Honors Theses

Honors College

4-2019

Cloning, Expression, and Purification of FadK and Its Application in CoA-RNA Capture

Jon-Michael L. Stork

Follow this and additional works at: https://aquila.usm.edu/honors_theses

Part of the Biochemistry Commons

The University of Southern Mississippi

Cloning, Expression, and Purification of FadK and Its Application in CoA-RNA
Capture

by

Jon-Michael Lee Stork

A Thesis
Submitted to the Honors College of
The University of Southern Mississippi
in Partial Fulfillment
of Honors Requirements

May 2019

Approved by

_____

Faqing Huang, Ph.D., Thesis Advisor

Department of Chemistry and Biochemistry

The University of Southern Mississippi

_____

Bernd Schroeder, Ph.D., Department Chair

School of Mathematics and Natural Sciences

The University of Southern Mississippi

_____

Ellen Weinauer, Ph.D., Dean

Honors College

The University of Southern Mississippi

Abstract

Coenzyme A (CoA) is an important enzyme cofactor involved in acyl transfer reactions. Recently, CoA and its various thioesters were found to exist at the 5' end of RNA molecules. Although the function of these molecules is still unknown due to difficulties in their isolation, their existence at the 5' RNA end reveals potentially novel biological roles of RNA. We are exploiting the broad substrate tolerance of the medium chain fatty acid-coenzyme A ligase (FadK), hypothesizing that this enzyme will accept biotinylated medium chain fatty acid substrates, creating biotin-tagged CoA-RNA. This modified CoA-RNA can later be captured using streptavidin affinity chromatography, allowing for the analysis and eventual determination of the sequence of CoA-RNA in bacterial cells. The focus of this project is to adequately clone, express, and purify the enzyme, followed by subsequent analysis of enzymatic activities, including the substrate spectrum and enzyme kinetics. Using recombinant DNA technology, we have successfully cloned, expressed, and purified FadK. Although there is no direct and simple way to analyze the activity of the purified enzyme, we hypothesize that AMP deaminase (AMPD) could be used to convert adenosine monophosphate (AMP) released during the thioesterification reaction to inosine monophosphate (IMP), which would cause UV absorbance changes that are easily monitored by UV spectrophotometry. When fully developed, this technique could be used to analyze and quantitate AMP, and therefore may have broad applications in analyzing enzyme-catalyzed reactions involving carboxyl group activation such as amino acid activation and fatty acid metabolism.

Key Terms: FadK, substrate tolerance, acyl transfer reaction, CoA-RNA, recombinant

DNA technology, UV spectrophotometry

Table of Contents

List of Abbreviations

CoA	coenzyme A
ATP	adenosine triphosphate
AMP	adenosine monophosphate
RNA	ribonucleic acid
mRNA	messenger ribonucleic acid
miRNA	micro ribonucleic acid
siRNA	small interfering ribonucleic acid
dpCoA	dephospho-coenzyme A
DNA	deoxyribonucleic acid
PCR	polymerase chain reaction
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
dNTP	deoxyribonucleotide triphosphate
EDTA	ethylenediaminetetraacetic acid
TEMED	N,N,N',N'-tetramethylethylenediamine
SOC	super optimal broth with catabolite repression
IPTG	isopropyl β-D-1-thiogalactopyranoside
$NiCl_2$	nickel (II) chloride
Na_2HPO_4	disodium phosphate
NaH_2PO_4	monosodium phosphate
DI	deionized

Chapter I: Introduction

CoA is an enzyme cofactor seen ubiquitously in biochemistry due to its various functional roles such as those in fatty acid metabolism and pyruvate oxidation, where it is often an acyl group carrier. The further presence of CoA thioesters in numerous metabolic reactions only strengthens the clearly recognizable importance of CoA in biochemical pathways. *In vivo*, CoA is synthesized from pantothenic acid, or vitamin B_5, over five steps; however, this process is somewhat expensive to the cell, requiring four molecules of ATP including the incorporation of an AMP residue via adenylation. Interestingly, the retained presence of adenine within CoA, CoA thioesters, and other enzyme cofactors is indicative of its importance in and alongside coenzyme functioning[8]. Indeed, according to White, CoA is a surviving metabolic molecule of a pre-protein RNA world, so it is only natural that CoA has such a presence in biochemical pathways[15].

Recently, CoA has been found to exist at the 5' end of certain short length RNA molecules[9], making it evident that further research must be performed to uncover the role of CoA-linked RNA and question current scientific understanding of the functions CoA is capable of carrying. As with instances of RNA alteration such as m^7G-capped mRNA seen in eukaryotes, CoA-linked RNA may be capable of new regulatory functions of genes or different forms of catalysis[13]. However, the issues of determining the structure and function of CoA-linked RNA are found in the lack of an effective way to isolate the molecule. Because structure and function are so closely related in terms of biomolecules, accurate isolation methods must be developed to allow for a better understanding of the

synthesis, degradation, alteration, and roles of CoA-linked RNA. In this regard, the enzyme FadK shows potential to remedy this issue by way of its substrate specificity.

FadK, a fatty acyl-CoA ligase, uses ATP to activate fatty acids into fatty acyl-CoA products for use in fatty acid metabolism through a two-step esterification reaction. As with all other fatty acyl-CoA ligases, this esterification reaction takes place through a fatty acyl-AMP intermediate. Important about FadK is that it functions with medium chain fatty acid substrates between 6 and 12 carbons in length but operates most efficiently using 6 and 8 carbon fatty acids[12]. Due to the broad substrate specificity of FadK, it may therefore be possible to exploit the protein's enzymatic activity in the presence of both a biotinylated medium chain fatty acid and 3'-dpCoA to yield a biotinylated fatty acyl-3'-dpCoA molecule. Using the capability of adenosine derivatives to act as transcription initiators *in vitro*[1,5], an RNA product containing a biotinylated fatty acyl-3'-dpCoA integrated at its 5' end can be synthesized. The use of 3'-dpCoA is critical in forming a biotinylated fatty acyl-CoA molecule, because the phosphate group normally found on the 3' end of CoA may interrupt transcription in the 3' direction[5]. The biotinylated fatty acyl-3'-dpCoA-linked RNA can then be captured via streptavidin affinity chromatography, where the biotin tag would allow for the isolation of the desired RNA molecules for further study.

