An enzyme engineering approach to the biosynthesis and utilization of FAD nucleobase analogues

विद्या वाचस्पति की
उपाधि की अपेक्षाओं की आंशिक पूर्ति में प्रस्तुत शोध प्रबंध

A thesis submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy

द्वारा / By

अतीक शाह / Ateek Shah

पंजीकरण सं. / Registration No.:

20162025

शोध प्रबंध पर्यवेक्षक / Thesis Supervisor:

डॉ. अमृता बी. हाजरा / Dr. Amrita B. Hazra

भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान पुणे

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

2023

CERTIFICATE

Certified that, the work incorporated in the thesis entitled, "**An enzyme engineering approach to the biosynthesis and utilization of FAD nucleobase analogues**" submitted by **Ateek Shah** was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or institution.

Date: 1st September 2023

Amrita Hazra

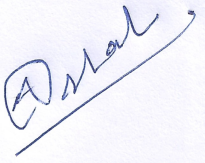
Dr. Amrita B. Hazra

Pune (MH), India

(Supervisor)

DECLARATION

I declare that this written submission represents my ideas in my own words and where others' ideas have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.



1st September, 2023

Ateek Shah

20162025

Acknowledgements

The road has been lengthy, but I have finally arrived at this point. I did not undertake this journey alone; there were individuals around me who provided support and motivation to help me reach this significant milestone. It is time to document these memories and express my gratitude to those who played a role in this journey.

I would like to commence by expressing my heartfelt gratitude to my supervisor, Amrita Hazra, who has been an exceptional teacher and the most outstanding mentor I have ever had. When I initially joined IISER, intending to delve into intriguing electrochemistry, it was Amrita's teaching that inspired me to explore the field of chemical biology. I vividly recall our first meeting regarding joining her lab, during which I had serious doubts about my ability to delve into chemical biology due to my limited knowledge of biology. Her reassuring words were, "Do not worry, you will do great." My primary aspiration was to engage in scientific explorations, and she expertly guided me in this pursuit. Amrita has been a steadfast support in my professional journey and has provided unwavering assistance during my personal highs and lows. I consider myself immensely fortunate to have had the opportunity to work with her.

I am equally appreciative of Gayathri Pananghat and S. G. Srivatsan for their roles as members of my research advisory committee (RAC). Their contributions have been invaluable, providing me with profound insights and guidance for my research endeavours. Their unwavering motivation and support during our meetings have been helpful in my progress. I distinctly remember my nervousness during our initial meeting; their encouragement was a source of great comfort and inspiration. I will not miss to thank Nishad Matange for his valuable input regarding my project's homologous recombineering and antibiotic stress studies.

Now, let us turn our attention to the extensive group of individuals within the Whytamin lab. Yashwant established the nucleotide specificity project in the lab, and then I joined him as a student. He has proven to be an exceptionally composed mentor. He patiently guided me through the intricate experiments of molecular biology and enzymatic assays, starting from the basics and meticulously explaining every bit of the detail. Moreover, beyond his mentorship, Yashwant has also been a wonderful friend and a source of invaluable emotional support throughout our journey together. Throughout this journey, I have been fortunate to have many lab members who have been invaluable sources of biological insights and support. Rupali's

extensive experience and diverse skills in microbiology, along with Yamini's profound knowledge of biochemistry, were instrumental in helping me troubleshoot various experiments. Reman's computational expertise laid my foundation for enzyme engineering, and Janhvi played a role in protein purification and characterization. Rohan's contribution even is a part of one of my publications. Vruta and Adiya expanded the nucleotide specificity project, further enriching our research. I fondly recall engaging in interesting and enjoyable conversations with Prathmesh, and Mrutyunjay has consistently been a calm and polite friend. Aniket's boundless enthusiasm continues to inspire me, and his humour-laden presentations always bring a smile. I have cherished my conversations with Shweta, which have allowed us to share our ups and downs. Riya's kindness has been a constant source of comfort. Interactions with newcomers to the lab have been both stimulating and enlightening. Saswata, Sheryl, Vibhor, Neel, Gina Welsing, Vyankatesh, Rajdip, and others have contributed directly or indirectly to this project. The unforgettable memories of our lab trips and outings have added a rich layer to this remarkable journey.

I want to sincerely thank the other labs that offered valuable assistance and graciously allowed me to work within their research spaces. Ajay from Harinath Chakrapani's lab played a pivotal role in introducing me to the realm of chemical synthesis. Not only is he a senior figure, but he has also become a dear friend throughout our journey. Furthermore, Pulak from Srivatsan's lab, a fellow member of my batch, provided essential guidance in establishing the chemical synthesis procedures for flavin analogues. Jayshree Rajput from Srinivas Hotha's lab generously provided reagents and offered invaluable assistance in deacetylation reactions. I must also express my gratitude to Shabana Khan's lab, which allowed me to utilize their handheld UV chamber for TLC visualization, enabling me to capture fluorescent TLC images. I thank Alok Das, Jeet Kalia, and Nirmalya Ballav Labs for welcoming me for my one-month lab rotations.

I am deeply grateful for the NMR and LC-MS facilities at IISER Pune. Nitin provided invaluable support in conducting NMR analysis, while Saddam offered expert assistance in mass spectrometry analysis. Their contributions were pivotal to the success of my research, and I appreciate their dedication and expertise.

I extend my heartfelt gratitude to IISER Pune for awarding me the institute fellowship, which provided crucial financial support during my research journey. Additionally, I am thankful to IISER for providing the necessary infrastructure, both in terms of living

accommodations and research facilities, that allowed me to carry out my work effectively. I would like to acknowledge the chemistry and biology departments at IISER for their support and for facilitating my research endeavours, where the excellent infrastructure and the collaborative sharing of instruments and facilities have immensely benefited my work. I want to thank the chemistry, biology and academic section staff members who have been there for help.

In the field of research, some moments leave you feeling exhausted and frustrated. In times like these, the IISER community serves as a supportive family. I have had the privilege of forging many wonderful friendships during my time here. Right from the beginning, Prakash and Manish have been among my closest friends with whom I have shared my professional and personal journeys. Prakash's intellect and inspiring thoughts have always motivated me, while Manish has imparted valuable lessons in diplomacy. IISER Has a diverse range of sports communities, and I have been an active member of the cricket and volleyball communities. Through these sports, I have had the pleasure of meeting and befriending numerous fascinating individuals. Umashankar, in particular, has held a special place in my heart, and his unwavering courage has been a constant source of admiration. Through Umashankar, I had the pleasure of getting to know two remarkable friends, Ravi and Tumpa. I am profoundly grateful for my involvement with the Furious XI cricket team. Although we may not have clinched an IISER Premier League (IPL) victory, we certainly won the hearts of our peers. Over the past year, Shuvam, Pratishruti, Devyani, and Abhishek from the volleyball community have become integral parts of this journey. During the final stages of my thesis, Pratishruti offered invaluable mental support and assisted me in the writing process.

Beyond the IISER community, my family has been my unwavering support throughout my years in research. I deeply appreciate having parents, Tarabi Bano and Naseer Shah, who not only encouraged me to pursue significant goals but also made personal sacrifices for my well-being. Additionally, my academic progress has been bolstered by my siblings, Kadeer, Tabassum, and Majeed, who have played pivotal roles. During challenging moments, my friends Shiv Kumar and Sachin have stood by me, and I am confident they will continue to do so in the future.

List of Publications

Commentary

Double agent indole-3-acetic acid: mechanistic analysis of indole-3-acetaldehyde dehydrogenase AldA that synthesizes IAA, an auxin that aids bacterial virulence

Ateek Shah*, Yamini Mathur*, Amrita B. Hazra [* Shared first author]

Bioscience Reports 41 (8), **2021**

DOI: https://doi.org/10.1042/BSR20210598

Research article

Efficient chemical and enzymatic syntheses of FAD nucleobase analogues and their analysis as enzyme cofactors

Ateek Shah*, Yashwant Kumar*, S. Rohan, Amrita B Hazra

ChemBioChem 24 (11), **2023**

DOI: https://doi.org/10.1002/cbic.202300055

Under preparation

An enzyme engineering approach to the biosynthesis and utilization of FAD nucleobase analogues in *Escherichia coli*

Ateek Shah, Reman Kumar, Amrita B. Hazra

TABLE OF CONTENTS

List of Figures

List of Tables

ABSTRACT

The nucleoside triphosphates (NTPs: ATP, GTP, CTP, and UTP) are the building blocks of RNA, and apart from that, they also serve specific functions in cellular metabolism, performing cellular processes such as energy production and signalling. These nucleotides share a common triphosphate and ribose moiety and are only identified based on the nucleobases. Despite the similar structure, a high degree of nucleotide selectivity of enzymes exists for various reactions. For instance, ATP is known as energy currency for all types of cells, and GTP is an intermediate in transmembrane signalling pathways, serving as a substrate for the G-protein family. Additionally, adenosine moiety from ATP serves as a recognition handle in coenzymes like flavin adenine dinucleotide (FAD), nicotinamide adenine dinucleotide (NAD), S-adenosyl methionine (SAM), and coenzyme A, while the other nucleobase analogues of these cofactors are not trackable in organisms signifying the specificity of ATP for the enzymes involved in the cofactor biosynthesis. For example, FAD is synthesized from vitamin B_2 (riboflavin), where riboflavin is converted to FMN (flavin mononucleotide), followed by the adenylation of FMN to FAD by the enzymes riboflavin kinase (RFK) and FMN adenylyltransferase (FMNAT), where only ATP is utilized in both the steps among the NTPs. In prokaryotes, both steps are performed by a bifunctional FAD synthetase enzyme possessing two separate modules – RFK and FMNAT module. The best example is the structurally well-explored bifunctional FAD synthetase from *Corynebacterium ammoniagenes* (*Ca*FADS).

This thesis focused on the FAD biosynthesis in *Escherichia coli* by FAD synthetase (*Ec*FADS), a homolog of *Ca*FADS. We synthesized FAD nucleobase analogues via new chemical and enzymatic approaches requiring no more than three steps and yielding moderate amounts (10-57%). Then, the analogues were subjected to evaluate their binding and functional capability with glutathione reductase flavoprotein and found to be active as reducing agents similar to FAD. The RFK module is ATP-specific, and the FMNAT module shows promiscuity towards each NTP, except GTP, resulting in FAD analogues (FCD, FUD, and dFAD). Despite this, *E. coli* shows only FAD production. We tuned the efficiency of the FMNAT module with each NTP by using computational and structural homology approaches. Specifically, we swapped the loop around the adenine of ATP within the FMNAT domain with a loop of a cytidylyl transferase enzyme. The LoopSwap mutant is efficiently active with GTP and shows increased activity with each NTP relative to *Ec*FADS. Notably, the expression of the LoopSwap mutant in *E. coli* resulted in the synthesis of FAD analogous (FGD, FCD and FUD) within the cell. Furthermore, the shortening of the same loop established its key role in the ATP specificity of the FMNAT module, producing FCD and FUD within the cell while showing decreased or diminished *in vitro* activity with each NTP. Swapping the *Ec*FADS expressing essential *ribF* gene with the LoopSwap (*ribF25*) mutant gene resulted in the biosynthesis of FGD, FCD, and FUD. Interestingly, the *ribF25* mutant strain shows similar growth as the wild-type strain. Importantly, the mutant strain exhibits antibiotic tolerance against aminoglycosides and β-lactams, outperforming the wild-type strain. Due to low FAD, the mutant strain shows slower growth under anaerobic conditions than the wild-type strain. Yet, the antibiotic tolerance is observable anaerobically.

This study proposes that manipulating the structure of vital loops is a potent strategy for fine-tuning enzyme activity without specifically targeting residues within the active site. The biosynthesis of FAD nucleobase analogues has potential applications in understanding FAD's molecular role in cellular metabolism as well as in biotechnology and synthetic biology similar to the applications developed using nucleobase analogue (NCD) of the redox cofactor NAD to make biorthogonal biological systems. Understanding how unnatural molecules, such as FAD analogues, aid antibiotic resistance will help increase antibiotics' efficacy.

Chapter 1

Bioinformatic and molecular efforts to study and alter nucleobase specificity of enzymes

1.1 Introduction

On the level of molecules, the specificity of a substrate for a protein is best understood through the interactions that occur between them. The interaction energies associated with hydrogen bonds of the substrate with the protein and the prevalence of certain protein residues in the vicinity of the substrate have a crucial role in deciding the enzyme's preference for a particular substrate.[1] Moreover, weaker bonds such as electrostatic and Vander Waals in the protein with its substrates significantly contribute to the specificity of the substrate for the enzymes, especially in the case of differentiation between closely related substrates.[2] For instance, the four nucleoside triphosphates (NTPs), ATP, GTP, CTP, and UTP, are the building blocks of RNA, while their deoxy versions (dATP, dGTP, dCTP, and dTTP in place of dUTP) help to store the genetic information as DNA. Besides this, these NTPs also have specific roles to play in cellular metabolism, while these nucleotides share common triphosphate and ribose moiety and differ only on the basis of nucleobase (Figure 1.1). For example, ATP is the energy currency for all the life domains and provides energy for processes to build molecules and tissues by facilitating enzyme activities, transporting molecules across the cell membrane, and facilitating motion.[3,4] GTP is an active substrate of the G-protein family, an intermediate of the transmembrane signalling pathway.[5] CTP is a constituent of CDP-choline, which forms phosphatidylcholine, a major phospholipid part of cell membranes in most nucleated cells.[6] UTP is used to create UDP-sugars such as UDP-glucose, which takes part in sugar metabolism and biosynthetic reactions, for example, glycogen biosynthesis.[7–9]

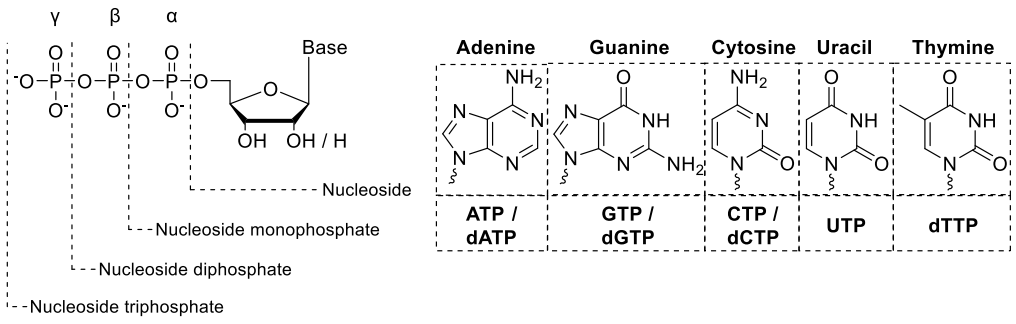


Figure 1.1 Structure of nucleotides.

Even though nucleotide triphosphates differ only based on the nucleobases, high selectivity exists for the choice of nucleotide in various enzymatic reactions. For example, the kinase family generally uses ATP for phosphorylation reactions (Figure 1.3A). Also, the source of adenosine moiety in coenzymes, like S-adenosyl methionine (SAM), nicotinamide adenine dinucleotide (NAD), flavin adenine dinucleotide (FAD), and coenzyme is ATP (Figure 1.2).[10] However, it has been seen that enzyme homologs from various organisms can have different nucleotide utilization for the same function. For example, phosphopantothenoyl-L-cysteine synthetase (PPCS) from *Escherichia coli* is CTP specific, and the same enzyme from *Homo sapiens* is ATP specific (Figure 1.3B).[11,12] Another enzyme, succinyl-CoA synthetase (SCS) from *Sus scrofa*, is GTP specific, while the same from *Blastocystis sp.* utilizes ATP (Figure 1.3C).[13,14] However, the same enzyme from *E. coli* is promiscuous to GTP and ATP, using both as a substrate for the same reaction.[15,16] The family of kinases depends on NTPs for the phosphorylation of a variety of substrates, while the majority of the population is ATP-specific. Besides, there are examples where the same kinase enzyme homolog can utilize GTP or CTP to phosphorylate the same substrate. In this instance, Riboflavin kinase is an example that is ATP-specific generally in eukaryotes and prokaryotes, but the same kinase from archaea *Methanocaldococcus jannaschii* and *Methanococcus maripaludis* was found to be CTP specific, and that from *Rhizopus oryzae* is GTP-specific.[17–22] For another instance, ATP is preferred by Diacylglycerol kinases from *E. coli*, and the homolog from *S. cerevisiae* prefers CTP.[23,24]

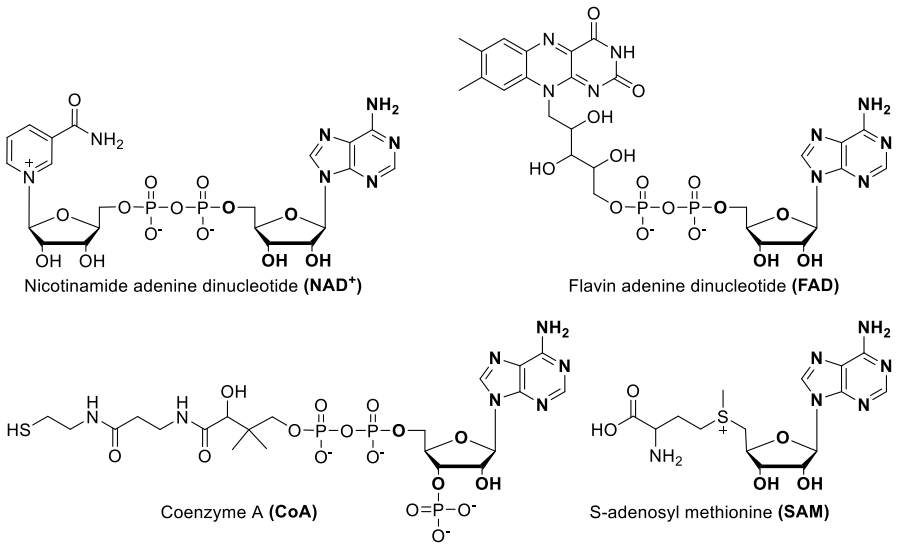


Figure 1.2 Structure of cofactors having the same adenosine moiety (shown as bold) coming from ATP.

The K_m value of NTPs for corresponding enzyme can also play a role in the nucleotide specificity. The nucleotides present in a cell have different concentrations according to their metabolic roles. ATP is most abundant in the cell as it acts as an energy currency as well as serves as cofactors for various reactions. In the nucleotide pool of *E. coli* mid log phase, ATP, GTP, UTP, and CTP are ~3.5, 1.6, 0.7, and 0.3 mM respectively.[25] However, for humans the GTP and UTP concentration drops to around 0.5 mM.[26] If the K_m values are higher than the cellular concentration of NTP, then the nucleotide selectivity can arise. However, there is one study which shows that ATP and GTP both binds to the adenylate kinase while only ATP is used to phosphorylate the AMP.[27] In case of ATP binding a large conformational change happens and enzyme gets into closed structure to facilitate the reaction. When GTP binds to the enzyme, it arrests the enzyme in open structure. This phenomenon balances the energy homeostasis as well as protects the GTP pool from adenylate kinase turnover as GTP serves for various important cellular regulations.

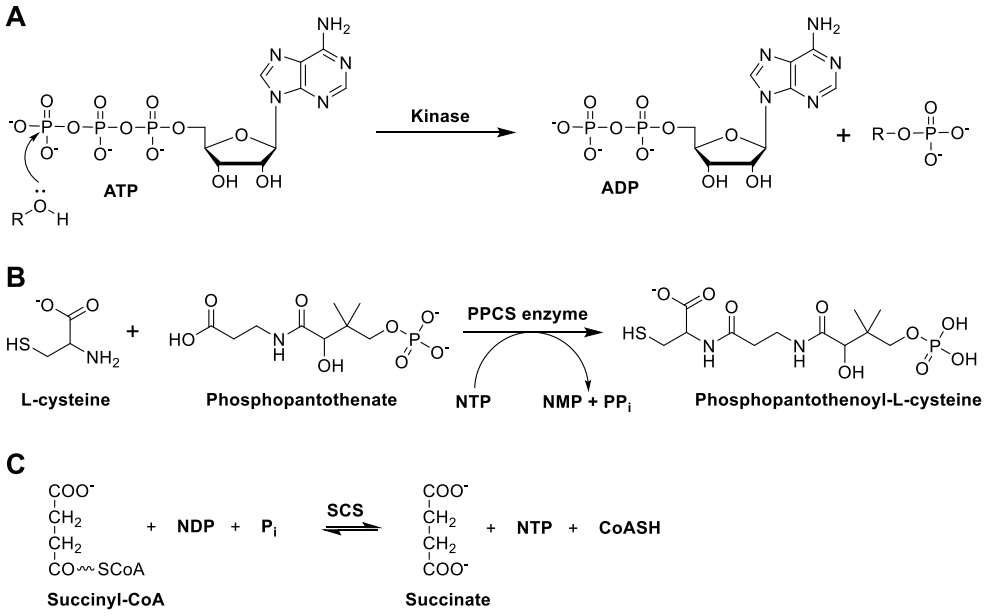


Figure 1.3 Some examples of nucleotide specificity based on the organism. (A) kinase enzyme employs ATP as a hydroxyl group phosphorylation source. **(B)** reaction scheme for phosphopantothenoyl-L-cysteine synthetase (PPCS) enzymatic reaction. **(C)** A scheme for enzymatic reaction of succinyl-CoA synthetase (SCS).

All of these observations and studies lead to several open questions regarding nucleobase specificity, which need to be addressed;

1. What is the molecular basis of this nucleotide specificity?
2. Why does a homologous enzyme from various organisms utilize different NTPs?

3. Why are only adenosine moiety-containing cofactors known, while this moiety does not directly contribute to the catalysis?
4. Can the nucleotide selectivity of an enzyme be altered? Can other nucleoside analogues of cofactors be created in the cell?

1.2 The conserved motifs for nucleotide specificity

Nucleotides play a central role in cellular metabolism, where most reactions utilize the triphosphate moiety of NTPs. For example, γ-phosphate of ATP is used to phosphorylate substrates by a kinase, and the GTP/GDP cycle goes on for signalling. Most enzyme interactions with its nucleotide substrate happen through the triphosphate moiety. However, the nucleotide specificity cannot be attributed directly to these interactions in the ATP and GTP binding enzymes, which share a common phosphate-binding loop (P-loop) but not in all of them (Table 1.1).[28,29] The lysine with a serine or threonine are well-conserved residues in the P-loop, and its structure is glycine-rich. In the G-protein family, N/TKxD is a conserved domain for guanine recognition.[30] However, there is no conserved domain for recognising adenine moiety, and the folds surrounding the nucleotide also vary.[31] Other efforts towards understanding the difference in adenine/guanine recognition have been made, and a notable difference in the conditions surrounding the two ligands has been documented. For instance, adenine is often found to be solvent-exposed, where the hydrogen bonding is fulfilled through water molecules, while guanine moiety is mostly situated deep within the active site, utilizing residues over water molecules to establish hydrogen bonds.[1] Besides this, the protein residues interacting with either nucleobase differ by the number and type of residues around nucleobase. Cytidylyltransferase superfamily also possesses a conserved motif HxGH for CTP binding, and some adenylyltransferases also have a similar motif.[32] The motif residues interact with phosphate oxygen of CTP or ATP via hydrogen bonds.

Table 1.1 Conserved residues in NTP utilising enzymes.

NTP	Conserved residues interacting with phosphates of NTP	Conserved residues interacting with the nucleobase of NTP
ATP	GxxxxGK	
GTP	GxxxxGK and DxxG	N/TKxD
CTP	HxGH	

Adenine binding has no significant sequence motif; still, it follows a specific trend of backbone interactions in ATP and other cofactors like FAD, NAD, and SAM (Figure 1.4).[10,31] It follows backbone polar and nonspecific hydrophobic interactions unrelated to the characteristic of side chains of protein residues. The motif for adenine-binding commonly has a loop composed of 3 amino acid residues (referred to as sites I, II, and III), accompanied by an extra non-polar residue located near the adenine ring. This loop can bind to the adenine moiety via two adenine-binding motifs: the "direct" or "reverse" motif (Figure 1.4). From the adenine ring, the N-1 and the NH_2 group of N-6 participate in the essential hydrogen bonding with the loop. The backbone NH group of the 3rd amino acid (site III) within the loop forms a hydrogen bond with the N-1 atom of adenine within the "direct" motif. Concurrently, the backbone carbonyl group of the 1st amino acid (site I) establishes the hydrogen bond with the NH_2 group at the N-6 site of the adenine. Conversely, within the "reverse" motif, N-1 and N-6 atoms of the adenine moiety establish hydrogen bonds with the 2nd amino acid (site II) in the loop through the backbone's amide and carbonyl group. Another motif called the "Asp" motif is often found in NAD-binding proteins, where a negatively charged amino acid side chain interacts with adenine's NH_2 group at the N-6 position (Figure 1.4).

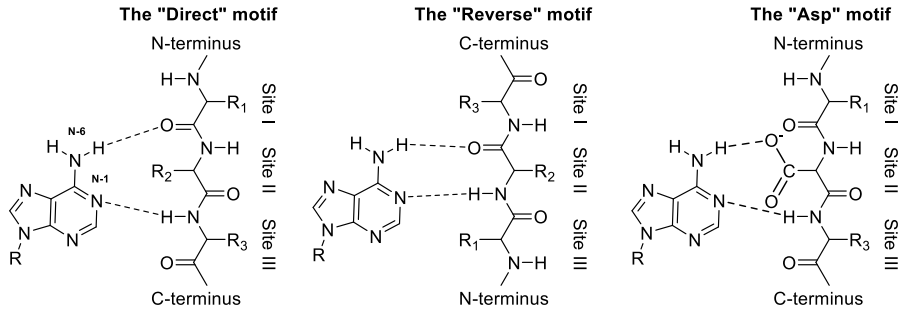


Figure 1.4 The trend of backbone interaction of adenine moiety of ATP and enzyme cofactors to enzyme residues.

1.3 Electrostatic potential approach to understand the adenine/guanine selectivity

The other method to understand nucleotide specificity is to determine the electrostatic potential (ESP) for active sites in the nucleotide-binding proteins rather than looking at the conserved motifs and sequence analysis. One study has discriminated between adenine/guanine-specific active sites in non-redundant proteins by expressing the two ESP of the adenine/guanine active sites for the dataset as a linear combination and then aligning the ESP vectors roughly with molecular dipoles of adenine/guanine and established high

correlation between adenine/guanine selectivity in the active sites and the electrostatic replacement energies of native substrate by the other non-native substrate.[2] Later, another study employed homology modelling, docking and molecular simulations for ESP determination and predicted the specificity of a nucleotide in *Blastocystis* succinyl-CoA synthetase (SCS).[33] This study indicated that ATP and GTP could be incorporated in the active binding site, while GTP is not the native substrate. The Lys42 and Lys110 destabilize the GTP by electrostatic dipole interactions, making this SCS enzyme ATP-specific.

1.4 Larger pocket for larger modified nucleobases

The phosphorylation of a protein is an essential phenomenon which participates in governing a part of all cellular signalling pathways. Homology modelling and genetic screening are a rapid way of identifying novel protein kinases. However, protein kinases show overlapping and redundancy for their substrates, making it challenging to find their native protein substrates. Kevan Shokat and co-workers engineered a protein tyrosine kinase, v-Src, to accommodate ATP analogues that do not bind to wild-type v-Src and this engineered kinase was used to screen its direct substrates in the presence of other ATP-utilizing kinases.[34] They created a pocket around adenine moiety by mutating bigger residues V323 and I338A with smaller alanine residues to accommodate bulky N6-modified ATP analogues such as N6-(cyclopentyl) ATP. Later, they found that I338 residue alone is responsible for this specificity and this site is well conserved in Src family protein kinases with similar residues.[35] Hence, they applied glycine mutation in v-Src (I338G) and other member Fyn kinase (T339G) and among many bulky N^6-modified ATP analogues, N^6-(benzyl) ATP was found to be the best substrate. Another study was done to probe G-Protein signalling by engineering unnatural nucleotide specificity where H-Ras G-protein was engineered to create a pocket to accommodate bulky guanine-modified GTP/GDP analogues by mutating L19 and N116 to alanine or glycine.[36] This mutant is unaffected by the mutation and works similarly to the wild-type enzyme with GTP/GDP. Moreover, this mutant is regulated by a pair of bulky GTP/GDP analogous, where it can be deactivated by one analogue and further activated by the other one.

1.5 The role of gatekeeper residues towards nucleotide selectivity

Kevan Shokat's group introduced the molecular gate concept (now called gatekeeper residue) for substrate specificity. They claimed that the partial pocket to accommodate bulky N^6-modified ATP already exists, and the point mutation (I338G or I338A in v-Src) creates a gate to enter these bulky substrates in the active site.[37] These types of ATP analogues, modified at various sites, have been widely used for protein kinase research.[38,39] Later, another group engineered nucleotide specificity of *Blastocystis sp.* SCS from ATP to GTP by mutating the gatekeeper residues.[40] They identified two gatekeeper residues guarding at the entry point of the nucleotide-active site as charged as well as solvent-exposed and claimed that these charged amino acids control the entry of the substrate in the binding site by creating electrostatic interactions with the non-native ligand GTP. They called this mechanism of control of access the "electrostatic gatekeeper effect". These gatekeeper lysine amino acids (KK) were replaced with two negatively charged residues (ED), glutamic acid and aspartic acid, which gave access to GTP in the active site. However, this ED mutant did not restrict the ATP from entering. Further mutation of leucine (L) to phenylalanine (F) and valine (V) to leucine (L) made this enzyme specific to GTP only. All these mutations were done based on the sequence alignment of *Blastocystis* SCS with a GTP-specific SCS from *Sus scorfa*. In another study, the gatekeeper residue of APH(2″)-IIa and APH(2″)-Iva, known as aminoglycoside 2″-phosphotransferase, was replaced with a bulky residue, as both these enzymes can incorporate ATP and GTP as co-substrates.[41] This mutation was based on APH(2″)-IIIa, which is GTP-specific and blocks the entry of ATP through a bulky amino acid named tyrosine. The M85Y mutation within the APH(2″)-IIa caused a ~320 times increase to the K_m of ATP and a ~10 times decrease to the K_m of GTP. However, in the case of F95Y mutation of APH(2″)-IVa, a slight decline of K_m for both ATP and GTP can be observed because the F95Y mutant has a larger active site pocket, allowing the movement of tyrosine residue away from the active site, which makes the entry of ATP and GTP to the active site possible.

1.6 Modulation of nucleotide selectivity through entire loop alterations or single Point mutations

In addition to factors like electrostatic potential, a large pocket, or gatekeeper residues, nucleotide preference can also be tuned by a point mutation. For example, a single amino acid substitution, D138N, in an elongation factor Tu (EF-Tu) of *E. coli*, lowered the substrate

affinity of the protein for GDP and enhanced its activity with xanthosine 5'-diphosphate (XDP) dramatically. This D138 position is close to the C-2 carbon of purine nucleobases, where GDP has an NH_2 group at the C-2 carbon, directly interacting with the Asp residue (D138); complementarily, Asn in place of D138 amino acid interacts with the 2-carbonyl group of XDP.[42] Later, similar point mutation engineered many GTP/GTP utilizing enzymes to accommodate XTP/XDP.[43] Similarly, cGMP/cAMP kinase active sites differ via threonine or alanine, which are important for the specificity of cGMP and cAMP.[44] In cGMP kinase, the Hydroxy group of threonine interacts with the NH_2 moiety of cGMP via hydrogen bonding, while the same interaction cannot happen with cAMP due to lack of 2-amino group; hence, cAMP kinase possesses alanine at the same position instead of threonine. The mutation A334T in cAMP kinase from bovine testis resulted in a high affinity for cGMP.[45] In another instance, a temperature-sensitive allele of FtsZ, FtsZ84, was found to show an Mg^{2+}-dependent ATPase activity *in vitro* because of a point mutation G105S while the wild-type FtsZ is a GTPase.[46]

The other way of changing substrate recognition is to swap a part of the whole sequence of a protein with its homolog, which is promiscuous to other substrates. In this regard, a *Homo sapiens* hypoxanthine–guanine phosphoribosyl transferase from (*Hs*HGPRT) utilizes the phosphoribosyl pyrophosphate (PRPP) for the phosphoribosylation of hypoxanthine to inosine monophosphate (IMP) and guanine to guanosine monophosphate (GMP). Moreover, the phosphoribosyl transferase from the *Plasmodium falciparum* (*Pf*HGXPRT) recognises xanthine as an additional substrate for the phosphoribosylation. The replacement of 49 residues of *Hs*HGPRT with the corresponding residues from the *Pf*HGXPRT resulted in a chimeric enzyme with additional specificity for xanthine, while this sequence is unrelated to the active site.[47] For another instance, a human guanine deaminase (*h*GDA) was computationally designed to introduce specific enzyme–substrate interactions to an unnatural substrate ammelide by directed loop remodelling near the active site.[48] This was done by comparing the active sites of *h*GDA with the bacterial cytosine deaminase. The loop of *h*GDA was remodelled by introducing the critical interactions of cytosine with the active site of *Pf*HGXPRT, and the residues directly critical for the catalysis were not touched. Relative to wild-type *h*GDA, the redesigned enzyme was found to be $2.5x10^4$-fold less active with native substrate guanine and $1x10^2$-fold more active with ammelide, which is a structural intermediate between guanine and cytosine.

1.7 Objective of dissertation

For my doctoral thesis, I have focused on the biosynthesis of FAD cofactor derived from vitamin B_2 (riboflavin), which utilises two molecules of ATP. The specific questions that we wish to address are:

1. What is the molecular basis of ATP specificity in FAD biosynthesis?

2. Can the nucleotide preference of FAD synthesising enzymes be altered by modifying the active site?

3. Is the biosynthesis of other nucleobase analogues of FAD (FGD, FCD, and FUD) possible in the cell by mutating the genes involved in FAD biosynthesis, and do these analogues affect cell physiology?

References

(1) Nobeli, I.; Laskowski, R. A.; Valdar, W. S. J.; Thornton, J. M. On the Molecular Discrimination between Adenine and Guanine by Proteins. *Nucleic Acids Res* **2001**, *29* (21), 4294–4309. https://doi.org/10.1093/nar/29.21.4294.

(2) Basu, G.; Sivanesan, D.; Kawabata, T.; Go, N. Electrostatic Potential of Nucleotide-Free Protein Is Sufficient for Discrimination Between Adenine and Guanine-Specific Binding Sites. *J Mol Biol* **2004**, *342* (3), 1053–1066. https://doi.org/10.1016/j.jmb.2004.07.047.

(3) Fontecilla-Camps, J. C. The Complex Roles of Adenosine Triphosphate in Bioenergetics. *ChemBioChem* **2022**, *23* (10), e202200064. https://doi.org/10.1002/cbic.202200064.

(4) Dunn, J.; Grider, M. H. *Physiology, Adenosine Triphosphate*; Treasure Island (FL): StatPearls Publishing, 2023.

(5) Gilman, A. G Proteins: Transducers Of Receptor-Generated Signals. *Annu Rev Biochem* **1987**, *56* (1), 615–649. https://doi.org/10.1146/annurev.biochem.56.1.615.

(6) Cornell, R. B.; Ridgway, N. D. CTP:Phosphocholine Cytidylyltransferase: Function, Regulation, and Structure of an Amphitropic Enzyme Required for Membrane

Biogenesis. *Prog Lipid Res* **2015**, *59*, 147–171. https://doi.org/10.1016/j.plipres.2015.07.001.

(7) Abi-Hanna, A.; Saavedra, J. M. GALACTOSE*. In *Encyclopedia of Human Nutrition*; Elsevier, 1998; pp 377–383. https://doi.org/10.1016/B0-12-226694-3/00144-7.

(8) Lecca, D.; Ceruti, S. Uracil Nucleotides: From Metabolic Intermediates to Neuroprotection and Neuroinflammation. *Biochem Pharmacol* **2008**, *75* (10), 1869–1881. https://doi.org/10.1016/j.bcp.2007.12.009.

(9) Ralevic, V. UDP-Glucose☆. In *Reference Module in Biomedical Sciences*; Elsevier, 2015. https://doi.org/10.1016/B978-0-12-801238-3.09699-9.

(10) Denessiouk, K. A.; Rantanen, V.-V.; Johnson, M. S. Adenine Recognition: A Motif Present in ATP-, CoA-, NAD-, NADP-, and FAD-Dependent Proteins. *Proteins: Structure, Function, and Genetics* **2001**, *44* (3), 282–291. https://doi.org/10.1002/prot.1093.

(11) Strauss, E.; Kinsland, C.; Ge, Y.; McLafferty, F. W.; Begley, T. P. Phosphopantothenoylcysteine Synthetase from Escherichia Coli. *Journal of Biological Chemistry* **2001**, *276* (17), 13513–13516. https://doi.org/10.1074/jbc.C100033200.

(12) Manoj, N.; Strauss, E.; Begley, T. P.; Ealick, S. E. Structure of Human Phosphopantothenoylcysteine Synthetase at 2.3 Å Resolution. *Structure* **2003**, *11* (8), 927–936. https://doi.org/10.1016/S0969-2126(03)00146-1.

(13) Fraser, M. E.; James, M. N. G.; Bridger, W. A.; Wolodko, W. T. Phosphorylated and Dephosphorylated Structures of Pig Heart, GTP-Specific Succinyl-CoA Synthetase. *J Mol Biol* **2000**, *299* (5), 1325–1339. https://doi.org/10.1006/jmbi.2000.3807.

(14) Huang, J.; Nguyen, V. H.; Hamblin, K. A.; Maytum, R.; van der Giezen, M.; Fraser, M. E. ATP-Specificity of Succinyl-CoA Synthetase from *Blastocystis Hominis*. *Acta Crystallogr D Struct Biol* **2019**, *75* (7), 647–659. https://doi.org/10.1107/S2059798319007976.

(15) Wolodko, W. T.; Fraser, M. E.; James, M. N.; Bridger, W. A. The Crystal Structure of Succinyl-CoA Synthetase from Escherichia Coli at 2.5-A Resolution. *Journal of Biological Chemistry* **1994**, *269* (14), 10883–10890. https://doi.org/10.1016/S0021-9258(17)34141-8.

(16) Fraser, M. E.; James, M. N. G.; Bridger, W. A.; Wolodko, W. T. A Detailed Structural Description of Escherichia Coli Succinyl-CoA Synthetase 1 1Edited by D. Rees. *J Mol Biol* **1999**, *285* (4), 1633–1653. https://doi.org/10.1006/jmbi.1998.2324.

(17) Manstein, D. J.; Pai, E. F. Purification and Characterization of FAD Synthetase from Brevibacterium Ammoniagenes. *J Biol Chem* **1986**, *261* (34), 16169–16173.

(18) Santos, M. A.; Jiménez, A.; Revuelta, JoséL. Molecular Characterization of FMN1, the Structural Gene for the Monofunctional Flavokinase of Saccharomyces Cerevisiae. *Journal of Biological Chemistry* **2000**, *275* (37), 28618–28624. https://doi.org/10.1074/jbc.M004621200.

(19) Mashhadi, Z.; Zhang, H.; Xu, H.; White, R. H. Identification and Characterization of an Archaeon-Specific Riboflavin Kinase. *J Bacteriol* **2008**, *190* (7), 2615–2618. https://doi.org/10.1128/JB.01900-07.

(20) Kumar, Y.; Singh, R. K.; Hazra, A. B. Characterization of a Novel Mesophilic CTP-Dependent Riboflavin Kinase and Rational Engineering to Create Its Thermostable Homologues**. *ChemBioChem* **2021**, *22* (24), 3414–3424. https://doi.org/10.1002/cbic.202100211.

(21) Tachibana, S. Discovery of a New Flavin Phosphate Synthetase Which Requires Guanosine-5′-Triphosphate. *J Vitaminol (Kyoto)* **1967**, *13* (2), 89–92. https://doi.org/10.5925/jnsv1954.13.89.

(22) Tachibana, S. GTP-Dependent Flavin Phosphate Synthetase. In *Methods in Enzymology*; 1971; pp 553–555. https://doi.org/10.1016/S0076-6879(71)18118-9.

(23) Walsh, J. P.; Bell, R. M. Diacylglycerol Kinase from Escherichia Coli; 1992; pp 153–162. https://doi.org/10.1016/0076-6879(92)09019-Y.

(24) Fakas, S.; Konstantinou, C.; Carman, G. M. DGK1-Encoded Diacylglycerol Kinase Activity Is Required for Phospholipid Synthesis during Growth Resumption from Stationary Phase in Saccharomyces Cerevisiae. *Journal of Biological Chemistry* **2011**, *286* (2), 1464–1474. https://doi.org/10.1074/jbc.M110.194308.

(25) 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.

(26) Traut, T. W. Physiological Concentrations of Purines and Pyrimidines. *Mol Cell Biochem* **1994**, *140* (1), 1–22. https://doi.org/10.1007/BF00928361.

(27) Rogne, P.; Rosselin, M.; Grundström, C.; Hedberg, C.; Sauer, U. H.; Wolf-Watz, M. Molecular Mechanism of ATP versus GTP Selectivity of Adenylate Kinase. *Proceedings of the National Academy of Sciences* **2018**, *115* (12), 3012–3017. https://doi.org/10.1073/pnas.1721508115.

(28) Leipe, D. D.; Koonin, E. V.; Aravind, L. Evolution and Classification of P-Loop Kinases and Related Proteins. *J Mol Biol* **2003**, *333* (4), 781–815. https://doi.org/10.1016/j.jmb.2003.08.040.

(29) Longo, L. M.; Jabłońska, J.; Vyas, P.; Kanade, M.; Kolodny, R.; Ben-Tal, N.; Tawfik, D. S. On the Emergence of P-Loop NTPase and Rossmann Enzymes from a Beta-Alpha-Beta Ancestral Fragment. *Elife* **2020**, *9*, e64415. https://doi.org/10.7554/eLife.64415.

(30) Kjeldgaard, M.; Nyborg, J.; Clark, B. F. C. The GTP Binding Motif: Variations on a Theme. *The FASEB Journal* **1996**, *10* (12), 1347–1368. https://doi.org/10.1096/fasebj.10.12.8903506.

(31) Narunsky, A.; Kessel, A.; Solan, R.; Alva, V.; Kolodny, R.; Ben-Tal, N. On the Evolution of Protein–Adenine Binding. *Proceedings of the National Academy of Sciences* **2020**, *117* (9), 4701–4709. https://doi.org/10.1073/pnas.1911349117.

(32) Bork, P.; Holm, L.; Koonin, E. V.; Sander, C. The Cytidylyltransferase Superfamily: Identification of the Nucleotide-Binding Site and Fold Prediction. *Proteins: Structure, Function, and Genetics* **1995**, *22* (3), 259–266. https://doi.org/10.1002/prot.340220306.

(33) Hamblin, K.; Standley, D. M.; Rogers, M. B.; Stechmann, A.; Roger, A. J.; Maytum, R.; van der Giezen, M. Localization and Nucleotide Specificity of *Blastocystis* Succinyl-CoA Synthetase. *Mol Microbiol* **2008**, *68* (6), 1395–1405. https://doi.org/10.1111/j.1365-2958.2008.06228.x.

(34) Shah, K.; Liu, Y.; Deirmengian, C.; Shokat, K. M. Engineering Unnatural Nucleotide Specificity for Rous Sarcoma Virus Tyrosine Kinase to Uniquely Label Its

Direct Substrates. *Proceedings of the National Academy of Sciences* **1997**, *94* (8), 3565–3570. https://doi.org/10.1073/pnas.94.8.3565.

(35) Liu, Y.; Shah, K.; Yang, F.; Witucki, L.; Shokat, K. M. Engineering Src Family Protein Kinases with Unnatural Nucleotide Specificity. *Chem Biol* **1998**, *5* (2), 91–101. https://doi.org/10.1016/S1074-5521(98)90143-0.

(36) Vincent, F.; Cook, S. P.; Johnson, E. O.; Emmert, D.; Shah, K. Engineering Unnatural Nucleotide Specificity to Probe G Protein Signaling. *Chem Biol* **2007**, *14* (9), 1007–1018. https://doi.org/10.1016/j.chembiol.2007.08.006.

(37) Liu, Y.; Shah, K.; Yang, F.; Witucki, L.; Shokat, K. M. A Molecular Gate Which Controls Unnatural ATP Analogue Recognition by the Tyrosine Kinase V-Src. *Bioorg Med Chem* **1998**, *6* (8), 1219–1226. https://doi.org/10.1016/S0968-0896(98)00099-6.

(38) Blethrow, J. D.; Glavy, J. S.; Morgan, D. O.; Shokat, K. M. Covalent Capture of Kinase-Specific Phosphopeptides Reveals Cdk1-Cyclin B Substrates. *Proceedings of the National Academy of Sciences* **2008**, *105* (5), 1442–1447. https://doi.org/10.1073/pnas.0708966105.

(39) Anthony, T. M.; Dedigama-Arachchige, P. M.; Embogama, D. M.; Faner, T. R.; Fouda, A. E.; Pflum, M. K. H. ATP Analogs in Protein Kinase Research. In *Kinomics*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2015; pp 137–168. https://doi.org/10.1002/9783527683031.ch6.

(40) Vashisht, K.; Verma, S.; Gupta, S.; Lynn, A. M.; Dixit, R.; Mishra, N.; Valecha, N.; Hamblin, K. A.; Maytum, R.; Pandey, K. C.; van der Giezen, M. Engineering Nucleotide Specificity of Succinyl-CoA Synthetase in Blastocystis : The Emerging Role of Gatekeeper Residues. *Biochemistry* **2017**, *56* (3), 534–542. https://doi.org/10.1021/acs.biochem.6b00098.

(41) Bhattacharya, M.; Toth, M.; Smith, C. A.; Vakulenko, S. B. Bulky "Gatekeeper" Residue Changes the Cosubstrate Specificity of Aminoglycoside 2″-Phosphotransferase IIa. *Antimicrob Agents Chemother* **2013**, *57* (8), 3763–3766. https://doi.org/10.1128/AAC.00381-13.

(42) Hwang, Y. W.; Miller, D. L. A Mutation That Alters the Nucleotide Specificity of Elongation Factor Tu, a GTP Regulatory Protein. *Journal of Biological Chemistry* **1987**, *262* (27), 13081–13085. https://doi.org/10.1016/S0021-9258(18)45170-8.

(43) Bishop, A.; Buzko, O.; Heyeck-Dumas, S.; Jung, I.; Kraybill, B.; Liu, Y.; Shah, K.; Ulrich, S.; Witucki, L.; Yang, F.; Zhang, C.; Shokat, K. M. Unnatural Ligands for Engineered Proteins: New Tools for Chemical Genetics. *Annu Rev Biophys Biomol Struct* **2000**, *29* (1), 577–606. https://doi.org/10.1146/annurev.biophys.29.1.577.

(44) Weber, I. T.; Shabb, J. B.; Corbin, J. D. Predicted Structures of the CGMP Binding Domains of the CGMP-Dependent Protein Kinase: A Key Alanine/Threonine Difference in Evolutionary Divergence of CAMP and CGMP Binding Sites. *Biochemistry* **1989**, *28* (14), 6122–6127. https://doi.org/10.1021/bi00440a059.

(45) Shabb, J. B.; Ng, L.; Corbin, J. D. One Amino Acid Change Produces a High Affinity CGMP-Binding Site in CAMP-Dependent Protein Kinase. *Journal of Biological Chemistry* **1990**, *265* (27), 16031–16034. https://doi.org/10.1016/S0021-9258(17)46182-5.

(46) RayChaudhuri, D.; Park, J. T. A Point Mutation Converts Escherichia Coli FtsZ Septation GTPase to an ATPase. *Journal of Biological Chemistry* **1994**, *269* (37), 22941–22944. https://doi.org/10.1016/S0021-9258(17)31600-9.

(47) Sujay Subbayya, I. N.; Sukumaran, S.; Shivashankar, K.; Balaram, H. Unusual Substrate Specificity of a Chimeric Hypoxanthine–Guanine Phosphoribosyltransferase Containing Segments from the Plasmodium Falciparum and Human Enzymes. *Biochem Biophys Res Commun* **2000**, *272* (2), 596–602. https://doi.org/10.1006/bbrc.2000.2816.

(48) Murphy, P. M.; Bolduc, J. M.; Gallaher, J. L.; Stoddard, B. L.; Baker, D. Alteration of Enzyme Specificity by Computational Loop Remodeling and Design. *Proceedings of the National Academy of Sciences* **2009**, *106* (23), 9215–9220. https://doi.org/10.1073/pnas.0811070106.

Chapter 2

Biosynthesis of FAD cofactor and nucleotide specificity

2.1 Introduction

Flavin mononucleotide (FMN) and Flavin adenine dinucleotide (FAD) are two very important coenzymes, and their biosynthesis relies on the vitamin B_2 known as riboflavin.[1] These coenzymes are responsible for important cell functions such as cellular respiration, energy metabolism, and facilitating growth and development.[2,3] Moreover, FAD and FMN contribute to the tryptophan conversion into niacin (vitamin B_3) and vitamin B_6 into its coenzyme form, pyridoxal 5'-phosphate, respectively.[4,5] The regulation of riboflavin, FMN, and FAD levels is well correlated to a range of conditions such as diabetes mellitus, metabolic irregularities, neurodegenerative disorders, and inborn metabolism errors such as multiple acyl-CoA dehydrogenation deficiency (MADD).[6–10] The FAD cofactor is derived from riboflavin precursor through two enzymatic reactions consuming an ATP molecule in each reaction. In the initial step, riboflavin kinase (RibK) catalyses the phosphate group transfer to riboflavin, resulting in flavin mononucleotide (FMN), utilizing ATP, and subsequently, the FMN adenylyl transferase (FMNAT) converts FMN into FAD with the aid of another ATP molecule (Figure 2.1).[11–13] Prokaryotes, such as *Corynebacterium ammoniagenes*, employ a bifunctional FAD synthetase (FADS) enzyme to perform both steps.[11] This enzyme combines the action of RibK and FMNAT into one catalytic unit having two separate modules. The prokaryotic RibK modules and eukaryotic RibKs exhibit sequence and structural similarity.[14,15] Conversely, the prokaryotic FMNAT module differs from eukaryotic FAD synthetase enzymes, highlighting the potential of bacterial FAD synthetases as viable targets for drug development.[16–18]

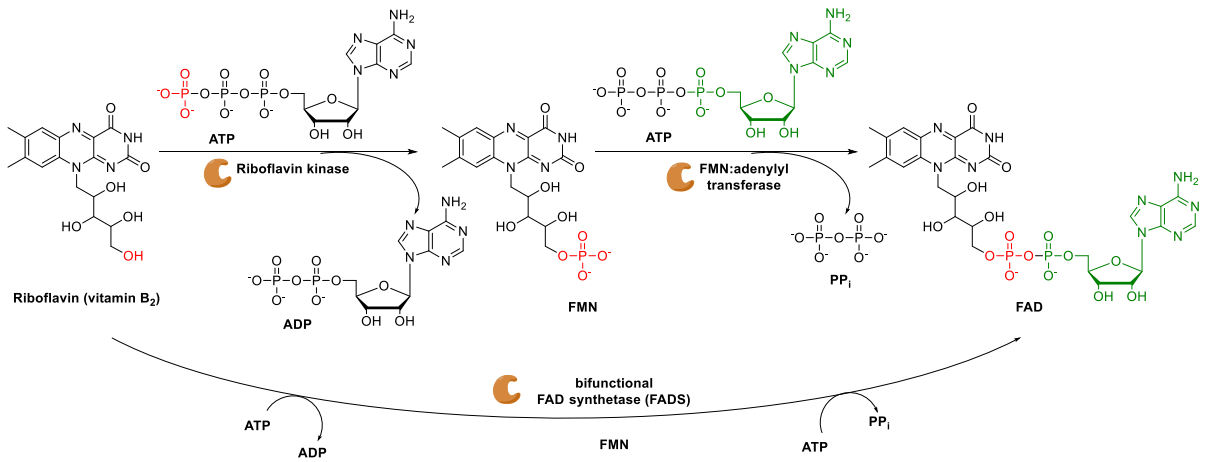


Figure 2.1 Scheme for flavin adenine dinucleotide (FAD) biosynthesis. Within the cell, riboflavin (RF) undergoes an initial transformation into flavin mononucleotide (FMN) through the action of riboflavin kinase (RibK). Subsequently, FMN adenylyl transferase (FMNAT) further modifies it through adenylation to generate Flavin adenine dinucleotide (FAD).where both these enzymes use adenosine triphosphate (ATP) as the second substrate. In prokaryotes, both steps are derived by a bifunctional FAD synthetase enzyme.

