Charles University

Faculty of Science

Study programme: Cell Biology

Bc. Jan Říha

tRNA synthetases as potential RNA capping enzymes

tRNA syntetázy jako potenciální RNA čepičkovací enzymy

Diploma thesis

Supervisor:

Ing. Hana Macíčková Cahová, Ph.D.

Prague, 2024

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 29.4.2024

Jan Říha

Acknowledgements:

I am grateful to my supervisor, Dr. Hana Macíčková Cahová, for her advice and enthusiasm throughout this project. I extend my thanks to all members of the Cahová group for their assistance, with special appreciation to Valentina Serianni, Dr. Ondřej Nešuta, and Dr. Pavel Vopálenský for their guidance and support in the laboratory. I would like to express my appreciation to Dr. Anton Škríba for LC-MS measurements and his expertise in chromatographic methods, as well as to Dr. Martin Lepšík for his contribution to molecular dynamics simulations and to Dr. Radek Pohl for conducting the NMR measurements.

Abstract (EN)

A novel type of non-canonical 5′ RNA caps, dinucleoside polyphosphate RNA caps, has recently been discovered in our laboratory[1]. To elucidate the physiological role of these caps, more detailed information about their synthesis is essential. It has been already described that dinucleoside polyphosphate RNA caps (Np_nN-RNA caps) can be accepted by RNA polymerase during transcription as a non-canonical initiating nucleotide (NCIN)[5,6]. However, the possibility that Np_nN-RNA caps are created post-transcriptionally cannot be ruled out. The best candidates for post-transcriptional capping enzymes are aminoacyl-tRNA synthetases, which, besides their crucial role in translation, are responsible for the synthesis of free dinucleoside polyphosphates[13]. In this thesis, four *E. coli* tRNA synthetases have been selected, cloned into plasmids, and purified using fast protein liquid chromatography (FPLC). Subsequently, selected tRNA synthetases have been tested for the production of free dinucleoside polyphosphate. These experiments have identified the optimal conditions for production of free dinucleoside polyphosphates, diadenosine tetraphosphate (Ap_4A) particularly. The tRNA synthetases were then tested for their capabilities to form an RNA cap on *in vitro* transcribed radioactively labelled RNA. We found that tRNA synthetases are unable to form an Ap_4A-RNA cap and, therefore, do not act as an RNA capping enzyme. These results were confirmed by analysis on liquid chromatography coupled with mass spectrometry.

Keywords: Non-canonical RNA caps, dinucleoside polyphosphates, Ap_4A, Ap_4A-RNA, aminoacyl-tRNA synthetases, LysU

Abstrakt (CZ)

V naší laboratoři byl nedávno objeven nový typ nekanonických RNA čepiček, obsahujících dinukleosid polyfosfáty. Abychom mohli objasnit jejich fyziologickou roli, je potřeba získat více informací o jejich syntéze v buňkách. Na základě nedávno publikovaných výsledků, dinukleosid polyfosfátové RNA (Np_nN-RNA) mohou vznikat při transkripci přímou inkorporací dinukleosid polyfosfátů pomocí RNA polymerázy na počátek nově vznikajícího transkriptu (takzvaný nekanonický iniciační nukleotid). Nelze vyloučit i jiné možnosti syntézy Np_nN čepičkované RNA, především jejich enzymatickou post-transkripční syntézu. Jedny z nejpravděpodobnějších kandidátů pro enzym vytvářející Np_nN-RNA jsou aminoacyl-tRNA syntetázy. Tyto enzymy jsou, kromě své esenciální role v translaci, odpovědny za syntézu volných dinukleosid polyfosfátů v buňce. V této práci jsme vybrali čtyři tRNA syntetázy, klonovali je do plazmidů a purifikovali pomocí proteinové kapalinové chromatografie (fast protein liquid chromatography – FPLC). Následně byly tRNA syntetázy testovány pro syntézu volných dinukleosid polyfosfátů, díky čemuž jsme zjistili optimální reakční podmínky. Za daných reakčních podmínek jsme následně testovali, zda jsou tRNA syntetázy schopny vytvořit Np_nN-RNA, a to pomocí reakce tRNA syntetáz s radioaktivně značenou RNA. Získané výsledky naznačují, že tRNA syntetázy nemohou vytvářet RNA čepičkovanou s dinukleosid polyfosfáty. Tyto výsledky byly následně potvrzeny pomocí analýzy založené na kapalinové chromatografii spojené s hmotnostní spektrometrií.

Klíčová slova: Nekanonické RNA čepičky, dinukleosid polyfosfáty, Ap_4A, Ap_4A-RNA, aminoacyl-tRNA syntetázy, LysU

List of abbreviations

aaRS	aminoacyl-tRNA synthetase (also referred to as tRNA synthetase)
LysRS	lysyl-tRNA synthetase
	(for other aminoacyl-tRNA synthetases, similar abbreviation can be created using three letter name of an amino acid and RS, which stands for the tRNA synthetase; e. g. MetRS - methionyl-tRNA synthetase)
Ap_4A	diadenosine tetraphosphate
Ap_3A	diadenosine triphosphate
Ap_4G	adenosine-guanosine tetraphosphate
Ap_4C	adenosine-cytidine tetraphosphate
Ap_4U	adenosine-uridine tetraphosphate
Gp_4G	diguanosine tetraphosphate
Np_nN	dinucleoside polyphosphate
ATP	adenosine 5′-triphosphate
GTP	guanosine 5′-triphosphate
CTP	cytidine 5′-triphosphate
UTP	uridine 5′-triphosphate
aa	amino acid
aa-AMP	amino acid – adenosine 5′-monophosphate intermediate
tRNA	transfer ribonucleic acid
mRNA	messenger RNA
rRNA	ribosomal RNA
NAD	nicotinamide adenine dinucleotide
FAD	flavin adenine dinucleotide
CoA	coenzyme A
NCIN	non-canonical initiating nucleotide
CBC	cap binding complex
RppH	RNA 5′ pyrophosphohydrolase
NudiX	nucleotide diphosphate linked to X
HIT	histidine triad
DXO	decapping exoribonuclease
ApaH	Ap_4A hydrolase

MSC	multi synthetase complex
KARS	human lysyl-tRNA synthetase
IMPDH	inosine-5′-monophosphate dehydrogenase
AdT	amidotransferase
MITF	microphthalmia-associated transcription factor
MAP	mitogen-activated protein kinase
PPAT	phosphopantetheine adenylyltransferase
RIG-I	retinoic acid-inducible gene I
HCV	hepatitis C virus
ACB	anticodon binding domain
CBS	cystathionine β-synthase domain
FPLC	fast protein liquid chromatography
IMAC	immobilized metal ion affinity chromatography
SEC	size exclusion chromatography
HPLC	high-performance liquid chromatography
LC-MS	liquid chromatography coupled with mass spectrometry
SPE	solid phase extraction
IVT	*in vitro* transcription
PAGE	polyacrylamide gel electrophoresis
NMR	nuclear magnetic resonance
MD	molecular dynamics
LB media	Luria–Bertani medium (or lysogeny broth)
EDTA	ethylenediaminetetraacetic acid
IPTG	isopropyl β-D-1-thiogalactopyranoside
OD	optical density
DMSO	dimethyl sulfoxide
DTT	dithiothreitol
SDS	sodium dodecyl sulphate
TBE	Tris/Borate/EDTA
TEAA	triethyl ammonium acetate
APS	ammonium persulfate
TEMED	tetramethyl ethylenediamine

Table of Contents

Introduction

1. Non-canonical RNA caps

Ribonucleic acid (RNA) molecules represent one of the most important biopolymers within the cells of every living organism. They play a pivotal role in the processing of genetic information from DNA into functional proteins. RNA is consisting of four nucleotides: adenosine, guanosine, cytidine, and uridine, but to fulfil specialized functions, chemical modifications of RNA molecules are essential. Currently, more than 170 RNA modifications have been identified[14]. Research in our group focuses on the modifications of 5′ end of RNA. These modifications are called RNA caps and have an important function in regulating RNA stability, translation, or nuclear export. Historically, only one modification of the 5′ end was known, namely the eukaryotic canonical messenger RNA (mRNA) cap. Recently, new group of the 5′ end modifications was discovered. These modifications are mostly composed of metabolite molecules on 5′ end, therefore they are called non-canonical metabolite RNA caps.

1.1. The canonical mRNA cap

The canonical eukaryotic mRNA cap is composed of guanosine methylated at position 7 connected by three phosphates to the 5′ end of RNA in a 5′ to 5′ orientation (m^7Gp_3N cap – structure of mRNA cap can be seen in Figure 1.)[15]. This cap plays an important role in mRNA metabolism, nuclear export and start of translation.

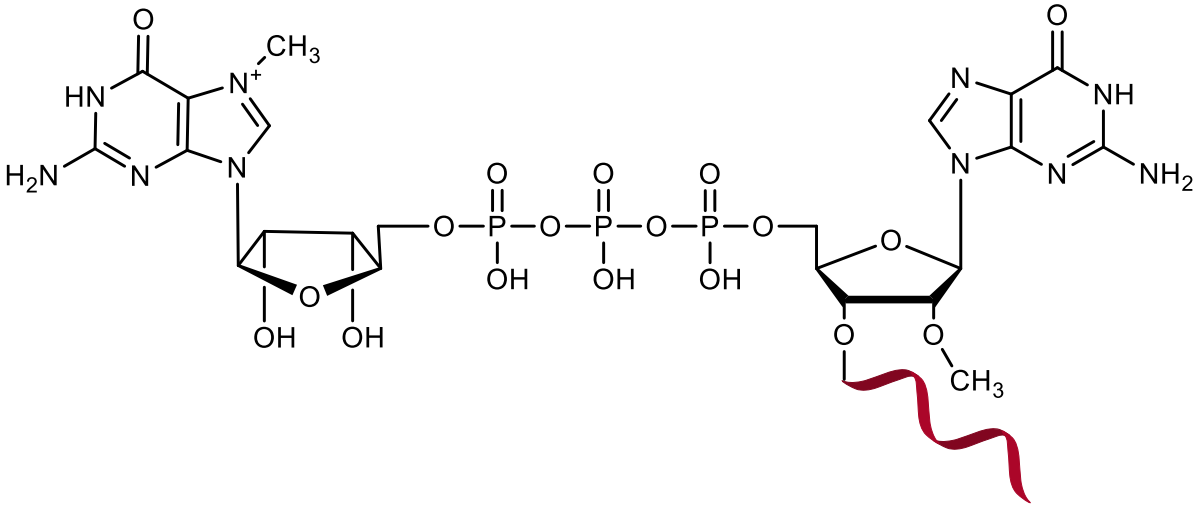


Figure 1.: Canonical mRNA cap (cap 1)

Following the initiation of mRNA synthesis by RNA polymerase II, triphosphate on 5′ end of mRNA is immediately modified by enzymes associated with RNA polymerase

II[16]. First, RNA triphosphatase removes one phosphate, leaving diphosphate RNA, after which mRNA guanylyltransferase attaches GTP in 5′ to 5′ orientation. In the end, guanosine is methylated at position 7 by mRNA methyltransferase.

The canonical mRNA cap is then recognized by the cap binding complex (CBC), which is responsible for the export of mRNA from the nucleus to the cytoplasm[17]. In the cytoplasm, CBC is replaced by eukaryotic initiation factors eIF4E and eIF4G. The mRNA bound to these initiation factors is then recognized by other initiation factors and ribosome, initiating translation. The canonical mRNA cap also plays an important role for mRNA stability, protecting mRNA from degradation by exonucleases and regulates mRNA stability based on presence of bound initiation factors or presence of decapping complexes inside the cell[18].

There are several structural variants of the canonical mRNA cap: Eukaryotic mRNA cap 0 is the one already described before (m^7Gp_3N cap)[19]. Cap 1 and cap 2 consist of cap 0 and 2′ O ribose methylation on the first and the first two nucleotides of the RNA, respectively (m^7Gp_3Nm is cap 1 and m^7Gp_3NmNm cap 2). On small nuclear RNAs (snRNAs), there is a particular type of cap, called trimethyl guanosine cap (TMG cap). This cap consists of cap 1 dimethylated on NH_2 group at position 2 of cap-forming guanosine ($m3Gp_3Nm$ cap).

Till recently, it was believed that in bacteria, mRNA molecules lack a specific mRNA cap on 5′ end of RNA. Instead, *E. coli* mRNAs supposed to have a triphosphate at 5′ end. This triphosphate can be cleaved by enzyme RNA 5′ pyrophosphohydrolase (RppH) to monophosphate, which serves as degradation signal[20]. Furthermore, homologs of eukaryotic mRNA capping enzymes have not been identified in bacteria[21].

1.2. NAD-RNA cap

In 2009, Liu *et al*. discovered the first non-canonical RNA cap, a nicotinamide adenine dinucleotide RNA cap (NAD-RNA cap)[22]. They isolated total RNA from *E. coli* and *S. venezuelae*, and by liquid chromatography coupled with mass spectrometry (LC-MS), they discovered NAD molecule linked to RNA. Subsequent experiments, focused on RNA degradation by nucleases, provided evidence that NAD is linked to 5′ end of RNA.

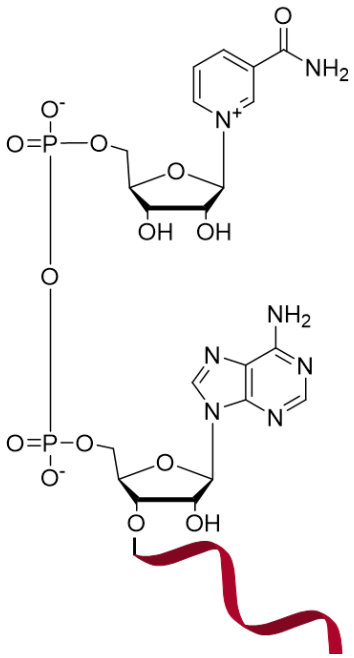


Figure 2.: NAD RNA cap

A breakthrough in the field of non-canonical RNA caps occurred with the development of NAD captureSeq method by Cahová *et al.*[23]. This method selectively enriches RNAs bearing NAD cap, enabling subsequent analysis by sequencing. With this method, it was revealed that NAD-RNA caps in *E. coli* are present mostly on small RNAs. It was shown that this cap is potentially stabilizing RNA because it protects RNA from action of RppH and therefore protects it from degradation. The study also identified an enzyme, NudC, that is responsible for degradation of the NAD-RNA cap. These observations indicate that NAD-RNA cap plays a role in RNA stability.

The synthesis of the NAD-RNA cap is co-transcriptional[24]. During transcription, RNA polymerase can take a molecule of NAD as the first (initiating) nucleotide, which is called non-canonical initiating nucleotide (NCIN). Till today, NAD-RNA cap has been identified in yeast[25], mammals[26], plants[27] and even in archaea[28]. It was shown, that in bacteria, NAD cap does not have any effect on translation[29], but in eukaryotes, NAD-RNA cap cannot be translated, only canonical cap can be translated[26,30]. However, there are some exceptions, NAD-RNA cap can be translated in mitochondria[30], as mitochondria have bacterial type translation machinery, and one report claims that NAD-RNA capped mRNAs can be translated in *Arabidopsis thaliana*[27].

In recent study from Wolfram-Schauerte *et al.*[31], a possible explanation of NAD-RNA caps function was proposed. During T4 bacteriophage infection in *E. coli*, bacteriophages utilize a protein modification known as ADP-ribosylation to reprogram

host's genetic system. Enzymes responsible for ADP-ribosylation use NAD molecule as a substrate. It was shown that these enzymes can utilize NAD-RNA as a substrate and covalently attach RNA to proteins. This modification called RNAylation was shown to be present *in vivo* on specific ribosomal proteins and it is crucial for bacteriophage infection. They demonstrated that only specific RNAs were attached to proteins, the same RNAs that were reported to be NAD capped, which could explain the NAD-RNA cap function in *E. coli*. However, the function of NAD-RNA cap is still not fully understood, and more research is necessary to understanding its role in cells.

1.3. Other types of non-canonical RNA caps

The dephosphocoenzyme A RNA cap (**dpCoA-RNA cap**) was identified by LC-MS analysis of total isolated RNA from *E. coli*, similarly to the discovery of NAD-RNA cap[32]. This cap consists of molecule of dephosphocoenzyme A attached to 5′ end of RNA. However, the abundance of this cap in bacteria is significantly lower compared to NAD cap. The mechanism of synthesis of this molecule was proposed to be the same as for NAD cap[24] (i.e. by non-canonical initiating nucleotide incorporation), but because of the low abundance of dephospho-CoA inside the cell, alternative mechanisms were proposed. One of them is the attachment of CoA to RNA by enzyme phosphopantetheine adenylyltransferase (PPAT), a component of the CoA biosynthetic pathway. In this reaction, the enzyme would take RNA instead of ATP molecule and create an RNA cap[33]. A problem of this proposed mechanism is that RNA is highly unfavourable substrate in comparison to ATP, nonetheless, this mechanism could create a low number of CoA-RNA cap observed in the cell. Reliable CoA capped RNA capturing method was not developed till today, and so the details about physiology of this cap are not known.

Flavin adenine dinucleotide RNA cap (**FAD-RNA cap**) was recently discovered on the 5′ end of RNA of Hepatitis C virus[34]. It was shown that HCV RNA dependent RNA polymerase can accept FAD as NCIN and create FAD cap. This cap was shown to be present on (-) strand in some HCV strains, and on both (-) and (+) strands in other

strains. In one HCV strain, 75% of all HCV RNA was capped, which is the highest ratio of all non-canonical capped transcript known so far. The authors of this study also found that none of human endogenous RNA transcripts were FAD capped. It has been showed that FAD-RNA cap has no effect on RNA stability, but rather protect HCV from immune response (RIG-I induced response against triphosphate 5′ RNA end). However, another recent study has shown that non-canonical RNA caps, including the FAD-RNA cap, do trigger RIG-I associated immune response[35].

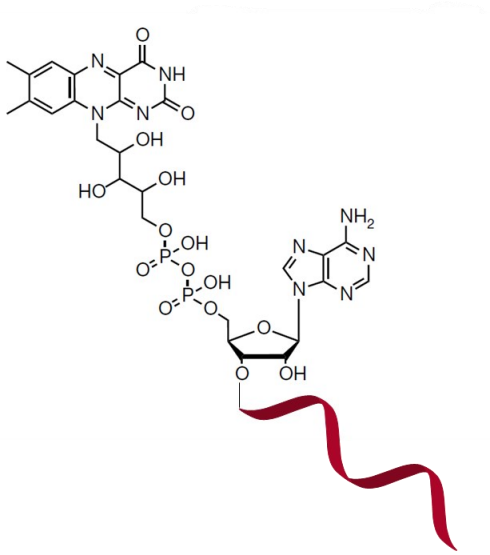


Figure 3.: FAD-RNA cap

Another novel type of non-canonical RNA caps was recently discovered in our laboratory[1]. Those caps consist of molecules of dinucleoside polyphosphates on 5′ end of RNA, therefore they are called **dinucleoside polyphosphate RNA caps.** As they are main subject of this thesis, the next chapters will be entirely devoted to their introduction. First, basic knowledge about free dinucleoside polyphosphates is crucial for understanding dinucleoside polyphosphate RNA caps, therefore they will be introduced in chapter 2. of introduction. In chapter 3. of introduction, I will finally focus on the characterization of dinucleoside polyphosphate RNA caps.

2. Dinucleoside polyphosphates

Dinucleoside polyphosphate molecules were discovered more than 50 years ago[13]. They consist of two nucleosides connected by several phosphates in 5′ to 5′ orientation (Figure 4.). The nucleoside base could be adenosine, guanosine, cytidine, or uridine, with purine bases being more common. The number of phosphates could vary from 2 to 6. The most common and widely studied dinucleoside polyphosphate is diadenosine tetraphosphate (Ap_4A), composed of two adenosines connected by four phosphates. In spite of over 50 years of research, physiological role of dinucleoside polyphosphates remains largely unresolved. It is known that they function as alarmones during stress response in bacteria[36,37]. In eukaryotes, they are linked to DNA damage response and Ap_4A acts as a second messenger during the activation of mast cells in mammals. Most recent studies have shown that they act as non-canonical RNA caps in both bacteria and eukaryotes[1,38].

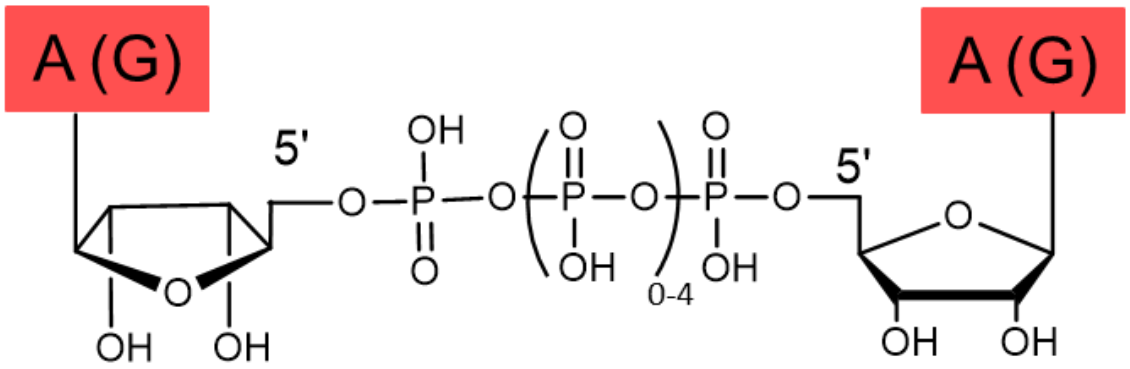


Figure 4.: Chemical structure of dinucleoside polyphosphates.

There are several dinucleoside polyphosphates discovered in living cells. First discovered dinucleoside polyphosphate was diadenosine tetraphosphate (Ap_4A), which was discovered by Zamecnik *et al.*[13] by analysis of products from reaction of lysyl-tRNA synthetases. As already mentioned, Ap_4A is the most widely studied dinucleoside polyphosphate[39]. Other dinucleoside polyphosphates discovered in living cells are Ap_3A, Ap_4G, Ap_3G or Gp_4G[40]. Their function is probably to some extent the same to Ap_4A, for example Ap_3A, Ap_4G and Ap_3G could act as alarmones in *E. coli* as well[41]. On the other hand, some studies on eukaryotes have shown that Ap_4A and Ap_3A could have an antagonistic function[42]. Gp_4G is a special case among dinucleoside polyphosphates, it has been discovered only in embryos of *Artemia* shrimps, where it acts as an energy and purine storage[43,44].

In both bacterial and eukaryotic cells, dinucleoside polyphosphates are synthesized by action of tRNA synthetases[36]. These enzymes are known for their important role in translation - the synthesis of dinucleoside polyphosphates is their moonlighting activity. There are around 20 tRNA synthetases in every living organism, each of these enzymes have different rate of Np_nN synthesis[45]. The enzyme most proficient in synthesizing Np_nN is the lysyl-tRNA synthetase and is likely to be responsible for synthesis of majority of Ap_4A inside a cell[46]. Ubiquitin and ubiquitin-like activating enzymes, DNA ligase III, T4 RNA ligase or several other enzymes have been shown to possess the ability to synthetize Ap_4A as well[47-50]. The last chapter of the introduction (chapter 5.) will be dedicated to detailed description of dinucleoside polyphosphates synthesis by tRNA synthetases.

2.1. Enzymatic degradation of dinucleoside polyphosphates

The enzymes responsible for degradation of dinucleoside polyphosphates can be classified into two groups, specific and non-specific hydrolases[51]. Specific hydrolases belong to several different families, with the most important being the NudiX hydrolase family. Other hydrolases are from histidine triad (HIT) family, DXO family and from ApaH-like hydrolase family. These enzymes can be further divided into two groups, based on their mode of action: symmetric hydrolysis, where Ap_4A is degraded to two ADP molecules, or asymmetric hydrolysis, where Ap_4A is degraded to ATP and AMP.

NudiX enzymes are responsible for hydrolysis of organic polyphosphates like in NAD, FAD, capped RNA or Np_nNs[51]. Their diverse function in cells include the metabolism of molecules with phosphate chain, cleaning the cell out of toxic nucleotides (like 8-OH-GTP) or they are involved in RNA decapping processes. For example, NUDT20, better known as DCP2, is responsible for cleavage of canonical mRNA cap in humans, or in bacteria, NudC is cleaving the non-canonical NAD-RNA cap[23]. There are many NudiX enzymes encoded in genomes of “model” organisms (13 NudiX enzymes in *E. coli*, 24 in humans and 28 in *Arabidopsis thaliana*), which marks their importance for the cellular physiology[52].

ApaH hydrolase is the most important enzyme for Np_nN degradation in *E. coli*. It is responsible for controlling their concentration levels and its mutations lead to a huge increase in the intracellular concentration of Np_nNs[53]. ApaH also has the ability to cleave Np_nN-RNA caps[1]. HIT proteins are eukaryotic enzymes also connected to Np_nN

metabolism[54] and DXO is enzyme responsible for NAD-RNA or Np_nN-RNA cap degradation in eukaryotes[38,55]. Non-specific enzymes degrading Np_nN, like Snake Venom Phosphodiesterase, have been widely used in assay for measuring Np_nN concentration[56].

2.2. Dinucleoside polyphosphates in bacteria

In bacteria, dinucleoside polyphosphates (but mostly Ap_4A), act as alamones regulating stress response[41]. Basal concentration of Ap_4A in *E. coli* cell is around 2 μM[40]. During stress conditions, concentration increases 50 times to 100 μM[40]. Stress conditions, that increase Ap_4A concentration include heat shock, oxidative stress (from various compound like H_2O_2, quinones, cadmium ions or from depletion of glutathione[57]) or even exposure to antibiotics (kanamycin)[58,59]. Moreover, deletion of *ApaH* hydrolase or *LysU* overexpression leads to an increased concentration of Ap_4A[58].

Most of these observations were already made around year 1980, but the molecular target of Ap_4A has not been identified for a very long time,. There was a long-lasting discussion if those molecules are “friends” or “foes”[60]. In other words, if they are responsible for regulating stress response, or they are just toxic metabolites created during stress conditions. Some studies showed that Ap_4A is not needed for activation of stress response elements and does not affect cell growth and survival under stress conditions[61]. However, recent studies showed that Ap_4A might really act as an alarmone[37].

In a study from Giammarinaro *et al.*[37], authors have identified, that Ap_4A binds to *Bacillus subtilis* enzyme Inosine-5′-monophosphate dehydrogenase (IMPDH) and regulates its activity. IMPDH is composed of a catalytic domain and an ATP binding CBS domain. The enzyme can be found in two oligomeric states, either the more active tetramer, or the less active octamer. Ap_4A can bind CBS domain of two tetramers and create a less active octamer. Lower activity of IMPDH then leads to higher activity of adenosine synthesis pathway at the expense of lower activity of guanosine synthesis pathway. Lower concentration of GTP leads to lower proteosynthesis by downregulating expression of rRNA. In contrast, ATP is crucial for function of protein chaperones, and increased ATP concentration can enhance their function. This should improve proteostasis during stress conditions and cell survival. Authors of this study

claim that in *B. subtilis*, IMPDH regulation should be primarily linked to Ap_4A and not to ATP. It should be noted that previous studies on *E. coli* also found interactions of Ap_4A to IMPDH but claim that Ap_4A binding is not specific and that the regulation of IMPDH is linked to ATP[61]. Another argument for dinucleoside polyphosphates being "friendly" alarmones is the discovery of dinucleoside polyphosphate RNA caps, that could explain their physiological function.

2.3. Dinucleoside polyphosphates in eukaryotes

Stress conditions in eukaryotes cause elevated Ap_4A concentration similarly to bacteria[62]. Basal concentration of Ap_4A in eukaryotes is lower than in bacteria, around 0.6 μM[40]. Stress conditions, like oxidative stress or heat shock, elevate Ap_4A concentration about 2 times[40]. The rise of Ap_4A concentration is also connected to DNA damage. In an article by Marriot *et al.*[48], authors have shown that Ap_4A acts as an inhibitor of DNA replication initiation during stress conditions.

Ap_4A acts as a second messenger during activation of mast cells[63]. Mast cell activation is a process, when inactivated mast cell binds IgE antibody and as a result releases histamine and other immunologically active compounds. IgE binds to FcεRI receptors, and this activates mitogen-activated protein (MAP) kinase pathway. MAP kinase Erk1 then phosphorylates lysyl-tRNA synthetase on Ser207. This leads to changes in

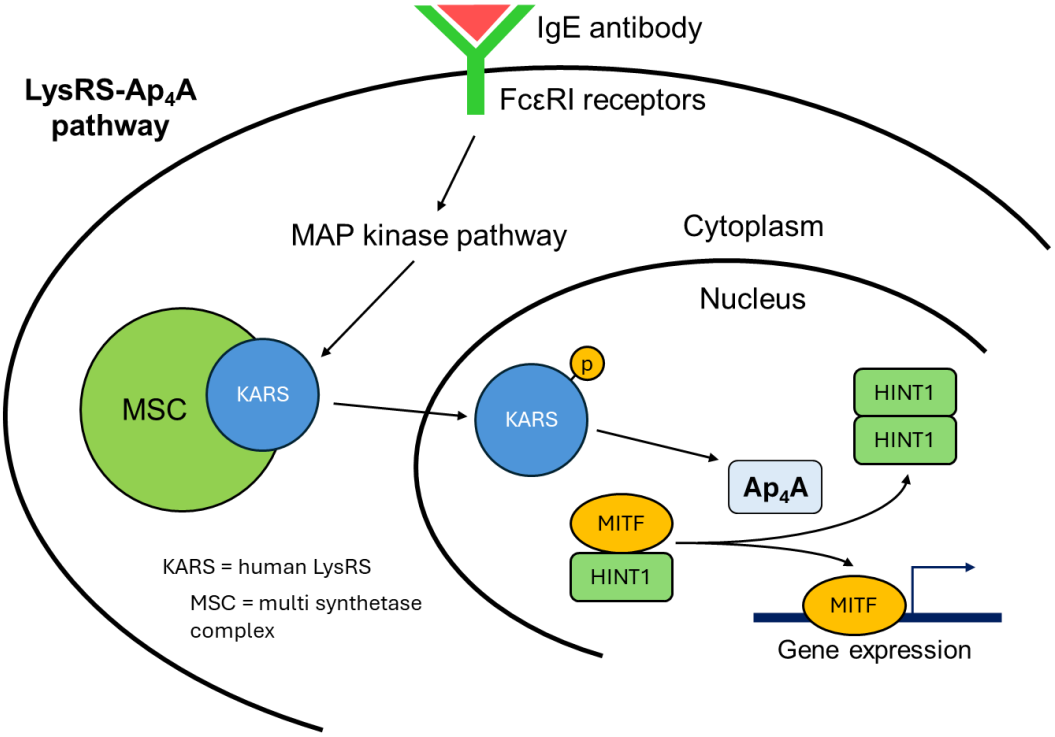


Figure 5.: LysRS-Ap_4A pathway during the activation of mast cells.

structure of LysRS and dissociation from multi synthetase complex (MSC)[64]. Phosphorylated LysRS then translocate to nucleus, where it starts producing Ap_4A. In the nucleus, there is a complex of inactive Microphthalmia-associated transcription factor (MITF) bound to HINT1 protein. When Ap_4A concentration in the nucleus rises, Ap_4A starts binding HINT1, therefore HINT1 starts to dimerize leading to release and activation of MITF transcription factor. MITF regulates genes involved in proliferation, cell migration or histamine and cytokine release (Figure 5.).

3. Dinucleoside polyphosphate RNA caps

Dinucleoside polyphosphate RNA caps were discovered in bacteria by Cahová group and Belasco group simultaneously in 2019[1,2]. This type of non-canonical RNA caps consists of molecule of dinucleoside polyphosphate on 5′ end of RNA. They were first observed by mass spectrometry-based analysis. RNA from *E. coli* was extracted, washed out of all non-covalently linked molecules, and then digested by nuclease P1. Nucleotides and dinucleosides polyphosphates were separated by liquid chromatography and analysed by mass spectrometry. This led to the discovery of several different caps (unmethylated and methylated) on 5′ end of small RNAs.: Ap_3A, Ap_3G, Ap_5A, m^6Ap_3A, mAp_4G, m^7Gp_4Gm, mAp_5A, mAp_5G, $2mAp_5G$ and $2mGp_3G$. Quantification of novel RNA caps showed that they are more abundant during stationary phase than in exponential phase (same as free dinucleoside polyphosphates in stationary phase). Stationary phase represents a stress condition for bacteria. Interestingly, abundance of cap methylation increased during stationary phase as well. This suggest that methylation could be used as protection against RNA degradation. This hypothesis is also supported by ability of enzyme RppH to cleave non-methylated caps, but it is unable to cleave caps with methyl groups.

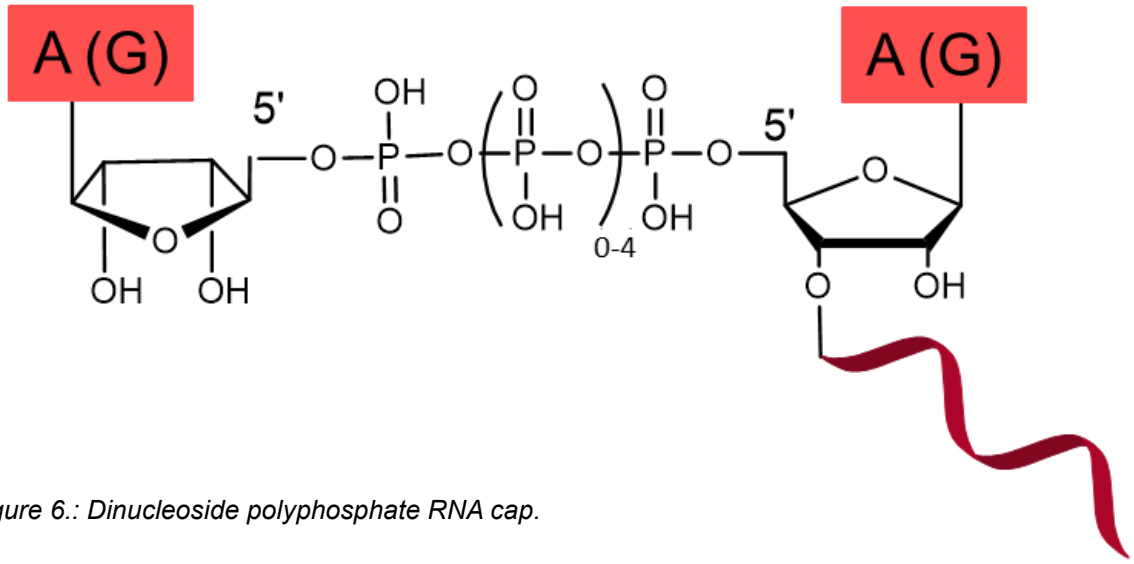


Figure 6.: Dinucleoside polyphosphate RNA cap.

