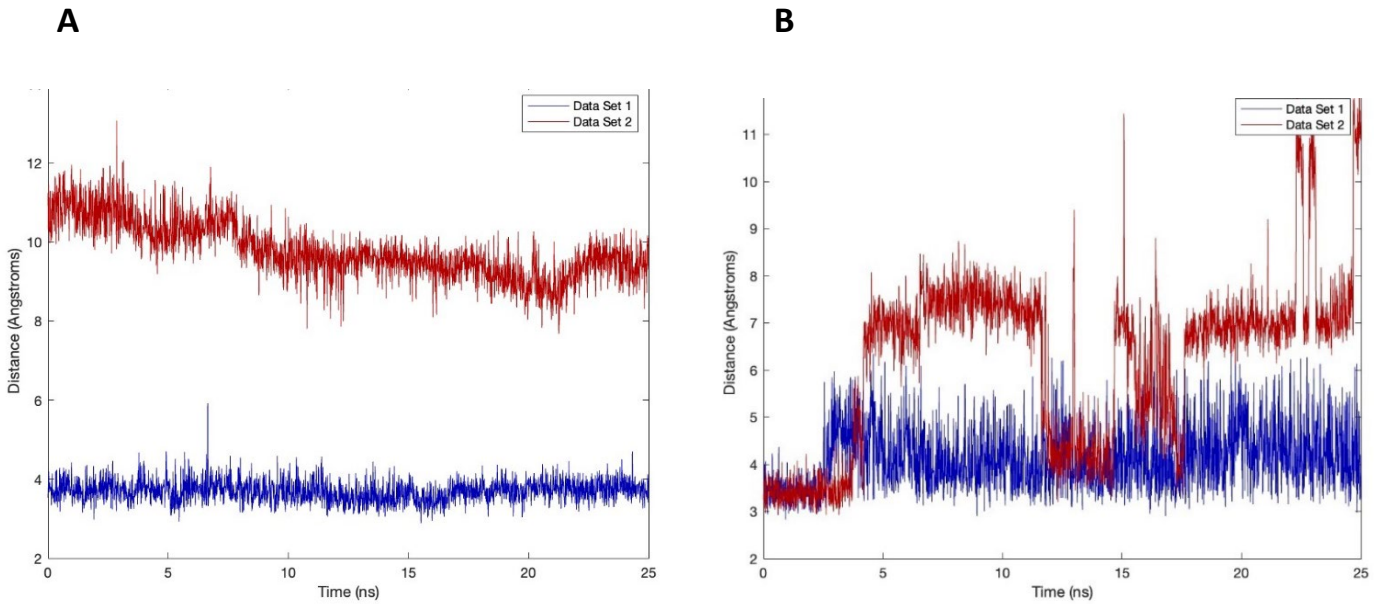


Figure 62: Distance plots of ECR showing (A) ethane distance to Ni with WT ECR and (B) ethane distance to Ni when ECR without PTMs was used for MD. Data sets 1 and 2 refer to ECR active sites 1 and 2

4.3.3 Effect of F430 modification on active sites of ANME-1 MCR and ECR

MD runs were also performed on MCR from ANME-1 and on ECR with their native F430 and with the regular F430.

4.3.3.1 Effect of F430 modification on ECR

When regular F430 was used in the production run of ECR instead of 17,17[2]-dimethyl F430, no significant difference was observed in nitrogen and Met 181 distance to Ni. Distance plots of lower axial Met 181 distance to Ni (Figure 63) and of tetrapyrrole nitrogen distance to Ni (Figure 64) was similar to that of the wildtype. Both plots show that the F430 was fairly well coordinated to its ligands and stable for the duration of the production run.

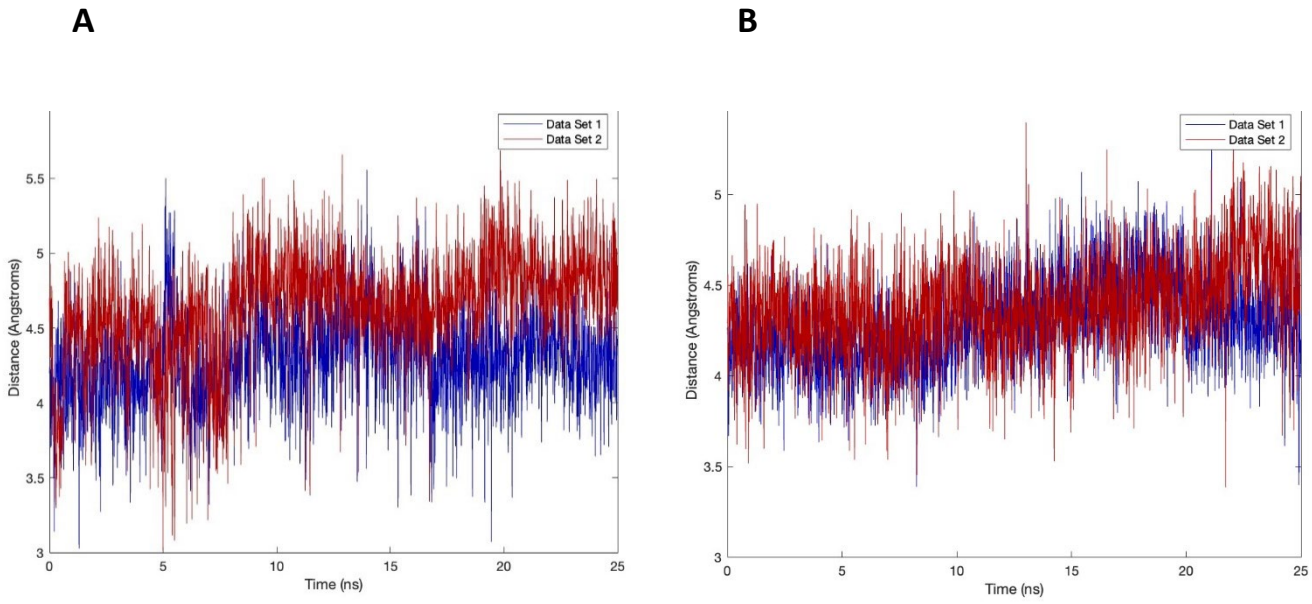


Figure 63: Distance plots of ECR showing (A) Met 181 distance to Ni with WT ECR and (B) Met 181 distance to Ni when ECR with regular F430 was used for MD. Data sets 1 and 2 refer to ECR active sites 1 and 2

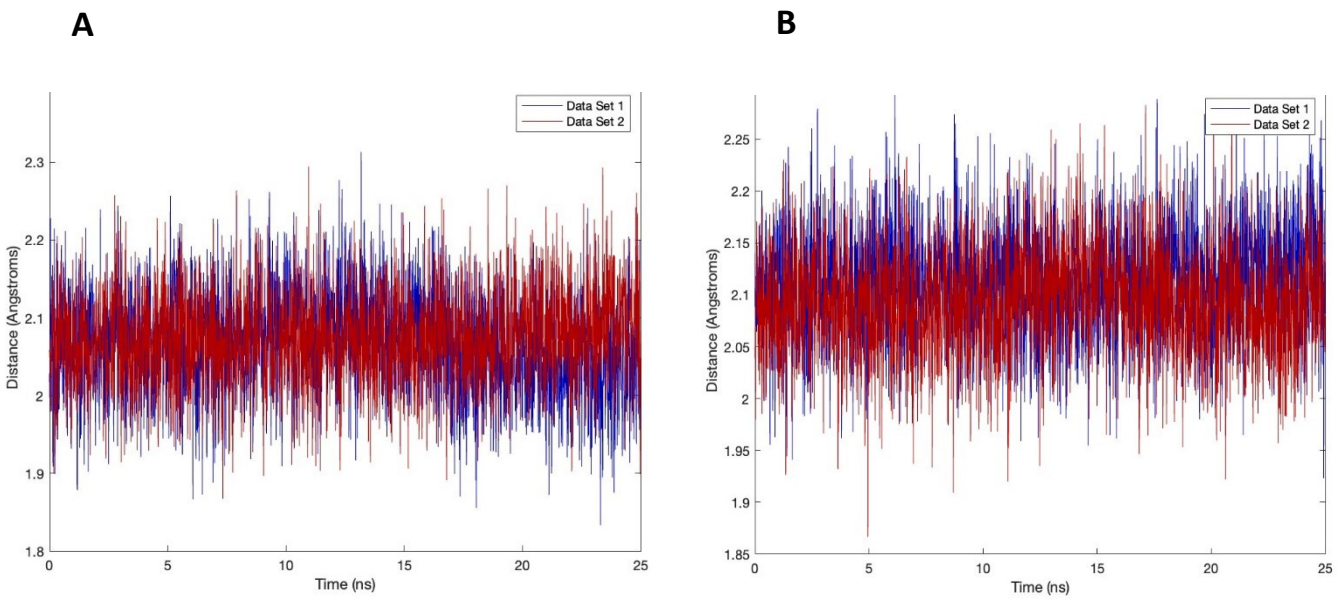


Figure 64: Distance plots of ECR showing (A) tetrapyrrole nitrogen distance to Ni with WT ECR and (B) tetrapyrrole nitrogen distance to Ni when ECR with regular F430 was used for MD. Data sets 1 and 2 refer to ECR active sites 1 and 2