Identification of Coenzyme A Linked RNA

Previously, knowledge about the structures of RNA molecules naturally present in eukaryotes and prokaryotes was largely confined to messenger RNAs, transfer RNAs, and ribosomal RNAs, as well as their derivatives such as capped mRNA. While new information about these various RNAs is being discovered, the general functions of each of the molecules have already been detailed.

On the other hand, little is known about small molecule RNAs. Recently, two new methods for the detection of RNA conjugation with small molecules were established by Kowtoniuk, et al. In the first method, a nucleophile or base was used to identify small molecules separated from RNA; the second method developed, however, monitored mononucleotides instead. Interestingly, in developing the second method, researchers discovered a number of novel small molecule conjugates including CoA-linked RNA and its numerous derivatives, all of which were only observed with RNAs around 200 or fewer nucleotides in length[9]. Based upon the structure of 3'-dephospho-CoA, a biological precursor to CoA, it was also suggested that the small molecules (namely CoA) existed at the 5' end of RNA molecules, implying a potential for an unspecified and undetermined role of RNA within the world of biology[9,10].

To confirm that the CoA was a part of the RNA conjugate and not simply a contaminate that had persisted through RNA purification, experimentation was repeated with added CoA-containing thioesters and compounds such as benzoyl-CoA and butyryl-CoA. The levels of 3'-dephosphorylated CoA species did not considerably increase

whilst the experiment was performed, indicating that any CoA detected beforehand was indeed linked to an accompanying RNA molecule[9]. To determine the location of the CoA-RNA linkage, mass spectrometry was used. Mass-shift motifs observed in a mass spectrum taken of the products of an RNA-nuclease P1 interaction in the presence of oxygen-18 enriched water allowed for molecules of the 5' and 3' ends of RNA to be clearly distinguished[9].

The existence of CoA-linked RNA has implications in many biochemical pathways. Current knowledge of RNA is expansive; however, this discovery unveils many possibilities for future research. While the function of CoA-RNA is still unknown, one can look to other sources of RNA alteration to understand the functional possibilities of CoA-RNA. In eukaryotes, mRNA is seen to contain a 7-methylguanyltae cap in addition to a poly(A) tail; these RNA alterations are important in RNA stabilization, preventing degradation from 5' exonucleases before translation into protein. Other RNAs, such as miRNA or siRNA, interfere with certain biological processes occurring within a cell such as gene expression. CoA-RNA may have similar functions in regulation of genes or protein products, or it may serve some protectory function similar to the m^7G cap observed in eukaryotes[13].

Medium Chain Fatty Acid-Coenzyme A Ligase (FadK)

Medium chain fatty acid-coenzyme A ligase, known as FadK, is encoded by the gene of the same name (*FadK*). FadK is an acyl-CoA synthetase, utilizing ATP to activate fatty acids. Fatty acid activation enables the molecules to undergo β-oxidation and eventually become incorporated into the citric acid cycle, providing an alternate

source of energy for organisms. As indicated by its name, FadK operates most efficiently when used in conjunction with medium chain fatty acids, which are fatty acids typically containing aliphatic tails between 6 and 12 carbons in length. The low substrate specificity of FadK is important to this research, as it is hypothesized that the proteins activity over a large range of substrates will allow for significant modification to potential substrates by way of biotin tagging.

FadD, another acyl-CoA synthetase, is useful for studying FadK, as the enzymes often have similar functionality and follow the same general reaction scheme. In research performed by Morgan-Kiss et al., FadK was explored alongside FadD in terms of gene expression, protein expression and purification, and enzymatic activity[12]. The esterification reaction that FadK catalyzes is shown in Figure 1.1 and Figure 1.2 below. A generic six carbon fatty acid is used in the scheme for simplicity.
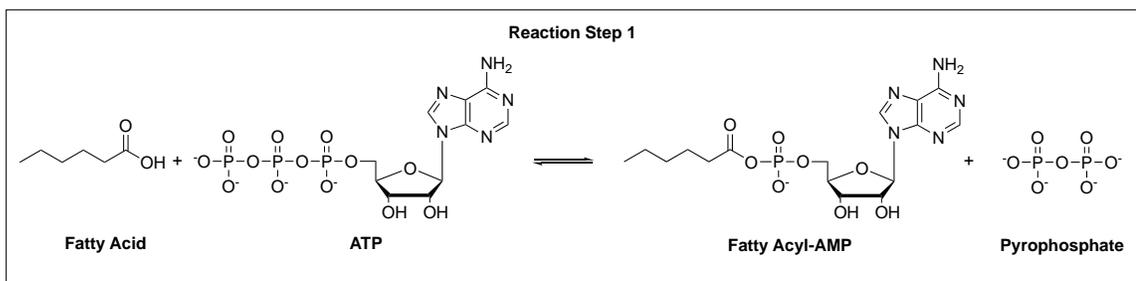

