Exploring the nucleotide promiscuity of *Escherichia coli* methionine adenosyltransferase mutants

A thesis submitted to
Indian Institute of Science Education and Research Pune
In partial fulfillment of the requirements for the
BS-MS Dual Degree Programme

by

Aditya Bhattacharyya

Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

April, 2023

Supervisor: Dr. Amrita B. Hazra

Certificate

This is to certify that this dissertation entitled "Exploring the nucleotide promiscuity of *Escherichia coli* methionine adenosyltransferase mutants" towards the partial fulfillment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Aditya Bhattacharyya at IISER Pune under the supervision of Dr. Amrita B. Hazra, Associate Professor, Department of Chemistry, IISER Pune, during the academic year 2022-2023.

Aditya Bhattacharyya

(IISER – Pune)

(20181165)

Dr. Amrita B. Hazra

(IISER – Pune)

(Supervisor)

Declaration

I hereby declare that the matter embodied in the report entitled "Exploring the nucleotide promiscuity of *Escherichia coli* methionine adenosyltransferase mutants" are the results of the work carried out by me at the Department of Chemistry, Indian Institute of Science Education and Research, Pune, under the supervision of Dr. Amrita B. Hazra and the same has not been submitted elsewhere for any other degree.

Aditya Bhattacharyya

(IISER – Pune)

(20181165)

Dr. Amrita B. Hazra

(IISER – Pune)

(Supervisor)

This thesis is dedicated to Dada, my biggest inspiration for pursuing scientific research

Acknowledgments

It gives me immense pleasure to thank those who aided in the completion of this thesis.

Firstly, I would like to thank my supervisor, Dr. Amrita B. Hazra, for her active mentorship throughout the past fifteen months I have worked in her lab. Especially, I owe her a debt of gratitude for backing me, an undergrad, to assume responsibility and lead a research project by myself. Working on the MAT project has been a memorable experience that has pivotally shifted my research interests and outlooks. I also thank Dr. Gayathri Pananghat for being on my thesis advisory committee and providing interesting suggestions throughout the thesis.

This thesis could not have been completed without the continuous support of my fellow lab members. I would like to thank Vruta for being an extraordinarily supportive mentor and friend and for patiently teaching me various molecular biology and enzymology techniques when I first joined the lab; Yamini for her invaluable insights on the MAT project and for being an inspiring role model over the last one year; Anjali for her assistance in some crucial experiments and engaging questions regarding MAT literature; Aniket for fun discussions and assistance during HPLC troubleshooting; Rohan and Ateek for helpful suggestions on designing nucleotide-promiscuous mutants; and Shweta, Neel, Rajdip, Rupali, Vibhor, Riya and Nikita for providing a vibrant lab environment.

I thank Prof. Robert Blumenthal from the University of Toledo for kindly providing the *E. coli* MOB1490 strain and Prof. Squire J. Booker from Pennsylvania State University for gifting the strain that overexpresses *Mj*MAT.

I also would like to thank the SFB autoclave facility, the LC-MS facility, and the Departments of Chemistry and Biology at IISER Pune.

Lastly, my friends at IISER Pune, especially Mihir, Kushan, Pavitra, and Bishal, and my family – mother, father, and grandmother have been an inexhaustible source of moral support and motivation. I cannot thank them enough for their positive impact over the last five years.

Contents

List of figures and tables

List of figures:

List of tables:

List of abbreviations

ATP	Adenosine-5’-triphosphate
AUC	Area under the curve
β-ME	β-mercaptoethanol
CTP	Cytidine-5’-triphosphate
***Ch*MAT**	*Cryptosporidium hominis* methionine adenosyltransferase
DNA	Deoxyribonucleic acid
***Ec*MAT**	*Escherichia coli* methionine adenosyltransferase
***Ec*GR**	*Escherichia coli* glutathione reductase
EIC	Extracted ion chromatogram
ESI	Electrospray ionization
FAD	Flavin adenine dinucleotide
FMN	Flavin mononucleotide
GTP	Guanosine-5’-triphosphate
HPLC	High-performance liquid chromatography
***Hs*MAT2a**	*Homo sapiens* methionine adenosyltransferase 2A isoform
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITP	Inosine-5’-triphosphate
LB	Luria-Bertani
LC-MS	Liquid chromatography – Mass spectrometry
MAT	Methionine adenosyltransferase
***Mj*FMNAT**	*Methanocaldococcus jannaschii* Flavin mononucleotide adenylyltransferase
***Mj*MAT**	*Methanocaldococcus jannaschii* MAT
MS	Mass spectrum
MTA	5’-methylthioadenosine
MTase	Methyltransferase
MTC	5’-methylthiocytidine
MTG	5’-methylthioguanosine
MTN	5’-methylthionucleoside
MTU	5’-methylthiouridine
NAD	Nicotinamide adenine dinucleotide

NCD	Nicotinamide cytosine dinucleotide
NTP	Nucleoside-5’-triphosphate
OD	Optical density
ONB-SAM	O-nitrobenzyme S-adenosyl methionine
PCR	Polymerase chain reaction
PDB	Protein data bank
P_i	Orthophosphate
PMSF	Phenylmethysulfonyl fluoride
PP_i	Pyrophosphate
PPP_i	Triphosphate
Q-TOF	Quadrupole-time of flight
SAH	S-adenosyl-L-homocysteine
SCH	S-cytidyl-L-homocysteine
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SGH	S-guanosyl-L-homocysteine
SNH	S-nucleosidyl-L-homocysteine
SUH	S-uridyl-L-homocysteine
SAM	S-adenosyl-L-methionine
SCM	S-cytidyl-L-methionine
SGM	S-guanosyl-L-methionine
SNM	S-nucleosidyl-L-methionine
SUM	S-uridyl-L-methionine
***Tk*MAT**	*Thermococcus kodakarensis* MAT
UTP	Uridine-5’-triphosphate
WT	Wild-type

Abstract

The cofactor S-adenosyl methionine (SAM) is a critical player in cellular metabolism, involved in processes like methylating proteins and nucleic acids and biosynthesizing polyamines and specific vitamins. Its role as a versatile cofactor makes the *in vivo* generation of its analogs highly desirable, especially for designing biorthogonal metabolic pathways inside the cell.

SAM is synthesized in the cell from adenosine–5'-triphosphate (ATP) and L-methionine by the enzyme methionine adenosyltransferase (MAT), also known as SAM synthetase. MAT is a cytosolic enzyme with access to additional non-cognate nucleoside-5'-triphosphates (NTPs), further motivating investigations into expanding the nucleotide substrate scope of the enzyme to synthesize nucleobase analogs of SAM (SNMs).

In this study, we have attempted to enhance the nucleotide promiscuity of the MAT from *Escherichia coli* (*Ec*MAT). We implemented a rational design approach where *Ec*MAT was mutated based on differences between its active site and that of an NTP-promiscuous homolog of MAT from *Methanocaldococcus jannaschii.* We observe that all the seven *Ec*MAT variants we characterized formed SAM *in vitro,* but only one – I102L *Ec*MAT exhibited robust *in vitro* gains in activity with guanosine-5'-triphosphate (GTP) and uridine-5'-triphosphate (UTP). All the mutants retained SAM synthetase activity *in vivo* when heterologously expressed inside *E. coli.* The I102L mutant, when complemented with an *E. coli metK (*gene encoding for MAT) knockout strain, also catalyzed the formation of detectable amounts of S-guanosyl methionine (SGM) inside the organism. These results provide insights for designing novel mutagenesis-based strategies to improve *Ec*MAT's nucleotide promiscuity. The study also provides the first evidence for the formation of SNMs inside a microbial cell.

1. Introduction

1.1 Promiscuous behavior of enzymes

Enzymes are proteins that catalyze chemical reactions between biomolecules inside the cell. They are generally considered extremely specific regarding their catalytic activity and substrate pool, so much so that elementary biochemistry textbooks often compare enzyme-substrate binding to a lock fitted with its key. However, several instances of enzymes deviating from this notion of specificity have been reported in the literature. These enzymes are termed 'promiscuous.' [1]

Enzyme promiscuity can be broadly classified into two categories: a) catalytic and b) substrate promiscuity.[2] The former case is when an enzyme can catalyze additional reactions different from its biologically relevant one. For example, one of the earliest reports of catalytic promiscuity was the enzyme L-asparaginase. Its primary role involves the hydrolysis of the amide bond in L-asparagine to form L-aspartate. However, it can also convert β-cyanoalanine into aspartate and ammonia, thereby acting as a nitrilase[3] (Figure 1A). On the other hand, substrate-promiscuous enzymes catalyze the same chemical transformation for multiple substrates. Lipases, which hydrolyze the glyceryl esters of several long-chain fatty acids, belong to this category[1] (Figure 1B). Lastly, some enzymes may belong to both categories. For example, the serine protease chymotrypsin exhibits substrate promiscuity by hydrolyzing amide linkages in a broad range of peptides. It is also catalytic promiscuous by showing secondary activity as a phosphotriesterase[4] (Figure 1C).

Enzyme promiscuity has generated much interest in the biocatalysis research community.[5] In particular, substrate-promiscuous enzymes have been extensively studied because they are active with a vast pool of natural but also synthetic molecules. Consequently, they can be treated purely as a catalyst to manufacture chemicals of interest.[6] However, these enzymes' natural ability to accept synthetic substrates is generally not high enough for practical purposes, in which case they have been modified to enhance promiscuity for the desired substrates.[7] Bacterial strains expressing these promiscuous enzymes have also been engineered for *in vivo* production of significant compounds.[8]

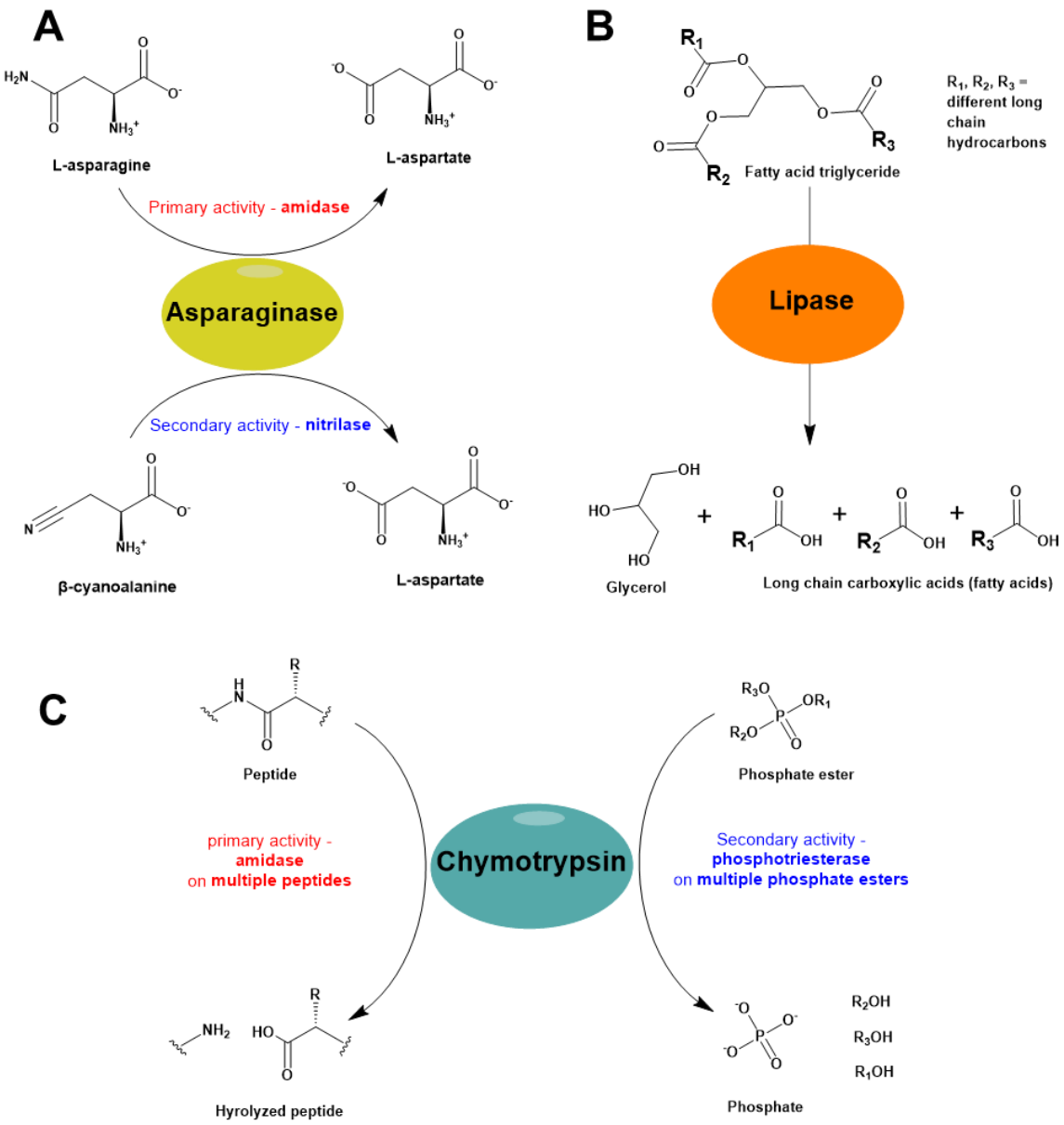


Figure 1: The different types of enzyme promiscuity. (A) Asparaginase exhibits catalytic promiscuity by being capable of acting both as an amidase (its primary activity) and as a nitrilase (secondary activity). **(B)** Lipases are substrate promiscuous as they can catalyze reactions with several fatty acid triglycerides. **(C)** Chymotrypsin is both catalytic and substrate promiscuous as it can exhibit both amidase (primary role) and phosphotriesterase (secondary role) activity on a large pool of peptides and phosphate esters, respectively.

1.2 Substrate promiscuity of cofactor-biosynthesis enzymes

Cofactors are small molecules or ions that aid enzymes in catalyzing their reactions. Some examples of small molecule cofactors include nicotinamide adenine dinucleotide (NAD) and flavin adenine dinucleotide (FAD), which perform hydride

transfers to facilitate enzymatic oxidation-reduction reactions.[9] Methylcobalamin (the bioactive form of Vitamin B_{12}) is another small molecule cofactor that assists the enzyme methionine synthase in methylating homocysteine to biosynthesize methionine.[10] Ionic cofactors include metal ions like Mg^{2+} and K^{+} that facilitate the phosphorylation reaction catalyzed by kinases by stabilizing the negatively charged phosphate group.[11]

Generally, a particular cofactor may assist enzymes from diverse biochemical pathways. Thus, if utilizable analogs of these cofactors can be produced *in vivo*, then biorthogonal variants of several metabolic pathways (which utilize only the cofactor analog) could be subsequently designed. However, synthesizing cofactor analogs *in vivo* would require the cofactor biosynthesis enzymes to demonstrate substrate promiscuity. Consequently, there is a motivation to investigate and enhance the substrate promiscuity of different cofactor biosynthesis enzymes.

A handful of research groups have extensively worked on this topic to biosynthesize analogs of various cofactors. Most notably, the Zhao group from the Dalian Institute of Chemical Physics modified NAD biosynthesis enzymes to generate nicotinamide cytosine dinucleotide (NCD) (Figure 2A), then engineered several NAD-dependent oxidoreductases to make them NCD-compatible and finally expressed the modified enzymes inside *Escherichia coli* resulting in a strain that can bio-orthogonally produce and utilize NCD.[12]Similarly, our group achieved the enzymatic synthesis of FAD nucleobase analogs by utilizing the substrate-promiscuous flavin mononucleotide adenylyl transferase from *Methanocaldococcus jannaschii* (*Mj*FMNAT) (Figure 2B), demonstrated the utilization of these analogs by *E. coli* glutathione reductase (*Ec*GR), and then established the production of the analogs inside *Mj*FMNAT-expressing *E. coli*.[13] Our group has also investigated the substrate scope of the B_{12} biosynthesis enzyme CobT.[14] This enzyme activates 5,6-dimethylbenzimidazole (the lower ligand of B_{12}) for attachment to the rest of the B_{12} molecule. However, it can also activate other benzimidazoles, purines, and phenolic derivatives resulting in B_{12} analogs (called cobamides) with different lower ligands (Figure 2C).

Thus, substrate-promiscuous enzymes involved in NAD, FAD, and B_{12} biosynthesis can be used to generate their analogs. This approach could subsequently be extended to create analogs of other cofactors.

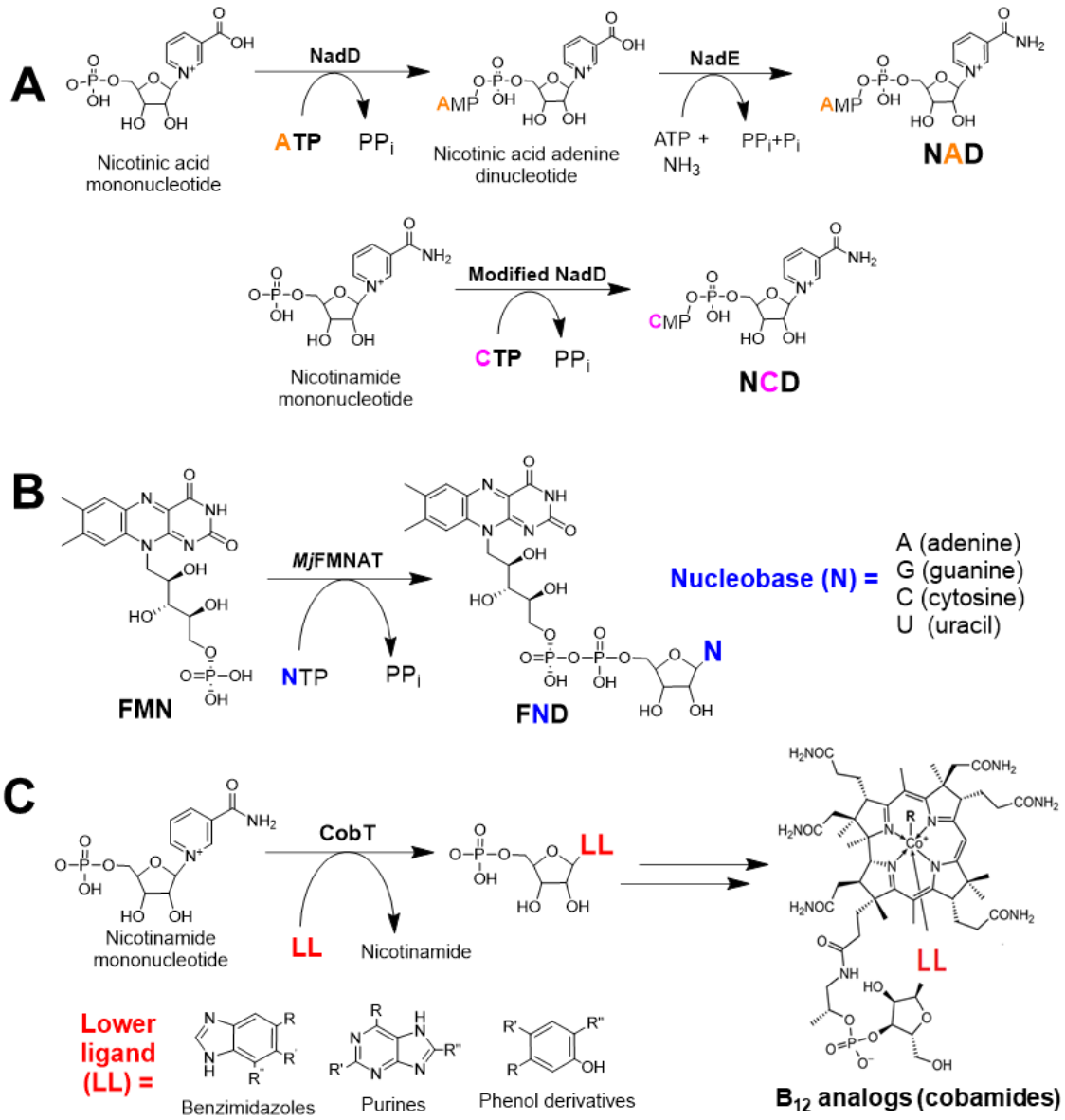


Figure 2: Substrate promiscuity of cofactor biosynthesis enzymes can be used to produce cofactor analogs. (A) The conversion of NAD from nicotinic acid mononucleotide occurs via two enzyme-catalyzed steps – NadD converts nicotinic acid mononucleotide to nicotinic acid adenine dinucleotide, which then undergoes amidation in the presence of NadE to form NAD. The Zhao group has engineered NadD to be promiscuous to nicotinamide mononucleotide and CTP leading to the one-step formation of NCD. **(B)** FMNAT catalyzes the transfer of the AMP group from ATP to flavin mononucleotide (FMN), forming FAD. *Mj*FMNAT can catalyze this reaction with other NTPs creating nucleobase analogs of FAD (FNDs). **(C)** CobT activates the lower ligand (LL) of B_{12} - 5,6-dimethylbenzimidazole (not shown here) for attachment to the rest of the molecule by transferring a phosphoribosyl group from nicotinamide mononucleotide. The ability of certain CobT homologs to accept other benzimidazoles, purines, and phenolic derivatives leads to the generation of B_{12} analogs with different lower ligands called cobamides.

1.3 S-adenosyl methionine (SAM) is a versatile cofactor

The cofactor S-adenosyl-L-methionine (SAM) is one of the most fascinating molecules in biological chemistry. Its structure consists of three moieties: methyl, adenosyl, and α-amino butanoate attached to a sulfur atom, generating a unique sulfonium moiety.[15] This positively-charged sulfur center withdraws electron density from the three adjacent carbon atoms, making them potent electrophilic centers, a feature exploited by various SAM-utilizing enzymes to catalyze a vast multitude of biochemical transformations[16,17] (Figure 3).

