Engineering *Escherichia coli* FAD synthetase for the biosynthesis of FAD nucleotide analogues results in aminoglycoside antibiotic tolerance

Ateek Shah[1#], Prathith Bhargav[1], Shreya Mahato[2], Reman Kumar Singh[1†], Yashwant Kumar[1§], Pratishruti Panda[1], Janhavi Borkar[2], Lavisha Parab[3], Nishad Matange[2], Arnab Mukherjee*[1, 4], Amrita B. Hazra* [1,2]

[1] Department of Chemistry, Indian Institute of Science Education and Research Pune, Dr. Homi Bhabha Road, Pune – 411008, INDIA

[2] Department of Biology, Indian Institute of Science Education and Research Pune, Dr. Homi Bhabha Road, Pune – 411008, INDIA

[3]Department of Bioscience and Biotechnology, Indian Institute of Technology Kharagpur, Kharagpur – 721302, West Bengal, INDIA

[4] Department of Data Science, Indian Institute of Science Education and Research Pune, Dr. Homi Bhabha Road, Pune – 411008, INDIA

Present address:

[#]National Institutes of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892, USA

[†]Department of Chemistry, St. Xavier's College, Dr. Camil Bulcke Path, Ranchi, Jharkhand – 834001, INDIA

[§]Discipline of Chemical Engineering, Indian Institute of Technology Gandhinagar, Gandhinagar – 382355, INDIA

* correspondence to amrita@iiserpune.ac.in , arnab.mukherjee@iiserpune.ac.in

ABSTRACT

The presence of the adenosine moiety in flavin adenine dinucleotide (FAD), nicotimanide adenine dinculeotide, coenzyme A, and S-adenosylmethionine cofactors is intriguing as it appears to play no direct catalytic role. In FAD biosynthesis, the bifunctional FAD synthetase catalyzes FMN adenylation using ATP via its FMN adenylyltransferase (FMNAT) domain. Substituting ATP with alternative nucleotides in this reaction would enable the synthesis of FAD nucleotide analogues. This study explores nucleotide selectivity of the *Escherichia coli* FMNAT domain. Mutant enzyme variants were engineered which produce FAD analogues. Chromosomal substitution of the wild-type *E. coli* FAD synthetase with a mutated version resulted in the intracellular synthesis of FAD analogues and conferred increased tolerance to aminoglycoside antibiotics. These findings validate the FMNAT domain loop as a viable site for cofactor engineering, with potential synthetic biology applications. Finally, the observed correlation between FAD analogue production and antibiotic tolerance reveals that there exist novel metabolic mechanisms underlying bacterial resistance.

INTRODUCTION

Flavin adenine dinucleotide (FAD) is an essential cofactor involved in redox biology across the domains of life. The biosynthesis of FAD originates from the nucleotide metabolism pathway - guanosine triphosphate converts to 5-amino-6-(D-ribitylamino)-uracil via pyrimidine intermediates and ribulose-5-phosphate, forming the pteridine intermediate 6,7-dimethyl-8-ribityllumazine.[1] In the final step, two molecules of the pteridine intermediate are dismutated by riboflavin synthase, forming the isoalloxazine ring that yields riboflavin (vitamin B_2) along with 5-amino-6-(D-ribitylamino)-uracil.[1] Subsequently, riboflavin undergoes phosphorylation by riboflavin kinase to generate flavin mononucleotide (FMN).[2–4] FMN is then adenylated to form FAD by the action of FMN adenylyltransferase (FMNAT) (Figure 1A).[2,3] The redox properties of FAD can be attributed to the isoalloxazine ring whereas the adenosine part is currently viewed as a binding and recognition handle.[5–7]

The adenosine moiety, also found in other cofactors such as nicotinamide adenine dinucleotide (NAD), S-adenosyl methionine (SAM), and coenzyme A, has been considered as a possible remnant of covalently bound coenzymes in an RNA-centric world.[7,8]. Interestingly, hypoxanthine nucleobase analogues of FAD and NAD have been confirmed in the cell extracts of marine bacterium *Desulfobacula toluolica* Tol2, along with acetyl-inosino-CoA and benzoyl-inosino-CoA, the inosine analogues of downstream products of coenzyme A.[9] In FAD, the adenine is found to be at a distance of ~11-16Å from the isoalloxazine ring when bound to flavoproteins, and appears to play no direct role in its catalytic functon.[10,11] Hence, replacing the adenine by a host of alternate ligands such as naturally occurring and non-natural nucleobases have applications in fluorescence quenching and crystallography studies.[12–15] Similarly, novel analogues of SAM and NAD modified at the nucleobase or nucleoside region exist, with applications including imaging and structural probes, antimicrobials, and enzyme inhibitors. For example, fluorescent nucleobase analogues such as DAPSAM have been used

to determine the K_d of the SAM riboswitches.[16] Several NAD analogues are potent inhibitors of the NADase activity of signalling enzyme CD38, with the guanine analogue ara-F-NGD being most effective.[17] Besides this, the cytosine analogue NCD has been used as a synthetic biology tool for driving reductive carboxylation of pyruvate to L-malate, leading to CO_2 fixation.[18] On these lines, FAD nucleobase analogues with guanine, cytosine, uracil and deoxyadenosine, namely FGD, FCD, FUD, and dFAD, are active redox cofactors for glutathione reductase, and can be envisioned as potential bio-orthogonal cofactors for modified enzymes within a cell, similar to NCD.[11]

FAD is produced in the cell from riboflavin via two steps that each consume a molecule of ATP (Figure 1A).[2,3] While the change of nucleotide substrate for riboflavin kinase in the first step should not affect the product outcome as it is a phosphorylation reaction, the incorporation of other nucleotides (GTP/ CTP/ UTP) by the FMNAT will result in FGD, FCD, or FUD (Figure 1B). Riboflavin kinase and FMNAT tend to exist as two distinct enzymes in eukaryotes and are fused together to create the bifunctional FAD synthetase (FADS) in prokaryotes.[2–4,19] There exist some studies exploring the nucleotide specificity of riboflavin kinase and FMNAT enzymes. While most riboflavin kinases have been experimentally tested only with ATP, the other nucleotide triphosphates behave as substrates in some cases. Homologues of *Bacillus subtilis* and *rat intestinal mucosa* riboflavin kinase display the ability to use a range of NTPs to varying extents to make FMN[20,21], whereas some archaeal homologues exhibit specificity for CTP over ATP[22,23], and a fungal *Rhizopus javanicus* homologue utilizes GTP specifically while ATP gives no reaction.[24] In the case of FMNAT, the *Methanocaldococcus jannaschii* FMNAT shows *in vitro* activity with nucleotides other than ATP resulting in FAD and its nucleobase analogues.[11,25] As ATP, GTP, CTP, UTP, and the corresponding dNTPs exist naturally in the cell, FAD nucleotide analogues can, in principle, be produced in cells if the FMNAT enzyme/domain can accept them as alternate substrates. FAD analogues have been

synthesized and have demonstrated *in vitro* applications, opening up avenues for applied/ synthetic biology studies similar to those that exist for other cofactor analogues.[11,12,18,26–28] Also, unnatural adenine analogues of FAD have been demonstrated to bind to a host of flavoproteins with applications such as biosensing, covalent immobilization of flavoproteins, and as enzyme electrodes.[14,29–32]

The goal of this study was to create an *E. coli* strain that can synthesize nucleotide analogues of FAD *in vivo*. This strain will act as a model system to explore the physiological role of these FAD analogues for their potential applications as unnatural/ biorthogonal cofactors. For this, first, we altered the FMNAT domain of the *E. coli* FAD synthetase enzyme *in vitro* to produce FAD analogues and accordingly engineered the enzyme within the *E.coli* chromosome to produce FAD and its nucleobase analogues. We then examined its metabolism with a focus on redox functions and found that the strain exhibits antibiotic tolerance towards aminoglycosides such as kanamycin, amikacin, streptomycin, and gentamicin. As redox biochemistry plays a central role in antibiotic tolerance and resistance, we attribute the acquired tolerance to a possible change in proton-motive force or novel metabolism arising from the presence of the FAD analogues. Our study not only establishes a general method to create a strain that produces cofactor analogues but also reveals an unprecedented link to antibiotic tolerance, which requires detailed mechanistic inspection.