The approach used by the Belasco group differed from that used in our group[2]. First of all, they isolated RNA from cadmium stressed *E. coli* cells. Several selected transcripts (not known, why these particular transcripts were selected) were then cleaved by deoxyribozyme to create uniform 3′ ends. Afterwards, RNA was analysed by boronate acrylamide gel electrophoresis, which was developed for analysis of NAD capped

RNAs[65]. Subsequent analysis on thin layer chromatography (TLC) and comparison with artificially prepared Np_nN capped RNA showed, that new type of capped RNA seen on boronate gel is a dinucleoside polyphosphate cap. Their approach enabled to quantify percentage of capped RNA among selected transcripts (mRNAs). For example, 74% of *yeiP* transcripts is capped during cadmium stress, but in case of *ApaH* deletion (ApaH is besides RppH one of two most important enzymes in degradation of Np_nN caps) only 36% of transcripts are capped. Among selected transcripts, percentages of capped transcripts varied greatly, some transcripts (*yeiP* or *efp*) are capped in more than 70% during cadmium stress, but other are almost not capped at all (*lpp* or *adk*).

Recently, our group discovered Ap_4A-RNA caps also in mammalian cells[38]. Experimental approach was similar to that used in discovery of dinucleoside polyphosphate caps in bacteria. From HEK293T cells, long fraction of RNA was isolated, washed from all non-covalently bound molecules and then analysed on liquid chromatography coupled with mass spectrometry (LC-MS). Subsequent experiments showed that Ap_4A RNA was degraded by NUDT2 and DXO in mammals, Ap_4A-capped mRNA was not translated and did not initiate immune response, which suggests it is a natural component of the cell.

3.1. Biodegradation of dinucleoside polyphosphate RNA caps

To elucidate the role of dinucleoside polyphosphate caps in cells, we need to acquire more details about their metabolism. There are several enzymes that were shown to be able to cleave Np_nN-RNA caps in bacteria. RppH is enzyme that cleaves phosphates from 5′ end of RNAs in bacteria, as already mentioned earlier. RppH can cleave Np_nN-RNA caps leaving monophosphate on 5′ end as well. Mechanistic studies have revealed that this process is more complicated, because RppH first cleaves the Np_nN cap after α′-phosphate (first phosphate from side of cap forming adenosine – in Figure 6., it is the phosphate furthest to the left) leaving triphosphate, which is then cleaved to monophosphate[66]. As already mentioned, RppH is not able to cleave methylated caps.

The second important enzyme in bacterial Np_nNs metabolism is ApaH. This enzyme is known for hydrolysis free dinucleoside polyphosphates, but it can cleave Np_nN-RNA caps as well, leaving diphosphate on 5′ end of RNA. ApaH can degrade even

methylated Np_nN-RNA caps. NudC, an enzyme responsible for degradation of NAD-RNA caps, was tested for Np_nN-RNA caps cleavage, but it has been shown that it is not able cleave them[1].

Based on known data about abundance of Np_nN-RNA caps, their degradation and expression profiles of enzymes responsible for their degradation, a model of Np_nN-RNA caps function in bacteria was proposed (Figure 7.)[1]. In the exponential phase, Np_nN caps are cleaved by RppH and then RNA is degraded by nucleases. During stress conditions like stationary phase, Np_nN caps are methylated by an unknown methyltransferase and cannot be cleaved by RppH anymore. This stabilizes RNA during stress conditions and cell do not waste energy on creating new RNAs. When the stress condition is over, ApaH cleaves Np_nN caps from RNA and they are then degraded by nucleases.

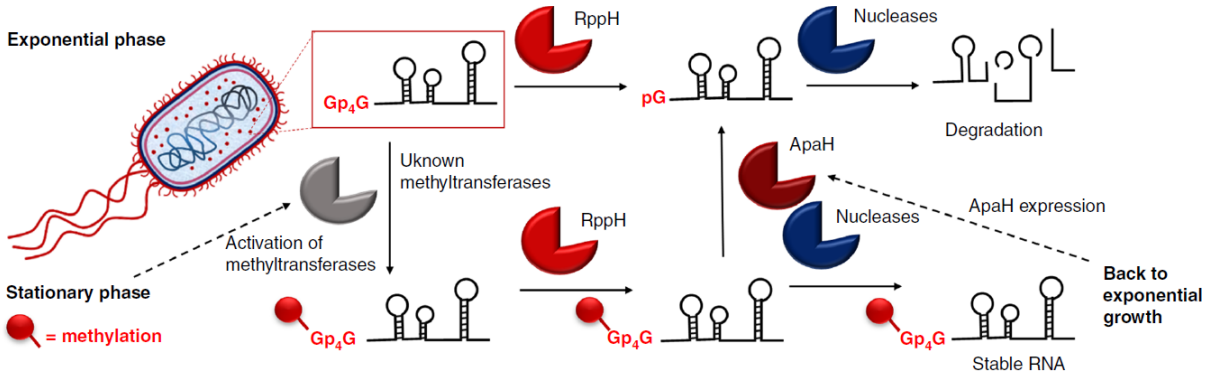


Figure 7.: Model of Np_nN-RNA cap function in E. coli based on Hudeček et al.[1]

In mammals, NUDT2 enzyme, which is responsible for degradation of free Ap_4A in a cell, was shown to be able to cleave Ap_4A-RNA, leaving monophosphate on 5′ end. DXO, which is known to cleave NAD-RNA cap, but not free dinucleoside polyphosphate, is also able to cleave Ap_4A-RNA in mammals. Interestingly, it cleaves within the polyphosphate chain of the cap, leaving monophosphate on 5′ end of RNA, which contrasts with cleavage of NAD-RNA cap, where whole NAD molecule is cleaved out of RNA[38].

3.2. Synthesis of dinucleoside polyphosphate RNA caps

Synthesis of Np_nN-RNA caps is another important aspect of their metabolism. There are two options how Np_nN caps could be synthetized in both bacteria and eukaryotes. First one is the incorporation of free dinucleoside polyphosphate by RNA polymerase as a non-canonical initiating nucleotide (NCIN) during transcription, similar mechanism as NAD-RNA cap is synthetized. This mechanism has already been proven functional

both *in vitro* and *in vivo* (in *E. coli*)[5,6]. Regulation of this mechanism could be achieved by two proposed ways: change of concentration of free dinucleoside polyphosphates and by the sequence of the promoter. It has been shown, that manipulating promotor sequence have effect on efficiency of incorporation of free Np_nN to RNA[6]. Greatest changes are achieved when position -1 on promoter is mutated. Position -2 or -3 also have effect on incorporation efficiency, but relatively minor in comparison to effect of -1 position. It was proposed, that dinucleoside polyphosphate can pair with base in position -1 on promoter. When A is present on -1 position (T on the complementary strand), highest capping potential with Ap_4A is achieved. Similarly, when G is on -1 position, highest capping potential is observed with Gp_4A. Base pairing of dinucleoside polyphosphate with base on -1 position of promoter could be canonical Watson-Crick base pairing[6], or non-canonical pairing of adenine on Ap_4A and thymine on complementary strand of promoter (adenine pairs with O2 and N3 on thymine)[5].

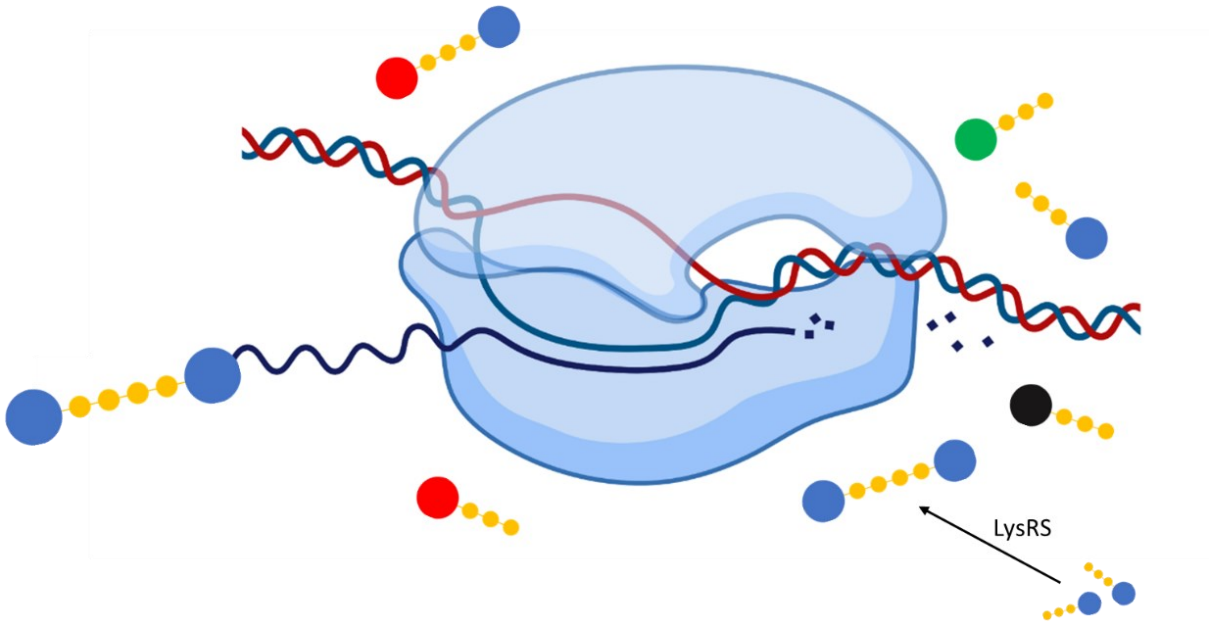


Figure 8.: Illustration of incorporation of free dinucleoside polyphosphates by bacterial RNA polymerase to RNA. Created using BioRender.

Abundance of free dinucleoside polyphosphates is another factor that influence Np_nN incorporation by RNA polymerase. Concentration of Ap_4A or other dinucleoside polyphosphates increases during stress conditions, which corelates to increased abundance on Np_nN-RNA caps in same stress conditions. However, one problem can be seen, when we compare types of discovered Np_nN caps and observed types of free dinucleoside polyphosphates in bacteria. Among discovered Np_nN caps, there is Gp_4G and Gp_3G cap, but Gp_4G of Gp_3G have not been observed in free form in *E. coli*.

This leads us to the second possible mechanism of cap acquisition, and that is post-transcriptional synthesis of dinucleoside polyphosphate cap. This mechanism would require, that some unknown enzyme would interact with triphosphorylated RNA and directly synthetize Np_nN-RNA cap. Ideal candidate enzymes are tRNA synthetases. They are one of the most important enzymes connected to protein synthesis, but they are also responsible for synthesis of dinucleoside polyphosphates[13]. General information about tRNA synthetases will be presented in next chapter (chapter 4.) and detailed description of synthesis of Ap_4A by tRNA synthetases with possibility to create Np_nN caps in following parts of this thesis (chapter 5.).

4. Aminoacyl-tRNA synthetases

Aminoacyl-tRNA synthetases are enzymes essential for every cellular organism, as they are responsible for attaching cognate amino acid to tRNA. This is crucial step in decoding information from mRNA into functional protein. In fact, they are the only enzymes capable of implementing genetic code[67]. There are approximately 20 different tRNA synthetases in each cellular organism, with one dedicated to each amino acid. tRNA synthetases have other functions besides the aminoacylation reaction. They can edit misacylated tRNAs or perform various functions unrelated to the tRNA aminoacylation. For instance, they play a role in transcription regulation in bacteria or immune response in mammals[68,69]. One of their moonlighting activities is the synthesis of dinucleoside polyphosphates[13]. This non-canonical reaction is in the focus of this thesis and next chapter (chapter 5.) will be dedicated to the synthesis of dinucleoside polyphosphates by tRNA synthetases.

For each proteogenic amino acid, there is generally different tRNA synthetase. All tRNA synthetase genes exhibit low sequence similarity, and for long time, it seemed that there is no relation between individual tRNA synthetases. In 1990, Eriani *et al.*[70] discovered that tRNA synthetases can be separated to two classes based on their structure (Figure 9.). Class I tRNA synthetases have a catalytic domain composed of Rossman fold (five parallel β-strands connected by α-helices) and conserved sequence motifs HIGH and KMSKS. Class II tRNA synthetases have a unique fold of catalytic domain composed of seven antiparallel β-strands and three conserved sequence motifs[71,72].

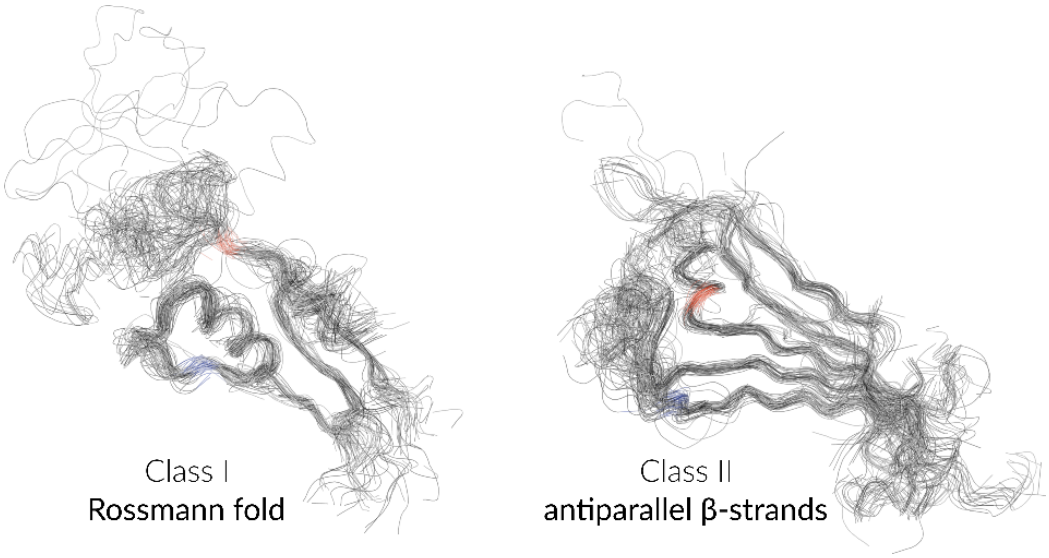


Figure 9.: Basic structural characteristics of two different classes of tRNA synthetases[3]

In addition to these basic structural characteristics, each class of tRNA synthetases have distinct biochemical and structural attributes[71]. Class I tRNA synthetases bind the tRNA-accepting stem from the minor groove side (except for TrpRS and TyrRS, which bind the tRNA-accepting stem from the major groove side) and subsequently aminoacylate 2′-OH group of 3′ terminal adenosine of tRNA. In contrast, class II enzymes bind the tRNA-accepting stem from the major groove and aminoacylates 3′-OH group of terminal adenosine (with an exception of PheRS, which aminoacylates 2′-OH group). The binding of an ATP molecule differs between these two classes. Class I enzymes bind ATP in an extended configuration, while class II enzymes bind ATP in an unusual bend configuration, where γ-phosphates bend back towards adenosine ring. This distinction leads to different reaction kinetics, where class I enzymes have aminoacyl-tRNA (aa-tRNA) release as the rate limiting step, whereas for class II, amino acid activation is the rate limiting step of the reaction. Another common characteristic is that class I tRNA synthetases are typically monomers, but class II enzymes form dimers (conserved motif 1 is responsible for dimerization).

Within both classes, tRNA synthetases can be further divided into subgroups (a, b and c in both classes). Enzymes in a subgroup sometimes uses chemically similar amino acids as a substrate. For example, class Ia enzymes catalyse the aminoacylation of long hydrophobic or thiolated amino acids (Leu, Ile, Val, Met), class Ic catalyse the reaction of only aromatic Tyr and Trp or class IIb catalyse the reaction of charged amino

Table 1.: Table of all tRNA synthetases from E. coli divided to classes (archaeal LysRS-I, which is not present in E. coli, is also included in the table. In E. coli, only LysRS-II is present)

Class I tRNA synthetases				**Class II tRNA synthetases**			
Subgroup	tRNA synthetase	Oligomeric structure	Number of aa in *E. coli*	Subgroup	tRNA synthetase	Oligomeric structure	Number of aa in *E. coli*
Ia				**IIa**			
	LeuRS	α	860		**SerRS**	α_2	430
	IleRS	α	937		**ThrRS**	α_2	642
	ValRS	α	951		**ProRS**	α_2	572
	MetRS	α or α_2	642		**HisRS**	α_2	424
	CysRS	α	461		**GlyRS**	$\alpha_2\beta_2$	303 and 689
	ArgRS	α	577				
Ib				**IIb**			
	GluRS	α	471		**LysRS-II**	α_2	505
	GlnRS	α	551		**AspRS**	α_2	590
	LysRS-I	α	N/A		**AsnRS**	α_2	467
Ic				**IIc**			
	TyrRS	α_2	424		**AlaRS**	α_4	875
	TrpRS	α_2	334		**PheRS**	$\alpha_2\beta_2$	327 and 795

acid such as Lys, Asp or Asn (however, not all subgroups use chemically similar amino acids; for example, class IIc catalyse aminoacylation of distinct amino acid Ala and Phe). The classification of all tRNA synthetases into classes and subgroups can be seen in the Table 1.1. (data from Giege *et* Springer[72]). The archaeal LysRS-I, which is not present in *E. coli*, is also included in the table. This is for illustration, that two different lysyl-tRNA synthetases for lysine exist, which will be discussed in chapter about evolution of tRNA synthetases (chapter 4.2). Whenever lysyl-tRNA synthetase is mentioned in the text, it means the class II bacterial enzyme (LysRS-II is also present in eukaryotes, which will be discussed later as well). When archaeal class I enzyme is mentioned in the text, it is explicitly specified.

4.1. Structure of tRNA synthetases and aminoacylation reaction

All tRNA synthetases are composed of at least two different domains, the catalytic domain, and the anticodon binding domain (ACB domain) (Figure 10.). Many tRNA synthetases have additional domains, with the most common being an editing domain inserted into the catalytic domain, or N or C terminal extensions, which are common in eukaryotic tRNA synthetases and play roles in their non-canonical functions[69]. Catalytic domains have two different folds, that separate tRNA synthetases to two different classes. As their name suggest, they are responsible for amino acylation, but they take part in tRNA recognition as well. ACB domains are responsible for binding an anticodon of tRNA. ACB domain folds are variable among different tRNA synthetases and could

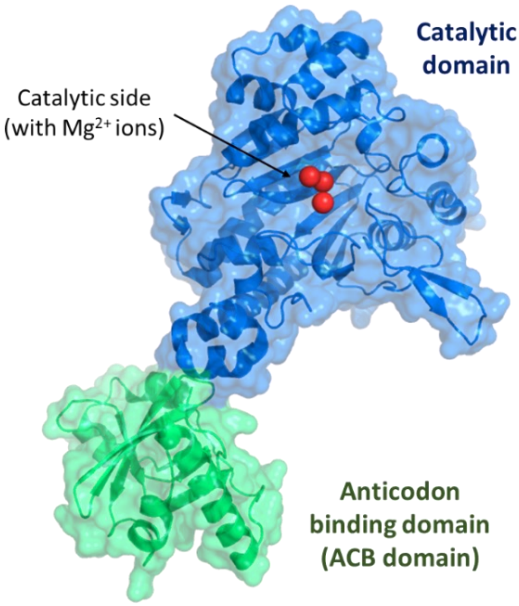


Figure 10.: Domains of E. coli lysyl-tRNA synthetase (PDB code: 1E22[8], edited using PyMOL)

be composed of only α-helices, only β-sheets or α/β folds. They are usually conserved through the evolution[7].

tRNA synthetases are well adapted for binding cognate tRNAs. There are several key residues on both tRNA and tRNA synthetase, that ensure right recognition[73]. On the tRNA synthetase, they are located in the ACB domain, near the active site, and even in the body of the tRNA synthetase, where they recognize nucleotides or phosphates of tRNA D-loop. This recognition is crucial for right positioning and discrimination among different tRNAs.

In eukaryotes, tRNA synthetases form large cytoplasmatic complex called multi synthetase complex (MSC)[74]. In *Saccharomyces cerevisiae*, the MSC is composed of only two tRNA synthetases (MetRS and GluRS) and one auxiliary protein, that connects them. In mammals, the MSC is composed from nine tRNA synthetases (LysRS, MetRS, IleRS, ArgRS, GlnRS, GluRS, ProRS, AspRS, LeuRS) and three accessory interacting multifunctional proteins (AIMPs 1-3). The exact role of the MSC is not fully understood, but it has been proposed that it interacts with several signalling pathways. For example, LeuRS and IleRS from MSC interacts with mTOR (LeuRS acts as an intracellular leucine sensor and activator of the mTORC1 complex)[71]. Also, in the previously mentioned LysRS-Ap_4A pathway, the MSC plays important role. In the inactivated mast cell, LysRS is part of the MSC, which is mediated by interaction with AIMP2. Upon mast cell activation, LysRS is phosphorylated. This leads to change in structure of LysRS, which prevents binding to AIMP2. LysRS is released from the MSC and translocated to the nucleus[75].

Aminoacylation reaction is the most important function of tRNA synthetases, therefore a more detailed description is necessary (Figure 11.). Within the catalytic domain, there are two pockets. One of them recognizes cognate amino acids and the other recognizes molecule of ATP. In the first step of the reaction, oxygen from carboxyl group of the amino acid attacks the α-phosphate of ATP forming an aminoacid-AMP intermediate. Pyrophosphate is released from the first step of the reaction. In the second step, the tRNA binds to the enzyme and is positioned in a way, that 3′ end is in the active site. Hydroxyl group from 2′ or 3′ carbon on the ribose (depends on tRNA synthetase class) attacks the carboxyl group of the amino acid and amino acid is transferred to 3′ end of the tRNA. AMP is a leaving group of the second step of aminoacylation reaction. Magnesium ions are required for this reaction[71].

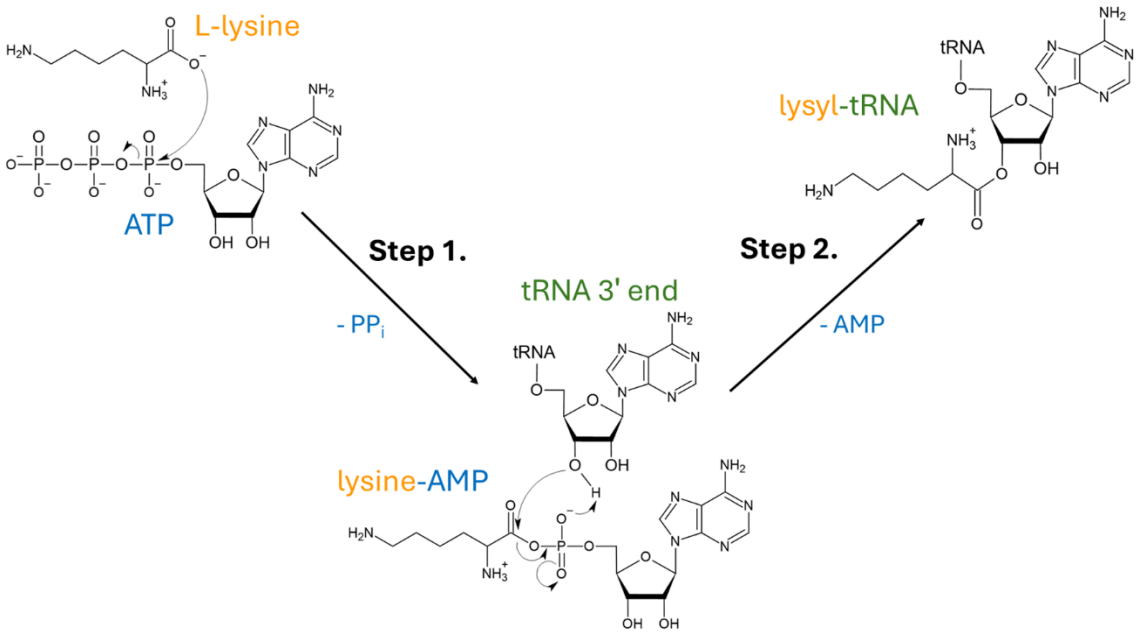


Figure 11.: Reaction of aminoacyl-tRNA synthetases on the example of lysyl-tRNA synthetase. In the first step, aa-AMP intermediate is created (this step is also called amino acid activation), and in the second step, amino acid is transfer to 3' end of tRNA.

4.2. Evolution of tRNA synthetases

The evolution of tRNA synthetases presents a challenging yet crucial field of study, essential for understanding the molecular basics of life and the origin of life on Earth. An almost complete set of tRNA synthetases was already present in earliest organisms[67,76], which gives this field a potential for studying earliest known organisms. In this section, some of the fundamental concepts underlying evolution of tRNA synthetases will be discussed.

The reconstruction of phylogenetic trees by Woese *et al*.[67] showed that horizonal gene transfer played an important role in the evolution of tRNA synthetases. Their analysis revealed that for some tRNA synthetases, phylogenetic tree follows the pattern of phylogenetic tree of rRNA. However, there are certain tRNA synthetases that do not follow the classical phylogenetic tree of rRNA. This could be illustrated on the lysyl-tRNA synthetase (Figure 12.). There are two different types of lysyl-tRNA synthetase, one belongs to class I and it is present in most of archaea (archaeal type). Second belongs to class II and it is present in both eukaryotes and bacteria (bacterial type). It has been proposed that through horizontal gene transfer, bacterial LysRS-II was

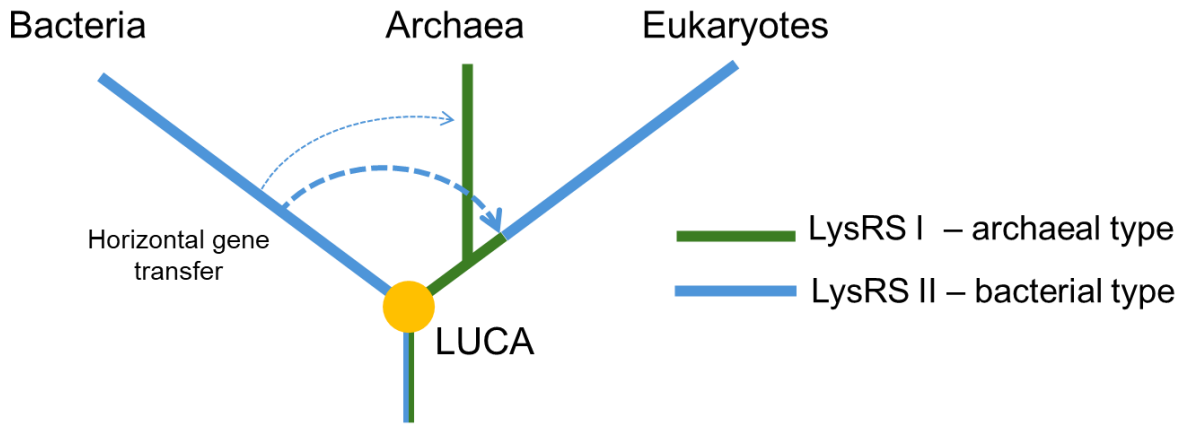


Figure 12.: Evolution of lysyl-tRNA synthetases. Bacterial type LysRS-II replaced the archaeal type LysRS-I in the ancestors of eukaryotes and some other archaeal species by means of horizontal gene transfer. Inspired by O'Donoghue et Luthey-Schulten[7].

transferred to archaeal ancestors of eukaryotes and replaced the archaeal type LysRS-I. This independently occurred multiple times in different archaeal groups[76,77].

The search for the origin of tRNA synthetases has led researchers to focus on their simplified versions. These enzymes are called urzymes ("ur" meaning primitive, original or earliest) and they are composed only of catalytic cores of tRNA synthetases, but still capable of aminoacylating tRNAs. The first urzyme was prepared from TrpRS[78]. This enzyme containing only catalytic domain was able to aminoacylate tRNA at almost the same speed as normal TrpRS, although the specificity of the reaction was significantly lower[79]. The search for urzymes is connected to the Rodin-Ohno hypothesis about the origin of tRNA synthetases.

The Rodin-Ohno hypothesis proposes that both classes of tRNA synthetases originated from a single ancestral gene, which was read bidirectionally[80]. Sequence alignments of HIGH and KMSKS motives with antiparallel motif 2 and motif 1 showed hight sequence similarity. This theory is further supported by minor or major groove tRNA-acceptor stem binding and by proposed similar age and similar codon usage in genes of both classes. This theory was later supported by large bioinformatic analyses and is now widely accepted as the best model for the origin of both classes of tRNA synthetases[81].

Some bacteria and archaea lack a complete set of 20 tRNA synthetases. In most cases, GlnRS and AsnRS are absent, and they are replaced by the Non-Discriminating (ND) GluRS and AspRS. These enzymes can bind both Glu/Asp and Gln/Asn tRNA. However, when they bind Gln/Asn tRNA, they aminoacylate the tRNA with Glu/Asp and

then form a complex with amidotransferase (AdT) enzyme[82]. This enzyme catalyses transformation of Glu/Asp to Gln/Asn directly on tRNA. A similar scenario occurs with cysteine in some archaea, where it is synthesized from O-phosphoserine tRNA directly on tRNA[83]. It has been proposed that GlnRS evolved later in evolution, probably in eukaryotes, and was acquired by some bacterial species through horizontal gene transfer[84].

4.3. tRNA processing

tRNAs are small RNAs essential for translation of mRNA into proteins. They contain an anticodon sequence corresponding to a specific amino acid codon. Their structure is unique among RNAs. It can be visualised as a cloverleaf or as a T structure in 3D space (Figure 13.). Besides the anticodon stem, they contain a D loop (typical for presence of dihydrouridines), a T loop (typically contains thymine od pseudothymine) and an acceptor stem. The acceptor stem terminates with a CCA sequence on 3′ end. The terminal adenosine of CCA sequence is the residue that is aminoacylated by tRNA synthetases. Aminoacylated tRNA is transferred to the ribosome, where it binds complementary codon on mRNA. Amino acid is then transferred from tRNA to the growing peptide[85].

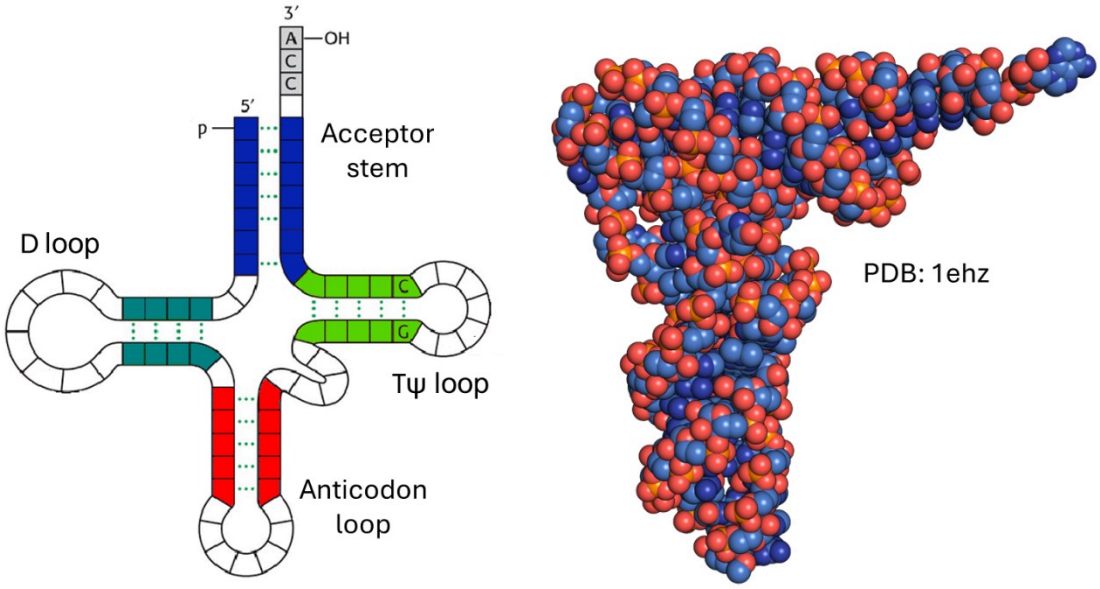


Figure 13.: Cloverleaf representation and 3D structure of tRNA. Source of tRNA cloverleaf representation is Wikimedia commons[45] and 3D structure was modelled in PyMOL, based on PDB structure 1EHZ[11].

In bacteria, tRNAs are usually transcribed in long polycistronic transcripts containing multiple tRNAs. As a typical bacterial RNA polymerase transcript, they have

triphosphate on 5′ end. Individual tRNAs are then cleaved from pre-tRNA long transcript by RNase P[86]. RNase P is a ribozyme with a protein component called C5[87]. RNase P together with C5 recognize a sequence of five nucleotide that ends with uridine before the 5′ end of tRNA, after which the cleavage takes place (for higher reaction specificity, RNase P also recognizes structures on 3′ end and the acceptor stem). This process leaves a monophosphate on the 5′ end of the tRNA. At the 3′ end, RNase E or RNase III cleaves transcripts between individual tRNAs and then RNase D degrades the transcript until it reaches CCA[88]. There it stops and the acceptor stem is formed (in eukaryotes, tRNA processing differs. tRNAs are transcribed by RNA polymerase III without the acceptor stem ending, which is then synthesized by enzyme CCA tRNA nucleotidyltransferase). This process is extremely fast, and it is hard to spot pre-tRNAs or processing intermediates in living cell.

There are also several tRNAs, which are transcribed as monocistronic transcripts. These transcripts still have 5 nucleotides upstream of mature tRNA, that are processed by RNase P and C5. Only difference is that they have a triphosphate on 5′ end, which prevents binding C5 protein[89]. Therefore, this triphosphate is first cleaved by the RppH and then tRNAs are further processed. Deletions of *RppH* lead to difficulties in processing of monocistronic tRNAs, but for polycistronic tRNAs, *RppH* deletion has no effect[89].