2.2 Nucleotide selectivity of riboflavin kinases from different organisms

Both the phosphorylation and adenylation reactions are known to be specific for the use of ATP as substrate. However, the use of other NTPs by riboflavin will not influence the resultant product FMN. Yet, there are exceptions where RibK enzymes from various organisms depend on other NTPs to produce FMN, and the NTP utilization by RibKs varies across different species (Table 2.1). Unlike other ATP-related RibKs, the archaeal RibKs from the thermophile *Methanocaldococcus jannaschii* and mesophilic *Methanococcus maripaludis* are specific to CTP.[19–21] Similarly, GTP is utilized by the RibK from the fungi *Rhizopus javanicus* for the phosphorylation reaction, forming FMN.[22] This enzyme exhibits minor activity with UTP but does not respond to ATP and CTP. The same author also reported a cyclic FMN synthetase from *Rhizopus oryzae* that relies on GTP and has minor activity with UTP and ITP.[23] Another CTP-dependent protein known as RbkR with a dual function of a transcriptional regulator and a kinase enzyme was found in archaea *Thermoplasma acidophilum*.[24] CTP-dependent RibKs show conservation across the archaeal family, with certain instances being bifunctional, combining their enzymatic role with a fused transcriptional regulator domain.

Other than this, some ATP-dependent RibKs exhibit versatility, being capable of utilizing other nucleotides along with ATP to varying extents. For example, in RibK homolog from the rat intestinal mucosa, UTP is comparably effective to ATP in transferring phosphate to riboflavin, while CTP and GTP are poor phosphate donors, and similarly, in *Lactobacillus arabinosus*, UTP shows 60 % activity in comparison to ATP while no activity was observed with other nucleotides.[25,26] The ATP-dependent RibK (ribR) from *Bacillus subtilis* can also utilize CTP (50%), UTP (31%), and dATP (54%) for its function, while a reducing environment is necessary for the phosphorylation reaction.[27] Another thesis from our lab also reported the utilization of other nucleotides (CTP, GTP, and UTP) by RibKs from eukaryotes named *Homo sapiens*, *Schizosaccharomyces pombe*, and *Neurospora crassa,* which are ATP-dependent.[28–31] Besides this, even the ADP can also be used as a phosphate donor, as shown by the RibK from *Glycine max* (soybean) and *Peptostreptococcus elsdenii* which catalysed the reactions with 10 % and 22 % efficiency relative to ATP, respectively.[32,33]

Table 2.1 List of some riboflavin kinases from various organisms and their nucleotide selectivity. This table has been adapted from a thesis by Kumar, Yashwant in our lab.[28]

Organism	ATP	GTP	CTP	UTP
Bacillus subtilis[27]	100	1	50	31
Escherichia coli[34]	100	-	-	-
Streptococcus agalactiae[35]	100	-	-	-
Peptostreptococcus elsdenii[33]	100	-	-	-
Lactobacillus arabinosus[26]	100	0	0	60
Schizosaccharomyces pombe[31]	100	AD*	AD*	AD*
Saccharomyces cerevisiae[36]	100	-	-	-
Mus musculus (from rat liver)[13]	100	>1	>1	-
Mus musculus (from intestinal mucosa)[25]	100	PD*	PD*	100
Homo sapiens[30]	100	AD*	AD*	AD*
Neurospora crassa[29]	100	AD*	AD*	AD*
Rhizopus oryzae (cell-free lysate)[23]	0	100	0	36
Rhizopus javanicus (cell-free lysate)[22]	0	100	0	AD*
Glycine max (Soybean)[32]	100	0	0	0
Vigna radiata (Mung bean)[37]	100	-	-	-
Thermoplasma acidophilum[24]	-	-	100	-
Methanocaldococcus jannaschii[20]	-	-	100	-
Methanococcus maripaludis[21]	-	-	100	-
Note: *N*umbers represent % activity. PD: Poor donor; AD: Active donor; *: values of activity not reported				

2.3 Promiscuity of the FMN adenylyl transferase enzymes (or modules) for other NTPs and corresponding FAD nucleobase analogues

The alteration of RibK's promiscuity towards different NTPs does not impact the product outcome, whereas employing NTPs other than ATP in FMNAT enzymes will lead to the creation of nucleobase analogous of FAD known as FND (FGD, FCD, and FUD). Nevertheless, these analogues of FAD do not occur naturally in organisms, suggesting that these enzymes strongly prefer ATP over other NTPs. One ATP-dependent archaeal FMNAT from *Methanocaldococcus jannaschii* (*Mj*FMNAT) has shown *in vitro* activity with CTP and GTP, resulting in FAD analogues FCD and FGD.[38] Nevertheless, the synthesis of these analogues inside a cell remains unexplored, and their use in cell physiology is a mystery. However, substitution of adenine nucleobase with others should not change the redox property of isoalloxazine ring in FAD as the the isoalloxazine and adenine rings typically adopt an

extended conformation, positioned around 11-16 Å apart (measured between terminal backbone atoms) when bound to flavoenzymes.[39]

This thesis is aimed at the biosynthesis of FND analogues, and the following questions were tried to be answered:

1. *Can the nucleotide specificity of the FMNAT enzyme (or module) be engineered, resulting in the FAD nucleobase analogues?*
2. *Can these engineered FMNAT mutants synthesize the FAD analogues inside the cell?*
3. *Do these analogues have any significant role to play in cell physiology?*

2.4 The mechanistic and structural understanding of FMNAT and FAD synthetase enzymes

The evolutionary connection between bacterial and eukaryotic FMNATs remains unclear due to their limited sequence similarity and distinct classification in protein databases like SCOP and Pfam.[40–44] The FMNAT module in the N-terminal of bifunctional FAD synthetase is categorized within the HxGH motif-containing nucleotidylyl transferase superfamily. Conversely, the FMNATs from eukaryotes belong to the family like 3′-phosphoadenosine 5′-phosphosulfate reductases, which belong to a superfamily of the adenine nucleotide α hydrolases. Structurally, FMNATs from bacteria and eukaryotes can differentiated evidently by their substrate-binding mechanisms and active site configurations.[45] The bacterial FMNATs display signature motifs, setting them apart from the conformations seen in eukaryotic FMNATs. The conserved motifs underpinning catalytic mechanisms also differ. Bacterial FMNATs feature motifs like HxGH and ISSTxxR, interacting with the β and γ phosphates of ATP nucleotide.[45,46] Eukaryotic FMNATs possess no such equivalent motifs serving comparable functions. Furthermore, analysis of product inhibition unveiled marked distinctions where the yeast *Candida glabrata* FMNAT exhibited robust Inhibition by FAD (a reaction product), while its D181A mutant experienced reduced product-based inhibition.[47]

The FMNAT found in *M. jannaschii* also possesses the HxGH motif.[38] The flavin binding site of *Mj*FMNAT is distinct and differs from the flavin binding site found in non-archaeal counterparts. While the gene responsible for encoding *Mj*FMNAT is conserved among

archaea, the absence of the crystal structure for the archaeal FMNAT enzyme has hindered a thorough investigation of their mechanism of substrate-binding and catalysis.

The *C. ammoniagenes* FAD synthetase (*Ca*FADS) has undergone extensive research as a representative of the prokaryotic FADS family. The *Ca*FADS and its various mutants have crystal structures available within the PDB database and have been thoroughly investigated.[14,48–52] This enzyme comprises the RibK (C-terminal) and FMNAT (N-terminal) modules. The structural analysis indicates an arrangement of a dimer of trimers, where each trimer demonstrates a head-to-tail configuration of protomers. This arrangement places the catalytic sites of RibK and FMNAT from adjacent protomers in close proximity.[14,49–51] This spatial arrangement, in accordance with active site residues, leads to the interdependency of one module's activity on the other.[52,53] The independently expressed C-terminal module can drive the phosphorylation process, while the N-terminal module alone does not display any adenylation reaction.[48] The activity of the truncated RibK module closely resembles that of the complete enzyme. Additionally, the RibK module displays inhibition by an excess of its substrate RF and by the excess of FMN production, operating independently of the FMNAT module.[54] The steady-state kinetic analysis of *Ca*FADS proposed that even though FMN serves as a bridge in the synthesis of FAD from riboflavin, it must be unbound first from the RibK module, followed by incorporation in the FMNAT module for the adenylation.[55] Moreover, the FMNAT module's adenylation process was found to be reversible, forming FMN and ATP from FAD and pyrophosphate. Besides this, *Ca*FADS has also been used as a catalyst for green chemistry applications, facilitating the synthesis of numerous isoalloxazine-based FAD analogous.[56,57] However, due to its high specificity for ATP, it has not succeeded in generating any FAD nucleobase analogous.[11]

2.5 Our choice of *Escherichia coli* FAD synthetase (*Ec*FADS) to probe the biosynthesis of FAD analogues

GTP and ribulose-5-phosphate are the precursor of riboflavin biosynthesis which is further converted to FMN and FAD. In prokaryotes, riboflavin transportation and biosynthesis is regulated by FMN riboswitch which directly binds to FMN.[58] The overproduction of riboflavin, FMN, and FAD in *E. coli* has already been demonstrated by overexpressing RF biosynthesis genes in low copy number plasmid and *ribF* gene in high copy number plasmid.[59] We opted to work with *Ec*FADS, a homolog of *Ca*FADS, which is mechanistically and

structurally well-explored in the literature. A technique patented in 1992 facilitated flavin nucleotide production using *Ec*FADS.[60] In this method, the G23 residue was modified with arginine, alanine, or aspartic acid, effectively terminating the FMNAT module activity within the recombinant DNA. Consequently, microorganism cultures containing this modified DNA yielded FMN exclusively. Another investigation involving *Ec*FADS exhibited its compatibility with the antibiotic roseoflavin, which is a riboflavin analogue.[61] This interaction led to the generation of roseoflavin versions of FMN and FAD, called roFMN and roFAD. Nevertheless, the *Ec*FADS enzyme has not been much explored except in some of these instances. Yet, these factors made us choose *E. coli* as this thesis involves the genomic substitution of FAD synthetase with its engineered mutant for the biosynthesis of FAD nucleobase analogues to study their impact on cell physiology and potential applications; the comprehensively studied genetic manipulation and metabolomics of *E. coli* genetics make the research easy.

We initiated an exploration of the molecular factors governing nucleotide base specificity within the FMNAT module of *Ec*FADS by synthesising nucleobase analogues of the FAD cofactor using chemical and enzymatic methods. The enzyme engineering approach was employed to introduce the FAD analogue in the cell by utilizing other NTPs. For this, the earlier approaches were considered, such as introducing charged residues based on the computational studies, tuning the size of the active site, and structural homology discussed in Chapter 1.

The forthcoming overview outlines the sections of this thesis that delve into our inquiries regarding the *in vivo* synthesis of FAD nucleobase analogues and the utilisation of these analogues in cell physiology.

2.5.1 Exploring the FAD nucleobase analogues synthesis via chemical and enzymatic approach and their potential as enzyme cofactors - Chapter 3: We demonstrated the successful synthesis of FAD nucleobases analogous by employing two distinct methodologies – a chemical synthesis pathway as well as the enzymatic approach. These approaches involve not more than three steps, yielding in the range from 10% to 57%. Subsequently, we evaluated the *E. coli* glutathione reductase for its binding capability and performance in conjunction with these analogous. Our findings indicated that all analogous exhibit binding capacity and can function as reducing agents similar to FAD. Lastly, considering that NTPs and riboflavin are inherent substrates within cellular systems, we undertook the heterologous expression of the versatile *Mj*FMNAT, demonstrating the potential for the intracellular synthesis of FAD nucleoside analogous.

2.5.2 An enzyme engineering approach to the biosynthesis of FAD nucleobase analogues in *Escherichia coli*- Chapter 4: This investigation has demonstrated that the RKF module within *Ec*FADS predominantly exhibits specificity towards ATP, whereas the FMNAT module displays versatility for various NTPs, excluding GTP. Leveraging computational analysis and structural homology, we engaged in enzyme engineering of the FMNAT module within *Ec*FADS. Through a strategic swap of the loop neighbouring the adenine base of ATP in the FMNAT module with a corresponding loop from a cytidylyl transferase, we augmented its capacity for accommodating multiple NTPs, including GTP. Notably, upon heterologous expression of this LoopSwap mutant, the production of FAD analogous was achieved within the cellular environment. Additionally, through loop shortening, we established the role of this loop in conferring ATP specificity within the cell.

2.5.3 Biosynthesis of FAD analogues and their role in cell physiology- Chapter 5: This chapter presents the genomic substitution of the essential *Ec*FADS-expressing gene, *ribF*, with the LoopSwap mutant expressing *ribF25* mutant gene, resulting in the creation of the *ribF25* mutant strain. Remarkably, this *ribF25* mutant strain exhibits comparable survival and growth rates to the wild-type strain and synthesizes FAD analogous (FGD, FCD, and FUD). Intriguingly, the mutant strain showcases tolerance against aminoglycoside and beta-lactam antibiotic stress. Under anaerobic growth conditions, the mutant strain experiences slower growth, attributed to the reduced concentration of FAD compared to the wild-type strain. Nevertheless, the mutant strain continues to display heightened responsiveness to antibiotic stress.

2.5.4 Findings, implications, and future directions- Chapter 6: This chapter has focused on providing a comprehensive overview of adenine recognition, considering the adenine moieties present in ATP, FAD, NAD, CoA, and SAM. Based on prior literature and our research, we propose that adenine binding likely emerged independently on multiple occasions during the course of evolution, indicative of convergent evolution processes.

References

(1) Barile, M.; Anna Giancaspero, T.; Brizio, C.; Panebianco, C.; Indiveri, C.; Galluccio, M.; Vergani, L.; Eberini, I.; Gianazza, E. Biosynthesis of Flavin Cofactors in Man: Implications in Health and Disease. *Curr Pharm Des* **2013**, *19* (14), 2649–2675. https://doi.org/10.2174/1381612811319140014.

(2) Massey, V. The Chemical and Biological Versatility of Riboflavin. *Biochem Soc Trans* **2000**, *28* (4), 283–296. https://doi.org/10.1042/bst0280283.

(3) Tomiki, T.; Saitou, N. Phylogenetic Analysis of Proteins Associated in the Four Major Energy Metabolism Systems: Photosynthesis, Aerobic Respiration, Denitrification, and Sulfur Respiration. *J Mol Evol* **2004**, *59* (2), 158–176. https://doi.org/10.1007/s00239-004-2610-2.

(4) Bender, D. A. NIACIN | Physiology. In *Encyclopedia of Food Sciences and Nutrition*; Elsevier, 2003; pp 4119–4128. https://doi.org/10.1016/B0-12-227055-X/00827-0.

(5) McCormick, D. B. Two Interconnected B Vitamins: Riboflavin and Pyridoxine. *Physiol Rev* **1989**, *69* (4), 1170–1198. https://doi.org/10.1152/physrev.1989.69.4.1170.

(6) Reddi, A. S. Riboflavin Nutritional Status and Flavoprotein Enzymes in Streptozotocin-Diabetic Rats. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1986**, *882* (1), 71–76. https://doi.org/10.1016/0304-4165(86)90057-7.

(7) Saedisomeolia, A.; Ashoori, M. Riboflavin in Human Health: A Review of Current Evidences. In *Advances in Food and Nutrition Research*; 2018; Vol. 83, pp 57–81. https://doi.org/10.1016/bs.afnr.2017.11.002.

(8) Marashly, E. T.; Bohlega, S. A. Riboflavin Has Neuroprotective Potential: Focus on Parkinson's Disease and Migraine. *Front Neurol* **2017**, *8* (333). https://doi.org/10.3389/fneur.2017.00333.

(9) Sinha, T.; Naash, M. I.; Al-Ubaidi, M. R. Flavins Act as a Critical Liaison Between Metabolic Homeostasis and Oxidative Stress in the Retina. *Front Cell Dev Biol* **2020**, *8* (861). https://doi.org/10.3389/fcell.2020.00861.

(10) Mereis, M.; Wanders, R. J. A.; Schoonen, M.; Dercksen, M.; Smuts, I.; van der Westhuizen, F. H. Disorders of Flavin Adenine Dinucleotide Metabolism: MADD and Related Deficiencies. *Int J Biochem Cell Biol* **2021**, *132*, 105899. https://doi.org/10.1016/j.biocel.2020.105899.

(11) Manstein, D. J.; Pai, E. F. Purification and Characterization of FAD Synthetase from Brevibacterium Ammoniagenes. *Journal of Biological Chemistry* **1986**, *261* (34), 16169–16173. https://doi.org/10.1016/S0021-9258(18)66693-1.

(12) Barile, M.; Brizio, C.; Valenti, D.; De Virgilio, C.; Passarella, S. The Riboflavin/FAD Cycle in Rat Liver Mitochondria. *Eur J Biochem* **2000**, *267* (15), 4888–4900. https://doi.org/10.1046/j.1432-1327.2000.01552.x.

(13) Merrill, A. H.; McCormick, D. B. Affinity Chromatographic Purification and Properties of Flavokinase (ATP:Riboflavin 5'-Phosphotransferase) from Rat Liver. *Journal of Biological Chemistry* **1980**, *255* (4), 1335–1338. https://doi.org/10.1016/S0021-9258(19)86034-9.

(14) Herguedas, B.; Martínez-Júlvez, M.; Frago, S.; Medina, M.; Hermoso, J. A. Oligomeric State in the Crystal Structure of Modular FAD Synthetase Provides Insights into Its Sequential Catalysis in Prokaryotes. *J Mol Biol* **2010**, *400* (2), 218–230. https://doi.org/10.1016/j.jmb.2010.05.018.

(15) Serrano, A.; Frago, S.; Herguedas, B.; Martínez-Júlvez, M.; Velázquez-Campoy, A.; Medina, M. Key Residues at the Riboflavin Kinase Catalytic Site of the Bifunctional Riboflavin Kinase/FMN Adenylyltransferase From Corynebacterium Ammoniagenes. *Cell Biochem Biophys* **2013**, *65* (1), 57–68. https://doi.org/10.1007/s12013-012-9403-9.

(16) Huerta, C.; Borek, D.; Machius, M.; Grishin, N. V.; Zhang, H. Structure and Mechanism of a Eukaryotic FMN Adenylyltransferase. *J Mol Biol* **2009**, *389* (2), 388–400. https://doi.org/10.1016/j.jmb.2009.04.022.

(17) Serrano, A.; Ferreira, P.; Martinez-Julvez, M.; Medina, M. The Prokaryotic FAD Synthetase Family: A Potential Drug Target. *Curr Pharm Des* **2013**, *19* (14), 2637–2648. https://doi.org/10.2174/1381612811319140013.

(18) Sebastián, M.; Anoz-Carbonell, E.; Gracia, B.; Cossio, P.; Aínsa, J. A.; Lans, I.; Medina, M. Discovery of Antimicrobial Compounds Targeting Bacterial Type FAD Synthetases. *J Enzyme Inhib Med Chem* **2018**, *33* (1), 241–254. https://doi.org/10.1080/14756366.2017.1411910.

(19) Ammelburg, M.; Hartmann, M. D.; Djuranovic, S.; Alva, V.; Koretke, K. K.; Martin, J.; Sauer, G.; Truffault, V.; Zeth, K.; Lupas, A. N.; Coles, M. A CTP-Dependent Archaeal Riboflavin Kinase Forms a Bridge in the Evolution of Cradle-Loop Barrels. *Structure* **2007**, *15* (12), 1577–1590. https://doi.org/10.1016/j.str.2007.09.027.

(20) Mashhadi, Z.; Zhang, H.; Xu, H.; White, R. H. Identification and Characterization of an Archaeon-Specific Riboflavin Kinase. *J Bacteriol* **2008**, *190* (7), 2615–2618. https://doi.org/10.1128/JB.01900-07.

(21) Kumar, Y.; Singh, R. K.; Hazra, A. B. Characterization of a Novel Mesophilic CTP-Dependent Riboflavin Kinase and Rational Engineering to Create Its Thermostable Homologues**. *ChemBioChem* **2021**, *22* (24), 3414–3424. https://doi.org/10.1002/cbic.202100211.

(22) Tachibana, S. GTP-Dependent Flavin Phosphate Synthetase. In *Methods in Enzymology*; 1971; pp 553–555. https://doi.org/10.1016/S0076-6879(71)18118-9.

(23) TACHIBANA, S. DISCOVERY OF A NEW FLAVIN PHOSPHATE SYNTHETASE WHICH REQUIRES GUANOSINE-5′-TRIPHOSPHATE. *J Vitaminol (Kyoto)* **1967**, *13* (2), 89–92. https://doi.org/10.5925/jnsv1954.13.89.

(24) Rodionova, I. A.; Vetting, M. W.; Li, X.; Almo, S. C.; Osterman, A. L.; Rodionov, D. A. A Novel Bifunctional Transcriptional Regulator of Riboflavin Metabolism in Archaea. *Nucleic Acids Res* **2017**, *45* (7), 3785–3799. https://doi.org/10.1093/nar/gkw1331.

(25) Kasai, S.; Nakano, H.; Maeda, K.; Matsui, K. Purification, Properties, and Function of Flavokinase from Rat Intestinal Mucosa. *The Journal of Biochemistry* **1990**, *107* (2), 298–303. https://doi.org/10.1093/oxfordjournals.jbchem.a123042.

(26) Snoswell, A. FLAVOKINASE OF LACTOBACILLUS ARABINOSUS 17.5. *Australian Journal of Experimental Biology and Medical Science* **1957**, *35* (5), 427–436. https://doi.org/10.1038/icb.1957.45.

(27) Solovieva, I. M.; Tarasov, K. V.; Perumov, D. A. Main Physicochemical Features of Monofunctional Flavokinase from Bacillus Subtilis. *Biochemistry (Moscow)* **2003**, *68* (2), 177–181. https://doi.org/10.1023/A:1022645327972.

(28) Kumar, Y. Molecular Determinants of Stability and Nucleobase Specificity in Flavin Biosynthesis Enzymes. Doctoral Thesis, IISER Pune, 2021.

(29) Rajeswari, S. R.; Jonnalagadda, V. S.; Jonnalagadda, S. Purification and Characterization of Flavokinase from Neurospora Crassa. *Indian J Biochem Biophys* **1999**, *36* (3), 137–142.

(30) Karthikeyan, S.; Zhou, Q.; Mseeh, F.; Grishin, N. V.; Osterman, A. L.; Zhang, H. Crystal Structure of Human Riboflavin Kinase Reveals a β Barrel Fold and a Novel Active Site Arch. *Structure* **2003**, *11* (3), 265–273. https://doi.org/10.1016/S0969-2126(03)00024-8.

(31) Bauer, S.; Kemter, K.; Bacher, A.; Huber, R.; Fischer, M.; Steinbacher, S. Crystal Structure of Schizosaccharomyces Pombe Riboflavin Kinase Reveals a Novel ATP and Riboflavin-Binding Fold. *J Mol Biol* **2003**, *326* (5), 1463–1473. https://doi.org/10.1016/S0022-2836(03)00059-7.

(32) MITSUDA, H.; TOMOZAWA, Y.; KAWAI, F. STUDIES ON PLANT FLAVOKINASE. *J Vitaminol (Kyoto)* **1963**, *9* (2), 142–148. https://doi.org/10.5925/jnsv1954.9.142.

(33) Mayhew, S. G.; Wassink, J. H. A Continuous Fluorometric Assay for Flavokinase Properties of Flavokinase from Peptostreptococcus Elsdenii. *Biochimica et Biophysica Acta (BBA) - Enzymology* **1977**, *482* (2), 341–347. https://doi.org/10.1016/0005-2744(77)90247-9.

(34) KATAGIRI, H.; YAMADA, H.; IMAI, K. BIOSYNTHESIS OF FLAVIN COENZYMES BY MICROORGANISMS. *J Vitaminol (Kyoto)* **1959**, *5* (2), 129–133. https://doi.org/10.5925/jnsv1954.5.129.

(35) Clarebout, G.; Villers, C.; Leclercq, R. Macrolide Resistance Gene *MreA* of *Streptococcus Agalactiae* Encodes a Flavokinase. *Antimicrob Agents Chemother* **2001**, *45* (8), 2280–2286. https://doi.org/10.1128/AAC.45.8.2280-2286.2001.

(36) Santos, M. A.; Jiménez, A.; Revuelta, JoséL. Molecular Characterization of FMN1, the Structural Gene for the Monofunctional Flavokinase of Saccharomyces Cerevisiae. *Journal of Biological Chemistry* **2000**, *275* (37), 28618–28624. https://doi.org/10.1074/jbc.M004621200.

(37) Das-Panja, K.; Jonnalagadda, V.; Jonnalagadda, S. Orthophosphate Is a Non-Essential Activator of Vigna Radiata Flavokinase. *IUBMB Life* **1999**, *47* (4), 547–554. https://doi.org/10.1080/15216549900201583.

(38) Mashhadi, Z.; Xu, H.; Grochowski, L. L.; White, R. H. Archaeal RibL: A New FAD Synthetase That Is Air Sensitive. *Biochemistry* **2010**, *49* (40), 8748–8755. https://doi.org/10.1021/bi100817q.

(39) Kuppuraj, G.; Kruise, D.; Yura, K. Conformational Behavior of Flavin Adenine Dinucleotide: Conserved Stereochemistry in Bound and Free States. *J Phys Chem B* **2014**, *118* (47), 13486–13497. https://doi.org/10.1021/jp507629n.

(40) Fox, N. K.; Brenner, S. E.; Chandonia, J.-M. SCOPe: Structural Classification of Proteins—Extended, Integrating SCOP and ASTRAL Data and Classification of New Structures. *Nucleic Acids Res* **2014**, *42* (D1), 304–309. https://doi.org/10.1093/nar/gkt1240.

(41) Sebastián, M.; Arilla-Luna, S.; Bellalou, J.; Yruela, I.; Medina, M. The Biosynthesis of Flavin Cofactors in Listeria Monocytogenes. *J Mol Biol* **2019**, *431* (15), 2762–2776. https://doi.org/10.1016/j.jmb.2019.05.029.

(42) Leone, P.; Galluccio, M.; Brizio, C.; Barbiroli, A.; Iametti, S.; Indiveri, C.; Barile, M. The Hidden Side of the Human FAD Synthase 2. *Int J Biol Macromol* **2019**, *138*, 986–995. https://doi.org/10.1016/j.ijbiomac.2019.07.138.

(43) Anoz-Carbonell, E.; Rivero, M.; Polo, V.; Velázquez-Campoy, A.; Medina, M. Human Riboflavin Kinase: Species-specific Traits in the Biosynthesis of the FMN Cofactor. *The FASEB Journal* **2020**, *34* (8), 10871–10886. https://doi.org/10.1096/fj.202000566R.

(44) Lohithakshan, A.; Narayanasamy, R.; Potteth, U. S.; Keshava, S.; Nagaraja, V.; Usharani, D.; Kumar, R. Molecular Insights into the Mechanism of Substrate Binding and Catalysis of Bifunctional FAD Synthetase from Staphylococcus Aureus. *Biochimie* **2021**, *182*, 217–227. https://doi.org/10.1016/j.biochi.2021.01.013.

(45) Huerta, C.; Borek, D.; Machius, M.; Grishin, N. V.; Zhang, H. Structure and Mechanism of a Eukaryotic FMN Adenylyltransferase. *J Mol Biol* **2009**, *389* (2), 388–400. https://doi.org/10.1016/j.jmb.2009.04.022.

(46) Wang, W.; Kim, R.; Yokota, H.; Kim, S.-H. Crystal Structure of Flavin Binding to FAD Synthetase of Thermotoga Maritima. *Proteins: Structure, Function, and Bioinformatics* **2004**, *58* (1), 246–248. https://doi.org/10.1002/prot.20207.

(47) Huerta, C.; Grishin, N. V.; Zhang, H. The "Super Mutant" of Yeast FMN Adenylyltransferase Enhances the Enzyme Turnover Rate by Attenuating Product Inhibition. *Biochemistry* **2013**, *52* (21), 3615–3617. https://doi.org/10.1021/bi400454w.

(48) Frago, S.; Martínez-Júlvez, M.; Serrano, A.; Medina, M. Structural Analysis of FAD Synthetase from Corynebacterium Ammoniagenes. *BMC Microbiol* **2008**, *8* (1), 160. https://doi.org/10.1186/1471-2180-8-160.

(49) Herguedas, B.; Martínez-Júlvez, M.; Frago, S.; Medina, M.; Hermoso, J. A. Crystallization and Preliminary X-Ray Diffraction Studies of FAD Synthetase from *Corynebacterium Ammoniagenes*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **2009**, *65* (12), 1285–1288. https://doi.org/10.1107/S1744309109044789.

(50) Marcuello, C.; Arilla-Luna, S.; Medina, M.; Lostao, A. Detection of a Quaternary Organization into Dimer of Trimers of Corynebacterium Ammoniagenes FAD Synthetase at the Single-Molecule Level and at the in Cell Level. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2013**, *1834* (3), 665–676. https://doi.org/10.1016/j.bbapap.2012.12.013.

(51) Serrano, A.; Sebastián, M.; Arilla-Luna, S.; Baquedano, S.; Herguedas, B.; Velázquez-Campoy, A.; Martínez-Júlvez, M.; Medina, M. The Trimer Interface in the Quaternary Structure of the Bifunctional Prokaryotic FAD Synthetase from Corynebacterium Ammoniagenes. *Sci Rep* **2017**, *7* (1), 404. https://doi.org/10.1038/s41598-017-00402-6.

(52) Lans, I.; Seco, J.; Serrano, A.; Burbano, R.; Cossio, P.; Daza, M. C.; Medina, M. The Dimer-of-Trimers Assembly Prevents Catalysis at the Transferase Site of Prokaryotic FAD Synthase. *Biophys J* **2018**, *115* (6), 988–995. https://doi.org/10.1016/j.bpj.2018.08.011.

(53) Serrano, A.; Sebastián, M.; Arilla-Luna, S.; Baquedano, S.; Pallarés, M. C.; Lostao, A.; Herguedas, B.; Velázquez-Campoy, A.; Martínez-Júlvez, M.; Medina, M. Quaternary Organization in a Bifunctional Prokaryotic FAD Synthetase: Involvement of an Arginine at Its Adenylyltransferase Module on the Riboflavin Kinase Activity. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2015**, *1854* (8), 897–906. https://doi.org/10.1016/j.bbapap.2015.03.005.

(54) Sebastián, M.; Serrano, A.; Velázquez-Campoy, A.; Medina, M. Kinetics and Thermodynamics of the Protein-Ligand Interactions in the Riboflavin Kinase Activity of the FAD Synthetase from Corynebacterium Ammoniagenes. *Sci Rep* **2017**, *7* (1), 7281. https://doi.org/10.1038/s41598-017-07875-5.

(55) Efimov, I.; Kuusk, V.; Zhang, X.; McIntire, W. S. Proposed Steady-State Kinetic Mechanism for *Corynebacterium Ammoniagenes* FAD Synthetase Produced by *Escherichia Coli*. *Biochemistry* **1998**, *37* (27), 9716–9723. https://doi.org/10.1021/bi972817j.

(56) Spencer, R.; Fisher, J.; Walsh, C. Preparation, Characterization, and Chemical Properties of the Flavin Coenzyme Analogues 5-Deazariboflavin, 5-Deazariboflavin 5'-Phosphate, and 5-Deazariboflavin 5'-Diphosphate, 5' → 5'-Adenosine Ester. *Biochemistry* **1976**, *15* (5), 1043–1053. https://doi.org/10.1021/bi00650a015.

(57) Mishanina, T. V.; Kohen, A. Synthesis and Application of Isotopically Labeled Flavin Nucleotides. *J Labelled Comp Radiopharm* **2015**, *58* (9), 370–375. https://doi.org/10.1002/jlcr.3313.

(58) Pedrolli, D.; Langer, S.; Hobl, B.; Schwarz, J.; Hashimoto, M.; Mack, M. The *RibB* FMN Riboswitch from *Escherichia Coli* Operates at the Transcriptional and Translational Level and Regulates Riboflavin Biosynthesis. *FEBS J* **2015**, *282* (16), 3230–3242. https://doi.org/10.1111/febs.13226.

(59) Liu, S.; Diao, N.; Wang, Z.; Lu, W.; Tang, Y.-J.; Chen, T. Modular Engineering of the Flavin Pathway in *Escherichia Coli* for Improved Flavin Mononucleotide and Flavin Adenine Dinucleotide Production. *J Agric Food Chem* **2019**, *67* (23), 6532–6540. https://doi.org/10.1021/acs.jafc.9b02646.

(60) Kitatsuji, K.; Ishino, S.; Teshiba, S.; Arimoto, M. Process for Producing Flavine Nucleotides. European Patent Office, 1992.

(61) Langer, S.; Hashimoto, M.; Hobl, B.; Mathes, T.; Mack, M. Flavoproteins Are Potential Targets for the Antibiotic Roseoflavin in Escherichia Coli. *J Bacteriol* **2013**, *195* (18), 4037–4045. https://doi.org/10.1128/JB.00646-13.

Chapter 3

Exploring the FAD nucleobase analogues synthesis via chemical and enzymatic approach and their potential as enzyme cofactors

3.1 Introduction

The coenzyme Flavin adenine dinucleotide (FAD) is crucial for cell metabolism, aiding various flavoenzymes in facilitating the redox chemistry of the cell. FAD is essential for various organisms, including humans, and is synthesized from riboflavin (vitamin B_2) through an enzymatic two-step process that requires an adenosine triphosphate (ATP) molecule in each step. Initially, the enzyme riboflavin kinase employs ATP to phosphorylate riboflavin, forming flavin mononucleotide (FMN). Subsequently, the succeeding enzyme, FMN adenylyl transferase (FMNAT), employs an additional ATP molecule to convert FMN into FAD (Figure 3.1A).[1–3] The isoalloxazine moiety is responsible for FAD's redox potential, while the two linked phosphate groups, the ribityl chain, and the adenosine moiety, are involved in enzyme binding and molecular recognition.[4–7] Particularly, the adenosine group recurs in other cofactors like coenzyme A, SAM, and NAD, suggesting a parallel role in recognition and binding.[5]

Due to the essential significance of FAD in the cell, various analogues were developed to explore its functional and structural characteristics and serve as investigative tools for understanding flavoenzyme functionality. These FAD analogues primarily fall into three major categories: those with alterations in the isoalloxazine moiety, ribityl side chain, and the adenosine component (Figure 3.1A). Variants of FAD containing fluorine, thiol, or other modifications in the isoalloxazine ring were employed to investigate the active site environment of proteins that bind flavins.[8,9] Additionally, two antibacterial agents, roseoflavin and 5-FDQD, exhibit noteworthy pharmacological impacts by interacting with FMN riboswitches.[10,11] Moreover, various halogenated isoalloxazine ring derivatives demonstrate antimalarial effects by inhibiting glutathione reductase.[12] The ribityl chain analogues, such as 2′-deoxy-FAD, arabino-FAD, and 2′-F-arabino-FAD, exhibit modified catalytic activities when they are engaged within the enzyme's binding site, such as mercuric reductase, glutathione reductase, and lipoamide dehydrogenase.[13] Conversely, nucleoside analogues of FAD have not received much comprehensive investigation. A particular photophysical study delves into the nucleoside's impact on FAD's structural and spectroscopic characteristics. This study employs

analogues of FAD involving substituted with purine or pyrimidine ribonucleosides/deoxyribonucleosides, and it suggests that purine analogues demonstrate a more pronounced intramolecular complex formation between the isoalloxazine and nucleobase.[14] Very few additional investigations have focused on evaluating the *in vitro* activity of diverse analogues that have undergone adenine modifications in conjunction with D-amino acid oxidase.[15–19] Nonetheless, the impact of FAD's nucleoside substitutions, particularly concerning their involvement in use or biosynthesis within the enzyme or cells, remains relatively underexplored.

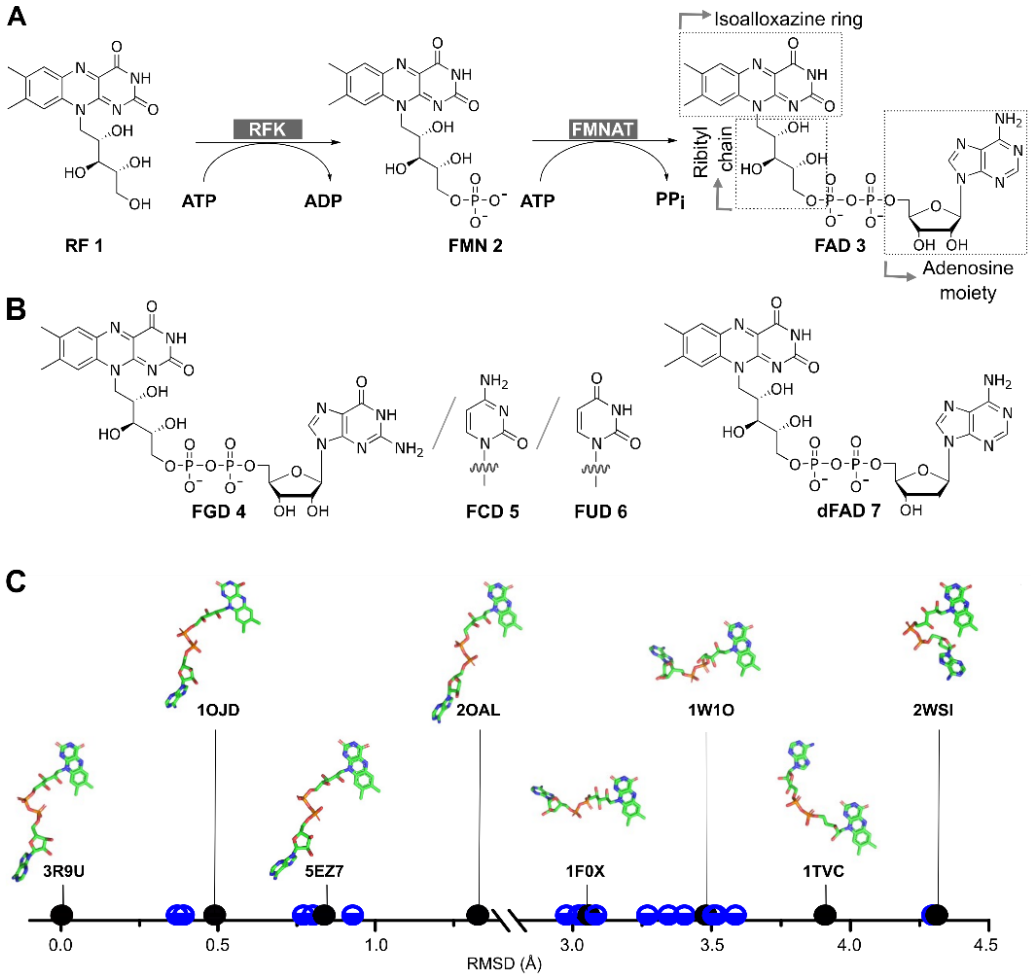


Figure 3.1: The biosynthesis of FAD and its nucleoside analogues in the cell followed by conformational analysis of FAD in flavoproteins. (A) flavin mononucleotide (FMN) is synthesized initially from Riboflavin (RF) by the enzyme riboflavin kinase (RibK). Subsequently, FMN is adenylated to create flavin adenine dinucleotide (FAD) through the action of FMN adenylyl transferase (FMNAT). Both of these enzymatic steps involve the utilization of adenosine triphosphate (ATP) as the secondary substrate. Modifications to the isoalloxazine ring, ribityl chain, or adenosine component can lead to the synthesis of various FAD analogues. **(B)** Structures of FAD nucleoside analogues based on the adenine/adenosine substitution by guanine, cytosine, uracil, and deoxyadenosine, resulting in FGD, FCD, FUD, and dFAD, respectively. These FAD analogues do not naturally exist within cellular systems, while the FMN and NTP substrates are already present in cells. Therefore, it is theoretically possible to generate this array of FAD nucleobase/nucleoside analogues within cells by utilizing appropriate RibK/FMNAT mutant enzymes. **(C)** A comprehensive analysis to understand the structural

conformation of FAD within FAD-bound proteins. 24 FAD-bound crystal structures from various flavoenzymes were taken randomly, and Root Mean Square Deviation (RMSD) was computed for each bounded FAD, using 3R9U as an arbitrarily chosen standard with an RMSD of 0.0 Å where 3R9U correspond to *Campylobacter jejuni* thioredoxin-disulfide reductase (as indicated in Table 3.2). Despite the distinct RMSD values observed in the FAD-bound conformations within different enzymes, they all share an elongated arrangement, with the isoalloxazine and adenine rings being primarily separated by around 11-16 Å. The representation showcases the RMSD values of 8 select structures using black-filled circles, highlighting the separation of the two rings, while the blue circles depict the RMSD values of the remaining 16 structures that were chosen for the analysis. Figure adapted from *ChemBioChem* (**2023**), Shah *et al*.

The literature contains various inventive methodologies for synthesizing FAD analogues, yet these approaches have certain constraints, including multistep processes, low yields, and/or challenges in obtaining initial materials. The initial synthesis of FAD was performed by coupling diverse metal salts of FMN with an adenosine-5'-benzyl phosphorochloridate derivative, followed by subsequent deprotection of the protective groups.[20] The original approach suffered from a slow process involving multiple steps, resulting in a notably lower product yield (approximately 6%). Using trifluoroacetic acid anhydride (TFAA) as the coupling agent, a more efficient method was developed as a refinement, involving a one-step coupling of FMN salts with adenosine monophosphate (AMP).[21] This approach, employed for thFAD nucleobase analogue synthesis, suffers from remarkably low yields, typically ranging between 0.5% to 1%, due to self-condensation involving the substrates.[15] However, despite its limited yield, this one-step method has been employed to generate nucleobase FAD and NAD analogues for some instances.[22,23] In a similar approach, carbodiimide was employed as a coupling agent to synthesize FAD from AMP and FMN. Nonetheless, this modification did not significantly improve the yields, which continued to be low.[16] An attempt to enhance the yield entailed activating the phosphate group of one substrate to a phosphoramidate or phosphoromorpholidate and then its coupling with the second nucleotide substrate.[14,17,24,25] This method resulted in an improved yield within the range of 25% to 40%. However, due to the limited reactivity of the phosphoromorpholidate or phosphoramidate, the duration of the reaction is extended from 4 to 7 days. When diphenyl phosphorochloridate was used as an activator of monophosphate to synthesize the FAD, it significantly increased the product to 70-80% yield, accomplishing the reaction in a few hours.[26] Nonetheless, when this method was applied to synthesize FAD and NAD analogues, the yield decreased to 25-30%.[18,19,27] An alternate approach involves employing riboflavin phosphorothioate along with the silver salt of AMP, leading to a yield of 51% achieved within a 5-hour timeframe.[28] However, the drawback of this method is to obtain the starting material,

which is unavailable commercially. Interestingly, a method for synthesizing NAD nucleobase analogues has been achieved by activating nicotinamide mononucleotide phosphate using the redox pair triphenylphosphine/2,2′-dipyridyl disulfide ($Ph_3P/(PyS)_2$) with the assistance of a nucleophilic catalyst, 1-methylimidazole, within a solvent mixture of DMSO/DMF, followed by subsequent reaction with the other mononucleotide.[29] However, this approach cannot be directly adapted for FAD synthesis due to the insolubility of FMN in DMSO/DMF. Each of these chemical methods for synthesizing FAD nucleobase analogues is restricted by various factors, including the number of steps, the product yield, synthesis duration, or the availability of starting substrates (as summarized in Table 3.1). *Corynebacterium ammoniagenes* FAD synthetase enzyme (*Ca*FADS) has undergone comprehensive optimization for the production of the isoalloxazine ring and the ribityl chain-based FAD analogues; nevertheless, this approach is not applicable for synthesizing nucleobase analogues of FAD due to its high selectivity for ATP.[1,30,31]

This chapter presents an effective way of synthesizing the nucleobase analogues of FAD using two distinct approaches: the chemical synthesis approach and the enzymatic approach. Both approaches involve a maximum of three steps, resulting in moderate product yields ranging from 10% to 57%. We propose that altering the nucleobase is not going to impact the redox capability of the isoalloxazine ring of analogues, as the isoalloxazine and adenine rings typically adopt an extended conformation, positioned around 11-16 Å apart (measured between terminal backbone atoms) when bound to flavoenzymes (depicted in Figure 3.1C).[32] Consequently, we conducted experiments to evaluate the redox capability of these analogues with *Escherichia coli* glutathione reductase (*Ec*GR) in the presence. Our findings demonstrated that all these analogues can bind and serve as reducing enzyme cofactors with distinct efficiencies. Subsequently, by capitalizing on the fact that substrates NTP and riboflavin are naturally present in cells, we utilized the versatile FMN adenylyltransferase from *Methanocaldococcus jannaschii* (*Mj*FMNAT) to demonstrate the synthesis of FAD nucleoside analogues within a living cell by heterologous expression (structures shown Figure 3.1B). Our analysis implies that nucleobase analogues of FAD can be utilized potentially to investigate the intricate function of FAD and the corresponding flavoenzymes in the cell. Similar to the applications formulated with nucleobase analogue (NCD) of the redox cofactor NAD, The FAD nucleobase analogues can have potential applications in the field of biotechnology and synthetic biology.[29,33,34]

3.2 Materials and Methods

3.2.1 Materials

Jena Biosciences supplied the order of 2'-deoxy adenosine triphosphate (dATP) and all nucleoside triphosphates (NTPs). Reagents and chemicals for synthesis and enzymatic assays were purchased from RANKEM Chemicals, TCI Chemicals, and Merck. All reagents and components of media used in bacterial cultures were ordered from TCI Chemicals, HiMedia, and SRL Chemicals. DSS Takara Bio USA supplied Taq polymerase, Dpn1, and the PrimeSTAR GXL DNA polymerase. Purification kits for DNA and plasmid were obtained from Agilent and Qiagen. We considered Merck, SD Fine Chemicals, and RANKEM for all solvents ordering to use in LC-MS and HPLC analysis. *Methanocaldococcus jannaschii* DSM 2661 genomic DNA was kindly provided by Biswarup Mukhopadhyay, Virginia State University. The National BioResource Project of ASKA collection was the source for the *Ec*GR-pCA24N plasmid transformed *E. coli* K-12 AG1 strain.

3.2.2 Synthesis of FMNAc (2', 3', 4' -triacetyl-riboflavin 5' -monophosphate)

In a round-bottomed flask (RB), vacuum drying of FMN (0.63 mmol, 300 mg) was performed for 30 minutes to remove residual moisture. Subsequently, 100-150 µL of 55% perchloric acid (3-5 drops) of acetic anhydride (21 mmol, 2 mL) was introduced. At room temperature, the solution was kept stirring for an hour, and the FMN dissolved completely. The reaction mixture was then made to precipitate by the addition of diethyl ether. After filtration, 5 mL of water was used to redissolve the precipitate. Initially, 10 mL of chloroform was employed three times to eliminate the acetylated riboflavin byproduct, followed by the purification of the product through C18-reversed-phase chromatography. The purification process employed solvent A as water and solvent B as MeOH for a 60-minute method at a 10 mL/min solvent flow. The gradient method used: 10 min with 5% B; 30 min linear gradient to 50% B; 50 min linear gradient to 100% B; 60 min at 100% B. 280 and 365 nm wavelengths were employed to record the chromatograms. The high vacuum drying of the final product was done. Characterization was achieved through NMR and LC-MS analysis, with a yield of 33%. The NMR spectra and literature reports were the same (NMR Figure 3.9-3.10).[35]

FMNAc. **^{1}H NMR** (400 MHz, D_2O) δ 7.59 (s, 1H), 7.57 (s, 1H), 5.52 (m, 2H), 5.40 (m, 1H), 3.98 - 4.23 (m, 2H), 2.51 (s, 3H), 2.34 (s, 3H), 2.26 (s, 3H), 2.23 (s, 3H), 1.69 (s, 3H).

^{31}P NMR (162 MHz, D_2O) δ 0.59 (s). LC-MS (ESI-TOF) m/z: calculated for $C_{23}H_{27}N_4O_{12}P$ $[M-H]^-$ 581.1285, found 581.1286.

3.2.3 FNDAc Synthesis

The procedure began by drying 0.15 mmol NMP with a high vacuum for 30 minutes in a flask. The dry condition using nitrogen was used for all reactions; all reagents and solvents were anhydrous. Next, the NMP was mixed with acetonitrile (400 µL), triethylamine (1.25 mmol, 175 µL) and N,N-dimethylaniline (0.5 mmol, 60 µL), followed by stirring the solution for 10 minutes in an ice-water bath. To this NMP mixture, a cooled trifluoroacetic anhydride (TFAA, 1.25 mmol, 175 µL) and acetonitrile (160 µL) mixture was subsequently introduced and kept stirring for 15 minutes at RT. After the disappearance of the NMP, the unreacted TFAA was cautiously removed through a high vacuum, followed by cooling and redissolving the resulted pale-yellow syrup in acetonitrile (100 µL). A cooled solution of N-methylimidazole (0.4 mmol, 30 µL) in acetonitrile (40 µL) was introduced to the trifluoroacetylated NMP solution, followed by stirring for 15 minutes in an ice-water bath. This resulting N-methylimidazole-activated NMP solution was introduced to another solution of FMNAc (0.2 mmol, 115 mg) in acetonitrile (1 mL) and tributylamine (0.2 mmol, 48 µL) in the presence of 3Å molecular sieves. The solution was stirred, and thin-layer chromatography (TLC) was employed to monitor the reaction progress. After 2 hours, 250 mM ammonium acetate (10 mL) was used to quench the reaction. First, dichloromethane (10 mL) was used two times to extract the aqueous phase, followed by purification through C18-reversed-phase chromatography. This purification process employed solvent A as 0.1% triethylammonium acetate and solvent B as MeOH with a 10 mL/min flow rate. The gradient method used: 10 min with 5% B; 50 min linear gradient to 30% B; 60 min linear gradient to 100% B; 70 min at 100% B. 280 and 365 nm wavelengths were employed to record the chromatograms. The high vacuum drying of the final product was done. Characterization was achieved through NMR (NMR Figure 3.14-3.16) and LC-MS analysis, with a yield of 31%.

FCDAc. **^{1}H NMR** (400 MHz, D_2O) δ 7.96 (d, J = 7.7 Hz, 1H), 7.78 (s, 1H), 7.70 (s, 1H), 6.06 (d, J = 7.7, 1H), 5.84 (d, J = 3.7 Hz, 1H), 5.55 – 5.57 (m, 2H), 5.45 (br, 1H), 5.16 (br, 1H), 4.97 (br, 1H), 4.32 – 4.14 (m, 7H), 2.58 (s, 3H), 2.43 (s, 3H), 2.27 (s, 3H), 2.25 (s, 3H), 1.72 (s, 3H). **^{13}C NMR** (100 MHz, D_2O) δ 172.85, 172.53, 165.08, 160.87, 157.25,

157.22, 156.29, 150.58, 149.94, 141.59, 139.26, 134.25, 134.09, 130.91, 130.77, 116.21, 96.14, 89.16, 82.56 (d, J_{CP} = 33.4 Hz), 74.30, 70.75 (d, J_{CP} = 21.8 Hz), 70.44, 69.84, 69.05, 64.42 (d, J_{CP} = 17.4 Hz), 63.78 (d, J_{CP} = 17.5), 45.05, 20.81, 20.63, 20.25, 19.70, 18.56. **^{31}P NMR** (162 MHz, D_2O) δ -10.94 (s). LC-MS (ESI-TOF) m/z: calculated for $C_{32}H_{39}N_7O_{19}P_2$ $[M-H]^-$ 886.1698, found 886.1699.