RESULTS

***E. coli* FADS shows *in vitro* activity with several nucleotides.** The *E. coli* FADS has been functionally and kinetically characterized for its ability to incorporate riboflavin, FMN and its analogues, and ATP .[33] The catalytic role of the individual domains of *E. coli* FADS has been briefly explored as a biotechnology route to improve FMN production, however, no nucleotide

other than ATP was tested for this reaction.[34] In contrast, the *Corynebacterium ammoniagenes* FADS (*Ca*FADS) has been extensively studied structurally and functionally as a model of prokaryotic FADS enzymes.[35–38] The *Ca*FADS riboflavin kinase module can be individually expressed and catalyzes the phosphorylation of riboflavin to produce FMN, while the FMNAT module when expressed individually appears to be non-functional.[35] Sequence alignment of *E. coli* FADS and *Ca*FADS shows 31% identity and 50% similarity, and the active sites for riboflavin, FMN, and both ATPs are well conserved (Figure 2A). Since no experimental structure is available for *E. coli* FADS, we used homology modelling to obtain a high confidence score structure (Figure 2B).[39]

We purified *E. coli* FADS and optimized the reaction conditions for riboflavin to FAD conversion with ATP as the co-substrate (Figure S1A, B, C). We then tested a panel of ribonucleotide and deoxyribonucleotide co-substrates ATP, GTP, CTP, UTP, and dATP, dGTP, dCTP, and dTTP at 3.5 mM each, which is the cellular concentration of ATP in *E. coli*.[40] Fluorescence thin-layer chromatography (TLC) and high performance liquid chromatography (HPLC) analysis showed the formation of FMN with ATP, CTP, and dATP and formation of FAD and dFAD (Figure 2C, D, Figure S1D). None of the other nucleotides appear to produce FMN and/ or FAD analogues (Figure S1D). Interestingly, when the cellular concentrations of each of the nucleotides were used with riboflavin as a substrate, dATP continued to show the formation of dFAD (Figure S1E). [40]

Next, we used FMN as a substrate with the panel of nucleotides to test only the FMNAT activity. We observed that the co-substrates ATP and dATP show high conversion to FAD and dFAD, while CTP, UTP, and dCTP showed partial conversion to FCD, FUD, and dFCD, respectively (Figure 2E, F). Based on these results, as the nucleotides and dATP showed significant conversion in both experiments, we proceeded with the ribonucleotides and dATP as our subset of co-substrates for the subsequent reactions.

Mutational analysis to alter the nucleotide specificity of the *E. coli* FADS FMNAT domain. The *E. coli* FADS FMNAT domain interacts with ATP via phosphate, the ribose ring, and the adenine base (Figure S2A, green, yellow and red residues, respectively). To probe the interactions important for ATP specificity and predict mutations, docking followed by molecular dynamics simulations were conducted. S165K was predicted to lower the enzyme's interaction energy with ATP, GTP and CTP by increasing interaction with the phosphate, and the G23K mutant was predicted to have a decreased preference for ATP and an increased preference for GTP as compared to the wt *E. coli* FADS (Figure S2B). However, when these mutants were cloned, expressed, purified and tested, an overall decrease in activity with ATP as compared to the wild-type FADS was observed and no additional promiscuity with nucleotides was seen (Figure S2C). Next, we created the G23S mutant corresponding to a similar mutation G105S in the FtsZ enzyme, which converts it from a GTPase to an ATPase (Figure S2D, the glycine is colored in red).[41] However, we found it completely inactive (Figure S2E).

A nucleobase region loop in the *E. coli* FADS FMNAT domain plays a role in nucleotide selectivity. Finally, a structure-based approach was employed to identify other nucleotidyltransferases that are homologous to *Ca*FADS and do not utilize ATP. A DALI search with a stringent z-score cutoff of 10 yielded four candidate cytidyltransferase structures (Table S1).[42] Interestingly, the *E. coli* FMNAT domain and cytidylyltransferases fall within the HxGH motif-containing nucleotidylyltransferase superfamily and display high overall structural similarity with a distinct difference in a loop adjoining to the nucleobase (Figure 3A). Glycerol-3-phosphate cytidylyltransferase has specificity for CTP or dCTP, hence, we hypothesized that changing the FMNAT domain loop with that of the cytidyltransferase may result in CTP specificity.[43] We swapped the *E. coli* FMNAT domain loop QTFCEGGVRIS sequence with the PRTEGIS loop from the glycerol-3-phosphate cytidylyltransferase and

purified the resultant *E. coli* LoopSwap-FADS enzyme. Surprisingly, on testing its activity with the nucleotide panel, the *E. coli* LoopSwap-FADS displayed similar efficiency at producing FAD as the wild-type *E. coli* FADS, and increased ability to utilize GTP, CTP, UTP, and dATP to produce FGD, FCD, FUD, and dFAD, respectively (Figure 3B, S3). Compared to the *E. coli* FADS, the LoopSwap-FADS is 17x, 1.2x, and 1.4x as efficient with GTP, CTP, and UTP, respectively (Figure 3B).

Next, we heterologously expressed the *E. coli* FADS and LoopSwap-FADS enzymes in *E. coli* BL21(DE3) and their lysates were analyzed by liquid chromatography-mass spectrometry (LC-MS), with the empty vector as a control. While the lysates for the empty vector and the FADS-expressing cells show that FAD is formed, the cells expressing the LoopSwap-FADS additionally produce FCD, FUD, and FGD along with FAD (Figure 3C). Our *in vitro* and heterologous expression data indicates that the FMNAT domain loop plays an essential role in nucleotide selectivity.

The FMNAT loop sequence is important for catalysis and nucleotide specificity. To probe the role of the residues in the *E. coli* FADS FMNAT domain loop, we created truncated variants by comparing the sequence, structure and length with the glycerol-3-phosphate cytidylyltransferase loop sequence. The shortened loop (ShortLoop) FADS enzyme mutants 1-6 – QTFCEGIS, QTFCGIS, QTFCIS, QTCGIS, QTFIS, and QTIS – were created (Figure 4A). The QTFCGIS enzyme shows formation of FAD, FUD, and FCD between 20-45% of that of the LoopSwap enzyme, whereas the QTFIS and QTIS enzymes showed significant activity with only ATP and dATP (Figure 4B, S4). Overall, a decrease in the formation of FAD and its analogues as compared to the wild-type and LoopSwap enzymes was noted (Figure 4B, S4). The decrease in the activity of ShortLoop enzymes may be attributed to their low interaction energy with the nucleotides, as calculated for the ShortLoop mutant QTFCGIS with ATP, CTP, and GTP (Figure S2B). Next, we heterologously expressed the ShortLoop enzymes in *E. coli*

BL21(DE3) cells to understand their cellular nucleotide selectivity profile. To our surprise, while the *in vitro* activity for the ShortLoop mutant enzymes was low, we detected FCD and FUD formation along with FAD by LC-MS in all the cell lysates containing the ShortLoop enzyme plasmids (Figure 4C). These results further strengthen the loop's role in ATP selectivity within the cell.