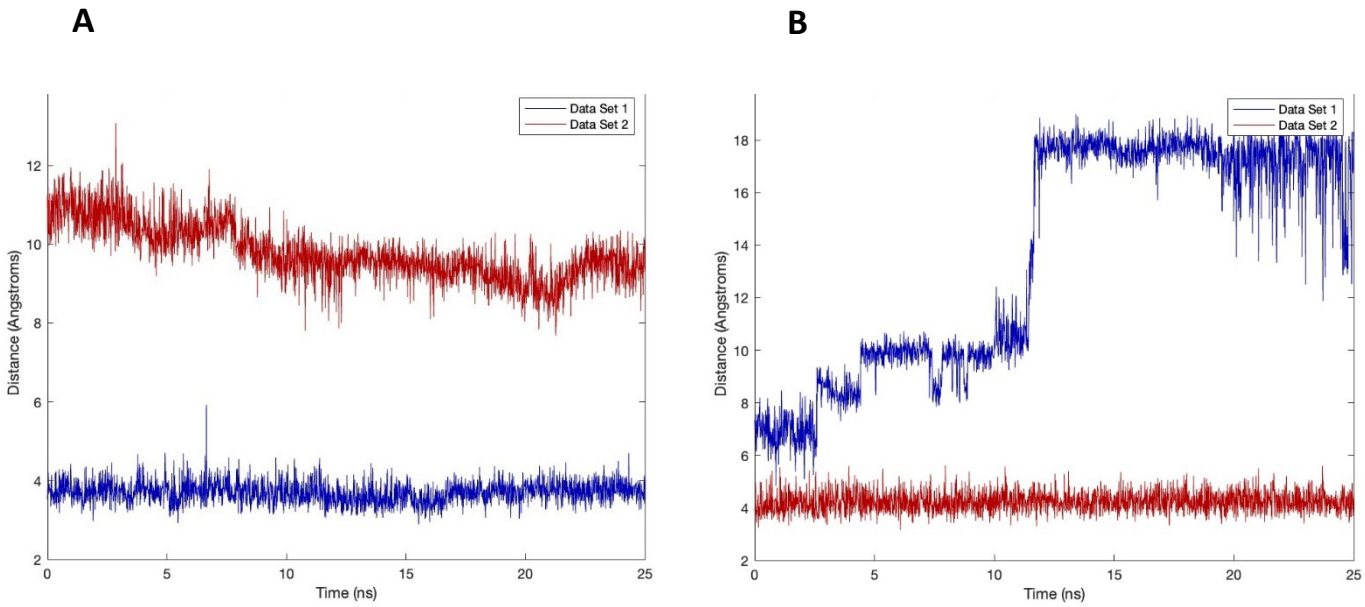


Figure 65: Distance plots of ECR showing (A) ethane distance to Ni with WT ECR and (B) ethane distance to Ni when ECR with regular F430 was used for MD. Data sets 1 and 2 refer to ECR active sites 1 and 2

Distance plots of ethane distance to Ni in the wildtype ECR and in ECR with regular F430 (Figure 65) show consistency with half-sites reactivity previously reported (4.3.1). With regular F430 however, while one ethane stays in close proximity to Ni, the other ethane molecule starts out pretty close (approximately 6 Å distance from Ni[30]) and escapes to a distal site towards the end of the simulation, where it is getting ready to dissociate from the enzyme. Inference can be made that the distal ethane is not bound tightly to the active site. This data suggests that the F430 modification could conceivably play a role in modulating the affinity of the Ni to ethane.

4.3.3.2 Effect of F430 modification on ANME-1 MCR

For WT ANME-1 MCR and ANME-1 MCR with regular F430, the tetrapyrrole nitrogen was stable and in close distance to Ni in both active sites (Figure 67).

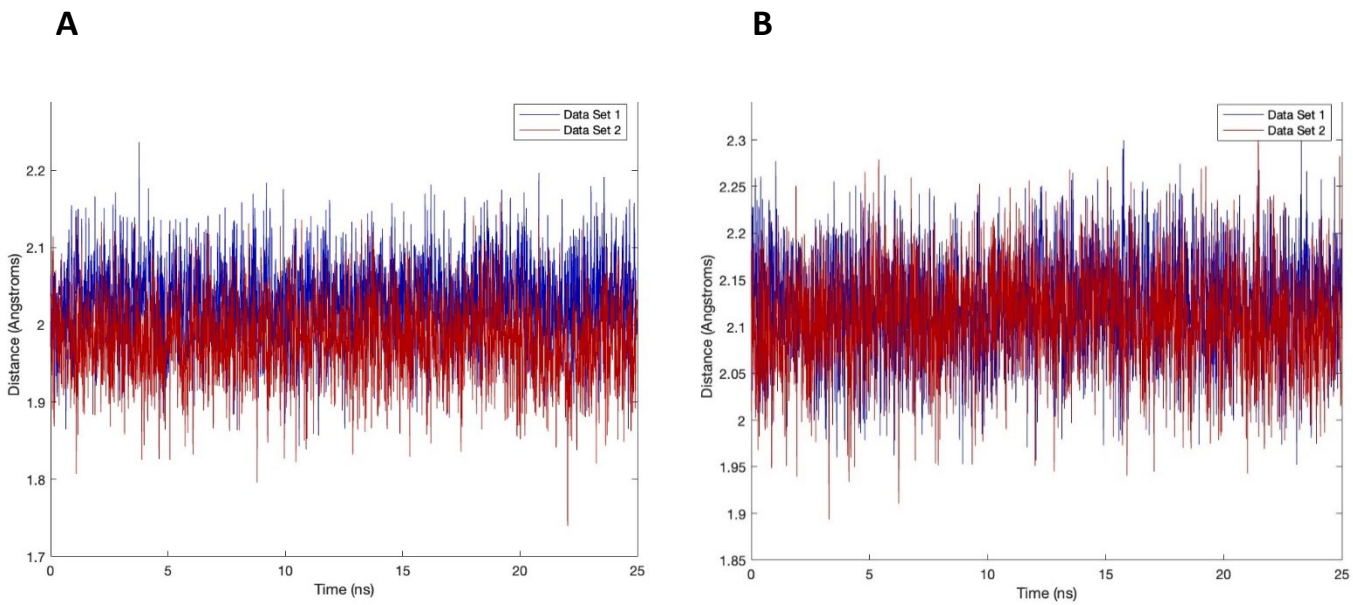


Figure 67: Distance plots of ANME-1 MCR showing (A) tetrapyrrole nitrogen distance to Ni with WT MCR and (B) tetrapyrrole nitrogen distance to Ni when MCR with regular F430 was used for MD. Data sets 1 and 2 refer to MCR active sites 1 and 2

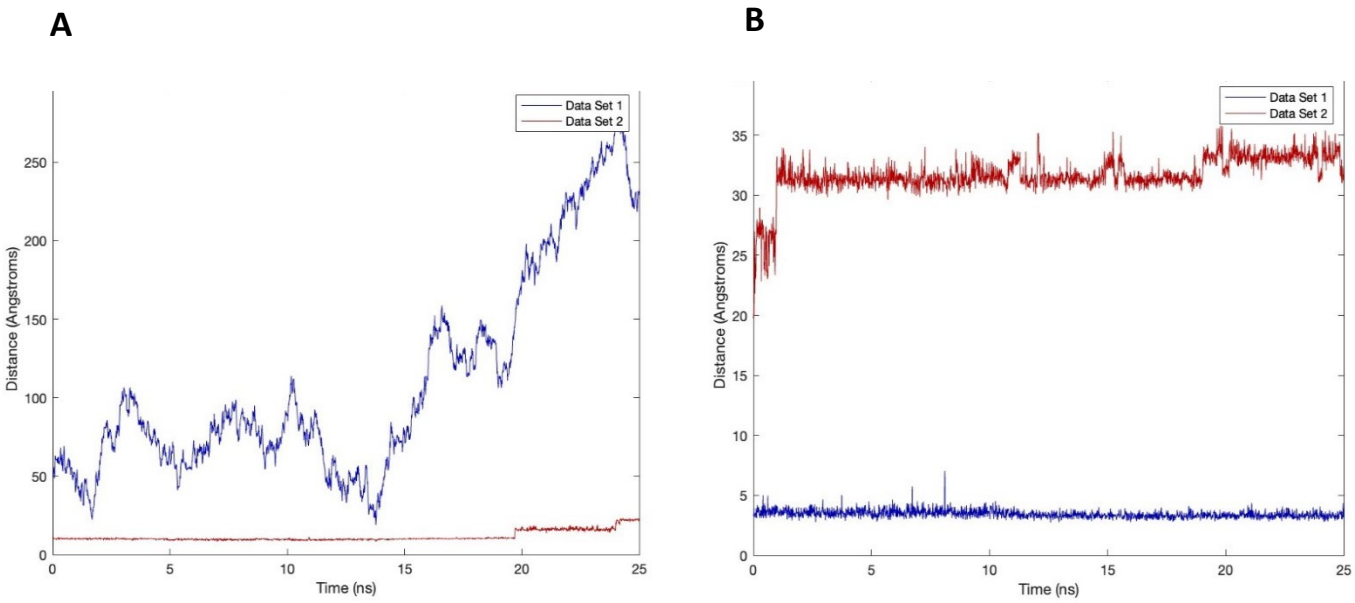


Figure 68: Distance plots of ANME-1 MCR showing (A) methane distance to Ni with WT MCR and (B) methane distance to Ni when MCR with regular F430 was used for MD. Data sets 1 and 2 refer to MCR active sites 1 and 2

In analyzing methane distance to Ni in both WT ANME-1 MCR and in ANME-1 MCR with regular F430, half site reactivity is observed. For MCR with regular F430, the distant methane stays at a fairly stable position, possibly in a distal binding pocket, for the entire duration of the run (Figure 68 and Figure 69). This is similar to what was observed for ethane binding in ECR with regular F430 (Figure 65). For the wildtype, the distant methane is not tightly bound and was completely dissociated from MCR at the end of the run.