Figure 1.1. Formation of a fatty acyl-AMP intermediate in the FadK esterification reaction.
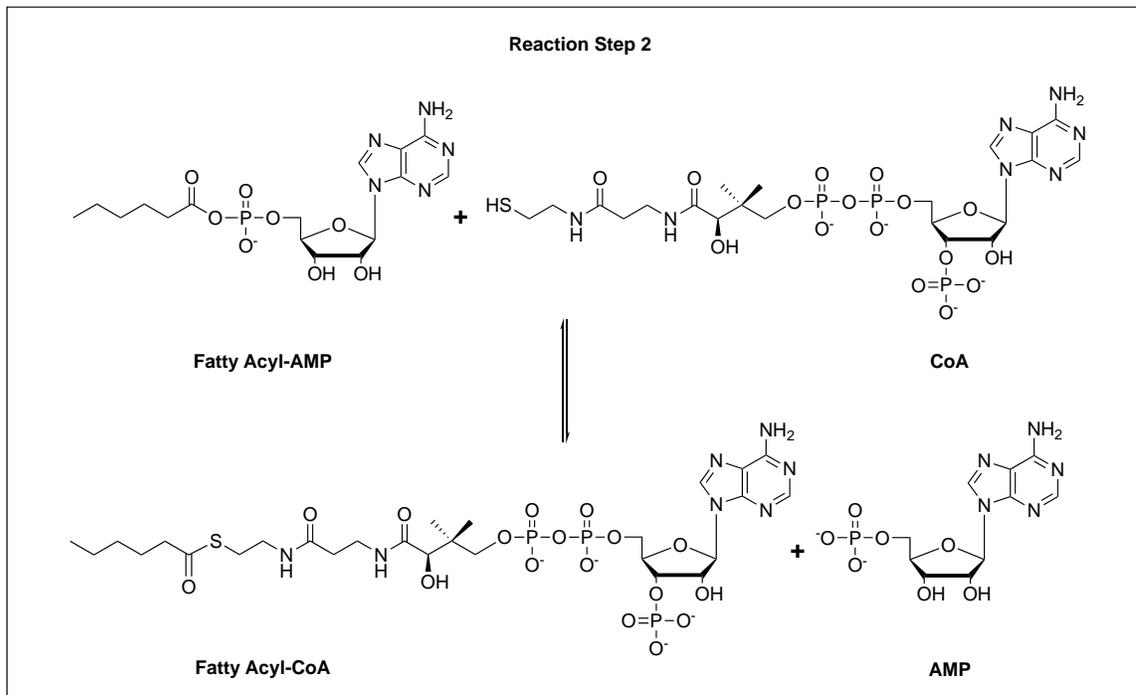
Figure 1.2. Formation of a fatty acyl-CoA from the fatty acyl-AMP intermediate.

Figure 1.1 shows the formation of the fatty acyl-AMP intermediate as well as the

resulting inorganic pyrophosphate. The intermediate formed during this reaction is the

same intermediate observed in various other biochemical reactions such as in the

transportation of fatty acids. The conversion of the fatty acyl-AMP into the target fatty

acyl-CoA, important in fatty acid metabolism, is seen in **Figure 1.2**.

Although **Figure 1.1** and **Figure 1.2** show the incorporation of a 6 carbon

substrate, FadK shows optimal activity with any fatty acids between 6 and 12 carbons in

length. Of these substrates, FadK was most active when used in conjunction with 6

carbon fatty acids as well as 8 carbon fatty acids. It is for these reasons that FadK appears

to be the perfect candidate for use in the isolation of CoA-RNA via the formation of a

biotinylated CoA-RNA molecule. In quantitating the activity of FadK, researchers used

the following measurement: nanomoles of product formed per minute of reaction per mg

of enzyme. In doing so, a direct comparison between the enzyme's activity and the

amount of enzyme could be made. With 6 carbon fatty acids, FadK had an activity of

2.75 nm/min/mg, whereas with 8 carbon fatty acids its activity was 3.75 nm/min/mg.

Reaction conditions for FadK favored a pH between 7.7 and 8.2[12]. Because enzyme

kinetics depend on a combination of factors like substrate and enzymatic concentration,

temperature, and pH, values outside of these ranges drastically affect the ability of the

enzyme to perform to standards. Temperature and concentration are key factors here, as

FadK appears to have stability issues when either of these values is higher.

Cloning, Expression, and Purification of FadK

Newly developed *in vivo* laboratory cloning methods allowed for quick and

simple cloning of *FadK*, with a functional plasmid being obtained within 2-3 days of

work. The underlying principles of the cloning method relied on the use of *E. coli* to

provide the machinery needed for cloning. DNA fragments designed to have between 9

and 25 overlapping nucleotides and produced via PCR resulted in DNA that was easily

taken up by *E. coli* bacteria via transformation, resulting in the desired plasmids for

experimentation[7].

Purification of FadK is a complicated task, as cold conditions are important for

the enzymatic stability of FadK in solution. In this experiment, it was important to follow

the procedure outlined by Morgan-Kiss, et al. as accurately as possible[12]. For instance,

FadK had to be purified at cold temperatures (5° Celsius or colder). Performing

purification at normal room temperatures had potential to yield an aggregated protein as

opposed to a soluble one. As well, a six-membered histidine tag $(His)_6$ was placed on the N-terminal of the FadK protein via plasmid design for purification. Tagging the protein with a histidine tag was important in the isolation of the protein, acting as a tether to separate it from solution; however, it is possible that the inclusion of the $(His)_6$ tag may interrupt protein migration through SDS-PAGE, hindering analysis[12].

FadK activity, while at acceptable levels for the overall experimental goal, was not up to the level of activity seen with the related FadD protein. It is thought that the reason for the lower activity level observed with FadK is due to the study of FadK as a protein in solution as opposed to a membrane-soluble protein, as is the case with FadD. However, further research needs to be performed in order to resolve this issue. Activity measurement of FadK also proved to be difficult, as protein concentrations that approached 2 μg/mL resulted in decreased enzymatic activity while protein concentrations generally greater than 5 μg/mL resulted in protein insolubility. Similar losses in protein activity were seen in longer storages at temperatures colder than -20º Celsius[12]. The sensitivity of FadK to slight changes in concentration or temperature once again may be a result of inconsistencies between our methods of studying the protein and its actual interactions *in vivo*.