The *E. coli* LoopSwap strain produces FAD nucleobase analogues. The *ribF* gene encodes the wild-type FADS enzyme in *E. coli*. We substituted *ribF* with the altered gene sequence that would result in the LoopSwap gene product and created the *E. coli* LoopSwap strain. LC-MS analysis of the LoopSwap strain lysate revealed the presence of the FAD analogues FGD, FCD, and FUD (Figure 5A). We then quantified the FAD in the LoopSwap versus wild-type strain - the wild-type strain contains 1452 ± 411 pM FAD while the LoopSwap strain contains 918 ± 223 pM FAD, and the LoopSwap strain also shows FCD, FGD, and FUD production (Figure 5B, Figure S5A).

Probing the physiology of the *E. coli* LoopSwap strain. After confirming the presence of the FAD nucleotide analogues in the LoopSwap strain, we turned our attention to their effect on cellular function. In rich and minimal medium, the LoopSwap strain displays similar growth to the wild-type strain (Figure 6, S5B). An RNA sequencing analysis was conducted to compare the wild-type and LoopSwap strains grown in rich medium. With a stringent cutoff of $\log_2$ (fold change) > 3.3 and < -3.3, we observed that a wide range of physiological functions were altered (Figure 5C). There is an upregulation of genes involved in motility (bacterial flagellar assembly, chemotaxis) and nitrogen metabolism, and downregulation of sulfur (taurine and hypotaurine) metabolism among others.

The *E. coli* LoopSwap strain exhibits antibiotic tolerance. As nitrogen and sulfur metabolism pathways are redox regulated as well as involved in cellular redox regulation, we

investigated whether oxidative stress would affect the LoopSwap strain differently than the wild-type strain. The wild-type and the LoopSwap strains were grown in the presence of H_2O_2 to induce oxidative stress - their growth was suppressed to a similar extent with 2 mM H_2O_2, and 5 mM H_2O_2 was the lethal concentration (Figure S6A). Next, we studied their growth with a representative set of antibiotics as the source of oxidative stress.[44] Interestingly, we noticed that the LoopSwap strain grew better than the wild-type when treated with the aminoglycoside class of antibiotics (kanamycin, streptomycin, gentamycin, and amikacin) whereas colistin, ciprofloxacin, trimethoprim, rifampicin, piperacillin, ampicillin, and tetracycline affected both strains equally (Figure 6A-D, S6B-H). To confirm the aminoglycoside tolerance phenotype, we also subjected the *E.coli* LoopSwap-Kan strain (which was an intermediate strain with kanamycin resistance while creating the LoopSwap strain) to all the oxidative stress agents and obtained identical results (Figure S7). Following this, an extensive antibiotic MIC survey was done for the LoopSwap and wild-type strains to assess their tolerance profile. The LoopSwap strain has a ~1.6x MIC50 value as compared to the wild-type strain for the aminoglycosides, otherwise, both wild-type and LoopSwap strains were found to have similar MIC50 values for all other tested antibiotics (Figure 6E, S6I).

DISCUSSION

In this study, we explored and altered the nucleotide specificity of the *E. coli* FADS enzyme to create a range of mutant enzymes that can produce FAD and its nucleotide analogues *in vitro* as well as within cells. Surprisingly, we find that the intracellular production of FCD, FGD, and FUD is linked with the emergence of antibiotic tolerance.

Firstly, *in vitro* reactions with the wild-type bifunctional *E. coli* FADS enzyme revealed that its two functional domains, riboflavin kinase and FMNAT, display a difference in the

choice of the nucleotide co-substrate. Our studies show that the *E. coli* riboflavin kinase domain converting riboflavin to FMN is specific for ATP and dATP, whereas the FMNAT domain converting FMN to FAD can accept ATP, CTP, UTP, dATP, and dCTP, and can synthesize the corresponding FAD analogues (Figure 2C-F). This is intriguing as the conversion of riboflavin to FMN involves a phosphate transfer and thus, in principle, the reaction is chemically indifferent to the identity of the nucleotide/ phosphate source, whereas the FMN to FAD conversion introduces the nucleotide, and thus its identity is important. However, we observed that in spite of the *in vitro* promiscuity of the FMNAT domain, FADS is specific for ATP inside the *E. coli* cell and produces only FAD under native or heterologous expression conditions (Figure S1F). This may be attributed to the cellular ATP concentration being the highest among all nucleotides or to other cellular factors that regulate the entry/ availability of ATP as the co-substrate.

Predicting the amino acid residues that dictate adenine nucleobase specificity and altering it to make the enzyme promiscuous was challenging as their interactions appear to be via the amide backbone.[7,45] Based on interaction energy calculations and existing literature, mutations at the G23 and S165 positions in the FMNAT domain were conducted to produce a promiscuous enzyme. However, G23K and S165K mutants resulted in specificity for ATP accompanied by an overall decrease in activity, and the G23S mutant produced an inactive protein (Figure S2). This suggested that these residues do alter the flexibility of the active site in a significant manner.

Our structural alignment search provided significant insights - notably, we found that the FMNAT domain aligns structurally with cytidylyltransferases with a loop around the nucleobase region that is different.[43] Swapping the FMNAT loop (QTFCEGGVRIS) with that of the glycerol-3-phosphate-cytidyltransferase (PRTEGIS) increased promiscuity for the various nucleotides, that is, overall, the LoopSwap mutant enzyme produced FGD, FCD, FUD,

and dFAD with higher efficiency than the *E. coli* FADS (Figure 3B and 4B). Also, the overall efficiency of FAD production was retained *in vitro* while the activity with GTP was greatly enhanced as compared to the wild-type enzyme (Figure 4B). As the cellular concentrations of nucleotides in *E. coli* follows the order ATP > GTP > UTP >CTP, we hypothesize that the original loop sequence may have enforced ATP binding owing to a lower binding affinity for CTP, UTP, and GTP.[40] The swapping of the loop may now allow for lower concentrations of nucleotides to bind and act as substrates for the FMNAT reaction. Probing the length of the loop sequence by creating the ShortLoop enzyme mutants additionally confirmed that the entire loop was essential for the most efficient promiscuous enzyme (Figure 4B, 4C, S4).

Another interesting result was that the heterologous expression on plasmids of the LoopSwap and ShortLoop mutants in *E. coli* led to the formation of FCD, FUD, and FGD along with FAD within the cell, as compared to the strict specificity of FAD production with the wild-type FADS enzyme or an empty vector (Figure 3C, 4C). These observations substantiate the role of the FMNAT loop sequence in *E. coli* FADS, in its complete structure and sequence, for ATP specificity within the cell. The QTFCEGGV sequence of the FMNAT loop appears to be variable whereas the RIS sequence shows high conservation across organisms, as illustrated by a position specific substitution matrix derived by aligning 5000 FADS homologues (Figure S8). To get a sense of how often a loop structure appears in the ATP- binding region of a protein, we looked for the presence of such loops in 80 representative PDB structures from the available published list of 985 crystal structures bounded to ATP, ADP, AMP, CoA, FAD, NAD, SAM and their analogues.[45] We defined the loop as at least four consecutive residues present around the nucleobase, and loops that are positioned away from the nucleobase in the crystal structure were not considered. Our findings revealed that 62% of proteins utilizing ATP and its derivatives featured this distinct loop structure with varying sizes (Table S2).