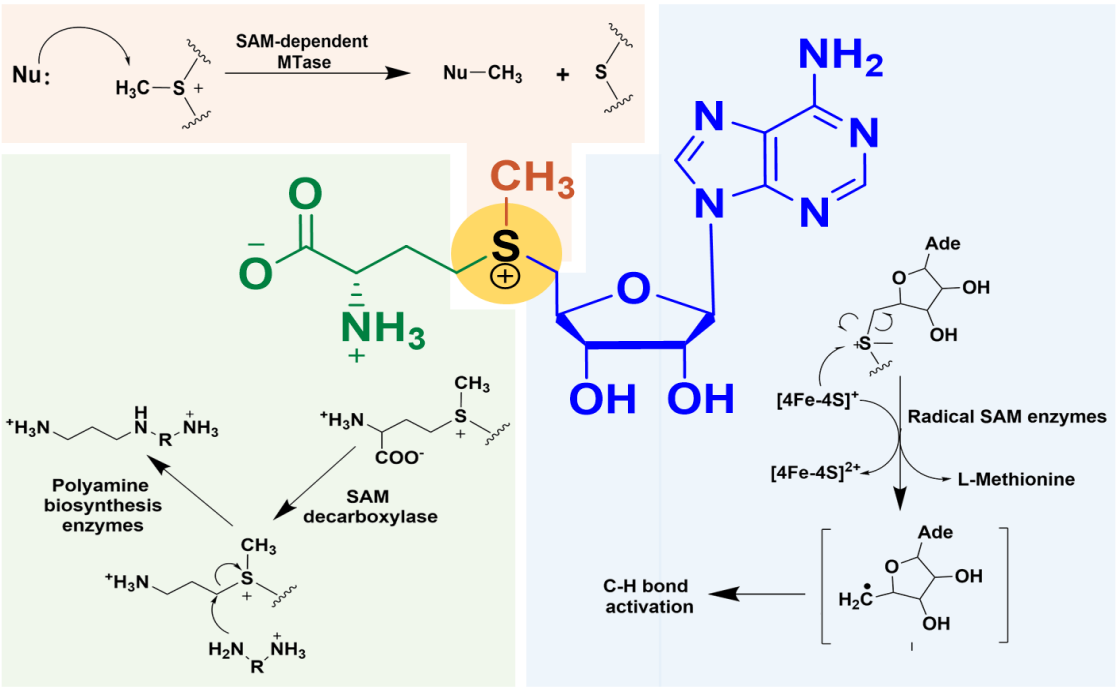


Figure 3: S-adenosyl methionine (SAM) is integral to cellular metabolism. SAM contains three structural moieties attached to a sulfonium center (highlighted in yellow) - methyl, adenosyl, and α-amino butanoate (highlighted in brown, blue, and green, respectively), each of which gets utilized in critical biochemical processes.

For example, some SAM-dependent methyltransferases (MTases) catalyze the S_N2 attack of biological nucleophiles on the methyl carbon of SAM to transfer the methyl group to different biomolecules.[18] These methylations play significant roles in gene expression, protein regulation, and the metabolism of natural products. On the other hand, radical SAM enzymes, using [Fe-S] clusters, exploit the electron-poor nature of the 5'-adenosyl carbon to facilitate the reductive homolytic cleavage of its bond with the sulfur, generating 5'-deoxyadenosyl radicals.[19] These radicals then activate C-H

bonds for performing unique biochemical transformations which play a critical role in several metabolic pathways, including lysine and pyruvate metabolism, DNA repair, and biosynthesis of cofactors like biotin, thiamin, heme, lipoic acid, and vitamin B_{12}. A third class of enzymes involved in the biosynthesis of bioactive molecules like polyamines, ethylene (in plants only), and quorum-sensing agents (in bacteria only) catalyzes nucleophilic attacks on the γ-carbon of the α-amino butanoate moiety.[16]

Thus, the presence of the sulfonium center in SAM leads to each of its structural moieties being utilized in several critical biochemical processes. This kind of versatility consequently motivates investigations into the substrate promiscuity of the SAM biosynthesis enzyme in order to create SAM analogs.

1.4 Methionine adenosyltransferase (MAT) biosynthesizes SAM

The biosynthesis of SAM is catalyzed by the gene product of *metK* called methionine adenosyltransferase (MAT), which is often referred to as SAM synthetase.[20] MAT utilizes adenosine-5'-triphosphate (ATP) and L-methionine as substrates to produce SAM, along with one molecule each of orthophosphate (P_i) and pyrophosphate (PP_i) as byproducts (Figure 4). A divalent and a monovalent metal cofactor (usually Mg^{2+} and K^+) are also required for catalysis, primarily to stabilize ATP's negatively charged triphosphate (PPP_i) moiety.[21]

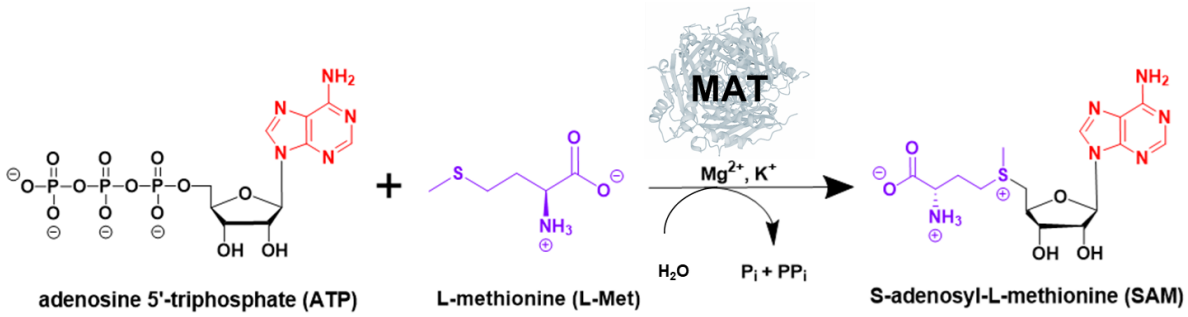


Figure 4: Methionine adenosyltransferase (MAT) catalyzes the reaction between ATP and methionine to biosynthesize SAM. SAM analogs differing at the methionine (lilac) or the nucleobase (red) moieties could be enzymatically synthesized by substrate promiscuous MATs.

Substrate promiscuity in MATs can be considered for either methionine or ATP. The enzymatic (using a promiscuous MAT) as well as the chemical synthesis of SAM methionine analogs (especially those where the terminal methyl group of methionine

is replaced with a fluorescent[22], photocleavable[23,24], or a biorthogonal[25] moiety), is well explored in the literature. Such SAM analogs have been used to probe cellular methylation[25,26] and study MTase activity[27] and inhibition[28,29]. On the other hand, we were interested in investigating the ability of MATs to accept other nucleoside triphosphates (NTPs), as they are cytosolic proteins with access to other NTPs like guanosine-5'-triphosphate (GTP), cytidine-5'-triphosphate (CTP), and uridine-5'-triphosphate. Consequently, an NTP-promiscuous MAT could produce SAM analogs differing at the nucleobase (aka S-nucleosidyl-L-methionine or SNMs) inside the cell.

1.5 Proposed catalytic mechanism for MAT suggests the enzyme is nucleotide promiscuous

The catalytic mechanism for bacterial MATs was proposed by the Markham and Takusagawa groups in 2004[30] after a series of studies in the late 1990s and early 2000s on the structure elucidation and active site characterization of *Escherichia coli* MAT (*Ec*MAT)[31–35] (Figure 5). Firstly, ATP and methionine bind to the enzyme's catalytically inactive (open) state. There is no order in the binding of substrates. However, based on reported dissociation constants and cellular concentrations of the substrates, ATP likely binds first, followed by methionine.[36] The substrate binding event causes the gatekeeper loop to fold into a helix to close the active site, resulting in the enzyme attaining its catalytically active (closed) state.[37] This conformational change results in the ATP ribose ring changing conformation from C4'-endo to C3'-exo, enabling the L-methionine sulfur center to simultaneously undergo an S_N2 attack on the 5'-carbon center of the ribose in ATP, leading to the formation of SAM and release of PPP_i.[30] The imidazole sidechain of a catalytic histidine residue (His14 in *Ec*MAT) stabilizes the transition state to facilitate the nucleophilic substitution further. PPP_i subsequently undergoes hydrolysis at the γ-phosphate position yielding P_i and PP_i, providing free energy to revert the enzyme conformation to its open state.[36] Finally, the newly-formed products leave the active site to complete the catalytic cycle. SAM leaves first, followed by the release (in no particular order) of PP_i and P_i.

Notably, the adenine moiety in ATP is relatively far from the reaction center and does not play an active role in the reaction mechanism. This fact hints that MATs could catalyze reactions with other NTPs (differing at the nucleobase position) to generate SNMs.

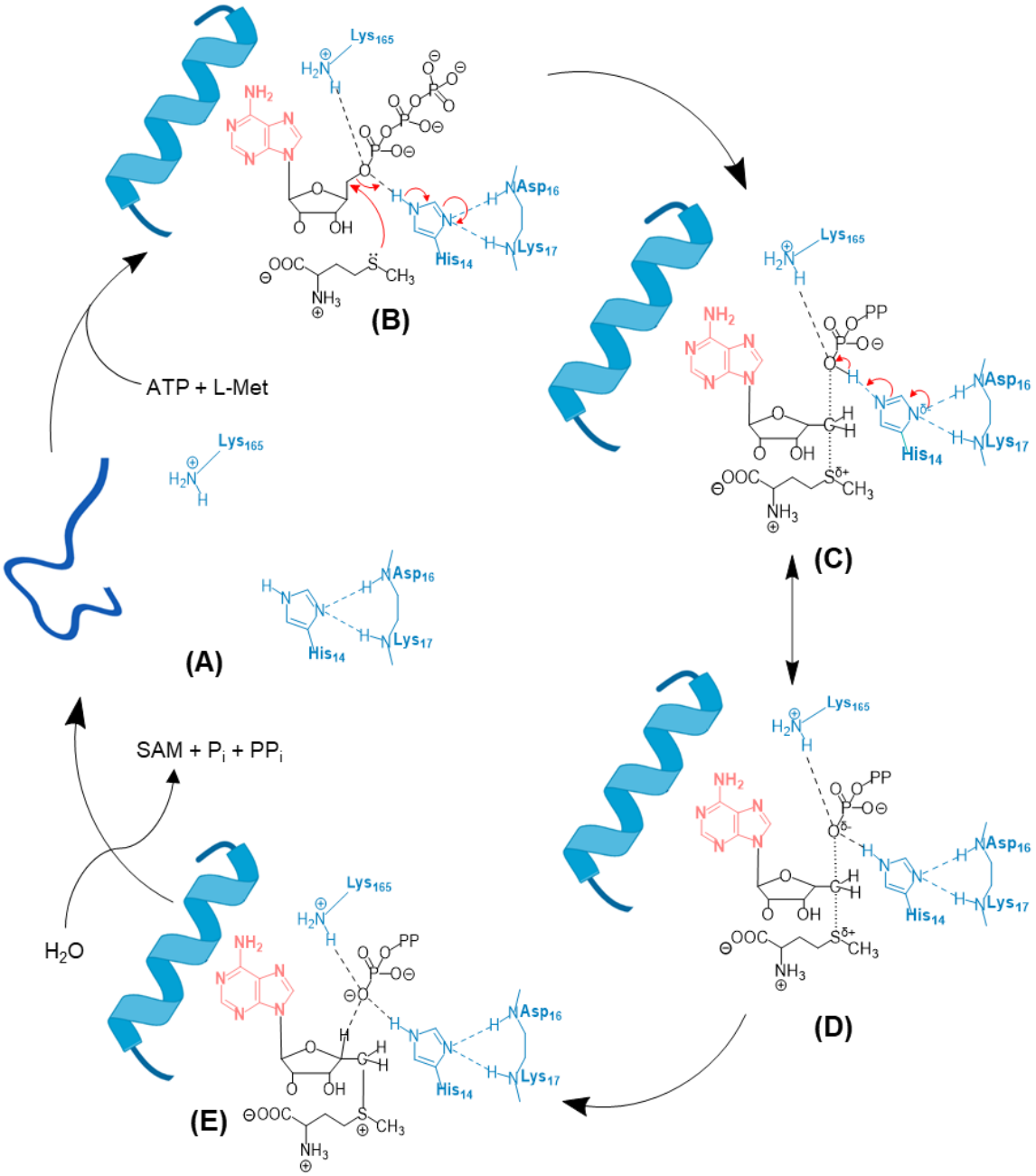


Figure 5: Proposed catalytic cycle for *Ec*MAT (A) MAT in its inactive open state. The gatekeeper loop (dark blue) is in a disordered conformation. **(B)** Upon ATP and methionine binding, the gatekeeper loop folds into a helix closing the active site. L-Met undergoes an S_N2 attack on the 5'-carbon of the ribose. **(C)** and **(D)** represent the transition state stabilized by His14. This residue is appropriately oriented via hydrogen bonds with Asp16 and Lys17. Lys165 stabilizes the leaving PPP_i group. **(E)** Formation of SAM and PPP_i. The latter hydrolyzes into PPi and Pi, triggering the unraveling of the gatekeeper helix and the release of the products. The transition from **(B)** to **(E)** occurs concertedly but is shown in distinct steps for clarity. The adenine moiety (pink) does not seem to play any role in the mechanism. This figure is adapted from Komoto et al. (2004) *Biochemistry*, *43* (7), 1821–1831.

1.6 Some MATs are promiscuous to naturally occurring NTPs

Investigations into the nucleotide promiscuity of MATs have been limited. Even among these, most studies have reported promiscuous behavior towards synthetic ATP analogs rather than naturally occurring nucleotides. These findings suggest that despite their reaction mechanism allowing for substrate promiscuity, most MATs seem to have evolved ATP specificity among the NTPs present in the cell. Nevertheless, four MAT homologs have been reported to show activity with naturally occurring non-cognate NTPs.

The 2002 study by Lu and Markham characterizing the MAT from the thermophilic archaeon *Methanocaldococcus jannaschii* (*Mj*MAT) is one of the earliest reports of a nucleotide-promiscuous MAT.[38] The authors performed reactions of various NTPs and $^{14}CH_3$-methionine with *Mj*MAT and detected $^{14}CH_3$-SNM formation using a cation-exchange filter-binding method. *Mj*MAT was reported to be active with several natural and synthetic NTPs, including ATP, GTP, CTP, UTP, and ITP (inosine-5'-triphosphate). *Mj*MAT's ability to react with ATP, GTP, CTP, and UTP and form the respective SNMs was also verified in our lab using HPLC (Figure 6B) and LC-MS.[39]

Like *M. jannaschii*, *Thermococcus kodakarensis* is a thermophilic archaeon whose MAT (*Tk*MAT) is nucleotide promiscuous. *Tk*MAT shows activity with ATP, CTP, and ITP. The full substrate scope of this enzyme is yet to be investigated.[40]

The MAT from the protist *Cryptosporidium hominis* (*Ch*MAT) and the 2A isoform of human MAT (*Hs*MAT2a) also exhibit nucleotide promiscuity. Apart from ATP, both *Ch*MAT and *Hs*MAT2a can accept a variety of non-natural ATP analogs.[41] *Ch*MAT is also active with ITP, while *Hs*MAT2a shows activity with GTP, CTP, and UTP.[42] Interestingly, the 1A isoform of human MAT (*Hs*MAT1a) is ATP-specific despite having 85% sequence identity with *Hs*MAT2a.[43]

It is important to note that all the mentioned studies are reports of '*in vitro'* promiscuity. The only evidence of MATs being promiscuous *in vivo* was the detection of 5'-methylthioguanosine (MTG), a degradation product of S-guanosyl methionine (SGM), in human liver cancer cell lysate.[42] However, the formation of SNMs by nucleotide-promiscuous MATs inside living organisms has not yet been conclusively demonstrated.

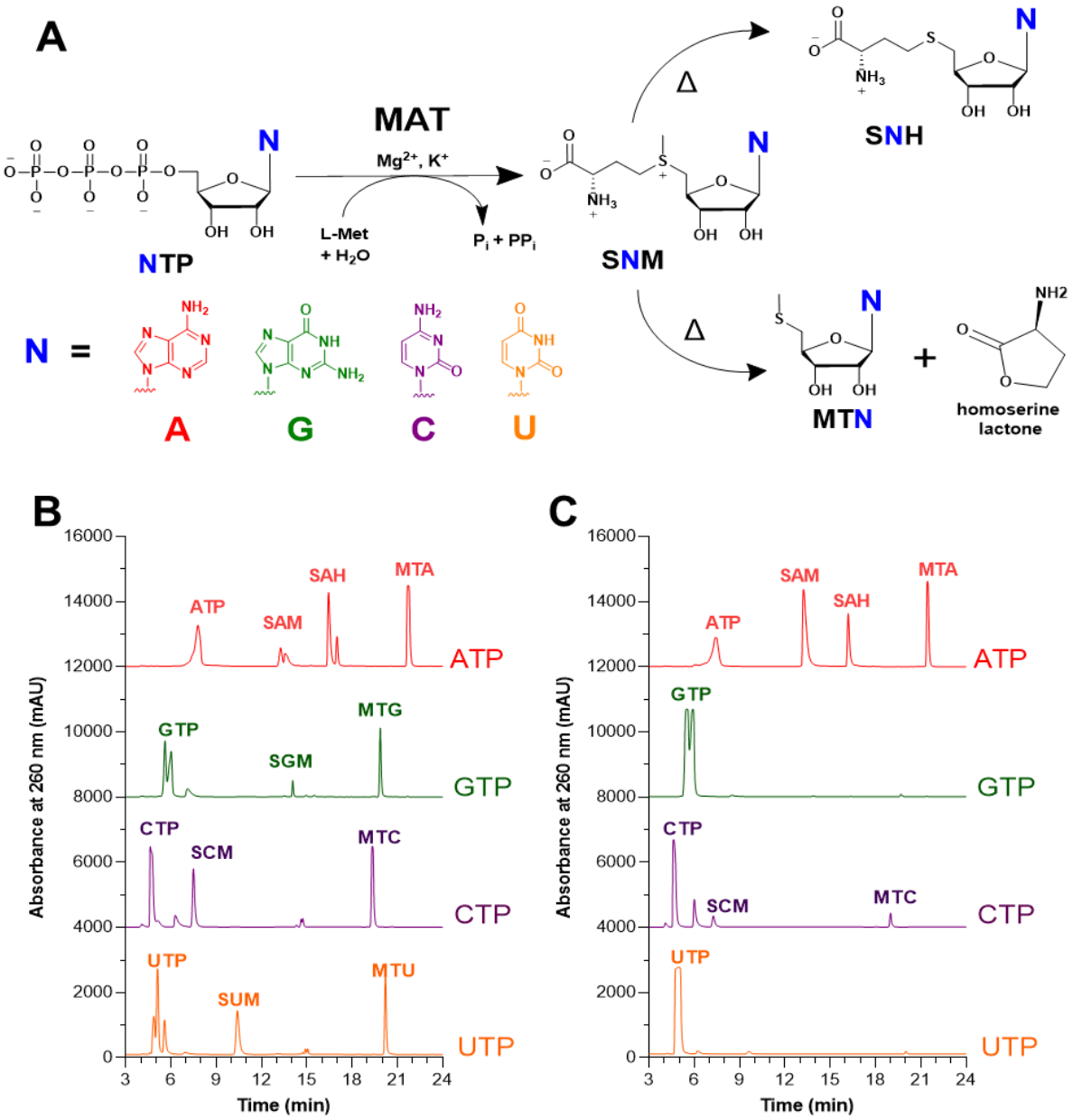


Figure 6: *In vitro* nucleotide promiscuity of *Mj*MAT and *Ec*MAT. (A) MATs catalyze the reaction of various NTPs with methionine to produce SNMs. Thermal degradation of SNMs produces S-nucleosidyl-L-homocysteines (SNHs) and 5'-methylthionucleosides (MTNs), which are also visible on the HPLC trace of the reaction. **(B)** HPLC traces of reactions of *Mj*MAT with ATP, GTP, CTP, and UTP show that the enzyme is active with all these NTPs. **(C)** HPLC traces of reactions of *Ec*MAT with the same four NTPs indicate that the enzyme is practically ATP-specific.

1.7 *Escherichia coli* MAT (*Ec*MAT) exhibits low nucleotide promiscuity

Escherichia coli has arguably been the most popular model organism for several decades due to its fast growth, versatility in growth conditions, and the availability of

tools for easy genetic manipulation.[44] These features also make this bacterium ideal for studying the effects of SAM nucleobase analogs on cellular physiology and metabolism, as well as designing new biorthogonal metabolic pathways that utilize these analogs. The MAT from *E. coli* (*Ec*MAT) displays *in vitro* promiscuity towards several synthetic ATP analogs.[45] However, among naturally occurring nucleotides, it is practically ATP-specific[38,39,42,46] (Figure 6C). The enzyme does show limited CTP conversion; however, the activity is 1/8th that of *Mj*MAT with CTP. *Ec*MAT's activity with GTP and UTP is dramatically lower than that of *Mj*MAT, resulting in indiscernible product peaks on the HPLC trace.

Thus, wild-type (WT) *Ec*MAT is not nucleotide-promiscuous enough to produce sufficient amounts of SNMs *in vitro*, let alone inside the organism. Consequently, *Ec*MAT needs to be mutated to enhance its ability to accept naturally occurring non-cognate NTPs, before it can be introduced inside the cell to test *in vivo* effects of SNMs.

1.8 Scope of the Study

MAT's catalytic mechanism is indiscriminate towards any particular nucleotide. Hence, it would be reasonable to expect that the nucleotide promiscuity of MATs is influenced by how well non-cognate NTPs bind to the enzyme's active site. Thus, the difference in NTP promiscuity of *Ec*MAT and *Mj*MAT might be due to the dissimilarities between the two enzymes' active site residues. This study aims to pinpoint the residues that form the basis behind this variation in nucleotide promiscuity. We first determine the differences in the two actives sites by comparing the crystal structures of the two enzymes. Based on this analysis, we then engineer *Ec*MAT point mutants where the parent residue is replaced by the corresponding one from *Mj*MAT. The *in vitro* ability of these mutants to accept ATP, GTP, CTP, and UTP is then tested. Based on the observed changes, we develop hypotheses on why specific intermolecular interactions are more significant in generating promiscuity than others. Finally, these mutants have also been introduced inside the *E. coli* organism to examine whether they can produce SNMs *in vivo.*

2. Materials and Methods

2.1 Materials

The pET15b (WT *EcmetK*) construct was cloned in our lab, as previously reported.[47] This vector produces N-6X-His-*Ec*MAT using the T7 expression system. *Escherichia coli* BL21(DE3) strain containing pET19b (WT *MjmetK*) (that expresses N-10X-His-*Mj*MAT) was a gift from Prof. Squire J. Booker's lab at Pennsylvania State University. *E. coli* MOB1490, a SAM auxotrophic strain, was kindly provided by Prof. Robert Blumenthal at the University of Toledo.[48] It contains Δ*metK* (*metK::kan*), plasmid pMW1402 harboring the SAM transporter from *Rickettsia prowazekii* (and Amp^R), and pMW1484, a pBAD-derived plasmid encoding *rpoB* (Rif^R). MOB1490 cells containing pUC57(WT *EcmetK*) were also provided by Dr. Blumenthal, from which the pUC57 vector was extracted.