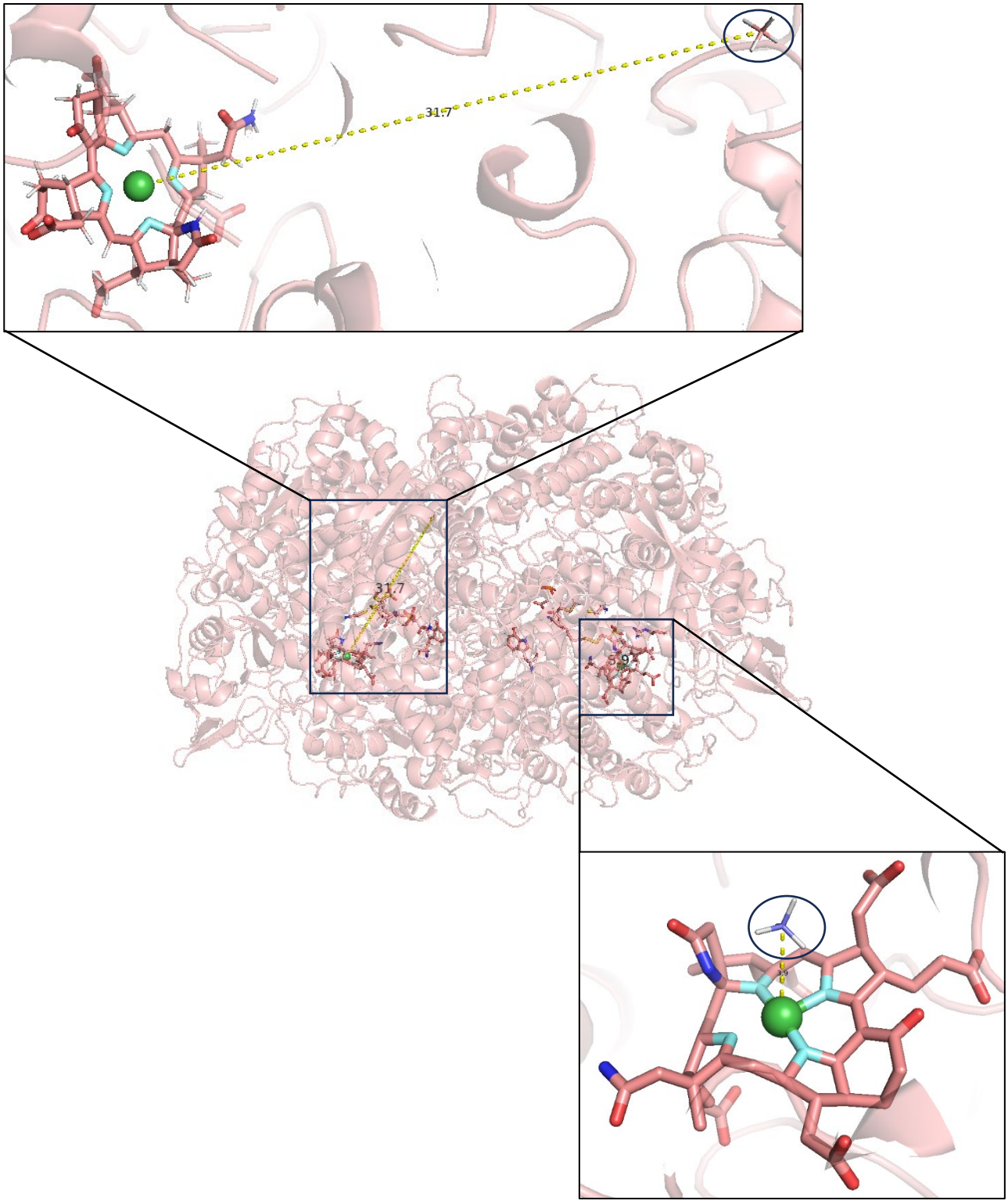


Figure 69: Closest to average distance structure of ANME-1 MCR with regular F430 displaying active site asymmetry. Methane distance to Ni in one active site is 3.9 Å, while it is 31.7 Å in the other active site

4.3.4 MCR post translational modifications play a role in the dynamism of McrA residues

Comparison of RMSF plots of MCR from *M. marburgensis* with and without its PTMs show that without PTMs, the movement of the amino acid residues in McrA were more attenuated. As shown in the plots (Figure 70), both McrA subunits are attenuated when the PTMs were deleted.

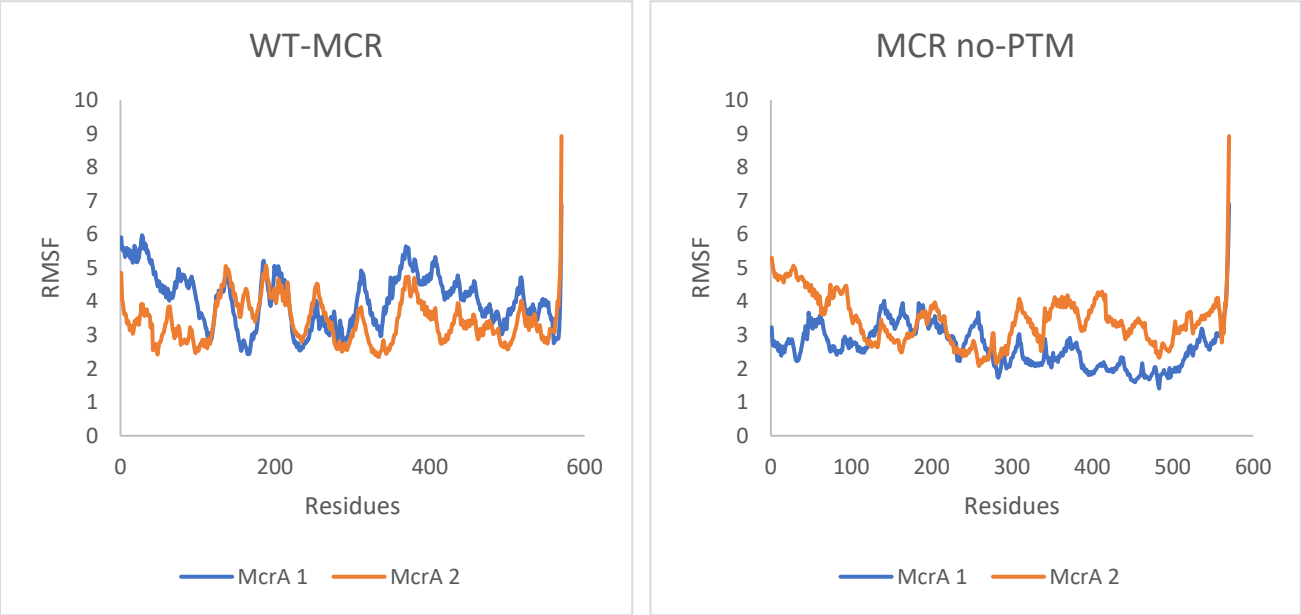


Figure 70: RMSF plots of post MD MCR with and without PTMs

4.4 Discussion

4.4.1 MCR active sites exhibit half sites reactivity

Half sites reactivity, a type of cooperativity, between both MCR active sites was strongly suggested by data from the MD runs. Cooperativity is usually exhibited by enzymes with more than one binding site and is characterized by substrate binding to one active site influencing substrate binding to the other active site. In this case, it was illustrated by the fact that methane was bound close to the reactive Ni center in only one of the active sites. In the other site, methane was generally found in a remote binding pocket, waiting its turn, as shown in the closest to average structures (Figure 55 and Figure 69). More specifically, the data indicates active site asymmetry, given that the simulation started with roughly symmetrical crystal structures of MCR. This was consistent with all MCR structures analyzed in this study and was corroborated by the distance plots in Figure 56 and Figure 57. We can infer from the data that both MCR active sites communicate with each other via some form of allosteric communication mechanism facilitated by the PTMs.

4.4.2 Post translational modifications may play a role in allosteric communication between MCR active sites

For both MCR, ECR and ANME-1 MCR, absence of the PTMs had no impact on the distance of the tetrapyrrole nitrogen to Ni. This makes sense when you consider that the Ni-bound coenzyme F430 is an intact unit and can exist as a separate entity from the rest of the MCR protein. While as a catalytic center of MCR, it can be influenced by the presence and absence of substrates, it stands to reason that it will not be impacted on by the slight variations accruing from the absence of the PTMs on the implicated residues. This is in contrast with the results of the study by Duin and McKee[125] where it is proposed that one of the tetrapyrrole nitrogen (ring C) is uncoordinated from the Ni during the catalytic cycle, leading to Ni being out of plane and coordinated by two tetrapyrrole nitrogen, methane, and CoM. It is worth mentioning that DFT calculations were used

for that study while MD was used here. Also, it is possible that both studies did not simulate MCR in the same states.

A close look at the methane/ethane to Ni binding pattern in post MD MCR and ECR without the PTMs show that PTMs coordinate allosteric communication between MCR active sites under the conditions tested. For MCR without PTMs from *M. marburgensis*, data (closest to average structure and distance calculations) show that while one active site methane was within 3.5 Å distance from Ni and at a stable position for the entire duration of the run, the other methane fluctuated between 4 – 6 Å for the most part, and stayed in the area around the enzyme active site (Figure 71). For ECR without PTMs, the fluctuation of one ethane molecule between being in close distance to Ni and being far away at a distal binding site was more dramatic, while the other ethane molecule stayed within the 6 Å of the active site.