5. Synthesis of dinucleoside polyphosphates by tRNA synthetases

Besides the canonical role of aminoacylating tRNAs, tRNA synthetases have ability to synthetize Ap_4A and other dinucleoside polyphosphates. This reaction was discovered by Zamecnik *et al.*[13] in 1966 by analysing products of *E. coli* lysyl-tRNA synthetase reaction. This experiment was also the first observation and characterization of dinucleoside polyphosphates, therefore tRNA synthetases and dinucleoside polyphosphates are tightly connected areas of research. In their article, they have already shown basic principles of how Ap_4A is created from ATP through amino acid-AMP intermediate. More detailed information about synthesis of Ap_4A by tRNA synthetases were described in papers from 80s (mostly by researchers lead by Sylvain Blanquet[90]). They described the differences of reaction rates among different tRNA synthetases and the ability of metals to catalyse the reaction. In recent years, some of mechanistic details about the reaction were concluded based on crystal structures of tRNA synthetases. Despite extensive research on this reaction, the reason why tRNA synthetases catalyse synthesis of dinucleoside polyphosphates, in addition to their essential role in aminoacylation of tRNAs, remains a mystery.

5.1. Different capabilities of tRNA synthetases to synthesize Ap_4A

It has been demonstrated that most tRNA synthetases are capable of synthetizing dinucleoside polyphosphates[50]. However, there are several tRNA synthetases that are unable to synthetize dinucleoside polyphosphates. For example, arginyl- and glutaminyl-tRNA synthetase from all tested organisms[91] or tryphophanyl- and tyrosyl-tRNA synthetases from *Lupinus luteus,* are unable to synthesize Ap_4A[92]. These tRNA synthetases are known to create the amino acid-AMP intermediate only when the tRNA is present[72,93], therefore amino acid is immediately transferred to tRNA, leaving no space for the generation of Ap_4A. There are several tRNA synthetases that have not been tested for the production of Ap_4A, like archaeal lysyl-tRNA synthetase (LysRS-I), but their reaction mechanism is the same as in arginyl-tRNA synthetase, and therefore they are not expected to produce Ap_4A[94].

tRNA synthetases, that can produce Ap_4A, differ dramatically in the rate of its synthesis. These differences are caused by the ability of several tRNA synthetases to be stimulated by metals, mostly by zinc ions. A complete list of all tRNA synthetases that are able to produce Ap_4A or different dinucleoside polyphosphates can be found in an article from Fraga *et* Fontes[50] or in a book '*Ap_4A and other dinucleoside*

polyphosphates' from McLennan[39]. The most capable tRNA synthetases in Ap_4A production are lysyl-, alanyl- and phenylalanyl-tRNA synthetases[45]. All these tRNA synthetases are stimulated by zinc. Without the zinc stimulation, they have similar rate of reaction as other tRNA synthetases. This shows the importance of Zn^{2+} ions for this reaction.

There are four tRNA synthetases, that are relevant for this thesis (*E. coli* LysRS from genes *LysU* and *LysS* (both class II enzymes), MetRS and HisRS). The initial rates of Ap_4A production of these tRNA synthetases were found in the literature. Based on these data from literature, LysU was described to be the enzyme most efficient in Ap_4A synthesis. LysS was reported to be only two times slower than LysU (differences between both LysRS in *E. coli* will be discussed later)[95]. HisS was described to be approximately 100-fold and MetG 1000-fold slower than LysU. Data from the literature can be seen in Table 2.[36,45,95,96].

Table 2.: Comparison of relevant tRNA synthetases for this thesis, based on data found in the literature. Data represent the initial rates of Ap_4A synthesis by tRNA synthetases in s^{-1} units. Most relevant data are highlighted in yellow.

LysU	0,02	0,4	4,2	2,5	0,0235	2,030	23	10
LysS							11	
HisS	0,0106		0,0092					4,1
MetG	0,0021		0,0021	0,0021	0,0021	0,0021		
Conditions	No Zn^{2+}	10 µM Zn^{2+}	150 µM Zn^{2+}	1mM Zn^{2+}	No Zn^{2+}	150 µM Zn^{2+}	150 µM Zn^{2+}	Without Zn^{2+}
Units	s^{-1}				s^{-1}		s^{-1}	$100*s^{-1}$
Source	Blanquet *et al.,* 1983				Brevet *et al.,* 1989		Brevet *et al.,* 1995	Goerlich *et al.,* 1982
Result	HisS 450x slower and MetG 2000x slower than LysRS				MetG 1000x slower than LysRS		LysS 2x slower than LysU	HisS 2,5x slower than LysRS withou Zn^{2+} stimulation

Lysyl-tRNA synthetase have been tested for the production of different dinucleoside polyphosphate. From the first reaction observations by Paul Zamecnik, it is known that LysRS is able to synthetize Ap_3A as well[13]. For the first step of reaction, ATP molecule is necessary, but the second step is less substrate selective. Therefore, when a mix of different nucleotides is present in the reaction, Ap_nN can be synthetized. Interestingly, not all Ap_nN are synthetized in equal rates[45]. Synthesis of Ap_4A is always faster than synthesis of other Ap_nN. The rate of synthesis of Ap_3A, Ap_4G, Ap_3G and Ap_4C is about the third of the rate of Ap_4A synthesis. Other Ap_nN like Ap_4U are synthetized even more slowly, the rate of reaction with UTP is about one fifth of Ap_4A synthesis reaction rate. With *E. coli* AlaRS, the reaction with GTP is about 50% slower than with ATP, the reaction with CTP is about 40% slower, but the reaction with UTP is almost undetectable[91].

5.2. Mechanism of dinucleoside polyphosphate synthesis

The canonical reaction of tRNA synthetases occurs in two steps. The first is reaction of an amino acid with ATP, forming an amino acid-AMP intermediate. The second step is transfer of the amino acid to tRNA and the release of AMP. For the synthesis of dinucleoside polyphosphate, the first step of reaction is identical. However, the second step differs, and the aa-AMP intermediate reacts with a molecule of ATP instead[13]. Ap_4A is created and the amino acid leaves the reaction. Other dinucleoside polyphosphates could be synthetized, when the aa-AMP reacts with a different nucleotide instead of ATP. A simplified scheme of reaction can be seen on Figure 14..

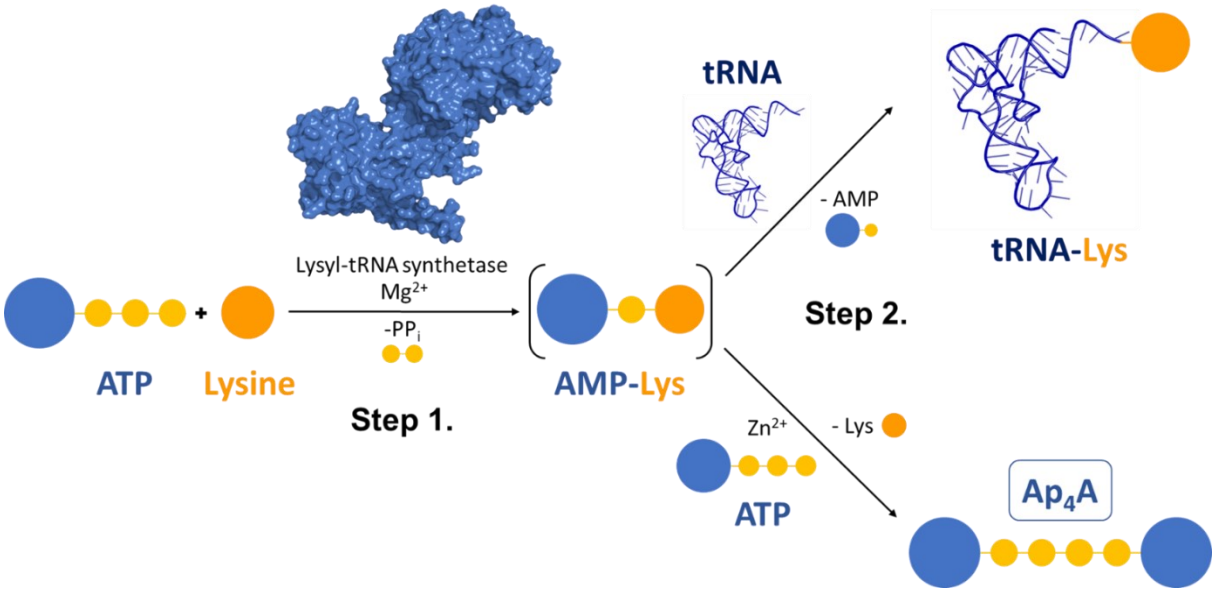


Figure 14.: A simplified scheme of Ap_4A synthesis by LysRS. In the second step, aa-AMP intermediate interacts with ATP instead of tRNA creating Ap_4A.

Metal ions play an important role in both steps of the reaction. For the first step, Mg^{2+} ions are essential. They are needed for ATP binding and catalyse intermediate formation. For class I tRNA synthetases, only one Mg^{2+} ion per active site is needed, while for class II enzymes, three Mg^{2+} ions bind to the active site. In the second step, Zn^{2+} ions and some other metal ions can catalyse the reaction. This catalysis occurs only with a few tRNA synthetases, as mentioned earlier (LysRS, AlaRS, PheRS and *E. coli* ProRS). Mn^{2+} and Cd^{2+} ions are also able to catalyse the synthesis of Ap_4A, but they are not as effective as Zn^{2+} ions[90]. However, the catalysis by metal ions could be species dependent, as research on *Myxococcus xanthus* has shown. *M. xanthus* LysRS has the highest rate of Ap_4A synthesis in the presence of Mn^{2+} ions, while Zn^{2+} ions inhibit the reaction[97].

The analysis of mutations of amino acids near the active site of lysyl-tRNA synthetase (bacterial class II enzyme) by Chen *et al*.[9] suggested the mechanism of Ap_4A synthesis. Mutations in Glu264, His270, Arg262, Glu414, Glu421 and Arg480 showed highest reduction in capabilities of Ap_4A synthesis (Figure 15.). They identified, that Arg262, Glu414, Glu421 and Arg480 are responsible for binding magnesium ions or directly binding phosphates. His270 binds both the γ-phosphate of the first ATP and makes a stacking interaction with the ring of adenosine. The most important amino acid in Ap_4A formation is Glu264. Mutation of this amino acid reduced the rate of reaction to levels without Zn^{2+} stimulation. The authors of this study claim that Glu264 is responsible for the binding of Zn^{2+}. The proposed Ap_4A reaction mechanism also includes so-called N7-N7 bridge (this is connection of both adenosine rings on Ap_4A by Zn^{2+} through binding to N7 of each adenosine[98]). The second molecule of ATP binds with γ and β phosphate to the same position as γ and β phosphate of the first ATP were binding (both of these phosphates left the active site as a pyrophosphate leaving group), but in the opposite direction (second ATP sticks out of active site). Zn^{2+} interacts with both ATP molecules via N7-N7 bridge and simultaneously with Glu264. Zn^{2+} then catalyse the attack of γ-phosphate of the second ATP to the aa-AMP intermediate. Ap_4A is formed and the amino acid leaves the reaction.

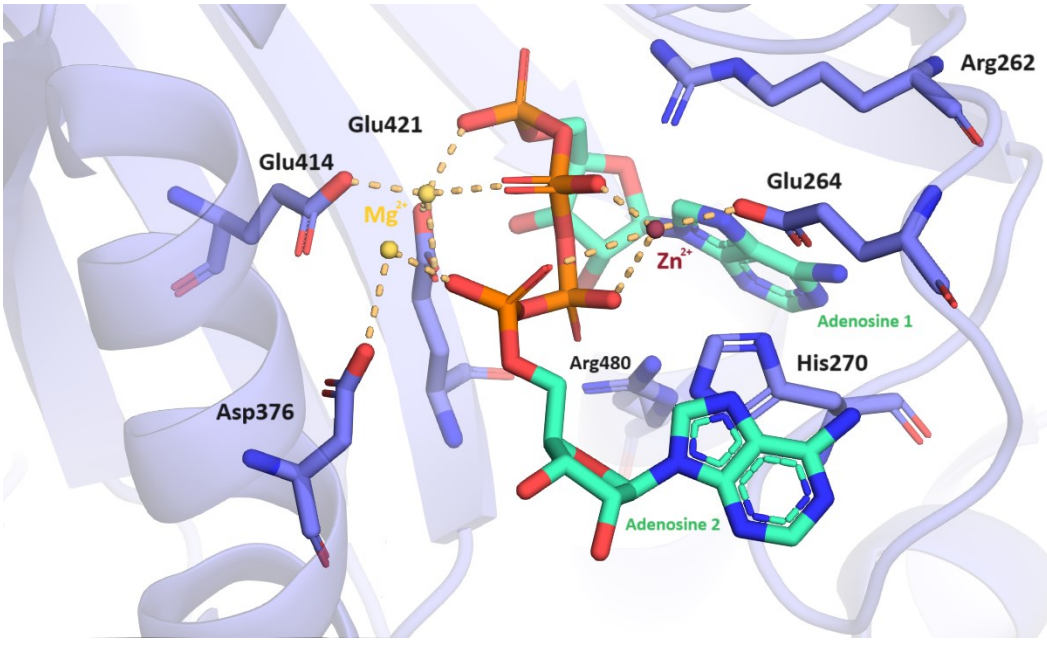


Figure 15.: Active side of LysU with important amino acid residues for Ap_4A synthesis highlighted. Model created in PyMOL based on structure 5YZX and information from Chen et al.[9]

ATP, amino acid and Mg^{2+} are essential for the first step of the reaction (an excess amount of Mg^{2+} and amino acid is used to boost Ap_4A formation)[99]. Pyrophosphate is a leaving group from the first step of reaction and because of the reaction reversibility,

higher amounts of pyrophosphate inhibit the Ap_4A production[100]. Therefore, inorganic pyrophosphatase is used to degrade the pyrophosphate and boost the reaction. Zn^{2+} ions are used to catalyse Ap_4A formation. It has been shown that only small amount of Zn^{2+} ions is needed for catalysis of Ap_4A formation (the best effect is observed in 0.1 mM concentration of $ZnCl_2$)[45]. tRNA is another factor that inhibits Ap_4A synthesis, because aminoacylation reaction seems to be preferred over Ap_4A formation[36,92]. The reaction is usually held at a slightly basic pH (around pH 8) in 37 °C (for *E. coli* or mammalian enzymes).

All tRNA synthetases create Ap_4A trough the aa-AMP intermediate. There is only one exception, and that is the human glycyl-tRNA synthetase[101]. It has been shown that this enzyme does not need an amino acid for Ap_4A synthesis, and the reaction is rather a direct condensation of two molecules of ATP. This is achieved by a small domain insertion into the active site of GlyRS, which binds the second molecule of ATP. On the other hand, the reaction speed of GlyRS is way slower than the reaction of LysRS, therefore the authors of the study proposed role of GlyRS in Ap_4A homeostasis (rather than in the response to stress conditions). The formation of Ap_4A by the bacterial GlyRS is amino acid dependent[102].

5.3. *E. coli* lysyl-tRNA synthetases LysS and LysU

In *E. coli*, there are two genes encoding lysyl-tRNA synthetase. They have 89% sequence similarity, and when they are structurally aligned, they are almost identical (RSMD = 0.85 and TM-score 0.99; structural alignment can be seen on Figure 16.).

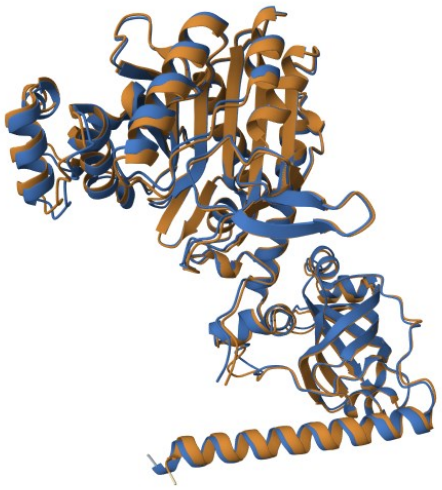


Figure 16.: Structural alignment of LysU (orange) and LysS (blue) by RSCB PDB Pairwise Structure Alignment. LysU and LysS structures were predicted by AlphaFold 2[10]

However, they both have different biochemical properties. LysS is about 8 times more active in the aminoacylation reaction, LysU is two times faster in synthetizing Ap_4A[95].

Their expression is differently regulated. *LysS* is constitutively transcribed and acts as a housekeeping gene. *LysU* is expressed only in stress conditions like heat shock, low pH, anaerobiosis or nutrient shortage[72]. The regulation of *LysU* includes factors such as *rpoH* (sigma factor σ32 responsible for regulation of heat shock genes), *lrp* (protein involved in Leu-Lrp operon, that mediates response to leucin), and the *rlu* element (*regulation of LysU*), which function is not yet known. Based on these observations, it was proposed that *LysS* act as a housekeeping gene for aminoacylation of tRNA, while *LysU* is gene mostly responsible for synthesis of Ap_4A during stress conditions. Therefore, LysU is suspected to be the main Ap_4A synthetizing enzyme in *E. coli*. However, it is important to note that it is not only hypothesis that explains, why two LysRS genes are present in *E. coli*. For instance, based on biochemical analysis, it has been proposed that LysU may be more suitable to carry housekeeping aminoacylation function in stress conditions such as low pH[95].

5.4. tRNA synthetases as RNA capping enzymes

The ability of tRNA synthetases to create Ap_4A makes them excellent candidates for post-transcriptional capping enzymes. The second step of the Ap_4A formation is known to be less substrate selective, which makes it possible that RNA could be used as a substrate. In this scenario, tRNA synthetase could utilize triphosphorylated RNA instead of ATP molecule in the second step of the reaction and create Ap_4A-RNA cap. Post-transcriptional capping by tRNA synthetases could thus explain the regulation of Ap_4A-RNA cap synthesis during stress, as dinucleoside polyphosphate caps are more abundant in stress conditions. In addition, in the same conditions, LysU, which is enzyme producing Ap_4A the most efficiently, is expressed. Therefore, LysU seems to be the most suitable candidate as the Ap_4A-RNA capping enzyme.

The possibility, that LysU could create Ap_4A-RNA cap was already reported by Luciano *et al*.[2]. They purified LysU from *E. coli* and tested if it is able to cap triphosphorylated *yeiP* transcript. Their experiment demonstrated that LysU is capable of synthetizing Ap_4A-RNA cap on *yeiP* (results can be seen on Figure 17.). However, further details about the reaction, such as the rate of cap synthesis or substrate specificity have not been tested. Also, the ability of different tRNA synthetases to create Ap_4A-RNA has not been investigated. Lastly, their experiment was performed only once without any replicates, therefore the ability to create Ap_4A-RNA must be verified.

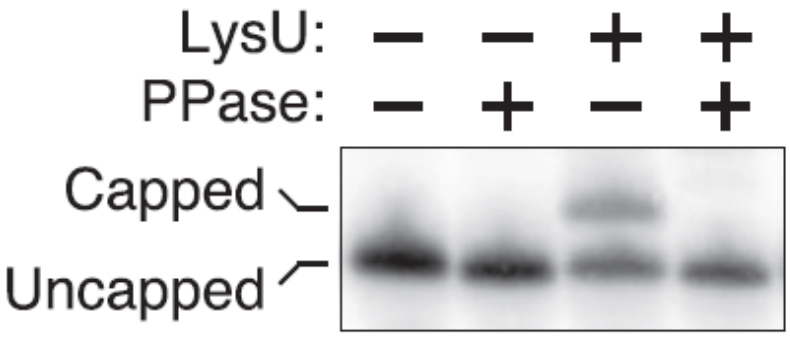


Figure 17.: Experiment showing the possibility of LysU to create Ap_4A-RNA conducted by Luciano[233]. RNA capping was analysed on boronate gel, retarded migration of LysU sample is marking capped RNA. Presence of Ap_4A-RNA cap was verified by treatment with RppH decapping enzyme.

Aims of the thesis

- Select four *E. coli* tRNA synthetases, subclone them to plasmids and express them in *E. coli.*

- Purify selected tRNA synthetases using fast protein liquid chromatography (FPLC).

- Test tRNA synthetases for production of free dinucleoside polyphosphates and find optimal reaction conditions (analysis will be performed on high-performance liquid chromatography).

- Test RNA capping potential of tRNA synthetases on *in vitro* transcribed radioactively labelled RNA.

Methods

1. Materials

1.1. Oligonucleotides

Name	Code	Sequence
LysU Forward	JaR-1	GATCATATGTCTGAACAAGAAACACG
LysU Reverse	JaR-2	ATTCTCGAGTTATTTCTGTGGGCGCATC
LysS Forward	JaR-3	GATCATATGTCTGAACAACACGCAC
LysS Reverse	JaR-4	ATTCTCGAGTTATTTTACCGGACGCATC
MetG Forward	JaR-9	GCCATTAATATGACTCAAGTCGCGAAG
MetG Reverse	JaR-10	ACTCTCGAGTTATTTCACCTGATGACC
HisS Forward	JaR-7	GCGCATATGGCAAAAAACATTCAAGC
HisS Reverse	JaR-8	ATTCTCGAGTTAACCCAGTAACGTGCG
yeiP IVT Forward	yeiP_126_F	AAAAGAATTCCAAATTAATACGACTCACTATT**A**GATTTTTGACATTTTCGACTACAGG
yeiP IVT Reverse	yeiP_126_R	TGGGCGACTGAATATCAATATCC
35mer IVT coding		CAGTAATACGACTCACTATT**A**GGGAAGCGGGCATGCGGCCAGCCATAGCCGATCA
35mer IVT complementary		TGATCGGCTATGGCTGGCCGCATGCCCGCTTCCCTAATAGTGAGTCGTATTACTG

1.2. Plasmids

Plasmids
pET-28a(+) + LysU
pET-28a(+) + LysS
pET-28a(+) + MetG
pET-28a(+) + HisS

1.3. *E. coli* strains

E. coli strains
DH5α
BL21 (DE3)

1.4. Enzymes

Enzyme	Manufacturer	Catalog number
T7 RNA polymerase	NEB	M0251L
DNase I (RNase-free)	NEB	M0303L
T4 DNA ligase	Thermo Scientific	931603
NdeI	Thermo Scientific	702390
AseI	NEB	R0526S
XhoI	NEB	R0146S
Pyrophosphatase, inorganic	Roche	10108987001
RNase T1	Thermo Scientific	EN0541
Phusion HF DNA polymerase	Thermo Scientific	F530L

1.5. Chemicals

Chemical	Manufacturer	Catalog number
Magnesium chloride ($MgCl_2$)	Sigma-Aldrich	M8266-100G
Zinc chloride ($ZnCl_2$)	Sigma-Aldrich	31650-250G
L-Methionine	Sigma-Aldrich	64319-25G-F
L-Histidine	Sigma-Aldrich	53319-25G
L-Lysine	Sigma-Aldrich	L5626-100G
EDTA	Sigma-Aldrich	EDS-500G
Tris Hydrochloride (Tris-HCl)	Fisher	BP153-1
Sodium Chloride (NaCl)	Lach-Ner	7647-14-5
Sodium phosphate dibasic (Na_2HPO_3)	Sigma-Aldrich	S0876-100G
Imidazole	Sigma-Aldrich	56750-100G
Sodium dodecyl sulfate (SDS)	Sigma-Aldrich	L3771-100G
Ammonium acetate	Sigma-Aldrich	A1542-250G
β-mercaptoethanol	Sigma-Aldrich	M6250-10ML
Dithiothreitol (DTT)	Thermo Scientific	A39255
Dimethyl sulfoxide (DMSO)	Fisher	D/4120/PB08
Triton X-100	Sigma-Aldrich	X100-100ML
Kanamycin (50mg/ml)	Sigma-Aldrich	K0254-20ML
Glycerol, anhydrous	Lach-Ner	56-81-5
Triethylammonium Acetate buffer (1 M)	Panreac AppliChem	A3846,1000
Acetonitrile	Fisher	A/0627/17
Agarose SERVA for DNA electrophoresis	SERVA	11404.05
ROTIPHORESE®Sequenziergel-Konzentrat 25%	Carl Roth	3043.1
30% Acrylamide/Bis Solution, 37.5:1	Bio-Rad	1610158
Ammonium persulfate (APS)	Sigma-Aldrich	A3678-25G
TEMED	Sigma-Aldrich	T9281-50ML
Boronate	Synthetized by M.-B. Mititelu (Cahová group)	
ATP 100 mM	New England BioLabs (NEB)	N0451A
GTP 100 mM	NEB	N0452A
CTP 100 mM	NEB	N0453A
UTP 100 mM	NEB	N0454A
[alpha-P32]GTP, 3000 Ci/mmol, 10 mCi/ml	HARTMANN ANALYTIC	FP-208
Diadenosine tetraphosphate (Ap_4A)	Sigma-Aldrich	D1262-50MG
NEBuffer r3.1	NEB	B6003S
10x T4 DNA Ligase Reaction Buffer	NEB	B0202S
10x RNA Polymerase Reaction Buffer	NEB	B9012S
10x DNase I Reaction Buffer	NEB	B0303S
RNA Gel Loading Dye (2X)	Thermo Scientific	R0641
Gel Loading Dye Purple (6x)	NEB	B7024A
SERVA DNA Stain Clear G	SERVA	39804.1
PageRuler Prestained Protein Ladder, 10 to 180 kDa	Thermo Scientific	26616
1 kb DNA Ladder	NEB	N3232L
GeneRule Ultra Low Range DNA Ladder	Thermo Scientific	SM1212
PMSF Protease Inhibitor	Thermo Scientific	36978
Pierce Protease Inhibitor Mini Tablets	Thermo Scientific	A32953
Adenosnie-$^{15}N_5$ 5'-triphosphate	Sigma-Aldrich	707783-10MG

2. Plasmid preparation

Genes encoding the selected *Escherichia coli* tRNA synthetases were PCR amplified using primers specified in table with used oligonucleotides, from *E. coli* cDNA (prepared by Dr. Pavel Vopálenský from 400 µg of total RNA isolated by RNAzol Reagent from exponential phase of growth of *E. coli* BL21 and reverse-transcribed by Luna Super RT Mix (NEB)). PCR products were subsequently analysed on 1 % agarose gel (1 h and 30 min at 40 V), visualised by SERVA stain G. Band corresponding to the length of selected genes were excised from the agarose gel and cleaned by Monarch DNA gel extraction kit (NEB). Cleaned PCR products were then cleaved with *NdeI* (*AseI* in case of *MetG*) and *XhoI* restriction enzymes. Plasmid pET-28a(+) was cleaved with *NdeI* and *XhoI* as well. Plasmid cleavage products were analysed on 0.9 % agarose gel (overnight at 20 V) and bands corresponding to plasmid with cleaved out previous construct (Microphthalmia-associated transcription factor (MITF)) excised from gel and cleaned by Monarch DNA gel extraction kit (NEB). Digested PCR amplicons with coding sequences for tRNA synthetases were then ligated to pET-28a by T4 DNA ligase (NEB, incubation at 16 °C overnight). Prepared constructs contain 6x histidine tag (His-tag) on N-terminus.

Ligation mixtures were transformed into chemocompetent ($CaCl_2$ method) *E. coli* strain DH5α. 5 µl of ligation solution was added to 50 µl of pre-prepared solution of *E. coli* DH5α used for transformation. *E. coli* was left for 30 min on ice and then put to 42 °C water bath for 30 s. Afterwards 1 ml of Luria–Bertani medium (LB medium) was added, and cells were incubated in 37 °C for 1 h, and then plated on agar plates with kanamycin. Colonies of transformed *E. coli* were selected and incubated overnight at 37 °C in LB medium with kanamycin (50 µg/mL). Plasmids were isolated from cell culture by GenJET Plasmid Miniprep Kit and transformation was checked by PCR and Sanger sequencing (SeqMe). The highest quality plasmids were selected transformed to *E. coli* BL21 with same procedure as previously described and checked again by Sanger sequencing (SeqMe). A stock with transformed bacteria was prepared by mixing 200 µl glycerol and 800 µl cell culture and frozen in -80 °C.

3. Protein purification

Buffers:

Buffer A – 50 mM Na_2HPO_4, 300 mM NaCl, 3 mM β-merkaptoethanol, 5 % glycerol, pH 7.4 (prepared in 1 l)

Buffer B – 500 mM imidazole in buffer A (prepared in 200 ml)

SEC buffer – 50 mM Tris-HCl, 100 mM NaCl, 1 mM β-merkaptoethanol, 20 % glycerol, pH 7.5 (prepared in 1 l, RNase-free)

Sonification buffer – 4% buffer B, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1x Protease Inhibitor (PI), all in buffer A (10 ml for 1g of pellet)

Lysis buffer – 1 mM PMSF, 1x PI in buffer A

His-tagged tRNA synthetases were expressed in *E. coli* BL21 transformed with plasmid pET-28a containing the selected tRNA synthetases. The test of induction was conducted by adding isopropyl β-d-1-thiogalactopyranoside (IPTG) in final concentration 1 mM (0.5 mM and 0.75 mM concentrations were tested for HisS as well) to 4 ml of cell culture (LB media, kanamycin 50 μg/mL, 37 °C) at exponential phase (optical density OD = 0.4), 4 ml of cell culture was left as control. After 3 h of incubation at 37 °C, the cell culture was centrifuged (5 min, 5000 g, 4 °C) and pellet was resuspended at 250 μl of lysis buffer. Samples were sonicated for 3 mins (5s on/10s off) on ice on Hielscher digital ultrasonic generator UP200St and then centrifuged (10 min, 20 000 g, 4 °C) to separate soluble fraction from cell debris. The resulting pellet was again resuspended at 250 μl of lysis buffer. Subsequently, samples were analysed on 8 % SDS-PAGE gel[103] (10 min at 90 V, 1 h 30 min at 110 V) and stained with Coomassie Brilliant Blue. Based on these experiments, the optimal conditions for the induction of all tRNA synthetases were determined (1 mM IPTG, 3h incubation, 37 °C).

For large-scale expression, tRNA synthetases were expressed in 1 l of LB media containing kanamycin (50 μg/mL) and induced with 1 mM IPTG at 37 °C for 3 h. The cell culture was then centrifuged (20 min, 5000 g, 4 °C) and the pellet was resuspended in 40 ml of cold phosphate-buffered saline (PBS). After another centrifugation step (10 min, 5000 g, 4 °C), the supernatant was discarded, and the pellet was stored at -80 °C.

Protein purification was performed using immobilized metal ion affinity chromatography (IMAC) followed by size exclusion chromatography (SEC) on ÄKTA pure (GE Healthcare). The pellet from the frozen cell culture was resuspended in sonification buffer and sonicated for 15 min (10s on/10s off) on ice on Hielscher digital ultrasonic generator UP200St. The sample was then centrifuged (20 min, 7197 g, 4 °C), and the supernatant was transferred to a new falcon tube. For IMAC, a HisTrap FF 1 ml column (Cytiva) was conditioned with 96 % Buffer A and 4 % Buffer B (flow rate 1 ml/min), and the lysate supernatant was applied on the column. After washing out unbound proteins, the his-tagged protein was eluted by increasing concentration of Buffer B. The 2 ml fractions were collected and analysed on 8% SDS-PAGE gel. Fractions containing desired protein were concentrated using an Amicon Ultra Centrifugal Filter 30 kDa MWCO (Merck) to a volume of less than 2 ml. Subsequently, samples were further purified by SEC using a HiLoad 16/600 Superdex 75 pg column (Cytiva) equilibrated with the SEC buffer (flow rate 1 ml/min). The 2 ml peak fractions were collected and analysed on 8 % SDS-PAGE gel, and fractions containing the selected tRNA synthetase were concentrated as described above. The protein concentration was measured using a DC protein assay (Bio-Rad) with bovine serum albumin as the standard. Glycerol was added up to the final concentration of 40 %, and protein preparation was aliquoted by 300 µl and stored in -80 °C freezer.

4. Ap_4A synthesis by tRNA synthetases

Ap_4A was synthesized in a reaction mixture composed of 100 mM Tris-HCl (pH 8), 10 mM $MgCl_2$, 0.1 mM $ZnCl_2$, 1 mM ATP, 1.5 mM L-lysine, 0.2 U PPase (Roche) and 1 µM LysU. Total reaction volume was 15 µl. Reaction was incubated at 37 °C for 15 min to 4 hours and quenched by 5 mM ethylenediaminetetraacetic acid (EDTA). Same reaction conditions were employed for reactions involving LysS. Under the similar condition, reactions with MetG and HisS, were prepared, except for the addition of $ZnCl_2$ and L-lysine. Indeed, two different amino acid were added, respectively L-methionine for MetG and L-histidine for HisS. For the synthesis of other Ap_4Ns, a mixture of ATP:NTP in a 1:1 ratio was used, while for the Ap_3A synthesis a mixture of ATP:ADP in a ratio 1:1 was prepared. Then the reaction was analysed on HPLC by method previously described in Mititelu *et al.*[52].

5. *In vitro* transcription

For *in vitro* transcription (IVT), a 35 bp long oligonucleotide from Hudeček *et al.*[1] was used. Two complementary DNA strands (listed in the table with used oligonucleotides) were annealed by mixing 40 µM of each, heated to 90 °C and slowly cooled down to 20 °C. The second IVT template was adopted from Luciano *et al.*[2] and modified for purposes of this thesis. A 126 bp long region of 5′ end of *yeiP* gene was PCR amplified from *E. coli* cDNA (prepared by Dr. Pavel Vopálenský as described in methods section 2.). The PCR primers are listed in table with oligonucleotides, with the forward primer including a T7 promoter. The PCR product was purified by Monarch PCR and DNA Clean up kit (NEB) and checked on 2 % agarose gel (1 h 10 min at 80 V, stained by SERVA stain G). The cleaned PCR product was directly used as template for the IVT.

The IVT mixture was containing T7 RNA polymerase buffer (NEB), 5% DMSO, 4.8 mM $MgCl_2$, 0.12 % Triton-X 100, 12 mM DTT, 1 mM ATP, 1 mM UTP, 1 mM CTP, 0.8 mM GTP, 0.4 µl [αP^{32}]GTP (activity 9.25 MBq/250 µCi in 25 µl), 62.5 units of T7 RNA polymerase (NEB) and approx. 120 ng of 35mer template or approx. 30 ng of yeiP_126 template. For the synthesis of uncapped transcripts, ATP concentration was increased to 2.6 mM, and for capped transcripts, 1.6 mM Ap_4A was added. The total reaction volume was 25 µl. The IVT reaction was incubated for 2 hours at 37 °C. Afterwards, DNase treatment was carried out by adding 3 µl of 10x DNase buffer (NEB), 1 µl of DNase I (NEB) and water (molecular biology grade) was added to reach the final volume of 30 µl. The mixture was incubated for 45 min at 37 °C and then for 10 min at 75 °C to inactivate the enzyme. Samples were cleaned by Micro Bio-Spin P-6 Gel Colums, Tris buffer (Bio-Rad) and the RNA concentration was measured by Qubit RNA High Sensitivity (HS) kit (Invitrogen). Non-radioactive IVT was prepared in the same manner, and the ratio of radioactive GTP was substituted by the addition of 1 mM GTP. Non-radioactive IVT samples were cleaned by ZymoClean RNA clean and concentrator (Zymo Research).