Finally, to understand the physiological effect of the loop, chromosomal substitution of the wild-type E. coli FAD synthetase gene with the LoopSwap gene sequence was done to generate the LoopSwap strain. The LoopSwap strain (i) produces ~63% FAD as compared to the wild-type strain, and (ii) additionally produces the FAD analogues (Figure 5A, B). Interestingly, it exhibited a phenotype of 1.6x aminoglycoside antibiotic tolerance as compared to the wild-type strain (Figure 6E). This could occur due to the lower FAD content in the *E. coli* LoopSwap strain which results in lowering the proton-motive force (PMF). As FAD plays an integral role in the electron transport chain, lowering its concentration may be associated with persisters with higher antibiotic tolerance.[46] However, the *E. coli* LoopSwap strain grows just as well as the *E. coli* wild-type strain hence this is an unlikely possibility (Figure S5B, C). Thus, we hypothesize that the FAD analogues may play a role analogous to FAD in the cell, or be potentially involved in other metabolic pathways. To test this further, we conducted computational studies to check whether the FAD analogues have any interactions with *E. coli* proteins. To this end, we prepared the "*E.coli* K-12 MG1655 protein structural master database", which is an exhaustive compilation of all *E.coli* protein structures as described in the methods section (Figure S9, Table S3). This database was used to dock FAD and its nucleotide analogues to all *E. coli* proteins and compare their binding affinities. First, we conducted docking studies of the FAD analogues with the FAD-utilizing enzyme glutathione reductase which has been experimentally shown to exhibit activity with the FAD analogues (Figure S10).[11,47,48] We found that FUD and FGD have similar binding energy as FAD. Additionally, FGD, FCD, and FUD have very similar docking poses as FAD within the glutathione reductase active site. Thus, it is plausible that the FAD analogues may act as potential cofactors for other FAD-binding enzymes. We commenced our proteome-wide docking analysis of FAD and its analogues by validating our docking approach with FAD-

binding proteins (as described in the methods section) and found it to be satisfactory (Figure S11A-C, Table S4).

Next, we docked the 74 FAD-binding proteins in the master database with FAD, FCD, FGD, and FUD, and calculated the value ΔA (FAD analogue) = A (FAD analogue) – A(FAD), where A is the protein-ligand binding energy. Docking results from the 11 FAD binding proteins with holo structures resulted in ΔA values between -1.04 to 2.61 kcal/ mol, indicating similar binding (Table S5). We found that the binding poses of FAD and the FAD analogues are similar, with comparable interactions around the catalytic isoalloxazine ring, while the difference in affinity arises from the interactions with the different nucleobases attached to the ligands (Figure 6F). We extended this computation of binding affinities to all *E. coli* FAD-utilizing proteins (apo structures from the PDB and modelled structures from AlphaFold DB, full list in Table S6), and the results indicate that 78% FAD-binding proteins bind similarly to the FAD analogues, with ΔA values ranging from -2 to 2 kcal/mol (Figure S12). Our docking studies show that FCD, FGD and FUD bind better to some of the FAD-binding proteins as compared to FAD, and this may lead to some functions being performed differently in the LoopSwap strain as compared to the wild-type strain (Figure S12).

Finally, it is also possible that the FAD analogues bind to non-FAD binding proteins causing downstream effects leading to tolerance. To investigate this, we docked FAD and the FAD analogues with the master database (as described in the methods section). Interestingly, we found 7 candidate proteins that bind stronger to the FAD analogues as compared to FAD, which may be interesting to test experimentally (Figures S13, S14, Tables S7, S8). Specific examples include the outer membrane porin F (OmpF) and the tryptophan-specific transport protein which are also among the highest upregulated/ downregulated genes in the RNA-Seq analysis (Figure S14). The significance and metabolic role of the proteins hits in inducing aminoglycoside tolerance is unexplored and requires further molecular level investigations.

CONCLUSION:

In conclusion, this study presents a structural and functional analysis of *E. coli* FAD synthetase, with a focus on elucidating its nucleotide selectivity determinants. Engineering a cytidyltransferase-derived loop into the FMNAT domain enabled broadening of the enzyme's substrate specificity to accommodate alternative nucleotides. The resulting *E. coli* LoopSwap strain synthesizes both canonical FAD and its nucleotide analogues, while exhibiting enhanced tolerance to aminoglycoside antibiotics. These findings suggest that targeted modification of the loop region proximal to the nucleobase-binding site can modulate nucleotide specificity in adenylating enzymes. This study advances our understanding of FAD biosynthesis and its metabolic roles, including its unexpected link to antibiotic resistance. Finally, it highlights the potential for developing bio-orthogonal redox systems in synthetic biology, thus expanding the biochemical toolkit for metabolic engineering and biotechnological innovation.

ACKNOWLEDGEMENTS

Financial support was provided by the Ministry of Science and Technology, Govt. of India – SERB Core Research Grant CRG/2019/003270 for this research to A.B.H. and for salary to A.S. The authors acknowledge the DST fund for Improvement of S&T Infrastructure (SR/FST/LSII-043/2016) to the IISER Pune Biology Dept for the Biological Mass Spectrometry Facility and the IISER Pune NMR and MS facilities and their managers. The support and the resources provided by 'PARAM Brahma Facility' under the National Supercomputing Mission, Government of India at the Indian Institute of Science Education and Research (IISER) Pune are gratefully acknowledged. We thank DBT, Govt. of India (Grant no. BT/PR34215/AI/133/22/2019) to A. M. for partial funding of the computational resources used in this work.

REFERENCES

1. Bacher, A., Eberhardt, S., Fischer, M., Kis, K. & Richter, G. Biosynthesis of Vitamin B_2 (Riboflavin). *Annu Rev Nutr* **20**, 153–167 (2000).
2. Manstein, D. J. & Pai, E. F. Purification and characterization of FAD synthetase from Brevibacterium ammoniagenes. *J Biol Chem* **261**, 16169–73 (1986).
3. Barile, M., Brizio, C., Valenti, D., De Virgilio, C. & Passarella, S. The riboflavin/FAD cycle in rat liver mitochondria. *Eur J Biochem* **267**, 4888–4900 (2000).
4. Merrill, A. H. & McCormick, D. B. Affinity chromatographic purification and properties of flavokinase (ATP:riboflavin 5'-phosphotransferase) from rat liver. *Journal of Biological Chemistry* **255**, 1335–1338 (1980).
5. Walsh, C. Flavin coenzymes: at the crossroads of biological redox chemistry. *Acc Chem Res* **13**, 148–155 (1980).
6. Moodie, S. L., Mitchell, J. B. O. & Thornton, J. M. Protein Recognition of Adenylate: An Example of a Fuzzy Recognition Template. *J Mol Biol* **263**, 486–500 (1996).
7. 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* **44**, 282–291 (2001).
8. Jadhav, V. R. & Yarus, M. Coenzymes as coribozymes. *Biochimie* **84**, 877–888 (2002).
9. Bruns, S. *et al.* A Novel Coenzyme A Analogue in the Anaerobic, Sulfate-Reducing, Marine Bacterium *Desulfobacula toluolica* Tol2 ^{T}. *ChemBioChem* **24**, e202200584 (2023).
10. Kuppuraj, G., Kruise, D. & Yura, K. Conformational Behavior of Flavin Adenine Dinucleotide: Conserved Stereochemistry in Bound and Free States. *J Phys Chem B* **118**, 13486–13497 (2014).
11. 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* **24**, e202300055 (2023).
12. 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* **4**, 504–10 (1965).
13. McCormick, D. B. & Opar, G. E. Syntheses of 8-bromo-5'-adenylate-containing nucleotides. *J Med Chem* **12**, 333–334 (1969).
14. 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* **238**, 519–528 (1996).
15. Koberstein, R. 8-Azidoadenine Analogs of NAD+ and FAD. Synthesis and Coenzyme Properties with NAD+ -Dependent and FAD-Dependent Enzymes. *Eur J Biochem* **67**, 223–229 (1976).

16. 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* **396**, 280–283 (2010).

17. Wang, S. *et al.* Design, Synthesis and SAR Studies of NAD Analogues as Potent Inhibitors towards CD38 NADase. *Molecules* **19**, 15754–15767 (2014).

18. Guo, X. *et al.* Non-natural Cofactor and Formate-Driven Reductive Carboxylation of Pyruvate. *Angewandte Chemie* **132**, 3167–3170 (2020).