Media components were purchased from HiMedia. Antibiotics were obtained from TCI chemicals and Sigma-Aldrich. 100 mM NTP solutions were procured from Jena Biosciences. L-methionine was obtained from TCI Chemicals. L-arabinose was obtained from SRL Chemicals. All other chemicals used were from Sigma-Aldrich (Merck). Protein ladders (10 to 250 kDa) were purchased from Genetix and ThermoFischer Scientific. DNA ladders (250 – 10,000 bp) were procured from NextGen. High-fidelity PrimeSTAR GXL Polymerase and DpnI were purchased from DSS TaKaRa Bio USA. Plasmid and DNA purification kits were obtained from Qiagen, Agilent, and Promega. Solvents for HPLC were procured from SD Fine Chemicals, Rankem, and Sigma-Aldrich. LC-MS grade methanol was purchased from J. T. Baker.

2.2 Experimental Methods

2.2.1 Restriction-free site-directed mutagenesis of EcMAT

Q98H, G117N, and I232T *Ec*MAT mutations were generated using megaprimer-based site-directed mutagenesis. Firstly, a polymerase chain reaction (PCR) is performed on pET15b (WT *EcmetK*) plasmid using flanking and mutagenic primers to generate a DNA fragment (called the megaprimer) containing the mutation at the desired location. In the case of Q98H and G117N mutants, the mutagenic primer is the forward primer, and the commercially available T7 terminator primer is the flanking primer. However,

for the I232T mutant, the mutagenic primer is the reverse primer, while the T7 promoter primer is the flanking primer. The successful formation of the megaprimer is confirmed on a 1% agarose gel, after which it is purified using protocols described in the Qiagen PCR Purification Kit. A second PCR is then performed on the pET15b (WT *EcmetK*) plasmid using the purified megaprimer to obtain the entire plasmid with the desired mutation.

R229V, I102L, T227A, D238S, and S186A mutations were introduced via Quickchange-based site-directed mutagenesis.[49] In this method, only a single PCR is performed on the pET15b (WT *EcmetK*) plasmid using appropriately designed forward and reverse Quickchange mutagenic primers, to amplify the entire plasmid with the desired mutation. The introduction of I102L mutation into pUC57 (WT *EcmetK*) plasmid was also achieved by the Quickchange method.

The product obtained from Quickchange PCR (or the second PCR in the megaprimer method) is then incubated with Dpn1 at 37 °C for 2.5 hours to digest the methylated template plasmid. Subsequently, a heat-shock transformation of the digestion product into chemically competent *E. coli* DH5α cells is performed. The transformed cells are plated on LB (Luria-Bertani) agar plates containing 100 µg/mL ampicillin and 0.2% glucose and incubated overnight at 37 °C. Colonies obtained on the plate are substreaked on a fresh LB agar plate containing 100 µg/mL ampicillin and incubated at 37 °C for 7-8 hours. Single colonies are picked from the substreak plate and inoculated into primary cultures containing LB Broth and 100 µg/mL ampicillin, which are then grown overnight at 37 °C with 180 rpm shaking. The plasmid is extracted from the harvested cells using a plasmid purification kit. The successful introduction of the mutation in the plasmid is confirmed via Sanger sequencing. The pET15b-derived plasmids were sequenced using the T7 promoter and terminator, and the pUC57 vectors were sequenced using M13 primers.

The cloning of Q98H *Ec*MAT was successfully performed by Vruta Gupte. Anjali Pattathil contributed to the cloning of D238S and S186A *Ec*MAT mutants.

All primers used for cloning and sequencing are listed in Table 1.

Sr no.	Name of primer	Sequence (5' → 3')
1	Q98H_MP_F	ATCGGCAAACATTCTCCTGACATCAAC
2	G117N_MP_F	GAACAGGGCGCGAACGACCAGGGTCTG
3	I232T_MP_R	ATTGGGCCACCGGTAACGAAACGA
4	R229V_QC_F	CGGTGTGTTCGTTATCGGTGGCCCAATGGG
5	R229V_QC_R	ATAACGAACACACCGGTCGGGTTGATGAAGAATTTGG
6	I102L_QC_F	CCTGACCTGAACCAGGGCGTTGAC
7	I102L_QC_R	CTGGTTCAGGTCAGGAGACTGTTTGCCGATAGCG
8	T227A_QC_F	CCCGGCGGGTCGTTTCGTTATCGGTGGCCC
9	T227A_QC_R	CGACCCGCCGGGTTGATGAAGAATTTGGTGGCAG
10	D238S_QC_F	TGGGTTCTTGCGGTCTGACTGGTCGTAAAATTATCG
11	D238S_QC_R	CCGCAAGAACCCATTGGGCCACCGATAACG
12	S186A_QC_F	GCTTGCGACTCAGCACTCTGAAGAGATCGACC
13	S186A_QC_R	CTGAGTCGCAAGCACGACAGCATCGATACCAAC
14	T7_promoter_F	TAATACGACTCACTATAGGG
15	T7_terminator_R	GCTAGTTATTGCTCAGCGG
16	M13_promoter_F	TGTAAAACGACGGCCAGT
17	M13_terminator_R	CAGGAAACAGCTATGACC

Table 1: Primers used for generating and sequencing *Ec*MAT mutants. The mutagenic primers have been named in the following manner: the mutation _ mutagenesis method (megaprimer (MP) / Quickchange (QC)) _ forward (F) / reverse (R). The codon encoding the mutation is highlighted in red.

2.2.2 Overexpression and purification of wild-type and mutant EcMATs

The overexpression and purification of *Ec*MATs were performed per the following previously reported protocol.[39] First, the pET15b plasmid carrying the WT or mutant *EcmetK* gene is incorporated into chemically competent *E. coli* BL21(DE3) cells via heat-shock transformation. The transformed cells are plated and grown overnight at 37 °C on LB agar plates containing 100 µg/mL ampicillin. A single colony is then

inoculated into a primary culture containing LB broth and 100 µg/mL ampicillin, which is then grown overnight at 37 °C with 180 rpm shaking. Next day, 2L LB secondary culture is inoculated with 1% of the primary culture and grown at 37 °C with 180 rpm shaking till the OD_{600} reaches 0.6-0.8. After that, 200 µM isopropyl β-D-1-thiogalactopyranoside (IPTG) is added, and the culture is incubated at 18 °C with 180 rpm shaking for 18 hours, followed by centrifugation at 5500 rpm for 1 hour in a swinging bucket rotor at 4 °C to harvest cells. The harvested cell pellet is stored at -80 °C.

For MAT purification, the cell pellet was resuspended in lysis buffer (50 mM Tris-Cl pH 8.0, 10 mM imidazole, 300 mM NaCl, 0.5 mM phenylmethylsulfonyl fluoride (PMSF), and 0.025% β-mercaptoethanol (β-ME)). The cell suspension was subjected to lysis by sonication (60% amplitude, 1s on 3s off cycle, 10 minutes). The cell lysate was centrifuged at 18000 rpm for 20 mins at 4 °C, and the supernatant was loaded onto a Ni-IDA column equilibrated with lysis buffer. The column was washed with 10 column volumes of wash buffer (50 mM Tris-Cl pH 8.0, 30 or 15 mM imidazole, 300 mM NaCl, 0.025% β-ME). The protein is eluted in elution buffer (50 mM Tris-Cl pH 8.0, 250 mM imidazole, 300 mM NaCl, and 0.025% β-ME). The presence of *Ec*MAT in the eluted fractions is verified using SDS-PAGE. Fractions containing the protein are pooled together, loaded onto a Bio-Rad Econo-Pac 10DG desalting column, and then desalted using a desalting buffer (50 mM Tris-Cl pH 8.0, 150 mM NaCl, and 0.025% BME). A 12% glycerol solution of the protein is prepared, flash-frozen with liquid N_2, and finally stored at -80 °C.

Protein concentration was determined using UV-vis spectrophotometry by measuring the absorbance at 280 nm. The molar extinction coefficient (ε = 37880 $M^{-1}\,cm^{-1}$) of WT and mutant *Ec*MATs was predicted using the Expasy ProtParam web tool.[50]

2.2.3 Reactions of MATs with various NTPs

Reactions of WT and mutant *Ec*MATs and WT *Mj*MAT with ATP, GTP, CTP, and UTP were performed. The reaction mixture comprised 10 mM NTP, 15 mM L-Met, 20 µM MAT, 50 mM HEPES-NaOH pH 8.0, 50 mM $MgCl_2$, and 20 mM KCl. The *Ec*MAT reactions were run at 37 °C, while the *Mj*MAT ones were carried out at 55 °C. All reactions were performed for 24 hours. The reactions are quenched by adding 9%

formic acid, followed by centrifugation at 14000 rpm for 40 minutes at 4 °C to precipitate the enzyme. The supernatant is neutralized by adding 0.1 M NaOH and then injected into HPLC to detect and quantify SNM formation.

2.2.4 Analysis of MAT reaction mixtures using HPLC

The reverse phase HPLC of reaction mixtures was performed on an Agilent 1260 Infinity II UPLC system paired with a UV-Vis DAD detector, using an Agilent Zorbax Eclipse Plus C18 5µ 250mm x 4.60mm column. The solvent system comprised 10 mM pH 6.5 ammonium acetate (solvent A) and methanol (solvent B). The mobile phase gradient used was: 0-2 min 100% A, 2-15 min 100% A to 70% B, 15-20 min 70% B, 20-20.5 min 70% B to 100% B, 20.5-25 min 100% B, 25.0-25.5 min 100% B to 100% A, 25.5-35 min 100% A.

The chromatograms were recorded at 260 nm, as all four nucleobases strongly absorb at that wavelength (Figure 7A). Consequently, any molecule possessing the nucleobase moiety is visible on the chromatogram. These include thermal degradation products of SNMs - SNHs (nucleobase analogs of S-adenosyl-L-homocysteine (SAH)) and MTNs (nucleobase analogs of 5'-methylthioadenosine (MTA)), which can also be used to infer SNM generation. MTNs are the primary degradation product of SNMs. Their peak is usually more prominent than that of SNM itself. SNHs are a minor degradation product and hence are detectable only in reactions with very high NTP turnover (like in the case of ATP). The HPLC traces were analyzed using OpenLabCDS and plotted using GraphPad Prism 8.4.3

2.2.5 Quantification of MAT nucleotide promiscuity

The formation of SAM from the reaction of MATs with ATP is quantified in the following manner. ATP solutions of concentrations in the range of 312.5 – 5000 µM are prepared via serial dilution. The HPLC of these solutions is performed to prepare an 'area under the curve' (AUC) vs. concentration calibration curve (Figure 7B). The final ATP concentration in the reaction mixture is then calculated by substituting the AUC value of the ATP peak into the standard curve equation. This value is subtracted from the starting ATP concentration (10 mM) to obtain the amount of SAM formed.

In the case of other SNMs, the turnover is very low, leading to detector saturation for the NTP peak. Consequently, their formation could not be quantified using the above method. Instead, the activity of *Ec*MAT mutants with the other NTPs relative to WT *Mj*MAT is determined. This value is obtained by calculating the ratio of the AUC sum of the SNM, SNH, and MTN peaks from the *Ec*MAT mutant-NTP reaction trace to that obtained from the *Mj*MAT-NTP reaction trace. This method assumes that the extinction coefficients of SNM, SNH, and MTN (for a particular nucleobase) are identical, as all three possess the same absorbing moiety.

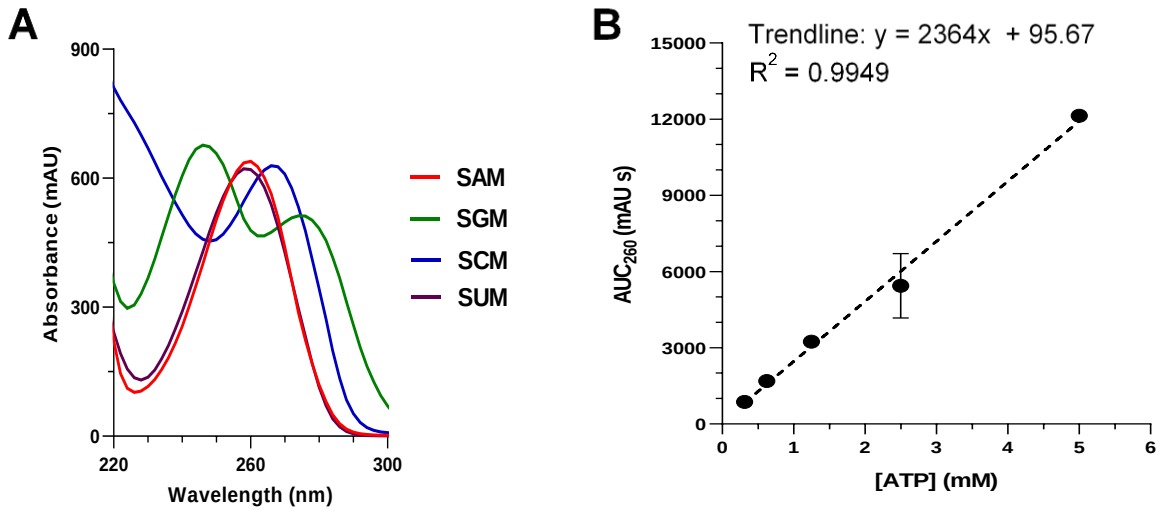


Figure 7: HPLC analysis of *Ec*MAT reaction mixtures. (A) Absorbance spectra of SAM, SGM, SCM, and SUM. The HPLC chromatograms were recorded at 260 nm as all four molecules absorb strongly at that wavelength. **(B)** The calibration curve used for determining ATP concentration in the reaction mixtures

2.2.6 Complementation of E. coli MOB1490 strain with mutant EcMAT genes

The pUC57 (I102L *EcmetK*) plasmid is introduced into chemically competent *E. coli* MOB1490 cells via heat-shock transformation. Following heat shock, the cells are recovered in LB broth containing 50 µM SAM. These cultures are then plated on LB agar plates containing 50 µg/mL ampicillin, 25 µg/mL kanamycin, 50 µg/mL rifampin, and 5 mM L-arabinose and incubated at 37 °C for 20 hours. A single colony is then inoculated into a primary culture containing LB Broth and 50 µg/mL ampicillin, 25 µg/mL kanamycin, 50 µg/mL rifampin, and 5 mM L-arabinose, which is then grown overnight at 37 °C. Glycerol stocks are prepared from the primary culture.

2.2.7 Extraction of metabolites from E. coli cultures

The BL21(DE3) cells containing pET15b empty vector, pET15b (mutant or WT *EcmetK*), and pET19b (WT *MjmetK*) are grown in separate 50 mL LB cultures containing 100 µg/mL ampicillin at 37 °C overnight with 180 rpm shaking. Also, parental MOB1490 cells and those complemented with the WT and I102L *EcmetK* are cultured in 50 mL LB broth containing 50 µg/mL ampicillin, 25 µg/mL kanamycin, 50 µg/mL rifampin, and 5 mM L-arabinose at 37°C for 20 hours with 180 rpm shaking. As the parental MOB1490 strain is a SAM auxotroph, 50 µM SAM is added only to its medium to support its growth.

Cell pellets are obtained after centrifugation of the cultures at 6500 rpm for 10 minutes at 4 °C, which are then resuspended in 2 mL of cold lysis solution (methanol, acetonitrile, and water (2:2:1 by volume) with 0.1 M formic acid). The resuspended cells are subjected to lysis by sonication (100% amplitude, 1s on 1s off cycle, 3 minutes). Next, the lysates are passed through a 0.22 µm filter, and the solvent is removed from the flowthrough using a Centrivap. Finally, the samples are redissolved in 100 µL of lysis solution and injected into LC-MS for detecting *in vivo* SNM formation.

2.2.7 Analysis of reaction mixtures and cell lysates using LC-MS

LC-MS was performed on ScieX X500R-QTOF mass spectrometer system attached to Exion-LC series UHPLC. The LC method is identical to the HPLC method described in Section 2.3.4. The MS method is as follows: Electrospray ionization (ESI) was executed in the positive mode with +4.5 kV ion spray voltage, medium de-clustering potential (80V), and low collision energy (5V) at 400 °C. SNMs and their degradation products were identified using a quadrupole time of flight (QTOF) detector followed by targeted MRM-HR analysis using the following precursor and fragment ion masses: ATP (508.0, 508.0030), GTP (524, 523.9979), CTP (484, 483.9918), UTP (484.1, 484.0758), methionine (150.1, 150.0583), (SAM (399.1, 399.1445), SAH (385.1, 385.1215), MTA (298.1, 298.0968), SGM (415.1, 415.1416), SGH (400.1, 400.1086), MTG (314.1, 314.0918), SCM (375.1, 375.1333), SCH (361.1103), MTC (274.1, 274.0856), SUM (376.1, 376.1173), SUH (362.1, 362.0943), MTU (275.1, 275.0696). Data were processed on the ScieX-OS Explorer software to obtain extracted ion chromatograms (EIC), peak integrations, and the mass spectra of desired analytes.

The mutant is considered active with a particular NTP if either SNMs or their degradation products get detected on LC-MS. A peak corresponding to the masses of SNMs and MTNs was detected for most reactions. SNHs are sometimes not found as they are a minor degradation product.

2.3 Computational Methods

2.3.1 Structural alignment of EcMAT and MjMAT

Three crystal structures of SAM-bound WT *Ec*MAT have been reported previously. Among these, the one corresponding to the PDB ID 7LOO[42] is used for the structural alignment as it has the highest resolution (1.95 A°). As of March 2023, no crystal structure of SAM-bound WT *Mj*MAT has been reported. However, o-nitro benzyme SAM (ONB-SAM) bound to L147A/I347A *Mj*MAT double mutant has been crystallized with a resolution of 2.05 A° (PDB ID: 7P84)[24], which was used for the structural alignment.

The structural alignment was performed and visualized on UCSF Chimera[51] using the inbuilt MatchMaker tool. The similarities or differences between the aligned amino acid residues are checked for every active site residue. (For this study, the active site is defined as all residues lying within a zone of 5 A° from the bound SAM/ONB-SAM ligand.)

2.3.2 Multiple sequence alignment of bacterial and archaeal MATs

A protein BLAST (***B***asic ***L***ocal ***A***lignment ***S***earch ***T***ool — available at https://blast.ncbi.nlm.nih.gov/Blast.cgilink)[52] search was performed on the FASTA sequences of WT *Ec*MAT and WT *Mj*MAT, respectively. The maximum number of search results was increased to 5000. Out of the BLAST results, a list of 54 and 93 archaeal and bacterial MATs was manually curated. Multiple sequence alignment of these lists of archaeal and bacterial MATs was achieved using the MUSCLE (***MU***ltiple ***S***equence ***C***omparison by ***L***og-***E***xpectation)[53] alignment algorithm. The alignment was visualized using Jalview.[54] The conservation of active site residues in *Ec*MAT and *Mj*MAT among the listed bacterial and archaeal MATs was analyzed.

3. Results and Discussion

3.1 Comparison of *Ec*MAT and *Mj*MAT active sites

MATs generally possess a homo-oligomeric quarternary structure, with the active site formed at the interface of two peptide chains.[24,34] This active site can be divided into four binding domains (Figure 8). The methionine-binding domain, as the name suggests, aids in the attachment of methionine. The NTP-binding region is divided into three binding domains based on the moiety they interact with - the triphosphate-binding domain, the ribose-binding domain, and the adenine-binding domain. The MAT gating loop, wherein some adenine-binding residues are also present, encloses the active site once ATP and methionine are bound.

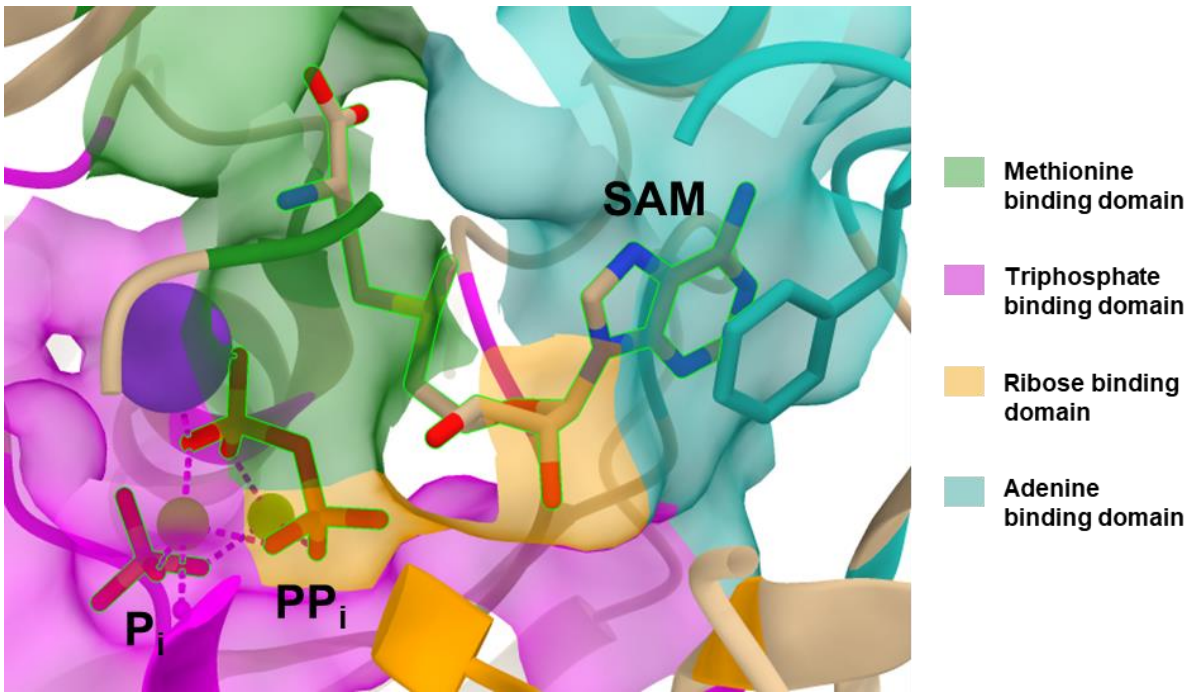


Figure 8: SAM, PP_i, and P_i bound to the active site of *Ec*MAT (PDB ID: 7LOO). The enzyme is in its catalytically active 'closed' conformation. The four binding domains constituting the MAT active site have been highlighted with different colors.