3.2.4 FNDAc deacetylation reaction

The entire procedure was conducted within a nitrogen atmosphere; all reagents and solvents were anhydrous. The methanol (1 mL) was added to FNDAc (~10 mg) in a round-bottom flask, followed by initiation of the reaction by adding 0.5 M sodium methoxide (100 µL). In about 5 minutes, precipitation of the product was observed due to the insolubility of FND in methanol. The HCl (0.5 M, 100 µL) was introduced to the reaction to quench further product degradation. The resultant product was subjected to purification through C18-reversed-phase chromatography. This purification process employed solvent A as 0.1% triethylammonium acetate and solvent B as methanol with a 10 mL/min flow rate. The gradient method used: 10 min with 5% B; 50 min linear gradient to 20% B; 70 min linear gradient to 100% B; 80 min at 100% B. 280 and 365 nm wavelengths were employed to record the chromatograms. The high vacuum drying of the final product was done. Characterization was achieved through NMR (NMR Figure 3.17-3.19) and LC-MS analysis, with a yield of 17% concerning CMP as the starting material.

FCD. ^{1}H NMR (400 MHz, D_2O) δ 7.98 (d, J = 7.7 Hz, 1H), 7.79 (s, 1H), 7.63 (s, 1H), 6.07 (d, J = 7.7 Hz, 1H), 5.79 (d, J = 4.0 Hz, 1H), 5.05 (dd, J = 12.8, 10.4 Hz, 1H), 4.68 (dd, J = 12.8, 1.2 Hz, 1H), 4.44 – 4.37 (m, 1H), 4.31 – 4.08 (m, 8H), 4.00 (dd, J = 7.2, 4.9 Hz, 1H), 2.50 (s, 3H), 2.38 (s, 3H). **^{13}C NMR** (150 MHz, D_2O) δ 160.70, 159.19, 156.09, 151.20, 149.10, 148.24, 140.86, 137.82, 132.65, 132.20, 129.99, 128.65, 115.38, 94.05, 87.67, 81.16 (d, J_{CP} = 37.1 Hz), 72.76, 70.86, 69.54 (d, J_{CP} = 31.0 Hz), 67.52, 67.32, 65.56 (d, J_{CP} = 22.0 Hz), 62.70 (d, J_{CP} = 20.7 Hz), 45.91, 19.15, 17.02.**^{31}P NMR** (162 MHz, D_2O) δ -10.62 (d, J_{PP} = 49.5 Hz,), -11.39 (d, J_{PP} = 49.5 Hz). LC-MS (ESI-TOF) m/z: calculated for $C_{26}H_{33}N_7O_{16}P_2$ $[M-H]^-$ 760.1381, found 760.1376.

3.2.5 Alternate approach to synthesize FCD

A nitrogen atmosphere was used to execute all the reactions with anhydrous reagents and solvents. The FMN (0.63 mmol, 300 mg) was acetylated with acetic anhydride (21 mmol, 2 mL) in the presence of ~100 µL perchloric acid (55%, 3-5 drops), followed by stirring this solution for 1 hour at RT. Diethyl ether was employed to precipitate the product, which facilitated the removal of acetic acid and anhydride, followed by filtering and redissolving the precipitate in 5 mL of water. Three rounds of 10 mL chloroform were used to extract the aqueous phase for the elimination of acetylated riboflavin. This yielded a mixture of acetylated FMN, which was directly employed in the subsequent coupling reaction.

CMP (free acid, 50 mg, 0.15 mmol) underwent vacuum drying for 30 minutes within an RB flask. The dry condition using nitrogen was used for all reactions; all reagents and solvents were anhydrous. Next, the CMP was mixed with acetonitrile (400 µL), triethylamine (1.25 mmol, 175 µL) and N,N-dimethylaniline (0.5 mmol, 60 µL), followed by stirring the solution for 10 minutes in an ice-water bath. To this CMP mixture, a cooled trifluoroacetic anhydride (TFAA, 1.25 mmol, 175 µL) and acetonitrile (160 µL) mixture was subsequently introduced and kept stirring for 15 minutes at RT. After the disappearance of the CMP, the unreacted TFAA was cautiously removed through a high vacuum, followed by cooling and redissolving the resulted pale-yellow syrup in acetonitrile (100 µL). A cooled solution of N-methylimidazole (0.4 mmol, 30 µL) in acetonitrile (40 µL) was introduced to the trifluoroacetylated CMP solution, followed by stirring for 15 minutes in an ice-water bath. This resulting N-methylimidazole-activated CMP solution was introduced to another solution of acetylated FMN mixture in acetonitrile (1 mL) and tributylamine (0.2 mmol, 48 µL), where acetylated FMN mixture and CMP were taken in the 3:1 ratio in the presence of 3Å molecular sieves. After 2 hours, 250 mM ammonium acetate (10 mL) was used to quench the reaction. First, dichloromethane (10 mL) was used two times to extract the aqueous phase, followed by drying the reaction mixture under a high vacuum. The dried material was deacetylated using 50 mM sodium methoxide in methanol for 3-5 minutes. This led to the precipitation of FCD and FMN in methanol. However, under these conditions, some of the FCD converted to cyclic FMN and CMP, a phenomenon similar to that observed for FAD.[36] The FCD was then purified through C18-reversed-phase chromatography by following the same method mentioned earlier. This method resulted in a yield of 13%.

3.2.6 FUD Synthesis

UMP (free acid, 50 mg, 0.15 mmol) was vacuum-dried for 30 minutes within a round-bottom (RB) flask. The dry condition using nitrogen was used for all reactions; all reagents and solvents were anhydrous. Next, the UMP was mixed with acetonitrile (400 µL), triethylamine (2.5 mmol, 350 µL) and N,N-dimethylaniline (0.5 mmol, 60 µL), followed by stirring the solution for 10 minutes in an ice-water bath. To this UMP mixture, a cooled TFAA (1.25 mmol, 175 µL) and acetonitrile (160 µL) mixture was subsequently introduced and kept stirring for 15 minutes at RT. After the disappearance of the UMP, the unreacted TFAA was cautiously removed through a high vacuum, followed by cooling and redissolving the resulted pale-yellow syrup in acetonitrile (100 µL). A cooled solution of N-methylimidazole (0.4 mmol, 30 µL) in acetonitrile (40 µL) was introduced to the trifluoroacetylated NMP solution, followed by stirring for 15 minutes in an ice-water bath. This resulting N-methylimidazole-activated UMP solution was introduced to another solution of FMNAc (0.2 mmol, 115 mg) in acetonitrile (1 mL) and tributylamine (2 mmol, 500 µL) with 3Å molecular sieves. The solution was stirred, and TLC was used to monitor the reaction progress. After 2 hours, 250 mM ammonium acetate (10 mL) was used to quench the reaction. First, dichloromethane (10 mL) was used two times to extract the aqueous phase, followed by purification through C18-reversed-phase chromatography. This purification process employed solvent A as 0.1% triethylammonium acetate and solvent B as MeOH with a 10 mL/min flow rate. The gradient method used: 10 min with 5% B; 50 min linear gradient to 30% B; 60 min linear gradient to 100% B; 70 min at 100% B. 280 and 365 nm wavelengths were employed to record the chromatograms. The high vacuum drying of the final product was done. Characterization was achieved through NMR (NMR Figure 3.20-3.22) and LC-MS analysis, with a yield of 11%.

^{1}H NMR (600 MHz, D_2O) δ 7.84 (s, 1H), 7.82 (s, 1H), 7.72 (s, 1H), 5.85 – 5.78 (m, 2H), 5.11 – 5.02 (m, 1H), 4.44 – 4.37 (m, 1H), 4.32 – 4.08 (m, 8H), 4.04 – 3.96 (m, 1H), 2.52 (s, 3H), 2.40 (s, 3H). **^{13}C NMR** (150 MHz, D_2O) δ 164.02, 159.57, 156.10, 149.61, 148.84, 148.28, 139.50, 137.56, 132.85, 132.41, 130.03, 128.62, 115.11, 100.57, 86.48, 81.27, 72.02, 70.75, 69.27, 67.69, 65.39, 65.38, 63.02, 45.81, 18.97, 16.81. **^{31}P NMR** (243 MHz, D_2O) δ -8.07 (br), -8.75 (br). LC-MS (negative mode ESI-TOF) m/z: calculated for $C_{26}H_{32}N_6O_{17}P_2$ $[M-H]^-$ 761.1226, found 761.1228.

3.2.7 Restriction-free cloning of *Mj*FMNAT

The FMN adenylyl transferase gene (MJ1179) from the genomic DNA of the archaea *Methanocaldococcus jannaschii* was introduced into the pET28a vector by restriction-free cloning.[37] This insertion was done between the vector's NdeI and BamHI restriction sites. The MJ1179 gene was amplified through polymerase chain reaction (PCR) using the following primers: Forward 5'-ATGAAAAAGAGGGTAGTTACCGC-3' and Reverse 5'-TTAGATTTTAATCTCTTTATTGCAG-3'. Subsequently, a second round of PCR was carried out to incorporate complementary sequences for the NdeI and BamHI sites at the ends of the PCR-amplified gene, which was done with the forward primer 5'-GTGCCGCGCGGCAGCCATATGAAAAAGAGGGTAGTTACCGC-3' and reverse primer 5'-CGACGGAGCTCGAATTCGGATCCTTAGATTTTAATCTCTTTATTGCAG-3'.

This gene, possessing the terminal flanking regions complementary to the plasmid, is called a megaprimer. The megaprimer was then employed to perform a third round of PCR, resulting in the amplification of the pET28a vector with the gene incorporated. Subsequently, the original plasmid was treated with Dpn1 (NEB), and the vector containing MJ1179 was introduced into *E. coli* DH5α chemically competent cells, allowing the plasmid's isolation.

3.2.8 Purification of *Mj*FMNAT

The isolated plasmid containing the MJ1179 gene encoding the *Mj*FMNAT enzyme was introduced into chemically competent *E. coli* BL21(DE3) cells. The culture was grown in Luria-Bertani (LB) broth at 37 °C and 180 rpm. Once OD_{600} reached the range from 0.6 to 0.8, IPTG (1 mM) was added to induce the culture at 30 °C. The culture was then allowed to shake for 8 hours. After this period, the cells were harvested and subjected to lysis using a lysis buffer consisting of Tris-HCl (100 mM, pH 8.0), PMSF (0.1 mM), beta-mercaptoethanol (0.025%), and NaCl (300 mM). Following lysis, the lysate underwent centrifugation (18000 rpm x 30 minutes). The resulting supernatant was loaded onto a Ni-NTA column pre-equilibrated with a buffer consisting of Tris-HCl (100 mM, pH 8.0), beta-mercaptoethanol (0.025%), and NaCl (300 mM). The elution of protein was done using a buffer including Tris-HCl (100 mM, pH 8.0), imidazole (250 mM), beta-mercaptoethanol (0.025%), and NaCl (300 mM). This protein concentration measurement was employed at A_{280}, considering the predicted molar extinction coefficient of the protein to be 5960 M^{-1} cm^{-1}.

3.2.9 Optimization of *Mj*FMNAT activity and synthesis of FAD analogues

The enzymatic activity of *Mj*FMNAT was done using various nucleotide substrates, including all NTPs and dATP. The reaction conditions involved mixing Tris-HCl (100 mM, pH 8.0), FMN (200 µM), MgCl2 (7 mM), NTP (5 mM), and DTT (14 mM) with 25 µM enzyme and making the volume to 200 µL. The mixture was then performed at 70 °C, and the progress of the reaction was analyzed by quenching this reaction at different time points (0.5, 1, 3, and 12 hours) with acetic acid (60 mM). The resulting FAD and its analogues were analyzed using fluorescence TLC, HPLC, and LC-ESI-MS. The m/z values of the synthesized analogues were consistent with what was reported previously for the organic synthesis: FGD ($C_{27}H_{33}N_9O_{16}P_2$) m/z $[M-H]^-$ calculated 800.1442, found 800.1425; dFAD ($C_{27}H_{33}N_9O_{14}P_2$) m/z $[M-H]^-$ calculated 768.1544, found 768.1545; dFCD (C26H33N7O15P2) m/z $[M-H]^-$ calculated 744.1437, found 744.1438.

To conduct the FAD analogue synthesis using *Mj*FMNAT, a 1 mL reaction was set up with FMN (400 µM) and each NTP (5 mM), along with Tris-HCl (100 mM, pH 8.0), DTT (14 mM), $MgCl_2$ (7 mM) and 25 µM enzyme. The resulting FAD analogues were purified using the same HPLC method and collected. After purification, the analogues were concentrated to dryness using the Centrivap DNA concentrator, followed by storing at -20 °C. Each analogue was dissolved in 500 µL of water, followed by concentration determination using HPLC with a standard curve of FAD at 440 nm, assuming that all analogues have the same molar coefficient as FAD. Based on the substrate's initial concentration used in the reaction, the percentage yield of the analogue was calculated.

3.2.10 Purification of glutathione reductase of *E. coli* (*Ec*GR)

The *E. coli* K12 AG1 cells transformed with pCA24n-*Ec*GR were cultivated in LB broth at a temperature of 37°C. When the OD_{600} reached the 0.6 to 0.8 range, IPTG (0.5 mM) was introduced to induce the protein overexpression at a lower temperature of 28°C. This induced culture was then kept shaking for 8 hours. Subsequently, the cells were collected and subjected to lysis using a buffer consisting of Tris-HCl (100 mM, pH 8.0), PMSF (0.1 mM), beta-mercaptoethanol (0.025%), and NaCl (300 mM). Following lysis, the lysate underwent centrifugation (18000 rpm x 30 minutes). The resulting supernatant was loaded onto a Ni-NTA resin pre-equilibrated with a buffer consisting of Tris-HCl (100 mM, pH 8.0), beta-mercaptoethanol (0.025%), and NaCl (300 mM). The elution of protein was done using a buffer including Tris-HCl (100 mM, pH 8.0), imidazole (250 mM), beta-mercaptoethanol (0.025%), and NaCl (300 mM).

To obtain the apoprotein, a saturated solution of ammonium sulfate was employed.[38] To 1 mL of protein solution, an ice-cold KBr (1 mL, 3 M) solution was introduced in a centrifuge tube. This mixture was then subjected to precipitation by adding saturated ammonium sulfate solution (1 mL, pH 2.2) and gentle swirling. After allowing the mixture to sit for 1 minute, additional saturated ammonium sulfate (4 mL) was added. The resulting solution was centrifuged at 18000 rpm for 30 minutes, forming a precipitate. This precipitate was re-dissolved in 0.1 M phosphate buffer (1 mL, pH 7), yielding the apoprotein, followed by further centrifugation to remove any remaining denatured protein.

For the reconstitution of the apoprotein, the enzyme was combined with an excessive amount (3 times the enzyme quantity) of FAD or FAD analogues. This mixture was incubated at 25 °C for 15 minutes. Subsequently, the surplus cofactor was eliminated by passing the protein through a desalting column equilibrated with 0.1 M phosphate buffer at pH 7, and the reconstituted protein was used to find its binding efficiency with the analogues. When performing an assay with the enzyme, the apoenzyme was reconstituted by incubating it with three times the amount of the cofactor (FAD or FAD analogues) at 25 °C. The enzyme could then be directly used in assays without removing any excess cofactor.

3.2.11 Glutathione reductase (*Ec*GR) activity

To assess the functionality of the FAD nucleoside analogues, the *Ec*GR enzyme was employed. In a concise description, this procedure encompasses the conversion of FAD to $FADH_2$ through the utilization of NADPH, causing NADPH to be oxidized to NADP. Subsequently, the $FADH_2$ participates in reducing the enzyme's disulfide bridge, which facilitates the reduction of GSSG to GSH (Figure 3.7A).[39] NADPH exhibits absorbance at 340 nm, while $NADP^+$ does not exhibit absorbance at this particular wavelength. Therefore, we tracked the advancement of the *Ec*GR reaction by observing the decline in NADPH's absorbance at 340 nm in a 96-well plate. This decrease signifies the oxidation of NADPH into $NADP^+$ as facilitated by GSSG reduction.[40] The reaction mixture consisted of a 200 µL final volume containing potassium phosphate buffer (0.1 M, pH 7), along with 0.1 mM NADPH, 1 mM GSSG, and 1 mM EDTA. The reaction was initiated by adding 0.1 µM of the reconstituted enzyme(final concentration), and the entire process occurred at 30 °C. Over a duration of 20 minutes, the reduction in absorbance at 340 nm was continually observed. This observation allowed us to evaluate the activity of *Ec*GR in the presence of each analogue.[40] For each curve, a logistic fit function in Origin was employed for fitting purposes. The slope was then determined by the best-fitting line at the inflexion point 'x0'. This process yielded values

indicating the rate of reduction of GSSG by 1 µM enzyme, expressed in micromoles per minute (µmol/min), utilizing FAD and its analogues as cofactors. The activity values were measured in micromoles per minute per micromole (µmol min^{-1} $µM^{-1}$) and are as follows: FAD bound purified-*Ec*GR= 0.59, apoprotein= 0.12, apoprotein+FAD= 0.72, apoprotein+FCD= 0.62, apoprotein+FUD= 0.60, apoprotein+FGD= 0.52, apoprotein+dFAD= 0.60. To determine the relative activities compared to FAD, the activity of Apoprotein+FAD was taken as the reference value and set to 1.

3.2.12 Synthesizing FAD nucleobase analogues within the *E. coli* cells by *Mj*FMNAT

In LB broth, a 50 mL secondary culture was started by adding 1% of an overnight grown *E. coli* BL21 (DE3) strain harboring either pET28a, *Ec*FADS-pET28a, or *Mj*FMNAT-pET28a constructs with kanamycin resistance, followed by the growth at 37 °C and 180rpm. Only the BL21(DE3) cell culture was grown without Kanamycin. Instead of inducing with IPTG, we relied on the naturally occurring low-level expression from the pET28a plasmid to avoid significant disruption of cellular metabolism. Following an incubation period of 18 hours, a 50 mL portion of the culture was subjected to centrifugation for 15 minutes at 4 °C and 5500 rpm. The resulting pellet was suspended in 90% methanol (2 mL) and subjected to sonication for 3 minutes using alternating 1-second pulses on and off at 100% amplitude. The final supernatant was obtained by centrifuging the lysate for 15 minutes at 4°C and 5500 rpm), followed by filtration through 0.22-micron filters. For LC-MS analysis, the supernatant was concentrated under high-vacuum conditions.

3.2.13 TLC, HPLC, and LC-MS

Silica gel TLC was employed to analyze all reactions, and the outcomes were detected under ultraviolet light with a wavelength of 365 nm. The mixture of water, acetic acid, and n-butanol in a ratio of 5:3:12 was employed as the mobile phase.

HPLC (Agilent 1260 Infinity-II series) linked with a UV-Vis diode array detector was utilized to assess enzymatic reactions. The analysis used a Phenomenex Gemini reversed-phase column of 250 x 4.6 mm dimensions with a particle size of 5 µm NX-C18. The solvent-A was ammonium acetate buffer (10 mM, pH 6.0), and methanol was used as solvent-B). A method spanning 60 minutes was applied, starting with 5% of solvent B at 0 minutes, maintaining this

ratio until 15 minutes, followed by a gradient increase to 80% of solvent B until 40 minutes, then a return to 5% solvent B until 55 minutes, and finally, maintaining at 5% solvent B until 60 minutes. Another method spanning 30 minutes was used, starting with 5% of solvent B at 0 minutes, maintaining this ratio until 2.5 minutes, followed by a gradient increase to 80% of solvent B until 20 minutes, then a return to 5% solvent B until 27 minutes, and finally, maintaining at 5% solvent B until 30 minutes.

A Mass spectrometer (ScieX X500R-QTOF) was employed for LC-MS analyses, where a UHPLC (Exion-LC series) was used to separate the flavins. The flavins were detected in negative mode at a temperature of 400 °C. The parameters are as follows: ion spray voltage was -4500 V, the de-clustering potential was -80V and collision energy was 5V. The ScieX-OS Explorer software facilitated data acquisition, processing, and the generation of metrics such as extracted ion chromatogram (EIC), total ion chromatogram (TIC), and the m/z values of targeted molecules. The same column and method as used in HPLC were utilized.

3.2.14 RMSD of FAD molecules in FAD-binding proteins

To quantitatively compare the different conformations of FAD within FAD-binding enzymes, we employed the Root Mean Square Deviation (RMSD) approach, where RMSD is calculated as: $RMSD= [(\sum(x_e-x_0)^2)/n]^{0.5}$ (where x_e= expected value, x_o= observed value and n= total number of values).[41] RMSD of atomic positions is a quantitative measure of the average distance between the corresponding atoms of two superimposed molecules (FAD in this context). For this analysis, a total of 24 FAD-binding protein structures available in the PDB (Protein Data Bank) were utilized, where the PDB can be accessed at https://www.rcsb.org.[42] The "match command" in UCSF Chimera was employed to execute optimal alignment (superposition) of designated atoms. This process involved aligning the atoms of one FAD molecule onto those of another FAD molecule, followed by the computation of RMSD values to quantify the structural disparities between the aligned molecules.[43] In these calculations, a total of 53 atoms within an FAD molecule (excluding hydrogen atoms) were considered. The process involved pairwise RMSD calculations for 24 FAD molecules of this kind. Among them, the FAD molecule associated with thioredoxin disulfide reductase in *Campylobacter jejuni* (PDB ID: 3R9U) was chosen arbitrarily as the reference molecule, assigned a reference RMSD value of 0.0 Å. Subsequently, RMSD values for each of the remaining 23 enzyme-

bound FAD molecules related to the reference molecule were calculated and depicted in a plot (Figure 3.1C).

3.3 Results and Discussion

3.3.1 An effective synthesis approach for FAD and its analogues

In order to establish a reliable method for synthesizing nucleobase analogues of FAD, we focused on investigating the direct condensation pathway of nucleoside monophosphates (NMP) with FMN as a means to generate flavin nucleoside dinucleotides (FND) by utilizing alternate activating agents. Except for yielding low quantities, the direct condensation routes offered several advantages, including a reduced number of reaction steps, shorter reaction duration, and easy accessibility to initial materials. To enhance the condensation efficiency of the reaction, we opted for the activating agent N-methylimidazole, which was influenced by a previously reported deoxynucleoside 5′-triphosphate (dNTP) synthesis.[44] In this described reaction, the hydroxyl groups on the phosphate of deoxynucleoside monophosphates (dNMP) were subjected to acylation using TFAA (trifluoroacetic anhydride). This was succeeded by the formation of the corresponding dNMP-N-methylimidazolide using N-methylimidazole. Ultimately, coupling with inorganic pyrophosphate resulted in the generation of deoxynucleoside triphosphates (dNTP) with a notable yield range of 89-92%. This approach has subsequently been employed to synthesize compounds such as various sugar nucleotides, UDP-α-D-galactofuranose, and metabolites containing di/triphosphate groups.[45–48] In this investigation, we successfully applied the aforementioned approach to synthesize nucleobase analogues of FAD. Our adaptation allowed for a three-step reaction, resulting in an approximate yield of around 20%. Additionally, a two-step reaction was also developed, yielding roughly 10%. Importantly, both of these reactions were accomplished within a remarkably short and less than 5-hour timeframe.

3.3.2 Synthesis of FMNAc (acetylated FMN)

The initial compounds used for synthesizing nucleobase analogues of FAD through this approach are FMN **2** and NMP **9**. Due to the insolubility of FMN **2** in organic solvents, we introduced acetylation to FMN **2** using acetic acid anhydride and perchloric acid. This

procedure led to the creation of tri-acetylated FMN, designated as FMNAc **8** (as depicted in Figure 3.2).[20] The reaction resulted in a mixture of acetylated FMN products where FMN **2** was acetylated at one, two, or all three hydroxy positions. Additionally, acetylated riboflavin was also formed due to the presence of riboflavin impurity in the starting material, FMN **2** (as shown in Figure 3.3D). Following the removal of acetylated riboflavin using chloroform, the FMNAc **8** compound was purified using C18-reversed-phase chromatography, yielding a 33% overall product. Its identity was confirmed through various characterization techniques, including TLC, HPLC, NMR, and LC-MS (Figures 3.3 and 3.4).

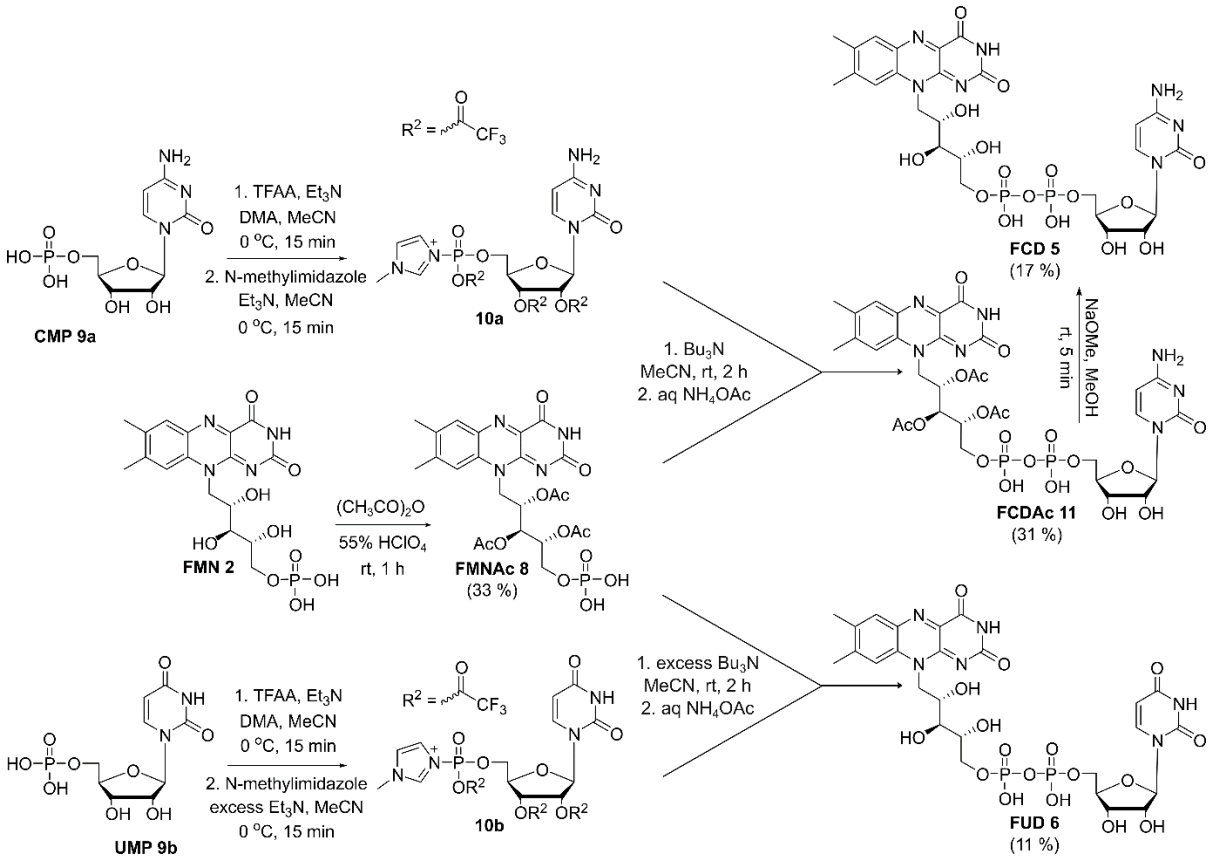


Figure 3.2 Synthesis scheme for FAD analogues (FCD and FUD). The process initiates with the activation of the nucleotide monophosphate (NMP), CMP/UMP. This activation is achieved through trifluoroacetic anhydride (TFAA) treatment, followed by N-methylimidazole. Subsequently, the acetyl-protected FMN (FMNAc) is coupled with the activated NMP, creating the corresponding acetylated flavin nucleoside diphosphate (FNDAc). Following this, FNDAc undergoes deacetylation using sodium methoxide to form FND. As an alternative approach, an excess of tributylamine can be introduced to bypass the FNDAc formation step. This adjustment streamlines the process by reducing the deacetylation stage, albeit with a trade-off of yielding lower quantities, as depicted for FUD synthesis. Figure adapted from *ChemBioChem* (**2023**), Shah *et al*.

Efforts were made to enhance the yield by performing the reaction at an elevated temperature of 80 °C for a duration of 24 hours, which was adapted from published literature.[35]

Unfortunately, despite the prolonged reaction time and increased temperature, the outcome remained consistent with forming a mixture of acetylated FMN compounds. Regrettably, this approach did not lead to an increase in the quantity of FMNAc **8**. Another approach to increase the yield we explored was neutralizing the acid derivative of acetylated FMN using sodium bicarbonate. This was aimed at reducing losses during the subsequent workup with chloroform. However, this attempt presented unexpected challenges when we generated the acetylated FMN salt; sodium salts of both perchloric acid and acetic acid were also formed in the process. These additional products complicated the purification process, posing difficulties in obtaining a pure sample of the desired acetylated FMN compound, FMNAc **8**.

3.3.3 Synthesis of FND by Coupling of FMNAc with NMP

To effectively couple FMNAc with CMP **9a**, the phosphate group activation in **9a** was initiated using TFAA, followed by a secondary activation as the N-methylimidazolyl compound **10a** using N-methylimidazole (as illustrated in Figure 3.2). Subsequently, in the presence of tributylamine, **10a** and FMNAc **8** coupled, followed by purification by C18-reversed phase column chromatography, yielding approximately 31% pure FCDAc **11** (as shown in Figures 3.3 and 3.4). The conversion of FCDAc **11** to FCD **5** was pursued through acid or base catalyzed deacetylation reaction. However, due to the instability of FCD **5** in acidic or basic conditions, an alternative approach for the deacetylation of FCDAc **11** was performed using a solution of sodium methoxide (50 μM) in methanol for short durations (3-5 minutes), followed by quenching with an equimolar HCl solution. Given the insolubility of FCD **5** in methanol, the precipitation of FCD, along with other FMN derivatives, was observed. The subsequent purification of FCD **5** was achieved through reverse-phase chromatography. The identity of FCD **5** was verified using fluorescence TLC and HPLC (as depicted in Figure 3.4A, B). Additionally, the confirmation of FCDAc **11** and FCD **5** was done through NMR and LC-MS, resulting in a 17% final yield (illustrated in Figures 3.3-3.4 and NMR Figures 3.17-3.19). This method represents an advancement over previous techniques as reaction and purification steps are reduced, and the starting materials are commercially accessible while the yield is still modest. Furthermore, this approach can be extended to synthesizing various nucleobase analogues of FAD by altering the mononucleotides.

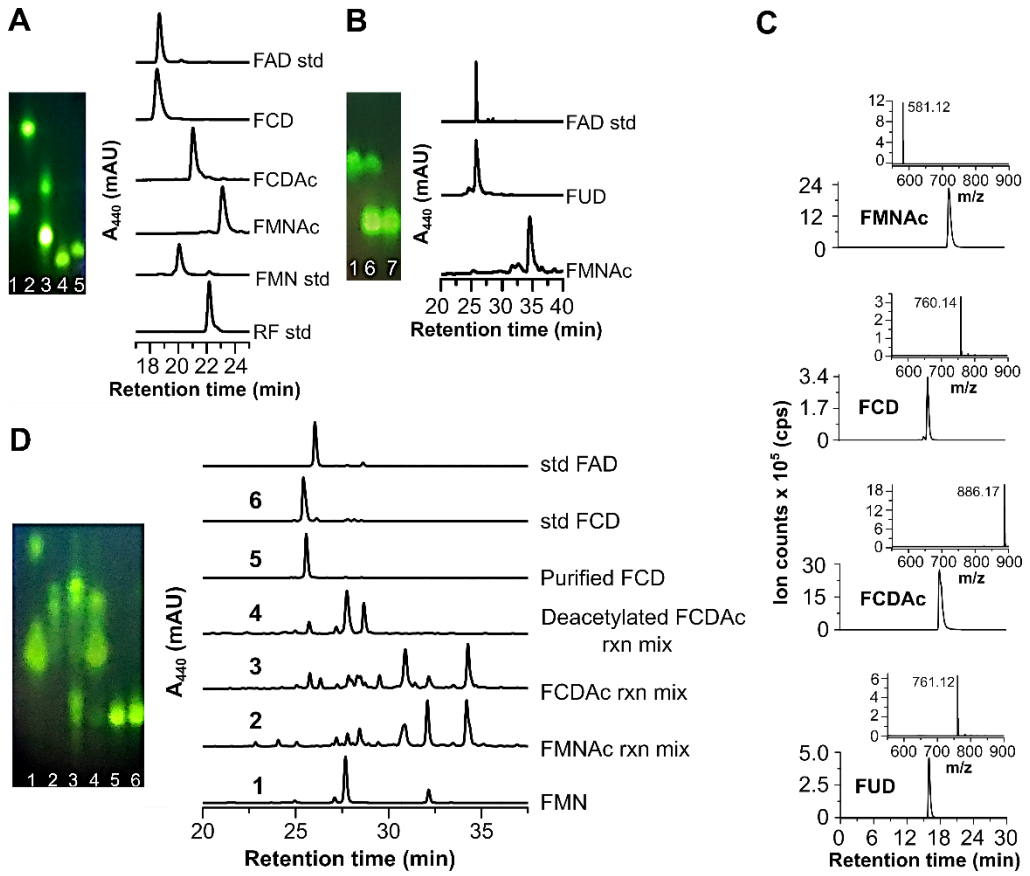


Figure 3.3 Characterization of FCD and FUD and their synthesis intermediates. (A) For FCD synthesis: Fluorescence thin-layer chromatography (TLC) Lane 1-FMN, Lane 2-RF, Lane 3-FMNAc (upper spot) and FCDAc (lower spot), Lane 4-FCD, Lane 5- standard FAD. The adjacent C18 - reversed-phase high-performance liquid chromatogram (HPLC) shows the reagent FMN, intermediate, and FCD. **(B)** For FUD synthesis: Fluorescence TLC Lane 1-FMN, Lane 6-FMN and FUD, Lane 7-FUD). The adjacent C18 reversed-phase HPLC trace shows the reagent and purified product of FUD synthesis. **(C)** Negative mode LC-MS extracted ion chromatograms (EIC) of purified intermediates and products FCD and FUD and their m/z spectra in the inset. **(D)** A direct approach to synthesizing FCD from FMN via multiple versions of acetylated FMN and FCD on various hydroxy groups. Each intermediate in the FCD synthesis process is depicted on a fluorescence TLC as follows: Lane 1 shows the bands for flavin mononucleotide (FMN, lower band) and riboflavin (RF, upper band) as an impurity of FMN, lane 2 illustrates the reaction mixture for acetylated flavin mononucleotides, Lane 3 presents the reaction mixture for multiple acetylated flavin cytidine dinucleotide versions, lane 4 displays the reaction mixture for the deacetylation of FCDAc, lane 5 depicts purified FCD, and lane 6 represents the standard FCD for reference. A stepwise HPLC chromatogram shows a progressive view of the reaction intermediates: The FMNAc reaction mixture showcases numerous peaks corresponding to a mixture of acetylated FMN. Similarly, FCDAc also exhibits peaks reflecting acetylation at distinct stages. On subsequent deacetylation, these stages lead to the final FCD product, along with cyclic FMN (a degradation product) and FMN. This analytical representation provides an insightful visualization of the synthesis progression and the distinct stages of intermediates in the synthesis of FCD. Figure adapted from *ChemBioChem* (**2023**), Shah *et al*.

Curiously, when applying this method to synthesize FAD 3 via the intermediate FADAc (verified by NMR Figures 3.11-3.13), a reduction in FADAc yield was observed compared to FCDAc (20% versus 31%). FAD was also obtained within the 3-5% range during deacetylation. In the case of FGDAc synthesis, the product formed was notably minimal and challenging to

quantify. Consequently, this synthesis of FGD was not pursued. The underlying cause for the poor yields of the acetylated purinyl analogues within this approach remains unclear.

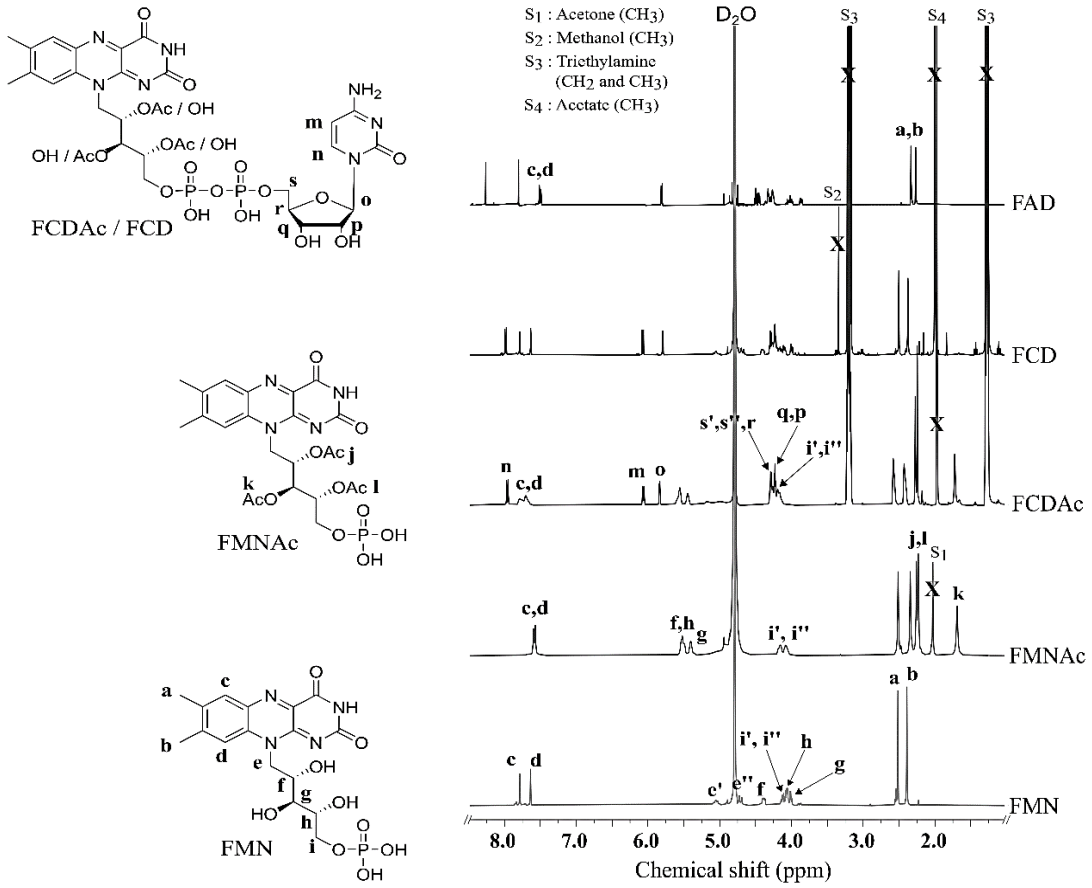


Figure 3.4 Comparison by ^{1}H-NMR of starting material FMN, intermediate FCDAc, and product FCD spectra. The FMNAc spectrum exhibits acetyl peaks (designated as j, k, and l), and compared to FMN, the corresponding proton peaks CH in the ribityl chain (marked as f, g, and h) manifest a downfield shift. The FCDAc spectrum portrays extra peaks originating from the adenosine component, with two aromatic peaks (represented as m and n). Notably, in the FCD spectrum, the trio of acetyl peaks j, k, and l is absent, and the peaks of CH proton f, g, and h revert to the upfield, akin to FMN. The peaks labelled S1 to S4 correspond to solvent residuals: acetone (CH3), methanol (CH3), triethylamine (CH2 and CH3), and acetate (CH3). Additionally, the 'and" symbol applied to protons in identical positions indicates different shielding due to adjacent chiral carbons. Figure adapted from *ChemBioChem* (**2023**), Shah *et al*.

In an effort to reduce the number of steps, an alternative approach was explored. Instead of isolating and subsequently deacetylating FNDAc, an attempt was made to perform the direct deacetylation during the coupling step by employing an excess of tributylamine and triethylamine (as illustrated in Figure 3.2). To synthesize FUD **6**, a similar method as that of FCD **5** up was employed, where N-methylimidazole was used to activate the UMP **9b**. However, triethylamine and tributylamine were used in excess during the coupling reaction. This simultaneous action resulted in the deacetylation of FMNAc **8** and FUDAc, converting

them into FMN **2** and FUD **6**. Then, the purification of FUD **6** was similar to FCD **5** and was validated through LC-MS and NMR analyses (Figures 3.3 and NMR Figures 3.20-3.22). This approach needs one step less than the earlier method, which requires three. The reduction in steps represents an enhancement in terms of time and resource efficiency for synthesis and purification. Nonetheless, the yield experienced a decline to 11%. Due to this yield reduction, this method was not pursued with other nucleotides.

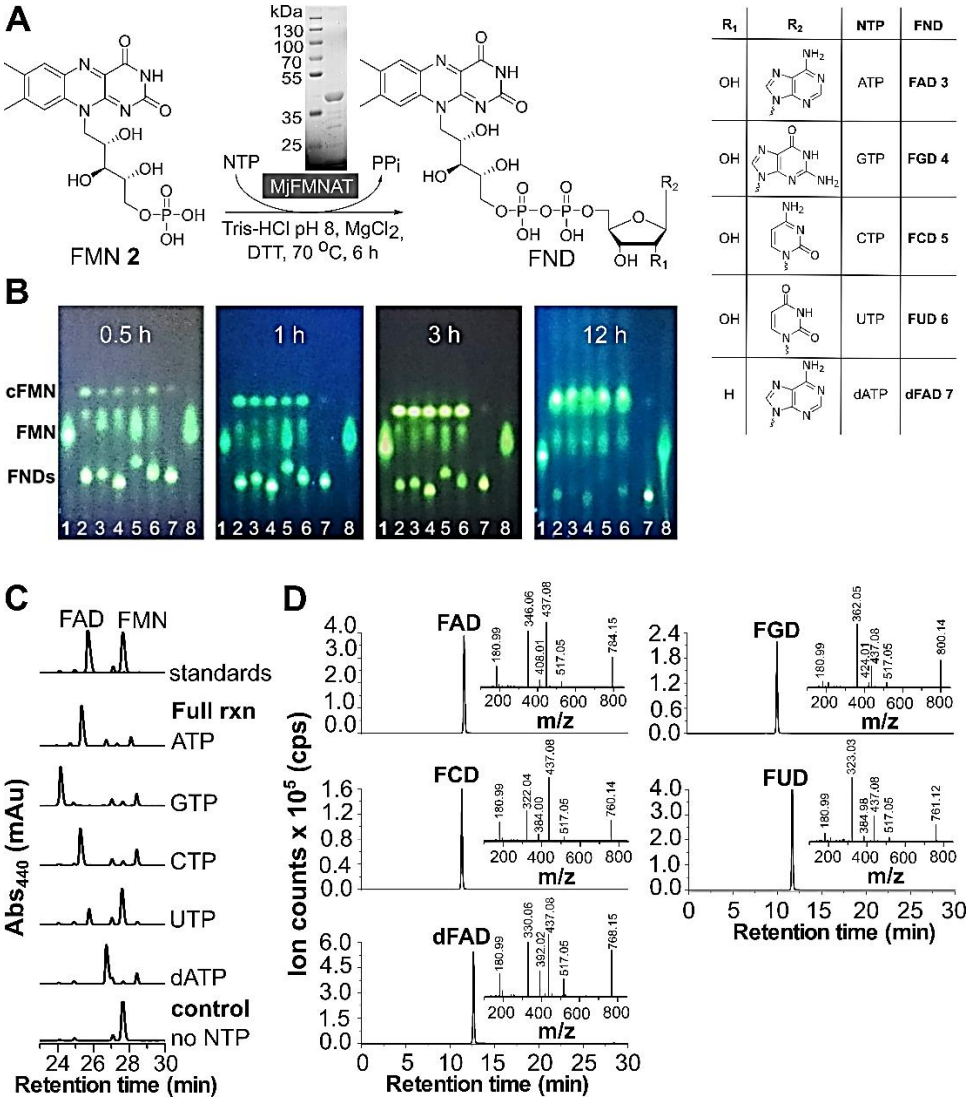


Figure 3.5 The enzymatic synthesis of FAD nucleoside analogues using *Mj*FMNAT. (A) *Mj*FMNAT is characterized as a stable and versatile enzyme capable of interacting with various NTPs. It efficiently catalyzes the synthesis of nucleoside analogues of FAD by utilizing FMN and NTP/dNTP substrates. **(B)** Reaction time optimization for *Mj*FMNAT with all NTPs and dATP. Fluorescence thin-layer chromatography (TLC) was conducted over a time course using the *Mj*FMNAT enzyme to establish a standardized conversion of FMN into various flavin nucleobase dinucleotides (FND) and dFAD using different NTPs and dATP, respectively. The TLC results at different time intervals are for 0.5 hours, 1 hour, 3 hours, and 12 hours. Each lane of the TLC plate is attributed as follows: Lane 1: FMN standard, Lane 2: FMN + ATP reaction, Lane 3: FMN + GTP reaction, Lane 4: FMN + CTP reaction, Lane 5: FMN + UTP reaction, Lane 6: FMN + dATP reaction, Lane 7: FAD standard, and Lane 8: No NTP control. In lanes 2 to 6, the spot located below the standard FMN spot corresponds to FND, while the spot located above the FMN spot represents cyclic FMN, recognized as the degradation product of FND.

This analysis provides a visual depiction of the conversion progress of FMN into various flavin nucleobase dinucleotides and dFAD under different conditions and over varying time intervals. **(C)** The HPLC data shows the formation of various FNDs resulting from the FMN and the corresponding nucleotide, which include ATP, GTP, CTP, UTP, and dATP. We see a decrease in the FMN peak (28 min) and the appearance of the corresponding FND peak (between 24-27.5 min) in all NTP reactions. **(D)** The EICs of the FNDs with m/z are shown in the inset. The fragmentation of FAD and FCD matches with previous reports.[22] The m/z peaks 180.99, 437.08 (cFMN), and 517.05 represent the common peaks fragmenting from the FMN moiety of FND. The specific characteristic peaks are as follows: For FAD: Peaks at m/z 346.06 (AMP), 408.01 (ADP with H_2O removed), 784.15 (FAD). For FGD: Peaks at m/z 362.05 (GMP), 424.01 (GDP with H_2O removed), 800.14 (FGD). For FCD: Peaks at m/z 322.04 (CMP), 384.00 (CDP with H_2O removed), 760.13 (FCD). For FUD: Peaks at m/z 323.02 (UMP), 384.98 (UDP with H_2O removed), 761.12 (FUD). For dFAD: Peaks at 330.06 (dAMP), 392.02 (dADP with H_2O removed), 768.15 (dFAD).

Since the coupling occurs among the phosphates of activated NMP and FMNAc, all variants of acetylated FMN (acetylated at one, two, or all three hydroxy positions) should theoretically generate a mixture of acetylated FNDs. Consequently, subjecting the mixture of acetylated FMNs to a reaction and deacetylating the reaction mixture should result in both the byproduct FMN and FND (as depicted in Figure 3.3D). Nonetheless, our observations revealed a significant reduction in yield during the deacetylation step of the acetylated FND mixture by sodium methoxide. This decline in yield can be attributed to FND degradation and the formation of a resembling riboflavin cyclic-4',5'-phosphate, commonly referred to as cyclic FMN or cFMN.[36] Hence, based on our findings, we can draw the conclusion that the purification of the tri-acetylated FMN (FMNAc **8**) holds a crucial role in achieving enhanced yields of the ultimate product.

3.3.4 The enzymatic approach for synthesizing FAD nucleobase analogues

An alternative and efficient method for producing FAD nucleobase analogues involves leveraging enzymes within the FAD biosynthesis pathway. Specifically, *in vitro* studies have demonstrated that FMN adenylyl transferase from *Methanocaldococcus jannaschii* (*Mj*FMNAT) exhibits a notable level of promiscuity when interacting with CTP, leading to the FCD **5** formation, confirmed through LC-MS analysis and chemically synthesized FCD **5** serving as a reference (Figure 3.5A).[22] Our investigation assessed *Mj*FMNAT's reactions with nucleotides ATP, GTP, CTP, UTP, and dATP. Notably, our findings unveiled that beyond CTP, *Mj*FMNAT displays promiscuity towards several other nucleotides, including dATP (as demonstrated in Figure 3.5). Subsequently, we standardized the reaction conditions and timing to establish an effective enzymatic approach catalyzed by *Mj*FMNAT for synthesizing FAD

analogues and yields were reported (Figure 3.5B). A pure *Mj*FMNAT enzyme was procured via heterologous expression in *E. coli*, followed by the reaction of FMN with NTPs/dATP. Remarkably, we observed a more than 90% conversion rate to the corresponding FAD analogues for all nucleotides except UTP (Figure 3.5C). Upon successfully isolating FAD **3**, FGD **4**, FCD **5**, FUD **6**, and dFAD **7** through semi-preparative C18-reversed-phase HPLC purification, the subsequent yields were as follows: 54% for FAD **3**, 51% for FGD **4**, 57% for FCD **5**, 48% for FUD **6**, and 45% for dFAD **7**.

Until this point, enzymatic syntheses of FAD analogues had exclusively employed *C. ammoniagenes* FAD synthetase. However, due to the enzyme's strict selectivity for ATP, the enzymatic synthesis of nucleobase analogues remained unexplored. This study marks a significant analysis by presenting the initial enzymatic synthesis as well as a comprehensive characterization of a diverse range of nucleobase analogues of FAD. Notably, the *Mj*FMNAT enzyme facilitated this achievement, resulting in moderate yields, and this accomplishment expands the possibilities for synthesizing FAD nucleobase analogues enzymatically and brings forth a new avenue for research in this field.

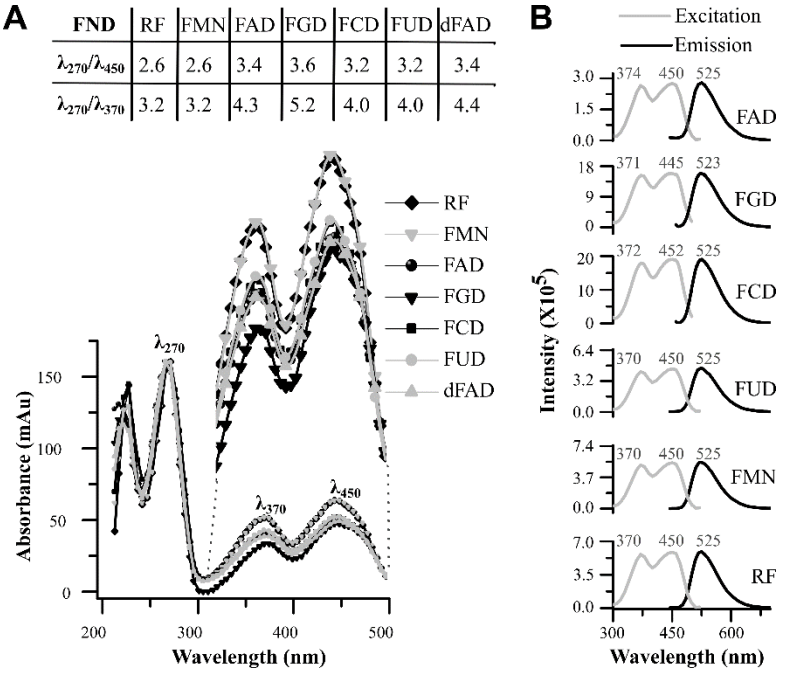


Figure 3.6 UV-visible and fluorescence study of FAD analogues. (A) The ultraviolet-visible spectra of riboflavin (RF), flavin mononucleotide (FMN), flavin adenine dinucleotide (FAD), and the synthesized analogues of FAD were examined. Interestingly, the ratios of λ_{270}: λ_{370} and λ_{270}: λ_{450} for these molecules exhibit variation based on the type of nucleobase present. The calculated values for these ratios are presented in the provided table of the figure. Notably, the ratios for RF and FMN are the same, as are those for FCD and FUD, which contain pyrimidine nucleobases. Conversely, FAD and dFAD display significant spectral overlap. Of particular interest is

FGD, which incorporates guanine as the nucleobase, showing the highest ratio among the tested molecules. In addition, a comparative analysis of the fluorescence spectra (excitation and emission) of FAD and its nucleobase analogues is presented alongside those of RF and FMN. Figure adapted from *ChemBioChem* (**2023**), Shah *et al*.