The Broad Functionality of CoA-RNA

Due to the reactivity of free thiol groups, like that of CoA, thiol-containing molecules are prime targets for utilization in larger molecule conjugates involving, for instance, synthesized macromolecules, biological macromolecules, and even gold nanoparticles[13]. Thiol activated RNA is one such example of a thiol-containing biological

macromolecule. With the recent discovery of CoA-RNA, it is apparent that this CoA-linked RNA can have similar functions. The implications of the inherent functionality of CoA's thiol group are significant: thiol-containing molecules can be used in the aforementioned molecule conjugates, allowing them to have the ability to be directed in different ways both *in vitro* and *in vivo*. For instance, a therapeutic biomolecule conjugated with folate has the ability to target cancer[16]. Moreover, other biomolecules can be conjugated with gold for immobilization[11]. If CoA-RNA were found to have gene expression functionalities, conjugated CoA-RNA could potentially be directed to sites in the body to alter gene expression or hinder protein activity. Of course, depending on the different functions that CoA-RNA has, this may or may not be possible. Regardless, the significance of the free thiol group of CoA opens up many future possibilities for the molecule.

The use of polymers in conjugation with thiol-containing RNA and other biomolecules is another area in research currently being explored. RAFT, or reversible addition-fragmentation transfer polymerization, is used with aminolysis techniques to form larger conjugates with biomolecules. Generally, the linkage between the two conjugate members is a disulfide[2]. Interestingly, the presence of the disulfide bond allows the overall conjugate to essentially fall apart under reducing conditions, dropping the biomolecule off at the location of reduction. This can also be seen in conjugates based from living systems, such as thiol-adenosine conjugates[6]. Thus, if the biomolecule also serves to be therapeutic, there are obvious benefits to this method of conjugation, as some cytoplasmic processes contain molecules capable of reduction[2]. The key to the use of RAFT and aminolysis, then, is to find a molecule that will aid in the delivery of the

therapeutic agent to its target zone, but this is not quite so simple[4]. Gold nanoparticles are one method to deliver therapeutic thiol-containing molecules such as thiol-RNA to their destination; when seen in conjunction with RNA molecules, the gold nanoparticles act to improve RNA stability[3,14]. While other methods of delivery exist, the use of gold nanoparticles and RAFT technology appears to be quite effective when used in conjunction with thiol-containing RNA and potentially CoA-RNA.

Chapter III: Materials and Methods

Reagents

DNA oligonucleotides were obtained from Integrated DNA Technologies. PCR and DNA recombination reagents (dNTPs, plasmid backbones, DNA markers, Taq polymerase) were obtained from New England Biolabs. Bacto^{TM} Agar was purchased from BD. Chemical and biochemical reagents (Tris, Tris-HCl, Borate, EDTA, ethanol, TEMED, SOC media, IPTG, $NiCl_2$, Na_2HPO_4, NaH_2PO_4, kanamycin, agarose, SDS, acrylamide, sodium chloride, tryptone, yeast extract, and imidazole) were purchased from Sigma-Aldrich. GelCode^{TM} Blue Safe Protein Stain and Top10 cells were obtained from Thermo Fisher Scientific. DI water was supplied by the University of Southern Mississippi.

1.1 Preparation of Petri Dishes

For all petri dish plating performed over the course of experimentation, agar gels were made by mixing 3.0 g of tryptone, 1.5 g of yeast extract, 3.0 g of NaCl, and 4.5 g of Bacto^{TM} Agar with 200 mL of water. The broth was stirred for 5 minutes and another 100 mL of water was added. After the broth was autoclaved, kanamycin was added and the petri dishes were filled with a thin layer of agar broth, which then cooled to form a gel proper for bacterial plating. When plating, ethanol-soaked beads were rinsed with water to remove residual ethanol and used to spread bacteria sample around evenly on the petri dish, making sure to re-soak the beads with ethanol afterwards.

1.2 Preparation and Running of Gels

Agar gels were made using 0.3 g of agarose in 30 mL of TBE buffer. The gels were heated in a beaker within a microwave, making sure not to let the contents of the beaker spill over the edges. Gels were stained afterwards with 3 μL of ethidium bromide dye and then poured into a comb-containing mold to cool.

SDS-PAGE resolving gels were made with 4.1 mL H_2O, 3.3 mL acrylamide, 2.5 mL Tris-HCl, 100 μL 10% SDS, 10 μL TEMED, and 32 μL APS. TEMED and APS were added immediately before pouring the gel to prevent it from polymerizing too quickly. Between 2 to 3 centimeters were left above the resolving gel for the stacking gel. Ethanol was used to cover the gel and prevent it from drying out. After waiting roughly 30 minutes for the resolving gel to polymerize, SDS-PAGE stacking gels were made with 6.1 mL H_2O, 1.3 mL acrylamide, 2.5 mL Tris-HCl, 100 μL 10% SDS, 10 μL TEMED, and 100 μL APS. Combs were added and the gel was allowed time to polymerize.

For DNA gel electrophoresis, the gel was run over a period of 30 minutes with the following readings: voltage: 100 volts; current: 40 milliamps; power: 4 Watts. For SDS-PAGE, gels were left to run at 150 volts and 400 milliamps for 50 minutes. After completion, gels were washed at least 3 times with water to remove any residual SDS and then stained with GelCode^{TM} Blue Safe Protein Stain. Enough protein stain was used to barely cover the gels. The gels were stained for an hour, then washed at least 3 times with water. Afterwards, the gels were left overnight on a shaker at a gentle rocking speed and viewed the following day.

1.3 Preparation of Polymerase Chain Reaction (PRC) Mixtures

PCR mixtures were prepared using a 11.5 µL mixture containing 10.0 µL of 2x laboratory stock, 0.5 µL of 10f RE-DRA$_{a1}$, and 1.0 µL of ET28$_{f1}$/BK$_{rr}$ primers. The PCR was initially run in the Bio-Rad thermal cycler for 2 minutes at 95º Celsius, followed by 8 cycles of the following: 10 seconds at 98º Celsius, then another 10 seconds at 62º Celsius, and finally 160 seconds at 72º Celsius. Gel electrophoresis allowed for the confirmation of the desired BK$_0$ PCR products.