The structures of a MAT's binding domains would strongly influence which NTPs it accepts. Hence, structural alignment of *Ec*MAT with *Mj*MAT was performed to compare each binding domain of the former with the corresponding one of the latter (Figure 9). Ligand-bound crystal structures were used for the alignment so that exact differences between the two MATs' binding domains could be located while the

enzymes were in their catalytically active 'closed' conformations. These differences are described in the upcoming subsections.

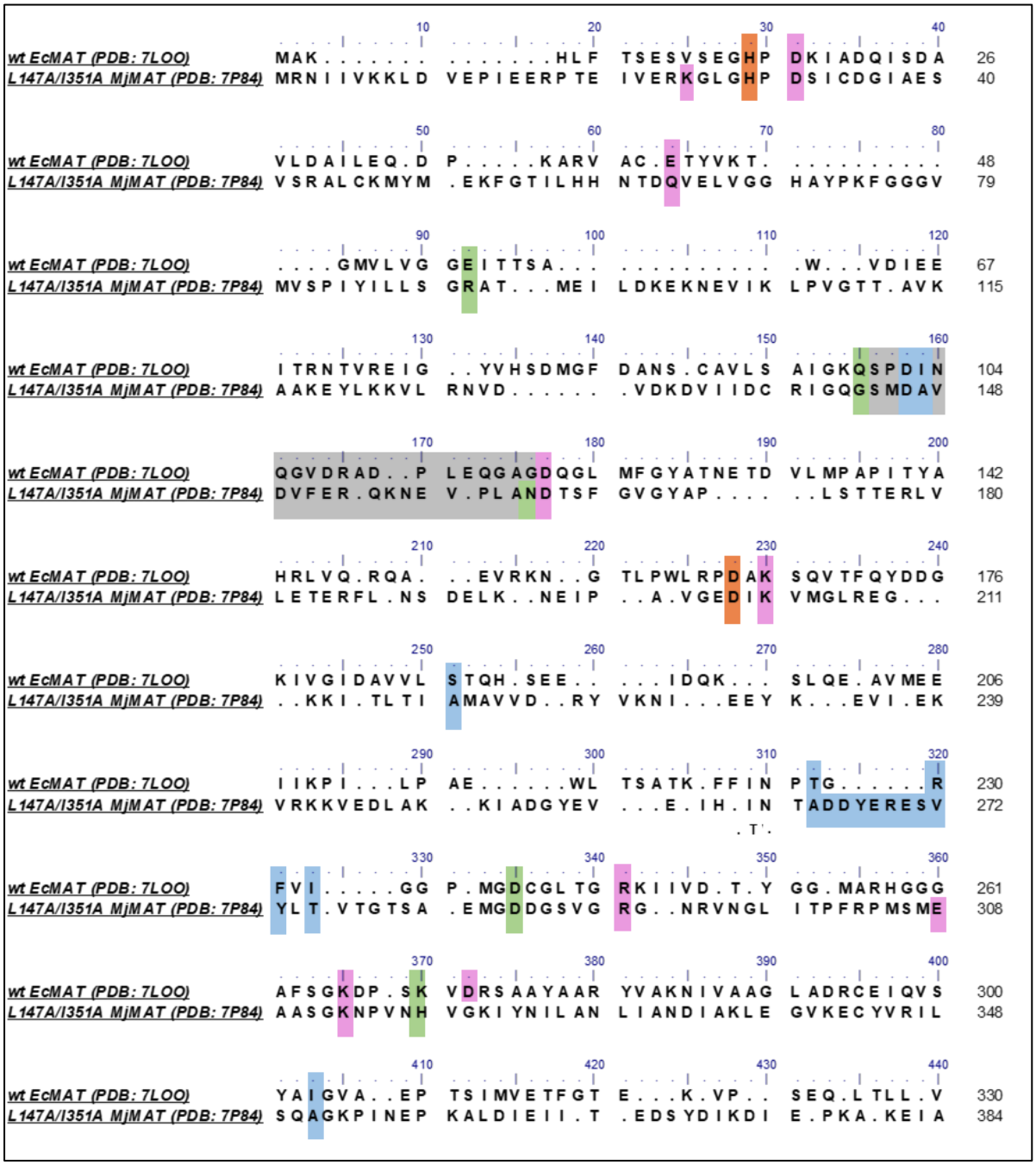


Figure 9: Alignment of ligand-bound structures of WT *Ec*MAT (PDB: 7LOO) and L147A/I351 *Mj*MAT (PDB: 7P84) Residues from the four bindings domain are color-coded in the following manner: methionine-binding domain – green, triphosphate-binding domain – pink, ribose binding domain – orange, and adenine-binding domain – blue. The gating loop sequence in both enzymes is highlighted in grey. Gaps are denoted with a point (.). Residues 330-386 of *Ec*MAT and 384-406 of *Mj*MAT are not shown in the figure as they do not contain any active site residues. This alignment was edited using BioEdit.

3.1.1 The methionine-binding domain

The methionine-binding domain of *Ec*MAT primarily consists of four residues – Q98, K269, E55, and D238 which form hydrogen bonds with L-methionine. The first two form hydrogen bonds with methionine's carboxylate moiety using their sidechain amide and amine groups, respectively. E55 and D238 form hydrogen bonds with the methionine's amine group using their sidechain carboxylate groups.

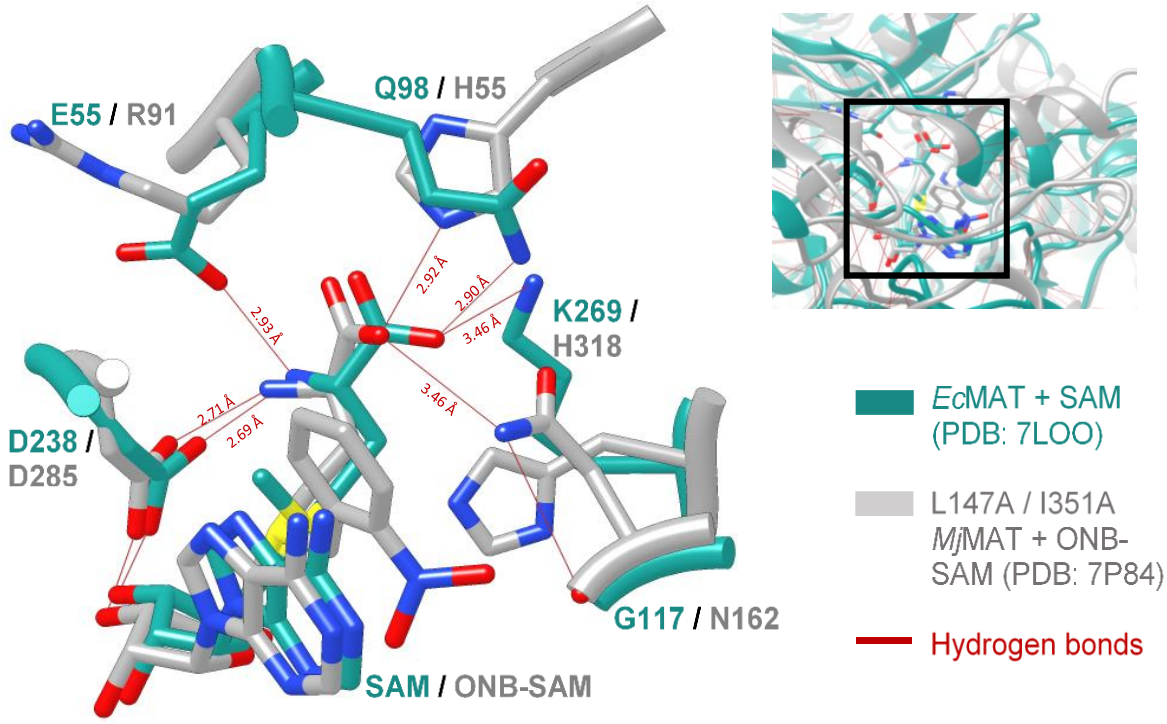


Figure 10: Comparison of methionine-binding domains from *Ec*MAT and *Mj*MAT. Hydrogen bond lengths are mentioned for the ones involving the methionine moiety. Active site water molecules have been omitted for the sake of clarity.

This domain in *Mj*MAT is significantly different from the one in *Ec*MAT (Figure 10). Of the four residues mentioned previously, only D238 is aligned to an identical residue in *Mj*MAT (D285). D285, like its *Ec*MAT counterpart, forms a hydrogen bond with the amine group of methionine. The sidechain of *Ec*MAT's Q98 is aligned with *Mj*MAT's H55. Thus, the imidazole moiety replaces the amide to form a hydrogen bond with methionine's carboxylate in *Mj*MAT. K269 and E55 are aligned to H318 and R91 in *Mj*MAT. Both these residues do not seem to interact with methionine, unlike their *Ec*MAT counterparts. Additionally, the sidechain amide of Asn162 in *Mj*MAT (aligned to Gly117 in *Ec*MAT) undergoes hydrogen bonding with methionine's carboxylate moiety – an interaction that is absent in *Ec*MAT.

3.1.2 The triphosphate-binding domain

The part of the *Ec*MAT active site primarily consists of charged amino acid residues that form an intricate network of hydrogen bonds and electrostatic interactions. The negatively charged residues D16, E42, D118, D238, and D271 contribute to binding one potassium and two magnesium ions. The bound metal cofactors join forces with the positively charged residues K165, R244, K245, and K265 to stabilize the highly negatively charged triphosphate moiety.

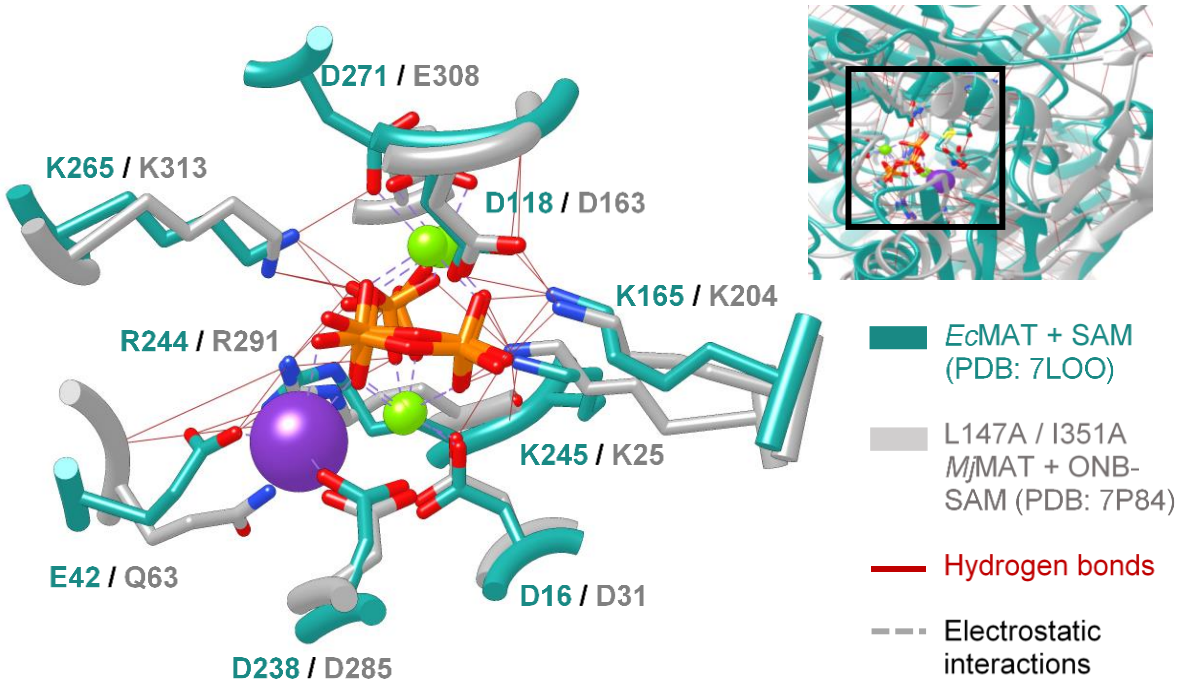


Figure 11: Structural comparison of triphosphate-binding domains from *Ec*MAT and *Mj*MAT. Active site water molecules have been omitted for the sake of clarity.

The triphosphate-binding domain in *Mj*MAT is very similar to *Ec*MAT's (Figure 11). Except for E42 and D271, all other residues mentioned above are aligned to identical residues in *Mj*MAT. These two are instead aligned to *Mj*MAT's Q63 and E308.

3.1.3 The ribose-binding domain

The ribose-binding domain of *Ec*MAT contains two aspartate residues- D163 and D238, which undergo hydrogen bonding with the ribose ring's 2' and 3' hydroxyl groups. It also contains H14, the catalytic histidine residue[31], located just below the 5' ribose carbon to stabilize the transition state formed upon the SN^2 attack of the methionine sulfur center.

*Mj*MAT's ribose-binding domain is identical to *Ec*MAT's (Figure 12). D163 and D238 in *Ec*MAT are aligned with D202 and D285 of *Mj*MAT. H14 is aligned to H29, the catalytic histidine residue of *Mj*MAT.

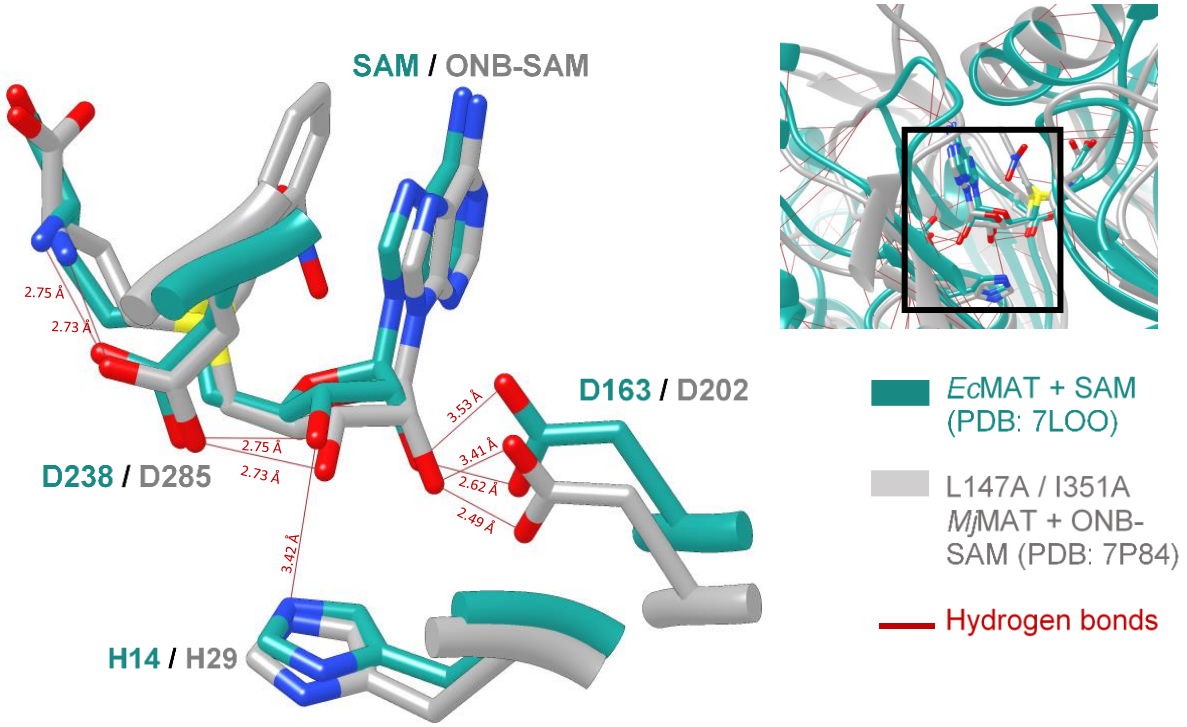


Figure 12: Structural comparison of ribose-binding domains from *Ec*MAT and *Mj*MAT. Hydrogen bond lengths are mentioned for the ones formed by SAM/ ONB-SAM. Active site water molecules have been omitted for the sake of clarity.

3.1.4 The adenine-binding domain

The *Ec*MAT adenine-binding domain notably contains very few hydrogen bonding interactions. Indeed, the adenine moiety only forms one direct hydrogen bond with the peptide chain – between the N6 amine and the backbone carbonyl of R229. Additionally, adenine participates in one water-mediated hydrogen bond – between the N1 nitrogen and the sidechain hydroxyl group of T227. The inherent minimal nucleotide promiscuity of WT *Ec*MAT might be due to this scarcity of hydrogen bonding with the nucleobase.

Instead, the binding domain consists primarily of residues that stabilize the nucleobase via van der Waals interactions. The most notable among these is the pi-stacking interaction between adenine and the aromatic sidechain of F230. Adenine also forms van der Waals contacts with D101, I102, S186, T227, I232, and I302.

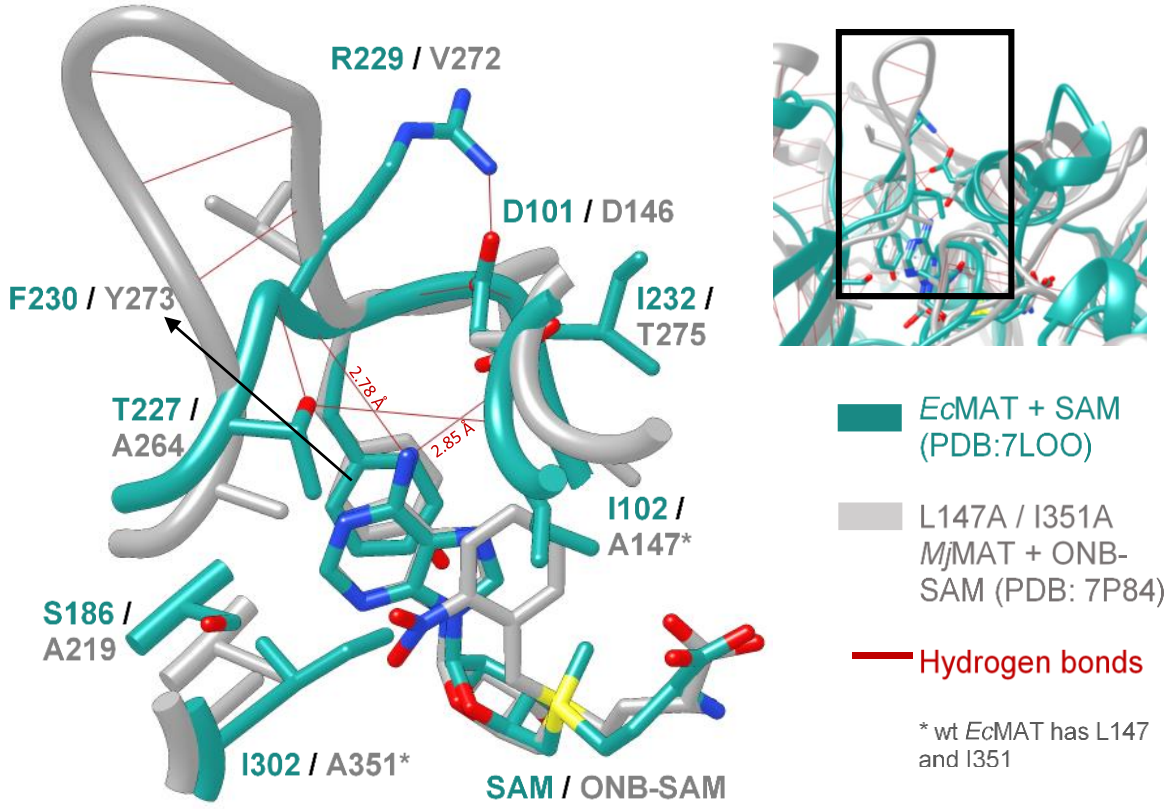


Figure 13: Structural comparison of adenine-binding domains from *Ec*MAT and *Mj*MAT. The 9-residue-long extra loop in *Mj*MAT is present at the top left of the figure. Hydrogen bond lengths are mentioned for the ones directly between adenine and the protein. Active site water molecules have been omitted for clarity.

The nucleobase-binding domain of *Mj*MAT is quite different from that of *Ec*MAT (Figure 13). Firstly, R229 of *Ec*MAT is aligned with V272 of *Mj*MAT. The N6 amine of adenine does not interact with the backbone carbonyl of Val272 but forms a hydrogen bond with the sidechain carboxylate of D146 (aligned to D101 in *Ec*MAT). Additionally, the residue pi-stacking with the nucleobase in MjMAT is a tyrosine (Y273) instead of phenylalanine (F230) as in *Ec*MAT. The phenolic hydroxyl group in Y273 forms a water-mediated hydrogen bond with the 3'-hydroxyl of the ribose ring. Some residues forming hydrophobic contacts with adenine in *Mj*MAT also differ from those in *Ec*MAT. These include L147, A219, A264 and T275 (aligned to I101, S186, T227 and I232 respectively in *Ec*MAT). Finally, *Mj*MAT possesses a loop of 9 residues (A264-V272) that is absent in *Ec*MAT.

A summary of the comparison between the two active sites is provided in Table 2.