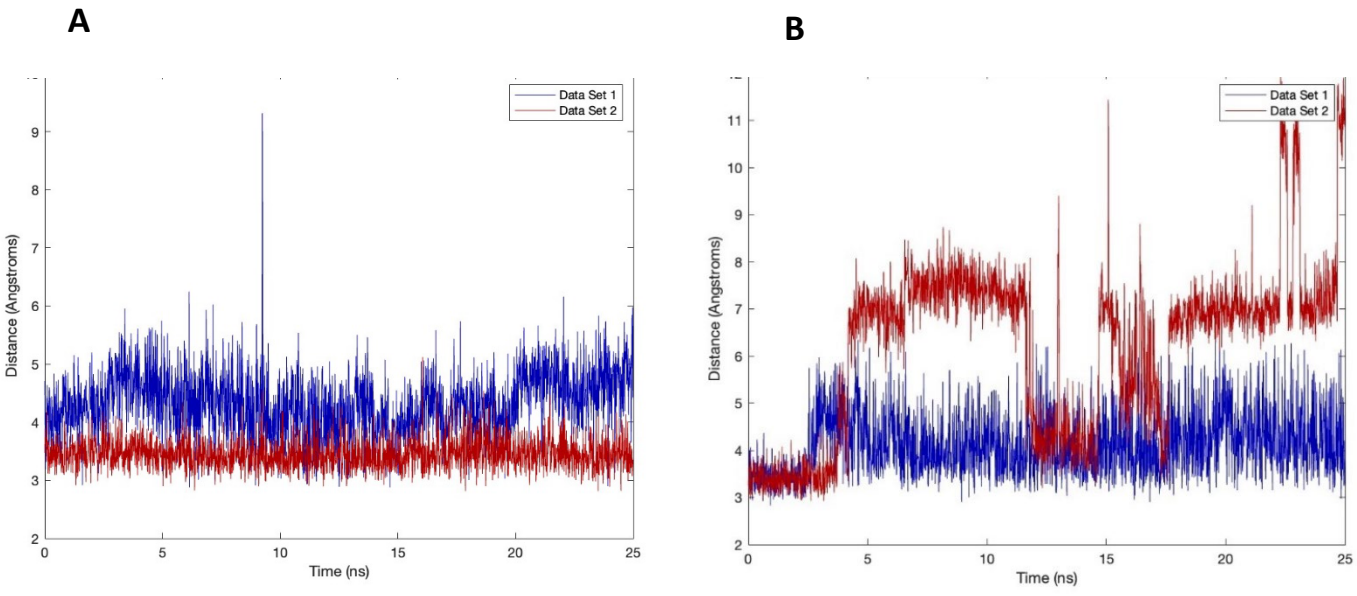


Figure 71: Distance plots showing (A) methane distance to Ni in MCR without PTMs and (B) ethane distance to Ni in ECR without PTMs. Data sets 1 and 2 refer to MCR/ECR active sites 1 and 2

From the data in Figure 58 and Figure 62, an argument can be made that the PTMs posit a geometric hindrance to the binding of one substrate in one active site, thereby playing a role in allosteric communication of the active sites, leading to the observed half sites reactivity. Without the PTMs, both substrates migrate freely to their active sites. Inference can be made that the PTMs facilitated some form of communication between MCR active sites, and this communication was lost in the absence of the PTMs. Previously, Deobald and colleagues had posited[92] that the methylation of arginine in MCR from *M. acetivorans* resulted in the formation of a salt-bridge between the arginine and a nearby glutamate of the McrB subunit. Without the methylation, the arginine residue is more flexible and can protrude into the substrate channel[31,92]. Conformational changes such as this may explain how the PTMs function to aid cooperativity in MCR active sites, and how their absence can result in the loss of cooperativity.

The effect of the PTMs on MCR was further illustrated by the RMSF plots (Figure 70) showing the relative distance of each residue from the average moved by each residue during the run. All PTMs are located in the McrA subunit, and the RMSF plot is a glimpse into the movement of each residue located in McrA during the course of the run. In comparing the wildtype MCR to that without PTMs, it can be inferred that the diminished movement of residues in MCR without PTMs impacted the allosteric communication between both active sites, leading to loss of negative cooperativity. This finding corroborates the hypothesis by Nayak and colleagues[62] showing that the PTMs were necessary for the enzyme's adaptation to mesophilic conditions. Movement of an enzyme's amino acid residues is essential to active site interaction with its substrate and to catalysis in general. Absence of MCR PTMs diminishes the movement of its residues which limits its activity, and this is more pronounced when the enzyme is exposed to lower temperatures. Methanogens are very old organisms that probably existed under thermophilic environments. It is

very possible that proto-MCR did not possess PTMs and that the PTMs on MCR are an adaptation to cooler temperatures as the earth evolved. To remain functional at lower temperatures, PTMs were introduced to the enzyme to enhance its dynamics and fine-tune the cooperativity of the enzyme.

In conclusion, for MCR, ANME-1 MCR and ECR studied under the various conditions, one thing was consistent - the half-sites reactivity of MCR active sites observed. PTMs were also found to influence allosteric communication and are potentially important for the adaptation of MCR to mesophilic conditions. For ECR, the data suggests that the dimethyl modification of F430 function to enhance the binding affinity of Ni to ethane. Future work should expand this study to include studying the effect of PTMs on ANME-1 MCR and calculating the dissociation constant (Kd) of methane/ethane to Ni using a program like MMPBSA. Cooperativity of MCR active sites can also be studied through enzyme kinetic assays using wildtype MCR and MCR without its PTMs.

Chapter 5

Discussion and future work

5 Chapter 5

5.1 Discussion and future work

Research into methyl coenzyme M reductase MCR is riddled with so many unknowns about the maturation and assembly process, the enzyme mechanism, and the exact function of its post translational modifications (PTMs). Considering its unique role in methane production and consumption, which is implicated in global warming and in the recycling of natural gas, strategies geared towards alleviating these burdens require a comprehensive, in-depth knowledge of the enzyme, and ways of effectively regulating it.

Data from my work gives insight into the possible functions of MCR PTMs and how they impact MCR activity. While previous work has demonstrated the essentiality of coenzyme F430 to MCR activity [30,63,65,67], the exact function of the PTMs has not yet been established, though the general consensus is that the PTMs neither impact the overall geometric structure of the enzyme, nor its catalytic ability[33,53,55,62,126]. Nayak et al.[62] postulated that the methylcysteine modification might be important to the enzyme's adaptability to mesophilic conditions, and work by Radle and colleagues[33] opines that the methylarginine modification improves the thermal stability. However, the question does arise on the rationale warranting the existence of these energy-intensive modifications in nature for any reason short of its contribution to the enzyme function.

Data from this research provides useful information on the function of MCR PTMs. Molecular dynamics (MD) simulations of MCR from *M. marburgensis*, and ANME-1 and of ethyl-coenzyme M reductase (ECR) from *Ca. E. thermophilum* showed evidence of negative cooperativity between active sites. Allosteric communication between active sites in MCR without its PTMs was significantly attenuated. In addition, PTMs appear to improve the dynamism of McrA residues as shown in Figure 70. Thus, the introduction of the PTMs may have been one of the strategies utilized

for the adaptation of proto-MCR from thermophilic early-earth conditions to mesophilic growth as the earth temperatures cooled, by enabling the proper enzyme dynamics at lower temperatures and absence of PTMs led to reduced movement of McrA residues. These findings present interesting opportunities for further research to substantiate MCR active site cooperativity and how it is affected by PTMs.

Active site cooperativity and the influence of PTMs on it can be determined via MCR kinetic assays using wildtype MCR and MCR variants without PTMs. Activity of MCR without PTMs is expected to differ from that of wildtype MCR. Furthermore, hydrogen deuterium exchange (HDX) experiments can be used to determine the dynamics of the enzyme, and changes in solvent accessibility. Deletion of the PTMs should decrease the dynamics of MCR, which should lead to a net decrease in the rate of exchange/deuterium uptake.

The gene responsible for methylhistidine is yet to be discovered, and while the *mm4* gene was investigated in this study due to its annotation as a methyltransferase, this study has successfully established that it is not responsible for the methylhistidine modification. Analysis of the protein structure also revealed that the protein is more likely to be a dehydrogenase or an acyl-coA transferase due to its close structural alignment with proteins in those classes. The gene candidate *MA0673* may be a more likely candidate for the methylhistidine modification.

In studying the role of *mcmA* gene in the methylcysteine modification of MCR, structural prediction analysis show that the McmA protein requires the cofactor Zn for activity. The presence of several cysteine residues on McmA led to the inference that the protein would purify better anaerobically. Expression of the cell line with zinc supplementation and purification of McmA anaerobically led to an increased yield of the McmA protein. Expression and purification of soluble McrA for use in *in vivo* assays was also successfully carried out, with McmA protein co-eluting

with McrA in the cell line harboring both genes. Furthermore, enzyme activity assays, *in vivo* experiments, and structural prediction studies, provided data which strongly suggests that the methylcysteine modification is co-translational rather than post-translational.

Finally, attempts to express MCR in a heterologous host such as *Escherichia coli* resulted in the successful expression of soluble McrA and McrG proteins in different non-methanogenic cell lines used for this study. It unearthed evidence of interaction between McrG and the Cfb proteins, giving credence to the discovery by the Mansoorabadi group that CfbE catalyzes the last step in F430 biosynthesis[10,80], interacts with McrD and may be activated by it[80]. The interaction of Cfb proteins with McrG hints at the possibility of the existence of a coenzyme F430 synthase complex which coordinate with McrD (which serves as a chaperone) to insert the coenzyme F430 into MCR. Future work can explore expressing the Cfb proteins with McrABCDG singly and in combination, as well as using x-ray crystallography and HDX experiments to obtain and study crystal structures and/or stable complexes of the proteins to ascertain the exact nature of the interaction between Cfb proteins and the Mcr proteins.