1.4 Preparation of Lysogeny Broths (LB Broths)

LB broths were made using tryptone, yeast extract, sodium chloride, and DI water. The standard recipe for 1.0 L of lysogeny broth called for 10 g of tryptone, 5 g of yeast extract, 10 g of NaCl, and 1 L of DI water. After measuring each ingredient, DI water was added, and the solution was stirred until homogenous. Afterwards, the broth was autoclaved. Kanamycin was added to broths when needed, in an amount proportional to the total volume (lab-made kanamycin was typically at a concentration of 50 mg/mL).

1.5 Preparation of the Washing Buffer, Elution Buffer, and Nickel Column

For protein elution, a washing buffer and elution buffer were made. A 50 mM EDTA solution, to be used in the elution buffer, was made by mixing 0.66 g of EDTA in 35.0 mL of distilled water. A washing buffer (pH 8.0), hereby referred to as Buffer 1, was made using 0.67 g of Na$_2$HPO$_4$, 37 mg (0.037 g) of NaH$_2$PO$_4$, 1.8 g of NaCl, 10 mM imidazole, and 100 mL of distilled water. An elution buffer (pH 8.0), hereby referred to as Buffer 2, was made using 10 mM Tris, 10 mM EDTA (diluted from 50 mM stock),

0.5% SDS (500 μL or 0.5 mL of 10% SDS), 250 mM imidazole, and distilled water. A nickel column was prepared by washing with three full column volumes of distilled water, one column volume of NiCl$_2$, and then two more column volumes of distilled water.

2.1 Cloning of the *FadK* Gene and Growth of Bacterial Colonies

A 50.0 μL aliquot of Top10 cells were thawed on ice until the last ice crystal melted. At the same time, SOC media was removed from the -20º Celsius freezer and incubated in 37º Celsius. Afterwards, 1.0 μL of BK plasmid backbone and 1.0 μL of FadK DNA were pipetted into the cell aliquot, and the mixture was mixed by gentle tapping. It is important to not pipette the mixture up and down, and not to vortex it, as this can potentially lyse the cells. The mixture was incubated on ice for 30 minutes exactly. The bacteria were then heat shocked at 42º Celsius for a duration of 35 seconds, and then chilled on ice for 2 minutes. SOC media was added to the mixture in the amount of 120 μL, and the entire solution was incubated at 37º Celsius for 1 hour on a shaker. The bacteria solution was then spread on an LB/Kan agar plate and grown overnight at 37º Celsius. The next day, a number of colonies were picked up, labeled, and then used to inoculate multiple LB/Kan media-containing test tubes. Each colony was allowed to grow in its respective media for roughly 8 hours.

2.2 Isolation of Bacterial Plasmids and Subsequent Confirmation via Restriction Digestion

Each tube of the inoculated media was then centrifuged over multiple turns of resuspension, lysis, neutralization, wash, and elution buffers in accordance with the protocol outlined in the Epoch Life Science DNA Purification Kit. The restriction enzymes NdeI and SmaI were then used to digest the DNA obtained from the centrifuged cells. Gel electrophoresis was performed as outlined in **Section 1.2** of **Materials and Methods**.

3.1 Expression of FadK within Bacterial Cells and Confirmation of Protein Expression

Any colony that contained the correct FadK plasmid, as confirmed by use of restriction digestion, was allowed to grow in roughly 5.0 mL of an LB/Kan media at 37º Celsius in a shaker until an optimal optical density had been reached (between 6 to 8 hours of growth). To the 5.0 mL of LB/Kan media, 2.5 μL of 1.0 M IPTG was added (to give an overall molarity of 0.5 mM), and the solution was put in a shaker at 17º Celsius and left overnight to allow for optimal protein expression. The following day, the sample(s) were spun down and the supernatant was discarded. It is important to save some sample in the event that there is protein expressed or an error in experimentation. An SDS-PAGE was performed as outlined in **Section 1.2** of **Materials and Methods**, confirming the presence of protein. It is important to use a relatively small volume of cells (roughly 50 μL), as higher volumes tend to lead to stickier samples and difficulty in loading gels.

3.2 Large-Scale Expression of FadK and Subsequent Centrifugation for Protein Harvesting

For large-scale protein expression, an LB broth was made in the amount of 2 L using the procedure outlined in **Section 1.4** of **Materials and Methods**. Around 400 mL of the broth was transferred to an Erlenmeyer flask, and 400 µL of 50 mg/mL kanamycin was added. Stored bacteria from the previous day were then dumped into the flask, and the inoculated LB/Kan broth was allowed to grow in a shaker at 37º Celsius until an optimal optical density had been reached. Once the desired optical density was reached, 200 µL of 1.0 M IPTG was added to the flask to induce expression, and the flask was left overnight in a 16º Celsius shaker. To harvest the protein, the broth was first split into two equal volumes of 200 mL and centrifuged at 4,000 x G for 10 minutes to remove the cells from solution. Following centrifugation, 20 mM of cold Tris-HCl was used to resuspend the pelleted cells; the solution was then recentrifuged at 4,000 x G again, but for 5 minutes. Any supernatant was drained, and the pellets were stored in -80º Celsius until the protein could be eluted.