19. Krupa, A., Sandhya, K., Srinivasan, N. & Jonnalagadda, S. A conserved domain in prokaryotic bifunctional FAD synthetases can potentially catalyze nucleotide transfer. *Trends Biochem Sci* **28**, 9–12 (2003).

20. Kasai, S., Nakano, H., Maeda, K. & Matsui, K. Purification, Properties, and Function of Flavokinase from Rat Intestinal Mucosa. *The Journal of Biochemistry* **107**, 298–303 (1990).

21. Solovieva, I. M., Tarasov, K. V. & Perumov, D. A. Main physicochemical features of monofunctional flavokinase from Bacillus subtilis. *Biochemistry (Moscow)* **68**, 177–181 (2003).

22. Mashhadi, Z., Zhang, H., Xu, H. & White, R. H. Identification and Characterization of an Archaeon-Specific Riboflavin Kinase. *J Bacteriol* **190**, 2615–2618 (2008).

23. 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* **22**, 3414–3424 (2021).

24. Tachibana, S. GTP-dependent flavin phosphate synthetase. in *Methods in Enzymology* 553–555 (1971). doi:10.1016/S0076-6879(71)18118-9.

25. Mashhadi, Z., Xu, H., Grochowski, L. L. & White, R. H. Archaeal RibL: A New FAD Synthetase That Is Air Sensitive. *Biochemistry* **49**, 8748–8755 (2010).

26. 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* **94**, 3565–3570 (1997).

27. 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* **6**, 1219–1226 (1998).

28. Bishop, A. *et al.* Unnatural Ligands for Engineered Proteins: New Tools for Chemical Genetics. *Annu Rev Biophys Biomol Struct* **29**, 577–606 (2000).

29. Posthuma-Trumpie, G. A. *et al.* Reconstitution of apoglucose oxidase with FAD conjugates for biosensoring of progesterone. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1774**, 803–812 (2007).

30. Krzek, M., van Beek, H. L., Permentier, H. P., Bischoff, R. & Fraaije, M. W. Covalent immobilization of a flavoprotein monooxygenase via its flavin cofactor. *Enzyme Microb Technol* **82**, 138–143 (2016).

31. Narayanan, M., Singh, V. R., Kodali, G., Moravcevic, K. & Stanley, R. J. An Ethenoadenine FADAnalog Accelerates UV Dimer Repair by DNA Photolyase. *Photochem Photobiol* **93**, 343–354 (2017).

32. Willner, I. *et al.* Electrical Wiring of Glucose Oxidase by Reconstitution of FAD-Modified Monolayers Assembled onto Au-Electrodes. *J Am Chem Soc* **118**, 10321–10322 (1996).

33. Langer, S., Hashimoto, M., Hobl, B., Mathes, T. & Mack, M. Flavoproteins Are Potential Targets for the Antibiotic Roseoflavin in Escherichia coli. *J Bacteriol* **195**, 4037–4045 (2013).

34. Kitatsuji, K., Ishino, S., Teshiba, S. & Arimoto, M. Process for producing flavine nucleotides. European Patent Office. (1992).

35. Frago, S., Martínez-Júlvez, M., Serrano, A. & Medina, M. Structural analysis of FAD synthetase from Corynebacterium ammoniagenes. *BMC Microbiol* **8**, 160 (2008).

36. 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* **65**, 1285–1288 (2009).

37. 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* **400**, 218–230 (2010).

38. Lans, I. *et al.* The Dimer-of-Trimers Assembly Prevents Catalysis at the Transferase Site of Prokaryotic FAD Synthase. *Biophys J* **115**, 988–995 (2018).

39. 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* **10**, 845–858 (2015).

40. Buckstein, M. H., He, J. & Rubin, H. Characterization of Nucleotide Pools as a Function of Physiological State in *Escherichia coli*. *J Bacteriol* **190**, 718–726 (2008).

41. RayChaudhuri, D. & Park, J. T. A point mutation converts Escherichia coli FtsZ septation GTPase to an ATPase. *Journal of Biological Chemistry* **269**, 22941–22944 (1994).

42. Holm, L. Dali server: structural unification of protein families. *Nucleic Acids Res* **50**, W210–W215 (2022).

43. 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* **268**, 16648–16654 (1993).

44. Dwyer, D. J. *et al.* Antibiotics induce redox-related physiological alterations as part of their lethality. *Proceedings of the National Academy of Sciences* **111**, E2100–E2109 (2014).

45. Narunsky, A. *et al.* On the evolution of protein–adenine binding. *Proceedings of the National Academy of Sciences* **117**, 4701–4709 (2020).

46. Allison K.R. *et al.* Metabolite-enabled eradication of bacterial persisters by aminoglycosides. *Nature* **473**, 216-220 (2011)

47. Corso, G. *et al.* Deep Confident Steps to New Pockets: Strategies for Docking Generalization. *ArXiv* (2024).

48. McNutt, A. T. *et al.* GNINA 1.0: molecular docking with deep learning. *J Cheminform* **13**, 43 (2021).

FIGURES

Engineering *Escherichia coli* FAD synthetase for the biosynthesis of FAD nucleotide analogues results in aminoglycoside antibiotic tolerance

Ateek Shah[1#], Prathith Bhargav[1], Shreya Mahato[2], Reman Kumar Singh[1†], Yashwant Kumar[1§], Pratishruti Panda[1], Janhavi Borkar[2], Lavisha Parab[3], Nishad Matange[2], Arnab Mukherjee*[1, 4], Amrita B. Hazra* [1,2]

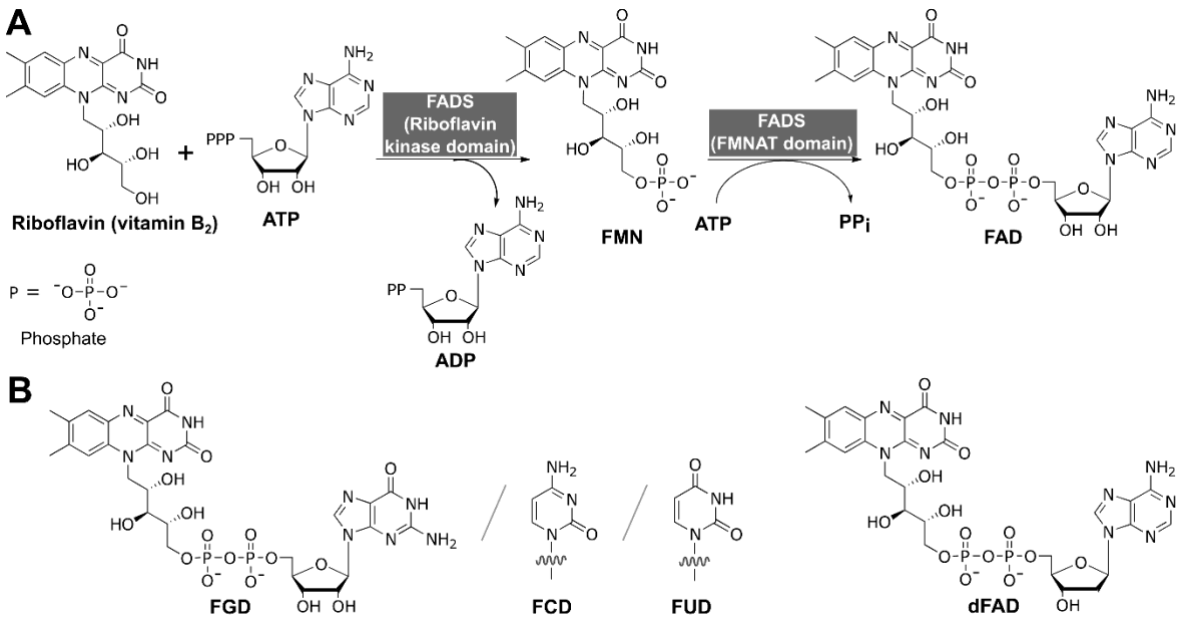


Figure 1. Biosynthesis of flavin adenine dinucleotide (FAD) and its nucleoside analogues. (A) The biosynthesis of FAD involves conversion of riboflavin (vitamin B_2) to flavin mononucleotide (FMN) followed by adenylation to produce flavin adenine dinucleotide (FAD). Both steps use adenosine triphosphate (ATP) as a co-substrate and produce adenosine diphosphate (ADP) and inorganic pyrophosphate (PP_i) as byproducts. **(B)** FAD nucleotide analogues - flavin guanine dinucleotide (FGD), flavin cytosine dinucleotide (FCD), and flavin uracil dinucleotide (FUD). When adenosine is