Appendix

6 Appendix

6.1 Making of competent cells using calcium chloride

A single cell colony from an agar plate was inoculated into a 15 mL tube containing 5 mL of lysogeny broth (LB) and left to shake overnight at 37 °C. 200 µL of the overnight culture is used to inoculate 20 mL of LB and left to shake for 2 - 3 hours at 37 °C. the cells are placed on ice after 3 hours for 10 minutes and collected by centrifugation at 6000 rpm for 3 minutes. Supernatant is decanted and the cells are gently resuspended in 2 mL cold 0.1 M calcium chloride solution. Cells are incubated on ice for 20 minutes, centrifuged at 6000 rpm for 3 minutes and the supernatant is decanted. Cells are gently resuspended in 1 mL of cold 0.1 M calcium chloride solution containing 15 % glycerol. Finally suspended cells are dispensed in 100 µL in 1.5 mL tubes, frozen immediately in liquid nitrogen and stored in the -80 C refrigerator. Competent cells are subsequently used for cell transformation.

6.2 Gibson Assembly protocol

Freshly purified DNA fragments from PCR are incubated in a plasmid:gene ratio of 1:3 to a total volume of 10 µL. 10 µL of the Gibson Assembly Master mix is added to the reaction mix, and mixture is incubated for 15 minutes at 50 °C. When large gene fragments are being assembled, 2 µL of DMSO is added to the reaction mix. Following incubation, samples are stored on ice or at –20 °C in readiness for transformation of a competent cell.

6.3 Transforming competent cells

About 2 – 5 µL of plasmid from a ligation reaction such as Gibson Assembly or freshly extracted from a cell line is added to 100 µL of competent cells and incubated on ice for 30 minutes. Cells are subjected to heat shock at 42 °C for 30 seconds and put back on ice for two minutes. After two minutes, 200 µL of Super Optimal broth with Catabolite repression (SOC) media is added to the cell and cells are incubated with shaking for one hour at 37 °C. incubated cells are subsequently plated on LB agar plates containing the requisite antibiotics and left overnight at 37 °C.

6.4 Distance plots of MCR from *Methanothermobacter marburgensis*

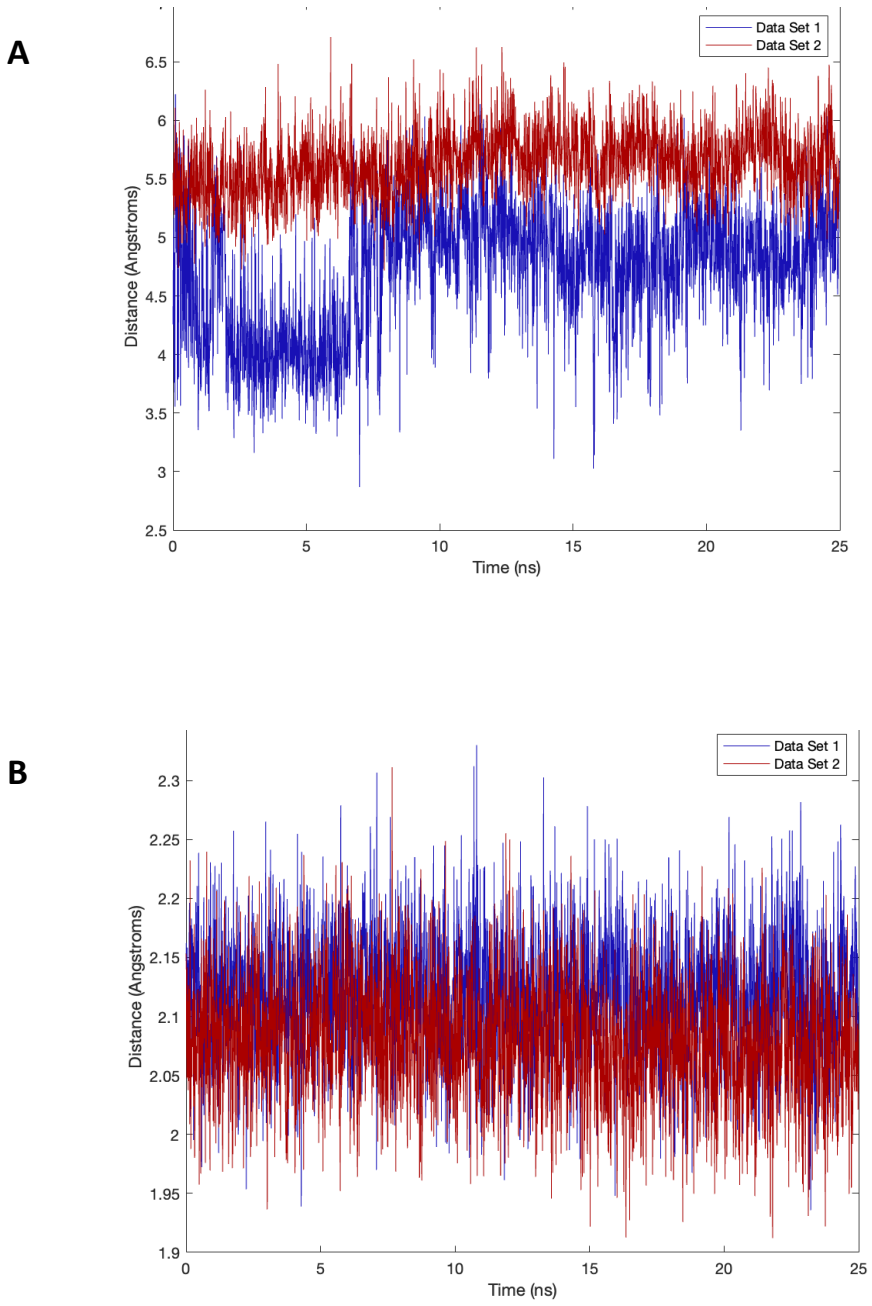
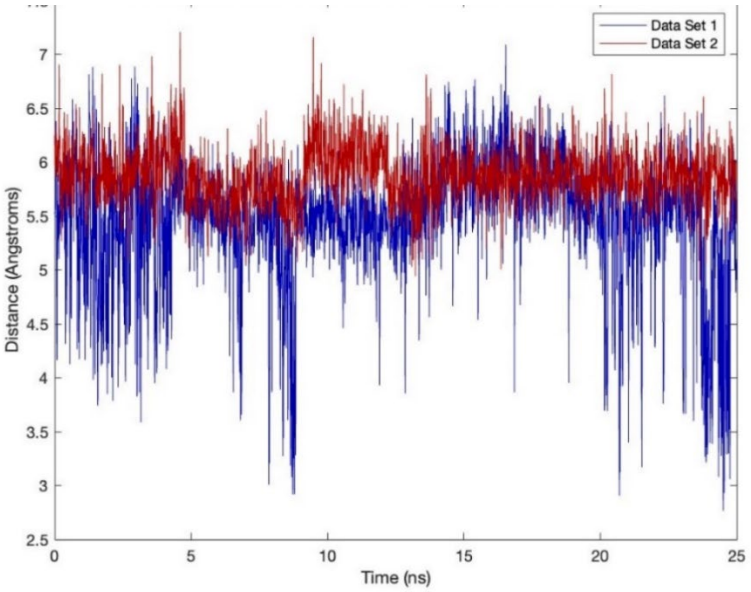


Figure 72: Distance plots of post MD WT MCR from Methanothermobacter marburgensis showing (A) Gln 147 distance to Ni and (B) tetrapyrrole nitrogen distance to Ni. Data sets 1 and 2 refer to MCR active sites 1 and 2

A



B

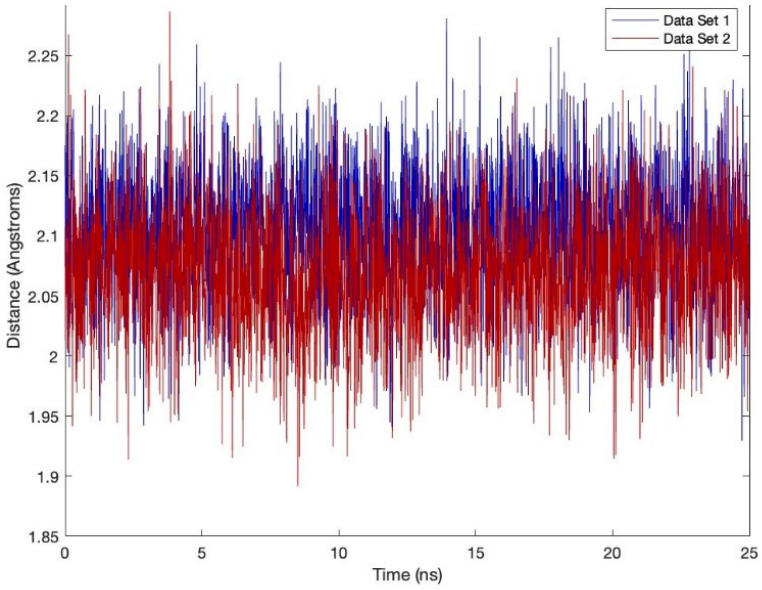


Figure 73: Distance plots of post MD MCR from Methanothermobacter marburgensis without PTMs showing (A) Gln 147 distance to Ni and (B) tetrapyrrole nitrogen distance to Ni. Data sets 1 and 2 refer to MCR active sites 1 and 2

7 References

1. Shindell D, Bréon F m, Collins W, et al. *Anthropogenic and Natural Radiative Forcing. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I*.; 2013.

2. Pinto JP, Brühl CH, Thompson AM. The Current and Future Environmental Role of Atmospheric Methane: Model Studies and Uncertainties. In: *Atmospheric Methane: Sources, Sinks, and Role in Global Change*. Springer Berlin Heidelberg; 1993:514-531. doi:10.1007/978-3-642-84605-2_21

3. Hook SE, Wright ADG, McBride BW. Methanogens: Methane producers of the rumen and mitigation strategies. *Archaea*. 2010;2010. doi:10.1155/2010/945785

4. Pachauri RK, Meyer LA. *Summary for Policymakers*. Vol 9781107025.; 2014. doi:10.1017/CBO9781139177245.003

5. Thauer RK. Biochemistry of methanogenesis : a tribute to Marjory Stephenson. *Microbiology (N Y)*. Published online 1998:2377-2406.