3.3 Elution and Purification of FadK Protein

Due to the relatively low protein stability of FadK, steps taken towards elution and purification were carried out between 0º Celsius and 5º Celsius. To begin protein harvesting, pellets were obtained from -80º Celsius and thawed on ice. Once thawed, cells were sonicated for membrane lysis. Using a previously prepared nickel column, protein-containing fractions were washed with five column volumes of Buffer 1. About

two or three column volumes of Buffer 2 were used to elute FadK, and the resulting

fractions were stored in 1.5 mL microcentrifuge tubes. FadK was then purified of any

contaminants and then stored in glycerol for protein security. A gel containing each

elution fraction and elution wash was ran to confirm the presence of protein, as well as

for troubleshooting purposes in the case of experimental error.
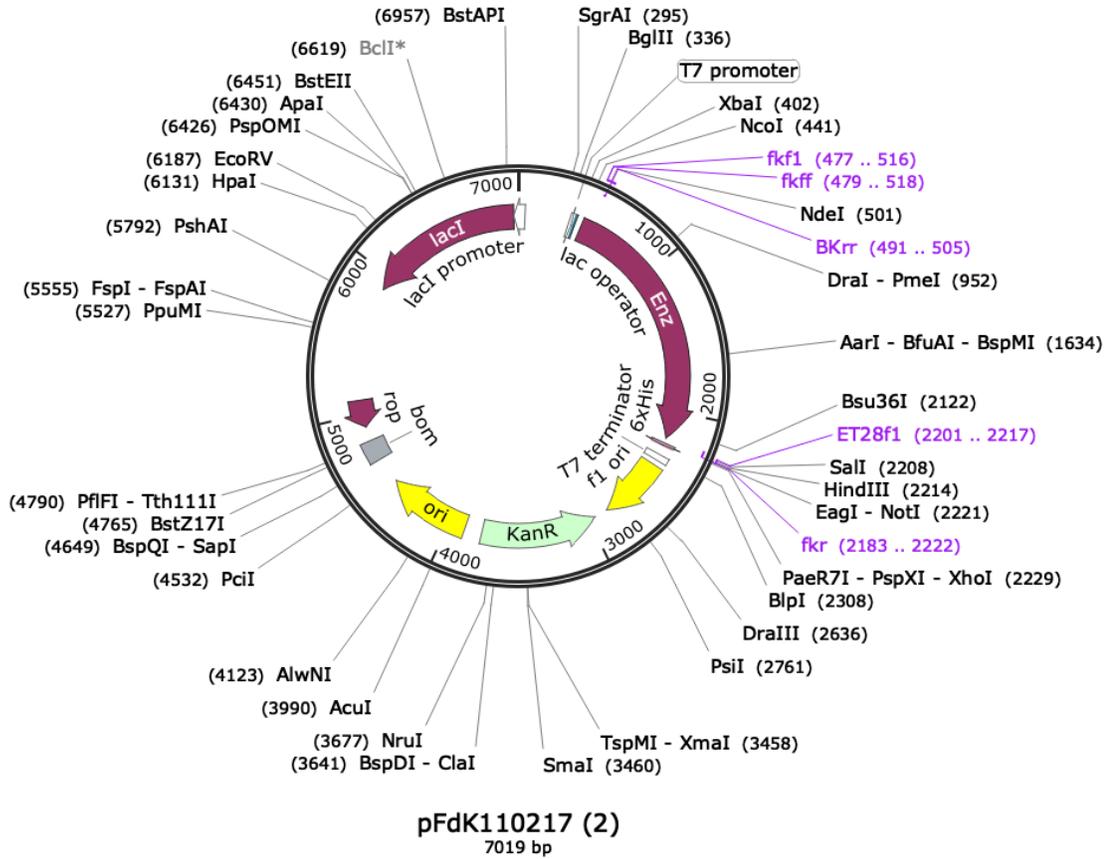
Chapter IV: Data and Results

Figure 2.1. Designed FadK plasmid.

Figure 2.1 shows the FadK plasmid designed for experimentation with FadK. The overall plasmid size is just over 7 kb. LacI was included to give us the ability to induce expression of *FadK* upon addition of IPTG. KanR was included to provide plasmid-containing bacteria the ability to resist kanamycin. Restriction sites of NdeI (at base pair 501) and SmaI (at base pair 3460) can be seen in the image.
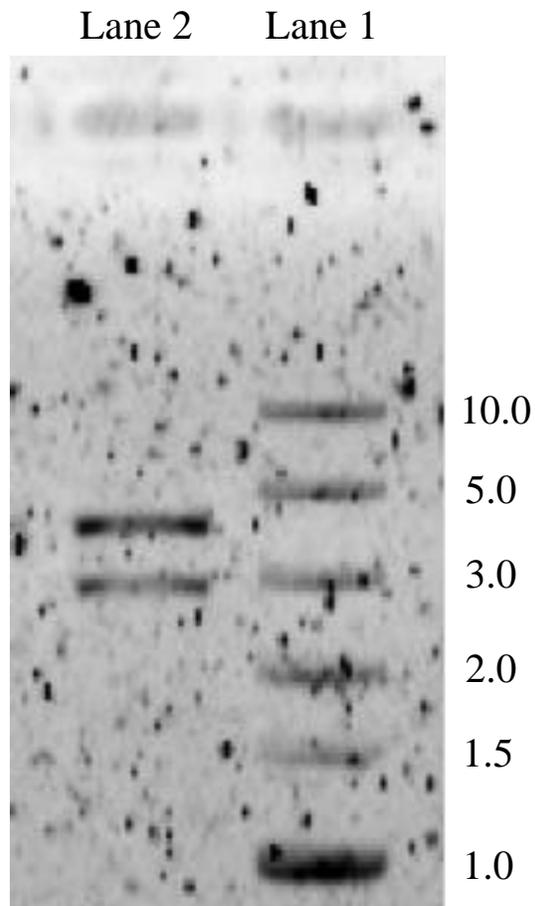
Lane 2   Lane 1

10.0

5.0

3.0

2.0

1.5

1.0

Figure 2.2. Agarose gel image of 7 kb FK2 (FadK plasmid) restriction digestion.