3.3.5 UV-visible and fluorescence spectroscopy of FAD nucleobase analogues

The nucleobase analogues FAD exhibit peaks at similar wavelengths in their UV-visible spectra. However, we have observed that the ratios of intensities at λ_{270}:λ_{370} and λ_{270}:λ_{450} offer distinctiveness associated with the nucleobase (Figure 3.6A). Interestingly, FAD and dFAD showcase identical spectra, and the similarity of the spectra can also be observed in pyrimidine nucleobase-containing FCD and FUD. Notably, FGD, which incorporates guanine as its nucleobase, stands out by displaying the highest ratio. The specific values for these ratios are presented in Figure 3.6A's accompanying table. These distinctive ratios can potentially be used for identifying the nucleobase analogues of FAD. The fluorescence spectra of the FAD analogues closely resemble that of FAD. However, minor shifts are evident at the excitation (λ_{ex}) and emission (λ_{em}) wavelengths (Figure 3.6B).

3.3.6 FAD nucleobase analogues as active cofactors

Our subsequent objective involved the evaluation of the potential applicability of these analogues of FAD as redox cofactors within enzymatic reactions. There are 94 FAD-binding flavoproteins in *E. coli* which may potentially bind to these FAD nucleobase analogues (Table 3.3). We opted to focus on glutathione reductase from *E. coli* (*Ec*GR) as our selected enzyme (Figure 3.7A). The enzyme employs $FADH_2$ as a means to catalyze the reduction of oxidized glutathione to its monomeric state, simultaneously generating FAD as a product. In the conducted assay, the enzyme-facilitated conversion of FAD to $FADH_2$ takes place through the participation of NADPH, resulting in the formation of $NADP^+$ (Figure 3.7A). This enzymatic transformation leads to a decline in absorbance at 340 nm, attributed to the conversion of NADPH to $NADP^+$, a change that is effectively used to track the progress of this reaction which make us choose this enzyme as our protein of interest to evaluate the applicability of these analogues as redox cofactors.[40] We employed this approach to ascertain the enzymatic capability of *Ec*GR to utilize the synthesized nucleobase analogues of FAD to facilitate its catalytic reaction.

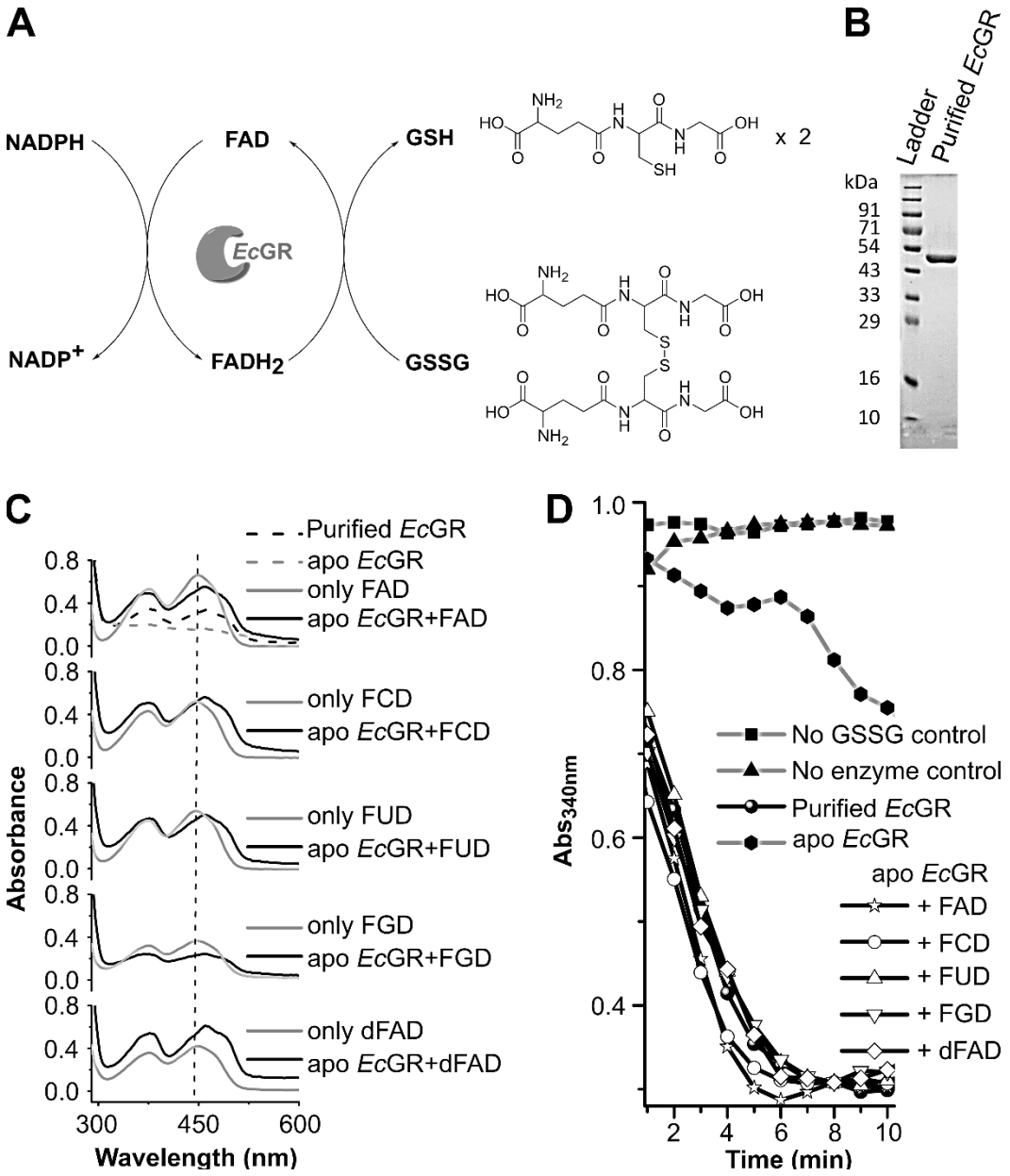


Figure 3.7 The glutathione reductase (*Ec*GR) activity with FAD and its nucleobase analogues. (A) The enzyme *Ec*GR converts oxidized glutathione (GSSG) into its reduced form (GSH), utilizing $FADH_2$ as a cofactor responsible for the reduction process. Concurrently, the oxidized variant of FAD, tightly bound to *Ec*GR, is reduced to form $FADH_2$, with NADPH serving as the electron donor and being transformed into $NADP^+$. Notably, NADPH exhibits absorbance at 340 nm, whereas $NADP^+$ lacks absorbance at this specific wavelength. Consequently, the advancement of the *Ec*GR catalytic reaction is tracked and observed through the decline of the NADPH absorbance at 340 nm. **(B)** This SDS-PAGE gel displays the isolated protein with a mass of 51 kDa, corresponding to *Ec*GR. **(C)** When FND is associated with apo-*Ec*GR, its absorbance peak at 450 nm experiences a redshift. This shift signifies the tightly binding of FAD and its analogues with the apo-enzyme. **(D)** The tight binding facilitates the catalytic reduction of GSSG into GSH. NADPH subsequently conducts the reduction of FND, and the observed decline in absorbance at 340 nm is due to $NADP^+$ formation. The enzymatic activity of *Ec*GR, when each analogue is employed, was quantitatively determined as the reduction rate of GSSG in µmol/min per 1 µM of *Ec*GR enzyme.[40] Activity values in µmol min^{-1} μM^{-1} are as follows: FAD bound purified-*Ec*GR: 0.59, apoenzyme: 0.12, apoenzyme+FAD: 0.72, apoenzyme+FCD: 0.62, apoenzyme+FUD: 0.60, apoenzyme+FGD: 0.52, apoenzyme+dFAD: 0.60. Figure adapted from *ChemBioChem* (**2023**), Shah *et al*.

Due to the strong attachment of FAD as a prosthetic group, *Ec*GR retains FAD during the elution process. To remove the FAD from *Ec*GR, the enzyme was precipitated by saturated ammonium sulfate, subsequently reconstituting the enzyme with different FAD nucleobase analogues. The difference in the absorbance spectra of FAD analogues with enzyme compared to those of the unbound analogues confirms their binding with the enzyme.[49,50] When the enzyme is bound with either FAD or the FAD analogue, a noticeable shift towards longer wavelengths, known as a redshift, is observed in comparison to the absorbance spectra of the unbound analogues, particularly evident at the 440 nm peak (Figure 3.7C). Subsequently, upon reconstitution of *Ec*GR with FAD and various nucleobase analogues, it was determined that all the analogues have the capacity to reduce the oxidized glutathione. However, their efficiencies in this regard vary, with relative activity compared to FAD being as follows: 0.86 for FCD, 0.83 for FUD, 0.72 for FGD, and 0.83 for dFAD. Notably, this was observed despite the enzyme being reconstituted with the cofactor to a range of 80-90% efficiency (Figure 3.7D and Table 3.4).

3.3.7 Engineering glutathione reductase to create a FAD analogue specific mutant

Since, *Ec*GR can accommodate and use each FAD analogue as cofactor, these FAD analogues with engineered flavoprotein can be used to create a bio-orthogonal system for applications in synthetic biology and biotechnology similar to cytosine-based NAD analogue, NCD.[29,33,34] Zhao group used saturation mutagenesis for residue (L310, Q401, L404) of maleic enzyme surrounding adenine base of NAD and got NCD specific mutants.[29] Many L310R or L310K based mutants with Q401 variations were able to use NCD better than NAD. Sequence alignment of maleic enzyme with various NAD utilizing enzyme resulted a common adenine binding motif, zGxGxxG where z is L/V/I belonging to L310 position. Substitution of z residue with R residue in two other enzymes resulted NCD specific mutants. To verify similar system for *Ec*GR with FCD, we tried to mutate *Ec*GR and found similar adenine base surrounding residues I10 and I270 where I10 belongs to zGxGxxG motif (Figure 3.8A). However, the *Ec*GR I10R was found to be insoluble. The I270 residue is just above the amine group of adenine base and creates enough space to accommodate a two membered ring easily (Figure 3.8A). Substitution of I270 with larger residue can fill up the space not allowing the adenine ring fit into cavity while pyrimidine-based FAD analogue, FCD and FUD, can be accommodated. However, I170L, I270K, and I270H based *Ec*GR mutant were found to be active with FAD and its FAD nucleobase analogues similar to the wild-type *Ec*GR (Figure 3.8).

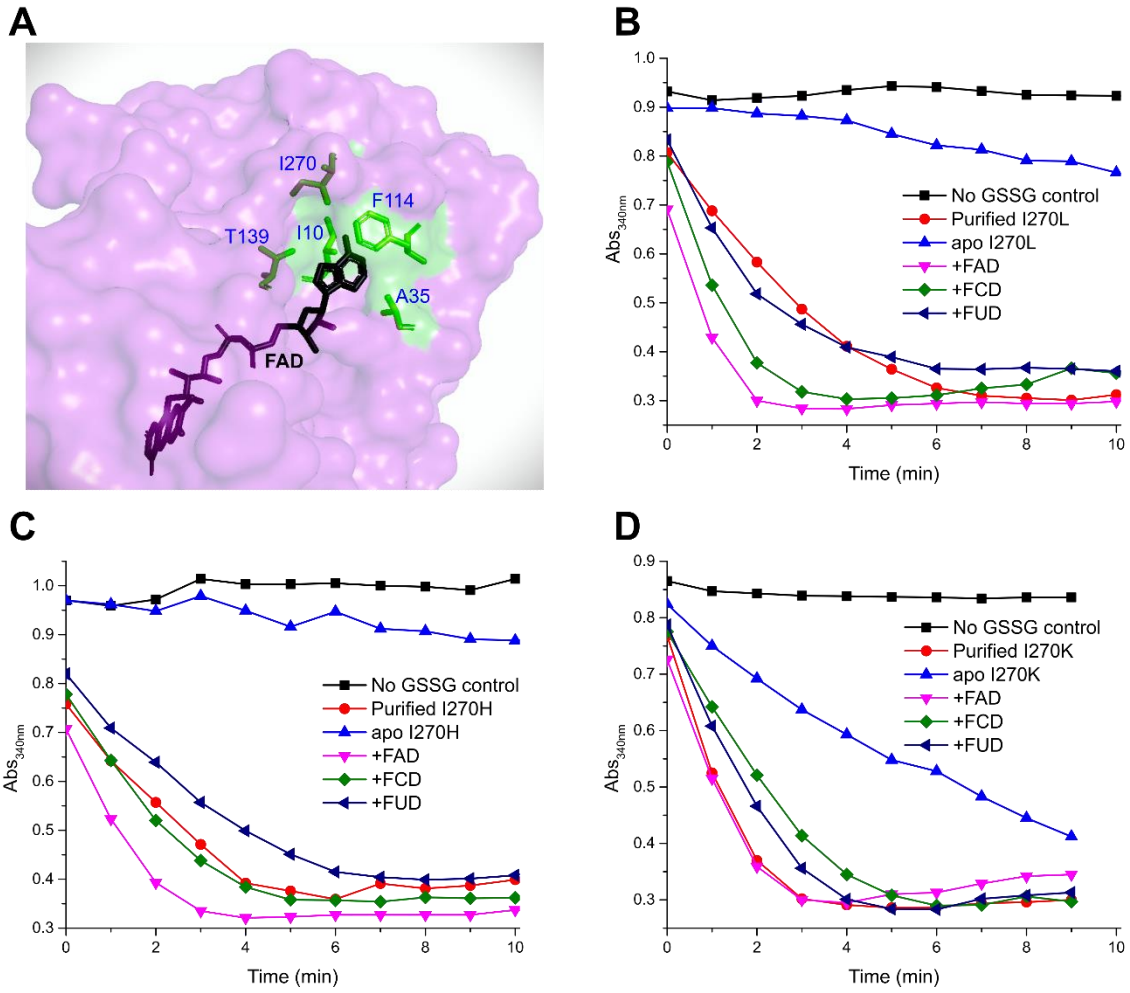


Figure 3.8 mutational analysis in the attempt to create FAD analogue specific *E. coli* glutathione reductase (*Ec*GR). (A) Crystal structure of *Ec*GR showing the surrounding residue to accommodate adenine nucleobase of FAD. The adenine base binds to enzyme through backbone interactions. **(B)** activity assay of *Ec*GR I270L mutant. **(C)** activity assay of *Ec*GR I270H mutant. **(D)** activity assay of *Ec*GR I270K mutant.

3.3.8 Synthesis of FAD analogues within *E. coli* cells

In pursuit of employing these FAD analogues as cofactors within cellular systems, we aimed to ascertain their synthesis within the cell. We conducted heterologous expression of *Mj*FMNAT from pET28a plasmid in *E. coli*. Notably, *E. coli* already possesses the *ribF* gene responsible for encoding FAD synthetase (*Ec*FADS), a pivotal gene that cannot be knocked out. In order to ensure comprehensive evaluation, we incorporated control setups involving an empty plasmid as well as the plasmid containing the *ribF* gene. We suggested that upon the expression of the *Mj*FMNAT-plasmid in the cell and successful synthesis of the FAD analogues,

cellular growth should be comparably hindered in relation to the counterpart *ribF*-plasmid expressing *Ec*FADS.

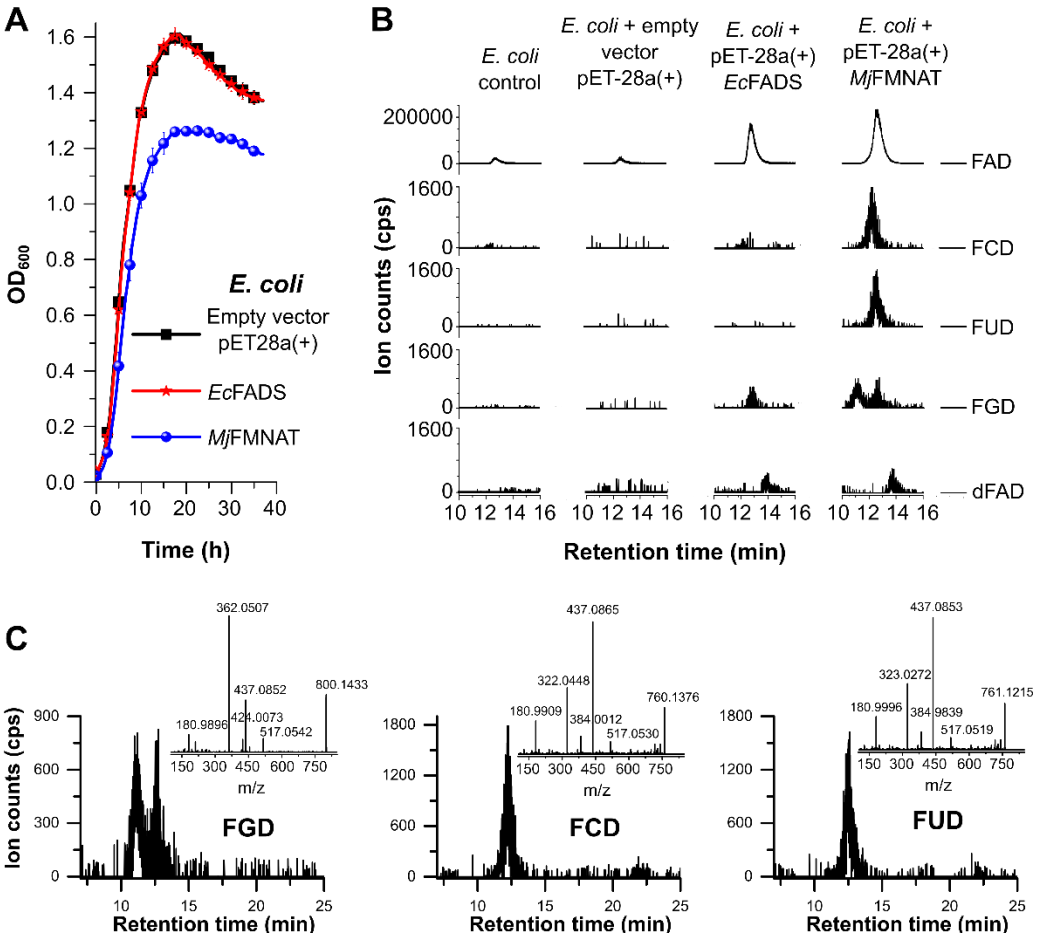


Figure 3.9 FAD nucleobase analogue synthesis in the *E. coli* cells and their impact on the growth of the cell. **(A)** The growth of *E. coli* that has undergone heterologous expression of *Mj*FMNAT using a plasmid exhibits reduced growth compared to that of unmodified *E. coli* or heterologously expressing *Ec*FADS *E. coli* cells. **(B)** LC-MS analysis of the *E. coli* strain expressing *Mj*FMNAT, showing the extracted ion chromatograms (EICs), unveils the synthesis of FAD nucleoside analogues inside the cell. Notably, strains heterologously expressing *Ec*FADS expression and those with empty plasmid did not exhibit the synthesis of these FAD analogues as indicated by the analysis. **(C)** Confirmation of the FAD nucleobase analogues produced in an *E. coli* strain using negative mode LC-ESI-MS analysis. The EIC and MS/MS spectra confirm the identity of FAD nucleobase analogues FGD, FCD, and FUD. The fragmentation of FNDs is a match with the fragmentation of the synthesized standards (Figure 3.5D), where the peak of mass 180.99, 437.08 (cFMN), and 517.05 (riboflavin 5'-pyrophosphate with water removed) are common peaks sourced from FMN moiety of FNDs. For the individual molecules, the characteristic peaks are as follows: FGD: 362.05 (GMP), 424.01 (GDP with H_2O removed), and 800.14 (FGD). FCD: 322.04 (CMP), 384.00 (CDP with H_2O removed), and 760.13 (FCD). FUD: 323.02 (UMP), 384.98 (UDP with H_2O removed), and 761.12 (FUD). Figure adapted from *ChemBioChem* (**2023**), Shah *et al.*

Indeed, we observed lower growth that could arise due to the competition for riboflavin as well as for NTPs within the shared cellular pool. When *E. coli* goes to stationary phase from log phase, the concentration of NTPs (except for UTP) and riboflavin lowers in the cellular

pool which makes the competition higher for these molecules and that may be the reason the growth is affected more in the stationary phase.[51] These FAD analogues might not effectively engage with the cell's FAD-dependent enzymes, leading to metabolic burden or potentially acting as FAD mimics that hinder cellular processes. A growth comparison was conducted among three cultures: the empty vector, the *Ec*FADS-vector, and the *Mj*FMNAT-vector harbouring *E. coli* BL21(DE3) cells. The *Mj*FMNAT-strain demonstrated approximately 0.8 times the growth compared to the growth of the empty vector as well as the *Ec*FADS-vector control (Figure 3.8A). Subsequently, the lysis of the aforementioned cultures and LC-MS analysis of supernatant resulted in interesting observations. *Mj*FMNAT-containing cells show the FAD nucleobase analogue synthesis (FUD, FCD, and FGD), in contrast to the empty-vector and *Ec*FADS-vector strains (Figure 3.8B). At the same time, FAD concentration in cells expressing the *Ec*FADS and *Mj*FMNAT enzymes was relatively higher than in the empty vector cell, as both enzymes contribute to FAD production (Figure 3.8B). Since, *Mj*FMNAT is a thermophilic enzyme, the low concentration of FAD analogues can be attributed to the low activity of the enzyme at 37 °C. However, the *Mj*FMNAT containing strain is producing higher amount of FAD compared to control strain. There can be three more factors for low FAD analogue in the strain. cellular concentration of nucleotides varies in the cell following the order ATP>GTP>UTP>CTP where K_m of the nucleotides can affect the analogue formation.[51] The other factor can be the utilisation of FAD over its analogues by flavoproteins as prosthetic group where FAD is tightly bound to flavoprotein which provides extra stability to FAD while its analogues cannot get this stability via to the flavoproteins. As ATP is native substrate and highest in concentration, enzyme prefers ATP over other nucleotides synthesizing FAD the most.

A comparison of the retention time and fragmentation patterns of the extracted analogues from *E. coli* against the synthesized FAD analogues validates the activity of *Mj*FMNAT in the cell (Figure 3.8C). As a result, we conclude that an appropriate enzyme has the capability to effectively synthesize FAD nucleobase analogues within cellular systems.

3.4 Conclusion

This study shows the successful synthesis of FAD nucleobase analogues while simultaneously investigating their potential to act as catalysts for redox reactions and their capability to be synthesized within cellular systems. Previous routes to synthesize FAD

analogues show challenges such as the availability of starting materials, multistep processes, and low yields. In contrast, our study employed both chemical and enzymatic methodologies to synthesize nucleobase analogues of FAD (FGD, FUD, FCD, and dFAD) within a manageable not more than three steps, resulting in varying yields of 10-51% over a span of 3-5 hours. Moreover, the enzymatic approaches offer advantages over chemical alternatives due to their precision in product formation, the need for small catalytic enzyme amounts, and environment-friendly purification procedures. Furthermore, we successfully demonstrated the practicality of these FAD nucleobase analogues as catalysts within a flavoprotein, glutathione reductase. Lastly, by exploiting the *in vitro Mj*FMNAT promiscuity for NTPs, we showed its capability of utilizing the naturally present NTPs and FMN in the cell, leading to the synthesis of FAD nucleobase analogues in the cellular milieu.

The synthesis methodologies detailed in this study serve as foundational approaches for generating artificial fluorescent and unconventional variants of FAD. Much like analogous applications and studies involving SAM and NAD^+ analogues documented in existing literature, these synthesized FAD analogues hold the potential for similar applications within drug development, cellular imaging, and various biotechnological contexts, as well as bioremediation.[50,52–59] As an illustration, various SAM nucleobase analogues were synthesized and subsequently assessed for their compatibility as cofactors of methyltransferase enzymes that play crucial roles in the methylation of small molecules, and RNA.[59] NCD, the cytosine analogue of NAD was employed as biorthogonal redox systems where in one case, metabolic systems were integrated with engineered NCD-specific enzymes, namely formate dehydrogenase and malic enzyme, to facilitate carbon dioxide (CO_2) fixation within *E. coli*.[29,33,34] Similarly, we anticipate the utilization of FAD nucleobase analogues as a tool to investigate cellular metabolism. Additionally, these analogues hold potential as biorthogonal catalysts, enabling the examination of specific metabolic pathways through enzyme engineering that enables the utilization of an FAD nucleobase analogue in lieu of natural FAD.

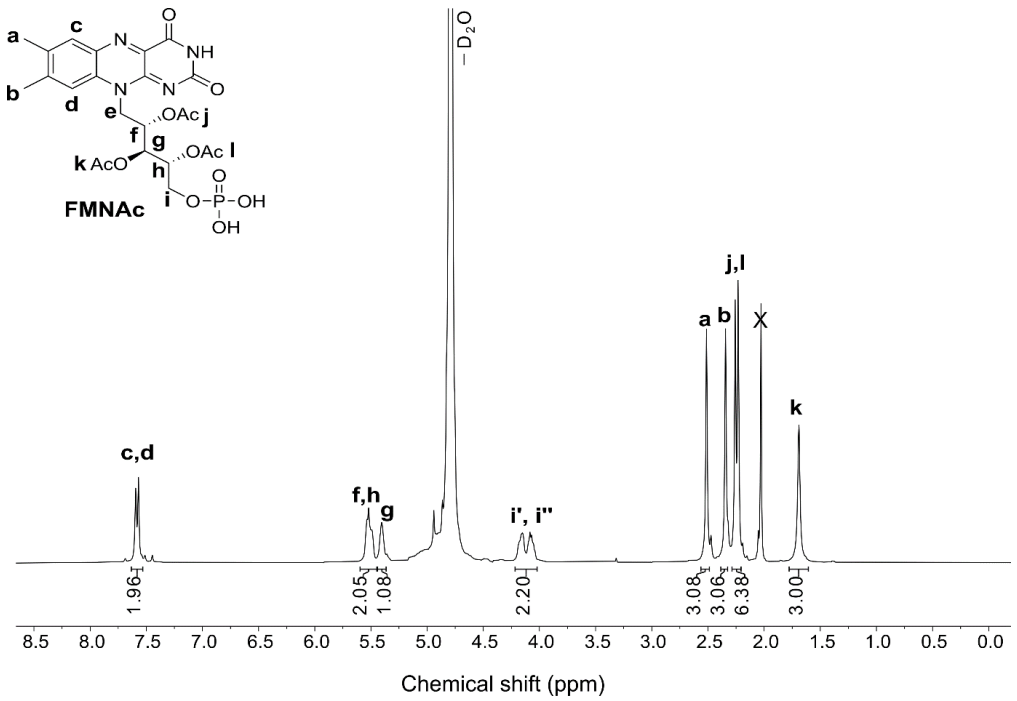


Figure 3.9 ^{1}H NMR (400 MHz, D_2O) of FMNAc (8). The peak with an 'X' is from acetic acid.

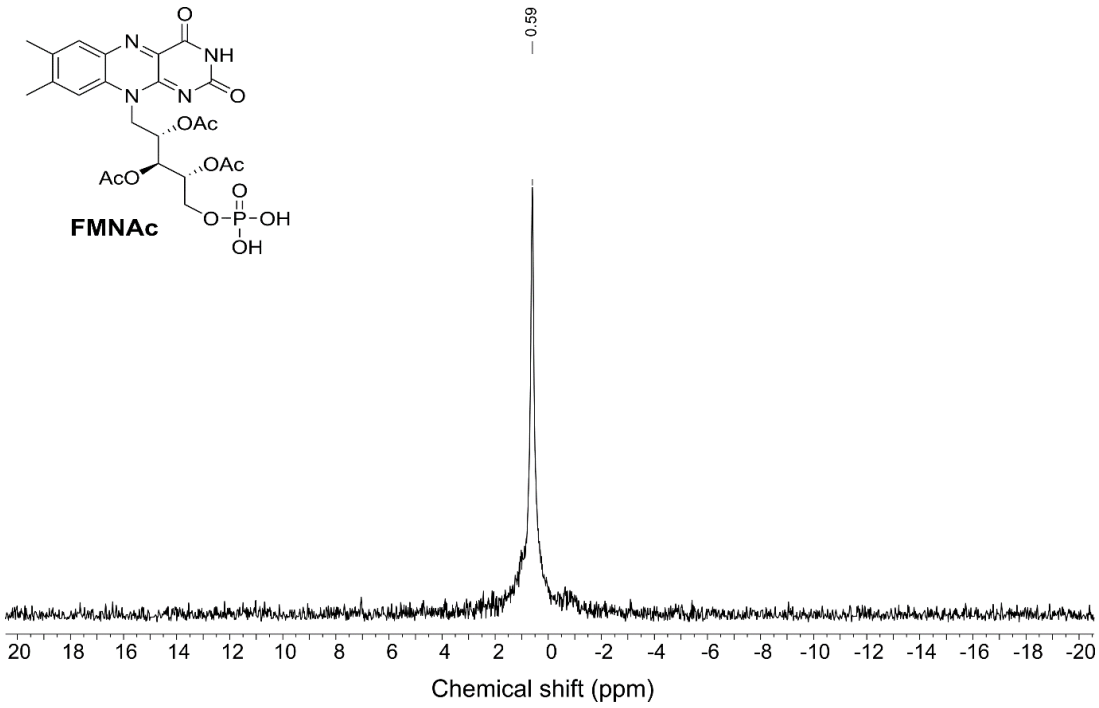


Figure 3.10 ^{31}P NMR (162 MHz, D_2O) of FMNAc (8)

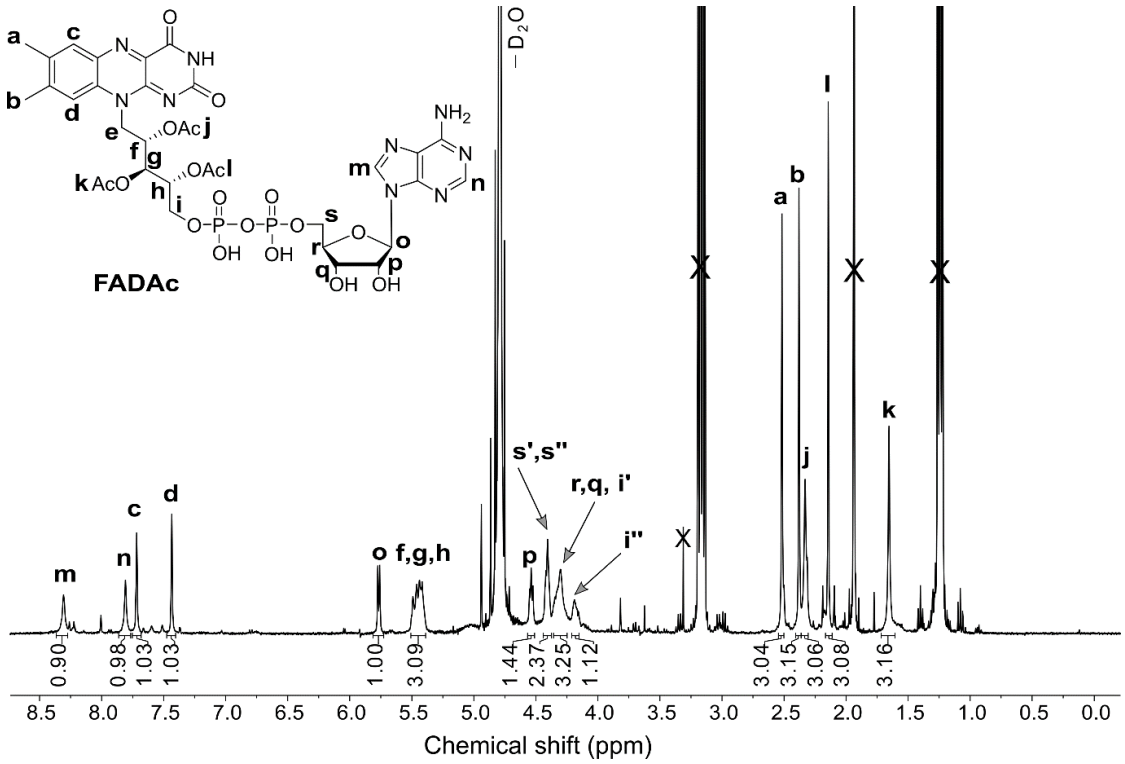


Figure 3.11 ^{1}H NMR (400 MHz, D_2O) of FADAc. The peaks marked with a large 'X' indicate signals arising from triethylammonium acetate, and the peak marked with a small 'x' at 3.31 ppm belongs to a trace amount of methanol solvent.

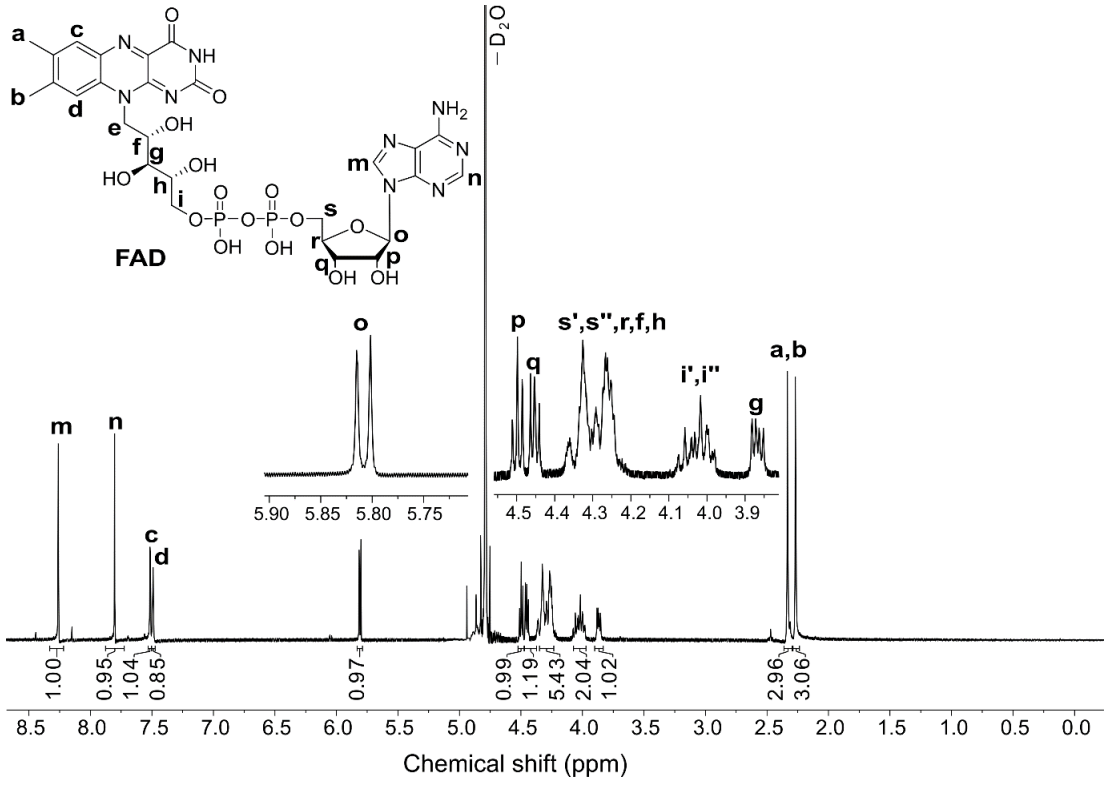


Figure 3.12 ^{1}H NMR (600 MHz, D_2O) of FAD (3)

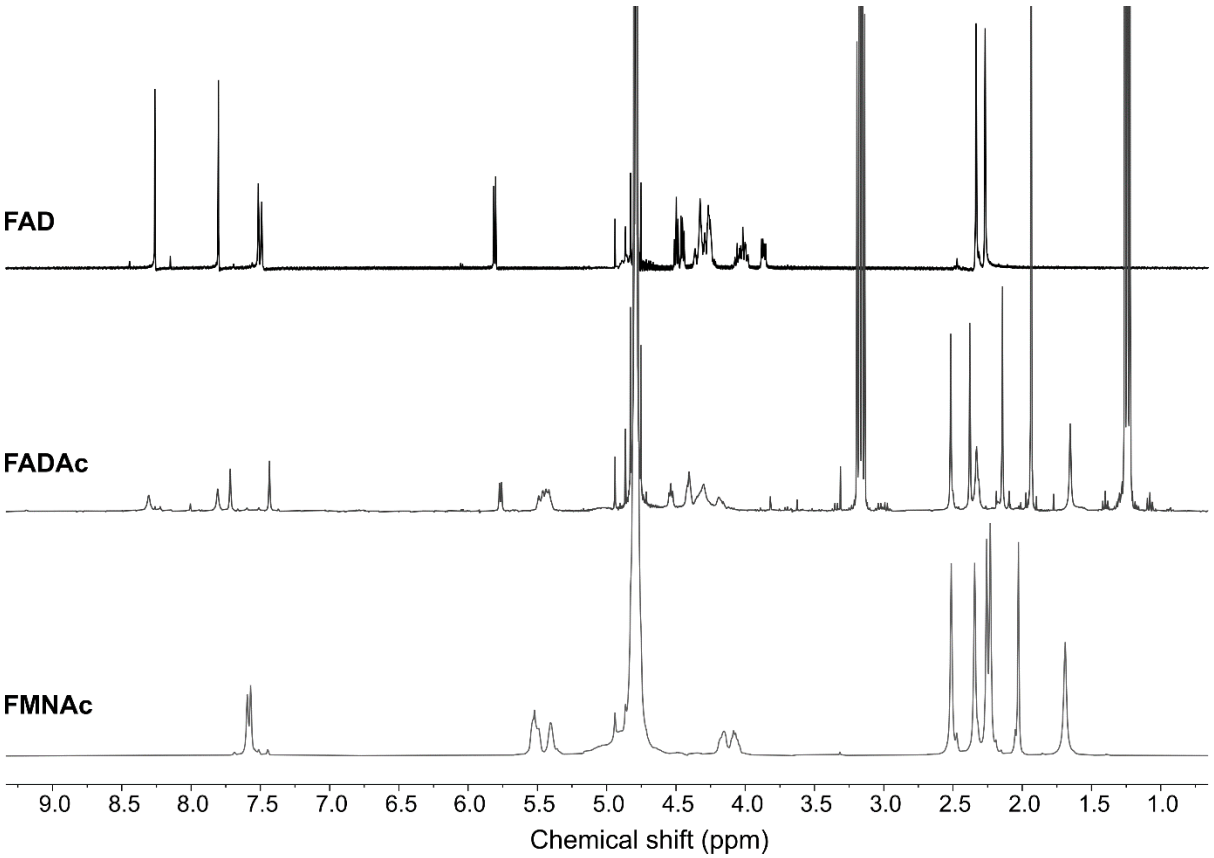


Figure 3.13 Comparison of NMR spectra of intermediates (FMNAc and FADAc) and the final product FAD (overlay of Figures 3.9, 3.11, and 3.12 for ease of comparison)

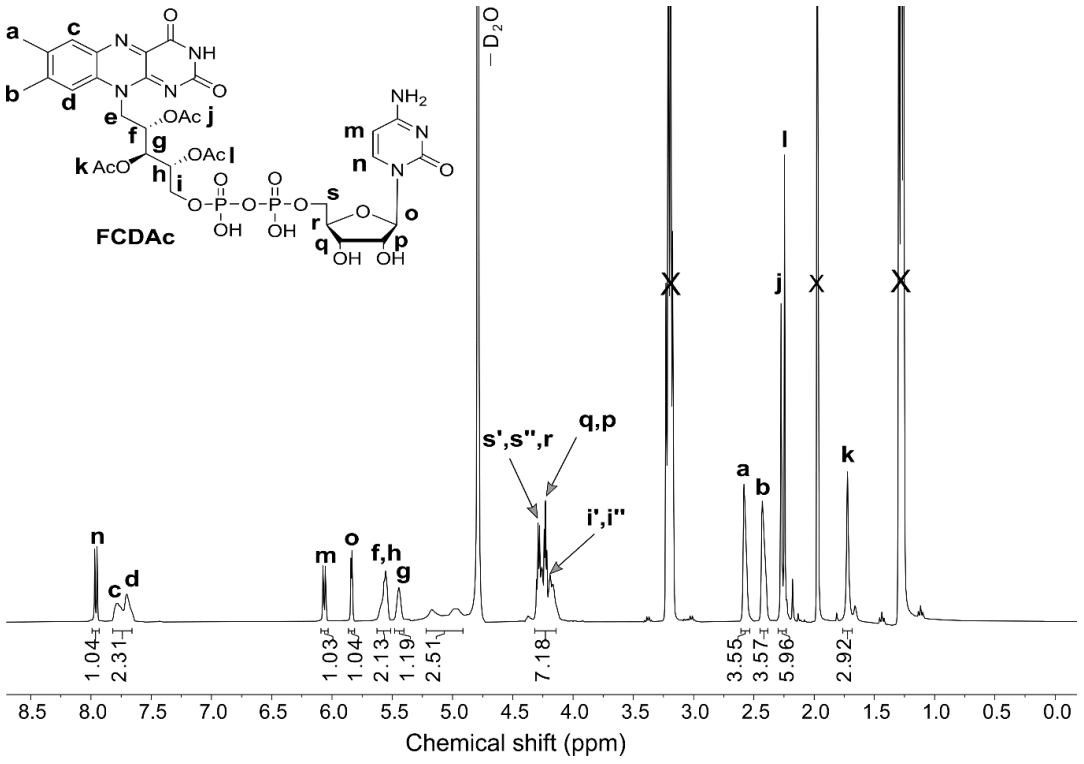


Figure 3.14 ^{1}H NMR (400 MHz, D_2O) of FCDAc (11). The 'X' peaks belong to the triethylammonium acetate.

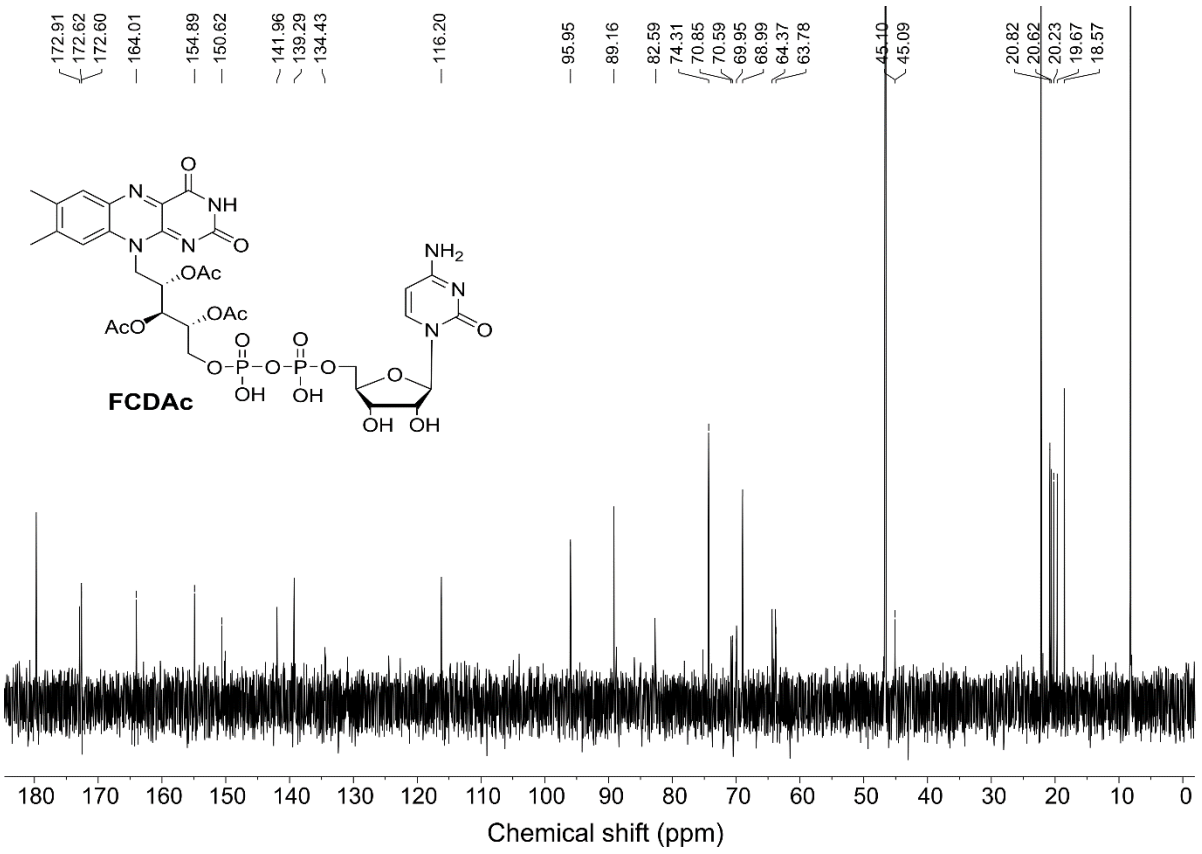


Figure 3.15 ^{13}C NMR (100 MHz, D_2O) of FCDAc (11)

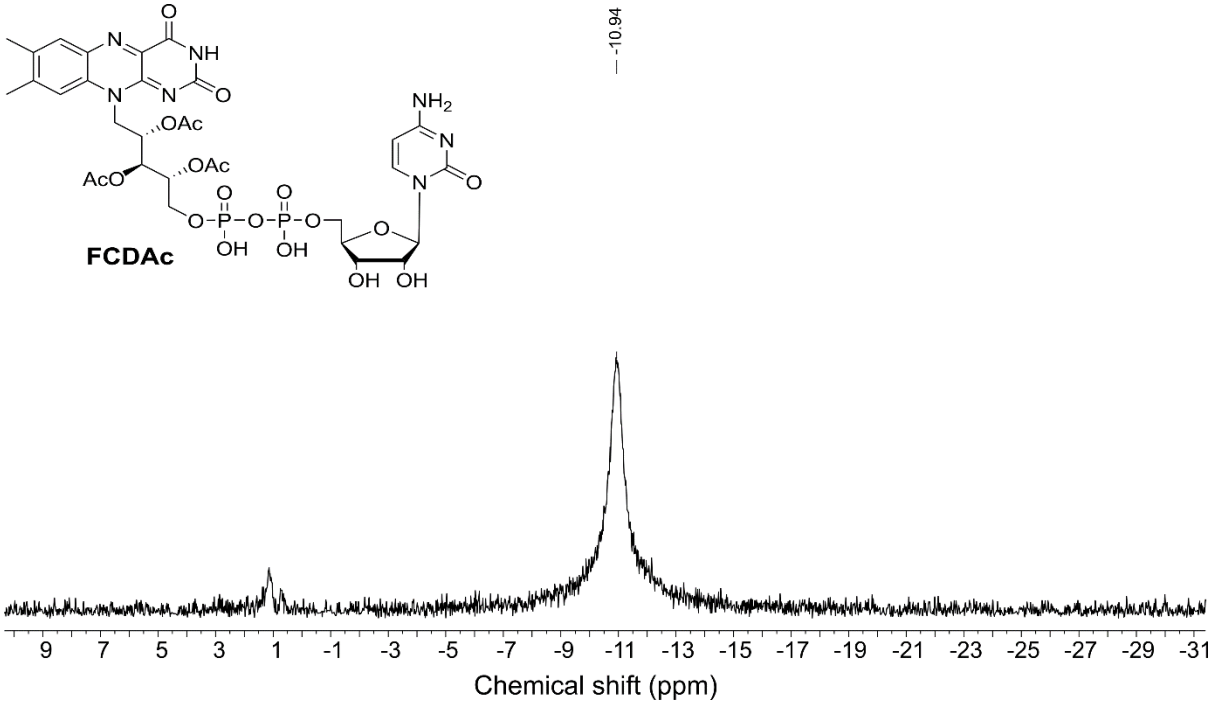


Figure 3.16 ^{31}P NMR (162 MHz, D_2O) of FCDAc (11)

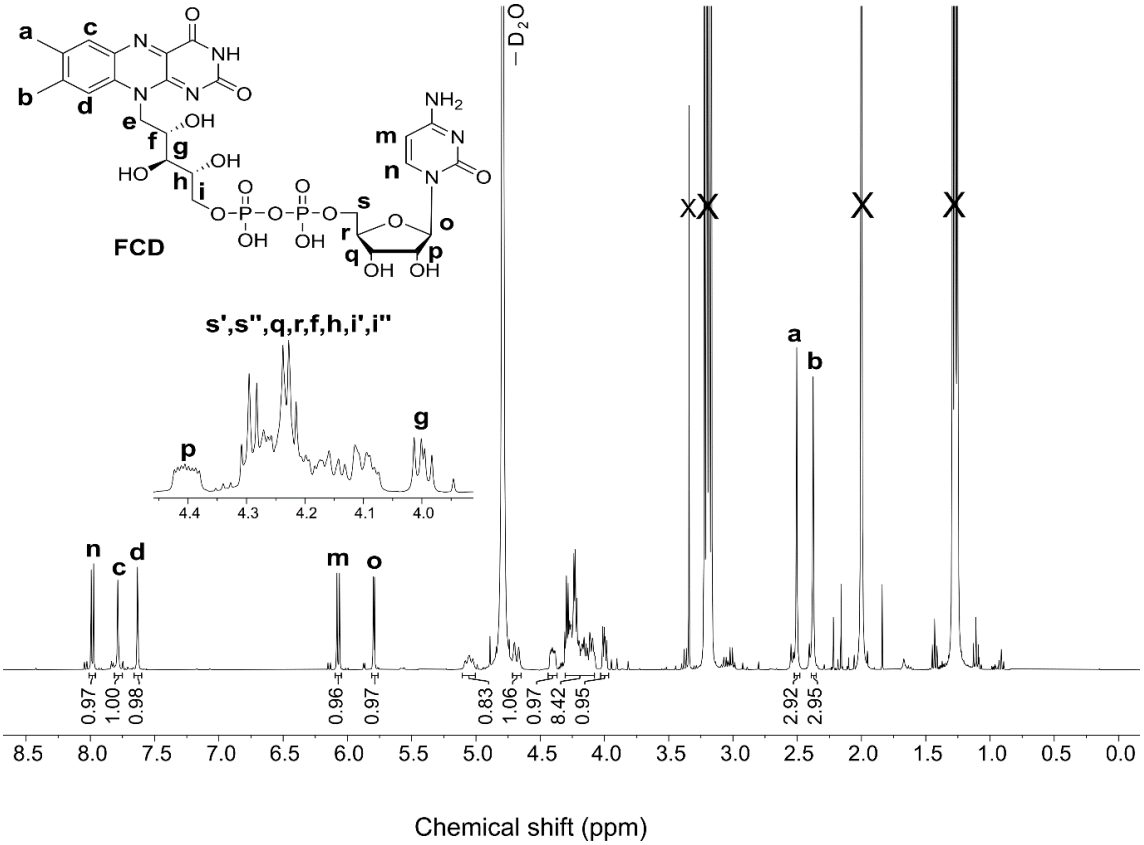


Figure 3.17 ^{1}H NMR (400 MHz, D_2O) of FCD (5). The peak with 'X' belongs to the triethylammonium acetate, and the peak marked with a small 'x' at 3.3 ppm belongs to the trace amount of methanol solvent.