replaced by deoxyadenosine, deoxy-flavin adenine dinucleotide (dFAD) is formed. Even though these FAD analogues do not naturally occur in cellular systems, the starting substrates FMN and nucleotides are available in cells. Hence, this set of FAD nucleobase/nucleoside analogues can, in principle, be generated in cells using the appropriate riboflavin kinase (RFK) and FMN adenylyltransferase (FMNAT) enzymes and their mutants.

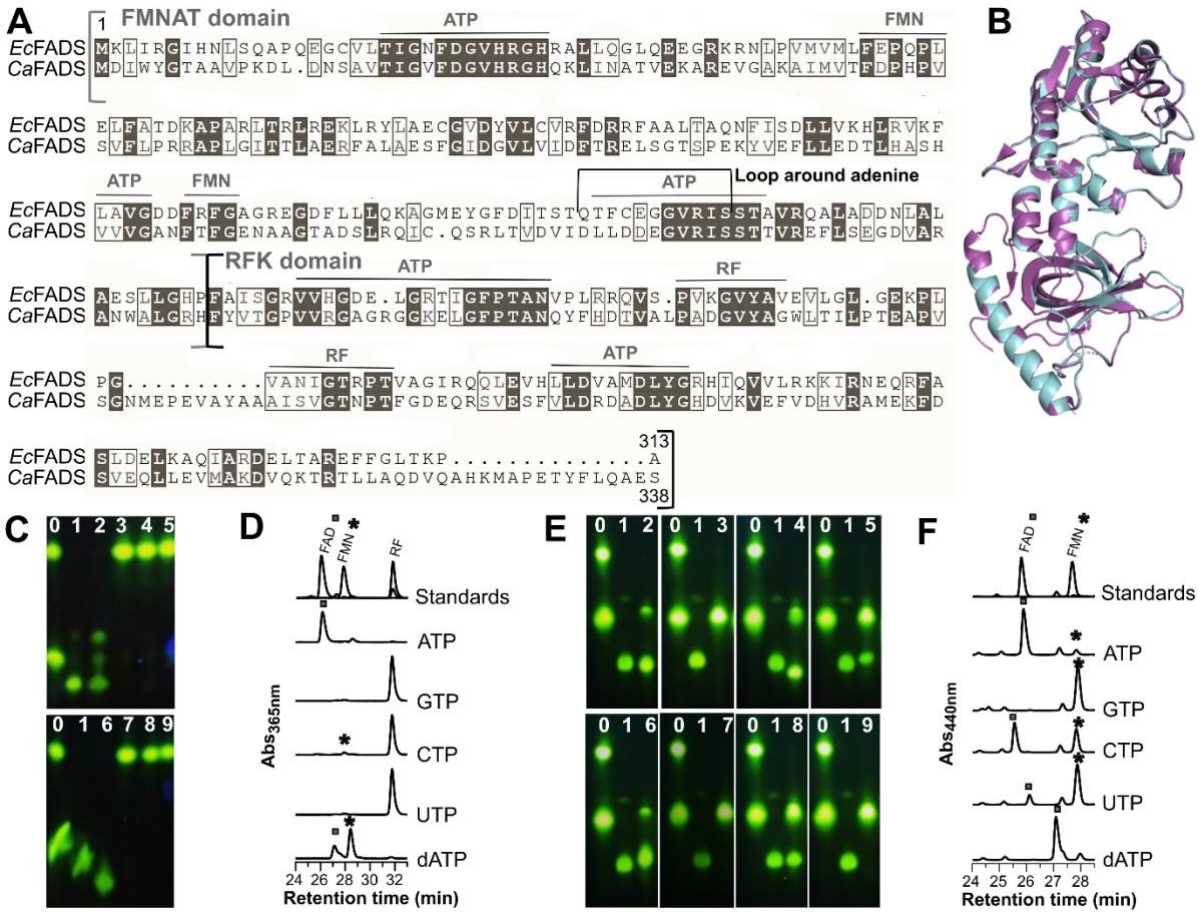


Figure 2. Sequence and modelled structure of *Ec*FADS and its activity with nucleotides. (A) Sequence alignment of *Ec*FADS with *Ca*FADS shows 31 % sequence similarity and 50 % homology. The active site residues are well conserved as shown in black boxes. The square brackets indicate the two domains FMNAT and riboflavin kinase (RFK). The grey horizontal lines indicate the FMN, riboflavin, and ATP-binding regions **(B)** The alignment of modelled *Ec*FADS (shown in cyan) with crystal structure of *Ca*FADS (PDB 2x0k) (shown in magenta) shows complete overlap, and the EcFADS protein model shows a 100% confidence score. **(C)** Activity of *Ec*FADS monitored with fluorescence thin-layer chromatography (TLC) with 3.5 mM nucleotides (ATP, GTP, CTP, UTP, dATP, dGTP, dCTP and dUTP) and riboflavin (RF)‡. **(D)** HPLC analysis of the *Ec*FADS activity with RF and nucleotides shows activity with ATP and dATP (* indicates the FMN and ■ indicates FAD and its analogues). **(E)** Fluorescence TLC of

*Ec*FADS activity with nucleotides and FMN. Signals for FAD, FCD, FUD, dFAD, and dFCD were observed. **(F)** HPLC analysis of the *Ec*FADS activity with FMN and nucleotides. The FMNAT site was found to be active with all nucleotides and dATP except GTP‡ (* indicates the FMN and ■ indicates FAD and its analogues). ‡**(Key for C and E)** Fluorescent TLC of *Ec*FADS activity Lane 0- RF and FMN standard, 1- FAD standard, Lanes 2-9, reaction with 2 - ATP, 3 - GTP, 4 - CTP, 5- UTP, 6 - dATP, 7- dGTP, 8- dCTP, 9 – dTTP.