6. RNA capping by tRNA synthetases

Reaction conditions for testing RNA capping potential were the same as for production of free Ap_4A described above but with the addition of 20 ng to 150 ng of *in vitro* transcribed radio-labelled RNA. Reaction was incubated for 2 or 4 hours and quenched by addition of 5 mM EDTA. The 2x RNA loading dye was added to the solution to reach

the final volume of 33.2 µl, and samples were heated for 10 min at 75 °C and analysed on polyacrylamide gel electrophoresis (PAGE).

The 35mer RNA (35 nt long) was separated on a 12 % PAGE gel and the yeiP_126 RNA (126 nt long) on 8% Boronate-PAGE gel[65]. 12 % PAGE gel was prepared by mixing 21 ml of 2x TBE buffer, 21 ml of 25 % Rotiphorese, 340 µl of 10 % APS and 16.8 µl TEMED and poured to middle sized sequencing gel form and kept overnight in 4 °C to solidify. Before loading samples, the gel was pre-runned at 600V for 20 min. The 8 % Boronate-PAGE gel was prepared by mixing 21 ml of 2x TBE, 14 ml of 25 % Rotiphorese, 7 ml of 7 M urea, 150 mg of boronate, 340 µl of 10 % APS and 16.8 µl of TEMED and prepared in the same way as for 12 % PAGE gel.

The RNA was separated after 4 hours at 600 V. Radioactivity of samples was measured by leaving the gel in GE Healthcare Storage Phosphor Screen cassette for 2 hours to overnight, based on radioactivity levels, and then screened by phosphor imaging on Amersham Typhoon Biomolecular Imager.

7. LC-MS analysis of RNA capping

Non-radioactive *in vitro* transcribed RNA (200 ng) was incubated for 2 hours with LysU under the same conditions as before and then cleaned using ZymoClean RNA clean and concentrator. The RNA was subsequently digested with 1000 units of RNase T1 (Thermo Scientific) in reaction containing 2.5 mM ammonium acetate (pH 7.5) and 25 µM EDTA. Total reaction volume was 40 µl. The reaction was incubated at 37 °C for 30 min and then analysed on Liquid Chromatography coupled with Mass Spectrometry (LC-MS) by Dr. Anton Škríba.

For the chromatographic separation, HPLC (Acquity H-class, Waters) equipped with Xbridge Premier BEH amide column (2.5 µm, 4.6 mm X 150 mm, Waters) was used. Mobile phase A contained 10 mM ammonium acetate in 90% acetonitrile and mobile phase B was composed of 10 mM ammonium acetate in an ultrapure water. Flow rate was set to 1 ml/min and the gradient of separation was following: 0 min 20 % B, 2 min 20 % B, 8 min 50 % B, 10 min 60 % B, 13 min 60 % B, 13.1 min 20 % B, 20 min 20 % B. Autosampler was kept at 22°C and 50 µl of sample was injected. Mass spectrometry detection was performed with electrospray ionization source equipped Xevo G2-XS QTof mass spectrometer (Waters) in 300-1600 m/z range.

8. Molecular dynamics simulations

LysU structures were downloaded from the Protein Data Bank (www.rcsb.org) − PDB codes 5YZX and 1E22 – and were prepared for molecular dynamics (MD) by CHARMM-GUI Solution Builder[104]. 5YZX chain B was read with B4P ligand (Ap_4A) and without crystallographic ions. Solution pH was set to 8.0, missing residues were automatically modelled and ASP and GLU residues were set to be deprotonated. 1E22 chain A was read with ACP ligand (non-hydrolysable ATP analogue) and three Mg^{2+} crystallographic ions (with other parameters being the same as for 5YZX). Both structures were fitted into a cubic water box extending 10 Å from the solute. For 5YZX, Na^+ and Cl^- were added up to 0.15 M and neutralizing the system, while for 1E22, the ions were Mg^{2+} and Cl^- and the concentration was 0.01 M (Monte-Carlo placing method). Topology, coordinate and parameter files for MD simulation in GROMACS were generated by CHARMM-GUI Force Field Converter (CHARMM36m + WYF (cation-pi interactions)).

MD simulations were performed by Dr. Martin Lepšík from High Performance Computing Core Facility of IOCB Prague using GROMACS 2022 (https://www.gromacs.org). The systems were run in NpT ensemble at 1 bar pressure and 303.15 K temperature. Movements of protein and ligands were restrained, only ions and water molecules were free to move. After initial minimization and equilibration, the production simulation run was 40 ns.

Results from MD were visualised using Visual Molecular Dynamics (VMD)[105]. Occupancies of ion sites were analysed using the VOLMAP plugin of VMD. MD trajectories were aligned to the crystallographic structures for visualisation of ion sites. Occupancies were set to isovalue of 0.15 for visualization of stable ion binding sites and to isovalue 0.03 for visualization of all possible ion binding sites. Pictures were rendered using Tachyon[106].

9. Purification of isotopically labelled Ap_4A

Diadenosine-$^{15}N_{10}$ tetraphosphate ($^{15}N_{10}$-Ap_4A) was synthesized by LysU from adenosine-$^{15}N_5$ triphosphate under the same reaction conditions as already described, with the total volume increased to 450 µl. The reaction was incubated for 2 hours 30 minutes and quenched by 5 mM EDTA. Two scaled up reactions were separated by HPLC using the method published in Mititelu *et al*.[52] (injection volume was increased to 150 µl). The fraction corresponding to Ap_4A (from 9.1 to 9.6 min) was collected and separated by solid phase extraction (SPE) from the triethylammonium ions used in HPLC (Otto SPEcialist Positive Pressure Manifold from Waters with Oasis WAX Cartridge 1cc,/10mg, 30µM 100/PK column was used for SPE). In the final step of the SPE, samples were eluted in 500 µl of mixture of acetonitrile, H_2O and NH_4OH (73:25:2). The elution solution was evaporated on SpeedVac Vacuum Concentrator (36 °C for approx. 7 hours). The obtained product was sent to nuclear magnetic resonance (NMR) spectroscopy for validation (^{1}H NMR and ^{31}P NMR – performed by Dr. Radek Pohl from NMR Spectroscopy group at IOCB Prague). After the NMR analysis, the sample was overnight lyophilised to purify it from remaining acetonitrile, weighted and stored in -80 °C freezer.

Results

1. Plasmid preparation

Genes of selected tRNA synthetases were PCR amplified from *E. coli* cDNA library (Figure 18.), cleaved with restriction enzymes (*NdeI*, *AseI* in case of *MetG* and *XhoI)* and subsequently ligated to pET-28a plasmid. Four out of five selected genes were successfully transformed to *E. coli* BL21 (*AlaS* was not transformed and therefore not used in later experiments). Plasmid transformation was checked by Sanger sequencing. Alignment of obtained sequences to *E. coli* genome by BLAST[107] showed that all transformed genes were 100 % identical to naturally occurring protein sequences. Maps of prepared plasmid can be seen on Figures 20. to 23..

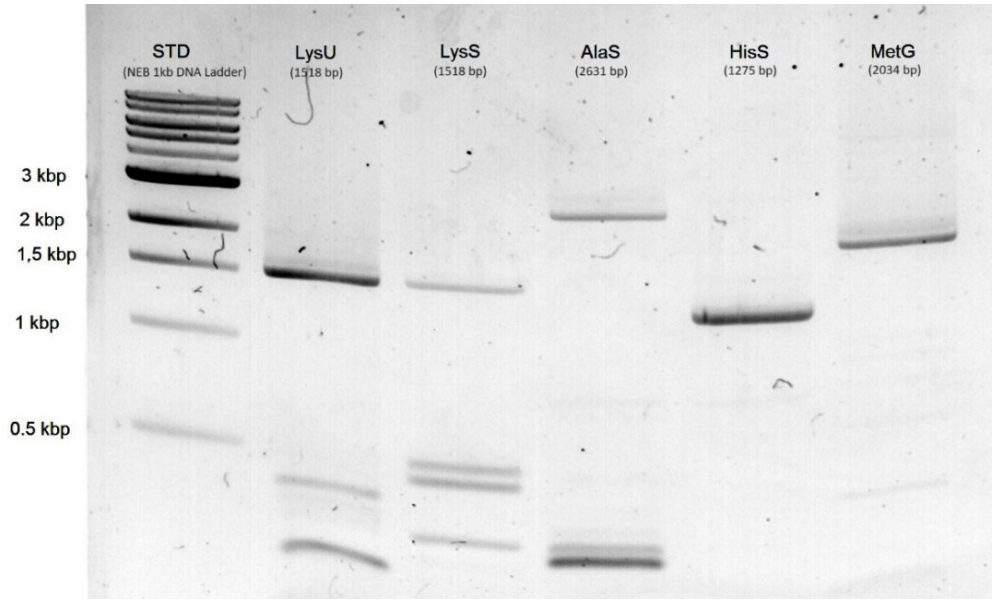


Figure 18.: 1 % agarose gel of PCR amplified tRNA synthetases.

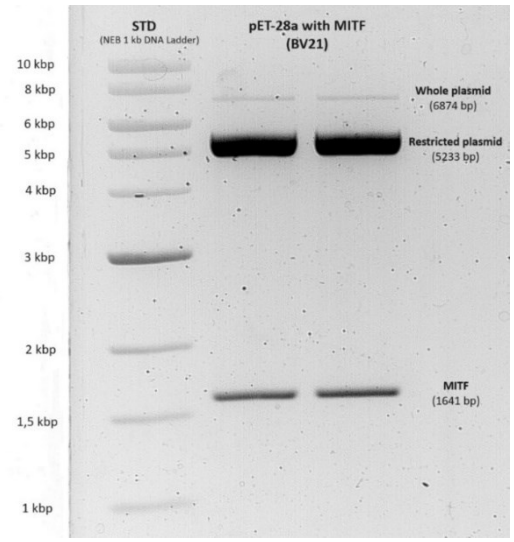


Figure 19.: 0.9 % agarose gel with cleaved pET-28a.

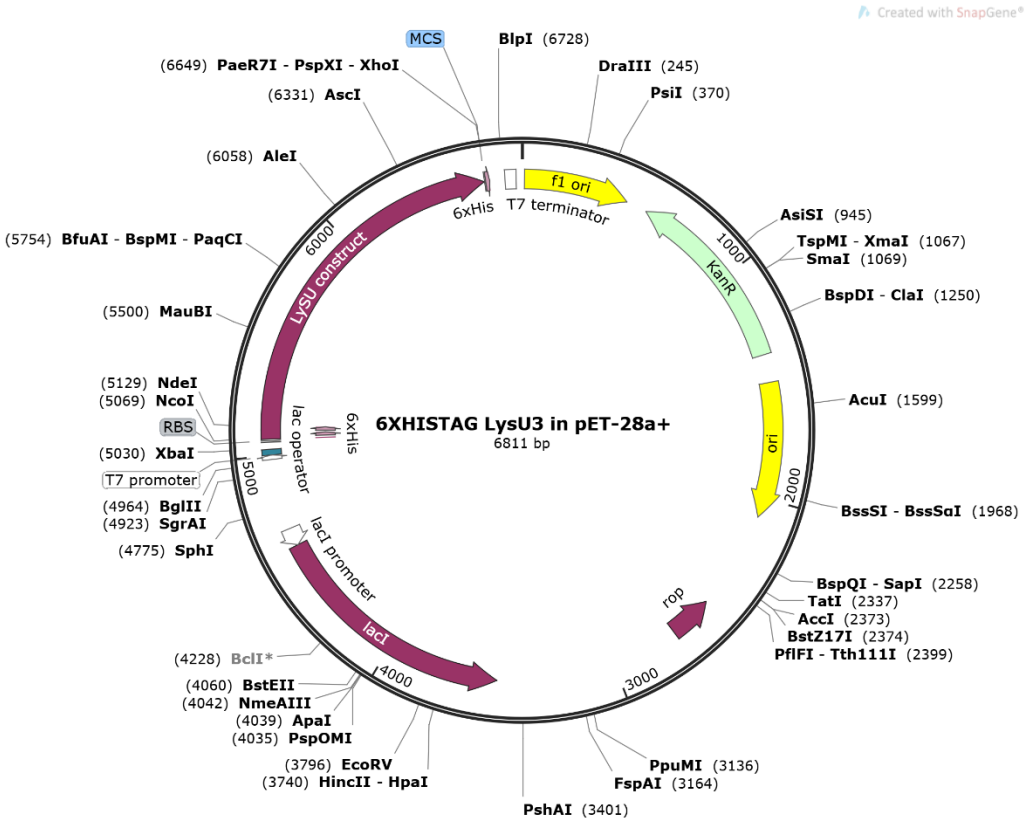


Figure 20.: Plasmid map of pET-28a with LysU.

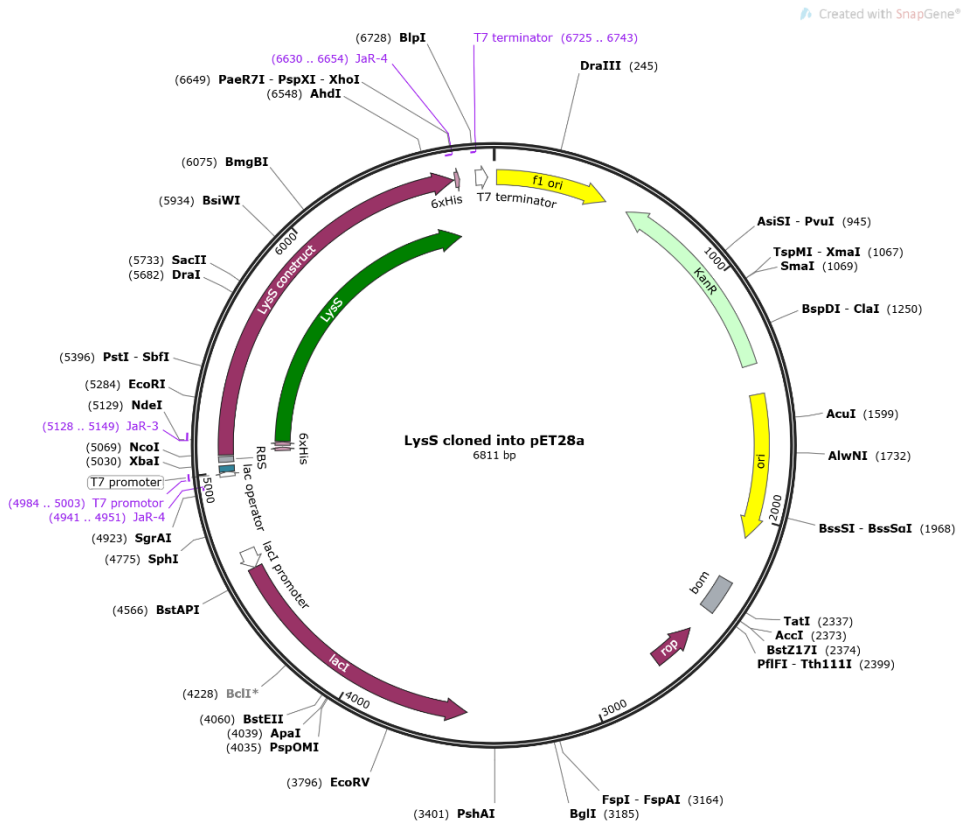


Figure 21.: Plasmid map of pET-28a with LysS.

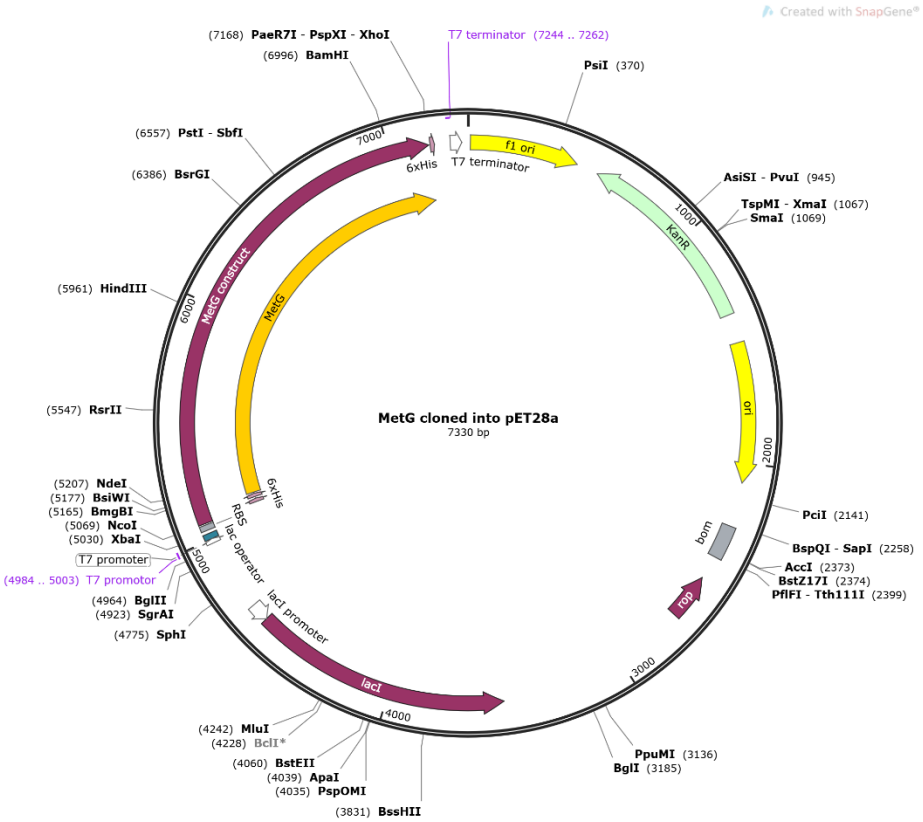


Figure 22.: Plasmid map of pET-28a with MetG.

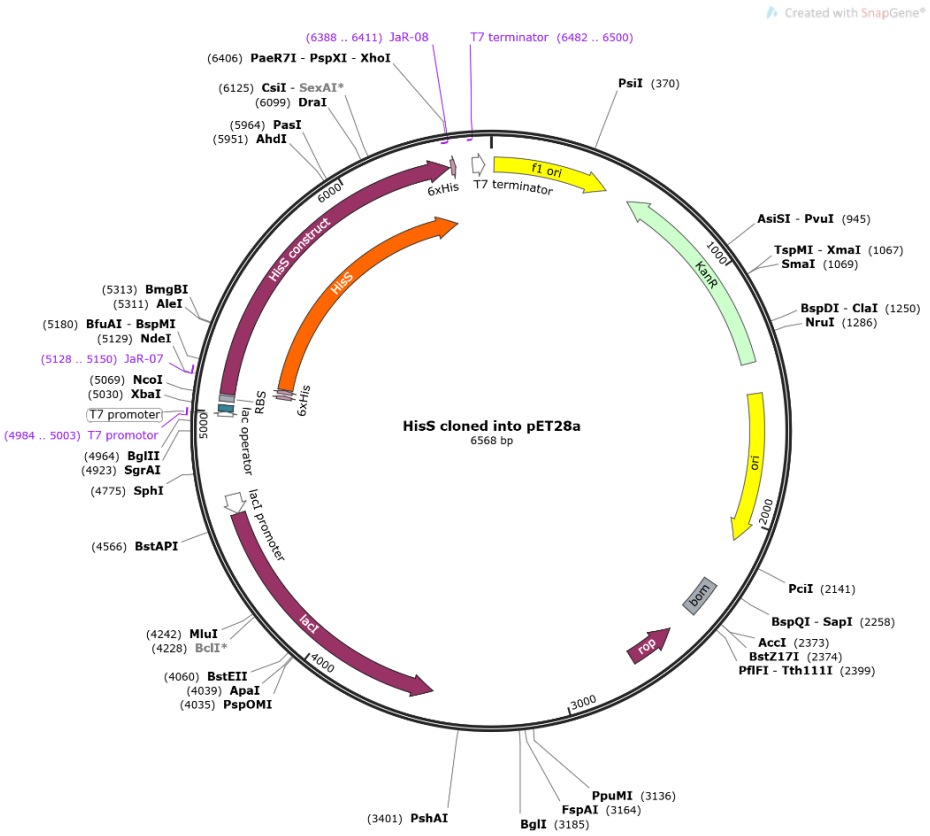


Figure 23.: Plasmid map of pET-28a with HisS.

2. Protein purification

E. coli BL21 culture was grown in LB medium to exponential phase (OD = 0.4), to which IPTG in final concentration 1 mM was added (in case of HisS, 0.5 mM, and 0.75 mM final concentrations of IPTG were tested as well). Cell culture was incubated for 3 hours, soluble fraction was then separated from cell debris and analysed on SDS-PAGE (Figures 24. and 25.). Based on these experiments, we found that the best conditions for expression of all tested tRNA synthetases are 1 mM IPTG and 3 hours of incubation in LB medium at 37 °C.

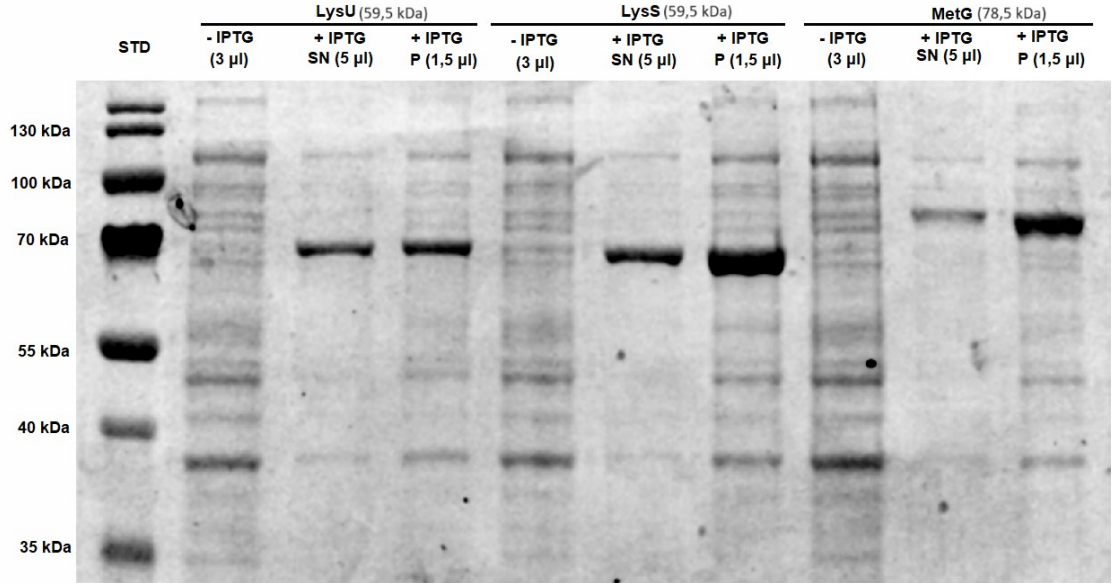


Figure 24.: 8 % SDS-PAGE gel. Test of induction of LysU, LysS and MetG. First line for each protein represents control without addition of IPTG, second line represents soluble fraction (SN = supernatant), and third line represents cell debris (P = pellet). Volume of analysed samples is indicated in brackets.

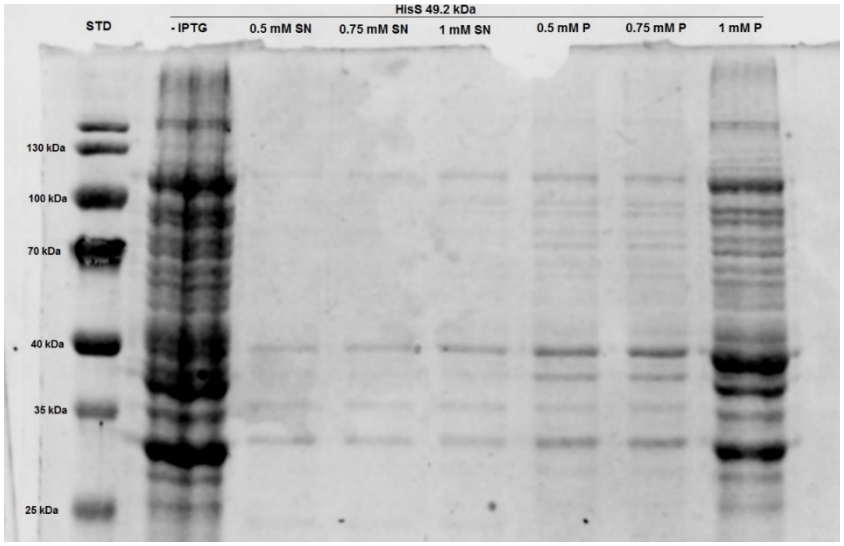


Figure 25.: 8 % SDS-PAGE. Test of induction of HisS. First line after standard represents control without IPTG, lines 3 to 5 represents soluble fraction (SN = supernatant) and lines 6 to 8 represent cell debris (insoluble pellet).

All selected tRNA synthetases were successfully purified employing fast protein liquid chromatography (FPLC) composed of two steps, immobilized metal ion affinity chromatography (HisTrap) followed by size exclusion chromatography (SEC) on ÄKTA pure (Figures 27. and 28.). According to the SDS-PAGE analysis of SEC purified enzymes, the purity was sufficient for further experiments. The protein yield of LysU was 12.5 mg; of LysS was 12.1 mg; of MetG was 12.5 mg and of HisS was 12.75 mg. With the successful purification of tRNA synthetases, we can proceed to the next step, which is the synthesis of free dinucleoside polyphosphates by tRNA synthetases.

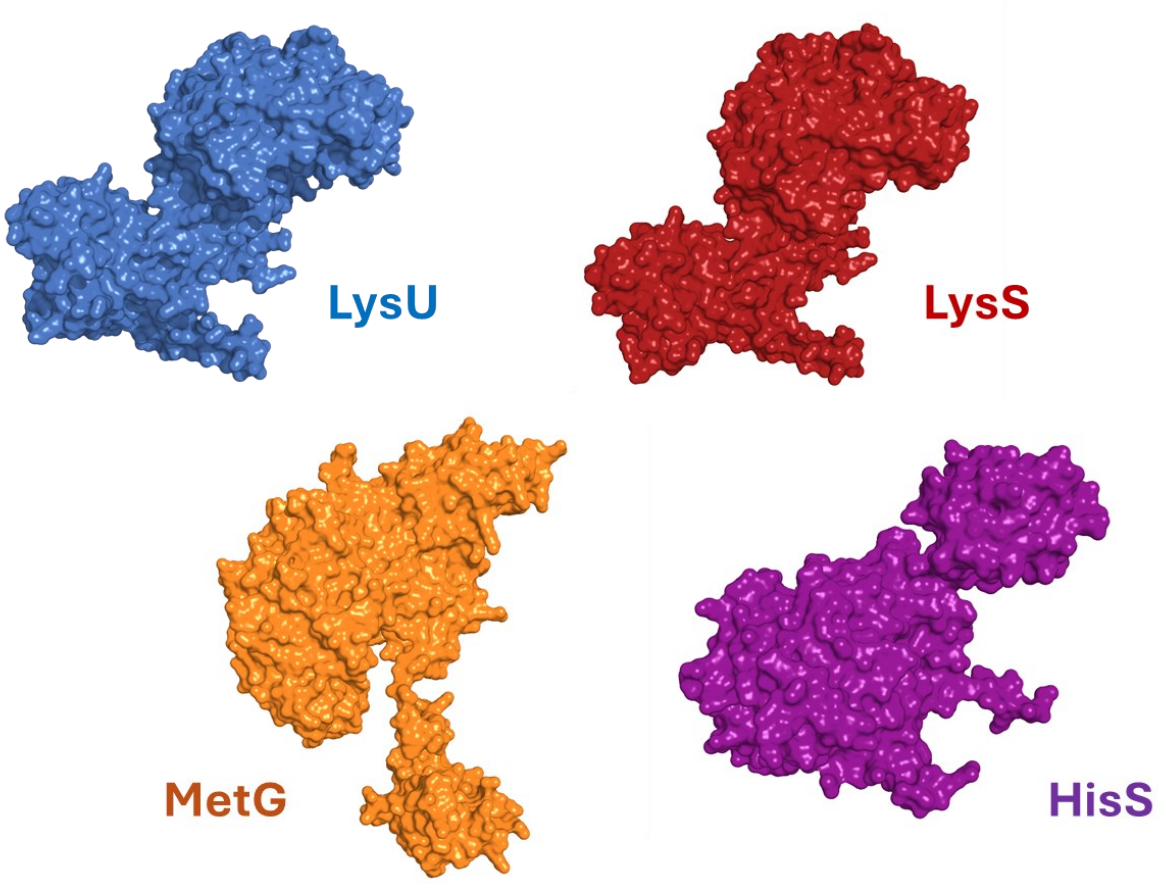


Figure 26.: Illustration of structures of purified tRNA synthetases. All structures were predicted by AlphaFold2[10] (using ColabFold[12]) and visualised in PyMOL.

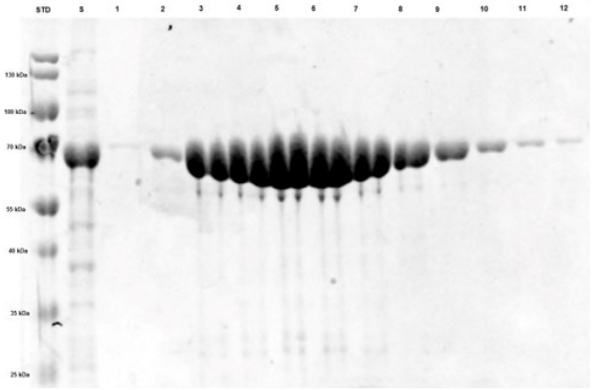
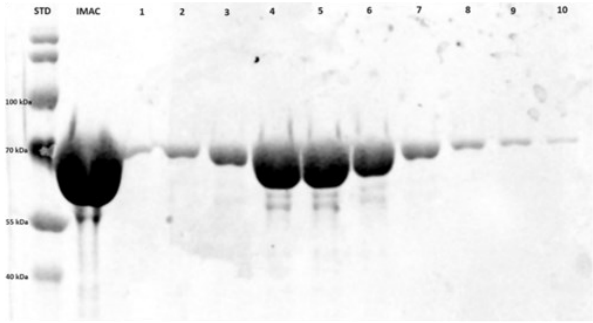
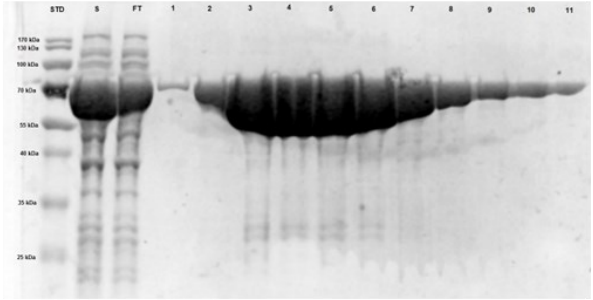
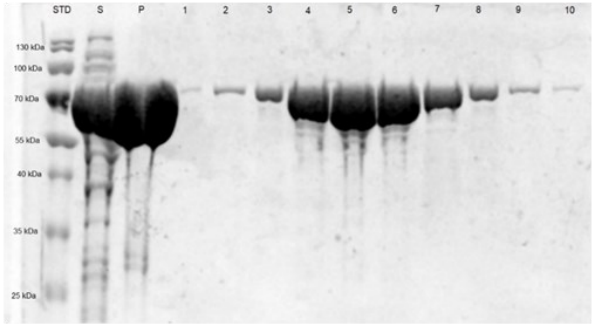
tRNA synthetase	HisTrap (SDS-PAGE)	SEC (SDS-PAGE)	Total protein yield
LysU (59.5 kDa)			**12.5 mg**
LysS (59.5 kDa)			**12.1 mg**

Figure 27.: Protein purification of LysU and LysS. In the second row, SDS-PAGE gel after IMAC (HisTrap) is present, in the third row, there is SDS-PAGE gel after SEC, and in the last row, total protein yield is indicated. SDS-PAGE gel represents fractions collected during purification. For a comparison, gels also include the standard (PageRuler Prestained Protein Ladder, 10 to 180 kDa from Thermo Scientific) and all proteins isolated from cells, or proteins after first step of purification.