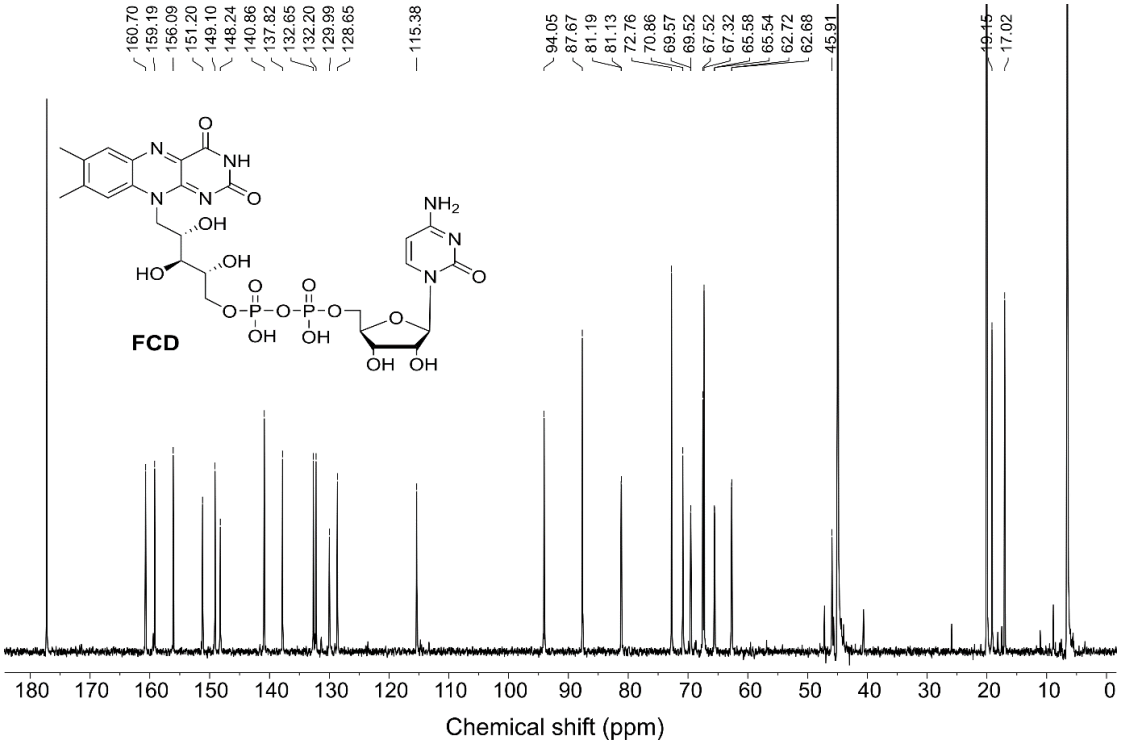


Figure 3.18 ^{13}C NMR (150 MHz, D_2O) of FCD (5)

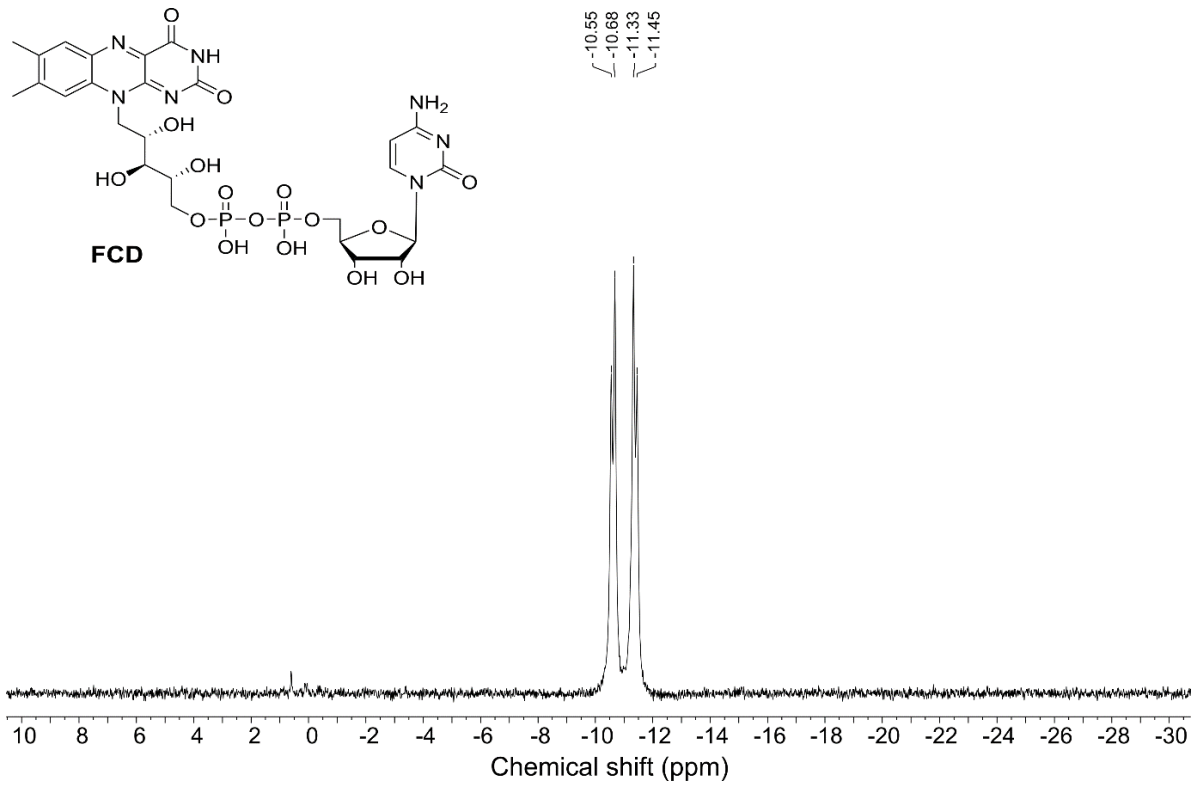


Figure 3.19 ^{31}P NMR (162 MHz, D_2O) of FCD (5)

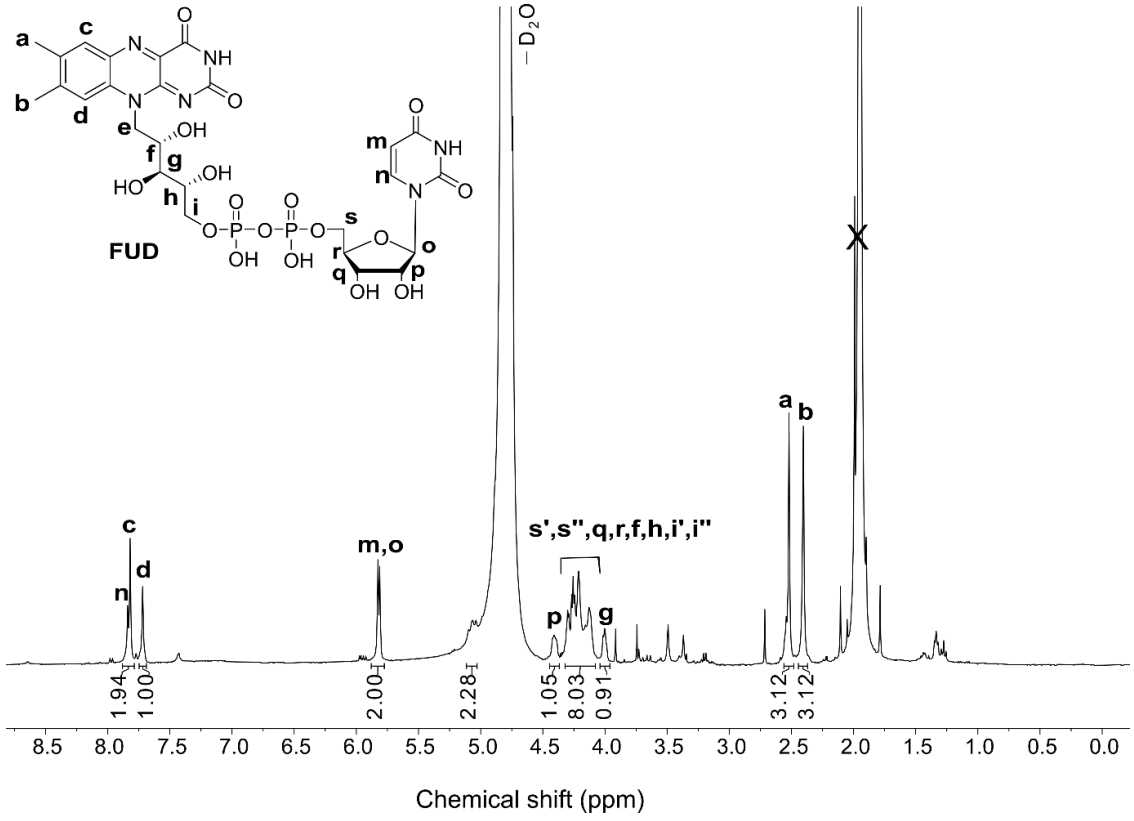


Figure 3.20 ^{1}H NMR (400 MHz, D_2O) of FUD (6). The crossed peak belongs to acetate.

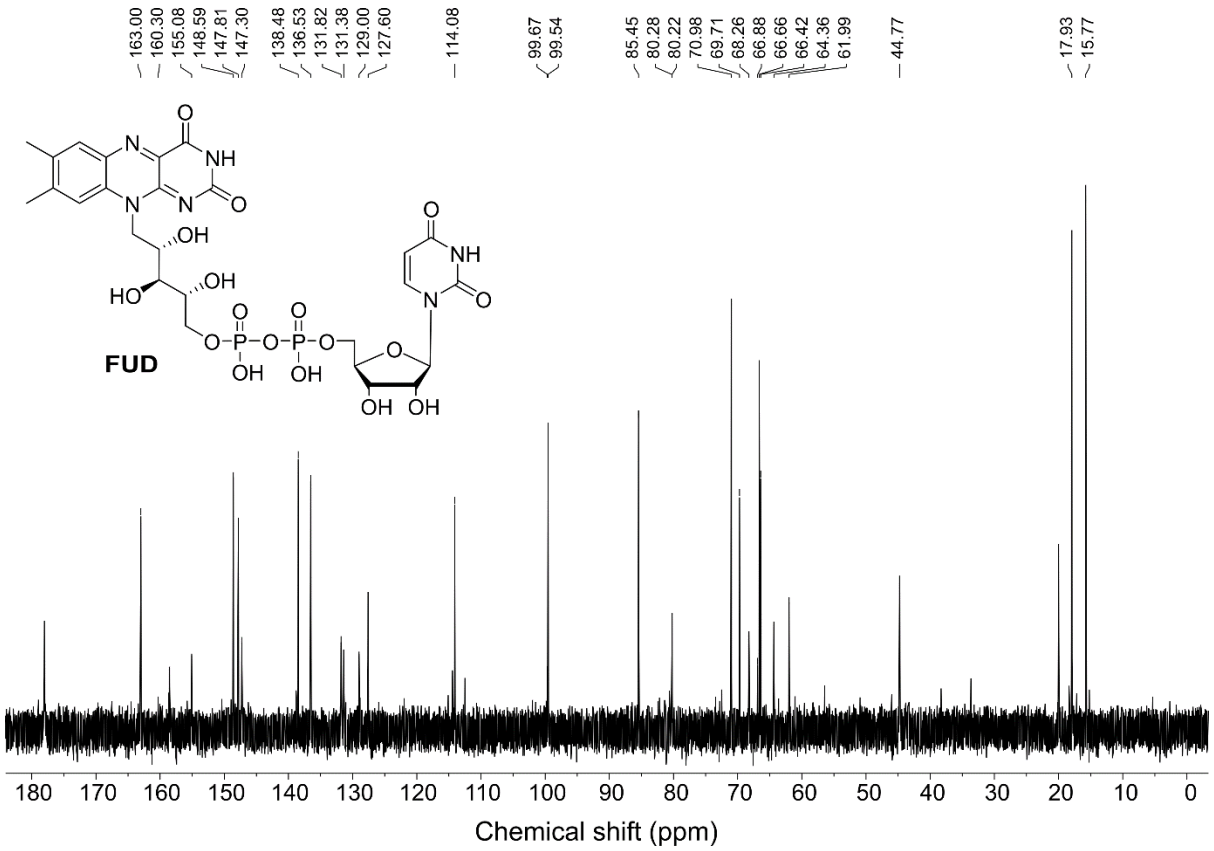


Figure 3.21 ^{13}C NMR (150 MHz, D_2O) of FUD (6)

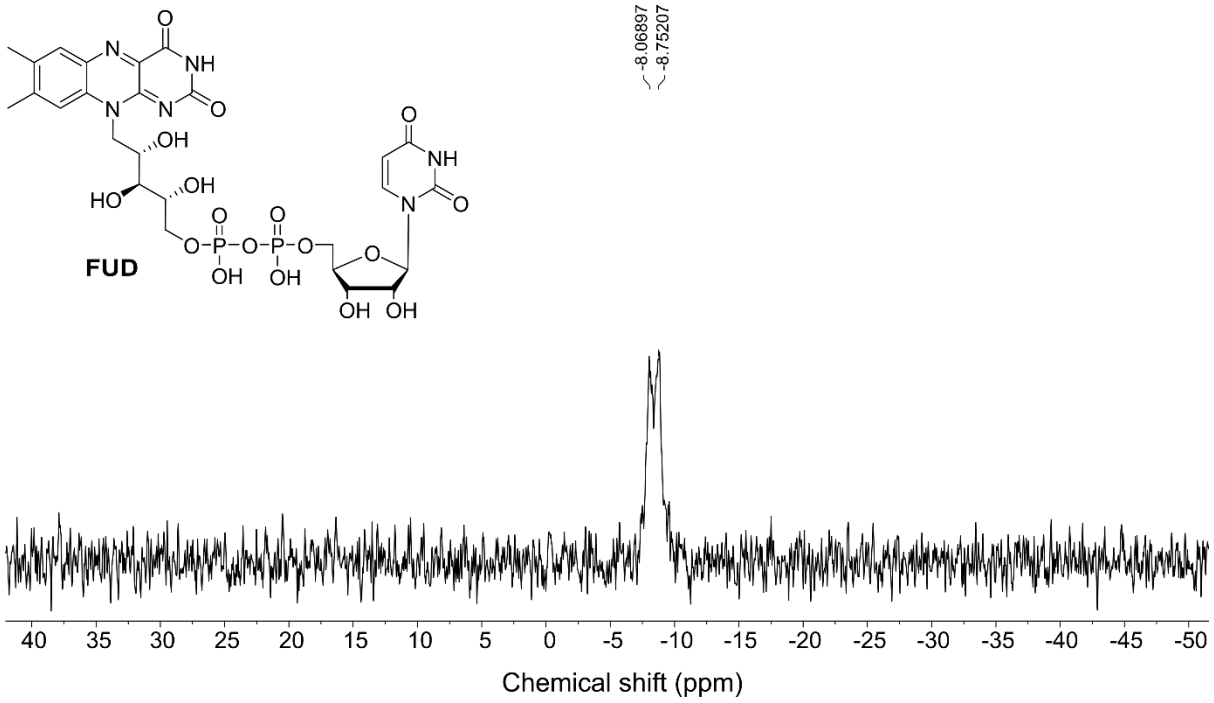


Figure 3.22 ^{31}P NMR (243 MHz, D_2O) of FUD (6)

Table 3.1: Other reported methods for synthesis of FAD and its nucleoside analogues.

S. No.	Method	description	Molecule	Ref.	yield	time	Steps
1	Chemical synthesis	Condensation of 2': 3'-O-isoPropylideneadenosine-5' Benzyl Phosphorochloridate and monothallous riboflavin-5' phosphate in phenol followed by removal of thallous chloride with dry ether. Deprotection in acid	FAD	Christie, Kenner, and Tod. *J. Chem. Soc.* (1954)	6.6%	10 h	4
2	Chemical synthesis	Coupling of FMN and AMP using di-p-tolyl carbodiimide in pyridine	FAD FID	Huennekens and Kilgour. *JACS* (1955)	<4%	24 h	1
3	Large-scale chemical synthesis	Direct condensation of FMN and NMP in the presence of trifluoroacetic acid anhydride	FAD FGD FCD FUD	DeLuca and Kaplan. *JBC* (1956)	<1%	16 h	1
			FCD	Mashhadi et al. *Biochemistry* (2010)	Not reported	18 h	1
4	Chemical synthesis	Coupling of adenosine-5' phosphoramidate or phosphoromorpholidate and riboflavin-5' phosphate in a mixture of pyridine and o-chlorophenol	FAD	Moffatt and Khorana. *JACS* (1958)	40%	4 d	2
			FdeAD FGD, FdeGD FCD, FdeCD FTD FHD, FdeHD FXD, FdeXD	McCormick, Chassy and Tsibris. *Biochimica et Biophysica Acta* (1964)	-	-	2
			FdeAD FGD, FdeGD FCD, FdeCD FTD FHD, FdeHD FXD, FdeXD	Tsibris, McCormick, and Wright. *Biochemistry* (1965)	-	-	2

			Flavin-8-bromoadenine dinucleotide	McCormick and Opar. *Journal of Medicinal Chemistry* (1969)	28%	7 d	2
5	Chemical synthesis	Mono (tri-n-octylammonium) salt of Adenosine-5' phosphate coupled with Diphenyl phosphochloridate DMF and dioxan followed by attack of mono (tri-n-octylammonium) flavin mononucleotide in DMF and pyridine.	FAD	Michelson. *Biochimica et Biophysica Acta* (1964)	70-80%	5 h	2
			5-deazaflavin adenine dinucleotide	Smit. *Recueil des Travaux Chimiques des Pays-Bas* (1984)	-	12-15 h	3
			Azido-FAD	Koberstein. *European Journal of Biochemistry* (1976)	25%	5 h	2
			N^6-(6-Carboxyhexyl)-FAD	Stocker, Hecht, And Buckmann. *European Journal of Biochemistry* (1996)	30%	12-17 h	4
6	Chemical synthesis	Coupling of nucleoside S'-phosphorothioate with the ammonium salt of FMN in dry pyridine	FAD	Hata and Nakagawa. *JACS* (1970)	51%	5 h	1
7	Enzymatic synthesis	Coupling of FMN and NTP using *Ca*FADS	-	Spencer, Fisher, and Walsh. *Biochemistry* (1976)	-	4-18 h	1

Table 3.2: List of 24 FAD-binding enzymes and their RMSD values.

S. No.	PDB ID	Name of Enzyme	Organism Name	RMSD (wrt PDB: 3R9U) (Å)	Distance (Å) between terminal backbone atoms (N9A-N10)
1	1EGC	Medium Chain Acyl CoA dehydrogenase	*Human*	2.974	13.807
2	2G37	L-Proline dehydrogenase	*Thermus thermophilus*	3.401	12.171
3	3R9U	Thioredoxin disulfide reductase	*Campylobacter jejuni*	0	15.222
4	6BZ0	Dihydrolipoyl dehydrogenase	*Acinetobacter baumannii*	0.801	15.156
5	2VFR	Alditol Oxidase	*Streptomyces coelicolor*	3.267	12.845
6	4I58	Cyclohexylamine Oxidase	*Microbacterium oxydans*	0.388	15.557
7	1TZL	Pyranose-2-oxidase	*Peniophora sp*	0.77	15.137
8	1F0X	D-lactate dehydrogenase	*E. coli*	3.036	12.942
9	1OJD	Monoamine Oxidase B	*Human*	0.488	15.334
10	2BVH	6-hydroxy-D-nicotine oxidase	*Paenarthrobacter nicotinovorans*	3.584	12.127
11	7PBG	4-ethyl phenol oxidase	*Gulosibacter chungangensis*	3.08	13.735
12	6PXS	iminodiacetate oxidase	*Chelativorans sp. BNC1*	0.927	16.225
13	3PQB	GilR	*Streptomyces griseoflavus*	3.342	12.487
14	1E8G	Vanillyl Alcohol Oxidase	*Penicillium simplicissimum*	3.056	13.491
15	1W1O	Cytokinin Dehydrogenase	*Zea mays*	3.477	12.312
16	4E0H	Sulfhydryl oxidase erv1	*Saccharomyces cerevisiae S288C*	3.505	11.508
17	4DQK	Cytochrome P450 BM3	*Priestia megaterium*	3.019	12.532
18	6ECI	MSMEG_5243	*Mycolicibacterium smegmatis MC2 155*	4.295	8.786
19	5EZ7	oxidoreductase PA4991	*Pseudomonas aeruginosa PAO1*	0.816	15.982
20	2I0Z	FAD utilizing dehydrogenase	*Bacillus cereus*	0.368	15.416
21	1TVC	methane monooxygenase reductase	*Methylococcus capsulatus*	3.907	13.32
22	1JR8	Sulfhydryl oxidase Erv2p	*Saccharomyces cerevisiae*	3.515	11.422
23	2WSI	FAD Synthetase	*Saccharomyces cerevisiae*	4.31	8.163
24	2OAL	Tryptophan halogenase-RebH	*Lentzea aerocolonigenes*	1.324	16.159

Table 3.3: List of all *E. coli* FAD-binding flavoprotein obtained from UniProt database.[60]

UniProt Entry	UniProt Entry Name	Protein names	Gene Names	Length
P00363	FRDA_ECOLI	Fumarate reductase flavoprotein subunit (EC 1.3.5.1) (Quinol-fumarate reductase flavoprotein subunit) (QFR flavoprotein subunit)	frdA b4154 JW4115	602
P00393	NDH_ECOLI	Type II NADH:quinone oxidoreductase (EC 1.6.5.9) (Cupric reductase) (EC 1.16.1.-) (NADH dehydrogenase-2) (NADH dh II) (NDH-2)	ndh b1109 JW1095	434
P00893	ILVI_ECOLI	Acetolactate synthase isozyme 3 large subunit (EC 2.2.1.6) (AHAS-III) (ALS-III) (Acetohydroxy-acid synthase III large subunit)	ilvI b0077 JW0076	574
P00914	PHR_ECOLI	Deoxyribodipyrimidine photo-lyase (EC 4.1.99.3) (DNA photolyase) (Photoreactivating enzyme)	phrB phr b0708 JW0698	472
P06149	DLD_ECOLI	Quinone-dependent D-lactate dehydrogenase (EC 1.1.5.12) ((R)-lactate:quinone 2-oxidoreductase)	dld b2133 JW2121	571
P06715	GSHR_ECOLI	Glutathione reductase (GR) (GRase) (EC 1.8.1.7)	gor b3500 JW3467	450
P07003	POXB_ECOLI	Pyruvate dehydrogenase [ubiquinone] (EC 1.2.5.1) (Pyruvate oxidase) (POX) (Pyruvate:ubiquinone-8 oxidoreductase)	poxB b0871 JW0855	572
P08142	ILVB_ECOLI	Acetolactate synthase isozyme 1 large subunit (AHAS-I) (EC 2.2.1.6) (Acetohydroxy-acid synthase I large subunit) (ALS-I)	ilvB b3671 JW3646	562
P08201	NIRB_ECOLI	Nitrite reductase (NADH) large subunit (EC 1.7.1.15)	nirB b3365 JW3328	847
P08373	MURB_ECOLI	UDP-N-acetylenolpyruvoylglucosamine reductase (EC 1.3.1.98) (UDP-N-acetylmuramate dehydrogenase)	murB yijB b3972 JW3940	342
P09546	PUTA_ECOLI	Bifunctional protein PutA [Includes: Proline dehydrogenase (EC 1.5.5.2)	putA poaA b1014 JW0999	1320
P09831	GLTB_ECOLI	Glutamate synthase [NADPH] large chain (EC 1.4.1.13) (Glutamate synthase subunit alpha) (GLTS alpha chain) (NADPH-GOGAT)	gltB aspB b3212 JW3179	1486
P0A6J5	DADA_ECOLI	D-amino acid dehydrogenase (EC 1.4.99.-) (D-alanine dehydrogenase)	dadA dadR b1189 JW1178	432
P0A6U3	MNMG_ECOLI	tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG (Glucose-inhibited division protein A)	mnmG gidA trmF b3741 JW3719	629
P0A9C0	GLPA_ECOLI	Anaerobic glycerol-3-phosphate dehydrogenase subunit A (G-3-P dehydrogenase) (EC 1.1.5.3)	glpA b2241 JW2235	542
P0A9P0	DLDH_ECOLI	Dihydrolipoyl dehydrogenase (EC 1.8.1.4)	lpdA lpd b0116 JW0112	474
P0A9P4	TRXB_ECOLI	Thioredoxin reductase (TRXR) (EC 1.8.1.9)	trxB b0888 JW0871	321
P0AC41	SDHA_ECOLI	Succinate dehydrogenase flavoprotein subunit (EC 1.3.5.1)	sdhA b0723 JW0713	588
P0AEN1	FRE_ECOLI	NAD(P)H-flavin reductase (EC 1.5.1.41) (FMN reductase) (Ferrisiderophore reductase C) (NAD(P)H:flavin oxidoreductase)	fre fadI flrD fsrC ubiB b3844 JW3820	233
P0AEP9	GLCD_ECOLI	Glycolate oxidase subunit GlcD (EC 1.1.99.14) (Glycolate dehydrogenase subunit GlcD)	glcD gox yghM b2979 JW2946	499
P0AEY5	MDAB_ECOLI	NADPH:quinone oxidoreductase MdaB (EC 1.6.5.10) (Modulator of drug activity B)	mdaB mda66 b3028 JW2996	193
P0AEZ1	METF_ECOLI	5,10-methylenetetrahydrofolate reductase (MTHFR) (EC 1.5.1.54) (NADH-dependent 5,10-methylenetetrahydrofolate reductase)	metF b3941 JW3913	296

P0DP90	ILVG_ECOLI	Acetolactate synthase isozyme 2 large subunit (AHAS-II) (EC 2.2.1.6) (ALS-II) (Acetohydroxy-acid synthase II large subunit)	ilvG	548
P10902	NADB_ECOLI	L-aspartate oxidase (LASPO) (EC 1.4.3.16) (L-aspartate:fumarate oxidoreductase) (EC 1.5.99.-) (Quinolinate synthetase B)	nadB nicB b2574 JW2558	540
P12008	AROC_ECOLI	Chorismate synthase (CS) (EC 4.2.3.5) (5-enolpyruvylshikimate-3-phosphate phospholyase) (EPSP phospholyase)	aroC b2329 JW2326	361
P13035	GLPD_ECOLI	Aerobic glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)	glpD glyD b3426 JW3389	501
P17444	BETA_ECOLI	Oxygen-dependent choline dehydrogenase (CDH) (CHD) (EC 1.1.99.1) (Betaine aldehyde dehydrogenase) (BADH) (EC 1.2.1.8)	betA b0311 JW0303	556
P24232	HMP_ECOLI	Flavohemoprotein (Flavohemoglobin) (HMP) (Hemoglobin-like protein) (Nitric oxide dioxygenase) (NOD) (EC 1.14.12.17)	hmp fsrB hmpA b2552 JW2536	396
P25534	UBIH_ECOLI	2-octaprenyl-6-methoxyphenol hydroxylase (EC 1.14.13.-)	ubiH visB b2907 JW2875	392
P25535	UBII_ECOLI	2-octaprenylphenol hydroxylase (EC 1.14.13.240) (2-polyprenylphenol 6-hydroxylase)	ubiI visC b2906 JW2874	400
P27306	STHA_ECOLI	Soluble pyridine nucleotide transhydrogenase (STH) (EC 1.6.1.1) (NAD(P)(+) transhydrogenase [B-specific])	sthA sth udhA b3962 JW5551	466
P28861	FENR_ECOLI	Flavodoxin/ferredoxin--NADP reductase (EC 1.18.1.2) (EC 1.19.1.1)	fpr mvrA b3924 JW3895	248
P33224	AIDB_ECOLI	Putative acyl-CoA dehydrogenase AidB (EC 1.3.99.-)	aidB b4187 JW5867	541
P35340	AHPF_ECOLI	Alkyl hydroperoxide reductase subunit F (EC 1.8.1.-) (Alkyl hydroperoxide reductase F52A protein)	ahpF b0606 JW0599	521
P37339	LHGD_ECOLI	L-2-hydroxyglutarate dehydrogenase (L2HG dehydrogenase) (EC 1.1.5.13) (L2HG:quinone oxidoreductase)	lhgD lhgO ygaF b2660 JW2635	422
P37747	GLF_ECOLI	UDP-galactopyranose mutase (UGM) (EC 5.4.99.9) (UDP-GALP mutase) (Uridine 5-diphosphate galactopyranose mutase)	glf yefE b2036 JW2021	367
P38038	CYSJ_ECOLI	Sulfite reductase [NADPH] flavoprotein alpha-component (SiR-FP) (EC 1.8.1.2)	cysJ b2764 JW2734	599
P40874	MTOX_ECOLI	N-methyl-L-tryptophan oxidase (MTOX) (EC 1.5.3.-)	solA b1059 JW1046	372
P42593	FADH_ECOLI	2,4-dienoyl-CoA reductase [(2E)-enoyl-CoA-producing] (DCR) (EC 1.3.1.34) (2,4-dienoyl-coenzyme A reductase [NADPH])	fadH ygjL b3081 JW3052	672
P52073	GLCE_ECOLI	Glycolate oxidase subunit GlcE (EC 1.1.99.14) (Glycolate dehydrogenase subunit GlcE)	glcE gox yghL b4468 JW5487	350
P75990	BLUF_ECOLI	Blue light- and temperature-regulated antirepressor BluF (Blrp)	bluF blrp ycgF b1163 JW1150	403
P76081	PAAE_ECOLI	1,2-phenylacetyl-CoA epoxidase, subunit E	paaE ydbR b1392 JW1387	356
P77182	MNMC_ECOLI	tRNA 5-methylaminomethyl-2-thiouridine biosynthesis bifunctional protein MnmC	mnmC yfcK b2324 JW5380	668
P77212	RCLA_ECOLI	Probable pyridine nucleotide-disulfide oxidoreductase RclA (Reactive chlorine resistance protein A)	rclA ykgC b0304 JW5040	441
P77324	PAOB_ECOLI	Aldehyde oxidoreductase FAD-binding subunit PaoB (EC 1.2.99.6)	paoB yagS b0285 JW0279	318
P77748	D2HDH_ECOLI	D-2-hydroxyglutarate dehydrogenase (D2HGDH) (EC 1.1.99.39)	ydiJ b1687 JW1677	1018

Q46871	YQJH_ECOLI	NADPH-dependent ferric-chelate reductase (EC 1.16.1.9) (Ferric siderophore reductase)	yqjH b3070 JW3041	254
Q47146	FADE_ECOLI	Acyl-coenzyme A dehydrogenase (ACDH) (Acyl-CoA dehydrogenase) (EC 1.3.8.7) (EC 1.3.8.8)	fadE yafH b0221 JW5020	814
P0AB85	APBE_ECOLI	FAD:protein FMN transferase (EC 2.7.1.180) (Flavin transferase)	apbE yojK yojL b2214 JW5368	351
P0AG40	RIBF_ECOLI	Bifunctional riboflavin kinase/FMN adenylyltransferase [Includes: Riboflavin kinase (EC 2.7.1.26)	ribF yaaC b0025 JW0023	313
P33940	MQO_ECOLI	Malate:quinone oxidoreductase (EC 1.1.5.4) (MQO) (Malate dehydrogenase [quinone])	mqo yojH b2210 JW2198	548
P37596	NORW_ECOLI	Nitric oxide reductase FlRd-NAD(+) reductase (EC 1.18.1.-) (Flavorubredoxin reductase) (FlRd-reductase) (FlavoRb reductase)	norW flrR ygbD b2711 JW2681	377
P50466	AER_ECOLI	Aerotaxis receptor	aer air yqjJ b3072 JW3043	506
P60584	CAIA_ECOLI	Crotonobetainyl-CoA reductase (EC 1.3.8.13) (Crotobetaine reductase) (Crotonobetainyl-CoA dehydrogenase)	caiA yaaO b0039 JW0038	380
P75728	UBIF_ECOLI	3-demethoxyubiquinol 3-hydroxylase (EC 1.14.99.60) (2-octaprenyl-3-methyl-6-methoxy-1,4-benzoquinol hydroxylase)	ubiF yleB b0662 JW0659	391
P75824	HCR_ECOLI	NADH oxidoreductase HCR (EC 1.-.-.-)	hcr ybjV b0872 JW5117	322
P76352	YEEO_ECOLI	Probable FMN/FAD exporter YeeO	yeeO b1985 JW1965	547
P77397	MHPA_ECOLI	3-(3-hydroxy-phenyl)propionate/3-hydroxycinnamic acid hydroxylase (3-HCI hydroxylase) (3-HPP hydroxylase) (EC 1.14.13.127)	mhpA b0347 JW0338	554
P77650	HCAD_ECOLI	3-phenylpropionate/cinnamic acid dioxygenase ferredoxin--NAD(+) reductase component (EC 1.18.1.3)	hcaD hcaA4 phdA yfhY b2542 JW2526	400
Q46800	XDHB_ECOLI	Putative xanthine dehydrogenase FAD-binding subunit XdhB (EC 1.17.1.4)	xdhB ygeT b2867 JW2835	292
B1X935	MNMC_ECODH	tRNA 5-methylaminomethyl-2-thiouridine biosynthesis bifunctional protein MnmC	mnmC ECDH10B_2486	668
B1X9K2	AROC_ECODH	Chorismate synthase (CS) (EC 4.2.3.5) (5-enolpyruvylshikimate-3-phosphate phospholyase)	aroC ECDH10B_2491	361
B1XCN8	NORW_ECODH	Nitric oxide reductase FlRd-NAD(+) reductase (EC 1.18.1.-) (Flavorubredoxin reductase) (FlRd-reductase) (FlavoRb reductase)	norW flrR ECDH10B_2879	377
B1XE52	BETA_ECODH	Oxygen-dependent choline dehydrogenase (CDH) (CHD) (EC 1.1.99.1) (Betaine aldehyde dehydrogenase) (BADH) (EC 1.2.1.8)	betA ECDH10B_0298	556
C4ZVM2	AROC_ECOBW	Chorismate synthase (CS) (EC 4.2.3.5) (5-enolpyruvylshikimate-3-phosphate phospholyase)	aroC BWG_2103	361
P0DP89	ILVGP_ECOLI	Putative acetolactate synthase isozyme 2 large subunit (AHAS-II) (ALS-II) (Acetohydroxy-acid synthase II large subunit)	ilvG b4488 b3767	327
P31574	FIXB_ECOLI	Protein FixB	fixB yaaR b0042 JW0041	313
Q46908	YGCR_ECOLI	Putative electron transfer flavoprotein subunit YgcR	ygcR b2770 JW5441	259
Q46911	YGCU_ECOLI	Uncharacterized FAD-linked oxidoreductase YgcU (EC 1.-.-.-)	ygcU ygcT ygcV b4463 JW5442	484
B1X8A6	MQO_ECODH	Probable malate:quinone oxidoreductase (EC 1.1.5.4) (MQO) (Malate dehydrogenase [quinone])	mqo ECDH10B_2367	548

B1X9H2	MTOX_ECODH	N-methyl-L-tryptophan oxidase (MTOX) (EC 1.5.3.-)	solA ECDH10B_1130	372
B1X9W9	MNMG_ECODH	tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG (Glucose-inhibited division protein A)	mnmG gidA ECDH10B_3928	629
B1XA76	DADA_ECODH	D-amino acid dehydrogenase (EC 1.4.99.-)	dadA ECDH10B_1242	432
B1XB17	HCAD_ECODH	3-phenylpropionate/cinnamic acid dioxygenase ferredoxin--NAD(+) reductase component (EC 1.18.1.3)	hcaD ECDH10B_2709	400
B1XBG4	CAIA_ECODH	Crotonobetainyl-CoA reductase (EC 1.3.8.13) (Crotonobetainyl-CoA dehydrogenase)	caiA ECDH10B_0040	380
B1XBG7	FIXB_ECODH	Protein FixB	fixB ECDH10B_0043	313
B1XBJ4	MHPA_ECODH	3-(3-hydroxy-phenyl)propionate/3-hydroxycinnamic acid hydroxylase (3-HCI hydroxylase) (3-HPP hydroxylase) (EC 1.14.13.127)	mhpA ECDH10B_1359	554
B1XBX3	STHA_ECODH	Soluble pyridine nucleotide transhydrogenase (STH) (EC 1.6.1.1) (NAD(P)(+) transhydrogenase [B-specific])	sthA udhA ECDH10B_4151	466
C4ZPW5	CAIA_ECOBW	Crotonobetainyl-CoA reductase (EC 1.3.8.13) (Crotonobetainyl-CoA dehydrogenase)	caiA BWG_0037	380
C4ZPW8	FIXB_ECOBW	Protein FixB	fixB BWG_0040	313
C4ZRZ9	MTOX_ECOBW	N-methyl-L-tryptophan oxidase (MTOX) (EC 1.5.3.-)	solA BWG_0907	372
C4ZTN0	DADA_ECOBW	D-amino acid dehydrogenase (EC 1.4.99.-)	dadA BWG_1014	432
C4ZU53	MQO_ECOBW	Probable malate:quinone oxidoreductase (EC 1.1.5.4) (MQO) (Malate dehydrogenase [quinone])	mqo BWG_1984	548
C4ZXB7	HCAD_ECOBW	3-phenylpropionate/cinnamic acid dioxygenase ferredoxin--NAD(+) reductase component (EC 1.18.1.3)	hcaD BWG_2306	400
C4ZYV6	NORW_ECOBW	Nitric oxide reductase FlRd-NAD(+) reductase (EC 1.18.1.-) (Flavorubredoxin reductase) (FlRd-reductase) (FlavoRb reductase)	norW flrR BWG_2447	377
C4ZZ19	MNMG_ECOBW	tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG (Glucose-inhibited division protein A)	mnmG gidA BWG_3432	629
C5A0R1	STHA_ECOBW	Soluble pyridine nucleotide transhydrogenase (STH) (EC 1.6.1.1) (NAD(P)(+) transhydrogenase [B-specific])	sthA udhA BWG_3630	466
P0A9U8	YDIO_ECOLI	Probable acyl-CoA dehydrogenase YdiO (EC 1.3.-.-)	ydiO b1695 JW5275	383
P68644	FIXC_ECOLI	Protein FixC	fixC yaaS b0043 JW0042	428
P76201	YDIQ_ECOLI	Putative electron transfer flavoprotein subunit YdiQ	ydiQ b1697 JW5276	254
P77337	YDIS_ECOLI	Probable electron transfer flavoprotein-quinone oxidoreductase YdiS (EC 1.5.5.-)	ydiS b1699 JW1689	429
P77378	YDIR_ECOLI	Putative electron transfer flavoprotein subunit YdiR	ydiR b1698 JW1688	312
Q46904	YGCN_ECOLI	Probable electron transfer flavoprotein-quinone oxidoreductase YgcN (EC 1.5.5.-)	ygcN b2766 JW2736	423
Q46907	YGCQ_ECOLI	Putative electron transfer flavoprotein subunit YgcQ	ygcQ b2769 JW5440	286

Table 3.4: Cofactor binding of purified, apo, and reconstituted *E. coli* glutathione reductase (*Ec*GR).

	Enzyme concentration (μM)	Cofactor concentration (μM)	Cofactor binding (%)
Purified *Ec*GR	199.7	195.96	98.13
apo *Ec*GR	86.97	5.34	6.14
*Ec*GR + FAD	93.25	83.44	89.47
*Ec*GR + FCD	98.23	83.91	85.43
*Ec*GR + FUD	82.72	74.9	90.55
*Ec*GR + FGD	Not calculated		
*Ec*GR + dFAD	Not calculated		

References

(1) Manstein, D. J.; Pai, E. F. Purification and Characterization of FAD Synthetase from Brevibacterium Ammoniagenes. *Journal of Biological Chemistry* **1986**, *261* (34), 16169–16173. https://doi.org/10.1016/S0021-9258(18)66693-1.

(2) Barile, M.; Brizio, C.; Valenti, D.; De Virgilio, C.; Passarella, S. The Riboflavin/FAD Cycle in Rat Liver Mitochondria. *Eur J Biochem* **2000**, *267* (15), 4888–4900. https://doi.org/10.1046/j.1432-1327.2000.01552.x.

(3) Merrill, A. H.; McCormick, D. B. Affinity Chromatographic Purification and Properties of Flavokinase (ATP:Riboflavin 5'-Phosphotransferase) from Rat Liver. *Journal of Biological Chemistry* **1980**, *255* (4), 1335–1338. https://doi.org/10.1016/S0021-9258(19)86034-9.

(4) Dym, O.; Eisenberg, D. Sequence-Structure Analysis of FAD-Containing Proteins. *Protein Science* **2001**, *10* (9), 1712–1728. https://doi.org/10.1110/ps.12801.

(5) Denessiouk, K. A.; Rantanen, V.-V.; Johnson, M. S. Adenine Recognition: A Motif Present in ATP-, CoA-, NAD-, NADP-, and FAD-Dependent Proteins. *Proteins: Structure, Function, and Genetics* **2001**, *44* (3), 282–291. https://doi.org/10.1002/prot.1093.

(6) Serrano, A.; Frago, S.; Velázquez-Campoy, A.; Medina, M. Role of Key Residues at the Flavin Mononucleotide (FMN):Adenylyltransferase Catalytic Site of the Bifunctional Riboflavin Kinase/Flavin Adenine Dinucleotide (FAD) Synthetase from Corynebacterium Ammoniagenes. *Int J Mol Sci* **2012**, *13* (12), 14492–14517. https://doi.org/10.3390/ijms131114492.

(7) Marcuello, C.; Arilla-Luna, S.; Medina, M.; Lostao, A. Detection of a Quaternary Organization into Dimer of Trimers of Corynebacterium Ammoniagenes FAD Synthetase at the Single-Molecule Level and at the in Cell Level. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2013**, *1834* (3), 665–676. https://doi.org/10.1016/j.bbapap.2012.12.013.

(8) Macheroux, P.; Kojiro, C. L.; Schopfer, L. M.; Chakraborty, S.; Massey, V. Fluorine-19 NMR Studies on 8-Fluoroflavins and 8-Fluoro Flavoproteins. *Biochemistry* **1990**, *29* (11), 2670–2679. https://doi.org/10.1021/bi00463a008.

(9) ZANETTI, G.; MASSEY, V.; CURTI, B. FAD Analogues as Mechanistic and "Binding-Domain" Probes of Spinach Ferredoxin-NADP+ Reductase. *Eur J Biochem* **1983**, *132* (1), 201–205. https://doi.org/10.1111/j.1432-1033.1983.tb07348.x.

(10) Lee, E. R.; Blount, K. F.; Breaker, R. R. Roseoflavin Is a Natural Antibacterial Compound That Binds to FMN Riboswitches and Regulates Gene Expression. *RNA Biol* **2009**, *6* (2), 187–194. https://doi.org/10.4161/rna.6.2.7727.

(11) Vicens, Q.; Mondragón, E.; Reyes, F. E.; Coish, P.; Aristoff, P.; Berman, J.; Kaur, H.; Kells, K. W.; Wickens, P.; Wilson, J.; Gadwood, R. C.; Schostarez, H. J.; Suto, R. K.; Blount, K. F.; Batey, R. T. Structure–Activity Relationship of Flavin Analogues That Target the Flavin Mononucleotide Riboswitch. *ACS Chem Biol* **2018**, *13* (10), 2908–2919. https://doi.org/10.1021/acschembio.8b00533.

(12) Becker, K.; Christopherson, R. I.; Cowden, W. B.; Hunt, N. H.; Schirmer, R. H. Flavin Analogs with Antimalarial Activity as Glutathione Reductase Inhibitors. *Biochem Pharmacol* **1990**, *39* (1), 59–65. https://doi.org/10.1016/0006-2952(90)90648-5.

(13) Murthy, Y. V. S. N.; Massey, V. Chemical Modification of the N-10 Ribityl Side Chain of Flavins EFFECTS ON PROPERTIES OF FLAVOPROTEIN DISULFIDE OXIDOREDUCTASES. *Journal of Biological Chemistry* **1995**, *270* (48), 28586–28594. https://doi.org/10.1074/jbc.270.48.28586.

(14) TSIBRIS, J. C.; MCCORMICK, D. B.; WRIGHT, L. D. STUDIES ON DONOR-ACCEPTOR COMPLEXES RELATING TO THE INTRAMOLECULAR ASSOCIATION OF THE RIBOFLAVIN AND ADENOSINE MOIETIES OF FLAVIN- ADENINE DINUCLEOTIDE. *Biochemistry* **1965**, *4*, 504–510. https://doi.org/10.1021/bi00879a020.

(15) DeLuca, C.; Kaplan, N. O. LARGE SCALE SYNTHESIS AND PURIFICATION OF FLAVIN ADENINE DINUCLEOTIDE. *Journal of Biological Chemistry* **1956**, *223* (1), 569–576. https://doi.org/10.1016/S0021-9258(18)65165-8.

(16) Huennekens, F. M.; Kilgour, G. L. A NEW CHEMICAL SYNTHESIS OF FLAVIN-ADENINE-DINUCLEOTIDE AND ANALOGS 1,2. *J Am Chem Soc* **1955**, *77* (24), 6716–6717. https://doi.org/10.1021/ja01629a136.

(17) McCormick, D. B.; Chassy, B. M.; Tsibris, J. C. M. Coenzyme Specificity of D-Amino Acid Oxidase for the Adenylate Moiety of FAD. *Biochimica et Biophysica Acta (BBA) - Specialized Section on Enzymological Subjects* **1964**, *89* (3), 447–452. https://doi.org/10.1016/0926-6569(64)90070-7.

(18) KOBERSTEIN, R. 8-Azidoadenine Analogs of NAD+ and FAD. Synthesis and Coenzyme Properties with NAD+ -Dependent and FAD-Dependent Enzymes. *Eur J Biochem* **1976**, *67* (1), 223–229. https://doi.org/10.1111/j.1432-1033.1976.tb10653.x.

(19) Stocker, A.; Hecht, H.-J.; Buckmann, A. F. Synthesis, Characterization and Preliminary Crystallographic Data of N6-(6-Carbamoylhexyl)-FAD-d-Amino-Acid Oxidase from Pig Kidney, a Semi-Synthetic Oxidase. *Eur J Biochem* **1996**, *238* (2), 519–528. https://doi.org/10.1111/j.1432-1033.1996.0519z.x.

(20) Christie, S. M. H.; Kenner, G. W.; Todd, A. R. Nucleotides. Part XXV. A Synthesis of Flavin–Adenine Dinucleotide. *J. Chem. Soc.* **1954**, 46–52. https://doi.org/10.1039/JR9540000046.

(21) Shuster, L.; Kaplan, N. O.; Stolzenbach, F. E. THE REACTION OF PYRIDINE NUCLEOTIDES WITH TRIFLUOROACETIC ACID ANHYDRIDE. *Journal of Biological Chemistry* **1955**, *215* (1), 195–209. https://doi.org/10.1016/S0021-9258(18)66028-4.

(22) Mashhadi, Z.; Xu, H.; Grochowski, L. L.; White, R. H. Archaeal RibL: A New FAD Synthetase That Is Air Sensitive. *Biochemistry* **2010**, *49* (40), 8748–8755. https://doi.org/10.1021/bi100817q.

(23) Craven, D. B.; Harvey, M. J.; Dean, P. D. G. The Synthesis of N^6 -(6-Aminohexyl)-NAD^+ and Its Application to Affinity Chromatography. *FEBS Lett* **1974**, *38* (3), 320–324. https://doi.org/10.1016/0014-5793(74)80082-7.

(24) Moffatt, J. G.; Khorana, H. G. Nucleoside Polyphosphates. VIII. 1 New and Improved Syntheses of Uridine Diphosphate Glucose and Flavin Adenine Dinucleotide Using Nucleoside-5' Phosphoramidates 2. *J Am Chem Soc* **1958**, *80* (14), 3756–3761. https://doi.org/10.1021/ja01547a073.

(25) McCormick, D. B.; Opar, G. E. Syntheses of 8-Bromo-5'-Adenylate-Containing Nucleotides. *J Med Chem* **1969**, *12* (2), 333–334. https://doi.org/10.1021/jm00302a037.

(26) Michelson, A. M. Synthesis of Nucleotide Anhydrides by Anion Exchange. *Biochimica et Biophysica Acta (BBA) - Specialized Section on Nucleic Acids and Related Subjects* **1964**, *91* (1), 1–13. https://doi.org/10.1016/0926-6550(64)90164-1.

(27) Smit, P.; Stork, G. A.; van der Plas, H. C.; den Hartog, J. A. J.; van der Marel, G. A.; van Boom, J. H. Synthesis of 5-Deazaflavine Adenine Dinucleotide (5-DFAD) by a Modified Triester Approach. *Recueil des Travaux Chimiques des Pays-Bas* **1984**, *103* (3), 101–103. https://doi.org/10.1002/recl.19841030307.

(28) Hata, T.; Nakagawa, I. Synthesis of Nucleotide Coenzymes via Nucleoside 5'-Phosphorothioate Intermediates. *J Am Chem Soc* **1970**, *92* (18), 5516–5516. https://doi.org/10.1021/ja00721a037.

(29) Ji, D.; Wang, L.; Hou, S.; Liu, W.; Wang, J.; Wang, Q.; Zhao, Z. K. Creation of Bioorthogonal Redox Systems Depending on Nicotinamide Flucytosine Dinucleotide. *J Am Chem Soc* **2011**, *133* (51), 20857–20862. https://doi.org/10.1021/ja2074032.

(30) Spencer, R.; Fisher, J.; Walsh, C. Preparation, Characterization, and Chemical Properties of the Flavin Coenzyme Analogues 5-Deazariboflavin, 5-Deazariboflavin 5'-Phosphate, and 5-Deazariboflavin 5'-Diphosphate, 5' → 5'-Adenosine Ester. *Biochemistry* **1976**, *15* (5), 1043–1053. https://doi.org/10.1021/bi00650a015.

(31) Mishanina, T. V.; Kohen, A. Synthesis and Application of Isotopically Labeled Flavin Nucleotides. *J Labelled Comp Radiopharm* **2015**, *58* (9), 370–375. https://doi.org/10.1002/jlcr.3313.

(32) Kuppuraj, G.; Kruise, D.; Yura, K. Conformational Behavior of Flavin Adenine Dinucleotide: Conserved Stereochemistry in Bound and Free States. *J Phys Chem B* **2014**, *118* (47), 13486–13497. https://doi.org/10.1021/jp507629n.

(33) Guo, X.; Liu, Y.; Wang, Q.; Wang, X.; Li, Q.; Liu, W.; Zhao, Z. K. Non-natural Cofactor and Formate-Driven Reductive Carboxylation of Pyruvate. *Angewandte Chemie* **2020**, *132* (8), 3167–3170. https://doi.org/10.1002/ange.201915303.

(34) Liu, Y.; Li, Q.; Wang, L.; Guo, X.; Wang, J.; Wang, Q.; Zhao, Z. K. Engineering <scp>d</Scp> -Lactate Dehydrogenase to Favor an Non-natural Cofactor Nicotinamide Cytosine Dinucleotide. *ChemBioChem* **2020**, *21* (14), 1972–1975. https://doi.org/10.1002/cbic.201900766.

(35) SCOLA-NAGELSCHNEIDER, G.; HEMMERICH, P. Synthesis, Separation, Identification and Interconversion of Riboflavin Phosphates and Their Acetyl Derivatives: A Reinvestigation. *Eur J Biochem* **1976**, *66* (3), 567–577. https://doi.org/10.1111/j.1432-1033.1976.tb10583.x.

(36) Forrest, H. S.; Mason, H. S.; Todd, A. R. 479. Nucleotides. Part XI. Some Preliminary Experiments on a Projected Synthesis of Flavin-Adenine Dinucleotide. New Methods for the Preparation of Riboflavin-4? : 5? Phosphate. *Journal of the Chemical Society (Resumed)* **1952**, 2530. https://doi.org/10.1039/jr9520002530.

(37) van den Ent, F.; Löwe, J. RF Cloning: A Restriction-Free Method for Inserting Target Genes into Plasmids. *J Biochem Biophys Methods* **2006**, *67* (1), 67–74. https://doi.org/10.1016/j.jbbm.2005.12.008.