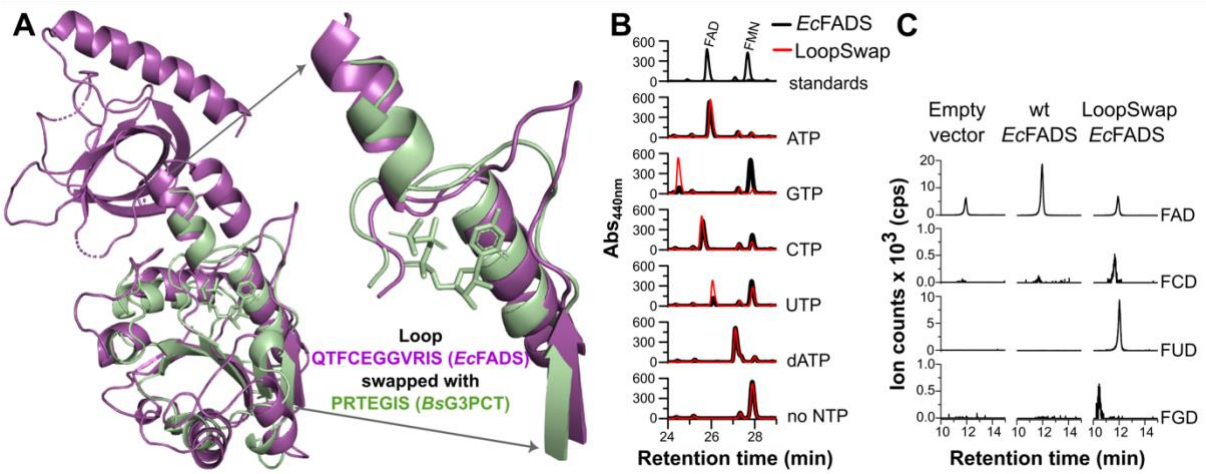


Figure 3. Testing the promiscuity of LoopSwap *Ec*FADS mutant *in vitro* and by heterologous expression. (**A**) Structural alignment of *Ec*FADS (magenta) with *Bacillus subtilis* glycerol 3-phosphate cytidylyl transferase (*Bs*G3PCT: green) (PDB 1coz). The *Ec*FADS FMNAT domain shares high structural similarity with the cytidylyl transferase and very low sequence similarity. The loops surrounding the nucleotide in *Ec*FADS-FMNAT domain and *Bs*G3PCT have different sizes. **(B)** HPLC data showing *in vitro* activity of the LoopSwap enzyme. The LoopSwap enzyme utilizes CTP, UTP and GTP as substrates, while the wild-type *Ec*FADS shows lower activity with CTP and UTP and no activity with GTP. (**C**) Extracted ion chromatograms (EIC) of FAD and its analogues

extracted from in the strains heterologously expressing the empty pET28a(+) vector, or the wild-type *Ec*FADS or LoopSwap enzyme. The strain expressing the LoopSwap enzyme shows signals for FAD nucleobase analogues along with FAD while the one expressing the wild-type *Ec*FADS shows only FAD.

A

Plasmid name	Loop	Deleted residues
*Ec*FADS	QTFCEGGVRIS	None
LoopSwap-FADS	PRTEGIS	QTFCEGGVRIS
ShortLoop1	QTFCEGIS	GVR
ShortLoop2	QTFCGIS	EGVR
ShortLoop3	QTFCIS	EGGVR
ShortLoop4	QTFCGIS	FEGVR
ShortLoop5	QTFIS	CEGGVR
ShortLoop6	QTIS	FCEGGVR

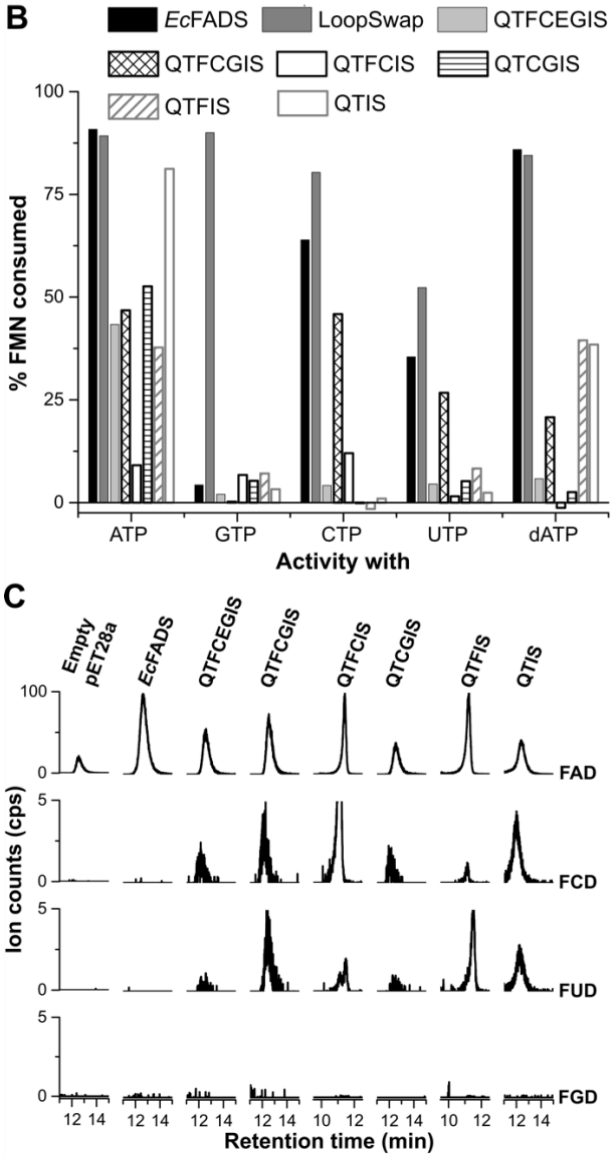
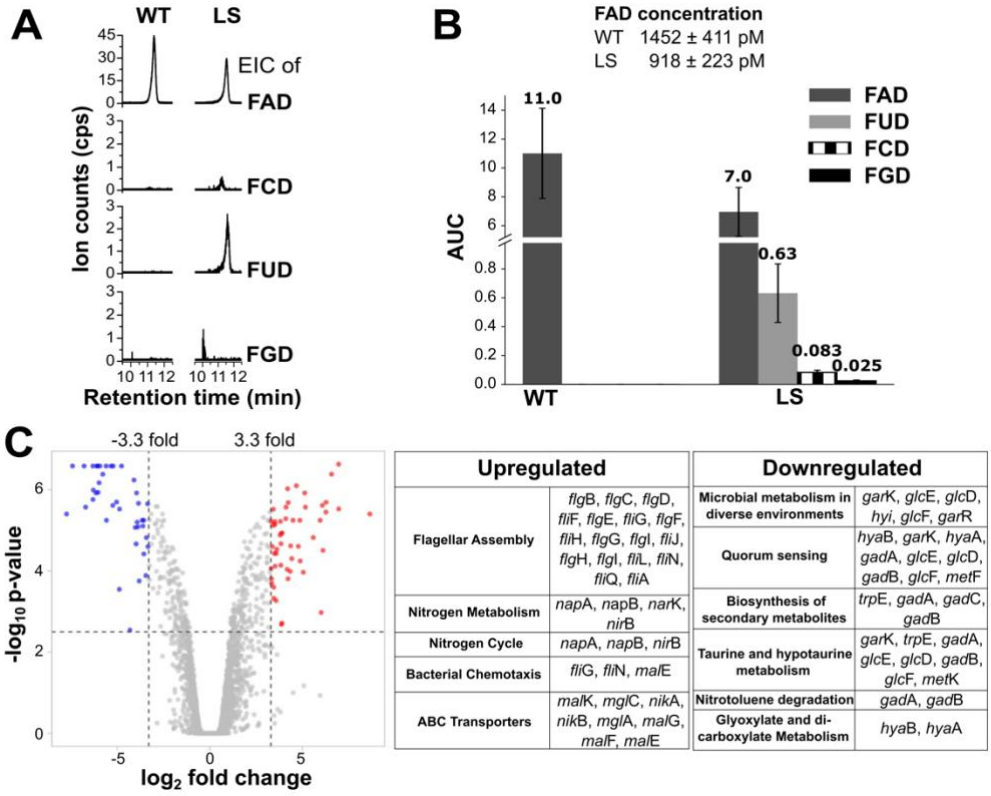


Figure 4. Comparison of the *in vitro* activity of the LoopSwap and ShortLoop mutants (A) The list of loops those present in these mutants and the corresponding deleted residues from wild-type *Ec*FADS. **(B)** A quantitative analysis of FAD and its analogues formed by wild-type *Ec*FADS, LoopSwap and ShortLoop mutants with

various nucleotides, derived from HPLC data. The LoopSwap enzyme shows highest activity with other nucleotides as compared to *Ec*FADS. In contrast, the ShortLoop mutants show a decrease in activity with ATP and other nucleotides. The QTFCGIS mutant is interesting because it shows loss of activity overall but mirrors the LoopSwap enzyme's preference for ATP, CTP, and UTP. **(C)** Extracted ion chromatograms (EIC) of FAD and its analogues extracted from the strains heterologously expressing the empty pET28a(+) vector, or the wild-type *Ec*FADS or the ShortLoop enzymes. Note that the ShortLoop mutants, being less efficient *in vitro*, can yet synthesize FAD nucleobase analogues in the cell when heterologously expressed. This shows that the loop plays a primary role in nucleotide specificity in the cell. The list of loops, deleted residues, and plasmid names is given in the table.