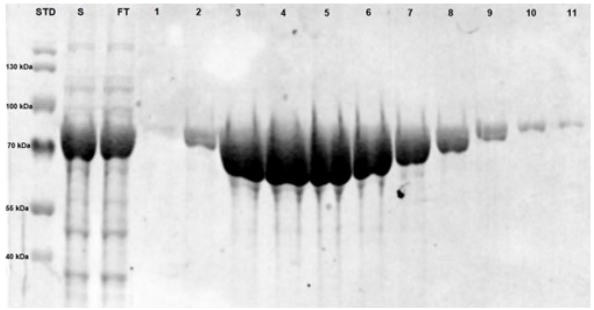
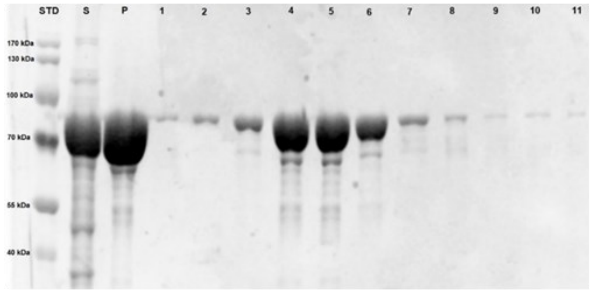
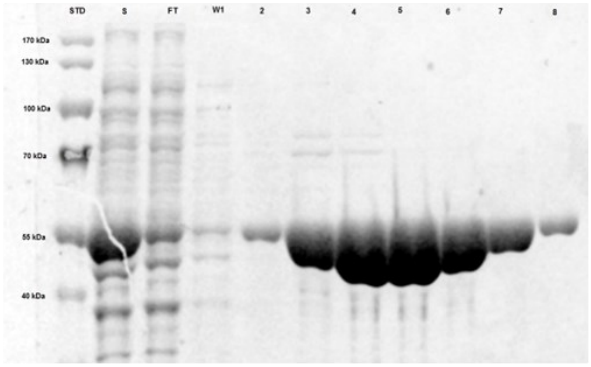
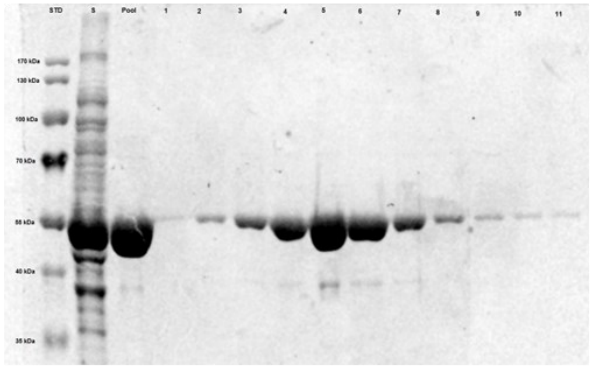
tRNA synthetase	HisTrap (SDS-PAGE)	SEC (SDS-PAGE)	Total protein yield
MetG (78.5 kDa)			**12.5 mg**
HisS (47 kDa)			**12.75 mg**

Figure 28.: Protein purification of MetG and HisS. In the second row, SDS-PAGE gel after IMAC (HisTrap) is present, in the third row, there is SDS-PAGE gel after SEC, and in the last row, total protein yield is indicated. SDS-PAGE gel represents fractions collected during purification. For a comparison, gels also include the standard (PageRuler Prestained Protein Ladder, 10 to 180 kDa from Thermo Scientific) and all proteins isolated from cells, or proteins after first step of purification.

3. Synthesis of Ap_4A and other Ap_nNs by tRNA synthetases

First, optimal reaction conditions for Ap_4A synthesis were tested on LysU tRNA synthetase. Indeed, LysU was able to synthesize Ap_4A under published conditions[2,45]. Ap_4A synthesis was followed for 4 hours (Figure 29.). The amount of formed Ap_4A grown in the following way: 187 µM (± 47 µM) in 15 min; 260 µM (± 83 µM) in 30 min; 333 µM (± 79 µM) in 1 h; 370 µM (± 48 µM) in 2h; 355 µM (± 100 µM) in 3h and 323 µM (± 128 µM) in 4 h. During first 2 hours of incubation, all ATP was already transformed into Ap_4A. During incubation, side product of reaction like ADP and AMP slowly appeared (slight reversibility of the reaction leads to Ap_4A hydrolysis[100]). Ap_3A could be detected after longer incubation time (Figure 30.). Ap_3A is potentially synthetized from lysyl-AMP intermediate reacting with ADP[13,100]. According to this experiment, 2 hours of incubation with 1 µM enzyme were concluded as optimal reaction conditions.

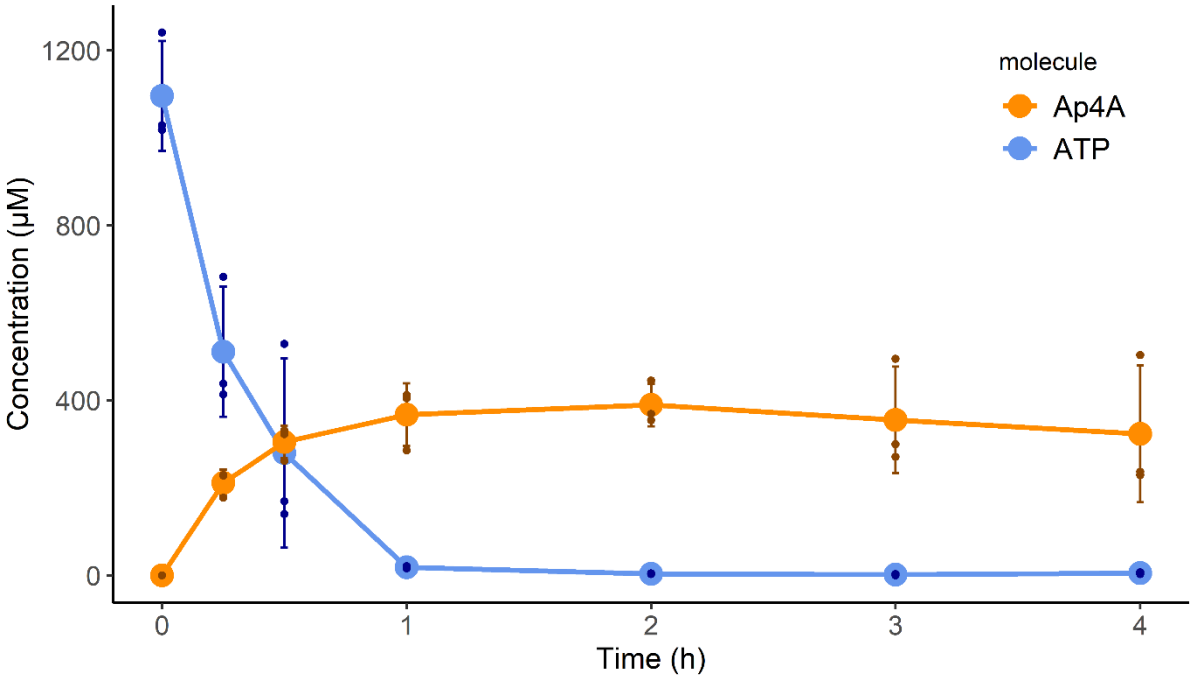


Figure 29.: Graf showing changes of concentration of ATP (blue) and Ap_4A (orange) over the time in reaction with 1 µM LysU. The error bar represents the standard deviation and dots are individual values from each replicate. Performed in triplicates.

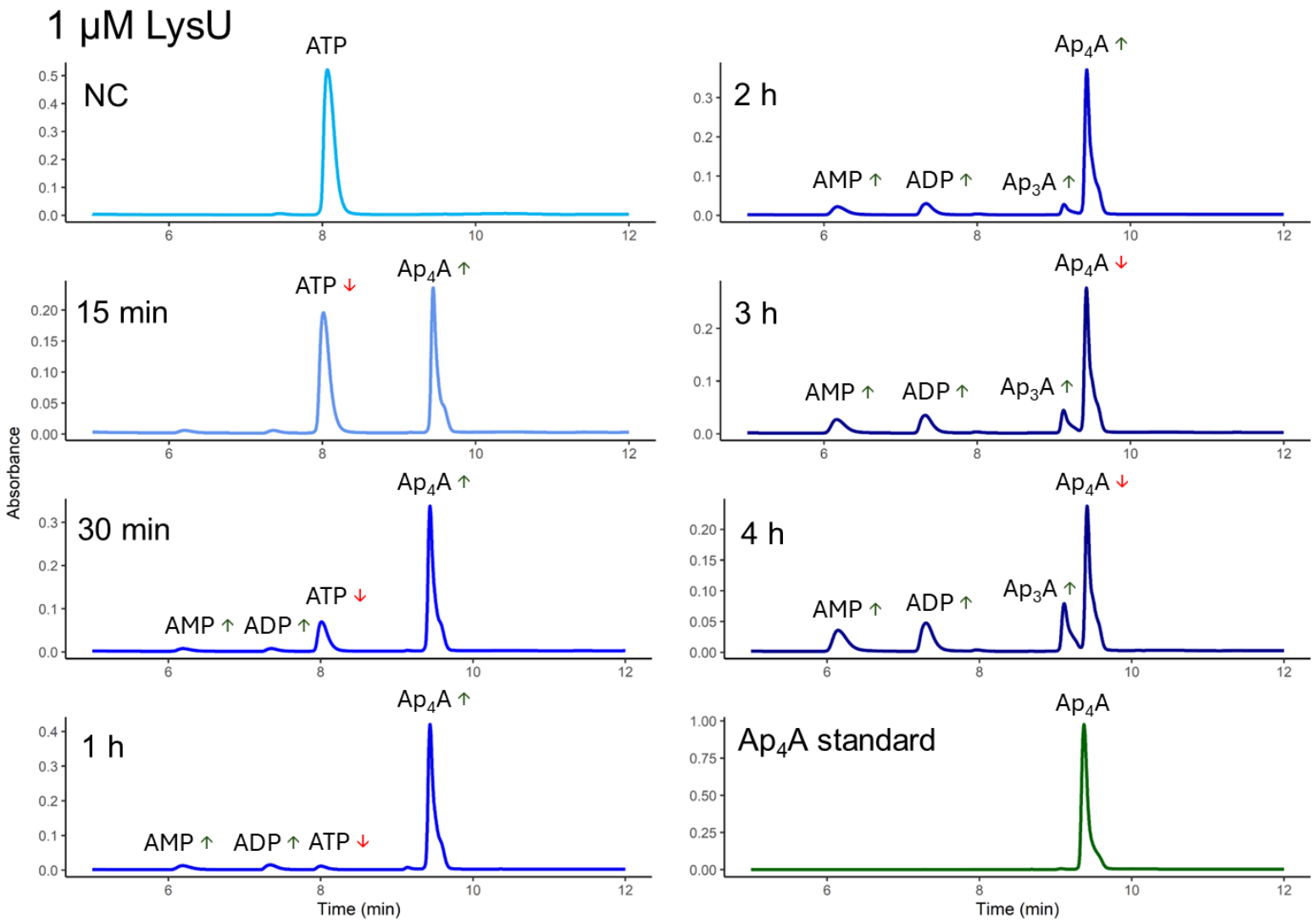


Figure 30.: HPLC chromatograms from Ap_4A synthesis by 1 µM LysU (NC, 15min, 30 min, 1 h, 2 h, 3 h, 4 h and Ap_4A standard). The arrows indicate the change of concentration from previous measurement (↑ - concentration increased, ↓ - concertation decreased).

The efficiency of Ap_4A formation was compared for all purified tRNA synthetases (Figure 31.). After 2 hours of incubation, concentration of Ap_4A reached 370 μM (± 48 μM) in reaction with LysU; 171 μM (± 64 μM) in reaction with LysS; 3.7 μM (± 0.7 μM) in reaction with HisS; and 0.19 μM (± 0.06 μM) in reaction with MetG. We observed that LysS is approximately 2.2 times slower in Ap_4A synthesis than LysU, HisS is approx. 101 times slower and MetG is approx. 1947 times slower than LysU. These observations are in accordance with known data from literature. LysU and LysS are therefore enzymes, that have the highest probability to be able to synthetize Ap_4A-RNA cap.

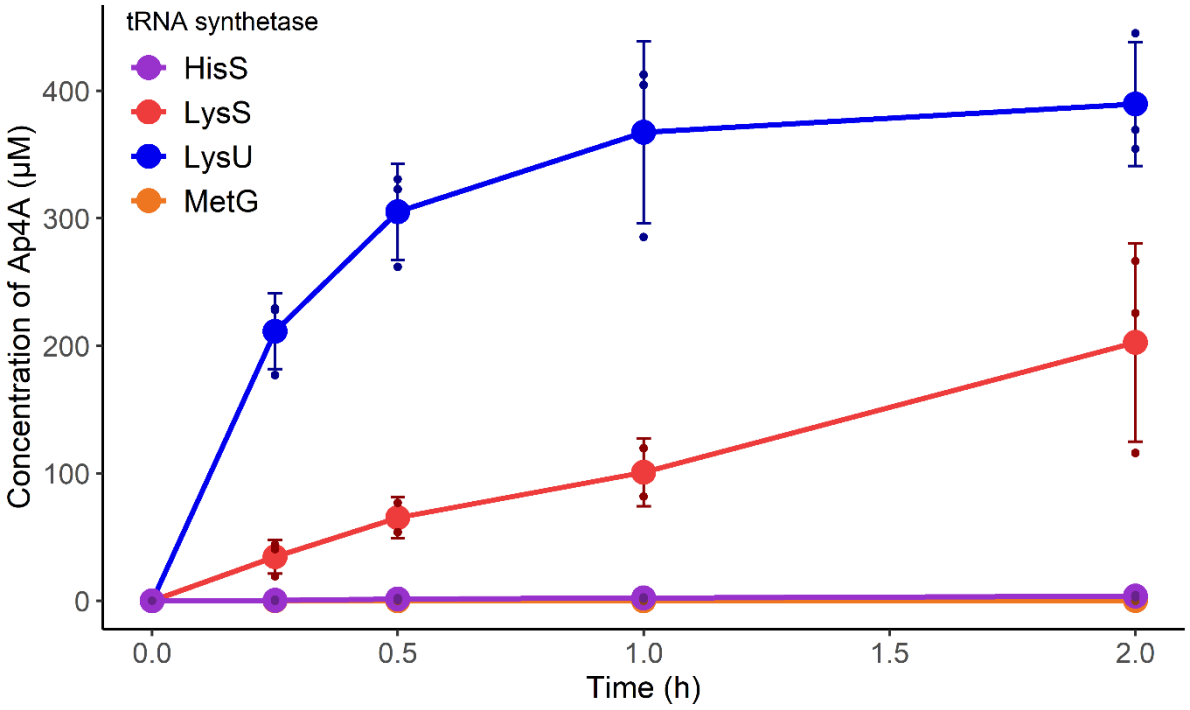


Figure 31.: Comparison of capability to synthetise Ap_4A by all selected tRNA synthetases (1 μM). LysU is in blue, LysS in red, HisS in purple and MetG is in orange. The error bar represents the standard deviation and dots are individual values from each replicate. Performed in triplicates.

We tested whether LysU can form also other Ap_nNs than Ap_4A. We observed that other Ap_nNs have been successfully synthetized from mixtures containing ATP and other nucleotide in 1:1 ratio. Specifically, Ap_3A, Ap_4G, Ap_4C and Ap_4U have been detected (Figures 32. to 35.). Synthesis of Ap_nNs is always slower than synthesis of Ap_4A (Figure 37.), which is in accordance with the data in literature. Since tRNA synthetases needs ATP for the first step of the reaction, they are not able to synthetize Np_nNs that do not contain adenine. This was confirmed by inability of LysU to create Gp_4G from reaction containing only GTP (Figure 36.).

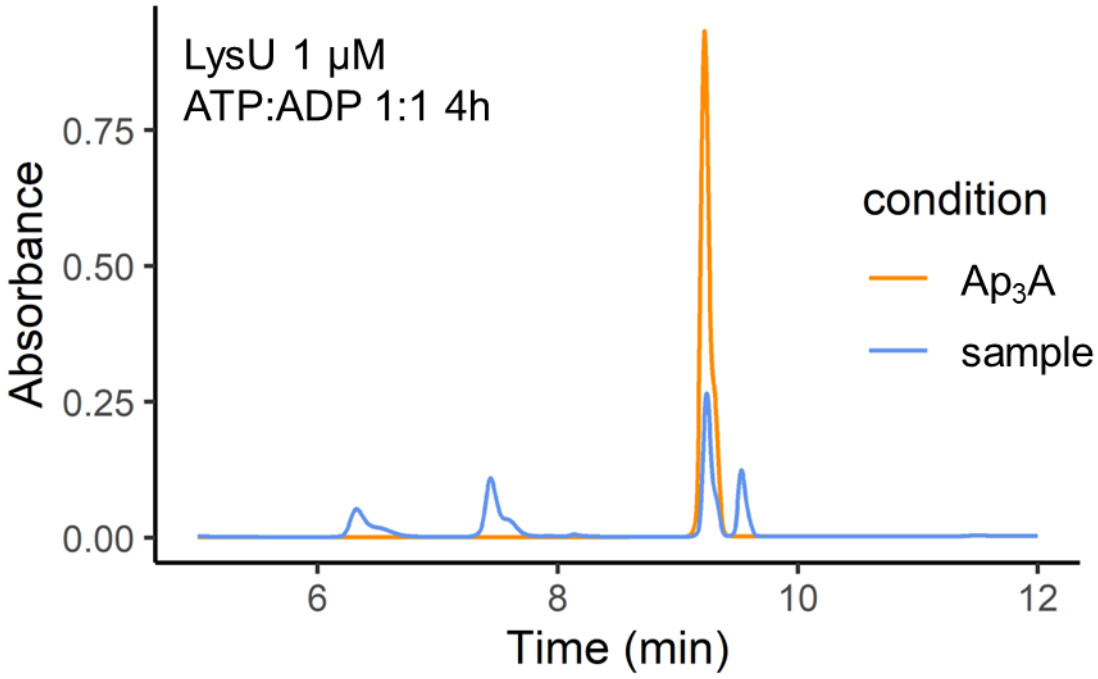


Figure 32.: Synthesis of Ap_3A by LysU from ATP:ADP mixture in 1:1 ratio. Ap_3A was identified based on available standard.

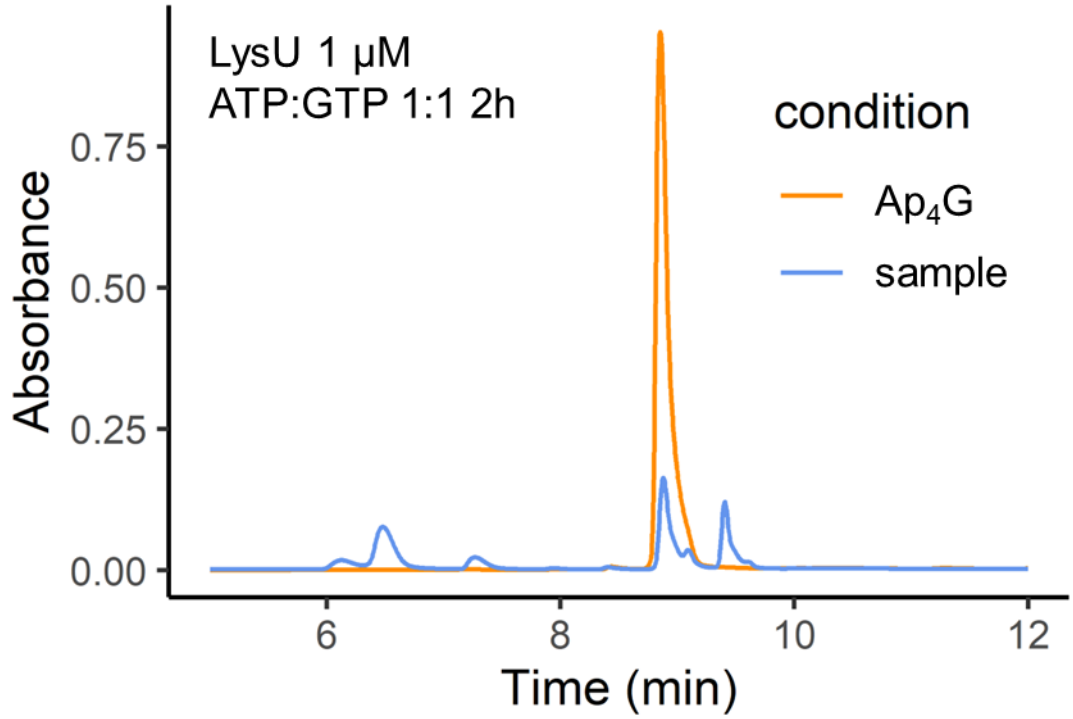


Figure 33.: Synthesis of Ap_4G by LysU from ATP:GTP mixture in 1:1 ratio. Ap_4G was identified based on available standard.

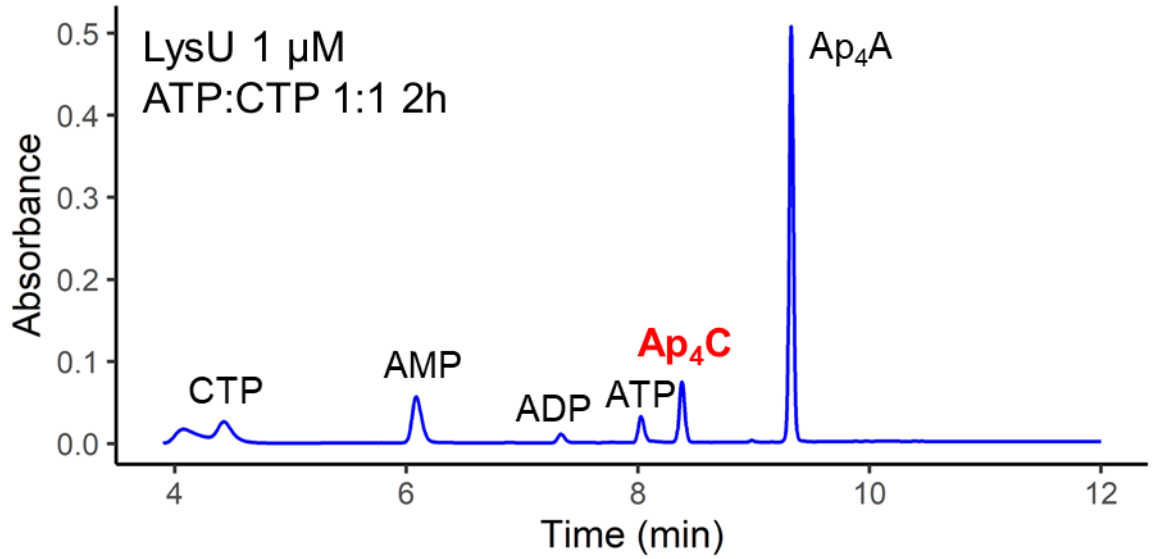


Figure 34.: Synthesis of Ap_4C by LysU from ATP:CTP mixture in 1:1 ratio. Ap_4C was identified based on exlusion of all other peaks, because Ap_4C standard was not available. Please note, that because R_f of CTP is before 5th minute, the x axis of this graph have been changed compared to other chromatograms.

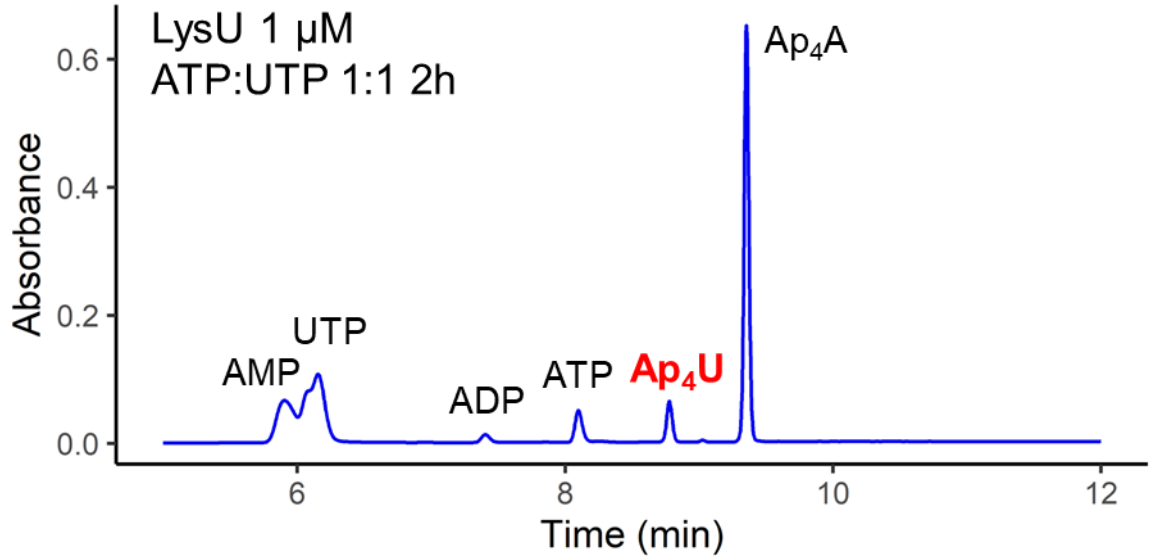


Figure 35.: Synthesis of Ap_4U by LysU from ATP:UTP mixture in 1:1 ratio. Ap_4U was identified based on exclusion of all other peaks, because Ap_4U standard was not available.

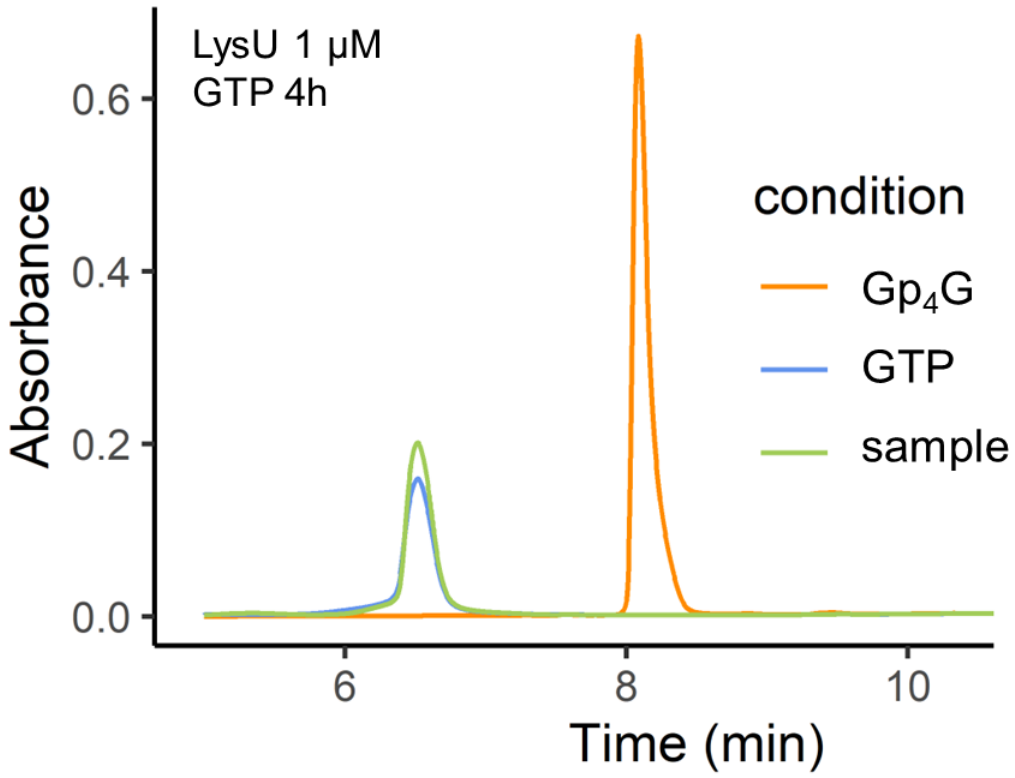


Figure 36.: LysU reaction with only GTP. Synthesis of Gp_4G was not observed.

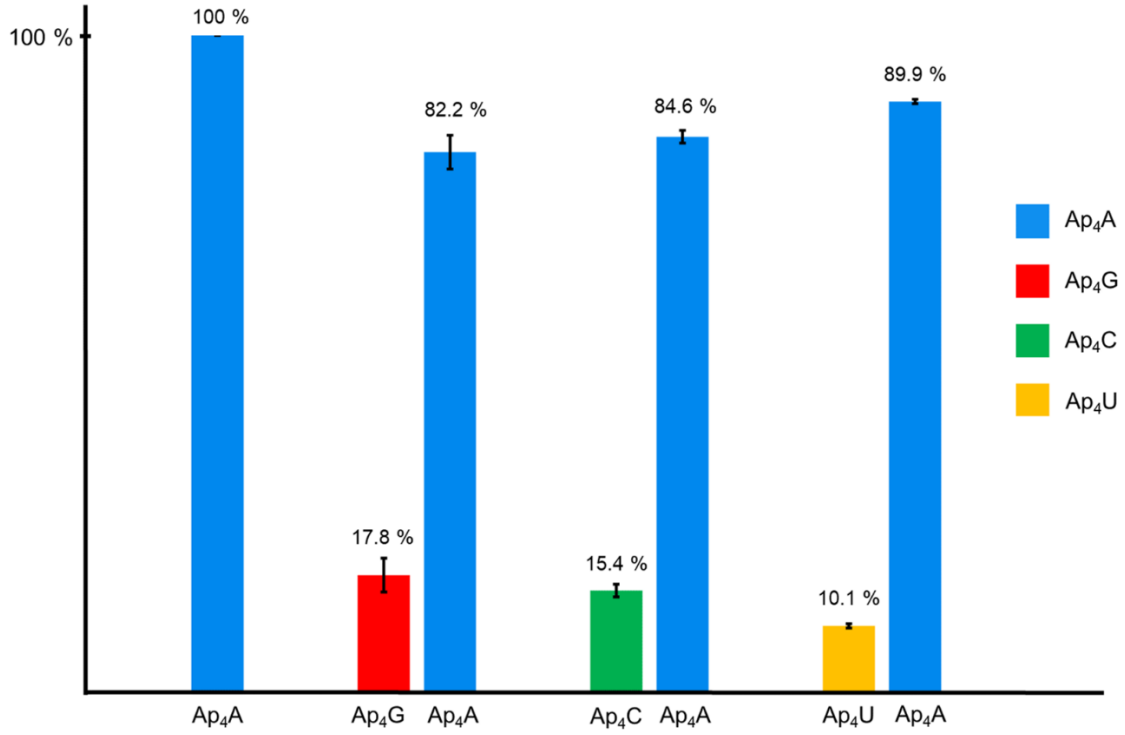


Figure 37.: Bar plot showing different rates of Ap_4N synthesis. Mixture of ATP:NTP in 1:1 ration was incubated for 2 hours with 1 µM LysU. Absorbance areas from HPLC chromatograms of Ap_4A and Ap_4N were summed and percentage of Ap_4N and Ap_4A was calculated. The error bar represents the standard deviation. Performed in triplicates.

4. RNA capping by tRNA synthetases

Once the conditions for formation of free Ap_4A were established, we applied them on *in vitro* transcribed RNA. tRNA synthetases were tested for the ability to create an Ap_4A-RNA cap on RNA 35mer bearing triphosphate at 5′ end. Tested conditions included 2 and 4 hours of incubation with 1 μM enzyme. In case of LysU and LysS, we also tested incubation of RNA with increased enzyme concentration (5 μM and 10 μM). The results of the reaction were analysed on a 12 % PAGE gel and compared with an Ap_4A-RNA standard created by IVT, as well as with negative controls (Figures 38. and 39.). The experiment was performed in a duplicate, while the experiment with an increased concentration of LysU and LysS was performed only once (Figure 40.).

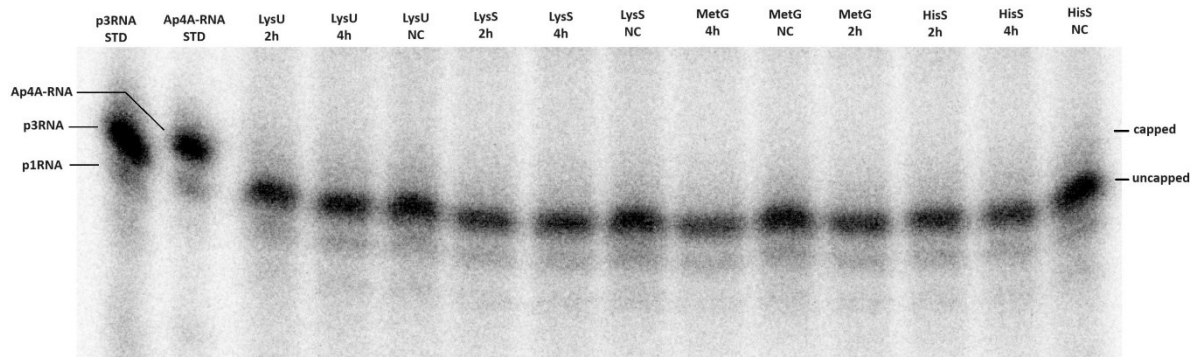


Figure 38.: 12 % PAGE gel. Analysis of RNA capping by tRNA synthetases on 35mer. In the first line, there is a triphosphate RNA standard, and in the second line, there is Ap_4A-RNA standard. Standards are followed by reaction with tRNA synthetase for 2 hours, 4 hours and negative control (in case of MetG, order was accidentally swapped). Position, where capped transcripts would be present is indicated on label on the left. The presence of Ap_4A in the samples was confirmed by HPLC analysis. First replicate.

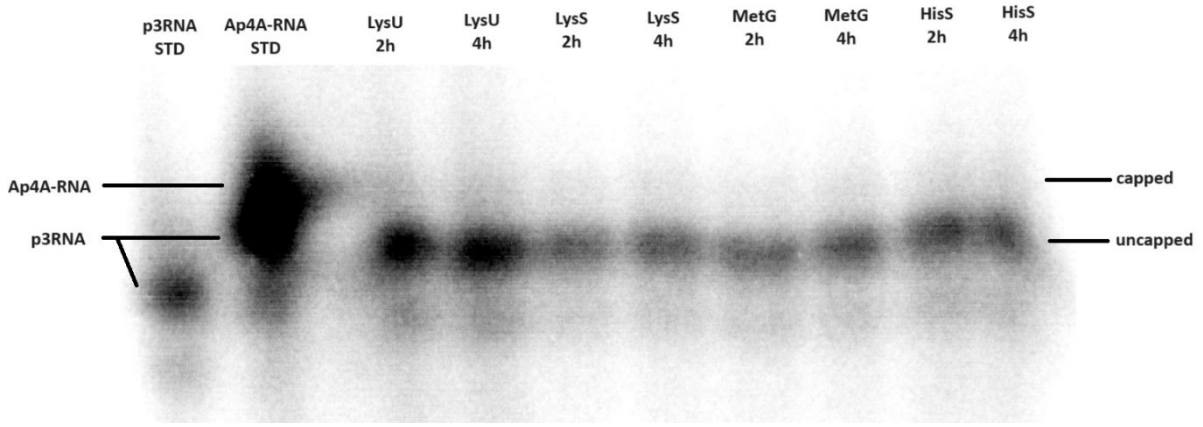


Figure 39.: 12 % PAGE gel. Analysis of RNA capping by tRNA synthetases on 35mer. In fist two lines, there is p_3RNA and Ap_4A-RNA respectively. Standards are followed by samples of p_3RNA incubated with tRNA synthetase (for 2 h and 4 h). Second replicate.