Colonies were picked up and labeled FK1 and FK2. A 1 kb fast DNA ladder was

used for reference and is seen in Lane 1. FK 2 is the only colony that was positive for the

colony and is shown in Lane 2 in **Figure 2.1**. Restriction enzymes NdeI/Sma1 were used

for digestion, giving two DNA fragments of 3 kb and 4 kb. Lane 1 corresponds to the
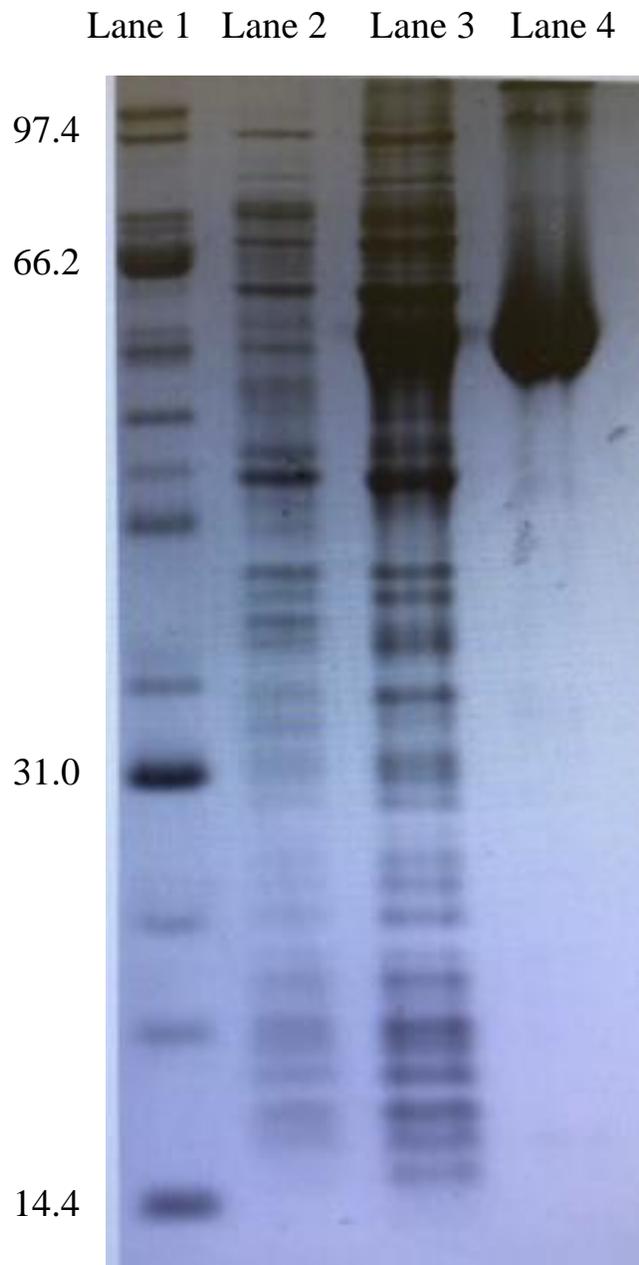
DNA ladder.

Lane 1   Lane 2   Lane 3   Lane 4

97.4
66.2
31.0
14.4

Figure 2.3. FadK purification (SDS-PAGE).

Lane 1 corresponds to the protein ladder, whereas Lane 2, Lane 3, and Lane 4 correspond to uninduced sample, induced sample, and the purified FadK protein. The protein ladder is measured in kDa. FadK is seen to have a molecular weight of roughly 62 kDa, corresponding to accepted values.

Chapter V: Discussion and Conclusion

The discovery of CoA attached to the 5' end of RNA has given rise to many questions, most of which are still largely unanswered. Although it shares structural similarities (namely, the adenosine moiety) with other coenzyme-linked RNAs, isolation of CoA-RNA has shown to be difficult in comparison. FadK, due to its broad substrate specificity and relative simplicity in expression, is a promising solution to the problem we currently face in effectively isolating CoA-linked RNA. If successfully isolated, CoA-linked RNA can be structurally elucidated and its functions identified.

In the cloning, expression, and purification of FadK, gel scans were the primary means by which data was recorded. To confirm the presence of the desired *FadK* gene plasmid in bacterial colonies, a few colonies were first selected and grown in media. The plasmids were isolated from these samples and analyzed using gel electrophoresis and restriction digestion, as seen in **Chapter IV: Data and Results**. SDS-PAGE was utilized to confirm the occurrence of protein expression in bacterial cells that were established to have the correct plasmid. Finally, SDS-PAGE was again used to confirm that the protein was present in solution after protein harvesting and elution. The designed FadK plasmid can be seen in **Figure 2.1**. LacI was included in the plasmid to give bacteria responsiveness to IPTG (a lactose analog), allowing for the expression of *FadK*. Kanamycin resistance was also added in the form of a KanR open reading frame, allowing for the isolation of bacteria containing only the desired plasmid. Restriction digestion was performed using NdeI and SmaI, shown in **Figure 2.2**. Cutting at base pairs 501 and 3460, respectively, these enzymes gave DNA plasmids of 3 kb and 4 kb.

Gel electrophoresis showed the presence of DNA fragments at 3 and 4 kb on the DNA latter, which is to be expected with FadK, thus confirming the presence of the plasmid. **Figure 2.3** shows purified FadK alongside a protein ladder, an uninduced sample, and an induced sample. A strong band is visible in lane 4, belonging to the purified FadK; however, a clear migration was not seen in the band, most likely due to the presence of the $(His)_6$ tag[12]. Despite this, there is a noticeable difference between the purified protein and the induced and uninduced samples.

Despite our ability to successfully purify FadK, limitations in experimentation still existed. For instance, the expression of FadK was initially unsuccessful due to a defective reagent (IPTG); this issue was corrected, only for further complications to arise in the stability of FadK during protein elution. Later in the experiment, difficulties in direct measurement of FadK activity meant that we were unable to assess enzymatic efficiency. Nevertheless, the goals of the experiment were achieved, and it was performed to satisfaction, proving it to be successful. Although the purpose of this experiment was to effectively clone, express, and purify FadK, we experienced the added benefit of not only advancing scientific understanding of FadK, but also providing resources in the current work to elucidate the structure of CoA-linked RNA.