***Ec*MAT residue**	***Mj*MAT residue**	**Type of interaction with ligand**	**Conservation among bacterial MATs**	**Conservation among archaeal MATs**
Methionine binding domain:				
E55	**R91**	EcMAT: H-bond between Met NH_2 and sidechain COO^- MjMAT: No interaction	100%	94%
Q98	**H55**	H-bond between Met COO^- and sidechain amide/imidazole	100%	100%
G117	**N162**	EcMAT: No interaction MjMAT: H-bond between COO^- and sidechain amide	100%	100%
D238*****	D285*****	Sidechain COO^- forms H-bond with Met NH_2 and ribose 3'-OH, also electrostatic interaction with K^+	100%	100%
K269	**H318**	EcMAT: H-bond between Met COO^- and sidechain NH_2 MjMAT: No interaction	100%	87%
Triphosphate binding domain:				
D16	D31	Electrostatic interaction with Mg^{2+}	100%	100%
E42	**Q63**	Electrostatic interaction with K^+	100%	33%
D118	D163	Electrostatic interaction with Mg^{2+}	100%	93%
K165	K204	Electrostatic interaction with PP_i	100%	91%
R244	R291	Electrostatic interaction with P_i and PPi	100%	89%
K245	K25	Electrostatic interaction with PP_i	100%	100%

K265	K313	Electrostatic interaction with P_i	100%	89%
D271	**E308**	Electrostatic interaction with Mg^{2+}	100%	89%
Ribose binding domain				
H14	H29	Catalytic histidine residue	100%	100%
D163	D202	H-bond between 2'-OH and sidechain COO^-	100%	100%
Adenine binding domain				
D101	D146	*Ec*MAT: van der Waals interaction, *Mj*MAT: H-bond between adenine N6-NH_2 and sidechain COO^-	100%	100%
I102	**L147**	van der Waals interaction	100%	96%
S186	**A219**	van der Waals interaction	100%	80%
T227	**A264**	*Ec*MAT: water-mediated H-bond between N6-NH_2 and the backbone *Mj*MAT: van der Waals interaction	100%	63%
R229	**V272**	*Ec*MAT: H-bond between N6-NH_2 and the backbone carbonyl *Mj*MAT: van der Waals interaction	77%	33%
F230	**Y273**	pi-pi stacking with adenine	100%	88%
I232	**T275**	van der Waals interaction	100%	100%
I302	I351	van der Waals interaction	99%	96%

***** also part of triphosphate and ribose binding domains.

Table 2: Structural alignment of *Ec*MAT and *Mj*MAT active site residues. The conservation of each *Ec*MAT and *Mj*MAT active site residue among bacterial and archaeal MATs is also reported. *Ec*MAT residues aligned to a different residue in *Mj*MAT are indicated in bold. The *Ec*MAT residues mutated in this study are highlighted in blue.

3.2 Approaches to rationally enhance *Ec*MAT's nucleotide promiscuity

Upon investigating the *Ec*MAT and *Mj*MAT active sites, we devised four potential mutagenesis-based strategies to enhance the nucleotide promiscuity of *Ec*MAT:

1. Mutations in the adenine-binding domain*:* Differences in the residues comprising adenine-binding domains of *Ec*MAT and *Mj*MAT might contribute to the latter stabilizing guanine, cytosine, and uracil better than the former. Thus, we decided to mutate *Ec*MAT residues in this domain with their correspondingly aligned ones from *Mj*MAT. The following factors were considered for choosing which EcMAT residues are to be mutated:

a) The proximity of the residue to the nucleobase: This is because most interactions in this domain are van der Waals forces that are effective only at short range.

b) Differences in sidechain composition and polarity between the residue and its *Mj*MAT counterpart: Extreme differences (like a shift from a hydrophobic to a charged residue) might lead to more significant substrate promiscuity changes.

c) Conservation of the residue and its *Mj*MAT counterpart among bacterial and archaeal MATs: A high conservation percentage indicates that the particular residue is present in most bacterial/archaeal MATs, implying that the residue has gotten naturally selected and might be a significant residue for catalytic activity.

Based on this criterion, six mutations in the adenine-binding domain were shortlisted: **R229V, I102L, T227A, I232T, S186A,** and **F230Y**. The R229V mutation was proposed by Vruta Gupte in her master's thesis [39], while the other ones were chosen by me.

2. Mutations in the methionine-binding domain: The binding of ATP to *Ec*MAT has a 32-fold improvement once methionine is already bound to the enzyme.[36] Thus, mutations that lead to stronger methionine binding might indirectly improve the binding of all NTPs (including GTP, CTP, and UTP) to *Ec*MAT. *Mj*MAT's N162, which forms hydrogen bonds with the methionine carboxylate, is aligned to G117 of *Ec*MAT. Thus the mutation **G117N** was decided to introduce an extra methionine-binding interaction. The mutation **Q98H** was also devised by replacing the *Ec*MAT residue with the

corresponding *Mj*MAT one. Both mutations were also proposed by Vruta in her master's thesis.

Comparing the active sites of both enzymes yields two other potential mutations in this domain – E55R and K269H. However, the *Mj*MAT residues aligned to E55 and K269 do not interact with methionine. Consequently, introducing these mutations would probably lead to a loss in methionine binding interactions rather than improve them. Hence, these mutations were not explored further.

3. Mutations in the ribose-binding domain: The ribose moiety is a 5-membered ring to which the 2' and 3' hydroxyl and adenine groups are attached. As a result of this cyclic structure, changes in the hydroxyl group's position (achieved by mutating the aspartate residues in the ribose-binding domain) would translate to slight changes in the nucleobase's position. This change might lead to the adenine attaining a slightly altered binding pocket that also accommodates other nucleobases. In this way, modifications in the ribose-binding domain may enhance nucleotide promiscuity.

To test this hypothesis, we decided to mutate *Ec*MAT's D238, which interacts with the substrate methionine as well as the 3’-hydroxyl group of the ribose moiety. Two mutations – **D238E** and **D238S** are proposed. The functional group is maintained in the former case, but the sidechain is enlarged. Conversely, the sidechain size is decreased slightly in the latter case, and the functional group has changed from carboxylate to hydroxyl.

4. Mutations in the triphosphate-binding domain*:* Introducing more positively charged residues in this region might improve triphosphate binding, thereby enhancing the binding of all NTPs with *Ec*MAT. Thus, such a mutant would theoretically show improved activity with ATP but also GTP, CTP, and UTP. However, this strategy was rejected because modifications in this region may also deteriorate the binding of the metal cofactors, which might directly inactivate the enzyme.

Thus 10 *Ec*MAT point mutants were planned in total, among which all except two (F230Y and D238E) were successfully generated by site-directed mutagenesis.

3.3 Purification of WT and mutant *Ec*MATs

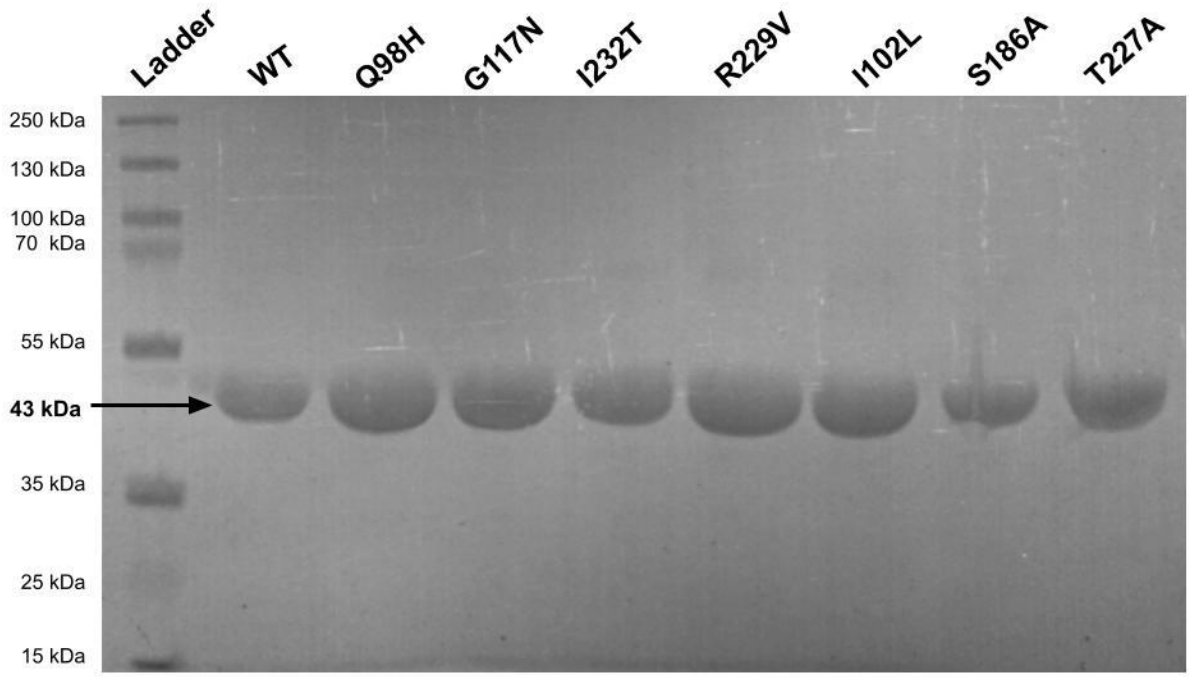


Figure 14: SDS-PAGE gel exhibiting all the purified *Ec*MATs. From left to right: wild-type (WT), Q98H, G117N, I232T, R229V, I102L, S186A, and T227A. The molecular weight of the *Ec*MAT monomer is around 43 kDa.

D238S *Ec*MAT could not be successfully extracted. The protein precipitated during the desalting step of the purification. D238 was one of the residues interacting with the monovalent metal cofactor (usually K^+). Modifying this residue into serine most probably hampered the binding of the metal ion, which is necessary for the protein to fold correctly. Consequently, the protein denatured due to the sudden drop in salt concentration during the desalting protocol.

3.4 *In vitro* characterization of *Ec*MAT mutants

24-hour reactions of each purified *Ec*MAT mutant with ATP, GTP, CTP, and UTP were performed and then analyzed on HPLC and LC-MS to investigate changes in nucleotide promiscuity caused by the mutation.

3.4.1 Mutations in the methionine-binding domain – Q98H and G117N

Despite being less efficient than the WT *Ec*MAT, both Q98H and G117N mutants can catalyze reactions with ATP to form SAM (Figure 15). The ATP activity of the Q98H

mutant is more than twice that of the G117N mutant (94 % versus 46% of WT *Ec*MAT's activity). (Fig 19). The formation of SAM by both mutants was confirmed by LC-MS.

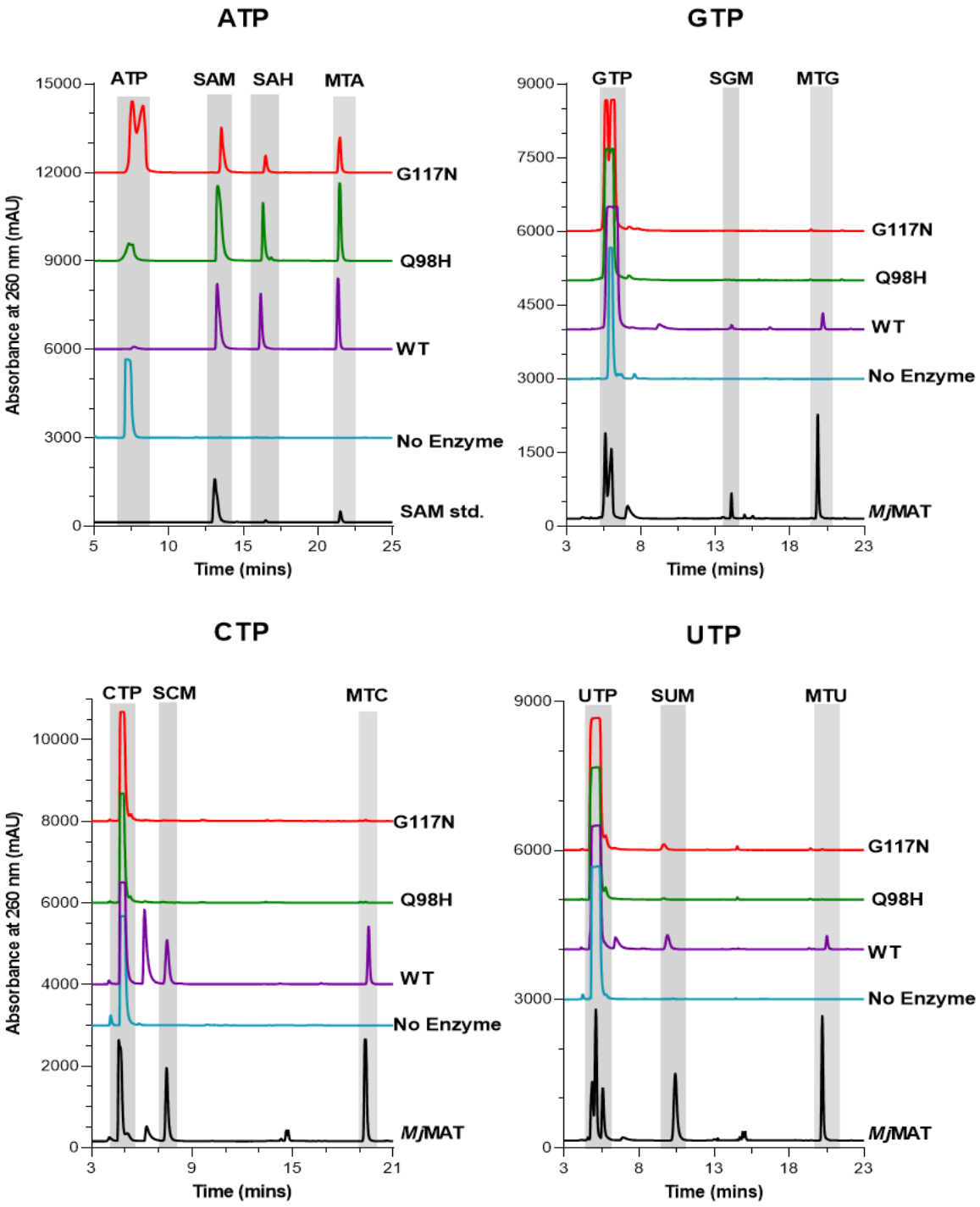


Figure 15: HPLC traces of reactions of NTPs with WT, Q98H, and G117N *Ec*MAT. In the case of GTP, CTP, and UTP, *Mj*MAT reactions with them were run as standards. 10 µL of the reaction mixture was injected for all reactions involving ATP or *Mj*MAT. 50 µL was injected for the rest of the reactions. Peaks for NTP, SNM, SNH, and MTN have been highlighted in grey.

Interestingly, both mutants have become even more ATP-specific than WT *Ec*MAT. Catalytic activity with GTP, CTP, and UTP is not detectable on HPLC for the Q98H mutant. However, LC-MS of these reaction mixtures revealed that this mutant shows trace GTP activity. HPLC of G117N *Ec*MAT reactions with GTP, CTP, and UTP suggested slight SNM formation only with UTP. LC-MS confirmed this UTP activity and revealed the mutant to also react scantily with CTP to form SCM.

As mentioned in Section 3.1.1, *Ec*MAT's Q98 residue forms hydrogen bonds with the carboxylate moiety of methionine using its sidechain amide. Thus, the shift from glutamine to histidine in Q98H *Ec*MAT would have slightly altered this interaction, resulting in the methionine attaining a slightly different conformation which might hamper its S_N2 attack on the ribose 5'-center. As a result, this mutant shows a drop in activity for all four NTPs.

It is unknown whether introducing an additional methionine-binding residue by replacing G117 with asparagine improved methionine binding. The sidechain amide might interact with the protein backbone, morphing the binding region and weakening methionine binding, which would explain the drastic drop in ATP activity and reduced nucleotide promiscuity. On the other hand, N117 may also improve methionine binding but alter its conformation, inhibiting the reaction. Isothermal calorimetry (ITC) of G117N *Ec*MAT – methionine complexes would shed more light on that aspect.

Another factor to consider is the difference between glycine and asparagine sidechains. The former's sidechain consists of a single hydrogen atom. On the other hand, asparagine's sidechain contains a two-carbon chain capped with an amide group. Homology modeling of G117N *Ec*MAT predicts the sidechain of N117 to be pointed toward the binding pocket, which would reduce the overall size of the active site. As a result, the binding of purine nucleotides might get inhibited more than pyrimidine nucleotides. This hypothesis helps explain why G117N *Ec*MAT retains some activity with CTP and UTP while being GTP-inactive and showing a considerable loss in ATP activity.

These results indicate that modifications in the methionine binding domain can translate to significant changes in NTP activity. However, replacing residues with those of the NTP-promiscuous *Mj*MAT did not improve nucleotide promiscuity.

3.4.2 Mutations in adenine-binding domain – R229V, S186A, T227A, I232T and I102L.

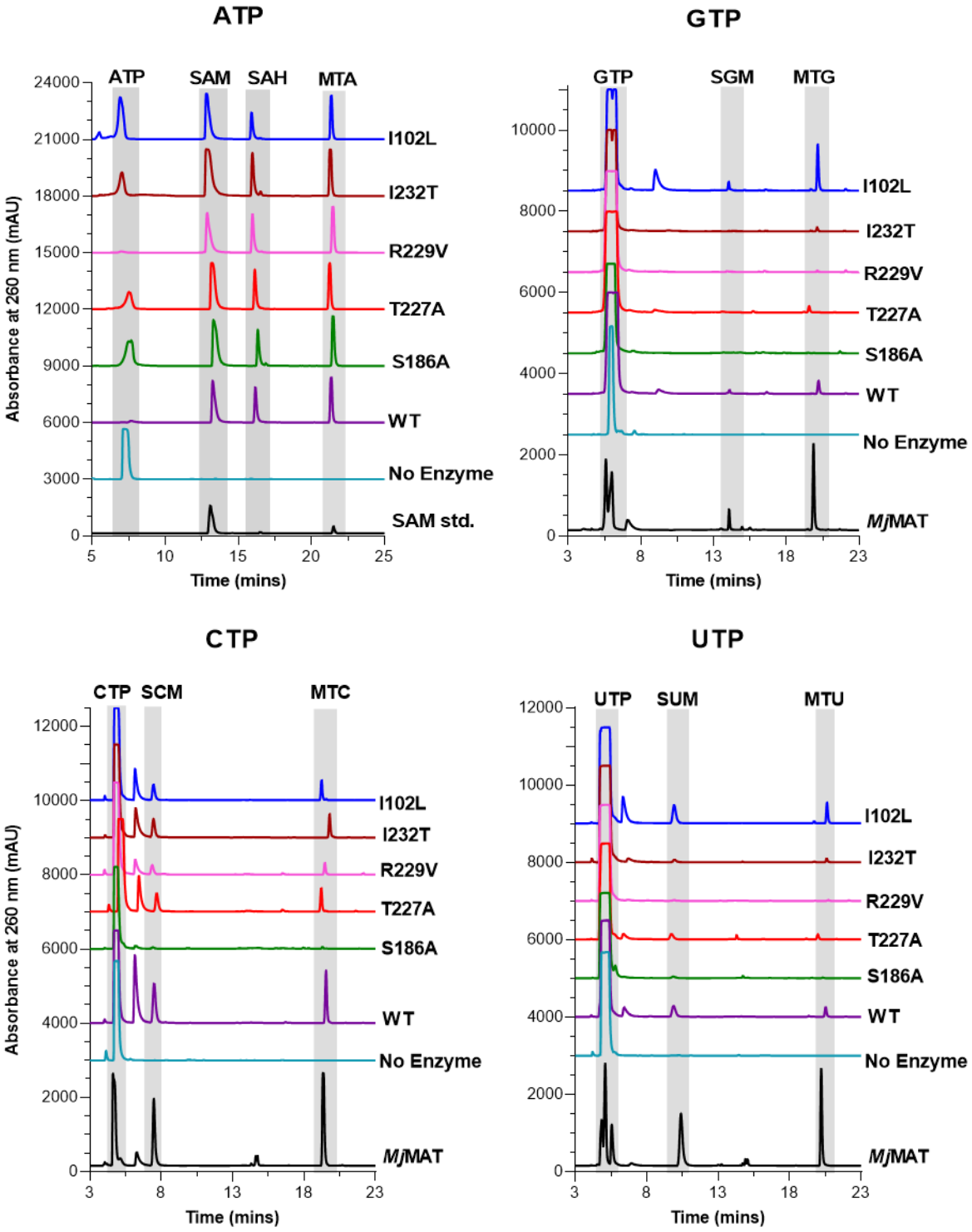


Figure 16: HPLC traces of reactions of NTPs with WT, S186A, T227A, R229V, I232T, and I102L *Ec*MAT. In the case of GTP, CTP, and UTP, *Mj*MAT reactions with them were run as standards. 10 µL of the reaction mixture was injected for all reactions involving ATP or *Mj*MAT. 50 µL was injected for the rest of the reactions. Peaks for NTP, SNM, SNH, and MTN have been highlighted in grey.

All five nucleobase-binding domain mutants exhibit some activity with the four NTPs (Figures 16 and 17), which was confirmed with LC-MS. Among these mutants, the R229V one was the most anticipated, as R229 is the only *Ec*MAT residue to form direct hydrogen bonds with the adenine moiety. Unfortunately, this mutation causes a slight rise (~2%) in ATP conversion but drastic drops (GTP – 93%, CTP – 86%, UTP – 93%) in activities with the other NTPs, making it the most ATP-specific among all the purified mutants. In the *Ec*MAT structure, R229's sidechain is pointed out of the active site, forming electrostatic interactions with D101 from the neighboring peptide chain to close the interfacial cavity. Replacing arginine with valine would remove this interaction, probably causing the nucleobase binding cavity to expand slightly, which might weaken specific van der Waals interactions that stabilize the smaller pyrimidine nucleobases, resulting in a deterioration in CTP and UTP activity. On the other hand, the drop in GTP activity might be due to the absence of a long-range electrostatic interaction between R229's guanidinium moiety and the oxygen atom of guanine's C4-carbonyl group.

With more than a 90% drop in SGM, SCM, and SUM formation, the S186A mutation also enhanced ATP specificity rather than nucleotide promiscuity. The S186A mutant also exhibits a 16% drop in ATP activity. The crystal structure shows S186 to be beside the C2 center of the adenine ring. Consequently, the loss in activity with the three non-cognate NTPs suggests that S186's sidechain hydroxyl group assists in GTP, CTP, and UTP binding by forming hydrogen bonds with the C2-carbonyl of cytosine and uracil and the C2-amine of guanine. The drop in ATP catalysis being less severe than other NTPs might be because adenine does not possess a hydrogen bonding moiety at the C2 position.