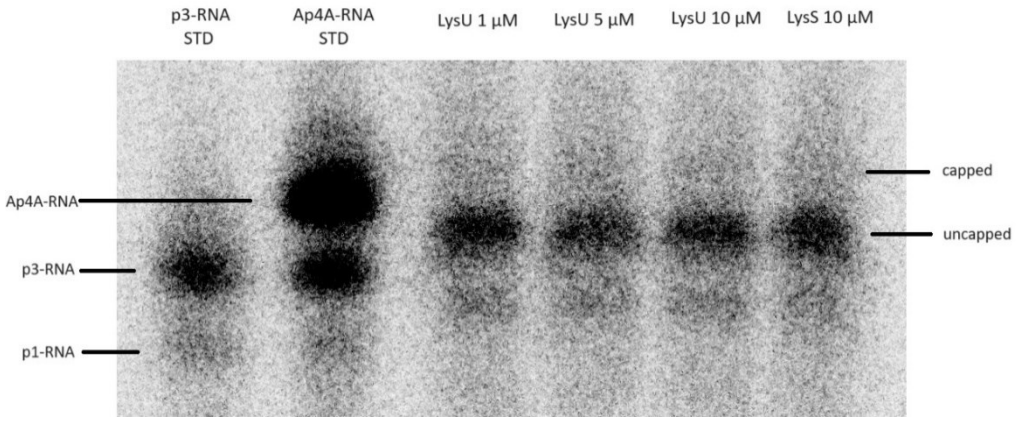


Figure 40.: 12 % PAGE gel. The analysis of RNA capping with increased concentration of LysU and LysS on 35mer RNA. Standards are shifted due to elution in water, instead of reaction buffer (different ionic strength makes RNA migrate differently).

Under any tested conditions, tRNA synthetases were not able to form the Ap_4A-RNA cap. To exclude the possibility that this reaction is sequence specific, we performed same experiment with different RNA. We used *yeiP* RNA, since it was shown by Luciano *et al.*[2] to be capped by LysU.

The IVT template from *yeiP* gene was prepared by PCR amplification of first 126 bp of the gene from *E. coli* (forward primer included T7 promotor; Figure 41.). The template could not be created from whole gene (573 bp) due to limitations in distinguishing capped and uncapped long RNAs on boronate gel. The radioactively labelled RNA (*yeiP*_126 RNA) produced by IVT was incubated with LysU and LysS and subsequently analysed on 8% boronate PAGE. Both tRNA synthetases, LysU and LysS, were not able to form the Ap_4A-RNA cap on *yeiP*_126 RNA (Figure 42. and 43.).

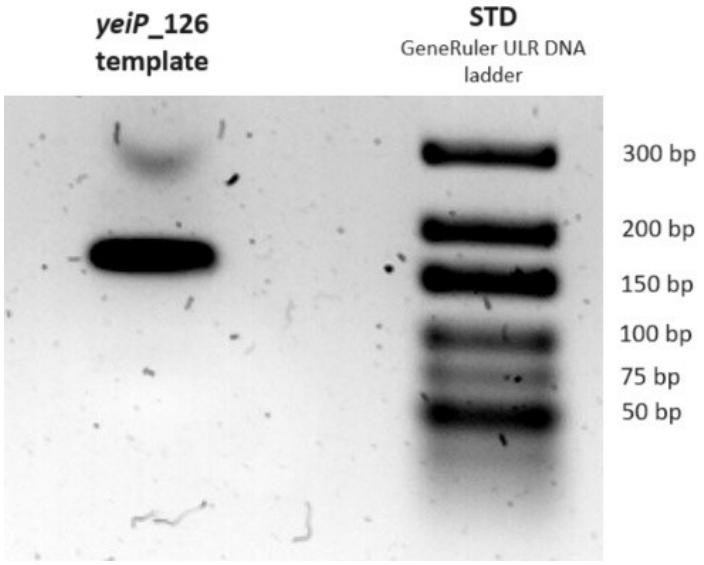


Figure 41.: 2 % agarose gel. PCR amplification of yeiP_126 IVT template. Length of PCR fragment is 158 bp (126 bp of yeiP + T7 promotor).

Different conditions, such as increased enzyme concentration or the absence of zinc or inorganic pyrophosphatase were tested as well, but they did not have any effect on RNA capping. The possibility, that LysU can form RNA cap and in the same it can also degrade it, was investigated. However, there was no difference in band intensity of Ap_4A-RNA incubated with LysU, indicating that LysU is not able to degrade the Ap_4A-RNA cap. The experiment of LysU and LysS RNA capping was performed in duplicate. These experiments again indicate that tRNA synthetases are not RNA capping enzymes. To further confirm this observation, we used more sensitive method based on LC-MS. Mass spectrometry-based methods are very accurate, and they can detect even very low amounts of capped RNA.

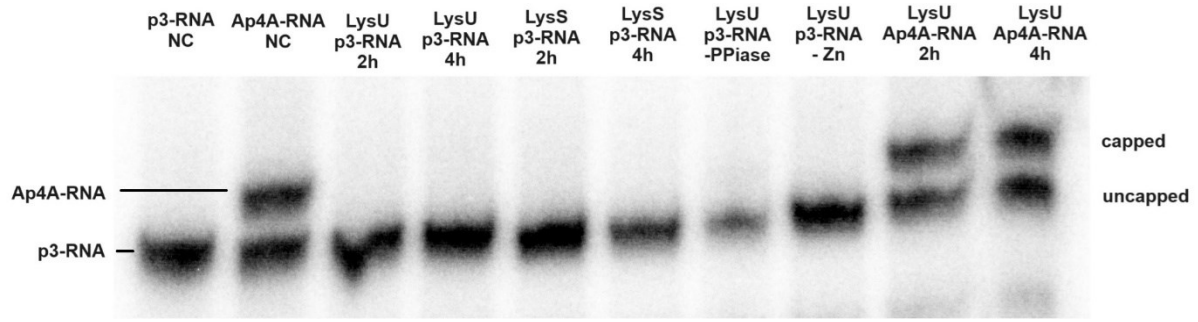


Figure 42.: 8 % boronate PAGE gel. Analysis of RNA capping by LysU and LysS on yeiP_126. In first line, there is an p_3RNA standard and in the second line, Ap_4A-RNA standard is present. In lines 3 to 6, there is a reaction of LysU and LysS respectively with p3RNA. In following lines, different conditions were tested – reaction without addition of pyrophosphatase and without addition of zinc. In last two lines, Ap_4A-RNA standard was incubated with LysU to address the possibility of Ap_4A-RNA degradation.

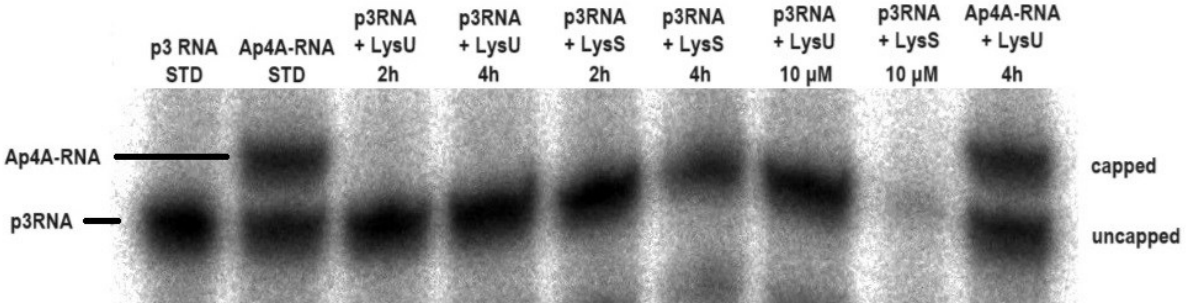


Figure 43.: 8 % boronate PAGE gel. Analysis of RNA capping by LysU and LysS on yeiP_126. Replicate of previous experiment. In lines 7 and 8, higher concentration of LysU and LysS enzymes were tested.

5. LC-MS analysis of RNA capping

To confirm the observation that tRNA synthetases are unable to form Ap_4A-RNA cap, more precise LC-MS analysis was employed. Non-radioactive RNA prepared from both templates, 35mer and *yeiP*_126, was incubated with LysU and subsequently digested by RNase T1 (Figure 44.). This enzyme cleaves RNA after every guanosine, generating RNA fragments that can be analysed on LC-MS[108]. The analysis focused on the Ap_4ApGp fragment (m/z = 1262.069 $[M-H]^-$) (capped RNA - both transcripts start with AG), on pppApGp (m/z = 933.017 $[M-H]^-$) and ApGp fragments (m/z = 693.118 $[M-H]^-$) (uncapped RNA), and Ap_4A (m/z = 837.056 $[M-H]^-$) (free diadenosine tetraphosphate created by LysU from free ATP).

The Ap_4ApGp fragment was not observed in any reaction of RNA with LysU, further indicating that tRNA synthetases are not RNA capping enzymes (Figures 45. and 46.). Moreover, Ap_4A was observed in samples with LysU, indicating that the enzyme is working. On the other hand, we have identified only small amount of pppApGp fragment in samples incubated with the enzyme. Instead, we have identified ApGp fragment with 5′-OH group, which is unusual on *in vitro* transcribed RNAs. Moreover, sequence AG is present only on 5′ end of 35mer RNA. In *yeiP*_126 RNA, this fragment could be also generated from the internal region of the RNA sequence. Experiment was performed in a duplicate.

The reason, why we have identified only small amount of pppApGp fragment, but identified ApGp instead is not known. We have proposed that it could be due to action of RNase T1, which hypothetically could cleave the 5′ terminal phosphates, or it could be due to the action of Zn^{2+} ions. The results and limitations of this experiment will be further discussed in the discussion chapter. Based on all performed experiments, it can be concluded that tRNA synthetases are unable to create Ap_4A-RNA cap and do not act as RNA capping enzymes.

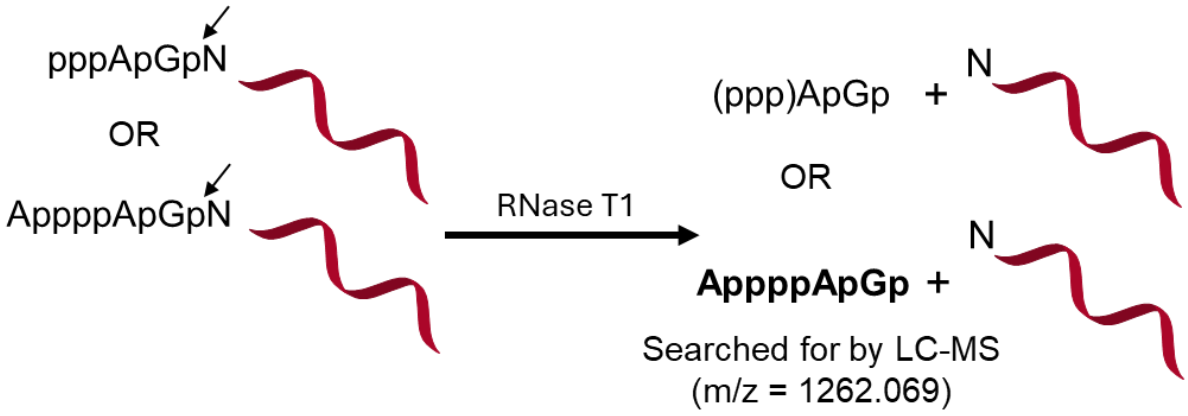


Figure 44.: Scheme of the LC-MS analysis of RNA capping. RNA was degraded by RNase T1, and fragments were analysed on LC-MS. We specifically search for Ap_4ApGp fragment, p_3ApGp fragment, ApGp fragment and Ap_4A.

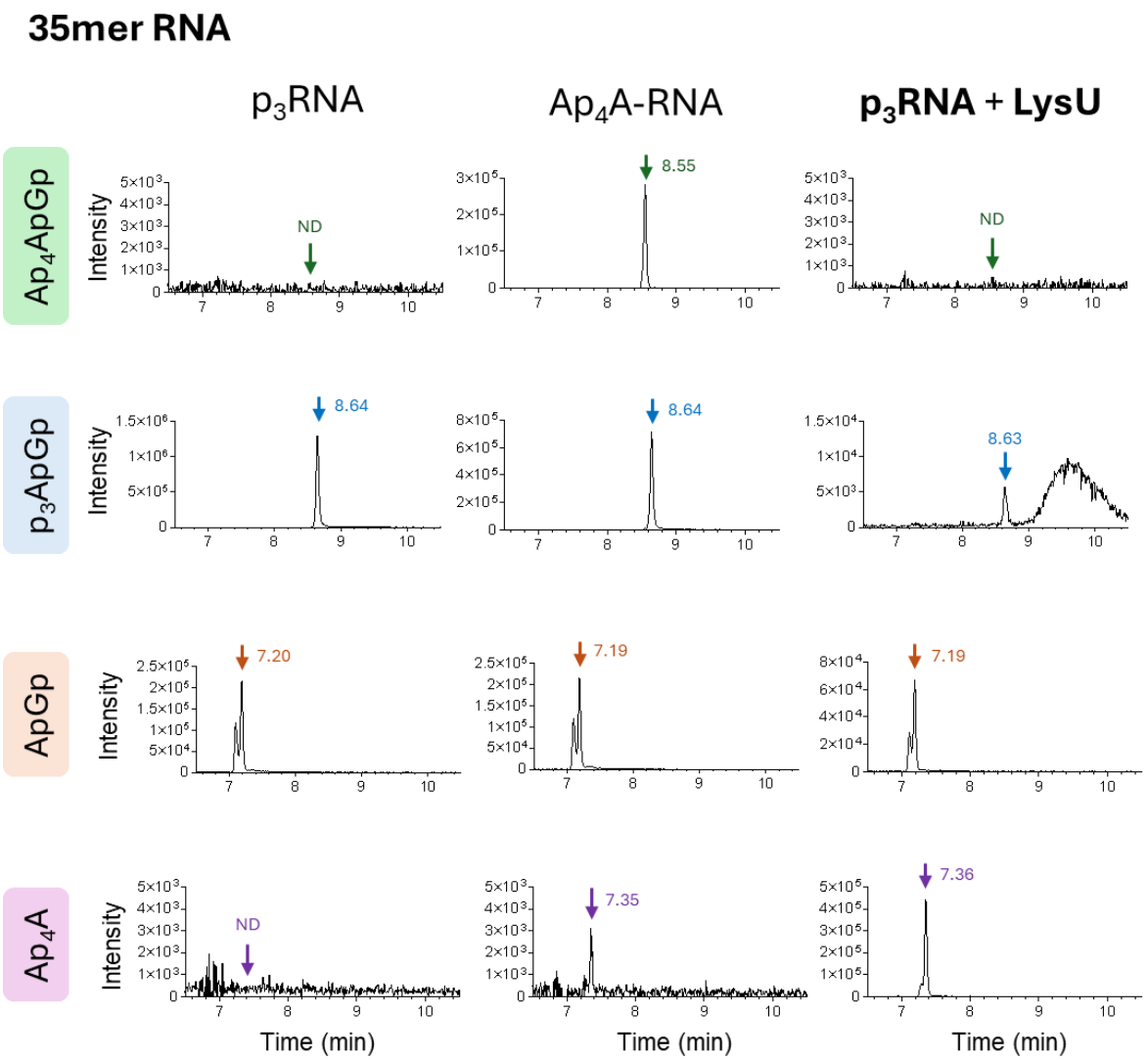


Figure 45.: Results from LC-MS analysis of 35mer RNA. Results represent LC-MS extracted ion chromatograms of Ap_4ApGp fragment (m/z = 1262.069), p_3ApGp fragment (m/z = 933.017), ApGp fragment (m/z = 693.118) and Ap_4A (m/z = 837.056). Arrows represent the elution time of a given fragment – arrows accompanied by elution time value means the fragments were present in the sample; arrow with ND (Not detected) represent point, where fragment should elute, but it was not detected. Total 3 samples were analysed, p_3RNA standard, Ap_4A-RNA standard and p_3RNA incubated with LysU.

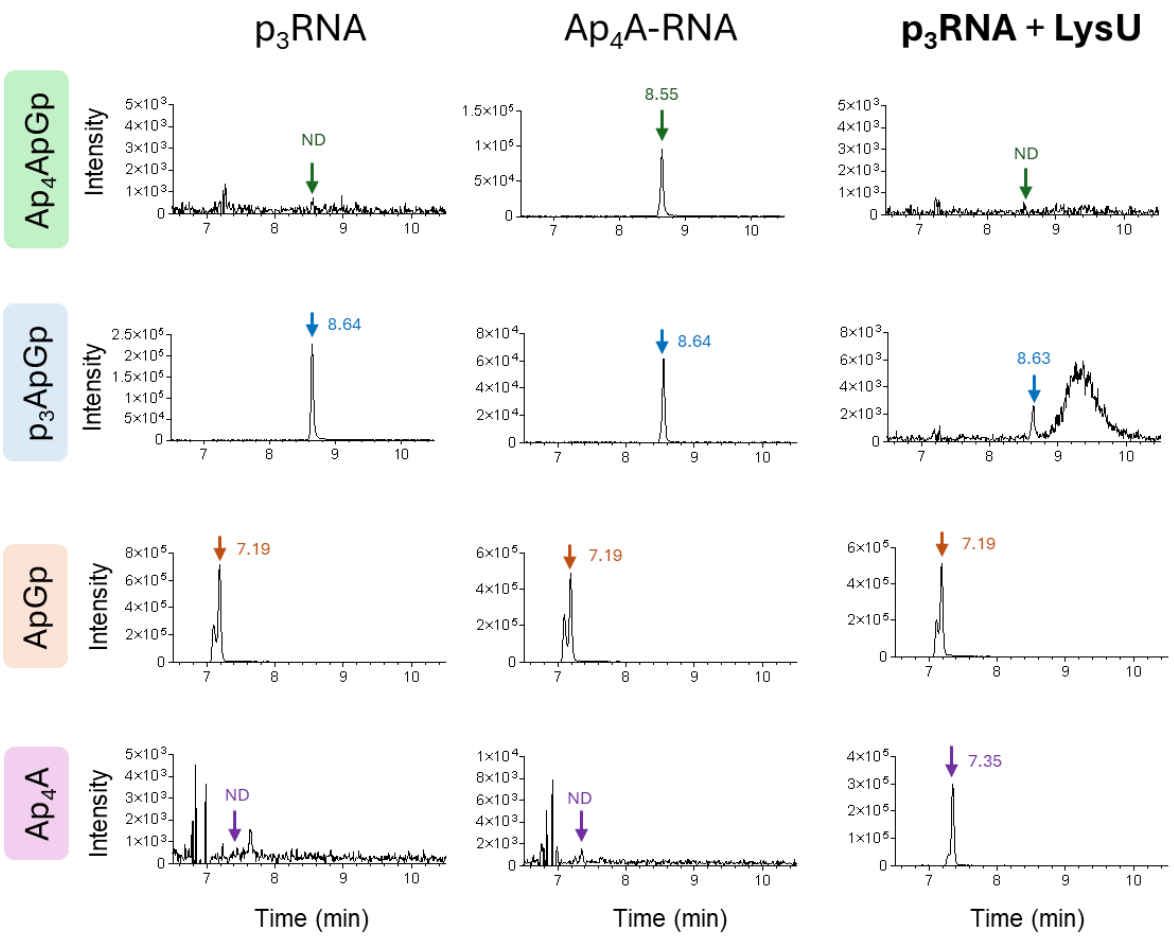


Figure 46.: Results from LC-MS analysis of yeiP_126 RNA. Results represent LC-MS extracted ion chromatograms of Ap_4ApGp fragment (m/z = 1262.069), p_3ApGp fragment (m/z = 933.017), ApGp fragment (m/z = 693.118) and Ap_4A (m/z = 837.056). Arrows represent the elution time of a given fragment – arrows accompanied by elution time value means the fragments were present in the sample; arrow with ND (Not detected) represent point, where fragment should elute, but it was not detected. Total 3 samples were analysed, p_3RNA standard, Ap_4A-RNA standard and p_3RNA incubated with LysU.

6. Molecular dynamics simulations of ion binding to LysU

To investigate possible structural reasons why tRNA synthetases are unable to accept RNA as a substrate, we investigated the structure of LysU with bound Ap_4A (PDB: 5YZX) in PyMOL software (The PyMOL Molecular Graphics System, Version 2.4.1 Schrödinger, LLC) with APBS Electrostatics plugin[109]. Our examination revealed a negatively charged area on the protein surface near the active site. This region is also the place, where RNA extending from 3′ OH group of Ap_4A would be positioned. RNA is a negatively charged molecule because of phosphates in its backbone. We propose a hypothesis that LysU cannot utilise 5′ triphosphate RNA as a substrate due to electrostatic repulsion from this negatively charged area (Figure 47.).

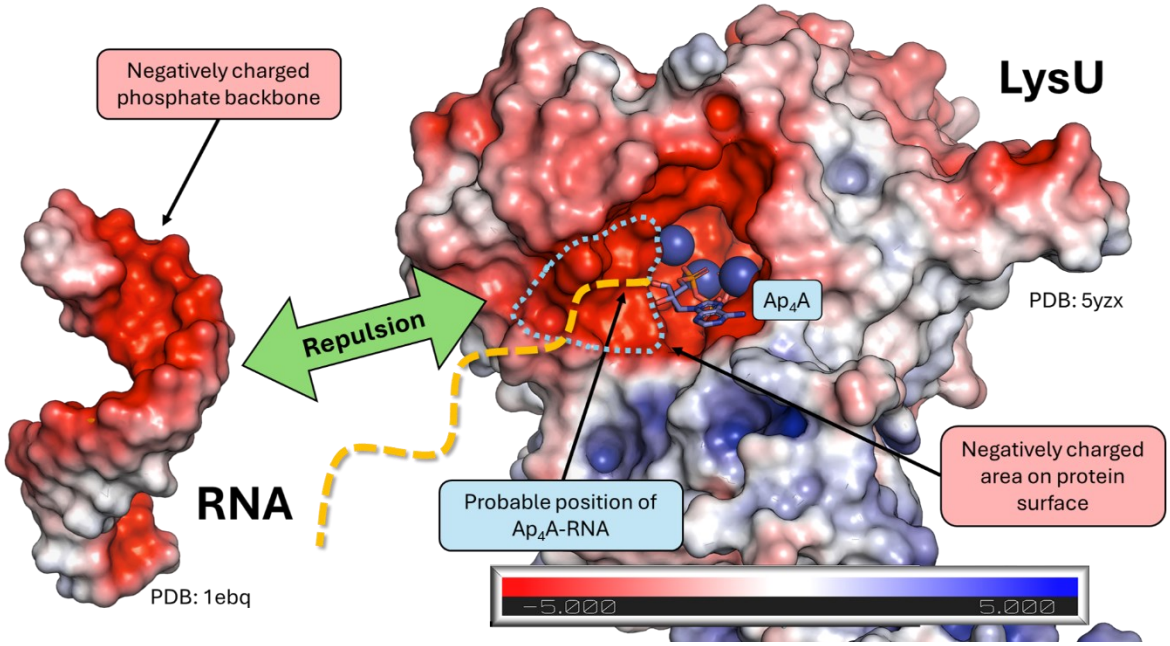


Figure 47.: Structural hypothesis for the inability of tRNA synthetases to create Ap_4A-RNA cap. The negatively charged area near the active site is also the probable place, where Ap_4A-RNA would bind (RNA must continue from 3' OH from this adenosine as the other is pointed inward into the protein). Because of the negatively charged phosphate backbone of RNA, the RNA cannot bind to LysU during Ap_4A synthesis. Based on structure 5YZX and visualised using PyMOL.

The structure of lysyl-tRNA synthetase in complex with tRNA has yet not been solved, however, we can explore the position of tRNA binding based on closely related aspartyl-tRNA synthetase (PDB: 1C0A)[110]. According to this structure, we can assume that tRNA binds to the active site from the opposite side to where Ap_4A-RNA would probably be located. Therefore, tRNA does not interact with this negatively charged region. Moreover, interaction with tRNA is far more complex and mediated through other interactions as well, which suggests that negatively charged surface near the active site should not inhibit the tRNA binding.

In the recent study by Giacobelli *et al*.[111], it was demonstrated that the interaction between negatively charged regions on the protein surface and RNA is feasible in the presence of positively charged ions. They found that Mg^{2+} or K^{+} ions could mediate the interaction between the negative protein surface and RNA. To address this issue, molecular dynamics (MD) simulations of ion binding to the protein surface were performed with cooperation from Dr. Martin Lepšík (MD simulation was performed by Dr. Lepšík, preparation of PDB structures for MD simulation with CHARMM-GUI and the visualisation in VMD was performed by me with assistance of Dr. Lepšík).

Two 40 ns MD simulations were conducted using GROMACS with the protein and ligands restrained. Structure 1E22 in system containing 0.01 M $MgCl_2$ was used for MD simulations of magnesium binding. Visualisation in VMD using volmap tool revealed the occupancy of Mg^{2+} molecules in the system – areas with high Mg^{2+} occupancies show the probable Mg^{2+} binding sites. The VMD volmap tool creates a volumetric map of atom occupancy, a 3D grid, where each point is set to 0 or 1, depending on whether it contains atoms or not. When averaged over all simulation frames, it will provide the fractional occupancy of each grid point.

By adjusting the volmap occupancy isovalue (value from 0 to 1 which sets the visible fractional occupancy), we were able to identify areas with varying probabilities of Mg^{2+} binding. At an isovalue set to 0.15, several biding sites were identified, but no ions were observed binding to the area of interest (the negatively charged region). At an isovalue set to 0.03, it was observed that Mg^{2+} ions could bind to the region of interest. This means that Mg^{2+} can bind to the region of interest, but only with low affinity (Figure 48. A). Throughout the simulation, Mg^{2+} ions were consistently present in the active site around ACP ligand (phosphoaminophosphonic acid-adenylate ester – non-hydrolysable ATP analogue) near the crystallographic positions (Figure 48. B).

MD simulation with structure 5YZX (with bound Ap_4A) was performed in presence of 0.15 M NaCl using GROMACS. Like the simulation with Mg^{2+} ions, the volmap plugin revealed the occupancies of Na^{+} ions in the system and probable ion binding sites on the protein surface. With an isovalue of 0.15, we searched for biding sites with high affinity for Na^{+} ions, while with an isovalue of 0.03, we searched for all possible binding sites. Our findings indicate that Na^{+} ions bind to the site of interest only with a low probability (Figure 49. A). Throughout the simulation, six Na^{+} ions were present in the active site around Ap_4A ligand (Figure 49. B).

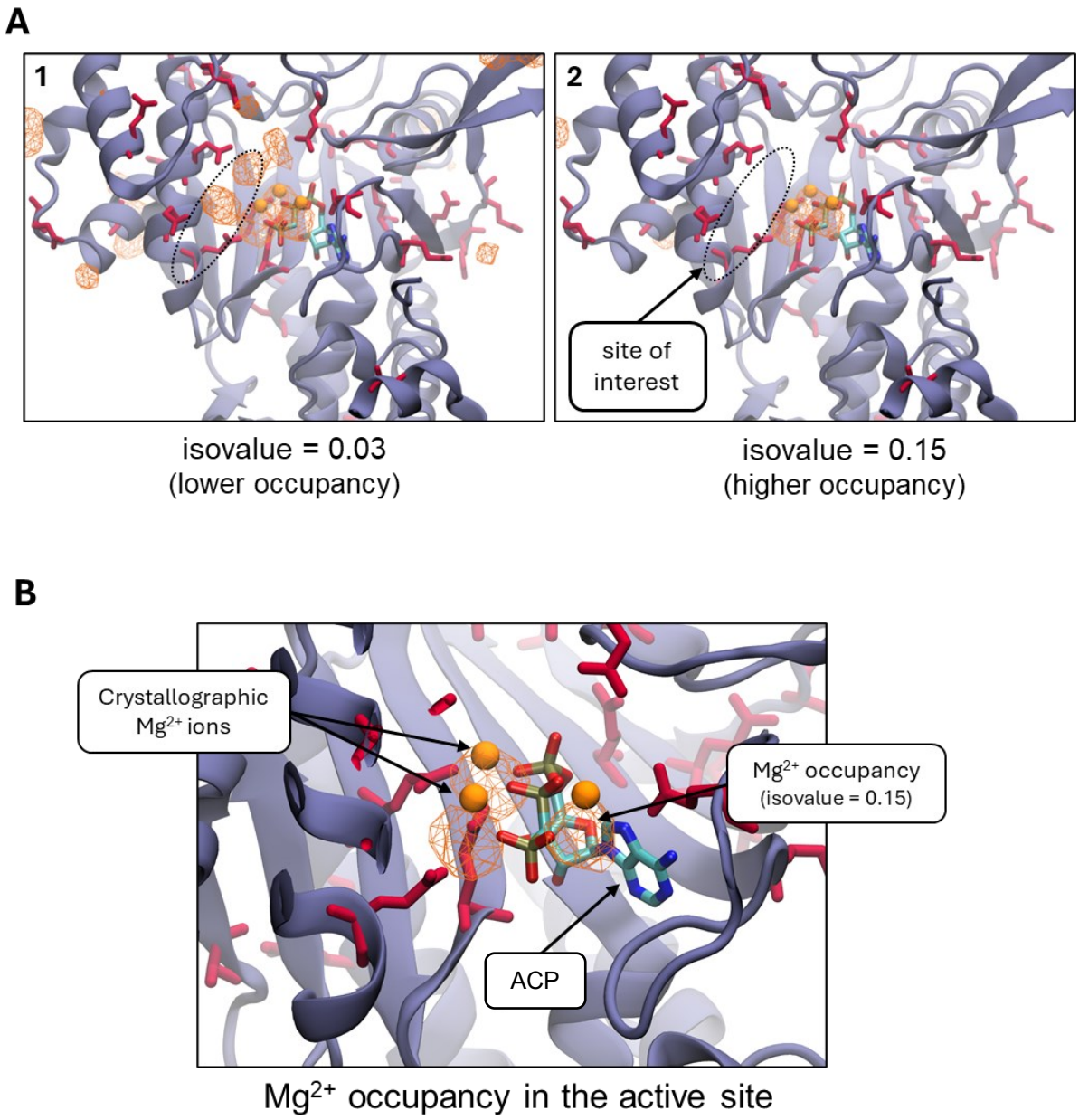


*Figure 48.: MD simulation of Mg^{2+} binding to LysU (1E22). GLU residues are highlighted in red, occupancy areas are visualised in orange **A** – occupancy of Mg^{2+} ions with isovalues 0.03 (1) and 0.15 (2). **B** – Mg^{2+} occupancy in the active site. Occupancies from MD correspond very well to the crystallographic ones.*

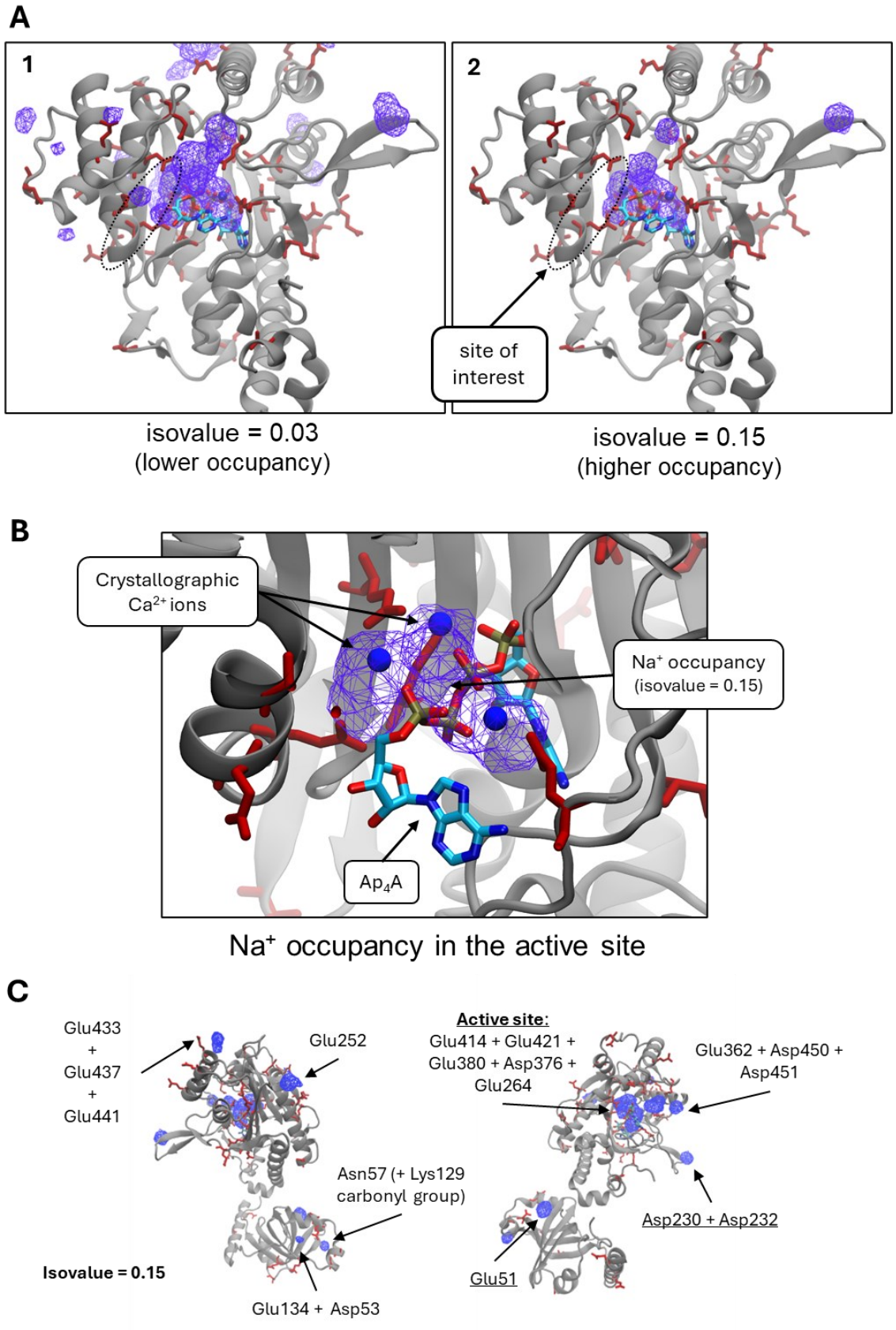


Figure 49.: MD simulation of Na^+ ions biding to LysU (5YZX). GLU residues are highlighted in red, occupancy areas are visualised in blue. ***A*** *– occupancy of Na^+ ions LysU with isovalues 0.03 (1) and 0.15 (2).* ***B*** *– ions occupancy in the active side around Ap_4A,* ***C*** *– most probable ion binding sites on the surface of LysU. The <u>underlined resides</u> were visible with an isovalue of 0.9.*

7. Synthesis and purification of diadenosine-$^{15}N_{10}$ tetraphosphate

The inability of tRNA synthetases to create Ap_4A-RNA caps lead us to the decision to use LysU for synthesis of commercially unavailable diadenosine-$^{15}N_{10}$ tetraphosphate. This compound is a useful standard for measurement of Ap_4A concentration on LC-MS in isolated RNA and availability of this standard, will allow for the quantification of Ap_4A-RNA in precious samples. Diadenosine-$^{15}N_{10}$ tetraphosphate was successfully synthetized by LysU from adenosine-$^{15}N_5$ triphosphate and subsequently purified using HPLC and solid phase extraction (SPE). The presence of $^{15}N_{10}$-Ap_4A and the purity of the sample was measured by ^{1}H NMR and ^{31}P NMR. We observed slight contamination by acetonitrile and small amount of triethylammonium, acetate and butanol (Figure 51.). These volatile compounds were removed by lyophilisation. The presence of other phosphate containing molecules was measure by ^{31}P NMR, and the results showed that the sample was composed only from Ap_4A (Figure 52.). The total yield of $^{15}N_{10}$-Ap_4A was 0.180 mg, which is 47.9 % of the theoretical yield (0.376 mg).