6. Stephenson M, Strickland LH. THE BACTERIAL FORMATION OF METHANE BY THE REDUCTION OF ONE-CARBON COMPOUNDS BY MOLECULAR HYDROGEN. *Biochem J*. 1933;27:1517-1527. doi:10.1042/bj0271517

7. Springer E, Sachs MS, Woese CR, Boone DR. Partial gene sequences for the A subunit of methyl-coenzyme M reductase (mcrI) as a phylogenetic tool for the family Methanosarcinaceae. *Int J Syst Bacteriol*. 1995;45(3):554-559. doi:10.1099/00207713-45-3-554

8. Bobik TA, Olson KD, Noll KM, Wolfe RS. Evidence that the heterodisulfide of coenzyme M and 7-mercaptoheptanoylthreonine phosphate is a product of the methylreductase reaction in Methanobacterium. *Biochem Biophys Res Commun*. 1987;149(2). doi:10.1016/0006-291X(87)90389-5

9. Ellermann J, Hedderich R, Böcher R, Thauer RK. The final step in methane formation. *Eur J Biochem*. 1988;172:669-677. doi:10.1111/j.1432-1033.1988.tb13941.x

10. Mansoorabadi SO, Zheng K, Ngo PD. F430 Biosynthesis and Insertion. *Encyclopedia of Inorganic and Bioinorganic Chemistry*. Published online 2017:1-12. doi:10.1002/9781119951438.eibc2488

11. Enzmann F, Mayer F, Rother M, Holtmann D. Methanogens: biochemical background and biotechnological applications. *AMB Express*. 2018;8(1). doi:10.1186/s13568-017-0531-x

12. Welte C, Deppenmeier U. Bioenergetics and anaerobic respiratory chains of aceticlastic methanogens. *Biochim Biophys Acta Bioenerg*. 2014;1837(7):1130-1147. doi:10.1016/j.bbabio.2013.12.002

13. Thauer RK, Kaster AK, Seedorf H, Buckel W, Hedderich R. Methanogenic archaea: Ecologically relevant differences in energy conservation. *Nat Rev Microbiol*. 2008;6(8):579-591. doi:10.1038/nrmicro1931

14. Krupenye C, Kano F, Hirata S, Call J, Tomasello M. Great apes anticipate that other individuals will act according to false beliefs. *Science (1979)*. 2016;354(6308):110-114. doi:10.1126/science.aaf8110

15. Burke SA, Krzycki JA. *Reconstitution of Monomethylamine:Coenzyme M Methyl Transfer with a Corrinoid Protein and Two Methyltransferases Purified from Methanosarcina Barkeri**.; 1997. http://www.jbc.org

16. Galagan JE, Nusbaum C, Roy A, et al. The genome of M. acetivorans reveals extensive metabolic and physiological diversity. *Genome Res*. 2002;12(4):532-542. doi:10.1101/gr.223902

17. Thauer RK, Hedderich R, Fischer R. Reactions and Enzymes Involved in Methanogenesis from CO2 and H2. *Methanogenesis*. 1993;2:209-252. doi:10.1007/978-1-4615-2391-8_5

18. Dridi B, Fardeau ML, Ollivier B, Raoult D, Drancourt M. Methanomassiliicoccus luminyensis gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. *Int J Syst Evol Microbiol*. 2012;62(8):1902-1907. doi:10.1099/ijs.0.033712-0

19. Iino T, Tamaki H, Tamazawa S, et al. Candidatus methanogranum caenicola: A novel methanogen from the anaerobic digested sludge, and proposal of Methanomassiliicoccaceae fam. nov. and Methanomassiliicoccales ord. nov., for a methanogenic lineage of the class Thermoplasmata. *Microbes Environ*. 2013;28(2):244-250. doi:10.1264/jsme2.ME12189

20. Garcia JL, Patel BKC, Ollivier B. Taxonomic, phylogenetic, and ecological diversity of methanogenic Archaea. *Anaerobe*. 2000;6(4):205-226. doi:10.1006/anae.2000.0345

21. Bräuer SL, Cadillo-Quiroz H, Yashiro E, Yavitt JB, Zinder SH. Isolation of a novel acidiphilic methanogen from an acidic peat bog. *Nature*. 2006;442(7099):192-194. doi:10.1038/nature04810

22. Mathran1 IM, Boone,'t DR, Mah RA, Fox GE, Lau2 PP. *Methanohalophilus Zhilinae Sp. Nov. , an Alkaliphilic, Halophilic, Methylotrophic Methanogen*.; 1988.

23. Wagner T, Kahnt J, Ermler U, Shima S. Didehydroaspartate Modification in Methyl-Coenzyme M Reductase Catalyzing Methane Formation. *Angewandte Chemie - International Edition*. 2016;55(36):10630-10633. doi:10.1002/anie.201603882

24. Conway De Macario E, Maeder DL, Macario AJL. Breaking the mould: Archaea with all four chaperoning systems. *Biochem Biophys Res Commun*. 2003;301(4):811-812. doi:10.1016/S0006-291X(03)00047-0

25. Sowers KR, Baron SF, Ferry JG. Methanosarcina acetivorans sp. nov., an Acetotrophic MethaneProducing Bacterium Isolated from Marine Sediments. *Microbiology (N Y)*. 1984;47(5):971-978.

26. Tashiro Y, Desai SH, Atsumi S. Two-dimensional isobutyl acetate production pathways to improve carbon yield. *Nat Commun*. 2015;6. doi:10.1038/ncomms8488

27. Jasso-Chávez R, Santiago-Martínez MG, Lira-Silva E, et al. Air-adapted Methanosarcina acetivorans shows high methane production and develops resistance against oxygen stress. *PLoS One*. 2015;10(2). doi:10.1371/journal.pone.0117331

28. Ellefson WL, Wolfe RS. Component C of the methylreductase system of Methanobacterium. *Journal of Biological Chemistry*. 1981;256(9):4259-4262.

29. Diekert G, Jaenchen R, Thauer RK. Biosynthetic evidence for a nickel tetrapyrrole structure of factor F430 from Methanobacterium thermoautotrophicum. *FEBS Lett*. 1980;119(1):118-120.

30. Ermler U, Grabarse W, Shima S, Goubeaud M, Thauer RK. Crystal structure of methyl-coenzyme M reductase: The key enzyme of biological methane formation. *Science (1979)*. 1997;278(5342):1457-1462. doi:10.1126/science.278.5342.1457

31. Grabarse W, Mahlert F, Shima S, Thauer RK, Ermler U. Comparison of three methyl-coenzyme M reductases from phylogenetically distant organisms: Unusual amino acid modification, conservation and adaptation. *J Mol Biol*. 2000;303(2):329-344. doi:10.1006/jmbi.2000.4136

32. Kahnt J, Buchenau B, Mahlert F, Krüger M, Shima S, Thauer RK. Post-translational modifications in the active site region of methyl-coenzyme M reductase from methanogenic and methanotrophic archaea. *FEBS Journal*. 2007;274(18):4913-4921. doi:10.1111/j.1742-4658.2007.06016.x

33. Radle MI, Miller D V., Laremore TN, Booker SJ. Methanogenesis marker protein 10 (Mmp10) from methanosarcina acetivorans is a radical S-adenosylmethionine methylase that unexpectedly requires cobalamin. *Journal of Biological Chemistry*. 2019;294(31):11712-11725. doi:10.1074/jbc.RA119.007609

34. Prakash D, Wu Y, Suh SJ, Duin EC. Elucidating the process of activation of methyl-coenzyme M reductase. *J Bacteriol*. 2014;196(13):2491-2498. doi:10.1128/JB.01658-14

35. Ankel-Fuchs D, Hüster R, Mörschel E, Albracht SPJ, Thauer RK. Structure and function of methyl-coenzyme M reductase and of factor F430 in methanogenic bacteria. *Syst Appl Microbiol*. 1986;7(2-3):383-387. doi:10.1016/S0723-2020(86)80038-8

36. Hinrichs KU, Boetius A. The anaerobic oxidation of methane: new insights in microbial ecology and biogeochemistry. In: Wefer G, Billet D, Hebbeln D, Jorgensen BB, Schluter M, Weering T van, eds. *Ocean Margin Systems.* Springer-Verlag; 2002:457-477.

37. Raghoebarsing AA, Pol A, Van De Pas-Schoonen KT, et al. A microbial consortium couples anaerobic methane oxidation to denitrification. *Nature*. 2006;440(7086):918-921. doi:10.1038/nature04617

38. Knittel K, Boetius A. Anaerobic Oxidation of Methane: Progress with an Unknown Process. *Annu Rev Microbiol*. 2009;63(1). doi:10.1146/annurev.micro.61.080706.093130

39. Hallam SJ, Putnam N, Preston CM, et al. Reverse methanogenesis: Testing the hypothesis with environmental genomics. *Science (1979)*. 2004;305(5689):1457-1462. doi:10.1126/science.1100025

40. Beal EJ, House CH, Orphan VJ. Manganese- and iron-dependent marine methane oxidation. *Science (1979)*. 2009;325(5937):184-187. doi:10.1126/science.1169984

41. Haroon MF, Hu S, Shi Y, et al. Anaerobic oxidation of methane coupled to nitrate reduction in a novel archaeal lineage. *Nature*. 2013;500(7464):567-570. doi:10.1038/nature12375

42. Reeburgh WS. *Microbial Growth on C1 Compounds*. (Lidstrom ME, Tabita F. R., eds.).; 1996. doi:10.1007/978-94-009-0213-8

43. Hinrichs KU, Hayes JM, Sylva SP, Brewer PG, DeLong EF. Methane-consuming archaebacteria in marine sediments. *Nature*. 1999;398:802-805.