Upregulated	
Flagellar Assembly	*flg*B, *flg*C, *flg*D, *fli*F, *flg*E, *fli*G, *flg*F, *fli*H, *flg*G, *flg*I, *fli*J, *flg*H, *flg*I, *fli*L, *fli*N, *fli*Q, *fli*A
Nitrogen Metabolism	*nap*A, *nap*B, *nar*K, *nir*B
Nitrogen Cycle	*nap*A, *nap*B, *nir*B
Bacterial Chemotaxis	*fli*G, *fli*N, *mal*E
ABC Transporters	*mal*K, *mgl*C, *nik*A, *nik*B, *mgl*A, *mal*G, *mal*F, *mal*E

Downregulated	
Microbial metabolism in diverse environments	*gar*K, *glc*E, *glc*D, *hyi*, *glc*F, *gar*R
Quorum sensing	*hya*B, *gar*K, *hya*A, *gad*A, *glc*E, *glc*D, *gad*B, *glc*F, *met*F
Biosynthesis of secondary metabolites	*trp*E, *gad*A, *gad*C, *gad*B
Taurine and hypotaurine metabolism	*gar*K, *trp*E, *gad*A, *glc*E, *glc*D, *gad*B, *glc*F, *met*K
Nitrotoluene degradation	*gad*A, *gad*B
Glyoxylate and di-carboxylate Metabolism	*hya*B, *hya*A

Figure 5. The LoopSwap (LS) strain vs *E. coli* WT strain. (A) EICs of FAD and its analogues in the *E.coli* wild-type (WT) and LoopSwap (LS) strains. **(B)** Area under the

curve (AUC) for qualitative measure of FAD and its analogues in the *E. coli* WT and LS strains. **(C)** Volcano plot for RNA sequencing of the *E.coli* wild-type compared to the LoopSwap strain showing 92 upregulated (red)/ downregulated (blue) genes (with a stringent cut-off criterion, $\log_2$(fold change) above 3.3 and below -3.3, respectively, with P-values below ~0.003), and their potential functions.

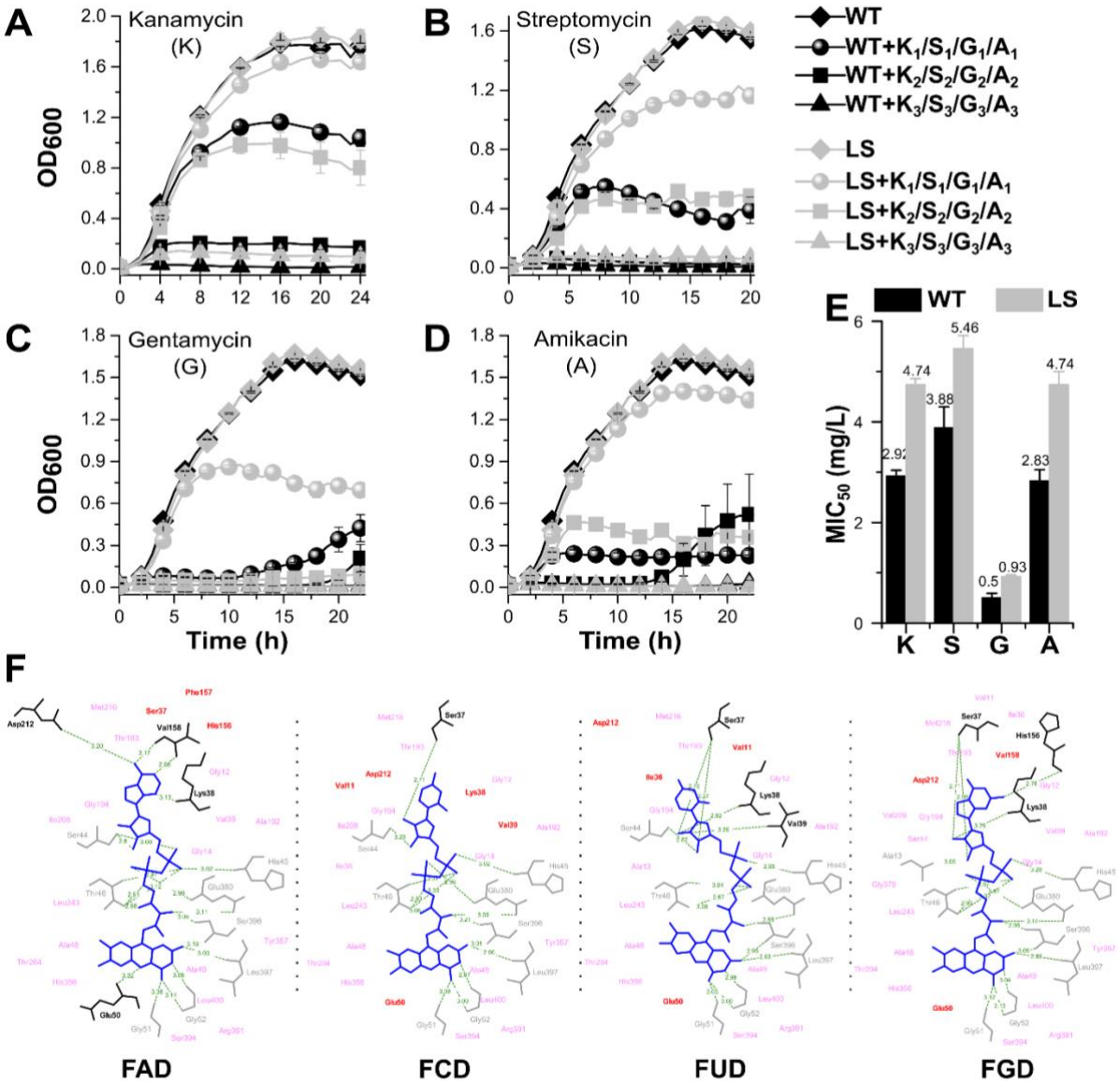


Figure 6. The LoopSwap (LS) strain exhibits antibiotic tolerance against aminoglycosides kanamycin, streptomycin, gentamicin, and amikacin. (A, B, C, D) Growth curve of *E. coli* WT and LS strains in presence of Kanamycin (K)*,

Streptomycin (S)**, Gentamycin (G)[&], Amikacin (A)[#]. K*- Kanamycin (K1= 2.5, K2= 5, K3= 10 µg/mL), S**- Streptomycin (S1= 5, S2= 10, S3= 15 µg/mL), G[&]- Gentamycin (G1= 1, G2= 2.5, G3= 5 µg/mL), A[#]- Amikacin (A1= 2.5, A2= 5, A3= 15 µg/mL). **(E)** MIC_{50} of aminoglycosides for WT and LS strains. **(F)** Interactions of docked FAD and its analogues with fumarate reductase flavoprotein subunit (UniProt ID: P00363) obtained through LigPlot+[15], table showing the association energy difference (ΔA) of holo FAD binding proteins. Differing hydrogen bonding interactions are highlighted in black while differing hydrophobic interactions between analogues are highlighted in red. Common hydrogen bonds are shown in grey, while common hydrophobic interactions between analogues is shown in pink.