(38) Zanetti, G.; Beretta, C.; Malandra, D. Properties of Rabbit Liver Glutathione Reductase Reconstituted with FAD Analogs. *Arch Biochem Biophys* **1986**, *244* (2), 831–837. https://doi.org/10.1016/0003-9861(86)90652-1.

(39) Pai, E. F.; Schulz, G. E. The Catalytic Mechanism of Glutathione Reductase as Derived from X-Ray Diffraction Analyses of Reaction Intermediates. *Journal of Biological Chemistry* **1983**, *258* (3), 1752–1757. https://doi.org/10.1016/S0021-9258(18)33050-3.

(40) Mannervik, B. Measurement of Glutathione Reductase Activity. *Curr Protoc Toxicol* **1999**, *00* (1). https://doi.org/10.1002/0471140856.tx0702s00.

(41) Maiorov, V. N.; Crippen, G. M. Significance of Root-Mean-Square Deviation in Comparing Three-Dimensional Structures of Globular Proteins. *J Mol Biol* **1994**, *235* (2), 625–634. https://doi.org/10.1006/jmbi.1994.1017.

(42) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res* **2000**, *28* (1), 235–242. https://doi.org/10.1093/nar/28.1.235.

(43) 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.

(44) Bogachev, V. S. Synthesis of Deoxynucleoside 5'-Triphosphates Using Trifluoroacetic Anhydride as an Activating Reagent. *Bioorganicheskaya Khimiya* **1996**, *22* (9), 699–705.

(45) Marlow, A. L.; Kiessling, L. L. Improved Chemical Synthesis of UDP-Galactofuranose. *Org Lett* **2001**, *3* (16), 2517–2519. https://doi.org/10.1021/ol016170d.

(46) Vincent, S. P.; Gastinel, L. N. An Efficient Synthesis of GDP-Hexanolamine, a Key Tool for the Purification of Fucosyltransferases. *Carbohydr Res* **2002**, *337* (11), 1039–1042. https://doi.org/10.1016/S0008-6215(02)00105-2.

(47) Mohamady, S.; Jakeman, D. L. An Improved Method for the Synthesis of Nucleoside Triphosphate Analogues. *J Org Chem* **2005**, *70* (25), 10588–10591. https://doi.org/10.1021/jo0518598.

(48) Timmons, S. C.; Jakeman, D. L. Stereospecific Synthesis of Sugar-1-Phosphates and Their Conversion to Sugar Nucleotides. *Carbohydr Res* **2008**, *343* (5), 865–874. https://doi.org/10.1016/j.carres.2008.01.046.

(49) Massey, V.; Ghisla, S.; Moore, E. G. 8-Mercaptoflavins as Active Site Probes of Flavoenzymes. *Journal of Biological Chemistry* **1979**, *254* (19), 9640–9650. https://doi.org/10.1016/S0021-9258(19)83564-0.

(50) KRAUTH-SIEGEL, R. L.; SCHIRMER, R. H.; GHISLA, S. FAD Analogues as Prosthetic Groups of Human Glutathione Reductase. Properties of the Modified Enzyme Species and Comparisons with the Active Site Structure. *Eur J Biochem* **1985**, *148* (2), 335–344. https://doi.org/10.1111/j.1432-1033.1985.tb08844.x.

(51) 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.

(52) Pergolizzi, G.; Butt, J. N.; Bowater, R. P.; Wagner, G. K. A Novel Fluorescent Probe for NAD-Consuming Enzymes. *Chemical Communications* **2011**, *47* (47), 12655. https://doi.org/10.1039/c1cc15499k.

(53) Feldmann, J.; Li, Y.; Tor, Y. Emissive Synthetic Cofactors: A Highly Responsive NAD + Analogue Reveals Biomolecular Recognition Features. *Chemistry – A European Journal* **2019**, *25* (17), 4379–4389. https://doi.org/10.1002/chem.201805520.

(54) Depaix, A.; Kowalska, J. NAD Analogs in Aid of Chemical Biology and Medicinal Chemistry. *Molecules* **2019**, *24* (22), 4187. https://doi.org/10.3390/molecules24224187.

(55) Pljevaljcic, G.; Pignot, M.; Weinhold, E. Design of a New Fluorescent Cofactor for DNA Methyltransferases and Sequence-Specific Labeling of DNA. *J Am Chem Soc* **2003**, *125* (12), 3486–3492. https://doi.org/10.1021/ja021106s.

(56) 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.

(57) Zhang, L.; Lin, H. Using Clickable NAD+ Analogs to Label Substrate Proteins of PARPs. In *Methods in Molecular Biology*; 2017; Vol. 1608, pp 95–109. https://doi.org/10.1007/978-1-4939-6993-7_8.

(58) Anmangandla, A.; Ren, Y.; Fu, Q.; Zhang, S.; Lin, H. The Acyl-CoA Specificity of Human Lysine Acetyltransferase KAT2A. *Biochemistry* **2022**, *61* (17), 1874–1882. https://doi.org/10.1021/acs.biochem.2c00308.

(59) 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). https://doi.org/10.1002/cbic.202200511.

(60) Bateman, A.; Martin, M.-J.; Orchard, S.; Magrane, M.; Agivetova, R.; Ahmad, S.; Alpi, E.; Bowler-Barnett, E. H.; Britto, R.; Bursteinas, B.; Bye-A-Jee, H.; Coetzee, R.;

Cukura, A.; Da Silva, A.; Denny, P.; Dogan, T.; Ebenezer, T.; Fan, J.; Castro, L. G.; Garmiri, P.; Georghiou, G.; Gonzales, L.; Hatton-Ellis, E.; Hussein, A.; Ignatchenko, A.; Insana, G.; Ishtiaq, R.; Jokinen, P.; Joshi, V.; Jyothi, D.; Lock, A.; Lopez, R.; Luciani, A.; Luo, J.; Lussi, Y.; MacDougall, A.; Madeira, F.; Mahmoudy, M.; Menchi, M.; Mishra, A.; Moulang, K.; Nightingale, A.; Oliveira, C. S.; Pundir, S.; Qi, G.; Raj, S.; Rice, D.; Lopez, M. R.; Saidi, R.; Sampson, J.; Sawford, T.; Speretta, E.; Turner, E.; Tyagi, N.; Vasudev, P.; Volynkin, V.; Warner, K.; Watkins, X.; Zaru, R.; Zellner, H.; Bridge, A.; Poux, S.; Redaschi, N.; Aimo, L.; Argoud-Puy, G.; Auchincloss, A.; Axelsen, K.; Bansal, P.; Baratin, D.; Blatter, M.-C.; Bolleman, J.; Boutet, E.; Breuza, L.; Casals-Casas, C.; de Castro, E.; Echioukh, K. C.; Coudert, E.; Cuche, B.; Doche, M.; Dornevil, D.; Estreicher, A.; Famiglietti, M. L.; Feuermann, M.; Gasteiger, E.; Gehant, S.; Gerritsen, V.; Gos, A.; Gruaz-Gumowski, N.; Hinz, U.; Hulo, C.; Hyka-Nouspikel, N.; Jungo, F.; Keller, G.; Kerhornou, A.; Lara, V.; Le Mercier, P.; Lieberherr, D.; Lombardot, T.; Martin, X.; Masson, P.; Morgat, A.; Neto, T. B.; Paesano, S.; Pedruzzi, I.; Pilbout, S.; Pourcel, L.; Pozzato, M.; Pruess, M.; Rivoire, C.; Sigrist, C.; Sonesson, K.; Stutz, A.; Sundaram, S.; Tognolli, M.; Verbregue, L.; Wu, C. H.; Arighi, C. N.; Arminski, L.; Chen, C.; Chen, Y.; Garavelli, J. S.; Huang, H.; Laiho, K.; McGarvey, P.; Natale, D. A.; Ross, K.; Vinayaka, C. R.; Wang, Q.; Wang, Y.; Yeh, L.-S.; Zhang, J.; Ruch, P.; Teodoro, D. UniProt: The Universal Protein Knowledgebase in 2021. *Nucleic Acids Res* **2021**, *49* (D1), D480–D489. https://doi.org/10.1093/nar/gkaa1100.

Chapter 4

An enzyme engineering approach to the synthesis of FAD nucleobase analogues inside *Escherichia coli*

4.1 Introduction

Adenosine triphosphate (ATP) has been an energy currency for all cells and plays various roles in cell metabolism besides being a constituent of RNAs. Other than these roles, ATP provides an adenosine recognition handle to essential cofactors like coenzyme A, NAD, FAD, and SAM.[1] This adenosine handle takes no part in the catalytic activity except anchoring and binding to the enzyme when incorporated with the cofactors. Much research has been done on creating and using the analogues of these cofactors modified on or near the catalytic centre. However, recently, researchers have started focusing on the adenine moiety to create effective analogues of these cofactors, which can be utilized as chemical biology tools, structural probes, antimicrobials, enzymatic inhibitors, and fluorescence imaging. For example, enzymatically synthesized stereospecific SAM fluorescent nucleobase analogues (DAPSAM, 2-AP-SAM, formycinylmethionine, and 8-aza-SAM) were found to be biologically active with EcoRI methyl transferase and fluorescence properties of DAPSAM were utilized to determine its K_d for binding with the SAM-III riboswitch RNA.[2] A series of adenosine-modified NAD analogues like 2′-deoxy-2′-fluoroarabinosyl-β-nicotinamide adenine dinucleotide (ara-F NAD) were found to be potent inhibitors of NADase activity of a signalling enzyme CD38, and the ara-F NGD was found to be the most potent than others.[3]

The adenylyl transferase enzymes for the biosynthesis of these adenosine-based cofactors are ATP-specific that no other NTPs (like GTP, CTP, or UTP) dependent FAD nucleobase analogue has been known to be present in nature in our knowledge, while the substrates to synthesize them are already available in the cell. However, recently, the marine bacterium *Desulfobacula toluolica* Tol2^T has shown the natural presence of adenosino-CoA analogue, inosino-CoA, along with the hypoxanthine-based coenzyme analogues (NHD and FHD) of NAD and FAD.[4] Besides this, Inosino-CoA was found to be incorporated with many acyl groups as inosino-CoA thioesters, suggesting that the nucleobase analogues can not only be produced in the cell but also can be utilized by the cell. Hence, the *in vivo* synthesis of FAD nucleobase analogues should be possible by utilizing the existing nucleoside triphosphate (NTP) and FMN as a substrate. The biosynthesis of these cofactor nucleobase analogues opens

the field to use them to create biorthogonal systems in the cell for chemical and synthetic biology applications.[5] The nicotinamide cytosine dinucleotide (NCD), an NAD analogue, was synthesized to be used specifically by engineered malic enzyme mutants for L-malate to pyruvate conversion by oxidative decarboxylation.[6] Later, several NAD-utilising enzymes, such as formate (FDH), phosphite, and D-lactate dehydrogenases, were engineered to use NCD over NAD selectively.[7–10] NCD was used to create a biorthogonal system for an atom-economic strategy of CO_2 fixation where nucleotide transporters NDT2 and NTT4 from *Arabidopsis thaliana* and *Protochlamydia amoebophila* UWE25 are used to transport NCD within the cell.[8,11]. In *E. coli* cells, NCD/NCDH cycles keep going as the FDH mutant converts formate to CO_2 using NCD, converted to NCDH, and The ME mutant uses this liberated CO_2 and NCDH as a substrate for the carboxylation of pyruvate to malate further generating NCD. However, nucleotide transporter caused a reduction in cell growth, and NCD transportation is expensive to grow cell cultures. Hence, NadD, the NAD synthetase, was engineered to get a mutant with strong CTP preference over ATP, and the mutant also directly incorporates nicotinamide mononucleotide (NMN) in the active site instead of its native substrate.[12,13] Overexpression of this mutant alone resulted in the production of NCD up to the concentration of 5 mM, which can directly be used by the NCD-based enzymatic cascades for bioorthogonal applications.[13] Similarly, FAD nucleobase analogues (FGD, FCD, FUD, dFAD) have been found to be active with glutathione reductase, and these analogues can be potential bio-orthogonal cofactors, just like NCD.[14]

FAD is produced from the riboflavin via an enzymatic two-step reaction consuming one ATP molecule in each reaction (Figure 3.1A). First, the phosphorylation of riboflavin takes place by riboflavin kinase (RFK), resulting in flavin mononucleotide (FMN) using ATP, which is adenylated to FAD by another enzyme FMN adenylyl transferase (FMNAT) using another ATP molecule (Scheme 1).[15–17] Generally, in prokaryotes, a bifunctional FAD synthetase enzyme does both these steps; for example, *Corynebacterium ammoniagenes* FAD synthetase (*Ca*FADS).[15] The FAD is an essential molecule contributing to various important cellular activities and neither single gene knock-out nor a mutant of FAD biosynthesis genes has been reported. The Keio collection study of *E. coli* single gene knockouts reported the *ribF* gene as an essential gene which expresses a bifunctional FAD synthetase responsible for FMN and FAD biosynthesis.[18] Generally, Riboflavin either synthesised or transported in the cell which is converted to FMN and FAD.[19,20] However, no transporter has been reported for FAD. The only approach to synthesize FAD analogues in the cell is to engineer a FAD synthetase enzyme

and either doing heterologous expression of the mutant in the cell through a plasmid or homologous expression by substituting the native chromosomal *ribF* gene with the mutant gene.

*Ca*FADS is an extensively explored enzyme representing the prokaryotes' bifunctional FADS family. The crystal structure of *Ca*FADS and its mutants are also available in the PDB database and are well-studied.[21–26] It possesses two modules, RFK (C-terminal) and FMNAT (N-terminal). *Ca*FADS has also been used as a green chemistry catalyst to synthesize many isoalloxazine-based FAD analogues.[27,28] However, being very specific for ATP, it failed to produce any FAD nucleobase analogue.[15] A FAD synthetase from a thermophilic methanogen *Methanocaldococcus jannaschii* (*Mj*FMNAT) can synthesize these analogues *in vitro* and within the cell.[14,29] However, its heterologous expression slows the growth of the cell. Being a thermophilic enzyme, it cannot replace the native gene in the bacterial genome for producing these nucleobase analogues of FAD. Hence, for biosynthesis of FAD nucleobase analogues, modifying a native bifunctional FAD synthetase to exhibit promiscuity towards different NTPs is a more reliable choice. Inside a cell, these analogues could potentially explore the role of FAD in important cellular processes and as bio-orthogonal cofactors for synthetic biology use similar to NCD.[6,8,10] For this purpose, we chose to work with *E. coli* FAD synthetase (*Ec*FADS) as it is a homolog of *Ca*FADS, and *E. coli* genetics are well-studied to manipulate it genetically and study its metabolism. However, *Ec*FADS has been studied in very few instances. For example, a method was patented in 1996 for producing flavin nucleotides using *Ec*FADS, where G23 residue was mutated with arginine, alanine, or aspartic acid to terminate the FMNAT domain activity in the recombinant DNA, and then only FMN was produced by microorganism cultures containing the recombinant DNA.[30] Another study of *Ec*FADS has been done where it is active with the antibiotic roseoflavin (an analogue of riboflavin.[31] It produces a roseoflavin version of FMN and FAD, known as roFMN and roFAD, in the cell by feeding roseoflavin exogenously. These cofactors were found to bind with 37 out of 38 overexpressed flavoproteins in the *E. coli* strain.[31] These studies suggest that the enzyme engineering of *Ec*FADS is possible for the biosynthesis of these FAD nucleobase analogues inside the cell.

This Chapter shows the optimization of *Ec*FADS activity, where we found that the RKF module of *Ec*FADS is mainly specific for ATP, while the FMNAT module is promiscuous for other NTPs except for GTP. Using computational study and structure homology, we have engineered the FMNAT module of *Ec*FADS to make it more promiscuous towards other NTPs.

We swapped the *Ec*FADS loop surrounding the adenine base in the FMNAT domain with a cytidylyl transferase loop, making it promiscuous for GTP. Heterologous expression of this LoopSwap mutant produces FAD analogues inside the cell. Other than this, by shortening the loop, we have shown that this loop is mainly the reason for ATP specificity inside the cell.

4.2 Materials and Methods

4.2.1 Materials

Jena Biosciences supplied the order of 2'-deoxy adenosine triphosphate (dATP) and all nucleoside triphosphates (NTPs). Reagents and chemicals for the enzymatic assays were purchased from RANKEM Chemicals, TCI Chemicals, and Merck. All reagents and components of media used in bacterial cultures were ordered from TCI Chemicals, HiMedia, and SRL Chemicals. DSS Takara Bio USA supplied Taq polymerase, Dpn1, and the PrimeSTAR GXL DNA polymerase. Purification kits for DNA and plasmid were obtained from Agilent and Qiagen. We considered Merck, SD Fine Chemicals, and RANKEM for all solvents ordering to use in LC-MS and HPLC analysis.

4.2.2 Cloning and purification of *Ec*FADS and mutants of the FMNAT module

The cloning of the *ribF* gene for the *Escherichia coli* FAD synthetase (*Ec*FADS) was done in a pET28a(+) plasmid between BamHI and NdeI restriction sites using restriction-free cloning.[32] The first round of PCR was done for amplification of the *ribF* gene from *E. coli* genomic DNA using the first set of primers (EcFADS_F1 and EcFADS_R1; Table 4.1). The second round of PCR adds flanking sequences at terminals of the gene that are complementary to the insertion sites of the pET28a vector. This step was done using the second set of primers (*Ec*FADS_F2 and *Ec*FADS_R2; Table 4.1). The gene obtained in the second round of PCR is called megaprimer, which was used to amplify the whole plasmid vector with the gene inserted between the restriction sites mentioned above. The parent plasmid was digested using Dpn1, and chemical-competent *E. coli* DH5α cells were transformed with the amplified nicked plasmid, followed by the isolation of the cloned vector. For the cloning of mutants, the wild-type *ribF*-plasmid vector was employed as a template, and the list of the corresponding primers is given in Table 4.1. These primers contain the altered codon at the site of mutation. The mutants were cloned by PCR in two rounds. The first round of PCR was done with the

corresponding forward primer of the mutant and the T7 terminator universal primer (Table 4.1). It amplified the fragment of plasmids from the T7 terminator site to the mutation site. This fragment, called megaprimer, was utilized in the other round of PCR to amplify the whole plasmid introduced with mutation, which was transformed into DH5α chemical competent cells followed by plasmid isolation.

Table 4.1: The primer List used for the cloning

#	Primer name	Sequence
1	*Ec*FADS_F1	5'-ATGAAGCTGATACGCGGCATAC-3'
2	*Ec*FADS_R1	5'-TTAAGCCGGTTTTGTTAGCCC-3'
3	*Ec*FADS_F2	5'-GTGCCGCGCGGCAGCCATATGAAGCTGATACGCGGCATAC-3'
4	*Ec*FADS_R2	5'-CGACGGAGCTCGAATTCGGATCCTTAAGCCGGTTTTGTTAGCCC-3'
5	G23K_F	5'-GCTGACTATTAAAAATTTCGACG-3'
6	S165K_F	5'-GCGCATCAAAAGCACCGC-3'
7	G23S_F	5'-GCTGACTATTAGCAATTTCGACG-3'
8	ΔGVR_F	5'-ACGCAAACTTTTTGCGAAGGTATCAGCAGCACCGCC-3'
9	ΔEGVR_F	5'-CCAGTACGCAAACTTTTTGCGGTATCAGCAGCACCGCC-3'
10	ΔFEGVR_F	5'-TCACCAGTACGCAAACTTGCGGCATCAGCAGCACCG-3'
11	ΔFCEGGVR_F	5'-GATATCACCAGTACGCAAACTATCAGCAGCACCGCCGTGC-3'
12	LoopSwap_F	5'-GGCTTCGATATCACCAGTACGCCGCGCACCGAAGGCATCAGCAGCAC-3'
14	T7 Terminator_R	5'-GCTAGTTATTGCTCAGCGG-3'

The *ribF* or the mutant gene containing isolated plasmids were introduced to the BL21(DE3) competent cells. Overnight grown primary culture was utilized to grow the secondary culture in autoclaved Luria Bertani (LB) broth with kanamycin (35 µg/mL), followed by the growth at 37 °C and 180 rpm. Protein overexpression was initiated at 28 °C by adding 0.1 M isopropyl-β-D-1-thiogalactopyranoside (IPTG) once the culture's OD_{600} reached a range of 0.6 to 0.8. The culture was then grown for 8 hours, following which the cellular components were separated into pellets through centrifugation and subsequently stored at a temperature of -80 °C. The Ni-NTA affinity chromatography was used to purify the N-terminally hexa-histidine-tagged protein. This method utilized the Akta pure FPLC system from GE Healthcare. The initial buffer used for equilibration contained Tris-HCl (50 mM) at pH 8.0, β-mercaptoethanol (0.025%), imidazole (20 mM), and NaCl (300 mM). Elution of the protein was accomplished using a buffer consisting of Tris-HCl (50 mM) at pH 8.0, β-mercaptoethanol (0.025%), imidazole (500 mM), and NaCl (300 mM). To determine protein

concentration, the Bradford assay was performed with standard Bovine serum albumin (BSA). An identical procedure was used for all *Ec*FADS variants. 12% SDS-PAGE gel electrophoresis was used to assess the enzymes' molecular weight and purity.

4.2.3 Activity assay of *Ec*FADS and its mutants

The activity of *Ec*FADS was optimized using Tris-HCl (100 mM, pH 8.0) buffer, riboflavin (200 μM), ATP (3.5 mM), and enzyme (5 μM) in final concentration with or without 20 mM DTT (dithiothreitol) and varying concentrations of $MgCl_2$ (0-10 mM). For different divalent metal ions, 3 mM concentration of M^{2+} ion was used instead of $MgCl_2$ in the presence of 20 mM DTT while keeping all other components the same. The time optimization was done with 3 mM $MgCl_2$ for several time points in the presence of DTT.

The final assay for the activity of *Ec*FADS and the mutants was as follows: Tris-HCl (100 mM, pH 8.0), riboflavin or FMN (200 μM), ATP (3.5 mM), DTT (20 mM), $MgCl_2$ (3 mM) and 5 μM of the enzyme in final concentration. The activity with other NTPs (UTP, CTP, and GTP) and deoxy nucleotides (dATP) was checked similarly at 3.5 mM concentration. All the reactions were performed for 6 h at 37 °C. TLC was used to visualize the fluorescent flavins excited at 365 nm. The mixture of water, acetic acid, and n-butanol was used in the ratio 5:3:12 as a solvent system.

For *Ec*FADS activity at the cellular concentration of NTPs, 1.5 mM GTP, 0.8 mM UTP, 0.5 mM CTP, or 0.2 mM dATP was used, keeping the other components the same. Aliquots at various time points were taken and observed on TLC.

4.2.4 Fluorescent analysis of flavins with DTT

The fluorescence of flavins was observed on FluoroMax-4 with excitation at 450 nm followed by emission at 525 nm. The fluorescence of 40 μM RF, FMN, and FAD was recorded first, and then varying concentrations of DTT (50, 100, and 200 mM) were added, followed by immediate fluorescence recording. 200 mM β-mercaptoethanol and 20 mM sodium dithionite were used as a negative and positive control, respectively.

4.2.5 Heterologous expression and lysate study by LC-MS

50 mL LB broth was inoculated with the 1% overnight grown BL21(DE3) strain transformed with empty pET28a, *Ec*FADS-pET28a, or *mutant*-pET28a constructs and the culture was grown with kanamycin (35 µg/mL) at 37 °C followed by continuous shaking on 180 rpm. Only the *E. coli* BL21(DE3) culture was grown without Kan. We did not induce the culture with IPTG and instead relied on the leaky expression of the pET28a plasmid to minimally perturb the cellular metabolism. After 14-16 h, the culture was centrifuged at 5500 rpm for 15 min at 4 °C resulting in a pellet. 2 mL of 90 % methanol was used for the resuspension of the pellet followed by sonication at 100 % amplitude for 3 min with 1s pulse on and 1s pulse off. The final supernatant was obtained by centrifuging the lysate at 4 °C with 5500 rpm for 15 min, followed by the filtration via 0.22-micron filters. The supernatant was kept at a high vacuum to concentrate it directly for analysis on LC-MS.

4.2.6 Computational analysis

For sequence alignment, a web-based tool, T-Coffee, was used.[33] All structure visualization and structure alignment were done with PyMOL. The homology modelling was done using a web-based tool, Phyre2 and a software called Modeller.[34,35] We used the DALI server to search the structural homologs of *Ca*FADS.[36] PDB90 search was employed to eliminate redundant PDBs with sequence similarity above 90%, and a list is provided shorted by z-score till 10 (Table 4.2).

Collaborating with Reman Singh and Arnab Mukherjee from IISER Pune, we conducted docking and MD simulations for *Ec*FADS and its predicted mutants. Utilizing Modeller, we constructed models for both wild-type EcFADS and its mutants.[35] The Autodock 4 software facilitated the docking of ligands such as FMN, ATP, GTP, and CTP.[37,38] First, our docking process involved ATP, GTP, and CTP with the *Ec*FADS structure, followed by subsequent molecular dynamics (MD) simulation. In this endeavour, the AMBER99SB force field was applied to the protein.[39] To generate the AMBER force field parameters for ATP, CTP, GTP, and FMN, we employed the antechamber module of AmberTool. We began with a quantum mechanical optimization of ATP, GTP, and CTP using the HF/6-31G* basis set in GAUSSIAN03 software.[40] Subsequently, AmberTool was utilized to compute restricted electrostatic potential (RESP) charges for the atoms within ATP, GTP, CTP, and FMN. The resulting topology and coordinates, produced by AmberTool, underwent conversion into GROMACS format using the amb2gmx.pl Perl program. Our molecular dynamics simulations

were executed using the GROMACS package (version 4.5).[41] We encapsulated the NTP-FMN bound protein within a cubic box measuring 70 Å in length. This system was then solvated with TIP3P water molecules. We introduced a solution of 150 mM MgCl2 ions and optimized the system using the steepest descent method. The subsequent steps involved heating the system to 300 K using the Berendsen thermostat with a coupling constant of 0.2 ps under NVT conditions. After this, a 10 ns NPT simulation was performed at 300 K and 1 bar pressure, adjusting the box size by employing the Berendsen thermostat and barostat with a coupling constant of 0.4 ps. Finally, a production run lasting 100 ns at 300 K was conducted using the Nose-Hoover thermostat. All subsequent analysis was carried out based on the trajectory from the production run.

4.2.7 TLC, HPLC, and LC-MS

Silica gel TLC was employed to analyze all reactions, and the outcomes were detected under ultraviolet light with a wavelength of 365 nm. The mixture of water, acetic acid, and n-butanol in a ratio of 5:3:12 was employed as the mobile phase.

HPLC (Agilent 1260 Infinity-II series) linked with a UV-Vis diode array detector was utilized to assess enzymatic reactions. The analysis used a Phenomenex Gemini reversed-phase column of 250 x 4.6 mm dimensions with a 5 µm particle size of NX-C18. The solvent-A was ammonium acetate buffer (10 mM, pH 6.0), and methanol was used as solvent-B. A method spanning 60 minutes was applied, starting with 5% of solvent B at 0 minutes, maintaining this ratio until 15 minutes, followed by a gradient increase to 80% of solvent B until 40 minutes, then a return to 5% solvent B until 55 minutes, and finally, maintaining at 5% solvent B until 60 minutes. Another method spanning 30 minutes was used, starting with 5% of solvent B at 0 minutes, maintaining this ratio until 2.5 minutes, followed by a gradient increase to 80% of solvent B until 20 minutes, then a return to 5% solvent B until 27 minutes, and finally, maintaining at 5% solvent B until 30 minutes.

A Mass spectrometer (ScieX X500R-QTOF) was employed for LC-MS analyses, where a UHPLC (Exion-LC series) was used to separate the flavins. The flavins were detected in negative mode at a temperature of 400 °C. The parameters are as follows: ion spray voltage was -4500 V, the de-clustering potential was -80V and collision energy was 5V. The ScieX-OS Explorer software facilitated data acquisition, processing, and the generation of metrics such

as extracted ion chromatogram (EIC), total ion chromatogram (TIC), and the m/z values of targeted molecules. The same column and method as used in HPLC were utilized.

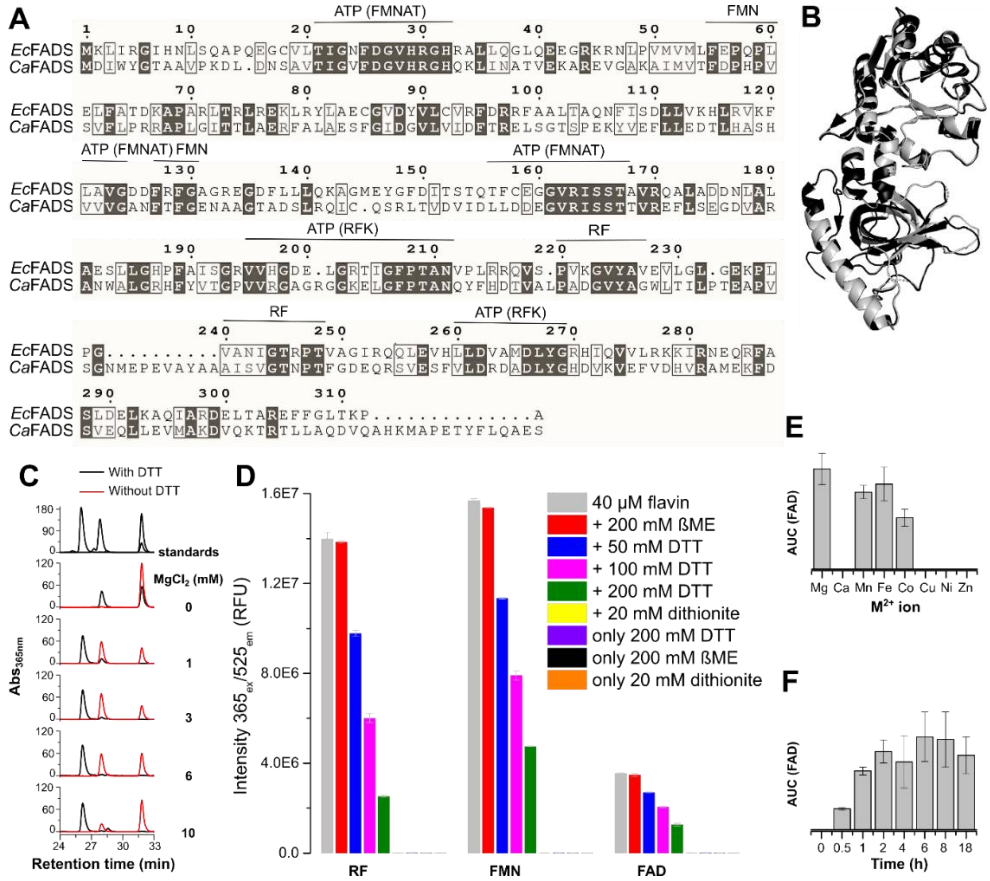


Figure 4.1 Homology modelling of *Ec*FADS and its activity optimization. (A) Sequence alignment of *Ec*FADS with *Ca*FADS shows 31 % sequence similarity and 50 % homology. Conserved residues are shown in black rectangles. **(B)** The alignment of modelled *Ec*FADS (grey) with the crystal structure of *Ca*FADS (black) shows complete overlap. **(C)** HPLC chromatogram of *Ec*FADS activity with RF and ATP in the presence of various $MgCl_2$ concentrations with DTT (black) and without DTT (red). Without DTT, only FMN is formed, while with DTT, both FMN and FAD are produced. 3 mM $MgCl_2$ is sufficient to complete the reaction with DTT, while without DTT till 3 mM $MgCl_2$, FMN formation increases, and after that, the *Ec*FADS activity decreases at higher concentrations of $MgCl_2$. **(D)** Fluorescence analysis of flavins to verify that DTT provides reduced flavins, the active substrate for *Ec*FADS. The decrease in fluorescence of flavins in the presence of DTT is due to the reduction of flavins. β-mercaptoethanol was used as a negative control as it does not reduce flavins, and sodium dithionite was used as a positive control, which makes the fluorescence of flavins vanish completely. **(E)** Area under the curve (AUC) for FAD peak extracted from HPLC chromatograms of *Ec*FADS activity with ATP, RF, and various M^{2+} metal ions. Only Mg, Mn, Fe, and Co metal ions are accepted by the enzyme in which the Mg^{2+} ion is the best for the activity. **(F)** Area under the curve (AUC) for FAD peak extracted from HPLC chromatograms of *Ec*FADS activity with ATP and RF showing time optimization for the reaction assay.

4.3 Results and Discussion

4.3.1 *Ec*FADS is homologous to *Ca*FADS

We did a sequence alignment of *Ec*FADS with *Ca*FADS using T-Coffee and found that *Ec*FADS has 31 % sequence similarity and 50 % homology to *Ca*FADS. Also, the active sites for RF, FMN, and both ATPs are well conserved, as marked in Figure 4.1A. Since no crystal structure is available for *Ec*FADS, we did homology modelling using the web-based tool phyre2 and software modeller. On structural alignment, *Ec*FADS completely overlaps with *Ca*FADS (Figure 4.1B). Like *Ca*FADS, *Ec*FADS possesses two modules for performing the phosphorylation and adenylation separately. Unlike *Ca*FADS, the FMNAT domain of *Ec*FADS needs reducing conditions for adenylation reaction.[42,43] DTT was used to fulfil the reducing conditions where, without DTT, the RFK domain produces a small amount of FMN (Figure 4.5C). The addition of the DTT increases FMN production, and FAD formation can also be seen. In the same analysis, we observed that Mg^{2+} ion also aids in RFK activity, and for FMNAT activity, Mg^{2+} ion is necessary along with DTT (Figure 4.1B). In the titration with various concentrations of Mg^{2+} ion, 3 mM was found suitable for both modules.

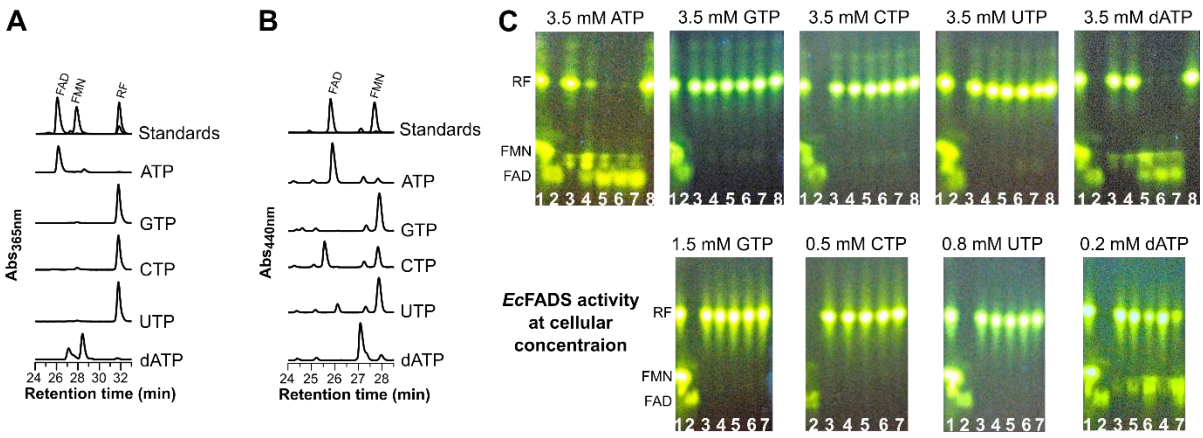


Figure 4.2 *Ec*FADS activity with nucleotides (NTPs) and dATP. (A) HPLC chromatogram showing *Ec*FADS activity with NTPs where RF is a starting substrate. *Ec*FADS is not active with other NTPs except dATP. **(B)** HPLC chromatogram showing *Ec*FADS activity with NTPs where FMN is a starting substrate. The FMNAT domain is active with all NTPs and dATP except GTP. **(C)** TLCs for *Ec*FADS activity with RF and NTPs at cellular concentrations. Lane 1- FMN and RF, 2- FAD, 3- reaction for 1 min, 4- reaction for 10 min, 5- reaction for 1 h, 6- reaction for 6 h, 7- reaction for 15 h, 8- no enzyme control reaction for 15 h. At a 3.5 mM concentration of NTPs, *Ec*FADS showed activity with ATP and dATP only, and activity with other NTPs was not detectable. However, dATP was also found active at the cellular concentration of 0.2 mM, showing a small spot for dFAD formation, while other NTPs still failed to give any activity.

DTT and β-mercaptoethanol have been known to reduce the disulfide bond, while DTT can also potentially reduce flavins and β-mercaptoethanol cannot.[44–46] On addition of DTT to flavins (RF, FMN, and FAD), a decrease in fluorescence was observed with an increase in DTT concentration (Figure 4.1D). This decrease was due to the reduction of flavins, as confirmed by the disappearance of fluorescence on the addition of sodium dithionite, which ultimately reduces the flavins.

FAD synthetase from *M. jannaschii* depends on divalent metals for its functioning. However, it was found that the most effective performance was with $Co^{2+,}$ which is 4-fold higher compared to its performance with Mg^{2+}.[29] We also found that *Ec*FADS is not specific only to Mg^{2+} ions, and it also accepts Mn^{2+}, Fe^{2+}, and Co^{2+} for its activity (Figure 4.1E). However, Mg^{2+} is still the best for its activity, following the order of activity Mg^{2+}> Fe^{2+}> Mn^{2+}> Co^{2+}. We also optimized the time for *Ec*FADS activity to be ~6 h (Figure 4.1F).

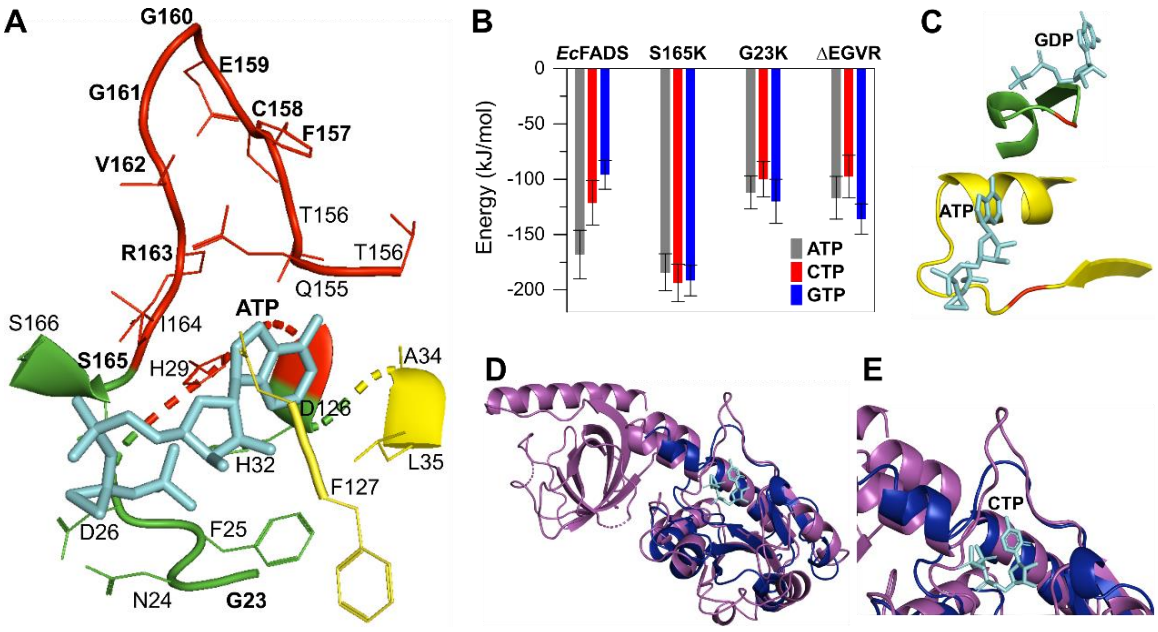


Figure 4.3. Mutations to alter the nucleotide selectivity of *Ec*FADS. (A) This modelled structure of *Ec*FADS shows the closest residues to ATP where red residues are near the adenine base, green surrounding the phosphates, and yellow are close to the ribose ring. (B) Interaction energy for the binding of *Ec*FADS and its mutants with ATP, CTP, and GTP where two residues (G23K and S165K) near the phosphates were mutated with lysine (K), making the negative charge and position of phosphates stable. ΔEGVR mutant was created by shortening the loop (red) around adenine, which shows lower interaction energy for each NTP. (C) glycine (G23) to serine (S) mutation (G23S) for GTP specificity based on a point mutation G105S in FtsZ, a GTPase, in the loop surrounding the phosphates. FtsZ is shown in green, bound to GDP, and the G105 residue is shown in red. *Ec*FADS is in yellow, bound to ATP and G23 residue in red. (D) Structure alignment of *Ec*FADS with *Bacillus subtilis* glycerol 3-phosphate cytidylyl (*Bs*G3PCT) transferase. FMNAT domain of *Ec*FADS has structural similarity with cytidylyl transferases. (E) Close view of the Loops surrounding the nucleobase in *Ec*FADS and *Bs*G3PCT. Both enzymes share the same structure near the nucleotide except for the Loop size.

4.3.2 The activity of *Ec*FADS with other NTPs and dNTPs

Since *Ec*FADS has two modules (RFK and FMNAT), we tried its activity with both the substrates, RF and FMN. On treating *Ec*FADS with other NTPs and dATP in place of ATP, it does not accept other nucleotides except dATP when RF is used as the substrate (Figure 4.2A). However, with FMN, it is active with each nucleotide (ATP, CTP, UTP, and dATP) except GTP and synthesizes FAD nucleobase analogues (FCD, FUD, and dFAD) (Figure 4.2B). The activity with dATP is similar to ATP, which can be due to similar structural features except for the hydroxy group. The activity with CTP and UTP was lesser than the native substrate ATP, which can be credited to lower interaction energy for CTP with the FMNAT domain relative to ATP (Figure 4.3B).

Since GTP has the lowest interaction energy, it does not give activity with *Ec*FADS. The other possible explanation is the difference in cellular concentration of each NTP, where ATP is produced more significantly (~3.5 mM) than others (GTP ~1.5, CTP ~0.5, UTP ~0.8, and dATP ~0.2 mM) in *E. coli*.[47,48] At cellular concentrations, only dATP actively participated in the reaction forming dFAD, and other NTPs had no detectable activity with *Ec*FADS (Figure 4.2C). However, dFAD can be produced at cellular concentration, and the FMNAT domain can produce FCD, FUD, and dFAD; the LC-MS analysis of *E. coli* cell lysate could not detect the presence of any FAD analogue (Figure 4.7C).

4.3.3 Potential residues for ATP specificity

The enzyme interacts with three moieties of ATP: the phosphate moiety (interacting with green residues), the ribose ring (yellow residues), and the adenine base (red residues), as shown in Figure 4.3A. Since most of the interactions of ATP are through the phosphate region, we did a computation analysis to create mutations around it to make it GTP-specific. G23 and S165 were mutated with lysine (K) as its positive charge can stabilize phosphate's negative charge and shows higher mutant-GTP interaction energy than GTP with wild-type (Figure 4.3B). However, the *in vitro* activity was decreased with ATP and diminished with other NTPs, making these mutants ATP-specific; however, they were less efficient than wild-type *Ec*FADS (Figure 4.4). The decline in activity can be mainly due to the tight binding of ATP, CTP and GTP to the S165K mutant, as suggested by the energy profile of CTP and GTP binding to the mutant in comparison to the wild-type enzyme (Figure 4.3B). In the case of the G23K mutant, the binding should be very poor to succeed in the reaction, as suggested by the interaction

energies of NTPs with this mutant (Figure 4.3B). A GTPase, FtsZ, was converted to an ATPase with a point mutation G105S (Figure 4.3C, FtsZ in green and G105 in red).[49] Since, the G23 residue of *Ec*FADS (Figure 4.3C, *Ec*FADS in yellow and G23 residue in red) was also predicted as a potential residue for GTP specificity computationally and lies at a similar position as in FtsZ, we created *Ec*FADS G23S mutant. However, we found it completely inactive (Figure 4.4C).

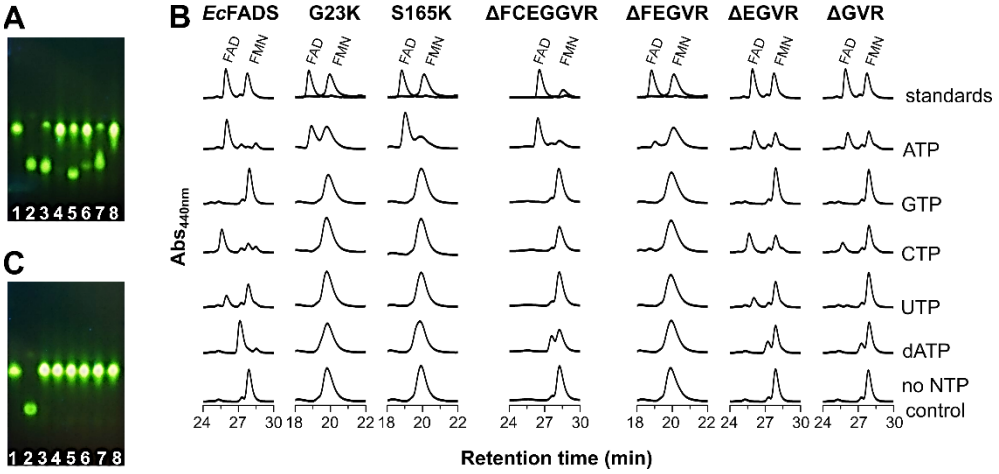


Figure 4.4. The activity of *Ec*FADS and its FMNAT site mutants with FMN and various NTPs in optimized conditions. (A) TLC for the activity of *Ec*FADS with FMN and NTPs. Lane 1- FMN, 2- FAD, 3- reaction with ATP, 4- reaction with GTP, 5- reaction with CTP, 6- reaction with UTP, 7- reaction with dATP, 8- control reaction without NTP. **(B)** HPLC chromatogram showing activity assay of *Ec*FADS and its mutants with NTPs. *Ec*FADS can form FCD, FUD, and dFAD besides FAD. The lysine mutation of G23 and S165 makes the FMNAT domain of *Ec*FADS specific for ATP. Deleting some residues in the loop region around adenine binding decreased activity with each NTP in the case of ΔEGVR and ΔGVR, while activity completely vanished with CTP and UTP for ΔFCEGGVR. However, the FEGVR shows almost no activity with any NTP except very little for ATP. **(C)** TLC for the activity of G23S mutant with FMN and NTPs. The FMNAT domain of the G23S mutant is entirely inactive, showing no activity with any NTP. The lanes are the same as in Figure 4.4A.

4.3.4 Loop of cytidylyl transferase in *Ec*FADS makes it promiscuous to GTP

Since adenylyl transferase is from the nucleotidylyl transferase family, the DALI server was used to find nucleotidylyl transferase structural homologs of *Ca*FADS.[36] The PDB90 feature of DALI eliminated the redundant protein structures by considering structure homologs with less than 90% sequence identity. We found some cytidylyl transferases (such as *Bacillus subtilis* glycerol-3-phosphate cytidylyl transferase (pdb: 1COZ), *Rattus norvegicus* phosphocholine cytidylyl transferase (pdb: 4MVC), and *Homo sapiens* phosphoethanolamine cytidylyl transferase (pdb: 3ELB)) available in PDB database. FMNAT domain of the Bacterial FAD synthetase is categorized within the HxGH motif-containing nucleotidylyl transferase superfamily, and cytidylyl transferases also fall in this category. The cytidyltransferases do not

show a similar sequence to *Ec*FADS but structurally align well with the FMNAT domain of *Ec*FADS (Figure 4.3D, Table 4.2). The main difference between these cytidylyl transferase enzymes and the FMNAT domain of *Ec*FADS is the loop size near the nucleobase region (Figure 4.3E). Hence, tuning the loop of *Ec*FADS according to the cytidyl transferase may result in a promiscuous or specific enzyme for other NTPs.

The glycerol-3-phosphate cytidylyl transferase (G3PCT) is specific to CTP or dCTP.[50] We swapped the whole loop (QTFCEGGVRIS) of the FMNAT domain with the loop (PRTEGIS) of glycerol-3-phosphate cytidylyl transferase and looked for its activity with the NTPs (Figure 4.5, 4.6A). Unlike *Ec*FADS, This FMNAT domain of LoopSwap mutant can utilize GTP along with ATP, CTP, UTP, and dATP and produces FAD and FAD analogues (FGD, FCD, FUD, and dFAD). Compared to *Ec*FADS, this LoopSwap mutant is more efficient with each NTP (Figure 4.7B). In most the cofactors such as ATP, FAD, NAD, SAM, and CoA, the adenine base binds to the enzyme through backbone interactions instead of side chain interactions. the activity is dependent on the shape of the active site providing the specific backbone interaction rather than on side chain interactions. Breaking these backbone interactions by loop alteration in the FMNAT module of *Ec*FADS did not change the shape of the active site; However, the selectivity towards the ATP has broken down.

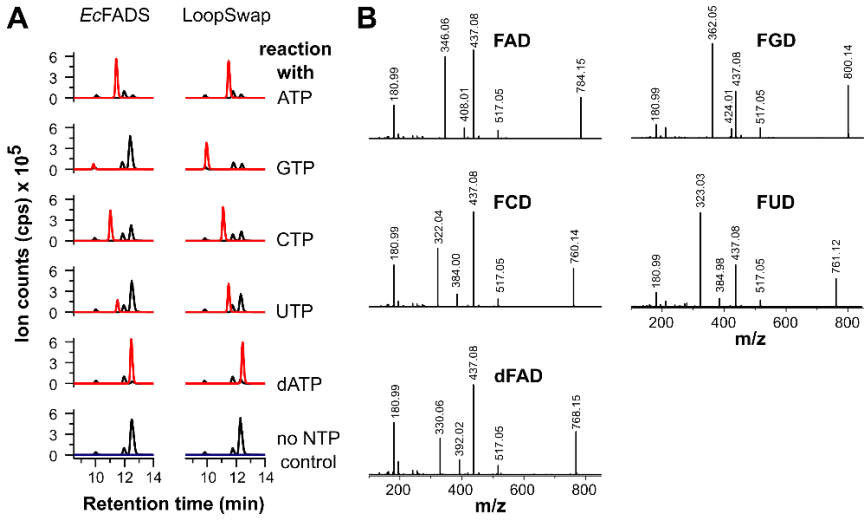


Figure 4.5. Activity of LoopSwap mutant with NTPs and their LC-MS analysis. (A) EICs for the activity of LoopSwap mutant with NTPs compared to *Ec*FADS. FMN was used as the starting substrate; its EIC is black. The EICs for produced FNDs from corresponding NTP have been shown in red. LoopSwap mutant accepts all the NTPs and dATP as substrate and is more efficient than *Ec*FADS. **(B)** Mass fragment analysis of all the FNDs matches the reported fragment masses of synthesized FNDs (Chapter 3).

The other interesting observation is that the mutant was also active with GTP with similar efficiency as *Ec*FADS with ATP (Figure 4.7B). Due to the *in vitro* activity of *Ec*FADS and its mutant with other NTPs, we tried heterologous expression of *Ec*FADS and its LoopSwap mutant in *E. coli* BL21(DE3) for FAD nucleobase analogues formation. We used lysate of completely grown cultures of these strains for LC-MS analysis, taking pET28a transformed BL21(DE3) cells as our negative control (Figure 4.6B). Empty pET28a-containing cells did not show any FAD analogue in the LS-MS analysis. *Ec*FADS-expressing cells also did not show any FAD analogue, but more FAD production was observed because of the extra expression of *Ec*FADS from the plasmid. However, this LoopSwap mutant strain showed FAD production and some FCD, FUD, and FGD, where FUD had a higher amount than others (Figure 4.6B). These results indicate that this loop goes through conformational dynamics, essential for nucleotide selectivity inside the cell.