44. Schreiber L, Holler T, Knittel K, Meyerdierks A, Amann R. Identification of the dominant sulfate-reducing bacterial partner of anaerobic methanotrophs of the ANME-2 clade. *Environ Microbiol*. 2010;12(8):2327-2340. doi:10.1111/j.1462-2920.2010.02275.x

45. Boetius A, Ravenschlag K, Schubert CJ, et al. *A Marine Microbial Consortium Apparently Mediating Anaerobic Oxidation of Methane*. Vol 407.; 2000. www.nature.com

46. Fairbrother WG, Yeh RF, Sharp PA, Burge CB. Predictive identification of exonic splicing enhancers in human genes. *Science (1979)*. 2002;297(5583):1007-1013. doi:10.1126/science.1073774

47. Orphan VJ, House CH, Hinrichs KU, McKeegan KD, DeLong EF. *Methane-Consuming Archaea Revealed by Directly Coupled Isotopic and Phylogenetic Analysis*. Vol 91.; 1997. www.sciencemag.org/cgi/content/full/293/5529/482/DC1.

48. Knittel K, Lösekann T, Boetius A, Kort R, Amann R. Diversity and distribution of methanotrophic archaea at cold seeps. *Appl Environ Microbiol*. 2005;71(1):467-479. doi:10.1128/AEM.71.1.467-479.2005

49. Reeburgh WS. Oceanic methane biogeochemistry. *Chem Rev*. 2007;107(2):486-513. doi:10.1021/cr050362v

50. Selmer T, Kahnt J, Goubeaud M, et al. The biosynthesis of methylated amino acids in the active site region of methyl-coenzyme M reductase. *Journal of Biological Chemistry*. 2000;275(6):3755-3760. doi:10.1074/jbc.275.6.3755

51. Krüger M, Meyerdierks A, Glöckner FO, et al. A conspicuous nickel protein in microbial mats that oxidize methane anaerobically. *Nature*. 2003;426(6968):878-881. doi:10.1038/nature02207

52. Shima S, Krueger M, Weinert T, et al. Structure of a methyl-coenzyme M reductase from Black Sea mats that oxidize methane anaerobically. *Nature*. 2012;481(7379):98-101. doi:10.1038/nature10663

53. Nayak DD, Mahanta N, Mitchell DA, Metcalf WW. Post-translational thioamidation of methyl-coenzyme M reductase, a key enzyme in methanogenic and methanotrophic archaea. *Elife*. 2017;6(I):1-18. doi:10.7554/eLife.29218

54. Mahanta N, Liu A, Dong S, Nair SK, Mitchell DA. Enzymatic reconstitution of ribosomal peptide backbone thioamidation. *Proc Natl Acad Sci U S A*. 2018;115(12):3030-3035. doi:10.1073/pnas.1722324115

55. Lyu Z, Chou X wen, Shi H, Patel R, Duin EC, Whitman WB. Mmp10 is required for PTM arginine in active site of MCR. *Microbiology (N Y)*. Published online 2017.

56. Sofia HJ, Chen G, Hetzler BG, Reyes-Spindola JF, Miller NE. *Radical SAM, a Novel Protein Superfamily Linking Unresolved Steps in Familiar Biosynthetic Pathways with Radical Mechanisms: Functional Characterization Using New Analysis and Information Visualization Methods*. Vol 29.; 2001. http://www.ncbi.nlm.nih.gov/BLAST/

57. Pierre S, Guillot A, Benjdia A, Sandstrom C, Langella P, Berteau O. Thiostrepton tryptophan methyltransferase expands the chemistry of radical SAM enzymes. *Nat Chem Biol*. 2012;8:957-959.

58. Frey PA, Booker S. Radical intermediates in the reaction of lysine 2,3-aminomutase. *Advances in free radical chemistry*. 1999;2:1-43.

59. Landgraf BJ, McCarthy EL, Booker SJ. Radical S-Adenosylmethionine Enzymes in Human Health and Disease. *Annu Rev Biochem*. 2016;85:485-514. doi:10.1146/annurev-biochem-060713-035504

60. Zhang Q, Van Der Donk WA, Liu W. Radical-mediated enzymatic methylation: A tale of two SAMS. *Acc Chem Res*. 2012;45(4):555-564. doi:10.1021/ar200202c

61. Bauerle MR, Schwalm EL, Booker SJ. Mechanistic diversity of radical S-adenosylmethionine (SAM)-dependent methylation. *Journal of Biological Chemistry*. 2015;290(7):3995-4002. doi:10.1074/jbc.R114.607044

62. Nayak DD, Liu A, Agrawal N, et al. Functional interactions between posttranslationally modified amino acids of methyl-coenzyme M reductase in Methanosarcina acetivorans. *PLoS Biol*. 2020;18(2):1-23. doi:10.1371/journal.pbio.3000507

63. Wongnate T, Sliwa D, Ginovska B, et al. *The Radical Mechanism of Biological Methane Synthesis by Methyl-Coenzyme M Reductase*.; 2016. https://www.science.org

64. Horng YC, Becker DF, Ragsdale SW. Mechanistic studies of methane biogenesis by methyl-coenzyme M reductase: Evidence that coenzyme B participates in cleaving the C - S bond of methyl-coenzyme M. *Biochemistry*. 2001;40(43):12875-12885. doi:10.1021/bi011196y

65. Ragsdale SW, Raugei S, Ginovska B, Wongnate T. Biochemistry of Methyl-Coenzyme M Reductase. In: Zamble D, Rowinska-Zyrek M, Kozlowski H, eds. *The Biological Chemistry of Nickel*. ; 2017:149-169.

66. Dey M, Kunz RC, Lyons DM, Ragsdale SW. Characterization of alkyl-nickel adducts generated by reaction of methyl-coenzyme M reductase with brominated acids. *Biochemistry*. 2007;46(42):11969-11978. doi:10.1021/bi700925n

67. Dey M, Telser J, Kunz RC, Lees NS, Ragsdale SW, Hoffman BM. Biochemical and spectroscopic studies of the electronic structure and reactivity of a methyl-Ni species formed on methyl-coenzyme M reductase. *J Am Chem Soc*. 2007;129(36):11030-11032. doi:10.1021/ja074556z

68. Pelmenschikov V, Blomberg MRA, Siegbahn PEM, Crabtree RH. A mechanism from quantum chemical studies for methane formation in methanogenesis. *J Am Chem Soc*. 2002;124(15):4039-4049. doi:10.1021/ja011664r

69. Chen SL, Pelmenschikov V, Blomberg MRA, Siegbahn PEM. Is there a ni-methyl intermediate in the mechanism of methyl-coenzyme M reductase? *J Am Chem Soc*. 2009;131(29):9912-9913. doi:10.1021/ja904301f

70. Pelmenschikov V, Siegbahn PEM. Catalysis by methyl-coenzyme M reductase: A theoretical study for heterodisulfide product formation. *Journal of Biological Inorganic Chemistry*. 2003;8(6):653-662. doi:10.1007/s00775-003-0461-8

71. Zheng K, Ngo PD, Owens VL, Yang XP, Mansoorabadi SO. The biosynthetic pathway of coenzyme F430 in methanogenic and methanotrophic archaea. *Science (1979)*. 2016;354(6310). doi:10.1126/science.aag2947

72. Allen KD, Wegener G, White RH. Discovery of multiple modified F430 coenzymes in methanogens and anaerobic methanotrophic archaea suggests possible new roles for F430 in nature. *Appl Environ Microbiol*. 2014;80(20):6403-6412. doi:10.1128/AEM.02202-14

73. Mayr S, Latkoczy C, Krüger M, et al. Structure of an F430 variant from archaea associated with anaerobic oxidation of methane. *J Am Chem Soc*. 2008;130(32):10758-10767. doi:10.1021/ja802929z

74. Hahn CJ, Lemaire ON, Kahnt J, Engilberge S, Wegener G, Wagner T. *Crystal Structure of a Key Enzyme for Anaerobic Ethane Activation*.

75. Scheller S, Goenrich M, Thauer RK, Jaun B. Methyl-coenzyme M reductase from methanogenic archaea: Isotope effects on the formation and anaerobic oxidation of methane. *J Am Chem Soc*. 2013;135(40):14975-14984. doi:10.1021/ja406485z

76. *Methane Emissions in the Oil and Gas Industry Quantifying Emissions and Distinguishing between Different Methane Sources*.; 2018. www.americangeosciences.

77. Shaftel H, Jackson R, Callery S, Bailey D. Global Warming vs. Climate Change | Resources – Climate Change: Vital Signs of the Planet. NASA- California Institute of Technology. Published 2020. Accessed June 2, 2020. https://climate.nasa.gov/resources/global-warming-vs-climate-change/

78. Change global climate. Effects | Facts – Climate Change: Vital Signs of the Planet. *Nasa*. Published online 2020. Accessed June 2, 2020. https://climate.nasa.gov/effects/

79. Ferry JG. Methane : Small Molecule , Big Impact. *Science (1979)*. 1997;278(5342):1413-1414. http://www.jstor.org/stable/2894748

80. Zheng K, Ngo PD, Owens VL, Yang XP, Mansoorabadi SO. The biosynthetic pathway of coenzyme F430 in methanogenic and methanotrophic archaea. *Science (1979)*. 2016;354(6310).