A similar hypothesis may explain the 8%, 73%, 68%, and 60% reduction in ATP, GTP, CTP, and UTP activity caused by the T227A mutation. The sidechain hydroxyl of T227 aids in adenine binding by forming a water-mediated hydrogen bond with its C6-amine of adenine. In the case of GTP, UTP, and CTP, a hydrogen bond might form with the C6-carbonyl, C4-carbonyl, and C4-amine of the respective nucleobases. The drastic drop in their activity (compared to ATP's) suggests they form a direct hydrogen bond with T227A's sidechain rather than a water-mediated one.

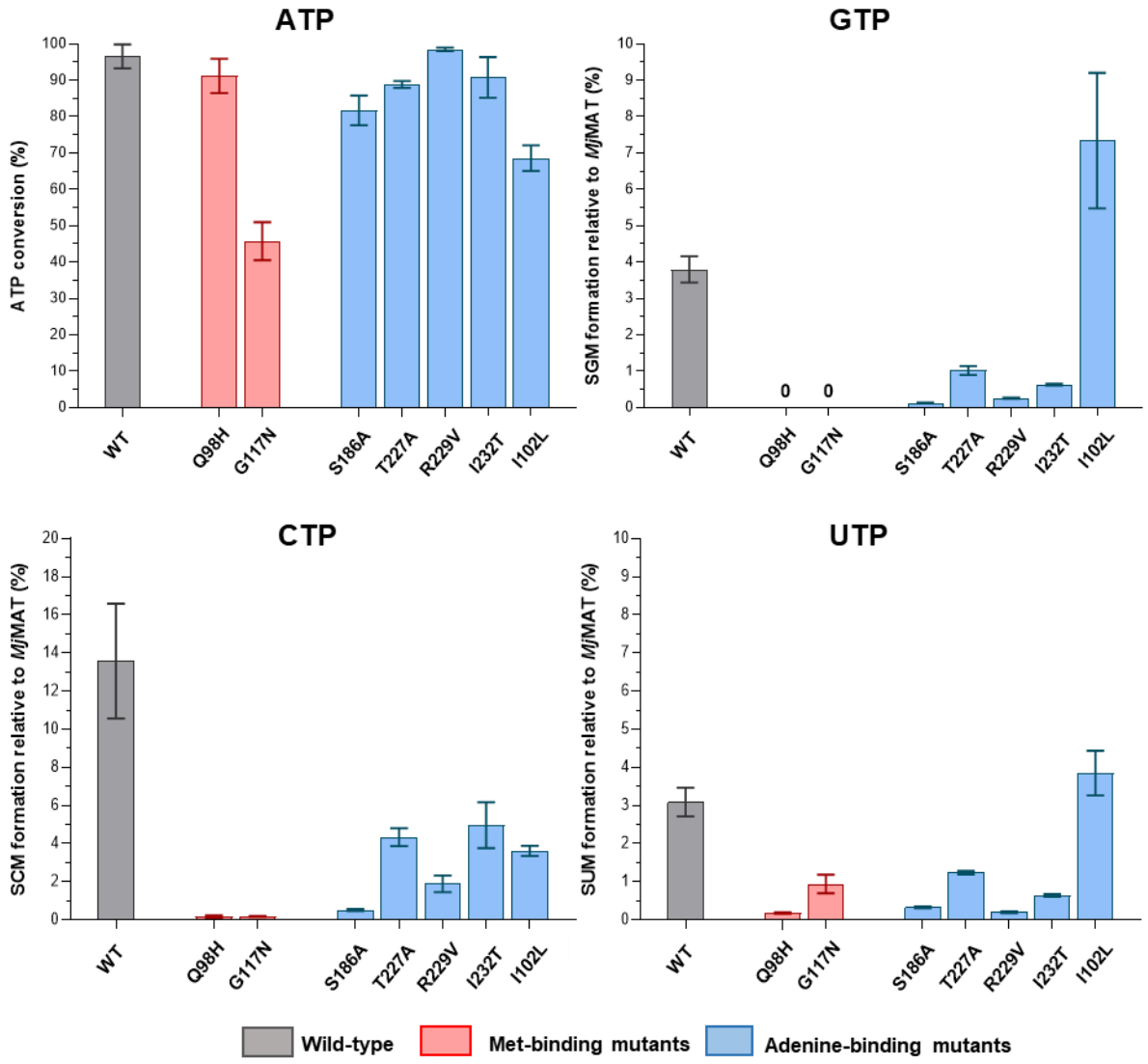


Figure 17: Comparison of *in vitro* ATP, GTP, CTP, and UTP activities between wild-type and mutant *Ec*MATs. ATP activities of all enzymes are obtained by quantifying the loss of ATP throughout the reaction. Activities with other NTPs are measured by comparing the formation of SNMs relative to WT *Mj*MAT. All quantification is performed using HPLC. The mean and standard deviation for three technical replicates (n=3) is plotted.

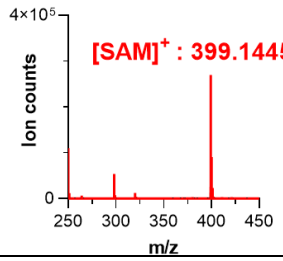
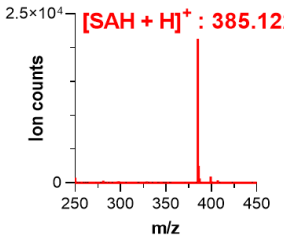
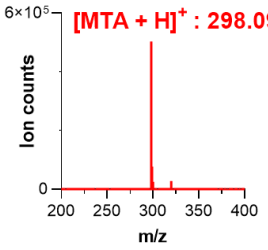
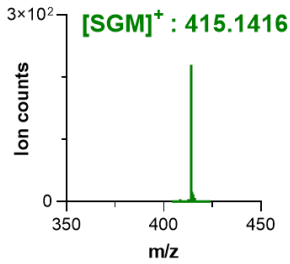
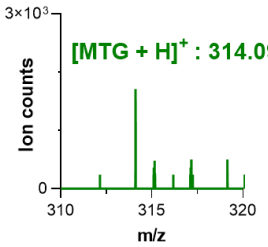
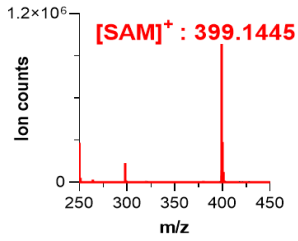
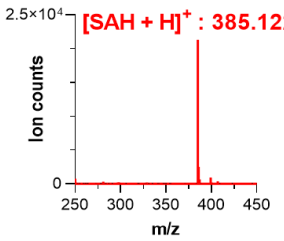
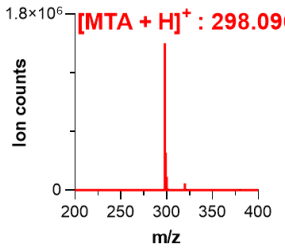
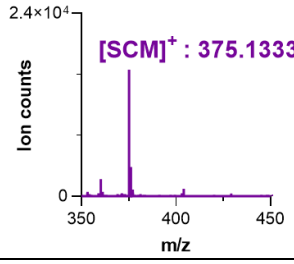
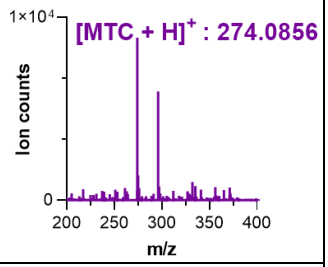
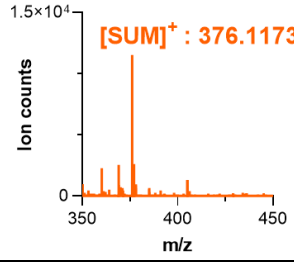
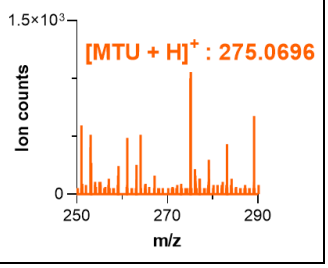
Unlike T227A and S186A, the I232T mutation introduces a hydroxyl group into the active site instead of removing one. However, this mutation also caused a reduction in ATP, GTP, CTP, and UTP conversion by 6%, 83%, 63%, and 80%, respectively. The reason behind these activity drops is challenging to hypothesize, as I232 only forms a van der Waals contact with the nucleobase through its backbone. One potential explanation might be the inhibition of hydrogen bonding between the sidechain hydroxyl of S99 (from the neighboring peptide) and T232's backbone by its

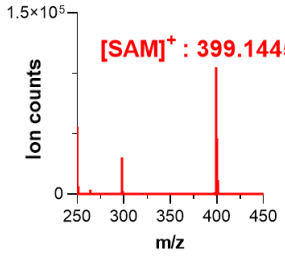
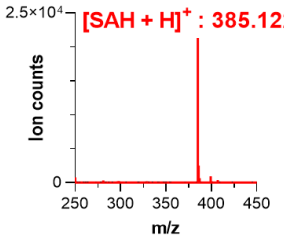
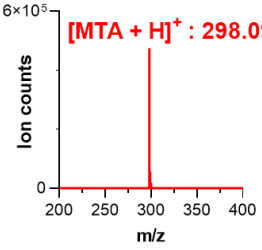
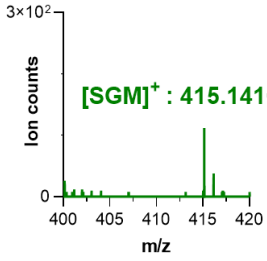
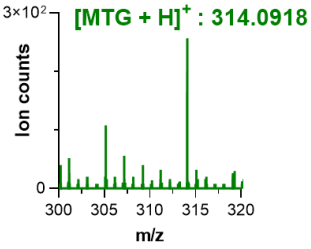
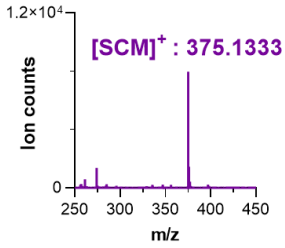
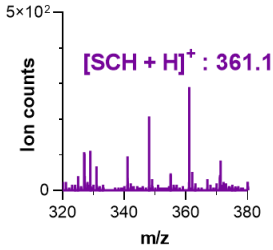
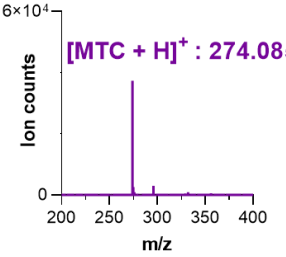
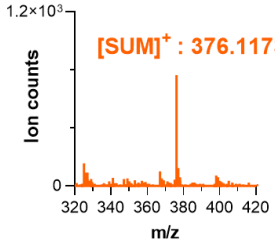
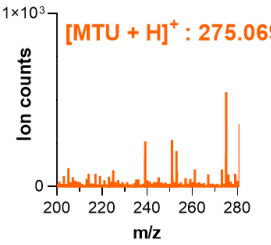
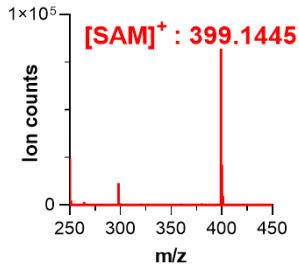
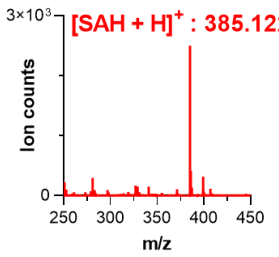
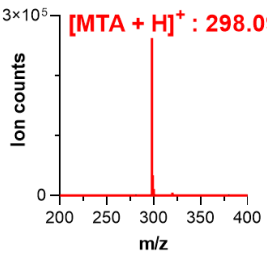
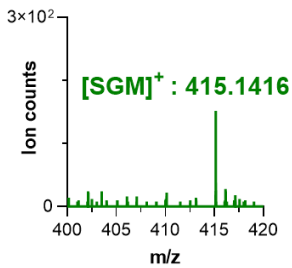
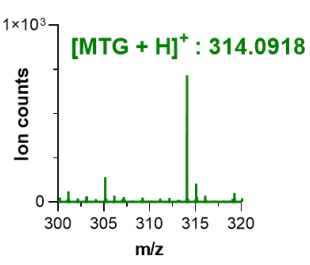
sidechain hydroxyl group, which slightly opens up the interfacial binding site. Consequently, the van der Waals interaction between T232's sidechain and the nucleobase is lower, resulting in weaker binding.

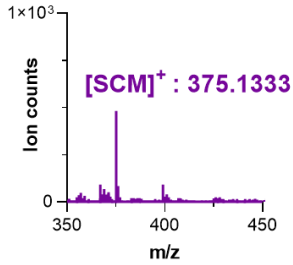
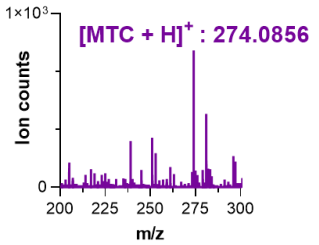
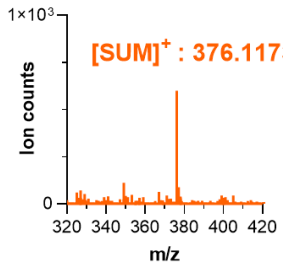
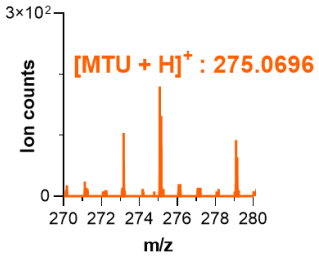
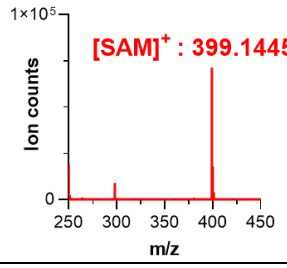
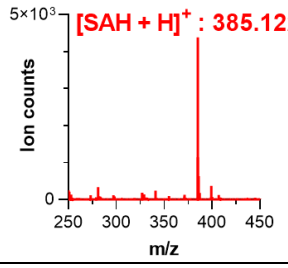
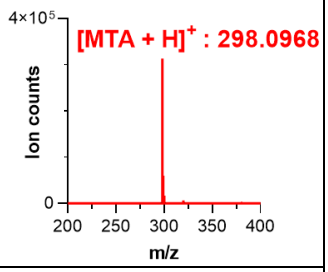
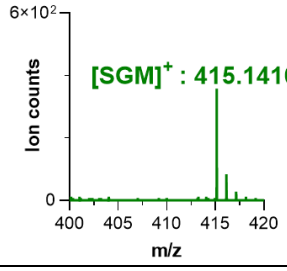
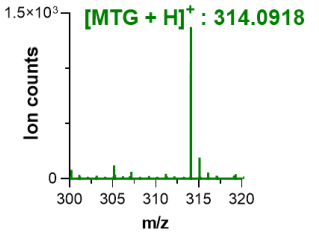
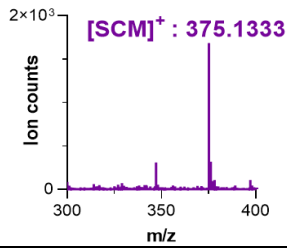
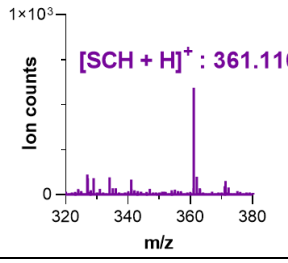
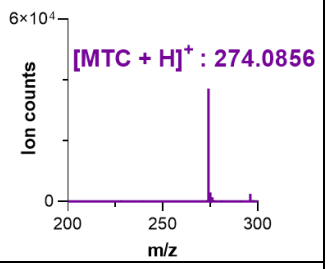
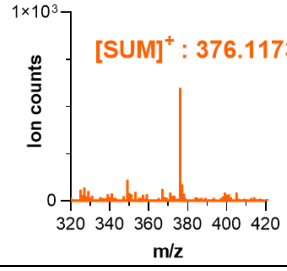
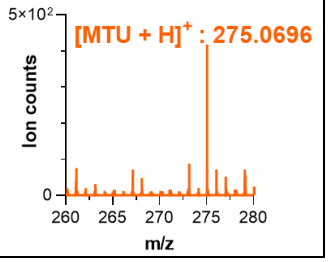
Among all purified *Ec*MAT variants, I102L *Ec*MAT is the most promising one in terms of nucleotide promiscuity. This mutation caused a 29% and 73% reduction in ATP and CTP conversion but also enhanced activity with GTP and UTP by 98% and 25%, respectively. In other words, the shift from isoleucine to leucine improves the binding of C4/C6 carbonyl-containing nucleotides like GTP and UTP at the expense of that of C4/C6 amine-containing nucleotides like ATP and CTP. It might be counterintuitive that the mutation with the slightest sidechain difference yielded the most drastic shifts in promiscuity. However, the proximity of I102 to the nucleobase (the terminal carbon of I102L's sidechain is only 3.4 Å from adenine's N7 center) is such that variations in van der Waals interactions caused due to shifting from isoleucine to leucine can result in significant changes in nucleotide binding.

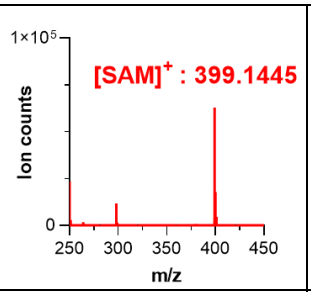
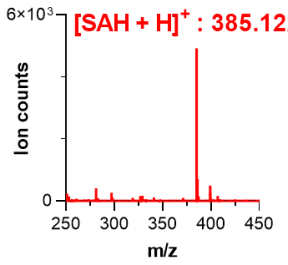
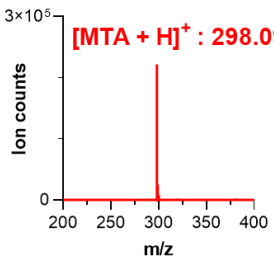
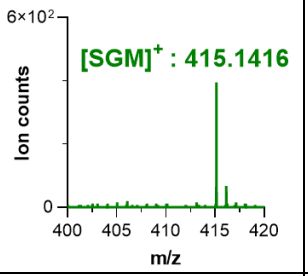
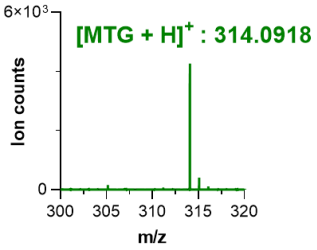
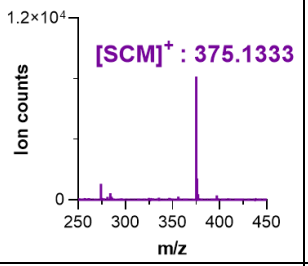
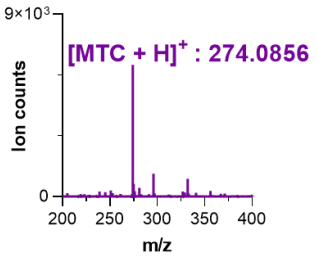
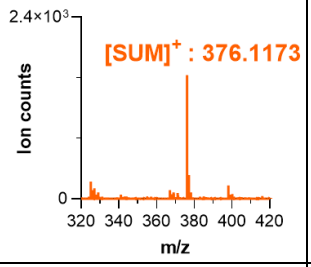
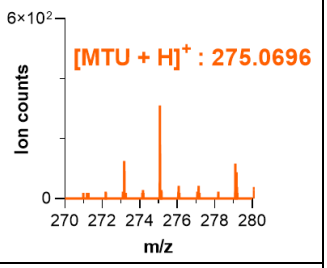
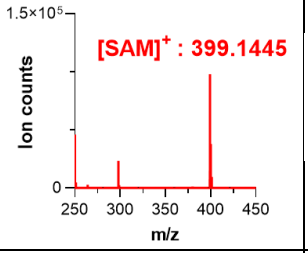
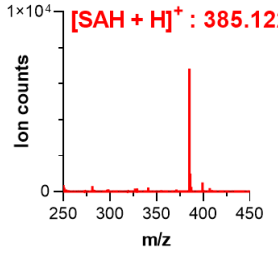
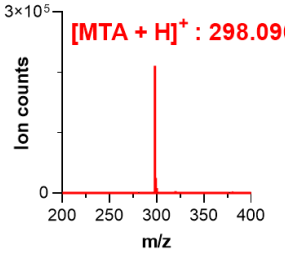
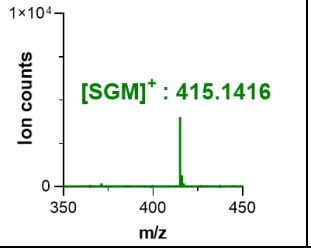
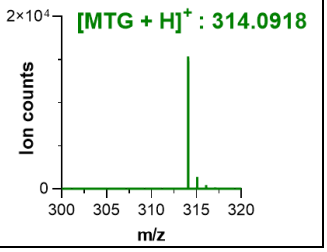
Table 3 reports MS traces for SNMs and their degradation products obtained from the LC-MS of *Ec*MAT mutant-NTP reaction mixtures. As mentioned in Section 3.4.1, LC-MS of Q98H and G117N *Ec*MAT reaction mixtures revealed additional activities of these mutants with specific non-cognate NTPs that were undetectable on HPLC. However, in the case of other mutants, the nucleotide promiscuity results obtained via LCMS are consistent with those obtained from HPLC.

Overall, the *in vitro* characterization of rationally engineered *Ec*MAT mutants mainly yielded disappointing results, with only I102L *Ec*MAT exhibiting minor improvements in GTP and UTP promiscuity. On the flip side, considering all mutations were made in the active site, it is remarkable that none of the mutants completely lost their primary activity with ATP, highlighting *Ec*MAT's robustness to mutagenesis. The results also suggest potential interactions between *Ec*MAT and non-cognate NTPs (like the hydrogen bond with S186's sidechain) that are absent in the ATP/SAM-bound enzyme crystal structure, thereby providing a rough picture of how the non-cognate NTPs bind to the enzyme and also assisting future mutagenesis studies to obtain a nucleotide-promiscuous *Ec*MAT.