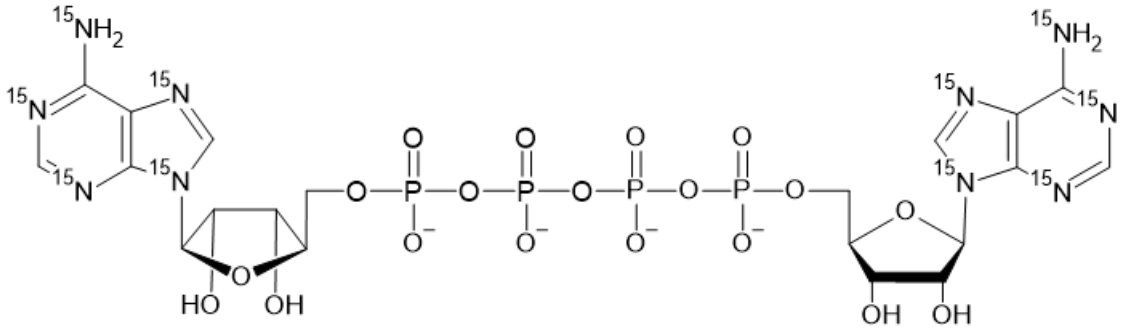


Figure 50.: Structure of diadenosine-$^{15}N_{10}$ tetraphosphate.

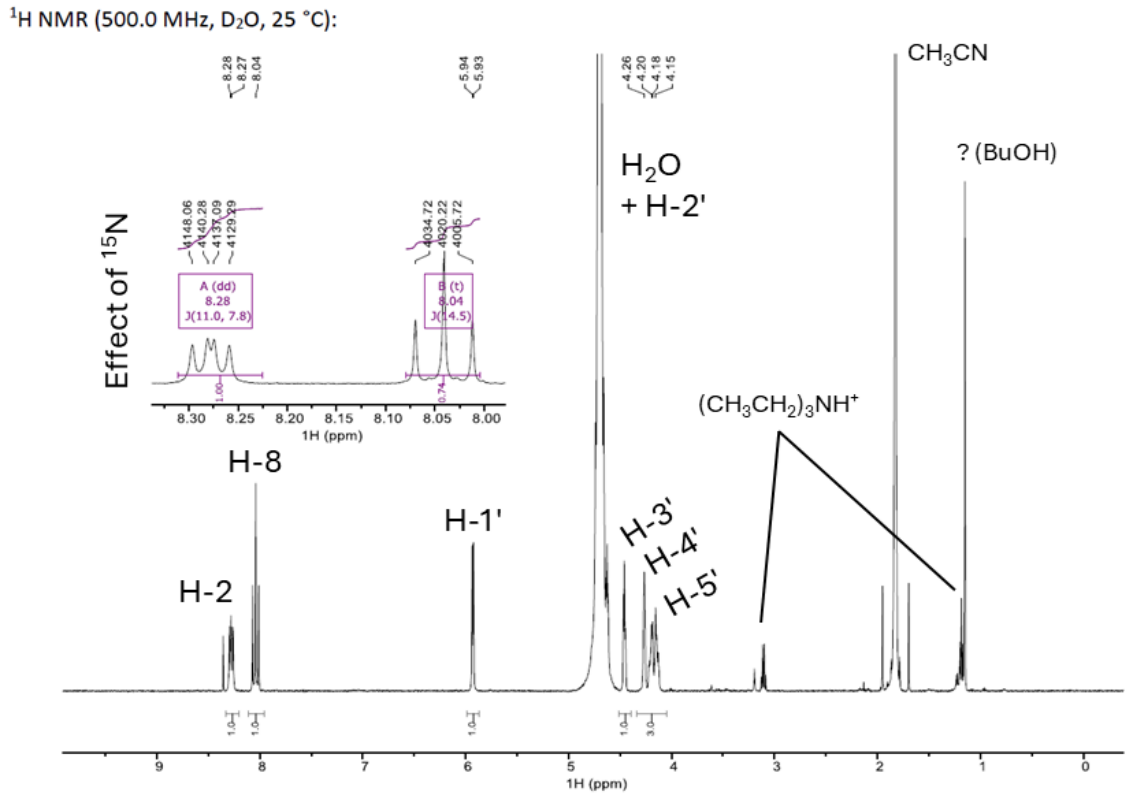


Figure 51.: ^{1}H NMR spectra of purified sample. H-2 and H-8 are hydrogens located on the purine ring of adenosine. These hydrogens are affected by the presence of ^{15}N, therefore their peaks are split into quartet and triplet (in normal Ap_4A, both hydrogen forms singlets).

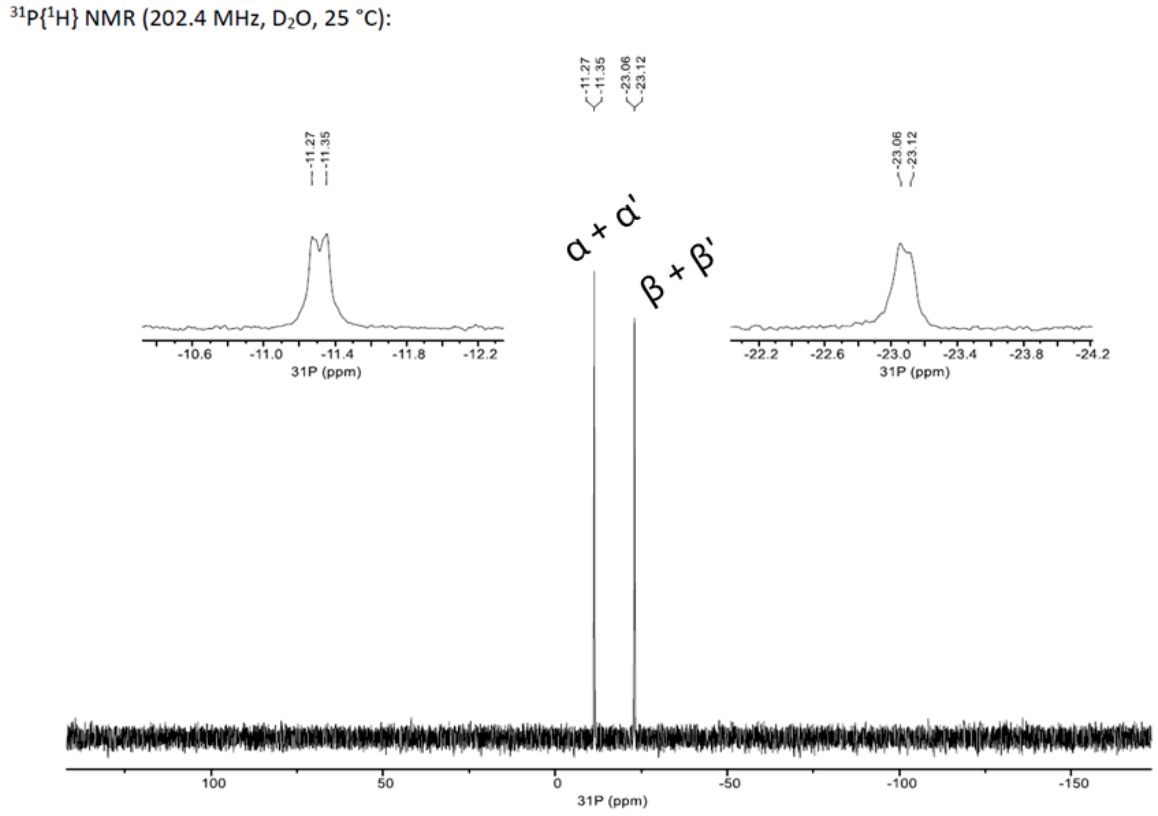


Figure 52.: ^{31}P NMR spectra of purified sample.

Discussion

1. Selection of tRNA synthetases

The selection of tRNA synthetases were based upon their ability to synthetize Ap_4A[50]. Lysyl-tRNA synthetase LysU was an obvious choice, as it is the best documented tRNA synthetases in its ability to synthesize Ap_4A. Moreover, it is the only enzyme, that have been reported for the ability to form Ap_4A-RNA cap. Therefore, LysU was the enzyme with the highest probability being an RNA capping enzyme in *E. coli*. Other tRNA synthetase, we have selected, lysyl-tRNA synthetase LysS is a homolog of LysU and its ability to synthetise Ap_4A is also well documented. It has been reported that LysS synthetize Ap_4A 2 times slower than LysU[95], therefore a comparison of reaction rates on Ap_4A and Ap_4A-RNA between LysU and LysS could provide more insights about the reaction specificity.

There were several other tRNA synthetases reported to synthetise Ap_4A in literature[50]. We selected histidyl-tRNA synthetase (HisS) and methionyl-tRNA synthetase (MetG). HisS has been reported to be one of the fastest enzymes in Ap_4A synthesis without zinc addition[36] (Table 2.). MetG was selected as a representant of class I tRNA synthetases (all other selected tRNA synthetases are class II enzymes) and for its documented synthesis of Ap_4A *in vivo*[96]. We also selected AlaS, as an enzyme that is very efficient in synthesis of Ap_4A (together with LysRS and PheRS is stimulated by zinc). We did not select PheRS, because it forms $\alpha_2\beta_2$ tetramers encoded by two different genes. Unfortunately, we were not able to transform AlaS into *E. coli*, therefore it was not used in following experiments.

2. Plasmid preparation

Genes for selected tRNA synthetases were PCR amplified, cleaved with restriction enzymes, and ligated to plasmid pET-28a. This plasmid is suitable for protein expression. It contains His-tag on N-terminus of protein construct; its expression is regulated by *lac* operator and plasmid contains gene for kanamycin resistance. The prepared plasmids were transformed into *E. coli* BL21 strain, which is an appropriate strain for protein expression. Transformed plasmids in BL21 were checked by Sanger sequencing. The obtained sequences were translated into protein sequences and checked by BLAST. Therefore, we have confirmed that all prepared plasmid contains

naturally occurring protein variants. With a successful plasmid transformation, we could proceed to protein expression and purification.

3. Protein purification

Selected tRNA synthetases were expressed in *E. coli* BL21 cells. For all tRNA synthetases, optimal conditions for their expression were 3 hours of incubation in LB media, after the addition of IPTG in final concentration 1 mM. Proteins were purified using fast protein liquid chromatography (FPLC), composed of two steps of purification, immobilized metal ion affinity chromatography using Ni^{2+} ions (HisTrap) and size exclusion chromatography. The results of both purification steps were analysed on SDS-PAGE.

Purification of all tRNA synthetases was successful, with an appropriate purity and yields for purposes of this thesis. The protein purification process proceeded smoothly without encountering any issues, which is partly because tRNA synthetases are soluble cytoplasmatic proteins naturally occurring in *E. coli*, and also because of the expertise of Dr. Ondřej Nešuta, who supervised me during the protein purification. With a successful purification of all selected tRNA synthetases, we can move to the next step, which is synthesis of free dinucleoside polyphosphate by tRNA synthetases.

4. Synthesis of free dinucleoside polyphosphates

The reaction conditions for synthesis of dinucleoside polyphosphates were based on the conditions described in Luciano *et al.*[2]. In an initial experiment, optimal reaction conditions of LysU were tested (Figure 29.). The reaction products were analysed on HPLC by method described in Mititelu *et al.*[52]. Highest concentration of Ap_4A was measured in 2 hours of reaction. In longer reaction times, Ap_4A was slowly hydrolysed to ADP and AMP (Figure 30.). Subsequently, from ADP reacting with aa-AMP intermediate, Ap_3A was formed[100]. In conclusion, best condition for Ap_4A synthesis were 2 hours incubation of 1 μM LysU.

Afterwards, all purified tRNA synthetases were compared in their ability to synthetize Ap_4A. After two hours of reaction, Ap_4A concentration in reaction with LysS was approx. 2 times lower than in reaction with LysU. In reaction with HisS, concentration was approx. 100 times lower, and in reaction with MetG, concentration was approx. 2000 times lower than in reaction with LysU. These results are in accordance with literature (Table 2.). MetG and HisS are very slow in synthesis of Ap_4A, which indicates that

probability for creating Ap_4A-RNA is relatively low. On the other hand, both LysRS are efficient in Ap_4A synthesis, and they are good candidates for an RNA capping enzyme.

LysU was tested for ability to create different Ap_nNs than Ap_4A[99]. Ap_3A was obtained from reaction, where ATP and ADP were in 1:1 mixture (Figure 32.), or by prolonged reaction with ATP only (as already described above, it is due to hydrolysis of Ap_4A to ADP, and then ADP reacting with aa-AMP intermediate). Ap_4G was obtained from 1:1 mixture of ATP and GTP (Figure 33.), Ap_4C from 1:1 mixture of ATP and CTP (Figure 34.) and Ap_4U from 1:1 mixture of ATP and UTP (figure 35.). We observed that the reaction with different nucleotide than ATP was always slower (Figure 37.). After 2 hours of incubation, approx. 80 to 90 % of dinucleoside polyphosphates in the reaction mixture was Ap_4A. The slower rate of the reaction with other nucleotides than ATP is in accordance with literature, but in our observations, the reaction is even slower than what was described in literature[45]. tRNA synthetases are unable to synthesize non-adenylated dinucleoside polyphosphates because they require ATP in first step of the reaction (Figure 11.). That means, that they are unable to synthetize Gp_4G or Gp_3G. This was experimentally verified by incubation of LysU with GTP, where product formation was not observed (Figure 36.).

In summary, it was shown that best conditions for Ap_4A synthesis by LysU is two hours of incubation with 1 μM enzyme concentration, LysU and LysS being much more efficient in Ap_4A synthesis than HisS and MetG, and that LysU is able to synthetize other Ap_nNs that Ap_4A. Once we found optimal conditions under which tRNA synthetases form Ap_nN, we proceed to testing their ability to form Ap_4A-RNA cap.

5. RNA capping by tRNA synthetases

5.1. PAGE analysis of RNA capping

tRNA synthetases were incubated with triphosphorylated RNA in the same reaction conditions as described above. The RNA for this experiment was prepared by *in vitro* transcription from a 35 nucleotide long artificial template (35mer)[1], supplemented with a radioactive [α-^{32}P]GTP, which enabled subsequent visualisation on the gel. All tRNA synthetases were incubated for 2 hours and 4 hours.

Unfortunately, formation of Ap_4A-RNA cap was not observed by any of the selected tRNA synthetases (Figures 38. and 39.). The experiment was performed in duplicate with all purified tRNA synthetases, the results were compared with the IVT prepared

standards of Ap_4A-RNA and p_3RNA and with negative controls as well. We also tested higher concentrations of enzymes, which led to significant increase in rate of free Ap_4A synthesis in previous experiments, but we did not observe any positive effect on Ap_4A-RNA synthesis (Figure 40.). In parallel, we tested the reaction with the same conditions without addition of radioactive RNA for presence of free Ap_4A. Indeed, free Ap_4A was present in these samples indicating that tRNA synthetases performed as expected, which excludes any mistakes during reaction preparation. This suggests that tRNA synthetases are unable to utilize RNA as a substrate for the second step of the reaction.

One possible explanation, why tRNA synthetases were unable to form Ap_4A-RNA cap, could be a substrate specificity towards RNA. The 35mer is an artificial template, which do not occur naturally in the cell. Therefore, we decided to test the different RNA. We chose a *yeiP* transcript, which is RNA used by Luciano *at al.*[2]. An IVT template was created from first 126 nt of the transcript, because of the inability to distinguish capped and uncapped transcripts in longer RNAs on boronate-PAGE.

The *yeiP*_126 RNA was incubated with LysU and LysS enzymes and analysed on 8 % boronate-PAGE[65]. Again, we did not observe formation of an Ap_4A-RNA (Figure 41. and 42.). This observation again indicates, that tRNA synthetases are unable to form Ap_4A-RNA cap. We also tested the possibility, that Zn ions or inorganic pyrophosphatase could influence the RNA capping. When both compounds were excluded from the reaction, we did not observe any effect on RNA capping. Furthermore, we increased enzyme concentrations without any effect on the RNA capping. Additionally, we explored the possibility, that tRNA synthetases can degrade Ap_4A-RNA. LysU was incubated with Ap_4A-RNA standard but change in band intensity of Ap_4A-RNA was not observed. This indicates that tRNA synthetases do not cleave Ap_4A-RNA.

The results of these experiments suggest that tRNA synthetases are unable to synthetize the Ap_4A-RNA cap and do not act as RNA capping enzymes. Experiments were performed with two different RNA substrates in duplicates and different reaction conditions were employed with any effect on Ap_4A-RNA cap creation. To verify that tRNA synthetases are not RNA capping enzymes, we decided to use LC-MS based techniques to analyse RNA capping.

5.2. LC-MS analysis of RNA capping

Methods based on liquid chromatography coupled with mass spectrometry are very precise and enable us identification of small amount of capped RNA. Both previously used transcripts, 35mer and *yeiP*_126, were incubated with LysU for 2 hours and then cleaved by RNase T1. This enzyme cleaves RNA specifically after guanosine, which creates RNA fragments that can be analysed by LC-MS[108]. RNase digestion enabled us to distinguish between Ap_4A-RNA and free Ap_4A, created from reaction with ATP. In LC-MS analysis, we searched for Ap_4ApGp fragment (capped RNA), p_3ApGp and ApGp fragments (uncapped RNA) and Ap_4A. In the analysis, we included standards of p_3RNA and Ap_4A-RNA of both transcripts.

The LC-MS analysis revealed that Ap_4ApGp fragment was not detectable in the samples, where p_3RNA was incubated with LysU (Figures 45. and 46.). This confirms our hypothesis that tRNA synthetases are not RNA capping enzymes. Moreover, we identified huge amounts of Ap_4A, even after one purification step. This confirms that the enzyme was working properly and suggests that ATP is the preferred substrate.

Surprisingly, the p_3ApGp fragment was identified in the samples incubated with LysU in significantly lower amount compared to digested p_3RNA standard. Moreover, we identified ApGp peak, which should not be present in *in vitro* transcribed 5′ end of RNA. In case of 35mer, this fragment could originate only from the 5′ end of RNA (in *yeiP*_126 RNA sample, this fragment originated from the inside of the RNA sequence as well as from the 5′ end). We also searched for different 5′ ends of RNA (ppApGp and pApGp fragments) or for different 5′ RNA end modifications (Ap_3ApGp and lys-p_3ApGp fragments), which could be potentially formed by tRNA synthetases, but we did not detect any of them.

Since the ApGp fragments were also identified in the standards (peak intensity was the same across samples from the same transcript), we cannot exclude a possibility that the loss of triphosphate from 5′ end occurred during the RNase T1 digestion. In general, we observed decreased amount of RNA in the samples incubated with the LysU enzyme, because they underwent one more cleaning step. This implies, that lower amount of RNA, but same amount of detected ApGp fragment in samples incubated with LysU as in p_3RNA and Ap_4A-RNA standards, results in low amount of p_3ApGp fragment in the samples incubated with LysU (in other words, the reaction rate

of triphosphate RNA end degradation by RNase T1 is same across all samples – samples with lower amount of p_3RNA will result in lower amount of detectable p_3ApGp fragment). On the other hand, cleavage of triphosphate end of RNA by RNase T1 was not observed in literature, and reaction mechanism of RNase T1 suggests, that 2′ OH group from ribose is needed for RNA restriction[112].

Other suggestion, why p_3ApGp was not identified, is that LysU is modifying the triphosphate 5′ end of RNA in an unknown way (search for Ap$_3$ApGp fragment and lys-p_3ApGp was unsuccessful). Another possibility is that zinc ions create complexes with triphosphate RNA, making it difficult to identify fragment on LC-MC. However, only small amounts of Zn ions were used in the reaction. Therefore, digestion of 5′ RNA triphosphate by RNase T1, despite the fact that it was not reported in the literature, seems as the most probable explanation, why we detected low amounts of p_3ApGp in the samples incubated with LysU.

In conclusion, PAGE and LC-MS analysis were used to study the potential of tRNA synthetases to form Ap$_4$A-RNA. We did not detect Ap$_4$A-RNA formation by any of these methods; therefore, we conclude that tRNA synthetases are unable to form Ap$_4$A-RNA and do not act as an RNA capping enzymes. These results are inconsistent with published literature, which will be discussed in following section (discussion section 5.4).

5.3. Possibility to form Ap$_4$A-RNA cap on tRNAs

tRNA synthetases are well adapted for cognate tRNA binding. When tRNA binds to the tRNA synthetase, it is positioned in a special way, that 3′ end is located in the active site. Amino acid is then transferred to the OH group at on 3′ or 2′ carbon of terminal adenosine. The aminoacylation reaction is an essential reaction for all living organism and appears to be preferred over Ap$_4$A synthesis[36,39].

However, we hypothesized that Ap$_4$A-RNA cap could, for example, stabilize tRNAs during stress conditions. As described in the introduction chapter 4.3., all tRNAs are processed in the similar way by RNase P, which leaves monophosphate at 5′ end. tRNA synthetases cannot use monophosphorylated nucleotide as a substrate for Ap$_n$N synthesis (unable to take AMP and create Ap$_2$A[45]), which suggest that tRNA synthetases would not be able to create Ap$_4$A-RNA cap on tRNAs. Moreover, tRNA is

positioned in a way that only 3′ end is in the active site and 5′ is left outside the active site, further indicating that tRNA synthetases cannot from RNA cap on 5′ end.

However, pre-tRNAs could theoretically still be capped by tRNA synthetases. On the contrary, pre-tRNAs are rapidly processed in cells and do not interact with tRNA synthetases[87], indicating that pre-tRNAs should not be post-transcriptionally capped by tRNA synthetases. Additionally, as demonstrated with monocistronic tRNAs, 5′ triphosphate prevents proper tRNA processing[89]. An Ap_4A-RNA cap is an even larger molecular structure than a triphosphate, which suggests that same problems could occur with Ap_4A capped pre-tRNA transcripts. These theoretical assumptions suggest that 5′ Ap_4A-RNA cap formation on tRNAs by tRNA synthetases is not possible, and therefore we directed our experimental focus elsewhere.

5.4. Comparison of results with literature

We found that tRNA synthetases are unable to create Ap_4A-RNA cap *in vitro*, and therefore do not act as RNA capping enzymes. This result is inconsistent with the findings from Luciano *et al.*[2], where LysU was shown to be capable of creating a dinucleoside polyphosphate cap on triphosphorylated *yeiP* RNA.

The experiment conducted by Luciano *et al.*[2] was an initial test to explore the possibility of creating an Ap_4A-RNA cap by tRNA synthetases. Their experiment involved only one tRNA synthetase, tested only one condition and crucially, it was conducted only once without a replicate (it is not indicated anywhere in their publication that the experiment was repeated). Our study involved four different tRNA synthetases, all experiments were performed in duplicates, and we utilized two different RNA substrates. Furthermore, to confirm our results, we utilized LC-MS approach to detect RNA capping by tRNA synthetases. Therefore, we are convinced that our data are robust in comparison to the initial test conducted by Luciano *et al.*[2].

Additionally, in the following publication by Luciano *and* Belasco[6], the authors analysed the effect of promoter changes on dinucleoside polyphosphate cap synthesis via the NCIN mechanism. Their results pointed out that the sequence of the promoter is an important factor in regulation of Ap_4N-RNA cap synthesis. They further discussed that effect of tRNA synthetases on the abundance of Ap_4N capped transcripts should be relatively minor compared to effects of promotor changes, but they did not disprove the possibility that post-transcriptional mechanism exists in cells.

In a recent publication from our group, the ability of human lysyl-tRNA synthetase (KARS) to create Ap_4A-RNA was also tested *in vitro*[38]. Human lysyl-tRNA synthetase was unable to create Ap_4A-RNA caps *in vitro*. KARS was initially tested for production of Ap_4A, which was successful, but following analysis of reaction with 35mer RNA showed that KARS was not able to cap triphosphorylated RNA. The reaction conditions in this experiment were similar to the conditions used in this thesis, only the enzyme concentration was significantly higher. The results from the experiments with KARS further support conclusion of this thesis that tRNA synthetases do not act as RNA capping enzymes.

Taking into consideration results from this thesis and available literature sources, we can exclude the possibility that tRNA synthetases act as RNA capping enzymes. Apart from one initial experiment reported by others, tRNA synthetases were not able to create Ap_4A-RNA cap *in vitro* under any tested conditions. Therefore, it is highly probable that the incorporation of Np_nN during transcription as NCIN is the primary mechanism of dinucleoside polyphosphate cap synthesis. This process might be regulated by the sequence of promoter, but further experiments are necessary to elucidate the incorporation process and its regulation.

The only remaining problem is the synthesis of Gp_4G-RNA cap, which is probably not synthesized by incorporation during transcription, as Gp_4G is not present in free form in *E. coli* cells[53]. Even if tRNA synthetases would be capable of creating RNA cap, it would not solve this problem, as they are able to create only adenosine-containing RNA caps (tRNA synthetases require ATP molecule in the first step of reaction). There are several other candidates for post-transcriptional capping enzyme, which will be discussed later in the discussion chapter 7.. As the next step in our research, we would like to understand, why tRNA synthetases are unable to create Ap_4A-RNA cap.

6. Molecular dynamics simulation of ion binding to LysU

We wanted to explore why tRNA synthetases can take ATP as a substrate to create Ap_4A, but they are unable to perform the same reaction with 5′ triphosphate RNA. To understand this experimental observation, we examined the crystal structure of LysU tRNA synthetase. We studied LysU crystal structure with bound Ap_4A (PDB code 5YZX) with APBS Electrostatics plugin[109] and we found that there is a negatively charged area on the protein surface near the 3′ OH group of second adenosine from

Ap_4A (Figure 47.). This could mean that RNA molecule continuing from the position of second adenosine of Ap_4A would be in contact with this area. Since RNA is highly negatively charged molecule (phosphate backbone), it would imply that RNA and the protein might be electrostatically repelled. This could indicate that LysU cannot take triphosphorylated RNA at all, or it is a very unpreferable substrate compared to ATP.

Here, we present a hypothesis, that LysU is unable to synthesize Ap_4A-RNA, because p_3RNA substrate is repelled by the negatively charged area on the protein near the active site of LysU. This negatively charged area is composed of two glutamic acids, threonine and several carbonyl groups from the peptide bonds. However, in recent publication from Giacobelli *et al*.[111], it was shown that interaction between negatively charged areas on protein surface could promote RNA binding through interaction mediated by metal ions (Mg^{2+} or K^+ ions). We used here an approach based on molecular dynamics modelling to assess the potential binding of Mg^{2+} ions to this negatively charged area. The lack of Mg^{2+} binding sites ions observed in MD could support our hypothesis, that RNA is repulsed by negatively charged surface near the active site.

There are several limitations for simulations of Mg^{2+} binding to the protein surface[113]. Mg^{2+} ions are small and carry a positive +2 charge. Their high charge density results in their strong hydration. This leads to slow binding of magnesium ions to potential binding sites on the protein (ion stays in its hydration shell instead of binding to the protein surface). On the other hand, once Mg^{2+} ions bind to the protein surface, they remain there for a long time, which may create false positive Mg^{2+} binding sites. To address this issue, we also performed MD simulation with Na^+ ions, which carry +1 positive charge (weaker interactions with water molecules) and have a similar mass as magnesium ions.

The molecular dynamics simulation showed us that metal ions Mg^{2+} and Na^+ may bind to the site of interest with only a low probability. This suggests that the interaction between negatively charged protein surface and the RNA is unlikely to be mediated through positively charged ions, and RNA might thus be repelled by the protein surface.

7. Possible RNA capping enzymes (other than tRNA synthetases)

tRNA synthetases were considered to be the best candidates for the RNA capping enzymes. This hypothesis was disproved by results presented in this thesis. However, there is still one possible tRNA synthetase, that could be working as an RNA capping enzyme. That is eukaryotic glycyl-tRNA synthetase (GlyRS)[101]. This enzyme can synthetize Ap_4A independently on amino acid, because it has inserted ATP binding domain. GlyRS is relatively slow in Ap_4A synthesis, compared to LysRS, but it could mean that RNA is a preferred substrate. The unique feature of this tRNA synthetases could hypothetically be used for an RNA capping.

There are many other enzymes that can synthetize Ap_4A, and therefore hypothetically act as RNA capping enzymes[50]. Some of these enzymes are part of metabolic pathways, such as acetyl CoA syntase[114], UTP:glucose-1-phosphate uridylyltransferase[115], luciferase[116] or propionate kinase[117]. We did not consider these enzymes to be the best candidates for RNA capping enzymes, because they are mostly constitutively transcribed housekeeping genes, whereas dinucleoside polyphosphate caps are related to the stress response. However, any moonlighting activity of these enzymes cannot be ruled out. Ubiquitin ligases have been showed to create Ap_4A as well, but their reaction rates are very slow[47].

Among the enzymes, with a high probability of being RNA capping enzymes are considered RNA ligases. Firstly, these enzymes have been demonstrated to create Ap_4A[49]. Next, their reaction mechanism involves the formation of an enzyme-lysine-AMP intermediate, where AMP is then transferred to monophosphorylated RNA, resulting in the creation of an AppRNA intermediate (which can be viewed as an Ap_2N-RNA cap). AppRNA intermediate is then used for ligation to 3′ end of different RNA molecule[118]. This illustrates that RNA ligases are capable of formation of molecular structures similar to dinucleoside polyphosphate caps. On the top of that, RNA ligase, DNA ligases and eukaryotic mRNA guanylyltransferase have been shown to form a single protein family[21,119]. However, RNA ligases can ligate only monophosphorylated RNA on 5′ end and not triphosphorylated RNA, which would be required for Np_4N-RNA synthesis. Only one engineered artificial RNA ligase was able to ligate triphosphorylated RNA, which demonstrates that it is theoretically possible[120,121].

RNA ligases RtcA and RtcB were tested for their RNA capping potential by Luciano *and* Belasco[6]. They measured the abundance of dinucleoside polyphosphate RNA caps after the deletion of both RtcA and RtcB but did not observe any significant change in their abundance. Therefore, RtcA and RtcB can be dismissed from the list of potential RNA capping enzymes. DNA ligases were also shown to be able synthetize Ap_4A[48], and their reaction mechanism is similar to that of RNA ligases, but RNA ligases seem more probable to be RNA capping enzymes, because they can specifically bind RNA molecules.

Thg1 is a guanylyltransferase that attaches single guanosine to the 5' end of $tRNA^{His}$ in 3′ to 5′ orientation[122]. This essential enzyme is present in eukaryotes, archaea, and bacteria (not present in *E. coli*)[123]. Its reaction mechanism is similar to that of both tRNA synthetases and RNA ligases, but it does not belong to either of these protein families[122,124]. Thg1 first creates an App-tRNA intermediate and then binds molecule of GTP. Subsequently, the 3′-OH group of GTP attacks the intermediate and guanosine is attached to 5′ end of tRNA (and triphosphate 5′ end on tRNA is subsequently cleaved to monophosphate 5′ end)[125]. As this enzyme attaches guanosine in 3′ end to 5′ end orientation, it is not a potential RNA capping enzyme, however, its reaction mechanism and function illustrates that there might be a homologous protein capable of performing a similar reaction in a 5′ to 5′ orientation (theoretically, a different positioning of GTP in the active site could result in the attack from 5′ triphosphate group of GTP, instead of 3′-OH group, leading to formation of Gp_4G-RNA cap).

In a hypothetical scenario, there could be an unknown *E. coli* RNA ligase or a homolog of RNA ligases that could utilize GTP in the first step of reaction and triphosphorylated RNA in the second step of reaction. This enzyme would then create Gp_4G-RNA and therefore act as an RNA capping enzyme in *E. coli*. Considering that in 2018, there were 1264 genes with unknown functions in *E. coli*[126], it is possible that among them is an enzyme with these hypothetical characteristics.

8. Synthesis and purification of isotopically labelled Ap_4A

The inability of tRNA synthetases to create Ap_4A-RNA caps lead us to the decision to use LysU for synthesis of commercially unavailable diadenosine-$^{15}N_{10}$ tetraphosphate. Chemical synthesis of all dinucleoside polyphosphates is challenging with several purification steps and modest yields[127]. In contrary, enzymatic synthesis of

dinucleoside polyphosphates is a one-step reaction with one purification step, which leads to higher yields and purity of the product. Isotopically labelled molecules are used as standards in LC-MS based quantification. These molecules have the same chemical properties as the quantified molecules but differs in their molecular weight. In this case, $^{15}N_{10}$-Ap$_4$A can be used in Ap$_4$A-RNA cap quantification in RNA isolated from living cells.

The $^{15}N_{10}$-Ap$_4$A was synthetized by LysU from $^{15}N_5$-ATP and then separated by HPLC and SPE. The product was subsequently measured by NMR, which confirmed the presence of $^{15}N_{10}$-Ap$_4$A and the purity of the sample. The final yield of $^{15}N_{10}$-Ap$_4$A was 0.180 mg. The amount of isotopically labelled standard needed for each measurement is very low, therefore this amount of $^{15}N_{10}$-Ap$_4$A could be probably used for several years.