81. Thauer RK. Methyl (Alkyl)-Coenzyme M Reductases: Nickel F-430-Containing Enzymes Involved in Anaerobic Methane Formation and in Anaerobic Oxidation of Methane or of Short Chain Alkanes. *Biochemistry*. 2019;58(52):5198-5220. doi:10.1021/acs.biochem.9b00164

82. Mansoorabadi SO, Zheng K, Ngo PD. F430 Biosynthesis and Insertion. *Encyclopedia of Inorganic and Bioinorganic Chemistry*. Published online 2017:1-12. doi:10.1002/9781119951438.eibc2488

83. Nayak DD, Liu A, Agrawal N, et al. Functional interactions between posttranslationally modified amino acids of methyl-coenzyme M reductase in Methanosarcina acetivorans. *PLoS Biol*. 2020;18(2):1-23. doi:10.1371/journal.pbio.3000507

84. Wagner T, Kahnt J, Ermler U, Shima S. Didehydroaspartate Modification in Methyl-Coenzyme M Reductase Catalyzing Methane Formation. *Angewandte Chemie - International Edition*. 2016;55(36):10630-10633. doi:10.1002/anie.201603882

85. Radle MI, Miller D V., Laremore TN, Booker SJ. Methanogenesis marker protein 10 (Mmp10) from methanosarcina acetivorans is a radical S-adenosylmethionine methylase that unexpectedly requires cobalamin. *Journal of Biological Chemistry*. 2019;294(31):11712-11725. doi:10.1074/jbc.RA119.007609

86. Lyu Z, Chou CW, Shi H, et al. Assembly of Methyl Coenzyme M Reductase in the Methanogenic Archaeon *Methanococcus maripaludis*. Metcalf WW, ed. *J Bacteriol*. 2018;200(7). doi:10.1128/JB.00746-17

87. Nayak DD, Metcalf WW. Genetic techniques for studies of methyl-coenzyme M reductase from Methanosarcina acetivorans C2A. In: *Methods in Enzymology*. Vol 613. ; 2018:325-347. doi:10.1016/bs.mie.2018.10.012

88. Zheng K, Mansoorabadi S. *A Study of Methyl-Coenzyme M Reductase Maturation: Coenzyme F430 Biosynthesis and Post-Translational Modifications*.; 2018.

89. Ngo PD, Mansoorabadi S. *A Story About Light and Energy: Mechanistic and Biosynthetic Investigations of Bioluminescence and Methanogenesis*.; 2016.

90. Du F, Liu YQ, Xu YS, et al. Regulating the T7 RNA polymerase expression in E. coli BL21 (DE3) to provide more host options for recombinant protein production. *Microb Cell Fact*. 2021;20(1). doi:10.1186/s12934-021-01680-6

91. *PET System Manual Novagen 1*.; 1992.

92. Deobald D, Adrian L, Schöne C, Rother M, Layer G. Identification of a unique Radical SAM methyltransferase required for the sp3-C-methylation of an arginine residue of methyl-coenzyme M reductase. *Sci Rep*. 2018;8(1). doi:10.1038/s41598-018-25716-x

93. Hogan KB, Hoffman JS, Thompson AM. *Methane on the Greenhouse Agenda*.; 1991.

94. Muller S, Hoege C, Pyrowolakis G, Jentsch S. Sumo, ubiquitin's mystery cousin. *Nat Rev Mol Cell Biol*. 2001;2:202-210.

95. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-589. doi:10.1038/s41586-021-03819-2

96. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nat Methods*. 2022;19(6):679-682. doi:10.1038/s41592-022-01488-1

97. Schrodinger L. The PyMOL Molecular Graphics System, Version 1.3r1. Published online 2010.

98. Ihsanawati, Nishimoto M, Higashijima K, et al. Crystal Structure of tRNA N2,N2-Guanosine Dimethyltransferase Trm1 from Pyrococcus horikoshii. *J Mol Biol*. 2008;383:871-884.

99. Jansson A, Niemi J, Lindqvist Y, Mäntsälä P, Schneider G. Crystal structure of aclacinomycin-10-hydroxylase, a S-adenosyl-L- methionine-dependent methyltransferase homolog involved in anthracycline biosynthesis in Streptomyces purpurascens. *J Mol Biol*. 2003;334(2):269-280. doi:10.1016/j.jmb.2003.09.061

100. Vryer R, Saffery R. Metabolic Regulation of DNA Methylation in Mammalian Cells. In: *Handbook of Epigenetics*. Elsevier; 2017:293-305. doi:10.1016/B978-0-12-805388-1.00019-5

101. Holm L. Dali server: structural unification of protein families. *Nucleic Acids Res*. 2022;50:W210-W215.

102. Aksnes H, Ree R, Arnesen T. Co-translational, Post-translational, and Non-catalytic Roles of N-Terminal Acetyltransferases. *Mol Cell*. 2019;73(6):1097-1114. doi:10.1016/j.molcel.2019.02.007

103. Arnesen T, Anderson D, Baldersheim C, Lanotte M, Varhaug JE, Lillehaug JR. Identification and characterization of the human ARD1-NATH protein acetyltransferase complex. *Biochemistry*. 2005;386:433-443.

104. Gautschi M, Just S, Mun A, et al. The yeast N(alpha)-acetyltransferase NatA is quantitatively anchored to the ribosome and interacts with nascent polypeptides. *Mol Cell Biol*. 2003;23:7403-7414.

105. Schlicke H, Richter A, Rothbart M, Brzezowski P, Hedtke B, Grimm B. Function of Tetrapyrroles, Regulation of Tetrapyrrole Metabolism and Methods for Analyses of Tetrapyrroles. *Procedia Chem*. 2015;14:171-175. doi:10.1016/j.proche.2015.03.025

106. Warren MJ, Smith AG. *Tetrapyrroles : Birth, Life, and Death*. Austin, Tex. : Landes Bioscience ; New York, N.Y. : Springer Science & Business Media; 2009.

107. Scheller S, Goenrich M, Thauer RK, Jaun B. Methyl-coenzyme M reductase from methanogenic archaea: Isotope effects on the formation and anaerobic oxidation of methane. *J Am Chem Soc*. 2013;135(40):14975-14984. doi:10.1021/ja406485z

108. Wagner T, Wegner CE, Kahnt J, Ermler U, Shima S. Phylogenetic and structural comparisons of the three types of methyl coenzyme M reductase from Methanococcales. *J Bacteriol*. 2017;199(16):1-15. doi:10.1128/JB.00197-17

109. Brenner MC, Zhang H, Scott RA. Nature of the low activity of S-methyl-... determined by active site titrations - Brenner. *Journal of Biological Chemistry*. 1993;268(25):18491-18495.

110. Karplus M, McCammon A. Molecular Dynamics Simulations of Biomolecules. *Nature structural biology* . 2002;9(9):646-652. doi:10.1038/nsb0902-646

111. Mccammon JA, Gel BR, Karp Ius M. *Dynamics of Folded Proteins*. Vol 267. Nature Publishing Group; 1977.

112. Hollingsworth SA, Dror RO. Molecular Dynamics Simulation for All. *Neuron*. 2018;99(6):1129-1143. doi:10.1016/j.neuron.2018.08.011

113. Donnan PH, Mansoorabadi SO, Harris JM. *Computational Studies of Dinoflagellate Luciferase and Radical S-Adenosyl-L-Methionine Enzymes*.; 2022.

114. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold-Making protein folding accessible to all Predicting the three-dimensional structure of a protein from. doi:10.1101/2021.08.15.456425

115. Sahrmann PG, Donnan PH, Merz KM, Mansoorabadi SO, Goodwin DC. MRP.py: A parametrizer of post-translationally modified residues. *J Chem Inf Model*. 2020;60(10):4424-4428. doi:10.1021/acs.jcim.0c00472

116. Jakalian A, Jack DB, Bayly CI. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem*. 2002;23(16):1623-1641. doi:10.1002/jcc.10128

117. Wang J, Wang W, Kollman PA, Case DA. Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model*. 2006;25(2):247-260. doi:10.1016/j.jmgm.2005.12.005

118. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general Amber force field. *J Comput Chem*. 2004;25(9):1157-1174. doi:10.1002/jcc.20035

119. Case Ross C Walker UC San Diego DA, Thomas Cheatham AE. *Amber 2022 Reference Manual Principal Contributors to the Current Codes*. https://ambermd.org/contributors.html

120. Li P, Merz KM. MCPB.py: A Python Based Metal Center Parameter Builder. *J Chem Inf Model*. 2016;56(4):599-604. doi:10.1021/acs.jcim.5b00674

121. Ponder JW, Caset D. *FORCE FIELDS FOR PROTEIN SIMULATIONS*.; 2003.

122. Loncharich RJ, Brooks BR, Pastor RW. Langevin dynamics of peptides: The frictional dependence of isomerization rates of N-acetylalanyl-N'-methylamide. *Biopolymers*. 1992;32(5):523-535. doi:10.1002/bip.360320508

123. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput*. 2013;9(7):3084-3095. doi:10.1021/ct400341p

124. Wongnate T, Ragsdale SW. The reaction mechanism of methyl-coenzyme M reductase. *Journal of Biological Chemistry*. 2015;290(15):9322-9334. doi:10.1074/jbc.M115.636761

125. Duin EC, McKee ML. A New Mechanism for Methane Production from Methyl-Coenzyme M Reductase As Derived from Density Functional Calculations. *J Phys Chem B*. 2008;112(8):2466-2482. doi:10.1021/jp709860c

126. Lyu Z, Shao N, Chou CW, et al. Posttranslational Methylation of Arginine in Methyl Coenzyme M Reductase Has a Profound Impact on both Methanogenesis and Growth of *Methanococcus maripaludis*. Becker A, ed. *J Bacteriol*. 2019;202(3). doi:10.1128/JB.00654-19