Ec*MAT* mutant	NTP	MS trace		
		SNM	SNH	MTN
Q98H	ATP			
	GTP		Not detected	
	CTP	Not detected	Not detected	Not detected
	UTP	Not detected	Not detected	Not detected
G117N	ATP			
	GTP	Not detected	Not detected	Not detected
	CTP		Not detected	
	UTP		Not detected	

R229V	ATP			
	GTP		Not detected	
	CTP			
	UTP		Not detected	
S186A	ATP			
	GTP		Not detected	

	CTP		Not detected	
	UTP		Not detected	
T227A	ATP			
	GTP		Not detected	
	CTP			
	UTP		Not detected	

I232T	ATP			
	GTP		Not detected	
	CTP		Not detected	
	UTP		Not detected	
I102L	ATP			
	GTP		Not detected	

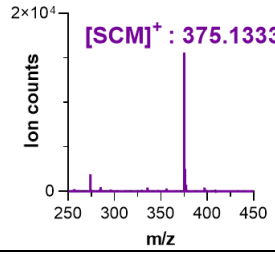
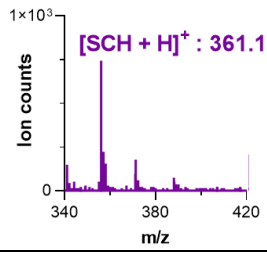
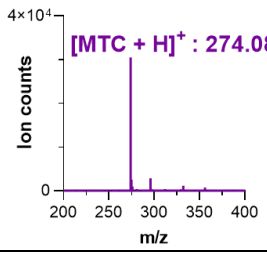
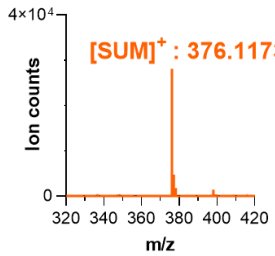
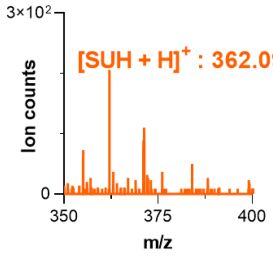
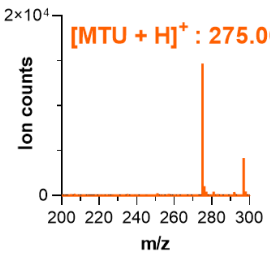
	CTP			
	UTP			

Table 3: Detection of SNM formation in *Ec*MAT mutant-NTP reaction mixtures using LC-MS. The MS traces for detected SNMs, SNHs, and MTNs are reported and color-coded based on the nucleobase moiety (adenine - red, guanine - green, cytosine - violet, uracil – orange).

3.5 *In vivo* production of SAM by *Ec*MAT mutants

As all the *Ec*MAT mutants retained a considerable fraction of the wild-type enzyme's *in vitro* ATP activity, we were curious whether these mutants could also function inside the *E. coli* microorganism. We also wanted to check whether the *in vitro* GTP and UTP promiscuity enhancement in I102L *Ec*MAT translates to increased SGM and SUM production in a cellular context. Lastly, we were also interested in investigating the *in vivo* nucleotide promiscuity of wt *Mj*MAT because that has not yet been explored in the literature.

Consequently, we heterologously expressed *Mj*MAT, *Ec*MAT, and the seven mutants in *E. coli* BL21(DE3). WT and all mutant *Ec*MAT genes are present in a pET15b plasmid, while the WT *Mj*MAT gene is in a pET19b vector. Both these vectors possess similar copy numbers in *E coli.* Consequently, comparisons could be made between the *in vivo* activities of each enzyme. As a control, BL21(DE3) cells containing the pET15b empty vector were also grown to account for SAM production by the genomic MAT. Cells harvested from all these cultures were lysed, followed by LC-MS of the lysate to check for the cellular production of SNMs.

SAM and MTA were successfully detected in the lysates of all the BL21(DE3) cultures (Figure 18A). The EIC of SAM obtained from the cell extracts of T227A and WT

*Ec*MAT- expressing BL21(DE3) show an abnormal peak with an earlier retention time compared to the standard, probably due to SAM getting stuck on minute cellular debris, which has persisted in the sample despite filtration. For the rest of the cell lysates, clean EIC peaks were obtained for both SAM and MTA.

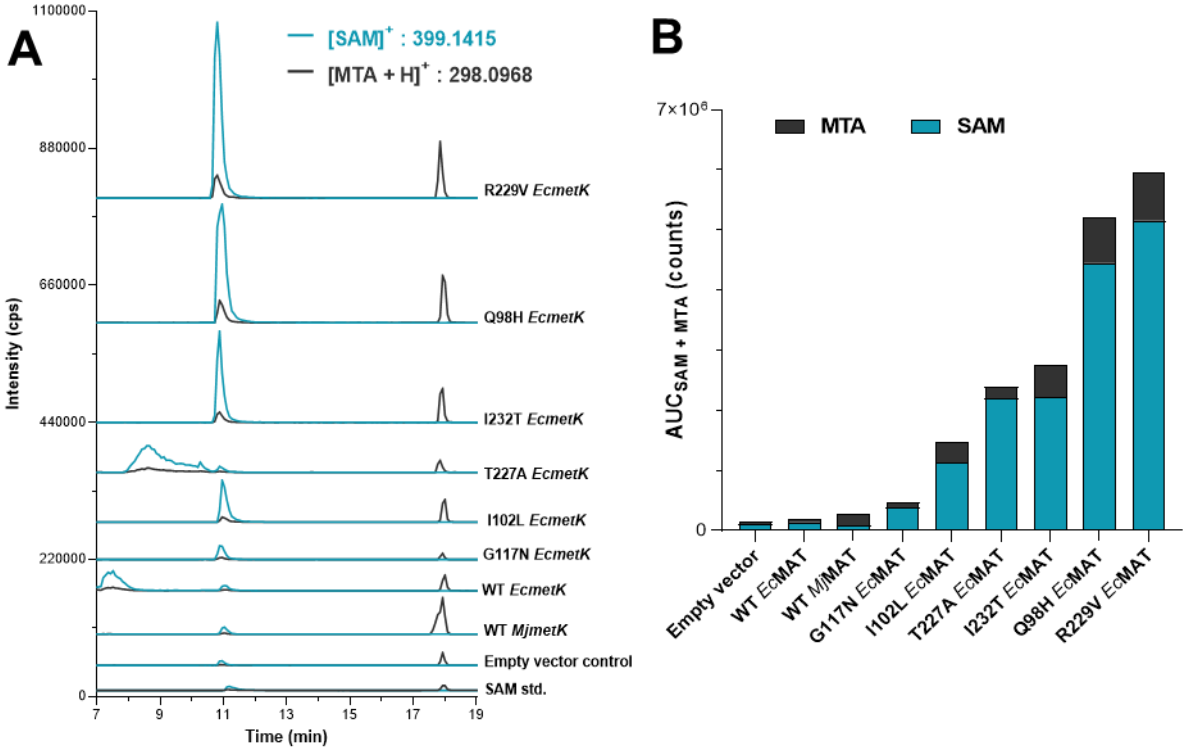


Figure 18: Production of SAM inside *E. coli* by heterologously expressed MATs. (A) EICs corresponding to $[SAM]^+$ (blue) and $[MTA+H]^+$ (black) from the LC-MS analysis of BL21(DE3) cells heterologously expressing different MATs. **(B)** The sum of AUCs of SAM and MTA peaks from the EICs shown in **(A)**.

The amount of SAM detected in the empty vector control was lower than all the others, indicating that all the *Ec*MAT mutants exhibit SAM formation inside the cell. (Figure 18B) The AUC values for SAM and MTA were then compared between mutants to obtain a relative order of *in vivo* ATP activity (Figure 16B). R229V *Ec*MAT displayed the highest activity inside the cell, followed by Q98H, I232T, T227A, I102L, and finally, G117N *Ec*MAT.

Interestingly, the mutants also follow this very trend for *in vitro* ATP activity (Figure 15), suggesting that these mutations affect *Ec*MAT's activity inside the cell in the same way as they do *in vitro*. However, the difference between mutants' SAM production abilities seems more significant *in vivo* than *in vitro*. For example, the least ATP-active mutant - G117N, is 53% less active *in vitro* but 92% less active *in vivo* than R229V -

the most ATP-active mutant. This observation might be arising due to the better growth of BL21(DE3) cells expressing more ATP-efficient mutants, creating a positive feedback loop that would result in significantly higher amounts of SAM being detected for the same culture volume. Studying the growth of these heterologous MAT-expressing *E coli* over time would provide further insight and help substantiate this hypothesis.

Surprisingly, WT *Mj*MAT and *Ec*MAT showed lower cellular SAM formation than all the *Ec*MAT mutants. The thermophilic nature of the former could explain its lack of activity inside the mesophilic *E. coli,* which were grown at 37 °C. The lack of SAM and MTA detected in the latter case might be due to inefficient cell lysis during the metabolite extraction.

Despite showing *in vitro* nucleotide promiscuity, *Mj*MAT does not seem to form detectable amounts of other SNMs inside *E. coli,* which might be due to a) the enzyme being thermophilic and b) concentrations of GTP, CTP, and UTP inside the microorganism (1.67, 0.32, and 0.67 mM respectively during mid-log phase)[55] being lower than the one (10 mM) used during the *in vitro* assays. The latter reason also explains the lack of *in vivo* SGM and SUM formation by I102L *Ec*MAT, despite its enhanced *in vitro* promiscuity towards UTP and GTP. WT and the other mutant *Ec*MATs also do not produce detectable amounts of SGM, SCM, or SUM, which is unsurprising considering their *in vitro* nucleotide promiscuity is much lower than *Mj*MAT's.

Finally, it must be mentioned that the MAT heterologous expression experiment has been performed only once and thus must be repeated at least twice more to confirm the robustness of the results.

3.6 I102L *Ec*MAT seems to show GTP activity inside MOB1490 cells

The presence of the genomic MAT inside BL21(DE3) cells might be causing I102L *Ec*MAT not to get expressed enough to make detectable amounts of SGM and SUM. Hence we wished to introduce this mutant inside an *E. coli metK* (the gene encoding for MAT) knockout strain. However, due to SAM's critical role in metabolism, *metK* is an essential gene for the organism's survival.[56] Consequently, constructing a Δ*metK* variant is exceptionally challenging.

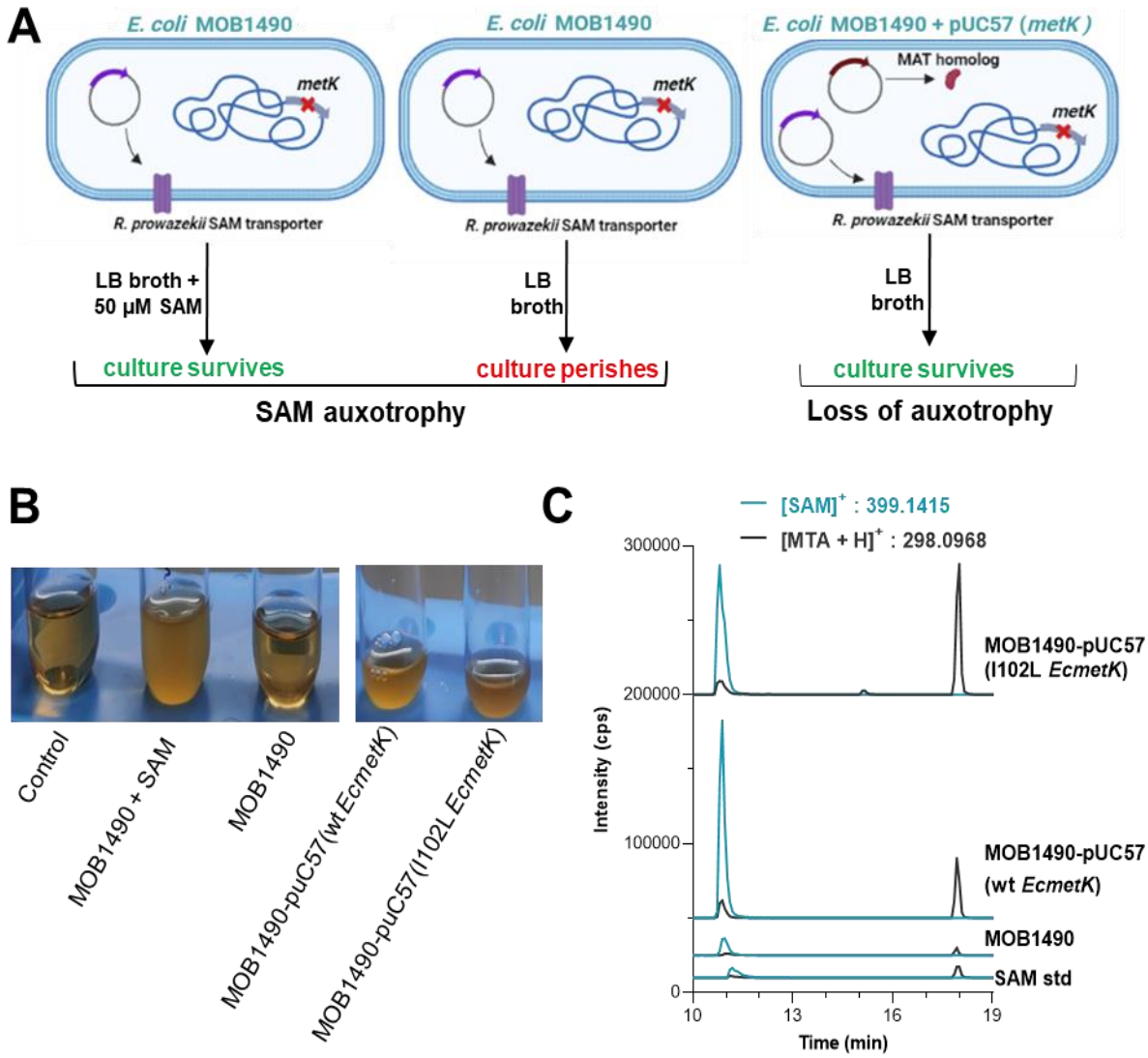


Figure 19: *E. coli* MOB1490 strain is a SAM auxotroph. (A) Parental MOB1490 cells require exogenous SAM for survival. Upon complementation with a *metK* gene, MOB1490 ceases to behave as a SAM auxotroph. **(B)** MOB1490 cultures grown in LB broth containing ampicillin, kanamycin, rifampicin, and l-arabinose. From left to right: No inoculation control, MOB1490 with 50 µM exogenous SAM, MOB1490, MOB1490-pUC57(WT *EcmetK),* MOB1490-pUC57(I102L *EcmetK)*. **(C)** EICs corresponding to $[SAM]^+$ and $[MTA+H]^+$ from the LC-MS analysis of parental, WT *EcmetK*-complemented, and I102L *EcmetK*-complemented MOB1490 cell lysates.

In 2004, Dr. David Wood's group at the University of South Alabama successfully developed such a strain by first incorporating a plasmid encoding the SAM transporter from *Rickettsia prowazekii* into the microbe before removing the *metK* gene.[57] This strain, called MOB1490, can thus survive despite lacking an in-house MAT by uptaking SAM from its growth medium. In other words, it is a SAM auxotroph. However,

complementing parental MOB1490 with an external *metK* gene would lead to a loss of auxotrophy, i.e., the cells would longer need exogenous SAM for survival (Figure 19A).

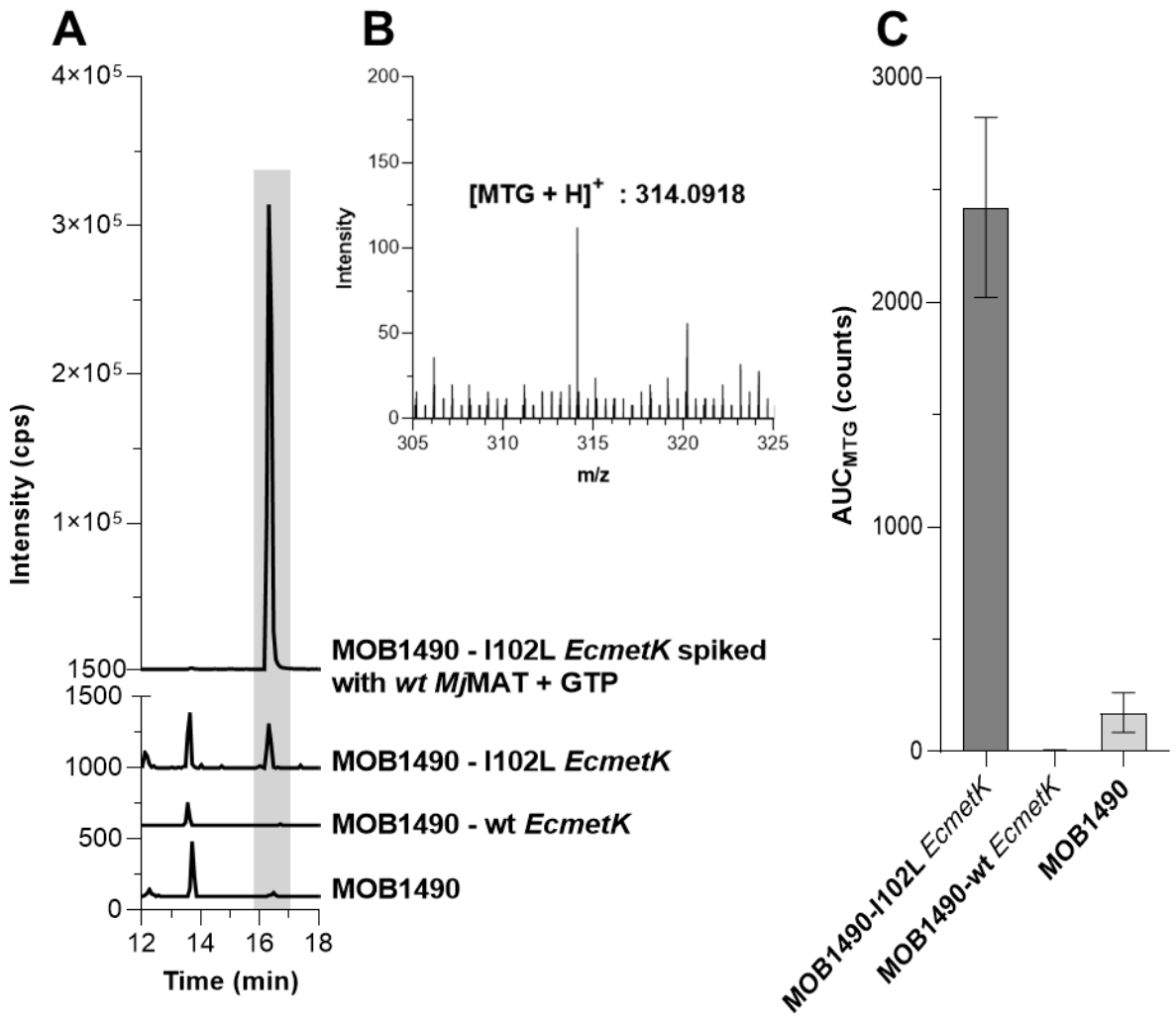


Figure 20: MTG detection in I102L *EcmetK*-complemented MOB1490. (A) EICs corresponding to the mass of $[MTG+H]^+$ (314.0918) from the LC-MS analysis of parental, WT *EcmetK*-complemented, and I102L *EcmetK*-complemented MOB1490 cell lysates, and the latter spiked with *Mj*MAT-GTP reaction mixture. The peak corresponding to MTG is highlighted in grey. **(B)** MS trace of I102L *EcmetK*-complemented MOB1490 cell extract showing peak for MTG. **(C)** AUC of the MTG peak from the EICs shown in **(A)**. The mean and standard deviation for three biological replicates (n=3) is plotted.

Parental and WT *EcmetK*-complemented MOB1490 strains were obtained from Dr. Robert Blumenthal's lab at the University of Toledo. The growth phenotypes of both strains were first verified. LB culture of MOB1490 does not survive without adding SAM

into the medium. However, MOB1490 harboring pUC57(WT *EcmetK*) grows without exogenous SAM in the medium (Figure 19B). We then performed LC-MS of their cell lysates (Figure 19C). The EIC peaks of SAM and MTA were shorter in uncomplemented MOB1490 than in the *EcmetK*-complemented one, which indicates that the former uptakes a bare minimum amount of exogenous SAM required for its sustenance. On the other hand, the latter possesses an in-house MAT and thus synthesizes much higher amounts of SAM, leading to higher peaks. Thus, SAM uptake in parental MOB1490, and SAM formation in *metK*-complemented MOB1490 were respectively confirmed.

We then complemented I102L *EcmetK* into parental MOB1490 and confirmed the loss in SAM auxotrophy. (Figure 19B, first tube from the right) LC-MS of the cell extracts was subsequently performed to check for SAM, SGM, and SUM formation. SAM formation was lesser in these cells than in the ones complemented with the wild-type enzyme, possibly due to the lower ATP activity exhibited by the I102L mutant. (Figure 19C)

SGM, SUM, and MTU could not be detected in the cell lysate of I102L *EcmetK*-complemented MOB1490. However, the EIC corresponding to the ion mass of MTG showed a peak at the retention time of MTG (Figure 20). This information was confirmed by spiking the cell lysate sample with the *Mj*MAT-GTP reaction mixture. Interestingly, this peak is absent in the cell lysate of MOB1490 harboring wt *EcmetK*, which further validates that this peak is indeed MTG.

The breakdown of SGM remains the only known process of producing MTG inside the cell. Thus it is reasonable to infer that I102L *Ec*MAT is getting sufficiently expressed in MOB1490 to form SGM, which degrades to detectable amounts of MTG inside the microorganism. Thus, after the detection of MTG in human liver cancer cells by Gade et al.[42], this result provides the first evidence of SNMs in a microbial cell.

4. Conclusions and Future Directions

This study attempted the engineering of highly nucleotide-promiscuous EcMAT variants by using a rational approach involving the replacement of the enzyme's active site residues with corresponding ones from *Mj*MAT. However, the results show that despite performing mutations at diverse locations in the *Ec*MAT active site, all but one caused a loss in *in vitro* promiscuity rather than a gain. Moreover, the one mutant with enhanced promiscuity - I102L *Ec*MAT is still practically ATP-specific. Thus, the rational design strategy has yet failed to generate highly promiscuous mutants, which could be due to the following reasons.

A) *Mj*MAT is from a thermophilic organism. Thus, its active site has evolved to exhibit catalytic activity at higher temperatures. Consequently, introducing parts of that active site into a mesophilic enzyme like *Ec*MAT might not induce nucleotide promiscuity in the latter.