Future researchers have many angles to improve upon the current knowledge available about FadK and CoA-RNA. FadK can be relatively unstable, depending on enzyme concentrations present in solution and the temperature of its environment[12]. Methods to improve FadK stability may prove to be fruitful, having applications with other enzymes beyond FadK. Furthermore, because there exists no simple way to quantitatively monitor the enzymatic activity of FadK, there is a need for more effective

methods of assessing the catalytic ability of it and similar proteins. As such, we hypothesized that AMP deaminase may be used to measure the activity of FadK as it activates fatty acids into their fatty acyl-CoA counterparts, as seen in **Figure 1.2**. AMP deaminase converts AMP into IMP, resulting in an altered conjugation that allows for the measurement of detectable changes in UV absorbance. Research into the cloning, expression, and purification of AMP deaminase could have broad applications in quantitating the activity of enzymes involved with fatty acid metabolism and other carboxyl group activation mechanisms. The purification of AMP deaminase was briefly explored in our laboratory; however, in addition to lack of research surrounding AMP deaminase purification, the cloning and expression of the enzyme proved to be too difficult due to the nature of the enzyme. Other acyl-CoA synthetases, such as FadD, can also be exploited in the similar ways to how we are using FadK, assuming there exists a comparable broad substrate specificity. Of course, this is in no way an exhaustive list: other approaches may be taken in respect to the isolation and structural elucidation of CoA-RNA, depending on the available literature as well as laboratory resources.

References

(1) Bird, J. G.; Zhang, Y.; Tian, Y.; Panova, N.; Barvík, I.; Greene, L.; Liu, M.; Buckley, B.; Krásný, L.; Lee, J. K.; et al. The Mechanism of RNA 5′ Capping with NAD+, NADH, and Desphospho-CoA. *Nature* **2016**, *535* (7612), 444–447.

(2) Boyer, C.; Bulmus, V.; Davis, T. P. Efficient Usage of Thiocarbonates for Both the Production and the Biofunctionalization of Polymers. *Macromol. Rapid Commun.* **2009**, *30* (7), 493–497.

(3) Giljohann, D. A.; Seferos, D. S.; Prigodich, A. E.; Patel, P. C.; Mirkin, C. A. Gene Regulation with Polyvalent SiRNA−Nanoparticle Conjugates. *J. Am. Chem. Soc.* **2009**, *131* (6), 2072–2073.

(4) Heredia, K. L.; Nguyen, T. H.; Chang, C.-W.; Bulmus, V.; Davis, T. P.; Maynard, H. D. Reversible SiRNA–Polymer Conjugates by RAFT Polymerization. *Chem. Commun.* **2008**, *0* (28), 3245–3247.

(5) Huang, F. Efficient Incorporation of CoA, NAD and FAD into RNA by in Vitro Transcription. *Nucleic Acids Res.* **2003**, *31* (3), e8–e8.

(6) Huang, F.; Shi, Y. Synthesis of Symmetrical Thiol-Adenosine Conjugate and 5′ Thiol-RNA Preparation by Efficient One-Step Transcription. *Bioorg. Med. Chem. Lett.* **2010**, *20* (21), 6254–6257.

(7) Huang, F.; Spangler, J. R.; Huang, A. Y. In Vivo Cloning of up to 16 Kb Plasmids in E. Coli Is as Simple as PCR. *PLOS ONE* **2017**, *12* (8), e0183974.

(8) Jäschke, A.; Höfer, K.; Nübel, G.; Frindert, J. Cap-like Structures in Bacterial RNA and Epitranscriptomic Modification. *Curr. Opin. Microbiol.* **2016**, *30*, 44–49.

(9)   Kowtoniuk, W. E.; Shen, Y.; Heemstra, J. M.; Agarwal, I.; Liu, D. R. A Chemical Screen for Biological Small Molecule-RNA Conjugates Reveals CoA-Linked RNA. *Proc. Natl. Acad. Sci.* **2009**, *106* (19), 7768–7773.

(10) Leonardi, R.; Zhang, Y.-M.; Rock, C. O.; Jackowski, S. Coenzyme A: Back in Action. *Prog. Lipid Res.* **2005**, *44* (2), 125–153.

(11) McCarthy, J. E. G.; Marsden, S.; von der Haar, T. Biophysical Studies of the Translation Initiation Pathway with Immobilized mRNA Analogs. *Methods Enzymol.* **2007**, *430*, 247–264.

(12) Morgan-Kiss, R. M.; Cronan, J. E. The Escherichia Coli FadK (YdiD) Gene Encodes an Anerobically Regulated Short Chain Acyl-CoA Synthetase. *J. Biol. Chem.* **2004**, *279* (36), 37324–37333.

(13) Sapkota, K.; Huang, F. Efficient One-Pot Enzymatic Synthesis of Dephospho Coenzyme A. *Bioorg. Chem.* **2018**, *76*, 23–27.

(14) Vázquez-Dorbatt, V.; Tolstyka, Z. P.; Chang, C.-W.; Maynard, H. D. Synthesis of a Pyridyl Disulfide End-Functionalized Glycopolymer for Conjugation to Biomolecules and Patterning on Gold Surfaces. *Biomacromolecules* **2009**, *10* (8), 2207–2212.

(15) White, H. B. Coenzymes as Fossils of an Earlier Metabolic State. *J. Mol. Evol.* **1976**, *7* (2), 101–104.

(16) York, A. W.; Huang, F.; McCormick, C. L. Rational Design of Targeted Cancer Therapeutics through the Multiconjugation of Folate and Cleavable SiRNA to RAFT-Synthesized (HPMA-s-APMA) Copolymers. *Biomacromolecules* **2010**, *11* (2), 505–514.