Conclusion

A novel type of non-canonical RNA caps was recently discovered in bacteria[1]. They are composed of a molecule of dinucleoside polyphosphate at the 5′ end of RNA, therefore they are called dinucleoside polyphosphate RNA caps. Their physiological function has yet to be elucidated, therefore more details about their metabolism are crucial for understanding their role in the cell. One of the important aspects of their metabolism is their synthesis. It has already been showed that Np_nN-RNA caps can be formed by incorporation of a dinucleoside polyphosphate molecule at the 5′ end of RNA during transcription as a NCIN[5,6]. However, post-transcriptional enzymatic synthesis of these caps cannot be ruled out. tRNA synthetases are enzymes known for their essential role in translation, but they are also the enzymes responsible for the synthesis of dinucleoside polyphosphates in the cell[13]. This makes them the best candidates for post-transcriptional capping enzyme in bacteria. The aim of this thesis was to test the hypothesis, that tRNA synthetases are able to form Ap_4A-RNA cap *in vitro* and therefore act as RNA capping enzymes.

In this thesis, four tRNA synthetases from *E. coli* were selected – LysU, LysS, HisS and MetG. Selected tRNA synthetases were subcloned to plasmids, expressed in *E. coli,* and subsequently purified by fast protein liquid chromatography (FPLC). Purified tRNA synthetases were tested for production of free dinucleoside polyphosphates and the reaction products were analysed by high-performance liquid chromatography (HPLC). Subsequently, all purified tRNA synthetases were tested and compared in their ability to create Ap_4A. LysU and LysS were shown to be very efficient in Ap_4A synthesis, however, HisS and MetG were found out to be relatively slow in its synthesis. LysU was tested for production of different Ap_nNs as well and it was shown that it is able to synthetize Ap_3A, Ap_4G, Ap_4C and Ap_4U, but with slower reaction rates than Ap_4A.

In the subsequent experiments, tRNA synthetase were tested for formation of Ap_4A-RNA. However, we found that tRNA synthetases were unable to form Ap_4A-RNA. The reaction was tested on two different RNAs, an artificial 35mer and *yeiP* transcript (beginning of an *E. coli* gene). The reaction was performed using several different reaction conditions, but we did not observe formation of capped RNA. These results were then confirmed by LC-MS. In conclusion, tRNA synthetases do not form Ap_4A-RNA cap *in vitro* and do not act as RNA capping enzymes.

These results are inconsistent with findings reported by Luciano *et al*.[2]. In their experiment, an initial test for the ability to form Ap_4A-RNA cap by LysU was successful. We are confident that our results are more significant, as we conducted the experiment with four tRNA synthetases on two different transcripts and verified our findings by LC-MS. Furthermore, results presented in this thesis are further supported by the inability of human lysyl-tRNA synthetases to form Ap_4A-RNA cap *in vitro*, as reported by Potužník *et al.*[38].

We present a hypothesis that the inability to take an RNA as a substrate is due to the negatively charged region near the active site of LysU, which respells the negatively charged RNA. To exclude the hypothesis that the interaction of negatively charged protein with RNA is mediated by magnesium ions, we employed molecular dynamics simulation of Mg^{2+} binding to protein surface. It was found that ions were binding with only low probability to negatively charged area. This suggests that it is unlikely that the interaction is mediated through ion binding, which makes our hypothesis more reliable. In the end, we used LysU to synthesize commercially unavailable isotopically labelled $^{15}N_{10}$-Ap_4A.

In summary, four *E. coli* tRNA synthetases were cloned to plasmids, purified by FPLC and successfully tested for production of free dinucleoside polyphosphates. Subsequent experiments on RNA showed that tRNA synthetases are unable to form dinucleoside polyphosphate RNA cap and therefore do not act as RNA capping enzymes.

Citations

1 Hudecek, O. *et al.* Dinucleoside polyphosphates act as 5'-RNA caps in bacteria. *Nat Commun* **11**, 1052 (2020). https://doi.org:10.1038/s41467-020-14896-8

2 Luciano, D. J., Levenson-Palmer, R. & Belasco, J. G. Stresses that Raise Np(4)A Levels Induce Protective Nucleoside Tetraphosphate Capping of Bacterial RNA. *Mol Cell* **75**, 957-966 e958 (2019). https://doi.org:10.1016/j.molcel.2019.05.031

3 Enauz. https://commons.wikimedia.org/wiki/File:Aars_alignment.png (Wikimedia Commons, 2018).

4 Yersinia~commonswiki. https://commons.wikimedia.org/wiki/File:TRNA_sekundaarstrukt.png (Wikimedia Commons, 2006).

5 Benoni, R., Culka, M., Hudecek, O., Gahurova, L. & Cahova, H. Dinucleoside Polyphosphates as RNA Building Blocks with Pairing Ability in Transcription Initiation. *ACS Chem Biol* **15**, 1765-1772 (2020). https://doi.org:10.1021/acschembio.0c00178

6 Luciano, D. J. & Belasco, J. G. Np(4)A alarmones function in bacteria as precursors to RNA caps. *Proc Natl Acad Sci U S A* **117**, 3560-3567 (2020). https://doi.org:10.1073/pnas.1914229117

7 O'Donoghue, P. & Luthey-Schulten, Z. On the evolution of structure in aminoacyl-tRNA synthetases. *Microbiol Mol Biol Rev* **67**, 550-573 (2003). https://doi.org:10.1128/mmbr.67.4.550-573.2003

8 Desogus, G., Todone, F., Brick, P. & Onesti, S. Active site of lysyl-tRNA synthetase: structural studies of the adenylation reaction. *Biochemistry* **39**, 8418-8425 (2000). https://doi.org:10.1021/bi0006722

9 Chen, X. *et al.* Multiple catalytic activities of Escherichia coli lysyl-tRNA synthetase (LysU) are dissected by site-directed mutagenesis. *Febs j* **280**, 102-114 (2013). https://doi.org:10.1111/febs.12053

10 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021). https://doi.org:10.1038/s41586-021-03819-2

11 Shi, H. & Moore, P. B. The crystal structure of yeast phenylalanine tRNA at 1.93 A resolution: a classic structure revisited. *Rna* **6**, 1091-1105 (2000). https://doi.org:10.1017/s1355838200000364

12 Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nature Methods* **19**, 679-682 (2022). https://doi.org:10.1038/s41592-022-01488-1

13 Zamecnik, P. C., Stephenson, M. L., Janeway, C. M. & Randerath, K. Enzymatic synthesis of diadenosine tetraphosphate and diadenosine triphosphate with a purified lysyl-sRNA synthetase. *Biochem Biophys Res Commun* **24**, 91-97 (1966). https://doi.org:10.1016/0006-291x(66)90415-3

14 Cappannini, A. *et al.* MODOMICS: a database of RNA modifications and related information. 2023 update. *Nucleic Acids Research* (2023). https://doi.org:10.1093/nar/gkad1083

15 Shatkin, A. J. Capping of eucaryotic mRNAs. *Cell* **9**, 645-653 (1976). https://doi.org:10.1016/0092-8674(76)90128-8

16 McCracken, S. *et al.* 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev* **11**, 3306-3318 (1997). https://doi.org:10.1101/gad.11.24.3306

17 Gonatopoulos-Pournatzis, T. & Cowling, V. H. Cap-binding complex (CBC). *Biochem J* **457**, 231-242 (2014). https://doi.org:10.1042/bj20131214

18 Evdokimova, V. *et al.* The major mRNA-associated protein YB-1 is a potent 5' cap-dependent mRNA stabilizer. *Embo j* **20**, 5491-5502 (2001). https://doi.org:10.1093/emboj/20.19.5491
19 Cahova, H. Prokaryotes also Protect their RNAs – New Types of 5' RNA Caps. *Chemické listy* **111**, 414–423 (2017).
20 Hui, M. P., Foley, P. L. & Belasco, J. G. Messenger RNA degradation in bacterial cells. *Annu Rev Genet* **48**, 537-559 (2014). https://doi.org:10.1146/annurev-genet-120213-092340
21 Shuman, S. & Lima, C. D. The polynucleotide ligase and RNA capping enzyme superfamily of covalent nucleotidyltransferases. *Curr Opin Struct Biol* **14**, 757-764 (2004). https://doi.org:10.1016/j.sbi.2004.10.006
22 Chen, Y. G., Kowtoniuk, W. E., Agarwal, I., Shen, Y. & Liu, D. R. LC/MS analysis of cellular RNA reveals NAD-linked RNA. *Nat Chem Biol* **5**, 879-881 (2009). https://doi.org:10.1038/nchembio.235
23 Cahova, H., Winz, M. L., Hofer, K., Nubel, G. & Jaschke, A. NAD captureSeq indicates NAD as a bacterial cap for a subset of regulatory RNAs. *Nature* **519**, 374-377 (2015). https://doi.org:10.1038/nature14020
24 Bird, J. G. *et al.* The mechanism of RNA 5′ capping with NAD+, NADH and desphospho-CoA. *Nature* **535**, 444-447 (2016). https://doi.org:10.1038/nature18622
25 Walters, R. W. *et al.* Identification of NAD+ capped mRNAs in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **114**, 480-485 (2017). https://doi.org:10.1073/pnas.1619369114
26 Jiao, X. *et al.* 5' End Nicotinamide Adenine Dinucleotide Cap in Human Cells Promotes RNA Decay through DXO-Mediated deNADding. *Cell* **168**, 1015-1027.e1010 (2017). https://doi.org:10.1016/j.cell.2017.02.019
27 Wang, Y. *et al.* NAD(+)-capped RNAs are widespread in the Arabidopsis transcriptome and can probably be translated. *Proc Natl Acad Sci U S A* **116**, 12094-12102 (2019). https://doi.org:10.1073/pnas.1903682116
28 Gomes-Filho, J. V. *et al.* Identification of NAD-RNA species and ADPR-RNA decapping in Archaea. *Nat Commun* **14**, 7597 (2023). https://doi.org:10.1038/s41467-023-43377-x
29 Wiedermannova, J., Babu, R. & Yuzenkova, Y. 5'-NADylation of RNA is well-tolerated by bacterial cell. *bioRxiv*, 2023.2003.2031.535138 (2023). https://doi.org:10.1101/2023.03.31.535138
30 Sharma, S., Yang, J., Favate, J., Shah, P. & Kiledjian, M. NADcapPro and circNC: methods for accurate profiling of NAD and non-canonical RNA caps in eukaryotes. *Commun Biol* **6**, 406 (2023). https://doi.org:10.1038/s42003-023-04774-6
31 Wolfram-Schauerte, M. *et al.* A viral ADP-ribosyltransferase attaches RNA chains to host proteins. *Nature* **620**, 1054-1062 (2023). https://doi.org:10.1038/s41586-023-06429-2
32 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 U S A* **106**, 7768-7773 (2009). https://doi.org:10.1073/pnas.0900528106
33 Sapkota, K. *et al.* Post-transcriptional capping generates coenzyme A-linked RNA. *RNA Biol* **21**, 1-12 (2024). https://doi.org:10.1080/15476286.2023.2288740

34 Sherwood, A. V. *et al.* Hepatitis C virus RNA is 5′-capped with flavin adenine dinucleotide. *Nature* **619**, 811-818 (2023). https://doi.org:10.1038/s41586-023-06301-3
35 Schweibenz, B. D., Solotchi, M., Hanpude, P., Devarkar, S. C. & Patel, S. S. RIG-I recognizes metabolite-capped RNAs as signaling ligands. *Nucleic Acids Res* **51**, 8102-8114 (2023). https://doi.org:10.1093/nar/gkad518
36 Goerlich, O., Foeckler, R. & Holler, E. Mechanism of synthesis of adenosine(5')tetraphospho(5')adenosine (AppppA) by aminoacyl-tRNA synthetases. *Eur J Biochem* **126**, 135-142 (1982). https://doi.org:10.1111/j.1432-1033.1982.tb06757.x
37 Giammarinaro, P. I. *et al.* Diadenosine tetraphosphate regulates biosynthesis of GTP in Bacillus subtilis. *Nature Microbiology* **7**, 1442-1452 (2022). https://doi.org:10.1038/s41564-022-01193-x
38 Potuznik, J. F. *et al.* Diadenosine Tetraphosphate (Ap4A) Serves as a 5' RNA Cap in Mammalian Cells. *Angewandte Chemie International Edition* **n/a**, e202314951 https://doi.org:https://doi.org/10.1002/anie.202314951
39 McLennan, A. G. *Ap4a and Other Dinucleoside Polyphosphates*. (Taylor & Francis, 1992).
40 Říha, J. *Intracellular dinucleoside polyphosphates and methods of their detection* Bachelor thesis, Charles university, Faculty of Science, (2022).
41 Lee, P. C., Bochner, B. R. & Ames, B. N. AppppA, heat-shock stress, and cell oxidation. *Proc Natl Acad Sci U S A* **80**, 7496-7500 (1983). https://doi.org:10.1073/pnas.80.24.7496
42 Murphy, G. A., Halliday, D. & McLennan, A. G. The Fhit tumor suppressor protein regulates the intracellular concentration of diadenosine triphosphate but not diadenosine tetraphosphate. *Cancer Res* **60**, 2342-2344 (2000).
43 Liu, J. J. & McLennan, A. G. Purification and properties of GTP:GTP guanylyltransferase from encysted embryos of the brine shrimp Artemia. *J Biol Chem* **269**, 11787-11794 (1994).
44 Finamore, F. J. & Warner, A. H. The occurrence of P1, P4-diguanosine 5'-tetraphosphate in brine shrimp eggs. *J Biol Chem* **238**, 344-348 (1963).
45 Blanquet, S., Plateau, P. & Brevet, A. The role of zinc in 5',5'-diadenosine tetraphosphate production by aminoacyl-transfer RNA synthetases. *Mol Cell Biochem* **52**, 3-11 (1983). https://doi.org:10.1007/bf00230583
46 Plateau, P. & Blanquet, S. Zinc-dependent synthesis of various dinucleoside 5',5' ' '-P1,P3-Tri- or 5'',5' ' '-P1,P4-tetraphosphates by Escherichia coli lysyl-tRNA synthetase. *Biochemistry* **21**, 5273-5279 (1982). https://doi.org:10.1021/bi00264a024
47 Götz, K. H. *et al.* Formation of the Alarmones Diadenosine Triphosphate and Tetraphosphate by Ubiquitin- and Ubiquitin-like-Activating Enzymes. *Cell Chem Biol* **26**, 1535-1543.e1535 (2019). https://doi.org:10.1016/j.chembiol.2019.08.004
48 Marriott, A. S. *et al.* Diadenosine 5', 5'''-P(1),P(4)-tetraphosphate (Ap4A) is synthesized in response to DNA damage and inhibits the initiation of DNA replication. *DNA Repair (Amst)* **33**, 90-100 (2015). https://doi.org:10.1016/j.dnarep.2015.06.008
49 Atencia, E. A., Madrid, O., Günther Sillero, M. A. & Sillero, A. T4 RNA ligase catalyzes the synthesis of dinucleoside polyphosphates. *Eur J Biochem* **261**, 802-811 (1999). https://doi.org:10.1046/j.1432-1327.1999.00338.x

50 Fraga, H. & Fontes, R. Enzymatic synthesis of mono and dinucleoside polyphosphates. *Biochim Biophys Acta* **1810**, 1195-1204 (2011). https://doi.org:10.1016/j.bbagen.2011.09.010
51 McLennan, A. G. The Nudix hydrolase superfamily. *Cell Mol Life Sci* **63**, 123-143 (2006). https://doi.org:10.1007/s00018-005-5386-7
52 Mititelu, M.-B. *et al.* Arabidopsis thaliana NudiXes have RNA-decapping activity. *RSC Chemical Biology* **4**, 223-228 (2023). https://doi.org:10.1039/D2CB00213B
53 Farr, S. B., Arnosti, D. N., Chamberlin, M. J. & Ames, B. N. An apaH mutation causes AppppA to accumulate and affects motility and catabolite repression in Escherichia coli. *Proc Natl Acad Sci U S A* **86**, 5010-5014 (1989). https://doi.org:10.1073/pnas.86.13.5010
54 Barnes, L. D. *et al.* Fhit, a putative tumor suppressor in humans, is a dinucleoside 5',5"'-P1,P3-triphosphate hydrolase. *Biochemistry* **35**, 11529-11535 (1996). https://doi.org:10.1021/bi961415t
55 Abele, F. *et al.* A Novel NAD-RNA Decapping Pathway Discovered by Synthetic Light-Up NAD-RNAs. *Biomolecules* **10** (2020). https://doi.org:10.3390/biom10040513
56 Ogilvie, A. Determination of diadenosine tetraphosphate (Ap4A) levels in subpicomole quantities by a phosphodiesterase luciferin--luciferase coupled assay: application as a specific assay for diadenosine tetraphosphatase. *Anal Biochem* **115**, 302-307 (1981). https://doi.org:10.1016/0003-2697(81)90009-9
57 Lee, P. C., Bochner, B. R. & Ames, B. N. Diadenosine 5',5"'-P1,P4-tetraphosphate and related adenylylated nucleotides in Salmonella typhimurium. *J Biol Chem* **258**, 6827-6834 (1983).
58 Ji, X. *et al.* Alarmone Ap4A is elevated by aminoglycoside antibiotics and enhances their bactericidal activity. *Proc Natl Acad Sci U S A* **116**, 9578-9585 (2019). https://doi.org:10.1073/pnas.1822026116
59 Ji, X. *et al.* Diadenosine tetraphosphate modulated quorum sensing in bacteria treated with kanamycin. *BMC Microbiol* **23**, 353 (2023). https://doi.org:10.1186/s12866-023-03113-3
60 McLennan, A. G. Dinucleoside polyphosphates-friend or foe? *Pharmacol Ther* **87**, 73-89 (2000). https://doi.org:10.1016/s0163-7258(00)00041-3
61 Despotovic, D. *et al.* Diadenosine tetraphosphate (Ap4A) - an E. coli alarmone or a damage metabolite? *FEBS J* **284**, 2194-2215 (2017). https://doi.org:10.1111/febs.14113
62 Garrison, P. N., Mathis, S. A. & Barnes, L. D. In vivo levels of diadenosine tetraphosphate and adenosine tetraphospho-guanosine in Physarum polycephalum during the cell cycle and oxidative stress. *Mol Cell Biol* **6**, 1179-1186 (1986). https://doi.org:10.1128/mcb.6.4.1179-1186.1986
63 Lee, Y.-N., Nechushtan, H., Figov, N. & Razin, E. The Function of Lysyl-tRNA Synthetase and Ap4A as Signaling Regulators of MITF Activity in FcεRI-Activated Mast Cells. *Immunity* **20**, 145-151 (2004). https://doi.org:10.1016/S1074-7613(04)00020-2
64 Yu, J. *et al.* Second messenger Ap4A polymerizes target protein HINT1 to transduce signals in FcεRI-activated mast cells. *Nature Communications* **10**, 4664 (2019). https://doi.org:10.1038/s41467-019-12710-8
65 Nübel, G., Sorgenfrei, F. A. & Jäschke, A. Boronate affinity electrophoresis for the purification and analysis of cofactor-modified RNAs. *Methods* **117**, 14-20 (2017). https://doi.org:10.1016/j.ymeth.2016.09.008

66 Levenson-Palmer, R. *et al.* A distinct RNA recognition mechanism governs Np(4) decapping by RppH. *Proc Natl Acad Sci U S A* **119** (2022). https://doi.org:10.1073/pnas.2117318119
67 Woese, C. R., Olsen, G. J., Ibba, M. & Söll, D. Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol Mol Biol Rev* **64**, 202-236 (2000). https://doi.org:10.1128/mmbr.64.1.202-236.2000
68 Pang, Y. L., Poruri, K. & Martinis, S. A. tRNA synthetase: tRNA aminoacylation and beyond. *Wiley Interdiscip Rev RNA* **5**, 461-480 (2014). https://doi.org:10.1002/wrna.1224
69 Wakasugi, K. & Yokosawa, T. Non-canonical functions of human cytoplasmic tyrosyl-, tryptophanyl- and other aminoacyl-tRNA synthetases. *Enzymes* **48**, 207-242 (2020). https://doi.org:10.1016/bs.enz.2020.04.001
70 Eriani, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* **347**, 203-206 (1990). https://doi.org:10.1038/347203a0
71 Rubio Gomez, M. A. & Ibba, M. Aminoacyl-tRNA synthetases. *RNA* **26**, 910-936 (2020). https://doi.org:10.1261/rna.071720.119
72 Giege, R. & Springer, M. Aminoacyl-tRNA Synthetases in the Bacterial World. *EcoSal Plus* **7** (2016). https://doi.org:10.1128/ecosalplus.ESP-0002-2016
73 Tamaki, S., Tomita, M., Suzuki, H. & Kanai, A. Systematic Analysis of the Binding Surfaces between tRNAs and Their Respective Aminoacyl tRNA Synthetase Based on Structural and Evolutionary Data. *Front Genet* **8**, 227 (2017). https://doi.org:10.3389/fgene.2017.00227
74 Rho, S. B. *et al.* Interaction between human tRNA synthetases involves repeated sequence elements. *Proc Natl Acad Sci U S A* **93**, 10128-10133 (1996). https://doi.org:10.1073/pnas.93.19.10128
75 Ofir-Birin, Y. *et al.* Structural switch of lysyl-tRNA synthetase between translation and transcription. *Mol Cell* **49**, 30-42 (2013). https://doi.org:10.1016/j.molcel.2012.10.010
76 Chaliotis, A. *et al.* The complex evolutionary history of aminoacyl-tRNA synthetases. *Nucleic Acids Res* **45**, 1059-1068 (2017). https://doi.org:10.1093/nar/gkw1182
77 Wolf, Y. I., Aravind, L., Grishin, N. V. & Koonin, E. V. Evolution of aminoacyl-tRNA synthetases--analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res* **9**, 689-710 (1999).
78 Pham, Y. *et al.* A minimal TrpRS catalytic domain supports sense/antisense ancestry of class I and II aminoacyl-tRNA synthetases. *Mol Cell* **25**, 851-862 (2007). https://doi.org:10.1016/j.molcel.2007.02.010
79 Pham, Y. *et al.* Tryptophanyl-tRNA synthetase Urzyme: a model to recapitulate molecular evolution and investigate intramolecular complementation. *J Biol Chem* **285**, 38590-38601 (2010). https://doi.org:10.1074/jbc.M110.136911
80 Rodin, S. N. & Ohno, S. Two types of aminoacyl-tRNA synthetases could be originally encoded by complementary strands of the same nucleic acid. *Orig Life Evol Biosph* **25**, 565-589 (1995). https://doi.org:10.1007/bf01582025
81 Carter, C. W., Jr. *et al.* The Rodin-Ohno hypothesis that two enzyme superfamilies descended from one ancestral gene: an unlikely scenario for the origins of translation that will not be dismissed. *Biol Direct* **9**, 11 (2014). https://doi.org:10.1186/1745-6150-9-11

82 Curnow, A. W. *et al.* Glu-tRNAGln amidotransferase: a novel heterotrimeric enzyme required for correct decoding of glutamine codons during translation. *Proc Natl Acad Sci U S A* **94**, 11819-11826 (1997). https://doi.org:10.1073/pnas.94.22.11819
83 Ambrogelly, A. *et al.* Cys-tRNACys formation and cysteine biosynthesis in methanogenic archaea: two faces of the same problem? *Cell Mol Life Sci* **61**, 2437-2445 (2004). https://doi.org:10.1007/s00018-004-4194-9
84 Lamour, V. *et al.* Evolution of the Glx-tRNA synthetase family: the glutaminyl enzyme as a case of horizontal gene transfer. *Proc Natl Acad Sci U S A* **91**, 8670-8674 (1994). https://doi.org:10.1073/pnas.91.18.8670
85 Shepherd, J. & Ibba, M. Bacterial transfer RNAs. *FEMS Microbiol Rev* **39**, 280-300 (2015). https://doi.org:10.1093/femsre/fuv004
86 Altman, S. & Smith, J. D. Tyrosine tRNA precursor molecule polynucleotide sequence. *Nat New Biol* **233**, 35-39 (1971). https://doi.org:10.1038/newbio233035a0
87 Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**, 849-857 (1983). https://doi.org:https://doi.org/10.1016/0092-8674(83)90117-4
88 Li, Z. & Deutscher, M. P. RNase E plays an essential role in the maturation of Escherichia coli tRNA precursors. *Rna* **8**, 97-109 (2002). https://doi.org:10.1017/s1355838202014929
89 Bowden, K. E., Wiese, N. S., Perwez, T., Mohanty, B. K. & Kushner, S. R. The rph-1-Encoded Truncated RNase PH Protein Inhibits RNase P Maturation of Pre-tRNAs with Short Leader Sequences in the Absence of RppH. *J Bacteriol* **199** (2017). https://doi.org:10.1128/jb.00301-17
90 Plateau, P., Mayaux, J. F. & Blanquet, S. Zinc(II)-dependent synthesis of diadenosine 5', 5''' -P(1) ,P(4) -tetraphosphate by Escherichia coli and yeast phenylalanyl transfer ribonucleic acid synthetases. *Biochemistry* **20**, 4654-4662 (1981). https://doi.org:10.1021/bi00519a021
91 Traut, T. W. Synthesis of hybrid bisnucleoside 5',5'''-P1,P4-tetraphosphates by aminoacyl-tRNA synthetases. *Mol Cell Biochem* **75**, 15-21 (1987). https://doi.org:10.1007/bf00231604
92 Jakubowski, H. Synthesis of diadenosine 5',5'''-P1,P4-tetraphosphate and related compounds by plant (Lupinus luteus) seryl-tRNA and phenylalanyl-tRNA synthetases. *Acta Biochim Pol* **30**, 51-69 (1983).
93 Delagoutte, B., Moras, D. & Cavarelli, J. tRNA aminoacylation by arginyl-tRNA synthetase: induced conformations during substrates binding. *Embo j* **19**, 5599-5610 (2000). https://doi.org:10.1093/emboj/19.21.5599
94 Ibba, M. *et al.* Substrate recognition by class I lysyl-tRNA synthetases: a molecular basis for gene displacement. *Proc Natl Acad Sci U S A* **96**, 418-423 (1999). https://doi.org:10.1073/pnas.96.2.418
95 Brevet, A., Chen, J., Lévêque, F., Blanquet, S. & Plateau, P. Comparison of the enzymatic properties of the two Escherichia coli lysyl-tRNA synthetase species. *J Biol Chem* **270**, 14439-14444 (1995). https://doi.org:10.1074/jbc.270.24.14439
96 Brevet, A., Chen, J., Lévêque, F., Plateau, P. & Blanquet, S. In vivo synthesis of adenylylated bis(5'-nucleosidyl) tetraphosphates (Ap4N) by Escherichia coli aminoacyl-tRNA synthetases. *Proc Natl Acad Sci U S A* **86**, 8275-8279 (1989). https://doi.org:10.1073/pnas.86.21.8275

97 Kimura, Y., Tanaka, C. & Oka, M. Identification of Major Enzymes Involved in the Synthesis of Diadenosine Tetraphosphate and/or Adenosine Tetraphosphate in Myxococcus xanthus. *Curr Microbiol* **75**, 811-817 (2018). https://doi.org:10.1007/s00284-018-1452-x
98 Tanner, J. A., Abowath, A. & Miller, A. D. Isothermal titration calorimetry reveals a zinc ion as an atomic switch in the diadenosine polyphosphates. *J Biol Chem* **277**, 3073-3078 (2002). https://doi.org:10.1074/jbc.M106588200
99 Randerath, K., Janeway, C. M., Stephenson, M. L. & Zamecnik, P. C. Isolation and characterization of dinucleoside tetra- and tri-phosphates formed in the presence of lysyl-sRNA synthetase. *Biochem Biophys Res Commun* **24**, 98-105 (1966). https://doi.org:10.1016/0006-291x(66)90416-5
100 Wright, M., Boonyalai, N., Tanner, J. A., Hindley, A. D. & Miller, A. D. The duality of LysU, a catalyst for both Ap4A and Ap3A formation. *FEBS J* **273**, 3534-3544 (2006). https://doi.org:10.1111/j.1742-4658.2006.05361.x
101 Guo, R. T., Chong, Y. E., Guo, M. & Yang, X. L. Crystal structures and biochemical analyses suggest a unique mechanism and role for human glycyl-tRNA synthetase in Ap4A homeostasis. *J Biol Chem* **284**, 28968-28976 (2009). https://doi.org:10.1074/jbc.M109.030692
102 Led, J. J., Switon, W. K. & Jensen, K. F. Phosphorolytic activity of Escherichia coli glycyl-tRNA synthetase towards its cognate aminoacyl adenylate detected by 31P-NMR spectroscopy and thin-layer chromatography. *Eur J Biochem* **136**, 469-479 (1983). https://doi.org:10.1111/j.1432-1033.1983.tb07765.x
103 Voleníková, B. *Studies of protein and 5´terminus RNA interactions*, UCT Prague, (2023).
104 Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem* **29**, 1859-1865 (2008). https://doi.org:10.1002/jcc.20945
105 Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, 27-38 (1996). https://doi.org:10.1016/0263-7855(96)00018-5
106 Stone, J. An Efficient Library for Parallel Ray Tracing and Animation. (1997).
107 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J Mol Biol* **215**, 403-410 (1990). https://doi.org:10.1016/s0022-2836(05)80360-2
108 Vopalensky, P. *et al.* Exploring RNA modifications in infectious non-coding circular RNAs. *bioRxiv*, 2024.2003.2012.584625 (2024). https://doi.org:10.1101/2024.03.12.584625
109 Jurrus, E. *et al.* Improvements to the APBS biomolecular solvation software suite. *Protein Sci* **27**, 112-128 (2018). https://doi.org:10.1002/pro.3280
110 Eiler, S., Dock-Bregeon, A., Moulinier, L., Thierry, J. C. & Moras, D. Synthesis of aspartyl-tRNA(Asp) in Escherichia coli--a snapshot of the second step. *Embo j* **18**, 6532-6541 (1999). https://doi.org:10.1093/emboj/18.22.6532
111 Giacobelli, V. G. *et al.* In Vitro Evolution Reveals Noncationic Protein-RNA Interaction Mediated by Metal Ions. *Mol Biol Evol* **39** (2022). https://doi.org:10.1093/molbev/msac032
112 Loverix, S., Winqvist, A., Strömberg, R. & Steyaert, J. Mechanism of RNase T1: concerted triester-like phosphoryl transfer via a catalytic three-centered hydrogen bond. *Chem Biol* **7**, 651-658 (2000). https://doi.org:10.1016/s1074-5521(00)00005-3

113 Allnér, O., Nilsson, L. & Villa, A. Magnesium Ion-Water Coordination and Exchange in Biomolecular Simulations. *J Chem Theory Comput* **8**, 1493-1502 (2012). https://doi.org:10.1021/ct3000734
114 Guranowski, A., Günther Sillero, M. A. & Sillero, A. Adenosine 5'-tetraphosphate and adenosine 5'-pentaphosphate are synthesized by yeast acetyl coenzyme A synthetase. *J Bacteriol* **176**, 2986-2990 (1994). https://doi.org:10.1128/jb.176.10.2986-2990.1994
115 Guranowski, A., de Diego, A., Sillero, A. & Günther Sillero, M. A. Uridine 5'-polyphosphates (p4U and p5U) and uridine(5')polyphospho(5')nucleosides (Up(n)Ns) can be synthesized by UTP:glucose-1-phosphate uridylyltransferase from Saccharomyces cerevisiae. *FEBS Lett* **561**, 83-88 (2004). https://doi.org:10.1016/s0014-5793(04)00126-7
116 Ortiz, B., Sillero, A. & Günther Sillero, M. A. Specific synthesis of adenosine(5')tetraphospho(5')nucleoside and adenosine(5')oligophospho(5')adenosine (n > 4) catalyzed by firefly luciferase. *Eur J Biochem* **212**, 263-270 (1993). https://doi.org:10.1111/j.1432-1033.1993.tb17658.x
117 Simanshu, D. K., Savithri, H. S. & Murthy, M. R. Crystal structures of Salmonella typhimurium propionate kinase and its complex with Ap4A: evidence for a novel Ap4A synthetic activity. *Proteins* **70**, 1379-1388 (2008). https://doi.org:10.1002/prot.21626
118 Shuman, S. RNA Repair: Hiding in Plain Sight. *Annu Rev Genet* **57**, 461-489 (2023). https://doi.org:10.1146/annurev-genet-071719-021856
119 Fresco, L. D. & Buratowski, S. Active site of the mRNA-capping enzyme guanylyltransferase from Saccharomyces cerevisiae: similarity to the nucleotidyl attachment motif of DNA and RNA ligases. *Proc Natl Acad Sci U S A* **91**, 6624-6628 (1994). https://doi.org:10.1073/pnas.91.14.6624
120 Seelig, B. & Szostak, J. W. Selection and evolution of enzymes from a partially randomized non-catalytic scaffold. *Nature* **448**, 828-831 (2007). https://doi.org:10.1038/nature06032
121 Haugner, J. C., 3rd & Seelig, B. Universal labeling of 5'-triphosphate RNAs by artificial RNA ligase enzyme with broad substrate specificity. *Chem Commun (Camb)* **49**, 7322-7324 (2013). https://doi.org:10.1039/c3cc44454f
122 Jackman, J. E. & Phizicky, E. M. tRNAHis guanylyltransferase catalyzes a 3'-5' polymerization reaction that is distinct from G-1 addition. *Proc Natl Acad Sci U S A* **103**, 8640-8645 (2006). https://doi.org:10.1073/pnas.0603068103
123 Heinemann, I. U., Randau, L., Tomko, R. J., Jr. & Söll, D. 3'-5' tRNAHis guanylyltransferase in bacteria. *FEBS Lett* **584**, 3567-3572 (2010). https://doi.org:10.1016/j.febslet.2010.07.023
124 Patel, K. J., Yourik, P. & Jackman, J. E. Fidelity of base-pair recognition by a 3'-5' polymerase: mechanism of the Saccharomyces cerevisiae tRNA(His) guanylyltransferase. *Rna* **27**, 683-693 (2021). https://doi.org:10.1261/rna.078686.121
125 Heinemann, I. U., Nakamura, A., O'Donoghue, P., Eiler, D. & Söll, D. tRNAHis-guanylyltransferase establishes tRNAHis identity. *Nucleic Acids Res* **40**, 333-344 (2012). https://doi.org:10.1093/nar/gkr696
126 Karp, P. D. *et al.* The EcoCyc Database. *EcoSal Plus* **8** (2018). https://doi.org:10.1128/ecosalplus.ESP-0006-2018

127 Mohamady, S. & Taylor, S. D. General Procedure for the Synthesis of Dinucleoside Polyphosphates. *The Journal of Organic Chemistry* **76**, 6344-6349 (2011). https://doi.org:10.1021/jo200540e