B) Residues not located at the active site might be more significant in influencing NTP promiscuity in MATs. Gade et al. proposed this hypothesis to explain why *Hs*MAT2a (2a isoform of human MAT) is NTP-promiscuous while *Ec*MAT is ATP-specific despite both possessing near-identical active sites.[42] The results from the present study hint that the same is probably also valid in the case of *Mj*MAT and *Ec*MAT.

Nevertheless, the systematic development and analysis of *Ec*MAT mutants performed in this study do provide insights that can be applied to engineer more promiscuous MATs in the future. For example, *Ec*MAT's S186 residue seems to form hydrogen bonds with the C2-functional groups of guanine, cytosine, and uracil but not with adenine. Hence, site saturation mutagenesis[58] at that particular position could help yield a mutant strongly binding to non-cognate NTPs leading to nucleotide promiscuity. Saturation mutagenesis could also be executed at the I102 site to see whether mutants more promiscuous than I102L can be obtained. Also, the fact that I102 is part of the *Ec*MAT gating sequence suggests that the gating loop dictates which NTPs can enter the active site. Consequently, trying other rationally designed mutations in the gating sequence, or even swapping the entire loop with that from *Mj*MAT or *Hs*MAT2a, might bring forth exciting shifts in promiscuity. Finally, highly promiscuous MATs could

also be engineered using directed evolution.[59] In such a case, I102L *Ec*MAT could be chosen as the starting enzyme, then mutated randomly. The activity towards non-cognate NTPs of the numerous mutants generated could be screened efficiently on a 96-well plate using a simple colorimetric assay like the malachite green assay.[60] The ones that show enhancement in promiscuity would then be selected for the following round of selection. This process would be repeated till a sufficiently promiscuous mutant has been obtained.

Once a sufficiently *in vitro* promiscuous *Ec*MAT mutant has been engineered, its ability to generate SNMs inside the cell would then have to be determined. This thesis establishes the workflow for testing the activity of the EcMAT mutants inside *E. coli* and also shows the I102L mutant to form SGM inside the organism. Even a weakly promiscuous variant like I102L *Ec*MAT forming detectable amounts of SNMs *in vivo* generates optimism that a more nucleotide promiscuous variant designed in the future could produce significant amounts of SNMs inside the cell.

Lastly, for engineering novel SNM-utilizing metabolic pathways inside the cell, SAM-utilizing enzymes should be able to accept other SNMs as substrates. Thus substrate promiscuity of these enzymes will have to be investigated. A handful of papers on this topic have been published in recent years.

For example, a 2016 study by Ji et al. showed that the radical SAM enzyme NosL which catalyzes the synthesis of 3-methyl-2-indolic acid from tryptophan, could accept non-cognate SNMs, notably SCM and SGM.[61] Another radical SAM enzyme, NosN, which is involved in nosiheptide biosynthesis, was shown to accept SGM by Ding et al. in 2017.[62] Over the last three years, the Rentmeister group at the University of Münster has studied the SNM scope of several MTases by performing MAT-MTase cascade reactions using nucleotide promiscuous MATs that generate the SNMs *in situ*.[27,41] The Laurino group also investigated the promiscuity of two DNA MTases – Taq1 MTase and HhaI MTase towards SUM, SGM, and SCM, respectively.[63]

However, this is just the tip of the iceberg with regard to mapping the SNM scope of all SAM–utilizing enzymes. Notably, the SNM promiscuity of SAM decarboxylase, polyamine biosynthesis enzymes, acyl-homoserine lactone synthases, and vitamin biosynthesis enzymes are yet to be characterized.

Thus, this project has plenty of potential research avenues that could be explored in the future. However, developing a highly NTP-promiscuous *Ec*MAT variant is the foremost hurdle to unlocking those research directions. One hopes that some of the proposed mutagenesis approaches can successfully generate *Ec*MAT mutants that depict exciting shifts in nucleotide promiscuity.

5. References

1. Hult, K.; Berglund, P. Enzyme Promiscuity: Mechanism and Applications. *Trends Biotechnol.* **2007**, *25* (5), 231–238. https://doi.org/10.1016/j.tibtech.2007.03.002.

2. Copley, S. D. Shining a Light on Enzyme Promiscuity. *Curr. Opin. Struct. Biol.* **2017**, *47*, 167–175. https://doi.org/10.1016/j.sbi.2017.11.001.

3. Jackson, R. C.; Handschumacher, R. E. Escherichia Coli L-Asparaginase. Catalytic Activity and Subunit Nature. *Biochemistry* **1970**, *9* (18), 3585–3590. https://doi.org/10.1021/bi00820a013.

4. O'Brien, P. J.; Herschlag, D. Catalytic Promiscuity and the Evolution of New Enzymatic Activities. *Chem. Biol.* **1999**, *6* (4), R91–R105. https://doi.org/10.1016/S1074-5521(99)80033-7.

5. Humble, M. S.; Berglund, P. Biocatalytic Promiscuity. *Eur. J. Org. Chem.* **2011**, *2011* (19), 3391–3401. https://doi.org/10.1002/ejoc.201001664.

6. Fessner, W.-D. Systems Biocatalysis: Development and Engineering of Cell-Free "Artificial Metabolisms" for Preparative Multi-Enzymatic Synthesis. *New Biotechnol.* **2015**, *32* (6), 658–664. https://doi.org/10.1016/j.nbt.2014.11.007.

7. Bilal, M.; Iqbal, H. M. N. Tailoring Multipurpose Biocatalysts via Protein Engineering Approaches: A Review. *Catal. Lett.* **2019**, *149* (8), 2204–2217. https://doi.org/10.1007/s10562-019-02821-8.

8. Valle, A.; Bolívar, J. Chapter 8 - Escherichia Coli, the Workhorse Cell Factory for the Production of Chemicals. In *Microbial Cell Factories Engineering for Production of Biomolecules*; Singh, V., Ed.; Academic Press, 2021; pp 115–137. https://doi.org/10.1016/B978-0-12-821477-0.00012-X.

9. McDonald, A. The Enzyme List Class 1—Oxidoreductases. *Nomencl. Comm. Int. Union Biochem. Mol. Biol. NC-IUBMB* **2019**.

10. Banerjee, R. V.; Frasca, V.; Ballou, D. P.; Matthews, R. G. Participation of Cob(I)Alamin in the Reaction Catalyzed by Methionine Synthase from Escherichia Coli: A Steady-State and Rapid Reaction Kinetic Analysis. *Biochemistry* **1990**, *29* (50), 11101–11109. https://doi.org/10.1021/bi00502a013.

11. Bossemeyer, D. Protein Kinases — Structure and Function. *FEBS Lett.* **1995**, *369* (1), 57–61. https://doi.org/10.1016/0014-5793(95)00580-3.

12. Wang, X.; Feng, Y.; Guo, X.; Wang, Q.; Ning, S.; Li, Q.; Wang, J.; Wang, L.; Zhao, Z. K. Creating Enzymes and Self-Sufficient Cells for Biosynthesis of the

Non-Natural Cofactor Nicotinamide Cytosine Dinucleotide. *Nat. Commun.* **2021**, *12* (1), 2116. https://doi.org/10.1038/s41467-021-22357-z.

13. Shah, A.; Kumar, Y.; Rohan, S.; Hazra, A. B. Efficient Chemical and Enzymatic Syntheses of FAD Nucleobase Analogues and Their Analysis as Enzyme Cofactors. *ChemBioChem* **2023**, *n/a* (n/a), e202300055. https://doi.org/10.1002/cbic.202300055.

14. Mathur, Y.; Vartak, A. R.; Hazra, A. B. Guardian of Cobamide Diversity: Probing the Role of CobT in Lower Ligand Activation in the Biosynthesis of Vitamin B_{12}; and Other Cobamide Cofactors. *Methods Enzymol.* **2022**, *668*, 25—59. https://doi.org/10.1016/bs.mie.2022.01.001.

15. Markham, G. D. S-Adenosylmethionine. In *Encyclopedia of Life Sciences*; John Wiley & Sons, Ltd, 2010. https://doi.org/10.1002/9780470015902.a0000662.pub2.

16. Fontecave, M.; Atta, M.; Mulliez, E. S-Adenosylmethionine: Nothing Goes to Waste. *Trends Biochem. Sci.* **2004**, *29* (5), 243–249. https://doi.org/10.1016/j.tibs.2004.03.007.

17. Loenen, W. A. M. S-Adenosylmethionine: Jack of All Trades and Master of Everything? *Biochem. Soc. Trans.* **2006**, *34* (2), 330. https://doi.org/10.1042/BST20060330.

18. Abdelraheem, E.; Thair, B.; Varela, R. F.; Jockmann, E.; Popadić, D.; Hailes, H. C.; Ward, J. M.; Iribarren, A. M.; Lewkowicz, E. S.; Andexer, J. N.; Hagedoorn, P.-L.; Hanefeld, U. Methyltransferases: Functions and Applications. *ChemBioChem* **2022**, *23* (18), e202200212. https://doi.org/10.1002/cbic.202200212.

19. Broderick, J. B.; Duffus, B. R.; Duschene, K. S.; Shepard, E. M. Radical S-Adenosylmethionine Enzymes. *Chem. Rev.* **2014**, *114* (8), 4229–4317. https://doi.org/10.1021/cr4004709.

20. Cantoni, G. L. S-ADENOSYLMETHIONINE; A NEW INTERMEDIATE FORMED ENZYMATICALLY FROM l-METHIONINE AND ADENOSINETRIPHOSPHATE. *J. Biol. Chem.* **1953**, *204* (1), 403–416. https://doi.org/10.1016/S0021-9258(18)66148-4.

21. Markham, G. D.; Pajares, M. A. Structure-Function Relationships in Methionine Adenosyltransferases. *Cell. Mol. Life Sci. CMLS* **2009**, *66* (4), 636–648. https://doi.org/10.1007/s00018-008-8516-1.

22. Goyvaerts, V.; Van Snick, S.; D'Huys, L.; Vitale, R.; Helmer Lauer, M.; Wang, S.; Leen, V.; Dehaen, W.; Hofkens, J. Fluorescent SAM Analogues for Methyltransferase Based DNA Labeling. *Chem Commun* **2020**, *56* (22), 3317–3320. https://doi.org/10.1039/C9CC08938A.

23. Townsend, A. P.; Roth, S.; Williams, H. E. L.; Stylianou, E.; Thomas, N. R. New S-Adenosyl-l-Methionine Analogues: Synthesis and Reactivity Studies. *Org. Lett.* **2009**, *11* (14), 2976–2979. https://doi.org/10.1021/ol9009859.

24. Peters, A.; Herrmann, E.; Cornelissen, N. V.; Klöcker, N.; Kümmel, D.; Rentmeister, A. Visible-Light Removable Photocaging Groups Accepted by MjMAT Variant: Structural Basis and Compatibility with DNA and RNA Methyltransferases. *Chembiochem Eur. J. Chem. Biol.* **2022**, *23* (1), e202100437. https://doi.org/10.1002/cbic.202100437.

25. Islam, K.; Chen, Y.; Wu, H.; Bothwell, I. R.; Blum, G. J.; Zeng, H.; Dong, A.; Zheng, W.; Min, J.; Deng, H.; Luo, M. Defining Efficient Enzyme–Cofactor Pairs for Bioorthogonal Profiling of Protein Methylation. *Proc. Natl. Acad. Sci.* **2013**, *110* (42), 16778–16783. https://doi.org/10.1073/pnas.1216365110.

26. Bergman, S. J. H.; Comstock, L. R. N-Mustard Analogs of S-Adenosyl-l-Methionine as Biochemical Probes of Protein Arginine Methylation. *Bioorg. Med. Chem.* **2015**, *23* (15), 5050–5055. https://doi.org/10.1016/j.bmc.2015.05.001.

27. Erguven, M.; Cornelissen, N. V.; Peters, A.; Karaca, E.; Rentmeister, A. Enzymatic Generation of Double-Modified AdoMet Analogues and Their Application in Cascade Reactions with Different Methyltransferases. *ChemBioChem* **2022**, *23* (24), e202200511. https://doi.org/10.1002/cbic.202200511.

28. Zhang, J.; Zheng, Y. G. SAM/SAH Analogs as Versatile Tools for SAM-Dependent Methyltransferases. *ACS Chem. Biol.* **2016**, *11* (3), 583–597. https://doi.org/10.1021/acschembio.5b00812.

29. Rudenko, A. Yu.; Mariasina, S. S.; Sergiev, P. V.; Polshakov, V. I. Analogs of S-Adenosyl-L-Methionine in Studies of Methyltransferases. *Mol. Biol.* **2022**, *56* (2), 229–250. https://doi.org/10.1134/S002689332202011X.

30. Komoto, J.; Yamada, T.; Takata, Y.; Markham, G. D.; Takusagawa, F. Crystal Structure of the S-Adenosylmethionine Synthetase Ternary Complex: A Novel Catalytic Mechanism of S-Adenosylmethionine Synthesis from ATP and Met. *Biochemistry* **2004**, *43* (7), 1821–1831. https://doi.org/10.1021/bi035611t.

31. Taylor, J. C.; Markham, G. D. The Bifunctional Active Site of S-Adenosylmethionine Synthetase: ROLES OF THE BASIC RESIDUES *. *J. Biol. Chem.* **2000**, *275* (6), 4060–4065. https://doi.org/10.1074/jbc.275.6.4060.

32. Taylor, J. C.; Markham, G. D. The Bifunctional Active Site of S-Adenosylmethionine Synthetase: ROLES OF THE ACTIVE SITE ASPARTATES *. *J. Biol. Chem.* **1999**, *274* (46), 32909–32914. https://doi.org/10.1074/jbc.274.46.32909.

33. Reczkowski, R. S.; Taylor, J. C.; Markham, G. D. The Active-Site Arginine of S-Adenosylmethionine Synthetase Orients the Reaction Intermediate. *Biochemistry* **1998**, *37* (39), 13499–13506. https://doi.org/10.1021/bi9811011.

34. Takusagawa, F.; Kamitori, S.; Markham, G. D. Structure and Function of S-Adenosylmethionine Synthetase: Crystal Structures of S-Adenosylmethionine Synthetase with ADP, BrADP, and PPi at 2.8 Å Resolution,. *Biochemistry* **1996**, *35* (8), 2586–2596. https://doi.org/10.1021/bi952604z.

35. Reczkowski, R. S.; Markham, G. D. Structural and Functional Roles of Cysteine 90 and Cysteine 240 in S-Adenosylmethionine Synthetase (∗). *J. Biol. Chem.* **1995**, *270* (31), 18484–18490. https://doi.org/10.1074/jbc.270.31.18484.

36. McQueney, M. S.; Anderson, K. S.; Markham, G. D. Energetics of S-Adenosylmethionine Synthetase Catalysis. *Biochemistry* **2000**, *39* (15), 4443–4454. https://doi.org/10.1021/bi992876s.

37. Taylor, J. C.; Takusagawa, F.; Markham, G. D. The Active Site Loop of S-Adenosylmethionine Synthetase Modulates Catalytic Efficiency. *Biochemistry* **2002**, *41* (30), 9358–9369. https://doi.org/10.1021/bi025851t.

38. Lu, Z. J.; Markham, G. D. Enzymatic Properties of S-Adenosylmethionine Synthetase from the Archaeon Methanococcus Jannaschii. *J. Biol. Chem.* **2002**, *277* (19), 16624–16631. https://doi.org/10.1074/jbc.M110456200.

39. Gupte, V. Engineering SAM Synthetase for the Production of SAM Nucleobase Analogs, 2022.

40. Popadić, D.; Mhaindarkar, D.; Dang Thai, M. H. N.; Hailes, H. C.; Mordhorst, S.; Andexer, J. N. A Bicyclic S-Adenosylmethionine Regeneration System Applicable with Different Nucleosides or Nucleotides as Cofactor Building Blocks. *RSC Chem Biol* **2021**, *2* (3), 883–891. https://doi.org/10.1039/D1CB00033K.

41. Cornelissen, N. V.; Michailidou, F.; Muttach, F.; Rau, K.; Rentmeister, A. Nucleoside-Modified AdoMet Analogues for Differential Methyltransferase Targeting. *Chem Commun* **2020**, *56* (14), 2115–2118. https://doi.org/10.1039/C9CC07807J.

42. Gade, M.; Tan, L. L.; Damry, A. M.; Sandhu, M.; Brock, J. S.; Delaney, A.; Villar-Briones, A.; Jackson, C. J.; Laurino, P. Substrate Dynamics Contribute to Enzymatic Specificity in Human and Bacterial Methionine Adenosyltransferases. *JACS Au* **2021**, *1* (12), 2349–2360. https://doi.org/10.1021/jacsau.1c00464.

43. Chouhan, B. P. S.; Gade, M.; Martinez, D.; Toledo-Patino, S.; Laurino, P. Implications of Divergence of Methionine Adenosyltransferase in Archaea. *FEBS Open Bio* **2022**, *12* (1), 130–145. https://doi.org/10.1002/2211-5463.13312.

44. Blount, Z. D. The Natural History of Model Organisms: The Unexhausted Potential of *E. Coli*. *eLife* **2015**, *4*, e05826. https://doi.org/10.7554/eLife.05826.

45. Ottink, O. M.; Nelissen, F. H. T.; Derks, Y.; Wijmenga, S. S.; Heus, H. A. Enzymatic Stereospecific Preparation of Fluorescent S-Adenosyl-l-Methionine Analogs. *Anal. Biochem.* **2010**, *396* (2), 280–283. https://doi.org/10.1016/j.ab.2009.09.013.

46. Markham, G. D.; Hafner, E. W.; Tabor, C. W.; Tabor, H. S-Adenosylmethionine Synthetase from Escherichia Coli. *J. Biol. Chem.* **1980**, *255* (19), 9082–9092.

47. Mathur, Y.; Sreyas, S.; Datar, P. M.; Sathian, M. B.; Hazra, A. B. CobT and BzaC Catalyze the Regiospecific Activation and Methylation of the 5-Hydroxybenzimidazole Lower Ligand in Anaerobic Cobamide Biosynthesis. *J. Biol. Chem.* **2020**, *295* (31), 10522–10534. https://doi.org/10.1074/jbc.RA120.014197.

48. Parungao, G. G.; Zhao, M.; Wang, Q.; Zano, S. P.; Viola, R. E.; Blumenthal, R. M. Complementation of a MetK-Deficient E. Coli Strain with Heterologous AdoMet Synthetase Genes. *Microbiol. Read. Engl.* **2017**, *163* (12), 1812–1821. https://doi.org/10.1099/mic.0.000565.

49. Tseng, W.-C.; Lin, J.-W.; Wei, T.-Y.; Fang, T.-Y. A Novel Megaprimed and Ligase-Free, PCR-Based, Site-Directed Mutagenesis Method. *Anal. Biochem.* **2008**, *375* (2), 376–378. https://doi.org/10.1016/j.ab.2007.12.013.

50. *The Proteomics Protocols Handbook*; Walker, J. M., Ed.; Humana Press: Totowa, N.J, 2005.

51. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25* (13), 1605–1612. https://doi.org/10.1002/jcc.20084.

52. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215* (3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

53. Edgar, R. C. MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Res.* **2004**, *32* (5), 1792–1797. https://doi.org/10.1093/nar/gkh340.

54. Waterhouse, A. M.; Procter, J. B.; Martin, D. M. A.; Clamp, M.; Barton, G. J. Jalview Version 2—a Multiple Sequence Alignment Editor and Analysis Workbench. *Bioinformatics* **2009**, *25* (9), 1189–1191. https://doi.org/10.1093/bioinformatics/btp033.

55. Buckstein, M. H.; He, J.; Rubin, H. Characterization of Nucleotide Pools as a Function of Physiological State in *Escherichia Coli*. *J. Bacteriol.* **2008**, *190* (2), 718–726. https://doi.org/10.1128/JB.01020-07.

56. Wei, Y.; Newman, E. B. Studies on the Role of the MetK Gene Product of Escherichia Coli K-12. *Mol. Microbiol.* **2002**, *43* (6), 1651–1656. https://doi.org/10.1046/j.1365-2958.2002.02856.x.

57. Driskell, L. O.; Tucker, A. M.; Winkler, H. H.; Wood, D. O. Rickettsial *MetK*-Encoded Methionine Adenosyltransferase Expression in an *Escherichia Coli MetK* Deletion Strain. *J. Bacteriol.* **2005**, *187* (16), 5719–5722. https://doi.org/10.1128/JB.187.16.5719-5722.2005.

58. Siloto, R. M. P.; Weselake, R. J. Site Saturation Mutagenesis: Methods and Applications in Protein Engineering. *Biocatal. Agric. Biotechnol.* **2012**, *1* (3), 181–189. https://doi.org/10.1016/j.bcab.2012.03.010.

59. Wang, Y.; Xue, P.; Cao, M.; Yu, T.; Lane, S. T.; Zhao, H. Directed Evolution: Methodologies and Applications. *Chem. Rev.* **2021**, *121* (20), 12384–12444. https://doi.org/10.1021/acs.chemrev.1c00260.

60. Baykov, A. A.; Evtushenko, O. A.; Avaeva, S. M. A Malachite Green Procedure for Orthophosphate Determination and Its Use in Alkaline Phosphatase-Based Enzyme Immunoassay. *Anal. Biochem.* **1988**, *171* (2), 266–270. https://doi.org/10.1016/0003-2697(88)90484-8.

61. Ji, X.; Li, Y.; Xie, L.; Lu, H.; Ding, W.; Zhang, Q. Expanding Radical SAM Chemistry by Using Radical Addition Reactions and SAM Analogues. *Angew. Chem. Int. Ed.* **2016**, *55* (39), 11845–11848. https://doi.org/10.1002/anie.201605917.

62. Ding, W.; Wu, Y.; Ji, X.; Qianzhu, H.; Chen, F.; Deng, Z.; Yu, Y.; Zhang, Q. Nucleoside-Linked Shunt Products in the Reaction Catalyzed by the Class C Radical S-Adenosylmethionine Methyltransferase NosN. *Chem Commun* **2017**, *53* (37), 5235–5238. https://doi.org/10.1039/C7CC02162C.

63. Gade, M.; Gardner, J. M.; Jain, P.; Laurino, P. Nucleoside-Driven Specificity of DNA Methyltransferase. *bioRxiv* **2023**, 2022.10.26.513954. https://doi.org/10.1101/2022.10.26.513954.

.....