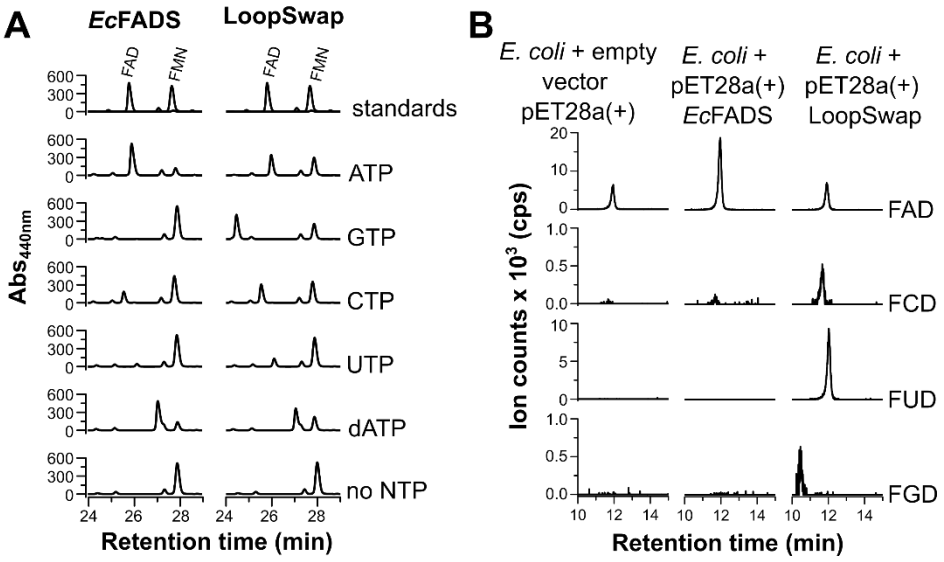


Figure 4.6. The LoopSwap mutant shows *in vitro* promiscuity for each NTP and produces FNDs inside the cell. (A) HPLC data showing *in vitro* activity of LoopSwap mutant compared to wild-type. LoopSwap is not only more efficient with the *Ec*FADS, but it is also active with GTP. (B) EICs of FAD and its analogues in the strains transformed with empty pET28a and pET28a-*Ec*FADS or pET28a-LoopSwap. The LoopSwap mutant synthesizes FAD nucleobase analogues inside the cell, while *Ec*FADS is very specific for ATP, producing FAD only.

4.3.5 Changing the Loop dynamics by shortening the loop surrounding the nucleobase, which makes *Ec*FADS promiscuous for NTPs inside the cell

To further clarify the role of the adenine surrounding loop in nucleotide specificity, we created various mutants by shortening the loop (QTFCEGGVRIS). For this, 3-7 amino acids were deleted, creating four mutants – ΔGVR, ΔEGVR, ΔFEGVR, and ΔFCEGGVR by

considering the loop (PRTEGIS) of the G3PCT (Figure 4.7A). The ΔGVR mutant was created by deleting the three residues GVR from the top of the loop. However, a decrease in the activity of this mutant with ATP, CTP, and dATP was observed, getting no activity with GTP and UTP (Figure 4.4B, 4.7B), and this loop is one residue bigger than the loop of G3PCT (Figure 4.7A). We further deleted four EGVR residues from the EGGVR fragment, generating a mutant ΔEGVR with a loop of the same size as in G3PCT (Figure 4.7). One of the Gs residues was not deleted from the EGGVR fragment to keep the GIS sequence (QTFC**GIS**) of the remaining loop similar to the loop (PRTE**GIS**) of G3PCT. The ΔEGVR mutant also shows decreased activity with the NTPs relative to *Ec*FADS, while it is more efficient than the ΔGVR in the case of CTP and UTP (Figure 4.4B, 4.7B). The homology-modelled structure of the ΔEGVR mutant shows a bend in loop structure in comparison to G3PCT, where the residue F bends towards the nucleobase, restricting the bigger adenine base more than pyrimidine bases in the active site (Figure 4.7A). On further deleting residue F, the ΔFEGVR mutant shows diminished activity for each nucleotide except very little with ATP (Figure 4.4B, 4.7B). We created the ΔFCEGGVR mutant by shortening the loop to two residues, which was found to be specific for adenine base, synthesizing FAD and dFAD only. The decrease in the activity of ShortLoop mutants can be attributed to the low interaction energy of these mutants with NTPs, as depicted for the ΔEGVR mutant with ATP, CTP, and GTP (Figure 4.3B).

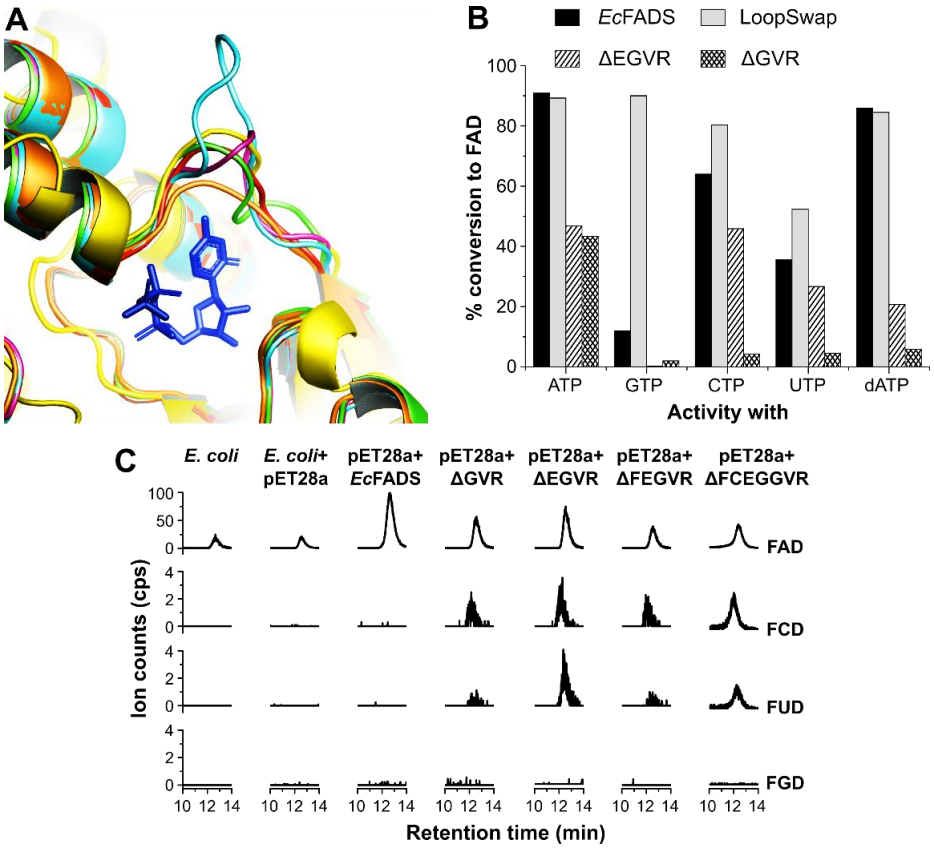


Figure 4.7. The ShortLoop mutants and their activity inside the cell, compared to wild-type *Ec*FADS, talk about the importance of loop in nucleotide specificity. (A) The loop structure comparison of ShortLoop mutants relative to glycerol-3-phosphate cytidylyl transferase (G3PCT) bound to CTP (blue colour). The colour of these mutants is as follows: *Ec*FADS- cyan, G3PCT- yellow, ΔGVR- magenta, ΔEGVR- green, ΔFEGVR- red, and ΔFCEGGVR-orange. **(B)** Quantitative analysis of FAD and its analogues formed by *Ec*FADS and its mutants from various nucleotides, respectively. LoopSwap is more active with other nucleotides compared to *Ec*FADS. At the same time, the short loop mutants show a decrease or complete demolition of activity with ATP and other nucleotides **(C)** EICs of FAD and its analogues in the strains transformed with empty pET28a and pET28a-*Ec*FADS or pET28a-ShortLoop mutants compared to *E. coli* cells. ΔEGVR mutant can synthesize FAD nucleobase analogs, while *Ec*FADS fails to produce the same, which shows that the loop plays a primary role in nucleotide specificity for the cell.

To understand the activity of the loop responsible for the ATP selectivity inside a cell, we heterologously expressed these mutants in the BL21(DE3) cells and studied their lysate on LC-MS (Figure 4.7C). Besides, these mutants have decreased or no *in vitro* activity for NTPs; still, we observed FCD and FUD formation along with FAD in all these mutant strains; however, FGD was not detected. These results further conclude that the loop dynamics is essential for this nucleotide selectivity inside the cell, despite the deleted amino acids not interacting directly with the nucleobase of NTPs.

4.4 Conclusion

In this study, we have shown that *Ec*FADS can individually accept RF or FMN as substrate and perform both reactions: phosphorylation of RF and adenylation of FMN. Besides, the RFK domain is specific for ATP out of all NTPs, and the FMNAT domain is promiscuous for CTP, UTP, and dATP and can synthesize corresponding FAD analogues while accepting ATP as a primary substrate. Besides the *in vitro* promiscuity of the FMNAT domain, *Ec*FADS remains specific for ATP inside the cell. Mutations in the phosphate binding region suggest that the surrounding residues are particular for the activity as the efficiency of the enzyme for each NTP decreases compared to the native enzyme. However, the G23K and S165K mutations make the FMNAT domain specific for ATP. This specificity can be reasoned as the tight phosphate binding with lysine in the S165K mutant and very loose binding in the G23K as per the interaction energy (Figure 4.3B). These mutations take away the flexibility of other ligands to adopt the conformation necessary for the active site to make the reaction happen. The other mutant, G23S, creates an entirely inactive FMNAT domain, which can be attributed to the loss of flexibility of the corresponding loop. Predicting the residues corresponding to ATP specificity is challenging as the adenine interaction with protein residues happens through the

backbone. The structure alignment of the FMNAT domain with cytidylyl transferases suggests that the loop size can be one of the reasons for ATP selectivity. The LoopSwap mutant supports this fact, as the loop exchange broke the ATP selectivity, accepting other NTPs efficiently, and FGD was obtained along with FAD, FCD, FUD, and dFAD with better efficiency. Besides this, the *Ec*FADS is very specific for ATP inside a cell, as no FAD analogue was observed in the LC-MS analysis for *Ec*FADS expressing cells. However, this specificity vanishes for the LoopSwap mutant when heterologously expressed inside the cell. The role of the loop size in ATP specificity was also confirmed by the ShortLoop mutants, where the *in vitro* activity decreases with each NTP, while the heterologous expression of each short loop mutant shows FAD analogues (FCD and FUD) inside the cell. These observations show that this loop goes through some sort of loop dynamics for this ATP specificity *in vivo*.[51] To further understand the role of the loop in the adenine recognition, we looked for the presence of a similar loop in 80 PDB structures randomly taken from the already available published list of 985 crystal structures bounded to ATP, ADP, AMP, CoA, FAD, NAD, SAM and their analogues (Table 4.3).[52] In this context, we have taken into consideration those loops that are intricately folded around the nucleobase, involving at least four consecutive residues. Loops that are positioned away from the nucleobase were not considered. Our findings revealed that 62 % of ATP and its derivatives featured this distinct loop structure, showcasing varying sizes of folding. Additionally, a subset of 4 out of 8 SAH (S-adenosyl-L-homocysteine) molecules also displayed the loop surrounding the adenine base, with a length spanning 7 to 10 amino acids. On the other hand, for molecules like CoA, FAD, NAD, and their derivatives, interactions with the loop occurred either at a distant fold or involved loop sizes of less than 4 amino acids. Due to the distinct folding observed around ATP and its derivatives compared to other cofactor molecules, there is a possibility that this type of loop is common around ATP-binding proteins.

The wild type *Ec*FADS and its mutant enzymes behave very differently *in vitro* and *in vivo*. *Ec*FADS shows *in vitro* activity with CTP, UTP, and dATP while its expression in the cell results only extra production of FAD and no FAD analogue. Similarly, ShortLoop mutants show decreased or no activity with each nucleotide, still their expression in the cell shows FCD and FUD. This phenomenon suggests that the Loop dynamics are important for the specificity and disruption of that dynamics break this specificity. There can be three more factors for low concentration FAD analogue in the strain. cellular concentration of nucleotides varies in the cell following the order ATP>GTP>UTP>CTP where K_m of the nucleotides can affect the analogue formation. The other factor can be the utilisation of FAD over its analogues by

flavoproteins as prosthetic group. FAD is tightly bound to flavoprotein which provides extra stability to FAD while the analogues cannot get this stability by the flavoproteins. As ATP is native substrate and highest in concentration, enzyme prefers ATP over other nucleotides synthesizing FAD the most. These can be the reason that the LoopSwap, being very efficient enzyme, produces low amount of FAD analogues.

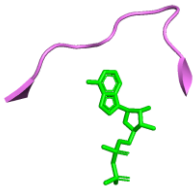
In summary, we have created a potential LoopSwap mutant with great promiscuity for NTPs producing FAD analogues *in vitro* and inside the cell. We have also created *in vitro* ATP-specific *Ec*FADS mutants like S165K and G23K. Our study shows that some sort of loop dynamics is vital for this ATP specificity inside the cell. In essence, this study proposes that mimicking the behaviour of proteins by manipulating vital loop structures' conformational changes is a potent strategy for fine-tuning enzyme activity, and this can be accomplished without specifically targeting residues within the active site.

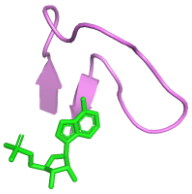
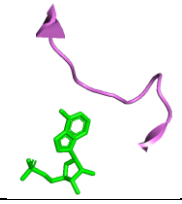
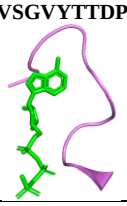
Table 4.2 List of structural homologs of *Ca*FADS.

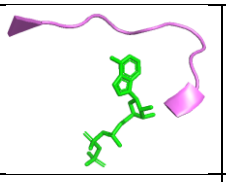
S. No.	PDB id-Chain	z-score	RMSD	% identity	Structure	Description
1	2x0k-A	55.2	0	100	PDB	RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBF
2	3op1-B	31.5	2.2	30	PDB	MACROLIDE-EFFLUX PROTEIN
3	1s4m-A	28.2	2.6	32	PDB	RIBOFLAVIN KINASE/FMN ADENYLYLTRANSFERASE
4	3q10-A	14.4	5.1	16	PDB	PANTOATE--BETA-ALANINE LIGASE
5	3uk2-A	14.4	5.5	17	PDB	PANTOTHENATE SYNTHETASE
6	1n06-B	14.2	2.3	22	PDB	PUTATIVE RIBOFLAVIN KINASE
7	3inn-C	14.2	5.3	12	PDB	PANTOTHENATE SYNTHETASE
8	3ag6-A	14	5.1	14	PDB	PANTOTHENATE SYNTHETASE
9	5ucr-A	13.9	6.4	18	PDB	PANTOTHENATE SYNTHETASE
10	1iho-A	13.7	4.3	15	PDB	PANTOATE--BETA-ALANINE LIGASE
11	1q9s-A	13.5	2.8	27	PDB	HYPOTHETICAL PROTEIN FLJ11149
12	2ejc-A	13.4	5.5	14	PDB	PANTOATE--BETA-ALANINE LIGASE
13	7tot-A	13.4	3.8	14	PDB	PANTOTHENATE SYNTHETASE
14	3mxt-A	13.3	5.3	16	PDB	PANTOTHENATE SYNTHETASE
15	3mue-A	13.3	3.8	15	PDB	PANTOTHENATE SYNTHETASE
16	3n8h-A	13.2	8.2	14	PDB	PANTOTHENATE SYNTHETASE
17	3q10-D	13	3	16	PDB	PANTOATE--BETA-ALANINE LIGASE
18	1coz-A	12.7	2.5	16	PDB	PROTEIN (GLYCEROL-3-PHOSPHATE
19	1ufv-A	12.5	3.7	16	PDB	PANTOATE-BETA-ALANINE LIGASE
20	3bnw-A	11.9	2.8	29	PDB	RIBOFLAVIN KINASE, PUTATIVE
21	3ivc-A	11.8	3.7	16	PDB	PANTOTHENATE SYNTHETASE
22	2b7l-A	11.3	2.7	17	PDB	GLYCEROL-3-PHOSPHATE **CYTIDYLYLTRANSFERASE**
23	1v47-A	11.1	2.9	15	PDB	ATP SULFURYLASE

24	3k9w-A	11	3.5	18	PDB	PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE
25	7wgj-A	11	3.3	14	PDB	PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE
26	3otw-A	10.9	3.3	19	PDB	PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE
27	3l93-A	10.9	3.2	19	PDB	PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE
28	8i8j-B	10.8	3.1	15	PDB	PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE
29	1o6b-A	10.7	3.7	13	PDB	PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE
30	5h7x-E	10.6	3.5	14	PDB	PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE
31	4mvd-B	10.5	4.2	11	PDB	CHOLINE-PHOSPHATE **CYTIDYLYLTRANSFERASE**
32	3glv-A	10.5	3.3	15	PDB	LIPOPOLYSACCHARIDE CORE BIOSYNTHESIS PROTEIN
33	3nbk-D	10.4	3.4	14	PDB	PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE
34	4nah-B	10.4	3.3	13	PDB	PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE
35	5o08-A	10.3	3.5	14	PDB	PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE
36	5x3d-A	10.3	3.3	17	PDB	PHOSPHOENOLPYRUVATE PHOSPHOMUTASE;
37	3nd5-A	10.2	3.2	16	PDB	PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE
38	4zcs-A	10.2	3.9	13	PDB	CHOLINE-PHOSPHATE **CYTIDYLYLTRANSFERASE**
39	1vlh-B	10.2	3.6	16	PDB	PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE
40	5jbn-A	10.2	3.5	18	PDB	PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE
41	5ts2-A	10.2	3.5	18	PDB	PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE
42	1od6-A	10.1	3.3	16	PDB	PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE
43	4f3r-B	10	3.1	16	PDB	PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE
44	1f9a-A	10	3.1	13	PDB	HYPOTHETICAL PROTEIN MJ0541
45	3dlo-D	9.9	3.3	15	PDB	UNIVERSAL STRESS PROTEIN

Table 4.3 The list of 80 PDB structures from the PDB database bound to ATP, ADP, AMP, CoA, FAD, NAD, SAM, and their derivatives.

PDB ID	Ligand	Loop sequence around the nucleobase	N1 H-bonds	N3 H-bonds	N6 Watson-Crick H-bonds	N6 Hoogsteen H-bonds	N7 H-bonds
2FT0	ACO		HOH/O		HOH/O	HOH/O	
3MQG	ACO						HOH/O
5KF9	ACO		HOH/O				
1M15	ADP				S122/O	S282/OG	
2C9O	ADP		V40/N		V40/O	Y366/OH	
2FNA	ADP		F15/N	K8/NZ	F15/O	HOH/O	HOH/O
2PYW	ADP	**RYLEPPH** 	L119/N	HOH/O	R117/O		HOH/O
2XZO	ADP	**GLPDLN**			D470/O		Q475/NE2
3D36	ADP	**TGVGM**	HOH/O	HOH/O	D352/OD2	HOH/O	HOH/O
3EHH	ADP	**DGTFKG**	HOH/O	K325/N	D320/OD2	HOH/O	HOH/O
3LL3	ADP		HOH/O		HOH/O	HOH/O	
3TTC	ADP				E296/O	HOH/O	HOH/O

3W0O	ADP	**RAQG**	A95/N	HOH/O	R93/O	HOH/O	HOH/O
4NE2	ADP		HOH/O			HOH/O	K303/NZ
4QPM	ADP	**ELYSYG**	Y869/N		E867/O		HOH/O
5FBS	ADP	**MISPE**	I186/N		Q184/O		
5HE9	ADP				HOH/O	HOH/O	HOH/O
5IZ4	ADP	**ADVG**	V67/N	HOH/O	D66/OD1	HOH/O	HOH/O
5LB3	ADP	**FGFDSFK**	HOH/O	HOH/O	S28/O		Q34/NE2
5LTJ	ADP						HOH/O
5LY3	ADP		HOH/O	HOH/O	HOH/O/HOH/O		HOH/O
12AS	AMP	**RPDEDRLSPLHSV** 	S111/N		S111/O	E103/OE1	
1CJA	AMP	**ELVR**	V165/N		E163/O		
1KHT	AMP				G104/O	T91/OG1	
1S68	AMP	**KIGH**	HOH/O			E34/O	I36/N
2GXQ	AMP	**GLTTP**	HOH/O		T23/O		Q28/NE2
3LFR	AMP		I74/N	HOH/O	I74/O	R96/O	HOH/O
4HG0	AMP	**RSQMIT** 	I80/N		I80/O	R102/O	
5D0N	AMP	**SGIPRSSPSRNGRL**					
5GMD	AMP	**GLAS**			HOH/O		HOH/O
5NC8	AMP	**GDAA**	A457/N	H437/N	D456/OD1		
1P3D	ANP		HOH/O		HOH/O	N295/OD1	N295/ND2
3A99	ANP	**RPEP**		HOH/O	E121/O		HOH/O
3C1M	ANP	**VSGVYTTDP** 	Y236/N		Y236/O	HOH/O	
3RC3	ANP		HOH/O	HOH/O			
4D2I	ANP				I460/O		
5FLG	ANP		V146/N		V146/O	E31/OE2	HOH/O
5IX1	ANP	**NGNGM**			D67/OD1		HOH/O
1HP1	ATP	**PFGN**	N431/ND2				HOH/O
1QHH	ATP	**AHLN**		Q16/NE2			
2A5Y	ATP	**NVPKQMTCYI**			Y131/O	HOH/O	
3FKQ	ATP	**APRYEHA**	Y331/N	N307/ND2	P329/O	N307/OD1	HOH/O

							
4CYI	ATP	**FHPYI**	H340/N		D338/O	HOH/O	HOH/O
4FFL	ATP	**FDVIK**	V50/N	N73/ND2	D49/OD1		K32/N
5J1S	ATP		F70/N		F70/O	T106/O	HOH/O
2EIS	COA					HOH/O	
4LRT	COA		HOH/O		HOH/O	HOH/O	HOH/O
4PSW	COA						
4U89	COA			HOH/O		K78/O	HOH/O
2CUL	FAD		A91/N	Q34/N	A91/O	HOH/O	HOH/O
2IPI	FAD		V203/N		V203/O	S67/OG	S67/OG
2QCU	FAD		A172/N	A34/N		HOH/O	
3DME	FAD		L173/N	A35/N	L173/O	HOH/O	HOH/O
3EC6	FAD			HOH/O			
3EWK	FAD						
3F8D	FAD		V92/N	E46/N	V92/O		
3JQQ	FAD					HOH/O	
3NKS	FAD		V257/N	S35/N	V257/O		HOH/O
4EQS	FAD		V81/N	K34/N	V81/O	HOH/O	HOH/O
4H4R	FAD		A82/N	D40/N	A82/O	HOH/O	HOH/O
4X9M	FAD		V177/N	K34/N	V177/O		
5JCI	FAD		I96/N	K40/N	I96/O	HOH/O	HOH/O
5KF6	FAD				T393/O		
1KAE	NAD			HOH/O			
4CPD	NAD						
5TT5	NAD		K290/NZ	HOH/O	E113/OE2	L114/O	L116/N
2C0C	NAP		HOH/O	HOH/O	HOH/O	HOH/O	
2Y5D	NAP		HOH/O				
3LZW	NAP		HOH/O	HOH/O			
4DPL	NAP		HOH/O	HOH/O			
4R3N	NAP		HOH/O				
5DP2	NAP		HOH/O	HOH/O	HOH/O/HOH/O	HOH/O	HOH/O
2HA8	SAH	**IPQQGIIRSL** 	I149/N		I149/O		L158/N
2JJQ	SAH		D326/N	S300/N	D326/OD1	HOH/O	
3H2B	SAH		I150/N		HOH/O		HOH/O
3NDC	SAH	**RASWPDQ**	A193/N		A193/O	HOH/O/P10/O	HOH/O/HOH/O
3TOS	SAH	**TFTGFPDV**	V167/N	T108/N	D166/OD1	HOH/O	
4JWJ	SAH		L232/N		HOH/O	R244/O	L246/N

5CCB	SAH		V164/N	F136/N	D163/OD1		
5MGZ	SAH	**GDAQLLAG**	A97/N	L71/N	D96/OD1	E4/OE1	

ACO: Acetyl coenzyme A; ADP: Adenosine diphosphate; AMP: Adenosine monophosphate; ATP: Adenosine triphosphate; CoA: Coenzyme A; FAD: Flavin adenine dinucleotide; NAD: Nicotinamide adenine dinucleotide; NAP: Nicotinamide adenine dinucleotide phosphate; SAH: S-adenosylhomocysteine

References

(1) Denessiouk, K. A.; Rantanen, V.-V.; Johnson, M. S. Adenine Recognition: A Motif Present in ATP-, CoA-, NAD-, NADP-, and FAD-Dependent Proteins. *Proteins: Structure, Function, and Genetics* **2001**, *44* (3), 282–291. https://doi.org/10.1002/prot.1093.

(2) 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.

(3) Wang, S.; Zhu, W.; Wang, X.; Li, J.; Zhang, K.; Zhang, L.; Zhao, Y.-J.; Lee, H.; Zhang, L. Design, Synthesis and SAR Studies of NAD Analogues as Potent Inhibitors towards CD38 NADase. *Molecules* **2014**, *19* (10), 15754–15767. https://doi.org/10.3390/molecules191015754.

(4) Bruns, S.; Cakić, N.; Mitschke, N.; Kopke, B. J.; Rabus, R.; Wilkes, H. A Novel Coenzyme A Analogue in the Anaerobic, Sulfate-Reducing, Marine Bacterium *Desulfobacula Toluolica* Tol2 ^{T}. *ChemBioChem* **2023**, *24* (2), e202200584. https://doi.org/10.1002/cbic.202200584.

(5) Chen, X.; Li, S.; Liu, L. Engineering Redox Balance through Cofactor Systems. *Trends Biotechnol* **2014**, *32* (6), 337–343. https://doi.org/10.1016/j.tibtech.2014.04.003.

(6) Ji, D.; Wang, L.; Hou, S.; Liu, W.; Wang, J.; Wang, Q.; Zhao, Z. K. Creation of Bioorthogonal Redox Systems Depending on Nicotinamide Flucytosine Dinucleotide. *J Am Chem Soc* **2011**, *133* (51), 20857–20862. https://doi.org/10.1021/ja2074032.

(7) Wang, L.; Ji, D.; Liu, Y.; Wang, Q.; Wang, X.; Zhou, Y. J.; Zhang, Y.; Liu, W.; Zhao, Z. K. Synthetic Cofactor-Linked Metabolic Circuits for Selective Energy Transfer. *ACS Catal* **2017**, *7* (3), 1977–1983. https://doi.org/10.1021/acscatal.6b03579.

(8) Guo, X.; Liu, Y.; Wang, Q.; Wang, X.; Li, Q.; Liu, W.; Zhao, Z. K. Non-natural Cofactor and Formate-Driven Reductive Carboxylation of Pyruvate. *Angewandte Chemie* **2020**, *132* (8), 3167–3170. https://doi.org/10.1002/ange.201915303.

(9) Liu, Y.; Feng, Y.; Wang, L.; Guo, X.; Liu, W.; Li, Q.; Wang, X.; Xue, S.; Zhao, Z. K. Structural Insights into Phosphite Dehydrogenase Variants Favoring a Non-Natural Redox Cofactor. *ACS Catal* **2019**, *9* (3), 1883–1887. https://doi.org/10.1021/acscatal.8b04822.

(10) Liu, Y.; Li, Q.; Wang, L.; Guo, X.; Wang, J.; Wang, Q.; Zhao, Z. K. Engineering <scp>d</Scp> -Lactate Dehydrogenase to Favor an Non-natural Cofactor Nicotinamide Cytosine Dinucleotide. *ChemBioChem* **2020**, *21* (14), 1972–1975. https://doi.org/10.1002/cbic.201900766.

(11) Wang, L.; Liu, B.; Liu, Y.; Sun, Y.; Liu, W.; Yu, D.; Zhao, Z. K. *Escherichia Coli* Strain Designed for Characterizing *in Vivo* Functions of Nicotinamide Adenine Dinucleotide Analogues. *Org Lett* **2019**, *21* (9), 3218–3222. https://doi.org/10.1021/acs.orglett.9b00935.

(12) Wang, X.; Zhou, Y. J.; Wang, L.; Liu, W.; Liu, Y.; Peng, C.; Zhao, Z. K. Engineering Escherichia Coli Nicotinic Acid Mononucleotide Adenylyltransferase for Fully Active Amidated NAD Biosynthesis. *Appl Environ Microbiol* **2017**, *83* (13), e00692-17. https://doi.org/10.1128/AEM.00692-17.

(13) 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.

(14) 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**, *24* (11), e202300055. https://doi.org/10.1002/cbic.202300055.

(15) Manstein, D. J.; Pai, E. F. Purification and Characterization of FAD Synthetase from Brevibacterium Ammoniagenes. *Journal of Biological Chemistry* **1986**, *261* (34), 16169–16173. https://doi.org/10.1016/S0021-9258(18)66693-1.

(16) Barile, M.; Brizio, C.; Valenti, D.; De Virgilio, C.; Passarella, S. The Riboflavin/FAD Cycle in Rat Liver Mitochondria. *Eur J Biochem* **2000**, *267* (15), 4888–4900. https://doi.org/10.1046/j.1432-1327.2000.01552.x.

(17) Merrill, A. H.; McCormick, D. B. Affinity Chromatographic Purification and Properties of Flavokinase (ATP:Riboflavin 5'-Phosphotransferase) from Rat Liver. *Journal of Biological Chemistry* **1980**, *255* (4), 1335–1338. https://doi.org/10.1016/S0021-9258(19)86034-9.

(18) Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H. Construction of *Escherichia Coli* K-12 In-frame, Single-gene Knockout Mutants: The Keio Collection. *Mol Syst Biol* **2006**, *2* (1), 2006.0008. https://doi.org/10.1038/msb4100050.

(19) Rodionov, D. A.; Hebbeln, P.; Eudes, A.; ter Beek, J.; Rodionova, I. A.; Erkens, G. B.; Slotboom, D. J.; Gelfand, M. S.; Osterman, A. L.; Hanson, A. D.; Eitinger, T. A Novel Class of Modular Transporters for Vitamins in Prokaryotes. *J Bacteriol* **2009**, *191* (1), 42–51. https://doi.org/10.1128/JB.01208-08.

(20) Jaehme, M.; Slotboom, D. J. Diversity of Membrane Transport Proteins for Vitamins in Bacteria and Archaea. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2015**, *1850* (3), 565–576. https://doi.org/10.1016/j.bbagen.2014.05.006.

(21) Frago, S.; Martínez-Júlvez, M.; Serrano, A.; Medina, M. Structural Analysis of FAD Synthetase from Corynebacterium Ammoniagenes. *BMC Microbiol* **2008**, *8* (1), 160. https://doi.org/10.1186/1471-2180-8-160.

(22) Herguedas, B.; Martínez-Júlvez, M.; Frago, S.; Medina, M.; Hermoso, J. A. Crystallization and Preliminary X-Ray Diffraction Studies of FAD Synthetase from *Corynebacterium Ammoniagenes*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **2009**, *65* (12), 1285–1288. https://doi.org/10.1107/S1744309109044789.

(23) Herguedas, B.; Martínez-Júlvez, M.; Frago, S.; Medina, M.; Hermoso, J. A. Oligomeric State in the Crystal Structure of Modular FAD Synthetase Provides Insights into Its Sequential Catalysis in Prokaryotes. *J Mol Biol* **2010**, *400* (2), 218–230. https://doi.org/10.1016/j.jmb.2010.05.018.

(24) Marcuello, C.; Arilla-Luna, S.; Medina, M.; Lostao, A. Detection of a Quaternary Organization into Dimer of Trimers of Corynebacterium Ammoniagenes FAD Synthetase at the Single-Molecule Level and at the in Cell Level. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2013**, *1834* (3), 665–676. https://doi.org/10.1016/j.bbapap.2012.12.013.

(25) Serrano, A.; Sebastián, M.; Arilla-Luna, S.; Baquedano, S.; Herguedas, B.; Velázquez-Campoy, A.; Martínez-Júlvez, M.; Medina, M. The Trimer Interface in the Quaternary Structure of the Bifunctional Prokaryotic FAD Synthetase from Corynebacterium Ammoniagenes. *Sci Rep* **2017**, *7* (1), 404. https://doi.org/10.1038/s41598-017-00402-6.

(26) Lans, I.; Seco, J.; Serrano, A.; Burbano, R.; Cossio, P.; Daza, M. C.; Medina, M. The Dimer-of-Trimers Assembly Prevents Catalysis at the Transferase Site of Prokaryotic FAD Synthase. *Biophys J* **2018**, *115* (6), 988–995. https://doi.org/10.1016/j.bpj.2018.08.011.

(27) Spencer, R.; Fisher, J.; Walsh, C. Preparation, Characterization, and Chemical Properties of the Flavin Coenzyme Analogues 5-Deazariboflavin, 5-Deazariboflavin 5'-Phosphate, and 5-Deazariboflavin 5'-Diphosphate, 5' → 5'-Adenosine Ester. *Biochemistry* **1976**, *15* (5), 1043–1053. https://doi.org/10.1021/bi00650a015.

(28) Mishanina, T. V.; Kohen, A. Synthesis and Application of Isotopically Labeled Flavin Nucleotides. *J Labelled Comp Radiopharm* **2015**, *58* (9), 370–375. https://doi.org/10.1002/jlcr.3313.

(29) Mashhadi, Z.; Xu, H.; Grochowski, L. L.; White, R. H. Archaeal RibL: A New FAD Synthetase That Is Air Sensitive. *Biochemistry* **2010**, *49* (40), 8748–8755. https://doi.org/10.1021/bi100817q.

(30) Kitatsuji, K.; Ishino, S.; Teshiba, S.; Arimoto, M. Process for Producing Flavine Nucleotides. European Patent Office, 1992.

(31) Langer, S.; Hashimoto, M.; Hobl, B.; Mathes, T.; Mack, M. Flavoproteins Are Potential Targets for the Antibiotic Roseoflavin in Escherichia Coli. *J Bacteriol* **2013**, *195* (18), 4037–4045. https://doi.org/10.1128/JB.00646-13.

(32) van den Ent, F.; Löwe, J. RF Cloning: A Restriction-Free Method for Inserting Target Genes into Plasmids. *J Biochem Biophys Methods* **2006**, *67* (1), 67–74. https://doi.org/10.1016/j.jbbm.2005.12.008.

(33) Madeira, F.; Pearce, M.; Tivey, A. R. N.; Basutkar, P.; Lee, J.; Edbali, O.; Madhusoodanan, N.; Kolesnikov, A.; Lopez, R. Search and Sequence Analysis Tools Services from EMBL-EBI in 2022. *Nucleic Acids Res* **2022**, *50* (W1), W276–W279. https://doi.org/10.1093/nar/gkac240.

(34) Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E. The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nat Protoc* **2015**, *10* (6), 845–858. https://doi.org/10.1038/nprot.2015.053.

(35) Sali, A. Comparative Protein Modeling by Satisfaction of Spatial Restraints. *Mol Med Today* **1995**, *1* (6), 270–277. https://doi.org/10.1016/S1357-4310(95)91170-7.

(36) Holm, L. Dali Server: Structural Unification of Protein Families. *Nucleic Acids Res* **2022**, *50* (W1), W210–W215. https://doi.org/10.1093/nar/gkac387.

(37) Goodsell, D. S.; Morris, G. M.; Olson, A. J. Automated Docking of Flexible Ligands: Applications of Autodock. *Journal of Molecular Recognition* **1996**, *9* (1), 1–5. https://doi.org/10.1002/(SICI)1099-1352(199601)9:1<1::AID-JMR241>3.0.CO;2-6.

(38) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J Comput Chem* **2009**, *30* (16), 2785–2791. https://doi.org/10.1002/jcc.21256.

(39) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Structure, Function, and Bioinformatics* **2006**, *65* (3), 712–725. https://doi.org/10.1002/prot.21123.

(40) Frisch, M. J.; et al. Gaussian 03, Revision C.02. *Gaussian Inc; Wallingford CT; 2004* **2004**.

(41) Abraham, M. J.; van der Spoel, D.; Lindahl, E.; Hess, B.; the GROMACS development team. GROMACS User Manual Version 2019. **2019**.

(42) Serrano, A.; Frago, S.; Velázquez-Campoy, A.; Medina, M. Role of Key Residues at the Flavin Mononucleotide (FMN):Adenylyltransferase Catalytic Site of the Bifunctional Riboflavin Kinase/Flavin Adenine Dinucleotide (FAD) Synthetase from Corynebacterium Ammoniagenes. *Int J Mol Sci* **2012**, *13* (12), 14492–14517. https://doi.org/10.3390/ijms131114492.

(43) Serrano, A.; Frago, S.; Herguedas, B.; Martínez-Júlvez, M.; Velázquez-Campoy, A.; Medina, M. Key Residues at the Riboflavin Kinase Catalytic Site of the Bifunctional Riboflavin Kinase/FMN Adenylyltransferase From Corynebacterium Ammoniagenes. *Cell Biochem Biophys* **2013**, *65* (1), 57–68. https://doi.org/10.1007/s12013-012-9403-9.

(44) Iyer, K. S.; Klee, W. A. Direct Spectrophotometric Measurement of the Rate of Reduction of Disulfide Bonds. *Journal of Biological Chemistry* **1973**, *248* (2), 707–710. https://doi.org/10.1016/S0021-9258(19)44430-X.

(45) Loechler, E. L.; Hollocher, T. C. Reduction of Flavins by Thiols. 1. Reaction Mechanism from the Kinetics of the Attack and Breakdown Steps. *J Am Chem Soc* **1980**, *102* (24), 7312–7321. https://doi.org/10.1021/ja00544a027.

(46) Loechler, E. L.; Hollocher, T. C. Reduction of Flavins by Thiols. 2. Spectrophotometric Evidence for a Thiol-C(4a) Flavin Adduct and the Kinetics of Deprotonation of the -SH Group of the Dithiothreitol Adduct. *J Am Chem Soc* **1980**, *102* (24), 7322–7327. https://doi.org/10.1021/ja00544a028.

(47) Bochner, B. R.; Ames, B. N. Complete Analysis of Cellular Nucleotides by Two-Dimensional Thin Layer Chromatography. *Journal of Biological Chemistry* **1982**, *257* (16), 9759–9769. https://doi.org/10.1016/S0021-9258(18)34138-3.

(48) 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.

(49) RayChaudhuri, D.; Park, J. T. A Point Mutation Converts Escherichia Coli FtsZ Septation GTPase to an ATPase. *Journal of Biological Chemistry* **1994**, *269* (37), 22941–22944. https://doi.org/10.1016/S0021-9258(17)31600-9.

(50) Park, Y. S.; Sweitzer, T. D.; Dixon, J. E.; Kent, C. Expression, Purification, and Characterization of CTP:Glycerol-3-Phosphate Cytidylyltransferase from Bacillus

Subtilis. *Journal of Biological Chemistry* **1993**, *268* (22), 16648–16654. https://doi.org/10.1016/S0021-9258(19)85467-4.

(51) Corbella, M.; Pinto, G. P.; Kamerlin, S. C. L. Loop Dynamics and the Evolution of Enzyme Activity. *Nat Rev Chem* **2023**, *7* (8), 536–547. https://doi.org/10.1038/s41570-023-00495-w.

(52) Narunsky, A.; Kessel, A.; Solan, R.; Alva, V.; Kolodny, R.; Ben-Tal, N. On the Evolution of Protein–Adenine Binding. *Proceedings of the National Academy of Sciences* **2020**, *117* (9), 4701–4709. https://doi.org/10.1073/pnas.1911349117.

Chapter 5

Biosynthesis of FAD analogues and their role in cell physiology

5.1 Introduction

Oxidative stress arises from an imbalance between oxidant and antioxidant levels, arising from the persistent elevation in Reactive Oxygen Species (ROS) generation, which induces chemical alterations in various macromolecules, including RNA, DNA, proteins, and lipids through oxidation.[1] These ROS have the capability to degrade cell membranes, inhibit the activity of crucial enzymes, slow down essential cellular processes, hinder regular cell division, and disrupt energy production.[2] However, the cell employs enzymatic and non-enzymatic antioxidants to fight back against the oxidative stress by ROS, as glutathione peroxidase and catalase are enzymatic antioxidants, while vitamin E, vitamin C, and glutathione (GSH) are non-enzymatic antioxidants.[2] in oxidative conditions, the oxidized glutathione (GSSG) tends to increase intracellularly, and the GSH/GSSG ratio serves as a reliable indicator of an organism's oxidative stress level. Glutathione plays a protective role against oxidative stress through multiple mechanisms, where it acts as a substrate for various detoxifying enzymes, aids in the transportation of amino acids, directly neutralizes singlet oxygen and hydroxyl radicals, and facilitates the active state renewal of vitamins C and E.[3]

The efficacy of the glutathione-based defense system against oxidative stress markedly relies upon the presence and function of flavins where the glutathione reductase employs FAD as a cofactor for replenishing intracellular GSH levels through the GSSG reduction and the NADPH acts as the electron donor.[4] Additionally, we have also established the FAD nucleobase analogues (FCD, FUD, FGD, and dFAD) as active cofactor for *Escherichia coli* glutathione reductase.[5] Our study and the dependency of GSH on FAD for maintaining the redox state of cells led us to explore the impact of these analogues on cellular physiology under oxidative stress.

Chapters 3 and 4 show that the heterologous expression of *Mj*FMNAT and the mutants of *Ec*FADS can synthesize the FAD analogues inside the cell. However, their effect on FAD production and cell physiology cannot be studied directly without having the controls, as the mutants also contribute directly to an increase in the FAD concentration of the cell.

This chapter shows the genomic swapping of the *Ec*FADS expressing essential gene, *ribF*, with the LoopSwap mutant gene, *ribF25*. The *ribF25* mutant strain survives and grows like the wild-type strain, producing FAD analogues (FGD, FCD, and FUD). The oxidative stress generated through H_2O_2 did not create any difference between the mutant and wild-type strain. Given that antibiotics are also recognized as contributors to oxidative stress, our investigation involved analyzing the growth patterns of both strains under the influence of diverse antibiotic-induced stressors.[6] We have shown that the mutant strain can survive better under the antibiotic stress of aminoglycosides and β-lactams than the wild-type strain. When grown anaerobically, the mutant strain has slow growth because of the low concentration of FAD produced compared to the wild-type strain. However, the mutant strain can still respond better under the stress of the streptomycin antibiotic. Other findings also show that an elevated presence of some other internal metabolites, aiding in oxidative stress, such as hydrogen sulfide, indole, nitric oxide, and gaseous ammonia, is linked to bacterial resistance against antibiotics.[7–11]

5.2 Materials and Methods

5.2.1 Materials

Jena Biosciences supplied the order of 2'-deoxy adenosine triphosphate (dATP) and all nucleoside triphosphates (NTPs). Reagents and chemicals for the enzymatic assays were purchased from RANKEM Chemicals, TCI Chemicals, and Merck. All reagents and components of media used in bacterial cultures were ordered from TCI Chemicals, HiMedia, and SRL Chemicals. DSS Takara Bio USA supplied Taq polymerase, Dpn1, and the PrimeSTAR GXL DNA polymerase. Purification kits for DNA and plasmid were obtained from Agilent and Qiagen. We considered Merck, SD Fine Chemicals, and RANKEM for all solvents ordering to use in LC-MS and HPLC analysis.

5.2.2 Homologous recombination of *Ec*FADS-kan and its LoopSwap-kan mutant

Both E. coli K-12 MG1655 *ribF-kan* and *ribF25-kan* mutant strains were generated using recombination by the λ Red recombineering system.[12] *Ec*FADS or LoopSwap mutant expressing genes (*ribF* and *ribF25,* respectively) were amplified from pET28a(+) plasmid using following primers (forward primer 5'-

GACCGCTGTACAAGGTATACTCGGACGATTTTCACTGTTTTGAGCCAGACATGAAG CTGATACGCGGCATACATA-3' and reverse primer 5'-ACTTCGAAGCAGCTCCAGCCTACACTTAAGCCGGTTTTGTTAGCCCAAAA-3'), where forward primer has 50 extra complementary bases for *E. coli* genome downstream to *ribF* gene and the reverse primer has 25 extra complimentary bases for pKD4 vector. Further, the gene was directly stitched to *kan* gene (having FRT sites) from pKD4 by taking the amplified *ribF* or *ribF25* gene as forward primer and the other primer (5'-TCACTCATCAGATTCTCGGTTCCGTATTTCGGTTTGATTACATAACAGGCATATGAA TATCCTCCTTAGTTCCTA-3') which has 50 complementary bases to the *E. coli* genome upstream to the *ribF* gene.

5 mL of *E. coli* K-12 MG1655 transformed with pKD46 (which expresses λ red recombinase and is temperature sensitive) was grown overnight at 28 °C with ampicillin (100 µg/mL) in Luria Bertani (LB) media. 30 mL of secondary culture was grown with 1% overnight grown primary culture, 100 µg/mL Ampicillin, and 0.2 % arabinose in LB media for 3 h at 28 °C (till OD_{600} reached 0.3-0.5). The cells were pelleted down by centrifugation (4000 rpm for 10 minutes) followed by resuspension in 10 mL of autoclaved milli-Q water already kept on ice. Then, the cells were again centrifuged, followed by resuspension in 600 µL of autoclaved chilled milli-Q water. Then, the resuspended cells were transformed with the stitched fragment gene (*ribF-kan* or *ribF25-kan* gene) by electroporation. The electroporated cells were grown at 37 °C for 5 h in 1 mL LB media, followed by growth on an LB agar plate with 35 µg/mL kanamycin antibiotic. The obtained colonies were screened by colony PCR using the following primers (forward primer 5'-CTCAATCGCCGGTTAACCTT-3' and reverse primer 5'-CGGCAAATTCAGGGTTGATTTATAG-3'), and the amplified fragment was confirmed by sequencing.

5.2.3 LC-MS analysis for FAD analogues in the lysate of *ribF-kan* and its *ribF25-kan* mutant strain

Inoculation of 50 mL LB broth was done with overnight grown primary culture of both the strains, and the culture was grown with kanamycin (35 µg/mL) at 37 °C and 180 rpm. After 14-16 h, the culture was centrifuged at 5500 rpm for 15 minutes at 4 °C resulting in a pellet. 2 mL of 90 % methanol was used for the resuspension of the pellet followed by sonication at 100 % amplitude for 3 minutes with 1s pulse on and 1s pulse off. The final supernatant was obtained

by centrifuging the lysate at 4 °C at 5500 rpm for 15 minutes, followed by the filtration via 0.22-micron filters. The supernatant was kept at a high vacuum to concentrate it directly for analysis on LC-MS.

5.2.4 Aerobic growth of *ribF-kan* and its *ribF25-kan* mutant strain under antibiotic stress

The primary culture for each strain was grown overnight from glycerol stock in the presence of the 35 μg/mL kanamycin in LB media at 37 °C, 180 rpm. Then, in LB media, 200 μL secondary culture was grown under the stress of various antibiotics with varying antibiotic concentrations at a starting OD_{600} of 0.1 in a 96-well plate equipped with a lid, and the plate was incubated aerobically at 37 °C, with ∼300 rpm orbital shaking, for 24 to 48 h. OD_{600} readings were taken after every shaking cycle of 1 hour inside a BMG Labtech CLARIOstar® Plus microplate reader.

5.2.5 Flipping out of *kan* cassette from *ribF-kan* and its *ribF25-kan* mutant strain

The *kan* cassette was removed by transforming the mutant strain with PCP20, which has yeast Flp recombinase known to detect the FRT sites. Both the strains were grown in 5 mL LB media till OD_{600} reached 0.4 - 0.5. The cultures were centrifuged (4000 rpm for 10 minutes) at 4 °C, followed by discarding the supernatant. The pellet was again resuspended in 0.1 M calcium chloride (3 mL) and kept at 4 °C for 10 minutes. The culture was again centrifuged at 4 °C, and the supernatant was discarded, followed by resuspending it in 0.1 M calcium chloride (1 mL). This time, the resuspended cells were kept at 4 °C in calcium chloride for 30 minutes to make the cells chemically competent. These competent cells were transformed with PCP20 plasmid by heat sock at 42 °C for 1 minute, and then the addition of 0.3 mL LB media was done. The transformed cells were kept shaking at 30 °C for 4 hours, and the grown cells were spread on LB agar with ampicillin antibiotic (100 μg/mL) and kept at 30 °C till colonies were observed on the plate. To remove the temperature-sensitive PCP20 vector from the strains, the obtained colony was grown in LB media at 42 °C for 5-6 hours. The culture was diluted 10^{-6} times, and around 100 μL was spread on LB agar, incubating overnight at 37 °C. Four colonies were picked up and streaked on four different LB agar plates as follows: (i) only LB agar plate at 37 °C, (ii) LB agar plate with ampicillin (100 μg/mL) at 30 °C, (iii) LB agar plate with kanamycin (35 μg/mL) at 37 °C, and (iv) LB agar plate with ampicillin (100 g/mL) and

kanamycin (35 μg/mL) at 30 °C. Only the LB agar plate without antibiotics had the grown colonies, confirming the removal of the *kan* cassette and PCP20 vector.

5.2.6 Aerobic growth of MG1655 and its *ribF25* mutant strain under antibiotic stress

Primary culture was grown overnight from glycerol stock in LB media at 37 °C, 180 rpm. Then, in LB media, 200 μL secondary culture was grown under various antibiotic stresses of varying antibiotic concentration at a starting OD_{600} of 0.1 in a 96-well plate equipped with a lid, and the plate was incubated aerobically at 37 °C, with ~300 rpm orbital shaking, for 24 to 48 h. OD_{600} readings were taken after every shaking cycle of 1 hour inside a BMG Labtech CLARIOstar® Plus microplate reader.

5.2.7 Anaerobic growth of MG1655 and its *ribF25* mutant strain under antibiotic stress

The primary culture of each strain was grown overnight from glycerol stocks in LB media at 37 °C, 180 rpm. The anaerobic secondary culture was done inside an anaerobic glove bag chamber (~ 170 ppm oxygen level), where 2 mL secondary culture was grown in LB media at 37 °C with 1 % primary culture added under streptomycin antibiotic stress of varying concentrations. 200 μL of the same culture was aliquoted in a 96-well plate with a lid, and the plate was incubated for aerobic growth at 37 °C, with ~200 rpm shaking as for comparison. The OD_{600} of aerobic and anaerobic cultures was recorded after each 12-hour using a BMG Labtech CLARIOstar® Plus microplate reader.

5.2.8 LC-MS

A Mass spectrometer (ScieX X500R-QTOF) was employed for LC-MS analyses, where a UHPLC (Exion-LC series) was used to separate the flavins. The flavins were detected in negative mode at a temperature of 400 °C. The parameters are as follows: ion spray voltage was -4500 V, the de-clustering potential was -80V and collision energy was 5V. The ScieX-OS Explorer software facilitated data acquisition, processing, and the generation of metrics such as extracted ion chromatogram (EIC), total ion chromatogram (TIC), and the m/z values of targeted molecules. The analysis used a Phenomenex Gemini reversed-phase column of 250 x 4.6 mm dimensions with a 5 μm particle size of NX-C18. The solvent-A was ammonium

acetate buffer (10 mM, pH 6.0), and methanol was used as solvent-B. A method spanning 30 minutes was used, starting with 5% of solvent B at 0 minutes, maintaining this ratio until 2.5 minutes, followed by a gradient increase to 80% of solvent B until 20 minutes, then a return to 5% solvent B until 27 minutes, and finally, maintaining at 5% solvent B until 30 minutes.

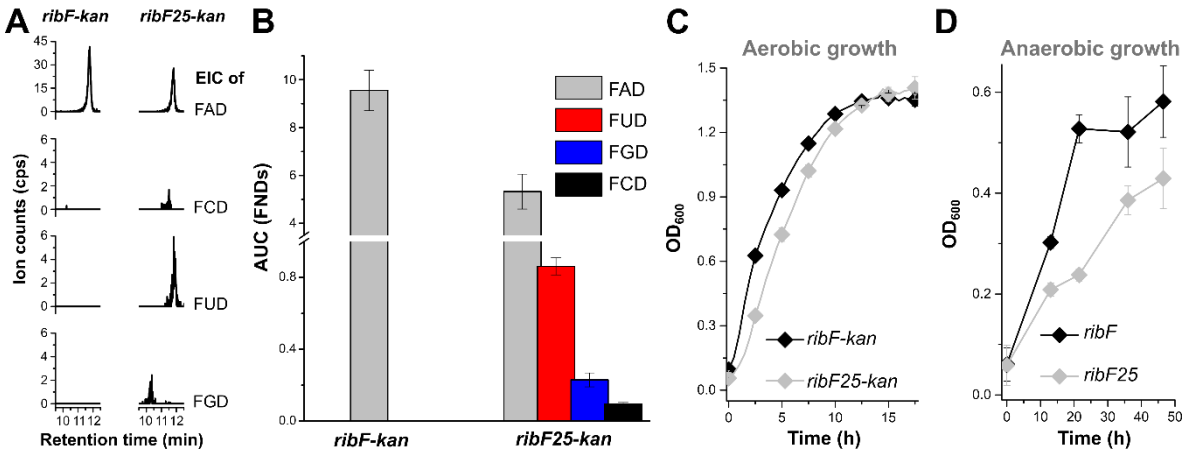


Figure 5.1 Biosynthesis of FAD analogues by *ribF25-kan* mutant strain and the growth assay. (A) EICs of FAD and its analogues in the *ribF-Kan* and *ribF25-kan* mutant strain grown in minimal media. The *ribF-Kan* mutant strain is efficient in the biosynthesis of FAD analogues: FUD, FCD, and FGD. **(B)** The area under the curve **(**AUC) was extracted for FNDs in both the strains, and these AUCs were extracted from EICs where each peak was normalized to OD_{600} = 1 and volume of culture = 1 mL. The mutant strain produces lesser FAD than the wild-type strain. **(C)** Aerobic growth of *ribF-kan* and *ribF25-kan* mutant strain. Swapping the *ribF* gene with the *ribF25* mutant gene does not affect the growth, where the lower FAD production and biosynthesis of FAD analogues make no difference in the growth. **(D)** Yet, the growth difference between wild-type *ribF* (*E. coli* K-12 MG1655) and mutant *ribF25* strain can be observed due to the lower concentration of FAD in the mutant strain.

5.3 Results and Discussion

5.3.1 LoopSwap gene mutant is responsible for FAD nucleobase analogues and grows as wild type.

As shown in Chapter 4, the LoopSwap mutant has the potential to produce the FAD analogues inside the cell along with FAD; the *Ec*FADS expressing essential gene *ribF* in the genome can be swapped by this LoopSwap expressing gene *ribF25*. Hence, the homologous recombination was done to replace *ribF* with the *ribF25* mutant gene linked with the *kan* cassette, resulting in *ribF25-kan* strain. The *kan* cassette was also inserted upstream to the *ribF* gene in the wild-type strain, similar to the mutant strain, resulting in the *ribF-kan* strain being used as a control.

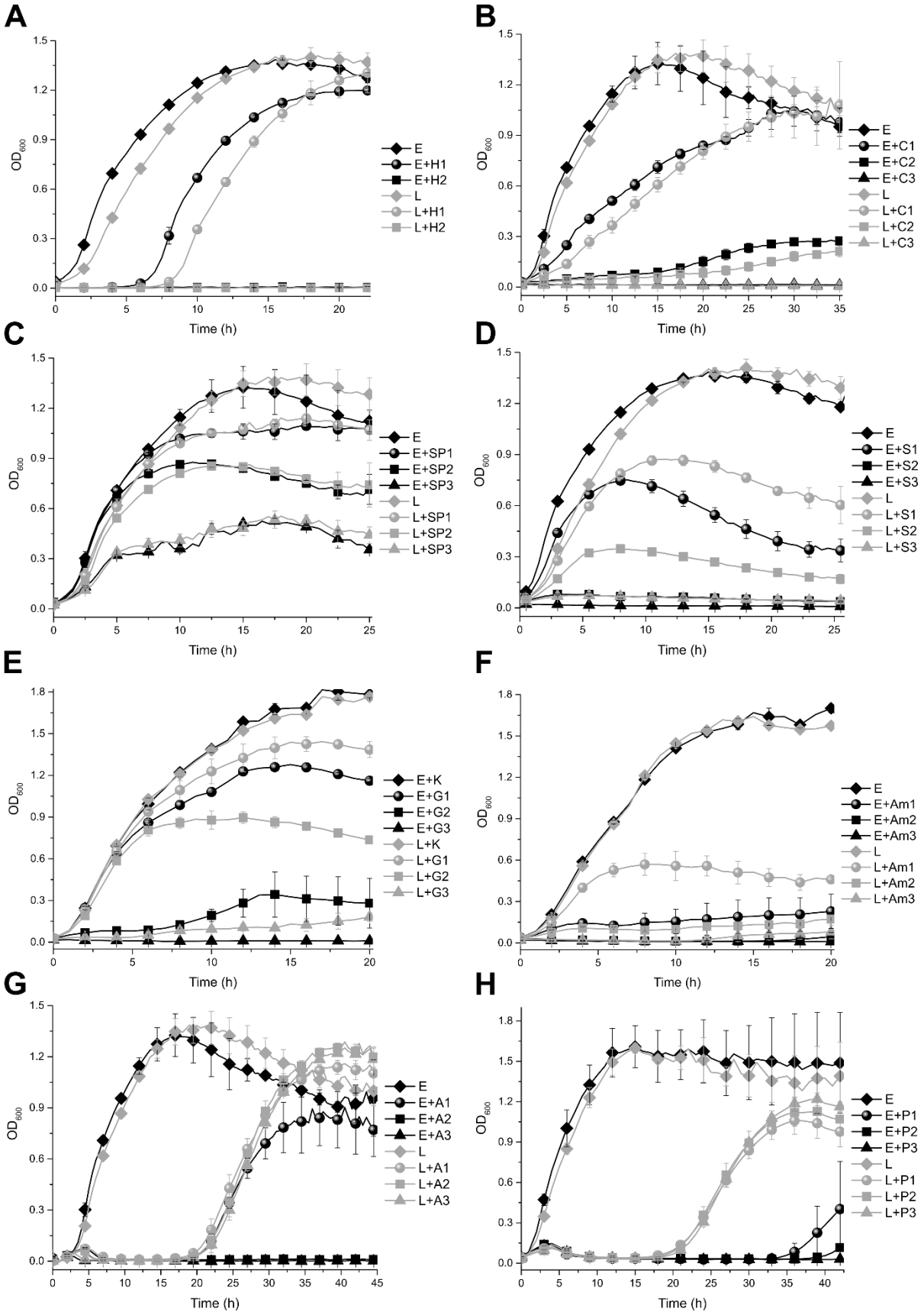


Figure 5.2 Growth of *ribF25-kan* mutant strain under stress of various antibiotics. Where, E= *ribF-kan* control strain, L= *ribF25-kan* mutant strain, **(A)** H= H_2O_2 (H1= 2 µM, H2= 5 µM) **(B)** C= chloramphenicol (C1= 1, C2= 2, C3= 6 µg/mL), **(C)** SP= spectinomycin (SP1= 2, SP2= 4, SP3= 10 µg/mL), **(D)** S= Streptomycin (S1=

5, S2= 10, S3= 15 µg/mL), **(E)** G= gentamycin (G1= 0.5, G2= 2, G3= 5 µg/mL), **(F)** Am= amikacin (Am1= 5, Am2= 10, Am3= 15 µg/mL), **(G)** A= Ampicillin (A1= 5, A2= 10, A3= 20 µg/mL), and **(H)** P= piperacillin (P1= 10, P2= 15, P3= 20 µg/mL). The *ribF25-kan* mutant strain has antibiotic tolerance to aminoglycosides (such as streptomycin, gentamycin, and amikacin) and β-lactams (such as ampicillin and piperacillin) compared to the *ribF-kan* strain. However, the mutant strain is ineffective in some aminoglycosides, such as chloramphenicol and spectinomycin.

Interestingly, in LB media, this mutant strain survives and grows similarly to the original control strain, as it can produce enough FAD to grow (Figure 5.1). Moreover, when the lysate of the mutant and the control strain was studied by LC-MS, FAD analogues (FGD, FCD, and FUD) were observed, suggesting that this mutant strain could also do the biosynthesis of these analogues along with FAD and does not lose its viability (Figure 5.1A). We completely grew both strains in minimal media and studied their lysate on LC-MS to quantify the FAD and its analogues in the mutant strain. The EICs of each FND from both strains were normalized to OD_{600}= 1 and volume= 1 mL (Figure 5.1B). When the area under the curve (AUC) was extracted for each FND, FAD formed by the *ribF25-kan* mutant strain was lesser than FAD in the *ribF-kan* control strain (Figure 5.1B). This lower amount of FAD in the mutant strain does not affect its growth aerobically compared to the wild-type strain, while in the anaerobic conditions, the mutant strain has slower growth than the wild-type (Figure 5.1).

5.3.2 FAD nucleobase analogues aid the cell in antibiotic tolerance

After confirming the presence of these FAD nucleobase analogues in the *rib25-kan* mutant strain, we focused on the effect of these analogues in the cells under oxidative stress, as FAD is important for maintaining the redox chemistry of the cell. We created oxidative stress by treating both the strains with H_2O_2, and we observed that the mutant strain did not show any difference in growth compared to the wild type in the presence of H_2O_2; however, both strains struggled to survive under this stress (Figure 5.2A). As antibiotics are also known to cause oxidative stress, we studied the growth of both strains under the stress of various antibiotics.[6] Both strains have similar growth in the presence of chloramphenicol and spectinomycin (Figure 5.2B and C). However, the other aminoglycoside antibiotics (streptomycin, gentamycin, amikacin) are more effective for the *ribF-kan* control strain, while the mutant strain grows better and can survive at higher concentrations of the same antibiotic(Figure 5.2D-F). The same phenomenon was observed for the β-lactam class of antibiotics (ampicillin and piperacillin),

where the control strain survives at a low concentration of antibiotic while the mutant strain still survives at a higher concentration of antibiotic for which the control strain dies (Figure 5.2G-H). In the case of the beta-lactam antibiotic, a considerable lag of 20-25 h is observed, and then the strain starts growing.

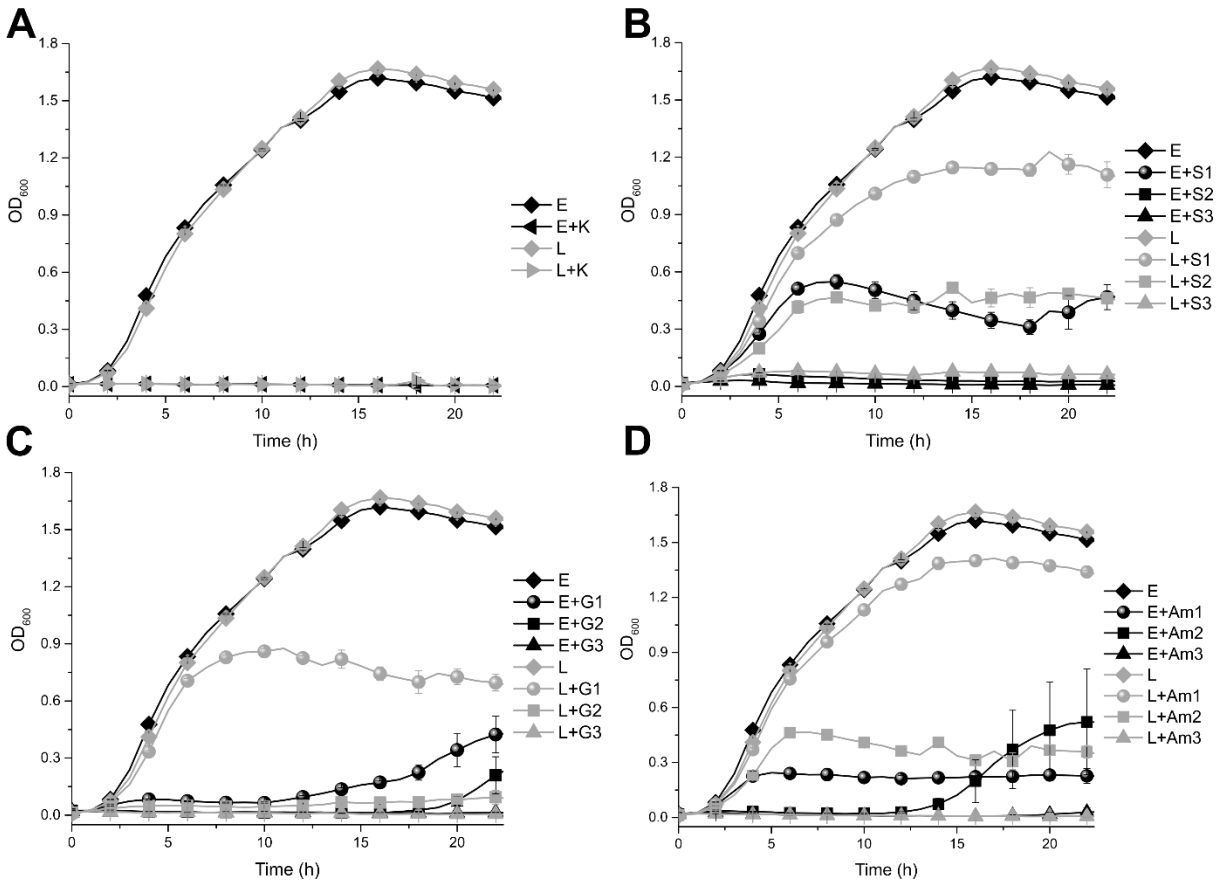


Figure 5.3 Growth of *ribF25* mutant strain under stress of various antibiotics. Where E= wild-type MG1655, L= *ribF25* mutant strain, **(A)** K= 35 µg/mL kanamycin, **(B)** S= Streptomycin (S1= 5, S2= 10, S3= 15 µg/mL), **(C)** G= gentamycin (G1= 1, G2= 2.5, G3= 5 mg/L), **(D)** Am= amikacin (Am1= 2.5, Am2= 5, Am3= 15 mg/L). *ribF25* mutant strain also has antibiotic tolerance to aminoglycosides such as streptomycin, gentamycin, and amikacin compared to the *ribF-kan* strain.

5.3.3 Flipping out the *kan* cassette, followed by the aerobic growth of wild-type MG1655 and rib25 mutant strain under antibiotic stress

The *kan* gene from PKD4 plasmid expresses the o-phosphotransferase to phosphorylate the kanamycin for antibiotic resistance.[13] To cancel out the possibility that the FAD analogues and the *kan* gene are cumulatively responsible for the antibiotic tolerance against the antibiotic stress by aminoglycoside such as streptomycin, we knocked out the kanamycin gene from *ribF-*

kan and *ribF25-kan* mutant strain by using Flp recombinase. Both the kan-knocked-out strains, MG1655 and *ribF25* mutant, do not grow in the presence of an optimal kanamycin concentration, confirming the flipp out of the *kan* cassette (Figure 5.3A). Moreover, the *ribF25* mutant strain still grows better under the stress of aminoglycoside antibiotics such as streptomycin, gentamycin, and amikacin than the wild-type MG1655 strain (Figure 5.3B-D). These results indicate that the *kan* gene has no role in the antibiotic tolerance of the *ribF25* mutant strain for aminoglycosides.

5.3.4 Anaerobic growth of MG1655 and rib25 mutant strain under antibiotic stress

As in anaerobic conditions, FAD is used excessively; the *ribF25* mutant strain has slow growth anaerobically as it produces lesser FAD than the wild-type strain (Figure 5.4A). The low concentration of streptomycin (2 µg/mL) does not affect the growth of the mutant strain, and the growth decreases on increasing the streptomycin (7.5 µg/mL) in the culture (Figure 5.4A). However, the decrease in growth of the wild-type strain is higher than the mutant strain. When the same strain from the anaerobic culture was grown aerobically, the growth of the mutant and wild-type was the same, and the growth of the mutant strain was better in the presence of streptomycin (Figure 5.4B). These results confirm that the mutant strain also has antibiotic tolerance in aerobic and anaerobic conditions and suggest that oxygen has no role in the tolerance to aminoglycoside antibiotics such as streptomycin.

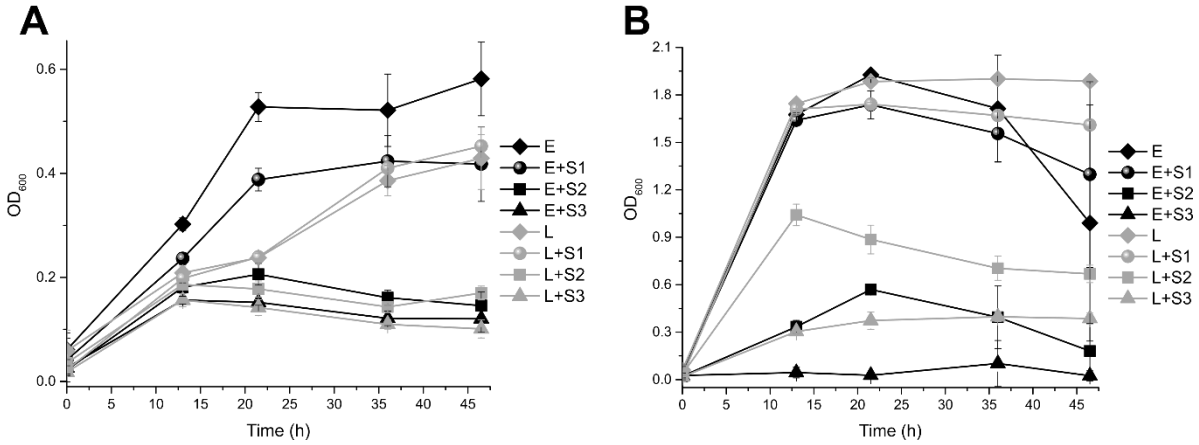


Figure 5.4 Comparison of aerobic and anaerobic growth of *ribF25* mutant strain under stress of streptomycin antibiotic. Where, E= wild-type MG1655, L= *ribF25* mutant strain, S= Streptomycin (S1= 2, S2= 7.5, S3= 20 µg/mL. **(A)** Anaerobic growth of *ribF25* mutant strain. The mutant strain has slow growth in comparison to the wild-type strain. However, the streptomycin does not affect the mutant strain as much as it

affects the wild-type strain. **(B)** Aerobic growth of *ribF25* mutant strain. The mutant strain grows similarly to the wild-type strain when grown aerobically. However, streptomycin still sustains its effects on the wild-type strain more than the mutant strain.

5.4 Conclusion

To understand the true potential of this LoopSwap mutant compared to the wild type inside a cell, we incorporated this mutant gene into the genome of K-12 MG1655 in place of the wild-type gene *ribF*. This *ribF* is an essential gene, and its knockout is not achievable. However, mutations in this gene are achievable as the *ribF25* mutant strain survived due to its capability to synthesize FAD. Interestingly, the mutant strain also has the potential to produce the FAD analogues, which can have potential applications in synthetic biology. The other interesting observation is that the FAD produced by the *ribF25* mutant strain was observed to be lesser than the FAD in the wild-type strain, and aerobically, the mutant strain still grows similarly to the wild-type strain. One of the possibilities for this similar growth can be attributed to the utilization of FAD analogues by the cell to aid FAD. However, the mutant strain could not compete with the wild-type strain in anaerobic conditions due to lesser FAD production. We envision employing nucleobase analogues of FAD as tools to investigate cell metabolism and as distinct catalysts to explore particular metabolic processes related to FAD. This involves modifying flavoenzymes to preferentially use a FAD nucleobase analogue instead of the regular FAD, allowing us to scrutinize specific metabolic stages, which has been shown in the case of NCD analogue for NAD-utilizing enzymes.[14–18]

This study also explored whether this LoopSwap mutant can benefit the cell in oxidative stress as FAD and its analogues possess redox properties. No difference was observed in the growth of the mutant strain relative to the wild-type strain in the presence of oxidative stress by H_2O_2. However, antibiotics are also known to create oxidative stress. When grown under antibiotic stress, we observed that the mutant strain grows better under the stress of some aminoglycoside antibiotics like streptomycin, gentamycin, and amikacin compared to wild-type strain. This mutant strain also has better growth than the wild-type strain in the presence of β-lactams like ampicillin and piperacillin. However, a 20-25 h lag in growth can be seen in the case of β-lactam antibiotic stress. The rise and dissemination of bacteria that are resistant to antibiotics pose a significant danger to human well-being and may also profoundly affect the welfare and continuation of other organisms.[19,20] Conventional methods for creating antibacterial agents rely on blocking crucial mechanisms associated with antibiotic resistance, which makes it

important to understand how unnatural molecules, such as FAD nucleobase analogues, aid antibiotic resistance, which can help increase antibiotics' efficacy.

The bacteria can show various types of defences towards antibiotic presence, such as ribosomal mutation, ribosome modification by methyltransferases, aminoglycoside modifying enzymes, and the active transport of aminoglycoside out of cells through efflux pumps.[21] However, the mechanism of antibiotic tolerance through the FAD nucleobase analogues remains an open question. There are many ways by which these aminoglycoside modifying enzymes can change the structure of aminoglycosides, like o-phosphorylation, adenylation, or acetylation.[22] We have listed some of the direct and indirect possibilities by which FAD analogues might be aiding to the antibiotic tolerance.

(I) since the *kan* gene translates an o-phosphotransferase, which utilizes ATP as phosphor transfer to phosphorylate kanamycin. The possibility that o-phosphotransferase may also modify the other aminoglycoside with the aid of these FAD analogues was ruled out by flipping the *kan* gene out from the genome. The antibiotic tolerance of the mutant strain persisted after knocking out the *kan* gene, and the growth of the mutant strain remained the same as the wild-type strain.

(II) The other possibility is that the FAD analogues may take part in some type of aminoglycoside modification, such as N-formimidoylation/-iminoacetylation of aminoglycosides carried with the help of the FAD cofactor similar to the N-formimidoylation of glycinothricin by FAD.[23] However, this reaction requires molecular oxygen, which terminates the possibility of N-formimidoylation of aminoglycosides by FAD analogues, as the mutant strain had a lesser effect on its growth of the presence of streptomycin as compared to the MG1655 strain.

(III) LoopSwap mutant might directly interact with the aminoglycoside to nullify the antibiotic effect aiding in antibiotic tolerance.

(IV) Since, aminoglycosides cause the antibiotic stress by binding to ribosome units and interrupting protein synthesis, these FAD analogues might be interacting with protein, or a regulator involved in protein synthesis machinery to aid in antibiotic tolerance against aminoglycosides.

(V) The FAD analogues might be triggering the regulation of the effector molecules such as guanosine-3', 5'-bis(diphosphate) (ppGpp) and guanosine-5'-triphosphate 3'-diphosphate (pppGpp) which have been known to regulate several bacterial processes such as antimicrobial resistance, persistence, and virulence.[24]

There are various ways by which a bacteria attain antibiotic resistance or tolerance. There is no direct relation of FAD known to any mechanism by which antibiotic resistance can be attained. We are collaborating with Dr. Nishad Matange, IISER Pune who is working in the field of antibiotic resistance, and we are also collaborating with Dr. Arnab Mukherjee, IISER Pune who works in bioinformatics. Here we are looking for the docking of FAD and its analogues with all proteins present in *E. coli* which will provide the potential targets by which these FAD analogues aiding the cell in antibiotic tolerance. We are also working on RNA sequencing of both the strains to map the gene regulation in both strain and figuring out the possible genes responsible for the antibiotic tolerance. The other study we planned to work on is to look at the metabolite difference in both the strains and looking at the potential molecules related to the antibiotic tolerance. To rule out the possibility of LoopSwap mutant directly interacting with the antibiotic, we planned to do activity assays for EcFADS and LoopSwap mutant in presence of various concentrations of antibiotic and seeing if the reaction is hindered or not.

In summary, we have created a new MG1655 mutant strain that can do biosynthesis of FAD analogues and grow similarly to the wild-type strain in aerobic conditions. However, the difference in growth can be seen anaerobically. Besides this, the new mutant strain is tolerant towards the antibiotic stress of various aminoglycosides and β-lactams.

References

(1) Markkanen, E. Not Breathing Is Not an Option: How to Deal with Oxidative DNA Damage. *DNA Repair (Amst)* **2017**, *59*, 82–105. https://doi.org/10.1016/j.dnarep.2017.09.007.

(2) Kurutas, E. B. The Importance of Antioxidants Which Play the Role in Cellular Response against Oxidative/Nitrosative Stress: Current State. *Nutr J* **2015**, *15* (1), 71. https://doi.org/10.1186/s12937-016-0186-5.

(3) Birk, J.; Meyer, M.; Aller, I.; Hansen, H. G.; Odermatt, A.; Dick, T. P.; Meyer, A. J.; Appenzeller-Herzog, C. Endoplasmic Reticulum: Reduced and Oxidized Glutathione Revisited. *J Cell Sci* **2013**, *126*, 1604–1617. https://doi.org/10.1242/jcs.117218.

(4) Beutler, E. Effect of Flavin Compounds on Glutathione Reductase Activity: In Vivo and in Vitro Studies. *Journal of Clinical Investigation* **1969**, *48* (10), 1957–1966. https://doi.org/10.1172/JCI106162.

(5) 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**, *24* (11), e202300055. https://doi.org/10.1002/cbic.202300055.

(6) Dwyer, D. J.; Belenky, P. A.; Yang, J. H.; MacDonald, I. C.; Martell, J. D.; Takahashi, N.; Chan, C. T. Y.; Lobritz, M. A.; Braff, D.; Schwarz, E. G.; Ye, J. D.; Pati, M.; Vercruysse, M.; Ralifo, P. S.; Allison, K. R.; Khalil, A. S.; Ting, A. Y.; Walker, G. C.; Collins, J. J. Antibiotics Induce Redox-Related Physiological Alterations as Part of Their Lethality. *Proceedings of the National Academy of Sciences* **2014**, *111* (20), E2100–E2109. https://doi.org/10.1073/pnas.1401876111.

(7) Gusarov, I.; Shatalin, K.; Starodubtseva, M.; Nudler, E. Endogenous Nitric Oxide Protects Bacteria Against a Wide Spectrum of Antibiotics. *Science (1979)* **2009**, *325* (5946), 1380–1384. https://doi.org/10.1126/science.1175439.

(8) Lee, H. H.; Molla, M. N.; Cantor, C. R.; Collins, J. J. Bacterial Charity Work Leads to Population-Wide Resistance. *Nature* **2010**, *467* (7311), 82–85. https://doi.org/10.1038/nature09354.

(9) Shatalin, K.; Shatalina, E.; Mironov, A.; Nudler, E. H_2S: A Universal Defense Against Antibiotics in Bacteria. *Science (1979)* **2011**, *334* (6058), 986–990. https://doi.org/10.1126/science.1209855.

(10) Vitko, N. P.; Spahich, N. A.; Richardson, A. R. Glycolytic Dependency of High-Level Nitric Oxide Resistance and Virulence in Staphylococcus Aureus. *mBio* **2015**, *6* (2), e00045-e15. https://doi.org/10.1128/mBio.00045-15.

(11) Cohen, N. R.; Ross, C. A.; Jain, S.; Shapiro, R. S.; Gutierrez, A.; Belenky, P.; Li, H.; Collins, J. J. A Role for the Bacterial GATC Methylome in Antibiotic Stress Survival. *Nat Genet* **2016**, *48* (5), 581–586. https://doi.org/10.1038/ng.3530.

(12) Thomason, L. C.; Costantino, N.; Li, X.; Court, D. L. Recombineering: Genetic Engineering in *Escherichia Coli* Using Homologous Recombination. *Curr Protoc* **2023**, *3* (2), e656. https://doi.org/10.1002/cpz1.656.

(13) Wright, G. D. Aminoglycoside Phosphotransferases: Proteins, Structure, and Mechanism. *Frontiers in Bioscience* **1999**, *4* (1–3), d9-24. https://doi.org/10.2741/Wright.

(14) Wang, L.; Ji, D.; Liu, Y.; Wang, Q.; Wang, X.; Zhou, Y. J.; Zhang, Y.; Liu, W.; Zhao, Z. K. Synthetic Cofactor-Linked Metabolic Circuits for Selective Energy Transfer. *ACS Catal* **2017**, *7* (3), 1977–1983. https://doi.org/10.1021/acscatal.6b03579.

(15) Guo, X.; Liu, Y.; Wang, Q.; Wang, X.; Li, Q.; Liu, W.; Zhao, Z. K. Non-natural Cofactor and Formate-Driven Reductive Carboxylation of Pyruvate. *Angewandte Chemie* **2020**, *132* (8), 3167–3170. https://doi.org/10.1002/ange.201915303.

(16) Liu, Y.; Feng, Y.; Wang, L.; Guo, X.; Liu, W.; Li, Q.; Wang, X.; Xue, S.; Zhao, Z. K. Structural Insights into Phosphite Dehydrogenase Variants Favoring a Non-Natural Redox Cofactor. *ACS Catal* **2019**, *9* (3), 1883–1887. https://doi.org/10.1021/acscatal.8b04822.

(17) Liu, Y.; Li, Q.; Wang, L.; Guo, X.; Wang, J.; Wang, Q.; Zhao, Z. K. Engineering <scp>d</Scp> -Lactate Dehydrogenase to Favor an Non-natural Cofactor Nicotinamide Cytosine Dinucleotide. *ChemBioChem* **2020**, *21* (14), 1972–1975. https://doi.org/10.1002/cbic.201900766.

(18) 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.

(19) King, A. M.; Reid-Yu, S. A.; Wang, W.; King, D. T.; De Pascale, G.; Strynadka, N. C.; Walsh, T. R.; Coombes, B. K.; Wright, G. D. Aspergillomarasmine A Overcomes Metallo-β-Lactamase Antibiotic Resistance. *Nature* **2014**, *510* (7506), 503–506. https://doi.org/10.1038/nature13445.

(20) Chakradhar, S. Reservoirs of Resistance: To Understand Why Antibiotics Fail, Geneticists Chase the "Resistome." *Nat Med* **2016**, *22* (10), 1069–1071. https://doi.org/10.1038/nm1016-1069.

(21) Garneau-Tsodikova, S.; Labby, K. J. Mechanisms of Resistance to Aminoglycoside Antibiotics: Overview and Perspectives. *Medchemcomm* **2016**, *7* (1), 11–27. https://doi.org/10.1039/C5MD00344J.

(22) Ramirez, M. S.; Tolmasky, M. E. Aminoglycoside Modifying Enzymes. *Drug Resistance Updates* **2010**, *13* (6), 151–171. https://doi.org/10.1016/j.drup.2010.08.003.

(23) Wang, Y.-L.; Chang, C.-Y.; Hsu, N.-S.; Lo, I.-W.; Lin, K.-H.; Chen, C.-L.; Chang, C.-F.; Wang, Z.-C.; Ogasawara, Y.; Dairi, T.; Maruyama, C.; Hamano, Y.; Li, T.-L. N-Formimidoylation/-Iminoacetylation Modification in Aminoglycosides Requires FAD-Dependent and Ligand-Protein NOS Bridge Dual Chemistry. *Nat Commun* **2023**, *14* (1), 2528. https://doi.org/10.1038/s41467-023-38218-w.

(24) Das, B.; Bhadra, R. K. (P)PpGpp Metabolism and Antimicrobial Resistance in Bacterial Pathogens. *Front Microbiol* **2020**, *11*, 563944. https://doi.org/10.3389/fmicb.2020.563944.

Chapter 6

Findings, implications, and future directions

6.1 Understanding the adenine recognition

Ligand binding is one of the important and most common functions for protein evolution to recognise the substrate.[1] To understand the process of protein evolution of recognizing and binding the substrate is a fundamental goal in the field of evolution. One of the ways it can be approached is by focusing on ancient substrates or a specific fragment of such substrate. Understanding various proteins of a common evolutionary origin and their interaction patterns with these substrates can provide insights into the evolutionary development of these interactions. For this instance, Nucleotides (NTPs: ATP, GTP, CTP, UTP) are a common type of ligands in living organisms and differ only based on the nucleobases. NTPs as enzyme cofactors are particularly widespread, taking part in important metabolic reactions. The NTP binding is an example of protein-ligand interactions, which are one of the most conserved and the oldest, possibly due to the prevalence of catalytic RNA molecules in the early stages of life on Earth.[2,3] Adenine moiety from ATP is a fragment of most enzyme cofactors, such as FAD, NAD, SAM, and coenzyme A, which plays a crucial role in protein-nucleotide interactions.[4] Adenine possesses various groups that enable different types of interactions, including electrostatic, hydrogen bonding, aromatic, and nonpolar interactions.[1]

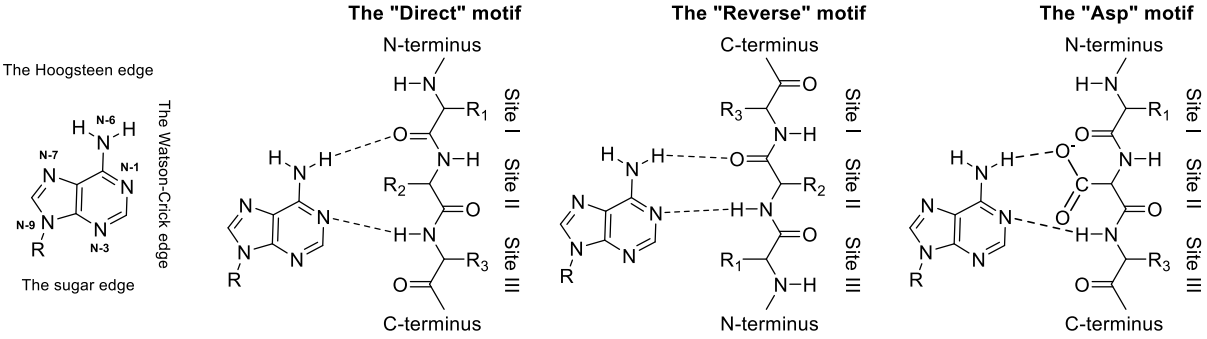


Figure 6.1 The trend of backbone interaction of adenine moiety of ATP and enzyme cofactors to enzyme residues.

Due to the planer triangle-like structure of adenine, it has three edges to participate in hydrogen bonding with its surrounding residues and molecules. These edges are the Watson–Crick edge, the Hoogsteen edge, and the sugar edge (Figure 6.1). Subsequent studies confirmed

that adenine-binding predominantly involves interactions with adenine's Watson–Crick edge.[5–8] The adenine of ATP, SAM, NAD, and FAD was found to be interacting with the protein via three variants of a motif: the "direct motif," the "reverse motif," and the "Asp motif" (Figure 6.1).[4] In the first two variants, the backbone of the protein loop interacts with adenine, implying that these modes may be ancient and driven by structural constraints. A negatively charged amino acid side chain interacts with adenine's N6 in the "Asp motif", which was often observed in NAD-binding proteins (Figure 6.1). Later, another study, based on a larger dataset, also confirmed these patterns, revealing interactions with adenine's all three edges.[9] For this study, 985 protein-nucleotide complexes were used in this large and comprehensive dataset from the Protein Data Bank (PDB) (Table 6.1). ATP preferably binds to the "direct" and "reverse" motif, while SAM and NAD prefer the "Asp" motif, and FAD has a preference for the "reverse" motif. The study concentrated on two primary perspectives of protein-adenine interactions. First, it explored the physicochemical nature of the ligand-binding, focusing on interactions such as the hydrogen bonding of proteins with its ligands as well as the structural motifs supporting the ligand-binding. Secondly, it investigates evolutionary traces remained in the ligand-binding by examining recurring sequences called "themes" in adenine-binding proteins. The study connected the understanding of hydrogen-bonding interactions observed in adenine-binding proteins with evolutionary trends, particularly in terms of their theme composition. Differences in structural geometries, as well as ligand-binding patterns across many proteins, suggested that adenine binding mainly emerged independently at various times in the evolution process. The study also directs towards future research efforts to uncover an evolutionary path that explains how an initial set of short themes resulted in larger "themes" and corresponding binding sites.

Table 6.1 Quantitative measure of occurrence for the "direct," "reverse," and "Asp" variations of the adenine-binding motif. This table is taken from Narunsky *et al*'s work.[9]

Cofactor	The "direct" motif	The "reverse" motif	The "Asp" motif	Other	Total
ATP	77 (19 %)	74 (18 %)	10 (2.5 %)	245 (60 %)	406
FAD	0	128 (63 %)	6 (3 %)	69 (34 %)	203
SAM	1 (0.5 %)	29 (17 %)	107 (61.5 %)	37 (21 %)	174
NAD	0	3 (2.5 %)	42 (34 %)	78 (63.5 %)	123
CoA	2 (2.5 %)	0	0	77 (97.5 %)	79
Total	80	234	165	506	**985**

6.2 The structural fold of adenine binding loop and nucleobase selectivity

In Chapter 4, We established the role of a specific loop (Figure 6.2) in the FMNAT domain of the enzyme *Ec*FADS (*Escherichia coli* FAD synthetase) in relation to its ATP selectivity where LoopSwap mutant broke this selectivity and produced FGD along with FAD, FCD, FUD, and dFAD with better efficiency. The *Ec*FADS is very specific for ATP inside a cell, and when the large loop around the adenine was shortened or swapped with a short loop from another enzyme (glycerol-3-phosphate cytidylyl transferase), the specificity of *Ec*FADS for ATP was lost. The observations suggest that the loop's conformational dynamics might be integral to its role in ATP binding. The loop might undergo changes in its structure or flexibility to accommodate the ATP molecule, and altering the loop's properties could lead to changes in the enzyme's substrate selectivity.[10]

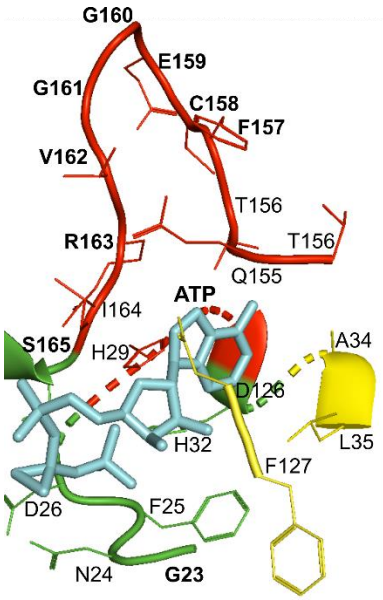


Figure 6.2 The loop structure of the FMNAT domain in *Ec*FADS. It shows the closest residues to ATP, where red residues are near the adenine base, green surround the phosphates, and yellow are close to the ribose ring.

To further understand the role of the loop in the adenine recognition, we looked for the presence of a similar loop in 80 PDB structures randomly taken from the already available published list of 985 crystal structures bounded to ATP, ADP, AMP, CoA, FAD, NAD, SAM and their analogues (Table 4.3).[9] Our findings revealed that 62 % of ATP and its derivatives featured this distinct loop structure, showcasing varying sizes of folding. Additionally, a subset of 4 out of 8 SAH (S-adenosyl-L-homocysteine) molecules also displayed the loop surrounding

the adenine base, with a length spanning 7 to 10 amino acids. On the other hand, for molecules like CoA, FAD, NAD, and their derivatives, interactions with the loop occurred either at a distant fold or involved loop sizes of less than 4 amino acids. Due to the distinct folding observed around ATP and its derivatives compared to other cofactor molecules, there is a possibility that the ligand-binding pattern could have emerged independently on various occasions.

The question of why enzyme cofactors such as FAD, NAD and others possess only adenosine moiety over nucleoside moieties coming from other NTPs is yet to be answered.

6.3 FAD nucleobase analogues in cell physiology

In Chapter 3, we demonstrated the synthesis of FAD nucleobase analogous—FGD, FCD, and FUD. These analogous exhibited comparable activity to that of native cofactors FAD in the enzymatic assay of glutathione reductase. Furthermore, these analogous displayed the capability to be synthesized intracellularly. This study highlights the functional efficacy of these analogous within cellular physiology.

In Chapter 5, we successfully developed a novel MG1655 mutant strain capable of biosynthesizing the FAD analogous while maintaining similar growth rates to the wild-type strain under aerobic conditions. However, distinct growth behaviour between the mutant and wild-type strains can be seen under anaerobic conditions due to the low production of FAD in the mutant strain. Additionally, the new mutant strain exhibits resistance to antibiotic stress induced by various aminoglycosides and β-lactams when grown aerobically and anaerobically. Other findings also show that an elevated presence of some other internal metabolites like hydrogen sulfide, indole, nitric oxide, and gaseous ammonia is linked to a greater occurrence of bacterial resistance to antibiotics.[11–15]

However, the direct or indirect mechanism by which these FAD analogues aid the cell under antibiotic stress is yet to be explored. Can these analogues replace the native substrate FAD and fulfil its place, or are there other roles that these analogues can play in cell physiology?

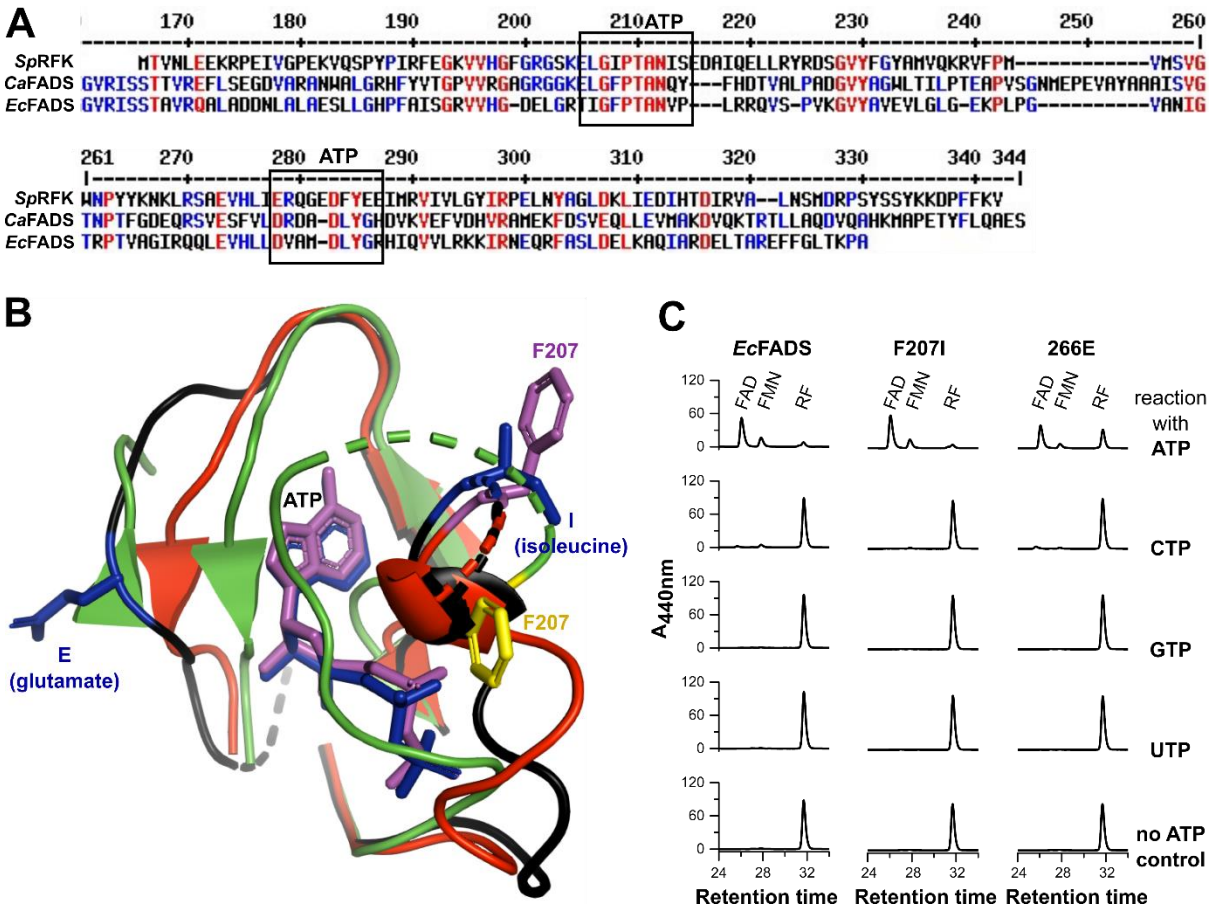


Figure 6.3 Modulating ATP-specificity of RFK module in *Ec*FADS. (A) sequence alignment of *Sp*RFK with *Ec*FADS and *Ca*FADS. RFK module and *Sp*RFK share common ATP binding motifs, while two key differences can be seen in the adenine binding site marked in black boxes. The RFK modules have a large phenylalanine (F) residue, and SpRFK has a small isoleucine (I) residue. Secondly, SpRFK possesses an aspartate (E) residue while absent in others. **(B)** The structural alignment of *Sp*RFK with *Ca*FADS as apoenzyme and ADP-bound enzyme. The ADP-bound *Ca*FDAS possesses a helix near the adenine ring, while the apoenzyme has a flexible loop consisting of the same residues. (C) HPLC chromatograms of activity of mutants with NTPs in comparison to *Ec*FADS.

6.4 Engineering RFK module of *Ec*FADS

The RFK module of *Ec*FADS is ATP-specific, which can be a limiting factor for the low production of FAD nucleobase analogues in the *ribF25* mutant strain. Therefore, we decided to modify the RFK module based on sequence and structural similarities with the *Schizosaccharomyces pombe* riboflavin kinase (*Sp*RFK), which has demonstrated promiscuity in accommodating various NTPs in our lab (Yashwant's thesis). It is worth noting that eukaryotes and prokaryotes share structurally similar RFKs. Upon conducting a sequence alignment, we identified two key amino acid differences in the ATP binding site, particularly

around the nucleobase (Figure 6.3A). Firstly, in the conserved domain GxFTAN, the X residue is phenylalanine in *Ec*FADS and *Ca*FADS, while a smaller residue, isoleucine, is present in *Sp*RFK (Figure 6.3A, B). Secondly, the glutamate residue found in *Sp*RFK is absent in the structures of the other two enzymes (Figure 6.3A, B). As mentioned in Chapter 1, these two residues are likely acting as gatekeeper residues. The smaller size of isoleucine in *Sp*RFK allows other NTPs to enter its active site, and the electrostatic interaction of the aspartate residue with the nucleobase groups of NTPs directs the substrate inside the active site. When we generated mutants with these two changes, namely F207I and 266E, we observed that FAD and FMN were not produced using other NTPs (Figure 6.3C). Additionally, the activity of the 266E mutant was lower compared to the other two mutants. It suggests that these two residues are not responsible for the nucleotide specificity.

However, one more difference we observed is that the *Ca*FADS bound to ADP shows a helix formation near the adenine base while the helix is absent in the apoenzyme (Figure 6.3B). We postulate that the helix disruption can lead to the break of nucleotide specificity of the RFK module in *Ec*FADS. This postulate is supported by two studies in our lab where Yashwant's thesis reports similar helix formation in *Methanocaldococcus jannaschii* RFK, which was disrupted by introducing proline residues in the helix and found the mutant promiscuous. Similarly, Aditya's thesis in our lab also reported a similar helix formation phenomenon in *E. coli* methionine adenylyl transferase (*Ec*MAT).

In future, we would tune the nucleotide specificity of the RFK module in EcFADS by disrupting the adenine binding helix based on SpRFK, followed by a similar mutation in the genome of the ribF25 mutant strain (LoopSwap mutant strain) to enhance the FAD nucleobase analogue in vivo production.

References

(1) Kessel, A.; Ben-Tal, N. *Introduction to Proteins*; Chapman and Hall/CRC, 2018. https://doi.org/10.1201/9781315113876.

(2) Gilbert, W. Origin of Life: The RNA World. *Nature* **1986**, *319* (6055), 618–618. https://doi.org/10.1038/319618a0.

(3) Caetano-Anollés, G.; Wang, M.; Caetano-Anollés, D.; Mittenthal, J. E. The Origin, Evolution and Structure of the Protein World. *Biochemical Journal* **2009**, *417* (3), 621–637. https://doi.org/10.1042/BJ20082063.

(4) Denessiouk, K. A.; Rantanen, V.-V.; Johnson, M. S. Adenine Recognition: A Motif Present in ATP-, CoA-, NAD-, NADP-, and FAD-Dependent Proteins. *Proteins: Structure, Function, and Genetics* **2001**, *44* (3), 282–291. https://doi.org/10.1002/prot.1093.

(5) Cappello, V.; Tramontano, A.; Koch, U. Classification of Proteins Based on the Properties of the Ligand-Binding Site: The Case of Adenine-Binding Proteins. *Proteins: Structure, Function, and Bioinformatics* **2002**, *47* (2), 106–115. https://doi.org/10.1002/prot.10070.

(6) Shulman-Peleg, A.; Nussinov, R.; Wolfson, H. J. Recognition of Functional Sites in Protein Structures. *J Mol Biol* **2004**, *339* (3), 607–633. https://doi.org/10.1016/j.jmb.2004.04.012.

(7) Kuttner, Y. Y.; Sobolev, V.; Raskind, A.; Edelman, M. A Consensus-Binding Structure for Adenine at the Atomic Level Permits Searching for the Ligand Site in a Wide Spectrum of Adenine-Containing Complexes. *Proteins: Structure, Function, and Bioinformatics* **2003**, *52* (3), 400–411. https://doi.org/10.1002/prot.10422.

(8) He, W.; Liang, Z.; Teng, M.; Niu, L. Lib <scp>ME</Scp> —Automatic Extraction of 3D Ligand-binding Motifs for Mechanistic Analysis of Protein–Ligand Recognition. *FEBS Open Bio* **2016**, *6* (12), 1331–1340. https://doi.org/10.1002/2211-5463.12150.

(9) Narunsky, A.; Kessel, A.; Solan, R.; Alva, V.; Kolodny, R.; Ben-Tal, N. On the Evolution of Protein–Adenine Binding. *Proceedings of the National Academy of Sciences* **2020**, *117* (9), 4701–4709. https://doi.org/10.1073/pnas.1911349117.

(10) Corbella, M.; Pinto, G. P.; Kamerlin, S. C. L. Loop Dynamics and the Evolution of Enzyme Activity. *Nat Rev Chem* **2023**, *7* (8), 536–547. https://doi.org/10.1038/s41570-023-00495-w.

(11) Gusarov, I.; Shatalin, K.; Starodubtseva, M.; Nudler, E. Endogenous Nitric Oxide Protects Bacteria Against a Wide Spectrum of Antibiotics. *Science (1979)* **2009**, *325* (5946), 1380–1384. https://doi.org/10.1126/science.1175439.

(12) Lee, H. H.; Molla, M. N.; Cantor, C. R.; Collins, J. J. Bacterial Charity Work Leads to Population-Wide Resistance. *Nature* **2010**, *467* (7311), 82–85. https://doi.org/10.1038/nature09354.

(13) Shatalin, K.; Shatalina, E.; Mironov, A.; Nudler, E. H $_2$ S: A Universal Defense Against Antibiotics in Bacteria. *Science (1979)* **2011**, *334* (6058), 986–990. https://doi.org/10.1126/science.1209855.

(14) Vitko, N. P.; Spahich, N. A.; Richardson, A. R. Glycolytic Dependency of High-Level Nitric Oxide Resistance and Virulence in Staphylococcus Aureus. *mBio* **2015**, *6* (2), e00045-e15. https://doi.org/10.1128/mBio.00045-15.

(15) Cohen, N. R.; Ross, C. A.; Jain, S.; Shapiro, R. S.; Gutierrez, A.; Belenky, P.; Li, H.; Collins, J. J. A Role for the Bacterial GATC Methylome in Antibiotic Stress Survival. *Nat Genet* **2016**, *48* (5), 581–586. https://doi.org/10.1038/ng.3530.